Lexical Predictability and Speech–Brain Synchrony in Compressed Speech

Irina Saranskaia (irinasaranskaia@gmail.com)[1], Olesia Dogonasheva[2], Denis Zakharov[1]
[1]Centre for Cognition and Decision Making, HSE University, Moscow, Russia, [2]Institut de l'Audition, Institut Pasteur, Paris, France

Speech comprehension depends on the coordination between acoustic input, neural temporal sampling, and top-down linguistic prediction [2]. Time-compressed speech provides a direct test of this coordination: as compression increases, syllabic information moves beyond the θ range (3-8 Hz), where cortical activity is thought to support syllabic parsing [3]. Contextual predictability may partly compensate for this temporal degradation. Here, we ask whether lexical predictability supports comprehension under strong compression and whether this benefit is reflected in speech-brain synchrony.

Native Russian listeners heard and repeated synthesised sentences while EEG was recorded. Sentences varied in lexical entropy and compression, yielding a 2×6 within-participant design: two entropy conditions and six compression levels (1×, 2×, 2.5×, 2.75×, 3×, 3.5×). Lexical entropy was defined as contextual uncertainty over upcoming lexical material. It was computed with the RuGPT-3 model by estimating next-token probability distributions at each word position, calculating Shannon entropy, and averaging values across the sentence. The stimulus set contained 440 sentences: 220 low-entropy and 220 high-entropy sentences. The two sets differed strongly in entropy (Welch's t-test, $p = 6.5 \times 10^{-87}$), but were matched on sentence length and syllabic density. Mean sentence duration was 2301 ms, and the uncompressed syllable rate was approximately 5.2 Hz. Responses were scored using word error rate (WER) and character error rate (CER), interpreted as indices of sentence recognition rather than process-pure measures of comprehension.

Preliminary mixed-effects analyses of data from 30 participants showed that repetition errors increased nonlinearly with compression and that this increase depended on lexical entropy (Fig. 1). Models including the entropy × compression interaction fit significantly better than main-effect models for both WER, $\chi^2(5) = 376.795$, $p < .001$, and CER, $\chi^2(5) = 362.018$, $p < .001$. Errors remained low at mild compression levels, but the entropy conditions diverged from approximately 2.75× onward. High-entropy sentences produced an earlier and steeper increase in errors than low-entropy sentences, suggesting that contextual predictability supports recognition when speech is strongly compressed.

The EEG analysis will test whether compression-related performance decline is accompanied by changes in envelope–EEG phase consistency. θ synchrony will test syllabic tracking, whereas δ synchrony will index slower temporal grouping or sensory chunking of articulated speech [1]. β activity and cross-frequency interactions will be analysed as candidate markers of top-down predictive support under temporal degradation [5]. We expect lexical entropy to modulate compression-related performance decline: low-entropy sentences should remain easier to recognise at strong compression levels. If this advantage is accompanied by changes in β activity or cross-frequency coupling, it would support a predictive-support account. In contrast, if low-frequency envelope–EEG synchrony remains detectable despite behavioural breakdown, this would be more consistent with the view that speech tracking can partly reflect stimulus-driven evoked responses rather than successful linguistic parsing [4].

References

[1] V. J. Boucher, A. C. Gilbert, and B. Jemel. The role of low-frequency neural oscillations in speech processing: Revisiting delta entrainment. *Journal of Cognitive Neuroscience*, 31(8):1205–1215, 2019. doi: 10.1162/jocn_a_01410. URL https://doi.org/10.1162/jocn_a_01410.

[2] O. Dogonasheva, K. B. Doelling, D. Zakharov, et al. Rhythm–based hierarchical predictive computations support acoustic–semantic transformation in speech processing. *Nature Computational Science*, 5(10):915–926, 2025. doi: 10.1038/s43588-025-00876-9. URL https://doi.org/10.1038/s43588-025-00876-9.

[3] A.-L. Giraud and D. Poeppel. Cortical oscillations and speech processing: Emerging computational principles and operations. *Nature Neuroscience*, 15(4):511–517, 2012. doi: 10.1038/nn.3063. URL https://doi.org/10.1038/nn.3063.

[4] E. Lalor and A. R. Nidiffer. On the generative mechanisms underlying the cortical tracking of natural speech: a position paper. 2025. doi: 10.31219/osf.io/xf8ay_v1. URL https://osf.io/preprints/osf/xf8ay_v1. Preprint.

[5] M. Pefkou, L. H. Arnal, L. Fontolan, and A.-L. Giraud. θ-band and β-band neural activity reflects independent syllable tracking and comprehension of time-compressed speech. *Journal of Neuroscience*, 37(33):7930–7938, 2017. doi: 10.1523/JNEUROSCI.2882-16.2017. URL https://doi.org/10.1523/JNEUROSCI.2882-16.2017.

Figures and Tables

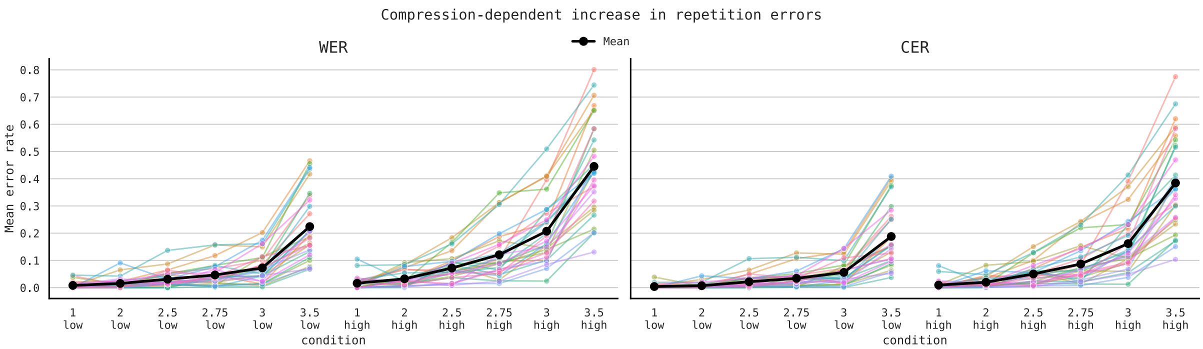

Figure 1: **Subject-wise behavioural performance.** Each coloured line shows one participant's condition-wise mean error rate; the black line shows the grand mean. Both WER and CER increased nonlinearly with compression, with a stronger increase for high-entropy sentences than for low-entropy sentences. The clearest difference between conditions emerged from approximately 2.75× onward.