

# The challenges of viewpoint-taking when learning a sign language: Data from the ‘frog story’ in British Sign Language

Second Language Research

1–33

© The Author(s) 2020



Article reuse guidelines:

[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)

DOI: 10.1177/0267658320906855

[journals.sagepub.com/home/slr](https://journals.sagepub.com/home/slr)**Sannah Gulamani**

Deafness, Cognition and Language Research Centre, University College London, UK

**Chloë Marshall** 

Deafness, Cognition and Language Research Centre, and Institute of Education, University College London, UK

**Gary Morgan**

Deafness, Cognition and Language Research Centre, University College London, and Division of Language and Communication Science, City University of London, UK

## Abstract

Little is known about how hearing adults learn sign languages. Our objective in this study was to investigate how learners of British Sign Language (BSL) produce narratives, and we focused in particular on viewpoint-taking. Twenty-three intermediate-level learners of BSL and 10 deaf native/early signers produced a narrative in BSL using the wordless picture book *Frog, where are you?* (Mayer, 1969). We selected specific episodes from part of the book that provided rich opportunities for shifting between different characters and taking on different viewpoints. We coded for details of story content, the frequency with which different viewpoints were used and how long those viewpoints were used for, and the numbers of articulators that were used simultaneously. We found that even though learners’ and deaf signers’ narratives did not differ in overall duration, learners’ narratives had less content. Learners used character viewpoint less frequently than deaf signers. Although learners spent just as long as deaf signers in character viewpoint, they spent longer than deaf signers in observer viewpoint. Together, these findings suggest that character viewpoint was harder than observer viewpoint for learners. Furthermore,

---

## Corresponding author:

Chloë Marshall, UCL Institute of Education, University College London, 20 Bedford Way, London, WC1H 0AL, UK.

Email: [c.marshall@ioe.ac.uk](mailto:c.marshall@ioe.ac.uk)

learners were less skilled than deaf signers in using multiple articulators simultaneously. We conclude that challenges for learners of sign include taking character viewpoint when narrating a story and encoding information across multiple articulators simultaneously.

## Keywords

character viewpoint, co-speech gesture, narrative, observer viewpoint, sign language, viewpoint

## 1 Introduction

Narrating a story is one of the most challenging aspects of language production in the acquisition of one's first and subsequent languages, and its development spans a protracted period (for first language acquisition, see Berman and Slobin, 1994; for second language acquisition, see, amongst others, Ellis, 1987; Ortega, 2000; Robinson, 1995). Constructing a narrative is challenging because language and cognitive resources come to bear at the same time: several higher-level language and cognitive skills are necessary to form cohesive, coherent and structured narratives (Bamberg and Damrad-Frye, 1991). These include the mastery of a variety of linguistic (lexical, syntactic and pragmatic) skills, the ability to remember and sequence a series of events, and the ability to establish and maintain the different viewpoints of a range of characters (Norbury et al., 2014). As such, narrative is an interesting area to study in second language acquisition because it can provide a rich source of data on a range of linguistic and cognitive phenomena and can be used as an index of language fluency when late learners' performance is compared with that of native/early learners.

In the current study, we investigate how well hearing adults learning British Sign Language (BSL) are able to produce a signed narrative. Our particular focus is on how learners take on and shift between character or observer viewpoints during narrative production. Despite growing numbers of learners (e.g. Furman et al., 2010), there is a paucity of research into the learning of sign languages by hearing adults, who are learning not just a new language, of course, but a new language that is produced with a set of articulators not used for spoken language.<sup>1</sup> Our study aims to contribute to this novel but growing literature on learners of a sign language by focusing on several aspects of narrative production. We would expect some aspects of narrative to be challenging for second language learners whatever the language being learnt and whatever the modality. An obvious example is that we would predict hearing learners to produce less detailed narratives than deaf fluent signers, missing out details of important information such as the characters, settings, initiating events, internal responses, plans and consequences. Such omissions would likely stem from increased demands being placed on working memory by accessing a less developed language system (Miyake and Friedman, 1998; Service, 1992). Other challenges might be more specific to telling a signed story; for example, the requirement to encode language information across multiple articulators, i.e. the two hands, body, eyebrows, mouth, eyes and head (Sandler et al., 2011). Another aspect of narrative that we predict to be challenging is the use of different viewpoints. There are, as we explain below, both similarities and differences in how viewpoints are expressed across language modalities. We then review additional relevant research literature in order to motivate our predictions for which aspects of narrative learners will find challenging.

## II Narrative production

### *I Expressing viewpoint in spoken language and co-speech gesture*

When speakers recount a story, they can adopt different viewpoints, by being either inside the scene taking the viewpoint of the characters themselves, or outside the scene observing the characters. They can also switch between viewpoints during their narrative, i.e. from one character to another, and from character to observer viewpoint and vice versa. There is also the possibility of combining two viewpoints, i.e. to use character and observer viewpoints simultaneously (an option we term ‘mixed viewpoint’ in the current study). Viewpoint can be expressed linguistically (e.g. by using direct and indirect speech) and through gesture. Speakers can (and frequently do) use co-speech gesture with their spoken narratives, and these co-speech gestures vary in form depending on whether they accompany speech in character viewpoint (where the form of the gesture shows that the speaker is acting as the character in the scene) or whether the gestures represent an observer’s viewpoint (the form of the gesture shows that the speaker is observing the scene as though it is unfolding on a smaller scale in front of them) (McNeill, 1992).

For example, during the course of narrating a story, a speaker might hold her fists out as though gripping the handlebars of a bike and move her fists to show the rider wobbling and then falling (character viewpoint), or she might use a flat hand which moves in the space in front of her to show the wobbling and then falling bike on a smaller scale (observer viewpoint). Alternatively, she might use both viewpoints sequentially, perhaps starting in character viewpoint for the wobbling event and then finishing in observer viewpoint for the falling event. Furthermore, manual gestures are not the only gestural resources that can be recruited by narrators. For example, facial gestures to show emotional affect – such as expressions of surprise, pain or sympathy – might accompany the telling of the unfortunate cyclist’s fall. Likewise, when reporting dialogue between two or more characters, speakers have been reported to use facial expression, eye gaze, and torso position, in addition to manual gestures (Clark, 2016; Stec et al., 2017). While co-speech gesture is a rich resource for speakers to use when recounting a narrative, using articulators beyond the voice to portray viewpoints is optional and not conventionalized (i.e. speech can describe viewpoint shifting without gesture, and there is no grammar constraining how gesture should or should not be used). In contrast, for the same discourse function – viewpoint-taking – movements of the body, face and hands are both necessary and conventionalized in a set of human languages produced uniquely in the visual modality, i.e. in sign languages.

### *2 Expressing viewpoint in sign languages*

Like speakers, narrators who use a sign language rather than a spoken language can also adopt character and observer viewpoints<sup>2</sup> when relating a narrative (Özyürek and Perniss, 2011), and many of the devices that they employ look similar to those used by speakers in their (optional) co-speech gesture. For example, Cormier et al. (2012a) have argued that the handling classifiers used by signers to indicate what the referent’s hands are doing during an event are visually similar to the character viewpoint gestures used by

speakers. Several sign language researchers use the term ‘constructed action’ to describe the signer showing the character’s viewpoint through an acting out of the actions of a referent (e.g. Perniss, 2007; Stec, 2012). Thus, the signer portrays character viewpoint by the use of constructed action and this is visually and functionally similar to the set of co-speech gestures a speaker uses when acting as the character in a scene as described previously. The way in which signers use their body to show not only the character’s actions but also the emotional reactions (Metzger, 1995) has similarities to the gesturer who is in character viewpoint. And like speakers, signers also use eye gaze and torso position to shift between different roles when in character viewpoint (Engberg-Pedersen, 1995; Loew, 1984; Padden, 1986). Similarly, the entity classifiers or ‘depicting signs’ (Perniss, 2007; Pyers et al., 2015) are used by signers to represent an entire referent and to describe its orientation, location, and movement on a smaller scale in front of them. The presence of these forms portrays the observer viewpoint, and they are equivalent to speakers’ observer viewpoint gestures (Cormier et al., 2012a).

Signers are therefore likely to do quite similar things to co-speech gesturers when describing the wobbling and falling bike incident mentioned earlier (e.g. the American Sign Language example of a motorcyclist discussed by Dudis, 2004), and studies that have directly compared signers and speakers narrating the same materials have shown that they do indeed have a similar-looking repertoire of forms (Earis and Cormier, 2013; Quinto-Pozos and Parrill, 2015). For example, Earis and Cormier (2013) found that the two spoken language and two sign language narrators in their study used similar facial expressions of affect. Meanwhile, the co-speech gesturers and signers in Quinto-Pozos and Parrill’s (2015) study tended to use character or observer viewpoint for descriptions of the same events. This finding suggests that properties of the events may provoke the same viewpoint choice from both groups of language users. Similarities between signers and co-speech gesturers are perhaps not surprising, because they are likely the result of how sign languages evolved and how human beings conceptualize space for communication (Emmorey et al., 2000).

While there are similarities between viewpoint-taking in co-speech gesture and sign languages, there are also some important differences. There are cross-modality differences as to which viewpoint is preferred. Özyürek and Perniss (2011) found that signers preferred to use character viewpoint more than observer viewpoint, a finding consistent with the signers in Earis and Cormier’s study (2013). This preference appears to be in contrast to speakers, who prefer observer viewpoint (Earis and Cormier, 2013). Parrill (2010) also found a preference for observer viewpoint in speakers, but this was modulated by linguistic, event and discourse structure. Taken together, these findings might suggest that learners of sign need to learn to use character viewpoint to a greater extent than they would be used to doing in spoken language, and this is a potential challenge that needs to be overcome. Furthermore, although mixed viewpoint has been previously described in signed languages (Aarons and Morgan, 2003; Quinto-Pozos and Parrill, 2015), it is reported to be rarer in co-speech gesturers (Parrill et al., 2016).

A further difference between co-speech gesture and sign languages is conventionalization (Perniss et al., 2015). As Earis and Cormier (2013) argue when comparing narratives in English and BSL, both languages have a pool of resources they can draw on, but some of the elements used by signers to depict characters’ actions are conventionalized

in a way that those used by speakers are not. They note that although the depiction of characters using expressive elements such as co-speech gesture does not always occur in spoken English, depicting characters with character viewpoint is a very important element of storytelling in sign. Among many comparisons between their two speaking and two signing participants, Earis and Cormier (2013) discuss the example of a speaker who used eye gaze at one point in the narrative in the same way that signers did, and yet did not do this consistently throughout their narrative.

A major difference between co-speech gesture and sign, therefore, is that co-speech gesture is both optional and unconstrained by its own grammar. While gesture does pattern systematically, the system it follows is that of the spoken language in which it is embedded (Kita, 2014; Kita and Özyürek, 2003). In contrast, the manual and non-manual articulators in sign languages are the only ones available to signers for expressing viewpoint. They are not combined with or subordinate to another language system, and because of this obligation the patterns for how these articulators are used – often simultaneously – have become both conventionalized (shared across a community of signers) and constrained by the individual's sign grammar (the language-internal regularities developed during first language acquisition).

An obvious question then arises because of the similarities and differences in how viewpoint is used across modalities: what happens when hearing adults, who likely use co-speech gesture when narrating, choose to learn a sign language where these devices have become conventionalized? What can sign language learners borrow from gesture directly, what needs to be adapted, and what completely novel-to-them language tools need to be learned? Compared with spoken second language research there have been relatively few studies of L2 learners of sign. There are a small number of studies, however, which can be drawn upon for making predictions of how learners will produce different viewpoints in a signed narrative. In the remainder of the introduction we review some relevant studies in order to motivate our research questions and our predictions.

### ***3 What might learners of a sign language find challenging with respect to narrative production?***

Anecdotal evidence from sign language tutors and learners suggests that taking on different viewpoints is something that hearing learners of sign find particularly challenging. These challenges have also been noted by Quinto-Pozos (2008) and Schornstein (2005). On the other hand, given the afore-mentioned similarities in the strategies used by signers and co-speech gesturers, hearing learners of sign do presumably have a set of existing and appropriate gesture-based forms on which to draw when narrating a story in sign.

At a gross level, it has been argued that despite the modality difference between signed and spoken language, there is the opportunity of transfer from learners' spoken L1 to their new signed L2, both of abstract linguistic entities (e.g. pragmatic constraints on reference, Bel et al., 2015) and of gesture (Chen Pichler and Koulidobrova, 2016). Alongside transfer is the possible existence of language learning universals, i.e. factors that will influence L2 learning whatever the L1 or the new language being learned (Chen Pichler and Koulidobrova, 2016). One such universal tendency widely reported in the L2 spoken language narrative literature concerns referential cohesion (Halliday and Hasan,

1976), and this has also been the subject of two investigations in hearing learners of sign (Bel et al., 2015; Frederiksen and Mayberry, 2019). The narrator needs to construct their narrative in such a way that the listener can follow which entities are involved in which actions as the story progresses. Achieving cohesion requires the narrator to introduce referents and refer back to them appropriately as needed, and it involves a range of different linguistic forms, including nominal phrases, pronouns and null elements, whose use is, to a certain extent, language-specific.

L2 learners of spoken languages tend to adopt a strategy of overproducing overt forms when marking reference in order to reduce ambiguity (Gullberg, 2006; Hendriks, 2003). Bel et al. (2015) found that L2 learners of sign also used this over-redundancy strategy. They overused overt pronouns and underused null pronouns for maintaining reference, in comparison to native signers, although these group differences did not reach statistical significance. In another study of reference tracking, Frederiksen and Mayberry (2019) found that L2 learners of sign did not reliably overuse nominals and pronouns, nor underuse null elements compared to the native signers. They did, however, use classifiers a little differently to native signers, using them to reintroduce referents when native signers did not.

As for the particular forms that learners of sign use for entity classifiers, two studies provide some initial evidence that these elements can be challenging for learners (Boers-Visker and van den Bogaerde, 2018; Marshall and Morgan, 2015). Marshall and Morgan (2015) found that even within a highly structured elicitation task that did not require production of a narrative, learners frequently selected an incorrect classifier handshape. Using a different elicitation task, Boers-Visker and van den Bogaerde (2018) also found that learners experienced difficulties with entity classifiers. For example, learners frequently expressed referents sequentially instead of simultaneously, and, as a result, did not always establish spatial relationships between referents correctly.

Also on the theme of expressing spatial relationships, findings by Ferrara and Nilsson (2017) are relevant. They asked learners of sign to describe spatial layouts of routes and buildings with which they were familiar, and found that learners not only used fewer classifiers than native signers but also struggled to coordinate classifiers within the signing space and relative to their own bodies. Intriguingly, and relevant to the previous discussion about viewpoints, Ferrara and Nilsson's learners made different choices to native signers when signing about routes. Whereas native signers chose a route viewpoint (describing the spatial layout as though moving through it themselves), learners chose to describe the layout from above, statically. This is consistent with the aforementioned findings of Earis and Cormier (2013) and Parrill (2010) for co-gesturing speakers, who showed a preference for observer over character viewpoint.

The picture that emerges from the literature reviewed here is that there are likely to be challenges for learners of sign in taking on different viewpoints in signed narratives. Despite multi-modal articulation occurring in spoken language narratives, it is not conventionalized in the same way that it is for sign languages, and this is likely to be difficult for learners. There are likely, for example, to be challenges posed by the cognitive complexity of having to maintain several viewpoints whilst selecting which linguistic resources to use at any one time.

In this study we examine how hearing second language learners take on viewpoints during the telling of a BSL narrative (based on the widely-used *Frog story*), in comparison

to native/early signers of BSL. Firstly, we ask two fundamental questions about the quantity and quality of their narratives. We ask:

1. Does the length of the narratives produced by both groups differ?
2. Do learners differ from deaf signers in the amount of detail they provide in their narratives?

We then ask two questions related to investigating how successfully learners use different viewpoints:

3. We look at two aspects of viewpoint use, namely how frequently learners adopt each viewpoint, and the total time spent in each viewpoint. To do this we counted each time a signer expressed information from a specific viewpoint; and when in a specific viewpoint we timed how long this viewpoint was adopted for. Specifically, we ask: Do learners differ from deaf signers in the frequency with which they use character viewpoint, observer viewpoint and mixed viewpoint, and do they differ in the total amount of time they spend in each of those viewpoints?
4. Do learners differ from deaf signers in how much information they can express across different articulators simultaneously?

On the basis of the literature discussed in this introduction, we predict that learners will be able to adopt and shift between character and observer viewpoints in their narratives, and that they will do so using manual and non-manual articulators. However, we also predict that they will find these aspects of narrative challenging, as evidenced by less detail in their narratives, a reluctance to use character viewpoint (evidenced, for example, by lower frequency and/or shorter total duration of character viewpoint compared to deaf signers), and lower levels of simultaneous articulator use. We make no *a priori* predictions about the length of their narratives, because the challenges that learners face might result in them needing longer in which to sign their narratives, or alternatively they might produce shorter narratives in order to avoid these challenges.

### III Method

#### I Participants

Twenty-three hearing adults who were learning BSL (4 males) and ten deaf native/early<sup>3</sup> users (5 males), aged 19–53 years (deaf: mean = 30.7, SD = 8.6; hearing: mean = 32.2, SD = 9.7)<sup>4</sup> participated in the study. They were recruited via social media, university networks, university programs involving sign language studies, BSL learner forums, BSL classes, connections via the deaf community, deaf schools and a database of people who had given their consent to be contacted regarding opportunities for participating in sign language research. They gave informed consent for their narratives to be video-recorded for later coding and analysis, and the study was given ethical approval by the Ethics Review Board of City, University of London. Participants were remunerated for their time.

Narrative requires a certain level of BSL ability, and we wanted our learner participants to have at least some knowledge of how to use different viewpoints. Learners were therefore recruited either from sign language university programs after they had received at least 120 hours of BSL instruction, or, if they were not studying BSL at university, they had a minimum of a Level 2 Certificate in BSL awarded by Signature ([www.signature.org.uk](http://www.signature.org.uk)), and so had likely received around 160 hours of BSL instruction. All learners could be classed as ‘intermediate learners’ of BSL. They did, however, vary in their usage of BSL; for most it was daily, for others no more than once or twice a month. Further details of the participants in both groups are given in Table 1.

## 2 Materials

The 30-page wordless picture book *Frog, where are you?* (Mayer, 1969) was used to elicit signed narratives. The story concerns a boy and his dog who are searching for a missing pet frog and who encounter various forest animals that in some way interfere with their search. After several such encounters, the boy and dog eventually find the frog with a mate and a family of baby frogs. The story ends with the boy and dog leaving for home with one of the babies as their new pet frog. This book has been used extensively in cross-linguistic work (Berman and Slobin, 1994), including previous studies of sign languages (e.g. Australian Sign Language: Hodge et al., 2019; BSL: Morgan, 2002; Danish Sign Language: Engberg-Pederson, 2003).

Although participants were required to tell the whole story (see Section III.3) the analysis focused on just one portion of the story, namely the third part (pages 17–23), because it contained multiple opportunities for narrators to use different viewpoint and to potentially express two or more characters at the same time. In this part of the story, an owl chases the boy all the way to a large rock. The boy climbs onto the rock and grabs onto some branches so that he does not fall, but the branches turn out to be a deer’s antlers, and the deer picks the boy up onto his head. The deer starts running with the boy on his head, and the dog runs alongside them. When they arrive at the edge of a cliff, the deer stops suddenly and boy and the dog both fall into the pond below. The third part ends with both the boy and the dog hearing a familiar sound (which turns out to be the pet frog). In this portion of the story there are therefore episodes involving more than two protagonists simultaneously, which we expected to challenge learners.

## 3 Procedure

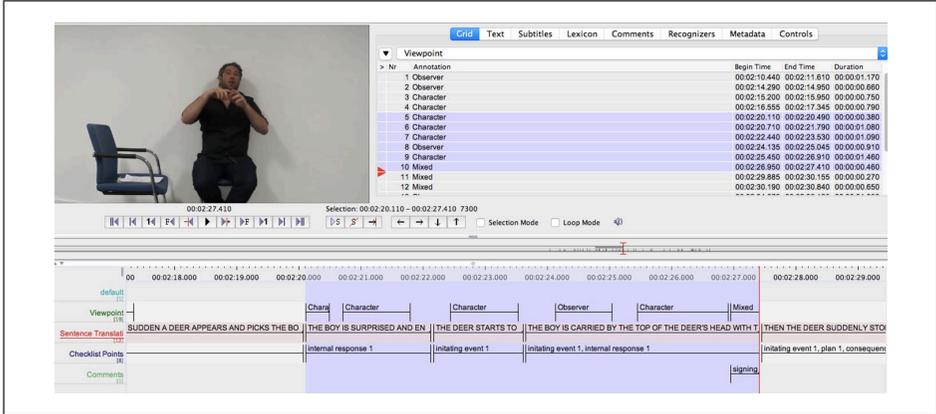
Participants were tested individually. A deaf experimenter (first author) conducted the study and the filming in BSL, taking care not to disclose too much information about the aims of the study so as not to influence participants’ signing behavior and response style. Prior to the filming, participants were asked to fill in a questionnaire requesting background information, and to provide verbal and written consent for the study to be conducted. Participants were given a hard copy of the *Frog story* book and asked to first browse through the whole book. In order to ease memory load, the story was split into four parts (where natural breaks occurred so as to minimize the disruption of narrative

**Table 1.** Participant details.

Participant code	Deaf (D) or Hearing (H)	Sex (male or female)	Age at test (years)	Age started learning BSL (years)	Number of years using BSL	Level of BSL	Regularity of BSL use
D001	D	F	19	4	15	Fluent	Daily
D002	D	M	34	Birth	34	Fluent	Daily
D003	D	F	35	Birth	35	Fluent	Daily
D004	D	M	33	Birth	33	Fluent	Daily
D005	D	F	31	4	27	Fluent	Daily
D006	D	F	29	Birth	29	Fluent	Daily
D007	D	M	Missing data			Fluent	Daily
D008	D	M	26	1	25	Fluent	Daily
D009	D	F	48	Birth	48	Fluent	Daily
D010	D	M	21	1.5	20	Fluent	Daily
H001	H	F	20	18	2	3rd year student	Daily
H002	H	M	25	23	2	2nd year student	Daily
H003	H	F	50	45	5	4th year student	Daily
H004	H	F	34	19	15	Level 2	Twice a month
H005	H	F	36	22	6	Pre-Level 3	Once a month
H006	H	F	29	25	4	Level 3	3 days a week
H007	H	F	32	26	6	Pre-Level 6	Daily
H008	H	F	28	Birth	28	4th year student	Daily
H009	H	F	Missing data		9	4th year student	Daily
H010	H	F	21	19	2	2nd year student	Daily
H011	H	F	29	17	12	2nd year student	Daily
H012	H	M	21	17	4	4th year student	Daily
H013	H	F	20	16	4	4th year student	Daily
H014	H	F	30	28	2	Level 2	Weekly
H015	H	M	31	25	6	Level 3	Daily
H016	H	F	44	16	28	Pre-Level 6	Daily
H017	H	F	53	40	13	Level 2	Daily
H018	H	F	35	26	9	Level 3	Daily
H019	H	F	50	45	5	Level 2	Daily
H020	H	M	28	19	9	Level 3	Daily
H021	H	F	38	14	24	Level 3	Daily
H022	H	F	27	25	2	Level 2	Weekly
H023	H	F	28	18	10	Level 3	Daily

Note. BSL = British Sign Language

flow) for retell. Participants had the opportunity to browse through the scenes of each specific part before retelling it to the experimenter. When the study had finished, the experimenter spent some time with the participants to explain further details of the



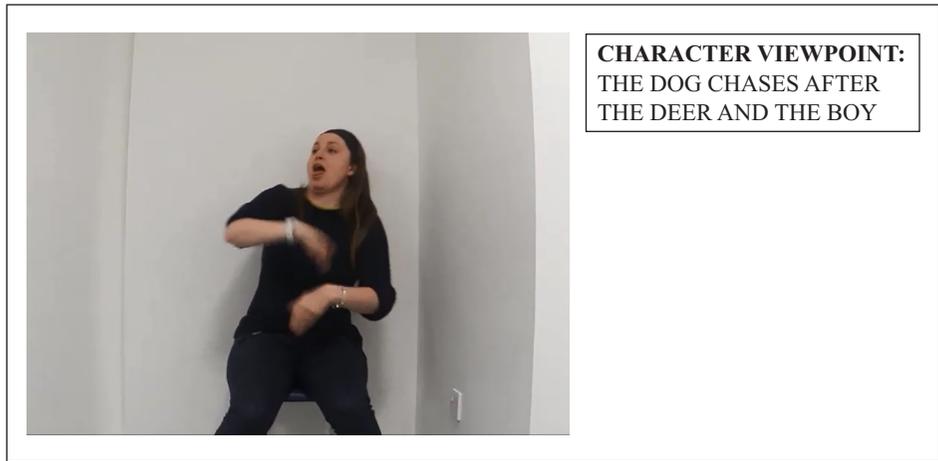
**Figure 1.** Screenshot of the ELAN tiers used for coding the information content and viewpoint of each narrative.

research aims and to provide the opportunity to raise any queries. Overall, the testing period lasted approximately one hour.

#### 4 Coding

Coding was done using ELAN 4.9.4 (<https://tla.mpi.nl/tools/tla-tools/elan>), software that allows for the creation of complex annotations on video resources. A template ELAN file was created with tiers and controlled vocabularies, and this template was applied to all files to enable consistency in tier structures and naming. Tiers were designed to provide information pertinent to the planned analyses: information content and the level of detail provided at a macrostructure level; frequency and duration of the three different viewpoints (character, observer and mixed); and production of meaningful<sup>5</sup> simultaneous articulators (dominant hand, non-dominant hand, body, eyebrows, mouth, eyes and head). To retrieve data for each analysis, information was extracted from each participant file using the Annotation Statistics function in ELAN.

*a Coding information content.* We developed a coding procedure based on the story script for *Frog, where are you?* from the manual for the Systematic Analysis of Language Transcripts (SALT, 2009: 249–50) and from the story elements coded by Squires et al. (2014). This involved counting unique tokens of six different macrostructure elements: characters, settings, initiating events, internal response, plans and consequences. A checklist of possible matches or combinations (single words or sentences) for each of these six elements was drawn up (for details, see Appendix 1). Two ELAN tiers were created: Sentence Translation and Checklist Points (see Figure 1). In the Sentence Translation tier we created a series of English glosses, as faithful to the original BSL as possible, in order to help with the coding of the informational content. In the Checklist Points tier we coded the number of items of informational content provided (according to the list provided in Appendix 1). A score of 1 was awarded for each



**Figure 2.** Example of a deaf signer using character viewpoint.

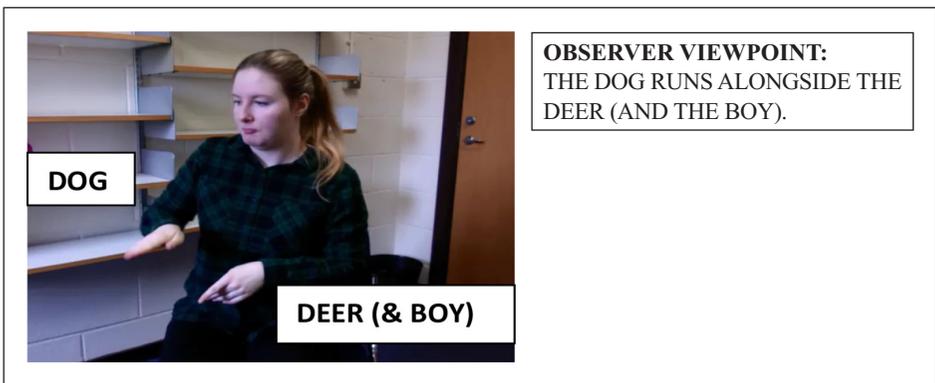
full item on the checklist that was provided, with a score of 0.5 for each item that was produced but not fully.<sup>6</sup>

*b Coding viewpoints.* In order to code viewpoints, we referred to Parrill's (2010) definitions for co-speech gesture, and adapted them for BSL (for further details, see Appendix 2). A new annotation segment was created on the Viewpoint tier from the start to the end of each viewpoint (see Figure 1). Viewpoint, therefore, could occur over a single sign, a series of signs, or across one or more whole sentences. If no viewpoint occurred, no annotation segment was created, and therefore it was considered that the signer was in a narrative (default) mode. A new annotation was added after to a previous annotation if there was a switch in viewpoint, or where the same viewpoint was maintained but where the information and production changed (e.g. a different character was introduced or where the same character was maintained but there was a change in articulators).

A controlled vocabulary was used, with the terms 'character', 'observer' and 'mixed'. Character viewpoint was coded when the signer was producing an internal, first person point of view which prompted the use of actions, dialogue, handling, body portioning productions, and displays of affect of a character. Generally in this viewpoint the character's head and torso are mapped to the signer's body and the size of the projected space is life-sized (see Figures 2 and 3 for two examples from deaf signers). Observer viewpoint was coded when the signer was producing an external, vantage, third person point of view which prompted the use of entity classifiers and topographic space and which gave information about path, trajectory, speed and location. In observer viewpoint the signer is not part of the event (in contrast to when they are when in character viewpoint), and so the event space is reduced and projected onto the area of space (a smaller scale) in front of the signer's body (see Figures 4 and 5 for two examples from learners). Mixed viewpoint simultaneously encodes both a character and observer point of view. For example, on occasion the deer and boy characters were represented differently: one hand



**Figure 3.** Example of a deaf signer using character viewpoint.



**Figure 4.** Example of a learner using observer viewpoint.

used an entity classifier to represent the boy on top of the deer and the other represented the antlers of the deer whilst the signer simultaneously depicted the physical movement of the deer's head. Such examples were coded as mixed viewpoint. We give a more detailed example in Section III.4.c to illustrate the richness of the information that signers were able to convey by mixing character and observer viewpoint.

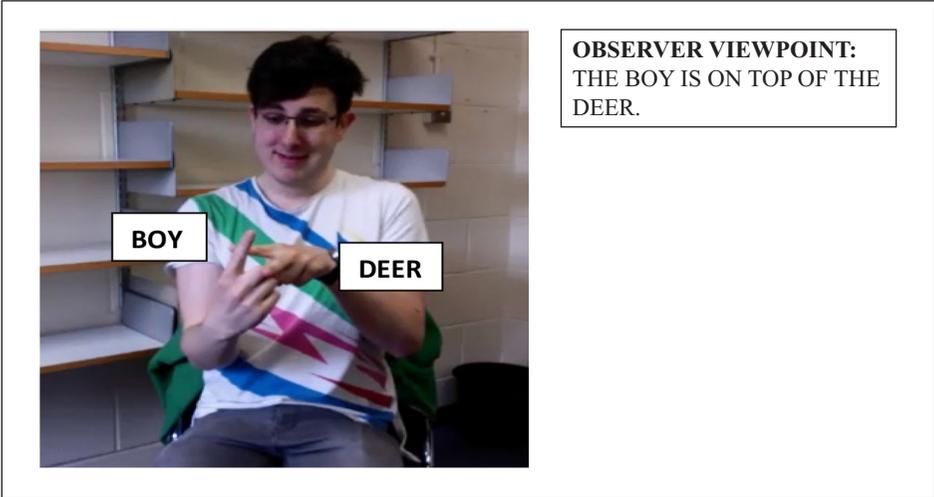


Figure 5. Example of a learner using observer viewpoint.

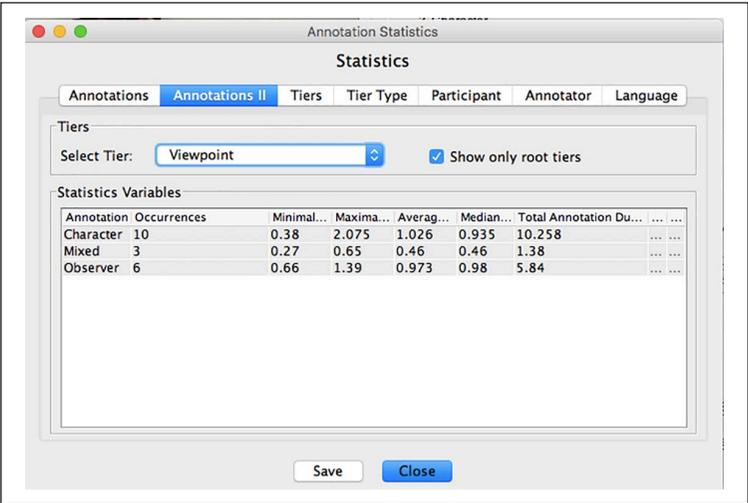
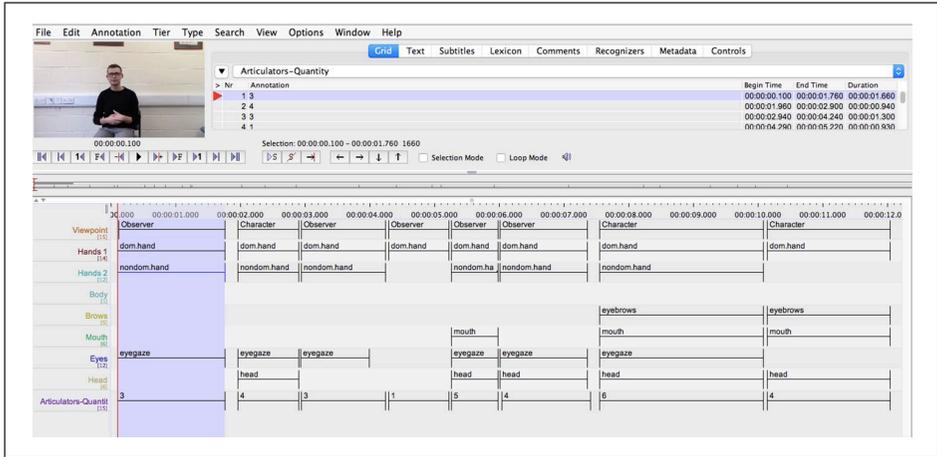


Figure 6. Screenshot of ELAN Annotation Statistics showing how occurrences and duration of the three different viewpoints were calculated.

The Annotation Statistics tool in ELAN was used to calculate, for each participant, the frequency of occurrence of each viewpoint (i.e. the number of annotations for each viewpoint) and the duration of each viewpoint in seconds (see Figure 6).

c *Coding articulators.* Tiers were developed for each of the seven articulators (i.e. dominant hand, non-dominant hand, body, eyebrows, mouth, eyes, head). See Appendix 3. Coders



**Figure 7.** Screenshot of the ELAN tiers used for coding the articulators used in each viewpoint.

inspected each viewpoint to mark which articulators were present and meaningful. An Articulator-Quantity tier with a controlled vocabulary of ‘1’-‘7’ was used to record the total number of articulators per viewpoint (see Figure 7). If any articulator was present in a particular viewpoint and continued into the next viewpoint with the same meaning and no additional information/description, then this articulator was only coded in the first viewpoint. To illustrate how articulators were used simultaneously, we provide two examples in Figures 8 and 9. These show a deaf signer and a learner, respectively, narrating the part of the story where the boy finds himself on top of the deer. While the deaf signer is using all seven articulators to represent the boy (dominant hand, mouth, eyes and eyebrows) and the deer (non-dominant hand, body and head), and is using mixed viewpoint, the learner is using just her hands, one for the boy and one for the deer (and is using only observer viewpoint).

All the data were coded by a fluent BSL user (first author). Any uncertainties were marked and a description of these was included in the Comments Tier. Uncertainties were resolved when revisiting files at a later date or in discussion with a second coder during reliability coding. A proportion of the data was additionally coded by a second fluent BSL user who was otherwise uninvolved with the study (existing annotations were hidden before this second coder started coding, so that their coding was independent), in order to calculate inter-rater reliability. Inter-rater reliability was calculated on the basis of data from 3 learners and 3 deaf signers, respectively representing 13% and 30% of participants in those groups. For the detail captured by the narratives, an intraclass correlation to measure consistency of coding using a two-way random effects model (where both people effects and measures effects are random) revealed excellent inter-rater consistency for the learners’ data, intraclass correlation = .995 (95% CI = .986 – .998),  $p < .001$ , and also for the deaf signers’ data, intraclass correlation = .963 (95% CI = .901 – .986),  $p < .001$ . For the frequency of viewpoints, there was again excellent consistency: for learners, intraclass correlation = .998 (95% CI = .991 – 1.000),  $p < .001$ , and for deaf signers, intraclass correlation = .997 (95% CI = .986 – .999),  $p < .001$ . Consistency

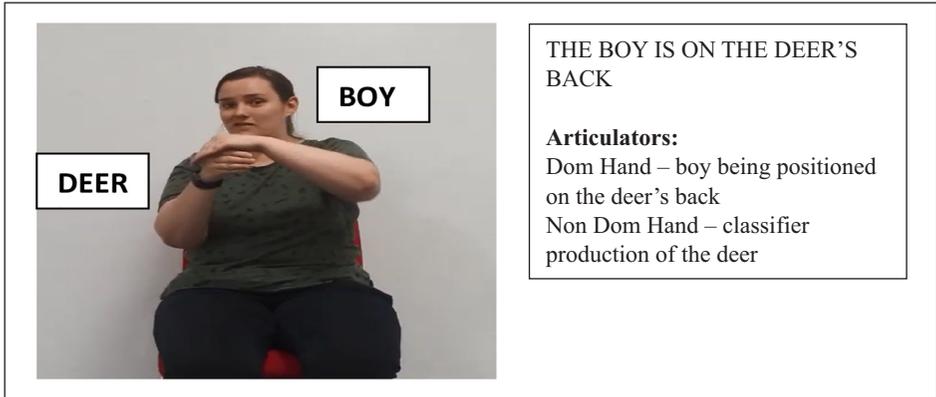


**THE BOY IS ON THE DEER'S BACK**

**Articulators:**

- Dom Hand – boy being positioned the deer's back
- Non Dom Hand – deer
- Body – showing motion of the deer moving
- Head – showing motion of the deer's head moving
- Mouth – showing reaction of the boy when being unsettled on the deer's back
- Eyes – looking upwards and then across to show view of the boy
- Eyebrows – frowning to show boy's reaction

**Figure 8.** Example of a deaf signer using 7 articulators simultaneously.



**Figure 9.** Example of a learner using 2 articulators simultaneously.

was also excellent for the number of articulators for both learners, intraclass correlation = .971 (95% CI = .929 – .988),  $p < .001$ , and for deaf signers, intraclass correlation = .981 (95% CI = .953 – .992),  $p < .001$ . Any areas of disagreement were revisited and discussed until final agreement was achieved.

## IV Results

We compared the two groups on a number of different measures. Unless data were non-normally distributed, we used parametric tests (independent measures  $t$ -tests, ANOVA) in our analyses, despite the uneven group sizes, because running the analyses with non-parametric versions of these tests gave the same results.

### *Analysis 1: Duration of the narratives*

First, we report on the duration of the section of the narrative that was coded for analysis. For the deaf group, the mean duration was 41.80 seconds (SD = 13.46; range = 26–70), and for the learners it was 50.48 seconds (SD = 24.12; range = 16–119). An independent samples  $t$ -test revealed that this difference in duration was not significant,  $t(31) = -1.06$ ,  $p = .297$ . The similarity in overall durations means that our analyses could proceed using raw counts for frequency and total duration of viewpoints, without needing to administer a correction for differing opportunities to produce the measures of interest. Although the duration of the narratives varied greatly between participants, this was the case within both groups.

### *Analysis 2: Information content in the narratives*

Next, we report on the information contained in the two groups' narratives. These data are presented in Table 2, together with the results of a set of non-parametric independent samples  $t$ -tests. The analysis shows that the learners expressed less information than the deaf signers with respect to all the features of the narrative measured in this analysis, namely the numbers of characters, settings, initiating events, internal responses, plans and consequences.

**Table 2.** Information content in the narratives.

Information contained in the narrative	Deaf signers (n = 10)			Learners (n = 23)			Results of non-parametric independent samples tests							
	Mean	SD	Median	Min	Max	Mean	SD	Median	Min	Max	Mann-Whitney U	Wilcoxon W	Z	p
Number of characters	3.95	0.69	4	3	5	2.96	0.72	3	1	4	39.5	315.0	3.16**	.002
Number of settings	4.70	0.95	5	3	6	3.35	1.46	3	1	7	47.0	323.0	2.73**	.006
Number of initiating events	8.85	2.29	8.5	6	14	5.67	2.12	5	2	9	35.5	311.5	3.13**	.002
Number of internal responses	3.80	1.87	3	1	7	1.85	1.29	2	0	4	44.5	320.5	2.81**	.005
Number of plans	1.70	0.48	2	1	2	.76	0.89	1	0	3	38.0	314.0	3.14**	.002
Number of consequences	2.00	0.71	2	1	3	.63	0.63	1	0	2	19.0	295.0	3.90***	< .001

Notes. \*p < .05. \*\*p < .01. \*\*\*p < .001.

**Table 3.** Frequency with which different viewpoints were used by the two groups.

Viewpoint	Deaf signers (n = 10)			Learners (n = 23)			Results of parametric independent samples t-tests	
	Mean	SD	Min–Max	Mean	SD	Min–Max	t	p
Character	18.90	6.31	10–27	9.43	5.99	1–23	4.11***	< .001
Observer	6.90	3.41	1–11	5.91	2.45	2–10	.94	.353
Mixed	2.10	2.02	0–7	.57	0.99	0–4	2.95**	.006

Note. \* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

### Analysis 3: Frequency and total duration of different viewpoints

The frequency with which the two groups chose the different viewpoint options is presented in Table 3. A 3(viewpoint: character, observer, mixed)  $\times$  2(group: deaf signers, learners) repeated measures ANOVA revealed main effects of viewpoint,  $F(2,62) = 74.03$ ,  $p < .001$ , partial eta squared = .705; and of group,  $F(1,31) = 22.10$ ,  $p < .001$ , partial eta squared = .416; and an interaction between viewpoint and group,  $F(2,62) = 9.97$ ,  $p < .001$ , partial eta squared = .243. To investigate the main effect of viewpoint within each group, a set of paired samples  $t$ -tests were carried out which revealed that character viewpoint was used more frequently than observer viewpoint by both the deaf signers and the learners,  $t(9) = 5.37$ ,  $p < .001$  and  $t(22) = 2.46$ ,  $p = .022$ , respectively; character viewpoint was used more frequently than mixed viewpoint,  $t(9) = 8.97$ ,  $p < .001$  and  $t(22) = 7.46$ ,  $p < .001$ , respectively; and observer viewpoint was used more frequently than mixed viewpoint,  $t(9) = 3.26$ ,  $p = .010$ , and  $t(22) = 8.28$ ,  $p < .001$ , respectively.

A set of independent samples  $t$ -tests comparing the two groups for each viewpoint is presented in Table 3. The deaf signer group used the character viewpoint and mixed viewpoint significantly more frequently than the learner group. In contrast, there was no significant group difference in the frequency with which observer viewpoint was used.

Investigating the total duration of different viewpoints used by the two groups reveals a different pattern (Table 4). For duration, there was again a significant main effect of viewpoint,  $F(2,62) = 35.19$ ,  $p < .001$ , partial eta squared = .532. However, there was no main effect of group,  $F(1,31) = .08$ ,  $p = .779$ , partial eta squared = .003; and no interaction between viewpoint and group,  $F(2,62) = 1.95$ ,  $p = .151$ , partial eta squared = .059. The main effect of viewpoint was driven by participants spending significantly longer in character viewpoint than observer viewpoint,  $t(32) = 3.84$ ,  $p = .001$ ; significantly longer in character viewpoint than mixed viewpoint,  $t(32) = 7.97$ ,  $p < .001$ ; and significantly longer in observer viewpoint than mixed viewpoint,  $t(32) = 7.56$ ,  $p < .001$ .

Despite the interaction between viewpoint and group not reaching significance, the effect size was medium (Cohen, 1988), so we ran a set of independent-samples  $t$ -tests to test for group differences in the amount of time spent in the different viewpoints. The results of this analysis are presented in Table 4. The learner group spent significantly longer in observer viewpoint than the deaf signer group. In contrast, there were no significant group differences in the amount of time spent in character viewpoint or in mixed viewpoint.

**Table 4.** Total duration, in seconds, of different viewpoints used by the two groups.

Viewpoint	Deaf signers (n = 10)			Learners (n = 23)			Results of parametric independent samples t- tests	
	Mean	SD	Min–Max	Mean	SD	Min–Max	t	p
Character	17.55	8.39	8.98–35.25	14.07	11.62	0.86–50.13	.85	.401
Observer	4.84	3.09	0.66–9.18	8.20	3.63	2.43–15.57	-2.55*	.016
Mixed	1.79	2.37	0.0–8.31	.67	1.25	0.0–5.38	1.79	.083

Note. \* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

The main findings of this third analysis can be summarized as follows: (1) deaf signers adopted character viewpoint significantly more frequently than learners did, and (2) although deaf signers and learners adopted observer viewpoint with similar frequency, learners remained in observer viewpoint for longer than deaf signers did.

#### *Analysis 4: Number of separate articulators used simultaneously within each viewpoint*

In this final analysis, we investigated how many articulators were used simultaneously while the narrator was in a particular viewpoint. Up to seven articulators could be used at any one time, namely, the dominant hand, non-dominant hand, body, eyebrows, mouth, eyes and head. Table 5 presents the mean frequency with which participants in each group used 1–7 articulators. Non-parametric comparisons were used because some cells had the value zero. The analyses reveal that learners used 1–3 articulators more frequently than the deaf signers; for 4 articulators, there was no difference; and deaf signers used 5–7 articulators more frequently than learners. For learners, the number of articulators used simultaneously peaked at 3–4, whereas for deaf signers it peaked at 6.

To further contextualize the data, we provide details on how frequently each of the 7 articulators was used. We present these data in Table 6, as percentages calculated as the number of times the articulator was used in a viewpoint divided by the number of viewpoints. While it is not fully appropriate to carry out inferential statistics on data we did not plan to analyse, the descriptive statistics show that both groups use non-manual articulators to a lesser extent than the hands.

## **V Discussion**

There is relatively little quantitative and qualitative research into how hearing adults learn sign languages as an L2, despite the practical importance of this topic for sign language pedagogy and its theoretical importance for theories of second language learning. In the current study we collected signed narratives (using the *Frog story*) in a group of intermediate level learners and a group of deaf native/early learners of BSL. Our aims were to reveal some of the challenges that learners of a sign language face when producing a signed narrative. We discovered that learners produced narratives that were just as long as the deaf signers' narratives, but that contained less information. Although learners were able to take on character viewpoint, they seemed less comfortable than deaf signers in

**Table 5.** Frequency which different numbers of articulators were used simultaneously.

Number of articulators	Deaf signers (n = 10)				Learners (n = 23)				Results of non-parametric independent samples tests			
	Mean	SD	Median	Min-Max	Mean	SD	Median	Min-Max	Mann Whitney U	Wilcoxon W	Z	p
1	0	0	0	0-0	0.65	.83	0	0-2	65.00	120.00	-2.42*	.016
2	.20	0.42	0	0-1	2.48	2.41	2	0-10	30.00	85.00	-3.44**	.001
3	1.70	1.34	1.5	0-5	3.91	1.93	4	1-8	36.50	91.50	-3.12**	.002
4	3.80	2.10	4.5	0-6	3.70	1.66	4	1-8	99.50	375.00	-.62	.536
5	5.60	1.71	6	3-9	2.61	1.77	2	0-7	26.50	302.50	-3.51***	< .001
6	9.40	3.17	9.5	4-15	1.78	2.78	1	0-10	12.00	288.00	-4.10***	< .001
7	7.00	3.92	8	0-11	0.74	1.45	0	0-5	20.00	296.00	-3.96***	< .001

Notes. \* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

**Table 6.** Mean (SD) percentages of articulator use.

	Dominant hand	Non-dominant hand	Body	Eyebrows	Mouth	Eyes	Head
Deaf signers	99.70 (0.95)	91.20 (7.08)	51.30 (22.30)	58.70 (7.15)	80.20 (10.48)	88.00 (10.14)	84.50 (15.35)
Learners	97.43 (4.02)	77.52 (11.09)	26.91 (19.77)	21.48 (14.83)	59.92 (20.87)	62.83 (21.23)	33.22 (18.23)

using it. Specifically, learners used character viewpoint less frequently than deaf signers, and although they spent just as long as deaf signers in that viewpoint, they spent longer than deaf signers in observer viewpoint, suggesting that they found observer viewpoint easier than character viewpoint. Finally, learners were limited compared to deaf signers in the numbers of articulators that they used simultaneously. We discuss each of these findings in relation to whether they represent challenges that have been previously documented in the second language research literature as universal learner patterns, or whether they might be particularly unique to learning a sign language.

Our participants were able to produce narratives that did not differ significantly in duration from those of the deaf signers. There was great variability in the duration of narratives within both groups, as has been found in a previous study that compared the narratives of deaf signers and of hearing adults who were learning Catalan Sign Language (Bel et al., 2015). Nevertheless, the learners produced less informational content in their narratives with respect to all the features that we analysed, namely the numbers of characters, settings, initiating events, internal responses, plans and consequences. This suggests that our learners had not yet mastered all the linguistic elements involved in narrative to a high enough degree of automaticity to bring them together simultaneously. In summary, they operated following universal learner patterns and at this point in their learning reduced the linguistic detail they produced in order to cope with the high task demands (Service, 1992; Miyake and Friedman, 1998). Narrative is a high-level task for second language learners because it requires them to understand the stimuli, plan the message, select linguistic items, attend to pragmatic rules, monitor their own performance and – in the case of signing – maintain their signing rate. All of these general cognitive demands are likely to influence why L2 signers, as would be the case for L2 learners more generally, produce qualitatively poorer narratives than fluent signers.

Other challenges might be more specific to learning to produce narrative in the signed modality. One potential challenge is the requirement to encode language information across multiple articulators, i.e. the two hands, body, eyebrows, mouth, eyes and head (Sandler, 2012). We purposely chose to code a section of the *Frog story* which featured more than two characters and which we therefore expected to challenge learners with the need to express information about multiple characters, potentially at the same time. The linguistic demands of constructing a coherent signed narrative are high. At the conceptual level two aspects of space need to be coordinated and expressed through character or observer viewpoint choices (Perniss, 2007; Emmorey and Tversky, 2002; Emmorey et al., 2000). This complexity is revealed by differences in deaf and L2 learners. For example, complex motion events involving the deer, boy and dog were expected to challenge learners' ability to represent all three characters. Learners were able to use multiple articulators simultaneously, but not nearly to the same extent as deaf signers. The differences were striking: whereas learners most commonly used 3–4 articulators at the same time, deaf signers most commonly used 6, and frequently 7. This type of learner pattern is arguably specific to learning a sign language, as in spoken languages learners need to master elements within a more constrained set of articulators (the speech production system) rather than across the hands, face, eyes and torso. Of course, the underlying reason for reducing the number of articulators may well be because of a general learner constraint: the cognitive load to manage all of these articulators simultaneously is high and production accuracy is therefore fragile in L2 learners.

In Figures 8 and 9 earlier we showed an example from a deaf signer and from a learner of a particular scene in the narrative where the boy finds himself on top of the deer, and we used this example to illustrate that the deaf signer used a greater number of articulators at the same time compared to the learner. The screenshots in Figures 10 and 11 are from shortly afterwards in the narrative. Now the learner (Figure 11) is using the eyebrows and eyes in addition to both hands to represent how the boy grabs on to what he thinks are branches, but which are actually the deer's antlers. The deaf signer (Figure 10) continues to use all 7 articulators, which, as is the case for the learner, all represent the boy. These figures illustrate the greater richness of the deaf signers' narratives, and the greater opportunity, through the use of a greater number of simultaneous articulators, to produce more detail (e.g. the boy's emotional reaction to the movement of the 'branches' is shown by the deaf signer in Figure 10, but not by the learner in Figure 11).

The use of more than one articulator simultaneously is of course relevant to expressing viewpoint, which was the principal focus of our study. Two findings lead us to conclude that learners find expressing character viewpoint more challenging than portraying observer viewpoint. Firstly, they used character viewpoint less frequently than deaf signers. Secondly, they spent longer in observer viewpoint than deaf signers. These findings might reflect speakers' preferences for observer viewpoint that have been documented by other researchers (Earis and Cormier, 2013; Parrill, 2010), and suggest that learners transfer their viewpoint preferences from their L1 English. This is strictly speaking not an error as both groups use these two viewpoints; what differs is the over-use of the observer viewpoint by the learners, and this could be interpreted as pragmatically inappropriate (Serratrice et al., 2004).

Figures 2 and 4, earlier, show an event in the story for which learners tended to use observer viewpoint with entity classifiers, but deaf signers tended to use character viewpoint with constructed action. Here the participants were narrating the part where the dog is running alongside/chasing the deer and the boy. The deaf signer (Figure 2) chose to represent the dog using character viewpoint with constructed action, whereas the learner (Figure 4) used observer viewpoint, representing the dog with one classifier handshape and the deer (and boy) with another. Figures 3 and 5, earlier, illustrate this same difference in choices – for representing the boy on top of the deer, the learner (Figure 5) chose to use classifiers, and therefore observer viewpoint, whereas the deaf signer (Figure 3) chose character viewpoint with constructed action. In previous studies learners have been shown to find the formational properties of classifiers difficult (Boers-Visker and van den Bogaerde, 2018; Marshall and Morgan, 2015), yet in our study they appear willing to use them in the context of a narrative, including for events which signers would preferentially use character perspective. Nevertheless, the learners have not chosen the conventional handshapes that more fluent signers would choose. For example, in Figure 4 the learner has used a classifier handshape for dog that would more appropriately be used for vehicles or flat objects, while her other hand combines both the direction of movement of the deer and the conventional classifier handshape for person in a way that fluent signers of BSL would judge as inappropriate. She has in effect run out of hands to express the three characters.

The use of mixed viewpoint, i.e. the simultaneous use of character and observer or two characters' viewpoints was much less frequent in both groups. Nevertheless, we did observe this viewpoint choice in both deaf signers and learners (see Figure 8, where the



**THE BOY HOLDS ON TO WHAT HE THINKS ARE BRANCHES, BUT HE IS SHAKEN FROM SIDE TO SIDE**

**Articulators:**

- Dom Hand – boy holding on to the branches
- Non Dom Hand – boy holding onto the branches
- Body – showing (unexpected) movement of the boy as the deer appears
- Head – showing movement of the boy as deer appears
- Mouth – showing shocked reaction of the boy
- Eyes – looking at the branches and then directly to show the boy's focus
- Eyebrows – furrowed brows to show the boy's worry and fear

**Figure 10.** Example of a deaf signer using 7 articulators simultaneously.



**BOY**

THE BOY HOLDS ON TO WHAT HE THINKS ARE BRANCHES, BUT THEY START TO MOVE UPWARDS

**Articulators:**  
Dom Hand – boy grabbing hold of the branches  
Non Dom Hand – preparing to grab the other branches  
Eyebrows – slight raise to show motion of grabbing  
Eyes – showing the boy's focus on the branches

Figure 11. Example of a learner using 4 articulators simultaneously.

signer is using one hand to demonstrate the boy on top of the deer's head (observer viewpoint) and the other hand to demonstrate the deer's antlers (character viewpoint)). Deaf signers used mixed viewpoint significantly more frequently than learners, although group differences in the amount of time spent in that viewpoint did not reach significance. Mixed viewpoint has been previously described in sign languages other than BSL, including South African Sign Language (Aarons and Morgan, 2003) and American Sign Language (Quinto-Pozos and Parrill, 2015), and it is reported to be rarer in co-speech gesture (Parrill et al., 2016). Its infrequent use in our learners compared to our deaf signers might therefore be another example of transfer of learners' L1 preferences to their L2.

A strength of our study is that we recruited a larger number of sign learners ( $n = 23$ ) than has been the case in previous research studies, for example  $n = 8$  (Frederiksen and Mayberry, 2019),  $n = 12$  (Ferrara and Nilsson, 2017; Marshall and Morgan, 2015) and  $n = 13$  (Bel et al., 2015). There is individual variability in sign L2 learners just as there is in spoken language L2 learners, and bigger samples are more likely to capture that variability, while in experimental studies small samples are likely to be under-powered to find significant differences between groups (as acknowledged, for example, by Frederiksen and Mayberry, 2019).

Our study has limitations too, of course. Most notably is that, given the time-consuming nature of sign language coding, we selected just one set of episodes from the narrative to code; it might have been the case that some learners used a more fluent set of sign narrative elements in the scenes other than ones we selected. In addition, the method for data elicitation was a fictitious story about an escaped frog. We did not elicit narratives of signers' personal experiences, and we chose not to do this because we wanted instead to elicit narratives with the same characters and events so that the data within and across groups were directly comparable. But personal narratives, by drawing on events that the narrator has directly experienced, might potentially make it less challenging for learners of sign to take on character viewpoint. Stec et al. (2017), in a micro-analysis of one personal spoken narrative involving two characters, showed that in the quoted dialogue sequences the narrator used multi-modal resources such as facial expression, manual gesture, eye gaze and posture change (including head movements) to distinguish between the two characters. It might be the case that transferring these resources from speech to sign is less challenging in the context of a personal narrative than when relating a previously unfamiliar story.

Future studies are needed to replicate the study reported in the current article (including in sign languages other than BSL), but also to extend it in various ways. An obvious question concerns how learners of sign express viewpoint in their spoken narratives, and whether their individual preferences for character versus observer viewpoint are found in both their spoken and their signed narratives. Another question concerns the developmental aspects of narrative, and how the use of character and observer viewpoint changes as learners' BSL fluency increases. An obvious prediction is that as learners progress they become more confident using character viewpoint and therefore move towards using it with the same frequency as deaf signers. However, this prediction remains to be tested.

Importantly, however, our study contributes to the research literature by revealing that despite M2L2 learners having a range of gestural resources available to them, the conventionalized nature of sign languages means that they have difficulties expressing viewpoint in signed narratives. Character viewpoint is particularly challenging for them, and coding information across multiple articulators simultaneously is also difficult. These findings

suggest avenues for future research but also have implications for the teaching of sign languages and for where teachers might wish to focus some of their instruction.

### Acknowledgements

We thank our collaborators in the Leverhulme Trust-funded international network for the learning of signed languages for discussions about the data presented here. We also thank Jordan Fenlon for his help in recruiting participants, and Hugh Mulloy for his help with the data coding. Finally, we are very grateful to all the people who gave up their time to participate in this study.

### Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Leverhulme Trust (Grant number IN-2015-048, awarded to the last author).

### ORCID iD

Chloë Marshall  <https://orcid.org/0000-0003-2405-1999>

### Notes

1. These learners are therefore sometimes termed ‘second modality, second language’ learners, or M2L2 learners for short (e.g. Chen Pichler and Koulidobrova, 2016; Ferrara and Nilsson, 2017; Williams and Newman, 2016).
2. Like Parrill et al. (2016), we have chosen to use the terms observer and character viewpoint when talking about sign language (even though these terms come from the gesture literature) because they are more transparent for readers with little or no knowledge of sign language.
3. We follow Cormier et al. (2012b) in calling ‘early signers’ those who were not exposed to BSL from birth but who were exposed before the age of 8 years. As can be seen from Table 1, all our early signers were first exposed to BSL aged 4 years or younger.
4. Age data are missing for one deaf and one hearing participant (see Table 1).
5. Meaningful articulator = the articulator has a specific meaning assigned to it, e.g. an eye gaze to show the placement of the boy on the left, an eyebrow raise to show surprise, mouthing to show frustration.
6. For example, when the signer referred to only one character doing an action when in fact the picture clearly showed that two characters were involved.

### References

- Aarons D and Morgan RZ (2003) Classifier predicates and the creation of multiple perspectives in South African Sign Language. *Sign Language Studies* 3: 125–56.
- Bamberg M and Damrad-Frye R (1991) On the ability to provide evaluative comments: Further explanations of children’s narrative competencies. *Journal of Child Language* 18: 689–710.
- Bel A, Ortells M, and Morgan G (2015) Reference control in the narratives of adult sign language learners. *International Journal of Bilingualism* 19: 608–24.

- Berman R and Slobin (1994) *Relating events in narrative: A crosslinguistic developmental study*. Hillsdale, NJ: Erlbaum.
- Boers-Visker E and van den Bogaerde B (2018) There's a goat on the roof: The acquisition of classifier constructions by adult learners of NGT. Unpublished poster presented at the International Conference on Sign Language Acquisition, Istanbul, Turkey.
- Chen Pichler D and Koulidobrova H (2016) Acquisition of sign language as a second language. In: Marschark M and Spencer PE (eds) *The Oxford handbook of deaf studies in language*. Oxford: Oxford University Press, pp. 218–30.
- Clark HH (2016) Depicting as a method of communication. *Psychological Review* 123: 324–47.
- Cohen J (1988) *Statistical power analysis for the behavioral sciences*. 2nd edition. Hillsdale, NJ: Erlbaum.
- Cormier K, Quinto-Pozos D, Sevcikova Z, and Schembri A (2012a) Lexicalisation and delexicalisation processed in sign languages: Comparing depicting constructions and viewpoint gestures. *Language and Communication* 32: 329–48.
- Cormier K, Schembri A, Vinson D, and Orfanidou E (2012b) First language acquisition differs from second language acquisition in prelingually deaf signers: Evidence from grammatical processing of British Sign Language. *Cognition* 124: 50–65.
- Dudis P (2004) Body-partitioning and real-space blends. *Cognitive Linguistics* 15: 223–38.
- Earis H and Cormier K (2013) Point of view in British Sign Language and spoken English narrative discourse: The example of 'The Tortoise and the Hare'. *Language and Cognition* 5: 313–43.
- ELAN (Version 5.0.0-beta) [Computer software]. (2017, April 18). Nijmegen: Max Planck Institute for Psycholinguistics. Available at: <https://tla.mpi.nl/tools/tla-tools/elan> (accessed January 2020).
- Ellis R (1987) Interlanguage variability in narrative discourse: style shifting and use of the past tense. *Studies in Second Language Acquisition* 9: 1–19.
- Emmorey K (2002) *Language, cognition, and the brain: Insights from Sign Language Research*. Mahwah, NJ: Lawrence Erlbaum.
- Emmorey K and Tversky B (2002) Spatial perspective choice in ASL. *Sign Language and Linguistics* 5: 3–25.
- Emmorey K, Tversky B, and Taylor HA (2000) Using space to describe space: Perspective in speech, sign, and gesture. *Journal of Spatial Cognition and Computation* 2: 157–80.
- Engberg-Pedersen E (1993) *Space in Danish Sign Language: The semantics and morphosyntax of the use of space in a visual language*. Hamburg: Signum.
- Engberg-Pedersen E (1995) Point of view expressed through shifters. In: Emmorey K and Reilly J (eds) *Language, gesture, and space*. Hillsdale, NJ: Lawrence Erlbaum, pp. 133–54.
- Engberg-Pedersen E (2003) Descriptions of falls in Danish Sign Language. In Emmorey K (ed.) *Perspectives on classifier constructions in sign languages*. Mahwah, NJ: Lawrence Erlbaum, pp. 311–32.
- Ferrara L and Nilsson A-L (2017) Describing spatial layouts as an M2 signed language learner. *Sign Language and Linguistics* 20: 1–26.
- Frederiksen AT and Mayberry R (2019) Reference tracking in early stages of different modality L2 acquisition: Limited overexplicitness in novice ASL signers' referring expressions. *Second Language Research* 35: 253–83.
- Furman N, Goldberg D, and Lusin N (2010) *Enrollments in languages other than English in United States institutions of higher education, Fall 2009*. New York: Modern Languages Association. Available at: [http://www.mla.org/2009\\_enrollmentsurvey](http://www.mla.org/2009_enrollmentsurvey) (accessed January 2020).
- Gullberg M (2006) Handling discourse: Gestures, reference tracking, and communication strategies in early L2. *Language Learning* 56: 155–96.

- Halliday MAK and Hasan R (1976) *Cohesion in English*. London: Longman Group.
- Hendriks H (2003) Using nouns for reference maintenance: A seeming contradiction in L2 discourse. In: Ramat AG (ed.) *Typology and second language acquisition*. Berlin: Walter de Gruyter, pp. 291–326.
- Hodge G, Ferrara L, and Anible BD (2019) The semiotic diversity of doing reference in a deaf signed language. *Journal of Pragmatics* 143: 33–53.
- Kita S (2014) Production of speech-accompanying gesture. In: Goldrick M, Ferreira V, and Miozzo M (eds) *Oxford handbook of language production*. Oxford: Oxford University Press, pp. 451–59.
- Kita S and Özyürek A (2003) What does cross-linguistic variation in semantic co-ordination of speech and gesture reveal? Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language* 48: 16–32.
- Lillo-Martin D (1995) The point of view predicate in American Sign Language. In Emmorey K and Reilly J (eds) *Language, gesture, and space*. Hillsdale, NJ: Lawrence Erlbaum, pp. 155–70.
- Loew RC (1984) Roles and reference in American Sign Language: A developmental perspective. Unpublished PhD thesis, University of Minnesota, Minnesota, MN, USA.
- Marshall CR and Morgan G (2015) From gesture to sign language: Conventionalization of classifier constructions by adult hearing learners of British sign language. *Topics in Cognitive Science* 7: 61–80.
- Mayer M (1969) *Frog, where are you?* New York: Dial Press.
- McNeill D (1992) *Hand and mind: What gestures reveal about thought*. Chicago, IL: Chicago University Press.
- Metzger M (1995) Constructed dialogue and constructed action in American Sign Language. In: Lucas C (ed.) *Sociolinguistics in deaf communities*. Washington, DC: Gallaudet University Press, pp. 255–71.
- Miyake A and Friedman NP (1998) Individual differences in second language proficiency: Working memory as language aptitude. In: Healy AF and Bourne LE (eds) *Foreign language learning: Psycholinguistic studies on training and retention*. Mahwah, NJ: Lawrence Erlbaum, pp. 339–64.
- Morgan G (2002) The encoding of simultaneity in children's BSL narratives. *Journal of Sign Language and Linguistics* 5: 127–61.
- Morgan G and Woll B (2007) Understanding sign language classifiers through a polycomponential approach. *Lingua* 117: 1159–68.
- Norbury C, Gemell T, and Paul R (2014) Pragmatic abilities in narrative production: A cross-disorder comparison. *Journal of Child Language* 41: 485–510.
- Ortega L (2000) Understanding syntactic complexity: The measurement of change in the syntax of instructed L2 Spanish learners. Unpublished PhD thesis, University of Hawai'i at Manoa, Manoa, HI, USA.
- Özyürek A and Perniss P (2011) Event representation in sign language: A crosslinguistic perspective. In: Bohnemeyer J and Pederson E (eds) *Event representation in language: Encoding events at the language-cognition interface*. Cambridge: Cambridge University Press, pp. 84–107.
- Padden C (1986) *Verbs and role-shifting in ASL*. In Padden C (ed.) *Proceedings of the 4th National Symposium on Signing Research and Teaching*. Las Vegas, NV / Washington, DC: The National Association of the Deaf, pp. 44–57.
- Parrill F (2010) Viewpoint in speech-gesture integration: Linguistic structure, discourse structure, and event structure. *Language and Cognitive Processes* 25: 650–68.
- Parrill F, Stec K, Quinto-Pozos D, and Rimehaug S (2016) Linguistic, gestural, and cinematographic viewpoint: An analysis of ASL and English narrative. *Cognitive Linguistics* 27: 345–69.

- Perniss P (2007) Achieving spatial coherence in German Sign Language narratives: The use of classifiers and perspective. *Lingua* 117: 1315–38.
- Perniss P, Özyürek A, and Morgan G (2015) The influence of the visual modality on language structure and conventionalization: Insights from sign language and gesture. *Topics in Cognitive Science* 7: 2–11.
- Pyers JE, Perniss P, and Emmorey K (2015) Viewpoint in the visual–spatial modality: The coordination of spatial perspective. *Spatial Cognition and Computation* 15: 143–69.
- Quinto-Pozos D (2008) Teaching American Sign Language to hearing adult learners. *Annual Review of Applied Linguistics* 31: 137–58.
- Quinto-Pozos D and Parrill F (2015) Signers and co-speech gesturers adopt similar strategies for portraying viewpoint in narratives. *Topics in Cognitive Science* 7: 12–35.
- Robinson P (1995) Task complexity and second language narrative discourse. *Language Learning* 45: 99–140.
- SALT (2009) *Assessing language production using SALT software: A clinician’s guide to language sample analysis*. Madison, WI: SALT Software. Available at: <http://saltsoftware.com/ssdownload/BookPDF/SALTReferenceBook.pdf> (accessed January 2020).
- Sandler W (2012) Dedicated gestures in the emergence of sign language. *Gesture* 12: 265–307.
- Sandler W, Meir I, Dachkovsky S, Padden C, and Aronoff M (2011) The emergence of complexity in prosody and syntax. *Lingua* 121: 2014–33.
- Schornstein RA (2005) Teaching ASL in the university: One teacher’s journey. *Sign Language Studies* 5: 398–414.
- Serratrice L, Sorace A, and Paoli A (2004) Crosslinguistic influence at the syntax–pragmatics interface: Subjects and objects in English–Italian bilingual and monolingual acquisition. *Bilingualism: Language and Cognition* 7: 183–205.
- Service E (1992) Phonology, working memory and foreign-language learning. *Quarterly Journal of Experimental Psychology* 45: 21–50.
- Squires KE, Lugo-Neris MJ, Pena E et al. (2014) Story retelling by bilingual children with language impairments and typically-developing controls. *International Journal of Language and Communication Disorders* 49: 60–74.
- Stec K (2012) Meaningful shifts: A review of the viewpoint markers in co-speech gesture and sign language. *Gesture* 12: 327–60.
- Stec K, Huiskes M, Wieling M, and Redeker G (2017) Multimodal character viewpoint in quoted dialogue sequences. *Glossa: A Journal of General Linguistics* 2: 39.
- Williams J and Newman S (2016) Phonological substitution errors in L2 ASL sentence processing by hearing M2L2 learners. *Second Language Research* 32: 347–66.

## Appendix I

### Coding for information content

The coding procedure involved checking unique tokens of characters, settings, initiating events, internal responses, plans, and consequences. These tokens were as follows:

Characters

- boy
- deer
- dog

- frog
- owl

### Settings

- outside
- rock
- cliff
- pond
- tree
- log
- bush
- branches
- grass
- weeds

### Initiating events

- The boy climbs up on branches/deer's antlers.
- The owl is sitting in a tree.
- The branches aren't really branches! They are the deer's antlers.
- The deer picks up the boy on his head.
- The boy is being thrown about.
- The dog is sniffing below on the ground.
- The deer starts running with the boy still on his head.
- The dog runs along too.
- The dog, running with his back to the cliff, is chasing the deer with the boy on top.
- They are getting close to a cliff.
- The deer stops suddenly.
- The boy and the dog fall over the edge of the cliff.
- Rocks from the edge crumble down.
- There is a pond below the cliff.
- They land with a splash right on top of one another.
- The dog cradles the boy/climbs on his head and rests his paws.
- They hear a familiar sound.

### Internal responses

- The boy is shocked/surprised/startled/confused when he holds onto the branches.
- The owl is staring curiously at the boy.
- As the boy is holding on, he wonders what is happening.
- The boy is frightened/confused when the deer is running.
- The dog is sniffing below, keeping himself to himself.
- The dog is chasing deer and boy, unsure what was happening.
- The dog looks in horror as he chases the deer and the boy.

- The dog looks sad.
- The dog doesn't like being wet so he sits on the boy's head.
- The boy smiles as he hears a familiar sound.

### Plans

- The boy thinks the deer's antlers are a branch/branches.
- The dog decides to chase the deer with the boy on top.
- The deer stops so he will not fall over the cliff.
- The dog decides to climb on top of the boy's head.

### Consequences

- The deer stops suddenly so the boy falls off down the cliff.
- The dog is running so fast and looks behind that he doesn't see the cliff edge.
- They both land in the pond wet.
- Luckily the water is shallow so the boy is able to sit down in the water but still gets wet.

## Appendix 2

### *Coding viewpoint*

We adapted Parrill's (2010) definitions of viewpoints for BSL:

In character viewpoint the signer:

- Portrays the actions of a character via the signer's movements and displays of affect
- May represent simulation from an internal, first person point of view
- Describes spaces using a single scale
- Uses constructed action,\* handling and body portioning productions
- Assumes the role of the character in the event, such that at least the character's head and torso are mapped to the signer's body, and the size of the projected space is life-sized

In observer viewpoint the signer:

- Depicts the entire character (or objects) and their movements in a smaller scale in front of them
- Uses the hands to represent the character, and the trajectory traced by the hand that represents the character's path
- Depicts the action as though observing it from afar
- Projects event space onto sign space from an external vantage point
- The signer is not part of the represented event, and the event space is reduced in size, projected onto the area of space in front of the signer's body

- May represent simulation from an external, third person point of view
- Describes space with varying scales
- Uses entity classifiers\*\*

In mixed viewpoint\*\*\* the signer:

- Simultaneously encodes both a character's and an observer's point of view
- Produces constructed action and entity classifiers at same time

*Notes.* \* Constructed action: Signer uses his/her face, head, body, hands, and/or other non-manual cues to represent a referent's actions, utterances, thoughts, feelings and/or attitudes (Metzger, 1995). Constructed action is also known as role shift (Padden, 1986), referential shift (Emmorey, 2002; Engberg-Pedersen, 1993), or point of view predicate (Lillo-Martin, 1995).

\*\* Entity classifiers: The hands represent a person or object from a distance. Entity classifiers are found in spatial verbs. They use topographic space, and give information about the path, trajectory, speed and location (Morgan and Woll, 2007).

\*\*\* Parrill (2010) termed this 'dual viewpoint' but we prefer the term 'mixed viewpoint'.

## Appendix 3

### *Coding articulators*

Tiers were developed for each articulator, as shown below. An Articulators-Quantity tier consisted of a controlled vocabulary with the numerals 1–7, and was used to record the total number of articulators per viewpoint.

Articulator tiers:

- Hands 1 (refers to dominant hand)
- Hands 2 (refers to non-dominant hand)
- Body
- Eyebrows
- Mouth
- Eyes
- Head
- Articulators-Quantity

Looking at each viewpoint shift, coders annotated which articulators were meaningful and present. When all articulators were coded for a particular shift then they coded how many articulators appeared in the Articulators-Quantity tier.