International Journal of Language & Communication Disorders

INT J LANG COMMUN DISORD, XXXX 2018,
VOL. 00, NO. 0, 1–15

Research Report

Speech understanding in noise in elderly adults: the effect of inhibitory control and syntactic complexity

Eline C. van Knijff†, Martine Coene†‡ and Paul J. Govaerts†‡

†Language and Hearing Center Amsterdam, Vrije Universiteit Amsterdam, the Netherlands
‡The Eargroup, Antwerp, Belgium

(*Received May 2017; accepted January 2018*)

Abstract

Background: Previous research has suggested that speech perception in elderly adults is influenced not only by age-related hearing loss or presbycusis but also by declines in cognitive abilities, by background noise and by the syntactic complexity of the message.
Aims: To gain further insight into the influence of these cognitive as well as acoustic and linguistic factors on speech perception in elderly adults by investigating inhibitory control as a listener characteristic and background noise type and syntactic complexity as input characteristics.
Methods & Procedures: Phoneme identification was measured in different noise conditions and in different linguistic contexts (single words, sentences with varying syntactic complexity). Additionally, inhibitory control was measured using a visual stimulus–response matching task. Fifty-one adults participated in this study, including elderly adults with age-related hearing loss ($n = 9$) and with normal hearing ($n = 17$), and a control group of normal hearing younger adults ($n = 25$).
Outcomes & Results: The analysis revealed that elderly adults with normal hearing and with hearing loss were less likely to identify successfully phonemes in single words than younger normal hearing controls. In the context of sentences, only elderly adults with hearing loss had a lower odds of correct phoneme perception than the control group. Additionally, in elderly adults with hearing loss, phoneme-in-sentence perception was linked to age-related declines in inhibitory control. In all participants, phoneme identification in sentences was influenced by both noise type and syntactic complexity.
Conclusions & Implications: Inhibitory control and syntactic complexity might play a significant role in speech perception, especially in elderly listeners. These factors might also influence the results of clinical assessments of speech perception. Testing procedures thus need to be selected and their results interpreted carefully with these influences in mind.

Keywords: speech-in-noise, speech perception, syntactic complexity, inhibitory control, hearing loss, presbycusis.

What this paper adds

What is already known on the subject

Successful speech perception in the elderly population has been linked to several factors other than hearing loss, including executive functioning. Also, speech perception may be influenced by the syntactic complexity of the message. However, it is not yet clear what the exact contribution of such linguistic factors is on the perception of speech-in-noise in elderly adults, nor what is the precise role of executive functioning.

What this paper adds to existing knowledge

Contrary to previous studies, we have found that non-auditory factors such as linguistic complexity and inhibitory control influence speech perception in both elderly adults and a younger control group. These results underline that

Address correspondence to: Martine Coene, Department of Language, Literature and Communication, Language and Hearing Center Amsterdam, VU Free University Amsterdam, De Boelelaan 1105, NL-1081 HV Amsterdam, the Netherlands; e-mail: m.m.r.coene@vu.nl.

ISSN 1368-2822 print/ISSN 1460-6984 online
DOI: 10.1111/1460-6984.12376

speech perception is a complex, multifactorial process, of which not all elements are generally taken into account in clinical settings.

What are the potential or actual clinical implications of this work?
Factors such as inhibitory control and syntactic complexity may influence speech perception outcomes in elderly adults and other clinical populations. This is relevant for speech perception in real-life settings, but also in clinical assessments. Clinicians should, therefore, keep these issues in mind when choosing speech-perception assessments and interpreting their outcomes to help ensure the appropriate procedure for treatment and/or rehabilitation is selected.

Introduction

For many ageing adults, understanding speech can be quite challenging.[1] Apart from the fact that increasing age is often associated with some form of hearing loss, research indicates that other factors could play a role as well. Many previous studies have been dedicated to such non-auditory listener characteristics influencing speech understanding. Listeners use their knowledge of their language(s) and the real world to predict and fill in missing auditory information (Benichov *et al.* 2012). In elderly adults, an age-related weakening of executive functions, more specifically working memory and inhibitory control, is thought to play a role as well (Bialystok *et al.* 2004, Borella *et al.* 2008, Sweeney 2001).

Furthermore, it is well known that the listening environment and the linguistic features of the speech signal can negatively affect speech perception. First, perceptual abilities are negatively affected by background noise. Successful speech understanding is dependent on the signal-to-noise ratio (SNR) and the type of noise (Festen and Plomp 1990, Füllgrabe *et al.* 2006, Lamoré and Houtgast 2010). Second, in natural conversations, speech sounds are embedded in larger contexts such as words and sentences. In the literature, such linguistic contexts have been shown to aid the perception of speech (Benichov *et al.* 2012, Pichora-Fuller 2008). Yet, linguistic context characterized by high syntactic complexity may hinder speech understanding (Carroll and Ruigendijk 2013, Uslar *et al.* 2011).

Summarizing, sentence understanding depends on established listener-related factors such as hearing loss, executive functioning and inhibitory control, on the one hand, and speech-related factors such as syntactic complexity, the mere presence of background noise as well as the particular type of background noise, on the other hand. This study will take both types of factors into consideration and investigate to what extent each contributes to phoneme understanding in elderly listeners.

Aims

The main aim is to understand better the age-related changes in speech perception by investigating the influence of hearing loss and cognitive decline in elderly adults on speech understanding of phonemes in words and in sentences with varying levels of syntactic complexity. When time progresses, every person faces some form of cognitive decline, broadly defined as an age-related deterioration of the biological framework that underlines the ability to think and reason. Behavioural signs of such age-related changes in cognitive ability include, amongst others, forgetfulness, a decreased ability to maintain focus and decreased problem-solving capability. In several studies, working memory capacity has been shown to be a crucial factor in speech in noise understanding (Rönnberg *et al.* 2010, Rudner *et al.* 2012, Gordon-Salant and Cole 2016). Recent work has shown that next to working memory, inhibitory control also plays an important role in segregating target speech from distracting background noise. Older individuals with a hearing loss have been found to have more difficulties at switching their attention from irrelevant to relevant stimuli than normally hearing younger adults (Stenbäck 2016). Within this particular context, our study focuses on the potential negative effect of declining inhibitory control on speech-in-noise understanding. The target population includes normal hearing elderly adults and elderly adults with hearing loss and a control group of normal hearing younger adults. In addition, another somewhat under-researched non-auditory aspect under investigation in this study concerns the potential effect of linguistic complexity of the test stimulus on speech understanding. This is assessed by comparing speech understanding based on constructions with low syntactic complexity such as subject–verb–complement (SVC) clauses to syntactically more complex relative clauses.

Research questions and hypotheses

Against this combined audiological and (psycho-)-linguistic background, we will address the following research questions:

- Is phoneme identification *in single words* against a background of speech noise influenced by inhibitory control? If so, is the effect of inhibitory control reinforced by previously established factors such as age and hearing loss?

- Do stimulus-related characteristics such as the grammatical features of the language system (syntactic complexity) and the acoustic characteristics of masking noise (noise fluctuations), be it on their own or in combination, affect the speech perception accuracy of word-initial phonemes that are part of a *carrier sentence*?

With respect to the first question, we expect phoneme-in-word understanding in noise to be negatively influenced by lowered inhibitory control and this effect to be more pronounced in elderly listeners than in younger adults and even more so if those elderly listeners are suffering from hearing loss. As for the second question, we expect that the negative effect of syntactic complexity on speech understanding may be partially cancelled by introducing spectro-temporal fluctuations in the speech noise masker.

In what follows, we will provide a motivation for both hypotheses based on the current state of the art concerning the role of auditory and non-auditory factors in speech understanding.

Age-related cognitive decline and hearing performance in the elderly

Elderly adults are often confronted with age-related hearing loss or presbycusis, which is characterized by a shift in hearing threshold, most notably in the higher frequencies, a decline in dynamic range and decreased frequency selectivity (Kapteyn and Lamoré 2012, Tun *et al.* 2012). Due to these changes, individuals with presbycusis experience a decreased ability to understand speech in noise, although often speech sounds are still heard. However, not all speech perception difficulties in elderly adults can be explained by hearing loss. A large portion of elderly individuals with normal hearing seem to exhibit perception difficulties as well (Wingfield *et al.* 2006). These have been related to age-related changes in general cognitive abilities (Benichov *et al.* 2012, Houtgast and Festen 2008, Tun *et al.* 2012).

The link between age-related cognitive decline and hearing performance is further confirmed by populations of elderly listeners who exhibit characteristics of pathological cognitive ageing. A recent study has shown that (self-reported) hearing loss is associated with a lower baseline Minimal Mental State Examination (MMSE) score and a greater decline in cognitive performance during a 25-year follow-up period (Amieva *et al.* 2015). Several papers seem to indicate that the close relationship between hearing impairment and age-related cognitive impairment holds in the other direction as well. Older adults with hearing loss are more likely to develop Alzheimer's disease and other types of dementia compared with hearing peers (Lin 2011, Lin *et al.* 2011). In contrast, hearing rehabilitation such as the use of hearing aids has been associated with better cognitive performance and reduced cognitive decline (Amieva *et al.* 2015, Dawes *et al.* 2015). Summarizing, the current state of the art points in the direction of a mutual reinforcement of age-related decline in cognitive abilities and in hearing performance. In what follows, we will focus on inhibitory control, a particular aspect of executive functioning that is taken to play a key role in age-related cognitive processes influencing speech understanding.

Executive functioning

Despite considerable variation between individuals, cognitive abilities such as executive functioning have been shown to decline with age (Bialystok *et al.* 2004, Borella *et al.* 2008, Sweeney 2001, Lubbe and van der Verleger 2002, Zelazo *et al.* 2004). Executive functions can be defined as the cognitive processes that plan, coordinate and regulate other cognitive processes including working memory, attention and inhibitory control (Borella *et al.* 2008). Inhibitory control is an important aspect of executive functioning that is hypothesized to help prevent irrelevant competing information from overloading working memory, as it allows people to eliminate distracting factors and focus on relevant task goals. As an important factor that determines the capacity limits of working memory, insufficient inhibitory control may turn listening in non-optimal conditions (e.g., in the presence of background noise and/or hearing loss) into an effortful activity. Interestingly, inhibitory control has been shown to decline with age (Bialystok *et al.* 2004, Lubbe and van der Verleger 2002), perhaps even more strongly than other aspects of executive functioning (Sweeney 2001). Non-verbal cognitive skills may help the listener to process the message when the accessibility to speech is compromised due to interference from background noise (Benichov *et al.* 2012). When confronted with speech against a noisy background, a listener's ability to eliminate the distractions of the noise is therefore expected to aid speech understanding. In this study, we will therefore be particularly concerned with the potential influence of inhibitory control on speech perception in noise in elderly adults.

Noise masking

In everyday life, background noise is present in almost all listening situations. For SNRs at or below the listener's threshold, speech perception becomes more challenging, especially for individuals with hearing loss. Elderly listeners with hearing loss need a more advantageous SNR than normal hearing adults to achieve similar speech understanding in stationary speech-shaped noise (George *et al.* 2007).

Speech understanding has been shown to be differently affected by distinct types of noise. For normal hearing individuals, speech reception thresholds have been shown to improve when small silent gaps occur within the noise signal (so-called 'fluctuating noise'; e.g., Füllgrabe *et al.* 2006). Significant spectro-temporal modulations in the masker may create glimpses in the speech-noise background that listeners may use to identify speech more easily. The observed 'masking release' effect is influenced by several different factors, including the nature of the fluctuations and the noise itself (Cook 2006), the cognitive abilities of the listener (Füllgrabe *et al.* 2006, George *et al.* 2006, Rönnberg *et al.* 2010) and hearing loss (Festen and Plomp 1990). Masking release has been consistently found in young, normal hearing listeners for different types of speech stimuli (Festen and Plomp 1990, George *et al.* 2006). Whereas normal hearing listeners have the ability to use the noise gaps to complete missing speech cues by using, for example, acoustic and/or lexical redundancies in the speech signal (Füllgrabe *et al.* 2006), hearing impaired listeners are less likely to benefit from such gaps in noise (e.g., George *et al.* 2006). As speech perception against a background of modulated noise has been shown to require additional cognitive resources (Clarke *et al.* 2014, Francart *et al.* 2011) the ability to 'dip listen', i.e., to make optimal use of the acoustic glimpses of the speech signal when the background noise temporarily drops, is especially affected when cognitive capacities are reduced (Rönnberg *et al.* 2010).

Syntactic complexity

Apart from the aforementioned factors, it is important to note that contextual information, even when limited, can influence speech perception. Listeners use their syntactic, (meta-)linguistic and general knowledge to make up for perception difficulties, at different linguistic levels of speech perception (Benichov *et al.* 2012). Although the influence of linguistic knowledge is rarely considered during clinical assessments of hearing loss, recent work shows that in speech audiometric assessments an important portion of consonant confusions are explained by such linguistic factors (Coene *et al.* 2015). At the level of syntax, in structurally simple 'canonical' utterances (exemplified for Germanic languages such as English or Dutch by subject–verb–object (SVO) structures, e.g., 'Sarah has bought bananas') a listener may fill in a potentially missed word based on information provided by the language system (Benichov *et al.* 2012, Rönnberg *et al.* 2013).

However, natural conversation includes sentences in a wide variety of syntactic structures, including structures that may not be supportive to speech perception and may even put a burden on cognitive processing. Psycholinguistic research has shown that the human processor follows an efficiency principle that builds on the size of the syntactic domain in which a given grammatical relation can be processed. The human mind has a clear preference for the smallest possible domains, i.e., exhibiting the smallest structural distance between interrelated items (Hawkins 1994, 2014).

The idea that the cognitive-processing complexity of an utterance is proportionally related to the size of its syntactic domain can be conveniently exemplified by transitive verbs that may take either a noun or an embedded clause as their complement. For example, an utterance such as 'The message claimed her attention' will be less difficult to process than 'He claims to know her' because in the latter case the object of the *claim* is an infinitive clause that enlarges the syntactic domain with its own internal clausal structure ('he' and 'her' being respectively the subject and the object of the verbal predicate 'to know'). Syntactic complexity partially erases the positive effects of context on speech understanding when combined with other perceptual challenges, such as speech noise (Coene *et al.* 2016, Uslar *et al.* 2013) or increased presentation speed, especially for older individuals with hearing loss (Carroll and Ruigendijk 2013, Uslar *et al.* 2011, Wingfield *et al.* 2006). Perception becomes more difficult in sentences with a non-canonical word order such as passive clauses (e.g., 'Information will be provided tomorrow') or in sentences with more than one possible verb argument structure (e.g., sentences with the verb to send: 'Sheldon sent the letter' and 'Sheldon sent Debbie the letter') (Uslar *et al.* 2011: 623–624).

This is also the case in complex sentences with a centre-embedded object relative clause (e.g., 'The man who the bear attacked ran away') (Wingfield *et al.* 2006). More specifically, recent work revealed that younger adults show significantly better speech-in-noise perception in simple canonical German sentences in comparison with the aforementioned complex structures, while adults between 41 and 54 years of age (normal hearing and with hearing loss) seem to be unable to benefit from the simpler sentence structures to the same degree (Uslar *et al.* 2011).

Materials and methods

Participants

Seventeen elderly adults with hearing loss (HL, age range 65–86 years), nine elderly adults with normal hearing (NH, age range 65–81 years) and 25 younger normal

Table 1. Demographic information per participant

Participant number	Age (years)	PTA	Participant number (continued)	Age (years)	PTA
	Elderly NH			*Control*	
1	65	16.25	27	31	13.75
2	65	18.75	28	33	12.50
3	66	15.00	29	33	12.50
4	67	20.00	30	34	12.50
5	67	21.25	31	37	10.00
6	68	21.25	32	37	13.75
7	72	23.75	33	38	16.25
8	75	22.50	34	39	10.00
9	81	22.50	35	41	6.25
			36	55	17.50
Elderly HL			37	43	8.75
10	65	28.75	39	48	12.50
11	66	31.25	40	50	21.25
12	68	33.75	41	51	10.00
13	69	43.75	42	53	22.50
14	70	48.75	43	53	10.00
15	71	30.00	44	53	21.25
16	71	51.25	45	54	18.75
17	72	63.75	46	55	17.50
18	72	60.00	47	56	17.50
19	75	25.00	48	58	18.75
20	75	32.50	49	59	21.25
21	77	33.75	50	60	21.25
22	78	28.75	51	61	21.25
23	80	45.00			
24	81	31.25			
25	84	28.75			
26	86	61.25			

Note: PTA, pure tone average (decibel hearing level, dB HL) over 0.5, 1, 2 and 4 kHz in the better ear; HL, hearing loss; NH, normal hearing.

Table 2. Overall demographic information, mean (standard deviation; range)

Group	*N*	Age (years)	PTA
Elderly HL	17	74 (6.1; 65–86)	40 (12.9; 25–63.75)
Elderly NH	9	70 (5.4; 65–81)	20 (3.0; 15–23.75)
Control	25	47 (9.6; 31–61)	15 (4.8; 6.25–22.5)
Total	51	60 (15; 31–86)	24.24 (13.9; 6.25–63.75)

Note: PTA, pure tone average (decibel hearing level, dB HL) over 500, 1000, 2000 and 4000 Hz in the better ear; HL, hearing loss; NH, normal hearing.

hearing adults (control, age range 31–61 years) were recruited, resulting in a total of 51 monolingual Dutch participants (28 women, 23 men). Normal hearing was defined as a pure tone average (PTA) over 0.5, 1, 2 and 4 kHz) of less than 25 dB HL (decibel hearing level) at the better ear (World Health Organisation (WHO) 2017). For more detailed demographic information, see tables 1 and 2.

Ethical approval for this study was granted by the ethical committee of the Vrije Universiteit Amsterdam (approval number VU EC02.14b). Participation was voluntary and there were no rewards, costs or risks attached to participating. All participants provided written informed consent prior to partaking in the study.

Pure tone audiometry

Pure tone audiometry was conducted using a portable screening audiometer (Interacoustics AS608e, Interacoustics, Middelfart, Denmark). Hearing thresholds were obtained for 0.5, 1, 2 and 4 kHz in each ear by means of the Hughson–Westlake 'ascending' procedure (Newhart and Reger 1945). The PTA of the better ear was used for subsequent analyses. Whenever hearing loss was present, participants were informed and advised to visit an audiologist.

Speech perception task

Participants' understanding of words and sentences was measured in silence, in stationary noise and in modulated noise. The target words consisted of 54 highly frequent, monosyllabic nouns, selected from the Spoken Dutch Corpus (Corpus Gesproken Nederlands—CGN) (Oostdijk 2003). Twenty-seven minimal pairs, one for each listening condition, were selected according to their word-initial consonants, yielding the following 14 phonemic contrasts: /p/-/b/, /t/-/d, /f/-/v/, /r/-/l/, /b/-/v/, /t/-/s/, /m/-/b/, /d/-/n/, /m/-/n/, /p/-/t/, /b/-/d/, /b/-/k/ and /t/-/k/. Some pairs occurred multiple times throughout the task, but only once per condition and represented through different target words. Apart from /f/, each phoneme occurred at least twice.

Eighteen of the selected target words were presented as single words, while the remaining pairs were divided over seven-word sentences with low or high syntactic complexity. Within these sentences, the target phoneme was always part of the fourth word of a sentence. Sentence length was kept stable as variations in length are known to influence working memory demands. The syntactically 'simple' condition included sentences with either SVC word order or with topic–verb–subject word order. Within these constructions, the target noun could take up different syntactic functions including that of object of a transitive verb, e.g., 'Marie kocht een *paard* voor haar man' (Mary bought a *horse* for her husband), prepositional complement of an (in)transitive verb, e.g., 'In het grote *bad* zwommen drie eendjes' (literally, In the big *bathtub* were swimming three little ducks) or indirect object of a ditransitive verb, e.g., 'Dirk gaf de *kok* een lange brief' (literally, Dirk gave the *cook* a long letter). In the syntactically 'complex' condition, the target word functioned as the antecedent of a subject-relative clause, such as 'Dat is het *boek* dat bijna viel' (That is the *book* that almost fell). The proposed complexity ranking (words in isolation < canonical SVO or SVC sentences < complex sentences with an embedded relative clause)

Table 3. Minimal phonemic pairs for the word and sentence repetition tasks

	Complexity level		
Listening condition	Isolation	Simple	Complex
Silence	/m/–/n/	/p/–/t/	/b/–/d/
	/r/–/l/	/b/–/v/	/t/–/s/
	/p/–/b/	/t/–/d/	/f/–/v/
Stationary noise	/b/–/d/	/b/–/k/	/b/–/d/
	/t/–/s/	/m/–/b/	/r/–/l/
	/p/–/b/	/b/–/p/	/t/–/d/
Modulated noise	/t/–/k/	/b/–/d/	/b/–/d/
	/d/–/n/	/r/–/l/	/m/–/b/
	/t/–/d/	/t/–/d/	/b/–/p/

Note: For each listening condition and for each linguistic complexity level, three pairs of phonemic contrasts were used.

builds on Hawkin's (1994) efficiency principle reflecting the increasing size of the syntactic domain from left to right. An overview of the phonemic contrasts used per listening condition and per linguistic complexity level is given in table 3.

All words and sentences were recorded using a Plantronics audio 400 DSP headset by a female speaker, using a flat intonation pattern and a normal conversational speaking rate. A masker whose spectrum equalled that of the long-term average speech spectrum was applied to the 36 speech stimuli at an SNR of 0 dB. The noise was kept stationary for one-third of the stimuli (18 stimuli) and amplitude modulated for another one-third using an 8 Hz sine wave modulation rate (18 stimuli) with modulation depth fixed at 100%. The remaining 18 stimuli were presented without noise.

The task was presented using a Dell Lattitude E4200 laptop equipped with a 30 W Sweex 2.0 Speaker set. Using an A SPL meter (the t.meter MPAA1), the speaker volume was adjusted to ensure stimuli presentation at approximately 70 dB SPL at a distance of 1 m from the speakers. The words and sentences were presented using the Auditory Speech Sound Evaluation (A§E) psychoacoustic test suite (Govaerts *et al.* 2006). Stimuli were grouped in different subtests for each listening condition (silence, stationary or modulated noise) for the word and both sentence conditions. Each participant first completed all stimuli in the silent condition, followed by the stimuli in modulated noise and finally stationary noise. The order of the different complexity conditions was randomized within each noise condition by the A§E test suite, but the order of words within each sub-condition was the same for each participant.

Participants were instructed to repeat everything they heard, even in the case of uncertainty or seemingly odd words or sentence structures. Bearing in mind that minimal pairs of target words were formed at the phonemic level, repetition accuracy was scored based on the correct repetition of the initial phoneme of the target word.

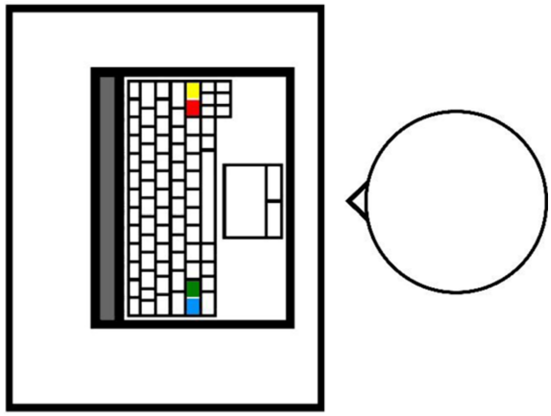

Figure 1. Location of response keys for the Simon task on a laptop keyboard.

Simon task

Inhibitory control was tested by administering the Simon task (Simon and Wolf 1963). The outcome measure of this task was the difference in reaction time (RT) between incongruent and congruent trials (ΔRT = incongruent RT – congruent RT) of coloured stimuli that were presented either on the left or the right side of the test screen (Bialystok *et al.* 2004, Lubbe and van der Verleger 2002). In the original task, stimuli vary between two colours (e.g., red, green) and two locations (left, right). In congruent trials, a coloured circle is presented on the same side (e.g., a red circle on the left) as its corresponding response button (left), while in incongruent trials this coloured circle is projected on the opposite side of the screen (right). Participants are asked to correctly identify the stimulus' colour while ignoring its location.

Here, an adapted version of this task was created by using four coloured circles (blue, green, red and yellow) to increase task demands (Bialystok *et al.* 2004). Each of the four colours was linked to one of two possible response keys: the left and right shift keys of the laptop keyboard were marked as response keys by a blue–green and a red–yellow sticker respectively (figure 1). Participants were instructed to press the blue–green key if a blue or green circle appeared on the screen, and the red–yellow key if a red or yellow circle appeared, as quickly as possible. All stimuli were presented on the same laptop with a 31 cm (12.1 inch) screen width. The task was programmed using PEBL software (Mueller 2012), by adapting an existing template for the Simon task.

Each trial started by presenting a white fixation cross on a black background in the middle of the screen for 500 ms, after which a coloured circle appeared on either

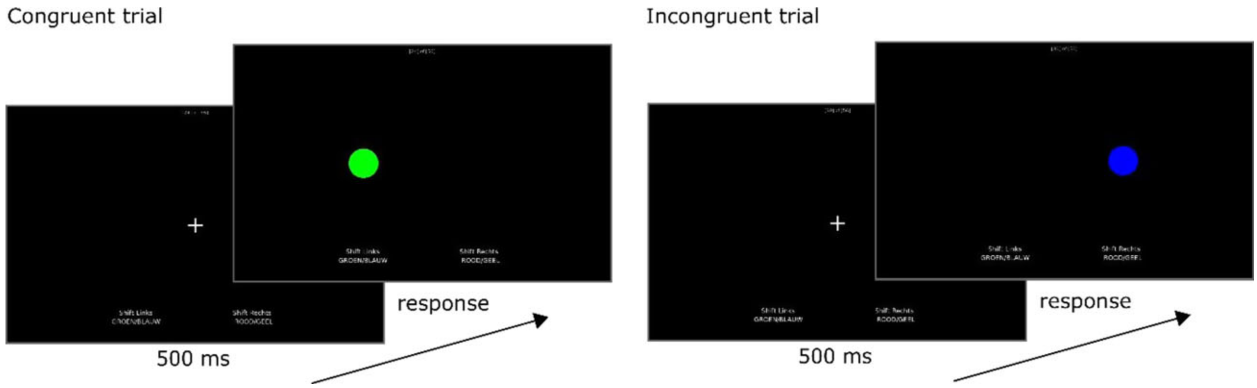

Figure 2. Presentation order of the trials in the Simon task for blue/green stimuli (left response key).

the left or right side of the screen. In congruent trials, the circle appeared on the same side as the correct response key for the colour of the circle (figure 2). In incongruent trials, the location of the response key and the location of the circle were opposites, e.g., a blue circle—with the left shift key as response key—presented on the right side (figure 2). Each of the eight location–colour combinations was repeated seven times in a random order, resulting in 56 trials in total. All visual test stimuli were presented at 0° azimuth. Participants wearing glasses or hearing devices were asked to keep them on during the Simon task. Instructions were presented on screen and orally supported by the experimenter. Before starting the test trials, participants were required to complete eight practice trials correctly, one for each colour–location combination. Participants were allowed to take a short self-timed break after 28 trials. Only correct responses were used for analysis. Subsequently, the so-called Simon effect was calculated by subtracting the mean RT for the congruent trails from the RT for the incongruent trials(ΔRT). This difference represents the added RT participants needed to cope with incongruent trials, with a lower ΔRT indicating better inhibitory control.

Procedure

At each testing session, the participant was first explained the general procedure. Participants were then asked to read and sign the informed consent form and provide relevant demographic information. Subsequently, the audiometry was performed to determine the existence and specifics of any hearing loss. Participants were shown their results directly afterwards and received information about audiological professionals in the area if applicable.

Following the audiometry, participants performed the speech perception task and the Simon task, in a random order. During the speech perception task, participants were seated at approximately 1 m from the speaker. Participants with hearing aids were requested to remove their devices during the audiometry and the speech perception task. During the Simon task, participants were seated behind the laptop.

Statistical analysis

A generalized estimating equations (GEE) approach was used to assess the various factors influencing correct or incorrect speech understanding at the word and sentence level. This approach allows for the modelling of clustered data, which is applicable to the current dataset as each participant has multiple (repeated) scores with a binary outcome. A total of 612 observations was analyzed for phoneme-in-word understanding, representing 12 observations for 51 participants. The total number of observations in sentence contexts was 1224, with 24 observations per participant.

In the GEEs, the quasi-likelihood under independence model criterion (QIC) was used to select the most efficient working correlation matrix (Pan 2001). Following the analysis of Scholtens *et al.* (2017), the corrected quasi-likelihood under independence model criterion (QICC) was used to select the best fitting combination of predictors and interactions, starting from a model only including main effects, adding interactions terms one by one. For both QIC and QICC, lower values indicate a better fit (Pan 2001). Coefficients (*B*) denote the estimate of change in probability of correct perception that can be attributed to the change of one unit in a predictor variable. Odds ratios (ORs) denote the change in odds of correct perception that can be attributed to change of one unit in a predictor variable. Significance levels were set at $p < .05$. All analyses were performed in IBM SPSS Statistics 24.

Results

Table 4 presents a descriptive summary of the phoneme perception scores for words and sentences in noise and silence. It shows a ceiling effect in the silence condition for both control group and normal hearing elderly

Table 4. Five-point summary of phoneme perception in words and in sentences

	Minimum	Q1	Median	Q3	Maximum
Phoneme-in-words in silence (% correct)	33	100	100	100	100
Elderly HL	33	67	83	100	100
Elderly NH	83	100	100	100	100
Control	100	100	100	100	100
Phoneme-in-sentences in silence (% correct)	17	100	100	100	100
Elderly HL	17	42	92	100	100
Elderly NH	75	100	100	100	100
Control	92	100	100	100	100
Phoneme-in-words in noise (% correct)	0	33	50	58	75
Elderly HL	0	0	33	42	58
Elderly NH	33	42	50	50	67
Control	25	50	58	67	75
Phoneme-in-sentences in noise (% correct)	0	29	54	63	75
Elderly HL	0	4	21	33	50
Elderly NH	8	42	46	54	63
Control	46	58	63	67	75

Note: HL, hearing loss; NH, normal hearing; Q1, percentile 25; Q3, percentile 75.

Table 5. Descriptive mean (standard deviation) reaction times (RTs) for the different types of trials in the Simon task

Group	*N*	Congruent	Incongruent	Simon effect
Elderly HL	17	805.05 (200.29)	868.36 (205.75)	63.31 (55.25)
Elderly NH	9	732.72 (122.99)	806.25 (156.81)	73.53 (56.47)
Control	25	572.60 (108.96)	598.42 (95.25)	25.82 (37.65)
All groups	50	678.34 (180.38)	725.08 (195.00)	46.74 (51.08)

Note: HL, hearing loss; NH, normal hearing; Simon effect = ΔRT incongruent/congruent trials (ms).

listeners. The conclusion to be drawn from this ceiling effect is that the linguistic content of the test items did not show a measurable contribution to speech understanding in such an 'easy' listening context. As it is not possible to capture potential differences amongst groups of listeners in this particular listening condition, it was decided to exclude it from further analysis.

Table 5 shows the descriptive statistics for the inhibitory control outcomes (the Simon effect). A one-way analysis of variance (ANOVA) confirmed that inhibitory control was significantly different between groups, $F(2,48) = 4.886$, $p = .012$, $\eta_p^2 = .169$. Bonferroni corrected post-hoc comparisons revealed that both elderly adult groups showed weaker inhibitory control than the control group (elderly HL versus control: $\Delta M = 37.48$, SE = 14.93, $p = .047$; elderly NH versus control: $\Delta M = 47.71$, SE = 18.47, $p = .039$), while no difference was found between elderly normal hearing adults and elderly adults with hearing loss.

Phoneme-in-word understanding against a speech noise background

The association between correct or incorrect phoneme-in-word understanding, and group and Simon effect was assessed using a binomial logistic GEE model. The lowest QICC was reached by including main effects for participant group and inhibitory control, as well as the interaction between these two terms. Group (Wald χ^2(d.f.) = 11.346 (2), $p = .004$), inhibitory control (Wald χ^2(d.f.) = 5.422(1), $p = .020$), and the group by inhibitory control interaction (Wald χ^2(d.f.) = 11.568 (2), $p = .003$) each added significantly to the model. The coefficients and ORs of the final model are presented in table 6.

As shown in table 6, the odds of correct word-in-noise perception for both elderly adult groups significantly decreased in comparison with the control group. The OR for the elderly adults with hearing loss (B = –0.725, OR = 0.484) was lower than in normal hearing elderly adults (B = –0.510, OR = 0.600). The main effect for inhibitory control did not reach significance. Inhibitory control did, however, significantly interact with group. More specifically, the odds of elderly adults with hearing loss correctly perceiving phonemes-in-words against a noise background significantly decreased as their ability to inhibit task irrelevant cues decreased (B = –0.009, OR = 0.991).

Phoneme-in-sentence understanding against a speech noise background

Participant group, Simon effect, noise type and syntactic complexity were included as factors in a binomial logistic GEE model to predict correct phoneme-in-sentence understanding. The lowest QICC value (1436) was reached for a model including the main effects group (Wald χ^2(d.f.) = 45.990 (2), $p < .001$), Simon effect (Wald χ^2(d.f.) = 7.306 (1), $p = .007$), syntactic

Table 6. Coefficients (*B*) and odds ratios (ORs) of correct phoneme-in-word perception in noise

Independent variable	*B* (SE)	OR	95% CI of OR		Wald χ^2	d.f.	*p*-value
			Lower	Upper			
(Intercept)	0.249 (0.136)	1.282	0.981	1.676	3.323	1	0.068
Group							
Elderly HL	−0.725 (0.245)	0.484	0.300	0.782	8.787	1	0.003*
Elderly NH	−0.510 (0.196)	0.600	0.409	0.881	6.792	1	0.009*
Control	0	1					
Inhibitory control	−0.001 (0.003)	0.999	0.994	1.004	0.272	1	0.602
Group Inhibitory control*							
Elderly HL*Inhibitory control	−0.009 (0.004)	0.991	0.983	0.999	4.565	1	0.033*
Elderly NH*Inhibitory control	0.003 (0.003)	1.003	0.997	1.008	1.073	1	0.300
Control*Inhibitory control	0	1					

Note: Analysis was performed using generalized estimating equations (GEE). Working correlation matrix structure = exchangeable: quasi-likelihood under independence model criterion (QIC) = 790. Corrected quasi-likelihood under independence model criterion (QICC) = 792. *Significant *p*-values.

Table 7. Coefficients (*B*) and odds ratios (ORs) of correct phoneme-in-sentence perception in noise

Independent variable	*B* (SE)	OR	95% CI of OR		Wald χ^2	d.f.	*p*-value
			Lower	Upper			
(Intercept)	0.258 (0.115)	1.294	1.033	1.620	5.047	1	0.025*
Group							
Elderly HL	−1.041 (0.227)	0.353	0.226	0.551	21.034	1	< 0.001*
Elderly NH	−0.099 (0.352)	0.905	0.454	1.805	0.080	1	0.778
Control	0	1					
Inhibitory control	0.001 (0.002)	1.001	0.998	1.004	0.243	1	0.622
Noise							
Modulated	0.521 (0.128)	1.685	1.312	2.163	16.701	1	< 0.001*
Stationary	0	1					
Syntactic complexity							
Complex	0.795 (0.189)	2.214	1.528	3.209	17.646	1	< 0.001*
Simple	0	1					
Group Inhibitory control*							
Elderly HL*Inhibitory control	−0.012 (0.003)	0.988	0.982	0.995	11.244	1	0.001*
Elderly NH* Inhibitory control	−0.007 (0.005)	0.993	0.983	1.003	1.884	1	0.170
Control*Inhibitory control	0	1					
Syntactic complexity Noise*							
Complex*Modulated	−1.593 (0.235)	0.203	0.128	0.323	45.794	1	< 0.001*
Complex*Stationary	0	1					
Simple*Modulated	0	1					
Simple*Stationary	0	1					
Group Syntactic complexity*							
Elderly HL*Complex	−0.748 (0.220)	0.473	0.307	0.728	11.581	1	0.001*
Elderly HL*Simple	0	1					
Elderly NH*Complex	−0.497 (0.225)	0.608	0.391	0.944	4.903	1	0.027*
Elderly NH*Simple	0	1					
Control*Complex	0	1					
Control*Simple	0	1					

Note: Analysis was performed using generalized estimating equations (GEE). Working correlation matrix structure = exchangeable; quasi-likelihood under independence model criterion (QIC) = 1440. Corrected quasi-likelihood under independence model criterion (QICC) = 1436. *Significant *p*-values.

complexity (Wald χ^2(d.f.) = 19.632 (1), $p < .001$) and noise type (Wald χ^2(d.f.) = 5.777 (1), $p = .016$), as well as interactions between group and Simon effect (Wald χ^2(d.f.) = 11.987 (2), $p = .002$), noise type and syntactic complexity (Wald χ^2(d.f.) = 45.794 (1), $p < .001$), and group and syntactic complexity (Wald χ^2(d.f.) = 13.113 (2), $p = .001$). The coefficients and ORs for the final model are presented in table 7.

There was a significant main effect of group: elderly adults with hearing loss had significantly smaller odds of correct phoneme-in-sentence perception than the control group ($B = –1.041$, OR = 0.353). There was no significant main effect of inhibitory control (Simon effect). Yet, a significant interaction between inhibitory control and phoneme-in-sentence perception by elderly adults with hearing loss was found. Poorer inhibitory control was significantly related to poorer odds of phoneme-in-sentence perception, each increase in ms. of the Simon effect thus lowering the odds of correct phoneme-in-sentence understanding ($B = -0.012$, OR = 0.988).

A significant main effect of noise was present. In comparison with stationary noise, modulated noise increased the odds of correct perception ($B = 0.521$, OR = 1.685). In addition to this masking release effect, a main effect of syntactic complexity was found: surprisingly, the odds of correct phoneme-in-sentence repetition were higher in complex sentences than in simple sentences ($B = 0.795$, OR = 2.214). Also, noise type and syntactic complexity significantly interacted: The lowest ORs for correct repetition were found for complex sentences in modulated noise.

Table 8 summarizes the above main effects and interactions of the categorical variables, showing their significant effects the interactions between them. The values below can be interpreted as follows: in simple sentences, correct perception was more likely to occur in modulated noise than in stationary noise. This pattern was reversed in complex sentences, when the odds of correct perception were higher in stationary noise than in modulated noise. There were also differences between the participant groups. In simple sentences, normal hearing elderly adults and the normal hearing control group showed an equal likelihood of correct perception, which was higher for simple sentences in modulated noise than in simple sentences in stationary noise. Yet, in complex sentences, modulated noise had a stronger negative effect on sentence perception in elderly adults as compared with the control group. Overall, the odds of correct perception were lowest in case of increased age and hearing loss, especially combined with a condition of high syntactic complexity and modulated noise. Note that for elderly adults with hearing loss, the above described patterns were negatively influenced by weakened inhibitory control: for each ms. increase in the Simon effect, the likelihood of correct perception decreased ($B = -0.012$, OR = 0.988).

Discussion

Age-related hearing loss

In line with the literature, a group effect was present for speech understanding, as measured for words presented in isolation and in sentences with different syntactic complexity (Carroll and Ruigendijk 2013, Uslar *et al.* 2011, Wingfield *et al.* 2006). For phoneme-in-word understanding, both elderly adult groups were significantly less likely to give a correct response as compared with the younger controls, while in phoneme-in-sentence contexts, this was only the case for elderly adults with hearing loss. These results may be taken to indicate that the contextual information provided by sentences allowed normal hearing elderly adults to perform like the control group. Yet, when the sentence context was not present, perception became more difficult for elderly adults. The observed discrepancy between the results for phoneme-in-word and phoneme-in-sentence understanding against a speech noise background in elderly adults with normal hearing was in line with earlier work (e.g., Pichora-Fuller 2008), who found that that elderly adults seem to experience benefit from context during speech perception. Those results indicated that a beneficial effect of context is present in both young and old adults, and may be larger in the latter group. This difference is ascribed to experience, not necessarily in terms of general language experience, but more specifically in terms of experience with using compensation strategies to counter perception difficulties in less advantageous listening conditions (Pichora-Fuller 2008).

Inhibitory control

The current results for the Simon task were in line with the notion that including inhibitory control is susceptible to age-related decline. This was consistent with the results found by Lubbe and van der Verleger (2002) and Bialystok *et al.* (2004) discussed above in the introduction. The present study revealed a small but significant relation between inhibitory control and speech perception for elderly adults with hearing loss. Elderly adults with hearing loss and better inhibitory control in the Simon task were more likely to perform well on the speech perception task. The ability to diminish the effect of distractions may thus be relevant for speech perception in noise in ageing populations and in populations of individuals with a hearing impairment.

The present outcomes may also provide indirect evidence that normally hearing elderly adults were able to use a sentence context to cope with the presence of noise in speech perception tasks, without needing to rely on inhibitory control. For individuals facing additional challenges, e.g., a perceptual disadvantage in the form of hearing loss, inhibitory control becomes of added value to increase the odds for successful speech perception. The observed relationship between the outcomes on the tasks for inhibitory control and speech perception in elderly listeners with hearing loss should be investigated further in future research, perhaps by using a dual tasks set-up to measure processing effort during speech perception.

Table 8. Summed significant coefficients (*B*) of correct phoneme-in-sentence perception in noise for different situations, excluding interaction between participant group and inhibitory control

Situation	Intercept	Group	Context	Noise	Group*Context	Noise*Context	Sum (*B*)	OR
EHL Com Mod	0.258	−1.041	0.795	0.521	−0.748	−1.593	−1.808	0.164
EHL Com. Stat.	0.258	−1.041	0.795	0[b]	−0.748	0[b]	−0.736	0.479
EHL Sim. Mod.	0.258	−1.041	0[b]	0.521	0[b]	0[b]	−0.262	0.770
EHL Sim. Stat.	0.258	−1.041	0[b]	0[b]	0[b]	0[b]	−0.783	0.457
ENH Com. Mod.	0.258	0[a]	0.795	0.521	−0.497	−1.593	−0.516	0.597
ENH Com. Stat.	0.258	0[a]	0.795	0[b]	−0.497	0[b]	0.556	1.744
ENH Sim. Mod.	0.258	0[a]	0[b]	0.521	0[b]	0[b]	0.779	2.179
ENH Sim. Stat.	0.258	0[a]	0[b]	0[b]	0[b]	0[b]	0.258	1.294
Control Com. Mod.	0.258	0[b]	0.795	0.521	0[b]	−1.593	−0.019	0.981
Control Com. Stat.	0.258	0[b]	0.795	0[b]	0[b]	0[b]	1.053	2.866
Control Sim. Mod.	0.258	0[b]	0[b]	0.521	0[b]	0[b]	0.779	2.179
Control Sim. Stat.	0.258	0[b]	0[b]	0[b]	0[b]	0[b]	0.258	1.294

Note: OR, Odds ratio or the exponential of the summed *B*; EHL, elderly with hearing loss; ENH, elderly with normal hearing; Sim., syntactically simple; Com., syntactically complex; Mod., modulated noise; Stat., stationary noise.
[a]Coefficient is set to zero because its effect is not significant ($p > .05$).
[b]Coefficient is set to zero because it is the reference category.

Syntactic complexity and noise

The present study showed that both the grammatical features of the language system and the acoustic characteristics of masking noise play a role in speech perception. First, in simple sentences, there was an effect of masking release, i.e., the overall odds of success were better for phoneme perception in modulated noise than in stationary noise. These findings are in line with previous work by, for example, Füllgrabe *et al.* (2006), Festen and Plomp (1990), and George *et al.* (2006). Surprisingly however, our expectation that the release of masking effect due to spectro-temporal fluctuations in the speech noise masker would cancel the negative effect of syntactic complexity on phoneme perception in sentence understanding was not borne out. On the contrary, in the complex sentence condition, the odds for accurate perception in stationary noise were better than in modulated noise. Masking release seemed to be present when the speech material consisted of sentences with a frequent, predictable syntactic structure, but not in the linguistically more complex context. A possible explanation for this effect could be the depletion of cognitive resources. As pointed out by Uslar *et al.* (2013), listening in fluctuating noise requires acoustic processing skills to benefit from noise gaps, and perhaps even increased general cognitive processing (Koelewijn *et al.* 2012). At the same time, successful linguistic processing of complex syntactic structures requires cognitive resources as well. When listeners are confronted with complex syntactic structures in fluctuating noise, the cognitive resources used to benefit from the noise gaps may therefore leave insufficient capacity to process these structures successfully, since two cognitively demanding listening conditions are combined.

As pointed out by an anonymous reviewer, our suggested capacity-based explanation seems to be in line with the outcomes of a research study comparing phoneme in noise recognition in automated speech recognition systems to those obtained from human listeners. Even if sufficient information exists to support phoneme identification in the spectro-temporal glimpses of a background noise masker, human listeners—contrary to automated speech recognizers—are facing the additional cognitive challenge to identify which parts of a noisy signal should be treated as glimpses of the target speech (Cook 2006).

Nevertheless, there is no reason to assume that the same cognitive resources are necessarily competing in both the acoustic and linguistic processing of the speech materials. If they are not, a different picture might arise. This is precisely what has been found in previous research—albeit based on younger adults—where masking release effects were observed in complex sentence structures (Uslar *et al.* 2013: fig. 6). The different results might be related to a difference in methodology with the current study. Speech-in-noise perception in the study by Uslar *et al.* (2013) was estimated using the speech reception threshold (SRT), by adapting noise levels until a SNR was found at which around 80% of words per sentence were correctly repeated. Yet, in the present study, the SNR at which stimuli were presented was the same for all participants. The number of words heard per sentence might therefore have varied between participants instead of being around 80%. As participants were only scored on the repetition of the target phoneme within the relevant word, there is also no direct estimation of the percentage of words per sentence repeated to accurately compare performance or stimulus difficulty between the two conditions. The present studies stimuli presentation levels may have resulted in a SRT with less than 80% correct, as the normal hearing control group on average correctly identified only 63%

of the target words in simple sentences (NH elderly adults = 49% and elderly adults with HL = 24%). Therefore, one could hypothesize that the present effects might for example only appear when SNR levels are neither too high nor too low, e.g., when competition arises between the cognitive resources needed to cope with the noise and resources needed for syntactic processing (Meister *et al.* 2016). Further research into the effects of syntactic complexity in relation to masking release at varying levels of speech perception performance (for example comparing SRTs at 50% and 80%) is necessary to confirm this hypothesis.

An additional pattern arose in the present data: successful phoneme perception in syntactically complex sentences such as 'John repaired the roof that really leaked' was more likely than in canonical sentences as 'Mick has a bump on his head'. This differed across groups: both the normal hearing elderly adults and the elderly adults with hearing loss seemed to have more difficulty with the complex sentences when compared with the normal hearing control group. Yet, across all groups, phonemes in sentences with high syntactic complexity were more likely to be perceived correctly. These positive effects of syntactic complexity on speech perception in the present study are not entirely in line with previous studies in which canonical constructions have been found to be better understood than marked ones (Uslar *et al.* 2011, Wingfield *et al.* 2006, Carroll and Ruigendijk 2013). Although the types of the syntactic constructions under investigation in our study are not identical the ones used in English and German studies, the rather unexpected absence of a release-of-masking effect in syntactically complex sentences called for further inspection.

A possible additional factor influencing sentence perception may be found in the pragmatic domain. Regardless of the syntactic structure of the sentence, some elements may stand out in running speech thanks to specific pragmatic cues. As a matter of fact, words that are associated with new or contrastive information in the sentence often carry prosodic stress (Selkirk 1984, 1995). Such prosodic marking is implemented phonetically be means of changes in fundamental frequency (F_0), duration and intensity of the speech signal. Fluctuations in F_0 have traditionally been described as a primary cue for prosodic prominence in many languages, including Dutch (Rietveld and van Heuven 2013). Elements carrying a pitch accent are thought to aid listeners in speech perception by directing attention to the message's critical sections (Cutler and Foss 1977). Relative clauses are typically focus marked constructions, in which the antecedent carries prosodic stress.

In the present study, in the complex sentence condition the word containing the target phoneme was always the antecedent of the relative clause. Although the target words of these complex sentences were Root-Mean-Square (RMS) loudness balanced at 70 dB SPL and recorded with an intonation pattern carrying as little fluctuations as possible in F_0, there might have been phonetic remnants of a prosodic bias on the target word. To verify whether the unexpected positive effect of syntactic complexity on speech understanding could be related to unwanted pragmatic salience, we measured the minimum and maximum F_0 for the target words using the phonetics analysis program PRAAT (Boersma and Weenink 2016), comparing the absolute difference between minimum and maximum F_0 values of each target word. Should prosodic markedness of the target noun relate to better outcomes in speech perception, it could be assumed that pragmatic salience outweighs syntactic complexity. This would imply that target words in relative clauses yield better repetition scores than those in simple sentences. Therefore, an independent-samples *t*-test was performed to compare the ΔF_0 of the target words in complex and simple sentences. Contrary to expectations, no significant difference in pragmatic prominence between the different sentence types was found, $t(22) = 0.742$, $p = .466$. These results thus indicated that the potential prosodic bias on the target words of the complex sentences did not cause the present results.

An alternative explanation is thus required. Some previous results for young adults seem to mirror the present pattern: such that subject relative clauses (SR) result in similar perception scores as topic–verb–object sentences (Coene *et al.* 2016) or better perception than SVO sentences, in fluctuating, and especially in stationary noise (Uslar *et al.* 2013: fig. 6). In the present study, influences from stimuli characteristics including sentence length, word frequency or prosody could not readily explain this effect. This leads us to conclude that the syntactic complexity of the sentence structure selected for the complex condition in the present study might not have been complex enough to elicit the hypothesized effects. During sentence processing, sentence elements are theorized to be activated in working memory until their function or role in relation to other units in the sentence is clarified. This means that processing demands increase as more items are added to working memory or an element needs to remain activated longer until uncertainty is resolved (Gibson 1998). In the present study, right-branching subject relative clauses, as 'Ze brak de *pot* die buiten stond' (She broke the *pot* that stood outside) were selected for the syntactically complex condition, where the word containing the target phoneme was always the antecedent of the relative clause. As suggested by an anonymous reviewer, this implies that the actual processing complexity might occur only *after* the relativized noun.

As was shown in previous research (Carroll and Ruigendijk 2013, Uslar *et al.* 2013) the syntactic complexity effects on sentence perception are likely more pronounced in centre embedded subject relative clauses as 'De *pot* die buiten stond, is gebroken' (The pot that stood outside is broken)—or centre embedded object relative clauses as 'De *pot* die ik kocht, is gebroken' (The *pot* that I bought is broken). Although a right-branching subject relative clause is more syntactically complex than an utterance exhibiting a SVC order, the difference in complexity between these two sentence types may not have been large enough. This may have influenced the results and only elicited the hypothesized effect in the just challenging enough conditions: when participants perceived enough of the speech stimulus for the sentence structure to interfere with participants' predictions. In a similar vein, the observed differences between previous studies and ours might also relate to the fact that this study was concerned with the identification of word-initial phonemes embedded in sentence contexts instead of sentence understanding of syntactic structures. As the present results are based on a relatively small sample and a limited set of stimuli and sentence types, we believe that further research to identify additional factors that might explain the observed variation in outcomes.

Conclusions

In sum, previous research has indicated that speech perception in elderly adults is not only influenced by hearing loss, but also by declines in different cognitive abilities, background noise, and syntactic complexity of the speech material. The present research aimed to gain further insight in how these different factors interact and influence speech perception in elderly adults. We investigated how inhibitory control, age, hearing acuity and syntactic complexity influence the perception of phonemes in words and in sentences against different types of speech noise. The main findings were as follows:

- In line with the existing literature, elderly adults were significantly less likely to perceive correctly word-initial phonemes against a background of speech noise than younger normal hearing controls. When the words with the target phoneme were embedded in a sentence context, only elderly adults with hearing loss showed this effect.
- For elderly adults with hearing loss, speech perception accuracy was significantly related to inhibitory control. This indicated that those elderly participants with weakened inhibitory control were less likely to perceive speech correctly.
- Syntactic complexity of the test stimulus and/or the fluctuations of the speech noise masker during speech perception tasks significantly influenced the perception of word-initial phonemes in sentences. The influence of syntactic complexity was different across groups and noise types and therefore probably dependent on the difficulty of the listening condition. However, the exact impact of syntactic complexity remains unclear and needs to be further investigated.

The present results emphasize that speech perception is a complex, multifactorial process. When hearing loss is assessed in clinical settings, factors including inhibitory control and syntactic complexity may inadvertently influence speech perception outcomes. Clinicians should be aware of these influences. They can select assessment materials to avoid the influence of for example syntactic complexity or interpret results in which inhibitory control or syntactic complexity may play a role with these factors in mind. This in turn might help ensure the most optimal process for rehabilitation or treatment is chosen.

Acknowledgements

The authors thank all the participants for their participation in this study. Furthermore, they thank the reviewers for their suggestions and helpful comments made on an earlier version of the manuscript. The study received funding from the European Union's Seventh Framework Programme (FP7-PEOPLE-2012-IAPP) (grant agreement number 324401). ***Declaration of interest:*** The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

Note

1. Throughout, speech understanding is used to refer to human speech identification, i.e., the correct labeling of speech items by the listener, in contrast to speech comprehension, i.e., a deeper (semantic) understanding of the stimulus (Rietveld and van Heuven 2013). In agreement with the field of language technology, speech recognition will refer to non-human (automated) speech recognition.

References

Amieva, H., Ouvrard, C., Giulioli, C., Meillon, C., Rullier, L. and Dartigues, J. F., 2015, Self-reported hearing loss, hearing aids, and cognitive decline in elderly adults: a 25-year study. *Journal of the American Geriatrics Society*, **63(10)**, 2099–2104. http://doi.org/10.1111/jgs.13649

Benichov, J., Cox, L. C., Tun, P. A. and 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

Bialystok, E., Craik, F. I. M., Klein, R. and Viswanathan, M., 2004, Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychology and Aging*, **19(2)**, 290–303. http://doi.org/10.1037/0882-7974.19.2.290

Boersma, P. and Weenink, D., 2016, *Praat: Doing Phonetics by Computer* [computer program] (University of Amsterdam, Amsterdam).

Borella, E., Carretti, B. and De Beni, R., 2008, Working memory and inhibition across the adult life-span. *Acta Psychologica*, **128(1)**, 33–44.

Carroll, R. and Ruigendijk, E., 2013, The effects of syntactic complexity on processing sentences in noise. *Journal of Psycholinguistic Research*, **42(2)**, 139–159.

Clarke, J., Gaudrain, E., Chatterjee, M. and Başkent, D., 2014, T'ain't the way you say it, it's what you say: perceptual continuity of voice and top-down restoration of speech. *Hearing Research*, **315**, 80–87.

Coene, M., van der Lee, A. and Govaerts, P. J., 2015, Spoken word recognition errors in speech audiometry: a measure of hearing performance? *BioMed Research International*, **2015(932519)**, 1–8. http://doi.org/10.1155/2015/932519

Coene, M., Krijger, S., Meeuws, M., De Ceulaer, G. and Govaerts, P. J., 2016, Linguistic factors influencing speech audiometric assessment. *BioMed Research International*, **2016(7249848)**, 1–14. http://doi.org/10.1155/2016/7249848

Cook, M., 2006, A glimpsing model of speech perception in noise. *Journal of the Acoustical Society of America*, **119(3)**, 1562–1573.

Cutler, A. and Foss, D. J., 1977, On the role of sentence stress on sentence processing. *Language and Speech*, **20**, 1–10.

Dawes, P., Emsley, R., Cruickshanks, K. J., Moore, D. R., Fortnum, H., Edmondson-Jones, M., . . . Munro, K. J., 2015, Hearing loss and cognition: the role of hearing aids, social isolation and depression. *Plos One*, **10(3)**, e0119616. http://doi.org/10.1371/journal.pone.0119616

Festen, J. M. and Plomp, R., 1990, Effects of fluctuating noise and interfering speech on the speech reception threshold for impaired and normal hearing. *Journal of the Acoustical Society of America*, **88(4)**, 1725–1736.

Francart, T., van Wieringen, A. and Wouters, J., 2011, Comparison of fluctuating maskers for speech recognition tests. *International Journal of Audiology*, **50(1)**, 2–13.

Füllgrabe, C., Berthommier, F. and Lorenzi, C., 2006, Masking release for consonant features in temporally fluctuating background noise. *Hearing Research*, **211(1–2)**, 74–84.

George, E. L. J., Festen, J. M. and Houtgast, T., 2006, Factors affecting masking release for speech in modulated noise for normal hearing and hearing-impaired listeners. *Journal of the Acoustical Society of America*, **120(4)**, 2295.

George, E. L. J., Zekveld, A. A., Kramer, S. E., Goverts, S. T., Festen, J. M. and Houtgast, T., 2007, Auditory and non-auditory factors affecting speech reception in noise by older listeners. *Journal of the Acoustical Society of America*, **121(4)**, 2362–2375. http://doi.org/10.1121/1.2642072

Gibson, E., 1998, Linguistic complexity: locality of syntactic dependencies. *Cognition*, **68(1)**, 1–76.

Gordon-Salant, S. and Cole, S. S., 2016, Effects of age and working memory capacity on speech recognition performance in noise among listeners with normal hearing. *Ear and Hearing*, **37(5)**, 593–602.

Govaerts, P. J., Daemers, K., Yperman, M., De Beukelaer, C., De Saegher, G. and De Ceulaer, G., 2006, Auditory speech sounds evaluation (A§E®): a new test to assess detection, discrimination and identification in hearing impairment. *Cochlear Implants International*, **7(2)**, 92–106. http://doi.org/10.1002/cii.302

Hawkins, J., 1994, *A Performance Theory of Order and Constituency* (Cambridge, UK: Cambridge University Press).

Hawkins, J. A., 1994, *Efficiency and Complexity in Grammars* (Oxford: Oxford University Press).

Hawkins, J. A., 2014, *Cross-Linguistic Variation and Efficiency* (Oxford: Oxford University Press).

Houtgast, T. and Festen, J. M., 2008, On the auditory and cognitive functions that may explain an individual's elevation of the speech reception threshold in noise. *International Journal of Audiology*, **47(6)**, 287–295.

Kapteyn, T. S. and Lamoré, P. J., 2012, *Audiologieboek—7.2.6(2). Gehoor en leeftijd (presbyacusis)* (available at: http://www.audiologieboek.nl/htm/hfd7/7-2-6.htm).

Koelewijn, T., Zekveld, A. A., Festen, J. M., Rönnberg, J. and Kramer, S. E., 2012, Processing load induced by informational masking is related to linguistic abilities. *International Journal of Otolaryngology*, **2012(865731)**, 1–11. http://doi.org/10.1155/2012/865731

Lamoré, P. J. and Houtgast, T., 2010, *Audiologieboek—2.9.1(2). Spraakverstaan in aanwezigheid van stoorgeluid* (available at: http://www.audiologieboek.nl/htm/hfd2/2-9-1.htm#2917).

Lin, F. R., 2011, Hearing loss and cognition among older adults in the United States. *Journal of Gerontology*, **66(10)**, 1131–1136 (available at: http://doi.org/10.1093/gerona/glr115).

Lin, F. R., Metter, E. J., O'Brien, R. J., Resnick, S. M., Zonderman, A. B. and Ferrucci, L., 2011, Hearing loss and incident dementia. *Archives of Neurology*, **68(2)**, 214–220 (available at: http://doi.org/10.1001/archneurol.2010.362).

Lubbe, R. H. J. and van der Verleger, R., 2002, Aging and the Simon task. *Psychophysiology*, **39(1)**, 100–110.

Meister, H., Schreitmüller, S., Ortmann, M., Rahlmann, S. and Walger, M., 2016, Effects of hearing loss and cognitive load on speech recognition with competing talkers. *Front Psychol*, **7**, 301. https://doi.org/10.3389/fpsyg.2016.00301

Mueller, S. T., 2012, *PEBL: The Psychology Experiment Building Language* (available at: http://pebl.sourceforge.net) [computer program].

Newhart, H. and Reger, S. N. (eds), 1945, Syllabus of audiometric procedures in the administration of a program for the conservation of hearing of school children. *Transactions of the American Academy of Ophthalmology and Otolaryngology*, **April (Suppl. 1)**, 1–28.

Oostdijk, N., 2003, Het Corpus Gesproken Nederlands: veelzijdig onderzoeksinstrument voor o.a. taalkundig en taal- en spraaktechnologisch onderzoek. *Link*, **14(1)**, 3–6.

Pan, W., 2001, Akaike's information criterion in generalized estimating equations. *Biometrics*, **57(1)**, 120–125.

Pichora-Fuller, M. K., 2008, Use of supportive context by younger and older adult listeners: balancing bottom-up and top-down information processing. *International Journal of Audiology*, **47(Suppl. 2)**, S72–S82.

Rietveld, A. C. M. and van Heuven, V. J., 2013, *Algemene fonetiek*, 3rd edn (Bussum: Uitgeverij Coutinho).

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

Rönnberg, J., Rudner, M., Lunner, T. and Zekveld, A. A., 2010, When cognition kicks in: working memory and speech understanding in noise. *Noise and Health*, **12(49)**, 263–269. http://doi.org/10.4103/1463-1741.70505

Rönnberg, J. *et al.*, 2013, The ease of language understanding (ELU) model: theoretical, empirical, and clinical advances. *Frontiers in Systems Neuroscience*, **7**, 31. https://doi.org/10.3389/fnsys.2013.00031

Rudner, M., Lunner, T., Behrens, T., Sundewall Thorén, E. and Rönnberg, J., 2012, Working memory capacity may influence perceived effort during aided speech recognition in noise. *Journal of the American Academy of Audiology*, **23**, 577–589.

Selkirk, E., 1984, *Phonology and Syntax: The Relation between Sound and Structure* (Cambridge, MA: MIT Press).

Selkirk, E., 1995, Sentence prosody: intonation, stress, and phrasing. In J. Goldsmith (ed.), *Handbook of Phonological Theory* (Oxford: Basil Blackwell), pp. 550–569.

Simon, J. R. and Wolf, J. D., 1963, Choice reaction time as a function of angular stimulus–response correspondence and age. *Ergonomics*, **6**, 99–105.

Stenbäck, V., 2016, *Speech Masking Speech in Everyday Communication. The Role of Inhibitory Control and Working Memory Capacity*. Studies from the Swedish Institute for Disability Research No. 80, Linköping University Medical Dissertations No. 1559) (Linköping: Linköping University).

Sweeney, J., 2001, Inhibitory control of attention declines more than working memory during normal aging. *Neurobiology of Aging*, **22(1)**, 39–47.

Tun, P., Williams, V., Small, B. and Hafter, E., 2012, The effects of aging on auditory processing and cognition. *American Journal of Audiology*, **21(December)**, 344–351. http://doi.org/10.1044/1059-0889(2012/12-0030)b

Uslar, V. N., Ruigendijk, E., Hamann, C., Brand, T. and Kollmeier, B., 2011, How does linguistic complexity influence intelligibility in a German audiometric sentence intelligibility test? *International Journal of Audiology*, **50(9)**, 621–631. http://doi.org/10.3109/14992027.2011.582166

Uslar, V. N., Carroll, R., Hanke, M., Hamann, C., Ruigendijk, E., Brand, T. and Kollmeier, B., 2013, Development and evaluation of a linguistically and audiologically controlled sentence intelligibility test. *The Journal of the Acoustical Society of America*, **134(4)**, 3039–3056. http://doi.org/10.1121/1.4818760

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

World Health Organisation (WHO), 2017, *Deafness and Hearing Loss. Fact Sheet* (available at: http://www.who.int/mediacentre/factsheets/fs300/en/).

Zelazo, P. D., Craik, F. I. M. and Booth, L., 2004, Executive function across the life span. *Acta Psychologica*, **115(2–3)**, 167–183.