Indirect discourse as mixed quotation? An experimental investigation

Sebastian Walter*

Abstract. The results of an experimental rating study are reported suggesting that *self-pointing gestures* aligned with a third-person pronoun are acceptable in German *indirect discourse* (ID) utterances. Following a proposal by Ebert & Hinterwimmer (2022) for self-pointing gestures in *free indirect discourse* (FID), self-pointing gestures in ID are interpreted as *character viewpoint gestures* (CVGs) quoted from the matrix subject. Crucially, it is argued that in ID, a perspective shift to the matrix subject can take place. It is proposed that ID is an instance of *mixed quotation* involving a demonstration (cf. Clark & Gerrig 1990, Davidson 2015) where self-pointing is quoted from the matrix subject's original utterance.

Keywords. Perspective taking; gesture semantics; indirect discourse; mixed quotation

1. Introduction. Speakers have several options to report what someone else thought or said. For example, they can do so by giving a verbatim repetition of the reported speaker's thought or speech, i.e., they directly quote the words the other person uttered. This is commonly referred to as *direct discourse* (DD). Alternatively, they can simply paraphrase the propositional content of the reported speaker's original thought or utterance. This is called *indirect discourse* (ID) and is standardly assumed to be non-quotational. A third option is what is called *free indirect discourse* (FID), a way to report someone's thoughts or speech without any overt marking, which has been argued to be an instance of highly specialized mixed quotation (Maier 2015). Particularly interesting is the behavior of so-called *indexicals* in speech reports, an example being the first-person pronoun *I*. Broadly speaking, the reference of an indexical can vary from context to context, which sets them apart from ordinary expressions. Coming back to the first-person pronoun *I*, this means that it always refers to the current speaker. In DD utterances, however, it receives a so-called shifted interpretation, that is, it is not interpreted from the current speaker's (= reporting speaker's) point of view, but rather from the reported speaker's point of view, as first-person pronouns in DD utterances always refer to the reported speaker. In general, all indexicals shift in DD utterances. As will be shown below, the picture is more complex for ID and FID because there, not all indexicals have to shift.

Clark & Gerrig (1990) propose that quotations are demonstrational, meaning that they are depictive rather than descriptive. Furthermore, in this approach not only words can be quoted, but, for example, also gestures (including facial expressions), intonation, and even non-linguistic behavior. These ideas have been formalized by Davidson (2015). In a study investigating Maier's (2015) claim that FID is an instance of mixed quotation with pronouns and tenses being systematically unquoted, Ebert & Hinterwimmer (2022) found that a self-pointing gesture (= a pointing gesture

*Acknowledgments: This research was conducted in the DFG-funded project *Visual and non-visual means of perspective taking in language*, which is part of the Priority Program 2392 (*ViCom*). I thankfully acknowledge the financial support of the German Research Foundation (DFG). Furthermore, I would like to thank Cornelia Ebert and Stefan Hinterwimmer for their valuable comments on this work. Finally, I thank Lennart Fritzsche for his help with conducting the study reported in this paper. Authors: Sebastian Walter, Goethe University Frankfurt (s.walter@em.uni-frankfurt.de).

toward the speaker's body) co-referent with the reported speaker is acceptable in FID. They found the same for first-person pronouns in DD utterances. By unifying Davidson's (2015) account of quotation as demonstration with Ebert et al.'s (2020) approach to gesture semantics, they analyze the results as follows: self-pointing gestures are interpreted as gestures quoted from the reported speaker. Based on the assumption that ID is non-quotational, they hypothesized that self-pointing in ID is unacceptable. However, they found that self-pointing was also surprisingly acceptable in ID utterances. According to them, this suggests that a similar perspective-shifting strategy as in FID is available in ID as well. However, it seems to be more constrained and is not necessarily present. In other words, ID allows for the option of mixed quotation to be present. The study reported in this paper further investigates this claim. Previous research has shown that temporal and local adverbials, which are also indexicals, can shift in ID only if the reported speaker's perspective is prominent in the surrounding discourse (Plank 1986, Anderson 2019). Based on these findings, it was hypothesized that self-pointing is acceptable in ID only if the perspective of the reported speaker (= the matrix subject) is prominent, or, in other words, that mixed quotation is available in ID if this perspective is prominent. To test this, a rating study was conducted. The results go beyond the initial hypothesis, since they show that self-pointing in ID is acceptable regardless of whether the reported speaker's perspective is prominent on the speech level or not, thus suggesting that ID in general allows for mixed quotation.

This paper is structured as follows: Section 2 provides the relevant theoretical background on perspective in speech (Section 2.1), speech-accompanying gestures (Section 2.2), and quotation (Section 2.3). The experimental study is reported in Section 3. Section 4 concludes the paper.

2. Background.

2.1. PERSPECTIVE IN SPEECH. Many expressions depend on a perspective in their interpretation. Among these expressions are predicates of personal taste (*(be) tasty*), epithets (*that idiot*), relational expressions (*this*, *that*, *left*, *right*), and evaluative expressions (*fun*). The default for them is to be interpreted from the perspective of the current speaker (Harris 2012). However, this default changes in instances of speech reports. As will be shown below, depending on the type of speech report, they can receive what is commonly referred to as a *shifted interpretation*, i.e., an interpretation not from the current speaker's perspective, but from the perspective of the speaker whose thoughts or speech is currently being reported. Examples of different types of speech reports are given in (1).

(1) On her way home, Mary heard a song by Kendrick Lamar that she liked on the radio.
   a. She thought: "I will buy his new album tomorrow."
   b. She thought that she would buy his new album on the following day.
   c. She would buy his new album tomorrow.

(Hinterwimmer 2017, p. 284)

The utterance given in (1a) is an instance of DD. In DD, the words or thoughts of the reported speaker are directly quoted. Therefore, a complete perspective shift toward the reported speaker takes place and all perspective-dependent expressions are evaluated from their perspective. Thus, the first-person personal pronoun *I* refers to Mary in (1a) rather than to the current speaker. Moreover, temporal adverbials, such as *tomorrow* in (1a), are also evaluated in the context where the

reported speech event took place, meaning that *tomorrow* interpreted as referring to the day following Mary's thought.

In (1b), where the same thought as in (1a) is rendered as an utterance in ID, perspective-dependent expressions behave differently than in DD. ID is generally analyzed as a non-quotational report of some other individual's speech or thoughts. Consequently, the systematic shifting behavior of perspective-dependent expressions attested for DD does not fully generalize to ID. For example, personal pronouns (at least in German, English, and many other languages) never shift in ID (but see, e.g., Anand & Nevins 2004 for data from other languages where personal pronouns shift in ID). This is illustrated by (1b), where a third-person pronoun is needed to refer to Mary. Using a first-person pronoun would result in a different interpretation of the sentence.

By contrast, predicates of personal taste and evaluative expressions receive a shifted interpretation in ID. Temporal and local deictic expressions are normally interpreted from the reporting speaker's perspective, thus explaining why *on the following day* is used to refer to the day following the thinking event in (1b). However, these expressions can also receive a shifted interpretation if the reported speaker's perspective (i.e., Mary's in (1b)) is prominent in the discourse (Plank 1986, Anderson 2019).

Finally, the utterance in (1c) is an instance of FID—a way to report another individual's thoughts or speech without overt marking. In FID, all perspective-dependent expressions shift toward the reported speaker, the only exceptions being pronouns and tense markings. FID thus shares properties with both ID and DD and is often seen as a blend of the two. In formal semantics, the most popular treatment of FID is the so-called double context analysis (e.g., Schlenker 2004, Eckardt 2014). There is, however, also a proposal by Maier (2015), which treats FID as an instance of highly specialized mixed quotation, where a FID utterance is seen as a verbatim quote of an individual's thoughts or speech with pronouns and tense markings being systematically unquoted. A similar proposal will be made for ID in this paper.

2.2. SPEECH-ACCOMPANYING GESTURES. Gestures are defined as communicative movements of the hands, arms, and even the face which transport a speaker's thoughts, intentions, or emotions. Therefore, gestures that occur alongside speech add meaning to an utterance, or, more precisely, speech and gesture work together to convey a multimodal message (e.g., McNeill 1992, Kendon 2004, de Ruiter 2000, Ebert 2024). The following examples illustrate this:

(2) a. I brought <u>a bottle of water</u> to the talk. + BIG[1]
b. I brought <u>this bottle</u> to the talk. + POINTING TO BOTTLE

The iconic BIG-gesture in (2a) contributes size information of the bottle to the verbal part of the utterance, i.e., that it is big. The pointing gesture in (2b), by contrast, selects its referent directly and adds to the verbal part of the utterance that the bottle pointed at is the bottle the speaker brought to the talk. Crucially, in the semantic theory modeling the meaning contributions of speech-accompanying gestures by Ebert et al. (2020), it is argued that this is the only difference between pointing and co-speech gestures: while iconic gestures refer to an abstract individual, pointing gestures make reference to an individual in the world. Besides this, they behave alike

[1]Underlining in examples indicates gesture-speech alignment, small caps are used to indicate a speech-accompanying gesture.

from a semantic point of view in Ebert et al.'s (2020) account and are therefore both analyzed as making meaning contributions similar to that of appositives (cf. Potts 2005). They thus, unless aligned with a demonstrative as in (2b), make not-at-issue contributions by default (for experimental validation, see Ebert et al. 2020), meaning that they cannot be directly denied in discourse and project through semantic operators such as negation. Especially this first property has often been ascribed to not-at-issue content not being on the table for discussion (Farkas & Bruce 2010). The short discourse in (3) illustrates this.

(3) A: There is a window in Peter's living room. + ROUND
B_1: # That's not true! The window is not round.
B_2: Yes, true, but the window isn't round.

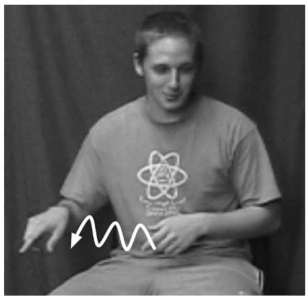
When trying to directly object the gesture's contribution of A's utterance as B_1 does, this results in infelicity. What one can do instead to target the not-at-issue content conveyed by the gesture is, for example, to assent with the main clause proposition followed by an adversative continuation (cf. diagnostic #1c in Tonhauser 2012). This explains why the response of B_2 is felicitous.

It has been shown in Section 2.1 already that perspective plays a decisive role when it comes to the interpretation of many lexical expressions. However, perspective can also be encoded in gesture (McNeill 1992), the most important distinction being the one between so-called *character viewpoint gestures* (CVGs) and *observer viewpoint gestures* (OVGs). When performing a CVG, the whole body is often involved in the gesture production as it depicts an event from an internal, first-person perspective. This contrasts with OVGs where normally only hands and arms are involved in the production of the gesture as it depicts an event from an outside, third-person perspective. The examples given in Figure 2a and Figure 2b illustrate this. The pictures are taken from a study reported in Parrill (2010) where participants described short cartoon clips to a friend they brought to the study. The two persons shown in Figure 2 describe the scene shown in Figure 1 where a skunk is hopping across a room. The person in Figure 2a produces a clear instance of a CVG as they mimic the skunk's body posture and also its facial expressions, thus adopting a first-person perspective. The person in Figure 2b, by contrast, describes the same hopping event by means of an OVG because they only use their right index finger to trace the hopping movement of the skunk. Thus, they adopt the external, third-person perspective that is representative of OVGs. From these examples, another interesting can be made: the two types of viewpoint gestures differ with respect to what they typically express. For example, OVGs encode trajectory more often than CVGs do (Parrill 2010). Another interesting observation is that CVGs have been shown to be more informative than OVGs (Beattie & Shovelton 2002). Several other differences between CVGs and OVGs have been attested. However, they will not be discussed here, as they would exceed the scope of the present paper.

2.3. QUOTATION AS DEMONSTRATION. Quotations, such as the DD utterance in (1a), have traditionally been argued to be verbatim repetitions of what someone else said (e.g., Quine 1940, Davidson 1979). The quoted words were assumed to be mentioned, but not used, suggesting they cannot interact with surrounding linguistic material. More recent research, however, has shifted away from a strict verbatim condition (e.g., Davidson 2015), and it is more common to assume that quoted elements are often simultaneously used and mentioned. This is evidenced by examples

Figure 1: Cartoon scene of a skunk hopping across a room. Taken from Parrill (2010)

(a) CVG used to depict the skunk in Figure 1 (Figure 3 in Parrill 2010, p. 652)



(b) OVG used to depict the skunk in Figure 1 (Figure 2 in Parrill 2010, p. 651)

Figure 2: Examples of a CVG and an OVG to depict the event shown in Figure 1

such as the following:

(4) Stiviano's lawyer has not denied the part about the gifts, although he says there is not a "peppercorn of a fact" that any fraud was involved.

(NYT, May 1, 2014, cited from Davidson 2015, p. 483)

Here, the quoted constituent *a peppercorn of a fact* is used and mentioned at the same time. It is mentioned because the author of the sentence in (4) indicates by use of the quotation marks that the words quoted are a verbatim repetition of what Stiviano's lawyer said. However, the phrase is also used as it needs to be integrated with the surrounding linguistic material to derive a full proposition. The content of the quotation thus also plays a role in this example. This means that the underlying report paraphrasing the original proposition expressed in the reported speech act can be retrieved when ignoring the quotation marks. The quotation marks add another level of meaning—namely that whatever stands in the quotation marks was part of the original utterance (Potts 2005, Maier 2015). This is commonly known as *mixed quotation*.

For FID utterances (cf. (1c)), an analysis has been proposed treating it as a highly specialized form of mixed quotation where the original thought or utterance of the reported speaker with pronouns and tense markings being systematically unquoted, thus accounting for their shifting behavior (Maier 2015). The unquotation mechanism is conceptualized as a pragmatic process. The quotational structure of (1c) is thus as in (5). Square brackets indicate unquotation.

(5) On her way home, Mary heard a song by Kendrick Lamar that she liked on the radio. "[She] [would] buy [his] new album tomorrow."

Note that the quotation marks and the square brackets are just used to illustrate (un)quotation. Normally, they are omitted in FID utterances.

Before turning to experimental evidence in favor of the mixed-quotational account of FID proposed by Maier (2015), let's briefly introduce the idea that quotation involves demonstration (Clark & Gerrig 1990). This means that quotations in general serve a depictive rather than a descriptive purpose and are thus iconic in their nature. Additionally, the idea that quotation involves demonstration implies that speakers can quote aspects of an utterance relating to its form and not only to its content.

(6) And so she said "[whispering] What are we going to do?"

(Clark & Gerrig 1990, p. 487)

Here, not only words are quoted, but also the manner in which they were uttered, that is, with a whispering voice.

The predictions of the demonstrational account to quotation go even further, as evidenced by (7). The *be like*-construction is also an instance of quotation (e.g., Davidson 2015) where special emphasis is put on quoting the form of an utterance. The quotation marks are again only inserted to highlight the quoted part of the utterance and are normally omitted.

(7) My cat was like "feed me!"

(Davidson 2015, p. 485)

By virtue of uttering (7), an event is demonstrated, i.e., quoted, where the speaker's cat behaves in a certain manner leading them to infer that it is hungry. This illustrates that when assuming that quotation is inherently demonstrational, even non-linguistic behavior can be quoted.

Davidson (2015) proposes a formal analysis for these observations and explicitly claims that also speech-accompanying gestures can be part of a demonstration, or, in other words, quoted. This leads Ebert & Hinterwimmer (2022) to combine Davidson's (2015) proposal with the formal proposal for the semantic contribution of gestures as proposed by Ebert et al. (2020) (cf. Section 2.2 of the present article). By means of a rating study, they test their claims by testing DD, ID, and FID speech reports in which a pronoun occurs that is co-referent with the reported speaker (first-person pronoun in DD, third-person pronoun in ID and FID). A sample item (translated from German) is given below. Capitalization indicates focalization.

(8) a. **FID:** Leona was extremely annoyed. Again, SHE had paid the bill for the entire group. + SELF-POINTING
b. **ID:** Leona was extremely annoyed. She was angry that SHE had paid the bill for the entire group again. + SELF-POINTING
c. **DD:** Leona was extremely annoyed. Angrily, she thought: "Now, I have paid the bill for the entire group again." + SELF-POINTING

(Ebert & Hinterwimmer 2022, p. 345)

Following Davidson (2015) in the assumption that quotation is demonstrational and therefore also gestures can be quoted, Ebert & Hinterwimmer (2022) hypothesize that self-pointing aligned with the first-person pronoun should be acceptable in the DD example (8c) because it is then interpreted as a CVG quoted from the reported speaker. If Maier's (2015) theory is correct and FID in fact involves mixed quotation, the self-pointing should also be acceptable in examples like (8a). Finally, for ID utterances as in (8b) they hypothesize that self-pointing should not be acceptable when aligned with a third-person pronoun as ID does not involve quotation and therefore the gesture cannot be interpreted as a quoted CVG. They thus predict that self-pointing should be judged significantly worse in ID than in FID and DD.

Their results were only partly in line with their hypothesis. Descriptively, the results show that ID was the condition that received the worst ratings. An ANOVA revealed, however, that this difference was only significant in the F1 and F2 analysis in the contrast between DD and ID. The contrast between FID and ID only reached significance in the F1 analysis, which they attribute to the surprisingly good ratings obtained for the ID condition. Overall, they interpret the results as follows: The good ratings for the FID condition show that it is indeed quotational, thus presenting evidence in favor of the approach proposed by Maier (2015). For ID, they argue, the results imply that a similar perspective-shifting strategy as in FID is available there, as well, meaning that self-pointing in ID can also be interpreted as a quoted CVG.

Ebert & Hinterwimmer's (2022) interpretation of the findings for ID thus suggests that it also is an instance of mixed quotation. It should thus receive a similar treatment as FID in the account proposed by Maier (2015). This aligns with the observation that some indexical expressions can shift or obligatorily shift toward the reported speaker in ID. However, the quotation mechanism should arguably be more constrained than in FID as more expressions are interpreted from the reporting speaker's perspective in ID than in FID. The study reported in this paper aims to test this hypothesis by adapting the study reported in Ebert & Hinterwimmer (2022). It tested whether altering the perspective prominent at the speech level in ID could affect the availability of interpreting a self-pointing gesture as a CVG quoted from the matrix subject. Following work suggesting that some indexicals can shift if the matrix subject's perspective is prominent (Plank 1986, Anderson 2019), it is hypothesized that self-pointing gestures aligned with a third-person pronoun referring to the matrix subject are acceptable in ID only if the matrix subject's perspective is prominent on the speech level. When the reporting speaker's perspective is prominent in the ID utterance, by contrast, self-pointing should not be acceptable as it then cannot be interpreted as a CVG quoted from the matrix subject.

3. Experimental study.

3.1. Method.

3.1.1. Participants. Self-reported native speakers of German (n = 60) were recruited via Prolific. All of them were naive with respect to the research question.

3.1.2. Materials. Sixteen videotaped experimental items were constructed. Each experimental item consisted of an opening statement, followed by the target sentence rendered in ID. The first sentence described how the protagonist, i.e., the matrix subject of the ID sentence, felt. The subsequent ID utterance explained why they felt this way. Crucially, in the ID sentence, either

the matrix subject's (cf. (9a)) or the speaker's perspective (cf. (9b)) was made prominent (factor PERSPECTIVE: matrix subject vs. speaker). This was tested in a pilot study. In addition, each ID sentence contained a focalized third-person pronoun which was co-referent with the matrix subject. This pronoun was either aligned with a self-pointing or a beat gesture (factor GESTURE: self-pointing vs. beat). Thus, the study was of a 2x2 design. An example is given in (9):

(9) a. **Matrix subject's perspective:** Pia ging es erbärmlich. Sie fragte sich, warum ihre beste Freundin Anna, diese gottverdammte Saufziege, gestern Abend mal wieder IHR zu viel Wein nachgeschüttet hat, obwohl sie doch so wenig verträgt. + SELF-POINTING/BEAT
'Pia was feeling miserable. She wondered why her best friend Anna, that damned lush, had poured HER too much wine again last night, even though she couldn't handle it.'

b. **Speaker's perspective:** Pia ging es erbärmlich. Sie fragte sich, warum ihre beste Freundin Anna, die aber eigentlich nur die letzte Pfütze aus der Weinflasche loswerden wollte, gestern Abend mal wieder IHR zu viel Wein nachgeschüttet hat, obwohl sie doch so wenig verträgt. + SELF-POINTING/BEAT
'Pia was feeling miserable. She wondered why her best friend Anna, who was just trying to get rid of the last drops in the wine bottle, had poured HER too much wine again last night, even though she couldn't handle it.'

In (9a), the matrix subject's perspective is prominent due to the appositive *diese gottverdammte Saufziege* ('that damned lush'). Under its most salient interpretation, the appositive conveys an attitude toward Anna that is most likely Pia's. It seems implausible to assume that this is the speaker's attitude in this context. By contrast, in (9b), the appositive *die aber eigentlich nur die letzte Pfütze aus der Weinflasche loswerden wollte* ('who was just trying to get rid of the last drops in the wine bottle') is most plausibly interpreted as additional information about the wine-pouring scenario to which only the speaker has access, not Pia. This makes the speaker's perspective more prominent in this example. To distract participants from the experimental manipulations, the experimental items were interspersed with 30 unrelated fillers.

3.1.3. PROCEDURE. Participants were first made familiar with the task by means of an introductory text. In this text, they were explicitly instructed to always pay attention to the audio and the videotape. In addition, they were informed about their data protection rights and had to give informed consent before starting to complete the questionnaire, which started with a training session consisting of two items. SoSci Survey (Leiner 2022) was used to create the questionnaire, an online platform which can be used free of charge for academic purposes. The questionnaire was distributed via Prolific using its pre-filtering function to exclude participants from the aforementioned pilot study and to ensure that only native speakers of German participate in the experiment. Each list was run independently in order to prevent participants from participating multiple times. Again, this was achieved by using Prolific's pre-filtering function.

The items were split up according to a Latin square design onto four lists. The 30 fillers and the two training items were included on every list. Experimental items and fillers were randomized for each participant. Their task was to rate the items on a 7-point Likert scale for acceptability (1 = completely unacceptable; 7 = completely acceptable). To check whether the participants paid attention to the stimuli, they were asked attention questions about the videos (e.g., they were asked

what the color of the background in the videos was) at the end of the questionnaire.

3.2. PREDICTIONS. Based on the hypothesis that self-pointing gestures in ID can be interpreted as CVGs quoted from the matrix subject when their perspective is prominent, it is assumed that self-pointing will be rated significantly higher in the matrix subject condition of the factor PERSPECTIVE. Since beat gestures synchronize with the speech rhythm in general (McNeill 1992), they are assumed to be highly acceptable regardless of the manipulation of the factor PERSPECTIVE. Therefore, an interaction between the factors PERSPECTIVE and GESTURE is predicted.

3.3. RESULTS. Statistical analysis was done using the programming language R (R Core Team 2022) inside the integrated development environment RStudio (RStudio Posit Team 2023). For data processing and visualization, the package 'tidyverse' (Wickham et al. 2019) was used. To test for significant effects, the results were analyzed using a cumulative link mixed effects model with the clmm() function in the R package 'ordinal' (Christensen 2023). Using effect coding, the two factors GESTURE and PERSPECTIVE and all their interactions were entered as fixed effects into the model, meaning that the intercept represents the unweighted grand mean and the fixed effects compare the factor levels to each other. The analysis script as well as the materials are available on osf: https://osf.io/h6rt5/.

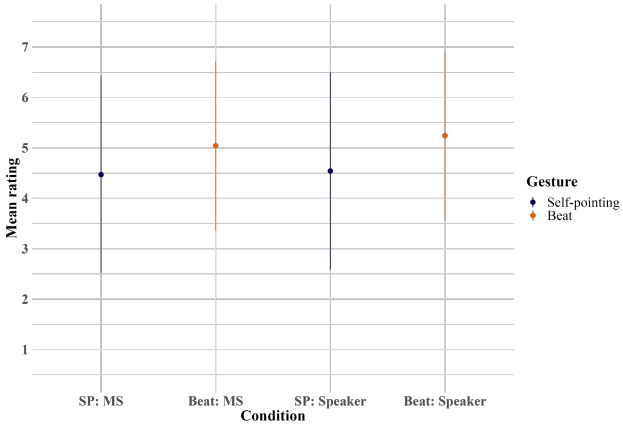


Figure 3: Mean values and standard deviations for each condition. (Abbreviations: SP = Self-pointing, MS = matrix subject's perspective prominent, Speaker = speaker's perspective prominent)

Figure 3 shows the mean values and standard deviations. The self-pointing condition was rated equally well in both conditions of the factor PERSPECTIVE (matrix subject: M = 4.47, SD = 1.95; speaker: M = 4.54, SD = 1.96). The same can be observed for the beat condition for GESTURE (matrix subject: M = 5.04, SD = 1.68; speaker: M = 5.24, SD = 1.67). In general, it can be observed that beat gestures were overall rated better than self-pointing gestures.

The output of the ordinal mixed-effects model is given in Table 1. It shows a main effect for the factor GESTURE, confirming the tendency observed in the descriptive data shown in Figure 3: Beat gestures were rated significantly higher than self-pointing gestures, irrespective of the manipulation of the factor PERSPECTIVE.

3.4. DISCUSSION. Contrary to the predicted interaction, the results only show a main effect for GESTURE. Despite this contradiction, the results are still in line with the hypothesis: as evidenced

	Estimate	Std. Error	z value	Pr(> |z|)
Gesture	-.757	.12	-6.212	5.24e-10 ***
Perspective	.162	.118	1.375	.169
Gesture:Perspective	-.181	.235	-.767	.443

Table 1: Ordinal mixed-effects model with Mode and Match as fixed effects and participants and items as random intercepts. Formula: choiceO ∼ Gesture * Perspective + (1|CASE) + (1|item) Significance codes: *** 0.001 | ** 0.01 | * 0.05 | . 0.1

by the descriptive data in Figure 3, self-pointing gestures are acceptable in ID. In a way, the results of the study reported here even go beyond the initial hypothesis because not only are self-pointing gestures in ID acceptable when the matrix subject's perspective is prominent on the speech level, but also when the speaker's perspective is made prominent. As will be elaborated below, the results are in line with the assumption that—in a similar vein as in FID (Ebert & Hinterwimmer 2022)—a perspective shift can also take place in ID utterances and consequently, self-pointing gestures aligned with third-person pronouns can be interpreted as CVGs quoted from the matrix subject.

4. General discussion and conclusion. The study reported by Ebert & Hinterwimmer (2022) tested whether FID utterances are an instance of mixed quotation (cf. Maier 2015). Under the assumption that quotation involves demonstration (Clark & Gerrig 1990, Davidson 2015), they argue that a self-pointing gesture aligned with a third-person pronoun co-referent with the reported speaker should be acceptable as it is then interpreted as a CVG quoted from them. The results supported the mixed-quotational account of FID put forth in Maier (2015). As a control condition, they also included ID utterances, hypothesizing that self-pointing should not be acceptable here, since ID utterances are normally argued to not involve quotation. Surprisingly, however, the ID condition was also rated surprisingly well, thus suggesting that some form of quotation is—or at least can be—present in ID, as well.

The study reported in this paper aimed to further investigate this. Previous research has shown that some indexical expressions can receive a shifted interpretation if the perspective of the matrix subject is prominent in the surrounding discourse (Plank 1986, Anderson 2019). Given the findings of Ebert & Hinterwimmer's (2022), this shifted interpretation could be re-analyzed as an instance of these expressions being quoted from the matrix subject, in line with their suggestions for the self-pointing data. Therefore, the present study manipulated the perspective prominent on the speech level (matrix subject's perspective vs. speaker's perspective), hypothesizing that self-pointing interpreted as a quoted CVG is only acceptable in ID if the perspective of the matrix subject is prominent on the speech level. The results go beyond this hypothesis, as self-pointing in ID was found to be acceptable regardless of the perspective prominent on the speech level. This suggests that mixed quotation is available in ID. This means that ID should receive a formal treatment along the lines of the one proposed for FID by Maier (2015). However, as not all elements can be quoted in ID, the quotation mechanism should be more constrained in ID in comparison to FID, as it remains unclear for now which elements in ID can be quoted and under which circumstances. The results of a study reported elsewhere (Walter 2024) suggest that face emoji, which normally receive an author-oriented interpretation by default (Grosz et al. 2023), can also receive

a shifted interpretation in ID. I argued that they should be analyzed as quoted facial expressions from the matrix subject. For emoji to shift, however, the matrix subject's perspective has to be prominent, highlighting the importance for future research to further investigate the influence perspective prominence in ID has on the availability of shifted interpretations and thus the availability of mixed quotation in ID.

References

Anand, Pranav & Andrew Nevins. 2004. Shifty operators in changing contexts. In Robert B. Young (ed.), *Proceedings of Semantics and Linguistic Theory (SALT) 14*, 20–37. Ithaca, NY: CLC Publications. 10.3765/salt.v14i0.2913.

Anderson, Carolyn J. 2019. Tomorrow isn't always a day away. In M. Teresa Espinal, Elena Castroviejo, Manuel Leonetti, Louise McNally & Cristina Real-Puigdollers (eds.), *Proceedings of Sinn und Bedeutung 23*, 37–56. Barcelona, Spain: Universitat Autònoma de Barcelona.

Beattie, Geoffrey & Heather Shovelton. 2002. An experimental investigation of some properties of individual iconic gestures that mediate their communicative power. *British Journal of Psychology* 93(2). 179–192. 10.1075/gest.1.2.03bea.

Christensen, Rune H. B. 2023. *ordinal—Regression Models for Ordinal Data*. https://CRAN.R-project.org/package=ordinal. R package version 2023.12-4.

Clark, Herbert H. & Richard J. Gerrig. 1990. Quotations as demonstrations. *Language* 66(4). 764–805. 10.2307/414729.

Davidson, Donald. 1979. Quotation. *Theory and Decision* 11(1). 27. 10.1007/bf00126690.

Davidson, Kathryn. 2015. Quotation, demonstration, and iconicity. *Linguistics and Philosophy* 38(6). 477–520. 10.1007/s10988-015-9180-1.

Ebert, Christian, Cornelia Ebert & Robin Hörnig. 2020. Demonstratives as dimension shifters. In *Proceedings of Sinn und Bedeutung 24*, 161–178. Osnabrück: University of Osnabrück.

Ebert, Cornelia. 2024. Semantics of gesture. *Annual Review of Linguistics* 10(1). 169–189.

Ebert, Cornelia & Stefan Hinterwimmer. 2022. Free indirect discourse meets character viewpoint gestures. In Sam Featherston, Robin Hörnig, Andreas Konietzko & Sophie von Wietersheim (eds.), *Proceedings of Linguistic Evidence 2020: Linguistic theory enriched by experimental data*, 333–349. Tübingen, Germany: University of Tübingen.

Eckardt, Regine. 2014. *The semantics of free indirect disocurse: How texts allow to mind-read and eavesdrop*. Leiden, The Netherlands: Brill. 10.1163/9789004266735.

Farkas, Donka F. & Kim B. Bruce. 2010. On reacting to assertions and polar questions. *Journal of Semantics* 27(1). 81–118. 10.1093/jos/ffp010.

Grosz, Patrick G., Gabriel Greenberg, Christian De Leon & Elsi Kaiser. 2023. A semantics of face emoji in discourse. *Linguistics and Philosophy* 46(4). 905–957.

Harris, Jesse A. 2012. *Processing perspectives*. Amherst, MA: University of Massachusetts Amherst dissertation.

Hinterwimmer, Stefan. 2017. Two kinds of perspective taking in narrative texts. In Dan Burgdorf, Jacob Collard, Sireemas Maspong & Brynhildur Stefánsdóttir (eds.), *Proccedings of Semantics and Linguistic Theory (SALT) 27*, 282–301. University Park, MD: University of Maryland. 10.3765/salt.v27i0.4153.

Kendon, Adam. 2004. *Gesture: Visible action as utterance*. Cambridge, UK: Cambridge University Press. 10.1017/CBO9780511807572.

Leiner, Daniel J. 2022. SoSci Survey (Version 3.3.14a). Available at https://www.soscisurvey.de.

Maier, Emar. 2015. Quotation and unquotation in free indirect discourse. *Mind & Language* 30(3). 345–373. 10.1111/mila.12083.

McNeill, David. 1992. *Hand and Mind: What Gestures Reveal about Thought*. Chicago, IL: University of Chicago Press.

Parrill, Fey. 2010. Viewpoint in speech-gesture integration: Linguistic structure, discourse structure, and event structure. *Language and Cognitive Processes* 25(5). 650–668. 10.1080/01690960903424248.

Plank, Frans. 1986. Über den Personenwechsel und den anderer deiktischer Kategorien in indirekter Rede. *Zeitschrift für germanistische Linguistik* 14(3). 284–308. 10.1515/zfgl.1986.14.3.284.

Potts, Christopher. 2005. *The logic of conventional implicatures*. Oxford, UK: Oxford University Press. 10.1093/acprof:oso/9780199273829.001.0001.

Quine, Willard V. 1940. *Mathematical logic*, vol. 4. Cambridge, MA: Haravrd University Press.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing Vienna, Austria. https://www.R-project.org/.

RStudio Posit Team. 2023. *RStudio: Integrated development environment for R*. Posit Software, PBC Boston, MA. http://www.posit.co/.

de Ruiter, Jan P. 2000. The production of gesture and speech. In David McNeill (ed.), *Language and gesture*, Cambridge, UK: Cambridge University Press. 10.1017/cbo9780511620850.018.

Schlenker, Philippe. 2004. Context of thought and context of utterance: A note on free indirect discourse and the historical present. *Mind & Language* 19(3). 279–304. 10.1111/j.1468-0017.2004.00259.x.

Tonhauser, Judith. 2012. Diagnosing (not-)at-issue content. In Elizabeth Bogal-Allbritten (ed.), *Proceedings of Semantics of Under-represented Languages of the Americas (SULA) 6*, 239–254.

Walter, Sebastian. 2024. Can face emoji receive a shifted interpretation in indirect discourse? Poster presented at *Architechtures and Mechanisms of Language Processing (AMLaP)* 30, Edinburgh, UK.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, Alex Hayes, Lionel Henry, Jim Hester, Max Kuhn, Thomas Lin Pedersen, Evan Miller, Stephan Milton Bache, Kirill Müller, Jeroen Ooms, David Robinson, Dana Paige Seidel, Vitalie Spinu, Kohske Takahashi, Davis Vaughan, Claus Wilke, Kara Woo & Hiroaki Yutani. 2019. Welcome to the tidyverse. *Journal of Open Source Software* 4(43). 1686. 10.21105/joss.01686.