

Relating Gesture to Speech: reflections on the role of conditional presuppositions*

Julie Hunter

Abstract

In his paper ‘Gesture Projection and Cosuppositions,’ Philippe Schlenker argues that co-verbal gestures convey not at-issue content by default and in particular, that they trigger conditional presuppositions. In this commentary, I take issue with both of these claims. Conditional presuppositions do not supply a systematic means for capturing the semantic contribution of a co-verbal gesture. Some gestures appear to contribute content inside of a negation when their associated speech content is likewise embedded; in other cases, co-verbal gestures arguably contribute unconditional content to the global level. When this happens, we can infer what might look like a conditional presupposition, but this inference follows naturally from general principles already at work in purely verbal discourse and does not justify the claim that gesture content is contributed to a conditional presupposition. Problems exposed in the discussion of conditional presuppositions show that we are not yet in a position to make a general claim about the at-issue status of co-verbal gestures.

1 Combining gestures and speech

When we speak in a face-to-face setting, we use our hands, shoulders and faces as well as verbal signals to communicate our messages. Many of these gestures are arguably meaningful, and may even add substantive content to a spoken message. But how? How does their content interact with the content of verbal moves?

In his paper ‘Gesture Projection and Cosuppositions,’ Philippe Schlenker attempts to tackle a part of this problem by offering a proposal for how co-verbal, iconic gestures combine with their co-occurring speech¹ (Schlenker, 2018), and in particular, how “gestural enrichments of elementary expressions are inherited by complex sentences”. His proposal is, in a nutshell, that the content of a co-verbal gesture must be entailed by the local context of its co-occurring speech. In other words, the gesture triggers a presupposition to the effect that if the content of the co-occurring speech holds, then the content expressed by the gesture holds.

This proposal is motivated in part by arguments from Ebert and Ebert (2014) meant to show that co-verbal, iconic gestures contribute not-at-issue meanings by default. Schlenker adopts this conclusion (allowing limited exceptions), but takes issue with Ebert & Ebert’s proposal regarding how the not-at-issue contributions of co-verbal gestures should be modelled. Whereas Ebert &

*I gratefully acknowledge support from Juan de la Cierva fellowship IJCI-2014-22059. I would also like to thank Nicholas Asher, David Beaver, Alex Lascarides, Cynthia Matuszek, Kate Thompson, and the students of the ESSLLI 2017 course on Situated Discourse for helpful discussion. Finally, I thank the reviewers and editor of *Linguistics & Philosophy* for their constructive feedback.

¹I will use *co-occurring* here to mean roughly *temporally simultaneous*, but this glosses over the fact that identifying the segment of speech to which a co-verbal gesture relates is actually a complex task (Alahverdzhieva et al., 2018; Ebert et al., 2011).

Ebert suggest that the contribution of a co-verbal gesture is best modelled as a supplement, such as a non-restrictive relative clause, Schlenker argues that co-verbal gesture content contributes to presuppositions of a particular sort.

Consider the contrast between (1) and (2) below. Example (1) follows Schlenker’s conventions so that ‘LARGE’ indicates the relevant iconic gesture—in this case, we imagine a speaker holding one hand above the other as though the bottom hand were supporting the bottom of a large water bottle and the top hand were marking the top of the large water bottle. The speech content that is meant to co-occur with the gesture is indicated in bold.

- (1) Nobody brought a **LARGE_bottle of water** to the talk.
- (2) # Nobody brought a bottle of water, which was large, to the talk.

According to Schlenker, there are cases in which an iconic, co-verbal gesture can felicitously co-occur with embedded content, as in (1), but it cannot felicitously be replaced with a supplement, as illustrated in (2). The conflict in (2) arises because the content of the supplement is entailed at the global level, but the existence of the water bottle that the supplement content is meant to modify does not. If (1) is acceptable, then, the contribution of the gesture must be modelled with a mechanism that would keep its content from being entailed at the global level while remaining not at-issue. Accordingly, Schlenker proposes that a co-verbal gesture triggers a presupposition to the effect that the gesture content conditionally depends on the content of the co-occurring speech. This is meant to allow that co-verbal gestures coupled with unembedded speech content will be entailed at the global level, just as supplements are, but those coupled with embedded speech content need not be. (1), then, does not entail that anything is actually large, but rather that if there are water bottles brought to the talk, they are large. (It just happens in this case that no such bottles were actually brought.²)

In what follows, I take issue with both the proposal that co-verbal gestures are best modelled as triggering conditional presuppositions and with the claim that co-verbal gestures contribute not-at-issue content by default. Section 2 takes on the first point, arguing that conditional presuppositions fail to provide an analysis of the contribution of co-verbal gestures. On the one hand, some gestures appear to contribute content inside of a negation when their associated speech content is likewise embedded, but Schlenker’s account cannot explain this. On the other hand, some co-verbal gestures arguably contribute unconditional content to the global level. When this happens, we can infer what might look like a conditional presupposition, but this inference follows naturally from general principles already at work in purely verbal discourse and does not justify the claim that gesture content always triggers a conditional presupposition. Section 3 argues that many of the issues brought out in section 2 show that we are not yet in a position to make a general claim about the at-issue status of co-verbal gestures.

2 Conditional presuppositions

Schlenker’s account predicts that the presence of a co-verbal iconic gesture systematically triggers a conditional whose antecedent consists of the co-occurring speech content and which is such that: (i) the co-verbal gesture necessarily contributes its (whole) content to the consequent of the conditional, ensuring that the gesture’s contribution remains local in embedded examples, and (ii) the conditional suffices to capture the truth-conditional contribution of the gesture in embedded environments. In this section, I argue that the proposal falls short on both counts.

²The contrast between (1) and (2) alone is insufficient to undercut a supplement analysis, and Schlenker offers a more nuanced argument in other work, but the takeaway point is that we need to consider how co-verbal gestures behave when they co-occur with speech content in the scope of operators.

2.1 On the sufficiency of the conditional presupposition

Consider the following example from the AMI corpus (dialogue ES2005b), discussed in Lascarides and Stone (2009b).

- (3) D: And um I thought not too edgy and like a box, more kind of hand-held more um . . . not as uh [computery] and organic, yeah, more organic shape I think.
When D says ‘computery,’ her right hand has fingers and thumb curled downward (in a 5-claw shape), palm also facing down, and she moves the fingers as if to depict typing.

As pointed out in Lascarides and Stone (2009b), (3) has a reading in which it is not only the content of “computery” that is understood to be under the scope of the negation in (3), but also the content of the gesture—the speaker doesn’t want something with a keyboard.³ But how do we get the gesture content under the scope of the negation in Schlenker’s account? Negation of the antecedent of a conditional does not justify negation of the consequent in any logic. And it does not matter where the conditional *if x is computery then x has a keyboard* is accommodated; for a conditional of the form $p \rightarrow q$, neither the formula $(\neg p \wedge (p \rightarrow q))$ nor the formula $\neg(p \wedge (p \rightarrow q))$ supports an inference to $\neg q$.

Part of the problem is that in an effort to ensure that the gesture content remains not-at-issue, Schlenker’s analysis ends up rendering the gesture discursively inert. His account would yield the following gloss of (3): *If it were computery, it would have a keyboard, and I don’t want something computery.* But this gloss is awkward: why bother explaining what something computery would be like if the speaker isn’t interested in having something computery in any case? How things *would be* if the speaker had a different set of desires is irrelevant. (3) could be more appropriately glossed as follows: *I don’t want something computery, and what I mean by that is that I don’t want something with a keyboard.* In this case, the fact that the gesture content elaborates on the speech content and serves to make it more precise is discursively relevant. By highlighting the relevant undesirable feature of computery things, it ends up explaining why the speaker does not want something so “computery”.

2.2 On the necessity of the conditional presupposition

Let’s begin by considering a set of examples from Tieu et al. (2017), which Schlenker discusses. In this set of examples, the speaker points up (or down) while saying something like:

- (4) None of these three girls will UP **use the stairs**.

The acceptability of (4) does not obviously motivate a cosuppositional analysis. If on acceptable readings of (4), the gesture is interpreted deictically as a (possibly deferred) demonstration of

³There is another reading of this example in which the typing gesture does not constrain the meaning of “computery” in the context, but is better understood as a kind of default gesture, illustrating a feature often associated with computers. On this reading, something very box-like might still qualify as “computery,” and therefore be rejected by the speaker, even if it doesn’t have a keyboard. It is difficult to see how Schlenker’s account could be applied to such uses of co-verbal gesture.

One reviewer suggested distinguishing a *concept* interpretation of an iconic gesture from an *object-related* interpretation (Ebert et al., 2011; Fricke, 2012) and evaluating Schlenker’s proposal only relative to object-related interpretations, where a concept interpretation is what I have described as a kind of default interpretation, and an object-related interpretation is tied to a particular occasion. I focus on object-related interpretations in the body of this paper; however, the possibility of the concept interpretation highlights an important question: how do we tease these readings apart empirically? That is, in a conversation, how can we tell or how do we decide which reading the speaker is committed to? And perhaps more importantly, how often does it actually matter? I return to this point in the discussion of at-issue content in Section 3.

where the students are meant to or expected to end up, then the contribution of the gesture will be globally, not merely locally, entailed. On this reading, (4) behaves like (5).

- (5) The girls need to go up, but none of them will use the stairs. (They only use them to go down.)

In (5), the global context for the content in the scope of *but* settles the question of which direction the girls will be going; this is why saying that the girls will not take the stairs is compatible with saying that they use the stairs to go down. Similarly, on a deictic reading, (4) is felicitous in a context in which it is antecedently expected or desired, etc., that the girls go upstairs, and the speech content serves to exclude the possibility that they will go up by taking the stairs. In this case, we can infer the conditional *If the girls use the stairs, they will use them to go up*, but we do not need a conditional presupposition to explain this: the conditional is entailed by other information in the context.

A parallel argument can be made for at least some examples involving iconic gestures. Suppose my daughter has been busy working on craft projects, and she suddenly comes in the room and asks (6), making a slight scribbling gesture as she says *paper*:

- (6) Do we have any more SCRIBBLE.paper?

Based on this utterance, I would probably infer something that looks like a conditional presupposition: *if we had paper, she (or someone?) would write/draw (on it)*. But just as with (4), a purely linguistic example would support the same inference:

- (7) I want to write/draw something. Do we have any paper?

From (7), we can infer *if we had paper, she would write/draw on it*, yet this example does not require a conditional presupposition; the conditional follows from reasoning about why the question was relevant in the context of the first sentence. The problem for Schlenker's account, then, is that if we assume that the content of the gesture in (6) is something along the lines of *I want to write/draw something*, then we do not need to posit that the gesture triggers a conditional presupposition for this case either. The gesture will contribute content at the global level, and as in (7), the conditional will follow for free.

In his discussion, Schlenker stipulates that the content of a gesture is given, and assumes that this stipulation is innocuous. The discussion of (6) and (7), however, shows that the choice is highly consequential because it can significantly change the way that we understand co-verbal gesture content as interacting with the content of its accompanying speech. Given that this is precisely the problem Schlenker set out to solve, the content of an iconic, co-verbal gesture is not something we can simply take for granted.

2.3 More than elaboration

In the introduction of his paper, Schlenker says that the “intuition we develop is that the context should guarantee that the co-speech gesture merely illustrates the expression it modifies”. Accordingly, he considers examples in which the gesture is meant to elaborate on, and thereby restrict the interpretation of, the content of its co-occurring speech.

While what it means to “illustrate an expression” is not clear, I take it that (6) is probably not the kind of example that Schlenker has in mind. I shifted to it because I find it more natural than Schlenker's examples (with the exception of (4), which I find natural, but perhaps only because I get a deictic reading of the gesture). Personally, I cannot get the desired reading for (1) (“Nobody brought a LARGE.bottle of water to the talk”), and I am not alone—I have

discussed these examples with numerous English speakers (native and non-native) who share my judgment. It is difficult for me to imagine a context in which I would naturally utter (1) while also producing the requisite gesture, unless I wanted the content of that gesture to be at-issue. For example, I might say, “Nobody brought a **LARGE.water bottle**. They only brought these **puny things**” and hold up a small water bottle while saying “puny things”. Otherwise, it might be possible to understand the gesture in (1) as, for example, illustrating the extremities of a water bottle without any commitment to size (even if the same gesture would naturally be taken to illustrate size in a different context). In this case, however, the gesture would arguably be a faint echo of the speech content, not an elaboration on or restriction of it, so this interpretation, too, would fail to serve as a good example for Schlenker’s proposal.

I find the examples in which gestures co-occur with verbs to be more acceptable, though still somewhat unnatural (especially if I try to interpret them as Schlenker suggests). (8) is one such example.

(8) John didn’t **SLAP.punish** his son.

In this case, Schlenker claims that we should infer not only that John didn’t punish his son, but also that if he were to punish him, he would slap him. More specifically, any event of John punishing his son would also be an event of him slapping his son.

As Schlenker notes, judgments about gestures can be especially delicate, and I am happy to assume the felicity of (8) for the sake of discussion. What I wish to emphasize here is how the discussion of (6) (“Do we have any more **SCRIBBLE.paper?**”) bears on the intended analysis of (8). Again, the conditional inference that we are meant to draw in (8) is: *if John were to punish his son, he would slap him*. The conditional inference inferred in (6) is of a different sort: *if we have some more paper, then I will write/draw on it*. In the latter conditional, the consequent does not restrict the content of the antecedent, but describes what the speaker will do if the antecedent is satisfied.⁴ For this reason, Schlenker could say that his account is not designed to handle (6), but that would be a shame. Thinking about (6) and how we might arrive at the associated conditional, rather than a conditional such as: *if we have paper, then we will have paper for writing/drawing*, brings out some general issues that are relevant for (8).

2.4 On the necessity of the conditional presupposition in (8)

The above discussion of (4) (“None of these three girls will **UP.use the stairs**”) and (6) showed that for many co-verbal gestures, we can understand the gesture as triggering the addition of *unconditionalized* content to the *global* discourse context in order to derive the requisite conditional inference. In other words, something that might at first look like a conditional presupposition triggered by the gesture might be better understood as a side effect of whatever semantic mechanisms the gesture content is actually contributing to.

We can extend this argument to (8). Consider the following examples.

(9) John didn’t punish his son. He didn’t slap him.

(10) Did John punish his son? Did he slap him?

Upon hearing (9), one might infer that if John were to punish his son, he would do so by slapping him. In fact, if we interpret the second clause as providing more information about

⁴One might object that the gesture does serve to restrict the meaning of *paper* so that we understand that the speaker wants drawing or writing paper. This might be a possible reading of (6), but I do not know what kind of conditional would be supported by this reading. To the extent that we naturally infer a conditional for (6), I think it is a conditional relevant to the purpose of the paper.

the first clause—i.e., *John didn't punish his son, and what I mean by that is that he didn't slap him*—this conditional inference is necessary. However, the conditional in this case follows from a non-conditional, elaboration relation that entails the truth of the two clauses. (10) on its own might support the same inference, at least in some contexts: a good reason for the speaker to follow up with a question about a more precise form of punishment is that the speaker thinks that punishment should generally involve slapping or she thinks this form of punishment would be a likely choice for John.

The question that immediately arises for (8) is: why not posit that the gesture in (8) allows us to infer content comparable or equivalent to *John didn't slap his son*, and then derive the conditional inference by appealing to the same mechanisms used to derive the conditional inference in (9)? The idea would be that in (8), recognizing that the gesture elaborates on the co-occurring speech content allows us to treat the gesture content as specifying the semantic speech content; that is, the event of punishing is specified to be an event of slapping. This has the consequence first that the gesture content must be assigned the same type as its co-occurring speech. A second consequence is that the negation must distribute over both the speech and gesture content; if events of punishing are specified or precisified to be events of slapping, then anything that is not an event of punishing is also not an event of slapping. So this proposal would entail that in such elaboration-style examples: (i) gesture content is a separate unit of discourse from the speech content; (ii) a logical operator like negation is interpreted distributively, (iii) the negated gesture content and the negated speech content are both inferrable from the elaboration, endowing the proposal with empirically distinct claims from Schlenker's.

This strategy would work for (3), as well, where the negation intuitively distributes over both *computery* and the content of the typing gesture. By contrast, the strategy would fail for (6) (“Do we have any more **SCRIBBLE.paper?**”): in this example, we do not want the interrogative operator to distribute over the gesture content. The speaker is not asking what she should do with the paper. This failure, however, does not undermine the suggestion of applying the strategy to examples like (8). In (6), we do not interpret the gesture as elaborating on the content of its co-occurring speech but as revealing the purpose that the paper is meant to serve. Accordingly, we infer a different sort of conditional (*if we had paper, she (someone?) would write/draw on it*). These interpretive differences arguably justify a different analysis.

One of the central claims of Lascarides and Stone (2009a) and Lascarides and Stone (2009b) is that co-verbal gestures contribute independent discourse units that can enter into a variety of discourse relations,⁵ and Lascarides and Stone provide a wide array of data to support this claim. Seen from the point of view of such a theory, the fact that an iconic gesture could interact with speech in different ways, which would in turn support different inferences, is to be expected.

What is also expected from the point of view of a discourse-level account is that the presence of an embedding operator can change the relation between the very same pair of contents. That is, where s is a unit of speech content and g is the content of an iconic gesture performed simultaneously with s , we should expect that the relation between s and g might change depending on whether s is embedded or not. Consider the following contrast between two purely linguistic examples.

- (11) The play is starting in 10 minutes. The lead actor just fainted from stage-fright.
 (12) The play is not starting in 10 minutes. The lead actor just fainted from stage-fright.

If we imagine a context in which the play is actually scheduled to start in ten minutes (in order to exclude a reading of (11) in which the start time has been pushed back due to the lead actor's

⁵For the purposes of this commentary, a discourse unit can be understood as a clause that is capable of entering into a discourse relation such as Explanation, Narration, Contrast, or Elaboration.

fainting), then I think the most accessible reading of (11) is one in which the actor’s fainting is understood as a *result* of his impending entrance. In (12), by contrast, the fainting is much more easily understood as an *explanation* of the state described by the preceding sentence.

In some cases, embedding one discourse argument under a logical operator might preclude a relation with another discourse unit altogether. This is the case for so-called *entity-elaborations*. One of the foundational claims of DRT is that at least in standard, non-contrastive cases, embedding a discourse referent under negation renders that discourse referent inaccessible for subsequent discourse units.

(13) I brought a bottle of water to the talk. It was large.

(14) No one brought a bottle of water to the talk. # It was large.

In this example, we cannot pick up on the discourse referent introduced by *a bottle of water* in such a way as to add information about the non-existent bottle of water. Nevertheless, we can elaborate on the negative state as a whole.

(15) No one brought a bottle of water to the talk. In fact, no one brought any drinks of any sort.

When an event description is embedded under a negation, it yields a state which will be accessible for discourse continuations. When an indefinite is truly in the scope of a negation, however, the discourse referent that it introduces will not generally be accessible for discourse continuations because there is no entity to talk about. Eventuality-elaborations and entity-elaborations therefore give rise to difference discursive effects (Prévoit et al., 2009).

This difference should lead us to expect a difference between examples in which a co-verbal gesture elaborates on an eventuality described in speech, as in (16), and examples in which a co-verbal gesture elaborates on an entity, as in (17).⁶

(16) John SLAP_punished his son.

(17) Someone brought a LARGE_bottle of water to the talk.

Specifically, we should expect that the negated variant of (16)—namely (8)—should be more felicitous than (1), the negated variant of (17). While Schlenker apparently finds (1) acceptable, the prediction that it is not conforms with my intuitions and those of other English speakers with whom I have discussed the example, as well as with Ebert & Ebert’s analysis.

Schlenker set out to determine how “gestural enrichments of elementary expressions are inherited by complex sentences”. What the foregoing discussion shows, however, is that from the point of view of a discourse-level account, the presupposition that co-verbal gesture content will interact with accompanying speech content in a uniform way and that this type of interaction will be inherited by complex sentences is unfounded. We should expect the content of a co-verbal gesture to be able to interact with speech content in a variety of ways, and we should expect these interactions to be influenced by the presence of embedding operators.⁷ The examples discussed in this commentary conform to this prediction.

⁶The requirement that the gesture contribute something of the same type as the argument on which it elaborates will only hold for eventuality elaborations. Entity elaborations, in which an eventuality is being used to elaborate on an object, will of course not have this requirement.

⁷Schlenker claims that Lascarides and Stone do not consider the projection problem for gesture content, but this is inaccurate. In a discourse theory of the sort Lascarides and Stone develop, projection of presupposed content is a special case of the more general phenomenon of discourse attachment, which is one of the main problems Lascarides and Stone are trying to solve for co-verbal gesture content.

The need to treat co-verbal gesture content as an independent discourse unit that can combine in different ways with speech content is concealed by Schlenker’s focus on elaboration-style examples. This artificial restriction obscures the fact that the conditionals he analyses as presuppositions are actually artefacts of the discourse relation at work. This becomes clear when we change the style of example and see that Schlenker’s account does not provide a systematic way of deriving the conditionals that we can infer from examples of co-verbal gesture. Conditional presuppositions therefore do not capture the fundamental relation between a co-verbal gesture and its accompanying speech but are more accurately viewed as a side effect of inferring such a relation.

3 Gestures and not at-issue content

I have raised the possibility that an example such as (18) might be more appropriately modelled by appealing to whatever mechanisms underlie the interpretation of (9), repeated as (19).

- (18) John didn’t SLAP **punish** his son.
(19) John didn’t punish his son. He didn’t slap him.

Now imagine that after someone, say Susan, utters (19), someone else, say Pam, responds with (20):

- (20) Yes, he did. (He took away his ice cream.)

On the assumption that Pam had heard and processed all of (19), “Yes, he did” would likely be understood to mean “Yes, he slapped him,” which would make the continuation “He took away his ice cream” awkward. (“He took away his ice cream, *too*” might be better.) The content of “He didn’t slap him” in (19) is thus arguably at-issue.

Now imagine that Susan produces (18), rather than (19), and Pam responds with (20). How would “Yes, he did” be interpreted? Personally, I do not have clear intuitions about how Pam’s response would be interpreted in this case; my approach will therefore be to consider the two obvious possibilities and the consequences that would follow from each. One possibility is that Pam’s response would be understood as “Yes, he slapped him”. In this case, (18) would behave like (19), and we could conclude that the content of SLAP in (18) is at-issue.

The other possibility is that (20) would be interpreted as, “Yes, he punished him,” and in fact, I would guess that this interpretation would be more accessible than “Yes, he slapped him,” although that is just a hypothesis. The question is, if this hypothesis were correct, would it justify the conclusion that the gesture indicated by ‘SLAP’ contributes not-at-issue content in (18)? Not immediately. There are at least three concerns that should be addressed first.

For one thing, if we treat the gesture content as an independent discourse unit, as I, following Lascarides and Stone, have suggested, then we will need to consider how the combination of the gesture content with speech content will influence discourse continuations. The order in which two discourse units update a discourse context can have implications for possible discourse continuations (Hunter and Asher, 2016; Jasinskaja, 2016; Syrett and Koev, 2015). For example, the set of discourse continuations supported by (19) might differ from those supported by (21):

- (21) John didn’t slap his son. He didn’t punish him.

Given that with co-verbal gesture, there are arguably two discourse units produced simultaneously, it is an open question how they will affect discourse structure. Were we to conclude that (18) functions discursively more like (21), then it would be true that the gesture contributes not-

at-issue content, but not in the sense that Schlenker or Ebert & Ebert have in mind. It would not be because the gesture contributes a presupposition or a supplement, but simply because it contributes a discourse unit that attaches to the discourse context in a way that renders it less accessible for continuations.

A second concern is that for gesture content to be considered not-at-issue in the discourse, it should be reasonable for us to assume that an interlocutor becomes committed to the gesture's content if she does not challenge it (on the assumption that she witnessed the gesture). It is not enough for a speaker to make a gesture; the gesture content needs to be incorporated into the shared discourse context in the right way. In other words, we need to take seriously the distinction between *failing to contribute at-issue content* and *successfully contributing not-at-issue content* when considering interpretive judgments about gesture content.

To illustrate the relevance of this distinction, consider the tests that Ebert & Ebert used to argue that co-verbal gestures contribute not-at-issue content by default. Imagine that I utter:

(22) I brought a **LARGE.bottle of water** to the talk.

Ebert & Ebert claim that a direct denial of the gesture content is unacceptable:

(23) # That's not true! You actually brought a small bottle.

Yet (23) could also be infelicitous if the gesture contributed a content other than the content that Ebert & Ebert associate with **LARGE** or if it contributed no content at all. This exchange therefore only supports the weaker claim that the gesture fails to contribute the at-issue content that the bottle was large, not the stronger claim that it succeeds in contributing the relevant content to the not-at-issue level. Similar remarks can be made for the exchange in (24) and (25). Ebert & Ebert claim that (24) cannot be felicitously followed with (25):

(24) I did not bring a **LARGE.bottle of water** to the talk.

(25) # A small one is enough for me.

Again, the infelicity of (25) is consistent with the gesture in (24) not actually providing the content that Ebert & Ebert associate with **LARGE** in the first place.⁸

While I do not mean to equate co-verbal gesture content with implicatures, it is worth noting that implicatures cannot be targeted by corrections such as "That's not true!". To take a classic example:

- (26)
- a. Dorothy: Does Frank have a girlfriend?
 - b. Ernest: He's been making a lot of trips to New York lately.
 - c. Dorothy: Do you mean that he has a girlfriend in New York?
 - d. Ernest: Yes.
 - e. George: That's not true!

George's response in (e) is only felicitous once Ernest reinforces in (d) the implicature that Dorothy associates in (c) with Ernest's utterance in (b). The correction in (e) would not have been acceptable immediately after (b), but this is not because the implicature is contributing not-at-issue content—if anything, it is the implicature that is being offered as a potential answer to Dorothy's question and so should not have not-at-issue status; (e) would be an unacceptable response to (b) because the targeted implicature is not entailed by (b).

⁸I note in passing that according to Schlenker's account, the gesture in (24) should trigger a conditional presupposition along the lines of *if I had brought a bottle, it would have been large*, and this should render (25) infelicitous for independent polarities. Ebert & Ebert's point can be brought into harmony with Schlenker's account if we simply reverse the polarities of (24) and (25).

Another way of looking at the problem can be brought out by (27):

(27) I lost a **LARGE_water bottle**. It was large/**LARGE_this large**.

If the content that Schlenker and Ebert & Ebert associate with ‘LARGE’ can be made explicit without redundancy, then either (i) interpreters are not actually associating the content of ‘LARGE’ with the gesture or (ii) the gesture content is not contributing the content of a supplement:

(28) I lost a water bottle, which was large. ?? It was large.

The felicity of (27) would, I suppose, be consistent with Schlenker’s analysis:

(29) If I were to bring a bottle of water to the talk, it would be large. I brought a bottle of water to the talk. It was large.

The first two sentences of (29) already form an awkward discourse to my ear, but I don’t find that it gets worse by adding the final sentence. (Though what should we make of “I didn’t bring a **LARGE_water bottle** to the talk, but if I had brought a water bottle, it would have been large”? I would think that given Schlenker’s analysis, it should sound equivalent to: “If I were to bring (had brought?) a bottle of water to the talk, it would be (would have been?) large. I didn’t bring a bottle of water to the talk, but if I had brought a water bottle, it would have been large”.) Still, the felicity of (27) does not amount to support for a presuppositional analysis because it is also consistent with option (i).

Because gestures are frequently performed unconsciously, and because they have such underspecified meanings on their own (Goldin-Meadow and Brentari, 2016), we need to be cautious when applying the tests that we employ to judge the acceptability of linguistically specified content. How clear can we expect our intuitions to be about movements that we generally don’t even realize that we’re making? And what are we supposed to conclude from acceptability judgments concerning co-verbal gestures if we aren’t sure what content, if any, an interpreter is associating with that gesture? (When making judgments involving co-verbal gestures, an interpreter should of course be viewing the gesture, not reading something like ‘LARGE’ on a sheet of paper, so it is an open question what content the interpreter is associating with the gesture when she makes her judgment.) The reliability of elicited judgments for discursive phenomena is already notoriously questionable; a creative enough interpreter can often find a context to make even the strangest of examples acceptable. We cannot take acceptability judgments about discourse level phenomena at face value, but must also understand the contextual features that support a judgment of (un)acceptability. The underspecified meanings of gestures and their generally unconscious production only compound this concern about the reliability and usefulness of acceptability judgments because these factors shift so much of the weight of interpretation on the addressee (similarly to the case of implicature).

To encourage a particular reading of a gesture or to urge an interpreter to pay attention to it, it perhaps helps to make the gesture discourse relevant.⁹ This might at least push an object-related interpretation over a concept interpretation (Ebert et al., 2011; Fricke, 2012, cf. footnote 3). Imagine that someone, Matthew, is looking for his lost water bottle. He says,

(30) I lost a **LARGE_water bottle**.

⁹The gesture associated with ‘LARGE’ in Ebert & Ebert’s slides is so exaggerated that it would probably encourage the desired reading, but I think it is important that the discussion also encompass less exaggerated gestures that conform to the description I gave for (1). I assume that many of the iconic gestures that we produce for size or shape are produced unconsciously and are therefore less exaggerated.

In this case, an interpreter might naturally take Matthew’s gesture to have an object-related interpretation that commits him to the content that his water bottle was roughly the size of the space between his hands, as such descriptive information would be helpful for finding the water bottle. In this case, I agree that “That’s not true! It was smaller than that” would be odd (although “No, it was more like MEDIUM.**this size**” doesn’t strike me as so bad). And perhaps Ebert & Ebert are correct that someone could react to (30) (or to (22)) with (31):

(31) Hey, wait a minute! Actually, the bottle was not that big.

“Hey, wait a minute!” strikes me as a bit grandiose for the circumstances, but I take the point that it is acceptable for someone to interrupt the flow of conversation with something like “Wait” or “Hold up” in order to clarify the size of the water bottle.

Given that the *Hey, wait a minute* test (von Stechow, 2004, cf. Shanon (1976)) is designed to target not-at-issue content, the discussion in the preceding paragraph suggests that the gesture in (30) actually succeeds in contributing not-at-issue content. Still, I think caution is in order. First of all, Schlenker (2018) maintains that post-speech gestures are generally at-issue, but consider:

(32) I brought a water bottle to the talk – LARGE
 a. That’s not true! It was smaller than that.
 b. Hey, wait a minute! It wasn’t that big!

The response in (b) strikes me as more natural than (a).¹⁰ At least, it is not nearly as clear that (a) is better than (b) as it would be if we had taken (a) and (b) as responses to (33).

(33) I brought a water bottle to the talk. It was LARGE.**this** large.

This is unexpected if the content associated with ‘LARGE’ is at-issue and the tests for at-issue content are working as they are meant to.

The *Hey, wait a minute* construction tells an interpreter to stop and take a step back. This makes it a useful tool when a speaker wants to go back and take issue with a kind of projective content that was antecedently added to the discourse. But it can also be useful when a speaker is unsure about the nature of the content under discussion. To take a very basic toy example as an illustration, consider (34):

(34) a. Andy: Where did Doris and Ernest meet?
 b. Bill: They met near a bank.
 c. Carmen: Hey, wait a minute, do you mean a financial bank or a river bank?¹¹
 d. Bill: a financial bank.
 e. Carmen: Well then that’s not true. They met near a river.

In (c), Carmen interrupts the flow of conversation to check that she is clear on what Bill means—if an interlocutor lacks confidence on what content a speaker is committing herself to, then she should stop and check. Similarly, in (31), “hey, wait a minute” might be a useful means of

¹⁰I think this intuition is even clearer with the SLAP examples:

(i) John punished his son – SLAP.
 a. ?That’s not true! He grounded him.
 b. Hey, wait a minute! Do you mean that he slapped him?

Response (b) seems far more natural than (a).

¹¹As for (31), “Hey, wait a minute” is probably too dramatic for the circumstances—something like “wait” or “hold up” might be more appropriate—but because I find it as natural here as I do in (31), I’ll leave aside this concern.

clarifying information about the size of the bottle while recognizing that the speaker of (22) is not fully committed to the content that the bottle was the size indicated by her gesture. This possibility must at least be factored in when drawing conclusions from the *Hey, wait a minute* test in examples like (31).

Putting aside concerns about intuitions, there is a third problem for the claim that gestures contribute not-at-issue content by default that is tied to Schlenker's particular analysis. Returning to the discussion of (18) and (20), one could imagine testing experimentally for the interpretation of (20) as follows. Give subjects examples of (18) followed by Pam's utterance of "Yes, he did" (i.e., (20) without the continuation in parentheses) and ask them if they can infer from the correction that John slapped his son or at least that Pam is committed to John's having slapped his son. If they do not infer this with a significant level of reliability, this result would be consistent with the conclusion that the gesture contributes not-at-issue content. Yet it might also mean that the subjects are simply ignoring the gesture content altogether, or that they are interpreting (18) in such a way that it does not entail: *if John punished his son, then he slapped his son*. Such a result would be inclusive at best for Schlenker's proposal. If, on the other hand, the subjects were to infer that John slapped his son, this would be consistent with the presence of the purported conditional presupposition. Unfortunately, it would also be consistent with the hypothesis that the gesture contributes the at-issue content that John slapped his son.

In fact, because Schlenker's account ensures that punishing entails slapping (in a context C), it will always be difficult to pry apart the content of a co-verbal gesture from the content of its accompanying speech in an elaboration-style example. If for some/all events e , the fact that e is an event of John punishing his son entails that e is an event of John slapping his son, then we can conclude that in any world compatible with the context C, if e is an event of John punishing his son, e is an event of John punishing his son and slapping his son. ($P \rightarrow Q$ entails $P \rightarrow (P \wedge Q)$.) But it's a logical truth that if e is an event of John punishing his son and slapping his son, then e is an event of John punishing his son, so we derive a biconditional dependence between John punishing his son and his punishing and slapping his son in C. Put another way, intuitively what we would infer from (18) (and (19)) is the conditional: *if John punished his son, he punished him by slapping him*. But if John punished his son by slapping him, this entails that John punished his son. This biconditional dependence will ensure that the two contents are interchangeable within the scope of operators like negation, conditionals or questions in C. From that it should follow that the co-verbal gesture content is at-issue just in case the speech content is.

Of course, to the extent that examples of co-verbal gesture support conditional inferences that have the form of conditional presuppositions—as I have agreed that they at least often do—*these conditionals* should certainly be not-at-issue. Where I have targeted Schlenker's proposal is in the claim that a co-verbal gesture will contribute the whole of its content to the consequent of such a conditional. This claim, which I argued against for independent reasons in the previous section, is motivated by the premise that co-verbal gestures contribute not-at-issue content by default. The goal of this section has been to call this motivating premise into question.

4 Moving forward

It is widely accepted that gesture and speech work together to form a unified message (Kendon (2004), McNeill (1992)), and there is certainly an important role for formal semantics to play in figuring out how they coordinate. I do not think, however, that Schlenker's account succeeds in modelling this interaction. Schlenker proposes that a co-verbal gesture contributes content that is conditional on its co-occurring speech content, but I have argued that his account does not outline a systematic way of predicting how these conditionals will look, and that the conditional

inferences that we draw arguably follow from more general discursive principles that are already at work in purely verbal discourse. A discourse-level approach also does a better job at capturing examples in which operators intuitively distribute over both a gesture and its co-occurring speech content. Finally, I have called into question one of the main premises that motivated the hypothesis that a co-verbal gesture contributes the whole of its content to the consequent of a conditional presupposition, namely the premise that co-verbal gestures contribute not-at-issue content by default.

A more general take-home message is that while attempting to capture systematic relations between speech and gesture content is certainly a worthwhile and much-needed enterprise, the set of tools that we have to choose from to model these interactions is larger than the set normally employed in formal semantics. Moreover, determining which of these tools is appropriate and when is going to require more empirical work (building on, e.g., Kendon, 2004; Lücking et al., 2013; McNeill, 1992) so that we have a better understanding of how gestures get endowed with content in the first place. Even experimental studies designed to test for a specific relation between speech and gesture, such as Tieu et al. (2017), need to be approached with caution unless the experimental designers can be sure that they know the content of the gestures they are dealing with. And the content will need to be determined naturally, not by stipulating a content that the designers would like the gesture to have—calling a horse’s tail a *leg* will not change the fact that a normal horse has four legs. It is necessary that we be able to reliably associate a reasonably unambiguous content with a gesture if we wish to test how this content interacts with speech content.

Compliance with Ethical Standards

1. Conflict of Interest: The author declares that she has no conflict of interest.
2. Research Involving Human Participants and/or Animals: The author declares that the research in this paper did not involve human participants or animals.
3. Informed Consent: (not applicable, as the research did not involve human participants)

References

- Alahverdzhieva, K., A. Lascarides, and D. Flickinger (2018). Aligning speech and co-speech gesture in a constraint based grammar. *Journal of Language Modelling* 5(3), 421–464.
- Ebert, C. and C. Ebert (2014). Gestures, demonstratives, and the attributive/referential distinction. Talk given at Semantics and Philosophy in Europe (SPE 7), Berlin, June 28, 2014.
- Ebert, C., S. Evert, and K. Wilmes (2011). Focus marking via gestures. In *Proceedings of Sinn and Bedeutung 15*, Saarbrücken.
- Fricke, E. (2012). *Grammatik multimodal: Wie Wörter und Gesten zusammenwirken*. Berlin/Boston: de Gruyter.
- Goldin-Meadow, S. and D. Brentari (2016). Gesture, sign and language: The coming of age of sign language and gesture studies. *Behavioral and Brain Sciences*.
- Hunter, J. and N. Asher (2016). Shapes of conversation and at-issue content. In M. Moroney, C. Little, J. Collard, and D. Burgdorf (Eds.), *Semantics and Linguistic Theory (SALT) 26*, pp. 1022–1042.

- Jasinskaja, K. (2016). Saliency and (not-)at-issue status of subordinate clauses. In *Sinn und Bedeutung* 21. forthcoming.
- Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.
- Lascarides, A. and M. Stone (2009a). Discourse coherence and gesture interpretation. *Gesture* 9(2), 147–180.
- Lascarides, A. and M. Stone (2009b). A formal semantic analysis of gesture. *Journal of Semantics* 26(4), 393–449.
- Lücking, A., K. Bergman, F. Hahn, S. Kopp, and H. Rieser (2013). Data-based analysis of speech and gesture: The bielefeld speech and gesture alignment corpus (saga) and its applications. *Journal of Multimodal User Interfaces* 7(1-2), 5–18.
- McNeill, D. (1992). *Hand and Mind: What Gestures Reveal about Thought*. Chicago: University of Chicago Press.
- Prévot, L., L. Vieu, and N. Asher (2009). Une formalisation plus précise pour une annotation moins confuse: la relation d’élaboration d’entité. *Journal of French Language Studies* 19(2), 207–228.
- Schlenker, P. (2018). Gesture projection and cosuppositions. *Linguistics and Philosophy*.
- Shanon, B. (1976). On the two kinds of presuppositions in natural language. *Foundations of Language* 14, 247–249.
- Syrett, K. and T. Koev (2015). Experimental evidence for the truth conditional contribution and shifting information status of appositives. *Journal of Semantics* 32, 525–577.
- Tieu, L., P. Rob, P. Schlenker, and E. Chemla (2017). Co-speech gesture projection: Evidence from inferential judgments. unpublished manuscript.
- von Stechow, K. (2004). Would you believe it? the king of france is back! In M. Reimer and A. Bezuidenhout (Eds.), *Descriptions and beyond*. Oxford: Oxford University Press.