

Negative Polar Question Types in English

Maribel Romero¹, Anja Arnhold², Bettina Braun¹ & Filippo Domaneschi³

¹University of Konstanz, ²University of Alberta, ³University of Genoa

1. Introduction

Consider the following Polar Question (PQ) forms. (1-a) is a positive question (PosQ), (1-b) a negative question with low negation (LowNQ), (1-c) a negative question with high negation (HiNQ) and (1-d) a positive question with the nuclear accent on the tensed verb and/or with the particle *really* (*really*-PosQ).

- | | | | |
|-----|----|--------------------------|---------------------|
| (1) | a. | Did John drink? | PosQ |
| | b. | Did John not drink? | LowNQ |
| | c. | Didn't John drink? | HiNQ |
| | d. | DID John (really) drink? | <i>really</i> -PosQ |

All these forms intuitively raise the same issue $\{p, \neg p\}$, that is, they ultimately induce a choice between p and $\neg p$. However, the forms cannot be used interchangeably. The choice of form has been argued to depend (possibly among other things) on two kinds of epistemic bias of the speaker (Ladd 1981, Büring & Gunlogson 2000, Sudo 2013, Domaneschi et al. 2017): original bias and contextual evidence bias.

Original bias is defined in (2) (Domaneschi et al. 2017). For example, a HiNQ *n't p?* mandatorily expresses an original speaker bias for proposition p that need not attend the PosQ counterpart (Ladd 1981). This is illustrated in (3): The PosQ version (3-a) can be used to ask an unbiased question at court but (3-b) would declare a bias for p .

- (2) *Original bias (for a proposition p):*
Speaker's belief or expectation –possibly private– that p is true, based on her epistemic state prior to the current situational context and conversational exchange.
- (3) *Scenario: Lawyer asking unbiased questions at court.*
- | | |
|----|---|
| a. | Did you see the culprit hit the victim? |
| b. | #Didn't you see the culprit hit the victim? |

Contextual evidence bias is defined in (4) (Büring & Gunlogson 2000). For example, a PosQ *p?* is compatible with contextual evidence for proposition *p* whereas a LowNQ *not p?* is incompatible with it, as shown in (5):

- (4) *Contextual evidence bias (for a proposition p)*:
Expectation that *p* is true (possibly contradicting prior belief of the speaker) induced by evidence that has just become mutually available to the participants in the current discourse situation.
- (5) *Scenario*: A enters S' windowless computer room wearing a dripping wet raincoat (contextual evidence for *p = it is raining*). S says:
- a. What's the weather like out there? Is it raining?
 - b. #What's the weather like out there? Is it not raining?

One more interesting aspect of HiNQs will be relevant. While expressing original bias for *p*, HiNQs of form *n't p?* have been argued to allow for two intuitive interpretations (Ladd 1981, Romero & Han 2004, Sudo 2013, a.o.): (i) an **outer** negation reading whereby the speaker double-checks *p*, and (ii) an **inner** negation reading by which the speaker double-checks $\neg p$. The reading is disambiguated by the presence of positive polarity items (PPIs) vs. negative polarity items (NPIs): PPIs like *some*, *already* and *too* enforce the outer negation interpretation, as in (6-a), while NPIS like *any*, *yet* and *either* secure the inner negation reading, as in (6-b):

- (6) a. Didn't John drink some_{PPI} beer? ⇒ Outer reading only
b. Didn't John drink any_{NPI} beer? ⇒ Inner reading only

In this paper, we mostly concentrate on the negative PQ forms, in (7). There is blatant disagreement in the current literature as to which negative PQ forms share the same final meaning –i.e., semantic reading and pragmatic bias profile– and which forms differ in their meaning contribution. Four different splits have been proposed. According to **split (i)**, LowNQs, inner HiNQs and outer HiNQs all form a single grammatical type (van Rooy & Safárová 2003). **Split (ii)** clusters LowNQs and inner HiNQs together as sharing the same bias profile, distinct from that of outer HiNQs (Krifka 2017). A third classification, **split (iii)**, distinguishes between LowNQs and HiNQs in term of bias profiles but clusters inner and outer HiNQs together as having the same semantic reading and bias profile (Ander-Bois 2011, Northrup 2014). Finally, **split (iv)** maintains that LowNQs, inner HiNQs and outer HiNQs need all to be distinguished from one another, since they differ in terms of semantic reading and/or bias profile (Romero & Han 2004, Repp 2013). These four splits are depicted in Table (8):

- (7) a. Did John not drink? LowNQ
b. Didn't John drink a / any beer? inner HiNQ
c. Didn't John drink a / some beer? outer HiNQ

Negative Polar Question Types in English

(8) *Four different splits in the literature*

Split (i)	LowNQ	InnerHiNQ	OuterHiNQ	[vR&S]
Split (ii)	LowNQ	InnerHiNQ	OuterHiNQ	[Kr]
Split (iii)	LowNQ	InnerHiNQ	OuterHiNQ	[A,N]
Split (iv)	LowNQ	InnerHiNQ	OuterHiNQ	[R&H,R]

The goal of the present paper is to assess the empirical plausibility of each of these four splits using experimental evidence. To this end, three experimental studies will be presented. **Study 1** tests the bias conditions of LowNQs and HiNQs (section 2). As we will see, the predictions of splits (i) and (ii) are falsified in several bias configurations, leaving only splits (iii) and (iv) in the running. After offering additional background on inner vs. outer HiNQs (section 3), we test splits (iii) and (iv) in the remaining studies. **Study 2** is concerned with the prosody of inner and outer HiNQs. It is found that the disambiguation of the two readings correlates with prosody; however, the attested prosodic difference can be explained both under splits (iii) and (iv). Finally, in **Study 3**, we reverse the polarity of the original speaker bias and compare *really*-PosQ vs. stacked negation HiNQs. As it will be shown, the selection of *really*-PosQ vs. stacked negation HiNQs directly follows from split (iv) but is unexpected under split (iii).

2. Study 1

In Study 1, we cross the two kinds of epistemic bias to investigate what negative question forms are produced in certain pragmatic bias configurations, focusing on the predictions of the analyses in split (i) and split (ii). More concretely, on the *x*-axis we list two values for original bias: original bias for *p* and no original bias, which we write as “n(neutral)”; on the *y*-axis we list two values for contextual evidence bias: n(neutral) and contextual evidence bias for $\neg p$. The result is the four pragmatic conditions in Table (9). Of these, the cell n/n is our control cell, for which all the splits involved predict that PosQ should be selected; the rest are our test cells.

(9) *Pragmatic conditions in Study 1*

	Original bias		
Contextual		p	n
Evidence	n		\checkmark PosQ
bias	$\neg p$		

We start with the predictions of **split (i)**, due to van Rooy & Safárová (2003). All three negative PQ forms –LowNQs, inner HiNQs and outer HiNQs– are treated uniformly as requiring that the utility value of the pronounced proposition $\neg p$ be higher than that for *p*. Concentrating on informativity-based utility (i.e., not on goal-based utility), the utility value of $\neg p$ is higher than that of *p* whenever the speaker’s expectation is towards *p*. This is because learning an unexpected $\neg p$ would be highly informative whereas learning an expected *p* would be less informative. This gives us the predictions for two of the test cells. In the p/n cell (original bias for *p*, no contextual evidence bias), the existing bias for *p* makes

the pronounced proposition $\neg p$ highly informative and thus LowNQs and (inner and outer) HiNQs should be all equally appropriate. Conversely, in the $n/\neg p$ cell, the existing bias for $\neg p$ makes the pronounced $\neg p$ less informative and thus LowNQs and (inner and outer) HiNQs should be all equally dispreferred. For the final cell, the $p/\neg p$ cell where conflicting biases are at stake, a further nuance is introduced. While informativity is typically checked with respect to the current epistemic state of the speaker, it may sometimes be checked with respect to a previous one. More concretely, in the case of LowNQs, informativity may be computed with respect to an epistemic state (favoring p) prior to encountering the contextual evidence (for $\neg p$), which licenses LowQs in this cell; in the case of HiNQs, informativity is always checked with respect the current epistemic state, and thus they are licensed in this cell as long as the contextual evidence has not substantially affected the speaker's original confidence towards p .¹ These predictions are summarized in Table (10). Crucially, LowNQs and HiNQs are predicted to pattern together across the board:²

(10) *Predictions by van Rooy & Safárová (2003)*

		Original bias	
		p	n
Contextual Evidence bias	n	✓LowNQ ✓HiNQ	✓PosQ
	$\neg p$	✓LowNQ ✓HiNQ	#LowNQ #HiNQ

We turn to the predictions of **split (ii)**, defended by Krifka (2017). While unbiased PQs like the PosQ in (1-a) have the structure in (11-a), LowNQs and inner HiNQs (double-checking $\neg p$) like (7-a)-(7-b) are treated uniformly as having the underlying representation (11-b). The idea is the following. A question updates the current commitment state c –roughly, the set of commitments accumulated up to the current point in discourse– by imposing on the addressee a restriction on possible future continuations. In the case of structure (11-a), the question operator QU builds the set $\{p, \neg p\}$ and the updated commitment state c allows the addressee to continue by asserting p or by asserting $\neg p$, hence permitting an unbiased choice between the two. In the case of (11-b), we have the speech act operators ASSERT and REQUEST, which combine with the sentence radical to update c by allowing only one possible continuation by the addressee, namely, the addressee is requested to assert $\neg p$. This single-choice allowance is argued to be felicitous, among other cases, when contextual evidence bias for $\neg p$ has been encountered. This predicts both LowNQs and inner HiNQs to be uniformly acceptable in the test cell $n/\neg p$, as indicated in Table (12):

¹The distinction between LowNQs and HiNQs in van Rooy & Safárová (2003) is based on their potential focus realizations: Negation can be focused in LowNQs but not in HiNQs. It is not explained, however, why focus vs. no focus on negation makes the previous vs. current epistemic state relevant.

²Recall that van Rooy & Safárová (2003) also allow for utility to be based on the speaker's goals rather than her information state. However, factoring in goal-based utility into any of the cells would not change the homogenous pattern, since it would affect LowNQs and HiNQs equally.

Negative Polar Question Types in English


- (11) a. [QU [p]]
 b. [REQUEST [ASSERT \neg p]]
- (12) *Predictions by Krifka (2017)*

		Original bias	
		p	n
Contextual Evidence bias	n		✓PosQ
	\neg p		✓LowNQ ✓HiNQ

To test these predictions, we ran an experimental study with 42 participants (English students at University College London: $M = 25.0$; $SD = 2.9$; 17 males, 25 fem), with the following design (Domaneschi et al. 2017). We created 42 written scenarios presenting ordinary fictional conversations. Each scenario contained two caption/picture pairs. The first caption/picture manipulated the original bias of the speaker (for p, neutral) and the second the contextual evidence bias (neutral, for \neg p), as illustrated in (13).³ Each scenario was followed by set of four questions –a PosQ, a *really*-PosQ, a LowNQ and a HiNQ– plus an “other” option, as illustrated in Table (14). The subjects were instructed to select the question that sounded most natural by producing it aloud or to select the “other” option if none of the questions was considered appropriate.


- (13) *Sample materials (caption 1 and caption 2) in Study 1*

CAPTION 1: If it doesn't rain tomorrow, you will surely go to the beach. The forecast for the next morning indicates:



90% raining

CAPTION 2: The day after your flat-mate Sam comes from the outside and enters your bedroom:



...with sunglasses and diver's mask

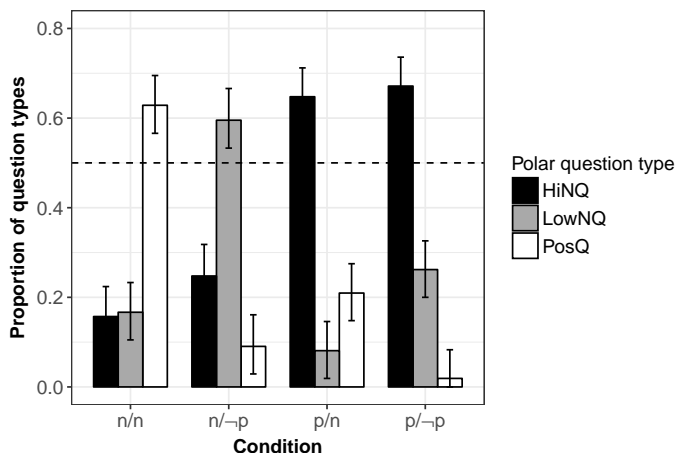
- (14) *Sample question forms for selection task in Study 1*

QUESTION TYPE	QUESTION
<i>PosQ</i>	Is it raining?
<i>Really-PosQ</i>	Really? Is it raining?
<i>LowNQ</i>	Is it not raining?
<i>HiNQ</i>	Isn't it raining?
<i>Other option</i>	Other ways of asking if it is raining

³The complete experiment included three values for original bias (for p, neutral, for \neg p) and three for evidence bias (for p, neutral, for \neg p). See Domaneschi et al. (2017) for details. Here we concentrate on the four bias crossings that are relevant for splits (i) and (ii).

The results of this study for the forms that concern us are depicted in (15). The preferred choice (PosQ in n/n , LowNQ in $n/\neg p$, HiNQ in p/n and $p/\neg p$) was significantly above 50% (p-values of separate t-tests for by-subjects and by-items aggregates all < 0.05).

- (15) *Average proportion of question types chosen in each condition (in all Figures, whiskers show 95% confidence interval)*



What do these results tell us about split (i) and split (ii)? In the control cell n/n , PosQ is the optimal choice, as predicted by the four splits in the literature. However, in the test cells $n/\neg p$, p/n and $p/\neg p$, the predictions of split (i) and split (ii) are not borne out.

Split (i) assigns LowNQs, inner HiNQs and outer HiNQs the exact same analysis for cells $n/\neg p$ and p/n , predicting them to be equally dispreferred in condition $n/\neg p$ and equally preferred in condition p/n . But the results show that LowNQs are clearly preferred in the former condition and HiNQs in the latter.⁴ Split (i) also predicts LowNQs and HiNQs to be acceptable in the third condition $p/\neg p$, although each in reference to a different \neg -previous vs. current-epistemic state. The attested preference for HiNQs in this condition might be due to the current epistemic state being accidentally more salient in the stimuli. Since we did not control for this potential factor, this cell yields inconclusive results.

According to split (ii), LowNQs and inner HiNQs have the exact same underlying representation (11-b), which makes them equally suitable for condition $n/\neg p$. However, the results show that LowNQs are clearly preferred over HiNQs in this condition.

In sum, neither the general parallelism between LowQs and HiNQs predicted by split (i) for the cells $n/\neg p$ and p/n nor the restricted parallelism predicted by split (ii) for cell $n/\neg p$ are borne out. Instead, LowNQs and HiNQs are each preferred in different conditions. This leaves only splits (iii) and (iv) in the running, which we investigate in the remaining two studies.

⁴One could try to explain part of the skewing via some pragmatic competition, e.g., saying that the syntactic form of LowNQs is less marked than that of HiNQs and that this makes the former preferred over the latter despite both having the same semantic and pragmatic characterization. That would take care of the skewing towards LowNQs in cell $n/\neg p$, but then we would need the opposite story for cell p/n . In the end, if any such pragmatic competition takes place, it will have to resource to some difference in the semantic/pragmatic analysis of LowNQs vs. HiNQs so that different preferences in different bias conditions can be derived.

3. Background for studies 2 & 3

As mentioned in the introduction, HiNQs like (16) have been noted to intuitively allow for an outer negation interpretation and an inner negation interpretation (Ladd 1981). In the outer negation interpretation, the S(peaker) originally believes p and –with or without new evidence to the contrary– wants to confirm p by means of double-checking p . This is illustrated with Ladd’s original example (17). In the inner negation interpretation, S originally believes p , has now tentatively inferred the opposite and wants to double-check $\neg p$, the new inference. Ladd’s original example is provided in (18):

(16) Isn’t there a vegetarian restaurant around here?

(17) A: You guys must be starving. You want to get something to eat?
S: Yeah, isn’t there a vegetarian restaurant around here?

(18) S: I’d like to take you guys out to dinner while I’m here –we’d have time to go somewhere around here before the evening session tonight, don’t you think?
A: I guess, but there’s not really any place to go in Hyde Park.
S: Oh, really, isn’t there a vegetarian restaurant around here?

Splits (iii) and (iv) differ in how this double interpretation is treated and derived.

Split (iv), defended a.o. by Romero & Han (2004) and Repp (2013), treats this double interpretation as genuine ambiguity. According to Romero & Han (2004), the preposing of negation introduces the conversational epistemic operator VERUM defined in (19). Roughly, [VERUM p] states that x (= the Speaker and/or the Addressee) is sure that according to x ’s conversational goals p should be added to the C(ommon) G(round). Negation may scope over VERUM, as in structure (20-a), or under VERUM, as in (20-b). By using structure (20-a), S asks the A(ddressee) for any information that may cast doubt on p . This means that the prejacent –the proposition the question is about– is p and suggests that S is still attached to her original belief p and wants to hear reasons to doubt p before revising her epistemic state. This corresponds to Ladd’s intuitive outer negation reading. By using structure (20-b), S asks A for conclusive evidence for $\neg p$. Hence, the prejacent is $\neg p$ and the question suggests that S is considering switching to the new inference $\neg p$ if full justification is provided. This matches Ladd’s inner negation reading. Finally, the two readings are disambiguated by PPIs/NPIs and possibly by other means.

(19) $[[\text{VERUM}]] = \lambda p_{\langle s,t \rangle} . \lambda w_s . \forall w' \in \text{Epi}_x(w) [\forall w'' \in \text{Conv}_x(w') [p \in \text{CG}_{w''}]]$

(20) a. [Q [\neg VERUM [p]]] OUTER READING
b. [Q [VERUM [$\neg p$]]] INNER READING

Split (iii), defended by AnderBois (2011) and Northrup (2014), contends that HiNQs have only one genuine reading –the outer negation interpretation– and that what feels like an inner negation interpretation is just a subcase of that unique reading in a certain prag-

matic configuration. AnderBois (2011) uses three main ingredients. First, besides classical negation, a special lexical entry for high negation is given, abbreviated \neg_{hi} , which guarantees that the speaker has an original bias towards p . Second, HiNQs unambiguously have the structure (21), that is, HiNQs invariably have p as the prejacent. Third, by default, S has a tendency towards retaining her original belief p . When this default tendency is maintained, we have a question double-checking p in which S is still leaning towards her original belief p . This corresponds to Ladd's intuitive outer negation reading. Deviating from this default tendency requires an NPI –whose pragmatic properties overwrite the default– or other overt marking. In this case, though the question is still double-checking the prejacent p , S might be considering switching to the new inference $\neg p$. This last combination of factors, it is argued, is what is perceived as Ladd's intuitive inner negation reading.

(21) [Q \neg_{hi} p]

4. Study 2

The question arises, what other overt marking helps distinguish the outer and inner interpretations. More concretely, does some prosodic cue correlate with the outer/inner distinction? And, if so, does the meaning contribution of this cue –to the extent that this meaning can be pinned down– allow us to favour one analysis over the other?

To investigate this question, we designed a study focusing on the cell $p/\neg p$ and crossing two variables: (i) “ownership” of the checked proposition (S 's original belief p vs. the proposition $\neg p$ implied by A) and (ii) certainty about checked proposition (high vs. low; see Vanrell et al. 2013, on prosodic effects of certainty in Catalan). We constructed dialogs setting up S 's belief p , A 's belief $\neg p$ and describing S 's wish to check p or $\neg p$ and her degree of certainty about the checked proposition (90% as high, 60% as low). This gives us four conditions: S checks p and overwhelmingly leans towards her original belief p (p +high certainty); S checks p and mildly leans towards p (p +low certainty); S checks $\neg p$ and mildly leans towards A 's implication $\neg p$ ($\neg p$ +low certainty); and S checks $\neg p$ and overwhelmingly leans towards $\neg p$ ($\neg p$ +high certainty). The subjects had to choose a HiNQ or a LowNQ to resolve the issue in these conditions and utter it aloud.

We recruited 24 participants (students at the University of Alberta, native speakers of Canadian English: $M=22.5$; $SD=3.9$; 8 males, 15 fem, 1 other). We had 16 experimental items, each appearing twice per list in different conditions (2 blocks), and 15 fillers, each appearing twice per list in different conditions, which amounted to 62 dialogues per list.

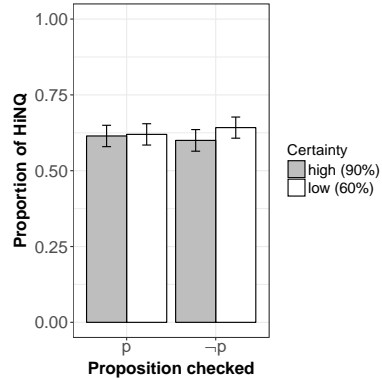
A caveat should be made at this point. Study 2 was conducted in Alberta, where uptalk is frequent among young speakers. Thus, the shape of prosodic cues may differ from those in other varieties of English.

We obtained two main results. First, in a setting where the S originally believes p and A implies $\neg p$, a HiNQ is preferred over a LowNQ no matter whether the S 's original belief p is checked or the new inference $\neg p$ is checked and no matter whether the certainty about the checked proposition is high or low. The proportion of selection of HiNQs is shown in (22); the remaining proportion corresponds to the selection of LowNQs. Results of a logistic mixed-effects regression model (Baayen et al. 2008) with proposition type and certainty as

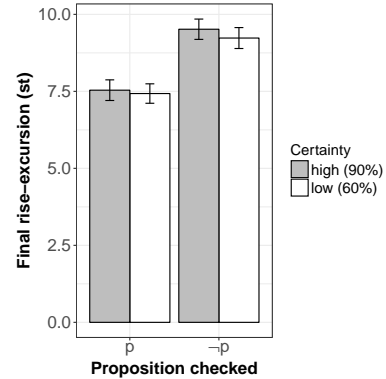
Negative Polar Question Types in English

fixed factors and participants and items as crossed random effects showed no main effects and no interaction (all p -values > 0.5).

(22) *Average proportion of HiNQ.*

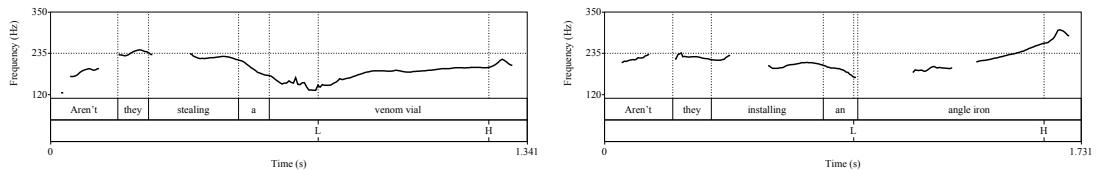


(23) *Average f_0 -excursion of final rise.*



Second, in analysing the HiNQ recordings, a correlation was found between the pitch excursion of the final rise and the checked proposition. The average rise excursion in the conditions checking p was 7.5 semitones (st) whereas the average excursion in the conditions checking $\neg p$ was approximately 9.5st. Results of a linear mixed effects regression model showed a main effect of proposition on the f_0 -excursion of the final rise ($p = 0.001$); there was no effect of certainty and no interaction (both p -values > 0.3). The results are plotted in (23). Two sample pitch tracks are given in (24): The one on the left illustrates a low final rise and the one on the right a high final rise.

(24) *Low rise when checking p (left) and high rise when checking $\neg p$ (right)*



Let us discuss these two results. The first result shows that, in the bias configuration [Original bias = p , Evidence bias = $\neg p$], both the outer negation interpretation and the inner negation interpretation are preferably expressed with a HiNQ rather than with a LowNQ. This confirms the results of Study 1 with respect to the $p/\neg p$ cell and, furthermore, it refines them. Since we did not control for Ladd's interpretations in Study 1, we did not know which one of Ladd's readings the subjects were expressing and, thus, we couldn't tell whether HiNQs were preferred over LowNQs for the outer negation reading, for the inner negation interpretation or for both. The present results show that, in the bias configuration $p/\neg p$, both interpretations are preferably expressed with a HiNQ rather than with a LowNQ.

The second result shows that Ladd's two intuitive interpretations are characterised by different final pitch contours: The outer negation interpretation tends to be pronounced with a low final rise \uparrow whereas the inner negation interpretation typically presents a high final rise \uparrow . Recall that these results come from an uptalk variety of English, so the low rise may correspond to a final falling contour and the high rise to a normal rising contour in other

English dialects. With this in mind, can we find a parallel prosodic contrast in some other English constructions? Can the meaning contribution proposed for the contrast in those other constructions be extended to our data and help us decide between splits (iii) and (iv)? We will see that parallel prosodic contrasts are indeed found, but that, at this point, both split (iii) and split (iv) are amenable to explaining our experimental results.

A first possibly related cue is the final boundary tone in falling and rising declaratives like (25). This final contour has been analysed by Gunlogson (2001) as indicating the ‘locus’ of the commitment or authority. In a nutshell, she proposes that declarative syntax signals commitment to the prejacent (= ‘it is raining’); a final fall ↓ indicates that the locus of the commitment/authority is the Speaker whereas a final rise ↑ indicates that the locus is the Addressee.

- (25) a. It is raining ↓.
 b. It is raining ↑?

This idea can be adapted to explain our experimental results on HiNQs under Romero & Han’s split (iv) analysis as follows. With a HiNQ [*n’t p?*], S brings up the issue {*p*, ¬*p*} while at the same time expressing an original bias for *p*. On the one hand, if S pronounces the HiNQ with a final low rise ↑, S indicates that she considers the locus of the authority on this issue to be herself. That is, even though the issue {*p*, ¬*p*} is open, S expresses a bias for *p* and signals that she puts herself in the authoritative position. Since, when maintaining one’s own belief, the outer negation structure [Q [¬ VERUM [*p*]]] is most appropriate, the final low rise and the outer negation structure are natural companions. On the other hand, if S pronounces a HiNQ with a final high rise ↑, S signals that she considers the locus of the authority to be the Addressee. Since, when S is considering switching to the A’s proposition ¬*p*, the inner negation structure [Q [VERUM [¬*p*]]] is most appropriate, the final high rise and the inner negation structure go together.

A second potentially related cue is the final contour in reversed polarity tags, which may be falling or rising, as in (26) (see, e.g., Ladd 1981). Farkas & Roelofsen (2017), building on Northrup (2014), analyse this prosodic cue as indicating the degree of S’s conditional commitment to the prejacent proposition *p* (= ‘that Amalia left’): A final fall ↓ signals that S’s conditional commitment is moderate to high and a final rise ↑ indicates that it is low.

- (26) a. Amalia left, didn’t she ↓?
 b. Amalia left, didn’t she ↑?

This idea can be applied to our experimental results on HiNQs within AnderBois’ split (iii) analysis as follows. A HiNQ [*n’t p?*] always has *p* as the prejacent: [Q ¬_{hi} *p*]. A low final rise ↑ indicates that S’s conditional commitment to prejacent *p* is high(ish). This matches the default tendency to retain one’s own belief and gives us the outer negation reading. A high final rise ↑, instead, overtly marks that S is deviating from this default tendency and that S’s conditional commitment to prejacent *p* is now low. This change in S’s attitude, combined with the invariable structure [Q ¬_{hi} *p*], leads to a perception of the sentence as yielding Ladd’s intuitive inner negation interpretation.

In sum, a prosodic cue has been found to correlate with the two intuitive interpretations of HiNQs: The outer reading is realised with a low final rise \uparrow , whereas the inner interpretation –be it a separate reading or a subcase of the former reading– is pronounced with a high final rise \uparrow . At this point, different meanings may be assigned to this cue, making the finding in principle compatible both with split (iii) and with split (iv).

5. Study 3

Recall the characterization of Ladd’s intuitive interpretations of HiNQs. In the outer negation reading, S originally believes p and –with or without new evidence to the contrary– wants to confirm her own belief p by means of checking p. In the inner negation interpretation, S originally believes p, has now tentatively inferred that the opposite might be true and wants to check $\neg p$, the new inference. The prosodic results of Study 2 seem to correlate with S’s *attitude* towards the expected resolution of the issue in the two interpretations: maintenance of S’s original belief vs. concession towards A’s proposition. Study 3 strives to tap directly into the shape of the *prejacent* in the two interpretations.

As we saw in section 3, in the bias configuration [Original bias = p, Evidence bias = $\neg p$], the split (iv) analysis by Romero & Han (2004) produces the two underlying structures in (20), both of which map to a surface realisation as a HiNQ. In the same bias configuration, the split (iii) analysis by AnderBois (2011) produces the unique structure (21), which is again realised as a HiNQ. The idea in Study 3 is to start with the opposite bias configuration, namely, [Original bias = p, Evidence bias = $\neg p$] and examine the underlying structures and surface realisations predicted by the two analyses.

We start with split (iv) à la Romero & Han (2004). Starting with S’s original bias for $\neg p$ and A’s implication against it, S may choose to use an outer negation structure double-checking her original belief $\neg p$, as in (27-a), or use (what in principle looks like) an inner negation structure double-checking A’s implication $\neg\neg p$, which simplifies to a structure double-checking p, as in (27-b). When the context primes Ladd’s outer reading, the former structure will be selected; since it contains one negation above and one below VERUM, it will be pronounced as a HiNQ with stacked negation, e.g., *Didn’t John not drink?*. When the context primes Ladd’s inner reading, the latter structure will be chosen; since it contains VERUM and no negation, it will crucially be realised as a *really*-PosQ with focus on the tense verb and/or the particle *really*, e.g., *DID John drink?*. This means that we expect an asymmetric realisation pattern between the conditions checking $\neg p$ vs. p.

(27) Romero & Han (2004)’s predictions with [Original bias = $\neg p$, Evidence bias = p]:

- a. [Q \neg VERUM $\neg p$] \Rightarrow stacked negation HiNQ, e.g., *Didn’t John not drink?*
- b. [Q VERUM \nearrow \nearrow p] \Rightarrow *really*-PosQ, e.g., *DID John drink?*

The predictions of split (iii) à la AnderBois (2011) are considered next. In his analysis, the same structure (28) with S’s original proposition $\neg p$ in the *prejacent* can be used in contexts leading to Ladd’s outer reading (default) and in contexts priming Ladd’s inner

interpretation (with some special surface marking, e.g., a final high rise). This means that, in principle, we expect a parallel selection pattern in the conditions checking $\neg p$ vs. p .

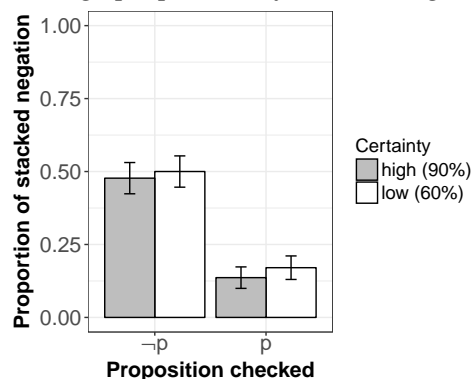
- (28) *AnderBois (2011)*'s predictions with [Original bias = $\neg p$, Evidence bias = p]:
 [Q $\neg_{hi} \neg p$] \Rightarrow stacked negation HiNQ, e.g., *Didn't John not drink?*

To test these predictions, we designed Study 3 as follows. We concentrate on the bias condition [Original bias = $\neg p$, Evidence bias = p], opposite from Study 2. Then, we cross the two same variables as in Study 2: ownership of checked proposition (Speaker's vs. Addressee's) and degree of certainty about checked proposition (high vs. low). Dialogs were built setting up S's belief that $\neg p$ and A's implication that p , describing S's wish to check p or $\neg p$ and stating her degree of certainty about the checked proposition (90% as high, 60% as low). The subjects were requested to choose a *really-PosQ* (e.g., *DID John drink?*) or stacked negation HiNQ (e.g., *Didn't John not drink?*) to resolve the conflict.

We recruited 30 participants (students at the University of Alberta, native speakers of Canadian English: $M = 21.5$; $SD = 2.76$; 7 male, 23 fem). We tested 16 experimental items, each appearing once per list in one of the conditions, and had 16 fillers with choice between other question types (HiNQ, LowNQ, declarative question, *really-PosQ*), resulting in 32 scenarios per list.

Let us turn to the results. The results show that stacked HiNQs were chosen at approximately the same rate as *really-PosQ*s in the checking $\neg p$ conditions but were significantly dispreferred in favour of *really-PosQ*s in the checking p conditions. The proportion of selection of stacked HiNQs is given in (29); the remaining proportion corresponds to *really-PosQ*s. Results of a linear mixed effects regression model showed a main effect of proposition ($p < 0.0001$), no effect of certainty ($p > 0.8$) and no interaction ($p > 0.15$).

- (29) *Average proportion of stacked negation HiNQ*



How do these results impact on splits (iii) and (iv)? In the conditions checking S's proposition $\neg p$ (outer reading), the stacked negation form *Didn't John not drink?* is in principle predicted to be preferred over the *really-PosQ* form *DID John drink?* by both approaches. The fact that the stacked negation form is selected only in approx. 50% of the

cases may be due to its inherent markedness. The two approaches differ, though, in their predictions for the conditions checking A's proposition p (inner reading).

Split (iv) predicts the *really*-PosQ form *DID John drink?* to be preferred over the stacked negation HiNQ form in the conditions checking A's proposition p. Taking the 50%-50% distribution of the two forms in the conditions checking S's proposition $\neg p$ as the baseline, a significant skewing in favour of the *really*-PosQ form is predicted by this analysis in the p conditions. This prediction is borne out.

Split (iii) does not predict a difference between the $\neg p$ conditions (genuine outer reading) and the p conditions (inner-like interpretations). In both cases, to express the combination of biases in the context (original bias for $\neg p$, contextual evidence for p), the structure [Q \neg_{hi} $\neg p$] should be equally available (with the a low rise \uparrow in the $\neg p$ conditions and a high rise \uparrow in the p conditions). Under this analysis, it is unclear at this point why the *really*-PosQ form is a stronger competitor in one condition than in the other.⁵

6. Conclusion

Four different splits of negative polar questions have been proposed in the literature, as depicted in Table (8). In this paper, we have presented three experimental studies investigating the predictions of these four splits.

Study 1 tested the preferred choice of question form in different bias conditions crossing the speaker's original bias and contextual evidence bias. The results showed that LowNQs and HiNQs are selected asymmetrically across different bias conditions. This goes contra the predictions of splits (i) and (ii), which treat LowNQs and (inner) HiNQs on a par and thus expect a parallel behaviour of the two.

Study 2 concentrated on the bias configuration [Original bias = p, Evidence Bias = $\neg p$] and investigated the prosodic realisation of HiNQs under the two intuitive readings discussed by Ladd (1981). The outer negation reading was found to correlate with a low final rising contour and the inner negation interpretation with a high final rising contour. Upon examination of parallel prosodic cues in other constructions, the prosodic findings turned out to be explainable both under split (iii) and (iv).

Study 3 inverted the initial bias configuration of Study 2 to [Original bias = $\neg p$, Evidence Bias = p] and tested stacked negation HiNQs and (*really*-)PosQs in scenarios lead-

⁵To explain our experimental results, defendants of split (iii) might try to argue as follows. The bundle [O(riginal) B(ias): $\neg p$, E(vidence) B(ias): p, checking $\neg p$, sticking to $\neg p$] matches the default and genuine (outer) reading of stacked HiNQ; the bundle [OB: $\neg p$, EB: p, checking $\neg p$, leaning towards p] matches the non-default inner-like interpretation of stacked HiNQ; and the bundle [OB: $\neg p$, EB: p, checking p, leaning towards p] cannot be expressed by a HiNQ to begin with, but by a *really*-PosQ. In other words, one could try to argue that, by specifying in the p conditions that the speaker wants to check p, the underlying representation [Q \neg_{hi} $\neg p$] cannot be used to begin with, since the proposition in the sentence radical –the proposition being checked– is $\neg p$. This would explain the skewing towards *really*-PosQ in the p conditions. However, if that were the case, a parallel result would have been found in Study 2. More concretely, then the bundle [OB p, EB $\neg p$, checking $\neg p$, leaning towards $\neg p$] should also not be expressible by a HiNQ, but by some other competing form, e.g., by a LowNQ. But that is not what the experimental findings of Study 2 show (see (22)): HiNQs were preferred over LowNQs both in the p conditions and –despite their (alleged) structure [Q \neg_{hi} p]– in the $\neg p$ conditions, and approximately at the same rate (i.e., no skewing in the $\neg p$ conditions).

ing to Ladd's outer and inner readings. The results show that the two forms are selected at different rates in outer reading scenarios vs. inner reading scenarios. This asymmetry is predicted by split (iv) but, at this point, remains unexpected under split (iii). This final result tentatively favours split (iv) over split (iii).

References

- AnderBois, Scott. 2011. Issues and alternatives. Doctoral dissertation, U. C. Santa Cruz.
- Baayen, R. Harald, Douglas J. Davidson, & Douglas M. Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59:390–412.
- Büring, Daniel, & Christine Gunlogson. 2000. Aren't positive and negative polar questions the same? Ms. UCLA & UCSC.
- Domaneschi, Filippo, Maribel Romero, & Bettina Braun. 2017. Bias in polar questions. evidence from English and German production experiments. *Glossa* 2:26. 1–28.
- Farkas, Donka, & Floris Roelofsen. 2017. Division of labor in the interpretation of declaratives and interrogatives. *Journal of Semantics* 1–53.
- Gunlogson, Christine. 2001. *True to form: Rising and falling declaratives as questions in english*. New York: Routledge.
- Krifka, Manfred. 2017. Negated polarity questions as denegations of assertions. In *Contrastiveness in information structure, alternatives and scalar implicatures*, ed. Chungmin Lee, Ferenc Kiefer, & Manfred Krifka, 359–398. Cham: Springer.
- Ladd, Robert. 1981. A first look at the semantics and pragmatics of negative questions and tag questions. In *Proceedings of Chicago Linguistic Society 17*, 164–171.
- Northrup, Oliver. 2014. Grounds for commitment. Doctoral dissertation, U.C. Santa Cruz.
- Repp, Sophie. 2013. Common ground management: Modal particles, illocutionary negation and verum. In *Beyond expressives: Explorations in use-conditional meaning*, ed. D. Gutzmann & H.-M. Gaertner, 231–274. Leiden: Brill.
- Romero, Maribel, & Chung-hye Han. 2004. On negative “yes/no” questions. *Linguistics and Philosophy* 27:609–658.
- Sudo, Yasutada. 2013. Biased polar questions in English and Japanese. In *Beyond expressives: Explorations in use-conditional meaning*, ed. D. Gutzmann & H.-M. Gaertner, 275–295. Leiden: Brill.
- van Rooy, Robert, & Marie Safárová. 2003. On polar questions. In *Proceedings of Semantics and Linguistic Theory 13*, ed. R. B. Young & Y. Zhou, 292–309. Ithaca: CLC.
- Vanrell, Maria M., Ignasi Mascaró, Francesc Torres-Tamarit, & Pilar Prieto. 2013. Intonation as an encoder of speaker's certainty: information and confirmation yes-no questions in Catalan. *Language and Speech* 56:163–190.

Maribel Romero, Anja Arnhold, Bettina Braun, Filippo Domaneschi
maribel.romero@uni-konstanz.de, arnhold@ualberta.ca, bettina.braun@uni-konstanz.de,
filippo.domaneschi@unige.it