Evie A. Malaia*, Joshua D. Borneman, Emre Kurtoglu, Sevgi Z. Gurbuz, Darrin Griffin, Chris Crawford and Ali C. Gurbuz

Complexity in sign languages

Linguistic and dimensional analysis of information transfer in dynamic visual communication

https://doi.org/10.1515/lingvan-2021-0005
Received January 12, 2021; accepted September 6, 2021; published online October 31, 2022

Abstract: Sign languages are human communication systems that are equivalent to spoken language in their capacity for information transfer, but which use a dynamic visual signal for communication. Thus, linguistic metrics of complexity, which are typically developed for linear, symbolic linguistic representation (such as written forms of spoken languages) do not translate easily into sign language analysis. A comparison of physical signal metrics, on the other hand, is complicated by the higher dimensionality (spatial and temporal) of the sign language signal as compared to a speech signal (solely temporal). Here, we review a variety of approaches to operationalizing sign language complexity based on linguistic and physical data, and identify the approaches that allow for high fidelity modeling of the data in the visual domain, while capturing linguistically-relevant features of the sign language signal.

Keywords: complexity; entropy; language processing; linguistic laws; sign language

1 Introduction

As language complexity, signal entropy, and information transfer in communicative signals have become a topic of intense interest to linguistics in the past decade, and have come to be viewed as a critical piece of the puzzle in the understanding of human language faculty, it is important to ensure that complexity metrics developed by language science are sufficiently robust as to be applicable to all types of human linguistic communication, including sign languages. Sign languages are communication systems that employ the visual modality, while being equivalent to spoken language in communicative efficacy and flexibility (Bellugi and Fischer 1972; Quer et al. 2021). Sign language signal simultaneously consists of specific handshapes, motion of the hands, arms, and non-manual components (positioning of eye brows and mouth shape, as well as movement of the head and body). Minimal changes in a sign's movement, location, or handshape can lead to changes in meaning; yet, sign languages differ in the inventory of allowable handshapes, preferred places of articulation, and in specific correlations between physical parameters of motion (range, speed, acceleration) and the linguistic features that those are linked to. The core difference between signing and speech is the

***Corresponding author: Evie A. Malaia**, Department of Communication Disorders, The University of Alabama, Tuscaloosa, AL, USA, E-mail: eamalaia@ua.edu. https://orcid.org/0000-0002-4700-0257
Joshua D. Borneman, Department of Speech, Language, and Hearing, Purdue University, West Lafayette, IN, USA
Emre Kurtoglu and Sevgi Z. Gurbuz, Department of Electrical and Computer Engineering, The University of Alabama, Tuscaloosa, AL, USA
Darrin Griffin, Department of Communication Studies, The University of Alabama, Tuscaloosa, AL, USA
Chris Crawford, Department of Computer Science, The University of Alabama, Tuscaloosa, AL, USA
Ali C. Gurbuz, Department of Electrical and Computer Engineering, Mississippi State University, Starkville, MS, USA

communication modality, as sign languages rely on visual-kinematic signal for information transfer. Since sign languages simultaneously use multiple articulators (i.e. both hands; non-manual markers such as eyebrow, eye, cheek, and mouth postures; as well as head and body position) to convey linguistically and emotionally rich messages, they are best construed as multi-channel systems.

Developing complexity metrics that can be applied to human languages across modalities is a difficult endeavor, as the term 'complexity' is used with different meanings across sub-fields of linguistics. As the main goal of operationalizing a variable is to allow comparative analysis to be rooted in empirical data (corpora or experiments), it means that metrics for operationalizing complexity need to be: 1) *modality-independent*: i.e. allowing for evaluation of either spoken or signed communication; 2) *quantifiable* in a way that can capture differences across languages with different structures, and across modalities; 3) *objective*, or theory- and measurement-independent: the results would be consistent across methods of data collection, and independent of theoretical assumptions about the underlying structure of data units (examples of objective metrics: frequency of item occurrence in corpora; temporal duration of the linguistic component).

In this review, we summarize approaches to quantification of the communicative signal in sign languages using linguistic (sign-based) and physical (signal-based) metrics. Both of these types of metrics are successfully used to characterize both speech and written language (Blumenthal-Dramé and Malaia 2019; Torre et al. 2019). We consider how modality- and dimensionality-related differences of sign language affect the applicability of existing complexity metrics to sign languages, and discuss development of new ones, with the goal of developing universally applicable complexity metrics for analysis across the full spectrum of human communication. First, we review the ways in which linguistic complexity has been conceptualized in sign language research, building on the structure of a sign language signal (Section 2). Then, we consider a variety of proposed signal-based complexity metrics for sign language, and apply dimensional analysis to assess what component of the multidimensional signing signal these the metrics capture (Section 3). We conclude that robust complexity metrics for visual communication take into account the linguistic structure of sign languages, while relying on high-fidelity spatio-temporal data.

2 Linguistic metrics of complexity in sign languages

Discussion about linguistic metrics of complexity for sign languages has to start with the description of the structure, or components of sign language. In this section, we review the building blocks of sign language, and then discuss linguistic metrics developed to measure complexity in the phonology and lexicon in sign languages.

2.1 The building blocks of sign language signals

Across a variety of sign language studies, the parameters of handshape, its location, and movement are universally employed as building blocks of sign language phonology, and allow for the creation of minimal pairs based on the change in one of these parameters (Quer et al. 2021). The skeleton of a sign, typically consisting of location-movement-location (L-M-L) is sequential; the movement, being at the core of a sign-syllable, constitutes the most information-dense portion of the sign, similar to the nucleus in spoken language syllables (Liddell 1984; Malaia and Wilbur 2020).

The gross component of articulator movement in a sign-syllable (primary, or path movement) typically consists of the change of location for the hand in signing space using elbow and shoulder joints. Movement in a number of signs also incorporates handshape change (secondary, or local movement), which consists of finger/wrist motion. Some signs simultaneously employ both primary and secondary movement. Additionally, between-sign motion (transitional motion) is necessary to move the hands from the final location of one sign to the starting location of another. The onset of the sign corresponds to the point when the sign-initial handshape of the dominant hand is formed, and the major movement begins; sign offset corresponds to the moment when the hand(s) begin movement away from sign-final location. The order of signs within a signed sentence depends on language-specific syntax (Quer et al. 2021), and message-specific information structure (Krebs et al. 2020). Additional information is often conveyed by non-manual components (facial expressions, position of head and body, etc.), which can contribute to semantic, syntactic, prosodic, or information-structural components of the message in sign language.

2.2 Complexity metrics in sign language phonology

Within the field of sign language phonology, complexity currently refers to two distinct concepts: that of articulatory complexity (Battison 1978; Eccarius and Brentari 2007), or that of phonological complexity (Brentari 2019). Articulatory complexity is provisionally defined as the assumed effort involved in production of a phonological form (primarily handshape), which suggests a potential overlap with quantitative metrics from the field of kinesiology.

In one of the earliest works on sign phonology dealing with complexity, Battison (1978) suggested a hierarchy of articulatory complexity for individual signs based on the amount of (presumed) effort for their articulation and perception. One-handed signs were assumed to be the simplest. Among two-handed signs, those with two active hands performing the same movement were the easiest; the movement could be either synchronized or alternating. The next level of articulatory complexity encompassed two-handed signs with one active and one passive hand, both with the same handshape; finally, two-handed signs with an active and a passive hand differing in handshapes were assumed to be the most complex in terms of articulation. It is important to note that Battison relied on the ceramic model of phonology (Stokoe 1960), in which the entire handshape was considered the minimal contrastive unit, while later models of sign language phonology singled out individual features of the handshape (cf. Brentari 1998).

Battison's model allowed for specification of universal articulatory constraints, which hold across known sign languages. For example, the symmetry condition specifies that if both hands move, then they have to move symmetrically or identically (i.e. in the same locations and with the same handshape and form of movement). The dominance condition specifies that if the two hands have different handshapes, then one is passive (with a restricted list of possible handshapes), and the other (dominant) articulates the movement. While the constraints appear to be empirically correct, the basis for these is still unknown. Assumed driving forces for linguistic universals tend to be the ease of production, perception, and/or redundancy of information encoding; however, lack of quantifiable metrics for articulatory complexity makes it difficult to empirically test falsifiable hypotheses in this domain (but see Bradley et al. 2022).

2.3 Complexity metrics in sign language lexicon

Historically, metrics designed to describe the density of information in speech signals have been assessed based on written representations. Attempts to quantitatively characterize relationships between linguistic levels in spoken languages (e.g. phoneme vs. lexicon, or word vs. text) resulted in formulation of the earliest linguistic laws that characterized language as a complex system: Zipf's law (Zipf 2013), Brevity law (or Zipf's rank-correlation law; Zipf 1949, 1935), and Menzerath-Altmann law (Altmann 1980; Altmann and Schwibbe 1989). Zipf's law, one of the more widely known linguistic observations, describes the frequency of word occurrence in relation to the word's frequency rank as:

$$f(r) \sim r^{-\alpha} \tag{1}$$

In other words, the frequency of the word's occurrence in the corpora falls off logarithmically with the decrease in the word's rank order. An extension of Zipf's law, the so-called brevity law (or Zipf's law of abbreviation) summarizes the observation that the more frequently used words tend to be shorter (Zipf 1935). The driving force behind Zipf's laws is best understood in terms of economy in the production of language: for high-frequency lexical items to be shorter than low-frequency items minimizes the time for information transfer during language production (Bybee 2017, 2003). In sign language research, Zipf's law has been confirmed on a Swedish Sign Language (SSL) corpus. Börstell et al. (2016) noted that Swedish Sign Language conforms to Zipf's law in demonstrating log-distribution of sign frequencies (i.e. sign frequency in the corpus falling off logarithmically with sign rank), with the most frequent lexical items being function-type signs (e.g. pointing). Brevity law is upheld as well: the most frequent items in SSL corpus are shorter in duration (Börstell et al. 2016).

Another modality-independent linguistic law—Menzerath-Altmann law—operates at the interfaces of subsystems of language. Menzerath-Altmann law relates the size of a language construct to the size of its constituents: the longer the construct (such as a word), the shorter the units of which it is constructed (e.g. phonemes) tend to be (Torre et al. 2019). Menzerath-Altmann law has been tested in physical units for Czech Sign Language (Andres et al. 2020). Duration of individual signs (in seconds), clauses (in signs), and sentences (in clauses) in Czech Sign Language were shown to be reciprocally related (Andres et al. 2020).

Consistently glossed corpora can be used for analysis of lexical and morphological complexity in sign languages. Such corpora exist for a number of sign languages, including American Sign Language (Morford and MacFarlane 2003); New Zealand Sign Language (McKee and Kennedy 2006); Australian Sign Language (Johnston 2012); British Sign Language (Fenlon et al. 2014), and Swedish Sign Language (Börstell et al. 2016). As objective metrics at the level of the lexicon, such as word duration and frequency, are modality-independent, complexity analysis at this level is consistent between speech and sign languages.

3 Signal-based complexity in sign language

Having reviewed linguistic measures of complexity, we now turn to the question of physical signal complexity in communication. It is important to note that physical and linguistic complexity, while connected, are not equivalent (in either sign languages or speech). Analysis of a physical signal necessarily proceeds at all levels of linguistic structure simultaneously; for example, a spectrogram of a speech utterance contains information at the phonetic, lexical, syntactic, and prosodic levels simultaneously. The same is true of sign languages; the physical sign language signal contains information from all levels of linguistic structure. However, due to the fact that sign language articulation unfolds both in time and in space, the question of *fidelity*, or accuracy of representing the raw sign language signal, becomes important for analysis of the applicability of complexity metrics to sign language data.

3.1 Entropy in sign language signal

The field of complex systems uses entropy analysis to describe the state of a system that involves a variety of spatial and temporal components, or to quantify the difference between states. Mathematically, entropy describes the potential information transfer by the signal (Shannon 1948). Thus, entropy measures of a signal describe that signal's predictability over time. The continuous signal of human motion can be viewed as a complex system, since it is based on interaction of neural and motor systems at a range of temporal and spatial scales (Buzsáki and Mizuseki 2014).

The techniques for estimating the entropy of a signal use mathematical representations of the signal as time series; thus, multi-scale recordings of a physical signal (such as motion capture data) lend themselves well to this analysis. For example, time-series motion capture data (i.e. displacement in time) for multiple articulators in ASL have indicated that information throughput is highest on the dominant hand, and independently significant for both the non-dominant hand and the head, lending further support to the multi-channel model of sign language communication (Malaia and Wilbur 2019). In contrast to linear recordings, multidimensional recordings of biological systems (such as video recordings of human motion) present a computational challenge for the analyses of complexity. Such analyses are limited by the fidelity of the recorded signal, that is, the accuracy and dimensionality with which the original signal can be recorded, as well as potential dimensionality reduction and fidelity loss during further analysis. Each step in the measurement and processing of the original signal acts as a filtering function, potentially reducing the fidelity of resulting data. Therefore, results of language signal complexity analysis depend on the fidelity of both the measurement tools and the analysis methods.

Spoken language is a one-dimensional time series signal, which is a function of temporal frequency [f_t], and the results of any recording and analysis are a reduced-fidelity signal [$S_1(f_t)$], which results from the original signal [$S_0(f_t)$], after it has been filtered through transfer functions associated with the electronics/ sampling [$T_e(f_t)$], and the processing method [$T_p(f_t)$]. All $T(f_t) \leq 1$, since $T(f_t) = 1$ would be a perfect unfiltered signal at f_t. Therefore, we then have that $S_0(f_t) \cdot T_e(f_t) \cdot T_p(f_t) = S_1(f_t)$.

In a continuous (1D) auditory signal, information density is contained entirely within the temporal component, and therefore requires high temporal sampling and analysis frequencies, as raw acoustic recordings may contain 20 k or more samples per second. The signal amplitude (i.e. volume) at a single point in time contains no information.

Similar considerations of information density are relevant for recording and processing signed languages. Signed languages, as a 3-dimensional spatial signal also varying in time, $S_0(f_x, f_y, f_z, f_t)$, are filtered in both spatial frequencies, f_x, f_y, f_z, and temporal frequencies, f_t (thus, $S_0 \cdot T_e = S_1$, where $S_1(f_x, f_y, f_z, f_t)$ is the recorded signal, and $S_1 \cdot T_p = S_2$, where $S_2(f_x, f_y, f_z, f_t)$ is the signal after data processing).

Sign languages use comparatively low temporal frequencies as compared to spoken languages. However, when compared to speech, signing also transfers information in additional spatial dimensions (3D, as opposed to 2D). For information transfer and complexity analysis, the information content of the spatial dimension of the sign language must be included.

2D video recordings are, in essence, a spatial frequency filter for sign language production, such that, $T_e(f_z) = 0$, and $T_e(f_x, f_y)$ are limited by the camera distance and resolution (the camera, ideally, being placed in a position such that the camera bandpass transfers all relevant articulators within the signing space), and $T_e(f_t)$ is limited by the frame rate of the recording. Thus, $S_{1,video}$ is a function of two spatial dimensions and time (f_x, f_y, f_t).

The characterization of language complexity based on the physical signal begins with parametrization of typical and potential physical parameters of the communicative signal. The physical space in which sign production occurs (termed signing space) is relatively small, falling within an 80 cm vertical and 60 cm horizontal space directly in front of the signer (Bosworth et al. 2019). In terms of physical characterization, this means that specific spatial parameters of the signer's image (such as spatial frequencies and entropy of the signal) might carry the most information for distinguishing sign language features.

In the earliest attempt at developing quantifiable metrics for a sign language signal, Bosworth et al. (2006) evaluated 100-unit sets of static images of signers signing in ASL, natural scene images, and face images, comparing mean amplitude and entropy of the amplitude across the image set as a function of spatial frequency (the unit employed to express spatial frequency is the number of cycles that fall within one degree of visual angle, the cycle being a resolvable high-contrast feature) and orientation. The differences in image statistics for ASL signs included more amplitude for vertical than for horizontal contours (indicating vertically-oriented arm articulators as primary sources of information), and entropy analyses for spatial frequency, which showed a peak in entropy at approximately 0.75 cycles/cm (assuming a viewing distance of five feet, this would convert to 1.9 cycles/degree). The result can be interpreted to mean that spatial frequencies around this peak contain the greatest amount of information (and would, for example, result in minimal cognitive load) for distinguishing an ASL sign from such an image. This result found confirmation in a previously-published work (Riedl and Sperling 1988), which presented native signers with spatial-frequency-filtered sign images (filtered at peaks of 0.05, 0.20, 0.35, and 0.75 cycles/cm to determine which spatial frequencies were used for sign recognition), and determined that the signs filtered at 0.75 cyc/cm spatial frequency elicited maximal behavioral accuracy for sign discrimination. The valuable contribution of this work consisted in a quantified description of the changes in the static parameters of a visual signal; however, as sign language signal is highly reliant on motion, higher-dimensional analyses for continuous signal recording, such as video, were necessary.

The proposed complexity metrics for analysis of multi-dimensional sign language signal vary in the fidelity of spatiotemporal signal quantification. For example, Brookshire et al. proposed an Instantaneous Visual Change (IVC) metric (Brookshire et al. 2017), which attempts to quantify sign language video data by measuring the relative pixel value change from frame to frame. The proposed analysis equation in Brookshire et al. (2017) is as follows (with the addition of notation here for the two spatial dimensions (x, y), which are conflated in the original source): $IVC(t) = \sum_{x,y} [S_1'(x, y, t) - S_1'(x, y, t-1)]^2$, where x and y are spatial dimensions in the 2D video, and t is time.

This way of calculating the metric acts as a spatial frequency filter. If instead of a function in linear space and time, $S_1'(x, y, t)$, we instead examine it as a function of spatial and temporal frequencies, $S_1(f_x, f_y, f_t)$, then the two spatial dimensions are reduced to a single scalar value versus time: $S_1(f_x, f_y, f_t) \cdot T_{IVC}(f_x, f_y, f_t) = IVC(f_t)$. Notably, the properties of such a filter can change frame by frame, depending on the colors and contrast within the scene, rather than the complexity of spatial and temporal components of sign language signal, where T_{IVC} acts as a low-pass spatial frequency filter. For instance, this method will be sensitive to large area changes in the image. The motion of a signer's arm will affect a large number of pixels, and will therefore result in a large IVC value, whereas complex finger motion will affect a small number of pixels and result in a low IVC value. Thus, in $IVC(f_t)$, all spatial information is both low-pass filtered, and then also reduced to a single scalar value versus time. The resulting metric might contain a small portion of the information inherent in the sign language signal; however, given the spatial nature of sign language, the majority of the information contained in sign language video recording is lost. Further, given low video frame rates, the processed signal would have a bandwidth of 30 samples per second: thus, use of IVC for sign language analysis is analogous to using a 30 Hz volume envelope of a spoken language signal for analysis of speech.

A higher-fidelity method of 2D analysis for sign language signal utilizes optical flow (OF) tracking (Borneman et al. 2018). This method tracks the velocity magnitude of each object (typically based on edge contrast values) in pixels per frame, $S_1'(vx, vy)$ where vx and vy are the velocity of pixels in the x and y directions respectively. This approach preserves the dimensionality and data rate of the input video, where in frequency space: $S_1(f_{vx}, f_{vy}, f_t) \cdot T_{OF}(f_{vx}, f_{vy}, f_t) = OF(f_{vx}, f_{vy}, f_t)$, where f_{vx}, f_{vy} are the frequency profiles of the velocities in two dimensions, x and y, respectively. Although optical flow analysis converts each frame to a velocity profile, it does not filter the spatial content of dimensions, as the resulting signal contains velocity per pixel versus time. Optical flow information may be further dimensionally reduced without reduction in data density by parameterizing the two spatial dimension velocities into a single dimension. For instance, 640×480 video will contain 307,200 samples per frame, and may be parameterized such that $OF(f_v, f_t)$, where f_v contains 307,200 samples per frame. Language complexity analysis from this point may be done on the full signal bandwidth, although in some cases further reduction through velocity binning may simplify the analysis. Use of optical flow, with spatio-temporal downsampling through binning (e.g. looking at spatial dimension velocity profiles from 0 to 0.4 pixels per frame over 200 velocity bins, cf. Borneman et al. 2018), yields a bandwidth of 6,000 samples per second, as compared to 30 samples per second of IVC (Brookshire et al. 2017). From the standpoint of a physical signal, the optical flow metric, which identifies pixel value change from frame to frame, track both the primary motion and the secondary (handshape change) motion in signs, which then contribute to calculation of signal entropy (Borneman et al. 2021; Malaia et al. 2021a, 2021b). However, as only a 2D projection of the 3D handshape is available in a video, the tracking of handshape change does not have full fidelity to the original signal.

A recording modality which allows for 3D recording of sign language with a lower loss of fidelity is frequency-modulated continuous wave (FMCW) radar. In the case of radar recordings, radar signal processing algorithms may be applied to extract range-Doppler (RD) maps (2D images of range vs. Doppler frequency) or micro-Doppler signature (Doppler frequency vs. time). In this case, $T_e(f_x, f_y)$ are conflated into Doppler velocity vector, and $T_e(f_t)$ is limited by the chirp rate (transmit frequency swept per pulse). Therefore, $S_{1,doppler}$ is a function of motion velocity, slant range, and time (f_v, f_z, f_t).

The optical flow of the range-Doppler maps, as a function of time, can be computed to derive representations of radial velocity and acceleration, as shown in Figure 1. Direct computation of optical flow, however,

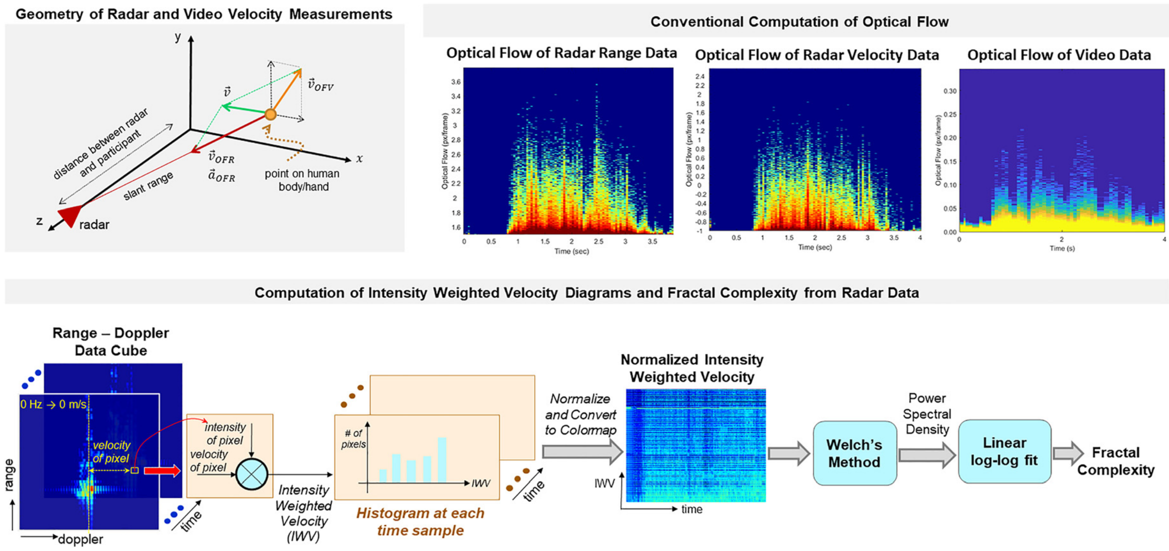

Figure 1: Radar-based computation of optical flow and fractal complexity metrics for sign language.

does not take into account the area size of the moving part. Since the power of the received radar return is related to area, a more equivalent representation is Intensity Weighted Velocity Diagrams (IWVD) in which the velocity corresponding to each pixel in the RD map is weighted according to its intensity and binned. In this way, we can ensure the information of both the velocity and the area of moving objects in the same speed range, which correspond to OF magnitude and the intensity in video OF diagrams, respectively, are preserved (cf. Gurbuz et al. 2020a, 2020c).

One entropy-based metric that has demonstrable consistency for evaluating information transfer in sign language signal (from either radar or video) is *fractal complexity*. To compute it, the power-spectral density for each velocity is first calculated using a Fourier transform of the velocity spectrum versus time resulting in a matrix $M(f_t)$, a function of temporal frequency $[f_t]$. This matrix can then be characterized by the general fractal complexity function $M(f_t) = \frac{\alpha}{f^{\beta}}$, where α and β are the spectral density amplitude and the fractal complexity fitting parameters, respectively, and f_t is temporal frequency. Fractal complexity, β, of the information transfer is found using a linear fit to the log-log function, $ln(M) = \beta \cdot ln(f_t) + \alpha$.

Fractal complexity-based evaluation of video data comparing American Sign Language with everyday activities (Borneman et al. 2018) demonstrated that sign language signal evidenced more complex underlying structure as compared to everyday motion. Another study, which used microdoppler radar measurements from participants enacting daily activities, and producing ASL signs, indicated that the average value of the complexity parameter for ASL was 201.8, with the equivalent metric for everyday activities averaging at 186.8 (Gurbuz et al. 2020b). Converging results between fractal complexity analysis of optical flow for ASL and everyday motion/gesture in two different recording modalities, video and radar, suggest that this metric is sensitive to spatio-temporal parameters of the information-bearing signal with high dimensionality, while being robust to some dimensionality reduction.

3.2 Motion analysis in sign languages: intersection between physical and language-based metrics of information transfer

Motion is an important component of sign language signal due to its centrality to the structure of signs. At the same time, it is a quantifiable physical measure of the sign language signal. Thus, a number of quantitative metrics of motion in sign languages have been developed, which connect linguistic and physical parameters of signing in a variety of sign languages. Optical flow analysis, discussed above, measures the entropy and the information content of the overall communication system (i.e. sign language). However, evaluation of motion profiles for individual linguistic units (signs, sign morphemes, sentences, etc.) allows one to make connections between the physical signal and the specific linguistic units under investigation. The research on this relationship is what we consider in this section.

Physical parameters of motion in sign languages are determined by rapid changes of position and handshape (primary and secondary motion) within the signing space. The time-varying movement primarily occurs in the range of 1/2 to 1 m/s, although velocities of up to 4 m/s have been recorded as part of natural sign language discourse (Bosworth et al. 2019; Malaia et al. 2008). One representation of dominant hand velocity in three dimensions from a motion capture recording is given in Figure 2. Here, the velocity of the dominant hand throughout a short connected narrative as recorded by a motion capture system with a 200 Hz frequency rate (i.e. 200 frames per second) is projected into the three orthogonal planes – *X*, *Y*, and *Z*. Note story-initial and story-final intervals of high-speed motion along the vertical axis (*Z*), which are transitional motions spanning the largest distances, from neutral (non-signing) position (hands-down, i.e. away from signing space), towards signing space in front of the signer; and back to hands-down position at the end of the narrative. Figure 2 shows that the articulatory signal in sign language is highly dynamic. What are the connections, then, between the physical parameters of motion, and the linguistic features of sign language?

The overall amplitude of motion, as well as duration of individual signs, which are the simplest visually available parameters, have been associated with markers of stress in a number of sign languages (Wilbur 1999). Tkachman et al. (2019) proposed a metric termed *visible amplitude*, calculated from optical flow in signing videos as the root-mean-square of a motion velocity time-series (Tkachman et al. 2019). This attempt to quantify phonological prominence for sign languages (i.e. equivalent to loudness in speech) did not take into account the visual filtering parameters of the source video recording, such as the distance to the signer in the video (but see Borneman et al. 2018 for a scaling method that accounts for both the distance to the participant, and the difference in individual articulator size).

Studies of articulator dynamics for sign languages, where objective motion capture data was collected, such as ASL (Bosworth et al. 2019; Malaia 2017), Croatian Sign Language (Malaia et al. 2013; Milković 2011), and Austrian Sign Language (Krebs et al. 2021) demonstrate a variety of strategies whereby sign language recruits kinematic features into linguistic systems. Among the investigated features, articulator velocity has been associated with adjective scale (Wilbur et al. 2012), sentential stress in ASL (Malaia and Wilbur 2012), and perfective/imperfective grammatical distinction in Croatian Sign Language (*Hrvatski znakovni jezik*, or HZJ; cf. Malaia et al. 2013).

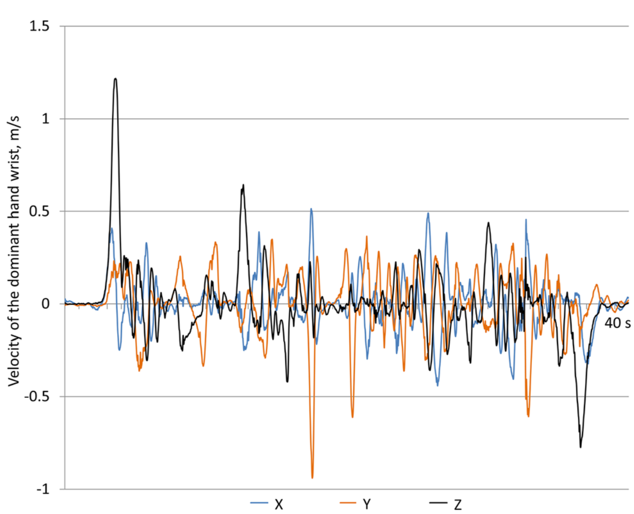

Figure 2: Velocity of the dominant hand during signing of a short narrative, along orthogonal axes: anterior-posterior (*X*), medio-lateral (*Y*), and vertical (*Z*).

Investigations based on motion capture data in HZJ have identified a grammatical strategy of recruiting speed of motion into a linguistic system: HZJ signers systematically use high deceleration of the dominant hand to produce differential kinematic profiles for perfective, as opposed to imperfective verb signs during regular sign production (Malaia et al. 2008, 2013). As aspect is conflated with telicity in HZJ, and spoken Slavic languages which it interacts with, the same physical feature that is observed corresponds to the semantic field of telicity in ASL, is used to denote grammatical aspect in HZJ.

In Austrian Sign Language (*Österreichische Gebärdensprache*, or ÖGS), physical parameters also differentiate between verb types. However, the list of those parameters differs. While in HZJ, telic and atelic signs can be differentiated based on the peak velocity of hand motion within the sign, ASL and ÖGS appear to rely on acceleration/deceleration parameters for differentiation of semantic features. This difference can be traced back to the differences in morphosyntactic structures of these languages. In HZJ, the grammaticalized – and frequent – event structure-marking parameter (peak velocity) had to be robust to prosodic effects, such as phrase-final lengthening (Malaia et al. 2013). For ÖGS and ASL, on the other hand, velocity is not the only physical parameter that distinguishes verbs with differing event structure: there are handshape and place of articulation differences, as well as motion trajectory. In ASL, instead, regularized variations in peak velocity are indicative of stress patterns (Wilbur 1999), meaning that while velocity is significant for the physical manifestation of linguistic features in ASL, it also relates to the level of prosody, rather than grammar. Since recent (Strickland et al. 2015) suggests that association of particular kinematic markers with specific semantic fields or grammatical features may have its roots in more general cognitive heuristics (such as event segmentation), the question of the inventory of dynamic features, and their linguistic roles across different sign languages, is a rapidly developing domain of inquiry for sign language research.

4 Conclusion

In this work, we have considered the applicability of multiple complexity metrics to sign languages, as dictated by the spatio-temporal parameters of a sign language communicative signal. After considering existing studies using linguistic (sign-based) and physical (signal-based) metrics, we show that sign language complexity is comparable to that reported for spoken languages, which have been characterized as complex systems that demonstrate power law behaviors. However, development and application of complexity metrics for sign language data requires careful consideration of the applicability of such metrics to both the linguistic structure of sign language, and to spatio-temporal nature of sign language communication, to preserve signal fidelity.

Further development of sensitive, yet robust metrics of complexity for sign language analysis (both as a physical communicative signal, and in terms of linguistic subsystems) is a critical step toward more computationally-intensive analyses of sign language for sign recognition and translation, as well as for applied fields of second language teaching and assessments of language proficiency. While lack of easily-processed symbolic representation (written form) has originally hampered research on sign language complexity, the development of cheap high-fidelity methods of recording, and corpora development for metric testing, are leading to rapid progress in the area.

Acknowledgment: Preparation of this article was partially funded by the National Science Foundation under grants #1932547, #1931861 and #1734938.

References

Altmann, Gabriel. 1980. Prolegomena to Menzerath's law. *Glottometrika* 2(2). 1–10.

Altmann, Gabriel & Michael H. Schwibbe. 1989. *Das menzerathsche gesetz in informationsverarbeitenden systemen*. Hildesheim: Georg Olms Verlag.

Andres, Jan, Jiři Langer & Vladimır Matlach. 2020. Fractal–based analysis of sign language. *Communications in Nonlinear Science and Numerical Simulation* 84. 105214.
Battison, Robbin. 1978. *Lexical borrowing in American sign language*. Silver Spring, MD: Linstok Press.
Bellugi, Ursula & Susan Fischer. 1972. A comparison of sign language and spoken language. *Cognition* 1(2–3). 173–200.
Blumenthal-Dramé, Alice & Evie Malaia. 2019. Shared neural and cognitive mechanisms in action and language: The multiscale information transfer framework. *Wiley Interdisciplinary Reviews: Cognitive Science* 10(2). e1484.
Borneman, Joshua D., Evie Malaia & Ronnie B. Wilbur. 2018. Motion characterization using optical flow and fractal complexity. *Journal of Electronic Imaging* 27(5). 051229.
Borneman, Sean, Julia Krebs, Ronnie B. Wilbur & Evie Malaia. 2021. Application of machine learning to signal entrainment identifies predictive processing in sign language. *Proceedings of the Annual Meeting of the Cognitive Science Society* 43(43).
Börstell, Carl, Thomas Hörberg & Robert Östling. 2016. Distribution and duration of signs and parts of speech in Swedish sign language. *Sign Language & Linguistics* 19(2). 143–196.
Bosworth, Rain G., Marian S. Bartlett & Karen R. Dobkins. 2006. Image statistics of American Sign Language: Comparison with faces and natural scenes. *JOSA A* 23(9). 2085–2096.
Bosworth, Rain G., Charles E. Wright & Karen R. Dobkins. 2019. Analysis of the visual spatiotemporal properties of American sign language. *Vision Research* 164. 34–43.
Bradley, Chuck, Evie A. Malaia, Jeffrey Mark Siskind & Ronnie B. Wilbur. 2022. Visual form of ASL verb signs predicts non-signer judgment of transitivity. *PLoS One* 17(2). e0262098.
Brentari, Diane. 1998. *A prosodic model of sign language phonology*. Cambridge, MA: MIT Press.
Brentari, Diane. 2019. *Sign language phonology*. Cambridge, UK: Cambridge University Press.
Brookshire, Geoffrey, Jenny Lu, Howard C. Nusbaum, Susan Goldin-Meadow & Daniel Casasanto. 2017. Visual cortex entrains to sign language. *Proceedings of the National Academy of Sciences* 114(24). 6352–6357.
Buzsáki, György & Kenji Mizuseki. 2014. The log-dynamic brain: How skewed distributions affect network operations. *Nature Reviews Neuroscience* 15(4). 264–278.
Bybee, Joan. 2003. *Phonology and language use* Cambridge, UK: Cambridge University Press.
Bybee, Joan. 2017. Mechanisms of change in grammaticization: The role of frequency. *The handbook of historical linguistics*, 602–623. Hoboken, NJ: Blackwell Publishing.
Eccarius, Petra & Diane Brentari. 2007. Symmetry and dominance: A cross-linguistic study of signs and classifier constructions. *Lingua* 117(7). 1169–1201.
Fenlon, Jordan, Adam Schembri, Ramas Rentelis, David Vinson & Kearsy Cormier. 2014. Using conversational data to determine lexical frequency in British sign language: The influence of text type. *Lingua* 143. 187–202.
Gurbuz, Sevgi, Ali Gurbuz, Evie Malaia, Darrin Griffin, Chris Crawford, Emre Kurtoglu, Mahbubur Rahman, Ridvan Aksu & Robiulhossain Mdrafi. 2020a. ASL recognition based on kinematics derived from a multi-frequency RF sensor network. In *2020 IEEE SENSORS conference*, 1–4.
Gurbuz, Sevgi, Ali Gurbuz, Evie Malaia, Darrin Griffin, Chris Crawford, Mahbubur Rahman, Ridvan Aksu, Emre Kurtoglu, Robiulhossain Mdrafi, Ajaymehul Anbuselvam, Trevor Macks & Engin Ozcelik. 2020b. A linguistic perspective on radar micro-doppler analysis of American sign language. In *2020 IEEE international radar conference (RADAR)*, 232–237.
Gurbuz, Sevgi, Ali Gurbuz, Evie Malaia, Darrin Griffin, Chris Crawford, Mohammad Rahman, Emre Kurtoglu, Ridvan Aksu, Trevor Macks & Robiulhossain Mdrafi. 2020c. American sign language recognition using RF sensing. *IEEE Sensors Journal* 21(3). 3763–3775.
Johnston, Trevor. 2012. Lexical frequency in sign languages. *Journal of Deaf Studies and Deaf Education* 17(2). 163–193.
Krebs, Julia, Evie Malaia, Ronnie B. Wilbur & Dietmar Roehm. 2020. Interaction between topic marking and subject preference strategy in sign language processing. *Language, Cognition and Neuroscience* 35(4). 466–484.
Krebs, Julia, Gerda Strutzenberger, Hermann Schwameder, Ronnie Wilbur, Evie Malaia & Dietmar Roehm. 2021. Event visibility in sign language motion: Evidence from Austrian sign language (ÖGS). In *Proceedings of the annual meeting of the Cognitive Science Society*, vol. 43.
Liddell, Scott K. 1984. THINK and BELIEVE: Sequentiality in American sign language. *Language* 60. 372–399.
Malaia, Evguenia, John Borneman & Ronnie B. Wilbur. 2008. Analysis of ASL motion capture data towards identification of verb type. In *Semantics in text processing. STEP 2008 conference proceedings*, 155–164.
Malaia, Evie. 2017. Current and future methodologies for quantitative analysis of information transfer in sign language and gesture data. *Behavioral and Brain Sciences* 40. E63.
Malaia, Evie, Linda Ford, Sean Borneman, Julia Krebs & Brendan Ames. 2021a. Salience of low-frequency entrainment to visual signal for classification points to predictive processing in sign language. In proceedings of 30th annual computational neuroscience meeting: CNS*2021. *Journal of Computational Neuroscience* 49(1 Suppl). 3–208.
Malaia, Evie & Ronnie B. Wilbur. 2012. Kinematic signatures of telic and atelic events in ASL predicates. *Language and Speech* 55(3). 407–421.
Malaia, Evie A. & Ronnie B. Wilbur & A. Malaia. 2020. Syllable as a unit of information transfer in linguistic communication: The entropy syllable parsing model. *WIREs Cognitive Science* 11(1). e1518.

Malaia, Evie, Ronnie B. Wilbur & Marina Milković. 2013. Kinematic parameters of signed verbs. *Journal of Speech, Language, and Hearing Research* 56. 1677–1688.
Malaia, Evie & Ronnie Wilbur. 2019. Visual and linguistic components of short-term memory: Generalized neural model (GNM) for spoken and sign languages. *Cortex* 112. 69–79.
Malaia, Evie, Sean Borneman, Julia Krebs & Ronnie Wilbur. 2021b. Low-frequency entrainment to visual motion underlies sign language comprehension. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 29. 2456–2463.
McKee, David & Graeme Kennedy. 2006. The distribution of signs in New Zealand sign language. *Sign Language Studies* 6(4). 372–390.
Milković, Marina. 2011. *Verb classes in Croatian sign language (HZJ): Syntactic and semantic properties*. Croatia: University of Zagreb PhD dissertation.
Morford, Jill P. & James MacFarlane. 2003. Frequency characteristics of American sign language. *Sign Language Studies* 3. 213–225.
Quer, Josep, Roland Pfau & Annika Herrmann. 2021. *The Routledge handbook of theoretical and experimental sign language research*. Milton Park, UK: Routledge.
Riedl, Thomas R. & George Sperling. 1988. Spatial-frequency bands in complex visual stimuli: American sign language. *JOSA A* 5(4). 606–616.
Shannon, Claude Elwood. 1948. A mathematical theory of communication. *The Bell System Technical Journal* 27(3). 379–423.
Stokoe, William. 1960. Sign language structure, an outline of the visual communications systems of American deaf. *Studies in Linguistics Occasional Paper* 8. 3–37.
Strickland, Brent, Carlo Geraci, Emmanuel Chemla, Philippe Schlenker, Meltem Kelepir & Roland Pfau. 2015. Event representations constrain the structure of language: Sign language as a window into universally accessible linguistic biases. *Proceedings of the National Academy of Sciences* 112(19). 5968–5973.
Tkachman, Oksana, Kathleen Currie Hall, Robert Fuhrman & Yurika Aonuki. 2019. Visible amplitude: Towards quantifying prominence in sign language. *Journal of Phonetics* 77. 100935.
Torre, Iván G., Bartolo Luque, Lucas Lacasa, Christopher T. Kello & Antoni Hernández-Fernández. 2019. On the physical origin of linguistic laws and lognormality in speech. *Royal Society Open Science* 6(8). 191023.
Wilbur, Ronnie B. 1999. Stress in ASL: Empirical evidence and linguistic issues. *Language and Speech* 42(2–3). 229–250.
Wilbur, Ronnie B., Evie Malaia & Robin A. Shay. 2012. Degree modification and intensification in American sign language adjectives. In *Proceedings of the 18th Amsterdam colloquim conference on logic, language and meaning*, 92–101.
Zipf, George Kingsley. 1935. *The psycho-biology of language: An introduction to dynamic philology*. Boston: Houghton Mifflin.
Zipf, George Kingsley. 1949. *Human behaviour and the principle of least-effort*, vol. 24. Reading: Addison-Wesley.
Zipf, George Kingsley. 2013. *Selected studies of the principle of relative frequency in language*. Cambridge, MA: Harvard University Press.