

COVER PAGE

Document downloaded by @DAEL

Sat Nov 28 09:04:13 2026
For personal use

When automatic English translation is provided, only the original document is authentic.
The EAA cannot be held responsible of any translation error

Bibliographical reference

Effects of Hearing Loss on Maintaining and Switching Attention, Adrian K. C. Lee, Eric Larson and Christi W. Miller, Acta Acustica **vol. 104** (Number 5), 2018, pp. 787-791

DOI
https://doi.org/10.3813/AAA.919224

Acta Acustica united with Acustica
Vol. 104 (2018) 787 – 791

DOI 10.3813/AAA.919224

Effects of Hearing Loss on Maintaining and Switching Attention

Adrian K. C. Lee[1,2], Eric Larson[2], Christi W. Miller[1]

[1] Department of Speech and Hearing Sciences, University of Washington, Seattle, WA 98195, USA.
aklee@uw.edu

[2] Institute for Learning and Brain Sciences, University of Washington, Seattle, WA 98195, USA

Summary

The ability to intentionally control attention based on task goals and stimulus properties is critical to communication in many environments. However, when a person has a damaged auditory system, such as with hearing loss, perceptual organization may also be impaired, making it more difficult to direct attention to different auditory objects in the environment. Here we examined the behavioral cost associated with maintaining and switching attention in people with hearing loss compared to the normal hearing population, and found a cost associated with attending to a target stream in a multi-talker environment that cannot solely be attributed to audibility issues.

© 2018 The Author(s). Published by S. Hirzel Verlag · EAA. This is an open access article under the terms of the Creative Commons Attribution (CC BY 4.0) license (https://creativecommons.org/licenses/by/4.0/).

PACS no. 43.66.-x

1. Introduction

Listeners with sensorineural hearing loss (HL) often experience difficulty understanding speech in the presence of competing talkers [1]. This observation has been well documented clinically [2] and is supported by both subjective ratings [3] and by empirical studies in laboratory settings [4]. Yet, uncovering the root cause of this listening difficulty in everyday settings has been elusive. Undoubtedly, audibility plays an important role in determining speech recognition ability [5]; but when audibility has been statistically accounted for [6], the influence of other cognitive factors on listeners' performance emerges (e.g., working memory [7]). A meta-analysis of twenty studies investigating the relationship between speech recognition in noise and cognitive abilities concluded that while there is some link between cognitive performance and speech reception, its exact nature is unclear [8]. This is perhaps complicated by the fact that the relationship of speech intelligibility to hearing and cognition varies for different speech perception tests [9], and thus disentangling the impact of HL from age-related cognitive factors is challenging [10]. Here, by only incorporating over-learned stimulus tokens in an established dual-stream paradigm [11, 12] that has previously been used to study maintaining and switching attention in listeners with normal hearing (NH), we investigated whether there is a HL-related behavioral cost associated with attention deployment.

Received 1 May 2018,
accepted 23 August 2018.

2. Methods

2.1. Subjects and inclusion criteria

Subjects were required to speak English fluently, have no active otologic disorders, and have symmetrical hearing with no conductive involvement (i.e., interaural hearing threshold levels within 15 dB and air-bone gaps within 10 dB). All subjects also underwent a cognitive screening, and they had to score 26 or better on the Montreal Cognitive Assessment [13] for inclusion to reduce the likelihood of including subjects ranging from mild to severe cognitive impairment.

Inclusion criteria for the NH group included hearing threshold levels of 25 dB HL or better from 250-8000 Hz. Inclusion criteria for the HL group included hearing threshold levels of 55 dB HL or better from 250–2000 Hz, 70 dB HL or better from 2000-4000 Hz, and unrestricted from 4000–8000 Hz. Ten subjects with sensorineural HL and 10 age-matched (±2 years) NH listeners participated, between the ages of 18 and 64 years. All subjects gave informed consent to participate in the study as approved by the University of Washington Institutional Review Board.

2.2. Dual-stream task

Subjects performed a task where they pressed a response button as quickly as possible whenever the target (one of two simultaneous talkers) said the letter "O" (Figure 1). Each talker spoke streams of four consecutive letters, with a 200, 400, or 600 ms silent gap (δ) separating the first two from the last two letters. One talker was male, the other female (see 2.4 for more information). Subjects were cued

© 2018 The Author(s). Published by S. Hirzel Verlag · EAA.
This is an open access article under the terms of the CC BY 4.0 license.

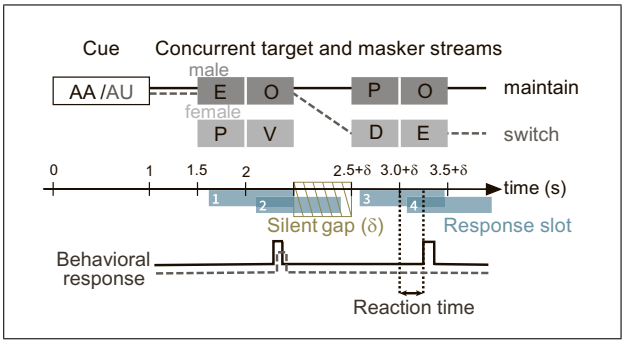

Figure 1. The dual-stream paradigm, with correct example maintain and switch attention responses below. The gap duration (δ) can be 200, 400, or 600 ms.

at the start of each trial by the initial target talker saying "AA" or "AU." "AA" indicated to attend to that talker for the entire trial (maintain-attention), and "AU" to switch attention to the other talker after the gap (switch-attention). An auditory cue was used in order to keep the visual field constant (fixation dot).

2.2.1. Training

Three training block types ensured that subjects understood and could perform 1) maintain-attention trials (only), 2) switch-attention trials (only), and 3) a mixture of maintain- and switch-attention trials. Each block type was repeated until the subjects were able to perform the task with a 100% hit rate and 0% false alarm rate on 6/10, 6/10, and 4/10 trials, respectively. The time for this task varied across participants depending on their performance.

2.3. Single-stream control task

Prior to performing the training and dual-stream task, subjects performed a single-stream control task. Individual letters (0.5-1.5 sec random inter-stimulus interval, intermixed male and female talkers) were presented, and subjects were instructed to press a button as soon as possible following any "O". The randomized stream consisted of 96 "O" and 155 non-target letters (see 2.4 below). This task served three purposes: 1) ensure task-specific audibility, i.e., ability for all our subjects to identify the target letter presented in isolation, 2) familiarize each subject with the stimuli and response interface, and 3) provide an individualized (baseline) single-stream reaction-time.

2.4. Stimuli

Stimuli were generated using tokens from one male and one female talker in the ISOLET v1.3 corpus [14], chosen such that all letters are as close as possible to, but no longer than, 500 ms in duration (trimmed of leading and trailing silence, and downsampled to 24414 Hz, 16 bits). Each letter was matched in intensity. The subset of letters used were {A, U} (cue); {O} (target); and {D, E, G, P, V} (non-target), with the mean pitch for each letter ranging from 108–131 Hz (male) and 180–202 Hz (female).

The stimulus for each dual-stream trial was comprised of a 2-letter cue, 500 ms of silence, then two simultaneous sequences of four letters each. A silent gap of 200, 400, or 600 ms was inserted between the first two and last two simultaneously spoken letters. "O" stimulus onsets were spaced by at least 1000 ms to ensure that reaction times could be disambiguated. There were 228 dual-stream trials and 38 trials per condition (maintain/switch x 3 gaps). In each trial, the probability of an "O" in each stimulus slot was 0.6, 0.4, 0.59, and 0.41, and non-target letters were randomly chosen. The 228 trials were broken into 10 blocks (24 trials for each block except blocks 9 and 10, which had 18 trials) such that each block contained an even distribution of trials from each condition. This took roughly 24 minutes independent of breaks.

Stimuli were presented diotically through Etymotic ER-2 insert earphones via a TDT RP2 real-time processor (Tucker Davis Technologies, Alachula, FL) to subjects seated in a sound-treated room. For NH listeners, we presented stimuli at 65 dB SPL (resulting in 39 to 52 dB SL). For HL listeners, we used an audible and comfortable listening level, aiming for a 20-30 dB SL referenced to their 500-4000 Hz pure tone average. Stimuli were presented at 65 dB SPL if the resulting SL was at least 20 dB, otherwise the level was increased to 75 dB SPL. For 6 listeners, 20 dB SL was not comfortably loud enough, so 75 dB SPL was used (yielding 32-43 dB SL). For 1 listener, 20 dB SL was too loud, so 10 dB SL (90 dB SPL) was used.

2.5. Behavioral measures

To compute d′ and reaction times, responses between 100 and 1000 ms after the onset of a target "O" were considered "hits" (and for a masker "O," as false alarms). By this definition, false alarms are taken to be when a listener assigns a letter to the incorrect stream (i.e., presses to an "O" in the masker stream) rather than when they fail to identify a letter in the target stream (e.g., press to the occurrence of an "E" in the target stream that they thought was an "O"). Only button presses in response to target "O" stimuli were used to compute reaction times. Normalized dual-stream values (here called ΔRT) were computed by subtracting the one single-stream (control) task reaction time value for each subject from all of their dual-stream reaction times.

2.6. Clinical measures

We performed a battery of clinically-validated tests related to listening in crowded environments, including the Multimodal Lexical Sentence Test in 4-talker babble (MLST; [15]), the short form of the Speech, Spatial and Qualities questionnaire (SSQ-12; [3, 16]), the Abbreviated Profile of Hearing Aid Benefit (APHAB; [2]), and pure-tone average (PTA).

3. Results

3.1. Single-stream (control) task

In single-stream (control) task, there was no significant difference between the NH and HL groups in error rate (two-sample t-test $p = 0.800$) – which was $< 6\%$ for all

subjects – or RT ($p = 0.733$), and no significant correlation between age and RT ($p = 0.239$).

3.2. Dual-stream task

In training, all subjects successfully learned the dual-stream task, taking at most 5 total blocks to pass the criteria for all 3 types of training blocks. There was no significant difference between NH vs. HL in the total number of training blocks, or number of training blocks of each type (t-test $p > 0.05$).

To evaluate task performance, we used a 4-way ANOVA (within factors: attention condition, gap duration, response slot; between factor: NH vs. HL) on both d′ and ΔRT. We found significant main effects on d′ and ΔRT of attention ($p < 0.001$ and $p = 0.002$, respectively), gap ($p \leq 0.001$ both), and slot ($p < 0.001$ both). We also found a significant main effect of NH vs. HL for ΔRT ($p = 0.005$). There were significant interactions for d′ and ΔRT of slot×attention ($p < 0.001$ both), slot×gap ($p < 0.001$ both), and slot×attention×gap ($p = 0.010$ and $p = 0.007$), but none involving NH vs. HL ($p > 0.05$ all).

Due to the significant main effect of gap duration, post-hoc paired t-tests were performed between each gap duration (collapsing across the other three dimensions). We found that 200 ms was significantly greater than 400 and 600 ms for ΔRT ($p < 0.005$ each, Bonferroni corrected), and that 200 and 400 were significantly less than 600 ms for d′ ($p < 0.001$ and $p = 0.015$ respectively). Because this was largely consistent with our previous results [11, 12], performance is shown in Figure 2 collapsed across gap durations.

Due to the significant main effect of slot, post-hoc paired t-tests were performed between all response slots (collapsing across the other three dimensions). There were significant (Bonferroni corrected $p < 0.05$) differences between all pairs of slots for d′ and ΔRT, except for d′ slot 3 vs 4 and ΔRT slot 2 vs 4.

To compare the dual- and single-stream tasks, we tested whether the reaction times were significantly different—i.e., whether ΔRT was significantly different from zero—as a function of slot, attention condition, and NH/HL while collapsing across gap durations. For NH, only slot 1 in the switch-attention condition was significant at the $p < 0.05$ level (Bonferroni corrected); for HL listeners, both attention conditions were significantly different for slots 1 and 3 (Figure 2).

We examined speed-accuracy trade-offs by correlating d′ with ΔRT and with single-stream (control) reaction time across subjects (collapsing attention and gap; using slot 3 as it was the most relevant for the switch condition), but these were not significant.

3.3. Exploratory correlations

We computed Pearson correlation coefficients for ΔRT (in slot 3, collapsed across gaps) in maintain and switch conditions as well as single-stream (control) RT with four clinical measures: PTA, SSQ, APHAB, and MLST. The

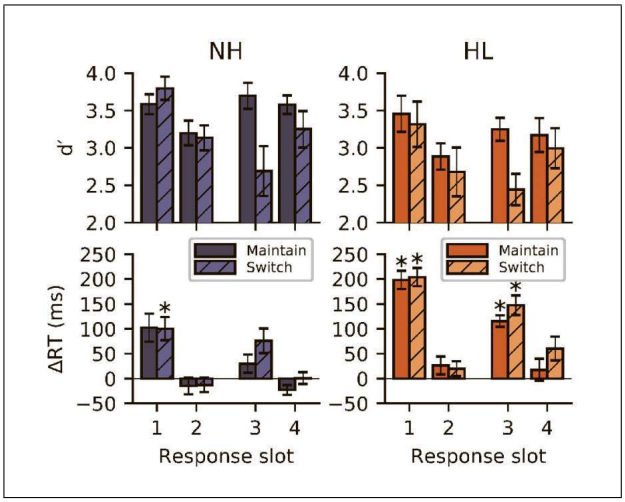

Figure 2. NH and HL dual-stream task performance for maintain- and switch-attention trials (dark solid and light hatched bars) in each response slot (abscissa). ΔRT asterisks indicate significant differences between the dual- and the single-stream reaction times. Although we measured responses for three gap durations, these have been collapsed here as the trends followed our previously published results [11].

only significant (uncorrected) correlations were PTA with $\Delta RT_{maintain}$ ($r = 0.525$, $p = 0.017$) and ΔRT_{switch} ($r = 0.609$, $p = 0.004$), and MLST in the auditory-only condition presented at 8 dB target-to-masker ratio ($r = 0.502$, $p = 0.024$), but these were not significant for the HL group by itself ($p > 0.7$, all).

In a post-hoc exploration, we created an index to quantify the difficulty of selecting a talker relative to following a talker in the attended stream by computing an index over ΔRT values in the switch-attention condition: index = $(slot_1 + slot_3) - (slot_2 + slot_4)$. This index was not correlated with the single-stream (control) RT ($p = 0.321$). For clinical measures, this index was not correlated with MLST ($p > 0.297$ for all four conditions), but was correlated with PTA ($r = 0.603$, $p = 0.005$) and SSQ ($r = -0.479$, $p = 0.033$). This index was also correlated with the APHAB (unaided) global score ($r = 0.595$, $p = 0.006$) and its sub-scales for ease of communication (EC; $r = 0.587$, $p = 0.006$), reverberation (RV; $r = 0.579$, $p = 0.007$), and background noise (BN; $r = 0.518$, $p = 0.019$). Performing correlations for only the HL group, significant correlations remained between this index and PTA ($r = -0.798$, $p = 0.006$), APHAB-EC ($r = 0.699$, $p = 0.024$), and APHAB-RV ($r = 0.639$, $p = 0.047$).

4. Discussion

The observed behavioral difference in ΔRT between NH and HL listeners suggests a greater cognitive effort is associated with the HL listeners attending to an auditory stream in a multi-talker environment. The comparable error rates in the single-stream (control) task and comparable sensitivity in the dual-stream task (d′ > 2.5 in both

groups) together suggest that audibility is not the main factor contributing to the group difference.

Our post-hoc analysis of the ΔRT versus single-stream control RT suggests that NH and HL listeners take longer to respond when they are first exposed to an auditory scene with competition (slot 1) but quickly adapt (slot 2). However, HL listeners struggle again after a silent gap (slot 3) while both groups adapt to the task again in slot 4. In both listening groups, there is a robust behavioral cost associated with switching attention, being particularly salient in the decreased d′ after the silent gap, replicating previous findings [11, 12, 17]. Together, this suggests that HL listeners have a similar switch cost compared to NH listeners, but there is an additional HL-related cost even when just maintaining attention.

Interestingly, none of our clinical test measures meaningfully correlated with the dual-stream ΔRT (i.e., where we observe a behavioral difference between groups). However, a derivative index for the difficulty of selecting a talker (relative to following a talker in an attended stream) did correlate most strongly with APHAB, in particular the ease-of-communication and reverberation sub-scores. These correlations were exploratory and not corrected for multiple comparisons; in future work, a larger sample size would help determine if these correlations persist in the general population.

This pattern of results is most consistent with the idea that listeners with HL experience degradation at the stage of perceptual organization [18, 19]. Future neuroimaging studies may illuminate neural correlates of this observed reaction-time difference, possibly pointing to neurobiological causes of the struggle and fatigue experienced by listeners with HL in these active listening situations. Future studies using letter stimuli could also be designed to use only vowel stimuli to relate more closely to nonlinguistic measures of spectral resolution, such as the one proposed in [20].

5. Conclusion

Using a dual-stream paradigm, we found a behavioral cost associated with HL listeners attending to a target stream in a multi-talker environment that cannot solely be attributed to audibility issues. The neural correlate of this behavioral cost is currently unknown, and may require future neuroimaging studies to illuminate the cortical circuitry involved in attention deployment to compensate for peripheral degradation due to hearing loss. Finally, exploratory correlations suggest that reaction-time measurements in this paradigm should be further investigated in future studies to tease apart aspects of clinical measurements that are potentially related to deficits realted to auditory attention deployment in these listeners.

Acknowledgement

We thank Nicole Whittle for extensive help with data collection. This work was funded by USA NIH-National Institute on Deafness and Other Communication Disorders R01DC013260 (AKCL) and R21DC016380 (CWM).

References

[1] E. Roverud, V. Best, C. R. Mason, J. Swaminathan, G. Kidd Jr: Informational masking in normal-hearing and hearing-impaired listeners measured in a nonspeech pattern identification task. Trends Hear **20** (2016) 1–17.

[2] J. A. Johnson, R. M. Cox, G. C. Alexander: Development of APHAB norms for WDRC hearing aids and comparisons with original norms. Ear Hearing **31** (2010) 47–55.

[3] S. Gatehouse, W. Noble: The speech, spatial and qualities of hearing scale (SSQ). Int J Audiol **43** (2004) 85–99.

[4] N. Marrone, C.R. Mason, G. Kidd Jr: The effects of hearing loss and age on the benefit of spatial separation between multiple talkers in reverberant rooms. J Acoust Soc Am **124** (2008) 3064–3075.

[5] L. E. Humes: Understanding the speech-understanding problems of the hearing impaired. J Am Acad Audiol **2** (1991) 599–69.

[6] D. R. Moore, M. Edmondson-Jones, P. Dawes, H. Fortnum, A. McCormack, R. H. Pierzycki, K.J. Munro: Relation between speech-in-noise threshold, Hearing Loss and Cognition from 40–69 Years of Age. PLoS One **9** (2014) e107720.

[7] T. Lunner, M. Rudner, J. Rönnberg: Cognition and hearing aids. Scand J Psychol **50** (2009) 395–403.

[8] M. A. Akeroyd: Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. Int J Audiol **47** (2008) S53–S71.

[9] A. Heinrich, H. Henshaw, M. A. Ferguson: The relationship of speech intelligibility with hearing sensitivity, cognition, and perceived hearing difficulties varies for different speech perception tests. Front Psychol **6** (2015) Article 782.

[10] H. Meister, S. Schreitmüller, M. Ortmann, S. Rählmann, M. Walger: Effects of Hearing Loss and Cognitive Load on Speech Recognition with Competing Talkers. Front Psychol **7** (2016) Article 301.

[11] E. Larson, A. K. C. Lee: Influence of preparation time and pitch separation in switching of auditory attention between streams. J Acoust Soc Am **134** (2013) EL165–EL171.

[12] D. R. McCloy, B. K. Lau, E. Larson, K. A. I. Pratt, A. K. C. Lee: Pupillometry shows the effort of auditory attention switching. J Acoust Soc Am **141** (2017) 2440–2451.

[13] Z. S. Nasreddine, N. A. Phillips, V. Bédirian, S. Charbonneau, V. Whitehead, I. Collin, J. L. Cummings, H. Chertkow: The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc **53** (2005) 695–699.

[14] R. A. Cole, Y. Muthusamy, M. Fanty, M.: The ISOLET spoken letter database. Technical Report CS/E 90-004 (1990) Oregon Graduate Institute, Hilsboro, OR.

[15] K. I. Kirk, L. Prusick, B. French, C. Gotch, L. S. Eisenberg, N. Young: Assessing spoken word recognition in children who are deaf or hard of hearing: A translational approach. J Am Acad of Audiol **23** (2012), 464–475.

[16] W. Noble, N. S. Jensen, G. Naylor, N. Bhullar, M. A. Akeroyd: A short form of the Speech, Spatial and Qualities of Hearing scale suitable for clinical use: The SSQ12. Int J Audiol **52** (2013) 409–412.

[17] V. Best, E. J. Ozmeral, N. Kopčo, B. G. Shinn-Cunningham: Object continuity enhances selective auditory attention. P Natl Acad Sci USA **105** (2008) 13174–13178.

[18] G. Kidd Jr, T. L. Arbogast, C. R. Mason, M. Walsh: Informational masking in listeners with sensorineural hearing loss. JARO **3** (2001) 107–119.

[19] B. G. Shinn-Cunningham, V. Best: Selective attention in normal and impaired hearing. Trends Amplif **12** (2008) 283–299.

[20] W. R. Drennan, E. S. Anderson, J. H. Won, J. T. Rubinstein: Validation of a clinical assessment of spectral-ripple resolution for cochlear implant users. Ear Hear **3** (2014) e92–e98.