

SPECIAL COLLECTION: PERSPECTIVE TAKING

Multimodal character viewpoint in quoted dialogue sequences

Kashmiri Stec, Mike Huiskes, Martijn Wieling and Gisela Redeker

University of Groningen, PO Box 716, 9700 AS Groningen, NL

Corresponding author: Kashmiri Stec (kashmiri.stec@gmail.com)

We investigate the multimodal production of character viewpoint in spoken American English narratives by performing complementary qualitative and quantitative analyses of two quoted dialogues, focusing on the storyteller's use of character viewpoint gestures, character intonation, character facial expression, spatial orientation and gaze. A micro-analysis revealed that the extent of multimodal articulation depends on (i) the quoted speaker, with different multimodal articulatory patterns found for quotes by the speaker's past self vs. a third-person character, and (ii) the position of the quoted utterance within the quoted dialogue, with mid-dialogue utterances garnering less co-articulation than initial or final utterances within the quoted dialogue. We further investigated these observations using a quantitative approach, which was based on *generalized additive modeling* (GAM). The GAM analysis revealed different multimodal patterns for each quoted character, as indicated by the number of co-produced multimodal articulators. These patterns were found to hold regardless of the quote's position within the narrative. We discuss these findings with respect to previous work on multimodal quotation.

Keywords: multimodality; co-speech gesture; viewpoint; direct speech; quotation

1 Introduction

A large part of our daily interactions concerns the recounting or narrating of prior experiences – a capability which is, as far as we know, unique to humans (cf. Turner 1998; Donald 2001; Zunshine 2006). The narrated events usually involve interactions which the narrator has witnessed or has been involved in and are often rendered in direct quotations of the characters' utterances and thoughts (Labov 1972; Li 1986; Tannen 1989). Past events are brought to life by shifting the viewpoint to the quoted characters, whose utterances are often dramatized with expressive intonation and facial, manual, or other bodily articulations (Polanyi 1989; Clark & Gerrig 1990). In this paper, we investigate how this *multimodal co-articulation* is used to distinguish between characters in extended stretches of quoted dialogues. With qualitative and quantitative analyses we show how one narrator, who produced the two longest quoted dialogue sequences in our corpus of semi-spontaneous narratives by American English speakers (collected by the first author; see Stec 2016), employed paraverbal and nonverbal means to embody and differentiate the characters in her story.

Multimodal co-articulation has been associated with the function of quotations as demonstrations (Clark & Gerrig 1990; Bavelas et al. 2014), depictions (Clark 2016), (re) enactments (Streek 2002; Sidnell 2006) and viewpoint shifts (McNeill 1992; Dancygier & Sweetser 2012; Stec et al. 2015; 2016). In studies of signed languages, the concepts of role shift, constructed dialogue, constructed action, or perspective shift are used for the representation of actions, thoughts and feelings of narrative characters via manual and

non-manual means, such as the systematic use of gaze, facial portrayals and sign space to manage narrative structure (see Metzger 1995; Quinto-Pozos 2007; Janzen 2012; comparative studies of speakers and signers include Rayman 1999; Marentette et al. 2004; Earis & Cormier 2013).

Recent studies of conversational storytelling have provided strong evidence that speakers make prolific use of multimodal co-articulation in quotations (Park 2009; Bavelas et al. 2014; Thompson & Suzuki 2014; Blackwell et al. 2015; Stec et al. 2015; 2016). Gaze, facial expression, body posture, character gestures, and in spoken narratives also intonation, have all been shown to signal the viewpoint shift involved in quotation. Sidnell (2006) notes that gaze specifically directed away from interlocutors can signal that a reenactment (by which is typically meant a quoted utterance) is taking place, and that other nonverbal actions such as gestures and facial portrayals, which can be evocative of the reenacted character or scenario, are used to highlight a *live as it happened* account for addressees. Sidnell further notes that multimodal production typically accompanies the quoted utterance at what he calls the left boundary of the quote (Sidnell 2006: 382). Bavelas and Chovil (1997) note that facial portrayals are commonly used when demonstrating characters' emotional reactions to different situations.

The use of iconic manual gestures in quotations has been studied by eliciting narratives based on cartoon stimuli (e.g. McNeill 1992; Holler & Wilkin 2009; Parrill 2010). These studies focus on the distinction between *character viewpoint gestures* (CVPT) such as, e.g., grasping hands moving upwards to demonstrate the character climbing a ladder, and *observer viewpoint gestures* (OVPT), where the same event is demonstrated from an onlooker's perspective by a stepwise upward movement of the index finger. Parrill (2010) found that certain events lend themselves more easily to character or observer viewpoint gestures. Discourse structure and the interactive context also appear to play a role: discourse-central events, but also new information, first mentions and re-introductions all tend to be accompanied by CVPT gestures, while given information and maintenance contexts tend to be accompanied by OVPT gestures (Gerwing & Bavelas 2004; Perniss & Özyürek 2015).

All of these studies point to individual multimodal articulators which are used during quotation. Stec et al. (2016) take these observations a step further and note that, in our corpus of 85 semi-spontaneous narratives told by 26 native speakers of American English, speakers often simultaneously use multiple multimodal articulators during quotation to achieve something like role shift as is typically described for users of signed languages. Similarly, Park (2009) describes rich multimodal co-articulation of quotations in Korean multiparty conversations. Note, however, that the majority of the quotations in Park's corpus concern interactions between participants in those conversations, that is, 1st- and 2nd-person quotes, so that much of the multimodal co-articulation serves the interactional management of the quotes among the co-narrators and co-participants. In our corpus, collected to elicit viewpoint shifts, the narratives were produced semi-spontaneously in a dyadic situation. They often concerned past interactions of the narrator and thus contained 1st-person quotes of the narrator's past self (395 of the 704 quotations), but only eight quotations were quoting the addressee (see Table A1 in the Appendix).

In addition to these aspects of multimodal production, research also indicates that the multimodal production of quotes may also be sensitive to the quotation environment. Stec et al. (2015) discuss the extent to which multimodal articulation can differentiate between single quotes, quoted monologues and quoted dialogues, although there are also some similarities – such as the use of gaze with character facial expression, or a general scarcity of CVPT gestures. To further explore the effect of sequential position of a quote and character alternation in extended sequences of quotations, we have selected the two

longest quoted dialogue sequences from our corpus (21 and 16 quotes respectively), as these exceptionally long sequences provide the ideal environment for investigating the extent to which multimodal actions can be used to distinguish quoted characters. When adopting character viewpoint, e.g., when successively quoting different characters in a dialogue, speakers might use co-articulated multimodal actions to signal to their addressees that a shift to contrastive character perspectives is taking place, and might further use these actions to differentiate characters within a narrative. Signers have been demonstrated to do this when quoting contrastive perspectives (Padden 1986 – but see Janzen 2012 for signers who use an alternate strategy). Stec et al. (2015) note that it is an open question what speakers would do in the same situation. It might also be the case that the multimodal actions used by speakers change over time as characters within a narrative are re-referenced, or depending on which character was quoted. As others have observed (So et al. 2009; Gunter et al. 2015; Perniss & Özyürek 2015), there is an important relationship between manual gesture, spatial location and repeated reference: whereas speakers have been shown to produce (and listeners to expect) relatively stable spatial locations for referents, the way in which the manual gestures associated with those referents are produced varies, with a gradual reduction in complexity and representation over time. Given that speakers do produce multimodal quoted utterances with some systematicity, it might be the case that as a character is re-quoted in an extended quoted dialogue sequence, the multimodal component associated with that character is used consistently. Alternatively, it might be the case that the complexity of the multimodal component gradually decreases with time.

In light of these considerations, we pose the following research questions: (i) Can multimodal co-articulation be used to differentiate characters in a spoken narrative? and (ii) If so, how consistent is that differentiation? We will answer these questions by investigating the use of character viewpoint gestures, character intonation, character facial expression, and changes in spatial orientation and gaze co-timed with quoted utterances in two extended quoted dialogue sequences. First, we provide a micro-analysis of multimodal behaviors occurring with quotations in initial, medial, and final position in the quoted sequences. We then use *generalized additive modeling* (GAM) to investigate how the use of multimodal behaviors changes with time.

2 Method

Our corpus consists of 85 semi-spontaneous narratives told by 26 native speakers of American English and collected and annotated by the first author. In previous work (Stec et al. 2015; 2016), we investigated multimodal quotation in the entire corpus. In this paper, we focus on one narrative which contains two exceptionally long quoted dialogue sequences. We present a micro-analysis of the multimodal behaviors used in these quotation sequences in Section 3 and a quantitative analysis of those same behaviors using generalized additive modeling in Section 4.

2.1 Overview of collection and annotation procedures

For the larger project this study is part of (Stec 2016), we collected semi-spontaneous autobiographical narratives from pairs of native speakers of American English. All pairs of speakers knew each other, and were asked to tell each other personal narratives they would be comfortable having videotaped. The 26 participants (17 female, 9 male) volunteered their time and consented to the use of the videotaped materials in our research and in publications. In the 85 narratives they told each other, we identified 704 quoted utterances that formed the corpus for the larger project (see Table A1 in the Appendix).

The quotations were annotated fully by the first author, and were assessed for inter-annotator validity using a consensus procedure for which annotations made on a subset of the data (10%) were compared with annotations made by the second author and three independent coders.¹ These comparisons involved discussions aimed at identifying and resolving underlying sources of disagreement. For the complex phenomena we are investigating, such a stepwise consensus procedure is more valuable than a purely quantitative assessment of inter-annotator agreement, e.g. with Kappa (cf. Stelma & Cameron 2007 and Gnsici et al. 2013 who caution against its use). More information about our methods – including annotation scheme, annotation procedure, annotated data, R scripts used for analysis, etc. – is available in a paper package hosted at the Mind Research Repository.

2.2 Quotations in the Airports story

The narrative we analyze is called *Airports* and is narrated by a woman whom we call *Black* for the color of shirt she wore at the time of recording. The story is 3 minutes and 16 seconds long. It contains 38 quoted utterances, 37 of which occur during two longer stretches of quotes from an encounter of Black's past self with airport officials (A). Those two quoted dialogue sequences were used in this study. The first sequence is comprised of 21 quotes (past self: 9 quotes; A: 12 quotes), and the second of 16 quotes (past self: 7 quotes; A: 9 quotes).

Before moving forward, we should note that most quoted dialogues in our corpus are much shorter, containing only three quoted utterances, while some contain as many as six. Previous research on this corpus (e.g. Stec et al. 2015) demonstrates the extent to which quoted dialogue sequences are accompanied by different kinds of multimodal actions compared to single quotes or quoted monologues. We do not know how common the extended quoted dialogue sequences analyzed here are in everyday talk. However, because of the repeated shifts in character perspective between Black's past self and the airport officials, they provide an ideal situation for investigating if and how multimodal articulators are used to uniquely identify and represent characters during maintained alternating perspective shifts.

2.3 Annotation

We used ELAN (Wittenburg et al. 2006) to implement our annotation scheme with a hierarchical arrangement of tiers (variables) and controlled vocabularies (values).² Only a subset of the annotation scheme described by Stec (2016: Chapter 4) is relevant for this analysis, and is presented in Table 1, showing the parallel annotations (the *tiers*) and the predefined categories within each tier (the *controlled vocabulary*).

We are interested in the degree to which different multimodal articulators contribute to the expression of multimodal viewpoint during direct speech quotes. Stec (2012: 351–353) identifies the articulators that can be used to express viewpoint shifts in co-speech gesture, and we used those observations to create an annotation scheme that captures the extent to which different articulators *actively* contribute to these viewpoint shifts.

First, we noted whether an utterance was a quotative. Only quoted utterances were annotated. Quotes were identified on the basis of quoting predicates such as *say* or *be like*, or on other indicators of direct speech such as a switch to first person or shifted

¹ As an initial step in this procedure, Cohen's Kappa was computed. Linguistic variables showed high inter-annotator reliability (all Kappa's above 0.8). Non-linguistic variables mostly obtained low Kappa values (0.2 to 0.6), indicating the need for a consensus procedure to ascertain the validity of this fine-grained coding.

² ELAN is a multimodal software tool developed for use by the Language Archive at the Max Planck Institute for Psycholinguistics, and is freely available at <http://tla.mpi.nl/tools/tla-tools/elan/>.

Category	Tier	Controlled vocabulary
Linguistic information	Transcript	Text
	Utterance type	<ul style="list-style-type: none"> – Quote (the utterance is an instance of direct speech) – Not a quote (utterance is not a quote, and will not be further annotated)
	Quoted character	<ul style="list-style-type: none"> – Speaker (the speaker quotes themselves) – Addressee (the speaker quotes their addressee) – Speaker+Addressee (the speaker quotes themselves + their addressee) – A–F (the letters A–F are used to identify other quoted characters in the narrative)
Multimodal articulators	Character intonation	<ul style="list-style-type: none"> – Present (speaker’s voice altered to demonstrate the quoted character) – Absent (speaker’s voice unchanged) – Unclear
	Hands	<ul style="list-style-type: none"> – Character viewpoint gesture (speaker’s hands demonstrate a manual action performed by another entity) – Other gesture (including beats, iconic gestures which are not character viewpoint, deictic gestures, emblems, etc.) – No gesture
	Character facial expression	<ul style="list-style-type: none"> – Present (speaker’s facial expression changes to demonstrate the quoted character) – Absent (speaker’s facial expression is unchanged) – Unclear
	Gaze	<ul style="list-style-type: none"> – Maintained with addressee (speaker’s gaze is directed to addressee throughout the quote) – Away from addressee (speaker’s gaze is not directed to the addressee throughout the quote) – Late change (speaker’s gaze moves away from the addressee after the quote started) – Quick shift (speaker’s gaze jumps around throughout the quote) – Unclear
	Posture change	<ul style="list-style-type: none"> – Horizontal (the speaker moves in a horizontal direction) – Vertical (the speaker moves in a vertical direction) – Sagittal (the speaker moves in a sagittal direction) – Unclear – None (the speaker’s body does not move)

Table 1: ELAN tiers and controlled vocabularies used in this analysis (adapted from Table 4.2 in Stec 2016).

temporal or locational deictics. See Buchstaller (2013) on identifying direct speech in discourse, and Dancygier & Sweetser (2012) on the multimodal expression of viewpoint. For each quote, we noted which character was quoted. Next, we noted which multimodal articulators were *actively* co-produced with the quoted utterance. Our notion of *active* articulators aims to capture the fact that there is a difference between a speaker who, for example, shows a neutral facial expression with a manual CVPT gesture for paddling a kayak (both hands clenching in fists while simultaneously making figure eights) and a speaker who makes the same gesture but uses their face to also depict an emotion such as excitement, terror, or determination. In both cases, the speaker’s entire body can be said to depict character viewpoint (cf. McNeill 1992), but only in the second case can we say that their face actively represents the character. The five active multimodal articulators identified in this project are: character intonation, manual CVPT gestures, facial expressions which depict the quoted character, any non-neutral use of space, and any meaningful use of gaze.

Previous work on the depiction of character perspective in speakers has identified the use of character intonation, character facial expression and, to a limited extent, manual character viewpoint gestures as being indicative of character perspective in both narrative (Earis & Cormier 2013) and quotative (Stec et al. 2015) environments. While body torque or body shift has previously been described as a means by which speakers negotiate activities within a given space (Schegloff 1998), no link has yet been made between such shifting and viewpoint shifts. Amongst signers, a shift in torso or body orientation is often associated with viewpoint shifts during quotation and more generally in constructed action sequences (Padden 1986). Thus, any change in orientation (coded in the Posture Change tier as any category except *none*) is considered active. Previous work on the use of gaze has indicated that speakers often look away at the start of a quoted utterance (Sidnell 2006), a finding which is comparable to some descriptions of constructed action in sign languages (Metzger 1995). As such shifts in gaze can be indicative of viewpoint shift, we identified possible values for the meaningful or active use of gaze as *looks away*, *late change* (speaker's gaze moves away from the addressee after the quote started) or *quick shift* (speaker's gaze jumps around throughout the quote) but not *maintains gaze with addressee*.

Finally, for the GAM analysis, we created two variables. Articulator Count counts the number of active articulators used by the narrator. Possible values range from 0 (no articulators active) to 5 (all articulators active). For example, an utterance with character intonation (present, 1), no manual gesture (0), character facial expression (present, 1), accompanied by non-neutral body movement (sagittal, 1) and maintained gaze with addressee (0) has an Articulator Count of 3. The second variable is Sequential Position. It treats each quote as an item in a sequence, and preserves narrative order.³ It ranges from 1 (the first quote in the first quoted dialogue) to 37 (the last quote in the second quoted dialogue).

3 Multimodal articulation in quoted dialogues

In this section, we offer a qualitative micro-analysis of the multimodal quotes produced in the narrative *Airports*. Black, on the right in the figures below, talks about the frustrations she experienced in airports as a dual citizen of the US and Ireland. The story focuses on one incident which happened the previous summer when Black visited friends and family in France and flew home to the US from Spain via a major German airport. Customs officials at that airport interrogated her when she accidentally showed her EU passport rather than her US passport when boarding her flight to the US. In the first quoted dialogue, Black recounts the official's attempts to understand her summer itinerary, ascertain her citizenship and determine where her family lives. In the second quoted dialogue, Black recounts the official's frustrated attempts to search her electronics and baggage – only to discover she only has a carry-on, and no electronic devices. Bemused, they let her board the plane. In both quoted dialogue sequences, Black distinguishes her past self from the quoted airport officials by using multimodal indicators of character viewpoint. We observe a general pattern whereby quotes by the airport officials are accompanied by more multimodal articulators than quotes by Black's past self. Additionally, we observe a positional difference in the use of these multimodal indicators across the quoted utterances in the dialogues, with beginnings and endings marked differently than middle sequences. We

³ When performing a GAM analysis, time can be included in several ways (actual time, relative time, or sequential position). We chose sequential position as our measure since this allowed us to describe the development of multimodal articulatory patterns via repeated mentions of the same characters rather than an overall progression of narrative time. We ran models using relative time or sequential position, and obtained similar results.

discuss illustrative examples here; a complete transcript of the quoted dialogues is provided in Table A3 in the Appendix.

3.1 Distinguishing quoted characters

Throughout the quoted dialogues, Black uses multimodal production to distinguish her past self from the airport officials. Typically, quotes of the airport officials are accompanied by more active multimodal articulators than quotes of her past self. One way Black distinguishes between the two characters in the quoted dialogues is by her use of facial expressions. This is shown in Figure 1 and Transcript 1, from the beginning of the second quoted dialogue. Each line of the transcript corresponds to an image in the figure, e.g., line 1 is co-articulated with the behaviors in image 1. Each transcript is formatted as follows: *Speaker_name: [quoted.speaker] quote*. *A* indicates the airport officials, and *past.self* indicates Black's past self in the airport encounter. At the end of each quote in the transcript, we indicate the total number of active articulators during that quote, so that, e.g., [4] after quote 24 (in lines 3–4) means that four articulators were active.

In this example, we see character facial expressions in each image – with more emphatic expressions in images 3 and 4, where Black quotes her past self. In addition, Black uses more of her gesture space in multimodal utterances accompanying her past self (images 3 and 4 in Figure 1): her head makes more pronounced movements with multiple movement phases, and both of her hands are used to gesture in an effortful way, as indicated by a comparison of their location and handshape in images 1 and 4. Both, 1st- and 3rd-person quoted characters get special intonation in this sequence, quotes of the airport officials are marked with a voice change: Black's voice takes on a deeper, authoritative quality during lines 1–2 of the transcript. In contrast, during the quotes on lines 3–4 Black's past self sounds puzzled. Both quotes are considered to have four active articulators: character facial expression, character intonation, meaningful use of gaze and body movement. In this example, the difference in the production of the two quotes is not the number of active articulators but rather the way that these articulators are used: sharper, controlled, authoritative movements accompanying the quotes by the airport officials, and emphatic puzzlement accompanying the quotes by Black's past self.

Another strategy is exemplified in Figure 2 and Transcript 2, from the end of the first quoted dialogue. Here, we observe character facial expressions for each quoted utterance,



Figure 1: Stills from *Airports*.

1	Black:	[A] they were like you were in Europe for three weeks	
2		what did you do with all your stuff	[2]
3	[past.self]	I was like I didn't bring stuff	
4		like I just have this backpack (0.3) I swear	[4]

Transcript 1: *Airports* 1:42–1:49, quotes 23–24.



Figure 2: Stills from *Airports*.

1 Black: [A] they were like *why do you have an Irish passport* [3]
 2 [past.self] I was like *I have this one too* [5]
 3 [A] and they were like *god why did you show us the Irish one* [5]

Transcript 2: *Airports* 1:15–1:20, quotes 19–21.

as well as CVPT gestures for each quoted utterance: in image 1, while quoting the airport officials, Black's left hand is raised, and then rises higher, demonstrating the airport officials' confusion; in image 2, her right hand moves forward, and then moves even more forward, to demonstrate Black's past self offering her US passport to the airport official; and in image 3, her left hand, initially held near her face, moves down towards her shoulder, demonstrating the airport officials' exasperation. In addition, Black changes the orientation of her head during lines 2–3 of the transcript (images 2–3, respectively, in Figure 2), marking the alternation between the past self and the 3rd-person characters, first moving right (image 2) and then down (image 3). The quote in line 1 has an Articulator Count of 3 (no character intonation or CVPT gesture), while the quotes in lines 2 and 3 each have an Articulator Count of 5, as all articulators are active.

In addition to bodily movements with multiple phases and an increase in the size of her gesture space in this example, Black's head movements become larger and swifter, and her voice takes on a different quality, together indicating a shift in perspective. This change in voice quality is one means by which Black distinguishes between the two quoted characters. When Black quotes her past self, her voice remains neutral, but when she quotes the airport officials, her voice becomes deeper and more resonant. Another means of differentiating characters are head movements, with lateral and vertical changes used to mark the shift from one character to the next.

As these examples show, the co-articulated multimodal actions involved in these sequences are evocative of character viewpoint (McNeill 1992) or reenactments (Sidnell 2006) insofar as manual gestures and non-manual articulators work together with the spoken utterances to visually embody different aspects of the quoted character: quotes of the airport officials are typically accompanied by movements which are controlled and authoritative, while quotes of Black's past self are typically accompanied by movements which evoke puzzlement and exasperation. We see not only character traits, but also contrasts, e.g. in the change in head orientation and movement, the vocal characteristics of the quoted characters, as well as their facial expressions and overall demeanor. Thus, in each of these examples, we observe a multimodal enactment of the quoted characters during Black's experience at the airport.

As we will show in Section 3.2, the multimodal articulators used in the beginning and end of quoted dialogues do not differ from the indicators used to in the middle of the quoted dialogues. However, the degree to which they are used and the manner in which the gestures are produced differs.

3.2 Moving in and out of quoted dialogues

One prominent pattern of multimodal utterance production in *Airports* concerns the way each quoted dialogue sequence begins and ends. One interactional task a storyteller faces when moving into quotation is the marking of continuing shifts of perspective as the story progresses from reporting about past events to the enactment of the characters interacting in an episode. At the end of a quotation sequence, the storyteller is faced with a related task, marking the end and climax of the episode and moving out of the quoted perspectives back to the present. In *Airports*, Black marks these shifts in a specific way, and this marking is more prominent in quotes appearing dialogue initially and finally than in quotes appearing in mid-dialogue (for the latter, see Figure 3).

As the sequences begin, quotes generally tend to be longer as Black sets the stage for the extended quoted dialogue sequence. When moving in or out of a sequence, the stroke of Black's manual gesture is comprised of several movements (Bressem & Ladewig 2011), and her gaze and head make several movements as well, e.g. her head tilts left and then farther left. In addition, Black makes use of an extensive gesture space – her gestures are normally comprised of small articulations made close to the body, but in these examples, we can see that she comfortably uses a larger gesture space. This was exemplified by Transcripts 1 and 2 and Figures 1 and 2 (Section 3.1).

Whereas beginnings and endings of quoted dialogue sequences seem to be marked by extensive use of multimodal markers, a different production strategy is evident as the quoted dialogue unfolds. Quoted utterances in mid-sequence occur without quoting verbs, e.g., as bare quotes (Matthis & Yule 1994), as Black swiftly and efficiently alternates between voicing the airport officials and her past self. Differentiation of the characters is maintained throughout. Quotes by Black's past self often contain only one active articulator. By contrast, Black tends to use marked character intonation for the airport officials (a lower almost masculine voice which sounds authoritative), and always makes a visible change with her body: sometimes a shift in gaze, sometimes a change in head or torso orientation, sometimes a tilt of her head or torso, sometimes a facial expression for one character or the other. However, these bodily actions are less pronounced here than they are at the beginning or end of quoted dialogues.

For example, consider Figure 3 and Transcript 3, which are taken from the first quoted dialogue. In this excerpt, direction of head movement is used to distinguish quoted characters (Figure 3, images 3, 4, 5, 6 and 7), as is the direction of gaze (towards the addressee in image 1, away from the addressee in image 2) and character facial expression (image 4). In addition, in image 1, Black's left hand moves from palm up to palm down.

There is a minimal marking of conceptual viewpoint shift across the quoted utterances in this example. From quote to quote, very few articulators are actively used to represent the quoted character – but the ones which are used are employed in a contrastive way, e.g. with vertical head movements for Black's past self and horizontal ones for the airport officials, or a shift in the direct of gaze which is co-timed with the onset of the quote. Once Black is in the middle of an extended quoted dialogue sequence, she is very consistent about this minimal production strategy.

A second example, this time from the second quoted dialogue, is given in Figure 4 and Transcript 4. The quoted dialogue sequence starts with Black maintaining gaze with her addressee. Following this, we see use of head movements to distinguish quoted characters (Figure 4, images 2, 3, 4, 6 and 8), and two right-handed gestures (one in image 5 and one in image 7, with the transition between them happening in image 6), both of which accompany utterances by the airport officials (A), and seem to indicate growing exasperation with the situation.

Again, we see a consistent differentiation of the quoted characters, with the airport officials becoming more emphatic and incredulous at the situation, and Black's past self



Figure 3: Stills from *Airports*.

1	Black:	[past.self] my dad	[1]
2		[A] where does he live	[2]
3		[past.self] France	[1]
4		[A] I thought you said you flew from Spain	[2]
5		[past.self] yeah	[1]
6		[A] is he Spanish or French	[3]
7		[past.self] he's Irish	[2]

Transcript 3: *Airports* 1:03–1:08, quotes 10–16.

giving in to the absurdity of their questions. Here we see head movements in alternate directions used for the quoted characters (horizontal left for the airport officials and horizontal right for Black’s past self, as well as vertical down for Black’s past self; note that these are not negative headshakes accompanying her denials), along with the kind of facial expressions and intonation patterns which have been evocative of both characters throughout the extended quoted dialogue sequences.

This suggests the following: first, quotes are co-produced with a number of articulators which work together in complex ways. Although most previous work on co-speech gesture has focused on the production of manual CVPT gestures, these examples show how flexible multimodal communication can be – and that given the right context, even the smallest of movements or multimodal actions can indicate important conceptual changes. Moving beyond the hands and investigating the contribution of other multimodal articulators is important if we want to document the extent to which language is multimodal. The extent to which multiple multimodal articulators contribute to multimodal utterance production, and how these articulators co-occur, should be investigated further. Second, we observe a consistent differentiation of the two quoted characters, Black’s past self and the airport officials, which is largely based on multimodal production. For each quoted utterance, the multimodal articulation differentiated the quoted characters. Different facial expressions and intonation patterns were used for the airport officials and Black’s past self, and each bodily action was made in a slightly different area of Black’s gesture space. Head movement, direction of gaze, facial expression, intonation and even manual gestures were used in particular ways to iconically represent



Figure 4: Stills from *Airports*.

1	Black:	[A] <i>you don't have a mobile phone</i>	[4]
2		[past.self] <i>no</i>	[2]
3		[A] <i>laptop</i>	[2]
4		[past.self] <i>no</i>	[1]
5		[A] <i>digital camera</i>	[3]
6		[past.self] <i>no</i>	[1]
7		[A] <i>iPod</i>	[2]
8		[past.self] <i>no</i>	[2]

Transcript 4: *Airports* 1:57–2:01, quotes 30–36.

each of the characters. Third, we saw a difference in multimodal production strategies which appears to vary with respect to position in the quoted dialogue. Earlier and later quotes were accompanied by more multimodal articulation – more multimodal articulators were active, and used more of Black's gesture space. Mid-dialogue quotes, on the other hand, were co-produced with fewer multimodal articulators compared to quotes at the beginning or end of the quoted dialogues, and used less of Black's gesture space. In general, we saw a pattern whereby quotes of the airport officials were accompanied by more simultaneously used multimodal articulators than quotes by Black's past self.

In summary, we find affirmative answers to our research questions: multiple multimodal articulators can be used to indicate a shift to character viewpoint, and these different articulatory patterns can be used to differentiate the quoted characters. Moreover, there appears to be an overall differentiation of characters across the extended quoted dialogues – e.g., the airport officials are always quoted with character intonation which is evocative of authority figures and character facial displays which indicate stern disbelief.

4 Modeling quoted dialogues

In this section, we present a model of the multimodal behaviors which accompany utterances quoting the airport officials and Black's past self by using GAM analyses. Previously, GAMs have been used to model psycholinguistic data, such as evoked-response potentials (Baayen 2010; Meulman et al. 2015) and the geographic distribution of dialects in the

Netherlands (Wieling et al. 2011). Here, we model the use of multimodal articulators accompanying quoted utterances from the airport officials (A) and the storyteller's past self. As GAMs have not previously been used to study co-speech gesture data, we provide a brief overview of the method. More information can be found in Wood (2006).

GAMs are an extension of generalized linear modeling (i.e. regression) which is able to assess non-linear relationships and interactions. GAMs model the relationship between individual predictor variables and dependent variables with a non-linear smooth function. The appropriate degree of smoothness of the non-linear pattern is assessed via cross-validation to prevent overfitting. For fitting the non-linear influence of our predictor of interest (Articulator Count), we use a thin plate regression spline (Wood 2003) incorporated in the *mgcv* package (Wood 2006; 2011; Wood et al. 2015) in R 3.2.0 (R Core Team 2014).⁴ We created a binary dependent variable for each quoted character (the airport officials, A, and Black's past self, speaker), which are inverse to each other (i.e. 1 for A means 0 for Black's past self and vice versa), and assessed the relationship between the number of active articulators and the quoted character. We also investigated if this relationship changed over time (via the Sequential Position of the quotes).

Figure 5 presents the results of our GAM analyses. The non-linear relationship between the Articulator Count and the probability of observing a quote of the airport officials (left plot) or of Black's past self (right plot) are shown. Note that the probability (including 95% confidence bands) is represented by logits, the log of the odds of seeing a quote from the airport officials versus Black's past self (and vice versa). Positive values indicate probabilities higher than 50%, while negative values indicate probabilities below 50% (0 indicates a 50% probability).

Overall, the two plots show that the airport officials' quotes are more likely to be accompanied by three or four articulators, and that quotes of Black's past self are more likely to be accompanied by a single articulator. Allowing for a non-linear interaction with time or Sequential Position (as suggested by our qualitative analysis; see Section 3.2) did not improve the model fit, indicating this pattern remains stable throughout both sequences. Plots of the raw Articulator Counts per utterance (see Figure 6) show that the

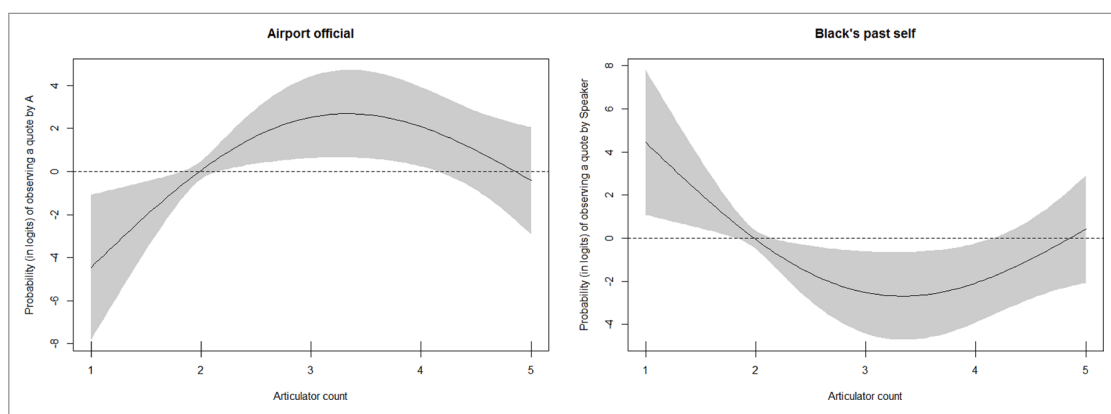


Figure 5: Probability curves obtained from the GAM analysis for the airport officials (left panel) and storyteller's past self (right panel).

⁴ We fitted identical models using the ML method, which is more conservative than the REML method reported here, and obtained similar results. As we are interested in the specific non-linearity, which is somewhat oversmoothed using the ML fitting method, we report the results on the basis of REML.

two characters' values tend to move in parallel maintaining a fairly stable difference in favor of the airport officials.

In sum, there is a systematic multimodal differentiation of both quoted characters, with quotes of the airport officials more likely to be accompanied by a variety of multimodal articulators, while Black's past self is more likely to be accompanied by fewer multimodal articulators. Table 2 shows the associated estimates of the model predicting from the Articulator Count if it is the airport officials that are quoted (the reverse prediction for Black's past self is redundant, as it yields the same estimates except for the sign of the intercept).

Overall, the results presented in this section demonstrate that some multimodal co-articulation is always present when Black quotes these two characters – thus, it is not a question of *whether* multimodal utterances occur, but of *how* they occur: which articulators are involved and how does their use change over time? The quantitative results confirm one aspect of our qualitative analysis: there is a differentiation of quoted characters in the number of articulators which are used, and this differentiation is maintained over the course of the quoted dialogue episode.

5 Discussion

In this case study of two extended quoted dialogue sequences, we have demonstrated that the narrator, Black, fluidly uses the multiple multimodal articulatory means available to her to not only iconically represent but also distinguish the two quoted characters. The quoted utterances were always accompanied by at least one multimodal articulator and often multiple articulators contributed to utterance production – although not every utterance was accompanied by full character embodiment (a finding in line with existing research, cf. Earis & Cormier 2013; Stec et al. 2016). Black's multimodal co-articulation

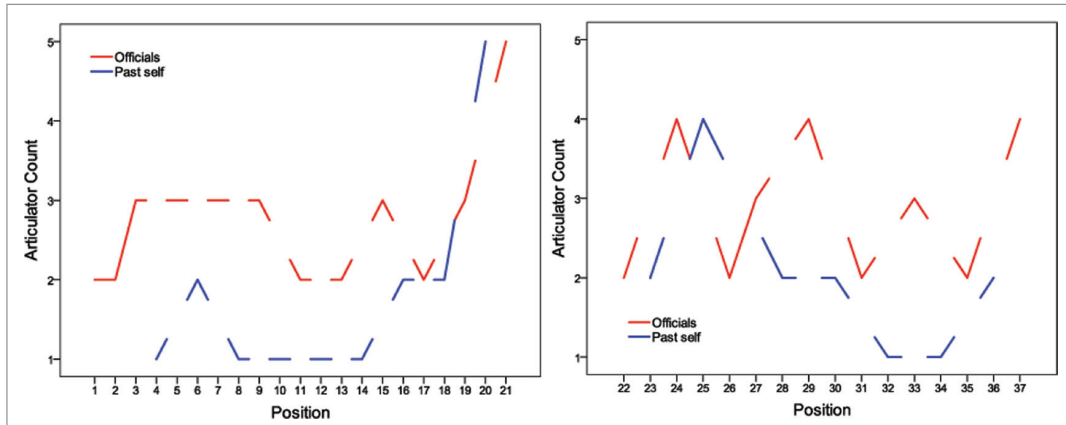


Figure 6: Articulator counts per quoted utterance from airport officials or Black's past self in the two dialogue sequences (for the text of the quoted utterances see Table A3).

Formula: $IsA \sim s(ArtCnt, k=3)$				
Parametric coefficients	Estimate	Std. Error	z-value	Pr(> t)
(Intercept)	0.09419	0.45672	0.206	0.837
Smooth terms	edf	Red.df	F	p-value
$s(ArtCnt)$	1.901	1.99	7.493	0.0234
R-squared (adjusted) = 0.404				

Table 2: GAM modeling the non-linear effect of the number of active articulators predicting quotes of the airport officials.

was systematic in the sense that its quantity and quality distinguished between the characters she quoted. This suggests that not only is language inherently multimodal – both in the sense that multiple modalities *and* different modes of production are involved throughout linguistic production – but that multimodal co-articulation can be used to achieve certain goals, such as the differentiation of characters within a narrative.

The qualitative micro-analysis demonstrated that Black used different multimodal articulation strategies to differentiate the two quoted characters in the extended quoted dialogue sequences investigated here. Quotes by the airport official were accompanied by multimodal articulators depicting authority and control. Quotes by Black's past self, on the other hand, were accompanied by multimodal articulators depicting puzzlement. Multimodal articulators were also used contrastively, e.g. the direction of head movement or gaze was used to differentiate characters, as were the facial expressions or intonation patterns which were evocative of each character. In this way, we saw a sustained differentiation of characters across the quoted dialogues.

The number and intensity of multimodal articulators differed between the quotes of the airport official and quotes of her own past self and also varied across the quoted sequences, with initial and final quotes of a sequences receiving multiple articulatory strokes (e.g. multiple head movements) and larger gesture spaces than quotes occurring in mid-sequence. In general, the activation was more exaggerated when more articulators were used, e.g., character facial expressions used more of the expressive qualities of Black's face, and character intonation was more pronounced. In the case of fewer articulators, articulators made only one stroke, and movements were made in a smaller space, closer to Black's body. Sequence-medial quotes, especially those quoting Black's past self, often used only gaze direction and head movements. Sometimes character facial expression and character intonation were used as well, but they were less pronounced.

The qualitative description of the data was complemented with a quantitative analysis that focused on the variety of different articulators, ignoring differences in intensity or repeated occurrences within one quote, thereby i.a. minimizing any correlation with the length of the utterances. Overall, the Articulator Count averaged 2.8 (see Table A2 in the Appendix), showing that the quotations in our corpus were commonly produced with multiple multimodal articulators. The GAM analysis of the two extended quoted dialogue sequences in this study demonstrated that Black generally distinguished the two most quoted characters in her narrative: three multimodal articulators often accompanied quotes of the airport official, while quotes by Black's past self were typically produced with a single multimodal articulator. The differentiation of characters was evident in the qualitative micro-analysis (e.g. shifts in head orientation or changes in the quality of character intonation from quote to quote) and the quantitative analysis showed that the pattern was indeed systematic and stable across the dialogue sequences.

Of course, there are limitations to our study: we analyzed one narrative, and only the quoted utterances of two characters (the only two quoted characters) within that narrative, as the two dialogue sequences we analyzed were the only ones of this considerable length in our entire corpus. While this case study has been instructive in several important ways, it also invites questions, such as: what happens in narratives with three or more characters? What do other speakers do? What would this speaker do in another narrative context? As we pointed out earlier, the number of quoted characters in semi-spontaneous narratives is variable, and often the speaker's past self is the most quoted. Investigating these questions might therefore entail an experimental design where the number of characters and the kind of quoted interactions can be manipulated. A limiting factor in our quantitative analysis is the fact that we modeled changes with respect to the number of articulators involved in multimodal utterances, not differences in how

those articulators were used or co-occurred. In other words, we were able to model categorical presence/absence rather than the fluid conversational dynamics which make personal narratives so compelling. While this can tell us something about the relative contributions of the body (i.e. that more or less of it contributed to multimodal utterance production), and can highlight the features discussed here, it masks the qualitative differences highlighted by our micro-analysis – such as differences in the amount of gesture space used, the degree to which active articulators were actually activated, or even the extent to which multiple multimodal articulators co-occur. At the same time, however, our analyses offered complementary perspectives on the situated practices used by Black throughout her narrative, and this complementarity paints an exciting picture in which more of the body – not only the hands – is involved in the articulation of multimodal utterances.

Another intriguing question concerns the generalizability of our finding that Black produces less multimodal co-articulation with quotes of her past self (1st-person quotes) than with quotes of the airport official (3rd-person quotes). This multimodal differentiation of characters might indicate that self-quotes are simply less marked than quotes of other characters, which is in line with findings from two previous studies on the linguistic realization of quotes (Golato 2002 on German self-quotes, and Rühlemann 2014 on English storytelling). Inspection of the Articulator Counts in our whole corpus of 704 quotes from 26 speakers telling 85 narratives shows that non-initial 3rd-person quotes were on average accompanied by more multimodal articulators than non-initial past self quotes (mean = 3.01 vs. mean = 2.66), while there was no difference for initial quotes (see Table A2 in the Appendix). This suggests that Black's differential treatment of past-self and 3rd-person quotes is not idiosyncratic and not limited to the specific setting of this narrative (e.g. the asymmetric roles of airport official and traveler). We can only speculate why the difference does not show in quotations of single utterances or in the initial quotes in quoted monologue and dialogue sequences. Possibly the task of initiating quotation, with the narrator lending their voice to a character (be it their own past self or a 3rd-person) obscures the differences in those initial quotes.

It should be noted that our finding may well be restricted to narratives that do not involve quotations of co-participants in a current interaction. In her Korean data, Park (2009) found that 1st- and 2nd-person quotes in multiparty conversations received much more multimodal co-articulation than the less frequent 3rd-person quotes did. Her study shows that multimodal co-articulation when quoting co-participants serves important interactional functions in the participants' orientation to what is essentially a joint narration. Park unfortunately does not differentiate between 1st-person quotes from earlier conversations with co-participants and those from conversations with others. Only the latter (much less frequent in her corpus) would be comparable to our data. Park's descriptions suggest that multimodal articulation in her data mainly supports the interactional management of the joint production and joint evaluation of the quoted dialogues. As most of her 1st-person quotes involve co-participants, those interactional functions can explain her finding of more multimodal articulation in 1st-person quotes than in 3rd-person ones. For our data, where only eight of the 704 quoted utterances quote the addressee of the narrative (see Table A1 in the Appendix), we have shown that multimodal co-articulation is used to signal viewpoint or role shift. With respect to this function, it seems plausible that self-quotes should be less marked than quotes of a 3rd-person character. As Sweetser (2012) notes, the human body is the best iconic representation of another human body. By extension, one's own body is the best representation of one's past self, and might therefore need fewer multimodal articulators to evoke itself in narrative contexts.

6 Conclusion

In summary, our results indicate that English speakers use multimodal utterances to differentiate characters in semi-spontaneous narrative by means of iconic representation, and at least one English speaker (Black) is able to maintain that differentiation over time. This iconic, multimodal representation may be more minimalistic or more fully embodied – but it is always present, supporting the view that language itself is multimodal. While we have demonstrated that English speakers are capable of using multiple multimodal articulators in a meaningful way, we do not yet know the extent to which people in general use this kind of iconic representation and differentiation during everyday communication, and the extent to which it aids the production or comprehension of quoted utterances or quoted sequences remains an open question. We hope that further research into the online grounding of multimodal perspective will address these issues.

Additional File

The additional files for this article can be found as follows:

- **Appendix.** [[Description]]. DOI: <https://doi.org/10.5334/gjgl.255.s1>

Abbreviations

CVPT = Character viewpoint gestures, OVPT = Observer viewpoint gestures, GAM = Generalized Additive Modeling

Acknowledgements

This paper has benefited from helpful comments and suggestions by three anonymous reviewers, whom we wish to thank for their thorough attention. The first author's contribution was financed by grant number 276.70.019, awarded by the Netherlands Organization for Scientific Research (NWO) to Esther Pascual.

Competing Interests

The authors have no competing interests to declare.

References

- Baayen, R. Harald. 2010. The directed compound graph of English. An exploration of lexical connectivity and its processing consequences. In Susan Olson (ed.), *New impulses in word-formation*, 383–402. Hamburg: Buske.
- Bavelas, Janet B., Jennifer Gerwing & Sarah Healing. 2014. Effect of dialogue on demonstrations: Direct quotations, facial portrayals, hand gestures, and figurative references. *Discourse Processes* 51(8). 619–655. DOI: <https://doi.org/10.1080/0163853X.2014.883730>
- Bavelas, Janet B. & Nicole Chovil. 1997. Faces in dialogue. In James A. Russell & Jose Miguel Fernandez-Dols (eds.), *The psychology of facial expression*, 334–346. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/cbo9780511659911.017>
- Blackwell, Natalia L., Marcus Perlman & Jean E. Fox Tree. 2015. Quotation as a multimodal construction. *Journal of Pragmatics* 81. 1–7. DOI: <https://doi.org/10.1016/j.pragma.2015.03.004>
- Bressem, Jana & Silva H. Ladewig. 2011. Rethinking gestural phases: Articulatory features of gesture movement? *Semiotica* 184(1/4). 53–91. DOI: <https://doi.org/10.1515/semi.2011.022>
- Buchstaller, Isabelle. 2013. *Quotatives: new trends and sociolinguistic implications*. Oxford: Wiley Blackwell. DOI: <https://doi.org/10.1002/9781118584415>

- Clark, Herbert H. 2016. Depicting as a method of communication. *Psychological Review* 123(3). 324–347. DOI: <https://doi.org/10.1037/rev000002>
- Clark, Herbert H. & Richard J. Gerrig. 1990. Quotations as demonstrations. *Language* 66(4). 764–805. DOI: <https://doi.org/10.2307/414729>
- Dancygier, Barbara & Eve Sweetser (eds.). 2012. *Viewpoint in language: A multimodal perspective*. Cambridge: Cambridge University Press.
- Donald, Merlin. 2001. *A mind so rare: The evolution of human consciousness*. New York, NY: W. W. Norton.
- Earis, Helen & Kearsy Cormier. 2013. Point of view in British Sign Language and spoken English narrative discourse: The example of ‘The Tortoise and the Hare’. *Language & Cognition* 5(4). 313–343. DOI: <https://doi.org/10.1515/langcog-2013-0021>
- Gerwing, Jennifer & Janet Bavelas. 2004. Linguistic influences on gesture’s form. *Gesture* 4(2). 157–195. DOI: <https://doi.org/10.1075/gest.4.2.04ger>
- Gnisci, Augusto, Fridanna Maricchiolo & Marino Bonaiuto. 2014. Reliability and validity of coding systems for bodily forms of communication. In Cornelia Müller, Alan Cienki, Ellen Fricke, Silva Ladewig, David McNeill & Sedinha Tessendorf (eds.), *Body language communication*, 879–892. Berlin: Mouton de Gruyter.
- Golato, Andrea. 2002. Self-quotation in German. In Tom Güldemann & Manfred von Roncador (eds.), *Reported discourse: A meeting ground for different linguistic domains*. Amsterdam: John Benjamins. DOI: <https://doi.org/10.1075/tsl.52.05gol>
- Gunter, Thomas C., J. E. Douglas Weinbrenner & Henning Holle. 2015. Inconsistent use of gesture space during abstract pointing impairs gesture comprehension. *Frontiers in Psychology* 6. 80. DOI: <https://doi.org/10.3389/fpsyg.2015.00080>
- Holler, Judith & Katie Wilkin. 2009. Communicating common ground: How mutually shared knowledge influences speech and gesture in a narrative task. *Language and Cognitive Processes* 24(2). 267–289. DOI: <https://doi.org/10.1080/01690960802095545>
- Janzen, Terry. 2012. Two ways of conceptualizing space. In Barbara Dancygier & Eve Sweetser (eds.), *Viewpoint in language: A multimodal perspective*, 156–175. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/cbo9781139084727.012>
- Labov, William. 1972. The transformation of experience in narrative syntax. In William Labov (ed.), *Language in the inner city: Studies in the Black English vernacular*, 354–396. Philadelphia, PA: University of Pennsylvania Press.
- Li, Charles N. 1986. Direct and indirect speech: a functional study. In Florian Coulmas (ed.), *Direct and indirect speech*, 29–45. Berlin/New York: Mouton de Gruyter. DOI: <https://doi.org/10.1515/9783110871968.29>
- Marentette, Paula, Natasha Tuck, Elena Nicoladis & Simone Pika. 2004. The effects of language, culture and embodiment on signed stories. Paper presented at *Theoretical Issues in Sign Language Research 8*. University of Barcelona, 30 September – 2 October.
- Matthis, Terrie & George Yule. 1994. Zero quotatives. *Discourse Processes* 18(1). 63–76. DOI: <https://doi.org/10.1080/01638539409544884>
- McNeill, David. 1992. *Hand and mind*. Chicago, IL: University of Chicago Press.
- Metzger, Melanie. 1995. Constructed dialogue and constructed action in American Sign Language. In Ceil Lucas (ed.), *Sociolinguistics in deaf communities*, 255–271. Washington, DC: Gallaudet University Press.
- Meulman, Nienke, Martijn Wieling, Simone A. Sprenger, Laurie A. Stowe & Monika S. Schmid. 2015. Age effects in L2 grammar processing as revealed by ERPs and how (not) to study them. *PLoS ONE* 10(12). e0143328. DOI: <https://doi.org/10.1371/journal.pone.0143328>


- Padden, Carol. 1986. Verbs and role shifting in American Sign Language. In Carol Padden (ed.), *Proceedings of the Fourth National Symposium on Sign Language Research and Teaching Sign*, 44–57. Silver Spring, MD: NAD.
- Park, Yujong. 2009. Interaction between grammar and multimodal resources: Quoting different characters in Korean multiparty conversation. *Discourse Studies* 11(1). 79–104. DOI: <https://doi.org/10.1177/1461445608098499>
- Parrill, Fey. 2010. Viewpoint in speech-gesture integration: Linguistic structure, discourse structure, and event structure. *Language and Cognitive Processes* 25(5). 650–668. DOI: <https://doi.org/10.1080/01690960903424248>
- Perniss, Pamela & Asli Özyürek. 2015. Visible cohesion: A comparison of reference tracking in sign, speech and co-speech gesture. *Topics in Cognitive Science* 7(1). 36–60. DOI: <https://doi.org/10.1111/tops.12122>
- Polanyi, Livia. 1989. *Telling the American story*. Cambridge, MA: MIT Press.
- Rayman, Jennifer. 1999. Storytelling in the visual mode: A comparison of ASL and English. In Elizabeth A. Winston (ed.), *Storytelling and conversation: Discourse in deaf communities*, 59–82. Washington, DC: Gallaudet University Press.
- R Core Team. 2014. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. <http://www.r-project.org/>.
- Rühlemann, Christoph. 2014. *Narrative in English conversation: A corpus analysis of storytelling*. Cambridge: Cambridge University Press.
- Schegloff, Emanuel A. 1998. Body torque. *Social Research* 65(3). 535–596.
- Sidnell, Jack. 2006. Coordinating gesture, talk and gaze in reenactments. *Research on Language and Social Interaction* 39(4). 377–409. DOI: https://doi.org/10.1207/s15327973rlsi3904_2
- So, Wing Chee, Sotaro Kita & Susan Goldin-Meadow. 2009. Using the hands to identify who does what to whom: Gesture and speech go hand-in-hand. *Cognitive Science* 33(1). 115–125. DOI: <https://doi.org/10.1111/j.1551-6709.2008.01006.x>
- Stec, Kashmiri. 2012. Meaningful shifts: A review of viewpoint markers in co-speech gesture and sign language. *Gesture* 12(3). 327–360. DOI: <https://doi.org/10.1075/gest/12.3.03.ste>
- Stec, Kashmiri. 2016. *Visible quotation: The multimodal expression of viewpoint*. Groningen: University of Groningen dissertation. <http://hdl.handle.net/11370/f6b13d04-fa53-4bcc-be5f-8e09d2da200d>.
- Stec, Kashmiri, Mike Huiskes & Gisela Redeker. 2015. Multimodal analysis of quotation in oral narratives. *Open Linguistics* 1(1). DOI: <https://doi.org/10.1515/opli-2015-0018>
- Stec, Kashmiri, Mike Huiskes & Gisela Redeker. 2016. Multimodal quotation: Role shift practices in spoken narratives. *Journal of Pragmatics* 104. 1–17. DOI: <https://doi.org/10.1016/j.pragma.2016.07.008>
- Stelma, Juurd H. & Lynne J. Cameron. 2007. Intonation units in spoken interaction: Developing transcription skills. *Text & Talk* 27(3). 361–393. DOI: <https://doi.org/10.1515/TEXT.2007.015>
- Streeck, Jürgen. 2002. Grammars, words, and embodied meanings: On the uses and evolution of *so* and *like*. *Journal of Communication* 52(3). 581–596.
- 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*, 1–23. Cambridge: Cambridge University Press.
- Thompson, Sandra A. & Ryoko Suzuki. 2014. Reenactments in conversation: Gaze and reciprocity. *Discourse Studies* 16(6). 1–31. DOI: <https://doi.org/10.1177/1461445614546259>
- Turner, Mark. 1998. *The literary mind: The origins of thought and language*. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780195126679.001.0001>

- Wieling, Martijn, John Nerbonne & R. Harald Baayen. 2011. Quantitative social dialectology: Explaining linguistic variation geographically and socially. *PLOS ONE* 6(9). e23613. <https://doi.org/10.1371/journal.pone.0023613>
- Wittenburg, Peter, Hennie Brugman, Albert Russel, Alex Klassmann & Han Sloetjes. 2006. ELAN: a Professional Framework for Multimodality Research. In *Proceedings of LREC 2006, Fifth International Conference on Language Resources and Evaluation*.
- Wood, Simon N. 2003. Thin-plate regression splines. *Journal of the Royal Statistical Society (B)* 65(1). 95–114. DOI: <https://doi.org/10.1111/1467-9868.00374>
- Wood, Simon N. 2006. *Generalized additive models: An introduction with R*. Boca Raton, FL: Chapman and Hall/CRC. <http://cran.r-project.org/web/packages/mgcv/index.html>.
- Wood, Simon N. 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society (B)* 73(1). 3–36. DOI: <https://doi.org/10.1111/j.1467-9868.2010.00749.x>
- Wood, Simon N., Yannig Goude & Simon Shaw. 2015. Generalized additive models for large data sets. *Journal of the Royal Statistical Society (C)* 64(1). 139–155. DOI: <https://doi.org/10.1111/rssc.12068>
- Zunshine, Lisa. 2006. *Why we read fiction: Theory of mind and the novel*. Columbus, OH: Ohio State University Press.

How to cite this article: Stec, Kashmiri, Mike Huiskes, Martijn Wieling and Gisela Redeker. 2017. Multimodal character viewpoint in quoted dialogue sequences. *Glossa: a journal of general linguistics* X(X): X.1–19, DOI: <https://doi.org/10.5334/gjgl.255>

Published: XX Month XXXX

Copyright: © 2017 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

 *Glossa: a journal of general linguistics* is a peer-reviewed open access journal published by Ubiquity Press.

OPEN ACCESS 