Doctoral Thesis

Age and hearing-loss effects on speech processing

Xaver Koch

ISBN: 978-90-76203-96-6

Cover photo: part of ‘Metamorphose’ high relief (Rüdiger Roehl, 1982)

Printed and bound by Ipskamp Drukkers b.v., Enschede

Age and hearing-loss effects on speech processing

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. Krieken,
volgens besluit van het college van decanen
in het openbaar te verdedigen op dinsdag 18 september 2018
om 14:30 uur precies

door

Xaver Koch

geboren op 2 september 1979
te Dresden, Duitsland

Promotor: Prof. dr. Mirjam Ernestus
Co-Promotor: Dr. Esther Janse

Manuscriptcommissie: Prof. dr. Roeland van Hout (voorzitter)
Prof. dr. Johan Verhoeven (Universiteit Antwerpen & London, City)
Prof. dr. Hugo Quené (Universiteit Utrecht)
Dr. Willemijn Heeren (Universiteit Leiden)
Dr. Margreet Langereis (Radboud UMC, Nijmegen)

Age and hearing-loss effects on speech processing

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. Krieken,
according to the decisions of the Council of Deans
to be defended in public on Tuesday, September 18, 2018
at 14:30 hours

by

Xaver Koch

born on 2 September 1979
in Dresden, Germany

Promotor:	Prof. dr. Mirjam Ernestus
Co-Promotor:	Dr. Esther Janse
Doct. Thesis Committee:	Prof. dr. Roeland van Hout (chair)
	Prof. dr. Johan Verhoeven (Antwerp University & London, City)
	Prof. dr. Hugo Quené (Utrecht University)
	Dr. Willemijn Heeren (Leiden University)
	Dr. Margreet Langereis (Radboud UMC, Nijmegen)

Contents

General Introduction

Most psycholinguistic theories and models are based on experimental findings collected under optimal and well controlled test conditions with a specific test population (young, healthy, well-educated university students). Given present and anticipated demographic developments there is a growing research interest in aging and its effects on cognition and communication. In the next sixty years, the proportion of adults in Europe aged over 65 years will rise further from the current 19 percent to 28 percent of the total population (i.e., 100 billion individuals) and there will be more than twice as many adults older than 80 years than at present (Eurostat, 2017; Kontis et al., 2017)

Epidemiological findings (e.g., Lin et al., 2013) show an exponential relationship between age and the prevalence of hearing loss. This implies that in the future, a larger proportion of the population will not only be older, but will also have reduced hearing abilities. This combination may impact not only on speech comprehension, but also on speech production. Theories of speech production (e.g., the Hierarchical State Feedback Control Model; cf. Hickok, 2012) and speech motor control (e.g., the Direction Into Velocity of Articulators Model; cf. Tourville and Guenther, 2011) have argued that auditory feedback (i.e., hearing back your own speech) plays a role in speech production in that speech production targets may be defined perceptually. As such, corrupted auditory feedback due to hearing loss potentially affects articulation of speech sounds. Age-related sensory decline, particularly in hearing, obviously impacts on ease of spoken communication for older adults. Furthermore, older age may negatively impact on speech processing through effects of cognitive decline, i.e., a slowdown of processing, but also reduced working memory capacity and decreased inhibitory control (cf. Thornton and Light, 2006; but also cf. Ramscar et al., 2014).

If current psycholinguistic theories of speech processing are based on young well-educated listeners, the question arises whether and how such theories should be modified in order to accommodate and predict speech processing in older listeners for whom sensory input as well as the processing of that input deviates from that of young students. Research with older participants with or without hearing loss may provide insights into how existing models of speech comprehension and production should be modified, analogous to the way in which the study of language processing in aphasic patients contributed substantially to the emergence of the field of psycholinguistics in the 19th century (cf. Wernicke, 1874; Lichtheim, 1885). Moreover, research into the effects of aging and hearing loss may help to face the challenges of ongoing demographic change, e.g., by improving diagnosis and programs for hearing-aid fitting as well as optimization of working conditions (cf. Charness and Bosman, 1994).

Hearing loss and cognitive decline are not only age-related but also independently associated (cf. Uhlmann et al., 1989; Fortunato et al., 2016; Lin et al., 2013; Pichora-Fuller, 2003; Baltes and Lindenberger, 1997; Lindenberger and Baltes, 1994). Thus, independent of age, increased hearing loss is often associated with poorer cognitive function. A number of (mutually non-exclusive) hypotheses have been proposed to account for this correlation between cognitive abilities and hearing acuity in older adults. These hypotheses are relevant to the work in this dissertation and will therefore be briefly introduced here, although the work reported here was not directly aimed at proving or disproving these assumptions.

First, the ‘common cause’ hypothesis assumes that hearing loss and cognitive decline share their etiology. In fact, diseases such as diabetes and hypertension may trigger both hearing decline and cognitive deterioration. In contrast, the ’deprivation hypothesis’ supposes that cognitive decline may be an indirect consequence of hearing loss. According to this ‘deprivation hypothesis’, hearing loss causes withdrawal and avoidance behavior, with reduced social participation leading to decreased cognitive challenges (e.g., Bassuk et al., 1999). The results of Amieva and colleagues (2015), showing that hearing aid use attenuates cognitive decline, are also in line with the notion that hearing acuity plays a central role in cognitive decline. As such, successfully restored hearing provides the listener with the necessary amount of speech input to keep up social interactions and may therefore preserve cognitive functioning. Third, the ‘information degradation hypothesis’ claims that peripheral auditory decline may directly impede cognitive processing, rather than indirectly as in the previous deprivation hypothesis. The direct effect of auditory decline on cognitive processing proposed by this ‘information degradation hypothesis’ is a reversible effect that dissolves if auditory information is restored. A fourth account of the association between hearing loss and cognitive decline in aging is the ‘cognitive load on perception hypothesis’. This account emphasizes the effects of imperfect listening situations on speech comprehension. Cognitive load, caused by difficult listening situations (e.g., reverberant room, background conversations, traffic noise) may impede speech processing. That is, the individual cognitive capacity may limit speech comprehension and/or increase the subjective listening effort in challenging listening situations, especially in older (hearing-impaired) adults.

Listening can be effortful for several reasons. Both reduced abilities of the individual listener (e.g., hearing loss, cognitive limitations) as well as (external) signal degradation or signal transmission issues (e.g., noise, reverberations, conversational speech, high speech rates) are included in Mattys et al.’s (2012) rather broad definition of adverse or ‘effortful’ listening conditions. In their ‘Framework for understanding effortful listening' (FUEL) Pichora-Fuller and colleagues (2016) describe how input-related demands, motivation, and cognitive capacity interact in understanding the speech message. The FUEL account (ibid.) is an adapted version of Kahnemann's (1973) 'Capacity Model of Attention'. The FUEL account assumes that arousal and cognitive capacity are associated in the sense that more capacity (i.e., processing resources) can be made available if arousal is higher. Accordingly, challenging (i.e., more effortful) listening situations would require increased arousal. The need for sustained arousal may explain listening fatigue symptoms by (older) hearing-impaired listeners in challenging listening situations. As such, adverse conditions offer "an ideal ground for examining speech and cognitive processes in combination" (Mattys et al., 2012, p. 967).

Models of spoken word recognition generally assume that presentation of speech input yields activation of matching sublexical and lexical representations, followed by selection of the best matching lexical candidate. Competition-based models of spoken word recognition, such as the Cohort Model (Marslen-Wilson and Komisarjevsky Tyler, 1980; Marslen-Wilson and Welsh, 1978), TRACE (McClelland and Elman, 1986), the Neighbourhood Activation Model (Luce and Pisoni, 1998), or Shortlist (Norris, 1994), would supposedly model the effect of hearing loss as resulting in less definite and/or delayed activation of sublexical and lexical representations upon auditory word presentation. It is less clear whether and how individual listeners' cognitive abilities, such as information processing speed or working memory, should be incorporated in models of spoken word recognition. Whereas most of the research literature linking cognitive skills to speech perception has focused on the perception of carefully read aloud sentences presented in noise (cf. Dryden et al., 2017 for a recent meta-analysis), the studies described in this thesis attempt to identify which individual abilities (such as cognitive or linguistic skills) predict speech-processing ability *across a range of speech materials* (from read speech to conversational speech) and in *different listening conditions*. In doing so, this thesis will provide stepping stones for refinement of models of spoken-word recognition. Similarly, by investigating whether individual hearing acuity and cognitive abilities are related to speakers' acoustic realization of speech sounds, this thesis contributes to elaboration of models of speech production.

The first part of this thesis contains two studies investigating the effects of several common adverse conditions on speech processing in listeners of varying age, namely the effects of reduced intelligibility of conversational speech, increased speech rate, and noise masking. To start with conversational speech, reductions and elisions as well as higher speech rates may account for reduced intelligibility for this type of speech material, as compared to read aloud speech (Assmann and Summerfield, 2004). Higher speech rates (as may be found in conversational speech) have been shown to make speech comprehension more challenging, particularly for older adults (the so called age × speech rate interaction effect; cf. Wingfield, 1996; Gordon-Salant and Fitzgibbons, 1999; Gordon-Salant et al., 2014; but cf. Schneider et al., 2005 and Gordon et al., 2009). This finding is supported by older adults' subjective reports of increased listening effort for fast speech. Moreover, it has been shown that speech rate effects are more pronounced if hearing loss is more severe (Wingfield et al., 2006). This age by speech rate finding is compatible with the 'information degradation hypothesis', as well as with the 'cognitive load on perception hypothesis'. Relatedly, the size of the speech rate effects in older adults has been shown to be associated to their speed of processing (Janse, 2009) – a result in line with the 'cognitive load on perception hypothesis'. However, it should be noted that most studies showing that older adults are differentially impacted by increased speech rate manipulated speech rate artificially. This is obviously a controlled way to manipulate rate, but time-compressed speech is processed differently from speech that is spoken fast (Janse, 2004). Chapter 2 therefore investigates the effect of increased speech rate as observed in fragments of conversational speech. The question is whether older adults are more impacted by increased speech rate in natural fragments than younger adults, and which individual sensory and cognitive abilities are associated with the added difficulty of the increased rate.

Chapter 3 focuses on the concept of listening effort in a different, more subjective way by investigating how much background noise listeners are willing to accept while listening to speech. This chapter is inspired by the fact that hearing aid success has been shown to be related to listeners' subjective, rather than objective, evaluation of how the device performs in noise. If we could identify problematic hearing-aid users before their hearing aids end up 'in the drawer', counseling could be attuned to set more realistic expectations and possible supply additional assistive listening devices. The puzzling finding that objective recognition of speech in noise cannot predict hearing-aid success led to the hypothesis that the subjective measure of perceived listening effort might contribute to the listener's satisfaction with the device (cf. McGarrigle et al., 2014). Obviously, such subjective measures seem to capture aspects of listening effort that objective measures do not cover (cf. Ohlenforst et al., 2017). The Acceptable Noise Level (ANL) Test (e.g., Nábělek et al., 2006), as a prime example of such a subjective measure, quantifies how much background noise listeners are willing to accept while listening to speech. The adverse condition present in the ANL test is a typical type of background noise (i.e., energetic masking) that approximates the acoustics of multitalker babble in a crowded room. In order to find out more about what it is the ANL may be capturing about listening effort, the question here was whether ANL outcome was sensitive to the type of speech material participants are presented with. Most previous ANL studies quantified listening effort using carefully-pronounced speech rather than everyday conversational materials. Does subjective listening effort, as quantified by the level of background noise listeners are willing to put up with during listening, differ between carefully pronounced read speech and conversational fragments?

Whereas the first part of this thesis (Chapters 2 and 3) focuses on the consequences of aging and age-related hearing loss on ease of spoken language understanding, the second part (Chapters 4 and 5) presents two studies that investigate the effects of age and hearing loss on speech production. This research contributes to the topic of the nature of the link between speech perception and speech production. Probably the most well-known example demonstrating how speech perception and speech production are intertwined is the Lombard effect (Lombard, 1911). Speech produced in loud background noise, that is if auditory feedback on one's own speech is missing or weakened, is characterized by longer segment durations and increased intensity compared to speech production in quiet. Two mutually non-exclusive accounts have been put forward to explain the Lombard phenomenon. First, speakers may increase their loudness in background noise to be able to hear themselves sufficiently. Alternatively, by way of switching to Lombard speech, the speaker may adapt to the adverse listening situation in order to optimize intelligibility for the listener (cf. Lane and Tranel, 1971). Although the second account seems compelling, speakers also produce Lombard speech in non-communicative tasks such as reading word lists, thus without an interlocutor being present. Moreover, is has been shown that the Lombard effect is robust and cannot easily be inhibited, which suggests that the effect is an automatic response, or reflex, to suboptimal auditory feedback rather than a reflection of volitional audience design (cf. Pick et al., 1989). The evidence for links between speech acoustics and the availability of auditory feedback in adults, as demonstrated by the Lombard effect, motivates our investigations into effects of hearing loss acquired later in life on speech production.

The classical 'perceptual loop theory' (Levelt, 1989) as well as the more recent 'Hierarchical State Feedback Control model' (Hickok, 2012; Guenther and Hickok, 2015) discuss the role of feedback systems for speech production. The perceptual loop theory postulates that speech is monitored via an internal and an external loop, corresponding to monitoring of speech that has not been and that has been realized overtly, respectively. Both loops are implemented to monitor and correct errors during speech production. Starting from a conceptual stage with the intended message being generated, information is fed forward to a lemma level which in turn activates the phonological level (generation of the phonological word). The necessary motor commands for the encoded phonological units are subsequently executed by the articulators. The internal loop is used to compare the assembled phonological word (cf. Levelt et al., 1999; Wheeldon and Levelt, 1995) with the phonological information of the intended lemma. An external loop is employed to compare the speech acoustics against an auditory target associated with the planned lemma.

The main difference between Levelts' 'perceptual loop theory' (Levelt, 1989) and Hickok's (2012) 'Hierarchical State Feedback Control' (HSFC) model of speech production is that the latter model builds on theories of general action control, as applied to e.g., movement of limbs. As such, the HSFC model provides more detail on the speech motor aspects of speech production. Similar to Levelts' model the HSFC model assumes internal and external feedback loops. In contrast to more limited models which exclusively zoom in on the speech motor control aspect of speech production such as DIVA (e.g., Tourville and Guenther, 2011), Hickok thus proposes a typical psycholinguistic model architecture with message formulation as the starting point for speech production. In general, state feedback control models of action control (such as the HSFC) involve the prediction of the consequences of an action (e.g., movement, speech). These predictions (or efference copies) can be compared against the actual sensory feedback to monitor and correct deviations from the intended movement. For speech production Hickok (2012) hypothesized two components of the prediction-feedback system: a primary auditory phonological system and a secondary motor-somatosensory system (somatosensory information being the internal information about where one's articulators are). The auditory phonological system is directly activated via the lemma level and subsequently activates the somatosensory system.

Both systems together constitute the internal model (prediction) of a word or an utterance. That is, before we utter a word we activate an auditory model of what the word will sound like and a somatosensory model of how it should feel to pronounce the word (e.g., two times lip contact and two jaw opening and closing gestures for the word 'pepper'). As this thesis is concerned with effects of hearing loss on speech production, this thesis focuses on the auditory model in speech production.

How does the HSFC system detect speech errors or pronunciations deviating from the prediction? If the speech motor selection routine produces an error (e.g., planning articulatory movements for 'ship' instead of 'sip'), this deviation from the internal model can be mirrored (and almost immediately be corrected) via the primary auditory phonological system which is connected to the somatosensory system. If, however, the error is not detected internally and it is only the auditory feedback of the speech output that does not match the prediction, this is monitored via the second feedback system: the external loop. The role of the external loop is most salient in first language acquisition. During the language acquisition period, external feedback is used to learn the relationships between the motor commands and their (auditory and somatosensory) consequences. In other words, the internal model is built up and shaped through external auditory feedback. Consequently, hearing loss during the phase of language acquisition obviously leads to deviations from normal speech production (e.g., Plant and Hammarberg, 1983). External feedback may also be used later in life to update our internal model (predictions) in the sense of a constant recalibration process. In fact, a number of studies demonstrate that participants instantaneously adapt their speech production to counteract acoustically manipulated feedback (cf. Houde and Jordan, 1998; Purcell and Munhall, 2006), even though individuals differ in the way and the extent to which they do so (Lametti et al., 2012; Schuerman, 2017). Thus, representations (the internal model in Hickok's 2012 model) may be consolidated or updated throughout the life span depending on the available (external) acoustic feedback of one's own speech and auditory input from speech of others. Impoverished or altered auditory feedback and impoverished auditory input may over time gradually affect the internal model, resulting in deviations of speech production, relative to normal hearing.

Indeed, adults who only became deaf after having acquired language while relatively normal-hearing as a child (i.e., post-lingually deafened), show deviant consonant and vowel productions, as well as reduced acoustic contrasts in comparison to normal-hearing participants (e.g., Waldstein, 1990; Schenk et al., 2003; Lane and Wozniak Webster, 1991; Lane et al., 2007). This indicates that missing auditory information, including missing auditory feedback from one's own speech, may change long-term speech representations and thus feedforward commands. In Chapter 4 we address the question of whether (mild forms of) age-related hearing loss may already affect speech production in a non-clinical population. Even though individuals with hearing loss can still rely on their knowledge of what production of speech sounds feels like (i.e., somatosensory information), the question is whether somatosensory feedback alone will be sufficient to keep articulation (feed-forward commands) stable in mild forms of hearing loss. Earlier studies on hearing loss acquired later in life always focused on 'severe cases', also because of reports of maintained speech intelligibility of post-lingually deafened adults. The study reported in Chapter 4 investigates consequences of subtler degrees of (age-associated) hearing loss on speech production, in order to get at the earliest stages of change in production.

Apart from effects of age-related hearing loss on production, advanced adult age may have its effects on (the stability of) speech motor commands (cf. Krampe, 2002; Mefferd and Corder, 2014; Tremblay et al., 2013; Bilodeau-Mercure et al., 2014). Additionally, or possibly related to age effects on motor control, it is conceivable that age-related cognitive decline, such as for example in working memory or in executive functions, may negatively affect speech articulation. As such, speech articulation accuracy may partly rely on memory ability. To compare one's own incoming speech to the auditory representations (i.e., the auditory subsystem of the internal model; cf. Hickok, 2012) via external feedback, one must keep the prediction and realized signal in auditory memory. If this comparison between realization and predicted outcome is less successful, speech sound representation may become less well-defined and articulation may become sloppy (Neger et al., 2015). Therefore, aging effects on speech articulation and resulting acoustics may arise through age-related sensory effects as well as age-related cognitive effects. However, age-related effects of cognitive decline on speech production have rarely been investigated (cf. review by Torre and Barlow, 2009). This thesis will try and tease apart potential relationships between speech acoustics and hearing acuity on the one hand, and acoustics and cognitive aspects of aging on the other.

As effects of age and age-related hearing loss on acoustic realization of speech are likely to be subtle, the focus will be on specific speech sounds. Sibilant fricatives are prime candidates as target sounds for acoustic analyses because their perception is most likely to be affected by age-related hearing loss, which starts off in the high frequencies. Sibilants also require precise articulatory control which may decline in older adulthood as indicated by studies demonstrating deteriorated fine motor control in aging (cf. Krampe, 2002; Mefferd and Corder, 2014). Thus, if age effects on hearing, speech motor control or on cognitive abilities alter the acoustic realization of sounds, these effects are most likely to be found for sibilant fricatives, which will be investigated in Chapter 4.

Chapter 5 investigates whether and how speech production changes in post-lingually deafened novice cochlear implant (CI) users after implantation. As these patients have often been hearing-impaired for years prior to implantation, this population offers an ideal test bed to investigate long-term effects of hearing loss on speech production prior to CI surgery, as well as flexibility of the speech processing system following cochlear implantation. If prolonged hearing loss prior to cochlear implantation impacts on speech production accuracy, this would provide empirical support for models of speech production which postulate an outstanding role of the auditory domain for the representation of speech (e.g., Hickok, 2012). The study in Chapter 5 also complements studies on short-term effects (i.e., within the experimental session) of changed auditory feedback on speech production in adults (cf. Houde and Jordan, 1998; Jones and Munhall, 2000; Purcell and Munhall, 2006), showing that adults instantaneously counteract such auditory manipulations. The time course of potential changes in production is important because of earlier suggestions that the speech of novice CI users might undergo an initial phase of speech deterioration directly after CI implantation, followed by a return to baseline after several weeks (Lane et al., 2007). Consequently, Chapter 5 presents a longitudinal study with post-lingually deafened adult CI candidates plus an age- and gender-matched control group, who repeatedly read carrier sentences containing target words with sibilants and vowels.

Research Questions and Outline

The first part of this thesis (Chapters 2 and 3) focuses on the effect of adverse listening conditions on speech perception for different listeners. More specifically, the studies in this part address whether and how age and age-associated changes in hearing and cognitive performance affect the perception of speech in ecologically valid listening conditions. The second part of this thesis consists of two studies (Chapters 4 and 5) examining the effects of age and hearing loss on adult speech production. Below, the objectives for each experimental chapter are summarized.

Chapter 2 investigates the effect of speech rate on spoken word recognition across the adult age range. The study addresses the question whether the common finding that older adults show stronger speech rate effects, compared to younger adults, if tested with time-compressed speech, can be replicated using conversational speech materials. Possibly, older adults' increased difficulty with fast speech should be attributed to the acoustic artifacts of artificial time compression rather than to increased speech rate per se. Conversely, even though natural fast speech does not contain acoustic artefacts, the high speech rate, in combination with acoustic reduction and hence reduced signal redundancy, may still impact older adults more than younger adults. The study presented in Chapter 2 therefore investigates effects of natural variation in speech rate as observed in fragments of conversational speech on speech perception in adult listeners of varying age. A number of cognitive predictors (including information processing speed), hearing thresholds and vocabulary size serve as covariates to evaluate whether hearing loss, linguistic and cognitive capabilities predict the impact of increased speech rate on speech processing of conversational materials. The study makes use of the visual-world eye-tracking paradigm. Participants listen to sentences and indicate (by way of a click response) which out of four visually presented words they heard in the presented speech fragment. This experimental method is analyzed to yield three different dependent variables. Firstly, the method yields the offline measure click-response time to target words presented in context, secondly, it provides a time-continuous measure of gaze behavior and thirdly pupil size data over time are collected. The combination of methods employed in Chapter 2 allows us to investigate the effects of speech rate for different listeners of varying adult age on both early (continuous eye gaze data) and later stages of speech processing (click response times), as well as on listening effort (as indexed by pupillometry).

The study in Chapter 3 focuses on material effects for the Acceptable Noise Level (ANL) test, a procedure argued to measure subjective listening effort. The study is set up to answer the question whether ANL results obtained with conversational materials differ from those obtained with standard audiological stimuli. More specifically, we ask whether meaningless or incoherent speech materials, which are often used in the clinical setting, yield differential ANL test outcomes than more ecologically valid conversational materials. In addition, in order to be able to relate the measure to cognitive listening effort, we investigate whether the finding that the ANL is associated with working-memory can be replicated (cf. Brännström et al., 2012). The ANL test has been used to predict hearing-aid success by quantifying the amount of noise a participant is willing to accept while listening to speech. However, previous studies have used a range of test materials and results regarding the predictive value of the ANL thus far are mixed. The study on the ANL material effects is designed as precursor to a clinical study with hearing-impaired participants. Therefore, a normal-hearing participant sample with an age range representative for hearing aid users is tested.

Chapter 4 presents a study investigating the research question whether age and/or age-related sensory and cognitive decline are associated with speech production. Sibilant sounds are chosen as speech production targets because their realization, which is characterized by a concentration of energy in the high-frequency domain, is most likely to be affected by age-related (high-frequency) hearing loss. In order to dissociate age-related sensory effects on speech acoustics from other age-related effects, we include indices of cognitive abilities as potential predictors of speech production. To test for possible effects of age-related high-frequency hearing loss on sibilant realization (operationalized as the acoustic measure Center of Gravity), a sample of participants spanning a wide adult age range read carrier sentences containing target words with word-initial sibilant sounds.

The production experiment in Chapter 5 represents a longitudinal clinical study. The study addresses the question whether prolonged hearing loss in cochlear implant candidates has affected acoustic realization of sibilant fricatives and vowels compared to acoustic realization by an age- and gendermatched control group. The CI group is recorded once before implantation and twice (two weeks, three months) after activation of the CI. Participants in the control group are recorded three times as well times, with the same amount of time elapsing between test sessions. By including multiple test points for patient and control group, vowel formants (F1, F2) and sibilants' Center of Gravity are analyzed as dependent variables taking into account effects of repeated testing. The study also investigates whether cochlear implantation affected acoustic realization by comparing pre-implant and post-implant realization, and contrasting the changes over time following cochlear implantation to changes over time due to multiple testing in a control group. The study thus addresses long-term effects of prolonged hearing loss on speech production prior to cochlear implantation, as well as effects of cochlear implantation. This line of research complements speech production experiments with short-term auditory feed-back manipulations (i.e., within the experimental session), which have shown that adults instantaneously counteract such auditory deviations.

Speech rate effects on the processing of conversational speech across the adult life span

Chapter 2

This chapter is based on:
Xaver Koch and Esther Janse (2016)
Speech rate effects on the processing of conversational speech across the adult life span
Journal of the Acoustical Society of America, 139(4), pages 1618–1636

This study investigates the effect of speech rate on spoken word recognition across the adult life span. Contrary to previous studies, conversational materials with a natural variation in speech rate were used rather than lab-recorded stimuli that are subsequently artificially time-compressed. It was investigated whether older adults' speech recognition is more adversely affected by increased speech rate compared to younger and middle-aged adults, and which individual listener characteristics (e.g., hearing, fluid cognitive processing ability) predict the size of the speech rate effect on recognition performance. In an eye-tracking experiment, participants indicated with a mouse-click which visually presented words they recognized in a conversational fragment. Click response times, gaze and pupil size data were analyzed. As expected, click response times and gaze behavior were affected by speech rate, indicating that word recognition is more difficult if speech rate is faster. Contrary to earlier findings, increased speech rate affected the age groups to the same extent. Fluid cognitive processing ability predicted general recognition performance, but did not modulate the speech rate effect. These findings emphasize that earlier results of age by speech rate interactions mainly obtained with artificially speeded materials may not generalize to speech rate variation as encountered in conversational speech.

1.1 Introduction

Older adults, particularly those who are hearing impaired, report that they face challenges in speech comprehension in adverse listening conditions, such as when there is background noise or talkers have accents, mumble, speak softly or rapidly. The effect of increased speech rate on older adults' speech comprehension performance has often been operationalized by using artificial time compression, which may approximate some of the difficulties reported with fast speech (e.g., Wingfield, 1996; Vaughan et al., 2006). Several studies have shown that artificially time-compressed speech makes comprehension and recall more difficult than normal-rate speech, and that this speech rate effect is larger for older, compared to younger adults (Wingfield, 1996; Gordon-Salant and Fitzgibbons, 1999; but cf. Schneider et al., 2005 and Gordon et al., 2009). Furthermore, speech rate effects seem to interact with the linguistic characteristics of the presented stimuli. Wingfield and colleagues (2003) have found that for older adults increased speech rate made listening particularly challenging if the presented sentences were also syntactically complex.

Before we provide a more detailed account of the literature on this finding that the effect of increased speech rate is larger for older than younger adults (henceforth, the *age × speech rate* interaction), we raise the point that results obtained with artificial time compression may either underestimate or overestimate the difficulty that listeners experience with naturally produced fast speech. Schmitt and Moore (1989) compared comprehension performance for time-compressed versus naturally produced faster speech rate in older adults. Their results showed generally better comprehension scores for naturally speeded up or slowed down materials than for unselectively compressed/expanded speech, suggesting that artificial time compression presents a more difficult listening condition than naturally increased speech rate. In contrast, a recent study (Gordon-Salant et al., 2014) has shown that the recognition of artificially time-compressed read sentences seems to overestimate the recognition of natural fast-rate speech (see also Janse, 2004). Gordon-Salant and colleagues found that both younger and older adults showed better sentence recognition performance for artificially speeded speech (originally read at a normal rate) than for natural fast-rate sentences read aloud by a talker at a very fast rate. However, what may be crucial is whether instructing talkers to read out sentences at their ceiling rate (as in Gordon-Salant et al., 2014 and Janse, 2004) is representative of rate variation as observed in conversational speech in which speakers themselves habitually speak or choose to speak at a particular rate. Unlike artificially time-compressed speech, instructing talkers to speak as fast as they can generally involves less clear articulation because most speakers are only able to speed up their speech rate through reduction of segments and syllables. The present study aims to investigate how naturally varying speech rate, as encountered in conversational speech materials spoken by different speakers, affects listening performance in younger, middle-aged and older adult listeners.

We now return to the accounts that have been provided for the *age* × *speech rate* interaction finding (as observed with artificially speeded speech) introduced above. Several studies have provided explanations for this differential rate effect on older adults' comprehension or recall performance interaction (e.g., Wingfield et al., 1999; Schneider et al., 2005). A first account for older adults' problems with speeded speech is the 'generalized slowing hypothesis', which is based on cognitive aging research (e.g., Cerella, 1990). Salthouse (1985, 1996) proposed that a reduction in processing speed leads to impairments in cognitive functioning ('processing-speed theory' of cognitive aging). A general slowing of brain functions in aging and thus a reduced processing speed will lead to comprehension problems if more information units are transmitted per unit of time than the processor can handle (Wingfield, 1996). Importantly, an individual's processing speed predicted the effect of speech rate on older listeners' performance in a study by Janse (2009) using artificially speeded speech. If domain-general slowing should be held responsible for older adults' problems with fast speech rates, then increased rates of visual text presentation can be expected to also differentially affect older adults, compared to younger adults. However, this was not the case in a study by Humes and colleagues (2007). In their study, effects of increased rate of visual presentation were similar for younger and older adults.

Age-related changes in hearing have been put forward as another possible explanation for the increased problems older adults may have with fast speech. Epidemiological data suggest that around 40 to 50 percent of the population aged between 50 and 90 years are affected by hearing decline defined as pure-tone average thresholds (averaged over 0.5, 1.0, 2.0, and 4 kHz) above 25 dB HL (Cruickshanks et al., 1998). Hearing impairment and age were found to independently contribute to deficits in recognizing temporally manipulated speech (Gordon-Salant and Fitzgibbons, 1993).

A third account for the *age* × *speech rate* interaction is that auditory processing ability may be impaired in older adults. Thus, apart from a gradual decline in absolute hearing sensitivity particularly for the higher frequencies, aging is accompanied by problems with central hearing, such as changes in temporal processing (Fitzgibbons and Gordon-Salant, 2010). Relatedly, older adults' problems with fast speech have been linked to longer neural adaptation periods in older listeners. Longer adaptation processes in older adults, as evidenced by e.g., higher gap detection thresholds in older than in younger adults (Gordon-Salant et al., 2006; Pichora-Fuller et al., 2006; Haubert and Pichora-Fuller, 1999), may negatively influence the perception of stop consonants in fast speech. In line with this auditory processing account, Schneider and colleagues (2005) argued that older adults process artificially time-compressed speech differently from younger listeners. Schneider and colleagues base their 'perceptual hypothesis' on the "notion that older adults find it more difficult to handle speed-induced acoustic distortions than do younger adults" (ibid., p. 268), thereby arguing for age-related differences in sensitivity to signal manipulations, such as artificial time compression. Schneider and colleagues (2005) compared the effects of a linear type of time compression (eliminating every third amplitude sample, the sampling method) and a selective time compression method that particularly compresses steady-state segments and leaves rapid transitions intact. Indeed, Schneider and colleagues' (2005) results showed that younger and older adult groups were equally affected by increased speech rate when speech was speeded in a way that produced minimal acoustic degradation.

More evidence for acoustic degradation induced by artificial time compression algorithms comes from Kusomoto and Vaughan (2004), who compared acoustic features of artificially speeded-up (Synchronous-OverLap-Add technique) and natural speech. Their results suggest that for higher compression rates durational cues for plosive and fricative consonants may differ from natural speech. As durational cues are exploited in speech perception (e.g., Klatt, 1976; Raphael and Dorman, 1980), artificial speeding techniques may complicate speech processing, particularly at higher compression rates. Thus, artificial time compression changes perceptually relevant durational cues, which impairs speech comprehension, and this effect may be more pronounced for older than younger listeners (e.g., Goy et al., 2013).

In sum, studies on age and individual differences in the effect of speech rate on speech perception so far have mainly focused on artificially time-compressed speech. Moreover, most studies have focused on sentences that were read aloud. Importantly, Wingfield et al. (1999) state that recall of auditorily presented speech passages drops significantly if the presented speech rates exceed "normal limits" (ibid., p. 385), particularly for older adults. Gordon and colleagues (2009) also state that *age* × *speech rate* interaction effects usually occur if materials are speeded to rates beyond those found in normal speech.

This raises the question as to which speech rates can be considered 'normal' and what is a 'normal' range? Speech rate is operationalized as the number of linguistic units (e.g., words, syllables, phones) per unit of time (e.g., minute, second). In contrast to 'articulation rate', 'speech rate' includes pauses. Krause and Braida (2004) state that clear speech involves speech rates of about 100 words per minute (wpm, i.e., 2.3 syll./s.[1]) and that conversational speech would easily involve a doubling of that tempo (i.e., 4.6 syll./s.). Greenberg's (1998) study of a spontaneous English dis-

---

[1] Taking Lamel et al.'s (1989) formula to convert words per minute into syllables per second (syll./s.) or minute.

course corpus showed a mean syllable duration of around 200 ms, i.e., an articulation rate of 5 syllables per second. For Dutch, Quené (2008) found a mean articulation rate of about 4.2 syllables per second in the interview part of the Spoken Dutch Corpus (Oostdijk, 2000). The unit of measurement in Quené (2008) was interpause chunks. The fastest speaker in this sample had a mean articulation rate of about 5.6 syllables per second and the slowest speaker a rate of 3.0 syllables per second. The highest articulation rate Quené (2008) found in an interpause chunk was 12.1 syllables per second (personal communication, August 26, 2014). In sum, a speech rate of about 4 to 6 syllables per second can be assumed typical for conversational speech in West Germanic languages such as English or Dutch. Speech rates roughly range between around 2 and 12 syllables per second. The *age* × *speech rate* interaction effect found by Janse (2009), for example, is based on the comparison of a rate that is 1.5 times normal rate (i.e., given that the normal rate in that study was 5.7 syllables per second, 1.5 × 5.7 syll./s.=8.6 syll./s.) and a rate that was twice the normal rate (i.e., 2.0 × 5.7 syll./s.=11.4 syll./s.). Both time-compressed conditions therefore, represent higher-than-typical speech rates. Speech rate studies have worked with higher-than-typical rates, and artificially speeding speech changes perceptually relevant durational cues (cf. Kusomoto and Vaughan, 2004). This raises the question whether experimental results obtained with artificial time compression generalize to processing of natural speech heard in everyday conversations. The present study therefore investigated how natural speech rate variation as found within and between speakers in a conversational speech corpus affects listening performance in adults of varying age (cf. Gordon et al., 2009).

As hypothesized by the perceptual and generalized slowing accounts of the *age* × *speech rate* interaction, the effect of speech rate on speech comprehension may interact with the listener's auditory, linguistic and cognitive abilities. We therefore included these participant-related variables into our modeling of perceptual performance. We investigated speech processing by employing the visual-world paradigm. This technique provides information on the time course of the recognition of a word embedded in a running sentence and yields complementary behavioral (click response times) and psychophysiological data (gaze data, pupil size data). Eye-tracking allows us to observe speech processing in real time as there "is no appreciated lag between what is fixated and what is processed" (Just and Carpenter, 1980, p. 331). The task-evoked pupil response reflects the cognitive demands of processing a stimulus (Zekveld et al., 2013). Speech rate is expected to affect ease of processing, and hence understanding faster stimuli is cognitively demanding. Cognitive demand affects the pupil response (e.g., Zekveld et al., 2013). We therefore hypothesized that processing effects that are related to increased speech rate should be reflected in click response times, gaze data and in the task-evoked pupil response.

We address the following three research questions:

1. Can we replicate speech rate effects on word recognition performance using conversational materials with naturally varying speech rates?
2. Do younger adults, middle-aged adults, and older adults differ in the effect of speech rate on their word recognition performance?
3. Which individual measures predict general word recognition performance and the effect of increased speech rate on recognition performance over the adult life span?

1.2 Method

1.2.1 Participants

Three age groups were included: older adults (aged over sixty years), middle-aged adults (between 30 and 60 years), and younger adults (between 18 and 30 years). None of the participants reported hearing difficulties. From the initial sample of 112 adults, 12 participants were excluded from the analyses for the following reasons. The semi-automatized eye-tracking calibration procedure was not successful for two participants (one older and one younger adult). The test session of one middle-aged participant was interrupted by construction noise. Furthermore, eight participants were excluded (seven older adults and one middle-aged) because hearing loss in one or both ears exceeded the Dutch prescription criterion for hearing aids (pure-tone average over 1, 2 and 4 kHz ($\mathrm{PTA}^{\mathrm{high}}$)>35 dB HL). One additional older adult was excluded because of very low task accuracy (less than nine percent of all 60 trials correct) while accuracy for the remaining participants ranged between 77 and 100% correct (M=97.1%, SD=3.3, see Analyses). The final sample consisted of 100 Dutch participants, 32 older adults (M^{age}=67 years, SD=4.7, 20 females), 33 middle-aged adults (M^{age}=50 years, SD=7.5, 21 females) and 35 younger adults (M^{age}=21 years, SD=2.5, 22 females).

1.2.2 Background measures

Participants' hearing was screened in both ears with air conduction pure-tone audiometry using the Hughson-Westlake procedure (Carhart and Jerger, 1959) for octave frequencies from 0.25 to 8 kHz, including two half-octave frequencies of 3 and 6 kHz, see Figure 1.1.

Figure 1.1: Mean audiometric pure-tone air conduction thresholds (for left and right ear) as a function of frequency for the younger, middle-aged and older adults. Error bars represent standard errors.

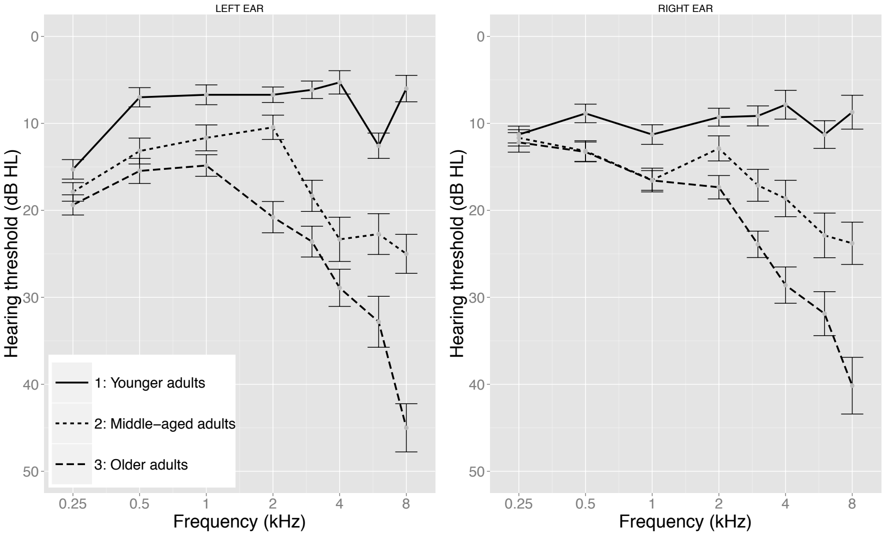


Table 1.1: Means and standard deviations of pure tone average measures in the better ear for younger adults (YA), middle-aged adults (MA) and older adults (OA) and results of test statistics investigating age group differences in pure-tone average measures (Mann-Whitney-Wilcoxon, significance levels corrected for multiple testing).

	Age group			Comparisons		
	YA	MA	OA	YA–MA	YA–OA	MA–OA
Hearing variable	*M(SD)*	*M(SD)*	*M(SD)*	*p*	*p*	*p*
PTA$^{\text{low}}$	8.62 (4.55)	13.54 (6.51)	13.65 (4.05)	***	***	ns
PTA$^{\text{high}}$	6.33 (4.53)	13.64 (6.71)	19.42 (5.48)	***	***	***
PTA$^{\text{HF}}$	7.14 (5.29)	20.86 (11.40)	32.76 (12.61)	***	***	***

Significance level notation: ***$p<.001$.

Audiometric thresholds for the better ear were entered as a covariate in our statistical modeling of word recognition performance. This was done as auditory presentation in the word recognition experiment was binaural: we assumed that hearing sensitivity in the better ear would at least partly compensate for hearing loss in the worse ear, such that taking the better ear, rather than the poorer ear, presents a conservative estimation of the effect of hearing loss on performance (cf. Chen et al., 2015). Four participants (one younger and three older adults) showed asymmetric hearing loss, defined as an interaural difference of more than 10 dB, averaged over 0.5, 1, 2, and 4 kHz (following Noble and Gatehouse, 2004). Table 1.1 lists descriptive and test statistics regarding the hearing sensitivity measures for the three age groups. Three different pure-tone average (PTA) measures were analyzed: (a) PTA^{low}: mean over 0.5, 1, and 2 kHz; (b) PTA^{high}: mean over 1, 2, and 4 kHz; and (c) high-frequency PTA (PTA^{HF}): mean over 3, 4, 6, and 8 kHz. Age groups particularly differed in the higher frequencies (cf. Table 1.1 for significant age group differences in PTA measures).

In addition to the assessment of hearing thresholds, all participants completed the following five tests: (a) a visual acuity test, (b) the Digit Symbol Substitution Test, (c) the vocabulary subpart of the Groningen Intelligence Test, (d) a visual Digit Span Test with backward recall, and (e) Raven's Standard Progressive Matrices Test. The five tests and the reasons for including them are described below.

1. Visual acuity test

Visual acuity was tested because all participants should be able to easily read the orthographic stimuli presented during the experiment (30 point Tahoma, i.e., approx. 0.8 cm height, see Procedure). Depending on whether participants wore their lenses or glasses during actual testing, their vision or corrected vision was tested to measure their (corrected) visual acuity. Acuity was assessed with the participant's head on a chinrest with constant 330 lux illumination. A standard Snellen visual acuity test chart was downscaled to be appropriate for the fixed test distance of 60 centimeter (being the fixed test distance during the eye-tracking experiment). Individual visual acuity was operationalized as the LogMAR equivalent (cf. Holladay, 1997) which is based on the logarithmic transformation of the Snellen fractions. Note that the LogMAR equivalent for normal vision is 0, with higher values representing poorer visual acuity. Mean visual acuity was 0.23 (*SD*=0.17) and ranged between 0 and 0.57. Crucially, all participants were able to correctly read the row with the largest font on the test chart which was half as large as the orthographic stimuli presented during the experiment (30 point Tahoma). As expected, visual acuity was poorer with higher age. All three age group comparisons showed significant age-related declines in visual acuity (cf. Table 1.2).

2. Digit-Symbol-Substitution Test

Participants' individual processing speed was assessed with the Digit-Symbol-Substitution Test (henceforth, DSST), which is a subpart to the Wechsler Adult Intelligence Test (2004). Salthouse (2000) found that scores on the DSST relate to processing and perceptual speed. Importantly, DSST performance was included as it predicted how much the individual listener was impacted by increased speech rate (Janse, 2009). Test performance was operationalized as specified in the test manual (number of correctly re-coded items within two minutes). Processing speed generally declines with age (Salthouse, 2000), which is also evidenced in our data (cf. Table 1.2).

3. Vocabulary Test

The vocabulary subpart measure of the Groningen Intelligence Test (Luteijn and van der Ploeg, 1983) was included as an index of individual linguistic ability to investigate whether word recognition, and the effect of speech rate on word recognition, is associated with vocabulary size. During the computerized multiple-choice test participants had to select correct synonyms for 20 words (choice out of four options for each word). There was no time pressure to complete the test. Test performance was operationalized as the number of correct responses. Younger adults showed poorer vocabulary scores than middle-aged and older adults (cf. Table 1.2).

4. Digit Span Test Backwards

Many studies have shown that recognition of spoken sentences in noise is associated with individual working memory ability, verbal working memory in particular (e.g., Rönnberg et al., 2013, 2008). Furthermore, Small and colleagues (1997) demonstrated that individual working memory capacity modulates speech rate effects on speech comprehension. We selected a digit span test with backward recall to tap simultaneous storage and manipulation of verbal information. A computerized visual version of the Wechsler Adult Intelligence Test (2004) digit-span test was administered. Participants had to recall 12 digit sequences after two practice trials. The digits in each sequence (two to seven items, increasing in length over trials) were presented one after another on a computer screen and participants were prompted to type in the digits in reverse order after presentation (digit-display time: 1000 ms, inter-stimulus interval: 200 ms). Individual performance was operationalized as the percentage of accurate trials. Middle-aged adults outperformed older adults in this task, but none of the other age group comparisons showed significant differences (cf. Table 1.2).

5. Raven's Standard Progressive Matrices Test

A test of non-verbal reasoning was included to investigate whether non-verbal intelligence (as opposed to verbal abilities measured by digit span performance) relates to speech processing performance. A modified version of the Raven's Matrices Test (Raven et al., 2003; henceforth, RAVEN) was administered in which a time limit was imposed to restrict the overall test session duration (cf. Wilhelm and Schulze, 2002). Participants were asked to complete as many items as possible within 10 minutes. Skipping items was prohibited. We modified the results form and enlarged the font sizes to 14 point as the original version had a rather small font size (9 point). The RAVEN score reflects the sum of correct responses for all five matrices sets. The maximal score that could be obtained was 60 (5 sets × 12 items). The results in Table 1.2 show that reasoning abilities differ between the age groups with younger participants outperforming the middle-aged and older groups.

Table 1.2: Means and standard deviations of non-auditory participant related variables for younger adults (YA), middle-aged adults (MA) and older adults (OA) and results of test statistics investigating age group differences (Mann-Whitney-Wilcoxon, significance levels corrected for multiple comparisons).

	Age group			Comparisons		
background	YA	MA	OA	YA–MA	YA–OA	MA–OA
variable	*M(SD)*	*M(SD)*	*M(SD)*	*p*	*p*	*p*
Visual acuity	0.10 (0.10)	0.23 (0.16)	0.37 (0.14)	.002**	***	***
Processing speed	87.26 (13.46)	76.12 (15.45)	64.56 (13.36)	.018*	***	***
Vocabulary	13.83 (2.04)	15.79 (1.60)	16.63 (2.06)	***	***	ns
Working memory	55.95 (18.81)	63.64 (23.83)	48.96 (17.93)	ns	ns	.012*
Reasoning	44.54 (5.60)	38.64 (5.99)	32.25 (8.20)	***	***	.002**
Fluid cognitive processing ability	0.76 (0.68)	-0.01 (0.78)	-0.83 (0.83)	***	***	***

Significance level notation: ***$p<.001$; **$p<.01$; *$p<.05$.

6. Correlations between background measures

We investigated possible intercorrelations between background measures and age using Spearman's rank-order correlation tests (cf. Table 1.3). A moderate-to-strong correlation was observed between the nonverbal intelligence measure and processing speed (RAVEN and DSST, respectively, r=0.58, p<.001) which may partly be due to a mental speed component in the speeded version of this reasoning task (cf. Wilhelm and Schulze, 2002). We ran a factor analysis using IBM SPSS Statistics 20 to derive a factor representing the common variance between the two cognitive measures. The use of this factor allowed us to avoid collinearity issues (redundancy) in our statistical modeling and enabled us to include a construct underlying both variables. The factor thus combines processing speed, which is linked to general (fluid) intelligence (e.g., Coyle et al., 2011), and reasoning abilities, which are thought to reflect general (nonverbal) intelligence. The analysis revealed an initial eigenvalue of the single factor explaining 79% of the variance with factor loadings of 0.89 both for processing speed and reasoning. Individual scores for each participant for the newly created composite variable 'fluid cognitive processing ability' (cf. Park et al., 2010) were included in the statistical analyses.

Table 1.3: Correlation matrix with correlation coefficients and significance levels for participant-related variables including age (Spearman's rank, significance levels corrected for multiple comparisons).

background variable	Age	PTA$^{\mathrm{low}}$	PTA$^{\mathrm{high}}$	PTA$^{\mathrm{HF}}$	Visual acuity	Processing speed	Vocabulary	Working memory	Reasoning
PTA$^{\mathrm{low}}$	0.44***								
PTA$^{\mathrm{high}}$	0.75***	0.59***							
PTA$^{\mathrm{HF}}$	0.80***	0.42***	0.83***						
Visual acuity	0.64***	-0.38**	-0.48***	-0.59***					
Processing speed	-0.60***	-0.41***	-0.57***	-0.56***	-0.48***				
Vocabulary	0.50***	0.27	0.32**	0.34**	0.31*	-0.13			
Working memory	-0.13	0.04	-0.12	-0.24	-0.06	0.11	0.01		
Reasoning	-0.59***	-0.41***	-0.52***	-0.51***	-0.38**	0.58***	-0.19	-0.02	
Fluid cognitive processing	-0.67***	-0.46***	-0.61***	-0.60***	-0.48***	0.88***	-0.17	0.04	0.89***

Significance level notation: $^{***}p$<.001; $^{**}p$<.01; $^{*}p$<.05.

1.2.3 Materials

1. Conversational stimuli

We specifically chose question-answer sequences (henceforth, QA sequences) for our test paradigm as these represent minimal conversational units which are “a reasonable proxy for turn-taking more generally” (Stivers et al., 2009, p. 10588). Conversational fragments were selected from the spontaneous dialogue part (face-to-face component) of the Spoken Dutch Corpus (Oostdijk, 2000). The following three primary criteria were defined to extract stimuli from the corpus on the basis of the corpus’ orthographic transcriptions and part-of-speech tagging: (a) the QA sequence had to consist of two speakers and one change of turns; (b) the minimal length of the question was two words (e.g., “Van wie?”, ‘Of whom?’); and (c) the minimal length of the answer had to be five words (e.g., “Ik ga een zon maken.”, ‘I will make a sun.’). The orthographic representations of the 1200 candidate QA sequences that met the criteria above were checked for coherence of question and answer by a native speaker of Dutch. Moreover, we selected QA sequences containing at least one Dutch mono- or disyllabic (trochaic) target noun in order to match syntactic category and length for the target words. To avoid prosodic boundary phenomena we only chose question-answer sequences in which the target word was neither the first word nor the last word in the answer portion of the QA sequence. The resulting set of QA sequences was narrowed down further by excluding conversations with speaker overlaps and stimuli with loud background noise, as well as QA sequences containing pauses longer than 0.2 s in the answer part of the second speaker. Application of these criteria led to a set of about 90 short question-answer conversation fragments. Out of those 90 QA sequences, 60 instances were selected as target stimuli, plus a set of 15 filler QA sequences, showing the same conversational features as the targets. An example target QA sequence is given in the orthographic transcription below with the target word underlined in the Dutch transcription and English translation (all QA sequences are listed in the Appendix A1).

Example 1

speaker 1: “Waar was het nou toch?”	‘Where was it again?’
speaker 2: “Waar die ten hemel steeg.”	‘Where he ascended to heaven.’

Table 1.4 shows the descriptive statistics of the variables related to the QA sequences used in our statistical analyses. Thirty-eight of the 60 target words were disyllabic, the remaining 22 target words were monosyllabic (monosyllabic structures varied in complexity from CVC to CCVCC; disyllabic nouns varied in complexity from CV–CV to CCCVC–CVC, see Appendix). We included the target words’ number of syllables measure as a variable in our analyses as the uniqueness point for

disyllabic target words may be earlier relative to word offset than for monosyllabic target words (see Analyses).

Target word duration ranged between 196 and 866 milliseconds (*M*=372 ms, *SD*=139). Mean CELEX word frequency (Baayen et al., 1993) for the 60 target words was 185 (occurrences per million tokens: English words having this frequency would be words such as *table*, *parents*, *evening*, *group*). Target word frequency values were logarithmically transformed to normalize their distribution (*M*=4.00, *SD*=1.78). Shapiro-Wilk normality tests showed that the log-transformed target word frequencies were normally distributed where the untransformed frequency values were not. Log-transformed target word frequency showed a statistically significant negative correlation with target word duration (Spearman's rank: *r*(58)=-0.27, *p*<.05). This relation was expected on the basis of Zipf's law, which predicts more frequent words to be shorter (Zipf, 1949, 1965). Log-transformed target word frequency was included as a control variable in our analyses, as we expected more frequent target words to be easier to recognize (see Analyses).

Table 1.4: Descriptives of the item-related variables used in statistical modeling.

Covariate	*M*	*SD*	Range
Speech rate (syllables/second)	5.91	1.80	2.93 – 11.22
Target word frequency (per 10^6 tokens, logtransformed)	4.00	1.78	0 – 7.22
Trial number (excl. 6 training trials)	45.13	21.40	7 – 81
Target word predictability	0.42	0.10	0.22 – 0.84
Target word position in the answer phrase	7.05	2.59	3 – 16
SNR (dB) for the answer phrase	23.79	5.34	12.43 – 37.42
Target word's number of syllables			1 | 2 *n*=22 : *n*=38

The position of the target word in the answer phrase of the QA sequence ranged between the third and the sixteenth word (*M*=7.05, *SD*=2.59). We included target word position as an item-related (control) predictor in the analyses as it can be interpreted as a predictability measure. Our hypothesis was that having a later position in the sentence would facilitate target word recognition.

For each target stimulus, speech rate of the second speaker's answer fragment was calculated in syllables per second from answer onset until the end of the target word. We based this calculation on the canonical (dictionary-based) number of syllables for each word in the target passage, rather than on the number of realized syllables. Speech rates are normally distributed over the stimulus set, and ranged between 2.93 and 11.22 syllables per second (*M*=5.91, *SD*=1.80). Obviously, target word duration and speech rate of the test items were strongly negatively correlated (Spearman's rank: $r(58)=-0.57$, $p<.001$). As an additional control covariate for our analyses, we also approximated the signal-to-noise ratio (henceforth, SNR) for all items separately. For each item the background noise intensity level (noise floor for the channel of the target speaker recording) was subtracted from the mean intensity of the respective answer part of the target speaker (*M*=23.79 dB, *SD*=5.34). Speech rate and SNR were not correlated (Spearman's rank: $r(58)=0.12$, $p>0.1$).

To investigate how the spectral content of our speech materials compared to standardized materials, we compared the long-term average spectrum (henceforth, LTAS) of our test stimuli to the LTAS of the International Speech Test Signal (Holube et al., 2010) which, in turn, has been shown to be comparable to the international long-term average speech spectrum (Byrne et al., 1994). This comparison did not show substantial differences between our 60 question-answer sequences and the ISTS material up to 4 kHz (mean difference 1 to 4 kHz over 100 Hz wide bins: 0.74 dB). We also checked whether the fragments with higher speech rate had a different spectral content from the lower speech rate fragments (by means of a mean split on speech rate). No systematic spectral differences were observed between the two sets of fragments.

The set of answer fragments in the 60 target stimuli involved 49 different speakers (age range: 19 to 76 years, M^{age}=37 yrs., *SD*=18.7). Eight target speakers were presented multiple times to the participants (maximally three times).

2. Orthographic stimuli

After the extraction of the target and filler QA sequences, orthographic stimuli were selected for the visual world paradigm employed in the eye-tracking experiment (McQueen and Viebahn, 2007). Three word categories were created for each target word: a semantic distractor, a phonetic distractor, and a phonetic distractor to the semantic distractor. The latter category was chosen to make the display symmetrical in that there were always two pairs of onset-overlapping words on the screen.

The semantic distractors were derived from the same semantic field as the respective target words (e.g., for the target “hemel” (‘sky’) the semantic distractor “aarde” (‘earth’) was selected). The phonetic distractors shared at least the initial phoneme with the respective target words but often also the following vowel (e.g., for the target “hemel” (‘sky’) the phonetic distractor “heling” (‘handling stolen goods’) was selected). The phonetic distractor to the semantic distractor stimuli minimally shared the initial phoneme with the respective semantic distractors (e.g., for the semantic distractor “aarde” (‘earth’) the semantic-phonetic distractor “aanhef” (‘salutation’) was selected). We verified that all orthographic distractors matched the morphosyntactic context of each individual sentence in terms of word class, number, and noun gender (as common gender nouns take a different definite article than neuter nouns in Dutch). For the 15 fillers the four orthographic stimulus categories were selected accordingly to ensure that participants could not tell upfront whether a stimulus was a target or a filler trial. Appendix A2 (p. 170) shows the set of distractors for the target words.

3. Assessing target predictability

All target QA sequences were tested for the predictability of the target word given the preceding conversational context. This was done in a separate test and allowed us to distinguish speech rate effects from effects of contextual target predictability (to be entered as control variable in our analyses). Note that this predictability measure differs from the position of the target word in the answer fragment introduced above: two target words that both occur as the fifth word in the phrase may still differ in how predictable they are given the prior words in the phrase (see correlations between measures below). Eighteen younger Dutch adults (M^{age}=19.6 years, 14 female) participated in this predictability rating experiment. Participants were presented with orthographic representations of the 60 test QA sequences up until the target word and had to rate all four orthographic word stimuli (the target, the semantic distractor, the phonetic distractor, and the phonetic distractor to the semantic distractor) for their match with the given context on a scale from 0 (does not fit at all) to 100 (fits perfectly). Participants gave their ratings in a text processing program on a computer. Target word predictability scores were calculated for each QA sequence in two ways: as an absolute predictability score for the target word given the pre-context, and as a proportional value, being the target word’s rating against the sum of ratings for all four orthographic representations (see Appendix A1 and A2; descriptive statistics for the proportional values see Table 1.4). The proportional predictability score was calculated to take into account how probable the target word was relative to the predictability ratings of the three orthographic distractors (cf. Brouwer et al., 2012). No significant correlations were found between target predictability and the following four item-related variables: speech rate in the answer fragment, target word frequency (CELEX frequency), target word position in the answer phrase and SNR (none of the r values exceeded |0.20|, all Bonferroni-corrected p values exceeded 0.1).

1.2.4 Procedure

We set up a word recognition experiment, using the visual world paradigm (Allopenna et al., 1998). On each experimental trial, participants had to click (with a computer mouse) the one out of four orthographically represented words they heard in a conversational speech fragment. We used a Desktop Mount Eyelink 1000 eye-tracker with a chinrest, a ViewSonic 22 inch screen monitor, circumaural Sennheiser HD 215 headphones plus a Hewlett Packard USB mouse. Testing took place in a sound-attenuating booth. Illumination was kept constant at 50 lux for all subjects during the whole test procedure to allow for pupillometry measurement (cf. Zekveld et al., 2010). Before participating in the eye-tracking experiment, all participants underwent a visual acuity check in the test booth with a near vision Snellen chart (cf. subsection 1.2.2, p. 20). Participants first read the experiment instructions outside the test booth and could ask questions if anything was unclear. Instructions were repeated by displaying them on the computer screen after the eye-tracking calibration process just before the practice trials were presented. Each trial consisted of three phases: a talker familiarization phase, a preview phase and a response phase. During talker familiarization participants listened to short fragments of the two speakers they would hear in the upcoming QA sequence. These audio fragments consisted of about two seconds of speech (minimally a six syllable utterance), and were not related, content-wise, to the test QA sequence. Each of the familiarization fragments was preceded by an announcement spoken by a female speaker whom they would be hearing next (‘speaker 1’ or ‘speaker 2’). The order in which speakers were introduced matched the order of the speakers in the upcoming test stimulus. After this familiarization phase a fixation cross was presented for 300 milliseconds centered on the screen. After talker familiarization and fixation cross, participants got a preview of the four (candidate) words on the screen for a period of three seconds. During this preview phase participants could read the four words silently to be prepared for the upcoming test conversation. These words (cf. subsection 1.2.3, p. 26 and Appendix A2, p. 170 for more details) were presented in a black sans serif Tahoma font (30 point, bold letters) on a white background in four equal sized quadrants of a centralized section of the display. Apart from the four quadrant areas, an additional click region was present on the screen for both target and filler trials. This region (a centralized smaller grey colored circle labeled ‘none of these words’) had to be clicked if none of the words on the screen had been perceived. After the three seconds preview period, listeners were presented with the question-answer sequence with the target word embedded in the second speaker's answer (for test trials) or without a target word (for filler trials). The familiarization and test stimuli were presented binaurally (same signal for both channels), at an intensity level of 70 dB SPL using headphones.

The four words were displayed throughout the entire preview and response period until the participants clicked one of the words or the 'none of these words' area on the screen. On each trial, the mouse cursor was reset to the screen center at the beginning of the preview phase. Each participant was presented with the same set of test and filler items. Participants were randomly assigned to one of four lists that comprised the 60 target stimuli plus 15 filler items in a different randomized order. Each of the word categories (target, phonetic distractor, semantic distractor, semantic-phonetic distractor) occurred equally often in each of the four quadrants on each of the randomization lists. Each participant got the same set of six practice trials before the experiment started (3 target trials and 3 filler trials) to familiarize them with the task. Test duration for the eye-tracking procedure was approximately 22 minutes.

1.3 Analyses

Two types of models were set up: age group comparison models and individual-differences models (across age groups). Speech rate was the continuous within-subjects variable (n=60) of interest in both types of models. Age group was entered as a between-subjects factor (younger adults, middle-aged adults, older adults) in the age group comparison models. For each of the investigated dependent variables (click response times, target word gaze probability, pupil dilation latency, pupil dilation amplitude) separate statistical regression models were run using linear mixed-effect models in the program `R` with the `lme4` package (Bates et al., 2013). As additional control variables (within-subjects predictors), we included the frequency of the target word (continuous), trial number (continuous), target word predictability (continuous), the target word's position in the answer phrase (continuous), SNR (continuous), and the target word's number of syllables (two levels: monosyllabic, disyllabic). Target word frequency (log-transformed) was included as an (item-related) control variable in our analyses because we hypothesized that more frequent targets would be easier to recognize. Trial number was included as control variable in the click response time (henceforth, click RT) and the pupil data analysis. Our hypothesis was that fatigue or practice effects would be covered by including trial number. To control for context effects we included the target word's predictability in our modeling. We hypothesized that items with a higher predictability would be easier to recognize. As noted above, the target word's position in the answer phrase can also be interpreted as a predictability measure. It is however a more local context measure than the predictability measure above as it only covers the number of words prior to the target in the answer sequence of the second speaker. Our hypothesis was that more prior words would facilitate word recognition. We also hypothesized the SNR of the items to have an effect on performance. We expected word recognition to be easier at higher SNRs. We included the target word's number of syllables as a variable in the model as the uniqueness point (i.e., that point where the transcription of the word makes it unique relative to all other words) of disyllabic target words may be earlier in the word than for monosyllabic words. Thus, if we measure from word offset, disyllabic words may show shorter click response times, faster looks to the target and faster pupil dilation responses than monosyllabic words.

Word recognition accuracy (as evidenced by clicking on the correct target) was investigated with generalized linear mixed-effect modelling (fixed effect: age group, random effects: participant, test item). Accuracy of the click responses was 98.1% for the younger adults, 98.3% for middle-aged adults, and 94.8% for the older adults. The analysis showed a significant age effect with older participants performing slightly worse than middle-aged (B=-1.39, $|z|$=4.76, p<.001)[2] and than younger adults (B=-1.24, $|z|$=4.45, p<.001).

Only correct responses for the target trials were included in the subsequent data analyses (97.1% of all 6000 target trials). The models contained all item-related control variables as fixed effects (i.e., target word frequency, trial number, target word predictability, position of the target word in the answer phrase, SNR, number of syllables of the target word, see Table 1.4 for descriptives). Trial number was included in the click response and the pupil data analysis. Trial number could not be included as a covariate into the gaze analysis as aggregated gaze data (over participants) did not contain information about the trial order anymore.

The age group comparison models also contained the critical interaction between the variables speech rate and age group. Individual-differences models also included participant-related variables (such as hearing, age in years, and scores on the five cognitive/linguistic tests) as fixed effects, as well as the critical interaction between speech rate and these participant-related variables.

The target word gaze probability data were analyzed with Growth Curve Analysis (Mirman, 2014) to capture the time course of participants' gaze behavior. Growth curve analysis (henceforth, GCA) is a type of multilevel regression with which variation in curve shapes over time can be modeled. Thus, GCA can model linear and non-linear behavior of a dependent variable. In addition, a main advantage of GCA is that it does not involve alpha inflation due to repeated comparisons for multiple analysis windows. Given our Growth Curve analysis approach for the eye gaze data,

---

[2] B here denoting an unstandardized coefficient

the first, second and third order orthogonal time polynomials (linear, quadratic and cubic time component) were included in our gaze data analysis as fixed effects. The first orthogonal polynomial (Time^1) describes linear change in the dependent variable over time, which is comparable to the slope in a linear regression model. The second time polynomial (Time^2) captures the change of a dependent variable over time that follows a quadratic function (flat vs. more bent curve). The third time polynomial (Time^3) encompasses the curve shape as a product of a cubic time fitting function. A cubic function involves an additional twist in the curve compared to the quadratic function. In the age group comparison model, interactions between each of these three time variables and age group were included, as well as between each of the time variables and speech rate. Secondly, three-way interactions between age, speech rate and each of these three time variables were included.

Performance of the younger adults (i.e., the group mapped on the intercept) served as baseline for the age group comparison analyses. The random-effect structure of the models included random intercepts for participants and items where possible (i.e., for the click response time analysis and for the pupillometry data).

Due to the sparseness of the gaze data for Growth Curve Analysis, we had to aggregate our data. For the age group comparison, we aggregated over participants within each age group. Consequently, the random structure of the gaze data age group comparison analysis had only item (and not participant) as a random intercept. No individual-differences model is reported here, as setting up such a model would have entailed aggregating over items (and hence over speech rate, being our variable of interest).

We allowed for the possibility that the effect of speech rate randomly differed across participants. We therefore added random slopes for speech rate to the random structure of our click response time and pupillometry data models. Additionally, for the gaze probability model, the orthogonal polynomial terms (time components) were added on the respective random intercept (test item). All continuous variables were z-transformed. As the linear mixed-effect models do not output significance levels, we obtained these test statistics by using the `Anova` function of the `R car` package which calculates Type II Wald χ^2 values. For models including age group (which is an ordinal categorical variable with three levels: younger, middle-aged, older), p values were obtained using the model's t values. The number of degrees of freedom was estimated via the formula df=J-m-1 (Hox, 2010), with J being the most conservative number of second-level units (number of items in our study, n=60) and m being the number of included predictors.

Below we will describe the dependent variables (click response time, target gaze probability, pupil dilation latency and pupil dilation amplitude) separately to elaborate on the necessary data transformation steps.

1. Click response time

A priori we expected increased speech rate to make word recognition more difficult such that click response times would be slower (cf. Janse, 2009). We measured click response times from target word offset such that we did not have to take word duration into account, which was correlated with speech rate. Outliers were removed per age group for all accurate trials: responses slower than 2.5 *SDs* above each respective age group's mean click response time (in milliseconds) were excluded (148 trials). The exclusion of inaccurate and outlier responses resulted in 95% of the recorded click response times being fed into the analyses (95% of the data points for the younger, 96% for middle-aged adults and 93% for the older adults). Mean click response time after outlier removal was 1030 ms after word offset (*SD*=507). As expected, middle-aged and older adults generally gave slower responses than younger participants (younger adults: M=877 ms; *SD*=444; middle-aged adults: M=1026 ms; *SD*=508; older adults: M=1205 ms; *SD*=515).

2. Target gaze probability

We expected that increased speech rate would make word recognition more difficult, and that the increased word recognition difficulty would result in a lower probability of correct target gazes (cf. Ben-David et al., 2011) and in a lower slope of the rise in target gaze probability in the analysis window.

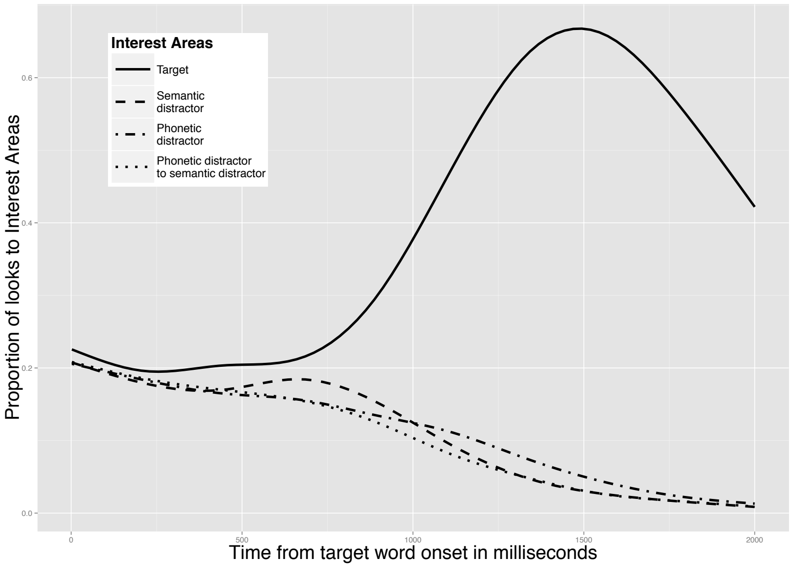
Gaze fixations to the five interest areas (i.e., to the four word quadrants and to the 'none of these words' area) were investigated in the time window between 200 and 1400 ms after target word onset. The onset of the analysis window was chosen because programming a saccade takes approximately 200 ms (cf. Barr, 2008). The window's upper limit was set to 1400 ms (given the mean click RT as measured from word onset).

Binomial gaze data for the interest areas were transformed to gaze probability on a log-odds scale (empirical logit, see Barr, 2008) over 24 consecutive time bins of 50 ms between 200 and 1400 ms after target word onset. As noted above, for the analysis of the *age* × *speech rate* interaction gaze data were aggregated over participants within each age group for each item (the items varying continuously in speech rate). Consequently, 1440 data points (24 time bins × 60 items/speech rates) were available per age group for growth curve analysis.

Mean probability (log odds ratio) of looking at the target over the 24 time windows for the three age groups was -0.85 (unit empirical logit, *SD*=1.10, range: -7.47 – 3.07) which corresponds to a mean probability of 29.9 percent (range: 0.06% – 95.6%).

This probability was very similar across age groups (younger adults: *M*=-0.88, *SD*=1.06, range: -7.47 – 1.67; middle-aged adults: *M*=-0.79, *SD*=1.10, range: -7.38 – 2.47; older adults: *M*=-0.87, *SD*=1.12, range: -7.38 – 3.07). Figure 1.2 shows the gaze probability curves for target and distractors from target word onset for the analysis interval (200 – 1400 ms after target word onset).

Figure 1.2: Grand mean of the fixation proportions to target and distractor words over time (measured from word onset).



3. Pupillometric data

Several pupil measures have been reported to reflect cognitive effort in language processing, such as mean pupil size and pupil peak latency (Andreassi, 2000; Zekveld et al., 2013; Kuchinsky et al., 2012; Schmidtke, 2014). We investigated the task-evoked pupil response (pupil peak latency, pupil peak amplitude) for the word recognition task starting from target word onset. Our hypothesis was that faster speech rates would result in higher processing demands, yielding a delayed dilation response with a higher peak amplitude (cf. Beatty, 1982; Zekveld et al., 2010). Pupil size data were recorded with a sampling rate of 500 Hz. For each trial the last 500 ms of the (silent) preview phase served as a baseline for the size of the pupil. This time interval was chosen for an item-individual baseline correction as visual input during baseline was the same as during listening in the test phase (having the 4 candidate words for this trial on the screen). Consequently, mean baseline pupil size could be subtracted from pupil size data points for the analysis interval. Trials with a high rate of missing values (i.e., more than 3.0 *SDs* above the mean) for the baseline interval were excluded from further processing (resulting in exclusion of 2.0% of all accurate trials, 117 trials). Missing values in the remaining baseline data were imputed by linear approximation (`na.approx` function, `R` package `zoo`). We then applied a locally weighted polynomial regression fitting algorithm in `R` (`loess` function, package `stats`, settings: span=0.1, degree=1). For each trial, we calculated a unique baseline pupil size value averaged over the fitted data in the baseline time window. For pupil size during the test window, the pupil size data in the time window between 500 ms before target word onset until the participant's click response was processed. Trials with a high incidence of missing values (i.e., more than 3.0 *SDs* above the mean) were excluded (resulting in exclusion of 1.1%, 65 trials). Missing values were imputed during the data smoothing and fitting procedure using a polynomial regression algorithm, which assigns less weight to outliers (`smooth.m`[3] with method `rloess`, span=0.2, `MATLAB`). Baseline correction per trial was accomplished by subtracting the baseline mean pupil size value from each of the samples of the smoothed and fitted test data. The resulting baseline-corrected data showed a mean peak dilation maximum at around 1000 ms after target word onset which is in line with the timing of the canonical pupil response for processing a stimulus (cf. Zekveld et al., 2010; Kuchinsky et al., 2012). As a reminder, for the gaze data the analysis window was set between 200 and 1400 ms after target word onset (1400 ms being the mean click RT measured from word onset). However, based on the literature (Privitera et al., 2008) and on visual inspection of the pupil response grand mean over all trials, we chose a different analysis window for the pupil dilation peak data. Privitera and colleagues (2008) report latencies of around 300 ms to 700 ms for the onset of the dilation phase. We therefore set our peak detection window between 500 ms and 1800 ms after target word onset (see Figure 5 for the pupil dilation curves per age group). For each trial peak latency and peak amplitude were automatically extracted (`peakdet.m`[4], delta=0.9, `MATLAB`). Whenever there were multiple peaks in the detection window, the highest-amplitude peak was chosen. Automatic pupil peak detection was successful in 69% of the trials (with similar percentages of included trials for the different age groups). Mean pupil peak latency for the three age groups was 796 milliseconds (*SD*=266, range: 224 – 1380; younger adults: *M*=756 ms, *SD*=263; middle-aged adults: *M*=792 ms, *SD*=262; older adults: *M*=842 ms, *SD*=267). Mean pupil peak amplitude for the three age groups was 300.82 (arbitrary unit, *SD*=200.80; younger adults: *M*=351.08, *SD*=232.10; middle-aged adults: *M*=280.47, *SD*=159.12; older adults: *M*=270.26, *SD*=194.11).

---

[3] www.mathworks.com/matlabcentral/fileexchange/6271-spec-file-reader/content/specreader/smooth.m

[4] www.mathworks.com/matlabcentral/fileexchange/32828-spectr-o-matic/content/peakdet.m

1.4 Results

1. Click response time

The response time analysis was conducted to answer Research Questions 1 (Can we replicate speech rate effect using conversational speech materials?) and 2 (Do the three age groups differ in the effect of speech rate on their word identification performance?).

The result of the statistical model testing for the critical interaction between the predictors speech rate and age group (plus the control variables discussed in section Analyses above) is shown in Table 1.5 and is illustrated in Figure 1.3.

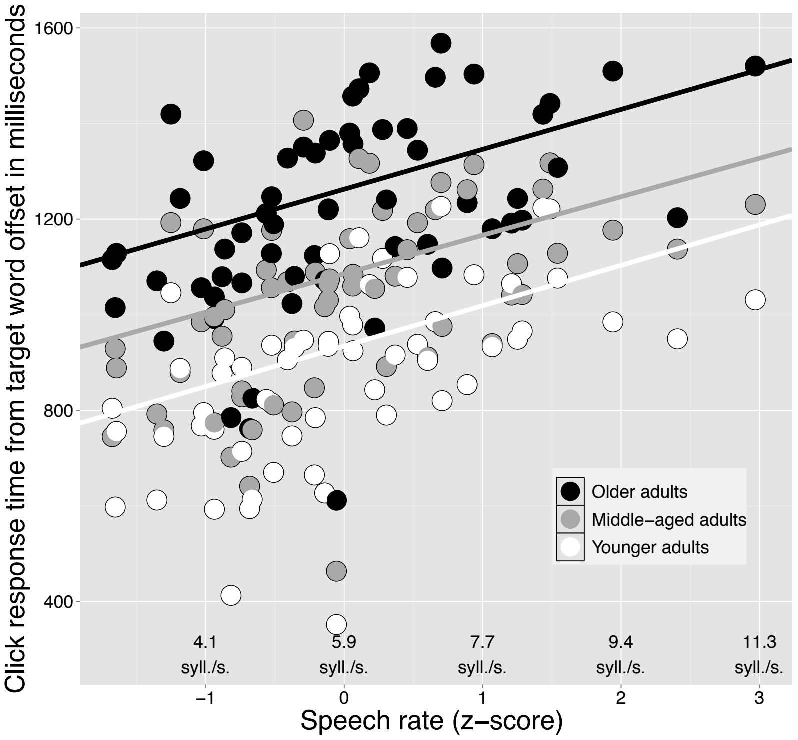
Age groups differed significantly in their click response times (t>2.62, p<.02, for both comparisons). As can be seen in the model in Table 1.5 and in the model plot in Figure 3 younger adults showed the fastest click responses (approx. 930 ms at the mean speech rate of 5.9 syll./s.) followed by middle-aged adults (approx. 150 ms slower than younger adults) with older adults having the slowest click RTs (approx. 330 ms slower than younger adults).

Table 1.5: Click response time data (in milliseconds): Model testing for the *age (group)* × *speech rate* interaction.

Fixed effects	Estimate	*SE*	|*t*|	*p*<
Intercept: Younger adults	934.97	49.60	18.85	
Middle-aged vs. Younger adults	151.43	57.65	2.63	.012*
Older adults vs. Younger adults	328.30	57.97	5.66	.001***
Speech rate: Younger adults	84.25	20.34	4.14	.001***
Speech rate: Middle-aged vs. Younger adults	-4.00	16.39	0.24	.809
Speech rate: Older adults vs. Younger adults	-1.08	16.04	0.07	.947
Target word frequency	18.86	18.13	1.04	.304
Trial number	-69.06	6.04	11.45	.001***
Target word predictability	-4.82	19.48	0.25	.806
Target word position	-16.53	18.39	0.90	.373
SNR	15.64	17.67	0.89	.381
Target word's number of syllables	-70.96	39.62	1.79	.080 .
Speech rate × Trial number	23.91	6.14	3.89	.001***

Significance level notation: ***p<.001; *p<.05, .p<0.1.

Figure 1.3: Model predictions for click response times in milliseconds per age group as a function of speech rate (*speech rate* × *age group* model). Points represent mean observed click response times per age group across speech rates (speech rate is given both as *z*-scores on the *x*-axis, and for illustration purposes also as actual syllables per second). The model predictions (fit lines) take the contribution of all control variables into account.



For younger adults (mapped on the intercept), the model predicted an increase in click response times of about 84 ms for an increase of one *z*-score in speech rate (cf. estimate for speech rate variable, |*t*|=4.14, *p*<.001). This corresponds to an increase of 47 ms in click response time for an increase of one syllable per second and sums up to an effect size of around 390 ms for the tested range in speech rate (2.93 – 11.22 syll./s.). Running the same model with the middle-aged group (rather than the younger group) mapped on the intercept, we observed a significant speech rate effect of 80 ms for an increase of one *z*-score in speech rate (|*t*|=3.59, *p*<.001). Mapping older adults on the intercept resulted in a significant speech rate effect of 83 ms for an increase of one *z*-score in speech rate (|*t*|=3.53, *p*<.001). As stated above, compared to younger adults, click response times of middle-aged and older adults were slower.

However, the three age groups showed similar speech rate slopes, as evidenced by the insignificant age group by speech rate interaction estimates ($|t|<1$, $p>0.1$, for both comparisons; see Figure 3 and Table 1.5).

None of the (item-related) control variables, except trial number, had a significant effect on click response time (trial number effect: $|t|=11.45$, $p<.001$). Click RTs decreased with increasing trial number (-69 ms for each z score unit increase, i.e., 248 ms over the experiment), suggesting task familiarization. The model also showed a marginally significant effect of the target word being monosyllabic or disyllabic: as expected, disyllabic words tended to be recognized earlier relative to word offset than monosyllabic words (effect size: approx. 70 ms).

Our third question was which individual abilities would modulate the effect of speech rate on word recognition performance. This was investigated in a model testing for interactions between speech rate and all participant-related predictors including chronological age (plus all control predictors related to item characteristics). We also tested for possible interactions between trial number and the participant-related variables to check for background variables that modulated the individual task familiarization effect of the participants. Table 1.6 displays the resulting model.

As before, statistically significant effects of speech rate, age and trial number were observed. Slower click RTs were observed for items with higher speech rates ($|t|=4.16$, $p<.001$) and for older compared to younger participants ($|t|=2.04$, $p<.05$). Click RTs decreased over trials ($|t|=11.50$, $p<.001$). None of the other control predictors affected click response time. Participants with better fluid cognitive processing ability and better vocabulary knowledge showed generally faster click RTs ($|t|>2.17$, $p<.05$). Importantly, however, none of the participant-related variables showed significant interactions with speech rate. The variable trial number showed an interaction with speech rate such that speech rate effects became larger for later trials ($|t|=3.89$, $p<.001$). This may relate to the general trial effect that participants speeded up their click responses over the experiment due to task familiarization. Possibly, stimulus-related effects, like speech-rate variation, become more apparent once response times are more closely time-locked to ongoing speech processing.

In sum, our click response time results confirmed that speech rate effects on word recognition performance can be found using conversational stimuli (Research Question 1). Secondly, the click response time data showed that the three age groups were equally affected by increased speech rate (Research Question 2). Concerning our third research question on individual differences in the effect of speech rate on word identification, none of the included cognitive, hearing-related or linguistic abilities was found to be associated with the size of the speech rate effect on click response times.

Table 1.6: Click response time data (in milliseconds): Model testing for interactions between speech rate and participant-related variables.

Fixed Effects	Estimate	*SE*	|*t*|	*p*<
Intercept	1148.98	71.09	16.16	
Age	92.55	45.43	2.04	.034*
Speech rate	83.31	20.01	4.16	.001***
Target word frequency	19.83	18.33	1.08	.280
Trial number	-69.40	6.04	11.50	.001***
Target word predictability	-3.48	19.70	0.17	.860
Target word position	-17.76	18.59	0.96	.340
SNR	16.59	17.86	0.93	.353
Target word’s number of syllables	-64.95	40.07	1.62	.106
Speech rate × Trial number	23.87	6.14	3.89	.001***
PTA^{HF}	20.05	33.89	0.59	.981
Visual acuity	16.73	29.14	0.57	.473
Fluid cognitive processing ability	-74.53	30.39	2.45	.003**
Vocabulary	-57.99	26.69	2.17	.048*
Working memory	17.43	22.47	0.78	.344
Speech rate × Age	-1.26	12.86	0.09	.923
Speech rate × PTA^{HF}	-15.57	9.36	1.66	.097 .
Speech rate × Visual acuity	2.56	8.03	0.32	.750
Speech rate × Fluid cognitive processing ability	-8.90	8.38	1.06	.289
Speech rate × Vocabulary	6.30	7.37	0.86	.393
Speech rate × Working memory	2.67	6.16	0.43	.665
Trial number × Age	8.08	11.04	0.73	.465
Trial number × PTA^{HF}	-13.91	8.25	1.69	.092 .
Trial number × Visual acuity	5.72	7.10	0.81	.421
Trial number × Fluid cognitive processing ability	4.85	7.41	0.66	.513
Trial number × Vocabulary	-3.42	6.52	0.53	.600
Trial number × Working memory	10.55	5.44	1.94	.053 .

Significance level notation: ***$p<.001$; **$p<.01$; *$p<.05$, .$p<0.1$.

2. Target word gaze probability

As indicated above, Growth Curve Analysis (GCA) was used to analyze the time course of the target gaze data from 200 ms to 1400 ms after target word onset. As data aggregation was necessary for GCA, and the continuous variable speech rate was our variable of interest, we only carried out the age group comparison analysis. This analysis tested for the critical *age (group)* × *speech rate* interaction on the time course of the target word fixations (probability of looking at the target word). We hypothesized that higher speech rates would result in overall less fixations on the target word and in a shallower slope of the target fixation probability. Table 1.7 provides the full resulting model.

The model showed a statistically significant effect of speech rate on the probability of looking at the target word (β=-0.13, $|t|$=3.42, p<.01)[5]. This means that speech rate affected the probability of fixating the target word, with higher speech rates leading to decreased target gaze probability.

With the middle-aged group on the intercept, the model outputs a speech-rate estimate of β=-0.10 ($|t|$=2.76, p<.01) and with older adults as reference group the speech-rate β is -0.14 ($|t|$=3.68, p<.001). We did not find age effects on target gaze probability ($|t|$<1, p>0.1, for both comparisons). However, the model shows that older adults have a higher linear increase (Time^1 component) of their target gaze probability over the analysis window (β=0.61, $|t|$=2.21, p<.05). As can be seen in Figure 1.4, older adults differ in their target gaze behavior from the other two age groups mainly in the very first two to three time bins (i.e., 200 – 350 ms after target word onset). This steeper linear increase in gaze probability may mainly be due to older adults' early gaze behavior (i.e., at the start of the analysis window). Note that target gaze probabilities were only around 10% for the older adults in the first two to three time windows (i.e., below the chance level of 20%).

While we did not find an effect of speech rate in interaction with the linear time term ($|t|$<1, p>0.1), the quadratic time term (Time^2, curvature) changed with increasing speech rate (β=0.17, $|t|$=2.05, p<.05). Thus, the higher the speech rate, the more bent the gaze probability curve was, indicating a delayed target fixation pattern.

Whereas we observed both a speech rate effect (Research Question 1) as well as a general age (group) effect on target gaze probability (generally steeper linear increase for older adults), the model did not provide evidence for an *age* × *speech rate* interaction (Research Question 2). Figure 1.4 shows the gaze curves of the three age groups broken down by speech rate (dichotomized for illustration purposes).

---

[5] β here denoting a standardized coefficient

Table 1.7: Target gaze probability data (empirical logit scale): Growth Curve Analysis model testing for the *age (group)* × *speech rate* interaction over time.

Fixed effects	Estimate	*SE*	$\|t\|$	*p*<
Intercept: Younger adults	-7.74×10^{-1}	8.42×10^{-2}	9.10	
Middle-aged vs. Younger adults	8.75×10^{-2}	9.12×10^{-2}	0.96	.342
Older adults vs. Younger adults	6.74×10^{-3}	9.12×10^{-2}	0.07	.941
Speech rate: Younger adults	-1.28×10^{-1}	3.76×10^{-2}	3.42	$.002^{***}$
Speech rate: Middle-aged vs. Younger adults	2.49×10^{-2}	5.12×10^{-2}	0.49	.630
Speech rate: Older adults vs. Younger adults	-9.67×10^{-3}	5.12×10^{-2}	0.19	.852
Target word frequency	2.76×10^{-2}	3.95×10^{-2}	0.70	.490
Target word predictability	2.21×10^{-2}	4.22×10^{-2}	0.52	.604
Target word position	8.61×10^{-2}	3.95×10^{-2}	2.18	$.035^{*}$
SNR	2.30×10^{-2}	3.84×10^{-2}	0.60	.554
Target word's number of syllables	-1.65×10^{-1}	8.55×10^{-2}	1.93	.060 .
Time^1	3.20	2.56×10^{-1}	12.52	$.001^{***}$
Time^2	1.30	1.83×10^{-1}	7.11	$.001^{***}$
Time^3	-1.78×10^{-1}	1.46×10^{-1}	1.21	.232
Time^1 × Speech rate (Younger adults)	-1.35×10^{-2}	1.14×10^{-1}	0.12	.907
Time^1 × Middle-aged vs. Younger adults	3.36×10^{-1}	2.77×10^{-1}	1.21	.231
Time^1 × Older adults vs. Younger adults	6.13×10^{-1}	2.77×10^{-1}	2.21	$.032^{*}$
Time^1 × Target word frequency	2.26×10^{-2}	1.20×10^{-1}	1.89	.066 .
Time^1 × Target word predictability	1.81×10^{-1}	1.28×10^{-1}	1.41	.165
Time^1 × Target word position	1.17×10^{-1}	1.20×10^{-1}	0.98	.334
Time^1 × SNR	7.92×10^{-2}	1.17×10^{-1}	0.68	.500
Time^1 × Target word's number of syllables	1.76×10^{-2}	2.59×10^{-1}	0.68	.500
Time^2 × Speech rate (Younger adults)	1.67×10^{-1}	8.19×10^{-2}	2.05	$.047^{*}$
Time^2 × Middle-aged vs. Younger adults	2.95×10^{-1}	1.98×10^{-1}	1.49	.144
Time^2 × Older adults vs. Younger adults	-1.48×10^{-1}	1.98×10^{-1}	0.75	.460
Time^2 × Target word frequency	-1.24×10^{-1}	8.59×10^{-2}	1.44	.157
Time^2 × Target word predictability	1.11×10^{-1}	9.19×10^{-2}	1.21	.232
Time^2 × Target word position	-7.11×10^{-2}	8.59×10^{-2}	0.83	.412
Time^2 × SNR	1.29×10^{-1}	8.36×10^{-2}	1.55	.128
Time^2 × Target word's number of syllables	-2.40×10^{-1}	1.86×10^{-1}	1.29	.203
Time^3 × Speech rate (Younger adults)	-4.29×10^{-2}	6.53×10^{-2}	0.66	.515
Time^3 × Middle-aged vs. Younger adults	2.37×10^{-1}	1.59×10^{-1}	1.49	.143
Time^3 × Older adults vs. Younger adults	2.61×10^{-1}	1.50×10^{-1}	1.64	.108
Time^3 × Target word frequency	-1.01×10^{-1}	6.53×10^{-2}	1.48	.146
Time^3 × Target word predictability	-1.65×10^{-2}	7.35×10^{-2}	0.23	.823
Time^3 × Target word position	-1.67×10^{-2}	6.87×10^{-2}	0.24	.809
Time^3 × SNR	-9.48×10^{-2}	6.69×10^{-2}	1.42	.163
Time^3 × Target word's number of syllables	-2.69×10^{-1}	1.49×10^{-1}	1.81	.077 .
Time^1 × Speech rate × Middle-aged vs. Younger adults	8.82×10^{-2}	1.55×10^{-1}	0.57	.573
Time^1 × Speech rate × Older adults vs. Younger adults	2.28×10^{-1}	1.55×10^{-1}	1.47	.149
Time^2 × Speech rate × Middle-aged vs. Younger adults	-2.51×10^{-2}	1.11×10^{-1}	0.23	.823
Time^2 × Speech rate × Older adults vs. Younger adults	-1.56×10^{-1}	1.11×10^{-1}	1.40	.168
Time^3 × Speech rate × Middle-aged vs. Younger adults	-1.67×10^{-2}	8.90×10^{-2}	0.19	.852
Time^3 × Speech rate × Older adults vs. Younger adults	-6.01×10^{-3}	8.90×10^{-2}	0.07	.947

Significance level notation: ***$p<.001$; *$p<.05$, .$p<0.1$.

Additionally, the model showed an effect of the control predictor target word position in the answer phrase: Items for which the target word was later in the answer phrase showed a higher probability of looks to the target (β=0.09, $|t|$=2.18, p<.05).

To conclude, our analysis of the time course of looking at the target word confirmed that speech rate effects on word recognition performance can be found using conversational stimuli (Research Question 1). The gaze data also showed that the gaze behavior pattern of the three age groups was equally affected by increased speech rate (Research Question 2).

Figure 1.4: Target fixation probability over the analysis interval (200 ms – 1400 ms) for the three age groups for low and high speech rate items (median split on speech rate). Error bars represent standard errors.

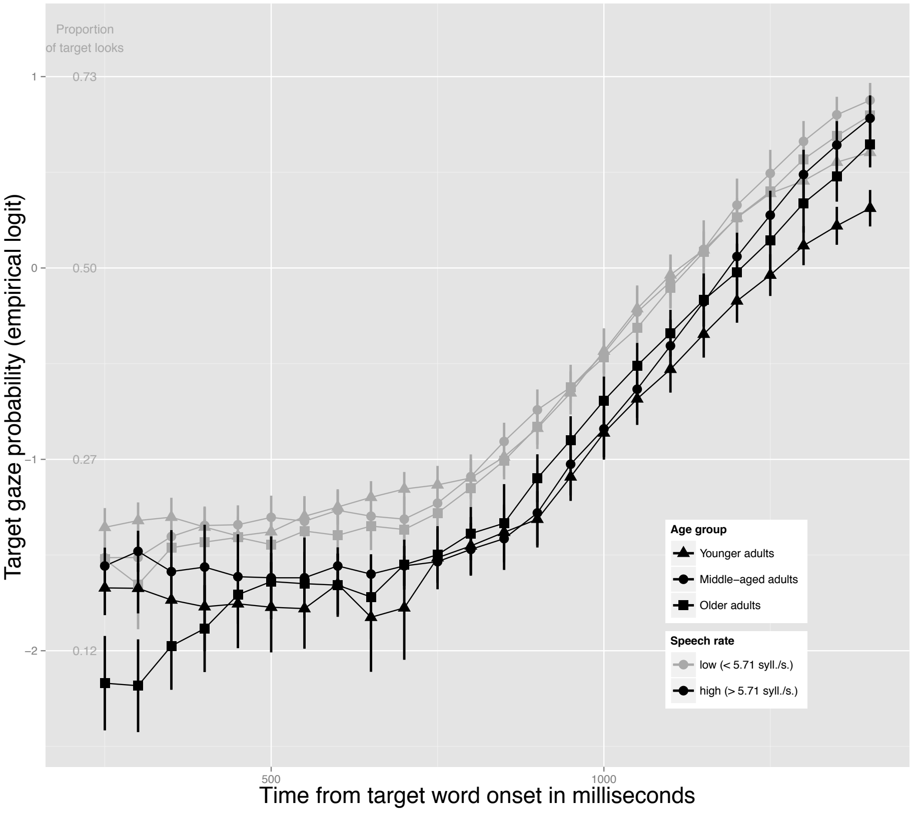
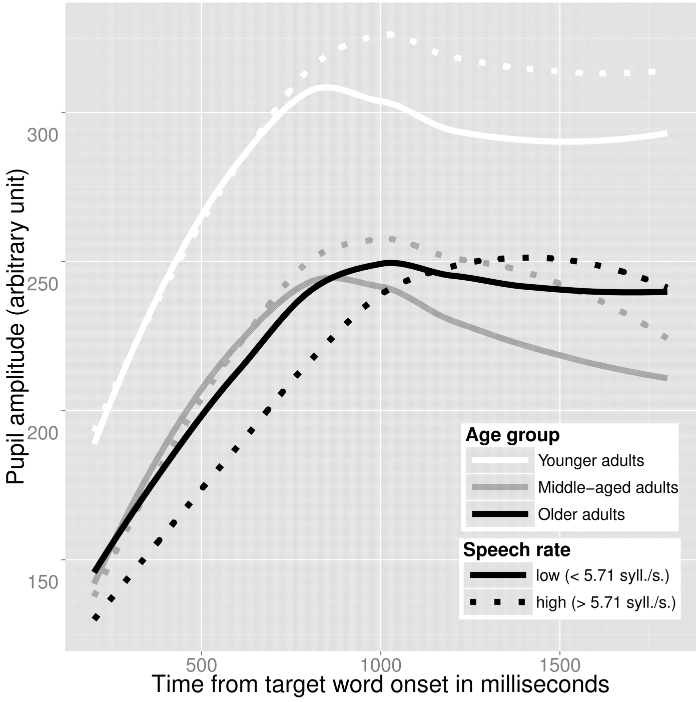


Figure 1.5: Task evoked pupillometry response per age group for low and high speech rate items (median split on speech rate). The window chosen for peak detection was from 500 ms to 1800 ms after word onset.



3. Pupillometric data

Two different analyses were conducted on the two pupillometry variables (pupil peak latency, pupil peak amplitude): one to address the age group comparison and the other to investigate individual differences. Figure 1.5 shows the time course of the pupil response per age group for low and high speech rates.

Age groups differed significantly in their pupil peak latency ($|t|>2.8$, $p<.05$, for both age group comparisons). Younger adults showed the fastest pupil peak dilation latency (approx. 900 ms at the mean speech rate of 5.9 syll./s.) followed by middle-aged adults (approx. 39 ms slower than younger adults) with older adults having the slowest pupil peak dilation response (approx. 106 ms slower than younger adults). However, even though Figure 1.5 suggests that pupil peak latency is affected by speech rate, the age group comparison model showed no significant speech rate effect on pupil peak latency ($|t|<1$, $p>0.1$). Furthermore, there was no evidence for a significant age group by speech rate interaction ($|t|<1$, $p>0.1$, for both comparisons; see Table 1.8). Additionally, the model shows significant effects of target word predictability, target word position in the answer phrase and trial number. Pupil peak latency was smaller for more probable items and if the target word came later in the answer phrase ($|t|>2.11$, $p<.05$, for both effects). Moreover, pupil dilation latency decreased over trials ($|t|=3.68$, $p<.001$). Thus, all three described control variables facilitated word recognition.

3.a. Pupil peak latency The result of the statistical model testing for the critical interaction between the predictors speech rate and age group (plus the control variables discussed in section Analyses above) is shown in Table 1.8.

Table 1.8: Pupil peak latency data (in milliseconds): Model testing for the *age (group)* × *speech rate* interaction.

Fixed effects	Estimate	*SE*	$\|t\|$	*p*<
Intercept: Younger adults	900.20	18.61	48.38	
Middle-aged vs. Younger adults	38.77	18.55	2.09	.042*
Older adults vs. Younger adults	106.41	18.63	5.71	.001***
Speech rate: Younger adults	5.04	10.66	0.47	.639
Speech rate: Middle-aged vs. Younger adults	9.90	10.63	0.93	.356
Speech rate: Older adults vs. Younger adults	3.01	10.57	0.29	.777
Target word frequency	-11.84	8.43	1.41	.167
Trial number	-17.90	4.86	3.68	.001***
Target word predictability	-19.67	8.81	2.23	.030*
Target word position	-17.78	8.39	2.12	.040*
SNR	-14.64	8.09	1.81	.077 .
Target word's number of syllables	1.34	18.13	0.07	.942
Speech rate × Trial number	2.53	4.99	0.51	.615

Significance level notation: $^{***}p<.001$; $^{*}p<.05$, $.p<.1$.

A second model (see Table 1.9) was set up to investigate which individual abilities might modulate the effect of speech rate on pupil peak latency (note though that the pupil peak latency model above showed no speech rate effect). We tested for interactions between speech rate and all participant-related predictors including chronological age (and included all control predictors related to item characteristics).

Table 1.9: Pupil peak latency data (in milliseconds): Model testing for interactions between speech rate and participant-related variables.

Fixed Effects	Estimate	*SE*	$\|t\|$	$p<$
Intercept	946.56	15.17	62.40	
Age	33.05	15.12	2.19	.025*
Speech rate	9.17	8.73	1.05	.272
Target word frequency	-12.09	8.43	1.43	.152
Trial number	-17.97	4.85	3.70	.001***
Target word predictability	-19.78	8.82	2.24	.025*
Target word position	-17.55	8.40	2.09	.037*
SNR	-14.73	8.10	1.82	.069 .
Target word’s number of syllables	0.80	18.14	0.04	.965
Speech rate × Trial number	2.04	4.98	0.41	.683
PTA^{HF}	-2.80	11.16	0.25	.818
Visual acuity	3.42	9.72	0.35	.732
Fluid cognitive processing ability	-24.93	10.04	2.48	.014*
Vocabulary	-8.78	8.88	0.99	.319
Working memory	1.96	7.41	0.26	.775
Speech rate × Age	17.29	9.15	1.89	.059 .
Speech rate × PTA^{HF}	-12.82	6.75	1.90	.058 .
Speech rate × Visual acuity	-2.18	5.82	0.37	.709
Speech rate × Fluid cognitive processing ability	4.87	6.03	0.81	.419
Speech rate × Vocabulary	-1.36	5.29	0.26	.797
Speech rate × Working memory	6.97	4.40	1.58	.114

Significance level notation: ***$p<.001$; *$p<.05$, .$p<.1$.

Again, age showed a significant effect on the timing of the event-related pupil peak ($|t|=2.19$, $p<.05$) with a slower pupil dilation response for older participants. Speech rate did not significantly affect pupil dilation latency ($|t|=1.05$, $p>0.1$). The only participant-related measure that significantly affected pupil peak latency (apart from age) was the composite factor fluid cognitive processing ability ($|t|=2.48$, $p<.05$). Importantly, however, none of the participant-related variables showed significant interactions with speech rate. As in the age group comparison model above, the individual-differences model showed significant effects of target word predictability, target word position in the answer phrase and trial number. Pupil peak latency was shorter for more probable items ($|t|=2.24$, $p<.05$) and the more words of the target speaker were available prior to the target word (|*t*|=2.09, p<.05). Pupil latency also decreased over trials (|*t*|=3.70, p<.001). In sum, contrary to the other dependent variables, our pupil peak dilation data did not show evidence that increased speech rate made spoken word recognition more difficult. Furthermore, none of the included participant-related variables was significantly associated with the size of the speech rate effect for the pupil peak latency.

3.b. Pupil peak amplitude The result of the statistical model testing for the critical interaction between the predictors speech rate and age group (including the control variables discussed in section Analyses above) for pupil peak amplitude is shown in Table 1.10.

Table 1.10: Pupil peak amplitude data (arbitrary unit): Model testing for the *age (group)* × *speech rate* interaction.

Fixed effects	Estimate	*SE*	|*t*|	*p*<
Intercept: Younger adults	361.12	22.53	16.03	
Middle-aged vs. Younger adults	-79.46	31.02	2.56	.014*
Older adults vs. Younger adults	-82.45	31.25	2.64	.012*
Speech rate: Younger adults	0.34	5.40	0.06	.951
Speech rate: Middle-aged vs. Younger adults	3.95	5.86	0.67	.504
Speech rate: Older adults vs. Younger adults	0.66	5.82	0.11	.911
Target word frequency	-10.46	4.06	2.58	.013*
Trial number	-8.65	2.64	3.28	.002**
Target word predictability	9.87	4.24	2.33	.024*
Target word position	12.86	4.05	3.17	.003**
SNR	-0.84	3.91	0.22	.830
Target word's number of syllables	12.83	8.77	1.46	.150
Speech rate × Trial number	-2.12	2.72	0.78	.440

Significance level notation: **p<.01; *p<.05.

Pupil peak amplitude differed considerably between the age groups (|*t*|>2.55, p<.05, for both comparisons with younger adults mapped on the intercept). This is in line with earlier reports of reduced pupil size and less task-evoked pupil dilation for older participants (van Gerven et al., 2004; Birren et al., 1950). Yet, older adults did not differ significantly from middle-aged adults (as shown in a similar model with the middle-aged group on the intercept). The pupil peak amplitude model showed no simple speech rate effect ($|t|<1$, $p>0.1$), nor an *age group* × *speech rate* interaction ($|t|<1$, $p>0.1$ for both age group comparisons). As found for pupil peak latency, trial number affected pupil peak amplitude ($|t|=3.28$, $p<.01$), suggesting task familiarization over the experimental trials. Additionally, the model showed significant effects of target word frequency, target word predictability and target word position in the answer phrase. Unexpectedly, pupil peak amplitude was higher for more probable items ($|t|=2.33$, $p<.05$) and for items that came later in the phrase ($|t|=3.17$, $p<.01$). As expected, we observed a smaller pupil peak amplitude for words with a higher word frequency ($|t|=2.58$, $p<.05$). A second pupil peak amplitude model was set up to test for interactions between speech rate and all participant-related predictors including chronological age (including all control predictors related to item characteristics; see Table 1.11).

In line with the age group analysis above, no speech rate effect was observed nor any significant interactions between speech rate and any of the participant-related variables. Similarly, consistent with the previous *age (group)* × *speech rate* model for the peak amplitude data, effects of trial number, target word frequency, the number of words prior to the target and the probability of the target word were observed (in the same direction). The individual-differences model showed no effect of (continuous) age. The discrepancy regarding the age effect between the age group analysis (see Table 1.10) and the individual-differences model in Table 1.11 suggests that multicollinearity was an issue in the latter more complex pupil peak amplitude model (Table 1.11). As can be seen in Table 1.3, age is correlated with most of the participant-related variables. If correlated variables are fed into the regression analysis simultaneously, variance is inflated resulting in higher standard errors and thus reduced statistical power. We also set up a more parsimonious individual-differences model, leaving out those participant-related variables which were considerably correlated with age ($|r|>0.60$): i.e., high-frequency hearing loss ($\mathrm{PTA}^{\mathrm{HF}}$), fluid cognitive processing ability and visual acuity. As expected, age effects reappeared in this model ($|t|=2.39$, $p<.05$), with reduced pupil dilation amplitudes for older participants. The more parsimonious model was similar to the model presented in Table 1.11 in all other respects. To follow up on this we also conducted the individual differences analyses separately for each age group (models not reported in detail here). In line with the overall model (Table 1.11), none of the age groups showed a speech rate effect. These separate age group models also showed that the effects of trial number, target word predictability and target word position (reported in Table 1.11) were driven mainly by the younger participants. This may relate to age differences in dynamic range of the task-evoked pupil reaction discussed above.

In sum, no speech rate effects were observed on the dependent variable pupil peak amplitude (Research Question 1). The data also did not show evidence for age group differences in the effect of speech rate (Research Question 2). Similarly, none of the included participant-related variables was associated with the size of the speech rate effect on the pupil peak amplitude.

Table 1.11: Pupil peak amplitude data (arbitrary unit): Model testing for interactions between speech rate and participant-related variables.

Fixed Effects	Estimate	*SE*	$\|t\|$	$p<$
Intercept	308.91	14.28	21.63	
Age	-24.37	26.66	0.91	.364
Speech rate	1.72	4.22	0.41	.684
Target word frequency	-10.47	4.07	2.57	.011*
Trial number	-8.65	2.64	3.28	.001***
Target word predictability	9.84	4.26	2.31	.021*
Target word position	12.81	4.06	3.15	.002**
SNR	-0.78	3.92	0.20	.842
Target word's number of syllables	12.81	8.79	1.46	.146
Speech rate × Trial number	-2.03	2.73	0.75	.456
PTA^{HF}	-6.11	19.90	0.31	.760
Visual acuity	2.51	17.18	0.15	.883
Fluid cognitive processing ability	11.17	17.84	0.63	.532
Vocabulary	-6.28	15.68	0.40	.683
Working memory	-2.02	13.20	0.15	.873
Speech rate × Age	2.66	5.04	0.53	.598
Speech rate × PTA^{HF}	-1.33	3.72	0.36	.721
Speech rate × Visual acuity	0.46	3.21	0.14	.887
Speech rate × Fluid cognitive processing ability	-0.44	3.33	0.13	.895
Speech rate × Vocabulary	-2.35	2.92	0.80	.422
Speech rate × Working memory	-3.32	2.43	1.37	.172

Significance level notation: ***p<.001; **p<.01; *p<.05.

1.5 Discussion and conclusion

Speech rate effects in aging have been addressed in numerous studies (e.g., Schmitt and Moore, 1989; Gordon-Salant and Fitzgibbons, 1999; Wingfield et al., 1999). Most of these studies have used artificial time compression to systematically vary speech rate. Possibly, the common observation that older adults show stronger speech rate effects than younger adults is (partly) due to signal degradation caused by time compression techniques (Schneider et al., 2005; Gordon et al., 2009; Kusomoto and Vaughan, 2004) or to the fact that many studies have compressed speech to rates that are higher than typically found in natural speech. The present study was set up to investigate speech rate effects on word recognition across the adult life span by using variation in speech rate within and between speakers as found in a corpus of conversational speech. In addition, to address the different accounts that have been put forward for the *age* × *speech rate* interaction, participant-related variables were collected to study which cognitive, perceptual and linguistic abilities may modulate the size of the speech rate effect on speech recognition. A word recognition task was embedded in a visual-world eye-tracking paradigm, such that multiple dependent variables were obtained at a time (click response times, eye gaze behavior and pupillometry measures). As expected, increased speech rate made word recognition more challenging as evidenced by longer click response times and delayed eye gaze behavior to the target word. Thus, even though our speech materials were less controlled than artificially speeded lab-recorded sentences, rate variation in our conversational stimuli affected ease of word recognition. Furthermore, age effects were observed on click response times, eye gaze behavior, and on the pupil measures, with slower click responses, slower and decreased pupil dilation responses and slightly delayed gaze behavior for the older adults. However, our main question was whether younger, middle-aged, and older adults differ in the ability to keep up with faster rates of speech. None of the dependent variables under investigation showed any convincing evidence that increased speech rate affected older or middle-aged adults more than younger adults. Relatedly, none of the participant-related measures (e.g., hearing sensitivity or fluid cognitive processing ability) modulated the speech rate effect on the different indices of word recognition.

Even though we found consistent effects of speech rate on click response times and gaze behavior, these effects were not found in the pupillometry measures. This may be due to our experimental procedure (i.e., the visual-world paradigm). The fact that participants moved their eyes because of the visual search task may have affected peak detection (resulting either in missing values or incorrect peaks). This may have reduced statistical power of our pupillometry analyses. Note that the pupil dilation curves (provided in Figure 1.5) suggest that pupil peaks are slightly higher and somewhat delayed for higher speech rates, indexing increased processing effort.

Higher natural speech rates not only present listeners with a higher information rate to keep up with (i.e., more words per minute), but also with word forms that are more reduced (Ernestus and Warner, 2011) and hence less redundant (cf. Aylett and Turk, 2004). This effect of less clear articulation was also particularly present in the study by Gordon-Salant and colleagues (2014), who found that older and younger adults had more difficulty with naturally produced fast speech than with artificially speeded speech. Whereas younger and older adults showed equal performance for the normal-rate speech in their study, older adults performed more poorly than younger adults both for the time-compressed and naturally fast materials (thus again showing an *age* × *speech rate* interaction). As argued in the introduction, speech obtained by instructing speakers to read aloud at their ceiling rate (as in Gordon-Salant et al., 2014) may be different from speech varying in tempo as encountered in everyday conversations. In our study, fragments were taken from a corpus of conversational speech in which speakers speak at their habitual rate or deliberately choose to speak at a particular tempo. Possibly, pushing speakers to speak faster than they would normally do (with no communicative intent) may yield more slurring and acoustic reduction than present in our materials. Only more extreme fast and slurred articulation might have affected older adults more adversely than younger adults.

The combined pattern of results thus converges on speech rate effects being similar across age groups for conversational speech fragments. Note that this may be because our older adults had relatively good hearing as they were not eligible for hearing aids. The different accounts of the *age* × *speech rate* interaction have either emphasized the role of age-related hearing loss or cognitive decline (cognitive slowing in particular). Hearing loss did not affect our dependent variables (in models in which age was also included), nor did it interact with the effect of speech rate. Fluid cognitive processing ability, measuring cognitive slowing, affected click response times and pupil dilation latency in the expected direction, but did not modulate effects of speech rate. Apart from cognitive and hearing abilities, we also expected linguistic abilities to facilitate word recognition. Participants with better vocabulary knowledge were shown to have faster click responses. Thus, speech processing may be facilitated by hearing, cognitive and linguistic abilities, but they were not found to modulate effects of speech rate. Therefore, our findings emphasize that earlier claims about *age* × *speech rate* interactions mainly obtained with artificial time compression may not generalize to natural speech rate variation as encountered in conversational speech, at least not for an older adult sample with relatively good hearing.

Type of speech material affects Acceptable Noise Level test outcome

Chapter 3

This chapter is based on:
Xaver Koch, Gertjan Dingemanse, André Goedegebure and Esther Janse (2016)
Type of speech material affects Acceptable Noise Level test outcome
Frontiers in Psychology 186 (7), pages 1–14

The acceptable noise level (ANL) test, in which individuals indicate what level of noise they are willing to put up with while following speech, has been used to guide hearing aid fitting decisions and has been found to relate to prospective hearing aid use. Unlike objective measures of speech perception ability, ANL outcome is not related to individual hearing loss or age, but rather reflects an individual's inherent acceptance of competing noise while listening to speech. As such, the measure may predict aspects of hearing aid success. Crucially, however, recent studies have questioned its repeatability (test-retest reliability). The first question for this study was whether the inconsistent results regarding the repeatability of the ANL test may be due to differences in speech material types used in previous studies. Second, it is unclear whether meaningfulness and semantic coherence of the speech modify ANL outcome. To investigate these questions, we compared ANLs obtained with three types of materials: the International Speech Test Signal (ISTS), which is non-meaningful and semantically non-coherent by definition, passages consisting of concatenated meaningful standard audiology sentences, and longer fragments taken from conversational speech. We included conversational speech as this type of speech material is most representative of everyday listening. Additionally, we investigated whether ANL outcomes, obtained with these three different speech materials, were associated with self-reported limitations due to hearing problems and listening effort in everyday life, as assessed by a questionnaire. ANL data were collected for 57 relatively good-hearing adult participants with an age range representative for hearing aid users. Results showed that meaningfulness, but not semantic coherence of the speech material affected ANL. Less noise was accepted for the non-meaningful ISTS signal than for the meaningful speech materials. ANL repeatability was comparable across the speech materials. Furthermore, ANL was found to be associated with the outcome of a hearing-related questionnaire. This suggests that ANL may predict activity limitations for listening to speech-in-noise in everyday situations. In conclusion, more natural speech materials can be used in a clinical setting as their repeatability is not reduced compared to more standard materials.

3.1 Introduction

One of the most frequent complaints of adult hearing aid users is that comprehending speech is challenging in noisy environments (Cord et al., 2004; Killion et al., 2004; Nábělek et al., 2006). Indeed insufficient benefit of hearing aids in noisy situations seems to be an important reason for people fitted with a hearing aid not to use it. Hearing rehabilitation could be better attuned to the needs of hearing-impaired individuals if audiologists were able to identify those hearing-impaired individuals who will have problems with accepting higher noise levels in everyday communication situations. Individualized counseling may help hearing-impaired individuals to set realistic expectations of hearing-aid benefit in noise. Furthermore, the use of assistive listening devices could then be applied early on for individuals who can be expected to be unsatisfied with hearing devices in noisy environments in order to ultimately minimize disappointment with the device, activity limitations and participation restrictions related to hearing disabilities (cf. Nábělek et al., 2006; Kim et al., 2015).

This raises the question of how to identify future hearing aid users who may be discouraged from using hearing aids because of difficulty listening in noise. One obvious approach would be to measure the individual's objective ability to understand speech in noise (e.g., the standard speech-reception threshold measure). However, such objective performance measures are not predictive of hearing aid benefit or success (Bender et al., 1993; Humes et al., 1996; Nábělek et al., 2006). In contrast, one subjective measure called 'Acceptable Noise Level' or 'tolerated SNR' (henceforth, ANL) seems to be predictive of hearing aid and cochlear implant success (Nábělek et al., 1991, 2006; Bender et al., 1993; Humes et al., 1996; Plyler et al., 2008; but cf. Olsen and Brännström, 2014). The ANL procedure involves the following two steps: listeners are first asked to indicate the loudness level they find most comfortable (henceforth, Most Comfortable Loudness Level (MCL), cf. Hochberg, 2008) for listening to a continuous speech signal. In a second step, listeners adjust the background noise level (henceforth, Background Noise Level (BNL)) to the maximum level they are willing to put up with while following the running speech presented at their individual MCL level. Subtracting the BNL value from the MCL value yields the ANL measure which typically ranges between -15 and 40 dB with a mean of around 5 to 12 dB (cf. Nábělek et al., 1991, 2006; von Hapsburg and Bahng, 2006; Eddins et al., 2013; Walravens et al., 2014). The lower the ANL value, the more noise the participant accepts while listening to speech. The ANL measure quantifies the individual's "willingness to listen to speech in background noise" (cf. Nábělek et al., 2006, p. 626). As such, it may be a better indicator of successful hearing aid uptake than the individual's objective ability to understand speech in noise as it is more telling about the individual's wishes, motivation, and intentions.

Speech perception is generally considered to involve an interaction between the processing of acoustic information (bottom-up processing) and linguistic and cognitive processing (top-down processing). An important question is how ANL outcome relates to this interaction, as participants are explicitly instructed to 'follow the speech' during the ANL task. Even though listeners may engage in setting up linguistic hypotheses about upcoming content when the signal is clear, top-down contextual support may be particularly helpful in reconstructing the message when the signal is presented in noise. It is unclear whether type of speech material affects ANL. The original ANL publications (e.g., Nábělek et al., 1991, 2006) used a standard stretch of read speech, making up a coherent story (the Arizona Travelogue passage). In contrast, Olsen and Brännström (2014) used the International Speech Test Signal (ISTS; Holube et al., 2010), which is non-meaningful by definition as the signal consists of roughly syllable-sized units from six different languages and speakers, concatenated into a continuous speech stream. Olsen and Brännström (2014) argue that the ISTS can be used to compare ANL values across languages. However, the use of the ISTS precludes top-down processing. In that sense, the question whether type of speech material affects ANL outcome is a question about the nature of the ANL task in the broader context of models of speech processing. Regarding the question of whether meaningfulness affects ANL outcome, ANLs obtained with unintelligible speech (i.e., reversed or unfamiliar speech) have been found to be higher (i.e., indicative of lower noise tolerance) than those obtained with intelligible speech (Gordon-Hickey and Moore, 2008). In contrast, Brännström et al. (2012) showed that ANLs were lower for the ISTS in comparison with meaningful speech stimuli. We investigate whether ANL depends on meaningfulness and coherence by using three different stimulus types that differ in meaningfulness (ISTS vs. concatenated sentences and fragments of conversational speech) and coherence (concatenated sentences vs. coherent conversational speech). If meaningfulness of the test material does not affect ANL outcome, listeners' acceptance of noise while following speech may mainly rely on bottom-up processing. Consequently, following speech in noise as captured by the ANL task would deviate from speech perception and comprehension. In line with Gordon-Hickey and Moore (2008), we expect to find increased ANL values for the non-meaningful ISTS material compared to the meaningful materials. Our hypothesis regarding the direction of a semantic coherence effect is that participants will accept more noise (i.e., show lower ANLs) for the conversational stimulus type in comparison with the passage of concatenated sentences as redundant information is available on the discourse level, which facilitates speech comprehension. Alternatively, however, the faster speech rate and less careful articulation observed in conversational speech may make listening harder than in the sentence materials and may yield lower noise acceptance.

In order for ANL to be a clinically useful tool in hearing rehabilitation, it is important to establish its repeatability (i.e., consistency over repeated measures or test-retest reliability with the exact same materials). Olsen and Brännström (2014) questioned the repeatability of the existing ANL procedures using the ISTS material. In the present study we investigate whether speech material type affects ANL outcomes and repeatability. Relatedly, repetition of the exact same materials may lead to substantial priming effects, especially for the meaningful materials. Consequently, participants would accept more noise upon repeated exposure, yielding a lower repeatability. We investigate whether the use of meaningful materials yields differential repeatability compared to non-semantic ISTS material. Nábělek et al. (2006) suggest that future hearing aid use can be predicted on the basis of ANL outcome for a majority of hearing aid candidates. Olsen and Brännström (2014), however, challenge the predictive value of ANL outcome for hearing-aid use, and report that results regarding the association between ANL and self-reported hearing-aid outcome measures have been mixed. These inconsistent findings may be caused by the multitude of variables that are possibly related to hearing-aid use, hearing-aid satisfaction and hearing-aid success, as reviewed by Knudsen et al. (2010) and McCormack and Fortnum (2013). Note, however, that self-reported hearing problems have been shown to be consistently associated with hearing-aid outcome measures obtained throughout the process of getting a hearing aid (help seeking, hearing-aid uptake, use, and satisfaction). We investigate whether ANL is associated with (specific components of) the Speech, Spatial, and Qualities of Hearing self-report questionnaire (SSQ; Gatehouse and Noble, 2004) and whether this relation depends on ANL test material type. Our expectation is to find differential correlations between the questionnaire outcome and ANL for three speech stimulus types with stronger associations for the more ecologically valid materials.

The central concept of the ANL measure is ‘Listening comfort’. Thus, individual ANLs are not necessarily linked to the listener’s objective ability to comprehend speech in noise, as shown in a number of studies (cf. Nábělek et al., 2004; Mueller et al., 2006; von Hapsburg and Bahng, 2006; Plyler et al., 2008, but cf. Gordon-Hickey and Morlas, 2015). Whether and how the concept of comfort in noisy listening situations relates to listening effort is unclear. The clinical meaning of the concept of listening effort has recently been discussed in several papers (McGarrigle et al., 2014; Rennies et al., 2014; Francis and Füllgrabe, 2015; Schulte et al., 2015). One way to quantify listening effort is to ask participants to fill in effort-related subscales of self-report questionnaires (cf. McGarrigle et al., 2014). We therefore investigate whether listening effort, as measured with specific questions of the SSQ

(Akeroyd et al., 2008) is associated with ANL. We hypothesize that ANL is associated with a listening effort-related subscale of the SSQ with more subjective listening effort related to lower noise acceptance (i.e., higher ANLs).

Listeners need cognitive capacity to map a noisy signal onto stored representations (McGarrigle et al., 2014), as laid out in the Ease of Language Understanding model (Rönnberg et al., 2008, 2013). Multiple studies have shown that hearing aid users' objective speech understanding in adverse conditions (such as background noise) is related to their working memory capacity, verbal working memory in particular (Akeroyd, 2008; Rudner et al., 2011; Ng et al., 2013, 2014). Given the relatively large amount of unexplained variance for individual ANLs, ANLs may also be associated with working memory. Brännström et al. (2012) found a significant correlation between working memory capacity and ANL for a sample of normal-hearing participants, with lower noise acceptance (i.e., higher ANLs) relating to poorer working-memory capacity. We investigate whether ANL outcomes obtained with the different types of speech materials relate to listeners' working memory capacity, where we expect to replicate the results of Brännström et al. (2012).

As ANL specifically asks listeners about their willingness to accept noise, ANL may be related to personality traits. Indeed, self-control abilities (i.e., the capability to control thoughts, feelings, impulses and performance; Baumeister et al., 1994), have been found to predict ANL outcomes (Nichols and Gordon-Hickey, 2012). We revisit the question to what extent ANL outcome relates to personality characteristics in this study. We expect to replicate effects of self-control on ANL with better self-control related to lower ANLs (cf. Nichols and Gordon-Hickey, 2012). Furthermore, even though earlier studies have not found a link between ANL and age (Nábělek et al., 1991; Moore et al., 2011), nor between ANL and pure-tone hearing thresholds (Nábělek et al., 1991; Freyaldenhoven et al., 2007; Plyler et al., 2007), or between ANL and speech perception accuracy in noise (Nábělek et al., 2004), we investigate whether our data replicate this pattern of results.

This study investigates whether speech material type affects ANL outcomes and repeatability for a reference sample of normal-hearing middle-aged and older participants. As addressing these questions on speech material and repeatability involves relatively long testing sessions with repeated ANL measurements, we tested a non-clinical population first so as not to burden a patient population. Future testing is then required to see whether material type effects generalize to a patient population and whether ANLs based on conversational materials better predict hearing aid success than ANL values obtained with more standard audiology materials (such as, e.g., ISTS).

The present study was set up to address the following four research questions:

(1) Does ANL outcome depend on the meaningfulness (1A) and semantic coherence (1B) of the speech materials?

(2) Does ANL repeatability differ across speech material types?

(3) Are ANLs differentially associated with self-report measures of listening effort and of hearing-related activity limitations for the different speech materials?

(4) Do participant characteristics such as working-memory (4A), and self-control abilities, age, hearing thresholds, and speech perception in noise predict ANL (4B)?

3.2 Materials and Methods

3.2.1 Participants

Seventy-one adults were recruited, all native speakers of Dutch, above 30 years of age (39 female, 33 male). From the initial sample, we excluded 10 participants whose hearing loss in one or both ears exceeded the Dutch health insurance criterion for partial reimbursement of hearing aids (i.e., pure-tone average over 1000, 2000, and 4000 Hz $\geq$ 35 dB HL in either ear). We also excluded two participants who suffered from tinnitus and one participant who showed significant binaural low-frequency hearing loss. One participant was excluded because she did not manage to perform the ANL task in the training phase. The 57 remaining participants (34 female, 23 male) ranged in age from 30 to 77 years with an overall mean of 60.7 years (*SD*=11.0). All participants indicated that they had no hearing impairment and did not use hearing aids. None of the participants had a history of a neurological disease. We followed the protocols of the Radboud University Ethics Assessment Committee for the Humanities. All participants provided written informed consent and were informed that they could withdraw from the study at any time.

3.2.2 Speech Stimuli

Three types of speech materials were used for ANL testing that differed in meaningfulness and semantic coherence: the unintelligible speech-like ISTS (Holube et al., 2010), a concatenated passage of meaningful Dutch sentences taken from speech material developed by Versfeld et al. (2000; henceforth, SENT), and conversational speech (henceforth, CONV) extracted from the Dutch conversational IFADV corpus (van Son et al., 2008). The 60 s long ISTS signal is made up of units that are roughly syllable sized, originating from six female speakers each reading a short standard passage in their native language (being Mandarin, Spanish, English, German, French, and Arabic). The ISTS signal had been developed on the basis of an automatic procedure to cut, concatenate and reassemble the roughly syllable sized segments from the original six recordings to create a smooth 60 s long speech-like signal including pauses at regular intervals (all pause durations being smaller than 600 ms). The resulting speech rate is approximately 4 syllables per second (Holube et al., 2010). Furthermore, the ISTS signal has been shaped to spectrally match the female international long-term-average speech spectrum (ILTASS, Byrne et al., 1994).

To create the second type of material (SENT), we concatenated fifty sentences from the female speaker of the materials of Versfeld et al. (2000) with intervals of 500 ms silence between sentences (total duration of the passage was 120 s). These sentences are all between five and eight words long and are semantically coherent. A translated example sentence is: ‘I hope to be able to catch the train’. The speech rate of the sentences ranges between 3.5 to 5.7 syllables per second (*M*=4.6 syllables/s, *SD*=0.6). In order to match the spectral properties of the SENT materials to the ISTS materials, the concatenated SENT material was filtered to the ILTASS (combination of male and female signal) using a finite impulse response (FIR) filter between 100 and 16000 Hz.

The third type of speech material was created by extracting two male and two female recordings from the conversational IFADV corpus (van Son et al., 2008). The Dutch open-source IFADV corpus consists of annotated high-quality recordings of dialogs on daily topics such as problems in public transport, leisure time activities or vacations. As we wanted to spectrally shape these materials, we selected four longer stretches of speech (CONV1 (female speaker), CONV2 (male speaker), CONV3 (male speaker), CONV4 (female speaker)) where only one speaker was speaking, without being interrupted by the dialog partner. These stretches were based on the available corpus annotations. In a few instances we cut out verbal backchannelling (e.g., ‘yes’, ‘hmm’) of the interlocutor, which did not overlap with the target speech. All pauses longer than 500 ms were shortened to 500 ms. The four resulting speech files ranged in duration between 63 and 75 s. Speech rate calculated over the breath groups (sequence of words between inhalations) ranged between 2.6 and 7.5 syllables per second (*M*=5.7 syllables/s, *SD*=1.2; CONV1: 6.10 syllables/s, CONV2: 5.10 syllables/s, CONV3: 5.79 syllables/s, CONV4: 5.89 syllables/s). In order to match the spectral contents of the conversational materials to the other types of materials, the four conversational fragments were also filtered to the ILTASS (combination of male and female signal) using a FIR filter between 100 and 16000 Hz.

3.2.3 Noise Material

The noise stimulus used throughout the ANL test procedure was a non-stationary eight speaker babble noise (BAB8; Scharenborg et al., 2014) filtered to the ILTASS (combination of male and female spectrum) using a FIR filter between 100 and 16000 Hz. In line with the idea of aiming to approximate realistic listening conditions, we used a multi-talker babble noise since it is a typical background sound encountered in daily life.

3.2.4 Experimental Procedure

Test Set-Up

All ANL test materials were presented in a sound-attenuated booth using an Alesis multimix 4USBFX device and Behringer MS16 loudspeakers in front of the listener (0° azimuth) at a distance of 1 m. Stimuli were presented in a custom application (cf. Dingemanse and Goedegebure, 2015) running in `MATLAB` (v7.10.0) on a MacBook Pro (type 9,1). Participants adjusted the sound level of the speech stimuli or the noise file using the up and down keys of a customized keyboard. The starting intensity for the MCL was 45 dB (SPL). The intensity of the speech file for the BNL task was set to the mean of the three measurements in the preceding MCL task. The step size for the intensity adjustment for both tasks was fixed at 2 dB per button press.

All speech and noise materials were scaled to have the same overall level in dB (RMS). Sound level calibration was done using a 2250 Brüel and Kjær real time sound analyzer and a 1000 Hz warble test tone with the same RMS-value as the ANL materials.

ANL Instructions

Participants were instructed to first adjust the level of the speech until it was too loud (i.e., up to the first deviation point), then to reduce the intensity until the speech became very soft (being the second deviation point) and lastly find the MCL. Then the participant's task was to select the maximum BNL they were willing to accept while following the speech at their MCL. They were instructed to use the same pattern of adjustments as described for MCL: turn up the volume of the noise until it was too loud to comfortably listen to the speech (i.e., the first deviation point), then to reduce the noise intensity until the speech became very clear (i.e., the second deviation point) and lastly to find the maximal background noise level they were willing to put up with while following the speech signal (BNL).

Familiarization Phase

In order to familiarize participants with the ANL procedure prior to actual testing, each participant was presented with a phonetically balanced Dutch training fragment. A 2-min-long recording of a female Dutch speaker reading a standard text passage ("Dappere fietsers" – 'Brave cyclists') served as training material. The noise stimulus (BAB8) used throughout the actual ANL test (BNL part) also served as background noise during the training session. Participants first received written instructions on the experimental task (which was a Dutch translation of the instruction provided in Nábělek et al., 2006, p. 639). The experimenter then demonstrated the task, using scripted instructions, which again followed the translation of Nábělek et al. (2006). A visual display was available during the familiarization phase that enabled the participant, as well as the experimenter, to see the course of the presentation level during the MCL and the BNL tasks. Each participant had to demonstrate the expected intensity pattern (up-down-final adjustments, cf. deviation points above) three times in a row for both MCL and BNL components before they could proceed with the test phase.

Test Phase

Unlike during the familiarization phase, visual output was available only to the experimenter during the ANL test sessions. Participants had to perform the MCL and BNL tasks for each of the six ANL test stimuli, and each of the two tasks was repeated three times in a row to decrease measurement error (cf. Brännström et al., 2014a; Walravens et al., 2014). The ANL for each fragment and for each participant was calculated by subtracting the mean BNL from the averaged MCL. Note that stimulus presentation was looped such that if participants had not provided their response before the end of the stimulus, the stimulus was automatically repeated. All participants managed to set the MCL and BNL levels within the stimulus duration in the test phase (minimal duration: 60 s for the ISTS).

Test Repetition

In order to test the repeatability of the ANL measures across the different materials, we asked the participants to do the ANL task twice for each stimulus type (ISTS, SENT, CONV) with exactly the same material. Note that we took into account that the repetition of the exact same materials across sessions could lead to substantial priming effects, especially for the meaningful materials, by including a control variable in our models to capture changes in ANL over test sessions. Participants first performed the ANL test with the different materials at the beginning of the test session, and again (approximately 1 h later) toward the end of the session. Participant characteristics data were collected in between these two ANL test sessions. During the first ANL session (session I), six different fragments were presented: ISTS, SENT, CONV1, CONV2, CONV3, and CONV4. To restrict testing time, we only presented one fragment for each of the three material types in the test repetition (session II): ISTS, SENT, and CONV4. We selected the CONV4 stimulus from the four conversational test fragments because it featured a female speaker (as was the case for the ISTS and the SENT material) and because its speech rate was typical for conversational speech (i.e., 5.89 syllables per second).

Randomization

We used a block-wise randomization procedure to minimize presentation order effects for the material types. Each participant was pseudorandomly assigned to one out of six possible block orders for the speech material types (ISTS, SENT, CONV). The order of the presented speech material types for the second test session (session II) matched the order of session I. The order in which the four conversational materials appeared in the first ANL test session was also randomized. Each participant was randomly assigned one out of 24 possible presentation orders for the conversational speech stimuli.

3.2.5 Tests of Participant Characteristics

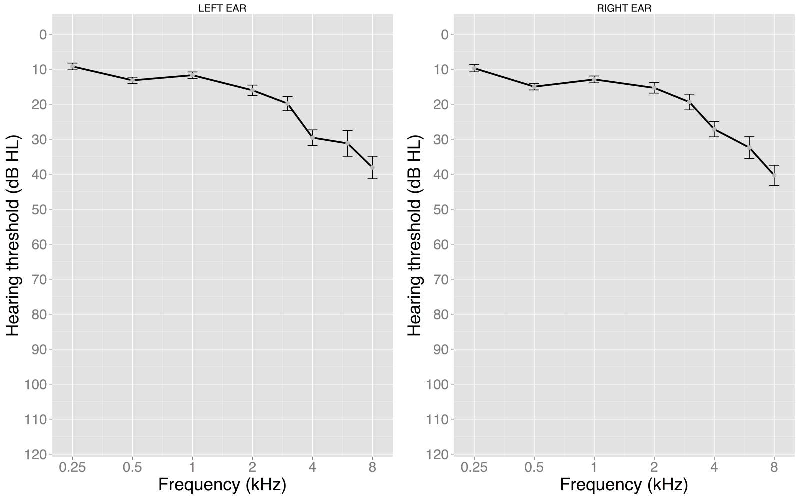
Hearing (Pure-Tone Average)

Hearing status was screened with air conduction pure-tone audiometry using the modified Hughson-Westlake technique for octave-frequencies between 250 and 8000 Hz, including two half-octave frequencies of 3000 and 6000 Hz (see Figure 3.1). Audiometric averaged thresholds were calculated for the better ear as auditory presentation of the ANL test was binaural. Seven participants showed an asymmetric hearing loss, defined as an interaural difference of more than 10 dB averaged over 500, 1000, 2000, and 4000 Hz (Noble and Gatehouse, 2004). In addition to the pure-tone average over 1000, 2000, and 4000 Hz, we calculated high-frequency $\mathrm{PTA}^{\mathrm{HF}}$ as the mean threshold over 3000, 4000, 6000, and 8000 Hz. Table 3.1 displays descriptives for the two PTA measures. Higher values indicate poorer hearing.

Speech Perception in Noise

Speech perception in noise was tested using a standard Dutch speech audiometry test, the CVC word material from Bosman and Smoorenburg (1992, 1995), which is common in clinical practice in the Netherlands. The test allows presenting the

Figure 3.1: Mean audiometric pure-tone air conduction thresholds (for left and right ear) as a function of frequency. Error bars represent standard errors.



materials at SNRs which are reasonably representative of noise levels during everyday communication (Smeds et al., 2015). This test material consists of meaningful monosyllables (e.g., "kaas", 'cheese') produced by a female speaker arranged in lists of 12 words. The material was presented in a sound-attenuated booth using Behringer MS16 loudspeakers placed in front of the listener (0° azimuth) at a distance of one meter. The CVC words were presented at an intensity level of 65 dB (SPL) mixed with a masking noise of the same intensity (long-term-average spectrum of the recorded speaker). The test score was based on the number of correctly reproduced phonemes (max. three per test item), discarding the first item of each list (which is considered a practice item). Based on Bosman and Smoorenburg's standardizations results, we expected a mean phoneme accuracy score of about 80-85% for normal hearing adult participants at an SNR of 0 dB (more favorable signal-to-noise ratios may thus lead to ceiling effects in performance). All participants were presented with five consecutive lists (list 31–35), which resulted in a maximum accuracy score of 165 phonemes correct (5 lists × 11 items × 3 phonemes). The speech perception in noise score reported here was quantified as the percentage of correct phonemes produced. Table 3.1 provides the descriptives for the perception in noise score. Higher values indicate better speech perception in noise.

Reading Span

We used a Dutch version of the well-established reading span test to index working memory (cf. Daneman and Merikle, 1996; Besser et al., 2013; Besser, 2015). The Dutch test consists of 54 grammatically correct sentences, consisting of a noun phrase plus verb phrase. The 54 sentences are divided in 12 sets of three, four, five, or six consecutive sentences. Half of the 54 sentences make sense (e.g., ‘The student sang a song’); the other half is absurd (e.g., ‘The daughter climbed the past’). The sentences were presented orthographically in chunks: first the subject noun phrase was presented (determiner-noun, e.g., ‘The student’), followed by the verb (e.g., ‘sang’), followed by the object noun phrase (determiner-noun, e.g., ‘a song’; cf. Besser, 2015, p. 173). We used `E-PRIME` (2.0, Psychology Software Tools) to present the chunks of the respective test sentences (Subject, Verb, and Object) consecutively on a computer screen (display time of each chunk: 800 ms, blank inter chunk interval: 75 ms). Font size was 36 pt (Verdana). The primary unspeeded task was to repeat back either the first or the last nouns of the respective test set ranging in length from three to six consecutive sentences. Thus, participants were visually prompted to (orally) recall either the subject noun phrases (first nouns) or the object noun phrases (last nouns) of the 12 test sets. The order in which participants recalled the first or last words was not taken into consideration for the scoring (cf. Besser et al., 2013). Additionally, participants were asked to perform a speeded plausibility judgment after each sentence as a secondary task. This task ensured that participants read and comprehended the sentences. Response time was restricted by imposing a time out of 1.75 s after a visual prompt appeared that initiated the plausibility judgment task. Participants gave their plausibility judgment by either pressing a red (i.e., absurd) or a green button (i.e., makes sense) on a customized standard keyboard. Participants received written task instructions and completed a training test set before the actual test started. Reading span score was quantified as the percentage of correctly recalled nouns across the 12 sets. Table 3.1 displays the descriptives for the reading span test. Higher values indicate better working memory capabilities.

Self Control

Participants filled in a Dutch translation of the Brief Self-Control Scale, a 13 items questionnaire using a five-point Likert scale (Tangney et al., 2004; cf. Kuijer et al., 2008). Individual test scores were quantified as the percentage of points out of the maximum of 65 points. Table 3.1 displays the descriptives for the self-control predictor variable. Higher values indicate better self-control abilities.

Table 3.1: Descriptives for the participant characteristics.

	M	*SD*	Range
Age (years)	60.72	11.04	30 – 77
PTA (dB HL)	16.05	8.16	0 – 31.67
PTA^{HF} (dB HL)	25.09	15.68	-1.25 – 56.25
Speech perception in noise (% correct)	88.22	6.79	67.88 – 96.36
Reading Span (% correct)	28.43	10.73	0 – 48.15
Self-Control Scale (% of maximum)	67.34	12.05	38.46 – 93.85
SSQ Part 1 ‘Speech hearing’ (mean score)	7.07	1.07	4.86 – 9.36
SSQ Part 3 ‘Qualities of hearing’ (mean score)	7.98	0.93	5.50 – 9.83
SSQ ‘effort and concentration’ (mean score)	6.55	1.71	3.00 – 9.50

SSQ Questionnaire

Prior to the ANL testing session, participants filled in an online (Dutch) version of the Speech, Spatial and Quality of Hearing Scale (SSQ, Gatehouse and Noble, 2004). The SSQ self-report scale, which consists of 49 items, is subdivided into three parts: Part 1: ‘Speech hearing’ (14 questions), Part 2: ‘Spatial hearing’ (17 questions), and Part 3: ‘Qualities of hearing’ (18 questions). Following Akeroyd et al. (2008), we extracted a factor related to listening effort covering question numbers 15 and 18 of the SSQ subscale ‘Qualities of hearing’ (‘Do you have to put in a lot of effort to hear what is being said in conversation with others?’; ‘Can you easily ignore other sounds when trying to listen to something?’). Hence, we calculated the SSQ ‘effort and concentration’ subscale by averaging scores over these two questions. We also calculated the average over the first and the third SSQ scale as these two were deemed most relevant. Table 3.1 presents the descriptive values for averaged SSQ ‘Speech hearing’ and ‘Qualities of hearing’ scores, as well as for the factor related to listening effort (SSQ ‘effort and concentration’). Higher values on the SSQ scale indicate fewer limitations in self-reported activity due to hearing problems. Table 3.2 provides a correlation matrix of all the participant-related characteristics.

Table 3.2: Correlation matrix with correlation coefficients and significance levels for participant characteristics (Spearman's rank, uncorrected).

	Age	PTA^{HF}	SPIN	RST	SCS	SSQ^{1}	SSQ^{3}	SSQ^{EC}
Age								
PTA^{HF}	.42**							
Speech perception in noise (SPIN)	-.48***	-.71***						
Reading span (RST)	-.35***	-.28*	.51***					
Self-control scale (SCS)	.08	.07	.01	-.06				
SSQ 'Speech hearing' (SSQ^{1})	-.19	-.08	.22	-.03	.39**			
SSQ 'Qualities of hearing' (SSQ^{3})	-.17	.01	.21	-.06	.39**	.65***		
SSQ 'effort and concentration' (SSQ^{EC})	-.10	-.07	.17	-.02	.34**	.54***	.64***	

Significance level notation: ***p<.001, **p<.01, *p<.05.

3.3 Analyses

RQ1

Two separate statistical regression models were run to investigate the effects of meaningfulness and coherence (RQ1) of the test material on ANL, using linear mixed-effect models with participants as random variable. The program `R` was used with the `lme4` package (Bates et al., 2013) and restricted maximum likelihood estimation. P-values were calculated using the `Anova` function of the `car` package which calculates type II Wald χ^2 values. The categorical within-subject variable meaning fulness included two levels: not meaningful (ISTS material) vs. meaningful (CONV and SENT material). The within-subject variable coherence featured two categories: coherent on sentence level (SENT material) vs. coherent on discourse level (CONV material). Block order (order a–f) was included as additional control variable in all models. For the model on meaningfulness (model 1A), we allowed for the possibility that the effect of meaningfulness differed across participants by including a random participant slope for meaningfulness. Similarly, we allowed for the possibility that the effect of semantic coherence differed across participants by including a random participant slope for meaningfulness in the 'coherence' analysis (model 1B). Note that we also included the interaction between session number and meaningfulness (in model 1A) or between session number and coherence (in model 1B), to allow for the possibility that ANLs may systematically change with session number due to semantic priming. Consequently, we also allowed for the possibility that the effect of session number differed across participants by including a random participant slope for both models (model 1A, model 1B).

RQ2

We first ran a linear mixed-effect model (with random intercepts for participants) with ANL differences between test sessions as dependent variable. The question was whether ANL values obtained for the three types of speech materials differed in their repeatability across test sessions. One outlier was excluded from repeatability analysis of the ISTS material as the ANL difference between sessions I and II of this participant exceeded a threshold of the sample mean plus three standard deviations.

Apart from the mixed-effect analysis described above, we followed the procedures described by Brännström et al. (2014a) to assess the repeatability of the three speech materials. Hence, we inspected the Bland-Altman plots (Bland and Altman, 1986; Vaz et al., 2013) as well as the coefficient of repeatability (henceforth, CR) for each of the three test materials for which two test sessions had been run. The CR measure is a repeatability (test-retest reliability) measure. It indicates the size of the measurement error in its original measured unit (i.e., dB). In our case, it represents the size of the difference between one measurement (session) and another measurement using the exact same material (with 95% confidence level). The Bland-Altman plots show for each of the three speech materials (ISTS, SENT, CONV4) each participant's mean ANL over the two sessions on the x-axis against the difference between the two sessions on the y-axis. The CR was calculated for each material by multiplying the standard deviation of the differences between ANLs (averaged over repetitions) for the two sessions with 1.96. Additionally, we calculated the coefficients of repeatability for all test materials (i.e., incl. CONV1, CONV2, and CONV3) over their three repetitions within test sessions (repetition 1 vs. repetition 2; repetition 2 vs. repetition 3). This enabled us to analyze whether repeatability changed within and across test sessions.

RQ3

To assess the question whether self-reported hearing related activity limitations and listening effort differentially predict ANL outcomes for the three different speech materials (RQ3) we set up four linear mixed-effect models that included a categorical speech material variable (ISTS, SENT, CONV) in interaction with one of three variables derived from the SSQ scale (SSQ Part 1, SSQ Part 3, SSQ 'effort and concentration'). Session number was added as categorical covariate to capture repetition effects due to semantic priming. Again, we allowed for the possibility that the effects of session number and speech material differed across participants and therefore added random slopes for the variable speech material and session number to the model.

RQ4

To investigate the effects of participant characteristics (age, hearing thresholds, speech perception in noise accuracy, working memory, and self-control abilities) on ANL for the three speech materials (RQ4) we performed 15 correlation analyses (Pearson's *r*) and Bonferroni corrected for multiple comparisons. ANL values were pooled across the two test sessions.

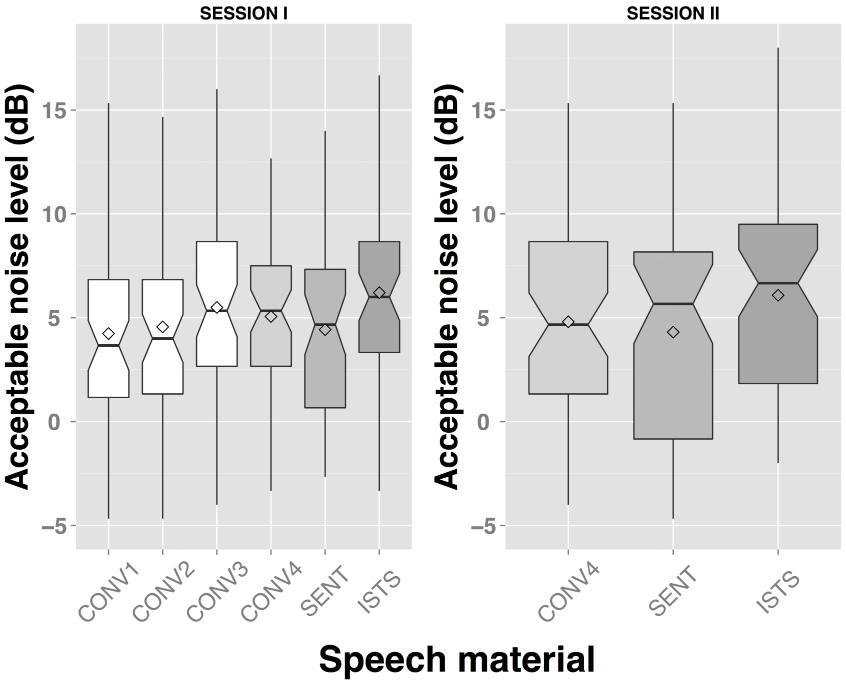
3.4 Results

Table 3.3 shows the ANL test results per speech material per test session for the three unrepeated conversational materials (CONV1-3) and the three repeated materials (CONV4, SENT, ISTS). Mean ANLs are higher for the ISTS material than for the meaningful materials. Figure 3.2 gives an overview of the ANL test results per test session including the conversational materials that were only presented in test session I (i.e., CONV1, CONV2, and CONV3).

Table 3.3: Acceptable noise level (ANL) descriptive statistics for the six speech materials and the two test sessions (in dB).

Test material	Test session I		Test session II	
	M	*SD*	*M*	*SD*
CONV1	4.06	4.59	–	–
CONV2	4.39	4.58	–	–
CONV3	5.50	4.29	–	–
CONV4	5.30	4.43	4.81	4.53
SENT	4.32	5.57	4.13	5.24
ISTS	6.25	4.90	5.84	5.25

Figure 3.2: Acceptable noise level (ANL) test results per speech material and per test session. Note that the notch plots include a marker for the mean (diamond symbol).



Research Question 1A: Does ANL outcome depend on the meaningfulness of the speech material?

The results of the statistical model (cf. Table 3.4) showed that ANLs for the meaningful materials (SENT, CONV) were significantly different from those for the non-meaningful ISTS material ($\chi^2(1, n{=}341){=}17.98, p{<}.001$). Participants showed 1.46 dB higher ANLs and thus less noise acceptance for the ISTS signal in comparison with the meaningful materials. The observed effect direction matched our a priori hypothesis that participants would accept less noise for the non-semantic ISTS material than for the meaningful materials. Block order of presentation did not influence ANL, nor did session number. These control variables also did not interact with the meaningfulness of the test material. The absence of a significant effect of session number on ANL suggests that ANL was stable over sessions and that no semantic priming occurred between sessions. This absence of priming held across material types as the meaningfulness × session number interaction was insignificant. Block order did not affect the ANL outcome, which suggests that our randomization procedure was adequate. For reasons of brevity block order is left out in the model presented below (the variable having six levels; $\chi^2(5, n=341)=2.13, p>.1$).

We also investigated the effect of meaningfulness including all conversational materials (this implies that it can only be assessed for session I). To that end, we averaged ANLs per participant over the conversational materials (CONV1 - CONV4). In line with the results presented in Table 3.4, this analysis showed an effect of meaningfulness on ANL with less noise acceptance for the non-meaningful ISTS material compared to the two types of meaningful materials ($\chi^2(1, n=170)=18.47, p<.001$).

Research Question 1B: Does ANL outcome depend on the semantic coherence of the speech material?

A significant effect of coherence was observed with higher ANLs for the material with coherence on discourse level, i.e., the conversational material ($\chi^2(1, n=227) =6.04, p<.05$) than for the concatenated sentences (cf. Table 3.5). Thus, for the conversational test material participants accepted less background noise. The size of the effect was 1.05 dB. The observed direction of the effect matched the hypothesis that participants would accept less noise for the conversational material, which was coherent at the discourse level, but may have been more difficult in terms of speech rate and speaking style than the concatenated sentences. Again, neither simple nor interaction effects (with the variable of interest, i.e., coherence) were found for the predictors session number and block order suggesting that the randomization procedures were appropriate and that there was no semantic priming from the first to the second session. The control variable block order is not included in the model below for reasons of brevity ($\chi^2(5, n=227)=2.62, p>.1$).

We also investigated whether the coherence effect can be generalized to different conversational speech fragments by replacing the conversational ANL values in the analysis above (CONV4) by the average ANL over the four conversational speech materials (CONV1 - CONV4) per participant (for the first session only). The results of this alternative analysis did not replicate the previous finding of a coherence effect on ANL ($\chi^2(1, n=113)=1.41, p>.1$). Thus, there is no clear evidence for a coherence effect on ANL in our data. We raised the possibility that speech rate may affect ANL outcomes and that the difference between the conversational and concatenated sentences material is not just about discourse coherence, but also about speech rate. To follow up on that, we tested whether speech rate differences between the four conversational fragments affected ANL outcome by setting up a linear mixed-effect model with speech rate as a continuous predictor of ANL (first session measurements only, only conversational fragments). Speech rate turned out not to be a significant predictor of ANL in this subset analysis ($\chi^2(1, n=228)=0.33, p>.1$).

Table 3.4: Model testing for the effect of meaningfulness on ANL (model 1A).

	β	*SE*
Intercept	4.79	0.62
Meaningfulness	1.46	0.44^{***}
Session number	-0.32	0.34
Meaningfulness × session number	-0.09	0.59

Significance level notation: $^{***}p<.001$.

Table 3.5: Model testing for the effect of semantic coherence on ANL (model 1B).

	β	*SE*
Intercept	4.25	0.72
Coherence	1.05	0.46^{*}
Session number	-0.12	0.43
Coherence × session number	-0.37	0.60

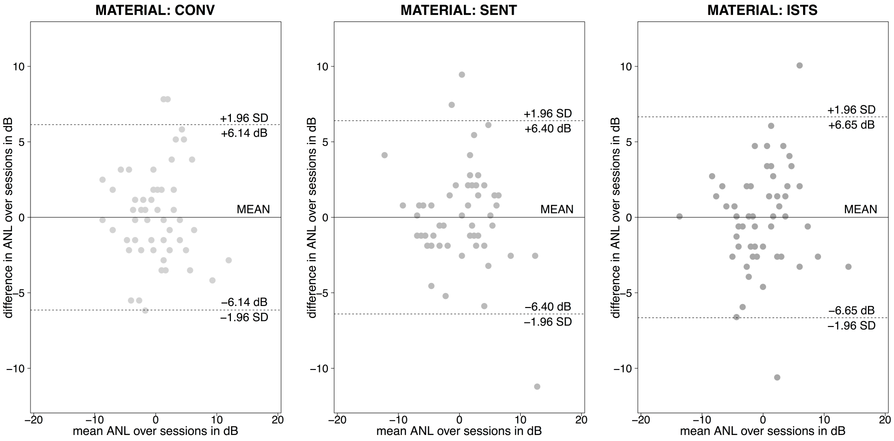
Significance level notation: $^{*}p<.05$.

Research Question 2: Does ANL repeatability differ across speech material types?

The mixed-model analysis did not show a significant speech material effect on repeatability of the ANL, quantified as the difference between the ANLs per participant for the two test sessions ($\chi^2(2, n=169)=0.57, p>.1$). In an additional analysis on repeatability across material types we used the statistical approach of the coefficient of repeatability (CR). Figure 3.3 displays the Bland-Altman plots for the three materials for which two test sessions had been run.

The highest coefficient of repeatability and thus the lowest repeatability was found for the ISTS material (CR=± 6.65 dB). Both the concatenated sentences material (SENT) as well as the conversational material showed lower coefficients of repeatability and thus numerically slightly better repeatability. For the concatenated sentences material (SENT) the CR was ± 6.40 dB. The best repeatability (numerically)

Figure 3.3: Bland-Altman plots for repeated ANL tests using conversational (CONV), concatenated sentence (SENT) and ISTS material. Horizontal lines represent the mean of the differences over the two test sessions as well as the boundaries for the 95% confidence interval per material type.



was found for the conversational test material with a CR of ± 6.14 dB. The combination of these two analyses suggests comparable repeatability across the speech materials.

In an additional step we calculated the coefficients of repeatability for all test materials over subsequent repetitions within test sessions. Table 3.6 shows that ANL repeatability increased numerically (i.e., CRs decreased) within test session I for all test materials except for CONV3. The same pattern of improved repeatability is seen for the CRs within test session II except for the SENT material. Overall, the repeatability in test session II does not seem to be numerically different from the repeatability in test session I. Note that repeatability seems to be most stable for the CONV4 material both within and across test sessions.

Research Question 3: Are ANLs differentially associated with self-report measures of listening effort and of hearing-related activity limitations for the different speech materials?

We first tested whether the first subscale of the SSQ self-report questionnaire (‘Speech hearing’) would be associated with ANL outcomes. The model showed significant material effects ($\chi^2(2, n{=}341){=}21.39$, $p<.001$) with highest ANLs found for the ISTS material and lowest ANLs for the sentence material (SENT). Importantly, this

Table 3.6: Coefficients of repeatability (in dB) for ANL for the six speech materials and the two test sessions contrasting subsequent repetitions.

Test material	Test session I		Test session II	
	Repetition 1 vs. 2	Repetition 2 vs. 3	Repetition 1 vs. 2	Repetition 2 vs. 3
CONV1	6.04	4.42	–	–
CONV2	6.87	5.29	–	–
CONV3	5.76	6.34	–	–
CONV4	4.98	4.75	5.50	5.07
SENT	6.38	4.65	4.32	6.06
ISTS	6.76	4.68	6.16	5.76

model showed a significant effect of the subjective questionnaire predictor SSQ (subscale ‘Speech hearing’) on ANL ($\chi^2(1, n{=}341){=}4.62, p{<}.05$, see Table 3.7). Higher scores on the SSQ subscale (i.e., fewer self-reported limitations due to hearing problems) were associated with more noise acceptance and thus lower ANLs. For an increase of 1 point on the SSQ ‘Speech hearing’ subscale the model predicted an ANL decrease of approximately 1 dB, which corresponds to an overall effect size of 4.4 dB (with the SSQ 'Speech hearing' subscale ranging from 4.86 to 9.36). However, the model did not show differential SSQ subscale effects on ANL for the three materials ($\chi^2(2, n{=}341){=}0.74, p{>}.1$).

We also investigated the association between the third subscale of the SSQ self-report questionnaire (‘Qualities of hearing’) and ANL. The model showed significant material effects with lowest ANLs for the sentence material ($\chi^2(2, n{=}341){=}21.31$, $p{<}.001$). However, we did not find an association between ANL and the third subscale of the SSQ self-report ($\chi^2(1, n{=}341){=}0.43, p{>}.1$), nor differential SSQ ‘Qualities of hearing’ effects on ANL for the three materials ($\chi^2(2, n{=}341){=}1.56, p{>}.1$).

In a third step we analyzed the association between the factor ‘Effort and concentration’ (questions number 15 and 18 of the ‘Qualities of hearing’ subscale of the SSQ) and ANL. As for the analyses above, the model showed significant material effects with lowest ANLs for the sentence material ($\chi^2(2, n{=}341){=}21.32$, $p{<}.001$). Yet, neither an association of ANL with the factor ‘Effort and concentration' ($\chi^2(1, n{=}341){=}1.80, p{>}.1$) nor differential 'Effort and concentration' effects on ANL for the three materials were found ($\chi^2(2, n{=}341){=}1.30, p{>}.1$).

Additionally, we explored the strength of the association between the SSQ self-report measures (subscale 'Speech hearing') and the ANLs (pooled over sessions) separately for the three materials by running correlation analyses. Only for the conversational material (CONV) a marginally significant correlation (r =-.23, p=.082, Pearson' r) was found.

Table 3.7: Model testing for differential associations between SSQ subscale scores and ANLs for three speech materials (CONV, SENT, ISTS).

	β	*SE*
Intercept (CONV material)	12.14	3.65
SENT material	-2.73	2.36
ISTS material	0.97	2.39
Session number	-0.98	0.52*
SSQ Part 1 ('Speech hearing')	-0.34	0.31
SSQ ('Speech hearing') × SENT material	0.26	0.33
SSQ ('Speech hearing') × ISTS material	0.003	0.33

Significance level notation: *p<.05.

Research Question 4: Do participant characteristics such as working memory (4A), and age, hearing thresholds, speech perception in noise, and self-control abilities predict ANL (4B)?

Again, ANLs were pooled over the two test sessions for each of the three materials. Working memory was not correlated with ANL ($p{>}.1$). Likewise, none of the other correlations (n=15) were statistically significant at an alpha level of .05 (i.e., not even before application of any correction required for multiple testing). Similarly, adding participant characteristics as continuous variables to either of the linear mixed-effect models discussed above (for research questions 1A and 1B) did not yield any significant effects of these participant-related variables.

3.5 Discussion

The clinical purpose of the ANL test is to predict self-reported hearing problems and future hearing aid success as reliably as possible. Therefore, it is crucial to know whether and how its clinical applicability depends on what speech material listeners are presented with and how the test is administered. Material effects on the outcome of the ANL test have been addressed in numerous studies (von Hapsburg and Bahng, 2006; Gordon-Hickey and Moore, 2008; Olsen et al., 2012a,b; Ho et al., 2012; Olsen and Brännström, 2014). In a number of recent publications (Brännström et al., 2012; Brännström et al., 2014a,b; Olsen et al., 2012a,b) – the ISTS (Holube et al., 2010) has been used, which is non-meaningful by definition. However, the original ANL test fragment used by Nábělek et al. (2006), in which ANL outcome was shown to be predictive of hearing aid uptake, was a meaningful and coherent read story, and thus linguistically different from the ISTS material. With the present study we investigated material effects on ANL to find out whether meaningfulness and coherence affect ANL (RQ1). In addition, we evaluated the repeatability of the ANL test across a range of test materials to check whether ecologically more valid materials yield a comparable repeatability as more standard audiology materials and the ISTS signal (RQ2). Further, we analyzed the association between ANLs and the outcome of a questionnaire that measures activity limitations due to hearing problems to elaborate on the connection between listening effort and ANLs. We also re-examined the association of working memory and self-control abilities and ANLs (RQ4) found in previous studies (Brännström et al., 2012; Nichols and Gordon-Hickey, 2012).

As expected, ANLs were higher for the ISTS material in comparison with the meaningful materials. Our interpretation of this effect is that the available redundancy for the meaningful materials facilitated speech processing (via top-down processing) and thus led participants to choose higher levels of acceptable noise (i.e., lower ANLs) than for the non-meaningful material. The unintelligible ISTS signal might have led participants to still want to hear as much as possible (i.e., relying more heavily on bottom-up processing). Furthermore, contrasting conversational ANL test materials with a passage of concatenated standard audiology sentences, we have not found convincing evidence for a semantic coherence effect on ANL. Possibly, the faster and more casual speaking style in the conversational material made listening more difficult, but this speaking style effect may have been offset by greater semantic coherence in the conversation, providing a form of discourse redundancy. The data did not provide clear evidence for priming effects across tests sessions (but note that Table 3.6 shows that coefficients of repeatability were largest between the first and second measurement within test session I). All in all, these results provide some evidence that top-down processing plays a role in ANL performance.

An important question was whether repeatability differs across the three speech materials. Neither the statistical modeling approach nor the analysis of the coefficient of repeatability (CR) showed statistically differential repeatability. Rather, repeatability was comparable for the three speech material types with CR values ranging between ±6.14 dB for the conversational material and ±6.65 dB for the ISTS material. Crucially, a coefficient of repeatability lower or equal to ±6 dB ensures that measurement error is lower than the distance between the two thresholds used to categorize hearing aid users as either successful or unsuccessful (≤ 7 and >13 dB, cf. Nábělek et al., 2006). Across test sessions, all three speech material types yielded CRs just above the critical ±6 dB threshold. With respect to ANL repeatability within test sessions, the conversational material (CONV4) yielded most stable CRs with values below ±6 dB. Our interpretation of the relatively high CR values across sessions is that listeners' internal criteria for MCL and BNL may be somewhat variable over time, particularly if they are engaged in other activities in-between test and retest measurements. As suggested by Brännström et al. (2014b), noise acceptance while following speech may best be considered a range (Acceptable Noise Range), rather than a specific level (ANL). The relatively poor repeatability of ANL may raise concerns about the clinical value of the ANL as an indicator for hearing aid use and success. However, if the ANL is used to compare two hearing aid conditions within one session, within-session reliability seems to be sufficient. For example, the ANL has been used successfully to show the effect of a noise reduction algorithm (Mueller et al., 2006; Peeters et al., 2009; Dingemanse and Goedegebure, 2015). Further research would be required to investigate whether Acceptable Noise Range may be a more reliable predictor of hearing problems and future hearing aid success than ANL.

Our analysis on the association of ANLs and the outcome of a subjective hearing-related questionnaire (RQ3) relates to recent discussion about the clinical meaning of concepts such as listening effort and fatigue in hearing-impaired individuals (McGarrigle et al., 2014). Our data showed a significant effect of participants' score on the subscale 'Speech hearing' of the Speech, Spatial, and Qualities of Hearing self-report (SSQ, Gatehouse and Noble, 2004) on ANL, particularly when listening to conversational speech. Participants who reported fewer listening problems also tolerated more noise while listening to speech (i.e., lower ANLs). Most questions of the 'Speech hearing' subscale are about conversation in noise. Both measurements (SSQ and ANL) are subjective judgments, where SRT measurements are not. This makes an association between ANL and SSQ more likely than an association between SRT and SSQ. The subscale 'Qualities of Hearing' was not significantly correlated with ANL. The between-participant differences of the 'quality of sound rating' were relatively small in this group of nearly normal-hearing participants. Possibly, perceived sound quality and ANL may be associated among hearing-impaired participants. No association was found between ANL and the subscale ‘Effort and Concentration’ This suggests that noise tolerance (as one aspect of listening comfort), is a different concept than the listening effort concept as formulated in these specific questionnaire questions. Further research should clarify differences and commonalities of both concepts.

The association between self-reported listening difficulties in noise and noise acceptance (i.e., ANL) only becomes evident when such an ANL test relates to everyday experiences. We think this result clearly makes a case for the use of ecologically valid conversational materials in clinical testing. Audiologists and speech researchers should think about how representative the type of noise and noise levels are of everyday listening, but they should also care about differences between read aloud speech and spontaneous conversation.

Further, the attempt to replicate working memory effects on ANL was unsuccessful. This suggests that noise tolerance, as one aspect of listening comfort, is not related to individual working memory capacity. Importantly, in line with previous studies (cf. Akeroyd, 2008), working memory was considerably correlated with speech perception in noise (cf. Table 3.2), with higher working memory relating to better speech perception. The failure to replicate working memory effects on ANL in our study can be accounted for in two ways. First, it may be due to the use of different test materials and test procedures to quantify working memory. The test that Brännström et al. (2012) used to quantify working memory was an auditory version of the reading span task in which the examiner presented the sentences orally, which may have increased the contribution of hearing. Alternatively, the lack of a correlation between ANL and working memory can be taken to underline that ANL and speech perception in noise are different in nature. The latter account ties in with our observation that ANLs did not relate to age, hearing thresholds, and speech-in-noise perception abilities. This held in the relatively good-hearing adult sample as tested here, but was also found by Nábělek et al., 1991, 2004; Freyaldenhoven et al., 2007; Plyler et al., 2007; and Moore et al., 2011 for both normal-hearing and hearing-impaired participants. Moreover, we have not found evidence for an association between ANL and self-control abilities reported in Nichols and Gordon-Hickey (2012). However, the latter study used a self-control scale containing 36 items in contrast to the Brief Self-Control Scale with 13 items that we asked our participant to fill in.

The combined pattern of results converges on material effects being present for the ANL test with better noise tolerance and slightly better and more stable repeatability, at least numerically, for meaningful stimuli. We have also shown that activity limitations due to hearing problems and ANLs are related, especially if conversational materials are used as ANL test material. More natural speech materials can thus be used in a clinical setting as repeatability is not reduced compared to more standard materials. We aim to conduct follow-up research to investigate whether ecologically valid test materials – such as the conversational speech material used in this study – can be used to improve the predictive power of the ANL test for hearing aid success, relative to more standardized speech materials.

Effects of age and hearing loss on articulatory precision for sibilants

Chapter 4

This chapter is based on:
Xaver Koch and Esther Janse (2015)
Effects of age and hearing loss on articulatory precision for sibilants
In *Proceedings of the 18th International Congress of Phonetic Sciences*

This study investigates the effects of adult age and speaker abilities on articulatory precision for sibilant productions. Normal-hearing young adults with better sibilant discrimination have been shown to produce greater spectral sibilant contrasts. As reduced auditory feedback may gradually impact on feedforward commands, we investigate whether articulatory precision as indexed by spectral mean for [s] and [ʃ] decreases with age, and more particularly with age-related hearing loss. Younger, middle-aged and older adults read aloud words starting with the sibilants [s] or [ʃ]. Possible effects of cognitive, perceptual, linguistic and sociolinguistic background variables on the sibilants' acoustics were also investigated. Sibilant contrasts were less pronounced for male than female speakers. Most importantly, for the fricative [s], the spectral mean was modulated by individual high-frequency hearing loss, but not age. These results underscore that even mild hearing loss already affects articulatory precision.

4.1 Introduction

Adult aging may lead to several changes in speech, such as spectral modifications, altered voice characteristics and decreased speech rate (Torre and Barlow, 2009). Generally, imprecise articulation is a prominent perceptual feature of older adults' speech (Hartman and Danhauer, 1976). Auditory sensory decline has been argued to be a possible cause of the changes in their speech (Cruickshanks et al., 1998), besides cognitive and anatomical changes.

The finding that post-lingually deafened adults produce less pronounced consonant contrasts than healthy controls (Matthies et al., 1994) emphasizes the role of auditory feedback for precise articulation. Further important evidence that perceptual differences are linked to production differences comes from Perkell and colleagues (2004), showing that participants with good sibilant discrimination abilities also produce greater spectral sibilant contrasts.

Thus far, investigations of age effects on speech production have mainly focussed on speech rate, fundamental frequency, vowel formant values, and voice onset time. However, sibilant fricatives, due to their spectral prominence in high-frequency ranges, can be expected to be the first to be impacted by age-related high-frequency hearing loss. Furthermore, sibilants are acquired relatively late in child language development and are often affected by speech disorders such as dysarthria or apraxia of speech.

We assume that the combination of the sibilants' complex articulatory movements and their dependence on precise auditory feedback relate to their vulnerability to disorders and to their relatively late acquisition age. A negative relation between sequence in language development and language decline has been shown for language impairments in dementia of the Alzheimer type (Emery, 2000), a neurodegenerative disease which has been linked to aging. Even healthy aging may be accompanied by reduced motor control that would be apparent particularly for sounds that are relatively difficult to produce, and that require high-frequency auditory feedback information.

The present study therefore investigates whether and how sibilant production may change across the adult life span (Research Question 1). Additionally, we investigate which individual cognitive, perceptual, linguistic and sociolinguistic speaker characteristics predict articulation precision (Research Question 2).

4.2 Speakers

Three age groups were included (107 participants in total): 38 older adults (M^{age}=67.1 yrs., *SD*=4.7, 22 female), 34 middle-aged adults (M^{age}=49.9, *SD*=7.6, 21 female), and 35 younger adults (M^{age}=21.4, *SD*=2.6, 22 female). None of the participants wore hearing aids although six of them (one middle-aged and five older adults) met the Dutch hearing-aid criterion (pure-tone average over 1, 2 and 4 kHz ≥ 35 dB HL in either ear).

The speakers were sampled from a participant pool. All lived in the Nijmegen area, but came from different Dutch regions. Participants were asked to fill out a questionnaire on their language background and regional dialect. Participants were also asked to specify whether they spoke a Dutch dialect in everyday life or not (regionality self-rating).

4.3 Procedure

4.3.1 Materials and speech recordings

Participants read ten monosyllabic target words (nine nouns, one adjective) embedded in a carrier phrase ("Ik zei __ tegen hem", 'I said __ to him'). The two target sounds [s, ʃ] appeared in five vocalic contexts (cf. Table 4.1).

Table 4.1: Target words.

	[s]			[ʃ]	
Saab	[saːp]	*car brand*	sjaal	[ʃaːl]	'scarf'
set	[sɛt]	'set'	chef	[ʃɛf]	'boss'
Sieb	[sip]	*name*	chic	[ʃik]	'modish'
sop	[sɔp]	'soap'	shop	[ʃɔp]	'shop'
soep	[sup]	'soup'	Sjoerd	[ʃuʁt]	*name*

Each target word was repeated five times. All stimulus pairs were near minimal pairs with the exception of one true minimal pair ("sop" vs. "shop"). Recordings were made in a sound-attenuated booth using a Samson QV head-set microphone and an Edirol R09 recorder (44.1 kHz sampling frequency, 16 bit resolution). Fifty filler sentences (ten nouns without sibilants in word-initial position, each repeated five times) were interspersed with the target sentences on a single pseudorandomized list. Sentences from this list were presented to participants one by one on a computer screen in a self-paced manner.

4.3.2 Speaker abilities

Whereas there is a wealth of studies on individual predictors for speech comprehension, very little research has looked into relationships between speaker abilities and speech output (Haley et al., 2010). We explore whether auditory, cognitive and linguistic abilities are associated with articulatory precision. The following five tests were administered:

1. *Pure-tone audiometry* to index hearing thresholds: → hearing level in decibel
2. *Digit Symbol Substitution Test* (Wechsler, 2004) performance to index processing speed: → number of correctly recoded symbols (within 2 min., out of 133)
3. *Vocabulary subpart of the Groningen Intelligence Test* (Luteijn and van der Ploeg, 1983) to index linguistic ability: → number of correct synonym answers (out of 20)
4. *Digit Span Test* (Wechsler, 2004) with backward recall to index working memory (visually administered): → percentage of correctly recalled items (12 items)
5. *Raven's Standard Progressive Matrices Test* (Raven et al., 2003) to index general non-verbal intelligence: → number of correct items (in 10 min., out of 60).

Table 4.2: Means and standard deviations of speaker abilities per age group.

Speaker abilities	Young adults *M (SD)*	Middle-aged adults *M (SD)*	Older adults *M (SD)*
2 kHz hearing threshold	7.14 (5.72)	10.29 (7.97)	20.39 (10.03)
4 kHz hearing threshold	4.43 (7.35)	17.65 (12.20)	29.74 (15.68)
6 kHz hearing threshold	10.57 (9.38)	20.88 (15.50)	34.89 (17.99)
8 kHz hearing threshold	6.43 (8.54)	25.15 (13.84)	45.53 (18.19)
Processing speed	87.26 (13.46)	76.15 (15.21)	64.05 (13.42)
Linguistic ability	13.83 (2.04)	15.76 (1.56)	16.68 (2.00)
Working memory	55.95 (18.81)	62.50 (24.38)	50.00 (18.78)
Non-verbal intelligence	44.54 (5.60)	38.82 (6.00)	31.58 (8.06)

4.3.3 Processing of acoustic data

The recorded sentences (107 speakers × 5 repetitions × 10 sentences = 5350) were pre-annotated using the automatic speech recognition plugin `praatalign` (Lubbers and Torreira, 2016). Sibilant and target word boundaries were checked and corrected (if necessary) manually for all productions but the forced alignment procedure was quite accurate (maximally 10-15 ms deviation from hand-annotated boundaries). Target word durations were extracted to model potential speech rate effects on articulatory precision.

We derived spectral density estimates from the sibilant signals (cf. Reidy, 2015; Forrest et al., 1988). The middle third of each respective sibilant section was chosen as analysis interval to minimize coarticulation effects on the measurements. All analyses were done using `Praat` (Boersma and Weenink, 2016). A pre-emphasis of 6 dB/octave for the frequencies above 80 Hz was applied to the analysis intervals. The resulting spectra were cepstrally smoothed (500 Hz) and spectral moments (in Hz) were calculated. Only the values of the first spectral moment (Center of gravity: henceforth, COG) of the sibilant productions were analyzed.

4.4 Results

Statistical regression models were run using linear mixed-effect models in the program `R` with the `lme4` package (Bates et al., 2014) for the dependent variable spectral mean (COG). We started from a model containing interactions between sibilant identity ([s] vs. [ʃ]) and all experiment related control variables (vocalic context (± round), repetition number (1:5), trial position in experiment (1:100), speech rate in syllables per second) and gradually simplified this model using a backwards stepwise model selection approach. The modeling procedure was based on likelihood ratio tests to evaluate which interactions (sibilant identity × control variable), or control variables could be taken out without significant loss of model fit. The optimal random effect structure consisted of random intercepts for participants and items as well as random slopes for speech rate (correlated with random intercept for participants) and sibilant identity (no correlation with random intercept for participants).

Firstly, 22 target phrase productions with hesitations or slips of the tongue were excluded from the analyses. Secondly, spectral mean values above 10 kHz were excluded. Subsequently, outliers were removed separately for [s] and [ʃ] productions: COG values higher than 2.5 *SD*s above the respective means were excluded. Analyses are based on a dataset containing 2656 [s] and 2550 [ʃ] productions.

The resulting basic model (not reported here in detail as all effects are replicated in later models) showed significant effects of sibilant identity, vocalic context and trial position plus an interaction of sibilant identity × vocalic context (± round):

i. *sibilant identity effect:* higher spectral mean values for [s] compared to [ʃ] productions

ii. *vocalic context effect:* lower spectral mean values for the sibilants in +round (+back) vocalic context compared to −round context

iii. *trial position effect:* higher spectral means for trials later in the experiment

iv. *sibilant × vocalic context interaction effect:* stronger anticipatory coarticulation effects for [s] sibilants than for [ʃ] sibilants.

Table 4.3: Model testing for age and gender effects in sibilant productions.

Fixed effects	β	*SE*	*p*<
Intercept	7125.26	105.11	
Sibilant identity	-1569.16	107.40	$.001^{***}$
Vocalic context	-950.59	65.51	$.001^{***}$
Trial position	0.66	0.27	$.016^{*}$
Gender	-812.00	152.31	$.004^{**}$
Gender × voc. context	240.19	39.85	$.038^{*}$
Sibilant identity × gender	716.51	145.39	$.001^{***}$
Sibilant identity× voc. context	708.07	96.76	$.001^{***}$
Sibilant identity × voc. context × gender	-361.55	55.36	$.001^{***}$

Reference levels: sibilant identity: [s], vocalic context:−round, gender: female; *P*-values were calculated using the `Anova` function of the `car` package (Type II Wald χ^2 test). Significance level notation: ***p<.001, **p<.01, *p<.05.

Modeling age and gender effects for sibilant productions

To test for basic speaker information effects on articulatory precision (Research Question 1), we added chronological age and gender of the speakers in one step to the basic model described above (simple effects, interactions with control variables and sibilant identity). Gender was included as a control variable as sibilant productions are known to differ between male and female speakers (Fuchs and Toda, 2010; Stuart-Smith, 2007). The most parsimonious model resulting from adding age and gender effects is presented in Table 4.3.

Age did not affect the sibilants' spectral mean, nor did it interact with any of the other predictor variables. However, adding participants' gender to the sibilant production model significantly improved the data fitting. In line with earlier studies

(Fuchs and Toda, 2010; Stuart-Smith, 2007), male speakers showed lower spectral means for [s] productions than female speakers. Consequently, the acoustic contrast between [s] and [ʃ] productions is smaller for male than for female participants (cf. the sibilant identity × gender interaction). Men also show smaller coarticulation effects for [s] productions in the +round vocalic context compared to female speakers.

Modeling speaker ability effects on sibilant production

To investigate the role of speaker characteristics on sibilant production beyond age and gender effects, we carried out a third series of model comparisons. We did not model age and effects of speaker abilities simultaneously because hearing, processing speed, non-verbal intelligence and vocabulary size were all considerably correlated with age (Spearman's rank-order correlation tests: $|r|>.50$, $p<.001$). Thus, all background variables were added to the previous model (excl. age), as well as their interactions with sibilant identity. The resulting model (not shown here) showed that participants who categorized themselves as dialect speakers produced significantly lower spectral means than non-dialect participants (β=-305.44, SE=121.96, $p<.05$). However, the absence of an interaction between regionality self-rating with sibilant identity implies that dialect speakers do not show reduced acoustic sibilant contrasts but rather shift both sibilants' acoustic spaces to lower frequencies.

Interactions with sibilant identity were found for the two continuous predictors processing speed and 8 kHz hearing loss, whereas no simple effects were observed for these predictors. On the basis of these two interactions and after visual inspection of the relationship between hearing and the two sibilants' spectral means, we decided to run separate analyses for each sibilant to further investigate effects of speaker abilities on production of the sibilants.

Our analysis of the [ʃ] production data did not substantiate the effects of processing speed, hearing loss or any other speaker ability measure on spectral mean. For the [s] productions only high-frequency hearing loss was a predictor of spectral mean frequency (and not processing speed or any other speaker ability measure): The higher participants' 8 kHz hearing threshold, the lower their spectral mean for [s]. This effect of high-frequency hearing loss (≈ 8 Hz decrease in spectral mean per loss of 1 dB HL at 8 kHz) on the spectral properties of the [s] productions is illustrated in Figure 4.1. Table 4.4 shows the most parsimonious model on speaker abilities for the [s] productions (random effects structure: random intercepts for participants and items as well as a random slope for speech rate which was correlated with random intercept for participants). To rule out that the hearing loss effect was solely due to the six participants who met the Dutch hearing-aid criterion, we also ran the above model on a dataset excluding these speakers. The effect of hearing acuity at 8 kHz on the spectral mean (COG) of the [s] productions was replicated in this subset.

Figure 4.1: Effect of high-frequency hearing loss on the spectral mean (COG) of [s] productions. Experimental data (*n*=2656) and model prediction shown, mean COG value for [ʃ] productions as reference.

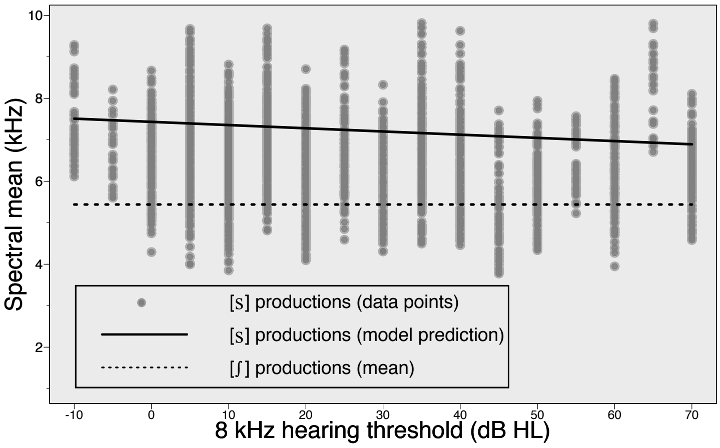


Table 4.4: Model testing for speaker ability effects on [s] productions.

Fixed effects	β	*SE*	$p<$
Intercept	7509.06	134.49	
Gender	-834.02	144.00	.001***
Vocalic context	-961.62	57.10	.001***
Regionality self-rating	-373.57	146.57	.011*
8 kHz hearing threshold	-7.72	3.24	.018*
Gender × voc. context	242.38	43.86	.001***

Reference levels: vocalic context: −round, gender: female, regionality self-rating: no dialect use; *P*-values were calculated using the `Anova` function of the `car` package (Type II Wald χ^2 test).

Significance level notation: ****p*<.001, **p*<.05.

4.5 Discussion

Numerous studies have investigated sibilant production addressing different questions (e.g., on speech production modeling (Perkell et al., 2004), or on sociophonetic variation (Fuchs and Toda, 2010; Stuart-Smith, 2007)). The present study was set up to investigate if changes over the adult life span influence articulation precision and to evaluate effects of individual speaker abilities on sibilant articulation. A standard sentence production paradigm was employed to elicit word-initial sibilant productions [s, ʃ] from a large sample of participants ($n > 100$), ranging in age between 18 and 78 years.

First, effects of vocalic context and speaker gender as found in other studies were replicated here, but the hypothesized age effect on sibilant articulation was not found (Research Question 1). Moreover, our data showed a gender by sibilant interaction effect, suggesting that the sibilant contrast was more pronounced for female than male speakers. Concerning our second research question on effects of speaker abilities on sibilant articulation, we found that high-frequency hearing loss modulated [s] productions. Thus, the sharpness of a speaker's [s] relates to the speaker's hearing acuity. Individual hearing acuity influences the auditory (feedback) information available from hearing one's own speech and from hearing other speakers. As we cannot be certain that the observed hearing acuity differences among speakers of our sample were actually acquired at an older age, our data indicate that high-fidelity auditory feedback is needed to either acquire or maintain precise articulation. Earlier research had shown effects of profound hearing loss on sibilant production. Our results, however, indicate that even mild (high-frequency) hearing loss modulates target production, particularly for targets with their distinct information in high-frequency spectral regions.

Changes in speech acoustics after cochlear implantation in post-lingually deafened adults

Chapter 5

This chapter is based on:
Xaver Koch, Thordis Neger, Wendy Huinck,
Emmanuel Mylanus, and Esther Janse (in preparation)
Changes in speech acoustics after cochlear implantation in post-lingually deafened adults

Parts of this work have been presented July 2017 at
the 7th International Conference on Speech Motor Control, in Groningen, The Netherlands
In *Stem-, Spraak-, Taalpathologie* 22, page 57

A prolonged deprivation of auditory feedback may negatively affect speech production accuracy, as shown by reports of reduced vowel space, changed formant values and reduced sound contrasts for challenging sounds such as sibilants in post-lingually deafened adults, compared to normal-hearing controls. Post-lingually deafened adults who undergo CI have been hypothesized to enter a phase of retuning of their feed-forward motor commands to account for the restored, altered auditory feedback. The present study investigates relatively early changes (two weeks to three months after cochlear implantation) in speech acoustics due to cochlear implantation. Post-lingually deafened novice CI users produced target words that contained the sibilant sounds [s] and [ʃ] and five vowels at three test sessions: before and directly after implantation, and approximately three months after they got their CI. Their acoustic realizations were compared to those of an age- and gender-matched control group (also tested three times). Spectral means, vowel formants and the average vowel space were analyzed. Results show that, prior to CI surgery, sibilant contrasts but not average vowel space were diminished for the deafened patients in comparison to the control group. Across test sessions, both groups enlarged the sibilant contrast, yet in different ways. The CI users enlarged the sibilant contrast mainly by lowering the spectral mean of the post-alveolar sound [ʃ], whereas the controls increased spectral means of the [s] productions as well as decreased spectral means for [ʃ] productions. Furthermore, duration of hearing loss predicted the amount of change in sibilant contrast after CI activation. The absence of auditory feedback thus impairs phonemic contrasts for sibilants over time in deafened participants. Once auditory feedback is available again sibilant contrasts improve immediately, and especially for participants with shorter duration of deafness.

5.1 Introduction

With approximately 50.000 newly implanted cochlear implants (henceforth, CIs) each year bionic hearing becomes an increasingly relevant research topic. CI recipients are presented with a novel, degraded and artificial signal delivered by the device's speech processor directly to their auditory nerve. The information that is processed by thousands of inner and outer hair cells in normal hearing is reduced to the stimulation of the auditory nerve via a limited number of channels (electrodes) resulting in unnaturally broad bandwidths and in frequency shifts due to the positioning of the electrode array in the cochlea. Unsurprisingly, given the nature of the dysfunction, the majority of CI research focuses on improvement in speech perception. Less is known about changes in speech production following cochlear implantation.

This study investigates effects of cochlear implantation on speech acoustics in post-lingually deafened adults. In contrast to pre-lingually deafened children, post-lingually deafened adults have had the opportunity to develop language normally and to establish a speech motor control system based on auditory feedback. Speech production research in CI users complements research on their speech perception for two reasons. The first practical reason is that CI users' speech intelligibility is important as it relates to their communicative success: CI users have to be understood by their interlocutors. Even though speech production, once acquired, is relatively robust and most post-lingually deafened adults maintain good intelligibility (e.g., Goehl and Kaufman, 1984; Lane and Tranel, 1971), deafened adults may produce noticeably 'deviant' speech. The second, theoretical, reason is that speech production research with CI recipients can elucidate the effects of long-term and short-term auditory changes on the speech production system and on the representation of speech sounds in general.

Normal speech production, as modeled for instance in the Directions Into Velocity of Articulators model (henceforth, DIVA; e.g., Tourville and Guenther, 2011) of speech motor control, involves stable feed-forward and feedback control systems to monitor and correct speech production. Articulatory movement is controlled by somatosensory feedback (i.e., the internal information about where e.g., one's tongue tip and one's jaw is, e.g., Ghosh et al., 2010), and auditory feedback (e.g., Perkell et al., 2004). Hearing loss, or the loss of auditory feedback, may gradually lead to altered feed-forward commands because the somatosensory feedback alone may not provide sufficient information to keep the feed-forward commands stable and precise. According to Lindblom's theory of Hyper and Hypo-articulation (Lindblom, 1990) speakers can generally choose to produce speech on a continuum from hyper-to-hypospeech by giving more or less priority to clarity over economy of effort. An example of hyperspeech would be the clear and careful speaking style while reading out a speech on a formal occasion. Hypospeech, which would be used in a colloquial conversation, is characterized by higher or more variable speech rates, reduced average vowel spaces and by more frequent sound reduction and deletions. For patients with severe hearing loss the principle of economy of effort may, over time, lead to a bias towards hypospeech (e.g., centralized vowels, less precise consonants) and may gradually change the feed-forward motor commands stored in long-term memory because corrective auditory feedback is poor or absent. In fact, a number of studies indicate that post-lingually deafened adults may show deviant consonant and vowel productions and decreased sound contrasts relative to a normal-hearing sample (e.g., Waldstein, 1990; Schenk et al., 2003; Lane and Wozniak Webster, 1991; Lane et al., 2007; but cf. Goehl and Kaufman, 1984). Sibilant fricatives can be expected to be the first to be impacted by hearing loss because their articulation involves complex articulatory movements, the fine-tuning of which depends on precise auditory feedback. For this reason, sibilants are acquired relatively late and are often affected in speech disorders. Some evidence that even mild forms of age-related hearing loss may affect sibilant production was found in a correlational study on a sample of older adults in which individual high-frequency hearing loss was found to predict acoustic realization of sibilants (Koch and Janse, 2015). Below, we will first elaborate on earlier findings concerning acoustic differences between speech of post-lingually deafened adults and controls, and then on the effect of cochlear implantation on speech production. Our literature overview will provide information on whether the studies included age- and gender-matched control groups.

Waldstein (1990) compared speech of post-lingually deafened speakers to that of an age- and gender-matched control group, focusing on differences in vowel formants, voice onset time and sound durations. Waldstein showed that both the vowel formants as well as the sound durations were more variable in deafened speakers compared to controls. Further, Waldstein found shorter VOTs for English voiceless stops in deafened speakers, resulting in acoustically less distinct consonant categories. Schenk et al. (2003) focused on German vowel acoustics, and observed vowel-specific changes in the first vowel formant and reduced vowel spaces for post-lingually deafened speakers compared to a control group (age-matched, not gender-matched). For English consonant production, Lane and Wozniak Webster (1991) observed less differentiated fricative and stop productions in three post-lingually deafened participants compared to an age- and gender-matched control group. Matthies et al. (1994) characterized three out of their sample of five post-lingually deafened participants as having poor acoustic sibilant categories, without presenting data for a control group. Note also that some previous studies included participants with meningitis as cause of the hearing loss (e.g., Schenk et al., 2003; Waldstein, 1990; Matthies et al., 1994). As pointed out by Sapir and Canter (1991), diseases such as meningitis not only lead to hearing loss but can also lead to neurological problems, such as motor- and speech problems. In sum, there is some indication, but little compelling evidence, due to methodological issues, that post-lingually acquired hearing loss affects speech acoustics.

So if we have some indication that speech is affected by prolonged hearing loss in post-lingually deafened adults, what do we know about the effect of cochlear implantation on speech? Several studies have investigated effects of cochlear implantation on segmental aspects of speech production (Perkell et al., 1992, 2000, 2005, 2007; Lane et al., 2001; Langereis et al., 1997, 1998, 1999; Svirsky et al., 1992; Matthies et al., 1994, 1996, 2008; Gould et al., 2001). CI patients showed increased vowel space following cochlear implantation (Lane et al., 2001), as well as greater acoustic contrasts for sibilants (as quantified with the Center of Gravity measure; Lane et al., 2007; Matthies et al., 1994). Importantly, some studies have argued that general changes in speech rate, fundamental frequency and sound pressure level following cochlear implantation might underlie acoustic changes after CI (Perkell et al., 1992, 2000). Indeed, profound hearing loss has been shown to affect temporal aspects of speech production, as evidenced by reduced articulation rates and increased pause frequency in deaf speakers in a number of studies (Lane and Wozniak Webster, 1991; Plant, 1983; Plant and Hammarberg, 1983; Leder and Spitzer, 1990). Even though results by Matthies et al. (1996) indicate that changes in segmental aspects of speech production after CI are not related to speech rate changes after CI, any changes in acoustic realization after cochlear implantation should obviously be considered in relation to potential changes in speech rate.

The investigation of effects of cochlear implantation on production obviously raises the question of how much time these potential changes in articulation require. Several studies have investigated short-term effects of the availability of auditory feedback by either simulating changed auditory feedback in normal-hearing participants (e.g., Casserley, 2015; Houde and Jordan, 1998), or by comparing speech production in CI on vs. CI off conditions in CI users (e.g., Perkell et al., 2007). Casserley's (2015) CI simulation lead to immediate vowel height shifts for normal-hearing participants, but not to acoustic changes for sibilants. Houde and Jordan (1998), who manipulated vowel formants online, demonstrated real-time compensation for the feedback alterations in normal-hearing participants. Perkell et al. (2007) investigated the effect of blocking or restoring auditory feedback via the CI. Their results did not indicate consistent effects of feedback availability on acoustic measures of vowel and sibilant contrasts. A number of other studies have investigated relatively short-term and longer-term changes in production following cochlear implantation (Langereis et al., 1997; Lane et al., 2007). A study by Langereis et al. (1997) investigated changes in vowel formant frequencies (F1, F2) in post-lingually deafened adults after cochlear implantation. Langereis and colleagues found vowel formant shifts toward normative values one year after implantation but not after three months, indicating that acoustics improvements for vowels take time to develop after CI. Lane et al. (2007) also investigated the time course of changes following CI activation. They argued that the new input via the CI leads to differential auditory cortex activity (compared to before the hearing loss) because of a change in frequency mappings. As such, active learning is necessary to utilize the new auditory space, to connect existing abstract sound representations to new auditory information, and to re-establish auditory goal regions for production. In contrast to the results of Langereis et al. (1997), Lane et al. (2007) observed that already after one month of CI use both vowel and sibilant contrasts improved. As their study did not include a post-activation session earlier than one month, it is unclear which changes in speech acoustics may have taken place in the initial weeks of CI use. Possibly, restructuring the auditory space may initially lead CI recipients to produce diminished sound contrasts, followed by recovery and improvement of the sound contrasts.

The present study therefore investigated the effect of CI on speech acoustics directly after CI activation, by measuring sound acoustics prior to surgery, approximately two weeks after CI, and at three months after CI activation. This way, we aim to investigate earlier and later changes in speech production following CI activation. In addition, we took into account that repeated production of the same materials across multiple sessions in the longitudinal design may affect acoustic realization, as previous research has shown that repeated mention of words may result in hypospeech, at least within a session (Lindblom, 1990; Baker and Bradlow, 2009). As we cannot rule out the possibility that speakers may remember the speech materials from earlier sessions, we aimed to disentangle possible repetition effects from effects of restored hearing in CI patients by including repeated test sessions for a control group as well.

Apart from the time that speakers need to alter their production after cochlear implantation, there is also a time aspect to their deafness prior to implantation. Previous studies have shown that the duration of hearing impairment is associated with speech perception success after CI (e.g., Blamey et al., 1992; Plant et al., 2016). This may be because a lack of auditory stimulation leads to a reduction of cells in the origin of the auditory nerve in the inner ear over time (cf. Nadol et al., 1989; but Blamey, 1997), such that speech perception can recover more easily with less cell loss. It is, less clear, however, whether duration of hearing loss also predicts speech production acoustics and improvements thereof after CI. The only study to date that investigated this link (Schenk et al., 2003) found no effect of the duration of hearing loss on vowel acoustics in deafened adults. However, it is unclear whether duration of hearing loss predicts sibilant articulation and potential changes in sibilant and vowel acoustics after CI.

This study was set up to investigate potential changes in sibilant and vowel acoustics after CI activation in post-lingually deafened adults. Restoring auditory feedback may affect sibilant and vowel production but it is not clear how much time such changes require. As mentioned above, we also took into account that repeated production of the same materials across multiple sessions may affect acoustic realization in both the patient group and a control group.

We summarize our two research questions below:

1. Do acoustic realizations of sibilants (1A) and vowels (1B) of adult post-lingually deafened candidates for cochlear implantation differ from those of an age- and gender-matched control sample?
2. How does cochlear implantation in severely hearing-impaired individuals affect sibilant (2A) and vowel acoustics (2B)? More particularly, what changes can be observed early after activation (i.e., approximately two weeks after CI activation), and after a period of three months after CI activation?

5.2 Study Design and Method

5.2.1 Participants

We tested nine post-lingually deafened native Dutch adults (M^{age}=54 yrs., 2 female) who underwent monaural cochlear implantation at the Radboud University Medical Centre ENT department (Nijmegen, The Netherlands). Participants had to be 18 years or older in order to be included in the study, with an onset of the deafness after 6 years of age. Participants with cognitive disabilities, non-normal (corrected) visual acuity, hearing loss due to meningitis or hearing loss due to syndromal deviations, or planned partial CI insertion, as well as participants who received special education were excluded from the study. The average duration of hearing loss was 23.2 years (*SD*=10.1 yrs.) with the earliest onset of hearing loss at 3 years of age (participant with the anonymized participant code OP, cf. Table 5.1 below) and latest at 49 years (participant EF). Pre-operation (aided) hearing status was assessed using pure-tone audiometry via air and bone conduction and CVC speech audiometry (see subsection 5.2.4 Background Tests). Speech intelligibility of the CI candidates, as rated subjectively by the third author during pre-operative counseling, was good for all participants, except that only for participant OP (see Table 5.1), subtle indications of speech characteristics of the deaf were reported (e.g., less accurate consonant productions). Table 5.1 shows demographic data as well as hearing-related characteristics of the CI group. An age- and gender-matched control group was included as a reference sample for the CI group (n=9, M^{age}=59 yrs., 3 female).

Table 5.1: Participant-related information (including CI device details) for the group of deafened adults that underwent cochlear implantation.

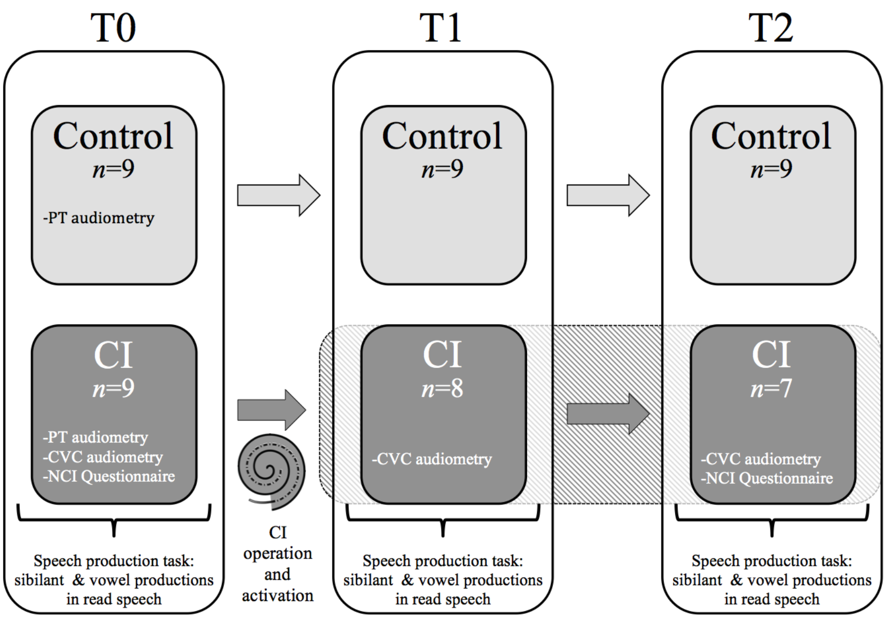
	Participants								
	AB	CD	EF	GH	IJ	KL	MN	OP	QR
Gender	male	male	male	male	male	male	male	female	female
Age	69	59	66	53	73	48	64	28	52
Hearing loss duration (yrs)	40	20	17	14	28	16	37	25	12
Hearing loss etiology	Otosclerosis	hereditary (DFNA9)	unknown	unknown	Otosclerosis	unknown/ hereditary	Otosclerosis	hereditary	unknown
CI model	Advanced Bionics	Cochlear Nucleus 422	Advanced Bionics	Cochlear Nucleus 522	Cochlear Nucleus 512	Advanced Bionics	Cochlear Nucleus 512	Cochlear Nucleus 422	Advanced Bionics
Stimulation Stragegy	HiRes Optima-S	ACE	HiRes Optima-S	ACE	ACE	HiRes Optima-S	ACE	ACE	HiRes Optima-S
Number of electrodes	16	22	16	22	22	16	22	22	16
CI stimulation frequency range (kHz)	0.25 – 8.7	0.19 – 7.9	0.25 – 8.7	0.19 – 7.9	0.19 – 7.9	0.25 – 8.7	0.19 – 7.9	0.19 – 7.9	0.25 – 8.7

5.2.2 Procedure

Both the CI group and the control group were tested three times (T0, T1, T2). This way we were able to disentangle mere repetition effects from the effects of restored auditory feedback (see Figure 5.1). The control group was re-tested twice after fixed intervals of four weeks. CI participants were tested once before they were operated on (T0) and twice after cochlear implantation (T1, T2).

- T0: pre-activation session in the clinic, mean time lag for session T0 to operation: 18 days, *SD*=13.5, median time lag T0 to activation of CI: 43 days[6]
- T1: first post-activation monitoring point in the clinic, on average 12 days after activation of the CI, mean time interval between T0 and T1 was 68 days, *SD*=46.5
- T2: third post-activation monitoring point in the clinic, on average 92 days after CI activation

Figure 5.1: Schematic of the procedure.



Three test recordings of the CI group could not be included in the analyses (one for T1 and two for T2) because of unexpected data loss (see Figure 5.1). Due to technical problems for one CI participant (with participant code QR) only three repetitions were recorded at measurement T0 instead of the five planned repetitions per target word. These recordings were included in the analyses.

---

[6] Participant CD's device was activated only 111 days after the operation, which contributed to this participants' extraordinarily long T0-T1 time lag of 153 days. The mean time lag between T0 and CI activation for the remaining sample (i.e., excluding participant CD) was 46 days (mean operation to CI-activation time lag: 31 days).

5.2.3 Stimulus material

Participants read ten monosyllabic target words (see Table 5.2) embedded in a carrier phrase (“Ik zei __ tegen hem”, ‘I said __ to him’) at all three test sessions (T0, T1, T2). The sibilant sounds [s, ʃ] appeared in five vocalic contexts [aː, ɛ, i, ɔ, u]. Each target word was repeated five times for the first test session (T0) and three times for the subsequent test sessions (T1, T2). All stimulus pairs were near-minimal pairs with the exception of one truly minimal pair (“sop” vs. “shop”). For each of the three test sessions 50 percent of the sentences were filler sentences (10 different fillers repeated multiple times) with the same carrier phrase structure but containing ‘target words’ without sibilant fricatives in word-initial position (e.g., “Ik zei *fiets* tegen hem”, ‘I said *bike* to him’). The target sentences were interspersed with the fillers on two pseudorandomized lists (T0: 50 test and 50 filler sentences, T1 and T2: 30 test and 30 filler sentences). The randomized lists for T0 and T1/T2 thus only differed in terms of number of repetitions of the 10 different target and 10 different filler sentences. Sentences from the lists were presented one by one in Arial font (36 pt) to participants in a self-paced manner centered on the screen of a 9.7-inch E-Reader. Participants were instructed to read the sentences at their habitual speed. Recordings were made using a Samson QV head-set microphone and an Edirol R-05 recorder (44.1 kHz sampling frequency, 16-bit resolution).

Table 5.2: Target words.

	[s]			[ʃ]	
Saab	[saːp]	*car brand*	sjaal	[ʃaːl]	‘scarf’
set	[sɛt]	‘set’	chef	[ʃɛf]	‘boss’
Sieb	[sip]	*name*	chic	[ʃik]	‘modish’
sop	[sɔp]	‘soap’	shop	[ʃɔp]	‘shop’
soep	[sup]	‘soup’	Sjoerd	[ʃuʁt]	*name*

The recorded sentences (n^{total}=1870, n^{T0}=880, n^{T1}=510, n^{T2}=480) were pre-annotated in the free open-source software `Praat` (Boersma and Weenink, 2016) using the automatic speech recognition plugin `Praatalign` (Lubbers and Torreira, 2016). Vowel, sibilant and target word boundaries were checked and corrected manually by the first author. Target word durations were extracted to model potential speech rate effects on acoustic realization of the sibilants and the vowels.

Sibilants

We derived spectral density estimates from the sibilant signals (cf. Reidy, 2015; Forrest et al., 1988). The middle third of each respective sibilant was chosen as analysis interval to minimize coarticulation effects from flanking vowels. We applied the Center of Gravity (COG) approach, which is based on discrete Fourier transform (DFT) to measure sibilant acoustics. The COG measure as is implemented in the software `Praat` was chosen because it is the typical approach to quantifying sibilant acoustics (Reidy, 2015). A pre-emphasis of 6 dB/octave for frequencies above 80 Hz was applied to the sibilant analysis intervals. The resulting spectra were cepstrally smoothed and spectral moments were then calculated. Only the values of the sibilants' first spectral moment (Center of Gravity, COG) were analyzed. Finally, all frequency values in Hertz were transformed to Equivalent Rectangular Bandwidth scale (ERB, Glasberg and Moore, 1990), to account for the frequency-dependent filter properties of the auditory system.

Vowels

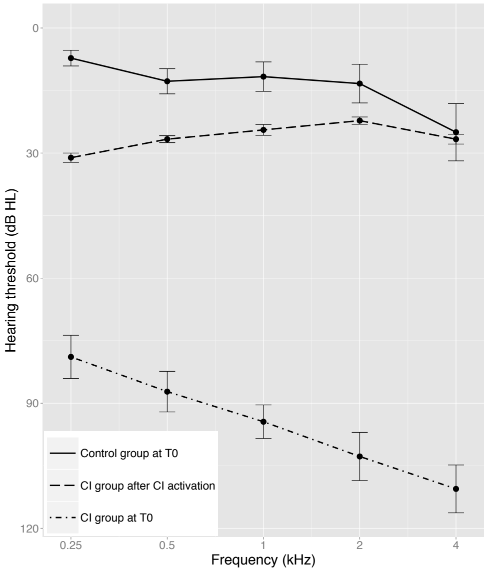
The vowel acoustics were analyzed using two alternative approaches. We analyzed the first two vowel formants for the different vowel categories. The first vowel formant (F1) is thought to reflect vowel height, the second vowel formant (F2) quantifies vowel frontness. In addition, the speakers' average vowel space (AVS, as a measure of vowel dispersion) was investigated, based on the vowel categories per speaker. The AVS approach entails averaging over vowel repetitions per speaker and per test session, while data analysis of the formant measurements allowed for full data analyses of F1 and F2 across the different vowels.The size of one's vowel space accounts for a considerable amount of speech intelligibility (cf. Turner et al., 1995). Vowel space is traditionally defined in a coordinate system formed by the properties vowel height on the vertical axis and vowel frontness on the horizontal axis.

We restricted our vowel analyses to target words starting with [s] to exclude coarticulation effects for the targets starting with [ʃ]. A linear predictive coding algorithm implemented in Praat was used to calculate the first and second vowel formants (in Hz). All formants were calculated at the vowel midpoint (analysis window length: 25 ms). Visual inspection of the individual vowel spaces and vowel formant values revealed quite some inaccurate measurements for F1 and F2 for the rounded vowels [u] and [ɔ] compared to reference values for Dutch vowel formants (Pols et al., 1973; Adank et al., 2004b; van der Harst, 2011). These productions (approximately 20 percent of all rounded vowels) were therefore manually corrected by inspecting vowel spectrum envelopes to derive reliable formant values. Lobanov's vowel normalization procedure as implemented in the `R` package `phonR` was used to to account for vocal tract length differences across speakers (Lobanov, 1971; Adank et al., 2004a). This procedure z-normalized each speakers' vowel formants F1 and F2 in Hertz at each test session (i.e., replacing each F1/F2 value by a z-score relative to the grand mean and standard deviation of all F1/F2 values for this vowel of this speaker). The `phonR` package was also employed to calculate average vowel spaces (AVS) per subject per test session. The area of the vowel space pentagon (in which each of the five vowels formed a corner) was calculated based on the normalized F1 and F2 values per speaker (as described above).

5.2.4 Background variables

Both prior to and after implantation, all CI participants completed a word recognition test plus standard pure-tone audiometry, and also filled in a questionnaire on hearing-related quality of life. These measures yielded both subjective and objective evaluations of improvements due to cochlear implantation. For the control group only pure-tone thresholds were tested at test session T0. The word recognition test and the questionnaire are described below.

Figure 5.2: Mean audiometric pure-tone air conduction thresholds as a function of test frequency. Error bars represent standard errors. The figure shows performance at T0 for the CI and the control group (CI group: ear to be implanted, control group: best ear). Additionally, thresholds for the implanted ear are shown for the CI group at approximately eight weeks after activation of the cochlear implant.



Pure-tone audiometry

Hearing of the CI candidates and of the control group participants was screened with air-conduction pure-tone audiometry. Figure 5.2 displays the results of the unaided pure-tone audiometry for octave frequencies from 0.25 to 4 kHz for both the CI group and the control group at test session T0 (i.e., prior to cochlear implantation for the CI group) as well as the retested pure-tone thresholds for the implanted ear approximately eight weeks after activation of the CI. Post-operative Pure-Tone Average results (averaged over 1, 2 and 4 kHz) were significantly better in the implanted ear than at test session T0 (paired *t*-test: $p<.001$).

CVC audiometry

CI participants' aided word recognition performance was tested at all three test sessions (T0, T1, T2) using a standard Dutch speech audiometry test, the CVC word material from Bosman and Smoorenburg (1992, 1995). This CVC test is common in clinical practice in the Netherlands. The test material, produced by a female native speaker of Dutch (as spoken in the Netherlands), consists of meaningful monosyllabic words (e.g., "naam", 'name') arranged in lists of 12 items. The material was presented through Behringer MS16 loudspeakers placed in front of the listener (0° azimuth) at a distance of one meter. The CVC words were presented at a fixed intensity level of 65 dB SPL without masking noise. In each test session (T0, T1, T2) participants were presented with five different consecutive lists, which resulted in a maximum accuracy score of 165 phonemes correct per session (5 lists × 11 items × 3 phonemes). Word recognition score was quantified as the percentage of correctly reproduced phonemes (max. three per test item), discarding the first item of each list (which is considered a practice item). Higher values indicate better word recognition. Table 5.3 provides the descriptive results for aided word recognition performance before cochlear implantation (T0) and at the two subsequent test sessions after implantation (T1, T2). The results of paired *t*-tests (cf. Table 5.3) show that word recognition improves at three months post CI activation (T2) compared to before cochlear implantation (T0), whereas no significant improvement in word recognition was found at T1 (relative to T0).

Subjective cochlear implant outcome

CI patients filled in the Nijmegen Cochlear Implant Questionnaire (NCIQ; Hinderinck et al., 2000) before cochlear implantation (T0) and at the test session around three months after CI activation (T2), except for participant IJ, who did not consent to repeated administration of the NCIQ. The NCI questionnaire has been shown to quantify improvements between pre- and post-CI measurements (ibid.). The material consists of 60 questions with ratings given on a five point Likert scale. Higher values indicate greater improvement in the respective domain. Participants rated their overall (hearing-related) quality of life to be significantly better (see Table 5.3) at three months after activation of the CI (T2) than prior to surgery (T0).

Table 5.3: Means and standard deviations of hearing-related variables for the CI group for the three test sessions (T0, T1, T2) and results of test statistics investigating differences due to cochlear implantation across test sessions (Paired *t*-tests, uncorrected *p*-values).

	Test session			Comparisons		
background	T0	T1	T2	T0–T1	T0–T2	T1–T2
variable	*M(SD)*	*M(SD)*	*M(SD)*	*p*<	*p*<	*p*<
CVC audiometry	60.47 (17.92)	73.13 (12.55)	79.73 (9.54)	$.122^{ns}$	$.024^{*}$	$.124^{ns}$
NCI-Questionnaire	52.30 (8.71)	–	75.95 (11.11)	–	$.001^{***}$	–

Significance level notation: ***$p<.001$, *$p<.05$.

5.2.5 Data analyses

Four acoustic variables were investigated: one measure of sibilant acoustics (the spectral Center of Gravity, or COG measure) and three vowel acoustics measures (AVS, F1, F2). The Center of Gravity (COG) measure informs about the highest peak in the spectral energy distribution of a sound (spectral mean). This measure was used to investigate the sibilant productions.

Statistical regression analyses were run in the program `R` for the four dependent variables. Simple linear regression models were calculated for the dependent variable AVS to investigate research questions 1B and 2B (1B: pre-surgery effects of severe hearing loss on vowel acoustics, 2B: effect of restored auditory feedback on vowel acoustics). All other analyses were run as mixed-effect regression models calculated with the `lme4` package (Bates et al., 2015). We employed the random-effects structure that was appropriate for the structure of the respective data set. For example, the dependent variable COG was analyzed using random intercepts for both subjects and items. Vowel identity was entered as fixed effect in the model structure for the dependent variables F1 and F2 and thus we modeled the formant data using random intercepts for subjects only (as items were confounded with the different vowels).

P-values were obtained following a twofold approach. Firstly, we used the `Anova` function of the `car` package (Type II Wald χ^2 test), which allows us to specify the significance of the contribution of a factor to a regression model fit independently of the factor level set to the model intercept. Secondly, we calculated *p*-values using the model's *t* values. The number of degrees of freedom was estimated via the Kenward-Roger approximation as implemented in the `R` package `pbkrtest`. Potential discrepancies between *p*-values via the `Anova`-approach and the *p*-values using the model's *t*-values originate from the contribution of factor levels set to the intercept. The more appropriate method to calculate *p*-values is the former (`Anova`) approach as it quantifies model fit given all model parameters. Nevertheless, we also included the *p*-values using the model's *t*, to be able to extract and visualize the unique contribution of factor levels as simple effects and/or in interaction with other variables

RQ1 – Pre-surgery hearing loss effects on sibilant and vowel acoustics

Production data of the first test session (T0) were analyzed to investigate whether the acoustic sibilant and vowel realizations of the CI candidates differed from those of the control sample. Number of observations for the sibilants was 450 for the control group (9 participants × 10 target words × 5 repetitions each) and 430 for the CI group (one out of nine participants produced three instead of five target word repetitions).

Sibilant acoustics: Vocalic context (± round, unrounded vowel context on intercept) and local speech rate for the target word were included as item-related control variables for the sibilant data. Speech rate was quantified as the inverse of the word duration (continuous variable, *z*-transformed) . Vowel roundedness, rather than vowel identity with all possible factor levels (i.e., [aː, ɛ, i, ɔ, u]), was entered in the analysis since only vowel roundedness was expected to considerably affect coarticulation for sibilants. The CI and the control group differed in speech rate at session T0, reflected in longer durations of the target words (approx. 60 ms longer) for the CI group than for the control group (Wilcoxon signed-rank test, $p<.001$). Speaker gender was included as a participant-related control variable. The predictors of interest were sibilant identity ([s] vs. [ʃ], with [ʃ] productions on intercept) and group affiliation (CI or CONTROL, the latter on the intercept). The sibilant model (RQ1A-COG) included the critical interaction of sibilant identity and group (sibilant × group) to test whether the groups differ in the sibilant contrast. We also included the interaction between sibilant identity and vowel context (sibilant × vowel context) and the interaction of speaker gender and sibilant identity (sibilant × gender). These interactions were included because coarticulation effects due to vowel roundedness have been shown to be stronger for [s] productions than for [ʃ] productions and because female participants produce enhanced acoustic sibilant contrasts compared to men (e.g., Koch and Janse, 2015).

Vowel acoustics: A linear regression model (RQ1B-AVS) was employed to test for group differences in vowel space (AVS). For the vowel formants F1 and F2 mixed-effect models were fitted (RQ1B-F1, RQ1B-F2) with vowel identity ([aː, ɛ, i, ɔ, u]) and group affiliation (CI or CONTROL, with CONTROL on the intercept) as predictors. Speech rate was included as item-related control variable for the analysis of F1 and F2. For the formant analyses (F1, F2) we set productions of the most open, central vowel [aː] on the model intercept. The F1 and F2 models included the interaction of vowel identity and group (vowel × group) to test whether potential group differences in acoustic vowel production were vowel-specific.

RQ2 – Effect of restored auditory feedback on sibilant and vowel acoustics

Sibilant and vowel production data for all three test sessions (T0, T1, T2) were analyzed to investigate the effect of restored auditory feedback on sibilant and vowel acoustics. The sibilant analysis was based on 990 observations for the control group and on 879 observations for the CI group. Again, vowel analyses were limited to the [s] target words only and thus included only half of the observations mentioned above.

Sibilant acoustics: To test whether possible changes in the CI groups' sibilant acoustics across test sessions differed from potential changes due to repeated reading of the same material in the control group, the linear mixed-effects model for RQ1 (RQ1A-COG) above was complemented by adding the interval-scale variable test session (T0, T1, T2; with T0 on intercept) in interaction with sibilant and group (i.e., sibilant × group × test session).

Vowel acoustics: Similarly, the effect of restored auditory feedback on average vowel space (AVS) was analyzed using a linear regression model (CI group: n=24, control group: n=27). Group and test session were included as fixed effects, as well as their interaction. For F1 and F2, linear mixed-effects model were built with test session in interaction with vowel identity and group (i.e., vowel × group × test session).

Effect of duration of hearing loss on sibilant and vowel acoustics after CI

To investigate whether duration of hearing loss related to changes in acoustic realization of sibilants and vowels, we analyzed the production data of the CI group separately. For each of the four dependent variables (COG, AVS, F1 and F2), we ran separate regression models including duration of hearing loss as predictor. Random effect structures were employed for COG, F1 and F2 (linear mixed-effect models), whereas for AVS we ran a simple regression analysis. We analyzed 879 observations for the sibilant data and 439 data points for the vowel acoustics.

Sibilant acoustics: As in the sibilant models introduced above, we included the item-related variables speech rate and vocalic context as well as the participant-related variable speaker gender as control variables in our analysis (for sibilant models only). The interaction between sibilant identity and vocalic context (sibilant × vocalic context) and the interaction of speaker gender and sibilant identity (sibilant × gender) were also included. The COG model was run with three-way interactions between the test session (T0, T1, T2; with T0 on intercept), sibilant identity ([s] vs. [ʃ], the latter on the model intercept) and the predictor duration of hearing loss (continuous variable, z-transformed).

Vowel acoustics: For the dependent variables AVS, F1 and F2 we modeled the crucial interaction of test session with duration of hearing loss. As in the models above speech rate was added as item-related control variable for the formant analyses.

5.3 Results

RQ1A-COG – Pre-surgery hearing loss effects on sibilant acoustics

We analyzed the spectral means (Center of Gravity, COG) of the sibilant productions at test session T0 (n=879, M=30.19, SD=1.61) to investigate whether CI candidates show deviant sibilant acoustics compared to a control sample. The result of the statistical model testing for the critical interaction between group and sibilant category (including the control predictors vocalic context, speech rate and gender) is shown in Table 5.4 and is illustrated in the data plot (Figure 5.3). Sibilants differed significantly in their spectral characteristics with higher spectral means for [s] than for [ʃ] target words (sibilant identity effect: $|\beta|$=1.65, SE=0.15, χ^2=270.07, p<.001). The sibilant's spectral mean was lower if rounded vowels followed the sibilant in comparison with unrounded vowels (vocalic context effect: $|\beta|$=0.20, SE=0.13, χ^2=18.42, p<.001). Vocalic context affected the spectral mean of [s] more than that of [ʃ] (vocalic context × sibilant identity interaction effect: $|\beta|=0.38$, $SE=0.19$, $\chi^2=4.27$, $p<.05$). This was interpreted as stronger anticipatory coarticulation during [s] than [ʃ].

Figure 5.3: Means and standard errors for the sibilants' Center of Gravity (measured in ERB) for CI and control group at test session T0 (pre-implant session). [s] productions symbolized as circles, [ʃ] productions as triangles.

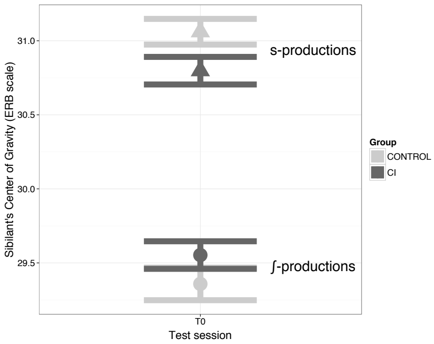


Table 5.4: Model testing for effects of group differences in sibilants' Center of Gravity at test session T0 (RQ1A, pre-implant session).

Fixed effects	β	SE	$p(t)<$	$p(\chi^2)<$
Intercept	30.72	0.45		
Sibilant identity	1.65	0.15	.001***	.001***
Vocalic context	-0.20	0.13	.141^{ns}	.001***
Speech rate	-0.03	0.05	.550^{ns}	.541^{ns}
Group	0.38	0.43	.396^{ns}	.773^{ns}
Gender	-1.92	0.48	.002**	.001***
Sibilant identity × voc. context	-0.38	0.19	.055 .	.039*
Sibilant identity × gender	0.32	0.13	.024*	.013*
Sibilant identity × group	-0.51	0.11	.001***	.001***

Reference levels: sibilant identity: [ʃ], vocalic context:−round, group: control group, gender: female.
P-values were calculated in two ways: 1. based on t-values applying the Kenward-Roger procedure to approximate degrees of freedom (`R pbkrtest` package) and 2. using the `Anova` function of the `car` package (Type II Wald χ^2 test).
Significance level notation: ***$p<.001$, **$p<.01$,*$p<.05$, . $p<.10$.

Female speakers produced generally higher spectral means than male speakers (gender effect: $|\beta|=1.92$, $SE=0.48$, $\chi^2=13.46$, $p<.001$). Female speakers also produced greater acoustic sibilant contrasts than male speakers (gender × sibilant identity interaction effect: $|\beta|=0.32$, $SE=0.13$, $\chi^2=6.26$, $p<.05$). No general group effect was observed on the COG of the sibilant productions. Crucially, however, we observed a significant group × sibilant interaction effect ($|\beta|=0.51$, $SE=0.11$, $\chi^2=19.85$, $p<.001$). This means that acoustic sibilant contrasts were smaller for the CI group prior to implantation compared to the control group (at test session T0).

RQ1B – Pre-surgery hearing loss effects on vowel acoustics

To investigate possible group differences in vowel acoustics we analyzed the average vowel space, as well as the first and the second vowel formant frequencies (F1, F2) at test session T0.

AVS: Only 18 data points remained for analysis ($M^{\mathrm{AVS}}=1.78$, $SD=0.47$) after averaging to a single AVS measure per participant for test session T0. A simple linear regression was run to predict average vowel space (AVS) based on group identity. The regression analysis did not provide evidence for AVS differences between groups at baseline test session T0 ($F(1,16)=1.175$, $p=.295$).[7]

F1: The mixed-effect regression model for the dependent variable F1 is shown in Table 5.5. Figure 5.4 illustrates the distribution of F1 values per speaker group for the vowel categories at test session T0. We observed a significant vowel identity effect ($\chi^2=4217.27$, $p<.001$), a significant group effect ($\chi^2=8.69$, $p<.01$), as well as a significant interaction effect of group × vowel identity for F1 ($\chi^2=9.70$, $p<.05$). As expected, F1 values differed depending on vowel identity. The highest F1 values were shown for the open vowel [aː] (on model intercept) and the lowest F1 values were found for the closed vowel [i]. According to the model, the CI group produced numerically higher F1 values than the control group, exept for the vowel [i]. The significant group × vowel identity interaction effect suggests that especially for the mid-open vowels [ɛ] and [ɔ] the participants in the CI group produced higher F1 frequencies than the control group. As a sanity check, an alternative simpler model was run, leaving out the interaction of vowel identity and group (note that this model had a poorer fit than the more complex model: $\chi^2=9.79$, $p<.05$). In this simpler model a significant group effect was observed ($\chi^2=8.69$, $p<.01$). These combined results suggest that the group difference in vowel realization is particularly driven by the groups' different realization of the vowels [ɛ] and [ɔ].

---

[7] The AVS results did not change if AVS was based on the raw, rather than the Lobanov-transformed, formant data.

Figure 5.4: Means and standard errors for F1 per vowel category for the CI and control group at test session T0 (RQ1B-F1, pre-implant test session).

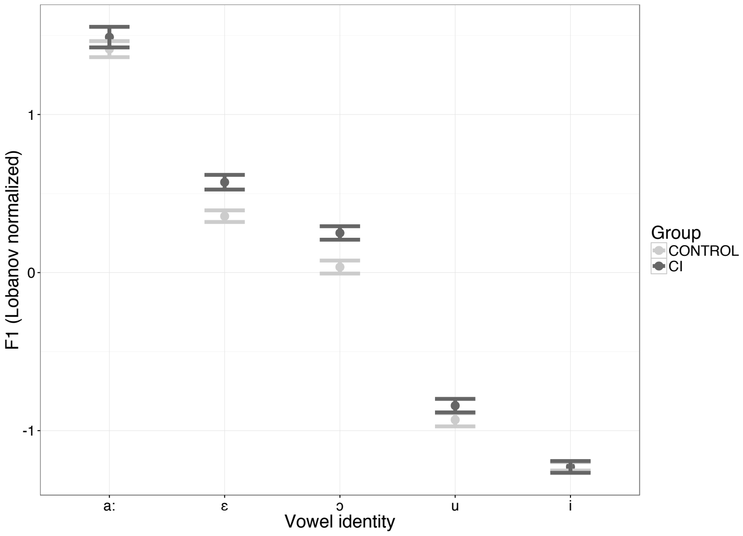


Table 5.5: Model testing for group differences in vowel F1 at test session T0 (before cochlear implantation, RQ1B-F1).

Fixed effects	β	SE	$p(t) <$	$p(\chi^2)<$
Intercept	1.43	0.05		
Vowel [ɛː]	-1.08	0.06	.001***	.001***
Vowel [i]	-2.67	0.06	.001***	
Vowel [o]	-1.40	0.06	.001***	
Vowel [u]	-2.37	0.06	.001***	
Speech rate	0.03	.020	.209^{ns}	.210^{ns}
Group	0.09	0.07	.213^{ns}	.004**
Vowel [ɛː] × group	0.15	0.09	.088.	.046*
Vowel [i] × group	-0.07	0.09	.418^{ns}	
Vowel [o] × group	0.14	0.08	.094.	
Vowel [u] × group	0.02	0.09	.772^{ns}	

Reference levels: vowel identity:[aː], group: control group.
P-values were calculated in two ways: 1. based on t-values applying the Kenward-Roger procedure to approximate degrees of freedom (`R pbkrtest` package) and 2. using the `Anova` function of the `car` package (Type II Wald χ^2 test).
Significance level notation: ***p<.001, **p<.01, *p<.05, . p<.10.

F2: A significant vowel identity effect (χ^2=11356.33, *p*<.001), as well as significant group × vowel identity interaction effect (χ^2=27.55, *p*<.001) was observed for F2 (cf. Table 5.6). Figure 5.5 illustrates the distribution F2 values per group for the vowel categories at test session T0. As expected, the most front vowel [i] showed the highest F2 values. Lower F2 values were observed for less fronted vowels ([i]>[ɛ]>[aː]>[ɔ]>[u]). The lack of a significant group effect (χ^2=2.51, *p*>.1) indicates that F2 was not simply shifted upwards or downwards across all vowels. The significant group × vowel identity interaction effect indicates that the CI candidates produced the vowels [aː] and [ɛ] in a more standard way (control alike) than the three vowels [i], [ɔ] and [u], which were produced with higher F2 and thus more fronted by the CI group than by the controls (cf. Figure 5.5).

Figure 5.5: Means and standard errors for F2 per vowel category for the CI and control groups at test session T0 (before cochlear implantation).

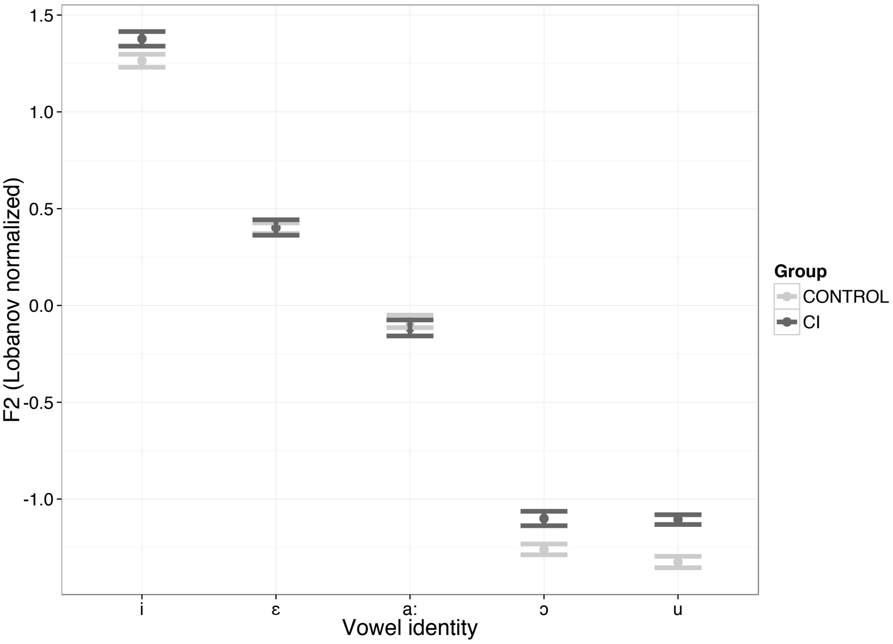
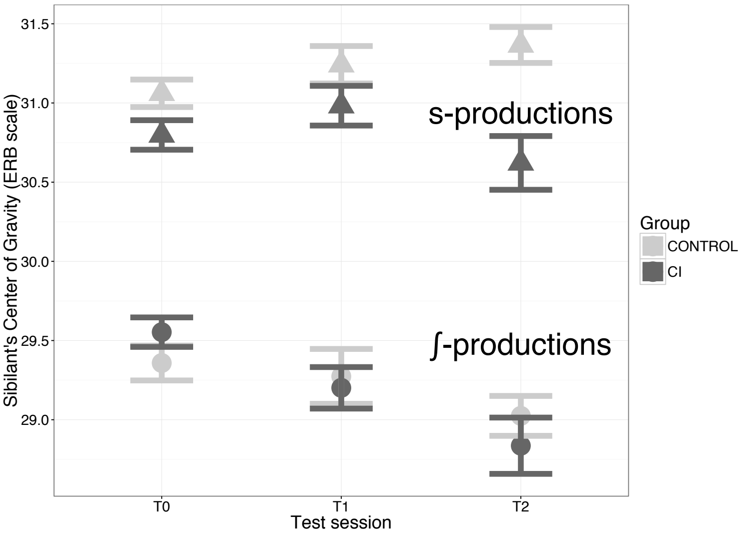


Table 5.6: Model testing for group differences in vowel F2 at test session T0 (before cochlear implantation, RQ1B-F2).

Fixed effects	β	*SE*	$p(t)<$	$p(\chi^2)<$
Intercept	-0.08	0.05		
Vowel [i]	1.34	0.05	.001***	.001***
Vowel [ɛː]	0.48	0.04	.001***	
Vowel [o]	-1.18	0.04	.001***	
Vowel [u]	-1.25	0.04	.001***	
Speech rate	0.01	0.02	.702^{ns}	.701^{ns}
Group	-0.03	0.07	.709^{ns}	.114^{ns}
Vowel [i] × group	0.15	0.06	.010*	.001***
Vowel [ɛː] × group	0.04	0.06	.492^{ns}	
Vowel [o] × group	0.19	0.06	.001***	
Vowel [u] × group	0.26	0.06	.001***	

Reference levels: vowel identity:[aː], group: control group.
P-values were calculated in two ways: 1. based on *t*-values applying the Kenward-Roger procedure to approximate degrees of freedom (`R pbkrtest` package) and 2. using the `Anova` function of the `car` package (Type II Wald χ^2 test).
Significance level notation: $^{***}p<.001$, $^{*}p<.05$, $^{ns}p>.10$.

Figure 5.6: Means and standard errors for the sibilants' Center of Gravity (in ERB) for CI and control group across the three test sessions (T0, T1, T2). [ʃ] productions symbolized as circles, [s] productions as triangles.



RQ2A – Effect of restored auditory feedback on sibilant acoustics

We compared changes in sibilant acoustics across the three test sessions for the CI and the control group (n=1869 in total, M^{COG}=30.13, SD=1.75). The result of the statistical model testing for the critical interaction between group, sibilant category and test session is shown in Table 5.7 and is illustrated in Figure 5.6.

This model replicates the previous effects of sibilant identity, vowel context and gender. In addition, it also replicates the earlier interaction effects between vowel context and sibilant identity, between gender and sibilant identity as well as between speaker group and sibilant identity. As can be seen from Figure 5.6, test session affected the sibilants' spectral mean (χ^2=14.91, p<.001, [ʃ] on model intercept) with lower COGs for sessions T1 and T2 compared to T0 (particularly between T0 and T2).

Table 5.7: Model testing for test session effects on sibilants' Center of Gravity in the two groups (RQ2A).

Fixed effects	β	SE	$p(t)<$	$p(\chi^2)<$
Intercept	30.90	0.43		
Sibilant identity	1.48	0.14	.001***	.001***
Vocalic context	-0.27	0.12	.021*	.001***
Speech rate	-0.02	0.03	.659^{ns}	.660^{ns}
Group	0.44	0.42	.299^{ns}	.992^{ns}
Gender	-2.14	0.46	.001***	.001***
Test session T1	-0.09	0.10	.379^{ns}	.001***
Test session T2	-0.33	0.10	.001***	
Sibilant identity × voc. context	-0.37	0.17	.031*	.031*
Sibilant identity × gender	0.56	0.09	.001***	.001***
Sibilant identity × group	-0.54	0.12	.001***	.001***
Test session T1 × group	-0.32	0.14	.028*	.001***
Test session T2 × group	-0.42	0.15	.005**	
Test session T1 × Sibilant identity	0.26	0.14	.065$_{.}$	.001***
Test session T2 × Sibilant identity	0.64	0.14	.001***	
Sibilant identity × group × Test session T1	0.28	0.20	.165^{ns}	.294^{ns}
Sibilant identity × group × Test session T2	0.04	0.21	.861^{ns}	

Reference levels: sibilant identity: [ʃ], Test session: T0, vocalic context:−round, group: control group, gender: female. P-values were calculated in two ways: 1. based on t-values applying the Kenward-Roger procedure to approximate degrees of freedom (`R pbkrtest` package) and 2. using the `Anova` function of the `car` package (Type II Wald χ^2 test). Significance level notation: ***p<.001, **p<.01, *p<.05, . p <.10.

We observed a gradual increase in sibilant contrasts over test sessions, as shown by a significant test session × sibilant identity interaction effect (χ=40.63, p<.001).

However, the critical three-way interaction between test session × group × sibilant identity, reflecting differential improvement patterns for the CI group compared to the reference sample, was not significant. Both groups show increased sibilant contrasts at later sessions (T1, T2) compared to baseline T0. However, even though not evident from a statistically significant three-way interaction, the data plot in Figure 5.6 suggests that the two groups show differences in the direction of change in sibilant acoustics. The control group seems to increase the sibilant contrasts mainly by raising the COG for the [s] targets words and partly (only for T2 vs. T0) by lowering the COG for the [ʃ] productions. In contrast, the CI group predominantly achieves enhanced sibilant contrast by lowering the COG for the [ʃ] productions.

To address the possibility that groups differed in what sibilant changed most across sessions, we fitted two additional models for the [s] and the [ʃ] sibilants, respectively (with session × group being the critical interaction for both models, see Table 5.8 and Table 5.9).

Table 5.8: Model testing for test session × group interaction effects on COG for [s] sibilant productions.

Fixed effects	β	*SE*	$p(t)<$	$p(\chi^2)<$
Intercept	32.51	0.44		
Vocalic context	-0.92	0.16	.001***	.001***
Speech rate	-0.02	0.04	.573^{ns}	.566^{ns}
Group	-0.09	0.42	.843^{ns}	.545^{ns}
Gender	-1.77	0.47	.002**	.001***
Test session T1	0.18	0.08	.053 .	.089 .
Test session T2	0.30	0.08	.003**	
Test session T1 × Group	-0.10	0.13	.442^{ns}	.001***
Test session T2 × Group	-0.53	0.13	.001***	
Vocalic context × Gender	0.39	0.11	.004**	.001***

Reference levels: vocalic context:−round, group: control group, gender: female.
P-values were calculated in two ways: 1. based on *t*-values applying the Kenward-Roger procedure to approximate degrees of freedom (`R pbkrtest` package) and 2. using the `Anova` function of the `car` package (Type II Wald χ^2 test).
Significance level notation: ***p<.001,**p<.01, . p<.10.

The model on the [s] data subset showed that, in contrast to the control sample participants, COG decreased for the CI group over test sessions (test session × group interaction effect: χ^2=17.05, p<.001; T1 × group: $|\beta|$=0.10, *SE*=0.13, T2 × group: $|\beta|$=0.53, *SE*=0.13). For the [ʃ] data subset we also observed differential changes in sibilant acoustics for the two groups over test sessions. Again, the CI participants showed a stronger decrease in COGs across sessions compared to the control group (test session effect: χ^2=52.25, p<.001; test session × group identity interaction effect: χ^2=9.52, p<.01).

Table 5.9: Model testing for test session × group interaction effects on COG for [ʃ] sibilant productions.

Fixed effects	β	SE	$p(t)<$	$p(\chi^2)<$
Intercept	30.91	0.51		
Vocalic context	-0.45	0.16	$.026^{*}$	$.048^{*}$
Speech rate	0.06	0.05	$.263^{ns}$	$.219^{ns}$
Group	0.46	0.50	$.392^{ns}$	$.595^{ns}$
Gender	-2.17	0.55	$.007^{**}$	$.001^{***}$
Test session T1	-0.08	0.09	$.438^{ns}$	$.001^{***}$
Test session T2	-0.33	0.09	$.010^{**}$	
Test session T1 × Group	-0.30	0.14	.071 .	$.001^{***}$
Test session T2 × Group	-0.41	0.14	$.027^{*}$	
Vocalic context × Gender	0.00	0.00	.065 .	$.026^{*}$

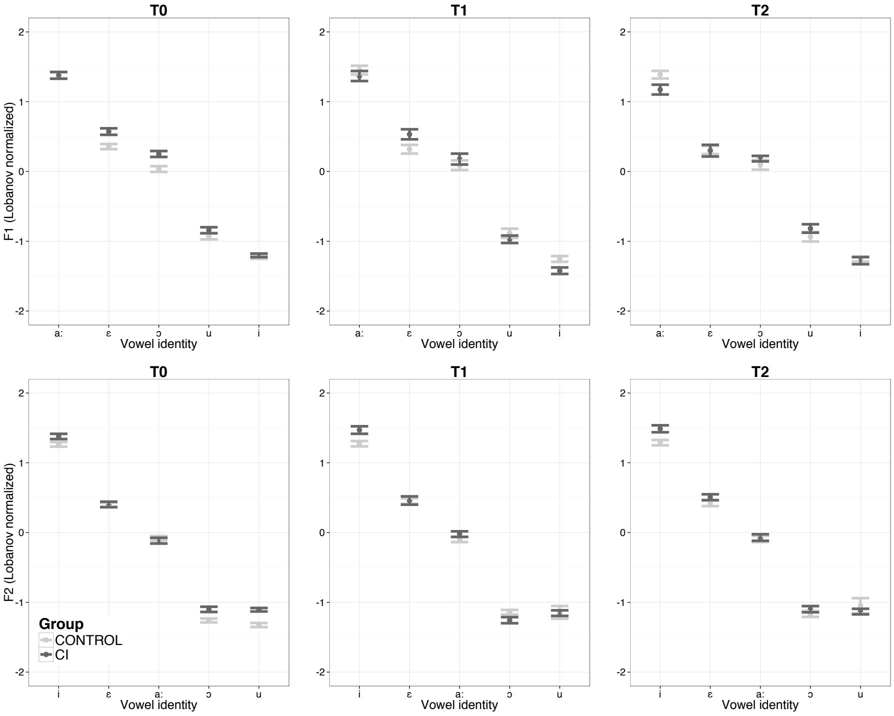
Reference levels: vocalic context:−round, group: control group, gender: female.
P-values were calculated in two ways: 1. based on t-values applying the Kenward-Roger procedure to approximate degrees of freedom (`R pbkrtest` package) and 2. using the `Anova` function of the `car` package (Type II Wald χ^2 test).
Significance level notation: ***p<.001, **p<.01, *p<.05, . p<.10.

RQ2B – Effect of restored auditory feedback on vowel acoustics

We tested for two-way interaction of test session × group in the average vowel space (AVS) measurements, and for a three-way interaction of vowel identity × test session × group in the first two vowel formants (F1, F2). As mentioned above, all vowel analyses were restricted to realizations of [s] target words. Observing the interaction for the AVS data would indicate group-specific changes in vowel space across test sessions, whereas observing the three-way interaction in the formant analyses would imply that the group-specific changes across test sessions were vowel-specific. The results of the statistical testing for each of the three dependent variables will be described below.

AVS: Fifty-one data points were available for analysis (M^{AVS}=1.76, SE=0.50) after averaging to a single AVS measure per participant per test session. AVS was neither affected by group identity (p>.1) nor by test session (p>.1). Furthermore, the test session × group interaction was also not significant (p>.1). Because of these null results the statistical model is not provided here in detail.[8]

Figure 5.7: Means and standard errors for normalized F1 (upper row) and F2 (lower row) as a function of vowel category. Each panel presents error plots for one test session (T0, T1, T2) with color coding for test groups (CI, control).



---

[8] Again, the AVS results did not change if AVS was based on the raw, rather than the Lobanov-transformed, formant data.

F1: 934 vowel measurements were included to analyze effects of test session and group (and their interaction) on the first vowel formant. The resulting model is presented in Table 5.10. The upper panels of Figure 5.7 show how F1 was affected by test session and group across the five vowel categories.

Besides an expected vowel identity effect (χ^2=8007.49, p<.001), a marginally significant effect of group (χ^2=2.99, p<.1), a significant effect of test session (χ^2=7.53, p<.05) and a significant interaction effect of vowel identity × group (χ^2=20.55, p<.001) were found. The CI group thus tended to produce generally higher F1 values (at T0), which corresponds to more open vowel productions, except for vowel [i]. The significant test session × group interaction shows that the CI group's F1 values changed differentially across test sessions as compared to those of the control group. The CI group shows a lowering of F1 at later sessions, particularly at T2 relative to the baseline at T0. The non-significant three-way interaction of vowel identity × session × group indicates that the F1 decrease across test sessions in the CI group was not vowel-specific.

For all vowels except for the closed vowel [i], numerically higher F1 values were found for the CI group when compared with the controls (at test session T0). The F1 decrease in the CI group may be interpreted as a recovery towards more normal vowel openness due to cochlear implantation.[9]

F2: Table 5.11 shows the model that investigates session effects on F2 for the two test groups (n=932). This model replicates the previous effects of vowel identity and the interaction effect of vowel context and speaker group. Furthermore, the model yielded a general effect of test session (χ^2=12.67, p<.01), with higher F2 values for test sessions T1 and T2 in comparison to test session T0. We did not observe a significant vowel identity × test session effect (χ^2=4.87, p>.1), indicating that the test session effect was not vowel-specific. The lack of a significant group × session effect (χ^2=2.04, p>.1) suggests that the two groups showed a similar pattern over sessions. However, the significant test session × group × vowel identity interaction (χ^2=30.34, p<.001) implies that the two groups differed across test sessions especially for the back vowels [u] and [ɔ]. Whereas the CI group produced these back vowels slightly more centrally at T0 than the controls, the CI group produced them less centrally at the later test sessions T1 and T2 compared to T0 baseline. This suggests a change, or recovery, towards more prototypical back vowel acoustics for F2 due to cochlear implantation (see Figure 5.7).

---

[9] A more parsimonious model, which did not include the three-way interaction of vowel identity × group × test session, confirmed the results of the more complex model regarding the simple effects of group and test session and the two-way interactions (vowel identity × group; test session × group).

Table 5.10: Model testing for effects of test session on the two groups' first vowel formant (RQ2B-F1).

Fixed effects	β	*SE*	$p(t)$<	$p(\chi^2)$<
Intercept	1.42	0.05		
Vowel [ɛː]	-1.07	0.06	.001***	.001***
Vowel [i]	-2.66	0.06	.001***	
Vowel [o]	-1.39	0.06	.001***	
Vowel [u]	-2.36	0.06	.001***	
Speech rate	0.01	0.01	.316^{ns}	.316^{ns}
Group	0.08	0.07	.227^{ns}	.084 .
Test session T1	0.04	0.07	.562^{ns}	.024*
Test session T2	0.004	0.07	.959^{ns}	
Test session T1 × group	-0.14	0.11	.197^{ns}	.005**
Test session T2 × group	-0.32	0.11	.003**	
Vowel [ɛː] × group	0.14	0.09	.113^{ns}	.001***
Vowel [i] × group	-0.07	0.09	.408^{ns}	
Vowel [o] × group	0.14	0.09	.116^{ns}	
Vowel [u] × group	0.02	0.09	.831^{ns}	
Vowel [ɛː] × Test session I	-0.07	0.10	.479^{ns}	.162^{ns}
Vowel [i] × Test session I	-0.07	0.10	.513^{ns}	
Vowel [o] × Test session I	0.01	0.10	.919^{ns}	
Vowel [u] × Test session I	0.002	0.10	.979^{ns}	
Vowel [ɛː] × Test session II	-0.05	0.10	.627^{ns}	
Vowel [i] × Test session II	-0.03	0.10	.783^{ns}	
Vowel [o] × Test session II	0.05	0.10	.617^{ns}	
Vowel [u] × Test session II	-0.01	0.10	.934^{ns}	
Vowel [ɛː] × Test session I × group	0.12	0.15	.431^{ns}	.169^{ns}
Vowel [i] × Test session I × group	-0.04	0.15	.792^{ns}	
Vowel [o] × Test session I × group	0.01	0.15	.962^{ns}	
Vowel [u] × Test session I × group	-0.05	0.15	.740^{ns}	
Vowel [ɛː] × Test session II × group	0.09	0.15	.561^{ns}	
Vowel [i] × Test session II × group	0.29	0.15	.053 .	
Vowel [o] × Test session II × group	0.19	0.15	.198^{ns}	
Vowel [u] × Test session II × group	0.35	0.15	.022*	

Reference levels: vowel identity: [aː], group: control group.
P-values were calculated in two ways: 1. based on *t*-values applying the Kenward-Roger procedure to approximate degrees of freedom (`R pbkrtest` package) and 2. using the `Anova` function of the `car` package (Type II Wald χ^2 test).
Significance level notation: ***p<.001, **p<.01, *p<.05, . p<.10.

Table 5.11: Model testing for effects of test session on the two groups' second vowel formant (RQ2B-F2).

Fixed effects	β	*SE*	$p(t)<$	$p(\chi^2)<$
Intercept	-0.08	0.05		
Vowel [i]	1.35	0.04	.001***	.001***
Vowel [ɛː]	0.48	0.04	.001***	
Vowel [o]	-1.18	0.04	.001***	
Vowel [u]	-1.24	0.04	.001***	
Speech rate	-0.001	0.01	.942^{ns}	.942^{ns}
Group	-0.03	0.07	.599	.115^{ns}
Test session T1	-0.02	0.05	0.723^{ns}	.002**
Test session T2	-0.01	0.05	0.842^{ns}	
Test session T1 × group	0.12	0.07	.088 .	.361^{ns}
Test session T2 × group	0.06	0.07	.409^{ns}	
Vowel [i] × group	0.15	0.06	.012*	.001***
Vowel [ɛː] × group	0.04	0.06	.512^{ns}	
Vowel [o] × group	0.19	0.06	.001***	
Vowel [u] × group	0.25	0.06	.001***	
Vowel [i] × Test session I	0.03	0.07	.700^{ns}	.772^{ns}
Vowel [ɛː] × Test session I	0.08	0.07	.235^{ns}	
Vowel [o] × Test session I	0.14	0.07	.041*	
Vowel [u] × Test session I	0.12	0.07	.077 .	
Vowel [i] × Test session II	0.03	0.07	.616^{ns}	
Vowel [ɛː] × Test session II	0.03	0.07	.667^{ns}	
Vowel [o] × Test session II	0.07	0.07	.120^{ns}	
Vowel [u] × Test session II	0.15	0.07	.007**	
Vowel [i] × Test session I × group	-0.03	0.10	.778^{ns}	.001***
Vowel [ɛː] × Test session I × group	-0.12	0.10	.237^{ns}	
Vowel [o] × Test session I × group	-0.38	0.10	.001***	
Vowel [u] × Test session I × group	-0.26	0.10	.007**	
Vowel [i] × Test session II × group	0.03	0.10	.745^{ns}	
Vowel [ɛː] × Test session II × group	0.03	0.10	.753^{ns}	
Vowel [o] × Test session II × group	-0.15	0.10	.136^{ns}	
Vowel [u] × Test session II × group	-0.25	0.10	.011*	

Reference levels: vowel identity: [aː], group: control group.
P-values were calculated in two ways: 1. based on *t*-values applying the Kenward-Roger procedure to approximate degrees of freedom (`R pbkrtest` package) and 2. using the `Anova` function of the `car` package (Type II Wald χ^2 test).
Significance level notation: ****p*<.001, ***p*<.01, **p*<.05, . *p*<.10.

Duration of hearing loss as predictor of changes in sibilant and vowel acoustics after CI

We specifically investigated whether the improvement in sibilant acoustics is related to the duration of hearing loss before CI (sibilant identity × test session interaction × duration of hearing loss before CI).

Table 5.12 shows the resulting model. As found in our earlier model (cf. Table 5.7), the acoustic sibilant contrast increases after cochlear implantation (sibilant identity × test session). The significant interaction of test session × duration of hearing loss indicates that changes in sibilant acoustics depend on the duration of the participant's hearing loss (χ^2=21.58, p<.001). Importantly, a significant three-way interaction of sibilant identity × test session × duration of hearing loss was observed (χ^2=8.96, p<.05). That is, the longer the duration of hearing loss was, the smaller the downward shift in COG for the [ʃ] productions resulting in higher COGs for [ʃ].

Note that because [ʃ] productions were used as reference level in our analyses, the model estimates below (cf. Table 5.12) show the co-modulating effect of the duration of hearing loss on the [ʃ] productions (test session × duration of hearing loss). An alternative model with [s] productions set as reference level, showed that in line with our earlier results (cf. Table 5.7, Figure 5.6) COG changes across test session for the CI group were mainly driven by the [ʃ] productions, as revealed by non-significant test session effects for the [s] productions (T0 vs. T1 and T0 vs. T2: t<1, p>.1). Crucially, the improvements were particularly driven by changes in the [ʃ] and not so much by changes in the [s] productions.

Duration of hearing loss as predictors of vowel acoustics after CI

We also investigated whether duration of hearing loss would predict vowel acoustics. Three separate models were run for the dependent variables AVS, F1, and F2. For AVS we ran a linear regression and for the vowel formant measures linear mixed-effect models. None of the models yielded significant interactions of test session and duration of hearing loss (p>.1). The only significant effect found was a higher AVS for participants with longer duration of hearing loss (F(1,18)=6.29, p<.05). However, this effect direction did not match our expectation. Previous studies have, on contrary, reported decreased vowel spaces in post-lingually deafened adults in comparison to controls.

Table 5.12: Model testing for the effect of duration of hearing loss on sibilants' acoustic change after cochlear implantation.

Fixed effects	β	*SE*	$p(t)<$	$p(\chi^2)<$
Intercept	31.13	0.53		
Sibilant identity	1.47	0.16	$.001^{***}$	$.001^{***}$
Vocalic context	-0.25	0.11	.054 .	$.001^{***}$
Speech rate	-0.02	0.05	$.755^{ns}$	$.745^{ns}$
Gender	-1.90	0.59	$.018^{*}$	$.001^{***}$
Test session T1	-0.39	0.10	$.008^{**}$	$.001^{***}$
Test session T2	-0.61	0.11	$.002^{**}$	
Duration of hearing loss	0.37	0.27	$.206^{ns}$	$.033^{*}$
Sibilant identity × duration of hearing loss	0.08	0.09	$.393^{ns}$	$.860^{ns}$
Sibilant identity × gender	-0.18	0.14	$.246^{ns}$	$.201^{ns}$
Sibilant identity × vocalic context	-0.22	0.15	$.198^{ns}$	$.151^{ns}$
Sibilant identity × Test session T1	0.47	0.14	$.014^{*}$	$.001^{***}$
Sibilant identity × Test session T2	0.44	0.15	$.025^{*}$	
Test session T1 × duration of hearing loss	0.26	0.10	$.046^{*}$	$.001^{***}$
Test session T2 × duration of hearing loss	0.59	0.12	$.003^{**}$	
Sibilant identity × Test session T1 × duration of hearing loss	0.01	0.14	$.937^{ns}$	$.012^{*}$
Sibilant identity × Test session T2 × duration of hearing loss	-0.45	0.16	$.030^{*}$	

Reference levels: sibilant identity: [ʃ], Test session: T0, vocalic context:−round, gender: female.
P-values were calculated in two ways: 1. based on *t*-values applying the Kenward-Roger procedure to approximate degrees of freedom (`R pbkrtest` package) and 2. using the `Anova` function of the `car` package (Type II Wald χ^2 test).
Significance level notation: $^{***}p<.001$, $^{**}p<.01$,$^{*}p<.05$.

5.4 Discussion

The current study was set up with two objectives: firstly, to investigate the effect of hearing loss on speech production in post-lingually deafened participants that undergo cochlear implantation; and, secondly, to assess early speech production changes after CI and the time course thereof, taking into account that repeated testing may affect acoustic realization. To this end, we have compared sibilant and vowel productions of a group of post-lingually deafened CI candidates, with an age- and gender-matched normal-hearing control group. Each of the two groups read carrier sentences containing target words featuring sibilant and vowel target sounds (sibilants: [s, ʃ]; vowels: [aː, ɛ, i, ɔ, u]) at three time points (prior to operation – T0, approx. 2 weeks after CI activation – T1, approx. 3 months after CI activation – T2).

Numerous studies have already addressed effects of hearing loss and restored hearing on speech perception (e.g., Blamey et al., 2013; Peelle et al., 2011; Hamzavi et al., 2003). Post-lingually deafened adults who undergo cochlear implantation are the ideal test population for the investigation of the role of auditory feedback for speech production. The study of the effect of hearing loss and restored hearing through cochlear implantation on speech production may inform models of speech production. Speech motor control models usually include feedback control modules that allow the speaker to monitor the auditory consequences of his/her articulation and compare these with stored representations. Healthy speakers generally change their articulation to counteract artificially modified auditory feedback (e.g., Houde and Jordan, 1998). One must assume, therefore, that a comparable mechanism applies for long-term changes to auditory feedback, resulting in at least subtle changes due to deteriorated auditory perception. Some evidence for this comes from a study that showed that an age-related reduction of high-frequency hearing acuity affects sibilant articulation (Koch and Janse, 2015). While reviewing the existing literature on speech production changes after CI we have found that methodological issues in some of the studies raise questions as to how to interpret their findings. Crucially, ignoring gender differences in the analysis of speech production may be problematic, in particular for comparisons of sibilant production. In contrast to vowel formant normalization procedures, which have been shown to minimize anatomical/physiological variation between speakers, no such procedure is available to account for the sociolinguistic phenomenon of sibilant (contrast) differences due to gender. Furthermore, despite the availability of normalization procedures to account for (sex-based) anatomical differences in vowel formants, these procedures have not been employed in all previous studies. In addition, potential effects of repeated testing on speakers' productions have largely been ignored in former studies. Because of these various methodological issues it is largely unclear whether and how hearing loss acquired later in life affects speech production, and whether and how hearing restoration via a cochlear implant affects speech production.

Changes in speech production due to changes in hearing status have been covered in the literature (e.g., Matthies et al., 1994; Perkell et al., 1992; Lane et al., 2007; Langereis et al., 1997), but less systematically and mostly with small test samples. Furthermore, in contrast to speech perception tasks, speech production tasks are not part of default test batteries at specialized cochlear implant centers, which complicates multicenter and retrospective studies.

In line with previous findings (e.g., Matthies et al., 1994; Lane et al., 2007), CI candidates showed reduced sibilant contrasts ([s] vs. [ʃ]) compared to the control group prior to cochlear implantation (i.e., at T0). Concerning vowel realization prior to cochlear implantation, as found by Schenk et al. (2003), we observed increased first vowel formants (F1) in the CI group (compared to controls). This F1 increase was vowel specific: The mid-open vowels [ɛ] and [ɔ] were produced with higher F1 and thus more open in the CI group, compared to the controls. The results for the second vowel formant (F2) suggest a vowel centralization tendency for the CI group prior to implantation. Especially the back vowel [u] was produced more fronted in the CI group than in the control group. Contrary to our expectations, we did not find differences in average vowel space due to hearing status. In sum, although the speech of all CI participants included in our study was rated as intelligible after a longer period of profound hearing loss, we have found reduced sibilant contrasts and subtle deviations in vowel characteristics for post-lingually deafened adults prior to CI surgery, in comparison to controls.

Concerning our second research question on the effect of cochlear implantation, we found increases in the sibilant contrast both for the CI as well as for the control group over test sessions. This corroborates our assumption that repeated testing affects speech realization (hyperspeech). A more interesting finding is that the direction of the spectral changes in sibilant acoustics differed for control and CI group. The participants in the control group predominantly shifted the COGs of their [s] production upwards for later test sessions (T1, T2), whereas CI participants' increased sibilant contrasts mainly resulted from a decrease in COGs of the [ʃ] productions.

We suggest that the control group shifted both sibilants' COGs in opposite directions (for [s] upwards, for [ʃ] downwards), following a symmetrical approach to hyperspeech. The differential production patterns of the CI group, however, can be explained in different ways. One possible account may be that frequency resolution of the cochlear implant device is reduced for high frequencies. The CI group shifted the COGs of the [ʃ] productions downwards in frequency, towards a region with narrower bandwidths and thus with better frequency resolution, to enhance the sibilant contrast. Shifting COGs for [s] productions upwards might not have been an op-tion in the CI group because the CI stimulation frequency range is limited for higher frequencies. There is an alternative ‘central processing’ explanation for the differential pattern of acoustic changes we observed in the CI group after implantation. Hearing loss often starts in the high frequency range. Mid and low frequencies are usually affected later and to a lesser extent by hearing loss (cf. Figure 5.2). Central processing capabilities for very high-frequencies in post-lingually deafened might be particularly diminished because auditory information in this frequency range has been longer absent or incomplete. Thus, after a longer period of severe hearing loss the central processing capabilities for the high frequency spectrum might not be able to recover completely or at least not within the time period after cochlear implantation investigated in this study. As such, our data suggest that the different sibilant production patterns for CI and control group at test sessions T1 and T2 reflect the differences in their auditory processing capacities. A peripheral or a central processing account (or a combination of both) may explain why CI users show the sibilant production pattern observed.

Secondly, concerning the effect of cochlear implantation on vowels, F1 decreased after cochlear implantation. This may be interpreted as a recovery due to restored hearing as this F1 increase was not found for the control group for repeated testing. Our analyses further yielded a vowel-specific recovery effect due to cochlear implantation for F2. The back vowels [u] and [ɔ] were produced less centralized and more prototypically after cochlear implantation compared to their production prior to CI surgery.

No evidence was found for a supposedly detrimental effect of cochlear implantation on sibilant or vowel acoustics in the early phase after implantation (approx. two weeks after CI activation) as compared to performance before implantation. This result is in conflict with the hypothesis of temporarily reduced articulation precision due to the re-tuning of feedforward commands after CI in post-lingually deafened adults (Lane et al., 2007). Our results suggest that perceptual adaptation and the resulting changes in production may not take as long as some earlier studies have suggested (e.g., Langereis et al., 1997).

In line with studies that investigate predictors of speech perception recovery after CI, we have found that sibilant production recovery after CI was related to duration of hearing loss: the longer the period of hearing loss before CI, the less improvement in the participants’ speech outcome after CI. This suggests that cell loss and processing changes in the central auditory system due to long-term hearing loss also affect the speech production system. The remapping of the new input via the CI to existing sound categories may be more successful if speech perception and with it the feedback control system suffered from shorter periods of auditory deprivation.

Our study has clearly demonstrated the need for a gender-matched control group, and for a similar test regime for test and control group. Future research should preferably avoid repetition of materials, or, if possible, include more filler material to hide the sound contrasts of interest. Ideally, future studies should also move away from reading aloud, and move to the use of a structured interview or dialogue task to get a more reliable impression of patients' spontaneous speech performance. Repetition effects may have obscured this studies' and previous studies' results, since we have demonstrated that repeated reading of target words embedded in carrier sentences lead to greater spectral sibilant contrasts in both the CI group and the controls.

To sum up, our results confirmed that hearing loss affects the acoustic realization of sibilants and vowels in post-lingually deafened adults. This implies that long-term speech sound representations gradually degrade if the speech production system is disconnected from feedback via auditory perception. Our results indicate that the speech production system quickly adapts to the restored auditory feedback via the CI. Importantly, however, we have not found evidence for a hypothesized initial retuning phase early after CI during which sound contrasts would be diminished. The combined pattern of results illustrates the plasticity of speech production across the adult life span. Speech sound representations are subject to changes due to longer periods hearing loss but may recover quite quickly once hearing is restored.

Summary & Conclusions

In this thesis I have studied effects of age and hearing loss on speech production and perception. The two studies in the first part of this thesis investigated whether and how age and age-associated changes in hearing and cognitive performance affect listening in ecologically valid conditions. Part two of this thesis contains two studies examining the effects of age and hearing loss on adult speech production. The present chapter first summarizes the main findings of this thesis and then discusses them in a broader context. Lastly, suggestions for future research will be made.

Chapter 2

Chapter 2 investigated the effect of increased speech rate on speech processing across the adult age range. A word-recognition experiment was set up using the visual-world paradigm to test whether the previous result that older adults show stronger speech rate effects than younger adults (e.g., Wingfield, 1996; Gordon-Salant and Fitzgibbons, 1999; Gordon-Salant et al., 2014) can be replicated using conversational speech materials (cf. Tucker and Ernestus, 2016) with a natural variation in speech rate. Previous studies showing that older adults are differentially impacted by increased speech rate mostly used artificial time-compression algorithms to modify speech rate. However, time-compressed speech is different and is processed differently from speech that is spoken fast (Janse, 2004), raising the question as to whether results obtained with artificial manipulations of speech rate generalize to more ecologically valid materials. Participants listened to short question-answer dialogues with speech rates ranging between three and eleven syllables per second and indicated with a mouse click which out of four words on a computer screen they heard in the dialogue. In addition, individual hearing thresholds, vocabulary size and a set of cognitive predictors, such as information processing speed, reasoning and working memory, served as predictors to systematically evaluate whether these participant characteristics are associated with the impact of speech rate on the processing of conversational materials (cf. Wingfield et al., 2006; Janse, 2009). Three different dependent variables were analyzed: click-response times, as well as the time-continuous measures of gaze direction and pupil size. Pupil size was included to serve as an index of listening effort (cf. Zekveld et al., 2013; Ohlenforst et al., 2017; Wagner et al., 2016). In line with our expectations, increased speech rate made word recognition more challenging, as indicated by longer response times and delayed eye gaze behavior to the target word. Furthermore, we found consistent age effects on all three dependent variables: Older adults showed slower click responses, a slightly delayed gaze behavior and a decreased amount of tasked-evoked pupil dilation compared to younger adults. However, there was no evidence that older listeners were affected differentially by higher speech rates than younger listeners. Although none of the participant characteristics predicted the size of the speech rate effect, speech processing was generally facilitated for participants with a larger vocabulary and with better fluid cognitive processing.

Three factors may explain why age or individual participant characteristics were not associated with the speech rate effects in our study. First, in contrast to the majority of previous studies, stimulus materials did not contain any acoustic artefacts, as we did not artificially speed-up our materials but extracted stimuli, ranging in speech rate, from a corpus. Second, our materials did not entail higher-than-typical speech rates. A number of previous studies (e.g., Janse, 2009; Gordon-Salant and Fitzgibbons, 1999) presented materials with relatively high speech rates, making use of artificial speeding techniques. These time-compression algorithms can manipulate speech signals resulting in intelligible signals at speech rates hardly or never found in natural speech. Third, our adult sample had been selected such that none of the participants was eligible for hearing aids, i.e., even though some showed mild hearing loss, their degree of hearing loss was such that they did not qualify to get partial financial compensation from their health insurance for hearing aids. This combination of choice of participant sample and naturalistic rate variation may account for the finding that neither age, nor hearing or cognitive capabilities were found to modulate effects of speech rate for conversational materials.

Chapter 3

The study in Chapter 3 employed the Acceptable Noise Level Test (ANL; Nábělek et al., 2006), a test applied in hearing rehabilitation, which has been suggested to relate to prospective hearing aid use (Nábělek et al., 2006; Wu et al., 2016; but cf. Olsen and Brännström, 2014). During the test, participants repeatedly indicate the level of noise they are willing to tolerate while following a speech signal played at their most comfortable loudness level. It has been argued that the subjective ability or willingness to accept noise while listening to speech may be an important factor for hearing aid uptake. Hearing aid candidates who tolerate more noise while listening to speech may benefit more from their devices, which in turn could influence their willingness to regularly use the amplification technology. Further, ANL may also measure aspects of listening effort, a concept which has lately received much research attention (cf. Ohlenforst et al., 2017).

The rationale behind the present investigation was to study ANL results systematically for a range of speech materials. Whereas there is ample evidence that semantic context information facilitates the recognition of sentences in noise (e.g., Kalikow et al., 1977), the effect of semantic and discourse context on ANL scores is less clear (cf. Gordon-Hickey and Moore, 2008; Brännström et al., 2012). It is also unclear whether, or to what extent, ANL scores relate to indices of cognitive listening effort. Lastly, for ANL to be a useful clinical tool, ANL test-retest reliability should be high. Our study was also set up to investigate whether ANL reliability differed across speech materials used for the ANL procedure.

To investigate whether ANL scores differ across speech materials, three materials were used: 1. The ISTS material (International Speech Test Signal; Holube et al., 2010), which is an unintelligible speech-like signal, 2. concatenated standard audiology sentences (meaningful, but no coherence above sentence level; cf. Versfeld et al., 2000), and 3. stretches of conversational speech extracted from a Dutch corpus (meaningful, coherence at sentence level and above; cf. van Son et al., 2008). We tested whether ANL results differed for conversational materials compared to concatenated standard audiology sentences, and compared to the meaningless ISTS material. The hypothesis was that conversational materials would yield the lowest ANL of all three test materials (i.e., increased noise acceptance), because top-down processing and context effects should be strongest for materials that allow for semantic integration at the sentence level and above. As argued in the previous paragraph, we also asked whether ANL test reliability would differ across the test materials. In order to investigate the relationship between ANL and cognitive listening effort, we tested whether working memory and ANL scores were correlated (cf. Brännström et al., 2012). Furthermore, we assessed whether ANL scores related to listening effort by investigating the relation between ANL and scores on a hearing-related questionnaire targeting self-reported activity limitations for listening to speech-in-noise in everyday situations.

ANLs were repeatedly measured for a sample of normal-hearing participants whose age was representative of that of a sample of hearing aid users (M^{age}=60.7 yrs.). Our results show that participants tolerated more noise when listening to stretches of conversational speech or concatenated standard audiology sentences compared to the unintelligible ISTS signal. The level of noise the participants accepted was comparable for concatenated standard audiology sentences and conversational corpus speech. To sum it up, ANL results were affected by meaningfulness of the speech material, but not by its semantic coherence above sentence level.

Thus, partly in line with our hypothesis, participants accepted less noise if the meaningless ISTS material was presented compared to meaningful speech (conversational material, concatenated standard audiology sentences). This result suggests that top-down processing affected ANL. With this in mind, the finding that concatenated standard audiology sentences and conversational materials yielded comparable ANLs is puzzling: Participants did not benefit from the additional discourse context that was available for the conversational materials.

Regardless of which test material was used for ANL testing, ANL scores were not found to be associated with working memory in our normal-hearing sample. Reliability of the ANL outcome was comparable but relatively poor, across the different test materials, challenging the practical use of the test methodology in clinical settings. Further, correlation analyses suggest that ANL scores obtained with conversational materials may be a better predictor of activity limitations related to listening effort in noise than ANL scores obtained with concatenated standard audiology sentences or meaningless test materials.

Chapter 4

Chapter 4 presents the first of two studies investigating the effect of age and hearing loss on speech production. In Chapter 4, we asked whether age, age-related hearing loss and cognitive abilities affected sibilant production. Firstly, adult age may have effects on the stability of speech motor control (cf. Krampe, 2002; Mefferd and Corder, 2014; Tremblay et al., 2013; Bilodeau-Mercure et al., 2014), with fine motor control being particularly important for the production of sibilant fricatives. Relatedly, advanced adult age also affects cognitive control. Possibly, advanced age may therefore impact sibilant production through effects of age-related cognitive control on the fine motor control required for sibilant production. Age-related cognitive decline could also contribute to changes in speech production because of age effects on auditory short-term memory, which may be indispensable for the comparison of one's own speech with long-term speech representations. Furthermore, hearing loss (be it age-related or not) may change long-term representations of speech sounds due to missing or altered auditory feedback and/or degraded auditory information about speech of others.

In order to investigate the effect of age, age-related hearing loss and cognitive abilities on sibilant production, participants across the adult age range read carrier sentences that contained target words with the sibilant onsets [s] and [ʃ], respectively. The dependent variable was the sibilants' first spectral moment, or Center of Gravity (COG), indexing the 'sharpness' of the sibilant sound. For all participants hearing thresholds, working memory capacity, information processing speed and reasoning capability were assessed.

Our analyses show that age by itself did not affect the sibilants' Center of Gravity. Further, none of the cognitive abilities affected the dependent variable. However, individual high-frequency hearing-loss at 8 kHz was associated with the COG for the sibilant [s]. Note that age-related hearing loss usually starts off and is most pronounced in the high frequencies and that COGs for the sibilant [s] usually range around 7 kHz (cf. Jongman et al., 2000; for English sibilants). The finding that the sharpness of a speaker's [s] production is related to the speaker's high-frequency hearing acuity but not speaker age indicates that speech production may undergo subtle changes due to age-related changes in hearing prior to the onset of severe hearing loss.

Chapter 5

Chapter 5 investigated effects of severe hearing loss and hearing rehabilitation on speech production acoustics in deafened participants. The study followed up on the findings of Chapter 4, where subtle changes in sibilant production were observed even within a sample of adults whose hearing ranged from normal hearing to moderate hearing loss. By means of cochlear implantation, hearing can be restored in deaf adults via direct stimulation of the cochlear nerve. Hence, post-lingually deafened CI candidates are an ideal test population to investigate both the long-term effect of hearing loss on speech acoustics, as well as the effect of hearing restoration on speech acoustics after cochlear implantation. CI candidates were asked to repeatedly produce target words containing the vowels [aː, ɛ, i, ɔ, u] and the sibilants [s] and [ʃ]. Both vowels and sibilants served as targets sounds for the study because severe hearing loss does not only affect the high-frequency range important for the discrimination of sibilants, but also the lower frequencies (below 3 kHz), which characterize the vowel spectrum. CI candidates were tested before the operation, shortly after activation of the CI (i.e., at the first clinical follow-up after operation) and at a clinical follow-up three months after CI activation. Vowel formants (F1, F2) and sibilants' COGs were analyzed longitudinally across the three test sessions. In addition, an age and gender matched control group was tested on the same stimuli at three test sessions equally spaced in time.

Acoustic analyses show that prior to CI surgery, sibilant contrasts were diminished for the deafened patients in comparison to the control group. Vowel formant results showed a similar trend towards centralization and diminished contrast for the CI patients compared to the controls, which is in line with previous results (e.g., Waldstein, 1990). This implies that severe long-term hearing loss affects speech production, with deviations being particularly evident in sibilant acoustics.

Across test sessions, both the CI group and the controls enlarged the acoustic sibilant contrast, which may be interpreted as hyperarticulation due to repetition of target words. However, these enlarged sibilant contrasts at later test sessions were achieved differently by the two groups. Control group speakers sharpened their [s] productions by increasing spectral means (COGs) of the [s] productions and also hyperarticulated their [ʃ] productions by decreasing spectral means for [ʃ] productions across test sessions. The novice CI users, however, enlarged the sibilant contrast mainly by lowering the spectral mean of the post-alveolar sibilant [ʃ]. Their [s] productions were relatively unaffected by the cochlear implant, which was interpreted by referring to the limitations of the signal transmitted by the CI. Interestingly, participants with shorter durations of hearing loss improved their sibilant contrast more after CI than participants with longer periods of deafness. Thus, prolonged deprivation of auditory input and hence deprivation of auditory feedback may impair phonemic contrasts in post-lingually deafened speakers with negative consequences for the restoration of the sound contrasts after cochlear implantation.

Theoretical implications

The findings summarized above have important implications for several current theoretical debates such as on models of speech perception and on the link between speech perception and speech production.

Models of speech perception

Most models of speech perception largely ignore the variation between language users in human language processing (cf. Levinson, 2012). Psycholinguistic studies using idealized test conditions (such as noise-free recordings of carefully read speech) and uniform test populations (well-educated university students, in which female participants are often overrepresented) may yield experimental data that neither resemble ecologically valid listening situations nor approximate the capabilities of representative listeners. This thesis has attempted to identify whether perceptual, cognitive and linguistics skills predict speech perception for a range of materials and listening situations, including conversational speech, higher speech rates and listening to speech in noise. If we can identify individual predictors of word processing, we can refine models of word recognition by indicating which processing parameters may be different among listeners. Farris-Trimble and colleagues (2014), for example, have been able to show that hearing loss and simulated hearing loss result in delayed lexical activations in a word recognition experiment using the visual-world paradigm. Similarly, using the eyetracking methodology we investigated the effect of speech rate on the time course of spoken word recognition in individual listeners in Chapter 2.

Although the results in Chapter 2 did not indicate that age, hearing loss or cognitive performance were related to the size of individual speech rate effects on perception, vocabulary size and fluid cognitive processing ability were found to relate to general word recognition performance. The finding of vocabulary size ties in with other studies investigating predictors of listening performance across a range of listening situations (McAuliffe et al., 2013; Bent et al., 2016; Avivi-Reich et al., 2015), including listening to dysarthric speech, speech in multitalker babble, and accented speech. Additionally, using a comparable setup as in Chapter 2, Carrol and colleagues (2017) investigated non-native processing of conversational speech and corroborated the finding that vocabulary size predicts word recognition in L2.

How can we incorporate these findings in a competition-based speech recognition model such as for example Shortlist A (Norris, 1994)? In this model, when a listener encounters a word initial sound [f], a list – the Shortlist - containing up to 30 lexical entries starting with this sound is automatically activated (e.g. ‘for’, ‘from’, ‘first’, ‘find’, ‘feel’, ‘family’, ‘four’, ‘few’, etc.). A larger vocabulary may either entail a more complete (i.e., exhaustive) or better structured list of candidates at this initial stage already. Alternatively, those listeners with larger vocabulary sizes may have more precise phonological representations in long-term memory (cf. Schmidtke, 2014, 2016). Having these more precise phonological representations facilitates the competition between activated candidates, as the degree of match or mismatch can be established in greater detail, which helps in singling out the best match. As such, vocabulary size may also play a role in determining word boundaries during competition. A more efficient singling out of the best match may also help to activate or inhibit alternative segmentation options, which prevents erroneous parallel processing.

Vocabulary knowledge may also be a proxy for language knowledge in general. Listeners with larger vocabularies may have a better knowledge of the frequencies with which words occur and co-occur, and may be better able to predict which words could follow in a given sentence or discourse context (e.g., van Berkum et al., 2005). These regularities may serve as priors, adjusting activation levels of following candidates. Shortlist A (Norris, 1994) does not integrate semantic information above the word level as a sort of pre-activation for spoken word recognition. Also Shortlist B, the Bayesian Shortlist version (cf. Norris and McQueen, 2008), only includes word frequencies to estimate prior word probabilities, although Norris and McQueen state that “word frequency represents only a fraction of the knowledge that listeners have at their disposal when recognizing continuous speech” (ibid., p. 359). A well-structured vocabulary, or better-structured language knowledge in general, may thus facilitate spoken word selection by adjusting the a-priori activation level (or prior probability; cf. Shortlist B, Norris and McQueen, 2008) of candidates in the Shortlist (cf. Huettig, 2015) based on sentence and discourse level information.

Apart from vocabulary effects, Chapter 2 showed a relationship between fluid cognitive processing ability and spoken-word recognition. As stated above, participants with better performance on this measure, which combines processing speed and speeded reasoning abilities, were faster at recognizing the target words in conversational speech. Should this best be modelled as the speed with which representations are activated or spread their activation? Generalized slowing (e.g., Birren, 1965; Salthouse, 1985), a symptom of age-related cognitive decline, has been proposed to negatively impact on speech processing in older adults (cf. Thornton and Light, 2006). Relatedly, TRACE simulations for visual-world word recognition data suggest that generalized slowing may account for individual differences in gaze behavior between (adolescent) controls and adolescents with cognitive and/or linguistic deficits (cf. McMurray et al., 2010).

The simulation work of McMurray and colleagues (2010) points towards cognitive capabilities already affecting relatively early processes of word recognition, namely the activation stage at the prelexical level, when phonemes or sub-phonemic units are activated based on their match with the available acoustic input. Spreading of activation from the prelexical level to the lexical level is assumed to be automatic. To date it is unclear as to whether fluid cognitive processing abilities may affect the speed of activation spreading. In principle, the mentioned simulation results (McMurray et al., 2010) do not exclude the possibility that also the propagation of activation from early (prelexical) to subsequent stages through the recognition system may be affected by cognitive slowing.

The subsequent competition stage, which involves winnowing down the set of activated lexical entries while constantly processing new speech input, seems to be a plausible level for the effects of fluid cognitive processing, because dealing with multiple candidates and new acoustic input (in parallel or serially) can be assumed to be cognitively demanding. We suggest that the speed of the selection of the best matching candidate depends on the general fluid cognitive performance, analogous to the lexical-activation rate findings of McMurray and colleagues (2010). Further, if a listener's processing is fast, more resources may become available for prediction of subsequent speech input. Fast recognition of a word may result in speeded activation of related words (increased prior probabilities). In other words, as soon as the word 'farm' is recognized, this may aid in rapidly recognizing the next word 'animals' because co-occurring words such as 'barn', 'corn' or 'animals' receive activations early on. In this way fluid cognitive processing may have a cascading positive effect on word recognition. Furthermore, Wagner and colleagues (2016) assume that "well-timed progress of information from sensory to pre-lexical and lexical stages of processing" (ibid., p. 1) is necessary for the seemingly effortless process of understanding speech. Consequently, if processing speed or fluid cognitive processing is limited, speech processing may become more effortful.

The study in Chapter 3 contributes to the discussion of whether and how adults pay attention to meaning of the spoken input they are presented with. Using the ANL procedure, we have found that more noise was accepted while participants followed meaningful materials as compared to following a meaningless speech-like material. This tolerance of a noisier signal for meaningful materials suggests that participants were attempting to actually comprehend the speech, also in the non-meaningful condition.

Even though the ISTS material was meant to represent unintelligible speech, upon close inspection of the ISTS signal, I identified some English, French and German words which participants may have recognized as well. Some example words include e.g., English ‘attempt’ and ‘around him’, German “nehmen” and “sollte”, and French “alors” and “chaque”. Recognizing these words from languages that participants might have been somewhat familiar with may have caused (Dutch) participants to listen more actively, and more analytically, than they would do listening to their native language, resulting in lower noise acceptance. In other words, they may have acted as if they turned down the volume of a noisy radio playing in the background in order to better follow a speaker that only from time to time produced intelligible snippets of speech (e.g., as if this were a poorly intelligible speaker with an unfamiliar dialect or L2 accent).

Regarding context effects on ANL I had hypothesized that listeners would accept more background noise for conversational materials compared to concatenated standard audiology sentences. The conversational passages presented a discourse context, which was expected to facilitate comprehension, whereas the concatenated sentences materials did not offer text cohesion. The hypothesis of discourse context facilitation was not confirmed in our study.

In line with our original hypothesis, Bentley and Ou (2017) found that less noise is accepted for concatenated standard audiology sentences (QuickSIN material, Killion et al., 2004) compared to the Arizona Travelogue signal, a longer passage of *read* travel descriptions. Note, however, that the QuickSIN sentences in their study were concatenated without a silent gap, whereas we included 500 ms silent gaps between sentences. The additional gaps in our sentence material may have provided our listeners with extra processing time making the task to follow the speech easier for the concatenated sentence material. This may explain why our study did not show benefits for noise tolerance of conversational materials compared to concatenated audiology sentences. Another explanation for the null result regarding context effects on noise tolerance in our study is that the relatively fast speech rate and less careful articulation (cf. Tiffin and Gordon-Hickey, 2017; Recker and Micheyl, 2017) observed in our conversational material (compared to the read out Arizona Travelogue signal) may have cancelled out coherence effects in our data, resulting in comparable noise acceptance for conversational materials and concatenated sentences, at least for the participants in our study who had normal hearing.

Models of speech production

In this thesis we asked whether and how age, cognitive aspects of aging, and mild forms of age-related hearing loss affect speech production as reflected in sibilant acoustics (Chapter 4). Furthermore, Chapter 5 investigated whether severe hearing loss affects vowel and sibilant acoustics and whether restored hearing would also impact on speech acoustics.

The results of both production studies (Ch. 4 and Ch. 5) provide empirical support for models of speech production which postulate an outstanding role of auditory feedback for the representation of speech in adults. Hearing loss was found to relate to sibilant acoustics in adult speakers and hearing restoration after a longer period of deafness lead to sibilant contrast improvements. This suggests that speech sound representations are not set in stone but rather interact with the auditory feedback one gets of one's own speech and/or with the auditory input of others (cf. Houde and Jordan, 1998; Purcell and Munhall, 2006; Villacorta et al., 2007; Lametti et al., 2012; Schuerman, 2017). Relatedly, Lyxell and colleagues (1998) have found that adults with severe to profound hearing loss show deteriorated phonological representations. In contrast to Lyxell and colleagues' findings, Meyer and colleagues (2003) argue that postlingually deafened adults "maintain a central representation of language that has a structure similar to the representation they had when they heard acoustically" (ibid, p. 613). However, similar representations do not necessarily entail that loss of auditory information and auditory feedback is not reflected in speech acoustics. Our results imply that, if hearing and thus auditory feedback becomes less reliable, due to gradually increasing or sudden hearing loss, or if feedback changes drastically after cochlear implantation, speech sound acoustics change, especially for sibilants. This sound group may be susceptible to the observed effect of altered auditory feedback because sibilants require fine-motor control and their perception is most likely to be affected by age-related hearing loss as well as by cochlear implantation.

The HSFC model (e.g., Hickok, 2012) includes an auditory as well as a somatosensory feedback control system. Auditory feedback is used to compare the spoken utterance with an internal auditory target, implying that lexical units are represented in auditory space. The somatosensory feedback system monitors the movements and positions of the speech articulators. Hickok (2012) posits that once auditory targets have been instantiated during language acquisition, somatosensory feedback primarily guides the articulatory movements because it is faster than the auditory feedback loop (cf. Chen and Watson, 2017). Only if the (slower) auditory feedback informs the speaker about errors or productions that are slightly off (e.g., via quasi real-time alterations of the auditory feedback as e.g., in Houde and Jordan, 1998 or Purcell and Munhall, 2006), motor programs have to be re-adjusted to ensure that the speech output again approximates the internal auditory targets. Note, however, that Nasir and

Ostry (2008) demonstrated that deaf adults compensate for mechanical perturbation of their speech while speaking with their CI turned off, which supports the claim that the somatosensory feedback system may be particularly important for the relatively preserved intelligibility in post-lingually deafened speakers. The crucial question in this dissertation was whether advanced age and/or hearing impairment would, through their effects on auditory representations, be reflected in speech acoustics.

Our findings of a close reflection of even subtle degrees of hearing loss in speech acoustics cannot be taken to contradict the original hypothesis of Hickok (2012) that the auditory feedback system played only a minor role in adulthood. One may argue that, contrary to Hickok's (2012) hypothesis, the auditory feedback system in adults seems to remain an active component that monitors and corrects also subtler deviations from the intended representations in auditory space. Feed-forward motor commands may have become less precise due to hearing loss, because regulatory, auditory feedback was insufficient and somatosensory feedback alone could not compensate. Alternatively, one can argue that sound representations in those with hearing loss simply may have come to reflect the absence of high-frequency information in the input a speaker gets from listening to others. In other words, our results cannot decide on the balance between auditory and somatosensory feedback systems in production, but do speak to the role of hearing loss on speech sound representations.

Our studies in Chapters 4 and 5 also spoke to the time course of hearing loss or hearing restoration effects on speech acoustics. We observed a gradual change in adults' speech acoustics due to changes in hearing acuity (cf. Chapter 4) and we have seen relatively immediate effects of restored hearing on speech production in the CI participants (cf. Chapter 5). Already two weeks after activation of the cochlear implant, the CI implantees in our sample showed improvements in the sibilant contrast (cf. Chapter 5). In contrast to what Lane and colleagues had hypothesized (2007) in a similar study with CI participants, we did not observe any detrimental effect of the CI on speech acoustics. Lane et al. (2007) concluded their study with the recommendation for future studies to test CI participants at an earlier test moment than they had (i.e., 1 month after CI activation), to test whether re-tuning of the auditory feedback control system would show initial decrements or absence of early improvements in speech acoustics. The new auditory percept and the re-mapping of auditory input to feed-forward commands in our test sample at two weeks after CI activation did not negatively affect speech acoustics. This implies that even after complete deprivation of auditory information due to severe hearing loss, the speech motor control system is flexible enough to quickly accommodate to new sensory input.

In line with studies investigating predictors of speech *perception* recovery after CI (e.g., Blamey et al., 1992; Plant et al., 2016; but cf. Meyer et al., 2003), we found that the duration of hearing loss predicted the improvement of speech *acoustics* after implantation (Ch. 5). This may be because over time a lack of auditory stimulation leads to a reduction of cells in the spiral ganglion – the origin of the auditory nerve in the inner ear (cf. Syka, 2002; Meyer et al., 2003; Pelle and Wingfield, 2016; but cf. Linthicum and Fayad, 2009). Such cell loss may cascade to the central auditory system with deleterious effects on auditory processing and on auditory feedback monitoring. Nevertheless, the brain metabolism in auditory cortex has been shown to recover after cochlear implantation in post-lingually deafened adults (cf. Syka, 2002). Duration of hearing loss in adult CI users may simply determine the speed with which their auditory processing recovers. Similarly, auditory evoked potentials, in children mature with respect to the duration of hearing loss before CI implantation (cf. Syka, 2002). Crucially, duration of hearing loss may also affect speech production recovery, because auditory feedback of one's own speech and auditory information about speech of others becomes available once auditory processing is re-established via the CI.

Whereas our results showed an association between sibilant acoustics and high-frequency hearing loss, others found sibilant acoustics to be related to sibilant discrimination (Perkell et al., 2004). Possibly, hearing loss, as assessed with pure-tone thresholds, is only part of the problem that those with more hearing loss have with sibilants. Age-related hearing loss goes hand in hand with decreased auditory processing abilities, such as greater spectral smearing (Clinard et al., 2010; Wingfield et al., 2005). Future studies on the relationship between sibilant production and perception should therefore preferably assess hearing thresholds, as well as perception of the sound contrast at hand.

Conclusions and outlook on future research

Chapter 2 showed that increased rates of speech made the processing of conversational speech more challenging. However, this effect of increased conversational speech rate is equal-sized for listeners across the adult life span. Moreover, none of the individual sensory and cognitive abilities were associated with the added difficulty of the increased rate. To follow up on this finding, future research could investigate effects of speech rate on turn-taking in natural conversation among listeners of varying ages (cf. Heffner et al., 2015). Even if speaking rate effects on listening were shown to be equal across age groups, it would be informative to test whether listening to faster speech impacts on older adults' ability to quickly formulate an appropriate response (cf. Bosker, 2017). Obviously, such a study should take age differences in word finding difficulty into account as well.

The results of Chapter 3 provided some indication that the use of conversational speech in speech processing measures better reflects everyday listening experience than standard audiology materials. Importantly, however, these results were obtained with the ANL, a subjective measure of willingness to keep up with noise, which is not related to objective performance measures, such as speech-in-noise word or sentence recognition. Future research should validate the suggestion that conversational speech better reflects everyday listening experience by comparing subjective and objective testing with *hearing aid users*, rather than the normal-hearing listeners included in our study. Objective measures such as speech reception thresholds or speech in noise recognition will always be administered because they capture the amount of acoustic and linguistic information that a listener can maximally derive from spoken input, while using a hearing device or without it. Nevertheless, subjective measures may do a better job than objective measures in capturing motivational aspects of listening to speech in noise. The motivation to regularly use hearing aids is the prerequisite for hearing aid success. Using conversational speech as diagnostic material may ultimately optimize individual hearing rehabilitation if it can improve the evaluation of individual noise sensitivity and the role of this sensitivity in hearing aid uptake.

Chapter 4 showed that sibilant acoustics are associated with (age-related) high-frequency hearing loss, but not with age, in a non-clinical population. From a communicative perspective, the result that speech production is affected by hearing acuity indicates that hearing devices may not only be beneficial for (older) adults with hearing-related comprehension problems, but also for their interlocutors, because speech production accuracy may be preserved if hearing is maintained or restored. The absence of age effects in our data is in conflict with studies suggesting that older adults' speech may be characterized by deteriorated neuromuscular or speech motor control (cf. Benjamin, 1997; Krampe, 2002; Mefferd and Corder, 2014; Knuijt et al., 2017). One explanation for the absence of an age effect in our data may be that the sentences elicited in our study were not very demanding in terms of speech motor control compared to tasks that entail maximum fast rate performance or for example metronome-paced syllable repetitions (cf. Mefferd and Corder, 2014). Relatedly, the absence of an age effect in our reading task may be because such effects become only more pronounced in older-old speakers (75+ yrs.). Follow-up research should therefore preferably move away from sentence reading, and include a cognitively more demanding and more naturalistic speech elicitation task (e.g., a conversational map-task or at least reading a dialogue), if only because our results point towards hyperarticulation due to the repetition of target words. This way we could distract the speakers from the segmental content of the utterances and obtain more representative speech samples at the expense of tightly controlled test conditions.

The Chapter 5 findings of relatively rapid improvements in sound contrasts in novice CI implantees clearly raise further questions on the time course of speech production changes after restoration of hearing. A follow-up study may investigate speech acoustics after CI for a longer period of time, e.g., 12 months, to test whether long-term hearing-impaired patients eventually catch up with short-term deafened participants after a longer period of time. Additionally, much more level of fine-grained detail on the time course of possible changes in speech production could be obtained by equipping novice implantees with portable recording devices (cf. https://www.lena.org/), in order to monitor their speech from day one after CI activation onwards. This kind of more continuous data would be needed to ultimately shed light on the time course of re-mapping of the novel auditory input to motor commands for production from CI activation onwards. Furthermore, such data would also allow the investigation of the role of conversation partners on CI users' performance. CI novices may need the auditory input from their interlocutors to rapidly learn how to relate the new signal delivered by the CI to the (existing) auditory speech sound representations and to improve articulation accuracy. Being able to analyze the quality and quantity of input CI users get and being able to monitor their speech over time during conversational turn-taking would provide a much more refined view on CI rehabilitation than can be achieved with repeated (clinical) testing.

Nederlandse samenvatting

Gegeven de huidige en verwachte demografische ontwikkelingen is er een groeiende interesse in onderzoek naar veroudering en haar effecten op cognitie en communicatie. Leeftijdsgerelateerde sensorische achteruitgang, met name wat betreft gehoor, heeft duidelijk invloed op het gemak van gesproken communicatie bij ouderen. Bovendien kan een hogere leeftijd een negatieve impact op de spraakverwerking hebben door effecten van cognitieve achteruitgang. De studies in dit proefschrift onderzoeken het effect van leeftijd en gehoorverlies op de spraakverwerking. Het eerste deel van dit proefschrift (Hoofdstukken 2 en 3) richten zich op de gevolgen van veroudering en leeftijdsgerelateerd gehoorverlies op spraakwaarneming. Het tweede deel (Hoofdstukken 4 en 5) bevat twee studies die bijdragen aan de kennis over de aard van het verband tussen spraakperceptie en spraakproductie. Deze twee studies onderzoeken de effecten van leeftijd en gehoorverlies op de spraakproductie.

Hoofdstuk 2 onderzocht het effect van toegenomen spraaksnelheid op de spraakverwerking bij volwassenen. In eerder onderzoek waarin gebruik werd gemaakt van kunstmatige tempoversnelling van spraakfragmenten was gevonden dat ouderen sterkere effecten van spreeksnelheid op hun spraakverwerking laten zien dan jongeren. Een woordherkenningsexperiment dat gebruikt maakte van het ‘visual world paradigma’ werd opgezet om te onderzoeken of deze bevinding gerepliceerd kan worden door gebruik te maken van gesproken conversatiefragmenten die (op een natuurlijke manier) varieerden in het tempo waarin ze uitgesproken waren. Proefpersonen luisterden naar korte vraag-antwoord dialogen en gaven aan met een muisklik welke van vier woorden op een computerscherm ze hoorden in de dialoog. Daarnaast werden individuele gehoordrempels, woordenschat en een aantal cognitieve vaardigheden, zoals de snelheid waarmee informatie wordt verwerkt, redeneren en werkgeheugen, als voorspellers gebruikt om te testen of deze proefpersoonkenmerken geassocieerd zijn met de impact van spreeksnelheid op de verwerking van conversaties.

Drie verschillende afhankelijke variabelen werden geanalyseerd: klik-responstijden, de proportie van oogfixaties en pupilgrootte. Zoals verwacht bemoeilijkte een versnelling van de spreeksnelheid de woordherkenning. Dit was te zien aan langere responstijden en vertraagde oogbewegingen naar het doelwoord. Verder vonden we consistente effecten van leeftijd op alle drie afhankelijke variabelen: Oudere volwassenen hadden tragere klik-responstijden, ietwat vertraagde oogbewegingen en een kleinere verwijding van de pupillen vergeleken met jongere volwassenen. Echter, er was geen bewijs dat oudere luisteraars verschillend beïnvloed werden door snellere spreeksnelheid dan jongere luisteraars. Wellicht is het de combinatie van de geselecteerde groep oudere deelnemers (uitsluiting van ouderen met een pure-tonen-gemiddelde van meer dan 35 dB gehoorverlies) en de natuurlijke tempovariatie die ervoor zorgde dat noch leeftijd, noch gehoorverlies of cognitieve vaardigheden geassocieerd waren met het effect van spreektempo in deze studie. De bevinding dat spraakverwerking makkelijker gaat voor deelnemers met een grotere woordenschat en betere cognitieve verwerking ('fluid cognitive processing') suggereert dat zowel cognitieve vaardigheden als linguïstische kennis (woordenschat) mee moeten genomen bij de formulering van toekomstige woordherkenningsmodellen. Het hebben van een grotere woordenschat zorgt ervoor dat tijdens de vroege stadia van woordherkenning er een completere of beter gestructureerde set woordkandidaten beschikbaar is. Een uitgebreidere woordenschat zorgt er tijdens de latere stadia van woordherkenning wellicht voor dat de competitie tussen woordkandidaten beter of sneller verloopt en men de juiste woordgrenzen sneller gevonden heeft. Over het algemeen is het wellicht zo dat luisteraars met een grotere woordenschat beter gebruik kunnen maken van taalspecifieke regelmatigheden, zoals kennis over de frequentie van woordcombinaties, waarmee ze semantische samenhang tussen woorden ook sneller oppakken. Betere cognitieve verwerking (fluid cognitive processing) versnelt zowel de activatie- als selectiestadia van woordherkenning.

Hoofdstuk 3 maakte gebruik van de Acceptable Noise Level (ANL) test. Dit is een test die gebruikt wordt in de gehoorrevalidatie en waarvan gesuggereerd wordt dat deze het toekomstige gebruik van een gehoorapparaat voorspelt. Tijdens de test geven proefpersonen herhaaldelijk aan, terwijl ze luisteren naar een spraaksignaal dat afgespeeld wordt op hun meest comfortabele luidheidsniveau, welke hoeveelheid ruis ze nog acceptabel vinden. Op deze manier geeft de ANL een indicatie van het negatieve effect van ruismaskering dat iemand ervaart bij het luisteren naar spraak. De achterliggende gedachte van de voorliggende studie was om ANL resultaten (de hoeveelheid ruis die men nog acceptabel vindt) systematisch te onderzoeken voor een verscheidenheid aan spreekmaterialen. Drie typen materiaal werden gebruikt: 1. De ISTS (International Speech Test Signal), een onverstaanbaar signaal dat lijkt op spraak, 2. aan elkaar geplakte standaard audiologische zinnen die per zin betekenisvol zijn maar samen geen coherent verhaal vormen en 3. stukken conversationele spraak die betekenisvol zijn en een coherent verhaal vormen. Verder werd geëvalueerd welke van de drie materialen de beste test-hertestbetrouwbaarheid zou hebben. Om de relatie tussen ANL en de cognitieve luisterinspanning te onderzoeken hebben we getest of de capaciteit van het werkgeheugen en ANL scores gecorreleerd waren. Evenzo werd een gehoorgerelateerde vragenlijst gebruikt om te onderzoeken of zelf-gerapporteerde beperkingen in het luisteren naar spraak-in-ruis in alledaagse situaties geassocieerd waren met ANL. Samengevat waren de resultaten betreffende ANL afhankelijk van het type spraakmateriaal, maar niet van de semantische coherentie (boven zinsniveau). Kortom, proefpersonen accepteerden minder ruis wanneer het betekenisloze ISTS materiaal werd afgespeeld dan wanneer ze naar betekenisvolle spraak luisterden (conversaties, aan elkaar geplakte audiologische zinnen). Deze resultaten suggereren dat top-down verwerking een invloed heeft op ANL. Dat men meer ruis accepteerde suggereert dat luisteraars daadwerkelijk probeerden het gesproken fragment te begrijpen. De betrouwbaarheid van de ANL uitkomstmaat was vergelijkbaar maar relatief klein over de drie test materialen genomen. Daardoor wordt de praktische bruikbaarheid van de testmethodologie in klinische settings in twijfel getrokken. Onafhankelijk van welk testmateriaal voor de ANL test gebruikt werd, waren de scores niet geassocieerd met werkgeheugen in onze groep van normaalhorenden, hetgeen eerdere ANL-bevindingen met zinnig spraakmateriaal tegenspreekt. Verder waren er einige indicaties dat het gebruik van conversationele spraak voor ANL tests beter het alledaagse luisteren weerspiegelt dan de standaard audiologische materialen.

In Hoofdstuk 4 onderzochten we of leeftijd, leeftijdsgerelateerd gehoorverlies en cognitieve vaardigheden een invloed hadden op de productie van sibilanten. Volwassen proefpersonen lazen zinnen die doelwoorden met de sibilantonsets [s] of [ʃ] bevatten. De afhankelijke variabele was het eerste spectrale moment van de sibilant, ofwel Center of Gravity (COG), die de ‘scherpheid’ van de sibilantrealisatie aangeeft. Voor alle deelnemers werden de gehoordrempels, de capaciteit van het werkgeheugen, de snelheid van informatieverwerking en redenatievermogen getest. Ten eerste zou gehoorverlies gevolgen kunnen hebben voor de representatie van spraakklanken in het langetermijngeheugen door verslechterde of ontbrekende auditieve feedback over de eigen spraak en/of door verslechterde auditieve informatie bij het luisteren naar de spraak van anderen. Ten tweede heeft leeftijd vermoedelijk een effect op de stabiliteit van de spraakmotorische controle en deze controle is vooral van belang bij het produceren van sibilanten. Daarnaast heeft voortschrijdende leeftijd ook effecten op cognitieve controle. Leeftijdsgerelateerde cognitieve controle zou kunnen bijdragen aan veranderingen in spraakproductie door leeftijdseffecten op auditief kortetermijngeheugen, dat een belangrijke rol speelt in het vergelijken van je eigen gerealiseerde spraak met de opgeslagen klankrepresentaties. Onze analyses laten zien dat leeftijd op zichzelf geen effect had op het Center of Gravity van de sibilanten. Verder had geen van de cognitieve vaardigheden een effect op de afhankelijke variabele. Echter, individueel gehoorverlies van hoge frequenties op 8 kHz was geassocieerd met de COG van de sibilant [s]. De bevinding dat de scherpte van de productie van [s] samenhangt met gehoorscherpte voor hoge frequenties van de spreker maar niet met de leeftijd van de spreker geeft aan dat spraakproductie subtiele veranderingen kan ondergaan door leeftijdsgerelateerde veranderingen in gehoorscherpte voorafgaand aan het begin van ernstig gehoorverlies. Feedforward motorcommando’s kunnen minder ‘precies’ geworden zijn onder invloed van gehoorverlies omdat de regulerende rol van auditieve feedback onvoldoende was en men niet voldoende had aan alleen somatosensorische feedback. Een alternatieve verklaring zou zijn dat de klankrepresentaties van personen met gehoorverlies eenvoudigweg een weerspiegeling vormen van de afwezigheid van hoogfrequente informatie in de spraakinput zoals deze persoon die waarneemt in het luisteren naar de spraak van anderen. Onze resultaten kunnen niet precies aangeven hoe groot de invloed van het luisteren naar eigen spraak in vergelijking met het luisteren naar spraak van anderen voor de opgeslagen klankrepresentaties is, maar bevatten wel belangrijke informatie over de algemene rol van gehoorverlies op spraakrepresentaties.

In Hoofstuk 5 werden post-linguaal dove CI kandidaten als testpopulatie genomen om het effect van langdurig gehoorverlies op de spraakakoestiek te onderzoeken, evenals het effect van gehoorherstel op spraakakoestiek na het plaatsen van een cochleair implantaat. CI kandidaten werd gevraagd om herhaaldelijk doelwoorden te produceren die de klinkers [aː, ɛ, i, ɔ, u] en de sibilanten [s, ʃ] bevatten. CI kandidaten werden getest voor de cochleair implantaatoperatie, kort na activering van de CI en tijdens een klinische vervolgafspraak drie maanden na CI activatie. De klinkerformanten en de COG's van de sibilanten werden geanalyseerd over de drie testsessies heen. Verder werd een controlegroep die in leeftijd en geslacht gematcht was getest op dezelfde stimuli in drie test sessies met evenveel tijd tussen de sessies. De akoestische analyses laten zien dat voorafgaand aan de CI operatie, de contrasten voor sibilanten kleiner waren voor de CI-kandidaten in vergelijking met de controlegroep. De resultaten met betrekking tot de klinkerformanten toonden een tendens van centralisatie en een verkleind klinkercontrast voor CI patiënten vergeleken met de controlegroep. Dit impliceert dat ernstig en langdurig gehoorverlies invloed heeft op de spraakproductie, met afwijkingen die met name evident zijn in de akoestiek van sibilanten. Over de testsessies genomen vergrooten zowel de CI groep als de controlegroep het akoestische contrast voor sibilanten. Echter, deze vergrote contrasten voor sibilanten in latere testsessies werden op een verschillende manier bereikt door de twee groepen. De sprekers uit de controlegroep verscherpten de productie van hun [s] door het spectrale gemiddelde (COGs) van de [s] producties te verhogen. Eveneens overarticuleerden zij hun [ʃ] producties door het spectrale gemiddelde voor [ʃ] over testsessies heen te verlagen. De beginnende CI gebruikers vergrooten echter voornamelijk het sibilantcontrast door het spectrale gemiddelde van de post-alveolaire [ʃ] te verlagen. Hun producties van [s] waren relatief gezien niet beïnvloed door het cochleaire implantaat. Deze bevinding werd toegeschreven aan de beperkingen van het CI-signaal. In tegenstelling tot de eerder geopperde verwachting dat cochleaire implantatie in eerste instantie spraak zou verslechteren hebben wij vrijwel onmiddellijke positieve effecten gevonden van hersteld gehoor op spraakakoestiek bij de CI-deelnemers. Al na twee weken nadat hun CI geactiveerd was lieten de CI-deelnemers een verbetering zien van hun sibilantcontrast. Deze snelle verbetering houdt in dat, zelfs na compleet verlies van auditieve informatie als gevolg van ernstig gehoorverlies, het spraakmotorisch systeem flexibel genoeg is om zich snel weer aan te passen aan het nieuwe sensorische aanbod. Een verdere interessante bevinding was dat deelnemers die minder lang ernstig gehoorverlies hadden gehad hun sibilantcontrast méér verbeterden na CI dan deelnemers met langere perioden van doofheid. Dus, aanhoudende ontbering van auditieve input en daarmee ontbering van auditieve feedback kan fonemische contrasten verminderen in postlingual doof geworden sprekers, met negatieve consequenties voor het herstel van klankcontrasten na het inbrengen van het cochleaire implantaat.

Dit proefschrift heeft aangetoond dat de effecten van leeftijd en spreeksnelheid op de spraakperceptie getest kunnen worden door gebruik te maken van conversationele spraakfragmenten (uit een corpus met spontane dialogen). Deze materialen zijn een betere weerspiegeling van alledaagse luistersituaties dan materiaal dat zorgvuldig en hardop voorgelezen is. Het eerder gevonden resultaat dat oudere luisteraars sterkere effecten van spreeksnelheid op hun spraakverwerking dan jongere luisteraars laten zien kon niet worden gerepliceerd. Dit suggereert dat ofwel onze keuze van deelnemers of het stimulusmateriaal een invloed hebben gehad op de uitkomst van ons experiment. Omdat geen van onze proefpersonen in aanmerking kwamen voor een gehoorapparaat, duidt het ontbreken van een significant interactie tussen leeftijd en spreeksnelheid op (leeftijdsgerelateerd) gehoorverlies en laat zien dat het niet perse leeftijd is dat verantwoordelijk is voor de eerdere bevinding dat oudere luisteraars verschillend beïnvloed worden door spreeksnelheid dan jongere luisteraars. Het feit dat onze studie geen interactie liet zien tussen leeftijd en spreeksnelheid op spraakwaarneming, waar eerdere studies dat wel hadden laten zien, lijkt erop te wijzen dat gehoorverlies een rol speelt in die interactie. Aangezien geen van onze deelnemers in aanmerking kwam voor een hoortoestel heeft onze studie wellicht geen kans gehad die interactie te vinden door de opgelegde relatief beperkte spreiding in gehoorverlies.

De resultaten van beide productiestudies leveren steun voor modellen van spraakproductie die een prominente rol voor auditieve feedback en auditieve input voor de representatie van spraak in ouderen aannemen. Ten eerste werd gehoorverlies bij hoge frequenties, maar niet leeftijd, bevonden gerelateerd te zijn aan de akoestische realisatie van sibilanten in gezonde volwassen sprekers. Ten tweede verkleinde een langere periode van doofheid de sibilantcontrasten in een groep van kandidaten voor een cochleair implantaat. Ten derde verbeterden de sibilantcontrasten in deze groep na het herstel van het gehoor door het CI. Daarbij werd bevonden dat de verbetering na het aanbrengen van de cochleair implantaat gerelateerd was aan de duur van gehoorverlies voorafgaand aan het aanbrengen van het CI van de proefpersonen. Concluderend kan worden gezegd dat de resultaten in dit proefschrift de rol van gehoorverlies in de representatie van spraakklanken benadrukken. Daarnaast vormen ze een stimulans voor verder onderzoek naar de rol van het luisteren naar eigen spraak en die van het luisteren naar spraak van anderen met betrekking tot de opgeslagen klankrepresentaties.

References

Adank, P., Smits, R., and van Hout, R. (2004a). A comparison of vowel normalization procedures for language variation research. *J. Acoust. Soc. Am.*, 116(5):3099–3107. https://doi.org/10.1121/1.1795335.

Adank, P., van Hout, R., and Smits, R. (2004b). An acoustic description of the vowels of northern and southern standard Dutch. *J. Acoust. Soc. Am.*, 116(3):1729–1738. https://doi.org/10.1121/1.1779271.

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. *Int. J. Audiol.*, Suppl.2(47):S53–S71. https://doi.org/10.1080/14992020802301142.

Akeroyd, M., Guy, F. H., Harrison, D. L., and Suller, S. L. (2008). A factor analysis of the SSQ (Speech, Spatial, and Qualities of Hearing Scale). *Int. J. Audiol.*, 53(2):101–114. https://dx.doi.org/10.3109/14992027.2013.824115.

Allopenna, P. D., Magnuson, J. S., and Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *J. Mem. Lang.*, 38(4):419–439. https://doi.org/10.1006/jmla.1997.2558.

Amieva, H., Ouvrard, C., Giulioli, C., Meillon, C., Rullier, L., and Dartigues, J. F. (2015). Self-reported hearing loss, hearing aids, and cognitive decline in elderly adults: A 25-year study. *J. Am. Geriatr. Soc.*, 63(10):2099–2104. https://doi.org/10.1111/jgs.13649.

Andreassi, J. L. (2000). *Psychophysiology. Human behaviour and physiological response*. Lawrence Erlbaum, Hillsdale, NJ.

Assmann, P. and Summerfield, Q. (2004). The perception of speech under adverse conditions. In Birren, J. E., Schaie, K. W., and Warner, K., editors, *Speech processing in the auditory system*, pages 231–308. Springer-Verlag. https://doi.org/10.1007/0-387-21575-1_5.

Avivi-Reich, M., Jakubczyk, A., Daneman, M., and Schneider, B. A. (2015). How age, linguistic status, and the nature of the auditory scene alter the manner in which listening comprehension is achieved in multitalker conversations. *J. Speech Lang. Hear. R.*, 58(5):1570–1591. https://tinyurl.com/AviviReichETAL2015.

Aylett, M. and Turk, A. (2004). The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Lang. Speech*, 47(1):31–56. https://doi.org/10.1177/00238309040470010201.

Baayen, R. H., Piepenbrock, R., and Rijn, H. (1993). The CELEX Lexical Database, Linguistic Data Consortium (University of Pennsylvania). [Electronic resource].

Baker, R. E. and Bradlow, A. R. (2009). Variability in word duration as a function of probability, speech style, and prosody. *Lang. Speech*, 52(4):391–413. https://doi.org/10.1177/0023830909336575.

Baltes, P. B. and Lindenberger, U. (1997). Emergence of a powerful connection between sensory and cognitive functions across the adult life span: A new window to the study of cognitive aging? *Psychol. Aging*, 12(1):12–21. https://doi.org/10.1037//0882-7974.12.1.12.

Barr, D. J. (2008). Analyzing 'visual world' eyetracking data using multilevel logistic regression. *J. Mem. Lang.*, 59(4):457–474. https://doi.org/10.1016/j.jml.2007.09.002.

Bassuk, S. S., Glass, T. A., and Berkman, L. F. (1999). Social disengagement and incident cognitive decline in community-dwelling elderly persons. *Ann. Intern. Med.*, 131(3):165–173. https://doi.org/chqm.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Software*, 67(1):1–48. https://doi.org/10.18637/jss.v067.i01.

Bates, D., Maechler, M., Bolker, B., and Walker, S. (2013). *lme4: Linear mixed-effect models using Eigen and S4 (R package version 1.0-5)*. https://tinyurl.com/RLME4-package.

Bates, D., Maechler, M., Bolker, B., and Walker, S. (2014). *lme4: Linear mixed-effect models using Eigen and S4 (R package version 1.1-7)*. https://tinyurl.com/RLME4-package.

Baumeister, R. F., Stillwell, A. M., and Heatherton, T. F. (1994). Guilt: An interpersonal approach. *Psychol. Bull.*, 115(2):243–267. https://doi.org/10.1037%2F0033-2909.115.2.243.

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychol. Bull.*, 91(2):276–292. https://doi.org/10.1037/0033-2909.91.2.276.

Ben-David, B. M., Chambers, C. G., Daneman, M., Pichora-Fuller, M. K., Reingold, E. M., and Schneider, B. A. (2011). Effects of aging and noise on real-time spoken word recognition: Evidence from eye movements. *J. Speech Lang. Hear. R.*, 54(1):243–262. https://doi.org/10.1044/1092-4388(2010/09-0233).

Bender, R. A., Niebuhr, D. P., Getta, J. P., and Anderson, C. V. (1993). Longitudinal study of hearing-aid effectiveness. I: Objective measures. *J. Speech Hear. Res.*, 36(4):808–819. https://doi.org/10.1044/jshr.3604.808.

Benjamin, B. J. (1997). Speech production of normally aging adults. *Semin. Speech Lang.*, 18(02):135–141. https://doi.org/10.1055/s-2008-1064068.

Bent, T., Baese-Berk, M., Borrie, S., and McKee, M. (2016). Individual differences in the perception of unfamiliar regional, nonnative, and disordered speech varieties. *J. Acoust. Soc. Am.*, 140(5):3775–3786. https://doi.org/10.1121/1.4966677.

Bentley, L. E. and Ou, H. (2017). Using QuickSIN speech material to measure acceptable noise level for adults with hearing loss. *AuD Capstone Projects - Communication Sciences and Disorders (Doctoral thesis)*, 5:1–20. https://ir.library.illinoisstate.edu/aucpcsd/5.

Besser, J. (2015). *The Connected Ear. Influences of Cognitive and Auditory-Temporal Processing on Speech Understanding in Adverse Conditions.* (Doctoral thesis), Vrije Universiteit Amsterdam, The Netherlands. http://dare.ubvu.vu.nl/handle/1871/52801.

Besser, J., Koelewijn, T., Zekveld, A. A., Kramer, S. E., and Festen, J. M. (2013). How linguistic closure and verbal working memory relate to speech recognition in noise – A review. *Trends Amplif.*, 17(2):75–93. https://doi.org/10.1177/1084713813495459.

Bilodeau-Mercure, M., Lortie, C. L., Sato, M., Guitton, M. J., and Tremblay, P. (2014). The neurobiology of speech perception decline in aging. *Brain Struct. Funct.*, 220(2):979–997. https://doi.org/10.1007/s00429-013-0695-3.

Birren, J. E. (1965). Age changes in speeded behavior: Its central nature and physiological correlates. In Welford, A. T. and Birren, J. E., editors, *Behavior, aging and the nervous system*, pages 191–216. Springfield, Ill.: Charles C. Thomas.

Birren, J. E., Casperson, R. C., and Botwinick, J. (1950). Age changes in pupil size. *J. Gerontol.*, 5(3):216–221. https://doi.org/10.1093/geronj/5.3.216.

Blamey, P. (1997). Are spiral ganglion cell numbers important for speech perception with a cochlear implant? *Am. J. Otolaryng.*, 18(Suppl.6):S11–S12. http://europepmc.org/abstract/MED/9391577.

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

Blamey, P., Pyman, B. C., Clark, G. M., Dowell, R. C., Gordon, M., Brown, A. M., and Hollow, R. D. (1992). Factors predicting postoperative sentence scores in postlingually deaf adult cochlear implant patients. *Ann. Otol. Rhinol. Laryngol.*, 101(4):342–348. https://doi.org/10.1177/000348949210100410.

Bland, J. M. and Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, 327(8476):307–310. https://doi.org/10.1016/s0140-6736(86)90837-8.

Boersma, P. and Weenink, D. (2016). Praat: Doing phonetics by computer [Computer program]. www.praat.org. Last accessed on 22-10-2016.

Bosker, H. R. (2017). How our own speech rate influences our perception of others. *J. Exp. Psychol. Learn.*, 43(8):1225–1238. https://doi.org/10.1037/xlm0000381.

Bosman, A. J. and Smoorenburg, G. F. (1992). Woordenlijst voor spraakaudiometrie. [Compact Disc].

Bosman, A. J. and Smoorenburg, G. F. (1995). Intelligibility of Dutch CVC syllables and sentences for listeners with normal hearing and with three types of hearing impairment. *Int. J. Audiol.*, 34(5):260–284. https://doi.org/10.3109/00206099509071918.

Brännström, K. J., Holm, L., Kastberg, T., and Olsen, S. Ø. (2014a). The acceptable noise level: The effect of repeated measurements. *Int. J. Audiol.*, 53(1):21–29. https://doi.org/10.3109/14992027.2013.824116.

Brännström, K. J., Lantz, J., Nielsen, L. H., and Olsen, S. Ø. (2012). Acceptable noise level with Danish, Swedish, and non-semantic materials. *Int. J. Audiol.*, 51(3):146–156. https://doi.org/10.3109/14992027.2011.609183.

Brännström, K. J., Olsen, S., Holm, L., Kastberg, T., and Ibertsson, T. (2014b). The effect of repeated measurements and working memory on the most comfortable level in the ANL test. *Int. J. Audiol.*, 53(11):787–795. https://doi.org/10.3109/14992027.2014.938781.

Brännström, K. J., Zunic, E., Borovac, A., and Ibertsson, T. (2012). Acceptance of background noise, working memory capacity, and auditory evoked potentials in subjects with normal hearing. *J. Am. Acad. Audiol.*, 23(7):542–552. https://doi.org/10.3766/jaaa.23.7.6.

Brouwer, S., Mitterer, H., and Huettig, F. (2012). Can hearing *puter* activate *pupil*? Phonological competition and the processing of reduced spoken words in spontaneous conversations. *Q. J. Exp. Psychol.*, 65(11):2193–2220. https://doi.org/10.1080/17470218.2012.693109.

Byrne, D., Dillon, H., Tran, K., Arlinger, S., Wilbraham, K., Cox, R. …., and Ludvigsen, C. (1994). An international comparison of long-term average speech spectra. *J. Acoust. Soc. Am.*, 96(4):2108–2120. https://doi.org/10.1121/1.410152.

Carhart, R. and Jerger, J. F. (1959). A preferred method for clinical determination of pure-tone thresholds. *J. Speech. Hear. Disord.*, 24:330–345. https://doi.org/10.1044/jshd.2404.330.

Carrol, R., Koch, X., Janse, E., and Ruigendijk, E. (2017). Speech rate effects on word recognition in non-native conversational speech. *Conversational Speech and Lexical Representations Workshop*. Nijmegen, The Netherlands. https://tinyurl.com/CONVERSWORKSHOP2017.

Casserley, E. D. (2015). Effects of real-time cochlear implant simulation on speech production. *J. Acoust. Soc. Am.*, 137(5):2791–2800. https://doi.org/10.1121/1.4916965.

Cerella, J. (1990). Handbook of the Psychology of Aging. In Birren, J. E. and Schaie, K. W., editors, *Handbook of the Psychology of Aging*, chapter Aging and information processing rate, pages 201–221. Academic Press, New York.

Charness, N. and Bosman, E. A. (1994). Age-related changes in perceptual and psychomotor performance: Implications for engineering design. *Exp. Aging Res.*, 20(1):45–59. https://doi.org/10.1080/03610739408253953.

Chen, D. S., Betz, J., Yaffe, K., Ayonayon, H. N., Kritchevsky, S., Martin, K. R. …., and Lin, F. R. (2015). Association of hearing impairment with declines in physical functioning and the risk of disability in older adults. *J. Gerontol. A Biol. Sci. Med. Sci.*, 70(5):654–661. https://doi.org/10.1093/gerona/glu207.

Chen, Y.-W. and Watson, J. (2017). Speech production and sensory impairment in mild parkinson's disease. *J. Acoust. Soc. Am.*, 141(4):3030–3014. https://doi.org/10.1121/1.4980138.

Clinard, C. G., Tremblay, K. L., and Krishnan, A. R. (2010). Aging alters the perception and physiological representation of frequency: Evidence from human frequency-following response recordings. *Ear Hear.*, 264(1-2):48–55. https://doi.org/10.1016/j.heares.2009.11.010.

Cord, M. T., Surr, R. K., Walden, B. E., and Dyrlund, O. (2004). Relationship between laboratory measures of directional advantage and everyday success with directional microphone hearing aids. *J. Am. Acad. Audiol.*, 37(15):353–364. https://doi.org/10.3766/jaaa.15.5.3.

Coyle, T. R., Pillow, D. R., Snyder, A. C., and Kochunov, P. (2011). Processing speed mediates the development of general intelligence (g) in adolescence. *Psychol. Sci.*, 22(10):1265–1269. https://doi.org/10.1177/0956797611418243.

Cruickshanks, K., Wiley, T., Tweed, T., Klein, B., Klein, R., Mares-Perlman, J., and Nondahl, D. (1998). Prevalence of hearing loss in older adults in Beaver Dam, Wisconsin. The epidemiology of hearing loss study. *J. Epidemiol.*, 148(9):879–886. https://doi.org/10.1093/oxfordjournals.aje.a009713.

Daneman, M. and Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychon. B. Rev.*, 3(4):422–433. https://doi.org/10.3758/bf03214546.

Dingemanse, J. G. and Goedegebure, A. (2015). Application of noise reduction algorithm ClearVoice in cochlear implant processing. *Ear Hear.*, 36(3):357–367. https://doi.org/10.1097/aud.0000000000000125.

Dryden, A., Allen, H. A., Henshaw, H., and Heinrich, A. (2017). The association between cognitive performance and speech-in-noise perception for adult listeners: A systematic literature review and meta-analysis. *Trends Hear.*, 21(5):1–21. https://doi.org/10.1177/2331216517744675.

Eddins, D. A., Arnold, M., Klein, A., and Ellison, J. (2013). Individual variability in unaided and aided measurements of the acceptable noise level. *Semin. Hear.*, 34(02):118–127. https://doi.org/10.1055/s-0033-1341348.

Emery, V. O. B. (2000). Language impairment in dementia of the Alzheimer type: A hierarchical decline? *Int. J. Psych. Med.*, 30(2):145–164. https://doi.org/10.2190/x09p-n7au-ucha-vw08.

Ernestus, M. and Warner, N. (2011). An introduction to reduced pronunciation variants. *J. Phonetics*, 39(3):253–260. https://tinyurl.com/JPhonErnestusWarner2011.

Eurostat (2017). *Population structure and ageing*. Retrieved from https://tinyurl.com/EurostatPopulationAging.

Farris-Trimble, A., McMurray, B., Cigrand, N., and Tomblin, J. B. (2014). The process of spoken word recognition in the face of signal degradation. *J. Exp. Psychol.*, 40(1):308–327. https://doi.org/10.1037/a0034353.

Fitzgibbons, P. J. and Gordon-Salant, S. (2010). Behavioural studies with aging humans: Hearing sensitivity and psychoacoustics. In Gordon-Salant, S., Frisina, R. D., Fay, R., and Popper, A. N., editors, *The aging auditory system. Springer Handbook of Auditory Research*, pages 111–134. Springer, Berlin.

Forrest, K., Weismer, G., Milenkovic, P., and Dougall, R. N. (1988). Statistical analysis of word-initial voiceless obstruents: Preliminary data. *J. Acoust. Soc. Am.*, 84(1):115–123. https://doi.org/10.1121/1.396977.

Fortunato, S., Forli, F., Guglielmi, V., Corso, E. D., Paludetti, G., Berrettini, S., and Fetoni, A. R. (2016). A review of new insights on the association between hearing loss and cognitive decline in ageing. *Acta Otorhinolaryngo.*, 36(3):155–166. http://www.actaitalica.it/issues/2016/3-2016/01-Fetoni-Berrettini.pdf.

Francis, A. L. and Füllgrabe, C. (2015). Research on listening effort: History and methods, theory, and practice. *J. Acoust. Soc. Am.*, 137(4):2209. https://doi.org/10.1121/1.4920043.

Freyaldenhoven, M. C., Plyler, P. N., Thelin, J. W., and Hedrick, M. S. (2007). The effects of speech presentation level on acceptance of noise in listeners with normal and impaired hearing. *J. Speech. Lang. Hear. R.*, 50(4):878–885. https://doi.org/10.1044/1092-4388(2007/062).

Fuchs, S. and Toda, M. (2010). Do differences in male versus female /s/ reflect biological or sociophonetic factors? In Fuchs, S., Toda, M., and Zygis, M., editors, *Turbulent sounds. An interdisciplinary guide*, pages 281–302. Berlin: Mouton de Gruyter. https://tinyurl.com/FuchsTodaPaper2008.

Gatehouse, S. and Noble, W. (2004). The Speech, Spatial and Qualities of hearing scale (SSQ). *Int. J. Audiol.*, 43(2):85–99. https://doi.org/10.1080/14992020400050014.

Ghosh, S. S., Matthies, M. L., Maas, E., Hanson, A., Tiede, M., Menard, L. . . . ., and Perkell, J. S. (2010). An investigation of the relation between sibilant production and somatosensory and auditory acuity. *J. Acoust. Soc. Am.*, 128(5):3079–3087. https://doi.org/10.1121/1.3493430.

Glasberg, B. R. and Moore, B. C. J. (1990). Derivation of auditory filter shapes from notched noise data. *Hear. Res.*, 47(1–2):103–138. https://tinyurl.com/GlasbergMoore1990.

Goehl, H. and Kaufman, D. K. (1984). Do the effects of adventitiously deafness include disordered speech? *J. Speech Hear. Disord.*, 49(1):58–64. https://doi.org/10.1044/jshd.4901.58.

Gordon, M., Daneman, M., and Schneider, B. (2009). Comprehension of speeded discourse by younger and older listeners. *Exp. Aging Res.*, 35(3):277–296. https://doi.org/10.1080/03610730902769262.

Gordon-Hickey, S. and Moore, R. E. (2008). Acceptance of noise with intelligible, reversed, and unfamiliar primary discourse. *Am. J. Audiol.*, 17(2):129–135. https://doi.org/10.1044/1059-0889(2008/06-0018).

Gordon-Hickey, S. and Morlas, H. (2015). Speech recognition at the acceptable noise level. *J. Am. Acad. Audiol.*, 26(5):443–450. https://doi.org/10.3766/jaaa.14079.

Gordon-Salant, S. and Fitzgibbons, P. J. (1993). Temporal factors and speech recognition performance in young and elderly listeners. *J. Speech Lang. Hear. Res.*, 36(6):1276–1285. https://doi.org/10.1044/jshr.3606.1276.

Gordon-Salant, S. and Fitzgibbons, P. J. (1999). Profile of auditory temporal processing in older listeners. *J. Speech Lang. Hear. R.*, 42(2):300–311. https://doi.org/10.1044/jslhr.4202.300.

Gordon-Salant, S., Yeni-Komshian, G. H., Fitzgibbons, P. J., and Barrett, J. (2006). Age-related differences in identification and discrimination of temporal cues in speech segments. *J. Acoust. Soc. Am.*, 119(4):2455–2466. https://doi.org/10.1121/1.2171527.

Gordon-Salant, S., Zion, D. J., and Espy-Wilson, C. (2014). Recognition of time-compressed speech does not predict recognition of natural fast-rate speech by older listeners. *J. Acoust. Soc. Am.*, 136(4):EL286–EL274. https://doi.org/10.1121/1.4895014.

Gould, J., Lane, H., Vick, J., Perkell, J. S., Matthies, M. L., and Zandipour, M. (2001). Changes in speech intelligibity of postlingually deaf adults after cochlear implantation. *Ear Hear.*, 22(6):453–460. https://tinyurl.com/GouldETAL2001Paper.

Goy, H., Pelletier, M., Coletta, M., and Pichora-Fuller, M. K. (2013). The effects of semantic context and the type and amount of acoustic distortion on lexical decision by younger and older adults. *J. Speech Lang. Hear. R.*, 56(6):1715–1732. https://doi.org/10.1044/1092-4388(2013/12-0053).

Greenberg, S. (1998). Speaking in shorthand – A syllable-centric perspective for understanding pronuniciation variation. In *Proceedings of the ESCA Workshop on Modeling Pronunciation Variation for Automatic Speech Recognition, Kekrade, the Netherlands*, volume 115, pages 47–56. https://tinyurl.com/Greenberg1998.

Guenther, F. H. and Hickok, G. (2015). Role of the auditory system in speech production. In Celesia, G. and Hickok, G., editors, *Handbook of Clinical Neurology. The Human Auditory System - Fundamental Organization and Clinical Disorders*, volume 129, pages 161–175. Elsevier. https://tinyurl.com/GuentherHickock2015.

Haley, K., Seelinger, E., Mandulak, K., and Zajac, D. (2010). Evaluating the spectral distinction between sibilant fricatives through a speaker-centered approach. *J. Phonetics*, 38(4):548–554. https://doi.org/10.1016/j.wocn.2010.07.006.

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

Hartman, D. and Danhauer, J. (1976). Perceptual features of speech for males in four perceived age decades. *J. Acoust. Soc. Am.*, 59(3):713–715. https://doi.org/10.1121/1.380894.

Haubert, N. and Pichora-Fuller, M. K. (1999). The perception of spoken language by elderly listeners: Contributions of auditory temporal processes. *Can. Acoust.*, 27(3):96–97. https://tinyurl.com/HaubertPFuller1999.

Heffner, C. C., Newman, R. S., Dilley, L. C., and Idsardi, W. J. (2015). Age-related differences in speech rate perception do not necessarily entail age-related differences in speech rate uses. *J. Speech Lang. Hear. R.*, 58(4):1341–1349. https://tinyurl.com/HeffnerETAL2015.

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nat. Rev. Neurosci.*, 13:135–145. https://doi.org/10.1038/nrn3158.

Hinderinck, J. B., Krabbe, P. F. M., and 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. *Otolaryng. Head Neck*, 123(6):756–765. https://doi.org/10.1067/mhn.2000.108203.

Ho, H.-C., Wu, Y.-H., Hsiao, S.-H., Stangl, E., Lentz, E. J., and Bentler, R. A. (2012). The equivalence of acceptable noise level (ANL) with English, Mandarin, and non-semantic speech: A study across the U.S. and Taiwan. *Int. J. Audiol.*, 52(2):83–91. https://doi.org/10.3109/14992027.2012.733422.

Hochberg, I. (2008). Most comfortable listening for loudness and intelligibility of speech. *Audiology*, 14(1):27–33. https://doi.org/10.3109/00206097509071720.

Holladay, J. T. (1997). Proper method for calculating average visual acuity. *J. Refract. Surg.*, 13(4):388–391. https://tinyurl.com/HolladayVisualAcuity.

Holube, I., Fredelake, S., Vlaming, M., and Kollmeier, B. (2010). Development and analysis of an International Speech Test Signal (ISTS). *Int. J. Audiol.*, 49(12):891–903. https://doi.org/10.3109/14992027.2010.506889.

Houde, J. F. and Jordan, M. I. (1998). Sensorimotor adaptation in speech production. *Science*, 279(5354):1213–1216. https://doi.org/10.1126/science.279.5354.1213.

Hox, J. J. (2010). *Multilevel analysis: Techniques and applications*. Routledge, New York. https://doi.org/10.4324/9780203852279.

Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Res.*, 1626:118–135. https://doi.org/10.1016/j.brainres.2015.02.014.

Humes, L. E., Burk, M. H., Coughling, M. P., Busey, T. A., and Strauser, L. E. (2007). Auditory speech recognition and visual text recognition in younger and older adults: Similarities and differences between modalities and the effects of presentation rate. *J. Speech Lang. Hear. R.*, 50(2):283–303. https://doi.org/10.1044/1092-4388(2007/021).

Humes, L. E., Halling, D., and Coughlin, M. (1996). Reliability and stability of various hearing-aid outcome measures in a group of elderly hearing-aid wearers. *J. Speech Hear. Res.*, 39(5):923–935. https://doi.org/10.1044/jshr.3905.923.

Janse, E. (2004). Word perception in fast speech: Artificially time-compressed vs. naturally produced fast speech. *Speech Commun.*, 42(2):155–173. https://doi.org/10.1016/j.specom.2003.07.001.

Janse, E. (2009). Processing of fast speech by elderly listeners. *J. Acoust. Soc. Am.*, 125(4):2361–2373. https://doi.org/10.1121/1.3082117.

Jones, J. A. and Munhall, K. G. (2000). Perceptual calibration of F0 production: Evidence from feedback perturbation. *J. Acoust. Soc. Am.*, 108(3):1246–1251. https://doi.org/10.1121/1.1288414.

Jongman, A., Wayland, R., and Wong, S. (2000). Acoustic characteristics of English fricatives. *J. Acoust. Soc. Am.*, 108(3):1252–1263. https://doi.org/10.1121/1.1288413.

Just, M. A. and Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychol. Rev.*, 87(4):329–354. https://tinyurl.com/JustCarpenter1980.

Kahnemann, D. (1973). *Attention and Effort*. Prentice-Hall, Englewood Cliffs.

Kalikow, D. N., Stevens, K. N., and Elliott, L. L. (1977). Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *J. Acoust. Soc. Am.*, 61(5):1337–1351. https://doi.org/10.1121/1.381436.

Killion, M. C., Niquette, P. A., Gudmundsen, G. I., Revit, L. J., and Banerjee, S. (2004). Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *J. Acoust. Soc. Am.*, 116(4):2395–2405. https://doi.org/10.1121/1.1784440.

Kim, J., Nam, K. W., Yook, S., Hong, S. H., Jong, D. P., and Kim, I. Y. (2015). Effect of the degree of sensorineural hearing impairment on the results of subjective evaluations of a noise-reduction algorithm. *Speech Commun.*, 68:1–10. https://doi.org/10.1016/j.specom.2015.01.001.

Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *J. Acoust. Soc. Am.*, 59(5):1208–1221. https://doi.org/10.1121/1.380986.

Knudsen, L. V., Öberg, M., Nielsen, C., Naylor, G., and Kramer, S. E. (2010). Factors influencing help seeking, hearing aid uptake, hearing aid use and satisfaction with hearing aids: A review of the literature. *Trends Amplif.*, 14(3):127–154. https://doi.org/10.1177/1084713810385712.

Knuijt, S., Kalf, J., van Engelen, B., Geurts, A., and de Swarts, B. (2017). Reference values of maximum performance tests of speech production. *Int. J. Speech Lang. Pa.*, Online preprint:1–9. https://doi.org/10.1080/17549507.2017.1380227.

Koch, X. and Janse, E. (2015). Effects of age and hearing loss on articulatory precision for sibilants. In The Scottish Consortium for ICPhS 2015, editor, *Proceedings of the 18th International Congress of Phonetic Sciences*. University of Glasgow. https://tinyurl.com/KochJanse2015.

Kontis, V., Bennett, J. E., Mathers, S. D., Li, G., Foreman, K., and Ezzati, M. (2017). Future life expectancy in 35 industrialised countries: Projections with a Bayesian model ensemble. *Lancet*, 389(10076):1323–1335. https://tinyurl.com/KontisETAL2017.

Krampe, R. T. (2002). Aging, expertise and fine motor movement. *Neurosci. Biobehav. R.*, 26(7):769–776. https://doi.org/10.1016/s0149-7634(02)00064-7.

Krause, J. C. and Braida, L. D. (2004). Acoustic properties of naturally produced clear speech at normal speaking rates. *J. Acoust. Soc. Am.*, 115(1):362–378. https://doi.org/10.1121/1.1635842.

Kuchinsky, S. E., Ahlstrom, J. B., Vaden, K. I., Cute, S. L., Humes, L. E., Dubno, J. R., and Eckert, M. A. (2012). Pupil size varies with word listening and response selection difficulty in older adults with hearing loss. *Psychophysiology*, 50(1):23–34. https://doi.org/10.1111/j.1469-8986.2012.01477.x.

Kuijer, R., de Ridder, D., Ouwehand, C., Houx, B., and van den Bos, R. (2008). Dieting as a case of behavioural decision making. Does self-control matter? *Appetite*, 51(3):506–511. https://doi.org/10.1016/j.appet.2008.03.014.

Kusomoto, A. and Vaughan, N. (2004). Comparison of acoustic features of time-compressed and natural speech. *J. Acoust. Soc. Am.*, 116(4):2600. https://doi.org/10.1121/1.4785377.

Lamel, L. F., Kassel, R. H., and Seneff, S. (1989). Speech database development: Design and analysis of the acoustic-phonetic corpus. In *Speech I/O Assessment and Speech Databases (ISCA), Noordwijkerhout, the Netherlands*, volume 115, pages 2161–2170. ISCA Archive. Retrieved from: https://tinyurl.com/LamelETAL1989-1.

Lametti, D. R., Nasir, S. M., and Ostry, D. J. (2012). Sensory preference in speech production revealed by simultaneous alteration of auditory and somatosensory feedback. *J. Neurosci.*, 32(27):9351–9358. https://doi.org/10.1523/jneurosci.0404-12.2012.

Lane, H., Matthies, M. L., Guenther, F., Perkell, J. S., Stockmann, E., Tiede, M. . . . ., and Zandipour, M. (2007). Effects of short- and long-term changes in auditory feedback on vowel and sibilant contrasts. *J. Speech Lang. Hear. R.*, 50:913–927. https://doi.org/10.1044/1092-4388(2007/065).

Lane, H., Matthies, M. L., Perkell, J. S., Vick, J., and Zandipour, M. (2001). The effect of changes in hearing status in cochlear implant users on the acoustic vowel space and CV coarticulation. *J. Speech Lang. Hear. R.*, 44(3):552–563. https://doi.org/10.1044/1092-4388(2001/043).

Lane, H. and Tranel, B. (1971). The Lombard sign and the role of hearing in speech. *J. Speech Lang. Hear. R.*, 14(4):677–709. https://doi.org/webbste10.1044/jshr.1404.677.

Lane, H. and Wozniak Webster, J. (1991). Speech deterioration in postlingually deafened adults. *J. Acoust. Soc. Am.*, 89(2):859–866. https://doi.org/10.1121/1.1894647.

Langereis, M. C., Bosman, A. J., van Olphen, A. F., and Smoorenburg, G. F. (1997). Changes in vowel quality in post-lingually deafened cochlear implant users. *Audiology*, 36(5):279–297. https://doi.org/10.3109/00206099709071980.

Langereis, M. C., Bosman, A. J., van Olphen, A. F., and Smoorenburg, G. F. (1998). Effect of cochlear implantation on voice fundamental frequency in post-lingually deafened adults. *Audiology*, 37(4):219–230. https://doi.org/10.3109/00206099809072976.

Langereis, M. C., Bosman, A. J., van Olphen, A. F., and Smoorenburg, G. F. (1999). Intelligibility of vowels produced by post-lingually deafened cochlear implant users. *Audiology*, 38(4):206–224. https://doi.org/10.3109/00206099909073025.

Leder, S. B. and Spitzer, J. B. (1990). A perceptual evaluation of the speech of adventitiously deaf adult males. *Ear Hear.*, 11(3):169–175. https://doi.org/10.1097/00003446-199006000-00001.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation.* MIT Press, Cambridge, MA. https://tinyurl.com/Levelt1989.

Levelt, W. J. M., Roelofs, A., and Meyer, A. S. (1999). A theory of lexical access in speech production. *Behav. Brain Sci.*, 22(01):1–75. https://doi.org/10.1017/s0140525x99001776.

Levinson, S. C. (2012). The original sin of cognitive science. *Top. Cogn. Sci.*, 4(3):396–403. https://doi.org/10.1111/j.1756-8765.2012.01195.x.

Lichtheim, L. (1885). On aphasia. *Brain*, 13:135–145. available via: https://doi.org/10.1093/acprof:oso/9780195177640.003.0020.

Lin, F. R., Yaffe, K., Xia, J., Xue, Q.-L., Harris, T. B., Purchase-Helzner, E. …., and Simonsick, E. M. (2013). Hearing loss and cognitive decline among older adults. *JAMA Intern. Med.*, 173(3):293–299. https://doi.org/10.1001/jamainternmed.2013.1868.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H & H theory. In Hardcastle, W. J. and Marchal, A., editors, *Speech Production and Speech Modelling*, pages 403–439. Kluwer. https://tinyurl.com/Lindblom1990.

Lindenberger, U. and Baltes, P. B. (1994). Sensory functioning and intelligence in old age. A strong connection. *Psychol. Aging*, 9(3):339–355. https://doi.org/10.1037//0882-7974.9.3.339.

Linthicum, F. H. and Fayad, J. N. (2009). Spiral ganglion cell loss is unrelated to segmental cochlear sensory system degeneration in humans. *Otol. Neurotol.*, 30(3):612–620. https://doi.org/10.1097/mao.0b013e31819a8827.

Lobanov, B. M. (1971). Classification of Russian vowels spoken by different speakers. *J. Acoust. Soc. Am.*, 49(2B):606–608. https://doi.org/10.1121/1.1912396.

Lombard, E. (1911). Le signe de l'élévation de la voix. *Ann. Mal. Oreille Larynx Nez Pharynx*, 37(2):101–119.

Lubbers, M. and Torreira, F. (2016). *Praatalign: An interactive Praat plug-in for performing phonetic forced alignment*. [Computer program/plugin]. Retrieved from https://github.com/dopefishh/praatalign.

Luce, P. A. and Pisoni, D. B. (1998). Recognizing spoken words: The neighboorhood activation model. *Ear Hear.*, 19(1):1–36. https://tinyurl.com/LucePisoni1998.

Luteijn, F. and van der Ploeg, F. A. E. (1983). *Handleiding Groninger Intelligentietest*. [Manual Groningen Intelligence Test] Lisse, The Netherlands:, Swets and Zeitlinger.

Lyxell, B., Andersson, J., Andersson, U., Arlinger, S., Bredberg, G., and Harder, H. (1998). Phonological representation and speech understanding with cochlear implants in deafened adults. *Scand. J. Psychol.*, 39(3):175–179. https://doi.org/10.1111/1467-9450.393075.

Marslen-Wilson, W. D. and Komisarjevsky Tyler, L. (1980). The temporal structure of spoken language understanding. *Cognition*, 8(1):1–71. https://tinyurl.com/MarslenWTyler1980.

Marslen-Wilson, W. D. and Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cogn. Pychol.*, 10(1):29–63. https://doi.org/10.1016/0010-0285(78)90018-x.

Matthies, M. L., Guenther, F. H., Denny, M., Perkell, J. S., Burton, E., Vick, J. …., and Zandipour, M. (2008). Perception and production of /r/ allophones improve with hearing from a cochlear implant. *J. Acoust. Soc. Am.*, 124(5):3191–3202. https://doi.org/10.1121/1.2987427.

Matthies, M. L., Svirsky, M., Perkell, J. S., and Lane, H. (1996). Acoustic and articulatory measures of sibilant productions with and without auditory feedback from a cochlear implant. *J. Speech Lang. Hear. R.*, 39(5):936–946. https://doi.org/10.1044/jshr.3905.936.

Matthies, M. L., Svirsky, M. A., Lane, H., and Perkell, J. S. (1994). A preliminary study of the effect of cochlear implants on the production of sibilants. *J. Acoust. Soc. Am.*, 96(3):1367–1373. https://doi.org/10.1121/1.410281.

Mattys, S. L., Davis, M. H., Bradlow, A. R., and Scott, S. K. (2012). Speech recognition in adverse conditions: A review. *Lang. Cognitive Proc.*, 27(7):953–978. https://doi.org/10.1080/01690965.2012.705006.

McAuliffe, M. J., Gibson, E. M. R., Kerr, S. E., Anderson, T., and LaShell, P. J. (2013). Vocabulary influences older and younger listeners' processing of dysarthric speech. *J. Acoust. Soc. Am.*, 134(5):1358–1368. https://doi.org/10.1121/1.4812764.

McClelland, J. L. and Elman, J. L. (1986). The TRACE model of speech perception. *Cogn. Pychol.*, 18(1):1–86. https://doi.org/10.1016/0010-0285(86)90015-0.

McCormack, A. and Fortnum, H. (2013). Why do people fitted with hearing aids not wear them? *Int. J. Audiol.*, 52(5):360–368. https://doi.org/10.3109/14992027.2013.769066.

McGarrigle, R., Munro, K. J., Dawes, P., Stewart, A. J., Moore, D. R., Barry, J. G., and Amitay, S. (2014). Listening effort and fatigue: What exactly are we measuring? A British Society of Audiology Cognition in Hearing Special Interest Group 'white paper'. *Int. J. Audiol.*, 53(7):433–440. https://doi.org/10.3109/14992027.2014.890296.

McMurray, B., Samelson, V. M., Lee, S. H., and Tomblin, J. B. (2010). Individual differences in online spoken word recognition: Implications for SLI. *Cogn. Psychol.*, 60(1):1–39. https://doi.org/10.1016/j.cogpsych.2009.06.003.

McQueen, J. M. and Viebahn, M. C. (2007). Tracking recognition of spoken words by tracking looks to printed words. *Q. J. Exp. Psychol.*, 60(5):661–671. https://doi.org/10.1080/17470210601183890.

Mefferd, A. S. and Corder, E. E. (2014). Assessing articulatory speed performance as a potential factor of slowed speech in older adults. *J. Speech Lang. Hear. R.*, 57(2):347. https://tinyurl.com/MefferdETAL2014.

Meyer, T. A., Frisch, S. A., Pisoni, D. B., Miyamoto, R. T., and Svirsky, M. A. (2003). Modeling open-set spoken word recognition in postlingually deafened adults after cochlear implantation: Some preliminary results with the Neighborhood Activation Model. *Otol. Neurotol.*, 24(4):612–620. https://tinyurl.com/MeyerETAL2003.

Mirman, D. (2014). *Growth curve analysis and visualization in R.* CRC Press, Boca Raton, FL.

Moore, R., Gordon-Hickey, S., and Jones, A. (2011). Most comfortable listening levels, background noise levels, and acceptable noise levels for children and adults with normal hearing. *J. Am. Acad. Audiol.*, 22(5):286–293. https://doi.org/10.3766/jaaa.22.5.5.

Mueller, G. H., Weber, J., and Hornsby, B. W. (2006). The effects of digital noise reduction on the acceptance of background noise. *Trends Amplif.*, 10(2):83–93. https://doi.org/10.1177/1084713806289553.

Nábělek, A. K., Freyaldenhoven, M. C., Tampas, J. W., Burchfield, S. B., and Muenchen, R. A. (2006). Acceptable noise level as a predictor of hearing aid use. *J. Am. Acad. Audiol.*, 17(9):626–639. https://doi.org/10.3766/jaaa.17.9.2.

Nábělek, A. K., Tampas, J. W., and Burchfield, S. B. (2004). Comparison of speech perception in background noise with acceptance of background noise in aided and unaided conditions. *J. Speech Lang. Hear. R.*, 47(5):1001–1011. https://doi.org/10.1044/1092-4388(2004/074).

Nábělek, A. K., Tucker, F. M., and Letowski, T. R. (1991). Toleration of background noises: Relationship with patterns of hearing aid use by elderly persons. *J. Speech Hear. Res.*, 34(3):679–685. https://doi.org/10.1044/jshr.3403.679.

Nadol, J. B., Young, Y.-S., and Glynn, R. J. (1989). Survival of spiral ganglion cells in profound sensorineural hearing loss: Implications for cochlear implantation. *Ann. Otol. Rhinol. Laryngol.*, 98(6):411–416. https://doi.org/10.1177/000348948909800602.

Nasir, S. M. and Ostry, D. J. (2008). Speech motor learning in profoundly deaf adults. *Nature Neurosc.*, 11(10):1217–1222. https://doi.org/10.1038/nn.2193.

Neger, T., Janse, E., and Rietveld, T. (2015). Correlates of older adults' discrimination of acoustic properties in speech. *Speech Lang. Hear.*, 18(2):102–115. https://doi.org/10.1179/2050572814y.0000000055.

Ng, E. H. N., Classon, E., Larsby, B., Arlinger, S., Lunner, T., Rudner, M., and Rönnberg, J. (2014). Dynamic relation between working memory capacity and speech recognition in noise during the first 6 months of hearing aid use. *Trends Hear.*, 18:1–10. https://doi.org/10.1177/2331216514558688.

Ng, E. H. N., Rudner, M., Lunner, T., Pedersen, M. S., and Rönnberg, J. (2013). Effects of noise and working memory capacity on memory processing of speech for hearing-aid users. *Int. J. Audiol.*, 52(7):433–441. https://doi.org/10.3109/14992027.2013.776181.

Nichols, A. C. and Gordon-Hickey, S. (2012). The relationship of locus of control, self-control, and acceptable noise levels for young listeners with normal hearing. *Int. J. Audiol.*, 51(4):353–359. https://doi.org/10.3109/14992027.2011.645074.

Noble, W. and Gatehouse, S. (2004). Interaural asymmetry of hearing loss, Speech, Spatial and Qualities of Hearing Scale (SSQ) disabilities, and handicap. *Int. J. Audiol.*, 43(2):100–114. https://doi.org/10.1080/14992020400050015.

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, 52(3):189–234. https://doi.org/10.1016/0010-0277(94)90043-4.

Norris, D. and McQueen, J. M. (2008). Shortlist B: A Bayesian model of continuous speech recognition. *Psychol. Rev.*, 115(2):357–395. https://tinyurl.com/NorrisMcQueen2008.

Ohlenforst, B., Zekveld, A. A., Jansma, E. P., Wang, Y., Naylor, G., Lorens, A. …., and Kramer, S. E. (2017). Effects of hearing impairment and hearing aid amplification on listening effort. *Ear Hear.*, 38(3):267–281. https://doi.org/10.1097/aud.0000000000000396.

Olsen, S. Ø. and Brännström, K. J. (2014). Does the acceptable noise level (ANL) predict hearing-aid use? *Int. J. Audiol.*, 53(1):2–10. https://doi.org/10.3109/14992027.2013.839887.

Olsen, S. Ø., Lantz, J., Nielsen, L. H., and Brännström, K. J. (2012a). Acceptable noise level (ANL) with Danish and non-semantic speech materials in adult hearing-aid users. *Int. J. Audiol.*, 51(9):678–688. https://doi.org/10.3109/14992027.2012.692822.

Olsen, S. Ø., Nielsen, L. H., Lantz, J., and Brännström, K. J. (2012b). Acceptable noise level: Repeatability with Danish and non-semantic speech materials for adults with normal hearing. *Int. J. Audiol.*, 51(7):557–563. https://doi.org/10.3109/14992027.2012.666362.

Oostdijk, N. (2000). The Spoken Dutch Corpus. Overview and first evaluation. In Gravilidou, M., Carayannis, G., Markantonatou, S., Piperidis, S., and Stainhaouer, G., editors, *Proceedings of the 2nd International Conference on Language Resources and Evaluation, Paris, France*, pages 887–893. ELRA. Retrieved from: http://www.lrec-conf.org/proceedings/lrec2000/pdf/110.pdf.

Park, J., Carp, J., Hebrank, A., Park, D. C., and Polk, T. A. (2010). Neural specificity predicts fluid processing ability in older adults. *J. Neurosci.*, 30(27):9253–9259. https://doi.org/10.1523/jneurosci.0853-10.2010.

Peelle, J. E., Troiani, V., Grossman, M., and Wingfield, A. (2011). Hearing loss in older adults affects neural systems supporting speech comprehension. *J. Neurosci.*, 31(35):12638–12643. https://doi.org/10.1523/JNEUROSCI.2559-11.2011.

Peeters, H., Kuk, F., Lau, C.-C., and Keenan, D. (2009). Subjective and objective evaluation of noise management algorithms. *J. Am. Acad. Audiol.*, 20(2):89–98. https://doi.org/10.3766/jaaa.20.2.2.

Pelle, J. E. and Wingfield, A. (2016). The neural consequences of age-related hearing loss. *Trends Neurosci.*, 39(7):486–497. https://doi.org/10.1016/j.tins.2016.05.001.

Perkell, J. S., Denny, M., Lane, H., Guenther, F., Mathies, M. L., Tiede, M. …., and Burton, E. (2005). Effects of masking noise on vowel and sibilant contrasts in normal-hearing speakers and postlingually deafened cochlear implant users. *J. Acoust. Soc. Am.*, 118(3):1963–1964. https://doi.org/10.1121/1.4781569.

Perkell, J. S., Guenther, F. H., Lane, H., Matthies, M. L., Perrier, P., Vick, J. …., and Zandipour, M. (2000). A theory of speech motor control and supporting data from speakers with normal hearing and with profound hearing loss. *J. Phonetics*, 28(3):233–272. https://doi.org/10.1006/jpho.2000.0116.

Perkell, J. S., Lane, H., Denny, M., Mathies, M. L., Tiede, M., Zandipour, M. …., and Burton, E. (2007). Time course of speech changes in response to unanticipated short-term changes in hearing state. *J. Acoust. Soc. Am.*, 121(4):2296–2311. https://doi.org/10.1121/1.2642349.

Perkell, J. S., Lane, H., Svirsky, M., and Webster, J. (1992). Speech of cochlear implant patients: A longitudinal study of vowel production. *J. Acoust. Soc. Am.*, 91(5):2961–2978. https://doi.org/10.1121/1.402932.

Perkell, J. S., Matthies, M. L., Tiede, M., Lane, H., Zandipour, M., Marrone, N. . . . , and Guenther, F. H. (2004). The distinctness of speakers' /s/– /ʃ/ contrast is related to their auditory discrimination and use of an articulatory saturation effect. *J. Speech Lang. Hear. R.*, 47(6):1259–1269. https://tinyurl.com/PerkellETAL2004.

Pichora-Fuller, M. K. (2003). Cognitive aging and auditory information processing. *Int. J. Audiol.*, 42(Suppl.2):26–32. https://doi.org/10.3109/14992020309074641.

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W. Y., Humes, L. E. . . . , and Wingfield, A. (2016). Hearing impairment and cognitive energy: The framework for understanding effortful listening (FUEL). *Ear Hear.*, 37(Suppl.1):5S–27S. https://doi.org/10.1097/aud.0000000000000312.

Pichora-Fuller, M. K., Schneider, B. A., Benson, N. J., Hamstra, S. J., and Storzer, E. (2006). Effect of age on detection of gaps in speech and nonspeech markers varying in duration and spectral symmetry. *J. Acoust. Soc. Am.*, 119(2):1143–1155. https://doi.org/10.1121/1.2149837.

Pick, H. L., Siegel, G. M., Fox, P. W., Garber, S. R., and Kearney, J. K. (1989). Inhibiting the Lombard effect. *J. Acoust. Soc. Am.*, 85(2):894–900. https://doi.org/10.1121/1.397561.

Plant, G. (1983). The effects of long-term hearing loss on speech production. *Speech Transmission Laboratory Stockholm, Quarterly Status and Progress Report*, 24(1):18–35. https://tinyurl.com/Plant1983.

Plant, G. and Hammarberg, B. (1983). Acoustic and perceptual analysis of the speech of the deafened. *Speech Transmission Laboratory Stockholm, Quarterly Status and Progress Report*, 24(1-3):85–107. https://tinyurl.com/PlantHammarberg1983.

Plant, K., McDermott, H., van Hoesel, R., Dawson, P., and Cowan, R. (2016). Factors predicting postoperative unilateral and bilateral speech recognition in adult cochlear implant recipients with acoustic hearing. *Ear Hear.*, 37(2):153–163. https://doi.org/10.1097/aud.0000000000000233.

Plyler, P. N., Bahng, J., and von Hapsburg, D. (2008). The acceptance of background noise in adult cochlear implant users. *J. Speech Lang. Hear. R.*, 51(2):502–515. https://doi.org/10.1044/1092-4388(2008/036).

Plyler, P. N., Madix, S. G., Thelin, J. W., and Johnston, K. W. (2007). Contribution of high-frequency information to the acceptance of background noise in listeners with normal and impaired hearing. *Am. J. Audiol.*, 16(2):149–156. https://doi.org/10.1044/1059-0889(2007/019).

Pols, L. C. W., Tromp, H. R. C., and Plomp, R. (1973). Frequency analysis of Dutch vowels from 50 male speakers. *J. Acoust. Soc. Am.*, 53(4):1093–1101. https://doi.org/10.1121/1.1913429.

Privitera, C. M., Renninger, L. W., Carney, T., Klein, S., and Aguilar, M. (2008). The pupil dilation response to visual detection. In Rogowitz, B. E. and Pappas, T. N., editors, *Proceeding of SPIE-IS&T Human Vision and Electronic Imaging XIII*, volume 6806, pages 68060T–1–68060T–11. The International Society for Optical Engineering. Retrieved from: http://cornea.berkeley.edu/pubs/186.pdf.

Purcell, D. W. and Munhall, K. G. (2006). Compensation following real-time manipulation of formants in isolated vowels. *J. Acoust. Soc. Am.*, 119(4):2288–2297. https://doi.org/10.1121/1.2173514.

Quené, H. (2008). Multilevel modeling of between-speaker and within-speaker variation in spontaneous speech tempo. *J. Acoust. Soc. Am.*, 123(2):1104–1113. https://doi.org/10.1121/1.2821762.

Ramscar, M., Hendrix, P., Shaoul, C., Milin, P., and Baayen, H. (2014). The myth of cognitive decline: Non-linear dynamics of lifelong learning. *Top. Cogn. Sci.*, 6(1):5–42. https://doi.org/10.1111/tops.12078.

Raphael, L. J. and Dorman, M. F. (1980). Silence as a cue to the perception of syllable-initial and syllable-final stop consonants. *J. Phonetics*, 8(3):269–275. http://www.haskins.yale.edu/Reprints/HL0317.pdf.

Raven, J., Raven, J. C., and Court, J. (2003). *The Manual for the Raven's Progressive Matrices and Vocabulary Scales. Section 1: General Overview*. San Antonio: Harcourt Assessment.

Recker, K. L. and Micheyl, C. (2017). Speech intelligibility as a cue for acceptable noise levels. *Ear Hear.*, 38(4):465–474. https://doi.org/10.1097/AUD.0000000000000408.

Reidy, P. F. (2015). A comparison of spectral estimation methods for the analysis of sibilant fricatives. *J. Acoust. Soc. Am.*, 137(4):EL248–EL254. https://doi.org/10.1121/1.4915064.

Rennies, J., Schepker, H., Holube, I., and Kollmeier, B. (2014). Listening effort and speech intelligibility in listening situations affected by noise and reverberation. *J. Acoust. Soc. Am.*, 136(5):2642–2653. https://doi.org/10.1121/1.4897398.

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B. …., and Rudner, M. (2013). The Ease of Language Understanding (ELU) model:

Theoretical, empirical, and clinical advances. *Front. Syst. Neurosci.*, 7:1–17. https://doi.org/10.3389/fnsys.2013.00031.

Rönnberg, J., Rudner, M., Foo, C., and Lunner, T. (2008). Cognition counts: A working memory system for ease of language understanding (ELU). *Int. J. Audiol.*, 47(Suppl.2):S99–S105. https://doi.org/10.1080/14992020802301167.

Rudner, M., Rönnberg, J., and Lunner, T. (2011). Working memory supports listening in noise for persons with hearing impairment. *J. Am. Acad. Audiol.*, 22(3):156–167. https://doi.org/10.3766/jaaa.22.3.4.

Salthouse, T. A. (1985). Speed of behaviour and its implications for cognition. In Birren, J. E., Schaie, K. W., and Warner, K., editors, *Handbook of the Psychology of Aging*, pages 400–426. Van Nostrand Reinhold, New York.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychol. Rev.*, 103(3):403–428. https://tinyurl.com/Salthouse1996.

Salthouse, T. A. (2000). Aging and measures of processing speed. *Biol. Psychol.*, 54(1-3):35–54. https://doi.org/10.1016/s0301-0511(00)00052-1.

Sapir, S. and Canter, G. J. (1991). Postlingual deaf speech and the role of audition in speech production: Comments on Waldstein's paper [R.S. Waldstein, J. Acoust. Soc. Am. 88, 2099-2114 (1990)]. *J. Acoust. Soc. Am.*, 90(3):1672–1673. https://doi.org/10.1121/1.401909.

Scharenborg, O., Sanders, E., and Cranen, B. (2014). Collecting a corpus of Dutch noise-induced slips of the ear. In *Proceedings of Interspeech 2014, Singapore*, pages 2600–2604. https://tinyurl.com/ScharenborgETAL2014.

Schenk, B. S., Baumgärtner, W.-D., and Hamzavi, J.-S. (2003). Effect of the loss of auditory feedback on segmental parameters of vowels of postlingually deafened speakers. *Auris Nasus Larynx*, 30(4):333–339. https://tinyurl.com/SchenkETAL2003.

Schmidtke, J. (2014). Second language experience modulates word retrieval effort in bilinguals: Evidence from pupillometry. *Front. Psychol.*, 5 (Article 137). https://doi.org/10.3389/fpsyg.2014.00137.

Schmidtke, J. (2016). The bilingual disadvantage in speech understanding in noise is likely a frequency effect related to reduced language exposure. *Front. Psychol.*, 7(Article 678):1–15. https://doi.org/10.3389/fpsyg.2016.00678.

Schmitt, J. F. and Moore, J. R. (1989). Natural alteration of speaking rate. *J. Speech Hear. Res.*, 32(2):445–450. https://doi.org/10.1044/jshr.3202.445.

Schneider, B. A., Daneman, M., and Murphy, D. R. (2005). Speech comprehension difficulties in older adults: Cognitive slowing or age-related changes in hearing? *Psychol. Aging*, 20(2):261–271. https://doi.org/10.1037/0882-7974.20.2.261.

Schuerman, W. L. (2017). *Sensorimotor Experience in Speech Perception*. (Doctoral thesis), MPI for Psycholinguistics: Nijmegen, The Netherlands. https://tinyurl.com/Schuerman2017.

Schulte, M., Krüger, M., Meis, M., and Wagener, K. C. (2015). Subjective listening effort. *J. Acoust. Soc. Am.*, 137(4):2236–2236. https://doi.org/10.1121/1.4920159.

Small, J. A., Andersen, E. S., and Kempler, D. (1997). Effects of working memory capacity on understanding rate-altered speech. *Aging Neuropsychol. C.*, 4(2):126–139. https://doi.org/10.1080/13825589708256641.

Smeds, K., Wolters, F., and Rung, M. (2015). Estimation of signal-to-noise ratios in realistic sound scenarios. *J. Am. Acad. Audiol.*, 26(2):183–196. https://doi.org/10.3766/jaaa.26.2.7.

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T. …., and Levinson, S. C. (2009). Universals and cultural variation in turn-taking in conversation. *P. N. A. S.*, 106(26):10587–10592. https://doi.org/10.1073/pnas.0903616106.

Stuart-Smith, J. (2007). Empirical evidence for gendered speech production: /s/ in Glaswegian. *Lab. Phon.*, 9:65–86.

Svirsky, M. A., Lane, H., Perkell, J. S., and Wozniak, J. (1992). Effects of short-term auditory deprivation on speech production in adult cochlear implant users. *J. Acoust. Soc. Am.*, 92(3):1284–1300. https://doi.org/10.1121/1.403923.

Syka, J. (2002). Plastic changes in the central auditory system after hearing loss, restoration of function, and during learning. *Physiol. Rev.*, 82(3):601–636. https://doi.org/10.1152

Tangney, J. P., Baumeister, R. F., and Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *J. Pers.*, 72(2):271–324. https://doi.org/10.1111/j.0022-3506.2004.00263.x.

Thornton, R. and Light, L. L. (2006). Language Comprehension and Production in Normal Aging. In Birren, J. E., Schaie, K. W., Abeles, R. P., Gatz, M., and

Salthouse, T. A., editors, *Handbook of Psychology of Aging*, pages 261–287. Academic Press. https://doi.org/10.1016/B978-012101264-9/50015-X.

Tiffin, S. and Gordon-Hickey, S. (2017). The influence of speech rate on acceptable noise levels. *J. Am. Acad. Audiol.*, page in press. https://doi.org/10.3766/jaaa.16159, Online preprint.

Torre, J. and Barlow, J. (2009). Age-related changes in acoustic characteristics of adult speech. *J. Commun. Disord.*, 42(5):324–333. https://doi.org/10.1016/j.jcomdis.2009.03.001.

Tourville, J. A. and Guenther, F. H. (2011). The DIVA model: A neural theory of speech acquisition and production. *Cogn. Process.*, 26(7):952–981. https://doi.org/10.1080/01690960903498424.

Tremblay, P., Dick, A. S., and Small, S. L. (2013). Functional and structural aging of the speech sensorimotor neural system: Functional magnetic resonance imaging evidence. *Neurobiol. Aging*, 34(8):1935–1951. https://doi.org/10.1016/j.neurobiolaging.2013.02.004.

Tucker, B. V. and Ernestus, M. (2016). Why we need to investigate casual speech to truly understand language production, processing and the mental lexicon. *The Mental Lexicon*, 11(3):375–400. https://doi.org/10.1075/ml.11.3.03tuc.

Turner, G. S., Tjaden, K., and Weismer, G. (1995). The influence of speaking rate on vowel space and speech intelligibility for individuals with amyotrophic lateral sclerosis. *J. Speech Lang. Hear. R.*, 38(5):1001–1013. https://doi.org/10.1044/jshr.3805.1001.

Uhlmann, R. F., Larson, E. B., Rees, T. S., Koepsell, T. D., and Duckert, L. G. (1989). Relationship of hearing impairment to dementia and cognitive dysfunction in older adults. *J. Am. Med. Assoc.*, 261(13):1916–1919. https://doi.org/10.1001/jama.1989.03420130084028.

van Berkum, J. J., Brown, C. M., Zwitserlood, P., Kooijman, V., and Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *J. Exp. Psychol. Learn. Mem. Cogn.*, 31(3):443–467. https://doi.org/10.1037/0278-7393.31.3.443.

van der Harst, S. (2011). *The Vowel Space Paradox. A Sociophonetic Study on Dutch.* (LOT Doctoral thesis), Utrecht, The Netherlands. https://tinyurl.com/vanderHarst2011.

van Gerven, P. W. M., Paas, F., van Merriënboer, J. J. G., and Schmidt, H. G. (2004). Memory load and the cognitive pupillary response in aging. *Psychophysiology*, 41(2):167–174. https://doi.org/10.1111/j.1469-8986.2003.00148.x.

van Son, R. J. J. H., Wesseling, W., Sanders, E., and van den Heuvel, H. (2008). The IFADV corpus: A free dialog video corpus. In *Proceedings of the Sixth International Conference on Language Resources and Evaluation, LREC 2008*, pages 501–508. ELRA. https://tinyurl.com/vanSonETAL2008.

Vaughan, N., Storzbach, D., and Furukawa, I. (2006). Sequencing versus nonsequencing working wemory in understanding of rapid speech by older listeners. *J. Am. Acad. Audiol.*, 17(7):506–518. https://doi.org/10.3766/jaaa.17.7.6.

Vaz, S., Falkmer, T., Passmore, A. E., Parsons, R., and Andreou, P. (2013). The case for using the repeatability coefficient when calculating test–retest reliability. *PLoS ONE*, 8(9):e73990(1–7). https://doi.org/10.1371/journal.pone.0073990.

Versfeld, N. J., Daalder, L., Festen, J. M., and Houtgast, T. (2000). Method for the selection of sentence materials for efficient measurement of the speech reception threshold. *J. Acoust. Soc. Am.*, 107(3):1671–1684. https://doi.org/10.1121/1.428451.

Villacorta, V. M., Perkell, J. S., and Guenther, F. H. (2007). Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception. *J. Acoust. Soc. Am.*, 122(4):2306–2319. https://doi.org/10.1121/1.2773966.

von Hapsburg, D. and Bahng, J. (2006). Acceptance of background noise levels in bilingual (Korean-English) listeners. *J. Am. Acad. Audiol.*, 17(9):649–658. https://doi.org/10.3766/jaaa.17.9.4.

Wagner, A. E., Toffani, P., and Başkent, D. (2016). The timing and effort of lexical access in natural and degraded speech. *Front. Psychol.*, 398(7):1–14. https://doi.org/10.3389/fpsyg.2016.00398.

Waldstein, R. S. (1990). Effects of postlingual deafness on speech production: Implications for the role of auditory feedback. *J. Acoust. Soc. Am.*, 88(5):2099–2114. https://doi.org/10.1121/1.400107.

Walravens, E., Keidser, G., Hartley, D., and Hickson, L. (2014). An Australian version of the acceptable noise level test and its predictive value for successful hearing aid use in an older population. *Int. J. Audiol.*, 53(Suppl.1):S52–S59. https://doi.org/10.3109/14992027.2013.862599.

Wechsler, D. (2004). *Wechsler Adult Intelligence Test.* [Test material, Dutch Version]. Amsterdam: Harcourt Assessment.

Wernicke, C. (1874). *Der aphasische Symptomencomplex. Eine psychologische Studie auf anatomischer Basis.* Max Cohn and Weigert, Lisse, The Netherlands.

Wheeldon, L. R. and Levelt, W. J. M. (1995). Monitoring the time course of phonological encoding. *J. Mem. Lang.*, 34(3):311–334. https://doi.org/10.1006/jmla.1995.1014.

Wilhelm, O. and Schulze, R. (2002). The relation of speeded and unspeeded reasoning with mental speed. *Intelligence*, 30(6):537–554. https://tinyurl.com/WilhelmSchulze2002-1.

Wingfield, A. (1996). Cognitive factors in auditory performance: Context, speed of processing, and constraints of memory. *J. Am. Acad. Audiol.*, 7(3):175–182. https://tinyurl.com/Wingfield1996.

Wingfield, A., McCoy, S. L., Peelle, J. E., Tun, P. A., and Cox, C. L. (2006). Effects of adult aging and hearing loss on comprehension of rapid speech varying in syntactic complexity. *J. Am. Acad. Audiol.*, 17(7):487–497. https://doi.org/10.3766/jaaa.17.7.4.

Wingfield, A., Tun, P. A., Koh, C. K., and Rosen, M. J. (1999). Regaining lost time: Adult aging and the effect of time restoration on recall of time-compressed speech. *Psychol. Aging*, 14(3):380–389. https://doi.org/10.1037/0882-7974.14.3.380.

Wingfield, A., Tun, P. A., and McCoy, S. L. (2005). Hearing loss in older adulthood. What it is and how it interacts with cognitive performance. *Curr. Dir. Psychol. Sci.*, 14(3):144–148. https://tinyurl.com/WingfieldETAL2005.

Wu, Y.-H., Ho, H.-C., Hsiao, S.-H., Brummet, R. B., and Chipara, O. (2016). Predicting three-month and 12-month post-fitting real-world hearing-aid outcome using pre-fitting acceptable noise. *Int. J. Audiol.*, 55(5):285–294. https://doi.org/10.3109/14992027.2015.1120892.

Zekveld, A. A., Festen, J. M., and Kramer, S. E. (2013). Task difficulty differentially affects two measures of processing load: The pupil response during sentence processing and delayed cued recall of the sentences. *J. Speech Lang. Hear. R.*, 56(4):1156–1165. https://doi.org/10.1044/1092-4388(2012/12-0058).

Zekveld, A. A., Kramer, S. E., and Festen, J. M. (2010). Pupil response as an indication of effortful listening: The influence of sentence intelligibility. *Ear Hear.*, 31(4):480–490. https://doi.org/10.1097/aud.0b013e3181d4f251.

Zipf, G. K. (1949). *Human behavior and the principle of least effort*. Addison-Wesley, Oxford.

Zipf, G. K. (1965). *The psychobiology of language. An introduction to dynamic philology*. MIT Press, Cambridge, MA.

Appendix

Appendix A1: Question-answer sequences, ordered by speech rate

Q-A-sequence (Dutch)	Q-A-sequence (English)	Speech rate (phrase)	Syllables (target word)	Target word position	Target word frequency	Predictability absolute	Predictability relative	SNR (dB)
1. Q: En is daar ook wat aan vernield dan of zo? A: Nou ik denk dat *uh* die die **palen** ...	Q: And is there also something destroyed then or so? A: Well, I think that *uh* those those **poles** ...	2.93	2	7	2.89	0.60	0.40	22.35
2. Q: En die hangen dan om beurten in - van die apparaten? A: Ja die strekken **spieren** ...	Q: And they hang then in turns in these machines? A: Yes they stretch **muscles** ...	2.97	2	4	0.69	0.68	0.57	18.40
3. Q: Maar wat was er nog meer niet praktisch? A: Nou dat voorste **kastje** ...	Q: But what else was there not practical? A: Well, the most front **cupboard** ...	2.98	2	4	1.61	0.77	0.38	26.03
4. Q: En ben je nog kwaad geworden? A: Nee want *uh* die twee **gasten** ...	Q: And did you get angry? A: No because *uh* the two **dudes** ...	3.50	2	5	4.04	0.80	0.43	28.54
5. Q: Van wie? A: Van de **dochter** ...	Q: Of whom? A: Of the **daughter** ...	3.59	2	3	4.80	0.39	0.36	13.21
6. Q: Kunnen we op de stopknop drukken? A: Nou ik denk een **stukje** ...	Q: Can we press the Stop button? A: Well, I think a **bit** ...	3.69	2	5	4.50	0.69	0.45	21.09
7. Q: Oh ja? A: Ja maar een **nichtje**...	Q: Oh yes? A: Yes, but a **niece** ...	3.80	2	4	2.48	0.76	0.32	22.62
8. Q: Je hebt al brood gehaald? A: Ja ik heb voor een **week** ...	Q: You have already fetched bread? A: Yes I have for a **week** ...	4.08	1	6	5.65	0.61	0.45	20.06
9. Q: Wat denk jij dan? A: Nou dat het is *uh* de kleine **ruimte** ...	Q: What are you thinking of then? A: Well that it is *uh* the small **space** ...	4.11	2	6	5.02	0.80	0.42	21.47
10. Q: Twee vrije plekjes? A: Twee vrije plekjes ja tweeënhalf want *uh* hun **zoontje**...	Q: Two free places? A: Two free places, yes, two and a half because *uh* their **son** ...	4.24	2	8	5.25	0.83	0.45	19.80
11. Q: Ik heb *uh* autodienst? A: Jij hebt autodienst zondag uit de **kerk** ...	Q: It is my duty to *uh* drive? A: It is your duty to drive on sunday from the **church** ...	4.25	1	7	5.33	0.84	0.45	20.12
12. Q: Want die kwamen natuurlijk wel eens op bezoek? A: Ome Neel was dat met **tante** ...	Q: Obviously they would come now and then for a visit? A: Uncle Neel was it with **aunt** ...	4.34	2	6	4.65	0.55	0.45	17.88
13. Q: Dus je denkt dat de Nederlanders zelf zich niet uit *zijn *uh* moeras had kunnen optrekken? A: Nee die waren wel doorgegaan met meer **drank** ...	Q: So you think that the Dutch would not have been able to *uh* lift themselves up by *his bootstraps? A: No they would have kept going with more **booze** ...	4.38	1	8	3.69	0.34	0.32	31.67
14. Q: Zaten daar veel vogels? A: *Uh* op dat **eiland** ...	Q: Were there many birds? A: *Uh* on that **island** ...	4.46	2	3	4.37	0.38	0.35	32.96
15. Q: Wat had die voor ene rol? A: Ja der was de **leider** ...	Q: What kind of role did he/she have? A: Yes, he was the **leader** ...	4.60	2	5	4.28	0.56	0.47	27.23
16. Q: En is ze goed? A: Ja ze heeft wel een *uh* volwassen **stem** ...	Q: And is she good? A: Yes she does have a *uh* mature **voice** ...	4.60	1	7	5.73	0.80	0.45	18.80
17. Q: Waar stond die eigenlijk die boom? A: Achter *uh* achter in de tuin tegen de **schuur** ...	Q: Where was it situated actually, that tree? A: In the back of *uh* in the back of the garden against the **shed** ...	4.70	1	8	3.14	0.80	0.57	24.73
18. Q: Terwijl het toch niet echt warm is, he? A: Nee, ja van die hele kleine **mugjes** ...	Q: Although is isnt' really warm, is it? A: No, yes these very small **midges** ...	4.73	2	7	1.95	0.76	0.36	28.29
19. Q: Waarom wil het dat zo graag? A: Nou omdat het kennelijk voor *de overheid vindt dat het nog te vaak voorkomt dat **mensen** ...	Q: Why does she want it so badly? A: Well, because it obviously for *the government thinks it happens too often that **people** ...	4.92	2	16	7.22	0.85	0.36	22.44
20. Q: Met datzelfde sop? A: Ja en dan *uh* ja kregen wel een paar **emmers** ...	Q: With the same soapy water? A: Yes and then *uh* yes we did get a couple of **buckets** ...	4.98	2	9	3.14	0.90	0.49	34.64

Appendix A1: Question-answer sequences, ordered by speech rate (continued)

Q-A-sequence (Dutch)	Q-A-sequence (English)	Speech rate (phrase)	Syllables (target word)	Target word position	Target word frequency	Predictability absolute	Predictability relative	SNR
21. Q: Doen ze 'm via de andere kant erin? A: Nou wat was volgens die **arts** ...	Q: They put him via the other side into it? A: Well, what was according to that **doctor** ...	4.98	1	6	4.54	0.59	0.84	29.06
22. Q: Die rijdt niet te hard? A: Die heeft nog nooit een *uh* **boete** ...	Q: He/she does not drive too fast? A: He/she has never had a *uh* **fine** ...	5.01	2	7	2.30	0.83	0.55	26.33
23. Q: Ja wat bedoel je dan met maken? A: Nou ik, wij vermaken een **kind** ...	Q: Yes, what do you mean by make? A: Well I, we entertain a **child** ...	5.19	1	6	6.87	0.92	0.37	25.91
24. Q: En Bas Van Meerakker Meerakker ook of *uh*? A: Ja die heeft, die gaf voorlichting op **school** ...	Q: And Bas Van Meerakker Meerakker (name) too, right *uh*? A: Yes he has, he presented in **school** ...	5.25	1	8	5.50	0.78	0.58	17.31
25. Q: Waren ze weg? A: Die waren weg dus ik heb nieuwe **kaartjes** ...	Q: Were they gone? A: They were gone so I have new **tickets** ...	5.28	2	8	2.30	0.80	0.37	16.75
26. Q: Wat heb je met je hand gedaan dan bij het pink? A: Ja ik *uh* ik heb een **ruzie** ...	Q: What happend to your hand there at this little finger? A: Yes I *uh* I had a **quarrel** ...	5.39	2	6	3.78	0.67	0.35	28.19
27: Q: Nou nee ja je hoeft ze ook niet allemaal zo heel goed te lezen toch? A: Nee maar ja ze heeft me een paar **vragen** ...	Q: Well, no, yes, you don't have to read them that precisely, do you? A: No but yes she has (given) me a couple of **questions** ...	5.53	2	9	6.17	0.82	0.46	20.96
28. Q: Maar van *uh* dinge ook? A: Ja die heeft wel zo een **training** ...	Q: But from *uh* things, too? A: Yes he/she did have such a **training** ...	5.55	2	7	2.94	0.77	0.31	26.68
29. Q: Met wie? A: Met met Wendy en dan nog twee **vrienden** ...	Q: With whom? A: With with Wendy and then another two **friends** ...	5.67	2	8	5.65	0.51	0.38	19.07
30. Q: Ga je daarvan kopen? A: *Uh* nou ik ga daar een **keyboard** ...	Q: You are going to buy from that? A: *Uh* well, I am going to (buy) a **keyboard** ...	5.71	2	6	0	0.78	0.47	31.74
31. Q: Ja heb je het gezien? A: Nou ik heb dat **gat** ...	Q: Yes did you see it (happen)? A: Well, I have (not seen) the **gap** ...	5.72	1	5	4.62	0.57	0.47	17.94
32. Q: Hebben ze dat nergens daar? A: *Uh* in Maleisie hebben ze wel normaal **schrift** ...	Q: Do they have it nowhere there? A: *Uh* in Malaysia the have normal **writing system** ...	5.73	1	7	2.40	0.74	0.38	18.14
33. Q: Of hebben jullie dat daar ook gekocht? A: Ja daar hadden ze wel eens van die kleine **bakjes** ...	Q: Or did you buy this also there? A: Yes, there they used to have these small **containers** ...	5.82	2	10	1.79	0.92	0.33	23.46
34. Q: Wat was de vraag? A: Wie speelde op het witte doek de **rol** ...	Q: What was the question? A: Who played in the movies the **role** ...	5.99	1	8	5.35	0.48	0.46	31.99
35. Q: En hoeveel heb je daarvan? A: Nou dat zijn al zes **stoelen** ...	Q: And how many do you have from these? A: Well, there are already six **chairs** ...	6.03	2	6	5.02	0.84	0.48	19.59
36. Q: Ja maar dat is toch ook niet leuk? A: Nee maar ja die **ouders** ...	Q: Yes, but this is also not great? A: No but yes those **parents** ...	6.03	2	5	5.37	0.85	0.38	23.34
37. Q: Dus in een supermarkt kunnen ze toch inpakken of zo? A: Ja maar daar hebben ze liever mensen die een **jaar** ...	Q: So in a supermarkt they can allright wrap or something? A: Yes but there they prefer people who (stay) a **year** ...	6.11	1	10	7.04	0.89	0.48	23.02
38. Q: Kijk het is wel weg, he? A: Ja maar ik ga dus niet elke ochtend met een staafmixer mijn **koffie** ...	Q: Look it is gone, right? A: Yes, but I am not going to (make) every morning with a blender my **coffee** ...	6.24	2	13	4.72	0.85	0.46	29.39
39. Q: Moeten we met die mensen ook nog afspreken? A: Ja ik heb nu die **datum** ...	Q: Do we also have to make appointments with these people? A: Yes I now have that **date** ...	6.31	2	6	3.18	0.81	0.50	28.89
40. Q: Ging het lekker? A: Dit was een leuke **tocht** ...	Q: Did it go well? A: This was a great **journey** ...	6.41	1	5	3.74	0.71	0.45	17.55

Appendix A1: Question-answer sequences, ordered by speech rate (continued)

Q-A-sequence (Dutch)	Q-A-sequence (English)	Speech rate (phrase)	Syllables (target word)	Target word position	Target word frequency	Predictability absolute	Predictability relative	SNR
41. Q: Zullen we zo ook gaan eten? A: Ja wachten wel effe tot mijn **broers** ...	Q: Shall we have dinner in a minute? A: Yes (we) wait just till my **brothers** ...	6.46	1	7	4.86	0.76	0.52	18.16
42. Q: Hoezo niet? A: Omdat je geen *nergens twee **straten** ...	Q: Why not? A: Because you none *nowhere (have) two **streets** ...	6.58	2	6	5.28	0.61	0.31	23.85
43. Q: Wat is er “oh jee” aan? A: “Oh jee” omdat dat natuurlijk niet in ons **schema** ...	Q: What is “oh dear” about it? A: “Oh dear” because that of course (fits) not in our **scheme** ...	6.73	2	9	3.26	0.47	0.36	20.32
44. Q: Staat 1927 op maar zo oud is ie toch niet? A: Ik heb geen flauw idee hoe oud die **poster** ...	Q: It says 1927 on it but it is not that old, is it? A: I don’t have a clue how old that **poster** ...	6.86	2	9	1.10	0.52	0.40	27.43
45. Q: Maar ik kan jou toch gewoon ook *uh* vrijdagochtend dan naar het station brengen? A: Ja maar ik wil ook een **fiets** ...	Q: But I can also (bring) you as usual *uh* Friday morning then to the station? A: Yes but I want also a **bicycle** ...	6.99	1	7	3.89	0.67	0.53	27.37
46. Q: Nee maar als je daarbij bent? A: Nee dat is wel een **poos** ...	Q: No, but if you are present? A: No this is indeed a **while** ...	7.09	1	6	3.85	0.88	0.41	16.87
47. Q: Wat ga jij nou doen? A: Ik ga een **zon** ...	Q: What are you going to do? A: I am going to (make) a **sun** ...	7.17	1	4	3.85	0.49	0.31	26.33
48. Q: Iedereen heeft toch vragen? A: Ja dan moet je een **afspraak** ...	Q: Everybody has questions, right? A: Yes then must (make) an **appointment** ...	7.18	2	6	3.99	0.82	0.45	12.43
49. Q: Is het te sterk? A: Nou het is tamelijk veel **tijm** ...	Q: Is it too strong? A: It is rather a lot of **thyme** ...	7.50	1	6	1.95	0.87	0.39	22.46
50. Q: Wat is dat toch? A: Is dat van *het **zeep** ...	Q: What is it again? A: It is from *the **soap** ...	7.59	1	5	2.83	0.81	0.38	27.53
51. Q: Welke denk je aan? A: *Uh* ik die en die andere die hartige **soesjes** ...	Q: Which ones do you have in mind? A: *Uh* I that one and that other the hearty **pastry** ...	7.83	2	10	0	0.85	0.36	37.42
52. Q: Dan wat waren ze dan aan het doen beneden? A: Ik heb het niet gezien want het was donker natuurlijk in de **zaal** ...	Q: Then, what were they doing down there? A: I did not see it because it was dark of course in the **hall** ...	8.08	1	13	3.99	0.88	0.49	20.34
53. Q: Hoezo zijn we al bijna klaar? A: Nee nou dan kunnen we eindelijk gaan wisselen met **huizen** ...	Q: Why are we almost finished? A: No, well then we can finally change (the) **houses**	8.15	2	10	6.45	0.67	0.41	29.62
54. Q: Natuurlijk want die tellen toch ook? A: Die tellen wel maar ik moet kijken wat ik allemaal nog in de **hand** ...	Q: Of course, because they count as well? A: They count as well but I have to look what I (have) altogether in the **hand** ...	8.21	1	14	6.94	0.79	0.40	21.09
55. Q: Waar was het nou toch? A: Waar die ten **hemel** ...	Q: Where was it again? A: Where he (ascended) to **heaven** ...	8.48	2	4	4.62	0.87	0.32	22.81
56. Q: Echt waar? A: Ja maar die heeft dus ook een **tijdje** ...	Q: Really true? A: Yes but she/he also had a **while** ...	8.57	2	8	6.99	0.6	0.41	29.78
57. Q: Ja en die uitkomst is dat hardback? A: Ze beginnen altijd inderdaad met een harde **kaft** ...	Q: Yes and the result is that hardback? A: They begin indeed always with a hard **cover** ...	8.67	1	8	1.10	0.78	0.55	21.37
58. Q: Wat dan? A: Dan kun je geen **dorpje** ...	Q: What then? A: Then you can not (buy) a **village** ...	9.38	2	5	4.93	0.66	0.22	29.51
59. Q: De vierentwintigste? A: En dan moeten we nog heel veel **yoghurt** ...	Q: The 24th? A: And then we have to (eat) a lot of **yoghurt** ...	10.22	2	8	1.39	0.32	0.30	20.55
60. Q: Zo vroeg al? A: Ja we moeten **auto** ...	Q: That early already? A: Yes we must **car** ...	11.22	2	4	5.34	0.67	0.22	24.52

Appendix A2: Orthographic representations used in the experiment

Item	target word (English translation in brackets)	phonetic distractor	semantic distractor	phonetic distractor to the semantic distractor
1	palen (poles)	pasen	balken	banen
2	spieren (muscles)	spiesen	pezen	perken
3	kastje (cupboard)	kaarsje	plankje	platje
4	gasten (dudes)	gangen	kerels	kegels
5	dochter (daughter)	dode	moeder	molen
6	stukje (bit)	stuurtje	eindje	eitje
7	nichtje (niece)	nietje	zusje	zuchtje
8	week	wees	dag	dam
9	ruimte (space)	ruiter	keuken	keuze
10	zoontje (son)	zorgen	leerling	lening
11	zondag (sunday)	kelk	mis	mist
12	tante (aunt)	taxi	opa	ober
13	drank (booze)	drab	wiet	wiel
14	eiland (island)	ijzer	plekje	plastic
15	leider (leader)	lijster	voorman	folder
16	stem (voice)	stek	klank	klas
17	schuur (shed)	schurk	hut	hulp
18	mugjes (midges)	muntjes	beestjes	beetjes
19	mensen (people)	meesters	vrouwen	vruchten
20	emmers (buckets)	enkels	bakken	banden
21	arts (doctor)	arm	pil	pijp
22	boete (fine)	boerka	deukje	deurtje
23	kind (child)	kilt	wicht	wilg
24	school	schoot	werk	werf
25	kaartjes (tickets)	kaasjes	plaatjes	planning
26	ruzie (quarrel)	rumba	wondje	wodka
27	vragen (questions)	vrachten	emails	iglo’s
28	training	trailer	cursus	kunststof
29	vrienden (friends)	friezen	meisjes	mijlen
30	keyboard	kiezel	orgel	order
31	gat (gap)	gas	punt	pus
32	schrift (writing system)	schroot	geld	geel
33	bakjes (containers)	ballen	schotels	schommel
34	rol (role)	rok	man	maan
35	stoelen (chairs)	stormen	kasten	kachels
36	ouders (parents)	aura’s	vaders	vazen
37	jaar (year)	jack	maand	maat
38	koffie (coffee)	kogel	water	wandje
39	datum (date)	daling	tijden	tijger
40	tocht (journey)	tolk	reis	rijst
41	broers (brothers)	broek	ooms	oogst
42	straten (streets)	stralen	wegen	wensen
43	schema (scheme)	schetsboek	rooster	rondje
44	poster	pony	foto	foetus
45	fiets (bicycle)	vis	bus	bult
46	poos (while)	pool	uur	urn
47	zon (sun)	zorg	wolk	wol
48	afspraak (appointment)	afweer	meeting	missie
49	tijm (thyme)	tijd	munt	mus
50	zeep (soap)	zeer	schuim	schuit
51	soesjes (pastries)	zoenen	taartjes	taakjes
52	zaal (hall)	zaag	hal	haan
53	huis (house)	huiden	kamers	kano’s
54	hand	haag	tas	tak
55	hemel (heaven)	heling	aarde	aanhef
56	tijdje (while)	teiltje	lifje	liftje
57	kaft (cover)	kamp	rug	ruit
58	dorpje (village)	doosje	huisje	hulpje
59	yoghurt	yoga	boter	bordjes
60	auto (car)	ouwe	wagen	wafels

Acknowledgments

My gratitude is due to many people who contributed to the success of this PhD-project. I will not be able to thank everyone individually. But I am grateful to each of them for everything they did.

First and foremost, I would like to thank my daily supervisor and co-promotor Esther Janse as well as my promotor Mirjam Ernestus. They have always been there for me with advice, support and patience. This dissertation would not have been possible without their help.

I would also like to thank the CLS, IMPRS and GSH for providing financial support and for organizing useful courses. In particular, I would like to thank the IMPRS team Els, Dirkje, and Kevin. Special thanks go to my colleagues in the CLSM for the great atmosphere at the workplace during the years: Annika, Ellen, Huib, Kim, Kimberley, Lisa, Louis, Mark, Martijn, Mybeth, Sascha, Sophie, and Thordis. For helping me to learn how to analyze eyetracking data, I would like to thank Sebastian Sauppe. I also want to thank Mark Noordenbos for his `MATLAB` support during the pupillometry analyses. Many thanks go to the student assistants Ilona and Willemijn for testing participants and for their indispensable help in the setup of the experiments. Thanks also to the study participants. I would also like to express gratitude to the MPI library staff and the MPI Technical Group. You have done amazing work. Furthermore, I would like to thank Gertjan Dingemanse and André Goedegebure for the productive collaboration on the ANL paper.

Thank you Kimberley and Thordis, for being my paranimfs. You have been invaluable in helping me to prepare and organize my defense from abroad. I am indebted to Roeland van Hout, Hugo Quéne, Johan Verhoeven, Willemijn Heeren and Margreet Langereis who formed my reading committee.

The Dutch Organisation for Scientific Research (NWO) funded this research (Grant No. 276-75-009 awarded to Esther Janse), for which I am grateful.

Finally, I thank my wife, family and friends for their love and support.

Curriculum vitae

Xaver Koch was born in 1979 in Dresden, Germany. After completing his studies in speech therapy her worked as speech and language therapist at a neurorehabilitation center. Subsequently, he studied Linguistics and Phonetics at Saarland University, Saarbrücken, where he worked in the laboratory of William Barry and Manfred Pützer. During this time he spent internships at Queen Margaret University (UK) working with Jim Scobbie and at the Max Planck Institute for Cognitive and Brain Sciences in Leipzig working with Corinna Bonhage in Angela Friederici's research group. He received a master's degree from Saarland University in 2012. In the same year he started a PhD project in Esther Janses NWO VIDI project 'What makes a good listener' in Nijmegen with Mirjam Ernestus as PhD promotor. Xaver is currently employed as research fellow at Humboldt University, Berlin.

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*
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 Carrion 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 Ho Lok 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 (LSA). *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. *William Laurence Schuerman*
123. Explorations of beta-band neural oscillations during language comprehension: Sentence processing and beyond. *Ashley Glen Lewis*

124. Influences on the magnitude of syntactic priming. *Evelien Heyselaar*
125. Lapse organization in interaction. *Elliott Michael Hoey*
126. The processing of reduced word pronunciation variants by natives and 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: Studies on sentence production and comprehension in Tagalog and German. *Sebastian Sauppe*
129. Learning from the (un)expected: Age and individual differences in statistical learning and perceptual learning in speech. *Thordis Marisa Neger*
130. Mental representations of Dutch regular morphologically complex neologisms. *Laura de Vaan*
131. Speech production, perception, and input of simultaneous bilingual preschoolers: Evidence from voice onset time. *Antje Stöhr*
132. A holistic approach to understanding pre-history. *Vishnupriya Kolipakam*
133. Characterization of transcription factors in monogenic disorders of speech and language. *Sara Busquets Estruch*
134. Indirect request comprehension in different contexts. *Johanne Tromp*
135. Envisioning Language - An exploration of perceptual processes in language comprehension. *Markus Ostarek*
136. Listening for the WHAT and the HOW: Older adults' processing of semantic and affective information in speech. *Juliane Kirsch*
137. Let the agents do the talking: On the influence of vocal tract anatomy on speech during ontogeny and glossogeny. *Rick Janssen*
138. Age and hearing-loss effects on speech processing. *Xaver Koch*