Sign language learning increases temporal resolution of visual attention

Serpil Karabüklü[1], Sandra Wood[2], Charles Bradley[1], Ronnie B. Wilbur[1], Evie A. Malaia[3]

[1]Department of Speech, Language, and Hearing, Purdue University, West Lafayette, IN 47406, USA
[2]Department of Linguistics, University of Southern Maine, Portland, ME 04104, USA
[3] Department of Communicative Disorders, University of Alabama, Tuscaloosa, AL 35401, USA

Corresponding Author:
Evie A. Malaia
Department of Communicative Disorders, University of Alabama,
Tuscaloosa, AL 35487-0231, USA
Email: eamalaia@ua.edu

Keywords: temporal resolution, visual cognition, ASL, sign language learning

ABSTRACT

Spatiotemporal parameters of sign language input differ substantially from those of everyday experience: the motion of articulators during signing is overtly visible, and extremely fast, in the range of 1-4 m/s (Bosworth et al., 2019; Malaia et al., 2016). And while for auditory modality, the ability to process temporal and spectral resolution of the signal is well established as critical for language development, spectrotemporal parameters of visual attention necessary for sign language comprehension have not been investigated. We studied the trajectory of attentional temporal resolution change in sign language learners to assess whether the physical parameters of sign language input results in changes in perceptual sampling among learners. Using a flicker paradigm (Farzin et al., 2011), we found that exposure to sign language leads to increased accuracy in identifying an out-of-phase visual flicker object at up to 60 Hz in 3rd semester learners of ASL. Remarkably, as sign language proficiency is gained beyond that, perceptual sampling appears to revert to normal levels recorded in adults (10-20 Hz). This suggests that temporal resolution of visual attention is malleable, and can be altered by environmental exposure, such as sign language learning.

Introduction

The motion of hand articulators during signing – i.e. visual linguistic communication - is very fast, and distinctly different from both everyday visual scenes, and non-linguistic hand motion (Bosworth et al., 2019; Malaia et al., 2016). Across the spatial frequency range, sign language communication relies on higher frequencies than those in everyday visual scenes (Bosworth et al., 2006). With regard to temporal parameters, high speed of motion and change of linguistically meaningful features, such as handshape, place of articulation, and hand orientation, generate visual scenes of high dynamic variability, which in turn results in a very specific perceptual experience for signers and those learning to sign (Borneman et al., 2018; Gurbuz et al., 2020; Malaia et al., 2016). For those who use sign language regularly, exposure to this specific visual input could modify spatiotemporal parameters of visual processing (Bosworth et al., 2019).

Sign language communication encompasses several information-carrying channels, which need to be processed simultaneously: while hands are the primary articulators, the position of the head and the body, as well as language-specific facial actions, also encode prosodic, pragmatic, semantic, and syntactic information (Malaia et al., 2018; Wilbur, 2000). Likely as a reflection of these physical properties of sign language communication, increased spatial attention to high-frequency stimuli (Bosworth & Dobkins, 2002a, 2002b), and enhancements of peripheral vision (Bavelier et al., 2006; Dye et al., 2009; Proksch & Bavelier, 2002) were documented in signers.

It has long been noted that signers and non-signers differ drastically in their perception of rapidly changing visual stimuli: e.g. Klima et al. (1999) showed that when signers and non-signers viewed point-light displays of 'writing', the users of different sign languages could differentiate between 'strokes' (information-bearing portions of point-light movement) from rapid 'transitions' (movement of the point-light from the end of one meaningful portion to the beginning of another), while non-signers could not. The difficulties of non-signers in correctly perceiving signed input might be due to lower temporal resolution of visual change perception as compared to signers, to differences in spectral allocation of visual attention, or to combination of the two (Proksch & Bavelier, 2002). Several researchers also noted that the age of sign language acquisition and the length of exposure to sign language appear to affect the sign production patterns and perceptual abilities of proficient signers (Bosworth et al., 2019; Malaia et al., 2020). In sign production by learners of American Sign Language, utterances, signs, and transitions between signs produced are longer in duration as compared to native signing, confirming findings that rhythmic characteristics of movement are important to making judgments of native fluency (Cull, 2014; Kantor, 1978). A study of early and late learners of Swiss German Sign Language also indicated that phrase-structure sensitive changes in signing dynamics (i.e. prosody) can be entirely absent in late learners (Braem, 1999).

Attentional allocation and perceptual modifications in visual modality resulting from exposure to and use of sign languages in native signers have been well-documented (Bavelier et al., 2006; Codina et al., 2017; Dye et al., 2009). However, for sign language learners, the exact nature of changes in allocation of visual attention and the timeline of these changes are not well understood. While non-signing adults cannot isolate and identify the individual flashes that compose the flicker beyond frequencies of about 10 Hz (Farzin et al. 2011), sign language users are able to make judgments based on much higher temporal resolution in biological motion (Klima et al. 1999).

In two studies presented below, we have manipulated spatial and temporal parameters of simple perceptual stimuli (square-wave flicker) to probe temporal resolution of visual attention in college-level learners of American Sign Language at various stages of proficiency. Based on prior research, we hypothesized that: 1) sign language learning is likely to result in ability to perceptually isolate objects at higher temporal resolutions; and 2) higher spatial frequency stimuli, which are closer to spectral properties of sign language input, are likely to allow individuation at higher frequencies in sign language learners at the same level of proficiency.

Methods

Participants

48 participants (F N=35, age M=22, SD =5; age range 17 to 43 years) took part in the study at two sites: Purdue University (IRB # 1212012994) and University of Southern Maine (IRB#...). Participants were students in classes of American Sign Language; course level at the date of data collection was coded as proficiency level (1-6, i.e. first to sixth semester of ASL). Participants were comfortably seated individually in a testing room in front of a computer screen. The stimuli were presented on a 15-in. LCD monitor with 60-Hz refresh rate. Participants were seated approximately 40 cm from the monitor. The task was programmed and presented using Presentation v. 22 (Neurobehavioral Systems).

Stimuli and procedure

Stimuli were composed of 4 blocks around a fixation point on the screen with neutral (grey) background (Fig 1A). Blocks changed colors between white and black, creating a flicker (Fig 1. B). In each trial, one of the blocks flickered in counter-phase with others: i.e. when three blocks were white, the fourth was black, and vice versa. Block color changes from black to white occurred at the rates of 10 Hz, 20 Hz, 30 Hz and 60 Hz. 16 trials were presented at each temporal frequency, in random order; a total of 64 trials were administered.

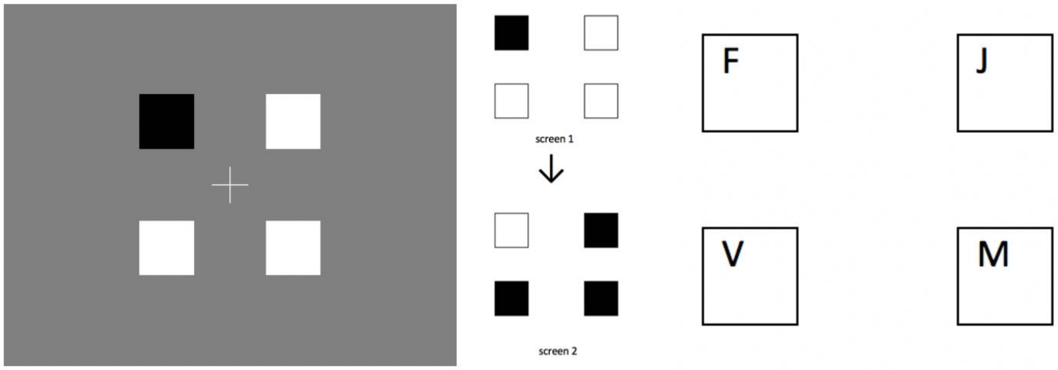

Figure 1. Stimuli presentation on the screen, flicker paradigm, and response keys.

Two versions of the stimuli were developed: for high visual frequency presentation (small blocks), block size was 2 sm on the screen, corresponding to .01 cyc/mrad visual frequency, and subtending ~3° of visual angle; for lower visual frequency presentation (large blocks), block size was 4 sm on the screen, corresponding to .005 cyc/mrad visual frequency, and subtending ~6° of visual angle. The order of stimuli presentation (small vs. large, or higher vs. lower visual frequency) was counterbalanced among participants, such that half the time the trial began with small block presentation. For a subset of participants (N=1), only large block data were collected due to equipment malfunction. This data is excluded from analysis.

The experiment was conducted as a four-alternative forced-choice (4-AFC) paradigm, where the participants were asked to look at the fixation point, try to decide which block is flickering out of synchronization, and press the corresponding key on the keyboard (Fig. 1 C). At slower rates of flicker, the target block is easier to identify, because participants are capable of individuation of the alternating black and white states. At frequencies above the threshold for phase individuation, however, all blocks appear to change color at the same time, and the out-of-phase block cannot be identified.

Participants' response times and accuracy were analyzed using 4 (temporal frequency) × 5 (ASL learning semester) repeated measures analysis of variance (ANOVA) separately for small (foveal field of view) and large blocks. Additionally, two-tailed t-tests ($p < .05$, Bonferroni corrected for multiple comparisons) were conducted to compare observed performance with chance-level performance (accuracy of 25%) at each temporal frequency and for each experience group.

Results

Accuracy

For higher spatial frequency stimuli (small blocks), a 4 (temporal frequency) × 5 (ASL level), repeated measures ANOVA revealed a significant main effect of temporal frequency, $F(3, 41) = 18.885$, $p <.001$, $\eta_p^2 = 0.580$, and a main effect of ASL level, $F(4, 40) = 3.80$, $p <.01$, $\eta_p^2 = 0.262$. For lower spatial frequency stimuli (large blocks), a 4 (temporal frequency) × 5

(ASL level) repeated measures ANOVA revealed a significant main effect of temporal frequency, $F(3, 41) = 9.647$, $p < .001$, $\eta_p^2 = 0.414$, and a main effect of ASL level, $F(4, 40) = 4.748$, $p < .003$, $\eta_p^2 = 0.306$. Additionally, the analysis revealed an interaction between temporal frequency and ASL level for these stimuli ($F(12, 129) = 2.012$, $p < .028$, $\eta_p^2 = 0.158$).

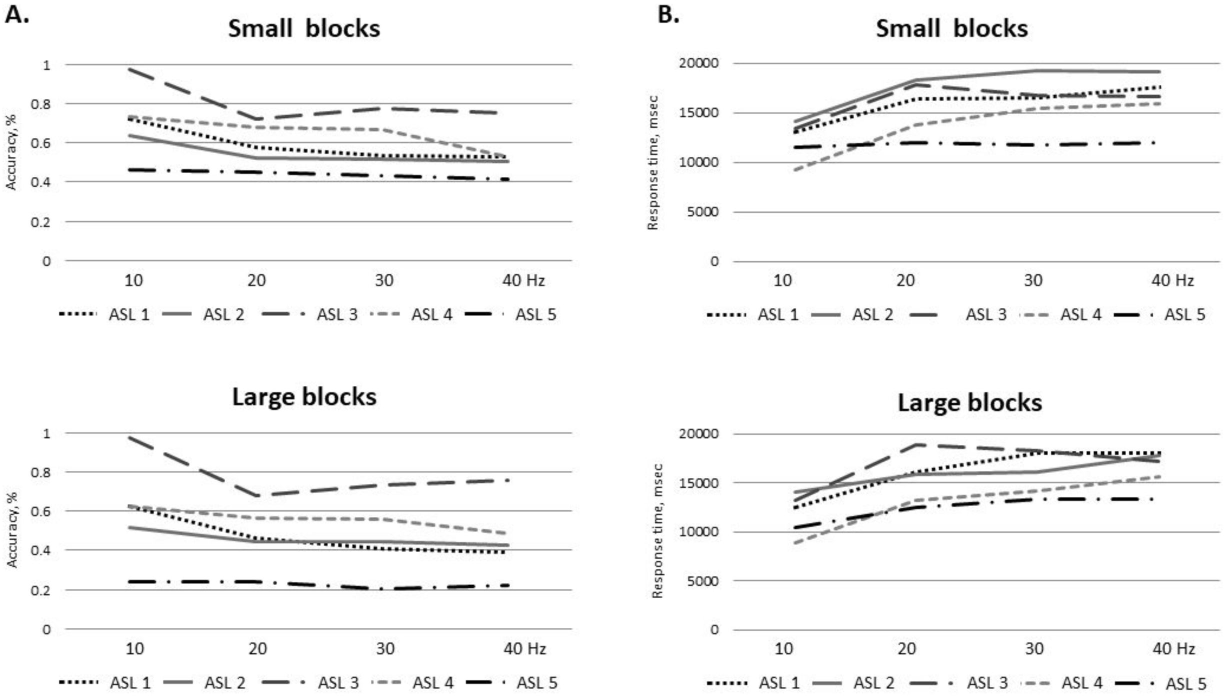

Figure 2. A) Average accuracy for ASL learners at each proficiency level. B) Average response time for ASL learners at each proficiency level.

Figure 2 presents average accuracy as a function of temporal frequency for ASL learners at each proficiency level for higher and lower spatial frequency stimuli (small vs. large blocks). ASL learners' accuracy in response to both spatial frequencies of stimuli across temporal frequencies increased between levels 1 and 3, and then decreased from those at level 3 at levels 4 and 6. Two-tailed t-tests to compare observed performance with chance-level performance (accuracy of .25) at each temporal frequency and for each experience group confirmed that for small squares (higher spatial frequencies), participants at all experience level determined the target square at all flicker rates above chance level ($p<.05$). However, for larger squares (lower spatial frequencies), performance above chance level was observed only for ASL level 1 at frequencies 10, 20, and 30; ASL level 3 (all temporal frequency levels); and ASL level 4 (frequency levels 10 and 20). Only ASL groups 3 showed accuracy higher than chance when the flicker rate of larger squares was at high temporal resolutions of 30 Hz and 60 Hz (30 Hz, ASL 3 group, t(6) =-4.779, p <.003; 60 Hz, ASL 3 group, t(6) = -6.879, p <.001).

Overall, the accuracy of performance appears to demonstrate an inverted U-shaped curve with experience: accuracy across all temporal and spatial frequencies reaches its peak at ASL level 3, and is lower before (ASL levels 1 and 2) and after (ASL levels 4 and 6).

Response time

For higher spatial frequency stimuli (small blocks), a 4 (temporal frequency) × 5 (ASL level) repeated measures ANOVA revealed a significant main effect of temporal frequency, $F(3, 41) = 9.957$, $p < .001$, $\eta_p^2 = 0.421$. For lower spatial frequency stimuli (large blocks), a 4 (temporal frequency) × 5 (ASL level) repeated measures ANOVA revealed a significant main effect of temporal frequency, $F(3, 41) = 6.965$, $p < .001$, $\eta_p^2 = 0.338$.

Figure 2B presents average response time as a function of temporal frequency for ASL learners at each proficiency level for higher and lower spatial frequency stimuli (small vs. large blocks). Overall, the response time to visual tasks appears to decrease with experience.

Discussion

Sign languages utilize extremely rapid changes of handshape, hand orientation, and place of articulation in visual production. The visual environment of a signer or sign learner thus differs drastically from that of everyday life. Our data suggests that sign language learners are able to increase, temporarily, visual sampling rate, likely as it becomes necessary to accurately and reliably identify visual parameters of articulation in proficient signers. This study is the first to characterize the sampling rate/temporal resolution of visual attention in sign language learners across proficiency levels.

We investigated whether the temporal resolution of visual attention, or attentional sampling rate at two different spatial frequencies can be affected by the process of learning American Sign Language. The results indicate that the attentional sampling rate at which visual events can be individuated increases rapidly during the first three semesters of ASL learning, and then decreases. Third-semester ASL learners demonstrated a marked increase in temporal resolution of visual sampling at both lower and higher spatial frequencies, performing flicker detection with above 70% accuracy at 60 Hz. This adaptive behavior is consistent with what is required by the spatiotemporal properties of visual events in sign language (Bosworth et al., 2019).

By presenting ASL learners with stimuli that had varying temporal frequency thresholds at which an out-of-phase flickering stimulus could be identified, we establish that the resolution of temporal visual attention in sign learners increases drastically as result of exposure to sign language: third-semester ASL students could individuate alternating states of flicker up to a rate of 60 Hz with accuracy above chance, while above-chance flicker threshold for first-semester learners was 10 Hz. The temporal resolution of attention, or visual sampling rate, of third-semester students was six times finer than that observed in first-semester learners, or attested in neurotypical adults in the literature. Our findings point to a rapid increase in resolution of temporal attention sign language learners, to the levels commensurate with recorded speeds of sign language articulator motion in everyday signing (Malaia et al., 2013; Malaia & Wilbur, 2012).

Another interesting finding in the present experiment is that resolution of temporal attention in sign language learners appears to decrease after approximately third semester of studying ASL, coinciding with further increase in ASL proficiency. This phenomenon might be related to increased mastery – ability to recognize discrete units - at multiple linguistic levels simultaneously (phonology, lexicon, syntax, and prosody), such that it becomes possible to 'fill in' visually under-sampled information from memorized units, and the higher rate of visual sampling becomes unnecessary.

Interestingly, spatial frequency of the stimuli affected accuracy of responses across proficiency levels. The stimuli with higher spatial frequency (small squares) elicited higher response accuracy at all levels of proficiency than stimuli with lower spatial frequency. Previous research indicated that high spatial frequencies dominate sign language experience (Bosworth et al., 2006), which might help explain the shifting of attentional focus to higher frequency stimuli. However, although temporal resolution of visual attention does appear to be higher for higher spatial frequencies, our data suggests that it is also malleable to increase for lower spatial frequencies in the process of exposure to sign language.

The finding that learners do increase their attentional resolution has implications for sign language teaching and learning, as well as for understanding sign language acquisition in infants, whose attentional temporal resolution lags far behind that of adults (Farzin et al., 2011). Human ability to sample the visual environment to identify individual events (such as change in handshape, or contact at the place of articulation) plays a crucial role in identifying phonological, semantic, and syntactic features in sign languages. Identification has to happen before perceptual binding of events across different parts of the visual field can take place for sign language comprehension (for example, a syntactic non-manual marker in sign language, such as eyebrow furrow, scopes over multiple manual signs in a signed sentence, and is critical for correct comprehension of ASL).

Our results also advance the understanding of perceptual sampling in adults, by demonstrating that attentional resolution of visual sampling is malleable, and can change depending on the needs of the individual and their current perceptual environment. Atypically high temporal resolution in mid-proficiency signers suggest that sign language learning has the potential to improve visual functions that require precise temporal sensitivity, such as attentional resolution for motion perception and tracking. While additional investigations would be needed to further characterize the relationship between temporal parameters of visual sampling/visual attention and the perceptual, cognitive, and motor skills of individuals learning sign language, identification of spatiotemporal frequency bands for attentional sampling due to experience is a promising venue.

Acknowledgements

Preparation of this manuscript was partially funded by grant #1734938 from the U.S. National Science Foundation to RBW and EM, grant #1932547 from the U.S. National Science Foundation to EM, and grant R01#108306 from the National Institutes for Health to RBW.

References:

Bavelier, D., Dye, M. W., & Hauser, P. C. (2006). Do deaf individuals see better? *Trends in Cognitive Sciences*, *10*(11), 512–518.

Borneman, J. D., Malaia, E., & Wilbur, R. B. (2018). Motion characterization using optical flow and fractal complexity. *Journal of Electronic Imaging*, *27*(05), 1. https://doi.org/10.1117/1.JEI.27.5.051229

Bosworth, R. G., Bartlett, M. S., & Dobkins, K. R. (2006). Image statistics of American Sign Language: Comparison with faces and natural scenes. *JOSA A*, *23*(9), 2085–2096.

Bosworth, R. G., & Dobkins, K. R. (2002a). The effects of spatial attention on motion processing in deaf signers, hearing signers, and hearing nonsigners. *Brain and Cognition*, *49*(1), 152–169. https://doi.org/10.1006/brcg.2001.1497

Bosworth, R. G., & Dobkins, K. R. (2002b). Visual field asymmetries for motion processing in deaf and hearing signers. *Brain and Cognition*, *49*(1), 170–181. https://doi.org/10.1006/brcg.2001.1498

Bosworth, R. G., Wright, C. E., & Dobkins, K. R. (2019). Analysis of the visual spatiotemporal properties of American Sign Language. *Vision Research*, *164*, 34–43. https://doi.org/10.1016/j.visres.2019.08.008

Braem, P. B. (1999). Rhythmic temporal patterns in the signing of deaf early and late learners of Swiss German Sign Language. *Language and Speech*, *42*(2–3), 177–208.

Codina, C. J., Pascalis, O., Baseler, H. A., Levine, A. T., & Buckley, D. (2017). Peripheral visual reaction time is faster in deaf adults and British sign language interpreters than in hearing adults. *Frontiers in Psychology*, *8*, 50.

Cull, A. (2014). *Production of Movement in Users of American Sign Language and its Influence on Being Identified as" Non-Native"*. Gallaudet University.

Dye, M. W., Hauser, P. C., & Bavelier, D. (2009). Is visual selective attention in deaf individuals enhanced or deficient? The case of the useful field of view. *PloS One*, *4*(5), e5640.

Farzin, F., Rivera, S. M., & Whitney, D. (2011). Time crawls: The temporal resolution of infants' visual attention. *Psychological Science*, *22*(8), 1004–1010.

Gurbuz, S. Z., Gurbuz, A. C., Malaia, E. A., Griffin, D. J., Crawford, C., Rahman, M. M., Aksu, R., Kurtoglu, E., Mdrafi, R., & Anbuselvam, A. (2020). A Linguistic Perspective on Radar Micro-Doppler Analysis of American Sign Language. *2020 IEEE International Radar Conference (RADAR)*, 232–237.

Kantor, R. (1978). Identifying native and second language signers. *Communication and Cognition Ghent*, *11*(1), 39–55.

Klima, E. S., Tzeng, O. J., Fok, Y. Y. A., Bellugi, U., Corina, D., & Bettger, J. G. (1999). From sign to script: Effects of linguistic experience on perceptual categorization. *Journal of Chinese Linguistics Monograph Series*, 96–129.

Malaia, E. A., Krebs, J., Roehm, D., & Wilbur, R. B. (2020). Age of acquisition effects differ across linguistic domains in sign language: EEG evidence. *Brain and Language*, *200*, 104708.

Malaia, E., Borneman, J. D., & Wilbur, R. B. (2016). Assessment of information content in visual signal: Analysis of optical flow fractal complexity. *Visual Cognition*, *24*(3), 246–251.

Malaia, E., Borneman, J. D., & Wilbur, R. B. (2018). Information transfer capacity of articulators in American Sign Language. *Language and Speech*, *61*(1), 97–112.

Malaia, E., & Wilbur, R. B. (2012). Kinematic signatures of telic and atelic events in ASL predicates. *Language and Speech*, *55*(Pt 3), 407–421. https://doi.org/10.1177/0023830911422201

Malaia, E., Wilbur, R. B., & Milkovic, M. (2013). Kinematic parameters of signed verbs. *Journal of Speech, Language, and Hearing Research: JSLHR*, *56*(5), 1677–1688. https://doi.org/10.1044/1092-4388(2013/12-0257)

Proksch, J., & Bavelier, D. (2002). Changes in the spatial distribution of visual attention after early deafness. *Journal of Cognitive Neuroscience*, *14*(5), 687–701.

Wilbur, R. B. (2000). Phonological and prosodic layering of nonmanuals in American Sign Language. *The Signs of Language Revisited: An Anthology to Honor Ursula Bellugi and Edward Klima*, 215–244.