Listening for the WHAT and the HOW:

Older adults' processing of semantic and affective information in speech

Juliane Kirsch

This research was funded by the European Union's Seventh Framework Program for research, technological development and demonstration under grant agreement no FP7-PEOPLE-2011-290000.

ISBN: 978-90-76203-94-2

Cover: Cora Linez Tattoo & Art and Nina Kurzenhäuser
Printed by: Ipskamp Printing, Enschede

Listening for the WHAT and the HOW:

Older adults' processing of semantic and affective information in speech

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens besluit van het college van decanen
in het openbaar te verdedigen op donderdag 5 juli 2018
om 12.30 uur precies
door

Juliane Kirsch
geboren op 26 mei 1986
te Saarbruecken (Duitsland)

Promotor
Prof. dr. M.T.C. Ernestus

Copromotoren
Dr. O.E. Scharenborg
Dr. E. Janse

Manuscriptcommissie
Prof. dr. A.C.M. Rietveld (voorzitter)
Prof. dr. D. Başkent (Universitair Medisch Centrum Groningen)
Prof. dr. B.W. Schuller (Universität Augsburg, Duitsland)
Dr. K.P. Truong (Universiteit Twente)
Dr. R.J.J.H. van Son (Antoni van Leeuwenhoek ziekenhuis)

Paranimfen
Sara Ahmadi
Mario Ganzeboom

Listening for the WHAT and the HOW:

Older adults’ processing of semantic and affective information in speech

Doctoral Thesis

to obtain the degree of doctor
from Radboud University Nijmegen
on the authority of the Rector Magnificus prof. dr. J.H.J.M. van Krieken,
according to the decision of the Council of Deans
to be defended in public on Thursday, July 5, 2018
at 12.30 hours
by

Juliane Kirsch
Born on May 26, 1986
in Saarbruecken (Germany)

Supervisor
Prof. dr. M.T.C. Ernestus

Co-supervisors
Dr. O.E. Scharenborg
Dr. E. Janse

Doctoral Thesis Committee
Prof. dr. A.C.M. Rietveld
Prof. dr. D. Başkent (University Medical Center Groningen)
Prof. dr. B.W. Schuller (University of Augsburg, Germany)
Dr. K.P. Truong (University of Twente)
Dr. R.J.J.H. van Son (Antoni van Leeuwenhoek ziekenhuis)

Paranymphs
Sara Ahmadi
Mario Ganzeboom

This thesis is dedicated to my grandmothers,

Luise & Herta

Table of contents

Chapter 1

General Introduction

The ability to communicate is crucial for social participation, integration, and interaction. This ability is particularly important for older adults, as they are more likely to suffer from social isolation due to changes in social networks (Dykstra, Van Tilburg, & De Jong, 2005), and social isolation may have a detrimental effect on health and well-being, especially in older adulthood (Bath & Deeg, 2005). A loss of communicative abilities may thus seriously impair quality of life in old age.

A central part of human communication is the recognition of sequences of meaningful units, such as words. The speech signal is often regarded as consisting of segmental and suprasegmental information. Segmental information consists of the segments of speech, i.e., the sounds, a speaker has produced. Segmental information thus contains information concerning *what* a speaker said. Suprasegmental information is information that is not bound to the segments but rather to larger-scale units such as words in the speech signal. An important source of suprasegmental information is prosody. Prosody provides information concerning *how* something is said. Prosodic information is used to differentiate between a statement and a question, can be used to highlight the important parts in a message, and can change the interpretation of segmental information. Take for instance the word “great”. “Great” has a positive connotation. Hence, a speaker telling you in a high-pitched voice that you have done a “great job” probably wants to tell you that you have done something of value. In contrast, “great job” uttered in a lower-pitched, more negative sounding voice may be interpreted as the exact opposite. Relatedly, an important function of prosody is the encoding of information on the emotional status of a speaker. Listeners may hear from someone's voice whether the speaker is happy, angry, or sad. The ability to adequately understand and respond to *what* is said and *how* it is said form an important part of social interaction.

With advanced adult age, understanding what is said and how it is said may be compromised (Dupuis & Pichora-Fuller, 2015), particularly in cognitively and acoustically challenging listening situations, and more so than for younger adults (Schneider, Daneman, & Pichora-Fuller, 2002). This age effect in speech comprehension has been linked to several factors. The two most prominent factors are differences between older and younger adults in auditory acuity and cognitive ability (Akeroyd, 2008; Benichov, Cox, Tun, & Wingfield, 2012; Sommers et al., 2011; Tun, McCoy, & Wingfield, 2009; Van Rooij & Plomp, 1992). This thesis investigates the effects of age and hearing acuity on the processing of what is said and with what emotion something is said. In other words, this thesis investigates the processing of semantic and affective meaning in the speech signal, taking on an individual differences approach.

Speech comprehension

To understand the message intended by a speaker, listeners need to perceive, interpret, and comprehend the speech signal. In normal listening conditions, speech comprehension is efficient and is perceived to be effortless. This efficiency builds on a complex interplay of several peripheral and central processes which are acknowledged by most abstract theories of speech comprehension (e.g., Marslen-Wilson, 1987). Figure 1 shows a visual representation of the supposedly central elements of the speech recognition process, as generally agreed on by several models of spoken word recognition (Luce & Pisoni, 1998; McClelland & Elman, 1986; Norris, 1994).

The basis of speech processing is formed by the input auditory speech signal1 containing both the segmental and suprasegmental information, which enters the ear of the listener (e.g., bat in Figure 1). Upon hearing the speech signal, first, the mental lexicon is accessed, and the acoustic signal is mapped onto stored linguistic knowledge (cf., *Cohort theory*, Marslen-Wilson, 1987; see also Figure 1). During this matching process,

[1] Communication is a multimodal process. It consists of both auditory and visual information. This thesis focusses solely on the auditory information stream in speech communication.

not only the segmental information is considered. All words that partially overlap with the input are then activated simultaneously (e.g., bat, cat, bag; Allopenna, Magnuson, & Tanenhaus, 1998; Gow & Gordon, 1995). Subsequently, all activated words are said to enter a competition phase. The candidate word that best matches the input and the higher-level information, such as the prosodic environment, the syntax, and the context, is selected and recognized (cf., Marslen-Wilson, 1987).

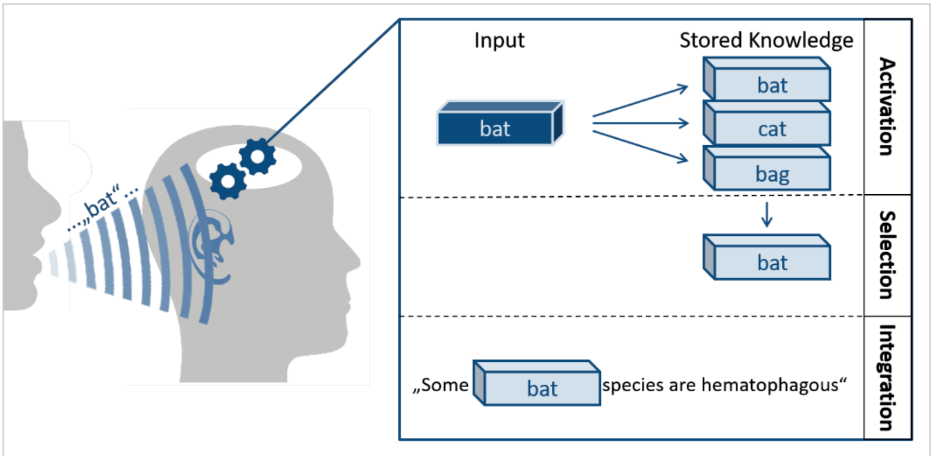

Figure 1. A simplified schematic of spoken word recognition. The acoustic signal activates candidate words which (partially) overlap with stored word representations (Activation). The word candidate which best matches the input and the higher-level information is then selected (Selection) and integrated into the linguistic context (Integration).

Semantic information

One of the reasons for the efficiency of speech comprehension is the way linguistic knowledge is stored in and retrieved from the brain, more specifically the mental lexicon. The mental lexicon contains lexical entries, or words, and additional information of these words, such as their pronunciation and meaning (i.e., semantic information). The mental lexicon is organized as a network of associations rather than a list of isolated words (see Figure 2). According to Collins and Loftus (1975), the incoming speech signal activates a word in the mental lexicon (e.g., *bat* in Figure 1), subsequently activation propagates through the network which means that not only the candidate word which partially matches the input speech is activated but also words that are semantically associated with the candidate words. For the example *bat* in Figure 2, semantic associates such as *vampire*, *night*, *black*, *baseball* would also be activated, and their recognition subsequently facilitated.

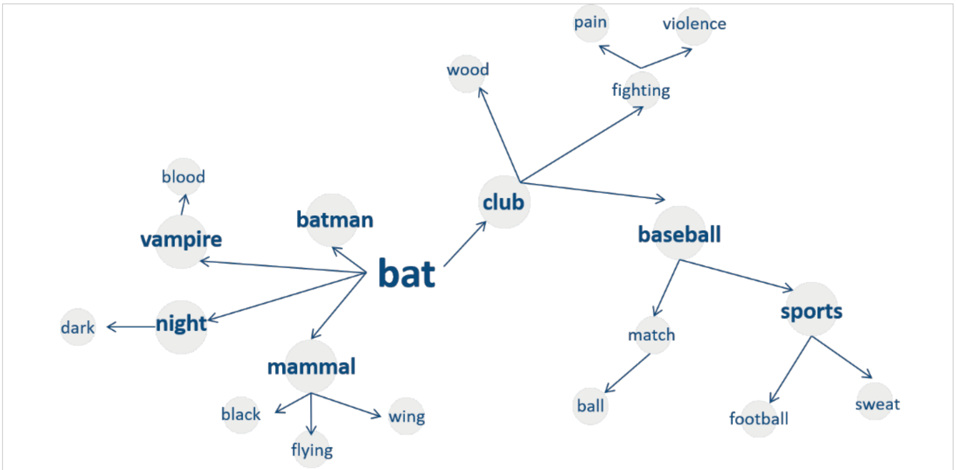

Figure 2. Semantic network of the word "bat".

Words are thus processed faster after hearing words that are semantically associated with them (cf., Swinney, Onifer, Prather, & Hirshkowitz, 1979). This cognitive process is known as semantic priming (Collins & Loftus, 1975; Gow & Gordon, 1995; McKoon & Ratcliff, 2012; Norris, Cutler, McQueen, & Butterfield, 2006; Tabossi, 1996), where the first word is the prime word (*bat* in this example), which primes the semantic associates.

An important factor that affects lexical and semantic processing and hence semantic priming is the ability to perceive an incoming signal. Word forms that are produced in running speech are often produced with less articulatory precision than if they are presented as single words in their citation form. ‘Reduced forms’ like *yeshay* for *yesterday* often occur in conversations. Research has shown that it takes more effort for these ‘reduced’ words to activate their semantic networks than for their unreduced counterparts, not only if the words are presented in isolation (Van de Ven, Tucker, & Ernestus, 2011), but also if they are presented in mid-sentence position which is where these reduced forms occur most frequently (Drijvers, Mulder, & Ernestus, 2016). High-frequency hearing loss, which is prevalent in older adults, compromises clarity of the incoming speech signal. In a study in which signal clarity was reduced by dampening the higher frequencies, which can be interpreted as a rough approximation to hearing loss, Aydelott and Bates (2004) found that a decrease in signal clarity inhibits lexical semantic access. Their findings suggest that low-pass filtered prime words are still recognized by the listener, but semantic facilitation is reduced. Similar results have been obtained with another type of signal degradation (noise-vocoded speech), where N400 amplitude, taken as a signature of integrating words into their semantic context, became more pronounced with improving speech intelligibility (Obleser & Kotz, 2011). These results thus emphasize how reduced signal clarity and increased listening effort may delay semantic comprehension.

Another important factor to the processing of semantic information is suggested to be attention. Listeners may not always want to pay attention to spoken input, or they may have limited cognitive resources to do so. Attention is a cornerstone in the Framework for Understanding Effortful Listening (the FUEL framework) developed by Pichora-Fuller and colleagues (Pichora-Fuller et al., 2016), in that listeners have a limited supply of mental resources (i.e., attention) to be allocated to the task at hand attention may play a role in realistic listening situations, for instance, listening to speech while having a radio turned on in the background. In a visual word recognition experiment, Smith and colleagues (Smith, Bentin,

& Spalek, 2001) found that no semantic activation occurred when attention was explicitly focused at the orthographic level, e.g., due to an attention demanding letter search task. They argue that the remaining attentional resources do not suffice to activate the prime and consequently, no semantic priming occurs. Similarly, Norris and colleagues (Norris et al., 2006) provide evidence that the size of the semantic priming effect may depend on whether the listeners' attention is drawn to (or away from) the prime words, such that the priming effect is larger when attention is drawn to the prime word. These results suggest that prime words are only processed deeply enough to elicit significant semantic priming if the participant's attention is concentrated on the prime. The importance of attention for semantic priming has been backed up by results found with dichotic listening (Dupoux, Kouider, & Mehler, 2003) and divided attention tasks (Otsuka & Kawaguchi, 2007) in which listeners did show priming in the single-task condition, and reduced priming under conditions of divided attention. Importantly, attentional control is affected by aging (Sylvain-Roy & Belleville, 2015; Titone, Prentice, & Wingfield, 2000). In attentionally demanding listening conditions where attention is divided over two tasks, older adults might then benefit less from semantic priming.

Affective prosodic information

Affect in speech is expressed by changes in various acoustic, prosodic parameters. The parameters include pitch (Hammerschmidt & Jürgens, 2007; Mozziconacci, 1998; Rodero, 2011), intensity (Aubergé & Cathiard, 2003; Schröder, Cowie, Douglas-Cowie, Westerdijk, & Gielen, 2001), spectral measures (Schröder, 2006; Tamarit, Goudbeek, & Scherer, 2008), and tempo (Mozziconacci & Hermes, 2000), where pitch plays a central role. As these acoustic parameters are mutually dependent, e.g., speaking louder often goes along with an increase in pitch (cf., Hammerschmidt & Jürgens, 2007), acoustic patterns conveying affect in natural speech may be complex. Listeners make use of these prosodic cues to derive a speaker's emotional state (Banse & Scherer, 1996; Coutinho & Dibben, 2013; Scherer, 2003).

Age has been found to influence perception of affective information in speech, resulting in increased misclassifications of emotion categories. For instance, Paulmann and colleagues (Paulmann, Pell, & Kotz, 2008) showed that young participants were significantly better at recognizing anger, disgust, fear, happiness, and sadness in speech than older adults. These age differences in affect perception may arise early, namely during auditory decoding of speech, or may only arise during processing stages beyond the decoding stage, or both. Evidence for a "processing" rather than an auditory account of the age difference in affect perception comes from a study by Ruffman and colleagues (Ruffman, Henry, Livingstone, & Phillips, 2008). They argue that age differences in emotion perception arise due to age-related structural and functional changes in the "social brain". Structural changes that are likely to affect emotion recognition are for example volume reductions in frontal and temporal brain areas. Brain areas related to the visual and auditory emotion integration (e.g., amygdala) have been reported to be less activated in older adults compared to younger adults when presented with audio-visual emotional, particularly negative emotional stimuli (Lawrence, Calder, McGowan, & Grasby, 2002). Functional changes due to advancing age are brought about by changes in the level of neurotransmitters (such as dopamine and noradrenaline) in the emotion processing regions. In other words, even if older adults can perceive the acoustic cues that signal affect, higher-order processing deficits affecting the processing or uptake of these cues may result in less differentiation of emotional content than in younger adults.

Nevertheless, acoustic cues to affect may sometimes be subtle. The observed age effects in affect perception may also, at least partly, result from older adults' differential auditory processing abilities compared to younger adults. Auditory processing abilities are generally assessed with psychoacoustic tasks in which participants are asked to discriminate between highly similar tones or to detect very short silent gaps in tone or noise stimuli. Older adults, even if normal-hearing, have been reported to be less sensitive to subtle pitch differences (He, Mills, & Dubno, 2007; Mitchell & Kingston, 2014; Souza, Arehart, Miller, & Muralimanohar, 2011), intensity differences (Harris, Mills, & Dubno, 2007), and temporal differences (Anderson, Parbery-Clark, White-Schwoch, & Kraus, 2012; Gordon-Salant & Fitzgibbons, 1999) than younger adults. Apart from these auditory processing differences between age groups, it is still unclear whether, or to what extent, age-related hearing loss may contribute to the age group difference in affect perception. Hearing loss is in fact also associated with elevated thresholds in tasks testing loudness processing abilities (Boettcher, Poth, Mills, & Dubno, 2001; Freigang et al., 2011; Koehnke, Culotta, Hawley, & Colburn, 1995). However, several studies failed to find a direct link between affect perception and individual hearing sensitivity (Lambrecht, Kreifelts, & Wildgruber, 2012; Orbelo, Grim, Talbott, & Ross, 2005), or between auditory processing abilities and emotion categorization performance (Dupuis & Pichora-Fuller, 2015). Still, as the perception of acoustic cues which convey affect in speech may be impacted by (high-frequency) hearing loss (cf., Boettcher et al., 2001), age-related hearing loss might moderate affect perception via the perception of the affect-related acoustic cues. Previous studies either have not included a large-enough range of hearing losses (Lambrecht et al., 2012) or have not related individual hearing loss to the acoustic properties of the stimuli (Orbelo et al., 2005; Dupuis and Pichora-Fuller, 2015). Consequently, the potential role of age-related hearing loss in affect perception is unclear. If the contribution of hearing loss to affect perception is not fully understood, it is also unclear whether the use of hearing aids can sufficiently restore information needed for affect perception in speech.

Aim and research questions

This thesis investigates how age effects and individual differences in hearing ability relate to the uptake and processing of semantic and affective information in older adults' speech comprehension. The aim of this thesis is addressed in three research questions.

Processing of semantic information in language comprehension has been suggested to be modulated by attentional resources. The **first research question** addressed in this thesis asks whether a cognitive load, induced by divided attention, affects semantic activation in speech processing in older adults, and whether this is modulated by individual differences in cognitive and hearing abilities. This question is addressed in Chapter 2 using an experiment where attention is divided over two tasks. Semantic activation was investigated using a continuous lexical decision task, in which response times to target words preceded by semantically related primes were analyzed. The cognitive load was induced using a simultaneous secondary task, i.e., a digit recall task, which continuously taxed working memory. Additionally, several cognitive tests, including cognitive tests measuring attention, such as the Trail Making Test (Reitan, 1958), a digit span task with backward recall (Wechsler, 2008), and digit/symbol coding test (Wechsler, 2008), were administered.

The second and third research questions focus on the processing of affective information. The experimental studies described in Chapters 3 and 4 are both concerned with the **second research question**, which is whether the observed age effects in the perception of affective information can be explained by differences in the use of prosodic cues by younger and older listeners, and if so, whether these age group differences should be attributed to hearing loss. Specifically, the focus is on the impact of mild to moderate hearing loss on the perception of acoustic cues associated with affective prosody. In both affect perception studies the same experimental set-up was used, i.e., the perception of affective information was investigated using an emotion rating task. Using a dimensional approach, two emotion dimensions were assessed separately: the emotion dimension arousal (the extent to which a stimulus is calm or expressive) and the emotion dimension valence (the extent to which a stimulus is positive or negative). The dimensional approach was chosen over the predominantly used categorical approach (Dupuis & Pichora-Fuller, 2015; Isaacowitz et al., 2007; Kiss & Ennis, 2001; Lambrecht et al., 2012; Lima, Alves, Scott, & Castro, 2014; Mitchell, Kingston, & Barbosa Bouças, 2011; Paulmann et al., 2008), where affect is described with concrete terms such as *happy*, *sad*, *angry*, because the latter approach has certain drawbacks. Emotion categories cause arousal and valence information to be confounded, e.g., the category *happy* may describe a positive state of high arousal. Moreover, emotion categories are less controllable in experimental settings as different participants might have different concepts of *happiness*, *sadness, anger,* etc. A more detailed and controlled description of affect can be achieved by the dimensional approach. The dimensional approach was particularly applicable to the current set of studies since affect-related prosodic parameters such as pitch, intensity, and tempo correlate with both arousal and valence (Pereira, 2000; Schröder et al., 2001). Hence, the dimensional approach allows a fine-grained investigation of the use of these parameters in emotion perception. While hearing loss did not directly impact on emotion categorization performance (e.g., Lambrecht et al., 2012; Orbelo et al., 2005), hearing loss might partially affect underlying emotion dimensions. There are hints in the literature (Grant, 1987) that hearing loss might influence the perception of prosodic parameters. Therefore, hearing loss might explain age-related differences in affect perception through its influence on the perception of prosodic parameters. Affect perception, in terms of arousal and valence ratings, is compared between several listener groups with varying age and hearing sensitivity. Crucially, the prosodic parameters cueing affect, age, and hearing sensitivity are related to the arousal and valence ratings for the different listener group.

The stimuli used in the emotion perception studies consisted of conversational rather than acted speech material. Ecological validity is highest in natural affective speech stimuli (Scherer, 2003). The motivation for choosing conversational speech over acted speech comes from the participant population under investigation. To investigate the influence of hearing ability on speech comprehension, older adults with different degrees of hearing loss were investigated. Prototypical acoustic patterns, as frequently encountered in acted speech, with exaggerated frequency contours may be relatively easy to perceive for people with hearing loss (Grant, 1987). Consequently, hearing loss might be less predictive of changes in the perception of affect in acted speech. Responses to conversational and acted speech may therefore differ, with more extreme affect realizations in the latter leading to more extreme responses (Wilting, Krahmer, & Swerts, 2006).

The **third research question** focuses on the effects of more severe hearing loss and hearing aid use in older adults on affect perception (Chapter 5) using the same experimental design as used for the previous research question. More specifically, I ask whether wearing a hearing aid makes listeners more sensitive to subtle differences in affective prosody. Moreover, affect perception of older adults using their hearing aids is compared to that of normal-hearing older adults. Hearing loss as well as reduced auditory processing abilities (such as detecting differences in loudness) may impact on the ability to use acoustic information for the interpretation of the speaker's affective state. Therefore, Chapter 5 not only investigates the effect of hearing aid use on affect rating, but also investigates potential relationships between individual hearing sensitivity and loudness processing ability on the one hand and affect rating on the other.

Outline

This thesis reports four studies. The first research question on the role of attention in semantic processing is addressed in Chapter 2. The second research question on the role of age and hearing loss in affect perception is addressed in Chapters 3 and 4. Chapter 3 reports an age group comparison on the perception of the emotion dimensions arousal and valence in fragments of conversational speech. This chapter focuses on the relationship between participants' affect dimension ratings and acoustic parameters known to be associated with arousal and valence. McCoy and colleagues (McCoy et al., 2005) showed that hearing loss and age only impair speech comprehension once both have reached a critical threshold. This hypothesis was tested in Chapter 4. Where the older adults in Chapter 3 had normal to mildly impaired hearing ability, the older adults in Chapter 4 had slightly worse hearing abilities, i.e., normal to moderately impaired hearing ability. Although some of the participants in Chapter 4 qualified for hearing aids, none were using hearing aids at the time of testing. Additionally, a more robust acoustic correlate of perceived intensity linked to the speaker's vocal effort was added to the prosodic parameters investigated in the previous chapter

The third research question is explored in Chapter 5. The focus of this study is on the arousal dimension as it is more reliably coded in the acoustic signal (cf., Chapter 3 and 4; Pereira, 2000; Schröder, 2001; Schröder, 2006). The influence of hearing aid use on arousal perception is investigated by comparing older hearing-aid wearing adults' perception of arousal while wearing their hearing aid to a condition when they were not. Moreover, the hearing-aid condition was compared to a group of normal-hearing peers. In Chapter 5, I specifically ask whether loudness processing ability modifies arousal ratings and/or the use of prosodic cues to arousal (such as intensity and vocal effort) beyond the effects of hearing loss.

This thesis ends with a general discussion in Chapter 6 which includes a summary of the main results, a discussion with respect to previous findings as well as theoretical implications, and an outlook on future research.

Chapter 2

Semantic Processing of Spoken Words under Cognitive Load in Older Listeners

This chapter is a reformatted version of:
Juliane Schmidt, Odette Scharenborg, Esther Janse (2015). Semantic processing of spoken words under cognitive load in older listeners. Proceedings of the International Congress of Phonetic Sciences, Glasgow, UK.

Abstract

Processing of semantic information in language comprehension has been suggested to be modulated by attentional resources. Consequently, cognitive load would be expected to reduce semantic priming, but studies have yielded inconsistent results. This study investigated whether cognitive load affects semantic activation in speech processing in older adults, and whether this is modulated by individual differences in cognitive and hearing abilities. Older adults participated in an auditory continuous lexical decision task in a low-load and high-load condition. The group analysis showed only a marginally significant reduction of semantic priming in the high-load condition compared to the low-load condition. The individual differences analysis showed that semantic priming was significantly reduced under increased load in participants with poorer attention-switching control. Hence, a resource-demanding secondary task may affect the integration of spoken words into a coherent semantic representation for listeners with poorer attentional skills.

Introduction

In speech comprehension, listeners decode acoustic information in order to access semantic information for the interpretation of the message. Consequently, processing of (target) words that are preceded by a semantically related (prime) word is facilitated, or primed (Collins & Loftus, 1975). As such, semantic priming is evidence that the prime has activated the semantic system.

Processing of semantic information is suggested to be modulated by attention. There is some evidence that the size of the semantic priming effect may depend on whether listeners' attention is drawn to (or away from) the prime words (Norris, Cutler, McQueen, & Butterfield, 2006). This suggests that prime words are only processed deeply enough to elicit significant semantic priming if participants' attention is concentrated on the prime.

Given the evidence for the importance of attention on semantic priming, cognitive load (CL) would be expected to reduce semantic priming, particularly for those with poorer attentional abilities. However, results of previous studies, obtained with student participants, have been inconsistent in whether or not CL decreased semantic priming (cf., Mattys & Wiget, 2011), perhaps due to methodological differences. Individual differences in attentional abilities may particularly be found in a population of older adults, as attentional abilities generally decline with age (Anderson, Craik, & Naveh-Benjamin, 1998; Naveh-Benjamin, Craik, Guez, & Kreuger, 2005), but not affecting all individuals to the same extent.

Apart from attentional factors, speech signal clarity might influence semantic activation. Speech signal clarity has been shown to affect lexical processing and hence semantic priming (Aydelott & Bates, 2004; Van de Ven, Tucker, & Ernestus, 2011). Processing of degraded input constitutes a perceptual load, which may occupy attentional resources that would otherwise be available for further processing of what has been heard

(Rabbitt, 1968, 1991). Particularly, the dampening of spectral information due to (age-related) hearing loss makes speech processing more effortful and may reduce semantic facilitation (Aydelott & Bates, 2004). Prime words in the acoustically degraded condition were recognized in Aydelott and Bates (2004), but processing lagged behind, relative to the clear-speech condition, such that activation had not spread fully to semantic associates.

Given that older adults are expected to present a more heterogeneous sample with respect to hearing acuity and attentional abilities, both related to semantic activation, this study focuses on speech processing by older adults. We first addressed the question whether the presence of CL induced by a dual-task paradigm loading verbal working memory generally decreases semantic activation. Importantly, our design ensured that working memory was continuously taxed and both prime and target were processed. Secondly, we investigated whether individual auditory and cognitive abilities modify the priming effect and the load effect on semantic priming. In addition to attentional and working memory abilities, we also investigated the effect of processing speed as the latter may also play a role in lexical processing (Janse, 2009) and spreading of activation. We expected to find an effect of CL on semantic priming, particularly for participants with poorer auditory and/or poorer cognitive skills.

Material and methods

Participants

Forty-six native Dutch older adults were recruited from the participant pool of the Max Planck Institute for Psycholinguistics in Nijmegen, The Netherlands. None of them wore hearing aids in daily life. Pure-tone (air conduction) thresholds were measured for both ears; the pure-tone average (PTA) of the better ear across 1 kHz, 2 kHz, and 4 kHz[1] was used as an index of hearing acuity (*M* = 22.17 dB HL; *SD* = 10.94). Working memory capacity was defined as the percentage of correct sequences in a digit span task with backward recall (Wechsler, 2008) with visually displayed digit sequences consisting of two to seven digits (Mean accuracy = 47.12%; *SD* = 22.40). Processing speed was assessed by a pencil-and-paper digit/symbol coding test (Wechsler, 2008). The higher the number of recoded symbols within 90 seconds, the faster the participant's processing (Mean number of recoded symbols = 46.46; *SD* = 9.59). The Trail-Making Test (Reitan, 1958) was administered as a measure of attention control, and the quotient of time in seconds the participant needed to complete part B (alternatingly connecting digits and letters) was divided by the time the participant needed to complete part A (just connecting digits in ascending order). A higher quotient (TrailB/TrailA) indicated poorer attention-switching control (*M* = 1.90; *SD* = 0.42). Nine participants were excluded on the basis of their outlier performance on the Trail-Making Test. The final sample consisted of 37 older adults aged between 60 and 84 years (21 females; mean age: 67.1 years, *SD* = 6.1). Participants were paid for their participation.

Material

Primary task

The primary task of the experiment consisted of an auditory lexical decision task. For this task, 72 semantically related word pairs consisting of Dutch nouns were selected. Each pair consisted of a prime and a target word of one to three syllables. Semantic-relatedness scores were

---

[1] In line with other studies on age-related hearing loss (e.g., Humes, 1996), we based our PTA on 1, 2, and 4 rather than on 0.5, 1, and 2 kHz because age-related hearing loss specifically affects the higher frequencies.

Individual hearing thresholds for the better ear (as judged by $PTA_{(1,2,4\ kHz)}$) are presented in Appendix Table A.

retrieved from the Dutch Word Association Database (henceforth: semantic relatedness) (De Deyne & Storms, 2008). We used log2-transformed scores from the “synonym search mode”, which considers the distributional overlap of the association responses of two cue words such that both direct associates and near neighbors are included. Association strength between the members of our set of 72-word pairs varied on a continuum from mildly related (e.g., *snor-wenkbrauw* 'moustache-eyebrow', log2 value of .26) to highly related (e.g., *appel-peer* 'apple-pear', log2 value of .64). As reaction times (RTs) are influenced by word frequency, log-transformed word frequencies of the target words were retrieved (Baayen, Piepenbrock, & Gulikers, 1995) and were entered as a control variable in our statistical analyses.

As priming should be implicit, the words of a pair were presented consecutively for continuous lexical decision and were mixed with fillers to hide their associative relationship. More than twice as many one-to three-syllable filler items (96 Dutch words, 240 phonotactically legal pseudo words) were included. A total of 480 stimuli were split into 24 blocks, consisting of 20 trials each. These blocks were split over the two load conditions. The order of the load conditions and blocks was balanced over 2 different lists.

Secondary task

The secondary task consisted of either variant of a load-inducing digit recall task: a low-load and a high-load condition. The complexity of the load manipulation was varied rather than comparing a load to a no-load condition to ensure that the same strategies were used in both conditions. In the low-load condition, one one-digit number was presented auditorily for recall during lexical decision trials; in the high-load condition, two two-digit numbers were presented auditorily.

In order to investigate whether and how task performance in the secondary (digit recall) task affected performance in the primary task, the difference between recall performance in the high-load and low-load condition (Recall Difference) was calculated for each subject. If participants were less affected by the increased cognitive load, the difference scores should be closer to zero.

Recording

Stimuli for both the primary and secondary task were read out at a normal rate by a male native speaker of Dutch and recorded with a Sennheiser K6 microphone at a sampling rate of 16 bit/44.1 kHz in a sound-attenuated booth.

Procedure

Participants were tested in a sound-attenuated booth and stimuli were controlled by means of E-Prime 2.0 and presented via closed headphones (Sennheiser HD 215). The volume was kept at a constant level (approximately 70 dB SPL). Half of the participants were first presented with 12 blocks in the low-load condition followed by 12 blocks in the high-load condition; for the other half the order of load conditions was reversed. Participants were allowed a short break in between the two load conditions. There were three consecutive phases for each block: digit presentation, auditory lexical decision and digit recall. First, after a blank screen (250 ms), participants heard either a one-digit number (low-load condition) or two two-digit numbers which were separated by a 50 ms pause (high-load condition). Following another blank screen (100 ms), the auditory lexical decision task started. On each trial, auditory presentation of each word was preceded by a fixation cross (500 ms) followed by a blank screen (100 ms). Participants had to decide whether the stimulus was a real Dutch word or not. They were instructed to make their choice as quickly and as accurately as possible using the keys ‘M’ (labeled ‘yes’) or ‘Z’ (labeled ‘no’) on the keyboard. Responses and RTs were measured from stimulus onset until key press. After the key press, the next stimulus was presented after a 1 second inter stimulus interval (ISI). If a participant did not respond within 4500 ms, a new trial started (cf., Van de Ven et al.,

2011) for similar timing parameters). Third, the participants were asked to recall the digits by entering them via the keyboard. After doing so, they proceeded to the next block.

Tests to assess hearing and cognitive skills were administered directly after the main task. The whole experiment session took approximately 60 minutes.

Results

Only responses to correctly identified target words preceded by correctly identified primes were analyzed. Mean accuracy in the auditory lexical decision task was at ceiling [1]in both the low-load (M = 95.5%, SD = 3.8) and high-load conditions (M = 95.1%, SD = 5.6), and did not differ significantly between the two load conditions; $t(36) = .36$, $p = .72$. In the digit recall task, mean accuracy in the low-load condition was high (M = 93.7%, SD = 12.9) and still reasonably high in the high-load condition (M = 73.2%, SD = 18.9). This difference in mean recall accuracy was significant; $t(36) = 6.12$, $p = .001$.

Lexical decision reaction time analysis

First, we investigated whether CL modifies semantic activation. Using linear mixed-effects regression modelling, log-transformed RTs (measured from auditory word onset) were entered as the dependent variable. Load condition (CL) and semantic relatedness (SemRel) were entered as fixed effects. Word frequency (per million words), word duration in ms (Duration target word), RT on the previous trial, block number, and trial number (within a block) served as control variables. Crucially, we tested for an interaction between load condition and semantic relatedness. We also allowed for the possibility that the load

[1] The exclusion criteria mentioned in the participant section (i.e., outlier performance on the Trail Making Test) also excluded participants with an overall auditory lexical decision accuracy < 80%.

effect might decrease over trials by including an interaction between load and trial. Continuous variables (such as SemRel) were centralized and the low-load condition was mapped on the intercept. As the effect of CL varied across participants, a random slope for load condition per participant was added to the best-fitting model.

Table 1: General model of the linear mixed-effects regression RT analysis

Fixed effects	β	*t*	
Intercept	6.99	310.30	***
CL	0.04	3.29	**
SemRel	-0.25	-2.40	*
Block number	-0.01	-3.43	***
Trial number	-0.00	-2.79	**
Previous RT	0.00	7.87	***
Duration $_{\text{target word}}$	0.00	8.26	***
CL × SemRel	0.14	1.68	+

Notes. *** $p < .001$, ** $p < .01$, * $p < .05$, + < .1

The general model (Table 1) showed a significant effect for CL. Moreover, there was a significant effect for semantic relatedness: target responses were facilitated when they were preceded by more strongly associated primes. These two findings show that RTs were sensitive to our load and semantic relatedness manipulations. Importantly, the interaction between CL and SemRel just missed significance: target facilitation only tended to be decreased in the high-load condition.

Individual differences

The second analysis investigated whether individual listener abilities modify lexical activation and the CL effect on lexical activation (Table 2).

The set-up of this analysis was similar to that of the first analysis, but now individual differences measures were added as fixed effects (all mean-centered) to our initial model: hearing sensitivity (PTA), cognitive processing speed (digit/symbol coding), attention-switching control (Trail-making task, TMT), working memory capacity (digit span) and the individual load effect on digit recall (Recall Difference). We tested whether the individual measures interacted with the load effect on performance and whether they modified a possible interaction between load and semantic relatedness.

Table 2: Individual differences model of the linear mixed-effects regression RT analysis

Fixed effects	β	*t*	
Intercept	6.99	308.15	***
CL	0.04	3.61	**
SemRel	-0.26	-2.42	*
Block number	-0.01	-3.47	***
Trial number	-0.00	-2.85	**
Previous RT	0.00	7.85	***
Duration $_{\text{target word}}$	0.00	8.26	***
TMT	-0.01	-0.13	
Recall Difference	-0.00	-1.16	
CL × SemRel	0.14	1.67	$^{+}$
CL × TMT	-0.01	-0.46	
SemRel × TMT	-0.28	-2.03	*
CL × Recall Difference	0.00	2.99	**
CL × SemRel × TMT	0.60	3.05	**

Notes. *** $p < .001$, ** $p < .01$, * $p < .05$, $^{+} < .1$

The best-fitting individual differences model replicated the effects of load and semantic relatedness and the marginally significant interaction between CL and semantic relatedness (CL × SemRel). The interaction between CL and Recall Difference was significant, i.e., older adults' lexical decision performance was more impacted by the increased load if they were also more affected by increased load in their digit recall. Second, in the low-load condition, those with poorer attention-switching control actually showed stronger semantic priming (SemRel × TMT) than those with better attention-switching skills. Importantly, however, under increased CL, participants with poorer attention-switching control showed significantly reduced semantic priming relative to the low-load condition (CL × SemRel × TMT).

General discussion

The question addressed in this study is whether the presence of a cognitive load (CL) modulates semantic priming, particularly for participants with poorer hearing or cognitive abilities. Our paradigm continuously taxed participants' working memory during the primary task, which is in contrast to earlier dual-tasking studies (Otsuka & Kawaguchi, 2007; Smith, Bentin, & Spalek, 2001) where two tasks had to be performed in succession. Furthermore, unlike Mattys and Wiget (2011), both our primary and secondary tasks taxed verbal working memory and were presented in the same auditory modality.

The results of our general analysis showed that significant semantic priming was found, as well as a clear effect of load on response times. Importantly, the hypothesized reduction of the priming effect in the high-load condition, compared to the low-load condition, did not reach significance. These results are similar to those of Mattys and Wiget (2011) but differ from those of Otsuka and Kawaguchi (2007), who found a significant reduction of the priming effect under divided attention, which they attributed to the cognitive load induced by their second task. This effect of cognitive load on semantic priming may, however, also be due to their experimental design. A prerequisite for semantic priming to occur is that prime words are actually processed (cf., Norris et al., 2006). However, participants in Otsuka and Kawaguchi (2007) only had to remember the pitch of the probe tone that was presented concurrently with the visual prime. Hence, participants may have opted to ignore the lexical content of the prime word, thereby cancelling the priming effect. In our set-up, ignoring the lexical content of the word was not an option, since participants had to decide on the lexical status of both the prime and the target (i.e., a continuous lexical decision task) which ensured processing of the prime. Nevertheless, no strong effect of cognitive load on semantic priming was found.

In a second analysis, we investigated effects of listener abilities, such as hearing sensitivity. Previous research has shown that perceptual load caused by degraded input, such as reduced (Van de Ven et al., 2011) or low-pass filtered speech (Aydelott & Bates, 2004), may hamper semantic activation in younger adults. However, we did not find an effect of hearing sensitivity on semantic activation, nor did it interact with CL. This may be related to the fact that our participants still had reasonably good hearing so that the perceptual load was still manageable without employing additional resources.

Working memory and processing speed were not found to play a role for lexical activation while individuals with poorer attention-switching control showed relatively stronger semantic facilitation in the low-load condition. We can only speculate that participants with poorer attention-switching control may have spent extra effort on the low-load condition. However, in the high-load condition, these participants were overtaxed, such that they were less able to process the prime deeply and quickly enough. These results confirm the attention modulation hypothesis (Smith et al., 2001), i.e., semantic priming depends on attention allocated to primes.

In realistic listening conditions, two tasks that compete for attentional resources are frequently encountered. This study suggests that such a secondary task or distraction may affect the integration of words into a coherent semantic representation, but only for participants with poorer attentional skills.

Chapter 3

Age, Hearing Loss and the Perception of Affective Utterances in Conversational Speech

This chapter is a reformatted version of:
Juliane Schmidt, Esther Janse, Odette Scharenborg (2014). Age, hearing loss and the perception of affective utterances in conversational speech. Proceedings of Interspeech 2014, Singapore.

Abstract

This study investigates whether age and/or hearing loss influence the perception of the emotion dimensions arousal (calm vs. aroused) and valence (positive vs. negative attitude) in conversational speech fragments. Specifically, this study focuses on the relationship between participants' ratings of affective speech and acoustic parameters known to be associated with arousal and valence (mean F0, intensity, and articulation rate). Ten normal-hearing younger and ten older adults with varying hearing loss were tested on two rating tasks. Stimuli consisted of short sentences taken from a corpus of conversational affective speech. In both rating tasks, participants estimated the value of the emotion dimension at hand using a 5-point scale. For arousal, higher intensity was generally associated with higher arousal in both age groups. Compared to younger participants, older participants rated the utterances as less

aroused, and showed a smaller effect of intensity on their arousal ratings. For valence, higher mean F0 was associated with more negative ratings in both age groups. Generally, age group differences in rating affective utterances may not relate to age group differences in hearing loss, but rather to other differences between the age groups, as older participants' rating patterns were not associated with their individual hearing loss.

Introduction

The speech signal does not only contain information on what has been said, but also contains information on how the speaker feels about the content. Affective prosody enhances emotional information processing (Pell, Jaywant, Monetta, & Kotz, 2011) and prosodic acoustic cues are crucial for the correct interpretation of certain affective expressions, e.g. irony (Cheang & Pell, 2008). Affective information can be described in several ways. The categorical approach, meaning that concrete terms such as happy, sad, neutral, bored, or angry are used to describe different affective expressions, seems to be predominant in emotion perception research on older populations (Isaacowitz et al., 2007; Kiss & Ennis, 2001; Orbelo, Grim, Talbott, & Ross, 2005). However, this approach has certain drawbacks as these discrete and rather static concepts do not capture emotion blends. Further, they may bias the responses of participants, or may not be consistent with the participant's own interpretation of a particular affective state (Scherer, 2003, 2005). A more detailed description of affect can be achieved by a dimensional and more continuous approach. Here, emotions are plotted into a two or three-dimensional space, where the most frequently used axes are arousal (calm-aroused) and valence (negative-positive) (Scherer, 2005). Several concepts have been proposed as a third dimension. Examples are tension, control, potency (Scherer, 2005), or dominance (Grimm, Kroschel, & Narayanan, 2008). Compared to the categorical approach, the dimensional approach disentangles the relative contribution of each dimension to emotion categories.

Affect in speech is expressed by physiological changes in articulation, resulting in a unique acoustic pattern for every type of affect expressed in speech (Banse & Scherer, 1996; Scherer, 1986, 2003). Certain acoustic parameters have been shown to correlate with specific emotion dimensions. Mean F0, mean intensity, and speech rate have been investigated most often. For arousal, Pereira (Pereira, 2000) reported higher mean F0 and increasing mean intensity with higher degrees of arousal for both male and female voices. This is supported by Schröder and colleagues (Schröder, Cowie, Douglas-Cowie, Westerdijk, & Gielen, 2001), who further report longer phrases and shorter pauses for aroused utterances. For valence, weaker correlations with acoustic measures were reported. Moreover, higher F0 for more positive utterances (for male speakers) (Pereira, 2000), increasing intensity and longer pauses for more negative utterances (Schröder et al., 2001) were reported.

Aging has been shown to result in changes in the perception of prosodic features (such as intensity, speech rate and F0), particularly of F0 (Souza, Arehart, Miller, & Muralimanohar, 2011). Age has indeed been found to influence perception of affective information in speech (Kiss & Ennis, 2001; Orbelo, Grim, Talbott, & Ross, 2005; Paulmann, Pell, & Kotz, 2008). For instance, Paulmann et al. (Paulmann et al., 2008) showed that young participants were significantly better at recognizing anger, disgust, fear, happiness, and sadness from prosodic information than older adults.

Moreover, hearing loss aggravates auditory processing difficulties for some of the acoustic parameters mentioned, e.g., intensity and speech rate (Cox, McCoy, Tun, & Wingfield, 2008). However, as older participants in (Kiss & Ennis, 2001; Souza et al., 2011) had near-normal hearing, the role of differences in individual hearing loss in this age effect on perception of affective information remains unclear. So far, these age effects have not been attributed to hearing loss, but this might be due to several methodological reasons. First, hearing loss was either not assessed properly in those studies (Kiss & Ennis, 2001; Paulmann et al., 2008), or it was not related to the acoustic parameters of the stimuli (Orbelo et al., 2005). Second, all three studies (Kiss & Ennis, 2001; Orbelo et al., 2005; Paulmann et al., 2008) used stimuli that were acted or artificial. Acted stimuli, however, might not reflect the acoustic details needed for a correct perception and classification of affect in everyday speech (Scherer, 1986). Though there are also disadvantages to natural speech material in emotion research (i.e., small number of speakers, short utterances, and poor recording quality), ecological validity is highest in natural speech stimuli (Scherer, 2003). Acted speech is likely to be overacted, particularly the acoustic cues for arousal (Scherer, 1986), resulting in the use of a more prototypical acoustic expression. This may lead to a more extreme realization of affective prosody in acted speech (Wilting, Krahmer, & Swerts, 2006). For instance, acted anger, for which mean intensity would be a prominent feature (Juslin & Laukka, 2003; Scherer, 1986), may be relatively easy to perceive if overdone, even for people with hearing loss. In natural speech, however, affective information is cued much more subtly (Scherer, 1986). Taken together, these findings are not conclusive as to whether hearing loss may influence the perception of affect in natural speech.

In this study, the role of age and hearing loss in the perception of affective utterances is investigated. We combine three aspects of affective speech perception that have not been combined in the studies described above: the perception of affective utterances is investigated by using a dimensional approach, by using natural (i.e., non-acted) speech stimuli, and by linking acoustic parameters directly to an individual's hearing loss. The focus is on arousal and valence because these emotion dimensions are used most consistently across studies. The first research question that we address is whether younger and older listeners differ in the perception of affect and in the way they make use of the corresponding acoustic parameters. This is investigated by comparing the associations between acoustic parameters and the affective ratings of the two age groups. The two age groups are compared using two separate one-dimensional affective rating tasks, one for arousal and one for valence. The second research question asks whether and how differences in hearing loss among the older adults impact their affective ratings.

Material and methods

Participants

Two groups of 10 native German participants were recruited (50% male participants in each group). The younger group consisted of students from Saarland University in Saarbruecken (age: M = 25.0; SD = 2.0; range: 22 – 28 years) and the older group was recruited from the greater area of Saarbruecken (age: M = 59.7; SD = 5.6; range 53 – 68 years). None of the participants used a hearing aid in daily life. Participants' hearing was assessed prior to testing by a professional hearing aid audiologist. Pure tone thresholds were retrieved for 0.25, 0.5, 0.75, 1, 1.5, 2, 3, 4, 6, and 8 kHz for both ears. The average of both ears was used for the analysis, as the degree of hearing loss did not differ significantly between the left and the right ear. Mann-Whitney U Test for independent samples showed that the younger group had significantly better hearing in the higher frequencies, i.e., from 1 kHz up to 8 kHz (ps < .05) than the older group. Individual pure-tone average (PTA) over participants' thresholds at 1, 2, and 4 kHz[1] over both ears was entered as an index of hearing loss in the analyses below. Mean PTA for the younger listener group was 3.61 dB HL (SD = 3.16, range: 1.67 – 13.33 dB HL) and for the older listener group was 16.33 dB HL (SD = 8.98, range: 6.67 – 31.67 dB HL).

Speech material

The stimuli were taken from the audio-only section of the audio-visual "Vera am Mittag" corpus (henceforth: VAM corpus) for authentic and affectively colored conversational speech (Grimm et al., 2008). The VAM audio corpus consists of 1018 affective utterances divided into two

---

[1] In line with other studies on age-related hearing loss (e.g., Humes, 1996), we based our PTA on 1, 2, and 4 rather than on 0.5, 1, and 2 kHz because age-related hearing loss specifically affects the higher frequencies.

Individual hearing thresholds for the better ear (as judged by $PTA_{(1,2,4\ kHz)}$) are presented in Appendix Table B.

subsets. Utterances were taken from a German TV talk show. The first subset (VAM-I) consists of 499 utterances produced by 19 different speakers (4 m/15 f). The second subset (VAM-II) consists of 519 utterances by 28 speakers (7 m/21 f). Moreover, the corpus provided mean affective ratings for the degree of arousal and valence for each utterance. These affective ratings were collected by means of a pictorial 5-point-scale (Lang, 1980), consisting of five line drawings of a human figure. Each figure expresses a different degree of arousal or valence through changes in attributes, i.e., indication of tremble or change of facial expression. A numeric value was attached to each point, ranging in 0.5 steps from -1 (very calm/very negative) to 1 (very aroused/very positive). Each utterance was evaluated by a group of presumably younger adults (N = 17 for VAM-I, N = 6 for VAM-II, their age is not documented). Their mean ratings per stimulus are treated as ground truth in our analysis.

According to the ground truth, the VAM corpus provides a good coverage of the emotional space (arousal range: -0.83 – 1.00; valence range: -0.80 – 0.77). However, due to the discussion topics within this TV format (relationship crises, jealousy, fatherhood questions, etc.), the emphasis within the corpus was found to be on neutral to more negative emotions (Grimm et al., 2008).

Subsets for the arousal and valence rating tasks

Stimuli were selected from both VAM-I and VAM-II. In order to not confuse or bias participants, only one emotion dimension per task was rated. Hence, we created two separate stimulus sets: one for arousal and one for valence that complied with the following three criteria. First, as the temporal window for information integration is limited (Pöppel, 2004), utterance duration had to be shorter than three seconds. Second, when interpreting an utterance, both verbal (what is said) and non-verbal (how something is said) information is used (Pell et al., 2011). In fact, the semantic meaning may change the emotional content of an utterance, e.g., when non-verbal information is negative while the verbal information is positive, as in sarcasm (Cheang & Pell, 2008). As we were exclusively interested in the non-verbal information, utterances had to be as semantically neutral as possible to minimize semantic interference (e.g., Du bist der Vater, 'You are the father'). Utterances were therefore presented in written form to three independent evaluators who rated whether these sentences were semantically neutral. Only the utterances labeled as neutral by at least two of the three raters were included in the final stimulus sets. Third, utterances had to have a ground truth value for either arousal or valence that was close to the value of the five points on the scale (1, 0.5, 0, -0.5, -1). In order to familiarize participants with the task, four additional utterances per rating task served as practice trials.

The final item set for the arousal rating task consisted of 9 utterances from VAM-I and 15 utterances from VAM-II (17 different speakers in total; 3 m/14 f; rating range: -0.66 – 0.94). The item set for the valence rating task included 7 utterances from VAM-I and 11 utterances from VAM-II (15 different speakers in total; 4 m/11 f; rating range: -0.80 – 0.77). There was an overlap of two utterances between the item sets; hence, two stimuli were rated for both dimensions.

Procedure

Ratings of degree of arousal and valence by a group of younger and a group of older adults were collected with a simple pen-and-paper version of the pictorial rating tool that was used in the VAM corpus (Lang, 1980): five line drawings of a figure depicted five states along the dimension of either arousal (calm – expressive) or valence (frowning – smiling).

Prior to each rating task, the emotion dimension at hand was explained to the participant. Further, the pictorial rating tool was introduced by describing the meaning of each point on the scale. Participants' attention was particularly drawn to the changing attributes of the figure. Finally, it was emphasized that listeners could give only one rating per trial and that they should indicate their choice by marking the figure on the scale. Additionally, written instructions were provided and there was the possibility to ask questions.

Both age groups completed the arousal rating task first, followed by the valence rating task, with a short break in between the two tasks. Participants were seated in a sound-attenuated booth and heard the utterances via closed headphones connected to a laptop. Each utterance was presented twice, i.e. as two separate trials. The order in which stimuli were presented was randomized for each participant, using the experimental program SCAPE (Grabowski & Bauer, 2004). Participants would always start with the practice trials. The utterances were played at the same fixed volume to both age groups. Listeners were allowed to listen to each trial several times before making a decision. For the arousal rating task, 29% of the utterances were listened to more than once, and for the valence ratings task, 20% of the utterances were repeated. Tasks were completed in the participant's own pace but completing both tasks did not take more than 30 minutes.

Results

First, we will report the acoustic measurements of the stimuli for both tasks and their relation to the ground truth to demonstrate that the acoustics were related to the two emotion dimensions. Second, to compare the age groups' ratings, we used linear mixed effects regression analyses. As each utterance was rated twice by the participants, we calculated the average rating per stimulus. The initial model allowed for interactions between age group and all acoustic parameters (with stimulus and participant as random effects). We arrived at the best fitting model by a stepwise exclusion of interactions and predictors with the highest non-significant *p*-values. Next, we investigated whether individuals' hearing loss was associated with their rating of affective information. Therefore, we checked for interactions between hearing loss and the acoustic parameters in the older adults' data. The model fitting procedure was identical to the one used in the group comparison.

Acoustic measurements section

Mean F0 and mean intensity for each utterance were measured using Praat (Boersma & Weenink, 2011). Articulation rate was calculated by dividing the number of syllables by file length minus pauses (i.e., pauses longer than 100 ms).

For arousal, we found strong positive correlations for mean F0 and mean intensity with the ground truth (VAM Value). Hence, higher mean intensity and higher mean F0 is associated with higher levels of arousal. The correlation between articulation rate and the ground truth was not significant. Moreover, there were no significant correlations between the ground truth for valence (VAM Value) and any of the acoustic parameters (see Table 1).

Table 1. Pearson correlation coefficients between acoustic predictors and reference ratings for arousal and valence.

		Mean Intensity	Articulation Rate	VAM Value
Arousal	Mean F0	.80 ***	- .38	.71 ***
	Mean Intensity	–	-.42 *	.91 ***
	Articulation Rate	–	–	-.38
Valence	Mean F0	.67 ***	- .16	-.35
	Mean Intensity	–	- .21	.06
	Articulation Rate	–	–	.20

Notes. * $p < .05$, *** $p < .001$

Rating analysis

Figure 1 shows participants' ratings (y-axis) compared to the ground truth (x-axis). Star symbols show the mean of the younger participants' ratings for each individual utterance, and triangles show the mean of the older participants' ratings for each individual utterance.

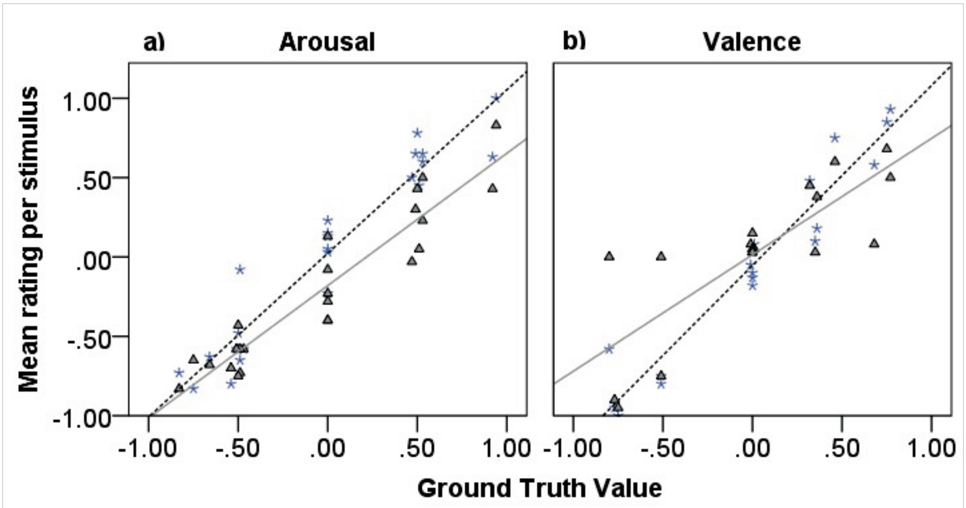

Figure 1. Mean ratings per utterance of the younger (star symbols, dashed lines) and older participants (triangle symbols, solid lines) as a function of ground truth values.

Fit lines have been added, where the dashed line depicts the fit line through the younger participants' ratings, and the solid line depicts the fit line through the older participants' ratings. Figure 1a shows the results for arousal, and Figure 1b for valence.

Analysis of arousal

Figure 1a shows that younger participants are more in agreement with the ground truth than the older participants, as the latter group diverges more from the diagonal. We carried out a statistical analysis to investigate whether this age group difference is significant, and to investigate the influence of the acoustic parameters on the ratings. The Arousal panel in

Table 2 displays the best fitting models for both the age group comparison and the separate analysis of the older adults group.

Table 2. Final models of the linear mixed-effects regression rating analyses.

		Age Group Comparison			Older Adults		
		β		***t***	**β**		***t***
Arousal	Group	-0.20	**	-3.27	–		–
	Mean intensity	0.11	***	10.49	0.09	***	7.90
	Group × Mean Intensity	-0.03	***	-3.95	–		–
Valence	Group	0.06		1.03	–		–
	Mean intensity	0.08		1.70	–		–
	Mean F0	-0.01	*	-2.41	-0.00	**	-3.42
	Group × Mean Intensity	-0.06	***	-4.87	–		–

Notes. * $p < .05$, ** $p < .01$, ***$p < .001$

The statistical analysis revealed significant effects of mean intensity and age group on perceived arousal and an age group by mean intensity interaction. This implies that the higher the stimulus' mean intensity was, the more aroused younger participants (mapped on the intercept) rated the utterance. Older adults generally rated the utterances as less aroused, and showed a smaller effect of intensity on their ratings.

As in the age group comparison, the analysis of the older adults' data only showed an effect of intensity on perceived arousal. Thus, higher intensity is perceived as more aroused among the older group. There was no effect of hearing loss.

Analysis of valence

Figure 1b shows that stimuli on both ends of the ground truth scale are rated as less extreme by older participants compared to younger participants. To investigate this, we carried out a similar statistical analysis as for arousal. The best fitting models are displayed in the Valence panel in Table 2.

The age group comparison for valence showed a significant effect of mean F0 and a significant interaction between age group and mean intensity. Hence, higher mean F0 led to more negative ratings in both age groups. While younger adults (mapped on the intercept) do not use mean intensity to rate valence, older adults seem to interpret increasing intensity as more negative.

Analogous to the analysis of the arousal task data, the data of the older adults were analyzed separately to investigate the effect of hearing loss. To that end, hearing loss was allowed to interact with the acoustic predictors. Results showed a significant effect of mean F0: higher mean F0 yields more negative ratings of valence. Though older adults seem to use intensity differently than younger adults in the age group analysis above, the effect of mean intensity is not strong enough to surface in the subset analysis here. Moreover, hearing loss did not affect valence ratings, nor did it interact with the acoustic parameters.

General discussion

This study was set up to investigate possible age effects in the perception of affect in speech in relation to several acoustic parameters of the stimulus. In contrast to earlier studies on age differences in affect perception, we combined three methodological aspects to investigate the perception of emotional content, i.e., linking the perception of acoustic parameters with individual hearing loss, conversational (rather than acted) speech, and a rating of emotional dimensions, rather than categorical classification.

The expression of affect has been related to acoustic parameters, which listeners use to interpret affect (Banse & Scherer, 1996). In our study, we investigated the relation between acoustic parameters and affective ratings in two ways. First, we investigated how the acoustic parameters in our stimuli correlated with the ground truth affective ratings that came with the corpus materials. Second, these acoustic parameters were entered as predictors of participants' ratings of the conversational stimuli.

Arousal stimuli showed positive correlations between the ground truth arousal ratings from the corpus and both mean F0 and mean intensity. These correlations are comparable to correlations reported in the literature (Pereira, 2000; Schröder et al., 2001). The association between intensity and arousal is further supported by participants' ratings. Our results showed that older adults perceived aroused utterances as less aroused than younger adults, and the effect of intensity on arousal ratings was smaller compared to its effect on younger adults' ratings. This finding agrees with the results of previous studies. Paulmann et al. (Paulmann et al., 2008), for instance, found that older adults classified a stimulus more often as sad when fear was the intended emotion and more often as happy when pleasant surprise was the intended emotion. If the positions of these emotions on the arousal axis are considered (e.g., Scherer, 2005), older adults more often choose the term that is linked to the less aroused emotion (sad, happy) while younger adults prefer the term that is related to the more aroused term (fear, pleasant surprise).

For valence, there were no significant correlations between the ground truth and the acoustic measures investigated in this study. This is not surprising, given that correlations between valence and acoustic parameters found in larger item sets were also less strong as compared to arousal (Pereira, 2000; Schröder et al., 2001). In line with the ground truth evaluator panel (see correlations with VAM ratings in Table 1), we did not find any evidence that the younger adults in our study based their valence ratings on mean intensity, nor on articulation rate. Older adults, however, seem to make more use of mean intensity when rating valence than younger adults though the intensity effect does not reach significance in the subset analysis of the data of the older adults. In addition, our data suggests that mean F0 plays a crucial role for both age groups when rating valence. In other words, independent of age, higher mean F0 is associated with more negative utterances. This result is in opposition to the finding of Pereira (Pereira, 2000), but note that her result held for male talkers only (the majority of our talkers being female).

Our second question concerned the impact of hearing differences among the older adults on their ratings of affect. All in all, age effects on affect perception seemed to outweigh hearing effects. Older age led to smaller-sized intensity effects in the affect dimension. However, the older adults' data do not provide any indication that hearing loss may account for this age effect. This is in line with the findings of Orbelo et al. (Orbelo et al., 2005) who also did not find that hearing measures predicted affective ratings. Note, that our sample size was small (10 participants per age group) and participants' hearing was still rather good. In order to better account for possible effects of hearing loss, future work should include more participants representing a broader range in hearing loss.

Conclusion

Taking a dimensional approach, we tried to link acoustic variation as found in natural conversational speech stimuli to ratings of affect in younger and older adults. Our results show age effects on affect perception, and show that older age led to smaller-sized effects of acoustic differences on affective ratings. No effect of hearing loss on affective rating was observed in this older adult sample. Future research should aim at including more participants covering a broader range of hearing loss to obtain a better picture of how hearing loss may influence perception of affect in conversational speech.

Chapter 4

Perception of Emotion in Conversational Speech by Younger and Older Listeners

This chapter is a reformatted version of:
Juliane Schmidt, Esther Janse, Odette Scharenborg (2016). Perception of emotion in conversational speech by younger and older listeners. Frontiers in Psychology, 7, 781.

Abstract

This study investigated whether age and/or differences in hearing sensitivity influence the perception of the emotion dimensions arousal (calm vs. aroused) and valence (positive vs. negative attitude) in conversational speech. To that end, this study specifically focused on the relationship between participants' ratings of short affective utterances and the utterances' acoustic parameters (pitch, intensity, and articulation rate) known to be associated with the emotion dimensions arousal and valence. Stimuli consisted of short utterances taken from a corpus of conversational speech. In two rating tasks, younger and older adults either rated arousal or valence using a 5-point scale. Mean intensity was found to be the main cue participants used in the arousal task (i.e., higher mean intensity cueing higher levels of arousal) while mean F0 was the main cue in the valence task (i.e., higher mean F0 being interpreted as more negative). Even though there were no overall age group differences in arousal or valence ratings, compared to younger adults, older adults responded less strongly to mean intensity differences cueing arousal and

responded more strongly to differences in mean F0 cueing valence. Individual hearing sensitivity among the older adults did not modify the use of mean intensity as an arousal cue. However, individual hearing sensitivity generally affected valence ratings and modified the use of mean F0. We conclude that age differences in the interpretation of mean F0 as a cue for valence are likely due to age-related hearing loss, whereas age differences in rating arousal do not seem to be driven by hearing sensitivity differences between age groups (as measured by pure-tone audiometry).

Introduction

Accurate emotion recognition is a crucial component of successful social interaction (Blair, 2003). One modality in which affective information is conveyed is speech. Affect in speech manifests itself in differences in prosodic, acoustic patterns, which are used by listeners to derive the emotion intended by the speaker (Banse & Scherer, 1996; Coutinho & Dibben, 2013; Scherer, 2003). Studies have shown that the perception of affect is influenced by age (Mitchell, Kingston, & Barbosa Bouças, 2011; Orbelo, Grim, Talbott, & Ross, 2005; Paulmann, Pell, & Kotz, 2008). For instance, Paulmann and colleagues (2008) have shown that young adults are significantly better at recognizing the emotion categories anger, disgust, fear, happiness, and sadness from prosodic, acoustic information than middle-aged adults. Middle-aged adults in turn outperform older adults in recognizing emotion categories from affective prosody (Kiss & Ennis, 2001). The current study investigates where this age difference originates.

Important acoustic cues for affect perception include pitch, intensity and articulation rate. Pitch is considered the most telling component of affective prosody (Hammerschmidt & Jürgens, 2007; Mozziconacci, 1998; Rodero, 2011). Emotions with higher levels of arousal, such as excitement, fear, and anger, have been shown to have higher mean F0 (Mozziconacci, 1998; Schröder, 2006). Other acoustic cues that signal to affective prosody are temporal aspects (Mozziconacci & Hermes, 2000), intensity (Aubergé & Cathiard, 2003; Schröder, Cowie, Douglas-Cowie, Westerdijk, & Gielen, 2001), and spectral measures (Schröder, 2006; Tamarit et al., 2008). Importantly, these acoustic parameters are mutually dependent in speech, e.g., spectral measures such as spectral slope reflect the energy distribution over the spectrum and correlate highly with intensity (cf., Banse & Scherer, 1996). Moreover, intensity shows a strong positive correlation with pitch (cf., Hammerschmidt & Jürgens, 2007). As a consequence, acoustic patterns conveying affect in speech may be complex.

Aging affects the perception of these acoustic cues. Older adults, even those without hearing loss, have been reported to be less sensitive to pitch differences (He, Mills, & Dubno, 2007; Mitchell & Kingston, 2014; Souza, Arehart, Miller, & Muralimanohar, 2011), intensity differences (Harris, Mills, & Dubno, 2007), and temporal differences (Anderson, Parbery-Clark, White-Schwoch, & Kraus, 2012; Gordon-Salant & Fitzgibbons, 1999) than younger adults. Given that these acoustic cues (pitch, intensity, and tempo) have been argued to convey affect in speech (Banse & Scherer, 1996), it might be hypothesized that the observed age effects in affect perception have their origins in the difference in the perception of acoustic cues compared to younger adults. However, few researchers have looked into the relationship between the use of affect-related acoustic information and age-related differences in affective prosody perception. An exception is a study by Lima and colleagues (Lima, Alves, Scott, & Castro, 2014) who investigated affect perception in vocalizations without verbal content.

Additionally, many other causes have been proposed to explain the apparent age difference in verbal affect perception. Examples are general age differences in cognitive abilities, emotion regulation, and personality (Lambrecht, Kreifelts, & Wildgruber, 2012; Lima et al., 2014; Orbelo et al., 2005). However, these cognitive or personality measures revealed either no effect (Lambrecht et al., 2012; Lima et al., 2014) or only a marginal effect (Orbelo et al., 2005) on differences in affect perception among participants. Prosodic emotion perception may be impaired at an auditory processing level (Mitchell & Kingston, 2014), including hearing loss. As the perception of acoustic cues such as intensity differences may be impacted by (high-frequency) hearing loss (cf., Boettcher, Poth, Mills, & Dubno, 2001), age-related hearing loss might moderate affect perception. Importantly, previous studies either have not included a large-enough range of hearing losses (Lambrecht et al., 2012; Schmidt, Janse, &

Scharenborg, 2014) or have not related individual hearing loss to the acoustic properties of the stimuli (Orbelo et al., 2005).

In this study, we investigate the question whether age differences in the perception of affect-related acoustic cues can explain the observed age differences in the perception of affect in verbal stimuli. The first research question of this study is therefore whether younger and older listeners differ in their use of acoustic cues for rating affect, and whether such a difference in cue use can explain the observed age difference in affect perception. This question is investigated by comparing the associations between acoustic parameters and affective ratings of a younger and an older listener group. We focus on three acoustic parameters: mean F0 (pitch cue), mean intensity (intensity cue), and articulation rate (tempo cue). These parameters are selected as they have been found to be important conveyers of affect and because sensitivity to pitch, intensity, and temporal changes has been found to be age-dependent. Moreover, in addition to the speech fragments' absolute intensity, which may have been related to how far the speaker happened to be away from the microphone, we include a spectral measure related to vocal effort (Sluijter & Heuven, 1996), i.e., the Hammarberg Index (Hammarberg, Fritzell, Gauffin, Sundberg, & Wedin, 1980).

The second research question addressed in this study revisits the question whether hearing sensitivity plays a role in affect perception. As noted above, several earlier studies, where affect perception is mostly operationalized as emotion categorization performance, have suggested that age differences in affect perception should not be attributed to age-related decline in auditory abilities (e.g., Dupuis and Pichora-Fuller, 2015). Our study addresses this question by investigating the link between hearing sensitivity and differences in the perception of acoustic cues using a group of older listeners with a wide range of hearing sensitivity. We restrict this study to older adults who are not using hearing aids yet, even though some qualify for them.

In order to relate acoustic cues signaling affect to an individual's perception of affect, researchers have frequently used two approaches: the categorical and the dimensional approach. In the categorical approach, concrete terms such as happy, sad, neutral, bored, or angry are used to describe different affectively colored utterances. Note, however, that the underlying affect concepts and interpretation of emotion terms may vary between individuals (Scherer, 2003, 2005). This is because category labels are numerous and require linguistic interpretation, i.e., higher level processing. The dimensional approach offers a more flexible and continuous description of affect (Wundt, 1905). Here, emotions are described by a two- or three-dimensional space, where the most frequently used axes are arousal (calm-aroused) and valence (negative-positive). Moreover, ratings do not depend on consistent interpretations of linguistic labels (such as bored or angry) because emotion dimensions are few, comprehensible, and easy to communicate to a participant in a linguistic (e.g., Likert scale, Likert, 1932) or non-linguistic manner (e.g., a pictorial self-assessment manikin, Bradley & Lang, 1994). In addition, the acoustic parameters pitch, intensity, and tempo have been shown to correlate with specific emotion dimensions (see, e.g., Sauter, Eisner, Calder, & Scott, 2010; Lima, Castro, & Scott, 2013). Sauter et al. (2010) investigated the acoustics of non-verbal vocalizations and found, for instance, that arousal ratings were predicted by durational, pitch, and spectral measures. Therefore, the dimensional approach is employed in this study.

Unlike many other studies on age differences in affect perception, we investigate affect perception using natural stimuli rather than acted stimuli (e.g., Lambrecht et al., 2012; Orbelo et al., 2005; Paulmann et al., 2008). Ecological validity is highest in natural affective speech stimuli (Scherer, 2003) but the use of natural stimuli has certain difficulties and has consequently been little used to investigate affect perception. Stimuli taken from a natural speech corpus generally vary not just in affect dimensions, but also in semantic meaning, utterance length, and inter-speaker variations. Consequently, stimulus numbers might be relatively small after controlling for these factors. Additionally, natural speech corpora often have poor recording quality. A means to reduce this variability is to use controlled stimuli such as manipulated or acted speech, which has its own drawbacks. The encoding of verbal affect, particularly in acted speech, may be more extreme and prototypical (Scherer, 1986; Wilting, Krahmer, & Swerts, 2006) compared to natural speech in which affect perception is cued more subtly. Consequently, responses to natural and acted speech may differ, with more extreme affect realizations in the latter leading to more extreme responses (Wilting et al., 2006). There is evidence that listeners are indeed sensitive to the authenticity of affect. McGettigan et al. (2015) showed that listeners show different neural responses to authentic amusement laughter compared to more controlled voluntary laughter. Moreover, prototypical acoustic patterns with exaggerated frequency contours may be relatively easy to perceive for people with hearing loss (Grant, 1987). Consequently, hearing loss might be less predictive of changes in the use of acoustic cues if emotions are cued prototypically.

In short, this study investigates the perception of affective utterances in younger and older adults using a dimensional approach and natural (i.e., non-acted) speech stimuli by linking acoustic parameters and individual hearing loss directly to participants' affective ratings. By doing so, we aim to investigate the origin of age differences in the perception of affect.

Material and methods

Participants

Two groups of participants were recruited to participate in the experiment: one younger group of students, and one older group. The younger group consisted of 20 native Germans who were students at Radboud University, Nijmegen (18 women, 2 men; age: M = 22.1 years, SD = 1.6, range: 19 – 24 years). The older group consisted of 20 native

Germans who were recruited from the greater area of Saarbruecken via local senior clubs. All participants were paid for their participation. Older participants completed the Montreal Cognitive Assessment Test (MoCA), a brief cognitive screening test in order to check for mild cognitive impairment (Nasreddine et al., 2005). Two participants were excluded because they had a MoCA score of 20 or lower (out of 30; Waldron-Perrine & Axelrod, 2012). None of the participants used hearing aids in daily life. All participants underwent a hearing sensitivity test. Pure-tone thresholds for octave frequencies were measured for both ears with an Oscilla® USB-300 PC-based screener audiometer (air conduction thresholds only). Individual pure-tone averages (PTA) over participants' thresholds for 0.5, 1, 2, and 4 kHz[1] in the better ear were used as an index of hearing sensitivity in the statistical analyses. Higher pure-tone average indicated poorer hearing sensitivity. One older adult reported to have tinnitus and was therefore excluded from the analyses. The final group of older adults then consisted of 17 individuals (14 women, 3 men; age: *M* = 72.6 years, *SD* = 5.4, range: 61 – 82 years). There is evidence that gender differences exist in affect perception (e.g., Lambrecht, Kreifelts, & Wildgruber, 2014). It is therefore important that both age groups have a similar distribution of female and male participants, with both age groups being skewed toward female participants. Mean PTA was 2.2 dB HL (*SD* = 3.8, range: -5.0 – 10.0 dB HL) for the younger listener group and 25.1 dB HL for the older listener group (*SD* = 12.0, range = 3.8 – 46.3 dB HL). Hearing

[1] To capture hearing losses across a broad range of frequencies we used an extended PTA measure here based on octave frequencies from 0.5 – 4 kHz. Individual hearing thresholds for the better ear (as judged by $PTA_{(0.5,1,2,4\ kHz)}$) are presented in Appendix Table C.

To compare the results on hearing and affect obtained in Chapter 4 with results obtained in Chapter 3, we re-ran the older-adults-only model as described in the results section with a PTA based on frequencies 1, 2, and 4 kHz. Results for arousal and valence were unchanged: As before, this (alternative) PTA measure did not surface in any simple effect or interaction in the arousal model, while replicating the effects observed for hearing in the valence model.

sensitivity differed significantly between the younger and older adults ($t = -7.34$, $p < .001$). Neither in the younger group ($r = 0.08$, $p = 0.74$) nor in the older group ($r = 0.12$, $p = 0.65$) did hearing sensitivity correlate with age.

Experimental design

The VAM corpus

The stimuli were taken from the audio-only section of the audio-visual "Vera am Mittag" (a German TV talk show; henceforth: VAM) corpus for authentic, affectively colored conversational speech (Grimm, Kroschel, & Narayanan, 2008). The VAM corpus consists of 1018 affective utterances divided into two subsets: VAM-Audio I and VAM-Audio II. VAM-Audio I consists of 499 utterances produced by 19 different speakers (4 male and 15 female). VAM-Audio II consists of 519 utterances by 28 speakers (7 male and 21 female). The corpus comes with mean reference values for the degree of arousal and valence for each utterance. These reference values were collected with the same pictorial 5-point scales, ranging from -1 (calm/negative) to +1 (aroused/positive), as employed in the present study. Each utterance was evaluated by a group of younger adults (VAM-Audio I: 17 evaluators, VAM-Audio II: 6 evaluators). Their mean ratings are treated as reference values in our analyses. According to the reference values, the VAM corpus provides a good coverage of the emotional space (arousal: min = -0.83, max = 1.00; valence: min = -0.80, max = 0.77). However, due to the discussion topics within this TV program (relationship crises, jealousy, fatherhood questions, etc.), the emphasis within the corpus was found to be on neutral to more negative emotions (Grimm et al., 2008).

Subsets for the arousal and valence rating tasks

Stimuli for the affect rating experiments were selected from both VAM-Audio I and II. In order to not overload, confuse, or bias participants, the two age groups were presented with two separate one-dimensional emotion rating tasks, i.e., participants rated only one emotion dimension per stimulus at the time. Separate stimulus sets for arousal and for valence were created, whereby both sets complied with the following three criteria. First, we only selected stimuli that did not exceed the perceptual window for information integration, i.e., utterances did not exceed three seconds including hesitations and pauses (cf., Pöppel, 2004). Moreover, longer utterances might be less consistent in their degree of arousal or valence thus making it harder for the participants to attribute the utterance to one of the five steps on the scale. Second, since semantic meaning may change the emotional content of an utterance, e.g., when non-verbal information is negative while the verbal information is positive as in sarcasm (Cheang & Pell, 2008), only semantically neutral utterances were selected to minimize semantic interference (e.g. '*Was hast du getan?*' 'What have you done?', '*Erzählst denn du*' '(What) do you say,' '*Hab ich mir doch gedacht*' 'That's what I have thought', '*Er ist relativ jung*' 'He is relatively young'). To that end, transcriptions of the utterances were presented to three independent evaluators who were asked whether they thought a particular utterance has a positive or negative connotation or whether it was semantically neutral. Only the utterances labeled as neutral by at least two of the three raters were included in the final stimulus sets. Third, only stimuli which had arousal or valence reference ratings closest to the values of the five steps on the scale (1, 0.5, 0, -0.5, -1) were included in the final test sets. In order to familiarize participants with the task, another four utterances per rating task were selected to serve as practice trials.

The final item set for the arousal rating task consisted of 24 utterances from 17 different speakers in total (3 male and 14 female speakers; minimum reference rating = -0.66, maximum reference rating = 0.94). The item set for the valence rating task included 18 utterances from 15 different speakers (4 male and 11 female speakers; minimum reference rating = -0.80, maximum reference rating = 0.77). Please note that as fragments were selected to represent a range of either arousal or valence reference ratings and due to the stimulus selection constraints outlined above, the stimulus sets differed for the two affect dimensions. There was, however, an overlap of two utterances between the item sets; thus, two stimuli were rated for both arousal and valence. There was a large overrepresentation of utterances with negative valence in the corpus due to the corpus's nature, making it hard to control for valence in the arousal sentences. In fact, no stimuli for the arousal task had positive valence (valence values ranged from −0.8 to 0.1, SD = 0.25). The arousal values for the valence sentences were more balanced (range: −0.8 to 0.9, SD = 0.48).

Acoustic measurements

Acoustic analyses were carried out for the stimuli. Acoustic measurements were related to the VAM reference values for the two emotion dimensions. Mean F0 and mean intensity were calculated (averaged over the phrase) using Praat (Boersma & Weenink, 2013). As a measure of tempo, articulation rate was calculated by dividing the number of syllables in the canonical transcription of the utterance by its file length excluding pauses longer than 100 ms. Spectral slope related to vocal effort is reflected in the spectral information described by the Hammarberg Index (Hammarberg et al., 1980). The Hammarberg Index is defined as the intensity difference between the maximum intensity in a lower frequency band [0 – 2000 Hz] versus a higher frequency band [2000 – 5000 Hz]. In this study, the Hammarberg Index was used as an energy distribution measure averaged across the entire utterance. Table 1 shows the Pearson correlation coefficients for the correlations between the acoustic parameters and for the correlations of the acoustic parameters with the reference ratings for arousal and valence.

As expected, for both arousal and valence, a positive correlation between mean F0 and mean intensity was found (arousal: r = .79; valence: r =.67). Articulation rate correlated with mean intensity for the arousal stimuli (r = -.42) but did not correlate with any of the other acoustic parameters, or with the reference affect ratings. For the Hammarberg index of vocal effort, we found significant, positive correlations with mean

F0 (r = .47) and mean intensity (r = .71) for the valence stimuli but not for the arousal stimuli. Banse & Scherer (1996) intercorrelated many different acoustic parameters from affective speech, including mean F0, the Hammarberg Index, and a measure related to mean intensity. Our findings are in agreement with theirs in terms of the direction of the correlations, though effect sizes differ slightly: they found correlations for mean F0 with mean intensity (r = 0.62), Hammarberg Index with mean F0 (r = 0.34), and Hammarberg Index with mean intensity (r = .60).

Table 1. Correlation coefficients per emotion dimension.

		Mean F0	Mean Intensity	Articulation Rate	VAM Reference Values
Arousal	Mean F0	–			.47 **
	Mean Intensity	.79 ***	–		.75 ***
	Articulation Rate	-.38	-.42 *	–	-.20
	Hammarberg Index	.25	.39	- .13	.39 **
Valence	Mean F0	–			-.35
	Mean Intensity	.67 **	–		.06
	Articulation Rate	-.16	-.21	–	.20
	Hammarberg Index	.47 *	.71 **	-.22	.05

Notes. Pearson correlation coefficients between the acoustic parameters and with the reference ratings for arousal and valence for the two sets of stimuli. For VAM reference values, Kendall’s Tau was used instead of Pearson’s correlation coefficients, as reference ratings for arousal were not normally distributed in our data set. $*p < .05$, $**p < .01$, $***p < .001$.

In general, correlations between acoustic parameters and the VAM reference values were stronger for arousal than for valence (cf. also Sauter et al., 2010; Lima et al., 2013). For arousal, positive correlations were found for mean F0, mean intensity, and the Hammarberg Index with the reference ratings, as has been found in other studies (e.g., Pereira, 2000;

Schröder et al., 2001; Schröder, 2006). In contrast, there were no significant correlations between the reference ratings for valence and any of the acoustic parameters. Other studies have found correlations between valence and acoustic parameters, but weaker than those for arousal (Pereira, 2000; Schröder et al., 2001; Schröder, 2006). Pereira (2000), for instance, only found a significant correlation between mean F0 and valence for male speakers (note that the majority of the speakers in our stimulus set were female).

Procedure

Prior to testing, all participants gave written informed consent to use their data for research purposes. Younger adults were tested individually in a sound-attenuated booth at the Centre for Language Studies Lab at the Radboud University in Nijmegen. Older participants were either tested in a quiet environment at their homes or in a quiet room at a senior club house. They were comfortably seated in front of a laptop. First, participants carried out the self-paced emotion ratings tasks (15-30 min for the two rating tasks). The order in which the two emotion rating tasks (Arousal, Valence) were presented was counterbalanced across participants. Subsequently, the older participants completed the MoCA test (15 min) and hearing sensitivity of both age groups was measured (15 min). The total experiment duration was about one hour.

Prior to each rating task, the emotion dimension at hand and the pictorial rating tool were explained to the participant (Figure 1). A printed version of the rating tool was provided which depicted the five steps for each emotion dimension. On the printed version, numbers from 1 to 5 were assigned to each step (arousal: very calm = 1, very aroused = 5; valence: very negative = 1, very positive = 5), replacing the values ranging from -1 to +1. The meaning of each step on the scale was described to the participant by the experimenter and the participant's attention was particularly drawn to the changing attributes of the figure, i.e., calm versus expressive (arousal) and smiling versus frowning (valence). Each stimulus was presented twice to increase statistical power. There was no break in between the two renditions of the stimulus set. The order of the utterances was randomized for each rendition. Participants were informed that each utterance occurred twice in the experiment. Furthermore, in addition to verbal instructions, written instructions were provided on the computer screen. Throughout the instructions, participants could ask questions. Each rating task was preceded by a practice session, which was identical to the set-up of the rating task. During the practice session, four items were presented in a randomized order in two renditions. Practice items were different for the arousal and the valence task.

Each trial started with the presentation of a fixation cross (250 ms) followed by a white screen (100 ms) in order to alert participants that a stimulus was coming up. Then the utterance was presented auditorily to both ears via circumaural headphones (Sennheiser HD 215). The mean presentation level was kept constant at 70 dB SPL for both participant groups. Participants entered the number corresponding to the intended step on the scale via the keyboard and proceeded to the next trial by pressing the return key. Participants were asked to rate the utterances as intuitively as possible. They could listen to an utterance multiple times by pressing the space bar on the keyboard. This option was provided in order to allow participants to fully process the auditory input, since the utterances were relatively short. However, they were encouraged to make their ratings as spontaneously as possible, i.e., to use the repeat function only if they thought they had missed crucial information to be able to rate the utterance. Collapsed over renditions, younger adults listened to the arousal stimuli on average 1.14 times (range: 1 – 6) and older adults 1.18 times (range: 1 – 3). For the arousal rating task, 87.1% of the utterances (younger adults: 90.0%, older adults: 83.7%) were rated on their first presentation. For the valence ratings task, collapsed over renditions, younger adults listened to the stimuli on average 1.13 times (range: 1 – 7) and older adults 1.27 times (range: 1 – 3). Of the utterances, 83.6%

(younger adults: 89.9%, older adults: 76.1%) were rated on their first presentation.

Results

Analysis

In order to investigate whether younger and older listeners use acoustic cues differently when rating affect in speech, the age groups' ratings of affect (the dependent variable) were compared using linear mixed-effects regression analyses (with random intercepts for stimulus and participant). Note that parametric tests like regression, including linear mixed-effects models, are robust against violations of the assumption of normal distribution. Moreover, linear mixed-effect models have been shown to be good models to analyze Likert scale data (cf., Norman, 2010). Nevertheless, we also analyzed whether results obtained with our linear mixed-effect regression models were replicated in analyses for ordinal data, although there are suggestions that the risk of finding a false positive (Type 1 error) are higher for the ordinal data analysis method compared to the linear mixed-effects method (cf., Kizach, 2014). The initial model allowed for two-way interactions between each of the acoustic parameters and age group and between each of the acoustic parameters and rendition (i.e., whether they rated the stimulus for the first or the second time); the latter serving as a control variable. Moreover, an interaction effect of age group and rendition was tested. The model with the lowest Akaike Information Criterion (AIC) was considered the best-fitting model. Interactions and predictors that did not improve model fit were removed using a stepwise exclusion procedure (interactions before simple effects, and those with the highest non-significant *p*-values first).

The second research question, concerning the impact of hearing sensitivity on the affect ratings, was investigated using the data from the group of older adults only, where differences in individual hearing sensitivity were more pronounced (see Section Participants).

Table 2. Mean arousal and valence ratings, with standard deviations, for the younger and older adults separately.

	Younger adults		Older adults	
	Mean	**SD**	**Mean**	**SD**
Arousal	-0.044	0.65	0.062	0.60
Valence	-0.013	0.60	0.037	0.60

Therefore, in a second analysis, hearing sensitivity was associated with the affect ratings by the older adults group. The set-up and model selection procedure of this analysis was similar to the first analysis except for the continuous hearing sensitivity measure (PTA) replacing the binomial age group factor in the previous analysis. Analyses were carried out for the arousal and the valence tasks separately. Table 2 lists the mean arousal and valence ratings, with standard deviations, for the younger and older adults separately.

Analysis of arousal rating

Figure 1 shows the relationship between mean intensity and the arousal ratings; more particularly the mean arousal ratings per stimulus for the younger (round symbols) and older (triangles) listener groups plotted against the mean intensity (on the x-axis).

Tables 3 and 4 show the best-fitting models for the two arousal analyses. Both younger and older adults associated higher mean intensity with a higher level of arousal (see the significant simple effect for mean intensity in Table 3), which is also shown by the upward sloping fit lines in Figure 1 (solid lines for the younger and dashed lines for the older participants). There was no general effect of age group in the arousal rating task (cf. Figure 1).

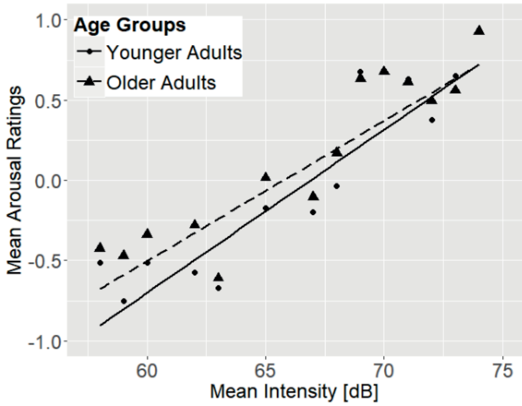

Figure 1. Younger (round symbols) and older participants' (triangular symbols) mean arousal ratings for each individual stimulus as a function of mean intensity of the speech fragments, and the fit lines for the younger (solid line) and older (dashed line) participants.

Table 3. Fixed effect estimates of the best-fitting models of performance for the group comparison of the arousal data; bold indicates significant results, number of observations D 1776, AIC D 1496.

	β	***SE***	***p***
Age Group	0.065	0.061	.29
Rendition	0.024	0.022	.28
Mean Intensity	0.102	0.011	< .001
Age Group × Rendition	0.083	0.033	.012
Age Group × Mean Intensity	- 0.014	0.004	< .001

If we were to remove the acoustic variables from our arousal rating analysis to only test for a simple age group difference across renditions, the Age Group effect also fails to reach significance (β = 0.106, SE = 0.058, *p* = 0.077). However, the older adults showed a less steep intensity increment (as shown by the interaction between Age Group and mean intensity), i.e. the older adults showed a smaller effect of mean intensity on their ratings than the younger adults. Moreover, we found a significant interaction between Age Group and Rendition, i.e., older participants rated the second rendition of the stimulus as more aroused. Importantly, acoustic measures for mean F0, articulation rate, and the spectral measure were not predictive of participants' ratings in the arousal task.

Table 4. Fixed effect estimates for the best-fitting models of performance for the analysis of the arousal data for the older adults only; bold indicates significant results, number of observations D 816, AIC D 804.6

	β	***SE***	***p***
Rendition	0.107	0.026	< .001
Mean Intensity	0.087	0.011	< .001

The analysis of the older adults' data (Table 4) to investigate the role of hearing sensitivity showed a similar picture to the age group comparison: Only significant effects for mean intensity and rendition were found. Thus, again, stimuli were rated as more aroused when rated for the second time and higher mean intensity was perceived as more aroused among the older group. Importantly, however, there was no simple effect of hearing sensitivity, nor was there an interaction between hearing sensitivity and interpretation of the acoustic measures[1].

Analysis of valence ratings

Figure 2 shows the relationship between mean F0 and the valence ratings in terms of the mean valence ratings per stimulus for the younger (round symbols) and older (triangles) listener groups plotted against the mean F0

[1] Analyses of the group comparison and the data of the older adults with a statistical method specifically for ordinal data (cumulative link mixed models, CLMMs, cf. Agresti,2002) showed similar results to those obtained with the linear mixed-effect models.

(on the x-axis). Fit lines for the younger (solid line) and older (dashed line) participants are also shown. Table 5 and 6 shows the best-fitting models for the two valence analyses.

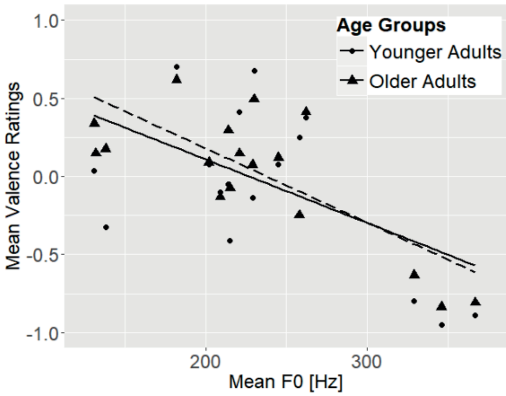

Figure 2. Younger (round symbols) and older participants' (triangular symbols) mean valence ratings for each individual stimulus as a function of mean F0 of the speech fragments, and the fit lines for the younger (solid line) and older (dashed line) participants.

Table 5. Fixed effect estimates of the best-fitting models of performance for the group comparison of the valence data; bold indicates significant results, number of observations D 1332, AIC D 1415.

	β	*SE*	*p*
Mean F0	-0.004	0.001	.008
Age Group	0.050	0.042	.24
Age Group × Mean F0	-0.001	3.271 10^{-4}	.038

The age group comparison for valence (Table 5) showed a simple effect for mean F0. Higher mean F0 of the stimuli was associated with more negative utterances in younger adults. The significant interaction between Age Group and mean F0 indicates that both age groups rated higher mean F0 as more negative, but the change in valence rating associated with each unit increase in F0 was larger for the older than for the younger adults (as also shown by the steeper slope of the fit line for the older adults in Figure 2). Rendition, mean intensity, articulation rate, and the spectral measure of vocal effort were not predictive of valence ratings, nor did they interact with age group.

Table 6. Fixed effect estimates for the best-fitting models of performance for the analysis of the valence data for the older adults only; bold indicates significant results, number of observations D 612, AIC D 767.8.

	β	***SE***	***p***
Mean F0	-0.005	0.001	< .001
Hearing Sensitivity	0.006	0.002	.018
Hearing Sensitivity × Mean F0	$-5.795\ \ 10^{-5}$	$2.117\ \ 10^{-5}$	.006

The analysis of the older listeners' valence data (Table 6) to investigate the role of hearing sensitivity showed a simple effect for mean F0. As was found in the age group comparison, higher mean F0 lead to more negative ratings. Importantly, there was a significant simple effect of hearing sensitivity: poorer hearing (i.e., higher PTA values) was associated with more positive valence ratings. Finally, there was an interaction between hearing sensitivity and mean F0: the change in valence rating associated with each unit increase in F0 was larger with increasing hearing loss[2] [3].

[2] Here, a typo from the original publication (Schmidt, Janse, Scharenborg, 2016) has been fixed by exchanging the term "hearing sensitivity" with "hearing loss".

[3] Analyses of the group comparison and the data of the older adults with CLMMs showed similar results to those obtained with the linear mixed-effect models, except for

General discussion

Previous research revealed age differences in the perception of verbal affect (Orbelo et al., 2005; Paulmann et al., 2008). The current study investigated the origin of this age difference. The first aim was to investigate whether younger and older listeners differ in the way they make use of affect-related acoustic cues in natural speech; more specifically, mean F0, mean intensity, articulation rate, and vocal effort. The second aim was to determine the impact of age-related hearing sensitivity differences on the use of these affective cues. Three methodological aspects were combined to investigate the perception of affect: the perception of acoustic parameters was linked to individual hearing sensitivity, conversational (rather than acted) speech was used in the rating tasks, and participants rated two emotional dimensions, arousal and valence, rather than classified affect categories.

The results showed that only two acoustic cues predicted our participants' ratings. Both the younger and the older age groups associated higher mean intensity with an increased level of arousal and associated higher mean F0 with more negative valence. As pitch is considered the most telling component of affective prosody (Hammerschmidt & Jürgens, 2007; Mozziconacci, 1998; Rodero, 2011), the finding that mean F0 was a good predictor of valence was not surprising. Others have found, however, that higher levels of arousal are also particularly related to higher pitch (Schröder, 2006). This was not borne out by our data and may be accounted for by the interplay between several acoustic cues: Mean F0 and mean intensity were highly correlated in the selected subset of arousal stimuli. For the present item sample, mean intensity was probably the most prominent acoustic cue and was therefore a better predictor for arousal ratings than mean F0. It is unclear why intensity differences were a stronger cue to arousal than vocal effort,

---

the interaction between Age Group and mean F0, which became marginal ($\beta=-0.003$, $SE=0.002$, $p=0.06$).

as the former, but not the latter, could simply relate to the speaker's distance from the microphone. Possibly, listeners rely on these salient and prototypical intensity differences because vocal effort may be more difficult to compare across multiple speakers in a situation in which a listener has to evaluate affect across multiple speakers (as is the case in our design).

The arousal data showed an effect of rendition in that the older (but not the younger) adults rated the same utterances as more aroused in the second rendition. Note that this effect was absent in the valence ratings. This finding suggests that rating behavior can change over the course of a rating task, which should encourage researchers to investigate block or rendition effects in their experimental designs. Possibly, listening to affective utterances raises the general level of arousal within the listener, thus slowly increasing the reference or resting level for arousal over time.

While age-related differences in the perception of emotion categories in speech are well documented (e.g., Mitchell et al., 2011; Orbelo et al., 2005; Paulmann et al., 2008), no such difference was observed with the dimensional approach in the current study. In other words, in our study, younger and older adults did not generally differ in the way they rated the emotion dimensions arousal or valence. This absence of a general age difference in our study could be due to lack of statistical power (given our relatively small sample size), or to our use of the dimensional approach instead of classification of emotion categories, even though age differences have been reported using that approach as well (e.g., Lima et al., 2014). Nevertheless, we found age differences in the use or interpretation of both mean intensity and mean F0, which was the focus of our study. For arousal, older adults' ratings were less affected by changes in mean intensity compared to younger adults. We will come back to this point below. For valence, differences in mean F0 affected the ratings of older adults more than those of younger adults. Hence, the effect of a mean F0 change on valence rating was more pronounced for the older than for the younger adults. Our finding may relate to a recent study by Lima and colleagues (2014). Lima and colleagues (2014) investigated affect perception in younger and older adults using 'affective bursts', which are vocalizations without verbal content, such as laughter, sobs, and sighs. Their study showed that pitch was used differently across age groups depending on whether age groups evaluated positive or negative affect. In Lima and colleagues (2014), mean F0 was a stronger predictor for rating fear and sadness (negative valence) in older adults than in younger adults. Conversely, F0 differences were associated with pleasure (positive valence) in younger but not in older adults. As our set of valence stimuli was skewed towards negative affect, our finding that older adults were more sensitive to F0 differences in the interpretation of valence agrees with their findings for fear and sadness. Thus, both younger and older adults used mean F0 as a cue to valence, but the age groups used mean F0 to a different extent. Nevertheless, as observed before, this differential use of mean F0 as a cue to valence did not lead to age differences in overall valence ratings. As argued earlier and below, the correlations between valence ratings and the acoustic cues were low. Possibly, older adults and younger adults also differed in their use of other, here not investigated, acoustic cues that cue valence, which counteract the differences in the use of mean F0.

Previous research has shown that hearing loss impacts intensity discrimination (Boettcher et al., 2001). Considering that mean intensity was identified as the main cue for rating arousal in the current study, deterioration in the perception of intensity due to age-related hearing loss may account for the observed age differences in the arousal rating task. However, this was not confirmed by the analysis of the older adults' data: Among older adults, individual hearing loss was not related to ratings of arousal. This confirms recent findings by Dupuis and Pichora-Fuller (2015) who also found that emotion categorization accuracy by younger and older adults was not correlated with their auditory abilities (i.e., neither with their hearing sensitivity, nor with measures of auditory processing, such as F0 or intensity difference limens). There are several possible interpretations of this finding. Older adults may be less willing than younger adults to use the entire rating scale while performing a rating task. As also argued by Lima et al. (2014), this ‘conservatism’ account is somewhat unlikely, however, considering that the older adults used a wider range of the scale than younger adults for the valence task. A second explanation could be that arousal perception is relatively robust against mild-to-moderate hearing loss because arousal is cued by several other acoustic parameters. In both the current study and previous work (Pereira, 2000; Schröder, 2006), arousal has been reported to show strong correlations with multiple acoustic parameters, including intensity and pitch measures. Hence, possibly, the perception of arousal in affective speech is more robust against mild sensory degradations due to the availability of clear acoustic cues which reliably signal arousal in the speech signal. Hearing loss might have played a role if our older adult sample had been (even) more diverse in hearing sensitivity. Third, the interaction between the use of mean intensity and age group may still have an auditory/ perceptual origin: possibly, age-related hearing decrements in auditory processing that are not apparent from the tone audiogram may relate to older adults’ smaller sized intensity effect. Finally, age differences in affect perception may be dissociated from hearing loss if they primarily arise at processing levels following auditory analysis. This account would be in keeping with the meta-analysis by Ruffman, Henry, Livingstone, and Phillips (2008) that age differences in emotion recognition arise due to age-related changes in the “social brain,” i.e., due to changes in volume of frontal and temporal brain areas, as well as changes in neurotransmitters. In other words, even if older adults are able to hear cues that signal affect, higher-order processing of these cues may result in less differentiation of emotional content than in younger adults.

Individual hearing sensitivity did, however, impact the interpretation of valence in our experiment, i.e., poorer hearing generally led to more positive valence ratings. This observation makes it less likely that the lack of a hearing sensitivity effect on arousal (discussed above) should be attributed to lack of statistical power due to a relatively small sample of older adults. In contrast to our findings, Orbelo and colleagues (2005) did not find such a global effect of hearing loss on the comprehension of affective prosody, even though the mean and standard deviation of their subjects' hearing sensitivity was comparable to the hearing loss observed in the present study. The difference in findings may, however, originate from a difference in the used materials. Orbelo and colleagues used acted affective speech material, whereas we used natural affective speech. As argued in the introduction, the more prototypical acoustic expression in acted compared to natural speech may lead to a more extreme realization of affective prosody (Scherer, 1986; Wilting et al., 2006), which may be relatively easy to perceive, even for people with hearing loss (Grant, 1987). With natural, and hence less extreme speech materials, as used in our study, those with poorer hearing may be less certain about their valence perception.

There was no impact of hearing sensitivity on mean intensity, articulation rate, or the spectral measure of vocal effort on the participant's valence ratings. This was not unexpected as none of these parameters predicted valence in general. Importantly, apart from the general hearing loss effect of valence rating, hearing loss also modulated listeners' use of the pitch cue for valence. Based on findings of poorer pitch discrimination in older compared to younger adults (He et al., 2007; Souza et al., 2011), one would expect pitch to affect ratings of older adults less than those of younger adults. We can only speculate on why our findings show the opposite result. Note again that mean F0 showed up as a significant predictor of valence ratings in our study, but Table 1 showed no correlation between mean F0 and the reference ratings that came with the conversational speech corpus. This shows that the relationship between the valence ratings and the acoustic measures we focused on here was not as strong and straightforward as for arousal. Participants in the current study relied on mean F0 when rating valence. However, the acoustic profile of affective speech is complex and is not only encoded in pitch, intensity and tempo of the utterance. Some variations in affective speech may be captured by alternative, perhaps more subtle, cues (Bänziger, Patel, & Scherer, 2014) that we did not include here. Voice quality, for example, is known to be used in verbal affect perception (Grichkovtsova, Morel, & Lacheret, 2012), and is related to the perception of valence in affective speech (Waaramaa & Leisiö, 2013). Possibly, these alternative cues for valence may have been less available to the older listeners with poorer hearing in our experiment, leading to a differential use of mean F0 in the current sample of older adults. Note also that this may then tie in with the account provided above on the similarity between our valence results and those by Lima et al. (2014). A different weighing of acoustic cues across age groups may result from age-related hearing loss, but not necessarily. Lima et al. (2014) found that age groups were equally efficient in using acoustic cues but that there were differences in the patterns of emotion-specific predictors. Lima et al. (2014) therefore argue, in line with Ruffman et al. (2008), that age-related differences in weighting of acoustic cues may reflect changes in higher-order processing. Clearly, follow-up research with more controlled or experimentally manipulated materials would be required to test this cue trading in more detail, and to see to what extent changes in cue use are driven by age-related changes in perception or in higher-order processing.

This study showed that both younger and older listeners base their affect ratings on acoustic cues in speech: mean intensity for arousal and mean F0 for valence. However, the extent to which these acoustic parameters are used for affect rating varies across age groups: intensity differences are used less by older adults for arousal ratings, while differences in mean F0 influence valence ratings by older adults more than they do those of younger adults. Arousal perception seems to be robust against mild-to-moderate hearing loss which may be explained by the availability of multiple clear acoustic parameters consistently signaling arousal. In conclusion, this study suggests that age differences in the perception of affect relate to differences in acoustic cue use. Moreover, differences in cue use for the two emotion dimensions suggests that future studies should treat the perception of arousal and valence separately.

Conclusion

This study suggests that age differences in the perception of affect relate to differences in acoustic cue use, and that age differences in cue use can only partly be explained by age-related changes in hearing sensitivity. Moreover, differences in cue use for the two emotion dimensions suggest that future studies should treat the perception of arousal and valence separately.

Chapter 5

Perceiving Arousal in speech: Effects of hearing loss, loudness processing ability, and hearing aid use

This chapter is a reformatted version of:
Juliane Schmidt, Odette Scharenborg, Diana Herzog & Esther Janse (in preparation). Perceiving Arousal in speech: Effects of hearing loss, loudness processing ability, and hearing aid use.

Part of the data reported in this chapter has been published in:
Juliane Schmidt, Diana Herzog, Odette Scharenborg & Esther Janse (2015). Do hearing aids improve affect perception? In van Dijk, P., Başkent, D., Gaudrain, E., de Kleine, E., Wagner, A., and Lanting, C. (Eds.), *Physiology, Psychoacoustics and Cognition in Normal and Impaired Hearing*. Advances in Experimental Medicine and Biology, 894 (pp. 47–55). Springer International Publishing, Cham.

Abstract
Listeners derive affective meaning from prosodic cues in the speech signal. Affect can be conceptualized as having two emotion dimensions: valence (negative vs. positive) and arousal (calm vs. expressive). The latter dimension is the focus of the current study. The perception of prosodic cues signaling arousal in speech (e.g., pitch, vocal effort, intensity) may be influenced by hearing loss and general loudness processing abilities such that arousal perception is impaired. This study investigated the influence of hearing aid use, individual hearing loss, and individual loudness

processing ability on the perception of arousal and the use of arousal-related prosodic cues in the speech signal. Arousal ratings by a group of older hearing aid users were compared for aided and unaided listening and were also compared to ratings by a group of older adults with age-normal hearing.

Intensity differences were the primary cue to arousal perception across participant groups and listening conditions. Wearing hearing aids counteracted the general effect of hearing loss on arousal perception: Compared to participants with age-normal hearing, hearing aid users in the aided listening condition generally showed the same pattern of affect ratings and were significantly better at using mean intensity as a cue to arousal. Individual loudness processing abilities did not explain any additional variance beyond hearing loss effects.

As impaired arousal perception may have consequences for the perception of fine differences in affect (e.g., hot versus cold anger), the current results underline the importance of hearing aids in the rehabilitation of affect perception.

Introduction

The ability to adequately understand and respond to emotional signals is crucial for social participation and communication. This ability is particularly important for older adults because social isolation may have a detrimental effect on health and well-being in older adulthood (Bath & Deeg, 2005). Affect is communicated via facial cues and information in the speech signal; specifically, prosodic parameters have been shown to cue a speaker's emotional state (Banse & Scherer, 1996; Coutinho & Dibben, 2013; Scherer, 2003). Although both visual and auditory information is available in many everyday communicative settings, there are situations in which the listener is deprived of visual information, (e.g., when not facing the speaker), and consequently has to solely rely on information coded in the speech signal.

Older adults have been reported to be less sensitive to affective prosody than younger adults, as evidenced by poorer accuracy in emotion recognition tasks (e.g., Kiss & Ennis, 2001; Mitchell, Kingston, & Barbosa Bouças, 2011; Orbelo, Grim, Talbott, & Ross, 2005; Paulmann, Pell, & Kotz, 2008). This difference in accuracy could potentially be influenced by age-related changes that have been observed in the perception of the prosodic parameters (He, Dubno, & Mills, 1998; He, Mills, & Dubno, 2007).

Multiple prosodic parameters correlate with the emotion dimension arousal (Pereira, 2000; Schmidt, Janse, & Scharenborg, 2014; Schröder, Cowie, Douglas-Cowie, Westerdijk, & Gielen, 2001; Schröder, 2006), which represents the extent to which a stimulus is calm or expressive. Arousal is positively correlated with pitch, vocal effort, and mean intensity; mean intensity being the most prominent prosodic cue to arousal (e.g., Lima, Castro, & Scott, 2013; Schmidt, Janse, & Scharenborg, 2016). Correlations between prosodic parameters and the rating of valence (the extent to which a stimulus is positive or negative), a second major emotion dimension, were weaker and inconsistent (Pereira, 2000; Schmidt et al., 2014; Schröder et al., 2001; Schröder, 2006). This study investigates the relation between hearing status and the use of prosodic cues to affect. Therefore, we only focus on the perception of arousal in this paper [1].

Hearing loss, which is prevalent in older adults, might be part of the reason of the observed age-related changes in the perception of prosodic parameters (Boettcher, Poth, Mills, & Dubno, 2001), and consequently might impair affect perception. Although hearing aids clearly improve speech intelligibility (e.g., Humes, 2007; Shanks, Wilson, Larson, & Williams, 2002), it is unclear to what extent hearing aids sufficiently restore information needed for affect perception in speech. Several studies with severely hearing-impaired children and adolescents indicate that hearing aid users perform poorly compared to their normal-hearing peers when rating affective prosody in speech (Most & Michaelis, 2012; Most, Weisel, & Zaychik, 1993). These findings, however, cannot be directly transferred to older hearing-aid wearing adults for several reasons. First, the type and onset of hearing loss in hearing-impaired younger and older adults differ. Older adults were often normal-hearing when they acquired language, and thus, will have learned to interpret the prosodic cues associated with affect. In contrast, hearing-impaired children with congenital hearing loss did not acquire language in the typical way, and hence might not acquire verbal affect perception in a typical way. Second, age-related hearing loss particularly affects the higher frequencies, whereas the audiogram in congenital hearing loss is not typically sloping. Third, younger and older adults differ in the perception of affective prosody, even if both groups have normal hearing (e.g., Kiss & Ennis, 2001).

So far, the influence of hearing loss on affective prosody perception was not confirmed by previous studies which have used a 'categorical approach' (Dupuis & Pichora-Fuller, 2015; Mitchell, 2007; Orbelo et al.,

---

[1] Results on valence perception can be found in Schmidt, Herzog, Scharenborg, & Janse, 2015.

2005), in which participants are asked to label or classify emotional content as emotion categories, such as for instance, happy, sad or angry. However, many emotion categories involve both dimensions arousal and valence (e.g., angry, being a negative and aroused emotion). As such, from the absence of hearing loss effects on emotion categorization, it is not clear whether hearing loss might affect perception of either valence or arousal. When arousal is investigated separately from valence information, hearing loss might be found to impact the perception of arousal (but cf. Schmidt et al., 2016), particularly for those with more severe levels of hearing loss who qualify for hearing aids. Finally, note that all studies (apart from our earlier study Schmidt et al., 2016) on the impact of hearing loss on the perception of affective prosody used acted speech material. The lack of a consistent effect of hearing sensitivity on affect perception could also be due to the more extreme prosodic expression of affect in acted compared to natural speech (Scherer, 1986; Wilting, Krahmer, & Swerts, 2006). More extreme expressions of affect may be relatively easy to perceive, even for people with hearing loss (Grant, 1987). The use of acted speech may thus obscure a possible influence of hearing sensitivity on affect perception in natural communicative settings.

In addition to hearing loss, general auditory processing abilities, such as loudness processing ability, may affect arousal perception. In order to estimate the degree of arousal in speech, listeners need to be able to detect small differences in intensity, i.e., the most relevant prosodic cue for arousal perception. However, both aging (e.g., Harris, Mills, & Dubno, 2007; He et al., 1998) and hearing loss (e.g., Boettcher et al., 2001; Koehnke, Culotta, Hawley, & Colburn, 1995) have been reported to cause elevated thresholds in tasks testing loudness processing abilities. Decreased loudness processing ability was already observed for mild hearing loss with mean thresholds for frequencies 0.25-8 kHz between 0 and 25 dB HL (Boettcher et al., 2001). Age-related decline is also observed for the perception of other prosodic parameters related to arousal, such as pitch (He et al., 2007; Mitchell & Kingston, 2014; Souza, Arehart, Miller,

& Muralimanohar, 2011). Dupuis and Pichora-Fuller (2015) found no link between auditory processing ability and emotion identification accuracy in older adults, but note that their study used acted speech, and that their sample consisted of older adults with normal hearing (Dupuis & Pichora-Fuller, 2015). We investigated whether individual ability to detect differences in intensity may predict arousal perception in a sample of older adults with more severe hearing loss.

The first aim of this study was to investigate the influence of hearing aid use and that of individual hearing loss on listeners' use of prosodic cues (i.e., use of pitch, intensity, and vocal effort variation in spoken fragments) to arousal. More specifically, we asked whether wearing a hearing aid makes listeners more responsive to subtle differences in the acoustic variation naturally present in the speech materials. To that end, older bilateral hearing aid users were tested on an arousal rating task, while wearing their hearing aids and without them. The speech materials consisted of natural conversational speech stimuli in order to mimic realistic listening conditions. By relating participants' ratings to the acoustic information present in the speech materials (i.e., variation in pitch, intensity and vocal effort), we investigated participants' use of the acoustic cues to arousal across hearing aid conditions. Secondly, we investigated whether the use of a hearing aid brings arousal perception to the level of normal-hearing older adults. To that end, the hearing-impaired listeners were compared to a group of age-matched normal-hearing older listeners. The third aim of this study was to investigate the influence of individual loudness processing ability on arousal ratings. More specifically, we asked whether loudness processing ability modifies arousal ratings and/or the use of prosodic cues to arousal (such as intensity and vocal effort) beyond the effects of hearing loss.

Material and methods

Participants

Two groups of older adults (age range 65 – 82) were tested. All participants were Swiss German native speakers and were financially compensated for their participation. One group consisted of 23 older hearing aid users with bilaterally symmetric hearing loss (M_{Age} = 73.5 years, SD_{Age} = 4.5; 17 men, 6 women), who were recruited via the Phonak AG participant database. These listeners have been wearing hearing aids bilaterally for at least two years ($M_{Years\ of\ Hearing\ Aid\ Use}$ = 15; $SD_{Years\ of\ Hearing\ Aid\ Use}$ = 9.9, range: 2 – 40 years). As can be seen in Appendix Table D presenting the individual hearing thresholds of the better ear (as judged by $PTA_{(0.5,1,2,4\ kHz)}$) are presented in Appendix Table A., most but not all hearing aid users had typical patterns of age-related hearing loss as evident from sloping audiograms. Participants with flat audiograms also had long duration of hearing aid use. Hearing aid fitting was checked prior to testing[2]. A second group consisted of 22 adults with age-normal hearing (M_{Age} = 70.8 years, SD_{Age} = 5.2; 10 men, 12 women), who were recruited via the Phonak human resource department and a local senior club in Staefa, Switzerland. Individual hearing thresholds of the better ear of these control participants (as judged by $PTA_{(0.5,1,2,4\ kHz)}$) are presented in Appendix Table E.

Participants' hearing acuity was assessed by means of pure-tone audiometry. The mean unaided pure-tone average (PTA) of the better ear

[1] Hearing aid fit was checked by comparing loudness processing abilities (see Section 'Loudness Scaling Task') in an aided (*M* = 55.8, *SD* = 8.1) versus an unaided condition (*M* = 40.7, *SD* = 11.7) across 0.5, 1, and 2 kHz. Significantly better overall loudness processing ability in the aided condition was taken as an indication of proper hearing aid fitting; $t(39.0) = 5.1$, $p < .001$. Moreover, aided loudness processing ability was compared to the performance of normal-hearing participants (*M* = 44.0, *SD* = 2.2). Overall loudness processing ability in aided hearing-impaired participants was significantly better compared to that in normal-hearing participants; $t(25.7) = 6.7$, $p < .001$.

across 0.5, 1, 2, and 4 kHz[1] for the hearing-impaired group was 49.8 dB HL (*SD* = 8.7, range: 32.5 – 68.8). Moreover, hearing-impaired participants stated that their hearing loss was not trauma-induced or hereditary (self-report). The normal-hearing participants had age-normal thresholds (as defined in the ISO 7029:2000 standards for this age group, International Organization for Standardization, 2000). Thresholds below the ISO's maximum pure-tone average threshold (across 0.5, 1, 2, and 4 kHz) were considered as normal hearing. For participants between 65 and 70 years, the maximum pure-tone average thresholds were applied according to their age and gender. Note that the ISO only provides thresholds for people up to the age of 70 for men (PTA = 33.5 dB HL) and women (PTA = 26.0 dB HL) and that thresholds of all our participants older than 70 years were below these thresholds. Additionally, participants underwent a brief cognitive screening test to rule out mild cognitive impairment. We used the German version of the Montreal Cognitive Assessment Test (Nasreddine et al., 2005) using a cutoff criterion of 67% accuracy (cf., Waldron-Perrine & Axelrod, 2012). The test was adjusted for hearing-impaired participants (Dupuis et al., 2014) by leaving out tasks in which auditorily presented items had to be memorized. All participants passed the test.

Emotion Task

Arousal perception was tested using the dimensional approach. In this approach, participants indicate the level of arousal (calm vs. aroused) on a rating scale. Stimuli were 24 short audio-only utterances from an audio-visual corpus of affective German conversational speech (Grimm, Kroschel, & Narayanan, 2008) produced by multiple speakers. Note that the identification of the meaning of emotional expressions correlates across languages, particularly for similar languages (cf., Scherer, Banse, & Wallbott, 2001). Therefore, the way affect is encoded in Swiss German

[1] To capture hearing losses across a broad range of frequencies we used an extended PTA measure here based on octave frequencies from 0.5 – 4 kHz.

(spoken by our participants) is not expected to differ considerably from that in German as spoken in Germany (and represented in the corpus) given the close relationship between German and Swiss German. The corpus comes with mean reference values (as provided by German raters) for the degree of arousal for each utterance. These reference values had been collected with a 5-step pictorial rating tool (Bradley & Lang, 1994), ranging from -1 (calm) to +1 (aroused). The same rating tool was used to obtain arousal ratings in the current study. Twenty-four utterances were selected from the corpus (with arousal reference value ranging from -0.66 to 0.94). All stimuli in our experiment were neutral regarding the content of what was said (e.g., 'Was hast du getan?' 'What have you done?'; 'Erzählst denn du' '(What) do you say'; 'Hab ich mir doch gedacht' 'That's what I have thought'; 'Er ist relativ jung' 'He is relatively young), to minimize semantic interference and were shorter than three seconds. Two randomized lists were created.

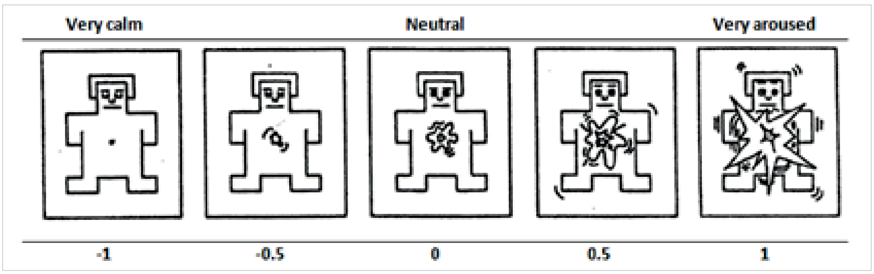

Figure 1. Pictorial 5-point scale for arousal (adapted from Bradley & Lang, 1994) with linguistic labels for both ends of the scale and their corresponding numeric values.

Participants were comfortably seated in a sound-treated room and were tested in the free field. The pictorial rating tool (see Figure 1) was displayed on a computer screen and stimuli were presented via a single loudspeaker (Meyer Sound MM-4XP) which was placed at head level in front of the participant (0˚ azimuth) at a distance of one meter. The mean presentation level of the stimuli was 70 dB SPL. Participants received written and oral instructions and performed four practice trials before proceeding to the test stimuli of the rating task. The rating task was completed at the participant's own pace. Utterances were rated one at a time and could be replayed if needed.

All participants performed the rating task in two listening conditions. For the hearing aid users, these two conditions were with (aided) and without their hearing aids (unaided). The normal-hearing participants completed the task in a normal listening condition and in a condition with simulated hearing loss (data of the latter condition are not reported here). In each listening condition, participants rated all stimulus utterances, so each participant rated each utterance twice. The order of listening conditions was counterbalanced across participants. The two different lists were used to present listeners with a different order of the stimuli in the two listening conditions. There was a short break between the two listening conditions.

Prosodic Parameters

Affect ratings provided by the participants in our study were related to three acoustic parameters which have been related to arousal in the literature: F0, intensity and vocal effort (Pereira, 2000; Schmidt et al., 2014; Schröder et al., 2001; Schröder, 2006). The ranges of the three prosodic parameters in our stimuli are presented in Table 1.

Table 1. Ranges of the prosodic parameters in our stimulus materials.

	MIN	MAX
Mean Intensity [dB SPL]	53	71
Vocal Effort	-30.1	8.0
Mean F0 [Hz]	116	344

Mean F0 and mean intensity were calculated for each utterance by averaging over the utterance using Praat (Boersma & Weenink, 2013). As mean intensity may be affected by situational factors such as the distance between the speaker's mouth and the microphone (cf., Sluijter & Heuven, 1996), we also included a more robust acoustic correlate of perceived intensity linked to the speaker's vocal effort.

Table 2. Intercorrelations between prosodic parameters and correlations between prosodic parameters and the reference ratings for arousal (Pearson's Correlation Coefficients) for our stimulus materials.

	Mean F0	Mean Intensity	Reference Arousal Ratings
Mean F0	–		.71 ***
Mean Intensity	.79 ***	–	.91 ***
Vocal Effort	.25	.39	.43 *

Notes. * $p < .05$, *** $p < .001$

In order to quantify vocal effort, we used a spectral measure (Tamarit, Goudbeek, & Scherer, 2008), namely the Hammarberg Index (Hammarberg, Fritzell, Gauffin, Sundberg, & Wedin, 1980). The Hammarberg Index is defined as the intensity difference between the maximum intensity in a lower frequency band [0 – 2 kHz] versus that in a higher frequency band [2 – 5 kHz]. In this study, the Hammarberg Index energy distribution measure (henceforth: vocal effort) was measured across the entire utterance. Table 2 shows the correlations between the different prosodic parameters in our stimulus set, and of the correlations of the prosodic variables and the reference arousal ratings that came with the corpus.

Loudness Scaling Task

Loudness processing ability was assessed by measuring individual hearing threshold level and loudness discomfort level. To that end, a loudness scaling task was administered in which participants listened to frequency-specific hearing assessment noise (FRESH noise[3], Aurical AUD, GN Otometrics) presented at different loudness levels, and were asked to rate the loudness of the stimulus. Noise was presented at three different center frequencies: 0.5 kHz, 1 kHz, and 2 kHz. These frequencies were chosen as intensity in the mid-frequencies contributes most to perceived loudness (Sluijter & Heuven, 1996). Before the start of the test phase, participants were familiarized with the noise at different frequencies. During the test phase, stimuli were presented following an adaptive procedure via an audiometer (Aurical AUD, GN Otometrics). Presentation levels were randomized over trials and ranged from 0 dB HL to 90 dB HL with a step-size of 5 dB. Moreover, the order of the three test frequencies was mixed over trials and presentation levels. On each trial, participants indicated their subjective loudness impression of the stimulus on a graphically defined loudness scale (see Figure 2) using a touch screen and a touch screen pen.

The scale included eleven steps (black bars are intermediate steps) spanning from “not heard” to “uncomfortably loud”, and participants were presented with a German translation of the scale. Both participant groups completed the task once using headphones (Sennheiser HD 600) and once using the same loudspeaker setting as described in the sub-section ‘Emotion Task’. The order of the two presentation modes was counterbalanced across participants. Hearing aid users were tested with their hearing aids in the loudspeaker mode, and were tested without their hearing aids only in headphone mode. The two modes were necessary to

[3] FREquency Specific Hearing assessment (FRESH) noise is narrow-band noise with extremely steep filter slopes. Testing with this type of noise allows frequency-specific measurement of auditory perception thresholds.

check hearing aid fitting (cf. sub-section 'Participants'). Only the headphone mode measurement was taken as an index of individual loudness processing ability. Hearing threshold levels and loudness discomfort levels were calculated per frequency, i.e., for 0.5 kHz, 1 kHz, and 2 kHz, for each individual in both listener groups. Larger ranges indicated better loudness processing abilities. The median range served as an overall measure of loudness processing ability. The median was preferred over the mean, as it includes information about the minimum and the maximum range across frequencies.

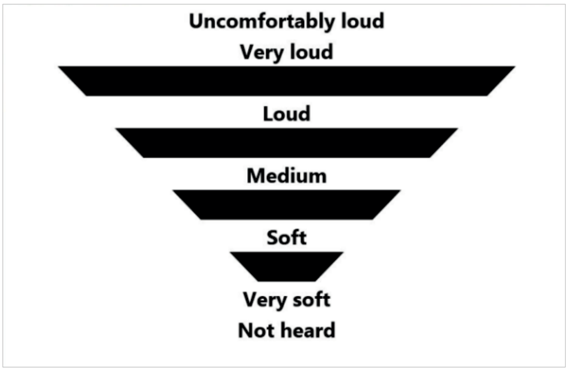

Figure 2. Graphic representation of the Loudness Scale. The scale included eleven steps (black bars are intermediate steps) spanning from "not heard" to "uncomfortably loud".

We analyzed whether the two listener groups differed in their loudness processing ability. For this between-group analysis, we excluded one participant from the group of 22 normal-hearing participants because of outlier performance on the loudness scaling task (i.e., the participant exceeded the normal-hearing listener group mean by more than two standard deviations). Loudness scaling range in the headphone condition (i.e., without hearing aid) was significantly poorer in the hearing aid users

(M = 40.7, SD = 11.7) compared to the normal-hearing listeners (M = 65.0, SD = 1.7); $t(23.1) = 9.9$, $p < .001$.

Table 3 shows correlations between the loudness processing measure, age, and hearing loss for both participant groups.

Table 3. Pearson's Correlation Coefficients between Loudness Processing Measure, Age and Hearing Ability (PTA) in Normal-hearing Participants (NH) and Hearing Aid Users (HA).

		PTA	**Loudness scaling range (median)**
NH adults	Age	.24	-.27
	PTA	–	-.45 *
HA users	Age	.29	-.17
	PTA	–	-.75 ***

Notes. * $p < .05$, *** $p < .001$

Age and hearing ability were not significantly correlated in either participant group, which may be due to how the two groups were recruited (as being either hearing-aid users, or having age-normal hearing). In the normal-hearing group, we found a moderate correlation between the overall measure for loudness scaling ability with hearing ability. This is in line with previous research (Boettcher et al., 2001; Koehnke et al., 1995). In hearing aid users, the overall measure for loudness scaling ability was highly correlated with hearing ability, such that poorer hearing was associated with a smaller loudness scaling range. However, loudness scaling ability was not correlated with age.

Results

Aided versus unaided listening

In order to assess whether wearing a hearing aid makes hearing aid users more responsive to subtle differences in acoustic parameters, we investigated the relation between arousal ratings (the dependent variable) and prosodic parameters, hearing aid use, and individual hearing loss (as predictor variables) using R (R Development Core Team, 2008). We used linear mixed-effects regression analyses with random intercepts for stimulus and participant. Note that parametric tests like regression, including linear mixed-effects models, are robust against violations of the assumption of normal distribution (cf., Norman, 2010). There are even some suggestions that the risk of finding a false positive (Type 1 error) is higher for ordered regression models for Likert scale data (i.e., cumulative link mixed models) compared to the linear mixed-effects method (cf., Kizach, 2014). Nevertheless, we also analyzed whether results obtained with the linear mixed-effect regression models were replicated using ordered regression models, i.e., cumulative link mixed models (CLMMs, cf., Agresti, 2002). Unless noted otherwise, analyses with CLMMs showed similar results (in terms of which effects were significant and the direction of the effects) to those obtained with the linear mixed-effect models.

The initial model tested for three-way interactions between listening condition (Condition: aided, unaided), individual hearing loss (PTA of the better ear across 0.5, 1, 2, and 4 kHz), and each of the acoustic parameters (Mean F0, Mean Intensity, Vocal Effort) to test the hypothesis that the use of the hearing aid may change the impact of the hearing loss on the listener's use of the prosodic information for arousal rating. Continuous predictors (i.e., all three acoustic parameters and hearing loss) were centralized. We arrived at the best fitting model by using a stepwise exclusion procedure. Interactions were removed before simple effects, and those with the highest non-significant *p*-values were excluded first.

Model fit was assessed using ANOVA. The most parsimonious model is presented in Table 4.

Table 4. Statistical model of the arousal ratings as a function of hearing aid use, acoustic parameters, and hearing loss in hearing aid users.

	β	***SE***	***p***
Condition (aided)	0.072	0.021	< .001
PTA	- 0.010	0.004	< .05
Mean Intensity	0.060	0.015	< .001
Vocal Effort	0.014	0.013	n.s.
Mean F0	0.001	8.583 10^{-4}	n.s.
Condition × PTA	0.007	0.002	< .01
Condition × Vocal Effort	0.014	0.004	< .001
PTA × Mean F0	4.464 10^{-5}	1.943 10^{-5}	< .05

The main cue for arousal, i.e., mean intensity, was stable across listening conditions as shown by the simple effect of Mean Intensity and by the lack of an interaction between Mean Intensity and Condition (Table 4). Higher intensity was rated as more aroused in both the aided and unaided condition. The simple effect for Condition suggests that ratings were generally more aroused in the aided condition than in the unaided condition (the latter mapped on the intercept). We were specifically interested in whether hearing aid use would modify listeners' use of prosodic cues. Indeed, a significant interaction of Condition and Vocal Effort was observed. Vocal effort affected arousal ratings such that higher vocal effort was associated with more aroused ratings in the aided listening condition (but not in the unaided condition). The finding was confirmed in a separate analysis where the aided condition was mapped on the intercept: the simple effect of Vocal Effort was then significant ($\beta = 0.028$, $p < .05$). This finding implies that the use of hearing aids changes the use of arousal-related acoustic cues: hearing-impaired participants make use of vocal effort, which they do not use when not wearing their hearing aids.

The simple effect of Hearing Loss (PTA) showed that participants with poorer hearing generally gave lower ratings of arousal compared to participants with better hearing (in the unaided condition). However, this was less the case in the aided condition, as a significant interaction between listening condition and hearing loss (Condition × PTA) is observed. This suggests that wearing hearing aids counteracted the effect of hearing loss on arousal perception and, eventually, made the rating patterns of poorer and better-hearing participants more alike. Furthermore, hearing loss seems to cause a shift in the use of mean F0. In general participants did not make significant use of mean F0 in either listening condition. Depending on the degree of hearing loss, however, participants used mean F0 (PTA × Mean F0) as a cue to arousal: Those with poorer hearing associated increased F0 more with higher arousal than those with better hearing across listening conditions. In other words, among hearing aid users, those with poorer hearing showed a different association between mean F0 and arousal ratings than those with relatively good hearing. Note, however, that no three-way interaction with Condition was observed, i.e., the influence of hearing loss on the use of mean F0 for arousal rating was the same across listening conditions. Similarly, neither of the other two acoustic parameters showed a three-way interaction (with Condition and PTA).

Hearing aid users versus normal-hearing controls

Secondly, we investigated whether the use of a hearing aid brings arousal rating to the level of normal-hearing older adults. To that end, we set up two separate analyses. First, we compared the normal-hearing group (mapped on the intercept) to the unaided hearing aid users to see whether groups differed in the first place. We then followed this up with a group comparison in which the hearing aid users were wearing their hearing aid. The analyses differed from the previous analysis in testing between groups, rather than within group, and consequently, did not involve the listening condition variable. The initial model allowed for two-way interactions between Group (hearing aid users vs. normal-hearing) and the acoustic parameters listed above to compare the two groups in their use of the acoustic parameters for arousal rating.

Unaided listening versus normal-hearing controls

The results of the statistical analysis are presented in Table 5. As before, the results showed a significant effect of mean intensity: higher intensity is associated with more aroused ratings. This was the case for both participant groups because no interaction between group and mean intensity was observed. Hence, mean intensity was a reliable cue when rating arousal and the use of it was not impaired in unaided hearing aid users.

Table 5. Statistical model of arousal ratings and cue use in unaided hearing aid users versus normal-hearing listeners.

	β	***SE***	***p***
Mean Intensity	0.060	0.014	< .001
Vocal Effort	0.023	0.012	.07
Mean F0	0.001	0.001	n.s.
Group (unaided hearing aid users)	-0.102	0.059	.09
Group × Vocal Effort	-0.009	0.004	< .05
Group × Mean F0	0.001	0.001	< .05

No clear group differences were observed given that the general effect of Group was only marginally significant. Note that the interaction between Group and Mean F0 was not replicated in the model for ordinal data (CLMM). Furthermore, neither group showed a significant use of mean F0 (as verified by mapping the hearing aid users group on the intercept), which makes it difficult to interpret this interaction.

Likewise, the interpretation of the results obtained for vocal effort was not straightforward. The effect of vocal effort on arousal rating did not reach significance in this analysis (the effect is marginal in Table 5 but was significant in the CLMM). The interaction between group and vocal effort in Table 5 indicates that the vocal effort effect (be it that this effect is non-significant for the normal-hearing group already) is diminished further for unaided hearing aid users (Group × Vocal Effort). This finding was supported by a separate analysis in which hearing aid users were mapped on the intercept: Vocal effort did not affect their arousal ratings ($\beta = 0.014$, $p = 0.26$).

Aided listening versus normal-hearing controls

Next, we compared affect ratings of hearing aid users and normal-hearing older adults. The result of the best fitting model is shown in Table 6. As in the previous arousal analyses, a significant simple effect of mean intensity was found: higher mean intensity was associated with more aroused ratings. Greater vocal effort was also associated with more aroused ratings but note that this effect was only marginally significant (as in the previous analysis, it did show up significantly in the CLMM for ordinal data). Combined with the results of the previous analysis, these results suggest that vocal effort tends to be a secondary cue to arousal in the normal-hearing group and in aided listening (mean intensity being the primary cue in our study), but not in unaided listening. As such, the use of a hearing aid restores the availability of the weaker vocal effort cue (which mirrors the results reported in Table 4).

Ratings of the hearing aid users did not generally differ significantly from the normal-hearing participants (mapped onto the intercept; no significant effect of Group). However, the use of mean intensity differed between the two listener groups as shown by the interaction between

Group and Mean Intensity: hearing aid users responded more strongly to differences in intensity than participants with age-normal hearing.

Table 6. Statistical model of arousal ratings and cue use in aided hearing aid users versus normal-hearing listeners.

	β	***SE***	***p***
Mean Intensity	0.062	0.014	< .001
Vocal Effort	0.025	0.013	.08
Group (aided hearing aid users)	- 0.030	0.053	n.s.
Group × Mean Intensity	0.009	0.004	< .05

Arousal rating and loudness processing ability

The third aim of this study was to investigate the influence of individual loudness processing ability on arousal ratings and, more particularly, on the use of prosodic cues. Given the limited variability in loudness processing ability in normal-hearing participants, this analysis only involved hearing aid users.

We hypothesized that participants need to be sensitive to differences in prosodic parameters that are related to intensity for successful arousal perception. We, therefore, investigated the effect of loudness processing ability and of hearing aid use on the use of intensity and vocal effort when rating arousal (dependent variable). The initial model allowed for three-way interactions between listening condition (aided/unaided) and the participant's median loudness scaling range (unaided) on the one hand, and the three prosodic parameters mean intensity, vocal effort, and F0 on the other (F0 is kept in this initial model to keep the model comparable to all previous models). The model stripping procedure was identical to the one reported for the previous analyses.

Table 7. Statistical model of the arousal ratings: Ratings given by hearing aid users as a function of hearing aid use, the median loudness scaling ability across 500 Hz, 1000 Hz, and 2000 Hz, the use of mean intensity, and the use of vocal effort.

	β	***SE***	***p***
Loudness Scaling Range	0.008	0.003	<.05
Mean Intensity	0.070	0.013	<.001
Vocal Effort	0.013	0.013	n.s.
Condition (aided)	0.072	0.021	<.001
Condition × Loudness Scaling	- 0.005	0.002	<.05
Condition × Vocal Effort	0.014	0.004	<.001

The result of the best fitting model is presented in Table 7. The significant simple effect of loudness scaling range suggests that arousal ratings are generally affected by the loudness processing ability of the hearing aid user. As expected, larger loudness scaling ranges, indicating better loudness processing ability, are associated with more aroused ratings in unaided hearing aid users. This effect, however, is diminished in the aided condition, as suggested by the interaction between Condition and Loudness Scaling. In other words, while individual loudness processing abilities may impact arousal perception in hearing-impaired listeners, the use of hearing aids cancels out this effect (as the loudness scaling effect turned non-significant if the aided condition was mapped on the intercept).

No interactions between Loudness Scaling Range and either Mean Intensity or Vocal Effort were observed, which means that loudness processing ability did not modulate the use of prosodic cues underlying arousal. Simple effects of Mean Intensity and Condition, as well as the interaction between Condition and Vocal Effort, were similar to the effects reported in the section on Aided versus unaided listening.

Hearing loss versus loudness processing ability

As loudness processing was significantly correlated with hearing loss, we also investigated whether inclusion of loudness processing ability explained any additional variance beyond the effects of hearing loss observed in our earlier analysis. To that end, the measure for loudness scaling ability was added to the model reported in Table 4. Loudness scaling range did not surface as a significant predictor in this extended model and was therefore stripped from the model. Consequently, the 'extended' model was identical to the hearing loss-only model reported in Table 4. This indicates that loudness processing ability did not explain additional variance beyond the effects of hearing loss.

General discussion

The current study investigated arousal rating in speech in older hearing-impaired and normal-hearing adults. The choice to focus on arousal was driven by the observation that arousal is strongly cued by prosodic parameters in the speech signal, while the link between affect perception and prosodic information is less strong for valence (Lima, Alves, Scott, & Castro, 2014; Schmidt et al., 2016). Therefore, hearing loss, i.e., the sensory deprivation of prosodic information, is more likely to hinder arousal perception than valence perception. Our aim was threefold: First, we focused on the influence of hearing aid use and hearing loss on listeners' use of the acoustic parameters cueing arousal. Second, we assessed whether the use of hearing aids brings arousal perception, in response to the acoustic information in the speech fragments, to the level of an age-matched group with age-normal hearing. Third, we investigated the influence of individual loudness processing ability on the use of the prosodic parameters cueing arousal for affect rating, particularly those related to perceived loudness, i.e., intensity and vocal effort.

Hearing aid users wearing their hearing aids generally showed the same pattern of affect ratings as participants with age-normal hearing.

This indicates that hearing aids restore the information necessary for arousal rating in older adults. Hence, while aided hearing-impaired children perform poorly compared to normal-hearing peers on affect rating tasks (Most & Michaelis, 2012; Most et al., 1993), there seems to be no such difference, at least not for arousal rating, between hearing-impaired and normal-hearing older adults. This finding is highlighting that affect perception results on hearing-impaired children indeed cannot be directly transferred to hearing-impaired older listeners.

Hearing aid use did, however, change the way the main prosodic parameter mean intensity contributed to arousal perception. In fact, hearing aid use improved the perception of mean intensity in hearing-impaired listeners compared to age-matched normal-hearing adults. Listeners wearing a hearing aid were actually more responsive to intensity differences than participants in the normal-hearing group. This may be because hearing in the normal-hearing group was normal for their age, but still implied elevated high-frequency thresholds. Consequently, older adults in the normal-hearing group were less responsive to at least some acoustic differences than the hearing aid users. The role of another intensity-related prosodic parameter, i.e., vocal effort, will need further investigation. Vocal effort was not a reliable predictor of the level of arousal, not in normal-hearing older adults, nor in unaided hearing aid users. Nevertheless, there was some evidence that hearing aid users did pick up on the vocal effort cue to arousal while wearing their hearing aids. Clearly, however, vocal effort is only a secondary cue to arousal in our materials and set up, with intensity being the stronger and more reliable cue.

Mean F0 was generally not predictive of arousal ratings in neither group nor listening condition. Nevertheless, the degree of hearing loss in hearing aid users influenced the use of mean F0 when rating arousal. Regarding the current results we can only speculate upon the origin of these findings. Possibly, mean F0 is only used when listeners are facing extreme auditory deprivation while rating arousal. This interpretation is supported by previous studies showing that mean F0 is a robust cue in listening situations where the input is degraded, e.g., due to hearing impairment (Bilger & Wang, 1976) or in degraded listening conditions (Winn, Chatterjee, & Idsardi, 2013).

Among the hearing aid users (while listening without their hearing aids), we found an effect of hearing loss on the perception of arousal, such that poorer hearing led to less aroused ratings. This effect was counteracted by the use of the hearing aid. This finding of hearing loss effects on affect perception contrasts with prior work on affect perception in older adults which showed no effect of hearing loss (Dupuis & Pichora-Fuller, 2015; Mitchell, 2007; Orbelo et al., 2005). Several factors may account for this difference in findings. First, it may be due to the fact that these prior studies investigated older adults with rather mild hearing loss, which could possibly even be rated as clinically normal for that age group. This interpretation is supported by a previous study which did not show an effect of hearing loss on arousal perception for older adults with relatively mild hearing loss (Schmidt et al., 2016). Second, the difference may be due to the fact that most prior work used acted speech material (cf., Grant, 1987), in which realization of affect may be 'overacted' and more prototypical, and hence, easier to pick up, than in conversational material. Third, the difference in results may be due to methodological factors. One of the methodological novelties of the current study was the use of a dimensional approach, in which the arousal dimension was investigated independently from the valence dimension. A second methodological novelty was the focus on relating the arousal dimension to the prosodic information coded in the speech signal. Previous work has studied the perception of emotion categories, where information on arousal and valence is blended, and consequently, where the link between the emotion category and acoustic information is less clear-cut than for the single dimension arousal. By presenting participants with a task in which they can focus on the arousal dimension, which is clearly linked to prosodic information encoded in the speech signal, we think we are better able to capture the consequences of sensory deficits on affect perception and rating. Possibly, hearing-impaired participants in previous studies may have compensated for their sensory deficit by relying more on valence information in order to classify emotional meaning.

Note that in realistic listening conditions, the effect of hearing loss on arousal perception may particularly cause misclassifications in pairs of emotional categories that have similar valence but different degrees of arousal such as hot versus cold anger, sadness versus despair, anxiety versus fear. Indeed, it had been documented that these emotion pairs are often misclassified, even in normal-hearing listeners (Banse & Scherer, 1996; Johnstone & Scherer, 2000) and it is conceivable that poorer hearing impairs the detection of this subtle distinction.

Interestingly, the effect of hearing loss on arousal perception could not be explained by decreased sensitivity to loudness cues, such as variation in mean intensity and vocal effort. Instead, hearing loss impacted the use of mean F0: Those with poorer hearing associated an increase in F0 more with higher arousal than those with better hearing across listening conditions. This may suggest that listeners with more severe hearing impairment turn to the use of alternative cues to compensate for their sensory deficit. However, caution is required when interpreting the latter finding. Although pitch cues are considered the most telling component of affective prosody (Hammerschmidt & Jürgens, 2007; Mozziconacci, 1998; Rodero, 2011), neither participant group showed a significant use of mean F0 when rating arousal. Moreover, the interaction with hearing loss was not replicated in the model for ordinal data. Further research would be needed to investigate this possible link between degree of hearing impairment and the use of (alternative) prosodic cues in the speech signal.

In order to evaluate the degree of arousal in speech, listeners need to be able to detect subtle differences in prosodic information. Given that loudness processing abilities have been reported to be impacted by hearing loss (Boettcher et al., 2001; Koehnke et al., 1995), we investigated whether individual hearing loss and individual loudness processing abilities modulated arousal ratings or the use of arousal-related cues in hearing users. A correlation between hearing loss and loudness scaling ranges was found in the data, for both normal-hearing subjects ($r = -.45$, $p < .05$) and hearing aid users ($r = -.75$, $p < .001$). Loudness processing abilities were found to be predictive of arousal ratings in the unaided listening condition: those who were more sensitive to loudness differences in the loudness scaling task generally provided more aroused ratings when listening without their hearing aid. This was, however, a general effect, as the way in which participants made use of mean intensity differences was not influenced by individual loudness processing ability. However, the use of the hearing aids cancelled out this loudness scaling effect on rating behavior among individuals. The effect of loudness processing ability on rating only surfaced when loudness processing abilities were investigated as only individual predictor. Once hearing loss was taken into account, loudness processing ability did not explain any additional variance beyond that explained by hearing loss. These results confirm the tight link between hearing loss and loudness processing ability, and their similar relationship regarding the interpretation of affect in speech.

In summary, the current study shows that older hearing aid users do not generally differ from their normal-hearing peers in their perception of arousal. Although we found effects of individual loudness processing abilities and the degree of hearing loss in unaided hearing-impaired listeners, only the effect of hearing loss seems to be relevant for arousal perception since loudness perception abilities did not explain any additional variance. The effect of hearing loss was cancelled once these listeners were wearing their hearing aids. Given that impaired arousal perception may have consequences for the fine differences in affect perception (e.g., hot versus cold anger), the current results underline the importance of hearing aids in the rehabilitation of affect perception.

Chapter 6

General Discussion

This thesis investigated how individual differences in age and hearing ability relate to the uptake and processing of semantic and affective information in older adults' speech comprehension. Below, I will discuss the findings on the processing of semantic information first, followed by the findings on the processing of affective information. I will end with a conclusion and an outlook on possible avenues for future research.

Processing of semantic information – Influence of cognitive load and mild hearing loss

In **Chapter 2**, I investigated the hypothesis that understanding spoken language might be hampered by cognitive load induced by a secondary task. This would be the case, for instance, if you are listening to someone talking to you while at the same time trying to remember the things you wanted to buy from the supermarket. Speech understanding in **Chapter 2** is quantified by way of the semantic priming effect, which refers to the fact that the recognition of a word is facilitated by an earlier occurring semantically related word. For instance, the response to a target word (e.g., vampire) is faster if it is preceded by a semantically related (prime) word (bat) as compared to an unrelated word (dog). Hence, speech understanding of a target word is facilitated if it occurs in a context of words associated in meaning.

The "attention modulation hypothesis" (Smith, Bentin, & Spalek, 2001) states that semantic priming is modulated by attention, i.e., words are only processed deeply enough to elicit significant semantic priming if participants' attention is concentrated on the prime. As such, the size of the semantic priming effect may depend on whether listeners' attention is drawn to (or away from) the prime words.

Listening effort may also influence how attentive a listener is when listening to speech and hence may influence semantic priming. Listening effort depends on listener-external factors, such as the presence of noise, and listener-internal resources, such as working memory capacity (cf.

Lemke & Besser, 2016). Listener-internal resources available for listening are also subject to age-related changes. For instance, listening to speech can be particularly effortful for hearing-impaired listeners (Pals, Palsanastasios, Van Rijn, & Başkent, 2015) and when cognitive processing resources are extensively taxed, such as when listeners are dual-tasking (Mattys & Palmer, 2015).

This link between hearing impairment and listening effort is also accounted for in the recent FUEL framework laid out by Pichora-Fuller and collaborators (Pichora-Fuller et al., 2016), which follows up on Kahneman's (Kahneman, 1973) Capacity Model of Attention. The FUEL framework describes how listening effort depends on hearing difficulties and task demands, as well as on the listener's motivation to spend effort on the listening task at hand. As such, the framework describes how fewer resources are left for speech processing in those with hearing impairment. Evidence for the relationship between effortful listening and available resources is provided by the study by Pals, Sarampalis and Başkent (Pals, Sarampalis, & Başkent, 2013) who measured participants' listening effort while listening to speech of systematically manipulated speech intelligibility (CI-simulated speech with varying vocoding settings resulting in differential intelligibility). Listening effort in their study was quantified by investigating RTs on a secondary task. Their results show that if listening is facilitated due to a less distorted signal, more resources were left for the processing of either linguistic or nonlinguistic stimuli.

The current work hypothesized that listening effort, caused by a cognitively challenging secondary task, does not only influence speech processing and comprehension, but also its underlying processes such as semantic priming. Moreover, it was investigated whether the interaction between additional task demands and semantic priming would also be modulated by age-related listener-internal resources, namely hearing sensitivity and cognitive abilities. That is, whether the impact of additional cognitive load on semantic priming (and hence on comprehension) would be greater for those for whom listening is already challenging to begin with because of impaired hearing and/or limited cognitive resources.

The influence of cognitive load

The results of the semantic activation study showed significant effects of semantic priming and of cognitive load on response latencies. This indicated that words were recognized faster when preceded by more strongly associated prime words compared to less strongly related words. In addition, words were recognized faster in the condition with the lower cognitive load. Contrary to the hypothesis, there was no significant interaction between cognitive load and semantic priming showing no evidence that cognitive load in itself influenced semantic priming.

Even though previous studies had found mixed results on whether cognitive load affected semantic activation, increased cognitive load was expected to affect semantic activation for two reasons. First, contrary to other designs (Otsuka & Kawaguchi, 2007; Smith et al., 2001), the setup used in the current study continuously taxed working memory. Second, in contrast to studies with student populations (Mattys & Wiget, 2011), my study investigated older adults, whom I expected to be impacted more by a cognitive load manipulation because of age effects on cognitive capacity.

The lack of an interaction between the cognitive load manipulation and semantic activation at the group level in my study could be argued to suggest that participants in my study were able to single-task, putting the memory task on hold while processing the word to make a lexical decision. However, the finding of a robust effect of the load manipulation on lexical decision latency makes this explanation unlikely. Possibly, then, the lack of an interaction could be attributed to the fact that the semantic priming effect was already small in the low-load condition. As such, there was limited power to observe a significant reduction of the semantic priming effect. The stimuli and stimulus design were based on an earlier study with English materials (Van de Ven, Tucker, & Ernestus, 2011) that were translated to Dutch (and adapted, when necessary). Follow-up research is needed to investigate whether the translation of the materials or the difference in cognitive loading caused the small semantic priming effect.

Despite the lack of a general interaction between the cognitive load manipulation and semantic activation at the group level, my second observation was that there was a triple interaction between the participant's attentional ability, the effect of increased cognitive load, and semantic priming. Somewhat unexpectedly, individuals with poorer attention-switching control showed stronger semantic facilitation in the low-load condition compared to those with better attention-switching control. This may suggest that they spent extra effort on the low-load condition, while being overtaxed in the high-load condition, and no longer able to process the prime deeply and quickly enough. Consequently, the difference in semantic priming between the low-load and high-load condition was larger for the participants with poorer than for those with better attentional skills. Contrary to the group-level finding, this effect provides some support for the attention modulation hypothesis (Smith et al., 2001), by showing that semantic priming relies on the ability to allocate sufficient resources to primes.

In sum, the individual differences analysis suggests that imposing a secondary task or distraction may affect the integration of words into a coherent semantic representation. Like Pals, et al. (2013), my results show that cognitive load modulates (semantic) lexical processing, mainly in participants with poorer attentional skills. My results extend their findings by showing that the amount of resources available modulates message comprehension through associations between words.

The influence of hearing loss

Concerning the relation between signal clarity and semantic comprehension, the process of integrating words into their semantic context has been shown to be hampered by decreasing speech intelligibility (Obleser & Kotz, 2011). Likewise, degradation of speech input by low-pass filtering has been shown to result in reduced semantic priming compared to unfiltered speech (Aydelott & Bates, 2004). As low-pass filtering was considered a rough approximation of high-frequency hearing loss, hearing loss in my study was expected to interfere with rapid prime processing and hence semantic activation. In order to compensate for degraded sensory input, those listeners with more hearing loss were expected to allocate more cognitive resources to recognizing the speech (Rönnberg et al., 2013), and they were expected to be more vulnerable to increased cognitive load. However, no effect of hearing loss was observed on word processing, nor did hearing loss interact with the cognitive load manipulation.

Possibly, in contrast to the drastic effect of filtering on speech intelligibility, the range of hearing losses in my sample was not large enough to observe a clear effect of hearing loss on word processing (mean PTA_{Best} = 22.17 dB HL, range = 1.67 – 43.33 dB HL). Consequently, even for those participants with the poorest hearing in my relatively good-hearing sample of older adults, the perceptual load was too subtle to observe any effects, or to see the need for recruitment of additional resources. As only few studies have looked at effects of mild hearing loss on speeded word recognition tasks, it was not clear what (subtle) amount of hearing loss would result in detectable effects on the processing of single words. In sum, future research with a larger range of hearing loss differences is required to investigate the question of the role of hearing loss on semantic activation.

Processing of affective information – Influence of age and mild to moderate hearing loss

Affect in speech is expressed through various acoustic-prosodic parameters. Listeners make use of this prosodic information to derive a speaker's emotional state (Banse & Scherer, 1996; Coutinho & Dibben, 2013; Scherer, 2003). General age differences in affect categorization have been reported in several studies (Dupuis & Pichora-Fuller, 2015;

Isaacowitz et al., 2007; Kiss & Ennis, 2001; Lambrecht, Kreifelts, & Wildgruber, 2012; Mitchell, Kingston, & Barbosa Bouças, 2011; Orbelo, Grim, Talbott, & Ross, 2005; Paulmann, Pell, & Kotz, 2008), with most studies focusing on age differences in the classification of emotional categories. These studies all showed that older adults were poorer in accurately recognizing affect from prosodic cues in speech compared to younger adults. However, whether age-related hearing loss may (partially) account for this age difference had so far remained somewhat unclear.

In Chapters 3 and 4, age differences in the perception of affect in speech were investigated by having younger and older adult listeners rate short audio fragments taken from a talk show. Similar to some earlier studies (e.g., Lima, Alves, Scott, & Castro, 2014), I specifically related younger and older adults' rating of utterances along the affect dimensions of arousal and valence to the prosodic realization of the utterances. Importantly, however, no earlier study has specifically linked individual hearing loss to the uptake of acoustic information from the emotional utterances. Furthermore, most studies so far have used acted speech materials in which realization of affect is more pronounced (and hence more evident) than in more natural stretches of conversational speech (Laukka, Neiberg, Forsell, Karlsson, & Elenius, 2011). The few studies on spontaneous affect expression in everyday speech (e.g., Campbell, 2005; Cowie & Cornelius, 2003) have reported that milder expression of affective state is much more common than pronounced expression of affective state in conversational speech.

The conversational speech fragments selected for the studies in this thesis came from the "Vera am Mittag" corpus, which is a speech corpus based on dialogues taken from a German talk show. The unscripted, spontaneous, and authentic nature of the dialogues (Grimm, Kroschel, & Narayanan, 2008), evidenced by the absence of a clear acting attitude in the guests made this talk show-based corpus material suitable as a proxy for the realization of emotional content in spontaneous dialogue.

The studies reported in Chapters 3 and 4 (investigating separate participant groups) set out to test for age differences in the uptake of acoustic/prosodic parameters for the rating of affect in spoken fragments. Additionally, the studies aimed to investigate a possible link between the uptake of acoustic-prosodic information in the speech fragments and individual hearing sensitivity differences among the older adults. Whereas the preliminary study reported in Chapter 3 contained small participant samples (with ten participants per age group), age groups were larger in Chapter 4 (with twenty participants per age group). Moreover, hearing abilities of the participants in the study in Chapter 4 were slightly worse, i.e., normal to moderately impaired hearing ability, compared to the sample investigated in Chapter 3. Based on the insights obtained in the preliminary study, a more elaborate analysis was carried out in Chapter 4 and findings were discussed in more detail.

The prosodic parameter that most reliably predicted arousal ratings in Chapters 3 and 4 was the spoken fragment's mean intensity. Correlation analyses showed that the reference arousal ratings (that came with the "Vera am Mittag" speech corpus) correlated with mean intensity, as well as the utterance's mean F0 and vocal effort. However, if all prosodic parameters are entered simultaneously into the statistical analysis, as was done in the present study, only intensity surfaced as the most reliable predictor of arousal ratings. This observation of intensity being the primary cue for arousal agrees with some, but not all, other studies on affect perception. Similar results were found by Lima and colleagues (Lima, Castro, & Scott, 2013), who applied multiple regression analyses to relate acoustic measurements of their acted speech materials to emotion category ratings, as well as to emotion dimension ratings (such as arousal, valence). The acoustic cue that explained most variance in Lima et al.'s arousal ratings was intensity (beta weight of .63), followed by a center of gravity measurement (beta weight of .53), the latter being a measure of spectral tilt and vocal effort, as was the Hammarberg Index employed in Chapters 3 and 4. Similarly, Laukka and colleagues (Laukka et al., 2011) showed that listeners' ratings of irritation in real-life voice-controlled telephone services correlated more strongly with intensity measures than with measures of spectral tilt.

On the other hand, as noted already in Chapter 4, it is somewhat unexpected that intensity turned out to be the primary cue to arousal, rather than a measure of vocal effort. Intensity could simply be influenced by speakers' distance from the microphone rather than their emotional state, whereas vocal effort is not susceptible to such recording factors. Sauter and colleagues (Sauter, Eisner, Calder, & Scott, 2010) obtained different results regarding which acoustic parameter is the primary cue for arousal. Their multiple regression analyses, relating listeners' arousal ratings to acoustic measures, showed higher beta weights for pitch and spectral tilt measures than for intensity measures. However, our results on intensity rather than spectral tilt being associated with arousal are in line with Goudbeek and Scherer (2010).

The lack of consistent findings on which acoustic cue best predicts listeners' rating of arousal could be due to multiple factors. First of all, studies differ in which acoustic measures they include, particularly in which measure is included as an index of vocal effort (Hammarberg Index, or spectral center of gravity, or the ratio between the amplitudes of F0 and those of higher harmonics or formants). Second, studies on affect perception in unscripted speech (like the present studies, and e.g., Laukka et al., 2011) do not have full control over the content of the utterances, such that semantic content may have played a role in the interpretation of affect. The latter possibility particularly holds for studies, such as the ones presented in this thesis, with a relatively limited set of utterances. Nevertheless, studies on the interpretation of affect across acted and spontaneous materials thus largely agree on the top three of most informative cues (intensity, spectral tilt and F0).

Arousal rating, age and hearing loss

Several studies have found that older adults are less accurate in the recognition of emotions from the voice than younger adults (e.g., Paulmann et al., 2008). In line with previous studies, significant age effects were observed in Chapter 3: older adults generally perceived utterances as less aroused than younger adults. This general age effect on arousal rating was, however, not replicated with the larger participant sample in Chapter 4. Nevertheless, Chapters 3 and 4 both showed an age difference in the interpretation of intensity information for arousal rating, such that the effect of intensity on arousal ratings was less pronounced in older adults' than in younger adults' ratings.

As argued in the General Introduction, age differences in affect perception may arise early, namely during auditory decoding of speech, but may also arise during processing stages beyond the decoding stage. Age differences in emotion processing have, for instance, been attributed to age-related structural and functional changes in the "social brain" (Ruffman, Henry, Livingstone, & Phillips, 2008). Hence, changes in higher-order processing, rather than or in addition to auditory problems in older adults, may result in less differentiation of emotional input than in younger adults. Here, I address the question whether (part of) the age difference could be accounted for by (age-related) differences in auditory decoding of speech. This was addressed by investigating whether individual differences in hearing sensitivity among the older adults related to their use of intensity differences for rating arousal. The results presented in Chapters 3 and 4 did not provide any evidence for hearing sensitivity differences playing a role in explaining the often-observed age difference in affect perception, in line with results of Orbelo et al. (2005). Chapter 4 argued that the perception of arousal may not be affected much by mild to moderate degrees of hearing loss, as present in my older adults samples, because arousal is robustly encoded by multiple parameters (cf. Chapter 4, Table 1; Pereira, 2000; Schröder, 2006).

To investigate the idea that the clear and redundant encoding of arousal in the acoustic signal makes arousal perception robust against mild hearing loss, the final Chapter (Chapter 5) of this thesis investigated whether more severe hearing loss affects arousal rating. In this study with hearing-impaired older adults, the degree of hearing loss varied across participants (range 32.5 – 68.8 dB HL). In contrast to the findings obtained with individuals with mild hearing losses in Chapters 3 and 4, a significant effect of hearing loss on arousal ratings was observed among participants (not wearing hearing aids) in Chapter 5.

Chapter 5 also specifically investigated whether wearing a hearing aid makes listeners more responsive (in their affect ratings) to subtle differences in the acoustic variation naturally present in the speech materials. Indeed, the results showed that wearing hearing aids counteracted the effect of hearing loss on arousal perception by decreasing the difference between rating patterns of poorer and better-hearing participants. Furthermore, Chapter 5 was set up to investigate whether the use of hearing aids brings arousal perception back to the level of an age-matched group with age-normal hearing. The present finding suggests that hearing aid user's ratings of arousal did not differ significantly from ratings given by their peers who had normal hearing for their age.

The third question that Chapter 5 addressed another aspect of acoustic decoding of prosodic cues to affect. Older adults' uptake of auditory information may not just be related to the mere audibility of the stimuli, which would follow from hearing sensitivity. It may also be related to their processing of auditory information. Auditory processing ability is often measured by tasks in which listeners have to discriminate between stimuli differing in signal duration, intensity, or frequency (i.e., temporal, intensity, or frequency resolution). Note that auditory processing ability and hearing sensitivity are often correlated. Chapter 5 assessed the possibility that auditory processing abilities, i.e., auditory deficits that do not show up in the pure-tone audiogram, influence affect perception, beyond the effect of hearing loss. The focus was on loudness processing ability, as the loudness dimension, rather than temporal or frequency resolution, was most clearly related to the rating of arousal. The results showed that when participants were not wearing their hearing aids, their loudness processing abilities affected arousal ratings. Furthermore, the effect of loudness processing ability on rating behavior was diminished when listeners used their hearing aids. Importantly, however, individual loudness processing abilities did not explain any additional variance in arousal ratings beyond hearing loss effects. As such, age-related decline in auditory processing abilities is not likely to contribute to age differences in arousal perception beyond the contribution of (severe forms of) hearing loss.

The above results thus clearly show that the frequently observed age effect in affect perception can partly, but not fully, be accounted for by age-related changes in hearing. Older adults' arousal perception is only affected by hearing loss if the degree of hearing loss is more severe. Hearing rehabilitation, through the use of hearing aids, improved the uptake of intensity variation in the utterances. Furthermore, hearing loss impacted the use of mean F0. These findings may suggest that listeners with more severe hearing loss, to compensate for their sensory deficit, shift their affect perception to pitch cues. I will return to this compensatory cue use below in the discussion of the valence results.

Valence rating, age and hearing loss

The affect dimension valence is less straightforwardly related to prosodic parameters. In contrast to strong correlations between arousal and several prosodic parameters reported in the literature, correlations were weaker and less consistent for valence (Goudbeek & Scherer, 2010; Pereira, 2000; Schröder, Cowie, Douglas-Cowie, Westerdijk, & Gielen, 2001; Schröder, 2006;). This in itself could entail that different groups of participants may rely on different cues when rating valence. The weak relation between valence and prosodic parameters for my materials is also clear from Table 1 in Chapter 4, which shows that none of the prosodic parameters was significantly associated with the reference ratings (most probably provided by young adults) that came with the speech corpus. Note that this lack of clear correlations between acoustic parameters and valence could be due to the limited size of the material subset (cf., Goudbeek & Scherer, 2010 for acoustic correlates of valence, such as spectral slope, given a much larger data set). Nevertheless, in Chapters 3 and 4, F0 turned out to be a predictor of valence ratings for my speech materials across both age groups, with higher mean F0 leading to more negative ratings.

Even though F0 did not show up significantly in the multiple regression analysis of valence ratings by Sauter and colleagues (Sauter et al., 2010), their results show a similar trend in that higher F0 was associated with more negative ratings. Spectral tilt (and variability thereof) showed up as the only reliable predictor of valence in their study, with mean F0 (nonsignificant) having the next highest beta weight. From the size of the beta weights reported in the study by Lima and colleagues (Lima et al., 2013), harmonics-to-noise ratio (as a measure of voice quality) was the most promising (marginally significant) predictor of valence, followed by mean F0. Thus, F0 showed up across these three studies, as a potential but weak acoustic cue to valence.

Chapter 4 showed that younger and older adults differ in their interpretation of F0 differences for their valence ratings, and that, among the older adults, hearing sensitivity modulates the effect of F0 on valence ratings. Older adults show larger effects of F0 than younger adults (which lines up with the lack of a correlation between F0 and the reference valence ratings provided by young adults), and among older adults, those with poorer hearing show larger effects of pitch on their valence rating.

This ties in with some of the findings on hearing loss and hearing rehabilitation in Chapter 5. Whereas F0 was generally not a cue to arousal, among the hearing aid users in Chapter 5, those with poorer hearing associated increased F0 more with higher arousal than those with better hearing, both in the listening conditions with and without hearing aids. Similarly, the age difference in how F0 predicts valence ratings can be accounted for by differences in hearing sensitivity. The only interaction between age and the use of prosodic cues in Chapter 3 was that between age and intensity: older adults interpreted higher intensity as more negative than younger adults. As mean intensity and mean F0 are mutually dependent and highly correlated, the seemingly different findings in Chapter 3 (an interaction between age group and mean intensity) and Chapter 4 (an interaction between age group and mean pitch) in fact both suggest that older adults attach different weights to the prosodic information in the speech signal in their interpretation of valence than younger adults.

Chapter 4 argued that the observation of age differences in the interpretation of prosodic information fits with data from Lima et al. (Lima et al., 2014), who also observed that mean F0 was a stronger predictor for ratings of fear and sadness (negative valence) in older than younger adults. Note, however, that Lima and colleagues did not investigate the impact of individual hearing thresholds on older adults' ratings. Lima and colleagues argue that even if older adults are equally efficient in using certain acoustic cues as younger adults, age groups may still differ in how they weigh the cues when rating affect. In fact, the results by Lima and colleagues show that the variance in emotion category ratings explained by acoustic cues was equally large for younger and older adults, such that age groups were equally efficient and consistent in interpreting acoustic information for their emotion ratings. However, age groups differed in which information they attend to. This differential pattern of weighing types of acoustic information may originate from age-related differences in higher-level processing of affect, such as age-related changes in neuroanatomy or neural processing of emotional content (cf. Ruffman et al., 2008). My observation that hearing loss does relate to the use of F0, both for arousal and valence rating, suggests that, in addition to potential age-related changes in higher-level processing of affect, age-related hearing loss may also make listeners less certain about what to pay attention to, leading them to focus more on some, relatively salient, prosodic parameters such as F0 changes compared to younger or better-hearing listeners.

Conclusion and outlook

The ability to successfully communicate with others contributes substantially to older adults' health and quality of life (Bath & Deeg, 2005). A considerable part of communication is the comprehension of speech, which concerns what is said and how it is said. This thesis investigated how age, cognitive abilities, and hearing acuity may influence the processing of semantic and affective information encoded in acoustic-prosodic information. My results show that both age and age-related changes in cognitive (attentional) and auditory skills show significant effects on the processing of these sources of linguistic information.

While imposing additional task demands on listeners did not influence semantic processing across listeners, semantic processing was delayed in listeners with poorer attentional skills. My results thus show that the amount of resources that listeners have available may modulate the rapid comprehension of a spoken message through the ease or speed of the spreading of activation between associated words. Future research could follow up on this finding of semantic processing by investigating older adults' speech comprehension in a more naturalistic listening situation, such as listening to audio books (with or without additional task demands). Work by Sommers and colleagues (Sommers et al., 2011) on older adults' comprehension of extended spoken passages showed that audibility alone (through amplification by hearing aids) may not be sufficient to maintain passage comprehension in older adulthood. Future research could follow up on this by experimentally manipulating the number of semantically related words (or the strength of the associations between word pairs) in the passages, and testing how hearing and attentional skills and task demands may modulate the comprehension of spoken passages.

This thesis particularly highlighted the importance of methodological choices for the study of affect processing. Affect perception in this thesis was investigated with non-scripted speech to investigate whether age and hearing loss impaired listeners' uptake of relatively subtle acoustic information to derive affective meaning from the spoken utterance. Chapters 3, 4 and 5 showed that both age and more severe degrees of hearing loss may indeed change the way listeners interpret prosodic-acoustic information. Both age and hearing loss may change the relative importance of specific acoustic cues to arousal or valence. The results thus provide evidence for both the auditory and central processing account of age differences in affect perception.

One of the major challenges for future research linking acoustic information to listeners' processing of affect will be the creation of sizeable corpora containing (semi-)spontaneous conversations. Only larger corpora will allow researchers to work with larger stimulus sets so that interference from semantic content may become less problematic. Future research should also include a more fine-grained and complete description of the emotional space by including additional dimension such as potency/control. Additionally, validity of the findings can be increased by having more participants rate the affective materials. Apart from the quantity of these corpora, the quality of item sets is also important as high-quality (and relatively noise-free) recordings are needed for detailed acoustic analysis.

Future research on age or hearing loss effects on processing of affective meaning should preferably also extend beyond affect rating or categorization tasks. Age-related effects on cognitive and hearing abilities may not just influence immediate recognition or evaluation of affective meaning, but may also be important for how well message content is remembered later on. Messages spoken in a positive or negative tone were less well remembered by older adults than messages spoken in a neutral tone, whereas this valence effect on message recall is absent for younger adults (Fairfield, Di Domenico, Serricchio, Borella, & Mammarella, 2017). Therefore, in order to get at the time course of affect processing for different groups of listeners, future research should preferably investigate immediate effects of acoustic variability in a spoken message on message evaluation, as well as later message recall.

Finally, the results on affect perception in hearing aid users are encouraging in that hearing aid use levels off the differences in rating behavior due to hearing sensitivity in the condition without hearing aids. Furthermore, hearing aid use make listeners even more responsive to intensity differences in the speech materials than the control group. The results on how hearing loss and hearing aid use make listeners more or less sensitive to specific acoustic cues should be validated in larger cohorts. Larger cohorts are also necessary to investigate different types of hearing loss, e.g., hearing losses particularly affecting the lower frequencies versus hearing losses in the higher frequencies typical for age-related hearing loss. Since the uptake of these acoustic cues may also contribute to signal quality (clarity, naturalness, etc.), which is strongly associated with hearing aid user's satisfaction (Kochkin, 2010), listening conditions with decreased signal clarity, e.g., noise, should be investigated in future research to get an idea of hearing aid users' performance in real-life situations.

In sum, the ability to adequately understand and respond to what is said and how it is said form an important part of social interaction. This thesis investigated older (and younger) adults' listening for the what and the how in speech. Concerning the what, I have shown that cognitive load affects semantic comprehension in speech processing in older adults, but only for those with poorer attentional abilities. Concerning the how, I have shown that age differences in the perception of affective information can partly, but not fully, be explained by age-related hearing loss. These results suggest that both age-related change in hearing and in higher-level processing of auditory input change the way listeners interpret prosodic cues to affect.

References

Agresti, A. (2002). *Categorical data analysis*. Hoboken, New Jersey: John Wiley & Sons.

Akeroyd, M. A. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *International Journal of Audiology*, *47*(Suppl 2), S53–S71. http://doi.org/10.1080/14992020802301142

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

Anderson, N. D., Craik, F. I. M., & Naveh-Benjamin, M. (1998). The attentional demands of encoding and retrieval in younger and older adults: 1. Evidence from divided attention costs. *Psychology and Aging*, *13*(3), 405–423.

Anderson, S., Parbery-Clark, A., White-Schwoch, T., & Kraus, N. (2012). Aging affects neural precision of speech encoding. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *32*(41), 14156–14164. http://doi.org/10.1523/JNEUROSCI.2176-12.2012

Aubergé, V., & Cathiard, M. (2003). Can we hear the prosody of smile? *Speech Communication*, *40*, 87–97.

Aydelott, J., & Bates, E. (2004). Effects of acoustic distortion and semantic context on lexical access. *Language and Cognitive Processes*, *19*(1), 29–56. http://doi.org/10.1080/01690960344000099

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). The CELEX lexical database. Philadelphia: University of Pennsylvania, Linguistic Data Consortium.

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, *70*(3), 614–636. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8851745

Bänziger, T., Patel, S., & Scherer, K. R. (2014). The role of perceived voice and speech characteristics in vocal emotion communication. *Journal of Nonverbal Behavior*, *38*(1), 31–52. http://doi.org/10.1007/s10919-013-0165-x

Bath, P. A., & Deeg, D. (2005). Social engagement and health outcomes among older people: Introduction to a special section. *European Journal of Ageing*, *2*(1), 24–30. http://doi.org/10.1007/s10433-005-0019-4

Benichov, J., Cox, L. C., Tun, P. A., & Wingfield, A. (2012). Word recognition within a linguistic context: Effects of age, hearing acuity, verbal ability and cognitive function. *Ear and Hearing*, *32*(2), 250–256. http://doi.org/10.1097/AUD.0b013e31822f680f.Word

Bilger, R. C., & Wang, M. D. (1976). Consonant confusions in patients with sensorineural hearing loss. *Journal of Speech, Language, and Hearing Research*, *19*, 718–748.

Blair, R. J. R. (2003). Facial expressions, their communicatory functions and neuro-cognitive substrates. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *358*(1431), 561–72. http://doi.org/10.1098/rstb.2002.1220

Boersma, P., & Weenink, D. (2011). Praat: Doing phonetics by computer. Retrieved from www.praat.org

Boersma, P., & Weenink, D. (2013). Praat: Doing phonetics by computer. Retrieved from www.praat.org

Boettcher, F. A., Poth, E. A., Mills, J. H., & Dubno, J. R. (2001). The amplitude-modulation following response in young and aged human subjects. *Hearing Research*, *153*, 32–42.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, *25*(1), 49–59. http://doi.org/10.1016/0005-7916(94)90063-9

Campbell, N. (2005). Getting to the heart of the matter: Speech as the expression of affect; Rather than just text or language. *Language Resources and Evaluation*, *39*(1), 109–118. http://doi.org/10.1007/s10579-005-2699-y

Cheang, H. S., & Pell, M. D. (2008). The sound of sarcasm. *Speech Communication*, *50*(5), 366–381. http://doi.org/10.1016/j.specom.2007.11.003

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407–428. http://doi.org/10.1037//0033-295X.82.6.407

Coutinho, E., & Dibben, N. (2013). Psychoacoustic cues to emotion in speech prosody and music. *Cognition & Emotion*, *27*(4), 658–684. http://doi.org/10.1080/02699931.2012.732559

Cowie, R., & Cornelius, R. R. (2003). Describing the emotional states that are expressed in speech. *Speech Communication*, *40*(1-2), 5–32. http://doi.org/10.1016/S0167-6393(02)00071-7

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

De Deyne, S., & Storms, G. (2008). Word associations: norms for 1,424 Dutch words in a continuous task. *Behavior Research Methods*, *40*(1), 198–205. http://doi.org/10.3758/BRM .40.1.198

Drijvers, L., Mulder, K., & Ernestus, M. (2016). Alpha and gamma band oscillations index differential processing of acoustically reduced and full forms. *Brain and Language*, *153-154*, 27–37. http://doi.org/10.1016/j.bandl.2016.01.003

Dupoux, E., Kouider, S., & Mehler, J. (2003). Lexical access without attention? Explorations using dichotic priming. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(1), 172–184. http://doi.org/10.1037/0096-1523.29.1.172

Dupuis, K., & Pichora-Fuller, M. K. (2015). Aging affects identification of vocal emotions in semantically neutral sentences. *Journal of Speech, Language, and Hearing Research*, *58*(June), 1061–1076.

Dupuis, K., Pichora-Fuller, M. K., Chasteen, A. L., Marchuk, V., Singh, G., & Smith, S. L. (2014). Effects of hearing and vision impairments on the Montreal Cognitive Assessment. *Aging, Neuropsychology, and Cognition: A Journal on Normal and Dysfunctional Development*, *22*(4), 413–437. http://doi.org/10.1080/13825585.2014.968084

Dykstra, P. A., Van Tilburg, T. G., & De Jong. (2005). Changes in older adult loneliness: Results from a seven-year longitudinal study. *Research on Aging*, *27*(6), 725–747. http://doi.org/10.1177/0164027505279712

Fairfield, B., Di Domenico, A., Serricchio, S., Borella, E., & Mammarella, N. (2017). Emotional prosody effects on verbal memory in older and younger adults. *Aging, Neuropsychology, and Cognition*, *24*(4), 408–417. http://doi.org/10.1080/13825585.2016.1219690

Freigang, C., Schmidt, L., Wagner, J., Eckardt, R., Steinhagen-Thiessen, E., Ernst, A., & Rübsamen, R. (2011). Evaluation of central auditory discrimination abilities in older adults. *Frontiers in Aging Neuroscience*, *3*(May), 6. http://doi.org/10.3389/fnagi.2011.00006

Gordon-Salant, S., & Fitzgibbons, P. J. (1999). Profile of auditory temporal processing in older listeners. *Journal of Speech, Language, and Hearing Research*, *42*(2), 300–311. http://doi.org/1092-4388/99/4202-0300

Goudbeek, M., & Scherer, K. (2010). Beyond arousal: Valence and potency/control cues in the vocal expression of emotion. The Journal of the Acoustical Society of America, 128(3), 1322. http://doi.org/10.1121/1.3466853

Gow, D. W., & Gordon, P. C. (1995). Lexical and prelexical influences on word segmentation: Evidence from priming. *Journal of Experimental Psychology. Human Perception and Performance*, *21*(2), 344–359. http://doi.org/10.1037/0096-1523.21.2.344

Grabowski, R., & Bauer, D. (2004). System for computer-aided perception experiments (SCAPE). Retrieved from http://www.coli.uni-sb.de/?doba/scape/

Grant, K. W. (1987). Identification of intonation contours by normally hearing and profoundly hearing-impaired listeners performance. *Journal of the Acoustical Society of America*, *82*(4), 1172–1178.

Grichkovtsova, I., Morel, M., & Lacheret, A. (2012). The role of voice quality and prosodic contour in affective speech perception. *Speech Communication*, *54*(3), 414–429. http://doi.org/10.1016/j.specom.2011.10.005

Grimm, M., Kroschel, K., & Narayanan, S. (2008). The Vera am Mittag German audio-visual emotional speech database. In *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME)* (pp. 865–868).

Hammarberg, B., Fritzell, B., Gauffin, J., Sundberg, J., & Wedin, L. (1980). Perceptual and acoustic correlates of abnormal voice qualities. *Acta Otolaryngologica*, *90*, 441–451.

Hammerschmidt, K., & Jürgens, U. (2007). Acoustical correlates of affective prosody. *Journal of Voice : Official Journal of the Voice Foundation*, *21*(5), 531–40. http://doi.org/10.1016/j.jvoice.2006.03.002

Harris, K. C., Mills, J. H., & Dubno, J. R. (2007). Electrophysiologic correlates of intensity discrimination in cortical evoked potentials of younger and older adults. *Hearing Research*, *228*(1-2), 58–68. http://doi.org/10.1016/j.heares.2007.01.021

He, N., Dubno, J. R., & Mills, J. H. (1998). Frequency and intensity discrimination measured in a maximum-likelihood procedure from young and aged normal-hearing subjects. *The Journal of the Acoustical Society of America*, *103*(1), 553–565. http://doi.org/10.1121/1.421127

He, N., Mills, J. H., & Dubno, J. R. (2007). Frequency modulation detection: Effects of age, psychophysical method, and modulation waveform. *The Journal of the Acoustical Society of America*, *122*(1), 467–477. http://doi.org/10.1121/1.2741208

Humes, L. E. (1996). Speech understanding in the elderly. Journal of the American Academy of Audiology, 7, 161–167.

Humes, L. E. (2007). The contributions of audibility and cognitive factors to the benefit provided by amplified speech to older adults. *Journal of the American Academy of Audiology*, *18*, 590–603. http://doi.org/10.3766/jaaa.18.7.6

International Organization for Standardization. (2000). *Acoustics—Statistical distribution of hearing thresholds as a function of age* (ISO Standard Number 7029).

Isaacowitz, D. M., Löckenhoff, C. E., Lane, R. D., Wright, R., Sechrest, L., Riedel, R., & Costa, P. T. (2007). Age differences in recognition of emotion in lexical stimuli and facial expressions. *Psychology and Aging*, *22*(1), 147–59. http://doi.org/10.1037/0882-7974.22.1.147

Janse, E. (2009). Processing of fast speech by elderly listeners. *The Journal of the Acoustical Society of America*, *125*(4), 2361–2373. http://doi.org/10.1121/1.3082117

Johnstone, T., & Scherer, K. R. (2000). Vocal communication of emotion. In M. Lewis & J. Haviland (Eds.), *The Handbook of Emotion* (Second Edi, pp. 220–235). New Yor: Guilford Press.

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: different channels, same code? *Psychological Bulletin*, *129*(5), 770–814. http://doi.org/10.1037/0033-2909.129.5.770

Kahneman, D. (1973). *Attention and effort*. *The American Journal of Psychology* (Vol. 88). Englewood Cliffs, NJ: Prentice Hall. http://doi.org/10.2307/1421603

Kiss, I., & Ennis, T. (2001). Age-related decline in perception of prosodic affect. *Applied Neuropsychology*, *8*(4), 251–254. http://doi.org/10.1207/S15324826AN0804_9

Kizach, J. (2014). Analyzing Likert-scale data with mixed-effects linear models: A simulation study. Poster Presented at Linguistic Evidence 2014,Tübingen.

Kochkin, S. (2010). MarkeTrak VIII : Consumer satisfaction. *The Hearing Journal*, *63*(1), 19–32.

Koehnke, J., Culotta, C. P., Hawley, M. L., & Colburn, H. S. (1995). Effects of reference interaural time and intensity differences on binaural performance in listeners with normal and impaired hearing. *Ear & Hearing*, *16*, 331–353.

Lambrecht, L., Kreifelts, B., & Wildgruber, D. (2012). Age-related decrease in recognition of emotional facial and prosodic expressions. *Emotion*, *12*(3), 529–539. http://doi.org/10.1037/a0026827

Lambrecht, L., Kreifelts, B., & Wildgruber, D. (2014). Gender differences in emotion recognition: impact of sensory modality and emotional category. Cognition & Emotion 28, 452–469. doi: 10.1080/02699931.2013. 837378

Lang, P. J. (1980). Behavioral treatment and bio-behavioral assessment: Computer applications. In J. B. Sidowski, J. H. Johnson, & T. A. Williams (Eds.), *Technology in Mental Health Care Delivery Systems* (pp. 119–137). Norwood, NJ.

Laukka, P., Neiberg, D., Forsell, M., Karlsson, I., & Elenius, K. (2011). Expression of affect in spontaneous speech: Acoustic correlates and automatic detection of irritation and resignation. *Computer Speech and Language*, *25*(1), 84–104. http://doi.org/10.1016/j.csl.2010.03.004

Lawrence, A. D., Calder, A. J., McGowan, S. W., & Grasby, P. M. (2002). Selective disruption of the recognition of facial expressions of anger. *Neuroreport*, *13*(6), 881–884. http://doi.org/10.1097/00001756-200205070-00029

Lemke, U., & Besser, J. (2016). Cognitive load and listening effort : Concepts and age-related considerations. *Ear & Hearing*, *37*(1), 77–84.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, *22*(140), 1–55.

Lima, C. F., Alves, T., Scott, S. K., & Castro, S. L. (2014). In the ear of the beholder: How age shapes emotion processing in nonverbal vocalizations. *Emotion (Washington, D.C.)*, *14*(1), 145–60. http://doi.org/10.1037/a0034287

Lima, C. F., Castro, S. L., & Scott, S. K. (2013). When voices get emotional: A corpus of nonverbal vocalizations for research on emotion processing. *Behavior Research Methods*, *45*(4), 1234–1245. http://doi.org/10.3758/s13428-013-0324-3

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, *19*(1), 1–36. http://doi.org/10.1097/MPG.0b013e3181a15ae8.Screening

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, *25*(1-2), 71–102. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18480284

Mattys, S. L., & Palmer, S. D. (2015). Divided attention disrupts perceptual encoding during speech recognition. *The Journal of the Acoustical Society of America*, *137*(3), 1464–1472. http://doi.org/10.1121/1.4913507

Mattys, S. L., & Wiget, L. (2011). Effects of cognitive load on speech recognition. *Journal of Memory and Language*, *65*(2), 145–160. http://doi.org/10.1016/j.jml.2011.04.004

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*(1), 1–86. http://doi.org/10.1016/0010-0285(86)90015-0

McCoy, S. L., Tun, P. A., Cox, L. C., Colangelo, M., Stewart, R. A., & Wingfield, A. (2005). Hearing loss and perceptual effort: Downstream effects on older adults' memory for speech. *The Quarterly Journal of Experimental Psychology: Section A*, *58*(1), 22–33. http://doi.org/10.1080/02724980443000151

McGettigan, C., Walsh, E., Jessop, R., Agnew, Z. K., Sauter, D. A., Warren, J. E., Scott, S. K. (2015). Individual differences in laughter perception reveal roles for mentalizing and sensorimotor systems in the evaluation of emotional authenticity. *Cerebral Cortex 25*, 246–257. doi: 10.1093/cercor/ bht227

McKoon, G., & Ratcliff, R. (2012). Aging and IQ effects on associative recognition and priming in item recognition. *Journal of Memory and Language*, *66*(3), 416–437. http://doi.org/10.1016/j.jml.2011.12.001

Mitchell, R. L. C. (2007). Age-related decline in the ability to decode emotional prosody: Primary or secondary phenomenon? *Cognition & Emotion*, *21*(7), 1435–1454. http://doi.org/10.1080/02699930601133994

Mitchell, R. L. C., & Kingston, R. A. (2014). Age-related decline in emotional prosody discrimination: Acoustic correlates. *Experimental Psychology*, *61*(3), 215–223. http://doi.org/10.1027/1618-3169/a000241

Mitchell, R. L. C., Kingston, R. A., & Barbosa Bouças, S. L. (2011). The specificity of age-related decline in interpretation of emotion cues from prosody. *Psychology and Aging*, *26*(2), 406–414. http://doi.org/10.1037/a0021861

Most, T., & Michaelis, H. (2012). Auditory, visual, and auditory–visual perceptions of emotions by young children with hearing loss versus children with normal hearing. *Journal of Speech, Language, and Hearing Research*, *55*(4), 1148–1162. http://doi.org/10.1044/1092-4388(2011/11-0060)1148

Most, T., Weisel, A., & Zaychik, A. (1993). Auditory, visual and auditory-visual identification of emotions by hearing and hearing-impaired adolescents. *British Journal of Audiology*, *27*, 247–253.

Mozziconacci, S. J. L. (1998). *Speech variability and emotion: Production and perception*. Technische Universiteit Eindhoven, The Netherlands.

Mozziconacci, S. J. L., & Hermes, D. J. (2000). Expression of emotion and attitude through temporal speech variations. In *Sixth International Conference on Spoken Language Processing* (pp. 373–378). Beijing, China.

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, *53*(4), 695–699. http://doi.org/10.1111/j.1532-5415.2005.53221.x

Naveh-Benjamin, M., Craik, F. I. M., Guez, J., & Kreuger, S. (2005). Divided attention in younger and older adults: Effects of strategy and relatedness on memory performance and secondary task costs. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *31*(3), 520–537. http://doi.org/10.1037/0278-7393.31.3.520

Norman, G. (2010). Likert scales, levels of measurement and the “laws” of statistics. *Advances in Health Sciences Education*, *15*(5), 625–632. http://doi.org/10.1007/s10459-010-9222-y

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

Norris, D., Cutler, A., McQueen, J. M., & Butterfield, S. (2006). Phonological and conceptual activation in speech comprehension. *Cognitive Psychology*, *53*(2), 146–193. http://doi.org/10.1016/j.cogpsych.2006.03.001

Obleser, J., & Kotz, S. A. (2011). Multiple brain signatures of integration in the comprehension of degraded speech. *NeuroImage*, *55*(2), 713–723. http://doi.org/10.1016/j.neuroimage.2010.12.020

Orbelo, D. M., Grim, M. A., Talbott, R. E., & Ross, E. D. (2005). Impaired comprehension of affective prosody in elderly subjects is not predicted by age-related hearing loss or age-related cognitive decline. *Journal of Geriatric Psychiatry and Neurology*, *18*(1), 25–32. http://doi.org/10.1177/0891988704272214

Otsuka, S., & Kawaguchi, J. (2007). Divided attention modulates semantic activation: Evidence from a nonletter-level prime task. *Memory & Cognition*, *35*(8), 2001–2011. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18265615

Pals, C., Palsanastasios, A., Van Rijn, H., & Başkent, D. (2015). Validation of a simple response-time measure of listening effort. *Journal of the Acoustical Society of America*, *138*(3). http://doi.org/10.1121/1.4929614

Pals, C., Sarampalis, A., & Başkent, D. (2013). Listening effort with cochlear implant simulations. *Journal of Speech, Language, and Hearing Research*, *56*(August), 1075–1084.

Paulmann, S., Pell, M. D., & Kotz, S. A. (2008). How aging affects the recognition of emotional speech. *Brain and Language*, *104*(3), 262–269. http://doi.org/10.1016/j.bandl.2007.03.002

Pell, M. D., Jaywant, A., Monetta, L., & Kotz, S. a. (2011). Emotional speech processing: Disentangling the effects of prosody and semantic cues. *Cognition & Emotion*, *25*(5), 834–853. http://doi.org/10.1080/02699931.2010.516915

Pereira, C. (2000). Dimensions of emotional meaning in speech. In *Proceedings of the ISCA Workshop on Speech and Emotion* (pp. 25–28).

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W. Y., Humes, L. E., Lemke, U., Lunner, T., Matthen, M., Mackersie, C. L., Naylor, G., Phillips, N. A., Richter, M., Rudner, M., Sommers, M., Tremblay, K. L., Wingfield, A. (2016). Hearing impairment and cognitive energy : The Framework for Understanding Effortful Listening (FUEL). *Ear & Hearing*, *37*(Supplement 1), 5–27.

Pöppel, E. (2004). Lost in time: A historical frame, elementary processing units and the 3-second window. *Acta Neurobiologiae Experimentalis*, *64*(3), 295–301. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15283473

R Development Core Team. (2008). R: A language and environment for statistical computing. [Computer Software]. Vienna, Austria: R Foundation for Statistical Computing.

Rabbitt, P. (1968). Repetition effects and signal classification strategies in serial choice-response tasks. *Quarterly Journal of Experimental Psychology*, *20*(3), 232–240. http://doi.org/10.1080/14640746808400157

Rabbitt, P. (1991). Mild hearing loss can cause apparent memory failures. *Acta Otolaryngologica*, *111*(s476), 167–176.

Reitan, R. M. (1958). Validity of the Trail Making test as an indicator of organic brain damage. *Perceptual and Motor Skills*, *8*, 271–276.

Rodero, E. (2011). Intonation and emotion: Influence of pitch levels and contour type on creating emotions. *Journal of Voice : Official Journal of the Voice Foundation*, *25*(1), e25–34. http://doi.org/10.1016/j.jvoice.2010.02.002

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., Dahlström, O., Signoret, C., Stenfelt, S., Pichora-Fuller, M. K., Rudner, M. (2013). The Ease of Language Understanding (ELU) model: Theoretical, empirical, and clinical advances. *Frontiers in Systems Neuroscience*, *7*(July), 31. http://doi.org/10.3389/fnsys.2013.00031

Ruffman, T., Henry, J. D., Livingstone, V., & Phillips, L. H. (2008). A meta-analytic review of emotion recognition and aging: Implications for neuropsychological models of aging. *Neuroscience and Biobehavioral Reviews*, *32*(4), 863–881. http://doi.org/10.1016/j.neubiorev.2008.01.001

Sauter, D. A., Eisner, F., Calder, A. J., & Scott, S. K. (2010). Perceptual cues in nonverbal vocal expressions of emotion. *The Quarterly Journal of Experimental Psychology*, *63*(11), 2251–2272. http://doi.org/10.1080/17470211003721642

Scherer, K. R. (1986). Vocal affect expression: A review and a model for future research. *Psychological Bulletin*, *99*(2), 143–165. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3515381

Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication*, *40*(1-2), 227–256. http://doi.org/10.1016/S0167-6393(02)00084-5

Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information*, *44*(4), 695–729. http://doi.org/10.1177/0539018405058216

Scherer, K. R., Banse, R., & Wallbott, H. G. (2001). Emotion inferences from vocal expression correlate across languages and cultures. *Journal of Cross-Cultural Psychology*, *32*(1), 76–92. http://doi.org/10.1177/0022022101032001009

Schmidt, J., Herzog, D., Scharenborg, O., & Janse, E. (2015). Do hearing aids improve affect perception? In van Dijk, P., Başkent, D., Gaudrain, E., de Kleine, E., Wagner, A., and Lanting, C. (Eds.), *Physiology, Psychoacoustics and Cognition in Normal and Impaired Hearing*. Advances in Experimental Medicine and Biology, 894 (pp. 47–55). Springer International Publishing, Cham. http://doi.org/10.1007/978-3-319-25474-6

Schmidt, J., Janse, E., & Scharenborg, O. (2014). Age, hearing loss and the perception of affective utterances in conversational speech. In *Proceedings of the 15th Annual Conference of the International Speech Communication Association* (pp. 1929–1933).

Schmidt, J., Janse, E., & Scharenborg, O. (2016). Perception of emotion in conversational speech by younger and older listeners. *Frontiers in Psychology*, *7*, 781. http://doi.org/10.3389/fpsyg.2016.00781

Schneider, B. A., Daneman, M., & Pichora-Fuller, M. K. (2002). Listening in aging adults: From discourse comprehension to psychoacoustics. *Canadian Journal of Experimental Psychology*, *56*(3), 139–152.

Schröder, M. (2006). Expressing degree of activation in synthetic speech. *IEEE Transactions on Audio, Speech, and Language Processing*, *14*(4), 1128–1136.

Schröder, M., Cowie, R., Douglas-Cowie, E., Westerdijk, M., & Gielen, S. (2001). Acoustic correlates of emotion dimensions in view of speech synthesis. In *Proceedings of the 7th European Conference on Speech Communication and Technology (EUROSPEECH '01)* (pp. 87–90).

Shanks, J. E., Wilson, R. H., Larson, V., & Williams, D. (2002). Speech recognition performance of patients with sensorineural hearing loss under unaided and aided conditions using linear and compression hearing aids. *Ear & Hearing*, *23*, 280–290. http://doi.org/10.1097/01.AUD.0000027401.05012.36

Sluijter, A. M. C., & Heuven, V. J. Van. (1996). Spectral balance as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America*, *100*(4), 2471–2485.

Smith, M. C., Bentin, S., & Spalek, T. M. (2001). Attention constraints of semantic activation during visual word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(5), 1289–1298. http://doi.org/10.1037//0278-7393.27.5.1289

Sommers, M. S., Hale, S., Myerson, J., Rose, N., Tye-Murray, N., & Spehar, B. (2011). Listening comprehension across the adult lifespan. *Ear & Hearing*, *32*(6), 775–781.

Souza, P., Arehart, K., Miller, C. W., & Muralimanohar, R. K. (2011). Effects of age on F0 discrimination and intonation perception in simulated electric and electroacoustic hearing. *Ear and Hearing*, *32*(1), 75–83. http://doi.org/10.1097/AUD.0b013e3181eccfe9

Swinney, D. A., Onifer, W., Prather, P., & Hirshkowitz, M. (1979). Semantic facilitation across sensory modalities in the processing of individual words and sentences. *Memory & Cognition*, *7*(3), 159–165. http://doi.org/10.3758/BF03197534

Sylvain-Roy, S., & Belleville, S. (2015). Interindividual differences in attentional control profiles among younger and older adults. *Aging, Neuropsychology, and Cognition: A Journal on Normal and Dysfunctional Development*, *22*(3), 259–279. http://doi.org/10.1080/13825585.2014.926305

Tabossi, P. (1996). Cross-modal semantic priming. *Language and Cognitive Processes*, *11*(6). http://doi.org/10.1080/016909696386953

Tamarit, L., Goudbeek, M., & Scherer, K. (2008). Spectral slope measurements in emotionally expressive speech. In *Proceedings for ISCA ITRW Speech Analysis and Processing for Knowledge Discovery*. Aalborg.

Titone, D., Prentice, K. J., & Wingfield, A. (2000). Resource allocation during spoken discourse processing: Effects of age and passage difficulty as revealed by self-paced listening. *Memory & Cognition*, *28*(6), 1029–1040. http://doi.org/10.3758/BF03209351

Tun, P. A., McCoy, S., & Wingfield, A. (2009). Aging, hearing acuity, and the attentional costs of effortful listening. *Psychology and Aging*, *24*(3), 761–6. http://doi.org/10.1037/a0014802

Van de Ven, M., Tucker, B. V, & Ernestus, M. (2011). Semantic context effects in the comprehension of reduced pronunciation variants. *Memory & Cognition*, *39*(7), 1301–16. http://doi.org/10.3758/s13421-011-0103-2

Van Rooij, J. C. G. M., & Plomp, R. (1992). Auditive and cognitive factors in speech perception by elderly listeners.II: Multivariate analyses. *Journal of the Acoustical Society of America*, *88*(6), 2611–2624.

Waaramaa, T., & Leisiö, T. (2013). Perception of emotionally loaded vocal expressions and its connection to responses to music. A cross-cultural investigation: Estonia, Finland, Sweden, Russia, and the USA. *Frontiers in Psychology*, *4*(June), Article 344. http://doi.org/10.3389/fpsyg.2013.00344

Waldron-Perrine, B., & Axelrod, B. N. (2012). Determining an appropriate cutting score for indication of impairment on the Montreal Cognitive Assessment. *International Journal of Geriatric Psychiatry*, *27*(11), 1189–1194. http://doi.org/10.1002/gps.3768

Wechsler, D. (2008). Wechsler adult intelligence scale–Fourth edition. TX: Pearson.

Wilting, J., Krahmer, E., & Swerts, M. (2006). Real vs. acted emotional speech. In *Proceedings of the 9th International Conference on Spoken Language Processing (Interspeech)* (pp. 805–808).

Winn, M. B., Chatterjee, M., & Idsardi, W. J. (2013). The roles of voice onset time and F0 in stop consonant voicing perception: Effects of masking noise and low-pass filtering. *Journal of Speech, Language, and Hearing Research*, *56*(4), 1097–1107. http://doi.org/10.1044/1092-4388(2012/12-0086).The

Wundt, W. (1905). *Grundzüge der physiologischen Psychologie*. Leipzig: Engelmann.

Appendix

Table A. Older adults' individual thresholds (dB HL air conduction for the better ear) at different octave frequencies in Chapter 2. The pure-tone average used as a hearing measure in this chapter is the average across 1, 2, and 4 kHz.

Participant	0.5 kHz	1 kHz	2 kHz	4 kHz
101	5	20	20	30
102	25	20	30	55
103	25	20	20	30
104	10	5	15	15
105	15	25	25	50
106	10	20	30	55
108	30	25	30	60
109	15	10	15	15
110	25	25	20	45
111	15	20	25	25
112	15	10	25	30
113	10	5	30	60
115	10	0	0	35
116	0	5	10	10
117	15	10	5	0
118	20	20	35	40
119	20	10	10	45
121	10	10	20	25
122	20	15	25	25
123	10	5	15	45
124	10	5	0	10
126	25	-5	0	10
127	10	5	5	10
129	25	40	40	35
130	35	45	25	60

Participant	0.5 kHz	1 kHz	2 kHz	4 kHz
131	20	20	35	55
132	20	15	20	45
133	20	15	0	15
134	10	10	5	40
135	5	10	25	30
136	10	0	0	20
137	20	20	35	15
138	25	30	30	45
139	15	5	10	30
141	10	15	10	35
144	20	25	10	25
146	20	10	40	55

Table B. Older adults' individual thresholds (dB HL air conduction for the better ear) at different octave frequencies in Chapter 3. The pure-tone average used as a hearing measure in this chapter is the average across 1, 2, and 4 kHz.

Participant	0.5 kHz	1 kHz	2 kHz	4 kHz
211	5	5	0	15
212	15	10	20	65
213	30	10	15	5
214	15	10	15	30
215	15	30	35	20
221	10	5	20	20
222	15	5	5	10
223	20	10	20	40
224	10	10	0	15
225	15	5	15	25

Table C. Older adults' individual thresholds (dB HL air conduction for the better ear) at different octave frequencies in Chapter 4. The pure-tone average used as a hearing measure in this chapter is the average across 0.5, 1, 2, and 4 kHz.

Participant	0.5 kHz	1 kHz	2 kHz	4 kHz
101	20	20	20	20
102	10	20	10	30
103	20	-10	15	-10
104	25	45	55	60
107	15	10	5	50
108	25	30	45	50
109	30	40	45	40
110	30	35	50	55
111	-10	5	0	45
112	20	15	10	40
113	15	10	20	35
115	15	10	35	65
116	20	20	15	5
118	20	25	35	50
119	10	5	10	10
120	20	20	35	55
122	25	25	35	30

Table D. Older adults' individual thresholds (dB HL air conduction for the better ear) at different octave frequencies of hearing-impaired in Chapter 5. The pure-tone average used as a hearing measure in this chapter is the average across 0.5, 1, 2, and 4 kHz.

Participants	0.5 kHz	1 kHz	2 kHz	4 kHz
101	15	35	80	75
102	35	45	55	60
103	35	55	40	55
104	30	35	50	45
105	20	30	65	80
106	45	55	60	80
107	25	40	55	65
108	40	55	55	50
110	10	15	50	60
111	45	50	65	85
112	15	30	50	75
113	30	45	55	60
114	50	55	60	50
115	40	50	80	80
116	30	45	65	60
117	25	40	50	75
118	50	60	60	70
119	25	50	60	55
120	65	65	75	70
121	50	45	65	55
122	20	35	60	70
123	35	35	55	65
124	5	20	50	55

Table E. Older adults' individual thresholds (dB HL air conduction for the better ear) at different octave frequencies of normal-hearing participants in Chapter 5. The pure-tone average used as a hearing measure in this chapter is the average across 0.5, 1, 2, and 4 kHz.

Participants	0.5 kHz	1 kHz	2 kHz	4 kHz
1	5	0	10	35
2	5	10	15	40
3	5	10	15	45
4	15	15	15	25
5	5	5	15	10
6	15	5	20	40
8	0	5	5	20
10	10	25	20	25
11	0	0	10	10
12	5	10	15	20
14	5	15	25	30
15	5	15	5	10
17	15	15	10	20
18	25	15	30	20
19	10	5	0	10
20	0	10	10	10
21	10	5	5	10
22	10	10	10	45
23	15	10	15	15
24	10	0	30	55
25	10	10	0	15
26	5	0	5	40

Nederlandse samenvatting

Stelt u zich eens voor dat u een telefoontje krijgt van een vriend en dat hij u als volgt begroet: “Fantastisch! Deze dag had niet beter kunnen beginnen!”. Of dit gezegd wordt in een context waarin uw vriend net te weten is gekomen dat hij de lotterij heeft gewonnen of in een context waarin hij zich verslapen heeft, te laat komt voor de les en daarnaast ook nog eens ziet dat hij een lekke band heeft, maakt veel uit voor de betekenis van deze zin, ook al worden precies dezelfde woorden gebruikt. Voordat u hem vraagt “Wat is er aan de hand?”, heeft u hoogstwaarschijnlijk al een inschatting gemaakt of de dag goed of slecht is begonnen voor uw vriend en of hij dolblij is of de wanhoop nabij. Dit laat zien dat als we communiceren het niet alleen belangrijk is WAT we zeggen, maar ook HOE we het zeggen. Het vermogen om de semantische informatie (WAT we zeggen) en de affectieve informatie (HOE we het zeggen) goed genoeg te begrijpen en daar adequaat op te reageren is een belangrijk onderdeel van sociale interactie. Stelt u zich verder voor dat uw vriend niet u heeft gebeld, maar zijn grootmoeder. Hoe zou het gesprek dan verlopen in vergelijking met het gesprek met een jonger iemand? Deze vraag is gerelateerd aan het onderzoek dat in dit proefschrift beschreven is. Dit proefschrift onderzoekt hoe leeftijdsgerelateerde effecten en individuele gehoorverschillen een rol spelen bij de opname en verwerking van semantische en affectieve informatie tijdens het begrijpen van spraak door oudere volwassenen.

Semantische informatie – Luisteren naar WAT er wordt gezegd

Een van de redenen waarom het begrijpen van spraak zo efficiënt is, is de manier waarop linguistische kennis wordt opgeslagen in en opgevraagd uit de hersenen; dit gebeurt in het mentale lexicon. Het mentale lexicon bevat lexicale lemma's, oftewel woorden, en aanvullende informatie over deze woorden, bijvoorbeeld over de uitspraak en betekenis ervan (semantische informatie dus). Bij het begrijpen van spraak is de verwerking van een (doel)woord dat voorafgegaan wordt door een semantisch gerelateerd (geprimed) woord makkelijker dan als dit woord voorafgegaan wordt door een semantisch ongerelateerd woord. Dit heet priming. Als uw vriend bijvoorbeeld tegen zijn grootmoeder zegt "Mijn auto heeft een lekke band", kan het woord "band" sneller opgevraagd worden uit het mentale lexicon omdat het voorafgaande woord "auto" er semantisch aan gerelateerd. Verder suggereert onderzoek dat aandacht een belangrijke factor is voor het verwerken van semantische informatie. Luisteraars willen of kunnen wellicht niet altijd hun volle aandacht besteden aan het verwerken van de gesproken input of ze hebben wellicht beperkte cognitieve middelen om dit te doen. De studie die in **Hoofdstuk 2** werd gepresenteerd, onderzocht of semantische activatie belemmerd wordt door een gelijktijdige secundaire taak met verschillende complexiteit waarbij het verbale werkgeheugen belast wordt. Is de semantische activatie bijvoorbeeld langzamer als grootmoeder, terwijl ze naar het verhaal over de lekke band luistert, een boodschappenlijstje uit het hoofd moet leren met daarop twee artikelen (lagere cognitieve belasting) versus een boodschappenlijstje met zeven artikelen (hogere cognitieve belasting)? Daarnaast onderzocht deze studie of individuele auditieve en cognitieve vermogens het semantische primingeffect en het effect van cognitieve belasting op semantische priming beïnvloeden. Wat gebeurt er als de grootmoeder van uw vriend een minder goed gehoor heeft en minder goed haar aandacht kan focussen? De resultaten laten slechts een marginaal significante vermindering zien in semantische priming toen de cognitieve belasting werd vergroot in vergelijking met een situatie met minder cognitieve belasting. De analyses toonden ook aan dat semantische priming significant verminderde in geval van een grotere cognitieve belasting in deelnemers die hun aandacht minder goed konden switchen. Daardoor kan een secundaire taak die veel resources vraagt de integratie van gesproken woorden in een coherente semantische representatie negatief beïnvloeden in luisteraars met minder goede aandachtvaardigheden. We hebben echter geen effect gevonden van gehoorverlies op semantische activatie en er was ook geen interactie met cognitieve belasting.

Affectieve informatie – Luisteren naar HOE iets wordt gezegd

Leeftijd beïnvloedt de perceptie van affectieve informatie in spraak. Dit resulteert in een groter aantal verkeerde classificaties van emotiecategorieën. Onderzoek heeft laten zien dat jongere deelnemers bijvoorbeeld significant beter zijn in het herkennen van boosheid, afkeer, angst, blijdschap en verdriet aan de hand van prosodische informatie in het spraaksignaal dan oudere volwassenen (Paulmann et al., 2008). In plaats van een beschrijving van affectieve informatie door middel van discrete categorieën als blij en verdrietig, kan een gedetailleerdere beschrijving van affectieve informatie verkregen worden door emotiedimensies te gebruiken. Dit betekent dat affectieve informatie geclassificeerd kan worden op een continuüm tussen positief en negatief (de valentiedimensie) en een continuüm tussen kalm en opgewonden (de arousaldimensie). De loterij winnen geeft hoogstwaarschijnlijk een gevoel van opgewonden blijdschap, een zeer positieve emotie met een hoge arousal, terwijl het hebben van een slechte dag een gevoel van wanhoop kan opwekken, een negatieve emotie met een gemiddelde tot lage arousal.

Affectieve informatie in spraak wordt uitgedrukt door verandering in verschillende akoestische, prosodische parameters. Tot deze parameters behoren o.a. toonhoogte, intensiteit, spectrale componenten en tempo, parameters die correleren met emotiedimensies. **Hoofdstuk 3** en **4** hebben betrekking op de vraag of het geobserveerde verschil in de perceptie van affectieve informatie in natuurlijke spraak tussen jongere en oudere volwassenen verklaart kan woorden door een verschil in de perceptie van aan affectieve informatie gerelateerde akoestische parameters tussen de twee leeftijdsgroepen. De tweede onderzoeksvraag is of een verschil in gehoorverlies onder oudere volwassen t hun rating van affectieve informatie beïnvloedt, en zo ja hoe. Is in grootmoeders perceptie de stem van uw vriend net zo aroused en positief/negatief als in uw perceptie? En is haar perceptie gerelateerd aan haar mate van gehoorverlies? **Hoofdstuk 3** laat zien dat het gebruik van de akoestische parameter intensiteit vergelijkbaar is in beide leeftijdsgroepen: Een luidere stem werd over het algemeen geassocieerd met een hogere arousal. Desalniettemin waren er verschillen tussen de jongere en oudere deelnemers in hun arousalratings: vergeleken met jongere deelnemers classificeerden oudere deelnemers de uitspraken als minder aroused en lieten ze een kleiner effect zien van intensiteit op hun arousalratings. Dit betekent dat zowel grootmoeder als u de opgewondenheid hoorden in de stem van uw vriend in het geval van de lotterij. Echter, in vergelijking met uw indruk, interpreteerde grootmoeder zijn stem als minder aroused and baseerde zij haar interpretatie niet zozeer op hoe luid hij sprak.

Wat valentie betreft, kwam toonhoogte naar voren als de meest prominente akoestische parameter: een hogere toonhoogte werd geassocieerd met een negatievere rating in beide leeftijdsgroepen. Dus zowel grootmoeder als u herkenden dat de klagende stem behoorde tot een negatievere emotie in het geval van de lekke band. In deze studie was het verschil in de rating van affectieve uitspraken tussen de leeftijdsgroepen niet gerelateerd aan het verschil in gehoorverlies tussen de beide groepen; wellicht omdat het gehoor van de oudere volwassenen nog te goed was. Onderzoek toont aan dat gehoorverlies en leeftijd het begrijpen van spraak alleen negatief beïnvloeden als een kritieke drempel is bereikt (McCoy et al., 2005). Daarom werden in **Hoofdstuk 4** alleen oudere deelnemers geselecteerd met een iets slechter gehoor. Dit onderzoek liet zien dat voor beide leeftijdsgroepen de gemiddelde intensiteit de voornaamste aanwijzing was voor arousal (waarbij een hogere gemiddelde intensiteit een indicatie was voor een hogere mate van arousal), terwijl de gemiddelde toonhoogte de voornaamste aanwijzing was voor valentie (waarbij een hogere gemiddelde toonhoogte geïnterpreteerd werd als negatiever). Net als voorheen, reageerden oudere volwassenen minder sterk op het verschil in gemiddelde intensiteit als aanwijzing voor arousal en had gehoorgevoeligheid geen invloed op het gebruik van gemiddelde intensiteit als aanwijzing voor arousal. Verder liet dit onderzoek zien dat oudere volwassenen sterker reageerden op het verschil in toonhoogte als aanwijzing voor valentie dan jongere volwassenen. Daarnaast beïnvloedde individuele gehoorgevoeligheid over het algemeen de valentieratings en veranderde dit het gebruik van de gemiddelde toonhoogte. Dus met een grotere mate van gehoorverlies is het mogelijk dat grootmoeder de stem van uw vriend niet zo positief of negatief ziet als u.

Maar wat gebeurt er als grootvader, die meestal gehoorapparaten draagt, de telefoon opneemt? Hoe zou hij de mate van arousal in de stem van zijn kleinzoon interpreteren als hij zijn gehoorapparaten wel en niet in heeft? En in hoeverre komt dit overeen met de interpretatie van arousal door een leeftijdsgenoot met een gehoor dat normaal is voor zijn/haar leeftijd, namelijk zoals grootmoeders perceptie van de mate van arousal? Deze vragen worden behandeld in **Hoofdstuk 5**. Dit is belangrijk om te weten omdat als de bijdrage van gehoorverlies aan de perceptie van affectieve informatie niet geheel duidelijk is, dan is het ook onduidelijk of het gebruik van gehoorapparaten de informatie die nodig is voor de perceptie van affectieve informatie in spraak voldoende kan herstellen. Daarnaast wordt een andere mogelijke oorzaak van leeftijdsverschil in de perceptie van emotie onderzocht, namelijk auditieve verwerking. Naast gehoorverlies dat voortkomt uit het slechter functioneren van het buiten-, midden- of binnenoor, kan een verslechterde auditieve verwerking in de hersenen de herkenning van emoties ook beïnvloeden. Gedurende de levensloop ondergaat het "sociale brein" structurele en functionele veranderingen, wat onder andere bestaat uit een vermindering in volume en activatie en veranderingen in het aantal neurotransmitters. Oudere volwassenen, zelfs als ze normaal horen, zijn minder gevoelig voor subtiele verschillen in akoestische aanwijzingen, bijvoorbeeld verschillen in intensiteit (Harris, Mills, & Dubno, 2007). Daarom onderzoekt **Hoofdstuk 5** ook de mogelijke relatie tussen gehoorverlies en het vermogen om verschil in luidheid te detecteren aan de ene kant en arousalrating aan de andere. Net als bij de voorgaande studies liet het onderzoek in dit hoofdstuk zien dat intensiteit de voornaamste aanwijzing voor de perceptie van arousal is voor beide groepen deelnemers (mensen met een gehoorapparaat vs. leeftijdsgenoten met een gehoor dat normaal is voor hun leeftijd) en voor beide luisteromstandigheden (met en zonder een gehoorapparaat). Zowel grootvader als grootmoeder zouden afgaan op de luidheid van de stem van hun kleinzoon voor het analyseren van zijn mate van arousal. Het maakt daarbij niet uit of grootvader zijn gehoorapparaten in heeft of niet. Het dragen van gehoorapparaten neutraliseerde het algemene effect van gehoorverlies op de perceptie van arousal: In vergelijking met grootmoeder, die normaal hoort voor haar leeftijd, liet grootvader met gehoorapparaten over het algemeen hetzelfde patroon zien in de rating van affectieve informatie en deed hij het zelfs beter dan grootmoeder wat betreft het gebruik van intensiteit als aanwijzing voor arousal. Aangezien een verslechterde perceptie van arousal consequenties kan hebben voor de perceptie van kleine verschillen in affectieve informatie (is de kleinzoon bijvoorbeeld enkel blij of opgewonden, is hij een klein beetje van streek of voelt hij zich echt ellendig), onderstrepen de huidige resultaten het belang van gehoorapparaten voor de rehabilitatie van de perceptie van affectieve informatie. Het individuele vermogen om luidheid te verwerken had echter geen invloed op de perceptie van arousal naast gehoorverlies.

Samenvattend, het begrijpen van spraak, wat betrekking heeft op WAT er wordt gezegd en HOE het wordt gezegd, is een aanzienlijk deel van communicatie. Dit proefschrift onderzocht hoe leeftijd, cognitieve vermogens en gehoorverlies de verwerking van in akoestisch-prosodische informatie versleutelde semantische en affectieve informatie kan beïnvloeden. De resultaten tonen aan dat cognitieve belasting in oudere volwassenen de spraakverwerking van WAT er gezegd wordt beïnvloedt, maar alleen in ouderen met een minder goed aandachtvermogen. Wat betreft HOE iets gezegd wordt, laten de resultaten zien dat het verschil in perceptie van affectieve informatie deels maar niet geheel verklaard kan worden door leeftijdsgerelateerd gehoorverlies.

Acknowledgements

One of the best things during my PhD was the travelling part. I got the chance to present my research in so many different countries, meeting and discussing it with many different people. I would like to seize the chance at the end of my journey in academia to thank all my fellow travelers. Without you climbing hills and walking through valleys wouldn't have been possible!

The academic journey

Let me start with my travel guides: my supervisors. Dear Odette. Thank you for choosing me for this project. For me, this endeavor has been the starting point of so many positive things, friendships, and eventually a career. Thank you for giving me that opportunity and believing in my research idea. You have always been ready to help me when I got stuck, provided feedback on my writing and had always the right punch line up your sleeve. Apart from the professional part you tried to make this journey a fun one. I will remember our discussions at INSPIRE dinners and conferences and the evenings out with your research group.

Dear Esther. You have been a great support throughout my PhD. Especially in the very critical phases I have been so grateful to have you by my side. You always had an open door and an open ear for any problem – not only for me but for all your PhDs. Doing (and finishing) a PhD can be quite nerve wrecking but with your calm and thoughtful personality you always got me back on track and saved me from more than one panic attack. Apart from your guidance in academia I will always remember your guidance in Dutch culture: the queen Juliana mug, the Sinterklaas dinners and the Dutch version of “something old, something new, something borrowed, something blue”... just to name a few things.

I owe special thanks to my promotor Mirjam Ernestus. Thank you for caring, for your enthusiasm, for reading all my chapters and papers and for providing critical and constructive criticism. Despite your busy schedule you always managed to make time for questions and discussions and I valued our monthly meetings with Odette and Esther.

My deepest gratitude goes to the reading committee Toni Rietveld, Deniz Başkent, Björn Schuller, Khiet Truong and Rob van Son who made time to read my thesis and provided valuable feedback.

I would also like send out a big thank you to all the members of the INSPIRE ITN. Being part of a such a big network from day one of my PhD was a great support, both academically and emotionally. Special thanks go to the seniors, especially to Sven Mattys, Henk van den Heuvel, Bert Cranen and Torsten Dau, for support, inspiring discussions and valuable and constructive criticism. Dear INSPIRE siblings: Alex, Cezara, Deepak, Dorina, Gustáv, Gustav, Huarda, Johannes, Kurt, Letizia, Lou, Mate, Nemanja, Petra, Ricard, Sara and Tom. I feel privileged to have been part of this group of fellows. Miss you all!

Within the INSPIRE network I got the chance to do my industry secondment at Phonak in Zuerich, Switzerland. I am greatly indebted to Peter Derleth for making my visit to Phonak possible. This has been a great experience and ensured me that I want to work in industry in the future. Diana, as my daily supervisor at Phonak you did an amazing job! Most of what I know about audiology and hearing aids I learned from you and you took really good care of me.

I would like to thank everybody who participated in my studies. I tested younger and older adults in three different countries: the Netherlands, Germany and Switzerland. Dear participants, thanks for your enthusiasm, your motivation and your patience. I enjoyed the nice chats, sharing your listening experiences, for the “d Spys” recipe and the “Basler Leckerli”. You made testing an exciting and enjoyable part of my PhD life. I am also very thankful for everybody who gave me a hand in testing.

Willemijn, I am particularly grateful for your support and guidance in testing. A piece of cake with you by my side!

I thank the Max Planck Institute for Psycholinguistics who hosted me during my first year as a PhD. Special thanks go to my first office mates Helen, Sho and Nienke, who gave me a warm welcome, patiently answered all the burning questions of a first year PhD student and shared their secret supply of cookies with me. I was particularly grateful for the opportunity to present my studies in different research groups at the MPI, having access to the MPI research facilities and participating in the IMPRS for Language Sciences. I thank the coordinators of the IMPRS, Dirkje, Els, and Kevin, and all the people who put together an interesting and enjoyable curriculum, not to forget the people who helped organizing the legendary IMPRS pizza evenings. I would also like to thank the former coordinator of the Graduate School for the Humanities at Radboud University, especially their former coordinator Tanja Döller. Tanja, you were a great source of encouragement in difficult times and took good care of me during all those years.

Last but not least in the academic journey section I would like to thank the (former) members of the Centre for Language Studies, the Speech Comprehension group and the Language and Speech Technology group. Juul, Polina, Remy, Cezara, Mario, Eric, Louis, Marjoke, Wessel, Ferdy, Claire, Mirjam, Mario and Sara, thank you for your support, the lunches, the coffee breaks. The final lines of this section I would like to dedicate to my paranymphs Sara and Mario. Mario, you have an intriguing personality – attentive, cheerful and always in for a joke. I owe you a dept of gratitude not only for your support in organizing my defense but for your positivity. I wish you and your family all the best for the future and hope we will stay in contact. Sara, words cannot express what you have done for me in the past years – as a colleague, as a paranymph, but first and foremost as my friend. Although coming from different cultures we immediately bonded and I am proud to call you my friend. Thank you for your support, for long

discussions, for sharing the secrets of Persian cuisine with me and for being such a wonderful person!

The personal journey

I would not have survived the academic journey without the support of my family and friends. First, I would like to thank the Cranecastle City Crew: my Esther-siblings Conny and Thordis. Conny, we almost became flat mates when I moved to Kranenburg and looking back I regret that we did not make that happen. I will always remember our daily drives to Nijmegen, the Sinterklaas festivities and our evening and weekend activities. Looking forward to seeing soon! Thordis I have always looked up to you – as a colleague and as a friend! Your thoughtfulness, your persistence and your way of managing things have been a great role model for me. Thanks for always having an open ear, the cosy dinners at your place, your encouraging words and for sharing your passion for tea with me. Antje, thanks for your emotional support and for being ready to join my spontaneous ideas, like the trip to Hungary.

Birthe, my Saarland compatriot. We moved to Kranenburg at the same time not knowing each other and both returned to Saarland, now being friends. Thanks for cooking dinner for me and for providing some Saarlandian nostalgia when I felt homesick! Above all I owe you my deepest gratitude for introducing me to Ivan, the fearful (not ‘the fearsome’). My four-legged friend has been a great source of support, a faithful companion, and the best Kranenburg souvenir I could possibly have.

I would also like to thank my belly dance teachers and fellow dancers: wherever I went, belly dance classes have been an important prerequisite for my wellbeing. Aziza (Nijmegen), you introduced me to many different dance styles; Nur el Ayun (Zuerich) I learned a lot about improvisation and expression from you; The Velvet Serpents (Saarbruecken), although I had officially left the group when I started my PhD in Nijmegen, you always made room for me when I returned to Saarbruecken during the summers; Anima Serpentis (Saarbruecken), you girls rock!

Throughout my PhD, I also received remote support from my friends Elisa, Florian, Pia, Pia, Marco, Andi, Julia, Björn, Kathrin, Gard, Nicole, Tim, Nicola, Daniela and Michael. Thank you for always being there and never letting me down. I also like to thank my colleagues at SemVox for their support during the final period of my PhD.

In this last paragraph I would like to thank my whole family. My grandmothers, to whom I dedicate this thesis: It breaks my heart that you can't live to see that I made it! Liebe Tante Marlies, deine Unabhängigkeit und deine Reiselust waren schon früh ein Vorbild für mich. Danke, dass du mich immer dazu ermuntert hast es dir gleich zu tun. Mama und Papa, ihr habt mich nicht nur während der Doktorarbeit unterstützt, sondern unterstützt mich schon mein ganzes Leben. Ihr habt bei jedem Schritt an mich geglaubt, vor allem wenn ich selbst an mir gezweifelt habe. Danke! Liebe Schwiegereltern, liebe Daniela, danke, dass ihr immer Verständnis für eure verrückte Schwiegertochter/Schwägerin gezeigt und meinen Mann in den fast vier Jahren Fernbeziehung unterstützt habt. Lucia, ein besonderes Dankschön an dich, für deine herzliche Art, die Carepakete und die SMS, die du mir vor jeder meiner Konferenzreisen geschickt hast.

Dear Patrick, you are not only my husband but also my closest friend and biggest supporter. I know that loving me is not an easy undertaking as I have been constantly on the move in the past ten years. Thank you for your patience, for traveling almost the length of the equator during the 3.5 years of our long-distance relationship, and for encouraging me to believe in myself. Off to the next chapter!

Curriculum Vitae

Juliane Kirsch, née Schmidt, was born in 1986 in Saarbruecken, Germany. In 2011, she received her Master's degree in Phonetics, English Linguistics and Developmental Psychology from Saarland University, Saarbruecken, Germany. In fall 2012, Juliane joined the Marie Curie training network "Investigating Speech Processing In Realistic Environments (INSPIRE)" and started her PhD project at the Centre for Language Studies, Radboud University, Nijmegen and the International Max Planck Research School for Language Sciences. Since 2016 Juliane is working as a project manager and linguistic consultant at SemVox GmbH in Saarbruecken who develop intelligent speech dialog solutions.

Publications

Schmidt, J., Scharenborg, O., Herzog, D., & Janse E. (in preparation). Perceiving Arousal in speech: Effects of hearing loss, loudness processing ability, and hearing aid use.

Schmidt, J., Janse, E., & Scharenborg, O. (2016). Perception of emotion in conversational speech by younger and older listeners. *Frontiers in Psychology*, *7*, 781. doi:10.3389/fpsyg.2016.00781

Schmidt, J., Herzog, D., Scharenborg, O., & Janse, E. (2016). Do hearing aids improve affect perception? In van Dijk, P., Başkent, D., Gaudrain, E., de Kleine, E., Wagner, A., and Lanting, C. (Eds.), *Physiology, Psychoacoustics and Cognition in Normal and Impaired Hearing*. Advances in Experimental Medicine and Biology, 894 (pp. 47–55). Springer International Publishing, Cham. http://doi.org/10.1007/978-3-319-25474-6

Schmidt, J., Scharenborg, O., Janse, E. (2015). Semantic processing of spoken words under cognitive load in older listeners. In M. Wolters, J. Livingstone, B. Beattie, R. Smith, M. MacMahon, J. Stuart-Smith, & J. Scobbie (Eds.), *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS 2015)*. London: International Phonetic Association.

Schmidt, J., Janse, E., & Scharenborg, O. (2014). Age, hearing loss and the perception of affective utterances in conversational speech. In *Proceedings of the 15th Annual Conference of the International Speech Communication Association* (pp. 1929–1933).

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 Carrión Castillo*

115. Music and language comprehension in the brain. *Richard Kunert*

116. Comprehending Comprehension: Insights from neuronal oscillations on the neuronal basis of language. *Nietzsche H.L. Lam*

117. The biology of variation in anatomical brain asymmetries. *Tulio Guadalupe*

118. Language processing in a conversation context. *Lotte Schoot*

119. Achieving mutual understanding in Argentine Sign Language. *Elizabeth Manrique*

120. Talking Sense: the behavioural and neural correlates of sound symbolism. *Gwilym Lockwood*

121. Getting under your skin: The role of perspective and simulation of experience in narrative comprehension. *Franziska Hartung*

122. Sensorimotor experience in speech perception. *Will Schuerman*

123. Explorations of beta-band neural oscillations during language comprehension: Sentence processing and beyond. *Ashley Lewis*

124. Influences on the magnitude of syntactic priming. *Evelien Heyselaar*

125. Lapse organization in interaction. *Elliott Hoey*

126. The processing of reduced word pronunciation variants by natives and foreign language learners: Evidence from French casual speech. *Sophie Brand*

127. The neighbors will tell you what to expect: Effects of aging and predictability on language processing. *Cornelia Moers*

128. The role of voice and word order in incremental sentence processing. *Sebastian Sauppe*

129. Learning from the (un)expected: Age and individual differences in statistical learning and perceptual learning in speech. *Thordis Neger*

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 Stoehr*

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*