



Recovering discourse relations: Varying influence of discourse adverbials

Hannah Rohde, Anna Dickinson, Chris Clark, Annie Louis and Bonnie Webber *University of Edinburgh*

Abstract

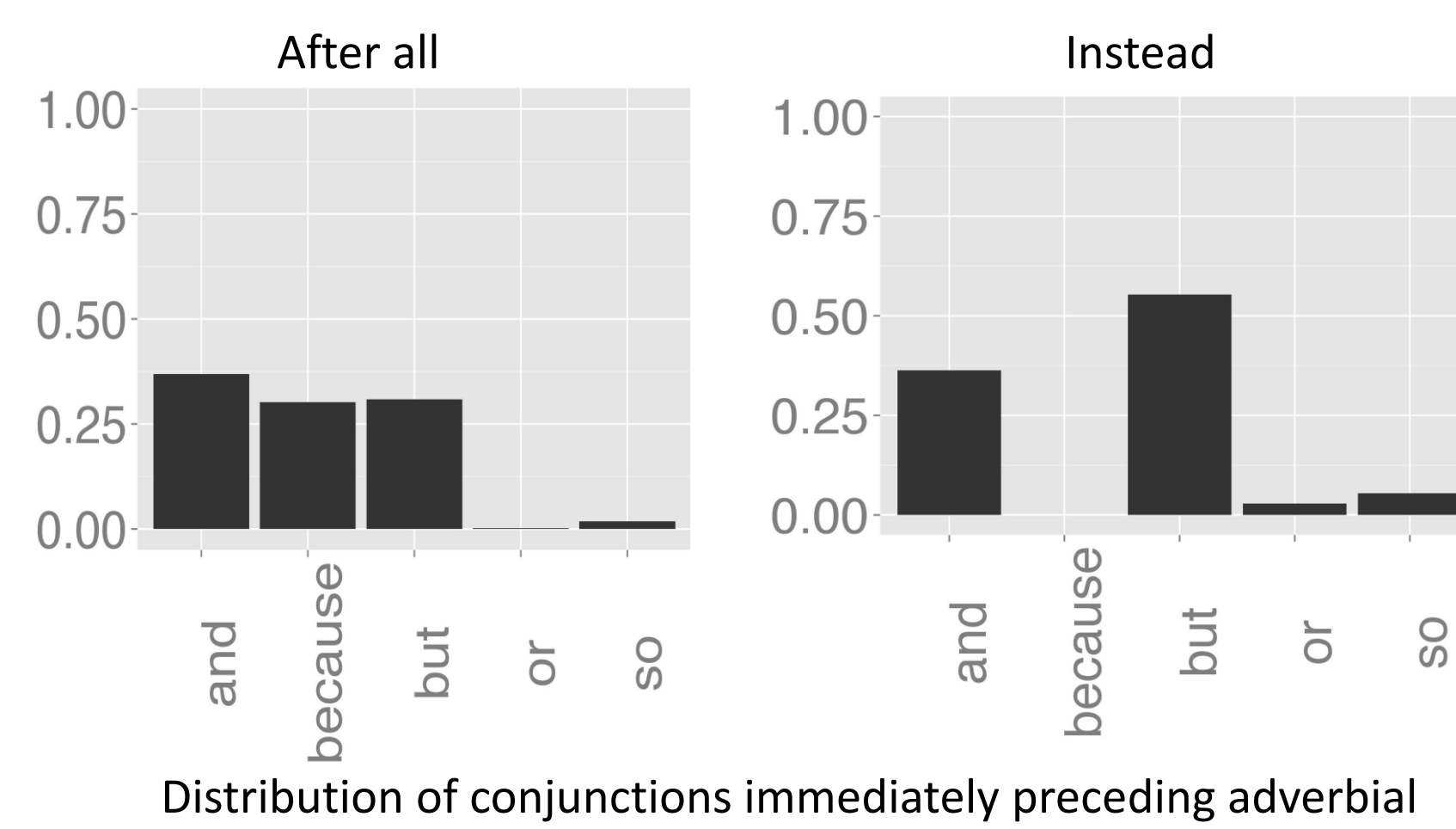
Discourse relations are a bridge between sentence-level semantics and discourse-level semantics. They can be signalled explicitly with discourse connectives or conveyed implicitly, to be inferred by a comprehender. The same discourse units can be related in more than one way, signalled by multiple connectives. But multiple connectives aren't necessary: Multiple relations can be conveyed even when only one connective is explicit. Here, we describe the initial phase in a larger experimental study aimed at answering two questions:

- 1) Given an explicit discourse adverbial, what discourse relation(s) do naive subjects take to be operative?
- 2) Can the relation be predicted on the basis of the explicit adverbial alone, or does it depend instead on other factors?

Google NGRAMs and distribution of conjunctions

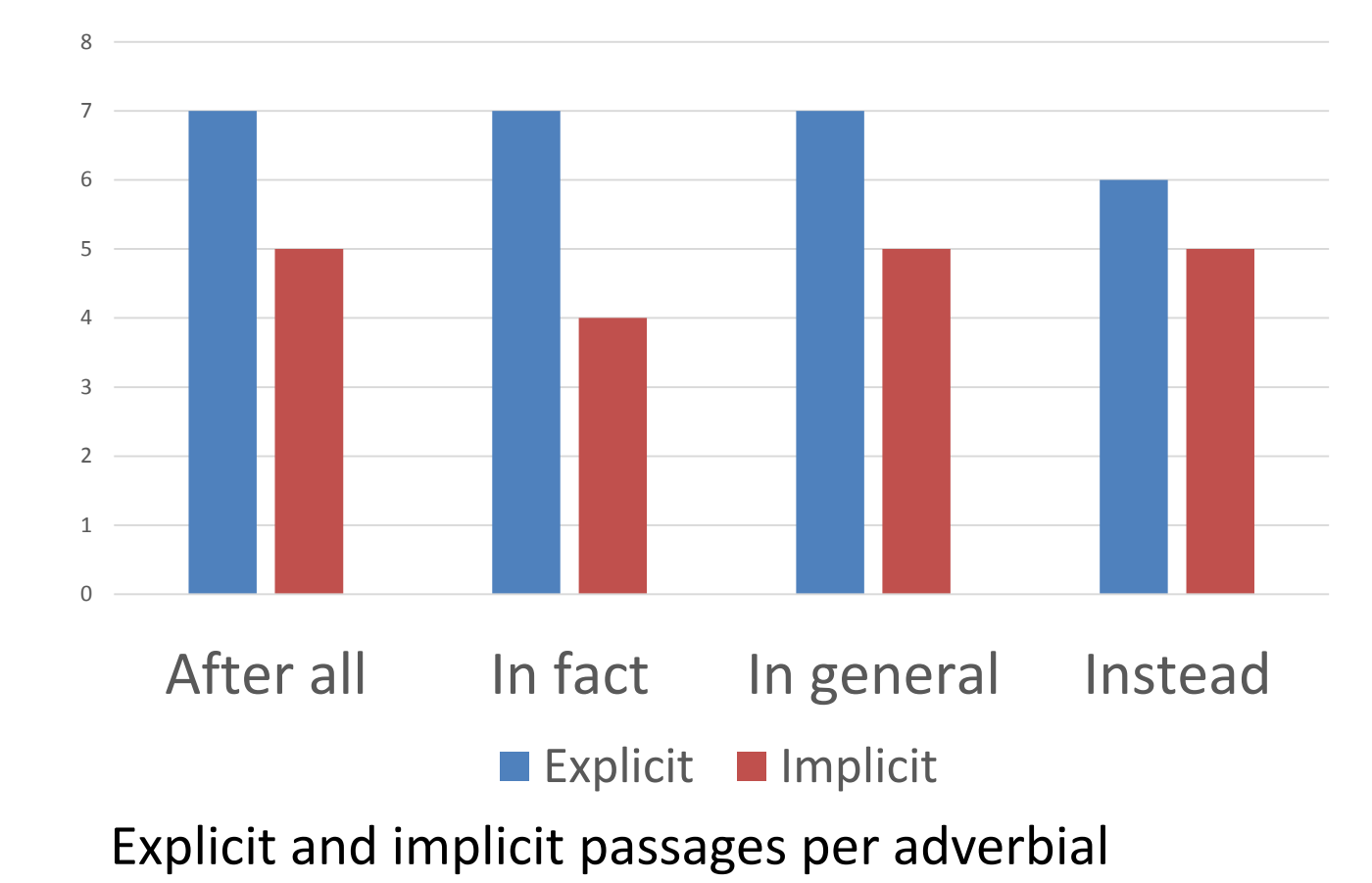
A survey of Google NGRAMs established overall frequency and preferred conjunction(s) of 73 adverbials.

The distribution of conjunctions is neither uniform for a given adverbial nor equivalent across adverbials.



Stimuli: Phase I

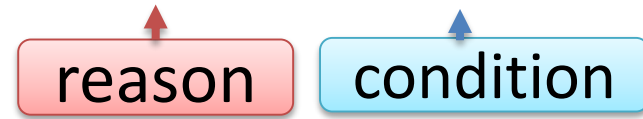
For our initial study, we used the four adverbials previously used by Jiang (2013). Stimuli were divided between *implicit* (no author-given conjunction) and *explicit* (author-given conjunction removed for experiment) passages.



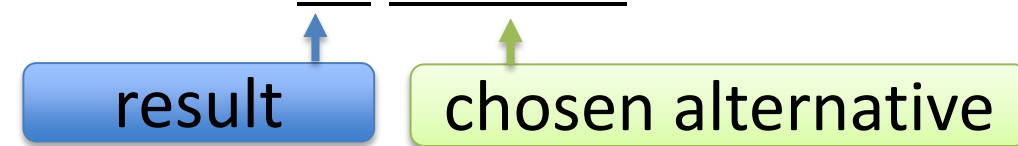
Discourse Relations

Previously, it was assumed that relations are conveyed implicitly when they are not signalled explicitly. But consider Ex. a-b, each with two explicit connectives conveying distinct relations:

a) Let's eat dinner now **because** otherwise we'll miss the film.



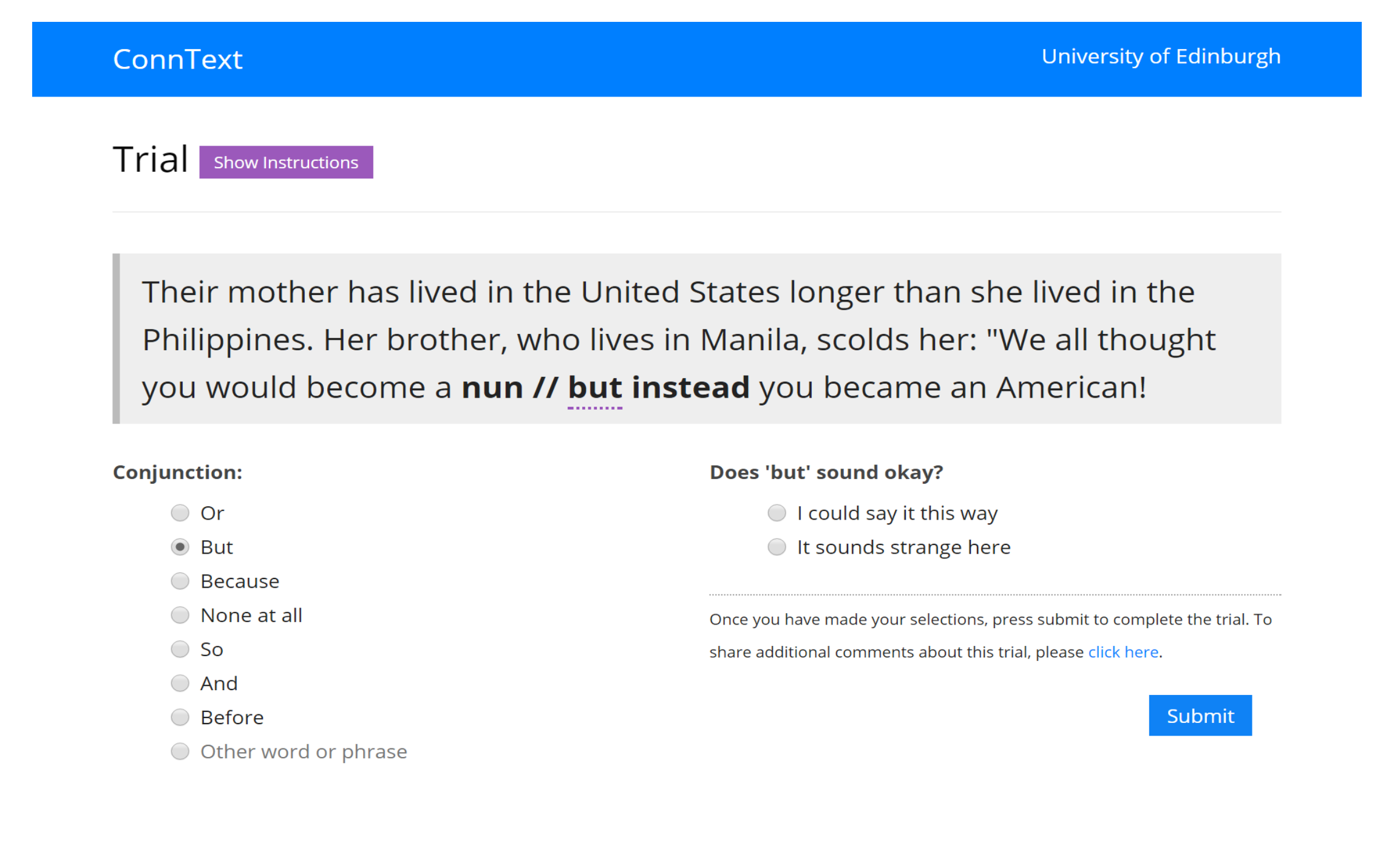
b) I can't walk 5 miles, **so** instead I'll take a taxi.



However, both relations may still be conveyed, even if only one is signalled explicitly, as in Ex. c-f:

- Let's eat dinner now. Otherwise we'll miss the film.
- I can't walk 5 miles. Instead I'll take a taxi.
- I can't walk 5 miles, so I'll take a taxi.
- Let's eat dinner now because we'll miss the film.

Interface



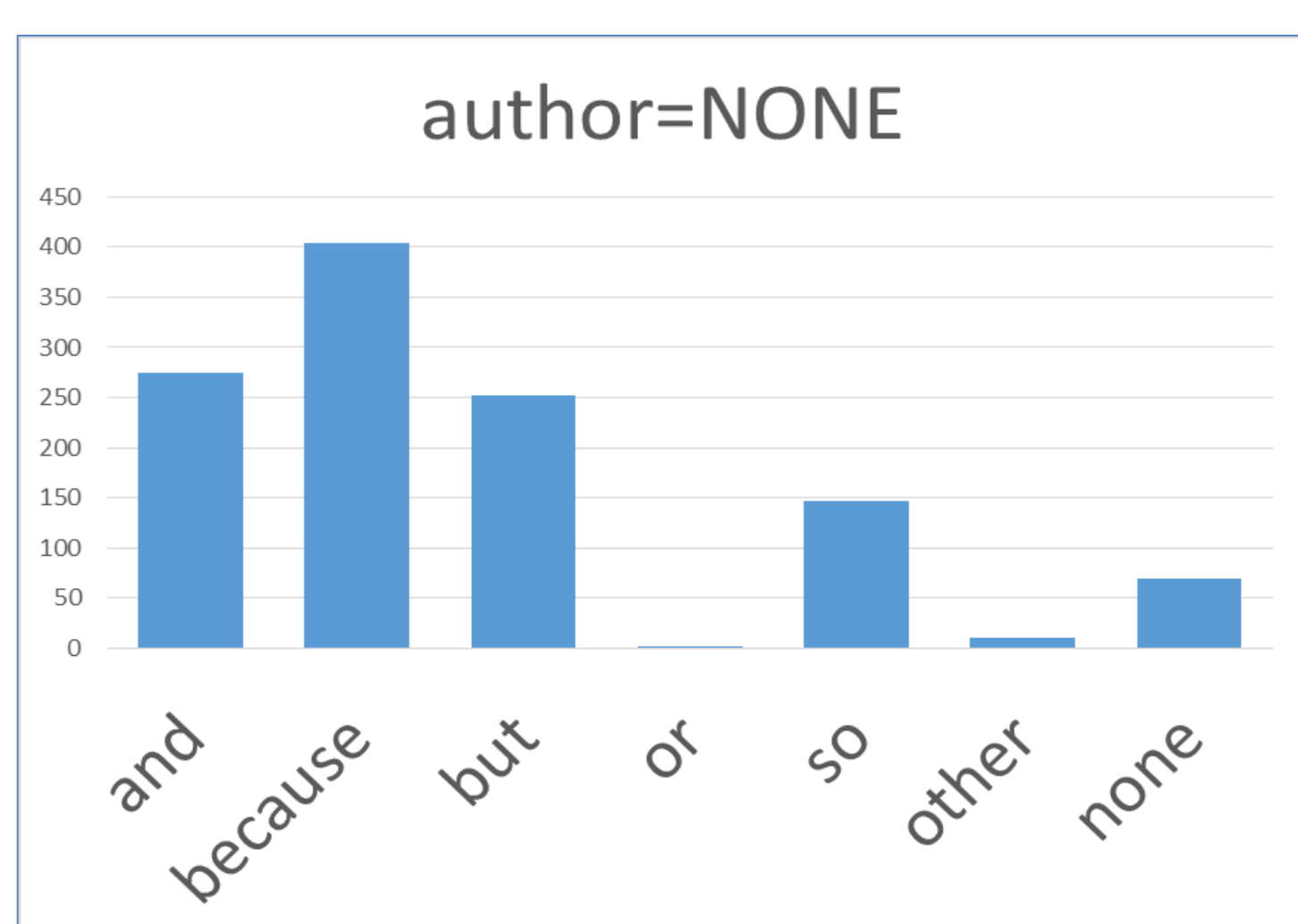
Results: author participant agreement

Confusion matrix for **explicit** passages.

		Author-selected conjunctions				
		AND	BECAUSE	BUT	OR	SO
Participant-selected conjunctions	And	189	14	81	5	33
	Because	60	105	60	2	9
	But	68	48	497	7	9
	Or	2	0	2	35	0
	So	125	1	25	2	56
	Other	3	1	8	2	0
	None	17	4	23	5	9

Participants and authors often agree on conjunction

Participant responses to **implicit** passages



Except when authors use no conjunction

Results: implicit preferences

Response distribution for implicit passages by adverbial:

	after all	in fact	in general	instead
and	50	87	118	20
because	245	35	86	38
but	16	83	50	103
or	1	0	0	0
so	4	3	21	119
other	5	3	2	0
none	26	20	13	10

after all favors because, whereas in fact, in general and instead show more variability

Results: explicit

Explicit response distributions. Participant responses in lower case versus author choice in CAPS.

	AND	BECAUSE	BUT	Total
And	18	6	30	54
Because	9	51	51	111
But	25	0	128	153
Or	0	0	0	0
So	0	0	3	3
Other	1	0	3	4
None	5	1	17	23

After all. Participants assigned because for author BUT and AND.

The adverbial may be biasing the inferred relation.

	AND	BUT	Total
And	16	1	17
Because	0	1	1
But	6	210	216
Or	0	2	2
So	92	17	109
Other	0	0	0
none	2	1	3

Instead. Author-selected AND got so responses.

Do participants attributing meaning select more specific so in preference to underspecified and?

	AND	BUT	SO	Total
and	102	23	4	129
because	50	4	8	62
but	36	85	3	124
or	2	0	0	2
so	33	1	41	75
other	2	0	0	2
none	7	3	2	12

In general. Higher convergence between participant- and author-selected conjunctions

In general not biasing inferred relation; responses depend on context or adjacent clause properties.

Conclusions

- 1) Naive subjects can infer an implicit conjunction alongside an explicit discourse adverbial, even for implicit passages where author used only explicit adverbial.
- 2) Subjects infer conjunctions reliably and systematically, depending on the adverb.
- 3) Discourse adverbials themselves are not indiscriminate with regard to the conjunction that they appear to favor.

Acknowledgments: This research has been supported by the Nuance Foundation.

The authors would also like to acknowledge advice and support from researchers on the Penn Discourse TreeBank.