Semantic Context Improves Speech Intelligibility and Reduces Listening Effort for Listeners with Hearing Impairment

Emma Holmes[1], Paula Folkeard[2], Ingrid S. Johnsrude[1,3], and Susan Scollie[2,3]

[1]Brain and Mind Institute, University of Western Ontario, London ON, Canada;

[2]National Centre for Audiology, University of Western Ontario, London ON, Canada;

[3]School of Communication Sciences and Disorders, University of Western Ontario, London ON, Canada

Keywords: psychoacoustics/hearing science; behavioural measures; hearing aids; speech perception

Abbreviations: s – seconds; dB – decibels; $\bar{x}$ – mean; *SD* – standard deviation; CST – Connected Speech Test; SNR – signal-to-noise ratio; RAU – rationalized arcsine units; ANOVA – analysis of variance; fMRI – functional magnetic resonance imaging.

Address correspondence to Emma Holmes, Wellcome Trust Centre for Neuroimaging, Institute of Neurology, University College London, 12 Queen Square, London WC1N 3BG, United Kingdom. Phone: (+1) 519 709-4499. E-mail: emma.holmes@ucl.ac.uk

Accepted Manuscript

Abstract

Objective: We investigated whether speech intelligibility and listening effort for hearing-aid users is affected by semantic context and hearing-aid setting.

Design: Participants heard target sentences spoken in a reverberant background of cafeteria noise and competing speech. Participants reported each sentence verbally. Eight participants also rated listening effort after each sentence. Sentence topic was either the same as, or different from, the previous target sentence.

Study sample: Twenty participants with sensorineural hearing loss were fit binaurally with Signia receiver-in-the-canal hearing aids. Participants performed the task twice: once using the hearing aid's omnidirectional setting and once using the 'Reverberant Room' setting, designed to aid listening in reverberant environments.

Results: Participants achieved better speech intelligibility for same-topic than different-topic sentences, and when they used the 'Reverberant Room' than the omnidirectional hearing-aid setting. Participants who rated effort showed a reliable reduction in listening effort for same-topic sentences and for the 'Reverberant Room' hearing-aid setting. The improvement in speech intelligibility from semantic context (i.e. same-topic compared to different-topic sentences) was greater than the improvement gained from changing hearing-aid setting.

Conclusions: These findings highlight the enormous potential of cognitive (specifically, semantic) factors for improving speech intelligibility and reducing perceived listening effort in noise for hearing-aid users.

Accepted Manuscript

Semantic Context Improves Speech Intelligibility and Reduces Listening Effort for Listeners with Hearing Impairment

Introduction

Listeners are frequently required to understand speech in reverberant environments and in the presence of multiple simultaneous talkers—a feat that is particularly challenging and effortful for people with hearing loss (e.g., Gatehouse & Noble, 2004). Reducing listening effort in these environments could potentially help improve quality of life in people with hearing impairment: the amount of effort experienced during conversation correlates positively with self-rated handicap (Gatehouse & Noble, 2004) and excessive listening effort can lead to social withdrawal and, eventually, isolation in people with hearing impairment (Gatehouse & Akeroyd, 2006; Noble & Gatehouse, 2006). Although a variety of non-acoustic factors are known to improve speech intelligibility in noise for listeners with normal hearing—such as semantic context (e.g., Dubno et al, 2000) and congruent visual lip movement information (e.g., Fraser et al, 2010)—the extent to which listeners with hearing impairment can use similar factors to reduce perceived listening effort is not fully understood. Here, we investigated whether maintaining a consistent topic across sentences improves accuracy and reduces subjective ratings of listening effort when hearing-aid users listen to speech in a reverberant background of competing talkers.

Anecdotally, many people with hearing impairment report that they find listening extremely effortful (e.g., Kramer et al, 2006). Listening can be effortful even for people with normal hearing when they listen in environments that are acoustically or cognitively demanding (for a review, see Mattys et al, 2012). For people with hearing loss, an additional source of effort originates from degradation of the acoustic signal at the auditory periphery (Mattys et al, 2012; Pichora-Fuller et al, 2016). A variety of different methods have been used to measure listening effort in previous studies, including self-reports of perceived effort,

Accepted Manuscript

pupil dilation measures (obtained using pupillometry), and performance on a secondary task (for a review, see McGarrigle et al, 2014).

There is some evidence that using hearing aids can reduce effort for listeners with hearing impairment (e.g., Noble & Gatehouse, 2006). However, it is unclear whether different hearing-aid settings—for example, directional compared to omnidirectional settings or different signal processing algorithms—affect listening effort in people with hearing impairment. When directional microphones were compared to omnidirectional microphones, one study found better performance on a secondary visual tracking task but no difference in self-reported listening effort (Desjardins, 2016), a different study found faster reaction times on a secondary visual task, but only for younger and not for older adults with hearing impairment (Wu et al, 2014), and another study found no differences in visual reaction times or self-reported listening effort (Hornsby, 2013). Some studies found a reduction in perceived effort with digital noise reduction algorithms (e.g., Bentler et al, 2008; Brons et al, 2013), although others found no difference (Desjardins & Doherty, 2014; Desjardins, 2016; Hornsby, 2013); in addition, some studies found better performance on a secondary visual tracking task with noise reduction (Neher et al, 2014; Desjardins & Doherty, 2014), whereas another found no difference (Desjardins, 2016). Thus, whether different hearing-aid settings affect listening effort is unclear.

Listening to speech in reverberant environments is particularly challenging for people with hearing impairment (e.g., Marrone et al, 2008). Recently developed hearing-aid programs have been designed specifically for reverberant environments (e.g. the 'Reverberant Room' setting on the Signia Pure™ primax devices, Sivantos Inc.) and could possibly reduce listening effort for hearing impaired people in reverberant settings. Given that many people with hearing impairment often report listening to be effortful and fatiguing, even when they

Accepted Manuscript

use hearing aids (e.g., Kramer et al, 2006), identifying novel factors that reduce listening effort could potentially improve their quality of life.

Recently, there has also been increased interest in cognitive factors related to hearing loss; utilizing cognitive factors (i.e. cues that are unrelated to the acoustic composition of the speech signal) could be a promising direction for reducing listening effort in people with hearing impairment. Listeners with normal hearing can use non-acoustic cues—such as congruent visual lip movement information (Sumby & Pollack, 1954), prior knowledge of talker characteristics (e.g., Johnsrude et al, 2013), and contextual information (e.g., Dubno et al, 2000; Davis et al, 2011)—to improve speech intelligibility. Listeners with hearing impairment can also use contextual (Pichora-Fuller et al, 1995; Desjardins & Doherty, 2014) and visual lip movement (e.g., Winn et al, 2013) information to improve speech intelligibility. For example, Pichora-Fuller *et al.* (1995) demonstrated that older adults with presbycusis reported words at the end of sentences more accurately when the final word was predictable from the preceding sentence context (e.g., 'The witness took a solemn oath') than when it was unpredictable (e.g., 'John hadn't discussed the oath').

Whether similar cognitive factors affect listening effort in listeners with hearing impairment is unclear. Within-sentence linguistic content has been found to reduce listening effort, measured using pupillometry, for cochlear-implant users (Winn, 2016). However, a study investigating within-sentence linguistic content in hearing-aid users found no effect of linguistic content on self-reported listening effort or performance on a secondary visual tracking task (Desjardins & Doherty, 2014). Also, performance on a secondary visual tracking task was found to be similar when participants with normal hearing and participants with mild-to-moderate hearing impairment recalled semantically-related and semantically-unrelated word lists (Tun et al, 2009). These discrepant findings could be explained by differences in how hearing-aid and cochlear-implant users utilise linguistic content or

Accepted Manuscript

different sensitivities of the listening effort measures used in these studies. Nevertheless, contextual information might be expected to reduce effort for listeners who use hearing aids, given that it improves *speech intelligibility* for listeners who use hearing aids (Desjardins & Doherty, 2014). Most previous studies that have investigated perceived effort have asked participants to rate effort across a block of trials (e.g., Mackersie & Cones, 2011; Fraser et al, 2010; Rudner et al, 2012). However, the current experiment asked participants to rate listening effort on a sentence-by-sentence basis. An advantage of this method is that it allowed us to examine changes in listening effort due to changes in the availability of prior semantic context between adjacent sentences.

The current experiment investigated the effects of two factors on speech intelligibility and perceived listening effort for hearing-aid users: (1) semantic context accumulated throughout a passage of speech, and (2) a hearing-aid setting ('Reverberant Room' setting on Signia Pure™ primax devices, Sivantos Inc.) designed for reverberant environments. We measured word report accuracy and subjective listening effort ratings for a modified version of the Connected Speech Test (CST; Cox et al, 1987), which was presented in a simulated reverberant environment with background talkers and cafeteria noise. Each sentence either had the same topic as the preceding sentence (i.e. congruent semantic context) or a different topic (i.e. incongruent semantic context). One possible outcome was that hearing-aid users would show no reduction in listening effort from congruent semantic context: perhaps the cognitive burden of the reverberant listening conditions, and that arising from degradation of the acoustic signal at the auditory periphery, are so great that listeners have few remaining cognitive resources to predict upcoming words based on their contextual probabilities. On the other hand, natural conversations are semantically rich, so we predicted that hearing-aid users would have learnt to rely on semantic context to help compensate for a degraded acoustic signal, thereby improving speech intelligibility and reducing listening effort. The

Accepted Manuscript

'Reverberant Room' hearing-aid setting used a combination of directional microphone, noise reduction, and de-reverberation signal processing—thus, we predicted it would improve speech intelligibility and reduce perceived effort compared to an omnidirectional hearing-aid setting.

Materials and Methods

Participants

We recruited 23 participants with bilateral sensorineural hearing loss. Three participants were excluded: one was unable to tolerate the sound levels of the stimuli without significant changes to the hearing aid or stimuli, one received an error in one of the audio files, and one was scored incorrectly due to an experimenter error. The remaining 20 participants (12 male) were aged 39–83 years (mean [$\bar{x}$] = 71.3 years, standard deviation [*SD*] = 10.8). Davis et al. (2011) reported a large benefit (corresponding to $\eta_p^2 = 0.91$; Faul et al., 2007) of sentence

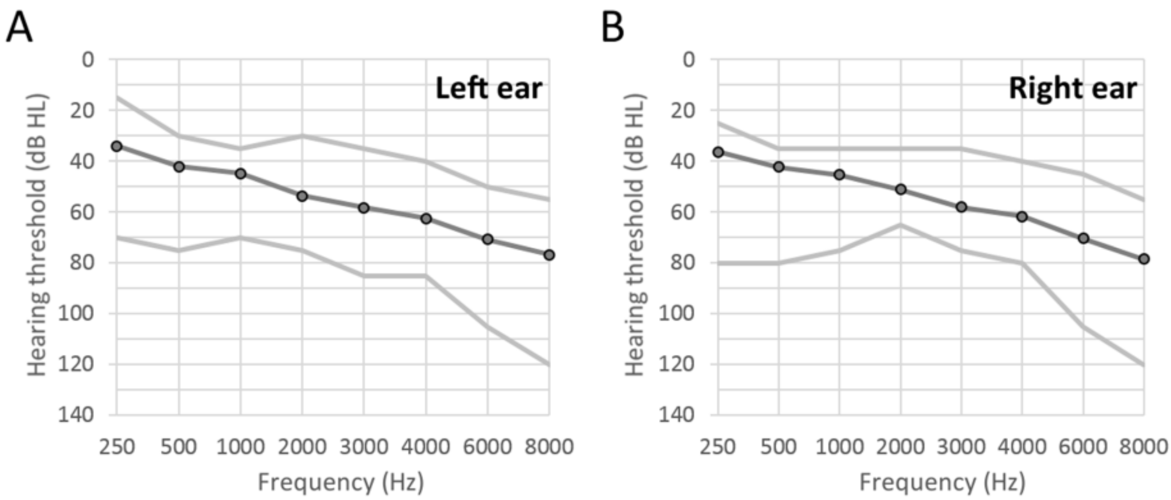

Figure 1. Pure-tone audiometric thresholds (dB HL) across participants, plotted separately for the left (A) and right (B) ears. The darker grey line illustrates the average and the lighter grey lines illustrate the highest and lowest thresholds across the group.

context on speech intelligibility in listeners with normal hearing; with 20 participants and an alpha level of 0.05, the estimated statistical power to detect within-subjects effect of this size is ~1.00 (Faul et al., 2007). Figure 1 illustrates their average pure-tone thresholds, which were measured at 8 frequencies (250, 500, 1000, 2000, 3000, 4000, 6000, and 8000 Hz) using a GSI-61 audiometer with ER-3A insert earphones. All participants had some experience with hearing aids: 3 were infrequent users and 17 were regular users. The experiment was cleared by Western University's Health Sciences Research Ethics Board. Informed consent was obtained from all participants and participants were compensated for their time.

Eight (8) of the 20 participants provided sentence-by-sentence listening effort ratings (described in more detail below). With 8 participants and an alpha level of 0.05, the estimated power to detect within-subject effects of the size reported by Davis et al. (2011) is ~1.00 (Faul et al., 2007).

Accepted Manuscript

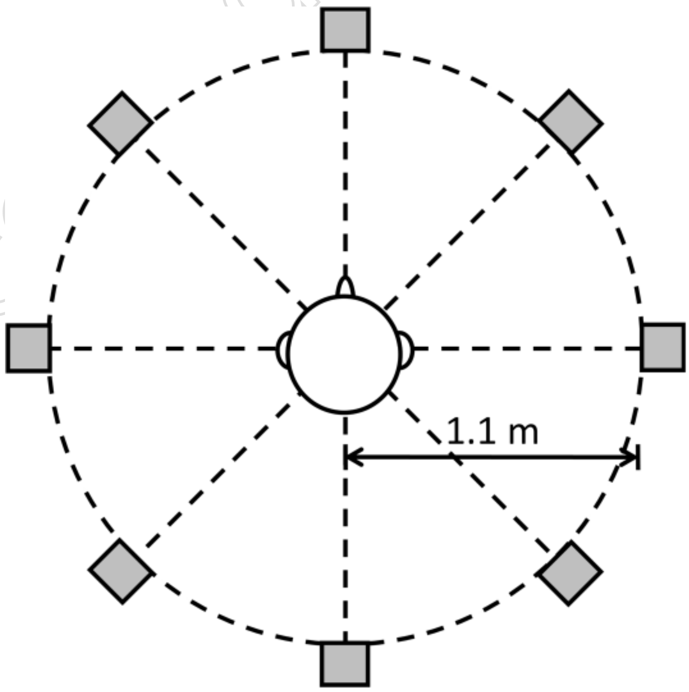

Figure 2. Layout of loudspeakers (grey squares) relative to a participant's head.

Apparatus and Stimuli

Participants were seated in an IAC acoustics (www.iacacoustics.com) double-walled sound booth. Stimuli were presented through eight Anthony Gallo Nucleus loudspeakers positioned at 0, 45, 90, 135, 180, 235, 270, and 315 degrees azimuth at a height of 1.2 metres and at a distance of 1.1 metres from the participant (Fig. 2). Participants sat facing the loudspeaker at 0-degrees azimuth.

Acoustic target stimuli were sentences from a modified version of the Connected Speech Test (CST; Cox et al. 1987; 1988). Sentences were original recordings of the CST, spoken by a female talker. The 220 CST sentences were separated into four sentence lists (see Supplemental Material, online), each containing 55 sentences with 139–143 key words and 6–7 topic changes. For each participant, three different sentence lists were selected: one for use with each hearing-aid setting and one for the practice run. Target CST sentences were presented from all eight loudspeakers.

Acoustic background noise consisted of a mixture of cafeteria noise and two custom-recorded speech passages spoken by a male talker. The background noise was presented from seven of the eight loudspeakers (excluding 0-degrees). One of the two passages came from the left (315° azimuth) and the other from the right (0° and 45° azimuth). The right passage was presented at the same level as the cafeteria noise (i.e. 0 dB SNR). The left passage was presented at 4 dB SNR relative to the cafeteria noise. The level of the background noise was fixed at a level of 55 dB(A) Leq(30 sec).

Stimulus files containing the target and background sounds were presented using Adobe® Audition (version CC 2015) and routed through an AudioFire12 Echo soundcard to a QSC amplifier. The amplifier was connected to eight Tucker-Davis Technologies PA5 attenuators—one for each loudspeaker. Loudspeakers were calibrated with a Larson Davis

Accepted Manuscript

824 Type 1 sound level meter prior to the project. Calibration checks prior to testing were completed with a MPT ST-805 Type II sound level meter.

Reverberation was applied to the target and background stimuli using Adobe® Audition's Studio Reverb effect (see Table 1).

Procedures

First, participants were fitted binaurally with commercially available Signia Pure™ primax (Sivantos Inc.) receiver-in-the-canal style hearing aids, using double domes to couple the receiver to the ear. Two hearing-aid programs were created using the Connexx 8 software: 'Universal', which used the omnidirectional setting, and 'Reverberant Room', which used a combination of directional microphone (amplifying sounds from in front and attenuating

Table 1. Reverberation settings applied to the acoustic stimuli in Adobe® Audition.

Setting	Value
Room size (%)	30
Decay (ms)	1600
Early reflections (%)	68.4
Stereo width (%)	30
High frequency cut (Hz)	1510
Low frequency cut (Hz)	60
Damping (%)	0
Diffusion (%)	20
Output level: Dry (%)	0
Output level: Wet (%)	100

Accepted Manuscript

sounds from behind the listener), noise reduction, and de-reverberation signal processing. The Reverberant Room setting provided a full-band directionality that attenuated noise from the rear by approximately 15 to 20 dB across frequencies, based on measures made in a clinical hearing aid analyzer (Audioscan VF2), along with approximately 5 dB low cut below 500 Hz. In contrast, the Universal program provided about 10 to 20 dB of directionality in a high frequency band that spanned approximately 2500 to 4000 Hz. The dereverberation signal processing used adaptive gain control to adjust to the level of the direct sound, which was assumed to reach the processor before reflected sounds and at a greater level than reflected sounds; adaptive compression times were set to optimise the ratio between the level of the direct sound and the level of the reverberation tail. Real ear verification of the frequency response of the hearing aids was completed using an Audioscan VF2 (version 4.4) and the aids were fine tuned to match the NAL-NL2 target at 65 dB.

Immediately after hearing-aid fitting, participants completed the CST task. They first completed a practice run, then they completed the full procedure twice: once using the ‘Universal’ hearing-aid setting and once using the ‘Reverberant Room’ setting. For each run, the background noise began approximately 3–4 s before the target sentences to give the adaptive features of the hearing aid time to adjust to the reverberant environment. The background sounds continued throughout the entire duration of the run, including the interval between target sentences. The inter-stimulus interval between adjacent target sentences was 5 s. After each target sentence, participants were required to, first, repeat the sentence and, second, verbally provide a numerical listening effort rating on a 7-point scale, where 1 indicated ‘No effort’ and 7 indicated ‘Extreme effort’ (Holube et al., 2016; Luts et al., 2010). Participants were familiarized with a visual version of the scale with descriptors (Fig. 3) at the start of the experiment. The range of the rating scale that individual participants used is reported in the Results section.

Accepted Manuscript

The practice run was completed using the ‘Universal’ hearing-aid setting. During the practice, the experimenter selected a signal-to-noise ratio (SNR) at which performance was not at floor or ceiling level. This level was used throughout the rest of the experiment. On average across participants, we used an SNR of +11.3 dB (*SD* = 4.3).

The order in which participants completed the main procedure using the ‘Universal’ and ‘Reverberant Room’ hearing-aid settings was counterbalanced across participants. Twelve of the 20 participants struggled to complete both the verbal sentence report and the verbal listening effort rating in the time between sentences during the practice run. The setup did not allow us to pause the experiment to provide participants with more time to respond verbally (because the target and background sounds for each condition were stored in single sound files, with equal durations, and ran concurrently), so these participants were instructed to only repeat the sentence (and not provide listening effort ratings) during the main procedure. Nevertheless, the inability of some participants to complete listening effort ratings in the available time was unrelated to age, audiometric thresholds, or speech intelligibility (as reported in the Results section), so it was unlikely to have biased the listening effort results.

epted Manuscript

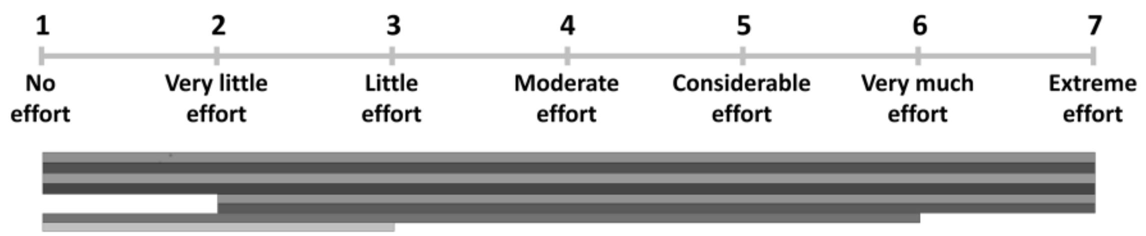

Figure 3. Visual representation of the listening effort scale, with which participants were familiarized at the beginning of the experiment. The bars below illustrate the range of listening effort ratings for individual participants (each row represents one participant).

Analyses

The number of correctly reported key words and the listening effort ratings were separated by context, depending on whether the sentence had the same topic as the previous sentence or a different topic. The first sentence from the list was included in the different-topic condition. We calculated the average percent correct score for sentences of each type (same topic / different topic) and for each hearing-aid condition (Universal / Reverberant Room) by comparing the number of key words reported correctly to the total number of key words in each condition. We then converted the percentages to rationalized arcsine units (RAU; Studebaker, 1985) before performing statistical analyses.

For participants who gave verbal reports of listening effort, we calculated the mean listening-effort rating in each condition, excluding trials in which no rating was given.

We conducted a 2 x 2 within-subjects ANOVA for speech intelligibility, with the factors Hearing-Aid Setting (Universal / Reverberant Room) and Context (same topic / different topic) as repeated measures. The speech intelligibility data met the assumptions of normality, as assessed using a combination of box plots, Q-Q plots, and the Kolmogorov-Smirnov test. For listening effort ratings, we used Wilcoxon signed-rank tests to compare the two Hearing-Aid Settings and two Context conditions.

To analyse sentence-by-sentence correlations between speech intelligibility and listening effort in individual participants, we calculated the percentage of key words reported correctly for each sentence. We included sentences belonging to all of the Hearing-Aid Setting and Context conditions in this analysis. We calculated Spearman’s rank correlation coefficients between listening effort ratings and speech intelligibility.

Accepted Manuscript

Results

Comparison of Participants Who Did and Did Not Report Subjective Listening Effort

More than half of participants did not complete the sentence-by-sentence listening effort ratings. We hypothesised that the ability to complete the ratings may relate to age, hearing thresholds, or speech intelligibility, so we conducted independent *t*-tests to compare participants who completed the listening effort ratings to those who did not. Participants who did not complete the ratings ($\bar{x}$ = 73.5 years, *SD* = 7.4) were slightly older than those who did ($\bar{x}$ = 67.6 years, *SD* = 14.1), although there was no significant difference in age between the groups [$t(18) = 1.27$, $p = 0.22$]. There was also no difference in 8-frequency average hearing thresholds (250, 500, 1000, 2000, 3000, 4000, 6000, and 8000 Hz, averaged across both ears) between participants who did not complete the ratings ($\bar{x}$ = 54.5 dB HL, *SD* = 7.4) and those who did ($\bar{x}$ = 56.2 dB HL, *SD* = 9.0) [$t(18) = 0.48$, $p = 0.64$].

Average speech intelligibility score did not differ between participants who did not complete the ratings ($\bar{x}$ = 69.2 RAU, σ = 10.2) and those who did ($\bar{x}$ = 71.1 RAU, σ = 10.1) [$t(18) = 0.42$, $p = 0.68$]. In addition, the magnitude of the speech intelligibility benefit (i.e. Same-Topic – Different-Topic) in the Universal [$t(19) = 0.65$, $p = 0.52$, $g_s = 0.28$] and Reverberant Room [$t(18.9) = 0.24$, $p = 0.82$, $g_s = 0.10$] hearing aid conditions did not differ between participants who did and did not complete the ratings. These results suggest that (1) the second task (i.e. reporting listening effort ratings verbally) did not reduce speech intelligibility or the magnitude of the intelligibility benefit gained from context, and (2) the comparison of listening effort ratings across Hearing Aid and Context conditions (below) was not biased by some participants not rating listening effort.

Speech Intelligibility

Figure 4A illustrates the mean percentage of key words participants correctly reported (RAU-

Accepted Manuscript

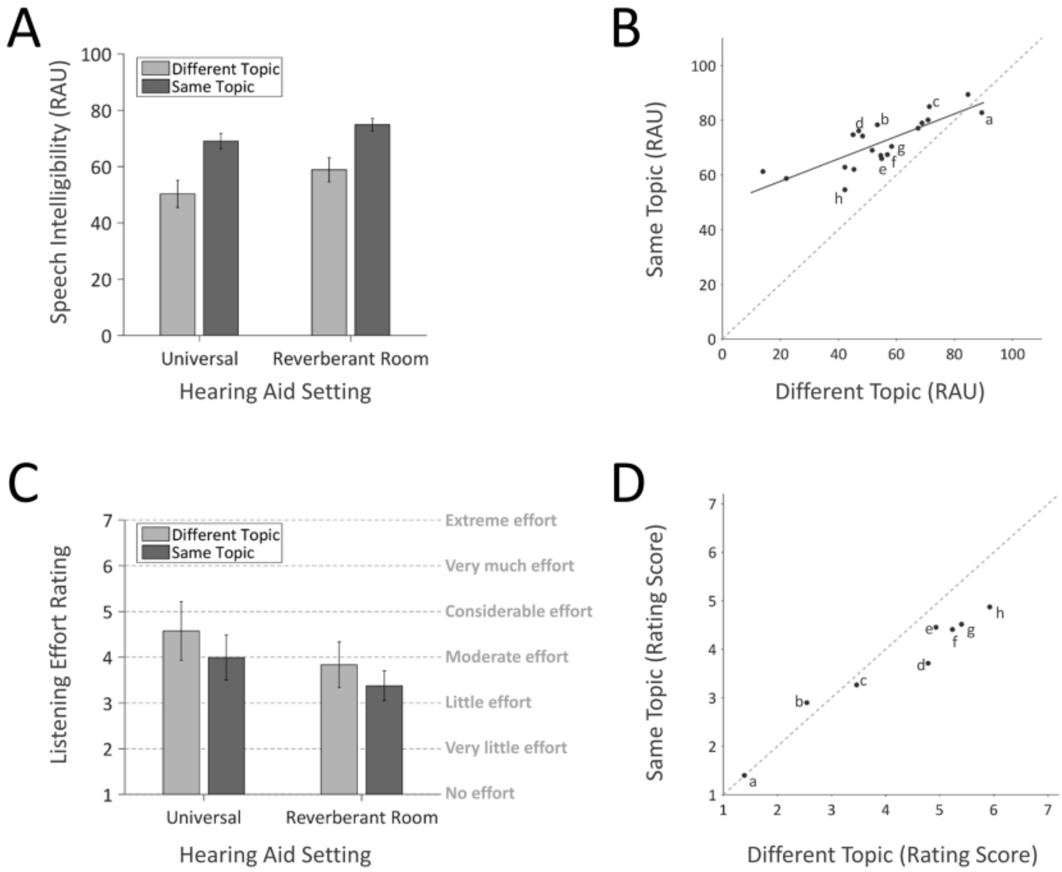

Figure 4. A, Average speech intelligibility (converted to rationalized arcsine units [RAU]) across participants for the two Hearing-Aid and two Context conditions. Error bars display one standard error of the mean. B, Relationship between speech intelligibility (RAU) for same-topic sentences and speech intelligibility for different-topic sentences. Each dot represents one participant. Letters a–h indicate participants who completed listening effort ratings (and correspond to the letters displayed in panel D). The grey dashed line is plotted at y = x. Dots above the grey dashed line show better speech intelligibility for same-topic than different-topic sentences. The solid black line illustrates the least-squares line of best fit, with the equation of the line of best fit and r^2 value resulting from a Pearson's product-moment correlation between different-topic and same-topic sentences displayed on the graph. C, Same as A, but for listening effort. D, Same as B, but for listening effort. Letters a–h beside the

dots show the mapping between individual scores in panels B and D. Dots below the grey dashed line represent lower listening effort for same-topic than different-topic sentences.

transformed) across the different Hearing-Aid Setting and Context conditions. A 2 x 2 within-subjects ANOVA showed significantly better speech intelligibility when the topic was the same as the previous sentence ($\bar{x}$ = 72.0 RAU, *SD* = 9.3) than when it was different ($\bar{x}$ = 54.6 RAU, *SD* = 18.0) [$F(1, 19) = 41.52$, $p < 0.001$, $\eta_p^2 = 0.69$]. Participants achieved better speech intelligibility when using the Reverberant Room hearing-aid setting ($\bar{x}$ = 66.9 RAU, *SD* = 13.8) than the Universal setting ($\bar{x}$ = 59.7 RAU, *SD* = 15.5), [$F(1, 19) = 5.61$, $p = 0.029$, $\eta_p^2 = 0.23$]. There was no significant interaction between Context and Hearing-Aid Setting factors [$F(1, 19) = 0.53$, $p = 0.47$, $\eta_p^2 = 003$].

On an individual basis, collapsed across Hearing-Aid Setting conditions, 19 out of the 20 participants achieved better speech intelligibility for same-topic than different-topic sentences (Fig. 4B). Across all participants, the average benefit gained from context was 17.4 RAU (*SD* = 12.1). Context improved speech intelligibility scores most for listeners who scored worst on different-topic sentences (see Fig. 4B).

Comparing the two hearing-aid settings, collapsed across Context conditions, 15 out of 20 participants achieved better speech intelligibility using the Reverberant Room setting than the Universal setting. The average benefit gained from the Reverberant Room setting was 7.2 RAU (*SD* = 13.6).

A paired samples *t*-test revealed that the improvement in speech intelligibility from congruent (compared to incongruent) semantic context ($\bar{x}$ = 17.4 RAU, *SD* = 12.1) was significantly greater than the improvement from the Reverberant Room (compared to Universal) hearing-aid setting ($\bar{x}$ = 7.2 RAU, *SD* = 13.6) [$t(19) = 2.35$, $p = 0.030$].

Accepted Manuscript

Listening Effort

First, we wanted to explore whether participants were using the whole listening effort rating scale (i.e. 1–7). Four (4) of the 8 participants who completed the listening effort ratings used the entire scale, 3 provided ratings that spanned most of the scale (1–6 or 2–7), whereas 1 participant used only the lower half of the scale (1–3; see Fig. 3 for a visual representation). On some trials, participants provided no listening-effort rating. The number of unrated sentences was very low and a Wilcoxon signed-rank test showed that the number of unrated sentences did not differ significantly between the Reverberant Room ($\bar{x} = 1.6$, $SD = 0.6$) and Universal ($\bar{x} = 2.4$, $SD = 1.3$) conditions [$T = 4.5$, $z = 0.18$, $p = 0.85$, $r = 0.07$].

Figure 4C illustrates the average listening effort ratings per sentence across the different Hearing Aid Setting and Context conditions. Listening effort was significantly lower when the sentence topic was the same as the previous sentence ($\bar{x} = 3.69$, $SD = 1.15$) than when it was different ($\bar{x} = 4.21$, $SD = 1.58$) [$T = 4$, $z = 1.96$, $p = 0.050$, $r = 0.69$]. Participants reported lower listening effort for sentences in the Reverberant Room ($\bar{x} = 3.61$, $SD = 1.17$) than the Universal ($\bar{x} = 4.29$, $SD = 1.59$) hearing-aid condition [$T = 2$, $z = 2.24$, $p = 0.025$, $r = 0.79$]. The reduction in listening effort from congruent context (i.e. Different-Topic – Same-Topic) did not differ significantly between the Reverberant Room ($\bar{x} = 0.46$, $SD = 0.53$) and Universal ($\bar{x} = 0.58$, $SD = 0.60$) conditions [$T = 12$, $z = 0.84$, $p = 0.40$, $r = 0.30$], showing that there was no interaction between Context and Hearing-Aid Setting factors.

Context reduced listening effort (i.e. same-topic sentences were rated as lower effort than different-topic sentences) for 6 out of 8 participants (Fig. 4D). On average the reduction in listening effort from context was 0.52 ($SD = 0.54$). Context reduced listening effort most for listeners who found different-topic sentences most effortful (see Fig. 4D).

Accepted Manuscript

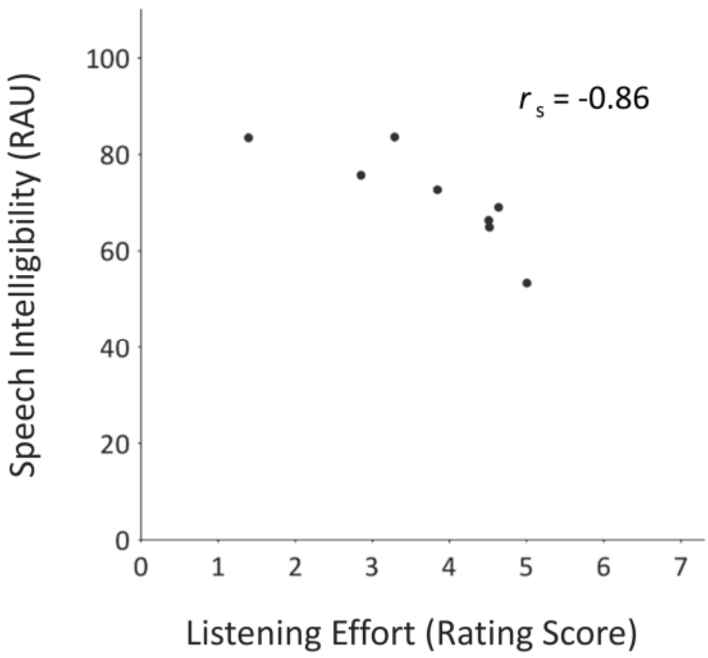

Figure 5. Relationship between listening effort ratings and speech intelligibility (in rationalized arcsine units [RAU]). Each dot represents one participant. r_s = Spearman's rho.

Seven (7) out of 8 participants reported lower listening effort in the Reverberant Room condition than the Universal condition. The average benefit gained from the Reverberant Room setting was 0.68 (*SD* = 0.66).

A Wilcoxon signed-rank test revealed that the benefit to listening effort gained from congruent (compared to incongruent) semantic context ($\bar{x}$ = 0.52, *SD* = 0.54) did not differ significantly from the benefit gained from the Reverberant Room (compared to Universal) hearing-aid setting ($\bar{x}$ = 0.68, *SD* = 0.66) [T = 13, z = 0.70, p = 0.49, r = 0.25].

Relationship between Speech Recognition and Subjective Listening Effort

Figure 5 illustrates, for participants who completed the listening effort ratings, the relationship between listening effort and speech intelligibility. Participants who achieved poorer speech intelligibility reported greater listening effort [r_s = -0.86, p = 0.011].

Accepted Manuscript

We also analysed the relationship between listening effort and speech intelligibility for each participant individually, to investigate whether listening effort ratings for individual sentences related to intelligibility of key words for that sentence. For each sentence, we calculated the percentage of key words the participant reported correctly. Spearman's rank correlations showed highly significant ($p < 0.001$) negative correlations between listening effort and speech intelligibility for all 8 participants, with a median *r*-value of -0.59 (range = -0.41 to -0.71; see Figure, Supplemental Digital Content). Thus, even within participants on a sentence-by-sentence basis, poorer speech intelligibility was coupled with greater perceived listening effort.

Discussion

Participants correctly identified more key words when the topic was the same as the previous sentence than when it was different. Participants who were able to report listening effort ratings in the available time showed a reliable reduction in perceived effort for same-topic than different-topic sentences. These results demonstrate that congruent semantic context, accumulated from one sentence to the next, helps hearing-aid users comprehend speech in the presence of multiple simultaneous talkers and background noise.

Participants also achieved better speech intelligibility, overall, when they used the Reverberant Room hearing-aid setting—designed to aid listening in reverberant environments—than when they used the Universal (omnidirectional) setting. Participants who reported listening effort ratings reliably reported reduced effort when using the Reverberant Room hearing-aid setting. Listening effort was reduced to a similar extent when the Reverberant Room setting was used (0.7 reduction in listening effort ratings) as when the previous sentence was the same topic (0.5 reduction in listening effort ratings). However, the improvement in speech intelligibility from congruent semantic context (17% improvement in

Accepted Manuscript

percent correct) was greater than the improvement from the Reverberant Room hearing-aid setting (7% improvement in percent correct).

Semantic context improved speech intelligibility and reduced perceived listening effort to a similar extent for both hearing-aid settings, showing that the benefit of semantic context for hearing-aid users can extend above and beyond the improvement gained from changing hearing-aid settings. Overall, these findings highlight the enormous potential of cognitive factors (specifically, those allowing listeners to benefit from semantic context) for improving speech intelligibility and reducing perceived listening effort in noisy acoustic environments for hearing-aid users. When semantic context benefits are combined with effective signal processing, the combined effect results in listening effort ratings in the range of little to moderate effort, whereas effort ratings are in the moderate to considerable range if no signal processing is used and the topic of the speech is not consistent with that of the previous sentence (Fig. 4C).

Semantic Context Improves Speech Intelligibility

The improvement in speech intelligibility for sentences with congruent semantic context is consistent with the results of several previous studies that found better speech intelligibility when the final word of a sentence was predictable than when it was unpredictable in listeners with normal hearing (e.g., Dubno et al, 2000; Davis et al, 2011) and in listeners with hearing impairment (Pichora-Fuller et al, 1995; Desjardins & Doherty, 2014). The improvement in speech intelligibility we observed is larger than that reported for listeners with normal hearing (Dubno et al, 2000; Davis et al, 2011) and is as large as one (Desjardins & Doherty, 2014) but not another (Pichora-Fuller et al, 1995) study reported for listeners with hearing impairment, although differences in age and audiometric thresholds may explain differing magnitudes of benefit. The method by which semantic context was manipulated in the current experiment—throughout a passage of speech with topic changes—

Accepted Manuscript

differs from these previous experiments, which varied the predictability of individual words within a sentence. The current results demonstrate that hearing-aid users can utilize semantic context over a longer time period (i.e. using the topic from the previous sentence) to improve speech intelligibility.

Semantic Context Reduces Perceived Effort

This experiment extends previous studies showing that semantic context improves speech intelligibility by demonstrating that semantic context also reduces the perceived effort of listening in noisy acoustic environments. Winn (2016) found that high final-word predictability was related to lower listening effort, measured using pupillometry, for people with cochlear implants. The current results extend this result by showing that (1) semantic context accumulated from the previous sentence topic reduces listening effort, (2) semantic context reduces listening effort for hearing-aid users, who typically have greater residual hearing than cochlear implant users, and (3) self-reports of listening effort are sensitive to semantic context.

The current results differ from those reported by Desjardins and Doherty (2014), who found no effect of word predictability on self-reported listening effort or performance on a secondary visual tracking task. The difference in findings could be due to at least two factors, which cannot be distinguished here: (1) the sentence-by-sentence measure of listening effort is more sensitive than effort ratings spanning multiple sentences and secondary-task performance; or (2) having a topic carried over from a preceding sentence reduces effort more than does having highly predictable words within a sentence. Future studies could directly compare different types of contextual information or the sensitivity of sentence-by-sentence listening effort ratings with that of other listening effort measures.

Accepted Manuscript

Individual Differences in Benefit Gained from Semantic Context

Our results demonstrate consistent benefits of congruent semantic context across the group of hearing-aid users (Fig. 4B, D). Nineteen of the 20 participants gained an improvement in speech intelligibility from congruent semantic context; also, 6 of the 8 participants who rated listening effort found semantic context reduced effort. Thus, semantic context can benefit speech perception across a range of participants who differ in age and whose average hearing thresholds span mild to severe hearing losses (see Fig. 1).

Participants who performed most poorly on different-topic sentences gained the greatest improvement in speech intelligibility and perceived effort from congruent semantic context. This finding may be trivial because speech intelligibility was already at or near ceiling for participants who achieved better speech intelligibility. On the other hand, the result demonstrates that even listeners who perform poorly in speech-in-noise have the cognitive abilities required to benefit, at least as much as the listeners who have better speech in noise performance, from provision of a consistent topic.

Hearing-Aid Setting Influences Intelligibility and Effort

Our results demonstrate that hearing-aid setting can influence perceived listening effort in people with hearing impairment. Previous experiments examining different hearing-aid settings in people with hearing impairment have produced mixed results (see Bentler et al, 2008; Brons et al, 2013; Desjardins & Doherty, 2014; Desjardins, 2016; Hornsby, 2013). Our results demonstrate that a hearing-aid program designed to aid listening in reverberant environments (which uses a combination of directional microphone, noise reduction, and de-reverberation signal processing) reduces sentence-by-sentence subjective listening effort ratings when people with hearing impairment listen to speech in a simulated reverberant environment—this reduction in effort was observed for 7 of the 8 participants. The improvement in speech intelligibility and listening effort might be attributable to the

Accepted Manuscript

attenuation of reflected sounds, which are particularly challenging for listeners with hearing impairment (e.g., Marrone et al, 2008). Although, at least some of the improvement may have arisen from the directional microphone and noise reduction algorithms that were used in the Reverberant Room hearing-aid setting, because these were absent from the Universal setting.

There are several possible explanations of the observed difference in listening effort, which was not found in some of the previous studies: (1) the 'Reverberant Room' hearing-aid setting (Sivantos, Inc.) is more effective at reducing listening effort than were hearing-aid settings tested in previous experiments; (2) the simulated reverberant environment was more acoustically challenging than the conditions used in previous experiments, meaning that participants relied on hearing-aid processing to a greater extent; or (3) self-reports of sentence-by-sentence listening effort are more sensitive than the measures of listening effort (ratings over multiple sentences and measures of secondary-task performance) used in previous studies.

Hearing-Aid Setting and Context Do Not Interact

We found no interaction between context and hearing-aid setting, consistent with the idea that these two factors may affect perceived listening effort independently. Different factors likely place demands on different parts of the auditory system (for reviews, see Mattys et al, 2012; Johnsrude & Rodd, 2016; Pichora-Fuller et al, 2016). Hearing loss degrades the speech signal at the auditory periphery and makes concurrent sounds more difficult to segregate, particularly in reverberant environments (e.g., Marrone et al, 2008). Applying different hearing-aid settings changes the acoustic signal, possibly improving the SNR of target sentences for the Reverberant Room condition. An improvement in SNR may have helped participants to better segregate the mixture of target and background sounds, reducing listening effort by allowing listeners to better focus on target sentences through

Accepted Manuscript

improved representations of the acoustic signal. In contrast, semantic context cannot be explained by acoustics and, instead, must arise from cognitive processes.

Semantic context may affect listening effort at a higher stage of auditory processing by facilitating word predictability. Functional magnetic resonance imaging (fMRI) studies have shown that activity in the left inferior frontal gyrus and superior temporal gyri relate to the restoration of degraded speech by linguistic expectations (i.e. greater for semantically coherent, e.g., “Her new skirt was made of denim”, compared to semantically anomalous, e.g., “Her good slope was done in carrot”, sentences; Davis et al, 2011) and word predictability (i.e., greater for sentences in which the final word is predictable than unpredictable from the preceding sentence context; Obleser & Kotz, 2010). Thus, hearing-aid processing and congruent semantic context are likely to reduce listening effort by different means. This idea implies that the greatest benefit for listeners with hearing impairment is likely to be obtained by using a combination of these factors, as we found in the current experiment.

In the current experiment, speech intelligibility and listening effort ratings were highly correlated. The extent to which hearing-aid setting and context would reduce perceived effort at equal levels of speech intelligibility is unclear. For example, would these factors reduce perceived effort when speech is highly intelligible? Future studies could aim to investigate this idea explicitly, to further tease apart effects on speech intelligibility and listening effort.

Evaluation of Listening Effort Measure

Here, we asked listeners to verbally report listening effort on a sentence-by-sentence basis. Most previous studies asked listeners to rate effort after a block of three (Krueger et al., 2017) or more (e.g., Mackersie & Cones, 2011; Rudner et al, 2012) sentences, although one previous study asked participants with normal hearing to rate listening effort after every

Accepted Manuscript

sentence (Picou et al, 2011). In the current experiment some, but not all, listeners with hearing impairment were able to verbally report perceived listening effort as well as word report in the time available (5 s) after each sentence. We found no difference in age, hearing thresholds, or speech intelligibility scores between participants who did and those who did not complete the ratings and it is, therefore, unclear why some participants needed more time to give both verbal responses. Possibly, participants who did not give listening effort ratings experienced greater levels of effort than those who completed the ratings. However, given that speech intelligibility correlated with listening effort ratings for participants who completed the ratings—and speech intelligibility scores did not differ significantly between participants who did and did not complete the ratings—this explanation seems unlikely. We did not encourage participants to complete listening effort ratings if they were not able to provide the ratings easily during practice trials. Possibly, with more training, all participants would be able to complete listening effort ratings on a sentence-by-sentence basis. Inflexibility of the inter-stimulus interval duration was a limitation of the current task and we suggest that future experiments using this method provide longer time intervals (i.e. > 5 s) for participants to respond, or wait to present the next sentence until both verbal responses have been collected.

The results demonstrate that the sentence-by-sentence measure of perceived effort we used is sensitive to different sentence contexts and hearing-aid settings and, therefore, has great potential for future listening effort studies. We found no difference between participants who did and those who did not complete the listening effort ratings in (1) overall speech intelligibility or (2) the magnitude of the congruent context speech intelligibility benefit. This finding demonstrates that providing verbal listening effort ratings after each sentence (i.e. the secondary task) did not interfere with word report (i.e. the primary task). Another advantage is that sentence-by-sentence listening effort ratings would be quick and

Accepted Manuscript

easy to administer in clinical settings as part of the regular assessment procedure, unlike dual-task procedures (e.g., Fraser et al, 2010), pupillometry (e.g., Koelewijn et al, 2015), or electro-encephalography (e.g., Wisniewski et al, 2015). Nevertheless, it is important to consider that different listening effort measures may assess different constructs and/or manifestations of listening effort (for a review, see McGarrigle et al, 2014). This idea is consistent with the mixed results found in previous studies when different listening effort measures were used.

In our instructions to participants, we emphasised that they should rate the effort they exerted to understand the sentence, rather than the words they were able to report. The descriptors that were used for the 7-point scale also explicitly stated that participants should report effort (see Fig. 3). Nevertheless, the interpretation of listening effort is subjective, may differ across participants, and may index a different component of listening than other measures of listening effort. Future research should aim to examine the similarities and differences between different listening effort measures and their relative sensitivities.

Some participants did not use the entire listening effort scale (i.e. 1–7). This would not affect the current within-subjects design, but would be problematic for a between-subjects design, in which each participant completes a sub-set of conditions. Future experiments using between-subjects designs could consider explicitly encouraging participants to use the full scale or normalising listening effort ratings across participants before performing statistical comparisons.

Accepted Manuscript

Conclusions

The current experiment demonstrates that hearing-aid users can use congruent semantic context between adjacent sentences to (1) improve speech intelligibility, and (2) reduce perceived listening effort when listening to speech in the presence of background noise. The magnitude of benefit gained from congruent semantic context was greater than the benefit gained from a hearing-aid setting designed to aid listening in reverberant environments (compared to an omnidirectional hearing-aid setting). Overall, these findings highlight the enormous potential of cognitive factors (specifically, those that allow listeners to utilise semantic context) for improving speech intelligibility and reducing perceived listening effort in noisy acoustic environments for hearing-aid users. The ability of participants with hearing impairment to use cognition to reduce listening effort could be key to improving future rehabilitation strategies.

Disclosure of Interest

This work was supported by Sivantos. The manufacturer provided stimuli, technical support, and covered the operating costs of the study. However, financial support by Sivantos is/was not related to the outcomes of this study. The authors have no financial interest in the commercialization of any product or knowledge arising from this work.

References

Bentler, R., Wu, Y.-H., Kettel, J., Hurtig, R., 2008. Digital noise reduction: outcomes from laboratory and field studies. *Int. J. Audiol.*, 47(8), p.447–60.

Brons, I., Houben, R., Dreschler, W.A., 2013. Perceptual effects of noise reduction with respect to personal preference, speech intelligibility, and listening effort. *Ear Hear.*, 34(1), p.29–41.

Cox, R.M., Alexander, G.C., Gilmore, C., 1987. Development of the Connected Speech Test (CST). *Ear Hear.*, 8, p.119S–126S.

Accepted Manuscript

Davis, M.H., Ford, M.A., Kherif, F., Johnsrude, I.S., 2011. Does semantic context benefit speech understanding through “top–down” processes? Evidence from time-resolved sparse fMRI. *J. Cogn. Neurosci.*, 23(12), p.3914–3932.

Desjardins, J.L., 2016. The effects of hearing aid directional microphone and noise reduction processing on listening effort in older adults with hearing loss. *J. Am. Acad. Audiol.*, 27(1), p.29–41.

Desjardins, J.L., Doherty, K.A., 2014. The effect of hearing aid noise reduction on listening effort in hearing-impaired adults. *Ear Hear.*, 35(6), p.600–610.

Dubno, J.R., Ahlstrom, J.B., Horwitz, a R., 2000. Use of context by young and aged adults with normal hearing. *J. Acoust. Soc. Am.*, 107(1), p.538–46.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.

Fraser, S., Gagné, J.-P., Alepins, M., Dubois, P., 2010. Evaluating the effort expended to understand speech in noise using a dual-task paradigm: the effects of providing visual speech cues. *J. speech, Lang. Hear. Res.*, 53(1), p.18–33.

Gatehouse, S., Akeroyd, M. A., 2006. Two-eared listening in dynamic situations. *Int. J. Audiol.*, 45, p.S120-4.

Gatehouse, S., Noble, W., 2004. The Speech, Spatial and Qualities of Hearing scale (SSQ). *Int. J. Audiol.*, 43, p.85–99.

Holube, I., Haeder, K., Imery, C., Weber, R., 2016. Subjective listening effort and electrodermal activity in listening situations with reverberation and noise. *Trends in Hearing*, 20, 1–15.

Hornsby, B.W.Y., 2013. The effects of hearing aid use on listening effort and mental fatigue associated with sustained speech processing demands. *Ear Hear.*, 34(5), p.523–34.

Accepted Manuscript

Johnsrude, I.S., Mackey, A., Hakyemez, H., Alexander, E., Trang, H.P., et al, 2013. Swinging at a cocktail party: voice familiarity aids speech perception in the presence of a competing voice. *Psychol. Sci.*, 24(10), p.1995–2004.

Johnsrude, I.S., Rodd, J.M., 2016. Factors that increase processing demands when listening to speech. In G. Hickok & S. L. Small, eds. *Neurobiology of Language*. Elsevier, pp. 491–502.

Koelewijn, T., de Kluiver, H., Shinn-Cunningham, B.G., Zekveld, A.A., Kramer, S.E., 2015. The pupil response reveals increased listening effort when it is difficult to focus attention. *Hear. Res.*, 323, p.81–90.

Kramer, S.E., Kapteyn, T.S., Houtgast, T., 2006. Occupational performance: Comparing normally-hearing and hearing-impaired employees using the Amsterdam Checklist for Hearing and Work. *Int. J. Audiol.*, 45(9), p.503–512.

Krueger, M., Schulte, M., Brand, T., & Holube, I. (2017). Development of an adaptive scaling method for subjective listening effort. *J. Acoust.. Soc. Am.*, 141(6), p.4680–4693.

Luts, H., Eneman, K., Wouters, J., Schulte, M., Vormann, M., et al., 2010. Multicenter evaluation of signal enhancement algorithms for hearing aids. *J Acoust. Soc Am.*, 127(3), p.1491-1505.

Mackersie, C.L., Cones, H., 2011. Subjective and psychophysiological indices of listening effort in a competing-talker task. *J Am Acad Audiol*, 22(2), p.113–122.

Marrone, N., Mason, C.R., Kidd, G., 2008. The effects of hearing loss and age on the benefit of spatial separation between multiple talkers in reverberant rooms. *J. Acoust. Soc. Am.*, 124(5), p.3064–75.

Mattys, S.L., Davis, M.H., Bradlow, A.R., Scott, S.K., 2012. Speech recognition in adverse conditions: A review. *Lang. Cogn. Process.*, 27(7–8), p.953–978.

Accepted Manuscript

McGarrigle, R., Munro, K.J., Dawes, P., Stewart, A.J., Moore, D.R., et al, 2014. Listening effort and fatigue: what exactly are we measuring? A British Society of Audiology Cognition in Hearing Special Interest Group “white paper.” *Int. J. Audiol.*, 53(7), p.433–45.

Neher, T., Grimm, G., Hohmann, V., 2014. Perceptual consequences of different signal changes due to binaural noise reduction: do hearing loss and working memory capacity play a role? *Ear Hear.*, (2013), p.213–227.

Noble, W., Gatehouse, S., 2006. Effects of bilateral versus unilateral hearing aid fitting on abilities measured by the Speech, Spatial, and Qualities of Hearing scale (SSQ). *Int. J. Audiol.*, 45(3), p.172–181.

Obleser, J., Kotz, S.A., 2010. Expectancy constraints in degraded speech modulate the language comprehension network. *Cereb. Cortex*, 20(3), p.633–640.

Pichora-Fuller, M.K., Kramer, S.E., Eckert, M.A., Edwards, B., Hornsby, B.W.Y., et al, 2016. Hearing impairment and cognitive energy: The framework for understanding effortful listening (FUEL). *Ear Hear.*, 37, p.5S–27S.

Pichora-Fuller, M.K., Schneider, B.A., Daneman, M., 1995. How young and old adults listen to and remember speech in noise. *J. Acoust. Soc. Am.*, 97(1), p.593–608.

Picou, E.M., Ricketts, T.A., Hornsby, B.W.Y., 2011. Visual cues and listening effort: individual variability. *J. Speech. Lang. Hear. Res.*, 54, p.1416–1430.

Picou, E.M., Ricketts, T. A., Hornsby, B.W.Y., 2013. How hearing aids, background noise, and visual cues influence objective listening effort. *Ear Hear.*, 34(5), p.e52-64.

Rudner, M., Lunner, T., Behrens, T., Thorén, E.S., Rönnberg, J., 2012. Working memory capacity may influence perceived effort during aided speech recognition in noise. *J. Am. Acad. Audiol.*, 23(8), p.577–89.

Accepted Manuscript

Studebaker, G.A., 1985. A “rationalized” arcsine transform. *J. Speech Lang. Hear. Res.*, 28(3), p.455.

Sumby, W., Pollack, I., 1954. Visual contribution to speech intelligibility in noise. *J. Acoust. Soc. Am.*, p.212–215.

Tun, P.A., Mccoy, S., Wingfield, A., 2009. Aging, hearing acuity, and the attentional costs of effortful listening. *Psychol. Aging*, 24(3), p.761–766.

Winn, M., 2016. Rapid release from listening effort resulting from semantic context, and effects of spectral degradation and cochlear implants. *Trends Hear.*, 20, p.1–17.

Winn, M.B., Rhone, A.E., Chatterjee, M., Idsardi, W.J., 2013. The use of auditory and visual context in speech perception by listeners with normal hearing and listeners with cochlear implants. *Front. Psychol.*, 4(NOV), p.1–13.

Wisniewski, M.G., Thompson, E.R., Iyer, N., Estepp, J.R., Goder-Reiser, M.N., et al, 2015. Frontal midline θ power as an index of listening effort. *Neuroreport*, 26(2), p.94–99.

Wu, Y.-H., Aksan, N., Rizzo, M., Stangl, E., Zhang, X., et al, 2014. Measuring listening effort: driving simulator vs. simple dual-task paradigm. *Ear Hear.*, 35(6), p.623–632.

Accepted Manuscript