

Alternative Questions between Logic and Discourse. An experimental study

Andrea Beltrama, Erlinde Meertens and Maribel Romero
University of Konstanz

1 Introduction

The current paper aims at comparing the discourse properties of these three question types: **(i)** Polar Questions, which mention only one possible alternative (henceforth, **PQs**); **(ii)** Negative Alternative Questions, which mention one alternative and its negation (henceforth **NAQs**); and **(iii)** Complement Alternative Questions (henceforth **CAQs**), which mention one alternative and its complement.

- (1) **Context:** I heard you have a new puppy!
- a. Is it a male? Polar Question (PQ)
 - b. Is it a male or not? Negative Alternative Questions (NAQ)
 - c. Is it a male or a female? Complement Alternative Questions (CAQ)

Based on results from two rating studies, we argue that each strategy has a distinct pragmatic distribution, which cannot be reduced to the other two. Specifically, NAQs are restricted to contexts in which a question is asked again; PQs are felicitous to ask a question for the first time; and CAQs are felicitous in both cases, emerging as the maximally flexible strategy among these three. These results do not align with the predictions made by accounts that seek to explain the discourse properties of Polar vs Alternative Questions on the basis of pragmatic (e.g., highlighting/prominence, van Rooij & Šafářová 2003) or semantic factors (e.g., exhaustivity/exclusivity, Biezma 2009) alone. Rather, they underscore the need for a more fine-grained theory of how these components interact to determine the message conveyed by a question, in line with what has been suggested by Biezma & Rawlins (To Appear). The paper is divided as follows. Section 2 summarizes previous empirical observations on the different discourse properties of polar and alternative questions; section 3 reviews the two major lines of analysis proposed to account for these differences; section 4 and 5 describe our experiments; section 6 provides a general discussion of the studies; section 7 concludes.

2 Polar vs Negative Alternative Questions: previous observations

It has been argued that questions with seemingly similar semantic content have significantly different pragmatic properties. In particular, Bolinger (1978) observed that PQs tend to have a broader distribution than the corresponding NAQs. For example, PQs have been reported to be more felicitous than NAQs in many non-canonical uses – e.g., when used to make invites, draw inferences, or pose rhetorical questions.

- (2) **Invites:**
- a. ✓Do you want something to drink?
 - b. # Do you want something to drink or not?
- (3) **Inferences:**
- a. A: I just saw David.
B: ✓Is he back from Toronto?
 - b. A: I just saw David.
B: # Is he back from Toronto or not?
- (4) **Rhetorical questions:**
- a. ✓Are you crazy?
 - b. #Are you crazy or not?

Concerning info-seeking contexts, Biezma (2009) observed NAQs are a felicitous strategy to ask a question for the first time, while PQs are excluded in this context.

- (5) **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.
- a. **You:** # Are you making pasta or not?
 - b. **You:** ✓Are you making pasta?¹

By contrast, NAQs are more felicitous than PQs when the speaker intends to re-ask an info-seeking question that previously went unanswered:

- (6) **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?
John: (Silence and dubitative faces)
- a. **You:** ✓Are you making pasta or not?
 - b. **You:** # Are you making pasta?

The contrast in the distribution of these two question types is summarized in Table 1 below.

Question type	Invite	Inference	Rhetorical	"Info-seek: 1st"	"Info seek: again"
PQ	✓	✓	✓	✓	#
NAQ	#	#	#	#	✓

Table 1: Comparative pragmatic distribution of NAQs and PQs

¹Crucially, this question is infelicitous only if prosody and focus structure are exactly the same as in the previously asked question. If one of these elements is different, a PQ is instead a felicitous strategy in this context as well. See (Beltrama *et al.* 2018) for further discussion.

3 Polar vs Negative Alternative Questions: two lines of analysis

To account for the different behavior of these questions, two distinct lines of analyses have been proposed: pragmatics-based accounts and semantics-based accounts. Pragmatic-based accounts argue that PQs and NAQs similarly denote a set containing a proposition and its complement, that is $\{p, \neg p\}$. What makes them different is that PQs, by virtue of pronouncing only one alternative, assign a special pragmatic status to such an alternative, signaling that it has higher Utility Value (van Rooij & Šafářová 2003). By contrast, NAQs, by pronouncing both alternatives, signal that both options are seen as equally important by the speaker. In this perspective, the divergent distribution of the two question types is explained by the fact that a non-canonical question $p?$ is normally asked in contexts in which p has indeed higher Utility than $\neg p$. While PQs, by highlighting p , are congruent with this goal, NAQs, by signaling indifference, are not. For example, when asking a question to check the inference that David is back from Toronto in (3), the PQ highlights p as the alternative that would allow us to confirm our hypothesis and explain the relevant state of affairs; by the same token, when offering someone a drink, as in (2), the PQ signals that p – that is, being willing to accept the invite – has higher Utility value for the hearer than $\neg p$ – that is, declining the invite. In both these cases, using an NAQ would not be able to convey the expected prominent status of p , thus resulting in incongruous pragmatic behavior.

Semantic accounts, by contrast, hold that PQs denote an open list of alternatives, which contains p and other unmentioned alternatives salient in discourse; alternative questions at large, by contrast, denote two exhaustive, mutually exclusive alternatives $\{p, q\}$ (Biezma 2009, Biezma and Rawlins 2012). What makes NAQs special, among alternative questions, is that they present *logically opposite* alternatives, that is $\{p, \neg p\}$. This semantic property imbues NAQs with a flavor of insistence, which Biezma calls *the cornering effect*. Since they offer no middle ground between the two alternatives, they emerge as a good strategy to force the addressee to respond to a previously unanswered question, as in (6); however, they are infelicitous when the illocutionary goal of the speaker is inconsistent with hard-pressing the listener, as is the case with info-seeking questions asked for the first time (5).

By exclusively contrasting PQs and NAQs, such accounts leave an issue open. While it is certainly possible that the restricted distribution of NAQs truly reflects general properties of alternative questions – e.g., be that highlighting or exclusivity/exhaustivity – the possibility remains that such restrictions might be driven instead by a specific property of NAQ. More precisely, it could be the case that the special distribution of NAQs is linked not to the principles above, but to the effects of spelling out the second disjunct via negation – i.e., as $\neg p$ –, rather than as a full complement proposition without negation. We suggest that a viable case study to shed light on this issue is represented by Complement Alternative Questions (CAQ), a type of alternative question that, similarly to NAQs, pronounce two logically opposite alternatives; but, contrary to NAQs, spell out the second alternative in full, as opposed to with "or not".

- (7) a. Is it a boy or a girl?
b. Is it heads or tails?

On the one hand, both the accounts discussed above would predict that CAQs

should feature similar distribution to NAQs. As far as the pragmatic analysis is concerned, CAQs pronounce both alternatives. As such, they should also communicate indifference between the two options, thus emerging as incongruous in non-canonical uses. As far as the semantic analysis is concerned, CAQs pose exclusive, exhaustive and logically opposite alternatives. As such, they should convey the same flavor of insistence as NAQs, and thus be limited to contexts in which the speaker asks the question again. On the other hand, if the distribution of NAQs is instead driven by the specific nature of the "or not" formulation, we predict that CAQs, by virtue of spelling out the second disjunct with a full proposition, should not have the same illocutionary distribution as NAQs. In the current paper, we aim to tease apart these two possibilities by comparing the illocutionary properties of NAQs and CAQs in two rating studies, which we now turn to discuss.

4 Experiment 1: non-canonical uses

In the first study, we compare the distribution of NAQs, CAQs and PQs in non-canonical contexts, that is, questions that are not purely info-seeking. The study has two goals. First, we intend to replicate in a quantitative study the introspective data from Bolinger (1978) and the subsequent literature, which suggested that NAQs are less felicitous than PQs in non-canonical contexts. Second, and most importantly, we aim to compare the behavior of CAQs vs NAQs with respect to the predictions made by previous accounts of the distributional contrast between NAQs and PQs. As can be recalled from the discussion in Section 3, both semantic and pragmatic lines of analysis predict that NAQs and CAQs should be equally infelicitous in non-canonical contexts.

4.1 Methods

4.1.1 Design

Two factors were crossed in a 3x4 design. Each trial consisted of a dialogue, preceded by a short description that made clear what the illocutionary goal of the speaker was (**Factor 1**). Three possible illocutionary goals were manipulated: drawing an inference, making an invitation, and asking a rhetorical question that is biased towards a negative answer. At the end of the dialogue, one participant would ask a question (**Factor 2**), which came in four levels: PQ, NAQ, CAQ and a control, which was completely unrelated to the conversation topic, and thus served as a negative baseline for the conversational exchange. Below is an example of an item for each illocutionary goal.

(8) a. **Inference**

Context: Right before the beginning of spring break, George sees camping equipment all around Joe's house and wonders why it is there. Thinking that Joe might be going camping during the break, George thus asks him:

Are you going camping for spring break?	PQ
Are you going camping for spring break or not?	NAQ
Are you going camping for spring break or are you doing something	

- else? CAQ
 Are you having a good day today? Control
- b. **Invite**
Context: It's very cold outside. Tom has an extra scarf in his backpack and wants to offer it to his friend Mark, who isn't wearing one. Tom thus turns to Mark and asks:
 Hey, do you want a scarf? PQ
 Hey, do you want a scarf or not? NAQ
 Hey, do you want a scarf or are you ok? CAQ
 Hey, do you want a beer? Control
- c. **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 good player and intends to deny the request. He thus asks:
 Are you a child? PQ
 Are you a child or not? NAQ
 Are you a child or an adult? CAQ
 Is there any soda in the fridge? Control

4.1.2 Procedure and Statistical analysis

Each subject saw 24 experimental items, 8 for each context type, together with 24 fillers.² The conditions were crossed in a Latin Square Design. 48 participants were recruited on Mechanical Turk and paid \$1.50 for participation. At the end of each trial, participants were asked to answer the following question with a value between 1 and 7: "How natural does the question sound in light of the goal of the speaker?" "1" indicated a completely unnatural question; "7" indicated a perfectly natural question. All items were presented in written form on a screen. For statistical analysis, for each illocutionary context we ran separate mixed-effects models on the ratings with Question Type as the fixed effect and random intercepts for Subjects and Items. The models were run with the *lmer* package (Kuznetsova *et al.* 2016). 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, we established NAQs as the reference level in the analysis, so as to verify.

4.2 Results

The results are plotted in Figure 1 below.

²The fillers contained dialogues with either pragmatically savvy or pragmatically odd polar questions, where felicity is exclusively grounded in the relevance of the question to the previous conversation.

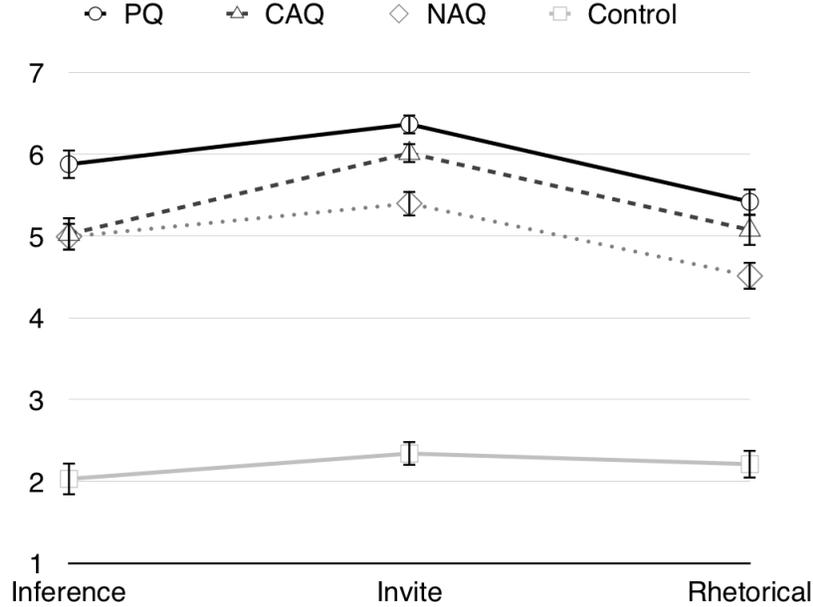


Figure 1: Average naturalness ratings for Experiment 1.

As predicted, the control condition turned out to be highly infelicitous across the board. We therefore removed it from the analysis. As for the other conditions, the models showed that PQs were significantly better than NAQs within each illocutionary context (Inference: $\beta=0.88$, $SE=0.28$, $p < .01$; Invite: $\beta=0.91$, $SE=0.23$, $p < .01$; Rhetorical: $\beta=0.92$, $SE=0.23$, $p < .01$). Remarkably, CAQs and NAQs patterned differently across illocutionary contexts. In Inference, CAQs and NAQs did not differ from one another ($\beta=0.03$, $SE=0.24$, $p > .1$). In Invites and Rhetorical, however, CAQs were rated higher than NAQs (Invites: $\beta=0.56$, $SE=0.23$, $p < .05$; Rhetorical: $\beta=0.56$, $SE=0.25$, $p < .05$).

4.3 Discussion

In Experiment 1, we compared the distribution of three types of questions: Polar Questions (PQs), Complement Alternative Questions (CAQs) and Negative Alternative Questions (NAQs). The study had two goals: replicating the introspective judgments provided in the literature, which reported that PQs had better status than NAQs in non-canonical contexts; and testing the hypothesis that NAQs and CAQs have the same distribution, as predicted by both semantic and pragmatic accounts (see Section 2). Concerning the former goal, the study does confirm that PQs are the highest rated choice and that they are significantly better than NAQs across all the tested non-canonical uses. The observations discussed in the literature are thus supported in a large-scale study. Concerning the latter goal, we observe that NAQs and CAQs do not pattern together across the tested uses. In particular, they feature equal (un)naturalness only in Inferences; in Invites and Rhetorical, instead, CAQs are rated more natural than NAQs. Since NAQs and CAQs both present logically opposite alternatives and pronounce both disjuncts, their different distribution sug-

gests that that the divergence of PQs and NAQs cannot be solely explained in terms of the semantic or pragmatic factors suggested in the previous literature. While in Section 6 we will suggest some avenues to account for these observations, we now turn to the comparison of NAQs and CAQs in info-seeking contexts.

5 Experiment 2: Info-seeking contexts

In this study, we compared the distribution of NAQs, CAQs and PQs in info-seeking contexts. Our goal is to test the prediction that NAQs and CAQs are subject to the same distributional constraints with respect to this illocutionary move as well. As can be recalled, Biezma (2009) observed that NAQs are restricted to contexts in which a question is asked again, and are instead not felicitous to ask a question for the first time; as an explanation for this behavior, she argued that, by asking logically opposite alternatives, NAQs force the hearer to provide an answer, and are thus limited to contexts in which being insistent is justified – i.e., asking a question again. The prediction of this analysis is that CAQs, by virtue of also posing logically opposite alternatives, should feature a similar restriction to ask-again contexts.

5.1 Methods

5.1.1 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 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 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. As in Experiment 1, it came in four levels: PQ, NAQ, CAQ and a control, which was completely unrelated to the conversation topic. Below is an example of an item for each illocutionary goal.

(9) a. **Ask first-time**

Context: Mary runs into Greg on the street. It's been one year since they last saw each other, so they want to catch up:

Greg: Hey, what's new?

Mary: I just got a puppy!

Greg:

Oh, is it a male?

Oh, is it a male or not?

Oh, is it a male or a female?

Oh! Do you like baseball?

PQ
NAQ
CAQ
Control

b. **Ask-again**

Context: Mark checks in at a hotel. After the receptionist hands him the 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

5.1.2 Procedure and Statistical analysis

Each subject saw 20 experimental items, 10 for the ask-first-time context and 10 for the ask-again context, plus 24 fillers. The conditions were crossed in a Latin Square Design. 48 participants were recruited on Mechanical Turk and paid \$1.50 for participation. 2 participants were excluded as they failed to complete the task. At the end of each trial, participants were asked to answer the following question with a value between 1 and 7: "How natural does the question sound in light of the goal of the speaker? "1" indicated a completely unnatural question; "7" indicated a perfectly natural question. All items were presented in written form on a screen. As in the first experiment, we ran separate mixed-effects models on the ratings of questions asked for the first time and asked again, with Question Type as the fixed effect and random intercepts for Subjects and Items. Again, the models were ran with the *lmer* package. Given the theoretical motivation of the study, we are especially interested in the comparison between NAQ and CAQ for each moment of the dialogue in which the question was asked. In light of this, we opted to establish NAQs as the reference level.

5.2 Results

The results are plotted in Figure 2 below.

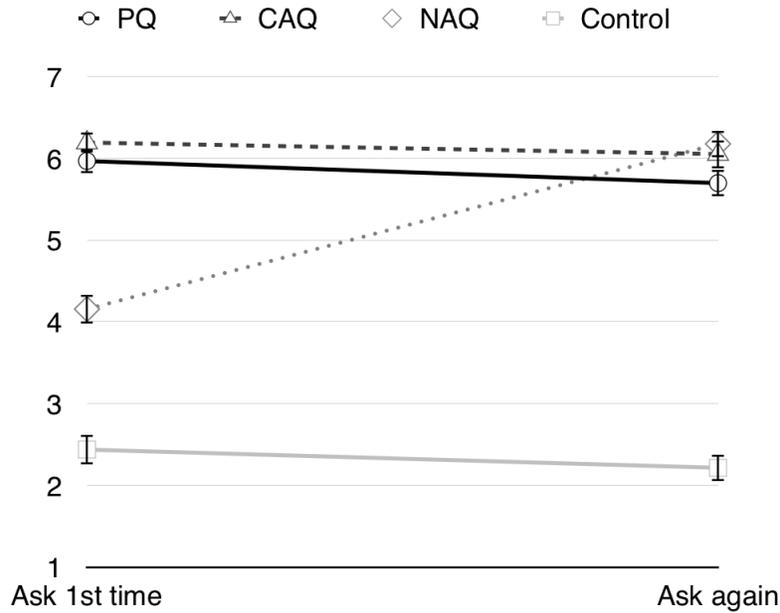


Figure 2: Average naturalness ratings for Experiment 2.

As predicted, the control condition turned out to be highly infelicitous across the board. We therefore removed it from the analysis. Remarkably, CAQs and NAQs patterned differently across these two contexts. When the question was asked for the first time, CAQs were rated higher than NAQs ($\beta=2.01$, $SE= 0.28$, $p <.0001$); when the question was asked again, instead, no difference emerged between NAQs and CAQs ($\beta=-.18$, $SE= 0.14$, $p =.2$). Concerning the contrast between PQs and NAQs, we observe that PQs were significantly better than NAQs when the question was asked for the first time ($\beta=1.78$, $SE= 0.32$, $p <.0001$); by contrast, NAQs were better than PQs when the question was being asked again ($\beta=.48$, $SE= 0.17$, $p <.01$).

5.3 Discussion

In Experiment 2, we compared the distribution PQs, NAQs and CAQs in info-seeking contexts. As a first observation, the results replicate Biezma’s observations with respect to the distinction between PQs and NAQs. PQs are more felicitous than NAQs to ask a question for the first time, whereas NAQs are more felicitous than PQs to ask a question again. Our findings, however, do not support the prediction that CAQs, by virtue of posing logically opposite alternatives, should also be felicitous only in situations in which they are used to ask a question again. Specifically, CAQs show remarkable flexibility across discourse-initial and non-discourse-initial uses. They are as felicitous as a PQ to ask a question for the first time; and they are as felicitous as a NAQ to ask a question again.

6 General discussion

In sum, the findings from both experiments suggest that the distribution of CAQs and NAQs is different, both in non-canonical and info-seeking contexts. We take this result to be evidence that the restricted distribution of NAQs cannot be solely explained on the basis of the factors that had been invoked in previous accounts. Pragmatic analyses, as can be recalled, suggested that the degraded status of NAQs in non-canonical contexts is linked to the fact that they pronounce both alternatives, thus failing to signal the higher utility of p that is normally expected to be in place in these uses. On this view, CAQs should be subject to the same restrictions as NAQs, since they also pronounce both disjuncts. However, this prediction is not borne out. CAQs appear to be felicitous to make invites and ask rhetorical questions, that is, in two of the three non-canonical contexts tested in the experiment. Semantic analyses suggested instead that the restriction of NAQs to contexts in which a question is asked again is related to their denotation, and specifically to the fact that they pose logically opposite alternatives. On this view, CAQs should be subject to the same restrictions as NAQs, since they also present logically opposite alternatives. However, this prediction is not borne out either. While PQs are indeed confirmed to be only felicitous to ask a question again, CAQs turn out to be equally felicitous to ask questions for the first time and to ask questions again. Taken together, these findings support the idea that the illocutionary restrictions on NAQs are crucially linked to the fact that in such questions the second disjunct is spelled out via negation, as opposed to via a full proposition – that is, the property that distinguishes these moves from logically similar alternative questions.

Having empirically established this point, the issue remains open as to why negating p in the second disjunct restricts the illocutionary range of an alternative question, while having the complement of p doesn't. While developing a full proposal would go beyond the scope of the current paper, we want to sketch out a promising line of analysis to shed light on this issue. Specifically, we argue that, *within* the category of alternative questions, different strategies to spell out the second disjunct convey additional information about the relative *prominence* of each alternatives. On the one hand, NAQs are moves that bring about a very specific effect. Besides forcing the addressee to pick between two mutually exclusive alternatives, they also express the second disjunct in terms of the negation of the first one, signaling that the speaker sees p as being more prominent than its alternative – that is, that p is "all that matters" (Biezma & Rawlins 2014; Biezma & Rawlins To Appear). On the other hand, CAQs are relatively flexible moves: they present opposite alternatives, thus pushing the addressee to choose one of them; but by virtue of spelling out both alternatives as a full proposition, they simultaneously convey to the addressee that each option has equal status. We suggest that, from an interactional perspective, NAQs' combination of forcing an answer and marking p as the prominent alternative is crucial to understand why these moves are restricted to ask-again contexts. Specifically, for NAQs to be felicitous, there needs to be a context in which the combination of pressuring the addressee and signaling prominence on p is consistent with the speaker's goal. Among the contexts tested in the experiments and discussed in the literature, this only happens when the speaker aims to re-ask a question about p to wrestle an answer from the listener. Let us consider one of our examples again.

- (10) Context: Mark checks in at a hotel. After the receptionist hands him the 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 or not?

In this context, the interlocutor's failure to respond justifies the speaker's determination to use a strategy that would force an answer; in addition, the fact that a polar question about p has already been asked (twice!) means that the speaker considers the proposition "There is a charge for it" as more prominent than its complement "It is free". If this were not the case, then a polar question like "is it free?" would have been asked instead in the previous conversational turns. In light of these two conditions, a NAQ is thus a consistent strategy to proceed. It forces an answer, consistently with the speaker's determination to obtain a response; and it preserves the prominence asymmetry between p and its alternatives, consistently with the fact that $?p$ has been asked before, and is thus the alternative that the speaker really cares about.

While providing a formal account of this intuitive explanation would go beyond the scope of the paper, we would like to mention two avenues that might be pursued. As a first possibility, we suggest that the notion of *bundling*, introduced by Biezma & Rawlins (2014); Biezma & Rawlins (To Appear), could be especially helpful to achieve this goal. In the authors' account, bundling refers to the particular strategy that a speaker adopts for "packaging" the available alternatives when asking a question. The authors, specifically, argue that every bundling choice made by a speaker is subject to a *Qualitative Constraint*: there must be some reason to group alternatives together as a strategy in a particular way, distinct from prior discourse. When it comes to NAQs, the authors argue that the only possible reason to re-organize the logical space of discourse around p is that that the interlocutors have already accepted a bias towards p in the exchange – that is, that p has already been asked. This derives NAQs' restriction to ask-again contexts. By contrast, bundling with p and its full complement does not presuppose that p is the most prominent alternative, and thus does not place the constraint that the interlocutors have accepted a bias for p . This would explain why CAQs are felicitous in discourse-initial contexts as well.

An alternative analytical route would be to derive the differential distribution of NAQs and CAQs in terms of a competition between marked and unmarked question strategies, following Horn 1984's influential theory of the *division of pragmatic labor*. According to this proposal, if two forms have the same logical content and different markedness status, the marked instantiation tends to have a more restricted distribution, while the unmarked one would instead become the default choice in all the other cases. Consider the contrast between "they stopped the car" and "they caused the car to stop", two truth-conditionally equivalent expressions in which the latter formulation, by means of being formally more complex, is typically regarded

as marked vis-a-vis the former. A possible application of this principle to our case would be the following: NAQs, by making use of negation and additionally conveying a prominence asymmetry between the disjuncts, bring about a more complex effect than CAQs, which spell out both propositions in the same way. This would make NAQs a marked alternative in comparison to CAQs, explaining the distributional specificity of the former and the larger flexibility of the latter. The proposal would however face two challenges. First, it is not obvious that the relevant competition would boil down to a binary opposition between NAQs and CAQs. Polar questions, at least according to some accounts, would also be logically equivalent to these two moves, and would therefore enter this dynamics as well, making it difficult to make clearcut predictions on the basis of a single contrast. Second, there are other plausible ways of determining the markedness of a question, according to some of which CAQs would qualify as the marked variant. For example, CAQs are typically longer, hence more prolix forms than NAQs. While none of these problems necessarily undermines this analysis, more work is needed to frame the findings of our study within the dynamics of this pragmatic competition.

Regardless of what line of analysis one opts to develop, the distribution of CAQs and NAQs in non-canonical contexts remains in need for an explanation. We would like to provide several informal considerations that might provide a useful lead in this direction. As can be recalled, the following pattern emerged from Experiment 1: NAQs and CAQs are both less felicitous than PQs when the speaker intends to draw an inference; but CAQs are more felicitous than NAQs, and as felicitous as PQs, when it comes to make an invite or ask a rhetorical question. As far as the equal status of these moves in inference-drawing contexts is concerned, we suggest that Van Rooij and Šafářová's notion of *Utility Value* provides a viable explanation. Consider the following example.

- (11) Context: Right before the beginning of spring break, George sees camping equipment all around Joe's house and wonders why it is there. Thinking that Joe might be going camping during the break, George thus asks him:
- a. Are you going camping for spring break?
 - b. # Are you going camping for spring break or not?
 - c. # Are you going camping for spring break or are you doing something else?

In such a context, p has clearly a higher utility value, since it is the alternative that would yield a viable explanation of the observed state of affairs. Any other proposition would be irrelevant in this regard, and thus be infelicitous if mentioned in the question, regardless of whether such an alternative surfaces via negation of p or via a different proposition. However, as we saw, Van Rooij and Šafářová's account cannot be straightforwardly extended to the other two types of non-canonical questions tested in the experiment – namely Invites and Rhetorical Questions. In these cases, the expected asymmetry in Utility Value between p and the alternative should predict that CAQs should be as infelicitous than NAQs, if not even more. Similar to what we did for info-seeking contexts, we suggest that the different ways in which CAQs and NAQs mark the prominence of the alternative could shed light on these cases as well.

As far as invites are concerned, we put forward that presenting the alternatives on a par could be a viable strategy for politeness reasons. Specifically, it would serve the purpose of making rejection of the invite look like a legitimate (or at least acceptable) option, thus making sure that the addressee does not feel compelled to accept the invitation. Using a NAQ, instead, would preserve the prominence on *p*. However, contrary to a Polar Question, would force the listener to provide an answer: the end result is a speech act that forces the listener to respond, and that highlights one response as more prominent than the alternative, pressuring the interlocutor to select it. We see this as a move that maximally restricts the discourse options of the addressee, and thus turns out to be at odds with the normally positive disposition towards the hearer that comes with an invitation.

- (12) Context: It's very cold outside. Tom has an extra scarf in his backpack and wants to offer it to his friend Mark, who isn't wearing one. Tom thus turns to Mark and asks:
- a. # Hey, do you want a scarf or not?
 - b. Hey, do you want a scarf or are you ok?

Concerning Rhetorical Questions, we suggest that the differential status of NAQs and CAQs can also be explained in terms of the implications of marking the prominence of one alternative versus presenting them on equal footing. Consider the following example again.

- (13) 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 good player and intends to deny the request. He thus asks:
- a. Are you a child?
 - b. # Are you a child or not?
 - c. Are you a child or an adult?

Recall that, the alternative highlighted in the NAQ – that an adult player is indeed a child – is very unlikely to be true, if not blatantly false. For this reason, as van Rooij & Šafářová 2003 suggest, this alternative, from an informational standpoint, has a much higher utility value than the other one; adding it to our belief state would entail a much more substantive informational gain than the other alternative – that is, the proposition that the player is indeed an adult. This is precisely why a PQ mentioning this alternative is an effective strategy to ask a rhetorical question. But there is also a sense in which the other alternative – i.e. being an adult – has actually high utility: it's also the alternative that picks out the state of the world desired by the speaker, as shown by the fact that the very reason to ask a rhetorical question is to push the interlocutor to adopt an adult-like behavior. Similar to invites, CAQs, by presenting the options on equal footing, would be a coherent strategy for a speaker that intends to convey that both of them have high value; NAQs, by contrast, would mark still one of them as more prominent, thus coming across as incoherent with the goal of highlighting the utility of both options.

In sum, the emerging generalization is as follows. For both invites and rhetorical questions, each alternative can be seen as being useful, though for different reasons. This has an important consequence: if a speaker decides to resort to an alternative question and thus mention both alternatives, it must be because they indeed want to convey that each of them has a high value in the context. As such, it is reasonable for them to resort to a CAQ, a questioning strategy that effectively preserves the equal status of such alternatives. By contrast, using a NAQ would be associated with a pragmatically incongruous behavior, resulting from two inconsistent moves: mentioning both options, and thus conveying that they have similar Utility Value; and conveying that one of them is actually more prominent.

7 Conclusion

In sum, the picture emerging from these studies is the following: purely semantic or pragmatic accounts of the meaning of Alternative Questions are not sufficient to derive the illocutionary restrictions on NAQs that have long been discussed in the literature. The flexible behavior of CAQs suggests that we need a more fine-grained theory of how pragmatic (e.g., highlighting/prominence) and semantic notions (e.g., exhaustivity/exclusivity) interact to determine the message conveyed by an alternative question, and, as a consequence, the illocutionary range of such a move. While we leave the elaboration of such a model to future work, we believe that semantic and pragmatic theory would have much to gain from this enterprise.

References

- Beltrama, A., E. Meertens, & M. Romero. 2018. Decomposing cornering effects in alternative questions: An experimental investigation. In *Proceedings of Sinn und Bedeutung 22*.
- Biezma, M. 2009. Alternative vs polar questions: The cornering effect. In *Proceedings of SALT 19*.
- Biezma, M., & K. Rawlins, 2014. Bundling questions and granularity in discourse. Talk at Semantics and Philosophy in Europe 7, ZAS Berlin, June 26–28.
- Biezma, M., & K. Rawlins. To Appear. Or what? *Semantics and Pragmatics*.
- Bolinger, D. 1978. Yes-no questions are not alternative questions. In *Questions*. Dordrecht: Reidel.
- Horn, L. 1984. Towards a new taxonomy of pragmatic inference: Q-based and r-based implicature. In *Meaning, form, and use in context: Linguistic applications*, 11–42. Washington: Georgetown University Press.
- Kuznetsova, A., P. B. Brockhoff, & R. H. B. Christensen, 2016. Tests in linear mixed effects models. Package 'lmerTest'.
- van Rooij, R., & M. Šafařová. 2003. On polar questions. *Semantics and Linguistic Theory* 13.