Working memory asymmetrically modulates auditory and linguistic processing of speech

Yiguang Liu[1,*], Cheng Luo[1], Jing Zheng[2], Junying Liang[3], Nai Ding[1,2]
[1]Zhejiang Lab, Hangzhou, China [2]Zhejiang University, Hangzhou, China
Email: lyg_1606@zju.edu.cn (YL)

Keywords: speech comprehension, working memory, neural tracking, auditory encoding, linguistic processing

Background

Speech comprehension entails multiple processing stages from early sensory encoding of acoustic features to higher-level linguistic processing, e.g., semantic integration and syntactic parsing. Evidence has suggested that this complex process suffers from working memory (WM) load imposed to listeners (e.g., Mattys & Wiget 2011), but it remains elusive which stage is the main locus of the WM effects on speech processing. One methodological challenge to addressing this issue lies in the difficulty in dissociating the mental representations of different processing stages of speech.

Methods

Here, we draw on a recently developed multiscale frequency-tagging paradigm, which allows to separate the neural responses to different linguistic units in the frequency domain (Ding et al. 2016). By doing so, we concurrently measure how the WM load modulates neural activity tracking three levels of linguistic units, i.e., syllables, phrases, and sentences. The neural tracking of syllable is closely linked to the encoding of speech envelope as an important acoustic feature (Ding et al. 2017), while neural tracking of phrases and sentences reflects rule-based linguistic processing (Jin, Lu & Ding 2020). Participants (N = 60) engage in a sentence comprehension task (*N of sentences* = 152) while recording their electroencephalogram (EEG) data, and the WM load is manipulated by asking them to memorize either auditory verbal sequences (Exp 1 and Exp 3) or visual patterns (Exp 2; see Fig. 1 for the speech stimuli and task design).

Results

It is found that verbal and visual WM load modulate speech processing in similar manners: Higher working memory load attenuates neural activity tracking of phrases and sentences, but enhances neural activity tracking of syllables (see Fig. 2 for the results). Since verbal and visual WM load similarly influence the neural responses to speech, such influences may derive from the domain-general component of WM system.

Discussion

Our data reveal that WM load asymmetrically modulates lower-level auditory encoding and higher-level linguistic processing of speech. The reversed direction of the observed WM effects possibly reveals a load-dependent reallocation of processing resources, which might function as a compensatory mechanism for degraded analyses of higher-level linguistic information. This compensation interpretation in line with the cue integration hypothesis, which posits that language comprehenders are capable of weighting signal-based and knowledge-based processing flexibly and adaptively against a given situation (Martin 2016).

References

Ding, Nai, Lucia Melloni, Hang Zhang, Xing Tian & David Poeppel. 2016. Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience* 19(1). 158–164.

Ding, Nai, Aniruddh D. Patel, Lin Chen, Henry Butler, Cheng Luo & David Poeppel. 2017. Temporal modulations in speech and music. *Neuroscience & Biobehavioral Reviews* 81. 181–187.

Jin, Peiqing, Yuhan Lu & Nai Ding. 2020. Low-frequency neural activity reflects rule-based chunking during speech listening. (Ed.) Tobias Reichenbach, Barbara G Shinn-Cunningham & Tobias Reichenbach. *eLife*. eLife Sciences Publications, Ltd 9. e55613.

Martin, Andrea E. 2016. Language Processing as Cue Integration: Grounding the Psychology of Language in Perception and Neurophysiology. *Frontiers in Psychology* 7. 120.

Mattys, Sven L. & Lukas Wiget. 2011. Effects of cognitive load on speech recognition. *Journal of Memory and Language* 65(2). 145–160.

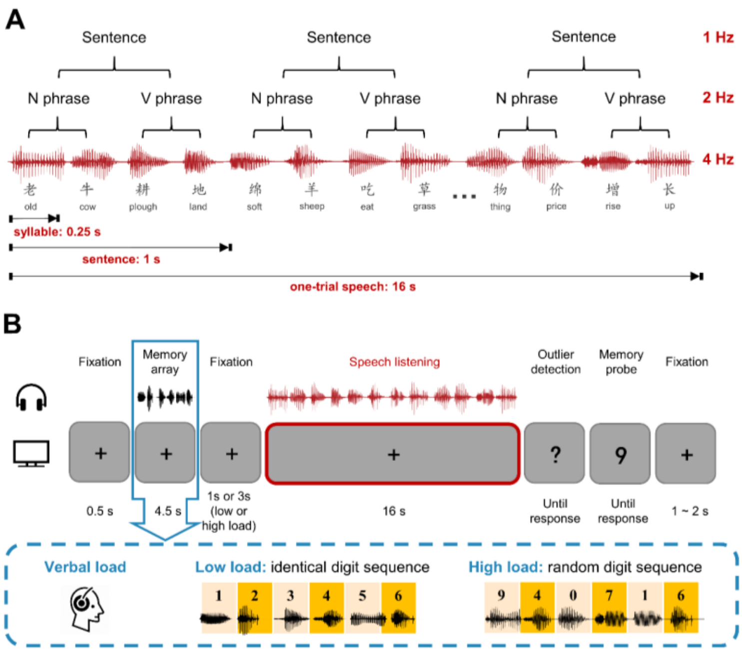

***Fig. 1: Stimulus and experiment procedure. A**, An isochronous sequence of syllables was presented at a constant rate of 4 Hz in the speech listening task. Two syllables grouped into a phrase and four syllables grouped into a sentence, and thus phrases and sentences were presented at 2 Hz and 1 Hz respectively. The stimuli were in Chinese and their English counterparts are shown for illustration. **B**, In Experiment 1, the working memory task was to memorize a string of six digits before speech listening and report afterwards whether a memory probe (a digit) was present in the memorized sequence. In a high working memory load condition, the six digits were randomly chosen while in a low working memory load condition, the digit sequence was fixed, i.e., "1, 2, 3, 4, 5, 6". EEG responses were only analyzed during the speech listening session (shown in red). The procedure of Exp 2 and Exp 3 are similar with that of Exp 1, except that visual patterns (Exp 2) and more complex verbal sequences (Exp 3) were memorized by the participants.*

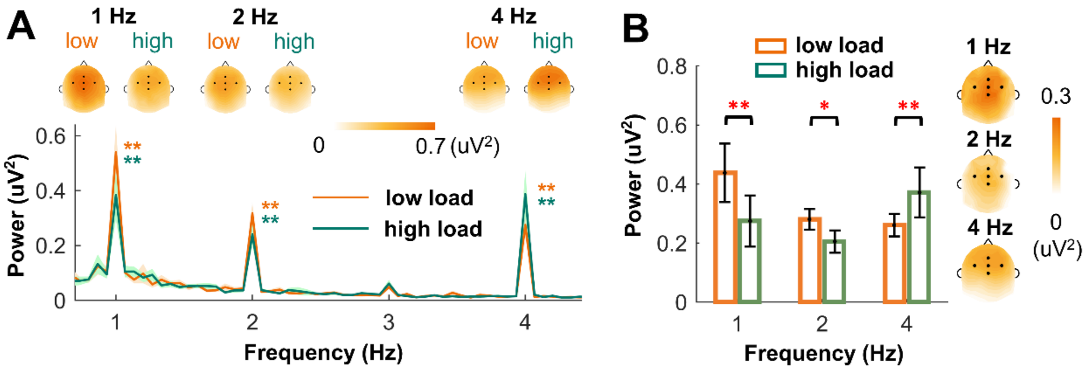

***Fig. 2: EEG response spectrum and topography in Exp 1**. **A**, The EEG response spectrum averaged over participants and channels. Three response peaks were observed at the sentential, phrasal, and syllable rates (1, 2, and 4 Hz, respectively). The shaded area indicates one SEM across participants. Stars in orange (low load) and green (high load) indicate significantly higher power at the target frequency than at neighboring frequency bins. The topographic plots are illustrated above the spectrum, showing a central-frontal distribution. **B**, Normalized EEG response at the target frequencies. The normalized power was generated by subtracting the mean power of four neighboring frequency bins. The error bars indicate SEM over participants. Significant cross-condition differences were observed at all target frequencies (i.e., 1, 2, and 4 Hz), as indicated by red stars. The topographic plots illustrate the difference between the low-load and high-load condition (1 and 2 Hz: low minus high; 4 Hz: high minus low), generally showing a central-frontal distribution. Five black dots in the topographies refer to the position of FCz (middle), Fz (upper), Cz (lower), FC3 (left), and FC4 (right). Similar patterns are shown in Exp 2 and Exp 3. *p < 0.05, **p < 0.01.*