Linguistic inferences without words: the case for pro-speech vocal gestures

Janek Guerrini (Institut Jean Nicod, Ecole Normale Supérieure)
Philippe Schlenker (Institut Jean Nicod, Ecole Normale Supérieure; NYU)

Schlenker, to appear a argues that the full typology of linguistic inferences (including scalar implicatures, presuppositions, supplements, homogeneity inferences, and expressives) can be replicated with pro- and post- speech gestures, i.e. gestures that fully replace some words, or follow words they modify. Tieu et al. 2018 provide experimental evidence for this claim for the first four inferential types. They also extend the claim by showing that visual animations in lieu of gesture can yield the same effects. Thus, subjects have access to productive algorithms that divide the informational content of entirely new semantic objects among the various components of the inferential typology – a non-trivial claim in view of current theories (e.g. for presuppositions, where the result argues for a general triggering algorithm).

But could some of these results be due to special properties of the visual modality? To address this question, we investigate the same issues with vocal gestures, and provide a clear replication with scalar implicatures, presuppositions and supplements, and a less clear one with homogeneity inferences. The phenomenon is thus much broader than was initially thought. We base our claims on the introspective judgments of 2 native speakers of English, and on comparable data in Italian (3 native speakers) (our examples are modeled after Tieu et al. 2018).

[1] A <u>direct scalar implicature</u> arises in (1): a single iteration of *BOOM* is indicative of a single punching event, possibly because it competes with the repeated version *BOOM_3*, realized by three unpunctuated (i.e. hard to separate) repetitions. An alternative analysis is that *BOOM_1* just means 'to punch once'. But this makes incorrect predictions for (2)b, which doesn't deny that Jenny will do exactly one punching, but rather that she will punch at all. In addition, (2)a. is a clear case of an <u>indirect implicature</u>: *not BOOM_3* competes with the more informative *not BOOM_1*, which is taken to be false - hence the inference that Jenny will do some punching.

(1)   Context: John the Alien has been training on the punching bag at the gym. // At last week's workout, John had a lot of energy. He was able to BOOM_3.
a. This week, John will BOOM.
b. This week, John BOOM_3

(2)   Jenny the Alien has been training on the punching bag at the gym. // In her first week of training, Jenny had a lot of energy. She was able to BOOM_3. But in the second week, Jenny did not BOOM.
a. This week, Jenny will not BOOM_3.
=> Jenny will punch somewhat but not a lot
b. This week, Jenny will not BOOM.
=> this week, Jenny will not punch

[2] <u>Presuppositions</u> can be generated by way of vocal gestures that evoke changes of state: whistling with a rising frequency evokes a ball being kicked up, whistling with a decreasing frequency evokes a ball being kicked down. Strikingly, in both cases the initial state (being down/being high) is presupposed. This is shown by the classical projection tests in (3): the presupposition projects out the question, and projects universally out of *none*-type contexts.

(3)   Context: A group of ten friends is playing. Five of them of them are on the ground, five are on a tree, and one ball is thrown back and forth.
a. Do you think the children will (i) WHISTLE-BLOW-LOW—>HIGH-PITCH? / (ii) WHISTLE-BLOW-HIGH—>LOW-PITCH?
=> the children are (i) on the ground (ii) in the tree
b. None of the children will (i) WHISTLE-BLOW-LOW—>HIGH-PITCH / (ii) WHISTLE-BLOW-HIGH— >LOW-PITCH
=> each of the children is (i) on the ground (ii) in the tree

[3] Tieu et al. argue that post-speech gestures are supplements, and they provide a partial, inferential argument. We replicate their result with a vocal gesture. (4)a is a baseline with a

standard supplement, which gives rise to a conditional inference. (4)b has a post- speech vocal gesture that exemplifies how the bothering will take place. It too gives rise to a similar

conditional inference, which is absent from the at-issue control in (4)c, with a 'like this' modifier.

(4)   Context: June has been misbehaving a lot on the playground these days, and her teachers are not very happy with her.
a. If June bothers a classmate today, which will involve punching her, she will get a detention.
=> if June bothers a classmate today, this will involve punching her
b. If June bothers a classmate today – GNEGNEGNEGNEGNE, she will get a
detention.[1] => if June bothers a classmate today, this will involve mocking/teasing her
c. If June bothers a classmate and does so like this - GNEGNEGNEGNEGNE today, she will get a detention.
≠> if June bothers a classmate today, this will involve mocking/teasing her.

Schlenker to appear a further argues that post- speech gestures pattern with supplements in being deviant in in some negative environments, such as (5)c(i). This result too extends to post-speech vocal gestures: (5)c(ii) is deviant, unlike the positive examples in (5)a(ii), b(ii) .

(5)   a June bothered her brother today (i), which angered him (ii) – GNEGNEGNEGNEGNE. b.
One of the girls bothered her brother today (i) , which angered him (ii) - GNEGNEGNEGNEGNE. c.
None of the children bothered her brother today (i) #, which angered him (ii) #-GNEGNE.

[4] <u>Homogeneity</u> (see Kriz&Spector, 2017) is a special inferential type triggered by definite plurals, which behave like (quasi-)universals in positive contexts and (quasi-)existentials in negative contexts. Thus *x found his presents* means that *x found all his presents*, whereas *x didn't find his presents* means that *x didn't find any of his presents*. Tieu et al. provide evidence that homogeneity inferences can be triggered by purely gestural means.

We submit that a homogeneity inference arises in (6). In (6)a, we obtain an inference that the three bottles of champagne will be opened. This could be because the second occurrence of PEMx3 (i) is interpreted as an indefinite ('three bottles'), or (ii) is interpreted as a definite plural ('the [three] bottles'). But (i) incorrectly predicts that (6)b should mean something too weak, i.e. *we won't open three bottles* (hence we might open two). (ii) makes the correct prediction, combined with the observation that definites under negation give rise to an existential reading.

(6)   a. The waiter brought four bottles overall: PEMx3 and TSCH. We opened PEMx3.
=> we opened all of the three bottles of champagne.
       b. The waiter brought four bottles overall: PEMx3 and TSCH. We did not open
PEMx3. => we didn't open any of the champagne bottles.

Informants were two native speakers of English (23- and 24-old females) and four native speakers of Italian (16-, 23-, and 68-old males). The judgements were all in line with expectations. The examples in Italian were adapted in order to deliver expressions that didn't require morphology, like "stare per" ("be about to") and "riuscire a" ("manage to"). For instance, instead of translating literally the English 'will BOOM' into an Italian future, we used the expression 'sta per BOOM'.

One interesting remark was passed by the Italian 16-years old informant, who pointed out that he found the inference we mention in (3) very strong, but just in case both WHISTLE-BLOW-LOW—>HIGH-PITCH and WHISTLE-BLOW-HIGH—>LOW-PITCH had been used in the conversation. He added that he wouldn't have necessarily interpreted the highness of the pitch as meaning-pertinent (as indicative of the provenance and directionality of the thrown ball) if the contrast to the alternative realization of the vocal gesture hadn't made this feature salient. The realizations that correspond to the vocal gestures can be found in form of audio files at the following links:

BOOM: https://goo.gl/gBqVV8, BOOM_3: https://goo.gl/vYyyij,
WHISTLE-BLOW-HIGH->LOW-PITCH: https://goo.gl/qM4QPm, WHISTLE-BLOW-LOW->HIGH-PITCH: https://goo.gl/A8dAoa, GNEGNEGNEGNEGNE: https://goo.gl/3T8JF2,
TSCH: https://goo.gl/945W3h, PEM:https://goo.gl/VkYv1R, PEMx3: https://goo.gl/cQ8xnt.

---

[1] The fact that children often make this sound when they tease one another has already been noted by Leonard Bernstein (cfr. p.16 Leonard Bernstein, The Unanswered Question: Six Talks at Harvard, Harvard University Press, 1976)

Bibliography

Kriz, M.; Spector, B., 2017, "Interpreting Plural Predication: Homogeneity and Non-Maximality". lingbuzz/003458, Manuscript, IJN.

Schlenker, P.: to appear a, "Gestural Semantics: Replicating the typology of linguistic inferences with pro- and post-speech gestures". *Natural Language & Linguistic Theory*.

Schlenker, P.: to appear b, "Iconic Pragmatics". *Natural Language & Linguistic Theory*.

Tieu, L.; Chemla, E.; and Schlenker, P., submitted. "Linguistic inferences without words: Replicating the inferential typology with gestures".