HEARING WITHOUT LISTENING

Running Head: HEARING WITHOUT LISTENING

Hearing without listening: attending to a quiet audiobook

Hettie Roebuck[a,*], Kun Guo[a], Patrick Bourke[a]

[a]School of Psychology, University of Lincoln, Lincoln, LN6 7TS, UK

E-mail: hroebuck@wisc.edu; kguo@lincoln.ac.uk; pbourke@lincoln.ac.uk

*Corresponding author: Hettie Roebuck, Waisman Center, 1500 Highland Ave, University of

Wisconsin Madison, 53706 USA

1

Abstract

Careful systematic tests of hearing ability may miss the cognitive consequences of sub-optimal hearing when listening in the real world. In Experiment One, sub-optimal hearing is simulated by presenting an audiobook at a quiet but discriminable level over 50 minutes. Recall of facts, words and inferences are assessed and performance compared to another group at a comfortable listening volume. At the quiet intensity, participants are able to detect, discriminate and identify spoken words but do so at a cost to sequential accuracy and fact recall when attention must be sustained over time. To exclude other interpretations, the effects are studied in Experiment Two by comparing recall to the same sentences presented in isolation. Here, the differences disappear. The results demonstrate that the cognitive consequences of listening at low volume arise when sustained attention is demanded over time.

Introduction

A standard hearing test allows us to establish the intensity level at which it is possible to detect different sounds, and is used to screen for hearing loss when intervention might be required (ASHA, 1995). However, there is increasing evidence to suggest that these measures may underestimate the difficulties experienced beyond the test. Bess, Dodd-Murphy and Parker (1998) found that children with "minimal" hearing loss had greater problems with reading, language, communication and attention than their peers. The educational impact of mild hearing losses are increasingly documented (e.g., Quigley & Thomure, 1968; Teaside & Sorensen 2007; Flexer, Millin & Brown 1990; Pakulski & Kaderavek, 2002). So why, if the hearing loss is indeed minimal, are cognitive difficulties still frequently reported?

Hearing ability is conventionally measured by presenting pure tones across different frequencies within the range of speech, and individual thresholds are obtained through repeated presentations of the tones. Standard hearing tests demonstrate high test retest reliability (e.g., Wertheimer, 1955; Ward, 1957), suggesting the levels achieved are an accurate measure of ability to detect a particular sound. Despite high reliability, these short tests with abstract stimuli may not accurately represent the additional demands of attending in everyday life with mild or slight hearing losses. Comparisons of children's speech and pure tone audiograms with pre-surgery 'glue ear' showed that 30% of children had worse speech audiograms than their pure tone audiogram suggested (Pringle, Thomson, & Reddy, 1993). The pure tone test in isolation does not appear to reflect the complexities of identifying speech.

Perhaps one reason the consequences of 'mild' hearing losses are underrepresented in conventional hearing tests is related to the greater processing demands that poor hearing imposes, rather than the mild perceptual disability per se. Notably a high proportion of individuals with hearing loss are also reported to receive diagnoses associated with problems

maintaining attention (Williams & Abeles, 2004). It has been suggested that in order to successfully perceive degraded stimuli a greater amount of attentional resources are required, leaving less resources for encoding and consolidation in long term memory (McCoy, Tun, Cox, Colangelo, Stewart, & Wingfield 2005). In the case of audition, when speech is quiet but detectable, additional attentional resources may be required to resolve perceptual identity. The impact of effortful listening in normal hearing listeners has also been observed in a controlled environment simply by reducing the perceptual intensity (Roebuck, Guo & Bourke, 2015). While such a method does not reflect all the challenges an individual with hearing loss will experience e.g., variability in hearing across frequencies, asymmetric hearing acuities, intermittent difficulty or impact of having a long term deficit the method allows experimental control to assess the role of effortful listening independently of these other factors. Studies with normal hearing listeners attending to quiet but detectable stimuli have shown that quiet stimuli take longer to identify, and are associated with increased errors of attention (missed targets) and inhibition (false alarms) (Roebuck, Guo &Bourke, 2015). However, the implications of trying to listen at a quiet intensity on everyday cognition without other difficulties associated with real hearing loss have yet to be explored. For example when attending to speech over an extended period of time.

The impact of this kind of 'effortful listening' on complex processing can be investigated by asking participants with normal hearing to attend to a story that is read to them at a quiet but detectable level. Failures in maintaining attention to long strings of information have been explored in the visual domain, i.e. when reading (Smallwood et al., 2008). Participants read a Sherlock Holmes novel and the impact of self-reported mind wandering on memory and understanding was explored. For fact recall, mind wandering at the point in which the fact was revealed, was related to the inability to successfully answer the fact-based question. In addition, participants were asked to form inferences including the

pseudonym used by the criminal. In readers that performed well on the task, mind wandering was less frequent at the point in which the text revealed a clue about the criminal's true identity. In the current study participants are asked to listen to the Sherlock Holmes story used by Smallwood et al., (2008). For this study the participants are asked to recall facts, repeat back phrases and make inferences regarding the events happening in the story. In Experiment One, participants listened to the story played to them at a normal speaking level (60dB) and their performance was compared against another group who listened to the same story played at a quiet level, i.e., the participants personal lowest detection level, set individually to establish the quietest level at which sentences could be heard fully. If as is expected, listening at a quiet volume introduces greater demands on attention, processing may be compromised as the resources available for memory consolidation may be reduced. We might predict reduced performance in recall of facts, inferences and repetition of sentences. In Experiment Two, to check a methodological interpretation for poorer performance, two new groups were tested using the same sentences, presented in isolation rather than embedded in the story. If poorer performance at quiet volume is simply due to a difficulty in discriminating words we would also predict poorer performance over a short time at the quiet intensity when recalling the individual sentences. If as we predict however, performance is exacerbated by attending over an extended period of time, we would instead expect to see the differences in performance between listening intensity condition to disappear in Experiment Two.

Experiment One

Method

Compliance with Ethical Standards

The authors declare they have no conflict of interests. This study was approved by the School

of Psychology Research Ethics Committee at the University of Lincoln. Informed written consent was received by all participants that took part in the study. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Participants

Forty four participants were tested for this study but the final sample consisted of 40 participants, 34 female and 6 male aged between 18 and 47, 22 ± 5 years (mean $\pm$ SD). All participants were native speakers of British English. Participants were selected from an undergraduate student population, and had not read the Sherlock Holmes story 'The Red-Headed League' (Conan-Doyle, 1892). No hearing difficulties were reported and a full hearing test was carried out to rule out any unidentified hearing loss using the Hughson-Westlake method (Carhart & Jerger, 1959). Any participant who could only hear sounds played at 25dB or above would be excluded. Participants' lowest hearing level ranged from 5 to 20dB. Twenty participants did the task at a 'normal' volume (60dB SPL) (18 females and 2 males aged between 18 and 47, mean 23 ± 7). Twenty participants did the task at their personal lowest detection level (procedure described below), 16 were female and 4 were male aged between 18 and 27, mean 21 ± 3 years. Participants were randomly assigned to either condition. The additional 4 participants' data were removed from the analysis due to answering in a retrospective self-report questionnaire that they were unable to hear the words in the story 'often'. They were excluded to ensure that errors made were not simply an artifact of a failure to hear the story.

Design

Participants listened to an audiobook lasting approximately 50 minutes, and were probed at approximately 5 minute intervals to recall the last phrase heard, followed by a fact based question. At the end of the experiment participants were asked to make inferences about the events in the story. In order to test inference-making ability the final part of the story was not played to the participant where Sherlock reveals the identity of the criminal and the clues to his identification. Participants either listened to the story at the listening level associated with normal speech, (60dB SPL, Pearsons, Bennett & Fidell, 1977) or their personal lowest detection level, set individually to establish the quietest level at which sentences could be heard fully. Participants were seated in a sound proof booth, with ambient sound levels complying with British Standards/European Norm (BS EN) ISO 8253-1:1998 specifications which allow thresholds as low as 0 dB HL to be established. The audiogram from the hearing test provided a starting point to further specify an appropriate level for the low-intensity condition. Because the dB range of speech is more complex than pure tones used in an audiogram, the audibility of the sentences were further defined. For this experiment, it was not only important that participants could detect a sound but identify what it was. To establish the personal lowest detection level, random sentences were selected from another Sherlock Holmes story "Silverblaze" (Conan-Doyle, 1892/2001), read by the same actor as the story in the experimental condition, at the same pace. Because of the variability in speech sounds, rather than presenting single words, participants were presented with whole sentences ranging from 12-16 words, allowing for the variability across frequencies of speech to be heard before participants were asked to make an assessment of the audibility. The level was achieved using a stepwise technique, consistent with methods used in a traditional hearing test for achieving a perceptual threshold. These were played at quiet intensities on a sliding scale of 5dB intervals. The sound levels were normalized using the software Audacity

(Audacity Team, http://audacity.sourceforge.net/). Participants were asked to state if they could hear all of the sentence typing "yes", some of the sentence typing "some", or if they could not hear the sentence clearly at all typing "no". They were then asked to type out the last word as a further clarification. If all of the sentence was heard clearly the intensity was decreased by 10dB until the sentence could not clearly be detected, known as the "5 up 10 down" technique, a stepwise technique used in audiometry. When the sentence was no longer clearly heard the volume intensity was increased by 5dB until a sentence could be heard completely again. Participants heard the same volume level twice for different sentences, and the quiet intensity for the experiment was selected if they stated that they were still able to hear all of the sentence with the "yes" option and correctly repeat the last word on both occasions at that dB level. Around a perceptual threshold there is likely a point where you can still hear most of what was being said, but not everything, making a simple yes/no choice more difficult. The addition of a third "some" category therefore allowed a more cautious measure of assessing if they could hear. A threshold was achieved only with the sentences with the "yes" all was clear response. The response "some" was regarded as not being able to accurately detect the sentence. Participants were then randomly assigned to a condition and listened to the story at one of the two levels, "quiet" or "normal volume"

Text. The story used for this experiment was *The Red-Headed League* (Conan-Doyle, 1892/2001), read by a single speaker, Sir Ralph Richardson. This was the same text used by Smallwood et al, (2008) but presented auditorily rather than visually. It was abridged to a length of approximately 8,400 words, excluding the explanation of the crime at the end. The story was presented through closed cup headphones. The narrative was paused a total of 10 times, approximately every 5 minutes for a question to be presented on the last section of text. The story always stopped naturally at the end of a sentence and varied from 4.49 to 5.24 minutes. This was designed to be regular, but not consistent so that the arrival of the question

was not overly anticipated. The questions appeared visually on the screen immediately after the story had been paused.

The first question was always the instruction to "repeat back as much of the last phrase as you can remember". Participants were made aware of the structure of the questions prior to the task, and were asked if they could not remember much of the last phrase to try and make sure they remembered the last word. An answer box was provided below the question and participants responded using the keyboard. The response box allowed the participant to see their response as they wrote it, and they were able to amend their answer before they submitted their response before proceeding. The questions were self-paced.

The second question was always a fact-based question about the last section of story heard. i.e., answers were described directly in the text (e.g. What is the job Jabez Wilson is required to do for the league?). After this question the participants initiated the next part of the story with a button press. These questions were designed to assess ability to recall facts that were explicitly described in the text,

At the end there were 5 inference-based questions. The answers to the questions were never mentioned explicitly in the abridged version of the story, and so could only be identified by inference of the narrative. These were designed to examine how lapses in attention interrupt the construction of an integrated narrative (Smallwood et al., 2008). The final inference was the key focus of the story, the identity of the criminal (i.e., who was the criminal John Clay earlier in the story?).

At the end of the experiment participants were also given a 5 scale Likert retrospective self-report questionnaire focusing on their subjective experience of the task related to their attention (i.e., My attention drifted from the story) and their hearing (i.e., I could not hear the words in the story) Options ranged from never to very often. The experiment lasted up to an hour and a half.

Results

Is fact recall affected when attending at a quiet volume?

Fact recall was not normally distributed in either condition (Kolmogorov-Smirnov test, *p* < .05), differences were compared with non-parametric tests. In line with expectation, a Mann Whitney U test showed that significantly more facts were remembered correctly in the normal intensity condition (median = 9) than in the quiet intensity condition (median = 7.5), U= 125.00, z =-2.08, p = .04, r = -.33, (Figure 1).

(Figure 1 about here)

Is immediate recall affected when attending at a quiet volume?

The responses for the recall of the last sentence heard were analyzed in three different ways, to separate capacity and accuracy effects. To measure capacity, the average total number of words that were attempted was calculated irrespective of whether the words correctly matched the transcript. No significant difference was observed in the total words attempted in the normal intensity condition (12.08 ±3.98) compared to the quiet intensity condition (10.29± 3.28) [t(38) =1.56, p = .126, r = 0.25, 95% CI, -.53, 4.10]. To measure accuracy, the total number of words recalled correctly was also calculated. There was no significant difference shown in the average number of correctly recalled words at normal intensity (10.12 ± 2.64) compared to quiet intensity (8.56 ± 2.52) [t(38) =1.91, p = .064, r = .30, 95% CI, -.09, 3.21]. However, the trend is in the same direction as predicted. To further explore accuracy, the average number of words was also calculated considering only the number of absolute correct words in sequence starting from the end of the sentence. In support of the prediction significantly fewer words were recalled correctly in order from the end of the sentence in the quiet intensity condition (3.82 ±1.97) than in the normal intensity

condition (5.54 ± 1.84) [t (38) =2.85, p = .007 r = .42, 95% CI, -.49, 2.93]; see Figure 2

(Figure 2 about here)

Is inference forming affected when attending at a quiet volume?

Inference recall was not normally distributed in either condition (Kolmogorov-Smirnov test, *p* < .05), differences were compared with non-parametric tests. Contrary to prediction, there was no significant difference in the number of correct inferences that were reported between the volume intensity conditions (U= 141.50, z = -1.61, p = .11, r = -.25). However, the number of correct inferences was in the same direction as that of the 'fact recall'. In the normal intensity condition, median correct was 3 of 5 and in the quiet intensity condition the median was 2. The fifth inference required putting together a number of smaller inferences, and was the key to solving the crime i.e., the identity of the criminal. Whether the overriding inference was affected by the experimental manipulation was examined. Pearson's chi-square indicated that participants' performance on the critical question was not affected by the intensity condition [χ2 (1,40) 0.42, p >.250]. Of those participants attending to the story at normal intensity 13 out of 20 (65%) successfully identified the villain. In the quiet condition 11 out of 20 (55%) of participants correctly identified the villain.

Does retrospective self-report correlate with task performance?

Self-perception of performance and approach to the task was assessed in a Likert style retrospective questionnaire with responses corresponding to never (1), rarely (2), sometimes (3), often (4) and very often (5). Spearmans correlations were carried out to assess the relationship between performance on the task across all participants, and their retrospective perceptions of their attention and hearing abilities. Correlations are summarised in Table 1.

Participant's retrospective self-report of attentional drift from the task was correlated with their ability to successfully recall facts, and to form inferences. There was no correlation with words accurately recalled in sequence. Participants were also asked for their perception of their ability to hear the words in the story. Report of not being able to hear the words in the story was not correlated with fact, or inference recall, but was correlated with ability to accurately repeat back words in the correct sequence.

(Table 1 about here)

To investigate if there were any differences in responses between the two intensity levels, Mann Whitney tests compared condition with responses from the self-report. There was a significant difference for the question 'I could not hear some of the words'. Participants at normal volume reported this less frequently with the median corresponding to 'Never' on the scale (Median = 1 Range =2) compared to 'sometimes' at the quiet intensity (Median = 3 Range =2) U= 51.50, z =-4.26, p < .001, r = -.67. There was no self-reported difference in attention drift between the two groups, U= 178.00, z = -.66, p >.250, r = -.67.

Discussion

Volume level affected the participants' ability to recall both facts and sequential words from the last section of text. Fewer facts and fewer sequential words were successfully recalled when attending at a quiet level. This would be consistent with the interpretation that the stimuli were never fully processed. This would happen if as hypothesized, attentional resources must be allocated to support perceptual processing and so there is less capacity for encoding a memory trace (McCoy et al., 2005). It may be especially critical in auditory

processing as sound information is progressive, unlike a visual sentence there is no opportunity beyond the echoic trace, to revisit the auditory sentence just presented as you can when reading (Smith & Pither, 1993). This problem is further compounded by the fact that processing low volume auditory stimuli takes longer (Roebuck et al., 2015) than those at normal volume. So, sustained allocation of attentional resources over a longer period of time is likely to be needed. This may not be possible given the need to allocate attention for the perceptual processing of the following sentence.

In contrast, there was no difference between intensity levels in the ability to form inferences about the story. Such a finding exemplifies that attending over an extended period of time to quiet but detectable information, may superficially appear not to impair performance. The meaning of the story is still integrated into a narrative, suggesting participants have not switched off their engagement entirely just because it is difficult to hear. These findings suggest that lapses in attention did not interrupt the understanding of the story as a whole but instead affected momentary details. Inferences are formed by piecing together several bits of information and being able to integrate the relevance. Whilst it may be dependent on remembering multiple pieces of information recalling these may be facilitated by the fact that they are not dependent on recalling a very specific detail which may only be referenced in a single moment i.e, knowing that Jabez Wilson found out that the Red Headed League had been dissolved by a note on the door. As previous research has demonstrated, drifts in attention at key moments in narrative significantly affect the ability to recall factual information if attention lapsed at the critical time point of the information, rather than the ability to recall information entirely (Smallwood et al., 2008). Inferences are also pertinent to the object of the story and thus attending to these details likely have conceptual salience, which may place more significance in both attending to and remembering these details.

In light of the current findings, it becomes comprehensible how the consequences of

mild hearing impairments may be easily missed because processing can still be impaired even when on the surface it may be possible to 'hear'. Contrary to expectation the hypothesized lack of resources for processing in the low volume condition did not reflect in the number of words participants attempted to recall, or how accurate they were on these words overall. Here there was no difference in attempted or accurate recall between intensity conditions. These findings suggest at least superficially, participants were able to continue to recall the story at both the quiet and normal intensities. To be able to repeat back words immediately does not imply that those words have also been fully encoded for meaning, they may be held in a short term phonological store without further processing (Badderley, 2000). In another study with elderly adults with hearing loss, the participants had poorer memory for words in a list than their normal hearing counterparts, even though the words could still be heard at the time of presentation (Rabbitt 1991). In the context of the current study poor fact recall alongside high sentence recall was observed in participants listening at a quiet intensity but with normal hearing. These findings support the notion that attentional resources are being successfully allocated to perceptual processing, but with greater processing effort to identify quiet information there is less capacity for encoding a memory trace. Despite high accurate recall, sequential recall was still affected by listening at the quiet intensity. Other studies testing list recall in participants with hearing loss have shown that while the final three words were recalled comparably to normal hearing listeners, words earlier in the sentence were not (McCoy et al., 2005). The sequential accuracy score in the current study may be a reflection of a smaller recency effect associated with greater processing effort.

It is also possible that the high recall between groups is being supplemented by the previous context of the story. Speech identification tasks show that prior context aids listeners in identifying the final word of a sentence embedded in noise (Pichora-Fuller, Schneider & Daneman, 1995). In this task participants are following the context of a continuous story, and

may be able to use this context to facilitate filling a misheard/misremembered word. The context of the preceding story, as well as the language used within it, may assist in inferring the real word, or recalling a semantically viable replacement. This suggestion is supported by the fact that participants often incorrectly recalled words in a sentence but replaced them with a word of a similar meaning. It is therefore difficult to establish how many of the recalled words were heard or constructed. As sequential recall is the most stringent measure of recall ability, a 'filled' word counts against the score by interrupting the sequence of accuracy. Given that the sequential recall measure is the only recall measure that is affected between intensity levels it seems likely that participants are using this strategy to assist their high recall at the quiet intensity. Considered alongside the capacity limited theories of processing, it is possible that listening at the quiet intensity influences the most stringent measures (i.e, sequential accuracy and fact recall), but not the less stringent ones (total recall and inference formation).

In the retrospective self-report questionnaire when asked about the frequency in which they 'could not hear some of the words', participants' modal response in the quiet condition was 'sometimes'. This may reflect (1) a direct reflection of physically detecting the sentences, (2) difficulty in detecting the sentences because of the availability of attention, or (3) a consequence of trying to attribute a cause to their perceived poor performance. Given the precautions taken, it seems likely that this impression of poor quality stimuli is a consequence of a lack of attentional resources. Indeed, there was no significant correlation with reports of not hearing the words and the ability to recall the number of facts. Lack of correlation here would suggest that poor fact recall in the quiet condition was not because they were unable to hear the words. Moreover, both fact and inference recall were significantly correlated with perceptions of attentional drift from the task. Together these findings, support the hypothesis that it is more difficult to sustain performance when

attending at a quiet but detectable level. To verify this interpretation the alternative explanation should be excluded, i.e. the volume is set too low for accurate response with the sentences used. Experiment Two tested whether participants have comparable performance at both intensity levels when the need for sustained attention is only required for a short predictable period. This method will allow perceptual ability to be assessed for the sentences used with reduced demands on attentional availability.

Experiment Two

Methods

Participants

Forty new participants aged between 18 and 58 (32 female, 8 male, mean age $\pm$ SD = 22 ± 8 years) took part in the study, 20 in the normal intensity condition and 20 at the quiet intensity (levels were set and inclusion criteria were the same as described in Experiment One).

Procedure

Text. The story used for this experiment was *The Red-Headed League* (Conan-Doyle, 1892/2001), read by Sir Ralph Richardson. In Experiment One participants were asked to recall 10 different sentences after a period of approximately 5 minutes of listening, over 50 minutes. In Experiment Two the same 10 sentences were used, but the sentences were presented in isolation without the rest of the story so that a shorter period of listening was required before testing recall. The order of the presentation was randomized between participants. The story was presented auditorily through headphones. Participants either completed the task at normal listening volume associated with speech levels set at 60dB (Pearsons et al., 1977) and lowest detection threshold (for each individual participant), the

lowest volume level at which all targets could just still be heard and repeated back with accuracy (as Experiment One).As in Experiment One, prior to the task, participants were given a hearing test in a soundproof booth. Participants then had to listen to the 10 sentences in one of two conditions, normal listening volume (60dB) or low but detectable listening volume. Low volume level was established using the same criteria as Experiment One. Participants then listened to the sentences one at a time presented in a random order at the specified dB level. After the sentence was presented the question appeared visually on the screen with an answer box below for them to type the answer using the keyboard. The instruction was consistent with Experiment One, to simply "repeat back as much of the last phrase as you can remember". The questions were self-paced and participants typed their answers using the keyboard. The answer box allowed the participant to see the whole of their response as they wrote it and they were able to amend their answer before they submitted their response to go on to the next sentence. The presentation of the following sentence was initiated by a button press from the participant. The experiment lasted approximately 10 minutes.

Results

Is direct recall affected when attending at a quiet volume for the immediate probe task in Experiment Two?

In line with expectations there was no difference in the total number of words attempted at normal (15.59± 4.31) and quiet intensities (14.10 ± 3.50) [t (38) =1.20, p = .238, r =.18, 95% CI, -1.03, 4.01] in the immediate probe task in Experiment Two. There was also no significant difference in the total correct words recalled at normal (12.81 ± 2.69) and quiet intensities (11.91 ± 2.58) [t (38) =1.70, p = .290, r =.27, 95% CI, -.79, 2.58]. Critically, total words recalled in sequence also displayed no significant difference between the normal (4.49

±1.96) and quiet intensity conditions (3.51 ± 1.86) [t (38) =1.60, p =.114, r =.25, 95% CI, -.84, 3.20].

Is direct recall affected when attending for the immediate probe task (Experiment One) compared to attending within the story (Experiment Two)?

Performance across experiments was compared. As predicted at the quiet intensity the total number of words attempted were higher in this immediate probed task (Experiment Two) (14.10 ± 3.50) than the quiet intensity task in Experiment One [t (38) =3.55, p = .001, r =.49, 95% CI, -1.63, 5.98] Figure 3A. This finding was also reflected in accuracy; more correct words were recalled in Experiment Two at the quiet intensity (mean = 11.91, SD = 2.58) than at the quiet intensity in Experiment One (8.56 ± 2.52) [t (38) =4.15, p <.001, r =.56, 95% CI, 1.72, 4.98] Figure 3B. Notably, there was no significant difference in the number of words correctly recalled in sequence between Experiment Two at the quiet intensity (3.51±1.86) and the quiet intensity in Experiment One (3.82 ± 1.97) [t (38) =-0.51, p = >.250, r = .25, 95% CI, -1.54, .92] Figure 3C.

This pattern of results was also found at the normal intensity. The total number of words attempted were higher in Experiment Two at the normal intensity (mean = 15.59, SD = 4.31) than the normal intensity condition in Experiment One (mean = 12.08, SD = 3.93) [t (38) =2.69, p = .011, r =.42, 95% CI, .86, 6.15] Figure 3A. More correct words were recalled at the normal intensity in Experiment Two (12.80 ± 2.69) than in the normal intensity condition in Experiment One (10.12 ±2.62) [t (38) =3.183, p = .003, r = .46, 95% CI, .98, 4.39] Figure 3B. There was no significant difference in the number of words correctly recalled in sequence between Experiment Two at the normal intensity (4.49 ± 1.96) and the normal intensity in Experiment One (5.54 ± 1.84) [t (38) =-1.74, p = .091, r =.27, 95% CI, -.2.26, .17] Figure 3C.

(Figure 3 about here)

Discussion

Experiment Two was designed to check a methodological interpretation for poorer performance, i.e., not being able to discriminate the words in the sentences because the volume was too low for identification. The results show that more words were recalled correctly in Experiment Two at both intensities than Experiment One. Here the impact of a longer period of listening in Experiment One can be seen. In Experiment Two higher recall was predicted because participants are probed immediately after the end of the phrase, recall at a specific time is highly anticipated, and focused attention is only required for a short duration relative to the story task. Shorter bouts of listening may also give the participant time to compensate for the difficulties in formulating the perceptual objects (Alain & Tremblay, 2007).

The results from Experiment Two show no significant differences in recall between the listening conditions. Total recall, correct recall and sequential recall were comparable at both the quiet and normal intensities. These findings support the view that the impaired performance represented in Experiment One is not a reflection on an inability to 'hear' and physically detect the sentences in any absolute sense. The findings also confirm that the method of setting the volume in Experiment One allowed accurate detection, when attention was maximally directed. Interestingly however, while there were enhancements in Experiment Two for both total recall, and correct recall, enhancement in performance was not observed for words reported in the correct sequence. Poorer sequential recall would not be predicted with a shorter period of attending. In Experiment Two however, participants heard phrases out of sequence but also without any knowledge about where the phrase was coming

from. Therefore replacing a misheard or misremembered word could not be drawn from a logical scenario (as in Experiment One). As the context of the rest of the story is absent in Experiment Two this gives further support to the interpretation that perceptual filling was being used in assisting 'recall' of words (Shinn-Cunningham & Wang, 2008) in Experiment One. Therefore shortening the listening duration in Experiment Two may not have exclusively reduced demands on attentional availability as it also implicitly reduces available contextual cues to facilitate recall. Participants had greater capacity to recall more words in the short task, but they did so at the cost of recalling words accurately before inferring an incorrect word.

Overall the results suggest that attending over a long period to a quiet story had a negative impact on processing. In Experiment One, recall of specific facts was significantly reduced and accurate sequential recall was affected. Listening condition did not affect inference formation or immediate recall. Together, these findings suggest that listening at a quiet level affects the ability to attend to specific aspects of the story, but does not appear to affect the understanding of the story as a whole (beyond the lapses of attention expected when the stimuli are clearly detectable). Experiment Two confirmed that the stimuli were physically detectable with the current method. The results suggest that in Experiment One, the reduced performance was not due to an underlying physical inability to detect the sound when attention was focused on the task. The effects may not be obvious over a short period, but emerge when greater demands are placed on attention. Hearing and listening are often intuitively thought of as part of one and the same thing, however, these findings suggest that to be able to identify and repeat information does not necessitate that the information is also fully processed/remembered for meaning.

These findings have significant implications for sub-optimal hearing in the real world where continuous speech will require sustained attentional support. Here the cognitive

consequences are indicated for populations with mild hearing losses. It hints at more substantial impact than a slight perceptual decrement might suggest. For children the relationship between hearing loss and behavioral problems becomes comprehensible. For everyone, mild hearing reduction appears to create novel cognitive challenges. When assessing listening abilities, it is likely to be beneficial to assess performance in tests which reflect everyday listening demands. These difficulties may not be apparent in standard measures of hearing and listening.

References

Alain, C., & Tremblay, K. (2007). The role of event-related brain potentials in assessing central auditory processing. *Journal of the American Academy of Audiology*, *18*, 573–589. American Speech-Language-Hearing Association (1997). *Guidelines for Audiological Screening*. American Speech-Language-Hearing Association. Retrieved from http://www.asha.org/policy/GL1997-00199/

Bess, F. H., Dodd-Murphy, J., & Parker, R. A (1998). Children with minimal sensorineural hearing loss: prevalence, educational performance, and functional status. *Ear and Hearing*, *19*(5), 339–354.

Carhart, R., & Jerger, J. (1959). Preferred method for clinical determination of pure-tone thresholds. *Journal of Speech & Hearing Disorders*, *24*, 330–345.

Conan-Doyle, A. (2001). *Adventures of Sherlock Holmes.* In E. Glinert (Ed.), *The adventures and memoirs of Sherlock Holmes*. London: Penguin. (Original work published 1892)

Flexer, C., Millin, J. P., & Brown, L. (1990). Children With Developmental Disabilities The Effect of Sound Field Amplification on Word Identification. *Language, Speech, and Hearing Services in Schools*, *21*(3), 177-182.

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

Pakulski, L. A., & Kaderavek, J. N. (2002). Children with Minimal Hearing Loss Interventions in the Classroom. *Intervention in School and Clinic*, *38*(2), 96-103.

Pearsons, K. S., Bennett, R. L., & Fidell, S. (1977). Speech levels in various noise environments. *Office of Health and Ecological Effects, Office of Research and*

Development, US EPA.

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

Pringle, M. B., Thompson, A., & Reddy, K. (1993). A comparison of speech audiometry and pure tone audiometry in patients with secretory otitis media. *The Journal of Laryngology & Otology*, *107*(09), 787-789.

Quigley, S., & Thomure, R. (1968). *Some effects of hearing impairment upon school performance.* Urbana, IL: University of Illinois. Institute for Research on Exceptional Children.

Rabbitt, P. (1991). Mild hearing loss can cause apparent memory failures which increase with age and reduce with IQ. *Acta oto-laryngologica*, *111*(sup476), 167-176.

Roebuck, H., Guo, K., & Bourke, P. (2015). Attending at a Low Intensity Increases Impulsivity in an Auditory Sustained Attention to Response Task. *Perception*, *44*(12), 1371-82.

Shinn-Cunningham, B. G., & Wang, D. (2008). Influences of auditory object formation on phonemic restoration. *The Journal of the Acoustical Society of America*, *123*(1), 295.

Smallwood, J., McSpadden, M., & Schooler, J. W. (2008). When attention matters: the curious incident of the wandering mind. *Memory & Cognition*, *36*(6), 1144–50.

Smith, P. R., & Pither, R. E. (1993). Long term follow-up of children with mild hearing impairment: pre and post intervention. *Journal of the British Association of Teachers of the Deaf*, *4*(17), 99–103.

Teasdale, T. W., & Sorensen, M. H. (2007). Hearing loss in relation to educational attainment and cognitive abilities: A population study: Hipoacusia en relación con los logros educativos y las habilidades cognitivas: estudio en una población. *International journal*

of audiology, *46*(4), 172-175.

Ward, W. D. (1957). The Single-Descent Group Audiometer. *Noise Control*, *3*(3), 15-18.

Wertheimer, M. (1955). The variability of auditory and visual absolute thresholds in time.

The Journal of General Psychology, *52*(1), 111-147.

Williams, C. R., & Abeles, N. (2004). Issues and implications of deaf culture in therapy.

Professional Psychology: Research and Practice, 35, 643.

Wingfield, A., Tun, P. A., & McCoy, S. L. (2005). Hearing loss in older adulthood what it is

and how it interacts with cognitive performance. *Current directions in psychological*

science, *14*(3), 144-148.

Table 1. Summary of r and *p* values (two-tailed) for Spearmans correlations for task performance, and responses on the retrospective self-report questionnaire. Significant values after Holm-Bonferroni adjustment for multiple comparisons have been highlighted in bold.

Items Recalled	"I could not hear some of the words"	"My attention drifted from the task"
Facts	-.31; = .050	**-.46; = .003**
Inferences	-.31; = .049	**-.36; = .022**
Words in sequence	**-.39; = .012**	-.29; = .072

Figure Legends

Figure 1. Box plots summarizing facts recalled in the normal and quiet intensity conditions. Black bars indicate median scores, the boxes incorporate the interquartile range, and the whiskers show the range of scores. Outliers are identified as falling outside 1.5 (°) times the interquartile range.

Figure 2. The mean number of recalled words attempted, correct and correct in sequence at normal and quiet intensities.

Figure 3. Recall in Experiment One and Experiment Two at normal and quiet intensities, 3A, Mean attempted words, Figure 3B, Mean correct words, Figure 3C, Mean correct words in sequence.