Asymmetries in voice-mismatched VP-ellipsis [1]

Till Poppels & Andrew Kehler *(Contact: tpoppels.github.io)*

December 2, 2018

[1] We thank audiences at CUNY 2018, CLS 54, MIT's Computational Psycholinguistics lab, Harvard's Language and Cognition group, as well as Helena Aparicio and Brian Dillon for helpful comments and discussion. We gratefully acknowledge support from NSF grant BCS-1456081.

1 *The "mismatch asymmetry"*

1.1 *An example*

Consider (1), adapted from Arregui et al. (2006).

antecedent clause — ellipsis clause

(1) a. The judge read the report first, and then the lawyer did too. `[A -> A]`
b. The report was first read by the judge, and then the confession was too. `[P -> P]`
c. The report was first read by the judge, and then the lawyer did too. `[P -> A]`
d. The judge read the report first, and then the confession was too. `[A -> P]`

Two empirical findings

- **Mismatch penalty:** (c-d) are less acceptable than (a-b) (Sag, 1976; Kehler, 2002; Kim et al., 2011; SanPietro et al., 2012; Kertz, 2013; note that the theoretical status of this fact, which is irrelevant for our purposes, remains controversial: Merchant, 2013; Kim & Runner, 2018; and many others)
- **Mismatch asymmetry:** `[A -> P]` less acceptable than `[P -> A]` (Arregui et al., 2006; Kim & Runner, 2018; Parker, 2017; but note null effect in Kim et al., 2011)

1.2 *A memory-based account: the Recycling Hypothesis* [2]

[2] The Recycling Hypothesis was first proposed in Arregui et al. (2006) and further defended in Frazier (2013).

- Two explanatory components:
  - **Grammar:** simple syntactic identity ⇒ rules out mismatches as ungrammatical
  - **Processing:** Recycling of non-ID antecedents ⇒ "acceptable ungrammaticality"
- **Memory asymmetry**: passive antecedents are more likely to be misremembered as active than vice versa (Mehler, 1963).
- memory asymmetry ⇒ asymmetric **illusion of grammaticality** ⇒ mismatch asymmetry: `[P -> A]` mismatches more likely to be misremembered as `[~~P~~A -> A]` matches.

(2) a. The report was first read by the judge, and then the lawyer did too. `[~~P~~A -> A]`
b. The judge read the report first, and then the confession was too. `[A -> P]`

1.3 *The remainder of this talk*

- **Thesis**: mismatch asymmetry can't be explained in terms of Recycling
- **Support**: evidence from 2 acceptability judgment experiments (Expts 1 and 2)
- **Next directions:** towards an alternative explanation (Expts 3 and 4)

2 *Experiment 1*

- **Goals:**
  - Replicate the mismatch asymmetry
  - Test voice-*matched* controls
- **Methods:** 2x2 design; 30 participants; 24 items like (3), 40 fillers; 5-point acceptability judgment task

2.1 *Stimuli (Expt 1)*

Expt 1 tested items like (1), repeated in (3). [3]

(3) a. The judge read the report first, and then the lawyer did. `[A -> A]`
b. The report was first read by the judge, and then the confession was too. `[P -> P]`
c. The report was first read by the judge, and then the lawyer did too. `[P -> A]`
d. The judge read the report first, and then the confession was too. `[A -> P]`

(4) a. The thief was arrested and his brother was as well. Acceptable filler
b. A proof that God exists doesn't. Unacceptable filler

[3] These items were adopted from Arregui et al. (2006). Half of them featured connective 'and' or 'and then', as shown in (3), and the other half employed 'after'.

2.2 *Predictions from Recycling (Expt 1)*

- **Mismatch penalty**: `[P -> A]` and `[A -> P]` less acceptable than `[A -> A]` and `[P -> P]`
- **Mismatch asymmetry**: `[P -> A]` more acceptable than `[A -> P]`
- **No "match asymmetry"**: `[A -> A]` and `[P -> P]` should be equally acceptable

2.3 *Results (Expt 1)*

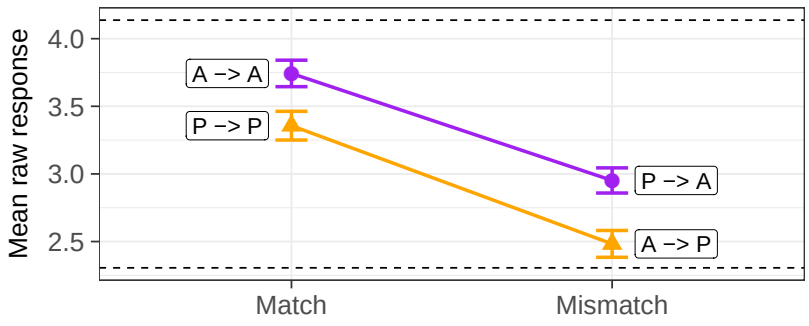

Figure 1: Results from Expt 1. There was a **mismatch penalty** ($\beta = -0.41, p < 0.001$), and a **passive penalty** ($\beta = -0.22, p = 0.001$), but **no interaction** between the two ($\beta = -0.02, p = 0.67$). Statistical results here and throughout are based on linear mixed-effects regression analyses with maximal random effect structure (Barr et al., 2013), and all p-values are derived from likelihood-ratio model comparisons. Dashed lines indicate mean ratings of (un)acceptable elliptical fillers.

- As predicted: **mismatch penalty** and **mismatch asymmetry**
- Contrary to Recycling: **"match asymmetry"**
- Mismatch asymmetry due to **penalty for passive ellipsis clauses**

3 *Experiment 2*

- **Goal:** Distinguish passive-penalty and Recycling explanations by manipulating clause order
- **Methods** (same as in Expt 1): 2x2 design; 30 participants; 24 items like (5), 40 fillers; 5-point acceptability judgment task

3.1 *Stimuli (Expt 2)*

(5) a. Before the lawyer did, the judge read the report first. `[A <- A]`
b. Before the confession was, the report was first read by the judge. `[P <- P]`
c. Before the lawyer did, the report was first read by the judge. `[A <- P]`
d. Before the confession was, the judge read the report first. `[P <- A]`

3.2 *Predictions (Expt 2)*

- **Passive penalty explanation:** if mismatch asymmetry is driven by passive penalty, `[P <- A]` should be **less acceptable** than `[A <- P]`
- **Recycling explanation:** if asymmetry is due to asymmetrically noisy memory for past material, `[P <- A]` should be **more acceptable** than `[A <- P]` (cf. (6) and (7))

Logic behind Recycling prediction:
Without cataphora, repeated from (2):

(6) a. The report was first read by the judge before the lawyer did too. `[P A -> A]`
b. The judge read the report first before the confession was too. `[A -> P]`

Under cataphora, the predictions flip:

(7) a. Before the lawyer did, the report was first read by the judge. `[A <- P]`
b. Before the confession was, the judge read the report first. `[P A <- A]`

3.3 *Results (Expt 2)*

- As in Expts 1: **mismatch penalty** and **passive penalty**
- Contrary to Recycling: **mismatch asymmetry did not flip**

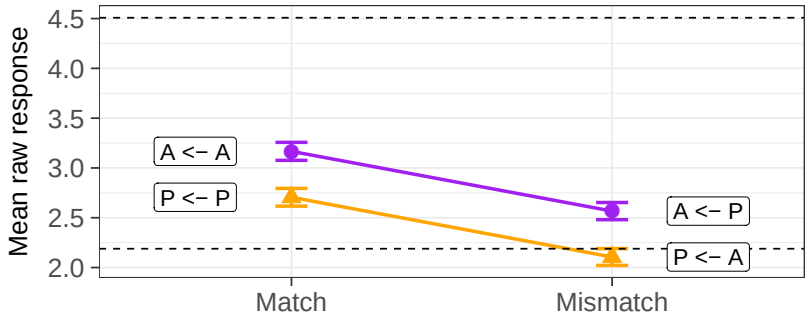

Figure 2: Results from Expt 2. There was a mismatch penalty ($\beta = -0.3, p < 0.001$), a passive penalty ($\beta = -0.23, p < 0.001$), and no interaction between the two ($\beta = 0.001, p = 0.99$). Dashed lines indicate mean ratings of (un)acceptable elliptical fillers.

3.4 *Discussion (Expt 2)*

- **Passive ellipsis clauses remain degraded** even when they *precede* the antecedent (despite being subject to misremembering), which is **inconsistent with Recycling** account
- Both Expts 1 and 2 found two independent main effects:
  - mismatch penalty
  - passive penalty
- Expt 3: are these effects ellipsis-specific?
- Expt 4: what about gapping and sluicing?

4 *Experiment 3*

- **Goals:**
  - Replicate Expt 1
  - Test for passive penalty in **non-elliptical** items
- **Methods: 2x2x2** design; **60** participants; 24 items like (8), 40 fillers; 5-point acceptability judgment task

4.1 *Stimuli (Expt 3)*

In addition to the VP-ellipsis variants in (3), Expt 3 tested their non-elliptical counterparts:

(8) a. The judge read the report first, and then the lawyer read it. `[A - A]`
b. The report was first read by the judge, and then the confession was read too. `[P - P]`
c. The report was first read by the judge, and then the lawyer read it too. `[P - A]`
d. The judge read the report first, and then the confession was read too. `[A - P]`

4.2 *Predictions (Expt 3)*

- Replicate results from Expt 1: mismatch and passive penalties
- If mismatch penalty and passive penalty are ellipsis-specific effects, they should **disappear in non-ellilptical variants**.

4.3 *Results (Expt 3)*

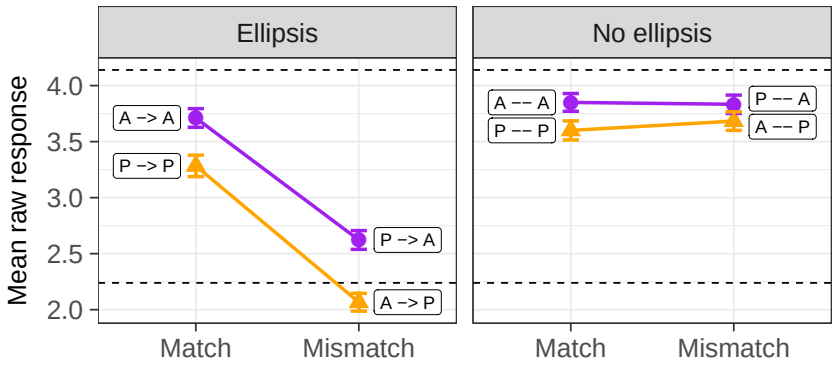

Figure 3: Results from Expt 3. Left: full replication of Expt 1. Right: Non-elliptical item variants revealed **no mismatch penalty**, and a substantially reduced, statistically **marginal passive penalty**: $\beta = -0.1, p = 0.068$. Dashed lines indicate mean ratings of (un)acceptable elliptical fillers.

- Full replication of Expt 1
- **No mismatch penalty** without ellipsis
- **No (or much reduced) passive penalty** without ellipsis

5 *Experiment 4*

- **Goal:**
  - Explore if passive penalty affects gapping and sluicing
- **Methods: 2x2** design; **26** participants; 12 sluicing items like (9), 12 gapping items like , 40 fillers; 5-point acceptability judgment task

5.1 *Stimuli (Expt 4)*

(9)
a. Someone read the report, but I don't know who. `elliptical [A -> A]`
b. The report was read by someone, but I don't know by whom. `elliptical [P -> P]`
c. Someone read the report, but I don't know who read it. `non-elliptical [A -> A]`
d. The report was read by someone, but I don't know by whom it was read. `non-elliptical [P -> P]`

(10)
a. Mary scolded Wilma, and Susan, Nancy. `elliptical [A -> A]`
b. Wilma was scolded by Mary, and Nancy, by Susan. `elliptical [P -> P]`
c. Mary scolded Wilma, and Susan scolded Nancy. `non-elliptical [A -> A]`
d. Wilma was scolded by Mary, and Nancy was scolded by Susan. `non-elliptical [P -> P]`

5.2 *Results (Expt 4)*

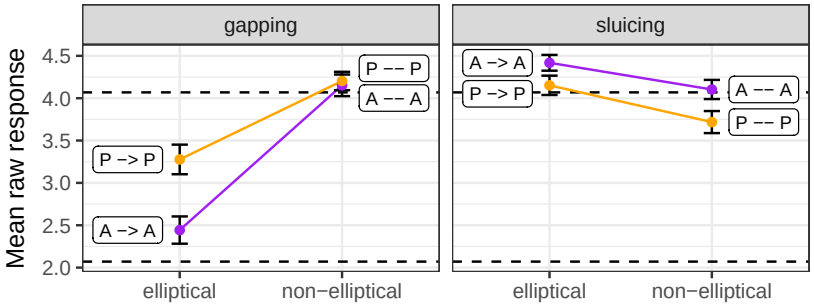

Figure 4: Results from Expt 4. Right: passive sluices were not significantly worse than active sluices ($\beta = -0.16, p = 0.17$), but ellipsis had a positive effect ($\beta = 0.19, p = 0.003$). Left: By contrast, acceptability suffered under gapping ($\beta = -0.65, p < 0.001$), which revealed an *active* penalty ($\beta = -0.2, p < 0.059$). Dashed lines indicate mean ratings of (un)acceptable elliptical fillers.

- **No passive penalty** for gapping or sluicing
- Improvement for sluicing under ellipsis could reflect "repeated-clause penalty"
- Open question: why the "active penalty" under gapping?

6 *Conclusion*

- **Main take-away**: mismatch asymmetry can't be explained in terms of Recycling
- Mismatch asymmetry is driven by passive penalty that also applies to voice-*matched* ellipsis
- Passive penalty is specific to VP-ellipsis: it doesn't affect non-elliptical utterances, gapping, or sluicing

References

[1] Ana Arregui, Charles Clifton, Lyn Frazier, and Keir Moulton. Processing elided verb phrases with flawed antecedents: The recycling hypothesis. *Journal of memory and language*, 55(2):232–246, 2006.

[2] Dale J Barr, Roger Levy, Christoph Scheepers, and Harry J Tily. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, 68(3):255–278, 2013.

[3] Lyn Frazier. A recycling approach to processing ellipsis. In Lisa Lai-Shen Cheng and Norbert Corver, editors, *Diagnosing Syntax*, pages 485–501. 2013.

[4] Andrew Kehler. *Coherence, reference, and the theory of grammar*. CSLI publications Stanford, CA, 2002.

[5] Laura Kertz. Verb phrase ellipsis: The view from information structure. *Language*, 89(3):390–428, 2013.

[6] Christina S. Kim, Gregory M. Kobele, Jeffrey T. Runner, and John T. Hale. The acceptability cline in vp ellipsis. *Syntax*, 14(4):318–354, Jul 2011.

[7] Christina S Kim and Jeffrey T Runner. The division of labor in explanations of verb phrase ellipsis. *Linguistics and Philosophy*, 41(1):41–85, 2018.

[8] Jacques Mehler. Some effects of grammatical transformations on the recall of english sentences. *Journal of verbal Learning and verbal Behavior*, 2(4):346–351, 1963.

[9] Jason Merchant. Voice and ellipsis. *Linguistic Inquiry*, 44(1):77–108, 2013.

[10] Daniel Parker. Navigating ellipsis structures in memory: New insights from computational modeling. LSA workshop: Corpus-based and Experimental Approaches to Ellipsis, 2017.

[11] Ivan A Sag. *Deletion and logical form*. PhD thesis, Massachusetts Institute of Technology, 1976.

[12] Steve SanPietro, Jason Merchant, and Ming Xiang. Accounting for voice mismatch in ellipsis. In *Proceedings of the 30th West Coast Conference on Formal Linguistics, ed. by Nathan Arnett and Ryan Bennett*, pages 303–312, 2012.

[13] Ming Xiang and Josef Klafka. Memory retrieval in comprehension is sensitive to production alternative. 31st Annual CUNY Sentence Processing Conference, 2018.