

A comparison of sign language with speech plus gesture¹

Commentary on Philippe Schlenker's

"Visible Meaning: Sign Language and the Foundations of Semantics"

Cornelia Ebert (Institut für Linguistik/Germanistik, Universität Stuttgart)

In the introduction to his target article Schlenker writes that „*sign languages provide overt evidence on crucial aspects of the Logical Form of sentences* that are only inferred indirectly in spoken language“ (p. 3) and furthermore that „*sign languages are strictly more expressive than spoken languages* because iconic phenomena can be found at their logical core“ (p.3). He further argues that one of the possible conclusions that can be drawn from these facts is that „spoken languages have comparable expressive mechanisms, but only when co-speech gestures are taken into account“ (p.3), as Goldin-Meadow & Brentari (2016) have recently claimed. In the following, I will elaborate on this possibility. Following Goldin-Meadow and Brentari (2016), I will show by reference to examples from the main text that a close comparison of sign language and spoken language under controlled conditions will not only tell us more about the semantics of gestures in spoken languages and semantics in general, but will also shed light on the notion of gesture within sign languages.

In section 6.1 Schlenker revisits some of the discussed sign language phenomena and draws parallels to cases of spoken language that is enriched by co-speech gestures. I will follow this path and discuss more such examples and the consequences that arise from this method. In particular, I will compare Role Shift in sign language to viewpoint gestures in spoken language, discuss loci and locative shifts in comparison to pointing gestures in spoken language, and finally speculate about the role of gestures in general (in sign and spoken language) and the semantic contribution they can make (concerning the semantic dimension they target, i.e. whether they are at issue or not).

1 Role Shift

In sections 3.1 and 5.3, Schlenker discusses the notion of Role Shift, an operation where the signer adopts the point of view of another individual (Attitude Role Shift) or reports in a particularly vivid way someone else's action (Action Role Shift) and marks this by turning the body in a different direction and/or by eyegaze shift. He quotes Sandler & Lillo-Martin (2006, p. 380), who note that Role Shift „can be described as quoting the thoughts of another or simply representing a scene from the point of view of another“ (pp. 22, 23, fn. 28) and argues, following Quer (2005), that Role Shift can be understood as a visible indication of context shift.

Interestingly, there is a branch in gesture research that is concerned with precisely this phenomenon, i.e. the overt indication of the perspective or the point of view the speaker adopts, which is made visible via an accompanying gesture (see Parill 2010 for an overview and Stec 2012 for a comparison of viewpoint markers in sign languages and with gestural means). It is assumed that gestures can reveal the perspective taken by a speaker to report a certain event (McNeill 1992). Iconic gestures that visualize certain aspects of an event can

¹ This manuscript has been submitted as a commentary to Philippe Schlenker's target article "Visible Meaning: Sign Language and the Foundations of Semantics" to *Theoretical Linguistics* on March 5 2017. I would like to thank Christian Ebert and Stefan Hinterwimmer for discussion and comments on an earlier version.

reveal a character viewpoint (C-VPT: the event is shown as if experienced by the character) or an observer viewpoint (O-VPT: the event is shown as if seen from a distance). For example, a sentence like *Jerry was running through the kitchen* can either be accompanied by an O-VPT gesture that indicates Jerry's trajectory, potentially in combination with an iconic gesture such as two fingers functioning as the legs imitating a running event, or by a C-VPT gesture mimicing the way Jerry ran, e.g. by quickly moving one's arms and legs as if one was running – with the entire body of the speaker embodying the character of the report. While there exists a large body of literature on the exact licensing conditions of O-VPT gestures vs. C-VPT gestures within one language (with e.g. thematic, narrative, or information structural restrictions) and across languages (with interesting differences in the frequency of the use of C-VPT gestures in Japanese and English, for example), what matters for us is the following: An O-VPT gesture indicates a bird's eye perspective, a C-VPT gesture suggests that the speaker adopts the perspective of the individual that is impersonated.

A C-VPT-gesture thus makes it directly visible what perspective the speaker is taking. If now Role Shift is an instance of context shift, but with the requirement of maximal iconicity, as suggested by Schlenker, and thus indicates that the speaker is taking over someone else's perspective (Sandler & Lillo-Martin 2006), C-VPT gestures seem to be the gestural co-speech equivalent to Role Shift in sign languages². It would now be interesting to see what happens to co-speech gestures in contexts that are known to give rise to context shift. In free indirect discourse, for example, a construction that most probably involves context shift (Banfield 1982, Schlenker 2004b), we would expect that iconic co-speech gestures are realized as C-VPT gestures that impersonate the protagonist of the thought context. As a consequence, it is predicted that pointing to oneself when uttering a 3rd-person pronoun would be felicitous in free indirect discourse (ex. 1a), while it should appear as odd in other contexts (ex. 1b).

- (1) a. Peter was upset. Why on earth had he done that?
 Why had [he] been so stupid?
 point to self
- b. Paula was annoyed with Peter. Why could he have done that?
 Why had [he] been so stupid?
 #*point to self*

Furthermore, it would be interesting to test whether the same mixing of perspectives is possible with co-speech gestures as Schlenker (2003) argues to be existent in Amharic and Russian or whether co-speech gestures rather shift together, as Anand and Nevins (2004) claim indexicals do in Zazaki. Potentially, languages even show various patterns of perspective-taking with co-speech gestures, just like there seem to be differences in the shifting behaviour of indexicals in spoken languages and different possibilities for mixing perspectives in sign languages.

Consequently, taking C-VPT gestures as indicators of context (or perspective) shift in spoken languages, the Role Shift data in (80)-(83) can be replicated with C-VPT vs. O-VPT co-speech gestures. Consider first examples (2a,b), which parallel Schlenker's Attitude Role Shift examples (80) and (81).

- (2) a. See that arrogant French swimmer? Yesterday he was angry.
 He said he [would leave].

² Note, however, that it is a well-known fact that narration not from the narrator's, but from a character's perspective, is much more prevalent in sign language than in spoken language. Consequently, Role Shift in sign language is more frequent than C-VPT gestures in spoken languages (see Rayman 1999).

O-VPT *WALK-gesture* (two fingers moving as to imitate walking)
:-)

b. See that arrogant French swimmer? Yesterday he was angry.

He said he [would leave].

C-VPT *WALK-gesture* (whole body impersonating a walking person)
:-)

The O-VPT gesture in (2a), which does not determine any fixed perspective, allows for combination with a happy face on the speaker's part. This facial expression can be interpreted as to mean that the speaker displays happiness about the fact that the French swimmer said he would leave. A happy face of the speaker in combination with the C-VPT gesture in (2b), however, appears infelicitous. This is even more notable for the fact that in principle it is possible to embody different characters with different parts of the body at the same time, e.g. handing something to oneself while uttering *she got something*, where the arm and hand embodies the giver and the rest of the body the receiver (cf. Cassell et al. 1999 for relevant discussion), or perform an O-VPT gesture and a C-VPT gesture simultaneously (Parill 2009).

By the same method, we can also replicate the Action Role Shift examples in (82) and (83) in spoken language with co-speech gestures.

(3) a. See that arrogant French swimmer? Yesterday he was angry.

He [left with energy].

O-VPT *walk-gesture* (two fingers moving as to imitate walking)
:-)

b. See that arrogant French swimmer? Yesterday he was angry.

He [left with energy].

C-VPT *walk-gesture* (whole body impersonating a walking person)
#:-)

Again, it seems that the C-VPT gesture in (3b) cannot be combined with a happy face, but an O-VPT (3a) gesture can. C-VPT gestures can thus be likened with Role Shift in sign languages and the same iconicity constraints seem to apply.

2 Loci

Turning to loci as another peculiarity of sign languages, we will see that, again, co-speech gestures can serve to detect certain parallels of sign and spoken languages that would not be observable otherwise. In section 6.1, Schlenker briefly discusses complement set anaphora and high loci in comparison to co-speech gestures. For high loci, he concludes that the correct view about height specifications might be to treat them as co-sign gestures that are merged with signs (see also Goldin-Meadow & Brentari 2016 for discussion). I would like to enlarge upon the discussion of plural anaphora and embedded loci and add a thought on locative shifts.

2.1 Embedded loci

Schlenker notes that the correct point of comparison for the cases of embedded plural loci of section 4.3 are not simple spoken language examples of complement set anaphora, but rather complement set anaphora cases where a speaker can additionally point to a diagram. For concreteness, let us revisit Schlenker's example (52) and its spoken variant in (4).

(4) Most students came to class. #They stayed home instead.

Let us furthermore assume that the speaker of (4) has a diagram as in (53) of the target article at her disposition where she can freely point to different areas in the picture. If now (4) is uttered with an additional pointing gesture to the set of students that did not come to class, the utterance suddenly becomes felicitous and thus comparable to the ASL example with embedded loci in (52b). With this assumption, *they* is not interpreted anaphorically any more, but deictically. Interestingly, this hypothesis is further supported by the fact that in German, one would have to use the d-pronoun/demonstrative *die* instead of the anaphoric pronoun *sie* if (4) is uttered and accompanied by a pointing gesture³. One could even go one step further and adopt the view that complement anaphora in general is a phenomenon that draws on (implicit) pointing. In those cases in which reference to the complement set is allowed (with the quantificational determiner *few* or *no* for example, cf. Nouwen 2003), German prefers the d-pronoun *die* over the ordinary pronoun *sie*, which might indicate that we are actually dealing with a demonstrative phenomenon and not with anaphora.

(5) Heute kamen nur ganz wenige Studenten zur Vorlesung. #Sie/Die sind lieber Eis essen gegangen.

Today, only very few students came to class. They preferred going for an ice-cream instead.

2.2 Locative shifts

Taking a closer look at locative shifts (section 4.5), we can again replicate the basic facts that Schlenker discusses for sign languages with speech-accompanying gestures. Having seen several sign language examples where a specific locus in the signer's signing area is assigned to a certain referent, it is worth noting that also in spoken language we have the possibility to let a referential expression *r* be accompanied by a pointing gesture to a specific location in the gesture space. This location is then associated with the referential expression *r* (cf. Kendon 2004). This procedure can be conceived of as an instance of Bühler's *deixis am Phantasma* (see Fricke 2009). Consider now example (6), that is designed along the lines of Schlenker's example (68).

(6) John owns an apartment in [a French city] and one in [an American city].
point to a *point to b*

- a. His apartment/One of his apartments is nice.
 ⇒ no inference about the apartment's location.
- b. [His apartment]/[One of his apartments] is nice.
point to b
 ⇒ John's American apartment is nice.

If the sentence *his apartment/one of his apartments is nice* follows the first sentence of (6), no inference about the apartment's location can be drawn. If this utterance is accompanied by a pointing gesture to the previously introduced location *b* as in (6b), it entails that it is the American apartment that is nice, just like the ASL expression in Schlenker's (68b) does.

A question that now arises and that will be discussed in the general case below is what status this additional information of specifying a location has – whether it is information that is at issue or not. In our example (6) it seems rather obvious that the additional information that is conveyed via the co-speech pointing gesture is information that is not at issue, as is generally assumed for co-speech gestures (Ebert 2008, Ebert & Ebert 2014, Schlenker to appear-a). To

³ Admittedly, the contrast could simply be due to the fact that d-pronouns are quite generally preferred in contexts where they refer back to less salient entities.

verify this, we can try to directly deny the information under discussion and we expect that such a move is deviant, because non-at-issue material is not directly deniable (Potts 2005). Hence, (7) should not be a felicitous response to (6b).

(7) No. His American apartment is actually horrible. His French apartment is nice.

And indeed, (7) appears to be an odd response to (6b). Rather, an objection to the not-at-issue information content of (6b) could be made via some discourse-interrupting objection such as by way of a ‘*hey, wait a minute*’-construal (see Potts 2012 for this test to address linguistic meaning components that are not at issue).

(8) Hey, wait a minute, it is in fact true that he has a nice apartment... But it is actually the French one that is nice, the American one is horrible.

This shows that the information about the location of the apartment is not at issue in the co-speech gesture example in (6b). The information status of the corresponding information in the ASL example in (68) still needs to be settled.

If the facts are similar in ASL as in English and the information about the location of the apartment is not at issue, Schlenker’s tentative analysis of ‘situation slices’ of individuals would have to take care about this fact. In other words, *IX-a* should not denote <John, Paris>, but rather formulate appropriate definedness conditions on its usage (along the lines proposed for high and low loci in section 4.4) or handle the non-at-issue status of the relevant information in some other way.

3 Information status

The distinction between sign and gesture in sign languages and the urge to develop strategies that tell them apart in systematic ways has recently been stressed in Goldin-Meadow & Brentari (2016). As gestures are performed in the same modality as signs, the two are often hard to tease apart. According to Goldin-Meadow & Brentari, a close comparison of spoken language enriched by gesture and sign language (which includes gestures) might be the key. Those forms generated by signers that resemble properties of gestures in spoken language, are likely to be gestures in sign language, as well. In the past, systematic comparative studies have been conducted with signers and “silent gesturers” (hearing non-signers that are asked to use their hands and not their voices to describe a certain event), but what is still missing are systematic studies that look at speech plus co-speech gestures in comparison to sign plus (yet to be identified) co-sign gestures. The cases that have been discussed above and in the target article are prime evidence that following this strategy – comparing sign and speech plus gestures in a systematic way, phenomenon by phenomenon – might indeed help to complete our picture about gestures and further our understanding of natural language semantics in general. For concreteness, we will make the following claims: 1. those information parts that are conveyed by speech-accompanying gestures in spoken language constructions are likely to be gestures in the corresponding sign language sequences, as well. And 2., gestures behave alike in sign and spoken language with respect to their information status.

Given this hypothesis, we expect that all information that is expressed via co-speech gestures in a given spoken language expression is necessarily non-at-issue also in the corresponding sign language construction, since co-speech gestures have been shown to contribute non-at-issue information in general, as discussed above. Schlenker (to appear-b), in a commentary to Goldin-Meadow & Brentari (2016), however, already showed that this is not the case, at least not always. In his commentary, he introduces the distinction between ‘internal and external enrichments’. Internal enrichments are defined as enrichments that operate on the form of an expression directly and modulate it in iconic ways. An example for an internal enrichment

would be the ASL GROW example of section 4.2 on p. 28 or the lengthening of a vowel in English (e.g. *looooooong*). Co-speech gestures on the other hand are external enrichments, because they iconically enrich the meaning of some external element. Schlenker goes on to suggest that internal enrichments at least can be at issue (although they are sometimes not), while external enrichments are generally not at issue. He then shows that the internal enrichments that are present in the different realizations of GROW (ex. 48, p.28) can be processed as at-issue information. He argues that

(9) IF POSS-1 GROUP GROW_{broad}, IX-b JOHN LEAD. (ASL, 33, 71; 2 trials)

has a prominent reading where the realization mode ('broad') is processed as at-issue material, leading to the reading: if the speaker's group grows a lot, John will lead it. If, however, we construct an English example with accompanying gesture along the lines of (9), the co-speech gesture would be interpreted non-at-issue.

Let us first replicate the facts of Schlenker's example (48) with the diverse realization forms indicated in (49) in English enriched by a co-speech gesture that accompanies the word *growing*.

(10) My group has been [growing].
GROW-gesture

Depending on the realization mode of the accompanying gesture (slow vs. fast movement, small vs. large distance between endpoints, ...), the same interpretational differences arise as observed for the ASL example in (48). Following Schlenker to appear-b, in English this would necessarily be an external enrichment, which is inevitably not at issue, while it would be an internal enrichment in ASL, which at least can be at issue, as was witnessed by example (9). And indeed the English counterpart to (9) does not seem to allow for an at-issue reading of the co-speech gesture.

(11) If my group [grows], John will lead it.
GROW_{broad-gesture}

(11) cannot mean that if the speaker's group grows a lot, John will lead it. But what do we make of this difference? Does this mean that co-speech gestures in fact behave in completely different ways from sign enrichments and that we should dismiss our attempt to parallel these two bearers of meaning altogether? I will argue that this conclusion would be all too hasty.

We know that meaning components that start out as non-at-issue material can end up being processed in the at-issue dimension. Ebert & Ebert (2014) treat demonstratives as dimension shifters that map speech-accompanying gestural material to the at-issue dimension (see also Fricke 2012 for an account of the German demonstrative *so* as place-holder for gestural material or Streeck 2002 for German *so* and English *like* as quotation markers that, among other things) serve to integrate gesture material into the semantic content of an utterance). Schlenker (to appear-a) discusses further cases, where information that starts out as non-at-issue information via a co-speech gesture can be turned into at-issue information by local accommodation. In particular, this is an option when there is contrastive focus on the expression that the gesture accompanies (see Esipova 2017 for an analysis). Ebert (2017) formulates restrictions on this possibility of turning non-at-issue meaning into at-issue meaning and hypothesizes that a co-speech gesture (without accompanying demonstrative) can only be mapped on the at-issue dimension if there are special facial markings like raising one's eyebrows and/or if the gesture is made especially expansive and eye-catching (cf. also Esipova 2017, fn. 2).

To verify this, consider (12), which allows for an at-issue reading of the co-speech gesture at least more readily than (11) does.

- Goldin-Meadow, Susan and Brentari, Diane: to appear, Gesture, sign and language: The coming of age of sign language and gesture studies. To appear as a target article in *Behavioral and Brain Sciences*.
- Kendon, Adam: 2008. *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.
- Nouwen, Rick: 2003. Complement Anaphora and Interpretation. *Journal of Semantics* 20/1: 73-113.
- Parrill, Fey (2010): Viewpoint in speech–gesture integration: Linguistic structure, discourse structure, and event structure, *Language and Cognitive Processes* 25/5: 650-668.
- Parrill, Fey (2010): Dual Viewpoint Gestures. *Gesture* 9/3: 271-289.
- Potts, Chris (2012): Conventional implicature and expressive content. In C. Maienborn, K. von Heusinger & P. Portner (eds.), *Semantics: An International Handbook of Natural Language Meaning*, Vol. 3, 2516-2536. Berlin: Mouton de Gruyter.
- Potts, Chris: 2012. Conventional Implicature and Expressive Content. In C. Maienborn, K. von Heusinger & P. Portner (eds.), *Semantics: An International Handbook of Natural Language Meaning*, Vol. 3, 2516-2536. Berlin: Mouton de Gruyter.
- Potts, Chris: 2005. *The Logic of Conventional Implicatures*. Oxford: Oxford University Press.
- Quer, Josep: 2005, Context shift and indexical variables in sign languages. In: *Proceedings of Semantic and Linguistic Theory (= SALT) XV*. Ithaca, NY: CLC Publications.
- Rayman, Jennifer (1999): Storytelling in the visual mode: a comparison of ASL and English. In E. Winston (ed.), *Storytelling and conversation: Discourse in deaf communities*, 59-82. Washington D.C., Gallaudet University Press.
- Sandler, Wendy & Lillo-Martin, Diane: 2006, *Sign Language and Linguistic Universals*. Cambridge University Press.
- Schlenker, Philippe: to appear-a. Gesture Projection and Cosuppositions, to appear in *Linguistics & Philosophy*.
- Schlenker, Philippe: to appear-b. Iconic Enrichments: Signs vs. Gestures [Short Commentary]. Commentary on Goldin-Meadow and Brentari's target article, to appear in *Behavioral and Brain Sciences*.
- Schlenker, Philippe: 2004b. Context of thought and context of utterance, *Mind & Language* 19/3: 279-304.
- Schlenker, Philippe: 2003. A Plea for Monsters, *Linguistics & Philosophy* 26: 29-120.
- Stec, Kashmiri: 2012. Meaningful shifts: A review of viewpoint markers in co-speech gesture and sign language. *Gesture* 12/3: 327–360.
- Streeck, Jürgen: 2002. Grammars, words, and embodied meanings. On the evolution and uses of so and like. *Journal of Communication* 52/3: 581-596.