Lexical representation doesn't fully capture low-frequency neural oscillations during sentence comprehension

Chia-Wen Lo & Jonathan R. Brennan (University of Michigan)
chiawenl@umich.edu

Introduction Neural responses appear to be entrained to hierarchical structures (Ding et al., 2016, 2017). Ding and colleagues (2016) observed cortical tracking of linguistic structures at evoked frequencies corresponding to phrasal (2 Hz) and sentence structures (1 Hz) levels of Mandarin structures when native Mandarin speakers listened to isochronic four-syllable sentences with each syllable in 250 ms. However, it is still unclear whether these entrainments at delta frequencies (0.5 - 3 Hz) can be reliably interpreted as reflexes of of hierarchical information (Martin & Doumas, 2017; Frank & Yang 2018). Specifically, the lexical representation model of Frank & Yang (2018) encodes word similarities and part of speech for each word, and represents stimuli as sequences of high-dimensional numeric vectors. The model simulates similar power spectra, suggesting that findings in Ding et al. (2016) may follow from the cortical tracking of lexical or part of speech information, not necessarily hierarchy. We conducted two EEG experiments using Ding's experimental paradigm and adopted Frank and Yang's (FY) model to examine whether the interaction between the temporal properties of stimuli and delta oscillations can be explained by the lexical representation approach.

Experiment 1 N=31 native speakers of Mandarin Chinese listened to trials consisting of ten 4-syllable sentences, following Ding's experimental paradigm. Four experimental conditions were included (Table 1). Crucially, Condition (4) maintains word-sequence patterns at 2 Hz and 1 Hz but the words do not form grammatical phrases. If neural tracking reflects sequence information, then oscillations at those bands should be equivalent between (4) and the regular sentences in (1). For each condition, we computed Evoked Power (EP) and Inter-trial phase coherence (ITPC) from 0.5 to 10 Hz in increments of 0.111 Hz (Fig. 1). No sentence or phrase-level peaks were found in condition (4). This suggests that oscillatory synchronization is modulated by linguistic hierarchical structures, not just word-sequences.

Simulations Following Frank & Yang (2018), word vectors were extracted for conditions (1) and (4). 12 simulated subjects and 50 randomized items modified from Ding et al. 2016 were tested in the model. First, we adopted FY's word vectors to create the word embeddings for the reversed phrases (condition (4)) by swapping columns, without changing other parameters. Inconsistent with our experimental results, model results show delta entrainment for reversed phrases (Fig. 2). Second, different word embeddings from a Chinese pre-trained word2vec (Park, 2016) and a Chinese pre-trained BERT (Cui et al., 2019) were encoded separately. Results also show delta entrainment at 2 Hz and 1 Hz for reversed phrases, inconsistent with our experimental results.

Experiment 2 Exp. 1 and previous studies tested syllable rates fixed at 4/sec (250 ms), but syllable, word, and sentence rates in natural speech are more variable. The next experiment used time-compressed speech to test temporal boundaries of hierarchical entrainment. N=27 native speaker of Mandarin speakers participated, and the same experimental paradigm was adopted. Five conditions that vary the length of syllables (250, 200, 100 ms) and the regularity between syllables were included (Table 2). Semi-regular items included two different syllable lengths in a fixed order (200-300-200-300 ms). The irregular condition included random syllable length (either 200 or 300 ms). We computed EP and ITPC from 0.1 to 20 Hz in increments of 0.1 Hz and a Hanning taper was applied after adding 10 seconds of zero-padding to each condition. Results show that peaks at the sentence, phrase, and syllable rate were observed in all conditions except 100ms/syl, which might show a sentence-level peak only. (Fig. 2).

Table 1: Stimuli design in Experiment 1.

Condition (1): Four-syllable sentence (ABCD) mian yang chi cau Cotton sheep eat grass ‘Sheep eat grass.’	**Condition (2): Semantically-mismatched sequences** jun hai ben cau Soldier child run grass
Condition (3): Two-syllable phrase (ABAB) lau niu qing cau Old cattle green grass	**Condition (4): Reversed phrase (BADC)** yang mian cau chi Sheep cotton grass eat

Table 2: Experimental conditions and frequency of interest in Experiment 2

Condition	Regular: 250 ms/syl (4Hz/syl)	Regular: 200 ms/syl (5Hz/syl)	Regular: 100 ms/syl (10Hz/syl)	Semi-regular: 300 200 300 200 ms	Irregular: 300 200 300 200, 300 300 200 200...
Frequency of interest	Sentence: 1 Hz Phrase: 2 Hz Syllable: 4 Hz	Sentence: 1.25 Hz Phrase: 2.5 Hz Syllable: 5 Hz	Sentence: 2.5 Hz Phrase: 5 Hz Syllable: 10 Hz	Sentence: 1 Hz Phrase: 2 Hz Syllable: 4 Hz	Sentence: 1 Hz Phrase: 2 Hz Syllable: 4 Hz

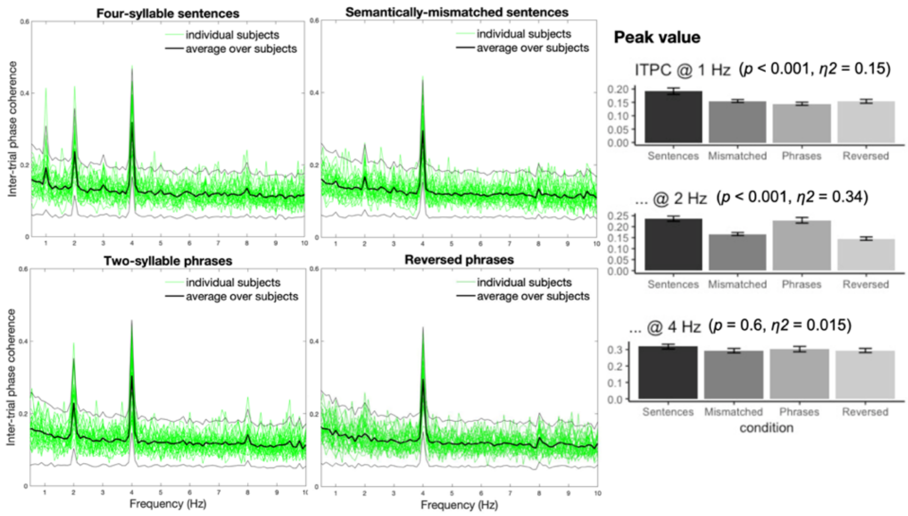

Figure 1. Inter-trial phase coherence (Experiment 1)

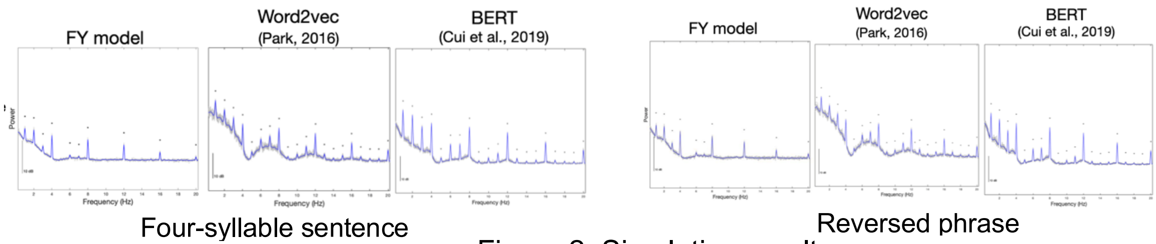

Figure 2. Simulation results

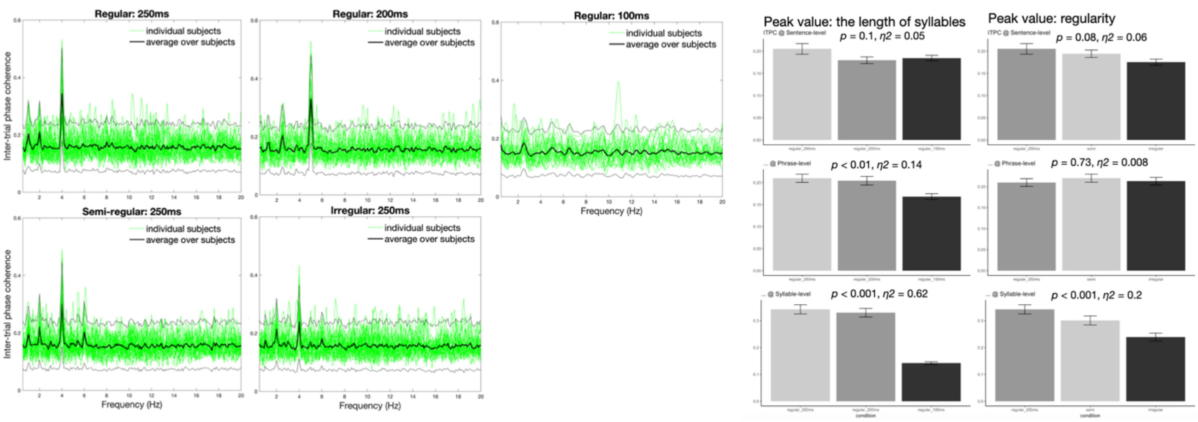

Figure 3. Inter-trial phase coherence (Experiment 2)