



Sato, Asha, Kirby, Simon and Flaherty, Molly. 2022. Language emergence can take multiple paths: Using motion capture to track axis use in Nicaraguan Sign Language. *Glossa: a journal of general linguistics* 7(1). pp. 1–17. DOI: <https://doi.org/10.16995/glossa.6151>



Open Library of Humanities

Language emergence can take multiple paths: Using motion capture to track axis use in Nicaraguan Sign Language

Asha Sato, Centre for Language Evolution University of Edinburgh, UK, Asha.Sato@ed.ac.uk

Simon Kirby, Centre for Language Evolution University of Edinburgh, UK, Simon.Kirby@ed.ac.uk

Molly Flaherty, Department of Psychology Davidson College, US, MoFlaherty@Davidson.edu

Research on emerging sign languages suggests that younger sign languages may make greater use of the z-axis, moving outwards from the body, than more established sign languages when describing the relationships between participants and events (Padden et al. 2010). This has been suggested to reflect a transition from iconicity rooted in the body (Meir et al. 2007) towards a more abstract schematic iconicity. We present the results of an experimental investigation into the use of axis by signers of Nicaraguan Sign Language (NSL). We analysed 1074 verb tokens elicited from NSL signers who entered the signing community at different points in time between 1974 and 2003. We used depth and motion tracking technology to quantify the position of signers' wrists over time, allowing us to build an automated and continuous measure of axis use. We also consider axis use from two perspectives: a camera-centric perspective and a signer-centric perspective. In contrast to earlier work, we do not observe a trend towards increasing use of the x-axis. Instead we find that signers appear to have an overall preference for the z-axis. However, this preference is only observed from the camera-centric perspective. When measured relative to the body, signers appear to be making approximately equal use of both axes, suggesting the preference for the z-axis is largely driven by signers moving their bodies (and not just their hands) along the z-axis. We argue from this finding that language emergence patterns are not necessarily universal and that use of the x-axis may not be a prerequisite for the establishment of a spatial grammar.

Glossa: a journal of general linguistics is a peer-reviewed open access journal published by the Open Library of Humanities. © 2022 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

OPEN ACCESS



1 Introduction

Cross-linguistically, sign languages make systematic, iconic use of space to encode relationships between participants and events in utterances, and track reference through discourse. One common way in which sign languages do this is through verb agreement or co-reference systems which are sometimes suggested to be akin to inflectional agreement systems in spoken language (Mathur & Rathmann 2012).¹

These systems are characterised by spatial modification of the citation form of a verb to indicate person and number agreement (Padden 1983). For example, the British Sign Language (BSL) translation equivalents of “I ask you” and “You ask me” differ in the direction of the path of the verb ASK, in the first instance starting at a location close to the body of the signer and moving away, and in the second instance starting at a distance and moving towards the signer, whilst the movement between two third person referents e.g. “She asks him” might have a path that moves side to side between two locations associated with those referents (Sutton-Spence & Woll 1999).

This kind of systematic, grammaticalised way of using of space is not necessarily present from the beginning in emerging sign languages, but is suggested to develop over time: In a study of two emerging sign languages, Israeli Sign Language (ISL) and Al-Sayyid Bedouin Sign Language (ABSL), Padden et al. (2010) report that signers of both languages make less consistent use of space to encode the relationships between participants than more established sign languages like American Sign Language (ASL). Specifically, they note that signers of ABSL and ISL show a preference for encoding certain events on the z-axis (moving outward from the body) (65% for ABSL and 54% for ISL) rather than spatially modulating them on the x-axis (horizontal in front of the body) (25% and 27% respectively), or what they refer to as z + x axis (diagonal). They additionally note that younger signers of both languages show this preference less strongly, producing more spatially modulated forms and therefore making use of the x-axis to a greater extent than older signers. This preference is taken to diverge from the norm in more established sign languages, where spatial modulations of the citation form of this type of verb are, as mentioned above, often interpreted as marking person agreement (Padden 1983). In spatial agreement systems, grammatical first person is generally associated with locations on the signer’s body (a pattern referred to as *body as subject* (Meir et al. 2007)), second person is associated with a location in the actual direction of the addressee (likely to be, but not necessarily, directly in front of the signer, along the z-axis) and third person is associated with some other location in space, either pointing towards the referent if physically present, or towards a referential or R-locus (Lillo-Martin & Klima 1990) which is usually first established lexically, and can then be referred back to later in the discourse. On the basis of their observations of ISL and ABSL, Padden et. al. suggest that the increasing use of the x-axis is indicative of the emergence of a spatial

¹ It is controversial whether these co-reference systems constitute agreement, we address this in more detail below.

grammar, i.e. verbal agreement, in these young sign languages, and that we might expect to see a comparable progression from predominant use of the z-axis to more use of the x-axis in other emerging sign languages.

Existing work on Nicaraguan Sign Language (NSL), a young sign language which began to emerge in the late 70s with the establishment of a special education school in Managua (Polich 2005) is suggestive of a similar pattern being present in NSL. Senghas & Coppola (2001) report that younger signers produce marginally more spatial modulations (defined in their work as signs produced in non neutral space) than older signers. In a study of the emergence of argument structure and spatial co-reference in NSL, Flaherty (2014) finds a similar pattern of axis use to that documented by Padden et. al. for ISL and ABSL: Older signers show a preference for use of the z-axis, with 60% of verbs produced on z-axis. This preference is not observed in younger signers (40% on z-axis). The apparent decrease in use of the z-axis is concomitant with an increase in use of the x-axis, though Flaherty notes that it is frequently not straightforward to determine from video data the axis upon which verbs are produced.

What factors might underlie a shift from predominantly encoding movement and spatial relationships along the z-axis towards encoding them on the x-axis? One explanation, favoured by Padden and colleagues, is that the use of axis reveals a competition between different kinds of iconicity. In ‘body-as-subject’ (Meir et al. 2007) the signer exploits the iconic possibilities of their own body as an animate subject, mapping the action of the subject onto their own body. This embodied iconicity is contrasted with a schematic or relational iconicity, where the signer makes use of the iconic possibility of the signing space in front of them for representing a scene with participants placed inside the scene like actors on a stage. Under this view, the shift towards use of the x-axis is taken to be a shift towards abstraction: indexing the grammatical role of subject in a referential locus removed from the iconicity of the body as an animate subject. However, abstraction through moving away from the body does not necessarily imply movement along the x-axis, as it is theoretically quite possible for loci to be established along the z-axis. Indeed Padden et. al. note several instances of loci being established along the z-axis in their own data. It is also typologically more common for sign languages to display object agreement, in which the end point of a path movement agrees with the R-locus of the object, than subject agreement (Börstell 2019). Indeed, object marking, but not subject marking, has been claimed to be obligatory in ASL (De Beuzeville et al. 2009), although this is not the case in some other sign languages (e.g. Engberg-Pedersen (1993)). A shift towards greater x-axis use cannot then be fully elucidated in terms of moving away from ‘body-as-subject’, but seems to require further explanation.

One observation is that the choice of location associated with a referent, whilst more abstract than embodied iconicity, is not fully arbitrary. The use of space can be ‘topographically’ motivated, where elements of the real/imagined world are mapped on to the signing space in meaningful ways (Cormier et al. 2015). The choice of location can also be ‘semantically loaded’

as physical proximity to the signer can be used to indicate preference or affinity (Engberg-Pedersen 1993). One reason the x-axis might then be preferred is that it can be used iconically to convey equal discourse weight of referents (Emmorey 2002) as the x-axis allows signers to locate them at equal distance from their body, whilst using the z-axis forces a choice of which referent to place closer to the body.

Another potential advantage of displacing the location of signs from the z-axis is that it increases their visual distinctiveness. Distinctions in location along the x-axis may be easier to perceive for an interlocutor, as they appear more visually distinct than distances of the same magnitude on the z-axis, which from the interlocutor's perspective will overlap, being distinguished only by their depth. A communicative pressure from the interlocutor could therefore exert a pressure for meanings to be encoded along the x-axis. However, if we take the interlocutor into account, consideration of the very notion of axis becomes more complex: visual distinctiveness along the x-axis is useful for the interlocutor only if the x-axis is defined in relation to the them, rather than the signer. When a signer is facing their interlocutor head on, the x-axis defined in relation to each signer's signing space coincides, but if they are positioned at an angle, the x-axis of the signer is offset from the x-axis of their interlocutor to the same degree (for an illustration of this, see **Figure 3**, where the position of the sensor can be conceptualised as the position of an interlocutor). To illustrate with an extreme example, if a signer were placed at a 90 degree angle to their interlocutor, the z-axis defined in relation to the signer's body would correspond to the x-axis defined in relation to the interlocutor's body. Returning to the idea of competing iconicities, one solution that preserves embodiment and its advantages, whilst also increasing distinctiveness on the x-axis would be to shift the body or rotate the torso, such that an embodied movement produced along the z-axis relative to the signer's torso now appears more distinct along the x-axis relative to the interlocutor.

This possibility highlights the fact that the signer's body is not a stable backdrop to the signing space but can be moved by the signer in meaningful ways. Indeed, in role shift, also known as constructed action/constructed dialogue, a signer "shifts" a third person into first person using a perceptible adjustment in the direction of their body, head and gaze for the duration of the role (Padden 1983; Lillo-Martin 2012). This kind of perspective taking is attested as a discourse marker similar to reported speech in many sign languages. Though it is a distinct phenomenon to verbal agreement, it may overlap and interact with such systems, as the direction of the body/head shift often points in the direction of the location associated with the referent being quoted (Emmorey 2002), and there is evidence that verb modification correlates with the presence of constructed action in BSL (Fenlon et al. 2018) and Australian Sign Language (Auslan) (De Beuzeville et al. 2009). This kind of body shifting has also been described in the co-speech gestures (Stec et al. 2016) and silent gesture (So et al. 2005; Motamedi et al. 2018) of hearing non-signers, as well as in homesign. It therefore seems important to take the orientation of the

body into account when determining the axis of motion of verb tokens. Role shift appears to be present in ISL (Meir & Sandler 2007) and NSL (Kocab et al. 2015), though it is reported not to be present in ABSL (Padden et al. 2010). It is unclear how or whether work describing axis use in these languages has taken the signer's body position into account.

The discussion so far has been framed in terms of the emergence of verbal agreement, but as mentioned earlier it is a matter of some controversy whether the schematic structures of spatial modulation under discussion in fact constitute verbal agreement (Lillo-Martin & Meier 2011). Building on Liddell (2000), Schembri et al. (2018) provide an alternate analysis of directional verbs using a construction grammar framework (Goldberg 2003). They argue that verbs that participate in directional modification, which they refer to as indicating verbs, constitute a blend of lexical signs with pointing gestures. This type of unimodal blend of language and gesture is suggested to be typologically unique to signed languages, though is perhaps comparable to the iconic use of prosody in spoken language (see e.g. Shintel et al. (2006); Perlman & Benitez (2010)). Directionality in these constructions is then understood a type of co-sign gesture analogous to co-speech gesture, pointing towards a mental space associated with a referent, rather than marking syntactic agreement. Under this kind of an analysis, there is no reason to expect ubiquitous directional modulation to emerge in sign languages rapidly, or at all (Schembri et al. 2018).

Independently of whether spatial modulation is best understood as verbal agreement or as gestural pointing, it is an empirical question how extensively more established sign languages like ASL make use of the x-axis. Though we are not aware of empirical research on ASL which looks directly at axis use, corpus studies² on the rate of modification of indicating verbs in two established sign languages, Auslan and BSL, are highly informative (De Beuzeville et al. 2009; Fenlon et al. 2018). Both of these studies find that indicating verbs are modified at a lower rate than would be expected. In the Auslan study, up to 63% of indicating verb tokens were coded as spatially modified. A similar figure of up to 68% of tokens were coded as spatially modulated in the BSL corpus. Intuitively, the rate of modification of indicating verbs might correspond somewhat to use of the x-axis, as unmodified verbs are described as those that do not differ from the citation form, which typically moves from a location near the signer to a location directly in front of the signer (i.e. along the z-axis). In fact these notions do not entirely overlap for several reasons. As already mentioned, it is possible for locations to be established along the z-axis, especially for locations corresponding to the thematic role of patient. Indeed, these authors distinguish between 'clearly' modified forms and 'congruent' forms. Congruent forms are cases where there is difficulty distinguishing between a modified

² It should be noted that there may be important differences between spontaneous conversational data as obtained from these corpus studies and elicited data like that in Padden et al. (2010).

and unmodified form, i.e. when the locations of the referents in question would be identical to the citation form. When taking just the clear cases of modification into account, only an average of 55% of tokens are coded as showing modification in Auslan. The second reason is that spatial modification can occur for agent or patient alone (presumably resulting in use of the $x + z$ axis) or for both (x -axis). In Fenlon et al. (2018), 27% of tokens were clearly modified for agent and 52% for patient. It is not stated how many were clearly modified for both. De Beuzeville et al. (2009) do not distinguish these cases. Both studies found that the presence of constructed action (role shift) was correlated with spatial modification, but because they include constructed action established on the basis of eye-gaze and head position alone (p.96), and additionally the position of the interlocutor in these data sets varied, it is unclear exactly how spatial modification is determined in relation to the orientation of the body. Neither study reported a change in rate of modification for younger and older signers, though such a change has been reported for Danish Sign Language, with younger signers reportedly modifying more than older signers (Engberg-Pedersen 1993). This patterns of results is relevant to the claim that more established sign languages make greater use of the x -axis, indicating that at least for Auslan and BSL, the x -axis is not necessarily preferred, though it must also be noted that the rate of modification does not map straightforwardly onto axis use.

In summary, the pattern of data reported for young and emerging sign languages, in which an initial preference for producing directional verbs along the z -axis appears to give way to increasing spatial modulation on the x -axis, has been suggested to contrast with a preference for use of the x -axis in established sign languages. However, the available empirical data on more established sign languages does not appear to show this pattern of x -axis preference, instead showing either a similar pattern of change towards increasing spatial modulation in younger signers (Engberg-Pedersen 1993), or a low rate of modification similar to what is reported for emerging sign languages (De Beuzeville et al. 2009; Fenlon et al. 2018). The position of the body is relevant in determining whether modification is present or not, and does not yet appear to have been taken into account beyond establishing a correlation between the presence of constructed action and the presence of spatial modification.

In this paper we report the results of an exploratory investigation into the use of space by signers of Nicaraguan Sign Language. Building on earlier work on NSL by Flaherty (2014), in which verb tokens were coded categorically by eye into x -axis, z -axis or $z + x$ -axis, we here aim to provide the first automated and continuous quantitative measure of axis use. We use a 3D depth and motion camera (Microsoft Kinect) to capture the relative position of several tracked joints in the body and wrists over time. This allows us to construct a detailed and fine-grained frame by frame picture of directional movement of the wrists, and also allows us to take into consideration the position of the signer's body. Our first goal is to see whether the pattern of axis-use previously described for emerging sign languages including NSL holds

when using this continuous measure. We suggested above that one reason the x-axis may be preferred is that it increases visual distinctiveness for the interlocutor. It follows from this that movement on the x-axis may be easier for researchers to perceive when coding from 2D video, which may lead to an under-estimation of the amount of z-axis movement present. On this basis we expect that our results might differ from previous findings. We also aim to further understand the relationship between the position of the body and directional movement of the wrists. We suggested that one way of increasing visual distinctiveness for the interlocutor whilst preserving a signer's preference for using the z-axis is to rotate the torso away from the signer. We therefore measure the use of axis from two perspectives. One is anchored to the camera (as a proxy for the interlocutor), whilst the other is relative to the orientation of the signer's body.

2 Methods and Procedure

The data set used in this study is a subset of that described in Flaherty et al. (submitted). For the reader's convenience, the methods and procedure for data collection are reiterated below.

2.1 Participants

Participants were recruited through the principal investigator's contacts in the community. Seventeen deaf signers took part (7 women, 10 men). All were native signers of Nicaraguan Sign Language, having been exposed to the language upon entry to school, before the age of 6. The year in which participants entered school spans from 1974 to 2003, giving an almost 30 year window into the different stages of the language to which signers were exposed, including signers of the first cohort (who created/were not exposed to an earlier version of the language). Participants received financial compensation for their participation in our study, and gave informed consent regarding their participation and use of their data. Ethical approval for this study was obtained from the University of Edinburgh's Ethics Board.

2.2 Stimuli

Participants viewed a series of 36 short video vignettes. The vignettes depicted a set of 18 actions (approach, crawl to, cycle to, feed, give ball to, hop to, jump to, poke, pull, punch, push, roll ball, run to, skip to, tap/touch, throw ball to, throw confetti on, cycle to) between two entities. These actions were chosen for their suitability for potential representation on the x-axis using directional movement between two abstract locations. Each action had two vignettes, one in which an animate entity (a man or a woman) acted on another animate entity (a man or woman) and one in which an animate entity acted on an inanimate entity (a chair or a plant). The actors/entities were always located on either side of the screen across from one another. The full set of video vignettes used are available in the supplementary materials.

2.3 Procedure

Participants viewed the 36 target event videos in one of two random orders on a laptop screen and were asked to describe what they saw to a signer of their peer group who could not see the laptop screen. The last frame of each vignette remained visible on the screen while participants produced their descriptions. Signers could provide clarification or repeat their description if their communication partner requested it, but only signers' first responses were included in our analysis. Responses were both videotaped and recorded using a Microsoft Kinect 3D camera (Figure 1). The Kinect tracked the position of signers' wrists by using depth and motion sensors to return the inferred XYZ position of 21 joints in the body at a target frame rate of 30 fps (Wang et al. 2015).

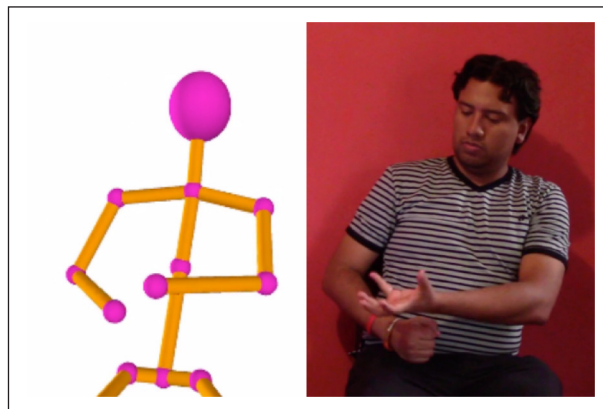


Figure 1: RGB and Kinect Frame.

3 Results

Prior to the analyses reported below, the body tracking data and video recordings were time-aligned in order to identify the correspondence between frames in the video and body tracking data. Time-aligning was necessary because the Kinect device records at a variable framerate and does not always achieve the target frame rate of 30 fps, meaning that for each second of video data there is a variable number of frames in the body tracking data. Missing frames were then interpolated from existing frames. The body tracking data was also filtered using median filtering to reduce noise (Microsoft 2005), and each participant's body tracking data was transformed by a scaling factor to remove any effect of differing body sizes. Each utterance was glossed and the first and last frame of each verb was identified as described in Flaherty et al. (submitted). After excluding one utterance due to failure of the device during recording, 1074 verb tokens were included in this analysis. For our analysis of axis-use, an additional step was undertaken in the linguistic coding to identify the handedness with which verbs were produced. This is because

our measure of axis (described below) is based on movement so it was important to know which hand(s) was(/were) moving meaningfully, so as for our measure not to be affected, for instance, by a non-dominant hand resting in the signer’s lap for the duration of the verb. Some of our signers were right or left-handed, but others showed variable handedness. We therefore coded handedness for each verb token. We identified three categories of handedness. Symmetrical verbs were those in which both hands produce the same movement. Asymmetrical verbs were coded as either left or right hand dominant verbs. For verbs produced with one hand, classification was straightforward. For asymmetric verbs produced with both hands, tokens were classified according to the hand whose movement had the longest path for that token.

In order to evaluate the use of axis, we constructed a measure r based on the variance of the tracked position of the wrists, given by the equation 1:

$$r = \ln \left(\frac{\sigma(x_1, x_2, \dots, x_n)}{\sigma(z_1, z_2, \dots, z_n)} \right) \quad (1)$$

where x_i and z_i are the tracked x and z coordinates of the wrist at the i th frame in a verb token, and σ is the usual function for standard deviation. For verbs with one dominant hand, we only take into account the coordinates of the dominant wrist. For symmetrical verbs, r is calculated separately and then averaged over both wrists. The ratio is log transformed to aid interpretation as log transformation results in symmetry around zero.³ Positive values indicate more variation on the x-axis, or more horizontal movement, whilst negative values indicate more variation on the z-axis, or sagittal movement outward from the body. Values closer to zero indicate more similar variation on both axes. This could represent diagonal movement, or other kinds of movement which make use of both axes, such as circular and zigzagging path movements. Example verb tokens from our dataset illustrating how r relates to axis use are shown in **Figure 2**. Additionally, we take the signer’s body position into account by calculating this measure from two perspectives: a signer-centric perspective and a camera-centric perspective. The two differ in whether the coordinate system is construed in relation to the signer’s body or in relation to the camera (interlocutor). For the camera-centric perspective, the origin of the coordinate system is at the camera, with the x-axis growing to the right of the camera, and the z-axis growing into the space in front of the camera (this is the perspective used in **Figure 2**). For the signer-centric perspective, the coordinate system is instead anchored to the signer’s body, with the x-axis passing through the signer’s shoulders, and the z-axis growing into the space in front of the signer’s body at each frame. The difference between the two perspectives is illustrated in **Figure 3**. Note that for a hypothetical token in which the signer’s body is facing the camera

³ Without log transformation, cases with higher x variation would range between (0,1), whilst cases with higher z variation would range between (1, ∞), making it difficult to compare x and z variation to one another.

perfectly, the two perspectives are equivalent in terms of the ratio of the variation on the two axes, differing only in the polarity of the z-axis. The two perspectives diverge only when the signer's shoulders rotate away from a neutral position relative to the camera.

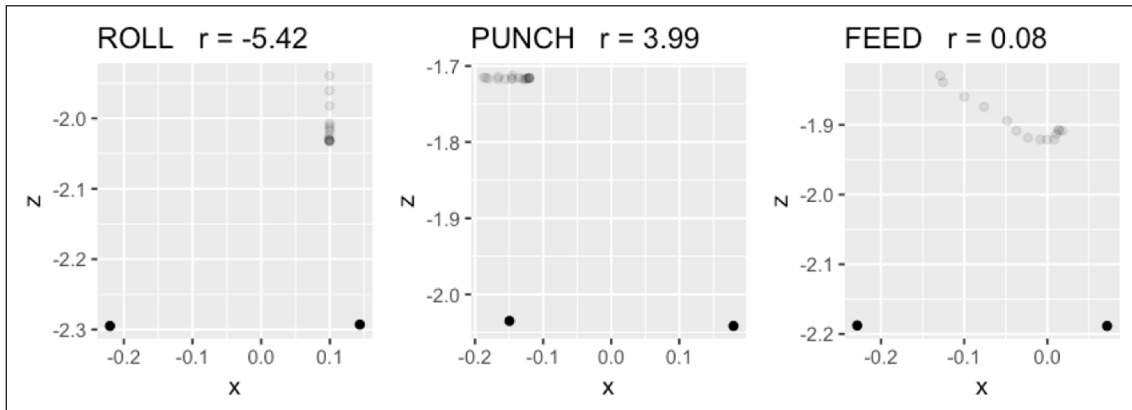


Figure 2: Each panel contains the overlaid frames from a single verb token from a bird's-eye view. The average position of the shoulders for each token is plotted in black, and the tracked position of the dominant wrist at each frame is overlaid with transparency. Negative values of r correspond to higher variability on the z-axis, positive values of r correspond to higher variability on the x-axis, and values closer to zero correspond to more similar variability on both axes.

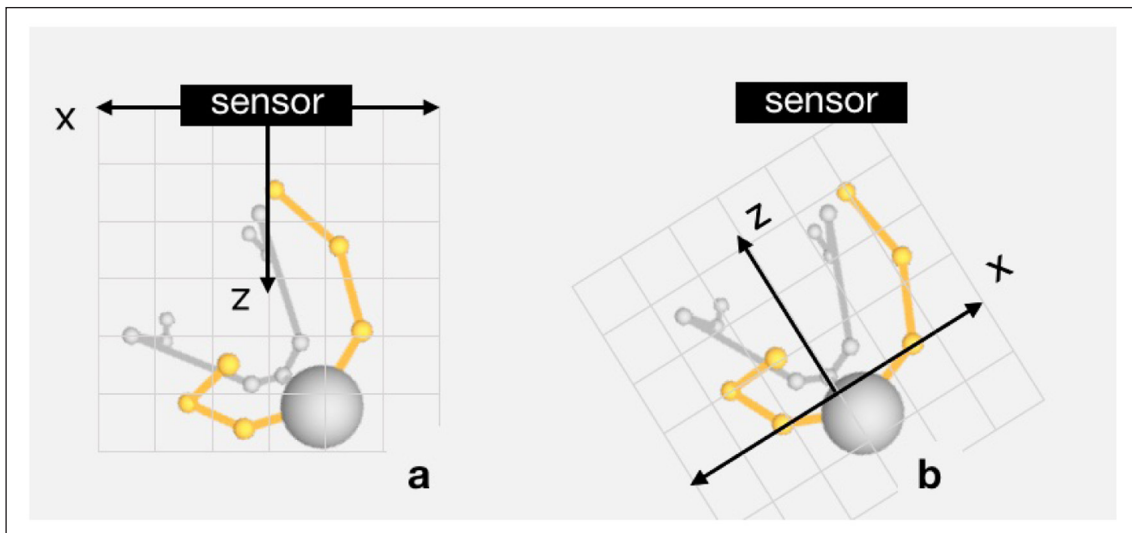


Figure 3: Our measure of axis use was calculated from two perspectives which differ with respect to the origin and orientation of the coordinate system for the same frame. The camera-centric perspective (**a**) is anchored at the Kinect sensor, whilst the signer-centric perspective (**b**) is anchored on the signer's body.

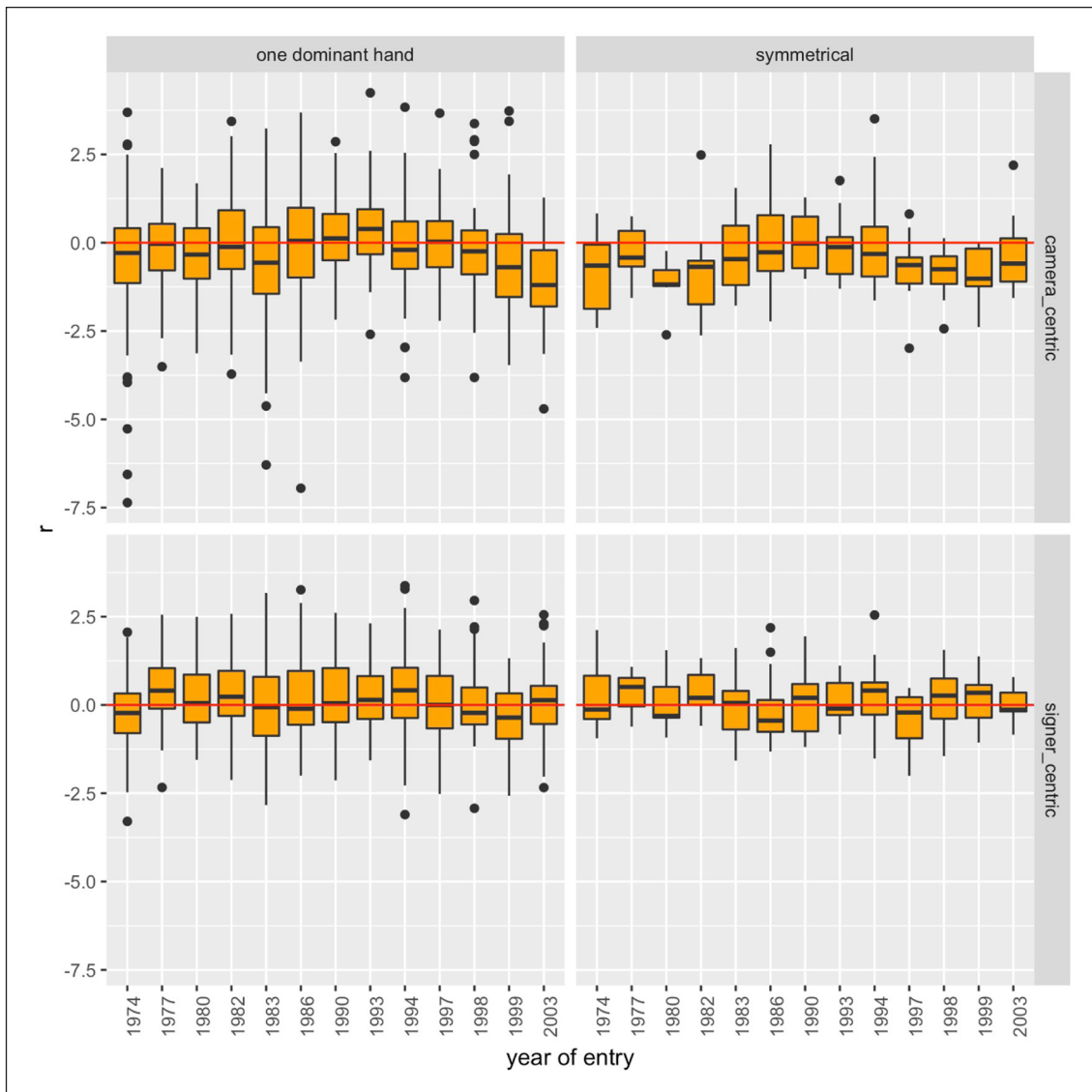


Figure 4: r by year of entry into the community and hand-dominance. A positive value of r indicates greater movement of the wrists in the x-axis, whereas a negative value indicates greater movement in the z-axis. We see some evidence for greater movement on the z-axis in the symmetrical verbs and from the camera (i.e. interlocutor) perspective. There is no effect of year of entry in our data.

Figure 4 shows values of r by year of entry into community. When calculated from a camera-centric perspective, we see slightly more variability on the z-axis, with an average r of -0.22 for tokens produced with one dominant hand, and -0.41 for symmetrical tokens produced with two hands. From the signer-centric perspective, we see an average r of 0.09 for tokens produced with one hand, and 0.04 for two-handed tokens. In order to assess the use of axis in our data set

we ran a mixed effects model predicting r with fixed effects of year of entry into the community (centered so that the intercept corresponds to the first year in our dataset), perspective (signer-centric or camera-centric), and hand dominance (symmetrical or one-dominant). Perspective and hand-dominance were sum coded so each level is compared to the mean for that variable. We included random intercepts per participant. Comparison to the null model by likelihood ratio testing confirms that our model explains the variance in our data better than chance ($X^2(3) = 46.89$, $p < .001$). Comparison to reduced models with each fixed effect removed reveals a main effect of perspective ($X^2(1) = 43.32$, $p < .001$), with an estimated decrease in r of -0.17 for the camera-centric perspective, relative to the model intercept of -0.15 ($\beta = -0.17$, standard error = 0.03). Recall that negative values of r correspond to greater variance on the z-axis; this result indicates that there is an overall preference for use of the z-axis. However, this preference is seen primarily from the camera (interlocutor) perspective. It therefore does not appear that signers are rotating their torso in order to increase visual distinctiveness for the interlocutor. Instead, signers appear to be making approximately equal use of both axes, when these are considered in relation to their own body. We also observe a main effect of hand-dominance, with symmetrical verbs showing more variance on the z-axis ($\beta = -0.06$, standard error = 0.03). There was no effect of year of entry into the community ($X^2(1) = 0.10$, $p = 0.74$), indicating that there is no change in the use of axis between younger and older signers.

4 Discussion

Our results indicate that signers of NSL show an overall preference for encoding movement along the z-axis. We do not see any evidence of a trend towards increasing use of the x-axis, nor do we see any evidence that signers rotate their bodies in such a way that increases visual distinctiveness on the x-axis for the interlocutor. Instead, when we look at the use of axis from a perspective anchored to the signer's body, we find that the apparent preference for movement along the z-axis disappears, and signers appear to be making use of the space in front of their body approximately equally on both axes.

The difference between the two perspectives suggests that the movement we see on the z-axis from the camera-centric perspective could be driven by signers shifting their body forwards. Moving the body allows the wrists to travel further in absolute space than would be reflected in the signer-centric measure, where the origin of the coordinate system moves with the signer. This calls attention to the fact that although our signer-centric measure was designed to take the signer's body into account, it does so through reference to the *orientation* of the body (the plane passing through the shoulders), and therefore does not capture shifts of the torso away from a neutral position. In other words, by anchoring our measure of axis to the signer's body, we gain information about how the space in front of the singer's body is used, but we lose information about how the body itself moves in space. We therefore checked the result

of computing r on the position of the body instead of the wrists. For this we use the midpoint between the shoulders at each frame, equivalent to the origin of the coordinate system in the signer-centric perspective. Using this point allows us to capture shifting or leaning (as opposed to rotation) of the torso. The result of this exploratory post-hoc analysis shows that signers are more likely to move their body back and forth than side to side (mean r of -0.7 ($t = -18.90$, $df = 1070$, $p < .001$)). This is in line with our intuition that the preference for z-axis we observe is driven by body shift, and further underscores the need for more nuanced consideration of the body in discussions of axis use.

Although the finding of a preference for the z-axis is in line with the reported pattern for ISL and ABSL, and from earlier work on NSL, our result also diverges from those reported in these studies in that we do not see any difference in axis use between younger and older signers. This may be because of how our notion of axis differs from that in these earlier studies, as our notion of axis is based primarily on the position and movement of the wrists. Even though we have incorporated the body into our analysis, it is important to note that the subjective perception of axis by the human eye comprises a more holistic judgement which may incorporate multiple other factors, such as the orientation of the face and direction of the gaze, the relative position of wrist and elbow, and so on. Additionally, some of these factors may be specific to individual sign languages. Although we do not observe a trend towards use of the x-axis, it is possible that there is indeed still a general trend from body as subject towards more abstract schematic iconicity in our data, but that this pattern is not discernible through looking at axis alone. As discussed in the introduction, it is possible for schematic iconicity to make use of the z-axis. Qualitatively, we see several instances of this in our data set. There are also other language-internal factors that may contribute to the degree to which the z-axis is preferred, such as the presence of body-anchored verbs which cannot be fully produced on the x-axis (e.g. the ASL verb TELL starts at the chin) (see Flaherty & Goldin-Meadow (in prep)).

A final consideration in comparing our results to those of previous studies is our use of motion tracking technology in determining axis. The use of motion tracking technology allows for fine-grained quantification of gradient phenomena in a way that is difficult to achieve with manual coding schemes. There are advantages to this kind of automation in terms of reducing researcher bias, as well as the potential for efficiently collecting and analysing large amounts of data. However, this needs to be calibrated against a careful consideration of how our measures relate to manually coded categorical measures of the same phenomena, as well as consideration of the ethical implications of the application of this technology to the study of minority languages by researchers like ourselves who are not members of those languages communities. We are particularly keen to encourage researchers to obtain this kind of data for a wider pool of sign languages not limited to the Global South.

Some researchers are hopeful that the development of new tools in the application of motion tracking technology will be crucial in deepening our understanding of continuous aspects of sign and gesture, perhaps comparable to the development of the spectrograph (Potter et al. 1966) for the spoken modality (Goldin-Meadow & Brentari 2017), though others are less convinced (Emmorey 2017). There are certainly limitations inherent in the use of motion tracking technology. Indeed, our measure of axis focused on the position of the wrists because the Kinect device we used does not offer precise data on the configuration or orientation of the hand. Even so, our data is informative not only in how NSL signers are using the space in front of the body, but also the relevance of the position and orientation of the body in considerations of axis use more broadly.

It is clear that the rich multi-modality of signed languages cannot be reduced to a simple digital signal and a lot of groundwork is still required in the application of motion tracking to the study of signed languages. Nonetheless, we believe that the use of this kind of technology holds great potential for fine-grained analyses of continuous data in the manual modality and we hope that our contribution will encourage other researchers in the sensitive application of technology in this line of investigation. Productive research in this area is ultimately likely to involve a collaborative approach, integrating the expert linguistic knowledge that is required in the implementation of traditional coding schemes with the benefits of automation and reduction in researcher bias offered by new technological advances.

Acknowledgements

We acknowledge Lynn Yong-Shi Hou and Diane Lillo-Martin for their thoughtful and insightful comments and suggestions on our original manuscript. Asha Sato acknowledges the Economic and Social Research Council for provision of funding (grant number ES/J500136/1) during the period this research was carried out.

Competing Interests

The authors have no competing interests to declare.

References

- Börstell, CA. 2019. Differential object marking in sign languages. DOI: <https://doi.org/10.5334/gjgl.780>
- Cormier, Kearsy & Fenlon, Jordan & Schembri, Adam. 2015. Indicating verbs in british sign language favour motivated use of space. *Open Linguistics* 1(open-issue). DOI: <https://doi.org/10.1515/opli-2015-0025>
- De Beuzeville, Louise & Johnston, Trevor & Schembri, Adam C. 2009. The use of space with indicating verbs in auslan: A corpus-based investigation. *Sign language & linguistics* 12(1). 53–82. DOI: <https://doi.org/10.1075/sll.12.1.03deb>
- Emmorey, Karen. 2002. *Language, cognition, and the brain insights from sign language research*. Mahwah, N.J.: Lawrence Erlbaum Associates. DOI: <https://doi.org/10.4324/9781410603982>
- Emmorey, Karen. 2017. How to distinguish gesture from sign: New technology is not the answer. *The Behavioral and brain sciences* 40. e54. DOI: <https://doi.org/10.1017/S0140525X15002897>
- Engberg-Pedersen, Elisabeth. 1993. Space in danish sign language: The semantics and morphosyntax of the use of space in a visual language.
- Fenlon, Jordan & Schembri, Adam & Cormier, Kearsy. 2018. Modification of indicating verbs in british sign language: A corpus-based study. *Language* 94(1). 84–118. DOI: <https://doi.org/10.1353/lan.2018.0002>
- Flaherty, M. & Goldin-Meadow, S. in prep. Argument structural relations in nicaraguan sign language.
- Flaherty, Molly. 2014. *The emergence of argument structural devices in nicaraguan sign language*: The University of Chicago dissertation.
- Flaherty, Molly & Sato, Asha & Kirby, Simon. submitted. Documenting a reduction in signing space in nicaraguan sign language using depth and motion capture. *Cognitive Science*.
- Goldberg, Adele E. 2003. Constructions: A new theoretical approach to language. *Trends in cognitive sciences* 7(5). 219–224. DOI: [https://doi.org/10.1016/S1364-6613\(03\)00080-9](https://doi.org/10.1016/S1364-6613(03)00080-9)

- Goldin-Meadow, Susan & Brentari, Diane. 2017. Gesture, sign, and language: The coming of age of sign language and gesture studies. *Behavioral and Brain Sciences* 40. DOI: <https://doi.org/10.1017/S0140525X1600039X>
- Kocab, Annemarie & Pyers, Jennie & Senghas, Ann. 2015. Referential shift in nicaraguan sign language: A transition from lexical to spatial devices. *Frontiers in psychology* 5. 1540. DOI: <https://doi.org/10.3389/fpsyg.2014.01540>
- Liddell, Scott K. 2000. Indicating verbs and pronouns: Pointing away from agreement. *The signs of language revisited: An anthology to honor Ursula Bellugi and Edward Klima* 303. 320.
- Lillo-Martin, Diane. 2012. Utterance reports and constructed action. *Sign language: An international handbook* 365. 387. DOI: <https://doi.org/10.1515/9783110261325.365>
- Lillo-Martin, Diane & Klima, Edward S. 1990. Pointing out differences: Asl pronouns in syntactic theory. *Theoretical issues in sign language research* 1. 191–210.
- Lillo-Martin, Diane & Meier, Richard P. 2011. On the linguistic status of ‘agreement’ in sign languages. *Theoretical linguistics* 37(3–4). 95–141. DOI: <https://doi.org/10.1515/thli.2011.009>
- Mathur, G. & Rathmann, C. 2012. Verb agreement. In Pfau, R., & Steinbach, M., & Woll, B (eds.), *Sign language: An international handbook*, chap. 7, 137–157. Berlin: De Gruyter Mouton. DOI: <https://doi.org/10.1515/9783110261325.136>
- Meir, Irit & Padden, Carol A. & Aronoff, Mark & Sandler, Wendy. 2007. Body as subject. *Journal of Linguistics* 43(3). 531. DOI: <https://doi.org/10.1017/S0022226707004768>
- Meir, Irit & Sandler, Wendy. 2007. *A language in space: The story of israeli sign language*. Psychology Press. DOI: <https://doi.org/10.4324/9780203810118>
- Microsoft. 2005. Microsoft kinect white paper. skeletal joint smoothing white paper. <https://msdn.microsoft.com/en-us/library/jj131429.aspx#ID4EQMAE>.
- Motamedi, Yasamin & Smith, Kenny & Schouwstra, Marieke & Culbertson, Jennifer & Kirby, Simon. 2018. The emergence of systematic argument distinctions in artificial sign languages. DOI: <https://doi.org/10.31234/osf.io/p6zy4>
- Padden, Carol. 1983. *Interaction of morphology and syntax in american sign language*: University of San Diego dissertation.
- Padden, Carol & Meir, Irit & Aronoff, Mark & Sandler, Wendy. 2010. The grammar of space in two new sign languages. In Brentari, Diane (ed.), *Sign languages: A cambridge language survey*, 570–592. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511712203.026>
- Perlman, Marcus & Benitez, Natalie J. 2010. Talking fast: The use of speech rate as iconic gesture. *Meaning, form, and body* 245. 262.
- Polich, Laura. 2005. *The emergence of the deaf community in nicaragua: "with sign language you can learn so much"*. Gallaudet University Press.
- Potter, Ralph K., et al. 1966. Visible speech.
- Schembri, Adam & Cormier, Kearsy & Fenlon, Jordan. 2018. Indicating verbs as typologically unique constructions: Reconsidering verb ‘agreement’ in sign languages. *Glossa: a journal of general linguistics* 3(1). DOI: <https://doi.org/10.5334/gjgl.468>

- Senghas, Ann & Coppola, Marie. 2001. Children creating language: How nicaraguan sign language acquired a spatial grammar. *Psychological science* 12(4). 323–328. DOI: <https://doi.org/10.1111/1467-9280.00359>
- Shintel, Hadas & Nusbaum, Howard C. & Okrent, Arika. 2006. Analog acoustic expression in speech communication. *Journal of Memory and Language* 55(2). 167–177. DOI: <https://doi.org/10.1016/j.jml.2006.03.002>
- So, Wing Chee & Coppola, Marie & Licciardello, Vincent & Goldin-Meadow, Susan. 2005. The seeds of spatial grammar in the manual modality. *Cognitive science* 29(6). 1029–1043. DOI: https://doi.org/10.1207/s15516709cog0000_38
- Stec, Kashmiri & Huiskes, Mike & Redeker, Gisela. 2016. Multimodal quotation: Role shift practices in spoken narratives. *Journal of Pragmatics* 104. 1–17. DOI: <https://doi.org/10.1016/j.pragma.2016.07.008>
- Sutton-Spence, Rachel & Woll, Bencie. 1999. *The linguistics of british sign language: an introduction*. Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9781139167048>
- Wang, Qifei & Kurillo, Gregorij & Ofli, Ferda & Bajcsy, Ruzena. 2015. Evaluation of pose tracking accuracy in the first and second generations of microsoft kinect. In *2015 international conference on healthcare informatics*, 380–389. IEEE. DOI: <https://doi.org/10.1109/ICHI.2015.54>

