

## Alternative questions: Distinguishing between negated and complementary disjuncts\*

Andrea Beltrama  
*University of Pennsylvania*

Erlinde Meertens  
*University of Konstanz*

Maribel Romero  
*University of Konstanz*

Submitted 2020-03-13 / First decision 2020-04-23 / Revision received 2020-05-28 /  
Accepted 2020-06-11 / Published 2021-04-22 / Final typesetting 2022-04-29

**Abstract** Polar Questions such as “is the light on?” (henceforth, PQs) and Negative Alternative Questions such as “is the light on or not?” (henceforth, NAQs), despite being seemingly semantically equivalent, are known to differ pragmatically (Bolinger 1978, van Rooy & Šafářová 2003, Biezma 2009). This paper explores the difference between these two question types and Complement Alternative Questions (henceforth, CAQs), a type of question that mentions two mutually exclusive propositions but makes no use of negation (e.g., “is the light on or off?”). Based on experimental evidence, we show that CAQs feature a different pragmatic profile from the other two question types. First, while PQs and NAQs are homogeneously felicitous or infelicitous as invitations, rhetorical questions or inference-corroborating questions, CAQs do not behave uniformly in such contexts. Second, while NAQs are confirmed to be infelicitous discourse-initially, neither NAQs nor CAQs are not necessarily limited to a discourse-final position. We take these findings to suggest that all accounts proposed cannot fully capture the full paradigm of these questions types and highlight several areas of future research to revise them.

**Keywords:** questions, alternatives, negation, complement, disjuncts, focus, Cornering Effect

---

\* This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) within project P3: RO 4247/4-2 “Alternative Questions” of the FOR 2111 “Questions at the Interfaces”. We would like to thank Maria Biezma, Ramona Wallner, the audiences at PLC 41, CLS 53, SuB 22, the Ohio State University Pragmatics reading group and fellow members of the Questions at the Interfaces research unit in Konstanz for insightful comments and questions. We’re also grateful to Kjell-Johan Saebo and three S&P anonymous reviewers for providing guidance throughout every step of the editorial process. All remaining errors are our own.

©2020 Andrea Beltrama, Erlinde Meertens, Maribel Romero  
This is an open-access article distributed under the terms of a Creative Commons Attribution License (<https://creativecommons.org/licenses/by/3.0/>).

## 1 Introduction

Consider the following two questions. (1) is a Polar Question (henceforth, *PQ*), which mentions only one proposition; (2) is a Negative Alternative Question (henceforth, *NAQ*), which mentions one proposition and its negated alternative.

- |     |                         |     |
|-----|-------------------------|-----|
| (1) | Is the light on?        | PQ  |
| (2) | Is the light on or not? | NAQ |

In a world in which light switches only have two possible states, a S(peaker) using either of these moves presents two possible resolutions: one in which the switch is on, and one in which the switch is off. Yet, despite their seeming logical equivalence, the two question types are known to differ pragmatically. First, PQs allow for a wider range of illocutionary acts than NAQs (Bolinger 1978). For example, while the PQ (3a) felicitously serves as a marriage proposal, the NAQ (3b) does not. Second, NAQs convey a sense of insistence, known as the *Cornering Effect*, that makes them perfectly felicitous in discourses like (4), where S is trying to force A to finally answer the question (Biezma 2009). PQs, instead, carry no such effect.

- (3) As marriage proposal:
- a. Will you marry me?
  - b. #Will you marry me or not?
- (4) S: Are you making pasta?  
 A: (Silence and dubitative faces)  
 S: Are you making pasta or not?

The divergent behavior of these moves raises an important question at the interface of semantics and pragmatics: what principles of interpretation allow us to account for the intuitive logical similarity and the distributional divergence between these moves? The current paper contributes a novel empirical perspective to this issue by exploring the difference between these two question types and Complement Alternative Questions (henceforth, *CAQs*), a type of question that mentions a — lexical or phrasal — complementary antonym in the second disjunct rather than negation, as in (5).<sup>1</sup>

<sup>1</sup> We assume a Hamblin-style semantic denotation for alternative questions (Hamblin 1973): They denote the set containing all and only the possible answers to the question. We also assume that all alternative questions — NAQs, CAQs and run-of-the-mill alternative ques-

- (5) Is the light on or off? CAQ

The contribution of the article is twofold. First, we present experimental evidence that CAQs feature a different pragmatic profile from the other two types: while PQs and NAQs are homogeneously felicitous or infelicitous across a variety of uses, CAQs do not behave uniformly in the same contexts; moreover, while NAQs cannot be used to raise an issue for the first time, they can be followed by further questions. Second, we observe that these findings are not immediately predicted by extant analyses of the contrast between PQs and NAQs, and provide suggestions for a revision of these proposals in future work.

The paper is structured as follows. Section 2 reports previous empirical observations on the difference between PQs and NAQs and introduces CAQs. Sections 3, 4, and 5 compare the behavior of PQs, NAQs and CAQs in three sets of experiments. Section 6 frames the experimental findings in light of analyses of the distribution of PQs vs. NAQs. Section 7 highlights several areas of future research to revise current proposals. Section 8 concludes.

## 2 The empirical landscape

### 2.1 Polar Questions vs “or not” Alternative Questions

As observed by Bolinger (1978), PQs are more flexible than the corresponding NAQs across different scenarios, including situations in which a speaker intends to make an invitation in (6), corroborate an inference in (7), or ask a rhetorical question in (8):<sup>2</sup>

- (6) *Invitation:*

*Scenario:* A wants to offer a drink to B.

---

tions — denote a set of complementary disjuncts in the following sense: the alternatives given are not just understood as mutually exclusive, but also as exhausting together the logical possibilities *at the current state of discourse*.

<sup>2</sup> In this paper we concentrate on Bolinger-style rhetorical contexts, which are biased towards a negative answer. Note, however, that rhetorical questions can also be biased towards a positive answer (see Caponigro & Sprouse 2007 for discussion); in such contexts, interestingly, NAQs can function rhetorically. We leave the investigation of these cases to further work.

- (i) (Commenting on someone’s childish behavior:) Are you an adult or not?

- a. A to B: ✓ Do you want something to drink?
  - b. A to B: # Do you want something to drink or not?
- (7) *Inference corroborating:*  
*Scenario:* B thought David was away, and wants to double check that he's back.  
 A: I just saw David.
- a. B: ✓ Is he back from Toronto?
  - b. B: # Is he back from Toronto or not?
- (8) *Rhetorical:*  
*Scenario:* B just did something very surprising. A intends to comment on that.
- a. A: ✓ Are you crazy?
  - b. A: # Are you crazy or not?

When it comes to standard information-seeking uses<sup>3</sup> NAQs also feature a constrained distribution known as the Cornering Effect. This restriction has been broken down into two parts (Biezma 2009). First, NAQs are inappropriate discourse-initially<sup>4</sup> as in (9); second, NAQs are necessarily discourse-final — they do not license follow-up questions/subquestions, as in (10). We label these constraints Part 1 and Part 2 of Cornering (Biezma 2009).

- (9) *Cornering, Part 1*  
*Scenario:* You are in charge of coordinating the cooks for the colloquium dinner. John is one of the cooks. Dinner is tomorrow and you need to know what is happening with the pasta.  
*You:* # Are you making pasta or not? NAQ
- (10) *Cornering, Part 2*  
*Scenario:* Same as (9).  
*You:* Are you making pasta?  
*John:* (Silence and dubitative faces)  
*You:* ✓ Are you making pasta or not? NAQ  
*John:* (Silence and dubitative faces)  
*You:* # Are you making pasta?

<sup>3</sup> We use *standard* to contrast these questions with those explicitly serving a special purpose (e.g., invitations, inference, rhetorical).

<sup>4</sup> Following Biezma (2009) et seq., we use the term *discourse-initial* to describe a question asked in a context where the same question has not been asked before.

## Alternative questions

It follows from these restrictions that NAQs are only felicitous in a context in which a question has already been asked and no other question follows it. Note that PQs are not subject to either component of Cornering. They are felicitous discourse-initially and can be followed by further questions, as shown in (10). Such contrasts are summarized in Table 1.

---

Question	Invitation	Inference	Rhetorical	I-S: D-initial	I-S: licenses f/u
PQ	✓	✓	✓	✓	✓
NAQ	#	#	#	#	#

---

**Table 1** Distribution of NAQs vs. PQs. Abbreviations. I-S = standard information-seeking; D-initial = discourse-initial; f/u= follow-up

---

## 2.2 Completing the paradigm: Complement Alternative Questions

Contrary to PQs and NAQs, virtually unexplored are Complement Alternative Questions, illustrated in (11):

- (11) a. Is the light on or *off*? (=5)  
 b. Did you pass or did you *fail*?  
 c. Are you a child or *an adult*

As a working example, let us consider the CAQ in (12c) within the broader paradigm in (12), which also includes the corresponding PQ (in (12a)) and NAQ (in (12b)). In a model where the light has only two possible states, all these questions share the same two potential resolutions: the light is on; the light is off (i.e., not on). Furthermore, the NAQ and the CAQ have the same Hamblin-style semantic denotation:  $\{p, \neg p\}$ .

- (12) a. *PQ*: Is the light on? (=1)  
 b. *NAQ*: Is the light on or not? (=2)  
 c. *CAQ*: Is the light on or off? (=5)

In light of these similarities, the question naturally arises as to how CAQs behave in discourse; and, more generally, as to how their pragmatic distribution compares to the one of the other members of the triple in (12). To shed

light on this issue, we present three sets of experiments that compare the acceptability of PQs, NAQs and CAQs across such contexts. These studies will also afford the welcome possibility of replicating in a controlled setting the introspective data reported by Bolinger and the subsequent literature on the contrast between PQs and NAQs.

### 3 Experiment 1: Invitations, inferences, and rhetorical questions

This study explores the distribution of NAQs, CAQs and PQs in three special uses described by Bolinger: making an invitation; asking a rhetorical question; and corroborating an inference.

#### 3.1 Methods

##### 3.1.1 Design

Since we are interested in assessing the relative naturalness of such question types within each type of communicative context, we ran separate experiments for each kind of scenario. Items across the three experiments featured a parallel structure. Each trial consisted of a dialogue, preceded by a short description that made clear what the goal of the speaker was. At the end of the dialogue, one participant would ask a question, which came in four levels: PQ, NAQ, CAQ and a control, which was completely unrelated to the topic, and thus served as a negative baseline for the ratings. Below is an example of an item for each experiment.

- (13) a. *Experiment 1A: Invitation*  
Context: Lena has a cavity. She is scared of going to the dentist, so her friend Lisa offers to go with her:  
*PQ:* Do you want me to come to the dentist with you?  
*NAQ:* Do you want me to come to the dentist with you or not?  
*CAQ:* Do you want me to come to the dentist with you or would you rather go alone?  
*Control:* What's the name of your landlord?
- b. *Experiment 1B: Rhetorical*  
Context: A football player complains that the drills in practice are too hard and asks for a day off. The coach wants to remind him that going through difficulties is an essential part of becoming a

## Alternative questions

good player and intends to deny the request. He thus asks:

*PQ*: Come on! Show me some effort! Are you a child?

*NAQ*: Come on! Show me some effort! Are you a child or not?

*CAQ*: Come on! Show me some effort! Are you a child or an adult?

*Control*: Come on! Show me some effort! Is there any soda in the fridge?

### c. *Experiment 1C: Inference*

Right after getting up, Fred is shivering. Brenda wonders what might have happened. She thus asks:

*PQ*: It's so cold in here. Did you leave the window open all night?

*NAQ*: It's so cold in here. Did you leave the window open all night or not?

*CAQ*: It's so cold in here. Did you leave the window open all night or did you close it?

*Control*: It's so cold in here. Did you see that there is a shoe rack by the front door?

### 3.1.2 Procedure and statistical analysis

Each experiment included 4 conditions and 24 items distributed into four lists with a Latin Square Design. All lists also included 24 fillers.<sup>5</sup> All experimental items are available in the supplementary file. 48 native speaker of English were recruited on Amazon MTurk and paid \$1.50. At the end of each trial, participants were asked to answer the following question with a value between 1 (completely unnatural) and 7 (perfectly natural): "How natural does the question sound in light of the goal of the speaker?". For each experiment we ran mixed-effects models on the ratings with Question Type as a fixed effect and random intercepts for Subjects/Items using the *lmerTest* package (Kuznetsova, Brockhoff & Christensen 2017). Given the theoretical motivation of the study, we are especially interested in the comparison between NAQ and CAQ within each illocutionary context. In light of this, the factors were treatment coded with NAQs as the reference level. This allows us to verify if, and to what extent, NAQs differ from CAQs on the one hand, and from

<sup>5</sup> The fillers contained dialogues with either pragmatically savvy or odd polar questions, where felicity is exclusively grounded in the relevance of the question to the conversation.

PQs — a type of question predicted to be felicitous across these illocutionary contexts — on the other.

### 3.2 Results

Summaries from the models from Experiments 1A-C are reported in Table 2.

---

Level	Exp 1A			Exp 1B			Exp 1C		
	$\beta$	SE	$p$	$\beta$	SE	$p$	$\beta$	SE	$p$
Int.	5.40	0.13	<.0001	4.69	0.17	<.0001	5.32	0.15	<.0001
PQ	0.98	0.11	<.0001	1.05	0.12	<.0001	0.81	0.13	<.0001
CAQ	0.72	0.11	<.0001	0.83	0.12	<.0001	-0.01	0.13	.93
Con.	-2.30	0.11	<.0001	-1.85	0.12	<.0001	-2.14	0.13	<.0001

---

**Table 2** Mixed effect model summary for Experiment 1A-C. Intercept: NAQ

---

We now present the results of each experiment separately.

#### 3.2.1 Experiment 1A: Invitations

The results from Experiment 1A are plotted in Figure 1.

The control condition turned out to be considerably worse than any other condition. As for the critical question types, PQs and CAQs are significantly better than NAQs.

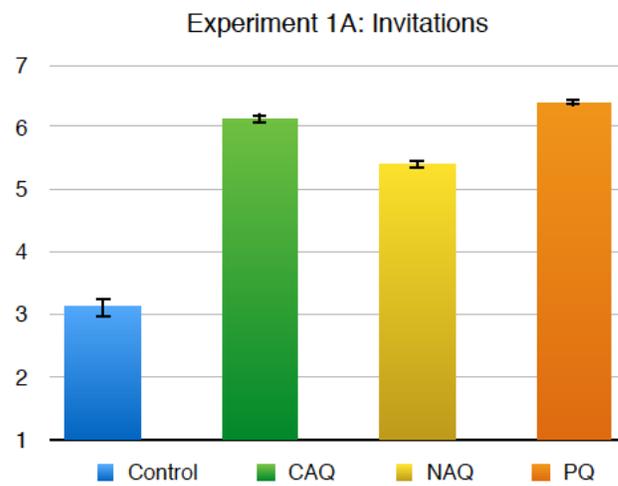
#### 3.2.2 Experiment 1B: Rhetorical

The results from Experiment 1B are plotted in Figure 2.

Again, the control condition turned out to be worse than the others. As for the critical conditions, PQs and CAQs are rated more natural than NAQs; no difference emerges between PQs and CAQs.

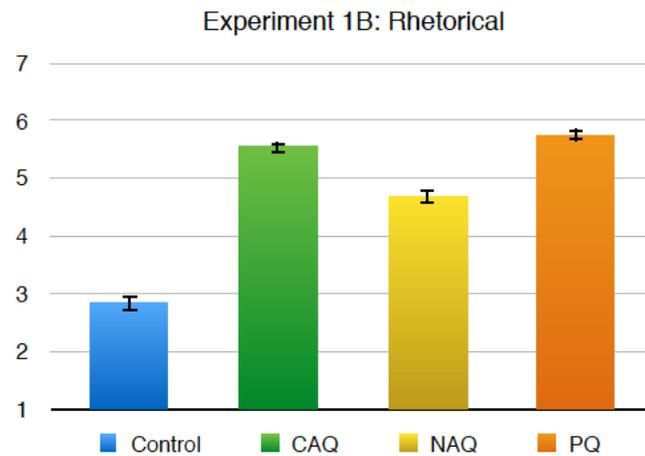
#### 3.2.3 Experiment 1C: Inferences

The results from Experiment 1C are plotted in Figure 3.



**Figure 1** Average naturalness ratings for invitations. Error bars represent standard error.

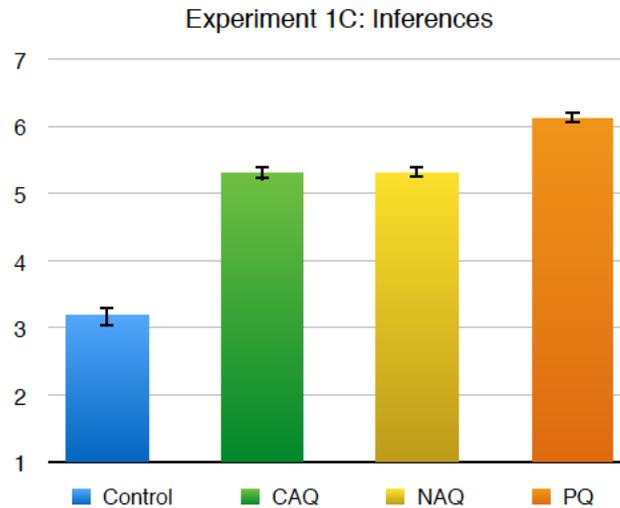
---



**Figure 2** Average naturalness ratings for rhetorical. Error bars represent standard error.

---

Again, the control condition turned out to be worse than the others. As for the critical conditions, PQs are rated more natural than NAQs and CAQs; no difference emerges between NAQs and CAQs.



**Figure 3** Average naturalness ratings for inferences. Error bars represent standard error.

---

### 3.3 Discussion

PQs are the highest rated choice and are significantly better than NAQs across all the tested uses, confirming the observations from the literature. However, NAQs and CAQs do not pattern together across the tested uses. They feature equal (un)naturalness only in Inferences; in Invitations and Rhetorical, instead, CAQs are rated more natural than NAQs.

## 4 Experiment 2: Cornering, Part 1

Experiment 2 compares PQs, NAQs and CAQs in discourse-initial position.

### 4.1 Methods

#### 4.1.1 Design

Two factors were crossed in a 2x4 design. Each trial consisted of a dialogue, at the end of which one participant would ask a standard information-seeking question. The first factor manipulated the moment of the dialogue in which the question is asked, with two levels: ask for the first time, in which the

## Alternative questions

question is asked discourse-initially; and ask-again, in which the question is asked for the third time, after the first two attempts failed to elicit a response. The second factor manipulated the type of question and came in four levels: PQ, NAQ, and CAQ, and a control, which was again completely unrelated to the topic, and thus expected to receive the lowest rating. Below is an example of an item set.

(14) a. *Ask first time*

Context: Mark checks in at a hotel. After the receptionist hands him the room keys, the following exchange ensues:

Receptionist: Sir, would you like to have breakfast directly served in your room?

Mark: Is there a charge for it? PQ

Is there a charge for it or not? NAQ

Is there a charge for it or is it free? CAQ

Is there cable tv in the room? Control

b. *Ask again*

Context: Mark checks in at a hotel. After the receptionist hands him the room keys, the following exchange ensues:

Receptionist: Sir, would you like to have breakfast directly served in your room?

Mark: Is there a charge for it?

Receptionist: It's a great service. Our customers love it.

Mark: Ok, but is there a charge for it?

Receptionist: You can also order food from the special menu.

Mark:

Is there a charge for it? PQ

Is there a charge for it or not? NAQ

Is there a charge for it or is it free? CAQ

Is there cable tv in the room? Control

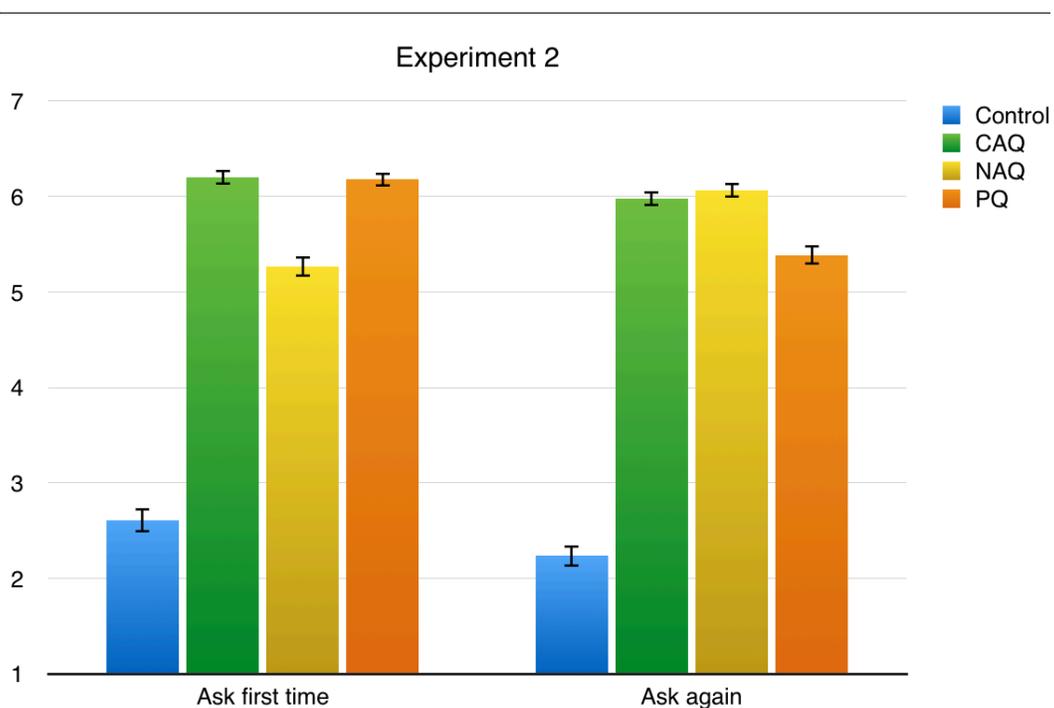
### 4.1.2 Procedure and statistical analysis

Each subject saw 24 experimental items. The conditions were crossed in a Latin Square Design. 96 native speakers of English were recruited on Mechanical Turk and paid \$2. 5 participants were excluded for failing to complete the task. At the end of each trial, participants were asked to answer

the following question with a value between 1 (completely unnatural) and 7 (perfectly natural): “How natural does the question sound in light of the goal of the speaker?”<sup>6</sup> All items were presented in written form on a screen. We ran a linear mixed-effects model on the ratings of questions, with Question Type and Context as fixed effects and random intercepts and slopes for Subjects/Items, using the *lmerTest* R package. The factors were treatment coded, with NAQs and Ask-first as reference levels.

## 4.2 Results

The results are plotted in Figure 4 below.



**Figure 4** Average naturalness ratings for Experiment 2.

<sup>6</sup> The fillers contained dialogues with a polar question asked in discourse-initial position, as well as dialogues in which the question was repeated multiple times, with the speaker explicitly signaling frustration and commenting on the listener’s lack of cooperation. All fillers were expected to be felicitous.

## Alternative questions

The model reveals a main effect of Context ( $\beta=0.79$ ,  $se= 0.12$ ,  $p <.0001$ ), main effects of Question Type (PQ:  $\beta=0.91$ ,  $se= 0.12$ ,  $p <.0001$ ; CAQ:  $\beta= 0.94$ ,  $se= 0.12$ ,  $p <.0001$ ; Control:  $\beta= - 2.31$ ,  $se= 0.12$ ,  $p <.0001$ ), and significant interactions Context\*Question Type (Repeat\*PQ:  $\beta= - 1.59$ ,  $se= 0.18$ ,  $p <.0001$ ; Repeat\*CAQ:  $\beta= - 1.03$ ,  $se= 0.18$ ,  $p <.0001$ ; Repeat\*Control:  $\beta= - 1.17$ ,  $se= 0.18$ ,  $p <.0001$ ). These results suggest that the status of PQs, CAQs and NAQs is crucially modulated by the discourse context, as indicated by the interactions between Question Type and Context. To compare NAQs and CAQs at each moment of the dialogue, we subsequently ran separate mixed-effects models for each level of the Context factor with Question Type as a fixed predictor and random intercepts for Subjects/Items. Results from each model are reported in Table 3.

---

Level	Ask first time			Ask again		
	$\beta$	SE	$p$	$\beta$	SE	$p$
Intercept	5.26	0.12	<.0001	6.06	0.12	<.0001
PQ	0.91	0.12	<.0001	- 0.67	0.12	<.0001
CAQ	0.95	0.12	<.0001	- 0.08	0.12	.47
Control	- 2.64	0.11	<.0001	- 3.82	0.12	<.0001

---

**Table 3** Mixed-effects models summary for Experiment 2. Intercept: NAQ

CAQs and NAQs patterned differently across these two contexts. When the question is asked for the first time, CAQs are significantly more felicitous than NAQs and as felicitous as PQs; when the question is asked again, no difference emerges between NAQs and CAQs, with both strategies being more felicitous than PQs.

### 4.3 Discussion

Replicating the observations from Biezma (2009), NAQs are more felicitous than PQs when used to ask a question again, while they are less felicitous than PQs when used discourse-initially. By contrast, CAQs show remarkable flexibility: they are as felicitous as PQs discourse-initially and as felicitous as NAQs in re-ask uses.

## 5 Experiment 3: Cornering, Part 2

In Experiment 3 we test whether CAQs, similar to what has been claimed for NAQs, have necessarily discourse-final status.

### 5.1 Methods

#### 5.1.1 Design

Each trial consisted of a dialogue in which one of the speakers would ask three questions, the first of which was always a PQ. Two factors were crossed in a 2x2(+1) design, resulting in 5 conditions. Factor 1 manipulated the type of second question, with NAQ and CAQ as levels; Factor 2 manipulated the type of the third question, with two levels: a question identical to the first PQ (i.e., “match”, abbreviated “M”); and a question different from the first PQ (i.e., “non-match”, abbreviated “NM”). Specifically, we ran two different sub-experiments, which were identical, except for the way in which the non-matching question was constructed. In Expt 3A the non-matching question was a PQ asked with emphatic tone (i.e., all caps);<sup>7</sup> in Expt 3B the non-matching question was a Wh-Question. Moreover, in each sub-experiment we had a control condition consisting of a sequence with a Wh-Question as the first question, a PQ as the second, and a NAQ as the third question. This sequence was predicted to be felicitous (see Biezma 2009). The item below illustrates the whole paradigm across the two sub-experiments.

- (15) *Expt3A*  
*Sophia and Rachel are about to play chess. The following exchange ensues.*
- Sophia: “Are you playing with black?” *PQ*
- Rachel: “Well, I can’t wait to play”
- Sophia: “Are you playing with black {or not? / or white?”  
*MidQ: {NAQ/CAQ}*
- Rachel: “I’m going to crush you!”
- Sophia: “{Are you playing with black?/ARE YOU PLAYING WITH BLACK?”  
*Match: {M/NM}*

<sup>7</sup> With respect to emphatic questions, participants received the following directions: “For questions WRITTEN IN ALL CAPS, assume they are being asked with a lot of emphasis.”

Alternative questions

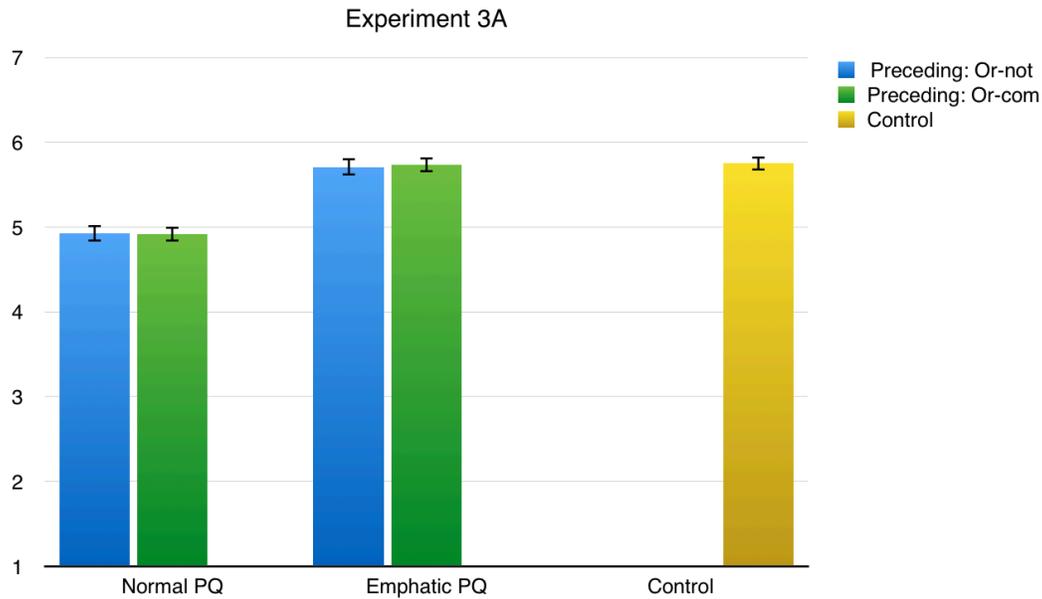
- (16) *Expt3B*  
*Sophia and Rachel are about to play chess. The following exchange ensues.*
- Sophia: “Are you playing with black?” *PQ*  
Rachel: “Well, I can’t wait to play”  
Sophia: ““Are you playing with black {or not? / or white?}”  
*MidQ: {NAQ/CAQ}*
- Rachel: “I’m going to crush you!”  
Sophia: “{Are you playing with black?/What color do you want to play with?}”  
*Match: {M/NM}*
- (17) *Control: same across Expt 3A and 3B*  
*Sophia and Rachel are about to play chess. The following exchange ensues.*
- Sophia: “What color do you want to play with?” *WhQ*  
Rachel: “I can’t wait to play”  
Sophia: “Are you playing with black?” *PQ*  
Rachel: “I’m going to crush you!”  
Sophia: “Are you playing with black or not?” *NAQ*

### 5.1.2 Procedure and statistical analysis

The items were distributed in 5 lists with a Latin Square Design. Each subject saw 30 experimental items, 6 for each condition. Each list also contained 20 fillers, identical to those used in Experiment 2. 48 native speakers of English were recruited in each sub-experiment via Mechanical Turk and paid \$2. At the end of each trial, participants were asked to answer the following question with a value between 1 (“completely unnatural”) and 7 (“perfectly natural”): “How natural does the last question of the conversation sound in light of the goal of the speaker?”. All items were presented in written form on a screen. We ran separate mixed-effects models on the ratings of questions, with MidQ type and Match as fixed effects, and random intercepts and slopes for Subjects/Items using the *lmerTest* package. The factors were treatment coded, with NAQ and Match as reference.

## 5.2 Results

The averages are plotted in Figure 5 and 6 below.



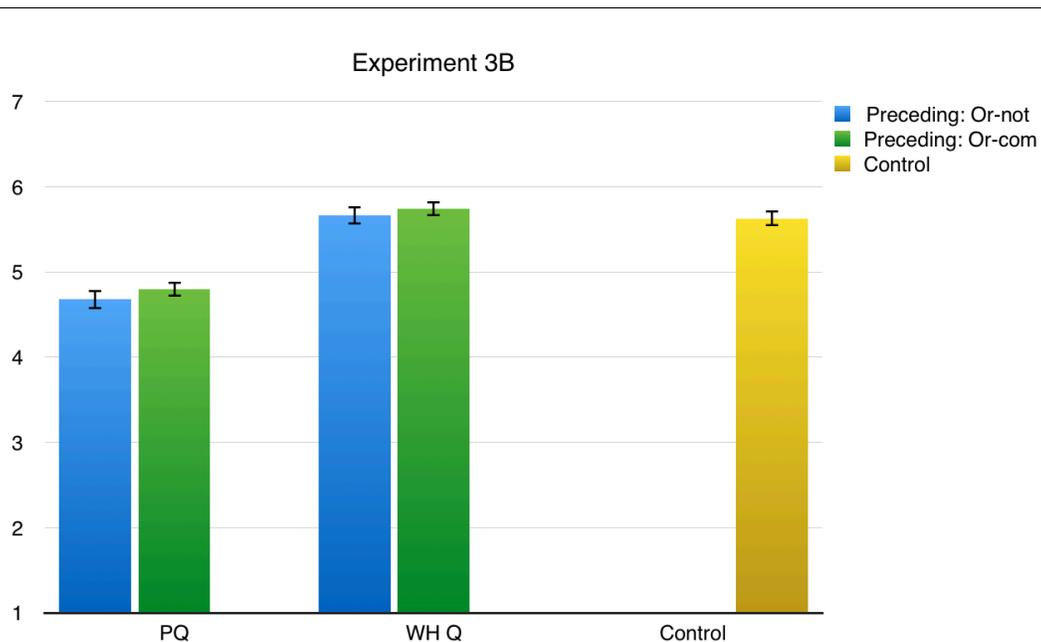
**Figure 5** Average naturalness ratings for Experiment 3A.

Summaries from the mixed-effects models from each experiment are provided in Table 4.

In both cases, we found a main effect of Match, with non-matching final questions rated as more natural than matching ones both when the mismatch is obtained prosodically (via emphasis, in Expt 3A) and syntactically (via a Wh-Question, in Expt 3B). Moreover, neither model shows a main effect of Mid-Question, or an interaction between Mid-Question Type and Match: whether the intermediate question is a NAQ or a CAQ has no implications for the naturalness of the subsequent question.

## 5.3 Discussion

These findings suggest that the status of a follow-up question in a conversation is not determined by whether the preceding move is a NAQ or a CAQ.



**Figure 6** Average naturalness ratings for Experiment 3B.

Specifically, neither NAQ nor CAQ can be followed up on by PQs that had already been used; by contrast, both NAQs and CAQs can be followed up on by new moves that had not been used in previous discourse, failing to provide evidence to Part 2 of Cornering.

#### 5.4 Interim summary

Three sets of experiments compared the distribution of PQs, NAQs, and CAQs. Experiment 1 confirms the hard split between PQs and NAQs across Bolinger's special uses, while suggesting that CAQs do not behave homogeneously in such environments: they are more felicitous than NAQs — and comparatively as felicitous as PQs — with invitations and rhetorical questions; but they are less felicitous than PQs — and comparatively as infelicitous as NAQs — for corroborating inferences. Experiment 2 also replicates the contrast between PQs and NAQs in discourse-initial contexts, while suggesting that, once again, CAQs have an independent distribution: similarly to PQs, they can be used to raise an issue discourse-initially; but similar to

Level	Expt 3A			Expt 3B		
	$\beta$	SE	$p$	$\beta$	SE	$p$
Intercept	4.92	0.21	<.0001	4.67	0.16	<.0001
CAQ	-0.05	0.08	0.94	0.12	0.11	.27
Non-match	0.78	0.22	<.001	0.99	0.11	<.0001
CAQ*Non-match	0.03	0.11	0.76	- 0.04	0.15	0.77

**Table 4** Mixed-effects models summary for Experiment 3. Intercept: NAQ and Match

NAQs, they can be used to return to an issue that had been previously raised in discourse. Finally, Experiment 3 suggests that both CAQs and NAQs can be followed up upon by further questioning moves that have not already been used in previous discourse – including Wh-Questions and emphatic PQs. This finding suggests that Part 2 of Cornering is not an inherent property of NAQs and that, more broadly, Part 1 and Part 2 of Cornering are not part and parcel of the same phenomenon, contrary to previous claims (see Section 6).

## 6 Previous accounts: Summary and predictions

Several lines of explanation have been proposed to derive the different pragmatic distribution of NAQs and PQs. In Section 6 we summarize such proposals and suggest that, in their current form, they don't appear to be immediately able to capture the behavior of CAQs. In Section 7, we proceed to sketch out several ways in which these analyses could be revised and extended accordingly in future work.

### 6.1 Same semantics, but different pragmatics

van Rooy & Šafářová (2003) capture the different distribution of PQs and NAQs in Bolinger's special contexts in terms of their pragmatics. Semantically, PQs and NAQs are assumed to both denote the set containing a proposition and its complement:  $\{p, \neg p\}$ . Pragmatically, the speaker's decision to explicitly mention one vs. two alternatives conveys a crucial difference concerning the *Utility Value* of the two propositions, as defined in (18).

## Alternative questions

PQs, by virtue of pronouncing only one proposition, convey that such an alternative has a higher Utility Value than the unmentioned disjunct. NAQs, by pronouncing both alternatives, signal that both options have equal Utility.

- (18) Given an agent's belief-desire state  $\langle P, U \rangle$  — where  $P$  represents the agent's beliefs and  $U$  (roughly) their preferences —, a proposition has high expected Utility Value if:
- a.  $p$  being true would cause a wide revision of the agent's beliefs, or
  - b.  $p$  being true would bring the agent closer to her goals.

The idea is that the more a proposition is informative (with respect to the agent's beliefs) and/or preferable (with respect to the agent's desires), the higher its Utility. In this perspective, the divergent distribution of the two question types is explained by the fact that Bolinger's special contexts typically presuppose that  $p$  has higher Utility Value than  $\neg p$ . For invitations like (6), the PQ signals that  $p$  — that is, being willing to accept the invitation — has higher Utility Value for the addressee (via (18b)) than  $\neg p$ . NAQs do not convey this asymmetry and thus lead to incongruous pragmatic behavior. Concerning inference-corroborating questions like (7),  $p$  is the alternative that would allow us to explain the relevant state of affairs. As such, it holds greater Utility Value (again via (18b)) than its complement. Again, PQs effectively convey this asymmetry; NAQs, by contrast, don't. Finally, concerning rhetorical uses,  $p$  has higher Utility Value (via (18a)) in that it would lead to a significant leap in the speaker's belief state: since the expectation is that the interlocutor is not crazy, learning that this is instead the case would result in considerable epistemic gain, whereas learning that this is not the case (i.e.,  $\neg p$ ) would have negligible informativity. Also in this case, only a PQ is able to convey the expected prominent status of  $p$ .

If we extend this account to cover CAQs, the prediction is that, by also mentioning both alternatives, these questions should signal that both alternatives have equal Utility, and thus be infelicitous across these contexts.

However, this is not borne out: while CAQs are less felicitous than PQs in inference-corroborating contexts, they are crucially more felicitous than NAQs — and comparatively as felicitous as PQs — in invitations and rhetorical questions. This suggests that, as it stands, this account cannot be straightforwardly extended to cover the distribution of CAQs as well.

## 6.2 Line of analysis 2: Different semantics, hence different pragmatics

The second line of analysis, first outlined by Biezma (2009) and Biezma & Rawlins (2012) and further developed by Biezma & Rawlins (2017), holds that PQs and NAQs conventionally encode different meanings. PQs denote the singleton  $\{p\}$ , leaving open implicit, contextually salient possibilities besides  $p$ . NAQs, by contrast, denote the two-membered set  $\{p, \neg p\}$ , which semantically *exhaustifies* the space of possibilities, leaving no other option available to the addressee. Biezma & Rawlins (2012) illustrate this difference in a context in which there are three epistemically possible meals: pasta, stew and fish. While the PQ in (19) puts forward one of the alternatives and leaves other (implicit) alternatives open to the addressee, the NAQ in (20) collapses together all the non-pasta alternatives into a single possibility, making only two alternatives available to the addressee.<sup>8</sup>

(19) Are you making pasta?

*Salient Alternatives in c*

---

You are making pasta  
You are making stew (implicit)  
You are making fish (implicit)

(20) Are you making pasta or not?

*Salient Alternatives in c*

---

You are making pasta  
You are not making pasta

The authors suggest that this semantically driven distinction ultimately determines the different distributional patterns of PQs and NAQs. However, they provide two non-trivially different accounts of the dynamics whereby these restrictions are derived.

### 6.2.1 Exhaustifying the epistemic space

Biezma (2009) and Biezma & Rawlins (2012) directly link the divergent distribution of PQs and NAQs to their different denotation. PQs, by presenting one

<sup>8</sup> The example is taken from Biezma & Rawlins 2012: 400; ex. (63).

among (many) possible alternatives, leave the possibility space wide open, granting the addressee a high degree of freedom in discourse. NAQs instead, by presenting two alternatives that exhaustify the possibility space, force the listener to choose one of these two possibilities, crucially restricting their room for maneuvering. It's precisely this particular way of segmenting the answer space that engenders the insisting flavor of NAQs, ultimately making their use highly constrained. On the one hand, collapsing multiple possible alternatives into a binary choice is motivated, and indeed pragmatically felicitous, to force the resolution of an issue that has been already raised in discourse: this explains why NAQs are a congruous strategy to follow-up on a polar question that hasn't been answered, as in (4) and (10) above. On the other hand, this move is inappropriate in questions that are typically used discourse-initially — that is, which occur in a context where previous discourse provides no cues that justify collapsing multiple alternatives into a binary choice, and multiple options are thus expected to remain open. Such moves include information-seeking questions raising an issue for the first time — see Part 1 of Cornering — as well as different special uses such as invitations, inference-corroborating and rhetorical questions, all of which are also typically effective only when used discourse-initially (see Biezma 2009: 52).

Finally, the distinctive semantic nature of NAQs similarly explains the observation that these moves do not license follow-up questions — that is, Part 2 of Cornering. Because the alternatives presented by NAQs exhaustify the answer space, no further questions can be used to zero in further on the question under discussion, making an answer on the part of the addressee the only possible move.

This account would lead us to predict that CAQs should align with NAQs — or, at least, behave more closely to NAQs than to PQs. By presenting complementary alternatives, these questions should indeed induce a similar effect of limiting the addressee's freedom, and thus should also require a valid justification to be used. However, the experimental findings do not support this idea. As Experiment 1 shows, CAQs are felicitous in at least a subset of the contexts in which NAQs are not; moreover, as Experiment 2-3 suggest, CAQs can be used discourse-initially, and can license follow-up questions.

### 6.2.2 Using negation as a label

In subsequent work, Biezma & Rawlins (2017) integrate Biezma (2009) and Biezma & Rawlins (2012)'s analysis, individuating two factors as conspiring to determine the discourse effects of a question: (a) the partition of the epistemic space induced by the disjuncts, as also put forward in their previous accounts; and (b) the particular *description* with which the speaker chooses to label each alternative — an element instead not invoked in previous proposals. In the case of NAQs, the second disjunct not only collapses or “bundles” multiple relevant epistemic possibilities into one discourse possibility; by using a negator in the description, it also groups all such possibilities by the characteristic of *not* being the first alternative. This, in turn, is argued to crucially convey a relevance asymmetry between the two disjuncts: the speaker signals that they are merely interested in the content proposition  $p$ , and thus one-sidedly eliminates alternatives bundled under “or not” from future relevance to the discourse.

To highlight the importance of ingredient (b), Biezma & Rawlins (2017) contrast NAQs with the discourse effects induced by questions such as (21) (Biezma & Rawlins 2017: 37, ex. (97)).

- (21) Do you want to go to the Green Door for dinner or to a place with meat?

In this case, potentially multiple restaurants are bundled under the label “a place with meat”. Similar to what happens with a NAQ, the second disjunct semantically collapses the options available to the responder, ultimately restricting their options. However, contrary to a NAQ, the options bundled under the second disjunct are crucially tagged with an “informative” label — one that doesn't merely identify them as *not* belonging to the first option, but describes them as sharing a positive, independent property. As a result, both disjuncts are said to remain relevant to the conversation, and the asymmetry that characterizes NAQs is claimed not to arise here.

On this view, Part 2 of Cornering is explained as follows: by eliminating the newly bundled alternative from future discourse relevance via a negative description, the speaker communicates that all they care about at this point is  $p$ , thus leaving the interlocutor with no other strategy than answering  $?p$ . As for Part 1, Biezma & Rawlins (2017) suggest that speakers need a reason to bundle all alternatives other than  $p$  and eliminate them from relevance for the rest of the conversation. This reason is clearly inferrable in repeat

contexts where the speaker already expressed interest in  $p$ , making NAQs felicitous; however, no reason is inferrable in discourse-initial contexts, when the issue is being raised for the first time. It follows that NAQs are unmotivated, and hence infelicitous, in this particular situation.

Concerning CAQs, Biezma & Rawlins (2017)'s account crucially makes different predictions than the proposal outlined in 6.2.1. If NAQs' restricted felicity is jointly driven by the exhaustivity of the alternatives and the act of describing the second disjunct as "not  $p$ ", it follows that CAQs like (11) should not feature the same distribution, since they spell out the second disjunct with a different proposition. This leads us to predict that CAQs should be felicitous in discourse-initial contexts, including standard information-seeking uses as well as invitations, inferences and rhetorical uses.<sup>9</sup> This prediction is partially borne out: CAQs are confirmed to be felicitous to make invitations and ask rhetorical questions, as well as to raise issues discourse-initially; the puzzle remains, however, as to why CAQs are not felicitous to corroborate inferences. A second prediction is that CAQs should allow for follow-up questions in the conversation. While this prediction is confirmed in the results, the findings from Experiment 3 are nevertheless problematic for this account: the fact that *both* NAQs and CAQs can be followed by a subsequent questioning move, indeed, does not support the empirical observation that NAQs are necessarily discourse final; more broadly, it also crucially questions the idea that Part 1 and Part 2 of the Cornering effect (see Section 2) are inherent properties of NAQs that can be jointly traced back to the same underlying phenomenon.

A final point merits discussion. While most of our experimental items feature CAQs with a *distinct* complementary label in the second disjunct, as in (11) above, some items describe the second disjunct with a *semantically bleached* complementary label, constructed by applying *else* or *other* to the property in the first disjunct, as in (22):

- (22) Are you going camping for spring break or are you *doing something else*?

While this sub-type of CAQ does not make explicit use of negation, it nevertheless spells out the second disjunct in terms of the first one. As such, one could make the case that, like NAQs, these CAQs also fail to provide an in-

<sup>9</sup> While Biezma & Rawlins (2017) primarily focus on standard information-seeking questions, their new analysis is meant to apply to a broader range of uses. See footnote 13 in Biezma & Rawlins 2017 for a discussion of rhetorical questions, for example.

formative cover term for the second disjunct, and should thus be subject to the same restrictions as NAQs. While our experiments were not designed to systematically explore the differential distribution of these two subflavors of CAQs, a cursory look at the ratings does not reveal a difference in this direction (see supplementary file), raising a further puzzle for this account. We leave further investigation of this issue to future work.

### 6.3 Experimental results and previous accounts: A summary

The picture emerging from the studies is one in which no proposal alone, as it stands, is able to fully derive the entire distributional spectrum of the question moves considered above. Table 5 reports the findings, comparing them with the predictions that can be extrapolated for CAQs. The critical findings are highlighted in red.

CONTEXT	vRS2003			Bo9/B&R12		
	PQ	NAQ	CAQ	PQ	NAQ	CAQ
Invitation	✓	#	#	✓	#	#
Inference	✓	#	#	✓	#	#
Rhetorical	✓	#	#	✓	#	#
I-S: D-Initial	NA	NA	NA	✓	#	#
I-S: Licenses f/u	NA	NA	NA	✓	#	#
CONTEXT	B&R 2017			Expt 1-3		
	PQ	NAQ	CAQ	PQ	NAQ	CAQ
Invitation	✓	#	✓	✓	#	✓
Inference	✓	#	✓	✓	#	#
Rhetorical	✓	#	✓	✓	#	✓
I-S: D-Initial	✓	#	✓	✓	#	✓
I-S: Licenses f/u	✓	#	✓	✓	#/✓	#/✓

**Table 5** Predicted distribution of CAQs. vRS 2003 = van Rooy & Šafářová 2003; Bo9 = Biezma 2009; B&R12 = Biezma & Rawlins 2012; B&R 2017 = Biezma & Rawlins 2017. I-S = standard information-seeking; D-Initial = discourse-initial; f/u= follow-up

## 7 Revising the proposals: Several suggestions

The observation that CAQs' distribution is not successfully predicted by the proposals in the literature calls for a refinement of such accounts, both with respect their explanation of the divergent behavior of NAQs and PQs, and the broader principles that they rely on to model the pragmatic effects of different types of questions. While proposing a full revision goes beyond the scope of this article, we nevertheless deem it worthy to sketch out two areas that emerge as especially relevant for future endeavors in this direction.

### 7.1 Special uses and illocutionary goals

The first area concerns the split behavior of CAQs in Bolinger's special uses. As discussed above, this result is puzzling: while the infelicity is NAQs across these contexts is predicted by all available accounts, no proposal predicts the observed asymmetry between inference-corroborating contexts vs. invitations and rhetorical questions. To shed light on this issue, we'd like to highlight a property according to which the latter might be indeed taken to form a natural class to the exclusion of the former. Specifically, in both invitations and rhetorical questions *both* alternatives identify worlds that are compatible with some objective of the speaker in the conversation, though for different reasons.

In invitations, the first alternative points out a course of action that, according to the speaker, would contribute to the well-being of the interlocutor (see also van Rooy & Šafářová 2003); the second alternative, however, serves an important face-saving purpose, as it contributes to minimizing imposition on the addressee, sparing them the burden of feeling hard-pressed to accept the invite (see also Brown & Levinson 1987 theory of *politeness*). In rhetorical uses, the first alternative is highly unlikely/surprising, and would thus entail a substantive informational gain for the interlocutor (see van Rooy & Šafářová 2003); the second one identifies worlds in line with the speaker's desires.<sup>10</sup> By contrast, a similar observation does not seem to hold for inference-corroborating questions. Here, the first disjunct crucially yields a viable explanation of the observed state of affairs and is thus fully in line with the speaker's objective of shedding light on an otherwise inexplicable fact; how-

<sup>10</sup> This also holds for rhetorical CAQs that are biased towards a positive answer. It is unclear, however, how previous accounts would derive the felicity of NAQs in such contexts (see footnote 2).

ever, it is difficult to see how the second alternative could actually fulfill any relevant purpose, contrary to the cases discussed above: it does not enhance politeness, and it does not identify any particularly desirable state of affairs.

In conclusion, the emerging trend is one in which CAQs appear to be felicitous in situations in which both disjuncts identify a state of affairs that, for a particular reason, would serve the interlocutors' purposes in the conversation; this suggests that, more generally, there is much to gain from considering from a closer angle the fine-grained differences between the specific goals behind such question moves, and how such goals can be achieved or undermined by the question form used by the speaker. While we leave the implementation of this idea to further work, we believe that both [van Rooy & Šafářová \(2003\)](#)'s notion of Utility Value and [Biezma & Rawlins \(2017\)](#)'s notions of bundling and discourse relevance could be viable analytical tools to refine the current accounts in this direction.

## 7.2 Re-thinking Cornering

The second area of future work highlighted by our results revolves around the understanding of the Cornering Effect previously invoked for NAQs. As can be recalled, the results from Experiment 3 call into question the claim that NAQs are necessarily discourse final — that is, part 2 of Cornering. Looking ahead, this finding thus calls for a revised account of Cornering — one that departs from previous accounts that treat it as a unified phenomenon, and instead pinpoints what independent discourse constraints underlie each of its components separately. We would like to highlight two domains of inquiry that could inform further work on the topic.

Concerning Part 2, one possibility suggested by the results of Experiments 3 is that whether a question could license a further subquestion is not tied to its particular form, but might instead depend on the nature of the follow-up. More specifically, a hypothesis to be tested in further work could be that a follow-up question cannot be used if the same question has already been unsuccessfully used and abandoned in previous discourse. Crucially, such a possibility would be consistent not only with the experimental result, but also with the more general observation that re-using unsuccessful discourse moves generally leads to infelicitous pragmatic sequences also outside the domain of questions: for example, it is odd to repeat a command after the previous one failed to achieve the intended goal.

## Alternative questions

- (23) A: Stop playing!  
B: [Keeps playing]  
A: Hey, can you stop playing?  
B: [Keeps playing]  
a. A: #Stop playing!  
b. A': I told you to stop playing!

A narrower possibility, also consistent with the results, could be that PQs in their plain — i.e., non-emphatic — form cannot follow either NAQs or CAQs, regardless of what the structure of the preceding discourse is. To rule out this possibility, one would need to look into the status of sequences in which a NAQ/CAQ is followed by a PQ *and* a PQ has not been used in previous discourse — e.g., WhQ → NAQ/CAQ → PQ. If the ban on repetition is indeed responsible for explaining the results from Experiment 3, we would predict that such sequences should indeed be felicitous.<sup>11</sup>

As for Part 1 of Cornering, our results support [Biezma & Rawlins \(2017\)](#)'s idea that the description of the second alternative matters: despite logical equivalence, NAQs with negation in the second disjunct are infelicitous discourse-initially whereas CAQs with a positive description are felicitous in these contexts. The question remains, however, as to why negation matters. While [Biezma & Rawlins \(2017\)](#) explain this in terms of the action of cutting off the bundled alternatives from future relevance, we would like to suggest the alternative possibility that the effect of negation in NAQs might be ultimately linked to polarity focus. In particular, since alternative questions in general carry focal stress on each disjunct ([Bartels 1999](#), [Han & Romero 2004](#)), negation in NAQs can be assumed to be focus marked. Then, if one assumes — following [Roberts \(2012\)](#) — that focus must always be licensed with respect to the Question under Discussion (QUD), the infelicity of discourse-initial NAQs may be attributed to the lack of an appropriate preceding QUD. If adequately developed, this idea would come with the additional benefit of unifying the discourse-initial infelicity of NAQs like (24b) with that of PQs with focus on the auxiliary like (24c), which also have been argued to involve focus on the polarity head ([Wilder 2013](#), [Goodhue 2018](#)):<sup>12</sup>

<sup>11</sup> We are grateful to an anonymous reviewer for this suggestion.

<sup>12</sup> Enforcing that focus in an utterance be licensed by the QUD is used by [Roberts \(2012\)](#) to explain, among other things, the felicity of the question sequences (i) and the infelicity of the question sequences in (ii). An alternative analysis, proposed among others in [Han & Romero 2004](#) and pointed out by an anonymous reviewer, would be to assume that focus within each disjunct of a run-of-the-mill AltQ can be licensed by the other disjunct, without

- (24) A: Jane got a new car!  
 a. B: Is it automatic?  
 b. B: #Is it automatic or NOT?  
 c. B: #IS it automatic?

## 8 Conclusion

The studies presented in this article provide an empirical characterization of the pragmatic behavior of CAQs, a hitherto unexplored type of question; these results, in turn, call for a more refined characterization of how logical, pragmatic and discourse factors differentially inform the distribution of question strategies with similar resolution conditions and/or denotations. We hope that this article serves as a first step towards accomplishing this project.

## References

- Bartels, Christine. 1999. *The intonation of English statements and questions*. New York, NY: Routledge.
- Biezma, Maria. 2009. Alternative vs polar questions: The cornering effect. *Semantics and Linguistic Theory (SALT)* 19. 37-54. <https://doi.org/10.3765/salt.v19i0.2519>.
- Biezma, Maria & Kyle Rawlins. 2012. Responding to alternative and polar questions. *Linguistics & Philosophy* 35(5). 361-406. <https://doi.org/10.1007/s10988-012-9123-z>.
- Biezma, Maria & Kyle Rawlins. 2017. Or what? *Semantics and Pragmatics* 10(16). 1-44. <https://doi.org/10.3765/sp.10.16>.

---

any reference to the QUD. However, under the latter approach, the contrast between (i.b) and (ii.b) is unexpected. For discussion of examples like (iii), which seemingly do not license focus via the QUD, see Roberts' (2012: 53 ff.) suggestion that metalinguistic use of prosody is involved.

- (i) a. Who called Amy? Did JOHN<sub>F</sub> call Amy?  
 b. Who called Amy? Did JOHN<sub>F</sub> or PAUL<sub>F</sub> call Amy?
- (ii) a. #Who called Amy? Did John call AMY<sub>F</sub>?  
 b. #Who called Amy? Did John call AMY<sub>F</sub> or SUSAN<sub>F</sub>?
- (iii) An AMERICAN<sub>F</sub> farmer was talking to a CANADIAN<sub>F</sub> farmer.

Alternative questions

- Bolinger, Dwight. 1978. Yes—no questions are not alternative questions. In Henry Hiz (ed.), *Questions*, 87–105. Dordrecht: Springer. [https://doi.org/10.1007/978-94-009-9509-3\\_3](https://doi.org/10.1007/978-94-009-9509-3_3).
- Brown, Penelope & Stephen C. Levinson. 1987. *Politeness*. Cambridge, MA: Cambridge University Press.
- Caponigro, Ivano & Jon Sprouse. 2007. Rhetorical questions as questions. *Sinn und Bedeutung* 11. 121–133. <https://doi.org/10.18148/sub/2007.v11i0.635>.
- Goodhue, Daniel. 2018. *On asking and answering biased polar questions*. Montreal, QC: McGill University dissertation.
- Hamblin, Charles L. 1973. Questions in Montague grammar. *Foundations of language* 10(1). 41–53.
- Han, Chung-hye & Maribel Romero. 2004. Disjunction, focus and scope. *Linguistic Inquiry* 35(2). 179–217. <https://doi.org/10.1162/002438904323019048>.
- Kuznetsova, Alexandra, Per B. Brockhoff & Rune Haubo Bojesen Christensen. 2017. *lmerTest* package: Tests in linear mixed effects models. *Journal of Statistical Software* 82(13). 1–26. <https://doi.org/10.18637/jss.v082.i13>.
- Roberts, Craige. 2012. Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics* 5(6). 1–69. <https://doi.org/10.3765/sp.5.6>.
- van Rooy, Robert & Marie Šafářová. 2003. On polar questions. *Semantics and Linguistic Theory (SALT)* 13. 292–309. <https://doi.org/10.3765/salt.v13i0.2887>.
- Wilder, Chris. 2013. English ‘emphatic do’. *Lingua* 128. 142–171. <https://doi.org/10.1016/j.lingua.2012.10.005>.

Andrea Beltrama  
Department of Linguistics  
University of Pennsylvania  
3401 Walnut Street — 300C  
Philadelphia, PA, USA 19104  
[beltrama@sas.upenn.edu](mailto:beltrama@sas.upenn.edu)

Erlinde Meertens  
Department of Linguistics  
University of Konstanz  
Universitaetsstrasse 10, Room H146  
Konstanz, Germany, 78457  
[erlindemeertens@protonmail.com](mailto:erlindemeertens@protonmail.com)

Maribel Romero  
Department of Linguistics  
University of Konstanz  
Universitaetsstrasse 10, Room G222  
Konstanz, Germany, 78457  
[maribel.romero@uni-konstanz.de](mailto:maribel.romero@uni-konstanz.de)