DOI: 10.61782/fa.2023.0656

forum acusticum 2023

INVESTIGATING THE RELIABILITY OF THE EVOKED PUPIL RESPONSE IN LISTENERS WITH HEARING IMPAIRMENT DURING A SPEECH-IN-NOISE TASK

Kirsten Maria Jensen-Rico[1,2*] **Abigail Anne Kressner**[1,2] **Mihaela-Beatrice Neagu**[1] **Dorothea Wendt**[1,3]

[1] Department of Health Technology, Technical University of Denmark, Kgs. Lyngby, Denmark
[2] Copenhagen Hearing and Balance Centre, Rigshospitalet, Denmark
[3] Erikholm Research Centre, Snekkersten, Denmark

ABSTRACT

Pupillometry is often used as a measure of listening effort associated with a task, where an enlarged pupil size is associated with increased mental effort. However, pupillometry has so far been mostly applied using group-level analysis with normal-hearing cohorts. This study investigated the individual reliability of the evoked pupil response among listeners with hearing impairment during a speech-in-noise task. The aim was to study which internal and external factors could impact the reliability of the pupil response in this group (i.e., hearing and cognitive abilities, normalization procedure, signal-to-noise ratio, SNR, and visit number, respectively). For this, two pupil features were measured (peak and mean pupil dilation) across three different visits. The results obtained in the current study showed that with increasing number of visits, the reliability of the pupil features increases, consistent with earlier studies with normal-hearing listeners. Moreover, the changes in the pupil response were reproducible after correcting for the individual's baseline pupil response and the individual dynamic range. There was evidence of a "give-up effect" among listeners with larger hearing loss in the more demanding tasks. These findings make an important contribution to the use of pupillometry in assessment of listening effort on an individual level.

*Corresponding author: kirsten.maria.jensen.rico@regionh.dk.

Keywords: *pupillometry, listening effort, speech-in-noise, hearing-impairment.*

1. INTRODUCTION

Pupillometry has been commonly used as a measure of listening effort during speech-in-noise tests. Listening effort has been defined as "the extent to which the demands imposed by the task at a given moment consume the resources available to maintain successful execution" [1]. Pupil dilation reflects, therefore, the amount of processing demands associated with a task, with pupil size enlarging with an increase of mental effort, cognitive load, and task complexity, as long as the listener is engaged in the task [2, 3]. Although there is not always a direct relation between the correct performance in a listening task and the exerted listening effort, pupillometry may offer insight into the cognitive demands and difficulties patients are experiencing even when they are able to complete tasks with high performance levels [4].

So far, pupillometry has mostly been applied on a group level, and most studies focus on normal-hearing population. In order to explore the potential for being a diagnostic tool for measuring individualized listening effort, there is a need for a more in-depth investigation into this method and its reliability in hearing-impaired listeners. Neagu et al. [5] found that the main factor that affects reliability in normal-hearing subjects is the normalization procedure, regardless of the task demand or number of visits.

2. METHODS

2.1 Participants

Eighteen hearing-impaired participants were recruited from the Technical University of Denmark's database (DTU) for this study and took part in three different experimental visits. The participants were Danish native speakers with ages ranging from 55 to 81 (mean age = 70.83 years, SD = 6.91). The audiometric criterion for the recruitment and inclusion of participants was symmetrical, mild-to-moderately-severe sensorineural hearing loss.

2.2 Stimuli and procedure

For this experiment, sentences from the Danish Hearing in Noise Test (HINT) corpus [6] were played through headphones to subjects while they were sitting inside a soundproof booth. A four-talker babble noise, consisting of two male and two female talkers, was presented at 70 dB SPL. During the audio playback, the pupil response was recorded. The subjects were then asked to repeat the sentence aloud. The intelligibility scores were recorded based on how many correct words were repeated back from 0 to 5. The speech was presented at four signal-to-noise ratios (SNRs) between -3 dB and +6 dB, distributed in steps of 3 dB.

2.3 Pupillometry and data processing

To analyse the reliability of the pupil features, an average baseline for each trial was calculated during the noise-only segment by calculating the mean pupil dilation value, measured 1 s prior to the sentence onset (as suggested by [2]). Two normalization procedures were applied on the pupil data: (i) Baseline-correction, where the baseline is subtracted from the data, and (ii) Baseline-range normalization, where the baseline-corrected data is divided by the overall dynamic range of each individual. Subsequently, two pupil features were extracted for each participant as indicators of listening effort. Peak and mean pupil dilation (PPD and MPD) were computed by taking the maximum and average values, respectively, from the sentence onset.

3. RESULTS

3.1 Normalization procedure

First, an analysis on the pupil dilation over time was done on a group level. The pupil traces from the baseline region (4-5 s), to the noise offset (approximately 9 s) averaged across all participant for each of the four test conditions, and for each of the three visits, are depicted in Fig. 1. A comparison between the two different normalization procedures is also shown.

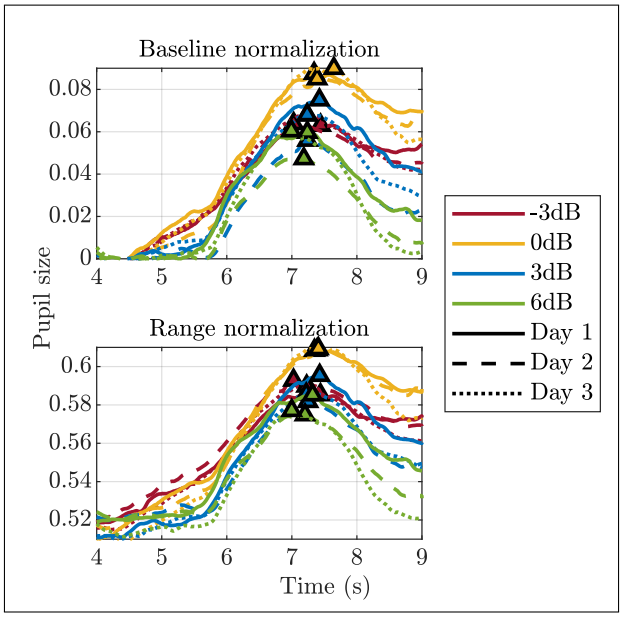

Figure 1. Comparison of pupil traces for baseline-correction and range-normalization. The 4 different SNRs are shown in different colors, and the different test visits are shown with different line-styles.The black triangles denote the PPD point for each curve.

Visual inspection shows a lower pupil dilation in the sentence onset for the conditions with highest and lowest task demand (−3 and 6 dB, respectively). This is aligned with previous research [5] where the giving-up effect can be seen. Additionally, the highest SNR results in the lowest pupil dilation over all four SNR conditions, which would suggest that, on average, this condition demands less cognitive resources.

3.2 Repeatability of pupil features across visits

The repeatability of PPD and MPD for two consecutive visits was calculated by a Spearman correlation. The results for the PPD and their respective coefficients are shown in Fig. 2, wherein the repeatability coefficients for each SNR and overall are shown for visits one and two and for visits two and three, as well as for both normalization procedures. The results are shown only for the PPD

since the trends of the PPD and MPD were very similar.

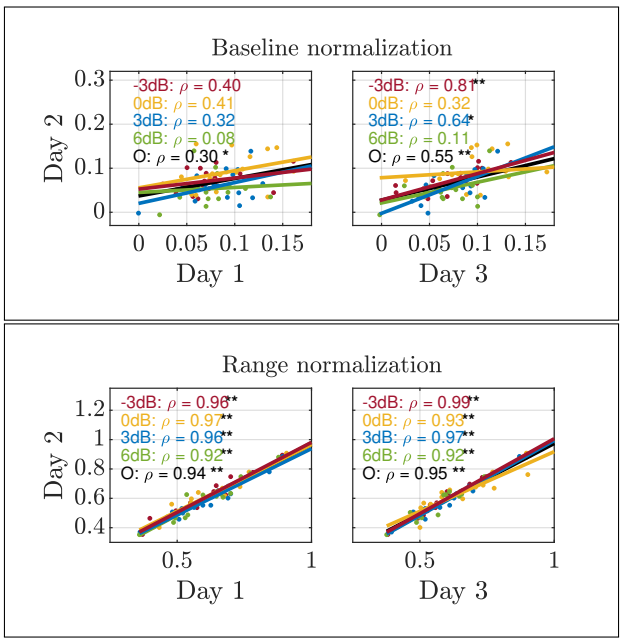

Figure 2. Consistency of PPD for two consecutive visits (day 1-2 and 2-3) for two normalization procedures. The Spearman Correlations are shown for each SNR (colors) and overall, "O", (black). Each dot represents each participant in each condition. *, and **, significant at $p < 0.01$ and $p < 0.001$, respectively.

For the baseline-normalization, the highest repeatability coefficient overall was obtained between the second and third visits, with a moderate correlation, compared to between the first and second visits, in which a weak correlation was observed, for both PPD and MPD. The -3 dB SNR condition was found to have a very strong correlation for the PPD. The PPD and MPD showed an overall weak relationship between the first and second days. No other significant correlations were found for the baseline-normalized data. For the range-normalization, a very strong correlation was found for each SNR condition and overall, for both pupil features and across all visits.

4. DISCUSSION

Regarding the analysis of the pupil traces on a group level, the conditions with lower task demand (6 dB) resulted in a lower pupil dilation. For the most difficult SNR condition (-3 dB), a positive correlation between intelligibility and pupil response was found indicating that a smaller pupil response is associated with lower performance. In other words, this might suggest a "giving up" effect in hard conditions, i.e., participants might disengage when performance drops which is in line with other studies [5, 7, 8]. The other conditions (0 and 3 dB) showed the largest evoked pupil response as was expected since this is around the 50% intelligibility point.

When having a look at the repeatability of the pupil features across the different test visits, the normalization procedure was found to have an impact on the correlations. There was an overall higher consistency of PPD and MPD between the latter visits compared to between the first and second, which could be attributed to learning effects. This means that, with an increased number of visits, there is an increase in correlation. It was expected that the conditions with lowest and highest task demands would give high correlation, but for the intermediate SNRs, there is a change in performance which could also lead to variability and low correlation from one day to another. Range-normalization gave strong correlations for all SNR conditions and overall for two consecutive visits, which would suggest that the normalization according to each participant's individual range could reduce this variability across participant and increase the repeatability of the pupil features.

Because in this study a less homogeneous group is analysed, some listener-specific factors were also measured in order to better understand their potential effect on the reliability of an individual's pupil response. Hearing abilities, working memory and chronic fatigue were some of the factors that were analysed by different questionnaires and additional tests. A correlation analysis between the change in pupil features and the chosen listener-specific factors was performed. Negative correlations were found between the Pure Tone Audiogram (PTA) and the pupil features, indicating lower pupil response with lower hearing abilities. The remaining chosen factors showed weak to non-existing correlations for both PPD and MPD and are therefore not shown.

5. CONCLUSION

This study looked at the reliability of the evoked pupil response of hearing-impaired population during a speech-in-noise task. The purpose was to investigate which factors could impact reliability in this specific group, taking into account not only their hearing abilities, but also their

cognitive abilities which have been shown to impact pupil response. It was shown that both the number of visits and the normalization procedure have an impact on reliability, the latter being the one with the strongest impact. The SNR did not have a major effect. There was, however, a tendency of giving up at the lowest SNR that was shown to be correlated with the PTA. The remaining chosen listener-specific factors were not found to be predictors of reliability, as they did not impact PPD or MPD. The results of this study are an important step towards the potential use of pupillometry as a diagnostic tool with hearing-impaired populations.

6. ACKNOWLEDGMENTS

This work was supported by the William Demant Foundation. Special thanks to the PUPILS project partners from DTU Hearing Systems and Eriksholm Research Centre.

7. REFERENCES

[1] M. Pichora-Fuller, S. Kramer, M. Eckert, B. Edwards, B. Hornsby, L. Humes, U. Lemke, T. Lunner, M. Matthen, C. Mackersie, G. Naylor, N. Phillips, M. Richter, M. Rudner, M. Sommers, K. Tremblay, and A. Wingfield, "Hearing impairment and cognitive energy: The framework for understanding effortful listening (fuel)," *Ear and hearing*, vol. 37 Suppl 1, pp. 5S–27S, 07 2016.

[2] M. B. Winn, D. Wendt, T. Koelewijn, and S. E. Kuchinsky, "Best practices and advice for using pupillometry to measure listening effort: An introduction for those who want to get started," *Trends in Hearing*, vol. 22, p. 2331216518800869, 2018. PMID: 30261825.

[3] A. Zekveld, S. Kramer, and J. Festen, "Pupil response as an indication of effortful listening: The influence of sentence intelligibility," *Ear and hearing*, vol. 31, pp. 480–90, 08 2010.

[4] S. E. M. Colby and B. McMurray, "Cognitive and physiological measures of listening effort during degraded speech perception: Relating dual-task and pupillometry paradigms," 04 2021.

[5] M.-B. Neagu, A. A. Kressner, H. Relaño-Iborra, P. Bækgaard, T. Dau, and D. Wendt, "Investigating the reliability of pupillometry as a measure of individualized listening effort," *Trends in Hearing*, vol. 27, p. 23312165231153288, 2023.

[6] J. Nielsen and T. Dau, "The danish hearing in noise test," *International journal of audiology*, vol. 50, pp. 202–8, 03 2011.

[7] D. Wendt, T. Koelewijn, P. Książek, S. E. Kramer, and T. Lunner, "Toward a more comprehensive understanding of the impact of masker type and signal-to-noise ratio on the pupillary response while performing a speech-in-noise test," *Hearing Research*, vol. 369, pp. 67–78, 2018. Aging Speech Communication 2017.

[8] B. Ohlenforst, D. Wendt, S. E. Kramer, G. Naylor, A. A. Zekveld, and T. Lunner, "Impact of snr, masker type and noise reduction processing on sentence recognition performance and listening effort as indicated by the pupil dilation response," *Hearing Research*, vol. 365, pp. 90–99, 2018.