



**HAL**  
open science

## Sign Language Semantics: Problems and Prospects

Philippe Schlenker

► **To cite this version:**

Philippe Schlenker. Sign Language Semantics: Problems and Prospects. *Theoretical Linguistics*, 2018, 44 (3-4), pp.295-353. 10.1515/tl-2018-0022 . hal-03052044

**HAL Id: hal-03052044**

**<https://hal.science/hal-03052044>**

Submitted on 10 Dec 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Sign Language Semantics: Problems and Prospects\*

Philippe Schlenker\*\*  
(Institut Jean-Nicod, CNRS; New York University)

Typos corrected, October 22, 2018. To appear in *Theoretical Linguistics*.

**Abstract.** 'Visible Meaning' (Schlenker 2018b) claims (i) that sign language makes visible some aspects of the Logical Form of sentences that are covert in spoken language, and (ii) that, along some dimensions, sign languages are more expressive than spoken languages because iconic conditions can be found at their logical core. Following nine peer commentaries, we clarify both claims and discuss three main issues: what is the nature of the interaction between logic and iconicity in sign language and beyond? does iconicity in sign language play the same role as gestures in spoken language? and is sign language Role Shift best analyzed in terms of visible context shift, or by way of demonstrations referring to gestures?

**Note:** While this article is intended to be self-contained, it originated in replies to peer commentaries on 'Visible Meaning', a target article in *Theoretical Linguistics* 2018; reference is made throughout to 'Visible Meaning', and to the commentaries.

Keywords: sign language, semantics, iconicity, gestures, role shift, co-speech gestures, pro-speech gestures

---

\*I am extraordinarily grateful to the ten colleagues who commented on 'Visible Meaning'; I learned a lot from their very detailed and remarkably interesting commentaries. Commentators were: Valentina Aristodemo and Mirko Santoro; Diane Brentari; Kathryn Davidson; Cornelia Ebert; Regine Eckardt; Karen Emmorey; Emar Maier; Josep Quer; Sandro Zucchi. Many thanks as well to Brian Buccola, Emmanuel Chemla, Rob Pasternak and Lyn Tieu for regular and patient help pertaining to my work on gestures; helpful remarks were also made by Masha Esipova. In addition to their great contribution as commentators, Diane Brentari and Kathryn Davidson also sent me very helpful comments on a preliminary version of the present piece. (The sections on Role Shift indirectly benefited from discussions about shifted indexicals with Amy Rose Deal.)

**ASL consultant for data cited in this article: Jonathan Lamberton.** Special thanks to Jonathan Lamberton, who provided exceptionally fine-grained data throughout this research; his contribution as a consultant was considerable. I am also very grateful to him for authorizing me to make public use of the video sketch in (13), and for correcting the transcriptions and translations of examples (21), (22) and (47).

**Grant acknowledgments:** The research leading to these results received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement N°324115–FRONTSEM (PI: Schlenker). Research was conducted at Institut d'Etudes Cognitives, Ecole Normale Supérieure - PSL Research University. Institut d'Etudes Cognitives is supported by grants ANR-17-EURE-0017 FrontCog and ANR-10-IDEX-0001-02 PSL.

\*\*Institut Jean-Nicod (ENS - EHESS - CNRS), Département d'Etudes Cognitives, Ecole Normale Supérieure, Paris, France; PSL Research University; New York University, New York.

<b>1</b>	<b>Claims and Clarifications</b>	<b>3</b>
1.1	Logical Visibility	3
1.2	Iconicity	4
1.3	Terminology	6
<b>2</b>	<b>Signs and Iconicity</b>	<b>9</b>
2.1	Does iconicity play the same semantic role in speech and in sign?	9
2.2	How can gradience be established?	10
2.3	Why is there less iconicity in speech than in sign?	11
2.4	Can iconic enrichments always be equated with gestures?	12
2.4.1	<i>Incorporated gestures</i>	12
2.4.2	<i>The 'like this' analysis of iconic modulations</i>	13
2.5	Iconicity and modularity	15
<b>3</b>	<b>Signs and Gestures</b>	<b>16</b>
3.1	Sign+iconicity ≠ speech+gesture	16
3.2	What does ellipsis do?	17
3.2.1	<i>Disregarding gestures under ellipsis</i>	17
3.2.2	<i>Further iconic enrichments under ellipsis</i>	17
3.3	Do some signs include incorporated gestures?	20
3.3.1	<i>The claim: Aristodemo 2017, Aristodemo and Santoro 2018</i>	20
3.3.2	<i>Cosuppositions triggered by purely iconic elements</i>	21
3.4	From co-speech to pro-speech gestures: gestural grammar	23
3.4.1	<i>When non-signers know some properties of sign language</i>	23
3.4.2	<i>Replicating some properties of sign language grammar with co-speech gestures?</i>	23
3.4.3	<i>The dawn of pro-speech gestures</i>	24
<b>4</b>	<b>Role Shift: Context Shift, Iconicity and Gestures</b>	<b>26</b>
4.1	Role Shift and context shift	26
4.2	Context shift vs. mixed quotation	27
4.3	Davidson's theory	28
4.4	Maier's proposal	29
4.5	Quer's observation	30
4.6	Eckardt's hypothesis	30
4.7	Role Shift in gestures	32
4.7.1	<i>Ebert on character viewpoint co-speech gestures</i>	32
4.7.2	<i>Role Shift in pro-speech gestures?</i>	33
<b>5</b>	<b>Conclusions</b>	<b>33</b>
	<b>References</b>	<b>35</b>

What is the place of sign language within Universal Semantics, the comparative study of meaning in human languages? Two general claims were made in 'Visible Meaning' (Schlenker 2018b) about the relation between sign (= sign language) and speech (= spoken language).

**Logical Visibility:** In some cases, sign makes visible some aspects of Logical Form which are covert in speech.

**Iconicity:** Along some dimensions, sign has greater expressive resources than speech due to the presence of iconic conditions at its logical core.<sup>1</sup>

In this piece, we clarify both claims (Section 1) and discuss three main issues: What is the nature of the interaction between logic and iconicity in sign language and beyond (Section 2)? Does iconicity in sign language play the same semantic role as gestures in spoken language (Section 3)? Is sign language Role Shift best analyzed in terms of visible context shift (hence as an instance of Logical Visibility), or by way of demonstrations referring to gestures (Section 4)? In several cases, the empirical and theoretical landscape changed quite a bit after 'Visible Meaning' was written, hence the discussion will go significantly beyond the claims made in that article. (For elicitation methods and transcription conventions, the reader is referred to 'Visible Meaning'.)

## 1 Claims and Clarifications

### 1.1 Logical Visibility

It is standard in linguistics to discover that a structure that is covert in one language is overt in another. For instance, syntacticians routinely posit empty pronouns in languages that do not realize them overtly; and as mentioned in 'Visible Meaning', Hungarian is sometimes taken to display an overt version of quantifier movement, which is covert in English (Szabolcsi 2001, Brody and Szabolcsi 2003). Some cases of Visibility in sign language are of this general kind. It was noted in 'Visible Meaning' that ASL (American Sign Language) can use the same pronoun with temporal, modal and individual denotations, thus revealing the existence of an abstract (domain-neutral) module of anaphoric reference.<sup>2</sup> There is nothing in spoken language that would prohibit such a case, but it has not been described in the literature, although other constructions with both individual and modal meanings have been investigated (Bittner 2001, Bhatt and Pancheva 2006).

On the other hand, the fact that logical indices (variables) are overtly realized in sign language might well be due to the modality, in part because this instance of Visibility requires a mechanism to produce an unbounded number of symbols. Now one could in principle imagine that speech did so by constructing longer and longer words, with *ba* corresponding to the first index, *baba* to the second, *bababa* to the third, etc. While it is not clear why such a system doesn't exist (or at least hasn't been described!), the signed modality has a different and apparently much easier way to create indices at will, namely by way of points in signing space. As we will see in Section 3.4.3, non-signers arguably know this form-to-meaning mapping, which might thus be part of Universal Grammar, or have a broader cognitive source.

Each case of Logical Visibility relies on analyses that could be called into question. As reported in 'Visible Meaning', the claim that loci realize discourse referents can be challenged, in particular along Kuhn's lines (Kuhn 2016): one could take loci to be uninterpreted grammatical features akin to gender features. An alternative, suggested by Davidson 2018, is that loci should really be connected to deictic pointing in speech, which would have the advantage of accounting for some of their iconic properties. But two problems must be raised in this connection.

---

<sup>1</sup> Note that, at this point at least, we take Logical Visibility and Iconicity to be largely independent conditions (Brentari 2018 takes us to mean that "logical visibility is made possible via iconicity", but this would be a more ambitious claim than we intended).

<sup>2</sup> Note, however, that the most common strategy in ASL is certainly to rely on *covert* temporal and modal anaphora: the cases of temporal and modal anaphora discussed in Schlenker 2013a cannot at all be considered as a default strategy.

First, loci have bound uses under quantifiers (e.g. Schlenker 2014). Deictic expressions that are analyzed as getting their denotation from a demonstration alone (and hence solely from the context parameter) do not have this property, and thus a pure demonstrative analysis cannot work (for this reason, when some demonstratives can have bound readings, they too must interact with assignment functions rather than just with the context parameter). Second, there is a genuine risk of circularity if the analysis is made more liberal so as to count as a 'demonstration' whatever can be realized with a pointing gesture in spoken language. For as we will see in Section 3.4.3, it is possible to create mixes of English words and gestures in which a pointing gesture fully replaces a pronoun. These pointing gestures turn out to share non-trivial properties with sign language pointing signs, including binding. But this also means that the broader problem is to explain how *both* types of pointing elements, in sign and in gesture, come to have the grammatical and semantic behavior that they do. Reducing aspects of signs to gestures or the other way around serves to highlight interesting similarities between these systems, but not to explain their behavior.

Role Shift is another case in which a Visibility account can be challenged. 'Visible Meaning' and earlier work by several researchers (e.g. Quer 2005, 2013, Schlenker, to appear a,b) argued that Role Shift is the visible realization of a covert context-shifting operation that was originally posited for some spoken language attitude reports (Schlenker 2003, Anand and Nevins 2004, Anand 2006). But two properties distinguished Role Shift from spoken language context shift. First, several sign languages don't just have Attitude Role Shift, which closely resembles context shift in attitude reports, but also Action Role Shift, which serves to report actions rather than attitudes, and has no obvious counterpart in spoken language.<sup>3</sup> Second, even within Attitude Role Shift, quotational conditions appear to play a stronger role in ASL and LSF (French Sign Language) than context shift does in spoken languages in which it has been described (for the latter case, see for instance Anand 2006). This fully justifies the search for alternatives that analyze Role Shift by other means than context shift, as forcefully argued by Davidson 2015, 2018 and Maier 2018. We return to this point in Section 4.

## 1.2 Iconicity

'Visible Meaning' argued that iconicity plays a role 'at the logical core' of sign language. In 'Visible Meaning' this meant, quite literally, that **within sign language, iconicity can be found in some uncontroversially logical expressions such as variables** (this was the case of high loci and loci with diagrammatic uses pertaining to plurals).<sup>4</sup> But the claim invited broader interpretations as well, and for clarity these should be stated, distinguished and evaluated.

First, the claim can be understood to pertain to the observation that **sign languages offer a unique integration between iconic and logical conditions**. Iconic variables display a special case of this integration, but the claim can be understood more broadly. While sign languages have the same kind of logical and grammatical resources as spoken languages, their iconic means are far richer. In fact, it might be that signers are in some ways iconicity experts, as hinted by Brentari 2018, due to their constant attention to the visual modality. In any event, if one wishes to study the full extent of Universal Semantics, spoken language might be a poor guide to iconic components of meaning, and in this respect one might do well to start with sign language (we revisit below the issue of the integration between iconic gestures and speech). We take this descriptive claim about the integration of logical and iconic conditions to be uncontroversial in view of the data.

Second, a methodological requirement follows from this empirical observation: **an appropriate semantics for sign language must be a semantics with iconicity**. Furthermore, the methodological requirement that semantics should be explicit, i.e. formal, entails that one needs a formal account of how iconicity works. 'Visible Meaning' only sketched such an account, but included very simple examples for which explicit rules can easily be devised. Such was the case of plural loci

---

<sup>3</sup> We come back in Section 4.5 to Joseph Quer's observation that Action Role Shift seems to be missing in his LSC (Catalan Sign Language) data.

<sup>4</sup> Thus 'Visible Meaning' stated that "there is a non-trivial interaction between grammar and iconicity at the logical core of sign language: one and the same expression – say, a singular or a plural locus – can display a logical behavior (e.g. as a bound variable) while also having an iconic function".

used in a diagrammatic fashion: when a small locus  $a$  is signed within a large locus  $ab$ , this automatically gives rise to a complement locus  $b$ . And one can posit an iconic semantic rule whereby relative complementation among loci is 'preserved' by the interpretation function, with the result that the denotation of  $b$  is the complement of the denotation of  $a$  relative to the denotation of  $ab$ . In the end, this iconic condition explained why ASL has a grammatical counterpart of the deviant English discourse: *#Most of my students came to class. They stayed home instead.* But this only happened in case a large locus  $ab$  was introduced for the students, and an embedded locus  $a$  for the students that came to class – which sufficed to create a locus  $b$  referring to the students that didn't come to class. The reasoning is diagrammatic, and it is fair to compare it to what one would do with gestures representing diagrams.<sup>5</sup> But irrespective of one's final analysis, it will have to include an iconic component. In the general case, iconicity is complex, and more general and sophisticated formal rules will be needed to capture its workings, as suggested in Greenberg's recent work on picture semantics (Greenberg 2013).

Third, one might understand something different by the claim that *iconicity plays a role 'at the logical core' of sign language*, namely that **iconic rules do not form a separate module within sign language**. But the question is what one calls a 'module'. By inspecting the way in which meaning is produced, it is clear that iconic conditions work very differently from standard logical conditions, as they involve the preservation of geometric properties of signs in the space of semantic denotations. In traditional linguistics, one posits very different rules for phonology and syntax, or within syntax for Binding Theory and for Case Theory, for instance. This means that certain sets of rules form natural classes and are in a sense separate 'modules'. 'Visible Meaning' was careful to state that the presence of iconicity at the 'logical core' of sign language "doesn't mean that a two-module theory couldn't be developed; but the relevant notion of 'module' would have to be appropriately abstract". Besides the internal description of the relevant rules, one could ask whether they are shared with other systems. It is almost immediate that iconic rules are shared with non-linguistic representation systems, and in that sense they can be taken to form a separate 'module' from other components of language.<sup>6</sup> A further question, addressed by Emmorey 2018, is what studies of brain implementation can tell us about this issue; we come back to this point in Section 2.5.

Fourth, however, and despite the clear differences between the workings of iconic and non-iconic semantic rules, there is a precise sense in which **iconic conditions are not separable from logical conditions: it is not in general possible to take the meaning of a sentence to stem from the conjunction of a logical and of an iconic component**. Such a 'separability hypothesis' is not at all absurd, but it is incorrect. To have a simple model of it, consider Potts's influential analysis of expressives in English (e.g. Potts 2005, 2007). For him, their main property is that they fail to interact with logical operators, and for this reason the expressive component of meaning is a completely separate semantic dimension. One could imagine that the same thing happens with iconic enrichments. This is not so.

- A weak version of the refutation involves the presuppositional effect of height specifications of pronouns. These appear to be presuppositional in nature. If so, they are expected to interact with logical operators. Just as *None of these ten guy knows that he is incompetent* triggers a presupposition that each is in fact incompetent (due to the non-trivial interaction between presupposition and *none*-type quantifiers [e.g. Chemla 2009]), one would expect the sentence in (1) to trigger a presupposition (easily satisfied in this case) that each giant is tall.

---

<sup>5</sup> The comparison with co-speech gestures and diagrams was in particular suggested by Adam Schembri, and it is independently discussed in Ebert 2018; see Section 3.4.2 below.

<sup>6</sup> Davidson 2018 contrasts her views with the present ones by writing: "Instead of placing iconicity at the "logical core" of a language, sign languages, like spoken languages when gesture is also included, may instead highlight the fluidity with which we are able to integrate language with other cognitive systems." We do not disagree with this point, but it only highlights the need for clear definitions of what counts as 'language', and as 'the same' or 'different' cognitive systems – an issue 'Visible Meaning' had nothing to contribute to. Since iconic rules are obviously shared with rules of visual representation, there is a sense in which Davidson's claim is entirely uncontroversial (although it should be broadened: the iconic effect obtained in the iconic modulation of *long* vs. *looong* is of the same type, and does not require gestures).

- (1) NO GIANT THINK IX-1 LIKE IX-a<sup>high</sup>  
'No giant thinks that I like him.' (ASL 11, 26; Schlenker et al. 2013)

If so, the presuppositional component requires access to the logical structure of the sentence (to obtain the universal presupposition), but one can still take the final meaning to be the conjunction of an iconicity-free sentence (without the high locus) and of the universal presupposition.

• But in the general case, this won't do. The at-issue contribution of the iconic modulation in *GROW* shows why: in (2), *GROW* was realized with broad endpoints (hence a 'grew a lot' meaning), and its contribution was interpreted in the scope of the *if*-clause.<sup>7</sup>

- (2) *Context*: we are discussing the future of the speaker's research group.

<sup>7</sup> IF POSS-1 GROUP GROW<sub>large</sub>, JOHN WILL LEAD.  
'If my group grows a lot, John will lead it.' (ASL, 34, 1944; 2 judgments)

In this case, there is simply no way to analyze the final result as the conjunction of *If my group grows*, *John will lead it* and of some iconic condition, for the simple reason that the final meaning is *weaker* than that of the first conjunct.

Let us add that **iconic and logical conditions only become more entangled when one considers anaphora**. The example mentioned above already showed that plural discourse referents can be created on iconic grounds. Schlenker and Lamberton, to appear, make a similar remark about iconically used plurals: the word *TROPHY-rep*, realized with an unpunctuated repetition, provides information about the size and the spatial arrangement of the denoted group of trophies.<sup>8</sup> But in addition, it creates singular discourse referents at the vertices: for a horizontal arrangement, one can point towards the right or left to refer to the right-most or left-most trophy (pointing in the middle is degraded); for a triangular arrangement, one can point to each of the vertices to refer to the trophy found in that position. A logical account of language wouldn't be complete without a proper account of anaphoric relationships; this, in turn, requires a detailed account of iconicity.

### 1.3 Terminology

Some terminological issues should be cleared to forestall ambiguities.

**Iconicity** is often characterized as a resemblance between a sign and what it denotes. But since the notion of resemblance is vague, it might be better to define iconicity as **a requirement that the possible denotations of an expression *E* systematically preserve some properties of the form of *E*** – and it is incumbent on semantic theories to say which properties must be so preserved, and how.<sup>9</sup> The example we gave above involving complement set loci was so simple that the preservation rules could be stated very precisely, but in the general case iconic conditions can be non-trivial to state, as shown by Greenberg's and Abusch's work, among others (e.g. Greenberg 2013, Abusch 2015).

As this example shows, **iconicity need not involve gradience**: in the case of plural loci used as diagrams, there was no gradience, but still a strong argument for iconicity. **Nor does gradience entail iconicity**: one's voice may become more high-pitched when one is more stressed, but it doesn't follow that high pitch correlates with stress because of a semantic rule; rather, it seems to be a biological fact that greater stress has certain consequences on one's vocal chords. Still, gradience is


<sup>7</sup> As in 'Visible Meaning', acceptability scores are on a 7-point scale, with 7 = best; they appear before the relevant sentences. Note that there was no trace of a presuppositional effect in (2), but see Section 3.1 below for Ebert's suggestion that this is an effect of local accommodation.


<sup>8</sup> For an earlier claim about the interaction between pluractionals and iconicity, see Kuhn 2015 and Kuhn and Aristodemo 2017.

<sup>9</sup> As an example, the verb *GROW* in ASL can be realized more or less quickly to indicate that the growth process was more or less quick, hence an iconic component, to the effect, roughly, that *the faster the realization of the sign, the fast the growth process*. Note that if a lexical sign happens to resemble its denotation, but no interpretive rule makes reference to the preservation of some geometric properties, it does not count as 'iconic' in this sense: it may result from 'fossilized' iconicity without displaying iconicity in action, so to speak. Only the latter needs to be incorporated into interpretative rules.

often a helpful way to establish iconicity, because it is absent from a standard logical semantics. If you can gradually modify the form of a sign, and a corresponding gradient modification of the meaning follows, then you *might* be seeing an iconic rule in action.

**Gestures** are not easy to define, especially when one aims for a class that includes gestures in sign as well as in speech. One defines things as one wants, but some definitions will prove hard to use in the present debates (see Emmorey 1999 for an in-depth discussion of gestures in sign language).

While one could define as gestural anything that has an iconic component, this makes it hard to *ask* on the basis of independent criteria whether iconic modulations of ASL signs behave in the same way as gestures in spoken language: the definition is too broad to capture this distinction. In other respects, such a definition is also too narrow: the 'thumbs up' gesture  doesn't have an iconic component, but it still makes sense to treat it as a gesture.

We propose, as a very first approximation, to **define gestures as elements that (i) semantically enrich or replace some words, and (ii) fully fall outside of the phonology of the language**; but we believe more sophisticated definitions will be needed in future research. On this initial definition, the 'thumbs up' gesture  indeed qualifies as a gesture, on the assumption that English phonology is vocal. Lengthening of *o* in *loong* satisfies criterion (i), but it doesn't fully satisfy criterion (ii): while this very long vowel is outside of the phonology of English<sup>10</sup>, it is clear that it is a *modification* of the phonology of the language. Thus it would make sense *not* to treat this lengthening as a gesture, and indeed we call it an 'iconic modulation'. For the same reason, the lengthening of ASL *GROW* doesn't count as a gesture.<sup>11</sup>


Brentari helpfully proposes a different set of criteria, summarized in (3).

- (3) Criteria for distinguishing linguistic and gestural elements in signed and spoken languages (Brentari 2018)
- i) Linguistic elements in a string can have either hierarchical or flat structures; gestures elements in a string have only flat structures.
  - ii) Linguistic elements exhibit less individual variation than gestural units.
  - iii) Linguistic units exhibit less gradience than gestural units.
  - iv) Producers and receivers are typically less consciously aware of gestures than language.

We believe that these criteria are suboptimal or follow from our proposed definition. Let us discuss them in turn.

(i) We do not see strong reasons to think that gestural elements only have flat structures. For starters, lots of cognitive elements have a hierarchical structure, including phrases, syllables, music, and perceived events (Lerdahl and Jackendoff 1983, Jackendoff 2009, Zacks et al. 2001), so there is no *a priori* reason to think that gestures couldn't have a hierarchical structure as well. In addition, one can create examples in which gestures arguably do have structure. Consider first a plural-related example from Schlenker, to appear *f* (a modified version of which gave rise to subtle inferential judgments in the experiments of Tieu et al. 2018)



*Notation:* We write + for , and *o* for an O-shaped gesture representing medallions. *o-rep<sub>3</sub>* refers to three iterations respectively of the medallion gesture, and similarly for *+rep<sub>3</sub>*. — indicates that the repetitions are arranged on a horizontal line. *a* in *[+rep<sub>3</sub>—]<sub>a</sub>* indicates that

<sup>10</sup> Thanks to Donca Steriade (p.c.) for discussion of this point.

<sup>11</sup> Three remarks should be added.

1. As an alternative, one could adopt a slightly more complex definition: gestures are elements that (i) semantically enrich or replace some words, (ii) do not conform to the phonological rules of the language, and (iii) **are separable and thus realizable without linguistic material**. Condition (iii) would clearly classify the lengthening of *loong* and of *GROW* outside of gestures.

2. Superficially, sign language classifier predicates would seem to us to have considerable phonological freedom, but it doesn't follow that they don't obey some phonological constraints: their lexical, arbitrary component certainly does. See Brentari and Eccarius 2010 for a discussion of phonological constraints on classifiers, and Emmorey 1999 for relevant discussion of tracing constructions vs. tracing gestures in ASL (as she argues, the former are subject to non-trivial linguistic constraints).

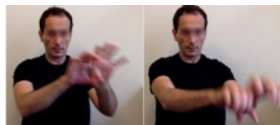
3. Our current definition, as well as the alternative in 1., might make it hard to state Aristodemo's (2017) hypothesis, according to which some LIS signs obligatorily contain incorporated gestures. We revisit this issue from a more substantive perspective in Section 3.3.



the gesture is made in position *a*, which we take to be realized roughly in front of the speaker, on the dominant side, while *b* represents a position on the speaker's non-dominant side. *IX-a* refers to an index pointing towards gestural locus *a*. *TAKE\_2-handed-a* is illustrated below and targets locus *a*.

- (4) *Context*: in a treasure hunt, the speaker is supposed to find medallions or crosses.

You will enter a room. You will see [o-*rep*<sub>3</sub>—]<sub>b</sub>, and also [*+*-*rep*<sub>3</sub>—]<sub>a</sub>, **and you will TAKE\_2-**



**handed-a**

=> the addressee will see some medallions (horizontally arranged), and some crosses (horizontally arranged), and the addressee will take (all) the crosses

Let us now modify this example in two ways. First, we will replace the unpunctuated repetition of the cross gesture, written above as *+rep*, with a punctuated repetition (produced horizontally), with each iteration clearly distinguished from the other, and written as *+++*. Second, we will add a repeated cross gesture in the final clause, as the object of the gestural verb, to mean that the addressee is to take the three crosses (if there are three iterations), or just two of them (if there are two iterations):<sup>12</sup>

- (5) *Context*: in a treasure hunt, the speaker is supposed to find medallions or crosses.

You will enter a room. You will see [o-*rep*<sub>3</sub>—]<sub>b</sub>, and also [*+* + *+*]<sub>a</sub>, **and you will**

a. **TAKE\_2-handed** [*+* + *+*]<sub>a</sub>.

=> the addressee will see some medallions (horizontally arranged), and three crosses (horizontally arranged), and the addressee will **take the three crosses**

b. **TAKE\_2-handed** [*+* + *+*]<sub>a</sub>.

=> the addressee will see some medallions (horizontally arranged), and three crosses (horizontally arranged), and the addressee will **take two of the three crosses**

Here the boldfaced sequence at the end includes three or four gestures, namely *TAKE\_2-handed*, followed by iterations of the cross gesture: three in (5)a, two (5)b. Now it's likely that these two or three iterations do not have internal structure: they might be a good example of Brentari's flat structure of gestures. But certainly they form a unit, to the exclusion of the gestural verb. And the next higher unit is certainly the gestural verb together with its gestural object. Standard syntactic tests could be applied to buttress this conclusion, such as pronominalization (the three crosses together are in effect pronominalized in the last clause of (4)), or possibly movement tests. If so, we end up with the syntactic structures depicted in (6):

- (6) a. [TAKE\_2-handed [*+* + *+*]]

b. [TAKE\_2-handed [*+* + *+*]]

For this reason, we do not take the absence of a hierarchical structure to be a good defining criterion of gestures.

(ii) It is probably true that linguistic elements exhibit less individual variation than gestural units, but (a) in the general case, this is probably because gestures do not conform to the phonological rules of the language, and (b) in any event, there might be exceptions, as in conventional gestures such as the 'thumbs up' gesture.

(iii) Gradience is neither a necessary nor a sufficient condition for gestures, as we saw above.

(iv) Lack of consciousness is neither a necessary nor a sufficient condition for gestures either. Pro-speech gestures, which fully replace spoken words, are certainly very conscious (as they play an ineliminable semantic role). Conversely, all sorts of linguistic elements, from traces to features, are not conscious and yet are linguistic and non-gestural.

<sup>12</sup> The repetition of [*+* + *+*]<sub>a</sub> is a bit redundant (as *TAKE\_2-handed-a* without an object would have yielded the same meaning), hence the example in (5)a might be slightly less acceptable than that in (5)b. This could be addressed by modifying the linguistic context and replacing *You will see [o-*rep*<sub>3</sub>—]<sub>b</sub>, and also [*+* + *+*]<sub>a</sub>* with: *You will see [o-*rep*<sub>3</sub>—]<sub>b</sub>, and also [*+*-*rep*<sub>3</sub>—]<sub>a</sub>* (with the cross gesture meaning: several crosses), or for that matter with: *You will see [o-*rep*<sub>3</sub>—]<sub>b</sub>, and also [*+* + *+* + *+*]<sub>a</sub>* (with cross gesture meaning: precisely four crosses). In these modified contexts, the gestural object of *TAKE\_2-handed* [*+* + *+*]<sub>a</sub> should not seem redundant.

**Language** has given rise to numerous attempts at definitions, and for scientific purposes, one should be suspicious when something essential turns on such definitional moves. **Human language is unique among so many dimensions that one can pick any number of them to define it**, with little insight gained in the process. Now one sometimes treats discrete infinity and conventionality as characteristic properties of language. Zucchi 2018 writes in this connection:

the effect of counting co-speech gestures as linguistic items is that one must give up both discrete infinity and conventionality as characteristic properties of spoken languages. The same would be true for sign languages, if locus height specifications were assimilated to co-speech gestures and then co-speech gestures were regarded as linguistic.

Once again, one can pick definitions as one chooses. But in any event, the reasonable thought is that language has discrete infinity and that it has conventional words, not that gradience and non-conventional effects are *excluded* from it. Intonation is probably gradient in its effects, and biological codes were brought to bear on certain universal effects of sound, as in Ohala 1994, according to which lower pitch is associated with larger body size. So the key definitional property of language from this perspective is that it has discrete infinity, not that it lacks gradience.

In fact, language-internal evidence as well as the comparison with related species suggests that gestures, iconic phenomena and standard linguistic elements should be studied as a natural class, and thus that one might need a broad concept of 'language'. First, it is increasingly clear that our closest relatives, chimpanzees and bonobos, have rich means of multi-modal communication, and can choose one modality or the other depending on what is most conducive to communicative success; for instance, as reported by Byrne et al. 2017, "chimpanzees and bonobos are more likely to use audible or silent visual gestures with an audience facing them, whereas for contact gestures no such effect is found". Second, the integration of facial and manual gestures with spoken language suggests that it is just a mistake not to analyze them as part of the same system (see for instance Joutteau 2004). Raised eyebrows were shown to be used to mark focus in speech (Dohen 2005, Dohen and Lovenbruck 2009), various means of pointing interact with grammar (McNeill 2005), and gestures of various kinds (co-speech, post-speech, pro-speech) fall within a well-established typology of linguistic inferences (Ebert and Ebert 2014, Schlenker, to appear d, to appear f). On the sign language side, iconic effects cannot be neatly excised from the rest of the language, especially in cases in which they make at-issue contributions, which can in no way be disregarded in the analysis of truth conditions (this was our point about non-separability in Section 1.2). Thus we would caution against arguing for or against an analysis on the basis of definitions, and in particular of definitions of 'language' that might be too narrow to capture the appropriate natural classes.

## 2 Signs and Iconicity

The existence of iconic effects that interact with the logical core of language in no way entails that iconicity *only* belongs to the linguistic system; as noted at the outset, this would be a nearly absurd position, since by its very nature iconicity functions as other representational systems, in particular visual representation for sign iconicity. But further issues are raised by the integration of iconicity with sign language: (i) Does iconicity play the same semantic role in the signed and in the vocal medium? (ii) Do relevant cases of iconicity in sign language display genuine gradience? (iii) Why is iconicity less prevalent in speech than in sign? (iv) Can all iconic enrichments be equated with gestures?








### 2.1 Does iconicity play the same semantic role in speech and in sign?

Our position is that **iconicity makes the same types of semantic contributions in speech and in sign, but that due to the representational limitations of the vocal medium, iconicity plays a more prevalent role in sign.**

We proposed in Schlenker, to appear d, a typology in which different kinds of contributions are made by different iconic enrichments depending on whether they are external to the relevant words (and are thus syntactically eliminable), and whether they have a separate time slot. This

typology is illustrated in (7). Without going into details, the crucial observation was that in speech and sign alike, iconic modulations – as in the lengthening of a vowel or a sign – can make at-issue contributions, while co-speech gestures and co-sign facial expressions contribute special varieties of presuppositions (called cosuppositions), and post-speech and post-sign gestures and facial expressions contribute supplements, i.e. make the same kind of contributions as appositive relative clauses.

(7) **Typology of iconic enrichments** (Schlenker, to appear d)

	External enrichments (= syntactically eliminable)		Internal enrichments (= syntactically ineliminable)	
	No separate time slot: <b>Co-speech/co-sign gestures</b>	Separate time slot: <b>Post-speech/post-sign gestures</b>	No separate time slot: <b>Iconic modulations</b>	Separate time slot: <b>Pro-speech/pro-sign gestures</b>
<b>Speech</b>	John  punished his son.	John punished his son – 	The talk was loooooong.	Your brother, I am going to 
<b>Sign</b>	IX-arc-b NEVER  [SPEND MONEY]	IX-arc-b NEVER SPEND MONEY] – 	POSS-1 GROUP GROW_  	[currently unclear]
<b>Meaning</b>	cosuppositions (= presuppositions of a special sort)	supplements	at-issue or not, depending on the case	at-issue, with an additional non-at-issue component in some cases

Importantly, with the exception of pro-sign gestures (i.e. sign-replacing gestures, whose existence and status is unclear to us), the same typology was taken to hold in speech and sign. Co-speech gestures arguably trigger conditionalized presuppositions (cosuppositions): *Will Ann LIFT help her son?*, with a lifting gesture co-occurring with *help*, yields the inference that *if Ann helps her son, lifting will be involved* (Schlenker 2018a). And it was claimed in Schlenker, to appear d, that the same general observation (pertaining to conditionalized presuppositions) holds of co-sign facial expressions in ASL. Post-speech gestures, for their part, were taken to trigger supplements, i.e. the same kind of contribution as appositive relative clauses; and the same claim was made about post-sign facial expressions in ASL. Finally, iconic modulations of *loong* and *GROW* were taken to be in principle very free, and in particular to have the ability to make at-issue contributions, irrespective of the modality.

This typology was taken to be constrained by the status of the iconic enrichments: syntactically eliminable ones were claimed not to make at-issue contributions; and enrichments that have their own time slot were claimed not to be presuppositional/cosuppositional. Be that as it may, the general typology was taken to be modality-neutral.

## 2.2 How can gradience be established?

As we noted, gradience is neither a necessary nor a sufficient condition for iconicity, but it is often a good indication for it. One could ask, following Zucchi 2018, how strong the evidence for gradience is in various areas of sign language semantics. In the general case, **gradience in semantics can be defined by the fact that between two realizations  $R_1$  and  $R_2$  of an expression, with meanings  $M_1$  and  $M_2$ , it is always possible to find a third realization  $R_3$  whose form is intermediate between  $R_1$  and  $R_2$ , and whose meaning is also intermediate between  $M_1$  and  $M_2$ .** The definition can only make sense if there is a pre-determined notion of order defined on the form and on the meaning side. For instance, different lengths of the vowel in *loong* (as in: *the talk was loong*, to mean that it was very long) could be taken to approximate a gradient behavior. If I say that *the talk was long*, and then that *the talk was loooooong*, I can find an intermediate length such as: *the talk was loong*.<sup>13</sup> The reason

<sup>13</sup> Needless to say, repetition of letters is a less gradient process than modulation of lengths, hence transcriptions are somewhat misleading in this case.

gradience can be suspected is that vowel length can be varied in a gradient fashion, and it is plausible that the corresponding meaning is varied in this way as well. But two obstacles immediately arise: first, we can per force only test a finite number of such distinctions, whereas gradience in a strict sense entails infinitely many such distinctions. Second, the distinctions must be producible and perceptible on the form side, and relevant on the meaning side. Constraints on production and perception suggests that, in practice, distinctions will be severely limited.

Still, if this makes for an insightful analysis, it is reasonable to posit a gradient system that interacts with performance limitations (on the production and perception side), which entail that, in practice, one can only use a discrete approximation of gradience ('quasi-gradience'). Given two loci, one can usually establish a third one between them, denoting a third individual (but see Steinbach and Onea 2016). It would not seem insightful to posit that there are, say 10 or 100 discrete pre-established positions in signing space that can be used as loci (for starters, the choice of the number of loci and of their precise position might look arbitrary). In a different domain, Emmorey and Herzig 'gradiently' varied the position of a small object classifier in ASL relative to a flat object handshape. But of course all the authors could do was to pick *enough cases* to make it plausible that they were approximating a gradient system.

The situation seems to us to be similar with locus height, except that fewer positions have been simultaneously studied. Zucchi 2018 writes of a 'three-way distinction' in locus height, but 'Visible Meaning' discussed an example with four gymnasts appearing in standing position at four different heights of a slanted bar (see also Schlenker 2015a). When the gymnasts were rotated downwards, four further heights appeared as well, but now under the bar, with the result that we had a total of eight height heights – which to us makes it implausible that the underlying system involves a fixed number of discrete height distinctions.

It must be added that the iconic character of the height system is also established very strongly by related exercises in which astronauts had to be rotated at various angles as part of their training (Schlenker 2014). Loci representing these astronauts were rotated accordingly. Thus the point is not just that loci can be established at different heights depending (roughly) on the position of the head of the denoted individuals. Rather, loci can simultaneously function as variables *and* as simplified pictures of what they denote. In the end, as proposed in 'Visible Meaning', one genuinely needs a synthesis of the theoretical insights of the 'grammatical camp' (loci are variables, e.g. Lillo-Martin and Klima 1990) and of those of the 'iconic camp' (loci are simplified pictures, e.g. Liddell 2003).

### 2.3 *Why is there less iconicity in speech than in sign?*

Another question is why iconicity does not play in speech the role that it does in sign (a point correctly raised by Zucchi 2018). As far as we can tell, the differences can plausibly be explained by two considerations: vocal iconicity is relatively limited in what it can represent besides size/duration and acoustic effects; and those effects might be harder to perceive than in sign language.

Still, sound symbolism is increasingly studied and appears to have deep relations with some iconic effects in sign language. In the notorious *bouba/kiki* effect (e.g. Ramachandran and Hubbard 2001, McCormick et al. 2015), nonce words with different mouth shapes and tongue movements are systematically associated with different denotations: if given a choice, subjects tend to take *bouba* to refer to a rounded shape, *kiki* to an angular one. For their part, Kuhn et al. 2018 explored vocal iconicity effects that originate in studies of 'Wilbur's generalization' in sign language. Wilbur (e.g. Wilbur 2003) had shown that in several sign languages, telic verbs are phonologically expressed by signs with sharp boundaries, unlike atelic verbs. Strickland et al. 2015 argued with experimental means that this distinction is 'known' by non-signers, who can in effect guess the telicity of sign language verbs. Kuhn et al. 2018 further show that, if given a choice between a count vs. a mass meaning in the nominal domain, non-signers preferably associate sharp boundaries to count meanings.<sup>14</sup> Importantly, they also find vocal counterparts of sharp boundaries by contrasting nonce

---

<sup>14</sup> Still, the distinction between sharp vs. non-sharp boundaries is not known to play a strong grammatical role in the count/mass distinction in sign language. But see Schlenker and Lamberton, to appear, for possibly relevant ASL data pertaining to iconically repeated mass terms.

words that end vs. do not end in stops. Sound symbolism is found in this case as well: words with stops are preferentially assigned telic or count meanings, presumably because stops create sharp boundaries in production and/or perception. It thus appears that Wilbur's generalization has extensions in spoken language.

Zucchi 2018 asks why speech doesn't make use of different vowel lengths to represent different degrees of height in *tall*, or at least why the option is not grammaticalized. But certainly the option is exercised, as in *loong* above, or possibly even in *John is taaall*. 'Grammaticalized' is a somewhat theory-internal category. On independent grounds, one might not expect agreement phenomena to arise in English with degrees (but one might still look for other cases in other languages in which an iconically modified sound might be expected to enter in agreement phenomena). By contrast, agreement phenomena are notoriously present in pronouns, and it is thus unsurprising to find them with iconically modified pronouns in sign language.

It should be added that there is a growing body of work on ideophones in spoken language, which do come with their own grammatical constraints (Dingemanse 2013, Dingemanse and Akita 2016, Dingemanse et al. 2016), and which might also suggest that iconic effects can in fact be grammaticalized in some ways; we leave a proper discussion for future research (and for more competent researchers).

#### 2.4 *Can iconic enrichments always be equated with gestures?*

One might be tempted to take all iconic effects in sign language (including the iconic modulations illustrated in (7)) to be analyzable as gestural elements. If we call 'gestural' any element that is iconic, then iconic enrichments are indeed gestures, but as mentioned, only confusion is gained in this terminological move: some gestures are conventional, and some iconic enrichments are modifications of words rather than external additions to them, as we saw above.

If we take gestures in something like the normal sense, then one can have an empirically defeasible *theory* of iconic enrichments according to which these stem from incorporation of some gestures in words.<sup>15</sup> But one immediately needs to be more precise: as we saw above, gestures come in different types, including co-speech, pro-speech and post-speech gestures, and each category has its own characteristic properties.

Two substantive proposals ought to be discussed in this connection: one takes some signs to include incorporated gestures; the other takes iconic modulations to include a component akin to the modifier *like this*, where *this* makes reference to the form of a sign *qua* gesture. Let us discuss these theories in turn.

##### 2.4.1 *Incorporated gestures*

Aristodemo 2017 and Aristodemo and Santoro 2018 provide a beautiful analysis of the adjective *FULL* in LIS (Italian Sign Language) in terms of an incorporated gesture that means 'completely', and is an integral part of the word. But the reason their analysis is so compelling is that they go to great lengths to display characteristic properties of co-speech gestures in the 'completely' component of *FULL*, notably the fact that (i) it triggers a cosupposition, and (ii) it can be disregarded under ellipsis (Brentari 2018 also discusses the interaction of iconic enrichments and ellipsis, a point to which we return in Section 3.2). Still, it is important to realize that this behavior – notably, the triggering of a cosupposition – does not seem to be the general case: several iconic modulations of signs seem to pattern like the lengthening of *loong* in making at-issue contributions. By contrast, the inferences triggered by co-speech gestures do *not* start out as being at-issue. One could of course posit that some mechanism akin to 'local accommodation' is responsible for the at-issue uses seen in sign language, as argued by Ebert 2018, but evidence would be needed that this is indeed the case. In Section 3.3, We will revisit Aristodemo's analysis, and we will argue that the behavior she uncovered might not be characteristic of co-speech gestures, but rather of a slightly broader class – which in turn might offer an alternative to the (interesting) 'gesture incorporation' analysis of her LIS examples.

---

<sup>15</sup> We say 'in something like the normal sense' because the definition of 'gestures' we gave in Section 1.3 will have to be refined if incorporated gestures obey the phonological rules of the language.

## 2.4.2 The 'like this' analysis of iconic modulations

Zucchi 2011, 2018 and Davidson 2015 proposed that classifier predicates in sign language have an iconic component that comes about because their semantics has a 'like this' component that makes reference to their own form. In Zucchi's words (2018),

the intuitive idea behind these proposals is that movement in classifier predicates of motion is a gesture which fixes the referent of a hidden demonstrative (i.e., a demonstration), in the same way in which the reference of [a] demonstrative may be fixed by a gesture:

- (8) The car moved in a way similar to this.

Davidson 2015 further proposed that sign language Role Shift makes iconic reference to expressions in its scope. For Attitude Role Shift, this yields a quotational reading. For Action Role Shift, which she takes to affect only single verbs (not larger constituents), this yields a reading on which the role-shifted verb is both used and quoted. Thus an example such as (9) is analyzed in simplified form in (10).

- (9) <sup>7</sup> RECENTLY WOLF IPHONE<sub>c</sub> FIND<sub>c</sub> HAPPY. SHEEP<sub>a</sub> IX-b b-CALL-a.  
'Recently the wolf was happy to find an iPhone. He called the sheep.'

RS<sub>b</sub>\_\_\_\_\_

IX-b IPHONE 1-SHOW-CL-a.  
=> *the wolf in fact showed the iPhone to the sheep*  
He [= the wolf] showed the iPhone to him.' (LSF, 39, 26)

- (10) [[IX-b IPHONE SHOW-CL rs [gesturally appropriate handshape and movement]]]  
=  $\exists e$  [agent(e,wolf) & theme(e,iphone) & showing (e) & demonstration( $d_1$ ,e)], where the value of  $d_1$  is given by the iconic properties of *SHOW-CL* (adapted from Davidson 2015 (68))

'Like this' analyses of iconic enrichments have one great virtue: they offer a theory of iconic modulations that is easy to grasp, comes with a simple semantic analysis (by way of modification), and highlight the at-issue status of some iconic modulations (since the modifier *like this* always makes an at-issue contribution). But they also have one drawback: they sweep under the rug the detailed contribution of the iconic component. The challenge for a theory of iconically-rich systems is to provide such an account, and leaving the explanation to gestural paraphrase won't quite do in the end. This need not be a deep criticism, however: one would only need to supplement the 'like this' analysis with a detailed account of how the demonstration works.

But when this is done, we believe the analysis might in some cases reveal itself to be lacking, although for different reasons in the case of classifier predicates and in the case of Action Role Shift. Starting with Action Role Shift, the problem is that the predicted truth conditions are either incorrect or insufficiently specified. Consider for instance (11), from Schlenker, to appear a:

- (11) <sup>7</sup> IX-a JOHN OFTEN MEET-MEET [POOR PEOPLE]<sub>b</sub>, IX-1 KNOW LOTS PEOPLE IX-arc-b IX-a  
FINISH  
'John often meets poor people. I know lots of people that he  
RS<sub>a</sub>\_\_\_\_\_  
1-GIVE-b-rep MONEY.  
has given money to.' (ASL, 18, 98)

The lexical words *GIVE* and *MONEY* appear under Role Shift. But certainly a giving of money doesn't resemble these expressions much. In fact, if one utters (12) in English, but using ASL words for the demonstration, the result would be very odd indeed: the ASL words are just not appropriate to serve as demonstrations.

- (12) Yesterday I gave money to people like this: RS<sub>a</sub>\_\_\_\_\_
- 1-GIVE-b-rep MONEY.

What Davidson means is that the giving resembles *in relevant respects* the realization of the words. But stating *what these relevant respect are* is precisely the goal of an explicit semantics with iconicity.<sup>16</sup> Precisely this goal motivated the statement of explicit preservation principles in earlier accounts of iconic modulations, for instance of *GROW* (Schlenker, to appear a<sup>17</sup>).

Due to their stronger iconic character, classifier predicates do not give rise to quite the same problems. But we believe that the 'like this' analysis of their semantic contribution might nonetheless prove inadequate in the end, although more research will be needed to establish this point. The reason is that some predicate classifiers appear to generate presuppositions even though controls with at-issue modifiers that come close to 'like this' expressions do not. This finding is discussed in detail in Schlenker 2018c, but the paradigm in (13) will suffice to sketch the main idea.

- (13) *Context*: our company has one helicopter and one airplane.  
 WITHIN 1-HOUR OUR COMPANY BIG HELICOPTER BOSTON<sub>a</sub> NEW-YORK<sub>b</sub> **MAYBE** a-\_\_\_\_-b.  
 'Within the next hour, maybe our company's big helicopter will \_\_\_\_ from Boston to New York.'  
 (ASL, 34, 3633; 3 judgments; Schlenker 2018c)  
 Video: <http://bit.ly/2FvF6mJ>

Condition (ASL, 34, 3633; 3 judgments)	Words (replacing ____ ) and acceptability	Translation
a. neutral path	<sup>7</sup> GO-helicopter-large_	fly
b. curved path	<sup>7</sup> GO-helicopter-large_∩_smooth	fly (with the assumption that this would involve a curved path)
c. at-issue control of curved path	<sup>5</sup> GO-helicopter-large_ WITH _∩_smooth	fly with a curved path
d. orthogonal detour	<sup>6,7</sup> GO-helicopter- large_ _no_acceleration	make an orthogonal detour on its way
e. at-issue control of orthogonal detour	<sup>5,3</sup> GO-helicopter-large_ WITH _ _no_acceleration	fly with an orthogonal detour on its way
f. pause in the middle	<sup>7</sup> GO-helicopter-large_ _ _	pause to hover on its way
g. at-issue control of pause in the middle	<sup>6,7</sup> GO-helicopter-large_ _ WITH PAUSE _ _ _	fly with a pause to hover on its way

All the sentences have the general meaning that within the next hour, our company's big helicopter might go from Boston to New York. But in (13)b,d,f, the path involves some modulations: it is curved, or involves an orthogonal detour, or a pause in the middle.

In the 'curved path' condition in (13)b, a (relatively weak) presupposition is triggered to the effect that *if the helicopter were to go from Boston to New York within the next hour, it would have a curved path*; technically, this is a cosupposition because it is conditionalized on the at-issue contribution of the expression (this point will matter in Section 3.3.2). In the 'orthogonal detour' and in the 'pause in the middle' conditions in (13)d,f, a (relatively weak) presupposition is triggered to the effect that *the helicopter will go from Boston to New York within the next hour*. Crucially, the controls with at-issue modifiers do not trigger such inferences. Rather, these controls make precisely the kind of at-issue contribution that one would expect with all sorts of explicit modifiers, including *like this*. Now our controls do not literally involve a 'like this' modifier, but rather display the relevant path after the word *WITH*. One could of course test closer analogues of 'like this' in ASL. But in any event the presuppositional or cosupposition behavior of some classifier predicates is not expected under the current *like this* analysis.<sup>18</sup>

While we do not mean to adjudicate the debate, we wish to highlight the importance of making detailed predictions about the semantic/pragmatic status of iconic enrichments. As we saw, in the case of Action Role Shift, the predictions of the 'like this' analysis might be wrong or

<sup>16</sup> One natural way to go, for Davidson and for alternative analyses, would be to specify in this case that *those aspects of sign form that are not conventional* are to be interpreted iconically.

<sup>17</sup> This manuscript was first uploaded on LingBuzz in April 2014.

<sup>18</sup> See Schlenker 2018c for a discussion of why the correct semantic behavior might be explained if these are treated as simplex constructions [rather than as covert 'like this' constructions].

underspecified, and in the case of predicate classifiers, this analysis might also be insufficiently fine-grained.

## 2.5 *Iconicity and modularity*

As mentioned at the outset, it is enough to inspect the semantic rules by which logical vs. iconic meanings are produced to note that they have very different properties: they just produce information in very different ways. On an abstract level, then, iconic and logical rules can be considered as separate modules – but this doesn't tell us anything about their implementation in the brain (similarly, different components of grammar have different formal properties and may thus constitute different 'modules', but this need not say much about issues of brain implementation). Emmorey 2018 summarizes neurolinguistic results that are directly relevant.<sup>19</sup>

Emmorey's general approach is congenial to that developed in 'Visible Meaning'. A verb like *GROW* was taken to have a lexical meaning, and an iconic component that can enrich it. This also corresponds to signers' intuition that there are 'neutral' forms that are not iconically enriched. A similar situation is the default in spoken language, where iconic modulation (as in *the talk was loong*) is the exception, not the rule. In Emmorey's words, "many lexical signs are of necessity interpreted and produced categorically, unless the signer intentionally exploits the iconic features of the sign (and the two modules thus interact)". On the basis of the PET study in Emmorey et al. 2013, she concludes that "classifier handshapes are categorical morphemes which are retrieved via left hemisphere language regions", whereas "the production of classifier constructions expressing spatial locations engaged regions in bilateral superior parietal cortex".

One general direction for the future would lie in conducting comparative studies of iconic enrichment in sign and in speech. For reasons outlined in Section 2.1, there are at least three different kinds of iconic enrichments: iconic modulations, co-speech/sign enrichments, post-speech/sign enrichments. These are distinguished both by formal considerations (syntactic eliminability and timing), and by semantic ones pertaining to the type of contribution made (at-issue, cosuppositional, or supplemental). Importantly, the natural classes obtained in this way needn't be those that come to mind without a linguistic analysis. Thus the iconic modulation of *GROW* is most directly related to the vocal modulation of *loong* rather than to the addition of a co-speech gesture to a spoken word. Co-speech gestures (and facial expressions) could be analogized to (non-grammatical) co-sign facial expressions, on the other hand; and similarly, post-speech gestures could be analogized to post-sign facial expressions. One can only hope that there will be more interaction between the formal semantic analysis of signs and gestures on the one hand, and detailed studies of brain implementation as outlined in Emmorey 2018.

---

<sup>19</sup> As Emmorey 2018 emphasizes, psycholinguistic and neurolinguistic results are also directly relevant to other aspects of the analysis of 'Visible Meaning'.

–As mentioned at the outset, Emmorey and Herzig's (2003) very careful experimental work on predicate classifiers established completely convincingly that these have a gradient component.

–Emmorey and Falgier's (2004) processing work on Locative Shift raises problems for analyses, such as Schlenker, to appear e, which take a locative-shifted locus to involve both an individual and a situation component: this would lead one to expect that both a space-denoting and an individual-location loci should be reactivated, but Emmorey and Falgier only found reactivation of the individual-denoted loci, hence a puzzle for future research.

–Similarly, results reported by Emmorey 2018 on the memory effects of locus inversion for arbitrary vs. spatially informative loci could bear on theoretical research (Emmorey et al. 1995). These results make it somewhat unlikely that one can always assign loci a (possibly very abstract) spatial component. Interestingly, precisely this view might be necessary if one wanted to solve the problem of 'bishop' sentences in sign language (described in Section 2.4 of 'Visual Meaning') within an E-type analysis, by taking the apparent symmetry between the antecedents *FRENCHMAN<sub>a</sub>* and *FRENCHMAN<sub>b</sub>* to be broken by way of an interpretation of the type: 'a Frenchman here' vs. 'a Frenchman there'.



### 3 Signs and Gestures

How should gestures be integrated to the sign/speech comparison? Goldin-Meadow and Brentari 2017 correctly argued that sign with iconicity should be compared to speech with gesture rather than to speech alone. This might suggest that speech with co-speech gestures is, on an abstract level, the same system as sign with iconicity. For reasons pertaining to the typology of iconic enrichments discussed in Section 2.1, we believe that this is in general incorrect. On the other hand, there might be special cases in which gestures are incorporated into signs, as argued by Aristodemo 2017 and Aristodemo and Santoro 2018; but as we will see in Section 3.3, there might be a non-gestural analysis of their data. Be that as it may, **we believe that some of most fruitful points of comparisons for signs with iconicity have remained understudied: outside of gestures, they pertain to iconic modulations of spoken words**, illustrated above with *looong*; **within the domain of gestures, a less close but possibly fruitful comparison involves pro-speech gestures**, which fully replace spoken words.<sup>20</sup>

#### 3.1 Sign+iconicity ≠ speech+gesture

As noted in Section 2.1, iconic modulations in speech (e.g. length modulations of English *long*) and sign (e.g. modulations of ASL *GROW*) can make at-issue contributions, whereas co-speech and co-sign gestures or facial expressions usually make non-at-issue contributions (note that we do not exclude that some iconic modulations can make non-at-issue contributions; in fact, the classifier predicates in (13) highlight precisely this point). Besides the empirical arguments mentioned in Section 2.1, this typology is conceptually motivated, since the iconic modulation of *looong* is indeed of the same type as that of *GROW*: it modifies the form of a conventional word in a way that provides iconic information. Co-speech gestures, by contrast, are external add-ons to conventional words, which might explain why they generally make a different type of semantic contribution.

On the positive side, this analysis suggests that **iconic modulations of signs might be more fruitfully compared to iconic modulations of spoken words rather than to co-speech gestures**, including for neurolinguistic studies (we will see below that pro-speech gestures might offer another point of comparison).

On the negative side, our analysis argues against the view, outlined by Ebert 2018, that the effect of iconic modulations can be emulated with co-speech gestures. But Ebert has a point: even co-speech gestures can be at-issue, although introspective and experimental studies suggest that this is a costly process (Schlenker 2018a, to appear d; Tieu et al. 2017, to appear), sometimes aided by contrastive focus (Esipova 2016a,b). The empirical question is thus a subtle one: can iconic modulations make by default, i.e. without a costly process of modulation, an at-issue contribution? The issue is all the more complex since all we claimed above is that iconic modulations can be at-issue, not that all of them are.

An important task for future research will thus be to conduct a comparison of at-issue contributions of iconic modulations and of co-speech gestures. Ebert proposes one example in which a co-speech gestures makes an at-issue contribution:

- |                           |                              |                    |              |
|---------------------------|------------------------------|--------------------|--------------|
| (14) If my group [grows], | John will lead it, but if it | [grows],           | he will not. |
| GROW-broad                |                              | GROW-broad         |              |
| gaze at hearer            |                              | gaze at hearer     |              |
| raised eyebrows...        |                              | raised eyebrows... |              |

There are two issues to keep in mind with this example, however. First, raised eyebrows can be a marker of focus/intensification in sign and to some extent in spoken language (Dohen 2005, Dohen and Lovenbruck 2009), and this might conceivably suffice to help modulate the word *grow* independently from the co-speech gesture; the reason is that focus/emphasis can be a means of intensification, in particular for gradable constructions (e.g. Bergen 2016). Second, Esipova 2016a,b noticed that presuppositional expressions in such contrastive environments can easily get at-issue readings (with local accommodation). In the case at hand, the point is that local accommodation is

<sup>20</sup> For Brentari 2018, the relevance of pro-speech gestures might stem from the fact that they are more salient than co-speech gestures and for this reason more similar to signs.

forced because, without it, we would obtain contradictory inferences, to the effect that if my group grows, John will both lead it (first conjunct) and not lead it (second conjunct; in addition, conditions on contrast won't be satisfied, as the two *if*-clauses should certainly be contrastive in some way).

We note that the typological view sketched above motivates part, but only part, of the following conjecture by Brentari (2018), at least if we understand the expression 'included in logical form' to mean that the relevant enrichments can be at-issue:

*when co-sign or co-speech gesture occurs within modality, the likelihood of being included in logical form is enhanced. The reverse is also true: when co-sign or co-speech gesture occurs in a different modality, the tendency to include elements in logical form is diminished.*


Co-sign facial expressions occur within modality but should count as external enrichments in our typology, and thus they should make non-at-issue contributions, something argued for in Schlenker, to appear d (and below in (21)-(23)). On the other hand, iconic modulations of signs and gestures should fall under Brentari's conjecture.

### 3.2 What does ellipsis do?

#### 3.2.1 Disregarding gestures under ellipsis

As noted in Schlenker 2015b, to appear b, co-speech gestures can be disregarded under ellipsis, as is illustrated in (15)b. Importantly, pro-speech gestures, by contrast, cannot be ignored in the course of ellipsis resolution: in (15)a, the elided VP is preferably understood as *hanged*, hence the need for a specific covert word: *hanged* but not *punished*.

- (15) a. A traitor should be HANG . A whistleblower shouldn't be.  
Preferred interpretation: whistleblowers shouldn't be hanged

- b. A traitor should be HANG  **punished**. A whistleblower shouldn't be.  
Preferred interpretation: whistleblowers shouldn't be punished  
(Schlenker, to appear b)

Why can co-speech gestures be ignored under ellipsis? The natural thought is that this is because the phonological form they modify can easily be separated from, and thus copied without, the co-speech gestures. These results can potentially serve as a test for what constitutes a co-speech gesture, although it must be handled with some care: there are other linguistic elements, such as features inherited through agreement, which can also be disregarded under ellipsis, as illustrated in (16)b. And in Section 3.3.2, we will consider the possibility that this 'disappearing act' is characteristic of a broader class of cosupposition-inducing expressions.

- (16) In my study group,  
a. Mary did her homework, and John did too.  
=> available bound variable reading in the second clause  
b. Mary  $\lambda_i t_i$  did her<sub>i</sub> homework, and John  $\lambda_i t_i$  did [*do ~~her~~ homework*] too  
c. Mary did her homework, and John did her homework too.  
=> no bound variable reading in the second clause

#### 3.2.2 Further iconic enrichments under ellipsis

Brentari 2018 discusses very interesting comparative data pertaining to iconic enrichments in English vs. ASL. Do they follow from the distinctions we have drawn up to this point?

As Brentari 2018 writes, preliminary results suggest that "the excited manner in the co-speech gesture can be ignored in English [in] [(17)b]— the meaning could be that John was simply pacing (not in an excited way)—but not in the co-sign ASL [in] [(18)b]—it is not possible to interpret *JOHN NOT* without also including *excited*". For Brentari, such a contrast might suggest that "visual prosody may be treated differently in spoken vs. signed language".

- (17) Ellipsis targeting visual prosody as expressive content (English)  
 Sylvester paced [facial expression: excitedly], ...  
 a. and John did [~~excited~~] too. [verb]  
 b. but John didn't [pace excited/~~excited~~] [verb+negation]
- (18) Ellipsis targeting visual prosody as expressive content (ASL)  
 CAT PACE [FACIAL EXPRESSION: EXCITEDLY], AND...  
 a. JOHN SAME [~~EXCITEDLY~~] TOO. [verb]  
 b. ?JOHN NOT [PACE EXCITED/?~~EXCITED~~] [verb+negation]


We believe that these results can be interpreted as follows. First, the behavior of the English co-speech facial expression is expected because co-speech gestures can be disregarded under ellipsis, as is illustrated in (15)b. As for the ASL facial enrichment, signers might, depending on the case, view them in two ways. On the one hand, they might treat the facial expression as a co-sign gesture, in which case it should be possible to disregard it. On the other hand, because the facial expression is in the same modality as the sign, it could also be treated as an iconic modulation of that sign, in which case it should be hard to ignore it. The first case should for instance arise when the facial expression provides information about the signer's evaluation of the described situations, rather than about the behavior of an individual in those situations. One would of course like to have further criteria to distinguish the two cases, an issue we leave for future research.

Brentari's other initial results can also be explained by way of the typology introduced up to this point. First, a co-speech gesture accompanying *grow* in English can be ignored under ellipsis. Second, an iconic modulation of ASL *GROW* cannot be ignored in this way, as shown in (19). This is expected from the present perspective because such a modulation cannot be easily excised from the expression it modifies (when we take into account the possibility that cosuppositions can be ignored under ellipsis, as in Section 3.3.2 below, we need to also posit that the modulation of *GROW* yields an at-issue enrichment and for this reason cannot be ignored under ellipsis).

- (19) VP Ellipses with iconic enrichments accompanying GROW (ASL)  
 POSS-1 GROUP GROW [slow/fast; small/medium/large],  
 a. #AND MINE SAME [GROW [slow/fast; small/medium/large; #~~slow/fast; small/medium/large~~] too.  
 [verb]  
 b. #BUT MINE NOT[GROW [slow/fast; small/medium/large; #~~slow/fast; small/medium/large~~]  
 [verb+negation]

Finally, Brentari 2018 notes that in ASL, a size gesture/sign<sup>21</sup> cannot be ignored under ellipsis, as shown in (20). For us, this is because such an element cannot be easily excised from the syntactic form copied by ellipsis. As Brentari herself notes, things are different with a co-speech size gesture in English co-occurring on a noun such as *bottle*: it may be ignored under ellipsis.



- (20) PHILOSOPHER BRING BEERWORKSHOP [Gesture: , ...  
 a. #AND GEORGE SAME [bring a large/#~~large~~ bottle of beer to the workshop] too.  
 b. #BUT GEORGE NOT [bring a large/#~~large~~ bottle of beer to the workshop].

We would suggest that the missing part of Brentari's typology pertains to facial expressions which are external to the signs they modify, and can thus count as co-sign gestures rather than as iconic modulations, especially when they provide information about the signer's attitude. We believe a simple example pertains to :-([*SPEND MONEY*] realized with a disgusted expression indicating that the signer finds the spending disgusting, as is illustrated in the summary table in (7), and discussed in Schlenker, to appear d: we predict that such a facial expression behaves in the same way as co-speech gestures in speech, and is thus easy to ignore under ellipsis.

Recent data suggest that this prediction is on the right track. (21)b has the verb *USE* co-occurring with a disgusted facial expression, and it appears under negation. This triggers a cosupposition to the effect that using the relevant fish would involve a disgusted expression. But an elided clause (without negation) pertaining to meat doesn't have the inference: the 'disgusted' component of the antecedent can be ignored under ellipsis. This is tested by way of the inferential

<sup>21</sup> D. Brentari tells me (p.c.) that she would label this a 'sign+gesture hybrid'.

questions in (23)a (conditional in form because it pertains to a negative sentence) and (23)b (unconditional in force because it pertains to a positive sentence): endorsement is strong for the antecedent (6 on a 7-point scale), and weak for the elided clause (2 on a 7-point scale). The other sentences of the paradigm are controls: (21)a has no disgusted facial expression, and (21)c involves an explicit at-issue modifier; (21)d-f are unelided versions of (21)a-c. The paradigm in (22) makes the same point, but with *EAT* replacing *USE*: the results are similar.<sup>22</sup>

*Conventions:* \:-( encodes a disgusted facial expression combined with a backwards body/head shift; the word it co-occurs with is boldfaced. We use the same facial expressions in some English translations, with the same boldfacing convention. (Throughout these two paradigms, no acceptability or inferential scores varied by more than two points across trials.)

- (21) [THAT FISH IX-1 WON'T \_\_\_\_ ]<sub>b</sub>. [THAT MEAT IX-1 WILL ... ]<sub>a</sub>.
- a. <sup>7</sup> \_\_\_\_ = USE ... = ∅  
'That fish, I won't use. That meat, I will.'
- b. <sup>6.3</sup> \_\_\_\_ = \:-( **USE** ... = ∅  
'That fish, I won't \:-( **use**. That meat, I will.'
- c. <sup>6</sup> \_\_\_\_ = USE WITH EXPRESSION \:-( ... = ∅  
'That fish, I won't use with this expression: \:-( . That meat, I will.'
- d. <sup>7</sup> \_\_\_\_ = USE ... = USE  
'That fish, I won't use. That meat, I will use.'
- e. <sup>6</sup> \_\_\_\_ = \:-( **USE** ... = \:-( **USE**  
'That fish, I won't use \:-( **use**. That meat, I will \:-( **use**.'
- f. <sup>7</sup> \_\_\_\_ = USE WITH EXPRESSION \:-( ... = USE WITH EXPRESSION \:-  
'That fish, I won't use with this expression: \:-( . That meat, I will use with this expression: \:-(' (ASL, [35,0170](#); 3 judgments)
- (22) [THAT FISH IX-1 WON'T \_\_\_\_ ]<sub>b</sub>. [THAT MEAT IX-1 WILL ... ]<sub>a</sub>.
- a. <sup>7</sup> \_\_\_\_ = EAT ... = ∅  
'That fish, I won't eat. That meat, I will.'
- b. <sup>7</sup> \_\_\_\_ = \:-( **EAT** ... = ∅  
'That fish, I won't \:-( **eat**. That meat, I will.'
- c. <sup>6.3</sup> \_\_\_\_ = EAT WITH EXPRESSION \:-( ... = ∅  
'That fish, I won't eat with this expression: \:-( . That meat, I will.'
- d. <sup>7</sup> \_\_\_\_ = EAT ... = EAT  
'That fish, I won't eat. That meat, I will eat.'
- e. <sup>6.7</sup> \_\_\_\_ = \:-( **EAT** ... = \:-( **EAT**  
'That fish, I won't \:-( **eat**. That meat, I will \:-( **eat**.'
- f. <sup>6.7</sup> \_\_\_\_ = EAT WITH EXPRESSION \:-( ... = EAT WITH EXPRESSION \:-  
'That fish, I won't eat with this expression: \:-( . That meat, I will eat with this expression: \:-(' (ASL, [35,0164](#); 3 judgments)

<sup>22</sup> In *EAT*, the dominant hand moves towards the mouth in a somewhat iconic fashion (*USE* just involves a *U* touching the non-dominant hand). One might thus expect the disgusted facial expression to be easily interpreted as an iconic modulation with *EAT* (see Meir 2010, Emmorey 2014 for a discussion of iconic implications of a somewhat similar sign in Israeli Sign Language; this point is also discussed in 'Visible Meaning'). Despite this, the disgusted expression triggers a cosupposition and can be disregarded under ellipsis.

(23) Inferential questions and inferential strength in (21)/(22)<sup>23</sup>

Does the sentence suggest that the following is the case? (1 = no inference; 7 strongest inference)

WON'T cosupposition: If the signer were use/eat the fish, he would display a disgusted expression.

WILL cosupposition: The signer will use/eat the meat with a disgusted expression.

Conditions / Strength of cosuppositions	WON'T cosupposition (21) (=ASL, 35, 0170)	WILL cosupposition (21) (=ASL, 35, 0170)	WON'T cosupposition (22) (=ASL, 35, 0164)	WILL cosupposition (22) (=ASL, 35, 0164)
a. neutral, ellipsis	1	1	1	1
b. co-sign facial expression, ellipsis	<b>6</b>	2	<b>6</b>	2
c. at-issue modifier control, ellipsis	2.7	<b>5</b>	2	4.7
d. neutral, explicit repetition	1	1	1	1
e. co-sign facial expression,	<b>5.7</b>	7	<b>5.7</b>	7
f. at-issue modifier control, ellipsis	2	7	2	7

In sum: as the highlighted line of results in (23)b shows, the disgusted facial expression triggers a cosupposition in the antecedent clause, and the facial expression can be disregarded in the elided clause.

What should we conclude from these results? Brentari 2018 offers a typology of the behavior of visual enrichments under ellipsis based on their differential salience in different modalities, with visual enrichments being less salient in the spoken modality.<sup>24</sup> We have a slightly different take: sometimes gestures cannot be disregarded under ellipsis in English, as in the of pro-speech gestures discussed in (15)a; sometimes a visual enrichment *can* be disregarded in ASL, as in the case of disgusted facial expressions. These analyses follow from a more abstract view, corresponding to the typology in (7): elements that can be construed as external enrichments to a word can be excised from it and for this reason do not have to be copied under ellipsis. By contrast, iconic modulations are not so easily excised and are harder to ignore (but iconic modulations that induce cosuppositions might be ignored under ellipsis, as we will suggest in Section 3.3.2).

### 3.3 Do some signs include incorporated gestures?

#### 3.3.1 The claim: Aristodemo 2017, Aristodemo and Santoro 2018

As summarized above, Aristodemo 2017 and Aristodemo and Santoro 2018 make the provocative but well-argued claim that some signs of LIS include incorporated gestures, ones that are an integral (and obligatory) component of the signs. As alluded to above, the logic is as follows: these iconic elements "(a) (...) are non-assertive; (b) they trigger a conditionalized presupposition [i.e. what we called above a 'cosupposition' -PS] of type 'if p then q' where p is the assertive component of the modified expression and q is the semantic contribution of the co-speech gesture/iconic component; (c) they can be disregarded under ellipsis". Part (b) is illustrated in (24).

(24) GIANNI GLASS IX-3-poss FULL NOT.

'Gianni's glass is not full.'

=> If Gianni's glass were full, it would be completely full.

Aristodemo and Santoro 2018 add relevant data pertaining to Action Role Shift: in (25), the modulation of *EAT-rep<sub>fast+dirty</sub>* can be disregarded under ellipsis.

(25)

PIETRO THINK^SURE YESTERDAY GIOVANNI<sub>a</sub> <sup>rs</sup> EAT-rep<sub>fast+dirty</sub>; MARIA<sub>b</sub> NOT.

'Pietro believes that yesterday Giovanni ate quickly getting dirty, Maria does not believe that yesterday Giovanni ate quickly getting dirty'

'Pietro believe that yesterday Giovanni ate quickly getting dirty, Maria does not believe that yesterday Giovanni ate '

We find Aristodemo's original data very convincing to establish her claim, but we believe that the new data involving Role Shift might give rise to several theoretical directions. First, the cited data

<sup>23</sup> Inferential scores at or above 5 are boldfaced.

<sup>24</sup> Specifically, Brentari 2018 writes: "... visual prosody may be treated differently in spoken vs. signed language. Why would this be the case? It could be a modality effect. The facial gestures in ASL occur in the same modality as the linguistic message and may therefore be more salient."

do not yet show that the contribution of *-rep<sub>fast+dirty</sub>* is cosuppositional or even just non-at-issue: embedding tests (e.g. in questions or under negation) would be needed to establish this point. Second, the ellipsis data could be analyzed as suggested by Aristodemo and Santoro 2018, but possibly in different ways as well.

First, the pluractional component contributed by *-rep* might behave like a plural in English, which ranges over singular and plural objects alike. This is the reason *Ann doesn't think that John ate pies* entails that she doesn't think he ate *any* pie, rather than the weaker claim that she doesn't think that he ate *at least two* pies (whether this is so with repetition-based plurals or pluractionals would have to be investigated in greater depth; for relevant plural data, see Koulidobrova 2018, Schlenker and Lamberton to appear, Kuhn 2015, Kuhn and Aristodemo 2017). Second, the *fast+dirty* component is due to two main elements: furrowed eyebrows, which usually mark intensification; and a mouthing that usually goes with LIS for *stain* (Aristodemo, p.c.). Thus it is conceivable that these expressions can be treated as co-sign gestures. If so, one should be able to disregard them under ellipsis (as we discussed in Section 3.2.1). Aristodemo's original data were in this respect different, since it was an iconic manual component that was responsible for the cosuppositional inference – and the sign without that component just didn't seem to be possible.

Aristodemo's original data thus require that some co-speech gestures should be obligatory (since *FULL* without its iconic component is ill-formed).<sup>25</sup> In addition, her proposal paves the way for analyses of further phenomena along the same lines. Height and locative specifications of sign language pronouns can be disregarded under ellipsis and *ONLY*, but the precise conditions are still a matter of debate (Schlenker 2014, Schlenker, to appear c). These specifications were originally analogized to *phi*-features of pronouns, and thus possibly inherited by an agreement phenomenon. But several examples suggest that they can be disregarded under ellipsis in a broader class of contexts. If so, the reason might (but need not) be that they are in fact incorporated gestures rather than feature-like elements, a view considered in Zuchi 2018, among others. A very detailed investigation will be needed to adjudicate the issue.

### 3.3.2 *Cosuppositions triggered by purely iconic elements*

We believe, however, that **there is an alternative analysis of Aristodemo's and Aristodemo and Santoro's data**. The heart of the matter is that **some purely iconic constructions can trigger cosuppositions although there is no evidence whatsoever that they are made of a gestural and a non-gestural part**. These constructions are of two kinds: predicate classifiers in ASL, and pro-speech gestures in English. In these cases, it is plausible that a conceptual analysis of their meaning contribution is responsible for their cosuppositional behavior. If so, this conceptual analysis might be available for signs as well. In the case of cosupposition-inducing pro-speech gestures, it also seems that their cosuppositional contribution can be disregarded under ellipsis (on the other hand, comparable ASL data do not bear this out as things stand).

The evidence for a cosuppositional effect stems from (13)b above, which unlike the at-issue control in (13)c, gives rise to a strong (5.7 out of 7) endorsement of the cosuppositional inference in (26):

(26) If the helicopter were to go from Boston to NYC within the next hour, it would make a smooth detour.

Our data do not show that this cosuppositional inference can be disregarded under ellipsis. We leave it open whether this is because of the ambiguity noted by Aristodemo and Santoro 2018 (= ignoring the cosupposition in the elided clause is optional), or possibly because our contexts did not sufficiently motivate why one should disregard this contribution.

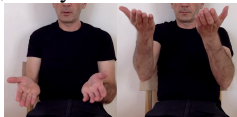
Importantly, there is no *prime facie* reason to take one part of the path displayed by the helicopter to be linguistic and another part to be gestural, especially when one notes that different modulations of the path can give rise to presuppositions rather than to cosuppositions: as discussed in Schlenker 2018c, the 'orthogonal detour' and 'pause in the middle' conditions in (13)d and (13)f trigger a (weak) presupposition that the helicopter will in fact go from Boston to New York.

The same conclusion can be drawn on the basis of pro-speech gestures in English (Schlenker

<sup>25</sup> This might cause difficulties for the alternative definition of gestures proposed in fn. 11.

2018c). (27) contrasts two realizations of a lifting gesture: a neutral lifting gesture, glossed as *LIFT* in (27)a; and a lifting gesture realized with difficulty (trembling hands), glossed as *LIFT-difficult* in (27)b. In addition, we include a gesture-free at-issue control, as in (27)c. Acceptability was rated by three native speakers of American English on a 7-point scale (ratings appear at the beginning of the relevant constructions). We simultaneously asked our informants to assess the strength (also on a 7-point scale) of the cosuppositional inference: *if the speaker were to lift the child, effort/difficulty would be involved*.

(27) This child, will you



a. <sup>6</sup> LIFT\_?

Strength of the cosupposition: 1

b. <sup>5.3</sup> LIFT-difficult?

Strength of the cosupposition: 4.7

c. <sup>5</sup> lift with difficulty?

Strength of the cosupposition: 1.3

(video 01, a,b,d; 3 informants)

The results suggest that we obtain a weak cosupposition under questions with *LIFT-difficult* in (27)b but not with the at-issue control in (27)c. Embedding under other operators confirms this pattern of projection (Schlenker 2018c).

It remains to determine whether this inference can be disregarded under ellipsis. The paradigm in (28) suggests that it can be.

(28) This [big]<sup>26</sup> adult, I won't \_\_\_\_\_, but this small child, I will.

a. <sup>6</sup> LIFT

Unelided cosupposition: 1

Elided cosupposition: 1

b. <sup>6</sup> LIFT-difficult

Unelided cosupposition: 6

Elided cosupposition: 1.3

c. <sup>5.3</sup> lift like this LIFT-difficult

Unelided cosupposition: 2

Elided cosupposition: 6.7

(video 14; 3 informants)

Here each sentence gave rise to two inferential questions, one for the unelided clause ('unelided cosupposition'), and one for the elided clause ('elided cosupposition'):

(i) Unelided cosupposition: If the speaker were to lift the big adult, the movement would be slow/difficult.

(ii) Elided cosupposition: The speaker will lift the small child with a slow/difficult movement.<sup>27</sup>

Endorsement strengths show that the unelided clause in (28)b gives rise to the (expected) cosupposition, but that the elided clause doesn't: the cosupposition can be disregarded under ellipsis. The opposite pattern is found in the at-issue control in (28)c: no cosuppositional inference is found under negation (which is expected since the contribution of the modifier is at-issue); but the contribution of the modifier cannot be disregarded under ellipsis – which is also expected given standard assumptions on ellipsis resolution (the entire VP must be copied, including the at-issue modifier).

<sup>26</sup> The survey contains 'fat' instead of 'big'.

<sup>27</sup> The two inferences had to be stated differently because the elided clause is assertive, and thus the cosupposition must be combined with the asserted component. Specifically, if we have both the assertion that *this small child, I will lift*, and the cosupposition that *if I were to lift this small child, the movement would be slow/difficult*, we also get the inference that *I will lift this small child with a slow/difficult movement*. One then has to test the latter, unconditional inference (the conditional inference alone would be underinformative, and might be rejected for this reason). Importantly, the 'elided cosupposition' inference does not distinguish between a genuine cosupposition and an at-issue modifier, which is the reason we get a strong endorsement of the inference in the elided part of (28)c. To assess the genuinely *cosuppositional* nature of the inference, we need to consider the strength of the inference pertaining to the antecedent.

**We conclude that a cosupposition triggered by a pro-speech gesture can in this case be disregarded under ellipsis.** Overall, we believe our data suggest an alternative analysis of Aristodemo and Aristodemo and Santoro's data. First, as we saw with the ASL predicate classifier in (13), some presuppositions are triggered by some path modifications, and there is no available justification for treating these as co-sign gestures (especially since other path modifications do not give rise to cosuppositions). Second, we reach the same conclusion on the basis of the pro-speech gesture *LIFT-difficult* in (28)b: not only does it trigger a presupposition, but it also allows it to be disregarded under ellipsis (something that our ASL data do not show). Here too, there is no independent evidence for treating the *difficult* component (realized by a trembling hand) as a 'co-gestural gesture'. Rather, these data suggest that cosuppositions can be triggered by a broader class than was initially thought, and that there might be something about cosuppositions that makes them easy to disregard under ellipsis (see Schlenker 2018c for some theoretical directions). If so, Aristodemo and Aristodemo and Santoro's data might not show that some LIS signs include incorporated gestures.<sup>28</sup>

### 3.4 From co-speech to pro-speech gestures: gestural grammar

We will now suggest that pro-speech gestures rather than just co-speech gestures ought to be investigated in order to determine which properties of signs can be replicated with gestures. (The limitations of these gestures should go without saying, since they neither come with a real lexicon nor with a developed grammar; at best, they can be expected to display some proto-properties of some signs. The fact that gestures have a proto-grammar in no way implies that signs are 'mere' gestures.)

#### 3.4.1 When non-signers know some properties of sign language

How much of sign language grammar can be replicated with gestures? As reported in Brentari 2018, Brentari et al. 2018 "found that if ASL sentences expressing commands, statements of permission, explanation, and advice are shown to ASL signers and to English-speaking non-signers, the two groups of participants have a similar pattern of judgments"; similar findings are reported in Kuhn and Chemla 2017. As Brentari notes, "the results from Strickland et al. (2015) might also be considered from the same perspective. Their study found that nonsigners were able to attribute telic and atelic meanings to forms produced with movements of the types predicted in Wilbur (2010); however one would have to ask if speakers be equally sensitive to these modifications of form if they occurred as co-speech gesture."

#### 3.4.2 Replicating some properties of sign language grammar with co-speech gestures?

How far can we go in this direction? Ebert 2018 sketches a systematic attempt based on co-speech gestures. Here we highlight some differences between Ebert's co-speech examples and our sign language data.

As mentioned above, ASL and LSF provide examples in which plural loci simultaneously function as grammatical elements and as diagrams that display the subset and complement set relations among their denotations; as a result, a 'complement set' locus *b* appears as soon as a large locus *ab* and a subset locus *a* have been introduced. Ebert notes that maybe "complement anaphora in general is a phenomenon that draws on (implicit) pointing", and she mentions as an argument that in German those cases that allow for complement set anaphora require the demonstrative-friendly pronoun *die* rather than the unmarked pronoun *sie*:

- (29) 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.'

---

<sup>28</sup> One important caveat is in order: as noted in Schlenker 2018c, there are alternative potential analyses of the cosuppositions triggered by pro-speech gestures and predicate classifiers, in particular in terms of more standard presuppositions. Furthermore, the behavior under ellipsis of iconic presuppositions which are not cosuppositional in nature is ill-understood and will require much more work.



But it isn't clear to us that this fact directly bears on the availability of iconically inferred discourse referents in sign language, for there might be two separate issues at stake. First, as pointed out by Nouwen 2003, context may sometimes suffice to make some discourse referents salient; it seems relatively unsurprising that in this case a demonstrative-friendly pronoun should be preferred. Second, the 'complement set' data in sign language involve a crucial use of iconicity: a plural pronoun, realized by way of ('arc') pointing, is insufficient to bring out the expected reading when default loci are used (Schlenker et al. 2013); but it is only in case a complement set locus can be iconically inferred that the desired reading arises: with *MOST*, the complement set is otherwise unavailable.

Ebert 2018 further argues that Locative Shift can be replicated in gestures. In sign language, a locus associated with a spatial location can sometimes be co-opted to refer to an individual found at that location. Ebert notes that co-nominal pointing in (30)b produces the same general information as Locative Shift in ASL.

*Notation for pointing gestures:* As in sign language, pointing gestures are alphabetized from the dominant to the non-dominant side from the speaker's perspective. *IX-a* encodes pointing with a finger towards position *a*, while *IX-hand-a* encodes pointing with an open hand, palm up, towards position *a*.

- (30) John owns an apartment in IX-a [**a French city**] and one in IX-b [**an American city**].  
 a. His apartment / One of his apartments is nice.  
 b. IX-b [**His apartment**] / IX-b [**One of his apartments**] is nice.

As Ebert notes, there are reasons to think that co-speech pointing yields presuppositional rather than at-issue inferences, contrary to what is claimed about some (though not necessarily all) instances of Locative Shift in 'Visible Meaning' and in Schlenker, to appear d. But as Ebert argues, one could posit that 'local accommodation' is responsible for these apparently at-issue uses.

Still, we believe that a natural analysis of Ebert's data involves locative rather than individual pointing, with different constraints from Locative Shift. One important finding about Locative Shift is that it cannot easily target indexical loci, in particular ones that correspond to the first and second person. But Ebert's co-nominal pointing can, we believe, co-occur with such pronouns:

- (31) Last year you spent your summer IX-hand-a [**in Paris**], and next year you'll spend it IX-hand-b [**in New York**]. I am sure IX-b [**you**] will have a good time.

Here we associate a locus with *in New York* by way of hand pointing, and retrieve this location-denoting locus by way a pointing gesture co-occurring with *you*. If this yielded a genuine equivalent of Locative Shift, we would expect the result to be degraded, but we don't think it is.

### 3.4.3 The dawn of pro-speech gestures

We believe that the most promising paradigms to replicate some grammatical properties of signs in gestures involving pro- rather than co-speech gestures. Let us give a few examples (but see Schlenker 2018d for a far more detailed argument, and Schlenker and Chemla 2018 for partial experimental data).

First, **a variety of third person loci can be introduced (by way of co-speech gestures) and retrieved by way of gestural pointing**, yielding readings that parallel those pronouns, as shown in (32). Furthermore, this finding extends to some examples in which pronouns are dependent on quantifiers, as seen in (33). Here we cite examples from Schlenker 2018d, who reports average scores (on a 7-point acceptability scale) from a survey with three informants who are not signers.

- (32) Yesterday I had a long conversation with IX-hand-a [**John**], then with IX-hand-b [**Mary**], then with IX-hand-c [**Sam**]. You know who the company's gonna promote?  
 a. <sup>6.7</sup> IX-a.    b. <sup>6.3</sup> IX-b.    c. <sup>6.7</sup> IX-c.    d. <sup>6</sup> IX-1.    e. <sup>6</sup> IX-2.  
 = John       = Mary       = Sam       = me       = you  
 (Video 3845; Schlenker 2018d)

- (33) Whenever I can hire IX-hand-a [**a mathematician**] or IX-hand-b [**a sociologist**], I pick  
 a. <sup>6.7</sup> IX-a. (= the mathematician)

- b.<sup>6.7</sup> IX-b. (= the sociologist)  
(Video 3927; Schlenker 2018d)

Second, **Strickland et al.'s (2015) findings on non-signers' ability to 'guess' the telic/atelic meaning of signs they've never seen before can be replicated with pro-speech gestures.** As Strickland et al. argue, non-signers know a version of Wilbur's Generalization, according to which verbal signs with sharp boundaries tend to have telic meanings. The same effect can be found with pro-speech gestures if we contrast a gesture representing a boundary-less version of a circling gesture representing as parachute, glossed as *CIRCLE-cont*, with a version with a sharp ending when the parachute reaches the ground, glossed as *CIRCLE-sharp*. The boundary-less *CIRCLE-cont* version is acceptable with *for five minutes* but not with *in five minutes*, which suggests that it is interpreted as atelic; the opposite pattern is found with *CIRCLE-sharp*, which appears to be interpreted as telic.

- (34) When skydiving tomorrow, you will  
a.<sup>6</sup> *CIRCLE-cont*  
b.<sup>1.8</sup> *CIRCLE-sharp*  
for five minutes.  
(Video 4033; Schlenker 2018d [which provides a link to an anonymized video])

- (35) When skydiving tomorrow, you will  
a.<sup>2</sup> *CIRCLE-cont*  
b.<sup>5.3</sup> *CIRCLE-sharp*  
in five minutes.  
(Video 4039; Schlenker 2018d [which provides a link to an anonymized video])

Third, **for some speakers at least, Locative Shift can be replicated with pro-speech pointing.** Some examples from Schlenker, to appear e are re-assessed the small survey in Schlenker 2018d. A basic case is illustrated in (36)c: instead of pointing towards the gestural locus introduced with *John*, the speaker points towards the gestural locus introduced with *New York*, but to refer to John-in-New York, so to speak.

- (36) Since IX-hand-a **John** can't seem to work/get along with IX-hand-2 **you**, I'll/I am going to have him transferred to IX-hand-c **New York**. And later, if I need to downsize, you know who I'll fire?<sup>29</sup>  
a.<sup>6.3</sup> IX-2. (= the addressee)  
b.<sup>4.7</sup> IX-a. (= John)  
c.<sup>4.3</sup> IX-c. (= John)  
(Video 3991; Schlenker 2018d)

But averages obscure important patterns in this case: Informant 1 obligatorily applies Locative Shift, Informant 2 prohibits it, Informant 3 allows it, as seen by the scores in (37).

- (37) Acceptability ratings for (36)

Video 3991	a.	b.	c.
Informant 1	5	2	5
Informant 2	7	6	1
Informant 3	7	6	7

When it is *denied* that John will be in New York, pointing towards the gestural locus associated with New York to refer to John entirely fails (for all informants), as seen in (38).

- (38) Since IX-hand-a **John** can't seem to work with IX-hand-2 **you**, I won't have him transferred to IX-hand-c **New York**. And if later I need/I later<sup>30</sup> need to downsize, you know who I'll fire?  
a.<sup>6.3</sup> IX-2. (= the addressee)  
b.<sup>6.3</sup> IX-a. (= John)  
c.<sup>1</sup> IX-c.  
(Video 3993; Schlenker 2018d)

<sup>29</sup> The model interchangeably used 'work with' and 'get along with' in different sentences; similar for 'I'll' vs. 'I am going to'.

<sup>30</sup> The model interchangeably used *if later I need* and *if I later need* depending on the sentence.

One important question is whether Locative Shift in pro-speech gestures resembles its sign language counterpart in being degraded with indexical loci (this is all the more important since the *absence* of this effect was an argument we gave *against* the equivalence between Ebert's co-speech examples and the original sign language data). The jury is still out on this matter: some informants disprefer applying Locative Shift in the second person gestural case, but the data are not robust as things stand (see Schlenker 2018d for discussion).

Despite these empirical difficulties, we hope that our examples highlight the fruitfulness of looking for counterparts of sign language grammar in pro-speech gestures, a program that could be systematically developed in the future (see Schlenker and Lamberton, to appear, for a recent application to repetition-based plurals).

Let us add for completeness that pro-speech gestures have recently been argued to be fruitful for purely semantic studies: as argued in Schlenker to appear f, nearly the full typology of natural language inferences (involving presuppositions, implicatures, supplements, expressives and so-called 'homogeneity inferences') can be replicated with just pro-speech and post-speech gestures, including gestures one might never have seen before. If so, some inferential properties that are often taken to originate in lexical entries, such as presuppositions, might be generated 'on the fly' for gestural 'words' one may encounter for the first time. This raises new questions about what is lexically encoded and what is productively generated in the inferential typology of natural language.

## 4 Role Shift: Context Shift, Iconicity and Gestures

### 4.1 Role Shift and context shift

'Visible Meaning' argues, following Quer 2005, 2013, that Role Shift in ASL and LSF presents another case of Logical Visibility. Several spoken languages have been argued to have a covert operation of context shift in the scope of some attitude reports, with the effect that the context of evaluation of some or all indexicals gets 'shifted' (e.g. Schlenker 2003, Anand 2006, Deal 2017). **Quer 2005, 2013 argued that context shift is overt in sign language and is realized by Role Shift.** 'Visible Meaning' makes several additional points, developed further in Schlenker, to appear a,b.

1. First, the typology of context-shifting operators in sign language parallels that found in spoken language: in both cases, some languages allow perspectives to be mixed, while others require that all indexicals 'shift together'. The former pattern ('Mixing of Perspectives') was for instance argued to hold for Amharic in Schlenker 2003 (with opposing views in Anand 2006), for LSC (Catalan Sign Language) in Quer 2005, 2013, for DGS (German Sign Language) in Herrmann and Steinbach 2012 and Hübl and Steinbach 2012, with additional examples from NGT (Sign Language of the Netherlands) in Maier 2016, discussed below. The latter pattern ('Shift Together') was argued to hold for Zazaki by Anand and Nevins 2004, Anand 2006, and is taken to be a characteristic property of context shift in Deal 2017. It is argued to hold for ASL and LSF in 'Visible Meaning' and in Schlenker, to appear a,b.

2. Second, an analysis based on context shift alone is incapable of accounting for quotational effects found under Attitude Role Shift, since such effects do not arise under standard indirect discourse. Schlenker, to appear b proposes that context shift should be supplemented with a rule of 'maximal iconicity' under Role Shift (it is maximal rather than 'exhaustive' because some instances of wh-extraction in ASL could not be handled within a pure quotational analysis).

3. Third, however, a pure quotational analysis would be incapable of accounting for a type of context shift not discussed in spoken language semantics, Action Role Shift, which involves action rather than attitude reports. In ASL and LSF, it also involves maximum iconicity, but quotation can't account for it because no thoughts or words are cited.

Davidson 2015 re-analyzes some of the data from Schlenker, to appear a,b in a different, and very interesting, fashion. She takes iconicity to be at the core of Attitude and Action Role Shift alike,

without context shift. Maier, a long-time proponent of quotational approaches to (apparent) context shift, goes in the same direction.

#### 4.2 Context shift vs. mixed quotation

Kaplan 1989 claimed for theory-internal reasons (having to do with his claim that indexicals are 'directly referential'), and on the basis of English data alone, that no operators can shift the context of evaluation of indexicals. Since quotation establishes a relation between an individual and some words rather than a proposition, no context-shifting operator is needed to derive the *appearance* of shifting in quotation, which therefore poses no problem for Kaplan. Early studies of shifted indexicals emphasized that they did not involve standard quotation of an entire clause because grammatical dependencies that cannot normally "cross quotation marks" were acceptable in these environments, most notably *wh*-extraction and NPI licensing.

But an alternative possibility was that the relevant constructions involved partial quotation of an indexical. As an example, the mixed quotation in (39)a is given, as a first approximation, a meaning akin to (39)b (here we follow Maier 2014a).

- (39) a. Quine said that quotation 'has a certain anomalous feature'.  
 b. Quine said that quotation has what he referred to as 'a certain anomalous feature'.

Analyses based on partial quotation deserve to be explored, but they encounter substantial problems with standard context shift. Still, the problems are different depending on the generalizations one assumes (and it might well be that different generalizations hold in different constructions).

1. The analyses of Schlenker 1999b 2003 took the correct generalizations to involve lexical specifications that allow an indexical to be shifted or not; when shifting was permissible, it was taken to be optional. For instance, it was claimed that in English the expression *the day before yesterday* cannot shift but *two days ago* can (a claim that later literature did not accept, e.g. Anand 2006). Assuming this generalization, the difficulty for a quotational analysis is that one must stipulate that some expressions cannot be partially quoted under attitude reports. Thus Schlenker 1999a proposed a theory in which apparent context shift really involves quotation of a Logical Form (following Larson and Ludlow 1993), with a quotational operator *QUOT*, and the assumption that "constituents of a quoted LF may be scoped out of *QUOT*" and interpreted by the mechanism of 'quantifying in' developed by Kaplan in his solution to Quine's 'Orcutt problem' (Kaplan 1968). But within this general quotation-based solution, one needs to stipulate that some expressions, namely unshiftable indexicals, must obligatorily escape the *QUOT* operator. This is arguably a bit less natural than alternative views on which different indexicals can be *indexed* with different context variables.<sup>31</sup>

Another argument was mentioned speculatively by Schlenker 1999b: a theory of partial quotation might have to allow for quotation of features, i.e. of elements that are so abstract that the notion of 'quotation' might get lost in the process. This was because, in very preliminary data from Amharic, it looked like something akin to *John and Peter said that we-are hero* can be used in a situation in which each thought: *I am a hero*. Incomparably more detailed work was conducted on this topic by LaTerza and colleagues (e.g. LaTerza et al. 2014), with related (but more complex) conclusions.

2. A different and more essential problem arises in cases in which all indexicals must 'shift together' (e.g. in Zazaki according to Anand and Nevins, and in all cases of genuine context shift according to Deal 2017). Anand 2006 showed that in Zazaki there is no requirement that under context shift non-

---

<sup>31</sup> There are precedents in grammar for having constraints on what can bind what (e.g. the syntactician's Binding Theory specifies that local binding of pronouns is prohibited, for instance). In fairness, work on positive polarity items makes use of conditions that prohibit certain elements from being in the scope of some operators (see for instance Szabolcsi 2004, Homer 2004). One could therefore imagine that unshiftable indexicals somehow do not want to be in the scope of quotation operators. One would still need to find a plausible variety of covert movement to account for the readings obtained (so that the unshifted indexicals can syntactically escape the scope of context-shifting operators).

indexical expressions should be quoted. On the other hand, if one indexical gets shifted, all other indexicals in the scope of the same operator must be shifted as well. This presents an essential difficulty for theories of partial/mixed quotation: it is entirely unclear how one could allow for partial quotation while forcing the quotation of one indexical to entail the quotation of all other indexicals in the same clause.

### 4.3 Davidson's theory<sup>32</sup>

Davidson 2015 made a very interesting proposal about the analysis of Role Shift, one that seeks to account for its iconic properties without positing context shift. As summarized in Davidson 2018, she

suggested that action role shift can be modeled within a formal semantic system (...) as a “body classifier” obligatory event modifier, where processes like action depiction and imitation, along with pragmatic constraints, determine the accuracy of the event modifier. For role shift, it also seems that a system of reference tracking by which we can determine which are the same and different actors and participants in a dialogue may play a role in how several iconic aspects of both attitude and action role shift get involved in interpretation.

Several remarks should be made in this connection.

(i) As noted in Section 2.4.2, Davidson's analysis, like other 'like this' analyses of iconicity, is either incorrect or insufficiently specified: the lexical form of the iconically modified words is not interpreted iconically, and thus one must have a semantics that specifies that not all aspects of a sign give rise to an iconic interpretation. Furthermore, the demonstrative *this* can refer to anything provided the context makes it sufficiently salient; not so with iconic rules, which have a logic of their own, as highlighted for instance in Greenberg's work (Greenberg 2013).<sup>33</sup>

(ii) Davidson 2015 adopts a quotational analysis of Attitude Role Shift, but this does not yield a unified analysis of Role Shift. Specifically, for Action Role Shift, she must claim, as was the case in Schlenker, to appear a,b, that the words are both used and mentioned (since they have an iconic component). But once this move is made for Action Role Shift, why couldn't the words in Attitude Role Shift also be simultaneously used and mentioned? If so, we would not have a pure case of quotation, but rather something very close to the analysis of Schlenker, to appear b, with indirect discourse that includes a strong iconic/quotational component – and a relatively unified analysis of Attitude and Action Role Shift (how unified the analysis should be is a matter of debate, however, in particular in view of Quer's data discussed in Section 4.5<sup>34</sup>).

(iii) Let us add that the quotational analysis of Attitude Role Shift won't easily extend to cases in which perspectives can be mixed under Attitude Role Shift, as in LSC, DGS and NGT, with an unshifted and a shifted indexical in one and the same role-shifted clause. Maier 2016, 2018 provides very nice further examples that go in the same direction in NGT:

- (40) a. Martine's original utterance: IX-1 BETTER SIGN THAN MACHA
- b. Macha's report: RS \_\_\_\_\_  
MARTINE IX-1 BETTER SIGN THAN [IX-1]  
'Martine said she is a better signer than Macha.' (NGT, Maier 2016)

<sup>32</sup> Many thanks to Kathryn Davidson for very helpful discussion of these matters.

<sup>33</sup> As mentioned above, we take the *like this* analysis to be methodologically imperfect because it is not explicit about the workings of iconicity; a long-term challenge for sign language semantics is precisely to integrate logical and iconic semantics, which is the reason 'Visible Meaning' sought to give precise examples of how a formal iconic semantics would look like.

<sup>34</sup> Quer's data suggest that LSC lacks Action Role Shift. Kathryn Davidson (p.c.) mentions this might be because ASL and LSF "allow the signer's body to be used as a classifier", whereas LSC does not – an interesting proposal to be investigated in future research.

Davidson leaves open the possibility that ASL differs from some other sign languages with respect to Mixing of Perspectives. This raises a question: do different varieties of Role Shift have something in common, or should they be analyzed in entirely different ways? Some earlier frameworks could be tweaked as to allow for Mixing of Perspectives or Shift Together depending on the language (see Schlenker 2011<sup>35</sup>); it is not entirely obvious how this would work in Davidson's system

(iv) Finally, in Davidson's implementation, Action Role Shift should only apply to verbs, not to larger structures. This is because for her the demonstration we find in Action Role Shift is part of the lexical specification of the verb. But at least some of the data reported in Schlenker, to appear a, b suggest that Action Role Shift can be applied to larger structures, as long as they do not contain full indexical pronouns. Such data should be tested further in future research.

#### 4.4 Maier's proposal

As noted, Maier 2018 observes that Mixing of Perspectives seems to be a genuine possibility under Role Shift in DGS, LSC, and in his own data pertaining to NGT. So he proposes to add to a Davidson-style theory of Role Shift a component of 'unquotation'; in his words, "the idea is that unquoted expressions specify a referent, while the surrounding quotation specifies the literal linguistic form" (see also Maier 2014a,b, 2016). Technically, (40)b is analyzed as in (41), where  $\hat{\ }^{\circ}$  is used for concatenation and  $\ulcorner \dots \urcorner$  for quotation.

(41)  $\exists e \exists e' \subseteq e$  [agent(e, martine)  $\wedge$  form(e,  $\ulcorner$ IX-1 BETTER SIGN THAN $\urcorner$ form(e'))  $\wedge$  referent(e') = macha  $\wedge$  demonstration(d,e)]

In Maier's words, (41) "asserts the existence of an event e with linguistic form given by the sequence of NGT signs *IX-1*, *BETTER*, *SIGN*, and *THAN*, followed by some underspecified sign(s) the reference of which is/are fixed by the unquoted material (*IX-1*) as signed by the report (Macha)".

Interestingly, this mechanism achieves independently, and in a different way, something that the quotation-based theory of Schlenker 1999a attempted (by way of scope) within Larson and Ludlow's (1993) theory of indirect discourse combined with Kaplan's theory of Quantifying In. That analysis started from structures with quotation of the embedded clause *qua* syntactic structure (so as to account for the fact that it seems to be grammatically integrated), but De Re terms were moved out of the scope of the quotation operator and gave rise, semantically, to the insertion of 'vivid names' for their denotations, as is sketched in (42):

(42) a. IX-1<sub>i</sub> [MARTINE (SAY) [IX-1 BETTER SIGN THAN t<sub>i</sub>]]  
 b. [ $\exists a$ : a is a vivid name of the speaker for Martine] Martine utters the Logical form: [IX-1 BETTER SIGN THAN a]

As in Maier's account, but through different technical means, the result was thus to allow for quotations with 'holes' in them corresponding to De Re terms.

As was the case for Schlenker 1999a, lexical differences across indexicals might prove a bit unnatural to state: if some indexicals are shiftable while others are not, one would need to enrich the framework with lexical stipulations concerning what is and what isn't quotable. Thus Herrmann and Steinbach 2012 note that in DGS temporal and locative indexicals yield ambiguities under Role Shift but indexical pronouns and agreement verbs do not. Similarly, Quer 2013 writes about LSC that "there is an important asymmetry between locative and temporal indexicals, on the one hand, and person indexicals, on the other, because the latter do not seem to be able to receive non-shifted interpretations, that is, they are always interpreted in the derived context." From the standpoint of quotational theories, these are surprising lexical differences.

But the key difficulty lies in explaining Shift Together in the languages that have it: why does the fact that one indexical is shifted make it impossible to unshift an indexical in the same clause? Shift Together follows neatly from operator-based analysis of context shift (Anand and Nevins 2004, Anand 2006, Deal 2017), but seems hard to capture in Maier's theory. In addition, typological

<sup>35</sup> The suggested unification in Schlenker 2011 cannot account for the patterns discussed in Deal 2017.

contrasts between those languages that enforce Shift Together (e.g. ASL and LSF) and those that don't (e.g. DGS, LSC, NG) seems hard to capture in Maier's framework.

Maier seeks to go in a purely pragmatic direction to constrain the mechanism of unquotation, but the cross-linguistic differences observed might suggest that more grammatical explanations must be explored.

#### 4.5 *Quer's observation*

Quer 2018 makes a very interesting observation pertaining to Role Shift: "LSC does not seem to allow for the action interpretation even in the absence of any indexical element in the scope of role shift". This calls into question fully unified accounts of Attitude and of Action Role Shift.

The first thing to note is that **Quer's observation makes LSC similar to spoken languages in which context shift has been described, since none of them has clear cases of shifted indexicals outside of attitude reports**. In other words, the typological distinction unearthed by Quer between LSC and ASL/LSF is already one that is thought to hold between Amharic and Zazaki (at least as described in the literature) and ASL/LSF. As a result, context-shifting theories of Role Shift could simply posit that whatever (ill-understood) difference underlies the contrast between ASL/LSF and Amharic/Zazaki also explains that between ASL/LSF and LSC.

Second, however, one might want to go through the exercise of tweaking extant theories of Role Shift for ASL/LSF to adapt them to the LSC case. What would one need to do?<sup>36</sup> In Schlenker, to appear a, the Role Shift operator *RS* was analyzed as a context shifter with the semantics in (43):

(43) Let *c* be a context, *s* an assignment function and *w* a world of evaluation, and let *F* be a clause, then:

$$\llbracket \text{RS}_a F \rrbracket^{c, s, w} = \lambda x'_e \lambda w'_s. \llbracket F \rrbracket^{<x', w>, s, w'}$$

The effect of the Role Shift operator was thus to abstract over the coordinates of the context parameter *c*, analyzed for simplicity as a pair of form <agent, world>. Due to its intensional nature, Action Role Shift (an extensional operation) required a non-standard Logical Form, illustrated in (44), also from Schlenker, to appear a:

(44) a. IX-a  $\frac{\text{RS}_a \text{_____}}{1\text{-WALK-WITH-ENERGY}(\text{CL-ONE})}$

b. **w\*** IX-a  $\text{RS}_a [1\text{-WALK-WITH-ENERGY}(\text{CL-ONE}) \emptyset_1]$

In a nutshell, due to the lexical entry illustrated in (43), the role-shifted clause is predicted to have an intensional type, and as a result Action Role Shift as in (44)a can only be interpreted if one inserts a covert world variable *w\**, boldfaced in (44)b.

Thus a rather technical (and possibly unilluminating) account of the cross-linguistic variation noted by Quer 2018 could go like this: maybe such covert variables are available in some sign languages but not in others. Be that as it may, the broader point is that Attitude and Action Role Shift are fairly different, since the former is extensional while the latter is intensional; this might provide enough room for maneuver to account for Quer's finding. In addition, there is no *prima facie* reason for the LSC Role Shift operator to be semantically identical to that found in ASL/LSF, which brings us back to our first point, pertaining to the existence of an LSC-style pattern in spoken language.

In sum, a more ambitious and explanatory typology of the relevant phenomena should be sought, and Quer's observation will no doubt contribute to it.

#### 4.6 *Eckardt's hypothesis*

Eckardt 2018 asks whether the obligation to shift indexicals under Attitude Role Shift in ASL and LSF could derive from a broader constraint, and she states the following hypothesis:

the use of indexicals under role shift is shaped by the epistemic limitations of the internal speaker. The signing and interpretation of indexicals under role shift must not be based on knowledge that only the external speaker

<sup>36</sup> See fn. 34 for a proposal by Davidson (p.c.) on how to incorporate Quer's findings to her analysis of Role Shift.

can possess. In order to respect the epistemic limitations of the internal speaker, it seems plausible that the assignment of loci to discourse referents undergoes a Reset under role shift.

The proposal could come in different versions depending on whether all context-dependent and all assignment-dependent elements undergo Reset, or only a subclass of them. In particular, Eckardt suggests that in DGS time and place indexicals escape obligatory Reset because they do not come with a locus, whereas in ASL all indexicals are subject to Reset.<sup>37</sup>

Eckardt's hypothesis rightly calls attention to simple but possibly understudied examples with Role Shift, with the schematic form in (45), with the role-shifted cause appearing with subscript *RS*:

(45) Jane<sub>a</sub> arrived. Tom<sub>b</sub> thought [he liked her<sub>a</sub>]<sub>RS</sub>.

Here presumably *he* would be expressed with a first person pronoun, and the question is whether the locus associated with Jane can be re-used within the role-shifted clause. In ASL, there appear to be cases of variables bound from outside Role Shift, as argued for one consultant's ASL in Schlenker, to appear a, which provided examples of *wh*-extraction out of a role-shifted clause:

(46) *Context*: The speaker is in NYC; the listener was recently in LA with John.  
BEFORE IX-a JOHN IN LA [= while in LA],

RS \_\_\_\_\_  
6.7 WHO IX-a SAY IX-1 WILL LIVE WITH HERE WHO

'While John was in LA, who did he say he would live with there?' (ASL, 14, 91; Schlenker, to appear a)

If these data are confirmed, a blanket Reset would be too strong. But this case does not involve the *overt* use within the role-shifted clause of a locus introduced outside of it, and hence Eckardt's question has yet to be answered. Clearly, Maier's data from NGT in (40)b also argue against Reset, since one first person pronoun is shifted while the other is not.

Still, preliminary fieldwork we conducted on ASL for the present piece suggests that Eckardt's hypothesis might be on the right track. In the partial paradigm in (47)<sup>38</sup>, the context guarantees that John and Mary are not present in the reported speech situation, so that the agent could not be referring within Role Shift to their actual positions. Locus cross-reference across Role Shift in (47)a is a bit deviant compared to the version without Role Shift in (47)b (but in neither case does the reading obtained indicate that the pointing signs are quoted).

(47) YESTERDAY ANN<sub>a</sub> THE-TWO-1,<sub>a</sub> SECRET MEETING DISCUSS JOHN<sub>b</sub> MARY<sub>c</sub>. IX-a TELL-1

'Yesterday, Ann and I had a secret meeting to discuss John and Mary. Ann told me

a. <sup>5</sup>RS<sub>a</sub> \_\_\_\_\_

IX-1 DISLIKE IX-b, SAME DISLIKE IX-c.

=> no inference that Ann pointed to refer to John and Mary

'I dislike him and I also dislike her'.', or: that she dislikes him and also dislikes her.'

b. <sup>6.3</sup>IX-a DISLIKE IX-b, SAME DISLIKE IX-c.

=> no inference that Ann pointed to refer to John and Mary

that she dislikes him and also dislikes her.'

(ASL, 34, 4573a,e; 3 judgments)

While the acceptability contrasts are subtle, the consultant's comments are particularly informative, as he explicitly mentions in this connection that loci within Role Shift are deviant unless they are linked to the reported (rather than to the reporting) conversation.<sup>39</sup>

<sup>37</sup> Eckardt's proposal is in some respects reminiscent of a view of context shift advocated for Free Indirect Discourse by Sharvit 2008. For Sharvit, unlike standard indirect discourse (and presumably context-shifting indirect discourse that allows for *wh*-extraction), Free Indirect Discourse abstracts over both the context parameter and the assignment function parameter. We leave a comparison between Sharvit's operator and Eckardt's Reset for future research.

<sup>38</sup> We do not transcribe non-manuals besides Role Shift.

<sup>39</sup> In the consultant's words (from our elicitation notes, second session, 18.08.17): "we would never see someone use [Role Shift] and pointing in this manner, at least not without describing the original conversation and how Ann's pointing fit into that conversation". The sentence would be "much better if loci weren't arbitrary. Let's



A further question is whether varieties of context and variable shift/reset can be derived from more fundamental cognitive considerations. Eckardt is right to raise this question, which she proposes to address as follows:

My commentary proposes the principle of respecting the internal speaker's epistemic limitations as an additional reason behind these patterns. Respect for the epistemic limitations can be implemented in various manners, thus potentially covering not only strictly iconic languages such as ASL and LSF, but also the patterns of DGS or LSC.

The challenge will be to connect the details of current formal analyses to such broader cognitive considerations. One of the difficulties is that loci assigned by the signer are, in non-deictic cases, rather arbitrary in nature, and thus it is not clear why these should be treated any differently from covert variables that might appear in Logical Forms. A further challenge will be to extend these constraints to Action Role Shift, which in our ASL and LSF data does not like to have any indexicals (apart from indexical agreement markers) within its scope.

#### 4.7 *Role Shift in gestures*

Finally, we should ask whether Role Shift can be found in gestures. Ebert 2018 proposes to replicate its properties by way of co-speech gestures; we will suggest that a possibly clearer case can be found in pro-speech gestures (Schlenker, to appear f).

##### 4.7.1 *Ebert on character viewpoint co-speech gestures*

Ebert's starting point is the distinction in gesture research between object viewpoint and character viewpoint gestures: "An O-VPT [= object viewpoint] gesture indicates a bird's eye perspective, a C-VPT [= character viewpoint] gesture suggests that the speaker adopts the perspective of the individual that is impersonated". She further proposes that "C-VPT gestures seem to be the gestural co-speech equivalent to Role Shift in sign languages". This seems to us to be insufficiently fine-grained, for three reasons. First, there are multiple points of view phenomena in language, and Role Shift corresponds to a strict subset of those (arguably related to context-shifting attitude verbs in spoken language). Second, Role Shift is usually realized in a particular way involving body shift and eyegaze shift (see Quer 2013 for further properties), and it would thus be best to seek analogues of these in gestures before establishing a parallel. Third, Role Shift has an effect not just on paralinguistic elements but also on linguistic ones, thus forcing the shifting of indexicals in ASL and LSF. It is not trivial to see how non-context-shifting constructions in English could display a similar behavior, even accompanied with gestures. On the other hand, as a detector of perspectival shift loosely construed (rather than Role Shift in the strict sense), character viewpoint gestures are extremely interesting and ought to be investigated in detail. Thus Ebert argues that a pointing gestures towards the self can co-occur with the attitudinal center in Free Indirect Discourse, but not with other individuals – definitely a fruitful observation.

Be that as it may, Ebert makes the very interesting claim that 'maximal iconicity' characterizes character viewpoint gestures, both in attitudinal environments (purported equivalents of Attitude Role Shift) and in action reports (purported equivalents of Action Role Shift). The two cases are illustrated in (48) and (49) respectively. In both cases, a happy expression cannot co-occur with the whole body impersonating gesture.

- (48) See that arrogant French swimmer? Yesterday he was angry. He said  
 he [would leave].  
 C-VPT WALK-gesture (*whole body impersonating a walking person*)  
 # :-)
- (49) See that arrogant French swimmer? Yesterday he was angry.  
 He [left with energy].

---

say we're in my classroom and John usually sits in the direction of locus b and Mary in the direction of locus c. Then, this sentence would make perfect sense."

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

A complicating factor is that, due to the whole body nature of the gesture, it would certainly be natural to take the facial expression to be part of the gesture. Things are in this respect somewhat different under Action Role Shift in sign language, since there is no way the entire role-shifted expression can be taken to be gestural: it contains *bona fide* words, which furthermore are not quoted. Thus more work will be needed to assess whether Ebert's examples are a gestural version of the 'maximal iconicity' requirement we saw at work in ASL and LSF Role Shift.

#### 4.7.2 *Role Shift in pro-speech gestures?*

We shall now suggest that a close analogue of Role Shift can be found with pro-speech gestures (we follow Schlenker 2018d, but see Lillo-Martin 2012 and Davidson 2015 for some earlier comparisons with English gestural examples, and Quinto-Pozos and Parrill 2015 for a comparison between signers' and speakers' strategies to encode viewpoint in narratives). The example in (50), which requires some explanation about our notations, exemplifies both gestural versions of both Attitude and Action Role Shift.

*Notation:*  $RS_i$  indicates that the speaker shifts his body to adopt the position of a fictional character found in gestural locus  $i$  (here we will have  $i = a$  or  $i = b$ ). The gesture that follows  $RS_i$  is realized from this shifted position.

(50) <sup>7</sup> I was standing next to IX-hand-a [**little Robin**] and IX-hand-b [**little Francis**], and I was holding a really yummy chocolate bar. And I asked: Who wants it? And so of course

(i) IX-hand-a [**little Robin**] goes:  $RS_a$  IX-1. And IX-hand-b [**little Francis**] goes:  $RS_b$  IX-1.

Next thing I know,

(ii) IX-hand-a [**little Robin**] turns to IX-hand-b [**Francis**] and  $RS_a$  SLAP-b. And so IX-hand-b [**Francis**]  $RS_b$  PUNCH-a.

(Video 4053; Schlenker 2018d [which provides an anonymized video])

As stated in Schlenker 2018d, (50)(i) displays an example reminiscent of Attitude Role Shift, here involving direct quotation. The first person pointing gestures are interpreted as quoted material: one understands that Robin and Francis each replied with a first person pointing gesture (if the gesture were repeated, we would understand that they used repeated gestures in their answers); and facial expressions that accompany this gesture will equally be attributed to the relevant child. By contrast, in (50)(ii) we are dealing with a construction reminiscent of Action Role Shift: the gesture is not quoted, but rather it is used to refer to an action. Thus no speech report is involved here, but rather an action report. Still, in (ii) as in (i), it seems to us that the non-manual expressions must be read as being an integral part of the gestures reflecting the character's rather than the speaker's emotions, which might confirm Ebert's hypothesis; this should be further tested, of course.

Be that as it may, we believe pro-speech gestures should in the future offer fertile ground to further explore Ebert's interesting ideas.

## 5 Conclusions

We end by listing some conclusions that can be drawn from this debate.

1. Sign languages have a special role to play for the foundations of semantics, for at least two reasons: they provide cases of Logical Visibility that are not found in spoken languages, and they offer a unique integration of iconic and logical resources.

2. The nature of the interaction between these iconic and logical resources ought to be investigated in much greater detail, both in linguistics and in psycho/neuro-linguistics. It is uncontroversial that iconic rules, by their very nature, are different from standard semantic rules, and share some properties with other cognitive systems such as visual representation. But there is also a precise sense in which iconic contributions cannot be excised from sign language sentences, as the resulting meanings cannot, in the

general case, be obtained by conjoining an iconicity-free sentence with a separate iconic contribution, particularly when at-issue iconic modulations are taken into account. In the end, arguments for a 'one-module' or a 'two-module' theory cannot be evaluated unless precise definitions of 'module' are offered. Neurolinguistic evidence should also be brought to bear on this debate.

3. A comparison between speech with gestures and sign with iconicity can only be developed if 'iconicity' and 'gestures' are conceptually distinguished. Co-speech gestures are unlikely to be an optimal point of comparison for iconic modulations in sign language because the latter often have a different (at-issue) pragmatic status, as well as different formal properties. By contrast, rare cases of iconic modulations of speech (as in the lengthened adjective *looong*) might make for a more fruitful comparison with iconic modulations of signs.

4. This comparison between speech with gestures and sign with iconicity must take into account the full typology of iconic enrichments, which includes iconic modulations, co-speech/signs gestures, post-speech/signs gestures, and pro-speech gestures, as well as the inferential types they trigger (e.g. at-issue, presuppositional, cosuppositional, supplemental).

5. Gestures have at least three contributions to make to current theoretical debates. First, their semantics can be compared to that of iconic enrichments in sign language (although one should not assume that they have the same semantic/pragmatic status). Second, some gestures (especially pro-speech gestures) might make it possible to replicate some grammatical properties of signs, a finding that might matter for studies of homesigners and sign language evolution. Third, gestures raise theoretical questions in their own right: the typology of inferential contributions made by co-, post- and pro-speech gestures ought to be derived; and as we briefly alluded to, different pro-speech gestures make different contributions that fall neatly within different slots of the inferential typology of human language.

6. More generally, the semantic contribution of iconic gestures and signs (such as classifier predicates) can plausibly be understood upon first exposure of a particular iconic form. Due to the workings of iconicity, it is unsurprising that the global information conveyed by these constructions can be deduced in this 'zero-shot' fashion. What is striking, on the other hand, is that this informational content can be divided 'on the fly' among various components of the inferential typology (e.g. at-issue, presuppositional cosuppositional – see Schlenker 2018c). Iconic inferences are thus a rich source of new semantic data and questions, pertaining in particular to the nature of the algorithms that make this classification possible.

7. Future research will require further informant work on sign languages and gestures, and possibly further experimental research as well. But the way forward is also formal: the fine-grained description of the data requires detailed use of categories of formal semantics; and new theoretical challenges raised by these data will require the construction of new algorithms.

### *References*

- Abusch, Dorit: 2015, Possible Worlds Semantics for Pictures. Manuscript, Cornell University.
- Anand, Pranav & Nevins, Andrew: 2004. Shifty Operators in Changing Contexts. R. Young (ed.), SALT XIV 20-37, Ithaca, NY: Cornell University.
- Anand, Pranav: 2006. *De De Se*. Ph.D dissertation, University of California, Santa Cruz.
- Aristodemo, Valentina: 2017, *Gradable Constructions in LIS*. PhD dissertation, Institut Jean-Nicod.
- Aristodemo, Valentina & Santoro, Mirko: 2018, Iconic Components as Gestural Elements: The Case of LIS. *Theoretical Linguistics*.
- Bergen, Leon. 2016. *Joint Inference in Pragmatic Reasoning*. PhD dissertation, MIT.
- Brentari, Diane and Eccarius, Petra. 2010. Handshape Contrasts in Sign Language Phonology. In: Brentari, Diane (ed.) *Sign languages: A Cambridge language survey*, ed. Diane Brentari, 284–311. Cambridge: Cambridge University Press.
- Bhatt, Rajesh and Pancheva, Roumyana: 2006. Conditionals. In *The Blackwell Companion to Syntax*, M. Everaert and H. van Riemsdijk (eds.), Vol 1: 638-687. Boston and Oxford: Blackwell.
- Bittner, Maria. 2001. Topical referents for individuals and possibilities. In: *Proceedings of Semantics and Linguistic Theory XI* (Rachel Hastings, Brendan Jackson, and Zsófia Zvolenszky, eds.), 36–55. CLC, Ithaca.
- Brentari, Diane: 2018, Modality and Contextual Salience in Co-sign vs. Co-speech Gesture. *Theoretical Linguistics*.
- Brentari, Diane; Falk, Joshua; Giannakiou, Anastasia; Herrmann, Annika; Volk, Elisabeth; Steinbach, Markus: 2018. Production and Comprehension of Prosodic Markers in Sign Language Imperatives. *Frontiers in Psychology: Language Sciences* (Special Issue on Visual Language)
- Brody, Michael & Szabolcsi, Anna: 2003. Overt scope in Hungarian. *Syntax* 6.1: 19–51.
- Byrne RW, Cartmill E, Genty E, Graham KE, Hobaiter C, Tanner J. Great ape gestures: 2017, Intentional communication with a rich set of innate signals. *Animal Cognition* 20(4):755–69.
- Chemla, Emmanuel: 2009, Presuppositions of quantified sentences: experimental data. *Natural Language Semantics*, 17(4):299-340.
- Davidson, Kathryn: 2015, Quotation, Demonstration, and Iconicity. *Linguistics & Philosophy*, 38: 477-520.
- Davidson, Kathryn: 2018, What Belongs in the “logical core” of Language? *Theoretical Linguistics*.
- Deal, Amy Rose: 2017, Shifty asymmetries: universals and variation in shifty indexicality. Manuscript, University of California, Berkeley.
- Dingemanse, Mark: 2013, Advances in the cross-linguistic study of ideophones. *Language and Linguistics Compass* 6.654–72. DOI: 10.1002/Inc3.361
- Dingemanse, Mark and Akita, Kimi: 2016. An inverse relation between expressiveness and grammatical integration: on the morphosyntactic typology of ideophones, with special reference to Japanese. *Journal of Linguistics*.
- Dingemanse, M. & Schuerman, W. & Reinisch, E. & Tufvesson, S. & Mitterer, H.: 2016, What sound symbolism can and cannot do: Testing the iconicity of ideophones from five languages. *Language* 92.2 (2016): e117-e133.
- Dohen, Marion. 2005. Deixis prosodique multisensorielle: Production et perception audiovisuelle de la Focalisation contrastive en français. Doctoral dissertation, Institut National Polytechnique de Grenoble.
- Dohen, Marion, and Hélène Loevenbruck. 2009. Interaction of audition and vision for the perception of prosodic contrastive focus. *Language & Speech* 52(2-3): 177-206.
- Ebert, Cornelia: 2018, A comparison of sign language with speech plus gesture. *Theoretical Linguistics*.
- Ebert, Cornelia and Ebert, Christian: 2014, Gestures, Demonstratives, and the Attributive/Referential Distinction. Handout of a talk given at Semantics and Philosophy in Europe (SPE 7), Berlin, June 28, 2014.
- Eckardt, Regine: 2018, Pronouns in Space. *Theoretical Linguistics*.
- Emmorey, Karen. 1999. Do signers gesture? In: Messing, L. S. & Campbell, R. (eds), *Gesture, speech, and sign*, pp. 133–59. Oxford University Press.
- Emmorey, Karen: 2014. Iconicity as structure mapping. *Phil. Trans. R. Soc.* 369(1651).

- <http://dx.doi.org/10.1098/rstb.2013.0301>
- Emmorey, Karen: 2018, Experimental Approaches to Studying Visible Meaning. *Theoretical Linguistics*.
- Emmorey, Karen & Brenda Falgier. 2004. Conceptual locations and pronominal reference in american sign language. *Journal of Psycholinguistic Research*, 33(4). 321-331.
- Emmorey, K., Corina, D., and Bellugi, U.: 1995. Differential processing of topographic and referential functions of space. In K. Emmorey & J. Reilly (Eds). *Language, Gesture, and Space*, pp. 43-62, Lawrence Erlbaum Associates: Hillsdale, NJ.
- Emmorey, K. & Herzig, M.: 2003. Categorical versus gradient properties of classifier constructions in ASL. In K. Emmorey (Ed.), *Perspectives on Classifier Constructions in Signed Languages*, Lawrence Erlbaum Associates, Mahwah NJ, pp. 222–246
- Emmorey, K., McCullough, S., Mehta, S.H., Ponto, L.B., & Grabowski, T.J. (2013). The biology of linguistic expression impacts neural correlates for spatial language. *Journal of Cognitive Neuroscience*, 25(4), 517-533.
- Esipova, Maria: 2016a, Alternatives matter: Contrastive focus and presupposition projection in standard triggers and co-speech gestures. Poster, MACSIM conference, CUNY, October 1, 2016.
- Esipova, Maria: 2016b, Presuppositions under Contrastive Focus: Standard Triggers and Co-speech Gestures. Manuscript, New York University.
- Goldin-Meadow, Susan and Brentari, Diane: 2017, Gesture, sign and language: The coming of age of sign language and gesture studies. *Behavioral and Brain Sciences*, doi:10.1017/S0140525X15001247
- Greenberg, Gabriel: 2013. Beyond Resemblance. *Philosophical Review* 122:2, 2013
- Herrmann, Annika & Steinbach, Markus: 2012. Quotation in Sign Languages – A Visible Context Shift. In I. van Alphen & I. Buchstaller (eds.). *Quotatives: Cross-linguistic and Cross Disciplinary Perspectives*.
- Homer, Vincent: 2004, *Polarity and Modality*. PhD dissertation, UCLA.
- Hübl, Annika & Steinbach, Markus: 2012. Quotation across Modalities: Shifting Contexts in Sign and Spoken Languages. Talk delivered at the workshop *Quotation: Perspectives from Philosophy and Linguistics*, Ruhr-University Bochum.
- Jackendoff, Ray: 2009, Parallels and nonparallels between language and music. *Music Perception*, 26(3), 195-204.
- Jouitteau, Mélanie: 2004, Gestures as Expletives, Multichannel Syntax. In B. Schmeiser, V. Chand, A. Kelleher and A. Rodriguez (eds.), *Proceedings of WCCFL 23*, Cascadilla Press, 422-435.
- Kaplan, David. 1968. Quantifying In. *Synthese* 19,1/2: 178-214.
- Kaplan, David: 1989. Demonstratives. In Almog, Perry, and Wettstein (eds.), *Themes from Kaplan*, Oxford University Press.
- Koulidobrova, Elena: 2018, Counting Nouns in ASL. Manuscript, Central Connecticut State University. <http://ling.auf.net/lingbuzz/003871>
- Kuhn, Jeremy: 2015, Cross-categorial singular and plural reference in sign language. Doctoral dissertation, New York University.
- Kuhn, Jeremy: 2016, ASL Loci: Variables or Features? *Journal of Semantics*. doi: 10.1093/jos/ffv005
- Kuhn, Jeremy and Chemla, Emmanuel: 2017, Facial expressions and speech acts in non-signers. Refereed poster, 6th Meeting of Signed and Spoken Language Linguistics (SSLL 6), Japan.
- Kuhn, Jeremy and Aristodemo, Valentina. 2017. Pluractionality, iconicity, and scope in French Sign Language. *Semantics & Pragmatics*
- Kuhn, Jeremy; Geraci, Carlo; Schlenker, Philippe; Strickland, Brent: 2018, Boundaries in space and time: iconic biases across modalities. Manuscript, Institut Jean-Nicod.
- Larson, Richard K., and Peter Ludlow: 1993, Interpreted Logical Forms. *Synthese* 95,3: 305-55.
- LaTerza, Chris, Ruth Kramer, Morgan Rood, Dustin Chacón, and Jen Johnson. 2014. Plural shifted indexicals are plural: evidence from Amharic. In *The Proceedings of the 44th Annual Meeting of the North East Linguistic Society (NELS 44)*, eds. Jyoti Iyer & Leland Kusmer. 259-269. Amherst: GLSA.
- Lerdahl, Fred and Ray Jackendoff: 1983, *A generative theory of tonal music*. Cambridge, MA: MIT Press.

- Liddell, Scott K.: 2003. *Grammar, Gesture, and Meaning in American Sign Language*. Cambridge: Cambridge University Press.
- Lillo-Martin, Diane: 2012. Utterance reports and constructed action. In R. Pfau, M. Steinbach, and B. Woll, *Sign Language: An International Handbook*. De Gruyter Mouton, 365-387.
- Lillo-Martin, Diane & Klima, Edward S.: 1990, Pointing out Differences: ASL Pronouns in Syntactic Theory. In Susan D. Fischer & Patricia Siple (Eds.), *Theoretical Issues in Sign Language Research*, Volume 1: Linguistics, 191-210. Chicago: The University of Chicago Press.
- Maier, Emar: 2014a. Mixed Quotation. Survey article for the *Blackwell Companion to Semantics*. Manuscript, U. of Groningen.
- Maier, Emar: 2014b. Mixed Quotation: The grammar of apparently transparent opacity. *Semantics & Pragmatics* 7(7) pp 1-67.
- Maier, Emar. 2016. A plea against monsters. *Grazer Philosophische Studien* 93: 363–395.
- Maier, Emar: 2018, Quotation, Demonstration, and Attraction in Sign Language Role Shift. *Theoretical Linguistics*.
- McCormick, Kelly, Jee Young Kim, Sara List, and Lynne C. Nygaard: 2015, Sound to Meaning Mappings in the Bouba-Kiki Effect. *CogSci 2015: Proceedings of the 37th Annual Meeting of the Cognitive Science Society*, Pasadena, CA, 22 Jul. 2015.
- McNeill, David: 2005, *Gesture and Thought*. Chicago University Press.
- Meir, Irit : 2010. Iconicity and metaphor: constraints on metaphorical extension of iconic forms. *Language* 86(4) : 865–896. doi:10.1353/lan.2010.0044
- Nouwen, Rick: 2003. *Plural pronominal anaphora in context*. Number 84 in Netherlands Graduate School of Linguistics Dissertations, LOT, Utrecht.
- Ohala, J. J.: 1994, The frequency code underlies the sound-symbolic use of voice pitch. In L. Hinton, J. Nichols & J. J. Ohala (Eds.), *Sound Symbolism*, 325- 347. Cambridge: Cambridge University Press.
- Potts, Christopher. 2005. *The Logic of Conventional Implicatures*. Oxford Studies in Theoretical Linguistics, Oxford: Oxford University Press.
- Potts, Christopher. 2007. The expressive dimension. *Theoretical Linguistics* 33(2): 165–197. Published with commentaries by several researchers, and replies by Potts.
- Quer, Josep: 2005, Context shift and indexical variables in sign languages. In: *Proceedings of Semantic and Linguistic Theory (= SALT) XV*. Ithaca, NY: CLC Publications.
- Quer, Josep: 2013, Attitude ascriptions in sign languages and role shift. In Leah C. Geer (ed.), *Proceedings of the 13<sup>th</sup> Meeting of the Texas Linguistics Society*, 12-28. Austin: Texas Linguistics Forum.
- Quer, Josep: 2018, On Categorizing Types of Role Shift in Sign Languages. *Theoretical Linguistics*.
- Quinto-Pozos, David and Parrill, Fey: 2015, Signers and co-speech gesturers adopt similar strategies for portraying viewpoint in narratives. *Topics in Cognitive Science* 7, 12–35
- Ramachandran, V. S., & Hubbard, E. M.: 2001, Synaesthesia—A Window Into Perception, Thought and Language. *Journal of Consciousness Studies*, 8(12), 3–34
- Schlenker, Philippe: 1999a, Propositional Attitudes as Indirect Quotation: Evidence from Embedded Indexicals. Handout of a talk given at the Linguistic Society of American's Annual Meeting, January 8, 1999, Los Angeles.
- Schlenker, Philippe: 1999b. *Propositional Attitudes and Indexicality: a Cross-Categorial Approach*. Doctoral dissertation, MIT.
- Schlenker, Philippe: 2003. A Plea for Monsters, *Linguistics & Philosophy* 26: 29-120.
- Schlenker, Philippe: 2011, Indexicality and De Se Reports. In *Semantics*, edited by von Heusinger, Maienborn and Portner, Volume 2, Article 61, pp. 1561-1604, Mouton de Gruyter.
- Schlenker, Philippe: 2013a. Temporal and Modal Anaphora in Sign Language (ASL). *Natural Language and Linguistic Theory* 31(1): 207-234.
- Schlenker, Philippe: 2014. Iconic Features. *Natural Language Semantics* 22(4): 299-356.
- Schlenker, Philippe: 2015a Gradient and Iconic Features in ASL (squib). *Snippets* (Issue 29) doi: 10.7358/snip-2015-029-schl
- Schlenker, Philippe: 2015b, Gestural Presuppositions (squib). *Snippets* (Issue 30) doi: 10.7358/snip-2015-030-schl
- Schlenker, Philippe: 2018a, Gesture Projection and Cosuppositions. *Linguistics & Philosophy* 41,

- 3:295–365.
- Schlenker, Philippe: 2018b, Visible Meaning: Sign Language and the Foundations of Semantics. *Theoretical Linguistics*.
- Schlenker, Philippe: 2018c, Iconic Presuppositions. Manuscript, Institut Jean-Nicod and New York University.
- Schlenker, Philippe: 2018d, Gestural Grammar. Manuscript, Institut Jean-Nicod and New York University.
- Schlenker, Philippe: to appear, a. Super Monsters - Part I. To appear in *Semantics & Pragmatics*.
- Schlenker, Philippe: to appear, b. Super Monsters - Part II. To appear in *Semantics & Pragmatics*.
- Schlenker, Philippe: to appear, c. Locative Shift. To appear in *Glossa*.
- Schlenker, Philippe: to appear, d. Iconic Pragmatics. To appear in *Natural Language & Linguistic Theory*.
- Schlenker, Philippe: to appear, e. Locative Shift. To appear in *Glossa*.
- Schlenker, Philippe: to appear, f. Gestural Semantics: Replicating the typology of linguistic inferences with pro- and post-speech gestures. *Natural Language & Linguistic Theory*.
- Schlenker, Philippe, and Chemla, Emmanuel: 2018, Gestural Agreement. *Natural Language & Linguistic Theory*. 36, 2: 87–625587. <https://doi.org/10.1007/s11049-017-9378-8>
- Schlenker, Philippe, Lamberton, Jonathan & Santoro, Mirko: 2013. Iconic Variables. *Linguistics & Philosophy* 36(2): 91-149.
- Schlenker, Philippe, and Lamberton, Jonathan, to appear. Iconic Plurality. *Linguistics & Philosophy*.
- Sharvit, Yael: 2008. The puzzle of free indirect discourse. *Linguistics and Philosophy* 31: 353 - 395.
- Steinbach, Markus and Onea, Edgar: 2016, A DRT analysis of discourse referents and anaphora resolution in sign language. *Journal of Semantics* 33:409–448.
- Strickland, B., Geraci, C., Chemla, E., Schlenker, P., Kelepir, M., and Pfau, R.: 2015. Event representations constrain the structure of language: Sign language as a window into universally accessible linguistic biases. *Proceedings of the National Academy of Sciences*, 112(19):5968–5973.
- Szabolcsi, Anna: 2001, The syntax of scope. In: Mark Baltin and Chris Collins (eds.), *Handbook of Contemporary Syntactic Theory*, pp. 607–633. Oxford: Blackwell.
- Szabolcsi, Anna: 2004, Positive Polarity – Negative Polarity. *Natural Language & Linguistic Theory* 22,2: 409–452.
- Tieu, Lyn, Robert Pasternak, Philippe Schlenker, & Emmanuel Chemla: 2017a. Co-speech gesture projection: Evidence from truth-value judgment and picture selection tasks. *Glossa: A journal of general linguistics* 2(1).
- Tieu, Lyn, Robert Pasternak, Philippe Schlenker & Emmanuel Chemla: to appear, Co-speech gesture projection: Evidence from inferential judgments. *Glossa: A journal of general linguistics* 2(1).
- Tieu, Lyn; Schlenker, Philippe; & Chemla, Emmanuel: 2018, Linguistic inferences without words: Replicating the inferential typology with gestures. Manuscript.
- Wilbur, Ronnie B.: 2003. Representations of telicity in ASL. *Chicago Linguistic Society* 39 : 354-368.
- Wilbur, Ronnie B.: 2010, The semantics-phonology interface. In Diane Brentari, ed., *Sign Languages: A Cambridge Language Survey*, 355-380. Cambridge, MA: Cambridge University Press.
- Zacks, Jeffrey M., Tversky, Barbara, and Iyer, Gowri: 2001, Perceiving, remembering, and communicating structure in events. *Journal of Experimental Psychology: General*, 130, 29–58.
- Zucchi, Sandro: 2011. Event Descriptions and Classifier Predicates in Sign Languages. Presentation given at FEAST in Venice, June 21, 2011.
- Zucchi, Sandro: 2018. Sign Language Iconicity and Gradient Effects. *Theoretical Linguistics*.