**Visual cues for predictive entrainment in sign language**

Evie A. Malaia\(^1\), Julia Krebs\(^2\)\(^3\), Joshua D. Borneman\(^4\), Ronnie Wilbur\(^4\), Dietmar Roehm\(^2\)\(^3\)

\(^1\)Department of Communicative Disorders, University of Alabama, Tuscaloosa, USA
\(^2\)Neurobiology of Language, Department of Linguistics, University of Salzburg, Austria
\(^3\)Centre for Cognitive Neuroscience (CCNS), University of Salzburg, Austria
\(^4\)Speech, Language, and Hearing Department, Purdue University, USA

When people listen to speech, neural activity is cued by the fluctuation in the acoustic envelope (Peelle, Gross, & Davis, 2012). For spoken languages, this cue-based entrainment is the basis of signal parsing and predictive processing (Ding et al., 2016). A growing body of research also indicates that humans are highly sensitive to motion differences in the visual signal (Strickland et al., 2015), and signers make neurolinguistic distinctions based on motion profiles of signs (Malaia et al., 2013). The temporal mechanism of predictive parsing in the visual modality, however, is not yet clear. We tested the hypothesis that in signers, as in speakers, the syllable-driven fluctuations (~4 syllables per second, or 4Hz) in the envelope of the signal are cues to predictive entrainment.

EEG data was collected in signers viewing signed sentences (meaningful stimuli), and the same sentences played in reverse (meaningless stimuli with rich spectrotemporal structure). We then assessed the optical flow in the visual stimuli, which is a validated measure of dynamic entropy in the overall signal (Borneman et al., 2018), as well as a measure of overall motion in time that marks sign-syllable boundaries in continuous signed discourse. We then computed cross-correlation between the variations in the optical flow of the visual signal, and neural activity of the participants.

15 native signers of Austrian Sign Language (ÖGS) (9 F, age M = 39.37, SD = 10.19), who were born Deaf or lost their hearing early in life, participated in the study. They viewed 40 sentences in Austrian Sign Language (ÖGS), as well as the same sentences, but reversed in the time domain, and rated comprehensibility of the stimuli on a Likert scale. Each video was 5 to 7 seconds in duration. 80 critical sentences were presented to the participants in pseudo-random orders, with 200 fillers to distract from strategic processing (total 280 sentences). Optical flow (OF) for each video was determined using MATLAB vision toolbox optical flow function (Malaia et al., 2016); the amplitudes across all velocity bins were added to calculate the magnitude of optical flow for each frame, thereby generating an optical flow timeseries. Coherence was calculated between the optical flow timeseries for each stimulus video, and the neural response in each electrode timeseries for each participant.

Behavioral data indicated that only sentences in the direct video condition were rated as linguistically acceptable (signing videos acceptability M=6.24, SD=.8; reversed videos M=1.7, SD=.89). Peak coherence between the stimuli and neural activity occurred between 100 msec and 250 msec post-stimulus onset in response to both meaningful and meaningless video stimuli conditions. Negative values of cross-correlation were observed in response to reversed videos, which were rated as not linguistically acceptable. Frequency-domain analysis revealed cued entrainment at 1.2 Hz, 2.4 Hz, and 4.2 Hz, in response to language stimuli (Figure 1). In response to non-comprehended reversed video stimuli, cued entrainment peaks were distributed at 1 Hz, 1.6Hz, 2.6Hz, 3.2Hz, and 4.2 Hz, indicating broad search for cues at lower frequencies, and robust syllable-frequency response (~4Hz) regardless of stimulus type.
The combined behavioral and EEG data reveal that linguistic acceptability and comprehension of signed sentences in rooted in cued entrainment of neural oscillations to the sign-syllable boundaries in the visual signal, as measured by optical flow metric. The findings demonstrate that the cortical tracking of spectro-temporal entropy of the signal is a modality-independent mechanism for predictive processing in humans. These results point to the likelihood of modality-independent evolution of language based on cortical entrainment that facilitates scene segmentation.

![Cross-correlation values between video stimuli and neural activity in response to signed video sentences (solid line) and reversed videos (dashed line), illustrating cued entrainment at 1.2 Hz, 2.4 Hz, and 4.2 Hz in response to language stimuli.](image)

**Figure 1.** Cross-correlation values between video stimuli and neural activity in response to signed video sentences (solid line) and reversed videos (dashed line), illustrating cued entrainment at 1.2 Hz, 2.4 Hz, and 4.2 Hz in response to language stimuli.

**References**


