Crowd-Sourcing Concurrent Relations

Anna Dickinson, Hannah Rohde, Annie Louis, Christopher N. L. Clark & Bonnie Webber

University of Edinburgh

Signalling discourse relations

▶ Conjunctions

> **"We've started trying just about anything to keep sales moving in the stores," says Kim Renk, a Swank vice president." <u>But</u> there are limits."** [WSJ]

▶ Adverbials

> **They both called it a "welcome home" gathering. <u>Nevertheless</u>, an ANC rally by any other name is still an ANC rally.** [WSJ]

This talk: Cases in which more than one relation can hold

Concurrent relations

In school he was assigned regular essays to write, **but instead** he wrote poetry. [NYTimes 2002]

Mentioning the word sex on daytime television was taboo at the time, she recalled, **so instead** she simply held up the book cover to the camera. [NYT 2001]

So many people have come that Jeff LeDuff, the police chief here, has said that city officials will no longer be able to transport and place evacuees within the city **and instead** they will be sent to adjacent regions. [NYT 2005]

Humans have lost some stereotypic calls **because instead** they can shout *Danger* or *Leopard* or *Watch out*, much more efficiently than calls [Bouchard, 2013]

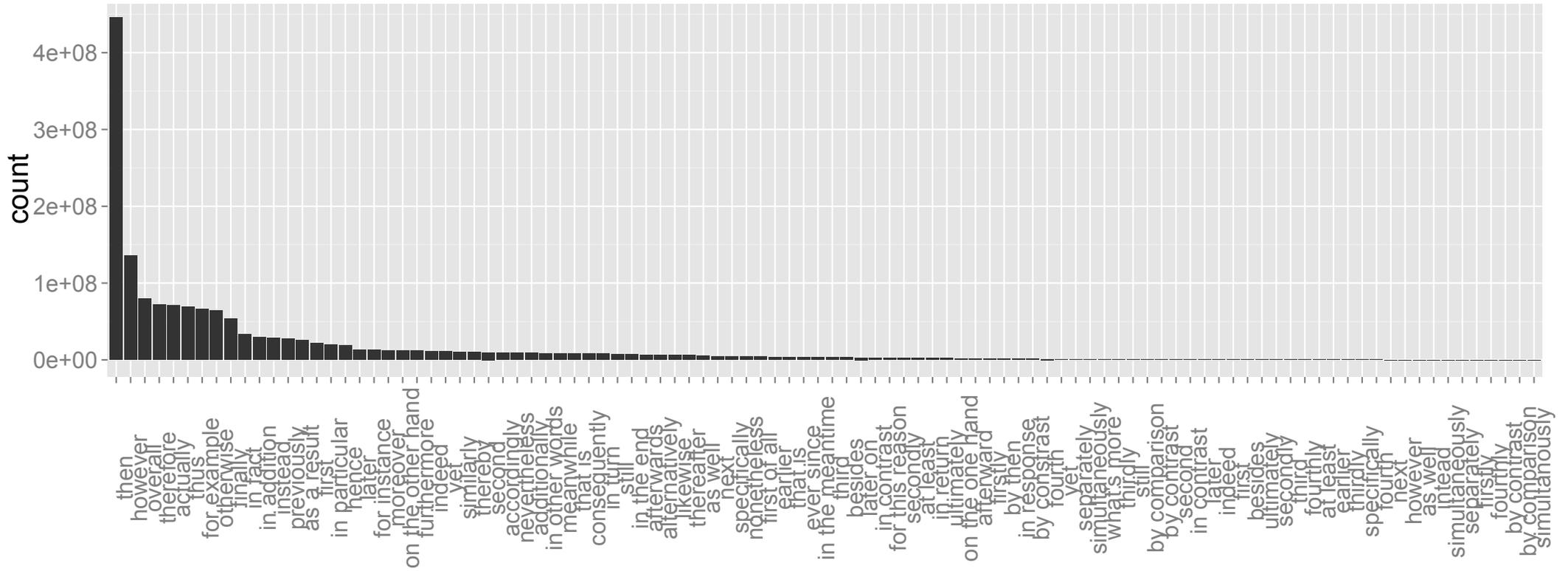–> Conjunction and adverbial can signal distinct relations

Motivations

▶ A lexically-based system like PDTB calls attention to the existence of multiple DRDs (see also Webber et al., 1999)

▶ These cases do not involve relations at different levels.

▶ Existing DRD theories do not predict such cases

▶ A better understanding of concurrent relations is relevant to

  ▶ Corpus annotation

  ▶ Psycholinguistics

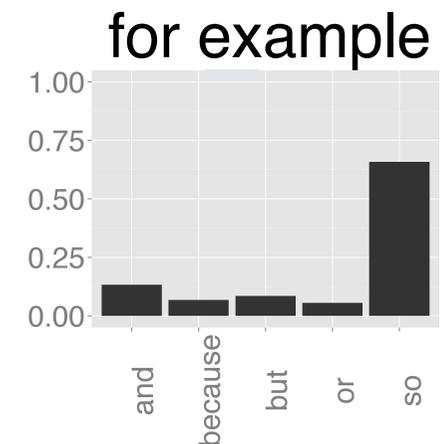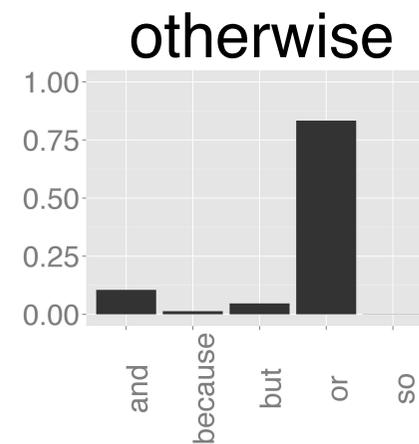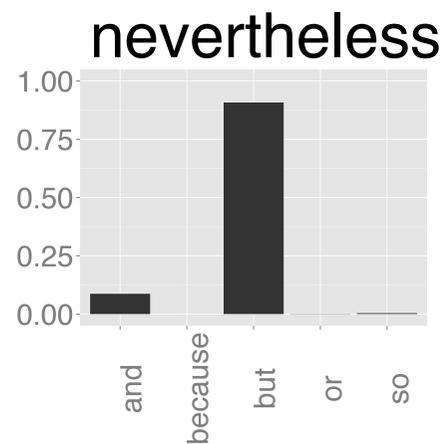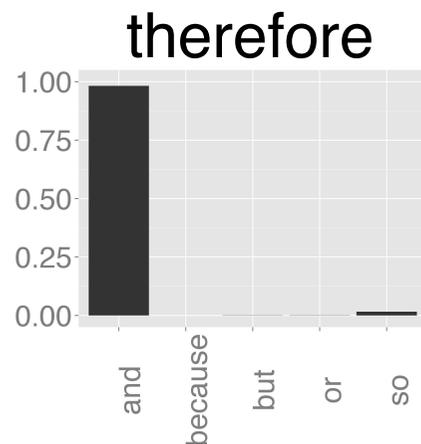 ▶ Language technologies

How to assess concurrent relations?

▸ Google Ngrams

    ▸ Frequency of discourse adverbials

    ▸ Frequency of adverbials accompanied by conjunction

    ▸ Possibility of same adverbial with different conjunctions

▸ Annotator judgments, collected via a web interface
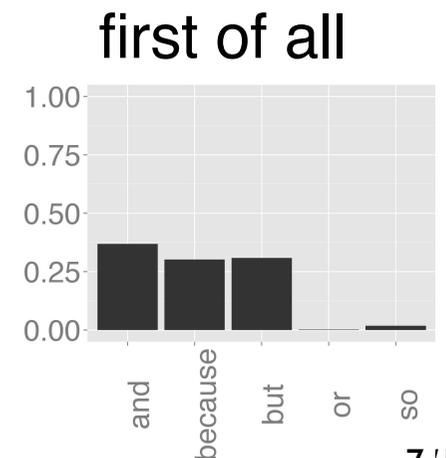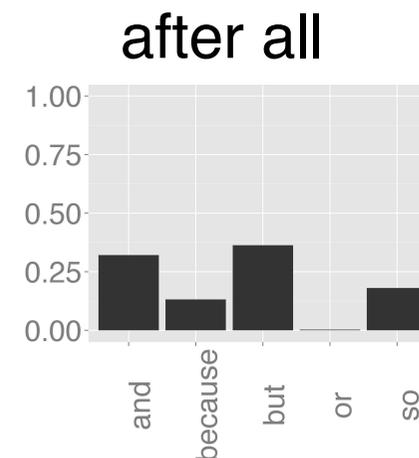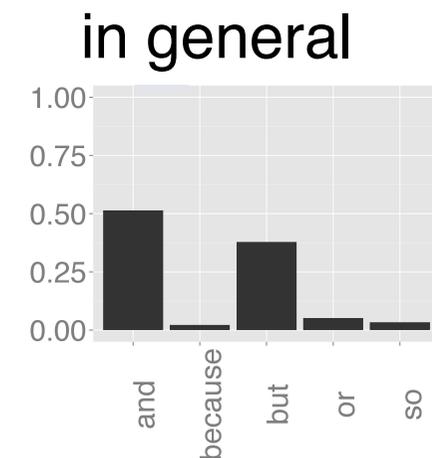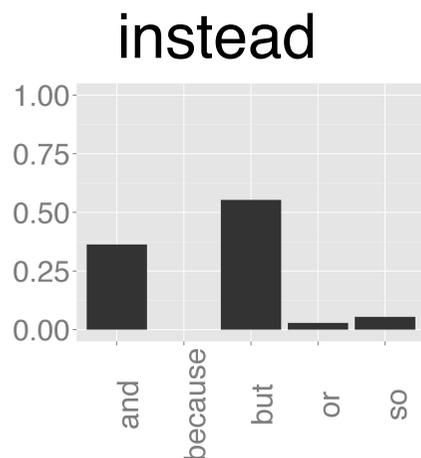[pilot results]

Frequency (Google Ngrams)

Adverbials and the conjunctions they prefer

▶ All adverbials occur most frequently alone.

▶ But with conjunctions, many have clear preferences.

▶ Whereas others have a broader distribution

What about bare adverbials?

> **Mentioning the word sex on daytime television was taboo at the time, she recalled, _____ instead she simply held up the book cover to the camera.**

- ▶ Jiang 2013: 4 adverbials

  - ▶ 80 passages, each viewed by 52 judges

  - ▶ Task: select a conjunction to fill a blank

- ▶ Jiang results: Adverbials differ from each other.

  - ▶ "After all" has an overall preference for "because"

  - ▶ "Instead" varies passage-by-passage

- ▶ No adverbial favored no conjunction.

- ▶ Shows that even without an explicit conjunction, two separate senses can be concurrently conveyed.

ConnText project

▸ Funded by small grant from Nuance Foundation

▸ **Goal:** Establish which discourse adverbials can co-occur with conjunctions and which conjunctions each adverbial licenses and favors

▸ **Pilot study:** 20 adverbials, 3 judges, 895 unique passages from the NYTimes Annotated Corpus (Sandhaus, 2008) and COCA (Davies, 2008)

ConnText

University of Edinburgh

Word Selection Show Instructions

> Traditionally, men do not do such **work //** _____ **instead** they stand around idly, waiting for something to fetch their attention away from the melancholia.

* Conjunction:

- ○ So
- ○ Before
- ○ And
- ○ Because
- ○ But
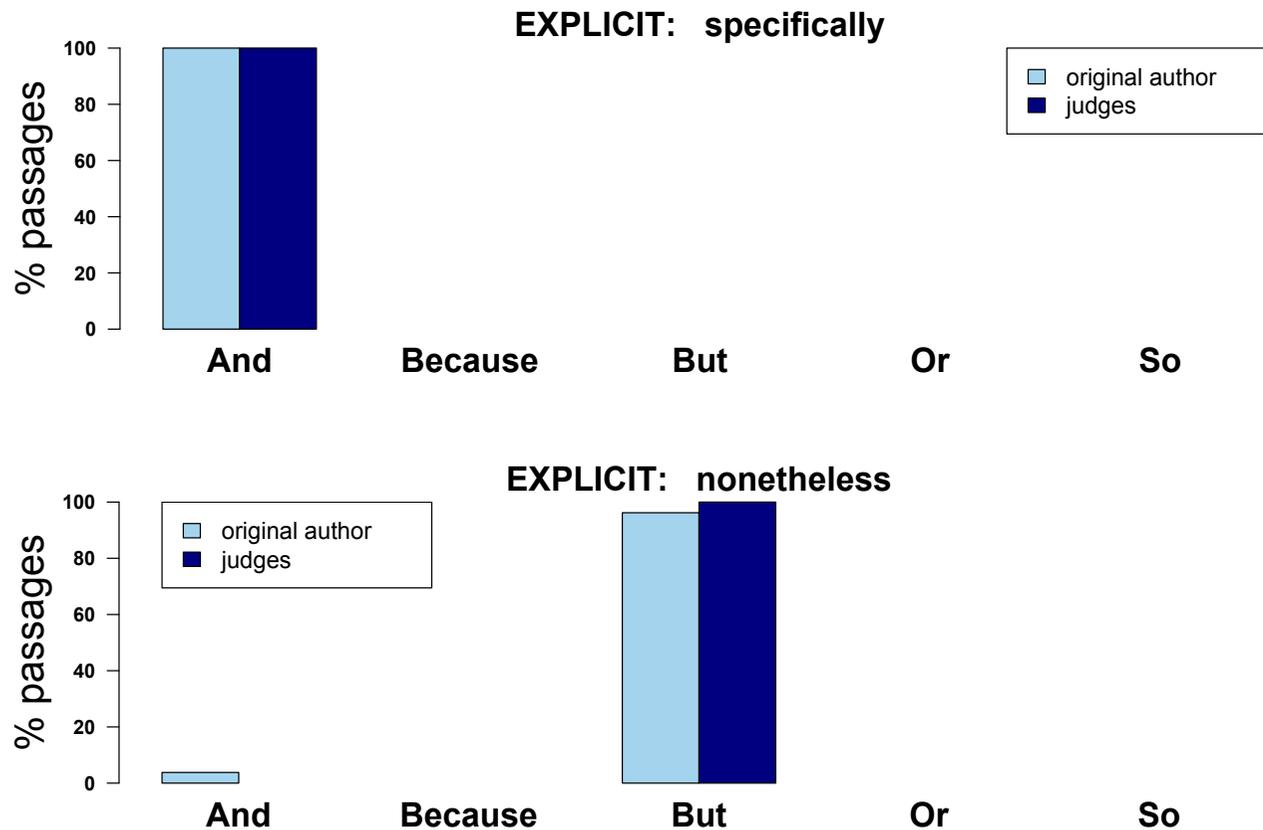- ○ Or
- ○ Other
- ○ None at all

Comments:

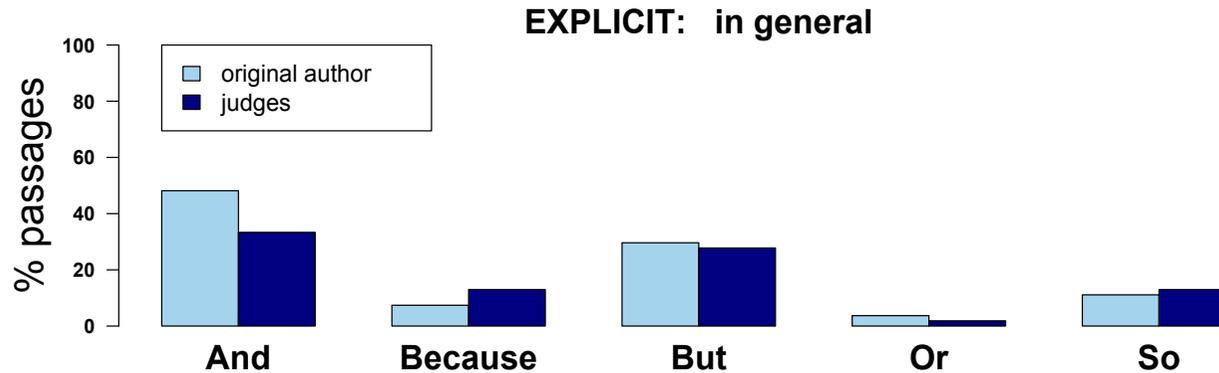*(Optional) Please share any comments you have about this trial*

Submit

Results: Explicits

▶ Dataset = 1507; Agreement with author: 19% ~ 100%
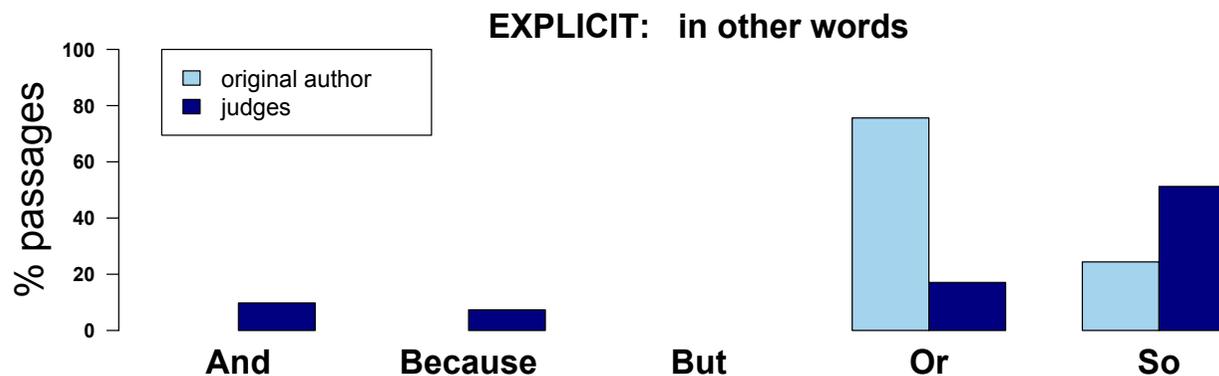
▶ Cases of strong agreement on single conjunction

Explicits

▶ Strong agreement on different conjunctions

EXPLICIT: in general

▶ Strong **dis**agreement with author

EXPLICIT: in other words

Example: disagreement among judges

'Ms. Morissette went through bouts of depression as the tour stretched on. ''Schedule-wise, my health and my peace of mind wasn't a priority,'' she said. ''There had been this dissonance in the midst of all the external success _____ **on the one hand**, I was expected to be overjoyed by it, and at the same time I was disillusioned by it.'' [NYTimes 1998]

[original: "because"]
[judges: "because", "and", "so"]

Results: Implicits

▶ Dataset=780 judgments

▶ Agreement across 3 judges: 64% ~ 90%

Some software automatically moves the pointer to the default button on a pop-up window _____ otherwise the user must move the pointer to that button or hit Enter (using the left hand or letting go of the mouse).

[original: none]
[judges: "but", "because", "none"]

Sources of judgment variation

▶ Interpretation of instructions (sense vs style)

"Your job is to make explicit the meaning that links the adjacent text spans. You must make a choice even if the insertion leads to an awkward or lengthy sentence, as long as you think the word brings out the meaning that links the two spans."

▶ Reduced context as basis for judgment

▶ Specificity of conjunctions (**and**/**so**)

▶ Idiosyncrasies (**or otherwise**/**because otherwise**)

You got to be nice to them _____ otherwise, they're not going to be nice to you. [COCA 1991]

[original: "or"]
[judges: "because", "because", "because"]

Summary (& open questions)

▶ In contrast to models of DRD usage that assume that each pair of clauses = one relationship, discourse adverbials:

   ▶ Do not always act alone (**because after all**)

   ▶ Need not license only one additional relation (**but**/**so**/**and**/**because instead**)

   ▶ Differ in the way they combine with possible conjunctions (single preference vs context-driven flexibility)

▶ Categories of discourse adverbials?

▶ Crowdsourcing as discovery tool?

▶ Other languages?

Thanks!