

## Chapter 1

# Psycholinguistic processing tasks and the study of question bias

E Jamieson

University of Southampton

Vinicius Macuch Silva

University of Birmingham

The study of biased questions is situated in a rich semantic tradition, based primarily on introspective cross-linguistic work. In this chapter we explore how psycholinguistic processing tasks can complement this existing work and help hone our understanding of the factors that influence question production, and the meanings these questions can carry, in this complex semantic-pragmatic domain.

## 1 Introduction

In recent years, a growing body of literature in formal semantics has sought to understand “non-canonical questions” (Dayal 2017), and particularly “biased questions” — interrogatives in which the questioner has some sort of expectation as to what the answer to the question might be.<sup>1</sup> The so-called “bias” may originate in an expectation based on, for example, epistemic or deontic information, such as when the questioner is knowledgeable about the state of affairs they are

---

<sup>0</sup>Joint first authors

<sup>1</sup>Questions with biases have also been of interest to researchers in the interactional linguistic tradition. This research focuses on how questions can reflect stances that users may choose to take in communicative interaction, in particular regarding the likelihood of the states of affairs that the questions refer to (e.g. Heritage & Raymond 2021; Raymond & Heritage 2021), and how responses to these questions may be organised (e.g. Heritage 2012; Lee 2015). Here we focus on what has been established in formal semantics, but acknowledge the links between the work conducted in both fields.

inquiring about or when they use the question as a form of directive. In this paper, we focus exclusively on epistemic biases. The main body of this work on non-canonical questions has relied on introspective judgments to delineate the pragmatic contexts in which these questions can be used, and to postulate how, semantically, the bias is incorporated into the question.

More recently, experimental work has begun in this domain, primarily focused on production tasks (Domaneschi et al. 2017). This work has begun to hone the wide array of epistemic and evidential biases that have been claimed to be relevant for biased question production, but a number of questions, including about how those interrogative forms are interpreted, remain.

In this paper, we illustrate how psycholinguistic processing experiments can complement the formal semantic research that has characterised biased question research to date. In section 2, we give a brief overview of the trajectory of the methodologies used to study question bias, from introspective observations to more recent experimental work. In section 3, we lay out some questions that can be addressed by approaching the topic of question bias through processing tasks, and what can be gained from these kinds of studies. With particular reference to the studies presented in Macuch & Jamieson (in prep) and Tian et al. (2021), we show how these questions have begun to be explored. Finally, we discuss some of the challenges of using processing methodologies to investigate a complex phenomenon like question bias, which spans the interfaces of semantics and pragmatics, as well as (morpho)syntax and prosody. We discuss how to design tasks to deal with these challenges, and thus to allow psycholinguistic research to contribute to this growing field.

## 2 Studying question bias: from introspection to production

### 2.1 Theoretical developments

Negation is a polyfunctional linguistic device which can be used pragmatically for purposes as varied as mitigating the meaning of an assessment (e.g. Fraenkel & Schul 2008; Giora et al. 2005; Krifka 2007) or modulating the acceptability of a request (e.g. Koike 1994). The rich formal semantic literature on question bias begins with an observation about the role of negation in polar questions, more specifically Ladd's (1981) observation that, in English, a question with a *negative* surface form can in fact indicate that the questioner has a belief of the *positive* proposition. This can be seen in 1-2, adapted from Ladd (1981), in which S uses a

negated form of the question to express their belief of the existence of a vegetarian restaurant.

- (1) *S is visiting A, and believes there is a good vegetarian restaurant in town.*  
A: You guys must be starving. Shall we get something to eat?  
S: Yeah, isn't there a vegetarian restaurant around here?
- (2) *S is visiting A, and believes there is a good vegetarian restaurant in town.*  
A: We should get something to eat, although I don't know if there's anything you'll like around here.  
S: Oh, isn't there a vegetarian restaurant around here?

Ladd proposes that there is an “ambiguity” in the usage of the negated question, whereby in situations like 1 it is used with the aim of confirming the questioner's existing belief of  $p$ , while in situations like 2 it is used with the aim of confirming the evidence the questioner has just received for  $\neg p$  in the face of their existing belief.

Ladd's initial observations about (American) English have spawned considerable further research reasoning about the pragmatic conditions that trigger the use of those particular type(s) of biased question, as well as if, and if so, how, the meaning of the question is semantically encoded into particular linguistic forms (Büring & Gunlogson 2000; van Rooij & Šafářová 2003; Romero & Han 2004; Asher & Reese 2007; AnderBois 2011; Northrup 2014). There has also been interest in how English biased questions are represented in terms of different syntactic forms. Early research focused on questions with syntactically “high” negation *-n't*, like those in 1-2, which, despite any potential differences in meaning, have the same form. “Low” negation *not* (as in 3) was not assumed to share these characteristics. Instead, low negation questions seemed to be used to genuinely question the truth of the negative proposition  $\neg p$  from current contextual evidence, with no involvement of any prior beliefs.<sup>2</sup>

- (3) Is there not a vegetarian restaurant around here?

The proposed distinction in the meanings of high and low negation questions has furthermore led to some syntactic proposals that these different negative markers *-n't* and *not* occupy different positions in the syntax (Cormack & Smith 2012; Krifka 2015).

---

<sup>2</sup>However, some later work incorporated “low negation” questions as having similar bias imports to “high negation” questions (van Rooij & Šafářová 2003; Asher & Reese 2007). Do-maneschi et al. (2017) propose that this may be the case in some contexts in German, but argue that it is not the case for English.

While much of the work on biased questions has focused on English, further work has indicated that biasing an interrogative, and having different forms to represent this, is also possible in other languages, with accounts for German (Büring & Gunlogson 2000), Japanese (Sudo 2013) and Hungarian (Gyuris 2016) contributing to Gaertner & Gyuris's attempt to delimit the possible bias profile space for questions cross-linguistically. This research has shown that biased questions may be marked syntactically, morphologically or prosodically, sometimes in addition to the use of a negation marker. For the remainder of this article, we use the term *question form* to refer to any linguistic form or realisation which is associated with a bias in the literature, regardless of whether the form itself consists of a morphosyntactic construction, a particle or morpheme, or an intonation pattern.

The rich introspective research on biased questions has led to various categorisations of the most important contextual phenomena that might license the use of these interrogatives, as well as the specific question forms associated with particular contexts. However, this has led to conflicting accounts regarding the role of the different factors, as well as the relationship between the semantic/pragmatic factors and the form of the interrogative (e.g. whether the negation is low or high in English, or whether particular particles are used in Japanese (Sudo 2013)). The literature has generally agreed that *original speaker bias* and *contextual evidence* are relevant factors for the licensing of these questions, but the ways in which those two types of influence have been understood and grouped together has varied considerably (see Romero 2020 for an overview). More recently, scholars have started gathering empirical data through experimental methodologies in order to tease apart the various proposals on the ways these biases and the linguistic forms that may express them interact. In the next section, we review the results of some of these studies. Though we acknowledge the need to explore this topic cross-linguistically, throughout this paper, we focus on the results of research on English biased questions, as this is where the most work has been done into understanding the interaction between the question form and the pragmatic factors that may license it.

## 2.2 Experimental tasks

A prolific research tradition in phonetics has investigated the realisation of polar questions across different languages, particularly in Romance (e.g. Crocco 2006; Escandell-Vidal 2017; Giordano 2006; Grice & Savino 1997; 2003; 2004; Henriksen et al. 2016). Plenty of the work in this tradition has not only been experimental in nature but it has also explicitly addressed question bias or speaker

certainty and epistemic stance, both in the perception (Armstrong & Prieto 2015; Vanrell et al. 2017; Orrico et al. 2019; Prieto & Borràs-Comes 2018) and production (Armstrong 2017; del Mar Vanrell et al. 2014) of polar questions. However, the (morpho)syntactic question forms that have driven formal semantic research into question bias have not been subject to the same quantity of experimental investigation.

For morphosyntactic question forms, the first experimental work to be conducted consisted of a series of acceptability judgment tasks investigating which interrogative structures participants find acceptable, given particular combinations of original speaker bias and contextual evidence. For example, Sailor (2013) finds no evidence for the sort of ambiguity of meaning for high negation in American English claimed by Ladd (1981). However, Sailor's study was small and focused on the acceptability of negative polarity items such as *either* in a high negation or a low negation construction, rather than the licensing of the interrogative structure itself. While the acceptability of negative polarity items can be extrapolated to the "biased" meaning of the question, the study does not directly address the production (or, interpretation) of the question forms in context.

Roelofsen et al. (2012) conduct an online acceptability judgment task investigating how natural different polar questions are when presented in various belief and bias contexts in English. Questions were shown as part of short cartoon contexts, with participants rating the items on a scale from 1 (natural) to 7 (unnatural). Roelofsen et al. (2012) found that high negation questions were acceptable in contexts with a preexisting speaker belief for *p*, and negative or no contextual evidence against *p* – as hypothesised. Perhaps surprisingly, they also found that low negation questions did not behave as differently as predicted to high negation questions, being preferred in contexts with prior speaker belief and negative or no contextual evidence.

While the study in Roelofsen et al. (2012) is broader in scope than Sailor (2013) and with a larger participant pool, there were some issues with the contextual evidence and speaker bias presented in their examples. In the negative cases, the evidence relied on a conversational implicature (for example, "Kate got a dog" does not entail that she did not also get a cat), which led to unexpected results (Roelofsen et al. 2012: 460). Secondly, as identified by Domaneschi et al. (2017: 6), the conditions with absence of speaker bias were *too* neutral. For example, Rose telling Jennifer that "Kate got a dog" out of the blue (without establishing, for example, that Kate prefers cats) renders any polar question somewhat unnatural.

Finally, we would point out that, for example, the belief that arises from a friend saying "I am *going to get* a dog" is different to "I *have got* a dog". It is

therefore questionable how strong the beliefs were in the various contexts, and to what extent this may have affected participants' rating of items. Roelofsen et al. (2012) find a "scale of speaker belief" in their results, and claim that "the neutral and negative [speaker bias] contexts in the experiment suggest the absence of positive speaker belief, but strictly speaking, they do not exclude it" (2012: 466). While this will always, to some extent, be the case, the speaker beliefs—or lack thereof—could have been strengthened.

Roelofsen et al.'s (2012) study presents an interesting potential counterpoint to the theoretical literature positing a true semantic difference between high and low negation questions. However, their materials included possible confounds which may have impacted on the results they saw.

We thus turn to Domaneschi et al. (2017), which builds on Roelofsen et al. (2012). Domaneschi et al. (2017) conduct lab-based production tasks in both English and German in order to "resolve" the conflicting claims in the literature as to which question forms (in these cases, low vs. high negation) are preferred in which contexts. Here, we discuss results from their English experiment only.

In their study, Domaneschi et al. (2017) asked participants to read through a short context that set up a prior belief for the participant, and then presented them with some contextual evidence that either supported, challenged, or was neutral in respect of that belief. Participants then chose one of a list of possible questions to ask, and produced this question into a microphone so that information about prosodic realisations could be recorded. The forms available to participants included positive polar questions (*Do you ...*), high and low negative polar questions (*Didn't you ... / Did you not ...*), as well as positive polar questions prefaced by a surprise marker (*Really? Did you ...*).

Domaneschi et al. (2017) find that participants have strong preferences as to which syntactic form of a question to produce given any one combination of prior belief and contextual evidence. Indeed, while low negation questions (LoNQ) are produced at a rate of 59% in contexts with a neutral bias and negative contextual evidence, high negation questions (HiNQ) are preferred in situations where there is a prior bias for the proposition expressed in the question, being produced at similar rates both in contexts with neutral evidence for the proposition (65%) and in contexts with evidence against the proposition (67%). The results for questions with negation are summarised in Table 1.

These results help us to understand what epistemic and evidential profile biased questions can have in English (and German). This study has since been replicated and extended upon by Maro et al. (2021), who find the same factors influencing production choice of biased questions in Italian.

Table 1: Attested pragmatic profile of polar questions with negation in English, from [Domaneschi et al. \(2017\)](#)

		speaker bias		
		p	neutral	$\neg p$
contextual evidence	p			
	neutral	HiNQ		
	$\neg p$	HiNQ	LoNQ	

However, these studies only begin to examine the complexities of licensing (and understanding) question bias. For example, [Maro et al. \(2021\)](#) note possible additional roles for the tense of the verb in the auxiliary construction.

Moreover, although [Domaneschi et al. \(2017\)](#) show that there are clear preferences for particular forms to be produced in each context, as summarised in Table 1, the distribution of choices for the remaining forms is not uniform, with certain “non-preferred” forms being produced above chance — indicating that perhaps these distinctions are not so clear cut. For example, in situations constructed so that participants had a prior bias towards the positive proposition  $p$ , and so that the contextual evidence was either neutral or in favour of  $\neg p$ , high negation questions were preferred. However, with a prior bias for  $p$ , when the contextual evidence was neutral, the second most preferred form (around 20% of the data) was a positive polar question, while when the contextual evidence was in favour of  $\neg p$ , around 25% of the data was a low negation question. So, although there seem to be clear preferences when people are asked to choose between different polar question forms, it is likely that in freer production the data would be considerably noisier, if including (polar) questions as the next relevant discourse move.

Furthermore, research by [Jamieson \(2018\)](#) indicates that the likelihood of a high negation question being produced in a given context appears to be stronger in certain contexts than others, despite the same basic bias profiles. Participants in Jamieson’s study were presented with 20 short contexts in a conversation with a friend. These contexts were designed to include a prior speaker belief and either negative or neutral contextual evidence. There were also 30 filler examples. Participants then chose between either a high negation question or a tag question to produce in response to these contexts.<sup>3</sup> All participants saw the same contexts, but the order was randomised.

<sup>3</sup>Tag questions are argued to be licensed in the same sort of belief and bias contexts as biased matrix questions ([Ladd 1981](#); [Reese & Asher 2006](#); [Malamud & Stephenson 2014](#); [Krifka 2015](#)).

Both of the contexts in 4 and 5 contain a prior belief, and the evidential context is neutral. We might, therefore, expect that high negation questions would be produced at around the same rate in both contexts. However, around 48% of participants chose to produce a high negation question in context 4, while only around 10% of the same participants did so in 5.

- (4) We are talking about a friend of ours who moved away. You are pretty sure she went to Berlin. You say:
  - Didn't she move to Berlin?
  - She moved to Berlin, didn't she?
- (5) A friend has said she is going to pick up her new pet. You are pretty sure it is a dog. You say:
  - Isn't she getting a dog?
  - She's getting a dog, isn't she?

This suggests that the choice to produce a high negation question may be more probabilistic than deterministic in any given context, even if the speaker has a prior belief of  $p$ . Any number of factors or normative expectations may be influencing the participant's choice of question at any one given time, including, potentially, other non-epistemic biases.

Finally, while data on language production and perception in both the phonetic and semantic traditions provide a strong indicator of overall usage distributions for particular question forms, given biases and contextual information – and thus can tap into intended meaning – language *comprehension* data may provide a very different picture. By looking at how participants process these questions with biases, we can gain a greater understanding of how the questions are interpreted, and thus speak to ongoing theoretical discussions of how these biases are expressed by the interrogative constructions at hand.

In the next section, we therefore explore how to design a processing study to investigate the real-time interpretation of biased questions. We firstly give an overview of related domains where (predictive) processing tasks have been used to explore issues of interpretation, and what sorts of issues surrounding the interpretation of question bias in particular can be addressed using these types of tasks. We then review two studies which have started to explore those issues: Tian et al. (2021) and Macuch & Jamieson (in prep). We do not present full results from these studies, but show how the methodologies can enhance the existing literature by investigating the question of whether the information structure of the question can allow participants to predict upcoming material, given a particular context.



Finally, we address the methodological choices that need to be made when designing a processing study to help understand *when* and *how* in interpretation biases are set for any given interrogative structure.

We continue to focus on English, and thus on how variation in (morpho)syntactic question forms may allow participants to predict upcoming material, but note that the topics and methodologies discussed in section 3 would extend to the other information structural cues seen in biased questions in other languages (see section 2.1).

## 3 Using processing tasks to explore question bias

### 3.1 What can processing studies bring to our general understanding of question bias?

Language production studies like those discussed in section 2.2 are a vital part of linguistic research, allowing us to shed a light on how language users plan and realise different sorts of linguistic elements, from individual phonemes to entire utterances. However, despite being deeply interconnected in conversation, the processes of producing and comprehending language are usually studied separately from one another, mostly in order to maximise experimental control and precision of measurement. For instance, given that language users may want to achieve various communicative goals with the bits of language they produce, focusing on how they choose to structure their linguistic signals given goals and (linguistic) contexts which are manipulated experimentally can shed light on the processes and motivations behind specific realisations. Conversely, focusing on how language users parse different types of linguistic signals can shed light on the inferences they draw when processing utterances in real time. Existent models of how sentences or otherwise non-sentential utterances are built up and how meaning is imparted in production therefore only tells us so much about how understanding is achieved in comprehension. In psycholinguistics, comprehension research sets about addressing the question of how language is understood based on the signal that is parsed – possibly including information from non-linguistic sources as well as multiple signaling modalities.

A key tenet of modern psycholinguistic research is that language processing is *incremental*; that is, comprehenders process information as they parse it, one chunk at a time (Altmann & Steedman 1988). While this means that to some extent processing happens on the fly for any individual lexical item, parsing a new word or chunk can both update previous beliefs about what the possible

meaning of the signal is (Levy et al. 2009), and, importantly, can help predict the interpretation of upcoming input in the remainder of the signal (Levy 2008; Kuperberg & Jaeger 2015).

Insights from at least three existing areas of processing research can help us conceptualise how to investigate question bias in terms of incremental (predictive) processing: information structure, extralinguistic context and negation.

Firstly, information structural cues like word order (e.g. Kaiser & Trueswell 2004; Yano & Koizumi 2018) and intonation (e.g. Kurumada et al. 2014; Roettger & Franke 2019) can allow the comprehender to make predictions about upcoming material in the signal, such as specific discourse referents or even non-referential lexical information, depending on context. The research on intonation is particularly relevant for the question bias agenda, as it complements the already extensive work in experimental phonetics referred to in section 2, which nevertheless does not focus on real-time processing and its impact on incremental interpretation.

To date, there has been considerably less work focused on the processing of interrogative structures compared to affirmatives, particularly in terms of how information structural cues might allow comprehenders to predict upcoming material in a question. However, for biased questions, we might assume that if a particular question form (for example, syntactically high negation in English) has a specific conventionalised meaning, its use in a particular pragmatic context may then facilitate processing of specific elements in the question through prediction.

Aside from work explicitly focused on pragmatic language interpretation, processing studies have generally focused on the predictability of particular interpretations given the immediate sentential context—that is, the context within the particular sentence being investigated, generally a specific word given the words preceding it. However, particularly pertinent for the study of question bias is the possible importance of extra-sentential context on predictability. Lemke et al. (2021) highlight the importance of investigating predictability based on extra-sentential context when investigating interrogatives, where predictability likely comes from the preceding discourse material, including the extralinguistic context, rather than from the form of the interrogative itself. For example, if two people are sharing a pizza, asking the question in 6 would most likely predict the continuation of *another slice* rather than, for example, *to go out on Saturday* or *some Christmas pudding*. This prediction is made purely on the extralinguistic context and not on any preceding material in the sentence.

- (6) Anna and Ben are sharing a pizza. Anna's plate is empty, and Ben says:  
Would you like (another slice/to go out on Saturday/some Christmas pud-

ding)?

Given what has been found in production data, therefore, it may be possible that based on a set of contextual beliefs and biases, comprehenders are able to predict either the form of the question itself (e.g. high or low negation in English), or further material in the main body of the question, based on the combination of context and question form.

While the example in 6 adapted from [Lemke et al. \(2021\)](#) demonstrates how extralinguistic context can influence the processing of positive polar interrogatives, the addition of negation makes the study of negated questions all the more complex.

The processing of negation has played a major role in the development of psycholinguistic research ([Kaup & Dudschig 2020](#)). Early studies (e.g. [Just & Carpenter 1971](#); [Fischler et al. 1983](#)) found that negative sentences are more troublesome for comprehenders than their positive counterparts, and are thus processed more slowly. Some more recent research ([Nieuwland & Kuperberg 2008](#); [Tian et al. 2010](#); [Dale & Duran 2011](#)) has indicated that the difficulty in processing of negation in these experiments may be purely circumstantial, due to a lack of contextual information provided to participants in the studies. Rather, if the negation is pragmatically licensed by a current question under discussion, processing proceeds unhindered as we would expect for affirmative sentences. However, even with high-level contextual support, evidence for the difficulties of processing negation still arises ([Darley et al. 2020](#)).

The role of negation in processing is thus another factor that must be taken into account when considering what processing studies can tell us about question bias, at least for languages where negation is part of the relevant question form, as in English and German. It may be that having a strong enough discourse context set up prior to the question itself resolves any issues that a negation marker may present; however, it may also be that the presence of negation—separate from the issue of the question form—could cause difficulty in processing.

Notably as negated biased questions relate to the affirmative proposition, it has been proposed that the negative marker in a high negation question is not a regular propositional negation, but an operator ([Repp 2009](#); [Romero 2015](#)) or a sort of metalinguistic negation ([Romero & Han 2004](#); [Krifka 2015](#)). Issues in the processing of low negation questions, but not in the processing of high negation questions, may be able to contribute to these debates, perhaps suggesting an operator based account.<sup>4</sup>

---

<sup>4</sup>There is little research on the processing of metalinguistic negation; however, limited results

Even removed from the study of negation, we might expect that if the interpretation of a biased question is tied to the conventionalised semantic meaning of a particular question form (such as an operator), then we might see different effects compared to if the interpretation is derived from pragmatic inferencing (as in e.g. [van Rooij & Šafářová 2003](#)).

In summary, on the basis of existing research in psycholinguistics, we can ask some of the following questions with regard to how question bias is processed:

1. Do biases originating from or relating to a speaker's original belief, or the contextual evidence, facilitate processing of particular types of biased interrogative structure?
2. Do information structural cues of a biased interrogative structure, such as high or low negation, facilitate processing downstream in the sentence, dependent on the context?
3. Does processing negation in a (biased) interrogative structure cause particular processing difficulties given any one question form or context?

Answering these questions about what comprehenders might or might not predict on the basis of the context and the question form can then contribute to the ongoing debate surrounding how bias is encoded into questions.

In the next section, we demonstrate, referring to studies from [Tian et al. \(2021\)](#) and [Macuch & Jamieson \(in prep\)](#), how (predictive) processing models and incremental interpretation can start to address these questions.

### **3.2 Exploring whether the question form allows comprehenders to predict upcoming material**

Combining the insights from research on predictability based on information structural cues with those from the research on extra-sentential cues and negation suggests that processing studies can shed light on where, and how, biases are incorporated into an interrogative, and whether particular linguistic forms encode these biases.

We firstly turn to a study from [Tian et al. \(2021\)](#). [Tian et al.](#) report results from a set of visual world eye tracking experiments in English and French — here, we focus on their English experiments. Participants were presented with

---

indicate that metalinguistic negation is processed in the same way as propositional negation ([Noh et al. 2013](#); [Blochowiak & Grisot 2018](#)).

a visual display containing four images, two distractors and two target images. Each target image represented a positive or a negative version of the same critical proposition, which always consisted of a binary predicate — e.g. an ironed shirt representing “The shirt is ironed” and a creased shirt representing “The shirt is not ironed”. Participants heard either a positive polar question 7, a question with high negation 8 or a question with low negation 9, along with a positive or negative answer to the critical question.

- (7) Has John ironed his father’s shirt? Yes, he has/ No, he hasn’t
- (8) Hasn’t John ironed his father’s shirt? Yes, he has/ No, he hasn’t
- (9) Has John not ironed his father’s shirt? Yes, he has/ No, he hasn’t

Participants’ task differed between two versions of the same experiment. We focus on the results from the experiment which involved not only listening to the question-answer pair and looking at the screen but also selecting the picture that best matched the dialogue. Participants’ gaze was measured throughout each trial as they listened to the dialogue, the measure of interest being the proportion of looks to both the  $p$  and  $\neg p$  targets during the question and the gap that ensued before the answer was played.

Tian et al. (2021) find that, across all question types, participants consider both  $p$  and  $\neg p$ , directing their gaze at both images. In the positive polar questions, participants show a bias for the  $p$  target starting at the noun (e.g. “shirt”) and continuing into the gap between the question and the answer. Though only reliable downstream in the sentence, descriptively, the bias seems to emerge as early as at the possessive (e.g. “his”), right after the main verb (e.g. “ironed”). In the high and low negation questions, on the other hand, participants show no reliable bias for either the  $p$  or  $\neg p$  targets, although, descriptively, the results for the high negation seem more akin to positive questions. Interestingly, the descriptive results in the high negation also suggest an early sensitivity to the  $p$  target at the main verb, although, here too, the differences are not reliable. All in all, the authors conclude that original speaker bias can account for the variation in gaze behaviour, with at least suggestive evidence for a bias towards  $p$  in high negation contexts, which would map onto the assumed bias that has been argued for in the literature. Regarding low negation questions, the authors tentatively conclude that low negation questions may also carry original speaker bias towards  $p$ , and that this may balance out any preference for the  $\neg p$  target, though the results are inconclusive.

While Tian et al. (2021) make a start at addressing the question of how questions with biases are processed, at this stage several open questions remain. First,

their study was designed to investigate whether polar interrogatives denote the set of all possible answers (Hamblin 1973; Karttunen 1977; Groenendijk & Stokhof 1984) regardless of their surface syntactic form, or whether different forms of interrogatives abstract from (Ginzburg & Sag 2001) or highlight (Roelofsen & Farkas 2015) one of the possible answers (e.g.  $p$  or  $\neg p$ ). Although they subsequently extrapolate to discussions of original speaker bias and contextual evidence, their materials were not designed to take this into account. In particular, the target questions were not situated within any context which set up biases and beliefs for the participant (as they were in, for example, Domaneschi et al. 2017) and so it is hard to make conclusions about how the different question structures might be processed in relation to these specific biases that have been highlighted in the formal semantic literature, which have otherwise been shown to affect controlled question production in the laboratory. Then, on the face of it, Tian et al.'s initial results clash with the existing claims and evidence from the literature, such that positive polar questions and high negation questions seem to have a similar processing profile. Lastly, it is unclear from the setup of their study whether any particular words were accentuated prosodically, which might provide contradictory cues to the morphosyntactic information manipulated explicitly.

We now turn to the study from Macuch & Jamieson (in prep). In this study, we conducted a set of self-paced reading experiments in English and German to investigate whether there is facilitation of processing one syntactic form of a negative question over another, given variations in original speaker belief with negative contextual evidence. We focus on the results from the English experiments.

In the study, participants were presented with short contexts, which set up either a prior belief of a given proposition  $p$ , or contained no specific prior belief. The context also contained some current evidence that suggested the truth of  $\neg p$ . The study therefore investigated the distinction between high negation questions and low negation questions found in the production study in Domaneschi et al. (2017), seen in the bottom row of Table 1. Recall that Domaneschi et al. (2017) found that, when presented with different polar question forms, participants preferred to produce high negation questions in contexts with an original belief of  $p$  plus contextual evidence for  $\neg p$ , while low negation questions were preferred in contexts with no specific original belief and contextual evidence for  $\neg p$ . An example of a context from Macuch & Jamieson (in prep) with a prior belief and contextual evidence for  $\neg p$  is given in 10, with the associated high and low negation test items.

- (10) Our friend is getting a new pet. You heard from her sister that it would be a cat. However, I tell you she is planning to take the pet for a lot of walks. You say:  
→ Hold on. Isn't she getting a cat?  
→ Hold on. Is she not getting a cat?

120 self-reported monolingual speakers of English read these fictional scenarios. Each scenario was followed by the test item, in which the words were masked with underscores. Participants were instructed to press the space bar to reveal one word at a time, following a moving window self-paced reading task design (Just et al. 1982). The time participants spent reading each word was measured, with any facilitatory processing expected to lead to shorter word reading times, and surprisal or difficulty leading to increased reading times.

Participants were instructed to read for comprehension, and were presented with comprehension questions after every 3 trials to ensure focus was kept on the task.

Macuch & Jamieson (in prep) find no reliable evidence for facilitation in reading either question form (so, *Isn't she...* or *Is she not...*) in either of the contexts tested, suggesting that although language users may have preferences as to which interrogative form to produce in context, these preferences may not translate into concrete expectations when it comes to the processing of the same question forms. Regarding the first of the questions set out in section 3.1, then, it does not appear that contextual beliefs and biases as put forward in the literature constrain the processing of either particular interrogative form, contra what the production biases in Domaneschi et al. (2017) might suggest.

However, unlike offline production data, which can be mapped more straightforwardly onto existing biased question accounts, no semantic account currently on offer provides concrete processing predictions, such that it remains unspecified *when* exactly in incremental interpretation contextual facilitation may occur. Indeed, even when a question is set against the belief and contextual evidence said to license a particular question form, there are still numerous topics that the actual biased question *could* be about. For example, in the above example in 10, the discourse continuation of *Isn't she...* could be about the cat, but it could also be about, for example, the friend being too busy with work to go out for regular walks. It could therefore be that a particular question form is preferred in context, but that facilitation only occurs once semantically constraining information like the main verb enters the frame, that is, once interpretation is further constrained by additional linguistic material in the signal. Facilitatory processing might be expected to occur not at the question form itself or immediately follow-



ing it but rather further downstream in the sentence once contentful information is processed and integrated into the unfolding discourse representation.

As at the question forms, Macuch & Jamieson (in prep) find no reliable evidence for facilitation in reading the main verb following the auxiliary verb construction for either high or low negation questions in either belief and bias context. However, we do find reliable evidence for a facilitatory effect at the noun (e.g. “cat”) in high negation questions which follow a discourse which has been set up to include a prior belief. Interestingly, this result seems to map onto the descriptive results reported by Tian et al. (2021), who find suggestive evidence that comprehenders have a preference to direct their gaze to images representing *p* when hearing the noun in a high negation question.

All in all, these results are indicative that comprehenders track epistemic information during the processing of negated polar questions, and that they might be sensitive to how that information interacts with the syntactic form of the question, such that they seem to expect specific lexical information downstream in the sentence when processing high negation questions. Ultimately, this might mean that there is no specific semantic meaning associated with the negation marker in a high negation question in English, but rather that the meaning of the auxiliary construction as a whole is inferred pragmatically in context.

However, this is only the tip of the iceberg regarding the study of processing biased questions. Firstly, Macuch & Jamieson (in prep) only investigate two cells from Table 1; how would low negation questions be processed in contexts with a prior belief but no contradictory evidence, for example?

Furthermore, it may be that word reading times are not a fine-grained enough measure of processing for this subtle interface phenomenon. Indeed, at the current level of analysis, the pragmatic expectations of interest are very categorical in nature – we expect there to be a preference for a particular question form, but not *how much* of a preference – which realistically means these expectations may not map well onto real-time processing demands and their accompanying behavioural correlates, unless there are strong enough pressures in place such that the relevant biases actually translate into measureable expectations.

In summary, the studies from Tian et al. (2021) and Macuch & Jamieson (in prep) illustrate how we can use processing tasks to help address the major topics of concern in the study of question bias, but there is much more that can be done. In the final section, we therefore reflect on some of the design decisions of these studies and overview how to potentially design new studies to help understand *when* and *how* in processing biases are set for any given interrogative structure.



### 3.3 Designing a processing study to investigate question bias

#### 3.3.1 The task

In order to study “online” interpretation processes, it is important to use tasks which measure participants’ reactions and understanding on a moment-by-moment basis, either in addition to or possibly without explicitly asking them to make “offline” decisions or judgments—in which various extralinguistic or metalinguistic pressures may affect the outcome of the task.

There are a variety of experimental tasks that can be used to investigate online processing, each with their own perks and pitfalls. The two studies we discussed in section 3.2 have employed two very different types of tasks: a self-paced reading task and a visual world eye tracking task. Here, we outline how to weigh up design choices for these two methodologies in particular, though a myriad of other psycholinguistic methods such as mouse tracking or ERP studies could be used to investigate question bias.

##### 3.3.1.1 Self-paced reading

Self-paced reading tasks (Just et al. 1982) allow for relatively cheap studies of psycholinguistic processing based on common computer equipment, making them particularly useful for gathering data both in the laboratory and online. In a standard “moving window” self-paced reading task, participants press a button to read a sentence one word or chunk at a time, the measurement of interest being the time spent between button presses. In general, self-paced reading tasks have been found to be comparable to other, more complex, cognitive measures of psycholinguistic processing (Marsden et al. 2018), which speaks to their usefulness as a method. A number of potential pitfalls arise, however, with using a self-paced reading task to understand a discourse phenomenon such as question bias.

Firstly, given that question bias is strongly situated in interactional contexts (e.g. Hennoste et al. 2017; Heritage & Raymond 2021), reading may not be particularly informative of how people actually process biased questions naturalistically. Of course, experimental psycholinguistics in general is not situated in rich interactive settings, and so this may not be a concern specific to this methodology, but the issue should nevertheless be taken into account. Constructing scenarios or (reported) dialogues which set up some form of a discourse context for the test sentence can go a certain way to solving this potential concern, at least under the assumption that one can extrapolate from reading processes to naturalistic comprehension in the first place.

A related concern when using self-paced reading, or any reading method for that matter, is that it is not possible to know how participants interpret the prosody of the read sentence. The licensing and interpretation of an interrogative can often be dependent on its prosodic structure (Pierrehumbert & Hirschberg 1990; Grice & Savino 1997; del Mar Vanrell et al. 2012; Hedberg et al. 2017), and in polar questions prosody can be an important factor whether or not intonation itself is the linguistic form said to mark any bias, as discussed earlier in section 2.2. Regarding negated biased questions where the bias is claimed to be marked primarily morphosyntactically, Domaneschi et al. (2017) find that while participants in their German production study consistently produce high negation questions with a final rise, participants in their English study produce high negation questions around 50% of the time with a final rise, and 50% of the time with a final fall.<sup>5</sup> Using a self-paced reading study, therefore, does not allow us to understand how prosody might impact the interpretation of a biased question, which could well be processed differently by different participants, depending on which contour they project on reading.

The obvious solution to the above issue is to present the stimuli auditorily to participants, thus making it possible to take account of effects of prosody on biased question interpretation. For that, one could retain the same general setup of a self-paced reading study but instead use a self-paced listening task (Ferreira, Henderson, et al. 1996), which follows the same format but with auditory rather than written stimuli. While self-paced listening is a rather underexplored psycholinguistic method, especially compared to self-paced reading, it has been shown to replicate the results of self-paced reading tasks (Ferreira, Anes, et al. 1996; Papadopoulou et al. 2013), and might thus be ideal for cases where written stimuli may not be so appropriate and where using other auditory delivery methods may not be feasible. One potential limitation of the method is that, compared to self-paced reading, self-paced listening might exacerbate even more the unnaturalness of the comprehension process, as it is very much unlike normal listening.

Secondly, recording the time it takes for a participant to press a button and read (or listen to) a particular segment in a test sentence may simply not be a fine-grained enough measure of online processing to allow us to determine how comprehenders process a complex semantic-pragmatic construction like a biased question, as discussed in the previous section. For instance, in Macuch & Jamieson (in prep), we find small differences in reading times at the noun when

---

<sup>5</sup>Domaneschi et al. (2017) propose that this intonational difference corresponds to the ambiguity in high negation described by Ladd (1981) and shown in examples 1-2.

comparing high negation questions in both bias conditions. It is likely that the effect size of interest is indeed very small, and so the granularity of a method such as self-paced reading may not allow capturing such subtle effects, which might otherwise be measureable using more fine-grained methods. Next, we discuss some of the trade-offs of using one such higher-resolution method, namely eye tracking.

### 3.3.1.2 Eye tracking

While self-paced reading still relies on the participant to actively engage in some element of the task beyond simply processing (i.e. pressing a button to move on to the next reading segment), eye tracking studies (Rayner 1978) allow for an even more natural reading process—given that eye movements are an essential part of processing (written) language. In eye tracking studies, eye movements and fixations are recorded as a participant reads through a sentence. Additionally, any “regressions” that the participant makes, such as re-reading a segment, can also be recorded. Where self-paced reading provides a singular measure for each word or chunk in a sentence, eye tracking can thus provide an extremely rich data set containing several measures for even one test item.

Eye tracking can rely on the same sort of measurement principles and assumptions as self-paced reading: namely, that when the eye is fixated on a particular target, that target is being considered, and that longer durations spent on any one target indicate that it is proving to be more difficult to process. However, unlike self-paced reading, eye tracking allows for more naturalistic reading processes, such as backtracking and re-reading. In the study of question bias, therefore, it may be possible using an eye tracking experiment to see, when a comprehender encounters the main contentful verb in the question, whether they then regress to the negation marker which they may have previously processed, indicating difficulty in integrating semantic information from verb. However, classical eye tracking studies involving the reading of sentences have the same issues as self-paced reading tasks in the study of biased questions: that questions are inherently interactional and so not so often read, and a participant’s interpretation of the sentence prosody cannot be controlled.

Visual world paradigms (Huettig et al. 2011) allow researchers to track participants’ eye movements towards particular (referential) targets, as they process auditory stimuli. This method was used to explore processing of different types of question forms in Tian et al. (2021), discussed in section 3. However, it does not trivially lend itself to investigating discourse situations which involve expect-

tations that might not be referential in nature<sup>6</sup>. In order to fully understand how biased questions are processed, the visual world paradigm would need to be combined with an effective way to present contextual evidence for the question to the participant. This could perhaps be done through video or audio story telling prior to the presentation of the visual world, or it could involve designing scenarios with epistemic expectations which can be encapsulated in concrete objects or actions as they are traditionally represented in visual world.

Both classic eye tracking and visual world paradigms can present researchers with an array of useful data that could help address major questions in the study of question bias, as outlined above. However, crafting useful designs where beliefs and biases can be straightforwardly operationalized might be a challenge for traditional setups. In the next section, we discuss some ways in which materials for such experiments can be designed.

### 3.3.2 Material design

Designing materials for any study of question bias is in itself a complex task, regardless of whether you are conducting a processing study or not. In this section, we focus on particular decisions that are important for processing studies—again looking in particular at self-paced reading and eye tracking—only touching on general decisions that must be made in any experimental study of question bias.

#### 3.3.2.1 Setting up the contexts

In order to capture potential interactions between speaker beliefs and contextual evidence in the processing of a biased question, the experimental stimulus must be combined with a discourse context which sets up beliefs and biases (as in e.g. [Domaneschi et al. 2017](#) and [Macuch & Jamieson in prep](#)). Ideally, these contexts should be as realistic and naturalistic as possible given the constraints of a controlled laboratory experiment, and importantly they must actually include the various biases that are being set up for the context, at an appropriate “strength”—for example, the biases should not be so strong that the very act of asking a question becomes irrelevant. In [Macuch & Jamieson \(in prep\)](#), we conducted an independent task prior to the main data collection in order to collect norms of the experimental discourse contexts, to understand the baseline likelihood of expecting certain epistemic beliefs in those contexts. Participants in the

---

<sup>6</sup>For research employing eye tracking while reading to investigate how spatial information can inform abstract language processing, see [Guerra & Knoeferle 2014; 2017](#)

norming study were asked to read a short scenario corresponding to the background contexts that would be used to set up the questions. Participants then answered a question pertaining to the scenario, which paraphrase the relevant negated question that would appear in the main study. An example can be seen in 11.

- (11) **Scenario:** Our friend is getting a new pet. You heard from her sister that it would be a cat.

**Q:** Do you think the person in this scenario is getting a cat?

Yes/No/Don't know

For contexts in which a prior speaker belief was to be established, we expected scenarios to show a bias for “yes” responses, while for scenarios which were to establish a belief, a bias against “yes” was expected.<sup>7</sup> By carrying out a norming study in advance, we can be more certain of ensuring that the beliefs and biases established in the scenarios map to the types of situations that have been detailed previously in the literature, and likewise that any expectations related to particular forms or contexts can be at least coarsely quantified. Yet, as discussed in Lemke et al. (2021), collecting study-specific norms (which are not necessarily generalisable to other data sets) might not be the best way of gauging variability or uncertainty in the expectations of interest; instead, relying on corpus estimates or norms which are open and freely available might be a more viable alternative.

### 3.3.2.2 Test stimuli

It is important that the test stimuli used in any processing study of biased questions meet the standard expectations of any well-conducted experimental paradigm: for example, in terms of appropriate numbers of stimuli, filler items and randomisation or counterbalancing procedures. However, there are other factors that must also be taken into account in a processing study.

For instance, in order to take valid measurements of the relevant moments in a processing task, each stimulus must follow the same basic frame so that the critical form is in a comparable position or time-window in every sentence. For example, in a straightforward declarative sentence, all items may follow the structure as in 12.

- (12) The boy walks to the shop.

DET NOUN VERB PREP DET NOUN

---

<sup>7</sup>Responses could be either in favour of “no” or “don't know” in the no belief conditions.

If the verb in the example in 12 is the critical form in the stimulus, to which the relevant facilitation or processing difficulties might be attributed, measurements can then be taken directly at the verb as well as at the words immediately following it in the case of reading, or a reasonable time-window around the verb can be normalised across different sentences in the case of unconstrained listening.

In conducting a processing task on question bias, we potentially run into problems given that biases can be associated with quite different question forms, as in the high and low negation forms in English. In this case, questions with low negation have one extra word compared to questions with high negation, which might directly impact both a reading and a listening experiment, given that the moment-by-moment unfolding of the signal cannot be matched as directly as when investigating biases which are mapped, for instance, onto different intonation contours over the same word or phrase. Furthermore, the position of the negation marker in the sentence is completely different (either first position in a high negation question, or after the subject in a low negation question). This makes direct comparison between the two types of structures difficult, and so while it may be intuitive to measure, for a given bias context, how each interrogative construction is processed, it is probably more convenient to instead compare, for each interrogative, how it is processed in each bias context.

A further potential problem arises in terms of the regions of interest in a sentence. In processing studies, it is important to avoid (as far as possible) having a critical region in the final region of the sentence. Participants' reading times often slow towards the end of a sentence (Conklin & Pellicer-Sánchez 2016), and the final region causes a general increase in processing times due to what Just et al. (1982) term the *sentence wrap up* effect, with comprehenders piecing together everything they've just processed.

Avoiding the final region may not be a major concern for the study of question bias in a language like English or German, but in a language like, for example, Japanese, where the negation or question particle comes at the end of the sentence, this would be an issue. Even in English and German, where fronted morphosyntactic constructions are the question forms of interest, cues at the end of a sentence might lead to qualitative shifts in interpretation, such as if the last word is accented in unexpected ways.

It is also important to avoid having critical regions at the very beginning of the stimulus. In English, this is a problem for potentially negated interrogatives, as the negation may be in the first region with a high negation question. Of course, many of these concerns are alleviated if one postulates that the question form itself serves as a cue to later information in the sentence, which is what the re-

sults from Macuch & Jamieson (in prep) and Tian et al. (2021) seem to suggest. Otherwise, issues with both the first and last regions may be accounted for by adding additional material on either side of the target sentence, which nonetheless introduces possibly unwanted sources of variation. In Macuch & Jamieson (in prep), we used short phrases like “hold on” or “wait a minute” before participants reached the question, which act as overt markers of surprise, perhaps making both question types more akin to negations prefaced by *really* in Domaneschi et al. (2017). If studying a language in which the negation or question particles came at the end of the sentence, it would be important that any additional material added still allowed the structure to be interpreted as an information seeking or confirmational question (and not, for example, a rhetorical question that did not require a response).

As well as the overall structure of the sentence, it is important to take into account factors such as the length and overall frequency of the individual words, and the predictability of the words in relation to the context, as these can affect measures of expectancy such as word reading times (Conklin & Pellicer-Sánchez 2016). Other factors which may also have an affect on the ease of interpretation of biased questions include the modality or tense of the question (Maro et al. 2021). If the stimuli is presented auditorily, it is also important that prosody is controlled for, as discussed above.

## 4 Conclusions

In this chapter we have provided an overview of what psycholinguistic processing studies can add to the existing literature on question bias. While introspective theoretical work has identified potential pragmatic constraints on the usage of different question forms, and production studies have narrowed down speakers’ preferences across those constraints, there are still major questions regarding if, and if so, how, a bias is semantically encoded into an interrogative form, as well as whether or not the choice of interrogative form is constrained by the combination of original speaker belief and contextual evidence. As we have shown, the empirical preferences for any particular form in different epistemic and evidential contexts seem to be much more gradient than originally discussed in the theoretical literature. Indeed, even in cases where one question form seems to be preferred over its potential alternatives, this preference is probabilistic in nature and dependent on a combination of factors that goes beyond the speaker bias/contextual evidence divide. Future work should address this variation, either explicitly attempting to capture it or at the very least providing an account as for



what the source of the natural gradience in the data might be.

With reference to [Macuch & Jamieson \(in prep\)](#) and [Tian et al. \(2021\)](#) in particular, we have shown that processing studies can add to this debate by investigating at what stage of the interrogative the bias appears to take hold, and whether there are specific relationships between the bias and the question form. However, the results from these studies are only the beginning of what psycholinguistic experimentation can add to the study of question bias, and so in section 3.3, we laid out some possible design decisions for processing studies of question bias, as well as what the choice of a method can bring to the topic. While we have centred our focus on processing studies, there are many other ways that experimental tasks could be employed to explore both the production and comprehension of question bias, such as through map tasks ([Anderson et al. 1991](#)), or rating the probabilities of beliefs ([Tonhauser 2016](#); [Degen et al. 2019](#)) given production of a question form. Ultimately, we believe that experimental psycholinguistic work should go hand in hand not only with theoretical work happening in formal semantics but also with other work, both empirical and theoretical, concerned with issues surrounding question bias, and we hope we have shown relevant ways to go about bridging these different approaches.

## Acknowledgements

Thanks to Hannah Rohde and colleagues at the Leibniz-Centre General Linguistics (ZAS) for helpful comments and discussion. E Jamieson gratefully acknowledges the support of the Economic and Social Research Council for award no. ES/T006919/1. Vinicius Macuch Silva gratefully acknowledges the support of the German Research Council via the Priority Program XPrag.de (DFG SPP 1727, FR 3482/1-2) as well as of the ERC Advanced Grant 787929 “Speech Acts in Grammar and Discourse” (SPAGAD).

## References

- Altmann, Gerry & Mark Steedman. 1988. Interaction with context during human sentence processing. *Cognition* 30. 191–238.
- AnderBois, Scott. 2011. *Issues and Alternatives*. UCSC. (PhD Thesis).
- Anderson, Anne H., Miles Bader, Ellen Gurman Bard, Elizabeth Boyle, Gwyneth Doherty, Simon Garrod, Stephen Isard, Jacqueline Kowtko, Jan McAllister, Jim Miller, Catherine Sotillo, Henry S. Thompson & Regina Weinert. 1991. The HCRC Map Task Corpus. *Language and Speech* 34(4). 351–366.



- Armstrong, Meghan E. 2017. Accounting for intonational form and function in Puerto Rican Spanish polar questions. *Probus* 29(1). 1–40.
- Armstrong, Meghan E & Pilar Prieto. 2015. The contribution of context and contour to perceived belief in polar questions. *Journal of Pragmatics* 81. 77–92.
- Asher, Nicholas & Brian Reese. 2007. Intonation and discourse: Biased questions. *Interdisciplinary Studies on Information Structure* 8. 1–38.
- Blochowiak, Joanna & Cristina Grisot. 2018. The pragmatics of descriptive and metalinguistic negation: Experimental data from French. *Glossa* 3(1). 1–23.
- Büring, Daniel & Christine Gunlogson. 2000. Aren't positive and negative polar questions the same? Ms., UCSC.
- Conklin, Kathy & Ana Pellicer-Sánchez. 2016. Using eye-tracking in applied linguistics and second language research. *Second Language Research* 32(3). 453–467.
- Cormack, Annabel & Neil Smith. 2012. The English Extended Verbal Projection. Available at [lingbuzz/001615](https://lingbuzz/001615).
- Crocco, Claudia. 2006. Prosodic and informational aspects of polar questions in Neapolitan Italian. In *3rd International Conference on Speech Prosody*. TUD Press.
- Dale, Rick & Nicholas D. Duran. 2011. The cognitive dynamics of negated sentence verification. *Cognitive Science* 35(5). 983–996.
- Darley, Emily J., Christopher Kent & Nina Kazanina. 2020. A 'no' with a trace of 'yes': a mouse-tracking study of negative sentence processing. *Cognition* 198.
- Dayal, Veena. 2017. *Questions*. Oxford: Oxford University Press.
- Degen, Judith, Andreas Trotzke, Gregory Scontras, Eva Wittenberg & Noah D. Goodman. 2019. Definitely, maybe: a new experimental paradigm for investigating the pragmatics of evidential devices across languages. *Journal of Pragmatics* 140. 33–48.
- del Mar Vanrell, Maria, Ignasi Mascaró, Francesc Torres-Tamarit & Pilar Prieto. 2012. Intonation as an encoder of speaker certainty: information and confirmation yes-no questions in Catalan. *Language and Speech* 56(2). 163–190.
- del Mar Vanrell, María, Francesc Ballone, Carlo Schirru & Pilar Prieto. 2014. Intonation and its interfaces in Sardinian polar questions. *Loquens*. 1.
- Domaneschi, Filippo, Maribel Romero & Bettina Braun. 2017. Bias in polar questions: Evidence from English and German production experiments. *Glossa* 2(1). 23.
- Escandell-Vidal, Victoria. 2017. Intonation and evidentiality in Spanish polar interrogatives. *Language and Speech* 60(2). 224–241.

- Ferreira, Fernanda, Michael D Anes & Matthew D Horine. 1996. Exploring the use of prosody during language comprehension using the auditory moving window technique. *Journal of Psycholinguistic Research* 25(2). 273–290.
- Ferreira, Fernanda, John M Henderson, Michael D Anes, Phillip A Weeks & David K McFarlane. 1996. Effects of lexical frequency and syntactic complexity in spoken-language comprehension: evidence from the auditory moving-window technique. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22(2). 324.
- Fischler, Ira, Paul A Bloom, Donald G Childers, Salim E Roucos & Nathan W Perry Jr. 1983. Brain potentials related to stages of sentence verification. *Psychophysiology* 20(4). 400–409.
- Fraenkel, Tamar & Yaacov Schul. 2008. The meaning of negated adjectives. *Intercultural Pragmatics* 5(4).
- Gaertner, Hans-Martin & Beáta Gyuris. 2017. On delimiting the space of bias profiles for polar interrogatives. *Linguistische Berichte* 251. 293–316.
- Ginzburg, Jonathan & Ivan Sag. 2001. *Interrogative investigations*. Oxford: OUP.
- Giora, Rachel, Ofer Fein, Jonathan Ganzi, Natalie Alkeslassy Levi & Hadas Sabah. 2005. On negation as mitigation: the case of negative irony. *Discourse Processes* 39(1). 81–100.
- Giordano, Rosa. 2006. The intonation of polar questions in two central varieties of Italian. In *Proceedings of speech prosody 2006*, paper 155.
- Grice, Martine & Michelina Savino. 1997. Can pitch accent type convey information status in yes-no questions? In Kai Alter, Hannes Pirker & Wolfgang Finkler (eds.), *Concept to Speech Generation Systems*, 29–38. Madrid: Universidad Nacional de Educacion a Distancia.
- Grice, Martine & Michelina Savino. 2003. Map tasks in Italian: Asking questions about given, accessible and new information. *Catalan Journal of Linguistics* 2. 153–180.
- Grice, Martine & Michelina Savino. 2004. Information structure and questions: evidence from task-oriented dialogues in a variety of Italian. In Peter Gilles & Jörg Peters (eds.), *Regional variation in intonation*, 161–189. Tübingen: Niemeyer.
- Groenendijk, Jeroen & Martin Stokhof. 1984. *Studies in the semantics of questions and the pragmatics of answers*. Amsterdam: University of Amsterdam. (PhD Thesis).
- Guerra, Ernesto & Pia Knoeferle. 2014. Spatial distance effects on incremental semantic interpretation of abstract sentences: evidence from eye tracking. *Cognition* 133(3). 535–552.

- Guerra, Ernesto & Pia Knoeferle. 2017. Visually perceived spatial distance affects the interpretation of linguistically mediated social meaning during online language comprehension: an eye tracking reading study. *Journal of Memory and Language* 92. 43–56.
- Gyuris, Beáta. 2016. New perspectives on bias in polar questions: A study of Hungarian -e. *International Review of Pragmatics* 8(1). 1–50.
- Hamblin, Charles Leonard. 1973. Questions in Montague English. *Foundations of Language* 10. 42–53.
- Hedberg, Nancy, Juan M. Sosa & Emrah Görgülü. 2017. The meaning of intonation in yes-no questions in American English: A corpus study. *Corpus Linguistics and Linguistic Theory* 13(2). 1–48.
- Hennoste, Tiit, Andriela Rääbis & Kirsi Laanesoo. 2017. Polar questions, social actions and epistemic stance. *STUF-Language Typology and Universals* 70(3). 523–545.
- Henriksen, Nicholas, Meghan E Armstrong & Lorenzo García-Amaya. 2016. The intonational meaning of polar questions in Manchego Spanish spontaneous speech. *Intonational grammar in Ibero-Romance: Approaches across linguistic subfields*. 181–205.
- Heritage, John. 2012. Epistemics in action: action formation and territories of knowledge. *Research on language & social interaction* 45(1). 1–29.
- Heritage, John & Chase Wesley Raymond. 2021. Preference and polarity: epistemic stance in question design. *Research on Language and Social Interaction* 54(1). 39–59.
- Huettig, Falk, Joost Rommers & Antje S Meyer. 2011. Using the visual world paradigm to study language processing: a review and critical evaluation. *Acta psychologica* 137(2). 151–171.
- Jamieson, E. 2018. Experimental evidence for the contextual acceptability of biased questions and tag questions. In *Proceedings of ConSOLE XXVI*, 331–349. Leiden: Leiden University Centre for Linguistics.
- Just, Marcel A. & Patricia A. Carpenter. 1971. Comprehension of negation with quantification. *Journal of Verbal Learning and Verbal Behavior* 10(3). 244–253.
- Just, Marcel A., Patricia A. Carpenter & Jacqueline D. Woolley. 1982. Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General* 111(2). 228–238.
- Kaiser, Elsi & John Trueswell. 2004. The role of discourse context in the processing of a flexible word-order language. *Cognition* 94(2). 113–147.
- Karttunen, Lauri. 1977. Syntax and semantics of questions. *Linguistics and Philosophy* 1. 3–44.

- Kaup, Barbara & Carolin Dudschig. 2020. Understanding negation: Issues in the processing of negation. In Viviane Déprez & M. Teresa Espinal (eds.), *The Oxford handbook of negation*, 634–655. Oxford University Press.
- Koike, Dale April. 1994. Negation in Spanish and English suggestions and requests: Mitigating effects? *Journal of Pragmatics* 21(5). 513–526.
- Krifka, Manfred. 2007. Negated antonyms: creating and filling the gap. In Uli Sauerland & Penka Stateva (eds.), *Presupposition and implicature in compositional semantics*, 163–177. London: Palgrave Macmillan.
- Krifka, Manfred. 2015. Bias in commitment space semantics: Declarative questions, negated questions and question tags. In Sarah D’Antonio, Mary Moroney & Carol Rose Little (eds.), *Proceedings of SALT 25*, 328–345. Stanford.
- Kuperberg, Gina R. & T. Florian Jaeger. 2015. What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience* 31(1). 32–59.
- Kurumada, Chigusa, Meredith Brown, Sarah Bibyk, Daniel F Pontillo & Michael K Tanenhaus. 2014. Is it or isn’t it: listeners make rapid use of prosody to infer speaker meanings. *Cognition* 133(2). 335–342.
- Ladd, Robert. 1981. A first look at the semantics and pragmatics of negative questions and tag questions. In Roberta Hendrick, Carrie Masek & Miller Mary Frances (eds.), *Papers from the Seventeenth Regional Meeting of the Chicago Linguistics Society* (CLS 17), 164–173. Chicago: Chicago Linguistics Society.
- Lee, Seung-Hee. 2015. Two forms of affirmative responses to polar questions. *Discourse Processes* 52(1). 21–46.
- Lemke, Robin, Lisa Schäfer & Ingo Reich. 2021. Modeling the predictive potential of extralinguistic context with script knowledge: the case of fragments. *PLOS ONE* 16(2). e0246255.
- Levy, Roger. 2008. Expectation-based syntactic comprehension. *Cognition* 106(3). 1126–1177.
- Levy, Roger, Klinton Bicknell, Tim Slattery & Keith Rayner. 2009. Eye movement evidence that readers maintain and act on uncertainty about past linguistic input. *Proceedings of the National Academy of Sciences* 106(50). 21086–21090.
- Macuch, Vinicius & E Jamieson. in prep. Processing negated polar questions with varying epistemic biases in English and German. Manuscript in preparation.
- Malamud, Sophia & Tamina Stephenson. 2014. Three ways to avoid commitments: Declarative force modifiers in the conversational scoreboard. *Journal of Semantics* 32(2). 275–311.
- Maro, Maria Di, Antonio Origlia & Francesco Cutugno. 2021. Polar express: polar question forms expressing bias-evidence conflicts in Italian. *International Journal of Linguistics*.

- Marsden, Emma, Sophie Thompson & Luke Plonsky. 2018. A methodological synthesis of self-paced reading in second language research. *Applied Psycholinguistics* 39(5). 861–904.
- Nieuwland, Mante S. & Gina R. Kuperberg. 2008. When the truth is not too hard to handle. *Psychological Science* 19(12). 1213–1218.
- Noh, Eun-Ju, Hyeree Choo & Sungryong Koh. 2013. Processing metalinguistic negation: evidence from eye-tracking experiments. *Journal of Pragmatics* 57. 1–18.
- Northrup, Oliver. 2014. *Grounds for commitment*. UCSC. (PhD Thesis).
- Orrico, Riccardo, Renata Savy & Mariapaola D’Imperio. 2019. The perception of speaker certainty in Salerno Italian intonation. In Sasha Calhoun, Paola Escuerdo, Marija Tabain & Paul Warren (eds.), *Proceedings of the 19th International Congress of Phonetic Sciences*, 2946–2950.
- Papadopoulou, Despina, Ianthi Tsimpli & Nikos Amvrazis. 2013. Self-paced listening. In *Research methods in second language psycholinguistics*, 66–84. London: Routledge.
- Pierrehumbert, Janet & Julia Hirschberg. 1990. The meaning of intonation contours in the interpretations of discourse. In Philip Cohen, Jerry Morgan & Martha Pollack (eds.), *Intentions in communications*, 271–311. Cambridge, MA: MIT Press.
- Prieto, Pilar & Joan Borràs-Comes. 2018. Question intonation contours as dynamic epistemic operators. *Natural Language & Linguistic Theory* 36(2). 563–586.
- Raymond, Chase Wesley & John Heritage. 2021. Probability and valence: two preferences in the design of polar questions and their management. *Research on Language and Social Interaction* 54(1). 60–79.
- Rayner, Keith. 1978. Eye movements in reading and information processing. *Psychological Bulletin* 85(3). 618–660.
- Reese, Brian & N Asher. 2006. Prosody and the interpretation of tag questions. In E Puig-Waldmüller (ed.), *Proceedings of Sinn und Bedeutung 11*, 448–462. Barcelona: Universitat Pompeu Fabra.
- Repp, Sophie. 2009. *Negation in Gapping*. Oxford: Oxford University Press.
- Roelofsen, Floris & Donka Farkas. 2015. Polarity particle responses as a window onto the interpretation of questions and assertions. *Language* 91(2). 359–414.
- Roelofsen, Floris, Noortje Venhuizen & Galit Weidman Sassoon. 2012. Positive and negative questions in discourse. In Eva Csipak & Hedde Zeijlstra (eds.), *Proceedings of Sinn und Bedeutung 17*, 519–536. Göttingen: University of Göttingen.

- Roettger, Timo B & Michael Franke. 2019. Evidential strength of intonational cues and rational adaptation to (un-) reliable intonation. *Cognitive Science* 43(7). e12745.
- Romero, Maribel. 2015. High negation in subjunctive conditionals and polar questions. In *Proceedings of Sinn und Bedeutung* 19, 499–516. University of Göttingen.
- Romero, Maribel. 2020. Form and function of negative, tag and rhetorical questions. In Viviane Déprez & M. Teresa Espinal (eds.), *The Oxford handbook of negation*. Oxford: Oxford University Press.
- Romero, Maribel & Chung-Hye Han. 2004. On negative yes/no questions. *Linguistics and Philosophy* 27(5). 609–658.
- Sailor, Craig. 2013. Questionable negation. Presentation given at LSA Annual Meeting, University of Michigan, 2nd January.
- Sudo, Yasutada. 2013. Biased polar questions in English and Japanese. In Daniel Gutzmann & Hans-Martin Gaertner (eds.), *Beyond Expressives: Explorations in Use-Conditional Meaning*, 275–296. Leiden: Brill.
- Tian, Ye, Richard Breheny & Heather J. Ferguson. 2010. Why we simulate negated information: a dynamic pragmatic account. *Quarterly Journal of Experimental Psychology* 63(12). 2305–2312.
- Tian, Ye, Bob van Tiel, Élise Clin & Richard Breheny. 2021. Representing polar questions. *Journal of Psycholinguistic Research* 50. 1535–1555.
- Tonhauser, Judith. 2016. Prosodic cues to presupposition projection. *Semantics and Linguistic Theory* 26. 934.
- Vanrell, Maria del Mar, Meghan E Armstrong & Pilar Prieto. 2017. Experimental evidence for the role of intonation in evidential marking. *Language and speech* 60(2). 242–259.
- van Rooij, Robert & Maria Šafářová. 2003. On polar questions. In R.B. Young & Y Zhou (eds.), *Proceedings of SALT XIII*, 292–309. Cornell University, Ithaca, NY: CLC Publications.
- Yano, Masataka & Masatoshi Koizumi. 2018. Processing of non-canonical word orders in (in)felicitous contexts: evidence from event-related brain potentials. *Language, Cognition and Neuroscience* 33(10). 1340–1354.