

# Free Indirect Discourse Meets Character Viewpoint Gestures

Cornelia Ebert & Stefan Hinterwimmer

Goethe University Frankfurt, University of Wuppertal

*ebert@lingua.uni-frankfurt.de, hinterwimmer@uni-wuppertal.de*

## 1 Introduction

In this paper we present a reconstruction of Davidson's (2015) demonstration account for quotations in the gesture semantics framework of Ebert & Ebert (2014). The paper has two related goals: the first is to provide support for Davidson's (2015) quotation as demonstration approach (as reconstructed by Maier, 2017; cf. Clark & Gerrig, 1990), motivating the additional demonstration condition postulated by Davidson (2015) with the general account of gestures by Ebert & Ebert (2014). The second goal is to provide empirical evidence for the quotation approach of free indirect discourse (FID) (Maier, 2015, 2017) and Davidson's (2015) demonstration approach for quotation in general by way of a rating study in which we looked at the interaction of perspective-taking at the linguistic level with perspective-taking at the level of gestures.

Davidson (2015) claims that certain sign language-specific constructions such as *role shift* and *classifier* constructions as well as quotation in spoken languages involve demonstrations. What Davidson characterizes as demonstrations can be considered gestures according to Ebert & Ebert's (2014) account of the semantics of co-speech gestures. In the first part of this paper we discuss additional data that corroborates Davidson's claim that quotations involve demonstrations (acted out as gestures) and propose an account of these data within the gesture semantics framework of Ebert & Ebert (2014).

In the second part, we present results from a rating study that aims to find empirical support for the claim that quotation involves demonstrations. In particular, we aim at showing that free indirect discourse (FID), as one instance of quotation, namely mixed quotation (Maier 2015, 2017, see also Dirscherl & Pafel, 2015), also involves demonstrational acts. According to Maier (2015, 2017), FID is a special, highly conventionalized form of mixed quotation in which pronouns and tenses are systematically unquoted. If quotation can involve a demonstrational act, as Clark & Gerrig (1990) and Davidson (2015) suggest, it should also be possible for FID to involve this type of demonstrational act, if it is indeed a form of mixed quotation. We found that direct discourse (DD) that involves demonstrations which are acted out as self-pointing gestures are judged as more natural than indirect discourse (ID) with such gestures. But more importantly, we also found preliminary evidence for a difference between ID and FID, namely that FID allows for speech accompanying self-pointing gestures more readily than ID. We argue that this difference between FID and ID can be explained if one adopts the view that FID involves quotation (Maier 2015, 2017), involves a perspective shift towards the individual whose thoughts or utterances are reported, and that self-pointing is a viewpoint gesture that shows the perspective of the (partially) quoted individual. This contrast between FID and ID is less obviously compatible with double context analyses of FID (Schlenker, 2004; Sharvit, 2008; Eckardt, 2014; see also Doron, 1991). We take this as initial empirical evidence for the mixed quotation approach.

The paper is structured as follows. In Section 2.1 Davidson’s (2015) analysis of quotation as demonstration is introduced and discussed. Section 2.2 introduces some novel observations going beyond Davidson (2015) that are in line with Ebert & Ebert’s (2014) proposal for gestures in general. Section 2.3 briefly summarizes the most important aspects of that proposal, which is then combined with Davidson’s (2015) approach in Section 2.4. Section 3 is the second part of the paper. In Section 3.1 FID is introduced as a form of speech or thought representation that combines features of DD with features of ID. Section 3.2 discusses two different analyses that have been proposed to capture the distinctive properties of FID: *double context analyses*, on the one hand, and the *mixed quotation approach*, on the other. Section 3.3 introduces the distinction between character viewpoint gestures and observer viewpoint gestures and their usage as forms of perspective-taking at the gestural level (Parrill, 2010; Stec, 2012 a.m.o.). In Section 4, we introduce and discuss an experimental study which tests the following prediction: Pointing gestures to the speaker’s body should be readily available with the use of first-person pronouns in DD constructions, since DD involves demonstration and these pointing gestures can be viewed as demonstrations realized as gestures that reveal the viewpoint the speaker adopts, here: the viewpoint of the quoted individual. Furthermore, self-pointing that is performed while uttering a third person pronoun referring to the person whose thoughts or utterances are reported should be interpretable as a demonstrational act that is part of the quotation in FID, but not in ID. This prediction naturally follows from two assumptions: First, in sentences reporting the thoughts or utterances of an individual distinct from the speaker, pointing gestures to one’s own body can be interpreted as character viewpoint gestures, which are demonstrations. Second, FID is a form of mixed quotation that is argued to involve such kinds of demonstrations.

## 2 Quotation as Demonstration

Davidson (2015), based on Clark & Gerrig (1990), claims that quotations involve demonstrations, i.e. acts that depict rather than describe certain aspects of an original event. These demonstrations can be acted out as bodily gestures. We present further data that corroborate Davidson’s approach and propose to capture these data in the gesture semantics framework of Ebert & Ebert (2014).

### 2.1 Davidson (2015)

In her seminal paper, Davidson (2015) argues that in sign languages, there are specific constructions that involve demonstrational acts (i.e. *classifier constructions*, *attitude role shift* and *action role shift*) and that, furthermore, *quotation* in signed and spoken languages is also a construction that features demonstrations. We turn to the three sign language-specific constructions and then quotation in turn.

*Classifiers* are very prevalent in sign languages. They can highlight specific details of an object or event. For example, they can reveal information about the size and shape of the object under discussion or about its movement by mimicking (some aspects of) the shape or movement of the objects to which they refer. In (1) you find an example from Davidson (2015) (her example (46), p. 494).<sup>1</sup>

- (1) a. *WOMAN CL-1-(LOCATE)*.  
‘There is a woman here.’

---

<sup>1</sup> *CL* is a classifier, realized by a certain handshape (here: the 1 handshape, i.e. a raised index finger, noted as “*CL-1*”), that is used for humans. If the subject were not a human, but a bicycle, for example, the handshape would be different, namely “*CL-3*” (see Davidson, 2015 for details). *LOCATE* and *MOVE* are demonstrations of (certain properties of) the agent’s location and movement.

b. *MAN CL-1-(LOCATE)*.

‘There is a man here.’

c. Right hand: *CL-1(MOVE)*

Left hand: *CL-1(MOVE)*

‘The two people walk toward each other.’

Overall interpretation: ‘The man and woman walked toward each other.’

(Davidson, 2015: 494)

Davidson (2015) notes that it is the combination of location and movement parameters of classifiers that presents a challenge for compositional semantics. This is because, although categorical meaning components are involved (the lexical sign for *MAN* and *WOMAN*, for example, plus the corresponding classifier), the way in which the two persons walk towards each other is entirely unspecified and the signer can choose to illustrate the walking properties in the way she wants and according to what she considers conversationally important, e.g. by using a zig-zag movement pattern. These more iconic meaning components are often considered gestural, as they cannot be lexically encoded and there are infinitely many possibilities for their realization.

*Attitude role shift* functions like quotation and the signer reports someone else’s speech or attitude and adopts the point of view of this individual. In *action role shift*, the signer reports in a particularly vivid way someone else’s action. Action role shift hence does not need to involve any kind of attitude verb. It is simply a perspective shift towards someone else’s thoughts or actions that can happen within a sentence. Both kinds of role shift are marked by the same specific marking strategies, e.g. body shifts or eye gaze break (see Sandler & Lillo-Martin, 2006; Quer, 2005 and many others).

In the following, we quote two examples of ASL (American Sign Language) from Schlenker (2017) (his examples (30a, b), p. 22). (2a) is an example for attitude role shift and (2b) for action role shift.<sup>2</sup>

(2) *SEE [THAT ARROGANT FRENCH SWIMMER] IX-a? YESTERDAY IX-a ANGRY<sub>a</sub>.*

‘Do you see that arrogant French swimmer? Yesterday he was angry.’

*RS<sub>a</sub>*

a. *IX-a SAY IX-1 WILL LEAVE.*

‘He said: ‘I will leave.’’

*RS<sub>a</sub>*

b. *IX-a 1-WALK-WITH-ENERGY(CL-ONE).*

‘He walked away with energy.’

(Schlenker, 2017: 22)

In (2a), the signer adopts the position of the swimmer (indicated by a body shift and marked as *RS<sub>a</sub>*) and reports what the swimmer said from the swimmer’s perspective. Note that the pronoun used in the reported speech is a first person pronoun (*IX-1*), which shows that we are not dealing with indirect speech.<sup>3</sup> In (2b), the speaker imitates the way in which the swimmer went away by way of using a classifier construction *WALK-WITH-ENERGY(CL-ONE)*, where the finger classifier makes a movement that is to represent that of the angry swimmer. No attitude verb is involved. Role shift starts after the third person pronoun *IX-a*. A first-person agreement marker inside the role-shifted clause is interpreted with respect to the shifted context.

In spoken language, the natural correspondent of attitude role shift is (direct) quotation and, according to Davidson (2015), the correspondent of action role shift is the *be-like* construction.

---

<sup>2</sup> *IX-a* is a third person pronoun, which is located at position *a*. Role shift starts from this location.

<sup>3</sup> It is sometimes argued that attitude role shift is not mere quotation either, because this construction has properties which are taken to be incompatible with quotation in general. For example, it allows for extraction and shifting of only some, but not all indexicals in one sentence (see Schlenker, 2017; Quer, 2005).

Davidson (2015) analyses example (3a) (her example (36)) as in (3b) and (4a) (a slightly modified version of her example (37)) as (4b).

- (3) a. Bob was like + GOBBLING GESTURE. (Davidson, 2015: 489)  
b.  $\exists e [\text{agent}(e, \text{bob}) \wedge \text{demonstration}(d_1, e)]$ ,  
with  $\text{demonstration}(d, e)$ :  $d$  is an event that reproduces some contextually salient properties of the event  $e$ ;  $d_1$  is a gobbling gesture.
- (4) a. Bob was like “This isn’t fair.” + WHINEY VOICE  
b.  $\exists e [\text{agent}(e, \text{bob}) \wedge \text{demonstration}(d_1, e)]$ ,  
with  $\text{demonstration}(d, e)$ :  $d$  is an event that reproduces some contextually salient properties of the event  $e$ ;  $d_1$  is the act of saying *This isn’t fair* in a whiney voice.

Davidson takes both the gobbling gesture and the whiney voice imitation to represent demonstrations. These must reproduce certain contextually salient properties of the event they are referring to. What exactly the contextually salient properties are is left for pragmatics. But crucially, what this means is that the demonstrational acts have to be *similar* (with respect to these contextually salient properties) to the original event  $e$ . She gives credit to Zucchi et al. (2011) for proposing a similar analysis of classifier constructions in sign languages. And interestingly, she explicitly argues that one might counter (4a) with (5).

- (5) No, I think he wasn’t whining when he said it wasn’t fair. (cf. Davidson, 2015: 489)

Likewise (3a) can be countered with

- (6) No, he did not gobble it down in such a way.

In (5), the hearer of (4a) does not take issue with the exact wording, but with the manner the words were presented (i.e. the whining). And in (6) it is the gobbling gesture of (3a) that is objected to. In recent literature it has been argued that direct denial as in (5) and (6) can be used as a diagnostic for the detection of at-issue material (Potts, 2005). We consider all material that the speaker or signer wants to present as the main core of her contribution as *at-issue* material, and all other information that comes with this, but is not directly relevant for answering the current question under discussion as *non-at-issue* material. In particular this is all non-asserted, but implicated or presupposed material. We thus note here that the gobbling gesture in (3a) and the presentation mode (whining) in (4a) seem to enter the semantic composition as at-issue material.

Summing up, we have seen that Davidson (2015) argues that action role shift and attitude role shift both involve demonstrations, but for different reasons. While action role shift functions like a classifier predicate in sign languages (and usually involves one), which makes use of an iconic demonstrational act, attitude role shift behaves like quotation, which has been analysed as involving demonstration elsewhere (Clark & Gerrig, 1990). Furthermore, Davidson argues that attitude role shift corresponds to direct quotation in spoken languages and action role shift can be likened to *be-like* constructions.

## 2.2 Novel observations

Taking a closer look at the at-issue/non-at-issue distinction, we would like to add some more data points that Davidson does not discuss in her paper. When we turn the *be-like* construction of (4a) into an ordinary report (as in (7a)), the utterance cannot be countered with (5) anymore, but only with an indirect denial along the lines of (7b).

- (7) a. Bob said: “This isn’t fair.” + WHINEY VOICE  
b. Yes, ok, he said that. But I don’t think he was whining when he said it wasn’t fair.

Likewise, if we add a verb phrase to (3a) and thus change the gobbling gesture plus *be-like* into an ordinary co-speech gesture (aligned with the verb phrase), things change accordingly. While

(6) (repeated as (8b)) is a felicitous reaction to (3a), it is inadequate as a reaction to (8a) and one would have to use (8c) instead.

- (8) a. Bob [ate the sausage]. + GOBBLING GESTURE  
b. #No, he did not gobble it down in such a way.  
c. Yes, he did eat the sausage, but he did not gobble it down in such a way.

This pattern is correctly predicted if we combine the two following assumptions (cf. also Ebert, Ebert & Hörnig, to appear). First, we adopt Ebert & Ebert's (2014) treatment of co-speech gestures and their dimension shifting analysis of demonstratives. On that analysis, demonstratives such as German *so* (Engl.: *so/such/like this*) shift gestural non-at-issue material to the at-issue dimension. Second, we take the act of uttering “*This isn't fair*” in a whiney voice and the gobbling demonstration to be co-speech gestures. Ebert & Ebert (2014) argue that co-speech gestures enter semantics as non-at-issue material by default. As a consequence, these gestural contributions cannot be directly countered (cf. (7b) or (8b), under the assumption that the whiney voice in (7a) is also a co-speech gesture). As opposed to these cases, in (3a) and (4a) a direct denial reaction is possible, because *like* acts as a demonstrative and makes the gesture contribution at-issue.

We build on work by Streeck (2002) who analyses *like* in *and he was like, I'm like* etc. as what he refers to as ‘body quotatives’, i.e. expressions that mark ‘body quotes’. The same analysis applies to German *so* in *und er so, er dann so*, etc. These are, according to Streeck, expressions that serve to make enactments or quotations of some other person's gestures salient. He writes that these words serve to “give these nonverbal behaviors grammatical status, anchoring them in unfolding structures of sentences and assigning them specific semantic-syntactic roles” (Streeck, 2002: 581). Based on these insights from Streeck, we treat English *like* and German *so* in the usages under discussion as similarity demonstratives for propositions or events.

In the following, we will briefly lay out Ebert & Ebert's (2014) proposal for the treatment of speech-accompanying gestures and their view of demonstratives. We will then show how this framework can also be used to account for the observed differences in the information status of the data we have discussed.

### 2.3 Ebert & Ebert (2014)

Ebert & Ebert (2014) argue that gesture meaning is not at issue by default (see also Schlenker, 2018; Esipova, 2018 for recent formal semantic approaches to the semantic contribution of gestures). For illustration purposes, consider an utterance of (9) with a simultaneous iconic gesture shaping an oval object (where underlining indicates co-occurrence of speech and gesture).

- (9) Peter bought a casserole dish <sub>‘oval’ gesture</sub>

The verbal and gestural meaning together communicate that Peter bought a casserole dish and that this casserole dish is oval. The gestural meaning is not at issue. (10) shows that gestural meaning cannot be directly denied ((10a) vs. (10b)<sup>4</sup>) and that it projects across operators such as negation ((10c) vs. (10d)). (10c) is an incoherent piece of discourse, while (10d) with the adjective conveying the shape information is perfectly acceptable.

- (10) a. #That's not true, the casserole dish isn't oval.  
b. Hey, wait a minute, the casserole dish isn't oval.

---

<sup>4</sup> Von Stechow (2004), based on Shanon (1976), uses *hey-wait-a-minute* statements as a tool to test for the presuppositional status of information units, but see Pearson (2010) and Potts (2015) for discussion about the test as a more general tool for diagnosing non-at-issue content and Syrett & Koev (2014) for critical discussion.

c. I would never buy a casserole dish.<sub>‘oval’ gesture</sub>. # Lasagna sheets would not fit in there.

d. I would never buy an oval casserole dish. Lasagna sheets would not fit in there.

As Ebert & Ebert have shown, however, if a demonstrative like German *so* is added, gesture information does become at-issue. In fact, the gesture contribution in (11a) can be denied directly as in (10a). Likewise, the German variant of (10c) with *so* in (11b) renders it synonymous to the initial sentence of (10d).

(11) a. *Peter hat sich so eine Auflaufform*.<sub>‘oval’ gesture</sub> *gekauft*.

‘Peter bought a casserole like that’.<sub>‘oval’ gesture</sub>

b. *Ich würde niemals so eine Auflaufform*.<sub>‘oval’ gesture</sub> *kaufen. Da passen Lasagneblätter nicht rein.*

‘I would never buy a casserole dish like that’.<sub>‘oval’ gesture</sub> Lasagna sheets would not fit in there.’

It has thus been proposed that *so* acts as a ‘dimension shifter’: it shifts non-at-issue meaning to the at-issue dimension. In the case of *so* in (11) we are dealing with a demonstrative operator that applies to an indefinite determiner (*so ein*, Engl.: *such a*). Hence it shifts all gestural contributions associated with the determiner *ein* (Engl.: *a*) (see also Ehlich, 1986; Fricke, 2012; Umbach & Gust, 2014 for semantic analyses of the German similarity demonstrative *so*).

Following Ebert & Ebert (2014), we also assume that there is no crucial difference between iconic and pointing gestures: both refer to an *intended referent* *g*. Formally, a gesture denotes the *rigid designator to the intended referent*. This intended referent *g* relates to the co-occurring speech signal in different ways. Temporal alignment of gesture and speech is taken to be meaningful and results in certain contributions to the semantic representation. Co-occurrence with a determiner expresses a type of similarity with the discourse referent introduced by the determiner. If the determiner is indefinite, it is similarity with respect to some contextually salient feature(s) (cf. Umbach & Gust, 2014 for uses with German *so*), if it is definite, it is strict identity. Co-occurrence with an NP results in an exemplification relation and the gesture has to exemplify the NP-concept (cf. Fricke, 2012; Lücking, 2013). Crucially, all these gestural meaning contributions are non-at-issue. The formal analysis is fleshed out in AnderBois et al.’s (2015) uni-dimensional, dynamic system, which keeps track of the dynamics of discourse referents (i.e. variables over individual concepts) as well as of propositional variables *p* and *p\** used to account for the division of at-issue and non-at-issue meaning, respectively, as indicated by corresponding subscripts on predicates<sup>5</sup>. Given the gesture referent *g*, *g* is the corresponding formal language expression that is interpreted as the individual concept with value *g* for all possible worlds. The formal representation of (9) comes down to (12) (assuming that the intended referent *g* is an oval shaped casserole):<sup>6</sup>

$$(12) [x] \wedge x = \text{PETER} \wedge [y] \wedge \text{CASSEROLE-DISH}_p(y) \wedge [z] \wedge z = \mathbf{g} \wedge \text{SIM}_{p^*}(y, z) \wedge \text{CASSEROLE}_{p^*}(z) \wedge \text{BUY}_p(x, y)$$

Following Farkas & Bruce (2014), we assume that at-issue material is “on the table” and thus up for discussion. It hence needs to be approved by the hearer to be added to the Common

<sup>5</sup> Since one of the main aims of Ebert & Ebert’s (2014) account is to identify discourse referents across different modalities and address properties to it via gesture and speech it is important to adopt a semantic account that takes care of dynamic bindings across dimensions. This makes Anderbois et al.’s (2015) framework to account for non-at-issueness phenomena the perfect approach also for the phenomena we are interested in and excludes other two-dimensional frameworks such as Potts (2005) or Gutzmann (2012), which do not allow modelling of dependencies across dimensions.

<sup>6</sup> In these representations,  $[x]$  stands for a reset of assignments at position *x*. We omit certain details of the formal system that are not essential to the understanding of our main ideas, but note that this essentially comes down to the introduction of a dynamic existential quantifier with bound variable *x* (see Anderbois et al. (2015) for details).

Ground, while non-at-issue material is silently imposed on the Common Ground. The interpretation of (12) then comes down to a proposal to add the information that Peter ( $x$ ) bought a casserole dish ( $y$ ) to the context set, while it imposes onto the context set that (1.) the casserole dish  $y$  is similar to the intended gesture referent  $g$  and (2.) the intended referent is a casserole dish. (1.), the similarity requirement, is a consequence of the temporal alignment of the indefinite article ‘ $a$ ’ and the oval shape gesture and (2.), the exemplification statement, is a consequence of the temporal alignment of gesture and NP.

As pointed out above, German *so* in (11) acts as a dimension shifter. Hence, one derives for (11a) the same interpretation as in (12) with the decisive difference that the similarity introduced via co-occurrence of gesture and indefinite determiner is at-issue (i.e.  $\dots \wedge \text{SIM}_p(y, z) \wedge \dots$ ). The similarity between gesture referent and the verbally introduced casserole then becomes part of the proposal, which explains why the denial in (10a) is coherent as an objection to (11a), but not to (9). For further details and experimental validation of this approach, we refer the reader to Ebert, Ebert & Hörnig (to appear).

## 2.4 Davidson (2015) combined with Ebert & Ebert (2014)

If we now adopt Ebert & Ebert’s (2014) account of co-speech gestures and demonstratives and combine it with Davidson’s (2015) approach to quotations, classifiers, and role shift, the facts in (3)-(11) can be derived straightforwardly. Davidson (2015) assumes that all mentioned constructions involve an act of demonstration. And demonstrations are events that are similar in certain contextually relevant respects to the events they are designed to demonstrate. Davidson furthermore assumes that, in sign languages, a semantic representation involving demonstrational acts is triggered by certain constructions such as classifiers and role shift. At the same time, she argues that in spoken languages, demonstrations are involved in quotations, including *be-like* constructions, and that *be-like* is of a type that expects a demonstration.

We disagree with Davidson only with respect to what we believe triggers the demonstration predicate and its interpretation as part of the semantic representation in spoken languages. While Davidson assumes that it is the lexical semantics of reporting verbs and *like* that introduce a predicate which requires that the demonstrational event “reproduces properties of  $e$  [(i.e. the event to be represented)] and those properties are relevant in the context of speech” (Davidson, 2015: 487), we take this to be a result of gesture-speech interaction. If a gesture co-occurs with speech, this triggers certain interpretational relations that have to hold between spoken and gestured material. As pointed out above, Ebert & Ebert (2014) have argued that indefinites plus gestures trigger a similarity relation, while definites plus gestures trigger identity between gesture and speech referent. We now extend this approach towards the assumption that expressions of type  $\langle s, t \rangle$ , i.e. assertional (or pseudo-assertional) sentences and clauses plus gesture, trigger the semantic contribution that the reported event must be similar to the event that the gesture demonstrates. This is essentially what Davidson considers to be a demonstration. It now follows that if this demonstration is accompanied by a demonstrative (here: the event type similarity demonstrative *like*) the gesture contribution (i.e. that the demonstrated event must be similar to the event that is talked about) is shifted to the at-issue dimension. With these assumptions in place we then predict that there is a similar contribution in (7a) and (8a) (without *like*) as there is in examples (3a) and (4a) (with *like*) from Davidson. But at the same time, it is predicted that these gesture contributions are at-issue in (3a) and (4a) and non-at-issue in (7a) and (8a), which is corroborated by the empirical data.

The gestures are co-speech in (7a) and (8a) and therefore enter the semantic interpretation as non-at-issue material, evidenced by the fact that they cannot be directly countered with *no* (via (5) and (6), respectively). In (3a) and (4a), however, the gobbling gesture and the whining demonstration are preceded by the demonstrative *like*. And this is the reason why the gestures function assertion-like and can be directly countered. Extending Ebert and Ebert’s (2014) account in the way sketched above, we arrive at the representations in (13).

- (13) a. (3a):  $[e] \wedge \text{agent}(e, \text{bob}) \wedge [z] \wedge z = \mathbf{g} \wedge \text{SIM}_p(e, z)$   
 b. (8a):  $[e] \wedge \text{agent}(e, \text{bob}) \wedge \text{ate-sausage}(e) \wedge [z] \wedge z = \mathbf{g} \wedge \text{SIM}_{p^*}(e, z)$   
 (where  $\mathbf{g}$  is the individual concept that stands for a gobbling gesture  $g$ )  
 c. (4a):  $[e] \wedge \text{agent}(e, \text{bob}) \wedge [z] \wedge z = \mathbf{g}' \wedge \text{SIM}_p(e, z)$   
 d. (7a):  $[e] \wedge \text{agent}(e, \text{bob}) \wedge \text{say}(e) \wedge \text{form}_p(e) = \text{“This isn’t fair”} \wedge [z] \wedge z = \mathbf{g}'$   
 $\wedge \text{SIM}_{p^*}(e, z)$   
 (where  $\mathbf{g}'$  is the individual concept that stands for a gestural act  $g'$  of uttering *This isn’t fair* in a whiney voice)

$\text{SIM}(e, z)$  is true iff  $e$  is similar to  $z$  in certain contextually relevant aspects (see Umbach & Gust, 2014 for a three-valued formally spelled out implementation of the SIM predicate).

In (13d), the quotation “*This isn’t fair*” is the complement of the *say* predicate and hence part of the at-issue semantics. At the same time, it is also a substantial part of the gestural act  $g'$ . We understand this as an implementation of the often-made claim that quotation involves an act of description (i.e. restricting  $e$ ) and one of depiction (i.e. gesture-like and restricting  $z$ ) at the same time or that it means using and mentioning the words in one unified act.

To sum up, in Ebert & Ebert’s (2014) account, the similarity condition SIM is conventionally triggered by the temporal alignment of gesture and speech. In the cases at hand, similarity has to hold between two events, while in the cases discussed in Ebert & Ebert (2014), where a gesture is temporally aligned with an indefinite, similarity must hold between two individuals. In other words, Davidson’s (2015) account of *be-like* constructions and quotation can be derived by extending Ebert & Ebert’s (2014) account to event gestures and similarity demonstratives for events. This, we think, yields further independent support for each of the two approaches. In particular, we observe the same at-issue/non-at-issue distinction and instances of shifting here as with other co-speech gestures.

This concludes the first part of the paper. In the second part, we turn to the interaction of perspective-taking at the linguistic level with perspective-taking at the level of gestures. We discuss an empirical study in which we tested a prediction that the analysis argued for in the first part of the paper makes when it is combined with the assumption that FID is a special form of mixed quotation (Maier, 2015, 2017). The prediction is that it should be possible to combine an utterance in FID, just like one in DD, with a pointing gesture to the speaker’s body that is designed to illustrate the viewpoint of the protagonist being (partially) quoted. For ID, in contrast, this should be rather awkward, since there is no quotation involved. Before we discuss that prediction and the rating study used to test it in detail, we will first provide background on FID and on co-speech gestures which illustrate the viewpoint of the individual talked about in the sentence that accompanies the respective gesture.

### 3 FID, Quotation, and Gestures

#### 3.1 Background on FID

FID is widely recognized as a special form of speech or thought representation that combines features of DD with features of ID (Hamburger, 1968; Rauh, 1978; Banfield, 1982; Doron, 1991; Schlenker, 2004; Sharvit, 2008; Eckardt, 2014; Maier, 2015, 2017). Compare the three short text segments in (14a-c).



- (14) a. Peter woke up in the middle of the night. He thought: ‘Tomorrow is Wednesday and I have not completed my damn homework yet’.
- b. Peter woke up in the middle of the night. He thought that the next day WAS Wednesday and that HE HAD not completed HIS damn homework yet.
- c. Peter woke up in the middle of the night. Tomorrow WAS Wednesday and HE HAD not completed HIS damn homework yet.

In all three cases, the second sentence reports that Peter had a thought with identical propositional content. In (14a) and (14b), which are instances of DD and ID, respectively, the propositional attitude verb in combination with the pronoun *he* functioning as its subject explicitly indicates that the sentence following the colon (in (14a)) and the embedded clause (in (14b)) is to be interpreted as the content of a thought by Peter. In the case of (14c), in contrast, this is not directly indicated by overt linguistic material. Rather, it is an inference drawn by the reader on the basis of the content combined with indirect linguistic signals to which we now turn.

Let us consider the underlined words in (14a). *Tomorrow*, *I* and *my* are all context-sensitive expressions that can usually only be interpreted relative to an utterance context (Kaplan, 1989): *I* and *my* refer to the speaker (the author of the context in Kaplan’s terms), while *tomorrow* refers to the day following the day on which the utterance took place. Similarly, the (present) tense markings on the verbs *is* and *have* require an utterance context for their interpretation. Finally, the evaluative expression *damn* can only be interpreted relative to the perspective of some individual, signalling the individual’s negative attitude towards the content of the sentence containing it (Potts, 2005; Gutzmann, 2012). (14a) is an instance of DD, where all context-sensitive expressions contained in the sentence following the colon are interpreted not with respect to the speaker’s or narrator’s context, but rather with respect to parameters of the situation in which Peter had the thought reported by that sentence. Consequently, that situation functions as the context with respect to which all context-sensitive expressions are interpreted. Likewise, the evaluative expression *damn* is interpreted with respect to Peter’s perspective, not the perspective of the speaker or narrator.

In the case of (14b), which is an instance of ID, all pronouns and tenses contained in the embedded clause are interpreted with respect to the speaker’s or narrator’s context (these are indicated by the words in small caps), with third person indicating distinctness from the speaker or narrator (and the addressee) and past tense indicating that the reported events and states precede the utterance or narration time. Concerning the context-sensitive temporal adverb *tomorrow*, it has to be replaced by *the next day* in order to receive the intended interpretation. It is only the evaluative expression *damn* which is interpreted from Peter’s perspective in the case of (14b).

Finally, in the case of (14c), which is an instance of FID, the context-sensitive expressions behave non-uniformly: While *tomorrow* is interpreted with respect to Peter’s context, just like in (14a), pronouns and tenses are interpreted with respect to the speaker’s or narrator’s context, just like in (14b). It is this non-uniform behaviour of context-sensitive expressions which, in combination with the content, gives rise to an interpretation of the second sentence as a thought of Peter rendered in FID: Otherwise, the sentence would be contradictory, due to incompatible requirements of *tomorrow* on the one hand and past tense on the other, since no event or state can be located in the future and in the past with respect to a single context. In the following section, we discuss two analyses that aim at capturing the peculiar behaviour of context-sensitive expressions in FID.

## 3.2 Two analyses of FID

There are two lines of analysis that have been proposed in the formal semantics literature to capture the distinctive properties of FID, which were also touched upon in the preceding section: Double context analyses and the mixed quotation approach. In the following two subsections, we will briefly discuss the two approaches in turn.

### 3.2.1 Double Context Analyses

The mixed behaviour of context-sensitive expressions in FID that we reviewed in Section 3.1 has led Schlenker (2004), Sharvit (2008) and Eckardt (2014) (whose analyses differ in many technical respects; see also Doron, 1991 for an early implementation of a similar idea in the framework of situation semantics) to postulate that sentences in FID are interpreted with respect to two different contexts. The following exposition is based on Eckardt's (2014) implementation of that basic assumption. Eckardt (2014) assumes that in narrative texts a protagonist's context  $c$  can optionally be introduced in addition to the narrator's context  $C$ , which corresponds to the speaker's context in oral communication. This second context  $c$  is implicitly introduced by the sentences preceding a sentence or stretch of discourse in FID mode. It consists of the protagonist functioning as the author of the context and the temporal and spatial location of that protagonist at the reference time of the ongoing story.

It is now lexically specified for each context-sensitive expression whether it can only be interpreted with respect to the speaker's or narrator's context  $C$ , or if it has to be interpreted with respect to a protagonist's context  $c$  whenever such a context is introduced (see Hinterwimmer, 2019; Hinterwimmer & Meuser, 2019 for detailed discussion of the conditions under which protagonists become available as authors of such an additional context  $c$ ). In the absence of  $c$  – i.e. whenever a sentence is not interpreted as FID – the respective context-sensitive expression is interpreted with respect to  $C$  by default.

The inconsistent behaviour of context-sensitive expressions in FID can now be captured straightforwardly: For pronouns and verbal tenses, it is lexically specified that they can only be interpreted with respect to  $C$ , while all other context sensitive expressions always have to be interpreted with respect to  $c$  whenever  $c$  is present. Additionally, Eckardt (2014) assumes that whenever a sentence is interpreted with respect to both  $C$  and  $c$ , the proposition it denotes is not added to the set of propositions characterizing the fictional worlds of the story directly. Rather, what is added is the proposition that the author of  $c$  believes the proposition denoted by that sentence. This latter assumption ensures that sentences or stretches of discourse in FID mode represent the beliefs of some prominent protagonist, not the narrator's beliefs. The second sentence in (14c) (repeated here as (15)) is thus interpreted as the content of a thought that Peter has at the time of his waking up in the middle of the night. While *tomorrow* is interpreted as referring to the day following the day of  $c$ , where the temporal parameter of  $c$  is the time of Peter's waking up, the pronouns *he* and *his* as well as the past tense markings of the verbs are interpreted with respect to the speaker's or narrator's context  $C$ .

(15) Peter woke up in the middle of the night. Tomorrow WAS Wednesday and HE HAD not completed HIS damn homework yet.

### 3.2.2 The Mixed Quotation Approach

Maier (2015, 2017; see also Dirscherl & Pafel, 2017) argues for a fundamentally different approach to FID. In his view, FID is a special, highly conventionalized form of mixed quotation: It consists of quotes of thoughts or utterances (just like in DD) that contain unquoted parts – namely pronouns and tenses. In contrast to mixed quotation as it is found in newspaper articles, for instance, where the quoted parts are typographically marked by quotation marks or italics, there is no typographical distinction between quoted and unquoted parts in FID.

The text segment in (15) accordingly corresponds to the schematic representation in (16a), where the second sentence is interpreted as a partial quotation of a thought that Peter has at the time of the event introduced by the first sentence. In (16b), the fully quoted, explicitly introduced version of the same thought that was already given in (14a) above is repeated for comparison.

- (16) a. Peter woke up in the middle of the night. ‘Tomorrow’ was ‘Wednesday and’ he had ‘not completed’ his ‘damn homework yet’.  
 b. Peter woke up in the middle of the night. He thought: ‘Tomorrow is Wednesday and I have not completed my damn homework yet’.

Maier (2015, 2017) assumes that a pragmatically driven convention that is also in effect in other forms of mixed quotation is responsible for the systematic unquotation of pronouns and tenses, as opposed to all other context-sensitive expressions, in FID. Note that in spite of the very different technical mechanisms by which they come about, the resulting interpretations in Maier’s analysis are very similar to those assumed by the double context analyses as far as context-sensitive items are concerned: The quoted context-sensitive items ultimately are interpreted with respect to the context of the protagonist whose utterance or thought is being partially quoted, while the unquoted pronouns and tenses are interpreted with respect to the narrator’s context.

Maier (2015, 2017) gives the following argument in favour of his analysis, however: FID involves more than a shifted interpretation of context-sensitive items. Rather, there are also cases in which a protagonist’s thoughts or utterances are rendered in the non-standard dialect spoken by that protagonist while the surrounding text is written in standard language. Such cases are unproblematic if FID is assumed to be mixed quotation. It is unclear, however, how they are to be captured by double context analyses.

### 3.3 Viewpoint Gestures

Research on co-speech gestures, i.e. gestures that speakers produce while describing events or scenes to their interlocuters, has shown that such gestures can reveal perspectives (McNeill, 1992; Parrill, 2010; Stec, 2012). There is one type of gesture, dubbed *character viewpoint gesture* (C-VPT), in which the speaker impersonates an individual participating in the event described by the sentence and enacts the event from that person’s point of view. In another type of gesture, dubbed *observer viewpoint gesture* (O-VPT), the speaker depicts the event described by the sentence as if it was observed from a distance. While C-VPTs often involve the speaker’s entire body and face, O-VPTs are usually performed with the hands exclusively. To see the difference between the two types of gestures consider the pictures in Figure 2 and Figure 3 (from Parrill, 2010), which show participants performing gestures while reporting an event of a skunk hopping across the room that they had seen in a cartoon (shown in Figure 1a).



Figure 1a: Original Scene

Figure 1b: O-VPT

Figure 1c: C-VPT

The gesture in Figure 1b in which the hand represents the skunk and the trajectory of the hand represents the path of the skunk is clearly an O-VPT since the event is depicted as if observed

from a distance. The gesture in Figure 1c, in contrast, in which the speaker enacts the hopping skunk, is clearly a C-VPT.

In the following section, we will discuss an experiment in which we look at the interaction of FID as an instance of perspective-taking at the linguistic level with perspective-taking conveyed by C-VPTs, based on the assumptions that FID is a form of mixed quotation (Maier, 2015, 2017) and that quotation can involve a demonstrational act which can be acted out as a gesture (Davidson, 2015). We consider situations where a speaker quotes another person and points to her own (= the speaker's) body while reporting the person's words or thoughts about herself (= the quoted person). This pointing gesture is clearly an instantiation of a C-VPT gesture as detailed above, since the speaker used her own body as standing proxy for the person she is quoting, which is used as a demonstration. In DD when the speaker uses the first-person pronoun 'I', it is expected that pointing to the speaker's body while uttering 'I' should be natural. In ID, however, which does not involve a demonstrational act and where the speaker uses a third person pronoun to refer to the person whose utterance she is reporting, pointing to oneself while uttering a third person pronoun should be odd. With regards to FID, matters are not as obvious. If FID involves quotation, as Maier (2015, 2017) claims, and if Davidson (2015) is correct that quotation involves a demonstrational act, it should be possible for a speaker to use C-VPT gestures to accompany FID. In particular, it is predicted that it should be possible to point to the speaker's body as a demonstrational act to refer to the reported character. Since in FID pronouns are interpreted with respect to the speaker's and not the protagonist's context, i.e. a third person pronoun will be used to refer to the protagonist, we expect that the use of a third person pronoun plus self-pointing should be legitimate in FID. In the following section, we will present an experiment that tests for this hypothesis.

#### 4 Experimental Study: Gestures in FID

Following Davidson (2015) in the assumption that quotation makes use of demonstrations, we expect that DD allows for self-pointing gestures accompanying the utterance of a first-person pronoun. If we also follow Maier (2015, 2017) and assume that FID involves quotation, we predict that FID, as in (17), can be accompanied by a self-pointing gesture, while this should not be possible in ID, as illustrated in (18), which do not involve direct quotation.

(17) Paul was annoyed. Why had [he] not been invited to Clara's stupid party?



(18) ??Paul asked himself why [he] had not been invited to Clara's stupid party.



In other words, the pattern in (17) and (18) is expected if the following holds: First, FID is a particular form of mixed quotation, with pronouns and tenses being unquoted, as argued for by Maier (2015, 2017). Second, it involves a perspective shift towards the reported speaker or thinker. Third, as it instantiates a form of quotation, it involves a demonstrational act, which can be acted out by an additional gesture (as argued for by Davidson, 2015 for direct quotes, see Section 2). Pointing to one's own body (plus potential face expressions, etc.) represents such a demonstrational act, i.e. a C-VPT gesture, which might serve to characterize the reported speaker or to signal that the speaker's body now acts as if it was the reported speaker's body. This does not necessarily imply that the (partially) quoted speech act has to have involved such a pointing gesture. In order to obtain empirical evidence for these assumptions, we conducted an experimental study in which we tested the following hypothesis: A self-pointing gesture to one's own body that is performed while uttering a first-person pronoun in DD should be rated significantly better than a self-pointing gesture that accompanies a third person pronoun in an

ID construction. Crucially, concerning FID, we predict that the use of a third person pronoun which refers to the person whose thoughts or utterances are reported can be accompanied by a self-pointing gesture and that such a combination is rated significantly better in FID than in ID.

#### 4.1 Method

In the study, participants (18 native speakers of German) were shown short video sequences where a speaker uttered a text segment consisting of two sentences. Each of the experimental items ( $n = 18$ ) came in three different conditions of the three-level factor MODE, where the opening sentence, which was identical in all conditions, introduced a topical referent. The second sentence always reported a thought of that protagonist. While the content of that thought was identical in all three conditions, it was rendered as FID in the first condition (19a), as ID in the second condition (19b) and as DD in the third condition (19c). In each condition, the speaker pointed at himself while uttering the pronoun referring to the topical referent in the clause reporting that referent's thought. In order to make the pointing as natural as possible, the items were constructed in such a way that the respective pronoun received a contrastive focal accent. An example is given in (19), where capitals indicate focal stress.

(19) a. **FID**

*Leona war stinksauer. Jetzt hatte schon wieder SIE die Rechnung*  
 Leona was extremely annoyed. Now had yet again she the bill.ACC  
*für die gesamte Gruppe übernommen.*  
 for the entire group.ACC paid.  
 'Leona was extremely annoyed. Again, SHE had paid the bill for the entire group.'

b. **ID**

*Leona war stinksauer. Sie ärgerte sich, dass schon wieder SIE*  
 Leona was extremely annoyed. She was angry self that yet again she  
*die Rechnung für die gesamte Gruppe übernommen hatte.*  
 the bill.ACC for the entire group.ACC paid had.  
 'Leona was extremely annoyed. She was angry that SHE had paid the bill for the entire group again.'

c. **DD**

*Leona war stinksauer. Verärgert dachte sie: "Jetzt habe schon wieder*  
 Leona was extremely annoyed. Angrily thought she: "Now have yet again  
*ICH die Rechnung für die gesamte Gruppe übernommen."*  
 I the bill.ACC for the entire group.ACC paid."  
 'Leona was extremely annoyed. Angrily, she thought: "Now, I have paid the bill for the entire group again."'

The experimental items were interspersed with 24 fillers, which did not involve self-pointing, but rather pointing to a location in the gesture space. Discourse referents were introduced by referential expressions accompanied by pointing gestures that anchored the referents in the gesture space. (e.g. *Gestern auf der Party hat Peter Linus beleidigt*. Engl.: *Yesterday at the party, Peter insulted Linus.*, plus pointing to a point left in the central gesture space in front of the speaker's body when uttering *Peter* and to a point right when uttering *Linus*.) These referents were then picked up via a combination of demonstrative pronouns (D-pronouns) and pointing gestures in a second sentence (*Der hat dann sofort angefangen zu weinen*. Engl.: *He (D-pronoun) then immediately started crying.*, plus pointing back to one of the previously introduced locations, while uttering *der*.) The fillers were constructed in such a way that pointing was correct and natural in one third of the cases: The gender of the pronoun in the second sentence matched the gender of the antecedent that was introduced in the location that the speaker pointed to while using the pronoun, and the demonstrative pronoun picked up the referent introduced by an object DP, which is the highly preferred option for D-pronouns as opposed to personal pronouns, which prefer subject antecedents (see, e.g. Bosch et al., 2007). In one third

of the cases they were incorrect: Gender did not match, i.e. the pointing gesture that accompanied the d-pronoun was directed to the location where an antecedent with mismatching gender features was introduced. (e.g. *Gestern auf der Party hat Peter Paula beleidigt. Die hat dann sofort angefangen zu weinen.* Engl.: *Yesterday at the party, Peter insulted Paula. She (D-pronoun) then immediately started crying.*, plus pointing to Peter's location when *die* was uttered.) In the final third of the cases, they were correct, but unnatural: There was no clash concerning gender information, but the demonstrative pronoun picked up the referent introduced by a subject DP.

Participants (who were told beforehand to pay attention to picture and sound) rated all videos on a scale from 1 to 10, where 1 corresponds to completely unnatural and 10 to completely natural.

## 4.2 Results

We excluded three participants from the analyses, who gave ratings for the fillers which were exactly the opposite of the expected ratings (probably due to scale reversal). Hence, our analysis is based on the judgements of 15 participants. The means for the three conditions were as follows (standard deviation in parentheses): FID: 7,18 (sd: 2,52), ID: 6,73 (sd: 2,65), and DD: 7,77 (sd: 1,63). The three-level factor MODE exerted a significant influence on perceived naturalness [ $F1(2,28) = 3.51, p < .05, \eta^2 = .200$ ;  $F2(2,34) = 3.93, p < .05, \eta^2 = .188$ ]. We specified two orthogonal contrasts such that condition ID was compared to both, FID and DD, with the first contrast addressing our main hypothesis.<sup>7</sup> ID was judged significantly less natural than DD ( $\bar{x} = 7.77$ ) in both analyses [ $F1(1,14) = 4.764, p = .047, \eta^2 = .254$ ;  $F2(1,17) = 5.498, p = .031, \eta^2 = .244$ ], as expected. Furthermore, our main hypothesis was confirmed only partly. ID ( $\bar{x} = 6.73$ ) was judged less natural than FID ( $\bar{x} = 7.18$ ), yet this difference is significant only in the  $F1$  analysis but not in the  $F2$  analysis [ $F1(1,14) = 5.849, p = .030, \eta^2 = .295$ ;  $F2(1,17) = 1.986, p = .177, \eta^2 = .105$ ].

## 4.3 Discussion

In our study we aimed to show that self-pointing gestures combine well for a speaker who utters a first person pronoun in a direct discourse although the speaker is not the originator of the content of the direct discourse. In case of an indirect discourse with the same content, in contrast, self-pointing gestures do not combine well with the third person pronoun uttered by the speaker. Our study established this difference. Our main prediction, however, addressed free indirect discourse. We argued that the third person pronoun in FID, compared to ID, does not stand in a similarly severe conflict with self-pointing gestures. The corresponding difference in naturalness ratings was significant in the  $F1$  analysis but not in the  $F2$  analysis, i.e., the analyses support a generalization to the population of speakers but not to the population of items. This inconclusive evidence might be partly attributed to the insufficient statistical power of the relatively small samples of participants and items. Moreover, judgements in the ID condition were surprisingly good.

We interpret this as follows: Participants tend to apply a perspective shifting strategy also in ID, contrary to what we had hypothesized beforehand. ID normally involves no perspective shift, which is indicated by the fact that for context-sensitive expressions contained in embedded clauses in ID in languages such as English and German (but see Schlenker, 2004 and Anand & Nevins, 2004 for evidence that languages such as Amharic and Zazaki behave differently in this respect), there is a tendency to interpret them with respect to the speaker's or narrator's context rather than the context of the attitude holder. Consequently, we expected that self-pointing cannot be interpreted as a C-VPT in the case of ID (or at least not as readily as in FID

---

<sup>7</sup> We used the contrast REPEATED in SPSS for the comparison of adjacent factor levels, FID, ID, DD.

constructions). However, Plank (1986) has shown that temporal and spatial expressions in German may in principle also be interpreted with respect to the attitude holder's context (see also Anderson, 2019 for similar conclusions with respect to *tomorrow*). If we assume that shifted interpretations of context-sensitive expressions in FID come about via partial quotation, it is rather natural to assume that the same strategy is in principle available in ID as well (cf. Maier, 2016), although it is dispreferred in comparison to FID. Consequently, the rather good ratings for the test items in the ID condition could be the result of the test persons applying that strategy to ID, albeit somewhat more reluctantly. Apart from this, the self-pointing gestures in our experiment were not contrasted with, for instance, 3<sup>rd</sup>-party pointing gestures, i.e., gestures to some other (imaginary) person in the gesture space. Including both kinds of gesture may raise the sensitivity of the participants about the appropriateness of either kind of gesture.

To summarize the experimental finding for FID, we consider the inconclusive evidence on our hypothesis as tentative/preliminary. Self-pointing gestures paired with FID pattern neither clearly with ID nor with DD. In a next step, we will compare self-pointing gestures with 3<sup>rd</sup>-party pointing gestures with larger samples of participants and items<sup>8</sup>.

## Acknowledgments

The research reported in this paper was funded by the DFG within the XPRAG.de program (first author) and within the SFB 1252 *Prominence in Language* (second author). We would like to thank Theresa Stender, Robert Müller, Janne Schmandt and Felix Jüstel for help with item generation and execution of the rating study reported in Section 4. We are particularly grateful to Robin Hörnig for his help with the experimental design and the statistical analysis as well as his input on earlier versions and our discussion of the results. Finally, we would like to thank the audience at *Linguistic Evidence 2020* (University of Tübingen) for questions and comments.

## References

- Anand, P., & Nevins, A. (2004). Shifty operators in changing contexts. In R. B. Young (Ed.), *Proceedings of Semantics and Linguistic Theory (SALT) XIV* (pp. 20-37). Ithaca, New York: Cornell University.
- AnderBois, S., Brasoveanu, A., & Henderson, R. (2013). At-issue proposals and appositive impositions in discourse. *Journal of Semantics*, 32(1), 93-138.
- Anderson, C. J. (2019). Tomorrow isn't always a day away. In M. T. Espinal, E. Castroviejo, M. Leonetti, L. McNally & C. Real-Puigdollers (Eds.), *Proceedings of Sinn und Bedeutung 23* (pp. 37-56). Konstanz: University of Konstanz. <https://ojs.ub.uni-konstanz.de/sub/index.php/sub/article/download/493/418>
- Banfield, A. (1982). *Unspeakable Sentences: Narration and Representation in the Language of Fiction*. Boston: Routledge.
- Bosch, P., Katz, G., & Umbach, C. (2007). The non-subject bias of German demonstrative pronouns. In M. Schwarz-Friesel, M. Consten & M. Knees (Eds.), *Anaphors in Text* (pp.145-164). Amsterdam/Philadelphia: John Benjamins.
- Clark, H. H., & Gerrig, R. J. (1990). Quotations as demonstrations. *Language*, 66(4), 764-805.

---

<sup>8</sup> To address this problem, we plan a follow-up study with two three-level factors SELF-POINTING and OTHER-POINTING and the same levels, FID, ID, and DD, as in the experiment reported in this paper. We expect self-pointings to be rated significantly worse than pointing to a third person for ID constructions in this follow-up experiment.

- Davidson, K. (2015). Quotation, demonstration and iconicity. *Linguistics & Philosophy*, 38(6), 477-520.
- Dirscherl, F., & Pafel, J. (2015). Die vier Arten der Rede- und Gedankendarstellung. Zwischen Zitieren und Referieren. *Linguistische Berichte*, 241, 3-47.
- Doron, E. (1991). Point of view as a factor of content. In S. Moore & A. Z. Wyner (Eds.), *Proceedings of Semantics and Linguistic Theory (SALT) I* (pp. 51-64). Ithaca, New York: Cornell University.
- Ebert, C., & Ebert, Ch. (2014). Gestures, demonstratives, and the attributive/referential distinction. Talk presented at the conference 'Semantics and Philosophy in Europe 7', 26-28 June 2014, ZAS, Berlin.
- Ebert, Ch., Ebert, C., & Hörnig, R. (to appear). Demonstratives as dimension shifters. *Proceedings of Sinn und Bedeutung 24*. Osnabrück: University of Osnabrück.
- Eckardt, R. (2014). *The Semantics of Free Indirect Discourse. How Texts Allow to Mind-Read and Eavesdrop*. Leiden: Brill.
- Ehlich, K. (1986). so-Überlegungen zum Verhältnis sprachlicher Formen und sprachlichen Handelns, allgemein und an einem widerspenstigen Beispiel. In I. Rosengren (Ed.), *Sprache und Pragmatik, Lunder germanistische Forschungen, Vol. 55* (pp. 279-298). Stockholm: Almqvist & Wiksell International.
- Esipova, M. (2018). Focus on what's not at issue: Gestures, presuppositions, appositives under contrastive focus. In U. Sauerland & S. Solt (Eds.), *Proceedings of Sinn und Bedeutung 22* (pp. 385-402). Berlin: ZAS.
- Farkas, D. F., & Bruce, K. B. (2010). On reacting to assertions and polar questions. *Journal of Semantics*, 27(1), 81-118.
- von Fintel, K. (2004). Would you believe it? The King of France is back! Presuppositions and truth-value intuitions. In M. Reimer & A. Bezuidenhout (Eds.), *Descriptions and Beyond* (pp. 315-341). Oxford: Oxford University Press.
- Fricke, E. (2012). *Grammatik multimodal: Wie Wörter und Gesten zusammenwirken*. Berlin/Boston: Walter de Gruyter.
- Gutzmann, D. (2012). *Use-Conditional Meaning: Studies in Multidimensional Semantics*. Oxford: Oxford University Press.
- Hamburger, K. (1968). *Die Logik der Dichtung*. Stuttgart: Ernst Klett Verlag.
- Hinterwimmer, S. (2019). Prominent protagonists. *Journal of Pragmatics*, 154, 79-91.
- Hinterwimmer, S., & Meuser, S. (2019). Erlebte Rede und Protagonistenprominenz. In S. Engelberg, C. Fortmann & I. Rapp (Eds.), *Rede- und Gedankenwiedergabe in narrativen Strukturen – Ambiguitäten und Varianz, Linguistische Berichte, Sonderheft 27* (pp. 177-200). Hamburg: Helmut Buske Verlag.
- Kaplan, D. (1989). Demonstratives. In J. Almog, J. Perry & H. Wettstein (Eds.), *Themes from Kaplan* (pp. 565-614). Oxford: Oxford University Press.
- Lücking, A. (2013). *Ikonische Gesten. Grundzüge einer linguistischen Theorie*. Berlin/Boston: Walter de Gruyter.
- Maier, E. (2015). Quotation and unquotation in free indirect discourse. *Mind & Language*, 30(3), 345-373.
- Maier, E. (2016). A plea against monsters. *Grazer Philosophische Studien*, 93(3), 363-395.
- Maier, E. (2017). The pragmatics of attraction. Explaining unquotation in direct and free indirect discourse. In P. Saka & M. Johnson (Eds.), *The Semantics and Pragmatics of Quotation* (pp. 259-280). Dordrecht: Springer.
- McNeill, D. (1992). *Hand and Mind. What Gestures Reveal About Thought*. Chicago: University of Chicago Press.



- Parrill, F. (2010). Viewpoint in speech-gesture integration: Linguistic structure, discourse structure, and event structure. *Language and Cognitive Processes*, 25(5), 650-668.
- Pearson, H. (2010). A modification of the “Hey, Wait a Minute” test. *Snippets*, 22, 7-8.
- Plank, F. (1986). Über den Personenwechsel und den anderer deiktischer Kategorien in der wiedergegebenen Rede. *Zeitschrift für germanistische Linguistik*, 14(3), 284-308.
- Potts, C. (2005). *The Logic of Conventional Implicatures*. Oxford: Oxford University Press.
- Potts, C. (2015). Presupposition and implicature. In S. Lappin & C. Fox (Eds.), *The Handbook of Contemporary Semantic Theory* (pp. 168-202). Hoboken, New Jersey/Oxford: Wiley-Blackwell Publishing.
- Quer, J. (2005). Context shift and indexical variables in sign languages. In E. Georgala & J. Howell (Eds.), *Proceedings of Semantics and Linguistic Theory (SALT) XV* (pp. 152-168). Ithaca, New York: Cornell University.
- Rauh, G. (1978). *Linguistische Beschreibung deiktischer Komplexität in narrativen Texten*. Tübingen: Narr Verlag.
- Schlenker, P. (2004). Context of thought and context of utterance: A note on free indirect discourse and the historical present. *Mind & Language*, 19(3), 279-304.
- Schlenker, P. (2017). Super monsters I: Attitude and Action Role Shift in sign language. *Semantics & Pragmatics*, 10(9), 1-65.
- Schlenker, P. (2018). Gesture projection and cosuppositions. *Linguistics & Philosophy*, 41(3), 295-365.
- Shanon, B. (1976). On the two kinds of presupposition in natural language. *Foundations of Language*, 14(2), 247-249.
- Sharvit, Y. (2008). The puzzle of free indirect discourse. *Linguistics & Philosophy*, 31(3), 353-395.
- Stec, K. (2012). Meaningful shifts: A review of viewpoint markers in co-speech gesture and sign language. *Gesture*, 12(3), 327-360.
- Streeck, J. (2002). Grammars, words, and embodied meanings: On the evolution and uses of *so* and *like*. *Journal of Communication*, 52(3), 581-596.
- Syrett, K., & Koev, T. (2014). Experimental evidence for the truth conditional contribution and shifting information status of appositives. *Journal of Semantics*, 32(3), 525-577.
- Umbach, C., & Gust, H. (2014). Similarity demonstratives. *Lingua*, 149, 74-93.
- Zucchi, S., Cecchetto C., & Geraci, C. (2011). Event descriptions and classifier predicates in sign languages. Talk presented at the conference ‘Formal and Experimental Advances in Sign Language Theory (FEAST)’, 20-22 June 2011, Università Ca’ Foscari.