Shifted Interpretation of Pointing Gestures in Speech Reports

Oliver Herbort (oliver.herbort@uni-wuerzburg.de)[1], Cornelia Ebert[2], Lisa-Marie Krause[1], Stefan Hinterwimmer[3] & Sebastian Walter[2]

[1]Department of Psychology, Julius-Maximilians-Universität Würzburg, Würzburg, Germany
[2]Institute of Linguistics, Goethe University Frankfurt, Frankfurt, Germany
[3]Department of German Studies, University of Hamburg, Hamburg, Germany

Abstract

We report an experiment on the interpretation of pointing gestures in speech reports. German speakers read a narrative in which a speaker reported another person's utterance using either direct or indirect discourse. The reporting speaker produced a pointing gesture whose interpretation depended on whether it was taken to be part of the reported speech event or a contribution of the current utterance. As only direct discourse is typically assumed to be quotational, we hypothesized that participants would interpret the gesture as part of the reported speech event more frequently in direct discourse. Participants identified the referent of the gesture and rated their certainty. The results confirmed the hypothesis: Participants were more likely to interpret the pointing gesture as part of the reported speech event in direct discourse. At the same time, gesture transformations were rare and associated with lower certainty across conditions, indicating competition from the canonical referential function of pointing.

Keywords: speech reports; pointing gestures; direct discourse; indirect discourse; quotation; multimodality

Introduction

Spoken language is inherently multimodal: In face-to-face communication, speakers routinely combine speech with gestures, gaze, and other non-verbal cues. A growing body of work has shown that such multimodal contributions are systematically integrated during interpretation and can affect reference resolution, perspective-taking, and discourse structure (McNeill, 1992; Kendon, 2004). From a semantic perspective, gestures are therefore not peripheral, paralinguistic elements, but form part of the meaning contributed by an utterance (e.g., Ebert et al., 2020).

The present study focuses on multimodal interpretation in speech reports, that is, direct and indirect discourse, a domain that is particularly well suited to probing how listeners integrate gestural meaning. Speech reports generally involve two speech situations—the context of the reported event and the current utterance context—and thus raise the question of how gestures are interpreted in these linguistic environments. Here, we investigate this question by comparing the interpretation of pointing gestures in direct versus indirect discourse, asking whether such gestures are interpreted as part of the reported speech event or instead as contributions of the reporting speaker in the current utterance context.

Pointing Gestures

Pointing gestures are canonically used to identify a referent in the shared environment and are therefore closely tied to demonstratives and referential expressions (Kaplan, 1989). Pointing gestures help to pick out or constrain a discourse referent, and they are usually interpreted relative to the current utterance context, with the pointed-at object serving as the demonstratum. At the same time, their interpretation depends on how they are integrated with speech (Ebert et al., 2020). In reported speech contexts, this gives rise to a potential ambiguity: A pointing gesture may either be interpreted canonically, as identifying a referent in the utterance context, or non-canonically, as part of a quoted or demonstrated event. Because the referential function of pointing gestures is so strong, we assume that interpretations in which they are integrated into a reported speech event are expected to be marked, for example, by the use of direct discourse.

Different Types of Speech Reports

When speakers report what someone else said or thought, they typically choose between direct discourse (DD) and indirect discourse (ID). While both report types serve the same communicative function—namely, representing a prior speech or thought event—they differ systematically in whether they predominantly encode the reporting or the reported speaker's perspective and how strongly they preserve features of the original utterance.

This contrast is illustrated by the minimal pair in (1).

(1) a. Yara said, "I will meet you here tomorrow."
b. Yara said that she would meet them there the following day.

The example in (1a) is a case of DD, whereas (1b) is an instance of ID. One of the most widely discussed differences between the two concerns the interpretation of perspective-dependent expressions such as personal pronouns, temporal adverbials, and spatial indexicals (Kaplan, 1989; Schlenker, 2003). In DD, these expressions are interpreted relative to the reported speech context and not in the current utterance, i.e., reporting, context. In ID, by contrast, they are typically interpreted relative to the reporting context.

In (1a), the first-person pronoun *I* thus refers to Yara, the addressee pronoun *you* refers to Yara's original addressee, *here* refers to where Yara was, and *tomorrow* is interpreted relative to the time of Yara's utterance and thus refers to the day following the day of Yara's utterance. In (1b), however, pronouns appear in the third person, and temporal expressions are adjusted relative to the time of reporting. These systematic differences have traditionally been taken to

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

reflect a deeper structural distinction: DD is assumed to preserve aspects of the original utterance event, whereas ID primarily conveys propositional content (Cappelen & Lepore, 1997; Davidson, 1979).

Perspective-Taking in Direct vs. Indirect Discourse

Direct Discourse Because DD preserves indexicals and other perspective-dependent expressions, it has often been treated as the clearest instance of quotation. In DD, all indexicals are interpreted from the perspective of the reported speaker. This includes personal pronouns, spatial expressions, and temporal adverbials (Davidson, 1979; Schlenker, 2003), as evidenced by (1a) already, meaning that the reporting speaker temporarily adopts the reported speaker's deictic center, allowing listeners to reconstruct the original perspective.

Importantly, perspective preservation in DD extends beyond indexicals. Evaluative expressions and expressive content are likewise interpreted as originating from the reported speaker:

(2) Mary said, "This movie was fantastic!"

In this example, the evaluative adjective *fantastic* is most naturally interpreted as presenting Mary's opinion of the movie, not the reporting speaker's, as it is part of the quote. This systematic preservation of perspective has motivated analyses treating quoted utterances as speech events that include both form and content (Maier, 2018). This means that DD can be defined as not merely describing what was said but presents aspects of the original utterance itself, i.e., how it was said. This observation has been recast in terms of the use–mention distinction and has led to the assumption that quoted words in DD are merely mentioned but not used (e.g., Quine, 1940), in that quotation makes aspects of an utterance's form accessible to interpretation rather than contributing its content. This contrasts it with ID, which is standardly analyzed as a non-verbatim report of the original utterance event (e.g., Hintikka, 1969).

Because DD foregrounds the reported speaker's perspective so strongly, it provides a natural environment for listeners to attribute accompanying cues—such as marked prosody or gesture—to the reported speaker rather than the reporting speaker (e.g., Clark & Gerrig, 1990; Wade & Clark, 1993).

Indirect Discourse In contrast to DD, ID has traditionally been analyzed as non-quotational. Under standard accounts, ID conveys the propositional content of a prior utterance without preserving its linguistic form (Hintikka, 1969). This difference is reflected in the behavior of perspective-dependent expressions in (1b), repeated here for convenience:

(1) b. Yara said that she would meet them there the following day.

In (1b), pronouns have to be in the third person to refer back to the reported speaker, and the temporal indexical from (1a) must be exchanged with a non-indexical temporal expression in order to refer to the day following Mary's utterance. Personal pronouns generally do not shift in ID in languages such as German and English. This has often been taken as evidence that ID lacks direct access to the original utterance context, in line with the assumption that is treated as a non-quotational report.

Effect of Speech Report Type on Gesture Interpretation Based on these considerations, we aim to test whether the type of speech report affects how accompanying pointing gestures are interpreted. Pointing gestures in ID reports are expected to be interpreted in the current utterance context, such that listeners preferentially treat them as directly identifying a referent in that context. That is, the referent is considered whatever the reporting person is pointing at.

In DD utterances, by contrast, an interpretation should be available where pointing gestures are integrated into a multimodal quotation. That is, the gesture might be interpreted as if it were performed by the reported speaker, taking into consideration the reported speaker's position and orientation, which may most likely differ from that of the reporting person.

Of course, the distinction between ID and DD is not as clear cut as traditional reasoning suggests. In the following, we thus discuss evidence that ID may sometimes recur to the reported speaker's perspective and that DD may not necessarily require a verbatim reproduction of a prior utterance. Nevertheless, we still expect that pointing gestures accompanying DD are more likely to be transformed to the position and orientation of the person whose speech is reported, that is, receive a *shifted* interpretation, than gestures accompanying ID.

The Reported Speaker's Perspective in ID

There are at least two lines of evidence showing that the reported speaker's perspective may be considered in ID. A growing body of literature suggests that shifted interpretations of pronouns and other indexicals are readily available in ID in numerous languages (Anand & Nevins, 2004; Deal, 2020; Schlenker, 2003, among others). Likewise, a growing body of empirical work suggests that perspectival anchoring in ID is more flexible than traditionally assumed, also in languages like German and English that, as argued above, allegedly do not permit indexical shift. Plank (1986) shows that local indexicals in German can receive shifted interpretations in ID, and Anderson (2019) reports similar findings for temporal indexicals in English.

(3) Hans war wütend. Er beschwerte sich, dass er morgen schon wieder mit dem Hund gehen musste.
'Hans was furious. He complained that he would have to walk the dog again tomorrow.'

In suitable discourse contexts, *morgen* ('tomorrow') can be interpreted relative to the reported speaker's perspective and thus from the reported context rather than the reporting one. These findings challenge the assumption that ID categorically excludes perspective shift.

Recent experimental work by Saure (2025) further demonstrates that perspective shift in German ID utterances is licensed under specific discourse conditions. In narrative environments where the reported speaker is sufficiently prominent, participants accept interpretations of indexicals that align with the reported speaker's viewpoint. These results suggest that perspective shift in ID is not exceptional but reflects general mechanisms governing how listeners track perspectival centers in discourse.

A second line of evidence is that natural language provides examples in which parts of an utterance clearly preserve features of the reported speaker's original wording. Such cases have been discussed under the label of mixed quotation (Davidson, 1979; Potts, 2007; Maier, 2017). Consider the following example (taken from Cappelen & Lepore, 1997, p. 429):

(4) Alice said that life "is difficult to understand."

In (4), the quoted material is syntactically integrated into the reporting clause while simultaneously marked as originating from the reported speaker. The quotation marks indicate that the speaker intends to preserve the form aspects of the original formulation rather than merely conveying its propositional content. At the same time, as is standardly argued for mixed quotation, the words enclosed by the quotation marks are also used, as their content is needed to derive a full proposition.

Mixed quotation demonstrates that the boundary between DD and ID is not absolute (e.g., Plank, 1986), yet it is often treated as such in formal semantics. Even within ID, speakers may selectively preserve linguistic material associated with the reported speaker.

In summary, the above examples suggest that aspects of the reported speech or the reported speaker may be preserved in ID. Our experiment would provide further support for these views if participants at least sometimes transformed pointing gestures accompanying ID.

Quotation as Demonstration

A growing line of research argues that quotation and thus also DD as one of its subclasses is fundamentally iconic rather than purely descriptive. Clark & Gerrig (1990) propose that quotation is a mode of communication they call demonstration. When demonstrating, speakers depict certain aspects of an event, for example. In cases of quotation and hence also in DD, the depicted event is some kind of speech event. Put differently, when rendering an utterance in DD, speakers iconically depict certain aspects of the original speech event. In quotation in general and DD in particular, demonstrations can include prosody, gesture, and facial expression. Crucially, demonstrations are only selectively depictive, meaning that quotation need not reproduce all features of the reported speech event.

Davidson (2015) provides central evidence for this view from *be like* constructions in English and role shift constructions in American Sign Language. Consider the following example:

(5) John was like, "The talk was soooooo boring."

In such cases, speakers do not necessarily attribute an exact utterance to John. Instead, they depict John's stance, attitude, reaction, or, in the case of (5), the marked prosody John used in the original speech vent. The reported material functions as a demonstration of the original event rather than a verbatim quotation. Crucially, *be like* constructions can include gestures, facial expressions, and prosodic cues that contribute to how the reported event is represented.

Davidson extends this analysis to DD more broadly. Even when speakers use canonical quotation formats, they need not necessarily reproduce prior utterances verbatim, as this, in Davidson's (2015) formal approach, is only a pragmatic principle rather than semantically hardwired. Instead, they present a depiction of the original event that highlights aspects deemed communicatively relevant. This is in line with experimental findings suggesting that interlocutors do not expect quoted utterances to be verbatim (Wade & Clark, 1993). The strict verbatimness requirement often posited in formal-semantic research is a mere artifact of the predominant focus on written data as the primary empirical source for theorizing.

Building on this framework, recent work (e.g., Walter & Hinterwimmer, 2025) proposes that ID can also license unmarked mixed quotation if it is seen as a demonstration. On this account, mixed quotation in ID is best understood as a demonstration of aspects of the reported speaker's original utterance rather than a purely descriptive report.

Experimental findings on gesture interpretation provide support for this proposal (Ebert & Hinterwimmer, 2022; Walter, 2025). When self-pointing gestures aligned with a third-person pronoun that is co-referent with the matrix subject occur in ID, participants often interpret them as quoted from the reported speaker.

(6) Mary said that [she] would win the race. + SELF-POINTING

Such interpretations suggest that listeners treat certain elements of ID utterances as demonstrations of the reported speaker's behavior. Perspective shift in ID thus arises when listeners interpret parts of an utterance as demonstrational. Crucially, such interpretations are optional and depend on contextual factors that make the reported speaker's perspective salient (cf. Saure, 2025; Walter & Hinterwimmer, 2025).

Regarding the experiment reported in this paper, the quotation-as-demonstration account by Clark & Gerrig (1990) predicts that pointing gestures can, in principle, be interpreted as part of the quotation and hence from the reported speaker's perspective. This radically sets it apart from other accounts of quotation because, to the best of our knowledge, it is the only theoretical account that allows elements besides words to be quoted. Since quotation is an obligatory feature of DD but only an optional one of ID, it moreover predicts that interlocutors should be more willing to interpret a pointing gesture from the reported speaker's perspective in DD as compared to ID.

Hypotheses

Building on this literature, we adopt the view that quotation is obligatorily present in DD but only optionally available in ID. If quotation is understood as demonstration (Clark & Gerrig, 1990; Davidson, 2015), this asymmetry leads to predictions about gesture interpretation.

In DD, listeners should be more inclined to interpret gestures as part of a multimodal quote because the reported utterance is presented as a depiction of the original event. In ID, gestures may more often be attributed to the reporting speaker unless contextual cues favor a quotational interpretation. To test this hypothesis, we showed participants how one person—an informer—reported the actions and speech of another person—the boss of a criminal syndicate. Importantly, the report and the reported events happened in the same room, but the positions and orientations of the informer and the boss in that room were opposed to each other (Figure 1). The report involves a segment about the boss referring to an object that is supported by a pointing gesture of the informer. Our primary interest is whether participants consider the actually pointed-at object as referent (the safe on the right in Figure 1) or an object that would have been pointed at, if the gesture was performed from the position and orientation occupied by the boss (the safe on the left in Figure 1). We expect the latter interpretation to be more frequent when the informer's report contains DD than when it entirely relies on ID.

Method

Participants

The sample size was determined using simulations based on the data of a pilot experiment. We decided to collect data from 400 valid participants, which allows us to reach a decision with the planned Bayesian test with probabilities of >95% and >85% if an effect is present and absent, respectively. Details are provided in the pre-registration. The experiment was advertised on Prolific to persons from Austria, Germany, Switzerland, and neighboring countries, who had an approval rating of at least 90, whose primary language was German, and who did not participate in a pilot experiment. We successfully collected data from 504 online participants, who received 1.20 GBP. The 411 participants whose data were analyzed were on average 32 years old (age range 18–73; 161 female, 246 male, 3 diverse/non-binary, 1 did not disclose). The experiment was pre-registered and in accordance with the guidelines of the ethics committee of the University of Würzburg. The pre-registration, experiment files, data, scripts, and supporting material are available on: https://osf.io/dwkgf/overview?view_only=fb6011b6527e4a0b9409810600155f4d

Stimulus and Apparatus

The experiment was programmed with lab.js (Henninger et al., 2022) and administered over the MindProbe-JATOS server. The primary stimuli were two versions of a verbal report. In one version, an informer used DD to report the speech of another person; in the other version, he used ID. The DD version of the report was:[1]

(1) Quick, we don't have much time. (2) I now tell you what I observed here yesterday. (3) I was examining this room yesterday, when I suddenly heard footsteps. (4) I just so managed to hide behind the crates. (5) From my hiding place, I could see a man with a satchel entering the room and looking around searchingly. (6) Out of the hallway, I heard a female voice that was bossy and angry, that shouted: (7) "… and just because of that snoop we now have to hide our plans down here!" (8) A woman stepped into the doorway and stood there. (9) Then I recognized her. She was the boss. (10) Standing in the doorway, the boss barked at the man: (11) "Don't be so pathetic and put the satchel with the plans in this safe!" (12) The man obeyed the command and the boss said less tensely: (13) "Good. And now let's go." (14) Then, she vanished from the doorway into the hallway behind, dragging the man with her. (15) It's a miracle that they didn't spot me. (16) Now you know where the plans are. My work is done. I'm off.

Figure 1: Segment 11 from the direct speech/informer left condition. Here, the informer uses direct discourse to report the speech of another person (the boss), who stood in the doorway, looked leftward, and would have pointed to the safe on the left side.

The ID version and the original German texts are provided on OSF. The report was broken into 16 segments that were presented individually in speech bubbles in a sequence of cartoon-like images (similar to Figure 1). All images showed the same storage room in which a person (the informer) stood on one side and was looking toward a door on the other side. The room contained one safe next to the door, another one behind the informer, as well as some boxes, crates, and barrels. The informer was presented in slightly varying postures with the hand beside his torso for all segments,

[1] We report English translations of the German text used in the experiment. Note that the German translation of "this" in segment 11 ("diesen") is distance-neutral.

except for segment 11, in which he was pointing in the direction of the far side of the room. To draw attention to the gesture, an animation was displayed between segments 10 and 11, in which the informer switched from a hands-lowered posture (displayed for 500 ms) to a pointing posture (displayed for 500 ms) and in which no speech bubble was shown. The animation was included because the pointing gesture remained unnoticed by many participants in a pilot experiment. A text beneath each image informed participants that they could continue by clicking or pressing the space bar.

Procedure

After giving informed consent and entering demographic information, the instructions were shown. Then, the informer's report was presented as a sequence of images. Participants self-paced transitions from one segment to the next by mouse click or hitting the space bar. After the last segment of the report, the room was shown without the informer and above the instruction to acquire the plans by clicking the object that contains the plans. Once participants clicked on a container (one of the safes or another container), they were asked to rate how certain they were that they chose the correct one on a 7-point scale with anchors "very uncertain" and "very certain" (German: "sehr unsicher" vs. "sehr sicher"). Finally, participants were asked to rate the correctness of two blocks of statements on a 5-point scale with options "certainly false", "somewhat false", "don't know", "somewhat true", and "certainly true" (German: "stimmt sicher nicht", "stimmt eher nicht", "weiß nicht", "stimmt eher", "stimmt sicher"). The first block contained nine statements about the content of the report (e.g., "The boss stood in the doorway"). The second block contained six statements about how the informer's report was perceived (e.g., "The informer used a pointing gesture"). The statements were intended to assess how well participants remembered the form and content of the report and specifically, whether participants picked up all relevant information to transform the pointing gesture. In addition, each block contained one attention check statement, in which participants were asked to respond in a certain way. Statements within blocks were presented in random order. After that, participants were debriefed. The experiment took on average six minutes.

We varied the factor *type of speech report*, i.e., whether the informer's report contained *DD* or *ID*, between participants. In addition, we counterbalanced which of two mirror-reversed versions of the images were presented (pointer left vs. pointer right).

Data Reduction and Analysis

The key dependent variable was the confidence-weighted choice of one of the safes (*gesture transformation score*). Specifically, it was computed by coding the certainty rating with values from 1 (very uncertain) to 7 (very certain) and giving it a positive sign if participants transformed the informer's gesture to the viewpoint of the boss (i.e., clicked on the safe behind the informer) and a negative sign if they did not transform the gesture (i.e., clicked on the safe next to the door). Agreement ratings were coded from 1 (certainly false) to 5 (certainly true).

We used Bayesian tests for data analysis and consider Bayes factors above 3 as evidence for an effect and below 1/3 as evidence against an effect.

Data of 93 participants whose mother tongue was not German (28), who did not answer both manipulation checks correctly (11), who did not agree with the statement "The informer used a pointing gesture" (57), or who did not click on either safe (16) were excluded from analysis. The analyzed data consisted of 206 participants in the DD group and 205 participants in the ID group.

Data was analyzed with R 4.4.1 (R Core Team, 2024) and statistical tests were conducted with the BayesFactor package (Morey & Rouder, 2025).

Results

Figure 2 shows histograms and means of gesture transformation scores. A Bayesian t-test (scale of prior: 0.707) provided strong evidence for an effect of the type of speech report (BF = 3674.4). That is, participants were more inclined to transform the gesture of a pointer to the viewpoint of the person whose speech the pointer was reporting, when the pointer used DD rather than ID. Choices conformed to a gesture transformation for 19% of participants in the DD group and 6% in the ID group.[2]

In addition, we analyzed how certain participants were, depending on the speech report type and whether they transformed the gesture with a Bayesian ANOVA (scale of prior: 0.5). Figure 3 shows that participants who transformed the gesture responded to be less certain about their choice (BF = 13262.8). In addition, Participants in the DD group were also more certain (BF = 3.1). The data were not conclusive with respect to whether there was an interaction (BF = 0.6).

An analysis of the ratings of various statements revealed the following key results. First, participants strongly agreed with statements that were objectively correct (m = 4.7, sd = 0.7) and clearly rejected statements that were objectively false (m = 1.3, sd = 0.6), showing that they attended the task. Second, participants picked up all the information necessary to transform the gesture. That is, participants agreed with the statement that the boss stood in the door (m = 4.6, sd = 0.8), agreed that the man carried out the order (m = 4.9, sd = 0.5), and noted the pointing gesture of the informer (m = 4.9, sd = 0.3; this was an inclusion criterion). The detailed responses are available on OSF.

[2] This pre-registered test was supported by two parameter-free tests, showing that participants in the DD condition were more likely to transform the gesture (Fisher's exact test, $p < .001$) and felt less certain about their response (Wilcoxon rank sum test, U = 18738, $p = .035$).

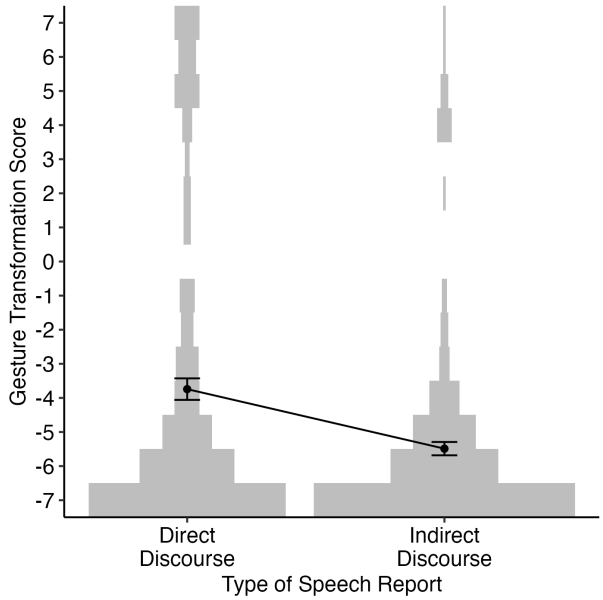

Figure 2: Gesture Transformation Score by type of speech report. Error bars show 1 s.e.m. The grey histograms show the distribution of responses within each group.

Discussion

The results provide clear evidence that speech report type affects the interpretation of pointing gestures. Pointing gestures were more frequently interpreted relative to the reported speech context in DD than in ID. This pattern supports the idea that DD by default involves quotation. It moreover suggests that quotation should be treated as an inherently multimodal phenomenon in spoken language, as not only spoken material but also accompanying gestures can be interpreted as part of the reported speech event and thus as quoted from the reported speaker. When a speaker uses DD, listeners are more willing to attribute multimodal material to the reported speaker rather than to the reporting speaker compared to ID reports.

By contrast, pointing gestures were interpreted relative to the reported context even less frequently in ID, suggesting that quotation is not a constant feature of ID. Note, however, that shifted interpretations of pointing gestures were not categorically ruled out in ID, indicating that ID can, under certain conditions, license a quotational component, in line with Walter & Hinterwimmer (2025), for example.

A second important finding is that pointing gestures were generally rarely interpreted as shifted, regardless of speech report type, and that such interpretations were associated with lower certainty ratings. We suggest that this reflects a general property of pointing gestures: Pointing has a strong and highly conventionalized communicative function, namely, to select a referent in the shared environment of speaker and listener (Kaplan, 1989; Ebert et al., 2020). Interpreting a pointing gesture as part of a reported speech event therefore requires overriding a strong default interpretation and is consequently dispreferred. This conclusion corresponds to the finding that instructing participants to change how they interpret pointing gestures only has a limited effect (Herbort & Kunde, 2018). This could be accounted for by positing a pragmatic principle that normally prevents pointing gestures from shifting, in line with previous findings that perspective shifting in general is cognitively demanding and thus often avoided unless necessary (e.g., Harris, 2021).

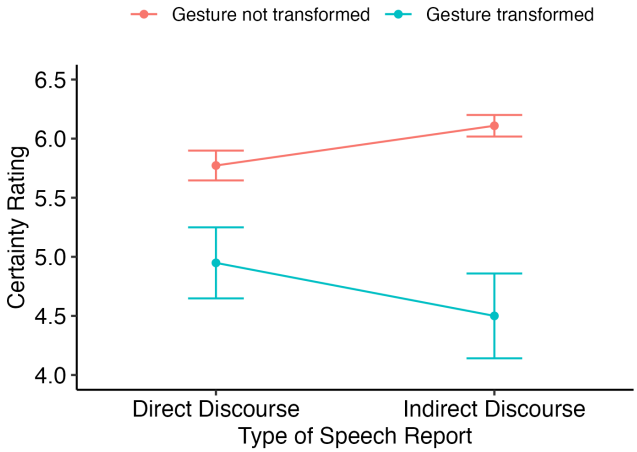

Figure 3: Certainty rating (1 = very uncertain; 7 = very certain) by type of speech report and whether participants transformed the gesture. Error bars show 1 s.e.m.

Finally, the results of our study are compatible with the quotation-as-demonstration account (Clark & Gerrig, 1990; Davidson, 2015). Recall that, on this view, quotation is selective: Not all features of the original speech event need to be demonstrated. A pointing gesture may therefore either be integrated into the quotation and thus receive a shifted interpretation, or it can be interpreted in the reporting context. Overall, the findings support an analysis of quotation as a multimodal phenomenon and show that investigating the interaction of speech and gesture yields new insights into the structure of reported speech.

References

Anand, P. & Nevins, A. (2004). Shifty operators in changing contexts. *Proceedings of Semantics and Linguistic Theory (SALT)*, *14,* 20–37. https://doi.org/10.3765/salt.v14i0.2913.

Anderson, C. J. (2019). Tomorrow isn't always a day away. *Proceedings of Sinn und Bedeutung (SuB)*, *23*, 37–56.

Cappelen, H. & Lepore, E. (1997). Varieties of quotation. *Mind*, *106*(423). 429–450. https://doi.org/10.1093/mind/106.423.429.

Clark, H. H. & Gerrig, R. J. (1990). Quotations as demonstrations. *Language*, *66*(4), 764–805. https://doi.org/10.2307/414729.

Davidson, D. (1979). Quotation. *Theory and Decision*, *11*(1), 27–40. https://doi.org/10.1007/bf00126690.

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

Deal, A. R. (2020). *A theory of indexical shift: Meaning, grammar, and crosslinguistic variation*. MIT Press. https://doi.org/10.7551/mit press/12374.001.0001.

Ebert, C., Ebert, C., & Hörnig, R. (2020). Demonstratives as dimension shifters. *Proceedings of Sinn und Bedeutung*, *24*, 161–178. https://doi.org/10.18148/sub/2020.v24i1. 859.

Ebert, C. & Hinterwimmer, S. (2022). Free indirect discourse meets character viewpoint gestures. *Proceedings of Linguistic Evidence 2020*, 333–349.

Harris, J. A. (2021). Extended perspective shift and discourse economy in language processing. *Frontiers in Psychology*, *12*, 613357. https://doi.org/10.3389/fpsyg.2021.613357.

Henninger, F., Shevchenko, Y., Mertens, U. K., Kieslich, P. J., & Hilbig, B. E. (2022). lab.js: A free, open, online study builder. *Behavior Research Methods*, *54*(2), 556-573. https://doi.org/10.3758/s13428-019-01283-5

Herbort, O., & Kunde, W. (2018). How to point and to interpret pointing gestures? Instructions can reduce pointer-observer misunderstandings. *Psychological Research, 82*(2), 395-406. https://doi.org/10.1007/s00426-016-0824-8

Hintikka, J. (1969). Semantics for propositional attitudes. In J. W. Davies, D. J. Hockney & W. K. Wilson (Eds.), *Philosophical Logic*, 21–45. Reidel.

Kaplan, D. (1989). Demonstratives. In J. Almog, J. Perry & H. Wettstein (Eds.), *Themes from Kaplan*, 565–614. Oxford University Press.

Kendon, A. (2004). *Gesture: Visible action as utterance.* Cambridge University Press. https://doi.org/10.1017/CB O9780511807572

Maier, E. (2017). The pragmatics of attraction. In P. Saka & M. Johnson (Eds.), *The semantics and pragmatics of quotation*, 259–280. Springer. https://doi.org.10.1007/978-3-319-68747-6_9.

Maier, E. (2018). Quotation, demonstration, and attraction in sign language role shift. *Theoretical Linguistics*, *44*(3-4). 265–276. https://doi.org/10.1515/tl-2018-0019.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought.* University of Chicago Press.

Morey, R. D., & Rouder, J. (2025). Computation of Bayes Factors for Common Designs. https://cran.r-project.org/web/packages/BayesFactor/

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

Potts, C. (2007). The dimensions of quotation. In C. Barker & P. Jacobson (Eds.), *Direct compositionality*. Oxford University Press. https://doi.org/10.1093/oso/978019920 4373.003.0012.

Quine, W. V. (1940). *Mathematical Logic.* Harvard University Press.

R Core Team. (2024). R: A Language and Environment for Statistical Computing.

Saure, C. (2025). *On the interpretation of perspective-dependent expressions in the complement clauses of propositional attitude verbs* [Doctoral dissertation, University of Wuppertal].

Schlenker, P. (2003). A plea for monsters. *Theoretical Linguistics*, *26*(1), 29–120.

Wade, E. & H. H. Clark. (1993). Reproduction and demonstration in quotations. *Journal of Memory and Language*, *32*(6), 805–819. https://doi.org/10.1006/jmla. 1993.1040.

Walter, S. (2025). Indirect discourse as mixed quotation? An experimental investigation. *Proceedings of Experiments in Linguistic Meaning (ELM)*, *3*, 423–434. https://doi.org/ 0.3765/elm.3.5845.

Walter, S. & Hinterwimmer, S. (2025). Shifted face emoji in indirect discourse: A mixed quotational approach. *Proceedings of Sinn und Bedeutung (SuB)*, *29*, 1662–1680. https://doi.org/10.18148/sub/2024.v29.1301.