There is reconstruction for Condition C in English questions*

Richard Stockwell, Aya Meltzer-Asscher & Dominique Sportiche

Christ Church, Oxford; Tel Aviv University; University of California, Los Angeles

1. Introduction

It is theoretically foundational that the Condition C effect in (1) (Chomsky 1981) persists after A-bar movement in (2) (Barss 1986, Heycock 1995, Fox 1999, Takahashi and Hulsey 2009, i.a.). That is, A-bar movement 'reconstructs' for Condition C:

(1) * He_i framed [the picture of Harry_i] .

(2) a. (*) [Which picture of Harry_i]$_j$ did he_i frame t_j ?
b. (*) [Which picture of Harry_i]$_j$ did Meghan say he_i framed t_j ?

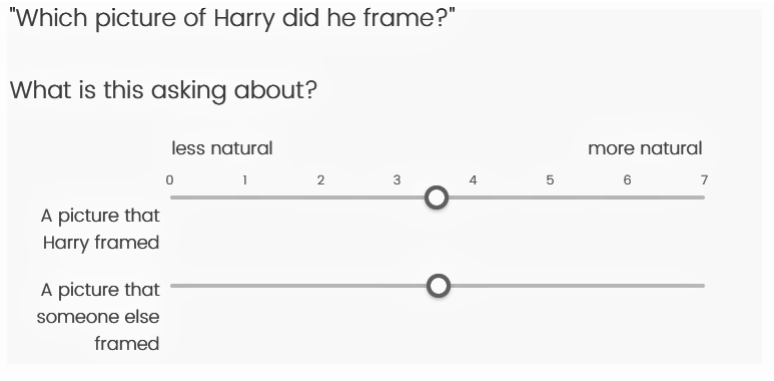
However, recent experimental work has questioned the existence of Condition C reconstruction in English (2a) (Bruening and Al Khalaf 2019), especially at a distance (2b) (Adger et al. 2017). On the contrary, this paper reports a formal, large-scale acceptability rating experiment which supports the claim that there is reconstruction for Condition C in English questions. The next two sections present our design and results. The discussion in section 4 reflects on distance, compares our experiment with previous studies, and speculates on a role for accent on the pronoun in alleviating Condition C effects. Section 5 concludes.

2. Design

Participants were tasked with imagining they were joining an ongoing conversation at a party — an 'eavesdropping' context, neutral as to co- or disjoint reference for the pronoun. The target item, a prompt, and two responses were presented simultaneously, as in (3):

*Many thanks to our audience at NELS 51, three anonymous NELS reviewers, and Ethan Poole and his UCLA seminar on reconstruction effects. Partial support for this study from a UCLA Academic Senate grant to Dominique Sportiche is gratefully acknowledged.

(3) *Task*



One response contained the same name as the question, indicating a coreferential reading for the pronoun; while the other contained *someone else*, indicating disjoint reference. Participants were asked to rate the naturalness of each response on separate 0-7 sliding Likert scales.

The 2x2x2 design of our experiment is laid out in (4), with an example set of items in (5)-(8):

(4) *Design*

Example	I. Condition C	II. Distance	III. Response
(5)	Yes	Short	Name/Else
(6)	Yes	Long	Name/Else
(7)	No	Short	Name/Else
(8)	No	Long	Name/Else

(5) *Yes, Short*
[Which picture of Harry] did he frame *t* ?
A picture that { Harry / someone else } framed.

(6) *Yes, Long*
[Which picture of Harry] did Meghan say he framed *t*?
A picture that { Harry / someone else } framed.

(7) *No, Short*
[Which picture of Harry] *t* made him laugh?
A picture that made { Harry / someone else } laugh.

(8) *No, Long*
[Which picture of Harry] *t* made Meghan say he has good taste?
A picture that suggests { Harry / someone else } has good taste.

The first factor was the potential for a Condition C effect to arise, given the base position of A-bar movement: Yes, with the base-position below the pronoun, as in (5) and (6); and No, with the base-position above the pronoun, as in (7) and (8).[1]

The second factor was Distance. In Short, monoclausal sentences, the name and pronoun were separated by a single word, *did* or *made*, as in (5) and (7); whereas in Long, biclausal sentences, the name and pronoun were separated by three words and a clause boundary, *did/made 'Name' say*, as in (6) and (8).[2]

The third factor was Response: Name, indicating a coreferential reading for the pronoun; and Else, indicating disjoint reference.

We created twelve sets of items using two of each of six pairs of noun and prepositional phrase complement: *picture of*, *statue of*, *side of*, *portrait of*, *report on*, and *letter to*.[3] Hence the target sentence was always a wh-question with a wh-phrase of the form *Which N P Name*.

---

[1]The ultimate base-position of the wh-phrases in (7) and (8) may be lower. An anonymous reviewer points out that in *make*-causatives, the causer reconstructs below the causee for Condition A. Indeed, the facts in (i) from Pesetsky (1995: 43f.) continue to hold after controlling for logophoricity with inanimates (Charnavel and Sportiche 2016) in (ii):

(i) a. Pictures of each other$_i$ made us$_i$ happy.
b. These stories about herself$_i$ made Mary$_i$ nervous.

(ii) a. Shortcomings in its$_i$ own design made the bridge$_i$ collapse.
b. The picture of itself$_i$ made the room$_i$ feel quirky.

What (i) and (ii) show, however, is that the causers have raised by A-movement from below the causees. A-movement can reconstruct for Condition A; but since A-movement does not have to reconstruct, it need not trigger a Condition C effect, as shown by the grammaticality of (iii):

(iii) [John$_j$'s mother]$_i$ seems to him$_j$ t_i to be lovely.

As far as Condition C is concerned, therefore, the base-position of A-bar movement indicated in (7) and (8) is the most relevant one. Though see section 4.1 for further discussion.

[2]Our long-distance questions positioned the pronoun in the lowest clause, as in (6). We did not test examples with the pronoun as the subject of the higher clause, like (i):

(i) [Which picture of Harry] did he say Meghan framed *t*?

Adger et al. (2017: ex. 14) tested examples of both kinds. They found a preference for coreference in examples like (i), and an even stronger preference for coreference in examples like (6). We chose to run the (6) pattern, where Adger et al. (2017) found 'less Condition C reconstruction', as it offered more to 'overcome'.

That said, there might be reasons to expect either of (6) or (i) with coreference to be worse than the other. Adger et al. (2017: ex. 10) discuss an account in terms of Vehicle Change (cf. Safir 1999) which predicts that (6) will be worse than (i). If names can count as pronouns in reconstructed positions, there would be no Condition C violation in either (6) or (i); but there would still be a Condition B violation in (6). As it happens, Adger et al.'s (2017) results showed the opposite, with (i) worse with coreference than (6). This pattern might follow from taking account of successive cyclic movement: the pronoun c-commands more intermediate traces of the wh-phrase in (i), which could cause more of a Condition C effect.

[3]We would be happy to supply a full list of stimuli upon request.

On the suspicion that intensionality may play a role in alleviating Condition C effects, we controlled for it by ensuring that all DP-taking verbs failed Moltmann's (1997) tests.

The items were distributed in a Latin square design across four lists. The order of presentation of the two response options was consistent for a given participant, but balanced across lists; i.e., two with Name positioned above Else, as in (3), and two with Else above Name.

Twelve baseline items were seen by every participant. In six items designed to be uncontroversially good with coreference, like (9), Name was rated appropriately highly. In six straightforward Condition C violations like (10), meanwhile, the low ratings for Name confirmed that our experiment was sensitive to Condition C:

(9) *Good*

[Which statue] did Flo say she bought *t*?		
A statue that Flo bought.	Name	6.29
A statue that someone else bought.	Else	1.94

(10) *Bad*

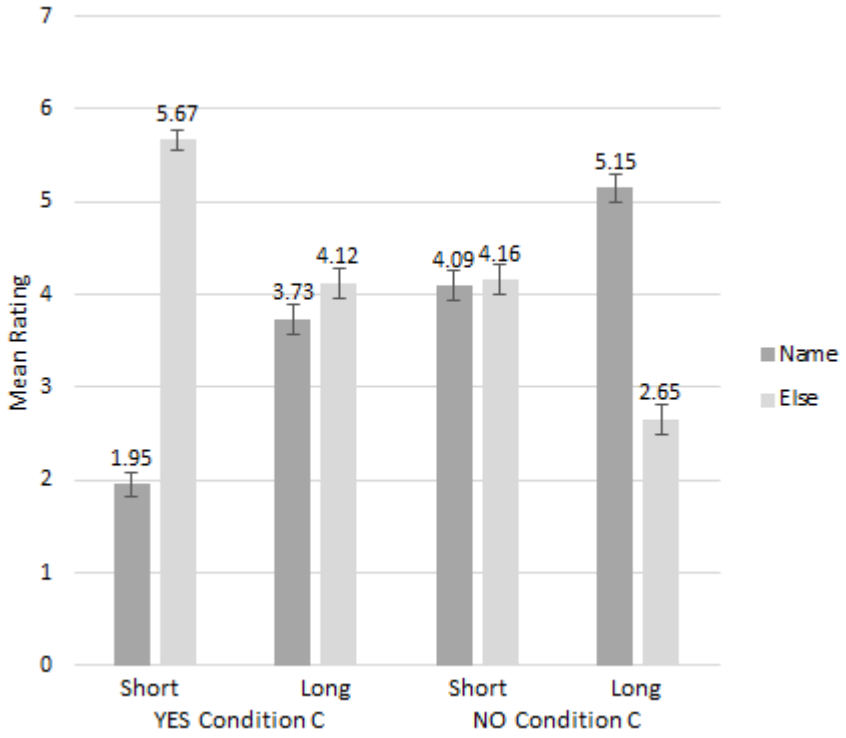
[Which statue] did he say Carol made Gary sell *t*?		
A statue that Gary was speaking about.	Name	1.61
A statue that someone else was speaking about.	Else	5.79

3. Results

Data from 223 native English-speaking undergraduates were analysed in the R programming environment (R Core Team 2013), modeled using mixed-effects linear regression with the lmerTest package (Kuznetsova et al. 2017).[4] Mean ratings for the four main conditions are plotted in (11), with +/-1 standard error of the mean. The darker bars represent Name responses, the lighter bars Else responses:

[4]There were 249 native English speaking undergraduate participants. Participants were excluded on two grounds. The first ground for exclusion was taking a time to complete the survey above the 97.5 percentile (5771 seconds $\approx$ 96 minutes) or below the 2.5 percentile (305 seconds $\approx$ 5 minutes). This resulted in the exclusion of 14 participants. The second was having an average rating of the 'Bad' baselines like (10) above one's mean rating across the experiment overall (i.e. z-score $>$ 0). This resulted in the exclusion of a further 12 participants. The remaining 223 participants were not evenly distributed across lists. List 1 had 60 participants, list 2 had 54, list 3 had 54, and list 4 had 55.

(11) *Results*



Descriptively, starting with Yes, we found strong evidence of Condition C reconstruction at a Short distance. In Long, the effect might at first look to have disappeared. But turning to No, there is an equal preference for Name and Else in Short, and a strong preference for Name in Long. In this light, we see that Condition C continues to have an effect at a distance, as it overturns the baseline Name preference in Long.

Statistically, there was a significant two-way interaction between Condition C and Response (estimate = -3.655, SE = 0.174, t (5122) = -21.017, $p < .001$). In Yes, Else responses increased ratings (Else 4.90, Name 2.84, $p < .001$); whereas in No, Else decreased ratings (Else 3.40, Name 4.62, $p < .001$). We take this to be strong evidence that there is Condition C reconstruction of preposed DPs in English. Further, the effect persisted with Distance. There was a significant three-way interaction between Condition C, Distance and Response (estimate = 0.76, SE = 0.24, t (5122) = 3.090, p = .002). In No, participants were equally happy with a Name or Else interpretation in Short (p = 0.616), and preferred Name in Long ($p < .001$); whereas in Yes, participants very strongly preferred Else in Short ($p < .001$), and continued to prefer Else in Long (p = .0013). That is, while the effect of Condition C reconstruction lessens with distance, it remains strong enough to flip the preference from Name to Else in Long.

4. Discussion

In sum, we found strong evidence for Condition C reconstruction of DPs in English questions, even at a distance. Our results thus differ from previous experimental work on English, which did not find robust evidence for Condition C reconstruction of DPs (Bruening and Al Khalaf 2019), especially at a distance (Adger et al. 2017). After speculating on the role of distance, this section sets our different findings in the context of differences in statistical power and methodology. Finally, we consider the potential role of accent on the pronoun in alleviating Condition C effects.

4.1 Distance

While we found a Condition C effect in Long, the effect was weaker than in Short. We speculate that the effect of Condition C diminishes with distance for processing reasons. While the reference of a DP is what matters for interpretation, its form is what matters for Binding Theory. Information about the form of a DP might decay with increased processing load, and greater distance brings more to process. Our Long items also included an additional discourse referent, the name of the 'say'-er; e.g. *Meghan* in (6).

At the same time, the strong preference for Name over Else in No, Long suggests a preference for resolving anaphora if at all possible. From this perspective, the lack of a similar preference in No, Short might be surprising. This lack of Name preference may be due to the ultimate base position of the wh-phrase being below the causee — recall the discussion in footnote 1. In No, Short, the ultimate base position of the wh-phrase would then be below the pronoun. The option of reconstruction to this position for Condition C could be responsible for the lower rating for Name in No, Short. In No, Long, on the other hand, the ultimate base position of the wh-phrase is still above the pronoun, giving no chance for a Condition C effect to arise. This speculation predicts that a version of Long with the pronoun as causee, namely (12), should pattern like our No, Short, with optional reconstruction below the pronoun precluding a preference for Name over Else:

(12) [Which picture of Harry] *t* made him (*t*) say Meghan has good taste?

Notably, if the rating for Name in No, Short is depressed for these reasons, then our results in the previous section undervalue the effect of Condition C.

4.2 Comparison with previous studies

One reason for the difference between our results and those of previous studies might be that our number of participants provided greater statistical power. Our 223 participants compares with 53/91/89 across the three experiments in Adger et al. (2017), and 75/75/70 in Bruening and Al Khalaf (2019). Statistical power may be important, considering many of the contrasts in Bruening and Al Khalaf (2019) trend towards significance.

More substantive are differences in methodology. Adger et al. (2017) forced a Yes/No response as to whether participants could use a sentence like (13) when the highlighted name and pronoun referred to the same individual:

(13) Which side of **Elizabeth** does **she** prefer? (Adger et al. 2017: ex. 14a)

The desire to resolve pronominal reference in the absence of any other salient options may have led to an over-representation of Yes responses. Biasing the same way, the directness of the task may have encouraged shallow processing of phi-feature match, ignoring the c-command configurations that are crucial to Condition C.

Bruening and Al Khalaf (2019), meanwhile, forced a choice between two intra-sentential referents for the pronoun in items like (14) (Bruening and Al Khalaf 2019: 251, ex. 15a):

(14) A female staffer told everyone which of the announcements that Hillary Clinton was running for president she had actually authorized.
Who authorized the announcement?
A: the staffer B: Hillary Clinton.

Their task thus probed referential preferences rather than possibilities, in sentences complex enough to house two potential antecedents.

By contrast, our methodology may be more sensitive to Condition C reconstruction. Our Else response option raises the possibility of disjoint reference to salience, rather than offering only one way to resolve the pronoun. Our task was also more indirect, asking about the naturalness of answers to a wh-question rather than directly about (co)reference. Finally, our task posed separate questions about referential possibilities, while keeping the target sentences relatively simple.[5, 6]

Our task was more similar the one reported in Georgi et al. (2018) and Wierzba et al. (to appear). They found robust Condition C reconstruction in German by posing separate questions about referential possibilities with items like (15) (Wierzba et al. to appear: exx. 7b, 8):

(15) Maria erzählt, [welche Statue von Anna] sie gesehen hat.
'Mary tells (of) which statue of Anna she saw.'
Q1: Kann man den Satz so verstehen, dass Maria eine Statue gesehen hat?
'Can the sentence be interpreted such that Mary saw a statue?'
Q2: Kann man den Satz so verstehen, dass Anna eine Statue gesehen hat?
'Can the sentence be interpreted such that Mary saw a statue?'

The differences between our task and that in Georgi et al. (2018) and Wierzba et al. (to appear) are more minor. First, they used forced choice Yes/No responses, while we gathered

[5]That said, while our task aimed to probe referential possibilities, participants seem to have responded, at least in aggregate, with preferences — each pair of bars in the graph in (11) adds up to around 8.

[6]Such pairwise presentation has also been found to be more sensitive and more statistically powerful than gathering ratings for sentences in isolation (Sprouse et al. 2013: 225, 228).

naturalness ratings on Likert scales. Second, their items were indirect wh-questions, where matrix subject position housed a second reference option; while our items were direct wh-questions, with disjoint reference to someone else not mentioned in the sentence as the second reference option.

4.3 Accent on the pronoun

Finally, our 'eavesdropping' context invited no special accent on the pronoun. We suspected that accent may play a role in alleviating reconstructed Condition C effects in English, based on reports that accenting a pronoun lessens Condition C effects in Italian (Cinque 2020: ch.2, fn.9) and in French and Greek (Angelopoulos and Sportiche to appear). Independently, Georgi et al. (2018) and Wierzba et al. (to appear) found no evidence of Condition C reconstruction with strong demonstrative pronouns in German.

Consistent with these observations, Yoshida et al. (2019) report clear experimental evidence of Condition C reconstruction in stripping, where ~~ellipsis~~ precludes accent on the stripped pronoun. While their primary concern was island repair, the contrast they report between (16) and (17) is attributable to Condition C reconstruction:

(16) A: Her$_i$ friends reported that the manager wrote to John.
B: No, [to Mary$_i$]$_j$ ~~her$_i$ friends reported that the manager wrote t_j~~.

(17) A: She$_i$ reported that the manager wrote to John.
B: *No, [to Mary$_i$]$_j$ ~~she$_i$ reported that the manager wrote t_j~~.

The base position of the ellipsis remnant *to Mary* is c-commanded by the coindexed subject pronoun *she* in (17), but not by *her* in (16).

5. Conclusion

Where previous work has questioned the existence of Condition C reconstruction with preposed DPs in English, this paper has shown that it is experimentally observable, even at a distance, plausibly validating the large theoretical literature that relies on its existence.

References

Adger, David, Alex Drummond, David Hall, and Coppe van Urk. 2017. Is there Condition C reconstruction? In *Proceedings of NELS 47*, ed. Andrew Lamont and Katerina Tetzloff, 21–31. GLSA.

Angelopoulos, Nikolaos, and Dominique Sportiche. to appear. Clitic dislocations and clitics in French and Greek: from interpretation to structure. To appear in *Natural Language and Linguistic Theory*.

Barss, Andrew. 1986. Chains and anaphoric dependencies. Doctoral dissertation, MIT.

Bruening, Benjamin, and Eman Al Khalaf. 2019. No argument-adjunct asymmetry in reconstruction for Binding Condition C. *Journal of Linguistics* 55:247–276.

Charnavel, Isabelle, and Dominique Sportiche. 2016. Anaphor binding: What French inanimate anaphors show. *Linguistic Inquiry* 47:35–87.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Cinque, Guglielmo. 2020. *The syntax of relative clauses: A unified analysis*. Cambridge University Press.

Fox, Danny. 1999. Reconstruction, binding theory, and the interpretation of chains. *Linguistic Inquiry* 30:157–196.

Georgi, Doreen, Martin Salzmann, and Marta Wierzba. 2018. Condition C reconstruction in German A'-movement: Methodological insights and theoretical implications. Condition C Workshop, Queen Mary University of London.

Heycock, Caroline. 1995. Asymmetries in reconstruction. *Linguistic Inquiry* 26:547–570.

Kuznetsova, Alexandra, Per B. Brockhoff, and Rune H. B. Christensen. 2017. lmertest package: Tests in linear mixed effects models. *Journal of Statistical Software* 82:1–26.

Moltmann, Friederike. 1997. Intensional verbs and quantifiers. *Natural Language Semantics* 5:1–52.

Pesetsky, David. 1995. *Zero syntax*. Cambridge, MA: MIT Press.

R Core Team. 2013. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Safir, Ken. 1999. Vehicle change and reconstruction in A'-chains. *Linguistic Inquiry* 30:587–620.

Sprouse, Jon, Carson T. Schütze, and Diogo Almeida. 2013. A comparison of informal and formal acceptability judgments using a random sample from Linguistic Inquiry 2001-2010. *Lingua* 134:219–248.

Takahashi, Shoichi, and Sarah Hulsey. 2009. Wholesale Late Merger: Beyond the A/A' distinction. *Linguistic Inquiry* 40:387–426.

Wierzba, Marta, Doreen Georgi, and Martin Salzmann. to appear. An experimental investigation of reconstruction for Condition C in German A'-movement. In *Proceedings of CLS 56*.

Yoshida, Masaya, David Potter, and Tim Hunter. 2019. Condition C reconstruction, clausal ellipsis and island repair. *Natural Language and Linguistic Theory* 37:1515–1544.

Richard Stockwell, Aya Meltzer-Asscher, Dominique Sportiche
richard.stockwell@chch.ox.ac.uk, ameltzer@tauex.tau.ac.il, sportich@g.ucla.edu