

Bibliographic Information:

Brookshire, G., Lu, J., Nusbaum, H. C., Goldin-Meadow, S., & Casasanto, D. (2017). Visual cortex entrains to sign language. *Proceedings of the National Academy of Sciences*, 114(24), 6352-6357.

Visual cortex entrains to sign language

Geoffrey Brookshire^{a,1}, Jenny Lu^a, Howard Nusbaum^a, Susan Goldin-Meadow^a, and Daniel Casasanto^{a,1}

^aDepartment of Psychology, University of Chicago, Chicago, IL 60637

This manuscript was compiled on April 11, 2017

Despite immense variability across languages, people can learn to understand any human language, spoken or signed. What neural mechanisms allow people to comprehend language across sensory modalities? When people listen to speech, electrophysiological oscillations in auditory cortex entrain to slow (<8 Hz) fluctuations in the acoustic envelope. Entrainment to the speech envelope may reflect mechanisms specialized for auditory perception. Alternatively, flexible entrainment may be a general-purpose cortical mechanism that optimizes sensitivity to rhythmic information regardless of modality. Here we test these proposals by examining cortical coherence to visual information in sign language. First, we develop a metric to quantify visual change over time. We find quasi-periodic fluctuations in sign language, characterized by lower frequencies than fluctuations in speech. Next, we test for entrainment of neural oscillations to visual change in sign language, using electroencephalography (EEG) in fluent speakers of American Sign Language (ASL) as they watch videos in ASL. We find significant cortical entrainment to visual oscillations in sign language below 5 Hz, peaking at about 1 Hz. Coherence to sign is strongest over occipital and parietal cortex, in contrast to speech, where coherence is strongest over the auditory cortex. Non-signers also show coherence to sign language, but entrainment at frontal sites is reduced relative to fluent signers. These results demonstrate that flexible cortical entrainment to language does not depend on neural processes that are specific to auditory speech perception. Low-frequency oscillatory entrainment may reflect a general cortical mechanism that maximizes sensitivity to informational peaks in time-varying signals.

sign language | EEG | oscillations | cortical entrainment

Languages differ dramatically from one another, yet people can learn to understand any natural language. What neural mechanisms allow humans to understand the vast diversity of languages, and to distinguish linguistic signal from noise? One mechanism that has been implicated in language comprehension is neural entrainment to the volume envelope of speech. The volume envelope of speech fluctuates at low frequencies (< 8 Hz), decreasing at boundaries between syllables, words, and phrases. When people listen to speech, neural oscillations in the delta (1–4 Hz) and theta bands (4–8 Hz) become entrained to these fluctuations in volume [1–4].

Entrainment to the volume envelope may represent an active neural mechanism to boost perceptual sensitivity to rhythmic stimuli [2, 5–7]. Although entrainment is partly driven by bottom-up features of the stimulus [8–10], it also depends on top-down signals to auditory cortex from other brain areas. Auditory entrainment is strengthened when people see congruent visual and auditory information [11, 12], and is modulated by attention [13] and by top-down signals from frontal cortex [4, 14].

Cortical entrainment is proposed to perform a key role in speech comprehension, such as segmenting out syllables from a continuous speech stream [1, 2, 15] or optimizing perceptual sensitivity to rhythmic pulses of sound [5–7]. However, the

mechanisms driving entrainment to speech remain unclear. We consider two hypotheses. First, flexible entrainment to quasi-periodic rhythms may be specific to auditory perception [6]; in visual perception, by contrast, cortical oscillations in the alpha band (8–12 Hz) may phase-lock only to consistent stimulus rhythms, without adjusting to variable stimulus rhythms [16]. Second, low-frequency cortical entrainment may be a general-purpose neural mechanism that helps optimize perception to time-varying stimuli regardless of the perceptual modality. Neural oscillations may allow the brain to rhythmically orient attention to quasi-periodic stimuli [5, 17] across sensory systems.

Because previous studies of cortical entrainment to rhythms in language have focused on oral speech, they have been unable to distinguish between these competing hypotheses. Here we test for low-frequency entrainment to a purely visual language: American Sign Language (ASL). Prior studies show that neural and behavioral oscillations in vision are preferentially entrained by stimuli that flicker in the alpha band [18–21]. Therefore, if flexible cortical entrainment to oral speech depends on modality-specific properties of auditory processing, then phase-locking to sign language should be concentrated in the alpha band, if it occurs at all [16]. Alternatively, if cortical entrainment is a generalized neural strategy to maximize sensitivity to rhythmic stimuli, then oscillatory activity in visual cortex should entrain at the frequency of informational changes in ASL.

To determine whether human cerebral cortex entrains to rhythmic information in sign language, first we developed a metric for quantifying the amplitude of visual change in sign, by analogy to the acoustic envelope of speech. Next, we characterized visual variability across four sign languages,

Significance Statement

Language comprehension is thought to rely on a combination of specialized and general-purpose neural mechanisms. When people listen to speech, low-frequency oscillations in cerebral cortex (<8 Hz) become entrained to quasi-rhythmic fluctuations in volume. Entrainment of auditory cortex to speech rhythms has been well documented, but its functional significance has remained unclear. By showing similar entrainment of visual cortex to sign language, we establish that this phase-locking is not due to specific properties of auditory cortex or of oral speech perception. Rather, low-frequency entrainment is a generalized cortical strategy for boosting perceptual sensitivity to informational peaks in time-varying signals.

G.B. and D.C. designed the research; G.B. and J.L. performed research; G.B. analyzed data; G.B. and D.C. wrote the paper with critical input from J.L., H.N., and S.G.M.

The authors declare no conflict of interest.

¹To whom correspondence should be addressed. E-mail: g_b@cal.berkeley.edu, casasantod@alum.mit.edu

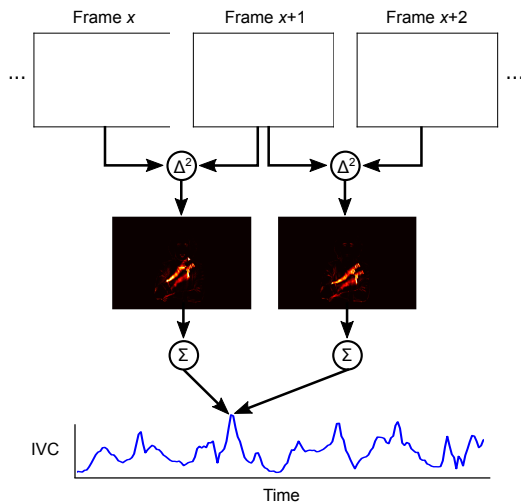


Fig. 1. Calculation of the Instantaneous Visual Change (IVC). The IVC summarizes total visual change at each point in time. First, the difference between adjacent grayscale video frames (top row) is calculated for each pixel. To aggregate over both increases and decreases in brightness, these pixel-wise differences are then squared (middle row). Finally, the brightness values in all pixels of the squared-difference images are summed to obtain a single value summarizing the magnitude of change between two video frames. Computation of this value for each adjacent pair of frames yields a time-series (bottom).

showing that this variability is quasi-periodic below 8 Hz. Finally, we demonstrated that cerebral cortex entrains to visual variability in sign language, and showed that entrainment is strongest around the frequencies of phrases and individual signs in ASL.

Results

Developing a metric for quantifying visual change. In order to examine neural entrainment to visual rhythms in sign language, we must first quantify the ‘amplitude envelope’ of a visual signal. The acoustic envelope is a highly reduced representation of sound, tracing extreme amplitude values in the time-varying signal. Oscillations in the envelope of speech depend on movements of various components of the vocal tract, including the rhythmic opening and closing of the mandible [22]. Sign language, in contrast, does not involve consistent oscillatory movements by any single effector [23]. However, quasi-periodic oscillations in sign language may arise from the coordinated movements of multiple effectors.

Here we present the Instantaneous Visual Change (IVC) as a metric that is conceptually similar to the acoustic envelope, summarizing the amplitude of change at each time point. The IVC is a time-series of aggregated visual changes between frames (Fig. 1, *Method and Materials*). This algorithm provides an automatic, objective alternative to human-coded methods of studying temporal structure in sign [24].

The amplitude of the IVC indexes the amount of visual change between two video frames. The largest peaks in the IVC therefore occur during large, quick movements. In the videos we analyzed, these changes corresponded primarily to movements of the signers’ hands and arms, but may also reflect movements of the face, head, and torso. For example, a quick arm movement results in a larger number of pixels changing in each adjacent frame – and a higher peak in the IVC – than

a slow arm movement. The IVC thus offers a heuristic index of new linguistic information in the visual signal. An example video illustrating the IVC is included in the online Supporting Information (Video S1).

Characterizing temporal structure in sign language. The IVC allows us to characterize one dimension of the temporal structure in sign language, and to directly compare the spectral signatures of amplitude variability across sign and oral speech. Visual examination of the raw IVC of sign language reveals quasi-periodic oscillations with irregularly-timed peaks (Fig. 2A). To characterize variability within and across sign languages, we computed the power spectra of the IVC from samples of four different sign languages: American Sign Language (ASL), German Sign Language (Deutsche Gebärdensprache, DGS), British Sign Language (BSL), and Australian Sign Language (Auslan). These languages developed independently of the oral languages spoken in these countries, and come from three genetically unrelated language families (BSL and Auslan from the BANZL family, ASL from the French Sign Language family, DGS from the German Sign Language family). In all four languages, power in the IVC decreases monotonically with increasing frequency, without any pronounced peaks in the spectrum (Fig. 2C). We tested for rhythmic components in the IVC by comparing these spectra against the $1/f$ spectrum characteristic of many signals in the natural world. Power that is higher than the $1/f$ function indicates periodicity at that frequency [22]. The IVC of sign language showed elevated power at approximately 2–8 Hz ($P_s < .01$). Individual signs in sign languages tend to occur at approximately 2–2.5 Hz [25, 26], on the lower end of the rhythmic components in the IVC. These analyses suggest that sign language involves weak, quasi-periodic rhythms with variable frequencies in the delta and theta range.

To explicitly compare temporal structure between sign and speech, we contrasted the IVC of sign with the broadband envelope of speech [22]. We computed the broadband envelope of samples from nine spoken languages representing five language families: English, French, Portuguese, Dutch, German, Hungarian, Japanese, Arabic, and Mandarin. After resampling the broadband envelopes and IVC signals to a common frequency (30 Hz) and standardizing the amplitude of each recording by dividing out its standard deviation, we compared the average spectra of the IVC of sign and the broadband envelope of speech (Fig. 2D). Spoken languages showed stronger modulations than sign languages above 2 Hz. This increased power may reflect modulation due to syllables in speech, which occur at approximately 2–10 Hz [22, 27]. Indeed, peaks from individual syllables are visible in the broadband envelope, and these peaks occur at about 4 Hz (Fig. 2B). These results indicate that visual motion in sign language is modulated at lower frequencies than auditory volume in spoken language. This difference is consistent with the slower movements in the articulators for sign (the hands) than in the articulators for speech (the vocal tract) [25, 26].

Cortical coherence to visual rhythms in sign language. We used electroencephalography (EEG) to examine cortical entrainment to quasi-rhythmic fluctuations in visual information in sign language. Fluent speakers of ASL watched videos of ASL stories against a static background. We tested for coherence of low-frequency electrophysiological oscillations to

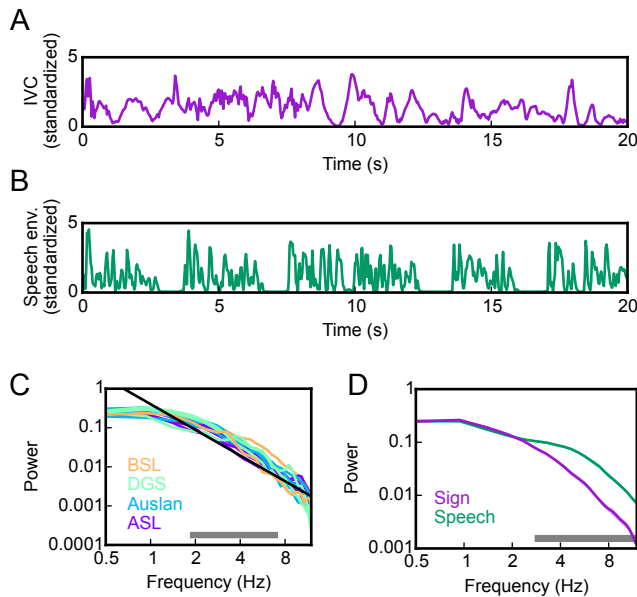


Fig. 2. Temporal structure in signed and spoken language. (A) Example trace of the IVC of ASL. (B) Example trace of the broadband envelope of English. (C) Spectrum of sign language plotted on log-log axes. Line color denotes language (BSL: British Sign Language; DGS: German Sign Language; Auslan: Australian Sign Language; ASL: American Sign Language). Each curve shows the spectrum of a separate video sample ($N = 14$, total duration 1:10:22). The black line shows the best-fit $1/f$ trend across all samples. The gray bar at the base of the plot shows where the IVC spectra are significantly greater than the $1/f$ fit ($P < .01$ by 1-sample t-tests). (D) Comparison of the mean spectra across signed and spoken languages. Shaded area depicts the standard error of the mean. The gray bar indicates significant differences between the two curves ($P < .01$ by independent-samples t-tests). Sign language samples are the same as in panel (C). Audio recordings were sampled from speech in 9 languages ($N = 12$, total duration 1:07:28). Amplitude in all analyses has been standardized by each recording's standard deviation.

quasi-periodic oscillations in the IVC.

Coherence was calculated separately at each EEG channel in partially overlapping, logarithmically spaced bins centered over 0.5–16 Hz (*Materials and Methods*). Because coherence provides no intrinsic measure of chance performance, we created a null distribution of coherence using a randomization procedure. To obtain each value in the null distribution, we time-shifted the IVC to a randomly-selected starting point, moving the portion of the IVC that remained after the final time point of the EEG signal to the beginning of the recording. This procedure preserves the spectral and temporal characteristics of the EEG and IVC recordings, but eliminates any relationship between these signals. Coherence was then computed between the EEG recordings and the randomly-shifted IVC.

A cluster-based permutation test indicated that coherence between cortical oscillations and the IVC of sign was stronger than would be expected by chance ($P = .0001$). Averaging the coherence spectrum across every EEG channel, coherence was above chance from 0.4–5 Hz, peaking at 1 Hz (Fig. 3A). Coherence emerged over a similar range of frequencies when we selected only occipital channels (0.8–5 Hz; $P = .0001$), primarily reflecting entrainment in visual cortex (Fig. 3B). In frontal channels, above-chance coherence was present from 0.4–1.25 Hz ($P = .0001$; Fig. 3C), revealing top-down control from frontal cortex. Examining the entire scalp distribution,

cortical coherence to the IVC of sign language was strongest over central and occipital channels (Fig. 3D).

To test whether cortical entrainment depends on linguistic knowledge, we examined coherence to sign language in people who did not know any ASL. Like signers, non-signers showed significant coherence to videos of ASL storytelling ($P < .0005$), with the strongest coherence over central and occipital channels from 0.8–3.5 Hz (Fig. S1). We then separately analyzed effects of linguistic knowledge on entrainment in occipital and frontal cortex. Although coherence at occipital channels did not significantly differ between groups (Fig. 4B), coherence at frontal channels was stronger in signers than in non-signers, indicating differences in top-down control based on familiarity with ASL ($P < .05$; Fig. 4A; Fig. S1).

Discussion

Cortical coherence to sign language. In this study, we find that electrophysiological oscillations in human cerebral cortex become entrained to quasi-periodic fluctuations of visual movement in sign language. In fluent signers, cortical entrainment to sign language was found between 0.4 and 5 Hz, peaking at about 1 Hz, and emerged most robustly over occipital and central EEG channels. These results show that the human brain entrains to low-frequency variability in language whether it is perceived with the ears or eyes. Visual cortex flexibly phase-locks to visible changes in sign language, analogously to the way auditory cortex phase-locks to amplitude changes in oral speech. Our findings argue that flexible entrainment depends on mechanisms that are not specific to any given effector or sensory modality.

Prior results suggest that auditory and visual perception are differentially modulated by rhythms at different frequencies [16]. Auditory sensitivity varies as a function of the power and phase of spontaneous 2–6 Hz rhythms [28], and these oscillations are entrained by sounds modulated at 3 Hz [29, 30]. Visual sensitivity, by contrast, depends on the power and phase of spontaneous alpha rhythms [31–33], and electrophysiological oscillations in visual cortex are robustly entrained by periodic stimulation around 10 Hz [19, 21]. When humans watch a light flicker at frequencies from 1–100 Hz, visual cortex shows the strongest entrainment around 10 Hz [18].

Although these rhythmic preferences are not absolute (visual cortex also shows rhythmic oscillations in the delta/theta range [11, 17, 34–37]), differences between the sensory modalities are apparent when different frequency bands are directly compared. Auditory detection sensitivity depends on the phase of underlying delta-theta but not alpha oscillations [28]. In response to aperiodic stimulation, visual cortex oscillates in the alpha band [20], whereas auditory cortex does not show consistent oscillatory activity [38].

If sensory-specific oscillatory preferences determined the spectrum of entrainment, then peak coherence to sign language should be observed at the frequencies preferred by visual cortex: around 10 Hz. Contrary to this prediction, we find that cortical coherence to sign language only emerges below 5 Hz. Cerebral cortex entrains to sign language around the frequencies of words and phrases in ASL.

Coherence across signers and non-signers. We find that cerebral cortex phase-locks to visual changes in ASL both in fluent signers and in people with no knowledge of sign language. In

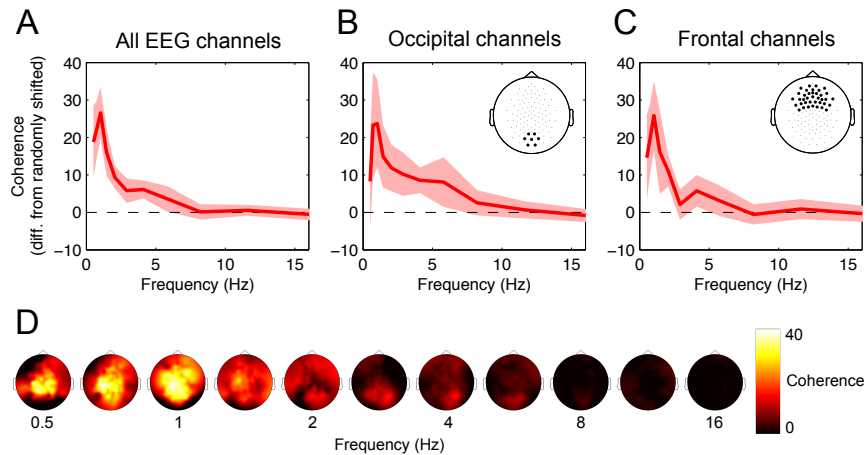


Fig. 3. Coherence between cortex and the IVC of sign language in fluent signers ($N = 13$). (A) Coherence spectrum averaged across all EEG channels. For each participant, we computed the difference in empirical coherence and a distribution of cortical coherence to randomly shifted IVC. The solid line shows the mean difference in empirical and randomized coherence across participants, and the shaded area shows the 95% CI around the mean. The dotted line shows chance levels. (B) Coherence spectrum averaged over occipital channels. Inset shows the location of selected channels. (C) Coherence spectrum averaged over frontal channels. Inset shows the location of selected channels. (D) Scalp topography of coherence in each frequency bin, averaged across participants.

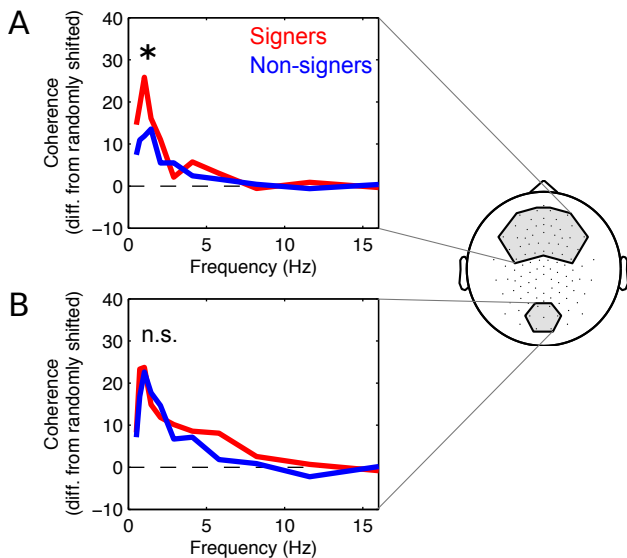


Fig. 4. Comparison of coherence in fluent signers ($N = 13$) and non-signers ($N = 15$). (A) Coherence at frontal channels (channel selection illustrated at right) was stronger in signers than in non-signers. (B) Coherence at occipital channels did not differ between groups. Data from signers is the same as in Fig. 3. *, $P < .05$; n.s., not statistically significant.

principle, coherence to ASL in non-signers could emerge for two reasons. Coherence could be driven either bottom-up, by sensory stimulation, or top-down, by non-linguistic temporal predictions. Human bodies move in predictable ways, and non-signers could entrain to sign language based on these regularities in human movement.

In frontal areas, fluent signers showed stronger coherence than non-signers. This difference in frontal coherence may reflect top-down sensory predictions based on knowledge of ASL. Alternatively, differences in coherence could reflect differences in attention to the videos. Cortical entrainment to oral speech decreases when people direct attention away from the speech stimulus [13, 39]. Reduced coherence in non-signers,

therefore, would also be predicted if non-signers do not attend to videos of ASL as strongly as fluent signers do. However, our findings at occipital channels argue against this possibility. If differences between groups were driven by attention, then occipital coherence should be stronger in signers than in non-signers. However, we find no evidence that occipital coherence depends on linguistic knowledge. Taken together, these results suggest that although linguistic knowledge is not necessary for entrainment, signers may leverage knowledge about ASL to sharpen temporal predictions during language comprehension. These sharpened predictions result in stronger entrainment in the frontal regions that exert top-down control over visual perception.

Specialization for speech? Syllables in oral speech occur at frequencies that largely overlap with cortical entrainment to the volume envelope. This overlap could be interpreted as evidence for a specialized oscillatory mechanism for speech comprehension. This type of speech-specific mechanism could evolve in at least two ways. First, “the articulatory motor system [may have] structured its output to match those rhythms the auditory system can best apprehend” [6]. Second, auditory mechanisms may have developed to comprehend speech based on the timing of preexisting oral behaviors. Non-human primates create vocalizations and facial displays that fluctuate at frequencies similar to human speech syllables [40], and their attention is preferentially captured by faces that move at these frequencies [41]; perhaps auditory processing evolved to fit the timing profile of these behaviors.

The data we report here, however, suggest that entrainment may not have any close evolutionary link to oral speech. Instead, a more general process may underlie cortical phase-locking to variability in language. Previous results are consistent with this interpretation as well. When participants watch videos of speech, entrainment emerges not only in auditory cortex, but also in visual cortex [11, 35, 36]. Furthermore, cortical rhythms entrain to rhythms in music [42] and to other rhythmic stimuli in audition [29, 30] and vision [19, 43]. These examples of low-frequency cortical entrainment to a broad

range of stimuli across sensory modalities suggest that the cortical mechanisms supporting entrainment to the volume envelope of speech may be a specialized case of a general predictive process.

Neural mechanisms of language comprehension across sensory modalities. Previous studies have shown that the functional neuroanatomy of speech largely overlaps with that of sign [44]. At the coarsest level of anatomical specificity, the left hemisphere is specialized for spoken language. The left hemisphere is also asymmetrically active during sign language perception [45] and production, regardless of which hand people use to sign [46, 47]. Left hemisphere damage, furthermore, results in linguistic deficits in signing patients [48].

Specific regions within the left hemisphere show similar involvement in processing both speech and sign. Across signed and spoken language, bloodflow increases to the LIFG and left inferior parietal lobe (IPL) during phonemic discrimination [49, 50] and morphosyntactic processing [51]. Similarly, word production in both signed and spoken languages activates LIFG, left IPL, and left temporal areas [52].

Differences in the cortical areas involved in sign and speech can often be attributed to differences in the form of these languages. For example, comprehension of sign language activates primary visual but not primary auditory cortex [45]. Consistent with the fact that sign language relies on spatial contrasts, inferior and superior parietal cortex is more strongly active during signed than during spoken language production [52] and perception [49].

Our findings go beyond functional neuroanatomy to examine neurophysiological processes that can arise in multiple cortical areas. We show that oscillatory entrainment to low-frequency variability in the stimulus occurs regardless of whether language is being processed using auditory cortex or visual cortex.

Our results differ from previous studies on entrainment to speech primarily in the scalp topography of coherence. Entrainment to auditory speech is strongest over auditory cortex [2, 8, 11, 35] and central frontal sites [14]. By contrast, our results show that entrainment to sign language is strongest at occipital and parietal channels, consistent with greater parietal activation during sign compared with speech [53]. This difference likely reflects increased visual and spatial demands of perceiving sign language.

The IVC quantifies temporal structure in visual perception.

The Instantaneous Visual Change (IVC) provides a novel method for examining gross temporal structure in natural visual stimuli. Analogously to the way the broadband envelope summarizes early stages of auditory processing, the IVC provides a first approximation of the magnitude of information available to the earliest stages of visual processing. At the first stage of auditory transduction, hair cells in the cochlea extract the narrowband envelope of sounds. Summing these narrowband envelopes together yields the overall auditory responses over time: the broadband envelope. In the retina, center-surround retinal ganglion cells respond to changes in the brightness of specific wavelengths of light. Summing the responses from these cells yields the overall visual responses over time, approximated by the IVC.

The IVC provides a coarse index of visual information in sign language, just as the broadband envelope provides a

coarse index of information in speech. For example, the volume envelope does not reflect small spectral differences that are crucial for discriminating vowels. The IVC, analogously, does not preserve information about which effectors are moving or their trajectories. Nevertheless, sign language comprehenders may use the IVC of sign heuristically, as listeners use the acoustic envelope of speech, to anticipate when important information is likely to appear.

In the current study, we use the IVC to characterize temporal structure in sign language, and to examine responses of the human brain to that temporal structure. The IVC could also be applied to study temporal structure in other domains, such as gesture, biological motion, and movement in natural scenes.

The functional role of entrainment to language. Oscillatory entrainment to language may be a specific case of a general cortical mechanism. In primates, spiking probability varies with the phase of low-frequency oscillations: neurons are most likely to fire at specific points in the phase of ongoing oscillations [54, 55]. Perhaps the cortex strategically resets the phase of ongoing neural oscillations to ensure that perceptual neurons are in an excitable state when new information is likely to appear [5–7, 14]. Oscillatory entrainment may constitute a cortical strategy to boost perceptual sensitivity at informational peaks in language. Our findings suggest that the brain can flexibly entrain to linguistic information regardless of the modality in which language is produced or perceived.

Materials and Methods

Participants watched approximately 20 minutes of naturalistic storytelling in American Sign Language (ASL) while EEG was recorded. Participants were instructed to watch the videos and remain still and relaxed. All procedures were approved by the Institutional Review Board of the University of Chicago. Detailed methods and analyses are available in the online Supporting Text.

Participants. Participants had corrected-to-normal vision and reported no history of epilepsy, brain surgery, or traumatic brain injuries. Informed consent was obtained before beginning the experiment, and participants were paid \$20/hour for their participation. We recorded EEG from 16 fluent signers of American Sign Language (ASL). Data from 2 participants were excluded before analysis due to excessive EEG artifacts, and data from 1 participant were lost due to experimenter error. All participants retained in the analyses began learning ASL by 5 years of age ($N = 13$; 3 female, 10 male; age 24–44; mean age of acquisition 1.1 years). Participants who used hearing aids or cochlear implants removed the devices before beginning the experiment. A fluent speaker of ASL (J.L.) answered participants' questions about the study. We recorded EEG from an additional 16 participants who had no prior exposure to ASL. These participants were recruited from the University of Chicago community through online postings. One participant who was currently learning ASL was excluded before analyses, leaving $N = 15$ non-signing participants (10 female, 5 male; age 18–31).

Instantaneous Visual Change (IVC). The IVC represents a time-series of aggregated visual change between frames (Fig. 1), and is computed as the sum of squared differences in each pixel across sequential frames of video:

$$\text{IVC}(t) = \sum_i [x_i(t) - x_i(t-1)]^2$$

where x is the grayscale value of pixel i at time t . Python code for the IVC is available at <http://casasanto.com/geoffbrookshire/ivc.py>.

EEG analysis. See Supporting Text for details about EEG acquisition and preprocessing. To compute coherence, IVC and EEG data were filtered into overlapping log-spaced frequency bins using phase-preserving forward-reverse Butterworth bandpass filters. Bins were centered on values from 0.5–16 Hz, and included frequencies in the range $(0.8f, 1.25f)$, where f is the center frequency $f = 2^n$ for $n \in \{-1, -0.5, 0, \dots, 4\}$. Instantaneous phase and power were determined with the Hilbert transform. Power was computed as the absolute value of the analytic signal, and phase as the angle of the analytic signal. These power and phase estimates were then used to calculate coherence:

$$Coh = \frac{|\sum_t (e^{i\theta t} \sqrt{P_{C,t} \cdot P_{V,t}})|}{\sqrt{\sum_t (P_{C,t} \cdot P_{V,t})}}$$

where t is the time point, θ is the phase difference between the IVC and EEG, P_V is power in the IVC, and P_C is power in the EEG recording [9]. Statistical significance of coherence was determined by a two-stage randomization procedure. First, the IVC was randomly shifted to obtain a null distribution of coherence between the two signals. Second, statistical significance was determined using cluster-based permutation tests [56] (Supporting Text).

ACKNOWLEDGMENTS. This research was supported by a William Orr Dingwall Neurolinguistics Fellowship (to G.B.), a research grant from the University of Chicago Center for Gesture, Sign, and Language (to G.B. and J.L.), the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-1144082 (to J.L.), the Institute of Education Sciences, U.S. Department of Education, through Grant R305B140048 (to J.L.) and McDonnell Scholar award 220020236 (to D.C.). Images from *The Rosa Lee Show* are used courtesy of Rosa Lee Timm.

- Ahissar E et al. (2001) Speech comprehension is correlated with temporal response patterns recorded from auditory cortex. *Proc Natl Acad Sci USA* 98(23):13367–13372.
- Luo H, Poeppel D (2007) Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron* 54(6):1001–1010.
- Gross J et al. (2013) Speech rhythms and multiplexed oscillatory sensory coding in the human brain. *PLoS Biol* 11(12):e1001752.
- Park H, Ince RA, Schyns PG, Thut G, Gross J (2015) Frontal top-down signals increase coupling of auditory low-frequency oscillations to continuous speech in human listeners. *Curr Biol* 25(12):1649–1653.
- Schroeder CE, Lakatos P (2009) Low-frequency neuronal oscillations as instruments of sensory selection. *Trends Neurosci* 32(1):9–18.
- Giraud AL, Poeppel D (2012) Cortical oscillations and speech processing: emerging computational principles and operations. *Nat Neurosci* 15(4):511–517.
- Peelle JE, Davis MH (2012) Neural oscillations carry speech rhythm through to comprehension. *Front Psychol* 3(320):1–17.
- Howard MF, Poeppel D (2010) Discrimination of speech stimuli based on neuronal response phase patterns depends on acoustics but not comprehension. *J Neurophysiol* 104(5):2500–2511.
- Doelling KB, Arnal LH, Ghitza O, Poeppel D (2014) Acoustic landmarks drive delta–theta oscillations to enable speech comprehension by facilitating perceptual parsing. *NeuroImage* 85:761–768.
- Ding N, Chatterjee M, Simon JZ (2014) Robust cortical entrainment to the speech envelope relies on the spectro-temporal fine structure. *NeuroImage* 88:41–46.
- Luo H, Liu Z, Poeppel D (2010) Auditory cortex tracks both auditory and visual stimulus dynamics using low-frequency neuronal phase modulation. *PLoS Biol* 8(8):e1000445.
- Crosse MJ, Butler JS, Lalor EC (2015) Congruent visual speech enhances cortical entrainment to continuous auditory speech in noise-free conditions. *J Neurosci* 35(42):14195–14204.
- Zion Golumbic EM et al. (2013) Mechanisms underlying selective neuronal tracking of attended speech at a “cocktail party”. *Neuron* 77(5):980–991.
- Kayser SJ, Ince RA, Gross J, Kayser C (2015) Irregular speech rate dissociates auditory cortical entrainment, evoked responses, and frontal alpha. *J Neurosci* 35(44):14691–14701.
- Ahissar E, Ahissar M (2005) Processing of the temporal envelope of speech in *The Auditory Cortex: A synthesis of human and animal research*, eds. König R, Heil P, Budinger E, Scheich H. (Lawrence Erlbaum Associates, Inc., London), pp. 295–313.
- VanRullen R, Zoefel B, Ilhan B (2014) On the cyclic nature of perception in vision versus audition. *Phil Trans R Soc B* 369(1641):20130214.
- Landau AN, Fries P (2012) Attention samples stimuli rhythmically. *Curr Biol* 22(11):1000–1004.
- Herrmann CS (2001) Human eeg responses to 1–100 hz flicker: resonance phenomena in visual cortex and their potential correlation to cognitive phenomena. *Exp Brain Res* 137(3-4):346–353.
- Mathewson KE et al. (2012) Making waves in the stream of consciousness: entraining oscillations in eeg alpha and fluctuations in visual awareness with rhythmic visual stimulation. *J Cognitive Neurosci* 24(12):2321–2333.
- VanRullen R, Macdonald JS (2012) Perceptual echoes at 10 Hz in the human brain. *Curr Biol* 22(11):995–999.
- Spaak E, de Lange FP, Jensen O (2014) Local entrainment of alpha oscillations by visual stimuli causes cyclic modulation of perception. *J Neurosci* 34(10):3536–3544.
- Chandrasekaran C, Trubanova A, Stillitano S, Caplier A, Ghazanfar AA (2009) The natural statistics of audiovisual speech. *PLoS Comput Biol* 5(7):e1000436.
- Meier RP (2002) Why different, why the same? Explaining effects and non-effects of modality upon linguistic structure in sign and speech. in *Modality and Structure in Signed and Spoken Languages*, eds. Meier RP, Cormier K, Quinto-Pozos D. (Cambridge University Press, Cambridge), pp. 1–26.
- Petitto LA, Holowka S, Sergio LE, Levy B, Ostry DJ (2004) Baby hands that move to the rhythm of language: hearing babies acquiring sign languages babble silently on the hands. *Cognition* 93(1):43–73.
- Bellugi U, Fischer S (1972) A comparison of sign language and spoken language. *Cognition* 1(2):173–200.
- Hwang SOK (2011) Ph.D. thesis (University of Maryland).
- Greenberg S, Carvey H, Hitchcock L, Chang S (2003) Temporal properties of spontaneous speech: a syllable-centric perspective. *J Phon* 31(3):465–485.
- Ng BSW, Schroeder T, Kayser C (2012) A precluding but not ensuring role of entrained low-frequency oscillations for auditory perception. *J Neurosci* 32(35):12268–12276.
- Henry MJ, Obleser J (2012) Frequency modulation entrains slow neural oscillations and optimizes human listening behavior. *Proc Natl Acad Sci USA* 109(49):20095–20100.
- Hickok G, Farahbod H, Saberi K (2015) The rhythm of perception entrainment to acoustic rhythms induces subsequent perceptual oscillation. *Psychol Sci* pp. 1006–13.
- Hanslmayr S et al. (2007) Prestimulus oscillations predict visual perception performance between and within subjects. *NeuroImage* 37(4):1465–1473.
- Van Dijk H, Schoffelen JM, Oostenveld R, Jensen O (2008) Prestimulus oscillatory activity in the alpha band predicts visual discrimination ability. *J Neurosci* 28(8):1816–1823.
- Mathewson KE, Gratton G, Fabiani M, Beck DM, Ro T (2009) To see or not to see: prestimulus α phase predicts visual awareness. *J Neurosci* 29(9):2725–2732.
- Busch NA, Dubois J, VanRullen R (2009) The phase of ongoing eeg oscillations predicts visual perception. *J Neurosci* 29(24):7869–7876.
- Park H, Kayser C, Thut G, Gross J (2016) Lip movements entrain the observers’ low-frequency brain oscillations to facilitate speech intelligibility. *eLife* 5:e14521.
- Power AJ, Mead N, Barnes L, Goswami U (2012) Neural entrainment to rhythmically presented auditory, visual, and audio-visual speech in children. *Front Psychol* 3(216):10–3389.
- Cravo AM, Rohenkohl G, Wyart V, Nobre AC (2013) Temporal expectation enhances contrast sensitivity by phase entrainment of low-frequency oscillations in visual cortex. *J Neurosci* 33(9):4002–4010.
- Ilhan B, VanRullen R (2012) No counterpart of visual perceptual echoes in the auditory system. *PLoS ONE* 7(11):e49287.
- O’Sullivan JA et al. (2015) Attentional selection in a cocktail party environment can be decoded from single-trial eeg. *Cereb Cortex* 25(7):1697–1706.
- Ghazanfar AA, Takahashi DY, Mathur N, Fitch WT (2012) Cineradiography of monkey lip-smacking reveals putative precursors of speech dynamics. *Curr Biol* 22(13):1176–1182.
- Ghazanfar AA, Morrill RJ, Kayser C (2013) Monkeys are perceptually tuned to facial expressions that exhibit a theta-like speech rhythm. *Proc Natl Acad Sci USA* 110(5):1959–1963.
- Doelling KB, Poeppel D (2015) Cortical entrainment to music and its modulation by expertise. *Proc Natl Acad Sci USA* 112(45):E6233–E6242.
- Lakatos P, Karmos G, Mehta AD, Ulbert I, Schroeder CE (2008) Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science* 320(5872):110–113.
- MacSweeney M, Capek CM, Campbell R, Woll B (2008) The signing brain: the neurobiology of sign language. *Trends Cogn Sci* 12(11):432–440.
- Leonard MK et al. (2012) Signed words in the congenitally deaf evoke typical late lexico-semantic responses with no early visual responses in left superior temporal cortex. *J Neurosci* 32(28):9700–9705.
- Corina DP, San Jose-Robertson L, Guillemin A, High J, Braun AR (2003) Language lateralization in a bimodal language. *J Cognitive Neurosci* 15(5):718–730.
- Gutierrez-Sigut E et al. (2015) Language lateralization of hearing native signers: A functional transcranial doppler sonography (ftcd) study of speech and sign production. *Brain Lang* 151:23–34.
- Hickok G, Bellugi U, Klima ES (1998) The neural organization of language: evidence from sign language aphasia. *Trends Cogn Sci* 2(4):129–136.
- MacSweeney M, Waters D, Brammer MJ, Woll B, Goswami U (2008) Phonological processing in deaf signers and the impact of age of first language acquisition. *NeuroImage* 40(3):1369–1379.
- Williams JT, Darcy I, Newman SD (2015) Modality-independent neural mechanisms for novel phonetic processing. *Brain Res* 1620:107–115.
- Newman AJ, Supalla T, Hauser P, Newport EL, Bavelier D (2010) Dissociating neural subsystems for grammar by contrasting word order and inflection. *Proc Natl Acad Sci USA* 107(16):7539–7544.
- Emmorey K, Mehta S, Grabowski TJ (2007) The neural correlates of sign versus word production. *NeuroImage* 36(1):202–208.
- Emmorey K, McCullough S, Mehta S, Grabowski TJ (2014) How sensory-motor systems impact the neural organization for language: direct contrasts between spoken and signed language. *Front Psychol* 5:484.
- Jacobs J, Kahana MJ, Ekstrom AD, Fried I (2007) Brain oscillations control timing of single-neuron activity in humans. *J Neurosci* 27(14):3839–3844.
- Lakatos P et al. (2005) An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *J Neurophysiol* 94(3):1904–1911.
- Maris E, Oostenveld R (2007) Nonparametric statistical testing of eeg-and meg-data. *J Neurosci Meth* 164(1):177–190.
- Lee R, Timm D (2008) *The Rosa Lee Show*.
- Oostenveld R, Fries P, Maris E, Schoffelen JM (2010) FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci* 2011.

Supporting Information:

Visual cortex entrains to sign language

Geoffrey Brookshire Jenny Lu Howard Nusbaum
Susan Goldin-Meadow Daniel Casasanto

SI Materials and Methods

Spectral analysis of speech and sign

We analyzed temporal structure in multiple spoken and signed languages. These analyses were performed using custom software written in Python.

Spoken language samples comprised audio recordings of stories in 9 languages from 5 languages families (N=12 samples, total duration 1:07:28; Table S1). Before computing the spectra, these recordings were trimmed to a maximum of 6 min (mean 5:37, SD 0:52), mixed down to 1 output channel, and downsampled to 22050 Hz.

Sign language samples were chosen to be as similar as possible to the spoken samples (N=14, total duration 1:10:22, mean 5:02, SD 3:19; Table S2), representing 4 languages from 3 families. All samples comprised forward-facing views of a single speaker. In 12 videos, the speaker told a story; in the remaining 2 videos, the speaker gave an informational speech. The IVC was calculated for each video.

All IVC and broadband envelope recordings were resampled to a common frequency of 30 Hz, and power was normalized by dividing out the SD of each signal. Power spectra were computed using Welch’s method. Signals were split into 2.13-s long segments (2^6 samples) that overlapped by 1.07 s (2^5 samples). A Hanning window was applied to each segment, and the linear trend was removed. Fast Fourier transforms (FFT) were then computed for each segment. The spectrum for each signal was obtained by averaging the spectra in all segments.

We fit sign language IVC spectra to a $1/f$ function using least squares regression by transforming the equation $Y = Af^{-\alpha}$ into the linear form $\log Y = \log A - \alpha \log f$, with power Y and frequency f [22]. We tested for deviations from the $1/f$ trend using 1-sample t-tests.

Broadband envelope

The broadband envelope was computed over samples of oral speech by adapting methods from previous studies [22]. First, the raw waveform was band-pass filtered into 25 logarithmically-spaced frequency bands from 0.1–10 kHz using least-squares filters with 501 taps. The narrowband envelope of each filtered signal was then calculated as the absolute value of the Hilbert transform. The narrowband envelopes were summed to obtain the broadband envelope for each recording.

Stimuli

Stimuli comprised two videos of ASL storytelling: “Kondima” and “Little Feet” from the Rosa Lee Show [57]. Each video was approximately 10 min long with a 30 Hz frame-rate, and depicted a native speaker of ASL telling a story against a static background. To ensure that the timing of the video was accurately matched with EEG recordings, a small white square flashed in the corner of the display once every 30 frames of video (out of view of the participant), and was registered by a photodiode connected to the EEG amplifier. In total, the session lasted 60–90 minutes.

EEG acquisition and preprocessing

EEG was recorded at 250 Hz using a 128-channel net (Electrical Geodesics, Eugene, OR). Impedances were reduced to $< 50 \text{ k}\Omega$ before participants watched each story. EEG analyses were performed in Matlab using custom software and the FieldTrip package [58]. Recordings from electrodes along

the face, beneath the ears, and at the base of the neck were excluded before any analysis, leaving 103 channels in all analyses. Electrode movement artifacts were manually identified and rejected by replacing the tagged region with zeros and applying a 4000-ms half-Hann taper to each side of the artifact. Artifacts from blinks and eye-movements were identified and removed using independent component analysis (ICA). To ensure that the IVC of each story was accurately matched to the EEG recordings, we used cubic spline interpolation to warp the IVC to the time-stamps registered by the photodiode. This process simultaneously resampled the IVC from 30 Hz to 250 Hz. Before computing coherence, EEG signals were re-referenced to the average mastoids.

EEG statistical testing

Statistical significance of EEG coherence to the IVC was determined by a two-stage randomization procedure. To obtain a null distribution of coherence, the onset of the IVC was circularly shifted to a randomly selected starting point, and coherence was computed between EEG signals and the shifted IVC. This procedure preserves the spectrotemporal characteristics of both signals, but eliminates any relationship between them. For each subject, we computed 100 randomly shifted baselines. Next, we tested for significant differences between the empirical and randomly shifted coherence using a cluster-based non-parametric permutation test [56]. This test looked for a difference in coherence between the empirical and randomly shifted data, contrasted with $N=10,000$ permutations in which the ‘empirical’ data was randomly selected from the group of empirical and randomly-shifted traces. For each frequency and each channel, a T-statistic was computed on the difference between empirical and randomly shifted data using a dependent-samples regression. The test statistic was computed as the maximum cluster size in each permutation (cluster threshold: $\alpha = .01$, two-tailed). The p-value was calculated as the proportion of permuted cluster statistics that were more extreme than the empirical value.

In previous studies, cortical entrainment to speech is often strongest over auditory cortex [2, 8, 11, 35] and frontal cortex [14]. We tested coherence to sign language in two analogous regions of interest (ROIs): at occipital channels, and at frontal channels.

To test whether knowledge of ASL influenced the strength of entrainment, we performed cluster permutation analyses on the difference in coherence between signers and non-signers at the occipital and frontal ROIs. For each participant, we normalized the data by computing the Z-score of their empirical coherence against the randomly shifted baseline coherence values. Z-scored empirical coherence was then compared between groups using cluster permutation tests.

Data deposition

All data, stimuli, and analysis scripts are stored in a repository at the University of Chicago, and can be obtained by contacting the corresponding author.

Figures

Videos

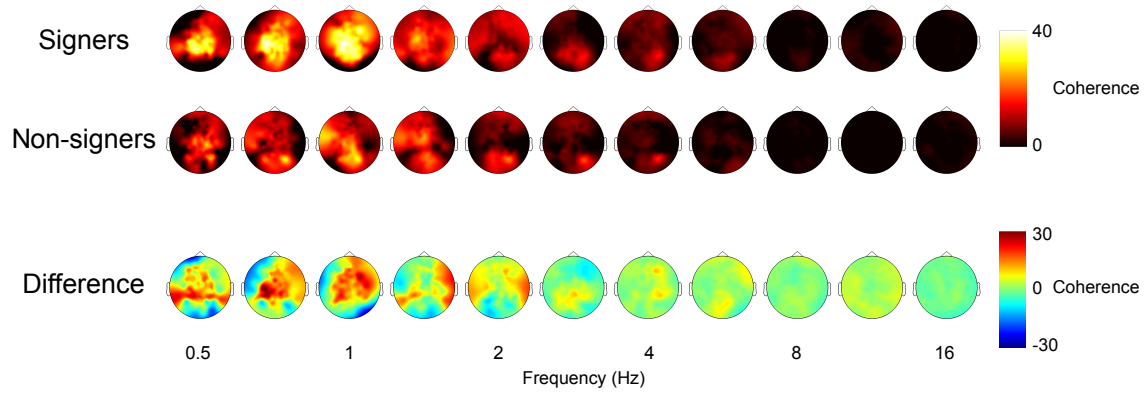
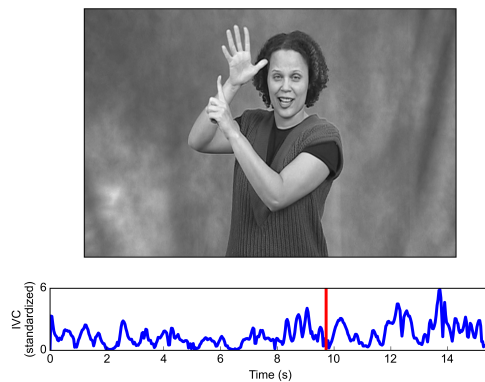


Fig. S1. Comparison of coherence across signing and non-signing participants. Coherence was robustly present in both groups of participants, regardless of experience with sign language. The top row shows coherence in each frequency bin for fluent signers, and the middle row shows coherence for non-signers. The bottom row shows the difference in coherence between the two groups. Positive values (shown in red) indicate stronger coherence in signers, and negative values (blue) indicate stronger coherence in non-signers.



Video S1. Example video illustrating how the Instantaneous Visual Change (IVC) tracks visual movement during sign language.