ASHA American Speech-Language-Hearing Association

JSLHR

Research Article

Informational Masking and Listening Effort in Speech Recognition in Noise: The Role of Working Memory Capacity and Inhibitory Control in Older Adults With and Without Hearing Impairment

Victoria Stenbäck,[a,b] Erik Marsja,[a] Mathias Hällgren,[c] Björn Lyxell,[a,d] and Birgitta Larsby[c]

[a] Disability Research Division, Department of Behavioural Sciences and Learning, Linköping University, Sweden [b] Division of Education, Teaching and Learning, Department of Behavioural Sciences and Learning, Linköping University, Sweden [c] Department of Otorhinolaryngology in Östergötland and Department of Biomedical and Clinical Sciences, Linköping University, Sweden [d] Department of Special Needs Education, University of Oslo, Norway

ARTICLE INFO

Article History:
Received December 22, 2021
Revision received June 3, 2022
Accepted July 19, 2022

Editor-in-Chief: Peggy B. Nelson
Editor: Yi Shen

https://doi.org/10.1044/2022_JSLHR-21-00674

ABSTRACT

Purpose: The study aimed to assess the relationship between (a) speech recognition in noise, mask type, working memory capacity (WMC), and inhibitory control and (b) self-rated listening effort, speech material, and mask type, in older adults with and without hearing impairment. It was of special interest to assess the relationship between WMC, inhibitory control, and speech recognition in noise when informational maskers masked target speech.
Method: A mixed design was used. A group (N = 24) of older (M_{age} = 69.7 years) individuals with hearing impairment and a group of age normal–hearing adults (M_{age} = 59.3 years, SD = 6.5) participated in the study. The participants were presented with auditory tests in a sound-attenuated room and with cognitive tests in a quiet office. The participants were asked to rate listening effort after being presented with energetic and informational background maskers in two different speech materials used in this study (i.e., Hearing In Noise Test and Hagerman test). Linear mixed-effects models were set up to assess the effect of the two different speech materials, energetic and informational maskers, hearing ability, WMC, inhibitory control, and self-rated listening effort.
Results: Results showed that WMC and inhibitory control were of importance for speech recognition in noise, even when controlling for pure-tone average 4 hearing thresholds and age, when the maskers were informational. Concerning listening effort, on the other hand, the results suggest that hearing ability, but not cognitive abilities, is important for self-rated listening effort in speech recognition in noise.
Conclusions: Speech-in-noise recognition is more dependent on WMC for older adults in informational maskers than in energetic maskers. Hearing ability is a stronger predictor than cognition for self-rated listening effort.
Supplemental Material: https://doi.org/10.23641/asha.21357648

Interpersonal communication involves two or more individuals intentionally transferring information among themselves. Successful communication in everyday situations depends on the individual's ability to hear, listen, and comprehend (Kiessling et al., 2003). According to Kiessling et al. (2003), hearing is a passive function allowing an individual to perceive sound and can be used to describe impairment. Listening involves intent and attention, extending the notion of hearing to a purposeful action. Comprehending is the endeavor of receiving the information or intention, before two-way communication can commence (Kiessling et al., 2003; Pichora-Fuller & Singh, 2006).

Sensorineural hearing loss (HL) is connected to deterioration in frequency discrimination, temporal processing, and perceiving speech in noise (Plomp, 1978). Sensorineural

---

Correspondence to Victoria Stenbäck: victoria.stenback@liu.se. ***Disclosure:*** *The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

hearing loss can be linked to deficits in the nerve pathways from the inner ear to the brain and/or the cochlea and loss of function of hair cells. This in turn can lead to difficulties perceiving sounds, especially in environments where background noise and/or speech is present (Committee on Hearing Bioacoustics and Biomechanics, 1988). According to the World Health Organization (2021), more than 1.5 billion people across the globe experience some degree of hearing loss (mild-to-complete HL) during their life span. It is approximated that at least 430 million of these people will require care (World Health Organization, 2021).

Various background noises carry different characteristics interfering with everyday communication. Background noise that shares similar acoustic energy as target speech can be considered *energetic*. Energetic maskers tax attentional resources and the ability to separate signal streams for successful recognition (Brungart, 2001; Durlach et al., 2003; Mattys et al., 2009, 2012; Schneider et al., 2007). When the background maskers contain information disrupting more central processes, further increasing cognitive load, it is referred to as informational maskers. According to, for example, Brungart (2001), Cooke and Lu (2010), Durlach et al. (2003), and Moore (2013), informational masking increases with similarities between target speech and background masker, making target speech more difficult to segregate. Cooke et al. (2008) describe that informational masking occurs when the effect of the energetic masker can be accounted for. Moreover, Cooke et al. further characterize three main consequences of informational masking: (a) when the masker competes with the target signal, increasing the difficulty to segregate the streams or selectively attend to the target signal; (b) lexical–semantic interference caused by, for example, a known language interfering with the target signal; and (c) limited processing resources causing an increase in cognitive load, making the main task more difficult to perform. According to Mattys et al. (2009), energetic masking can occur from informational masking, and vice versa, but informational masking can also occur alone.

Successful speech recognition is highly dependent on the signal-to-noise ratio (SNR). Speech recognition is also dependent on signal distortions by, for example, background noise, and/or an HI. Decreasing SNRs, background noise, a hearing impairment (HI). increase the amount of effort to recognize and comprehend a spoken message (Pichora-Fuller et al., 1995). Goy et al. (2013) also point out that the use of supportive contextual information can alleviate listening effort in adverse conditions and facilitate lexical access. Under adverse conditions, more cognitive resources, such as working memory (WM), inhibitory control, and lexical skills, are needed to piece incomplete information together (Koelewijn et al., 2015; Larsby et al., 2005, 2008, 2012; Lewis et al., 2021; Lunner & Sundewall-Thorén, 2007). Using more cognitive resources in adverse listening conditions leads to an increase in listening effort (e.g., Koelewijn et al., 2015). The Framework for Understanding Effortful Listening (FUEL; Pichora-Fuller et al., 2016) includes various components involved in effortful listening such as cognitive capacity, motivation of the listener to engage in the listening situation, and hearing difficulties.

The experienced difficulties with listening to speech in adverse conditions reported from older individuals with an HI might be caused by various factors, one being linked to worse hearing thresholds, and age- and hearing loss–related suprathreshold auditory processing deficits (Füllgrabe & Moore, 2018), whereas other factors may be linked to a general age-related decline in some cognitive skills (for a review, see Humes et al., 2012). Although there is evidence from individual studies that speech recognition in noise, especially for informational maskers, is relying on cognition (e.g., Michalek et al., 2018; Stenbäck et al., 2015, 2021), a review and meta-analysis by Füllgrabe and Rosen (2016) suggest that WM capacity (WMC) is not a strong predictor for young people with normal hearing.

Rönnberg et al. (2013, 2021) proposed the Ease of Language Understanding (ELU) model, where emphasis lies on the relevance of communicative skill and capacity theory for WM. The ELU model proposes WM as a central function for interpreting, using, and extracting explicit/implicit, meaningful information, where high WMC is beneficial for successful speech understanding (Rönnberg et al., 2013). The WM system describes how, when listening conditions are optimal, no effortful, conscious, explicit processes are engaged, and phonologically mediated lexical access is automatic. In contrast, a mismatch condition (i.e., when the speech signal is so distorted that it does not match long-term memory [LTM] representations) may occur due to various factors such as the context, talker, inaccurate phonological representations in LTM, and slow lexical access (Rönnberg et al., 2013). However, when the mismatch between incoming information and stored representations is too severe, cognitive abilities might not be enough to reach understanding (Rönnberg et al., 2019). The Rapid Automatic Multimodal Binding of Phonology (RAMBPHO) is a buffer mediating mismatch conditions by binding phonological and/or lexical information in LTM to stored representations. To elucidate the explicit and implicit WM processes during listening, Rönnberg et al. (2019) added postdiction and prediction to the model. Postdiction is involved when mismatch occurs when listening to speech, to some degree, and explicit, conscious WM functions are triggered to solve the mismatch. Prediction in the ELU involves priming and pretuning of RAMBPHO, engaging implicit, unconscious processes in WM. The prediction role of WM is to use phonological and semantic information to make predictions under adverse listening conditions. In communication situations, the postdiction and prediction states are involved interchangeably, co-occurring and cooperating for successful speech recognition (Rönnberg et al., 2019).

Stenbäck et al. (2015) found that performance in a test of inhibitory control, the Swedish Hayling task, significantly related to performance in a speech-in-noise task in young normally hearing (YNH) individuals. Moreover, Stenbäck et al. (2021) discuss that inhibitory control and WMC are closely linked and cooperate to suppress irrelevant information and store target information. Furthermore, Lewis et al. (2021) investigated relationships between neurocognitive skills, there among inhibition, and the identification of spectrally degraded speech in YNH individuals. They found that inhibition skills moderately correlated with degraded speech when presented auditorily, but not when presented audiovisually. Further research by Sommers and Danielsson (1999) investigated the importance of inhibitory control and semantic context on speech recognition and found that inhibitory control was especially important for older adults for speech recognition but that semantic context could alleviate some of the difficulties caused by decreasing inhibitory control.

In this study, speech-recognition-in-noise performance using the Hagerman test (Hagerman, 1982) and Swedish Hearing In Noise Test (HINT; Hällgren et al., 2006) was evaluated with four different maskers. Two of the maskers used were background noises with slight amplitude modulation, where one of the noises also contained temporal gaps but no temporal fine structure or comprehensible information (Francart et al., 2011); these maskers were considered energetic maskers. The other two maskers were considered informational and contained temporal gaps and temporal fine structure, and one of them also contained comprehensible information, increasing the similarities between masker and target speech (Arbogast et al., 2005; Francart et al., 2011). The speech materials were administered to older individuals with and without HI.

The aim of this study was twofold. First, we aimed to assess the relationship between speech-recognition-in-noise performance, masker type, WMC, and inhibitory control. Second, we aimed at examining the relationship between perceived listening effort, speech material, masker type, WMC, and inhibitory control. Older adults with age-normal hearing and with hearing impairment participated in the study. The current study extends the work in previous studies that address the relationship between inhibitory control and speech recognition in noise (e.g., Helfer & Freyman, 2014; Knight & Heinrich, 2017; Sommers & Danielson, 1999) and specifically investigate whether the role of inhibitory control varies depending on whether energetic or informational maskers are presented with the target speech. Regarding the aims, we hypothesized that performance in speech recognition in noise in different maskers would vary depending on WMC and inhibitory control. We further hypothesized that self-rated listening effort would differ depending on masker type, speech material, and individual WMC and inhibitory control.

Materials and Method

Participants

In this study, 24 older individuals between 59 and 80 years old (age: M = 69.8 years, SD = 5.4) with a mild-to-moderate HI participated (13 men, 11 women). Twenty-four older adults between 46 and 69 years old (age: M = 59.3 years, SD = 6.5, 4) with hearing thresholds within the normal range for their age (according to ISO 7029:2000, 2000; four men, 20 women) were also included (ISO 7029 specifies descriptive statistics for hearing thresholds for persons of various ages who are ontologically normal). The individuals with HI were recruited from the audiology clinic at Linköping University Hospital, using word-of-mouth, and by information letters to various organizations. The older adults with normal pure-tone hearing thresholds were recruited by word-of-mouth and social media. All participants signed a consent form containing all the information about the study design and aims and answered four questions regarding native language, education level, and self-reported neurological health. All participants reported Swedish as their native language and no neurological deficits and had self-reported normal or corrected-to-normal vision. The participants' names were coded, and they were given a number that could not be linked to a specific individual.

Procedure

The participants did not wear hearing aids during the auditory tests, as it was of importance to assess speech recognition performance and its relation to inhibitory control, WMC, and self-rated listening effort, in unaided conditions (i.e., without hearing aids). All participants using hearing aids were encouraged to wear hearing aids during the cognitive tests and when receiving information about the test procedure. Before the testing commenced, the test leader asked the participants to confirm that all information had been heard and understood. All tests were explained verbally by the test leader before each test started, and information regarding the cognitive tests was also provided written on a computer screen before testing. All auditory and cognitive tests started with test trials to familiarize the participants with the procedure. The target stimuli (the spoken sentences) were presented against the background maskers via an ECHO Audiofire 8 external PC soundcard at a fixed level of 70 dB SPL in a sound-attenuated room (CA Tegnér T-room model) via TDH39 headphones. The background noise was presented in an adaptive staircase procedure to obtain targeted speech recognition thresholds (SRTs). Both the Hagerman and HINT sentences were presented with a female voice. The Hayling task was administered in quiet, in the same sound-attenuated room, and also via headphones, whereas the reading span test was administered visually on a computer screen in a quiet office. During the reading span test, participants were seated approximately 50 cm from the computer screen.

Auditory Tests

Pure Tone Audiometry

Pure tone audiometry was performed in both ears separately for the frequencies of 125, 250, 500, 1000, 2000, 4000, 6000, and 8000 Hz. The average audiograms for the participants are shown in Figure 1.

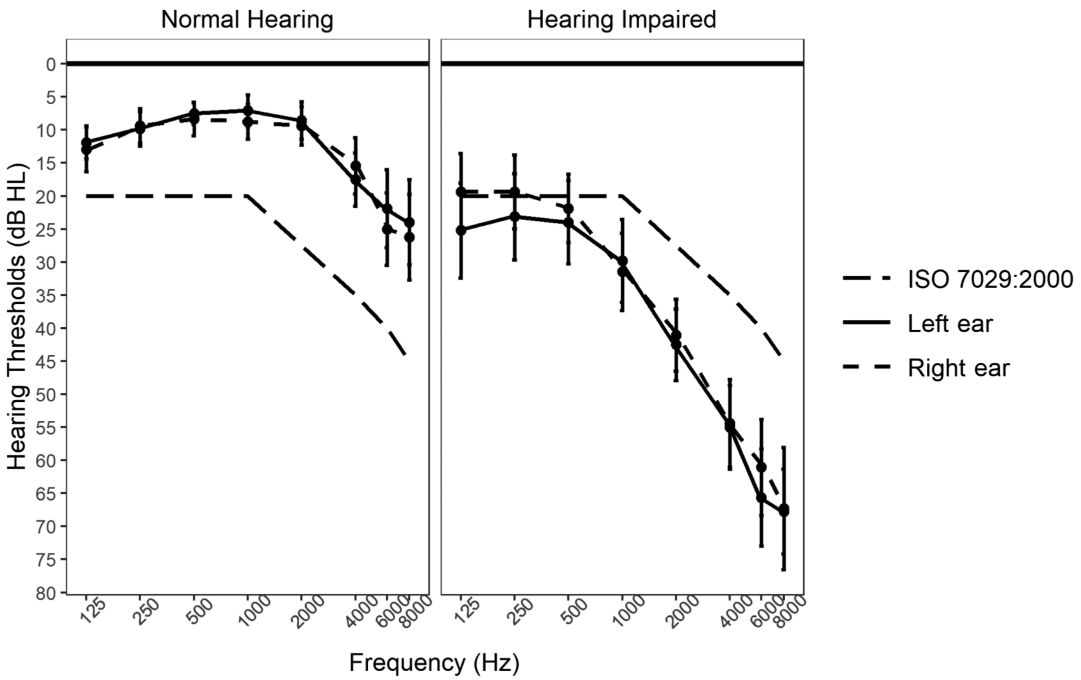
Hagerman Speech-in-Noise Test

The Hagerman speech-in-noise test (Hagerman, 1982) is a matrix-like sentence corpus with five-word sentences with the same grammatical structure. The material is composed of 11 lists, each with 10 sentences. The same 50 words are used in different combinations with the same word sequence: name – verb – number – adjective – noun (example: "Jonas has nine black baskets"). In this study, the sentences were presented in background noise in an adaptive staircase procedure targeting either SRTs for 50% speech recognition or for 80% speech recognition. The SNR either increased or decreased with 1 dB SNR depending on the number of correct repeated words. Two correct repeated words did not result in any increase or decrease in SNR. To minimize the risk of list effects, the lists were balanced so every list was presented an equal number of times in each background noise. The lists were also balanced so that half of the participants were presented with sentences targeting SRTs for 50% first, whereas the other half were presented with sentences targeting SRTs for 80% first. To familiarize the participants with the test procedure, they were presented with one list as a training session, which according to Hagerman and Kinnefors (1995) is enough to reduce any training effects.

Swedish HINT

The Swedish HINT sentences (Hällgren et al., 2006) are divided into 12 phonemically balanced lists, with a total of 250 sentences. The sentences consist of three to seven words (example: "Grandfather waxes the car") and are designed to be of higher semantic redundancy and more ecologically valid than the Hagerman sentences. In this study, the noise level was varied in an adaptive staircase procedure in steps of 2 dB to obtain an intelligibility

Figure 1. Audiograms for the elderly individuals with normally hearing, with the accepted hearing thresholds (dashed line) according to ISO 7029:2000 (2000) on the left and for the participants with hearing impairment on the right.



level targeting SRTs for 50% correct keywords or for 50% correct whole-sentence recognition (Larsby et al., 2012). Twenty sentences were presented before testing as a practice trial for the participant to get acquainted with the test procedure, as suggested by Hällgren et al. (2006). Within in each list, the sentences were presented in a randomized order and balanced so that half of the participants were presented with the speech recognition criterion for 50% correct keywords first and the other half with 50% correct whole-sentence recognition first.

Background Noises

In this study, four background noises, either considered energetic or informational, were used. The Hagerman sentences were presented in the Hagerman original, slightly (10%) amplitude modulated, stationary speech-shaped noise (SSN) with the same long-term average spectrum as the sentences. The HINT sentences were presented in the HINT original noise (also SSN; Hällgren et al., 2006), which was a random noise designed to match the long-term average spectrum of the HINT sentences. The additional three background noises for the two speech materials were the International Speech Test Signal (ISTS; see Holube et al., 2010, for the full creation of the speech test signal), which is composed of six women reading the same story in six different languages, where the stories are split into speech that constitute a single continuous speech stream, but only small fragments can be recognized as originating from a certain language or word, a single female talker reading a passage from a Swedish daily newspaper (FT), and a modulated noise constructed by modulating the HINT SSN with the envelope of the FT (for more detailed information about the method, see Stenbäck et al., 2021). In the current study, we were interested in the effects of informational versus energetic maskers. As research proposes (Brungart, 2001; Cooke & Lu, 2010; Durlach et al., 2003; Moore, 2013) that informational masking increases with similarities between target speech and masker, we considered the ISTS and the FT as informational maskers in the analyses, whereas the SSN and the modulated SSN from the FT were considered and analyzed as energetic maskers. The maskers were also counterbalanced to avoid order effects.

Cognitive Tests

Reading Span Test: Simultaneous Explicit Processing and Short-Term Storage

The reading span test was administered to all participants. The test was computerized and presented to the participants word-by-word on a computer screen. In this study, the test was composed of blocks of two to five sentences, each with three words. Each block had two trials. The subject's task was to, after each sentence had been presented, decide whether it made sense by pressing buttons corresponding to yes/no. After each block, the participants were asked to recall either the first or last word in each sentence. A correctly recalled word, regardless of recall order, scored 1 point, and the total score was used as a measure of WMC (Rönnberg et al., 1989).

The Hayling Task–Swedish Version: Initiation and Inhibition

The original version of the Hayling task (Burgess & Shallice, 1996) was developed to assess response inhibition in patients with frontal lobe lesions. The Swedish version of the task (Stenbäck et al., 2015, 2016) has been used as a measure of initiation and inhibitory control when assessing the importance of some cognitive abilities for successful speech recognition in noise. In the Swedish version, participants listen to three lists (conditions): Condition 1 (initiation), Condition 2, and Condition 3 (inhibition). There are 19 highly contextual sentences (sentences with high-cloze probability) in each list (e.g., "The captain wanted to sink with his…"), and the last word in each sentence is omitted. The subject's task in Condition 1 (initiation) is to, as quickly as possible, complete the sentence with the expected last word. In Condition 2 (inhibition), the subject is asked to finish new sentences with a grammatically correct but semantically nonrelated word. In Condition 3 (inhibition), the subject is presented with the sentences from Condition 1, but the instructions are the same as in Condition 2. Presenting the same sentences in Conditions 1 and 3 ensures that the final words have been previously activated and hence need to be actively suppressed (Borella et al., 2011; Stenbäck et al., 2015). In Conditions 2 and 3, an error score of zero is given if the words are semantically unrelated to the sentence or expected word, 1 point is given if it is semantically related (e.g., the opposite of the expected word), and 3 points are given if the sentence is completed with the expected word. The response on each sentence is recorded as a sound file in MATLAB. The response window starts 500 ms into the last word in the presented, unfinished sentences, therefore ensuring that the participants' responses are recorded even if they complete the sentence before the presentation of the sentence is finished (for a full review of the development of the Swedish Hayling task, see Stenbäck, 2016; Stenbäck et al., 2015). Response times (RTs) and error scores are used as a measure of inhibition by first subtracting the RTs in Condition 1 from the RTs in Conditions 2 and 3, respectively, to obtain a measure of "thinking time." The new RTs are divided by 19 to obtain the mean RT per sentence. The mean RTs are then divided with the total percentage of correct answers (total error score was 57, which means 57 corresponds to 0% correct answers). In this study, we used the values obtained by dividing the mean RT with the total percent correct answers from Condition 3 as we were interested to investigate inhibitory control for already activated words.

Perceived Listening Effort

The Borg CR-10 scale (Borg, 1982) was used to give the participants the opportunity to rate their perceived listening effort using numbers between 0 and 10 (or greater if necessary). The scale was primarily developed to register perceived physical effort but has also been used within hearing research (e.g., Hua et al., 2014; Kähäri, 2002). The scale combines number and category scaling by allowing the participants to use numbers that can be linked to a description of the perceived effort such as "nothing at all, weak, moderate, strong" (Borg, 1982). After each noise, the participants were asked to use the numbers or descriptions of perceived effort to rate how effortful they found the task of listening to, and repeating, the target sentences while ignoring/inhibiting the competing noise or speech.

Data Analysis

All data were preprocessed and analyzed using R statistical programming environment Version 4.1.1 (R Core Team, 2021). As some of the continuous variables were measured on different scales, they were scaled with zero as mean. To test the two hypotheses, we carried out linear mixed-effects modeling (LMEM) using the R package lme4 (Bates et al., 2015).

The models were fitted with a maximal random effects (RE) structure and justified by model comparison as well as the data to avoid the inflated risk of Type 1 errors in random intercept-only models (Barr et al., 2013). However, we only included theoretically sound effects, whether fixed or random, in our models as recommended by Matuschek et al. (2017). As a step in the aforementioned recommendations, we used likelihood ratio tests (Crainiceanu & Ruppert, 2004) to test the inclusion of REs and Satterthwaite's (1941) approximations of degrees of freedom to test the inclusion of fixed effects (FEs). These tests were carried out using the step() function in R package lmerTest (Kuznetsova et al., 2017). Statistical significance for the FEs in the final models was also calculated using Satterthwaite's approximations with lmerTest. For the first hypothesis, we fitted the LMEM using maximum likelihood to predict SNR with mask type (energetic vs. informational), reading span score, and inhibitory control. Moreover, to be able to see the effects of hearing versus cognition and of age versus cognition, pure-tone average 4 (PTA4; pure-tone average thresholds of 500, 1000, 2000, and 4000 Hz) and age were entered in the model as covariates. See Table 1 for an overview of the variables included in the models. The variance inflation factor (VIF) for multicollinearity was calculated for each of the variables considered for LMEM, using the performance package (Lüdecke, 2021). None of the included variables showed signs of multicollinearity (e.g., a VIF larger than 2.5; Johnston et al., 2018). In the models that we compared, we only included the random by-participant intercept and inhibitory control random by-participant intercept, as models using other REs did not converge due to overparameterization (i.e., the data did not support the RE structure). See Supplemental Material S1 for the results of the test of inclusion of the FEs. Concerning the inclusion of REs, the results from the LRTs showed that including the inhibitory control random by-participant slope was not significantly contributing to the model, $\chi^2(2) = 2.55$, $p = .28$. However, including the random by-participant intercept significantly contributed to the model, $\chi^2(1) = 271.85$, $p < .001$. This means that including participants as an RE ended up with a model significantly better than not including this RE.

Table 1. All variables considered for linear mixed-effects modeling (both hypotheses).

Variable	Variable type	Data type	Units/levels	In final model(s)
SNR	Dependent	Continuous	dB SNR	1
Effort	Dependent	Continuous	Self-rated	2
Material	Independent	Categorical	Hag or HINT^r	2
Mask type	Independent	Categorical	Informational or energetic^r	1, 2
Reading span	Independent	Continuous	Number of correctly recalled items	1
Inhibitory control	Independent	Continuous	RT/total percentage correct responses	1
Age	Independent	Continuous	Years	1, 2
PTA4	Independent	Continuous	dB HL	1, 2

Note. The superscript "^{r}" indicates reference in the model. "1" indicates the model testing the first hypothesis, and "2" indicates the model testing the second hypothesis. SNR = signal-to-noise ratio; Hag = Hagerman; HINT = Hearing In Noise Test; RT = response time; PTA4 = pure-tone average 4.

We used the same analytical strategy to investigate the second hypothesis (i.e., the maximal RE structure if the model[s] converge). Here, we fitted the LMEM using maximum likelihood to predict effort with mask type (energetic vs. informational), material (Hagerman vs. HINT), reading span score, and inhibitory control. The only RE included was a by-participant intercept. All other REs, for example, reading span, material, mask type, and inhibitory control random intercepts, resulted in models that failed to converge (i.e., the only RE included was by participant due to overparameterization). Again, we entered age and PTA4 as covariates in the model. The model comparisons (i.e., testing the significance of the different FEs) resulted in a final model only including mask type, material, PTA4, and age as FEs 27–29. See Supplemental Material S2 for model comparisons of the different models of interest (i.e., the inclusion of FEs).

In this article, we reported the results from the LMEMs following the guidelines of Meteyard and Davies (2020). The R scripts for preprocessing data, creating tables and figures, and data analysis can be found here: https://osf.io/76kp2/.

Results

WMC, Inhibitory Control, and Informational Maskers

To investigate the first hypothesis "that performance in speech-recognition-in-noise in different maskers would vary depending on WMC and inhibitory control," we fitted LMEM. The final model's total explanatory power is substantial (conditional $R^2 = .78$), and that of the part related to the FEs alone (marginal R^2) is .71.

The FE of mask type (informational) is statistically significant ($\beta = -1.30$, 95% confidence interval [CI] [−1.82, −0.78], $t(374) = -4.91$, $p < .001$). Moreover, the effect of reading span was statistically significant ($\beta = -0.63$, 95% CI [−1.17, −0.09], $t(378) = -2.28$, $p = .023$). The FE of age was statistically significant ($\beta = 0.09$, 95% CI [0.01, 0.18], $t(374) = 2.25$, $p = .025$). However, the effect of inhibitory control was not statistically significant ($\beta = -0.19$, 95% CI [−0.78, 0.40], $t(374) = -0.64$, $p = .525$). Finally, the interaction effect of Inhibitory Control × Mask Type (Informational) was statistically significant ($\beta = -.72$, 95% CI [−1.27, −0.17], $t(374) = -2.58$, $p = .010$). Concerning the covariates, the only interaction effect was found for PTA4 × Mask Type (Informational; $\beta = 2.94$, 95% CI [2.39, 3.49], $t(374) = 10.58$, $p < .001$). See Figure 2 for a visualization of the interaction effect. See Table 2 for parameter estimates, standard error, 95% CIs, and t and p values for the final model including the covariates age and PTA4.

Perceived Listening Effort, Mask Type, Material, WMC, and Inhibitory Control

To investigate the second hypothesis "that self-rated listening effort would differ depending on masker type, speech material, and individual WMC and inhibitory control," we fitted a second LMEM. The final model had

Figure 2. This figure shows the significant interaction between inhibitory control and signal-to-noise ratios (SNRs) in the two different masker types (energetic and informational). The error bars represent 95% confidence intervals.

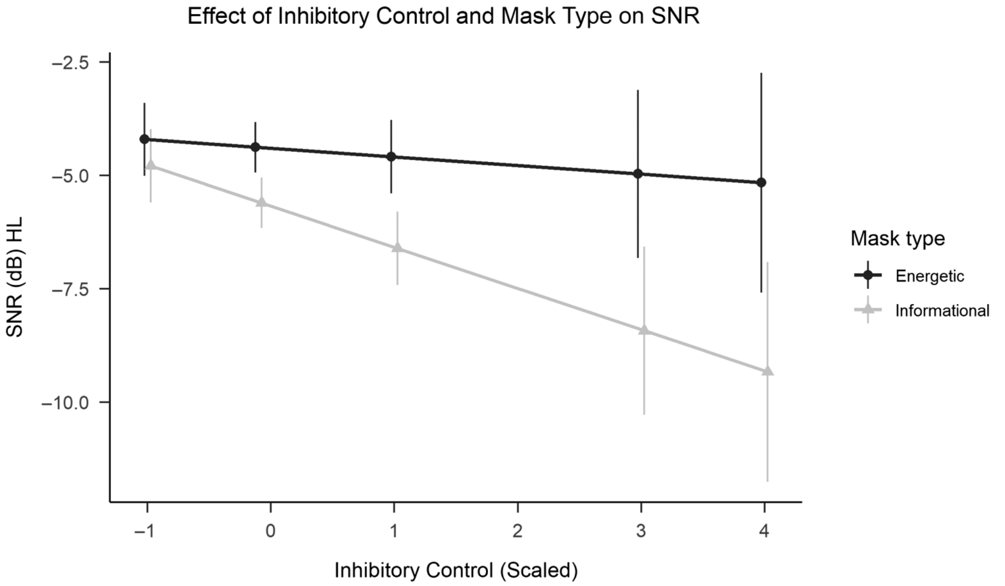


Table 2. Fixed effects estimates for signal-to-noise ratio (SNR) as the dependent variable.

	SNR				
Predictors	**Estimate**	***SE***	**CI**	***t***	***p***
Intercept	−10.45	2.71	−15.78, −5.13	−3.86	< .001
Mask type (informational)	−1.30	0.27	−1.82, −0.78	−4.91	< .001
Reading span (scaled)	−0.63	0.27	−1.17, −0.09	−2.28	.023
Age	0.09	0.04	0.01, 0.18	2.25	.025
Inhibitory control (scaled)	−0.19	0.30	−0.78, 0.40	−0.64	.525
PTA4 scaled	2.30	0.37	1.57, 3.04	6.15	< .001
Mask Type (Informational) × Inhibitory Control (Scaled)	−0.72	0.28	−1.27, −0.17	−2.58	.010
Mask Type (Informational) × PTA4 (Scaled)	2.94	0.28	2.39, 3.49	10.58	< .001
	Random effects				
	σ^2				6.74
	$\tau_{00\ \text{Participant}}$				2.12
	ICC				.24
	$N_{\text{Participant}}$				48
	Observations				384
	Marginal R^2/conditional R^2				.714/.782

Note. CI = confidence interval; PTA4 = pure-tone average 4; ICC = intraclass correlation.

substantial explanatory power (conditional R^2 = .51), and the marginal R^2 was .22. Concerning the FEs, the effect of mask type (informational) was statistically significant (β = 4.73, 95% CI [0.31, 9.16], t(370) = 2.10, p = .036), and the effect of material was statistically significant (β = 5.15, 95% CI [0.01, 10.29], t(370) = 1.97, p = .049). Moreover, the effect of PTA4 was statistically significant (β = 1.14, 95% CI [0.50, 1.78], t(370) = 3.48, p < .001), but the effect of age was not statistically significant (β = 0.08, 95% CI [−0.007, 0.17], t(370) = 1.81, p = .071). Concerning the interactions in the model, Material (HINT) × Mask Type (Informational) was statistically significant (β = −7.50, 95% CI [−13.74, −1.25], t(370) = −2.36, p = .019) as well as PTA4 × Material (HINT; β = −0.55, 95% CI [−1.04, −0.05], t(370) = −2.17, p = .031). However, Age × Mask Type (Informational) is statistically nonsignificant (β = −0.06, 95% CI [−0.13, 6.20e-03], t(370) = −1.79, p = .075) as well as Age × Material (HINT; β = −0.07, 95% CI [−0.14, 0.01], t(370) = −1.63, p = .105). However, the three-way interaction effect of Age × Mask Type (Informational) × Material (HINT) was statistically significant (β = 0.11, 95% CI [0.01, 0.21], t(370) = 2.25, p = .025). See Table 3 for parameter estimates and 95% CIs for the final model.

Table 3. Fixed effects estimates for effort as the dependent variable.

	Effort				
Predictors	**Estimate**	***SE***	**CI**	***t***	***p***
Intercept	0.24	2.93	−5.52, 6.00	0.08	.935
Mask type (informational)	4.73	2.25	0.31, 9.16	2.10	.036
Material (HINT)	5.15	2.61	0.01, 10.29	1.97	.049
PTA4 (scaled)	1.14	0.33	0.50, 1.78	3.48	.001
Age	0.08	0.05	−0.01, 0.17	1.81	.071
Mask Type (Informational) × Material (HINT)	−7.50	3.18	−13.74, −1.25	−2.36	.019
Material (HINT) × PTA4 (Scaled)	−0.55	0.25	−1.04, −0.05	−2.17	.031
Mask Type (Informational) × Age	−0.06	0.03	−0.13, 0.01	−1.79	.075
Material (HINT) × Age	−0.07	0.04	−0.14, 0.01	−1.63	.105
[Mask Type (Informational) × Material (HINT)] × Age	0.11	0.05	0.01, 0.21	2.25	.025
	Random effects				
	σ^2				3.51
	$\tau_{00\ \text{Participant}}$				2.11
	ICC				.38
	$N_{\text{Participant}}$				48
	Observations				382
	Marginal R^2/conditional R^2				.224/.515

Note. *SE* = standard error; CI = confidence interval; PTA4 = pure-tone average 4; HINT = Hearing In Noise Test; ICC = intraclass correlation.

Discussion

The aim of this study was to assess the relationship between (a) speech-recognition-in-noise performance, masker type, WMC, and inhibitory control and (b) self-rated listening effort, speech material, masker type, WMC, and inhibitory control. It was of special interest to assess inhibitory control when informational maskers were used to mask target sentences in speech recognition in noise, thereby extending the work from previous studies (e.g., Helfer & Freyman, 2014; Knight & Heinrich, 2017; Sommers & Danielson, 1999) addressing inhibitory control in relation to speech recognition in noise.

In this study, results showed that WMC and inhibitory control were related to speech-recognition-in-noise performance for informational maskers. The results also showed that type of masker, energetic or informational, influenced speech recognition. Further results showed that hearing ability, material, and mask type predicted self-rated listening effort.

WMC, Informational Maskers, and PTA4

Previous research (e.g., Larsby et al., 2012; Lunner & Sundewall-Thorén, 2007) has demonstrated that WMC is of importance for speech recognition in individuals with HI and that cognitive functioning is equally important for speech recognition for older individuals with normal hearing and older individuals with HI (Marsja et al., 2022). This study showed that individuals with lower WMC needed more favorable SNRs when listening to speech in informational maskers compared to listening in energetic maskers (see Figure 2). Furthermore, speech recognition in informational maskers was dependent on WMC, regardless of hearing ability (measured with PTA4). Although PTA4 significantly affected speech recognition in informational maskers, WMC was still an important factor (including PTA4 as a covariate resulted in a model with more explanatory power). This study found an effect of age showing that, with increasing age, more favorable SNRs were needed for successful speech recognition, as previously highlighted by, for example, Pichora-Fuller et al. (1995). The results also showed an interaction effect between inhibitory control and mask type, where better inhibitory control was beneficial when the masker was informational. According to the ELU model (Rönnberg et al., 2013, 2019, 2021), inhibitory control supports WM by inhibiting cue-activated mismatch and facilitating retrieval of information in semantic LTM. Research by Lewis et al. (2021) found that inhibition correlated with moderately degraded speech in an auditory modality. Further research, also supported by our results, by Sommers and Danielsson (1999) found that inhibitory control was especially important for older adults for speech recognition. Moreover, in broader contexts (i.e., not focusing on language understanding), it has been suggested that inhibitory control is, although a separate construct, subordinate to WM (e.g., Tiego et al., 2018; Vandierendonck, 2016). Other research, using factor analysis techniques, has shown when constructing a common factor that the inhibitory control factor explains no variance (Miyake & Friedman, 2012).

Stenbäck et al. (2021) found relationships between WMC and the Hagerman test in young individuals with normal hearing and older individuals with age-normal hearing. However, it is worth pointing out that, in a meta-analysis by Füllgrabe and Rosen (2016), WMC accounted for only 2% of the variance in YNH. In Stenbäck et al. (2021), age was significantly correlated with speech recognition in noise for the informational maskers in the Hagerman sentences, showing that SNRs increased with increasing age. This study found an effect of age on speech recognition in noise, although this study did not investigate differences between speech materials. In accordance with the ELU model proposed by Rönnberg et al. (2013, 2019), the individuals in this study with high WMC might have been able to use and extract meaningful information for successful sentence reconstruction in mismatch conditions more efficiently than the individuals with lower WMC. The interaction effect between inhibitory control and mask type could be interpreted as speech recognition being facilitated as interfering background speech could successfully be inhibited, thereby enabling retrieval of stored information in semantic LTM more efficiently.

Perceived Listening Effort, WMC, and Inhibitory Control

Under optimal listening conditions, lexical access is seemingly effortless, and no explicit processes are engaged (Rönnberg et al., 2013). If background noise or an HI is present, listening becomes more effortful and increases the demand on individual cognitive abilities such as WMC and inhibitory control (Larsby et al., 2012; Pichora-Fuller et al., 1995). The FUEL (Pichora-Fuller et al., 2016) states that cognitive abilities, hearing ability, and the individual's motivation to engage in an effortful listening situation play a role in how effortful it is perceived. In this study, type of masker and type of material affected listening effort. Depending on the structure of the speech material and the characteristics of the masker, greater mismatch may occur due to interference by competing speech on target speech. According to the ELU model (Rönnberg et al., 2013, 2019), the postdiction state triggers explicit WM processes to resolve a mismatch between signal and masker. In a material such as the Hagerman sentences, the structure is closed set and stylized with no semantic context. In mismatch conditions, cognitive abilities may not be enough to solve the conflict between target speech and competing masker, leading hearing ability to be the most important predictor for listening effort in speech recognition in noise in informational maskers (see Table 3). According to the ELU model, the prediction state makes use of phonological and semantic information to facilitate predictions in speech-in-noise recognition. The HINT being a material with semantic context and everyday sentences may offer cues to base predictions upon. The WM processes are faster and implicit, which facilitate retrieval from semantic LTM. However, although contextual information is offered in the HINT, informational maskers, contrary to energetic maskers, compete with the signal on a central level, increasing effort and cognitive load (Cooke et al., 2008). Cooke et al. (2008) describe that informational masking occurs when energetic masking has been accounted for and that there are three main consequences of informational masking, namely, competing signals, interference from known languages, and an increase in cognitive load. Even if there may be energetic masking from informational masking, or vice versa, increasing cognitive load, poorer hearing thresholds may lead to greater mismatch between target signal and competing masker, resulting in a mismatch too severe for cognitive abilities to resolve, thereby increasing effort. As previously discussed by, for example, Moore (2013) and Arbogast et al. (2005), informational masking increases with similarities between target speech and competing masker. Goy et al. (2013) also discuss that supportive contextual information may facilitate lexical access and alleviate listening effort. The contextual cues offered in the HINT may help mitigate listening effort, whereas no such contextual help is offered in the Hagerman test. The results in this study showed that hearing ability was important for listening effort, in part supporting the FUEL (Pichora-Fuller et al., 2016).

Conclusions

In conclusion, the individuals' capacity to process, store, and retrieve relevant information while inhibiting unwanted interference is related to successful speech recognition, only when informational maskers disrupt target speech. The individuals' hearing ability affects speech recognition; however, WMC is still a predictor for speech recognition in noise, and inhibitory control is important when maskers are informational. Listening effort is affected by hearing ability, type of masker, and the characteristics of the speech material.

Author Contributions

Victoria Stenbäck: Data curation (Lead), Formal analysis (Supporting), Investigation (Lead), Methodology (Lead), Writing – original draft (Lead). **Erik Marsja:** Data curation (Equal), Formal analysis (Lead), Writing – original draft (Equal). **Mathias Hällgren:** Conceptualization (Supporting), Methodology (Supporting), Supervision (Supporting), Writing – original draft (Supporting). **Björn Lyxell:** Conceptualization (Supporting), Supervision (Supporting), Writing – original draft (Supporting). **Birgitta Larsby:** Conceptualization (Lead), Funding acquisition (Lead), Methodology (Supporting), Supervision (Lead), Writing – original draft (Supporting).

Ethical Approval

The regional ethics committee in Linköping has approved this project (Dnr: 2012/105-31).

Data Availability Statement

Data from the current study are available on request from the authors. The R scripts for preprocessing data, creating tables and figures, and data analysis can be found here: https://osf.io/76kp2/.

Acknowledgments

This research was funded by the Swedish Research Council, Project 421-2009-1753. The authors would like to thank Niklas Rönnberg for invaluable advice and help. They would also like to specially thank all the participants for taking the time and energy to participate and for showing interest in this project.

References

Arbogast, T. L., Mason, C. R. & Kidd, G., Jr. (2005). The effect of spatial separation on informational masking of speech in normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America, 117*(4), 2169–2180. https://doi.org/10.1121/1.1861598

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

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

Borella, E., Ludwig, C., Fagot, D., & De Ribaupierre, A. (2011). The effect of age and individual differences in attentional control: A sample case using The Hayling test. *Archives of Gerontology and Geriatrics, 53*(1), e75–e80. https://doi.org/10.1016/j.archger.2010.11.005

Borg, G. A. V. (1982). Psychophysical bases of perceived exertion. *Medicine and Science in Sports and Exercise, 14*(5), 377–381. https://doi.org/10.1249/00005768-198205000-00012

Brungart, D. S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *The Journal of the Acoustical Society of America, 109*(3), 1101–1109. https://doi.org/10.1121/1.1345696

Burgess, P. W., & Shallice, T. (1996). Response suppression, initiation and strategy use following frontal lobe lesions. *Neuropsychologia, 34*(4), 263–272. https://doi.org/10.1016/0028-3932(95)00104-2

Committee on Hearing Bioacoustics and Biomechanics. (1988). Speech understanding and aging. *The Journal of the Acoustical Society of America, 83*(3), 859–895. https://doi.org/10.1121/1.395965

Cooke, M., Garcia Lecumberri, M. L., & Barker, J. (2008). The foreign language cocktail party problem: Energetic and informational masking effects in non-native speech perception. *The Journal of the Acoustical Society of America, 123*(1), 414–427. https://doi.org/10.1121/1.2804952

Cooke, M., & Lu, Y. (2010). Spectral and temporal changes to speech produced in the presence of energetic and informational maskers. *The Journal of the Acoustical Society of America, 128*(4), 2059–2069. https://doi.org/10.1121/1.3478775

Crainiceanu, C. M., & Ruppert, D. (2004). Likelihood ratio tests in linear mixed models with one variance component. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 66*(1), 165–185. https://doi.org/10.1111/j.1467-9868.2004.00438.x

Durlach, N. I., Mason, C. R., Kidd, G., Jr., Arbogast, T. L., Colburn, H. S., & Shinn-Cunningham, B. (2003). Note on informational masking. *The Journal of the Acoustical Society of America, 113*(6), 2984–2987. https://doi.org/10.1121/1.1570435

Francart, T., van Wieringen, A., & Wouters, J. (2011). Comparison of fluctuating maskers for speech recognition tests. *International Journal of Audiology, 50*(1), 2–13. https://doi.org/10.3109/14992027.2010.505582, 1

Füllgrabe, C., & Moore, B. C. J. (2018). The association between the processing of binaural temporal-fine-structure information and audiometric threshold and age: A meta-analysis. *Trends in Hearing, 22,* 2331216518797259. https://doi.org/10.1177/2331216518797259

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

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

Hagerman, B. (1982). Sentences for testing speech intelligibility in noise. *Scandinavian Audiology, 11*(2), 79–87. https://doi.org/10.3109/01050398209076203

Hagerman, B., & Kinnefors, C. (1995). Efficient adaptive methods for measuring speech reception threshold in quiet and in noise. *Scandinavian Audiology, 24*(1), 71–77. https://doi.org/10.3109/01050399509042213

Hällgren, M., Larsby, B., & Arlinger, S. (2006). A Swedish version of the Hearing In Noise Test (HINT) for measurement of speech recognition. *International Journal of Audiology, 45*(4), 227–237. https://doi.org/10.1080/14992020500429583

Helfer, K. S., & Freyman, R. L. (2014). Stimulus and listener factors affecting age-related changes in competing speech perception. *The Journal of the Acoustical Society of America, 136*(2), 748–759. https://doi.org/10.1121/1.4887463

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

Hua, H., Emilsson, M., Ellis, R. J., Widén, S., Möller, C., & Lyxell, B. (2014). Cognitive skills and the effect of noise on perceived effort in employees with aided hearing impairment and normal hearing. *Noise & Health, 16*(69), 79–88. https://doi.org/10.4103/1463-1741.132085

Humes, L. E., Dubno, R. J., Gordon-Salant, S., Lister, J. J., Cacace, A. T., Cruickshanks, K. J. K., Gates, G. A., Wilson, R. H., & Wingfield, A. (2012). Central presbycusis: A review and evaluation of the evidence. *The Journal of the Acoustical Society of America, 23*(08), 635–666. https://doi.org/10.3766/jaaa.23.8.5

Johnston, R., Jones, K., & Manley, D. (2018). Confounding and collinearity in regression analysis: A cautionary tale and an alternative procedure, illustrated by studies of British voting behaviour. *Quality & Quantity, 52*(4), 1957–1976. https://doi.org/10.1007/s11135-017-0584-6

Kähäri, K. R. (2002). *The influence of music on hearing. A study in classical and rock/jazz musicians* [Unpublished dissertation]. Göteborg University.

Kiessling, J., Pichora-Fuller, M. K., Gatehouse, S., Stephens, D., Arlinger, S., Chisolm, T., Davis, A. C., Erber, N. P., Hickson, L., Holmes, A., Rosenhall, U., & von Wedel, H. (2003). Candidature for and delivery of audiological services: Special needs of older people. *International Journal of Audiology, 42*(sup2), 92–101. https://doi.org/10.3109/14992020309074650

Knight, S., & Heinrich, A. (2017). Different measures of auditory and visual Stroop interference and their relationship to speech intelligibility in noise. *Frontiers in Psychology, 8,* 230. https://doi.org/10.3389/fpsyg.2017.00230

Koelewijn, T., de Kluiver, H., Shinn-Cunningham, G. S., Zekveld, A. A., & Kramer, E. S. (2015). The pupil response reveals increased listening effort when it is difficult to focus attention. *Hearing Research, 323,* 81–90. https://doi.org/10.1016/j.heares.2015.02.004

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTestPackage: Tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Larsby, B., Hällgren, M., & Lyxell, B. (2008). The interference of different background noises on speech processing in elderly hearing impaired subjects. *International Journal of Audiology, 47*(sup2), S83–S90. https://doi.org/10.1080/14992020802301159

Larsby, B., Hällgren, M., & Lyxell, B. (2012). Working memory capacity and lexical access in speech-recognition-in-noise. In T. Dau, M. L. Jepsen, J. C. Dalsgaard, & T. Poulsen (Eds.), *Proceedings of ISAAR 2011: Speech Perception and Auditory Disorders. 3rd International Symposium on Auditory and Audiological Research; August 2011, Nyborg, Denmark*. The Danavox Jubilee Foundation 2012.

Larsby, B., Hällgren, M., Lyxell, B., & Arlinger, S. (2005). Cognitive performance and perceived effort in speech processing tasks: Effects of different noise backgrounds in normal-hearing and hearing-impaired subjects. *International Journal of Audiology, 44*(3), 131–143. https://doi.org/10.1080/14992020500057244

Lewis, J. H., Castellanos, I., & Moberly, A. C. (2021). The impact of neurocognitive skills on recognition of spectrally degraded sentences. *Journal of the American Academy of Audiology, 32*(8), 528–536. https://doi.org/10.1055/s-0041-1732438

Lunner, T., & Sundewall-Thorén, E. (2007). Interactions between cognition, compression, and listening conditions: Effects on speech-in-noise performance in a two-channel hearing aid. *Journal of the American Academy of Audiology, 18*(07), 604–617. https://doi.org/10.3766/jaaa.18.7.7

Lüdecke, D. (2021). *sjPlot: Data visualization for statistics in social science*. https://cran.r-project.org/package=sjPlot

Marsja, E., Stenbäck, V., Moradi, S., Danielsson, H., & Rönnberg, J. (2022). Is having hearing loss fundamentally different? Multigroup structural equation modeling of the effect of cognitive functioning on speech identification. *Ear and Hearing, 43*(5), 1437–1446. https://doi.org/10.1097/AUD.0000000000001196

Mattys, S. L., Brooks, J., & Cooke, M. (2009). Recognizing speech under a processing load: Dissociating energetic from informational factors. *Cognitive Psychology, 59*(3), 203–243. https://doi.org/10.1016/j.cogpsych.2009.04.001

Mattys, S. L., Davis, M. H., Bradlow, A. R., & Scott, S. K. (2012). Speech recognition in adverse conditions: A review. *Language and Cognitive Processes, 27*(7–8), 953–978. https://doi.org/10.1080/01690965.2012.705006

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language, 94,* 305–315. https://doi.org/10.1016/j.jml.2017.01.001

Meteyard, L., & Davies, R. A. I. (2020). Best practice guidance for linear mixed-effects models in psychological science. *Journal of Memory and Language, 112,* 104092. https://doi.org/10.1016/j.jml.2020.104092

Michalek, A. M. P., Ash, I., & Schwartz, K. (2018). The independence of working memory capacity and audiovisual cues when listening in noise. *Scandinavian Journal of Psychology, 59*(6), 578–585. https://doi.org/10.1111/sjop.12480

Miyake, A., & Friedman N. P. (2012). *The nature and organization of individual differences in executive functions: Four general conclusions. Current Directions in Psychological Science, 21*(1):8–14. https://doi.org/10.1177/0963721411429458

Moore, B. C. J. (2013). *An introduction to the psychology of hearing*. Emerald Group Publishing Limited.

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. S., Tremblay, K. L., & Wingfield, A. (2016). Hearing impairment and cognitive energy: The Framework for Understanding Effortful Listening (FUEL). *Ear and Hearing, 37*(1), 5S–27S. https://doi.org/10.1097/AUD.0000000000000312

Pichora-Fuller, M. K., Schneider, A. B., & Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *The Journal of the Acoustical Society of America, 97*(1), 593–608. https://doi.org/10.1121/1.412282

Pichora-Fuller, M. K., & Singh, G. (2006). Effects of age on auditory and cognitive processing: Implications for hearing aid fitting and audiologic rehabilitation. *Trends in Amplification, 10*(1), 29–59. https://doi.org/10.1177/108471380601000103

Plomp, R. (1978). Auditory handicap of hearing impairment and the limited benefit of hearing aids. *The Journal of the Acoustical Society of America, 63*(2), 533–549. https://doi.org/10.1121/1.381753

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Rönnberg, J., Arlinger, S., Lyxell, B., & Kinnefors, C. (1989). Visual evoked potentials. *Journal of Speech and Hearing Research, 32*(4), 725–735. https://doi.org/10.1044/jshr.3204.725

Rönnberg, J., Holmer, E., & Rudner, M. (2019). Cognitive hearing science and ease of language understanding. *International Journal of Audiology, 58*(5), 247–261. https://doi.org/10.1177/1084713811409762

Rönnberg, J., Holmer, E., & Rudner, M. (2021). Cognitive hearing science: Three memory systems, two approaches, and the Ease of Language Understanding model. *Journal of Speech, Language, and Hearing Research, 64*(2), 359–370. https://doi.org/10.1044/2020_jslhr-20-00007

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., Dahlström, Ö., 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*(31), 1–17. https://doi.org/10.3389/fnsys.2013.00031

Satterthwaite, F. E. (1941). Synthesis of variance. *Psychometrika, 6*(5), 309–316. https://doi.org/10.1007/BF02288586

Schneider, B. A., Li, L., & Daneman, M. (2007). How competing speech interferes with speech comprehension in everyday listening situations. *Journal of the American Academy of Audiology, 18*(7), 559–572. https://doi.org/10.3766/jaaa.18.7.4

Sommers, M. S., & Danielsson, S. M. (1999). Inhibitory processes and spoken word recognition in young and older adults: The interaction of lexical competition and semantic context. *Psychology and Aging, 14*(3), 458–472. https://doi.org/10.1037/0882-7974.14.3.458

Stenbäck, V. (2016). *Speech masking speech in everyday communication—The role of inhibitory control and working memory capacity*. LiU-tryck.

Stenbäck, V., Hällgren, M., & Larsby, B. (2016). Executive functions and working memory capacity in speech communication under adverse conditions. *Speech, Language and Hearing, 19*(4), 218–226. https://doi.org/10.1080/2050571X.2016.1196034

Stenbäck, V., Hällgren, M., Lyxell, B., & Larsby, B. (2015). The Swedish Hayling Task, and its relation to working memory, verbal ability, and speech-recognition-in-noise. *Scandinavian Journal of Psychology, 56*(3), 264–272. https://doi.org/10.1111/sjop.12206

Stenbäck, V., Marsja, E., Hällgren, M., Lyxell, B., & Larsby, B. (2021). The contribution of age, working memory capacity, and inhibitory control on speech recognition in noise in young and older adult listeners. *Journal of Speech, Language, and Hearing Research, 64*(11), 4513–4523. https://doi.org/10.1044/2021_JSLHR-20-00251

Tiego, J., Testa, R., Bellgrove, M. A., Pantelis, C., & Whittle, S. (2018). A hierarchical model of inhibitory control. *Frontiers in Psychology, 9,* 1–25. https://doi.org/10.3389/fpsyg.2018.01339

Vandierendonck, A. (2016). A working memory system with distributed executive control. *Perspectives on Psychological Science, 11*(1), 74–100. https://doi.org/10.1177/1745691615596790

World Health Organization. (2021). *World report on hearing*. https://www.who.int/publications/i/item/world-report-on-hearing