The projection of co-speech sound effects

Robert Pasternak
Leibniz-Center for General Linguistics (ZAS)
pasternakrs@gmail.com

March 15, 2019

Much recent work has focused on the pragmatic contributions of co-speech gestures, gestures aligned with spoken content. Such gestures show a particular, seemingly not-at-issue projection pattern, with various theories being proposed to account for this inference profile (see Lascarides & Stone 2009, Ebert & Ebert 2014, Schlenker 2018, Esipova 2018, Hunter 2018).

A mostly unexplored question is which types of co-speech content generate this projection pattern. None of the following hypotheses (among others) is a priori eliminable:

1. Gestures have a unique role in language evolution, acquisition, and use, so this projection pattern is specific to them.

2. The projection pattern arises from embodied content, including gestures and “body noises” like vocal modulation or clapping.

3. The projection pattern results from multimodality: the visual content of the gesture projects when combined with speech. This excludes body noises and includes other visual imagery like pictures.

4. All co-speech content generates this projection pattern.

Co-speech sound effects make an informative case study, since they lack gestures’ special status, are not embodied content, and use the same modality as speech. Yet when looking at environments discussed by Schlenker (2018) and Tieu et al. (2017, 2018), we see the same inferential pattern arise as for gestures. To illustrate, we will look at examples with the upward-pointing gesture up and the explosion sound explode, focusing on four environments: unembedded and under might, negation, and none. The embedded examples can be contrasted with at-issue controls where \([VP]_{up}\) and \([VP]_{explode}\) are replaced with \(VP\) in \([this]\)_{up} direction and \(VP\) like this: \([ ]\)_{explode}. All sound effect examples and at-issue counterparts can be found at https://bit.ly/2Je6Sto.

In unembedded contexts, both gestures and sounds lead to interpretations in which the co-speech content is conjoined with the VP’s:

(1) a. Beatrice will [take the stairs]_{up}.
   \(\sim\) will take the stairs upward

   b. The soldier will [assassinate his target]_{explode}.
   \(\sim\) will assassinate his target via explosion

Under might, a conditional reading arises:

(2) a. Beatrice might [take the stairs]_{up}.
   \(\sim\) might take the stairs, and if she does she will take them upward
b. The soldier might [assassinate his target] explode.
   \[ \sim \text{might assassinate his target, and if he does it will be via explosion} \]

Under negation, we get counterfactual readings:

(3) a. Beatrice will not [take the stairs] upward.
   \[ \sim \text{will not take the stairs, but if she did it would be upward} \]

b. The soldier will not [assassinate his target] explode.
   \[ \sim \text{will not assassinate his target, but if he did it would be via explosion} \]

And under none this counterfactual inference projects universally:

(4) a. None of these three students will [take the stairs] upward.
   \[ \sim \text{No student will take the stairs, but for each, if she did it would be upward.} \]

b. None of these three soldiers will [assassinate his target] explode.
   \[ \sim \text{No soldier will assassinate his target, but for each, if he did it would be via explosion.} \]

These examples suggest that co-speech sound effects project like gestures, despite lacking many of the notable features of gestures. Thus, the “gesture” projection pattern might actually encompass a much broader variety of co-speech content.

References


