Perceptual optimization of language: evidence from American Sign Language

Naomi Caselli
Boston University, USA

Corrine Occhino
Syracuse University, USA

Bruno Artacho, Matthew Dye*, Andreas Savakis
Rochester Institute of Technology, USA

* Corresponding author: Matthew Dye, RIT/NTID, Rochester NY 14623, mwddls@rit.edu

Word Count: 3,760

Abstract

If language has evolved for communication, languages should be structured such that they maximize the efficiency of processing. What is efficient for communication in the visual-gestural modality is different from the auditory-oral modality, and we ask here whether sign languages have adapted to the affordances and constraints of the signed modality. During sign perception, perceivers look almost exclusively at the lower face, rarely looking down at the hands. This means that signs articulated far from the lower face must be perceived through peripheral vision, which has less acuity than central vision. We tested the hypothesis that signs that are more predictable (high frequency signs, signs with common handshapes) can be produced further from the face because precise visual resolution is not necessary for recognition. Using pose estimation algorithms, we examined the structure of over 2,000 American Sign Language lexical signs to identify whether lexical frequency and handshape probability affect the position of the wrist in 2D space. We found that frequent signs with rare handshapes tended to occur closer to the signer's face than frequent signs with common handshapes, and that frequent signs are generally more likely to be articulated further from the face than infrequent signs. Together these results provide empirical support for anecdotal assertions that the phonological structure of sign language is shaped by the properties of the human visual and motor systems.

Keywords: American Sign Language, language optimization, pose estimation, language production, language perception

1. Introduction

A longstanding debate within linguistics centers on whether or not languages have evolved for communication. Chomsky and others (Chomsky, 1975; Hauser et al., 2002) have argued that language has primarily evolved for thinking. Others argue that language has primarily evolved for communication (see Gibson et al., 2019 for a review), appealing to the idea that languages evolve to optimize efficiency, minimizing effort on the part of the listener and on the part of the speaker (e.g. Hockett, 1960; Zipf, 1949; Piantadosi & Fedorenko, 2017). This balance minimizes effort required for the speaker to generate the message, and minimizes the effort required for the listener to decode the message. For example, language that is more informative or less predictable tends to be lengthier than less informative or more predictable language, affording the listener more time to decode the message (e.g., Piantadosi et al., 2011; Aylett & Turk, 2004). To the extent that languages evolve for communication, the affordances and limitations of the human body may have left an imprint on the structure of language.

Sign languages offer a unique opportunity to ask how the body shapes the structure of language, because they are produced using different parts of the human body than spoken languages and lend themselves to a different set of perceptual and motoric capacities. What is efficient in one modality, may not be efficient in another. In most respects, sign languages conform to linguistic principles common to spoken languages, and include all the levels of linguistic structure that spoken languages have (e.g., Brentari, 1998; Mathur & Rathmann, 2012; Wilbur, 1987) and are processed in much the same way (see Emmorey, 2007, for a review). However, at a surface level there are clear differences between signed and spoken languages: sign languages are produced using the hands and body and, except in DeafBlind people, are generally perceived via the visual system. If languages evolve to maximize communicative efficiency, we might expect sign languages to be optimized for the manual-visual system.

Communicative efficiency in sign languages may reflect both ease of perception and ease of articulation. We focus this paper on ways perceptual demands may shape sign languages, but note throughout how articulatory demands may also play a part. Skilled signers look at the face of the person signing during sign perception (Agrafiotis et al., 2003; Muir & Richardson, 2005; Emmorey et al., 2009), and generally do not look at the hands, and therefore signs that are not produced near the face must be perceived using peripheral vision. The encoding of visual stimuli in the periphery is “lossy” and has access to relatively fewer neural resources compared to central vision. Together these facts about the visual system mean that signers may be better able to detect fine detail about signs that are produced near the face than signs produced at more peripheral locations in signing space. If sign languages evolve to match the perceptual abilities of comprehenders, signs that require fine perceptual discriminations should be more likely to be produced on or near the face (near the location of the viewer’s eye gaze) whereas those that do not require such discriminations should be more likely to be located further away in the sign space (Siple, 1978). Sign languages may also evolve to maximize the efficiency of articulation (e.g., with hands closer to resting position, using handshapes that are easier to generate). Repeated articulations of frequently used signs, for example, leads to routinized articulatory motor patterns, which in turn can become reduced, relative to older forms.

Thanks to the pioneering work of Frishberg, Battison, Siple, Woll and others it is commonly understood that the production of sign languages does indeed correspond to the visual perceptual abilities of those comprehending the signal. Frishberg (1976) identified a host of diachronic changes that occurred in American Sign Language over time, including spatial displacement in where a sign is articulated in order to make signs more perceivable (closer to the face). Woll (1987) noted shifts in the place of articulation of some signs in British Sign Language (BSL). For example, signs that were once produced in more distal body locations may change to locations immediately in front of the body (e.g., the BSL sign PERHAPS was historically produced on the forehead, but over time moved to a more central location in front of the chest). Similarly, Battison (1978) and Siple (1978) note diachronic reductions in movement and the phonological complexity of signs. However, because of technical limitations, the evidence for these claims was based on a small set of examples. There has been no rigorous, systematic investigation of these claims with more modern tools.

In sign languages, each sign has a handshape, and handshapes vary with respect to markedness. Markedness is a multifaceted construct that refers to features that, relative to unmarked features, are less frequent within and across languages, and harder to learn and process (see Rice, 2007, for a review). As we describe in the following section, the existing literature on sign languages characterizes handshape typicality in terms of markedness as a whole, but in the present study we focus more narrowly on one aspect of markedness: frequency within a language. Battison (1978) proposed a set of seven unmarked handshapes for sign languages on the basis that these handshapes are relatively more frequent, appear cross-linguistically, are easiest to perceive, and less restricted (i.e. can occur in two-handed asymmetrical signs) than other handshapes. Unmarked handshapes are more resistant to distortion by noise than marked handshapes (Lane et al., 1976), and children can acquire unmarked handshapes earlier (Siedlecki & Bonvillian, 1997; Marentette & Mayberry, 2000, Cheek et al., 2001; Clibbens & Harris, 1993; Karnopp, 2002; Morgan et al., 2007; von Tetzchner, 1984; Takkinen, 2003).

As Siple (1978) first proposed, if sign languages evolve to maximize efficiency on the part of comprehenders, signs that require fine perceptual discriminations, such as identifying marked handshapes, should be more likely to be produced on or near the face where they will be close to the center of the viewer's line of sight and thus easier to recognize. Fenlon et al. (2017) briefly report that in British Sign Language, a high proportion of signs with marked handshapes are produced on the head or neck, but it is unclear how this compares to signs with unmarked handshapes. Furthermore, it is not clear that using these place of articulation categories is optimal, as signs that are produced on the hand, arm, and in neutral space may or may not appear in the line of sight of the perceiver. For example, the sign BOOK (Figure 1) is not produced in the 'head' place of articulation, but is produced in front of the signer's face and so is likely to be directly in the perceiver's line of sight. Nevertheless, based on this and previous work, we predict that the spatial distribution of marked handshapes will be different to that of unmarked handshapes and more likely to appear nearer the face.[1]

Usage-based approaches to phonological change in spoken languages have shown that higher frequency words undergo sound change at a faster rate relative to low frequency words (Bybee, 1998; 2001; 2015). For ASL, one might make two predictions regarding frequency and its relation to phonological change. The first prediction involves location: frequent signs will be produced further from the face than infrequent signs. This pattern might arise because of perceptual efficiency, in that frequent lexical items are generally more predictable than infrequent lexical items (Bolinger, 1981; Fowler & Housum, 1987; Gregory et al., 1999), and so they may be free to occur in a location that is less easily perceived (i.e., further from the face).

[1] We note that one might also make the opposite prediction: signs with uncommon handshapes will be produced in more distal locations to highlight differences in their locations.

This pattern might alternatively arise because of articulatory efficiency: frequent signs should be produced with minimal physical exertion (i.e., with hands closer to a resting position[2]). The second prediction involves an interaction between location and handshape markedness: effects of handshape markedness may be more easily detected in frequent signs because frequent signs have a faster mutation rate and so will more readily select to maximize communicative efficiency than infrequent signs. In other words, signs with marked handshapes may be more likely to be produced closer to the face than signs with unmarked handshapes, but the difference will be more pronounced in high frequency signs.

In the present study, we used a new set of methodological tools to test these predictions. We selected signs from the ASL-LEX database (Caselli et al. 2017; Sevcikova Sehyr et al., 2021), which includes a video exemplar of each sign as well as detailed information about the hand configurations used in each sign. Markedness encompasses a number of factors (frequency within languages, frequency across languages, perceptual ease, articulatory ease, etc.), and in this paper we operationalized markedness by the frequency of the handshape within the language (we henceforth refer to this as handshape probability). Video exemplars were processed using a convolutional neural network called OpenPose (Cao et al. 2018) that generated a 2D representation of the major joints and body locations of the sign model. This allowed us to generate a spatial distribution map for the right wrist locations of all signs, and to compare those distributions[3] as a function of handshape probability and lexical frequency using linear mixed models. We tested three predictions:

1. Infrequent signs will be produced closer to the face than frequent signs.
2. Signs with uncommon handshapes will be produced closer to the face than signs with common handshapes.
3. Frequency and handshape probability will interact such that the effect of handshape probability will be larger for high frequency signs.

2. The ASL-LEX Dataset

[2] Here we consider the case for signs produced in isolation. However, for signs coarticulated within the context of a sentence, an efficient location might be closer to the location of the preceding or following sign.

[3] For visualization purposes, we present the spatial distributions of the right wrist. However, analytically we used the distance between the nose and the centroid of the right wrist across each sign's production as our dependent variable (see Section 3.3).

We drew 2,677 lexical signs from ASL-LEX[4], a lexical database that contains lexical and phonological information about signs in ASL (Caselli et al. 2017; Sevcikova Sehyr et al., 2021). The same deaf native signer produced all of the signs in ASL-LEX. She is female, middle-aged, White, born in the North-East USA, and resides in California. The signer produced the signs with mouthing of English words as she felt was natural and appropriate. The signs were exported at a frame rate of 29.93 frames per second, and edited into individual video clips. The videos begin and end with the signer's hands at rest. Sign onset and sign offset points, as coded in ASL-LEX, were used to trim the videos and remove the movements of the wrists to and from the signer's lap. Signs that were categorized as 'gestures' or 'violations' in the ASL-LEX database were removed, along with those signs that had missing values for our key predictor variables, resulting in a final dataset of 2,613 signs.

Nine phonological features were drawn from ASL-LEX. These include five handshape features (selected fingers, flexion, spread, thumb position, and thumb contact) as well as the *major location* of articulation and *sign type*. Sevcikova Sehyr et al., (2021) include a detailed description of the phonological coding procedure for these features. Briefly, *selected fingers* were defined as: (1) the group of fingers that move, (2) if none of the fingers move, the fingers that are not fully extended nor fully closed, (3) if neither of the first two rules apply, the fingers that are fully extended. *Flexion* at the sign onset included five values that fell into a roughly ordinal scale from fully extended to fully flexed plus one category for stacked (like the manual letter P) and one category for crossed (like the manual letter R). *Flexion Change* was coded as change or no change. *Spread* was coded as abducted or adducted at the sign onset. *Spread change* was coded as change or no change. *Thumb position* was coded as open or closed at the sign onset. *Thumb contact* was coded as making contact with the selected fingers at some point during the sign or not. The distribution of these phonological features is shown in Figure 2.

2.1. Handshape Probability Calculation

For each sign, we then computed the handshape probability by first calculating the sub-lexical frequency of each of the seven phonological features corresponding to handshape (the frequency of the value of a phonological feature / number of signs in the dataset). For example,

[4] The dataset for this study predated the publication of the second version of ASL-LEX, and the total number of sign videos that were available at the time is slightly different than what was published in Sevcikova Sehyr et al. (2021).

the selected finger in the sign APPLE is the index finger, which occurs in 676 signs in this dataset of 2,613 signs, so has a sub-lexical frequency of 0.259. We then averaged the sub-lexical frequencies of all seven handshape features to reach a handshape probability (in the case of APPLE, 0.346). We used handshape probability rather than using other measures of handshape markedness (e.g., Battison's marked handshapes) because many handshapes are dynamic and cannot be easily classified into these categories. Sign frequency and handshape probability were weakly, but significantly, correlated in this dataset ($r = 0.07$, $p < 0.001$).

3. Pose Estimation

We identified the position of the hand relative to the face using "deep learning" approaches to perform pose estimation on video and track the motion of joints over time.

3.1. Joint Detection

Videos of individual lexical signs from the ASL-LEX database were sampled at a common resolution of 640x480 and frame rate of 29.93 fps. We used the OpenPose (Cao et al. 2018) technique for pose estimation to determine the signer's joint locations in each video. OpenPose improves upon the popular Convolutional Pose Machine (CPM) method used in other works (Wei et al. 2016). In addition to the multi-stage structure in the CPM architecture that refines joint localization, OpenPose incorporates Part Affinity Fields (PAF) that use the detection of joints where confidence is high to better estimate the prediction of joints where the confidence is lower. In this way, OpenPose detects 18 key points in the body as well as facial landmarks in each video frame.

3.2. Normalization and Filtering

The distance between the neck and hip joints was used to normalize the body in order to account for differences in size and position of the signer across videos. Median (Tukey 1977) and Kalman (Kalman 1960) filters were then applied to the joint coordinate data. Median filtering replaces the detection location in the current frame by the median of the detections in adjacent frames, resulting in the suppression of outliers that may have occurred in cases when the detections fail. Kalman filtering estimates the probability distribution of the joint detections across frames with the objective of making predictions that reduce the effects of statistical noise representing detection errors. This type of post-processing, with Median and Kalman filtering in the temporal domain, eliminates outliers in the detection process and results in better estimates of the joint trajectories across time. Our variable of interest was the distance between the sign model's nose and the centroid of the right wrist location. However, the sign model was free to move her whole body during sign production and therefore the nose position varied from frame-to-frame. Therefore, the post-processed joint coordinates were translated with respect to the signer's nose location, which was thereby fixed at the same spatial coordinates for all videos and served as a reference point for all joints. Figure 3 shows the estimation of nose and wrist locations by OpenPose, and post-processing of the joint detections by median and Kalman filtering of the joint coordinates across video frames.

3.3. Spatial Position of Articulators

For each of the signs in ASL-LEX, the process detailed above produced a series of 2D spatial coordinates corresponding to sequential positions of the signer's right wrist. The (x,y) coordinates of the right (dominant) wrist was averaged for each sign, to compute a centroid for each exemplar. We then computed the Euclidean distance between the signer's nose and that wrist centroid.

4. Analysis

We first present a visualization of the distribution of signs in 2D space, and then report statistical analyses. As a validity check, we confirmed the data visualization technique by plotting the wrist position by the manually-tagged major locations from ASL-LEX (see Figure 4). The plotted wrist distributions match the distributions that might be expected (e.g., "head" locations are generally higher than "body" locations).

Next we visualized the distribution of wrist positions according to Sign Frequency and Handshape Probability. Sign Frequency and Handshape Probability were each divided into three quantiles. Figure 5 illustrates that low frequency signs tend to be produced slightly higher in signing space (and closer to the face) than high frequency signs, and that signs with low probability handshapes tend to be produced slightly higher in signing space (and closer to the face) than signs with high probability handshapes.

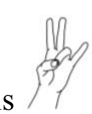
We conducted a linear regression predicting the nose to wrist distance. The critical predictor was an interaction between Sign Frequency and Handshape Probability. We also controlled for Iconicity, Lexical Class, and Sign Type (see Table 1), because these variables are related to the variables of interest and have been shown to affect sign production (e.g., Sehyr Sevcikova & Emmorey, in press). There was a significant interaction between Sign Frequency and Handshape Probability (see Figure 6). Simple slopes analysis probing this interaction indicated that signs with common handshapes were more likely to be produced further from the face than signs with uncommon handshapes, but only among medium and high frequency signs ($t_{MediumFrequency} = 2.89$, $p < 0.01$; $t_{HighFrequency} = 3.52$, $p < 0.001$); there was no effect of handshape probability for low frequency signs ($t_{LowFrequency} = 0.95$, $p = 0.34$). See Supplementary Figure 1 for a more detailed visualization of the interaction. To account for collinearity, we compared the full model to a model excluding the critical interaction between Handshape Probability and Sign Frequency using a log-likelihood test. The full model fit significantly better (AIC = 19,095) than a model excluding the interaction (AIC = 19,098 $p = 0.02$), indicating that the interaction had an independent effect above and beyond the other regressors.

Though not a primary question under investigation here, we also found that one handed signs are produced significantly closer to the nose than the three types of two handed sign signs, and that verbs and nouns were produced significantly closer to the nose than most other lexical classes, and nouns were produced significantly closer to the nose than verbs. Iconicity was not a significant predictor of nose to wrist distance.

5. Discussion

The goal of the research reported here was to bring empirical evidence to bear on three predictions that sign languages undergo diachronic changes whereby signs that are difficult to perceive and/or to produce will be articulated closer to the head of the signer. We applied novel human pose estimation techniques to a large corpus of lexical signs attested in ASL, allowing an assessment of a large number of signs without human location annotation and without assigning a priori regions of interest. All three of the predictions were borne out in the data. We found that among low frequency signs, those with uncommon handshapes tend to be produced closer to the face than signs with common handshapes. In addition, infrequent signs tend to be produced closer to the face than frequent signs. These main effects were qualified by an interaction between frequency and handshape probability, such that the effects of handshape probability were only seen in the most frequent signs and not in less frequent signs.

The finding that frequent signs tend to be produced further from the signer's face than infrequent signs could arise either because of perceptual or articulatory pressure, or both. High frequency signs may be so predictable and easy to perceive that perceivers do not need the added support of seeing them in the more central region of their visual field. Alternatively, the higher a signer has to raise their hand in order to articulate a sign, the less efficient the sign may be to produce. Signs that are produced often may be produced lower to preserve energy. These two possibilities are not mutually exclusive: articulatory and perceptual pressures might conspire together to push high frequency signs downward. We thank an anonymous reviewer for pointing out that research examining the z-axis may help to tease apart these various pressures, as variation in the z-axis could be driven by articulatory demands but is unlikely to be driven by perceptual demands. Further work using 3D pose estimation models will be needed to test this hypothesis.

While lexical frequency effects may be attributed to articulatory and/or perceptual pressures, the effects of handshape probability are more consistent with the idea that sign languages have evolved to maximize efficiency for the perceiver. Uncommon handshapes are less predictable and more difficult to perceive, and so signs with these handshapes are more likely to be produced in the line of sight of the perceiver where the visual system has the most acuity. We presume that common and uncommon handshapes do not systematically differ with respect to how efficient they are for the producer to articulate in various locations on the body (e.g., it is no more difficult to produce an uncommon handshape like this  than a common handshape like  in a distal location), and so we suggest that this pattern cannot be attributed to producer-based communicative efficiencies.

Our interpretation of the fact that effects of handshape probability were only observed in the high frequency signs is that ASL has evolved to preserve communicative efficiency. High frequency signs have a higher mutation rate (i.e., more opportunities to evolve) than infrequent signs, and afford more opportunities for communicative efficiency to shape the lexicon. Another compatible explanation is that high frequency signs are so predictable and easy to perceive that the added support of articulating signs in the perceiver's line of sight is not necessary.

These results align with evidence from spoken languages that vocal articulation space, mapped using the first two formants (F1 and F2), has evolved for communication. For example, words with many phonological neighbors tend to have larger vowel space (i.e., hyperarticulated vowels), perhaps to make it easier for the perceiver to discriminate between confusable words (e.g., Munson & Solomon, 2004, Wright, 2001, though Gahl et a., 2012 find a different pattern). Words that are very frequent tend to have smaller vowel space, perhaps because perceivers do not need clear articulation to correctly recognize the words or because they are highly routinized for the producers. The similarity between dispersion in signing space and dispersion in vowel space is striking, and points to hyperarticulation as a possible modality-general property of language.

One limitation of this study is that the data used are the videos from ASL-LEX, which include only a single exemplar of each sign, and does not reflect within and across signer variation in production. It is not clear whether these results should be taken to reflect diachronic patterns that are encoded in the lexicon, or synchronic patterns in how individuals produce signs. Nevertheless, this study presents a set of techniques that could be applied to answer such questions, though much larger datasets that include a diverse set of signers and productions are needed to help obtain reliable answers.

All of these effects, while significant, were quite small, reflecting a tendency rather than a primary factor driving the structure of sign languages. Effects may be small in part because signers shift their attention covertly to the lower visual field, potentially enhancing sensory processing of stimuli lower in the visual field (Stoll et al. 2018; Stoll & Dye 2019). Additionally, handshape probability and lexical frequency are likely just two of many factors that affect the position of the hands and so only account for a small portion of the variance in hand position.

6. Conclusion

Much of what we know about how languages evolve and are used is based on data from spoken languages. This study illustrates how studying signed languages can offer opportunities to test predictions and to disentangle properties that are language-general from those that are modality-specific. We show here that the perceptual and articulatory demands of manual-visual languages, which are different from that of auditory-oral languages, leave a distinctive imprint on the structure of the lexicon. In other words, languages conform to the bodies of their users.

Ethics and Consent

No human subjects were involved in the conduct of this work. Ethical approval was provided by the Institutional Review Board at Rochester Institute of Technology for the project Multimethod Investigation of Articulatory and Perceptual Constraints on Natural Language Evolution. All images are shared with the consent of those pictured.

Funding

This material is based upon work supported by the National Science Foundation under Grant No. BCS 1749376. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Acknowledgments

We are grateful to Donna Jo Napoli and Ann Senghas for discussions concerning this work.

Competing Interests

The authors have no competing interests to declare.

CRediT Statement

Naomi Caselli: Conceptualization, Methodology, Formal analysis, Investigation, Resources, Writing - Original Draft, Visualization, Funding acquisition **Corrine Occhino**: Writing - Original Draft, Funding acquisition **Bruno Artacho**: Software, Formal analysis, Writing - Review & Editing **Matthew Dye**: Conceptualization, Methodology, Writing - Review & Editing, Supervision, Project administration, Funding acquisition **Andreas Savakis**: Methodology, Resources, Writing - Review & Editing, Project administration, Funding acquisition

References

Agrafiotis, D., Canagarajah, N., Bull, D. & Dye, M. (2003). Perceptually optimised sign language video coding based on eye tracking analysis. *Electronics Letters*, *39*, 1703–1705. DOI: 10.1049/el:20031140

Aylett, M. & Turk, A. (2004). The smooth signal redundancy hypothesis: a functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech*, *47*, 31–56. DOI: 10.1177/00238309040470010201

Battison, R. (1978). *Lexical borrowing in American Sign Language*. Linstok.

Bolinger, D. (1981). *Two kinds of vowels, two kinds of rhythm*. Indiana University Linguistics Club.

Brentari, D. (1998). *A prosodic model of sign language phonology.* MIT Press.

Bybee, J. L. (1998). Emergent lexicon. *Chicago Linguistic Society: The Panels*, 34, 421–35.

Bybee, J. L. (2001). *Phonology and language use. (Cambridge Studies in Linguistics 94)*. Cambridge University Press.

Bybee, J. L. (2015). *Language change.* Cambridge University Press.

Cao, Z., Hidalgo, G., Simon, T., Wei, S., & Sheikh, Y. (2018). OpenPose: realtime multi-person 2D pose estimation using part affinity fields. *arXiv*, 1812.08008v2.

Caselli, N., Sevcikova Sehyr, Z., Cohen-Goldberg, A. & Emmorey, K. (2017). ASL-LEX: a lexical database of American Sign Language. *Behavior Research Methods*, *49*, 784–801. DOI: 10.3758/s13428-016-0742-0

Cheek, A., Cormier, K., Repp, A., & Meier, R. (2001). Prelinguistic gesture predicts mastery and error in the production of early signs. *Language*, *77*, 292-323. DOI: 10.1353/lan.2001.0072

Chomsky, N. (1975). *Reflections on language*. Pantheon.

Clibbens, J., & Harris, M. (1993). Phonological processes and sign language development. In D. Messer, & G. Turner (eds), *Critical influences on child language acquisition and development,* pp. 197–208. Palgrave Macmillan. DOI: 10.1007/978-1-349-22608-5_10

Emmorey, K. (2007). The psycholinguistics of signed and spoken languages: how biology affects processing. In M. Gareth Gaskell (ed.), *The Oxford handbook of psycholinguistics*, pp. 703–721. Oxford University Press. DOI: 10.1093/oxfordhb/9780198568971.013.0043

Emmorey, K., Thompson, R., & Colvin, R. (2009). Eye gaze during comprehension of American Sign Language by native and beginning signers. *The Journal of Deaf Studies and Deaf Education*, *14*, 237–243. DOI: 10.1093/deafed/enn037

Fenlon, J., Cormier, K., & Brentari, D. (2017). The phonology of sign languages. In S. J. Hannahs & A. Bosch (eds), *The Routledge handbook of phonological theory*, pp. 453-475. Routledge.

Fowler, C., & Housum, J. (1987). Talkers' signaling of "new" and "old" words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language*, *25*, 489-504. DOI: 10.1016/0749-596X(87)90136-7

Frishberg, N. (1976). Arbitrariness and iconicity: historical change in American Sign Language. *Language*, *51*, 696-719. DOI: 10.2307/412894

Gibson, E., Futrell, R., Piantadosi, S., Dautriche, I., Mahowald, K., Bergen, L., & Levy, R. (2019). How efficiency shapes human language. *Trends in Cognitive Sciences*, *23*, 389-407. DOI: 10.1016/j.tics.2019.02.003

Gregory, M. L., Raymond, W. D., Bell, A., Fosler-Lussier, E., & Jurafsky, D. (1999). The effects of collocational strength and contextual predictability in lexical production. *Chicago Linguistic Society*, *35*, 151-166.

Hauser, M., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science*, *298*, 1569-1579. DOI: 10.1126/science.298.5598.1569

Hockett, C. D. (1960). The origin of speech. *Scientific American*, *203*, 88-97. DOI: 10.1038/scientificamerican0960-88

Kalman, R. E. (1960). Contributions to the theory of optimal control. *Boletin de la Sociedad Matematica Mexicana*, *5,* 102-119.

Karnopp, L. (2002). Phonology acquisition in Brazilian Sign Language. In G. Morgan & B. Woll (eds), *Directions in sign language acquisition*, pp. 29–53. John Benjamins. DOI: 10.1075/tilar.2.05kar

Lane, H., Boyes-Braem, P., & Bellugi, U. (1976). Preliminaries to a distinctive feature analysis of handshapes in American Sign Language. *Cognitive Psychology*, *8*, 263–289. DOI: 10.1016/0010-0285(76)90027-X

Marentette, P., & Mayberry, R. (2000). Principles for an emerging phonological system: a case study of early ASL acquisition. In C. Chamberlain, J. Morford, & R. Mayberry (eds), *Language acquisition by eye*, pp. 71–90. Lawrence Erlbaum Associates.

Mathur, G., & Rathmann, C.. (2012). Verb agreement. In R. Pfau, M. Steinbach, & B. Woll (eds), *Sign language: an international handbook*, pp. 136–157. De Gruyter Mouton.

Morgan, G., Barrett-Jones, S., & Stoneham, H. (2007). The first signs of language: phonological development in British Sign Language. *Applied Psycholinguistics*, *28*, 3–22. DOI: 10.1017/S0142716407070014

Muir, L., & Richardson, I. (2005). Perception of sign language and its application to visual communications for deaf people. *The Journal of Deaf Studies and Deaf Education*, *10*, 390–401. DOI: 10.1093/deafed/eni037

Munson, B., & Solomon, N. P. (2004). The effect of phonological neighborhood density on vowel articulation. Journal of Speech, Language, and Hearing Research, 47(5), 1048-1058.

Piantadosi, S. T., & Fedorenko, E. (2017). Infinitely productive language can arise from chance under communicative pressure. *Journal of Language Evolution*, *2*, 141-147. DOI: 10.1093/jole/lzw013

Piantadosi, S., Tily, H., & Gibson, E. (2011). Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences*, *108*, 3526–3529. DOI: 10.1073/pnas.1012551108

Rice, K. (2007). Markedness in phonology. In P. de Lacy (ed), *The Cambridge handbook of phonology*, pp. 79-97. Cambridge University Press. DOI: 10.1017/CBO9780511486371.005

Sevcikova Sehyr, Z., Caselli, N., Cohen-Goldberg, A., & Emmorey, K. (2021). The ASL-LEX 2.0 project: a database of lexical and phonological properties for 2,723 signs in American Sign Language. *The Journal of Deaf Studies and Deaf Education*, *26*, 263-277. DOI: 10.1093/deafed/enaa038

Sehyr, Z. S., & Emmorey, K. (in press). The effects of multiple linguistic variables on picture naming in American Sign Language. Behavior Research Methods. https://psyarxiv.com/38qtx/

Siedlecki, T., & Bonvillian, J. (1997). Young children's acquisition of the handshape aspect of American Sign Language signs: parental report findings. *Applied Psycholinguistics*, *18*, 17–39. DOI: 10.1017/S0142716400009851

Siple, P. (1978). Visual constraints for sign language communication. *Sign Language Studies*, *19*, 95–110. DOI: 10.1353/sls.1978.0010

Stoll, C., & Dye, M. W. G. (2019). Sign language experience redistributes attentional resources to the inferior visual field. *Cognition*, *191*, e103957. DOI: 10.1016/j.cognition.2019.04.026

Stoll, C., Palluel-Germain, R., Gueriot, F.-X., Chiquet, C., Pascalis, O., & Aptel, F. (2018). Visual field plasticity in hearing users of sign language. *Vision Research*, *153*, 105-110. DOI: 10.1016/j.visres.2018.08.003

Takkinen, R. (2003). Variation of handshape features in the acquisition process. In A. Baker, B. van den Bogaerde, & O. Crasborn (eds), *Crosslinguistic perspectives in sign language research*, pp. 81-91. Signum.

Tukey, J. (1977). *Exploratory data analysis*. Addison-Wesley.

von Tetzchner, S. (1984). First signs acquired by a Norwegian deaf child with hearing parents. *Sign Language Studies*, *44*, 225-257. DOI: 10.1353/sls.1984.0007

Wei, S., Ramakrishna, V., Kanade, T., & Sheikh, Y. (2016). Convolutional pose machines. *arXiv*, 1602.00134

Wilbur, R. (1987). *American Sign Language: linguistic and applied dimensions*. College Hill.

Woll, B. (1987). Historical and comparative aspects of BSL. In J. Kyle (ed.), *Sign and school*, pp. 12-34. Multilingual Matters.

Wright, R. (2004). Factors of lexical competition in vowel articulation. Papers in laboratory phonology VI, 75-87.

Zipf, G. (1949). *Human behavior and the principle of least effort: an introduction to human ecology*. Ravenio Books.

Tables & Figures

Table 1. A linear regression predicting nose to wrist distance. The baseline level of Lexical Class was [Verb], and the baseline Sign Type was [One Handed].

	Nose to Wrist Distance		
Predictors	Estimates	CI	p
(intercept)	218.79	215.42 – 222.15	<0.001
Handshape Probability	2.23	0.70 – 3.75	0.004
Sign Frequency	4.41	2.03 – 6.79	<0.001
Iconicity	-0.96	-2.98 – 1.06	0.350
Lexical Class [Adjective]	7.86	2.60 – 13.11	0.003
Lexical Class [Adverb]	14.18	2.85 – 25.51	0.014
Lexical Class [Minor]	17.22	9.55 – 24.89	<0.001
Lexical Class [Name]	7.16	-5.23 – 19.55	0.257
Lexical Class [Noun]	-3.77	-7.22 – -0.32	0.032
Lexical Class [Number]	13.26	-2.79 – 29.31	0.105
Sign Type [Asymmetrical Different Handshape]	8.47	4.20 – 12.75	<0.001
Sign Type [Asymmetrical Same Handshape]	7.17	0.60 – 13.75	0.033
Sign Type [Symmetrical or Alternating]	17.85	14.40 – 21.31	<0.001
Handshape Probability : Sign Frequency	2.55	0.32 – 4.78	0.025
Observations	2613		
R2/R2 Adjusted	0.069/0.064		

Figure 1. The ASL sign book. Information about the sign and it's lexical properties can be found at https://asl-lex.org/visualization/?sign=book.

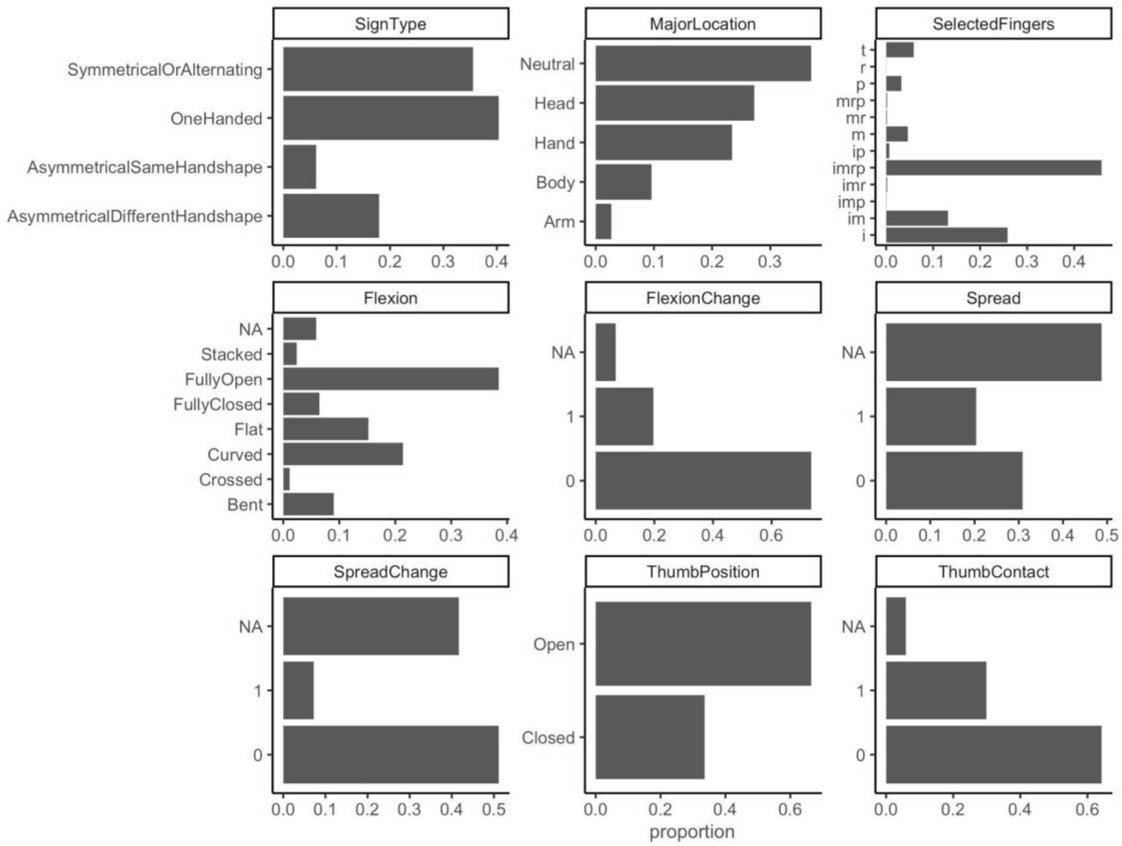


Figure 2. The distribution of phonological features. The facets correspond to each of the nine phonological features used in this study, the y-axis indicates the possible values for each feature, and the x-axis indicates the proportion of signs with each value.

Figure 3. Processing of pose estimation. The original detection from OpenPose is shown in the left image for the nose (green), left wrist (red), and right wrist (yellow) for a single sign. The center image shows the detections after normalization of the coordinates to the nose and application of a median filter, removing some of the outliers. The image on the right shows the final pose estimation after applying the Kalman filter, improving the detection of the trajectory for the wrists, and only accounting for detection during the act of signing. Note that, in the center and right images, the joint locations have been translated in the 2D plane such that wrist locations are all relative to a single estimate of the nose location. A video showing wrist and nose pose estimations superimposed upon the ASL sign tree is provided in Supplementary Materials online.

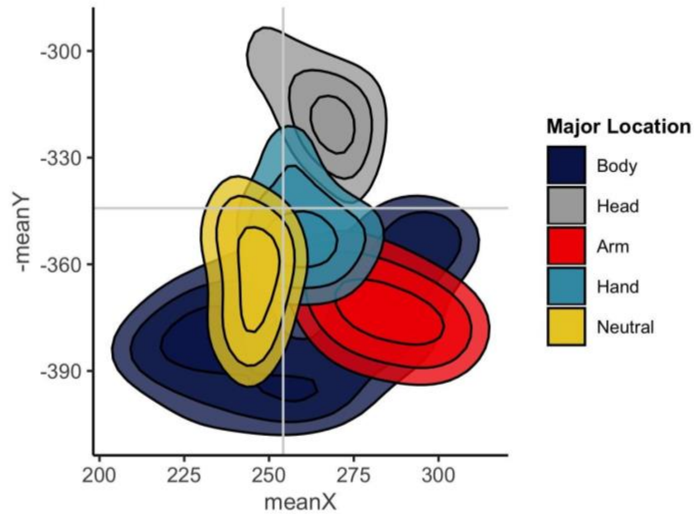


Figure 4. A validity check using 2D density plots to illustrate the distribution of signs in x-y space according to major location as coded in ASL-LEX. To minimize overlap, plots only include contours that enclose the highest density regions (75%, 85%, and 95%). The grey lines indicate the average x and y coordinates for all signs in the dataset.

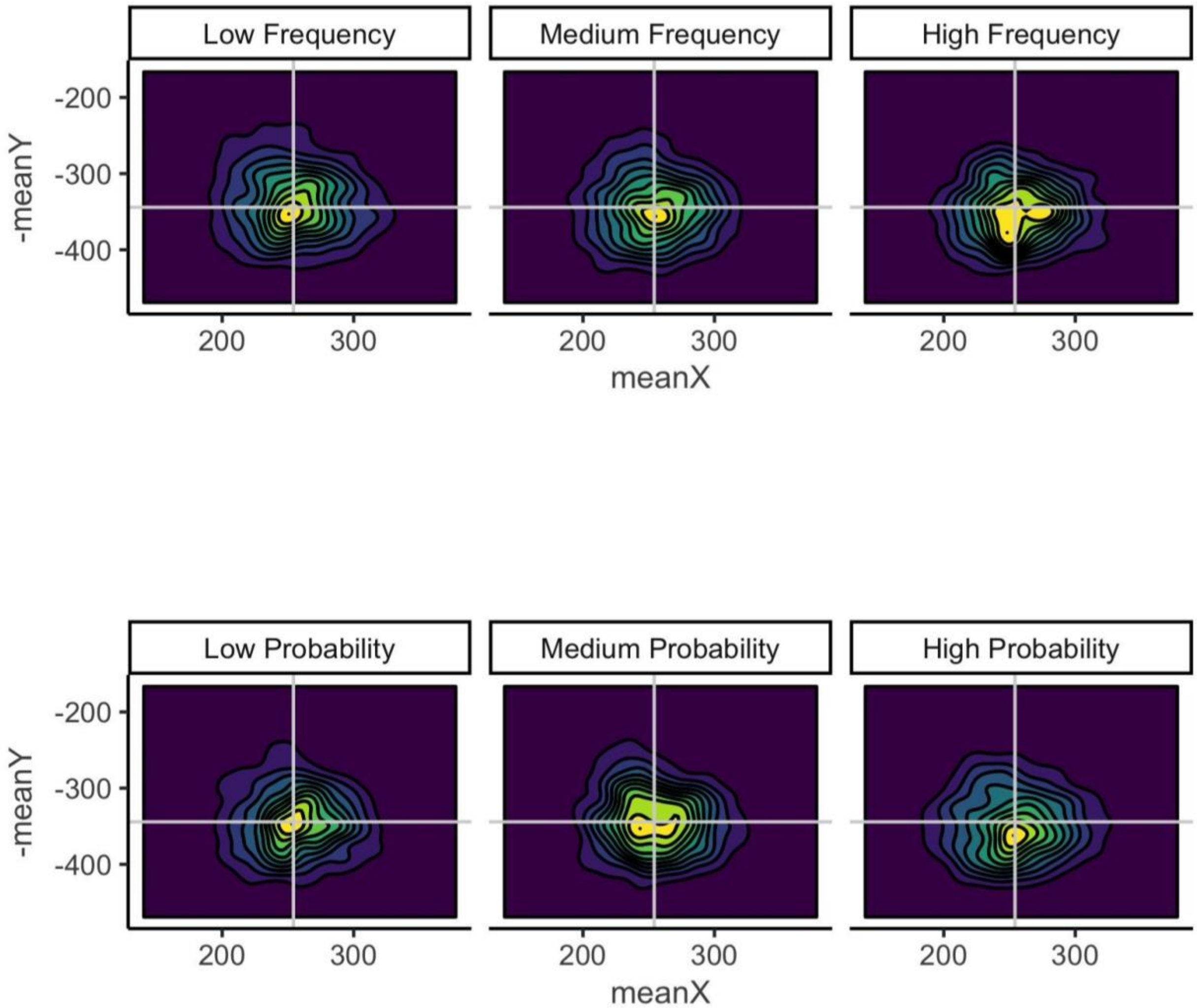


Figure 5. 2D kernel density plots illustrating distribution of signs in x-y space according to Sign Frequency (top row) and Handshape Probability (bottom row). The grey lines indicate the average x and y coordinates for all signs in the dataset. Each contour encloses an incremental 10% of the data. The higher a sign's frequency, the farther from the signer's nose it is articulated in the signing space. Similarly, signs that have higher probabilities are also articulated farther from the signer's nose.

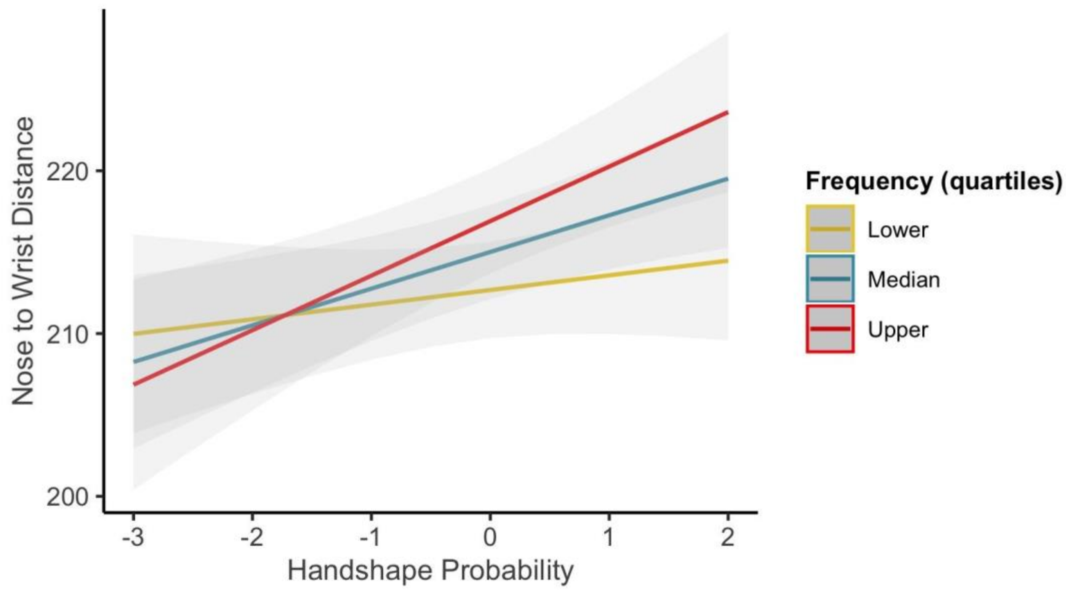
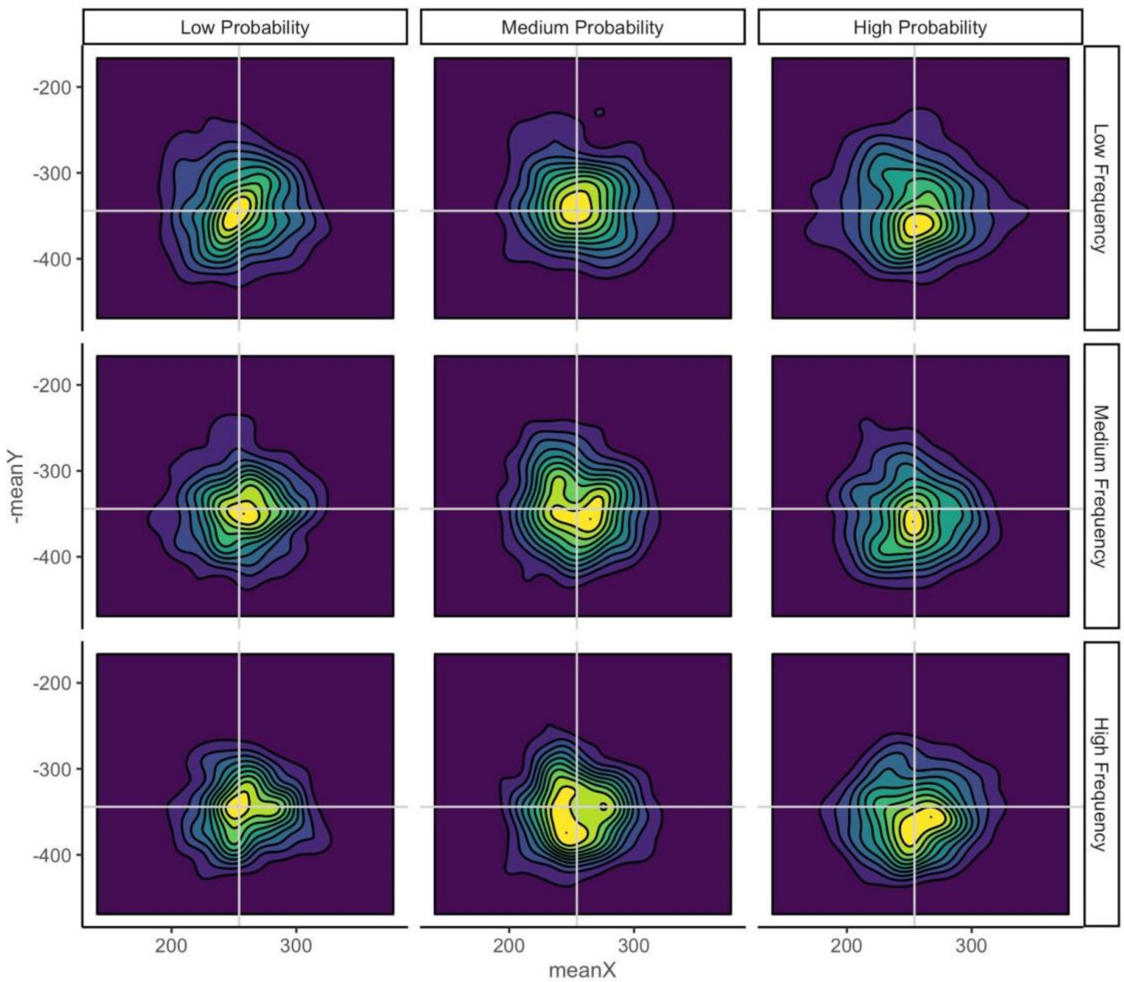


Figure 6. The interaction between handshape probability and frequency on nose to wrist distance. For high and medium frequency signs (red and blue lines respectively), lower probability handshapes were articulated closer to the signer's nose and higher probability handshapes were articulated farther away. This effect was not statistically significant for low frequency signs (yellow line).

Supplementary Material



Supplementary Figure 1. For low frequency signs (top row), handshape probability did not have a statistically significant effect on how far away a sign was articulated from the signer's nose. However, for medium frequency (middle row) and high frequency (bottom row) signs, as handshape probability increases (left-to-right) the sign was articulated farther away from the signer's nose (the typical location for interlocutor eye gaze fixation).

Supplementary Video

Supplementary Video 1. This video shows the ASL-LEX entry for the ASL sign tree (https://asl-lex.org/visualization/?sign=tree). The x-y locations of the 'joints' of interest are marked in green (nose) and red (right wrist).