

The effect of sentence predictability on speech recognition and listening effort using pupillometry

Dana Bsharat-Maalouf^{1,3}, Andreas Wölflé^{2,3}, Hanin Karawani^{1,3}, Esther Ruigendijk^{2,3}

¹Department of Communication Sciences and Disorders, University of Haifa, ²Department of Dutch Studies, Carl von Ossietzky University of Oldenburg, ³Cluster of Excellence “Hearing4all”.

Abstract

Understanding speech in everyday life often requires listeners to process degraded acoustic signals, particularly in background noise. Although speech perception may remain accurate under such conditions, listening can still be associated with increased effort [1]. One factor proposed to reduce listening effort is semantic context: highly predictable sentence contexts can constrain the range of possible lexical candidates and thereby facilitate perception [2]. However, previous findings on the effect of contextual predictability on listening effort remain mixed, particularly in studies using pupillometry [3, 4]. The present study therefore aimed to address this inconsistency by systematically examining whether sentence predictability influences both speech recognition and pupil-indexed processing effort. Specifically, we investigated not only overall recognition accuracy, but also the magnitude, temporal dynamics, and post-offset time course of the pupil response, providing a more comprehensive characterization of effort-related processing across behavioral and physiological measures.

Forty young adult German-speaking participants listened to German sentences (approximately 3 s in duration) varying in contextual predictability, with sentence-final target words embedded in either high-predictability ($n = 60$) or low-predictability ($n = 60$) contexts. Sentences were presented in both quiet and in background noise (speech-shaped noise at 0 dB SNR). Participants repeated each sentence aloud following a delayed response cue, allowing speech-recognition accuracy to be measured while pupil dilation was recorded continuously. Pupil responses were baseline-corrected prior to sentence onset and analyzed from sentence onset through the post-sentence interval preceding the response prompt.

Recognition accuracy was near ceiling in both quiet and noise for high- and low-predictability sentences. Despite these ceiling effects, pupil analyses revealed a robust effect of noise, with larger and delayed pupil responses in noise than in quiet. Sentence predictability also modulated the pupil response: low-predictability sentences elicited larger peak dilation than high-predictability sentences ($p = 0.03$), and peak dilation occurred later for low-predictability sentences ($p = 0.02$). These effects primarily reflected latency shifts in the pupil response profile rather than amplitude differences alone.

Time-course analyses further indicated that predictability-related differences emerged primarily in late stages of the pupil response. In quiet, these effects were brief and confined to a late time window (4348–4712 ms after speech onset). In noise, low-predictability sentences elicited greater pupil dilation than high-predictability sentences during a more sustained late window extending into the post-offset period (4227–5500 ms).

Overall, the findings suggest that semantic predictability modulates both the magnitude and timing of pupil-indexed processing during sentence recognition. High-predictability contexts were associated with reduced peak dilation and earlier peak timing, consistent with lower and more efficiently resolved processing demands. Predictability effects emerged primarily in late processing stages and were brief in quiet but more sustained in noise, suggesting that semantic predictability primarily shapes late-stage processing under more challenging listening conditions.

References

- [1] Bsharat-Maalouf, D., Schmidtke, J., Degani, T., & Karawani, H. (2025). Through the pupils' lens: Multilingual effort in first and second language listening. *Ear and Hearing*, 46(2), 494-511.
- [2] Strand, J., Simenstad, A., Cooperman, A., & Rowe, J. (2014). Grammatical context constrains lexical competition in spoken word recognition. *Memory & Cognition*, 42, 676–687. <https://doi.org/10.3758/s13421-013-0378-6>
- [3] Winn, M. B. (2016). Rapid release from listening effort resulting from semantic context, and effects of spectral degradation and cochlear implants. *Trends in Hearing*, 20, 2331216516669723.
- [4] Borghini, G., & Hazan, V. (2020). Effects of acoustic and semantic cues on listening effort during native and non-native speech perception. *The Journal of the Acoustical Society of America*, 147(6), 3783-3794.

Figures

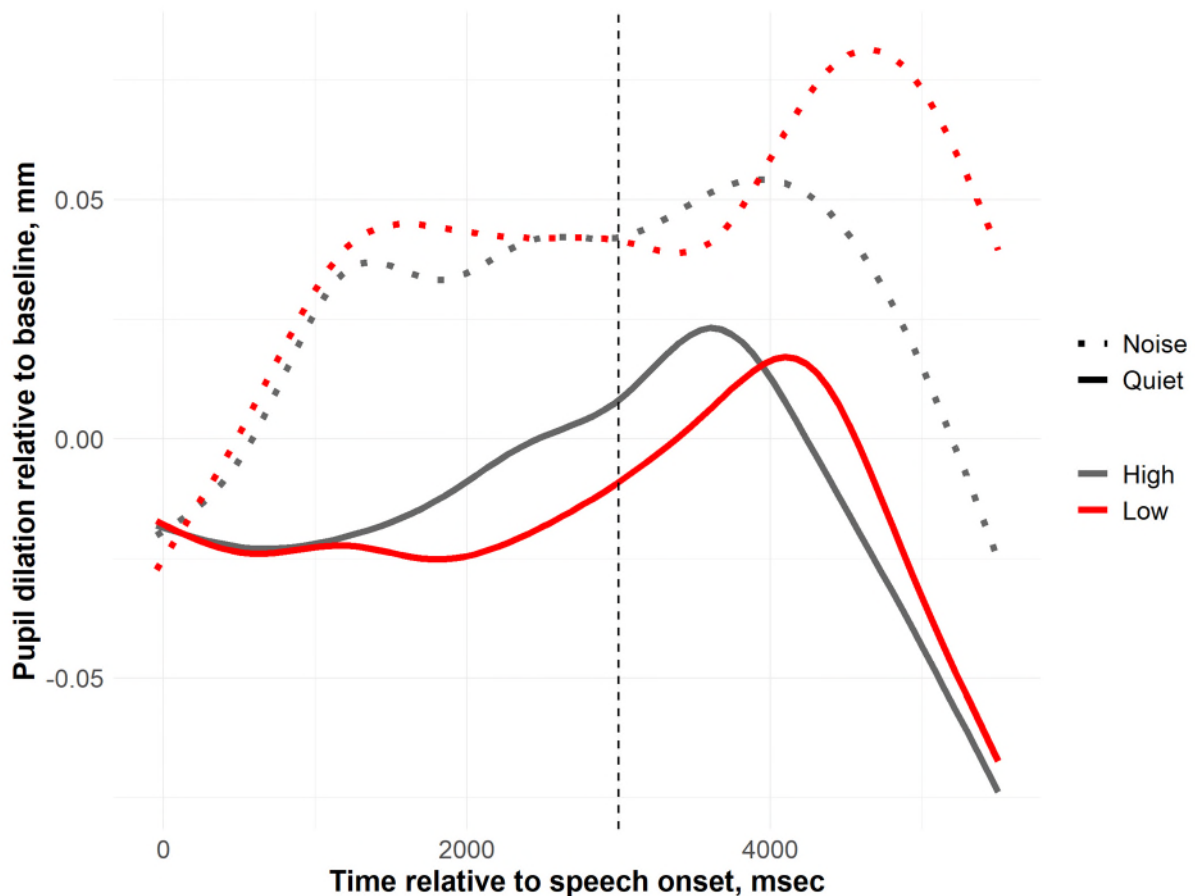


Figure 1. Time course of baseline-corrected pupil dilation during sentence processing in quiet and noise. Curves show responses to high-predictability contexts (grey) and low-predictability contexts (red), aligned to speech onset. The dashed vertical line marks sentence offset.