Correspondence

Low-frequency neural parsing of hierarchical linguistic structures

Check for updates

When a person is listening to speech, their cortical dynamics can track multiword linguistic structures[1]. Kazanina and Tavano provide an in-depth discussion about the interpretation of this phenomenon in a recent Perspective article (Kazanina, N. & Tavano, A. What neural oscillations can and cannot do for syntactic structure building. *Nat. Rev. Neurosci.* **24**, 113–128; 2023)[2].

In the literature, two hypotheses have been raised about how cortical dynamics track linguistic structures, leading to two lines of experimentation.

In the multiscale envelope tracking (MET) hypothesis, natural speech is parsed into discrete levels of units — for example, syllables, words and phrases — and each level relates to a unique frequency band in the speech envelope and the envelope-tracking neural response[3] (such relationships, however, have recently been questioned[4]). In particular, delta-band neural activity encodes spoken phrases.

In the hierarchical structure building (HSB) hypothesis, any incrementally constructed linguistic structure is tracked by a neural population in the following sense: when a new word is added to the structure or the unit closes, there is a corresponding change in the activity of the neural population[1,5]. Studies driven by this hypothesis generally use frequency-tagging to separate the neural responses to different linguistic units[1].

Kazanina and Tavano base their discussion on the MET hypothesis and argue that the delta-band neural oscillation (1–4 Hz) cannot possibly contribute to the parsing of hierarchical phrasal structures, as it can only segment a sequence into non-overlapping chunks and cannot encode phrases longer than 1 s. The HSB hypothesis, however, provides straightforward solutions to these issues.

First, nested structures are well encoded under the HSB hypothesis. For example, if the smaller structure 'new plans' is nested in the bigger structure 'new plans give hope', each level of structure is separately encoded by neural activity corresponding to the time scale of the structure[1].

Second, whether structure-tracking neural activity is bounded in frequency is an empirical question that can be experimentally studied, for example, by presenting phrases of different durations. There is no obvious reason to restrict the response frequency to be above 1 Hz (ref. 2), and many studies have already observed structure-tracking activity below 1 Hz (ref. 1).

Kazanina and Tavano propose that structure-tracking neural activity indicates sequence integration instead of sequence chunking, to allow the time scale of neural activity to be dissociated from the time scale of linguistic units[2]. For example, phrases presented at 0.5 Hz can be represented by neural activity at an arbitrary frequency of, for example, 2.1 Hz. The integration and chunking hypotheses, however, have equal power in explaining the neural responses that match the time scale of linguistic input, such as the responses observed in refs. 1,6–9.

In summary, Kazanina and Tavano's Perspective article starts from a neurophysiological concept — the delta oscillation — and analyses how the properties of delta oscillation constrain speech processing mechanisms. This approach is challenging since the properties of delta oscillations, such as the frequency range, that underlie neural networks and biophysical mechanisms remain elusive. A more feasible approach is to use the computational demand of linguistic structure building to constrain possible neural mechanisms[10], and to optimize the experimental design and analysis methods to dissociate linguistic structure building from related processes such as attention[6], word encoding[7,8] and prosodic processing[9].

There is a reply to this letter by Kazanina, N. & Tavano, A. *Nat. Rev. Neurosci.* https://doi.org/10.1038/s41583-023-00750-5 (2023).

Nai Ding ✉
Key Laboratory for Biomedical Engineering of Ministry of Education, College of Biomedical Engineering and Instrument Sciences, Zhejiang University, Hangzhou, China.
✉e-mail: ding_nai@zju.edu.cn

References

1. Ding, N., Melloni, L., Zhang, H., Tian, X. & Poeppel, D. Cortical tracking of hierarchical linguistic structures in connected speech. *Nat. Neurosci.* **19**, 158–164 (2016).
2. Kazanina, N. & Tavano, A. What neural oscillations can and cannot do for syntactic structure building. *Nat. Rev. Neurosci.* **24**, 113–128 (2023).
3. Keitel, A., Gross, J. & Kayser, C. Perceptually relevant speech tracking in auditory and motor cortex reflects distinct linguistic features. *PLoS Biol.* **16**, e2004473 (2018).
4. Zhang, Y., Zou, J. & Ding, N. Complex mapping between neural response frequency and linguistic units in natural speech. *J. Cogn. Neurosci.* **35**, 1361–1368 (2023).
5. Ding, N. A structure-based memory maintenance model for neural tracking of linguistic structures. Preprint at https://doi.org/10.48550/arXiv.2002.11870 (2020).
6. Jin, P., Zou, J., Zhou, T. & Ding, N. Eye activity tracks task-relevant structures during speech and auditory sequence perception. *Nat. Commun.* **9**, 5374 (2018).
7. Lo, C. W., Tung, T. Y., Ke, A. H. & Brennan, J. R. Hierarchy, not lexical regularity, modulates low-frequency neural synchrony during language comprehension. *Neurobiol. Lang.* **3**, 538–555 (2022).
8. Kalenkovich, E., Shestakova, A. & Kazanina, N. Frequency tagging of syntactic structure or lexical properties; a registered MEG study. *Cortex* **146**, 24–38 (2022).
9. Glushko, A., Poeppel, D. & Steinhauer, K. Overt and implicit prosody contribute to neurophysiological responses previously attributed to grammatical processing. *Sci. Rep.* **12**, 14759 (2022).
10. Poeppel, D. The maps problem and the mapping problem: two challenges for a cognitive neuroscience of speech and language. *Cogn. Neuropsychol.* **29**, 34–55 (2012).

Competing interests

The author declares no competing interests.