

Cognitive Linguistics Research

Editors

Dirk Geeraerts
Dagmar Divjak
John R. Taylor

Honorary editors

René Dirven
Ronald W. Langacker

Volume 55

Viewpoint and the Fabric of Meaning

Form and Use of Viewpoint Tools
across Languages and Modalities

Edited by
Barbara Dancygier
Wei-lun Lu
Arie Verhagen



DE GRUYTER
MOUTON

ISBN 978-3-11-036907-6
e-ISBN (PDF) 978-3-11-036546-7
e-ISBN (EPUB) 978-3-11-039307-1
ISSN 1861-4132

Library of Congress Cataloging-in-Publication Data

A CIP catalog record for this book has been applied for at the Library of Congress.

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie;
detailed bibliographic data are available on the Internet at <http://dnb.dnb.de>.

© 2016 Walter de Gruyter GmbH, Berlin/Boston
Typesetting: PTP-Berlin, Protago-T_EX-Production GmbH, Berlin
Printing and binding: CPI books GmbH, Leck
♾ Printed on acid-free paper
Printed in Germany

www.degruyter.com



Table of contents

List of contributors — vii

Arie Verhagen

Introduction: On tools for weaving meaning out of viewpoint threads — 1

Part I: The ubiquity of viewpoint

Barbara Dancygier and Lieven Vandelanotte

Discourse viewpoint as network — 13

Jeroen Vanderbiesen

Mixed viewpoints and the quotative-reportive cline in German: Reported speech and reportive evidentiality — 41

Katsunobu Izutsu and Mitsuko Narita Izutsu

Viewpoint fusion for realism enhancement in Ainu and Japanese narratives — 93

Elisabeth Engberg-Pedersen and Ditte Boeg Thomsen

The socio-cognitive foundation of Danish perspective-mixing dialogue particles — 125

Part II: Across languages

Kobie van Krieken, José Sanders, Hans Hoeken

Blended viewpoints, mediated witnesses: A cognitive linguistic approach to news narratives — 145

Wei-lun Lu and Arie Verhagen

Shifting viewpoints: How does that actually work across languages? An exercise in parallel text analysis — 169

Ad Foolen and Toshiko Yamaguchi

Perspective: Kawabata's *Beauty and Sadness* and its translations into English, German, and Dutch — 191

Part III: Across modalities

Chie Fukada

The dynamic interplay between words and pictures in picture storybooks: How visual and verbal information interact and affect the readers' viewpoint and understanding — 217

Eve Sweetser and Kashmiri Stec

Maintaining multiple viewpoints with gaze — 237

Maria Josep Jarque and Esther Pascual

Mixed viewpoints in factual and fictive discourse in Catalan Sign Language narratives — 259

Barbara Dancygier

Concluding remarks: Why viewpoint matters — 281

Index — 289

Eve Sweetser and Kashmira Stec

Maintaining multiple viewpoints with gaze

Abstract: Co-speech bodily gesture has remarkable flexibility in displaying or enacting viewpoint, since – unlike speech but like signed languages – it deploys multiple relatively orthogonal articulators, including head and gaze, two arms and hands, and torso posture. Combined with the viewpoints expressed in the linguistic track, this allows oral narrators to embody viewpoints of two characters at once, or to embody both narratorial viewpoint and an embedded character viewpoint simultaneously. This paper examines video data of semi-spontaneous personal narratives told by speakers of American English. We observe some of the ways in which gaze specifically is used to mark and maintain either the narrator's or some character's viewpoint (including the narrator's Past Self as a story character) even while other articulators may be marking a different viewpoint. These include discourse uses of gaze marking memory access, or "checking" for approval from an interlocutor, as well as content uses such as alternation between enacted characters' gazes. It is always the storyteller's own eyes and face doing the gaze-enaction, but the understood meaning attributes a particular gaze to one of a complex of narrative viewpoints. This is transparent to listeners/viewers because they have access to the complex set of mental spaces evoked, not just to the physical space.

1 Introduction

Unlike the linear sound sequence of spoken language, co-speech gesture and signed languages involve the simultaneous visible use of relatively independent articulators. Two hands, torso posture and orientation, head orientation, facial expression, and gaze are among the articulators which are regularly relevant in linguistic communication – especially for indicating viewpoint shift (see Parrill 2012 and Stec 2012 for reviews, and McClave 2000 for an indication of the multifunctionality of head movements in discourse). Although manual gesture is the most studied of these, and indeed gesture is often used specifically to mean manual gesture, we will be using a more inclusive definition of 'gesture' which includes all multimodal articulators. These multimodal articulators are both externally observable and centrally involved in everyday action and attention in the world, not just in communication. Looking at someone's bodily action and motion tells you what they are doing and attending to. And gaze, the topic of

this paper, plays a unique role in discerning communicative action, attention and intention (e.g. Rossano 2012, Sidnell 2006, Schegloff 1998).

Consider the case of a person working at her computer, who turns her head and gaze leftwards to address a colleague entering her office – but keeps her body facing the computer, with her hands held above the keyboard, as she speaks with the visitor. She is providing clear evidence to the visitor both that she is attending to him, and that she is embedding this temporary attention in a longer stretch of ongoing work which will be resumed. The effect would be extremely different if she instead also turned her whole body and gaze towards the visitor as she addressed him, removing her hands from the desk. And it would be different yet again if, for example, she kept her head, gaze and body turned towards a skype session on the computer, but perhaps silently stretched out a hand towards the (peripherally visible) visitor – to hand him a document she knows he wants, or instead greeted him with a palm held up to request that he wait. Perhaps most interestingly, even if she turned her body and face and arms towards the visitor, if her gaze repeatedly strayed towards the computer screen (or towards her watch), a sighted visitor would still quite strongly feel the effect of divided attention. This has been called *body partitioning* (Dudis 2004) or *body torque* (Schegloff 1998), and it shows not only what activities a person is currently engaged in (or currently engaged in representing, cf. the dual viewpoint gestures discussed by Parrill 2009), but also how ongoing discourse is re-structured as a result of such partitioning.

Of course, in a speaker/gesturer, the speech track is the dominant informational channel – which is why, unlike a signed language, co-speech gesture is not generally interpretable on its own. But in oral narrative, gesture provides very complex support for the viewpoint structure of the spoken content, in particular where a narrator is voicing and embodying multiple characters, while also maintaining interaction with her real-world interlocutor. This work is an initial attempt to pin down the role of gaze in this very multi-track multimodal set of affordances for an oral narrator. Our guiding questions are: How does gaze contribute to meaning differently from hands, body posture, position, and head movement? And what can this tell us about viewpoint in oral narrative?

We might think of narrative embedding – the embedded expression of a character's utterances and thoughts in a narrator's expression – as being a kind of combining and embedding of activities, somewhat like that of the person being interrupted by the visitor to her office. But the activities involved in narrative include representation of multiple – embedded, or contrasting – viewpoints. Instead of attention-sharing between the interactions with the computer and with the visitor, the writer of a narrative is dividing linguistic resources between author-reader interaction, narrator-reader interaction, and of course the content

of the story space, including different characters' viewpoints embedded in the narrator's viewpoint. It is fairly well understood by narratologists that combining linguistic forms appropriate to the narrator (e.g. 3rd person and past tense, since the narrator is recounting past events about other people) with forms appropriate to a character's point of view (descriptions such as *that rat* or *Daddy*; deictic forms like *now* and *here*) will result in a tight embedded viewpoint which is sometimes called Free Indirect Style or Free Indirect Speech and Thought (Genette 1980); other combinations of forms produce different narrative effects (Vandelanotte 2009, 2012; Dancygier 2012).

The oral narrator, however, is dividing not just her linguistic resources, but also her gesture space, her movements and actions, and her attention resources, between representing the content of the Story Space, including particular character viewpoints, and the real world interaction with her interlocutor. One way of talking about this is to talk about representation of different mental spaces, using the framework introduced by Gilles Fauconnier (Fauconnier 1994[1995]; Fauconnier and Turner 2002). Following other work in this framework on both gesture and signed language (e.g. Liddell 2003; Dudis 2004), we will refer to the shared physical space of the actual Speaker and Hearer as the Real Space. Gesture space is one aspect of Real Space.

Parrill and Sweetser (2004) noted that in gesture, as well as in signed languages (Liddell 2003; Dudis 2004) it is normal for sub-areas of physical gesture space to be devoted to different topics or characters. And Guntner et al. (2015) have used experimental paradigms with EEG to demonstrate that interlocutors not only recognise these distinct uses of gesture space, but expect them to be there after only minimal prompting. Sweetser and Sizemore (2008) and Stec and Sweetser (2013) have noted the functional division of gesture space, with the Speaker-Hearer line as the locus of Real Space interactional gestures, while a space to one side of that line would normally be allocated to the narrative.

Sweetser (2013) and Stec and Sweetser (2013) have argued that viewpoint-alternation and viewpoint-embedding can be achieved in extremely complex ways, since part of the narrator's body may be engaged in the part of the gesture space devoted to the narrative, while other simultaneous aspects of gesture may take place along the Speaker-Hearer line and engage in interaction in the Real Space.

Looking at gaze, Thompson and Suzuki (2014) have used the relationship between original and reenacted events to argue that gaze is an important means by which speakers jointly create reenactments with their listeners and differentiate events which took place in the here-and-now (or Base Space, to use our terminology) vs. in the reenactment (or Story Space), specifically to manage transitions between narrative proper and reenactment. Like Sidnell (2006), they found that

gaze is moved away from interlocutors during reenacted sequences, and returned to interlocutors at the end of such sequences. Moreover, they found that the direction of the averted gaze depends on the kind of reenactment, with gaze resting in central, neutral spaces during phone call or text reenactments, or more peripherally for face-to-face reenactments. At the same time, however, Holler et al. (2014) used an experimental task to demonstrate that listeners are sensitive to the direction of speaker-gaze: Listeners who were gazed at process speech better than listeners who are not gazed at. This suggests that there are comprehension costs inherent to speaker shifts in gaze. Nonetheless, as Thompson and Suzuki (2014), Sidnell (2006), and Park (2009) suggest – and as we show below – speakers do systematically avert and return gaze in ways which are related to the partitioning of Real Space, and to managing ongoing discourse.

Anchoring the body in one Real Space allows for easier transitions to other Real Spaces and therefore mental space. As we will show, gaze is one of the means by which this is accomplished. Of course, there are others. Sweetser (2014) argues that the orthogonality of gestural articulations allows the simultaneous physical maintenance of elements from different mental spaces, thus physically representing space embedding structures. As mentioned above, narratologists have long noted that the effect of combining linguistic markers appropriate to a higher space and an embedded space can give particular narrative effects (Genette 1980; Banfield 1982; Fleischman 1990; Sanders and Redeker 1996; Dancygier 2012; Vandelanotte 2009, 2012; Nikiforidou 2010, 2012). More specifically, Free Indirect Speech and Thought is conventionally represented by past tense and third-person reference to story characters, combined with character-based deictic forms (*here, now*) and character-based descriptions. Thus, a sentence like *She knew that by now he would be telling Daddy all about it* shows the Narrator's past tense and the Narrator's *she* to refer to the viewpoint character; but *Daddy* and *now* refer to the viewpoint character's father and temporal present, not the Narrator's. The result is an experience of active viewpoint embedding: the narrator's viewpoint is not abandoned, but rather maintained as constant background to the foregrounded viewpoint-character's space. Sweetser noted cases where a gesturer simultaneously maintains gestural depiction of a story character's body, and interacts (often with face/gaze as well as voice) with the real-world interlocutor, somewhat like the way Thompson and Suzuki (2014) and Park (2009) note that speakers can treat their interlocutors as Story Space characters for the purposes of reenactment. Both narratorial body and character body are enacted simultaneously, and the result is a vivid embedding of one embodied viewpoint within another.

Now that we have given an overview of the different issues involved in multimodal viewpoint embeddings, we will give an overview of the rest of this paper.

In Section 2, we briefly describe our data and the narratives used in our analysis. Following that, in Section 3, we discuss how gesture space is partitioned in our narratives. And then we discuss particular uses of speaker gaze to manage viewpoint embeddings within ongoing discourse: Character enactment (Section 4), Narratorial gaze (Section 5), Visual “checking” (Section 6), and to access memory (Section 7). Throughout, we will consider how gaze works to manage discourse functions whether in coordination with or separately from other multimodal articulators, and how it contributes to understanding oral narratives more broadly.

2 Data for the present study

Our data come from a video corpus of semi-spontaneous speech which was collected by the second author on the West Coast of the US in January 2012. All participants were native speakers of American English who brought a friend to the recording location, where they told and requested personal narratives in a relatively naturalistic way. Stec et al. (submitted) describes our recording procedure. Our corpus is comprised of 26 speakers (17F, 9M) and a total usable corpus length of approximately five hours. There are 85 narratives in the entire corpus – we only discuss three here, but the behaviours we note are present throughout the corpus.

In Stec and Sweetser (2012), we offered a quantitative analysis of the spatial distributions of gestures in these narratives, and correlations between different aspects of gesture, gesture space and mental spaces. In this paper, we will be focusing instead on qualitative analysis of a few complex clips, all of which involve the simultaneous representation of more than one viewpoint. We will be deepening the tentative analysis offered in Sweetser (2014) of partitioned gestural viewpoint as a representation of narrative embedding, but we will focus specifically on the role of gaze in such partitioned structures.

For this analysis, we will focus on three narratives from the corpus (about which more below): License, Snow, and Cats. For each narrative, we will identify participants by the colour of the shirt they wear. A representative screenshot from each narrative, with speakers circled, is provided in Figure 1. For relevant examples, we will provide a transcript and a figure composed of stills from the video-recording.



Figure 1: A representative still from each narrative, with primary narrators circled. License is shown in panel a, with Green identified on the left of the frame. Snow in panel b, with Grey identified on the right of the frame. And Cats in panel c, with Black identified on the right of the frame.

But first, a brief summary of the three narratives: In License, the narrator, Green, located on the left of the frame, tells the story of going to a town hall of a small town with another woman and requesting a marriage license from an official there. As she tells the story, she enacts her own Past Self and the Official, sometimes simultaneously – and she also interacts directly with her Real Space interlocutor. In Snow, Red on the left and Grey on the right are a married couple, and Grey is recounting the story of a long-past difficult drive through a blizzard which they made together, in more or less the same relative spatial configuration of the original event (since Red is seated to Grey's right in Real Space, and Grey was the driver in the Story Space). Grey is simultaneously reenacting both of their past selves, and also sharing enjoyment of this favourite story in the present with Red. In Cats, the narrator Black, located on the right of the frame, describes how her boyfriend likes to put a laundry basket over her cat to see what will happen. She enacts her Past Self, the Boyfriend, and even the Cat in this narrative, while also intermittently sharing appreciation of these events with her real interlocutor Pink.

As we shall see, in all narratives gaze plays a crucial role in allowing the narrator to simultaneously engage in more than one level of meaning building. License involves lengthy manual gesture holds and the use of gaze to facilitate transitions between the Story Space and Real Space. The rich co-narration involved in Snow, as the couple re-live their past experience, offers a great deal of gaze use as a way to maintain that shared experience in the present, even while one of them narrates the past.¹ Cats involves two very interactive friends, and offers special evidence of other ways in which meta-interaction takes place alongside narrative content. They also offer unusually vivid character-viewpoint depictions, with both positional alterations and partitioning of the body to depict multiple characters.

¹ For more information about the co-construction of gestural space in narratives, see Stec and Huiskes (2014).

3 Partitioning of the gestural body, and of gestural space

In one sense, gesture is simple: a human body moves in Real Space. But why does it move the way it does, and how does anyone interpret co-speech bodily activity as contributing to communicative meaning? The answer is both simple and complicated. First, much of gesture is interpreted as building content relative to mental spaces other than the current Base Space, the Speaker and Hearer's overall real world of which the Real Space is a part. The interpretive trick, then, is to figure out to which space a gesture is contributing meaning. This is, of course, in a way no different than the problem of interpreting language; someone listening to a story in the past tense has a good clue from the tense that the events described are not about the here-and-now Base Space, but about a Story Space whose content is past relative to the Base Space. In the case of gesture, things are a bit more complicated. Someone telling a past-tense narrative could gesture about running, as she described past running events; she could also put up a hand to stop interruption from the Real Space interlocutor, without any linguistic markers of doing so from the Story Space.

However, spatial partitioning gives crucial and complex clues. As is normal in co-speech gesture, our narrators regularly used the Real Space area directly between the Speaker and Hearer (the Speaker-Hearer line) as the locus for real-world interaction between them. They normally picked an area in physical space to one side of that line as the one used to represent the mental space of the narrative content (the Story Space). This partitioning of space is illustrated in Figure 2, where each dyad's use of gestural space in these three narratives is depicted.

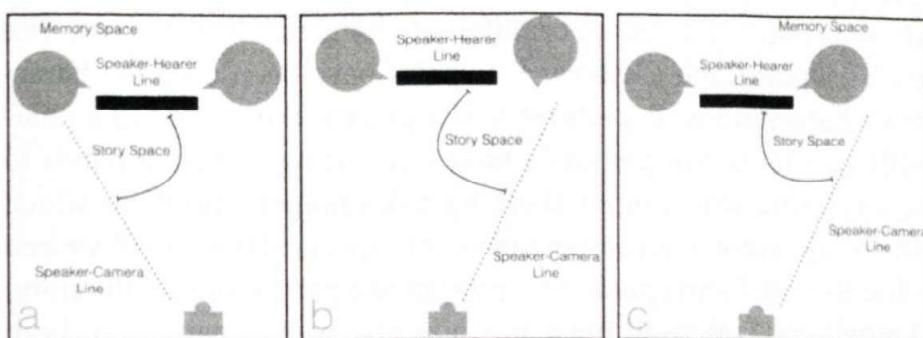


Figure 2: The partitioning of gestural space used by each dyad in the three narratives considered here. License appears in panel a, Snow in panel b and Cats in panel c. In each case, both interlocutors and the camera are identified, as well as the Speaker-Hearer line and the location of the narrator's story space. In addition, the Memory Spaces used by Green (panel a) and Black (panel c) are identified.

Gestures depicting events in the past tense Story Space occur in the physical part of the gesture space allotted to the story, while gestures towards the real-world interlocutor occur instead along the Real Space Speaker-Hearer line. Very often, given the physical setup, speakers chose to place the Story Space on the camera side of the Speaker-Hearer line, though not directly facing the camera. This had the practical advantage of not letting gestures be obstructed by the physical surroundings of the recording locations – as well as of treating the filming observer as a third participant. That is, if there had been a third person actively taking part in the conversation (as in, e.g., Özyürek 2002), the narrator would certainly not have chosen to locate the Story Space either right on that Speaker-Hearer line, or completely outside the shared three-way interactional zone, but rather in some neutral and accessible space in between. Because of this, the Speaker-Camera line was avoided as a primary Story Space locale. Özyürek (2002) has shown that narrators prefer to avoid the Speaker-Hearer line in miming at least some kinds of story event content. Özyürek asked participants to watch Looney Tunes cartoons and re-tell them to two naïve listeners. Placing her participants in a triangular array with the two listeners, she found that some subjects enacted event descriptions like *She threw the cat out the window* with a backwards over-the-shoulder throwing motion. This was a poor representation of the content being expressed, since in the video being re-narrated, the character threw the cat forwards out the window. However, the backwards-gesturing speaker thus avoided *throwing the cat into the personal gesture spaces of her hearers or onto the shared Speaker-Hearer communicative line*. Perhaps our storytellers, therefore, were avoiding both the communicative line between themselves and the filmed auditor, and also the “communicative line” between themselves and the camera.

Because gesture uses so many more independent articulators than speech, it is also less sequential: different gestural articulators of a narrator can simultaneously enact different parts of a scene, or different characters in the scene. The terms Character Viewpoint and Observer Viewpoint have been used to characterise the contrast between iconic gestures that represent the actions of a character's body with actions of the gesturer's body (e.g., using gesturing hands to mime grasping a steering wheel or climbing up a drainpipe), and those which use the gesturing body rather to represent objects or aspects of the overall viewed scene (e.g. two hands with facing palms to represent two participants in the story) (Parrill 2009; Liddell 2003; McNeill 1992). It is now also well recognized by both signed language and gesture analysts that body partitioning (Dudis 2004a, b) is an ordinary feature of communicative bodily action. For example, as we will discuss below, a speaker's gaze, head direction and facial expression may represent those of one narrative character, while that speaker's hands and body represent those of another character. Supposing these enactions to be accompanied

by utterances attributed to the two narrative characters, and to be carried out in the physical gesture area of the Story Space, neither the narrator nor the listener will have any trouble in appropriately interpreting them as representing the emotions and actions of the story character, rather than the emotions and actions of the speaker herself.

4 Character enactment

A narrator's gaze can readily represent the gaze of a character rather than that of the actual speaker, when it is in the Story Space rather than directed at the Real Space listener. This is equally true in ASL, as Liddell (1998) showed in his analysis of a signed recounting of a Garfield cartoon, where the signer's gaze alternately enacts the cat's upwards gaze towards the owner, and the hapless cat owner's downwards gaze at the cat. This phenomenon in co-speech gesture is well exemplified by the sequence in License where Green describes a past interaction with a town-hall bureaucrat. This is shown in Transcript 1 and Figure 3. In the story, she tells us that her birth certificate erroneously says "male" with a stapled-on correction saying "female" – this of course confuses the Official. The narrator's gaze alternately represents the Official's sceptical gaze (going back and forth between an imagined birth certificate, and the narrator's Past Self as they discuss it), and Green's Past Self's gaze (directed towards the official) and facial expression. Throughout the sequence, the narrator's hands represent the Official's hands, holding the document – even when her face and gaze are representing her Past Self, who is clearly not holding the document. In these cases, we have two distinct Character viewpoints, simultaneously represented by the narrator's partitioned body.

As can be seen in the stills in Figure 3, Green is using the space in front of her and to her right to enact the story; her Real Space interlocutor is to her left, away from the camera, and she turns head and gaze towards him for Real-Space interaction. Within the Story space in front of her, she alternately enacts the two story characters (Past Self and the Official).

Transcript 1: License, excerpt 1

- 1 Green: and finally someone turned over the stapled part of my birth certificate
- 2 and saw that it said female
- 3 and they were like nope that's why you can't



Figure 3: Stills from License. Images 1–2 correspond to line 1 of the transcript (Green enacts the Official manipulating her birth certificate). In image 2, Green underlines the part of the transcript that says “female” and in image 3, she shows the triumphant Official returning the birth certificate to her Past Self. Apart from looking at the document in images 1–2, Green looks at her interlocutor, Black, throughout this sequence.

Green’s interlocutor, Black, clearly has no trouble understanding Green’s partitioned representation as showing the two characters, and viewers to whom we have shown this material have also effortlessly reached the same interpretation. This interpretation depends first of all on the pre-constructed relationship between physical space and mental spaces; without the pre-establishment of a physical gestural subspace (to the right of the Speaker-Hearer line) devoted to the Story Space, these gazes would be not only meaningless but socially problematic as they might indicate that the speaker was not fully attending to the current speech interaction. After all, none of them are directed towards the real-world interlocutor, Black. And of course without the quoted speech sequence of the two characters within the Story Space, the actual physical gazes would mean nothing, since they are hard to interpret as the speaker’s current gaze.

Fascinatingly, note that once a narrator is “inside” the Story Space and enacting interaction between characters, a new Speaker-Hearer line develops inside the Story Space. That is, the Past Self and the Official, both enacted in turn by the narrator of License, are imagined to be at opposite ends of this communicative axis, and we mentally rotate/displace Green’s representations, blending them alternately with the Official’s end of the axis and the Past Self’s end. This looks very much like the kind of 180-degree virtual “rotation” between interlocutor-viewpoints described by Janzen (2013) in ASL storytelling.

Another example of this comes from Cats, and is shown in Transcript 2 and Figure 4. Here, the narrator Black, on the right of the frame, alternately enacts her past self and her boyfriend’s past self in narrated speech exchanges. Her head and gaze go upwards and to her right as she enacts her shorter self talking to the taller boyfriend, and downwards and leftwards as she enacts the boyfriend. Similarly, her gaze is downwards as she enacts both Past Self and Past Boyfriend looking downwards at the cat. This is shown in Figure 4.

Transcript 2: Cats, excerpt 1

- 1 Black: but she's not
- 2 sssssh
- 3 ok but she's not
- 4 sssssh (0.5)
- 5 ok but she's not doing anything
- 6 not (0.5) yet



Figure 4: Stills from Cats. Image 1 corresponds with line 1 in the transcript, image 2 with line 2, and so on. Throughout the series, Black is oriented towards her Story Space and makes a left/right and up/down distinction with her head movements and gaze to differentiate quotes made by her Past Self and her boyfriend in the story.

Again we must alternately blend the narrator's bodily behaviours as well as her speech track with the two interlocutors, who are of course not construed as being in the same location – despite the fact that the speaker does not shift her overall seated location. This looks very much like the most-described kind of signed-language role shift (e.g. Quinto-Pozos 2007), which involves actual partial body rotation (not imagined 180-degree rotation) when enacting a quoted character, and **partial rotation in the opposite direction** to enact that character's interlocutor, as well as character-viewpointed **signing and averted gaze**.

And, as we've noted, more than one character can be represented simultaneously by the same body – in License, Green keeps her hands in a document-holding position representing the Official holding the Document, even as she switches her gaze from representing that of the Official to that of Past Self, and even (as we shall see below) when her head and gaze revert momentarily to her present self in real-world interaction. (This is also true in signed language; cf.

Dudis 2004a,b and Dudis 2007.) In Cats, the Boyfriend's physical features, such as his height and position relative to Black's Past Self, are retained during enactment of Past Self's utterance and gaze, as well as during Black's performance of her Boyfriend's utterances and gaze.

We note that facial expression and gaze tend to go together here. License regularly involves representing character facial expression together with gaze – the Official examining the document with a dubious expression, the Past Self looking at and responding to the Official with a cheery, hopeful expression even while the speaker's hands remain those of the Official holding the Document. Similarly, in Cats, the alternation between representations of the two characters' speech coincides with alternations between enacting their gazes and their facial expressions simultaneously. This is something to examine closely in future work. Although gaze and facial expression are certainly separable in principle, in practice they seem tied – perhaps more strongly than gaze is tied to head direction, even. Or, as one reviewer put it: Body partitioning is fine – and even common (see, e.g., Schegloff 1998, Dudis 2004, or Parrill 2009) – but face partitioning is not.

5 Narratorial gaze as itself

Of course, the Real Space of actual interaction is always present in oral narration. This is why listeners don't feel offended by divided attention when the narrator's body and gaze vividly represent characters – they know that all of this is in the service of a meta level of attention to the narrative, which constitutes the current phase of the narrator's Real Space interaction with the listener. As evidence of this, Sidnell (2006) and Thompson and Suzuki (2014) document the extent to which reenactments (by which is generally meant, quotative utterances) are preceded and followed by gazes directed to interlocutors, but the reenactments themselves are produced with averted gaze. This behaviour is so entrenched that, as Thompson and Suzuki show, speakers purposefully seek their interlocutors' gaze to make sure proper attention is paid to the reenacted sequences. Normally, shared gaze is used to indicate shared attention (cf. the joint gaze discussion in Tomasello 2000) – so it is in some sense bizarre that joint gaze is sought specifically to enable averted gaze. The fact that it happens during these reenacted sequences with quoted utterances (and therefore, explicit conceptual viewpoint shifts, cf. Parrill 2012) makes some sense insofar as the pattern bears resemblance to standard role shift sequences in sign (e.g. Quinto-Pozos 2007) and enables the iconic representation of another viewpointed body (Sweetser 2012). But what about elsewhere?

The narrator's gaze can of course always remain (or revert to being) just the actual speaker's gaze within the current interaction – even while other parts of the speaker's body are not representing her real-world self. Often, but not always, this use of gaze is packaged with head movement, while the hands and body are the ongoing placeholder (or buoy, in the terminology of Liddell 2003) maintaining the embedded space, including embedded characters. This is how we would characterise the sequence in License in Transcript 1 and Figure 3, above. There, the narrator's hands continue to represent the Official's document-holding activity in the Story Space area, even when the narrator's face and gaze are turned away from the Story Space towards the Real Space interlocutor to answer an interpolated question. Just like the person at the desk whose body and hands remain directed towards her computer while she turns her head to respond to a visitor in her office, this narrator is showing clearly that she is still telling her story and will return to it after the meta-interaction is over.

Transcript 3: License, excerpt 2

- 1 Green: and so she scoots her chair back from her desk
- 2 like really abruptly
- 3 and just sits there for a second
- 4 like wondering what she's going to say



Figure 5: Stills from License. Green shows how the Official scooted their chair back (lines 2–3 in the transcript) and then sat there, wondering what to say (lines 3–4). Green gazes at her interlocutor throughout this sequence.

License is another good example – Green is enacting the described official's pushing back of her chair, while her gaze and head direction maintain direct connection between herself as speaker and the real-world addressee. This is shown in Transcript 3 and Figure 5. Similarly, later in the narrative, her hand is enacting the described character's telephoning a colleague, while gaze and head direction are Base Space interaction. This is shown in Transcript 4 and Figure 6. And in Cats, Black's hands hover in the Story Space area, prepared for narrative content

expression, while gaze and head direction are already directed towards the Story Space. This is shown in Transcript 5 and Figure 7. On line 2/image 2, Black's head moves towards Pink along the Real Space Speaker-Hearer interaction line, as the two actual interactants joke before the story starts – but her gaze and face maintain orientation towards her Story Space. This is shown in Figure 7. At the end of such a “meta” stretch, a narrator returning to narrative content often also linguistically marks the return with a linguistic resumptive marker (*anyway, OK, so*) as gaze and head direction return to the narrative space.

Transcript 4: License, excerpt 3

- 1 Green: so we went over there and in the meantime
- 2 she had phoned over to the
- 3 I think it was to the magistrate of the court



Figure 6: Stills from License. Green's head and gaze are oriented to her interlocutor while her hands produce character-viewpoint gestures which elaborate the Story Space.

Transcript 5: Cats, excerpt 2

- 1 Black: go to bed
- 2 eat your tea
- 3 Pink: eat your tea
- 4 Black: so (0.5)
- 5 so my sweetie and I

A point to note here is that we have observed an extremely tight link between gaze and facial expression; we haven't seen examples where a narrator's gaze represents one viewpoint entity, while her facial expression represents another character or entity. Gaze can be detached from direction of head/face, but not from facial emotion expression – as in License, where Green's gaze and facial expression shift together between representing the Official, the Past Self, and the present narrator, even while her hands and body direction are partitioned and may be representing another character or space. As we saw earlier: a body can be partitioned in multiple ways, but not a face.



Figure 7: Stills from Cats. Pink and Black joke before Black starts her narrative (line 1–3 in the transcript, images 1–2 in the figure). When Black says line 2, her head and gaze are already oriented towards the Story Space, and hands are already raised (image 3 in the figure). Black reclaims the floor in line 4, and in line 5 (image 4) manages to start her narrative. During the entire exchange, her gaze is oriented away from Pink and towards her Story Space.

And, as we said above, a vivid form of viewpoint embedding is achieved by this combination of bodily enactment of Character with Real-Space narratorial gaze and speech. It would be hard to argue that this is less complex than, say, Free Indirect Speech and Thought representation in language – it is certainly very rich, and we cannot even claim to have isolated all of the relevant parameters involved. It is certainly like Free Indirect Speech and Thought in that it is easy to process these multiple viewpoints simultaneously – interlocutors are not confused by the complexity. And gaze is a crucial component.

6 Visual “checking”

We mentioned above that at any point a narrator can choose to go “meta” and return from the narrative performance in the Story Space to the Real Space interaction. If she needs to address the listener in the midst of the narrative, she will almost certainly turn her head and gaze towards the listener during that time, even while perhaps maintaining hands and trunk in the physical area allotted to the Story Space. This does at least somewhat break the embedded storyline. However, we also observed instances of very brief gaze/head re-direction towards the addressee, while the verbal content and the rest of the gestural body remained related to the embedded Story Space. This was used as an interactional device of “checking” with the listener without interrupting the embedded verbal flow of narrative. Since no real interruption is involved, no linguistic resumptive devices such as *anyway* are needed either.

As we mentioned earlier, in Snow we see Grey telling the story of a past drive which he and Red, the current Real Space interlocutor, took together years before.

As Grey enacts both sides of a past conversation during the drive, he very briefly turns his head and meets Red's eyes, both right before and right after quoting her past utterance (lines 2 and 4 in the transcript). This is shown in Figure 8, images 2 and 4. At the second head-turn, he actually "checks" linguistically as well (line 4), but the first one has no linguistic concomitant.²

Transcript 7: Snow, excerpt 1

- 1 Grey: oh let's turn around
- 2 and [Red]'s going
- 3 no no no just go just go
- 4 remember



Figure 8: Stills from Snow. Grey turns his head and meets Red's gaze both right before (image 2) and right after (image 4) quoting her.

The closest parallel we can think of to this is back-channeling (or Japanese *aizuchi*), where small utterances (*yeah*, *OK*, *right*, *no way*, *wow*) or gestures (e.g., nods and headshakes) help to manage the listener's side of the interaction and maintain interparticipant connection, while not claiming the floor or interrupting the speaker. For example, during the narrative about their drive through the blizzard, Grey very briefly turns his head and meets Red's eyes, not claiming the floor or interrupting her interjection. In a similar way, one might say that this visual "checking" avoids self-interruption by avoiding actively returning the communicative floor to the Real Space interaction. And it allows both the overall linguistic track and the overall gestural track (with this momentary divergence) to stay in the narrative space, while still maintaining Real Space Speaker-Hearer connection.

² Park (2009) discusses quoting a co-present character, such as Red, in naturalistic Korean discourse. There, he notes that the co-present character withdraws their gaze from the speaker during the quoted utterance. We don't see that kind of behaviour with our English speakers.

7 Memory spaces

Averted gaze can often represent access to a Memory Space. McCarthy et al. (2008) demonstrated that the direction varies cross-culturally, with Canadian speakers preferring to look upwards in communicative contexts and downwards otherwise while Japanese speakers generally prefer to look downwards. And Glenberg et al. (1998) demonstrated that averted gaze improves cognitive processing in non-communicative situations, such as problem-solving. But as these and Holler et al. (2014)'s finding suggest, gaze is tricky – and not least of all for the huge individual variation in gaze patterns and use. For example, unlike McCarthy et al.'s Canadian speakers who varied gaze aversion patterns depending on context, the American speakers in our corpus – all of whom were in communicative contexts – averted gaze by looking either upwards or downwards. More important than the choice of direction was the fact of the action itself, which demonstrates “thinking” to the addressee and thus also holds the floor for the speaker. At the same time, it partitions gesture space – “thinking” doesn't happen on the Speaker-Hearer line or in the Story Space, but elsewhere.

For example, in Cats, Black glances up as she is remembering, before telling the story. This is shown in Transcript 6 and Figure 9. Often access to the Memory Space involves little head or no movement – it is specifically the eyes which move.

Transcript 6: Cats, excerpt 3

- 1 Pink: can I hear the cat and the laundry basket story
- 2 Black: yes yes you can



Figure 9: Stills from Cats. Black, the narrator, looks towards Pink as she requests the story in image 1 and then looks up as she prepares to tell the story in image 2.

These **memory** gazes bring to the fore an issue which pervades gesture studies overall: **how** intentional and how conscious is meaningful gestural behaviour? Although all of the behaviours we've been talking about are potentially observable, it's **not** clear exactly how conscious speakers or listeners are about them. It's true that a listener might subsequently say, if questioned, that the narrator was being **vivid**, or "acting out" past characters and scenes. But would they be able to bring **to** consciousness facts like the allotment of physical gesture space to different **functions** (narrative content, meta-interaction)? In the case of the upwards gaze for **gathering** memories together, however, we seem to be seeing a communicative **discourse** signal to the listener that the speaker is "searching" or "consulting" **her** memories, rather than (for example) just being silent. This is interestingly parallel to hesitation markers (see James 1973, Clark 1996 ch. 9) – while *um* or *uh* seem to simply indicate floor-holding during linguistic formulation, *oh* (as in, *There were, oh, about thirty people at the party*) specifically seems to mark memory **searching** – and hence to mark evidentially the fact that the speaker has personal **memories** to search.

8 Conclusions

We hope to have laid out in this paper at least some of the ways in which gaze contributes to the maintenance of multiple viewpoints in the gestural structure of our English oral narratives. Gaze in narrative would be uninterpretable without mental **space** mappings: whose gaze is understood as being involved or represented? An essential component in such interpretation is the partitioning of the gesture **space**: if the gaze is directed into the Story Space, we know it is not the Real Space narrator's gaze that is meant (since the Story Space sector of Real Space does not have real addressees in it). And the partitioning of the body itself is equally crucial: is the gaze to be interpreted as belonging to the same viewpoint as other gestural components?

Like anything about gesture, gaze is meaningful primarily because of how the mental spaces are constructed relative to the Real Space. Some of this is really general – for example, the Speaker-Hearer line of the Real Space or of an embedded discourse space, or the "up-there-ness" of memory. And some of it is very locally built – we always need to know where the speaker "put" the Story Space (in all our examples), and may need to know things like the relative heights of characters (e.g. the upwards and downwards head turns in Cats), as well as what they are saying/doing within the narrative, etc.

As we suggested in Stec and Sweetser (2013), particular aspects of gesture are not all equally orthogonal to each other. Hands and body positioning often go together. And as we have noticed in this paper, gaze and facial expression seem particularly tightly tied to each other. This is clearly just a beginning, since the richness of viewpoint structures in even these narratives has not been fully analysed – but we hope it is a helpful beginning as we start to look at the role of gaze in multimodal narrative viewpoint.

Acknowledgments

We thank members of the Gesture and Multi-Modality Group at UC Berkeley and participants at ISGS 2012, ICLC 2013 and ISGS 2014 for helpful comments made on earlier stages of this work. Kashmire Stec's contribution was funded by grant number 276.70.019, which was awarded to Dr. Esther Pascual by the Netherlands Organization for Scientific Research (NWO).

References

- Banfield, Ann. 1982. *Unspeakable Sentences: Narration and representation in the language of fiction*. Boston: Routledge and Kegan Paul.
- Clark, Herbert. 1996. *Using language*. Cambridge: Cambridge University Press.
- Dancygier, Barbara. 2012. *The language of stories: A cognitive approach*. Cambridge University Press.
- Dudis, Paul. 2004a. *Depiction of events in ASL: Conceptual integration of temporal components*. University of California, Berkeley dissertation.
- Dudis, Paul. 2004b. Body partitioning and real-space blends. *Cognitive Linguistics* 15(2), 223.
- Dudis, Paul. 2007. Types of depiction in ASL.Ms., Gallaudet University.
- Fauconnier, Gilles. 1994[1985]. *Mental Spaces*. Cambridge: Cambridge University Press.
- Fauconnier, Gilles & Mark Turner. 2002. *The way we think: Conceptual blending and the mind's hidden complexities*. New York: Basic Books.
- Fleischman, Suzanne. 1990. *Tense and narrativity: From medieval performance to modern fiction*. Austin: University of Texas Press.
- Genette, Jean. 1980. *Narrative discourse: An essay in method*. Ithaca: Cornell University Press.
- Glenberg, Arthur, M., Schroeder, Jonathan. L., & David A. Robertson. 1998. Averting the gaze disengages the environment and facilitates remembering. *Memory & Cognition* 26, 651–658.
- Gunter, Thomas C., J. E. Douglas Weinbrenner & Henning Holle. 2015. Inconsistent use of gesture space during abstract pointing impairs language comprehension. *Frontiers in Psychology* 6(80). doi: 10.3389/fpsyg.2015.00080

- James, Deborah. 1973. Another look at, say, some grammatical constraints on, oh, interjections and hesitations. *Papers from the Ninth Regional Meeting of the Chicago Linguistic Society*. 242–251. Chicago Linguistic Society.
- Holler, Judith, Louise Schubotz, Spencer Kelly, Peter Hagoort, Manuela Schuetze, & Asli Özyürek. 2014. Social eye gaze modulates processing of speech and co-speech gesture. *Cognition* 133(3). 692–697.
- Janzen, Terry. 2012. Two ways of conceptualizing space: Motivating the use of static and rotated Vantage point space in ASL. In Barbara Dancygier & Eve Sweetser (eds.), *Viewpoint in language: A multimodal perspective*, 156–176. Cambridge: Cambridge University Press.
- Liddell, Scott. 1998. Grounded blends, gestures and conceptual shifts. *Cognitive Linguistics* 9(3), 283.
- Liddell, Scott. 2003. *Grammar, gesture and meaning in American Sign Language*. Cambridge University Press.
- McCarthy, Anjanie, Lee, Kang, Itakura, Shoji & Darwin W. Muir. 2008. Gaze Display When Thinking Depends on Culture and Context. *Journal of Cross-Cultural Psychology* 39(6). 716–729.
- McClave, Evelyn Z. 2000. Linguistic functions of head movements in the context of speech. *Journal of Pragmatics* 32(7). 855–878.
- Nikiforidou, Kiki. 2010. Viewpoint and construction grammar: the case of past + now. *Language and Literature* 19(3). 265–284.
- Nikiforidou, Kiki. 2012. The constructional underpinnings of viewpoint blends: The Past + now in language and literature. In Barbara Dancygier & Eve Sweetser (eds.), *Viewpoint in language: A multimodal perspective*, 177–197. Cambridge: Cambridge University Press.
- Özyürek, Asli. 2002. Do speakers design their cospeech gestures for their addressees? The effects of addressee location on representational gestures. *Journal of Memory and Language* 46(4). 688–704.
- Parrill, Fey. 2009. Dual viewpoint gestures. *Gesture* 9(3). 271–289.
- Parrill, Fey. 2010. Viewpoint in speech-gesture integration: Linguistics structure, discourse structure, and event structure. *Language and Cognitive Processes* 25(5). 650–668.
- Parrill, Fey. 2012. Interactions between discourse status and viewpoint in co-speech gesture. In Barbara Dancygier & Eve Sweetser (eds.), *Viewpoint in language: A multimodal perspective*. 97–112. Cambridge: Cambridge University Press.
- Parrill, Fey & Eve Sweetser. 2004. What we mean by meaning: conceptual integration in gesture analysis and transcription. *Gesture* 4(1). 197–219.
- Park, Yujong. 2009. Interaction between grammar and multimodal resources: quoting different characters in Korean multiparty conversation. *Discourse Studies* 11(1). 79–104.
- Quinto-Pozos, David. 2007. Can constructed action be considered obligatory? *Lingua* 117(7). 1285–1314.
- Rossano, Federico. 2012. Gaze in social interaction. In Jack Sidnell & Tanya Stivers (eds.) *Handbook of Conversation Analysis*. 308–329. Malden, MA: Wiley-Blackwell.
- Sanders, Jose and Gisela Redeker. 1996. Perspective and representation in speech and thought in narrative discourse. In Gilles Fauconnier & Eve Sweetser (eds.), *Spaces, worlds and grammar*. 290–317. Chicago: University of Chicago Press.
- Schegloff, E. A. 1998. Body Torque. *Social Research*, 65(3). 535–596.
- Sidnell, Jack. 2006. Coordinating gesture, gaze and talk in reenactments. *Research on Language and Social Interaction* 39(4). 377–409.

- Stec, Kashmiri. 2012. Meaningful shifts: A review of viewpoint markers in gesture and sign language. *Gesture* 12(3). 327–360.
- Stec, Kashmiri & Mike Huiskes. 2014. Co-constructing referential space in multimodal narratives. *Cognitive Semiotics* 7(1).
- Stec, Kashmiri, Mike Huiskes, Alan Cienki & Gisela Redeker. Submitted. Annotating bodily indicators of perspective shifts in conversational narratives.
- Stec, Kashmiri & Eve Sweetser. 2012. Significant breaks: Space and viewpoint transitions in gesture. Paper presented at the 2012 International Gesture Studies Conference, in Lund, Sweden.
- Stec, Kashmiri & Eve Sweetser. 2013. Managing multiple viewpoints: Coordinating embedded perspective in multimodal narrative. Paper presented at the 12th International Cognitive Linguistics Conference, in Edmonton, Alberta.
- Sweetser, Eve. 2012. Introduction: Viewpoint and perspective in language and gesture, from the Ground down. In Barbara Dancygier & Eve Sweetser (eds.), *Viewpoint in language: A multimodal perspective*, 1–2. Cambridge: Cambridge University Press.
- Sweetser, Eve. 2014. Creativity across modalities in viewpoint construction. In Mike Borkent, Barbara Dancygier & Jennifer Hinnell (eds.), *Language and the creative mind*. 239–254. Stanford CA: CSLI Publications.
- Sweetser, Eve and Marisa Sizemore. 2008. Personal and interpersonal gesture space: functional contrasts in language and gesture. In Andrea Tyler, Yiyoung Kim & Mari Takada (eds.). *Language in the Context of Use: Cognitive and Discourse Approaches to Language and Language Learning*. 25–51. Berlin: Mouton de Gruyter.
- Thompson, Sandra A. & Ryoko Suzuki. 2014. Reenactments in conversation: Gaze and recipiency. *Discourse Studies* 16(6). 816–846.
- Vandelanotte, Lieven. 2009. *Speech and thought representation in English*. Berlin/New York: Mouton de Gruyter.
- Vandelanotte, Lieven. 2012. ‘Wait till you got started’: How to submerge another’s discourse in your own. In Barbara Dancygier & Eve Sweetser (eds.). *Viewpoint in language: A multimodal perspective*. 198–218. Cambridge: Cambridge University Press.
- Verhagen, Arie. 2005. *Constructions of intersubjectivity: Discourse, syntax and cognition*. Oxford: Oxford University Press.