Theoretical and empirical approaches to cleft constructions

Day 4: Clefts and discourse expectations

ESSLLI 2024, Leuven

Jun Chen & Swantje Tönnis

University of Stuttgart

August 1, 2024

Recap of day 3

Destruel and Velleman (2014)

- **Hypothesis:** Clefts are more acceptable the more strongly they conflict with the interlocutors' expectations.
- Next to expectations about the world, they also introduced expectations about where the discourse is going.
- They manipulated discourse expectations by manipulating at-issueness of the addressed proposition.
- Their results did not show a significant effect of at-issueness for the acceptability of clefts.

Today:

Formalize discourse expectations and test the predictions

Roadmap for day 4

1. Some more data
2. Background – Discourse expectations
3. Experiments (Tönnis and Tonhauser, 2022)
   - 3.1 Experiment 1 – Expectedness rating
   - 3.2 Experiment 2 – Preference rating
4. Discussion
5. Open issues

"Pizza puzzle" (Velleman et al., 2012, p. 449)

- In English and German, a cleft is less acceptable as an immediate answer to a question: ((1) is acceptable in French.)

(1) A: What did Mary eat?
B:?It was a Pizza that Mary ate.

- It improves with some intervening utterances:

(2) A: What did Mary eat?
C: I thought she said she was gonna get a pizza, but I might be wrong.
D: And did she also order a salad?
B: Guys, I was there. And C's right–it was a Pizza that Mary ate.

- Velleman et al. (2012) argue that clefts terminate the inquiry, which explains (2) but struggles with (1).

Peter puzzle

- In German, the cleft is almost always replaceable by the simpler canonical sentence. (3) is an exception (translated from German):

(3) *Yesterday at the party, Lena talked to some guy. They laughed a lot and agreed to meet again the next evening. Then, Lena went home happily.*

a. It was Peter she talked to.
b. ?She talked to Peter.

- Why is this? The reason is not just the number of intervening sentences (see example next slide).

Peter Puzzle II (Tönnis, 2021, p. 261–262)

- What makes the cleft in (4) less acceptable than previously?

(4) *Yesterday at the party, Lena talked to some guy. They laughed a lot and agreed to meet again the next evening.* ***Then, Lena even told him a secret.***

a. ?It was Peter she talked to.
b. ?She talked to Peter.

- And what fixes this in (5)?

(5) *Yesterday at the party, Lena talked to some guy. They laughed a lot and agreed to meet again the next evening.* ***Lena even told him a secret. Then, Lena went home happily.***

a. It was Peter she talked to.
b. ?She talked to Peter.

Solution

Clefts in German must represent a relatively less expected discourse move. (Tönnis, 2021; Destruel and Velleman, 2014)

- Conflict with expectations about where the discourse is going
- Expectations about the world (Destruel and Velleman, 2014) do not play a role in our examples:
  It doesn't have to be unexpected that Lena talked to Peter instead of anybody else.

Before we go through the examples again, let's make discourse expectations more explicit.

Discourse model by Roberts (2012)

- Roberts (2012) argues that discourse is a cooperative game with the ultimate goal to answer "What is the world like?".
- Interlocutors agree on answering subquestions to reach the goal step by step.
- Every sentence addresses an (implicit) question, which Roberts (2012) calls **Questions under discussion** (QUDs)
- QUDs are organized on a stack.
- When interlocutors agree on a question, it is added on top of the stack.
- The top-most question has to be addressed first.
- This mechanism determines what is a relevant discourse move, i.e., which QUD may be asked or answered.

Adding discourse expectations to the model

Tönnis (2021)

- Interlocutors have expectations about which questions are likely to be addressed (answerd or asked) in the next sentence.
- These expectations can be represented by a probability distribution over questions given a certain context (Kehler and Rohde, 2017).

(6) Yesterday I had a great dinner in Leuven.

		probability
a.	What did you have for dinner?	0.5
b.	What happened then?	0.3
c.	How did you end up at the dinner?	0.15
d.	How was the weather in Leuven?	0.05

- This replaces the QUD stack.
- A relevant discourse move must address a question that exceeds a certain probability threshold.

Formalization of expectedness

How are the question expectedness probabilities determined?

- Expectedness as a **function**, called f_e:
  - Input of f_e: A context and a question
  - Output of f_e: An expectedness value (EV) (probability)
- f_e is defined for all possible questions and for all possible contexts
- Given a context, the EVs of all questions sum up to 1 (probability distribution)
- Example
  - Context C: *Lena told Nina a secret.*
  - Question q_1: *What was the secret?*
  - Question q_2: *How did Nina like that?*

  Then, $f_e(C, q_1) = 0{,}6$ and $f_e(C, q_2) = 0{,}3$ (example values)

Recursion

- f_e is defined **recursively**, i.e. the output of $f_e(C)(q)$ again uses f_e.
- f_e incrementally evaluates the effect of the **most recent discourse move (DM)**, i.e., a question or an assertion.
- If the question is more expected after the newest DM, it adds some positive constant to the previous value.
- If the question is less expected after the newest DM, it subtracts some positive constant from the previous value.
- Consider the question q: *What was the secret?*, and the DM of asserting φ:

(7) Peter met Lena yesterday. (C) $\underbrace{\text{She told her a secret.}}_{\varphi}$ (C+φ)

- Then, $f_e(C + \varphi)(q) = f_e(C)(q) + \alpha$,

  where $\alpha > 0$ because φ makes q more expected.

Normalization

- For any given context C, $\sum_{x \in \mathcal{Q}} f_e(C)(x) = 1$,
  i.e., the expectedness values of all possible questions add up to 1 for a given context (probability distribution).
- This condition will capture the interaction of EVs of different questions, e.g., when the expectedness of one question rises, the EVs of the other questions are automatically lowered by normalization.

First step of recursion

- Assumption: At the beginning of a conversation we have certain prior expectations about which questions are likely to be addressed in a conversation.
- In the first step, f_e just assign each question its prior expectedness value (prior probability).
- In an out-of-the-blue context, question (8-a) will have a higher prior EV than question (8-b).

(8) a. How is the weather?

b. Will aliens take over the world?

Further steps of recursion

Take context C: *I met Lena yesterday.*

How can a discourse move φ affect question expectedness?

- **Potential question (Onea, 2016):** $\varphi =$ *"She told me a secret"*
  $\rightarrow$ EV of the PQ *"What is the secret?"* increases in $C + \varphi$.
- **Answer:** $\varphi =$ *"She told me a secret"*
  $\rightarrow$ EV of question *"What did Lena do?"* is reduced.
- **Subquestion:** $\varphi =$ *"What did she do?"*
  $\rightarrow$ EV of subquestion *"What did Lena do first?"* increases.
- **No change:** $\varphi =$ *"Lena told me a secret"*
  $\rightarrow$ EV of question *"How is the weather in Leuven?"* does not change.
- and more ...

EVs are adjusted for every q in $C + \varphi$ and then normalized.

simplified version of Tönnis (2021, 273)

When is a cleft acceptable?

Hypotheses

1. A German *es*-cleft addresses a question that is relatively less expected at that point in the discourse.

2. Those questions that are relatively more expected are preferably addressed with a canonical sentence.

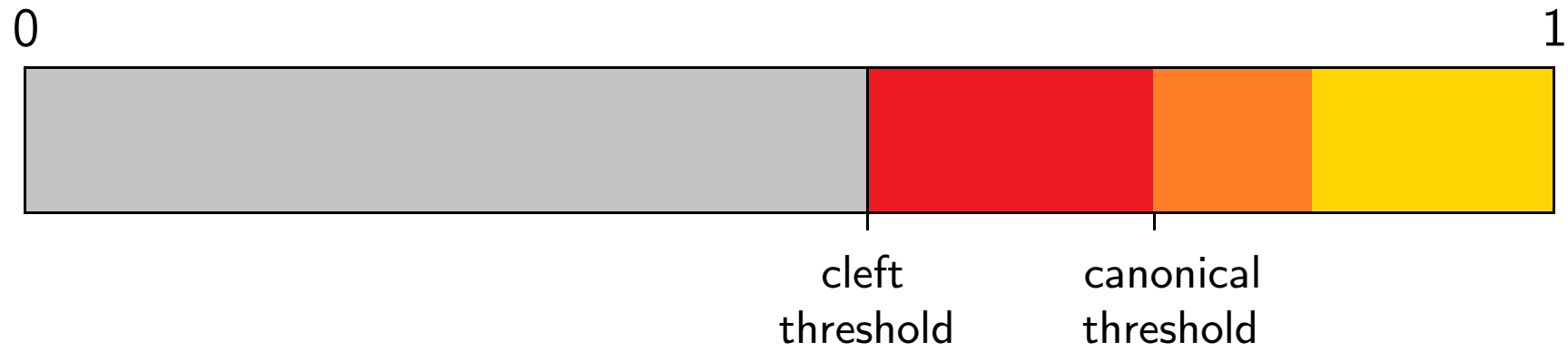
3. Those questions that are neither particularly expected nor particularly unexpected can equally well be addressed by a cleft or a canonical sentence.

(see Tönnis, 2021, 14)

Hypotheses – illustrated

- There are different **expectedness thresholds** for clefts and canonical sentences.

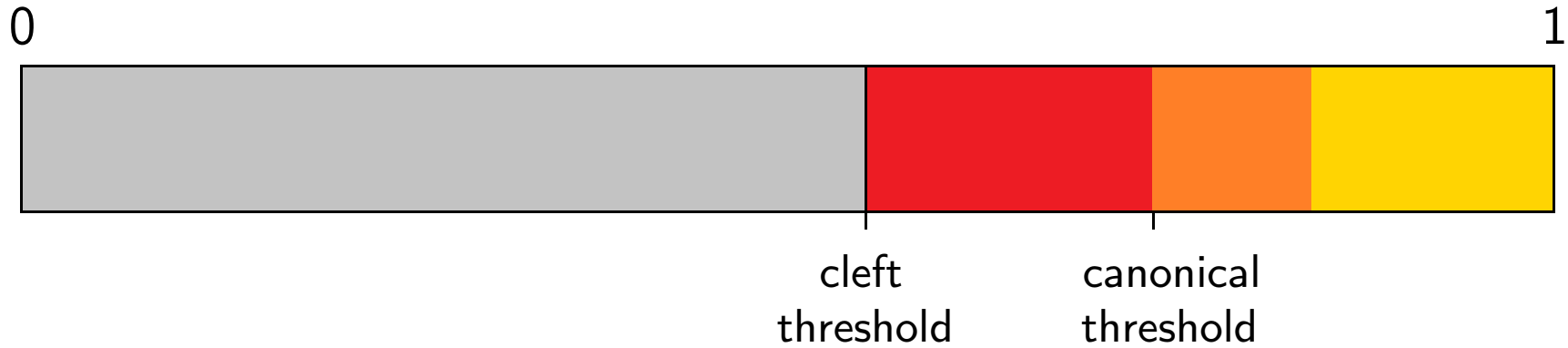
Possible expectedness values for questions



- If the addressed question falls above the respective threshold, the discourse move is acceptable.
- The exact threshold values must be determined empirically.

Predictions

Possible expectedness values for questions



- Yellow area: canonical sentence preferred over cleft
- Orange area: no preference between cleft and canonical sentence
- Red area: cleft preferred over canonical sentence

Let's go through the examples again!

Solving the "Pizza puzzle" (Velleman et al., 2012)

(1) A: What did Mary eat?
B:?It was a Pizza that Mary ate.

(2) A: What did Mary eat?
C: I thought she said she was gonna get a pizza, but I might be wrong.
D: And did she also order a salad?
B: Guys, I was there. And C's right–it was a Pizza that Mary ate.

Solving the Peter Puzzle

(3) *Yesterday at the party, Lena talked to some guy. They laughed a lot and agreed to meet again the next evening. Then, Lena went home happily.*

a. It was Peter she talked to.
b. ?She talked to Peter.

- Both (3-a) and (3-b) address Q1 (*Who did Lena talk to?*).
- Q1 is relatively unexpected in the context:
  - Q1 was raised by first sentence, and addressed after 2 intervening sentences.
  - There is a topic shift in the last sentence.
  - Partial answer to Q1 is given already.

$\rightsquigarrow$ Cleft is preferred over canonical sentence.

Solving Peter puzzle II

Tönnis (2021)

(4) *Yesterday at the party, Lena talked to some guy. They laughed a lot and agreed to meet again the next evening.* ***Then, Lena even told him a secret.***

a. ?It was Peter she talked to.
b. ?She talked to Peter.

(5) *Yesterday at the party, Lena talked to some guy. They laughed a lot and agreed to meet again the next evening.* ***Lena even told him a secret. Then, Lena went home happily.***

a. It was Peter she talked to.
b. ?She talked to Peter.

Let's test these predictions!

Two experiments by Tönnis and Tonhauser (2022)

Experiment 1 (80 participants)

- Identify two different contexts for a target question Q1
  1. Context makes Q1 relatively expected.
  2. Context makes Q1 relatively less expected.

Hypothesis

The expectedness of a question to be addressed in the next discourse move decreases with a higher number of intervening sentences to the sentence which raised the question. (Tönnis, 2021)

(see also Kehler and Rohde, 2017; Onea, 2016)

- New methodology which is suitable to investigate less expected questions, and subtle differences in discourse expectations: **Relative expectedness rating of Q1**
- Experiment 2 can then test the preference between cleft and canonical sentence depending on expectedness of Q1.

Experiment 1 – Stimuli

- 16 context-question pairs in German, each in 2 conditions.
- Example stimuli translated from German:

(9) **Condition 1: no intervening sentences**
When Lilly joined breakfast the bag of rolls was already empty.

Prediction: Q1 *Who ate the last roll?* relatively expected

(10) **Condition 2: two intervening sentences**
When Lilly joined breakfast the bag of rolls was already empty. There weren't any croissants or toast either. Hence, she went to the bakery nearby.

Prediction: Q1 *Who ate the last roll?* relatively less expected

Experiment 1 – Design

Condition 1: no intervening sentences translated from German

When Lilly joined breakfast the bag of rolls was already empty. ...

What do you expect that the next sentence will be about? Rate the following proposals and adjust the slider accordingly.

- About who ate the last roll. Q1
- About what Lilly could have for breakfast instead. Q2
- About what Lilly bought at the bakery. Q3
- About what Lilly did next. Q+
- About how the weather is in Colombia. Q-

Experiment 1 – Design

Condition 2: two intervening sentences translated from German

When Lilly joined breakfast the bag of rolls was already empty. There weren't any croissants or toast either. Hence, she went to the bakery nearby. ...

What do you expect that the next sentence will be about? Rate the following proposals and adjust the slider accordingly.

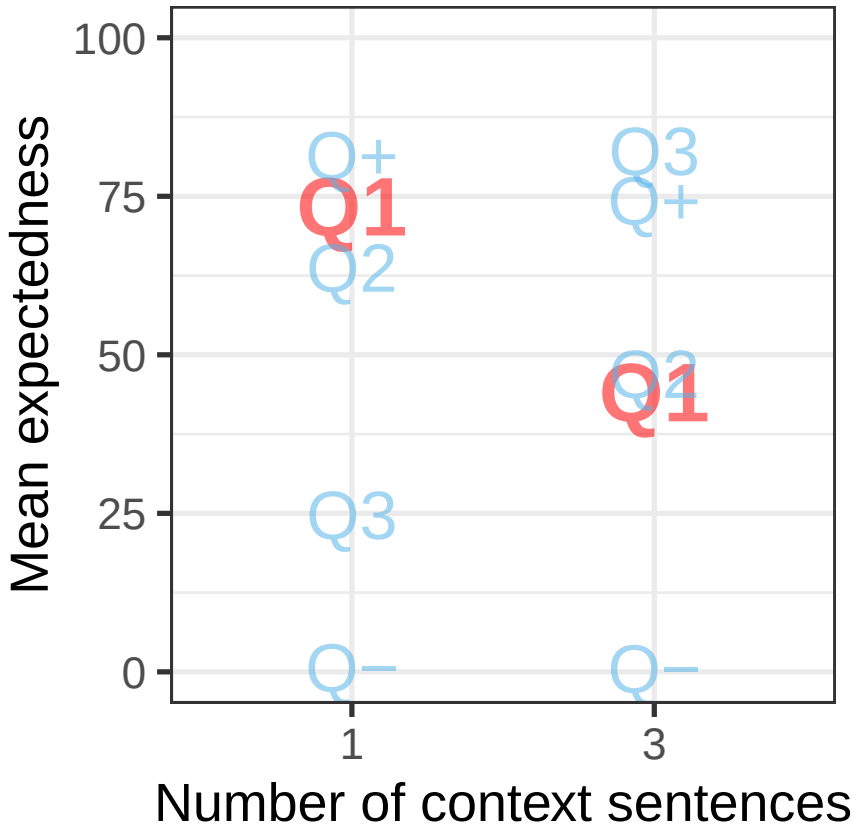
- About who ate the last roll. Q1
- About what Lilly could have for breakfast instead. Q2
- About what Lilly bought at the bakery. Q3
- About what Lilly did next. Q+
- About how the weather is in Colombia. Q-

Experiment 1 – Targets and baselines

- 16 context-question pairs (each with no and two intervening sentences)
- Target question: question raised by first sentence (*Who ate the last roll?*) (Q1)
- Each target question was presented with four other questions:
  - baselines: irrelevant question Q- (*How is the weather in Colombia?*) and *then*-question Q+ (*What did Lilly do then?*)
  - question raised by second sentence (*What could Lilly have for breakfast instead?*) (Q2)
  - question raised by the third sentence (*What did Lilly buy at the bakery?* (Q3)

(*When Lilly joined breakfast the bag of rolls was already empty. There weren't any croissants or toast either. Hence, she went to the bakery nearby.*)

Experiment 1 – Results



- **Significant difference!**
- Question Q1 is more expected with no intervening sentence (1-sentence context) than with two intervening sentences (3-sentence context).

- Linear mixed effects model with participant and item as random effects and a by participant slope (β = -29.3, SE = 2, t = -15, p $<$ .001).

Experiment 1 – Results by item

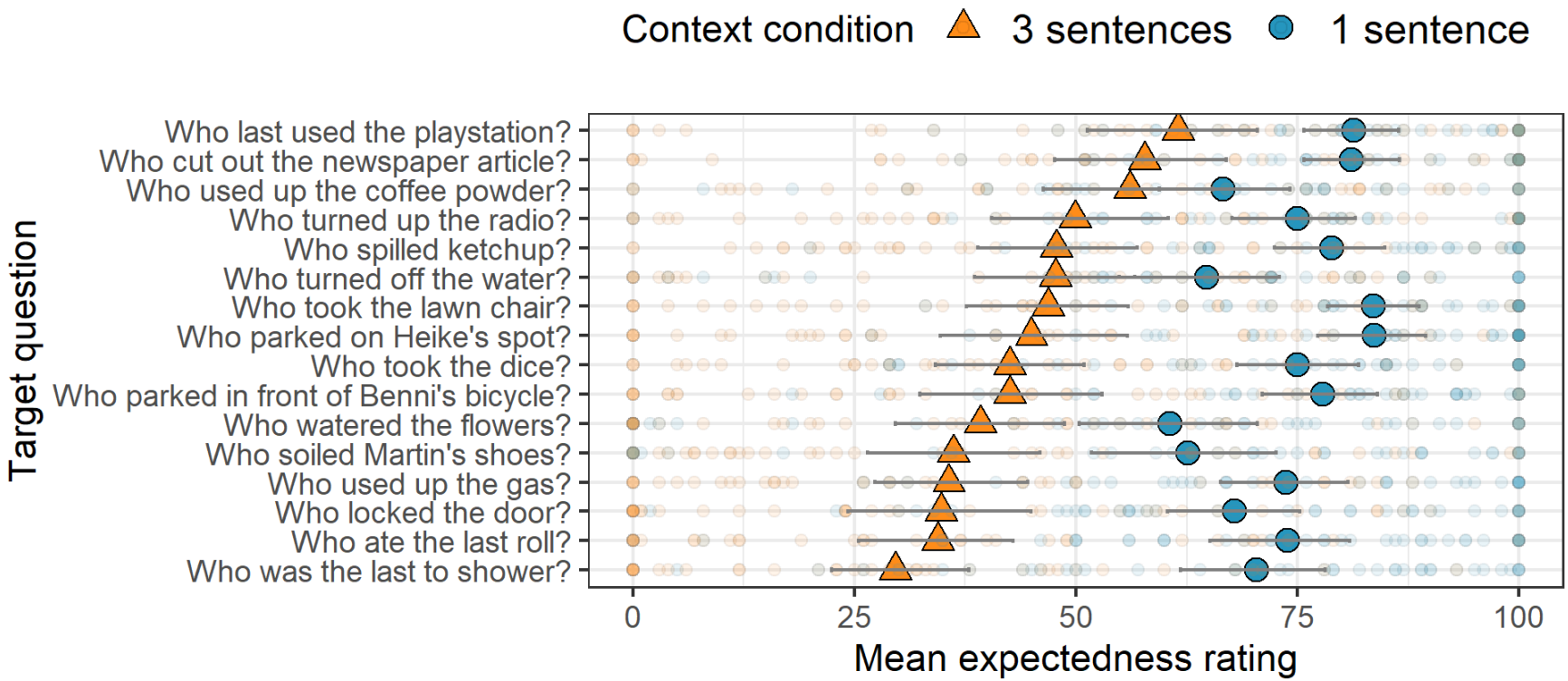


Figure: Mean expectedness of the target questions by context condition in the norming study. Error bars indicate 95% bootstrapped confidence intervals. Transparent dots indicate individual participants' ratings.

Experiment 2 (120 participants)

Hypothesis

German cleft sentences address relatively less expected questions while canonical sentences address relatively expected questions.

(Tönnis, 2021)

- **Method:** Preference rating between cleft and canonical sentence addressing respective Q1 in the 16 contexts from Experiment 1.
- **Context condition 1:** higher expectedness of Q1 (no intervening sentences)
- **Context condition 2:** lower expectedness of Q1 (2 intervening sentences)
- 4 targets and 8 fillers per participant
- **Prediction:** The preference for the cleft should be stronger in condition 2 than in condition 1.

Experiment 2 – Design

Condition 1: higher question expectedness translated from German

Read the text in the box. The next sentence of this text is illegible (marked by XXXX).

When Lilly joined breakfast the bag of rolls was already empty. **XXXX**

A. Benni ate the last roll.
B. It was Benni who ate the last roll.

Which of the sentences A or B would you prefer and how strongly? Adjust the slider accordingly.

(A much better) (both equally good) (B much better)

Experiment 2 – Design

Condition 2: lower question expectedness translated from German

Read the text in the box. The next sentence of this text is illegible (marked by XXXX).

When Lilly joined breakfast the bag of rolls was already empty. There weren't any croissants or toast either. Hence, she went to the bakery nearby. **XXXX**

A. Benni ate the last roll.
B. It was Benni who ate the last roll.

Which of the sentences A or B would you prefer and how strongly? Adjust the slider accordingly.

(A much better) (both equally good) (B much better)

Experiment 2 – Results

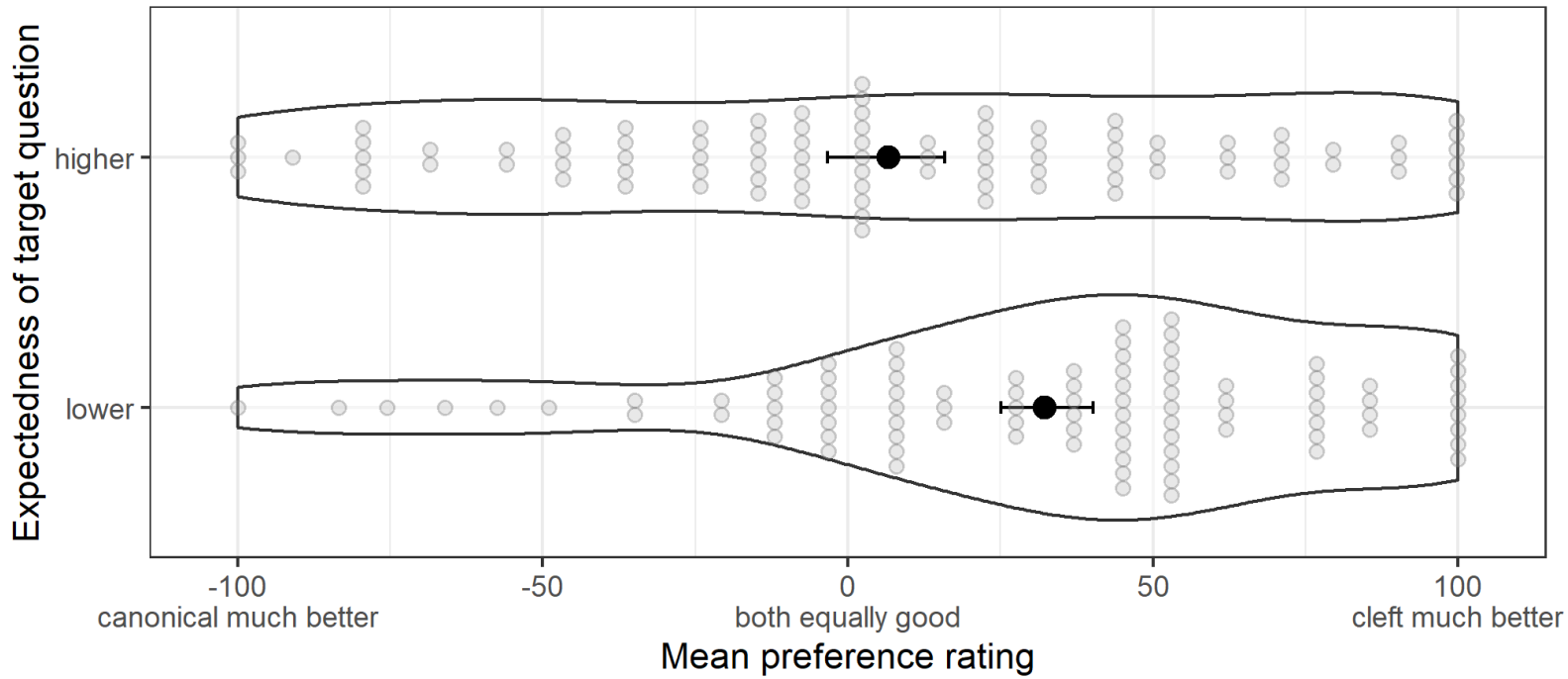


Figure: Mean preference rating with 95% confidence intervals by context condition.

- **Significant difference:** Cleft preference was stronger when measured expectedness of Q1 was lower.
- In the higher expectedness condition, there does not seem to be a clear preference.

Discussion I

- Tönnis' (2021) hypotheses are supported by our results.
  1. Question expectedness decreases with higher number of intervening sentences to the question-raising sentence. (Experiment 1)
  2. Clefts are more acceptable when they address a relatively unexpected question. (Experiment 2)
- However:
  - The contexts that were tested were very specific: identical discourse relations, tense, lexical verb classes...
  - That leads to less noise, but also to less insight: The effect could be due to a different reason which coincides with the expectedness level.
  - The cleft's sensitivity to discourse expectations might not be relevant for every use of the cleft.

Open issues

- Can other functions of the cleft be subsumed under its function/characteristics of addressing less expected questions? E.g., clefts establishing coherence or creating suspense?
- Can today's approach cover occurrences of clefts as corrections?
- Are expectations about the world also covered by today's approach or do we need to add something?

References I

Destruel, E. and L. Velleman (2014). Refining contrast: Empirical evidence from the English *it*-cleft. *Empirical Issues in Syntax and Semantics 10*, 197–214.

Kehler, A. and H. Rohde (2017). Evaluating an expectation-driven question-under-discussion model of discourse interpretation. *Discourse Processes 54*(3), 219–238.

Onea, E. (2016). *Potential Questions at the Semantics-Pragmatics Interface*, Volume 33 of *Current Research in the Semantics/Pragmatics Interface*. Leiden, Boston: Brill.

Roberts, C. (2012). Information Structure in Discourse: Towards an Integrated Formal Theory of Pragmatics. *Semantics and Pragmatics 5*(6), 1–69.

Tönnis, S. (2021). *German es-Clefts in Discourse. A Question-Based Analysis Involving Expectedness*. PhD thesis, Graz University.

References II

Tönnis, S. and J. Tonhauser (2022). German clefts address unexpected questions. In J. R. Starr, J. Kim, and B. Öney (Eds.), *Semantics and Linguistic Theory (SALT)*, Volume 32, pp. 661–684.

Velleman, L., D. I. Beaver, E. Destruel, D. Bumford, E. Onea, and L. Coppock (2012). *It*-clefts are IT (inquiry terminating) constructions. In *Proceedings of SALT*, Volume 22, pp. 441–460.