

# Iconic Pragmatics\*

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

November 5, 2017

To appear in *Natural Language & Linguistic Theory*

**Abstract:** Recent comparative semantics suggests that sign language makes use of the same logical resources as spoken language, but has richer mechanisms of iconic enrichment. It is often thought that when speech is analyzed in tandem with iconic gestures, 'expressive parity' is regained between speech and sign. But this comparison between sign-with-iconicity and speech-with-gestures turns out to be complex because iconic enrichments often have a non-at-issue status, as attested by their interaction with logical operators. We argue that the status of iconic enrichments is constrained by two parameters:  $\pm internal$ ,  $\pm separate\_time\_slot$ . If an enrichment is effected by the internal modification of an expression ( $+internal$ ) – e.g. by lengthening the word *loooong* in English, or the sign *GROW* in ASL – it can have any semantic status and can in particular be at-issue. If an enrichment is an external addition to an expression ( $-internal$ ) – as is the case of co-speech gestures in English – it does not make an at-issue contribution, but it may have the status of a presupposition or of a supplement. If an enrichment has a separate time slot ( $+separate\_time\_slot$ ), it may not be trivial (= presupposed), and must thus be at-issue or supplemental. The generalization is assessed on the basis of vocal iconicity in spoken language, iconic modulations in sign language, co-speech/co-sign as well as post-speech/post-sign gestures and facial expressions in spoken and sign language, and also gestures that fully *replace* words in spoken language. Our typology suggests that there are systematic expressive differences between sign-with-iconicity and speech-with-gestures, and also that the semantic status of iconic enrichments can in part be predicted by parameters pertaining to their form.

Keywords: sign language; gestures; iconicity; presuppositions; supplements; cosuppositions; co-speech gestures; post-speech gestures; post-speech gestures

---

\* **ASL consultant: 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 particularly grateful to the following colleagues for discussions and objections: Diane Brentari, Dylan Bumford, Emmanuel Chemla, Caterina Donati, Masha Esipova, Susan Goldin-Meadow, Gabe Greenberg, Jeremy Kuhn, Salvador Mascarenhas, Rob Pasternak, Matthew Stone, Brent Strickland, as well as the participants of my seminar at NYU in Fall 2016. Special thanks to Rob Pasternak for discussion of numerous data points. Thanks also to Lucie Ravoux for checking averages and working on the bibliography. All remaining errors are mine.

*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-10-LABX-0087 IEC et ANR-10-IDEX-0001-02 PSL\*.

<b>1</b>	<b>Introduction</b>	<b>3</b>
1.1	The importance of iconic enrichments	3
1.2	Cast of characters	4
1.2.1	<i>Spoken language</i>	4
1.2.2	<i>Sign language</i>	6
1.3	Typology and proposed generalization	8
1.4	Elicitation methods: ASL	9
<b>2</b>	<b>Iconic modulations</b>	<b>10</b>
2.1	At-issue iconic modulations	10
2.1.1	<i>Iconic modulations of spoken words</i>	10
2.1.2	<i>Iconic modulations of signs</i>	11
2.2	Non-at-issue iconic modulations	15
2.3	The connection with ideophones	16
2.4	Towards a formal analysis of elementary iconic enrichments	17
<b>3</b>	<b>Co-speech gestures: English</b>	<b>17</b>
3.1	Supplements or presuppositions?	17
3.2	Cosuppositions	18
3.3	Cosuppositional behavior: propositional cases	19
3.4	Cosuppositional behavior: quantified cases	20
3.5	Experimental approach	21
3.6	Cosuppositions triggered by facial expressions	22
<b>4</b>	<b>Co-sign gestures: ASL</b>	<b>23</b>
<b>5</b>	<b>Co-speech and co-sign gestures: further issues</b>	<b>26</b>
5.1	Deriving cosuppositions	26
5.2	Event semantics implementation	28
5.3	At-issue use of some co-speech gestures	28
5.4	At-issue use of some co-sign gestures?	29
<b>6</b>	<b>Post-speech and post-sign gestures: English and ASL</b>	<b>29</b>
6.1	Post-speech gestures: English	30
6.2	Post-sign gestures: ASL	31
6.3	Post-speech vocal gestures	33
<b>7</b>	<b>Post-speech and post-sign gestures: further issues</b>	<b>34</b>
7.1	An at-issue semantics?	34
7.2	A cosuppositional analysis (Krifka, Pasternak, p.c.)?	35
7.3	Intermediate conclusion	36
<b>8</b>	<b>Pro-speech gestures</b>	<b>37</b>
8.1	Pro-speech gestures can be at-issue	37
8.2	Pro-speech gestures need not be associated with covert spoken words	38
8.3	Pro-speech gestures can trigger presuppositions (just like normal words)	39
8.4	Pro-speech vocal gestures	40
<b>9</b>	<b>Typology and theoretical consequences</b>	<b>40</b>
9.1	Typology	40
9.2	Theoretical consequences	42
9.2.1	<i>Consequences for pragmatics</i>	42
9.2.2	<i>Consequences for the comparison between spoken and sign language</i>	42
<b>10</b>	<b>Conclusions</b>	<b>43</b>
	<b>References</b>	<b>44</b>
	<b>Supplementary Materials: Raw ASL Data</b>	<b>47</b>

## Introduction

### 1.1 *The importance of iconic enrichments*

Recent comparative work in semantics suggests that sign language makes use of the same general logical resources as spoken language, but that it has considerably richer mechanisms of iconic enrichment (e.g. Schlenker et al. 2013, Schlenker, 2017a, to appear a).<sup>1</sup> The situation may be summarized by the following two claims (e.g. Schlenker, to appear a).

(i) **Logical Visibility:** When iconic phenomena are disregarded, speech and sign share the same 'logical spine', including in cases where sign language makes visible the 'Logical Forms' of spoken language sentences – for instance by making overt use of logical indices realized as 'loci', whereas indices are mostly covert in spoken languages (Lillo-Martin and Klima 1990).

(ii) **Iconicity:** Sign language makes use of rich iconic resources, including at its logical core. For instance, sign language loci were argued in recent research to be *both* logical variables and simplified representations of what they denote (see Schlenker et al. 2013, and Liddell 2003 and Kegl 1977/2004 for some sources).

This theoretical situation could lead to two competing conclusions.

- One is that spoken language semantics is, along some dimensions, a 'simplified', iconically defective version of sign language semantics – because the iconic potential of the vocal stream is so limited (outside of the representation of sounds).
- Alternatively, it may be that the 'standard' conception of spoken language semantics was insufficiently inclusive, and that when sign is compared to speech-with-gestures rather than to speech alone, the two systems display similar expressive resources. In Goldin-Meadow and Brentari's words (to appear), "sign should not be compared with speech – it should be compared with speech-plus-gesture"; and it might well be that speech-with-gestures has in the end similar expressive resources as sign-with-iconicity.

Of course it is unlikely that sign language will be more expressive than spoken language along every dimension – for instance, vocal iconicity might be greatly defective in sign language.<sup>2</sup> In addition, spoken language might have all sorts of ways to *paraphrase* in more or less complex ways the iconic enrichments that are found in sign language. Thus the interesting question is whether signed and spoken sentences *which have comparable structures* can also have the same kinds of semantic contributions.

To clarify this debate, we need a semantic study of iconic enrichments across modalities.<sup>3</sup> But as soon as the analysis is made precise, it becomes apparent that not all iconic enrichments have the same semantic status. Notably, it has been argued in recent research that iconic modulations in sign language (e.g. the modulation of the size of the sign for *GROW* in ASL [American Sign Language]) can be at-issue (Schlenker 2013), whereas the enrichments afforded by co-speech gestures in spoken language are not usually at-issue: Ebert and Ebert 2014 take them to be supplements (similarly to the meaning of appositive relative clauses), whereas Schlenker 2015, to appear b claims that they are presuppositional.

We will argue that iconic enrichments afforded by co-speech gestures are not normally at-issue, whereas iconic modulations in sign language can (but need not always) be at issue. As a result, sign language might in certain respects be more expressive than spoken language. Of course the point is not to declare a winner in a somewhat dubious race, but to understand the extent and the sources of semantic variation across modalities. As we will see, a typology of iconic enrichments is thus needed for a proper

---

<sup>1</sup> On the terminological side, we follow Goldin-Meadow and Brentari, to appear, in sometimes referring to spoken language as 'speech' and to sign language as 'sign'.

<sup>2</sup> This claim is a bit less obvious than it might seem. First, there are numerous mouthings in sign language, which can be accompanied with sounds, including when produced by deaf signers. Second, assessing the claim would require that one study both deaf and hearing signers.

<sup>3</sup> Our claims are restricted to the projection of iconic enrichments. Thus we do not seek to explain how iconic meanings arise in the first place (but see for instance Ortega et al., to appear on similarities between sign and gesture iconicity). And we cannot exclude the possibility that non-iconic gestures make different kinds of contributions. For instance, Dohen and Loevenbruck 2009 show that raised eyebrows play a role in focus interpretation in spoken language; these are not iconic gestures and thus fall outside of our generalizations.

understanding of 'Universal Semantics', but it requires careful investigations, which we will begin to sketch in this piece.

## 1.2 *Cast of characters*

While some attention has been devoted to co-speech gestures, we will consider a larger cast of characters, which may be organized among two groups. We call an iconic enrichment *external* if it can be eliminated without affecting the integrity of the sign it modifies, and (therefore) the acceptability of the resulting sentence. An iconic enrichment is *internal* if this is not the case. At this point, we will leave the distinction on this intuitive level, although it should be formalized in future research. The key is that enrichments that are effected by adding to an expression an optional element in a different modality certainly count as external – as is the case for co-speech gestures. By contrast, modulations of a sign or word, obtained for instance by lengthening part of it, should certainly count as internal enrichments.

For simplicity, we will sometimes extend the term 'gesture' to facial expressions, as is relatively standard, but also to 'vocal gestures', so that the term will encompass iconic sounds, notably ones that can follow expressions or sometimes fully replace them. Our impression is that our generalizations about standard gestural enrichments extend to vocal enrichments, which will make this terminological move useful (see for instance Grenoble et al. 2015 for a related terminology). We will give some examples of these similarities, but a full study is left for future research.

### 1.2.1 *Spoken language*

In spoken language, attention is often restricted to co-speech gestures and facial expressions, illustrated in (2), with the notational conventions in (1)a. For legibility, gestures are transcribed in a non-standard font.

#### (1) **Notational conventions: spoken language**

a. A gesture that co-occurs with a spoken word (= a co-speech gesture) is written in capital letters or as a picture (or both) *preceding* the expression it modifies (which will be boldfaced, and enclosed in square brackets if it contains several words).

Examples:

John SLAP **punished** his son.



John SLAP\_ **punished** his son.



John **punished** his son.

b. A gesture that follows a spoken word (= a post-speech gesture) is written in capital letters or as a picture *following* the expression it modifies, and preceded by a dash: – .

Examples:

John punished his son – SLAP.



John punished his son – SLAP\_



John punished his son –


c. A gesture that replaces a spoken word (what we call a 'pro-speech gesture') is written in capital letters, if necessary with an onomatopoeic sound following it (with an 'underscore' connection \_ between the sound and the gesture, as for words modified by co-speech gestures).

Example:

Your brother, I will SLAP\_<phhh>.



Your brother, I will SLAP\_<phhh>.

Your brother, I will \_<phhh>.

d. The same conventions apply to facial expressions, but we also use the 'smiley' :-/ to refer to negative or disgusted expressions.

(2) **Co-speech gestures in spoken language: external enrichment - no separate time slot**


a. John SLAP\_  punished his son.


b. John brought LARGE\_  [a bottle of beer].


c. Sam went :-( [skiing with his parents].

While co-speech gestures are an instance of external enrichment, they are parasitic on the expressions they modify because they do not have their own time slot. In this respect, they differ from gestures that *follow* the expressions they modify – henceforth 'post-speech gestures'. Examples are given in (3), where it is apparent that post-speech gestures come with their own time slot (they are often accompanied with onomatopoeias, a point to which we return below).

(3) **Post-speech gestures in spoken language: external enrichment - separate time slot**

a. John punished his son – SLAP\_ \_<phhh>.

b. John brought [a bottle of beer] – LARGE\_ .

c. Sam went [skiing with his parents] – :-(.

While in simple examples post-speech and co-speech gestures appear to make the same kind of semantic contribution, we will see that upon closer inspection this is arguably not so. Specifically, we will argue that co-speech gestures trigger a variety of presuppositions, whereas post-speech gestures have the semantics of supplements (= the same kind of semantic contribution as appositive relative clauses).

The second subfamily of enrichments we consider are internal, in the sense that they cannot be separated from the spoken expressions they modify, either because they are an integral part of the expressions, or because they fully replace some words. The first case is illustrated by (4)a, where the length of the vowel gives an indication of the length of the talk, hence an intensified meaning.

(4) **Iconic modulations in spoken language: internal enrichment - no separate time slot**

- a. The talk was loooong.
- => the talk was very long
- b. ??The talk was shooooort.


The effect appears to be iconic, as is suggested by the fact that a similar modulation fails in (4)b, which pertains to the adjective *short*. Note that although the enrichment is internal, it shares with co-speech gestures the property of lacking its own time slot, since by definition the modulation co-occurs with the expression it modifies.<sup>4</sup>


<sup>4</sup> Okrent 2002 provides several examples of vocal iconicity:


- (i) a. It was a looooong time.
- b. The bird flew up [high pitch] and down [low pitch].
- c. Work, work, work, rest.

A remarkable use of 'internal enrichment' pertains to cases in which the iconic element cannot be separated from an expression it modifies because it fully replaces some words – and is ineliminable because it fulfills a grammatical function (see Slama-Cazacu 1976, Clark 1996, Fricke 2008, Ladewig 2011). We use the term 'pro-speech gesture' for such cases, as the gesture fully replaces a spoken expression (this is the prefix *pro* we find in *pronoun* and *proconsul*). Examples are given in (5).

(5) **Pro-speech gestures in spoken language: internal enrichment - separate time slot**

- a. Your brother, I am going to SLAP\_ \_ <phhh>.  
=> I am going to slap your brother

- b. In two minutes, our Chair will DOZE-OFF\_ .  
=> in two minutes, our Chair will fall asleep

- c. Dessert was DISGUSTING\_ .  
=> dessert was disgusting

- d. At 2pm yesterday, my car engine was about to <phhh>.  
=> at 2pm yesterday, my car engine was about to explode

It is worth noting that several post- and pro-speech gestures are preferentially accompanied with an onomatopoeia. One reason is probably that this just makes the iconic representation richer and more faithful to the original – and this preference is particularly unsurprising if vocal enrichments are themselves gestures of sorts. In addition, in the case of pro-speech gestures the presence of an iconic sound might help justify the absence of a spoken word, since the iconic sound couldn't co-occur with it (if it could, this would be a case of iconic modulation). An example is given in (5)a, where the gesture for *SLAP* is naturally accompanied with a sound that evokes the noise of a slapping. In such cases, if one wishes to minimally compare different enrichments, one should choose an onomatopoeia that has a non-specific meaning to ensure that most of the semantic contribution is due to the gesture. Importantly, it is also possible to find examples in which a pro-speech or post-speech gesture is silent, as in (5)b, which represents the action of falling asleep (the same gesture can also be used as a silent post-speech gesture). Facial expressions (whether silent or not) can also be used as pro-speech gestures, as is shown in (5)c. Finally, pure onomatopoeias – without special accompanying gestures or facial expressions – can replace speech as well, as illustrated in (5)d.

### 1.2.2 Sign language

Due to the existence of rich iconic mechanisms in sign language, and the fact that gestures and words occur in the same modality, the specific existence of gestures is harder to study in sign. But we will argue that at least part of the same typology of gestural enrichments can be found as in spoken language. One case will be absent at this point, that of pro-sign gestures, as we do not yet know of entirely clear criteria to distinguish pro-sign gestures from signs that are created 'on the fly'; but such criteria will hopefully be developed in the future.

(6) **Notational conventions: sign language**

- a. Standard conventions: sign language sentences are glossed in capital letters, as is standard. Expressions of the form *WORD*<sub>-i</sub> and *WORD*<sub>i</sub> indicate that the relevant expression is associated with the locus (= position in signing space) *i*. A suffixed locus, as in *WORD*<sub>-i</sub>, indicates that the association is effected by pointing; a subscripted locus, as in *WORD*<sub>i</sub> or [...*EXPRESSION*...]<sub>i</sub>, indicates that the relevant expression is signed in position *i*. Locus names are assigned from right to left from the signer's perspective; thus when loci *a*, *b*, *c* are mentioned, *a* appears on the signer's right, *c* on the left, and *b* somewhere in between. *IX* (for 'index') is a pointing sign towards a locus, while *POSS* is possessive; they are glossed as *IX*-*i* and *POSS*-*i* if they point towards (or 'index') locus *i*; the numbers 1 and 2

---

For recent discussions of vocal iconicity, see for instance Clark et al. 2013, Perlman and Cain 2015, and Grenoble et al. 2015.

correspond to the position of the signer and addressee respectively. *IX-i* is a standard way of realizing a pronoun corresponding to locus *i*, but it can also serve to *establish* rather than to *retrieve* one. Agreement verbs include loci in their realization – for instance the verb *a-ASK-I* starts out from the locus *a* and targets the first person locus *I*; it means that the third person individual denoted by *a* asks something to the signer. *IX-arc-i* refers to a plural pronoun indexing locus *i*, as it involves an arc motion towards *i* rather than a simple pointing sign.

b. Iconic modulations are notated as subscripts added to the relevant signs:

POSS-1 GROUP GROW\_ 

c. A facial expression (whether grammatical or not) that co-occurs with some expression is written before that expression (surrounded by square brackets if it contains several words).


Examples: IX-arc-b NEVER :-(  [SPEND MONEY].

IX-arc-b NEVER  [SPEND MONEY].

IX-arc-b NEVER :-( [SPEND MONEY].

d. A facial expression (whether grammatical or not) that follows some expression is written after the expression it modifies, and preceded by a dash: - .

Examples: IX-arc-b NEVER SPEND MONEY - :-( .







IX-arc-b NEVER SPEND MONEY -  .  
IX-arc-b NEVER SPEND MONEY - :-( .

The simplest case of iconic enrichment in sign language pertains to iconic modulations. To see an intuitively clear example, consider the verb *GROW* in (7), which can be realized in a variety of ways, six of which were tested in (8); in the 'slow movement' line, we have included pictures of the beginning and endpoint of *GROW*.<sup>5</sup>

(7) POSS-1 GROUP GROW.

'My group has been growing.' (ASL; 8, 263; Schlenker et al. 2013)

(8) Representation of *GROW*

	<b>Narrow endpoints</b>	<b>Medium endpoints</b>	<b>Broad endpoints</b>
<b>Slow movement</b>	small amount, slowly 	medium amount, slowly 	large amount, slowly 
<b>Fast movement</b>	small amount, quickly 	medium amount, quickly 	large amount, quickly 

As is seen in (8), the sign for *GROW* starts out with the two hands forming a sphere, with the closed fist of the right hand inside the hemisphere formed by the left hand; the two fists then move away from each other on a horizontal plane (simultaneously, the configuration of the right hand changes from closed to open position). The signer varied two main parameters in (8): the distance between the

<sup>5</sup> The discussion of *GROW* borrows from Schlenker et al. 2013.

endpoints; and the speed with which they were reached (only the first parameter is depicted). All variants were entirely acceptable, but yielded different meanings, illustrated in (8). Intuitively, there was a mapping between the physical properties of the sign and the event denoted: the broader the endpoints, the larger the final size of the group; the more rapid the movement, the quicker the growth process.

What about co-sign and post-sign gestures? We will initially focus on non-grammatical facial expressions, which have the advantage of being more easily distinguishable from signs than manual gestures are (this is not an all-or-nothing affair: as we will see, there are reasons to think that in some cases a non-grammatical facial expression can be viewed as an iconic modulation of a sign rather than as a co-sign gesture). In the simplified examples in (9) (which are discussed with a more complete context below), a disgusted facial expression follows a clause, and contributes the information that the fact referred to by that clause is unfortunate (with an alternative understanding for the co-sign case, namely that the action denoted by the VP is difficult to achieve).

(9) **Post-sign facial expressions in sign language: external enrichment - separate time slot**

<sup>7</sup>...[POOR STATE-rep IX-arc-b NEVER SPEND MONEY]<sub>b</sub> - :-().

"... Poor states never spend money (and that's unfortunate)." (ASL, 34, 1700b; 2 judgments)

(10) **Co-sign gestures in sign language: external enrichment - no separate time slot**






<sup>7</sup>...[POOR STATE-rep IX-arc-b NEVER :-( [SPEND MONEY].

"... Poor states never spend money (and it would be difficult or bad for poor states to spend money)." (ASL, 34, 1700d; 2 judgments)

### 1.3 Typology and proposed generalization

We summarize in (11) the typology of iconic enrichments we will propose in this piece.

(11) **Typology of iconic enrichments**

	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>
Speech	John  punished his son.	John punished his son -  .	The talk was loooooong.	Your brother, I am going to  .
Sign	IX-arc-b NEVER  [SPEND MONEY]	IX-arc-b NEVER SPEND MONEY] <sub>b</sub> -  .	POSS-1 GROUP GROW_ 	[currently unclear]
Meaning	<b>cosuppositions</b> (= presuppositions of a special sort)	<b>supplements</b>	<b>at-issue or not</b> , depending on the case	<b>at-issue</b> , with an additional non-at-issue component in some cases

Except for the last line, this table only summarizes the different cases we have discussed so far. The last line foreshadows the empirical claims we will be making, namely that across modalities:

- co-speech/co-sign gestures (including facial expressions) have a non-at-issue semantics, and more specifically that they contribute presuppositions of a particular sort, which we call 'cosuppositions'<sup>6</sup> (we

<sup>6</sup> Briefly, a cosupposition is a presupposition which is conditionalized on the at-issue contribution of the expression it modifies. For instance, *Did John SLAP punish his son?* does not trigger the presupposition that



differ in this from pioneering work on gesture projection in Ebert and Ebert 2014, who took co-speech gestures to contribute supplements; we follow instead the conclusions of Schlenker, to appear b);

- post-speech/post-sign gestures and facial expressions also have a non-at-issue semantics, but a non-presuppositional one: they trigger 'supplements', just as appositive relative clauses do;
- iconic modulations can have any status, and in particular they can make an at-issue contribution;
- pro-speech gestures usually make an at-issue contribution (as most words do), but in addition they may trigger other types of inferences, such as presuppositions.

We will tentatively propose that there is a connection between the form of iconic enrichments and their semantic status, as outlined in (12):<sup>7</sup>

(12) **Proposed generalization**

a.  $\pm$ internal

External enrichments (*-internal*) are not at-issue: because they are external, it should be possible to disregard them without affecting the main, at-issue content of the clause they appear in. By contrast, internal enrichments (*+internal*) can make any semantic contribution – just like standard words.

b.  $\pm$ separate time slot

Enrichments that have a separate time slot (*+separate time slot*) cannot be trivial (= presupposed): because they have their own time slot, they must make a non-trivial contribution to the sentence. By contrast, enrichments that do not have a separate time slot (*-separate time slot*) are not so constrained.

#### 1.4 Elicitation methods

English data were elicited informally, with standard introspective methods. The validity of these methods for spoken language linguistics has been positively assessed with experimental methods, for instance in Sprouse and Almeida, 2012, 2013, Sprouse, Schütze and Almeida, 2013. But it is fair to say that similar results have yet to be established for acceptability judgments on gestures. Still, by way of reassurance, we will summarize recent experimental results that broadly confirm early inferential judgments on co-speech gestures (in Section 3.5). Other results will have to be validated in a similar fashion in the future.

Let us turn to our sign language data, which all pertain to ASL. Data from earlier publications are cited as they initially appeared. New data were elicited using the 'playback method', with repeated quantitative acceptability judgments (1-7, with 7 = best) and repeated inferential (i.e. semantic) judgments (on separate days) on videos involving minimal pairs (see e.g. Schlenker et al. 2013, Schlenker 2014). In a nutshell, the playback method involves two steps. First, the sign language consultant signs sentences of interest on a video, as part of a paradigm (e.g. often with 2 to 6 sentences) signed as minimal pairs. Second, the consultant watches the video, provides quantitative acceptability ratings, and inferential judgments, enters them in a computer, and redundantly signs them on a video. The second step can be repeated on other days, sometimes with a considerable time delay. This method has the advantage of allowing for the precise assessment of minimal pairs (signed on the same video), in a quantitative, replicable way; its obvious limitation is that it solely assesses an individual's idiolect. While the judgments are obtained from just one consultant, the repetition of the task makes it possible to assess the stability of the judgments; and if necessary this method could be turned into an experimental one in the future, assessing the same videos with other signers.

For readability, only average judgments are provided (as superscripts before the sentences), as well as a summary of the relevant aspects of the inferential judgments (complete quantitative judgments are given when there is more than a 2-point difference in the judgments obtained for a given sentence). Unless otherwise noted, sentences that appear in the same numbered example were assessed as part of the same video. Raw data (obtained during elicitation sessions) are provided in the Supplementary

---

John slapped his son, but that *if* he punished his son, slapping was involved. A more detailed definition is given in Section 3.2.

<sup>7</sup> This is only a first approximation. One key issue in the future will be to distinguish optional modifiers from supplemental ones, since there is a sense in which *both* types can be eliminated without yielding syntactic unacceptability. In Schlenker 2010, 2013 it is argued that appositive relative clauses have a rather special syntax, which allows them to be attached to any propositional node that dominates their surface position – with the result that matrix appositives could be removed without changing anything to the syntactic tree to which they are attached. (By contrast, eliminating a standard modifier would remove a node from the expression it modifies.)

Materials, and specialists are invited to consult them when relevant. Notations such as *ASL, 34, 1550 a,e, 5 judgments* indicate that the relevant sentences appeared in ASL video 34, 1550, that only sentences *a* and *e* (i.e. the first and the fifth) from that paradigm are transcribed, and that averages are computed on the basis of 5 judgments (if no letters followed *34, 1550*, this would indicate that the entire paradigm was transcribed). When different inferential judgments were obtained on the same sentence, this is sometimes written with ratios, e.g. '3/5 judgments' referring to '3 judgments out of 5'.

## 2 Iconic modulations

We motivated our enterprise by highlighting the importance of a detailed comparison between the semantics of sign-with- iconicity and speech-with-gestures, as admonished by Goldin-Meadow and Brentari (to appear). One might initially think that 'expressive parity' between the two modalities can be regained when co-speech gestures are employed in spoken language to match the effects of iconic modulations in sign language. As announced at the outset, things are probably not so, for the following reason: iconic modulations can have all sorts of semantic effects, and in particular they may make at-issue contributions. By contrast, in most cases co-speech gestures are not at-issue; as we will argue below, they trigger a particular kind of presupposition, which we call 'cosupposition'. If this conclusion is correct, sign with iconic modulations displays rather different properties from speech with co-speech gestures. (A complicating factor is that the presuppositions triggered by co-speech gestures are often easy to 'locally accommodate' and thus to turn into part of the at-issue component; it is only when this further operation is taken into account that some cases of iconic modulations and co-speech gestures can yield the same effects.) Importantly, our point about expressive differences is not that referential information about the state of the world cannot be conveyed in similar ways in the two modalities, but rather that the information tends to be packaged differently across the at-issue, presupposition and supplemental dimensions.

### 2.1 At-issue iconic modulations

#### 2.1.1 Iconic modulations of spoken words

Let us go back to the simple iconic modulation we introduced in (4), involving *long*. It can modify the at-issue component of the clause, as can be checked if we embed this construction in the antecedent of a conditional. (Esipova 2016a,b argued that contrastive focus can force or facilitate local accommodation of presuppositions. For this reason, we try whenever possible to include examples that do not involve contrastive focus.)

- (13) I am normally rather patient. But if the talk is loooong, I'll leave before the end.  
 ≠> if the talk is long, the speaker will leave before the end  
 => if the talk is very long, the speaker will leave before the end.

Are there other iconic modulations of words? Several have been studied in the literature on ideophones, but these are often considered a distinguished category of words. Okrent 2002 provides the example in (14)a, with a vocal modulation, but unfortunately the iconic effect (= high-pitched for up, low-pitched for down) is redundant with the at-issue content of the modified expressions, hence its effects are hard to tease apart.<sup>8</sup>

- (14) The bird [flew up]\_<high\_pitch> and down\_low\_pitch.

But other types of examples can be constructed. Thus (15)b has a slightly different meaning from (15)a (pronounced in a neutral fashion): the latter leaves entirely open what kind of business is under discussion, whereas the former suggests that it is a serious kind of business (alternatively, it comes with a quotational component that suggests that John has a very deep voice).

---

<sup>8</sup> One could test whether an inference is obtained when a high pitch is applied to *fly away*, as in (i) – with the advantage that in this case the verb does not by itself trigger the inference that that might be independently obtained because of the iconic modulation.

(i) The bird [flew away]\_<high\_pitch>.

- (15) a. John wanted to talk about business.  
 b. John wanted to talk about business\_<low\_pitched>.  
 John wanted to talk about grave business (or: the speaker views business as unappealing/grave)

Similarly, to my ear (16)b suggests that Mary is going solve the problem in a way which is threatening or at least imposing (and it need not suggest that Mary has a low voice).

- (16) a. Mary is going to solve the problem once and for all.  
 b. Mary is going [solve the problem]\_<low\_pitched> once and for all.  
 => Mary is going to do something severe

These inferences are arguably due to a very general 'frequency code', discussed in linguistics by Ohala 1994, according to which lower pitch is associated with larger body size and derivatively imposing effects<sup>9</sup> (as Cross and Woodruff 2008 note, this correlation is at the source of important musical effects, a point also discussed in Schlenker 2017b, to appear e).

Can these inferences be at-issue? We believe that this is sometimes the case: to our ear, the examples in (17) allow for a reading in which *business* is interpreted as *grave business* within the scope of the relevant operators (and without an inference that *if the speaker/addressee talks about business, it has to be grave business*).

- (17) a. During this afternoon's meeting, I won't talk about business\_<low\_pitched>.  
 b. During this afternoon's meeting, I might talk about business\_<low\_pitched>.  
 c. If you talk about business\_<low\_pitched> during this afternoon's meeting, let me know.

But as Rob Pasternak (p.c.) notes, not all cases of vocal iconic modulations are at-issue (nor do our generalizations suggest that they must be). Thus if the speaker pronounces *talking* in (18)a with a very high pitch, or with great pitch excursions, one may take this to mean that Johnny speaks in a very high-pitched voice, or with great pitch excursions. Importantly, the same inference is or can be obtained in (18)b,c, despite the fact that *talk* is now embedded within the antecedent of a conditional and under *never* – which suggests that it has a non-at-issue meaning contribution.

- (18) a. Pay attention – Johnny is about to start talking\_<high\_pitched>.  
 => Johnny has a high-pitched voice  
 b. If Johnny starts talking\_<high\_pitched>, we'll all listen to him.  
 => Johnny has a high-pitched voice  
 c. Unfortunately, Johnny never talks\_<high\_pitched>.  
 => Johnny has a high-pitched voice

Since our claim is not that iconic modulations *must* be at-issue but only that they *can* be, we leave for future research a more complete typology.<sup>10</sup>

### 2.1.2 Iconic modulations of signs

We will now see that iconic modulations of signs can (but need not) make at-issue contributions. We shall briefly consider the case of verbal aspect, of plurals and of pluractionals, each of which has considerable independent interest.

<sup>9</sup> This is a sufficiently important inference that some animals apparently evolved mechanisms – specifically, laryngeal descent – to lower their vocal-tract resonant frequencies so as to exaggerate their perceived size (Fitch and Reby 2001).

<sup>10</sup> It is quite possible (as Pasternak observes) that high pitch can in this case trigger what we call a cosupposition, i.e. a conditionalized presupposition of the form: *if Johnny were to talk, he would talk with a high-pitched voice*. This is in particular suggested by the universal inference that is arguably obtained in (i):

- (i) None of these ten guys will start talking\_<high-pitched>.  
 => for each of these ten guys, if he talks, he does so in a high-pitched voice

□ Verbal aspect

In a series of influential papers, Wilbur (e.g. 2003, 2008) argued that the distinction between telic and atelic verbs is overtly realized in ASL. In Wilbur and Malaia's (2008) words, the observation was that telic verbs "have a sharper ending movement to a stop, presumably reflecting the semantic end-state of the affected argument (...). These end-states [are] overtly marked in ASL by several mechanisms: (1) change of handshape aperture (open/closed or closed/open); (2) change of handshape orientation; and (3) abrupt stop at a location in space or contact with a body part.". Schlenker, to appear argued that this is only part of the puzzle: in fact, telic and atelic verbs can be iconically modulated in such a way that the movement of the sign can be seen to represent the development of the denoted events. As we will now see, these iconic modulations can make an at-issue contribution, in line with our initial generalization.

To start with atelic verbs, the iconic modulations we observed in the ASL example in (4) can make at-issue contributions. (19) and (20) show that iconic enrichments of *GROW* can be interpreted in the scope of *IF* and *NOT* respectively – and do not give rise to presupposition-style (or supplement-style) inferences that 'project' out of the scope of these operators. Importantly, inferential judgments were given for individual sentences, not for discourses in which one form was contrasted with another (to completely rule out the role of contrast, we recorded the sentences on separate videos). This matters because Esipova 2016a,b shows that gestural cosuppositions (and presuppositions more generally) can become at-issue if this is mandated to satisfy certain pragmatic constraints on focus alternatives.

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

- IF POSS-1 GROUP \_\_\_\_\_, JOHN WILL LEAD.  
 a. <sup>7</sup>GROW\_neutral (ASL, 34, 1942; 2 judgments)  
 b. <sup>7</sup>GROW\_large (ASL, 34, 1944; 2 judgments)  
 c. <sup>7</sup>GROW\_small (ASL, 34, 1946; 2 judgments)

'If my group a. (really) grows / b. grows a lot / c. grows a little, John will lead it.'

(20) *Context*: we are discussing the recent history of the speaker's research group.

- LAST-YEAR POSS-1 GROUP NOT \_\_\_\_ .  
 a. <sup>7</sup>GROW\_neutral (ASL, 34, 1954; 2 judgments)  
 b. <sup>7</sup>GROW\_large (ASL, 34, 1956; 2 judgments)  
 c. <sup>7</sup>GROW\_small (ASL, 34, 1958; 2 judgments)

'Last year, my group didn't a. (really) grow / b. didn't grow a lot / c. didn't grow (even) a little.'

With other atelic verbs, the acceptability and meaning of iconically modulated forms varies, but examples with the verb *PLAY-PIANO* gave rise to rich iconic possibilities. Importantly, these enrichments were computed in the scope of the operators *IF* and *NEVER*, and they did not give rise to any presupposition-like (or supplement-like) patterns of projection. This suggests that these enrichments are in fact at-issue.

(21) IF JOHN<sub>a</sub> IX-a

'If John

- a. <sup>7</sup>PLAY-PIANO  
 plays the piano,  
 b. <sup>6.5</sup>PLAY-PIANO\_fast  
 plays often / at a quick speed,  
 c. <sup>7</sup>PLAY-PIANO\_slow  
 plays the piano slowly,  
 d. <sup>7</sup>PLAY-PIANO\_fast\_slow  
 plays the piano quickly and then slows down,  
 e. <sup>7</sup>PLAY-PIANO\_slow\_fast  
 plays the piano slowly and then speeds up,

IX-1 GIVE-a \$5.

I'll give him \$5.' (ASL, 34, 2088; 2 judgments)

=> no inference about the ways (e.g. speed) in which John will play if he does play the piano

(22) JOHN<sub>a</sub> IX-a NEVER

'John never

a. <sup>7</sup>PLAY-PIANO

plays the piano.'

b. <sup>6.5</sup>PLAY-PIANO<sub>fast</sub>

plays the piano quickly.'

c. <sup>7</sup>PLAY-PIANO<sub>slow</sub>

plays the piano slowly.'

d. <sup>7</sup>PLAY-PIANO<sub>fast\_slow</sub>

plays the piano changing tempo from quick to slow.'

e. <sup>7</sup>PLAY-PIANO<sub>slow\_fast</sub>

plays the piano while changing tempo from slow to quick.' (ASL, 34, 2090; 2 judgments)

=> no inferences about the way (e.g. speed) in which John will play the piano, apart from those that follow from the at-issue component (see the Supplementary Materials for details)

Related data can be found with telic constructions, although the extent of iconic modulations seems to be more limited than with atelic constructions. ASL examples involving the verb *HIT* are given in (23)-(24).<sup>11</sup>

(23) IF JOHN<sub>a</sub> IX-a

'If John

a. <sup>7</sup>a-HIT-2

hits you,

b. <sup>6.5</sup>a-HIT-2<sub>fast</sub>

hits you hard/fast,

c. <sup>6</sup>a-HIT-2<sub>slow</sub>

hits you softly/slowly/carefully,

IX-2 WILL 2-EQUAL-a.

the two of you will have done the same thing.' (ASL, 34, 2102; 2 judgments)

(24) JOHN IX-a WILL NEVER

'John will never

a. <sup>7</sup>a-HIT-2

hit you.'

b. <sup>6.5</sup>a-HIT-2<sub>fast</sub>

really hit you.'

c. <sup>6</sup>a-HIT-2<sub>slow</sub>

hit you softly/slowly.' (ASL, 34, 2104; 2 judgments)

We conclude that in ASL atelic and (possibly to a more limited extent) telic verbs can give rise to iconic modulations, ones that can be at-issue and thus take scope under operators such as *IF* and *NEVER*.

□ *Iconic plurals*

In ASL, plurals are optionally realized by repeating the relevant noun. These repetition-based plurals can also give rise to iconic modulations pertaining to the number and arrangement of the denoted objects (see Schlenker and Lamberton 2016 for a far more detailed analysis). In the paradigm in (25), there were either 3 or 4 iterations of the sign, which gave rise to slightly different inferences across the two conditions pertaining to the quantity of trophies referred to by the *if*-clause (in the '4 iterations'

<sup>11</sup> We do not claim that speed modulations of the sign necessarily trigger inferences about the speed of the event; less direct interpretations might be obtained pertaining for instance to the difficulty of accomplishing a particular action.

condition, it wasn't clear that the consequent was predicted to hold if only 3 trophies were present).<sup>12</sup> More strikingly, the arrangement of the signs was varied as well: they appeared as a line or as a triangle depending on the condition, as is illustrated in (26). In the 'triangular' case, the *if*-clause was understood to make specific reference to the arrangement of the corresponding objects, and it was only in case the apartment contained trophies *arranged as a triangle* that the signer committed to giving the addressee 20 additional dollars. Thus the shape-related iconic enrichment appears to take scope within the *if*-clause and thus to make an at-issue contribution<sup>13</sup> (the same conclusion holds for the modest semantic effect of the number of repetitions).

(25) *Context*: The speaker will be renting the addressee's apartment; he knows it contains trophies, but he hasn't seen them.

POSS-2 APT IF HAVE \_\_\_\_\_, IX-1 ADD 20 DOLLARS.

a. <sup>7</sup> TROPHY-rep-3<sub>horizontal</sub>

=> if there at least three or four trophies in a horizontal line, \$20 will be added.

b. <sup>6.7</sup> TROPHY-rep-3<sub>triangle</sub>

=> if there are at least 3 trophies forming a triangle, \$20 will be added.

c. <sup>6.7</sup> TROPHY-rep-≥4-horizontal

=> if there at least three or four or five trophies in a horizontal line, \$20 will be added.

d. <sup>6.5</sup> TROPHY-rep-≥4-triangle

=> if there are at least three or four or five trophies forming a triangle, \$20 will be added.

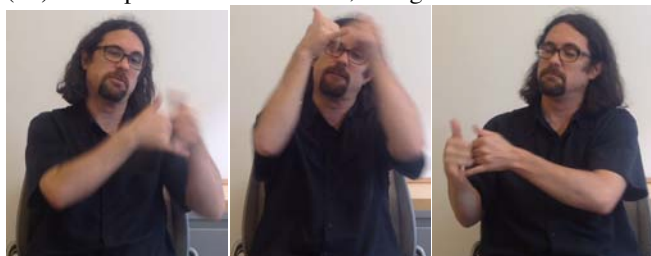
(ASL, 32, 0096 c, d, e, f; 4 judgments)

(26) Plural-based repetition in:

(25)a: 3 unpunctuated iterations, horizontal



(25)b: 3 unpunctuated iterations, triangular



#### □ *Pluractionals*

Kuhn 2015 and Kuhn and Aristodemo 2016 argue that repetition-based pluractionals in LSF (French Sign Language) also give rise to iconic effects, and that these too can make at-issue contributions. The difference in realization between the 'accelerating' and the 'decelerating' versions of LSF *GIVE* are represented in (28).

(27) a. MIRKO CHILD BOOK GIVE-rep-accelerating.

'Mirko gave the child a book at an accelerating pace.' (Kuhn and Aristodemo 2016)

<sup>12</sup> We disregard two further conditions, which involved repeated singular nouns; these were also interpreted in a highly iconic fashion.

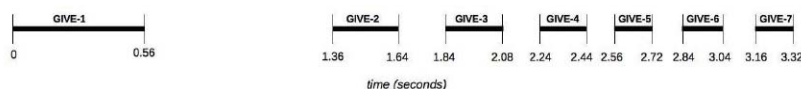
<sup>13</sup> We checked in the 4th judgment task that these sentences do not trigger any inference to the effect that *if* there are trophies, they should be arranged in a particular way; in other words, there is no 'projection' outside of the conditional of the inference pertaining to the arrangement of the relevant objects.

b. MIRKO CHILD BOOK GIVE-*rep*-deceleration.

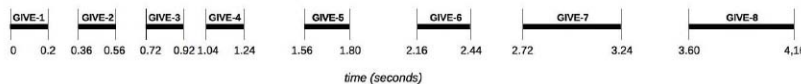
'Mirko gave the child a book at decelerating pace.' (Kuhn and Aristodemo 2016)

(28) Time-course diagrams of accelerating and decelerating *GIVE-*rep** (Kuhn and Aristodemo 2016)

a. Acceleration



b. Deceleration



Importantly for our purposes, Kuhn and Aristodemo 2016 show that this iconic enrichment can be at-issue and thus take scope within a conditional, as seen in (29).

(29) SECRETARY IX-a, IF MIRKO PAPERS IX-b GIVE-*rep*-*accel*, IX-a HAPPY.

'If Mirko gives papers at an accelerating rate, the secretary will be happy.' (LSF, Kuhn and Aristodemo 2016)

Their results extend to ASL - in this case with a slightly different repetition marker, which involves two-handed repetition (encoded with the suffix *-alt*) rather than one-handed repetition (encoded with the suffix *-rep*; both forms exist in ASL and LSF, and there are semantic and distributional differences between them which are discussed in Kuhn and Aristodemo 2016).

(30) IF JOHN PAPERS GIVE-*alt*-speeding-up, SECRETARY WILL HAPPY.

'If John gives papers at an accelerating rate, the secretary will be happy.' (ASL, Kuhn and Aristodemo 2016)

(While Kuhn and Aristodemo's description is suggestive of an at-issue contribution for the iconic component, one would still want to check that no projection phenomena are found in these cases, and in particular that one *fails* to get inferences that are typical of presuppositions or supplements.)

## 2.2 Non-at-issue iconic modulations

Our general claim is that some iconic modulations yield enrichments that may be at-issue, not that all *must* be. The intuition we pursue is that iconic modulations are an integral part of an expression, and thus have all the semantic freedom that lexical meanings do: they may be at-issue, presuppositional, etc. While we have not investigated the full range of inferences triggered by iconic modulations, we should mention that in some cases they make a presuppositional contribution. One such case pertains to 'loci', positions in signing space which often realize discourse referents, and are 'indexed' (i.e. pointed at) to realize pronominal meanings. While loci are normally realized on a horizontal plane (with different discourse referents located on different parts of the plane), high loci may be used when one talks about powerful or important individuals; low loci may be used to talk about short individuals. Importantly, the contribution of these vertical specifications of loci appears to be presuppositional in nature, as is suggested by the fact that the inferences project out of negation, as in the ASL example in (31).<sup>14</sup>

(31) **Vertical locus specifications in ASL** (Schlenker et al. 2013)

YESTERDAY IX-1 SEE R [= body-anchored proper name]. IX-1 NOT UNDERSTAND IX-a<sup>high / normal / low</sup>

'Yesterday I saw R [= body-anchored proper name]. I didn't understand him.' (11, 24)

a. <sup>7</sup>High locus. Inference: R is tall, or powerful/important

b. <sup>7</sup>Normal locus. Inference: nothing special

c. <sup>7</sup>Low locus. Inference: R is short

The ASL examples in (32)-(33) further show that the relevant inference projects out of negation (thus confirming the observations in (31)a) and *if*-clauses – as is expected for a presupposition.

<sup>14</sup> Schlenker et al. 2013 write that they "asked the consultant to provide acceptability judgments in the best context he could imagine for the sentence; this matters because the use of high or low loci in an 'out of the blue' context would normally yield deviance."

- (32) a. <sup>6</sup> POSS-1 YOUNG BROTHER WANT IX-1 REST. IX-1 UNDERSTAND IX-a<sup>high</sup>  
*Inference:* the speaker's younger brother is tall.  
 'My younger brother wants me to rest. I understand him.' (10, 41; 44)  
 b. <sup>6</sup> POSS-1 YOUNG BROTHER WANT IX-1 REST. IX-1 NOT UNDERSTAND IX-a<sup>high</sup>  
*Inference:* the speaker's younger brother is tall.  
 'My younger brother wants me to rest. I don't understand him.' (ASL; 10, 42; Schlenker et al. 2013)
- (33) <sup>7</sup> IX-1 LIKE POSS-2 BROTHER. IF IX-a<sup>high</sup> BECOME BASKETBALL PLAYER, IX-1 HAPPY.  
*Inference:* the speaker's younger brother is tall.  
 'I like your brother. If he becomes a basketball player, I'll be happy.' (ASL, 10, 60; Schlenker et al. 2013)

None of this should come as a surprise: as argued in Schlenker et al. 2013, vertical specifications of loci play very much the role of gender specifications of English pronouns, and certainly these trigger presupposition-like inferences, as illustrated in (34).<sup>15</sup>

- (34) a. I don't understand her.  
 => the third person under discussion is female  
 b. If she comes, I'll be happy.  
 => the third person under discussion is female

The upshot is that although ASL pronouns pointing upwards can be seen as iconic modulations of a sign (like the modifications of *GROW* discussed in Section 2.1.2), this iconic enrichment makes a presuppositional rather than an at-issue contribution. This might well be because the enrichment targets a pronominal expression: compositionally, the enrichment can prevent the pronoun from referring, but not add to it an at-issue, propositional component. This can be seen in the standard analysis of *phi*-features on pronouns. Evaluated under an assignment function  $s$ , a pronoun  $pro^F_i$  with index  $i$  and features  $F$  (for instance  $she_i$ , analyzed for simplicity as  $pro^{feminine}_i$ ) may either denote  $s(i)$ , or yield a failure if  $s(i)$  doesn't satisfy  $F$  (in our example, if  $s(i)$  fails to be female). In the first case, the feature makes no contribution at all; in the second case, it yields a presupposition failure; in neither case does it make an at-issue contribution. Be that as it may, it is clear that we certainly do not wish to claim that all iconic modulations give rise to at-issue contributions, only that they *may* do so; the case of high loci suggests that, in appropriate situations, they may give rise to presuppositional contributions as well.

### 2.3 The connection with ideophones

The sign language plurals and pluractionals discussed in Section 2.1.2 are rather different from the iconic modulation of *long* discussed in Section 2.1.1 in that they involve a *grammatical* marker, which both pluralizes a construction and enriches it with an iconic component. There are reasons to think that related constructions can be found in spoken language with ideophones, as was proposed by Henderson 2016 (for recent discussion of ideophones, see Dingemanse 2013, Dingemanse and Akita 2016, Dingemanse et al. 2016).

As an example, let us focus on what Henderson calls 'total reduplication' in the Mayan language Tzeltal, which targets ideophones, which he defines as a distinguished lexical category of expressions that have an iconic component (and fall under special rules, in particular on the phonological side). Importantly for our purposes, in 'total reduplication' in Tzeltal, "speakers can make multiple demonstrations using the same ideophone stem to demonstrate a plurality of events". In such cases, then, the manner of reduplication "iconically reproduces the temporal properties of the event-plurality". As a result, if we notate the ideophone as *IDF*, "*IDF IDF IDF* demonstrates events with a different temporal character than *IDF [pause] IDF [pause] IDF*". These facts are of course reminiscent of the sign language iconic plurals and pluractionals discussed in Section 2.2 – and Henderson does stress the connection with sign language pluractionals.

From the present perspective, a key question is whether ideophones have an at-issue, presuppositional or supplemental status; we do not know of literature that has performed the required tests in a systematic fashion (see Dingemanse et al. 2016, to appear for experimental work on the iconic

<sup>15</sup> Schlenker et al. 2013 provide preliminary data from ASL that suggest that height specifications of loci also resemble gender features of pronouns in that they trigger presuppositions that are indexical, i.e. evaluated with respect to the context of utterance rather than to the world of evaluation. We disregard this more subtle point here.



properties of ideophones – but one which does not address the issue of the at-issue or not-at-issue behavior).

#### 2.4 Towards a formal analysis of elementary iconic enrichments

Our focus in this piece is on the projection of iconic enrichments in complex sentences, not on how they are generated in elementary clauses. Still, a word should be added to explain how formal approaches can incorporate iconic conditions. The basis idea, developed in much greater detail in Greenberg's 'pictorial semantics' (e.g. Greenberg 2013), is that an iconic representation is true of an object just in case this object has certain structural properties of the representation. For this reason, iconic representations enrich Logical Forms by specifying that the objects that an expression is true of should also 'preserve' certain structural properties of that expression. In the case of visual representations, these are usually geometric properties. A formal treatment of the verb *GROW* discussed in (7) is developed in Schlenker, to appear d. Suffice it to say that a requirement is stated to the effect that the larger the realization of *GROW*, the greater the growth; and similarly for speed. Analogous principles could be applied to iconic modulations of *long* (the longer the word, the greater the length of the relevant object or event).





This part of the analysis would be essential for a complete semantic treatment of iconic enrichments; but for the narrower issue of the *projection* of these enrichments, we can take as given the contribution of iconic enrichments to elementary clauses, and seek to explain how they are inherited by complex sentences.

### 3 Co-speech gestures: English

Unlike iconic modulations, which are an integral part of the expressions they modify, co-speech gestures are external enrichments, and they do not have their own time slot since they are produced simultaneously with the expression they modify. We argue (following Schlenker, to appear b) that they do not normally make an at-issue contribution, and trigger a particular kind of presupposition, which we call 'cosupposition'. (We will argue in the next section that some instances of *co-sign gestures* can be found in sign language, and that they pattern with co-speech gestures.)

#### 3.1 Supplements or presuppositions?

Ebert and Ebert 2014 suggested that co-speech gestures should be analyzed as supplements. As we will see in Section 6, post-speech/sign gestures (not discussed by Ebert and Ebert) do arguably display a supplemental behavior, but in our view this makes it all the more striking that co-speech gestures do not. The point is made in Schlenker 2015 by focusing on negative environments in which supplements are degraded, but in which co-speech gestures appear rather freely. This is shown by the acceptability contrast between (35)a and (36)a (co-speech gestures) and (35)b and (36)b (appositives) (we include with the co-speech cases the intuitive inferences that are characteristic of cosuppositions, discussed below; inferences might be a bit weak in this case, hence our use of =>?). As we will see shortly, presupposition triggers also freely appear in these environments.

- (35) a. It's unlikely that the next speaker will bring LARGE\_  [a bottle of beer] to his talk.  
=>? if the next speaker brings a bottle of beer, it will be a large one
- b. #It's unlikely that the next speaker will bring a bottle of beer, which is LARGE\_  **this** large.  
(Schlenker 2015)
- (36) a. No philosopher brought LARGE\_  [a bottle of beer] to the workshop.  
=>? when a philosopher brings a bottle of beer, it is usually a large one
- b. #No philosopher brought a bottle of beer, which is LARGE\_  **this** large.  
(Schlenker 2015)


As mentioned in Schlenker, to appear b, a supplemental approach could deal with (35)a-(36)a by taking the gestures to behave like the appositives in (35)b-(36)b, but with *which would be* replacing *which is*. On the assumption that the resulting sentence is more acceptable, one would still need to ask why such an option should be available – and importantly why it *fails* to be available in the case of post-speech gestures, a point to which we will turn in Section 6.

### 3.2 Cosuppositions<sup>16</sup>

Following Schlenker, to appear b, we argue that a broad class of co-speech gestures should be analyzed within a presuppositional framework, albeit with a twist. In standard theories, a presupposition trigger  $pp'$  (e.g. *it stopped raining*) with presupposition  $p$  (*it rained*) and at-issue component  $p'$  (*it doesn't now rain*) comes with a requirement that  $p$  should be entailed by the local context of  $pp'$ . By contrast, an expression  $p$  co-occurring with a co-speech gesture  $G$  triggers a presupposition that  $p$  entails  $g$ . In other words, the co-speech gesture triggers an assertion-dependent presupposition, something we call a 'cosupposition'. To illustrate, the co-speech gesture *UP* in (37) definitely does not trigger the simple presupposition that some lifting will in fact take place, but rather the conditionalized presupposition or 'cosupposition' that if John helps his son, lifting will be involved.

(37) John will *UP*  **help** his son?

Why such a conditionalization? Schlenker, to appear b discusses several possible motivations. One goes like this: the context should guarantee that the co-speech gesture merely illustrates the expression it modifies, and thus that relative to that context the expression entails the content of the co-speech gesture. In unembedded cases, such as (37), one can posit that the relevant inference must follow from the context of the conversation. But in embedded cases a more sophisticated notion is needed, that of a *local context*. To see why it is needed, consider the example in (38).

(38) If little Johnny takes part in the competition, will his mother *UP*  **help** him?  
=> if little Johnny takes part in the competition, if his mother helps him, lifting will be involved

The requirement that the content of the gesture follow from the meaning of *help* relative to the global context would give rise to an overly strong inference, namely that *in general*, helping is understood to mean 'helping by lifting'. The inference which is in fact derived is narrower: besides the fact that it only applies to Johnny and his mother, it is relative to the hypothesis that Johnny will take part in a competition; and we thus infer that this kind of competition involves some kind of upward movement.

In modern theoretical parlance, the entailment need not hold with respect to the *global* context of the sentence, but only with respect to the *local* context obtained by 'updating' the global one with the antecedent of the conditional. The notion of a 'local context' is standardly used to motivate dynamic approaches to presupposition projection in the tradition of Heim 1983. As a first approximation, the local context of an expression recapitulates the semantic content already contributed by expressions that precede it, combined with the context of the conversation. In various theories of presupposition (e.g. Heim 1983, Beaver 2001, Schlenker 2009), an expression  $pp'$  triggering a presupposition  $p$  in a sentence  $S$  uttered relative to a Context Set  $C$  is acceptable only if  $p$  follows from the local context of  $pp'$  in  $S$  given  $C$ . This holds in particular if  $p$  is a cosupposition, of the form  $p = (p' \Rightarrow q)$ , with conditionalization on the at-issue component  $p'$ . This means that  $x$  *UP help*  $y$  can be analyzed as a standard presupposition trigger, which happens to trigger a presupposition of the rough form: *if x helps y, lifting is involved*.

Thus the cosuppositional analysis can be combined with standard theories of presupposition projection to account for complex patterns of gesture projection. Without getting into formal details that are developed in Schlenker, to appear b, we will now survey the main projection phenomena that



<sup>16</sup> This section, as well as the following two sections, borrow from Schlenker, to appear b.

are found, and we will emphasize their similarity with standard presuppositions by including whenever possible presuppositional controls.<sup>17</sup>



### 3.3 Cosuppositional behavior: propositional cases



In (39)a(i), a lifting gesture accompanies the verb *helped*, and yields roughly the same informational contribution as the example in (39)a(ii), in which an at-issue modifier is used, referring deictically to the same kind of gesture. The latter example is intended as a control in which the contribution of the co-speech gesture in (39)a(i) is made by a well-understood construction<sup>18</sup> (we could have used just as well the modifier *by lifting him*, but this would have made for a less minimal control). And of course the presuppositional control in (39)b gives rise to the factive inference that one expects given the presence of *realize*. We purposefully focus on a relatively weak presupposition trigger because, as mentioned above, we think that various co-speech gestures easily give rise to patterns of local accommodation.

*Notation:* When pairs of sentences (i) and (ii) are contrasted, we write (i), (ii) => ... if both sentence (i) and sentence (ii) give rise to inference ...; we write (i), (ii) => ... if sentence (i) but not sentence (ii) gives rise to inference ... .

- (39) a. John will (i)  **help** his son (ii) help his son like  **this**.  
 (i), (ii) => John will help his son by lifting him
- b. John will realize that his son is losing.  
 => John's son is/will be losing

Now the important observation is that when (39)a(i) and (39)a(ii) are embedded under logical operators, they display sharply different behaviors – and the inferences triggered by co-speech gestures pattern with the presupposition illustrated in (39)b. In questions and under negation, the relevant conditionalized presuppositions (i.e. the relevant cosuppositions) are 'projected' and thus inherited by the matrix clause, as shown in (40)-(41) (the latter is just a more controlled version of the initial example in (37)). The presupposition triggered by *realize* also projects, but of course it is not conditional in nature. We also include controls with *like this*, which *fail* to trigger the relevant inferences.<sup>19</sup>

- (40) a. John won't (i)  **help** his son. (ii) help his son like  **this**.  
 (i) => If John helped his son, he would do so by lifting him
- b. John won't realize that his son is losing.  
 => John's son is/will be losing

- (41) a. Will John (i)  **help** his son? (ii) help his son like  **this**?  
 (i) => If John helps his son, he will do so by lifting him


<sup>17</sup> One cautionary note might be helpful at the outset, however. Presuppositions can, at some cost, undergo 'local accommodation', with the result that their content is treated as an at-issue meaning component (e.g. Heim 1983). We come back to this issue in Section 5.3.

<sup>18</sup> Since *this* is a demonstrative, it denotes something which is made salient by the context – here the action evoked by the accompanying gesture. The semantics of *this* need not involve explicit gestures, however, and thus it must be handled by a theory of context dependency that does not reduce to gesture semantics.

<sup>19</sup> In some embedded environments, a sentence such as *John won't help his son like UP this* could trigger an implicature because it has a structurally simpler alternative *John won't help his son*. Since the latter is more informative, scalar reasoning yields the inference that it is false, hence an implicature that John will help his son. For simplicity, we disregard implicatures in the present discussion.

- b. Will John realize that his son is losing?  
=> John's son is/will be losing

To study gesture projection in upward-monotonic environments, it is sometimes useful to embed some of the examples in questions so as to distinguish the presuppositional component (which projects, i.e. is inherited by the question) from the at-issue component (which doesn't project). Thus (42) displays a more controlled version of our initial example in (38).

- (42) a. If little Johnny takes part in the competition, will (i) his mother UP\_  **help** him? (ii) his mother help

him like UP\_  **this**?



- (i) => if little Johnny takes part in the competition, if his mother helps him, lifting will be involved

b. If little Johnny cheats on the exam, will his mother realize that he committed an extremely serious offense?

- => if little Johnny cheats on the exam, he will have committed an extremely serious offense

As announced, (42)a,b show that in the consequent of a conditional, normal presuppositions and cosuppositions interact in non-trivial ways with the content of antecedent: it is neither the case that the presupposition 'projects' nor that it 'fails to project' in its original form. Rather, it projects *in modified form*, relativized to the antecedent of the conditional. This shows in particular that it won't do to say that co-speech gestures just trigger an inference to the effect that they illustrate the lexical meaning of the expression they modify. Rather, just as for normal presuppositions, the linguistic context must be taken into account, as is guaranteed by standard theories of presupposition.

It is worth noting that the acceptability of co-speech gestures under *unlikely*, as in (35), is unsurprising from a cosuppositional perspective. And it gives rise to the inferences that one would predict on the basis of standard presupposition triggers, as shown in (43): the presupposition/cosupposition is inherited by the matrix clause.

- (43) a. It is unlikely that John will (i) UP\_  **help** his son. (ii) help his son like UP\_  **this**.

- (i) => If John helps his son, he will do so by lifting him

b. It is unlikely that John will realize that his son is losing

- (i) => John's son is losing

Many other propositional examples are discussed in Schlenker, to appear b.

### 3.4 Cosuppositional behavior: quantified cases

Let us turn to some quantified examples (we henceforth omit pictures from the transcription of the gestures, unless these have not been exemplified before). Under universal quantifiers, we obtain universal projection of a presupposition/cosupposition, as seen in (44)

- (44) a. Did each of these 10 guys (i) UP **help** his son? (ii) help his son like UP **this**?

- (i) => for each of these 10 guys, if he helped his son, he did so by lifting him

b. Did each of these 10 guys realize that his son needed help?

- (i) => for each of these 10 guys, his son needed help.

Embedding under *none*-type quantifiers is particularly informative, for two reasons. First, presuppositions in their scope are believed to give rise to (positive) universal inferences, as illustrated in (45)b; this a characteristic behavior of presuppositions, established by Chemla 2009 with experimental data. Second, unlike presupposition triggers, supplements in the scope of *none*-type quantifiers seem to be degraded, as we saw in (36). As seen in (45), co-speech gestures pattern with presupposition triggers and yield universal cosuppositional inferences, as one would expect.

- (45) a. None of these 10 guys (i) UP **helped** his son. (ii) helped his son like UP **this**.

- (i) => none of these 10 guys helped his son; but for each of them, if he had helped his son, it would have

been by lifting him

b. None of these 10 guys realized that his son needed help.

=> for each of these 10 guys, his son needed help

Finally, going back to our initial examples in (35)a and (36)a, we can account for the inferences they trigger by postulating that *John brings LARGE [a bottle]* triggers the cosupposition that *if John brings a bottle, it will be a large one*.


### 3.5 Experimental approach

Up to this point, we have based our discussion on introspective judgments on the gesture semantics. But are these judgments reliable? Tieu et al. 2017, to appear develop an experimental approach designed to test exactly this. Their clearest results pertain to an inferential task (Tieu et al. 2017), performed in two separate experiments illustrated in (46) and (47) on the example of the quantifier *none*, with the types of inferential questions in (48)

#### (46) Experiment 1: Target Gestures

None of these three girls will  [use the stairs].

#### (47) Experiment 2: At-issue controls

None of these three girls will use the stairs in  [this direction].

#### (48) Inferential questions (for both experiments)

a. Existential inference

To what degree does this video suggest the inference below?

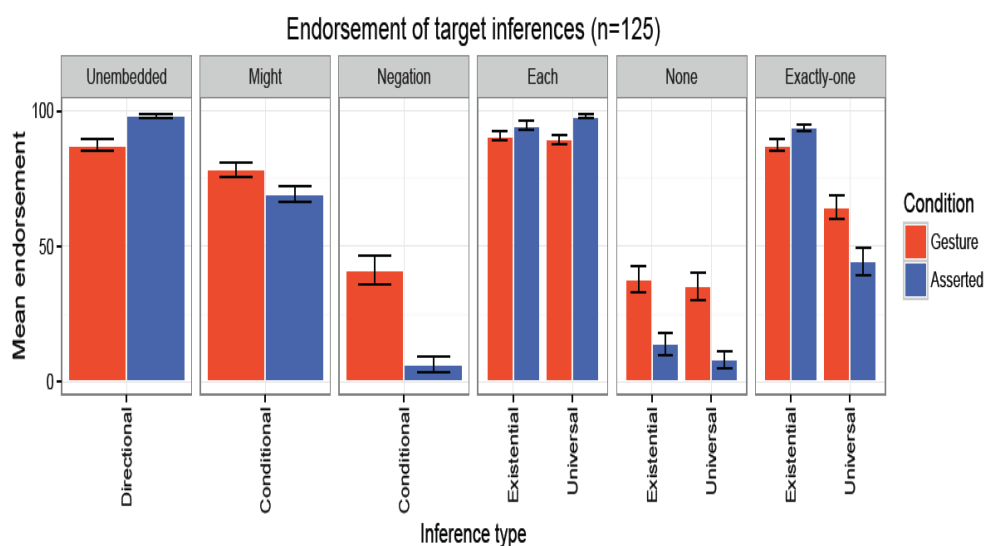
For at least one of these three girls, if she were to use the stairs, she would go up the stairs.

b. Universal inference

To what degree does this video suggest the inference below?

For each of these three girls, if she were to use the stairs, she would go up the stairs.

The expected cosuppositional inferences were tested in a variety of environments involving *might*, *not*, *each*, *none* and *exactly one*. Quantitative results are summarized in the graphs in (49); the results of the target experiment, involving co-speech gestures, are contrasted with those of the at-issue environments used in the control experiments. Quantified cases were tested both for universal inferences as (48)b and for existential inferences as in (48)a, in part because the data are not entirely clear, and in part because there is a debate between proponents of 'existential projection' vs. 'universal projection' in the presupposition literature (e.g. Beaver 2001).

(49) **Endorsement rates for different linguistic environments**

Without getting in too much detail, several points are worth noting.

(i) First, there are clear and significant differences between target gestural sentences and at-issue controls, which goes in the direction of the cosuppositional theory. Still, cosuppositional inferences are often far less strong than expected. As mentioned the outset, this might justify treating gestural triggers as very weak triggers, which easily give rise to local accommodation – a point to which we return in Section 5.3.

(ii) Second, in these inferential data there is some evidence of universal projection under *none* and under *exactly one*: despite the relative weakness of this projection behavior, it is significantly different from what is obtained with at-issue controls.

(iii) Third, the apparent availability of local accommodation turns out to be important in the detailed debate between the cosuppositional and the supplemental theory of co-speech gestures, as relevant controls with appositives do not seem to give rise to at-issue readings in the cases at hand. Thus Tieu et al. 2017, to appear argue that readings with local accommodation offer an important argument against supplemental analyses – a point we do not further develop in this piece.

### 3.6 *Cosuppositions triggered by facial expressions*

As mentioned in Schlenker, to appear b, the same generalizations arguably hold for some co-speech facial expressions. Here :-/ stands for an unhappy or disgusted face, illustrated in (50)a. Importantly, the disgust might be attributed to the speaker or to Sam in (50)a; but what matters for present purposes is that whatever inference is obtained in the unembedded clause is inherited by the complex sentences in the paradigm.

#### (50) **Propositional examples**



a. Sam went :-/ [skiing with his parents].

=> for Sam to go skiing with Sam's parents would be disgusting (from Sam's / from the speaker's standpoint)

b. Sam won't go :-/ [skiing with his parents].

=> for Sam to go skiing with Sam's parents would be disgusting (from Sam's / from the speaker's standpoint)

c. Did Sam go :-/ [skiing with his parents]?

=> for Sam to go skiing with Sam's parents would be disgusting (from Sam's / from the speaker's standpoint)

d. If Sam goes :-/ [skiing with his parents], I'll hear about it.

=> for Sam to go skiing with Sam's parents would be disgusting (from Sam's / from the speaker's standpoint)

We can also construct examples in which the disgusted facial expression appears in the consequent of a conditional, and the disgust can be relativized to the content of the antecedent, as shown in (51).

- (51) Sam loves skiing, but if the snow is as terrible as it is today, he won't go :- ( **skiing** tomorrow  
=> if the snow is as terrible as it is today, skiing will be disgusting tomorrow (from Sam's / from the speaker's standpoint)

In more complex examples, we can see the effect of a facial modifier co-occurring with an expression that contains a bound variable. The same generalizations hold as in our earlier examples involving manual gestures: in (52)a,b, an inference is triggered to the effect that for the speaker's friends to go skiing with their parents is somehow disgusting. Finally, we can test what happens when the facial expression spans the entire sentence, as in (52)c. While the facts are not entirely clear, our impression is that this can trigger the inference that it is disgusting *that* none of the relevant friends go skiing with his parents (we don't exclude that a reading is available to the effect that skiing with one's parents is disgusting). This possibility is predicted by the cosuppositional analysis: when the facial expression spans an entire proposition, we predict that the proposition should entail the content of the expression (hence disgust in the case at hand). As we will see, similar facts can be replicated with facial expressions in ASL.

(52) **Quantified examples**

a. Does each of your friends go :- ( **[skiing with his parents]**)?

=> for each of my friends, to go skiing with his/her parents would be disgusting (from the friend's / from the speaker's standpoint)

b. None of my friends goes :- ( **[skiing with his parents]**).

=> for each of my friends, to go skiing with his/her parents would be disgusting (from the friend's / from the speaker's standpoint)

c. :- ( **[None of my friends goes skiing with his parents]**).

#### 4 Co-sign gestures: ASL

Our initial evidence for cosuppositions comes from co-speech gestures in spoken language. It is non-trivial to extend the findings to sign language: since gestures are produced in the same modality as signs, teasing apart the difference between the two requires more sophisticated criteria than in spoken language. We will nonetheless propose that some (non-grammatical) facial expressions display the semantic hallmarks of co-speech gestures.<sup>20</sup>

While we will argue that some such expressions are indeed comparable to co-speech gestures, it must be noted at the outset that in *some* examples facial expressions are naturally interpreted as iconic modulations of a sign. Thus in (53)b the dominant (but not quite the sole) reading seems to be that John never plays the piano badly or with a bad attitude: the disgusted face appears to be preferentially integrated to the sign for *play the piano*, and it thus behaves like an internal enrichment.

(53) *Context*: John is an amateur musician. The speaker enjoys good music.

a. <sup>7</sup> JOHN IX-a NEVER PLAY-PIANO.

'John never plays the piano.'



b. <sup>6,4</sup> JOHN IX-a NEVER :- ( **[PLAY-PIANO]**).

<sup>20</sup> In principle, one could profitably investigate the role of vocal enrichments in sign language (in hearing native signers). These would genuinely be the analogue of co-speech gestures in spoken language. We leave this question for future research.

'John never plays the piano badly.'

*Meaning 1:* John never plays the piano badly or with a negative attitude, hence if he plays, he plays well/seriously (5/5 judgments, with 2 mentioning an ambiguity)

*Meaning 2:* John never plays the piano, but if were to play, the signer would hate it (2/5 judgments mentioning this as a possible meaning in addition to Meaning 1).

'John never plays the piano (badly)'

(ASL, 34, 1550 a.e, 5 judgments)

An alternative reading can to some extent be brought out in a minimally different example in which the signer bends backwards as he realizes the sign (something we write as \), possibly because in this case the facial expression can be seen as an external enrichment of the sign (see the Supplementary Materials for the complex judgments). The result is still ambiguous, but Meaning 2 is now easier to access:

(54) *Context:* John is an amateur musician. The speaker enjoys good music.



<sup>6,6</sup> JOHN IX-a NEVER \ :- ( [PLAY-PIANO].

*Meaning 1:* John never plays the piano badly, hence if he plays, he plays well/seriously (3/5 judgments mentioning this as a possible meaning, including 1 as the only possible meaning and 2 as a permissible (possibly preferred) meaning)

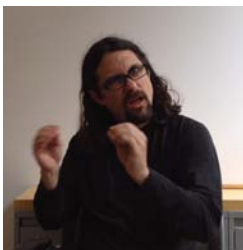
*Meaning 2:* John never plays the piano, but if he were to play, the speaker would hate it (4/5 judgments mentioning this as a possible meaning, including 2 as the only possible meaning and 2 as a permissible (possibly dispreferred) meaning).

(ASL, 34, 1546e; 5 judgments)

Still, neither of the foregoing examples provides clear evidence of cosuppositional inferences. Things are different in (55): even though the basic meaning of the facial expression is open to interpretation (in terms of disgust or, more marginally, in terms of difficulty), it gives rise to a universal inference out of the scope of the quantifier *NONE*.

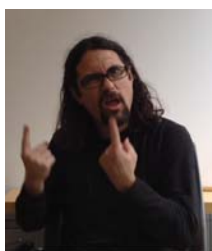
(55) AMERICA [RICH STATE-rep HELP PEOPLE]<sub>b</sub>.

a. <sup>6,7</sup> [POOR STATE-rep NONE IX-arc-a SPEND MONEY]<sub>a</sub>.



b. <sup>6</sup> [POOR STATE-rep :- ( [NONE IX-arc-a SPEND MONEY]]<sub>a</sub>.

=> it is good for poor and rich states to spend money



c. <sup>6</sup> [POOR STATE-rep NONE IX-arc-a :- ( [SPEND MONEY]]<sub>a</sub>

=> it is bad (3/4 judgments) or difficult (1/4 judgment) for poor states to spend money

(in one case, it is taken to be bad to for states in general to spend money; see the detailed data in the



Supplementary Materials)

'In the US, rich states help people. But no poor states spend money.' (ASL, 34, 1670a,c,d; 4 judgments)

More precisely, we obtain interpretations that suggest that a disgusted facial expression in the scope of the quantifier *NONE* gives rise to projection effects. To see this, let us consider the full paradigm. When the disgusted facial expression spans the entire clause, as in (55)b, what is taken to be disgusting is the fact that no poor state spends money. But when the facial expression spans the Verb Phrase only, one gets the inference that it would be disgusting (or difficult) for poor states to spend money (and sometimes this gets strengthened to an inference that also applies to rich states as well). This contrast replicates our impression of the English data discussed in (52)b-c.<sup>21</sup>

A related case of projection can be seen in (56), which has about the same truth conditions as (55), but with a very different structure, involving conditionals and adverbs rather than quantifiers.

- (56) AMERICA [IF STATE RICH HELP PEOPLE]<sub>b</sub>.  
a.<sup>7</sup> [IF STATE POOR NEVER SPEND MONEY]<sub>a</sub>.



- b.<sup>5,7</sup> -( [IF STATE POOR NEVER SPEND MONEY]<sub>a</sub>.

=> it is good for poor and rich states to spend money



- c.<sup>6,3</sup> [IF STATE POOR NEVER :- ( [SPEND MONEY]<sub>a</sub>.

=> it is bad (1/3 judgment) or difficult (2/3 judgments) for poor states to spend money

'In the US, if a state is rich, it helps people. If a state is poor, it never spends money.' (ASL, 34, 1712a,c,d; 3 judgments)

Unsurprisingly, we obtain in (56)b a negative assessment of the proposition that if a state is poor it never spends money – in other words: it would be *good* for poor states to spend money. World knowledge is probably responsible for the further inference that it is good for states *in general* to spend money: if it is good for poor states to spend money, *a fortiori* this should be the case for rich states (all the more so in the context of helping people). The interesting observation is that when the disgusted face only appears over *SPEND MONEY*, as in (56)c, the inference triggered by the facial expression 'projects' out of the scope of *NEVER*, since it is now inferred that it is difficult or bad for poor states to spend money (the ambiguity between 'difficult' vs. 'bad' makes things harder to test, but we can nonetheless ascertain that the inference projects as a cosupposition).

The observed inference can be explained if *SPEND MONEY* comes with a cosupposition of the form:  $\lambda x. x \text{ spends money} \Rightarrow x \text{ does something negative/difficult}$ . The environment is a bit more complex than those we discussed up to this point. The relevant Logical Form involves generalized quantification over situations, akin to: *No situation in which a state is poor is one in which it spends*

<sup>21</sup> As noted in (55)c, on occasion the inference is not just that it is bad/difficult for *poor* states to spend money, but for states in general. This might be part of a more general issue in the study of presupposition, called the Proviso Problem: in many cases, presuppositions that are predicted by dynamic semantics to be conditional in nature get strengthened (here, we predict an inference of the form: *if a state is poor, it is bad/difficult for that state to spend money*; this is strengthened to: *it is bad/difficult for states to spend money*). See Schlenker to appear b for further discussion, and see Beaver and Geurts 2014 and Schlenker 2016a for surveys of presupposition theory that discuss the Proviso Problem.

*money* (see for instance Lewis 1975 for a – by now standard – treatment of adverbial quantification in terms of generalized quantification). This is thus the same type of structure we investigated in (45)a(ii) and (52)b, repeated as (57)a,b.

(57) a. None of these ten guys UP [**helped his son**].

=> for each of these ten guys, if he had helped his son, he would have lifted him

b. None of my friends goes :-([**skiing with his parents**]).

=> for each of my friends, skiing with his parents wouldn't be fun

For these examples, we posited that the verbal predicate triggers a cosupposition of the form  $\lambda x. x \text{ helps } x\text{'s son} \Rightarrow x \text{ lifts } x\text{'s son}$  for (57)a, and  $\lambda x. x \text{ goes skiing with his parents} \Rightarrow x \text{ does something disgusting}$  for (57)b. We then made use of the fact that in structures of the form *No NP VP*, if a presupposition is triggered by the VP, it should be satisfied by every object that satisfies the NP – hence the inferences we obtained in (57)a,b. The case of quantification over situations is no different, as can be seen with gestural and non-gestural examples alike in (58): each situation that satisfies the *if*-clause should satisfy the presupposition of the main clause.

(58) a. If a kid goes on winter vacations with his parents, he never notices that skiing with them is ridiculous.

=> it is ridiculous for kids on winter vacations to ski with their parents

b. If a kid goes on winter vacations with his parents, he never goes :-([**skiing with his parents**]).

=> it is disgusting for kids on winter vacations to go skiing with their parents

Combining these results, we can understand why (56)c triggers the inference that if a state is poor, it is bad for it to spend money.

Finally, and rather unsurprisingly, the inferences seen in (55) and (56) can be replicated with co-speech facial expressions in spoken language, as seen in (59).<sup>22</sup>

(59) a. No poor state should :-([**spend money**]).

=> it is bad for poor states to spend money

b. If a state is poor, it should never :-([**spend money**]).

=> it is bad for poor states to spend money

Similarly, a disgusted facial expression co-occurring with the English VP *play the piano* gives rise to a cosuppositional reading, as seen in (60). Unlike what we saw in (53) and (54), there is no risk that the facial expression can be viewed as an iconic modulation of the sign: it is clearly analyzed as an external enrichment, and we obtain a relatively clear cosuppositional reading.

(60) *Context*: John is an amateur musician. The speaker enjoys good music.

John never :-([**plays the piano**]).

=> the speaker doesn't like John's piano playing

## 5 Co-speech and co-sign gestures: further issues

### 5.1 Deriving cosuppositions

A question left open by our empirical investigations is why co-speech/co-sign gestures should trigger cosuppositions. Three theoretical directions should be briefly mentioned, although we will not seek to adjudicate among them.

1. Within standard presupposition theory (e.g. Heim 1983), one could simply take the presuppositional behavior of co-speech gestures to be a lexical fact, encoded by way of entries such as (61), defined for the case in which *SLAP* co-occurs with a transitive verb (we assume a simple trivalent framework; adaptations would be needed for a full dynamic framework *à la* Heim 1983). The crucial part is

<sup>22</sup> The inference might be strengthened to 'it is bad for states in general to spend money', or even: 'spending money is bad'. As noted in fn. 21, this falls under the 'Proviso Problem', i.e. the observation that presuppositions are sometimes stronger than is predicted by standard dynamic analyses.

boldfaced: a presupposition failure is triggered if the argument modified by *SLAP* fails to entail the semantic content *slap'* of *SLAP*.

(61) **Trivalent entry for a co-speech gesture: *SLAP***

with the semantic contribution of *SLAP* written as: *slap'*

$[[SLAP]]^{f,w,s} = \lambda P_{\langle e, \langle e, st \rangle \rangle} . \lambda y_e . \lambda x_e . \# \text{ iff } P_w(y)(x) = \# \text{ or } \textit{slap}'_w(y)(x) = \# \text{ or } \mathbf{not: } P_w(y)(x) = 1 \Rightarrow$

$\textit{slap}'_w(y)(x) = 1; 1 \text{ iff } P_w(y)(x) = \text{ or } P_w(y)(x) = 1$

Besides the need to adapt the analysis to a full-fledged dynamic framework, one would have to state more general rules for cases in which *SLAP* modifies verbs that are not transitive.

2. One could state more directly a rule intended for dynamic analyses of presuppositions. A key insight of this framework is that *the presupposition of an expression must be entailed by the local context of that expression*. In classic statements of dynamic semantics, such as Heim 1983, the value of local contexts is computed by a recursive system: the meaning of expressions is taken to be evaluated with respect to a local context and to output a modified local context. In pragmatic reconstructions of local contexts such as Schlenker 2009, local contexts are computed by a pragmatic algorithm that starts from the syntax and the bivalent (i.e. classical) semantics of any sentence. In either framework, then, there are more contexts than meets the eye: besides the general context of the conversation, the local context of an expression may play a role in its evaluation. The key for an analysis of *SLAP* co-occurring with a transitive verb *V* is to require that *relative to its local context V* should entail the content of *SLAP* (this achieved by way of recursive rules in dynamic semantics, and post-semantically in frameworks with a pragmatic reconstruction of local contexts).

3. A different direction is explored in Schlenker 2016b, which goes back to a pragmatic analysis of presupposition called 'the Transparency Theory' (Schlenker 2008). Schlenker 2008 argues that the presupposition *d* of a (predicative/propositional) trigger *dd'* is a normal entailment that 'wants' to be articulated as a separate conjunct, a principle called 'Be Articulate'. To illustrate with the presupposition trigger *know*, one should say ... *it's raining and John knows it...* rather than ... *John knows that it's raining....*

*Be Articulate* is controlled by a Gricean principle of manner, *Be Brief*, which prohibits unnecessary prolixity, and takes precedence over *Be Articulate*. Importantly, *Be Brief* is sensitive to linear order: it prohibits material that can be determined to be redundant *when the beginning of the sentence is taken into account*. This has for instance the effect of ruling out the sentence *If it is raining, it is raining and John knows it*. The reason is that second occurrence of *it is raining* can be determined to be redundant when the beginning of the sentence (namely, *If it is raining*) is taken into account.

With precise statements of *Be Articulate* and *Be Brief*, Schlenker 2007 proves that this pragmatic analysis of presuppositions derives the results of Heim 1983 for a fragment with generalized quantifiers, *modulo* technical assumptions. (In turn, the Transparency theory is essentially equivalent to the theory of presupposition based on a pragmatic reconstruction of local contexts in Schlenker 2009).

Now one can observe that often the 'natural' way to articulate explicitly the content of a co-speech gesture is by way of a post-posed modifier rather than by way of an initial conjunct. Thus Schlenker 2016b takes the articulated version of (62)a to be a version of (62)b.

(62) a. None of these 10 guys *SLAP* **punished** his son.

b. None of these 10 guys punished his son like *SLAP* **this** / by slapping him.

But now to ensure that the latter is in violation of *Be Brief* because it is redundant, one has access to the entire beginning of the sentence, *including the VP*, in this case *punished his son*. As a result, as soon as it is assumed that for each of the relevant individuals *x*, *x punished x's son* entails *x slapped x's son*, the post-posed modifier will be redundant. Importantly, this analysis derives essentially the same predictions as a theory that postulates that co-speech gestures trigger conditionalized presuppositions.

Still, this analysis offers a relatively principled way of deriving the fact that cosuppositions as presuppositions that are *conditionalized* on the content of the expression they modify: in the end, this is due to the fact that gestural enrichments are naturally articulated *after* the expressions they modify.

(We write that the account is 'relatively principled' rather than 'principled' because we must still stipulate that the 'articulated' competitor of the VP-with-gesture in (62)a is the VP-with-post-posed-modifier in (62)b.)

## 5.2 Event semantics implementation

As mentioned in Schlenker, to appear b (following M. Krifka, p.c.), all three analyses should probably be implemented within an event or situation semantics in order to predict appropriate inferences. Consider again the sentences *John SLAP punished his son* or *Did John SLAP punish his son?* We predict a presupposition to the effect that *if John punished his son, some slapping occurred*. But this entailment does not specify what kind of connection, if any, there should be between the punishment and the slapping. We clearly want something stronger, namely: *if John punished his son, slapping was involved in the punishment in question*. As Krifka (p.c.) points out, this more adequate inference can be obtained by integrating cosuppositions to an event semantics, one in which *punished* holds true of events and individuals rather than just of individuals; we refer the interested reader to a slightly more detailed discussion of this point in Schlenker, to appear b.

## 5.3 At-issue use of some co-speech gestures

Presupposition triggers can often contribute to the at-issue component of a sentence if the circumstances are right. Thus *his girlfriend* normally triggers a presupposition that the relevant male has a girlfriend, and *x has stopped smoking* triggers the presupposition that x used to smoke; but not so in (63) and (64), where these contributions are at-issue and interpreted in the scope of a modal and a question operator respectively (see also Beaver 2010 for relevant examples).

(63) A: I wonder why that guy is looking so glum.

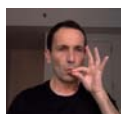
B: Maybe his girlfriend jilted him. (Fauconnier 1985, cited in Beaver and Geurts 2011)

(64) I notice that you keep chewing on your pencil. Have you recently stopped smoking? (Geurts 1994, cited in Simons 2001)

There are ill-understood differences among presupposition triggers; some triggers, called 'weak triggers', make it particularly easy to turn their presuppositional component into an-issue contribution (for analyses, see Abusch 2010 and Abrusan 2016). *Realize*, which we used as a presuppositional control in our discussions above, is an example.

We mentioned in Section 3.5 that recent experimental results suggest that co-speech gestures are indeed weak triggers. Another possibility, due to Esipova 2016a,b, is that quite generally when a presupposition trigger is focused, its presupposition is locally accommodated if this is necessary to ensure that various pragmatic conditions are satisfied (see also Simons et al., to appear). Esipova develops precise generalizations to that effect, and she argues that they hold both of standard presuppositions and of gestural cosuppositions. While her conclusions are very convincing, they do not suffice to derive the at-issue readings of examples discussed in Section 3.5, where focus seemed irrelevant; thus we might *also* need to posit that co-speech gestures are relatively weak triggers, as we have assumed in this piece.

In addition, there are cases in which there might be pragmatic pressure to use a co-speech gesture to express a part of the at-issue meaning that would be too 'dangerous' otherwise.<sup>23</sup> Thus in (65)a, there need not be a general inference that smoking involves pot. Rather, the co-speech gesture appears to make it into the at-issue contribution of the *if*-clause, possibly because one does not want to use the expression 'smoke pot', which refers to a prohibited action.



(65) If you need to  SMOKE\_WEED **smoke**, you should go outside.

<sup>23</sup> Thanks to Larry Horn for helpful remarks on this topic.

#### 5.4 *At-issue use of some co-sign gestures?*

The reader will have noticed that all of our co-sign gestures involve facial expressions. This is because co-sign gestures raise complex problems pertaining to their identification (how do we distinguish a gesture from a sign created 'on the fly') and to their projective status (at-issue or cosuppositional, in particular).

We believe that there might be cases in which co-sign manual gestures have an at-issue use, as is illustrated in (66), where the one-handed sign for *SCOLD* was accompanied with a gesture of slapping, performed with the non-dominant hand, either simultaneously or with a slight delay, or performed after *SCOLD* with the dominant hand. In all cases an at-issue meaning was obtained. Caution is needed, however: as a referee notes, co-articulation of signs is possible in ASL, and thus we cannot fully exclude at this point that we are dealing with two co-articulated signs (in this case, the fact that both make an at-issue contribution might be relatively unsurprising).

*Notation:* manual gestures are transcribed in lowercase letters, and they are subscripted before a sign if they are produced simultaneously with it (*slap*<sub>SCOLD</sub>), and subscripted after a sign if their beginning is slightly delayed (*SCOLD* *slap*). A gesture with its own time slot is transcribed in lowercase letters but is not subscripted.

- (66) JOHN IX-a POSS-a SON IX-b IX-a NEVER  
'His son, John never
- a. <sup>7</sup> SCOLD-b.  
scolds.'
- b. <sup>6,7</sup> *slap*-b SCOLD-b.  
slaps while (or right after) scolding him.'
- c. <sup>6</sup> SCOLD-b *slap*-b.  
scolds and/and then slaps with his left [= non-dominant] hand'.
- d. <sup>7</sup> SCOLD-b *slap*-b.  
scolds and/and then slaps.' (ASL, 34, 2116, 3 judgments)

We believe there are other cases that pattern differently<sup>24</sup>, but we cannot reach a firm conclusion on this matter, which requires further investigations. Be that as it may, the data in (66) raise two issues. First, can we establish that *slap* is a gesture rather than a (co-articulated) sign? Second, if *slap* is indeed a gesture, why does it make an at-issue contribution? There are two directions to explore.

1. One possibility is that manual gestures are more easily seen as integrated with the manual signs they modify, which would imply that they are often analyzed as internal modulations rather than as external enrichments. While this possibility is compatible with the present analysis, we would still expect to find *some* manual gestures that are not so easily viewed as internal; whether some can be found remains an open question.

2. An alternative possibility is that integration occurs in all cases, and that it is for semantic reasons that the expressive component contributed by facial expressions often gives rise to a cosuppositional-type reading.<sup>25</sup> If so, we would need to modify our typology to explain why we do not find instances of *bona fide* external enrichment in sign language: the potential absence of cosuppositional effects triggered by co-sign manual gestures could thus be a problem for our analysis.

## 6 Post-speech and post-sign gestures: English and ASL

Little attention has been devoted to iconic enrichments that come *after* the expressions they modify. We call these 'post-speech gestures' in spoken language, and 'post-sign gestures' in sign language, and

<sup>24</sup> We tested examples involving the one-handed sign for *COOL* performed with the gesture for a big belly, with some projection effects – but not ones that are captured by theories under discussion here.

<sup>25</sup> See Esipova 2016a,b for further discussion of this matter in connection with possibilities (or lack thereof) of local accommodation of inferences triggered by facial expressions.

argue that in both modalities they have the same semantic status as appositive relative clauses, at least in some salient cases. In the terms of Potts 2005, they are 'supplements' (we also explore an alternative view according to which they are just cosuppositions that are more constrained than those triggered by co-speech gestures).

### 6.1 *Post-speech gestures: English*<sup>26</sup>

Schlenker 2015 provides initial arguments against a supplemental analysis of co-speech gestures by focusing on negative environments in which supplements are degraded but co-speech gestures are acceptable; a fuller typology is discussed in Schlenker, to appear b, which we follow here. Our focus is on size-denoting gestures that modify NPs, as these make a comparison with appositives particularly straightforward. It is immediately apparent in the paradigms in (67)-(72) that the distribution of indicative appositive relative clauses does *not* mirror that of co-speech gestures, but rather that of post-speech gestures. It is of course important to pronounce the appositives as is appropriate for non-restrictive clauses; the relevant reading can be brought out by adding *by the way* right after the relativizer.<sup>27</sup> Overall, and with a possible exception in (72), the acceptability of post-speech gestures tracks that of appositive relative clauses. Since at this point we are interested in the distribution of these expressions rather than in the inferences they give rise to, we omit the latter from the present discussion.

- (67) a. Some/No philosopher brought **LARGE** [a bottle of beer].  
 b. Some/#No philosopher brought a bottle of beer, which (by the way) was **LARGE this** large.  
 (Schlenker 2015)  
 c. Some/#No philosopher brought a bottle of beer – **LARGE**.
- (68) a. It's likely/It's unlikely that the previous speaker brought **LARGE** [a bottle of beer] to his talk.  
 b. It's likely/#It's unlikely that the previous speaker brought a bottle of beer, which (by the way) was **LARGE this** large.<sup>28</sup>  
 c. It's likely/# It's unlikely that the previous speaker brought a bottle of beer – **LARGE**.
- (69) a. One/None of these 10 guys **UP helped** his son.  
 b. One/#None of these 10 guys helped his son, which (by the way) he did by lifting him.  
 c. One/#None of these 10 guys helped his son – **UP**.
- (70) a. It's likely/It's unlikely that John **UP helped** his son.  
 b. It's likely/#It's unlikely that John helped his son, which (by the way) he did by lifting him.  
 c. It's likely/#It's unlikely that John helped his son – **UP**.
- (71) a. One/None of these 10 guys **SLAP punished** his son.  
 b. One/#None of these 10 guys punished his son, which (by the way) he did by slapping him.  
 c. One/#None of these 10 guys punished his son – **SLAP**.
- (72) a. It's likely/It's unlikely that John **SLAP punished** his son.  
 b. It's likely/#It's unlikely that John punished son, which (by the way) he did by slapping him  
 c. It's likely/(?) It's unlikely that John punished his son – **SLAP**.

As mentioned in Schlenker 2015, a supplemental approach could deal with the negative version of the a-examples in (67)-(72)a by taking the gestures to behave like the appositives in the b-examples, but with the indicative appositive clause (e.g. *which was LARGE this large*) replaced with a subjunctive one (e.g. *which would have been LARGE this large*). On the assumption that the resulting sentence is more acceptable, one would still need to ask why such an option should be available. But a more crucial problem is that *this approach does not explain why such a strategy fails to be available in the case of post-speech gestures*, as can be seen in the c-examples, whose negative versions are

<sup>26</sup> This section borrows from Section 4.2.1 of Schlenker to appear b.

<sup>27</sup> Thanks to Miloje Despić for making this suggestion.

<sup>28</sup> Note that here and in the other b-examples one might expect a slight pragmatic oddity with *likely* because this expression implicates that it's not established that the relevant event took place, whereas the *which*-clause presupposes its existence. But in any event the version of the sentence with *unlikely* is far worse.

generally deviant (we come back below to a potential explanation). A more natural account is afforded by positing that post-speech gestures are supplements and that they cannot involve a covert counterfactual mood.

Importantly for purposes of comparison with sign language, related contrasts can be obtained on the basis of facial expressions. (73)a is an example we already discussed to highlight the presuppositional effect of facial expressions co-occurring with a predicate under the quantifier *no*. In (73)b, the same facial expression is post-posed. We believe this yields several readings, which can be paraphrased with different supplements, depending on the size of the constituent that the post-speech gesture modifies. One possible reading is that the fact that none of my friends goes skiing with his parents is assessed as negative. We believe that this reading is also available in (74)b, which will provide a convenient point of comparison for our sign language data.

(73) a. None of my friends goes :-([skiing with his parents]).

=> for each of my friends, skiing with his parents wouldn't be fun

b. None of my friends goes skiing with his parents – :-(.)

c. None of my friends goes skiing with his parents,

(i) which is sad [i.e. it is sad that none of my friends goes skiing with his parents];

(ii) which is unpleasant [i.e. it is generally unpleasant to go skiing with one's parents].

(74) a. No poor state :-([spends money]).

=> it would be bad for poor states (and states in general?) to spend money

b. No poor state spends money – :-(.)

=> it is bad that no poor state spends money, hence it would be good for poor states to spend money

## 6.2 Post-sign gestures: ASL

While several of the co-sign facial expressions we explored in ASL were a bit degraded, post-sign facial expressions were quite acceptable, and displayed the same kind of interaction with logical operators as in English. Importantly, there appear to be cases of ambiguity, which can be explained by allowing for different attachment sites for the post-sign gesture, as was discussed in connection with post-speech gestures in (73) in English.

Each of the paradigms in (53)-(56) can be complemented by considering a condition in which a version of the co-sign gesture appears as a post-sign gesture instead, at the end of the sentence. We note that they usually give rise to the same inferences as a co-sign facial expression over the *entire* sentence, a point we will revisit in Section 7.

Our initial examples involving piano playing give rise to somewhat ambiguous inferences, as shown in (75)-(76). The ambiguity found in (75)a and (76)a can be explained if there an attachment ambiguity pertaining to the post-posed facial expression-qua-supplement, which can be linked to the entire sentence or just to the verb *PLAY-PIANO* (on the other hand, we cannot account for the one judgment - out of five trials - of ambiguity with (75)b).

(75) Post-sign version of (75), with a co-sign control

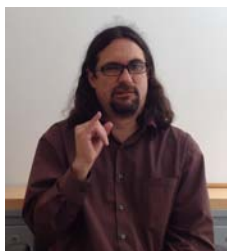
*Context:* John is an amateur musician. The speaker enjoys good music.



a. <sup>6.6</sup> JOHN IX-a NEVER PLAY-PIANO - :-(.)

=> John never plays the piano

=> ambiguous: speaker would be happy / disappointed if John played the piano (2/5 judgments: ambiguous; 2/5 judgments: speaker would be happy; 1/5 judgment: no inference)



b. <sup>6.8</sup> :- ( [JOHN IX-a NEVER PLAY-PIANO].

=> John never plays the piano

=> speaker would be happy if John played the piano (3/5 judgments); or: ambiguous (1/5 judgment):

likely: speaker would be disappointed, possibly: he would be happy (no inference on 1/5 judgment).

'John never plays the piano (badly)' (ASL, 34, 1550b,c, 5 judgments)

(76) Post-sign version of (54), with a co-sign control

Note: \ encodes a backwards movement of the signer

Context: John is an amateur musician. The speaker enjoys good music.



a. <sup>6.8</sup> JOHN IX-a NEVER PLAY-PIANO - \ :- (

=> John never plays the piano

=> the speaker would be happy if John played the piano (4/5 judgments; 1/5 judgment: John would also be disgusted, see Supplementary Materials)



b. <sup>6.6</sup> \ :- ( [JOHN IX-a NEVER PLAY-PIANO]

=> John never plays the piano

=> the speaker would be happy if John played the piano (5/5 judgments)

'John never plays the piano.' (ASL, 34, 1546b,c; 5 judgments)

For what it's worth, we believe that similar readings are available with disgusted or unhappy expressions in English sentences such as (77), although there might be an ambiguity between two readings:

(77) a. John never plays the piano – :- ( .

=> either (i) it is gross/weird that John never plays the piano, or (ii) by the speaker's or John's judgment, playing the piano is gross/weird

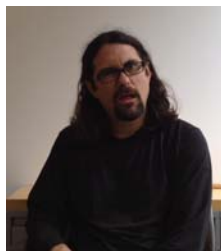
b. :- ( [John never plays the piano].

=> either (i) it is gross/weird that John never plays the piano, or (ii) by the speaker's or John's judgment, playing the piano is gross/weird

The post-sign (and co-sign control) versions of the examples in (55)-(56) above give rise to clearer inferences: the post-sign facial expression is understood to apply to the entire clause, and it gives rise to the same inferences as a co-sign facial expression spanning the entire clause. Strikingly, the inferences obtained in (78) and (79) are the opposite from those we saw in (55)c and (56)c when the disgusted facial expression only co-occurred with the VP *SPEND MONEY*. The latter case gave rise to the inference that it is bad for poor states to spend money. By contrast, in (78)-(79) we get the inference that is bad *that* no poor states spend money – in other words, it would be *good* for poor states to spend money.

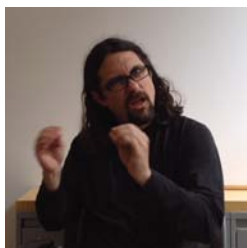


- (78) Post-sign version of (55), with the same co-speech control  
AMERICA [RICH STATES HELP PEOPLE]<sub>b</sub>.



- a. <sup>7</sup> [POOR STATES NONE IX-arc-a SPEND MONEY]<sub>a</sub> - :- ( ) .

=> the speaker would be happy if poor (and rich) states spend money



- b. <sup>6</sup> [POOR STATES :- ( ) [NONE IX-arc-a SPEND MONEY]]<sub>a</sub>.

=> it is good for poor (and rich) states to spend money

'In the US, rich states help people. But no poor states spend money.' (ASL, 34, 1670b,c; 4 judgments)

- (79) Post-sign version of (56), with a post-speech control  
AMERICA [IF STATE RICH HELP PEOPLE]<sub>b</sub>.



- a. <sup>7</sup> [IF STATE POOR NEVER SPEND MONEY]<sub>a</sub> - :- ( ) .

=> the speaker would be happy if poor (and rich) states spent money



- b. <sup>5.7</sup> :- ( ) [IF STATE POOR NEVER SPEND MONEY]<sub>a</sub>.

=> it is good for poor (and rich) states to spend money

'In the US, if a state is rich, it helps people. If a state is poor, it never spends money.' (ASL, 34, 1712b,c; 3 judgments)

### 6.3 Post-speech vocal gestures

For completeness, we mention that post-speech 'vocal gestures' seem to pattern with standard post-speech gestures (which are often accompanied with onomatopoeias to begin with). In particular, we believe that post-speech vocal gestures are acceptable after positive operators, but not so much after negative operators.

- (80) a. John encountered an obstacle – phh.  
=> there was a crash

- b. It's likely that John will encounter an obstacle – phh.  
=> if John encounters an obstacle, there will be a crash
- c. # It's unlikely that John will encounter an obstacle – phh.
- d. One of these ten guys encountered an obstacle – phh.  
=> the guy who encountered an obstacle crashed into it
- e. #None of these ten guys encountered an obstacle – pff.
- f. John often encountered obstacles – phh.  
=> when John encountered obstacles, he crashed into them
- g. #John never encountered obstacles – pff.

## 7 Post-speech and post-sign gestures: further issues

We explore two main analyses of post-speech and post-sign gestures. One, which had our preference in Schlenker, to appear b, posits that post-speech gestures have the semantics of supplements (just like appositive relative clauses), and thus do not introduce presuppositions. An alternative, suggested by Manfred Krifka and Rob Pasternak (p.c.), extends the cosuppositional analysis to post-speech/sign gestures, but posits that post-speech/sign gestures differ from co-speech/sign gestures in that they must recover their antecedent by an anaphoric mechanism. Before we get there, however, we should discuss an even simpler theory, according to which post-speech gestures have an at-issue semantics.<sup>29</sup>

### 7.1 An at-issue semantics?<sup>30</sup>

If post-posed gestures had an at-issue semantics, they should presumably be able to take scope under operators. The deviance of the negative version of the c-examples in (67)-(71) already suggests that under the negative quantifier *No NP* this is difficult. But one could argue that the post-speech gesture is in this case attached 'too high' to be in the scope of the quantifier, possibly because it can only be conjoined with full-fledged clauses. This, however, would fail to account for the cases of embedding under *unlikely*, since in this case it is possible for the post-speech gesture to attach to a clause in the scope of the modal adverb. Similarly, to the extent that (81)a is interpretable, it seems to us that it sharply differs from the controls with at-issue modifiers in (81)b-c: the latter are interpreted in the scope of *deny*, but the post-speech gesture *SLAP* isn't.

- (81) a. ?? I deny that John punished his son – SLAP.  
 b. I deny that John punished his son (and [that he] did so) like SLAP this.  
 c. I deny that John punished his son (and [that he] did so) by slapping him.

Similar remarks hold for (82)a: to the extent that it is acceptable, it seems to us that it does not raise the question but rather assumes that punishment would involve slapping – unlike the at-issue controls in (82)b-c.

- (82) a.?(?) Do you think that John punished his son – SLAP?  
 b. Do you think that John punished his son (and [that he] did so) like SLAP this?  
 c. Do you think that John punished his son (and [that he] did so) by slapping him?

An at-issue semantics for post-speech gestures would need to posit that they somehow cannot take scope under operators. This could be done by postulating that they are connected to their antecedent by an anaphoric device that strongly favors matrix resolution; why this should be so is currently unclear (the import of such an anaphoric mechanism is revisited in Section 7.2 in the context of a cosuppositional theory of post-speech gestures). In addition, it is not clear that this 'matrix only' requirement would account for the data in (83): it seems to us that a post-speech gesture in the antecedent of a conditional, as in (83)a, yields the same kinds of inferences as an appositive or parenthetical attached to the antecedent, as in (83)c,d (which sharply differ in their meaning from (83)b, which involves an at-issue modifier). By contrast, a post-speech gesture following the entire sentence, as in (83)e, does not yield a very coherent reading. But if matrix attachment is the only possibility there is for post-speech gestures, it is hard to see how the contrast between (83)a and (83)e can be derived.

<sup>29</sup> Thanks to Cornelia Ebert (p.c.) for discussion of this point.

<sup>30</sup> This section borrows from Section 4.2.2. of Schlenker to appear b.

- (83) a. If John punishes his son – SLAP, I might scream.  
 => if John punishes his son, slapping will be involved  
 b. If John punishes his son by slapping him, I might scream.  
 ≠> if John punishes his son, slapping will be involved  
 c. If John punishes his son, which will/would involve some slapping, I might scream.  
 => if John punishes his son, slapping will be involved  
 d. If John punishes his son (this would involve some slapping), I might scream.  
 => if John punishes his son, slapping will be involved  
 e. ?? If John punishes his son, I might scream – SLAP.

On the other hand, it is striking that the post-speech gestures in (67)-(71) have the same distribution as appositive relative clauses in the indicative mood. This led Schlenker 2015, to appear b to propose that post-speech gestures are supplements and thus display the same type of behavior as appositive relative clauses.

Still, it should be noted that full parentheticals also display a not-at-issue behavior, and hence a variant of this theory could treat post-speech gestures as parentheticals. For instance, (67)b, repeated as (84)b, displays the same behavior as (84)a, which involves a clausal parenthetical in lieu of an appositive relative clause. The point generalizes across examples and thus at this point it cannot be excluded that post-speech gestures are parentheticals.

- (84) a. Some/#No philosopher brought a bottle of beer (it was LARGE **this** large).  
 b. Some/#No philosopher brought a bottle of beer, which (by the way) was LARGE **this** large.  
 (Schlenker 2015)

We leave it for future research to distinguish between an appositive and parenthetical treatment of post-speech gestures.

## 7.2 A cosuppositional analysis (Krifka, Pasternak, p.c.)?<sup>31</sup>

An alternative was suggested by Manfred Krifka (p.c.), and independently by Rob Pasternak (p.c.). Krifka's proposal is that post-speech gestures have a cosuppositional semantics, but come with an anaphoric element that must be resolved – hence the deviance of the negative version of the c-examples in (67)-(71). This approach could immediately account for the similarity between the behavior of post-speech/sign facial expressions on the one hand, and facial expressions that span an entire sentence on the other: for the proposed analysis, the two cases yield similar readings, but the post-speech/sign version does by way of anaphora what the co-speech/sign version does by way of simultaneous production.

Without fully implementing the proposal, it could be made concrete along the lines of (85):

- (85) Cosuppositional analysis of post-speech gestures and facial expressions (informal statement)  
 a. A post-speech gesture or facial expression  $G$  takes a null anaphoric expression, hence has the form:  
 $G\ pro_i$   
 b. It yields a semantic failure unless, relative to the local context of  $G\ pro_i$ ,  $pro_i$  entails the (bivalent) content of  $G$  ('generalized entailment' must be used if these expressions are predicative rather than propositional).  
 c. If it does not yield a failure as in b.,  $G\ pro_i$  makes a trivial (tautologous) contribution: it has a purely cosuppositional meaning.

As an example, consider again example (73)b, repeated in (86):

- (86) None of my friends goes skiing with his parents – :-(.

There are at least two conceivable ways to resolve the anaphoric element  $pro_i$  introduced by the post-speech gesture.

- If  $pro_i$  is a proposition-denoting pronoun, it is natural to take it to be coindexed with the entire proposition that *none of my friends goes skiing with his parents*. This should yield a cosupposition to

<sup>31</sup> This section borrows from Section 4.2.2. of Schlenker to appear b.

the effect that this proposition entails disgust – hence it would be *good* for my friends to go skiing with their parents.

- If *pro<sub>i</sub>* is a property-denoting pronoun, it is natural to take it to corefer with the property of going skiing with one's parents. There is a technical issue involved at this point: one must determine (i) what the local context of the post-speech gesture is, and (ii) what it means for the property of going skiing with one's parents to entail disgust. Without fully deciding these questions, it is plausible that the local context of the post-speech gesture is the initial context updated with the main sentence. As a result, the predicted inference should be that skiing with one's parents in general should be disgusting (alternatively, one might want to derive the inference that *for my friends* to go skiing with their parents would be disgusting; it isn't quite clear how this weaker reading could be derived, however).

To the extent that (86) is ambiguous, the cosuppositional analysis of post-speech gestures could be maintained. But it encounters difficulties in several cases. In a nutshell, the problem is that post-speech gestures have a more restricted distribution than one might expect on a cosuppositional analysis.

(i) Our initial argument against Ebert and Ebert's (2014) supplement-based analysis of co-speech gestures was that these are less constrained than appositive relative clauses. But as we saw in (67)-(71), post-speech gestures are constrained in exactly this way. The cosuppositional analysis of post-speech gestures would have to posit that the antecedent of the covert pronoun is constrained to be propositional in nature. However this measure won't be enough for examples (68) and (70) above, where a full clause is embedded under *it is unlikely/it is likely that*, but can apparently not be targeted by the post-speech gesture. So one would need to posit that there is a general preference for resolving the null pronoun to a matrix proposition<sup>32</sup>. This proposal would raise the same issues for the paradigm in (83) as the at-issue analysis with resolution to the matrix proposition discussed in Section 7.2. We leave a more detailed development of this potential theory for future research.

(ii) In order to compare a cosuppositional analysis of post-speech gestures with the supplemental analysis, we would need to be more precise about the semantics and pragmatics of supplements. Some analysts take them, following Potts 2005, to be radically different from presupposition triggers in that supplements fail to interact scopally with logical operators. By contrast, Schlenker 2010, 2013 takes supplements to give rise to *bona fide* projection phenomena reminiscent of presuppositions, but with one important proviso: syntactically, appositive relative clauses tend to attach high, with the result that it is only in those rare cases in which they are in the scope of logical operators that genuine supplement projection (as opposed to mere 'high syntactic attachment') can be observed. Due to this potential similarity between supplement projection and presupposition projection, more research would be needed to compare the cosuppositional and the supplemental analysis of post-speech gestures.

(iii) There are important epistemic differences between presuppositions and supplements: the former can and usually are trivial, the latter should make a non-trivial contribution, as shown by (87).

(87) Lance Armstrong survived cancer. #When reporters interview Lance, who survived cancer, he often talks about the disease. (after Potts 2005)

A key question is whether the anti-triviality requirements are different for co-speech and post-speech gestures. Relevant data are discussed in Schlenker, to appear b, but they are not clear.

### 7.3 Intermediate conclusion

At this point, the supplemental (or parenthetical) analysis can explain a lot of the behavior of post-speech gestures. At-issue and co-suppositional analyses have difficulties accounting for the deviance of post-speech gestures in some contexts. They may posit that the antecedent of the post-speech gesture must be a matrix proposition, but this raises an explanatory question (why should this be the case?), and might run into difficulties with some conditional examples in which the antecedent rather than the matrix clause seems to serve as an antecedent.

<sup>32</sup> One possible way to develop the theory (possibly along M. Krifka's lines) would be to take the pronoun *pro<sub>i</sub>* in (85) to be a definite description of events, one that triggers a presupposition to the effect that the relevant events exist. (I believe this point was made by seminar participants at NYU in Fall 2016.)

## 8 Pro-speech gestures




In some cases, a gesture can *replace* rather than enrich a word (e.g. Slama-Cazacu 1976, Clark 1996, De Brabanter 2010). We call such expressions 'pro-speech gestures' (with *pro* meaning: 'replacing', as in *pronoun* and *proconsul*), and we argue that they should play a systematic role in gesture studies, for three reasons: (i) they enrich the typology of iconic inferences; (ii) they provide a way to create words 'on the fly', and thus to test how some inferences, notably presuppositions, are generated in the first place; (iii) in some limited cases, they make it possible to replicate within spoken language some properties of sign language. Here we focus on the role of pro-speech gestures in the typology of iconic inferences, and in the problem of presupposition generation; some morpho-syntactic properties of pro-speech gestures reminiscent of sign language agreement verbs are discussed in Schlenker and Chemla, to appear.




### 8.1 Pro-speech gestures can be at-issue

Unlike co-speech gestures, pro-speech predicates make an at-issue contribution. Thus (90)a contrasts with (90)b, which triggers a universal conditional presupposition (of the form: 'if punished, then hanged'). As is the case for post-speech gestures, acceptability seems increased when an onomatopoeia is produced concurrently with the gestural predicate, and thus makes the iconic representation more complete. In addition, in the case of pro-speech gestures this might help justify the absence of a spoken word. We can complete the paradigm by noting that, as expected, the post-speech version of the gesture yields deviance in this context, as shown in (90)c.

- (90) a. None of these ten traitors was HANG\_ .  
=>none of these ten traitors was hanged
- b. None of these ten traitors was HANG\_  punished.  
=>none of these ten traitors was punished, and for each, punishment would have meant hanging
- c. #None of these ten traitors was punished – HANG\_ .

There are ill-understood constraints on the syntax of pro-speech gestures. Our impression is that they are happiest at the end of a clause, and thus with transitive constructions it is helpful to move the object to a pre-verbal position. This makes it possible to extend the paradigm in (90) to cases involving *SLAP*. In unembedded cases, as in (91)b, we obtain inferences that are close but not identical to those obtained with co-speech and post-speech gestures – for the simple reason that the latter trigger an inference about the relation between the content of an overt predicate (e.g. *punish*) and the content of the gesture, whereas pro-speech gesture have no predicate to modify. In quantified examples, we regain the contrast we just observed with *HANG*: under the time quantifier *never*, *SLAP* has an at-issue contribution in (92)b, a cosuppositional contribution in (92)b, and it is deviant in (92)c.

- (91) a. Your brother, I am going to SLAP\_  punish.  
=> I will punish your brother by slapping him (hence: slapping will be a punishment)
- b. Your brother, I am going to SLAP\_ .  
=> I will slap your brother (no inference that this is a punishment)
- c. Your brother, I am going to punish - SLAP\_ .  
=> I will punish your brother by slapping him (hence: slapping will be a punishment)

- (92) a. Your brother, I will never SLAP\_  **punish**.  
=> if I punished your brother, this would involve some slapping
- b. Your brother, I will never SLAP\_  .  
=> I will never slap
- c. #Your brother, I will never punish - SLAP\_ .






## 8.2 Pro-speech gestures need not be associated with covert spoken words

Since speech normally involves spoken words rather than gestures, one might think that in the examples under investigation a covert word accompanies the pro-speech gesture. Two reductionist theories suggest themselves: according to one, the pro-speech gesture is in fact a *co-speech* gesture that accompanies a covert word; according to the other, the pro-speech gesture is a code for a spoken word. Neither theory is convincing, as we will now see.

First, are gestural predicates just *co-speech* gestures modifying covert words? This is unlikely to be the case, because the covert word would have to be as specific as the gesture, which in some cases is implausible due to its rich iconic contribution.

We start by noting that that unless the covert word in (90)b is *hanged* rather than *punished*, we won't explain why the gestural contribution is at-issue in (90)b but presuppositional in (90)a. With a covert *hanged*, the desired at-issue meaning is predicted, since *hanged* will produce the appropriate at-issue contribution, and the cosupposition triggered by the gesture will be vacuously satisfied. To obtain the right inferences, then, the covert word has to be as specific as the gesture.

Another argument yields the same conclusion. As noted in Schlenker 2015, *co-speech* gestures can be disregarded under ellipsis and in the focus dimension under *only*,<sup>33</sup> as is illustrated in (93)b. But pro-speech gestures cannot be ignored in the course of ellipsis resolution: in (93)a, the elided VP is preferably understood as *hanged*, hence the need for a specific covert word: *hanged* but not *punished*.

- (93) 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


The conclusion at this point is that *if* pro-speech gestures are analyzed as *co-speech* gestures that modify a covert expression, the latter must be as specific as the content of the gesture. However positing such specific covert words lacks plausibility in the general case because pro-speech gestures have iconic specifications that normal words usually don't. Thus the gestural predicate (94)a has an iconic ('rotating') contribution, but it is not clear which simple expression would have exactly the same one. Certainly 'take off' doesn't, as is suggested by the ellipsis facts in (94): in (94)a the *co-speech* gesture can be ignored in the course of ellipsis resolution, as expected – and thus the elided clause involving a plane is acceptable. Not so in (94)b, presumably because there is no covert *take off* that could be recovered by the elided clause (to the exclusion of the iconic gesture).

<sup>33</sup> In this respect, *co-speech* gestures resemble height specifications of sign language loci, as studied by Schlenker et al. 2013 and Schlenker 2014.

(94) a. This helicopter will soon  [take off], and this plane will too.

b. #This helicopter will soon , and this plane will too.

Now one could posit a more complex expression than ~~take-off~~ to play the role of a covert VP. But in the general cases, this is not plausible because the gesture might be just too specific, so that the covert VPs would have to become implausibly complicated. As an example, in (95) the gesture makes clear that the first two pies should be cut in a specific way (probably in 9 parts, since the gesture shows that pies will be sliced along two vertical and two horizontal lines); a very convoluted expression would be needed to approximate this content.

(95) The first pie should be , the second one should be as well, but the third one shouldn't be.

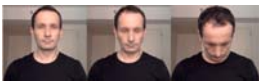
Similarly, in (96), the gesture for *SPREAD* may display *precisely* how the paint should be laid on the canvas, which will permit the students to perform the desired action more precisely than a description in words would. In this case the iconic contribution is gradient and thus very hard to emulate in words.


(96) *Context*: a painter shows his students how to complete a picture.  
When you have reached this point, the red paint, you should *SPREAD* [displaying the precise way in which the paint should be spread], but the blue paint, you shouldn't.

Turning to the second (code-based) theory that could come to mind, the considerations adduced in (94) and (95) also make it implausible that pro-speech gestures function as codes for spoken expressions. In these examples, it is just very unclear what expressions the gestures would be codes for.

### 8.3 *Pro-speech gestures can trigger presuppositions (just like normal words)*

While pro-speech gestures must have an at-issue component, nothing prevents them from also making a presuppositional contribution, just like normal words. In simple cases, their presuppositional component seems rather similar to that of words with a comparable bivalent content as in (97)-(99) (by 'bivalent content', we mean the truth conditions obtained by lumping together presupposition failure and falsity). For instance, the bivalent content of *John stopped smoking* is equivalent to: *John used to smoke and doesn't now smoke*, and similarly the bivalent content of *This helicopter took off* is equivalent to *this helicopter was on the ground and took off*.)

(97) a. In two minutes, our Chair might *DOZE-OFF* .  
=> our Chair is currently awake

b. Within a few minutes, the helicopter we took yesterday might *TAKE-OFF* .  
=> the helicopter we took yesterday is currently on the ground

(98) a. If in two minutes our Chair were to *DOZE-OFF*, everybody would notice.  
=> our Chair is currently awake

b. If in a few minutes the helicopter we took yesterday were to *TAKE-OFF*, I'd be surprised.  
=> the helicopter we took yesterday is currently on the ground

(99) a. None of our ten subjects will *DOZE-OFF*.

=> each of our ten subjects is awake

b. None of your five helicopters will *TAKE-OFF*

=> each of your five helicopters is on the ground

As with presuppositions, the relevant inferences fail to project if they are locally justified:

(100)a. If our Chair is awake, I am sure he'll soon DOZE-OFF.

⇒ our Chair is currently awake

b. If our Chair is awake, I am sure he'll soon fall asleep

⇒ our Chair is currently awake

Since it is unlikely that the presuppositions are triggered by covert spoken words, for reasons mentioned above, there are two main possibilities to explain how they are triggered – although a detailed investigation is beyond the scope of this study.

(i) One possibility is that the presupposition might be triggered by the iconic semantics of the gesture – possibly because its starting point corresponds to a presupposed state.<sup>34</sup> But in (101), there is no discernible inference corresponding to the initial state of the *SMOKE-WEED* gesture. Thus this theory would require a distinction among different kinds of gestural boundaries, possibly along the lines of Wilbur 2003, 2008 and Strickland et al. 2015.



(101)None of the ten guys will ever SMOKE-WEED.

(ii) An alternative (and possibly complementary) possibility is that some presuppositions are not triggered lexically, but are produced by a general 'triggering algorithm' that takes as input the bivalent content of atomic elements, be they words or gestures (Stalnaker 1974, Simons 2003, Abusch 2009, Schlenker 2010, Abrusan 2011). It would then be unsurprising that two atomic expressions with the same bivalent content give rise to the same presuppositions.

The formal study of pro-speech gestures is just in its infancy. For present purposes, it completes our typology by providing gestures that are not syntactically eliminable and seem, possibly for that reason, to come with an at-issue contribution.<sup>35</sup>

#### 8.4 Pro-speech vocal gestures

Unsurprisingly, pro-speech vocal gestures appear to have an at-issue component, as illustrated in (102) with an explosion-like onomatopoeia.

(102)a. Yesterday, as I was driving, suddenly my engine prr.

b. Hopefully this new engine will never prr.

c. Be careful – your engine might prr.

d. Don't worry – it's unlikely that your engine will prr.

We leave a study of vocal pro-speech gestures for future research.

## 9 Typology and theoretical consequences

### 9.1 Typology

We can now revisit the typology we announced in (11), highlighting this time the role of the hypothesis stated in (12) by organizing the table according to the two parameters we introduced:  $\pm$ internal enrichment,  $\pm$ separate time slot. As announced in (12), and repeated in (103), we believe there are two main constraints on this typology.

(103) **Proposed generalization**

a.  $\pm$ internal

External enrichments (-internal) are not at-issue: because they are external, it should be possible to disregard them without affecting the main, at-issue content of the clause they appear in. By contrast,

<sup>34</sup> In fact, *TAKE-OFF* in (97)b triggers another shape-related presupposition, namely that the subject denotes a helicopter: if *helicopters* is replaced with *aircraft* in (99)b, we get an inference that each of the aircraft is a helicopter.

<sup>35</sup> See Schlenker and Chemla, to appear for an attempt to use pro-speech gestures to replicate within English some properties of sign language object agreement verbs.










internal enrichments (+internal) can make any semantic contribution – just like standard words.

b. ±separate time slot

Enrichments that have a separate time slot (+separate time slot) cannot be trivial (= presupposed): because they have their own time slot, they must make a non-trivial contribution to the sentence. By contrast, enrichments that do not have a separate time slot (-*separate time slot*) are not so constrained.

Both parts raise difficulties that are further discussed in Section 9.2.2. For the moment, let us see how they explain parts of the table in (104).

(104) Typology	External enrichments (= syntactically eliminable) => not at-issue		Internal enrichments (= syntactically ineliminable)	
No separate time slot	<b>Cosuppositions</b>		<b>Anything goes: at-issue, etc.</b>	
	<b>Speech</b>	<b>Sign</b>	<b>Speech</b>	<b>Sign</b>
	Co-speech gestures and facial expressions  John  his son.	Co-sign facial expressions IX-arc-b NEVER  [SPEND MONEY].	Iconic vocal modulations The talk was loooong.	Iconic modulations of signs POSS-1 GROUP GROW 
Separate time slot => not presuppositional	<b>Supplements</b>		<b>Usually have an at-issue component</b>	
	<b>Speech</b>	<b>Sign</b>	<b>Speech</b>	<b>Sign</b>
	Post-speech gestures and facial expressions John punished his son – 	Post-sign facial expressions IX-arc-b NEVER SPEND MONEY]b – 	Pro-speech gestures Your brother, I am going to 	[currently unclear]

- The main observation is that the column called 'external enrichments' does not include at-issue meaning contributions, as is expected on the basis of (103)a: external enrichments that don't have their own time slot are cosuppositional, external enrichments that do have their own time slot have the status of supplements. (As noted in Section 5.4, some co-sign manual gestures either refute this generalization or are viewed as internal rather than external enrichments.)
- Within the column 'external enrichment', we can also predict on the basis of (103)b that post-speech gestures and facial expressions should not be presuppositional or cosuppositional. This constraint is indeed satisfied if our main theory is correct (it would not be satisfied if we took post-speech gestures and facial expressions to make a cosuppositional contribution, on the other hand).
- The column called 'internal enrichments' isn't much constrained by the generalization in (103), except for the part corresponding to pro-speech gestures: since these come with their own time slot, they shouldn't be entirely presuppositional.

Two salient observations are *not* derived by the generalization in (103). First, we do not *explain* why co-speech/sign gestures trigger presuppositions rather than supplements, although we do explain why they are not at-issue. The theoretical directions we explored in Section 5.1 go a bit further, without giving a full answer. Second, the generalization does not explain why all the pro-speech gestures discussed so far make an at-issue contribution (although the generalization does explain why they are not purely presuppositional). But since these pro-speech gestures replace full words and these rarely have a purely presuppositional contribution<sup>36</sup>, this is not a particularly surprising result.

<sup>36</sup> Exceptions include the presuppositional particles *too* and *again*.

## 9.2 Theoretical consequences

### 9.2.1 Consequences for pragmatics

In the extant literature, at-issue meanings, presuppositions and supplements are normally triggered by different expressions. Iconic enrichments are rather remarkable in that one and the same gesture – for instance *SLAP*, discussed at some length above – triggers a presupposition (more specifically: a cosupposition) if it co-occurs with an expression; it triggers a supplement if it follows it; and it makes (at least) an at-issue contribution if it fully replaces a word.

While these facts could be handled within a lexicalist account in which each construction (co-speech, post-speech, pro-speech) comes with a different interpretive rule, this might well miss a generalization, specifically the one we stated in (103): across modalities, enrichments that are external to an expression do not initially make an at-issue contribution (although further mechanisms, such as local accommodation, may eventually yield an at-issue contribution); and enrichments that come with their own time slot are not presuppositional. This could argue for a pragmatic analysis in which, in some cases at least, one and the same expression (such as *SLAP*) contributes at-issue, presuppositional or supplemental information depending on considerations of form (possibly following from maxims of manner).

We already discussed in Section 5.1 possible ways of deriving the fact that co-speech/sign gestures trigger cosuppositions. But it remains to explain how the supplemental contribution of post-speech/sign gestures could be connected with their cosuppositional contribution as co-speech/sign gestures. For theories such as Potts 2005, which take supplements to be an entirely separate class of meanings, this might require positing a covert lexical element to derive the desired meaning. This need not be a serious theoretical problem, since Potts 2005 already assumed that the 'comma intonation' was endowed with a lexical meaning that was crucial to the derivation of the meaning of appositive constructions. It would make excellent sense to posit, within this framework, that post-speech gestures also come with this comma intonation. Alternatively, one might take supplements and presuppositions to form a natural class, as argued in Schlenker 2010, 2013. On this view, supplements differ from presuppositions in that they should not be trivial in their local context. But they resemble presuppositions in that they give rise to the same kind of projection phenomena, at least when they are attached in the scope of logical operators (a syntactic part of the theory argues that appositives can be attached to any propositional node that dominates their surface position).<sup>37</sup>

### 9.2.2 Consequences for the comparison between spoken and sign language

The generalization we developed in (103) does not make specific reference to spoken or to sign language, but it makes use of categories that cut across modalities, namely whether an enrichment is internal or external; and whether it has a separate time slot. If this analysis is on the right track, there is no need for semantic or pragmatic stipulations tailored to one modality or to the other (although as noted in Section 5.4 there remains much work to be done on manual gestures in ASL).

From this, however, it does not follow that iconic enrichments behave in the same way across the two modalities. We noted that external enrichments do not normally make at-issue contributions. But since the iconic possibilities of the vocal medium are quite limited, most iconic enrichments in spoken language are effected by way of external enrichments, and hence they are not at-issue. As a consequence, *the expressive possibilities of sign and spoken language should in the end be quite different*. This will not come as a surprise to signers, who are often keenly aware of the fact that iconic enrichments play a greater role in sign than in spoken language.

Our generalizations also entail that iconic modulations in sign language are semantically closer to vocal modulations in spoken language than to co-speech gestures. In particular, a lengthened version of ASL *GROW* undergoes the same kind of iconic enrichment as a lengthened version of English *long*: both are iconic modulations, and in both cases the iconic enrichment is at-issue. By contrast, English

---

<sup>37</sup> The proposal in Schlenker 2013 was that supplements should be non-trivial in their local context, but they should be 'easy' to accommodate, in the sense that the global context C could 'easily' be turned into a strengthened context C<sup>+</sup> relative to which a supplement becomes locally trivial.

*grow* accompanied with co-speech gesture to indicate a 'large' growth should behave differently, as it is an instance of external enrichment. This slightly counter-intuitive typology should be kept in mind when iconic enrichments are studied across modalities, in particular from a neurolinguistic perspective.<sup>38</sup>

## 10 Conclusions

Our typology of iconic enrichments was stated in a modality-free fashion, by way of two main categories: the internal or external nature of the enrichments; their appearance with a separate time slot or not.

Despite this modality-free analysis, there remain differences between iconic enrichments in sign and speech. Many iconic enrichments in sign language are realized by way of iconic modulations. These exist in spoken language as well (as in the case of *loooong*), but due to the representational limitations of the vocal stream, they are by force limited in nature. Co-speech gestures and post-speech gestures, by contrast, have rich representational possibilities, but they are not initially at-issue. By contrast, pro-speech gestures make at-issue contributions, but they differ from iconic modulations in sign language in not being based on words. Thus there should in the end be significant expressive differences between spoken and sign language when iconicity is properly taken into account.

This study also raises questions for formal pragmatics. Standard accounts of presuppositions and supplements are currently based on lexical stipulations. Since these are relatively unconstrained, they can in principle be extended to account for our data. But this might be missing two generalizations, to the effect that (i) external enrichments are not (initially) at-issue, and (ii) enrichments that have their own time slot are not presuppositional. Such generalizations should be further explored and possibly derived in future pragmatic research.

Finally, we leave one important question for future work. Focus can notoriously affect the meaning contribution of an expression. We briefly discussed this in connection with the local accommodation of presuppositions, which in some cases seems to be facilitated by focus (Simons et al., to appear, Esipova 2016a,b). In addition, in 'out of the blue' contexts, focus on an expression can yield the impression that the meaning component which is 'given' (i.e. not in focus) is presupposed. Much research has argued that these 'givenness' effects are in fact different from presuppositions (e.g. Buring 2012). But we would need to study more controlled paradigms in order to rigorously isolate the potential contribution of focus. In particular, future studies might get some mileage out of the observation that enrichments that come with a separate time slot are easily interpreted as focused, whereas this is probably more difficult for expressions that do not have a separate time slot.

---

<sup>38</sup> It must be kept in mind, however, that our comparison between speech and sign was not fully minimal. In particular, we investigated a variety of co-speech gestures, but very few co-sign gestures: only facial expressions fulfilled this function. As emphasized by M. Esipova (p.c.), it cannot be excluded that these make a special contribution due to their semantics rather than to their co-sign status, and thus further gestures should be investigated in sign language.

### References

- Abrusan, Marta: 2016, Presupposition Cancellation: Explaining the 'Soft–Hard' Trigger Distinction. *Natural Language Semantics* 24 (2):165-202.
- Abusch, Dorit: 2010, Presupposition triggering from alternatives. *Journal of Semantics* 27(1): 37–80.
- Anvari, Amir: 2016, Iconic Gestures: Presuppositions vs. Cosuppositions. Manuscript, Institut Jean-Nicod.
- Beaver, David: 2001, *Presupposition and assertion in dynamic semantics*. Stanford: CSLI.
- Beaver, David: 2010, Have you noticed that your belly button lint colour is related to the colour of your clothing? In R. Bauerle, U. Reyle, and T. E. Zimmerman (Eds.), *Presupposition: Papers in Honor of Hans Kamp*. Bingley, UK: Emerald.
- Beaver, David I. and Geurts, Bart: 2014, Presupposition. *The Stanford Encyclopedia of Philosophy* (Winter 2014 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/win2014/entries/presupposition/>>.
- Büring, Daniel: 2012, Focus and Intonation. In: Gillian Russell and Delia Graff Fara (eds), *Routledge Companion to the Philosophy of Language*. London: Routledge. 103-115.
- Chemla, Emmanuel: 2009, Presuppositions of quantified sentences: experimental data. *Natural Language Semantics*,17(4):299-340.
- Clark, Herbert H.: 1996, *Using language*. Cambridge: Cambridge University Press.
- Clark N, Perlman M, Johansson Falck M: 2013, Iconic pitch expresses vertical space. In *Language and the creative mind* (eds B Dancygier, M Borkent, J Hinnell), pp. 393–410. Stanford, CA: CSLI Publications.
- Cross, I. and Woodruff, G. E.: 2008, Music as a communicative medium. In Botha, R. and Knight, C. (Eds.), *The Prehistory of Language*, Vol. 1, pp. 113–144.
- De Brabanter, Philippe: 2010, Uttering sentences made up of words and gestures. In *Explicit communication: Essays on Robyn Carston's Pragmatics*, ed. B. Soria and E. Romero, 199–216.
- Dingemans, Mark: 2013, Advances in the cross-linguistic study of ideophones. *Language and Linguistics Compass* 6.654–72. DOI: 10.1002/lnc3.361
- Dingemans, 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*.
- Dingemans, 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, 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 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.
- 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.
- Fauconnier, Gilles, 1985, *Mental Spaces: Aspects of Meaning Construction in Natural Language*, Cambridge, Massachusetts: MIT Press.
- Fitch, Tecumseh W., Reby, D.: 2001, The descended larynx is not uniquely human. *Proceedings of the Royal Society of London. Series B*, 268, 1669-1675.
- Fricke, Ellen: 2008, Grundlagen einer multimodalen Grammatik des Deutschen: Syntaktische Strukturen und Funktionen. Habilitation treatise, European University Viadrina, Frankfurt (Oder).
- Goldin-Meadow, Susan and Brentari, Diane: to appear, Gesture, sign and language: The coming of age of sign language and gesture studies. *Behavioral and Brain Sciences*
- Greenberg, Gabriel: 2013. Beyond Resemblance. *Philosophical Review* 122:2, 2013

- Grenoble, Lenore A.; Martinovic, Martina; Baglini, Rebekah: 2015, Verbal Gestures in Wolof. *Selected Proceedings of the 44th Annual Conference on African Linguistics*, ed. Ruth Kramer et al., 110-121. Somerville, MA: Cascadilla Proceedings Project.
- Heim, Irene: 1983, On the Projection Problem for Presuppositions. In D. Flickinger et al. (eds), *Proceedings of the Second West Coast Conference on Formal Linguistics*, 114-125. Reprinted in Davis 1991.
- Henderson, Robert: 2016, Pluractional Demonstrations. Manuscript. <http://ling.auf.net/lingbuzz/003127>
- Kegl, Judy: 1977/2004. ASL Syntax: Research in progress and proposed research. *Sign Language & Linguistics* 7:2. Reprint of an MIT manuscript written in 1977.
- Kuhn, Jeremy: 2015, *Cross-categorial singular and plural reference in sign language*. Doctoral dissertation, New York University.
- Kuhn, Jeremy and Aristodemo, Valentina: 2016, Pluractionality, iconicity, and scope in French Sign Language. Manuscript, Institut Jean-Nicod.
- Ladewig, Silva: 2011, *Syntactic and semantic integration of gestures into speech: Structural, cognitive, and conceptual aspects*. PhD thesis, Frankfurt (Oder): European University Viadrina.
- Lewis, David: 1975, Adverbs of Quantification. In Edward L Keenan (ed.), *Formal Semantics of Natural Language*. Cambridge: Cambridge University Press, pp. 3–15.
- Liddell, Scott K.: 2003. *Grammar, Gesture, and Meaning in American Sign Language*. Cambridge: Cambridge University Press.
- Lillo-Martin, Diane and 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.
- 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.
- Okrent, Arika: 2002, A modality-free notion of gesture and how it can help us with the morpheme vs. gesture question in sign language linguistics, or at least give us some criteria to work with. In R.P. Meier, D.G. Quinto-Pozos, & K.A. Cormier (eds). *Modality and structure in signed and spoken languages* (pp. 175-198). Cambridge: Cambridge University Press.
- Ortega, G., Schiefner, A., and Ozyurek, A.: to appear, Speakers' gestures predict the meaning and perception of iconicity in signs. In *Proceedings of the 39th Annual Conference of the Cognitive Science Society* (CogSci 2017).
- Perlman, M. and Cain, A.: to appear, Iconicity in vocalizations, comparisons with gesture, and implications for the evolution of language. *Gesture*.
- Potts, Christopher. 2005. *The Logic of Conventional Implicatures*. Oxford University Press.
- Schlenker, Philippe: 2007, Anti-Dynamics (Presupposition Projection Without Dynamic Semantics). *Journal of Logic, Language and Information* 16, 3: 325-256
- Schlenker, Philippe: 2008, *Be Articulate: A Pragmatic Theory of Presupposition Projection*. *Theoretical Linguistics*, 34, 3: 157-212
- Schlenker, Philippe: 2009, Local contexts. *Semantics and Pragmatics* 3:1–78.
- Schlenker, Philippe: 2010, Supplements Within a Unidimensional Semantics I: Scope. In: Maria Aloni, Harald Bastiaanse, Tikititu de Jager, and Katrin Schulz (eds) *Logic, Language and Meaning: 17th Amsterdam Colloquium*, Amsterdam, The Netherlands, December 16-18, 2009, Revised Selected Papers, Springer, 74-83.
- Schlenker, Philippe: 2013, Supplements Within a Unidimensional Semantics II: Epistemic Status and Projection. In S. Kan, C. Moore-Cantwell & R. Staubs (eds.), *Proceedings of the 40<sup>th</sup> Annual Meeting of the North East Linguistic Society*, Vol. 2. GLSA. Amherst, MA. 167–182.
- Schlenker, Philippe: 2014, Iconic Features. *Natural Language Semantics* 22, 4: 299-356.
- Schlenker, Philippe: 2015, Gestural Presuppositions (squib). *Snippets* (Issue 30) doi: 10.7358/snippet-2015-030-schl
- Schlenker, Philippe: 2016a, The Semantics/Pragmatics Interface. In: Aloni, Maria and Dekker, Paul (eds), *The Cambridge Handbook of Formal Semantics*, pp. 664-727 <http://dx.doi.org/10.1017/CBO9781139236157.023>
- Schlenker, Philippe: 2016b, Gesture Projection within the Transparency Theory. Manuscript, Institut

- Jean-Nicod and New York University.
- Schlenker, Philippe: 2017a, Sign Language and the Foundations of Anaphora. *Annual Review of Linguistics*, 3:149–77
- Schlenker, Philippe: 2017b, Outline of Music Semantics. *Music Perception: An Interdisciplinary Journal* 35, 1: 3-37 DOI: 10.1525/mp.2017.35.1.
- Schlenker, Philippe: to appear a, Visible Meaning: Sign Language and the Foundations of Semantics. Accepted for publication as a target article in *Theoretical Linguistics*.
- Schlenker, Philippe: to appear b, Gesture Projection and Cosuppositions. To appear in *Linguistics & Philosophy*.
- Schlenker, Philippe: to appear c, Super Monsters I. *Semantics & Pragmatics*.
- Schlenker, Philippe: to appear d, Super Monsters II. *Semantics & Pragmatics*.
- Schlenker, Philippe: to appear e, Prolegomena to Music Semantics. Accepted with revisions, *Review of Philosophy & Psychology*.
- Schlenker, Philippe, and Chemla, Emmanuel: to appear, Gestural Agreement. *Natural Language & Linguistic Theory*.
- Schlenker, Philippe and Lamberton, Jonathan. 2016. Iconic Plurality. Manuscript, Institut Jean-Nicod and New York University.
- Schlenker, Philippe, Lamberton, Jonathan, and Santoro, Mirko: 2013. Iconic Variables. *Linguistics & Philosophy* 36(2): 91-149.
- Simons, Mandy: 2003, Presupposition and Accommodation: Understanding the Stalnakerian Picture. *Philosophical Studies* 112: 251–278, 2003. DOI 10.1023/A:1023004203043
- Simons, Mandy; Beaver, David; Roberts, Craige; Tonhauser, Judith: to appear, The Best Question: Explaining the Projection Behavior of Factives. To appear in a special issue of *Discourse Processes*, edited by Anton Benz and Katja Jasinskaja.
- Slama-Cazacu, T.: 1976, Nonverbal components in message sequence: “Mixed syntax.” In W. C. McCormack & S. A. Wurm (Eds.), *Language and Man: Anthropological Issues* (pp. 217–227). The Hague: Mouton.
- Stalnaker, Robert: 1974, Pragmatic Presuppositions. In Munitz, M. and Unger, P. (eds.) *Semantics and Philosophy*. New York: New York University Press.
- Strickland, Brent; Geraci, Carlo; Chemla, Emmanuel; Schlenker, Philippe; Kelepir, Meltem; Pfau, Roland: 2015, Event representations constrain the structure of language: Sign language as a window into universally accessible linguistic biases. *PNAS* [www.pnas.org/cgi/doi/10.1073/pnas.1423080112](http://www.pnas.org/cgi/doi/10.1073/pnas.1423080112)
- Tieu, Lyn; Pasternak, Robert; Schlenker, Philippe; Chemla, Emmanuel: 2017, Co-speech gesture projection: Evidence from inferential judgments. Manuscript.
- Tieu, Lyn; Pasternak, Robert; Schlenker, Philippe; Chemla, Emmanuel: to appear, Co-speech gesture projection: Evidence from truth-value judgment and picture selection tasks. *Glossa*.
- Wilbur, Ronnie B. : 2003. Representations of telicity in ASL. *Chicago Linguistic Society* 39 : 354-368.
- Wilbur, Ronnie B.: 2008. Complex Predicates Involving Events, Time and Aspect: Is This Why Sign Languages Look So Similar?. In: Quer, J. (ed.) *Signs of the Time*, pp. 217–250. Signum, Hamburg.
- Wilbur, Ronnie B. & Malaia, Evie : 2008. Event Visibility Hypothesis: Motion Capture Evidence for Overt Marking of Telicity in ASL. Chicago, IL: Linguistic Society of America.

***Supplementary Materials: Raw ASL Data***

Raw ASL data can be downloaded in .docx format at the following URL:

<https://drive.google.com/file/d/0B7Mz-VKVeYNKNjBDNF95Zk9mUXM/view?usp=sharing>