Annotating Real-Space Depiction

Paul Dudis, Kristin Mulrooney, Clifton Langdon, Cecily Whitworth
Science of Learning Center on Visual Language and Visual Learning (VL2)
Dept. of Linguistics, Gallaudet University
800 Florida Avenue, NE
Washington D.C. 20002
E-mail: paul.dudis@gallaudet.edu, kristin.mulrooney@gallaudet.edu,
clifton.langdon@gmail.com, cecily.whitworth.gallaudet.edu

Abstract

"Shifted referential space" (SRS) and "fixed referential space" (FRS) (Morgan 2005) are two major types of referential space known to signed language researchers (see Perniss 2007 for a discussion of alternative labels used in the literature). An example of SRS has the signer's body representing an event participant. An example of FRS involves the use of "classifier predicates" to demonstrate spatial relationships of entities within a situation being described. A number of challenges in signed language text transcriptions identified in Morgan (2005) pertains to the use of SRS and FRS. As suggested in this poster presentation, a step towards resolving some of these challenges involves greater explicitness in the description of the conceptual make-up of SRS and FRS. Such explicitness is possible when more than just the signer's body, hands, and space are considered in the analysis. Dudis (2007) identifies the following as components within Real-Space (Liddell 1995) that are used to depict events, settings and objects: the setting/empty physical space, the signer's vantage point, the subject of conception (or, the self), temporal progression, and the body and its partitionable zones. We considered these components in a project designed to assist video coders to identify and annotate types of depiction in signed language texts. Our preliminary finding is that if we also consider the conceptual compression of space—which results in a diagrammatic space (Emmorey and Falgier 1999)—there are approximately fourteen types of depiction, excluding the more abstract ones, e.g. tokens (Liddell 1995).

Included in this poster presentation is a prototype of a flowchart to be used by video coders as part of depiction identification procedures. This flowchart is intended to reduce the effort of identifying depictions by creating binary (yes or no) decisions for each step of the flowchart. The research team is currently using ELAN (EUDICO Linguistic Annotator, www.lat-mpi.eu/tools/elan/) to code the depictions focusing on the relationship of genre and depiction type by looking at the depictions' length, frequency, and place of occurrence in 4 different genres: narrative of personal experience, academic, poetry, conversation. We also have been mindful that a good transcription system should be accessible in an electronic form and be searchable (Morgan 2005). In tiered transcription systems like ELAN the depiction annotation can simply be a tier of its own when it is not the emphasis of the research, or it can occupy several tiers when it is the forefront. In linear ASCII-style transcriptions the annotation can mark the type and beginning then end of the depiction. Our poster does not bring a complete bank of suggested annotation symbols, but rather the idea that greater explicitness as to the type of depiction in question may be beneficial to corpus work.

1. Introduction

This paper briefly describes a project aimed towards development of procedures to identify and annotate different ways users of any signed language create iconic representations. One main issue in the transcription of British Sign Language narratives identified by Morgan (2005) is the need for an effective way to demonstrate not only the interactions between what he calls Fixed Referential Space (FRS) and Shifted Referential Space (SRS), but also how linguistic items relate to them. We are reasonably certain that many researchers of other signed languages have similar concerns.

Our approach to this issue is based on Dudis' (2007) investigation of dynamic iconic representations, or what he terms **depiction**. We first review how the recognition of additional elements within the signer's conceptualization of her current environment as well as certain cognitive abilities leads to greater precision in describing the various types of depiction produced by signers. We then briefly describe our ongoing attempts to develop depiction identification procedures for purposes of coding and analysis.

2. Types of Depiction

To our knowledge, in their examination of depiction, most signed language researchers do not consider any elements within the signer's conceptualization of the immediate environment other than the signer, the manual articulators, and the surrounding space. Dudis (2007) demonstrates that there are additional Real-Space elements (Liddell 1995) that need to be recognized so as to describe the different ways signers depict things, settings, and events with greater precision. In all there are approximately five Real-Space elements that typically take part in depiction: the setting (or space), the vantage point, temporality, the subject (or the self—note that this does not refer to the clausal subject), and the body. Cognitive abilities also play a role in depiction. The cognitive ability underlying all instances of depiction is conceptual blending (Fauconnier & Turner 2002); see Liddell (2003) for demonstrations on how the conceptual blending model is used to describe "life-sized" blends (surrogates in Liddell's terms), depicting blends, and token blends. Depiction is the result of creating a network of mental spaces, one of which is Real Space. Another mental space in the network is one that has been built as

discourse proceeds, and it contains elements that correspond to Real-Space elements. The blending of these counterpart elements create the iconic representations that are of interest here, and the space in which they exist is called the blend.

Depictions of someone doing any type of activity involve the blending of several elements. The signer has two options here. First, a life-size blend could be created, one in which the Real-Space subject blends with the individual of the event being depicted. Since individuals exist in time and space, relevant counterpart elements also are blended. This type of depiction, which appears to be the SRS described by Morgan (2005), is represented by Figure 1. The box is the |setting|, the shaded figure is the |subject|, and the arrow represents |temporality|.

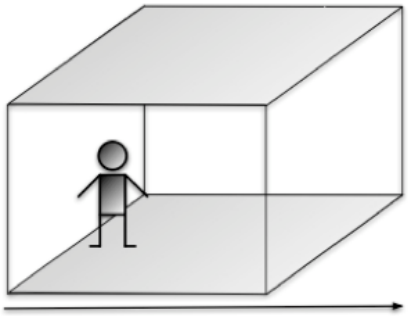

Figure 1

Often it is possible for the signer to choose to create a smaller-scaled depiction of the event. What contributes to this possibility is the cognitive ability to **compress** the setting of the depicted event onto a smaller portion of space, one that is in front of the signer. Since the space that takes part in the depiction does not include the space currently occupied by the signer, she (the Real-Space subject) is not part of the blend. This appears to be the FRS described by Morgan (2005). Figure 2 is a representation of this type of depiction. Since there is no |subject|, the signer is represented as a "regular" figure. The |setting| and |temporality| are represented by a smaller box and arrow. The time of the event being depicted can be compressed into a shorter span of "real time", but so far we see no compelling reason to include this information in our annotations of depiction. Also, we borrow the terms "viewer" and "diagrammatic" from Emmorey and Falgier (1999) to describe the life-sized versus compressed representations.

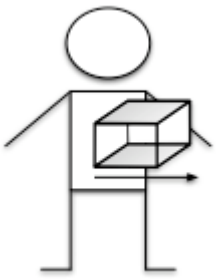

Figure 2

In the figure above we see that it is possible to select some but not all of the Real-Space elements that can take part in depiction. This appears to be what Fauconnier & Turner (2002) call **selective projection**. Dudis (2007) demonstrates that this cognitive ability contributes to the variety of depiction types that can be observed in everyday signed language discourse. As there is a dependency of sorts that certain Real-Space elements have on other elements, there appears to be a limited number of depiction types that signers can produce. For example, the subject must exist within a temporality and a setting, but as we have seen in Figure 2, it is possible to describe an event without creating a |subject| element.

Another cognitive ability that contributes to the variety of depiction types is **body partitioning** (Dudis 2007). The simultaneous activation of SRS and FRS depends on the ability to partition the manual articulators so that they can take part in the creation of representations distinct from the |subject|. We have observed that there are approximately four different types of depiction in which a |subject| is present (six if one wishes to distinguish between constructed dialogue and constructed action that does not involve partitioning; Winston (1991) and Metzger (1995) note that both appear to involve similar strategies). It is possible to depict dialogue and manual action with only the |subject| visible. It is also possible to depict action from the perspective of a patient, e.g. someone being punched, by partitioning off a manual articulator while keeping the |subject| activated; this type of depiction is represented in Figure 3. The manual articulators can also be partitioned off to produce simultaneous perspectives of the event being depicted. This type of depiction (Figure 4) has a participant of the depicted event represented using the Real-Space subject and one (or both) of the manual articulators. This allows the signer to depict, say, someone bumping into someone else by creating a viewer blend to depict specific features of only the patient while simultaneously creating a diagrammatic blend to depict the bump itself. A different type of depiction would be produced if the event is about an experiencer rather than a patient, e.g. someone seeing the bumping. Figure 5 represents this type of depiction. The thought balloon represents the psychological (as opposed to physical) experience one is having. Another example of this type of depiction is the expression of perceived motion (Valli and Lucas 2000). Morgan (2005) describes the possibility of creating "overlapped referential spaces" (p. 125), the co-activation of SRS and FRS. It seems clear that this involves the partitioning of the manual articulators.

Not all events involve an animate participant. It is possible to create a viewer blend to depict unobserved events such as a lightning hitting a tree in a forest. Because there is no animate participant to represent, no |subject| would be activated. Since this is a viewer blend, the location of the signer necessarily participates in the depiction. This location is the Real-Space vantage point. There are many (virtually infinite) locations within the setting of the event from which to view the event. One of these are selected and blended with the Real-Space vantage point, resulting in a blended |vantage point|. Figure 6 represents this type of depiction, with the dotted

figure representing the |vantage point|. We have already described above (Figure 2) another type of event depiction that does not have a |subject|. Because this involves a diagrammatic blend, the Real-Space vantage point is not integrated into the blend. However, this element is of course essential to the creation and development of the diagrammatic blend. After all, it is the limited portion of the space *in front of the signer* where the depiction takes place. We have also considered the ability to conceive of events apart from any specific setting in which they occur. However, as suggested by Langacker (1991), there is a dependency events have: they necessarily take place within a setting. While we were able to come up with expressions in which events are depicted without reference to specific settings, we have not determined whether it was useful to make a distinction between event depictions involving specific settings and those involving schematic settings. Also yet to be determined is the usefulness of identifying event depictions involving the cognitive ability of expansion, as opposed to compression. We can see this in the depiction of events occurring at, say, a subatomic level.

The rest of the types of depiction that we are currently concerned with here are setting depictions. They are non-temporal counterparts to the non-subject event depiction types just mentioned. A viewer blend can be created to depict where objects are located within a setting—say a light fixture in a kitchen. A diagrammatic blend can be created to depict the location of furniture within a room. Features of an object can be depicted apart from a specific setting. For example, the legs of an intricately carved wooden chair can be depicted in front of the signer rather than closer to the floor. Smaller objects can be expanded in size for more efficient depiction.

Classifier predicates (or what Liddell 2003 calls "depicting verbs") are a staple of depictions of objects, settings, and events. A discussion of how they relate to the types of depiction just described is not possible here, but suffice it to say that we view them (or their components) as being types of depiction themselves. For example, a verb that depicts a punch being thrown could be (but not always) considered to be an instance of a depiction involving a |subject|.

3. Depiction Identification and Annotation Procedures

One of our project's aims is to develop depiction identification and annotation procedures to assist video coders in their work. Among the introductory materials currently being developed, we are completing a flowchart of the types of depictions described in Section 2. The flowchart includes yes-no questions that eventually lead to coding instructions. For example, at one point in the flowchart the coder is asked whether there are two distinct visible entities that are life-sized (an example of this depiction is one that describes the event from the patient's viewpoint). If the brief description fits the type of depiction observed, then the coder is shown an illustration similar to those in the above section and is instructed to use a particular code. If the description does not fit, then the coder is instructed to move on to the next description. The flowchart has three major sections: depictions involving |subjects|, event depictions without |subject|, and setting depictions. In all there are between 8 to 14 types of depiction that we are currently interested at this stage of the project.

We use ELAN to annotate depiction observed in video texts. We currently are working with two tiers. One tier will be used to annotate instances of |subject| blends. Different types of |subject| blends will have their own code, and we are also determining a convenient way to identify blends that have been reactivated rather than created anew, as has been observed in narratives where an event is depicted from the viewpoints of multiple event participants. Another tier will be used to annotate instances of event depictions without a |subject| and of setting depictions. There are two reasons for having these two tiers. First, there are types of depictions that appear to be possible only when a |subject| is activated, e.g. those depicting dialogue and perception. The second reason is more well-known and has been documented in Morgan (2005) and elsewhere: signers often "move" between spaces. One of the things that might happen here, as described from a conceptual blending viewpoint, is that the depiction effectively becomes a setting depiction when the signer stops depicting the event to add information via linguistic items, e.g. nouns, that do not depict anything.

Future work will examine other types of depiction, including tokens, depictions that employ metaphor, and tokens, leading towards a more complete typology of depiction. While we begin with the analysis of depiction in simple narratives and related genres, we will eventually work with discourse in other settings. Testing depiction identification procedures in the coding of signed language discourse in academic settings, etc., are likely to reveal issues requiring the revision or refinement of these procedures. We also plan to ensure coder validity of the identification procedures. Ultimately, we hope that these procedures can be used to identify all types of depiction observed to occur any signed language discourse.

4. References

Dudis, Paul. 2007. Types of Depiction in ASL. MS.

Emmorey, Karen and Falgier, Brenda. 1999. Talking about Space with Space: Describing Environments in ASL. In E.A. Winston (ed.), *Story Telling and Conversations: Discourse in Deaf Communities*. Washington, D.C.: Gallaudet University Press, 3-26.

Fauconnier, Gilles & Mark Turner. 2002. *The Way We Think.* NY: Basic Books.

Langacker, Ronald. 1999. *Foundations of Cognitive Grammar, Vol II.: Descriptive Application*. Stanford, CA: Stanford University Press

Liddell, Scott K. 2003. *Grammar, Gesture and Meaning in American Sign Language.* Cambridge: Cambridge University Press

Liddell, Scott K. 1995. Real, surrogate, and token space: Grammatical consequences in ASL. In Karen Emmorey and Judy Reilly (eds.), *Language, Gesture, and Space*.

Hillsdale, NJ: Lawrence Erlbaum Associates, 19-41.
Metzger, Melanie. 1995. Constructed Dialogue and Constructed Action in American Sign Language. In Ceil Lucas (ed.) Sociolinguistics in Deaf Communities. Washington, D.C.: Gallaudet University Press
Morgan, G. (2005). Transcription of child sign language: A focus on narrative Journal of Sign Language and *Linguistics*, 8, 119-130
Valli, Clayton and Ceil Lucas. 1991. *Predicates of Perceived Motion in ASL*. In Fischer, Susan and Patricia Siple (eds.) Theoretical Issues in Sign Language Research. Chicago: University of Chiago Press.
Winston, E. A. (1991). Spatial referencing and cohesion in an American Sign Language text. Sign Language Studies, 1991, 73, Winter, 73(winter), 397-410.