Sources of evidence for acquiring discourse connectives: Explaining production-comprehension asymmetry in the acquisition of but

Connectives such as *but* are critical for building coherent discourse, and express meanings that do not fit neatly into standard semantics/pragmatics distinctions. How do children acquire them? Early corpus analyses concluded that *but* is **produced** on target and mastered by early pre-school[1,2]. However, **comprehension** experiments, both recent and old, show that even school-aged children struggle with *but*, particularly when used to indicate contrast with expectations (e.g., *It was freezing but Mary wore shorts*)[3-5].

What can explain this production-comprehension mismatch? We investigated (1) whether prior corpus analyses over-stated children's capability with *but,* and if so (2) whether acquisition is delayed because the meaning of *but* is hard to learn from caregiver speech. To do this, we used a variant of the human simulation paradigm[6] to measure whether children and adults use *but* in recognisably different contexts from the similar yet non-contrastive connective *and*.

Using corpora of parent-child conversations[7], we masked all instances of *but/and*, and then, for each instance spoken by a child, asked two trained annotators to judge which connective had been used (Figure 1, n=1329 annotations). We used signal detection analyses, comparing Hits (annotating *but* when the child said *but*) to False Alarms (annotating *but* when they said *and*) to generate a d' score, measuring sensitivity to the intended connective based on its context. We compared data from younger (M=38 months, range:18-48) and older (M=79 months, range:73-83) children.

Sensitivity to the intended word was always above chance (p<.001) but significantly worse (*p*<.01) for speech from younger children (d'=1.46, Hit rate=59%) than older children (d'=1.95, Hits=78%). Thus, younger children more often used *but* in contexts where annotators judged *and* to be more appropriate, suggesting they had not mastered its meaning.

Next we used the same technique to assess if caregivers use *but* in very different contexts from the word *and*. Importantly, they did not. While sensitivity to the intended word was above chance (n=4358 annotations, p<.001), it was not high (d'=1.79, Hits=72%) and no greater than for speech from older children. Thus, caregivers use *but* and *and* in similar contexts.

Finally, we examined which senses caregivers used when saying *but,* annotating for 1) the violation-of-expectations sense, 2) simple contrast (e.g., *Paul is big but David is short)*, 3) another word sense, or 4) used *but* mistakenly. Importantly, the violation-of-expectations sense used in comprehension tasks was extremely rare, occurring 14 times in 870 annotations (Figure 2).

These data provide important new context for the observed production-comprehension mismatch in children's discourse skills. First, they show that children's ability to produce connectives like *but* develops more slowly than previously thought, indicating that the mismatch may be more apparent than real. Second, they show that caregivers' speech provides a limited signal for learning *but's* meaning: Not only is it hard to distinguish *but* from *and* based on context, but some of *but's* core senses – like violation of expectations – are extremely rare. These factors may explain why the acquisition of *but* continues into school years. How it eventually develops is still open to question.

(Word count: 499)

Figure 1. Example of the Human Simulation Paradigm annotation task.

Target_Child : at the circus where I got that sword
Mother : oh yes
Target_Child : yeah once he did the steering wheel like that CONNECTIVE that CONNECTIVE
that CONNECTIVE that CONNECTIVE that CONNECTIVE that
Mother : CONNECTIVE what is this girl doing
Mother : she's
**Target_Child : CONNECTIVE CONNECTIVE that TARGET it CONNECTIVE that man was a
really um me um guy CONNECTIVE CONNECTIVE she has long hair CONNECTIVE the guy
had CONNECTIVE the guy at circus had**
Mother : mhm
Mother : it said what is she holding there
Target_Child : mm a soccer ball
Mother : that's not a soccer ball
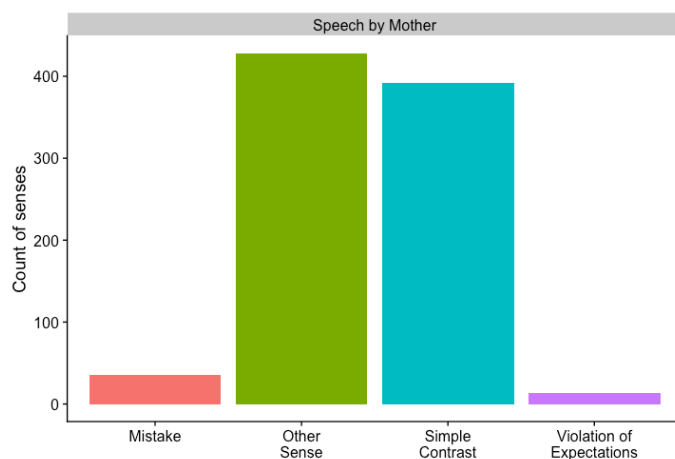Mother : soccer ball's round

Select Connective
○ But
○ And

Select confidence rating

○ Low Confidence
○ Moderate Confidence
○ High Confidence

Figure 2. Senses of *but* used by caregivers (all mothers) to children.

References
[1] Bloom, L., Lahey, M., Hood, L., Lifter, K., & Fiess, K. (1980). Complex sentences: Acquisition of syntactic connectives and the semantic relations they encode. *Journal of child language*, *7*(2), 235-261.

[2] Braunwald, S. R. (1985). The development of connectives. *J. of pragmatics*, *9*(4), 513-525.

[3] Skarabela, B., Cuthbert, N., Rees, A., Rohde, H. and Rabagliati, H. (under review). Learning dimensions of meaning: Children's acquisition of *but*.

[4] Kail, M., & Weissenborn, J. (1984). A developmental cross-linguistic study of adversative connectives: French 'mais' and German 'aber/sondern'. *J. of Child Language*, *11*(1), 143- 158.

[5] Cain, K., & Nash, H. M. (2011). The influence of connectives on young readers' processing and comprehension of text. *Journal of Educational Psychology*, *103*(2), 429.

[6] Gillette, J., Gleitman, H., Gleitman, L. and Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition*, 73: 135–176.

[7] Gelman, S. A., Taylor, M. G., & Nguyen, S. P. (2004). Mother-child conversations about gender: Understanding the acquisition of essentialist beliefs: III. How children and mothers express gender essentialism. *Monographs of the Society for Research in Child Development*.