THE UNIVERSITY OF CHICAGO

THE USE OF ROLE SHIFT WITH CLASSIFIER CONSTRUCTIONS IN AMERICAN SIGN LANGUAGE

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

DEPARTMENT OF PSYCHOLOGY

BY

JAMES MARZUOLA WALLER

CHICAGO, ILLINOIS

DECEMBER 2021

Table of Contents

Acknowledgements

This dissertation would not have been possible without the support of many, many people.

Thank you to my committee, Susan Goldin-Meadow, Diane Brentari, Alex Shaw, and Terra Edwards. To Susan Goldin-Meadow, thank you for all you have done these past five years, as an advisor and mentor. You taught me how to grapple with my ideas and arguments – how to sharpen them and untangle them. From you I learned how one can harness an intuitive fascination with language into tangible questions and project. And of course, you were always reminding me to pull myself out of the weeds, and look up (well, I'm still working on that one). Thank you for the years I've had in your lab.

To Diane Brentari, thank you for welcoming me into the world of sign linguistics. I arrived on campus in Psychology, a bit unsure (wary even) of the Linguistics department across the quad. But your door was always open, and as I write my dissertation, I am still realizing that so much of the foundation of my work is rooted in our many discussions, stretching back to the first year. You were an ally and mentor to me in many ways on campus, through all the years I've been here. Thank you.

Thank you, Alex Shaw. Your contributions were crucial to get me to really think through my study design. I was always amazed how you could piece my study design together after I'd manage to confuse everyone else, including myself.

To Terra Edwards. Your enthusiasm for your work and ideas is infectious. Every conversation we have has left me energized and ready to dive back into my work. Thank you.

Thank you to the researchers who provided feedback on various versions of the work presented in this dissertation. Pamela Perniss, in our few exchanges over the years, has left a lasting impact on the way I approach my work. My many discussions with Ryan Lepic pushed me more than once to rethink, revisit, and revise – and my work is stronger for it. But also, thank you for the stream of texts, memes, jokes, that remind me that linguistics is fun and weird and messy – and to simply embrace that, and to not take life too seriously.

Thank you to all the people of the Goldin-Meadow Lab who listened patiently to my presentations and talks, and my midday ramblings and mutterings. Thank you to my office-mates, Jenny Lu, Aron Marie, Yağmur Deniz Kısa, Casey Ferrara, and Claire Bergey for the constant discussions (sometimes too constant!), and tea, that flowed in and out of Green 517. It will always be a space I hold dear. Thank you to all the linguists – especially to Daniel Lam, Aurora Martinez Del Rio, and Mina Giannoula, who welcomed me with open arms.

Thank you to all of the people helped me create my experiment, collect data, and code the videos. Thank you to my many wonderful models, who patiently worked with the sometimes odd, always oddly specific, videos and pictures for two Studies: thank you to Joshua Falk, Lingwei Cheng, Kathryn Montemurro, Jenny Lu, Casey Ferrara, Brigid Roy, and Peter Waller. Thank you to the people who piloted my study and talked it over with me – Paige, Casey Johnson-Pasqua, Jacqui, Colin Lualdi. A huge thank you here to Zena Levan for all her help with so many stages of the project, from stimuli to collection to coding. And thank you to all the people in the Deaf community across the United States who participated in my study, who gave me their time and without whom this dissertation would simply not exist.

Thank you to the people of the Psychology department and Green Hall – Mimi Smiley, Ray Weathers, and Kristi Schonwald. Thank you to Mimi Maduff. Thank you also to the many interpreters I've worked with, particularly Rivka Hozinsky, Tramon Biciunas, and Amanda Grazian. I particularly want to thank those who were instrumental in the final push out the door –

Kristi Schonwald and Tramon Biciunas, your patience and determination made of world of difference.

Dissertations, and PhDs, can be long, winding, lonely, roads – but it's quite a bit less lonely with the right people there. And when I say there, I mean it literally – a huge reason I made it through this dissertation and PhD is because of how people have welcomed me into their homes during my many crunches – or popped up at my doorstep themselves.

In Chicago, thank you, thank you, to my sister Leona Waller, to Aron Marie, Casey Ferrara, Jenny Lu, Cat Harder, to my father Peter Waller, for coming through into my apartment and keeping me company in the journey. Conversely, thank you to those who let me into their homes (sometimes the same people) – in Chicago, thank you to Yağmur Deniz Kısa, Claire Bergey, Casey Ferrara, Jenny Lu, Natalie Dowling (and the Regenstein Library, and Charlie's Chicago), for opening your homes to me. In Rochester, thank you to Matthew Seita and Brennan Terhune-Cotter. In California, thank you to Brigid and David Roy, and thank you to Yoon. In Oklahoma, I thank my grandmother Gloria Marzuola and aunt Carol Marzuola – who's love keeps me grounded in the important things. In Washington D.C., thank you to Casey Johnson-Pasqua, Paige Watson, and Jacqui Wunderlich – who created a space that eight years later still feels like home to me. And of course, thank you to my parents, Susi Marzuola and Peter Waller for the many long months of the pandemic we weathered together.

All of these people – they made sure I moved forward, and ate, and slept, and kept my feet on the ground and my eyes open to the world around me, the whole darn way. I am so blessed.

To Jenny – it has been a long ride and we made it through. I can't even imagine my PhD without you – you were there for me every step of the way and I feel so lucky to have had your footsteps to guide me. From my very first visit, you always gave me a space in Chicago to be myself. Thank you so much. To Aron, thank you for stretching my mind and leaping with me through down long rabbit holes on language and life. You always had an incredible intuition for what others needed, and how to help. Thank you. To Laura Horton, for nights of laughter and mentorship, and wonderful gifs.

To Casey, Claire, Natalie, and Yağmur – I have many wonderful memories with you all from over the years. Today I want to thank you for these last few weeks, your patience, understanding, and support as I rocketed out of Chicago. You all make it so hard to leave. To Kathy, for her constant support.

To Brigid Roy – in the longest days of the pandemic, it was on our walks and in our conversation that I really realized, I was going to finish this darn thing.

To my family – I did it! Thank you to my mom, Susi Marzuola, for your patience. And thank you, for teaching me to be relentless in crafting the best works, and always reach for the big picture. To my dad, Peter Waller, who will get down in the trenches, with me again. To my sister, Leona Waller, who reminds me to really take care of myself – in the easy ways and the hard ways.

This dissertation is based on work supported by the National Science Foundation Graduate Research Fellowship Program (Grant No. DGE 1746045). 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.

Abstract

In my dissertation I examine the interaction between character role shift and the verb system of American Sign Language (ASL). Character role shift, also often referred to as constructed action, refers to the use of a signer's own body to enact the perspective, behavior, and/or experience of a character in their narrative – frequently explicitly marked by the use of non-manual markers such as enactment of the character's facial expression, or by shifting one's eye gaze or torso.

Two studies using data elicited from deaf ASL signers reveal several main findings. First, the data supports previous claims that some categories of verbs, specifically classifier constructions, are more amenable to role shift than other kinds of verbs. Secondly, different elements associated with character role shift display different distributions across classifier construction subtypes and contexts. For example, signers are quite particular in their use of facial enactment – using it often with handling classifier constructions, but less with entity classifier constructions and static classifier constructions. On the other hand, use of eye gaze shift is generally robust regardless of type of classifier construction.

On the basis of these findings and follow-up analyses, I offer a theoretical analysis of role shift. Broadly, my proposal contends that there are at least two distinct types of role shift that can be distinguished based on form and also structural requirements – each of these types of role shift can be distinguished by the way they each interact with the verbal system of American Sign Language. I end with observations on how these findings give insight into the various ways that iconicity structures communication.

List of Figures

List of Tables

1 Introduction

We spend a lot of time talking about other people – what they did, what they said, how they said it, what they were really thinking. Interlocutors can report the actions, emotions, and words of others. Or even recreate others as characters in their conversation, using their own body, hands, and voice to enact the character's own behavior and words. This dissertation essentially examines, how, when, and where this strategy is used in American Sign Language discourse.

Figure 1 shows a snapshot of a video in which a woman takes a tissue out of the box. In Figure 2, a signer describes the scene shown in Figure 1. In the original video, which is only a few seconds long, the tissue appears stuck and slips from the woman's fingers when she tries to pull it out, so she has to grab at the tissue multiple times to get it out. To show this, the signer makes a repeated grasping motion – as if grabbing for the tissue herself. This is an example of a handling classifier construction. It bears strong resemblance to an iconic gesture that a non-

Figure 1. Scene of woman getting a tissue from a box.

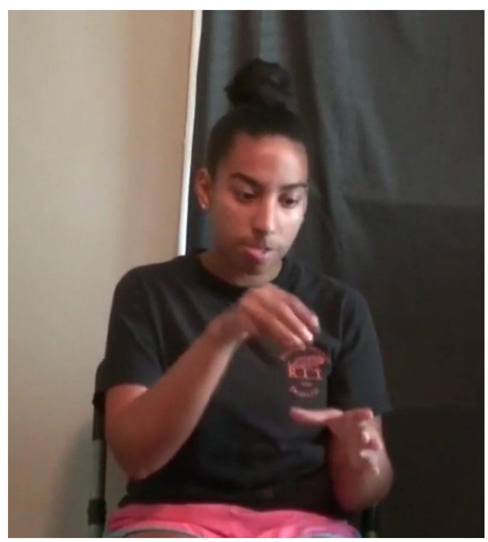


Figure 2. Handling classifier construction. Woman describes scene in Figure 1 (left) using a handling classifier construction on RH. American Sign Language.

Figure 3. Sign SEARCH with role shift. Signer produces sign SEARCH with facial enactment. American Sign Language.

signer might produce while speaking, at least at first glance. To complete the enactment, the signer's whole upper body is brought into the act: she crouches forward over the imagined tissue box, looking down on it and grimaces in mock frustration.

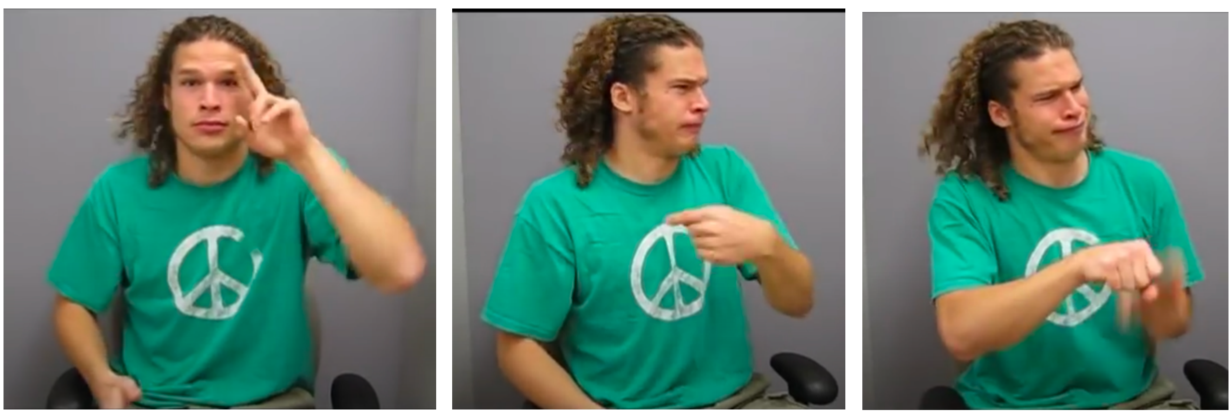
In a second example in Figure 3, a signer produces the sign SEARCH – a handshape in the shape of the letter C, shaken slightly in front of the face, in a description of a man searching through a book to find something, shown in Figure 3. The signer frowns as she looks down towards her non-dominant hand, as if searching through the book like the man in the video.

In both Figure 2 and 3, the signers are producing signs that describe an action being performed by a character. At the same time, they use their face, posture, and eye gaze to re-enact the behavior of the character. These are clear examples of character role shift, the phenomenon I investigate in this paper.

1.1 History of character role shift

Early research on character role shift took as a starting point the use of character role shift in contexts where signers were reporting conversations between multiple people, and the ‘shifts’ between the different interlocutors were marked in several ways including character enactment (Mandel, 1977).

To clarify what I mean, I give another example. In a story titled “Scary Chicken Story” told by Dack Virnig, a Deaf storyteller (2011)[1], Virnig recounts a conversation between a young



UNCLE PRO-1 CHOP++

‘ The uncle replied, “I’m chopping…”

Figure 4 (top row) and *Figure 5* (left). Example of reported dialogue with character role shift. Uncle’s statement and niece’s reply. American Sign Language. Reproduced from Virnig (2011).

- “Oh, I see”, replied the girl.

[1] Virnig, Dack (2011). *Scary Chicken Story* [Video]. YouTube. https://www.youtube.com/watch?v=xFbeQh0iuW0

girl and her uncle. The start of a quote from the dialogue is glossed in (1), and partially reproduced in Figure 4:

(1) UNCLE <PRO-1 CHOP HEAD…> rs_{uncle}
The uncle replied, "Oh that, I chop the [chicken] heads then …"
Note: PRO-1 is 1st person pronoun ("I" or "me" in English). The bracketed part is produced under role shift (in the role of the uncle character)

The example in (1) can be broken in two parts: the signer introduces the speaker of the quote UNCLE, followed by the actual quote itself. The signer shifts into the role of the uncle during the quote; looking at the pictures in Figure 4, there is a clear contrast between his body when signing UNCLE and when he starts the quote (PRO-1 and CHOP – "I chop"). The signer shifts his torso to his right and eye gaze to the left, as if he is looking towards the "niece" as he talks to her; his face assumes a 'grumpy' expression. The stretch of role shift is marked by the brackets labeled rs_{uncle}.

A moment later, the signer gives the niece's brief response: "Oh ok". The signer shifts to his left, looking upwards back to where the 'uncle' was a moment ago – he is now in the role of the niece.

In this example, we see instances of enactment, where the body is recruited to show aspects of the character's actions and experience: in Figure 5 (below Figure 4), from the signer's facial expression we can tell the niece is mildly bemused by her uncle's behavior and response.

We can also observe that role shifts can map onto the structure of the discourse. In the start of example (1), we see the signer starts out as the narrator – looking straight ahead at his 'audience'. Then he jumps into the role of the uncle, as he quotes his words, and then again jumps role to that of the niece, to give her response. When he's done with the conversation, he could then turn back to the camera and the audience behind it – letting us know he is 'out of character' and we are back to narrating the story. In this way the 'role shifts' track with the discourse, and whose voice or actions are being represented.

Many of the key elements of character role shift that formed the basis of early characterizations of the phenomenon are present in this example. First of all, role shift often involves overt shifts of the torso, which mark that the signer has changed 'role' – from one character to another, or from the narrator role to the character. Early work recognized that this torso shift, also termed 'body shift', is not truly character enactment per se, but rather a strategy for indexing which character is being enacted.

This distinction between using the body as an index to mark which character you're talking about, and using the body as iconic enactment, to describe the character is elegant, and I will return to it many times. In practice the distinction can be muddied when it comes to actual functions. Nonetheless, early researchers noted that body shift and enactment are overlapping phenomena, but not necessarily the same phenomena (Loew, 1984).

Early work on role shift often discussed the phenomenon in terms of theatre analogies and focused on long, elaborated examples of role shift – or 'role play'. This conversation started to shift in the late 1970's and 1980's. Padden (1986) made several observations about role shift that form the basis for much of our understanding of role shift today. First, she emphasizes that role shift can appear in short fragments integrated into sentences. As others note, these fragments could appear with some, all, or none of the markers associated with role shift to a character, including a break in eye gaze; enactment of the character's facial expression, posture, and bodily actions; a shift in the location of the torso, or more subtly, a tilt of the head to one side (Cormier et al., 2015).

Second, she observes that sometimes when role shift appears across a clause, its interpretation is ambiguous. Depending on context, the signs produced with role shift might not be interpreted as actual words said by the character – instead it could be interpreted as a description of the character's actions. When the words produced with role shift are a quotation of a character, we call this quotative role shift. When the words produced with role shift are a description of a person's actions rather than something they said, this is non-quotative role shift. Figure 2-3, in which signers use role shift with descriptions of actions, are examples of non-quotative role shift. Figure 4-5 is an example of quotative role shift.

Finally, Padden observes that certain kinds of verbs are particularly suited to non-quotative role shift. This is the line of research this dissertation is situated in. But before we can discuss how the use of role shift patterns across verb categories, we must begin with a discussion of the verb system in ASL.

1.2 Verb system in American Sign Language

American Sign Language, as is true for sign languages across the world, is known to have morphologically complex verbs and verbal constructions Moreover, not all verbs exhibit uniform behavior, and we can classify verbs by their internal structure (Padden, 1990).

Plain Verbs

Plain verbs do not display morphological inflection or agreement with subject or object, and are a single uninflected unit. Plain verbs like HAVE, EAT, and DRINK do not vary in their form based on the context they appear in and do not display any overt morphological inflection (Padden, 1990).

Classifier constructions

Classifier constructions, also variously referred to as depicting verbs/constructions and polycomponential signs, refer to a category of verbs typically used in spatial depictions or descriptions of events and scenes, and display some unique structural properties. Classifier constructions can generally be decomposed into two components: the handshape of the verb and the movement/location of the verb. Each of these two components makes an independent contribution to the meaning of the verb.

Classifier constructions are typically subdivided into several types. See Zwitserlood (2012) for a comprehensive overview of classifier constructions.

- *Entity classifier constructions (of motion)* describe how an entity moves through space: this entity can be a person, an animal, a vehicle, or an inanimate object such as a pen falling off a table. The choice of handshape is based on the type of entity, and the movement of the sign iconically matches the movement of the entity.
- *Handling classifier constructions* describe how a person handles an object. The signer's handshape indicates the type of object handled and how it was handled, and the signer's hand movement iconically matches the movement of the handling action.
- *Body part classifier constructions* describe how a person moved their body – head, eyes, legs, etc. The handshape is chosen based on the body part and the movement of the sign matches the movement of the body part.
- *Tracing classifier constructions* describe the shape of an object – the hand movement of the sign 'traces' the outline of the object. The handshape chosen is often an outstretched index finger, pointing to the outline.

Finally, researchers often mention the use of entity classifier constructions that describe the location of an entity rather than its motion. I separate location entity classifier constructions from entity classifier constructions of motion.

- *Location entity classifier constructions* describe the location of an object in a scene by 'placing' it in a corresponding position in the signing space. The handshape chosen is based on the type of entity. The sign movement is a short, punctuated movement with a final hold in the location indicating the real-world location of the object.

Classifier constructions have received much attention in the literature, often in the context of determining their status within the 'lexicon'. Classifier constructions clearly form part syntactic clauses, and the handshape category of the classifier construction corresponds to verb class: Benedicto and Brentari (2004) analyze how, in classifier constructions, the type of classifier handshape used predicts whether the classifier verb is transitive (handling classifiers) or intransitive (entity classifiers).

From a phonological perspective, classifier constructions behave systematically differently from plain lexical verbs. Rather than drawing from a discrete set of movements and locations, classifier constructions use the signing space to re-enact the motion of the event and the relative location of the entities.

Researchers have modeled the unique position of classifier constructions in several ways. Under Brentari and Padden (2001)'s model of the lexicon, there exist 'core' components of the lexicon and 'peripheral' components. In the peripheral components, the phonological constraints obeyed by the core lexicon may be weakened in specific ways. Brentari and Padden (2001) place classifier constructions in a peripheral component of the lexicon (the 'spatial lexicon'), indicating that to accommodate the visuo-spatial nature of classifier constructions, some core phonological constraints are weakened or do not apply to classifier constructions.

Classifier constructions, according to Davidson (2015), are unique because of their demonstrational component. That is, classifier constructions, through their iconic use of movement and location, appear to iconically demonstrate aspects of the event they refer to.

Agreement verbs and spatial verbs

Finally, there are two other categories of verbs that I do not discuss in depth. Agreement verbs are verbs that inflect for person and number agreement. Some agreement verbs can inflect for both the subject and the object of the verb; others only inflect for the object (Padden, 1990). Spatial verbs are superficially similar, except that they inflect for locative agreement.

1.3 Constraining character role shift

The use of character role shift (excluding cases of quotation) is constrained in terms of which kinds of verbs it can appear with. Lillo-Martin (1995/2013) and Davidson (2015) note that, for ASL, character role shift does not tend to co-occur with either nouns or pronouns, and occurs just with verbs. Moreover, Davidson also observes that particular kinds of verbs, classifier constructions, are best suited to appear under character role shift.

Davidson (2015) argues that body shift turns a verb into a predicate of demonstration (or depiction) that is produced simultaneously with the demonstration itself. For this alignment to happen, it is most effective/grammatical if the verb already contains an aspect of demonstration – as in the case of (motion) classifier constructions (Davidson analyzes the movement of these signs as demonstrations of actual movement). However, it has long been observed that these classifier constructions are not the only kinds of verbs to appear under role shift: plain lexical verbs occasionally do as well (see Padden, 1986; Engberg-Pedersen, 2003). SEARCH in Figure 3 would be one such example.

But it appears not all lexical verbs allow role shift. Schlenker (2017) provides two examples in which a signer describes how a person shows another person their phone. In one example, a classifier construction is used, in which the action looks similar to actually holding up a phone. In contrast, the second example uses a lexical verb for SHOW, in which the signer points their right finger to the palm of their left hand. The first example with the classifier construction is judged as acceptable[2] under character role shift by native signers. The second example is not. Davidson (2015) and Schlenker (2017) both claim that the distinction between verbs that allow character role shift and those that do not is attributable to whether the verb can be interpreted iconically; that is, these verbs have forms that resemble their meaning in some way. More iconic verbs are more amenable to body shift and character role shift generally. These researchers have begun to develop formal frameworks that make the role of iconicity explicit. However, more research in this field is needed – it is still unclear how signers use iconicity, the similarity between form and meaning, to decide if a verb can be produced under body shift.

In sum, nouns and pronouns do not appear to allow character role shift. Some verbs, namely classifier constructions, are amenable to character role shift. Finally, verbs that are not decomposable into a handshape and an iconic movement component seem to allow character role shift as a function of their degree of iconicity: more iconic verbs are more likely to co-occur with character role shift.

[2] In a standard acceptability judgement task using a 1-7 Likert scale (often used in grammaticality judgments).

In my first study, my goal is to take a broad view of the phenomenon, looking at the use of character role shift across both plain verbs and classifier constructions.

2 Study I: Character Role Shift across the Lexicon

The first study takes a broad look at character role shift in American Sign Language. I use video and picture stimuli to elicit descriptions that are likely to contain instances of character role shift in Deaf signers. I independently analyze each of three nonmanual markers (eye-gaze, facial expression, body shift) to understand whether they converge or diverge in their patterns of use. This study examines use of character role shift across, and within, two verb categories: classifier constructions and plain verbs.

Following the claim that Davidson (2015) and others have made that role shift is more amenable to predicates that are iconic, we would expect classifier constructions to appear with markers of character role shift more often than plain verbs. We would also expect variation in use of character role shift within verb categories to be explained by iconicity, particularly within the category of plain verbs.

2.1 Methods

Participants

Data were collected from 12 Deaf signers. All participants were exposed to American Sign Language by the age of 5 and use ASL everyday as their primary language.

Materials

A series of stimuli were presented on a computer screen. Each stimulus was either a video of about 5-10 seconds long, or a picture. Each stimulus showed one actor acting on some inanimate item (the 'patient'). Examples include a man breaking a stick into three parts, or a woman erasing a message on a whiteboard. The full list of stimuli is given in Appendix A.

The intended goal of the stimuli was to elicit utterances that contained noun phrases labeling the actor and the inanimate patient, as well as a transitive verb that described how the actor acted on the patient. Pictures and videos were used to elicit a wider variety of verbs. Videos were likely to elicit more classifier constructions because the actual behaviors of the actor are shown. In contrast, pictures were more likely to elicit plain lexical verbs because the details of the movement were not shown in the picture (although, in some cases, they were inferable from the picture).

In this way, we could collect relatively short utterances that contained nouns referring to different semantic roles – actor and patient (which closely map onto syntactic subject and object), as well as different kinds of verbs. Also, by keeping the utterances short and limited to only a few verbs, utterance-level analyses, particularly of relationships between verbs, is more feasible.

Procedure

Participants were instructed in ASL to watch each video then "give a short description of what happened". Participants saw the stimuli in one of three randomized orders, although the videos preceded the pictures for all participants. To increase the likelihood that participants used classifier constructions in description of the videos, we placed the pictures at the end of the study after the video. Since we expected pictures to elicit plain, monomorphemic, lexical verbs, we did not want to prime participants to use plain lexical verbs instead of classifier constructions for the videos as well.

2.2 Coding system

Two grammatical classes were coded: verbs and nouns. Each recorded sign was coded for the three types of non-manual markers of constructed action: eye gaze shift, changes in facial expression, and body shift. All coding is done using ELAN (Crasborn & Sloetjes, 2008)[1].

[1] ELAN (2021). http//:archive.mpi.nl/tla/elan

We classify all verbs[2] into one of several categories, based on the type of handshape used, and the movement of the sign. All verbs are also coded for co-occurrence with non-manual markers of character role shift. For nouns, only nouns labeling the actor or the patient were coded for use of character role shift.

Coding verb type

We categorized verbs on the basis of their form into one of several categories. To be assigned a category the verb must have a handshape and movement/location consistent with the given category.

- Handling classifier construction: the handshape of the sign matches the handshape of the character in the scene during the described event while handling the object. For this study, the movement of the sign must be an iconic match with the movement of the character's hands while handling an object.
- Entity classifier construction of motion: the handshape of the sign matches the classifier handshape for an entity in the scene. For this study, the movement of the sign must be an iconic match with the movement of the object during the scene.
- Body part classifier construction (verb of motion): the handshape of the sign matches the classifier handshape for a body part of the character in the scene (e.g., their eyes, or legs). The movement of the sign must be an iconic match with the movement of the object during the scene.
- Instrument classifier construction (verb of motion): the handshape of the sign matches the classifier handshape for an entity used as an instrument by the character. The movement

[2] With one exception: we exclude verbs that described the appearance of the actor (e.g., <has long hair>, description of clothing) which were a priori deemed not relevant to the current question.

of the sign must be an iconic match with the movement of the hand and instrument during the scene.

- Location entity classifier constructions: the handshape of the sign matches the classifier handshape for the entity in the scene. The movement of the sign is a short, punctuated movement ending at a location in the sign space, indicating the object was located in the corresponding location in the scene.
- Tracing classifier constructions: the handshape of the sign is shaped as if wrapped around the exterior of an object, or alternatively is tracing the exterior of an object with one or several fingers. The movement of the sign traces the shape of the object's exterior.
- Plain verb: the handshape, movement, and location are consistent with the citation form of an established plain, monomorphemic, lexical verb without inflection or modification. These will be referred to simply as 'plain verbs'.
- Noun: the handshape, movement, and location are consistent with the citation form of an established noun. Rather than code every instance of nouns, we only coded those that refer to the character in the scene (e.g., MAN, WOMAN) or the object handled/manipulated (e.g., BOOK, SPOON)

If a verb was consistent with multiple categories, it was classified as *ambiguous* and not used in the main analysis. A verb was classified as *unclassifiable* if they did not fit into any of our categories[3]. Since our data are mostly of people handling objects, the most common classifier

[3] A verb may be coded as unclassifiable for one of a few possible reasons. A verb might have a handshape and movement similar to a plain verb, but the movement is clearly embellished or modified to map on the action – so it no longer is a citation form. Also, a verb with any kind of inflection or agreement, such as aspectually inflected verbs or person-agreement verbs, would also be coded as unclassifiable.

constructions are expected to be handling classifier constructions. There may also be frequent use of plain verbs to describe the actions.

Coding non-manual markers of character role shift

As explained in the introduction, character role shift can be identified through three related but conceptually distinct markers: eye gaze shift, facial enactment changes, and body shift. We code each of these separately — which becomes relevant since they turn out to behave differently.

- Eye gaze shift coding*:* A shift in eye-gaze was coded if the signer broke eye-gaze with the camera and shifted their gaze towards the signing space and/or in the direction that the actor in the scene would have been looking[4].
- Facial enactment coding*:* If the signer changed their facial expression to assume a facial expression or a head movement that the actor in the video performed, this behavior was coded as facial enactment. The enactment does not need to be emotive; mouth actions such as pursed lips to indicate drinking from straw, are also coded.
- Body shift coding*:* If the signer shifted their torso and/or head from the signer's natural sitting position to a location that plausibly could be interpreted as the locus of the actor, this shift would be coded as a body/head shift[5]. 8 This shift is not an act that the actor in the scene actually performed; rather, it is a shift that simply places the signer in a new

[4] The actor's eye gaze is often directed towards an object, which the signer had established already in signing space, so it is often impossible to distinguish whether the signer is choosing to assume the actor's eye gaze or directing their attention to the objects shown in the space (see Engberg-Pedersen, 2003).

[5] Previous work distinguishes head and torso shifts (e.g., Hermann and Steinbach, 2012), but that distinction could not be coded reliably in our pilot data and was combined in this study.

'position' during a stretch of discourse[6].

2.3 Results

For our main analysis, we examined the following verb categories: handling classifier constructions, tracing classifier constructions, location classifier constructions, and plain verbs. Our analysis contains 447 verbs. 40% of the verbs that participants produced were plain verbs and 60% were classifiers (47% handling classifiers, 5% tracing classifiers, 7% locating classifiers).

We do not analyze body part classifier constructions, entity classifier constructions, verbs that are ambiguous between categories, or unclassifiable verbs that do not fit into any category. These categories either contained too few tokens, and/or were heterogenous in a way that it would be difficult to interpret the results.

Overall, 54.6% of verbs in our analysis appear with at least one marker of character role shift: eye gaze shift, facial enactment, or body shift. Eye gaze shift occurs on 28.0% of verbs and is the most common marker of the three. Facial enactment appears on 23.8% of verbs. Body shift appears on 22.4% of verbs. These differences are significant even after controlling for stimuli and participant effects – a log-likelihood ratio test shows the data are more likely to be predicted if type of non-manual marker is included ($\chi2(1) = 135.2$, $p < 0.001$). However, we are crucially

[6] Moving one's head can be read in one two ways – (1) the signer is copying the character's movement or (2) the signer is using the head shift to establish themselves in a new "spatial locus", just like body shift. This distinction emphasizes a broader point that "torso shift" and "facial expression" are not just distinctions in which articulators are used but also how they are used — 'facial expression' is only coded if the actions or attributes of the person are demonstrated through the face and/or head posture. 'Torso shift' is not enactment: the shift does not mean that the indexed agent moved their torso, rather the shift indicates that the reported agent is associated with a specific spatial locus. Eye gaze, however, is fully distinguished from the other markers on the basis of articulator — in large part because it can serve multiple functions, including both enactment and indication of the spatial loci of arguments, and we cannot disentangle these in the data.

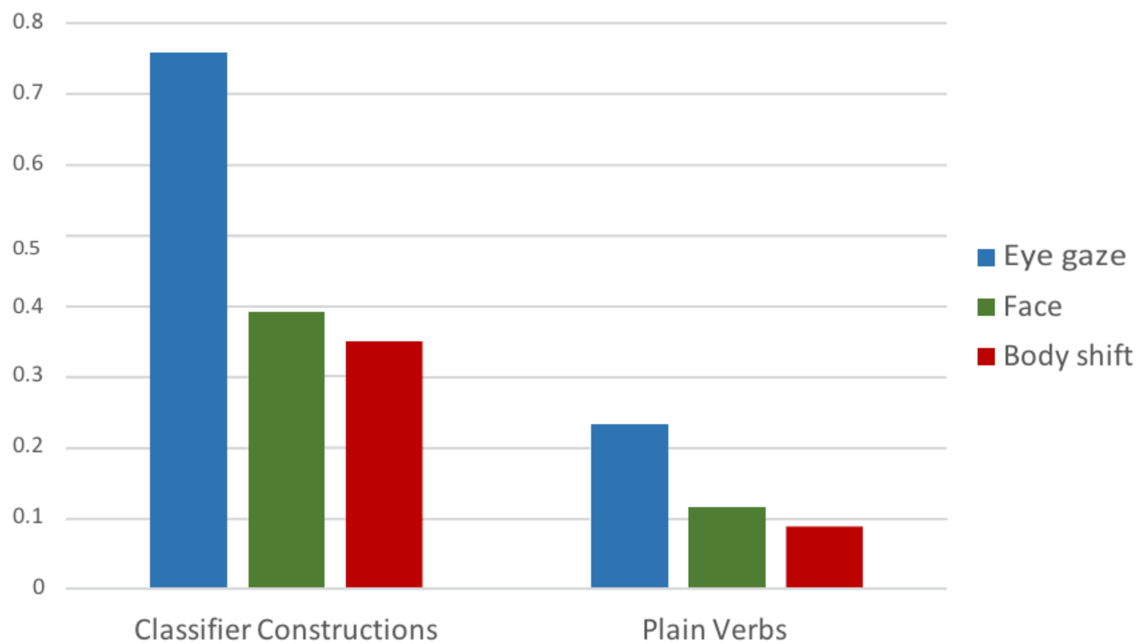


Figure 6. Proportion of classifier constructions and plain verbs with each marker of character role shift.

interested not in absolute rates of character role shift in our particular data set, but rather in how use of role shift varies as a function of type of verb.

The analyses are organized into three parts. First, we compare the three markers of role shift in plain verbs vs. classifier constructions. We then focus on variation in use of role shift within plain verbs. Finally, we turn to variation in use of role shift within classifier constructions.

Non-manual-markers of role shift in plain verbs and classifier constructions

Figure 6 shows the production rate of the three markers (eye gaze, facial expression, and torso shift) for both plain verbs and classifier constructions (as a single category): a rate of 1.0 indicates that the marker appeared on all the verbs in the category. Overall, all three markers are produced more often with classifier constructions than with plain lexical verbs: the logistic regression model predicts plain verbs are less likely to show markers of role shift (Wald Test, $z = -11.22$, $p < 0.001$). For plain verbs, no marker is used more than 25% of the time – eye gaze shift is at 22%, followed by facial enactment (12%) and body shift (9%). In contrast – 76% of classifier constructions appeared with eye gaze shift, 39% appeared with facial enactment, and 30% appeared with body shift. For both verb types, eye gaze shift is the most common marker, followed by facial enactment, and then body shift. However, the interaction of marker type and verb type is still significant ($\chi2(1) = 9.3$, $p < 0.01$): the rate of use of the three markers is more similar for plain verbs than for classifier constructions.

Role shift was thus used less often with plain verbs than with classifier constructions, as we might expect given the literature, and given that plain verbs are less iconic than classifier constructions. Nevertheless, all three markers of role shift were used in plain verbs, and in a similar distribution to classifier constructions (eye gaze shift > facial enactment > body shift). We next examine plain verbs more closely to explore the factors that determine when role shift is used with a plain verb.

Plain verbs with non-manual-markers of role shift: spreading and iconicity

To examine whether we might be able to explain the variation we see in use of character role shift with plain verbs, we see whether we can attribute some of this variation to the context of the utterance and/or to properties of the plain verb itself.

First, it is possible that the role shift 'spreads' from classifier constructions to plain verbs that accompany them. The hypothesis is that the markers of role shift originate with the classifier construction, but might extend to a plain verb that is produced in the same utterance. If so, plain verbs would receive role shift only when they accompany a classifier construction that is marked by role shift.

To test this hypothesis, we compare plain verbs in two contexts: those that appear in a description that also has a classifier construction (109 tokens), and those that appear in a description without classifier constructions (and without any verb that could be a classifier

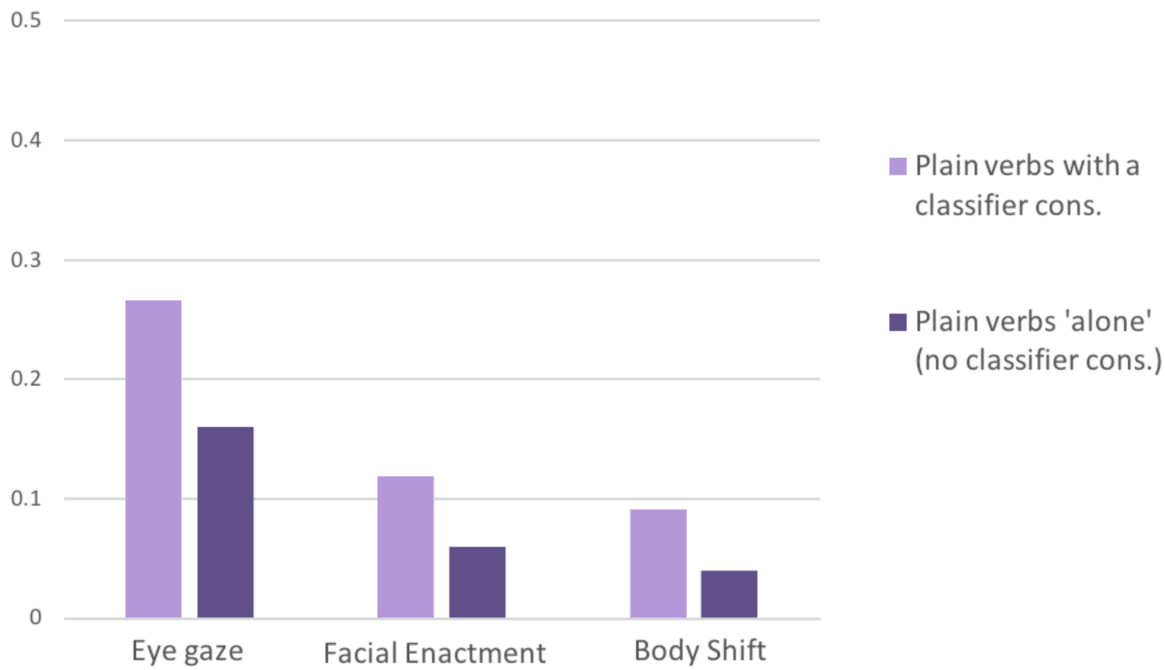


Figure 7. 'Spreading' of role shift from classifier constructions to plain verbs. Rate of use of each marker of character role shift with plain lexical verbs based on co-presence of a classifier construction.

constructions: 50 tokens[7]). Figure 7 presents the data, which are divided into three pairs of columns: one for each marker (eye gaze, face, and body shift). The first column in a pair represents the proportion of plain verbs that have the marker when they are not accompanied by a classifier; the second column represents the proportion of plain verbs that have the marker when they are accompanied by a classifier construction. Plain verbs that are 'alone' (without a classifier construction) in a description contain role shift 4-16% of the time. When the plain verb appears with a classifier construction, the rate rises to 9-27%. Overall, the three markers do appear with plain verbs more frequently when the verb is in a description that also contains a classifier construction, as indicated by the effect of the presence of classifiers in a model of the data in which subject is controlled ($z = 5.10$, $p = 0.023$). However, this effect is confounded with

[7] 22 lexical verb tokens were excluded because they the description did not have unambiguous cases of classifiers, but they did have ambiguous cases of classifiers.

type of stimulus: different kinds of stimuli both were more likely to elicit classifier constructions and also to elicit plain verbs with role shift, and the three effects cannot be fully distinguished. In other words, we do not know whether the effect of stimulus type on the use of role shift is mediated by the use of classifiers alongside plain verbs, or whether both are independently motivated by some other property of the stimulus. So, we tentatively suggest that there can be 'spread' of role shift that extends from a classifier construction to its neighboring plain verb. But it is clear that not all plain verbs conform to this pattern – a substantial number of plain verbs have markers of role shift despite the absence of a neighboring classifier construction.

Next, we examine the relationship between the degree of iconicity of the plain verb and use of role shift. Based on previous work (Davidson, 2015; Schlenker, 2017) we hypothesize that the more iconic a plain verb is, the more it is marked by role shift. To obtain a general measure of iconicity in plain verbs, we use data from the ASL-Lex project (asl-lex.org; Caselli et al., 2017), an online database of nearly 1,000 plain items in ASL. These verbs have 'iconicity ratings' on a scale of 1 to 7. The rating for each item is aggregated from scores of 25-30 hearing non-signers who, when shown the ASL sign, are asked to rate the degree of iconicity between the sign's form and its meaning, from 1 (not similar at all) to 7 (very similar)[8].

Our data contain 162 tokens coded as plain verbs (or 'instances of use' of lexical verbs), which represent 39 different "types" of plain lexical verbs (i.e., 39 different 'words', such as HAVE, EAT, SIT, etc.). ASL-Lex has iconicity ratings for 20 of the 39 verb types in our data (just above half), which account for 91 (56.2%) of all coded instances of lexical verbs. We show the results in Figures 2.3, 2.4, and 2.5, as three scatterplots – one for each marker of role shift.

[8] Iconic words were defined to participants as words that "look like what they mean", with examples like BALL and DRINK given; see Caselli et al. (2017) for a full script of the question.

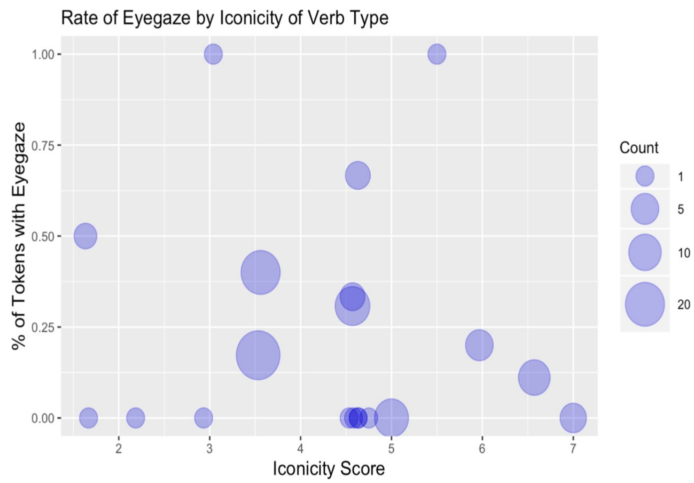


Figure 8. Relationship between rate of eye gaze and iconicity score for each verb type

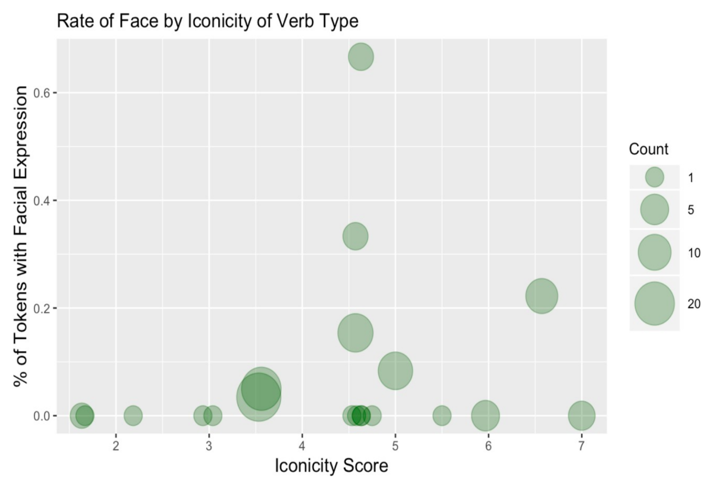


Figure 9. Relationship between rate of use of facial enactment ('facial expression') and iconicity score for each verb type

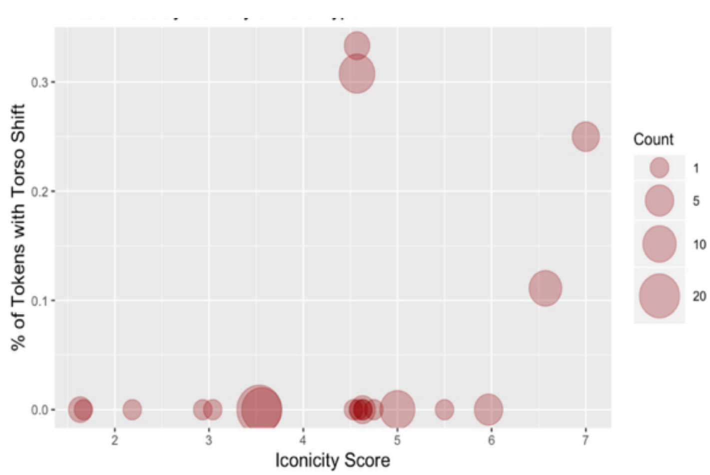


Figure 10. Relationship between rate of use of body shift ('torso shift) and iconicity score for each verb type

Figure 8 (blue) shows the relationship between eye gaze shift and iconicity score; Figure 9 (green) shows the relationship between facial enactment and iconicity score; Figure 10 shows the relationship between torso shift and iconicity score. In each scatterplot, a dot represents a single verb type (e.g. EAT) and the size of the dot shows how many instances of that verb type are in the data. The horizontal axis represents the iconicity score of the verb, and the y-axis shows the proportion of instances in which the verb appears with the given marker. Since all three graphs refer to the same set of plain verb tokens, the distribution of iconicity scores along the x-axis and the frequency of each token is identical across the graphs, but the rate of use of character role shift (represented on the y-axis) varies across the three graphs, depending on the type of nonmanual marker.

When analyzed at the level of tokens, iconicity ratings do predict not the use of role shift overall. But there is an interaction between slope of iconic effect and marker type (Main-Effect of Iconicity: $\chi 2(1) = 1.36$, $p > 0.2$; Interaction: $\chi 2(1) = 11.94$, $p = 0.008$). The interaction completely disappears when we examine the relationship between each verb type's iconicity score and use of role shift ($\chi 2(3) = 5.10$, $p > 0.7$). This indicates either that there is not enough variety in plain verb types (that have iconicity ratings) in our data to analyze, or that the effect we see at the token level may be the result of the particular kinds of verbs in our analysis.

In any case, the data still points to some possible patterns for further research. The interaction effect at the token level stems from the difference between eye gaze shift on one hand, and facial expression and torso shift on the other. Eye gaze shift, in Figure 8, does not show a clear pattern. appears on verbs with high iconicity scores and low iconicity scores. Facial expression, in Figure 9, only appears with verbs with relatively high iconicity scores – almost all the verbs that appear with facial expression are in the upper half of the range. Torso/body shift

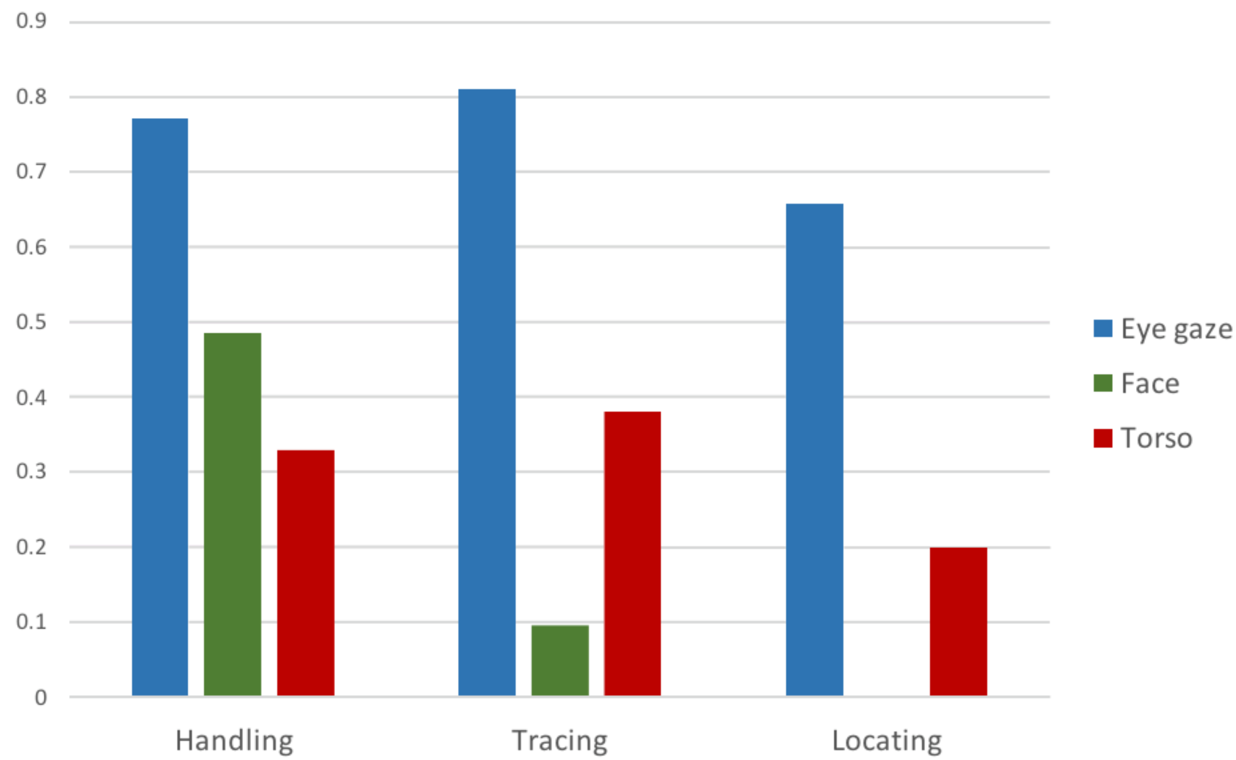


Figure 11. Proportion of each type of classifier construction (handling, tracing, locating) with each marker of character role shift (eye gaze, facial enactment, torso shift).

patterns similarly to facial expression: it also appears on some highly iconic words and on not words with low iconicity, but only 2-3 verbs used torso shift in our data, so more research is needed.

In sum, the results from plain verbs suggest that use of character role shift depends on the context the verb token appears in, and also the individual characteristics of the verb.

Classifiers constructions with non-manual-markers of role shift

Here we examine classifier constructions in detail, looking at the use of role shift with handling, tracing, and locating classifier constructions. Figure 11 displays the interaction of the 3 classifier types with the three marker types. The classifier constructions all have similarly high use of eye gaze shift (60-70%) and similarly low use of body shift (25-33%). But they behave in very distinct ways with respect to facial expression: handling classifiers appear with face markers 50% of the time, but tracing and locating classifiers appear with face less than 5% of the time each. The interaction between type of classifier construction and type of marker is highly significant ($\chi 2 = 55.9$, $p < 0.001$). In particular, there is a relatively large difference in rate of use of facial expression enactment when comparing handling classifier constructions with tracing and locating classifier constructions ($z = -3.92$, $p < 0.01$).

All three types of classifier constructions pattern similarly for use of eye gaze and torso shift to mark role shift, but only handling classifier constructions consistently appear with facial expression. It seems tracing and locating classifier constructions are generally not compatible with the use of facial expression. The three markers do not exhibit a uniform pattern across the classifier constructions.

2.4 Discussion

Overall, several patterns of character role shift and verb types appear to be with consistent previous observations that use of character role shift is correlated with iconicity in the verbal predicate. For example, handling event classifier constructions, the verb class with the highest rate of constructed action, are highly iconic: they iconically use movement to represent movement (changes in spatial relations over time), and also use the hand to represent a hand. Tracing and locating classifiers involve iconic representation of spatial relations, but not temporal relations or body-part relationships, and correspondingly they use fewer markers associated with character role shift. The least iconic category of verbs in the data, plain lexical verbs, use the least markers of character role shift – and the ones that do appear with character role shift are generally more iconic. Taken together, this suggests a systematic relationship between iconicity and character role shift across all predicates in the lexicon – including those in the spatial lexicon (e.g., classifier constructions) and even the core lexicon (e.g. plain, lexical, non-spatial, verbs).

However, once we examine the data more carefully, broad ideas of iconicity or similarity do not directly appear to account for all of the patterns at hand. We must account for the fact that the three markers of character role shift display different patterns across classifier verbs: a single account of generic iconicity cannot explain all three patterns at once. In particular, we want to be able to explain why signers produce eye gaze and body shift, but not facial enactment, during tracing and locating classifiers.

This puzzle can be broken down in two ways: first, tracing and location classifier constructions behave differently from handling classifier constructions, with respect to character role shift. Secondly, eye gaze shift and facial enactment behave differently in some contexts. In the following sections, I review literature that point to possible hypotheses that can explain each of these puzzles. These hypotheses form the basis of my second study, which aims to resolve the puzzles that this study raises.

3 Interlude

I outline some ways in which role shift, or specific markers of role shift, may be constrained or surprisingly, not constrained. Broadly speaking, I suggest we can re-examine the question of verbs under role shift from a few perspectives. First, rather than look at how iconicity as a whole constrains role shift, we can examine the relationship between the use of role shift and specific iconic mappings. Second, I review a common assumption in the literature that the character in the role shift is also the subject and agent of the verb produced with the role shift.

This section motivates the second study, which investigates whether the use of classifier constructions with role shift is affected by these different constraints rooted in iconicity, argument structure, and/or presence of character.

3.1 Iconicity and role shift

In this dissertation, the definition I give for iconicity is ‘resemblance between form and meaning’, or ‘the sign looks like what it means’. This definition does not restrict iconicity to a specific domain. It also indicates that the identification of iconicity is a subjective process – not all speakers or signers uniformly agree whether a given sign is iconic (Occhino et al., 2017; Caselli et al., 2017). The observation that iconicity constrains the use of role shift, then, is unsatisfying on its own – it is not immediately obvious how one might formalize this observation into a concrete set of criteria that determines when and where role shift is licensed.

However, in practice, iconic phenomena in sign languages can often be grouped by the type of iconic strategies they employ – and these individual strategies often allow for more precise characterization. For example, work by Meir, Padden, and others delineate several iconic strategies for sign language nouns, such as “manipulation” strategy (show how the referent should be handled) or “personification’ strategy (use one’s body to ‘stand in’ for the referent”).

Furthermore, they suggest that sign language lexicons and structures can be characterized by the presence of these specific iconic mappings. Different strategies of iconicity are used to represent different semantic categories of nouns (Padden et al., 2013; Padden et al., 2015). For example, "manipulation" strategies are typically used to represent tools (by showing how a tool is used), whereas "personification" strategies are typically used to represent animals (by portraying an animal with the body).

One way to tackle the question of iconicity and role shift is to identify specific iconic strategies or mappings that constrain role shift. In particular, I want to bring attention to three such iconic mappings: *body-as-body iconicity, temporal iconicity*, and *spatial iconicity*.

Body-as-body iconicity refers to constructions in which the signer iconically depicts the behavior of another person (the 'character' or 'referent') with their own body, in which their body (whole or part) stands in for the corresponding articulator of the referent. Facial enactment in character role shift is one such example – the signer's face represents or stands in for the character's face. Handling classifier constructions are another such example – the hand of the signer stands in for the hand of the referent.

This distinction between handling classifier constructions and entity classifier constructions has been noted in the literature (Dudis, 2004; Perniss, 2007; Cormier et al., 2012; Quinto-Pozos & Parrill, 2015; Engberg-Pedersen, 2015). Many researchers discuss the distinction in terms of perspective. Handling classifier constructions necessarily are produced 'from the perspective of the character' producing them. In contrast, entity classifier constructions can be produced in a compressed space, as if a signer is an observer looking at the event from above. This distinction between character and observer viewpoint originates with David McNeill's work on co-speech gesture in spoken discourse, who identifies two types of viewpoint gestures: character viewpoint and observer viewpoint gestures (McNeill, 1992). Under this framing, handling classifier constructions necessitate a character viewpoint in the scene and therefore also imply the use of character role shift.

I frame the constraint slight differently – as body-as-body iconicity – focusing on whether the hands are representing the same referent as the body in an iconic 1-to-1 mapping of the signer's whole body to the character's whole body. This is most closely aligned with Dudis's (2004) notion of 'unpartitioned depiction'.

Temporal iconicity refers to constructions in which the time course of the construction maps iconically onto the time course of the event. The beginning, middle, and end (and the points in between) of a temporally iconic construction correspond to the beginning, middle, and end of the event denoted. Under this definition, classifier constructions that denote motion events are temporally iconic. Static classifier constructions that indicate the location or shape (through tracing) are not temporally iconic – the beginning of the sign does not correspond to any "beginning of an event". In fact, in tracing classifier constructions, the temporal progression and path of the sign instead corresponds to the shape and spatial dimensions of an object. Temporal iconicity has been analyzed in classifier constructions (Grose et al., 2007; Dudis, 2012) and also in how signers express aspect inflection, in plain lexical signs (Wilbur, 2008; Kuhn, 2017) and discussed in person-agreement verbs (Dudis, 2010). It is an iconic mapping that appears across the lexicon and in various kinds of predicates.

Spatial iconicity refers to constructions in which the signing space, the area in front of a signer, is used to represent some actual space in the real world. Spatial iconicity, and how signers use their signing space to represent real-world space, has been studied extensively. In signs that recruit spatial iconicity, the 3-D signing space is recruited as a whole to represent some 3-D space of the real world (Perniss, 2007).

Spatial iconicity is a unifying factor across classifier constructions – all classifier constructions iconically represent space. For handling and entity classifier constructions, the movement of the sign reflects the real-world motion of the referent, and iconic movement implies iconic space[1]. For entity classifier constructions of location, the endpoint location of the sign indicates a corresponding real-word location where the object is. Often, entity classifier constructions of location only make sense either when some spatial mapping has already been established in discourse or in relation to other classifier constructions (of location or motion).

In Study I, I claim handling classifier constructions are more iconic than static classifier constructions. Now, we can refine this claim: handling classifier constructions contain iconic mappings not present in static classifier constructions. Static classifier constructions involve one of the above mappings, spatial iconicity, while handling classifier constructions involve all three mappings: spatial iconicity, temporal iconicity, and body-as-body iconicity.

One of the main motivations for Study II is to parse out the effect of these various iconic mappings, and in particular to separate the effects of body-as-body iconicity and temporal iconicity each from spatial iconicity. To this end, Study II focuses on classifier constructions. Classifier constructions are a very useful phenomenon through which we can study different iconic mappings since the organization of classifier constructions corresponds tightly to these iconic mappings, as shown in Table 1.

[1] Because motion is change in space over time (hence iconic motion implies both spatial and temporal iconicity)

Moreover, the compositional nature of these constructions means that these alternations in classifier constructions are often paradigmatic. For many handling classifier constructions that represent how an object is handled in some motion, there is a corresponding entity classifier construction varying only by handshape that denotes how an object moves independently (with the same motion), and also static classifier constructions that can be used to indicate the location of that same object.

3.2 Role shift and argument structure

However, there is a second possible set of explanations that still have to be differentiated from iconicity. In prototypical examples of character role shift, the referent of the character role shift is an argument of the verb(s) produced with role shift, specifically the subject/thematic agent of the verb. That is, the verb describes some action, and the agent of the action is the same as the referent of the character role shift. Consider Figure 1.3, reprinted here from Section 1.1.

The signer is signing SEARCH, with multiple elements of character role shift, including eye gaze shift and facial enactment. This picture is taken from a longer description, translated as "A man flips through a book, searching for something". The verb SEARCH describes the type of action that the man did, and correspondingly the character role shift enacts the behavior and expression of the man while performing the searching action.

	BODY-AS-BODY	***TEMPORAL***	***SPATIAL***
HANDLING	**YES**	**YES**	**YES**
ENTITY-MOTION	NO	**YES**	**YES**
ENTITY-STATIC	NO	NO	**YES**

Table 1. Types of iconic mappings (columns) present in each type of classifier construction (rows).

Both the character role shift and the verb are about the same character – more precisely, the syntactic subject of the verb SEARCH (the man) is also the referent of the character role shift. But this might not always be the case. Dudis (2004) discuss an example, in which a homeowner looks (glares?) at a faucet dripping water, perhaps irritated by the fact that the faucet stubbornly continues to leak. With his hands, the signer produces a verb, 'DRIP', showing the downward, rhythmic, dripping of water (from a faucet). With his eye gaze and face, the signer squints at his own hand, and we are invited to imagine that we are watching the homeowner himself, as he glares at the faucet dripping water.

The verb, DRIP, tells us only about the event of dripping water, and by itself gives no indication that there is any person is involved in the event. The referent of the character role shift, the homeowner, is not the subject or thematic agent of the verb. In fact, the homeowner is not an argument of the verb at all.

It is still an open question to what extent constructions like this one are allowed, and if so in what contexts. Besides this example, there are other instances where the referent of character role shift does not correspond with the subject or agent of the verb. In Janzen's (2004) work on role shift and perspective, he discusses how the signer can role shift to the syntactic object of a verb, showing how an event is perceived by the patient or recipient of an action. Both Napoli and Sutton-Spence (2010), and also Supalla (1982), generally observe that while it is possible to disassociate the verbal subject and referent of the role shift, generally the referent of the role shift is thematically involved in the event in some way.

In sum, there are at least two possible constraints on role shift originating in the argument structure and event semantics: 1) the referent of role shift must be the subject of the verb (syntactic constraint) and 2) the referent of the role shift must be a thematic role in the event expressed by the verb (semantic constraint). Thinking back to the results of Study I, both of these constraints would not restrict the use of handling classifier constructions with role shift but do constrain the use of static classifier constructions with role shift. Clearly, it can be difficult to tease apart evidence for iconicity-based constraints from evidence for argument-based constraints, and we have to be careful if we want to make claims about the type of constraints at play.

3.3 Role shift and presence of character

This study also tests the assumption that these markers of role shift actually are related to expressing a character's perspective. Specifically, it is possible that eye gaze shift may be used to express perspective without an actual character in the scene.

This is most directly studied by Ferrara (2019), in her work on Norwegian Sign Language. The study recruited participants from a sign language program: both the native signing teachers and their students (who are second language learners). Participants were given two spatial description tasks – they had to give directions on how to arrive at the "new campus" and describe the layout of the floor of the sign language classrooms.

In both, the native signers frequently looked towards their hands while describing the location of entities. Ferrara (2019) gives an example in which a signer shows the location of three classrooms, by marking three points in a line with the same entity classifier construction. As he does so, his eyes are (mostly) directed at the 'classrooms' represented by his hand, and his body is shifted to the side.

The author explains that the signer in this example has conceptualized the scene from a specific vantage point – looking down from the end of a hallway. According to the author, there is no evidence of a character present in the enactment. Rather, the signer "established a vantage point with his eye gaze that showed his own position in relation to the conceptualized scene".

Although the study focuses on eye gaze, the author notes that body shifts appear alongside many of the shifts in eye gaze. These examples of eye gaze shift and body shift, although they can look like our earlier examples of character role shift, are not connected to any actual character.

Ferrara is not the first to link the use of eye gaze to spatial viewpoint and the depiction of spatial relations. Engberg-Pedersen (2003) had also previously suggested the use of eye gaze to index the spatial relations between entities in Danish Sign Language. There has been work on the use of signing space to represent spatial layout in ASL (see Emmorey et al., 2000), but to my knowledge there has been no comparable analysis of eye gaze or torso shifts in these kinds of spatial descriptions for ASL. Dudis (2012) also discusses the notion of scene depictions with a 'vantage-point' – suggesting the need for perspective even in depictions without a character.

In Study I, all my stimuli had a character present in the scene, but it is possible that in some cases, the use of role shift did not actually implicate or require the character is present.

3.4 Summary

Study II examines the use of different markers of role shift with classifier constructions, and extends the results of Study I. By comparing different types of classifier constructions, Study II hopes to identify which specific iconic mappings are relevant for role shift, and also how the referent of the role shift must be related to the verb or event. Finally, we also investigate whether these markers appear in the absence of the character. Classifier constructions are suitable for this investigation because they can be clearly differentiated by iconic mappings, it is easier to reliably elicit the expected form without prompting, classifier constructions overall often make use of role shift (so where role shift is disfavored, the patterns are hopefully clear), and targeted manipulations in stimuli can elicit target manipulations in the classifier constructions.

Broadly speaking, this study is motivated by the possibility that different markers of role shift represent different underlying conceptual devices. If true, it should be possible to distinguish these different conceptual devices by the different distributions of their associated markers – since these different devices might obey different constraints with respect to iconicity, argument structure, and presence of character.

4 Study II: Character Role Shift with Classifier Constructions

4.1 Goal and design

Study II investigates the use of two markers of character role shift, eye gaze shift and facial enactment, and how they are used with different types of classifier constructions in various contexts.

I present deaf signers with a series of videos and picture stimuli for them to describe in American Sign Language. Each stimulus involves either a stationary object or an object that undergoes motion (handled or otherwise) and is intended to elicit a specific type of 'target' classifier construction that describes the motion or location of the object. Study II compares use of role shift across classifier constructions along four main axes:

1. Use of markers of character role shift with handling versus entity classifier constructions (of motion).
2. Use of markers of character role shift with events that do not involve the character as a participant.
3. Use of markers of character role shift with static versus motion classifier constructions.
4. Use of markers of character role shift with scenes with character present versus scenes without a character present at all.

We test for evidence of each possible constraint separately for facial enactment and eye gaze shift (in this study I did not code for body shift because it was so infrequent in Study I). I discuss each constraint in turn and, for each, I detail how this study tests for the given constraint (while controlling for the other constraints).

Handling versus entity classifier constructions

I will compare the use of nonmanual markers in handling and entity classifier constructions by comparing responses to several sets of video stimuli: “Motion-Handle” videos in which an actor moves an object, “Motion-Initiate” videos in which an actor acts on an object to initiate its movement but it then moves on its own; “Motion-Watch” videos in which a person observes an object moving. Motion-Handle stimuli elicit handling classifier constructions. Motion-Initiate and Motion-Watch stimuli elicit entity classifier constructions.

These sets of stimuli elicit different types of classifier constructions, but control for the following factors: the videos involve motion events, have a character present who either acts on the object or observes the object.

Suppose we observe a difference in use of role shift between handling and entity classifier constructions (of motion). In explaining this difference, we can rule out both explanations of temporal iconicity and spatial iconicity, since both handling classifier construction and entity classifier constructions of motion make use of these iconic mappings. Rather they are distinguished by whether they contain body-as-body iconicity – in handling classifier constructions (in which the hand represents a hand) the classifier construction and the body under role shift as a whole 'stand in’ for the character. This is not true for entity classifier constructions. Entity classifier constructions represent an entirely different referent[1].

There is a second explanation, however. Handling classifier constructions describe the actions of the referent of the character, and also the referent is the subject of the verb. Thus, the referent of role shift is also a syntactic and thematic argument of the handling classifier construction. Entity

[1] In some cases, the entity classifier constructions *can* represent the same referent as the role shift, such as whenusing the whole-entity handshape for a person to represent their movement – as they walk through a room, for example.

classifier constructions do not involve the character but only describe the emotion of the inanimate object, and the character is typically not an argument of the entity classifier constructions.

Event caused by character versus events observed by character

Next, I attempt to at least partially parse out these two accounts ('body-as-body' iconicity and referent-as-argument). To do so, I examine whether nonmanual markers are used more often if the character actually participates in the event (Motion-Initiate), as opposed to merely watching it (Motion-Watch). Both videos have a person in the video, but both elicit entity classifiers

This comparison specifically targets the notion of semantic argument – whether the character participates in the event. It is possible that entity classifier constructions are not restricted from being produced with role shift *per se* but are often used to describe events where the character has no involvement. This comparison allows us to isolate the effect of constraints due to the type of event described (the relationship between character and event) from constraints specific to the structure of the classifier construction (its iconic mappings and syntactic structure).

Static versus motion classifier constructions

In this analysis we compare the use of nonmanual markers across location entity classifier constructions and motion entity classifier constructions, by comparing signers' responses to Motion-Watch videos with their responses to Static-Watch pictures in which the person is shown watching a stationary object. The Motion-Watch stimuli elicit motion entity classifier constructions and the Static-Watch picture stimuli elicit location entity classifier constructions.

This comparison isolates the effect of temporal iconicity. Both static and motion entity classifier constructions lack body-as-body iconicity and both also have spatial iconicity. They are also both intransitive verbs in which the character observing the object is not an argument. Thus, other than temporal iconicity, we can rule out effects of other iconic mappings and constraints due to argument structure.

Character present versus character absent

We examine whether the use of nonmanual markers requires the presence of a character by comparing "Motion-Watch" stimuli to "Motion-Alone" stimuli, where an object moves without a person present. Both sets elicit motion entity classifier constructions of the object, they only differ in whether a character is present.

The classifier constructions in the two conditions are the same type (entity classifier construction of motion) and denote the same event – the independent movement of the object. They have the same argument structure and iconic mappings. When the character is present ("Motion-Watch") they are not involved in the event: they do not act on the object in any way. We can target precisely the effect of having a character present, and see if elements of role shift actually depend on the presence of a character at all.

4.2 Methodology

Participants

This study uses data collected from eight participants. All eight subjects were first exposed to ASL by the age of five and use ASL regularly in their daily lives. Participants were recruited through postings on public groups on social media (Facebook), and email blasts to people who previously participated in studies at University of Chicago and have agreed to being contacted about additional studies.

Materials

The study consists of 36 stimuli, including 24 videos and 12 pictures. The 36 stimuli are grouped under six different conditions consisting of six stimuli each. The conditions vary in terms of whether 1) it is a video of a moving object or a picture of a still object, 2) a human actor is present, and 3) a person is present, whether that person is interacting with the object and to what extent. There are 4 video conditions in which an object moves along a path – one in which there is no person at all, and three with a person present. In one condition, the person watches but does not touch the object; in another, the person initiates the object's movement (e.g. by pushing or dropping the object); and in the last video condition, the person holds onto the object and moves it along the path.

There are also two picture conditions: in one condition the object is resting alone without a person present; in the second picture condition, the only change is that a person is present looking at the object. The six conditions are listed below:

- Motion-Alone: Video of a moving object, no person present.
- Motion-Watch: Video of a moving object, a person watches the object as it moves.
- Motion-Initiate: Video of a moving object, a person initiates the movement of the object and then watches as it moves.
- Motion-Handle: Video of a person taking hold of an object and moving it along.
- Static-Alone: Picture of a still object, no person present
- Static-Watch: Picture of a still object, a person looks at the object

I describe example stimuli of each condition, using stimuli involving a toy truck either rolling across a table, stationary on the table.

The six examples are listed below:

- Motion-Alone: Video of toy truck rolling across a table, there is no person in the video.
- Motion-Watch: Video of a toy truck rolling across a table, while a woman watches.
- Motion-Initiate: A toy truck is sitting on a table. The woman gives the truck a push, and it rolls across the table.
- Motion-Handle: The woman grabs the toy truck and rolls it across the table.
- Static-Alone: Picture of the toy truck sitting on the table, there is no person in the picture.
- Static-Watch: Picture of the toy truck sitting on the table, while a woman stares at it.

All stimuli were designed to involve clearly identifiable objects, and the background is simple to encourage the subjects to direct their focus on the object at hand. In video stimuli, the movement of the object follows a single general path[2], and the videos last approximately 5-10 seconds. When a person is present, the person has a facial expression meant to convey an identifiable emotion, such as happiness or frustration – this is intended to encourage subjects to describe or enact the actor's facial expression in their descriptions of the stimuli), and the actors are always looking at the object.

The stimuli are also matched across conditions by object and 'scene' – that is, for each stimulus in a condition, there is a corresponding stimulus in each other condition involving the same object, person (if present), and was filmed in the same location. For example, we have six stimuli, one in each condition, that all are of a 'sinking-bowl' scene: they all include the same yellow bowl sunk in a vase of water – whether the bowl sunk by itself, was dropped into the

[2] Object might also have might have 'internal movement' (e.g., a rolling ball moves along in a straight line but also 'spins' as it rolls).

vase, or a person brought it all the way to the bottom. In the picture conditions, the plate is already resting at the bottom of the vase.

Stimuli from the same 'scene' have the following features in common:

- Involve the same object(s)
- Are filmed in the same place with the same background
- Involve the same actor with the same facial expression and clothes (if actor is present)
- For the four motion video stimuli, the objects follow the same motion path
- For the two static picture stimuli, the object is in the same location

There are six 'scene's in all. The objects and events in the scenes are listed below:

- A bowl sinking in water
- A toy truck rolling by
- A lightbulb swinging from the ceiling
- A poster falling off the wall
- A pen falling down
- A book falling over

Conceptually, the 36 stimuli used in this study are the result of fully crossing the six 'scenes' with the six conditions: for every possible combination of scene and condition, there is one stimulus.

The experiment was conducted on IbexFarm[3] and Zoom.

[3] Website: https://spellout.net/ibexfarm/ (no longer maintained). Created by Alex Drummond (2021).

Procedure

The entire study happens over one Zoom call. Each participant first completes a consent process and background interview that asks questions about language background. The interview takes approximately 5-10 minutes. They then receive a link to a website that contains the experimental stimuli, and instructions to "describe what you see". The stimuli are shown twice in two blocks: they see each stimulus once, before any of the stimuli are repeated. The order of each block is independently randomized. The experiment takes place over a Zoom call with the experimenter. The participant's video is recorded through the Zoom application.

4.3 Coding scheme

The coding process can be broken into two steps. The first step is to code for type of classifier construction type; the second step is to code for presence of each non-manual marker: eye gaze shift and facial enactment. The coding scheme is similar to the coding scheme from Study I, with some differences. All coding is done using ELAN (Crasborn & Sloetjes, 2008)[4].

First, only certain classifier constructions are coded for – handling classifier constructions, motion entity classifier constructions, location entity classifier constructions, and tracing classifier constructions. Plain verbs are not coded at all. Furthermore, the classifier construction must be describing the movement or location of the target object (or how it was handled). For example, if a signer used a location classifier construction to describe the location of another entity (e.g., the location of the table in the room), that classifier construction was not coded.

Second, I only code facial enactment and eye gaze shift. I do not code body shift – pilot data indicated that the camera quality in Zoom study is less consistent and that participants

[4] ELAN (2021). http//:archive.mpi.nl/tla/elan

Stimulus Condition	Target CLC
Motion-Alone	Motion Entity CLC
Motion-Watch	Motion Entity CLC
Motion-Initiate	Motion Entity CLC
Motion-Handling	Handling CLC (Motion)
Static-Alone	Location Entity CLC or Tracing CLC
Static-Watch	Location Entity CLC or Tracing CLC

Table 2. List of conditions and expected classifier construction (target CLC) for each

sometimes are constrained in their body movement by the limits of their camera screen, so body shift is harder to code for and possibly less common.

Coding for target classifier construction

For each stimulus condition, we were looking for a specific target classifier construction (CLC) (or one of two target constructions) in the signer's description, listed in Table 2. For each type of classifier construction that I code for, a sign must satisfy the criteria for both the movement and handshape parameters (that is, both the movement and handshape of the sign must match that of the target CLC) to be counted as an instance of the target CLC.

For the static picture conditions, which are meant to elicit static classifier constructions depicting the location of the object, both location entity and tracing classifier constructions[5] are accepted as possible target classifier constructions. The criteria for determining if a sign is compatible with each type of CLC is listed below:

- Criteria for motion entity CLC: the sign must use an appropriate entity classifier

[5] As long as the tracing classifier construction indicates the location of the object in the scene.

handshape that matches the object in the stimulus. The set of possible entity classifier handshapes depends on the object in the stimulus, and is drawn from the literature, my own intuitions (and consultation with native signers), and the handshapes signers used in the pilot data. The sign must also have a movement that is consistent with the movement of the actual object during the stimulus.

- Criteria for handling CLC: the sign must use an appropriate handling classifier handshape that matches how the object was handled in the stimulus. Again, the set of possible handshapes depends on the stimulus – both the kind of object and how it was handled. The sign must also have a movement that is consistent with the path along which the object was handled during the stimulus.
- Criteria for locative entity CLC: the criteria for handshape of the location entity CLC are the same as for motion entity CLC – it is determined by type of object (to be expected, as both CLC use entity classifier handshapes). The expected movement of the CLC is a short movement in a straight line, punctuated with a stop.
- Criteria for tracing CLC: the sign must use an appropriate tracing classifier handshape. The movement of the sign must trace the shape of the object.

For each description, we identified up to one target classifier construction – the first instance of a classifier construction that is consistent with the classifier construction category we expected for that stimulus. If a signer repeats the classifier construction multiple times in the description, we only use the first instance in our analysis.

Coding for nonmanual markers of character role shift

For every instance of a target classifier construction in a description, we code for the presence of two kinds of nonmanual markers related to character role shift: eye gaze shift and enactment of facial expression of character ('facial enactment').

- Eye gaze shift: a shift in eye-gaze is coded if the signer breaks eye-gaze with the camera and shifts their gaze towards the signing space, in plausibly the direction that the actor in the scene would have been looking (which was always towards the object). In stimuli where there is no actor, the eye gaze shift should be directed towards the object – where the actor is looking in the corresponding stimulus of the same event type with an actor present.
- Facial enactment: facial enactment is coded if the signer assumed a facial expression consistent with the facial expression used by that actor in that set of stimuli. Recall the stimuli were designed to elicit specific facial expressions for each condition. For example, in events with a bowl sinking in water, the actor, if present, always has a contented facial expression. We code for this facial expression in all bowl stimuli regardless of whether the actor is actually present. In Study I facial enactment could include mouth actions and even head movement or posture changes (e.g., crouching over). In Study II, the stimuli are specifically designed so that characters do not do much with their body or face other than make a facial expression showing a specific emotion; so facial enactment in Study II indicates specifically that the signer copied the character's facial expression.

4.4 Results

Overview

Overall, 443 classifier constructions are included in the main analysis. Table 4.2 shows the number of target classifier constructions elicited from each condition. Signers produced the expected classifier construction in their descriptions about 65-90% of the time.

Collapsed across the conditions, eye gaze shift appears on average on 59% of classifier constructions, whereas facial enactment only appears on 17% of target classifier constructions. When excluding conditions that did not include a character, use of facial enactment with the classifier construction increases to 25%. As in Study I, eye gaze shift is generally much more common than facial enactment across all classifier constructions.

Comparing handling and entity classifier constructions of motion

In the first analysis, we compare the use of both eye gaze shift and facial enactment with handling vs. entity classifier constructions. To do so, we compare three conditions: 1) the handling condition in which a person handles the object, 2) the initiate-motion condition in which the person initiates the movement of the object but does not hold onto for the duration of the event (e.g., they drop or push the object), and 3) the watch-motion condition in which the

Stimulus Condition	Motion Alone	Motion Watch	Motion Initiate	Motion Handling	Static Alone	Static Watch
Target CLC	Motion Entity CLC	Motion Entity CLC	Motion Entity CLC	Handling CLC (Motion)	Location or Tracing CLC	
No. Tokens	80	72	70	86	73	62
% of stimuli with target CLC	83%	75%	73%	90%	76%	65%

Table 3. Responses with target CLC by condition.

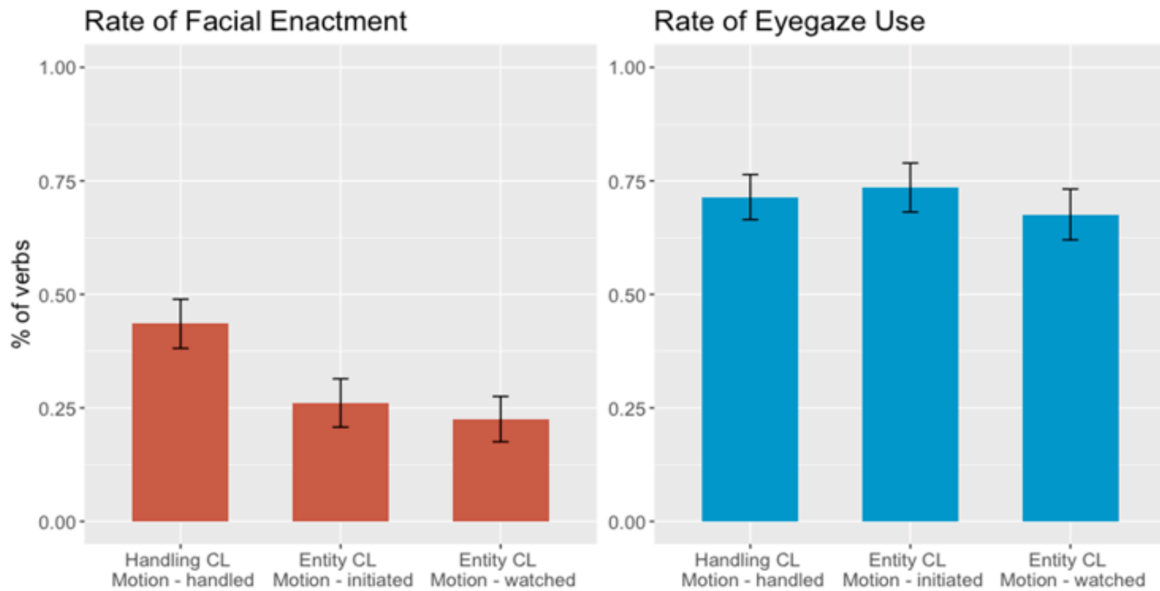


Figure 12. Test for effect of handling v. entity constructions and effect of have character as agent (causer) of event. Compares the use of eye gaze (blue) and facial enactment (red) between stimuli where a 1) person handles an object, 2) pushes/causes an object to move, 3) watches an object move.

object moves by itself while the person simply watches. The first condition elicited handling classifier constructions; the second two conditions elicited entity classifier constructions.

These three conditions test the predictions of two distinct hypotheses. First, according to the hypothesis that handling classifier constructions are more compatible with character role shift than entity classifier constructions, the handling condition should elicit the highest use of character role shift, compared to the other two. A separate second hypothesis suggests that if the character is involved in an event (e.g., by causing an event), then character role shift should be relatively more accessible than if they were not involved, even if the verb used does not change. This hypothesis predicts that the motion-initiate condition should elicit higher use of character role shift than the watch-motion condition even if both elicit entity classifier constructions. We analyze how use of each marker, eye gaze shift and facial enactment, vary across the three conditions. The results from these three conditions are shown in Figure 12.

The first three blue bars show the use of eye gaze shift with the target classifier construction in each of the three conditions (handling, motion-initiate, motion-watch). The height of the bars shows percent of the target classifier constructions occurring with eye gaze shift. Similarly, the three red bars show the use of facial enactment with the target classifier construction in each of the three conditions (in the same order: handling, motion-initiate, motion-watch).

From the graphs, it appears that eye gaze shift is consistently high for all three conditions, handling (71%), motion-initiate (74%), motion-watch (68%). Use of facial enactment is highest with the handling condition (43%), and lower for both motion-initiate (26%) and motion-watch (23%).

I first analyze the use of facial enactment across the three conditions. To test the effect of handling versus entity classifier constructions in use of facial enactment, we collapsed the motion-watch and motion-initiate conditions together and ran a logistic regression model with use of facial enactment in a trial as a binary response, with handling/entity as a two-level fixed factor and subject and setting as random factors. The effect of handling vs. entity conditions is significant (Wald test, $z(221) = 3.647$, $p < 0.001$), indicating that facial enactment is used more in the handling condition than the conditions that elicit entity classifier constructions.

Next, I test whether the motion-initiate and motion-watch conditions differ in use of facial enactment. In this logistic regression model, the response is again use of facial expression in a trial, with condition (motion-initiate or motion-watch) as a two-level fixed factor, and setting and subject as random factors. The effect difference between the motion-initiate and motion- watch conditions is not significant (Wald test, $z(137) = 0.490$, $p = 0.62$): there is no evidence that signers use more facial enactment with entity classifier constructions if the character actually caused the movement or not.

Finally, I compare the patterns in use of eye gaze shift across the three conditions. In contrast to facial enactment, there is no evidence that eye gaze shift is used more with handling classifier constructions than with entity classifier constructions (Wald test, $z(219) = 0.95$, $p = 0.34$). A model that incorporates both nonmanual markers shows that there is an interaction between the two nonmanual types (Wald test, $z(442) = 2.066$, $p = 0.038$). Facial enactment and eye gaze shift show different patterns across handling and entity classifier constructions: facial enactment is more common with handling classifier constructions than entity classifier constructions; eye gaze shift is used with both types of classifier constructions at similar rates. Use of eye gaze shift also does not appear to vary within entity classifier constructions as a function of the involvement of the character as an agent. When comparing entity classifier constructions in motion events where the character caused the object to move vs. entity classifier constructions in motion events where the object moved on its own, the use of eye gaze shift is the same (Wald test, $z(135) = 0.581$, $p = 0.56$). In this respect, eye gaze and facial enactment are similar, neither show a difference in use between motion-initiate and motion-watch conditions (Interaction: Wald test, $z(273) = -0.106$, $p = 0.92$).

Comparing motion and static classifier constructions

Another hypothesis to test is whether temporal iconicity modulates the use of role shift. To test this hypothesis, we can compare the entity classifier constructions elicited in the motion-watch condition to the static classifier constructions elicited in the static-watch condition. In both conditions, a character is present observing the object, but the character does not physically interact with the object at all. The two conditions are distinguished by event type: one condition elicits eventive (motion) verbs to describe motion events and the other elicits stative verbs to describe locations (and shapes) and temporal iconicity: the motion verbs are temporally iconic, the stative location/shape verbs are not.

Figure 13 shows the use of eye gaze shift (blue) and facial enactment (red) with each of the two conditions. The first blue bar represents the rate of use of eye gaze shift in the motion-watch condition, and the second blue bar represents the rate of use of eye gaze shift in the static-watch condition. Similarly, the two red bars represent the rate of use of facial enactment in the motion-watch and static-watch condition, respectively.

As noted earlier, eye gaze shift is consistently higher than facial enactment in both conditions. For the motion-watch data (repeated from Figure 12), eye gaze shift appears with 68% of the classifier constructions, but facial enactment only appears with 23% of the classifier constructions. For the picture-watch data, eye gaze shift appears with 49% of the classifier

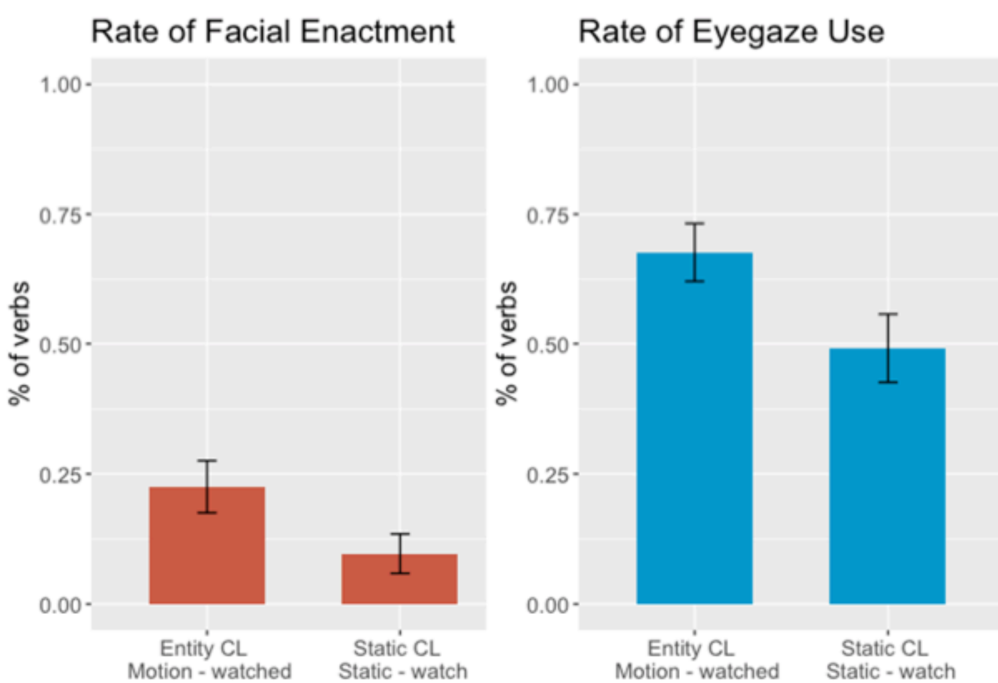


Figure 13. Test for effect of event v. static constructions. Compares the use of facial enactment (red) and eye gaze (blue) between stimuli where a person watches a moving object vs. stimuli where a person watches a stationary object.

constructions, and facial enactment appears with 10% of the classifier constructions. The static classifier constructions in the picture-watch condition elicited less use of both eye gaze shift and facial enactment than the motion entity classifier constructions in the motion-watch condition (Eye gaze shift: $z(126) = 2.501$, $p = 0.012$; Facial enactment: $z(129) = 2.04$, $p = 0.04$).

However, the pattern is the same and there is no interaction between the two markers ($z(257) = 0.253$, $p = 0.80$). Both facial enactment and eye gaze shift show a drop in use with static classifier constructions, compared to entity motion classifier constructions.

Use of 'character role shift' without a character

The results so far show again that use of eye gaze shift is quite robust with all the classifier constructions examined so far – although there are differences between motion classifier constructions and static classifier constructions, the use of eye gaze shift is relatively high across all four conditions (49-71%). What happens when we take the character out of the picture altogether (or out of the video, if you will)?

To isolate the effect of having the character present, we can compare use of eye gaze shift and facial enactment with classifier constructions in two conditions: 'motion-watch' and 'motion-alone'. In the motion-watch condition an object moves by itself without interference while a person watches. In the motion-alone condition the object moves by itself, and the person is not in the video. The only difference between the two conditions is that the person is removed in the motion-alone condition. The target classifier constructions in both conditions are entity classifier constructions.

Figure 14 shows the use of facial enactment and eye gaze shift in each of the two conditions, motion-watch and motion-alone, respectively. The motion-watch data is repeated from both Figure 12 and Figure 13: eye gaze shift appears with 68% of the classifier constructions and facial enactment only appears with 23% of the classifier constructions. For the motion-alone condition eye gaze shift appears with 49% of the classifier constructions and facial enactment only appears with 1.3% of the classifier constructions – exactly once across 73 instances.

Facial enactment behaves as one might expect – there is virtually no use of facial enactment to describe videos without a character present[6]. Eye gaze shift, however, happens

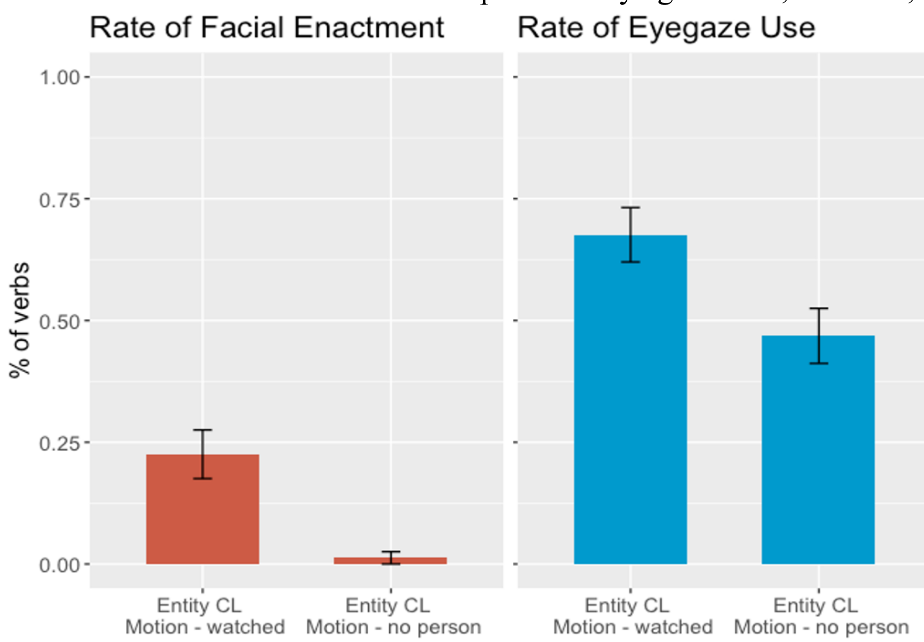


Figure 14. Test for effect of presence of character. Compares the use of facial enactment (red) and eye gaze (blue) between stimuli where a person watches a moving object vs. stimuli where an object moves (no person present)

---

[6] In some ways this is an artefact of the coding system. Remember that the stimuli across both conditions are matched by the type of object, how it moved, camera angle, where the filming happened, etc. To code for facial enactment in the motion-alone, we code whether they copied the facial expression of used by the character in the matching motion-watch condition. So, they have to not only produce a facial expression in their description of the stimulus, it has to match a particular facial expression of a character in another video.

This seems like an obviously impossible task – however, recall that the same level of stringency is applied to eye gaze shift as well – the signer must match where the character would be looking in stimuli where the character is present. However, eye gaze shift still is present, suggesting that eye gaze shift, even when it looks like enactment of a character's eye gaze is not. On the other hand, the data shows that facial enactment as it is coded in our system, quite a bit without a character present, although there is a significant drop when the character is removed (Wald Test, $z(146) = 2.776$, $p = 0.006$). When producing entity classifier constructions to show how an object moved, signers trace the path of their hands with their eyes, in nearly half of the responses in the motion-alone condition. This phenomenon, in which signers look at their hands while producing classifier constructions whether or not there's a character involved, is not just limited to motion events. The static-alone condition consists of simple pictures of objects, e.g., sitting on a table or hanging on a wall. When producing location or tracing classifier constructions to describe these scenes, signers produced 40% of the constructions with eye gaze shift.

It is clear that eye gaze shift is not just an indicator that a signer is enacting a character's behavior or experience. Signers look at their hands more during classifier constructions when describing scenes when there is a person present in the scene (as if imitating their eye gaze). But they also look at their hands a lot even when no one is present. This is in clear contrast to how one would expect character role shift to work, and in complete contrast to the results from facial enactment, which show that signers virtually never use facial enactment when the character is absent.

---

unambiguously requires a character to be present. This has the advantage of validating that our coding for facial enactment is indeed coding for facial enactment (and also that signers aren't including information from other previously presented stimuli in their descriptions).

Shaw (p.c.) wonders if coding for any kind of facial expression would show different patterns. It is possible. This is beyond the scope of this dissertation, and it requires a coding system to distinguish between uses of the face for affective and non-affective uses. Even affective uses of the face are ambiguous between expressing the emotion of the signer versus that of the character in the event. Nonetheless, it is possible there are affective uses of the face related to character role shift that are not precise enactments of a character's facial expression.

4.4 Discussion

Different types of classifier constructions clearly vary in how compatible they are with facial enactment and eye gaze shift. Facial enactment is common with handling classifier constructions, less common with entity classifier constructions of motion, and rare with static classifier constructions. For eye gaze shift, we do not observe any difference between handling and entity classifier constructions of motion, but eye gaze shift does drop off with static classifier constructions. In general, use of eye gaze shift is more robust across different types of classifier constructions, whereas facial enactment strongly favors certain classifier constructions over others. Nonetheless, neither is exclusive to handling classifier constructions, entity classifier constructions occur with both markers in some contexts.

Earlier, I proposed that role shift may require the predicate contain specific iconic mappings – such as temporal iconicity or body-as-body iconicity. There appears to be clear evidence the use of role shift is modulated by presence of temporal iconicity in the predicate. Both facial enactment and eye gaze shift are used more with temporally iconic motion classifier constructions than with static entity classifier constructions that are not temporally iconic. Moreover, facial enactment almost completely disappears with static classifier constructions. As for body-as-body iconicity, there is some evidence that handling classifier constructions appear with facial enactment (but not eye gaze shift) more than with entity classifier constructions.

This difference could alternatively be explained by differences in the argument structure between handling and entity classifier constructions. It does not seem to matter if the character is an agent of the event – i.e., if they caused the motion, since the pattern of role shift with entity classifier constructions is the same regardless of if the character involvement in the event.

However, another explanation is that handling classifier constructions specifically are compatible with facial enactment because they select the character as a syntactic subject.

In sum so far, the use of both markers is sensitive to temporal iconicity, and facial enactment in particular is very rare in predicates without temporal iconicity (eye gaze shift is still common with static classifier constructions). Facial enactment is also sensitive to the distinction between entity and handling classifier constructions – although it does appear with both types – which could be due to body-as-body iconicity or syntactic properties of the verb. We do not find any evidence eye gaze shift is affected by argument structure or body-as-body iconicity.

A final major takeaway is that eye gaze shift is not just about the character. The proportion of classifier constructions produced with eye gaze shift in the absence of a character was 40% in picture stimuli and 47% in video stimuli! Signers are looking at their hands a lot, even when there is no character to enact. To be clear, this is not a general property of signers: in data from Study I, with plain verbs signers shift their eye gaze 25% of the time and, with nouns, the rate of eye gaze shift is only 8.6% – even when describing scenes with characters. Signers typically spend most of their time in these studies looking towards their interlocutor, whether in person or on a screen.

One possible explanation is that eye gaze shift indicates that the signer has adopted a specific vantage-point from which they view the scene. In the literature, researchers note how depictions of scenes often involve establishing the signer's body as a kind of vantage point to view the scene, and some preliminary work by Ferrara (2019) suggests that eye gaze shift accompanies vantage-point constructions. I build on Ferrara (2019) and previous work to propose vantage point shift can be regarded as a type of role shift that overlaps in form with character role shift, but differ in function and also –I argue, in distribution across verb types.

This device, vantage-point shift, like character role shift, can make use of eye gaze shift and perhaps torso shift, but would not involve enactment of facial expression or posture of the character. It is compatible with any classifier construction that describes the location or motion of objects/hands in space, unlike character role shift. Character role shift requires a character to inhabit, vantage-point shift does not.

In any case, we are left with two distinct phenomena at hand. First, we have character role shift, which can involve both facial enactment and a shift in eye gaze – it is most commonly seen with handling classifier constructions, but also appears with other temporally iconic classifier constructions as well. But we also have a second phenomenon at play, which makes use of eye gaze shift and can be used with both any type of classifier construction (static, motion, handling) and is not dependent on the presence of a character. These two phenomena differ, but also overlap, with respect to their function, their form, and their distribution.

5 General Discussion

At the beginning of this investigation, I took as a starting point two major premises about role shift in ASL, based on previous literature on ASL and other sign languages. First, use of role shift is constrained in the kinds of words and contexts it can appear in (Lillo-Martin, 1995/2013). Second, a major constraint on role shit is the iconicity in a sign: a sign's compatibility with role shift is determined (at least in part) by its degree of iconicity – or the extent to which the sign's form resembles its meaning.

My data is clearly in line with previous literature with respect to the first premise. Signers are quite selective in when and where they use elements of role shift in their descriptions. Nouns are not compatible with any elements of role shift – and among verb constructions classifier constructions are largely much more compatible with elements of role shift than plain lexical verbs are.

Furthermore, there is some evidence to support the notion that use of role shift is constrained to highly iconic signs. Among plain lexical verbs, there is slight evidence suggesting more iconic verbs are also more likely to appear under role shift. Among classifier constructions, the most iconic ones, handling classifier constructions, are on average the most likely to appear with elements of role shift. This can be taken as evidence to support claims by Davidson (2015), that signs under role shift must be iconic enough to be interpretable as a demonstration of the event in question. A similar claim by Schlenker (2017), that iconic predicates are most compatible with role shift, also correctly predicts the patterns in the data.

However, this data also directly proposes a serious puzzle for both of these accounts, as well as for any account that attempts to give a single unified explanation for the distribution of role shift, as indicated by the presence of one of several markers. Simply put, different markers –

eye gaze shift, facial enactment, body shift – which across the literature are assumed by various researchers to indicate the presence of a common underlying phenomenon of 'role shift', display distinctly different distributions in my data. The contexts and type of signs in which facial enactment is used is not the same as the contexts in which eye gaze shift, or body shift, are used. It appears then, we cannot assume the existence of a single phenomenon of role shift that can be identified by the presence of any or all of these markers.

As an alternative, we might understand 'role shift' as referring to a family of phenomena that are conceptually related, often occur together, and/or make use of the body (the eyes, face, and torso) in similar ways. In the discussion for Study II, I suggested that the data cannot be accounted for with a single 'role shift' phenomenon, but that in fact there must be at least two types of role shift in my data.

I put forth a proposal of what these two role shift phenomena are, and how they can be distinguished by form, function, and distribution. One on hand, there is evidence of genuine character role shift, in which the signer enacts the character's behavior, thoughts, expressions, etc. with their body. On the other hand, my data also includes 'vantage-point' shift in which the signer uses their body to express the vantage-point from which a scene is viewed, without the depiction of an actual character, or their experience or actions. I elaborate of each of these types of role shift in turn.

5.1 Character role shift: form, function, distribution

In the beginning of the dissertation, I stated that character role shift recreates the actions, behavior, emotions, and/or experience of a character. Markers closely associated with character role shift include body shift, facial enactment, and eye gaze shift. This has been established in the literature and is not challenged by my data. Although the markers behave differently in some contexts, in the prototypical contexts that should be most amenable to character role shift (e.g., handling classifier constructions) you see high use of all three markers, frequently together.

However, several of the markers of character role shift are ambiguous by themselves, since they can be used in other types of shifts. In particular, eye gaze shift and body shift may not be reliable indicators of character role shift. The distribution of character role shift across verb categories and contexts is most reliably identified by the distribution of facial enactment. From the data with classifier constructions, there are a few factors that may make a verb most amenable to character role shift. I propose a crucial factor is temporal iconicity – classifier constructions cannot appear with facial enactment, or character role shift, if they are not temporally iconic. Also, character role shift may be more amenable to verbs in which the syntactic subject (or other syntactic argument) is the referent of the character role shift, and also be more amenable to verbs in which the hands of the signer iconically represent the hands of the referent (the body stands in for the body) – though neither factor constrains character role shift the way temporal iconicity does.

I argue that the proposals by Davidson (2015) and Schlenker (2017) can be refined. The predicates that appear with character role shift must be specifically temporally iconic (rather than simply 'iconic'). This proposal is in line with observations by Dudis (2012) that the use of facial enactment with some classifier constructions forces an eventive interpretation – and also a temporally iconic interpretation.

The account of temporal iconicity could explain not just the patterns we see with classifier constructions but also other predicates. For example, quotations are regarded as demonstrations of events (Clark & Gerrig, 1990; Davidson, 2015) – and crucially, they reproduce the temporal progression of the event they denote in real time. It also could explain why some plain lexical verbs tend to appear with facial enactment more than others, as seen in Study I. For many verbs in ASL, their telicity is visible in their form (Wilbur 2003, 2008). Telic verbs like STEAL or ARRIVE with a natural semantic endpoint contain some kind of change in the location at the end of the movement or in their handshape; atelic verbs like PLAY do not (cf. Davidson et al., 2019). Moreover, work by Wilbur (2008) show that aspectual inflections in ASL often iconically reflect the type of aspect they denote (durative, incompletive)[1]. This suggests that some instances of plain verbs in ASL contain temporal iconicity. Under the current proposal, these verbs may be more amenable to character role shift. This proposal is consistent with work by Dudis (2010), in which he argues that person-agreement verbs do not allow the use of facial enactment because they lack temporal iconicity.

In sum, while my current data focuses on the use of character role shift, the patterns of constraints on iconicity may very well extend across the entire lexicon of ASL – more work is need to verify this, however.

5.2 Vantage-point role shift and depiction of space

In instances in my data where signers use eye gaze shift and body shift, but without facial enactment and where character role shift is not expected, I propose these are cases of vantage point shift. Researchers have discussed the notion of 'vantage point' and its relevance for various types of spatial depiction, in ASL and other languages (ASL: Emmorey et al., 2000; Dudis, 2012; German Sign Language: Perniss, 2007, 2012; Norwegian Sign Language: Ferrara, 2019; Ferrara and Ringsø, 2019). This body of work, by and large, focuses on the various ways the body is positioned relative to the spatial depiction. For example, the body might be 'external' to

[1] Even more directly, Kuhn (2017) analyzes cases of modified signs in ASL in which he argues the signer is modifying the sign to iconically (and gradiently) reflect information about the progression of the event.

the spatial scene, such that the scene described is mapped onto the space in front of the signer, such that the signer is presenting a birds-eye view of the scene from above[2]. In contrast, the vantage point might instead be 'internal' to the scene (also 'life-sized'), and the entities in the signing space are arranged around the body, and the signer's body maps onto a identifiable point in the scene. Under this framework, the spatial scenes in my data are often 'internal' to the scene – entities are arranged around the signer, rather than in a condensed space in from them. To my knowledge, Ferrara (2019) is the only one who has directly studied the use of nonmanual markers to index the use of vantage point constructions. She shows there is extensive use of eye gaze shift with multiple types of vantage point constructions. My work extends her findings – static classifier constructions occur with eye gaze shift in the absence of a character in Norwegian Sign Language – to ASL, and also to classifier constructions of motion. In addition, in Study I, we find that body shift occurs frequently with static classifier constructions even though we expect that static classifier constructions block use of character role shift. It is possible then, body shift is not just limited to character role shift, but may also be a marker of vantage point shift, like eye gaze shift.

In my data, vantage point shift is not constrained in the way character role shift is. Not only can it occur in the absence of a character, but also it can occur regardless of the argument structure of the classifier construction, or whether or not the classifier construction exhibits temporal iconicity (or body-as-body iconicity). Facial enactment is absent from static classifier constructions, which indicates character role shift is not used with static classifier constructions. But eye gaze shift and body shift are compatible with static classifier constructions – if this

[2] In McNeill's distinction between character and observer viewpoint gesture and sign language work that draws on his work (Cormier at al., 2012), observer viewpoint gesture generally corresponds to gestures with external vantage points.

pattern cannot be explained by character role shift (due to the block on facial enactment) – this indicates that the alternative, vantage point shift is used with static classifier constructions. Vantage point shift is not completely insensitive to context. The results from Study I show the use of eye gaze shift and body shift is far higher with static classifier constructions than with plain verbs or nouns – suggesting that like character role shift, vantage point shift is also more restricted in its distribution with lexical verbs and nouns, compared to with classifier constructions.

Given that research on vantage point shift focuses on how it functions in depictions of space, it is likely that the high use of vantage point shift with classifier constructions is related to the spatial nature of classifier constructions. In particular, instead of body-as-body or temporal iconicity, the crucial element for the use of vantage point role shift could be spatial iconicity.

In this proposal I have sketched, vantage point shift is a type of role shift that can be distinguished from character role shift in form, function, and distribution. Unlike character role shift, it only involved eye gaze shift and possibly body shift – but not facial enactment. Its function is to establish a vantage point in a scene, and does not imply the presence of a character. It is not constrained to temporally iconic predicates, although it may still be subject to other constraints, such as spatial iconicity.

5.3 Relating the two types of role shift

I have sketched a proposal of two different phenomena at play in my data, and have emphasized where the two phenomena can be distinguished on the basis of form, function, and distribution. However, I do not believe that it is a coincidence that these two phenomena, character role shift and vantage point shift, overlap so much in form. Rather, both the literature and evidence from my data suggests that vantage point shift is best thought of as metaphorical extension of character role shift that has been stripped, or 'bleached', of certain elements.

The first part of this argument is that character role shift inherently expresses a vantage point in the scene. As Cormier et al. (2012) note, character role shift ('constructed action') is inherently 'character viewpoint', and utterances with construction action depict an event or scene from the perspective of the character in the scene. In this sense, character role shift involves the establishment of an internal vantage point, specifically conceptualizing the event from the vantage point of the character in the scene.

Under this view, character role shift encompasses vantage point shift in form and function. Like vantage point shift, it involves eye gaze shift and body shift but can also involve facial and body enactment. Like vantage point shift, it expresses a vantage point from which to view a spatialized scene, but also can enact or recreate the behavior or experience of a character.

The second part of this argument is that vantage point shift (e.g., the manifestation of vantage point through eye gaze shift) can be understood as a metaphorical extension of character role shift. The body and eye gaze no longer have to express the viewpoint of a specific character, but can index the use of any spatial viewpoint, whether it is internal or external to a scene.

Previous research often uses the metaphors of 'observer' and 'observation' to explain how vantage point is used in scene depictions, and I argue that this metaphorical mapping is not just helpful for explaining the phenomenon of vantage point shift, but might actually be the metaphorical mapping via which vantage point shift originated. As the use of role shift is metaphorically extended to be used for constructions with vantage point, other elements associated with character role shift are stripped away – specifically those elements associated with enactment of the character's experience and behavior, such as facial enactment. The lack of facial enactment in vantage point shift comes along with a loosening of constraints – no longer is vantage point shift tied to temporal iconicity or the presence of a character. This shift, in which metaphorical extension of a device comes with a loss of iconicity, is reminiscent of Meir's (2010) work on iconicity and metaphor. Meir notes that highly iconic signs tend to block metaphorical extensions of the sign meaning, if it would conflict with the iconicity in the sign[3]. The suppression of elements of character enactment from character role shift allows role shift to be extended to contexts without character present.

This story places these two types of role shift on two points along a pathway that resembles semantic bleaching – in which a 'bleached' construction has lost senses of meaning relative to another more saturated construction (Sweetser, 1988). Recent work has drawn connections between highly grounded bodily perspective and abstracted or grammaticalized expressions of perspective in language (Sweetser, 2012; Engberg-Pedersen, 2015) – the processes of abstraction and semantic bleaching here may lead towards grammaticalized variations of role shift (if they don't already exist!).

5.4 Future work

While I argue that my data convincingly shows that the various distributions of nonmanual elements across the lexicon indicates the existence of more than one kind of role shift phenomenon, further research can investigate this phenomenon further and test, and extend the claims laid out here.

First of all, my data does not have enough plain verb tokens, nor plain verb types, to do a full analysis of the ways in which role shift may be constrained with plain verbs and agreement

[3] An example would be using the sign FLY in American Sign Language, which resembles the flapping of wings, would not be metaphorically extended to expressions such as 'time flies' (as happens in English), due to the iconicity of the sign.

verbs (i.e., not classifier constructions) in American Sign Language. According to my proposal, we may be able to observe that specific iconic mappings, such as temporal iconicity or spatial iconicity, constrain the use of plain verbs with nonmanual markers of role shift[4]. While Dudis (2010) gives some examples suggesting this may be the case, a large-scale study of role shift and plain verbs could provide a fuller picture and test his observations.

In addition, it is still unknown the extent to which syntactic subject and body-as-body iconicity are relevant for the use of character role shift (or, why are handling classifier constructions so great for character role shift?). These two factors can be teased further apart by extending the study to include verbs that take the character as a syntactic subject but lack body-as-body iconicity.

Besides extending the research to give a more complete description of the phenomenon within a single sign language system, role shift is ripe for comparative work across sign languages, and also comparative work between sign language and gestures produced by hearing people. The use of role shift, with descriptions of actions and quotation, has been identified across many sign languages. The motivation for this study is based on work from American Sign Language, but also work from Norwegian Sign Language, French Sign Language, German Sign Language, British Sign Language, among others. But even as we find similarities in the use of role shift in sign languages, we also find variation (Schlenker, 2017; Quer, 2018). The use of role shift across different sign languages might be constrained by iconicity in different ways.

---

[4] It is also important to note that researchers have investigated the use of nonmanual articulators with lexical verbs and nouns, such as nouns. For example, Thompson et al. (2006) discusses the use of eye gaze shift with spatial-agreement verbs to mark locative agreement. It is an open question whether this phenomenon is related to the phenomenon discussed in this dissertation.

Generally, it is possible that the types of role shift that appear across sign languages vary in their form, function, and/or distribution.

We can also compare role shift phenomena in sign language to similar phenomena in gesture. A growing body of work is already investigating the parallels between types of viewpoint gestures and types of classifier constructions. Observer viewpoint gestures in particular adopt an external vantage point of the scene. We can ask if hearing people produce observer viewpoint gestures with eye gaze shift as signers do with vantage-point shift. If so, it would indicate that the use of eye gaze has been abstracted away from just enacting the eye gaze of a character, like in sign language.

In general, the evidence from my dissertation shows that signers decompose elements of role shift and also decompose the iconic mappings in their verbs, allowing for patterns in which individual markers associated with role shift can be associated with individual iconic mappings. Considering assumptions that gestures tend to be more 'holistic' compared to 'compositional' nature of language (Goldin-Meadow & Brentari, 2017), it is possible the gestures that hearing people produce behave as unanalyzed whole utterances, completely ruled by a single iconic strategy.

Figure 1. Scene of woman getting a tissue from a box.

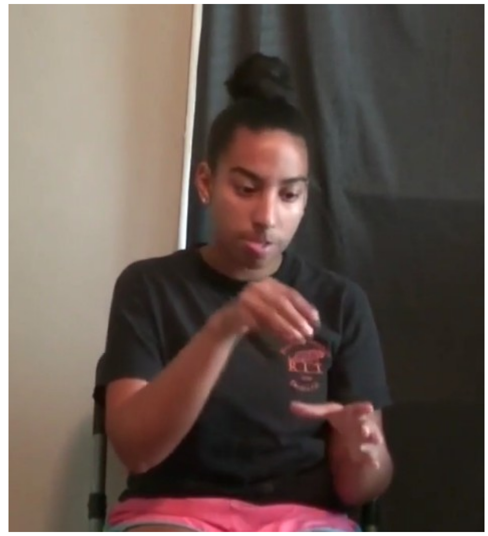


Figure 2. Handling classifier construction. Woman describes scene in Figure 1 (left) using a handling classifier construction on RH. American Sign Language.

5.5 Conclusion

I want to reconsider the example that I gave at the beginning of the dissertation, reproduced here. A signer in Figure 2 is producing a handling classifier construction, to describe the act of pulling a tissue out of a box (original stimulus in Figure 1). Taken by itself, this construction could be characterized as having a single holistic iconic strategy – many articulators of the body work together with the same goal – represent the body of the character with the signers own. The signer reproduces the behavior of the character's hands with her own hands, the movement of the character's eye with her own eyes, posture for posture, face for face. Examined by itself, it may appear a single iconic strategy can capture the whole phenomenon – as I suggested earlier, "an unanalyzed whole".

However, we know that for signers, this 'whole' is, in fact, well-analyzed. The iconic strategy of representing another's actions with one's own body involves at least three distinct mappings layered onto each other – temporal iconicity, spatial iconicity, and body-as-body iconicity. When signers combine character role shift with constructions that only match some of these iconic mappings but not others, we can see they have peeled apart these layers of mappings. Neither are other elements of the depiction as straightforward as they seem – eye gaze is not just eye gaze, but an expression of the character's perspective onto the scene in front them. We can see this clearly when eye gaze shifts are divorced from any character, merely representing a vantage point.

Thus, a single utterance, seemingly consistent with a simple holistic iconic strategy, starts to look like a collection of many structures and mappings – space-as-space, time-as-time, body-as-body, eyes-as-perspective (but also eyes), possibly face-as-experience, or face-as-emotion (but also face), a handling classifier handshape, and hand motion/location. The holistic construction has been thoroughly analyzed into the sum of mappings, and it has to be, otherwise we cannot explain the subtle ways character role shift interacts with the verb system, or how character role shift is distilled into vantage point shift. The iconic constructions are regularly more structured, and abstract, than they first appear (Taub, 2001; Wulfe & Dudis, 2005)

Note that when we break apart the utterance in Figure 2, we are not just breaking it down into discrete morphological units with lexically specified denotations. We are breaking iconic mappings into other iconic (and metaphorical) mappings.

We start to get a sense of how these various iconic mappings have to be coordinated with each other across the hands and body and with grammatical structures across the clause, or how iconic mapping can implicate several other types of iconic mappings. At times, iconic mappings may clash with non-iconic forms. In fact, sometimes iconic mappings clash with each other (Meir et al., 2007).

This dissertation shows that the use of role shift resists a straightforward characterization in American Sign Language. But hopefully, it offers a window into how we can approach the phenomena of role shift – the presence of iconicity does not preclude the possibility of structural analysis (following Taub, 2001; Dudis, 2004; Goldin-Meadow & Brentari, 2017). This has struck me as a surprise – often iconicity appears to exist in opposition to linguistic integration and grammaticalization: you must give up one to have the other (Dingemanse et al., 2015; Ortega & Morgan, 2015). Rather, the findings here align with the conclusions drawn by Meir et al. (2013): iconicity co-occurs with grammatical structures, and in fact can also shape them. Role shift is a case study of how a single phenomenon can be simultaneously iconic and also structured and integrated into linguistic structures – and structured by its iconicity, rather than in spite of it.

Works Cited

Benedicto, E., & Brentari, D. (2004). Where did all the arguments go?: Argument-changing properties of classifiers in ASL. *Natural Language and Linguistic Theory*, *22*(4), 743–810. https://doi.org/10.1007/s11049-003-4698-2

Brentari, D., & Padden, C. A. (2001). Native and foreign vocabulary in American Sign Language: A lexicon with multiple origins. In D. Brentari (Ed.), *Foreign Vocabulary in Sign Language: A Cross-linguistic Investigation of Word Formation* (pp. 87–119). Lawrence Erlbaum Associates.

Caselli, N. K., Sehyr, Z. S., Cohen-Goldberg, A. M., & Emmorey, K. (2017). ASL-LEX: A lexical database of American Sign Language. *Behavior Research Methods*, *49*(2), 784–801. https://doi.org/10.3758/s13428-016-0742-0

Clark, H. H., & Gerrig, R. J. (1990). Quotations as demonstrations. *Language*, *66*(4), 764–805. https://doi.org/10.2307/414729

Cormier, K., Quinto-Pozos, D., Sevcikova, Z., & Schembri, A. (2012). Lexicalisation and de-lexicalisation processes in sign languages: Comparing depicting constructions and viewpoint gestures. *Language & Communication*, *32*(4), 329–348. https://doi.org/10.1016/j.langcom.2012.09.004

Cormier, K., Smith, S., & Sevcikova-Sehyr, Z. (2015). Rethinking constructed action. *Sign Language & Linguistics*, *18*(2), 167–204. https://doi.org/10.1075/sll.18.2.01cor

Crasborn, O., Sloetjes, H. (2008). Enhanced ELAN functionality for sign language corpora. In: Proceedings of LREC 2008, Sixth International Conference on Language Resources and Evaluation.

Davidson, K. (2015). Quotation, demonstration, and iconicity. *Linguistics and Philosophy*, *38*(6), 477–520. https://doi.org/10.1007/s10988-015-9180-1

Davidson, K., Kocab, A., Sims, A. D., & Wagner, L. (To appear in Glossa). The relationship between verbal form and event structure in sign languages. *Glossa*. https://projects.iq.harvard.edu/files/meaningandmodality/files/davidsonkocabsimswagner-2018- manuscript.pdf

Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, iconicity, and systematicity in language. *Trends in Cognitive Sciences*, *19*(10),603–615. https://doi.org/10.1016/j.tics.2015.07.013

Donna Jo Napoli & Rachel Sutton-Spence. (2010). Limitations on simultaneity in sign language. *Language*, *86*(3), 647–662. https://doi.org/10.1353/lan.2010.0018

Dudis, P. G. (2004). Body partitioning and real-space blends. *Cognitive Linguistics*, *15*(2). https://doi.org/10.1515/cogl.2004.009

Dudis, P. G. (2010). Some observations on form-meaning correspondences in two types of verbs in ASL. In G. Mathur & D. J. Napoli (Eds.), *Deaf around the world: The impact of language*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199732548.003.0004

Dudis, P. G. (2012). The Body in Scene Depictions. In C. B. Roy (Ed.), *Discourse in signed languages* (pp. 3–66). Gallaudet University Press.

Emmorey, K., Tversky, B., & Taylor, H. A. (2000). *Using space to describe space: Perspective inspeech, sign, and gesture*. 24.

Engberg-Pedersen, E. (1999). *Eye gaze in Danish Sign Language monologues: Forms, functions, notation issues*. 3rd Intersign Workshop, Siena.

Engberg-Pedersen, E. (2003). *From Pointing to Reference and Predication: Pointing Signs, Eyegaze,and Head and Body Orientation in Danish Sign Language*. 16.

Engberg-Pedersen, E. (2015). Perspective in signed discourse: The privileged status of the signer'slocus and gaze. *Open Linguistics*, *1*(1). https://doi.org/10.1515/opli-2015-0010

ELAN (Version 6.2) [Computer software]. (2021). Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive. Retrieved from https://archive.mpi.nl/tla/elan

Ferrara, L. (2019). Coordinating signs and eye gaze in the depiction of directions and spatial scenes by fluent and L2 signers of Norwegian Sign Language. *Spatial Cognition & Computation*, *19*(3),220–251. https://doi.org/10.1080/13875868.2019.1572151

Ferrara, L., & Ringsø, T. (2019). Spatial Vantage Points in Norwegian Sign Language. *Open Linguistics*, *5*(1), 583–600. https://doi.org/10.1515/opli-2019-0032

Goldin-Meadow, S., & Brentari, D. (2017). Gesture, sign, and language: The coming of age of signlanguage and gesture studies. *Behavioral and Brain Sciences*, *40*, e46. https://doi.org/10.1017/S0140525X15001247

Herrmann, A., & Steinbach, M. (2012). Quotation in sign languages. *Quotatives: Cross-Linguistic and Cross-Disciplinary Perspectives. Amsterdam: John Benjamins*, 203–228.

Janzen, T. (2004). Space rotation, perspective shift, and verb morphology in ASL. *Cognitive Linguistics*, *15*(2). https://doi.org/10.1515/cogl.2004.006

Kuhn, J. (2017). *Telicity and iconic scales in ASL*. 30.

Lillo-Martin, D. (2013). The point of view predicate in American Sign Language. In *Language, gesture, and space* (pp. 165–180). Psychology Press. (Original work published 1995)

Loew, R. C. (1984). *Roles and reference in American Sign Language: A developmental perspective*. University of Minnesota.

Mandel, M. (1977). Iconic devices in American Sign Language. In L. A. Friedman (Ed.), *On the Other Hand: New Perspectives on American Sign Language* (pp. 57–107). Academic Press.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago press.

Meir, I. (2010). Iconicity and metaphor: Constraints on metaphorical extension of iconic forms. *Language*, *86*(4), 865–896. https://doi.org/10.1353/lan.2010.0044

Meir, I., Padden, C. A., Aronoff, M., & Sandler, W. (2007). Body as subject. *Journal of Linguistics*,*43*(3), 531–563. https://doi.org/10.1017/S0022226707004768

Meir, I., Padden, C., Aronoff, M., & Sandler, W. (2013). Competing iconicities in the structure oflanguages. *Cognitive Linguistics*, *24*(2). https://doi.org/10.1515/cog-2013-0010

Metzger, M. (1995). Constructed dialogue and constructed action in American Sign Language. *Sociolinguistics in Deaf Communities*, *1*, 255–271.

Occhino, C., Anible, B., Wilkinson, E., & Morford, J. P. (2017). Iconicity is in the eye of the beholder. *Gesture*, *16*(1), 100–126. https://doi.org/10.1075/gest.16.1.04occ

Ortega, G., & Morgan, G. (2015). Phonological Development in Hearing Learners of a Sign Language: The Influence of Phonological Parameters, Sign Complexity, and Iconicity: Phonological Development in Sign L2 Learners. *Language Learning*, *65*(3), 660–688. https://doi.org/10.1111/lang.12123

Padden, C. (1990). The relation between space and grammar in ASL verb morphology. *Sign Language Research: Theoretical Issues*, 118-132.

Padden, C. (1986). *Verbs and role-shifting in American Sign Language*. 44–57.

Padden, C. A., Meir, I., Hwang, S.-O., Lepic, R., Seegers, S., & Sampson, T. (2013). Patterned iconicity in sign language lexicons. *Gesture*, *13*(3), 287–308. https://doi.org/10.1075/gest.13.3.03pad

Padden, C., Hwang, S.-O., Lepic, R., & Seegers, S. (2015). Tools for Language: Patterned Iconicity in Sign Language Nouns and Verbs. *Topics in Cognitive Science*, *7*(1), 81–94. https://doi.org/10.1111/tops.12121

Perniss, P. M. (2007). Achieving spatial coherence in German Sign Language narratives: The use of classifiers and perspective. *Lingua*, *117*(7), 1315–1338. https://doi.org/10.1016/j.lingua.2005.06.013

Perniss, P. M. (2012). 19. Use of sign space. In R. Pfau, M. Steinbach, & B. Woll (Eds.), *Sign Language* (pp. 412–431). De Gruyter Mouton. https://doi.org/10.1515/9783110261325.412

Quer, J. (2018). On categorizing types of role shift in Sign languages. *Theoretical Linguistics*, *44*(3–4), 277–282. https://doi.org/10.1515/tl-2018-0020

Quinto-Pozos, D. (2007). Why Does Constructed Action Seem Obligatory? An Analysis of "Classifiers" and the Lack of Articulator-Referent Correspondence. *Sign Language Studies*, *7*(4),458–506. JSTOR.

Quinto-Pozos, D., & Parrill, F. (2015). Signers and Co-speech Gesturers Adopt Similar Strategies for Portraying Viewpoint in Narratives. *Topics in Cognitive Science*, *7*(1), 12–35. https://doi.org/10.1111/tops.12120

Schlenker, P. (2017). Super monsters II: Role Shift, iconicity and quotation in sign language. *Semantics and Pragmatics*, *10*(12), 1–67. https://doi.org/10.3765/sp.10.12

Supalla, Ted. (1982). *Structure and Acquisition of Verbs of Motion in American Sign Language*. University of California, San Diego.

Sweetser, E. E. (1988). Grammaticalization and semantic bleaching. *Proceedings of the FourteenthAnnual Meeting of the Berkeley Linguistics Society*, *14*, 39–405.

Sweetser, E. E. (2012). Introduction: viewpoint and perspective in language and gesture, from the Ground down. In Barbara Dancygier & Eve Sweetser (eds.), *Viewpoint in language: a multimodal perspective*, 1-22. Cambridge, UK: Cambridge University Press.

Taub, S. F. (2001). *Language from the Body: Iconicity and Metaphor in American Sign Language* (1sted.). Cambridge University Press. https://doi.org/10.1017/CBO9780511509629

Thompson, R., Emmorey, K., & Kluender, R. (2006). The Relationship between Eye Gaze and Verb Agreement in American Sign Language: An Eye-tracking Study. *Natural Language & LinguisticTheory*, *24*(2), 571–604. https://doi.org/10.1007/s11049-005-1829-y

Virnig, Dack (2011). *Scary Chicken Story* [Video]. YouTube. https://www.youtube.com/watch?v=xFbeQh0iuW0

Wilbur, R. (2003). Representations of telicity in ASL. *Proceedings from the Annual Meeting of theChicago Linguistic Society*, *39*, 354–368.

Wilbur, R. B. (2008). *Complex predicates involving events, time and aspect: Is this why sign languages look so similar?* 34.

Wulf, A., & Dudis, P. (2005). Body Partitioning in ASL Metaphorical Blends. *Sign Language Studies*, *5*(3), 317–332. https://doi.org/10.1353/sls.2005.0012

Zwitserlood, I. (2012). 8. Classifiers. In R. Pfau, M. Steinbach, & B. Woll (Eds.), *Sign Language* (pp. 158–186). De Gruyter Mouton. https://doi.org/10.1515/9783110261325.158

Appendices

Appendix A: List of all stimuli descriptions for Study I

Below is a list each stimulus the signers described for this analysis, which includes information about stimulus time (a video or still image) brief description of the stimulus, and where relevant, what kind of verb or verbs (and action) we had expected see to see in the description.

ID #	Type	Description of Stimulus	Target Verb
1	Video	A woman picks up a q-tip and breaks it into half	BREAK
2	Video	A man breaks a stick into three pieces and places it on a table.	BREAK
3	Video	A woman draws a single large star on a whiteboard.	DRAW
4	Video	A woman draws three asterisks on a whiteboard with a pen.	DRAW, WRITE
5	Video	A woman picks up a soda cup with a lid and straw, and drinks from it.	DRINK
6	Video	A woman takes a chip from a bowl of chips, dips it into a bowl of hummus, and takes a bite.	EAT
7	Video	A woman examines a pair of glasses, then puts them on.	PUT
8	Video	A woman bends a large plastic ladle/spoon.	*n/a*
9	Video	A woman reaches into a tissue box and takes out a tissue.	PICK, TAKE
10	Video	A woman picks up a pile of bookmarks, fans them out, then bunches them back up and places them into a cup.	PUT
11	Video	A man opens a book and flips through it to several different pages, as if skimming through it or searching for something.	READ
12	Video	A woman holds a sheet of paper in her hands – after a moment, she shakes her head, crumples the paper, and throws it to off to the side.	THROW
13	Picture	A woman has drawn a single large star on a whiteboard[1]	DRAW
14	Picture	A woman is putting a chip in her mouth to eat.[2]	EAT
15	Picture	A woman erases writing on a whiteboard (with an eraser).	ERASE
16	Picture	A man reads a small journal on couch	READ
17	Picture	A man stands over a sink, washing a bowl with a sponge.	WASH
18	Picture	A man writes in a blank notebook at a table.	WRITE

[1] In two cases, a picture stimulus is an image of the same scene as one of the video stimuli. The image for stimulus 14 is taken from video stimulus 3.

[2] The image for stimulus 15 is taken from video stimulus 6

Appendix B: List of lexical verbs found in Study I

Alphabetical list of all lexical verb glosses that were coded in the data, along with count of tokens of that glosses, the portion of verbs with each marker of constructed action, and iconicity score if available. In some cases (ERASE and THROW) two phonological variants with different handshape were observed, each variant is listed separately.

Lexical Verb Gloss	Token Count	Eye gaze shift	Facial/Bodily Enactment	Lateral Torso Shift	Iconicity Score
ANALYZE	2	0.50	0.50	0.	-
BECOME	1	0.	0.	0.	-
BITE	1	0.	0.	0.	-
BREAK	4	0.25	0.	0.	5.97
DECIDE	4	0.50	0.	0.	-
DO-SOMETHING	1	0.	0.	0.	-
DOUBT	1	0.	0.	0.	2.19
DRAW	12	0.	0.08	0.	5.00
DRINK	4	0.	0.	0.25	7.00
DROP	1	1.00	0.	0.	5.50
EAT	5	0.	0.	0.20	-
ERASE (HS: A)[3]	4	0.25	0.	0.	-
ERASE (HS: S)	1	0.	0.	0.	-
FACE-LOOK	2	0.	1.00	1.00	-
FINISH	1	0.	0.	0.	2.93
GET	1	0.	0.	0.	4.53
GO-AHEAD	9	0.	0.	0.	-
HAVE	29	0.17	0.03	0.	3.53
HOLD	5	0.20	0.	0.	-
LAUGH	1	0.	0.	0.	4.75
LOOK-LIKE	3	0.	0.	0.	-
MAKE	1	0.	0.	0.	4.63
MAKE-A-FACE	3	0.67	1.00	0.33	-
PICK	3	0.33	0.	0.33	-
PUT	1	1.00	0.	0.	-
READ	13	0.31	0.15	0.31	4.57
SEARCH	3	0.33	0.33	0.33	4.57

[3] ERASE appeared two variants in this data, one with the handshape 1 ('A' variant), and one with 6 ('S' variant)

SEE	3	0.67	0.67	0.	4.63
STRUGGLE	3	0.33	0.	0.67	-
SUCK	4	0.50	1.00	0.50	-
SUMMARIZE	1	1.00	0.	0.	3.04
THROW (HS: 1)[4]	3	0.33	0.33	0.	-
THROW (HS: U)	2	0.	0.50	1.00	5.51
TRY	2	0.50	0.	0.	1.63
UNDERSTAND	1	0.	0.	0.	4.63
USE	1	0.	0.	0.	1.67
WANT	1	0.	0.	0.	4.58
WASH	16	0.25	0.06	0.	-
WRITE	9	0.11	0.22	0.11	6.57

[4] THROW also appears with two variants: both begin with a fist handshape 6 but the movement-final handshape is either the index and middle finger extended together, without spread, ('U' variant) or an extended index finger alone ('1' variant).