Coordinating signs and eye gaze in the depiction of directions and spatial scenes by fluent and L2 signers of Norwegian Sign Language

Submitted to *Spatial Cognition & Computation*

Lindsay Ferrara
lindsay.n.ferrara@ntnu.no
NTNU, Norwegian University of Science and Technology

Abstract

The current study investigates the coordination of signs and eye gaze during depictions of directions and spatial scenes by fluent and second language (L2) signers of Norwegian Sign Language. First, findings show that fluent and L2 signers make different choices regarding the perspective they use to depict spatial scenes. Secondly, there is variation within and across groups in relation to how eye gaze is directed during these depictions. Findings suggest that eye gaze is used to establish a vantage point upon a depicted scene, but L2 learners may not always engage in this type of coordination. This study contributes to our understanding of how visual perspective is depicted in signed languages and has implications for signed language pedagogy.

Keywords: visual perspective, spatial language, depiction, Norwegian Sign Language, eye gaze

Introduction

Interaction is a multimodal activity, as people coordinate their hands, body, facial expressions, and eye gaze to communicate with each other. The current study looks at this type of multimodal coordination in samples of Norwegian Sign Language produced by fluent and second language (L2) learners.[1] In particular, a comparison will be made between how these two groups of signers give directions and describe spatial scenes. Previous studies showed that while beginner-intermediate L2 hearing signers are able to produce signs that depict locations and spatial relations, they still struggle to produce texts that make sense (Nilsson & Ferrara, 2015, 2016; Ferrara & Nilsson, 2017). To further investigate this matter, the current study draws on previous observations that signers direct their gaze and body to their hands, the signing space, or to invisible emerging scenes during periods of depiction (Engberg-Pedersen, 1993, 2016; Metzger, 1995; Dudis 2011). By doing this, signers establish a visual perspective with which the scene can be viewed and interpreted (e.g., Emmorey, Tversky, & Taylor, 2000; Janzen, 2004; Perniss, 2007). This study will combine these lines of research to investigate how fluent and L2 signers coordinate signs and eye gaze to establish visual perspective during periods of depiction.

Before moving on to the methods and findings of the study, a brief introduction to spatial language and its acquisition in spoken language contexts is presented. Then, the focus will turn to signed languages with a brief review of how signers show meanings related to locations and spatial relations through the movement and placement of iconic depicting signs in the signing space. In addition, the topographical use of the signing space and perspective are introduced, along with the role eye gaze plays in signed language discourse.

Spatial language and its acquisition

As mentioned in the introduction, this study is concerned with spatial language in Norwegian Sign Language. Investigations of spatial language have focused on how speakers of spoken languages talk about space, for example, documenting and comparing how different types of spatial terms and vocabulary invoke different spatial conceptualizations (e.g., Bowerman, 1996; Forker, 2012). Another line of research has considered the influence language has on spatial cognition. While some have seen the appeal for a non-linguistic basis for spatial cognition, there is now some convincing evidence that suggests a tighter connection between language (and culture) and spatial cognition (e.g., Bowerman, 1996; Levinson, 1996; Levinson, 2003; Landau, Dessalegn, & Goldberg, 2010).

Cross-linguistic research has also revealed immense variation in how languages carve up the spatial world and represent it linguistically (Levinson, 1996). As one example, studies have examined how different speech communities engage different reference frames—namely intrinsic, absolute, and relative—to specify locations. These frames of

[1] The use of the term 'multimodal' here refers to multiple sensory modalities (e.g., aural and visual) as well as to multiple dimensions within one sensory modality (e.g., manual actions and eye gaze behavior) (e.g., Ruiter *et al*., 2003, p. 408; Enfield, 2009, pgs. 17-18).

reference make use of coordinate systems based on different referents. In an intrinsic reference frame the speaker uses a ground object to locate a figure object. In an absolute frame of reference, in contrast, a speaker uses sets of fixed bearings to locate objects, often exemplified in terms like the English *north*, *south*, *east*, *west*. In this way, an absolute frame of reference lacks a viewpoint. Finally, relative frames of reference employ a coordinate system based on a particular viewpoint (e.g., that of the speaker vs. another viewer). While signed languages also have signs that lexically encode spatial relationships (e.g., the Norwegian signs for "on" or "to"), in this paper, we will be exploring how signers are also able to present spatial scenes through the placement and movement of depicting signs in the signing space. In this way, signers do not have to carve up their spatial world and represent it through lexical signs, but rather they must make decisions about how to directly map spatial relationships on to the space in front of them through the placement and movement of their hands.

The current study is also concerned with spatial language as it is used in a particular acquisition setting, namely hearing adult L2 contexts. While there is little research on spatial language acquisition in signed languages (see below), there has been some work done on the first language (L1) and second language (L2) acquisition of spatial language in spoken language contexts, which may help frame the current study. Within L1 acquisition research, it has been shown that some basic spatial terms are able to be learned early, but that full mastery has a protracted development (Coventry, Guijarro-Fuentes, Valdés, 2012; Shusterman & Li, 2016). There is also some evidence that children have an early preference for an intrinsic frame of reference, but that they will develop proficiency with a relative frame of reference if they need it for their language (Shusterman & Li, 2016).

Research also shows that spatial language can be difficult for adult L2 learners, often in relation to the use and comprehension of prepositions (e.g., Tyler & Evans, 2004; Coventry, Guijarro-Fuentes, Valdés, 2012). In a study of the English preposition *over*, Tyler & Evans (2004) suggested that L2 pedagogy that focuses on providing learners a coherent account of prepositions and their extended senses—rather than presenting them as a list of unrelated meanings or only focusing on central meanings—may improve language learning. Tyler (2012) later conducted a small-scale investigation of the L2 acquisition of prepositions and their extended senses and showed that learners (either L2 English or L2 Russian) struggled less with central senses and more with non-central senses (which may also present mismatches in the L1). She also commented that most L2 pedagogy focuses only on central senses. In a different study, Marotta & Mieni (2012) highlighted the fact that spatial prepositions are introduced early in many foreign language classrooms and that this input is assumed to play a role in students' learning. This position was supported in their findings from a study on the L2 acquisition of Italian prepositions, but they also suggested a possible influence from the students' L1. Both of these studies mentioned the influence of pedagogy on the acquisition process, a theme that will be taken up again in the discussion of findings from this study. They also focused on the lexical nature of spatial language, but once again, here, focus will not be on lexical prepositions but rather how signers depict spatial scenes in the signing space from different perspectives.

Introduction to depicting signs, the signing space, and perspective in signed languages

Depicting signs

There is one type of signed language sign that is central to a discussion of spatial language in signed languages. These signs go by many names in the literature (e.g., classifier predicates, verbs of motion and location) and have been analyzed in various ways (e.g., as morphemic, gestural, or a blend of the two). Here, the term *depicting sign* is adopted. Depicting signs are considered to be a type of partly lexical sign, because while some aspects of these signs' forms may be conventionally associated with particular meanings, e.g., type of referent or manner of movement, the sign as a whole is interpreted with additional spatial and mental mappings that arise as part of a particular instance of use. These signs are also seen as complex, iconic constructions that partially depict their meanings (Liddell, 2003; Johnston & Schembri, 2010).[2] Some linguists categorize depicting signs according to handshape, which is often considered the most conventionalized component of these signs (e.g., Supalla, 1986; Schick, 1990; Sandler & Lillo-Martin, 2006). In this study, however, categorization will follow the overall schematic meaning of the sign, aligning with cognitive-functional approaches to signed language structure (e.g., Liddell, 2003; Johnston & Schembri, 2007; Johnston, 2016). These categories and their gloss abbreviations are presented in Figure 1, along with examples from Norwegian Sign Language. As can be seen with the examples, depicting signs may involve imagery that aligns with more than one category. For data annotation and analysis purposes, the category chosen represents the main function of the depicting sign in context (as interpreted by the annotator). Thus, while the sign on the top left in Figure 1 depicts the movement of an entity, it also includes details about size and shape (and even location). However, the sign in context means, 'He rides along on his horse' and not, 'There was a man physically located on top of the horse.' Thus, the sign was categorized as depicting the movement of an entity, rather than a sign depicting size, shape, or location.

[2] As for alternative analyses of depicting signs, I direct the reader to work, for example, by Supalla, 1986; Schick, 1990; Sandler and Lillo-Martin, 2006. In this paper, the particular theoretical account of the composition of depicting signs is not essential to an analysis of eye gaze behavior with these signs. It only played a part in how the depicting signs in the data were labeled/annotated.

Figure 1. Examples of depicting signs in Norwegian Sign Language. Note that below each illustration are glosses for the sign as produced on either one or two hands. An English translation of these signs in context are also provided. (Adapted from Ferrara & Nilsson, 2017, p. 6)

Depicting signs can be used to express a range of meanings, from how entities act and move to what they look like or where they are located. However, an element of spatial meaning underlies all of these categories, because depicting signs often establish spatial relationships among discourse referents. Signers depict these spatial relationships iconically by positioning and moving the hands in the physical (empty) space around them. In addition, signers are able to manipulate the form of the depicting sign itself to reflect aspects of its meaning. For example, one could take the sign DSS:ROAD-EXTEND in Figure 1 and manipulate its path—e.g., straight, curved, arced—and this will change how the path of the road being discussed is conceptualized.[3]

Before describing how depicting signs are deployed in the signing space and their interaction with different types of visual perspective, a brief review of research on the acquisition of depicting signs is in order. This theme has received some research

[3] Depicting signs are represented in this paper by glosses written in small capitals that begin with a prefix for the type of depicting sign followed by a brief description of the sign's meaning in context.

attention but has been largely focused on single signs and their production, rather than how they are integrated into larger sequences of discourse. There is now a body of research that demonstrates a protracted development for depicting signs in L1 acquisition (e.g., Kantor, 1980; Slobin *et al.*, 2003; Beuzeville, 2006) and L2 acquisition (Marshall & Morgan, 2014; Ferrara & Nilsson, 2017, see also Smith & Cormier, 2014). During L1 acquisition, children have been observed to produce some types of depicting signs in simple settings correctly. For example, Slobin *et al*. (2003, p. 281) present an example of a deaf girl (aged 2;6) signing correctly in Dutch Sign Language a command to her mother with a sign that depicts putting a puzzle piece in the correct location. However, full mastery takes many more years. Researchers attribute this difficulty to the complexity of some depicting constructions, which may involve the two hands moving in relation to each other, simultaneous facial and other non-manual markers, and the perspective on the depicted scene (Supalla, 1982; Slobin *et al*., 2003; Beuzeville, 2006).

Even with these observations, there has been little follow-up work that takes into consideration the acquisition of these other features of depicting signs in use, in either L1 or L2 contexts. The majority of studies focus on the production and/or comprehension of single depicting signs, e.g., a focus on a depicting sign's formational parameters (handshape, location, orientation, and movement) (see Marshall and Morgan, 2014 for such a study in an L2 context). Only in a few instances have researchers considered non-manual actions, which are actions that occur on body parts other than the hands—e.g., head and torso movements, facial expression, mouth movements, and eye gaze (McIntire & Reilly, 1998). To date there has been only one study that focuses on the hearing L2 acquisition of depicting signs as they are used in larger scene depictions (Ferrara & Nilsson, 2017). Findings from that study provide further evidence that L2 learners not only struggle with the form of depicting signs, but that they also struggle with coordinating depicting signs within the signing space and in relation to their own bodies. The current study continues on from this recent work and focuses specifically on how depicting signs and eye gaze, one type of non-manual action, are coordinated to establish vantage points from a perspective on depicted scenes, and the differences seen between fluent and beginner-intermediate signers. First though, the signing space and visual perspective in signed languages is reviewed in the context of depicting signs.

Depicting signs and the signing space

Depicting signs are produced in the space in front of the signer and resulting spatial mappings help prompt meaning construction. Describing how signers use the signing space has been a major theme in signed language linguistics, as space is an essential aspect of signed language structure. For this paper, the topographic use of the signing space is in focus, which involves the signing space being used to express spatial meanings and relationships (as opposed to being simply a place where signs are articulated). Earlier work made a distinction between two kinds of topographic spaces, *diagrammatic space* and *viewer space*, which differ in scale, dimensions, vantage point, and placement (Emmorey, 2002). Diagrammatic space is scaled-down to the space in front of the signer and the scene is often conceptualized from a static vantage point positioned outside and above the scene. Scenes in viewer space, however, are conceptualized as life-sized and the vantage point is positioned within the scene and

can move over the course of the discourse. The characteristics of these two spaces are summarized in Table 1.

Table 1. Properties associated with diagrammatic and viewer space (based on Emmorey, Tversky, & Taylor, 2000, p. 167)

Property	Diagrammatic space	Viewer space
Scale	Signing space represents a map-like model of the environment	Signing space reflects an individual's view of the environment at a particular point in time and space
Dimensions	Space can have either a 2-D "map" format or a 3-D "model" format	Signing space is 3-D (normal-sized scale)
Vantage point	The vantage point does not change (generally a bird's eye view)	Vantage point can change
Placement	Relatively low horizontal signing space or a vertical plane	Relatively high horizontal signing space

Liddell (2003), working within a cognitive linguistics framework, describes the use of space in signed languages as instances of real-space blending, which is based on the more general Mental Spaces Theory and Blending Theory (Fauconnier, 1994, 1997; Fauconnier & Turner, 1998). Real space blending is a cognitive process that maps elements from the discourse onto the immediate physical environment of the signer. In this way, signers and interlocutors can conceptualize space and even the signer's hands and body as someone or something else. When some entity or character from the discourse is mapped partially onto the signer's body, the result is a surrogate blend within a life-sized *surrogate space*. Surrogate space is similar to the viewer space described by Emmorey and colleagues (2000), because spaces both are conceptualized as normal-sized, and both reflect a mobile vantage point within a scene. However, they are also different because viewer space emphasizes the viewed environment, while surrogate space profiles the actions of an enacted entity, i.e., a "character." Dudis (2011) also noted this difference during an expanded account of depiction in signed languages. He observed the possibility for a scene-focused or character-focused depiction. He explained that in life-sized depictions the body doesn't have to depict a body but may "[make] manifest the vantage point from which the scene is depicted" (Dudis, 2011, p. 27-28). This observation aligns with the description of viewer space, because an environment rather than a character is more salient.

Liddell (2003) proposed another type of blended space called *depicting space*, which is a "topographical real-space blend separate from the signer" (p. 367). Here, the signer is positioned outside the conceptualized scene that unfolds in front of the signer. Depicting space is often viewed from a more stationary vantage point, similar to diagrammatic space. Depicting space provides scaled-down representations—in contrast to the life-sized representations produced within viewer and surrogate space. The simultaneous and sequential use of depicting and surrogate space has also been observed and documented (Liddell, 2003; Dudis, 2004).

While this study primarily adopts the notions of diagrammatic and viewer space, aligning with Emmorey & Falgier (1999) and Emmorey, Tversky, & Taylor (2000), work by Liddell (2003) and Dudis (2004, 2011) highlight some important aspects of depiction

in signed languages. First, their work showed that signers build up complex scenes over time, which are conceptualized as physically present in the space in front of and around the signer. Furthermore, a signer can indicate and reference elements within these scenes with signs (e.g., pointing) and non-manual behaviors, such as eye gaze and body orientation. An example from Engberg-Pedersen's (2003) work on Danish Sign Language illustrates this possibility. She presents an example of a signer depicting water coming down the walls of her bathroom with her hands. "During the construction, [the signer] looked up, moving her head and her eye gaze from side to side as if looking at the wall, that is, an imaginary configuration in space" (p. 283).

A second aspect of depiction highlighted by Liddell (2003) and Dudis (2004, 2011) is that, at different points in time, depicted elements may be manifested visibly or invisibly in the signing space. In the Danish Sign Language example above, the water is depicted visibly by the signer's hands while her eye gaze and head indicate a wall, which is invisible but still conceptualized as present. Another example from the current study's Norwegian Sign Language dataset further illustrates how signers interact with visible and invisible entities within depicted scenes. The sign DSL:BUILDING-LOCATED-AT in Figure 2 visibly depicts an office building. The signer continues after this sign to explain that there is a parking lot next to the building (not pictured) and that to get to the new campus one begins by getting into a car and driving (out of "the parking lot") (illustrated by the last two images in Figure 2). During this utterance SIT DRIVE, "in your car, you drive out of the parking lot," the building is no longer visibly depicted with the signer's hand, but it is still conceptualized as present in the signing space. This is evident because the signer articulates the signs SIT and DRIVE to her left side (a noncanonical position), which is the space conceptualized as "the parking lot." During the sign DRIVE, she orientates her body and moves her hands away from both "the parking lot" and "the building" as she begins to describe the route to the old campus (illustrated by the final two images in Figure 2). This is similar to the Danish Sign Language example mentioned above in that the signer here indexes invisible elements within the scene through her eye gaze and body orientation. It should be stated that both visible and invisible elements are essential to the meaning of these depicted scenes.

Figure 2. Visible and invisible elements in a depicted scene.

The examples above demonstrate how both visible and invisible elements within depicted scenes (either in diagrammatic or viewer space) are conceptualized as present in the signer's physical environment and are indicated and referenced to with manual and non-manual actions. In this way, these examples demonstrate that depiction in signed languages is a whole-body affair.

Perspective

The relation between depicting signs and the topographical use of the signing space has been investigated during studies on perspective in signed language discourse. For example, Perniss (2007, 2012) examined the use of *character* and *observer perspective*, which places the signer within a life-sized event representation or outside of a scaled-down event representation, respectively. She also explained that signers achieve spatial coherence by using depicting signs with signing perspective (Perniss, 2007). Other researchers have echoed this observation, even though they use different terminology and work with different linguistic approaches (Schick, 1990; Morgan, 1999; Liddell, 2003; Slobin *et al*., 2003; Engberg-Pedersen, 2016). Some researchers have also made connections between signing perspective and McNeill's (1992) co-speech observer viewpoint gestures and character viewpoint gestures (e.g., Perniss, 2007; Cormier *et al*., 2012; Stec, 2012; Quinto-Pozos & Parill, 2015). Many of these observations are based on analyses of signed language narratives, and perhaps because of this they also emphasize the role of the signer in the represented events—as a character acting within a scene or as an outsider watching a scene unfold.

The signed language data analyzed for this study are not narratives but rather directions to places and descriptions of spatial scenes. In such settings, it has been shown that signers prefer to depict from an egocentric viewpoint, describing scenes from how they perceive it (e.g., Perniss, 2007; Pyers, Perniss, & Emmorey, 2015), similar to the relative frame of reference observed in spoken languages. In addition, researchers have used the terms *route*, *survey*, *gaze*, and *mixed perspectives* to describe the perspectives speakers and signers take on the spatial scenes they are describing, rather than using character and observer perspective (e.g., Linde & Labov, 1975; Ehricha & Kosterb, 1983; Emmorey & Falgier, 1999; Emmorey, Tversky, & Taylor, 2000). These terms are adopted in this paper and are clearly distinguished from each other through associations between particular vantage points and diagrammatic or viewer space. By doing this, it is assumed that this type of perspective in signed language discourse is fundamentally visual and spatial in nature.

A route perspective is adopted when a signer takes an interlocutor on a mental tour of the environment, with the vantage point reflecting movement within the conceptualized scene. Survey perspective, by contrast, locates the signer at a stationary vantage point outside and above the conceptualized scene. Emmorey, Tversky, and Taylor (2000) related route and survey perspective with viewer and diagrammatic space, respectively. They noted that a route perspective in viewer space results in a conceptualized life-sized scene, whereas a survey perspective in diagrammatic space scales down the depiction to fit in the space in front of the signer, which the signer can then look down upon, like a map on a table. They also observed that signers sometimes alternated between perspectives and spaces. Route and survey perspectives are summarized in Table 2.

Table 2. Summary of perspective, spatial format, and vantage point as described in Emmorey & Falgier, 1999 and Emmorey, Tversky, & Taylor, 2000.

Perspective	Spatial format	Vantage point
Route	Viewer	Moving, Within scene

Survey	Diagrammatic	Stationary, Outside/above scene
Gaze	Viewer	Stationary, Within/at edge of scene
Mixed (alternations between survey and route)*	Viewer & Diagrammatic	Moving and stationary, Within and outside scene

*While attested in the studies by Emmorey and colleagues, a mixed perspective was not observed in the current study and thus will not be mentioned further here.

Emmorey and Falgier (1999) also briefly discussed the use of a gaze-tour as a "description [that] does not describe movement through space; rather, the environment is described from a fixed vantage point from which a signer or speaker views the environment (see Ehrich and Koster, 1983)" (p. 18) A gaze tour effectively pairs viewer space with a fixed vantage point, combining elements typical of a route and survey perspective. Gaze-tour perspective appears similar to Dudis' (2011) observation that a signer's body in a life-sized scene depiction may simply provide a vantage point and is not always interpreted as an active participant within a depicted scene. Although Emmorey & Falgier (1999) mentioned the possibility of a gaze tour and even gave an example from their data, they did not integrate the use of gaze tour within their overall findings on perspective choice in their study. This perspective does prove useful for the current study's dataset however, and thus is included here (see Table 2). In the findings, these perspectives will be used to initially characterize some of the differences observed between the instructors' and the students' signing.

Eye gaze in signed (and spoken) language discourse

The brief introduction above to depicting signs and perspective highlighted some important aspects of signed language structure. It also implicated eye gaze as an important non-manual feature in these settings, although there has yet to be focused research on this type of coordination (but see Engberg-Pedersen, 2003). There is, however, a body of research that shows the important and varied functions eye gaze has to other aspects of signed language structure, including different grammatical purposes (e.g., Neidle, *et al.*, 2000; Tang & Sze, 2002). Some work has also considered eye gaze behavior during periods of enactment in signed (Padden, 1986; Reilly, 2000; see also Cormier, Smith & Sevcikova-Sehyr, 2015) and spoken languages (e.g., Sidnell, 2006). Research on co-speech gesture has also documented speakers directing their gaze towards gestures and the gesture space during spoken language interaction (e.g., Streeck, 1993; Hayashi, 2005).

This research suggests that eye gaze may indeed be important for signed language depiction. The current study aims to examine how eye gaze contributes to visual perspective and explore the challenges students have with this type of multimodal meaning-making. In the following sections, the methods for this exploratory study will be detailed and the data will be described. Findings and analysis will examine how fluent signers and learners give directions and describe spatial scenes with depicting signs and eye gaze. The multimodal nature of depicting from a particular visual perspective will be described, along with its challenges to L2 signers.

Methods

Participants

The data for this study come from video-recordings of deaf instructors and hearing students responding to prompt questions aimed to elicit directions or descriptions of locations. The students are 12 women, who were in their second year of a bachelor's program in Norwegian-Norwegian Sign Language interpreting. They all reported Norwegian (or in one case Swedish) as a first language, and only two reported having experience with Norwegian Sign Language before entering the program (although, it should be mentioned that these two students were still beginner signers upon starting the program).

At the time of the study, the students had completed one and a half years in the three-year program. During the first year, the students received approximately 12-16 hours of Norwegian Sign Language instruction each week (approx. 300 hours)—although, attendance was not counted nor obligatory. In the second year, the students received targeted signed language instruction approximately two hours a week (for approximately a total of 20 hours up to the point of data collection), with additional exposure in some of their other subjects and when they were out on practicum. The students reported in a questionnaire that they used Norwegian Sign Language outside of school contexts between one and 10 hours a week, with half the students reporting two hours or less a week (M=3.9, SD=2.9). Students also varied in how much they used Norwegian Sign Language as part of school activities, but outside of teaching hours. Here we find a range from zero to 10 hours, with most students again reporting two hours or less each week (M=2.9, SD=2.6). This group of second year students was chosen, because they were expected to be able to produce depicting signs in the context of giving directions and describing spatial layouts even though they still had not mastered this complex language setting. This study was in part an attempt to uncover some of the issues students still had with depicting in these contexts.

The students' primary Norwegian Sign Language instructors also participated in this study. The three instructors use Norwegian Sign Language in their daily life and are active members of the Norwegian Deaf community. Their workplace is multilingual, with face-to-face interaction occurring in Norwegian Sign Language. One of the instructors, DS1, was born deaf and began learning signed language early when he started attending deaf school at around six years old. He was in his early 60s at the time of the study. Another instructor, DS2, was also born deaf and began learning Norwegian Sign Language in kindergarten when he was around four years old. At the time of the study, he was in his late 40s. The third instructor, DS3, was born hearing and became deaf when she was 16 years old. She then began to learn Norwegian Sign Language. At the time of this project, she was in her late 30s. Her Norwegian Sign Language exhibits higher degrees of contact with Norwegian than the other instructors, but she is considered to be a fluent signer. In this study, data was collected from all three instructors, because they were the students' primary language models and because they demonstrate the linguistic diversity within the Norwegian Deaf community, where L2 acquisition happens.

Data elicitation and collection

The data comprise video-recorded responses by the students and instructors to two prompt questions in Norwegian Sign Language, which are summarized in (1).

> (1) a. How do you explain how to get to the new campus from the old one to someone who has never been there before? Either by walking, driving, or taking the bus.
>
> b. Can you please describe the third floor of the building where you have classes? For example, where do you find the large and small teaching rooms, the social area, etc.?

The questions were posed to the students by one of their instructors and formed part of a larger elicitation that targeted both the students' Norwegian Sign Language and their interpreting skills. These questions aimed to elicit samples of Norwegian Sign Language that included depicting signs and a topographic use of the signing space. Question 1a was designed to prompt the participants to give step-by-step instructions for how to walk, drive, or take the bus from one location to another, while Question 1b aimed to elicit descriptions of static spatial scenes.

All questions and responses were filmed with one high-definition video camera focused primarily on the participant. During the student elicitation, the instructor and the student were the only two people in the room. The researchers were present outside of the room if needed, however, and would check the recording equipment between elicitations. The instructor was free to repeat and re-word the questions if needed, although this option was not always exercised. The students could also clarify and follow up on the questions as needed. The two questions were also asked to the students' three Norwegian Sign Language instructors by both researchers, and their responses provide baseline data for the comparison and analysis presented here. This data collection and resulting research work has been approved by the Norwegian Centre for Research Data. All participants gave their consent to participate in the project and also consented to images of them being used in research publications.

In total 12 student responses were collected for Question 1a, with a mean length of 56 seconds (SD=42 seconds) and range between 17 seconds and 151 seconds. The three instructors' responses to this question took more time, with DS1 taking 89 seconds, DS2 taking 88 seconds, and DS3 taking 69 seconds. Twelve student responses to Question 1b were also recorded. These responses tended to be longer, with a mean duration of 80 seconds (SD=37 seconds) and a range between 39 seconds and 167 seconds. The instructors' responses to Question 1b where shorter than their responses to Question 1a, with DS1 signing for 50 seconds, DS2 for 82 seconds, and DS3 for 49 seconds.

Data annotation

Preliminary annotation of the resulting video data occurred during two previous studies (Nilsson & Ferrara, 2015, 2016, Ferrara & Nilsson, 2017). ELAN, a computer program developed at the Max Planck Institute of Psycholinguistics in Nijmegen, The

Netherland was used to do this work.[4] ELAN synchronizes video segments and annotations that are created on user-defined tiers (Crasborn & Sloetjes, 2008). First, the instructors' and students' responses were tokenized into single signs on right- and left-hand glossing tiers, with particular attention to depicting signs. These annotations were discussed, proofed, and amended over time across multiple passes by the two authors. In addition, a Norwegian Sign Language instructor (who did not participate in the study) was consulted on all of the depicting sign annotations. Depicting signs were identified and glossed according to the conventions used for the Auslan Corpus (Johnston, 2016). The code prefixes first introduced above in Figure 1 were used to identify depicting sign subtypes. An additional sub-type, glossed DS?, was also used, which indicated that there was uncertainty (on the part of the annotator(s)) regarding what the signer was trying to depict with the sign. Following the sub-type prefix, the meaning of the sign in context was given, e.g., DSS:ROAD-EXTENDS in Figure 1. Two-handed depicting signs received a gloss on both the right- and left-hand gloss tiers to reflect each hand's contribution to the sign as a whole. For example, the glosses DSM:MAN-RIDES-HORSE and DSM:HORSE-MOVES in Figure 1 indicate that the signer produces a two-handed sign depicting the movement of entities. The signer's dominant hand depicts a man riding a horse, while the non-dominant hand depicts a horse.

During these earlier studies, the data were also tagged for visual perspective. However, these annotations were revisited and checked again for the purposes of the current study. The perspective engaged by a signer was identified based on the characterizations presented in the sections above as well as by following the methods described in Emmorey & Falgier (1999) and Emmorey, Tversky, & Taylor, (2000) who identified a perspective by considering the general question of whether the response "felt more like a 'tour,' a birds-eye view description, or a mixture of both" (i.e., route, survey, and mixed perspectives respectively, Emmorey & Falgier, 1999, p.6). In addition, here, we also asked whether the response described an environment with a fixed vantage point within the scene (i.e., gaze tour, Emmorey & Falgier, 1999, p. 19). During this review of the data, two additional labels were created. One label identified non-target responses (i.e., the student did not give directions or provide descriptions of spatial scenes). The other label was used for target responses where the student did not engage with the signing space and did not depict a particular perspective.

In addition, all depicting signs produced by the instructors and students were revisited and tagged for eye gaze behavior on a dedicated tier. The categories annotated are summarized in Table 3 and illustrated in Figure 3. They can be summarized into four main categories: to interlocutor, towards signs and depicted spaces, away (from both the interlocutor and the signing space), or indeterminate.

Table 3. Eye gaze and body orientation tags.

DS-eye gaze tier	
tag	*Description*
to interlocutor	eye gaze directed towards camera/interlocutor
to sign	eye gaze directed towards a particular sign

[4] Please see http://www.lat-mpi.eu/tools/elan/ for more information regarding this free annotation software.

to depicted space	eye gaze directed towards an element (visible or invisible) in the depicted scene
away	not directed to audience or depicted space
closed	eyes are closed
(thinking)	eyes are directed away from both interlocutor and signing space, often directed up and to the side
?	The annotator was unable to confidently identify eye gaze in the given instance, even if a suggestion was provided. These cases are analyzed as indeterminate.

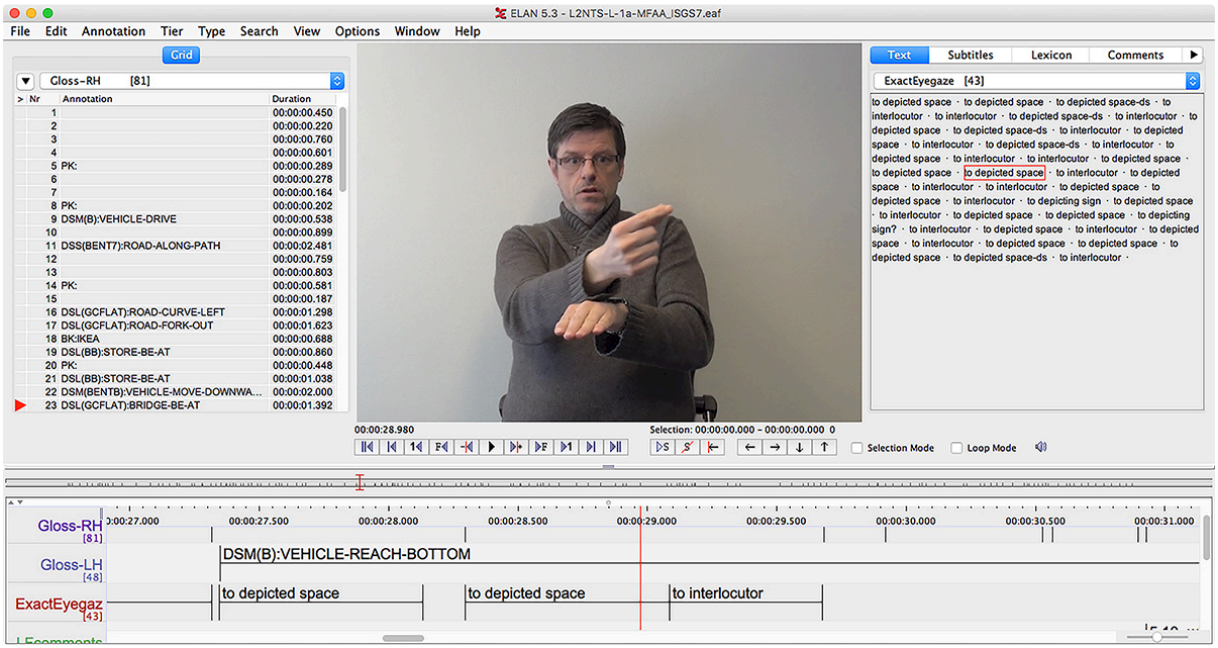

Figure 3. A screenshot of ELAN showing the annotation of signs and eye gaze during a segment depicting "a car driving down a hill; to the left is a bridge."

An example of the annotations is provided in Figure 3. Here, the signer directs his gaze to a depicted bridge located by the sign produced on the right hand. Thus, eye gaze was tagged in this moment as 'to depicted space.' Illustrations of the other types of eye gaze behavior are provided in examples detailed in later sections.

Findings related to multimodal perspective in fluent and L2 Norwegian Sign Language

The findings and examples that will now be presented are based on an analysis of the annotated data. An initial comparison of the perspectives adopted by the instructors and students is presented. Then findings related to eye gaze behavior during the production of depicting signs are outlined, before presenting some detailed examples that illustrate and contextualize these findings and the subsequent discussion.

Fluent and L2 perspective in scene depictions

First, the data show that fluent signers and learners do adopt visual perspectives while depicting directions and spatial scenes, aligning with observations made from previous

studies on other signed languages. For instance, all of the instructors' responses were able to be characterized as a particular visual perspective—here, either route or gaze. Most of the student responses were also able to be characterized for visual perspective—either survey, route, or gaze. Only two target responses lacked a visual perspective, which is explained by a heavy reliance on lexical signs instead of depictions of spatial scenes.

While the instructors and students did depict from a particular visual perspective, they differed as to which ones. For the prompt "How do you get from the old campus to the new campus?," all of the instructors depicted a scene from a route perspective (viewer space, moving vantage point within the scene). Among the students, only half answered this prompt question with the target response of directions from point A to point B.[5] Of those six target responses, four students attempted a survey perspective (diagrammatic space, stationary vantage point outside and above the scene) and two students chose to give directions lexically (without depicting a visual scene in the signing space, as mentioned in the previous paragraph). These findings are summarized in Figure 4.

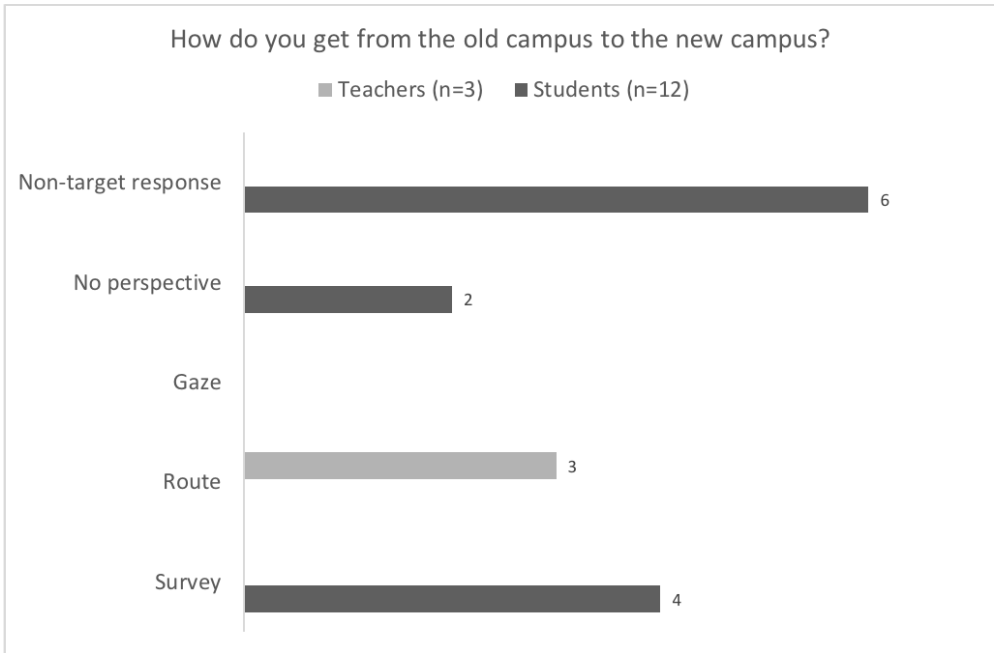

Figure 4. Perspectives of the student and instructor responses to the prompt question "How do you get from the old campus to the new campus."

For the prompt "Describe the floor with classrooms," all of the instructors chose to describe the floor from a gaze perspective (viewer space, stationary vantage point within the scene). One of the students also depicted with a gaze perspective. Another chose to depict from a route perspective (viewer space, moving vantage point within the scene). However, the rest of the students depicted from a survey perspective (diagrammatic space, stationary vantage point outside and often above scene). These findings are summarized in Figure 5.

[5] The students often misinterpreted this question by replying in the spirit of "I take the bus." Their interlocutor did not redirect the students in these cases but simply continued with the elicitation.

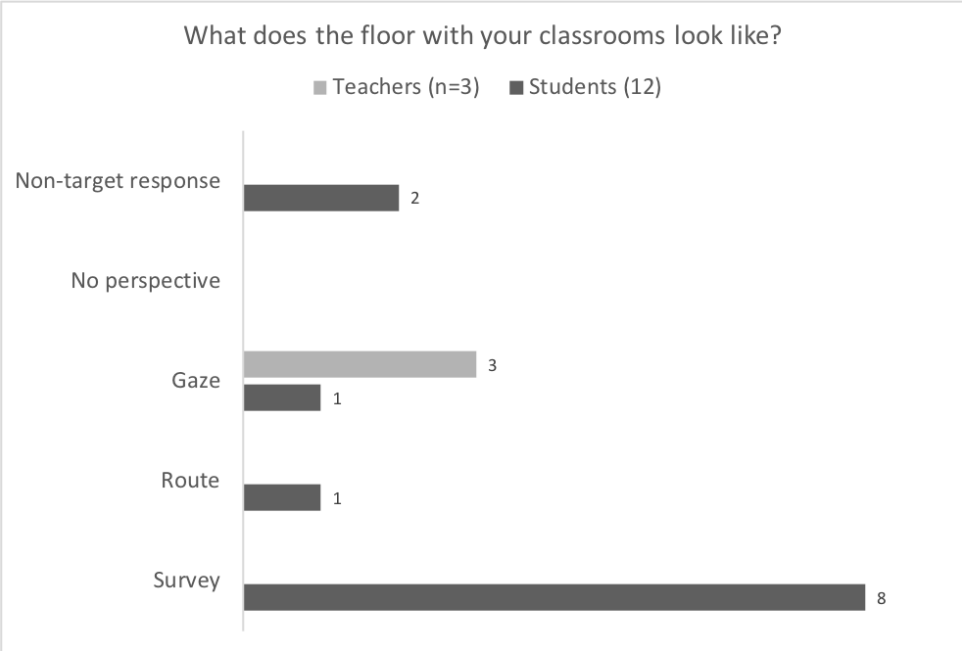

Figure 5. Perspectives of the student and instructor responses to the prompt "Describe the floor with your classrooms."

These initial observations show some interesting differences in the perspectives adopted by the instructors and students. The instructors, even though there were only three of them, consistently made similar decisions about how to best depict locations and give directions. The students, on the other hand, were much more heterogeneous in their choices, even though they favored a survey perspective across both elicitation tasks. More research into how deaf native signers engage different visual perspectives during depictions of various scenes and settings is needed to help establish more robust baselines with which learner data can be compared. For now, this initial data show that there may be interesting differences in the perspective fluent signers and learners choose for their depictions. These findings have implications for the coordination of eye gaze with depicting signs, which is more closely examined in the next section.

Eye gaze direction during depictions produced by instructors and students

An examination of the eye gaze behavior of instructors and students during periods of depiction show some qualitative and quantitative differences. Some of these differences appear to relate to the different choices the two groups of signers made regarding perspective. For example, the way a signer directs their eye gaze while depicting from a route perspective will differ from a depiction from a survey perspective, because the vantage points of such perspectives are different (one within and moving in the scene and one outside and stationary to the scene).[6] In addition though, it is suggested that the differences in the distribution of eye gaze behavior is also a reflection of the students' language skills.

First, the amount of time each group spent looking at their interlocutor, the depicted space (which includes looking at depicting signs), or away (including closed eyes) was

[6] It should be noted that in this study no instances of survey perspective were elicited from the instructors. This means we cannot make describe how particular perspectives affect eye gaze behavior. This will need to be followed up in a more dedicated study.

calculated. Figures show that both groups looked towards their interlocutor in similar proportions, around 30%. However, the students directed their gaze to the unfolding depicted space far less often than the instructors (29% vs. 55%, respectively). Instead, students directed their gaze most often away (36%)—to the side or upwards—from both the depicted space and their interlocutor. Indeterminate gaze accounts for 5% of the student data and 2% of the instructor data.

A more detailed examination of the data revealed variation between the two sets of responses as well as within and across the instructor and student groups in relation to the overall amounts of depiction as well as eye gaze behavior during these moments (see Figures 6 and 7).

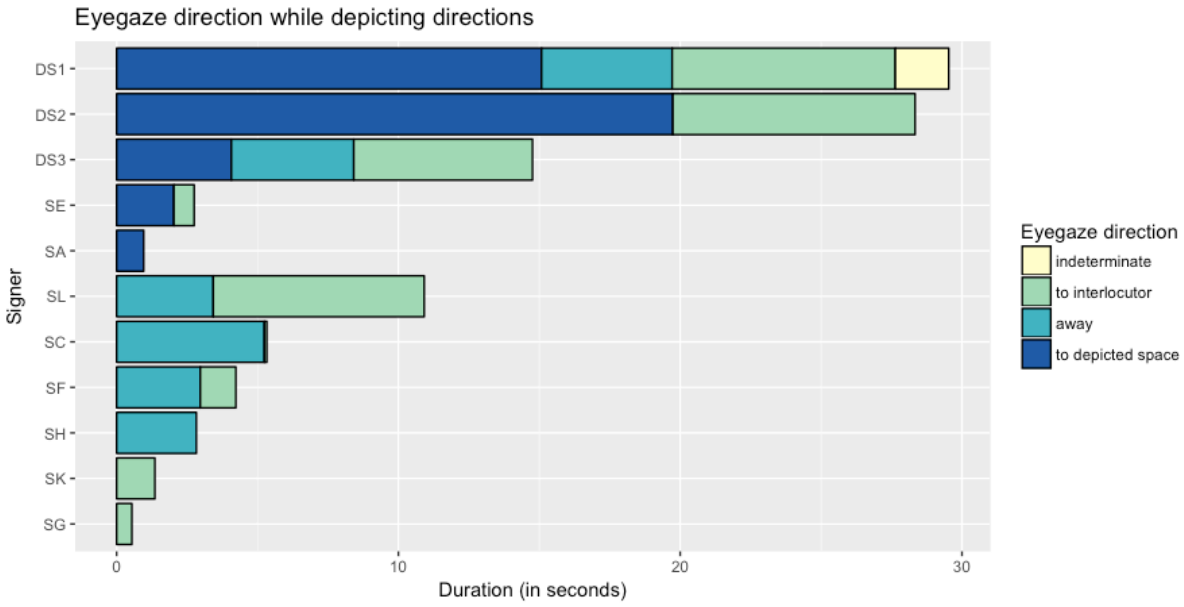

Figure 6. Eye gaze direction during depictions of directions (Note: the top three bars labeled DS1, DS2, and DS3 are the instructors' data).

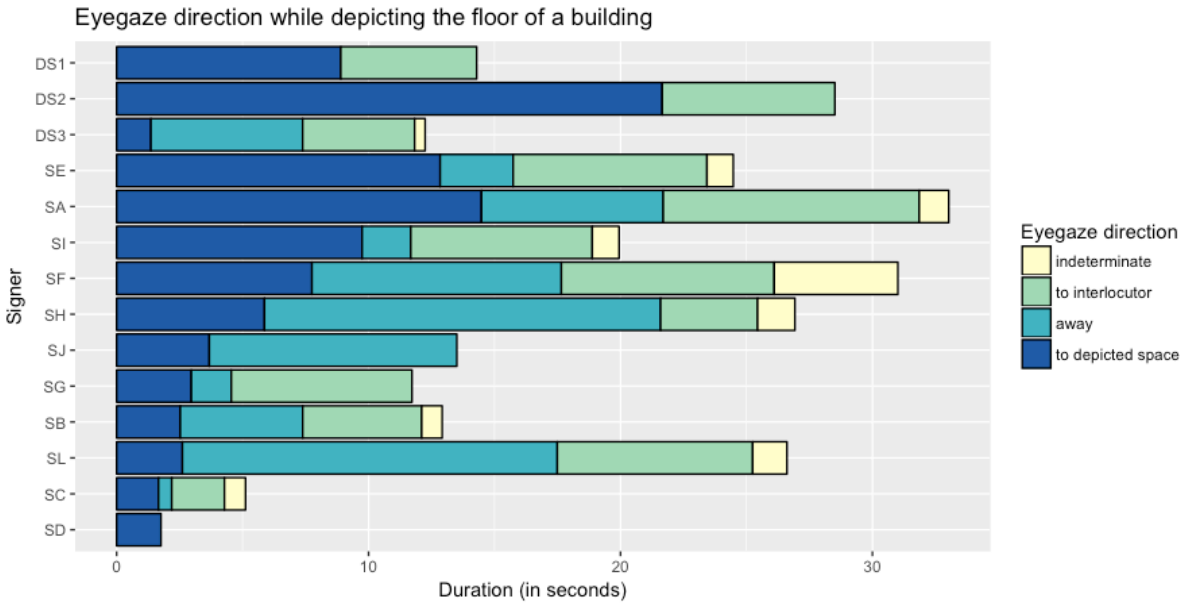

Figure 7. Eye gaze direction during depictions of the floor of a building (Note: the top three bars labeled DS1, DS2, DS3 are the instructors' data).

To begin, Figure 6 presents the distribution of where the instructors and students looked during their directions from the old campus to the new campus. The instructors, DS1, DS2, DS3, all gazed at the depicted space during their production of depicting signs. In the cases of DS1 and DS2, these instructors looked at the depicted space either half the time or more. The students on the other hand, except for SE and SA, do not look at the depicted space during their production of depicted signs. Instead, the students have higher proportions of time looking away—e.g., to the side or upwards—from both their interlocutor and the depicted space.

Eye gaze behavior in the second set of responses, about the spatial layout of the floor, presents a different picture (see Figure 7). First, all of the students looked towards their depictions at least some of the time during their depictions—in marked contrast to their directions to the new campus. This shows that they can engage with their own depictions non-manually. The students also continued to direct their gaze towards their interlocutors in similar proportions to the instructors. However, the distribution in Figure 8 also shows that, similar to the giving-directions responses, students varied in how much time they looked away during depicted sequences. Compare for example students SE and SI with SH and SL, who produce depictions of similar length (between 20-27 seconds each) (in Figure 8). SE and SI direct their gaze away 12% and 10% of the time, respectively, while SH and SL direct their gaze away 58% and 56% of the time, respectively.

Among the instructors, DS1 and DS2 never look away during depicted sequences and instead direct their eye gaze between their depicted spaces (62% and 76%, respectively) and their interlocutor (38% and 24%, respectively). In contrast, the instructor DS3 spent more time looking away (49%), with very little gaze towards her depicted space (11%). In this way, her eye gaze behavior patterns more similar to some of the students' eye gaze behavior.

These findings lead us to consider further eye gaze behavior directed away from both the depicted space and the interlocutor. For both sets of responses, students have a higher proportion of looking away than the instructors, with the possible exception of the responses by instructor DS3. One potential explanation for these behaviors could be that these questions were simply harder to answer—with the students and the instructor DS3 taking more time to think how to answer the prompts. In addition, it was clear from both sets of responses that students used a non-directed gaze to think about *how* to sign what they wanted to say, which reflects their beginner-intermediate skills in Norwegian Sign Language.

Both the instructors and students varied in how they directed their eye gaze during periods of depiction. The two instructors, who learned Norwegian Sign Language early in life, more often directed their gaze towards the depicted scene. The L2 learners did this less often, including the instructor who learned Norwegian Sign Language as a teenager (and so is also an L2 learner, albeit an advanced one). In the following sections, examples produced by the instructors and students are detailed to illustrate these different eye gaze behaviors. Discussion will focus on the role eye gaze plays in establishing a vantage point upon a depicted scene.

Comparing depictions produced by instructors and students

Depicting a floor with classrooms and hallways

A first set of examples comprise depictions from one instructor and one student as they describe the entrance to the floor of a building, an adjacent hallway, and three classrooms (Question 1b). The instructor engaged a gaze perspective for his depiction, while the student depicted from a survey perspective. The student also directed her gaze towards her depictions much less than the instructor, which illustrates the findings presented above.

An instructor's depiction of the floor

The first sequence is produced by the instructor DS1 and is depicted from a gaze perspective. It is composed of seven distinct signs, grouped into three to four utterances, and begins after the signer has already explained that there is a large classroom directly ahead of the front entrance. As Figure 8 shows, the signer then looks at his interlocutor as he begins to sign DS?:BACKWARDS, "moving back towards the front door." Midway through this sign, the signer moves his gaze towards the signing space and the emerging depicted scene, specifically the hallway.

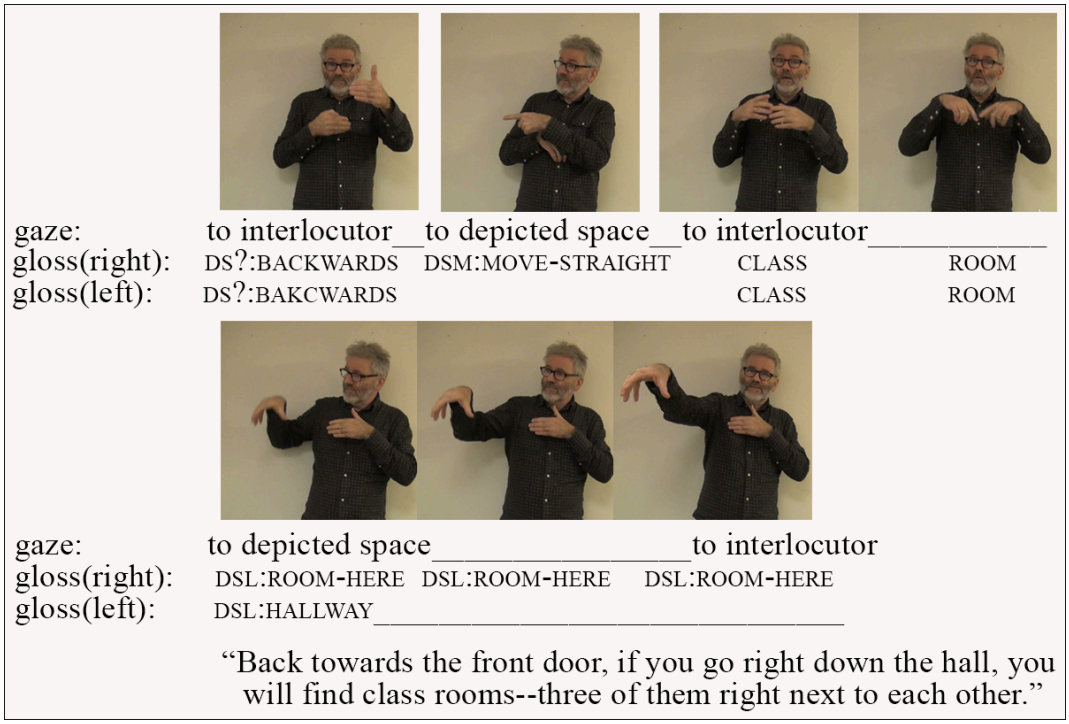

Figure 8. An instructor depicts a part of a floor in a building.

As he begins to produce a second depicting sign, he shifts his body to the right and his gaze turns towards the location the depicting sign indicates ("at the end of the hall", second image from the left in the top row of Figure 8). In this way, the signer uses his vantage point to look down the conceptualized hallway, even though this is not explicitly stated. He then produces the sign DSM:MOVE-STRAIGHT in a controlled manner, which is interpreted as both movement and direction ("as you go down the hall"). Together with this sign and eye gaze, the signer moves his vantage point from the front door to the end of the hall within the conceptualized scene.

The signer continues by producing the signs CLASS and ROOM, "classroom," while making eye contact with the interlocutor. Then he shifts his eye gaze and body back towards the depicted space and produces a series of depicting signs that locate three classrooms within the conceptualized scene. During this sequence, the signer depicts from a vantage point at the end of the hall, and so locates the classrooms as extending outward on his right side. This is reinforced as the signer's eye gaze follows the placement of the classrooms and "sees" where they are. At the end of the final depicting sign, the signer returns his gaze towards his interlocutor as a comprehension check.

The signer in this example consistently directed his gaze towards the emerging depicted scene and the various visible and invisible elements. It should be noted that the signer did not depict himself, or any other character, within the scene. However, he established a vantage point with his eye gaze that showed his own position in relation to the conceptualized scene. Such a position can be used by an interlocutor to build up their own conceptualization of the scene.

A student's depiction of the floor

This next example presents a depiction of the same part of the floor produced by a student (shown in Figure 9). In contrast to her instructors, this student depicted from a survey perspective and had trouble sizing her depiction, which ended up extending outside the normal signing space (see the final three images on the bottom row of Figure 9, where she leans forward and extends her arms to place the sign DSS:ROOM).

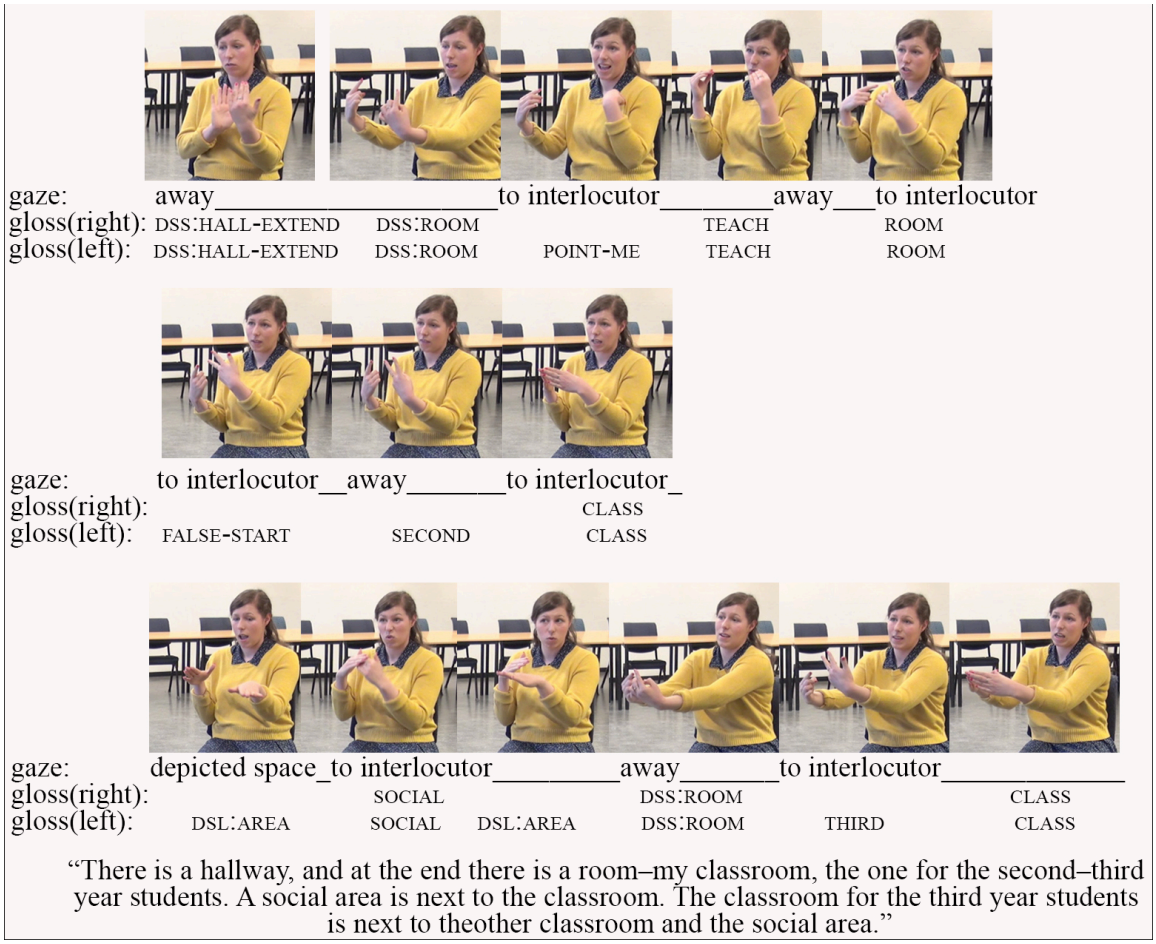

Figure 9. A student depicts a part of a floor in a building.

The student also does not look at her depiction in the same way that the instructor above did. She only looks to her depicted scene once during this approximately 14-second-long segment (see the first image glossed DSL:AREA in the bottom row of Figure 9). Instead she most often moves between looking at her interlocutor and away (e.g., see the signs produced during the beginning of this segment, shown on the top row in Figure 9). As a result, she creates a disconnect between herself and the emergent properties of the depiction and fails to establish a vantage point upon the conceptualized scene. The overall unified structure of the depicted scene must be inferred from only the placement of signs in the signing space.

Depicting directions through a city

An additional pair of examples are presented in this section, which include an instructor and student giving directions through the city (Question 1a). In these responses, the instructor depicted from a route perspective, and the student depicted from a survey perspective. Similar to the examples in the previous section, the instructor established a vantage point with his eye gaze during his depiction, which worked to lead an interlocutor through the conceptualized scene. Such coordination is lacking in the response by the student.

An instructor depicting directions from one place to another

The instructor DS2 in this example has just pointed to a location and signed "before you reach student union building." He then follows this utterance with a depiction of this meaning. He does this with two simultaneous depicting signs (illustrated in Figure 10). His right hand, DSL:BUILDING-EXTEND-HERE, depicts the student union building while his left hand, DSL:VEHICLE-BEFORE-BUILDING, depicts a car that has not yet reached the student union building but is driving towards it.

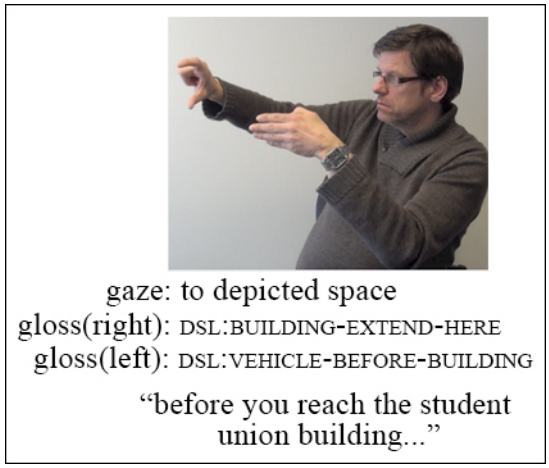

Figure 10. An instructor depicts directions through a city from a route perspective.

Unlike in the previous example where the instructor depicted a floor with classrooms from a gaze perspective, the fluent signer in this example gave directions from one place to another from a route perspective. This perspective was partly realized through the movement of the vantage point as it followed a depicted car through an environment. Specifically here, the placement of the two depicting signs at shoulder height and eye level also suggested a route perspective, because it allowed the signer to look around a conceptualized, three-dimensional, life-sized scene with his eye gaze. Producing signs

higher in the signing space has been suggested to help to establish a vantage point in a route perspective, because the signer can more easily align their line of vision with the depicted scene (Emmorey, Tversky, & Taylor, 2000, p. 168). When the signs and eye gaze are interpreted together, an interlocutor knows to conceptualize the scene from a vantage point located behind the car as it moves through the city.

A student depicting directions

A response by a student is presented here and is compared to the instructor's response presented in the previous section. The response is illustrated in Figure 11 and begins as the student signs BUS, "bus," followed by a sign glossed DSM:VEHICLE-MOVE?. This particular depicting sign is conventionally used in Norwegian Sign Language to depict people moving. However, the sign is ambiguous in this context, because it is preceded by the sign for "bus." The student may have meant to depict a person, a bus, or perhaps a person on a moving bus. In addition to this ambiguity, the student also fails to indicate the direction the bus actually moves, in relation to the city or any other landmarks. During these two signs the student does not gaze towards her depiction or her interlocutor. Instead, she has a non-directed gaze off the center of her signing space, thinking.

The student goes on to explain that one needs to change buses. Then after a pause she produces a final sign depicting that a bus (or herself) keeps moving in the same direction as before. The sign moves even farther away from the student's body and continues with a non-directed, thinking gaze. During this nearly eight seconds, the student only looks twice to her interlocutor—once during the lexical sign TO-CHANGE, "change," and then after her depiction when she signs HERE, "here." She does not look at her depicted scene during this segment, and as a result she does not establish a vantage point from which to conceptualize the scene.

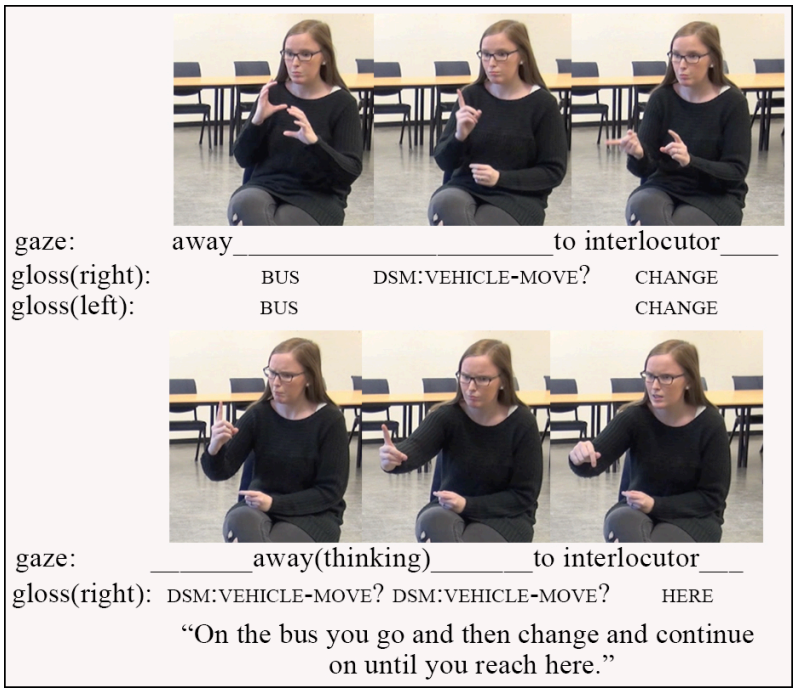

Figure 11. A student depicts the movement of a person/bus through a city.

A comparison of these two examples illustrate differences in how the student and the instructor relate to their depictions. The instructor depicted on a horizontal plane level with his own vision, establishing a vantage point that moved through the conceptualized, life-sized scene as the depicted car moved. The student on the other hand adopted a survey perspective, keeping her signing lower and farther away from her body. She also did not look around the scene with her interlocutor but rather simply placed signs in the signing space. Without a vantage point, the interlocutor must infer the relation between the signer's own body, her signs, and other conceptualized elements in the depicted scene.

Spatial language and signed language acquisition

The findings from this exploratory study suggest that these beginner-intermediate students do not yet fully coordinate non-manual and manual articulators to depict in Norwegian Sign Language, at least not in the same way as their three instructors do. In some way, these observations align with those from L1 signed language research. Slobin *et al*. (2003) commented that it takes children many years before they master the "simultaneous facial and postural markers" needed for depiction (p. 275). The data presented here indicated that a similar situation may hold for the L2 learners investigated in this study.

In addition, there are some parallels from this study's findings and those from spoken L2 acquisition research which show the difficulties L2 learners have with spatial language (e.g., Coventry, Guijarvo-Fuentes, & Valdés, 2012). Here, however, the matter is less about how the signed language lexicon 'carves up the spatial world,' but rather may be more about how signers map the spatial world onto the physical space around them with depicting signs and other non-manual actions. The hearing L2 speakers in this study were tasked with learning a new way of *depicting* space with a visual language. If they had been learning a new spoken language, their task would have been to reassign elements of the spatial world to the new language's lexicon and grammar. In this sense, the acquisition of a signed language by hearing adults is different from the acquisition of a spoken language.

In this study, L1 and L2 signers were tasked with giving directions and describing a floor's spatial layout. They all used depicting signs, albeit in varying degrees, to do this. During these periods of depiction, the three instructors were homogeneous in their choice of perspective—route perspective to give directions and gaze perspective to describe the floor. The students' responses, however, were more heterogenous and most often engaged a survey perspective. The students and instructors also varied in how they directed their eye gaze during depicted sequences. The instructors, and in particular the two L1 signers, directed their gaze more often to their depicted scenes. Such eye gaze behavior provided a vantage point with which to conceptualize the unfolding spatial scene (as Dudis, 2011 describes). The students and the L2 signing instructor looked to their depicted scenes less, although the individual variation here must be acknowledged. As a result, the L2 signers' scenes were not always depicted from a particular vantage point (except from what could be deduced from, for example, sign placement in the signing space and other non-manual actions such as body orientation). This meant that an interlocutor was tasked with inferring how scenes

were to be conceptualized along with their unified spatial structure, which may prove more or less difficult in any particular case.

The findings from this study showed that the L2 signing instructor patterned more similarly to the students in regard to eye gaze behavior. However, in reviewing her responses and her non-directed gaze, she seems be thinking about her response to the prompts. The students, on the other hand, appeared to be thinking both about how to respond the prompts and also *how* to sign their responses. In addition, while the L2 signing instructor directed her eye gaze less to her depicted scenes, she still positioned and structured her depiction in relation to her own body—which meant that a vantage point was still accessible. Looking through the student data showed that they did not orientate their body to their depicted scenes in the same way. This, along with the other aspects of depiction that the students struggle with, e.g., the form of depicting signs and the coordination of the two hands in the signing space (see Ferrara & Nilsson, 2017), means that they are less successful in depicting spatial scenes. They struggle to map entities and their spatial relationships onto the signing space and are less able to indicate the various visible and invisible elements within the scene. As a result of all of these issues together means that an interlocutor has more of the burden in making relevant inferences about how such scenes are to be conceptualized. These findings indicate that future studies would benefit from considering the interplay between multiple non-manual articulators (e.g., eye gaze as well as body orientation), as well as the role lexical signs may play in such spatial depictions.

Potential factors that may help to explain the observations made in this study relate to the natural process of acquiring a signed language as a second language as well as the influence of pedagogy on this process. Related to the former, it may be that hearing students learning their first signed language need time to develop an ability to coordinate multiple articulators (here, their two hands and eye gaze) during depictions of spatial scenes. It may be that beginner signers have an initial focus on the production of 'simple' depicting signs followed by the production of depicting signs coordinated with non-manual actions (which was also the suggested development for deaf, L1 signers by Slobin *et al*., 2003). This development may also reflect a difficulty in conceptualizing simple vs. complex spatial scenes within the signing space and then how to depict such scenes. Such a hypothesis received some initial support from the findings reported here, where students more often gazed towards their depictions of the floor of a building and less towards their depictions of directions through town. A floor of a building may be considered a more simple, delineated area of space, as it is smaller than a town and has more limited pathways through it. In addition, the floor was more easily depicted from a static perspective (survey or gaze) than the directions through town—which required a mobile perspective (route). Pedagogy may also play a role in L2 development. Signed language teachers should be aware of the various demands of depicting spatial scenes and tailor their teaching to focus on signs as well as non-manual aspects. Students may also benefit from practice depicting various types of spatial scenes. Such hypotheses should be investigated further by looking into L2 learners' language production over time as well as cross-linguistically, in a range of learning environments.

Conclusion

The findings from this study show that L1 and L2 signers coordinate their signs and eye gaze differently during depictions of spatial scenes. The L1 signing instructors often directed their eye gaze towards their signs and signing space and in this way established and maintained a vantage point with which to conceptualize the depicted scene, built up over the discourse. The students, as well as the L2 signing instructor, directed their gaze less frequently to their signing and the signing space, and as a result did not always establish a vantage point upon the conceptualized scene.

By comparing instructors and their students, this study contributes new knowledge about visual perspective and depiction in signed languages by detailing how vantage points are established during depictions and how such vantage points interact with the signing space to create perspective. Further research that investigates systematically when different visual perspectives are recruited is still needed, and this work will require more detail into how deaf, fluent signers coordinate their multimodal repertoire during depictions of spatial scenes (e.g., how other non-manual actions such as body orientation are used). The findings from this study also suggest that L2 learners of signed languages could benefit from more dedicated pedagogy related to this complex aspect of signed language structure. Students need diverse exposure and practice in coordinating their multiple articulators to establish vantage points and depict spatial scenes.

References

Beuzeville, L. de. (2006). *Visual and linguistic representation in the acquisition of depicting verbs: a study of native signing deaf children of Auslan (Australian Sign Language).* (Ph.D. dissertation), University of Newcastle, Australia.

Bowerman, M. (1996). Learning how to structure space for language: A cross-linguistic perspective. In P Bloom, M A Peterson, L Nadel, & M F Garrett (Eds.), *Language and space* (pp. 385-436). Cambridge, MA: MIT Press.

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

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 and Communication, 32*(4), 329-348. doi:http://dx.doi.org/10.1016/j.langcom.2012.09.004

Crasborn, O., & Sloetjes, H. (2008). Enhanced ELAN functionality for sign language corpora. In O Crasborn, E Efthimiou, T Hanke, E D Thoutenhoofd, & I Zwitserlood (Eds.), *The third workshop on the representation and processing of sign languages: Construction and exploitation of sign language corpora* (pp. 39-43). Paris: European Language Resources Association.

Coventry, K. R., Guijarro-Fuentes, P., & Valdés, B. (2012). On the first and second language acquisition of spatial language. *Spatial Cognition and Computation, 12*(4), 219-230. doi:10.1080/13875868.2012.713058

Dudis, P. (2004). Body partitioning and real-space blends. *Cognitive Linguistics, 15*(2), 223-238.

Dudis, P. (2011). The body in scene depictions. In C Roy (Ed.), *Discourse in signed languages* (pp. 3-45). Washington, D.C.: Gallaudet University Press.

Ehricha, V., & Kosterb, C. (1983). Discourse organization and sentence form: The structure of room descriptions in Dutch. *Discourse Processes, 6*(2), 169-195.

Emmorey, K. (2002). *Language, cognition, and the brain: Insights from sign language research*. Mahwah, NJ: Lawrence Earlbaum Associates.

Emmorey, K., & Falgier, B. (1999). Talking about space with space: Describing environments in ASL. In E Winston (Ed.), *Storytelling and conversation: Discourse in deaf communities* (pp. 3-26). Washington, D.C.: Gallaudet University Press.

Emmorey, K., Tversky, B., & Taylor, H. (2000). Using space to describe space: Perspective in speech, sign, and gesture. *Spatial Cognition and Computation, 2*(3), 157-180. doi:10.1023/A:1013118114571

Enfield, N. J. (2009). *The anatomy of meaning: Speech, gesture, and composite utterances*. Cambridge: Cambridge University Press.

Engberg-Pedersen, E. (1993). *Space in Danish Sign Language: The semantics and morphosyntax of the use of space in a visual language*. Hamburg: Signum Press.

Engberg-Pedersen, E. (2003). From pointing to reference and predication: pointing signs, eye gaze, and head and body orientation in Danish Sign Language. In S Kita (Ed.), *Pointing: Where language, culture and cognition meet* (pp. 269-292). Mahwah, NJ: Lawrence Erlbaum Associates.

Engberg-Pedersen, E. (2016). Perspective in signed discourse: The privileged status of the signer's locus and gaze. *Open Linguistics, 1*, 411-431. doi:10.1515/opli-2015-0010

Fauconnier, G. (1994). *Mental spaces: Aspects of meaning construction in natural language*. Cambridge: Cambridge University Press.

Fauconnier, G. (1997). *Mappings in thought and language*. Cambridge: Cambridge University Press.

Fauconnier, G., & Turner, M. (1998). Principles of conceptual integration. In J-P Koening (Ed.), *Discourse and cognition: Bridging the gap* (pp. 269-283). Stanford, CA: CSLI Publications.

Ferrara, L., & Nilsson, A-L. (2017). Describing spatial layouts as an M2 signed language learner. *Sign Language and Linguistics*, 20(1), 1-26. doi:10.1075/sll.20.1.01fer

Forker, D. (2012). Spatial relations in Hinuq and Bezhta. In L Filipović & K M Jaszczolt (Eds.), *Space and time in languages and cultures: Linguistic diversity* (pp. 15-34). Amsterdam/Philadelphia: John Benjamins.

Hayashi, M. (2005). Joint turn construction through language and the body: Notes on embodiment in coordinated participation in situated activities. *Semiotica, 156*(1-4), 21-53.

Janzen, T. (2004). Space rotation, perspective shift, and verb morphology in ASL. *Cognitive Linguistics, 15*(2), 149-174.

Johnston, T. (2016). *Auslan corpus annotation guidelines*. manuscript. Macquarie University. Sydney. http://www.auslan.org.au/about/corpus/

Johnston, T., & Schembri, A. (2007). *Australian Sign Language: An introduction to sign language linguistics*. Cambridge: Cambridge University Press.

Johnston, T., & Schembri, A. (2010). Variation, lexicalization and grammaticalization in signed languages. *Langage et société, 131*(March), 19-35.

Landau, B., Dessalegn, B., & Goldberg, A. M. (2010). Language and space: Momentary interactions. In V Evans & P Chilton (Eds.), *Language, cognition, and space: The state of the art and new directions* (pp. 51-77). London, UK: Equinox Publishing.

Levinson, S. C. (1996). Language and space. *Annual Review of Anthropology, 25*, 353-382.

Levinson, S. C. (2003). *Space in language and cognition: Explorations in cognitive diversity*. Cambridge, UK: Cambridge University Press.

Liddell, S. K. (2003). *Grammar, gesture, and meaning in American Sign Language*. New York: Cambridge University Press.

Linde, C., & Labov, W. (1975). Spatial networks as a site for the study of language and thought. *Language, 51*(4), 924-939. doi:10.2307/412701

Marotta, G., & Meini, L. (2012). Spatial prepositions in Italian L2: Universal and language-specific principles. In L Filipovic & K M Jaszczolt (Eds.), *Space and time in languages and cultures : linguistic diversity* (pp. 289-323). Amsterdam/Philadelphia: John Benjamins.

Marshall, C. R., & Morgan, G. (2014). From gesture to sign language: Conventionalization of classifier constructions by adult hearing learners of British Sign Language. *Topics in Cognitive Science*, 1-20. doi:10.1111/tops.12118

McIntire, M. L., & Reilly, J. S. (1998). Nonmanual behaviors in L1 & L2 learners of American Sign Language. *Sign Language Studies, 61*(Winter), 351-375. doi:10.1353/sls.1988.0034

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

University of Chicago Press.

Metzger, M. (1995). Constructed dialogue and constructed action in American Sign Language. In C Lucas (Ed.), *Sociolinguistics in deaf communities* (pp. 255-271). Washington, DC: Gallaudet University Press.

Morgan, G. (1999). Event packaging in British Sign Language discourse. In E Winston (Ed.), *Storytelling and conversation, discourse in deaf communities* (pp. 27-58). Washington D.C.: Gallaudet University Press.

Neidle, C., Kegl, J., MacLaughlin, D., Bahan, B., & Lee, R. G. (2000). *The syntax of American Sign Language: Functional categories and hierarchical structure*. Cambridge, MA: MIT Press.

Nilsson, A-L., & Ferrara, L. (2015, 18-21 August). *Depicting from experience: learning to give directions in space.* Paper presented at the The Fifth Conference of the Scandinavian Association for Language and Cognition (SALC V), NTNU, Trondheim, Norway.

Nilsson, A-L., & Ferrara, L. (2016, 4-7 January). *Meaningful use of space: easy or hard for hearing L2 learners?* Paper presented at the TISLR 12, La Trobe University, Melbourne, Australia.

Padden, C. (1986). Verbs and role-shifting in ASL. In Carol Padden (Ed.), *Proceedings of the Fourth National Symposium on SIgn Language Research and Teaching* (pp. 44-57). Silver Spring, MD: National Association of the Deaf.

Perniss, P. (2007). Achieving spatial coherence in German Sign Language narratives: The use of classifiers and perspective. *Lingua, 117*(7), 1315-1338.

Perniss, P. (2012). Use of sign space. In R. Pfau, M. Steinbach, & B. Woll (Eds.), *Sign language: An international handbook* (pp. 412-431). Berlin: Mouton de Gruyter.

Pyers, J. E, Perniss, P., & Emmorey, K. (2015). Viewpoint in the visual-spatial modality: The coordination of spatial perspective. *Spatial Cognition and Computation, 15*(3), 143-169. doi:10.1080/13875868.2014.1003933

Quinto-Pozos, D., & Parrill, F. (2015). Signers and co-speech gesturers adopt similar strategies for portraying viewpoint in narratives. *Topics in Cognitive Science, 7*, 12-35. doi:10.1111/tops.12120

Reilly, J. S. (2000). Bringing affective expression into the service of language: Acquiring perspective marking in narratives. In K Emmorey & H Lane (Eds.), *The signs of language revisited: An anthology to honor Ursula Bellugi and Edward Klima* (pp. 415-433). Mahwah, NJ: Lawrence Erlbaum Associates.

Ruiter, Jan P. de, Rossignol, S., Vuurpijl, L., Cunningham, D. W., & Levelt, W. (2003). SLOT: A research platform for investigating multimodal communication. *Behavior Research Methods, Instruments, & Computers, 35*(3), 408-419.

Sandler, W., & Lillo-Martin, D. (2006). *Sign language and linguistic universals*. New York: Cambridge University Press.

Schick, B. S. (1990). Classifier predicates in American Sign Language. *International Journal of Sign Linguistics, 1*(1), 15-40.

Shusterman, A., & Peggy, L. (2016). Frames of reference in spatial language acquisition. *Cognitive Psychology, 88*, 115-161. doi:http://dx.doi.org/10.1016/j.cogpsych.2016.06.001

Sidnell, J. (2006). Coordinating gesture, talk, and gaze in reenactments. *Research on Language & Social Interaction, 39*(4), 377-409. doi:10.1207/s15327973rlsi3904_2

Slobin, D. I, Hoiting, N., Kuntze, M., Lindert, R., Weinberg, A., Piers, J., Anthony, M., Biderman, Y., & Thumann, H. (2003). A cognitive/functional perspective on the acquisition of "classifiers". In K Emmorey (Ed.), *Perspectives on classifier constructions in sign language* (pp. 271-296). Mahwah, NJ: Lawrence Erlbaum Associates.

Smith, S., & Cormier, K. (2014). In or out? Spatial scale and enactment in narratives of native and non-native signing children acquiring BSL. *Sign Language Studies, 14*(3), 275-301. doi: 10.1353/sls.2014.0008

Stec, K. (2012). Meaningful shifts: A review of the viewpoint markers in co-speech gesture and sign language. *Gesture, 12*(3), 327-360. doi:10.1075/gest.12.3.03ste

Streeck, J. (1993). Gesture as communication I: Its coordination with gaze and speech. *Communication Monographs, 60*(4), 275-299.

Supalla, T. (1982). *Structure and acquisition of verbs of motion and location in ASL.* (Doctoral dissertation Doctoral dissertation), University of California, San Diego.

Supalla, T. (1986). The classifier system in American Sign Language. In C G Craig (Ed.), *Noun classes and categorization: Proceedings of a Symposium on Categorization and Noun Classification, Eugene, Oregon, October 1983*. Philadelphia: John Benjamins.

Tang, G., & Sze, F. Y. B. (2002). Nominal expressions in Hong Kong Sign Language: Does modality make a difference? In R Meier, K Cormier, & D Quinto-Pozos (Eds.), *Modality and structure in signed and spoken languages* (pp. 296-321). Cambridge, UK: Cambridge University Press.

Tyler, A. (2012). Spatial language, polysemy, and cross-linguistic semantic mismatches: Cognitive linguistics insights into challenges for second language learners. *Spatial Cognition and Computation, 12*(4), 305-335. doi:10.1080/13875868.2012.698670

Tyler, A., & Evans, V. (2004). Applying cognitive linguistics to pedagogical grammar: The case of over. In M Achard & S Niemeier (Eds.), *Cognitive Linguistics, Second Language Acquisition, and Foreign Language Teaching* (pp. 257-280). Berlin/New York: Mouton de Gruyter.