Learning from the (Un)Expected

Age and Individual Differences in Statistical Learning and Perceptual Learning in Speech

A Doctoral Thesis
by
Thordis Neger

Colophon

The research in this thesis was supported by a VIDI grant from the Netherlands Organization for Scientific Research (NWO) awarded to dr. Esther Janse.

Thesis Radboud University, Nijmegen, The Netherlands, with summary in English and Dutch.

Learning from the (Un)Expected: Age and Individual Differences in Statistical Learning and Perceptual Learning in Speech

Lay-out and cover design: **Thordis Neger and Alexey Klymov**

Printed by: **Ipskamp Printing, Enschede**

Learning from the (Un)Expected

Age and Individual Differences in Statistical Learning and Perceptual Learning in Speech

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens besluit van het college van decanen
in het openbaar te verdedigen op vrijdag 20 oktober 2017
om 10:30 uur precies

door

Thordis Marisa Neger
geboren op 24 februari 1986
te Goch (Duitsland)

Promotor
Prof. dr. A.C.M. Rietveld

Copromotor
Dr. E. Janse

Manuscriptcommissie
Prof. dr. M.T.C. Ernestus (voorzitter)
Prof. dr. A. van Wieringen (KU Leuven, België)
Prof. dr. P.P.G Boersma (Universiteit van Amsterdam)
Dr. S.J. Andringa (Universiteit van Amsterdam)
Dr. A. Hervais-Adelman (Max Planck Instituut voor Psycholinguistiek, Nijmegen)

Paranimfen
Cornelia Moers
Svetlana Gerakaki

Learning from the (Un)Expected

Age and Individual Differences in Statistical Learning and Perceptual Learning in Speech

Doctoral Thesis

to obtain the degree of doctor
from Radboud University Nijmegen
on the authority of the Rector Magnificus prof. dr. J.H.J.M. van Krieken,
according to the decision of the Council of Deans
to be defended in public on Friday, October 20, 2017
at 10:30 hours

by

Thordis Marisa Neger
born on 24 february 1986
in Goch (Germany)

Supervisor
Prof. dr. A.C.M Rietveld

Co-supervisor
Dr. E. Janse

Doctoral Thesis Committee
Prof. dr. M.T.C. Ernestus
Prof. dr. A. van Wieringen (KU Leuven, Belgium)
Prof. dr. P.P.G. Boersma (University of Amsterdam)
Dr. S.J. Andringa (University of Amsterdam)
Dr. A. Hervais-Adelman (Max Planck Institute for Psycholinguistics, Nijmegen)

Paranymphs
Cornelia Moers
Svetlana Gerakaki

Für Oma und Opa

Contents

1

General Introduction

1

General Introduction

"The measure of intelligence is the ability to change"
Albert Einstein

The ability to learn from experience is considered a key aspect of human intelligence. This ability is also central to human speech perception as we are constantly exposed to an extremely variable speech input. The same talker does not realize the same word twice in exactly the same way. If we consider different talkers, talkers may also vary in gender, age, speaking rate and articulatory clarity. This variability results in a rather ambiguous speech signal as, for instance, one person's /b/ may be another person's /p/. Additionally, our communication is not limited to familiar interlocutors. Therefore, our brain constantly has to adjust to new talkers and their speech idiosyncrasies. Despite this variability in the speech input, we generally perceive speech rather effortlessly. What enables us to deal with this large variability so easily?

One of the remarkable capabilities of our brain is to improve our perception of unfamiliar input by mere exposure to it. This process of perceptual learning (i.e., learning by perceiving) has recently been proposed as the central mechanism underlying robust and stable speech understanding (Kleinschmidt & Jaeger, 2015). For speech processing, we rely constantly on it: listeners adapt to the subtlest deviations in the speech signal such as a speaker's odd idiosyncratic articulation of a sound (Norris, McQueen, & Cutler, 2003) or a novel speaker's particular way of minimizing the contrast between two sounds (Clayards, Tanenhaus, Aslin, & Jacobs, 2008). Most of the times, we are not aware of this implicit process as the brain quickly adjusts to such slight deviations in the speech signal. We only start to notice the amazing flexibility and learning capability of our speech perception system if the system is challenged. That is, if speech input deviates too strongly from what we have encountered and stored in the past, our brain may need several minutes, days or even months to adjust to the unfamiliar speech input.

Given that more than 90 % of the Dutch population and more than half of all Europeans speak at least one additional language (European Commission, 2012), most of us have already experienced a more lengthy process of perceptual learning by listening to foreign-accented speech. Consider for example the following dialogue between a French hotel guest and the English speaking room service (derived from www.funny-joke-pictures.com):

>> Allo? Room service? Ici Monsieur Roux. I would like
some pepper.<<

>> Certainly, sir. Black or white? <<

>> Toilet. <<

In this written example, we need context information (toilet) to guide our understanding of the accented speech (pepper → paper). Similarly, when listening to accented speech, listeners' may engage additional resources to facilitate speech recognition and to aid the perceptual learning process. In this thesis, I aim to investigate which internal resources of listeners, i.e., which individual cognitive and perceptual abilities, predict flexibility in spoken language processing. In other words, what makes someone a good adapter?

It has been suggested that perceptual learning is mainly driven by detecting the underlying distributions and, thus, the statistical regularities in a speech input (Kleinschmidt & Jaeger, 2015). For example, a listener improves his speech perception by learning new cue distributions of the phonetic categories corresponding to a new talker or a new accent (e.g., how does this talker/accent differentiate between /b/ and /p/). Therefore, the current thesis particularly addresses the question whether individuals' ability to adapt to a novel speech condition is indeed related to their sensitivity to statistical regularities and whether both learning processes rely on the same cognitive and perceptual resources. This knowledge may help us to gain insights into processes underlying perceptual learning and may inform current models on perceptual learning and speech processing.

Current models on perceptual learning and statistical learning in speech

One of the first and most influential conceptual frameworks that accounted for changes in the perceptual system as a consequence of practice is the Reverse Hierarchy Theory (RHT) (Ahissar & Hochstein, 2004). The RHT was originally developed to explain phenomena of perceptual plasticity in the field of visual perception. Since then principles of this framework have also been applied to auditory perception (c.f., Amitay, 2009) and speech perception (e.g., Adank & Janse, 2009). The RHT argues that perceptual learning is a top-down guided process. In case a listener is exposed to a novel speech condition such as an unfamiliar accent, initial performance fails as the speech input can no longer be readily matched to pre-existing higher-level representations (e.g., word representations). Consequently, the listener cannot understand what is being said. According to the RHT, prolonged exposure to the novel speech input leads to modifications in the higher-level representations. The modified higher-level representations subsequently enable top-down guidance to retune weights at lower levels of the processing hierarchy: the weights of relevant input (e.g., cues relevant for the recognition of the word, syllable or phoneme) are increased and the weights of irrelevant input are pruned. This process of weight retuning starts at the highest level of the hierarchy (e.g., words) and continues gradually to the lower levels (e.g., syllables and phonemes) (i.e., the reverse hierarchy). When lower-level representations have been modified, performance under difficult conditions can be based on accessing these low-level representations. This is illustrated by findings that adaptation to acoustic speech degradations generalizes (from exposure words) to novel words (Hervais-Adelman, Davis, Johnsrude, & Carlyon, 2008), to non-words (Loebach, Bent, & Pisoni, 2008) and to the recognition of environmental sounds (Loebach, Pisoni, & Svirsky, 2009). These generalization findings suggest that perceptual learning in speech modifies representations at lower levels of the hierarchy, that is, representations at a sublexical level (Banai & Amitay, 2012; Hervais-Adelman et al., 2008).

The RHT has been influential in explaining behavioral observations in visual and auditory perceptual learning (Banai & Amitay, 2012; Cohen, Daikhin, & Ahissar, 2013; Nahum, Nelken, & Ahissar, 2010; Sabin, Clark, Eddins, & Wright, 2013). However, the RHT does not specify which processes take place in the initial stages of adaptation that enable the perceptual system to identify relevant cues in the input and to modify high-level representations. One of the basic principles in the RHT and other models of perceptual learning is the retuning of weights based on the relevance of features or dimensions for the specific task (Ahissar & Hochstein, 2004; Dosher & Lu, 1999; Goldstone, 1998; Petrov, Dosher, & Lu, 2005). This principle implies that stimuli have to share certain features, which can thus be considered task-relevant, for perceptual learning and for transfer of learning to occur. Accordingly, previous studies have highlighted the importance of structural regularities in the input for perceptual learning (e.g., Cohen et al., 2013; Nahum et al., 2010).

The importance of structural regularities for perceptual learning lies explicitly at the heart of a recently developed framework. The ideal listener framework (Kleinschmidt & Jaeger, 2015) has been put forward to explain how listeners are able to deal with the large variability in the speech signal. In this framework, robust speech perception is achieved by three complementary strategies: (1) recognition of the familiar; (2) generalization to the similar and (3) adaptation to the novel. The proposed strategies all rely on the same underlying principles. As the speech signal is variable, the meaning of the speech signal is always derived under a given uncertainty. Therefore, robust speech perception is probabilistic in nature and the speech perception system has to rely on distributional, i.e., statistical knowledge. As the statistical information may not always be available or predictable, a robust speech perception system has to regularly engage in statistical learning. That is, the system needs to be able to detect the frequencies with which sensory information co-occurs and to update the stored distributions accordingly. In fact, the ideal listener framework suggests that being an ideal adapter is the foundation of being an ideal listener. To put it in Kleinschmidt's and Jaeger's (2015) words: "In this way of looking at the speech perception system, learning or adaptation is a necessary part of normal speech perception" (p.169).

Kleinschmidt and Jaeger (2015) propose that adaptation in speech processing is a consequence of updating one's beliefs about the underlying distributions in speech. Listeners base their beliefs about the statistical properties of the speech signal on a finite number of observations. Consequently, listeners will never know the true underlying distribution of this talker or accent. If a new talker or accent clearly differs from previously encountered speech, the listeners' previous beliefs do not fit the input anymore. As a consequence, speech understanding drops, slows down and becomes more effortful. However, at the same time, listeners are presented with an overwhelming amount of new tokens that they can accrue to update their beliefs about the underlying statistical properties of this talker or accent. The listeners are engaged in perceptual learning. In this line of thinking, the strategies to recognize the familiar and to generalize to the similar may be viewed as short-cuts of the perceptual learning process. If listeners already possess knowledge about the statistical properties of a given auditory input (e.g., talking to a familiar person or to a person with a familiar accent), the speech perception system does not need to start from scratch but can rely on more or less accurate prior beliefs regarding this person or accent.

Observations from language acquisition support the notion that statistical learning may play a crucial role in perceptual learning and in learning to understand speech. After all, infants start to acquire language without any prior knowledge about the signal they are listening to. Yet, newborns of only one or two days old have been found to be sensitive to the statistical properties of incoming speech (Teinonen, Fellman, Naatanen, Alku, & Huotilainen, 2009). This implies that picking up on the frequencies with which sensory events co-occur is indeed central for building up a speech perception system. Only a couple of months later, babies make use of the transitional probabilities between syllables to segment speech into words (Saffran, Aslin, & Newport, 1996). Statistical learning of transitional probabilities between words has also been found to facilitate speech segmentation into phrases, thereby enabling syntax acquisition (Thompson & Newport, 2007). Indeed, by the age of five, children who perform better in statistical learning tasks have been found to show more progress in the acquisition of syntax (Kidd, 2012). Therefore, statistical learning has been proposed to be one of the central mechanisms in language acquisition, enabling young children to rapidly and successfully acquire their native language (Newport & Aslin, 2004).

Recent studies show that sensitivity to structural regularities in auditory input is not only important for language development. Adults who are more sensitive to input statistics show better language comprehension (Misyak & Christiansen, 2012), higher reading ability (Arciuli & Simpson, 2012) and more progress in learning to read in a second language (Frost, Siegelman, Narkiss, & Afek, 2013). Given that language consists of complex patterns of sequentially presented units such as phonemes, syllables and words (Conway & Christiansen, 2006), the involvement of statistical learning in online speech and language processing may only seem consequentially.

In conclusion, current models on perceptual learning implicitly or explicitly predict that sensitivity to statistical regularities is central to listeners' ability to quickly adapt to an unfamiliar speech input. Indeed, Bayesian models that incorporate functions to update the model's beliefs about the underlying cue distributions (i.e., models who show statistical learning) have shown good fit of human listeners' data on phonetic recalibration and selective adaptation (Kleinschmidt & Jaeger, 2015). However, studies on recalibration and selective adaptation typically investigate adaptation to single modified speech cues or idiosyncratic phonemes (embedded in isolated words at most). The current thesis aims to take one step forward by investigating the relationship between sensitivity to statistical regularities and speech adaptation in more natural conditions. First, globally deviant speech input (e.g., accented speech) may differ from previously encountered speech along a wide range of dimensions in contrast to single deviating phonemes or speech cues. Therefore, speech materials used in the current thesis are modified globally and thus present listeners with different levels of context information: speech materials range from words to sentences to conversational speech (cf. Chapter 4 and Chapter 6). Second, perceptual learning processes may take longer than a couple of minutes and, hence, cannot be observed over the course of a single experiment. Some natural occurring perceptual learning processes span several months as is the case in hearing rehabilitation (cf. Chapter 6). Third, current studies on perceptual learning mainly focus on group performance, i.e., does a group of listeners as a whole show adaptation to an unusual speech input. It is common knowledge, however, that individuals do not perform equally well on all tasks. Listeners' vary greatly in their speech perception performance and also in their ability to learn to understand unfamiliar speech. In this thesis, I aim to partly account for this natural variability in listeners' perceptual learning performance by linking it to listeners' statistical learning capacity. If sensitivity to statistical regularities is indeed a mainspring for perceptual learning as proposed by current models, then individuals' ability to detect such regularities should also be predictive of their perceptual learning performance in more natural perceptual learning conditions.

The role of age, hearing loss and individual differences

Most of what we know about speech processing in general and perceptual learning in particular comes from studies on highly educated, young adults. Obviously, this group is not representative of a large part of the population (this fact is known as the 'sophomore problem'). As such, we have to be cautious in generalizing findings from these studies to speech processing in general. Highly educated, young adults form a relatively homogenous group. Therefore, cognitive abilities that play a role in speech processing may go undetected when looking at this homogenous group of individuals. If we want to explain individual differences in perceptual and statistical learning to better understand the speech processing system, we may want to extend the scope of our research to more diverse samples.

In the current thesis, I aim to broaden our knowledge of statistical learning and perceptual learning in human speech processing by including younger and older adults (> 60 years). Older adults are of particular interest for studying perceptual learning for three main reasons: First, around 20 % of the European population (Eurostat, 2017) is currently above 65 years of age. This number is steadily growing as life expectancy gradually increases. Testing older adults thus offers insights into perceptual learning processes in approximately one fifth of the (Western) population. Second, aging is accompanied by perceptual and cognitive changes, but not to the same degree for all older adults. On the one hand, adults generally show a decline in their hearing sensitivity (e.g., Lin, Thorpe, Gordon-Salant, & Ferrucci, 2011) and in their performance on most psychometric tests requiring information processing (e.g., processing speed and working memory) (e.g., Salthouse, 2009) over the course of their lifespan. On the other hand, older adults outperform younger adults on tasks that tap into crystallized knowledge and experience such as vocabulary knowledge (Verhaeghen, 2003). Yet, the speed and the degree to which these perceptual and cognitive changes occur vary greatly among older individuals. This ensures that older adults show a great variability in their performance on perceptual and cognitive tasks. This is especially important as participants for psycholinguistic research often come from self-selected samples. For example, people who are confident about their language and cognitive skills are the ones who have an interest in language research and, hence, register as participants. In older adults, this self-selection bias is partially corrected by participants' individual performance changes due to aging. Third, it is especially older adults who face the need for perceptual learning to preserve and to restore successful communication. That is, hearing loss is one of the most common chronic conditions in older adults and it is estimated that one third of older adults are affected by disabling hearing loss (World Health Organization, 2017). By the age of 80, approximately 80 % of older adults suffer from it (Lin et al., 2011). Hearing loss is commonly treated by providing hearing devices such as hearing aids or cochlear implants (depending on factors such as the cause, the type and the degree of the hearing loss). Adaptation to a new hearing device is a long-term perceptual learning process that, in case of a cochlear implant, may span several months before patients reach stable speech processing performance (e.g., Vermeire et al., 2005). Crucially, patients vary greatly in their improvement in speech perception performance after implantation (e.g., Heydebrand, Hale, Potts, Gotter, & Skinner, 2007) and the benefit they experience with their new cochlear implant (e.g., Vermeire et al., 2005). Exploring which individual abilities play a role in successful adaptation processes may help clinicians to shape realistic expectations regarding progress and may pave the way for individualized training programs in audiological rehabilitation.

Outline

The goal of this thesis is to provide new insights into how listeners are able to adapt to unfamiliar speech input even if the speech input deviates so much from everything else they have encountered in the past that immediate recognition is barely possible. Particularly, is statistical learning involved in perceptual learning in speech as implied by recent theoretical frameworks? If both statistical learning and perceptual learning in speech are supposed to be implicit learning mechanisms, does that mean that they are both unaffected by perceptual and cognitive changes due to age? Which other perceptual and cognitive resources do listeners engage to recognize an unfamiliar speech signal? Are similar resources engaged in short-term adaptation processes (e.g., over the course of a single experiment) compared to long-term adaptation processes (e.g., in hearing rehabilitation)? The results of my experiments that aim to answer these questions will inform us on the processes that underlie successful statistical learning and perceptual learning in speech and may offer important clinical implications for the rehabilitation of speech- and hearing disorders.

In Chapter 2, I focus on the ability of auditory discrimination. Auditory discrimination is the ability to hear differences between two consecutive realizations of the same word or phrase. By investigating correlates of auditory discrimination performance, I aim to investigate which perceptual and cognitive abilities may particularly be involved in the matching component of perceptual learning. That is, for being able to update and adjust the stored representations on the basis of currently perceived input and thus, to perceptually learn, the current input first has to be matched onto the stored representations. In tasks of perceptual learning these two processing components of matching and updating are inevitably linked. As such, it is impossible to determine whether specific abilities are particularly involved in either the matching or updating component of perceptual learning, or in fact in both. For auditory discrimination, a current speech input has to be matched to the auditory memory trace of a recently encountered speech input. Auditory discrimination may therefore serve as a window into the processes involved in successful matching of auditory information. Note, however, that auditory discrimination and matching in perceptual learning are not exactly the same. In perceptual learning, a realization is matched to a representation stored in long-term memory. In auditory discrimination, two auditory presented realizations are compared in short-term memory.

Auditory discrimination is also of interest in the broader perspective of human flexibility in speech as auditory discrimination initiates changes in our speech behavior. Patients in speech therapy programs, such as the Lee Silverman Voice Treatment (Sapir, Ramig, & Fox, 2011) or E-learning based Speech Therapy (Beijer et al., 2010), are frequently asked to compare their own realizations to target realizations (e.g., presented by the therapist). Adequate self-perception is key to attain modifications of speech behavior (Schroter-Morasch & Ziegler, 2005). As such, auditory discrimination of different consecutive realizations is an essential component of speech and language therapy and a prerequisite for flexibility in speech behavior. Which perceptual and cognitive resources do listeners engage to detect differences between two consecutive realizations of the same utterance? To answer this question, I applied an individual differences approach. That is, I tested a large sample of participants on a variety of auditory, cognitive and linguistic tasks and explored whether individuals' performance on the auditory discrimination task was related to their performance on some of the other tasks. I particularly addressed this question in older adults as people who receive speech and language therapy often belong to

this age group.
Chapter 3 investigates the roles of modality, stimulus type and attention in statistical learning. This chapter sets the stage for testing the relationship between perceptual learning in speech and statistical learning (Chapter 4) by first answering important questions about statistical learning. Chapter 3 assesses whether statistical learning of temporal regularities is different for visual vs. auditory presentation, whether statistical learning differs depending on the linguistic or non-linguistic nature of the stimuli, and whether attentive processing of the stimuli is required for learning to occur. All of these aspects are critical with respect to the choice of statistical learning task in relation to perceptual learning of speech (in Chapter 4). Perceptual learning in speech naturally involves presenting auditory and linguistic stimuli that must be attended. The most rigorous test for a relationship between statistical learning and perceptual learning, as reported in Chapter 4, is therefore a statistical learning task that does not use auditory linguistic stimuli but rather reflects a general ability to implicitly detect regularities in an input. Chapter 3 investigates the roles of modality, stimulus type and attention in statistical learning by comparing learning across different variants of the same artificial grammar learning - serial reaction time paradigm. Only young adults were included in the study reported on in this chapter.
Chapter 4 examines whether younger and older listeners' adaptation to acoustically degraded speech can indeed be predicted by their statistical learning ability. To that end, participants were exposed to noise-vocoded speech, an artificial speech degradation that roughly imitates the speech signal transmitted by a cochlear implant. Furthermore, Chapter 4 investigates which other cognitive and linguistic tasks (i.e., working memory, attention switching control, processing speed or vocabulary knowledge) are engaged in successful statistical and perceptual learning performance and whether the contribution of these abilities to learning changes over the life span.
Chapter 5 follows up on Chapter 4's specific finding that perceptual learning and statistical learning were correlated for younger, but not older adults. This may well be because the group of older adults did

not show learning in the visual statistical learning task whereas the younger adults did. Previous reports of older adults' inability to learn probabilistic associations from visual input (e.g., Simon, Howard, & Howard, 2011) have been taken as evidence for a more general decrease in pattern sensitivity in older age (Negash, Howard, Japikse, & Howard, 2003). If older adults indeed have generally poorer pattern sensitivity than younger adults, then older adults' statistical learning performance should also be affected in a different (i.e., non-visual) modality. Given the importance of statistical learning for language and speech processing, Chapter 5 investigates whether auditory statistical learning is impaired in older compared to younger adults.

While the previous chapters examine statistical and perceptual learning in groups of healthy younger and older adults, the last experimental chapter investigates perceptual learning in a clinical population. More specifically, Chapter 6 aims to identify cognitive abilities relating to patients' initial adaptation progress with their cochlear implant. Adaptation to the signal of a cochlear implant may be considered a form of perceptual learning as cochlear implant recipients have to adapt to a novel listening situation. However, in comparison to traditional studies on perceptual learning, participants do not have to adapt to a linguistically degraded (e.g., accented speech) or acoustically degraded (e.g, noise-vocoded or sine-wave speech) speech signal but in fact have to adjust to an improved auditory signal. Investigating adaptation processes in a hearing-impaired population particularly offers insights into the time-scale of perceptual learning processes. Novice cochlear implant users take months to reach relatively stable speech perception performance (e.g., Vermeire et al., 2005) compared to the relatively quick perceptual learning processes that are typically the object of interest in perceptual learning studies. Exploring correlates of individuals' progress in speech understanding with a new cochlear implant is of clinical relevance as it may pave the way for future individualized rehabilitation programs. Chapter 6 additionally investigates the use of more naturalistic speech material (i.e., conversational speech), as compared to more traditional material to measure speech perception performance (i.e., single words and simple sentences), for the assessment of speech understanding performance in the rehabilitation of new cochlear implant recipients. The final chapter, Chapter 7, summarizes the results and provides a general discussion of the main findings of this thesis.

References

Adank, P., & Janse, E. (2009). Perceptual learning of time-compressed and natural fast speech. *The Journal of the Acoustical Society of America, 126*, 2649-2659. doi: 10.1121/1.3216914

Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences, 8*, 457-464. doi: 10.1016/j.tics.2004.08.011

Amitay, S. (2009). Forward and reverse hierarchies in auditory perceptual learning. *Learning & Perception, 1*(1), 59-68. doi: 10.1556/lp.1.2009.1.5

Arciuli, J., & Simpson, I. C. (2012). Statistical learning is related to reading ability in children and adults. *Cognitive Science, 36*(2), 286-304. doi: 10.1111/j.1551-6709.2011.01200.x

Banai, K., & Amitay, S. (2012). Stimulus uncertainty in auditory perceptual learning. *Vision Research, 61*, 83-88. doi: 10.1016/j.visres.2012.01.009

Beijer, L. J., Rietveld, A. C. M., van Beers, M. M. A., Slangen, R. M. L., van den Heuvel, H., de Swart, B. J. M., & Geurts, A. C. H. (2010). E-learning-based speech therapy: a web application for speech training. *Telemedicine Journal and E-Health, 16*, 177-180. doi: 10.1089/tmj.2009.0104

Clayards, M., Tanenhaus, M. K., Aslin, R. N., & Jacobs, R. A. (2008). Perception of speech reflects optimal use of probabilistic speech cues. *Cognition, 108*(3), 804-809. doi: 10.1016/j.cognition.2008.04.004

Cohen, Y., Daikhin, L., & Ahissar, M. (2013). Perceptual learning is specific to the trained structure of information. *Journal of Cognitive Neuroscience, 25*, 2047-2060. doi: 10.1162/jocn_a_00453

Conway, C. M., & Christiansen, M. H. (2006). Statistical learning within and between modalities: pitting abstract against stimulus-specific representations. *Psychological Science, 17*, 905-912. doi: 10.1111/j.1467-9280.2006.01801.x

Dosher, B. A., & Lu, Z. L. (1999). Mechanisms of perceptual learning. *Vision Research, 39*(19), 3197-3221. doi: 10.1016/S0042-6989(99)00059-0

European Commission. (2012). Europeans and their Languages *Special Eurobarometer 386, Wave EB77.1* (pp. 145). Brussels: TNS Opinion & Social.

Eurostat. (2017, 15.02.2017). Population by age group Retrieved 11.03.2017, from http://ec.europa.eu/eurostat/web/products-datasets/-/tps00010

Frost, R., Siegelman, N., Narkiss, A., & Afek, L. (2013). What predicts successful literacy acquisition in a second language? *Psychological Science, 24*(7), 1243-1252. doi: 10.1177/0956797612472207

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology, 49*, 585-612. doi: 10.1146/annurev.psych.49.1.585

Hervais-Adelman, A. G., Davis, M. H., Johnsrude, I. S., & Carlyon, R. P. (2008). Perceptual learning of noise vocoded words: effects of feedback and lexicality. *Journal of Experimental Psychology. Human Perception and Performance, 34*, 460-474. doi: 10.1037/0096-1523.34.2.460

Heydebrand, G., Hale, S., Potts, L., Gotter, B., & Skinner, M. (2007). Cognitive predictors of improvements in adults' spoken word recognition six months after cochlear implant activation. *Audiology & Neuro-Otology, 12*(4), 254-264. doi: 10.1159/000101473

Kidd, E. (2012). Implicit statistical learning is directly associated with the acquisition of syntax. *Developmental Psychology, 48*, 171-184. doi: 10.1037/a0025405

Kleinschmidt, D. F., & Jaeger, T. F. (2015). Robust speech perception: recognize the familiar, generalize to the similar, and adapt to the novel. *Psychological Review, 122*(2), 148-203. doi: 10.1037/a0038695

Lin, F. R., Thorpe, R., Gordon-Salant, S., & Ferrucci, L. (2011). Hearing loss prevalence and risk factors among older adults in the United States. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences, 66*, 582-590. doi: 10.1093/gerona/glr002

Loebach, J. L., Bent, T., & Pisoni, D. B. (2008). Multiple routes to the perceptual learning of speech. *The Journal of the Acoustical Society of America, 124*, 552-561. doi: 10.1121/1.2931948

Loebach, J. L., Pisoni, D. B., & Svirsky, M. A. (2009). Transfer of auditory perceptual learning with spectrally reduced speech to speech and nonspeech tasks: implications for cochlear implants. *Ear and Hearing, 30*, 662-674. doi: 10.1097/AUD.0b013e3181b9c92d

Misyak, J. B., & Christiansen, M. H. (2012). Statistical learning and language: an individual differences study. *Language Learning, 62*, 302-331. doi: 10.1111/j.1467-9922.2010.00626.x

Nahum, M., Nelken, I., & Ahissar, M. (2010). Stimulus uncertainty and perceptual learning: similar principles govern auditory and visual learning. *Vision Research, 50*, 391-401. doi: 10.1016/j.visres.2009.09.004

Negash, S., Howard, D. V., Japikse, K. C., & Howard, J. H. (2003). Age-related differences in implicit learning of non-spatial sequential patterns. *Aging, Neuropsychology, and Cognition, 10*, 108-121. doi: 10.1076/anec.10.2.108.14462

Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology, 48*(2), 127-162. doi: 10.1016/S0010-0285(03)00128-2

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology, 47*, 204-238. doi: 10.1016/S0010-0285(03)00006-9

Petrov, A. A., Dosher, B. A., & Lu, Z. (2005). The dynamics of perceptual learning: an incremental reweighting model. *Psychological Review, 112*, 715-743. doi: 10.1037/0033-295X.112.4.715

Sabin, A. T., Clark, C. A., Eddins, D. A., & Wright, B. A. (2013). Different patterns of perceptual learning on spectral modulation detection between older hearing-impaired and younger normal-hearing adults. *Journal of the Association for Research in Otolaryngology, 14*, 283-294. doi: 10.1007/s10162-012-0363-y

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*(5294), 1926-1928. doi: 10.1126/science.274.5294.1926

Salthouse, T. A. (2009). When does age-related cognitive decline begin? *Neurobiology of Aging, 30*(4), 507-514. doi: 10.1016/j.neurobiolaging.2008.09.023

Sapir, S., Ramig, L. O., & Fox, C. M. (2011). Intensive voice treatment in Parkinson's disease: Lee Silverman voice treatment. *Expert Review of Neurotherapeutics, 11*(6), 815-830. doi: 10.1586/ern.11.43

Schroter-Morasch, H., & Ziegler, W. (2005). Rehabilitation of impaired speech function (dysarthria, dysglossia). *GMS Current Topics in Otorhinolaryngology, Head and Neck Surgery, 4*, 15. doi: 10.1002/cncr.10668

Simon, J. R., Howard, J. H., & Howard, D. V. (2011). Age differences in implicit learning of probabilistic unstructured sequences. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences, 66*, 32-38. doi: 10.1093/geronb/gbq066

Teinonen, T., Fellman, V., Naatanen, R., Alku, P., & Huotilainen, M. (2009). Statistical language learning in neonates revealed by event-related brain potentials. *Bmc Neuroscience, 10*. doi: 10.1186/1471-2202-10-21

Thompson, S. P., & Newport, E. L. (2007). Statistical learning of syntax: the role of transitional probability. *Language Learning and Development, 3*(1), 1-42. doi: 10.1080/15475440709336999

Verhaeghen, P. (2003). Aging and vocabulary score: a meta-analysis. *Psychology and Aging, 18*(2), 332-339. doi: 10.1037/0882-7974.18.2.332

Vermeire, K., Brokx, J. P., Wuyts, F. L., Cochet, E., Hofkens, A., & Van de Heyning, P. H. (2005). Quality-of-life benefit from cochlear implantation in the elderly. *Otology & Neurotology, 26*(2), 188-195. doi: 10.1097/00129492-200503000-00010

World Health Organization. (2017, February 2017). Deafness and hearing loss Retrieved 09.03.2017, from http://www.who.int/mediacentre/factsheets/fs300/en/

2

Correlates of Older Adults' Discrimination of Acoustic Properties in Speech

This chapter is based on:

Neger, T. M., Janse, E., & Rietveld, T. (2015). Correlates of older adults' discrimination of acoustic properties in speech. Speech, Language and Hearing, 18(2), 102-115. doi:10.1179/2050572814Y.0000000055.

Abstract

Auditory discrimination of speech stimuli is an essential tool in speech and language therapy, e.g., in dysarthria rehabilitation. It is unclear, however, which listener characteristics are associated with the ability to perceive differences between one's own utterance and target speech. Knowledge about such associations may help to support patients participating in speech and language therapy programs that involve auditory discrimination tasks.

Discrimination performance was evaluated in 96 healthy participants over 60 years of age as individuals with dysarthria are typically in this age group. Participants compared meaningful words and sentences on the dimensions of loudness, pitch and speech rate. Auditory abilities were assessed using pure-tone audiometry, speech audiometry and speech understanding in noise. Cognitive measures included auditory short-term memory, working memory and processing speed. Linguistic functioning was assessed by means of vocabulary knowledge and language proficiency.

Exploratory factor analyses showed that discrimination performance was primarily associated with cognitive and linguistic skills, rather than auditory abilities. Accordingly, older adults' discrimination performance was mainly predicted by cognitive and linguistic skills. Discrimination accuracy was higher in older adults with better speech understanding in noise, faster processing speed, and better language proficiency, but accuracy decreased with age. This raises the question whether these associations generalize to clinical populations and, if so, whether patients with better cognitive or linguistic skills may benefit more from discrimination-based therapeutic approaches than patients with poorer cognitive or linguistic abilities.

Introduction

Adequate self-perception is considered a prerequisite to attain modifications of speech behavior (Schroter-Morasch & Ziegler, 2005). Patients in speech or language therapy programs are frequently asked to compare their own realizations with target speech, which is presented either by the therapist or by a speaker whose speech was recorded for that purpose. Auditory discrimination of speech stimuli is thus an essential component of speech and language therapy. Auditory discrimination of non-speech stimuli has been shown to rely on auditory and cognitive abilities (e.g., Humes, 2005). However, in therapy settings, patients typically have to discriminate words and sentences. This raises the question which abilities are involved in the discrimination of speech materials. After all, if patients are unable to perceive deviations in the acoustic properties in speech, their ability to adequately modify their speech production will be limited.

Recently, an auditory discrimination test [ADT] has been developed to assess the ability of individuals with dysarthria to discriminate acoustic properties in speech (Beijer, Rietveld, & van Stiphout, 2011). Outcomes of the test are used to evaluate whether patients qualify for E-learning based Speech Therapy [EST] (Beijer et al., 2010). This speech training offers patients with dysarthria secondary to neurological impairment (Parkinson's Disease or stroke) an opportunity for intensive and independent training at home. During training, patients are presented with sentences from a database – according to a personalized protocol – and have to repeat target utterances. The patients' task is to maximize the similarity between the targets and their own realizations (Beijer et al., 2011). To assess the success of their attempt, patients can replay their own utterance and compare it to the original sentence. Obviously, in order to benefit from this form of intervention, individuals with dysarthria should be able to detect the differences between their own realizations and the target utterances, so that they can adjust their speech production accordingly. As patients' performance on the ADT is used to (contra-) indicate participation in EST, it is important to know whether any other constraints in sensory or cognitive functioning may affect discrimination performance on language materials. That is, knowledge about the association between auditory, cognitive and linguistic abilities and auditory discrimination performance is of clinical relevance, as it may provide indications to support patients participating in programs of speech and language therapy involving auditory discrimination tasks.

Individuals with dysarthria are commonly over 60 years of age. Age-related changes in hearing sensitivity (Cruickshanks et al., 1998) and cognitive abilities (for a review see Park & Reuter-Lorenz, 2009) are prevalent in this age group. As auditory discrimination is a function of auditory processing , and measures of auditory processing have previously been associated with cognitive (e.g., Fogerty, Humes, & Kewley-Port, 2010; Humes, 2005) and auditory abilities (Cox, McCoy, Tun, & Wingfield, 2008; Humes, 2005), auditory discrimination performance may already be affected in a healthy older population. Therefore, our study was set up (1) to help establish reference data for clinical populations on those subtests of the ADT that represent the most general acoustic dimensions (i.e., loudness, pitch and speech rate) and (2) to determine how strongly variability in hearing abilities and cognitive and linguistic performance is associated with individual ADT performance among healthy older adults. Knowledge about associations between discrimination of acoustic properties in speech and both auditory and cognitive abilities may guide the therapist in choosing an appropriate therapeutic approach for individual patients: for example, if working memory turns out to play a role, the length of stimuli in auditory discrimination tasks may be adjusted.

Importantly, auditory processing is commonly assessed by asking participants to detect, discriminate or identify characteristics of *unfamiliar non-speech* stimuli such as discriminating pitch of pure tones or detecting gaps in modulated noise. Although performance on these tasks has been directly linked to older adults' speech processing abilities (e.g., Papakonstantinou, Strelcyk, & Dau, 2011), thereby highlighting the importance of auditory processing in speech perception, it remains unclear whether auditory processing of non-speech stimuli taps the same abilities as processing of speech stimuli.

Kidd and colleagues (2007) included identification of nonsense syllables, words and sentences in their analysis of auditory abilities. Based on the performance of 340 participants on nineteen auditory processing tests, they extracted four distinct auditory processing abilities. All measures regarding identification of speech stimuli in broadband noise loaded on a common factor, together with a task of familiar environmental sound identification. Kidd and colleagues (2007) concluded that the common factor described a distinct ability of auditory processing of *familiar* sounds including speech. As this factor was clearly distinct from the other auditory processing abilities (i.e., a loudness-duration factor reflecting sensitivity to changes in overall energy; an amplitude modulation factor resembling sensitivity to brief changes in level; and a pitch and time factor resembling sensitivity to patterns affecting longer stretches of sound), correlates of auditory processing of non-speech stimuli may not necessarily generalize to auditory processing tasks making use of speech materials. Given the differences between speech and non-speech stimuli - e.g., the duration of non-speech stimuli in discrimination tasks ranges between 100 and 250 ms (Kidd et al., 2007) whereas words and sentences span several hundreds of milliseconds - processing of speech and non-speech stimuli may indeed (partially) require different abilities.

The present study was therefore set up to investigate which cognitive processes are involved in auditory discrimination of acoustic properties in speech. As the first, anchor stimulus of a trial has to be stored over a brief period of time until its internal representation can be compared to the second, contrasting stimulus, memory skills are likely to come into play. Auditory short-term memory is needed for the immediate recall of auditory information (Gathercole, 1999). Moreover, working memory is required to simultaneously store and process auditory information, which is necessary to compare the two, anchor and contrast, stimulus (Gathercole, 1999). The high incidence of auditory processing deficits in patients with memory dysfunctions (e.g., Idrizbegovic et al., 2011) supports the idea that memory processes relate to auditory discrimination.

Furthermore, processing speed has been found to explain individual variation in temporal auditory processing (Harris, Eckert, Ahlstrom, & Dubno, 2010) and may also relate to auditory discrimination of speech characteristics . In speech, auditory temporal processing enables listeners to differentiate between phonemes such as stop consonants (Pichora-Fuller & Souza, 2003). Similarly, age-related declines in auditory temporal processing have been related to older adults‘ speech perception difficulties (Pichora-Fuller & Souza, 2003; B. Schneider & Pichora-Fuller, 2001), such as problems in discriminating certain phonemic contrasts that are distinguishable on the basis of durational cues such as the length of gaps (B. Schneider & Pichora-Fuller, 2001). Additionally, language proficiency and vocabulary knowledge may relate to the ease with which auditory speech discrimination can be performed as auditory processing involves two subtasks in this context: processing the speech stimuli and comparing the anchor with the contrast stimulus. Both subtasks demand cognitive processing. High language proficiency and better vocabulary knowledge may facilitate processing of the linguistic content, such that less cognitive resources are needed to perform the first subtask. Moreover, most older adults listen to a partially degraded auditory signal as they experience age-related hearing loss. Vocabulary knowledge has been shown to support comprehension of degraded speech such as dysarthric (McAuliffe, Gibson, Kerr, Anderson, & LaShell, 2013), noise-vocoded (see Chapter 4) or accented (Janse & Adank, 2012) speech. Consequently, more cognitive resources might be available to perform the second subtask of stimuli comparison, which is essential for auditory discrimination.

In sum, this study investigates how measures of auditory (i.e., hearing threshold, speech understanding in quiet and noise), cognitive (i.e., auditory short-term memory, working memory, processing speed) and linguistic abilities (i.e., vocabulary knowledge, language proficiency) are associated with auditory discrimination of acoustic properties in speech in healthy older people.

Method

Participants

Sufficient audibility is a prerequisite for auditory discrimination. Therefore, individuals with severe hearing loss in at least one ear were excluded from the current study. Applying the definition of the World Health Organization (2014), we classified severe hearing loss as a pure-tone average [PTA] over 0.5, 1, 2 and 4 kHz of at least 61 dB HL. As approximately 95% of older adults have a PTA up to 60 dB HL in both ears (Cruickshanks et al., 1998), hearing sensitivity observed in the present sample was representative of the hearing level typically observed in older adults (see below for more details). The initial sample of participants consisted of 97 native speakers of Dutch. One participant was excluded because of a recent stroke. The remaining 96 participants were neurologically intact and had no history of speech or language disorders except for one, who had received stuttering therapy in childhood. Participants were aged between 60 and 84 years and two-thirds were female. Participants‘ level of education was indicated on a 5-point-scale: primary school was coded as 1, secondary education was coded as 2, technical and vocational training for teenagers from 16-18 years old was coded as 3, upper technical and vocational training for adults was coded as 4, and university education was coded as 5. Most of the participants had followed upper technical vocational training for adults. Participant information including age, educational level and performance on all background tasks can be found in Table 2.1.

Table 2.1. Participant information, including age, educational level and performance on all tasks.

	M	*SD*	*Min*	*Max*
Discrimination accuracy (%)	94.95	5.6	71.11	100
Discrimination speed (s)	1.88	0.21	1.29	2.65
Hearing acuity (dB HL)	23	10.4	5	46.7
Speech reception in quiet	13.53	7.6	0	39
Speech reception in noise	-2.00	2.19	-5.43	4.50
Auditory short-term memory	36.53	6.32	17.12	45.98
Working memory (%)	47.4	18.31	8.33	100
Processing speed	48.88	10.6	27	70
Language proficiency	45.68	14.24	4	72
Vocabulary	0.87	0.06	0.63	0.98
Age (years)	68.5	5.7	60	84
Education level	4*	---	1	5

* median

Auditory measures

Pure tone thresholds

All experimental tasks were conducted in an sound-attenuating booth (Amplisilence) to minimize distraction. Peripheral auditory function was assessed by measuring air-conduction pure-tone thresholds with a PC-based diagnostic audiometer (Oscilla USB-300). Pure-tone thresholds were determined at 0.25, 0.5, 1, 2, 4 and 8 kHz in both ears. Higher thresholds reflected poorer hearing. Mean thresholds are given in Figure 2.1.

As age-related hearing loss particularly affects sensitivity to the higher frequencies, a high-frequency pure-tone average [PTA_{high}], calculated as the mean hearing threshold over 1, 2 and 4 kHz, was taken as an index of hearing acuity. According to the standard of insurance coverage in the Netherlands (PTA_{high} of one ear ≥ 35 dB HL), thirty-four older participants actually qualified for partially refunded hearing aids on the basis of their hearing thresholds. However, at the time of testing, none of the participants wore hearing aids in daily life (cf. Gopinath et al. (2011) on hearing aid uptake). To assess whether variation in audibility affects discrimination performance, we kept the presentation level of the auditory stimuli in our experiment constant. Twenty out of the 96 participants showed asymmetrical hearing loss, which was defined as a between-ears PTA difference (PTA calculated over 0.5, 1, 2 and 4 kHz) exceeding 10 dB HL (Noble & Gatehouse, 2004). Although hearing problems in these participants may not originate from age-related changes in the peripheral auditory system, we did not exclude listeners with asymmetrical hearing loss from the current study, as we wanted to assess whether common variability in older adults' auditory abilities (including speech reception thresholds, see below) affects individual ADT performance. Only the PTA_{high} of the best ear was considered in the analysis, as all auditory stimuli in the experimental tasks were presented binaurally via dynamic closed, circumaural headphones (Sennheiser HD 215).

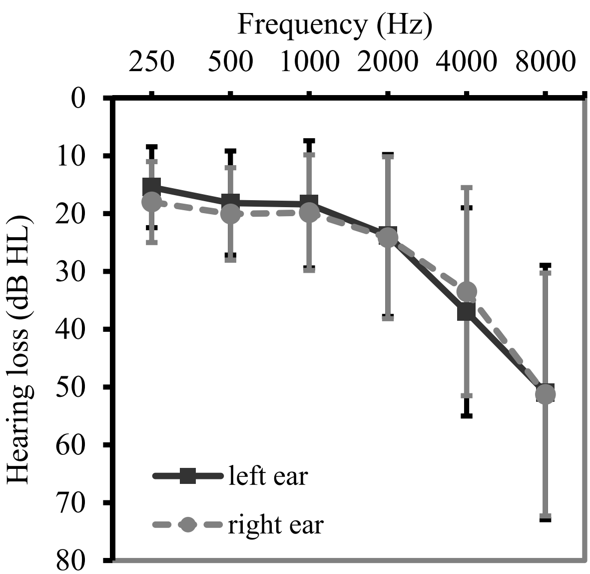

Figure 2.1. Average pure tone thresholds at 0.25, 0.5, 1, 2, 4 and 8 kHz for both ears (n = 96). Error bars represent one standard deviation of the mean.

Speech reception in quiet

As a second measure of participants' auditory abilities, we assessed participants' speech audiometry performance. The minimum hearing level at which a participant could understand 50% of phonetically balanced Dutch words (Bosman, 1989) was measured with the aid of an audiometer (Interacoustics Equinox$^{2.0}$ AC 440) and a noise-cancelling audiometric headset (Telephonics TDH 39 with Amplivox Audiocups). The speech reception threshold in quiet was indicated in terms of dB HL. Therefore, higher scores reflected poorer ability of participants to identify speech in quiet. As for the pure-tone thresholds, only the speech reception threshold of the better ear was considered in the analysis.

Speech reception in noise

Older adults frequently report difficulties in understanding speech in noisy conditions even if their hearing is clinically normal (Committee on Hearing, Bioacoustics and Biomechanics [CHABA], 1988). Therefore, speech understanding in noise adds important information on the process of auditory aging, specifically in those participants with asymmetrical hearing loss. Moreover, a measure of speech perception ability in noise may be of particular relevance given that patients following EST are likely to perform auditory judgments about speech in everyday noisy conditions.

In the current study, the speech reception threshold in noise was defined as the signal-to-noise ratio in dB at which listeners can correctly repeat 50% of keywords in sentences that are presented in speech-shaped noise. To determine the SRT, we applied an adaptive staircase procedure. That is, individual performance was kept around 50% correct by continually changing the noise level depending on the accuracy of the participant's previous response. In total, the task consisted of 15 test sentences each containing four keywords. All sentences were taken from standard Dutch-language audiology materials (Plomp & Mimpen, 1979). Participants were asked to repeat all the words in a sentence they had understood. Participants were encouraged to guess. After each trial, the number of correct responses (ranging from 0 - 4) was immediately scored online by an experimenter.

Cognitive measures

With the exception of auditory short-term memory, which is an auditory task in nature, all tasks to measure linguistic and cognitive skills were presented visually. This was done to avoid confounding effects of auditory sensitivity on older adults‘ performance on linguistic and cognitive tasks.

Auditory short-term memory

Performance on an auditory nonword repetition task is generally used as an index of verbal/auditory short-term memory (e.g., Gathercole, 1999). Note, however, that nonword repetition depends on accurate perception of stimuli and, consequently, also reflects differences in auditory abilities. In the auditory short-term memory task, participants were instructed to repeat nonwords. Each nonword was presented once at a fixed mean intensity of 80 dB SPL. Stimuli were provided via E-prime 1.2 (W. Schneider, Eschman, & Zuccolotto, 2002). The speaker was a female native speaker of Dutch. All items had been produced at a slow and clear speaking rate. The task consisted of 50 nonwords that were phonotactically legal in Dutch and varied in length from two to five syllables (including two practice trials). Nonwords were presented in mixed order (regardless of length), but the order was kept constant for all participants. Time between consecutive trials was 3 s. Verbal responses were digitally recorded and rated by a single native speaker of Dutch. If a nonword was repeated correctly, the item received a rating of 1. If not, the proportion of correctly repeated syllables was calculated. Consequently, participants could achieve a maximum score of 48. Higher scores reflected better performance. To assess reliability of the rating of the non-word repetition performance, we asked a second evaluator to rate a random subsample of 15 audio recordings (i.e., to rate task performance of 15 participants) from our participant pool. Tukey's test of nonadditivity showed no interaction between raters and subjects ($F = 1.69$, $df_1 = 1$, $df_2 = 13$, $p = .216$), indicating that the reliability analysis should yield unbiased results. Scoring of the nonword-repetition task had an excellent inter-rater reliability as assessed by the average measure intra class correlation coefficient in a two-way random model ($ICC = .968$, $df_1 = 14$, $df_2 = 14$, $p < .001$).

Working memory

Participants performed a digit span backward task as a measure of working memory capacity. The test was a computerized variant of the digit span backward task included in the Wechsler Adult Intelligence Scale Test (2004) and presented via E-prime 1.2 (W. Schneider et al., 2002). Participants were asked to report a series of visually presented digits back in reverse order; for example, if they received the sequence 5-9-2, participants had to type in 2-9-5. Digits were presented in black font (Arial, font size 100) against a white background. Each digit was presented for 1 s with an interval of 1 s between the consecutive digits of a series. The length of sequences increased stepwise from two to seven digits and performance on each length was tested by two different trials. The start of the measurement was preceded by two trials with a sequence-length of three to familiarize participants with the task. In total, participants had to recall 12 test sequences and individual performance was measured by the proportion of correctly reported sequences (out of 12).

Information processing speed

Information processing speed was assessed by means of a digit-symbol-substitution task, which was derived from the Wechsler Adult Intelligence Scale Test (2004). The digit-symbol-substitution task is a paper-and-pencil test in which participants have to convert as many digits as possible into assigned symbols (1 = '–'; 2 = '⊥'; 3 = '⊐'; 4 = '∟'; 5 = '⊔'; 6 = '○'; 7 = 'Λ'; 8 = '×'; 9 = '=') in a fixed amount of time. For example, in the box below a ‚9', participants have to fill in the ‚=' symbol. The digit-symbol key was printed at the top of the page and visible to the participant throughout the task. In total, the task consisted of 140 items of which the first seven items served as practice trials. Performance was measured by the number of correctly converted digits in 90 seconds, so that higher scores reflected higher information processing speed.

Linguistic measures

Language proficiency

A cloze test was administered to obtain a global measure of language competence. Cloze tests are assumed to measure integrative language proficiency, because knowledge of vocabulary, grammar, sentence structure, text structure and cohesion have to be integrated to perform such a task (e.g. Hanania & Shikhani, 1986). Our paper-and-pencil test consisted of three short texts, which did not overlap in content. Readability index scores of the three different paragraphs indicated that two paragraphs were relatively easy and that one was relatively difficult. Paragraphs were printed in order of ascending reading difficulty to create a measure of language proficiency that took reading speed as well as reading comprehension into account. The first and third cloze text contained 13 blanks and the second included 14 blanks. In total, 40 content words had been removed and participants were asked to fill in the blanks. Participants had to fill in one word for each missing item and they were free to choose appropriate words. They were only instructed not to reuse words. Participants were given five minutes to complete the task.

The maximum total score on the cloze test was 80 points. Participants obtained two points per item if the selected word matched the grammatical and semantic structure of the sentence. If a response was in line with either the grammatical or semantic context, it received a one-point credit. Items were scored as wrong (0 points) if more than one word had been inserted, a word had been reused or blanks had been left empty. All tests were rated by one single rater. To assess the reliability of the scoring of the cloze test, we had a second rater score a subset of 15 randomly selected tests. Tukey's test of nonadditivity was non-significant ($F = 4.10$, $df_1 = 1$, $df_2 = 13$, $p = .064$), indicating no interaction between raters and subjects and, hence, unbiased results of the reliability analysis. The scoring had a very high reliability assessed by the average measure intraclass correlation coefficient in a two-way random model ($ICC = .996$, $df_1 = 14$, $df_2 = 14$, $p < .001$).

Vocabulary

A vocabulary test in the form of multiple choice questions was administered to obtain a measure of linguistic knowledge (Andringa, Olsthoorn, van Beuningen, Schoonen, & Hulstijn, 2012). The computerized test was administered in Excel (Courier font size 15). Participants had to indicate which out of five possible answers was the correct meaning of Dutch low-frequency words, the last alternative always being ‚I don't know'. Words were not domain-specific and each target word was embedded in a different, neutral carrier phrase. The vocabulary test consisted of 60 items. There was no time limit or pressure to complete the test. Performance was measured by test accuracy, that is, the proportion of correct answers. Higher scores thus reflected greater vocabulary knowledge.

Speech discrimination test

Materials

We administered a shortened version of the auditory discrimination test [ADT] (Beijer et al., 2011). The original ADT consists of five subtests, each aiming to measure auditory discrimination of a specific speech dimension relevant for speech therapy, particularly the dimensions that are trained in E-learning based Speech Therapy (EST). Subtests include discrimination of speech segments, intensity, overall pitch, speech rate and intonation. EST focuses on these speech dimensions in order to improve speech intelligibility and to avoid voice strain, according to the principles of the Pitch Limiting Voice Treatment (de Swart, Willemse, Maassen, & Horstink, 2003) which is adapted from the Lee Silverman Voice Treatment (Ramig, Countryman, O'Brien, Hoehn, & Thompson, 1996) (for more details see Beijer et al., 2011). The shortened test included three out of five subtests, that is the subtests on ‚loudness', ‚overall pitch' and ‚speech rate' corresponding with the perception of the amplitude, frequency and time domains in speech. We excluded the subtests ‚segmental elements' and ‚intonation' because both subtests match specific therapy goals for dysarthria. Although these aspects may be clinically important, they were considered less general than the three dimensions loudness, pitch and speech rate[1].

The ADT consists of simple 'same-different' decisions. Per trial, participants hear two realizations of a word or sentence and have to indicate whether these are same or different by clicking on buttons that are marked on a keyboard (green = ‚same', red = ‚different'). The ADT contains two versions, one for men and women respectively. Versions differ only with respect to speaker gender in the subtest ‚overall pitch'. That is, male participants listen to a 66-year-old male speaker and female participants to a 62-year-old female speaker for this specific subtest. Both speakers are native speakers of Dutch. Speech rate has been equalized for all utterances to avoid speaker-intrinsic speech rate differences. Speech materials that are used in the subtests loudness and speech rate have been produced by the male speaker. The subtests loudness and overall pitch each contain seven words and eight sentences whereas discrimination of speech rate is only assessed with sentences. Sentences stem from screening materials typically used in auditory testing in the Netherlands that are of acceptable length (seven to nine syllables) and cover neutral semantic content (Plomp & Mimpen, 1979). Word stimuli in the ADT are commonly known Dutch words that occur with a minimum frequency of 50 in the CELEX corpus of written Dutch (H. Baayen, Piepenbrok, & Gulikers, 1995; Beijer et al., 2011).

1 The subtest 'segmental elements' focuses on the discrimination between a regular vowel or consonant and a deviant counterpart typically found in dysarthric speech (e.g. difference between regular vowel [a:] and a sound in between [a:] and [ə]). As monopitch is characteristic of (hypokinetic) dysarthric speech, the subtest intonation focuses on the discrimination of different versions of an F_0 peak, indicating different amounts of relative prominence of a sentence accent.

Each subtest is built up of 15 items: seven equal and eight unequal pairs. Unequal pairs are based on slightly adjusted Just Noticeable Differences (JND). That is, unequal test items were constructed to be detectable as different in at least 80% of the cases. In the subtest ‚loudness', mean intensity of stimuli has been adjusted to 65 dB. To create unequal pairs, mean intensity of the second stimulus has been raised to 71 dB (74 dB in one case). In the subtest ‚overall pitch', changes in pitch can be higher or lower. Higher pitch versions of a stimulus have been created by raising the original pitch by four semitones (three semitones in the female version). Naturally sounding pairs of lower pitch have been created by raising the first (default) stimulus by two and lowering the second stimulus by five semitones (four semitones in the female version) (for further details on the stimuli see Beijer et al., 2011). In the subtest ‚speech rate', natural sounding manipulations have been derived by asking the speaker to produce stimuli in a slightly faster or slower speech rate. These natural variations in speech rate have then been increased by additionally accelerating or decelerating the overall rate by 5%.

Procedure

The ADT was carried out in E-prime 1.2 (W. Schneider et al., 2002) and followed the procedure described in Beijer et al. (2011). The ADT is a self-paced test and decisions can be made before the end of the second realization. Once a decision is made, presentation of the second stimulus is terminated if needed, and the next trial starts after 2500 ms. The interstimulus interval between the two realizations of a word or sentence is 200 ms. Each subtest starts with an information screen to highlight the specific speech dimension that is going to be discriminated, after which two practice trials follow. In the practice trials, participants receive feedback whether the stimuli differ and if so, in which direction (e.g., for pitch: ‚The second item was lower than the first item'). All subtests and items are presented in a fixed order. The subtest ‚loudness' is followed by the subtests ‚overall pitch' and ‚speech rate'. Within each subtest, words are discriminated before sentences. Participants took approximately four minutes to complete a subtest. Overall, ten to fifteen minutes were sufficient to accomplish the shortened version of the ADT. Because one of the subtests assessed loudness discrimination, the presentation level of the stimuli ranged between 75 and 80 dB SPL. Presentation of stimuli in the subtests ‚overall pitch' and ‚speech rate' was set to a fixed volume of 75 dB SPL, ensuring sufficient audibility for all participants. All stimuli were presented binaurally via dynamic closed, circumaural headphones (Sennheiser HD 215).

Two different measures of performance were derived from the ADT: the overall percentage of correctly discriminated items (equal pairs stated as ‚same', unequal pairs stated as ‚different') as an index of *discrimination accuracy*, and reaction times for all correct same-different decisions as a measure of *discrimination speed*. We considered discrimination accuracy to be the most appropriate estimate to explore how speech discrimination ability relates to auditory and cognitive functioning in a factor analysis. Additionally, to investigate speech discrimination performance in more detail, we analyzed both discrimination accuracy and discrimination response times in a regression analysis. The latter analysis method allowed us to investigate whether performance differed for the subtests, and whether certain abilities were more predictive of performance in one subtest, compared to another. Note that we administered the ADT exactly as it is used in the clinical setting. That is, participants were informed that they could respond before the offset of the second stimulus, however, they were not specifically instructed to respond as quickly as possible. Therefore, participants' reaction times may not accurately reflect response speed and must be interpreted with caution. Still, although participants could wait for the second stimulus to finish, all participants responded before stimulus offset at least incidentally (52.1% of the responses to the sentence stimuli were given before stimulus offset). This suggests that the response times reflected the moment in time at which participants were confident of their discrimination decision. Discrimination accuracy was determined for overall performance. Discrimination speed was defined as the time difference between the onset of the second realization and the subsequent button press. Higher RTs indicated slower performance.

General procedure

Measurements were collected in two experimental sessions. Between both sessions was an interval of two months, because the measures were administered in the context of other studies. Measures of hearing acuity, auditory short-term memory, working memory, information processing speed and vocabulary were obtained during the first experimental session. Tasks of auditory discrimination, speech reception threshold in quiet and in noise and language proficiency were carried out during the second experimental session. Before the start of each task, participants received verbal and printed task instructions. Specifications of age and education level were available from the subject database of the Max Planck Institute for Psycholinguistics from which participants were recruited. All tasks were presented in a fixed order. Participants were compensated €8 per hour for their participation.

Data analyses

The study set out to explore the associations between auditory speech discrimination on the one hand and auditory, cognitive and linguistic abilities on the other. As a first step, we explored the correlations between auditory discrimination accuracy and auditory, cognitive and linguistic measures. As a second step, we determined the general structure underlying all measures by means of exploratory factor analysis. By assessing which measures load on the same latent constructs, we aimed to identify whether auditory discrimination of acoustic properties in speech is generally more closely related to auditory, cognitive or linguistic abilities, without targeting any a priori hypothesis. We used principal axis factoring as the measured variance in our data is a combination of shared and error variance. As education level was measured on an ordinal scale, we tested whether the use of polychoric correlations would be required. Education level had a skewness of -0.71 and a zero-centred kurtosis of -0.095. These indices were well below the critical value of 1 (in absolute value) (Muthén & Kaplan, 1985), such that the use of Pearson‘s correlation coefficients was warranted. Varimax rotation was applied in order to facilitate interpretation and labeling of the factors. Factor loadings can be considered as correlation coefficients, however, the standard error of factor loadings is greater than that of a typical correlation coefficient (Cliff & Hamburger, 1967). Therefore, we adopted a conservative alpha level of 1% to assess the significance of factor loadings. Given a sample size of 96 and a significance level of .01, only factor loadings greater than .262 were considered significant.

Before we conducted the factor analysis, we assessed whether the criteria for factor analysis were met. We tested for possible influential cases by estimating Cook‘s distance between auditory discrimination performance and each of the significantly correlated background measures. That is, Cook‘s distance was calculated on the basis of a linear regression model in which discrimination accuracy was predicted as a function of one background measure at a time. In all cases, Cook‘s distance was below the critical value of 1 (Cook & Weisberg, 1982), indicating no unduely influential participants. One data point of education level was missing and was replaced by the corresponding mean. Factorability of items was assessed by screening the one-tailed intercorrelations between variables. All measures correlated significantly with at least four other variables while no correlation was higher than 0.8, therefore implying reasonable factorability. Factorability of items was confirmed by the determinant of the correlation matrix. A determinant value of 0.17 (well above the critical value of .00001) suggested no singularity between included variables. Furthermore, Bartlett's test of sphericity was significant (χ^2_{55} = 337.250, p < .001), such that the assumption of uncorrelated measures in the analysis could be rejected. Sampling adequacy was verified by a Kaiser-Meyer-Olkin value of 0.79, indicating that factor analysis should yield distinct and reliable factors. Further, inspection of the anti-image matrix showed that sampling adequacy was sufficient for each given pair of variables as each pair had a value above the critical value of 0.5, such that the factor analysis could be conducted with all eleven variables included.

In a third step of the analysis, we investigated which individual auditory, cognitive and linguistic abilities predicted listeners‘ (a) discrimination accuracy and (b) discrimination speed by means of multiple regression analysis. In the linear regression analysis, we took into account that the associations between discrimination performance and individual abilities may be modulated by speech dimension (pitch, loudness, or speech rate) or length of speech stimuli (words vs. sentences). As the subtest speech rate only consisted of sentences, we first conducted an analysis to investigate possible interactions between background measures and the three speech dimensions, and we then conducted an analysis to investigate possible interactions between background measures and stimulus length.

Linear regression models are based on the assumption that predictors included in the analysis do not show collinearity (R. H. Baayen, 2012). This assumption was obviously not met as many predictor measures in the current study were intercorrelated (see section 3.2). However, if the aim is to assess the unique explanatory power of a predictor beyond that of other predictors, simultaneous inclusion of correlated measures in the analysis has been argued to be the way to proceed (e.g., Wurm & Fisicaro, 2014).

We implemented linear mixed-effects models using the lmer function from the lme4 package (Bates, Maechler, & Bolker, 2012) in R, as linear mixed-effect models can contain multiple independent variables including categorical and continuous predictors at the same time (Cunnings, 2012). Moreover, mixed-effect models allow for the inclusion of both participants and items as crossed random factors and, by including random slopes, it can additionally be assessed whether any fixed effects generalize over items/participants (R.H. Baayen, Davidson, & Bates, 2008). In the discrimination accuracy analysis, logistic regression was used (as an answer could be correct or not).

To identify the most parsimonious model explaining individual discrimination performance, we applied the same model fitting process across all regression analyses. First, we explored whether speech dimension (respectively stimulus length) had a general impact on discrimination performance by inserting subtest (respectively stimulus length) as a fixed categorical variable into the initial model.

In models predicting discrimination speed, we also included stimulus duration (in ms) as a control variable. In the second step, all predictor measures and their interactions with subtest (respectively stimulus length) were added simultaneously to the model. We then analyzed our data by means of a backward stepwise selection procedure, in which first interactions and then predictors were removed if they did not attain significance at the 5% level. Each change in the fixed effect structure was evaluated in terms of loss of model fit by means of a likelihood ratio test.

After we had determined the most parsimonious fixed-effect structure, we identified the maximal random slope structure to allow for the fact that different participants and different items may vary with regard to how sensitive they are with respect to the variable at hand (Barr, Levy, Scheepers, & Tily, 2013). Determining the maximal random slope structure reduces the probability of a type I error (Barr et al., 2013): if, e.g., speech dimension only matters for some, but not generally across participants, it should be removed from the fixed effect structure. First, the effect of subtest (respectively stimulus length) was allowed to vary by participant. Second, by-item slopes were tested for all predictors in the preliminary model. Changes in the random-slope structure were evaluated by means of changes in the Akaike Information Criterion (AIC). The model with the lower AIC value and, therefore, better model fit was retained. If fixed effects did not retain significance after the maximal random slope structure had been defined, these were excluded from the fixed-part of the model following the backward stepwise procedure. Results of the most parsimonious models are indicated in absolute effect sizes (*beta*), standard errors, *t* values (or *z* values in logistic regression) and *p* values. As it is currently unclear how to calculate the proportion of the variance that is accounted for by mixed-effect models with random slopes, no measure of explained variance can be reported.

The statistical software package SPSS (version 20.0.0.1) was used to obtain descriptive statistics, correlations, and to run the factor analysis. The R statistical program (version 2.15.1) was used to run linear regression models.

Results

Overall, participants achieved high accuracy scores on the auditory discrimination test. Mean proportion of correctly discriminated items was 94.95% (*SD* = 5.6), showing relatively little variation in overall performance (coefficient of variation = 0.06). On average, participants took 1.88 s (*SD* = 0.64) to correctly discriminate items, indicating reasonable variability in discrimination speed (coefficient of variation = 0.34). Mean discrimination performances in terms of accuracy and efficiency are displayed in Table 2.2 for the different stimulus lengths and subtests.

Table 2.2. Mean discrimination performance per stimulus length and subtest.

			Discrimination accuracy (in % correct)		Discrimination speed (in ms)	
		n	*M*	*SD*	*M*	*SD*
Stimulus length	Words	14	94.87	7.94	1338.73	578.64
	Sentences	31	94.99	5.94	2125.28	506.01
Subtest	Loudness	15	97.08	7.29	1670.02	554.74
	Overall pitch	15	93.82	9.29	1742.84	661.63
	Speech rate	15	93.96	7.98	2228.86	556.47

Associates of discrimination performance

Pearson correlation coefficients were obtained for all pairs of measures and coefficients are displayed in Table 2.3. Overall accuracy on the auditory discrimination task correlated significantly with measures of speech understanding in noise, language proficiency, information processing speed and age, and marginally with measures of working memory and vocabulary knowledge. We calculated partial correlations for discrimination accuracy and language proficiency, as discrimination performance and language proficiency were both significantly associated with processing speed, age and education. As the partial correlation between auditory discrimination and language

proficiency remained significant (r_{90} = .174, p = .048), auditory discrimination of acoustic properties in speech seems to have a consistent linguistic component.
With respect to intercorrelations between predictor measures, moderate correlations were found between nonword repetition and measures of auditory abilities. This supports the idea that nonword repetition, which is supposed to index memory abilities, also reflects

Table 2.3. Pearson's correlation coefficients between auditory, cognitive and linguistic measures (n = 96).

Measure	1.	2.	3.	4.	5.
1. Discrimination accuracy	1				
2. Hearing sensitivity	-.163	1			
3. Speech threshold (quiet)	-.147	.679**	1		
4. Speech threshold (noise)	-.265**	.622**	.561**	1	
5. Auditory memory	.097	-.588**	-.532**	-.555**	1
6. Processing speed	.255*	-.214*	-.198	-.302**	.121
7. Working memory	.194	-.151	-.134	-.234*	.280**
8. Language proficiency	.333**	-.307**	-.247*	-.433**	.247*
9. Vocabulary	.198	-.094	-.066	-.099	.203*
10. Age	-.238*	.394**	.388**	.502**	-.329**
11. Education[a]	.201	-.050	.012	.025	.025

Measure	6.	7.	8.	9.	10.
6. Processing speed	1				
7. Working memory	.266**	1			
8. Language proficiency	.496**	.164	1		
9. Vocabulary	.241*	.241*	.391**	1	
10. Age	-.501**	-.240*	-.340**	-.005	1
11. Education[a]	.178	.142	.267**	.488**	-.023

Note. [a] Spearman's rho, * $p < .05$, ** $p < .01$

auditory functioning if participants differ in hearing sensitivity. Less obvious correlations were observed between hearing thresholds in quiet and the measure of language proficiency and between hearing sensitivity and information processing speed. These intercorrelations disappeared after controlling for age. Age was significantly correlated to all measures except vocabulary knowledge and education level.

Discrimination performance in relation to auditory, cognitive and linguistic skills (factor analysis)

Three factors were extracted based on Kaiser's criterion of eigenvalues larger than 1. The number of latent constructs was in agreement with visual inspection of the scree plot. All factors together explained 49.5% of the variance. The factor loadings used to label the factors are given in Table 2.4. The first factor accounted for 22.4% of the total variance and was labeled „auditory functioning". The second factor accounted for 14.0% of the total variance and was labeled „processing efficiency". The third factor was considered to represent „linguistic skills" and accounted for 13.1% of the total variance.

Auditory discrimination performance showed cross loadings on the factors processing efficiency (= .316) and linguistic skills (= .323) but did not load on the factor of auditory functioning. This suggests that discrimination of acoustic properties in speech is generally more closely related to cognitive and linguistic processing than to auditory functioning.

However, two participants in the sample scored below the 80 % criterion in the auditory discrimination test (71.1 % and 73.3 % accuracy) and, thus, more than three standard deviations below the mean. Although our pre-analysis indicated no undue influential participants (see section on data analyses), we reran the factor analysis excluding both participants, as they may be considered as outliers. Exclusion of the two participants resulted in a loss of the significant relationships between discrimination accuracy and the background measures. Hence, these participants were driving the loadings of discrimination accuracy on the factors of ‚processing efficiency' and ‚linguistic skills'. As factor analysis is prone to the influence of outliers (with only one data point per subject), we applied mixed-effect models in a next step to assess the associations between discrimination performance and the background measures in more detail. Mixed-effect models are based on individual observations within a task and, hence, several data points are included per participant. Moreover, by including random slopes, we allowed for the possibility that participants and/or items were differentially affected by a given predictor. This approach is conservative and prevents effects to be driven by single participants or single items.

Table 2.4. Summary of principal factor analysis after varimax rotation (n = 96).

	Auditory functioning	Processing efficiency	Linguistic processing
Hearing (PTA_{high})	**.806**	-.169	-.078
Speech reception quiet	**.746**	-.171	-.003
Auditory short-term memory	**-.730**	.065	.127
Speech reception noise	**.692**	**-.380**	-.089
Age	**.372**	**-.659**	.048
Processing speed	-.079	**.711**	.238
Language proficiency	-.235	**.475**	**.433**
Auditory discrimination	-.108	**.316**	**.323**
Vocabulary	-.071	.032	**.771**
Education	.043	.127	**.635**
Working memory	-.181	.225	.252
Eigenvalues	2.468	1.540	1.442
% of variance explained	22.438	13.998	13.106

Note. Factor loadings > .262 are in boldface

Predictors of discrimination performance (linear regression analysis)

Overall, participants produced 4102 correct and 218 incorrect responses. Reaction times (RTs) were analyzed for accurate trials only. Valid RTs were restricted to those within three standard deviations from the mean RT (of all correct responses). Therefore, the final data set for predicting discrimination accuracy consisted of 4320 observations and the final data set for predicting discrimination speed consisted of 4076 observations.

Older adults' discrimination accuracy was modulated by speech dimension (cf. Table 2.5A): Listeners made more errors in discriminating pitch and speech rate, compared to their performance on the subtest intensity (placed on the intercept). Moreover, listeners who could perceive speech at less favorable signal-to-noise ratios tended to be better able to perceive acoustic differences in the speech stimuli. None of the other measures predicted discrimination accuracy or interacted with subtest. The maximal random slope structure showed that listeners differed in the degree to which their discrimination accuracy was affected by changes in speech dimension.

In contrast to speech dimension, stimulus length had no effect on discrimination accuracy (Table 2.5B). That is, listeners did not make more errors in discriminating sentences than in discriminating words across the subtests intensity and pitch. Overall, the only predictor of discrimination accuracy was language proficiency: the better listeners' language proficiency, the higher was their discrimination accuracy.

The most parsimonious model predicting response times on the basis of speech dimension included duration, speech dimension, information processing speed and an interaction between listeners' speech perception threshold in quiet and speech dimension (Table 2.6A). As expected, the longer the stimuli, the longer were the RTs. Moreover, listeners were slower in discriminating speech rate than in discriminating intensity, which makes sense as a longer stretch of speech is needed to evaluate speech rate than intensity. This effect was modulated by individuals' speech reception threshold in quiet, suggesting that listeners with better speech recognition in quiet sped

Table 2.5A. Statistical model for discrimination accuracy as a function of subtest. Standard error is indicated by *SE*.

	Subtest analysis				
Fixed effects	*β*	*SE*	*z*	*p*	
Intercept	4.83	0.36	13.31	< .001	
Pitch	-0.98	0.47	-2.09	.037	
Speech rate	-1.38	0.46	-3.02	.003	
Speech in noise	-0.10	0.06	-1.86	.063	
Random effects		*Variance*	*SD*	*Corr*	
Subject	Intercept	2.14	1.46		
	Pitch	2.81	1.68	-.657	
	Speech rate	2.35	1.53	-.797	.785
Item	Intercept	0.83	0.91		

Note. *n.s*. = p > 0.1

Table 2.5B. Statistical model for discrimination accuracy as a function of stimulus length. Standard error is indicated by *SE*.

	Stimulus length analysis			
Fixed effects	*β*	*SE*	*z*	*p*
Intercept	3.99	0.34	11.88	< .001
Sentences	-	-	-	*n.s.*
Language proficiency	0.04	0.01	3.46	< .001
Random effects		*Variance*	*SD*	
Subject	Intercept	1.04	1.02	
Item	Intercept	0.97	0.99	

Note. *n.s*. = p > 0.1

up more in discriminating the intensity of speech stimuli compared to their performance in discriminating speech rate and pitch, than those with worse speech recognition. Overall, participants with higher processing speed made accurate discrimination decisions faster. The maximal random slope structure indicated that the older adults differed in the degree to which their discrimination speed was affected by a change in speech dimension. Moreover, the maximal random slope structure included effects of processing speed, working memory and vocabulary knowledge on item, suggesting that the effects of these measures on the speed of discrimination decisions differed across speech stimuli.

The best fitting model predicting discrimination speed on the basis of stimulus length (words vs. sentences for the pitch and intensity stimuli) included stimulus length, processing speed, working memory and the interaction between stimulus length and these predictors (Table 2.6B). As can be expected, participants generally took longer to make the discrimination decision when they listened to sentences than to words. Further, the faster participants‘ processing speed, the more efficient their word discrimination was. Word discrimination performance was also marginally supported by working memory. However, these cognitive skills effects were modified by stimulus length: the contributions of processing speed and working memory observed in the discrimination of words were effectively cancelled in the discrimination of sentences. The maximal random slope structure indicated that listeners varied in the degree to which their discrimination speed was influenced by a change from words to sentences. Moreover, listeners' pure-tone thresholds showed a random slope on item, suggesting that discrimination of some speech stimuli was affected by hearing sensitivity[2].

2 We also reran the regression analyses without the two participants scoring below the 80% criterion of the auditory discrimination test. All associations between discrimination performance and both cognitive and linguistic measures remained consistent, indicating that the reported effects were not driven by single participants or single items.

Table 2.6A. Statistical model for the response times as a function of subtest. Standard error is indicated by *SE*.

	Subtest analysis				
Fixed effects	*β*	*SE*	*t*	*p*	
Intercept	1092.92	103.60	10.54	< .001	
Duration	0.38	0.06	6.58	< .001	
Pitch	-	-	-	*n.s.*	
Speech rate	236.84	95.65	2.48	.007	
Speech in quiet	-	-	-	*n.s.*	
Processing speed	-6.27	1.82	-3.44	< .001	
Pitch * speech in quiet	-	-	-	*n.s.*	
Speech rate * speech in quiet	-5.49	2.32	-2.37	.007	
Random effects		*Variance*	*SD*	*Corr*	
Subject	Intercept	3542.8	188.22		
	Pitch	1595.4	126.31	-.375	
	Speech rate	1767.2	132.93	-.330	.656
Item	Intercept	13813	371.66		
	Speed	8.16	2.86		
	Working memory	1.92	1.39		
	Vocabulary	16197	402.46	-.875	
Residual		8139.2	285.29		

Note. *n.s.* = *p* > 0.1

Table 2.6B. Statistical model for the response times as a function of stimulus length. Standard error is indicated by *SE*.

	Stimulus length analysis			
Fixed effects	*β*	*SE*	*t*	*p*
Intercept	1308.38	48.19	27.15	< .001
Sentences	695.35	60.47	11.50	< .001
Processing speed	-9.98	2.22	-4.50	< .001
Working memory	-2.16	1.28	-1.68	.092
Sentences * speed	7.53	1.74	4.33	< .001
Sentences * memory	2.65	1.00	2.64	0.009
Random effects		*Variance*	*SD*	*Corr*
Subject	Intercept	42208	205.45	
	Sentence	17767	133.30	-.582
Item	Intercept	37215	192.91	
	Hearing	6.87	2.62	-.640
Residual		83044	288.17	

Note. n.s. = p > 0.1

Discussion

This study was set up to assess the relationship between auditory discrimination performance on speech materials and measures of auditory, cognitive and linguistic functioning in healthy older people. Discrimination of intensity, pitch and speech rate were assessed in the context of meaningful words and sentences, as well-established methods in dysarthria rehabilitation, such as the Lee Silverman Voice Treatment (Ramig et al., 1996) and the Pitch Limiting Voice Treatment (de Swart et al., 2003), encourage patients to speak loud, low and/ or slow to increase their speech intelligibility. Ninety-four out of 96 healthy participants discriminated more than 80% of items correctly, thereby demonstrating that healthy older adults are generally able to perceive acoustic differences on relevant dimensions and on relevant materials for speech therapy. This shows that discrimination tasks can be applied in speech and language therapy in older adults.

Due to participants‘ high performance level on the auditory discrimination test, relatively little variation could be observed in discrimination accuracy. Despite this limitation, individual characteristics explained variability in discrimination performance, thus pinpointing abilities that differentiate between good and excellent performers. As these abilities are predictive of small performance differences here, they may be expected to also account for variation among more heterogeneous populations.

Overall, the current study suggests that perception of acoustic differences in speech is more closely related to cognitive processing and linguistic skills than to auditory abilities, as evidence for the association between auditory discrimination and both cognitive and linguistic skills converges from both the factor and the regression analysis. In the exploratory factor analysis, discrimination accuracy loaded on factors of processing efficiency and linguistic skills but was not associated with a third factor of auditory functioning. Even with mild to moderate hearing loss, older adults were able to reach high accuracy scores, suggesting that stimulus audibility was sufficient for successful discrimination. Importantly, discrimination speed was not affected by hearing abilities either. This implies that peripheral hearing sensitivity barely plays a role in discrimination tasks that are relevant for speech therapy, as long as participants‘ hearing is as good as in the present sample (note that approximately 95% of adults in this age range have a hearing status within the hearing range observed in the present study). Therefore, our results suggest that how suitable discrimination-based approaches in speech and language therapy will be for patients cannot simply be predicted from individuals‘ sensory functioning.

Apart from peripheral hearing loss, auditory aging is also characterized by difficulties in understanding speech in noisy conditions. Although we did not find an effect of hearing sensitivity, listeners‘ speech understanding in noise predicted their discrimination accuracy, indicating that central auditory processing, rather than peripheral auditory processing, relates to discrimination performance. However, as measures of speech understanding in noise and of auditory discrimination performance load on different factors, this suggests that auditory discrimination of speech indeed mainly taps cognitive and linguistic skills.

The association between auditory discrimination of speech materials and general cognitive processing is in line with previous literature on non-speech stimuli, implying that in behavioral tasks of auditory processing no clear distinction between auditory processing and cognitive functions can be made (e.g., Humes et al., 2012). Processing speed seems to be particularly important for discrimination performance, as it is associated with discrimination accuracy and discrimination speed. Further, participants with better working memory are faster in correctly discriminating words and tend to be more accurate in their overall performance. However, no effect of auditory short-term memory on discrimination performance was found. Thus, discrimination performance seems to be associated with complex processes of simultaneous storage and processing rather than auditory short-term memory which primarily reflects a storage component (Gathercole, 1999).

Note that processing speed and working memory only predicted discrimination speed in the discrimination of *words,* rather than sentences. This may be counterintuitive at first, as sentence processing may be expected to place more demands on processing speed and working memory. However, as participants were not instructed to respond as quickly as possible, discrimination of words may have been more sensitive to effects of processing speed than that of sentences because the latter are obviously longer in duration than words. Moreover, participants' task was to extract and compare representations of tested speech dimensions (i.e., intensity, overall pitch, and speech rate). Sentences provide a richer informational context than words to extract these characteristics. In word discrimination, participants have less time or structural information to derive a good internal representation of the speech dimension. Therefore, the use of longer speech stimuli such as simple sentences may be particularly recommended in the treatment of individuals with poorer working memory capacity or slower cognitive processing.

Furthermore, our results indicate that discrimination of speech stimuli not only engages general cognitive processes but also linguistic skills. Note that all discrimination tasks could have been performed in a foreign language, as the linguistic content of the speech signal was not relevant for the discrimination decision itself. Therefore, the finding that the perception of acoustic differences between speech stimuli is associated with individual linguistic performance suggests that listeners routinely process the available linguistic information. Proficient language users may perform the discrimination tasks with more ease as they need less cognitive resources for the task of linguistic processing even when processing high-frequency words or relatively common sentences. This cognitive trade-off effect may particularly emerge in older adults. Many participants in our study experienced some degree of hearing loss (be it slight), and were, therefore, exposed to a degraded speech signal. Previous research has shown that linguistic experience such as higher vocabulary knowledge facilitates the processing of degraded speech (Janse & Adank, 2012; McAuliffe et al., 2013; compare also Chapter 4) and that this effect is not influenced by word frequency (McAuliffe et al., 2013). In therapy settings, it may thus be beneficial for patients with poorer language proficiency to start with discrimination tasks with simple stimuli (e.g., evaluating the production of prolonged vowels or fricatives).

Our finding that perception of acoustic differences in speech seems to be more related to cognitive and linguistic skills than to auditory abilities may have some important clinical implications. Note that we observed these associations testing neurologically intact older adults in a non-distracting setting. The impact of these cognitive and linguistic skills on auditory discrimination ability may become even more prominent under less ideal listening conditions, such as performing auditory discrimination tasks in a clinical setting where therapy is unlikely to take place without interfering background noises, where patients are commonly asked to produce and compare utterances simultaneously, and where patients often have to analyze more than one speech dimension at the same time. Considering the relationship between speech understanding in noise and discrimination accuracy, a quiet non-distracting environment seems to be essential in order to apply auditory discrimination tasks in therapy effectively. Furthermore, the use of recording facilities may be recommended to record patients‘ speech during therapy. By doing so, we hypothesize that patients can first concentrate on producing the target utterances and, thereby, familiarize with the linguistic content. In a subsequent step, the speech and language therapist can ask patients to compare their own speech with the target speech. This approach is similar to the procedure implemented in E-learning based Speech Therapy (Beijer et al., 2010). Moreover, rehabilitation may be more effective if patients are asked to concentrate on just one defined speech dimension. As fewer errors were made in discriminating intensity than in discriminating both pitch and speech rate, this suggests that listeners are most familiar with the loudness dimension, which can therefore best be used as the dimension to familiarize patients with discrimination tasks.

The auditory discrimination test described in the current study was developed to assess whether individuals with dysarthria qualify for E-learning based Speech Therapy. It is crucial to take patient abilities into account when designing and implementing eHealth services, as patient abilities form a key human factor in utilization and acceptance of telemedicine programs (Brennan & Barker, 2008). By assessing which individual abilities are associated with performance on the auditory discrimination test in healthy older adults, the current study provides first starting points for the importance of patients‘ perceptual, cognitive and linguistic skills for E-learning based Speech Therapy. At the same time, the question arises whether the observed pattern of associations generalizes to clinical populations and if so, whether patients with better cognitive and linguistic performance and, hence better auditory discrimination skills, may benefit more from discrimination-based approaches of speech and language therapy than patients with poorer cognitive and linguistic performance. These aspects were not explored in the current study but may be considered for future clinical research.

References

American Speech-Language-Hearing Association. (1996). Central auditory processing: current status of research and implications for clinical practice. doi: 10.1044/policy.TR1996-00241

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. *Language Learning, 62*, 49-78. doi: 10.1111/j.1467-9922.2012.00706.x

Baayen, H., Piepenbrok, R., & Gulikers, L. (1995). The CELEX database consortium (CD-ROM). *Linguistic data.* Nijmegen, The Netherlands: Max Planck Institute for Psycholinguistics.

Baayen, R. H. (2012). *Analyzing Linguistic Data: A Practical Introduction to Statistics using R* (6th ed.). Cambridge: Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390-412. doi: 10.1016/j.jml.2007.12.005

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: keep it maximal. *Journal of Memory and Language, 68*, 255-278. doi: 10.1016/j.jml.2012.11.001

Bates, D., Maechler, M., & Bolker, B. (2012). lme4: Linear mixed-effects models using S4 classes (Version 0.999-0). Retrieved from http://CRAN.R-project.org/package=lme4

Beijer, L. J., Rietveld, A. C. M., van Beers, M. M. A., Slangen, R. M. L., van den Heuvel, H., de Swart, B. J. M., & Geurts, A. C. H. (2010). E-learning-based speech therapy: a web application for speech training. *Telemedicine Journal and E-Health, 16*, 177-180. doi: 10.1089/tmj.2009.0104

Beijer, L. J., Rietveld, A. C. M., & van Stiphout, A. J. L. (2011). Auditory discrimination as a condition for E-learning based Speech Therapy: a proposal for an auditory discrimination test (ADT) for adult dysarthric speakers. *Journal of Communication Disorders, 44*(6), 701-718. doi: 10.1016/j.jcomdis.2011.05.002

Bosman, A. (1989). *Speech perception by the hearing impaired.* Doctoral dissertation, Utrecht University, Utrecht, The Netherlands.

Brennan, D. M., & Barker, L. M. (2008). Human factors in the development and implementation of telerehabilitation systems. *Journal of Telemedicine and Telecare, 14*(2), 55-58. doi: 10.1258/jtt.2007.007040

Cliff, N., & Hamburger, C. D. (1967). Study of sampling errors in factor analysis by means of artificial experiments. *Psychological Bulletin, 68*(6), 430-445. doi: 10.1037/h0025178

Committee on Hearing, Bioacoustics and Biomechanics [CHABA]. (1988). Speech understanding and aging. *Journal of the Acoustical Society of America, 83*, 859-895. doi: 10.1121/1.395965

Cook, R. D., & Weisberg, S. (1982). *Residuals and influence in regression*. New York: Chapman & Hall.

Cox, L. C., McCoy, S. L., Tun, P. A., & Wingfield, A. (2008). Monotic auditory processing disorder tests in the older adult population. *Journal of the American Academy of Audiology, 19*(4), 293-308. doi: 10.3766/Jaaa.19.4.3

Cruickshanks, K. J., Wiley, T. L., Tweed, T. S., Klein, B. E., Klein, R., Mares-Perlman, J. A., & Nondahl, D. M. (1998). Prevalence of hearing loss in older adults in Beaver Dam, Wisconsin: The epidemiology of hearing loss study. *American Journal of Epidemiology, 148*, 879-886. doi: 10.1093/oxfordjournals.aje.a009713

Cunnings, I. (2012). An overview of mixed-effects statistical models for second language researchers. *Second Language Research, 28*, 369-382. doi: 10.1177/0267658312443651

de Swart, B. J. M., Willemse, S. C., Maassen, B. A. M., & Horstink, M. W. I. M. (2003). Improvement of voicing in patients with Parkinson's disease by speech therapy. *Neurology, 60*(3), 498-500. doi: 10.1212/01.WNL.0000044480.95458.56

Fogerty, D., Humes, L. E., & Kewley-Port, D. (2010). Auditory temporal-order processing of vowel sequences by young and elderly listeners. *The Journal of the Acoustical Society of America, 127*, 2509-2520. doi: 10.1121/1.3316291

Gathercole, S. E. (1999). Cognitive approaches to the development of short-term memory. *Trends in Cognitive Sciences, 3*(11), 410-419. doi: 10.1016/S1364-6613(99)01388-1

Gopinath, B., Schneider, J., Hartley, D., Teber, E., McMahon, C. M., Leeder, S. R., & Mitchell, P. (2011). Incidence and predictors of hearing aid use and ownership among older adults with hearing loss. *Annals of Epidemiology, 21*(7), 497-506. doi: 10.1016/j.annepidem.2011.03.005

Hanania, E., & Shikhani, M. (1986). Interrelationships among three tests of language proficiency: standardized Esl, cloze, and writing. *Tesol Quarterly, 20*(1), 97-109. doi: 10.2307/3586391

Harris, K. C., Eckert, M. A., Ahlstrom, J. B., & Dubno, J. R. (2010). Age-related differences in gap detection: effects of task difficulty and cognitive ability. *Hearing Research, 264*, 21-29. doi: 10.1016/j.heares.2009.09.017

Humes, L. E. (2005). Do ‚auditory processing' tests measure auditory processing in the elderly? *Ear and Hearing, 26*, 109-119. doi: 10.1097/00003446-200504000-00001

Humes, L. E., Dubno, J. R., Gordon-Salant, S., Lister, J. J., Cacace, A. T., Cruickshanks, K. J., . . . Wingfield, A. (2012). Central presbycusis: a review and evaluation of the evidence. *Journal of the American Academy of Audiology, 23*, 635-666. doi: 10.3766/jaaa.23.8.5

Idrizbegovic, E., Hederstierna, C., Dahlquist, M., Kampfe Nordstrom, C., Jelic, V., & Rosenhall, U. (2011). Central auditory function in early Alzheimer‘s disease and in mild cognitive impairment. *Age and Ageing, 40*(2), 249-254. doi: 10.1093/ageing/afq168

Janse, E., & Adank, P. (2012). Predicting foreign-accent adaptation in older adults. *Quarterly Journal of Experimental Psychology, 65*, 1563-1585. doi: 10.1080/17470218.2012.658822

Kidd, G. R., Watson, C. S., & Gygi, B. (2007). Individual differences in auditory abilities. *The Journal of the Acoustical Society of America, 122*, 418-435. doi: 10.1121/1.2743154

McAuliffe, M. J., Gibson, E. M., Kerr, S. E., Anderson, T., & LaShell, P. J. (2013). Vocabulary influences older and younger listeners‘ processing of dysarthric speech. *The Journal of the Acoustical Society of America, 134*(2), 1358-1368. doi: 10.1121/1.4812764

Muthén, B. O., & Kaplan, D. (1985). A comparison of some methodologies for the factor analysis of non-normal Likert variables. *British Journal of Mathematical and Statistical Psychology, 38*(2), 171-189. doi: 10.1111/j.2044-8317.1985.tb00832.x

Neger, T. M., Rietveld, A. C. M., & Janse, E. (2014). Relationship between perceptual learning in speech and statistical learning in younger and older adults. *Frontiers in Human Neuroscience, 8*, 628. doi: 10.3389/fnhum.2014.00628

Noble, W., & Gatehouse, S. (2004). Interaural asymmetry of hearing loss, speech, spatial and qualities of hearing scale (SSQ) disabilities, and handicap. *International Journal of Audiology, 43*(2), 100-114. doi: 10.1080/14992020400050015

Papakonstantinou, A., Strelcyk, O., & Dau, T. (2011). Relations between perceptual measures of temporal processing, auditory-evoked brainstem responses and speech intelligibility in noise. *Hearing Research, 280*, 30-37. doi: 10.1016/j.heares.2011.02.005

Park, D. C., & Reuter-Lorenz, P. (2009). The adaptive brain: aging and neurocognitive scaffolding. *Annual Review of Psychology, 60*, 173-196. doi: 10.1146/annurev.psych.59.103006.093656

Pichora-Fuller, M. K., & Souza, P. E. (2003). Effects of aging on auditory processing of speech. *International Journal of Audiology, 42*, S11-S16. doi: 10.3109/14992020309074638

Plomp, R., & Mimpen, A. M. (1979). Improving the reliability of testing the speech reception threshold for sentences. *Audiology, 18*, 43-52. doi: 10.3109/00206097909072618

Ramig, L. O., Countryman, S., O‘Brien, C., Hoehn, M., & Thompson, L. (1996). Intensive speech treatment for patients with Parkinson‘s disease: short-and long-term comparison of two techniques. *Neurology, 47*(6), 1496-1504. doi: 10.1212/WNL.47.6.1496

Schneider, B., & Pichora-Fuller, M. (2001). Age-related changes in temporal processing: implications for speech perception. *Seminars in Hearing, 22*(3), 227-240. doi: 10.1055/s-2001-15628

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-prime reference guide*. Pittsburgh, PA: Psychology Software Tools.

Schroter-Morasch, H., & Ziegler, W. (2005). Rehabilitation of impaired speech function (dysarthria, dysglossia). *GMS Current Topics in Otorhinolaryngology, Head and Neck Surgery, 4*, 15. doi: 10.1002/cncr.10668

Wechsler, D. (2004). *Wechsler Adult Intelligence Scale* (3rd ed.). Amsterdam: Harcourt Test.

World Health Organization. (2014). Grades of hearing impairment Retrieved 08.03.2014, from http://www.who.int/pbd/deafness/hearing_impairment_grades/en/

Wurm, L. H., & Fisicaro, S. A. (2014). What residualizing predictors in regression analyses does (and what it does not do). *Journal of Memory and Language, 72*, 37-48. doi: 10.1016/j.jml.2013.12.003

3

Effects of Modality, Stimulus Type and Attention on Statistical Learning

This chapter is based on:

Neger, T. M., Rietveld, T., Janse, E. (submitted). Effects of modality, stimulus type and attention on statistical learning

Abstract

Human sensitivity to underlying regularities is closely related to language acquisition and language processing. This ability of statistical learning, therefore, provides a unique test bed to investigate several questions on human information processing. The present study investigated the role of modality, stimulus type and attention on statistical learning. We created four variants of an artificial grammar learning - serial reaction time task which differed in terms of whether input was auditory or not, and linguistic or not, and in how much attention participants needed to pay to the input in order to perform a cover task. Our results indicate that modality and stimulus type influence statistical learning performance. Participants showed more statistical learning with non-linguistic compared to linguistic input, and when presented with auditory compared to visual information. These findings support recent theories that statistical learning is influenced by modality- and stimulus-specific processing constraints.

Introduction

The concept of statistical learning has gained increasing attention over the last couple of years. Picking up on the frequencies with which sensory events co-occur has been considered one of the core mechanisms of the brain to discover regularities in the environment and, consequently, to form predictions about it (Siegelman & Frost, 2015). Given that language consists of complex patterns of sequentially presented units such as phonemes, syllables and words (Conway & Christiansen, 2006), statistical learning has been shown to play a key role in speech and language processing. Newborns of only one or two days old have been found to be sensitive to the statistical properties of incoming speech (Teinonen, Fellman, Naatanen, Alku, & Huotilainen, 2009). Only a couple of months later, babies make use of the transitional probabilities between syllables to segment speech into words (Saffran, Aslin, & Newport, 1996). At a later age, children who perform better in statistical learning tasks have been found to show more progress in the acquisition of syntax (Kidd, 2012). Therefore, statistical learning has been proposed to be one of the central mechanisms in language acquisition, enabling young children to rapidly and successfully acquire their native language (Newport & Aslin, 2004).

Recent studies show that sensitivity to structural regularities in auditory input is not only important for language development. Adults who are more sensitive to input statistics show better language comprehension (Misyak & Christiansen, 2012), higher reading ability (Arciuli & Simpson, 2012), more progress in learning to read in a second language (Frost, Siegelman, Narkiss, & Afek, 2013), and adapt more quickly to unfamiliar speech input (see Chapter 4), suggesting that statistical learning is involved in online speech and language processing. As statistical learning is so closely related to language acquisition and language processing, it provides a unique test bed to investigate several questions on how information is processed. The present study focuses on the following three questions: Firstly, are co-occurrence frequencies easier to pick up on in the visual or auditory modality? Secondly, are participants more sensitive to linguistic information than to other types of information? Thirdly, is explicit attention required for the detection of regularities?

Statistical learning has been observed in multiple modalities such as the visual, auditory as well as the tactile domain (Conway & Christiansen, 2005). These observations have been taken to stress the generality of the learning mechanism. Accordingly, the view of a single amodal statistical learning system has been implemented in many influential models on statistical learning (e.g., Altmann, Dienes, & Goode, 1995; Reber, 1989). However, recent studies challenge the notion of a unitary learning system. Auditory statistical learning has been reported to outperform statistical learning of temporal regularities in the visual and tactile domains (Conway & Christiansen, 2005; Robinson & Sloutsky, 2013), which suggests that there is a modality-specific advantage for auditory statistical learning. Additionally, visual and auditory statistical learning have been shown to work in parallel but with grammatical regularities being learned only for the modality they are presented in (Conway & Christiansen, 2006). That is, if participants were simultaneously exposed to two artificial grammars, one being presented in tone sequences and one being presented in shape sequences, participants learned both grammars but grammatical regularities embedded in tone sequences did not generalize to shape sequences, and vice versa. This indicates that participants did not derive *abstract* patterns like AABACC, which could be applied to both tones and shapes, but that participants rather learned specific patterns for the stimulus set at hand. Furthermore, auditory and visual statistical learning have been reported to be differentially affected by changes in presentation speed (Emberson, Conway, & Christiansen, 2011). Taken together, these results indicate that statistical learning is influenced by modality-specific processes. Importantly, studies exploring modality effects in statistical learning often have to make use of different stimuli for each modality to investigate learning (e.g., vibrations vs. tones vs. shapes). With the current study, we aim to broaden our understanding of modality effects in statistical learning by addressing the question whether modality-specific differences still occur if the input is the same but only the modality in which target items are presented varies.

Statistical learning has not only been reported across different modalities but also across a wide range of stimuli such as shapes, colours, pictures, vibrations, tones, nonwords and words (e.g., Conway & Christiansen, 2005, 2006; Otsuka, Nishiyama, Nakahara, & Kawaguchi, 2013). Even though participants may show more statistical learning on the basis of some stimuli than others due to modality effects, it is rather unclear whether statistical learning within one modality is stimulus-specific. In the study by Conway and Christiansen (2006), regularities from one stimulus set did not generalize to a second set learned in parallel within the same modality if grammatical sets consisted of different stimuli (e.g., colour vs. shape sequences). This suggests that learning within the same modality is also specific for the stimulus set at hand. Additionally participants' performance on one statistical learning task does not necessarily correlate with their performance on another statistical learning task using different stimuli in the same modality (Siegelman & Frost, 2015). These findings suggest that statistical learning may not only be modality- but also stimulus-specific.

Two recent models account for modality-specific effects in statistical learning but only one explicitly predicts stimulus-specificity of learning. The mechanistic model of statistical learning (Frost, Armstrong, Siegelman, & Christiansen, 2015) assumes that similar neuronal and computational principles exist for statistical learning but that these principles are instantiated across separate neuronal networks rather than being implemented in a unitary learning system. Importantly, the learning principles are influenced by the specific constraints of the input which result from encoding and processing the input in different cortical areas (e.g., separate neuronal networks exist for processing auditory and visual information). Thus, the mechanistic model of statistical learning predicts stimulus-specific learning if stimuli are processed via different neuronal networks. The second model, the embodied learning mechanism (Emberson et al., 2011), hypothesizes that statistical learning is an integral part of the respective processing modality in the sense that perception itself is a prediction-based process. As the processing systems operate differently across modalities, this accounts for modality-specific effects in statistical learning. The latter may also imply that learning is different for different stimuli. In sum, even though the mechanistic model by Frost and colleagues (2015) is more explicit about it than the embodied learning one by Emberson and colleagues (2011), both models may account for stimulus-specificity in statistical learning. Observing learnability differences between different stimuli within the same modality may, therefore, not differentiate between these models but provide further evidence for them rather than for a unitary learning system.

Given the importance of statistical learning for language processing, we were specifically interested in the question whether humans may be more sensitive to *linguistic* stimuli than to other types of information. One of the few studies reporting statistical learning results for two different types of stimuli within the same modality implies no difference between auditory learning of syllables and learning of non-linguistic sounds (i.e., Macintosh alert signals) (Siegelman & Frost, 2015). In visual statistical learning, however, learning of non-adjacent dependencies has been shown to occur in linguistic stimuli (i.e., nonwords) but not in matched non-linguistic stimuli (i.e., black and white pattern matrices), indicating a learning advantage for linguistic input (Sturm & Smith, 2009). In contrast, auditory learning of a grammar consisting of harmonic chords has been observed to outperform learning of an auditory letter grammar (Bly, Carrion, & Rasch, 2009), which suggests that there is a learning advantage for non-linguistic stimuli. Although it is difficult in general to assess similarity of items across modalities, note that in the latter experiment the spoken names of the letters were fairly similar (consonants such as S, M, L, X), whereas the chords were fairly dissimilar (e.g., C Major, D minor, F major). As serial recall of auditory sequences is detrimentally affected by item similarity (e.g., Baddeley, 1966; Conrad, 1964; Williamson, Baddeley, & Hitch, 2010), the greater dissimilarity of the musical chords may have led to a processing advantage of the non-linguistic items in auditory short-term memory. In sum, quite some statistical learning research has focused on modality differences but there is only sparse evidence as to whether there are also stimulus-specific processing advantages in statistical learning. By directly comparing statistical learning of visual shapes to learning of visually presented nonwords, we aim to address this question specifically for the role of linguistic information. As language processing is such a vital component of the way humans interact with their environment, we may expect humans to be especially sensitive to underlying regularities in linguistic material.

Another aspect of statistical learning which is not yet fully understood is the role of attention. Statistical learning has been shown to occur when participants are passively exposed to auditory input (e.g., Toro, Sinnett, & Soto-Faraco, 2005) or visual input (e.g., Turk-Browne, Junge, & Scholl, 2005), which suggests that statistical learning is a rather automatic process. However, participants may attend to the input stream even if they have no explicit task instructions. One early finding taken as evidence that statistical learning does not require attention showed that both children and adults were able to learn statistical regularities in an auditory syllable stream that was played in the background when participants were actively focusing on a free drawing task (Saffran, Newport, Aslin, Tunick, & Barrueco, 1997). As participants were free to divide their attention between drawing and listening, the drawing task may not have been so attention-demanding as to impede attending the auditory input. In contrast, tasks which require participants to pay selective attention to competing input streams have been shown to impair learning of unattended information. For example, participants only showed statistical learning of picture sequences in a target colour that they were instructed to attend to but not in a picture stream of a non-attended colour (Turk-Browne et al., 2005). Similarly, learning of statistical properties in an auditory speech syllable stream was impaired if participants were simultaneously engaged in one of the following three tasks: if they had to detect stimulus repetitions in a competing input stream within the same modality (e.g., noises), if they had to detect stimulus repetition in a competing stream in a different modality (e.g., pictures), or if they had to detect changes in a competing dimension within the target stream (e.g., pitch) (Toro et al., 2005). These findings suggest that statistical learning requires at least some attentional resources. Note that in these experiments participants were explicitly asked to divert their attention from the target dimension and, thus, to guide their attention away from the input which was relevant for statistical learning. In the real world, however, as shown in the experiment by Saffran and colleagues (1997), humans are probably attending specific sensory input without actively ignoring competing, irrelevant information. Therefore, a question that remains highly relevant for statistical learning research is whether statistical learning may still occur in situations in which attention may not be focused on the stimuli themselves (Turk-Browne et al., 2005).

Further evidence that humans may be able to learn structural regularities even if they do not consciously attend the input comes from a recent study in which participants showed learning of symbol sequences under binocular rivalry (Kido & Makioka, 2015). That is, participants did not consciously attend the visual symbols, which were presented to one of the eyes only, as visual perception was suppressed by simultaneously presenting flash patterns to the other eye. To further investigate the role of attention in statistical learning, we aim to compare statistical learning under two conditions. In the one condition, participants have to process the input relevant for learning in order to be able to perform a cover task. In the other condition, participants are exposed to the same target stream but attentive processing of the stimuli is not required for performing the cover task. Based on the results of Saffran and colleagues (1997) and of Kido and Makioka (2015), we expect learning to take place in both conditions, but we also expect more learning in the condition in which the information is crucial for the cover task and, hence, participants have to attentively process the relevant information.

By adjusting a statistical learning paradigm that combines aspects of artificial grammar learning and serial reaction time paradigms (Misyak, Christiansen, & Tomblin, 2010), we aim to address the questions discussed above regarding the roles of modality, type of information (linguistic vs. nonlinguistic) as well as attention in statistical learning. As the paradigm makes use of response time data to track statistical learning, it enables us to measure statistical learning implicitly, contrary to most of the reported studies. Most importantly, the paradigm allows us to expose participants to exactly the same statistical learning tasks and grammars, the only difference being either the modality in which target items are cued, the stimulus type that is presented or the attention participants need to pay to the input stream in order to be able to perform a cover task.

Method

Participants

We administered four variants of a statistical learning task. Three variants were completed by sixty participants, one variant was completed by thirty participants (we will come back to these numbers below in the section on the statistical learning task). As some participants were tested on more than one variant, 132 university students in total participated in the current study. Fifty-nine participants completed exactly one of the four variants of the statistical learning task. Most participants (n = 73) completed more variants, as two variants were administered within the same experimental session. Five participants even performed three task variants. In these rare cases, the third task variant was administered more than a year after the first variant to prevent any carry-over effects between tasks.

All participants were native speakers of Dutch. Participants were aged between 18 and 27 years (M = 21.4 years, SD = 2.1 years) with the majority of participants being female (90 out of 132, 68.2 %). Participants had normal or corrected to normal vision and all had to pass a visual screening test with an acuity of at least 20/65. That is, participants had to be able to correctly identify letters of 2.5 mm in height at a distance of 60 cm on a Snellen chart. As one variant involved the presentation of auditory stimuli, hearing acuity was assessed in the participants on that task by means of pure-tone audiometry prior to testing (air-conduction thresholds only). All participants had normal hearing with a mean pure-tone threshold at 0.5, 1 and 2 kHz not exceeding 20 dB HL (mean pure-tone threshold = 6.11 dB HL, *SD* = 5.51) and were, therefore, allowed to participate in the study. Participants were recruited via the subject database of the Max Planck Institute for Psycholinguistics and were compensated €8 per hour for their time.

Statistical Learning Task

We adopted the artificial grammar learning - serial reaction time paradigm [AGL-SRT], which was developed to resemble processes of regularity detection in natural language learning (Misyak et al., 2010). To that end, the paradigm makes use of linguistic, auditory stimuli which are sequentially presented based on underlying regularities. The task has the advantage of eliciting statistical learning in a short amount of time. Moreover, statistical learning is assessed by means of changes in reaction time data. Therefore, the task allows insights into the online processing of statistical regularities and to measure statistical learning implicitly. That is, participants pay attention to a cover task in which they need to click on indicated items and are, in contrast to traditional artificial grammar learning designs, not explicitly instructed to remember the sequences.

To illustrate the rationale behind the statistical learning paradigm, we first describe the paradigm on the basis of one of the four task variants. The three other task variants are then specified at the end of this section (see also Figure 3.2). The task was designed as a two-alternative forced-choice serial reaction time paradigm. Participants saw two rows of nonwords on a computer screen. In the original task, each row consisted of three nonwords to explore learning of adjacent as well as non-adjacent dependencies. As we were focusing on statistical learning in general, the task in our study consisted of two rows of two nonwords only (see Figure 3.1). In each column, one nonword served as target and one as distractor item. Participants were instructed to click as fast as possible on the target items which were indicated by their auditory presentation one after the other. That is,

participants had to click on two target nonwords within each trial. The second target was only presented once the participant had correctly clicked on the first target. The first target item was always located on the left hand side of the screen (i.e., in the upper or lower row of the first column) and the
second target was always located on the right hand side of the screen (i.e., in the upper or lower row of the second column). Crucially, which of the two nonwords was going to be cued on the right hand side of the screen was predictable on the basis of the identity of the first target.

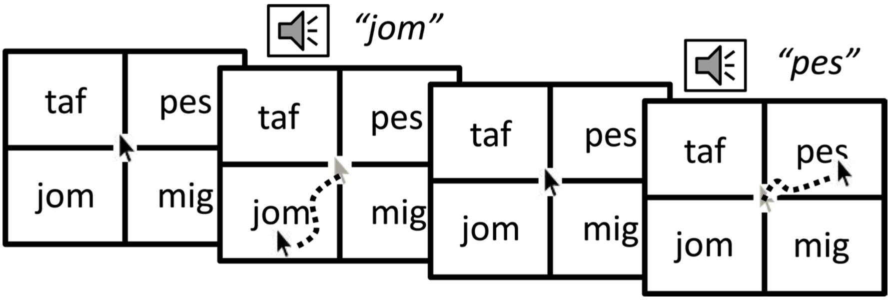

Figure 3.1. Overview of a trial of the auditory nonword matching task.

The grammar consisted of eight different nonwords which were separated into two grammatical sets of four nonwords. Within each set, two nonwords functioned as “leaders”, meaning that these nonwords always served as first (or ‚left hand‘) targets within a trial. The remaining two nonwords of a set were “followers”. These nonwords could only appear on the right hand side of the screen and always followed a leader item from the same set (and hence, served as second targets only). Accordingly, each set of two leaders and two followers could be presented as one of four legal combinations (c.f., Figure 3.2). This resulted in a total of eight legal combinations for the statistical learning task. As each leader could be followed by one of two different nonwords, the transitional probability from a leader to a specific follower was 0.5 within the grammar. Within each trial, however, the predictability of the second target was 1, as distractor items consisted of a legal combination from the competing grammatical set. Importantly, vertical target positions were randomly assigned so that it was impossible for participants to predict whether a specific nonword would appear on the upper or lower half of the screen.

The statistical learning task consisted of three phases. In the exposure phase, participants were repeatedly exposed to the legal target combinations. Overall, the exposure phase contained sixteen repetitions of each grammatical combination (each block presenting all legal combinations once), resulting in a total of 128 trials (16 blocks of 8 trials). By picking up on the statistical properties of the input, participants should become faster in clicking on the second target compared to clicking on the first target over the course of this exposure phase. However, faster response times can reflect both

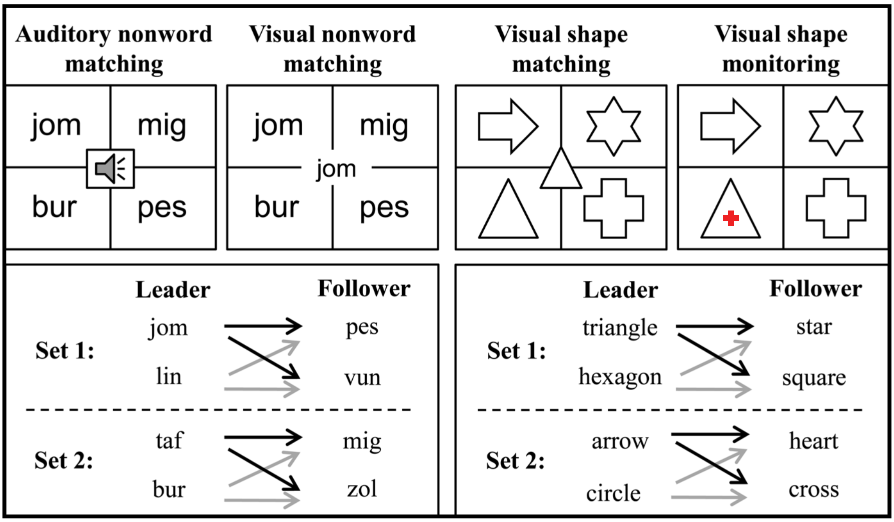

Figure 3.2. Overview of the four variants of the statistical learning task and their respective grammar.

statistical and task learning. In the test phase, the underlying regularities were therefore reversed to be able to disentangle statistical and task learning. Thus, the leaders of one set were now suddenly succeeded by the followers of the other set so that the second target was no longer predictable for the participants. If participants had detected the underlying regularities, they should show a drop in performance during the test phase. That is, participants would need to correct their predictions and, hence, would slow down in their response to the second target. Importantly, this slowdown can only be attributed to participants‘ implicit grammar sensitivity. To avoid that participants would start to adapt to the new regularities, we implemented a short test phase of two blocks of eight trials only. Statistical learning was then operationalized as the change in performance from the end of the exposure phase (i.e., blocks 13-16) to the test phase (blocks 17-18). Such a measure of learning is widely accepted in the literature on implicit learning (Janacsek & Nemeth, 2013). In the recovery phase, which contained two blocks of eight trials, the original regularities were reintroduced. This phase served as a control phase. By re-introducing the underlying regularities, participants‘ performance should not drop any further. In total, the statistical learning task consisted of 160 trials (20 blocks of 8 trials).

Importantly, performance on each trial was not quantified by participants’ absolute response times. Instead, performance was quantified as the facilitation participants experienced by being exposed to predictable stimuli. To that end, we calculated “facilitation scores” for each trial by dividing participants‘ response time to the first, unpredictable target by their response time to the second, predictable target. Using participants‘ response times to the first target within each trial as a baseline measure allowed us to minimize effects of individual differences in click response times throughout the statistical learning paradigm, and to account for general changes in click behaviour over the course of the experiment. If participants made no predictions about upcoming targets, they should be equally fast in clicking on both targets within a trial, which consequentially results in a facilitation score of 1. If participants predicted the identity of the second target, they should be faster in their response to the second compared to the first target, resulting in a facilitation score greater than 1. Thus, the facilitation score served as an index of participants' prediction of the second target within a trial.

To investigate the influence of input modality, linguistic vs. nonlinguistic information and attention on statistical learning, we implemented four variants of the statistical learning task. An overview of the different task variants is displayed in Figure 3.2. As described above, the first variant made use of the original stimuli and design with auditory input. That is, participants saw a display of nonwords on the screen and had to click on the target nonwords that they heard (referred to here as "auditory nonword matching"). To investigate the role of modality, we constructed the second variant such that the same nonwords were used but targets were now indicated by presenting them visually in the middle of the screen (referred to here as "visual nonword matching"). To assess whether linguistic information is specifically important for statistical learning, we designed the third variant such that the input stimuli were nonlinguistic, that is, geometrical shapes instead of nonwords. As in the visual nonword matching task, the targets were cued visually by smaller representation of the target shapes in the middle of the screen (this variant is referred to as "visual shape matching"). The fourth variant was implemented to specifically investigate the role of attention. In this "nonlinguistic" variant, the target shapes were indicated by a small visual marker in the form of a red cross inside the target shape. Importantly, in this variant, participants did not have to focus on the target shapes to solve the task (which is referred to as "visual shape monitoring") as participants only needed to look out for the red cross in order to know which quadrant to click. Consequently, participants may have been less attentive to the target shapes in this condition.

Thirty participants were tested for the auditory nonword matching task. All visual task variants were performed by sixty participants. As noted in the introduction, auditory statistical learning has been shown to outperform visual statistical learning. Given that smaller effects are more difficult to detect than bigger effects, we expected to need larger samples in the visual task variants to account for the difference in test sensitivity between the visual and auditory statistical learning variants.

Materials

For the "nonlinguistic" visual shape tasks, participants were presented with eight familiar, geometrical shapes drawn with a single, continuous black line (i.e., Set 1 leaders: *triangle*, *hexagon*; Set 1 followers: *star, square*; Set 2 leaders: *arrow, circle*; Set 2 followers: *heart, cross*). For the "linguistic" nonword tasks, participants were presented with eight monosyllabic CVC-nonwords (i.e., Set 1 leaders: *jom, lin*; Set 1 followers: *pes, vun*; Set 2 leaders: *taf, bur*; Set 2 followers: *mig, zol*) which had previously been implemented in a Dutch version of the AGL-SRT paradigm (Vuong, Meyer, & Christiansen, 2011). Nonwords were recorded by a 65 year old male speaker of Dutch. The auditory stimuli had an average duration of 442 ms ($SD = 60$).

Procedure

The statistical learning task was implemented in E-prime (Schneider, Eschman, & Zuccolotto, 2002) and started with five practice trials which consisted of legal combinations. At the start of each trial, participants saw a preview of the visual display which consisted of the four trial stimuli (i.e., either shapes or nonwords) and two grid lines which split the screen into two equally sized rows and columns. Shapes were displayed centred in these screen quadrants in a size of 75 x 70 mm. Nonwords were printed bold in Courier New with a font size of 24. The mouse cursor was located in the middle of the screen. After 500 ms, the first target was cued. Depending on the task, this cueing was done by the auditory presentation of the nonword recording (auditory nonword matching), the visual appearance of a smaller representation of the target nonword (printed bold in Courier New with a font size of 20) in the middle of the screen (visual nonword matching), the visual appearance of a smaller representation of the target shape (37.5 x 35 mm) in the middle of the screen (visual shape matching), or the visual appearance of small red cross (10 x 10

3

mm) in the middle of the target shape (visual shape monitoring). After participants had clicked on the first target item, the mouse cursor was automatically set back into the middle of the screen to ensure that the distance between the mouse and the targets was equal across all clicks. The second cue was presented 500 ms after the click to the first target. This small interval was implemented to allow for prediction effects to occur. Note that participants only proceeded to the next trial if they had clicked on the correctly cued target. Clicking on a wrong item or clicking outside the target area thus led to slower total response times for that click and, consequently, affected the facilitation score for that trial. A new trial started 500 ms after the click to the second target and automatic setback of the mouse cursor to its central position. After each block, a short break of 2500 ms was inserted to avoid fatigue effects. During this break, participants saw the block number as well as a reminder to click the target items as fast as possible on the screen. Participants took approximately 20 minutes to complete the task.

The visual nonword matching task and the visual shape matching task were administered within the same session due to time constraints. Though the tasks were very similar, we did not expect participants to transfer learning from one task to the other. No task order effects were found on learning in a within subject design in a visual statistical learning experiment comparing learning of orthographically presented words to learning of the same words displayed as line drawings (Otsuka et al., 2013). However, to account for possible order effects on learning, we counterbalanced the order in which participants performed the two tasks.

Legality Judgment Task

To assess participants' awareness of the underlying regularities and to derive a measure of their explicit knowledge of the target combinations, we additionally implemented a legality judgment task at the end of the experimental session. This also allowed us to compare our results to more traditional AGL designs and results. In this task, participants were presented with all legal and illegal leader-follower combinations and were asked to indicate whether they thought that they had observed the combination frequently or not. That is, participants were explicitly asked to recall which target combinations they had been presented with. As participants did not have to transfer their knowledge of the underlying regularities to new stimuli, as is the case in traditional AGL designs, performance above chance level cannot be taken as evidence for generalization but rather indicates participants‘ explicit knowledge of repeated fixed structures.

The legality judgment task was implemented in E-prime (Schneider et al., 2002) and consisted of sixteen trials. Within each trial, a follower-leader combination was printed in Courier New boldface with a font size of 24 (nonwords) on the screen. Shapes were displayed in a size of 75 x 70 mm. Follower and leader items were separated by a hyphen and presented in an order that was randomized for each participant. If participants recognised a combination from the statistical learning task, they were instructed to press the „1“ key on the keyboard. For unfamiliar combinations, participants pressed „2“. The assignment of the keys was continuously present on the screen. There was no time pressure to perform the task and it took participants approximately 5 minutes to complete the task. Participants who did both the visual nonword matching and the visual shape matching task completed the legality judgment task after the completion of both statistical learning tasks. In that case, participants did the two legality judgment tasks in the same order in which they had received the statistical learning tasks. The measure of interest was participants‘ accuracy to correctly identify legal and illegal target combinations (out of 16).

Data Analysis

Statistical Learning

We implemented linear mixed-effect models using the lmer function of the lme4 package (Bates, Maechler, & Bolker, 2012) in R (version 2.15.1) to analyse the contribution of different factors to statistical learning. Fixed variables of interest were (1) attention, indicating whether participants really had to focus on the targets or not (monitoring vs. matching tasks), (2) stimulus type, indicating whether participants had to process linguistic or non-linguistic stimuli (shape vs. nonword tasks), and (3) modality, indicating whether participants were exposed to auditory information (visual vs. auditory tasks). As statistical learning was defined as a drop in facilitation score from grammatical trials at the end of the exposure phase to ungrammatical trials in the test phase, the fixed categorical variable of phase (i.e., exposure vs. test) was the main predictor of interest, as well as the two-way interactions between phase and the respective variables attention, stimulus type and modality.

Additionally, two fixed control variables and their interaction were included in the analysis. Control variables were categorical measures of target position (i.e., did the first target appear upper or lower left?) and target alignment (i.e., were the two targets of a trial horizontally or diagonally aligned?). Regarding target position, we hypothesized that participants may tend to predict that the first target is located upper left instead of lower right due to the Western writing system. If that is the case, facilitation scores may be lower in trials in which the first target is located upper left (because of facilitation of the first click). Regarding alignment of the two targets, we expected participants to be faster in their response to the second target if both targets were aligned diagonally. Though it was pointed out to participants that the mouse would be automatically set back after each click, participants tended to move the mouse cursor back into the middle of the screen themselves. This compensatory movement may favour the diagonally aligned item and, hence, result in higher facilitation scores on diagonal trials.

All variables included in the analysis were categorical variables with two levels. Prior to the analysis, all variables were dummy coded such that higher values in the model indicated ungrammatical (test) trials relative to grammatical (exposure) trials (i.e., the phase variable), explicit attention to target relative to no explicit attention to target (i.e., the attention variable), nonwords relative to shapes (i.e., stimulus type variable), auditory information relative to visual information (i.e., the modality variable), upper left target position relative to lower left target (i.e., the position variable) and diagonal alignment relative

to horizontal alignment (i.e., the alignment variable). Additionally, all variables were centralized to allow for the model to calculate the average effects across all task variants. That is, the binomial variables were coded in such a way that their sum equalled 0 (e.g., for stimulus type, shapes were coded as -0.5 and nonwords were coded as + 0.5).
In the random effect structure, we included the random effect of participant, thereby assuming an individual facilitation baseline for each participant. Moreover, we included the random slope of phase on participant to account for the assumption that participants differ in their learning behaviour. To derive the best-fitting model, we followed a stepwise backward selection procedure in which first interactions and then single effects were removed if they did not attain significance at the 5% level. To that end, we assessed the loss of model fit by means of the anova function of the lme4 package.
Results of the analysis are indicated in absolute effect sizes (*β)*, standard errors, *t* values and *p* values. The current version of the lme4 package does not report *p* values as it is currently unknown how to calculate the appropriate number of degrees of freedom (Baayen, 2012). As 132 subjects and more than 9000 observations were included in the analysis, reported *p* values were, therefore, estimated on the basis of the respective *z* values. Moreover, we report the 95% confidence intervals of each fixed effect. Confidence intervals were calculated by means of the confint.merMod function of the lme4 package using the Wald-type confidence intervals which are appropriate for binomial variables.

Legality Judgment

We administered one-sample t-tests to assess whether participants were able to identify target combinations above chance level within each task variant. Chance level was 50% as participants had only two response options in the legality judgment task. As we performed four tests (one t-test for each variant) and multiple testing increases the chance of a type 1 error, we adjusted the significance level of the t-tests following the Holms correction for multiple comparisons (Holm, 1979).

Results

Statistical Learning Task

Only trials with valid click response times (i.e., not too short nor too long response times) were included in the analysis (Baayen & Milin, 2010). This led to the exclusion of seven trials with extremely long response times (> 2500 ms) and one trial with an extremely short response time (< 100 ms) in the relevant blocks (i.e., blocks 13-18). Facilitation scores were limited to those within 2.5 *SD* from the mean facilitation score within each task variant. In total, 9774 trials out of 10080 trials were kept in the analysis (97 % of the data). As noted before, participants only proceeded to the next trial if they had clicked on the correct item. If participants first clicked outside the target area, this was taken into account by delayed response times for that click (and hence affected facilitation scores). Overall, participants rarely clicked outside the target area (participants corrected less than 3 % of the responses to the first target and less than 3.5 % of the responses to the second target). Table 3.1 shows the mean response times and facilitation scores for the four different variants of the statistical learning task. In Figure 3.3, participants' performance on the four variants of the statistical learning task is presented over blocks.

We first performed a control analysis to assess whether facilitation scores changed over the course of the exposure phase. If participants started to implicitly detect the underlying regularities, they should become faster on their second click compared to their first click and, thus, facilitation scores should increase over exposure blocks. Indeed, the analysis indicated a main effect of block with no interaction between block and any of the four task variants. That is, a similar improvement in facilitation score could be observed during the exposure phase across all tasks, which indicates that participants learned throughout exposure. However, as we noted before, this improvement over blocks could either be due to task learning (i.e., learning of the cover task) or could be a true effect of statistical learning. To arrive at a better measure of statistical learning than improvement during exposure, we analysed the effect of phase on facilitation scores in all subsequent

Table 3.1. Mean response times (in ms) and facilitation scores on the four variants of the statistical learning task.

		1st target RT		2nd target RT		Facilitation score	
	n	*M*	*SD*	*M*	*SD*	*M*	*SD*
Auditory nonword matching	30	621.17	167.01	609.60	186.98	1.076	0.326
Visual nonword matching	60	730.72	188.15	742.61	216.04	1.031	0.285
Visual shape matching	60	672.75	174.87	664.55	199.78	1.063	0.303
Visual shape monitoring	60	494.21	128.36	474.63	127.58	1.081	0.290

3

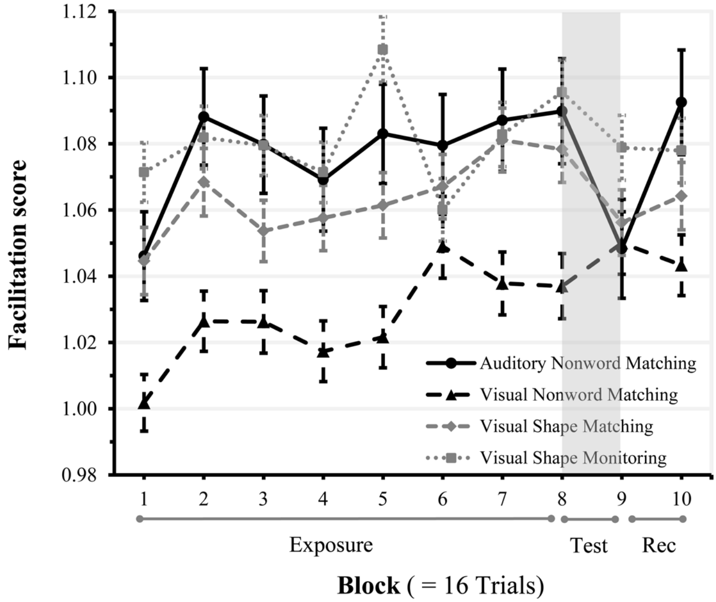

Figure 3.3. Mean facilitation score over blocks in the four variants of the statistical learning task. Error bars represent one standard error from the mean.

analyses of statistical learning. If the underlying regularities are removed during the test phase, participants should only show a drop in facilitation scores if they were sensitive to these regularities and, hence, if they showed statistical learning.

As the visual nonword matching task and the visual shape matching task were administered within the same session, we performed a subset analysis to investigate the effect of order on learning performance in these two tasks. The analysis showed no interaction between order and phase, suggesting that participants‘ amount of learning was similar for the first and second task. Only a main effect of order was observed, indicating that participants had overall higher facilitation scores at the end of the exposure phase in the second task they performed. This effect can simply be attributed to task learning effects. Therefore, the variable of order was not included in the main analysis. Estimates of the best model predicting participants‘ performance as a function of phase, attention, stimulus type and modality are displayed in Table 3.2.

The most parsimonious model showed significant statistical learning across the four task variants, as shown by a simple effect of phase. That is, participants‘ facilitation scores were significantly lower in the test phase compared to the end of the exposure phase, suggesting that participants were indeed affected by removal of the underlying regularities. Facilitation scores were overall lower in the matching tasks compared to the monitoring task (cf. the effect of attention), in the variants including nonwords compared to shapes (cf. the effect of stimulus type) and in the visual tasks compared to the auditory task (cf. the effect of modality). Importantly, the statistical learning effect was modified by stimulus type and by modality: Contrary to our expectation, participants were less sensitive to the statistical co-occurrence of the nonword combinations than of the shape combinations (cf. the interaction between phase and stimulus type). Regarding modality, participants showed a larger learning effect when the targets were cued auditorily (cf. the interaction between phase and modality), in line with previous findings in the literature on modality-specific advantages of auditory learning. The finding that attention did

Table 3.2. Best-fitting statistical model for the facilitation scores in the four task variants.

Fixed effects	*β*	*SE*	*t*	*p*	95 % CI
Intercept	1.081	0.008	135.58	< .001	[1.065, 1.096]
Phase	-0.028	0.011	-2.49	.013	[-0.050, -0.006]
Attention	-0.029	0.012	-2.43	.015	[-0.052, -0.006]
Stimulus Type	-0.024	0.008	-2.95	.003	[-0.041, -0.008]
Modality	0.034	0.014	2.39	.002	[0.006, 0.062]
Position	-0.059	0.006	-9.92	< .001	[-0.071, -0.047]
Alignment	0.072	0.006	12.04	< .001	[0.060, 0.083]
Phase x Attention	-0.018	0.018	-1.00	n.s.	[-0.053, 0.017]
Phase x Stimulus Type	0.039	0.017	2.35	.019	[0.007, 0.072]
Phase x Modality	-0.060	0.022	-2.78	.005	[-0.103, -0.018]
Position x Alignment	-0.036	0.012	-3.00	.003	[-0.059, -0.012]
Random effects		*Variance*	*SD*	*Corr*	
Subject	Intercept	0.003	0.052		
	Test phase	0.001	0.036	-0.08	
Residual		0.086	0.293		

3

not interact with phase indicates that participants were able to detect the underlying patterns in the monitoring task to the same degree as in the matching tasks.

Control variables of target position and target alignment influenced facilitation scores. As expected, participants had overall lower facilitation scores in trials in which the first target was located upper left (cf. the effect of position) and had higher facilitation scores in diagonally aligned targets (cf. the effect of alignment). The interaction between target position and target alignment indicated that the effect of diagonally aligned targets was less pronounced in case the first target was located upper left.

Note that the effects in the current overall analysis report the differences between all tasks. For example, the effect of modality describes the comparison between the auditory variant and all visual variants. We also performed analyses in which the tasks were contrast-coded. This allowed us to directly compare specific task variants with each other (e.g., the auditory nonword task to the visual nonword task). In these analyses, we obtained the same results as reported here: there was no difference between the visual shape monitoring and the visual shape matching task; participants showed more learning in the visual shape matching than in the visual nonword matching task; and participants showed more learning in the auditory nonword matching compared to the visual nonword matching task. As we would need to report multiple models to illustrate these effects, we only report the more straightforward dummy-coded model.

Legality Judgment Task

Table 3.3 presents participants‘ average accuracy in determining legal and illegal target combinations of the four different variants of the statistical learning task. Only in the auditory variant, participants were able to distinguish between legal and illegal target combinations above chance level.

Table 3.3. Mean accuracy (in %) on the legality judgment tasks and results of one-sample t-tests comparing participants' performance to chance level using Holm's-corrected significance level.

	n	*M*	*SD*	*t*	95 % CI	*Holm adjusted p*
Auditory nonword matching	30	56.67	11.71	3.12	[52.29, 61.04]	.016*
Visual nonword matching	60	53.75	12.82	2.27	[50.44, 57.06]	.081
Visual shape matching	60	47.40	12.03	-1.68	[44.29, 50.50]	.198
Visual shape monitoring	60	47.19	13.64	-1.60	[43.66, 50.71]	.198

3

Discussion

The current study was set up to broaden our understanding of the contribution of statistical learning in information processing. Specifically, we aimed to answer the questions whether modality-specific processing differences already emerge (1) if the target stream is held constant and only the modality of the cues to the targets is altered, (2) whether humans may be particularly sensitive to statistical properties in linguistic relative to nonlinguistic material, and (3) whether humans may also learn about underlying regularities if attention to targets is not strictly required for the task at hand.

In line with previous studies on modality-specific processing in statistical learning, we did indeed observe stronger statistical learning if information was presented auditorily. Importantly, however, the targets over which learning occurred were the same in both the auditory and the visual nonword matching task. This indicates that the availability of an auditory cue compared to a visual cue boosts statistical learning in the processing of temporal regularities and that this advantage is in fact an auditory modality advantage, rather than being due to the typical stimulus type selected in the auditory modality.

It should be noted, however, that the auditory statistical learning variant contained visual information throughout the whole task. That is, the written representations of the target and distractor items of a trial were continuously present on the screen. Therefore, it may be argued that the processing advantage we observed in the auditory nonword matching task did not emerge from the auditory nature of the cue but rather resulted from the multimodal integration of two information sources. That is, in the auditory nonword matching tasks, learning may have been facilitated by invoking two different processing routes (i.e., the visual and the auditory processing system) whereas in the visual nonword matching task, relevant information was only processed via the visual system. Multimodal integration has indeed been shown to influence statistical learning (Mitchel, Christiansen, & Weiss, 2014). In the latter study, multimodal integration was found to enhance learning performance if the integration of auditory and visual information altered the percept and, thereby, the statistical properties of the input (by means of the McGurk illusion). That is, participants attended videos of a speaker who produced a stream of CV syllables based on an underlying grammar. In the McGurk condition, the auditory and visual information was partially incongruent. For example, at specific places in the speech stream, participants heard the speaker say /mi/ while they saw the production of /gi/, leading participants to perceive /ni/. As a result, the underlying properties of the speech stream differed compared to an auditory-only condition and a congruent visual-auditory condition, such that integrating audio-visual information and, thereby, being “tricked” into the McGurk illusion allowed participants to learn word boundaries. Importantly, however, in case of a congruent auditory-visual input (where the auditory information matched the visual information), learning performance was lower than in an auditory-only condition in which participants only listened to the speech stream. This suggests that being presented with two congruent sources of information does not necessarily aid statistical learning performance, and, thus, that the processing advantage in the auditory nonword matching task may not readily be attributed to the bimodal processing of relevant information. The auditory processing advantage effect also becomes apparent from participants‘ explicit knowledge of grammatical target combinations. It was only for the auditory nonword matching task that participants were able to correctly identify target combinations above chance level in the legality judgment task.

Even though the modality effect was assessed using linguistic material only (i.e., nonwords), the modality effect seems not to be driven by the linguistic nature of the stimuli. This follows from the other task comparisons: though statistical learning in the auditory nonword matching task outperformed statistical learning in all visual variants, learning of linguistic regularities (i.e., nonword combinations in the visual variant) was not superior to learning of non-linguistic regularities (i.e., shape combinations). It remains unclear, however, whether the same advantage of auditory statistical learning may emerge in the processing of non-linguistic cues (e.g., by using pictures of musical instruments and cueing targets by their characteristic sounds).

Regarding possible processing differences between linguistic and non-linguistic stimuli in statistical learning, we observed a processing advantage for non-linguistic items (i.e., shapes). This contradicted our initial hypothesis, as we had expected participants to be especially sensitive to statistical properties in language given the importance of language processing for human interaction and the key role statistical learning plays in language development and language processing. First of all, it may be argued that the shapes we used in the current study were somewhat linguistic, as our shapes were familiar geometrical figures and were, as such, connected to a specific lemma. It is not obvious, however, that these lemma representations were activated during task performance. Visual representations of familiar geometrical shapes do not activate linguistic representations of the respective shape words in an interference paradigm (Sturz, Edwards, & Boyer, 2014). This provides evidence that processing of geometrical information is isolated from and, hence, can occur without linguistic processing of linguistic labels. Therefore, it is unlikely that the shapes, even though they were familiar to the participants, were processed linguistically. Further research is required to determine whether linguistic information does indeed contribute to statistical learning in the processing of familiar geometrical shapes, for instance by comparing task performance on the current visual shape matching task to performance on a visual variant in which shapes are cued by their orthographic representations. If only the latter variant activated lemma representations and if linguistic processing facilitates learning, learning should differ between both variants.

A possible account of why participants were more sensitive to co-occurrence patterns of non-linguistic, compared to linguistic stimuli, may be because sensitivity to these patterns was tested in the visual modality. That is, humans encounter statistical regularities in language for the most part by listening to auditory input streams. Linguistic stimuli could, therefore, be expected to show a processing advantage specifically in the auditory modality. However, this expectation is not supported by the literature, as a processing advantage for linguistic stimuli has been reported in the visual (Sturm & Smith, 2009) but not in the auditory modality (cf., Bly et al., 2009; Siegelman & Frost, 2015).

Although it is unclear what caused the current observation that there is more statistical learning with non-linguistic than linguistic input, this difference in learnability still holds if both types of stimuli are processed via the same modality. This indicates that there are not only modality-specific processes but also stimulus-specific processes in statistical learning at play. To the best of our knowledge, the only current theory that explicitly accounts for stimulus-specificity in statistical learning is the mechanistic model of statistical learning (Frost et al., 2015). In this model, modality- and stimulus-specific learning is obtained by applying similar learning mechanisms, which are defined as rather domain-general computational and neuronal principles, to modality- and stimulus-specific representations of the input. That is, different stimuli are processed via separate neural networks. This does not only apply to stimuli from different modalities (e.g., visual vs. auditory cortex), but also to different types of stimuli within a modality (e.g., different brain regions for the perception of faces, versus colours or written words in visual cortex). The internal representations which form the input for learning are, thus, modulated by the features of the stimulus-specific processing and encoding system. Consequently, learning is influenced by the specific constraints of the input.

It should be emphasized that the mechanistic model of statistical learning does not argue for a unitary learning system but for domain-general learning principles that are similarly applied across separate neuronal networks. One such underlying neuronal principle for learning may be the repetition suppression effect which is characterized by reduced neuronal activity upon repetition of stimuli (for a review see Grill-Spector, Henson, & Martin, 2006). The repetition suppression effect has consistently been observed (1) across the brain ranging from the level of single cell measurements in macaque monkeys (e.g., Sobotka & Ringo, 1996) to activation patterns of whole brain regions in humans (e.g., Jiang, Haxby, Martin, Ungerleider, & Parasuraman, 2000) and (2) across a wide range of different stimuli such as faces, animals or man-made objects (e.g., Grill-Spector et al., 1999). Importantly, this repetition suppression effect has been shown to be sensitive to the frequency with which repetitions occur in the environment: the repetition suppression effect is stronger in blocks in which repetitions occur frequently compared to blocks in which stimuli are repeated infrequently (Summerfield, Trittschuh, Monti, Mesulam, & Egner, 2008). This suggests that the repetition suppression effect is modulated by expectations and, hence, that predictability effects already play a role in the earliest stages of perception (Summerfield et al., 2008). A learning principle like repetition suppression may therefore bridge the gap between the mechanistic model of statistical learning (Frost et al., 2015) and the embodied theory of statistical learning (Emberson et al., 2011) as the latter theory proposes that perceptual and statistical learning mechanisms are not distinct from each other. Similar learning principles may be at work across different neuronal networks, as put forward by the mechanistic model of statistical learning. Importantly, these learning principles may not purely apply to modality- and stimulus-specific representations but may partially function at the level of sensory input already.

Regarding the role of attention in statistical learning, we did not observe differences in statistical learning between variants in which participants had to attentively process the target shapes to perform a cover task (i.e., in the matching tasks) and a task variant in which attention to the displayed shapes was irrelevant for the task at hand (i.e., in the monitoring task). As participants‘ attention was not actively guided away from the input over which learning could occur, we cannot exclude the possibility that participants were still directing their attention to the target shapes. Compared to more controlled tasks in which attention to the relevant input is either diverted (e.g., Toro et al., 2005; Turk-Browne et al., 2005) or actively suppressed (Kido & Makioka, 2015), we argue that a situation in which the information for learning is irrelevant but processing of the information itself is not hindered is more likely to represent statistical learning in the real world. Importantly, actively guiding attention away from the relevant input for learning (c.f. Toro et al., 2005; Turk-Browne et al., 2005) may not only limit statistical learning performance due to attentional processes but also by increasing participants‘ task load. That is, in order to show statistical learning in these experiments, participants have to implicitly perform a dual task (e.g., detect stimulus repetitions in noises and detect regularities in tones). Therefore, failure to observe statistical learning in these experiments may partially be accounted for by limited working memory capacity for learning. The current study, therefore, extends existing knowledge on the role of attention in statistical learning to a more ecologically valid task manipulation of implicit statistical learning, suggesting that attentive processing of the input is not required for successful statistical learning in everyday life.

Interestingly, participants in the „non-linguistic“ shape tasks could not identify target combinations above chance level, even though they showed implicit statistical learning in the artificial grammar learning - serial reaction time paradigm. This indicates that statistical learning is not directly linked to explicit knowledge of the target combinations which were frequently repeated throughout the learning task. Given that traditional AGL paradigms make use of unfamiliar target sequences to test generalization of the underlying regularities to unfamiliar items and, thereby, learning, such tasks may underestimate participants‘ actual statistical learning performance.

As we argued in the introduction, both language acquisition and language processing seem to be strongly intertwined with mechanisms of statistical learning. Findings on information processing in statistical learning may, therefore, inform us about general aspects of language processing. Modality-specific characteristics of statistical learning such as advantages of auditory learning for temporal regularities and advantages of visual learning for spatial regularities have been widely reported. As spoken languages are temporally structured, it is unclear, however, whether the sensitivity of the auditory statistical learning mechanism to temporal regularities is caused by language experience or whether the importance of auditory statistical learning for language processing is a consequence of its ability to detect temporal regularities. Findings of the current study suggest that the latter account is more likely. Auditory information boosts the learning of temporal regularities in visual displays, which highlights the importance of auditory input for deriving temporal structures. Notably, it is not the linguistic nature of an input that drives the detection of sequential patterns, as we should have found learning in the visual nonword matching task to be superior to learning of shapes then. Overall, our results stress the importance of auditory input for language learning as the implicit detection of temporal regularities and structures, which are at the heart of spoken language, will be specifically facilitated by the availability of auditory information.

In short, our results suggest that modality and stimulus type influence humans‘ performance on statistical learning tasks, thereby highlighting the involvement of these variables in the assessment of statistical learning. Explicit attention to the input was not necessary for participants to detect statistical regularities. Thus, statistical learning may take place even in situations in which humans do not attentively process the input over which learning occurs. Better statistical learning was associated with the processing of non-linguistic input compared to linguistic input and with the availability of auditory information relative to visual information. These findings support recent theories that statistical learning is influenced by modality- and stimulus-specific processing constraints.

References

Altmann, G. T. M., Dienes, Z., & Goode, A. (1995). Modality independence of implicitly learned grammatical knowledge. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 21*(4), 899-912. doi: 10.1037/0278-7393.21.4.899

Arciuli, J., & Simpson, I. C. (2012). Statistical learning is related to reading ability in children and adults. *Cognitive Science, 36*(2), 286-304. doi: 10.1111/j.1551-6709.2011.01200.x

Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research, 3*, 12-28. doi: 10.21500/20112084.807

Baddeley, A. D. (1966). Short-term memory for word sequences as a function of acoustic semantic and formal similarity. *Quarterly Journal of Experimental Psychology, 18*, 362-365. doi: 10.1080/14640746608400055

Bates, D., Maechler, M., & Bolker, B. (2012). lme4: Linear mixed-effects models using S4 classes (Version 0.999-0). Retrieved from http://CRAN.R-project.org/package=lme4

Bly, B. M., Carrion, R. E., & Rasch, B. (2009). Domain-specific learning of grammatical structure in musical and phonological sequences. *Memory & Cognition, 37*(1), 10-20. doi: 10.3758/Mc.37.1.10

Conrad, R. (1964). Acoustic confusions in immediate memory. *British Journal of Psychology, 55*(1), 75-&. doi: 10.1111/j.2044-8295.1964.tb00899.x

Conway, C. M., & Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 31*, 24-39. doi: 10.1037/0278-7393.31.1.24

Conway, C. M., & Christiansen, M. H. (2006). Statistical learning within and between modalities: pitting abstract against stimulus-specific representations. *Psychological Science, 17*, 905-912. doi: 10.1111/j.1467-9280.2006.01801.x

Emberson, L. L., Conway, C. M., & Christiansen, M. H. (2011). Timing is everything: changes in presentation rate have opposite effects on auditory and visual implicit statistical learning. *Quarterly Journal of Experimental Psychology, 64*, 1021-1040. doi: 10.1080/17470218.2010.538972

Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: the paradox of statistical learning. *Trends in Cognitive Sciences, 19*(3), 117-125. doi: 10.1016/j.tics.2014.12.010

Frost, R., Siegelman, N., Narkiss, A., & Afek, L. (2013). What predicts successful literacy acquisition in a second language? *Psychological Science, 24*(7), 1243-1252. doi: 10.1177/0956797612472207

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends in Cognitive Sciences, 10*(1), 14-23. doi: 10.1016/j.tics.2005.11.006

Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzchak, Y., & Malach, R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron, 24*(1), 187-203. doi: 10.1016/S0896-6273(00)80832-6

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6*(2), 65-70.

Janacsek, K., & Nemeth, D. (2013). Implicit sequence learning and working memory: correlated or complicated? *Cortex, 49*, 2001-2006. doi: 10.1016/j.cortex.2013.02.012

Jiang, Y., Haxby, J. V., Martin, A., Ungerleider, L. G., & Parasuraman, R. (2000). Complementary neural mechanisms for tracking items in human working memory. *Science, 287*(5453), 643-646. doi: 10.1126/science.287.5453.643

Kidd, E. (2012). Implicit statistical learning is directly associated with the acquisition of syntax. *Developmental Psychology, 48*, 171-184. doi: 10.1037/a0025405

Kido, K., & Makioka, S. (2015). Serial order learning of subliminal visual stimuli: evidence of multistage learning. *Frontiers in Psychology, 6(76)*. doi: 10.3389/Fpsyg.2015.00076

Misyak, J. B., & Christiansen, M. H. (2012). Statistical learning and language: an individual differences study. *Language Learning, 62*, 302-331. doi: 10.1111/j.1467-9922.2010.00626.x

Misyak, J. B., Christiansen, M. H., & Tomblin, J. B. (2010). Sequential expectations: the role of prediction-based learning in language. *Topics in Cognitive Science, 2*, 138-153. doi: 10.1111/j.1756-8765.2009.01072.x

Mitchel, A. D., Christiansen, M. H., & Weiss, D. J. (2014). Multimodal integration in statistical learning: evidence from the McGurk illusion. *Frontiers in Psychology, 5*, 407. doi: 10.3389/Fpsyg.2014.00407

Neger, T. M., Rietveld, A. C. M., & Janse, E. (2014). Relationship between perceptual learning in speech and statistical learning in younger and older adults. *Frontiers in Human Neuroscience, 8*, 628. doi: 10.3389/fnhum.2014.00628

Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology, 48*(2), 127-162. doi: 10.1016/S0010-0285(03)00128-2

Otsuka, S., Nishiyama, M., Nakahara, F., & Kawaguchi, J. (2013). Visual statistical learning based on the perceptual and semantic information of objects. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 39*(1), 196-207. doi: 10.1037/a0028645

Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology. General, 118*(3), 219-235. doi: 10.1037//0096-3445.118.3.219

Robinson, C. W., & Sloutsky, V. M. (2013). When audition dominates vision: evidence from cross-modal statistical learning. *Experimental Psychology, 60*, 113-121. doi: 10.1027/1618-3169/a000177

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*(5294), 1926-1928. doi: 10.1126/science.274.5294.1926

Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning: listening (and learning) out of the corner of your ear. *Psychological Science, 8*(2), 101-105. doi: 10.1111/j.1467-9280.1997.tb00690.x

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime User's Guide*. Pittsburgh, PA: Psychology Software Tools Inc.

Siegelman, N., & Frost, R. (2015). Statistical learning as an individual ability: theoretical perspectives and empirical evidence. *Journal of Memory and Language, 81*, 105-120. doi: 10.1016/j.jml.2015.02.001

Sobotka, S., & Ringo, J. L. (1996). Mnemonic responses of single units recorded from monkey inferotemporal cortex, accessed via transcommissural versus direct pathways: a dissociation between unit activity and behavior. *Journal of Neuroscience, 16*(13), 4222-4230.

Sturm, J. A., & Smith, K. E. (2009). *The role of Gestalt principles in the acquisition of non-adjacent dependencies in linguistic and non-linguistic sequences.* Paper presented at the Proceedings of the 31th Annual Conference of the Cognitive Science Society, Austin, TX.

Sturz, B. R., Edwards, J. E., & Boyer, T. W. (2014). Asymmetrical interference effects between two-dimensional geometric shapes and their corresponding shape words. *PloS One, 9*(3). doi: 10.1371/journal.pone.0092740

Summerfield, C., Trittschuh, E. H., Monti, J. M., Mesulam, M. M., & Egner, T. (2008). Neural repetition suppression reflects fulfilled perceptual expectations. *Nature Neuroscience, 11*(9), 1004-1006. doi: 10.1038/nn.2163

Teinonen, T., Fellman, V., Naatanen, R., Alku, P., & Huotilainen, M. (2009). Statistical language learning in neonates revealed by event-related brain potentials. *Bmc Neuroscience, 10*. doi: 10.1186/1471-2202-10-21

Toro, J. M., Sinnett, S., & Soto-Faraco, S. (2005). Speech segmentation by statistical learning depends on attention. *Cognition, 97*(2), B25-B34. doi: 10.1016/j.cognition.2005.01.006

Turk-Browne, N. B., Junge, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology. General, 134*(4), 552-564. doi: 10.1037/0096-3445.134.4.552

Vuong, L. C., Meyer, A. S., & Christiansen, M. H. (2011). *Simultaneous online tracking of adjacent and non-adjacent dependencies in statistical learning.* Paper presented at the Proceedings of the 33rd Annual Conference of the Cognitive Science Society, Austin, TX.

Williamson, V. J., Baddeley, A. D., & Hitch, G. J. (2010). Musicians' and nonmusicians' short-term memory for verbal and musical sequences: comparing phonological similarity and pitch proximity. *Memory & Cognition, 38*(2), 163-175. doi: 10.3758/Mc.38.2.163

4

Relationship between Statistical Learning and Perceptual Learning in Speech

This chapter is based on:

Neger, T. M., Rietveld, T., & Janse, E. (2014). Relationship between perceptual learning in speech and statistical learning in younger and older adults. Frontiers in Human Neuroscience, 8: 628. doi:10.3389/fnhum.2014.00628.

Abstract

Within a few sentences, listeners learn to understand severely degraded speech such as noise-vocoded speech. However, individuals vary in the amount of such perceptual learning and it is unclear what underlies these differences. The present study investigates whether perceptual learning in speech relates to statistical learning, as sensitivity to probabilistic information may aid identification of relevant cues in novel speech input. If statistical learning and perceptual learning (partly) draw on the same general mechanisms, then statistical learning in a non-auditory modality using non-linguistic sequences should predict adaptation to degraded speech. In the present study, 73 older adults (aged over 60 years) and 60 younger adults (aged between 18 and 30 years) performed a visual artificial grammar learning task and were presented with 60 meaningful noise-vocoded sentences in an auditory recall task. Within age groups, sentence recognition performance over exposure was analyzed as a function of statistical learning performance, and other variables that may predict learning (i.e., hearing, vocabulary, attention switching control, working memory, and processing speed). Younger and older adults showed similar amounts of perceptual learning, but only younger adults showed significant statistical learning. In older adults, improvement in understanding noise-vocoded speech was constrained by age. In younger adults, amount of adaptation was associated with lexical knowledge and with statistical learning ability. Thus, individual differences in general cognitive abilities explain listeners' variability in adapting to noise-vocoded speech. Results suggest that perceptual and statistical learning share mechanisms of implicit regularity detection, but that the ability to detect statistical regularities is impaired in older adults if visual sequences are presented quickly.

Introduction

Listeners' ability to rapidly learn to understand unfamiliar speech conditions such as accented, disordered or noise-vocoded speech is impressive. Within a few sentences, listeners learn to map a new type of speech input onto their old percept, some improving their speech recognition performance by more than 60 % (Eisner, McGettigan, Faulkner, Rosen, & Scott, 2010). However, listeners show great variability in the amount of such perceptual learning (Eisner et al., 2010). This raises the question which mechanisms underlie perceptual learning.

Perceptual learning can be defined as „relatively long-lasting changes to an organism's perceptual system that improve its ability to respond to its environment" (Goldstone, 1998, p. 585). As listeners are not able to describe the changes that led to their improved perception, perceptual learning is assumed to be a type of implicit learning (Fahle, 2006). A conceptual framework that accounts for changes in the perceptual system is the Reverse Hierarchy Theory (RHT) (Ahissar & Hochstein, 2004). The RHT argues that perceptual learning is a top-down guided process. When a listener is exposed to a novel speech condition, initial performance fails as the speech input can no longer be readily matched to higher-level representations such as word representations. According to the RHT, prolonged exposure modifies these higher-level representations, which subsequently enables top-down guidance to retune weights at lower levels of the processing hierarchy: the weights of task-relevant input are increased and the weights of task-irrelevant input are pruned. This process of weight retuning starts at the highest level of the hierarchy and continues gradually to the lower levels (i.e., the reverse hierarchy). When lower-level representations have been modified, performance under difficult conditions can be based on accessing these low-level representations. This is illustrated by findings that adaptation to noise-vocoded speech generalizes to novel words (Hervais-Adelman, Davis, Johnsrude, & Carlyon, 2008), to non-words (Loebach, Bent, & Pisoni, 2008) and to the recognition of environmental sounds (Loebach, Pisoni, & Svirsky, 2009). These generalization findings suggest that perceptual learning in speech modifies representations at lower levels of the hierarchy, that is, representations at a sublexical level (Banai & Amitay, 2012; Hervais-Adelman et al., 2008).

The RHT has been influential in explaining behavioral observations in visual and auditory perceptual learning (Banai & Amitay, 2012; Cohen, Daikhin, & Ahissar, 2013; Nahum, Nelken, & Ahissar, 2010; Sabin, Clark, Eddins, & Wright, 2013). However, the RHT does not specify which processes take place in the initial stages of adaptation that enable the perceptual system to identify task-relevant cues in the input and to modify high-level representations. One of the basic principles in the RHT and other models of perceptual learning is the retuning of weights based on the relevance of features or dimensions for the specific task (Ahissar & Hochstein, 2004; B. A. Dosher & Lu, 1999; Goldstone, 1998; Petrov, Dosher, & Lu, 2005). This principle implies that stimuli have to share certain features, which can thus be considered task-relevant, for perceptual learning and for transfer of learning to occur. Accordingly, several studies have highlighted the importance of structural regularities (Cohen et al., 2013) and of stimulus consistencies for perceptual learning (e.g., Nahum et al., 2010). In other words, for learning to occur, participants need to detect specific regularities in the input. Therefore, individual differences in sensitivity to such regularities may indicate why listeners differ in adapting to unfamiliar speech input.

An implicit learning mechanism that has been linked to pattern sensitivity is statistical learning. Statistical or probabilistic learning describes the ability to implicitly extract regularities from an input by detecting the probabilities with which properties co-occur (Misyak & Christiansen, 2012). Statistical learning has gained increasing attention over the past years in language research, as language itself is probabilistic in nature (Auer & Luce, 2005). Accordingly, co-occurrence probabilities of units have been shown to facilitate processing at various linguistic levels (e.g., effects of phonotactic probability (e.g., Vitevitch, Armbruster, & Chu, 2004) or transitional probability (e.g., Thompson & Newport, 2007). Statistical learning has been found to be of major importance in language acquisition (Saffran, 2003). Also in adulthood, individual differences in statistical learning have been shown to predict sentence processing performance (Misyak & Christiansen, 2012). Moreover, deficits in statistical learning ability have been reported for various language-related disorders such as specific language impairment (Evans, Saffran, & Robe-Torres, 2009), agrammatic aphasia (Christiansen, Kelly, Shillcock, & Greenfield, 2010) and language-based learning disabilities (Grunow, Spaulding, Gómez, & Plante, 2006). As statistical probabilities are provided and continuously updated by the input, relying on statistical probabilities actually enables language users to adapt to their environment, which is the essential characteristic of perceptual learning. Therefore, the present study aims to investigate whether statistical learning relates to perceptual learning in speech perception. If adaptation to a novel speech condition and statistical learning share general mechanisms of implicit regularity detection, then statistical learning performance in a non-auditory modality using non-linguistic stimuli should predict individuals‘ perceptual learning for speech comprehension.

Perceptual learning in speech and statistical learning may also draw (partly) on the same underlying cognitive abilities, such as working memory and attention. Therefore, we investigated whether both types of learning could be predicted from general cognitive and linguistic abilities. Ahissar and Hochstein (2004) proposed that attentional mechanisms may be engaged in choosing which neuronal populations pass on task-relevant information to the higher levels and in increasing the functional weights of these populations. Several frameworks of perceptual learning incorporate the idea that attentional mechanisms are involved in perceptual learning (e.g., B. A. Dosher, Han, & Lu, 2010; Fahle, 2006; Goldstone, 1998). A study on frequency discrimination found that perceptual learning even occurred after training with non-discriminable stimuli (Amitay, Irwin, & Moore, 2006). Apparently, training directed the participants‘ attentional focus to the relevant stimulus dimension, which was sufficient to access the relevant low-level representations during the test phase (Amitay et al., 2006). Moreover, performance on a selective attention task predicted the amount of learning in adaptation to accented speech (Janse & Adank, 2012). Further evidence that attention is involved in perceptual learning comes from studies in which listeners were simultaneously exposed to noise-vocoded speech and both auditory and visual distractors (Huyck & Johnsrude, 2012; Wild et al., 2012). Only listeners who attended the noise-vocoded stimuli showed improved performance in recognizing noise-vocoded speech. Similar effects of attentional focus arise in tasks of visual statistical learning. When observers are asked to attend to symbols of a certain color in a two-color symbol stream, statistical learning effects unfold for regularities within the attended color but not for regularities within the unattended color (Turk-Browne, Junge, & Scholl, 2005). These findings imply that only attended features are effectively learned. It has been proposed that training procedures that facilitate participants to *switch* their attention to appropriate perceptual features (e.g., fixed temporal presentation of multiple stimuli, repeated presentation) may particularly enhance perceptual learning (Ju. Zhang et al., 2008). Therefore, attention switching control may be involved in the process of distinguishing relevant from non-relevant features in tasks of implicit learning.

Another cognitive ability that may be involved in tasks of implicit learning is working memory, which is required to simultaneously store and process auditory or visual information (Gathercole, 1999). Performance on working memory tasks has been shown to predict performance in various speech reception tasks (for a review see Akeroyd, 2008) and, more specifically, there are indications that working memory relates to perceptual learning performance. Teenaged students with learning and reading disabilities who participated in perceptual learning tasks of frequency and duration discrimination showed improved working memory skills after training (Banai & Ahissar, 2009). Furthermore, the two students who failed to show perceptual learning were characterized by the poorest working memory capacity in the sample. During training, students were repeatedly presented with the same stimuli, which allowed them to access low-level representations, thereby improving frequency and duration discrimination. Thus, working memory may have aided perceptual learning by keeping stimuli accessible (also see Goldstone, 1998). In contrast to these findings, Erb and colleagues (2012) did not find an association between working memory and individual differences in adaptation to noise-vocoded speech. Note, however, in this study, working memory was measured by tasks that relied on immediate recall and, hence, on short term memory (i.e., nonword repetition task, digit span forward task). Possibly, more complex span tasks, that measure the ability to simultaneously store and process information, rather than just recall capacity, may be particularly associated with tasks of perceptual learning. With respect to statistical learning, recent studies reported correlations between working memory capacity and performance on implicit sequence learning tasks (Bo, Jennett, & Seidler, 2011, 2012). However, findings regarding the link between working memory and implicit learning of sequences are controversial (for a review see Janacsek & Nemeth, 2013) and it has been argued that working memory as an executive resource is not involved in tasks of implicit learning (Kaufman et al., 2010).

An additional cognitive ability that should be considered is processing speed. Processing speed reflects the efficiency of a processing system to perform simple operations (Kaufman et al., 2010) and as a general index of processing efficiency, may be assumed to facilitate perceptual learning. Previous research showed that processing speed correlates with performance on tasks of implicit sequence learning (Kaufman et al., 2010; Salthouse, McGuthry, & Hambrick, 1999). Higher efficiency of the processing system may be beneficial at various stages of the adaptation process. In the framework of the RHT, processing speed may reduce listeners‘ time to retrieve high-level representations and to initiate modification processes. Furthermore, processing speed may accelerate the process of weight retuning, thereby gaining faster access to low-level representations.

As the current study focuses on adaptation for spoken language understanding, perceptual learning may also draw on linguistic knowledge. Davis and colleagues (2005) presented data on how the so-called pop-out effect accelerates the process of perceptual learning:

if listeners knew the content of what was going to be said before they actually heard the sentence in its degraded form, this benefited their perceptual learning. In line with the Eureka effect in the RHT, in which a cue regarding the content of the stimulus can trigger direct perception of the stimulus and facilitates strong and long-lasting learning effects (Ahissar & Hochstein, 2004), this pop-out finding suggests that lexical knowledge facilitates access to higher-level representations, thereby initiating top-down processes that aid sublexical retuning (Davis et al., 2005). Accordingly, vocabulary, as a measure of lexical knowledge, has been found to predict the amount of perceptual learning in listeners who were adapting to an unfamiliar foreign-sounding accent (Janse & Adank, 2012), accents being linguistic degradations of the stimulus. If we assume that lexical knowledge aids perceptual learning by guiding the top-down search, effects of lexical knowledge should also arise in non-linguistic speech degradations. Therefore, we investigate whether linguistic knowledge, as indexed by vocabulary knowledge, may also facilitate shifting of attention to relevant features of acoustically degraded speech.

As we want to investigate which cognitive processes are involved in perceptual learning in speech, we also aim to test whether our findings generalize to a heterogeneous group of listeners. Older adults typically form a highly heterogeneous group, as perceptual and cognitive processing undergo changes over the life span. Age-related changes in hearing acuity (Lin, Thorpe, Gordon-Salant, & Ferrucci, 2011), processing speed, capacity on working memory tests, attentional control (for a review see Park & Reuter-Lorenz, 2009) but also lexical knowledge (Ramscar, Hendrix, Shaoul, Milin, & Baayen, 2014) may therefore help to identify relevant cognitive processes. Importantly, the ability to adapt to unfamiliar speech input is preserved throughout the life span (Adank & Janse, 2010; Golomb, Peelle, & Wingfield, 2007; Gordon-Salant, Yeni-Komshian, Fitzgibbons, & Schurman, 2010; Peelle & Wingfield, 2005). However, differences in the amount and pattern of perceptual learning over exposure between younger and older adults also indicate changes in the underlying processes. While younger and older listeners show the same amount of learning in the initial adaptation phase, older listeners' performance plateaus earlier in adapting to unfamiliar speech (Adank & Janse, 2010; Peelle & Wingfield, 2005), older adults show less transfer of learning to similar conditions (Peelle & Wingfield, 2005), and exhibit slower consolidation of learning (Sabin et al., 2013). Such differences illustrate that the interdependency between cognitive functions and implicit learning processes may change as a function of age. Cognitive abilities associated with adaptation to unfamiliar speech in younger adults may not be the same as in older adults. In order to gain more insights into individual abilities associated with adaptation to unfamiliar speech across the life span, we tested both younger and older adults.

In sum, this study investigates perceptual learning for spoken language understanding in younger and older adults. We use noise-vocoded speech, an acoustic degradation of the speech signal which simulates the auditory signal of a cochlear implant. In contrast to naturally occurring variability in speech (such as accents), participants do not encounter noise-vocoded speech in everyday life. As a consequence, all participants share the same naïve exposure level. We specifically study whether perceptual learning is associated with a general ability to implicitly detect statistical regularities. By testing participants' probabilistic sequence learning with visual non-linguistic stimuli, we apply a rigorous test of the association between the two types of implicit learning. Additionally, we investigate whether both types of implicit learning are associated with individual differences in attention switching control, working memory, information processing speed or lexical knowledge.

Method

Participants

In total, 60 younger and 73 older adults participated in the current study. All participants were native speakers of Dutch, neurologically intact and had no history of language disorders. One younger participant was excluded as he showed floor performance throughout the perceptual learning task (i.e., he did not understand the noise-vocoded speech at all). Younger adults were aged between 18 and 29 years (mean age 21years, sd 2.5 years) and older adults were aged between 60 and 84 years (mean age 68.4 years, sd 5.7 years). In both age groups, the majority of participants were female (53 out of 59 participants in the younger and 47 out of 73 participants in the older sample). Participants had normal or corrected-to-normal vision. Participants were recruited via the subject database of the Max Planck Institute for Psycholinguistics and were compensated €8 per hour for their time.

Auditory measure

Hearing thresholds

Age-related hearing loss is prevalent in older adults (Lin et al., 2011). Poorer hearing may affect perceptual learning as auditory input contains less detail, thereby interfering with accessing and retuning low-level representations. Participants‘ auditory function was assessed by measuring air-conduction pure tone thresholds with the aid of an Oscilla USB-300 screening audiometer. As age-related hearing loss particularly affects sensitivity to high frequencies, a high-frequency pure tone average [PTAH] was taken as index of hearing acuity. This PTAH was calculated as the mean hearing threshold over 1, 2 and 4 kHz (instead of the standard PTA over 0.5, 1, and 2 kHz). Only the PTAH of the best ear was entered in the analysis, as all auditory stimuli were presented binaurally. Twenty-seven older participants actually qualified for hearing aids on the basis of their hearing thresholds according to the standard of hearing-aid coverage in the Netherlands (PTAH of the worst ear ≥ 35 dB HL). None of the participants wore hearing aids in daily life, however. Higher thresholds reflected poorer hearing. Mean thresholds at different frequencies per age group are given in Figure 4.1.

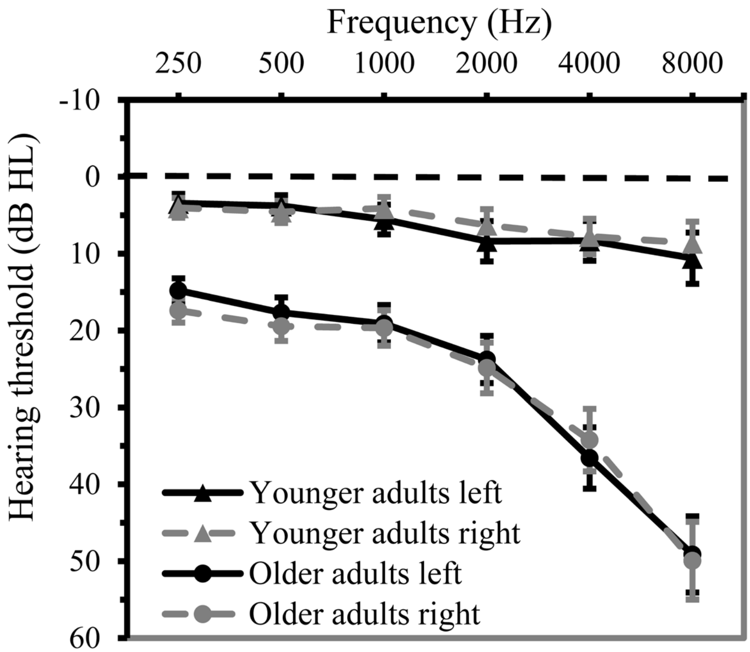

Figure 4.1. Mean hearing threshold (in dB HL) at 0.25, 0.5, 1, 2, 4 and 8 kHz for both ears in younger (n = 59) and older adults (n = 73). Error bars indicate two standard error from the mean.

Cognitive measures

Working memory

Participants performed a digit span backward task as an index of working memory capacity. The test was a computerized variant of the digit span backward task included in the Wechsler Adult Intelligence Scale Test (Wechsler, 2004) and presented via E-prime 1.2 (Schneider, Eschman, & Zuccolotto, 2002). Participants were asked to report back sequences of digits in reverse order. Digits were presented in a large white font (Arial, font size 100) against a black background. Each digit was presented for 1 s with an interval of 1 s between the consecutive digits of a sequence. Sequence length increased stepwise from two to seven digits and performance on each sequence length was tested on

4

two different trials (all participants were presented with all sequence lengths, regardless of their performance on earlier easier trials). The actual test trials were preceded by two practice trials with a sequence length of three to familiarize participants with the task. Participants had to recall 12 test sequences in total. Individual performance was operationalized as the proportion of correctly reported sequences (out of 12).

Processing speed

Information processing speed was assessed by means of a digit symbol substitution task. Participants had to convert as many digits as possible into assigned symbols in a fixed amount of time (90 s). The digit symbol substitution task is a paper-and-pencil test that was derived from the Wechsler Adult Intelligence Scale Test (Wechsler, 2004). Performance was measured by the number of correctly converted digits in 90 seconds, meaning that higher scores reflected higher information processing speed.

Attention switching control

The Trail Making Test was administered to obtain a measure of attention switching control. The paper-and-pencil test contained two parts. In Part A, participants were asked to connect numbers as quickly as possible in ascending order (i.e., 1-2-3…), the numbers being spread randomly over a white page. The Part B page had both numbers and letters randomly spread over the page. Participants now had to alternately join numbers and letters in ascending order (i.e., 1-A-2-B-3-C…). In both parts, 25 items had to be connected and the total time to complete each part was measured. We calculated the ratio between both parts (Part B / Part A) as measure of attention switching control (Arbuthnott & Frank, 2000), thereby taking general slowing into account (Salthouse, 2011; Verhaeghen & De Meersman, 1998). Higher scores indicated higher costs of switching between letters and numbers, therefore, poorer attention switching control.

Linguistic measure

Vocabulary knowledge

A vocabulary test in the form of multiple choice questions was administered to obtain a measure of linguistic knowledge (Andringa, Olsthoorn, van Beuningen, Schoonen, & Hulstijn, 2012). The computerized test was administered in Excel (Courier font size 15). Participants had to indicate which out of five possible answers was the correct meaning of Dutch low-frequency words, the last alternative always being ‚I don‘t know‘. Words were not domain-specific and each target word was embedded in a different, neutral carrier phrase. The vocabulary test consisted of 60 items. There was no time limit or pressure to complete the test. Performance was measured by test accuracy, that is, the proportion of correct answers (out of 60). Higher scores thus reflected greater vocabulary knowledge.

Statistical learning

Materials and design

To investigate statistical learning, we adopted the artificial grammar learning - serial reaction time paradigm (Misyak, Christiansen, & Tomblin, 2010a). This paradigm has typically been used in studies on statistical learning in language processing and has been found to link to individual language processing abilities (Misyak & Christiansen, 2012; Misyak et al., 2010a; Misyak, Christiansen, & Tomblin, 2010b). As artificial grammar learning simulates language learning processes, the task makes use of auditory presented sound sequences such as nonwords. However, as we wanted to investigate whether individuals' ability to adapt to an unfamiliar speech condition could be predicted by a general ability to implicitly detect regularities, we used visual and non-linguistic stimuli in the statistical learning task. That is, we applied a rigorous test for the relationship between statistical learning and perceptual learning by preventing that a relationship between both measures of learning was specific for auditory and linguistic processing.

Participants were presented with familiar, geometrical shapes in a 2x2 design on the computer screen (see Figure 4.2B), in which one shape on either side of the screen served as target and one as distractor item. Target shapes were sequentially highlighted by a visual marker and participants' task was to click as fast as possible on the highlighted target. The first target was always one on the left side of the screen (i.e., upper or lower one in the first column) and the second target was always on the right side of the screen (i.e., upper or lower one in the second column). The second target was only highlighted after the participant had clicked on the first target item. Crucially, which of the two items in the right-hand column would be highlighted was predictable on the basis of the first target (e.g., in b, a *triangle* would always be followed by a *star* or a *square (the latter is not in the display)*, but never by a *heart*).
Materials consisted of eight familiar, geometrical shapes drawn with a single, continuous black line. The shapes were divided into two grammatical subsets of four shapes each (i.e., Set 1: *triangle, hexagon, star, square;* Set 2: *arrow, circle, heart, cross)*. Within each set, two items were selected to appear as first targets (i.e., Set 1: *triangle, hexagon;* Set 2: *arrow, circle)* and were always followed by one of the other two items that served as second targets (i.e., Set 1: *star, square*; Set 2: *heart, cross)*. Therefore, four combinations

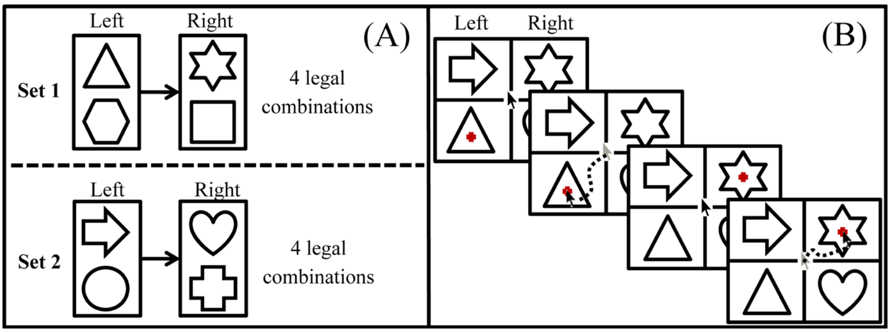

Figure 4.2. Structure of the statistical learning task. (A) Structure of the grammar in which the first target is always displayed on the left side of the screen and the second target is always displayed on the right side of the screen. (B) Procedure of a grammatical trial during the exposure phase.

of shapes were grammatical within each set, resulting in a total set of eight grammatical combinations (see Figure 4.2A). Target items were presented along with distractors in a rectangular grid display on the computer screen (see b). Distractor items were shapes from the subset that was currently not tested and the two distractor shapes on the screen formed a grammatical combination themselves. Thus, within a grammatical trial, the transitional probability from the first to the second target was 1, as the first target could only be followed by the target from the same subset. Within the grammar, however, the transitional probability between two adjacent items was 0.5, as a target was followed by a specific successor only half of the time (i.e., a *circle* being followed by either a *heart* or a *cross*, see a). Target positions were randomly assigned such that it was unpredictable whether a first or second target would be displayed in the upper or lower row of a particular column.

The artificial grammar learning task was composed of blocks and split into an exposure phase, a test phase and a recovery phase. During the exposure phase, participants could learn the grammar by picking up on the co-occurrence probabilities of the shapes. In total, the exposure phase consisted of sixteen grammatical blocks. Within each block, all grammatical combinations were repeated once, resulting in 128 exposure trials (8 x 16). The test phase consisted of two ungrammatical blocks (2x8 trials). In these ungrammatical blocks, the original grammar was reversed, such that a target was followed by targets of the other (competing) subset. Participants who implicitly learned the grammar should show a drop in performance as they would need to correct their predictions, resulting in a slowed response to the second target. This measure of learning is widely accepted in the literature on implicit learning (Janacsek & Nemeth, 2013): a drop in performance due to removing the underlying regularities can only be linked to grammar sensitivity, whereas learning measures in terms of improvement during the exposure phase cannot be teased apart into general task learning and statistical learning. Therefore, statistical learning was operationalized by the difference in task performance between the last four blocks of the exposure phase (blocks 13 –

16) and the subsequent ungrammatical test phase (blocks 17 – 18). The recovery phase again consisted of two grammatical blocks and serves as a control phase. If participants learned the grammar, by re-introducing the regularities in the recovery phase, participants' performance should not decrease any further. In total, the artificial grammar learning task thus contained twenty blocks and 160 trials (8 x 20).

Procedure

The artificial grammar learning task was presented in E-prime (Schneider et al., 2002) and started with five practice trials that were all grammatical. Participants were instructed to click as quickly as possible on target shapes that were marked by a small filled red cross (10 x 10 mm) in the center of the target shape. Participants were informed that they had to click on two successive targets and that the first target would be located in the first column and the second target would be located in the second column. Each trial started with the presentation of the visual display that consisted of the four shapes and two grid lines, marking the four quadrants on the screen. At the start of each trial, the mouse cursor was located in the center of the screen. Each shape was displayed in a size of 75 x 70 mm. The visual marker appeared in the middle of the first target shape 500 ms after the onset of the visual display, and was shown until the participant clicked on the marked picture. After the participant had responded, the mouse cursor was automatically set back to the center of the screen to ensure the same distance for all click responses. The second visual marker (same red cross now marking the second target shape) appeared 500 ms after the first click. This time interval was implemented in the design to allow for prediction effects, even in the adults who had slower processing. This time interval had been successfully applied in an earlier study on implicit sequence learning in older adults (Salthouse et al., 1999). Participants could not make errors: the experiment only proceeded if a participant clicked on the appropriate target shape. Clicking on a distractor shape or outside the target picture before giving a correct click resulted in a higher RT.

The intertrial-interval was 500 ms. After each block, a small break of 2500 ms was implemented to avoid fatigue effects. During this break, participants saw the block number of the upcoming block and a reminder to click as quickly as possible. It took approximately 20 minutes to complete the task.

To assess statistical learning, we measured latencies from target highlighting to the subsequent mouse response. Facilitation scores were calculated to index individuals' sensitivity to implicit regularities. The facilitation score was calculated by dividing the reaction time (RT) to the first, unpredictable target within a trial by the RT to the second, predictable target within the same trial. Thus, RT to the first target served as baseline performance within each trial. This was important to minimize biases of task learning and motor performance, particularly for those older adults who may have had little practice in using a computer mouse. During the course of the experiment, RTs may generally get faster as older adults get more experienced in using a mouse. By implementing a new baseline within each new trial, such motor learning should be accounted for. If participants cannot predict which target will be highlighted next, their RTs to both targets within a trial will be similar and will result in a facilitation score of 1. During the exposure phase, learning manifests itself in an increasing facilitation score. That is, if participants learn to predict the second target, RTs to the second item will be faster and, therefore, shorter compared to the first, unpredictable target RTs.

Perceptual learning

Materials and design

Sixty Dutch sentences were noise-vocoded to create an unfamiliar speech condition to which participants needed to adapt. In noise-vocoded speech, frequency information in the signal is replaced by noise while preserving the original amplitude structure over time. The speech signal was split into multiple non-overlapping frequency bands, which approximately matched equal distances on the basilar membrane (Greenwood, 1990). From each frequency band the smoothed amplitude envelope was derived and imposed on wide-band noise in the same frequency range. In a last step, these modulated noise bands were recombined, creating a speech signal that sounded like a harsh robot voice. All signal editing was done in Praat (Boersma & Weenink, 2011).

An important characteristic of noise-vocoded speech is that the comprehension level of the speech signal can easily be manipulated by varying the number of frequency bands. The more frequency bands are used to decompose the speech signal, the more detail of the original temporal and amplitude structure is preserved and the more intelligible the speech signal is. Previous research has shown that ten frequency bands are enough for naïve listeners to immediately understand more than 90 % of noise-vocoded speech (Sheldon, Pichora-Fuller, & Schneider, 2008). However, when presented with speech noise-vocoded with fewer bands, participants only reach this level of performance after a certain amount of exposure.

The maximal amount of learning or intelligibility improvement can be observed if the starting level is neither too high nor too low, so that sufficient information can be derived from the acoustic materials to initiate learning while at the same time allowing for sizeable improvement (see Peelle & Wingfield, 2005). We initially tried to provide participants with an individual starting level from which they could still show improvement. In a separate pilot study, we therefore assigned 23 older adults to a specific noise-vocoding condition (i.e., 4 or 6 bands) on the basis of their performance on a speech reception threshold (SRT) task in noise. Inspection of the data showed that participants‘ starting level clustered according to band condition. Older adults in the 4 band condition showed a very low starting level (on average they understood only 10 % of the sentences correctly), whereas older adults in the 6 band condition showed a very high starting level (on average they already understood 65 % of the sentences). Relatedly, the correlation between SRT result and initial performance on the noise-vocoded speech was weak. As our attempt to individualize starting levels on the basis of a speech-in-noise task was not successful, we aimed to provide a roughly similar starting level for both age groups. Based on the results of the pilot study, we decided to present older adults with speech that was vocoded with 5 bands (corner values using 5 frequency bands: 50-280-757-1742-3781-8000 Hz). As younger adults understand more when being exposed to the same degradation as older adults (Peelle & Wingfield, 2005; Sheldon et al., 2008), we presented younger adults with four-band speech (corner values using 4 frequency bands: 50-369-1161-3125-8000 Hz), thus, a more difficult speech condition than older adults (cf., Golomb et al., 2007). Consequently, we were able to see sizeable and comparable amounts of improvement over the course of exposure in both age groups.

Sentences were selected from audiological test materials (Versfeld, Daalder, Festen, & Houtgast, 2000) and were all produced by the same, male speaker. Each sentence had a length of eight or nine syllables and contained four keywords. Keywords in the selected set of sentences included a noun, verb and preposition. The fourth keyword was an adjective, adverb or a second noun. An example sentence *'De sneeuw glinstert in het maanlicht'* ('*The snow is glistening in the moonlight'*) contained the keywords '*sneeuw'*, '*glinstert'*, '*in'* and '*maanlicht*. Note that five additional sentences were selected for practice purposes, so that there was no overlap in sentence content between practice and test items. Practice sentences had the same length as test items (a list of all sentences used in the current study is provided in appendix A1).

Procedure

An auditory sentence identification task was administered to investigate perceptual learning using the experiment program E-prime (Schneider et al., 2002). Participants listened to the noise-vocoded sentences and were asked to identify and repeat these sentences. They were encouraged to guess if they were unsure. Participants were first presented with five practice trials. First, participants listened to three clear sentences to familiarize them with the task and the speaker. Moreover, these practice trials were used to check whether participants' memory span was sufficient to perform the task given clear input, which was the case for all participants. Then participants listened to two sentences that were noise-vocoded with only two frequency bands to present them with the type of degradation. This more difficult condition with fewer bands was chosen to make sure that no learning could occur during the practice phase (e.g., Ahissar & Hochstein, 1997; Liu, Mercado, Church, & Orduna, 2008; Pavlovskaya & Hochstein, 2004). Practice trials were identical for all participants and were presented in the same order. In contrast, the sixty test sentences were presented in random order for each participant, so that observed learning effects would be independent of inherent intelligibility differences between sentences (e.g., due to differences in semantic predictability). Participants heard a short (125 ms) 3.5 kHz tone to call their attention to the upcoming stimulus 500 ms before sentence onset. After each sentence, the researcher scored the number of correctly repeated keywords (0-4) online. The next trial started immediately after the researcher had confirmed the scoring of the previous trial. Auditory stimuli were presented binaurally via dynamic closed, circumaural headphones (Sennheiser HD 215), at a level of 85 dB SPL. Participants‘ answers were audiorecorded to allow for later checking of their responses.

Experimental procedure

Measures of younger adults were obtained in a single experimental session. Testing was spread over two sessions for the older adults, as they also participated in a different study. During the first session, older adults performed the background measures described above. The second session consisted of the statistical learning and the perceptual learning task and followed within a month on the first session. In both age groups, tasks were presented in a fixed order. Although the order differed between younger and older adults, the statistical learning task was always presented before the perceptual learning task. All participants were tested individually in a sound-attenuating booth to minimize distraction. Before the start of each task, participants received verbal and printed task instructions. Participants could ask questions at any time. Between tasks, participants were encouraged to take small breaks.

Data analysis

Statistical modeling

To assess learning performance, we implemented linear mixed-effects models using the lmer function from the lme4 package (Bates, Maechler, & Bolker, 2012) in R (version 2.15.1). In this way, both participants and items could be assessed as random factors and the maximal random slope structure of models could be defined to reduce the probability of a type 1 error (Barr, Levy, Scheepers, & Tily, 2013). First, we modeled statistical and perceptual learning performance as a function of age group to assess whether younger and older adults differed in their learning performance. Second, we analyzed the contributions of individual abilities in learning separately within each group as our focus was on individual differences within the respective age groups. Thus, the modeling process that is described here was applied to the statistical learning data and to the perceptual learning data of both age groups.

Linear regression models are based on the assumption that the predictors included in the analysis do not show collinearity (Baayen, 2012; Field, Miles, & Field, 2012). Although some predictor measures were intercorrelated (see the Results section on participants' performance on background measures), we did not control for these intercorrelations for two reasons. First, most correlations explained less than 20 % of the variance in the correlated measure (i.e., with correlation coefficients below 0.45). Only the correlation between age and speed in the older adults was moderately correlated ($r = -.562$). Second, simultaneous inclusion of correlated measures in the analysis has been shown to provide a more reliable interpretation of estimates than inclusion of residualized variables (Wurm & Fisicaro, 2014; York, 2012).

Statistical learning was defined as a drop in performance in the test phase (blocks 17 - 18) compared to the performance at the end of the exposure phase (blocks 13 - 16). Therefore, in models of statistical learning, the fixed categorical variable phase (exposure vs. test phase) was the variable of interest to predict individuals' facilitation scores and to indicate learning. Additionally, two control variables and the corresponding two- and three-way interactions with phase were included in models of statistical learning. Control variables were the categorical variable 'first target position' (was the first target displayed in the upper or lower row of the left column?) and the categorical variable 'target alignment' (were the two targets in a trial aligned horizontally or diagonally?). Given the directionality of Western writing systems, we expected a first target position effect as participants may click faster on a target in the upper left quadrant than in the lower left quadrant. We also expected the drop in facilitation score during the test phase to be less distinct in trials with the first target appearing in the upper left quadrant, such that target position was expected to interact with the amount of learning. Moreover, the alignment of targets was thought to affect second target RTs. Note that the experimental program always set the mouse back to the centre of the screen after each click. Despite this automatic mouse reset, participants tended to also move the mouse back to the centre of the screen. By doing that, participants unintentionally initiated a movement towards the diagonal shape. Therefore, we assumed that participants would be faster in responding to the second target if targets were arranged diagonally rather than horizontally (see), which would result in higher facilitation scores. This direction effect may interact with the effect of removing the regularities, such that the grammaticality effect be decreased for the diagonal movements.

In models of perceptual learning, the number of correctly repeated keywords per sentence served as index of recognition performance and was entered as numerical dependent variable into the model. As perceptual learning was defined as the improvement in speech understanding over exposure, we split the experiment into six blocks, containing ten sentences each and added block as numerical measure of exposure to the model. However, before block was included in the analysis, we performed a log-transformation of block, as perceptual learning has typically been described by fast initial learning that levels off with increasing exposure (see also Figure 4.4). The transformation of block therefore provided us with an index of exposure that took this non-linear improvement curve into account and converted the improvement over exposure into a linear scale[1].

In the first step of the analysis, we identified the maximal random slope structure of our models to allow for the fact that different participants or items may vary with regard to how sensitive they are with respect to the variables at hand (Barr et al., 2013; Cunnings, 2012): if, e.g., vocabulary knowledge only matters for the understanding of some sentences in the perceptual learning task but not for others, the effect of vocabulary should be modeled individually for each sentence and removed from the fixed effect structure. Changes in the random-slope structure were evaluated by means of the Akaike information criterion (AIC). The model with the lower AIC value (difference ≥ 2) and, therefore, better model fit was retained. As we were interested in the predictors of individual amount of learning, a random participant slope of phase was included in all models of statistical learning. Accordingly, in models of perceptual learning, a random participant slope of block was inserted. That is, all models calculated the learning effect (i.e., the effect of phase in statistical learning and the effect of block in perceptual learning, respectively) individually for each participant.

After determining the maximal random slope structure, we first performed an age group comparison by testing the interactions between age group and the respective index of learning (i.e., phase or block). As younger and older adults may differ with respect to the effects of target position and target alignment on their learning performance, all possible two-way interactions between grammaticality, age group, target alignment and first target position and the three-way interactions between (1) age group, grammaticality and target position and between (2) age group, grammaticality and target alignment were included in the age group comparison of statistical learning.

In a second step, we assessed which cognitive abilities may facilitate

1 Note that we ran a second analysis in which we kept the original index of block. The analysis resulted in the same best models and showed the same effects as the models reported here. However, models that included the log-transformed index of block showed an increased model fit, indicating non-linear learning behavior.

learning within the separate age groups. In the statistical learning analysis, the best model that explained the facilitation score on basis of the interactions between phase, target position and target alignment was taken as initial model. In the perceptual learning analysis, the initial model only contained block. Then, measures of age (in older adults only), hearing sensitivity (in models of perceptual learning only), statistical learning performance (in models of perceptual learning only), attention switching control, working memory, processing speed and vocabulary (all evaluated as numerical covariates) and their interaction with phase (in models of statistical learning) or with block (in models of perceptual learning) were added simultaneously to the initial model. This method of forced entry was preferred, as we had no prior theoretical assumptions about the relative importance of each predictor and aimed to identify those predictors that had unique exploratory power in predicting facilitation scores. All individual predictor measures were centered around their mean prior to inclusion. After we had entered the individual predictor measures, we adopted a backward stepwise selection procedure, in which first interactions and then predictors were removed if they did not attain significance at the 5 % level. Each change in the fixed effect structure was evaluated in terms of loss of model fit by means of a likelihood ratio test. Results of the analysis are indicated in estimated absolute effect sizes (β), standard errors, t-values and p-values. Note however that the current version of the lme4 package does not report p-values for t-tests in models with a maximal random slope structure, as it is presently unclear how to calculate the appropriate number of degrees of freedoms (Baayen, 2012). Reported p-values were, therefore, derived by performing a likelihood ratio test between a model that included the specific fixed effect or interaction and a model that did not while all other model parameters were kept constant. That is, p-values actually reflect the significance of loss in model fit if the effect or interaction was excluded from the model.

Individual measure of statistical learning performance

As we wanted to assess whether individual statistical learning performance predicts adaptation to noise-vocoded speech, we needed an index of statistical learning ability for each participant. We derived this index by calculating the random participant slopes of phase (individual adjustments to the general slope) on the basis of the most parsimonious model, in which facilitation scores were predicted only by phase and the control variables (i.e., we derived the measure of statistical learning ability before we included effects of individual predictor measures in the above mentioned analysis).

Thus, we determined an individual value for each participant with which the general effect of phase (in the fixed structure of the model) had to be adjusted to resemble his/her individual learning effect. The lower the value, the more negative was a participant‘s slope when changing from the end of the exposure phase to the test phase, indicating a steeper drop in facilitation score and, hence, more statistical learning.

Results

Performance on background measures

Mean performance of younger and older adults and age group differences on all background measures are displayed in Table 4.1. As expected, hearing acuity was better in younger adults (i.e., thresholds were lower) than in older adults. Moreover, younger adults showed faster processing and larger memory capacity than older adults. On average, older adults were able to correctly repeat 5.62 test sequences in the working memory test, which corresponds to a mean digit span of four. Younger adults correctly repeated 8.08 test sequences, corresponding to a mean digit span of five. No difference could be observed in attention switching control between age groups. Older adults outperformed younger adults on the vocabulary test. However, older adults also showed relatively little variation on the vocabulary test (coefficient of variation [*SD*/*M*] = 6.9 %). Statistical testing confirmed that the variance in older adults‘ vocabulary scores was

Table 4.1. Mean performance per age group and age group differences on cognitive, linguistic and auditory measures.

	Younger adults		Older adults		Age group difference	
Measure	*M*	*SD*	*M*	*SD*	*t*	*p*
Working memory	67.37	17.18	46.80	18.72	6.57	<0.001
Processing speed	68.10	9.44	48.73	11.01	10.88	<0.001
Vocabulary	0.68	0.08	0.87	0.06	-15.80	<0.001
Attention	1.97	0.44	2.05	0.62	-0.96	0.340
Hearing (PTAH)	0.90	5.56	23.31	10.28	-15.96	<0.001

Note. t-tests tested two-tailed

significantly lower than the variability in younger adults‘ data (coefficient of variation = 11.8 %) (Levene‘s Test: $F = 4.15$, $df_1 = 1$, $df_2 = 130$, $p = .044$).

Intercorrelations between background measures within each age group are reported in Table 4.2. In younger adults, significant correlations were observed between the cognitive measures of working memory and processing speed and between working memory and vocabulary. The same intercorrelations were also observed in the older adults. As expected, age correlated with hearing sensitivity and with processing speed within the older sample: older-older participants generally had poorer hearing and slower processing than younger-older participants. Moreover, processing speed was related to hearing sensitivity in older adults. However, when both measures (i.e., speed and hearing) were controlled for age, this correlation was no longer significant ($r = -.128$, $p = 0.279$, $df = 71$).

Table 4.2. Pearson's correlation coefficients between measures of cognitive, linguistic and auditory functioning per age group.

	Younger adults (n = 59)				
Measure	Attention	Memory	Speed	Vocabulary	Hearing
Attention switching	1				
Working memory	-0.109	1			
Processing speed	0.079	0.301*	1		
Vocabulary	-0.071	0.311*	-0.156	1	
Hearing	-0.097	-0.233	0.022	-0.119	1

	Older adults (n = 73)				
Measure	Attention	Memory	Speed	Vocabulary	Hearing
Attention switching	1				
Working memory	-0.211	1			
Processing speed	-0.187	0.311**	1		
Vocabulary	-0.076	0.250*	0.207	1	
Hearing	0.104	-0.170	-0.336**	-0.113	1
Age	0.059	-0.194	-0.562**	0.017	0.426**

Note. *p < .05, **p < .01 (tested two-tailed)

4

Statistical learning

Valid facilitation scores were restricted to those within 2.5 sd from the mean facilitation score within each age group. Table 4.3 shows the average performance of younger and older adults on the statistical learning task in terms of response times and facilitation score. As expected, younger adults were significantly faster in responding to the first target ($t = -84.302$, $df = 23249.45$, $p < .001$) and to the second target ($t = -104.343$, $df = 23585.75$, $p < .001$) than older adults. Note that all responses in the statistical learning task were accurate as the experimental task only proceeded when a participant had clicked on the correct shape. Figure 4.3A shows the average facilitation scores for both age groups over block[2]. Figure 4.3B displays the mean facilitation scores at the end of the exposure phase, in the test phase and in the recovery phase to illustrate the learning effect. Moreover, the range of statistical learning that was observed within each age group is displayed in Figure 4.3C. Estimates of the best model within each age group are displayed in Table 4.4A for younger adults and in Table 4.4B for older adults.

The age group comparison showed a significant effect of phase (beta = -0.137, SE = 0.045, $t = -3.03$, $p = .002$), indicating statistical learning in the group of younger adults, who were placed on the intercept. This effect of phase was modified by age group (beta = 0.125, SE = 0.061, $t = 2.06$, $p = .039$), suggesting that older adults learned less than younger adults and (given the almost equal beta values) that older adults were not affected by removal of the underlying regularities. This interaction between age group and phase tended to be less pronounced in diagonal trials (beta = -0.070, SE = 0.036, $t = -1.91$,

2 Note that the drop in performance that can be observed in the younger adults during the exposure phase (see Figure 3A, blocks 5 and 6) is not significant (beta = -0.018, SE = 0.010, $t = -1.71$, $p = 0.088$). This suggests that there was no general drop in performance across the group of younger adults. Likewise, the spread of the individual slope data (M = -0.018, SD = 0.019, Min = -0.078, Max = 0.027) also includes positive slope values (indicating improvement, rather than decreased performance). Moreover, a paired samples t-test shows that the size of the unexpected drop in the exposure phase is significantly smaller than the drop in the test phase that is considered to reflect learning ($t = 17.914$, $df = 58$, $p < .001$).

Table 4.3. Mean response times (in ms) and facilitation scores of younger adults (n = 59) and older adults (n = 73) on the statistical learning task.

	1st target RT		2nd target RT		Facilitation score	
	M	*SD*	*M*	*SD*	*M*	*SD*
Younger adults	499	130	473	129	1.103	0.328
Older adults	705	243	728	240	1.024	0.359

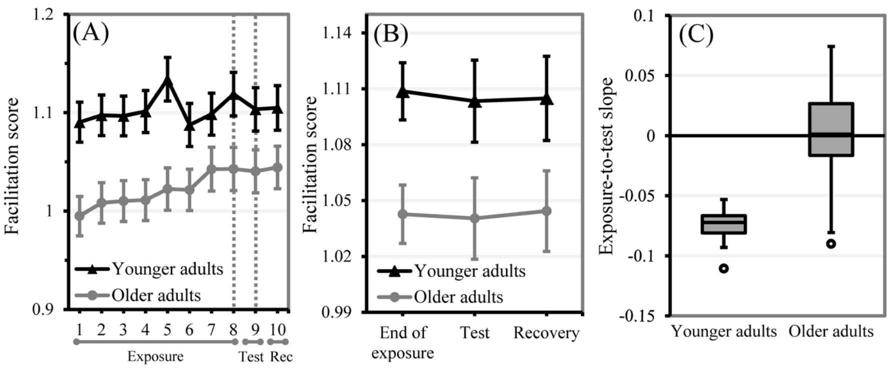

Figure 4.3. Performance on the statistical learning task. A drop in facilitation score from the end of the exposure phase (blocks 7 – 8) to the test phase (block 9) indicates learning. Error bars indicate two standard errors from the mean. (A) Mean statistical learning performance per age group and block. The area between the dotted lines represents where the effect of removing the underlying regularities should be observed. (B) Mean statistical learning performance per age group and phase. (C) Boxplot of statistical learning performance in younger and older adults (individual exposure-to-test slopes from the statistical model). More negative slopes reflect more learning.

p = .056). A fixed effect of age group indicated that, overall, older adults showed a lower facilitation score than younger adults (beta = -0.189, SE = 0.044, t = -4.28, p < .001). This effect of age group was influenced by both control variables. That is, the difference in facilitation score between younger and older adults was less distinct in diagonal (beta = 0.065, SE = 0.025, t = 2.54, p = .011) and in upper left trials (beta = 0.060, SE = 0.027, t = 2.25, p = .025). As expected, facilitation scores were higher in diagonal trials (beta = 0.120, SE = 0.019, t = 6.33, p < .001) and lower in trials, in which the first target appeared upper left (beta = -0.100, SE = 0.020, t = -4.98, p < .001). Moreover, the effect of phase was modified by both target position and by target alignment, implying that effects of statistical learning were less pronounced in diagonal trials (beta = 0.062, SE = 0.027, t = 2.29, p = .022) and in trials with an upper left target (beta = 0.064, SE = 0.027, t = 2.38, p =.017). The random slope structure indicated that participants differed in the degree to which they were affected by target position and by target alignment.

In the younger adults, the best-fitting model showed a significant effect of phase: the facilitation score of younger adults was lower in the test phase than at the end of the exposure phase, indicating that younger adults were affected by removing the underlying regularities. However, none of the individual listener characteristics interacted significantly with test phase, suggesting that amount of statistical learning was not associated with any of the selected measures of cognitive or linguistics abilities. Only processing speed showed a significant fixed effect on facilitation score, indicating that participants with higher processing speed had higher facilitation scores at the end of the exposure phase. As expected, facilitation scores were lower if the first target was displayed upper left and higher if targets were aligned diagonally. Both effects modulated learning in the anticipated direction: the effect of statistical learning was smaller in diagonal trials and in trials in which the first target was displayed upper left. In addition to the random slope of phase, the maximal random slope structure included random effects of first target position and target alignment on participant. Inclusion of these effects suggests

that younger participants differed in the degree to which they were affected by target alignment, that is, whether they had to move the cursor horizontally or diagonally. Removing the random slope of phase within subject from the maximal random slope structure did not result in a significant loss in model fit, indicating that the amount of statistical learning did not differ considerably among younger adults.
Overall, older adults showed no significant effect of test phase, suggesting that they generally did not pick up the subtle regularities in the input. Age was the only individual background measure that predicted performance: the older the participants were, the lower was their facilitation score at the end of the exposure phase. In older adults, facilitation score was mainly influenced by the control variables. That is, diagonal alignment of targets enhanced facilitation scores and upper left position of the first target decreased facilitation benefit. A significant interaction between target position and target alignment indicated that effects of one control variable were modified by the other control variable: the effect of the first target being located upper left was smaller when participants could make a diagonal mouse movement to the second target, respectively, the benefit in facilitation score based on a diagonal movement was decreased in case the first target was displayed in the upper left corner of the screen. The maximal random slope structure showed that older adults differed in the degree to which they were affected by changes in target position (random slope of first target position within subject) and target alignment (random slope of first target position within subject). However, in modeling the statistical learning data of the older adults, we had kept in a random slope of phase to allow that participants may vary in how much their performance was affected by removing the regularities (we also needed this random slope parameter as the individual measure of statistical learning). Importantly, inclusion of this random effect of phase did not increase the model fit, implying that older participants did not differ much in their sensitivity to statistical regularities: they were all relatively insensitive to the probabilistic sequence information. Note that older adults continued to show increased facilitation throughout the exposure phase (cf.,

Table 4.4A. Statistical model for the facilitation score of younger adults in the statistical learning task.

Fixed effects		β	*SE*	*t*	*p*
Intercept		1.098	0.015	73.68	< 0.001
Test phase		-0.069	0.022	-3.09	0.002
Target upper left		-0.099	0.015	-6.78	< 0.001
Diagonal alignment		0.122	0.020	6.23	< 0.001
Processing speed		0.002	0.001	2.13	0.034
Test phase x target upper left		0.066	0.025	2.63	0.009
Test phase x diagonal alignment		0.059	0.025	2.34	0.020
Random effects		*Varian.*	*SD*	*Corr*	*Corr*
Subject	Intercept	0.003	0.055		
	Test phase	0.001	0.024	-0.738	
	Diagonal	0.010	0.101	-0.150	0.778
Residual		0.096	0.310		

Table 4.4B. Statistical model for the facilitation score of older adults in the statistical learning task.

Fixed effects	β	*SE*	*t*	*p*
Intercept	1.004	0.015	65.55	< 0.001
Target upper left	-0.102	0.020	-4.98	< 0.001
Diagonal alignment	0.113	0.020	5.77	< 0.001
Age	-0.004	0.002	-2.36	0.020
Target upper left x diagonal	0.133	0.024	5.48	< 0.001

Random effects		*Varian.*	*SD*	*Corr*	*Corr*
Subject	Intercept	0.003	0.057		
	Test phase	0.003	0.058	-0.187	
	Diagonal	0.007	0.084	-0.300	-0.544
	Lower left	0.003	0.058		
	Upper left	0.005	0.069	-0.153	
Residual		0.120	0.346		

4

Figure 4.3A). As their performance was unaffected by the removal of the underlying regularities in the test phase, this suggests that the improvement over block in older adults reflects effects of task learning rather than effects of statistical learning.

Perceptual learning

As we wanted to include statistical learning performance as a predictor in the analyses of the perceptual learning data alongside the auditory and cognitive measures, we checked for intercorrelations between statistical learning ability and other individual background measures. In the older adults, no correlations were observed. In the younger adults, intercorrelations between statistical learning performance and both working memory ($r = -.263$, $p = .044$; $rho = -.297$, $df = 57$, $p = .022$) and information processing speed ($r = -.279$, $p = .032$; $rho = -.223$, $df = 57$, $p = .089$) were significant: more learning was associated with better working memory and with higher processing speed.

In Figure 4.4, the average recognition score per block is displayed to illustrate perceptual learning of the noise-vocoded speech within age group. Moreover, Figure 4.4 shows the range of perceptual learning that could be observed within each age group. Although younger adults were presented with a more difficult noise-vocoding condition (4 bands) than older adults (5 bands) and showed a lower starting performance, both age groups showed similar progress in perceptual learning. This indicates that speech conditions were appropriately selected to elicit sizeable and comparable amounts of improvement over the course of exposure in both age groups. Estimates of the best model to predict sentence identification performance within each age group are displayed in Table 4.5A for younger adults and in Table 4.5B for older adults.

The age group comparison showed a significant effect of block ($beta = 0.710$, $SE = 0.034$, $t = 20.780$, $p < .001$) that was not modified by age group ($beta = -0.071$, $SE = 0.046$, $t = -1.555$, $p = .120$), indicating that both age groups showed a similar amount of perceptual learning

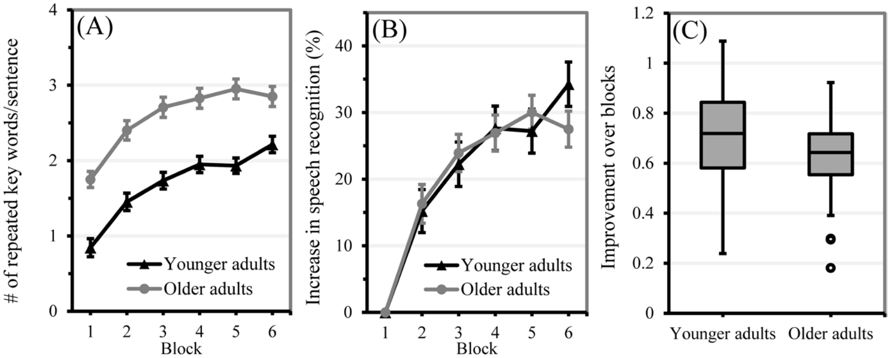

Figure 4.4. Performance on the perceptual learning task. Error bars indicate two standard errors from the mean. (A) Mean improvement in speech understanding per age group over block. (B) Improvement in speech understanding performance (in %) relative to baseline level. (C) Box plot of perceptual learning performance in younger and older adults (individual block slopes from the statistical model). More positive slopes reflect more learning.

over the course of the experiment[3]. As older adults were presented with an easier condition (5 instead of 4 band vocoded speech), a fixed effect of age group showed that older adults repeated more key words correctly than younger adults (beta = 0.971, SE = 0.103, t = 9.414, $p < .001$). Our results suggest that we were successful in providing older and younger adults with a starting level that allowed for a comparable amount of perceptual learning within both age groups.

In younger adults, none of the predictor measures showed a fixed effect, suggesting that none of the predictor measures could be used

3 Note that we also performed an age group comparison including the data of the pilot study. In this analysis, we compared younger adults' performance on the 4 bands speech (placed on the intercept) to older adults' performance across the different band conditions (i.e., 4 bands, 5 bands and 6 bands). In the 4 band condition, younger adults showed a higher starting performance than older adults (beta = -0.64, SE = 0.16, t = -3.991, $p < .0001$). More importantly, in the 4 band condition, we found an interaction between age group and improvement over blocks. That is, older adults showed a significantly smaller effect of block and, hence, less learning, than younger adults (beta = -0.10, SE = 0.03, t = -3.795, $p < .0001$). This result emphasizes that it is not possible to elicit similar amounts of perceptual learning in the two age groups by presenting younger and older adults with the same signal degradation condition.

to predict initial speech recognition performance (i.e., for block 1 performance, being on the intercept). The best fitting model showed that younger participants identified more keywords correctly with increasing exposure over blocks, indicating that they generally adapted to noise-vocoded speech. This effect of perceptual learning was modified by statistical learning ability: the more participants had picked up the implicit regularities in the statistical learning task (and thus the steeper their drop in performance in the test phase), the more they improved in understanding noise-vocoded speech. This result provides first evidence that perceptual learning and statistical learning are associated. Further, the effect of perceptual learning was modified by vocabulary knowledge: younger adults who had greater vocabulary knowledge showed faster speech adaptation over blocks, underscoring the involvement of linguistic knowledge in perceptual learning of speech. Note that we had excluded an interaction between block and processing speed during the modeling process, as its inclusion led only to a marginal improvement of model fit. This marginal interaction suggested that higher processing speed tended to be associated with faster adaptation. The maximal random slope structure included the effect of block within subject. Removing this effect from the maximal random slope structure reduced the model fit significantly, indicating that individuals differed considerably in perceptual learning ability. As random slopes of individual predictor measures within items did not improve the model fit, this indicated that the effects of predictor measures could be generalized across sentences.

In the older adults, initial sentence identification performance was associated with hearing sensitivity and processing speed: hearing loss considerably affected initial speech understanding, whereas those with higher processing speed showed better initial speech recognition performance. Like the younger adults, older adults showed perceptual learning of noise-vocoded speech, which was indicated by a significant improvement in identification performance over blocks. This block effect was modified by age, indicating that older adults within the older age group improved less over the course of exposure than younger older adults. As age in the older adult sample was intercorrelated with processing speed and hearing sensitivity (see intercorrelations in Table 4.2), we also investigated whether either variable would have surfaced as a predictor for adaptation if we left out age. The variance in amount of perceptual learning that was assigned to age was not taken over by any of the other predictors included in the analysis. This indicates that the effect of age explains unique variance in perceptual learning performance that is not captured by the included cognitive and perceptual predictors. Importantly, statistical learning ability did not facilitate the amount of improvement over the course of the experiment. The maximal random slope structure included effects of age and hearing sensitivity on item, suggesting that the effects of age and hearing sensitivity on recall of noise-vocoded sentences differed across sentences. That is, hearing and age affected speech understanding of some sentences more than of others, in addition to the general impact these predictors had on sentence recall. Moreover, inclusion of the random effect of block within participant significantly improved the model fit, implying that older participants differed in their improvement to understand noise-vocoded speech over the course of exposure.

Table 4.5A. Statistical model for sentence identification performance of younger adults in the perceptual learning task.

Fixed effects	*β*	*SE*	*t*	*p*
Intercept	0.917	0.155	5.90	< 0.001
Block	0.703	0.043	16.41	< 0.001
Statistical learning	6.649	8.966	0.74	*n.s*
Vocabulary	-0.921	1.255	-0.73	*n.s.*
Block x statistical learning	-9.223	4.058	-2.27	0.023
Block x vocabulary	1.489	0.568	2.62	0.009

Random effects		*Variance*	*SD*	*Corr*
Subject	Intercept	0.436	0.660	
	Block	0.049	0.221	-0.821
Item	Intercept	0.912	0.955	
Residual		1.284	1.133	

Note. *n.s.* = p > 0.05

Table 4.5B. Statistical model for sentence identification performance of older adults in the perceptual learning task.

Fixed effects		*β*	*SE*	*t*	*p*
Intercept		1.897	0.122	15.60	< 0.001
Block		0.624	0.037	16.73	< 0.001
Hearing		-0.029	0.005	-6.23	< 0.001
Processing speed		0.011	0.004	2.46	0.017
Age		0.004	0.011	0.38	*n.s.*
Block x Age		-0.017	0.007	-2.53	0.012
Random effects			*Variance*	*SD*	*Corr*
Subject	Intercept		0.122	0.350	
	Block		0.040	0.201	-0.509
Item	Intercept		2.770	1.664	
	Hearing		0.000	0.015	
	Age		0.001	0.022	-0.887
Residual			1.307	1.143	

Discussion

This study investigated the contribution of general cognitive abilities to listeners' capacity to adapt to novel speech conditions. In order to gain more insight into individual abilities associated with adaptation to unfamiliar speech across the life span, we tested both younger and older adults. Specifically, we aimed to test the hypothesis that listeners' improvement in understanding unfamiliar types of speech could be predicted from individual differences in statistical learning ability and in general cognitive skills.

The ability to implicitly learn has been argued to remain stable over the life span (Midford & Kirsner, 2005). In line with this, several studies reported that older adults are sensitive to probabilistic sequences (Campbell, Zimerman, Healey, Lee, & Hasher, 2012; Negash, Howard, Japikse, & Howard, 2003; Salthouse et al., 1999; Simon, Howard, & Howard, 2011) and found the ability to adapt to novel speech conditions to be preserved in older adults (Adank & Janse, 2010; Golomb et al., 2007; Gordon-Salant et al., 2010; Peelle & Wingfield, 2005). Our findings support the notion that perceptual learning ability remains stable over the life span, as both younger and older listeners showed significant improvement in understanding noise-vocoded speech over exposure. Moreover, the observed amount of learning was comparable in both age groups. This suggests that older adults can reach the same amount of perceptual learning as younger adults given better starting level intelligibility. However, only younger adults were sensitive to statistical regularities in the input. As we found a significant learning by age group interaction, this indicated age-related declines in the ability to detect statistical regularities if visual sequences are presented quickly.

Possibly, certain aspects of our statistical learning task may be responsible for the absence of a statistical learning effect in older adults. In particular, we had incorporated an inter-target interval of 500 ms (following Salthouse et al., 1999) between both clicks within a trial to allow for prediction effects, even in older adults with slower processing. As we tested statistical learning in a speeded computer mouse task, and movement control on computer mouse tasks is reduced in older adults (Smith, Sharit, & Czaja, 1999), the implemented inter-target interval may have been too short for older adults to show prediction effects. Moreover, to prevent associations between both measures of implicit learning due to modality-specific processing, we chose for a rigorous test of the association between the two types of learning by testing statistical learning ability in a non-auditory (i.e., visual) domain with non-linguistic stimuli. As older adults were able to implicitly learn in the auditory task, it may be argued that we did not observe implicit learning in the visual paradigm due to age-specific modality effects. In both implicit learning tasks, task-relevant information was presented sequentially (i.e., speech unfolding over time in the auditory task and successive highlighting of targets in the visual task). Visual stimuli have been shown to have less salient temporal relations than auditory stimuli (Kubovy, 1988). Consequently, auditory learning is superior to visual learning in sequence learning tasks (Conway & Christiansen, 2005). Additionally, a recent study found that statistical learning performance is decreased if visual stimuli are presented at a fast rate (Emberson, Conway, & Christiansen, 2011). Although stimuli presentation in our statistical learning task was not timed as it depended on participants' performance speed (i.e., participants who clicked faster, saw visual stimuli shorter), the time pressure induced by the speeded task, as well as relatively fast and sequential presentation of visual stimuli, may have interfered with statistical learning performance in older adults. That is, results of the current study suggest that older adults' statistical learning ability is affected if fast, sequential processing of visual stimuli is required. However, as previous studies have shown that older adults remain sensitive to probabilistic information in the input (Campbell et al., 2012; Negash et al., 2003; Salthouse et al., 1999; Simon et al., 2011), our failure to observe statistical learning in older adults should not be taken as evidence that older adults are generally insensitive to probabilistic information in the input, or that probabilistic information in the input is generally unimportant for perceptual learning in older adults. Obviously, further research is required to investigate possible links between statistical and perceptual learning in a setting where older adults do show both types of learning.

Overall, limited variability could be observed on the measure of statistical learning ability in both age groups and the amount of individual statistical learning could not be explained by individual differences in cognitive or linguistic abilities in our analyses. However, note the correlations between statistical learning on the one hand and speed and working memory on the other hand in the younger adults. These correlations suggest that, despite relatively little variation in statistical learning, there was some systematicity in younger adults‘ statistical learning differences. In contrast, participants showed great variability in the amount of adaptation to degraded speech and individual differences in learning to understand noise-vocoded speech could be associated with listeners‘ cognitive abilities. This finding supports the claim of the RHT that perceptual learning is a top-down guided process, implying that higher cognitive processes are indeed involved in the top-down search to identify task-relevant cues in the input. However, links between cognitive abilities and perceptual learning performance seem to undergo age-related changes, as different associations between perceptual learning ability and cognitive measures emerged in younger and older adults.

In younger adults, initial performance in identifying noise-vocoded speech was not predicted by general cognitive or linguistic abilities. However, differences in the amount of improvement over the course of exposure were associated with individual sensitivity to probabilistic information and with individual vocabulary knowledge. In line with our hypothesis, our results suggest that adaptation to novel speech conditions and statistical learning share mechanisms of implicit regularity detection. Our results contribute to earlier literature indicating a relationship between statistical learning performance and individual differences in language processing (Misyak et al., 2010a). As statistical learning was tested using visual and non-linguistic stimuli, this suggests that general abilities, that are neither modality-specific nor specific for language processing, drive this association.

As argued in the Introduction, the link between statistical learning and perceptual learning in speech can be twofold. On the one hand, statistical and perceptual learning may be associated as they draw on the same underlying abilities. Our findings do not support this “mediation account”: the observed association between perceptual learning in speech and statistical learning performance does not seem to be mediated by the specific cognitive abilities tested in the current study. On the other hand, perceptual learning processes may directly rely on statistical properties in the input. In novel speech conditions, perceptual learning may be facilitated by sensitivity to statistical properties as language itself conveys probabilistic information e.g., in terms of phonotactic (Vitevitch et al., 2004) and transitional probability (Thompson & Newport, 2007). Listeners have been shown to make use of this probabilistic information to segment speech streams into words (Saffran, Newport, Aslin, Tunick, & Barrueco, 1997). In the framework of the RHT, listeners who are more sensitive to statistical regularities may, hence, be faster in identifying subunits (e.g., words) in novel speech input, thereby facilitating faster access to high-level representations. Moreover, the information that is transferred from lower to higher levels of the hierarchy may itself be probabilistic in nature. Recent theories in visual perceptual learning argue that the process of input reweighting is based on such probabilistic decisions (e.g., Petrov et al., 2005; J. Zhang et al., 2010). For example, assuming that the information that is conveyed from lower levels to higher levels is normally distributed and that the mean of the distribution resembles the most relevant input, each incoming input could be reweighted based on its relative distance from the mean, with distance serving as index of informational relevance (J. Zhang et al., 2010). First evidence that probabilistic information may be encoded in the input from lower to higher hierarchical levels comes from a study in which neuronal network models that relied on probabilistic inferences could explain neurophysiological changes in early sensory areas in visual perceptual learning tasks that could not be accounted for by other models (Bejjanki, Beck, Lu, & Pouget, 2011).

The finding that improvement in understanding noise-vocoded speech in younger adults is predicted by participants‘ vocabulary size confirms the link between increased lexical knowledge and success in perceptual learning that has previously been reported in adapting to novel-accented speech (Janse & Adank, 2012). Thus, lexical knowledge is not only associated with adaptation to linguistic degradations, e.g., systematic phonological deviations in how a foreign-accented speaker pronounces words, but also relates to perceptual learning of acoustically degraded speech. Previous research has shown that younger and older individuals with higher scores on vocabulary tests also show better performance on measures of verbal fluency (e.g., Hedden, Lautenschlager, & Park, 2005; Kemper & Sumner, 2001). Thus, individuals with greater vocabulary knowledge may be more efficient processors of linguistic information (Kemper & Sumner, 2001), and linguistic knowledge may improve perceptual learning in speech by facilitating access to higher-level representations. As access to higher-level representations aids sublexical retuning by enabling and guiding top-down search processes (Ahissar & Hochstein, 2004), effects of lexical knowledge should in fact arise irrespective of type of systematic speech degradation. Given that Janse and Adank (2012) observed a relationship between vocabulary knowledge and adaptation (to accented speech) in older adults, this raises the question why older adults‘ perceptual learning performance was not predicted by their linguistic knowledge here. Older adults outperformed younger adults on the measure of lexical knowledge, but note that older adults also showed relatively little variation on the vocabulary test (see the Results section on participants’ performace on background measures). Consequently, there was less room to relate lexical knowledge to individual differences in perceptual learning ability in older adults than in younger adults. We checked the variation for older adults‘ vocabulary scores in the sample of Janse and Adank (2012) (coefficient of variation = 10.3%), which was close to the variation we now observed in the younger adults. Therefore, variation in older adults‘ vocabulary scores in the current study may have indeed been insufficient to predict perceptual learning.

In the older adult group, listeners‘ starting level in understanding noise-vocoded speech was associated with higher processing speed and affected by hearing loss, whereas listeners' age predicted how well they adapted to the novel speech condition. That is, younger listeners in the group of older adults showed more learning than older-older listeners. This effect of age had unique explanatory power that was not captured by the included cognitive and perceptual predictors. This finding seems to be consistent with previous research which reported declines in the general identification of noise-vocoded speech with increasing age that were independent of hearing sensitivity (Sheldon et al., 2008; Souza & Boike, 2006) and which may have reflected limited improvement over exposure. Importantly, the current design allowed us to differentiate between effects of individual predictors on both starting level speech identification performance and on amount of perceptual learning. Thus, our results complement earlier findings, suggesting that hearing loss affects initial recognition of noise-vocoded speech, whereas age-related deficits specifically constrain improvement in adaptation to a novel speech input. Younger adults generally outperform older adults when being exposed to the same speech degradation (Peelle & Wingfield, 2005; Sheldon et al., 2008). Importantly, providing younger and older adults with the same speech degradation also has consequences for age groups‘ ability to improve their performance over exposure (cf. our pilot result data discussed in the Results section on perceptual learning performance). In order to have similarly large amounts of perceptual learning for the two age groups, older adults have to be presented with an easier condition than younger adults (Golomb et al., 2007), which was also done in the current study. It is unclear, however, what the age effect on perceptual learning ability among the older adults reflects. A possible account may come from recent studies reporting that coherence between activated brain regions is decreased in older adults (Andrews-Hanna et al., 2007; Peelle, Troiani, Wingfield, & Grossman, 2010), relative to younger adults. Importantly, these deteriorations in connectivity were associated with declines in speech understanding performance under difficult listening conditions (Peelle et al., 2010) and with poorer performance on cognitive tasks (Andrews-Hanna et al., 2007). In the framework of the RHT, we may speculate that a reduced coordination between neuronal regions may hinder effective information flow between hierarchical levels, thereby constraining processes of input reweighting. Consequently, this decreased information flow would then impede modifications to the lower-level representations. Thus, an age-related decrease in the ability to coordinate activity between brain regions may affect adaptation to challenging novel speech input. In short, our results suggest that individual differences in general cognitive and linguistic abilities can explain listeners‘ variability in adaptation to noise-vocoded speech, thereby highlighting the involvement of listener-based abilities in perceptual learning. As noise-vocoded speech simulates the auditory signal of a cochlear implant, findings of the current study may provide valuable insights for aural rehabilitation in younger and older adults. Amount of adaptation over the course of exposure was specifically associated with vocabulary knowledge and with individuals‘ sensitivity to probabilistic regularities. These combined results emphasize the importance of pattern recognition and linguistic knowledge for perceptual learning and adaptation in speech processing.

References

Adank, P., & Janse, E. (2010). Comprehension of a novel accent by young and older listeners. *Psychology and Aging, 25*, 736-740. doi: 10.1037/a0020054

Ahissar, M., & Hochstein, S. (1997). Task difficulty and the specificity of perceptual learning. *Nature, 387*, 401–406. doi: 10.1038/387401a0

Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences, 8*, 457-464. doi: 10.1016/j.tics.2004.08.011

Akeroyd, M. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *International Journal of Audiology, 47*, 53-71. doi: 10.1080/14992020802301142

Amitay, S., Irwin, A., & Moore, D. R. (2006). Discrimination learning induced by training with identical stimuli. *Nature Neuroscience, 9*, 1446-1448. doi: 10.1038/nn1787

Andrews-Hanna, J. R., Snyder, A. Z., Vincent, J. L., Lustig, C., Head, D., Raichle, M. E., & Buckner, R. L. (2007). Disruption of large-scale brain systems in advanced aging. *Neuron, 56*(5), 924-935. doi: 10.1016/j.neuron.2007.10.038

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. *Language Learning, 62*, 49-78. doi: 10.1111/j.1467-9922.2012.00706.x

Arbuthnott, K., & Frank, J. (2000). Trail making test, part B as a measure of executive control: validation using a set-switching paradigm. *Journal of Clinical and Experimental Neuropsychology, 22*, 518-528. doi: 10.1076/1380-3395(200008)22:4;1-0;FT518

Auer, E., & Luce, P. A. (2005). Probabilistic phonotactics in spoken word recognition. In D. B. Pisoni & R. Remez (Eds.), *The Handbook of Speech Perception* (pp. 610-630). Oxford: Blackwell.

Baayen, R. H. (2012). *Analyzing Linguistic Data: A Practical Introduction to Statistics using R* (6th ed.). Cambridge: Cambridge University Press.

Banai, K., & Ahissar, M. (2009). Perceptual learning as a tool for boosting working memory among individuals with reading and learning disabilities. *Learning & Perception, 1*, 115-134. doi: 10.1556/LP.1.2009.1.9

Banai, K., & Amitay, S. (2012). Stimulus uncertainty in auditory perceptual learning. *Vision Research, 61*, 83-88. doi: 10.1016/j.visres.2012.01.009

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: keep it maximal. *Journal of Memory and Language, 68*, 255-278. doi: 10.1016/j.jml.2012.11.001

Bates, D., Maechler, M., & Bolker, B. (2012). lme4: Linear mixed-effects models using S4 classes (Version 0.999-0). Retrieved from http://CRAN.R-project.org/package=lme4

Bejjanki, V. R., Beck, J. M., Lu, Z., & Pouget, A. (2011). Perceptual learning as improved probabilistic inference in early sensory areas. *Nature Neuroscience, 14*, 642-648. doi: 10.1038/nn.2796

Bo, J., Jennett, S., & Seidler, R. D. (2011). Working memory capacity correlates with implicit serial reaction time task performance. *Experimental Brain Research, 214*, 73-81. doi: 10.1007/s00221-011-2807-8

Bo, J., Jennett, S., & Seidler, R. D. (2012). Differential working memory correlates for implicit sequence performance in young and older adults. *Experimental Brain Research, 221*, 467-477. doi: 10.1007/s00221-012-3189-2

Boersma, P., & Weenink, D. (2011). Praat: doing phonetics by computer (Version 5.1). Retrieved from http://www.praat.org/

Campbell, K. L., Zimerman, S., Healey, M. K., Lee, M. M. S., & Hasher, L. (2012). Age differences in visual statistical learning. *Psychology and Aging, 27*, 650-656. doi: 10.1037/a0026780

Christiansen, M. H., Kelly, M. L., Shillcock, R. C., & Greenfield, K. (2010). Impaired artificial grammar learning in agrammatism. *Cognition, 116*(3), 382-393. doi: DOI 10.1016/j.cognition.2010.05.015

Cohen, Y., Daikhin, L., & Ahissar, M. (2013). Perceptual learning is specific to the trained structure of information. *Journal of Cognitive Neuroscience, 25*, 2047-2060. doi: 10.1162/jocn_a_00453

Conway, C. M., & Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 31*, 24-39. doi: 10.1037/0278-7393.31.1.24

Cunnings, I. (2012). An overview of mixed-effects statistical models for second language researchers. *Second Language Research, 28*, 369-382. doi: 10.1177/0267658312443651

Davis, M. H., Johnsrude, I. S., Hervais-Adelman, A., Taylor, K., & McGettigan, C. (2005). Lexical information drives perceptual learning of distorted speech: evidence from the comprehension of noise-vocoded sentences. *Journal of Experimental Psychology. General, 134*, 222-241. doi: 10.1037/0096-3445.134.2.222

Dosher, B. A., Han, S., & Lu, Z. (2010). Perceptual learning and attention: Reduction of object attention limitations with practice. *Vision Research, 50*, 402-415. doi: 10.1016/j.visres.2009.09.010

Dosher, B. A., & Lu, Z. L. (1999). Mechanisms of perceptual learning. *Vision Research, 39*(19), 3197-3221. doi: 10.1016/S0042-6989(99)00059-0

Eisner, F., McGettigan, C., Faulkner, A., Rosen, S., & Scott, S. K. (2010). Inferior frontal gyrus activation predicts individual differences in perceptual learning of cochlear-implant simulations. *The Journal of Neuroscience, 30*, 7179-7186. doi: 10.1523/JNEUROSCI.4040-09.2010

Emberson, L. L., Conway, C. M., & Christiansen, M. H. (2011). Timing is everything: changes in presentation rate have opposite effects on auditory and visual implicit statistical learning. *Quarterly Journal of Experimental Psychology, 64*, 1021-1040. doi: 10.1080/17470218.2010.538972

Erb, J., Henry, M. J., Eisner, F., & Obleser, J. (2012). Auditory skills and brain morphology predict individual differences in adaptation to degraded speech. *Neuropsychologia, 50*, 2154-2164. doi: 10.1016/j.neuropsychologia.2012.05.013

Evans, J. L., Saffran, J. R., & Robe-Torres, K. (2009). Statistical learning in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 52*, 321-335. doi: 10.1044/1092-4388(2009/07-0189)

Fahle, M. (2006). Perceptual Learning. In L. Nadel (Ed.), *Encyclopedia of Cognitive Science* (pp. 1-5). Hoboken: John Wiley & Sons, Ltd.

Field, A., Miles, J., & Field, Z. (2012). *Discovering Statistics Using R*. London: SAGE Publications Ltd.

Gathercole, S. E. (1999). Cognitive approaches to the development of short-term memory. *Trends in Cognitive Sciences, 3*(11), 410-419. doi: 10.1016/S1364-6613(99)01388-1

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology, 49*, 585-612. doi: 10.1146/annurev.psych.49.1.585

Golomb, J. D., Peelle, J. E., & Wingfield, A. (2007). Effects of stimulus variability and adult aging on adaptation to time-compressed speech. *The Journal of the Acoustical Society of America, 121*, 1701. doi: 10.1121/1.2436635

Gordon-Salant, S., Yeni-Komshian, G. H., Fitzgibbons, P. J., & Schurman, J. (2010). Short-term adaptation to accented English by younger and older adults. *The Journal of the Acoustical Society of America, 128*, 200-204. doi: 10.1121/1.3486199

Greenwood, D. D. (1990). A cochlear frequency-position function for several species - 29 years later. *The Journal of the Acoustical Society of America, 87*(6), 2592-2605. doi: 10.1121/1.399052

Grunow, H., Spaulding, T. J., Gómez, R. L., & Plante, E. (2006). The effects of variation on learning word order rules by adults with and without language-based learning disabilities. *Journal of Communication Disorders, 39*, 158-170. doi: 10.1016/j.jcomdis.2005.11.004

Hedden, T., Lautenschlager, G., & Park, D. C. (2005). Contributions of processing ability and knowledge to verbal memory tasks across the adult life-span. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology, 58*(1), 169-190. doi: 10.1080/02724980443000179

Hervais-Adelman, A. G., Davis, M. H., Johnsrude, I. S., & Carlyon, R. P. (2008). Perceptual learning of noise vocoded words: effects of feedback and lexicality. *Journal of Experimental Psychology. Human Perception and Performance, 34*, 460-474. doi: 10.1037/0096-1523.34.2.460

Huyck, J. J., & Johnsrude, I. S. (2012). Rapid perceptual learning of noise-vocoded speech requires attention. *The Journal of the Acoustical Society of America, 131*, 236-242. doi: 10.1121/1.3685511

Janacsek, K., & Nemeth, D. (2013). Implicit sequence learning and working memory: correlated or complicated? *Cortex, 49*, 2001-2006. doi: 10.1016/j.cortex.2013.02.012

Janse, E., & Adank, P. (2012). Predicting foreign-accent adaptation in older adults. *Quarterly Journal of Experimental Psychology, 65*, 1563-1585. doi: 10.1080/17470218.2012.658822

Kaufman, S. B., De Young, C. G., Gray, J. R., Jimenez, L., Brown, J., & Mackintosh, N. (2010). Implicit learning as an ability. *Cognition, 116*(3), 321-340. doi: 10.1016/j.cognition.2010.05.011

Kemper, S., & Sumner, A. (2001). The structure of verbal abilities in young and older adults. *Psychology and Aging, 16*(2), 312-322. doi: 10.1037//0882-7974.16.2.312

Kubovy, M. (1988). Should we resist the seductiveness of the space-time, vision-audition analogy. *Journal of Experimental Psychology. Human Perception and Performance, 14*(2), 318-320. doi: 10.1037//0096-1523.14.2.318

Lin, F. R., Thorpe, R., Gordon-Salant, S., & Ferrucci, L. (2011). Hearing loss prevalence and risk factors among older adults in the United States. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences, 66*, 582-590. doi: 10.1093/gerona/glr002

Liu, E. H., Mercado, E., Church, B. A., & Orduna, I. (2008). The easy-to-hard effect in human (Homo sapiens) and rat (Rattus norvegicus) auditory identification. *Journal of Comparative Psychology, 122*(2), 132-145. doi: 10.1037/0735-7036.122.2.132

Loebach, J. L., Bent, T., & Pisoni, D. B. (2008). Multiple routes to the perceptual learning of speech. *The Journal of the Acoustical Society of America, 124*, 552-561. doi: 10.1121/1.2931948

Loebach, J. L., Pisoni, D. B., & Svirsky, M. A. (2009). Transfer of auditory perceptual learning with spectrally reduced speech to speech and nonspeech tasks: implications for cochlear implants. *Ear and Hearing, 30*, 662-674. doi: 10.1097/AUD.0b013e3181b9c92d

Midford, R., & Kirsner, K. (2005). Implicit and explicit learning in aged and young adults. *Aging, Neuropsychology, and Cognition, 12*, 359-387. doi: 10.1080/13825580500246894

Misyak, J. B., & Christiansen, M. H. (2012). Statistical learning and language: an individual differences study. *Language Learning, 62*, 302-331. doi: 10.1111/j.1467-9922.2010.00626.x

Misyak, J. B., Christiansen, M. H., & Tomblin, J. B. (2010a). On-line individual differences in statistical learning predict language processing. *Frontiers in Psychology, 1*, 31. doi: 10.3389/fpsyg.2010.00031

Misyak, J. B., Christiansen, M. H., & Tomblin, J. B. (2010b). Sequential expectations: the role of prediction-based learning in language. *Topics in Cognitive Science, 2*, 138-153. doi: 10.1111/j.1756-8765.2009.01072.x

Nahum, M., Nelken, I., & Ahissar, M. (2010). Stimulus uncertainty and perceptual learning: similar principles govern auditory and visual learning. *Vision Research, 50*, 391-401. doi: 10.1016/j.visres.2009.09.004

Negash, S., Howard, D. V., Japikse, K. C., & Howard, J. H. (2003). Age-related differences in implicit learning of non-spatial sequential patterns. *Aging, Neuropsychology, and Cognition, 10*, 108-121. doi: 10.1076/anec.10.2.108.14462

Park, D. C., & Reuter-Lorenz, P. (2009). The adaptive brain: aging and neurocognitive scaffolding. *Annual Review of Psychology, 60*, 173-196. doi: 10.1146/annurev.psych.59.103006.093656

Pavlovskaya, M., & Hochstein, S. (2004). Cross-dimensional transfer and specificity of perceptual learning effects. *Perception, 33*, 137-137. doi: 10.1068/v040041

Peelle, J. E., Troiani, V., Wingfield, A., & Grossman, M. (2010). Neural processing during older adults‘ comprehension of spoken sentences: age differences in resource allocation and connectivity. *Cerebral Cortex, 20*(4), 773-782. doi: 10.1093/cercor/bhp142

Peelle, J. E., & Wingfield, A. (2005). Dissociations in perceptual learning revealed by adult age differences in adaptation to time-compressed speech. *Journal of Experimental Psychology. Human Perception and Performance, 31*, 1315-1330. doi: 10.1037/0096-1523.31.6.1315

Petrov, A. A., Dosher, B. A., & Lu, Z. (2005). The dynamics of perceptual learning: an incremental reweighting model. *Psychological Review, 112*, 715-743. doi: 10.1037/0033-295X.112.4.715

Ramscar, M., Hendrix, P., Shaoul, C., Milin, P., & Baayen, H. (2014). The myth of cognitive decline: non-linear dynamics of lifelong learning. *Topics in Cognitive Science, 6*, 5-42. doi: 10.1111/tops.12078

Sabin, A. T., Clark, C. A., Eddins, D. A., & Wright, B. A. (2013). Different patterns of perceptual learning on spectral modulation detection between older hearing-impaired and younger normal-hearing adults. *Journal of the Association for Research in Otolaryngology, 14*, 283-294. doi: 10.1007/s10162-012-0363-y

Saffran, J. R. (2003). Statistical language learning: mechanisms and constraints. *Current Directions in Psychological Science, 12*, 110-114. doi: 10.1111/1467-8721.01243

Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning: listening (and learning) out of the corner of your ear. *Psychological Science, 8*(2), 101-105. doi: 10.1111/j.1467-9280.1997.tb00690.x

Salthouse, T. A. (2011). What cognitive abilities are involved in trail-making performance? *Intelligence, 39*, 222-232. doi: 10.1016/j.intell.2011.03.001

Salthouse, T. A., McGuthry, K. E., & Hambrick, D. Z. (1999). A framework for analyzing and interpreting differential aging patterns: application to three measures of implicit learning. *Aging, Neuropsychology, and Cognition, 6*, 1-18. doi: 10.1076/anec.6.1.1.789

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime User's Guide*. Pittsburgh, PA: Psychology Software Tools Inc.

Sheldon, S., Pichora-Fuller, M. K., & Schneider, B. A. (2008). Effect of age, presentation method, and learning on identification of noise-vocoded words. *The Journal of the Acoustical Society of America, 123*, 476-488. doi: 10.1121/1.2805676

Simon, J. R., Howard, J. H., & Howard, D. V. (2011). Age differences in implicit learning of probabilistic unstructured sequences. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences, 66*, 32-38. doi: 10.1093/geronb/gbq066

Smith, M. W., Sharit, J., & Czaja, S. J. (1999). Aging, motor control, and the performance of computer mouse tasks. *Human Factors, 41*, 389-396. doi: 10.1518/001872099779611102

Souza, P. E., & Boike, K. T. (2006). Combining temporal-envelope cues across channels: effects of age and hearing loss. *Journal of Speech, Language, and Hearing Research, 49*, 138-149. doi: 10.1044/1092-4388(2006/011)

Thompson, S. P., & Newport, E. L. (2007). Statistical learning of syntax: the role of transitional probability. *Language Learning and Development, 3*(1), 1-42. doi: 10.1080/15475440709336999

Turk-Browne, N. B., Junge, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology. General, 134*(4), 552-564. doi: 10.1037/0096-3445.134.4.552

Verhaeghen, P., & De Meersman, L. (1998). Aging and the Stroop effect: a meta-analysis. *Psychology and Aging, 13*, 120-126. doi: 10.1037/0882-7974.13.1.120

Versfeld, N. J., Daalder, L., Festen, J. M., & Houtgast, T. (2000). Method for the selection of sentence materials for efficient measurement of the speech reception threshold. *The Journal of the Acoustical Society of America, 107*, 1671-1684. doi: 10.1121/1.428451

Vitevitch, M. S., Armbruster, J., & Chu, S. Y. (2004). Sublexical and lexical representations in speech production: effects of phonotactic probability and onset density. *Journal of Experimental Psychology. Learning Memory and Cognition, 30*(2), 514-529. doi: 10.1037/0278-7393.30.2.514

Wechsler, D. (2004). *Wechsler Adult Intelligence Scale* (3rd ed.). Amsterdam: Harcourt Test.

Wild, C. J., Yusuf, A., Wilson, D. E., Peelle, J. E., Davis, M. H., & Johnsrude, I. S. (2012). Effortful listening: the processing of degraded speech depends critically on attention. *The Journal of Neuroscience, 32*, 14010-14021. doi: 10.1523/JNEUROSCI.1528-12.2012

Wurm, L. H., & Fisicaro, S. A. (2014). What residualizing predictors in regression analyses does (and what it does not do). *Journal of Memory and Language, 72*, 37-48. doi: 10.1016/j.jml.2013.12.003

York, R. (2012). Residualization is not the answer: rethinking how to address multicollinearity. *Social Science Research, 41*(6), 1379-1386. doi: 10.1016/j.ssresearch.2012.05.014

Zhang, J., Kuai, S., Xiao, L., Klein, S. A., Levi, D. M., & Yu, C. (2008). Stimulus coding rules for perceptual learning. *PLoS Biology, 6*, e197. doi: 10.1371/journal.pbio.0060197

Zhang, J., Zhang, G., Xiao, L., Klein, S. A., Levi, D. M., & Yu, C. (2010). Rule-based learning explains visual perceptual learning and its specificity and transfer. *Journal of Neuroscience, 30*, 12323-12328. doi: 10.1523/JNEUROSCI.0704-10.2010

5

Adult Age Effects in Auditory Statistical Learning

This chapter is based on:

Neger, T. M., Rietveld, T., & Janse, E. (2015). Adult age effects in auditory statistical learning. In M. Wolters, J. Livingstone, B. Beattie, R. Smith, M. MacMahon, J. Stuart-Smith, & J. Scobbie (Eds.), Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS 2015). London: International Phonetic Association.

Abstract

Statistical learning plays a key role in language processing, e.g., for speech segmentation. Older adults have been reported to show less statistical learning on the basis of visual input than younger adults. Given age-related changes in perception and cognition, we investigated whether statistical learning is also impaired in the auditory modality in older compared to younger adults and whether individual learning ability is associated with measures of perceptual (i.e., hearing sensitivity) and cognitive functioning in both age groups. Thirty younger and thirty older adults performed an auditory artificial-grammar-learning task to assess their statistical learning ability. In younger adults, perceptual effort came at the cost of processing resources required for learning. Inhibitory control (as indexed by Stroop color-naming performance) did not predict auditory learning. Overall, younger and older adults showed the same amount of auditory learning, indicating that statistical learning ability is preserved over the adult life span.

Introduction

Language has been argued to be probabilistic in nature (Auer & Luce, 2005). In line with this idea, frequencies with which units co-occur have been shown to play an important role in perception at various linguistic levels. At the word level, for example, sequences of phonological elements are not equally probable. Consider the phonological sequences /kæ/ as in *cat* and /hæ/ as in *hat*. Both sequences are legal word beginnings in English. However, /æ/ is more likely to follow /k/ than /h/. By the age of eight months, infants are already sensitive to these phonotactic probabilities and they, like as adults (Pitt & McQueen, 1998), make use of these statistical properties to segment fluent speech into words (Saffran, Aslin, & Newport, 1996). At the sentence level, transitional probabilities between words have been found to facilitate speech segmentation into phrases, thereby enabling syntax acquisition (Thompson & Newport, 2007).

The ability to implicitly extract such statistical regularities from input is called statistical learning (Misyak & Christiansen, 2012). As sensitivity to statistical regularities appears to be essential in online speech processing, language users with better statistical learning ability are expected to show better speech processing performance. Indeed, statistical learning ability has been shown to predict sentence processing performance in younger adults (Neger, Rietveld, & Janse, 2014). In older adults, however, deficits have been reported in the ability to learn probabilistic associations from visual input (Neger et al., 2014; Simon, Howard, & Howard, 2011). This has been taken as evidence for a more general decrease in pattern sensitivity in older age (Negash, Howard, Japikse, & Howard, 2003). If older adults indeed have generally poorer pattern sensitivity than younger adults, then older adults' statistical learning performance should also be affected in a different (i.e., non-visual) modality. Given the importance of statistical learning for language and speech processing, the present study investigated whether younger and older adults also differ in auditory statistical learning.

Importantly, age-related declines in perceptual and cognitive abilities may be expected to lead to an age group difference in auditory statistical learning. Many older adults suffer from high-frequency hearing loss (Lin, Thorpe, Gordon-Salant, & Ferrucci, 2011). That is, older adults' ability to extract acoustic information from the speech signal is not only poorer, but auditory processing also becomes more effortful, which may take resources that would otherwise be available for encoding the information in memory (McCoy et al., 2005). Reduced hearing sensitivity may therefore limit auditory learning. Moreover, the ability to inhibit irrelevant information is often reported to decline with age (Hasher, Zacks, & May, 1999). As older adults may be less able to ignore extraneous information, they may be less sensitive to relevant regularities. Therefore, the current study also investigated whether individual statistical learning ability is related to an individual's hearing sensitivity and inhibitory control.

Method

Participants

Thirty younger adults aged between 18 and 30 years ($M = 21.6$ years, $SD = 2.9$) and 30 older adults aged between 61 and 77 years ($M = 67.9$ years, $SD = 4.7$) participated in the current study. Participants were recruited via the participant pool of the Max Planck Institute for Psycholinguistics and were paid €8 per hour for their participation.

Hearing sensitivity

To assess participants' auditory functioning, we measured air-conduction pure-tone thresholds with an Oscilla USB-300 screening audiometer. The pure-tone average [PTA] was calculated over 1, 2 and 4 kHz to account for age-related high-frequency hearing loss. As auditory stimuli were presented binaurally, the PTA of the better ear served as index of hearing sensitivity, with higher values indicating poorer hearing. Younger adults had a mean PTA of 3.78 dB HL ($SD = 6.15$) and older adults of 18.22 dB HL ($SD = 6.81$). Pure-tone average thresholds differed significantly between age groups ($t(58) = 8.56$, $p < .05$).

Inhibitory control

Participants' performance on the Stroop color word test (Hammes, 1978) was taken as a measure of inhibitory control. The Stroop test consisted of three subtasks (I-III). Each subtask consisted of 100 stimuli regarding the colors blue, green, red and yellow: (I) color words printed in black ink, (II) colored patches and (III) color words printed in a conflicting color. Stimuli for each subtask were printed on a white A4 sheet (landscape orientation) and arranged in a 10x10 array. Participants were asked to read the color words printed in black (subtask I), name the color of the patches (subtask II) and name the ink color of the color words (subtask III) as quickly and accurately as possible. Participants' time to complete each task was measured (in seconds). An interference score was calculated for each individual by subtracting the time for completing subtask II from that of subtask III. The higher the score, the more difficult it was for participants to ignore the distracting incongruent information (i.e., color words) during color-naming.
On average, younger adults took 20.3 s (SD = 6.62) longer to name the 100 ink colors in the presence of distracting information. Older adults needed an additional 35.1 s (SD = 12.92). The difference in inhibitory control between age groups was significant (*Welch's* $t(1, 57.49) = 5.58$, $p < .05$).

Auditory statistical learning

We adopted the artificial grammar learning - serial reaction time paradigm (Misyak & Christiansen, 2012) as it was built to resemble statistical learning in online language processing. Materials contained eight monosyllabic Dutch CVC-nonwords (i.e., *lin, jom, taf, bur, zol, pes, mig, vun,*) used in previous studies on statistical learning (Vuong, Meyer, & Christiansen, 2011). Stimuli were recorded by a 65 year old male native speaker of Dutch. Mean stimulus duration was 442 ms (SD = 60).
On each trial, participants were presented with a visual display with four quadrants, and one printed nonword in each of the four

quadrants. Participants were instructed to click as quickly as possible on two target nonwords that would be presented auditorily one after the other (cf., Figure 5.1). The second target was only presented once the participant had clicked the correct first target. The first target was always located left (i.e., in the upper left or lower left quadrant) and the second target was always located right (i.e., in the upper right or lower right quadrant) but the specific target positions were randomly assigned. As such, within each column, one nonword served as target and one as distractor.

Participants had no possibility to anticipate the first target. Crucially, which of the two nonwords from the right column was going to be presented was dependent on the first target nonword. That is, nonwords were grouped into two grammatical sets. Within each set, two nonwords were selected as 'leaders' (Set 1: *jom*, *lin*; Set 2: *taf, bur*) which served as first targets only. The remaining two nonwords of a set were 'followers' (Set 1: *pes, vun*; Set 2: *mig, zol*), as they only appeared as second targets, following a leader nonword of the same set. Thus, four combinations of nonwords were legal within a set, resulting in a total of eight grammatical combinations (i.e., Set 1: *jom-pes, jom-vun, lin-pes, lin-vun;* Set 2: *taf-mig, taf-zol, bur-mig, bur-zol)*. Given that a target could only follow a nonword from the same set, the transitional probability from the first to the second target was 1.0 within a trial. Within the grammar, however, the transitional

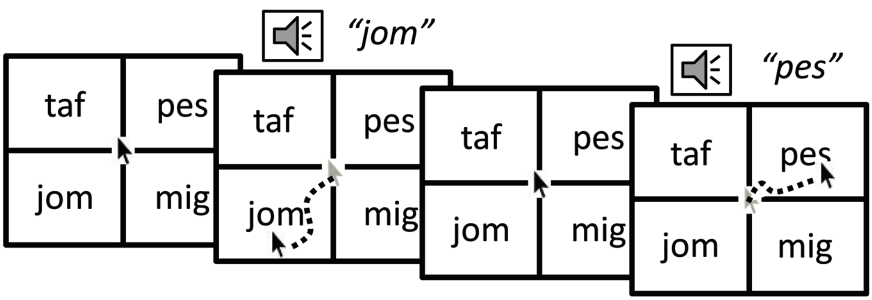

Figure 5.1. Procedure of a grammatical trial during the exposure phase of the statistical learning task.

probability between leaders and followers was 0.5 as a leader could precede two possible followers (cf., Figure 5.1, '*jom*' can be followed by '*pes*' or '*vun*' (the latter is not in the display), but never by '*mig*').
In total, the statistical learning task consisted of 20 blocks of eight trials each. The blocks were subdivided into three phases. The exposure phase spanned 16 blocks. Each grammatical combination was presented once in each block, such that participants were repeatedly exposed to the different grammatical combinations. Once participants start to implicitly detect the regularities in the input, they should become faster in clicking on the second target compared to the first target. This facilitation was measured by dividing participants' response time to the first *unpredictable* target by their response time to the second *predictable* target per trial. However, as participants may also speed up their click responses over trials, improvement during the exposure phase may partly reflect task learning.
To control for task learning effects, we implemented a test phase. In the test phase, which consisted of two blocks, the grammar was reversed. That is, leaders from one subset were now followed by followers from the other subset (e.g., *jom-mig*). If participants had detected the underlying patterns during exposure, they should show a drop in facilitation scores as they would need to correct their initial expectations. Thus, participants' statistical learning ability was operationalized as their drop in performance from the end of the exposure phase (i.e., blocks 13-16) to the test phase (blocks 17-18).
The last two blocks constituted the recovery phase and served as a control. In this phase, the grammatical combinations were re-introduced. By re-introducing the original grammar, participants' performance should not decrease any further.

Results

Age effects in auditory statistical learning

Participants' facilitation scores were analyzed by means of linear mixed-effect models using the lmer function from the lme4 package (Bates, Maechler, & Bolker, 2012) in R. Facilitation scores were restricted to those within 2.5 standard deviations from the age group's mean. Mean facilitation scores per age group and block are displayed in Figure 5.2.

To explore age group differences in statistical learning, we tested the influence of two predictors and their interaction on facilitation scores. These predictors of interest were the fixed categorical variables of age group (i.e., younger or older adults) and phase, which indicated whether a participant was exposed to grammatical trials at the end of the exposure phase (blocks 13-16) or to ungrammatical trials during the test phase. The position of the first target (i.e., upper left or lower left), the alignment of targets within a trial (i.e., horizontal or diagonal) and the interaction between target position and alignment were entered as fixed control variables.

In the random effect structure, participants were assumed to differ in their facilitation scores (random effect of participant) as well as in their amount of statistical learning (random slope of phase on participant). Moreover, it was tested whether individuals varied in

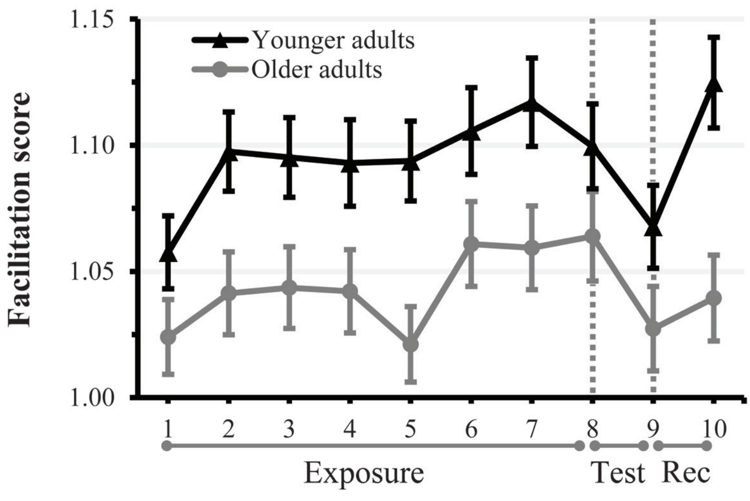

Figure 5.2: Statistical learning performance per age group and block (of 16 trials). Error bars indicate one standard error from the mean. The area between the dotted lines represents where the effect of removing the regularities is observed.

their sensitivity to target position and target alignment (random slopes of position and alignment on participant). In a stepwise selection procedure, interactions were removed before predictors if they did not attain significance at the 5% level.

The ensuing most parsimonious model explained facilitation scores as a function of phase, age group, target position and target alignment. Participants' facilitation scores were lower if the first target appeared upper left compared to lower left. This suggests that participants anticipated targets to appear in this position, probably due to influences of the Western writing system. Moreover, participants' facilitation scores were higher if targets were aligned diagonally. This implies that participants were biased towards a diagonal mouse movement. These effects of control predictors also emerged from all subsequent analyses. Going from the end of the exposure phase to the test phase resulted in a drop in facilitation scores ($\beta = -0.044$, $SE = 0.016$, $t = -2.76$, $p = .007$), thereby indicating statistical learning. Older adults showed overall lower facilitation scores than younger adults ($\beta = -0.05$, $SE = 0.021$, $t = -2.32$, $p = .021$). However, age group did not interact with phase, suggesting that the amount of statistical learning did not differ between younger and older adults.

Note that we also tested whether younger and older adults differed in their improvement over the course of the exposure blocks and in response to re-introducing the grammatical regularities in the recovery phase. This was not the case.

Individual predictors of auditory learning

Individual predictors of auditory statistical learning were identified within the separate age groups. We used the same approach as described in the previous section on age effects in auditory statistical learning, but instead of age group, hearing sensitivity, inhibitory control and their respective interactions with phase were included in the fixed-effect structure of the model. In the younger adults, the best-fitting model showed effects of target position, target alignment and phase ($\beta = -0.048$, $SE = 0.022$, $t = -2.18$, $p = .030$), the latter indicating statistical learning. Importantly, this effect of phase was modified by hearing sensitivity (β = 0.008, *SE* = 0.004, *t* = 2.16, *p* = .032): the poorer younger adults' hearing sensitivity, the less they were affected by removing the underlying regularities in the test phase and, hence, the less they learned.

In older adults, facilitation scores were explained by target position and alignment. Facilitation scores indicated a trend to drop in the test phase (β = -0.042, *SE* = 0.022, *t* = -1.95, *p* = .055). Though this effect just missed significance, the more powerful age group comparison showed learning in both younger and older adults. Overall, older adults with poorer hearing showed higher facilitation scores (β = 0.005, *SE* = 0.002, *t* = 2.91, *p* = .005). However, none of the participant characteristics interacted with phase and, thus, none were associated with older adults' amount of auditory statistical learning.

Discussion

Based on findings from visual statistical learning (Negash et al., 2003), older adults have been argued to be generally less sensitive to co-variation in the environment than younger adults. The current study investigated whether a reduced sensitivity to statistical properties can also be observed for *auditory* input, given the importance of statistical learning for speech processing. Our results showed the same amount of auditory statistical learning for both age groups. This result thus challenges the notion of a general age-related decline in pattern sensitivity. Even though hearing loss may impact on auditory statistical learning (as is evident from the younger adult data), the ability to implicitly detect regularities in an input is not affected by age per se. However, older adults apparently experience difficulties in deriving sequential patterns from the visual modality. This is in line with previous studies indicating that auditory learning is superior to visual learning in sequence learning tasks (Conway & Christiansen, 2005).

Our age group comparison also showed that the relation between the first and second click response differed between younger and older adults. Overall, older adults showed lower facilitation scores than younger adults. As we implemented a speeded computer mouse task, this was probably due to age effects on motor speed (Smith, Sharit, & Czaja, 1999).

A second aim of this study was to investigate the association between individual perceptual and cognitive abilities and auditory statistical learning performance. The results show that in both younger and older adults the amount of auditory learning was not predicted by individual inhibitory control. However, as we adopted a rather simple grammar in the current study, little task-irrelevant information was present. Under more natural conditions of auditory statistical learning, e.g., in speech processing, the input is less controlled and contains more distracting information. Therefore, inhibitory control might play a role in more demanding situations of auditory statistical learning.

In younger adults, those with poorer hearing (within a normal hearing range) showed smaller amounts of statistical learning. This suggests that perceptual effort comes at the cost of processing resources required for auditory learning. Although older adults' hearing was generally poorer (within a normal to near-normal range) than that of younger adults, this hearing effect on learning was not observed in the group of older adults. Possibly, this was due to the availability of supportive visual information throughout the task. Older adults may have implicitly compensated for the loss of acoustic detail by attending more to the visual information present. Learners have been shown to successfully integrate multimodal input during statistical learning (Mitchel, Christiansen, & Weiss, 2014). Therefore, we may speculate that increased attention of older adults to the written presentations of the nonwords and, thus, a better integration of the information from both modalities may have compensated for hearing loss effects on processing effort in the auditory modality. Better integration of the information from both modalities may also account for the finding that older adults with poorer hearing showed overall higher facilitation scores than older adults with better hearing. By paying more attention to the printed nonwords, participants may remember their positions better and are, hence, faster in locating the correct target.

Our results add to a growing body of studies on possible adult age effects on statistical learning (Negash et al., 2003; Simon et al., 2011).

In contrast to earlier findings on visual statistical learning, however, no evidence for an age-related decline in the sensitivity to statistical regularities was observed. Our findings suggest that the general ability of statistical learning is preserved over the adult life span, even though perceptual effort due to poorer hearing poses a challenge to auditory statistical learning.

References

Auer, E., & Luce, P. A. (2005). Probabilistic phonotactics in spoken word recognition. In D. B. Pisoni & R. Remez (Eds.), *The Handbook of Speech Perception* (pp. 610-630). Oxford: Blackwell.

Bates, D., Maechler, M., & Bolker, B. (2012). lme4: Linear mixed-effects models using S4 classes (Version 0.999-0). Retrieved from http://CRAN.R-project.org/package=lme4

Conway, C. M., & Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 31*, 24-39. doi: 10.1037/0278-7393.31.1.24

Hammes, J. (1978). De Stroop Kleur-woord Test (2nd ed.). Amsterdam: Pearson.

Hasher, L., Zacks, R. T., & May, C. P. (1999). Inhibitory control, circadian arousal, and age. In D. Gopher & A. Koriat (Eds.), *Attention and Performance XVII, Cognitive Regulation of Performance: Interaction of Theory and Application* (Vol. 17, pp. 653-675). Cambridge, MA: MIT Press.

Lin, F. R., Thorpe, R., Gordon-Salant, S., & Ferrucci, L. (2011). Hearing loss prevalence and risk factors among older adults in the United States. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences, 66*, 582-590. doi: 10.1093/gerona/glr002

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. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology, 58*, 22-33. doi: 10.1080/02724980443000151

Misyak, J. B., & Christiansen, M. H. (2012). Statistical learning and language: an individual differences study. *Language Learning, 62*, 302-331. doi: 10.1111/j.1467-9922.2010.00626.x

Mitchel, A. D., Christiansen, M. H., & Weiss, D. J. (2014). Multimodal integration in statistical learning: evidence from the McGurk illusion. *Frontiers in Psychology, 5*, 407. doi: 10.3389/Fpsyg.2014.00407

Negash, S., Howard, D. V., Japikse, K. C., & Howard, J. H. (2003). Age-related differences in implicit learning of non-spatial sequential patterns. *Aging, Neuropsychology, and Cognition, 10*, 108-121. doi: 10.1076/anec.10.2.108.14462

Neger, T. M., Rietveld, A. C. M., & Janse, E. (2014). Relationship between perceptual learning in speech and statistical learning in younger and older adults. *Frontiers in Human Neuroscience, 8*, 628. doi: 10.3389/fnhum.2014.00628

Pitt, M. A., & McQueen, J. M. (1998). Is compensation for coarticulation mediated by the lexicon? *Journal of Memory and Language, 39*(3), 347-370. doi: 10.1006/jmla.1998.2571

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*(5294), 1926-1928. doi: 10.1126/science.274.5294.1926

Simon, J. R., Howard, J. H., & Howard, D. V. (2011). Age differences in implicit learning of probabilistic unstructured sequences. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences, 66*, 32-38. doi: 10.1093/geronb/gbq066

Smith, M. W., Sharit, J., & Czaja, S. J. (1999). Aging, motor control, and the performance of computer mouse tasks. *Human Factors, 41*, 389-396. doi: 10.1518/001872099779611102

Thompson, S. P., & Newport, E. L. (2007). Statistical learning of syntax: the role of transitional probability. *Language Learning and Development, 3*(1), 1-42. doi: 10.1080/15475440709336999

Vuong, L. C., Meyer, A. S., & Christiansen, M. H. (2011). *Simultaneous online tracking of adjacent and non-adjacent dependencies in statistical learning.* Paper presented at the Proceedings of the 33rd Annual Conference of the Cognitive Science Society, Austin, TX.

6

Cognitive Predictors of Speech Perception Outcome after Cochlear Implantation

This chapter is based on:

Neger, T. M., Huinck W. J., Mylanus, E. A. M., Rietveld, T. and Janse, E. (in preparation). Cognitive predictors of speech perception outcome after cochlear implantation.

Abstract

Though cochlear implantation generally leads to considerable increases in patients' speech understanding performance, patients vary widely in their performance with a cochlear implant (CI). Recent studies suggest that cognitive abilities may underlie individual differences in adaptation to unfamiliar speech input. We therefore aimed to identify cognitive abilities relating to CI success in order to pave the way for future individualized rehabilitation programs. However, speech materials used to determine CI success (clearly articulated words and sentences) differ considerably from CI recipients' everyday speech input. Therefore, we also assessed patients' performance on more naturalistic speech material and how performance on different types of speech materials relates to patients' quality of life.

We assessed speech perception in nine postlingually deafened adults one week, three weeks and eleven weeks after implant activation by administering standard tests of word and sentence identification and a novel task in which words had to be detected in conversational speech. Participants were also tested on cognitive abilities of memory, processing speed, statistical learning and linguistic ability and filled in a quality of life questionnaire.

Correlational analysis was preliminary and limited to a description of effect sizes due to the small number of subjects. Speech perception improvement showed large associations with processing speed, vocabulary knowledge and memory capacities, particularly on the conversational speech test. Out of the different speech material types, performance and improvement on the conversational speech test over the first three months showed the strongest association with individuals' quality of life three months after implant activation as well as with patients' improvement therein.

Individual cognitive abilities in the ability to rapidly encode and process auditory information may underlie patients' initial adaptation success with a CI. Moreover, the current study highlights the potential that more naturalistic speech material offers for clinical practice.

Introduction

Postlingually deafened adults who are provided with a cochlear implant (CI) face a great challenge: they have to learn to interpret an unfamiliar and rather artificially sounding speech signal. Intriguingly, already within a couple of weeks, dramatic increases in patients' speech understanding performance have been reported. However, the benefit a patient experiences from a cochlear implant varies considerably between individuals. For example, individual improvement in spoken word recognition after six months of CI activation was found to range between 2% and 66% among CI recipients (Heydebrand, Hale, Potts, Gotter, & Skinner, 2007). This large variation in individual improvement raises the question which factors predict speech perception outcome after implantation. Knowledge about such factors is of clinical relevance as it may pave the way for individualized training and rehabilitation programs.

Over the last decades, impairment-related factors have been the main object of study as possible factors accounting for this inter-individual variability. Although age at onset of hearing loss (e.g., Kaplan, Shipp, Chen, Ng, & Nedzelski, 2003), duration of hearing loss (e.g., Chan et al., 2007; Hamzavi, Baumgartner, Pok, Franz, & Gstoettner, 2003; Holden et al., 2013; Oh et al., 2003) and to lesser extents etiology (e.g., Blamey et al., 1996; Geier, Barker, Fisher, & Opie, 1999) and age at implantation (e.g., Holden et al., 2013) have been shown to explain adults' variability in speech perception skills after implantation, these factors were found to explain only 10 to 21% of the variance in CI outcome (cf., Blamey et al., 1996; Blamey et al., 2013). This suggests that a large amount of variability between individual CI recipients remains unexplained.

Individual cognitive abilities may account for additional variance in CI recipients. Particularly in children with CIs, researchers have argued that individual differences in cognitive abilities such as attention, learning or memory abilities may explain the large variability in CI outcome (Ingvalson & Wong, 2013; NIH Consens Statement, 1995; Pisoni, Cleary, Geers, & Tobey, 1999). Indeed, it has been found that measures of behavioral inhibition (Horn, Davis, Pisoni, & Miyamoto, 2005), novel word learning ability (Davidson, Geers, & Nicholas, 2014), short-term memory (Cleary, Pisoni, & Kirk, 2000; Edwards & Anderson, 2014; Pisoni et al., 1999; Willstedt-Svensson, Lofqvist, Almqvist, & Sahlen, 2004) and working memory (Pisoni & Cleary, 2003; Willstedt-Svensson et al., 2004) were associated with language outcome in children with CIs. However, children are typically implanted at such a young age that it is not feasible to measure cognitive abilities prior to implantation. Moreover, children's cognitive and language development are highly intertwined and the relationship between cognitive and linguistic skills may be reciprocal. On the one hand, cognitive abilities may facilitate successful adaptation to a CI and hence language development. On the other hand, access to auditory and linguistic input after CI activation may stimulate children's cognitive development, for example by increasing children's attentional skills (Khan, Edwards, & Langdon, 2005) or by facilitating their verbal intelligence (Jacobs et al., 2016).

Testing postlingually deafened CI recipients, who acquired linguistic and cognitive skills normally, allows us to approach the predictive value of specific cognitive abilities for CI adaptation success more directly. Although the number of postlingually deafened CI recipients is rapidly increasing, only few studies have tried to link cognitive performance prior to surgery to speech perception outcome after CI activation. Early studies reported that working memory (Lyxell et al., 1998) and fast processing of visual sequences (Gantz, Woodworth, Knutson, Abbas, & Tyler, 1993; Knutson et al., 1991) relate to speech perception performance at nine to eighteen months post-surgery. More recently, verbal learning ability prior to implantation was found to be predictive of CI recipients' word recognition ability at six months after CI activation (Heydebrand et al., 2007). These research findings suggest that cognitive abilities play a role in CI outcome. However, studies have so far only linked cognitive abilities to medium- and long-term outcome. More knowledge about the influence of cognitive abilities on the *initial* adaptation process may be particularly valuable as CI recipients receive intensive auditory training during the first weeks after CI activation. Consequently, most improvement in speech perception can be expected to occur during the initial weeks after CI activation. For example, if memory abilities are associated with the amount of progress, pre-implant evaluation of memory capacities may help to shape realistic expectations regarding initial progress. Furthermore, treatment may be adapted for participants with poorer memory skills by extending the rehabilitation program.

Evidence for the potential role of cognitive abilities in the success of *initial* CI adaptation comes from studies on fast perceptual learning in speech. Participants in these studies are exposed to unfamiliar speech input such as accented speech or noise-vocoded speech. Participants generally improve rapidly in speech understanding performance over exposure (e.g., over 20 sentences). Speech recognition improvement has been reported to relate to measures of linguistic ability such as vocabulary knowledge (Janse & Adank, 2012; see Chapter 4), and to measures of cognitive ability such as selective attention (Banks, Gowen, Munro, & Adank, 2015; Janse & Adank, 2012), with more improvement over exposure for those with better vocabulary and better selective attention. As CI recipients also have to adapt to a novel listening situation, adaptation to the signal of a CI may be considered a form of perceptual learning. Supporting this assumption, children who performed well with a CI showed increased activity in brain areas (i.e., inferior frontal gyrus and angular gyrus) (Giraud & Lee, 2007) that have also been linked to perceptual learning performance (Eisner, McGettigan, Faulkner, Rosen, & Scott, 2010).

Several perceptual learning studies have highlighted the importance of structural regularities (e.g., Y. Cohen, Daikhin, & Ahissar, 2013), indicating that participants may have to be able to detect specific regularities in the input for perceptual learning to occur. This ability to implicitly extract regularities from input by means of the probabilities with which properties co-occur is called statistical learning (Misyak & Christiansen, 2012). We, therefore, aim to investigate whether initial adaptation in CI users is associated with statistical learning ability. Indeed, statistical learning ability has recently been put forward as one of the underlying processes that may predict individual differences in speech and language outcome following cochlear implantation (Pisoni, Kronenberger, Chandramouli, & Conway, 2016). That is, statistical learning has been found to relate to language outcome in children with a CI (Conway, Pisoni, Anaya, Karpicke, & Henning, 2011). In the latter study, implanted children who benefited more from the presence of underlying regularities in a visual span task (i.e., children who performed better if elements of color sequences could be predicted on the basis of the preceding color, compared to unpredictable color sequences) showed higher scores on standardized measures of language outcome than implanted children who were less sensitive to these regularities. These associations remained significant even when they were controlled for mediating measures of memory capacity. Additionally, statistical learning ability predicts perceptual learning success in adults (see Chapter 4). Participants who showed better statistical learning performance (i.e., were more sensitive to the frequency with which certain shapes tended to co-occur in visual input) were more successful in adapting to noise-vocoded speech. As noise-vocoded speech roughly simulates the auditory signal of a cochlear implant, we hypothesize that individual differences in statistical learning ability may be associated with initial CI adaptation success.

Most studies on predictors of CI outcome used simple consonant-vowel-consonant (CVC) words or simple sentences to quantify individual speech perception ability. However, in everyday life, CI recipients are generally confronted with *conversational speech*. In contrast to carefully designed audiological speech materials spoken by a single selected talker at a constant rate, conversational speech may consist of sentences from different speakers that vary in speech rate and articulatory clarity. That is, audiological testing materials, which never include hesitations and sloppily articulated speech, are clearly different from conversational speech and, hence, different from the speech input cochlear implant recipients are exposed to during the adaptation process. Processing of longer, more variable, stretches of speech may, thus, be more ecologically valid to measure patients' progress in speech understanding performance.

One advantage of highly variable speech material may be a wider performance range. In CI users, ceiling effects have been reported on standard speech materials such as simple sentences in quiet (Gifford, Shallop, & Peterson, 2008). Though ceiling effects may be avoided by presenting speech materials in noise, testing in quiet is an essential part of clinical practice to determine implant candidacy, to monitor performance over time or to evaluate new implant settings (King, Firszt, Reeder, Holden, & Strube, 2012). As speech understanding performance is decreased by variability in talkers (Sommers, 1997), speaking rate (Sommers & Barcroft, 2006) and speaking styles (Sommers & Barcroft, 2006), patients' accuracy scores on diverse speech materials may be expected to be lower and more variable than on current audiological testing materials. Indeed, normal-hearing adults listening to speech in noise (Gilbert, Tamati, & Pisoni, 2013) and cochlear implant persons listening to speech in quiet (King et al., 2012) showed high performance variability when presented with lists of TIMIT sentences (Lamel, Kassel, & Seneff, 1989), which vary in talkers, speaker gender, regional American English dialects and speaking rates. Thus, the use of conversational speech material may help to bypass ceiling effects that can be observed with current audiological testing material and may help to tap into perceptual and neurocognitive processes relevant for listening in everyday communication.

Assessing speech understanding performance by means of more naturalistic speech material may also help to bridge the gap between audiometric measures and patients' quality of life. Hearing loss has been shown to reduce quality of life due to communicative, emotional and social limitations, particularly in older patients (e.g., Hogan, O'Loughlin, Davis, & Kendig, 2009; Mulrow et al., 1990). Thus, improvement in perceived quality of life after implantation is equally important to indicate success of cochlear implantation as improvement in terms of hearing and speech perception scores. In patients with postlingual hearing loss, treatment with cochlear implants has been found to reduce depressive symptoms (Choi et al., 2016) and to lead to clinically relevant improvements in quality of life (e.g., Contrera et al., 2016; Damen, Beynon, Krabbe, Mulder, & Mylanus, 2007; Hinderink, Krabbe, & Van Den Broek, 2000; Klop et al., 2008). Though patients typically improve in speech perception performance and in quality of life after implantation as a group, studies report weak or inconsistent associations between these two aspects of cochlear implant outcome (Amoodi et al., 2012; Capretta & Moberly, 2016; Damen et al., 2007; Heo, Lee, & Lee, 2013; Hinderink et al., 2000; Straatman, Huinck, Langereis, Snik, & Mulder, 2014; Vermeire et al., 2005). This suggests that current clinical measures of speech perception do not effectively measure CI users' quality of life or improvements thereof, possibly because these measures do not reflect patients' everyday listening conditions. We might therefore expect that a speech recognition test that makes use of everyday speech is more strongly associated with patients' perceived quality of life than current standard speech recognition tests. A better relationship between speech perception tests and patients' quality of life would be valuable for clinical practice as it allows clinicians to evaluate or compare new implant settings under conditions that reflect CI users' experience in everyday life. Therefore, in the current study, individuals' quality of life after cochlear implantation and changes therein will be related to speech recognition performance in three conditions of increasing ecological validity: a standard CVC-word recognition task, a sentence recognition task (clear read aloud speech from one selected speaker) and a word identification task in conversational speech (with speech from multiple speakers).

In sum, this study aims to investigate the role of possible predictors of adaptation success during the initial adaptation phase after CI activation. On the basis of earlier reports, we focus on the general cognitive abilities of statistical learning, memory, processing speed and linguistic ability as predictors of speech recognition performance. Moreover, we expect individuals' speech recognition performance to be associated with their quality of life, particularly their performance on a more ecologically valid conversational speech task.

Method

Participants

In total, nine postlingually deafened adults, i.e., patients who became deaf after the age of six, participated in the current study. All participants were native speakers of Dutch and received a unilateral cochlear implant. In order to be eligible to participate in this study, participants had to be 18 years or older, to be a cochlear implant candidate and to have normal or corrected-to-normal vision. Patients with mental or visual disabilities, patients who had received special education and patients with partial insertion of the cochlear implant or hearing loss due to meningitis or as part of a syndrome were excluded from participation in the study. As impairment-related factors have been shown to explain part of the variance in CI outcome (cf., Blamey et al., 1996; Blamey et al., 2013), we also took correlations between speech perception and the following measures into account: age at implantation, duration of hearing loss, the degree of residual hearing prior to implantation by means of air conduction pure tone averages (0.5, 1, 2, 4 kHz) of the better ear and patients' best-aided CVC phoneme identification score at 70 dB(A) prior to implantation (tested binaurally). Demographic data of the participants are specified in Table 6.1.

Statistical learning ability

To assess participants' sensitivity to underlying regularities in input stimuli, we administered a visual variant of the artificial grammar learning – serial reaction time paradigm (Misyak, Christiansen, & Tomblin, 2010). In Figure 6.1, a trial and the grammar of the task are displayed. Within a trial, participants were presented with four familiar shapes, arranged in two rows and two columns. After a short preview time, a smaller representation of one of the two shapes on the left hand side of the screen appeared in the middle of the screen. Participants were instructed to click as fast as possible on the indicated target shape. After participants had clicked on the correct target shape,

Table 6.1. Demographic data of participants. HL = hearing loss. PTA = pure tone average.

Age	Gender	Etiology of HL	Duration of HL (yrs)	Preimplant residual hearing Air (dB HL)		Preimplant CVC identification (%)
				Left ear	Right ear	
69	m	otosclerosis	40	101	99	50
28	f	hereditary HL	25	100	100	55
59	m	hereditary HL	20	94	65	70
66	m	unknown	17	89	98	67
53	m	unknown	14	76	69	56
60	f	unknown	20	100	83	43
73	m	otosclerosis	28	113	58	81
48	m	hereditary HL	16	110	106	33
56	f	otosclerosis	37	115	99	90

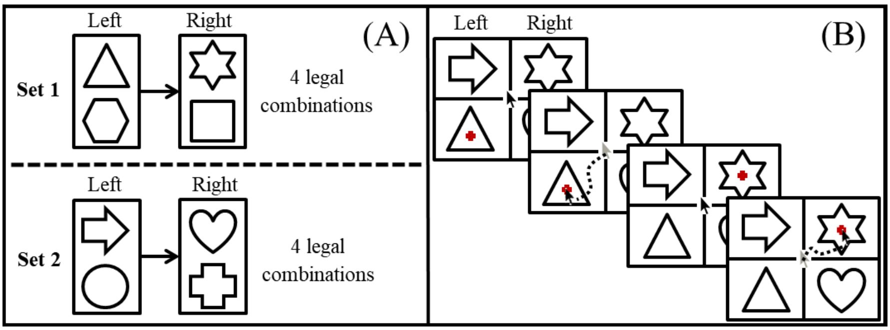

Figure 6.1. Structure of the statistical learning task. (A) Structure of the grammar in which the first target is always displayed on the left side of the screen and the second target is always displayed on the right side of the screen. (B) Procedure of a grammatical trial during the exposure phase.

a smaller representation of one of the two shapes on the right hand side of the screen appeared. Again, participants had to click as fast as possible on the indicated target shape. Crucially, which of the two right-hand side shapes was going to be cued was predictable on the basis of the identity of the first target.

The grammar consisted of eight geometrical shapes which were divided into two grammatical sets. Within each set, two shapes served as “leaders” and could only appear on the left hand side of the screen (Set 1: *triangle, hexagon*; Set 2: *arrow, circle*). As the remaining two shapes of a set always followed one of the leaders from the same grammatical set, they served as “followers” and could only appear on the right hand side of the screen (Set 1 *star, square*; Set 2: *heart, cross*). This resulted in a total of four legal combinations in a grammatical set and, hence, in a total of eight legal combinations for the statistical learning task (c.f., Figure 6.1). As each leader could be followed by one of the two followers from the same set, transitional probabilities between a leader and a follower were 0.5 within the grammar. Within a trial, however, the distractor shapes consisted of a legal combination from the competing subset. Therefore, the transitional probability between a leader and a follower was 1 within a trial. Importantly, vertical target positions were randomly assigned so that it was impossible for participants to predict whether a specific shape would appear on the upper or lower half of the screen.

The statistical learning task consisted of three phases. In the exposure phase of sixteen blocks, participants were exposed to sixteen repetitions of each legal combination (each block presenting each legal combination once). If participants started to pick up on the underlying regularities, they should become faster in clicking on the second target compared to the first target. However, faster response times can reflect both statistical and task learning. Therefore, a test phase was implemented in which the underlying regularities were reversed. That is, leaders from the one grammatical set were now followed by followers from the competing set. If participants had detected the underlying regularities, they should show a drop in performance, i.e., they should slow down in their response to the second target, as they needed to correct their initial predictions about the upcoming followers. Importantly, this slowdown can only be attributed to participants' implicit grammar sensitivity. To avoid that participants would start to adapt to the new regularities, we implemented a short test phase of two blocks of eight trials only. Statistical learning was then operationalized as the difference in performance between the end of the exposure phase (i.e., blocks 13-16) and the test phase (blocks 17-18). In the recovery phase, which served as a control phase, the original combinations were reintroduced, such that participants' performance should not drop any further. In total, the statistical learning task consisted of 160 trials (20 blocks of 8 trials).

Importantly, performance on each trial was quantified as the facilitation participants experienced by being exposed to predictable stimuli, instead of their absolute response times. To that end, we calculated facilitation scores for each trial by dividing participants' response time to the first, unpredictable target by their response time to the second, predictable target. Using participants' response times to the first target within each trial as a baseline measure allowed us to minimize effects of individual differences in click response times throughout the statistical learning paradigm, and to account for general changes in click behavior over the course of the experiment. If participants made no predictions about upcoming targets, they should be equally fast in clicking on both targets within a trial, thus, resulting in a facilitation score of 1. If participants predicted the identity of the second target, they should be faster in their response to the second compared to the first target, resulting in a facilitation score greater than 1. Thus, the facilitation score served as an index of participants' prediction of the second target within a trial.

The statistical learning task was implemented in E-prime (Schneider, Eschman, & Zuccolotto, 2002) and started with five practice trials. At the start of each trial, participants saw a preview of the visual display which consisted of the four trial shapes and two grid lines which split the screen into two equally sized rows and columns. Shapes were displayed centred in these screen quadrants in a size of 75 x 70 mm. The mouse cursor was located in the middle of the screen. After 500 ms, the first target was cued by the visual appearance of a smaller representation of the target shape (37.5 x 35 mm) in the middle of the screen. After participants had clicked on the first target item, the mouse cursor was automatically set back into the middle of the screen to ensure that the distance between the mouse and the targets was equal across all clicks. The second cue was presented 500 ms after the click to the first target. This small interval was implemented to allow for prediction effects to occur. Note that the experimental procedure only continued after participants had clicked on the correctly cued target. Clicking on a wrong item or clicking outside the target area led to slower response times and, therefore, to adjusted facilitation scores. A new trial started 500 ms after the click to the second target and automatic setback of the mouse cursor to its central position. After each block, a short break of 2500 ms was inserted to avoid fatigue effects. During this break, participants saw the block number as well as a reminder to click the target items as fast as possible on the screen. Participants took approximately 20 minutes to complete the task.

Memory

Verbal short-term memory

We administered a visual variant of an auditory nonword repetition task (Gathercole, Willis, Baddeley, & Emslie, 1994) to assess verbal short-term memory capacity without the confounding effect of hearing sensitivity. In this task, participants saw nonwords consisting of four CVC-syllables (e.g., "*kessurpardes*") and were asked to verbally repeat these nonwords after a short amount of time. The task was implemented in E-prime. At the beginning of each trial, a fixation cross appeared for 500 ms in the middle of the screen. Then the nonword was presented for 3000 ms printed boldface in Courier New with a font size of 18. After additional 3000 ms, in which participants were presented with a white screen, verbal repetition of the nonword was cued by means of an exclamation mark that appeared in the middle of the screen. The following trial started 100 ms after participants pressed the space bar to indicate that they had finished their response. After two practice trials, participants were asked to repeat 35 nonwords in total. Participants' responses were recorded and scored afterwards. We calculated the percentage of correctly repeated syllables to indicate verbal short-term memory performance. Thus, higher scores reflected better verbal short-term memory.

Short-term memory

Short-term memory performance was assessed by means of the (visually administered) digit span forward test (Daneman & Carpenter, 1980). Participants' task was to remember digit sequences ranging from two to nine digits in length and to report the sequences back by typing in the sequences in the correct order. The task was computerized and implemented in E-prime. At the start of each trial, participants saw a fixation cross for 250 ms in the middle of the screen. Digits, printed black in front of a white foreground in Arial boldface with a point size of 100, were then presented one after another for 1000 ms to the participants. Consecutive digits within a sequence were presented with an inter-stimulus-interval of 200 ms. After participants had typed in their responses, the next trial started after a short inter-trial-interval of 100 ms. Before the start of the test, participants performed two practice trials with a sequence length of three. In total, the task consisted of 16 test trials, two trials for each sequence length. Trials were presented in ascending order of sequences length. Participants' short-term memory capacity was operationalized as the percentage correctly repeated sequences. Thus, higher scores indicated better short-term memory capacity.

Working memory

Working memory performance was assessed by means of the digit span backward test (Daneman & Carpenter, 1980). The task was very similar to the digit span forward test only that participants had to report digit sequences back in *reverse* order. That is, if participants were presented with a sequence of 3-8-2, they had to report the sequence back as 2-8-3. Thus, participants' task was not only to remember the sequences but also to mentally manipulate them. The task consisted of 12 trials, as participants were only tested up to a sequence length of seven digits. The percentage of correctly repeated sequences (out of 12) served as index for working memory performance with higher scores indicating better working memory capacity.

Processing speed

Motor and mental speed

To assess motor and mental speed, we administered the digit symbol substitution test from the Wechsler Adult Intelligence Scale III (Wechsler, 2004). The goal of this paper and pencil test was to recode the digits from 1 to 9 as quickly as possible into preassigned symbols (i.e., (1 = '–'; 2 = '⊥'; 3 = '⊐'; 4 = '∟'; 5 = '⊔ '; 6 = '○'; 7 = 'Λ'; 8 = '×'; 9 = '='). The code was printed at the top of the page, such that it was continuously available to participants throughout the task. Participants performed eight practice trials prior to the start of the test. In total, the test contained 132 digits for recoding. Participants had two minutes to recode as many digits as possible. The total number of correctly recoded digits served as index of participants' processing speed. Thus, higher scores indicated faster motor and mental speed.

Scanning speed

We assessed participants' scanning speed by means of the letter comparison task (Salthouse & Babcock, 1991). In this task, participants had to indicate as quickly as possible whether two letter sequences which were simultaneously presented on the screen were same or different. Letter sequences contained consonants only and consisted of either three or six upper-case letters. The task was implemented in E-prime. At the beginning of each trial, a fixation cross was displayed for 500 ms in the middle of the screen. After a short blank screen of 100 ms, the two letter sequences were presented one above the other. Participants indicated their answer by either pressing 'm', which was marked green, if the two letter sequences were the same or by pressing 'x', which was marked red, if the two letter sequences were different.

In different trials, only one of the letters was altered. Alterations could appear on each position in the sequence. Before the start of the test, participants completed eight practice trials. Four practice trials of three-letter sequences and four practice trials of six-letter sequences. The test consisted of 24 trials of three-letter sequences and 24 trials of six-letter sequences. All participants responded correctly to more than 90% of the trials. Trials with a scanning speed of 2.5 *SD* above the mean scanning speed on accurate trials were excluded from the analysis. Participants' scanning speed performance was then defined as their average response time on accurate trials. As all other cognitive measures were coded in such a way that higher scores indicated better performance, we divided 1 by participants' response time in seconds. Thus, scanning speed reflected the number of decisions per second and compatible with all other cognitive background measures higher scores indicated faster scanning speed.

Linguistic ability

Vocabulary knowledge

To determine participants' linguistic knowledge, we administered a multiple choice vocabulary test (Andringa, Olsthoorn, van Beuningen, Schoonen, & Hulstijn, 2012). In this task, participants were asked to indicate the correct meaning of Dutch low-frequency words out of five alternatives, the last alternative always being "I don't know". The test was implemented in Excel (Courier font size 15) and consisted of two practice items and 60 test items. Each target word was embedded in a different, neutral carrier phrase. Participants could take as long as they wanted to respond. Vocabulary knowledge was measured by means of the proportion of correct answers (out of 60). Higher scores indicated greater linguistic knowledge.

Language proficiency

We performed a cloze test to obtain a global measure of language competence (see Chapter 2) as participants have to integrate knowledge of vocabulary, grammar, sentence structure, text structure, and cohesion to perform such a task (e.g., Hanania & Shikhani, 1986).

In this paper-and-pencil test, participants had to fill in missing words in three short text passages as quickly as possible. The texts did not overlap in content. The first two paragraphs were relatively easy and the third paragraph was relatively difficult to comprehend as indicated by readability index scores of the original texts. Passages were printed in order of ascending difficulty such that the measure of language proficiency took both reading comprehension as well as reading speed into account. In total, 40 content words had been removed from the original texts, i.e., 13 from the first and last paragraph and 14 from the second paragraph. Participants were required to fill in exactly one word for each missing item. Participants were free to choose appropriate words with the single constraint that they were not allowed to use a word twice. Participants had a time limit of five minutes to complete the task.

Participants could maximally obtain a test score of 80 points as each item was worth two points, i.e., one point if the selected word matched the grammatical structure of the sentence and one point if the selected word matched the content at the sentence and discourse level. Items received 0 points, if participants inserted more than one word, if a word was used for the second time or if a blank had been left empty. Language proficiency was measured by means of absolute test score. Higher scores, thus, reflected better overall language proficiency.

Quality of Life

We administered the Nijmegen Cochlear Implant Questionnaire [NCIQ] (Hinderink et al., 2000) to measure patients' health-related quality of life after cochlear implantation. The paper-and-pencil questionnaire contained 60 items regarding the six categories basic sound perception, advanced sound perception, speech production, self-esteem, activity limitations and social interactions. Each category was tested by ten questions. Items did not appear per category but were presented in mixed order. Participants were asked to provide an answer to the questions (e.g., "Are you able to recognize certain melodies in music?") on a 5-point scale ranging from "never", "sometimes", "regularly", "usually" to "always" with a sixth answer option of “not applicable”. Participants were instructed to answer the questions based on their experience with the cochlear implant. There was no time limit to complete the questionnaire.

Participants’ answers were transformed into scores from 1 (“never”) to 5 (“always). In case questions were phrased such that "never" would be considered more favorable than "always", answers were accordingly recoded into scores from 1 to 5. Higher scores thus indicated better subjective performance in the best-aided hearing condition. For each category, participants’ average outcome percentage was computed by adding together the ten item scores of each category, dividing by the number of applicable items and multiplying by five. We then obtained four measures of quality of life: “overall performance” was defined as the average QoL score across all categories; “perception performance” was defined as the average performance across the two categories basic sound perception and advanced sound perception; “production performance” reflected the average QoL score of the category “speech production”; and “social performance” was defined as the average QoL score across the two categories activity limitations and social interactions. As participants had filled in the NCIQ assessing their best-aided hearing experience as part of the regular clinical practice prior to surgery and twelve weeks after implant activation, we were also able to derive measures of participants’ subjective *improvement* in quality of life with their CI. To that end, we calculated the differences in QoL scored between participants’ pre-surgery and post-surgery evaluation of their hearing experience regarding “overall performance”, “perceptual performance”, “production performance” and “social performance”.

Speech perception

Word identification

Participants’ word perception ability was assessed by means of an auditory word identification task. Participants listened to Dutch CVC words and were asked to repeat the words. For this task, we made use of the standard word materials for speech audiometry in the Netherlands (Bosman, 1989). In each test session, participants were presented with five lists of 13 words, the first word of each list being a practice item. Participants' answers were audio recorded to allow for later response scoring. The average percentage of correctly identified phonemes (out of 180) was obtained as a measure of participants' word identification performance. Higher scores thus reflected better phoneme recognition in words.

Sentence identification

An auditory sentence identification task was administered to investigate sentence recognition performance. Participants listened to short sentences and were asked to identify and repeat these sentences. They were encouraged to guess if they were unsure. Sentences were selected from audiological test materials (Versfeld, Daalder, Festen, & Houtgast, 2000) and were all produced by the same, male speaker. Each sentence had a length of eight or nine syllables and contained four keywords. Keywords in the selected set of sentences included a noun, verb and preposition. The fourth keyword was an adjective, adverb or a second noun. An example sentence 'De sneeuw glinstert in het maanlicht' ("The snow is glistening in the moonlight") contained the keywords "sneeuw", "glinstert", "in" and "maanlicht" (a list of all test sentences used in the current study is provided in appendix A1).
The task was implemented in E-prime. In each test session, participants were first presented with ten practice trials. Practice sentences had the same length as test items but there was no overlap in sentence content between practice and test items. The twenty test sentences were then presented in random order for each participant. Participants heard a short (125 ms) 3.5 kHz tone to call their attention to the upcoming stimulus 500 ms before sentence onset. After each sentence, the researcher scored the number of correctly repeated keywords (0-4) online. The next trial started immediately after the researcher had confirmed the scoring of the previous trial. Participants' answers were audio recorded to allow for later checking of their responses. We calculated the average percentage correctly identified keywords (out of 80) as a measure of participants' sentence identification performance, with higher scores indicating better sentence recognition.

Word detection in conversational speech

To assess participants' word understanding ability in conversational situations, we administered a recently developed word detection task consisting of short conversational question-answer dyads taken from longer conversations between two speakers (Koch & Janse, 2016). Participants were instructed to click as quickly as possible on the one target word out of four visually presented words that was mentioned in the answer of the question-answer sequence. Participants also had a fifth answer option in the form of a grey square in the middle of the screen indicating that they had not detected any of the displayed word alternatives in the answer. Sequences were derived from spontaneous face-to-face dialogues in the Spoken Dutch Corpus (Oostdijk, 2000). Target words were either mono- or disyllabic and target word duration ranged between 196 and 866 ms (*M* = 372 ms, *SD* = 139).

Example.

Speaker 1: "*Iedereen heeft toch vragen?*"
"Everybody has questions, right?"

Speaker 2: "*Ja, dan moet je een afspraak maken met hem.*"
"Yes, then you should make an appointment with him."

In total, the test consisted of 60 target and 14 filler sequences (i.e., sequences in which none of the visually word alternatives occurred). The answer sequences were derived from 49 different speakers with a maximum of three answer sequences per speaker. The test was split into two parts to assess participants' word identification in conversation speech in the beginning and at the end of the three months adaptation process. That is, at both test moments participants listened to 30 target and 7 filler sequences. Sequence selection criteria, target word characteristics, and distractor word construction are described in detail in Koch & Janse (2016).
The task was implemented in Eprime. Before the start of the test, participants performed four practice trials consisting of three target items and one filler item. Each trial consisted of a speaker

familiarization phase, a preview of the four word alternatives on the screen, and a response phase. In the speaker familiarization phase, participants listened to two short utterances of approximately two seconds each that were produced by the two speaker of the subsequent test trial. The speakers were introduced in the same order in which they appeared in the upcoming stimulus (i.e., first the speaker asking a question, then the speaker answering the question). Prior to presentation of the familiarization utterances a female speaker announced the next speaker ("speaker 1" or "speaker 2"). The content of the familiarization utterances was not related to the content of the upcoming test trial. After speaker familiarization, a fixation cross was displayed for 300 ms in the middle of the screen followed by a 3000 ms preview of the four word alternatives. Words were printed in black (Courier New, font size18, bold) against a white background. Each word was presented in one of four equal-sized click regions located in each of the four quadrants of the screen. A smaller grey colored quadrant was located in the middle of the screen indicating an additional click region (labelled "none of the words"). At the beginning of each preview phase, the mouse cursor was automatically set to the centre of the screen. The response phase in which participants had to click on the target word started with the presentation of the answer-question sequence. The four word alternatives were continuously present on the screen. A trial was terminated once the participant clicked on one of the five answer alternatives but only if the key word had already occurred in the answer sequence. Stimuli were presented in random order and the location of target and distractor words was randomly assigned at the beginning of each trial. Participants' word detection ability in conversational speech was operationalized as the percentage of correctly answered trials (out of the 30 test trials).

Procedure

Participants performed all speech understanding tasks in their best-aided, bimodal hearing condition (i.e., with the hearing aid and CI they used in their daily life at that moment) to resemble patients' everyday listening experience. In all tasks, auditory stimuli were presented at a level of 70 dB(A). Stimuli were presented via two Behringer MS 16 monitor speakers that were stacked on top of each other one meter in front of the participants (0° azimuth). Prior to presentation of speech stimuli, the intensity level was controlled and calibrated with the aid of speech-shaped noise and a Monacor SM-1 sound level meter. Speech recordings were done with a Samson PM6 headset microphone that was connected to a Roland R-05 portable recorder (44.1kHz, 16 bit recordings).

Participants were recruited via the Radboud university medical center where they got their surgery. The study was formally approved by the medical ethical research committee and the clinical research centre Nijmegen (CMO # NL51331.091.14). All participants signed a written informed consent form prior to participation. Participants were tested at four points in time. The experimental design and the arrangement of tasks per session can be found in Figure 6.2. Testing of cognitive and verbal abilities required two hours and took place a couple of days prior to surgery. Patient's speech understanding performance was evaluated at follow-up meetings one week, three weeks and eleven weeks after activation of the cochlear implant. During all of these visits, participants performed the word identification and sentence identification task. Word detection in conversational speech was tested

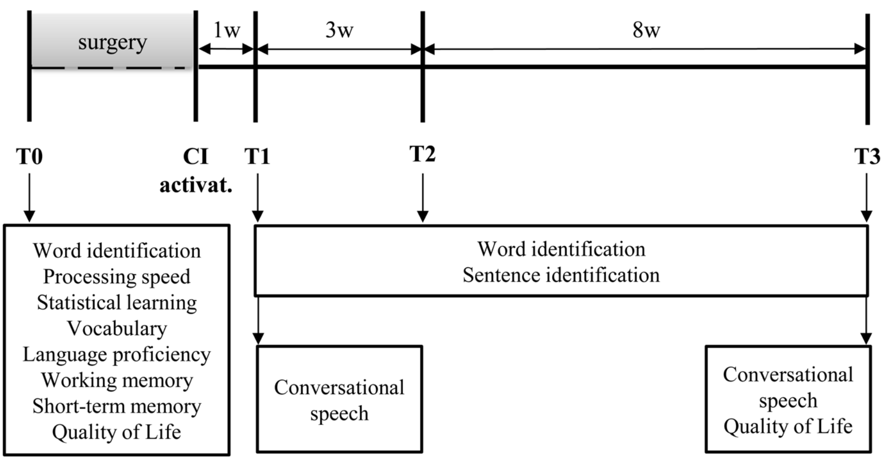

Figure 2. Experimental design.

one week and eleven weeks post activation. A follow-up session lasted either thirty (three weeks post activation) or sixty minutes (one week and eleven weeks post activation). At the end of the last test session, participants were also asked to fill in the Nijmegen Cochlear Implant Questionnaire. Participants received travel reimbursement for each visit and were compensated €40 for their time after completion of all test sessions.

Data analysis

As a first step in the correlational analysis, we determined individual speech perception improvement within the first three months after cochlear implant activation. To that end, we set up linear regression models that predicted speech recognition performance at the first and the third posttest as a function of the number of days since CI activation.[1] This way of analyzing allowed us to account for the fact that participants could not be tested at exactly the same points in time after CI activation and, hence, differed in the amount of exposure they had experienced from their CI at the moment of testing (e.g., the first posttest took place between two to fifteen days after activation, in one exceptional case 34 days after activation due to health problems of the participant; the third posttest took place between 78 to 117 days after CI activation and, hence, differed more than a month between participants). Models were run for each task and each participant separately. Participants' baseline corresponded to the modelled intercept at the day of activation and the slope corresponded to the modelled improvement on the speech perception task. That is, the unstandardized beta-coefficient of the predictor 'days since activation' represented individuals' absolute increase per day on the respective speech measure (i.e., improvement in percentage of

1 We also obtained measures of speech perception improvement that took performance at all measurement points into account as participants' speech recognition performance was measured at multiple points in time (three times for both word and sentence recognition and two times for word identification in conversational speech). To that end, we ran linear regression models that predicted speech recognition performance at all posttests by means of the number of days since CI-activation. The overall picture of the reported correlations remained the same.

correctly recognized phonemes per day for the word identification task; improvement in percentage of correctly understood keywords per day for the sentence identification task; and improvement in percentage correctly identified words in conversational speech), with a higher value indicating a steeper improvement.

As a second step of the correlational analysis, we administered Pearson's product-moment correlations to assess the relationship between cognitive abilities and patients' changes in speech perception performance as well as to assess the relationship between individuals' speech perception performance and their subjectively perceived benefit. Pearson's product-moment correlation is a suitable measure for the assessment of linear relationships in normally distributed data or data with moderate skewness and kurtosis even if the sample size is low (n = 10) (de Siqueira Santos, Takahashi, Nakata, & Fujita, 2014). All performance measures were normally-distributed as was evaluated by means of the Shapiro-Wilk test for normality. Due to the exploratory nature of the current study, effect sizes could not be determined upfront, nor was it possible to carry out a power analysis. Correlation coefficients were interpreted in terms of effect size rather than in terms of significance. That is, correlation coefficients of .5 or larger were considered of interest as they resemble large effects (J. Cohen, 1988, 1992) accounting for 25% of the variance and, thus, hint at important associations for follow-up research. However, due to the small sample size this also means that most of the observed correlations were nonsignificant. All statistical analyses were performed in the statistical software package SPSS (version 22.0.0.0).

Results

Improvement in speech perception

Speech perception performance on both the word and the sentence identification task in our within-subject design was analyzed with the GLM repeated measures procedure of SPSS. There was one fixed, within-subject factor of 'measurement point' consisting of three levels (i.e., 1 week, 4 weeks and 12 weeks post-implant activation). Due to health problems, one participant could only be tested four weeks after implant activation for the first time. The participant's missing data on the word and sentence identification task one week after implant activation were, therefore, imputed by means of the regression procedure in the missing value analysis in SPSS. Mauchley's test indicated that the assumption of sphericity had not been violated for the main effect of measurement point neither for the analysis of phoneme identification in words, χ^2 (2) = 2.42, p = .298, nor for the analysis of word identification in sentences, χ^2 (2) = 0.35, p = .839. The main effect of measurement point was significant at the 5% level for word identification, $F_{2, 16} = 5.89$, p = .012, MSE = 35.13, $\eta^2_{partial} = .424$, as well as for sentence identification, $F_{2, 16} = 3.72$, p = .047, MSE = 156.68, $\eta^2_{partial} = .317$, indicating an improvement in word and sentence identification over the first three months after implant activation. In both cases, trend analysis showed significant linear components for the effect of measurement point (i.e., word identification: $F_{1, 8} = 6.25$, p = .037; sentence identification: $F_{1, 8} = 5.53$, p = .047) whereas the quadratic component was non-significant (i.e., word identification: $F_{1, 8} = 4.80$, p = .060; sentence identification: $F_{1, 8} = 2.39$, p = .161). This suggests that performance on both the word and the sentence identification task increased linearly over time.

We performed a paired-samples t-test to assess participants' improvement on the conversational speech test, which was administered at only two points in time (i.e., one week and twelve weeks post implant activation). Note that three participants did not manage to finish the test at the first posttest as they experienced the test as too difficult. Scores of another participant were missing at the last posttest due to an experimenter's error. Thus, the sample size for the paired-samples t-test was limited to five pairs of data and, hence, had very small power. The paired-samples t-test indicated that participants did not improve in their word detection abilities in conversational speech within the first three months after implant activation ($t_4 = 0.70$, p = .523).

Table 6.2. Mean accuracy performance in % correctly identified speech units (best aided). CV = Coefficient of variation (*SD*/*M*).

Speech test	Post test	*M*	*SD*	CV	Min.	Max.
Words	1 week	71	12	.169	56	90
	4 weeks	79	8	.100	64	90
	12 weeks	80	10	.120	62	93
Sentences	1 week	71	23	.328	36	98
	4 weeks	86	12	.134	60	100
	12 weeks	84	16	.187	48	100
Conversational speech	1week	57	24	.410	33	93
	12 weeks	63	24	.378	33	100

Participants' mean performance on the three different speech tests at the three measurement points is displayed in Table 6.2. As the coefficient of variation (*SD* / *M*) suggests, participants' performance seemed to be more variable with increasing ecological validity of the task. In Figure 6.3, individual speech perception performance on the three different tasks is shown in terms of absolute performance (left panel) and in terms of modelled performance (right panel), the latter displaying participants' performance and performance changes that were used for the correlational analysis.

Improvement in Quality of Life

We performed paired-samples t-tests to assess participants' improvement in quality of life on the Nijmegen Cochlear Implant Questionnaire (NCIQ), which was administered before implantation and three months after implant activation. Note that for one of the participants, the NCIQ had not been administered as part of the standard intake procedure of the hospital prior to implantation. The paired-samples t-test indicated that within the first three months after implant activation participants experienced a significant improvement in their overall quality of life ($t_7 = 5.86$, $p < .001$, *Cohen's d* = 2.07)

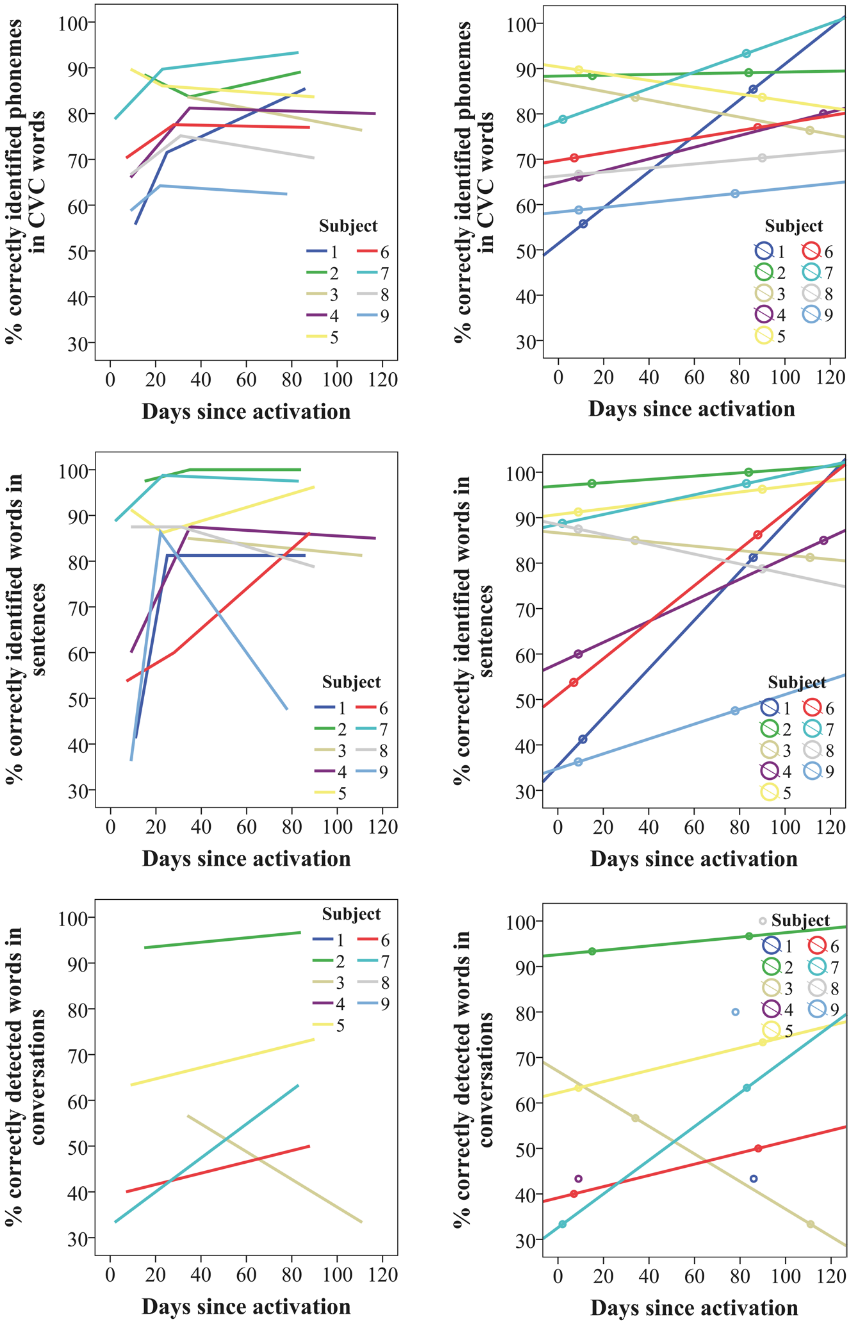

Figure 6.3. Individual speech perception in terms of observed performance (left panel) and modelled performance (right panel) regarding word identification (upper panel), sentence identification (middle panel) and word detection in conversational speech (lower panel) as a function of days since activation.

as well as in their perception (t_7 = 5.58, p < .001, *Cohen's d* = 1.97) and their social life (social life: t_7 = 5.02, p = .002, *Cohen's d* = 1.78). Regarding speech production, participants tended to show an improvement in their quality of life but it did not reach statistical significance (t_7 = 2.03, p = .082, *Cohen's d* = 0.72). Mean scores per measurement point are displayed in Figure 6.4.

Relationship between control measures and speech perception outcome

Table 6.3 displays the correlations between control measures and participants' baseline (modelled intercept at the day of activation) as well as their slope (modelled improvement) on the speech perception tasks. Participants' initial performance on all speech perception tasks seemed to be associated with participants' age in that older participants had more difficulties in performing these tasks in the first days after implant activation than younger participants. However, on the word identification task, older participants also tended to improve more than younger participants within the first three months after activation. This might be explained by the general

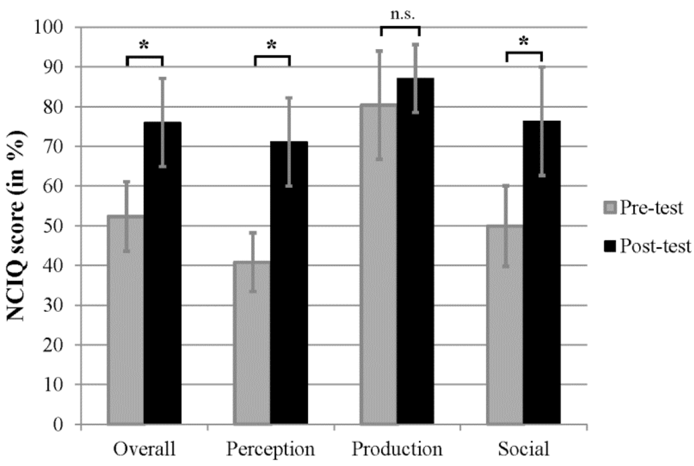

Figure 6.4. Mean scores on the Nijmegen Cochlear Implant Questionnaire (NCIQ) and three of its subcategories before implantation (grey bars) and three months after implant activation (black bars). Error bars represent one standard deviation from the mean.

relationship between starting level and improvement in adapting to a new listening condition: the maximal amount of learning can be observed if the starting level is neither too high nor too low, so that sufficient information can be derived from the acoustic materials to initiate learning while at the same time allowing for sizeable improvement (see Peelle & Wingfield, 2005). As participants' initial performance on the word identification task was relatively high (> 55 % correct, cf., Table 6.2), only participants who started with lower scores had room for improvement. Similarly to age, participants with longer durations of hearing loss initially showed lower scores on the word and the sentence identification task but also showed more improvement after implantation than cochlear implant users with shorter durations of hearing loss. Overall, impairment-related factors seemed to be particularly associated with initial performance on the word identification task: initial word identification ability tended to be decreased with poorer air conduction thresholds and with longer duration of hearing loss prior to implantation.

Participants' best-aided word identification abilities prior to implantation were unrelated to their speech perception performance after implantation. This suggests that participants' results in speech perception prior to implantation are not necessarily predictive of their CI outcomes.

Table 6.3. Pearson's product-moment correlations between control measures and participants' modelled performance on the speech perception tasks (effect sizes >.5 printed in bold).

	Words (n = 9)		**Sentences (n = 9)**		**Conversations (n = 5)**	
	Intercept	Improv.	Intercept	Improv.	Intercept	Improv.
Age	**-.515**	**.518**	**-.559**	.490	**-.918***	.290
Duration of HL	**-.584**	**.659**	**-.628**	**.531**	-.118	.376
Best Air	**-.579**	.269	-.377	.153	**.655**	-.123
Words (pre)	.032	-.070	-.173	-.071	-.231	.112

* p < .05

Cognitive predictors of speech perception outcome

The correlations between modelled speech perception performance after cochlear implantation (modeled intercept as index of initial performance and modeled slope as index of improvement) and cognitive measures are shown in Table 6.4. Overall, accurate perception of standard speech materials (i.e., CVC words and standard read aloud sentences) seemed to be mainly associated with measures of processing speed and memory. Participants with faster scanning speed showed steeper improvement in word identification than participants with slower scanning speed. Moreover, participants with greater working memory capacities and higher processing speed tended to show better sentence identification immediately after implant activation than participants with lower scores on these cognitive tests. Note, however, that participants with better working memory scores also seemed to improve less in sentence identification as they might have had less room for improvement due to the higher starting level.

Understanding high-variability speech material (i.e., conversational speech) was particularly linked to memory abilities and to vocabulary knowledge more so than performance on the other speech materials. All measured memory capacities as well as vocabulary showed a strong association with improvement in understanding conversational speech, such that individuals with better scores on these cognitive tasks were the ones to improve more in understanding conversational speech. However, better performance on these tasks was simultaneously linked to lower starting performance, again hinting at the relationship between starting level and improvement. In sum, these results suggest that working memory and processing speed are associated with patients' initial understanding of isolated words and clear sentences and that successfully improving in understanding conversational speech material with a CI is linked to memory and vocabulary knowledge.

Table 6.4. Pearson's product-moment correlations between cognitive measures and participants' modelled performance on the speech perception tasks (effect sizes >.5 printed in bold).

		Words (n = 9)		**Sentences (n = 9)**		**Conversations (n = 5)**	
		Intrcpt.	Improv.	Intrcpt.	Improv.	Intrcpt.	Improv.
Learning	Statistical learning	-.150	-.010	.038	-.401	.041	.024
Memory	Verbal memory	-.433	.261	-.287	.240	**-.704**	**.843**
	Short-term memory	.243	.216	.306	.217	**-.651**	**.923***
	Working memory	.480	-.394	**.691***	**-.630**	**-.772**	**.624**
Processing speed	Mental speed	.330	-.095	**.614**	-.417	.157	.136
	Scanning speed	-.378	**.693***	-.036	.243	.434	.173
Linguistic ability	Language proficiency	-.085	.024	.211	-.159	-.114	.187
	Vocabulary	-.028	.172	.079	.174	**-.952***	**.672**

* $p < .05$

6

Relationship between speech material and Quality of Life

For one of the participants, the NCIQ was not administered as part of the standard intake procedure of the hospital prior to implantation. Therefore, one data point regarding benefit in quality of life could not be calculated for this participant. In Table 6.5, correlations between speech perception performance and QoL twelve weeks after implant activation, as well as improvement in QoL are presented for word identification (Table 6.5A), sentence identification (Table 6.5B) and word detection in conversational speech (Table 6.5C). Overall, measures of participants' quality of life and changes therein were only related to participants' perception of more naturalistic speech material. That is, participants' improvements on standard speech perception tasks of CVC-word and sentence identification showed no associations with patients' quality of life three months after activation, nor with improvements therein. In contrast, participants who improved more in understanding conversational speech reported to be generally more satisfied with their cochlear implant three months after implantation, particularly in situations of social interaction and advanced perception (such as speech understanding). Additionally, patients who were better able to understand conversational speech immediately after implant activation also tended to experience greater improvement in quality of life from their cochlear implant. In sum, these findings suggest that ecologically valid speech materials are more promising than standard speech perception task in predicting participants' quality of life with their cochlear implant.

Table 6.5A. Pearson's product-moment correlations between subjective cochlear implant outcome (perceived CI benefit at 12 weeks post-surgery and improvement in quality of life from pre-implant to post-implant) and participants' modelled performance (Intercept referring to their initial modelled performance at the day of implant activation and Improvement referring to the modelled slope from performance one week post-implant to three months post-implant) on the word identification task (NCIQ = Nijmegen Cochlear Implant Questionnaire; effect sizes >.5 printed in bold).

		Words (n = 9)	
		Intercept	Improvement
NCIQ benefit (12 weeks post)	Perception	.040	.237
	Production	.377	-.376
	Social	.013	.356
	Total	-.006	.327
		Words (n = 8)	
		Intercept	Improvement
NCIQ improvement	Perception	-.006	-.252
	Production	.497	**-.531**
	Social	.127	-.187
	Total	.136	-.284

Table 6.5B. Pearson's product-moment correlations between subjective cochlear implant outcome and participants' modelled performance on the sentence identification task (NCIQ = Nijmegen Cochlear Implant Questionnaire; effect sizes >.5 printed in bold).

		Sentences (n = 9)	
		Intercept	Improvement
NCIQ benefit (12 weeks post)	Perception	.416	-.099
	Production	.275	.050
	Social	.229	.152
	Total	.270	.100
		Sentences (n = 8)	
		Intercept	Improvement
NCIQ improvement	Perception	.190	-.210
	Production	.411	-.351
	Social	.232	-.058
	Total	.272	-.173

Table 6.5C. Pearson's product-moment correlations between subjective cochlear implant outcome and participants' modelled performance on the conversational speech test (NCIQ = Nijmegen Cochlear Implant Questionnaire; effect sizes >.5 printed in bold).

		Conversations (n = 5)	
		Intercept	Improvement
NCIQ benefit (12 weeks post)	Perception	-.127	**.615**
	Production	-.241	.247
	Social	-.312	**.594**
	Total	-.347	**.677**
		Conversations (n = 4)	
		Intercept	Improvement
NCIQ improvement	Perception	**.672**	-.314
	Production	**.888**	**-.538**
	Social	**.706**	-.133
	Total	**.766**	-.245

6

Discussion

The aim of the current study was twofold. First, the study was set up to identify possible cognitive predictors of speech perception success in novice cochlear implant users. Second, we aimed to assess the relationship between CI-recipients' performance on different speech perception tests (varying in ecological validity) on the one hand and their perceived quality of life after cochlear implantation on the other. Speech perception was measured by means of standard tasks of word identification and sentence identification and by means of a novel word detection test with conversational speech. Within the first three months after implantation, CI users' performance on the word and sentence identification tests increased, reflecting improved perception of slowly and clearly articulated speech. Performance on the conversational speech test did not change. This may partially be due to the limited number of patients who completed the conversational speech test at both measurement points: only five out of nine participants completed the test twice. Four out of these five participants showed a higher test score three months after implant activation than at the start of the rehabilitation phase. Overall, performance on the conversational speech test was more variable than on the standard speech perception measures (i.e., word identification and sentence identification). This suggests that ceiling effects observed on standard measures such as identification of clearly articulated, short and simple sentences (Gifford et al., 2008) may indeed be avoided by the use of more variable and naturalistic speech material.

Regarding cognitive predictors of CI success, we identified some correlates of speech perception performance in novice CI recipients that seem to be promising candidates for future studies. Given that the use of more naturalistic speech material may particularly help to tap into the perceptual and neurocognitive processes underlying listening in everyday communication, individual cognitive abilities were indeed predominantly related to performance on the conversational speech test. Overall, patients with faster processing speed, higher memory capacities and greater vocabulary knowledge tended to show steeper improvement in speech perception performance after implant activation. On the one hand, this may be due to lower baseline scores these participants showed on some of the speech perception tasks: lower baseline performance allows for more improvement than high baseline performance provided that sufficient information can be derived from the acoustic materials to initiate learning (see Peelle & Wingfield, 2005). On the other hand, we may assume that these cognitive abilities actually play a role in effectively adapting to the speech signal transmitted by a CI. This provides further support that processing speed and memory capacity are positively associated with speech perception performance with a CI (cf. Gantz et al., 1993; Heydebrand et al., 2007; Knutson et al., 1991; Lyxell et al., 1998) and that linguistic knowledge predicts perceptual learning in speech (see Chapter 4). As argued in Chapter 4, linguistic knowledge may enhance the adaptation process as it facilitates access to higher-level representations.

Our results on cognitive predictors of CI success particularly tie in with recent findings showing that vocabulary knowledge, short-term memory and working memory differentiate between good and poor performers (Tamati, Gilbert, & Pisoni, 2013) on noise-vocoded stimuli of a recently developed speech recognition test for American English (PRESTO) (Gilbert et al., 2013). Like our current conversational speech test, PRESTO makes use of highly variable speech material (i.e., TIMIT sentences; Lamel et al., 1989) including utterances from different speakers varying in gender, speech rate and regional dialects. Taken together with our results, abilities that help to encode and maintain the auditory input (i.e., memory and linguistic capacity) seem to best predict learning to understand acoustically degraded and highly variable stretches of speech.

Working memory (Lyxell et al., 1998) and fast processing of visual information (Gantz et al., 1993; Knutson et al., 1991) have both been suggested to facilitate medium- to long-term speech perception outcome in postlingually deaf CI recipients. Our findings extend the current knowledge indicating that these abilities are also associated with *short-term* speech perception performance after implant activation and may, thus, play an important role over the whole rehabilitation period. Findings from hearing-aid fitting suggest that the role of working memory may even be more important during the early adaptation stages than after several months of aided hearing (Ng et al., 2014), possibly because the incoming acoustic signal deviates most strongly from stored phonological or lexical representations when patients are first exposed to the acoustic signal of a new hearing device. Consequently, the initial period of hearing aid use is the time which is most taxing with respect to the mapping of the new input onto old representations and, therefore, most likely requires the engagement of cognitive capacities (e.g., Rudner, Foo, Ronnberg, & Lunner, 2009; Rudner, Foo, Sundewall-Thoren, Lunner, & Ronnberg, 2008).

As memory capacities were most consistently related to speech perception success with a CI across speech materials, pre-implant assessment of memory capacities may be considered most promising for the evaluation and rehabilitation of CI recipients. That is, knowledge about pre-implantation memory abilities may help to shape realistic expectations regarding speech perception progress and outcome after implantation. Considering the role of memory capacity in speech processing and adaptation, it may also seem appealing to offer working memory training for patients with poorer memory scores to facilitate the adaptation process. Working memory training may have an additional long-term benefit by helping CI users to cope with listening effort. Thereby, memory training may prevent or diminish detrimental effects sustained effortful listening has on individuals' lives such as mental fatigue (Hornsby, 2013) or stress-related sick leave from work (Kramer, Kapteyn, & Houtgast, 2006). However, recent studies offering working memory training to older hearing-aid users (Ferguson & Henshaw, 2015) and older individuals with varying degrees of age-related hearing loss (Wayne, Hamilton, Jones Huyck, & Johnsrude, 2016) have failed to show transfer effects of memory training to speech perception performance. In contrast, auditory training and auditory-based cognitive training programs seem to be more effective for enhancing working memory capacities, processing speed, speech perception performance and self-reported benefit (Anderson, White-Schwoch, Choi, & Kraus, 2013; Anderson, White-Schwoch, Parbery-Clark, & Kraus, 2013; Ferguson, Henshaw, Clark, & Moore, 2014; Henshaw & Ferguson, 2014; Smith et al., 2009; Sweetow & Sabes, 2006). Yet, it is unclear whether such training programs are also effective for novice CI users. Three weeks of auditory training increased consonant discrimination abilities in a small group of experienced CI users, but it did not enhance discrimination of vowels or words in sentences (Stacey et al., 2010). Further research is needed to evaluate whether extending the rehabilitation program by auditory-cognitive training may help to facilitate adaptation to a CI, particularly in patients with poorer memory skills.

Our study focused on investigating correlates of speech perception improvement for postlingually deaf CI recipients. Yet, findings that memory capacity relates to language outcome in implanted children (e.g., Edwards & Anderson, 2014; Willstedt-Svensson et al., 2004) suggest that similar mechanisms might underlie children's' ability to acquire speech and language with their CI. Note, however, that we could not find evidence that statistical learning ability, an ability recently put forward as one of the underlying processes explaining CI success in children (cf. Pisoni et al., 2016), was related to postlingually adults' speech perception performance with a CI. Adaptation outcome in children may relate more to general learning capacities such as statistical learning ability than adaptation outcome in postlingually deaf patients. Prelingually deaf adults and children with a CI face auditory deprivation during sensitive periods of language learning and cognitive development. As auditory and linguistic input particularly carry information about temporal structures and regularities, limited exposure to auditory and language input in these sensitive periods may lead to impaired sequencing and implicit statistical learning skills.

Regarding control predictors of CI outcome, findings of the current study are consistent with studies showing that age at implantation (e.g., Holden et al., 2013), duration of hearing loss (e.g., Chan et al., 2007; Hamzavi et al., 2003; Holden et al., 2013; Oh et al., 2003) and pre-implant residual hearing (Gantz et al., 1993) explain some of the individual variability in speech perception performance after implantation. Older participants tended to initially score lower on all speech perception measures than younger participants. However, older patients also seemed to improve more than younger patients in word identification within the first three months after implant activation, indicating that age at implantation did not necessarily restrict rehabilitation success. Patients who suffered longer from hearing loss and who suffered from more profound hearing loss prior to implantation (in terms of both air and bone conduction) had more difficulties in identifying phonemes in CVC words. The perception of sentences and conversational speech, however, was unaffected by pre-implant audiometric scores. In contrast to the sentence and the conversational speech test, performance on the word identification task is not supported by the availability of context information. Moreover, CVC words are more easily confused than the keywords that were embedded in the sentence identification task (mainly being bi- and trisyllabic words) due to the larger number of phonological neighbors. This implies that auditory deprivation due to longer and more profound hearing loss may have particularly affected representations of fine-phonetic detail in auditory memory rather than higher-level representations.

The second aim of the current study was to assess the relationship between CI recipients' ability to understand different types of speech material and their quality of life after cochlear implant activation. We were specifically interested to find out whether more naturalistic speech material may show better associations with their quality of life (i.e., perceived quality of life twelve weeks after implant activation and improvement in quality of life from pre- to post implantation) compared to standard speech perception tests using CVC words or simple sentences. Overall, patients experienced a significant increase in their quality of life which was particularly evident in their sound perception and their social life. Even though patients also improved significantly on the two standard tests of word and sentence identification, which are the most commonly administered speech perception measures in clinical practice, the measures showed no associations with patients' quality of life. These results are in line with previous studies observing only inconsistent or no associations between standard speech perception measures and quality of life (Amoodi et al., 2012; Capretta & Moberly, 2016; Damen et al., 2007; Heo et al., 2013; Hinderink et al., 2000; Straatman et al., 2014; Vermeire et al., 2005), suggesting that current audiological measures do not effectively reflect patients' everyday experience with their cochlear implant.

The current study is the first to adapt more naturalistic speech material for audiological testing in Dutch, thereby following recent approaches to implement variable and naturalistic speech material in audiological practice in English (Gilbert et al., 2013; King et al., 2012; Tamati et al., 2013). It is also the first to report on the relationship between perception of conversational speech material and perceived cochlear implant benefit, showing that initial performance and improvement in understanding conversational speech seem to be predictive of CI users' quality of life and changes therein. Thus, in line with our hypothesis, the relationship between objective speech perception tasks and patients' quality of life was stronger with increasing naturalness of the speech material. This may partially be due to patients' higher performance variability on the conversational speech test compared to their performance variability on the word and sentence identification tasks. The current findings suggest that objective speech perception tests using more naturalistic speech material could indeed be used as indices of how well patients perform with their cochlear implant in everyday life. This may be beneficial for clinical practice as implants can be customized and new implant settings can be evaluated in the lab under conditions that reflect patients' daily experience. More naturalistic speech material may be used more widely, e.g., in the rehabilitation of CI recipients as training material, and as diagnostic assessment of their rehabilitation outcome. More naturalistic speech perception tests may also offer a vantage point for further studies exploring the perceptual and neurocognitive underpinnings of successful cochlear implant adaptation.

It is important to keep in mind that few significant associations could be observed due to the very small sample size and the great number of associations explored. Despite this limitation, we are confident that the general pattern of associations with strong effect sizes can hint at promising components contributing to speech perception outcome with a CI. Furthermore, we could only test for linear relationships. Future studies including larger sample sizes may also aim to test non-linear predictors of speech perception success. Adaptation processes often follow a non-linear learning curve with the steepest performance increase during initial exposure, and with a gradual levelling off over time. As the current study focused on the early adaptation process only, it may particularly be relevant to take a non-linear approach if predictors of mid- and long-term adaptation success are explored.

In short, our results suggest that individual cognitive abilities are promising in explaining patients' variability in adapting to a cochlear implant, thereby emphasizing the importance of listener-based abilities for rehabilitation. Speech perception outcome was associated with memory capacity, vocabulary knowledge and processing speed, indicating that particularly differences in the ability to rapidly encode and process auditory information may underlie patients' adaptation success with a CI. The current findings underline that similar processes are involved in perceptual learning in normal-hearing and hearing-impaired populations. Moreover, the current study highlights the potential that more naturalistic speech material offers for clinical practice.

References

Amoodi, H. A., Mick, P. T., Shipp, D. B., Friesen, L. M., Nedzelski, J. M., Chen, J. M., & Lin, V. Y. W. (2012). Results with cochlear implantation in adults with speech recognition scores exceeding current criteria. *Otology & Neurotology, 33*(1), 6-12. doi: 10.1097/MAO.0b013e318239e5a1

Anderson, S., White-Schwoch, T., Choi, H. J., & Kraus, N. (2013). Training changes processing of speech cues in older adults with hearing loss. *Frontiers in Systems Neuroscience, 7*, 97. doi: 10.3389/fnsys.2013.00097

Anderson, S., White-Schwoch, T., Parbery-Clark, A., & Kraus, N. (2013). Reversal of age-related neural timing delays with training. *Proceedings of the National Academy of Sciences of the United States of America, 110*(11), 4357-4362. doi: 10.1073/pnas.1213555110

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. *Language Learning, 62*, 49-78. doi: 10.1111/j.1467-9922.2012.00706.x

Banks, B., Gowen, E., Munro, K. J., & Adank, P. (2015). Cognitive predictors of perceptual adaptation to accented speech. *The Journal of the Acoustical Society of America, 137*(4), 2015-2024. doi: 10.1121/1.4916265

Blamey, P., Arndt, P., Bergeron, F., Bredberg, G., Brimacombe, J., Facer, G., . . . Whitford, L. (1996). Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants. *Audiology & Neuro-Otology, 1*(5), 293-306. doi: 10.1159/000259212

Blamey, P., Artieres, F., Başkent, D., Bergeron, F., Beynon, A., Burke, E., . . . Lazard, D. S. (2013). Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants: an update with 2251 patients. *Audiology and Neurotology, 18*(1), 36-47. doi: 10.1159/000343189

Bosman, A. (1989). *Speech perception by the hearing impaired.* Doctoral dissertation, Utrecht University, Utrecht, The Netherlands.

Capretta, N. R., & Moberly, A. C. (2016). Does quality of life depend on speech recognition performance for adult cochlear implant users? *The Laryngoscope, 126*(3), 699-706. doi: 10.1002/lary.25525

Chan, V., Tong, M., Yue, V., Wong, T., Leung, E., Yuen, K., & van Hasselt, A. (2007). Performance of older adult cochlear implant users in Hong Kong. *Ear and hearing, 28*(2 Suppl), 52S-55S. doi: 10.1097/AUD.0b013e318031509d

Cleary, M., Pisoni, D. B., & Kirk, K. I. (2000). Working memory spans as predictors of spoken word recognition and receptive vocabulary in children with cochlear implants. *The Volta Review, 102*(4), 259-280.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed. ed.). Hillsdale etc.: Erlbaum.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155-159. doi: 10.1037/0033-2909.112.1.155

Cohen, Y., Daikhin, L., & Ahissar, M. (2013). Perceptual learning is specific to the trained structure of information. *Journal of Cognitive Neuroscience, 25*, 2047-2060. doi: 10.1162/jocn_a_00453

Contrera, K. J., Betz, J., Li, L., Blake, C. R., Sung, Y. K., Choi, J. S., & Lin, F. R. (2016). Quality of life after intervention with a cochlear implant or hearing aid. *The Laryngoscope, 126*(9), 2110-2115. doi: 10.1002/lary.25848

Conway, C. M., Pisoni, D. B., Anaya, E. M., Karpicke, J., & Henning, S. C. (2011). Implicit sequence learning in deaf children with cochlear implants. *Developmental Science, 14*(1), 69-82. doi: 10.1111/j.1467-7687.2010.00960.x

Damen, G. W., Beynon, A. J., Krabbe, P. F., Mulder, J. J., & Mylanus, E. A. (2007). Cochlear implantation and quality of life in postlingually deaf adults: long-term follow-up. *Otolaryngology–Head & Neck Surgery, 136*(4), 597-604. doi: 10.1016/j.otohns.2006.11.044

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*(4), 450-466. doi: 10.1016/S0022-5371(80)90312-6

Davidson, L. S., Geers, A. E., & Nicholas, J. G. (2014). The effects of audibility and novel word learning ability on vocabulary level in children with cochlear implants. *Cochlear Implants International, 15*(4), 211-221. doi: 10.1179/1754762813Y.0000000051

de Siqueira Santos, S., Takahashi, D. Y., Nakata, A., & Fujita, A. (2014). A comparative study of statistical methods used to identify dependencies between gene expression signals. *Briefings in Bioinformatics, 15*(6), 906-918. doi: 10.1093/bib/bbt051

Edwards, L., & Anderson, S. (2014). The association between visual, nonverbal cognitive abilities and speech, phonological processing, vocabulary and reading outcomes in children with cochlear implants. *Ear and Hearing, 35*(3), 366-374. doi: 10.1097/AUD.0000000000000012

Eisner, F., McGettigan, C., Faulkner, A., Rosen, S., & Scott, S. K. (2010). Inferior frontal gyrus activation predicts individual differences in perceptual learning of cochlear-implant simulations. *The Journal of Neuroscience, 30*, 7179-7186. doi: 10.1523/JNEUROSCI.4040-09.2010

Ferguson, M. A., & Henshaw, H. (2015). Auditory training can improve working memory, attention, and communication in adverse conditions for adults with hearing loss. *Frontiers in Psychology, 6*, 556. doi: 10.3389/fpsyg.2015.00556

Ferguson, M. A., Henshaw, H., Clark, D. P., & Moore, D. R. (2014). Benefits of phoneme discrimination training in a randomized controlled trial of 50- to 74-year-olds with mild hearing loss. *Ear and Hearing, 35*(4), e110-121. doi: 10.1097/AUD.0000000000000020

Gantz, B. J., Woodworth, G. G., Knutson, J. F., Abbas, P. J., & Tyler, R. S. (1993). Multivariate predictors of audiological success with multichannel cochlear implants. *The Annals of Otology, Rhinology, and Laryngology, 102*(12), 909-916. doi: 10.1177/000348949310201201

Gathercole, S. E., Willis, C. S., Baddeley, A. D., & Emslie, H. (1994). The children's test of nonword repetition: a test of phonological working memory. *Memory, 2*(2), 103-127. doi: 10.1080/09658219408258940

Geier, L., Barker, M., Fisher, L., & Opie, J. (1999). The effect of long-term deafness on speech recognition in postlingually deafened adult CLARION cochlear implant users. *The Annals of Otology, Rhinology & Laryngology, 177*, 80-83. doi: 10.1177/00034894991080S416

Gifford, R. H., Shallop, J. K., & Peterson, A. M. (2008). Speech recognition materials and ceiling effects: considerations for cochlear implant programs. *Audiology & Neuro-Otology, 13*(3), 193-205. doi: 10.1159/000113510

Gilbert, J. L., Tamati, T. N., & Pisoni, D. B. (2013). Development, reliability, and validity of PRESTO: a new high-variability sentence recognition test. *Journal of the American Academy of Audiology, 24*(1), 26-36. doi: 10.3766/jaaa.24.1.4

Giraud, A. L., & Lee, H. J. (2007). Predicting cochlear implant outcome from brain organisation in the deaf. *Restorative Neurology and Neuroscience, 25*(3-4), 381-390.

Hamzavi, J., Baumgartner, W. D., Pok, S. M., Franz, P., & Gstoettner, W. (2003). Variables affecting speech perception in postlingually deaf adults following cochlear implantation. *Acta Oto-Laryngologica, 123*(4), 493-498. doi: 10.1080/0036554021000028120

Hanania, E., & Shikhani, M. (1986). Interrelationships among three tests of language proficiency: standardized ESL, cloze, and writing. *Tesol Quarterly, 20*(1), 97-109. doi: 10.2307/3586391

Henshaw, H., & Ferguson, M. A. (2014). Assessing the benefits of auditory training to real-world listening: identifying appropriate and sensitive outcomes. In T. Dau & S. Santurette (Eds.), *Auditory plasticity - listening with the brain* (pp. 45-52). Ballerup, Denmark: Danavox Jubilee Foundation.

Heo, J. H., Lee, J. H., & Lee, W. S. (2013). Bimodal benefits on objective and subjective outcomes for adult cochlear implant users. *Korean Journal of Audiology, 17*(2), 65-73. doi: 10.7874/kja.2013.17.2.65

Heydebrand, G., Hale, S., Potts, L., Gotter, B., & Skinner, M. (2007). Cognitive predictors of improvements in adults' spoken word recognition six months after cochlear implant activation. *Audiology & Neuro-Otology, 12*(4), 254-264. doi: 10.1159/000101473

Hinderink, J. B., Krabbe, P. F., & Van Den Broek, P. (2000). Development and application of a health-related quality-of-life instrument for adults with cochlear implants: the Nijmegen cochlear implant questionnaire. *Otolaryngology–Head and Neck Surgery, 123*(6), 756-765. doi: 10.1067/mhn.2000.108203

Hogan, A., O'Loughlin, K., Davis, A., & Kendig, H. (2009). Hearing loss and paid employment: Australian population survey findings. *International Journal of Audiology, 48*(3), 117-122. doi: 10.1080/14992020802449008

Holden, L. K., Finley, C. C., Firszt, J. B., Holden, T. A., Brenner, C., Potts, L. G., . . . Skinner, M. W. (2013). Factors affecting open-set word recognition in adults with cochlear implants. *Ear and Hearing, 34*(3), 342-360. doi: 10.1097/AUD.0b013e3182741aa7

Horn, D. L., Davis, R. A., Pisoni, D. B., & Miyamoto, R. T. (2005). Behavioral inhibition and clinical outcomes in children with cochlear implants. *The Laryngoscope, 115*(4), 595-600. doi: 10.1097/01.mlg.0000161340.00258.1d

Hornsby, B. W. (2013). The effects of hearing aid use on listening effort and mental fatigue associated with sustained speech processing demands. *Ear and Hearing, 34*(5), 523-534. doi: 10.1097/AUD.0b013e31828003d8

Ingvalson, E. M., & Wong, P. C. (2013). Training to improve language outcomes in cochlear implant recipients. *Frontiers in Psychology, 4*, 263. doi: 10.3389/fpsyg.2013.00263

Jacobs, E., Langereis, M. C., Frijns, J. H., Free, R. H., Goedegebure, A., Smits, C., . . . Vermeulen, A. M. (2016). Benefits of simultaneous bilateral cochlear implantation on verbal reasoning skills in prelingually deaf children. *Research in Developmental Disabilities, 58*, 104-113. doi: 10.1016/j.ridd.2016.08.016

Janse, E., & Adank, P. (2012). Predicting foreign-accent adaptation in older adults. *Quarterly Journal of Experimental Psychology, 65*, 1563-1585. doi: 10.1080/17470218.2012.658822

Kaplan, D. M., Shipp, D. B., Chen, J. M., Ng, A. H., & Nedzelski, J. M. (2003). Early-deafened adult cochlear implant users: assessment of outcomes. *The Journal of Otolaryngology, 32*(4), 245-249. doi: 10.2310/7070.2003.41601

Khan, S., Edwards, L., & Langdon, D. (2005). The cognition and behaviour of children with cochlear implants, children with hearing aids and their hearing peers: a comparison. *Audiology & Neuro-Otology, 10*(2), 117-126. doi: 10.1159/000083367

King, S. E., Firszt, J. B., Reeder, R. M., Holden, L. K., & Strube, M. (2012). Evaluation of TIMIT sentence list equivalency with adult cochlear implant recipients. *Journal of the American Academy of Audiology, 23*(5), 313-331. doi: 10.3766/jaaa.23.5.3

Klop, W. M., Boermans, P. P., Ferrier, M. B., van den Hout, W. B., Stiggelbout, A. M., & Frijns, J. H. (2008). Clinical relevance of quality of life outcome in cochlear implantation in postlingually deafened adults. *Otology & Neurotology, 29*(5), 615-621. doi: 10.1097/MAO.0b013e318172cfac

Knutson, J. F., Hinrichs, J. V., Tyler, R. S., Gantz, B. J., Schartz, H. A., & Woodworth, G. (1991). Psychological predictors of audiological outcomes of multichannel cochlear implants: preliminary findings. *The Annals of Otology, Rhinology, and Laryngology, 100*(10), 817-822. doi: 10.1177/000348949110001006

Koch, X., & Janse, E. (2016). Speech rate effects on the processing of conversational speech across the adult life span. *The Journal of the Acoustical Society of America, 139*(4), 1618. doi: 10.1121/1.4944032

Kramer, S. E., Kapteyn, T. S., & Houtgast, T. (2006). Occupational performance: comparing normally-hearing and hearing-impaired employees using the Amsterdam Checklist for Hearing and Work. *International Journal of Audiology, 45*(9), 503-512. doi: 10.1080/14992020600754583

Lamel, L. F., Kassel, R. H., & Seneff, S. (1989). *Speech database development: design and analysis of the acoustic-phonetic corpus.* Paper presented at the Speech Input/Output Assessment and Speech Databases.

Lyxell, B., Andersson, J., Andersson, U., Arlinger, S., Bredberg, G., & Harder, H. (1998). Phonological representation and speech understanding with cochlear implants in deafened adults. *Scandinavian Journal of Psychology, 39*(3), 175-179. doi: 10.1111/1467-9450.393075

Misyak, J. B., & Christiansen, M. H. (2012). Statistical learning and language: an individual differences study. *Language Learning, 62*, 302-331. doi: 10.1111/j.1467-9922.2010.00626.x

Misyak, J. B., Christiansen, M. H., & Tomblin, J. B. (2010). On-line individual differences in statistical learning predict language processing. *Frontiers in Psychology, 1*, 31. doi: 10.3389/fpsyg.2010.00031

Mulrow, C. D., Aguilar, C., Endicott, J. E., Velez, R., Tuley, M. R., Charlip, W. S., & Hill, J. A. (1990). Association between hearing impairment and the quality of life of elderly individuals. *Journal of the American Geriatrics Society, 38*(1), 45-50. doi: 10.1111/j.1532-5415.1990.tb01595.x

Neger, T. M., Janse, E., & Rietveld, A. C. M. (2015). Correlates of older adults' discrimination of acoustic properties in speech. *Speech, Language and Hearing, 18*(2), 102-115. doi: 10.1179/2050572814y.0000000055

Neger, T. M., Rietveld, A. C. M., & Janse, E. (2014). Relationship between perceptual learning in speech and statistical learning in younger and older adults. *Frontiers in Human Neuroscience, 8*, 628. doi: 10.3389/fnhum.2014.00628

Ng, E. H., Classon, E., Larsby, B., Arlinger, S., Lunner, T., Rudner, M., & Ronnberg, J. (2014). Dynamic relation between working memory capacity and speech recognition in noise during the first 6 months of hearing aid use. *Trends in Hearing, 18*. doi: 10.1177/2331216514558688

NIH Consens Statement. (1995). Cochlear Implants in Adults and Children (Vol. 13, pp. 1-30).

Oh, S. H., Kim, C. S., Kang, E. J., Lee, D. S., Lee, H. J., Chang, S. O., . . . Koo, J. W. (2003). Speech perception after cochlear implantation over a 4-year time period. *Acta Oto-Laryngologica, 123*(2), 148-153. doi: 10.1080/0036554021000028111

Oostdijk, N. (2000). The spoken Dutch corpus project. *ELRA Newsletter, 5*, 4-8.

Peelle, J. E., & Wingfield, A. (2005). Dissociations in perceptual learning revealed by adult age differences in adaptation to time-compressed speech. *Journal of Experimental Psychology. Human Perception and Performance, 31*, 1315-1330. doi: 10.1037/0096-1523.31.6.1315

Pisoni, D. B., & Cleary, M. (2003). Measures of working memory span and verbal rehearsal speed in deaf children after cochlear implantation. *Ear and Hearing, 24*. doi: 10.1097/01.AUD.0000051692.05140.8E

Pisoni, D. B., Cleary, M., Geers, A. E., & Tobey, E. A. (1999). Individual differences in effectiveness of cochlear implants in children who are prelingually deaf: new process measures of performance. *Volta Review, 101*(3), 111-164.

Pisoni, D. B., Kronenberger, W. G., Chandramouli, S. H., & Conway, C. M. (2016). Learning and memory processes following cochlear implantation: the missing piece of the puzzle. *Frontiers in Psychology, 7*, 493. doi: 10.3389/fpsyg.2016.00493

Rudner, M., Foo, C., Ronnberg, 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. *Scandinavian Journal of Psychology, 50*(5), 405-418. doi: 10.1111/j.1467-9450.2009.00745.x

Rudner, M., Foo, C., Sundewall-Thoren, E., Lunner, T., & Ronnberg, J. (2008). Phonological mismatch and explicit cognitive processing in a sample of 102 hearing-aid users. *International Journal of Audiology, 47*, S91-S98. doi: 10.1080/14992020802304393

Salthouse, T. A., & Babcock, R. L. (1991). Decomposing adult age differences in working memory. *Developmental Psychology, 27*(5), 763. doi: 10.3758/BF03197200

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime User's Guide*. Pittsburgh, PA: Psychology Software Tools Inc.

Smith, G. E., Housen, P., Yaffe, K., Ruff, R., Kennison, R. F., Mahncke, H. W., & Zelinski, E. M. (2009). A cognitive training program based on principles of brain plasticity: results from the improvement in memory with Plasticity-based Adaptive Cognitive Training (IMPACT) study. *Journal of the American Geriatrics Society, 57*(4), 594-603. doi: 10.1111/j.1532-5415.2008.02167.x

Sommers, M. S. (1997). Stimulus variability and spoken word recognition. II. The effects of age and hearing impairment. *The Journal of the Acoustical Society of America, 101*, 2278-2288. doi: 10.1121/1.418208

Sommers, M. S., & Barcroft, J. (2006). Stimulus variability and the phonetic relevance hypothesis: effects of variability in speaking style, fundamental frequency, and speaking rate on spoken word identification. *The Journal of the Acoustical Society of America, 119*(4), 2406-2416. doi: 10.1121/1.2171836

Stacey, P. C., Raine, C. H., O'Donoghue, G. M., Tapper, L., Twomey, T., & Summerfield, A. Q. (2010). Effectiveness of computer-based auditory training for adult users of cochlear implants. *International Journal of Audiology, 49*(5), 347-356. doi: 10.3109/14992020903397838

Straatman, L. V., Huinck, W. J., Langereis, M. C., Snik, A. F., & Mulder, J. J. (2014). Cochlear implantation in late-implanted prelingually deafened adults: changes in quality of life. *Otology & Neurotology, 35*(2), 253-259. doi: 10.1097/MAO.0b013e3182a4758e

Sweetow, R. W., & Sabes, J. H. (2006). The need for and development of an adaptive Listening and Communication Enhancement (LACE) Program. *Journal of the American Academy of Audiology, 17*(8), 538-558. doi: 10.3766/jaaa.17.8.2

Tamati, T. N., Gilbert, J. L., & Pisoni, D. B. (2013). Some factors underlying individual differences in speech recognition on PRESTO: a first report. *Journal of the American Academy of Audiology, 24*(7), 616-634. doi: 10.3766/jaaa.24.7.10

Vermeire, K., Brokx, J. P., Wuyts, F. L., Cochet, E., Hofkens, A., & Van de Heyning, P. H. (2005). Quality-of-life benefit from cochlear implantation in the elderly. *Otology & Neurotology, 26*(2), 188-195. doi: 10.1097/00129492-200503000-00010

Versfeld, N. J., Daalder, L., Festen, J. M., & Houtgast, T. (2000). Method for the selection of sentence materials for efficient measurement of the speech reception threshold. *The Journal of the Acoustical Society of America, 107*, 1671-1684. doi: 10.1121/1.428451

Wayne, R. V., Hamilton, C., Jones Huyck, J., & Johnsrude, I. S. (2016). Working memory training and speech in noise comprehension in older adults. *Frontiers in Aging Neuroscience, 8*, 49. doi: 10.3389/fnagi.2016.00049

Willstedt-Svensson, U., Lofqvist, A., Almqvist, B., & Sahlen, B. (2004). Is age at implant the only factor that counts? The influence of working memory on lexical and grammatical development in children with cochlear implants. *International Journal of Audiology, 43*(9), 506-515. doi: 10.1080/14992020400050065

7

Summary and Conclusions

Summary and Conclusions

When listening to speech, we constantly make implicit predictions. These predictions do not only concern *what* is going to be said but also about *how* it is going to be said: we may expect a Scotsman to speak differently from someone who grew up in London, we may expect shakiness in the voice of an older adult, and just by the look or by the name of someone we meet for the first time we may expect him or her to have a foreign accent. Often our predictions are met and we gain further support to strengthen our beliefs. Sometimes our predictions are incorrect. We are then presented with new information in the speech input that we can use to update our beliefs. In both situations, we are continuously learning: from the expected as well as from the unexpected.

In the current thesis, I investigated mechanisms that underlie this key ability to learn from experience and thereby ensure the flexibility of our speech processing system. I particularly focused on the concepts of statistical learning and perceptual learning in speech and how age and individual differences shape these learning processes. This Chapter summarizes and discusses the main findings of the dissertation. Broader implications of the findings for theoretical as well as clinical work will be discussed.

Summary of results

Perceptual learning, i.e., updating one's perceptual system by adjusting stored representations on the basis of currently perceived input, involves two components. First, the current input has to be matched onto the stored representations. Second, representations have to be established or updated on the basis of the input. These two processes are inevitably linked in tasks of perceptual learning. In Chapter 2, I aimed to investigate which perceptual and cognitive processes are particularly involved in matching two speech stimuli with each other in the auditory modality. Therefore, I looked at correlates of auditory discrimination performance. Auditory discrimination is the ability to hear whether two consecutive realizations of the same word or phrase are the same or different. This ability is different from perceptual learning with the latter focusing on the match between a realization and a representation stored in long-term memory. Nevertheless, auditory discrimination may serve as a window into the processes involved in successful matching of auditory information as it describes the ability to compare two auditory presented realizations in short-term memory. That is, a current speech stimulus has to be matched to the auditory memory trace of recently encountered speech input.

Furthermore, auditory discrimination is of interest in the broader context of human flexibility in speech as auditory discrimination ability is a prerequisite for changing our speech behavior. For example, patients in speech therapy programs, such as the Lee Silverman Voice Treatment (Sapir, Ramig, & Fox, 2011) or E-learning based Speech Therapy (Beijer et al., 2010), are frequently asked to compare their own realizations to target realizations (e.g., presented by the therapist). In doing so, adequate self-perception is key to attain favored modifications of speech behavior (Schroter-Morasch & Ziegler, 2005). As such, the ability to compare two consecutive utterances is crucial for adjusting our speech production.

My data on older adults' auditory discrimination performance indicated that auditory discrimination of speech stimuli is primarily associated with cognitive and linguistic skills, rather than auditory abilities. Discrimination accuracy was higher in older adults with better speech understanding in noise, faster processing speed, and better language proficiency, but accuracy decreased with age. Measures of processing efficiency (i.e., processing speed and a speeded language proficiency test) were best associated with older adults' ability to match two consecutive speech stimuli.

Chapter 3 investigated the contribution of modality, stimulus type and attention to statistical learning, thereby setting the stage for testing the relationship between perceptual learning in speech and statistical learning in Chapter 4. I specifically aimed to answer the questions whether statistical learning of temporal regularities is different for visual vs. auditory presentation, whether statistical learning differs depending on the linguistic or non-linguistic nature of the stimuli, and whether attentive processing of the stimuli is required for learning to occur. To that end, four variants of a serial reaction time – artificial grammar learning task were administered which differed as a function of modality (auditory or visual), stimulus type (linguistic or non-linguistic), and attention (well or no attention to the input required in order to perform the task). Stable statistical learning effects were observed in three out of four task variants.

In line with previous studies on modality-specific processing in statistical learning, stronger statistical learning was observed if information was presented auditorily. I showed that this auditory advantage is already present if the items on the basis of which learning occurred (i.e., visually presented nonwords) were the same in both the auditory and the visual nonword matching task. Regarding the question whether participants may be particularly sensitive to statistical properties in linguistic relative to nonlinguistic material, I found evidence to the contrary. That is, non-linguistic items (i.e., visual shapes) showed a processing advantage compared to linguistic items (i.e., written nonwords), suggesting that there are not only modality-specific processes but also stimulus-specific processes in statistical learning at play. Regarding the role of attention in statistical learning, differences in statistical learning were not observed between variants in which participants had to attentively process the target shapes to perform a cover task (i.e., a matching task in which targets were cued visually by smaller representation of the target shapes in the middle of the screen) and a task variant in which attention to the displayed shapes was irrelevant for the task at hand (i.e., a monitoring task in which the target was visually highlighted and, therefore, identification of the target shape itself was not necessary).

In Chapter 4, I aimed to investigate individual differences in statistical learning and perceptual learning with particular focus on the question whether individuals' ability to adapt to an unfamiliar speech condition can be predicted by a general ability to implicitly detect regularities. To prevent that a relationship between both measures of learning would be specific to auditory and linguistic processing, a visual statistical learning task was administered that made use of non-linguistic stimuli and that did not require attentive processing. This task variant had shown stable statistical learning effects in Chapter 3.

In line with previous studies on perceptual learning (Adank & Janse, 2010; Golomb, Peelle, & Wingfield, 2007; Gordon-Salant, Yeni-Komshian, Fitzgibbons, & Schurman, 2010; Peelle & Wingfield, 2005), younger and older listeners showed significant improvement in understanding noise-vocoded speech over exposure. However, only younger adults were sensitive to statistical regularities in the statistical learning task. In line with predictions of current models on perceptual learning (Ahissar & Hochstein, 2004; Kleinschmidt & Jaeger, 2015), differences in the amount of perceptual learning were indeed associated with individual sensitivity to probabilistic information in younger adults. Vocabulary knowledge was another predictor of the amount of adaptation younger adults showed over the course of exposure to noise-vocoded speech, thereby highlighting the involvement of listener-based abilities in perceptual learning.

For the research reported in Chapter 5, I designed a follow-up experiment on older adults' statistical learning ability as no stable statistical learning effects were observed in older adults in Chapter 4. My findings from Chapter 4 suggested that older adults' sensitivity to temporal regularities decreases if fast, sequential processing of visual stimuli is required. Previous reports of older adults' inability to learn probabilistic associations from visual input (e.g., Simon, Howard, & Howard, 2011) had been taken as evidence for a more general decrease in pattern sensitivity in older age (Negash, Howard, Japikse, & Howard, 2003). If older adults indeed have generally poorer pattern sensitivity than younger adults, then older adults' statistical learning performance should also be affected in a different (i.e., non-visual) modality. Given the importance of temporal regularities for language and speech processing, I aimed to address the question of whether younger and older adults also differ in auditory statistical learning. Results of the data analysis indicated no difference in auditory statistical learning performance between both age groups, thereby challenging the notion of a general age-related decline in pattern sensitivity.

In Chapter 6, I elaborated on the role of cognitive abilities in statistical learning and perceptual learning in a clinical population. Whereas Chapter 4 made use of the speech simulation of a cochlear implant [CI] to investigate speech adaptation during an experimental session, Chapter 6 was designed to investigate perceptual learning processes as observed in real life. To that end, nine postlingually deaf adults were tested prior to their cochlear implantation and followed during the first three months after implant activation. Overall, patients with faster processing speed, higher memory capacity and greater vocabulary knowledge tended to show steeper improvement in speech perception performance after implant activation.

Chapter 6 additionally investigated the use of more naturalistic speech material (i.e., conversational speech), as compared to more traditional material to measure speech perception performance (i.e., single words and simple sentences), for the assessment of speech understanding performance. A relationship between objective speech perception tasks and patients' quality of life was only observed for conversational speech material.

Theoretical implications

The findings of the current thesis have several implications for specific aspects of the adaptation process to unfamiliar speech. I will discuss implications for three theoretical domains. First, I will discuss the relationship between statistical learning and perceptual learning in speech in light of two theoretical models, namely the ideal listener framework (Kleinschmidt & Jaeger, 2015) and the Reverse Hierarchy Theory (Ahissar & Hochstein, 2004). Second, I will present ideas about the role of cognitive abilities in adaptation to unfamiliar speech input and how cognitive abilities may influence the perceptual learning process. Third, the current thesis investigated perceptual and statistical learning processes in younger as well as older adults. Therefore, results of this thesis shed light on learning processes over the life span.

Perceptual learning and statistical learning: two sides of the same coin?

In the recently developed ideal listener framework (Kleinschmidt & Jaeger, 2015), listeners' ability to derive and update the underlying distributions in speech cues from speech input (i.e., statistical learning) is indispensable to listeners' ability to adapt to an unfamiliar speech input such as foreign-accented or CI-simulated speech (i.e., perceptual learning). In this view, both types of learning may be considered to be two sides of the same coin in that perceptual learning is nothing else than a form of statistical learning. Supporting this assumption, implementation of functions to update a Bayesian model's beliefs about the underlying cue distributions (i.e., a statistical learning module) has resulted in a good fit between models' predicted and human listeners' actual performance (Kleinschmidt & Jaeger, 2015). However, listeners' performance in Kleinschmidt and Jaeger (2015) was only modelled on tasks of recalibration and selective adaptation and, hence, on perceptual learning tasks in which single isolated speech cues or phonemes are modified. In the current dissertation, I aimed to investigate the relationship between perceptual learning and statistical learning under more diverse conditions. If the two learning processes are two sides of the same coin, they should be interrelated under less controlled and more naturalistic conditions.

In Chapter 4, I indeed observed a relationship between individual differences in statistical learning ability and individual differences in perceptual learning of noise-vocoded speech. I applied a rigorous test that prevented that the observed relationship was specific for auditory or language processing. Furthermore, the observed relationship was not mediated by any of the tested cognitive and linguistic abilities (i.e., indices of working memory, processing speed, attention switching control and vocabulary knowledge). This implies that perceptual learning may indeed be a form of statistical learning or that perceptual learning may directly rely on statistical learning processes. Thus, my findings from Chapter 4 underline the importance of statistical learning for perceptual learning.

Inconsistent with the notion that perceptual learning and statistical learning are essentially the same, I found no indication that novel postlingually deaf cochlear implant users engage statistical learning processes to adapt to the speech input provided by their new implant (cf. Chapter 6). Note, however, that no significant statistical learning effect could be observed across the group of CI patients. Considering the small group of nine novel CI users, the sample size was probably too small to detect the rather subtle statistical learning effects in the visual variants of the statistical learning task (cf. Chapters 3 and 4). Moreover, the group of CI patients mainly consisted of participants older than 50 years. As we had found no statistical learning effect in a larger sample of 60 normal-hearing older adults either (Chapter 4), the non-significant statistical learning effect in the sample of CI patients may provide further evidence that there may be age-related declines in the visual processing of temporal regularities. Importantly, the failure to observe statistical learning in the current patient sample should neither be taken as evidence that novel CI users are generally insensitive to probabilistic information in the input, nor that probabilistic information in the input is generally unimportant for adaptation to a cochlear implant. In hearing-impaired children, sequence learning ability has repeatedly been found to predict CI outcome and, therefore, been put forward as one of the underlying processes explaining CI success in children (for an overview see Pisoni, Kronenberger, Chandramouli, & Conway, 2016). Obviously, further research is required to investigate possible links between statistical and perceptual learning in a setting where postlingually deaf patients do show both types of learning.

In the current thesis, I only applied one task of statistical learning performance, namely the artificial-grammar-learning serial-reaction-time paradigm presented in different variants. As is evident from my findings in Chapter 3 and from recent research (e.g., Conway & Christiansen, 2005, 2006; Frost, Armstrong, Siegelman, & Christiansen, 2015), statistical learning performance is influenced by modality and stimulus-specific effects. Moreover, statistical learning performance on one task does not necessarily correlate with performance on other tasks of statistical learning (Siegelman & Frost, 2015). This implies that statistical learning is not a unified capacity and that different statistical learning tasks may tap into different components of statistical learning ability (for an elaborate discussion see Siegelman, Bogaerts, Christiansen, & Frost, 2017). It is unlikely that the selected artificial-grammar-learning serial-reaction-time paradigm represents the complete underlying construct of statistical learning. As such, performance on a single statistical learning task may not necessarily predict perceptual learning equally well across conditions and populations. The current thesis particularly provides evidence that the ability to pick up on adjacent co-occurrence frequencies is involved in listeners' ability to perceptually learn and that this relationship is independent of other cognitive abilities.

Although I showed that statistical learning is an important process underlying perceptual learning success, I also showed that it is not the only ability perceptual leaning draws upon. Throughout different studies in this thesis (cf., Chapter 2, Chapter 4 and Chapter 6), particularly measures of processing speed and vocabulary knowledge emerged as correlates of auditory discrimination performance and of improvement in speech understanding performance over time when listeners were exposed to unfamiliar speech input. As I described earlier, perceptual learning includes a matching and an updating component. I assume that the matching component is mediated by the tested cognitive abilities (for an elaborate rationale see the later Discussion section on the role of cognitive abilities in perceptual learning): processing speed and linguistic knowledge may aid matching in perceptual learning by allowing for efficient online processing of the auditory input and for efficient access to stored representations. Moreover, higher working memory may support the matching process under taxing listening conditions by providing more spare capacity for online speech processing when listening is effortful. Note that working memory might also be expected to play a role in the updating component of perceptual learning given that updating has been described as the ability to simultaneously store and actively manipulate information in working memory (Miyake et al., 2000). However, updating in perceptual learning requires a system to continuously adjust representations in long-term memory by observations from the current input. Whereas recent findings suggest that speech segmentation by statistical learning is supported by processing time and domain-general attentional or working memory capacity (Palmer & Mattys, 2016), I found little to no evidence that working memory is linked to perceptual learning performance or to statistical learning performance. This indicates that individuals' working memory capacity does not play a substantial role in the general ability to quickly adapt to unfamiliar speech input and, hence, contradicts the assumption that working memory is driving the updating of representations in perceptual learning.

As Kleinschmidt and Jaeger (2015) argue, the updating process in perceptual learning has to be incremental in speech processing as speech unfolds over time. Statistical learning is such an incremental learning mechanism as statistical probabilities are provided and continuously adjusted by the input. That is, with each new observation the stored distribution is shifted slightly. In the artificial-grammar-learning serial-reaction-time paradigm, which I applied as measure of statistical learning performance throughout this thesis, each trial in the exposure phase provided additional evidence that specific items co-occurred. As expected for an updating mechanism, participants' got faster in clicking on the second, predictable item within a trial and, thus, started to anticipate upcoming targets on the basis of the underlying transitional probabilities over exposure. Thus, participants improved their task performance by representing this distributional knowledge. That this improvement over exposure was indeed a consequence of updated representations became evident in the test phase in which participants showed a drop in performance as the learned distributions no longer applied. Being sensitive to distributional information thus enables language users to update their stored representations and, hence, their prior beliefs, in perceptual learning for example about what words or sounds are supposed to sound like in a certain setting. I would, therefore, like to refine Kleinschmidt and Jaegers' view (2015) on perceptual learning being statistical learning by hypothesizing that statistical learning constitutes the updating component in perceptual learning.

As mentioned above, one limitation of the ideal listener framework (Kleinschmidt & Jaeger, 2015) is that it is based on data from perceptual learning tasks in which only single phonemes or isolated speech cue distributions are altered. The reverse hierarchy theory [RHT] may be instrumental in explaining how statistical learning may apply to more global speech alterations. The RHT suggests that updating does not take place on a single level (e.g., a single speech cue distribution) but that updating the stored representations proceeds from higher-level representations (e.g., words, syllables) to lower level representations (e.g., biphonemic clusters, phonemes, speech cues). By implanting the reweighting component of the RHT as the updating component in the ideal listener framework, the ideal listener framework can be expanded to include additional processing layers. Within each processing layer statistical learning is key to adjust the stored distributions. This may be the likelihood with which specific words co-occur on the word level or the likelihood how acoustic speech cues are produced on the speech cue level. At the same time, the information that is transferred between levels of the hierarchy may also be probabilistic in nature. That is, neuronal network models that rely on probabilistic inferences for providing information from lower to higher hierarchical levels explain neurophysiological changes in early sensory areas in visual perceptual learning tasks that cannot be accounted for by other models (Bejjanki et al., 2011).

The role of cognitive abilities in perceptual learning

Individuals do not perform equally well on perceptual tasks. Listeners' vary greatly in their speech perception performance and also in their ability to learn to understand unfamiliar speech. What makes someone a good adapter? In this thesis, I partly accounted for the natural variability in listeners' perceptual learning performance by linking it to listeners' statistical learning capacity. Though my findings underline the importance of statistical learning for perceptual learning, my data suggest that statistical learning is not the only cognitive ability involved in successful adaptation to unfamiliar speech input. In this section, I aim to bring evidence from the previous Chapters together to identify which cognitive resources listeners may engage at which processing stage to recognize and to adapt to an unfamiliar speech signal.

As described earlier in this Chapter, perceptual learning consists of two components. First, the current input has to be matched onto stored representations. Second, the stored representations are updated based on the input. In Chapter 2, auditory discrimination performance and, hence, the ability to match current auditory input to a recent representation in short-term memory, was associated with factors of processing efficiency and linguistic ability and in particular with processing speed. In subsequent studies, however, measures of processing speed did not predict perceptual learning success. I, therefore, hypothesize that processing speed might particularly be involved in listeners' ability to initially process an unfamiliar auditory input for the matching component, be it the matching of two subsequent realizations to each other, or the matching of incoming input onto stored representations.

Given the transient nature of spoken language, less efficient processors are likely to miss information in an unfamiliar speech signal comparable to a too low sampling rate for audio recording. Therefore, individuals with lower processing speed may have less information available for being able to match the current speech input onto their stored representations than their more efficient peers. This may result in a larger selected set of possible matches and, hence, in more uncertainty concerning the match. In line with this assumption, I found that processing speed predicted older adults' initial speech recognition level but not their improvement in understanding noise-vocoded speech in Chapter 4. Hence, starting level performance on the noise-vocoded speech was worse for those with poorer sampling of the message. Moreover, it follows from this line of thinking that individuals with slower processing speed may initially need auditory input which contains more information to allow for the same amount of matching and, consequentially, for the same amount of learning than faster individuals. Indeed, older adults, who as a group are argued to suffer from age-related cognitive slowing (Salthouse, 2009a, 2009b, 2010) have been shown to need a higher starting level (i.e., more identifiable information in the beginning) than younger adults to exhibit the same amount of perceptual learning (e.g., Peelle & Wingfield, 2005).

An important cognitive resource that emerged as correlate of perceptual learning across several studies reported in this thesis was linguistic ability. Linguistic ability was associated with auditory discrimination ability (Chapter 2) and predicted perceptual learning performance in both normal-hearing (Chapter 4) and hearing-impaired adults (Chapter 6). These observations confirm the link between increased lexical knowledge and success in perceptual learning that has previously been reported in perception (Banks, Gowen, Munro, & Adank, 2015; Bent, Baese-Berk, Borrie, & McKee, 2016; Borrie, Lansford, & Barrett, 2017) and in adaptation to unfamiliar speech (Baese-Berk, Bent, Borrie, & McKee, 2015; Janse & Adank, 2012). More importantly, the current thesis shows that lexical knowledge is not only associated with adaptation to linguistic degradations, e.g., systematic phonological deviations in how a foreign-accented speaker pronounces words, but also relates to perceptual learning of acoustically degraded speech (i.e., noise-vocoded speech and the speech signal transmitted by a CI).

As individuals with higher scores on vocabulary tests also show better performance on measures of verbal fluency (e.g., Hedden, Lautenschlager, & Park, 2005; Kemper & Sumner, 2001), individuals with greater vocabulary knowledge may be considered more efficient processors of linguistic information (Kemper & Sumner, 2001). Whereas processing speed may be particularly involved in the efficient online processing of the incoming auditory signal, linguistic knowledge may specifically represent facilitated access to stored representations and, thereby, aid the matching component of perceptual learning. This may best be illustrated by the so-called pop-out effect (Davis, Johnsrude, Hervais-Adelman, Taylor, & McGettigan, 2005): if listeners know the content of what is going to be said before they actually hear a sentence in its degraded form, this boosts their immediate speech recognition and, more importantly, benefits their subsequent perceptual learning. That is, by providing listeners with access to the higher-level representations of the upcoming speech (e.g., words), the ongoing unfamiliar speech input can easily be matched onto the stored representations. This, in turn, enables and guides top-down search processes for sublexical retuning (Ahissar & Hochstein, 2004). Recognition of subsequent speech input is then improved (compared to previous exposure without lexical information) as perception can additionally be based on updated lower level representations.

Working memory, which is the cognitive ability most commonly associated with language processing, only correlated with perceptual learning performance in novel, severely hearing-impaired CI users (Chapter 6), and not with perceptual learning in normal-hearing younger adults or older adults with mild hearing loss. This discrepancy may be explained by differences in listening demands. That is, working memory capacity is supposed to play a key role in speech understanding if the listening condition is taxing (Rönnberg, Rudner, Foo, & Lunner, 2008; Rönnberg, Rudner, Lunner, & Zekveld, 2010). Still, given that perception of noise-vocoded speech has also been linked to increased cognitive demands for listening (e.g., Başkent, 2012; Chatterjee, Peredo, Nelson, & Başkent, 2010; Pals, Sarampalis, & Başkent, 2013), why did working memory not predict perceptual learning performance of noise-vocoded speech for our non-clinical groups (Chapter 4)? I speculate that this finding may be due to group differences in the duration of exposure to adverse listening. In normal-hearing or mildly hearing-impaired adults, perceptual learning tasks that span only a couple of minutes are unlikely to evoke similar listening effort as in novel CI users. After all, novel CI users experienced an extensive period of auditory deprivation before implant activation and, therefore, have to expend effort just in picking up auditory information. Accordingly, studies on CI outcome often report fatigue due to listening effort as a frequent complain in CI users (Hughes & Galvin, 2013). In line with this assumption that working memory is of particular importance for more taxing listening conditions, working memory has been shown to aid speech recognition in noise in hearing-impaired (e.g., Foo, Rudner, Ronnberg, & Lunner, 2007; Rudner, Foo, Ronnberg, & Lunner, 2009; Rudner, Foo, Sundewall-Thoren, Lunner, & Ronnberg, 2008; Rudner, Ronnberg, & Lunner, 2011) but not in normal-hearing adults (Füllgrabe, Moore, & Stone, 2014; Füllgrabe & Rosen, 2016a, 2016b; Schoof & Rosen, 2014). Therefore, individual differences in working memory may not always predict speech perception performance under challenging listening conditions. Rather, the involvement of working memory seems to depend on the hearing status and the age of the listener (Füllgrabe & Rosen, 2016b). Thus, working memory may modulate perceptual learning performance if listening is particularly effortful as is the case in older, hearing-impaired populations. Individual differences in cognitive capacity will then determine how much effort it requires listeners to process speech in otherwise equal listening conditions, and conversely, how much spare capacity is available for matching the speech input onto stored representations.

In sum, I showed that individual cognitive abilities can partly explain why some listeners are better perceptual learners than others. I propose that processing speed is involved in efficient processing of the incoming unfamiliar speech input for matching and that linguistic ability aids perceptual learning by facilitating access to the stored representations for matching. Working memory may only be associated with perceptual learning in case older, hearing-impaired individuals are tested for whom listening conditions are particularly taxing and effortful for a prolonged period of time.

Learning over the life span

If we learn from exposure, for example by listening to unfamiliar speech input, learning proceeds without our intent or conscious awareness (Kaufman et al., 2010). We cannot communicate the changes that take place and that enable us to improve in perception. Therefore, perceptual learning and statistical learning are both considered implicit learning processes. In contrast to explicit abilities such as working memory, implicit learning abilities have been argued to stay rather intact throughout the life span (e.g., Midford & Kirsner, 2005). In line with this, several studies report that older adults remain sensitive to probabilistic sequences (Campbell, Zimerman, Healey, Lee, & Hasher, 2012; Negash et al., 2003; Salthouse, McGuthry, & Hambrick, 1999; Simon et al., 2011) and that older adults are able to improve in understanding novel speech input over exposure (Adank & Janse, 2010; Golomb et al., 2007; Gordon-Salant et al., 2010; Peelle & Wingfield, 2005). However, many studies also report age-related changes in implicit learning. On the one hand, studies report age-related declines such as less learning under cognitive load (Vandenbossche, Coomans, Homblé, & Deroost, 2014), less learning with increased complexity of the task (Bennett, Howard, & Howard, 2007) and less learning with longer training (Adank & Janse, 2010; Peelle & Wingfield, 2005; Simon, Vaidya, Howard, & Howard, 2012). On the other hand, older adults have been reported to show more learning from non-attended input than younger adults (Campbell et al., 2012). To investigate possible age-related changes in implicit learning in this thesis, I tested statistical learning and perceptual learning in both younger and older adults and investigated whether these learning abilities are affected by age.

Regarding statistical learning, I found older adults to pick up on temporal regularities equally well as younger adults (Chapter 5). If the same learning paradigm was administered in the visual modality, however, older adults were unable to detect these regularities (Chapter 4). This suggests that sensitivity to statistical patterns is generally preserved in older adults but that older adults suffer from modality-specific deficits in processing of sequential information. This observation is in line with previous studies indicating advantages of auditory learning for temporal regularities and advantages of visual learning for spatial regularities (e.g., Conway & Christiansen, 2005). Thus, older adults seem to experience a loss of sensitivity to statistical properties in the modality that is less specialized for a given input. Arguably, this loss should not be considered a result of cognitive decline with aging but a result of learning. That is, based on their life-long experience, older adults have gained increased knowledge about informative processing routes (e.g., auditory processing for temporal information) and, therefore, become less sensitive to less informative processing routes (e.g., processing temporal information via the visual modality).

Regarding perceptual learning, I showed that both younger and older listeners are able to significantly improve in understanding noise-vocoded speech over exposure (cf. Chapter 4). The observed amount of learning was comparable in both age groups given age-matched levels of difficulty of noise-vocoded speech. Thus, older adults reached the same amount of perceptual learning as younger adults given better starting level intelligibility. Similarly, in Chapter 6, age was negatively correlated with novel CI users' initial speech understanding performance but not with their improvement in understanding the unfamiliar speech input over time. Considering the hypothesized twofold nature of perceptual learning (i.e., first matching to and then updating representations), I hypothesize that listeners' ability to establish and update speech representations remains stable over the life span but that older adults particularly experience difficulties in matching unfamiliar speech input onto stored representations.

There are two possible explanations why older adults may be disadvantaged in matching novel speech input onto stored representations. First, older adults are less efficient processors than younger adults due to general perceptual and cognitive slowing (e.g., Salthouse, 2009a, 2009b, 2010). As argued above concerning the role of processing speed in perceptual learning, older adults are therefore likely to miss information in an unfamiliar speech signal and probably experience difficulties in matching the current speech input onto their stored representations. Second, it may be assumed that older adults possess about much more linguistic knowledge than younger adults due to their life-long exposure to language. Consequentially, older adults may have more difficulties in matching a current speech input to their stored representations as more lexical candidates are available. In this view, older adults' difficulties in matching are caused by a higher information-processing cost due to increased knowledge (Ramscar, Hendrix, Shaoul, Milin, & Baayen, 2014). Note, however, that in younger adults and in hearing-impaired middle-aged adults, I found higher vocabulary knowledge to predict better not worse perceptual learning performance. This implies that the first account on decreased processing efficiency with older age explaining age-related differences in perceptual learning seems more plausible than the second account on higher processing cost due to increased knowledge.

In sum, findings from this dissertation support the idea that both statistical learning and perceptual learning are generally preserved over the life span. However, the current thesis also indicates changes in implicit learning due to aging. Older adults seem to be less efficient in perceptual learning as they may only reach the same amount of learning as younger adults given higher starting level intelligibility for the older adults. Furthermore, older adults seem to be less efficient in processing statistical patterns via the less specialized modality as indicated by a deficit to detect temporal regularities in visual input.

Clinical implications

Perceptual learning is an essential process in hearing rehabilitation as patients have to adapt to unfamiliar speech input when they are provided with a new hearing device. Moreover, perceptual learning processes in patients' relatives and caregivers may support communication with patients whose speech is characterized by reduced intelligibility. Therefore, clinical implications of the current work for the rehabilitation of hearing and speech disorders will be discussed in the two subsequent sections.

Implications for hearing rehabilitation

Since a couple of years, individual differences in patients' performance gain increasing attention in rehabilitation practice. This trend manifested 2001 in the publication and the official endorsement of the International Classification of Functioning Disability and Health [ICF] (World Health Organization, 2001). In contrast to former health classification systems, the ICF aims to describe persons in their entity with particular focus on an individual perspective. What has led to this perspective change in rehabilitation? First, investigating individual differences in patients' performance may help to identify factors aggravating given impairments such as longer duration of severe hearing loss limiting CI outcome (Blamey et al., 2013). Second, even if patients suffer from the exact same impairment, e.g., their hearing impairment being equally severe, patients still perceive their conditions differently depending on for example their environment and their social network. Therefore, treatment should ideally be individualized to patients' needs and capabilities to achieve the best possible outcome.

In Chapter 6 of this thesis, I looked into individual abilities that may be involved in the ability of hearing-impaired persons to adapt to a new hearing device. Such adaptation for successful communication is important to preserve independence and self-reliance in older adults (Di Nardo, Anzivino, Giannantonio, Schinaia, & Paludetti, 2014). Hearing loss and, hence, reduced ability to adequately perceive speech has been shown to reduce quality of life due to communicative, emotional and social limitations (e.g., Hogan, O'Loughlin, Davis, & Kendig, 2009). Treatment with hearing devices such as cochlear implants have been shown to increase quality of life again (Contrera et al., 2016; Damen, Beynon, Krabbe, Mulder, & Mylanus, 2007; Hinderink, Krabbe, & Van Den Broek, 2000; Klop et al., 2008) and to reduce depressive symptoms (Choi, Betz, Li, & et al., 2016). However, the speech perception benefit patients experience from a cochlear implant varies considerably between individuals (e.g., Heydebrand, Hale, Potts, Gotter, & Skinner, 2007).

The data described in this thesis suggest that individual cognitive abilities are promising in explaining individuals' variability in adapting to the new type of speech input provided by a cochlear implant. Importantly, individual differences in listeners' linguistic abilities did not only predict performance in short-term perceptual learning, which took place over the course of a single experimental task and made use of CI simulated speech, but were also associated with patients' improvement in speech understanding performance in the weeks following cochlear implantation. Therefore, the current findings underline that similar processes are involved in short-term perceptual learning processes in normal-hearing and in long-term perceptual learning processes in hearing-impaired populations. Speech adaptation success with a CI was mainly associated with memory capacity, vocabulary knowledge and processing speed (Chapter 6). This indicates that differences in the ability to rapidly encode and process auditory information may underlie patients' adaptation success with a CI.

How may findings from the current thesis inform clinical practice in hearing rehabilitation? First, knowledge about pre-implantation abilities such as memory capacity, vocabulary knowledge and processing speed may help to shape realistic expectations regarding speech perception progress and outcome after implantation. Therapists may be able to provide patients with more individualized prospects on the probable outcome, preparing patients for more or less effortful adaptation trajectories, or pointing out the importance of extensive and regular CI use if needed. Furthermore, monitoring and keeping patients' motivation up (e.g., by individually extending the rehabilitation program by additional counselling from a psychologist) may be particularly indicated if patients are expected to experience less progress in speech perception after implantation.

Second, it seems to make sense to facilitate the adaptation process by training the identified cognitive capacities such as for example offering working memory training. Working memory training may have an additional long-term benefit by helping CI users to cope with listening effort. Thereby, memory training may prevent or diminish detrimental effects sustained effortful listening has on individuals' lives such as mental fatigue (Hornsby, 2013) or stress-related sick leave from work (Kramer, Kapteyn, & Houtgast, 2006). However, recent studies offering working memory training to older hearing-aid users (Ferguson & Henshaw, 2015) and older individuals with varying degrees of age-related hearing loss (Wayne, Hamilton, Jones Huyck, & Johnsrude, 2016) have failed to show transfer effects of memory training to speech perception performance. This suggests that working memory may not be feasible to improve patients' initial adaptation process. In contrast, auditory training and auditory-based cognitive training programs seem to be more effective for enhancing working memory capacities, processing speed, speech perception performance and self-reported benefit (Anderson, White-Schwoch, Choi, & Kraus, 2013; Anderson, White-Schwoch, Parbery-Clark, & Kraus, 2013; Ferguson, Henshaw, Clark, & Moore, 2014; Henshaw & Ferguson, 2014; Smith et al., 2009; Sweetow & Sabes, 2006). Yet, it is unclear whether such training programs are also effective for novice CI users. Three weeks of auditory training increased consonant discrimination abilities in a small group of experienced CI users, but did not enhance discrimination of vowels or words in sentences (Stacey et al., 2010). Further research is needed to evaluate whether extending the rehabilitation program by auditory-cognitive training may help to facilitate adaptation to a CI, particularly in patients with poorer memory skills.

Third, the current thesis highlights the potential of more naturalistic speech material for clinical practice in audiological rehabilitation. That is, in Chapter 6, I adapted conversational speech material for audiological testing in Dutch, thereby following recent approaches to implement more variable and naturalistic speech material in audiological practice in English (Gilbert, Tamati, & Pisoni, 2013; King, Firszt, Reeder, Holden, & Strube, 2012; Tamati, Gilbert, & Pisoni, 2013). Importantly, the study described in Chapter 6 is the first to report on the relationship between perception of conversational speech material and perceived cochlear implant benefit, whereas this relationship was absent for more traditional materials. Objective speech perception tests using more naturalistic speech material may therefore be beneficial for clinical practice as they may be used more generally, e.g., as training material, as diagnostic assessment of rehabilitation outcome, and to customize and evaluate implant settings in the lab under conditions that reflect patients' daily experience. More naturalistic speech perception tests may also offer a vantage point for further studies exploring the perceptual and neurocognitive underpinnings of successful cochlear implant adaptation.

Implications for the treatment of speech disorders

The work described in the current thesis also has implications for the treatment of speech disorders. In contrast to hearing rehabilitation, in which the patient undergoes perceptual learning processes to improve speech perception, treatment of speech disorders may benefit from perceptual learning in patients' relatives or caregivers for enhancing communication. Patients with speech disorders such as dysarthria are difficult to understand as the motor control of the muscles that are involved in speech production is affected. Dysarthric speech is characterized by abnormal movements of the articulators, voice quality, rhythm, pitch and speech rate resulting in the impression of slurred and mumbled speech (American Speech-Language-Hearing Association). Compared to noise-vocoded or foreign-accented speech, dysarthric speech deviates in a rather unsystematic way from standard speech production. Nevertheless, exposure to short passages of dysarthric speech has been shown to improve listeners' recognition of dysarthric speech (Borrie, McAuliffe, & Liss, 2012; Borrie, McAuliffe, Liss, Kirk, et al., 2012; Borrie, McAuliffe, Liss, O'Beirne, & Anderson, 2012, 2013; Kim & Nanney, 2014).

Comparable to studies on foreign-accented and acoustically degraded speech, listeners have been shown to vary greatly in their perception (e.g., Borrie, 2015) as well as in their improvement to understand dysarthric speech (e.g., Bent et al., 2016; Borrie et al., 2017). Given that individuals who are better able to understand dysarthric speech also show better recognition of foreign-accented speech (Baese-Berk et al., 2015; Bent et al., 2016), it may be assumed that specific skills of individual listeners aid recognition and perceptual learning of unfamiliar speech irrespective of the type of speech degradation. Indeed, vocabulary knowledge, which I found to consistently predict speech adaptation success in the current thesis, has repeatedly been reported to predict listeners' ability to perceive (e.g., Bent et al., 2016; Borrie et al., 2017) as well as to improve in the perception of dysarthric speech (Baese-Berk et al., 2015). The important role of lexical knowledge for perceptual learning of dysarthric speech is also apparent from studies in which orthographic representations of the lexical content of dysarthric speech input have been shown to accelerate perceptual learning compared to auditory exposure only (Borrie, McAuliffe, Liss, Kirk, et al., 2012; Kim & Nanney, 2014). It may therefore be helpful to ask patients to read short passages of text to their caregivers, who are also provided with a copy of the text, to immediately enhance speech understanding. Dysarthric patients may also be advised to start dialogues by addressing the topic of what they are going to say (e.g., grocery, lunch, weather) to facilitate listeners' access to relevant representations of lexical items in long-term memory.

Listeners' ability to adapt to dysarthric speech offers possibilities for listener-centered rather than more traditional patient-centered approaches in the treatment of speech disorders. That is, in addition to teaching patients to improve the intelligibility of their speech, relatives and caregivers of patients may be trained to better understand the patients' speech. Listener-centered approaches may be particularly valuable in the treatment of dysarthria as cognitive impairments are relatively frequently observed in conditions that cause dysarthria such as Parkinson disease (Aarsland et al., 2010), amyotrophic lateral sclerosis (Ringholz et al., 2005; Rippon et al., 2006), multiple sclerosis (Chiaravalloti & DeLuca, 2008) and stroke (Douiri, Rudd, & Wolfe, 2013). Due to this prevalence of cognitive decline in dysarthria, some dysarthric patients may be limited in their capacity to learn and to improve their speech production through speech therapy. Furthermore, dysarthria is often caused by progressive diseases in which patients' general condition gradually decreases and with it their speech intelligibility over time. Therefore, improving relatives' and caregivers' understanding of the deviant speech may be of additional help to preserve successful communication. In the end, maintenance and improvement of communication is the ultimate goal of speech and language therapy.

In addition to implications for listener-centered approaches in the rehabilitation of speech disorders, the current thesis also provides implications for patient-centered treatment of dysarthria. In Chapter

2, I investigated correlates of auditory discrimination performance. Auditory discrimination of speech stimuli is an essential component of speech and language therapy, as patients are frequently asked to compare their own realizations with target speech, which is presented either by the therapist or by a speaker whose speech was recorded for that purpose. This is particularly the case in intervention programs for the treatment of dysarthria such as the Lee Silverman Voice Treatment (Sapir et al., 2011) or E-learning based Speech Therapy (Beijer et al., 2010).

My finding that perception of acoustic differences in speech is more related to cognitive and linguistic skills than to auditory abilities has some important clinical implications. First, my result suggests that individuals' sensory functioning does not determine whether discrimination-based approaches in speech and language therapy will be suitable for patients. Second, the associations between cognitive and linguistic skills on the one hand and auditory discrimination ability on the other, observed in a quiet non-distracting setting, may become even more prominent under less ideal listening conditions, such as performing auditory discrimination tasks in a clinical setting. In a clinical setting, patients are commonly asked to produce and compare utterances simultaneously, and patients often have to analyze more than one speech dimension at the same time. In order to help patients focus their attention, the use of recording facilities may be recommended to record patients' speech during therapy. By doing so, I hypothesize that patients can first concentrate on producing the target utterances and, thereby, familiarize with the linguistic content. In a subsequent step, the speech and language therapist can ask patients to compare their own speech with the target speech. This approach is similar to the procedure implemented in E-learning based Speech Therapy (Beijer et al., 2010). Moreover, rehabilitation may be more effective if patients are asked to concentrate on just a single speech dimension. As the results of Chapter 2 showed fewer errors in discriminating intensity than in discriminating both pitch and speech rate, this suggests that listeners are most familiar with the loudness dimension, which can therefore best be used as the dimension to

familiarize patients with discrimination tasks.
The auditory discrimination test described in Chapter 2 was developed to assess whether individuals with dysarthria qualify for E-learning based Speech Therapy. My findings thus underline the notion that it is crucial to take patient abilities into account when designing and implementing eHealth services. That is, patient abilities form a key human factor in utilization and acceptance of telemedicine programs (Brennan & Barker, 2008). At the same time, the question arises whether the observed pattern of associations generalizes to clinical populations and if so, whether patients with better cognitive and linguistic performance and, hence better auditory discrimination skills, may benefit more from discrimination-based approaches of speech and language therapy than patients with poorer cognitive and linguistic performance. These aspects were not explored in the current thesis but may be considered for future clinical research.

Conclusion and closing remarks

Statistical learning and perceptual learning in speech are key to stable speech understanding. As human listeners we are continuously engaged in these learning processes to improve our perception of unfamiliar input by merely listening to it. The current series of studies has provided insights into the nature of adaptation to unfamiliar speech input, with a particular focus on adaptation to cochlear implants and to CI-simulated speech. In doing so, this thesis provides a valuable contribution to the area of psycholinguistic research with clinical implications for the fields of hearing rehabilitation and speech and language therapy. The research described in this thesis has highlighted that individual differences in listeners' ability to improve in the perception of unfamiliar speech input can partially be explained by listeners' cognitive and linguistic abilities. Thereby, the current thesis emphasizes the role of individual differences in speech processing and highlights the possibilities of both individualized patient-centered and listener-based approaches to rehabilitation. The current thesis suggests that statistical learning, processing speed and vocabulary knowledge predict successful perceptual learning in speech. Moreover, the research in this thesis adds to the study of speech processing over the life span indicating that both perceptual learning and statistical learning are generally preserved over the life span but that learning may be less effective in older adults under certain circumstances.

References

Aarsland, D., Bronnick, K., Williams-Gray, C., Weintraub, D., Marder, K., Kulisevsky, J., . . . Emre, M. (2010). Mild cognitive impairment in Parkinson disease: a multicenter pooled analysis. *Neurology, 75*(12), 1062-1069. doi: 10.1212/WNL.0b013e3181f39d0e

Adank, P., & Janse, E. (2010). Comprehension of a novel accent by young and older listeners. *Psychology and Aging, 25*, 736-740. doi: 10.1037/a0020054

Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences, 8*, 457-464. doi: 10.1016/j.tics.2004.08.011

American Speech-Language-Hearing Association. Dysarthria Retrieved 20.04.2017, from http://www.asha.org/public/speech/disorders/dysarthria/

Anderson, S., White-Schwoch, T., Choi, H. J., & Kraus, N. (2013). Training changes processing of speech cues in older adults with hearing loss. *Frontiers in Systems Neuroscience, 7*, 97. doi: 10.3389/fnsys.2013.00097

Anderson, S., White-Schwoch, T., Parbery-Clark, A., & Kraus, N. (2013). Reversal of age-related neural timing delays with training. *Proceedings of the National Academy of Sciences of the United States of America, 110*(11), 4357-4362. doi: 10.1073/pnas.1213555110

Baese-Berk, M., Bent, T., Borrie, S. A., & McKee, M. (2015). *Individual differences in perception of unfamiliar speech.* Paper presented at the 18th International Congress of Phonetic Sciences [ICPhS 2015], Glasgow.

Banks, B., Gowen, E., Munro, K. J., & Adank, P. (2015). Cognitive predictors of perceptual adaptation to accented speech. *The Journal of the Acoustical Society of America, 137*(4), 2015-2024. doi: 10.1121/1.4916265

Başkent, D. (2012). Effect of speech degradation on top-down repair: phonemic restoration with simulations of cochlear implants and combined electric-acoustic stimulation. *Journal of the Association for Research in Otolaryngology, 13*(5), 683-692. doi: 10.1007/s10162-012-0334-3

Beijer, L. J., Rietveld, A. C. M., van Beers, M. M. A., Slangen, R. M. L., van den Heuvel, H., de Swart, B. J. M., & Geurts, A. C. H. (2010). E-learning-based speech therapy: a web application for speech training. *Telemedicine Journal and E-Health, 16*, 177-180. doi: 10.1089/tmj.2009.0104

Bennett, I. J., Howard, J. H., & Howard, D. V. (2007). Age-related differences in implicit learning of subtle third-order sequential structure. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences, 62*, P98-103. doi: 10.1093/geronb/62.2.P98

Bent, T., Baese-Berk, M., Borrie, S. A., & McKee, M. (2016). Individual differences in the perception of regional, nonnative, and disordered speech varieties. *The Journal of the Acoustical Society of America, 140*(5), 3775. doi: 10.1121/1.4966677

Blamey, P., Artieres, F., Başkent, D., Bergeron, F., Beynon, A., Burke, E., . . . Lazard, D. S. (2013). Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants: an update with 2251 patients. *Audiology and Neurotology, 18*(1), 36-47. doi: 10.1159/000343189

Borrie, S. A. (2015). Visual speech information: a help or hindrance in perceptual processing of dysarthric speech. *The Journal of the Acoustical Society of America, 137*(3), 1473-1480. doi: 10.1121/1.4913770

Borrie, S. A., Lansford, K. L., & Barrett, T. S. (2017). Rhythm perception and its role in perception and learning of dysrhythmic speech. *Journal of Speech, Language, and Hearing Research, 60*(3), 561-570. doi: 10.1044/2016_JSLHR-S-16-0094

Borrie, S. A., McAuliffe, M. J., & Liss, J. M. (2012). Perceptual learning of dysarthric speech: a review of experimental studies. *Journal of Speech, Language, and Hearing Research, 55*, 290-305. doi: 10.1044/1092-4388(2011/10-0349)

Borrie, S. A., McAuliffe, M. J., Liss, J. M., Kirk, C., O'Beirne, G. A., & Anderson, T. (2012). Familiarisation conditions and the mechanisms that underlie improved recognition of dysarthric speech. *Language and Cognitive Processes, 27*(7-8), 1039-1055. doi: 10.1080/01690965.2011.610596

Borrie, S. A., McAuliffe, M. J., Liss, J. M., O'Beirne, G. A., & Anderson, T. J. (2012). A follow-up investigation into the mechanisms that underlie improved recognition of dysarthric speech. *The Journal of the Acoustical Society of America, 132*, EL102-108. doi: 10.1121/1.4736952

Borrie, S. A., McAuliffe, M. J., Liss, J. M., O'Beirne, G. A., & Anderson, T. J. (2013). The role of linguistic and indexical information in improved recognition of dysarthric speech. *The Journal of the Acoustical Society of America, 133*, 474-482. doi: 10.1121/1.4770239

Brennan, D. M., & Barker, L. M. (2008). Human factors in the development and implementation of telerehabilitation systems. *Journal of Telemedicine and Telecare, 14*(2), 55-58. doi: 10.1258/jtt.2007.007040

Campbell, K. L., Zimerman, S., Healey, M. K., Lee, M. M. S., & Hasher, L. (2012). Age differences in visual statistical learning. *Psychology and Aging, 27*, 650-656. doi: 10.1037/a0026780

Chatterjee, M., Peredo, F., Nelson, D., & Başkent, D. (2010). Recognition of interrupted sentences under conditions of spectral degradation. *The Journal of the Acoustical Society of America, 127*(2), EL37-EL41. doi: 10.1121/1.3284544

Chiaravalloti, N. D., & DeLuca, J. (2008). Cognitive impairment in multiple sclerosis. *The Lancet. Neurology, 7*(12), 1139-1151. doi: 10.1016/S1474-4422(08)70259-X

Choi, J. S., Betz, J., Li, L., & et al. (2016). Association of using hearing aids or cochlear implants with changes in depressive symptoms in older adults. *Otolaryngology–Head & Neck Surgery, 142*(7), 652-657. doi: 10.1001/jamaoto.2016.0700

Contrera, K. J., Betz, J., Li, L., Blake, C. R., Sung, Y. K., Choi, J. S., & Lin, F. R. (2016). Quality of life after intervention with a cochlear implant or hearing aid. *The Laryngoscope, 126*(9), 2110-2115. doi: 10.1002/lary.25848

Conway, C. M., & Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 31*, 24-39. doi: 10.1037/0278-7393.31.1.24

Conway, C. M., & Christiansen, M. H. (2006). Statistical learning within and between modalities: pitting abstract against stimulus-specific representations. *Psychological Science, 17*, 905-912. doi: 10.1111/j.1467-9280.2006.01801.x

Damen, G. W., Beynon, A. J., Krabbe, P. F., Mulder, J. J., & Mylanus, E. A. (2007). Cochlear implantation and quality of life in postlingually deaf adults: long-term follow-up. *Otolaryngology–Head & Neck Surgery, 136*(4), 597-604. doi: 10.1016/j.otohns.2006.11.044

Davis, M. H., Johnsrude, I. S., Hervais-Adelman, A., Taylor, K., & McGettigan, C. (2005). Lexical information drives perceptual learning of distorted speech: evidence from the comprehension of noise-vocoded sentences. *Journal of Experimental Psychology. General, 134*, 222-241. doi: 10.1037/0096-3445.134.2.222

Di Nardo, W., Anzivino, R., Giannantonio, S., Schinaia, L., & Paludetti, G. (2014). The effects of cochlear implantation on quality of life in the elderly. *European Archives of Oto-Rhino-Laryngology, 271*(1), 65-73. doi: 10.1007/s00405-013-2396-1

Douiri, A., Rudd, A. G., & Wolfe, C. D. (2013). Prevalence of poststroke cognitive impairment: South London Stroke Register 1995-2010. *Stroke, 44*(1), 138-145. doi: 10.1161/STROKEAHA.112.670844

Ferguson, M. A., & Henshaw, H. (2015). Auditory training can improve working memory, attention, and communication in adverse conditions for adults with hearing loss. *Frontiers in Psychology, 6*, 556. doi: 10.3389/fpsyg.2015.00556

Ferguson, M. A., Henshaw, H., Clark, D. P., & Moore, D. R. (2014). Benefits of phoneme discrimination training in a randomized controlled trial of 50- to 74-year-olds with mild hearing loss. *Ear and Hearing, 35*(4), e110-121. doi: 10.1097/AUD.0000000000000020

Foo, C., Rudner, M., Ronnberg, J., & Lunner, T. (2007). Recognition of speech in noise with new hearing instrument compression release settings requires explicit cognitive storage and processing capacity. *Journal of the American Academy of Audiology, 18*(7), 618-631. doi: 10.3766/jaaa.18.7.8

Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: the paradox of statistical learning. *Trends in Cognitive Sciences, 19*(3), 117-125. doi: 10.1016/j.tics.2014.12.010

Füllgrabe, C., Moore, B. C., & Stone, M. A. (2014). Age-group differences in speech identification despite matched audiometrically normal hearing: contributions from auditory temporal processing and cognition. *Frontiers in Aging Neuroscience, 6*, 347. doi: 10.3389/fnagi.2014.00347

Füllgrabe, C., & Rosen, S. (2016a). Investigating the role of working memory in speech-in-noise identification for listeners with normal hearing. In P. van Dijk, D. Başkent, E. Gaudrain, E. de Kleine, A. Wagner & C. Lanting (Eds.), *Physiology, Psychoacoustics and Cognition in Normal and Impaired Hearing* (pp. 29-36). Cham: Springer International Publishing.

Füllgrabe, C., & Rosen, S. (2016b). On the (un)importance of working memory in speech-in-noise processing for listeners with normal hearing thresholds. *Frontiers in Psychology, 7*(1268). doi: 10.3389/fpsyg.2016.01268

Gilbert, J. L., Tamati, T. N., & Pisoni, D. B. (2013). Development, reliability, and validity of PRESTO: a new high-variability sentence recognition test. *Journal of the American Academy of Audiology, 24*(1), 26-36. doi: 10.3766/jaaa.24.1.4

Golomb, J. D., Peelle, J. E., & Wingfield, A. (2007). Effects of stimulus variability and adult aging on adaptation to time-compressed speech. *The Journal of the Acoustical Society of America, 121*, 1701. doi: 10.1121/1.2436635

Gordon-Salant, S., Yeni-Komshian, G. H., Fitzgibbons, P. J., & Schurman, J. (2010). Short-term adaptation to accented English by younger and older adults. *The Journal of the Acoustical Society of America, 128*, 200-204. doi: 10.1121/1.3486199

Hedden, T., Lautenschlager, G., & Park, D. C. (2005). Contributions of processing ability and knowledge to verbal memory tasks across the adult life-span. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology, 58*(1), 169-190. doi: 10.1080/02724980443000179

Henshaw, H., & Ferguson, M. A. (2014). Assessing the benefits of auditory training to real-world listening: identifying appropriate and sensitive outcomes. In T. Dau & S. Santurette (Eds.), *Auditory plasticity - listening with the brain* (pp. 45-52). Ballerup, Denmark: Danavox Jubilee Foundation.

Heydebrand, G., Hale, S., Potts, L., Gotter, B., & Skinner, M. (2007). Cognitive predictors of improvements in adults' spoken word recognition six months after cochlear implant activation. *Audiology & Neuro-Otology, 12*(4), 254-264. doi: 10.1159/000101473

Hinderink, J. B., Krabbe, P. F., & Van Den Broek, P. (2000). Development and application of a health-related quality-of-life instrument for adults with cochlear implants: the Nijmegen cochlear implant questionnaire. *Otolaryngology–Head and Neck Surgery, 123*(6), 756-765. doi: 10.1067/mhn.2000.108203

Hogan, A., O'Loughlin, K., Davis, A., & Kendig, H. (2009). Hearing loss and paid employment: Australian population survey findings. *International Journal of Audiology, 48*(3), 117-122. doi: 10.1080/14992020802449008

Hornsby, B. W. (2013). The effects of hearing aid use on listening effort and mental fatigue associated with sustained speech processing demands. *Ear and Hearing, 34*(5), 523-534. doi: 10.1097/AUD.0b013e31828003d8

Hughes, K. C., & Galvin, K. L. (2013). Measuring listening effort expended by adolescents and young adults with unilateral or bilateral cochlear implants or normal hearing. *Cochlear Implants International, 14*(3), 121-129. doi: 10.1179/1754762812Y.0000000009

Janse, E., & Adank, P. (2012). Predicting foreign-accent adaptation in older adults. *Quarterly Journal of Experimental Psychology, 65*, 1563-1585. doi: 10.1080/17470218.2012.658822

Kaufman, S. B., De Young, C. G., Gray, J. R., Jimenez, L., Brown, J., & Mackintosh, N. (2010). Implicit learning as an ability. *Cognition, 116*(3), 321-340. doi: 10.1016/j.cognition.2010.05.011

Kemper, S., & Sumner, A. (2001). The structure of verbal abilities in young and older adults. *Psychology and Aging, 16*(2), 312-322. doi: 10.1037//0882-7974.16.2.312

Kim, H., & Nanney, S. (2014). Familiarization effects on word intelligibility in dysarthric speech. *Folia Phoniatrica et Logopaedica, 66*(6), 258-264. doi: 10.1159/000369799

King, S. E., Firszt, J. B., Reeder, R. M., Holden, L. K., & Strube, M. (2012). Evaluation of TIMIT sentence list equivalency with adult cochlear implant recipients. *Journal of the American Academy of Audiology, 23*(5), 313-331. doi: 10.3766/jaaa.23.5.3

Kleinschmidt, D. F., & Jaeger, T. F. (2015). Robust speech perception: recognize the familiar, generalize to the similar, and adapt to the novel. *Psychological Review, 122*(2), 148-203. doi: 10.1037/a0038695

Klop, W. M., Boermans, P. P., Ferrier, M. B., van den Hout, W. B., Stiggelbout, A. M., & Frijns, J. H. (2008). Clinical relevance of quality of life outcome in cochlear implantation in postlingually deafened adults. *Otology & Neurotology, 29*(5), 615-621. doi: 10.1097/MAO.0b013e318172cfac

Kramer, S. E., Kapteyn, T. S., & Houtgast, T. (2006). Occupational performance: comparing normally-hearing and hearing-impaired employees using the Amsterdam Checklist for Hearing and Work. *International Journal of Audiology, 45*(9), 503-512. doi: 10.1080/14992020600754583

Midford, R., & Kirsner, K. (2005). Implicit and explicit learning in aged and young adults. *Aging, Neuropsychology, and Cognition, 12*, 359-387. doi: 10.1080/13825580500246894

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. *Cognitive Psychology, 41*, 49-100. doi: 10.1006/cogp.1999.0734

Negash, S., Howard, D. V., Japikse, K. C., & Howard, J. H. (2003). Age-related differences in implicit learning of non-spatial sequential patterns. *Aging, Neuropsychology, and Cognition, 10*, 108-121. doi: 10.1076/anec.10.2.108.14462

Palmer, S. D., & Mattys, S. L. (2016). Speech segmentation by statistical learning is supported by domain-general processes within working memory. *Quarterly Journal of Experimental Psychology, 69*(12), 2390-2401. doi: 10.1080/17470218.2015.1112825

Pals, C., Sarampalis, A., & Başkent, D. (2013). Listening effort with cochlear implant simulations. *Journal of Speech, Language, and Hearing Research, 56*, 1075-1084. doi: 10.1044/1092-4388(2012/12-0074)

Peelle, J. E., & Wingfield, A. (2005). Dissociations in perceptual learning revealed by adult age differences in adaptation to time-compressed speech. *Journal of Experimental Psychology. Human Perception and Performance, 31*, 1315-1330. doi: 10.1037/0096-1523.31.6.1315

Pisoni, D. B., Kronenberger, W. G., Chandramouli, S. H., & Conway, C. M. (2016). Learning and memory processes following cochlear implantation: the missing piece of the puzzle. *Frontiers in Psychology, 7*, 493. doi: 10.3389/fpsyg.2016.00493

Ramscar, M., Hendrix, P., Shaoul, C., Milin, P., & Baayen, H. (2014). The myth of cognitive decline: non-linear dynamics of lifelong learning. *Topics in Cognitive Science, 6*, 5-42. doi: 10.1111/tops.12078

Ringholz, G. M., Appel, S. H., Bradshaw, M., Cooke, N. A., Mosnik, D. M., & Schulz, P. E. (2005). Prevalence and patterns of cognitive impairment in sporadic ALS. *Neurology, 65*(4), 586-590. doi: 10.1212/01.wnl.0000172911.39167.b6

Rippon, G. A., Scarmeas, N., Gordon, P. H., Murphy, P. L., Albert, S. M., Mitsumoto, H., . . . Stern, Y. (2006). An observational study of cognitive impairment in amyotrophic lateral sclerosis. *Archives of Neurology, 63*(3), 345-352. doi: 10.1001/archneur.63.3.345

Rönnberg, J., Rudner, M., Foo, C., & Lunner, T. (2008). Cognition counts: A working memory system for ease of language understanding (ELU). *International journal of audiology, 47*(sup2), S99-S105. doi: 10.1080/14992020802301167

Rönnberg, J., Rudner, M., Lunner, T., & Zekveld, A. (2010). When cognition kicks in: working memory and speech understanding in noise. *Noise & Health, 12*(49), 263-269. doi: 10.4103/1463-1741.70505

Rudner, M., Foo, C., Ronnberg, 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*(5), 405-418. doi: 10.1111/j.1467-9450.2009.00745.x

Rudner, M., Foo, C., Sundewall-Thoren, E., Lunner, T., & Ronnberg, J. (2008). Phonological mismatch and explicit cognitive processing in a sample of 102 hearing-aid users. *International Journal of Audiology, 47*, S91-S98. doi: 10.1080/14992020802304393

Rudner, M., Ronnberg, J., & Lunner, T. (2011). Working Memory Supports Listening in Noise for Persons with Hearing Impairment. *Journal of the American Academy of Audiology, 22*(3), 156-167. doi: 10.3766/Jaaa.22.3.4

Salthouse, T. A. (2009a). Decomposing age correlations on neuropsychological and cognitive variables. *J Int Neuropsychol Soc, 15*(5), 650-661. doi: 10.1017/S1355617709990385

Salthouse, T. A. (2009b). When does age-related cognitive decline begin? *Neurobiology of Aging, 30*(4), 507-514. doi: 10.1016/j.neurobiolaging.2008.09.023

Salthouse, T. A. (2010). Selective review of cognitive aging. *Journal of the International Neuropsychological Society, 16*(5), 754-760. doi: 10.1017/S1355617710000706

Salthouse, T. A., McGuthry, K. E., & Hambrick, D. Z. (1999). A framework for analyzing and interpreting differential aging patterns: application to three measures of implicit learning. *Aging, Neuropsychology, and Cognition, 6*, 1-18. doi: 10.1076/anec.6.1.1.789

Sapir, S., Ramig, L. O., & Fox, C. M. (2011). Intensive voice treatment in Parkinson's disease: Lee Silverman voice treatment. *Expert Review of Neurotherapeutics, 11*(6), 815-830. doi: 10.1586/ern.11.43

Schoof, T., & Rosen, S. (2014). The role of auditory and cognitive factors in understanding speech in noise by normal-hearing older listeners. *Frontiers in Aging Neuroscience, 6*, 307. doi: 10.3389/fnagi.2014.00307

Schroter-Morasch, H., & Ziegler, W. (2005). Rehabilitation of impaired speech function (dysarthria, dysglossia). *GMS Current Topics in Otorhinolaryngology, Head and Neck Surgery, 4*, 15. doi: 10.1002/cncr.10668

Siegelman, N., Bogaerts, L., Christiansen, M. H., & Frost, R. (2017). Towards a theory of individual differences in statistical learning. *Philosophical Transactions of the Royal Society. B, Biological Sciences, 372*(1711). doi: 10.1098/rstb.2016.0059

Siegelman, N., & Frost, R. (2015). Statistical learning as an individual ability: theoretical perspectives and empirical evidence. *Journal of Memory and Language, 81*, 105-120. doi: 10.1016/j.jml.2015.02.001

Simon, J. R., Howard, J. H., & Howard, D. V. (2011). Age differences in implicit learning of probabilistic unstructured sequences. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences, 66*, 32-38. doi: 10.1093/geronb/gbq066

Simon, J. R., Vaidya, C. J., Howard, J. H., & Howard, D. V. (2012). The effects of aging on the neural basis of implicit associative learning in a probabilistic triplets learning task. *Journal of Cognitive Neuroscience, 24*, 451-463. doi: 10.1162/jocn_a_00116

Smith, G. E., Housen, P., Yaffe, K., Ruff, R., Kennison, R. F., Mahncke, H. W., & Zelinski, E. M. (2009). A cognitive training program based on principles of brain plasticity: results from the improvement in memory with Plasticity-based Adaptive Cognitive Training (IMPACT) study. *Journal of the American Geriatrics Society, 57*(4), 594-603. doi: 10.1111/j.1532-5415.2008.02167.x

Stacey, P. C., Raine, C. H., O'Donoghue, G. M., Tapper, L., Twomey, T., & Summerfield, A. Q. (2010). Effectiveness of computer-based auditory training for adult users of cochlear implants. *International Journal of Audiology, 49*(5), 347-356. doi: 10.3109/14992020903397838

Sweetow, R. W., & Sabes, J. H. (2006). The need for and development of an adaptive Listening and Communication Enhancement (LACE) Program. *Journal of the American Academy of Audiology, 17*(8), 538-558. doi: 10.3766/jaaa.17.8.2

Tamati, T. N., Gilbert, J. L., & Pisoni, D. B. (2013). Some factors underlying individual differences in speech recognition on PRESTO: a first report. *Journal of the American Academy of Audiology, 24*(7), 616-634. doi: 10.3766/jaaa.24.7.10

Vandenbossche, J., Coomans, D., Homblé, K., & Deroost, N. (2014). The effect of cognitive aging on implicit sequence learning and dual tasking. *Frontiers in Psychology, 5*, 154. doi: 10.3389/fpsyg.2014.00154

Wayne, R. V., Hamilton, C., Jones Huyck, J., & Johnsrude, I. S. (2016). Working memory training and speech in noise comprehension in older adults. *Frontiers in Aging Neuroscience, 8*, 49. doi: 10.3389/fnagi.2016.00049

World Health Organization. (2001, 27.01.2017). International Classification of Functioning, Disability and Health (ICF) Retrieved 19.04.2017, from http://www.who.int/classifications/icf/en/

8

Nederlandse samenvatting
Acknowledgements
Curriculum Vitae
List of Publications
MPI series in Psycholinguistics

Nederlandse samenvatting

Als we naar spraak luisteren, maken we onbewust voorspellingen over datgene wat we horen. We voorspellen niet alleen *wat* de ander waarschijnlijk gaat zeggen, maar voorspellen ook *hoe* iets zal worden uitgesproken. Zo verwachten we dat een Schot anders zal klinken dan iemand die in Londen opgroeide, we verwachten dat de stem van een wat ouder persoon een beetje krakerig zal klinken, en alleen al op basis van de naam of van het uiterlijk van iemand die we voor het eerst zien, verwachten we dat hij of zij met een accent praat. In de meeste gevallen kloppen onze voorspellingen en ondersteunen daarmee onze bestaande opvattingen. Maar soms kloppen onze voorspellingen niet. In die gevallen ontvangen we een heleboel nieuwe informatie die we kunnen gebruiken om onze opvattingen aan te passen. In beide gevallen leren we zowel van het verwachte als van het onverwachte.

In dit proefschrift heb ik dit proces om van ervaring te leren onderzocht op het gebied van de spraakwaarneming. Als we naar spraak luisteren gebruiken we deze vaardigheid namelijk regelmatig We wennen bijvoorbeeld makkelijk aan subtiele afwijkingen in het spraaksignaal zoals een ietwat vreemd uitgesproken klank (Norris, McQueen and Cutler, 2003). Meestal beseffen we niet eens dat onze hersenen aan het leren zijn terwijl we luisteren, omdat onze hersenen heel snel wennen aan kleine veranderingen in het spraaksignaal. We merken de verbazingwekkende flexibiliteit en leervaardigheid van onze hersenen pas op als ons spraaksysteem wordt uitgedaagd. Als spraakinput namelijk te veel afwijkt van alles wat we ooit hebben gehoord, hebben onze hersenen enkele minuten, dagen of zelfs maanden nodig om aan deze ongewone spraakinput te wennen en de spraak makkelijk te kunnen verstaan.

Uitgaand van het feit dat meer dan 90% van de Nederlandse bevolking en meer dan 50% van de Europese bevolking aangeeft meer dan één taal te kunnen spreken (Europees Commissie, 2012), is de kans groot dat de meesten van ons al een keer een langer gewenningsproces hebben meegemaakt door te luisteren naar iemand die met een vreemd accent spreekt. Neem bijvoorbeeld het volgende gesprek tussen een

Franse hotelgast en de Engelstalige roomservice (ontleend aan www.funny-joke-pictures.com):

>> Allo? Room service? Ici Monsieur Roux. I would like some pepper. <<

>> Certainly, sir. Black or white? <<

>> Toilet. <<

In dit voorbeeld hebben we contextinformatie (*toilet*) nodig om de tweedetaalspreker te kunnen begrijpen (*pepper* → *paper*). Als we naar spraak met een vreemd accent luisteren, vallen we mogelijk eveneens terug op extra middelen om de ongewone spraak te kunnen begrijpen en er vervolgens aan te kunnen wennen. In deze dissertatie wilde ik daarom onderzoeken welke interne middelen van luisteraars, meer specifiek welke individuele cognitieve en perceptuele vaardigheden, mogelijk voorspellen hoe goed iemand kan wennen aan ongewone spraakinput. In andere woorden, wat maakt iemand een goede luisteraar en vooral een goede aanpasser.

De vraag hoe we in staat zijn om onze spraakwaarneming aan te passen puur op basis van datgene wat we horen staat centraal in deze dissertatie. Dit proces van perceptueel leren, oftewel leren door waarnemen (perceptie), is uiteindelijk wat ons zo flexibel maakt in de spraakherkenning en waarborgt dat we spraak makkelijk kunnen verstaan. Korte tijd geleden werd voorgesteld dat het oppikken van onderliggende distributies in het spraaksignaal en dus het oppikken van statistische regelmatigheden in het spraaksignaal ten grondslag ligt aan onze vaardigheid om perceptueel te leren (Kleinschmidt & Jaeger, 2015). Ik heb me in dit proefschrift daarom vooral gericht op de concepten van perceptueel leren en van statistisch leren en hoe leeftijd en individuele verschillen mogelijk deze leerprocessen beïnvloeden.

Perceptueel leren bestaat uit twee delen. Ten eerste moet de ongewone spraakinput afgestemd worden op bestaande (woord)representaties in het langetermijngeheugen. Ten tweede moeten de representaties aangepast worden op basis van de input. Deze twee processen zijn onvermijdelijk met elkaar verbonden in taken van perceptueel leren. Hoofdstuk 2 had als doel om uit te vinden welke perceptuele en cognitieve processen met name betrokken zijn in het eerste onderdeel bij het afstemmen en vergelijken van auditieve informatie. Ik onderzocht daarom de auditieve discriminatievaardigheid van mensen. Auditieve discriminatie is immers de vaardigheid om verschillen tussen twee achter elkaar aangeboden auditieve (spraak)stimuli waar te nemen. Auditieve discriminatie zou dus inzicht kunnen bieden in processen die bij de succesvolle afstemming van auditieve informatie betrokken zijn. Auditieve discriminatie is echter niet gelijk te stellen met perceptueel leren omdat het voor perceptueel leren vooral belangrijk is om de spraakinput te koppelen aan representaties in het langetermijngeheugen. Auditieve discriminatie vereist daarentegen het vergelijken van twee auditieve stimuli in het kortetermijngeheugen. Dat wil zeggen, een binnenkomende spraakstimulus moet vergeleken worden met de "auditieve indruk" die een net beluisterde spraakinput zojuist heeft achtergelaten in het kortetermijngeheugen.

Auditieve discriminatie is ook in een bredere context van belang omdat auditieve discriminatie een voorwaarde is om spraakgedrag aan te kunnen passen. Daardoor bevordert auditieve discriminatie dus de menselijke flexibiliteit in spraak. Aan patiënten die spraaktherapieprogramma's doorlopen, zoals het Lee Silverman Voice Treatment (Sapir, Ramig, & Fox, 2011) of de E-learning gebaseerde SpraakTherapie (Beijer et al., 2010), wordt bijvoorbeeld vaak gevraagd om hun eigen uitingen te vergelijken met doelrealisaties van die uitingen (bv. uitgesproken door een logopedist). Adequate zelfwaarneming is dus de sleutel om gewenste veranderingen in het spraakgedrag aan te brengen (Schroter-Morasch & Ziegler, 2005) en de vaardigheid om twee op elkaar volgende uitingen met elkaar te vergelijken is essentieel om de eigen spraakproductie aan te passen.

Mijn resultaten met betrekking tot de auditieve discriminatievaardigheid van ouderen lieten zien dat de auditieve discriminatie van spraakstimuli vooral samenhangt met cognitieve en linguïstische vaardigheden en minder met auditieve vaardigheden. Ouderen waren nauwkeuriger in het onderscheiden of twee spraakstimuli gelijk of verschillend

waren als ze spraak in ruis beter konden verstaan, sneller waren in het verwerken van zowel visuele als linguïstische informatie en hoger scoorden op een test van hun algemene taalvaardigheid. De discriminatienauwkeurigheid nam echter af met toenemende leeftijd. Vooral maten die de verwerkingssnelheid van de proefpersonen weergaven waren geassocieerd met hun vaardigheid om twee op elkaar volgende spraakstimuli succesvol van elkaar te onderscheiden.
In hoofdstuk 3 heb ik onderzocht in hoe verre modaliteit (auditief of visueel), het soort stimulus en aandacht bijdragen aan statistisch leren. Dit hoofdstuk vormt daarmee het startpunt om de relatie te kunnen testen tussen statistisch leren en perceptueel leren in hoofdstuk 4. Hoofdstuk 3 probeert specifiek drie vragen te beantwoorden: (1) is het leren van temporele regelmatigheden anders als informatie auditief of visueel gepresenteerd wordt in een statistische-lerentaak, (2) maakt het voor het statistisch leren uit of de input linguïstische informatie bevat of niet en (3) is aandacht bij de verwerking van de input noodzakelijk voor het statistisch leren. Om die vragen te kunnen beantwoorden werden vier varianten van een "serial reaction time – artificial grammar learning taak" aangeboden die verschilden in (1) de modaliteit van de stimuli (auditief of visueel), (2) het soort stimulus (linguïstisch of niet-linguïstisch) en (3) aandacht (wel of geen gerichte aandacht nodig voor de input om de taak uit te kunnen voeren). Duidelijke effecten van statistisch leren werden geobserveerd in drie van de vier taak varianten.
Overeenkomstig met eerdere studies omtrent de effecten van modaliteit in statistisch leren, werden sterkere leereffecten gevonden als informatie auditief werd aangeboden. Ik kon laten zien dat dit voordeel van auditieve verwerking ook aanwezig is als de informatiestroom, waarin geleerd wordt, constant wordt gehouden (een serie van altijd vier visueel aangeboden pseudowoorden). De taak verschilde alleen in hoe de doelwoorden binnen de aangeboden pseudowoorden werden aangeduid (met behulp van een auditieve weergave of door een visuele orthografische representatie in het midden van het scherm). Met betrekking to de vraag of mensen met name gevoelig zijn voor statistische regelmatigheden in *linguïstisch*e (vergeleken met niet-

linguïstische) input, vond ik eerder bewijs voor het tegendeel. Dat wil zeggen, non-linguïstische items (visueel aangeboden meetkundige figuren) lieten een voordeel in de verwerking zien in vergelijking met linguïstische stimuli (uitgeschreven pseudowoorden). Dit suggereert dat niet alleen modaliteitspecifieke processen maar ook stimulusspecifieke processen betrokken zijn bij statistisch leren. Wat betreft de invloed van aandacht op statistisch leren, konden geen verschillen geobserveerd worden tussen de twee taakvarianten die in ons experiment zijn gebruikt. In de ene taakvariant moesten proefpersonen expliciet op de doelfiguren letten om de taak uit te kunnen voeren (een afstemmingstaak waarin de doelfiguren visueel werden aangeduid door een kleinere afbeelding van die figuur af te beelden in het midden van het scherm). In de andere taakvariant was aandacht voor het doelfiguur zelf niet noodzakelijk omdat het voor het uitvoeren van de taak niet nodig was om het doelfiguur te identificeren (d.w.z. men moest het figuur aanklikken dat visueel werd gemarkeerd door de verschijning van een rood kruisje in het midden).

In hoofdstuk 4 heb ik individuele verschillen in statistisch en perceptueel leren onderzocht aan de hand van de vraag of de vaardigheid om aan een ongewone spraakinput te wennen voorspeld zou kunnen worden door een algemene gevoeligheid voor statistische regelmatigheden. Om te voorkomen dat overeenkomsten in de auditieve of linguïstische verwerking van de stimuli ten grondslag zouden liggen aan een samenhang tussen de twee soorten van leren, werd specifiek díe variant van de statistische-lerentaak aangeboden die gebruik maakte van visuele, niet-linguïstische stimuli en waarvoor geen specifieke aandacht voor de stimuli nodig was. Deze taakvariant had duidelijke leereffecten laten zien in hoofdstuk 3.

Evenals in studies omtrent perceptueel leren (Adank & Janse, 2010; Golomb, Peelle, & Wingfield, 2007; Gordon-Salant, Yeni-Komshian, Fitzgibbons, & Schurman, 2010; Peelle & Wingfield, 2005) lieten zowel jongere als oudere deelnemers een significante vooruitgang zien in het verstaan van de noise-vocoded spraak, die geacht wordt een simulatie te zijn van spraak zoals waargenomen via een cochleair implantaat [CI]. Alleen jongere proefpersonen bleken echter gevoelig voor de statistische regelmatigheden in de statistische leertaak. Zoals verwacht werd op basis van huidige modellen van perceptueel leren (Ahissar & Hochstein, 2004; Kleinschmidt & Jaeger, 2015) konden verschillen in de mate van perceptueel leren daadwerkelijk gecorreleerd worden aan de individuele vaardigheid van jongere proefpersonen om statistische regelmatigheden op te pikken. Woordenschat was een tweede voorspeller van de mate waarin jongere proefpersonen konden wennen aan CI-gesimuleerde spraak. Deze resultaten benadrukken de betrokkenheid van de individuele vaardigheden van een luisteraar om een ongewone spraakinput te kunnen leren verstaan.

Voor het onderzoek in hoofdstuk 5 werd een follow-up experiment opgezet met betrekking tot de statistische leervaardigheid van ouderen. In hoofdstuk 4 kon immers geen statistisch leren geobserveerd worden in de groep van oudere deelnemers. De resultaten van hoofdstuk 4 deden vermoeden dat de gevoeligheid voor statistische regelmatigheden afneemt als ouderen snel op elkaar volgende visuele stimuli moeten verwerken. Op basis van eerdere bevindingen dat ouderen geen statistisch leren laten zien in visuele taken (bv. Simon, Howard, & Howard, 2011) werd aangevoerd dat de algemene vaardigheid om regelmatigheden waar te nemen afneemt op oudere leeftijd (Negash, Howard, Japikse, & Howard, 2003). Als met leeftijd daadwerkelijk de vaardigheid vermindert om statistische regelmatigheden op te merken, dan zou de statistische leervaardigheid van ouderen ook aangetast moeten zijn in een andere (niet-visuele) modaliteit. Gezien het feit dat temporele regelmatigheden van bijzonder belang zijn in de taal- en spraakverwerking, wilde ik in hoofdstuk 5 onderzoeken of jongere en oudere proefpersonen ook verschillen in auditief statistisch leren. De resultaten lieten geen verschil tussen jongeren en ouderen zien wat betreft hun leervaardigheid. Dit resultaat spreekt tegen een algemene leeftijdsgebonden afname in de vaardigheid om statistisch te leren.

In hoofdstuk 6 ben ik verder ingegaan op de rol van cognitieve vaardigheden voor statistisch en perceptueel leren, maar dan in een klinische populatie. Waar hoofdstuk 4 gebruik maakte van CI-simulatiespraak om te onderzoeken hoe mensen binnen een experimentele sessie kunnen wennen aan een ongewone spraak input,

was hoofdstuk 6 opgezet om een perceptueel leerproces “in het echte leven” te volgen. Te dien einde werden negen, postlinguaal doof volwassenen voor hun cochleaire implantatie getest en werden gevolgd tijdens de eerste drie maanden na de activatie van hun implantaat. Over het algemeen lieten patiënten met een hogere verwerkingssnelheid, hoger werkgeheugen capaciteit en grotere woordenschat een betere vooruitgang in hun spraakverstaan zien na implantaatactivatie.
Hoofdstuk 6 had bovendien als doel om het gebruik van meer natuurlijk spraakmateriaal (spraakfragmenten uit conversaties) te evalueren voor gehoorrevalidatie. Dit werd gedaan in vergelijking met standaard spraakmateriaal (losse woorden en eenvoudige zinnen) dat normaalgesproken in de klinische praktijk gebruikt wordt om spraakverstaan te meten. Een samenhang tussen objectieve maten van spraakverstaan en de door patiënten aangegeven kwaliteit van leven kon alleen geobserveerd worden voor conversatiespraak. Deze bevinding benadrukt het potentieel dat meer natuurlijk spraakmateriaal zou kunnen hebben voor gehoorrevalidatie. Objectieve maten van de spraakwaarneming die gebaseerd zijn op meer natuurlijk spraakmateriaal zouden gebruikt kunnen worden om beter te evalueren hoe patiënten in het alledaagse leven functioneren met hun CI.

Acknowledgements

Completing a PhD-project is a combined effort. During the last couple of years, my PhD-project allowed me to meet numerous generous, open-minded and decent people in- and outside of academia who supported me on my way.

First and foremost, I would like to thank my both supervisors Toni and Esther. Toni, I am so grateful to you that you gave me the opportunity to pursue my dream of obtaining a doctoral degree. I already got to know you during my masters and you taught me most of what I know about phonetics and statistics. I still remember your words when the master projects were allotted. You said that during the work on their master thesis few students manage to evolve from students to colleagues. I always wanted to be one of those exceptional cases and you always gave me the feeling that I was. I am still secretly proud that you remembered my name though you often said that your memory for student names is far from excellent. One reason could be that particularly my last name is relatively special considering its meaning in Dutch and German (and the reason why I am the only person you write down with her full name in your agenda). However, I always felt that it was me you remembered, that I somehow stood out for you. You always showed and told me explicitly that you believe in me and my qualities. You never had any doubts that I am able to complete my PhD to the fullest of your expectations. This feeling of unconditional support is what I value most about working with you the last couple of years. It motivated me more than anything else as I always aimed to stand up to your high standards.

Dear Esther, it is hard for me to decide what I am most grateful for the last couple of years as there are so many things on my mind. That is, you are an amazing supervisor and I never could have wished for a better one. I was your first official PhD student (though not the first to defend as Sophie and Conny beat me by a couple of days) and as such we both had to find our way into unfamiliar roles. I want to assure you that you did an excellent job from the beginning. Your door has always been open for me and my fellow PhD students. I think the only days that we didn't speak to each other were the days one of us was not around. Most importantly, you have the rare ability to guide your students on both a professional and a personal level. On a professional level, your profound knowledge and your analytical skills surely helped to improve my writing and this thesis. Your feedback was always quick, constructive and motivating. You followed each line of thought thoroughly and you were able to pick out ideas that were only vaguely worked out with ultimate precision. As I was working on your VIDI grant, I was working on a project you had preconceived. Nevertheless, you gave me the opportunity, the freedom and the support to realize my own ideas and to make the project my own. On a personal level, you always allowed your PhD students insight into the current state of your work, your career and your private life. You not only shared your successes but also your drawbacks and how to deal with them. You always had an honest interest in us, supported us and backed us and I think it is primarily your accomplishment that we all graduate. During the last years, you have been my most important role model. This thesis was not possible without you.

I am also grateful to the members of my manuscript committee, Mirjam Ernestus, Paul Boersma, Astrid van Wieringen, Sible Andringa and Alexis Hervais-Adelman who took their time and effort to read my thesis and who provided me with valuable feedback.

Moreover, I am indebted to all the participants, patients and their relatives. Without these committed individuals willing to spend their time to participate in my studies, none of this work was possible. Running my experiments, I was allowed to meet so many special people who were honestly interested in language research and my project. However, running studies on individual differences requires testing relatively large numbers of participants. Without the help of our student assistants Willemijn van den Bergh and Ilona van der Linden, I would probably still be testing. Thank you.

I would also like to thank my colleagues in the speech comprehension group. Annika, Arina, Ellen, Esther H., Helen, Huib, Iris, Kim, Kimberley, Laura, Lisa, Louis, Mark, Martijn, Mirjam Mybeth, Nienke, Rian, Robert, Sascha, Sho, Sophie and Xaver, I enjoyed working, socializing, having lunch and sharing my ideas with you. In

particular, I would like to thank my office mates: Menno, you made me feel at home during the initial months at the institute and helped me to gain back confidence in my English skills (which had seriously declined after speaking Dutch as a foreign language for several years only). Louis, I could always discuss my statistical analysis with you and you are such a cordial person I always love to talk to. Laura, your visits to our various offices were always a delight and we never ran out of conversation topics. I admire you for your assertiveness, pursuing a successful athletic career as well as your PhD. Sascha and particularly Lisa, you were my longest office mates and, therefore, you had to get most accustomed to my habits and quirks (e.g., brewing tea and storing it in a bleeping thermos flask, having the window open all day). I want to thank you for always bearing with me. Lisa, you are one of the most sociable persons I know and as such it was always a pleasure having you around. Ellen, Kim, Malte and Sophie, though we never actual shared an office, we shared so many friendly chats and coffee breaks throughout the last years that it felt as if we had.

Over the past years, I took part in meetings of the Psychology and Language group (MPI), the last meetings of the Language Comprehension group (MPI) and combined meetings with the Language and Speech, Learning and Therapy group (CLS) and the Language and Speech Technology group (CLS). As I had the opportunity to present my research projects and my data, I would like to thank all the (former) members of these groups for offering helpful suggestions, comments and ideas. Within and outside of these groups, I was allowed to meet so many wonderful colleagues along the various stages of my PhD-project. In particular, I would like to thank Antje S., Christina, Florian, Franziska, Gabriela, Ilse, Job, Laura H., Laurence, Linda, Louise, Mijntje, Odette, Polina, Remy, Salomi, Sarah, Wencui and Zeshu for all the nice little encounters at the workplace.

I thank Rachel, Dirkje, Els and Kevin as the coordinators of the IMPRS for Language Sciences for organizing so many valuable and diverse courses, workshops and social activities. You have always been there in the background, helping, taking care and checking up on me. I would also like to thank Tanja Döller as the former coordinator

of the Graduate School for Humanities. We started working at the RU at the same time and you supervised me and my progress during all the years. Meetings with you were always motivating and I enjoyed every single conversation that we had as you always really cared for your students.
The last chapter in this thesis is somewhat closest to my heart. As a speech and language therapist, I desire to apply my findings to the real world and by testing cochlear implant patients I got the opportunity to transfer some of my outcomes into the clinical setting. This had not been possible without the help of many great colleagues at the Radboud UMC. Wendy, Andy and Emmanuel, thank you for giving me the opportunity to run my study in the CI-team. Wendy, you not only supervised me and helped me with all the organizational aspects, e.g., gaining ethics approval, but you also helped to improve the study design and the chapter by your valuable feedback. Special thanks also go to the secretaries for helping with scheduling the test sessions, booking rooms and contacting patients. Your help was absolutely essential. Moreover, I would like to thank all the wonderful speech and language therapists in the CI-team. You do an amazing job and your honest interest in my research and its possible impact for your work made me always feel welcome and appreciated.
I was trained as a speech and language therapist at the HAN. I started to work there and former lecturers became colleagues. During the first years of my PhD-project, I continued to work at the HAN and I would like to thank all of my dear HAN-colleagues. You formed me to the speech and language therapist I am. You supported and believed in me. I feel immediately at home when I return to the department for a visit. Particularly, I would like to thank Antje, Lena and Sabine: for your friendship, your company and for always keeping me up to date.
A very special thanks goes to all the (former) members of the MPI volleyballteam. You are such a gorgeous group of people and I enjoyed every Thursday night with you. With you, I could totally relax (first from work, later from baby duties) and just have fun. The worst thing during my pregnancies was waiving our games. I would like to thank all of you personally but I am too much afraid that I will

forget someone. Therefore, I restrict myself to the people who formed – at least for me - the heart of the volleyball team in the last years: Ad, Alex, Dirkje, Haiteng, Kees, Ine, Maurice, Reiner, Sebastian thank you for all the nice games, the friendly mocking and the fun we had together.

Xaver and Jule, my dear Esther-siblings, together with Conny, we form the first batch of Esther's PhD children. It's something that knits us all together. I will always gladly remember our Sinterklaas festivities, our several conference trips and the many walks together to the Erasmus building. We shared the ups and downs of our work and of our private life. Now we are scattered all around Germany and barely see each other but I will always be happy to hear from you.

Sometimes fellow colleagues become friends and I'm lucky that two of them agreed to stand by my side today. Conny, we started around the same time and we finish around the same time. You were always there (with the exception of your 8-months stay in Chicago that felt like years). We share so many memories that my whole PhD-project feels like a Conny-retrospect. I will never forget our US trip during the governmental shutdown, which forced us to improvise a lot, the conference trips and shared hotel rooms, the many coffee breaks and the frequent visits in evenings and weekends. We immediately formed a bond and every day since you are in Ulm I miss you and your generosity, sense of humor, warmth and empathy.

Svetlana, you are such a welcoming and affectionate person. I am so grateful that our children currently grow up together and learn from each other. You accompanied me as a friend during the whole PhD-trajectory but particularly in the last phase you were often my support. As we were both struggling to finish our dissertations while having to take care of our little ones, you always had time, tea and warm words for me. I admire you for your determination and I am confident that I will be reading your acknowledgements soon.

Katja and Kimberley you are both my dear friends and paranymphs in heart. Katja, we got to know each other during the perception course and that made the course totally worth it. Being your paranymph was a great honor. I am so glad that we are staying in contact and that

you even manage to visit me sometimes. You are a loyal and faithful friend. When we talk it immediately feels as if we've seen each other just the other day. Kimberley, I still remember our first conversation at the PIF workshop in Berg and Dal. We spoke about future plans and how a family may fit into these lives. Today, you are setting an extraordinary example yourself. You are such an exceptionable and devoted woman handling both a scientific career and a fruitful family life while having to commute from Amsterdam to Nijmegen. One may think you don't have any time left for your friends but that is far from true. You are always approachable and the game tournaments you organize at your place are legendary.
I am grateful to all of my other wonderful friends for always being there and for staying in contact with me. My dear friends, you enrichen my life: Christin, Daniela, Florian, Ilka, Julia, Julia & Jeroen, Anna & Roman, Wanxun & Arnold, Fang & Ronald, and Rosanne.

To my family:
Liebe Ella, im Endspurt zu dieser Thesis warst du wohl die wichtigste Person. Ohne dich und deine Wochenendbesuche, in denen du dich 24 Stunden am Tag um Malina gekümmert hast, hätte ich diese Thesis bei weitem nicht so zügig zu Ende schreiben können. Gerade während ich diese Worte schreibe, höre ich Malina lauthals mit dir zusammen lachen. Aber ich will dir auch für deine Herzlichkeit und Fürsorglichkeit danken. Ich fühle mich bei dir immer zu Hause und bei jedem Besuch verwöhnst du uns mit Köstlichkeiten. Du bist die tollste babuschka, die ich mir für unsere Kinder wünschen kann.
Liebe Dina, deinen größten Wunsch habe ich dir schon erfüllt, der zweitgrößte Wunsch ist in Arbeit, und die Dissertation ist abgeschlossen. Jetzt muss ich nur noch besser Russisch lernen und ich hoffe, dass ich dafür in Zukunft mehr Zeit habe. Meine Großeltern können diesen besonderen Moment leider nicht mehr miterleben, aber ich bin sehr froh und dankbar, dass du dabei sein kannst, denn ich weiß dass du genauso stolz auf mich bist, wie sie es wären.
Liebe Alina, lieber Oliver, Mischa und Stella, bei euch fühle ich mich immer wohl und willkommen. Egal wie spontan unsere Besuche

sind, die Tür zu eurem kunterbunten Leben steht immer für uns offen. Wir freuen uns auf noch ganz viele gegenseitige Besuche und Silvesterabende mit euch.
Liebe Birthe, lieber Sebastian, euch verdanke ich die Urlaube der letzten Jahre und ich habe jeden einzelnen davon sehr genossen. Ich hoffe, dass wir auch in Zukunft noch oft zusammen wegfahren werden (auch wenn wir mittlerweile wesentlich größere Ferienwohnungen anmieten müssen ☺). Noch mehr aber hoffe ich, dass wir immer so eng verbunden bleiben werden. Lieber Sebastian, du bist ein toller Schwager und wie ein Bruder für mich. Ich bin sehr froh dass du zu unserer Familie gehörst. Liebe Birthe, als große Schwester hast du mich immer beschützt. Egal was passiert, ich weiß, dass du immer für mich und meine Familie da sein wirst und das bedeutet mir mehr als du dir vorstellen kannst. Besonders aber hoffe ich, dass du immer für mein zweites Kind da sein wirst. Das ist schließlich die Aufgabe einer Patentante.
Liebe Asbirg, meine große, kleine Schwester. Ich bin wahnsinnig stolz auf dich. Ich bewundere wie du dich um deine Mitmenschen kümmerst. Du hast immer ein offenes Ohr für deine Freunde und hilfst ihnen wie du nur kannst. Obwohl du seit einigen Jahren in Lübeck wohnst, habe ich nie das Gefühl, dass du weit weg bist. Du hältst immer den Kontakt, egal wie stressig es manchmal bei uns ist, und lässt uns an deinem Leben teilhaben. Du beißt dich durch dein Studium und gibst nie auf. In vielerlei Hinsicht sind wir uns nicht sehr ähnlich, aber mit dir habe ich immer die meisten Interessen und Hobbies geteilt, vom Klarinette spielen übers rätseln, puzzeln und lesen. Mit dir zusammen macht allerdings alles mehr Spaß.
Liebe Mama, lieber Papa, ihr habt mich immer unterstützt und zu dem Menschen geformt, der ich bin. Euer Ziel war es immer eure Kinder zu selbstständigen Menschen heranzuziehen und ich hoffe, dass ich das bei euren Enkelkindern annähernd so gut schaffen werde. Ihr habt mir immer alle Möglichkeiten gegeben und mich nie in meinen Entscheidungen eingeschränkt. Ihr habt mir immer das Gefühl und das Selbstvertrauen gegeben alles im Leben erreichen zu können. Ohne euch stünde ich nicht hier. Danke.

Dear Malina, dear little one. You two are the most wonderful diversions of a PhD project I can think of. No matter how proud I am of my dissertation, it will never compare to how proud I am of the work my body did and how proud I am of you. Though my PhD project is completed, the most important project just started: raising you to happy, caring, tolerant and independent individuals.
Dear Alex, I studied and investigated speech, language and communication for several years and, yet, it is difficult for me to find words that describe how much you mean to me. You are my partner, the father of my children and my best friend. You are my biggest support and my worst distraction. With you I feel precious and safe. You have the ability to always make me laugh and to always cheer me up. You are perfect with all your little imperfections and I am fond of every single one. The chances that we met were close to zero but still we are here. I could not be happier about it. Ich liebe dich.

Curriculum Vitae

Thordis Neger was born in 1986 in Goch, Germany. In 2009, she obtained her Bachelor's degree in speech and language therapy from the Hogeschool van Arnhem en Nijmegen (HAN). She received her Master's degree in Linguistics with a specialization in speech and language pathology from Radboud University in 2010 (cum laude). In the same year, she started to work as a lecturer at the speech and language therapy department of the HAN. In 2011, Thordis joined the Speech Comprehension PI group of the Centre for Language Studies at Radboud University to work on her PhD project.

List of Publications

- Neger, T. M., Rietveld, T., & Janse, E. (2014). Relationship between perceptual learning in speech and statistical learning in younger and older adults. Frontiers in Human Neuroscience, 8: 628. doi:10.3389/fnhum.2014.00628.

- Neger, T. M., Janse, E., & Rietveld, T. (2015). Correlates of older adults' discrimination of acoustic properties in speech. Speech, Language and Hearing, 18(2), 102-115. doi:10.1179/2050572814Y.0000000055.

- Neger, T. M., Rietveld, T., & Janse, E. (2015). Adult age effects in auditory statistical learning. In M. Wolters, J. Livingstone, B. Beattie, R. Smith, M. MacMahon, J. Stuart-Smith, & J. Scobbie (Eds.), Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS 2015). London: International Phonetic Association.

- Neger, T. M., Rietveld, T., Janse, E. (submitted). Effects of modality, stimulus type and attention on statistical learning

- Neger, T. M., Huinck W. J., Mylanus, E. A. M., Rietveld, T. and Janse, E. (in preparation). Cognitive predictors of speech perception outcome after cochlear implantation.

MPI Series in Psycholinguistics

1. The electrophysiology of speaking: Investigations on the time course of semantic, syntactic, and phonological processing. *Miranda van Turennout*
2. The role of the syllable in speech production: Evidence from lexical statistics, metalinguistics, masked priming, and electromagnetic midsagittal articulography. *Niels O. Schiller*
3. Lexical access in the production of ellipsis and pronouns. *Bernadette M. Schmitt*
4. The open-/closed-class distinction in spoken-word recognition. *Alette Haveman*
5. The acquisition of phonetic categories in young infants: A self-organising artificial neural network approach. *Kay Behnke*
6. Gesture and speech production. *Jan-Peter de Ruiter*
7. Comparative intonational phonology: English and German. *Esther Grabe*
8. Finiteness in adult and child German. *Ingeborg Lasser*
9. Language input for word discovery. *Joost van de Weijer*
10. Inherent complement verbs revisited: Towards an understanding of argument structure in Ewe. *James Essegbey*
11. Producing past and plural inflections. *Dirk Janssen*
12. Valence and transitivity in Saliba: An Oceanic language of Papua New Guinea. *Anna Margetts*
13. From speech to words. *Arie van der Lugt*
14. Simple and complex verbs in Jaminjung: A study of event categorisation in an Australian language. *Eva Schultze-Berndt*
15. Interpreting indefinites: An experimental study of children's language comprehension. *Irene Krämer*
16. Language-specific listening: The case of phonetic sequences. *Andrea Weber*
17. Moving eyes and naming objects. *Femke van der Meulen*

8

18. Analogy in morphology: The selection of linking elements in Dutch compounds. *Andrea Krott*
19. Morphology in speech comprehension. *Kerstin Mauth*
20. Morphological families in the mental lexicon. *Nivja H. de Jong*
21. Fixed expressions and the production of idioms. *Simone A. Sprenger*
22. The grammatical coding of postural semantics in Goemai (a West Chadic language of Nigeria). *Birgit Hellwig*
23. Paradigmatic structures in morphological processing: Computational and cross-linguistic experimental studies. *Fermín Moscoso del Prado Martín*
24. Contextual influences on spoken-word processing: An electrophysiological approach. *Daniëlle van den Brink*
25. Perceptual relevance of prevoicing in Dutch. *Petra M. van Alphen*
26. Syllables in speech production: Effects of syllable preparation and syllable frequency. *Joana Cholin*
27. Producing complex spoken numerals for time and space. *Marjolein Meeuwissen*
28. Morphology in auditory lexical processing: Sensitivity to fine phonetic detail and insensitivity to suffix reduction. *Rachèl J. J. K. Kemps*
29. At the same time...: The expression of simultaneity in learner varieties. *Barbara Schmiedtová*
30. A grammar of Jalonke argument structure. *Friederike Lüpke*
31. Agrammatic comprehension: An electrophysiological approach. *Marlies Wassenaar*
32. The structure and use of shape-based noun classes in Miraña (North West Amazon). *Frank Seifart*
33. Prosodically-conditioned detail in the recognition of spoken words. *Anne Pier Salverda*
34. Phonetic and lexical processing in a second language. *Mirjam Broersma*
35. Retrieving semantic and syntactic word properties. *Oliver Müller*
36. Lexically-guided perceptual learning in speech processing. *Frank Eisner*

37. Sensitivity to detailed acoustic information in word recognition. *Keren B. Shatzman*
38. The relationship between spoken word production and comprehension. *Rebecca Özdemir*
39. Disfluency: Interrupting speech and gesture. *Mandana Seyfeddinipur*
40. The acquisition of phonological structure: Distinguishing contrastive from non-contrastive variation. *Christiane Dietrich*
41. Cognitive cladistics and the relativity of spatial cognition. *Daniel B.M. Haun*
42. The acquisition of auditory categories. *Martijn Goudbeek*
43. Affix reduction in spoken Dutch. *Mark Pluymaekers*
44. Continuous-speech segmentation at the beginning of language acquisition: Electrophysiological evidence. *Valesca Kooijman*
45. Space and iconicity in German Sign Language (DGS). *Pamela Perniss*
46. On the production of morphologically complex words with special attention to effects of frequency. *Heidrun Bien*
47. Crosslinguistic influence in first and second languages: Convergence in speech and gesture. *Amanda Brown*
48. The acquisition of verb compounding in Mandarin Chinese. *Jidong Chen*
49. Phoneme inventories and patterns of speech sound perception. *Anita Wagner*
50. Lexical processing of morphologically complex words: An information-theoretical perspective. *Victor Kuperman*
51. A grammar of Savosavo, a Papuan language of the Solomon Islands. *Claudia Wegener*
52. Prosodic structure in speech production and perception. *Claudia Kuzla*
53. The acquisition of finiteness by Turkish learners of German and Turkish learners of French: Investigating knowledge of forms and functions in production and comprehension. *Sarah Schimke*
54. Studies on intonation and information structure in child and adult German. *Laura de Ruiter*
55. Processing the fine temporal structure of spoken words. *Eva Reinisch*

56. Semantics and (ir)regular inflection in morphological processing. *Wieke Tabak*
57. Processing strongly reduced forms in casual speech. *Susanne Brouwer*
58. Ambiguous pronoun resolution in L1 and L2 German and Dutch. *Miriam Ellert*
59. Lexical interactions in non-native speech comprehension: Evidence from electro-encephalography, eye-tracking, and functional magnetic resonance imaging. *Ian FitzPatrick*
60. Processing casual speech in native and non-native language. *Annelie Tuinman*
61. Split intransitivity in Rotokas, a Papuan language of Bougainville. *Stuart Robinson*
62. Evidentiality and intersubjectivity in Yurakaré: An interactional account. *Sonja Gipper*
63. The influence of information structure on language comprehension: A neurocognitive perspective. *Lin Wang*
64. The meaning and use of ideophones in Siwu. *Mark Dingemanse*
65. The role of acoustic detail and context in the comprehension of reduced pronunciation variants. *Marco van de Ven*
66. Speech reduction in spontaneous French and Spanish. *Francisco Torreira*
67. The relevance of early word recognition: Insights from the infant brain. *Caroline Junge*
68. Adjusting to different speakers: Extrinsic normalization in vowel perception. *Matthias J. Sjerps*
69. Structuring language. Contributions to the neurocognition of syntax. *Katrien R. Segaert*
70. Infants' appreciation of others' mental states in prelinguistic communication: A second person approach to mindreading. *Birgit Knudsen*
71. Gaze behavior in face-to-face interaction. *Federico Rossano*
72. Sign-spatiality in Kata Kolok: how a village sign language of Bali inscribes its signing space**.** *Conny de Vos*
73. Who is talking? Behavioural and neural evidence for norm-based coding in voice identity learning. *Attila Andics*

74. Lexical processing of foreign-accented speech: Rapid and flexible adaptation. *Marijt Witteman*

75. The use of deictic versus representational gestures in infancy. *Daniel Puccini*

76. Territories of knowledge in Japanese conversation. *Kaoru Hayano*

77. Family and neighbourhood relations in the mental lexicon: A cross-language perspective. *Kimberley Mulder*

78. Contributions of executive control to individual differences in word production. *Zeshu Shao*

79. Hearing speech and seeing speech: Perceptual adjustments in auditory-visual processing. *Patrick van der Zande*

80. High pitches and thick voices: The role of language in space-pitch associations. *Sarah Dolscheid*

81. Seeing what's next: Processing and anticipating language referring to objects. *Joost Rommers*

82. Mental representation and processing of reduced words in casual speech. *Iris Hanique*

83. The many ways listeners adapt to reductions in casual speech. *Katja Poellmann*

84. Contrasting opposite polarity in Germanic and Romance languages: Verum Focus and affirmative particles in native speakers and advanced L2 learners. *Giuseppina Turco*

85. Morphological processing in younger and older people: Evidence for flexible dual-route access. *Jana Reifegerste*

86. Semantic and syntactic constraints on the production of subject-verb agreement. *Alma Veenstra*

87. The acquisition of morphophonological alternations across languages. *Helen Buckler*

88. The evolutionary dynamics of motion event encoding. *Annemarie Verkerk*

89. Rediscovering a forgotten language. *Jiyoun Choi*

90. The road to native listening: Language-general perception, language-specific input. *Sho Tsuji*

91. Infants'understanding of communication as participants and observers. *Gudmundur Bjarki Thorgrímsson*

92. Information structure in Avatime. *Saskia van Putten*
93. Switch reference in Whitesands. *Jeremy Hammond*
94. Machine learning for gesture recognition from videos. *Binyam Gebrekidan Gebre*
95. Acquisition of spatial language by signing and speaking children: a comparison of Turkish sign language (TID) and Turkish. *Beyza Sümer*
96. An ear for pitch: on the effects of experience and aptitude in processing pitch in language and music. *Salomi Savvatia Asaridou*
97. Incrementality and Flexibility in Sentence Production. *Maartje van de Velde*
98. Social learning dynamics in chimpanzees: Reflections on (nonhuman) animal culture. *Edwin van Leeuwen*
99. The request system in Italian interaction. *Giovanni Rossi*
100. Timing turns in conversation: A temporal preparation account. *Lilla Magyari*
101. Assessing birth language memory in young adoptees**.** *Wencui Zhou*
102. A social and neurobiological approach to pointing in speech and gesture. *David Peeters*
103. Investigating the genetic basis of reading and language skills. *Alessandro Gialluisi*
104. Conversation Electrified: The Electrophysiology of Spoken Speech Act Recognition**.** *Rósa Signý Gisladottir*
105. Modelling Multimodal Language Processing. *Alastair Smith*
106. Predicting language in different contexts: The nature and limits of mechanisms in anticipatory language processing. *Florian Hintz*
107. Situational variation in non-native communication. *Huib Kouwenhoven*
108. Sustained attention in language production. *Suzanne Jongman*
109. Acoustic reduction in spoken-word processing: Distributional, syntactic, morphosyntatic, and orthographic effects. *Malte Viebahn*
110. Nativeness, dominance, and the flexibility of listening to spoken language. *Laurence Bruggeman*
111. Semantic specificity of perception verbs in Maniq. *Ewelina Wnuk*

112. On the identification of FOXP2 gene enhancers and their role in brain development. *Martin Becker*
113. Events in language and thought: The case of serial verb constructions in Avatime. *Rebecca Defina*
114. Deciphering common and rare genetic effects on reading ability. *Amaia Carrión Castillo*
115. Music and language comprehension in the brain. *Richard Kunert*
116. Comprehending Comprehension: Insights from neuronal oscillations on the neuronal basis of language**.** *Nietzsche H.L. Lam*
117. The biology of variation in anatomical brain asymmetries. *Tulio Guadalupe*
118. Language processing in a conversation context. *Lotte Schoot*
119. Achieving mutual understanding in Argentine Sign Language. *Elizabeth Manrique*
120. Talking Sense: the behavioural and neural correlates of sound symbolism. *Gwilym Lockwood*
121. Getting under your skin: The role of perspective and simulation of experience in narrative comprehension. *Franziska Hartung*
122. Sensorimotor experience in speech perception. *Will Schuerman*
123. Explorations of beta-band neural oscillations during language comprehension: Sentence processing and beyond. *Ashley Lewis*
124. Influences on the magnitude of syntactic priming. *Evelien Heyselaar*
125. Lapse organization in interaction. *Elliott Hoey*
126. The processing of reduced word pronunciation variants by natives and foreign language learners: Evidence from French casual speech. *Sophie Brand*
127. The neighbors will tell you what to expect: Effects of aging and predictability on language processing. *Cornelia Moers*
128. The role of voice and word order in incremental sentence processing. *Sebastian Sauppe*
129. Learning from the (un)expected: Age and individual differences in statistical learning and perceptual learning in speech. *Thordis Neger*