Syntactically-mismatched questions are easy to sluice if you know how

Till Poppels & Andrew Kehler, *UC San Diego*

UCSD

Ellipsis & Voice

Sluicing: a form of **ellipsis** that targets **clauses** under interrogative wh-phrases.

(1)
  a. Joe was murdered,
     but we don't know by who ~~he was murdered~~.
  b. Someone murdered Joe,
     but we don't know who ~~murdered him~~.
  c. # Joe was murdered,
     but we don't know who ~~murdered him~~.
  d. # Someone murdered Joe,
     but we don't know by who ~~he was murdered~~.

antecedent clause / ellipsis remnant / elided clause

It appears that voice-mismatched clauses cannot be elided.

This contrasts with cases of **VP-ellipsis** that tolerate voice mismatches:

(2) This problem should have been looked into, but nobody did ~~look into it~~. (Kehler, 2002)

Syntactic identity theories

(Chung, 2006; 2013; Merchant, 2013; Rudin, 2019)

Central claim: Ellipsis is grammatical if and only if the elided material is **syntactically identical** to an antecedent constituent in the linguistic context.

Merchant (2013): VP-ellipsis tolerates voice mismatches and sluicing doesn't because of the position of VoiceP.

XP   ⇒ Ø: voice mismatch *disallowed*
  VoiceP
  Voice   YP   ⇒ Ø: voice mismatch *allowed*
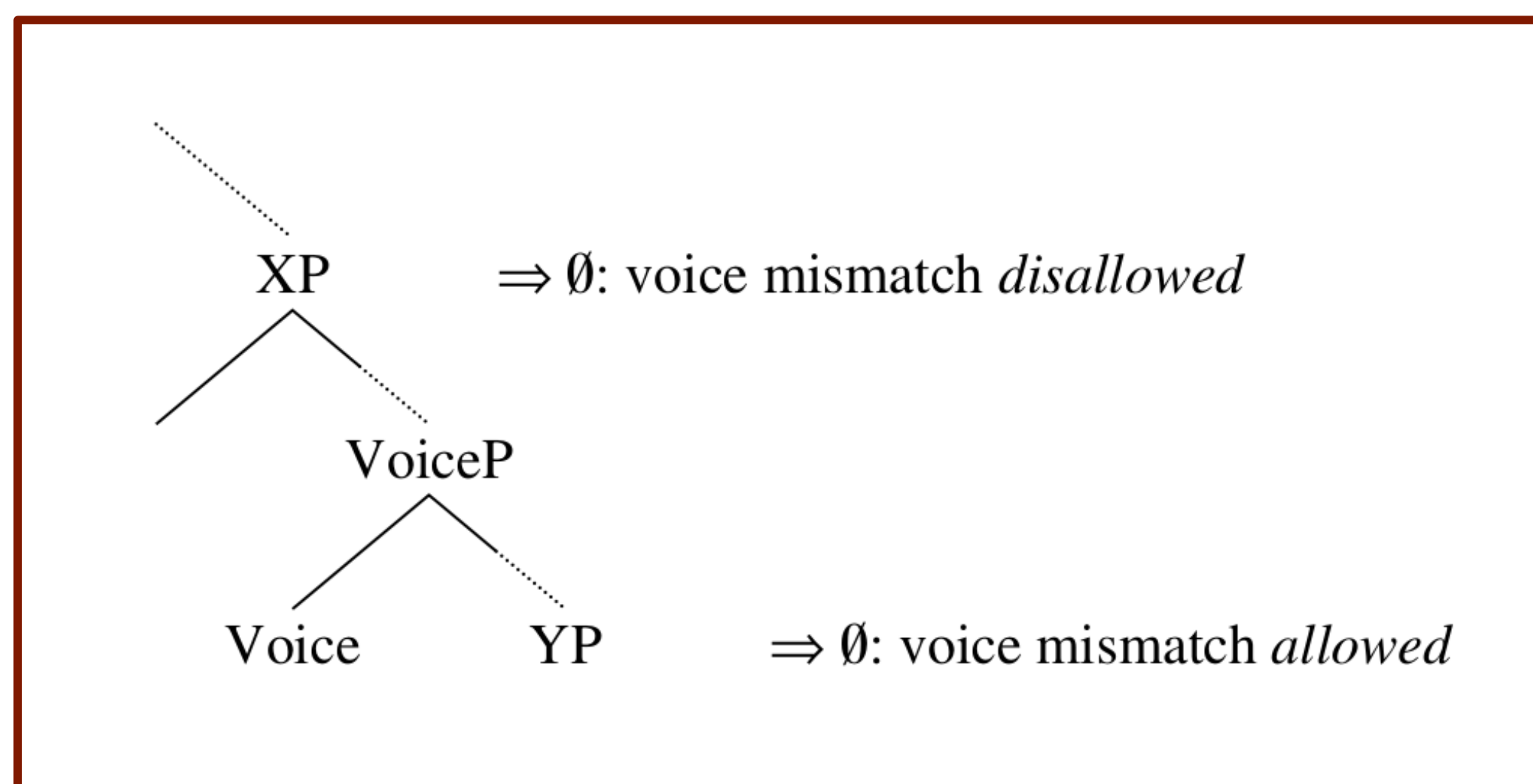
Figure 1. Schematic representation from Merchant (2013, Fig. 1). Prediction: Ellipsis below VoiceP tolerates voice mismatches, ellipsis above VoiceP is ungrammatical.

This analysis captures (1) and (2), but the picture is more complicated…

A closer look at mismatches

1. Voice-mismatched VP-ellipsis isn't categorically acceptable (Kehler, 2002; Kertz, 2013). For example, Kertz (2013) found that mismatches based on "tough movement" can be acceptable, as in (3a), but aren't always, as in (3b).

(3)
  a. Venomous snakes are easy to identify,
     and most experienced hikers can ~~identify them~~.
  b. # It's easy to identify venomous snakes,
     and poisonous plants are ~~easy to identify~~ as well.

Whatever is causing this difference, it cannot be explained by 'ellipsis scope' analysis shown in Fig. 1.

2. Sluicing permits other types of mismatches.

(4)
  a. I remember meeting him, but I don't remember when ~~I met him~~. (**finiteness mismatch**; Merchant, 2001, ex. 33)
  b. Your favorite plant is alive, but you can never know how long ~~it will be alive~~. (**tense mismatch**; Rudin, 2019, ex. 22)
  c. Customers should be given lower rates, but it's hard to see how ~~they could be given lower rates~~. (**modality mism.**; Rudin, 2019, ex. 23c)
  d. Either John didn't do an extra credit problem, or he didn't mark which one ~~he did~~. (**polarity mismatch**; Kroll, 2016, ex. 3)
  e. Can I get a few autographs? ---Sure, how many ~~do you want~~? (**verb mismatch**; Poppels & Kehler, 2019)

3. Voice-mismatched sluicing is hard to prove.
Since sluicing remnants are minimal compared to VP-ellipsis, it is **difficult to "force"** a voice mismatch:

(5) The toilet **needs** to be fixed but I don't know how…
  a. ...~~to fix it~~.   voice mismatch
  b. ...it **can** ~~be fixed~~.   no voice mismatch

Mismatch	Sluicing	VP-ellipsis
Finiteness	✔	✔
Tense	✔	(✔)
Modality	✔	(✔)
Polarity	✔	(✔)
Verb	✔	✔
Voice	?	✔

Table 1. Partial list of types of mismatches that are possible under sluicing and VP-ellipsis. Checkmarks in brackets indicate that mismatch occurs outside the elided material.

Research Question: does the unacceptability of (1c) and (1d) generalize to other argument-structure mismatches beyond voice?

Methods

Goal: Test acceptability of mismatched sluicing under *tough alternation*.

Participants: **41 native English speakers** from Amazon's Mechanical Turk.

Stimuli: 24 items like (6); **2x2x3 design**; 48 fillers (2:1 ratio).

(6)
  a. Brake fluid is easy to replace if you know {how|when|where}.
  b. It's easy to replace brake fluid if you know {how|when|where}.
  c. Brake fluid is easy to replace if you know {how|when|where} to replace it.
  d. It is easy to replace brake fluid if you know {how|when|where} to replace it.

Task: acceptability judgment task (1-5 Likert scale)

Brake fluid is pretty easy to replace if you know when.

(unacceptable) [1] [2] [3] [4] [5] *(fully acceptable)*

Use number keys or click boxes to answer.

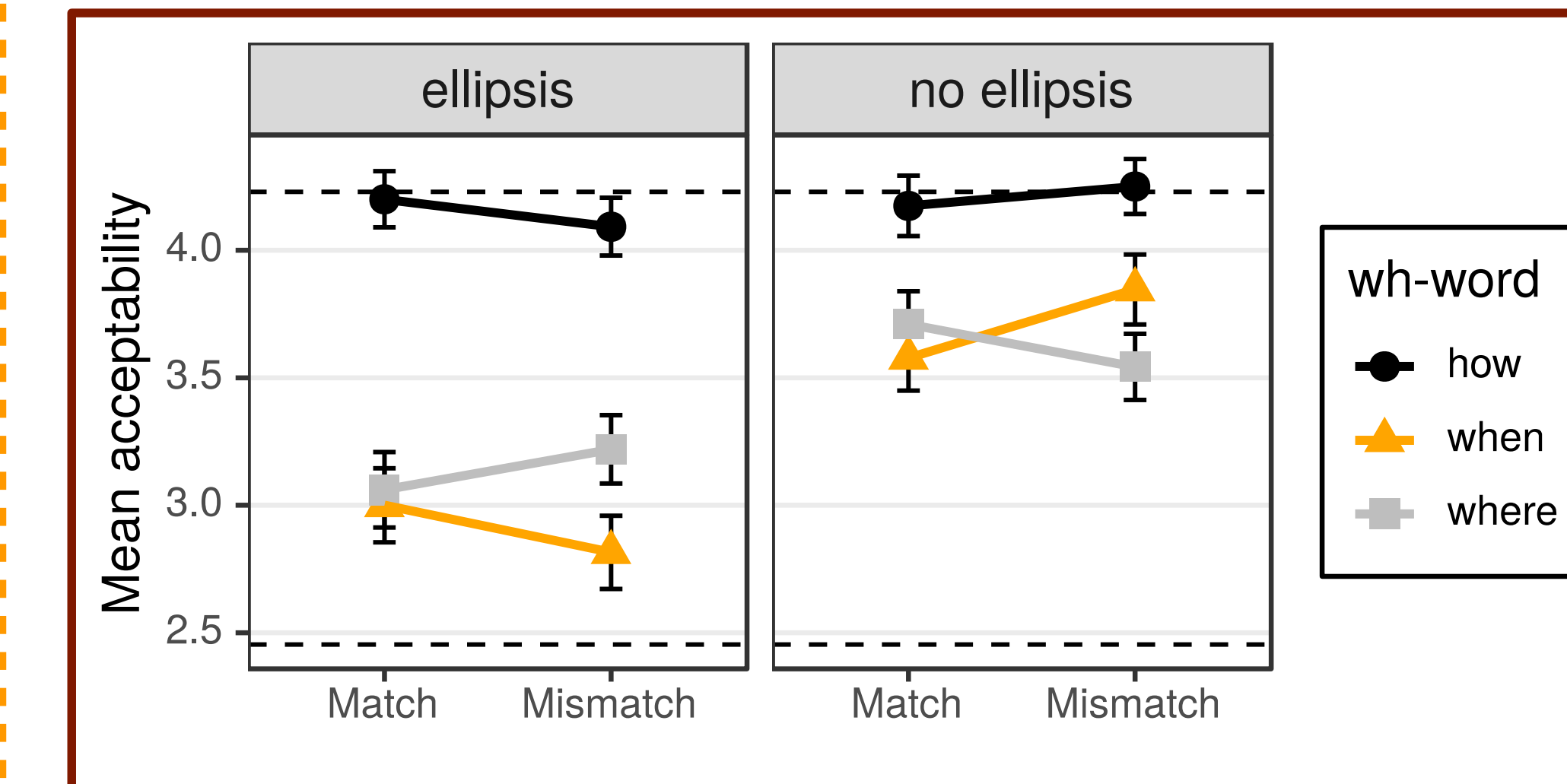*Figure 2.* Screenshot of sample trial.

Results

Figure 3. Mean acceptability of tough alternation items across 3 factors: match vs. mismatch; ellipsis vs. no ellipsis; how vs. when vs. where. Dashed lines show mean acceptability of upper- and lower-bound fillers.

We modeled the results in an **ordinal regression** analysis with maximal random effects (Barr et al., 2013) using the brms R package (Bürkner, 2017):

```
rating ~ ellipsis*mismatch*wh.word +
  (1 + ellipsis*mismatch*wh.word | subject) +
  (1 + ellipsis*mismatch*wh.word | item)
```

Figure 4 shows posterior samples of the population-level coefficients, which suggest the following 4 results:

1. No mismatch penalty. The tough-mismatched sentences were statistically indistinguishable from their non-mismatched counterparts ($\beta = 0.19$, $P(\beta < 0) = 0.31$), both under ellipsis and without ellipsis.

2. No overall ellipsis penalty. Sluicing did not affect acceptability ($\beta = -0.12$, $P(\beta < 0) = 0.63$).

3. 'When/where' items were degraded. Compared to their 'how' counterparts, 'when' and 'where' items received lower ratings ($\beta = -1.57$, $P(\beta < 0) = 1$; $\beta = -1.24$, $P(\beta < 0) = 0.99$, respectively).

4. 'When/where' penalty greater under ellipsis. The penalty for 'when' and 'where' items compared to 'how' variants was greater for sluiced items compared to their non-elliptical counterparts ($\beta = -1.15$, $P(\beta < 0) = 0.96$; $\beta = -1.55$, $P(\beta < 0) = 0.99$, respectively).
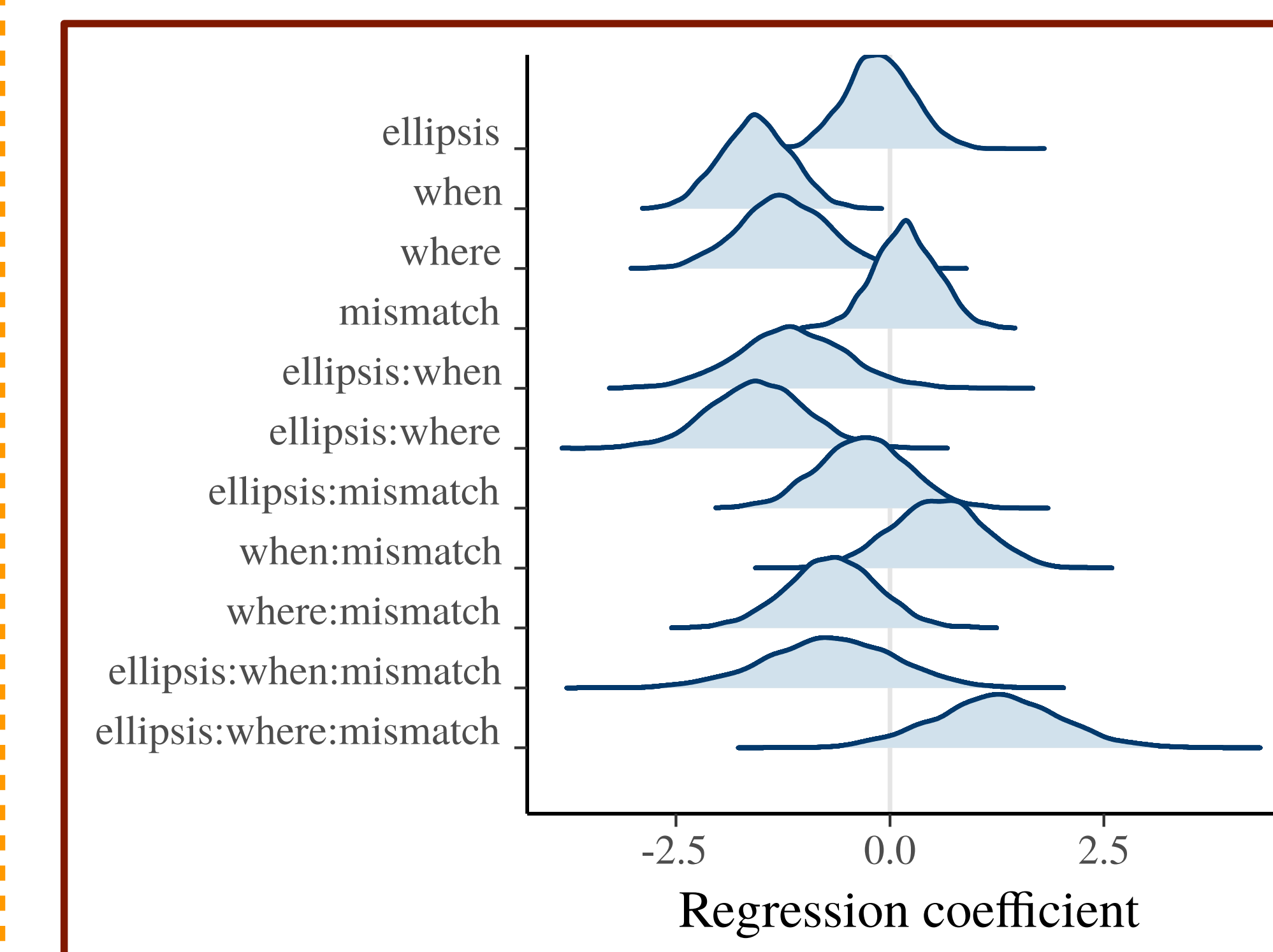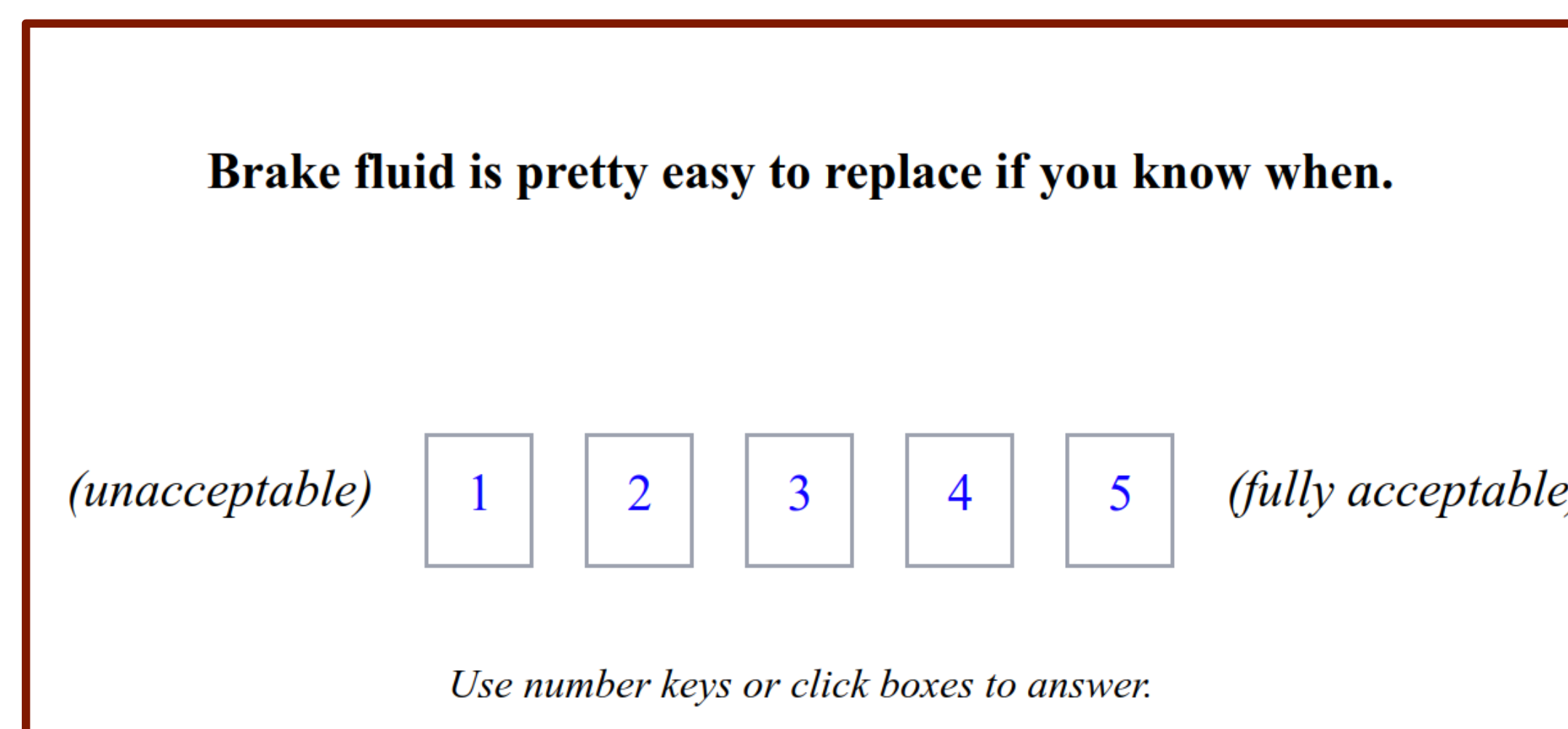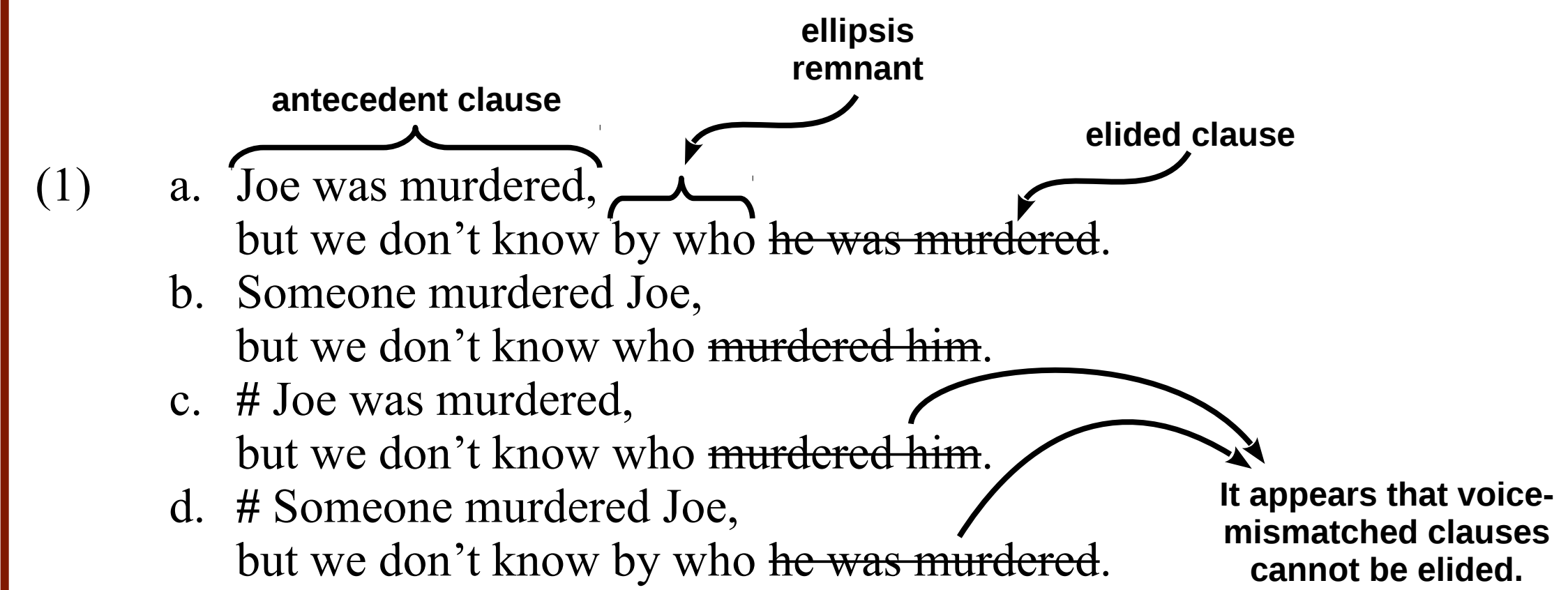

Figure 4. Probability density of posterior samples of population-level regression coefficients.

Discussion

Our results indicate that sluicing is **insensitive** to argument-structure mismatches that result from tough movement.

Such mismatches can be handled by existing "fine-tuning" accounts that relax the Identity condition to allow lexical mismatches as long as they have **identical indices** (e.g. Merchant, 2001; Rudin, 2019):

(7) Mismatches$_i$ are easy **to elide t_i** if you know how ~~to elide mismatches$_i$~~.

Merchant (2001) and Rudin (2019) argue that this revision of the Identity condition is required for examples like (8), among others:

(8) **Who$_i$** did the suspect call t_i and when ~~did they call them$_i$~~. (Merchant, 2001, ex. 112b)

However, those Identity approaches are problematic for 3 reasons:

1. "Over-fitting:" In the absence of independent motivation, fine-tuning the Identity condition to capture specific observations isn't (by itself) explanatory.

2. Descriptive adequacy: Even the latest Identity account (Rudin, 2019), which improves on previous attempts by allowing acceptable mismatches like those in (4a-d), falsely rules out lexical mismatches like the one in (4e).

3. Can't explain 'when'/'where' results: We found an ellipsis-specific penalty for 'when' and 'where' variants relative to their 'how' counterparts. Since antecedent and elided material are held constant within items, any Identity condition that licenses the 'how' variant will also predict the 'when'/'where' variants to be acceptable. The cause of this effect is an important question for future research.

Final remarks

The acceptability of non-elliptical items is "mirrored" in the ellipsis condition (see Fig. 3). If this pattern isn't spurious, it may hint at a "speaker-choice effect:" whatever is causing an elliptical utterance to be degraded could improve the acceptability of non-elliptical alternatives that convey the same meaning and vice versa. Further research is needed to test this possibility.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: keep it maximal. *Journal of memory and language*, 68(3), 255–278.
Bürkner, Paul-Christian (2017). Brms: an r package for bayesian multilevel models using stan. *Journal of Statistical Software*, 80(1), 1–28.
Chung, S. (2006). Sluicing and the lexicon: the point of no return. In *Proceedings of the annual meeting of the Berkeley Linguistics Society* (pp. 73–91).
Chung, S. (2013). Syntactic identity in sluicing: how much and why. *Linguistic Inquiry*, 44(1), 1–44.
Kehler, A. (2002). Coherence, reference, and the theory of grammar. CSLI publications, Stanford, CA.
Kertz, L. (2013). Verb phrase ellipsis: the view from information structure. *Language*, 89(3), 390–428.
Kroll, M. (2016). Polarity Reversals under Sluicing. In *Proceedings of Sinn und Bedeutung*, Eds. R. Truswell, C. Cummins, C. Heycock, B. Rabern, and H. Rohde, (21), 713-729.
Merchant, J. (2001). The syntax of silence: sluicing, islands, and the theory of ellipsis. Oxford University Press on Demand.
Merchant, J. (2013). Voice and ellipsis. *Linguistic Inquiry*, 44(1), 77–108.
Poppels, T. & Kehler, A. (2019). Sluicing inferred propositions. *Sluicing and Ellipsis At 50*. Chicago, IL.
Rudin, D. (2019). Head-based syntactic identity in sluicing. Linguistic Inquiry, 50(2), 253–283.

tpoppels.github.io