GLOSSA psycholinguistics

Razguliaeva, M., Onoeva, M., Šimík, R., Meyer, R., & Hrdinková, K. (2026). Processing of Russian and Czech polar questions: Evidence for the effect of question bias. *Glossa Psycholinguistics*, 5(1): X, pp. 1–50. DOI: https://doi.org/10.5070/G6011.42487

eScholarship University of California

Processing of Russian and Czech polar questions: Evidence for the effect of question bias

Mariia Razguliaeva, Humboldt-Universität zu Berlin and Leibniz-Zentrum Allgemeine Sprachwissenschaft (ZAS), mariia.razguliaeva.1@hu-berlin.de

Maria Onoeva, Charles University and Humboldt-Universität zu Berlin, onoevam@ff.cuni.cz

Radek Šimík, Charles University, Radek.Simik@ff.cuni.cz

Roland Meyer, Humboldt-Universität zu Berlin, roland.meyer@hu-berlin.de

Kateřina Hrdinková, Prague, Czechia, kacka.hrdinkova@gmail.com

While bias in polar questions has generated an increasing body of cross-linguistic research, relatively little is known about their processing. Adapting the experimental design of Tian et al. (2016, 2021), we conducted a visual world eye-tracking study of assertions, polar questions with and without verb inversion, and polar questions containing particles *razve* and *snad*, respectively, in Russian and Czech. While participants heard small dialogues, their fixations on pictures showing either the positive outcome implied in the stimulus or its opposite were recorded. While, in assertions, participants focused their attention on the picture corresponding to the expressed polarity, fixations in questions became concentrated on the positive picture, largely independently of the polarity expressed in the stimulus. In questions with particles *razve* and *snad*, attention remained balanced between the positive and the negative pictures. We argue that the preference for the positive picture in questions is related to pragmatic processing, specifically, to the influence of a positive epistemic bias in polar questions.

1. Introduction

Research on polar questions (PQs) (and the corresponding assertions; see (1)) has yielded a rich landscape of theories attempting to account for their structural properties, meaning, and use. Most literature relies on introspective evidence or evidence from offline experiments (such as acceptability ratings, e.g., Geist & Repp, 2023; Roelofsen et al., 2012; truth or felicity judgments, e.g., Domaneschi et al., 2017; sentence–picture matching tasks, and the like). Comparatively little is known about how PQs are processed during online comprehension. A processing perspective may not just provide a window into the mental representation of PQs and speech acts more generally, but supply evidence relevant for semantic theories, to the extent that these theories yield predictions about processing.

(1) a. *Polar questions*
   (i) Did Kate clean the shoes? / Kate cleaned the shoes? [positive]
   (ii) Didn't Kate clean the shoes? / Did Kate not clean the shoes? / Kate didn't clean the shoes? [negative]

   b. *Corresponding assertions*
   (i) Kate cleaned the shoes. [positive]
   (ii) Kate didn't clean the shoes. [negative]

Using eye-tracking and, in particular, the visual world paradigm, we explore how PQs and related utterances are processed by Russian and Czech speakers. Building on and partially replicating recent research on English and French (Tian et al., 2021), we further support the finding that PQs elicit looks to a picture that corresponds to the positive prejacent, independently of the polarity of the PQ. This means that all PQs like (1a), regardless of their polarity, draw the comprehender's attention primarily to a picture with clean (rather than dirty) shoes. We also show that this behavior contrasts with the processing of assertions, where the polarity maps more transparently to the picture fixated on: an assertion like (1b-i) yields looks to clean shoes, but its negative counterpart (1b-ii), to dirty shoes. We argue that this result is, *prima facie*, at odds with some dominant theoretical accounts of PQs, in particular: (i) the idea that a PQ denotes a set of the two propositions which correspond to its possible answers (Hamblin, 1973) and (ii) the idea that a PQ denotes a singleton set whose sole member is the proposition proffered by the PQ, i.e., either the positive or the negative one, depending on the polarity of the PQ (e.g., Biezma & Rawlins, 2012). Instead, our eye-tracking evidence suggests the prominence of the speaker's so-called *positive epistemic bias*, i.e., the speaker's implicit assumption that the positive prejacent of the PQ holds, regardless of the polarity of the PQ. In positive PQs, positive bias has been argued to represent the speaker's default assumption or expectation (e.g., AnderBois, 2019). By asking (1a-i), for instance, the speaker may convey her expectation that Kate was supposed to clean her shoes, and the PQ serves to check whether the expectation is met. Also, negative PQs have been

argued to give rise to positive epistemic bias, although of a different kind. More specifically, they imply that the speaker believed – at least until shortly before uttering the PQ – that the positive prejacent is true (e.g., Romero & Han, 2004). That is, (1a-ii) implies that the speaker believed that Kate did, in fact, clean the shoes (even though the currently relevant evidence might suggest otherwise). Given the dominant looks to the positive picture, we suggest that positive epistemic bias is a conventional part of PQ meaning.

The motivation for investigating Russian and Czech is that these languages exhibit theoretically relevant intra- and cross-linguistic variation in the formal encoding of PQs. Russian has two main strategies for forming default PQs: (i) verb fronting accompanied by the particle *li* attached to the fronted verb (2a), and (ii) a specific prosodic tune accompanied by prosodic prominence on the finite verb, which appears in the post-subject position (2b) (the typical word order for declarative clauses).[1] The particle *li* is morphologically related to the disjunction *ili* 'or', a cross-linguistically common source for PQ particles (Metslang et al., 2017). Korotkova (submitted) treats *li* as a binary alternative operator that can combine with questions, so we hypothesized that the particle makes both propositional alternatives implied in Hamblin's (1973) PQ semantics highly salient (as compared to the Czech counterpart, which exhibits no particle; see below), which in turn predicts more balanced looks to the pictures representing the two alternatives (see note 2 for our linking hypothesis considerations). We will see tentative evidence in favor of this hypothesis in our results.

(2) *Russian polar questions*

a. (Ne) Počistila li Katja utrom sapogi?
NEG cleaned LI Katja morning boots
'Did Katja (not) clean the boots in the morning?' V1 + *li*

b. Katja (ne) POČISTILA utrom sapogi?
Katja NEG cleaned morning boots
'Did Katja (not) clean the boots in the morning?' V2

The genealogically related Czech makes use of two strategies which are superficially similar to Russian, but also different in important respects. (i) The verb-fronting strategy, accompanied by final rising intonation (indicated by ↑), represents the default way of forming a PQ in Czech (3a). Crucially, however, Czech uses no particle, and thus provides an ideal background against

[1] In this study, we investigate Russian V2 PQs with prosodic prominence on the verb. In contrast to rising declaratives in English, they are natural in a neutral, "out-of-the-blue" context (Esipova, 2025). Generally, PQs with declarative word order are predominant in spoken Russian (Onoeva & Staňková, 2025), compared to V1 PQs with *li*, which also speaks in favor of them representing a canonical PQ marking strategy. Another type of PQs, usually marked with a pitch accent on the last stressed syllable (Esipova, 2025), which are potentially more similar to English rising declaratives (Munteanu & Kiss, 2025), will not be discussed in this article.

which to evaluate the contribution of *li* in Russian, on top of any effect that results from the interrogative (V1) word order. (ii) PQs with declarative word order are also available in Czech (3b), but are typical for contexts suggesting the truth of the prejacent (so-called *evidential bias*; Staňková & Šimík, 2025; Sudo, 2013). In this regard, they resemble rising declaratives in English (Gunlogson, 2002, 2008; Malamud & Stephenson, 2015). The comparison between (2b) and (3b) could, therefore, reveal the effect of evidential bias on the processing of PQs. We will see no evidence of this type of bias in our results. This is particularly evident in the processing of negative V2 PQs, which still trigger looks to the positive picture. We interpret this as the processing prominence of the speaker's prior epistemic bias, opposite in polarity to the negative evidential bias.

(3) *Czech polar questions*

a. (Ne-) Umyla Dita včera holínky↑?
NEG cleaned Dita yesterday rubber.boots
'Did Dita (not) clean the rubber boots yesterday?' V1

b. Dita (ne-) umyla včera holínky↑?
Dita NEG cleaned yesterday rubber.boots
'Dita (didn't) clean(ed) the rubber boots yesterday?' V2

The idea that bias plays a pivotal role in the processing of PQs will also receive preliminary support from an experiment involving particles – *razve* in Russian and *snad* in Czech – which convey the conflict of the epistemic and evidential biases.

This study is organized as follows. In Section 2, we discuss some prominent theories of polar questions and their implications for online processing. Section 3 provides an overview of the experimental paradigm and setup that we have used. Sections 4–7 report on the individual experiments – on the processing of assertions (Section 4), questions (Section 5), and questions with particles (Section 7). Section 6 presents an analysis comparing the results of the experiments on assertions and questions. Section 8 discusses the overall results and highlights some avenues for future research.

2. Theories of polar questions and their implications for processing

In this section, we outline some prominent theoretical approaches to the meaning of polar questions. We concentrate not so much on the technical details of the theories, but rather on the implications they have for PQ processing. We discuss the traditional and still dominant proposition-set approach, the more recent singleton-set or highlighted alternative approach, and, finally, approaches that capitalize on the importance of question biases. For more detailed surveys, the reader is referred to Cross & Roelofsen (2014), Krifka (2011), Krifka & Trinh (2025), Roelofsen (2019), and Romero (2024).

PQs (and questions more generally) differ from the corresponding assertions in that they "are not naturally construed as being true or false" (Roelofsen, 2019), which, in turn, means that their meaning cannot straightforwardly be captured in terms of truth conditions (Davidson, 1967). The traditional way to deal with this problem is to assume that "a question sets up a choice-situation between a set of propositions, namely, those propositions that count as answers to it" (Hamblin, 1973, p. 48). The meaning of questions is, thus, reduced to the meaning of something more familiar – propositions and their truth conditions. Our example PQ – *Did Kate clean the shoes?* – therefore denotes the set {'Kate cleaned the shoes', 'Kate didn't clean the shoes'} (or $\{p, \overline{p}\}$, for short). This core idea has found its way into most semantic theories of questions, such as the various incarnations of alternative semantics (Dayal, 2016; Hamblin, 1973; Karttunen, 1977), partition semantics (Groenendijk & Stokhof, 1984), the structured meaning theory (Krifka, 2001, 2006; Szabolcsi, 1981), and inquisitive semantics (AnderBois, 2019; Ciardelli et al., 2019). What are the implications of this basic theoretical assumption for processing? When a discourse participant is exposed to a PQ, she is expected to consider both propositional alternatives that constitute the PQ's denotation. In terms of our experimental design, the participant is expected to pay attention to both pictures representing these propositions – the picture of shoes that are clean (p) and the picture of shoes that are not clean ($\overline{p}$).[2] We point out that this approach predicts no difference in the processing of positive and negative PQs, as both denote the same set of propositions.[3] At the same time, however, the proposition-set approach can be, and typically is, complemented with additional assumptions about question bias, which can modulate the predictions significantly (see below).

Empirical problems with the proposition-set theory – esp. the problematic consequences of the assumption that positive and negative PQs have equivalent denotations – have led to new proposals. The singleton-set approach postulates that a PQ denotes the singleton set whose sole member is the proposition that is proffered in the PQ; that is, a positive PQ denotes $\{p\}$ and a negative PQ $\{\overline{p}\}$ (Abels, 2007; Biezma & Rawlins, 2012; Roberts, 1996/2012). A related proposal is that while a PQ denotes a set of two propositions, the proffered proposition is, in some sense, more prominent or salient. van Rooy & Šafářová (2003) argue that the proffered alternative has a higher "utility value" than its implicit polar alternative. Roelofsen & van Gool (2010) and Roelofsen & Farkas (2015) consider the proffered alternative "highlighted" and argue that highlighting manifests itself, for instance, in the interpretation of short answers to PQs (– *Did*

[2] As noticed by an anonymous reviewer, it is important to specify how we model the link between eye-tracking measurements and theoretical semantic concepts (*linking hypothesis*, see Magnuson, 2019, for an overview). Following Tian et al. (2021), we did not seriously consider testing specific linking hypotheses, but we believe that our experimental design, which included a picture selection task, is compatible with the Referential Belief linking assumption, as presented in Degen et al. (2021). In the context of our study, we interpret it as linking the proportions of fixations on a picture to the participants' belief that the referent depicted on that picture renders the predication made by the speaker true.

[3] More precisely, a PQ of the form $?\overline{p}$ denotes the set $\{\neg p, \neg\neg p\}$, which is equivalent to $\{\neg p, p\}$ and to $\{p, \neg p\}$.

Kate clean the shoes? – Yes, she did/#didn't. vs. *– Did Kate not clean the shoes? – Yes, she did(n't).*). One piece of empirical evidence for these approaches comes from a systematic comparison of PQs with the corresponding alternative questions (*Did Kate clean the shoes or not?*). Only the latter type makes both alternatives equally salient (which is evident in the impossibility of using short answers, such as *yes*). Biezma & Rawlins (2012) also notice that alternative questions, as opposed to PQs, involve an overt disjunction (*or*), which they consider to be the operator responsible for deriving the set of two alternatives (following Alonso-Ovalle, 2006).[4] Both the singleton-set approach and the "imbalanced-set" approach predict that the participants in our experiment will pay closer attention to the picture representing the proffered alternative. The processing of PQs is, thus, expected to be qualitatively similar to the processing of the corresponding assertions.

Most current approaches to the meaning of PQs build on the nowadays standard assumption that linguistic meaning is multi-dimensional (see, e.g., Gutzmann, 2015). It is thus routinely assumed that the basic meaning of PQs (often Hamblin-style) is complemented with secondary meanings referred to as *question biases*. Pre-theoretically speaking, by choosing to ask ?p rather than ?$\overline{p}$, the speaker conveys that she considers p more informative, desirable, contextually supported, and the like. The prediction of this type of approach, pursued in different ways by Büring & Gunlogson (2000), van Rooy & Šafářová (2003), AnderBois (2019), a.o., matches the one discussed in the previous paragraph. Deeper research on question bias has shown, however, that the situation is more complex. In particular, while negative PQs are biased towards the negative answer, they typically also convey that the speaker thought or believed the positive alternative to be true. A PQ like *Did Kate not clean the shoes?* likely reacts to a situation suggesting that Kate did not clean the shoes – so-called *contextual* or *evidential bias* –, but, at the same time, conveys that the speaker believed (or thought, hoped, expected) the opposite (i.e., that Kate did clean the shoes) to be true – so-called *speaker-related* or *(original/prior) epistemic bias*; see AnderBois (2019), Büring & Gunlogson (2000), Gärtner & Gyuris (2017), Geist & Repp (2023), Goodhue (2022), Romero (2024), Romero & Han (2004), and Sudo (2013). For the sake of explicitness, we define epistemic and evidential bias as in (4) and (5), respectively.[5]

(4) Epistemic bias [a.k.a. original speaker bias] for a proposition p:
Belief or expectation of the speaker that p is true, based on her epistemic state prior to the current situational context and conversational exchange.
(Domaneschi et al., 2017; cited from Romero, 2024, p. 281)

[4] It is this assumption that triggers the expectation that a PQ particle morphologically related to the disjunction – in our case, the Russian particle *li* – gives rise to the Hamblin-style denotation.

[5] The division into two bias types is a simplification. Epistemic bias, for instance, is definitely not a monolith: there are arguably different flavors of epistemic bias (related to beliefs, desires, etc.; see Sudo, 2013) and possibly also different forces (considering something necessary vs. just possible; see Šimík to appear).

(5) Evidential bias [a.k.a. contextual evidence bias] for a proposition *p*:
Expectation that *p* is true (NB: possibly contradicting prior belief of the speaker) induced by evidence that has just become mutually available to the participants in the current discourse situation.
(Büring & Gunlogson, 2000; cited from Romero, 2024, p. 282)

Given that a single PQ can convey two conflicting biases (one for p, the other for $\overline{p}$), it is not straightforward to translate this type of approach into concrete predictions about processing. The situation is further complicated by the fact that the bias profile of PQs is affected in complex ways by contextual and formal manipulations (the latter involving word order alternations, different prosodic profiles, or particles).

Tian et al.'s (2021) experimental evidence points to the prominence of positive bias, regardless of the polarity of the PQ: both positive and negative PQs elicit more looks to the positive picture. One way of interpreting this result is that the epistemic bias is more salient than the evidential bias. That negative questions – esp. in questions with high negation – typically imply positive epistemic bias, even in the absence of contextual support, is generally agreed upon (see the references above). This assumption is controversial for positive questions, which are sometimes even argued to convey a negative epistemic bias, particularly in situations where they convey a positive evidential bias. However, without any contextual support, i.e., in an "out-of-the-blue" context, which corresponds to Tian and colleagues' and our experimental design, positive questions do, indeed, seem to convey positive bias, which is arguably of the epistemic type. This can be exemplified by questions like *Is John left-handed?* or *Is today the seventeenth?*, which imply – even in the absence of contextual evidence – that the speaker has a reason to think or hope that John is left-handed / that it is the seventeenth (AnderBois, 2019; Bolinger, 1978).[6] For the purposes of our experiment, we hypothesize that in the absence of strong contextual support (or another kind of support, e.g., by particles – see below), epistemic bias is more salient than evidential bias: While the former is arguably inherently associated with PQs and is positive ("by default" in positive PQs and due to negation in negative PQs), the latter arises in specific situations or contexts.

Question bias – and, particularly, the truth-conditional conflict between epistemic and evidential bias – surfaces prominently if certain particles are used. Russian *razve* is a case in point: a PQ introduced by *razve* is felicitous in a context in which it has been suggested that the proffered alternative (positive in (6)) is true, and conveys that the speaker expected (before the contextual evidence became apparent) the opposite ($\overline{p}$) to be true (Geist & Repp, 2023; Korotkova, 2023, submitted; Restan, 1969). Geist & Repp (2023) approximate PQs with *razve* to English PQs

[6] This type of bias seems particularly salient in questions whose prejacent is stereotypically unlikely – being left-handed is less likely than being right-handed, for instance. Correspondingly, the question *Is John right-handed?* has been argued to be completely unbiased (AnderBois, 2019). The results of Tian et al. (2021), as well as our results in this article, seem to suggest a more general validity of this type of "default" bias of positive PQs.

with conversational *really* (Bill & Koev, 2022; Romero & Han, 2004); however, later in Repp & Geist (2025), they conclude that the questions are not the same in meaning. When negated, the former allow both outer- and inner-negation interpretations, whereas the latter are compatible with inner negation only. Furthermore, Korotkova (submitted) suggests that the speaker asks the PQ with *razve* in (6) to merely resolve a given epistemic conflict between her belief (i.e., Kostja did not pick the flowers) and the evidence (it appears he did), with no preferred option; for PQs with *really*, there is a strong conviction that $\overline{p}$.

(6) Razve Kostja narval na pole cvety?
RAZVE Kostja picked on field flowers
'Did Kostja pick flowers on the field? (It appears he did [evidential bias]; I thought he didn't [epistemic bias])'

If question bias affects the processing of PQs, we expect it to be particularly salient with particles like *razve*.

3. Experiments and their overall design

Our experiments[7] partially replicate Tian and colleagues' (2016, 2021) experiments on English and French, but in one unified setup. In the first study, Tian et al. (2016) explored the processing of assertions without context, using an eye-tracking visual world paradigm. They asked participants to listen to positive and negative assertions in simple and in cleft form, as in (7), while looking at visual scenes containing a representation of the implied state (a closed window in (7)) and of the opposite state (an open window), along with unrelated distractors. The results show a difference in processing between positive and negative sentences in the simple conditions (7a/b), but not in the cleft conditions (7c/d). While listening to simple positive sentences, participants directed their attention to the picture corresponding to the positive state as soon as the verb was pronounced. During the processing of simple negative sentences, participants hesitated for a few hundred milliseconds between the positive and negative states before shifting their attention to the negative one.

(7) a. Matt has shut his dad's window.
b. Matt hasn't shut his dad's window.
c. It is Matt who has shut his dad's window.
d. It is Matt who hasn't shut his dad's window.

The aim of Tian et al. (2021) was to investigate the online processing of different forms of English and French PQs, once again in a contextless visual world paradigm setup. All test sentences appeared in three conditions in English, as in (8), and two conditions in French, as in (9).

[7] The study was preregistered on Open Science Framework, see Šimík et al. (2023).

(8) *English*

a. Has Anne closed her mom's umbrella? [positive]
b. Hasn't Anne closed her mom's umbrella? [high-negation]
c. Has Anne not closed her mom's umbrella? [low-negation]

(9) *French*

a. Jean a-t-il repassé cet après-midi sa chemise? [positive]
   Jean has-he ironed this afternoon his shirt
   'Has Jean ironed his shirt this afternoon?'

b. Jean n'a-t-il pas repassé cet après-midi sa chemise? [negative]
   Jean NEG-has-he NEG ironed this afternoon his shirt
   'Hasn't Jean ironed his shirt this afternoon?'

The study found that after an initial phase of attention split between the pictures representing the positive and the negative state (an ironed and wrinkled state, respectively), participants shifted their attention to the picture representing the positive state in all positive PQs, high-negation questions in English, and negative questions in French. No preference in look patterns was observed for English low-negation PQs. This suggests that PQs are processed differently from assertions: The positive proposition remains salient throughout, and more salient than the negative one in most cases. Tian et al. (2016, p. 1551) argue that the changing distribution of attention corresponds to pragmatic processing following semantic processing. In positive and high-negation PQs, there is a positive epistemic bias which may be held responsible for the shift of attention to the positive picture. In English low-negation PQs, the positive epistemic bias is optional; French PQs are ambiguous between a low-negation and a high-negation reading. This may counter-balance and, thus, reduce the amount of attention to the positive picture across the experiment.

In our study, we conceptually replicate the gist of Tian and colleagues' experiments, but within a single setup, allowing for direct statistical comparison between assertions and questions.

3.1 Method and stimuli

The experimental arrangement, overviewed in **Table 1**, consisted of four subexperiments, which constituted mutual fillers. Experiments 1 and 2 were designed as parallel in Czech and Russian (with only minor differences in particular lexical realizations).[8] Experiment 1 conceptually replicates the gist of Tian et al. (2016) and the design of Experiment 2 follows Tian et al. (2021), facilitating cross-linguistic comparison. Experiment 3 is an exploration of the effect of bias-inducing particles. Filler stimuli involved small-scale exploratory experiments (not reported

[8] The R scripts, notebooks, and data used for the analysis are available on OSF; see Onoeva et al. (2025).

here). Items from all experiments were combined, amounting to a total of 80 stimuli for each participant. For each experimental list, item order was pseudo-randomized, so that two items from the same experiment could not appear one after another and two conditions from the same experiment could not be presented one after another, even when separated by an item from another experiment. We use the standard Latin Square design, so participants heard each condition with eight items, but only one condition per given item.

Table 1: Overview of reported experiments.

EXPERIMENt	VARIABLES	NR OF ITEMS
E1 Assertions	Polarity: pos × neg	16
E2 Questions	1) Verb position: V1 × V2 2) Polarity: pos × neg	32
E3 Particles in PQs Ru *razve*, Cz *snad*	Polarity: pos × neg	16
Fillers		16
Total		80

The main task was to listen to prerecorded mini-dialogues between a man (M) and a woman (W) and simultaneously look at a visual scene containing four pictures – two target pictures and two distractors – as in **Figure 1**. An English translation of an actual stimulus for **Figure 1** is available in (10). The auditory stimulus went off 1 second after the visual part appeared. After listening to the mini-dialogue, participants selected, by pressing the corresponding button, the

Figure 1: Visual stimulus for (10).

picture they thought constituted the best match to the dialogue's conclusion; for the example in (10), it is the bottom left picture of the dirty boots in **Figure 1**. The general shape for all experimental items was the same, as shown in (11). The auditory stimuli were recorded by four native speakers in the phonetic laboratory at the Faculty of Arts, Charles University.

(10) a. [M:] Did Kate clean the boots in the morning?
b. [W:] I think she did.

(11) a. [M:] Assertion/Question
b. [W:] 'I think that yes.'/'I think that no.' (literally)

Our picture materials were taken from Tian et al. (2021) and only slightly adapted. Contrary to Tian and colleagues' design, our visual scene did not involve a visual representation of the agent of the activity described in M's utterance. We were concerned that looks to the agent would attract gaze and thereby reduce the amount of looks to the targets, which could lead to unnecessary information loss, esp. in the interrogative (VS) orders, where the effect of interrogativity on processing is expected to step in very early on.

The whole dialogue was divided into seven interest periods (IPs): four in the first utterance, a one-second gap, and two in the response (see **Table 2**). The first utterance consisted of the subject (a common male or female name), a verb in the past tense, a temporal or locative adverbial introduced to provide more processing time,[9] and the object, which corresponded to the entity depicted in the targets of the visual scene. The relative positions of subject and verb were varied in the PQ experiments. The response was split into 'I think that' and either 'yes' or 'no'.

3.2 Participants, procedure, and equipment

We tested 52 participants per language and provided them with a compensation of €11 (Berlin) / 350 CZK (Prague). All participants had uncorrected or corrected-to-normal vision. The reliability of the participants was checked via the picture selection task in the positive V1 question condition. A participant was considered reliable if, in 6 or more trials out of 8, they chose the picture corresponding to the polarity of the response particle.

The Russian part of the study included 51 participants (one was excluded as unreliable; 38 females, mean age: 30.5 years). They were Russian native speakers who had recently relocated, recruited via social networks. The experiment took place in Zentrallabor of the Faculty of Linguistics and Literature at Humboldt-Universität zu Berlin. The Czech study covers 48 participants (four excluded; 37 females) who were native Czech speakers and mostly students

[9] It plays a role comparable to Tian et al.'s (2021) Saxon possessive in the object phrase. The reasons for not using possessives have to do with their grammar in Slavic languages, where prenominal possessives do not always sound completely natural.

recruited via an online university psycholinguistics pool. It was carried out in Prague, in the LABELS lab (affiliated with the Faculty of Arts, Charles University, and the Institute of Psychology of the Czech Academy of Sciences).

Upon arrival at the laboratories, participants filled in a short questionnaire and signed a consent form. After they received all instructions and underwent calibration, all participants tried out two test items to make sure they understood the task. A short break was enforced after completing 40 stimuli (i.e., in the middle of the session). The whole procedure took approximately one hour.

Table 2: Interest periods.

M (question/assertion)				**Gap**	**W (response)**	
P1	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**
> Verb(+*li*) [± neg] > Subj	> Subject > Verb [± neg]	Adverb	Object	Gap	'I think that'	> yes > no

The experiment was performed on the EyeLink Portable Duo eye-tracker in Berlin and on the EyeLink 1000 Plus eye-tracker in Prague. Experiment Builder (SR Research Ltd., 2020) and R (R Core Team, 2022) were employed for data collection and statistical analysis. We have obtained full approval from the Ethics Committees at Humboldt-Universität zu Berlin and Charles University for all experiments.

3.3 Dependent variable

Independent variables were different for each experiment, and will be described below. The dependent variable, following Tian et al. (2021), was the natural logarithm of the ratio of the probability of looks to the p picture over the probability of looks to the $\overline{p}$ picture:

$$\ln\left(\frac{P_p}{P_{\overline{p}}}\right)$$

This metric is commonly used in visual world analyses that compare fixations on two areas of interest (see Tian et al., 2016, and references therein). To compute the probabilities of looks, total fixation durations to the p and $\overline{p}$ pictures during each interest period were divided by the duration of the interest period. Time boundaries of individual interest periods were obtained by annotation of the auditory stimuli, using Praat software (Boersma & Weenink, 2009). Czech stimuli were annotated manually. For Russian stimuli, WebMAUS Basic service (Kisler et al., 2017) was applied to generate automatic annotations that were then manually corrected. To prevent division by zero in cases where there were no fixations on the $\overline{p}$ picture, 0.5 ms was

added to all values (total fixation durations and interest period durations). The natural logarithm transformation aims at making the distribution of the data less skewed and, therefore, more suitable for the statistical method we applied (Baayen, 2008, p. 31).

The dependent variable (from now on referred to as LOG-RATIO) is centered around 0. Since the natural logarithm of 1 is 0, a zero LOG-RATIO indicates an equal probability of looks to the p and $\overline{p}$ pictures. A positive LOG-RATIO corresponds to a higher probability of looks to the p picture than to the $\overline{p}$ picture. The higher the positive value, the larger the preference. A negative value indicates a preference for the negative alternative, and lower values represent a larger difference.

This derivation of the dependent variable implies that only data for fixations on the pictures p and $\overline{p}$ were used for the analysis. Although fixations on the distractors were also measured, they are not of interest to our research hypotheses and will not be reported.

Picture selection was an auxiliary dependent variable.

3.4 Data manipulation and statistical analysis

Data preprocessing was performed in R[10] using the eyelinkReader (Pastukhov, 2024), vctrs (Wickham et al., 2023), plyr (Wickham, 2011), here (Müller, 2020), and tidyverse (Wickham et al., 2019) packages. During this step, fixations that lasted less than 80 ms were excluded, because short fixations are commonly attributed to false saccade planning (Bosker et al., 2014; Tian et al., 2021). All interest periods were offset by 200 ms to account for the time needed to plan an eye movement (Hallett, 1986; Tian et al., 2021). Trials in which no fixations on p or $\overline{p}$ pictures were registered during the whole auditory stimulus (a question or an assertion, the gap, and the response) were excluded from the analysis.

The rationale of our statistical analysis largely follows Tian et al. (2016, 2021). We fitted linear mixed-effects (LME) models using lme4 (Bates et al., 2015), with statistical inference performed using lmerTest (Kuznetsova et al., 2017). The analysis consisted of two parts. First, LOG-RATIO was compared to 0 for each interest period in each separate condition. The dependent variable being significantly different from 0 would imply a significant difference between the probabilities of looks to the p and $\overline{p}$ pictures for that condition and interest period. This would allow for an indirect comparison between conditions, based on when a significant preference for one of the alternatives emerged. For this analysis, we followed the guidelines in Ito & Knoeferle (2022). For each interest period, LME models were fitted to data from each condition. Fixed effects were, therefore, excluded from these models, leaving only the intercept and the random

[10] Parts of code for data preprocessing and analysis were generated with the help of ChatGPT (OpenAI, 2025). Those parts were then manually tested and corrected, if needed.

intercepts, as shown in the formula below.[11] Since each condition was analyzed separately, fitting random slopes would not be justified. If a model did not converge or returned a singularity warning, one of the random intercepts was removed.

$$\text{log\_ratio\_transf} \sim (1 \mid \text{item}) + (1 \mid \text{participant})$$

Second, the impact of fixed predictors on LOG-RATIO was tested. LME models were fitted to each interest period. All predictors for fixed effects were categorical and binary and were effect-coded (exact values for each analysis are given in the corresponding subexperiment sections). When fitting random effects of item and participant, we started from the maximal model that was supported by the design (Barr et al., 2013). To prevent conversion and singularity issues, correlation parameters were excluded from all the models. This is justified by the fact that the research hypotheses only concerned fixed effects and there were no specific expectations about variance between participants or items (Bates et al., 2018). Furthermore, the control parameter was set to *control = lmerControl(calc.derivs = FALSE)* (Brown, 2021). If a model still did not converge without a singularity warning, the number of random slopes and, if needed, intercepts was gradually reduced up until successful convergence. Plots were generated based on model predictions obtained with help of the ggeffects package (Lüdecke, 2018).[12]

4. Experiment 1: Assertions

4.1 Design and hypotheses

The goal for this experiment was to investigate the processing of positive and negative assertions, so POLARITY was the only manipulated variable. An example of a Russian item (its linguistic part) is provided in (12). Czech items were fully analogous.

(12) *Russian*

a. Anja otkryla v kafe knigu.
Anja opened in café book
'Anja opened a book in a café.' (positive)

b. Anja ne otkryla v kafe knigu.
Anja NEG opened in café book
'Anja didn't open a book in a café.' (negative)

We hypothesized that the gaze would be affected by the polarity of the assertion: positive assertions would instigate fixations on the picture representing the positive alternative (also

[11] The name of the dependent variable, log_ratio_transf, indicates that it has been transformed as described in 3.3. The same naming principle applies to all other statistical formulas in this article.

[12] For other visualizations, we used DescTools (Signorell, 2024), ggplot2 (Wickham, 2016), ggalt (Rudis et al., 2017), ggh4x (van den Brand, 2024).

called *positive fixations/picture*, or simply p, below), in this case, an open book, and negative assertions would lead to fixations on the picture representing the negative alternative (*negative fixations/picture*, or simply $\overline{p}$), i.e., a closed book.

4.2 Descriptive findings

Figure 2 shows how the proportion of fixations on p and $\overline{p}$ evolved over time during the dialogue (the beginning of all interest periods was offset by 200 ms; see 3.4). To generate this and the corresponding visualizations in the following subsections, each trial was split into 1000 subsequent fragments of equal duration. For each fragment, it was calculated if there was a fixation on p or on $\overline{p}$. The proportion of fixations, thus, corresponds to the proportion of experimental trials in which the positive or negative picture was fixated on in each fragment.

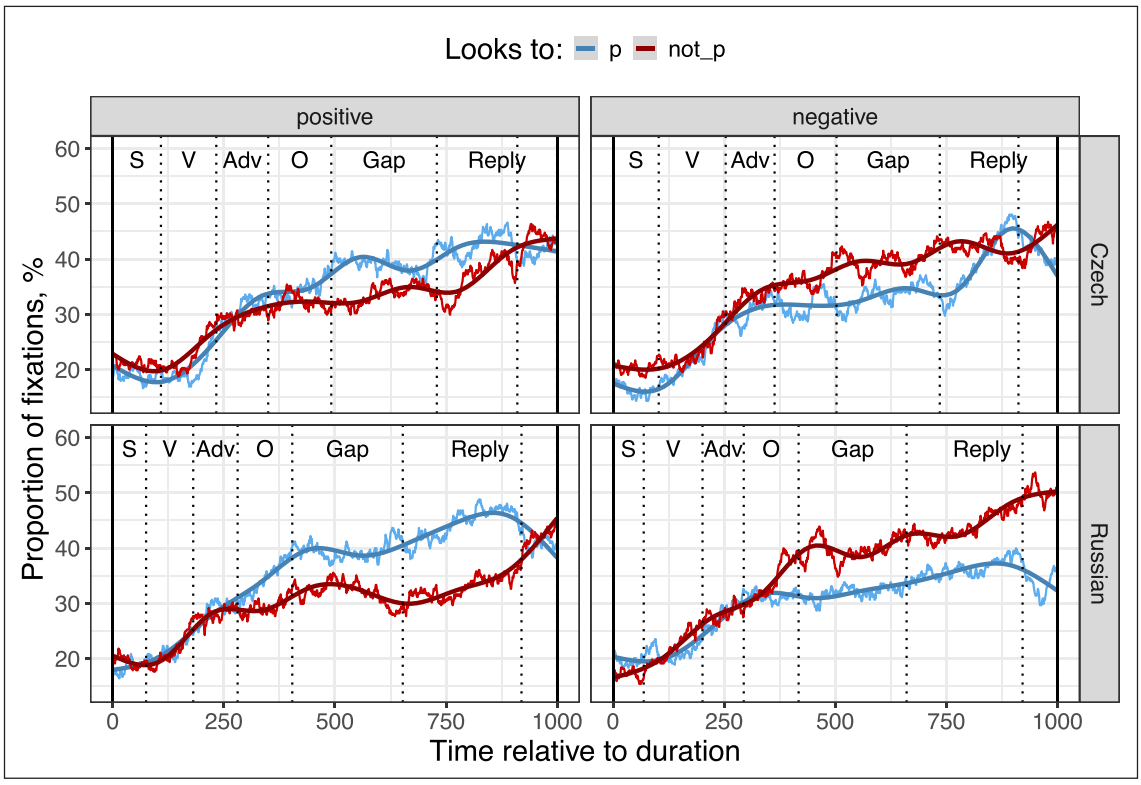

Figure 2: Proportions of looks to p and $\overline{p}$ in Czech and Russian assertions, normalized for sound duration across conditions (IP boundaries also normalized) with geom_smooth.

The proportions of looks both to p and $\overline{p}$ grow as information about the target object is revealed to, and processed by, the participants. However, this increase is not equal for the positive alternative and the negative alternative, and is conditioned by the polarity of the assertion. In both Czech and Russian, positive assertions are characterized by a higher proportion of looks to p, and negative assertions, to $\overline{p}$. This tendency seems to establish itself in late periods of the

assertion (adverbial in Czech negative assertions, object in the other cases), and to become most apparent in the gap or the first part of the answer. The next subsections report on the statistical analysis that estimates the significance of those trends.

4.3 Models

In this experiment, the sole predictor was POLARITY, a categorical variable with two levels: positive (coded as 1) and negative (coded as –1).

As a starting point, the maximal reduced LME models were fitted to each interest period.[13] Aside from the fixed predictor, each model included random intercepts and random slopes for items and participants. If needed, the random effects structure of individual models was gradually simplified to resolve convergence and singularity issues (see 3.4).

4.4 Results

Figure 3 shows mean LOG-RATIO (on the *y*-axis) over the interest periods (on the *x*-axis) across languages. On each facet, the black horizontal line indicates zero LOG-RATIO, which would correspond to equal mean proportions of fixations on p and $\overline{p}$. Green dots mark combinations of interest period and condition in which LOG-RATIO was significantly different from 0, according to the LME analysis of the intercept described in 3.4.

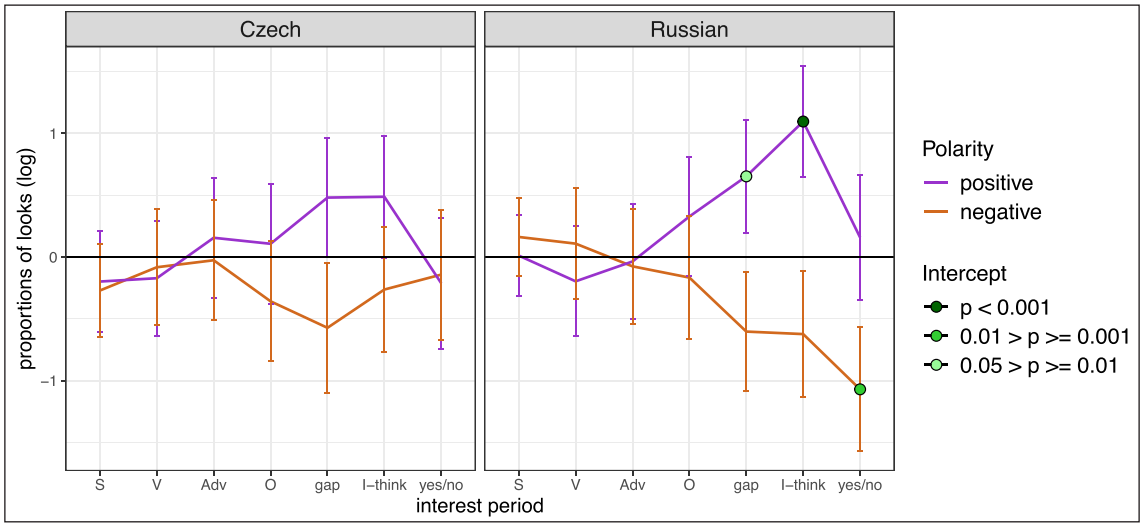

Figure 3: Average progression of LOG-RATIO of proportions of fixations on p over $\overline{p}$ by interest period in Czech and Russian assertions. Error bars indicate 95% confidence intervals. Green dots show where LOG-RATIO was significantly different from 0.

[13] Maximal model formula: log_ratio_transf ~ polarity_contr + (1 + polarity_contr || item) + (1 + polarity_contr || participant). The addition of "contr" to POLARITY specifies that this variable has been contrast-coded manually, to enable the exclusion of the correlation terms (see 3.4). The latter is encoded by the double bar signs. The same formalism holds for the rest of the formulas in the article.

It is evident that assertions tend to trigger a preference for the pictures which correspond to their polarity. In the positive condition (represented by the violet line), mean LOG-RATIO is above 0 from the adverbial to the first part of the response in Czech, and from the object on in Russian. Negative assertions (orange line) show the contrary tendency. From the adverbial on, mean LOG-RATIO in the negative condition is below 0 in both languages, indicating longer average fixation times on the negative picture. However, the statistical comparison of the dependent variable to 0 rendered only few significant differences, all after the target utterance in Russian.

Although the preference for the picture with the corresponding polarity may be characterized as rather moderate, the difference between positive and negative assertions is robust enough to be supported by the statistical models. In both languages, a significant effect of POLARITY was found in the gap (Russian: $t = 3.381, p < 0.01$; Czech: $t = 2.873, p < 0.01$) and in the first part of the answer (Russian: $t = 3.692, p < 0.001$; Czech: $t = 2.059, p = 0.04$). **Figure 4** shows predicted values of the dependent variables, based on the LME models fitted for these interest periods.

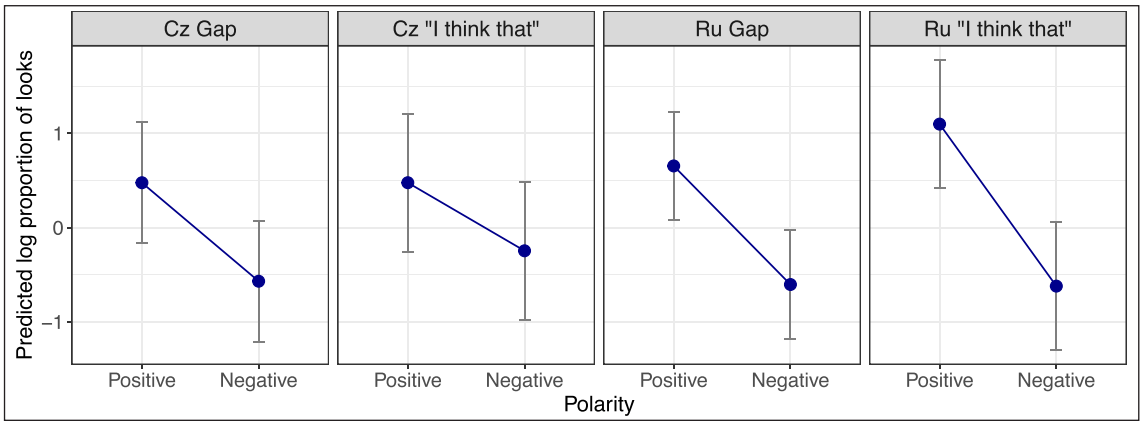

Figure 4: LOG-RATIO predicted by LME models in selected interest periods of Czech and Russian assertions.

It is worth mentioning that Russian and Czech exhibit different patterns in the second part of the response ('yes' or 'no'). In Czech, mean LOG-RATIO in this interest period is relatively close to 0 in both the positive and the negative conditions. This indicates roughly balanced proportions of looks to p and $\overline{p}$, which is in accordance with the equal amount of stimuli containing 'yes' and 'no' answer particles in each experimental list. A similar tendency can be observed for Russian positive assertions. In contrast, the negative condition in Russian shows a clear preference for the negative picture during the answer particle period, LOG-RATIO being estimated as significantly lower than 0 by the model. A more complete picture can be achieved if we observe the outcome for 'yes' and 'no' response particles separately (see **Figure 5**). In Czech, 'yes' seems to be associated with the p picture and 'no' with the $\overline{p}$ picture, after both positive and negative assertions. In Russian, however, both response particles seem to trigger fixations on p after a negative assertion. These fixation patterns are in accordance with the results of the picture selection task (see **Figure 6**).

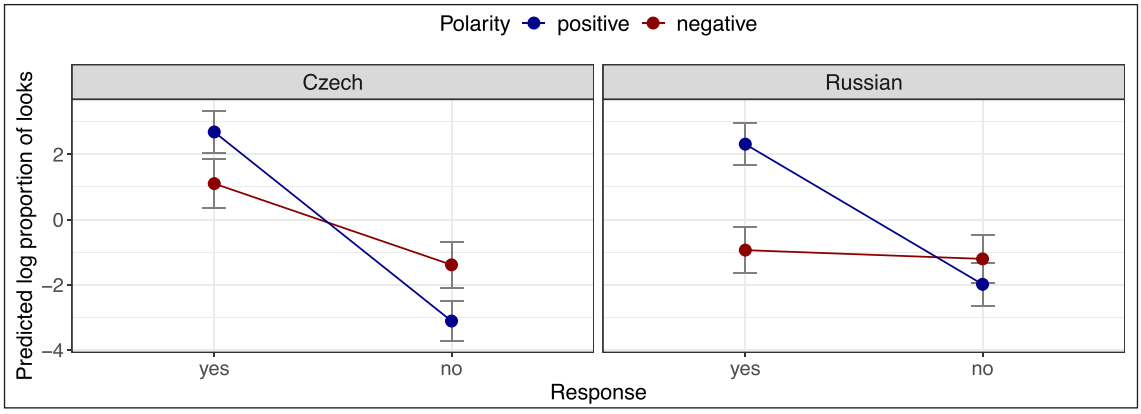

Figure 5: LOG-RATIO in the second part of the response ('yes'/'no') in Czech and Russian assertions.

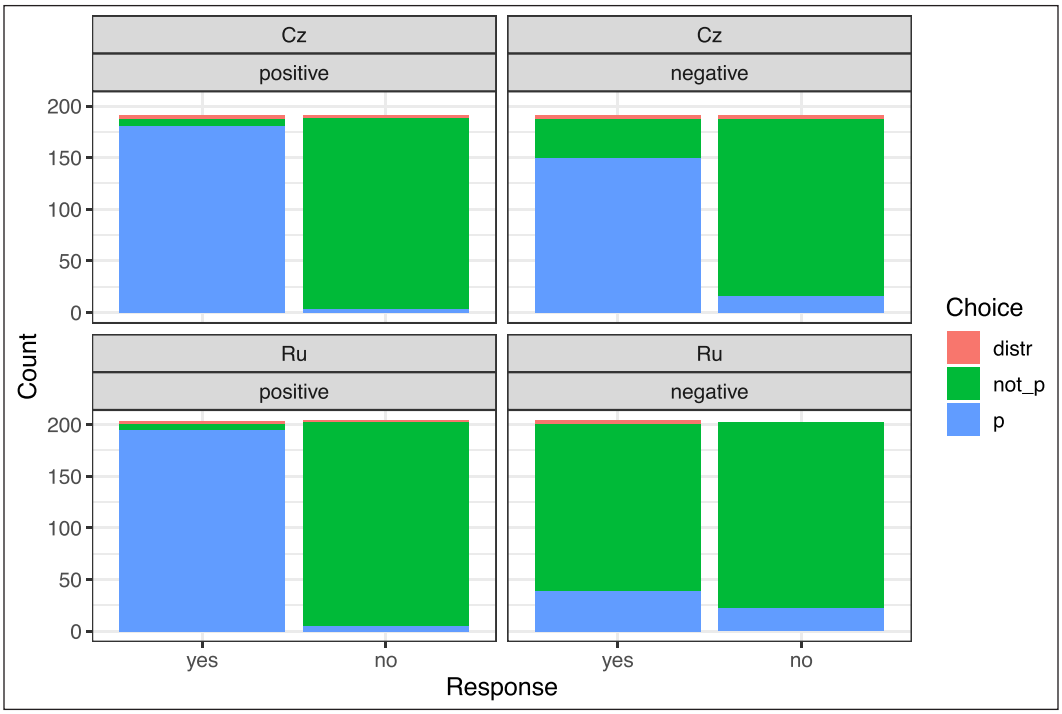

Figure 6: Picture choices in assertions.

4.5 Discussion

In both languages, positive assertions instigated a higher proportion of looks to the p picture over looks to the $\overline{p}$ picture than negative assertions. The difference became significant in the gap period, i.e., right after the assertion had been fully heard, and continued through the first part of the response ('I think that'). Comparisons of the LOG-RATIO measure to the 0 line showed

that looks to the picture corresponding to the polarity of the assertion diverged significantly from a balanced distribution in this same period for Russian positive assertions, and for Russian negative assertions upon hearing the answer particle.

This result is in line with Tian et al. (2016), who found that simple English positive vs. negative assertions significantly favored looks to the picture corresponding to their polarity from about 900 ms after the end of the verb period on.[14]

The result is consistent with the idea that the semantic interpretation of the utterance is built up incrementally during processing, with a time lag between the cue and the zooming in on the corresponding visual stimulus. Even though the visual stimuli represented only the objects to which the action had (not) been applied, not the agents, verb meaning itself already allowed for a prediction as to which object might fit. Thus, the somewhat delayed fixation on the expected visual stimulus may come as a surprise. As argued in some detail in Tian et al. (2016), there is good reason to assume that the positive proposition remains active in processing for a rather long time in both positive and negative utterances.

In addition to testing our main hypotheses, the findings obtained also provide some insights into the processing of the response particles. Previous linguistic analyses have shown that the status of the response particle in reactions to a negated assertion differs between Czech and Russian. In Czech, the negative particle *ne* 'no' encodes the negative polarity of the response and *ano* 'yes', agreement with the positive part of the negative prejacent (Hrdinková & Šimík, 2025). In Russian, both the positive *da* 'yes' and the negative *net* 'no' can mark agreement with a negative prejacent, while only the combination of a negative response particle and a following positive verb can indicate rejection (Esipova, 2021, ex. (5)).

Picture choices confirm this analysis (see **Figure 6**): In Russian, both the positive and the negative response particles were associated with $\overline{p}$. This reasoning explains the difference in fixation patterns we observe during the second part of the response to the negative assertions. While in Czech, 'yes' can agree with the positive part of the prejacent and instigate positive looks, in Russian, it conveyed agreement with the negative statement, triggering negative fixations.

5. Experiment 2: Questions

5.1 Design and hypotheses

We used a 2 × 2 fully-crossed factorial design for the PQs experiment. The variable VERB POSITION represents two major interrogative strategies in Czech and Russian – verb-initial order (V1) and verb-second order (V2) combined with question-specific intonation. The second variable

[14] Tian et al. (2016) did not fit a linear mixed-effects model, but conducted an ANOVA over the LOG-RATIO measure.

is POLARITY; PQs were either positive or negative. Both variables were manipulated within 32 items and within subjects. The linguistic stimuli are illustrated in (13) by one Czech item in all conditions. Russian stimuli were fully parallel and differed only in condition V1, in that the verb is followed by the particle *li*.

(13) *Czech*

a. Umyla Dita včera holínky?
cleaned Dita yesterday rubber.boots
'Did Dita clean the rubber boots yesterday?' (positive V1)

b. Dita umyla včera holínky?
Dita cleaned yesterday rubber.boots
'Dita cleaned the rubber boots yesterday?' (positive V2)

c. Neumyla Dita včera holínky?
NEG.cleaned Dita yesterday rubber.boots
'Didn't Dita clean the rubber boots yesterday?' (negative V1)

d. Dita neumyla včera holínky?
Dita NEG.cleaned yesterday rubber.boots
'Dita didn't clean the rubber boots yesterday?' (negative V2)

Factoring in the results of Tian et al. (2021), we hypothesized that the processing of PQs would be accompanied by looks to the *p* picture, independently of their polarity, reflecting the interpretation of positive bias.

We further assume that this tendency will be modulated by our experimental and cross-linguistic manipulations. First, because V1 questions in Russian are accompanied by the *li* particle (attached to the verb), a salient question marker morphologically related to the disjunction *ili* 'or' (see Section 1), we expect the looks to be more evenly distributed between p and $\overline{p}$ than in the V2 condition, where question meaning is conveyed by means of intonation. This expectation also follows from the assumption that V1 questions in Russian are unbiased (Šimík, to appear; Korotkova, submitted). The effect of the *li* particle can be set against the background of Czech, where we expect to observe a "pure" effect of interrogative word order; the results can, thus, reasonably be expected to be similar to what Tian et al. (2021) observed for English, i.e., dominant looks to p. V2 questions in Czech are less neutral than in Russian, in the sense that they are (evidentially) biased (Staňková & Šimík, 2025), which, in turn, can be expected to lead to increased fixations on the proffered alternative.

The effect of polarity is expected to differ depending on word order. Negation in V1 questions in both Russian and Czech arguably maps to positive epistemic bias (Šimík, to appear). The expectations for V2 questions are not straightforward. On the one hand, negation could instigate

more looks to $\overline{p}$ (esp. in Czech, where V2 is accompanied by negative evidential bias); on the other, negation triggers positive epistemic bias even in V2 questions (possibly leading to a conflict of both biases), which might neutralize the effect of negation.

The word order manipulation in both languages is further expected to be reflected in the different timing of where question marking affects the fixations: the preference for *p* developing earlier in V1 than in V2. More specifically, in Russian V1, it becomes evident that the processed utterance is a question in the initial (verb) period; the effect is reinforced by the *li* particle. However, *li* as a marker of unbiased PQs (see above) may also lead to the opposite effect on timing, causing a more balanced distribution of fixations in the early periods. The situation in Czech is different, because Czech (being pro-drop) allows for non-interrogative V1 utterances; but as soon as the postverbal overt subject is encountered, the likelihood of question interpretation reaches near-certainty (see Šimík, to appear). In Russian V2 questions, interpretation is conveyed by the dedicated pitch accent on the verb; question interpretation should, therefore, step in soon after that. In Czech V2, it is the utterance-final rising tone that is considered to be a reliable question cue. Question interpretation is, thus, expected to step in fairly late.

5.2 Descriptive findings

Figures 7 and **8** display the temporal progression of fixation proportions on *p* and $\overline{p}$ for Czech and Russian PQs, respectively. Almost all conditions seem to exhibit a clear preference for the positive alternative which manifests itself starting from the second to fourth period of the question. Only in Russian negative V2 questions do the looks tend to be distributed in a more balanced way.

5.3 Models

The statistical analysis included two categorical predictors: POLARITY and VERB POSITION. Positive POLARITY and V1 VERB POSITION were coded as 1, negative POLARITY and V2 VERB POSITION as –1. The random effects structure includes random intercepts for item and participant and random slopes for polarity, verb position, and their interaction both for item and participant.[15]

To explore potential differences between Russian and Czech in PQs processing, additional models were fitted to selected interest periods of V1 (verb, subject, and adverbial) and V2 (verb, adverbial, object, and gap) questions separately. Those models included POLARITY and LANGUAGE as predictors, with positive POLARITY and Russian LANGUAGE coded as 1, and negative POLARITY and Czech LANGUAGE as –1. The maximal models were fitted to the data

[15] Model formula: log_ratio_transf ~ polarity_contr*v_position_contr + (1 + polarity_contr*v_position_contr || item) +(1 + polarity_contr*v_position_contr || participant).

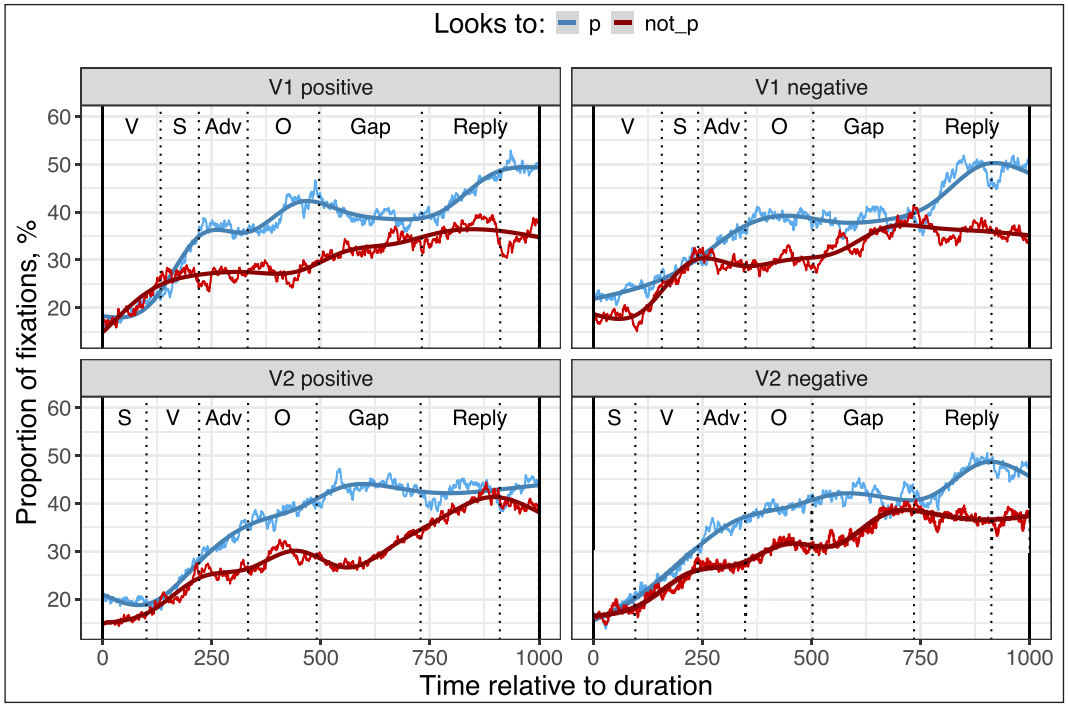

Figure 7: Proportions of looks to p and $\overline{p}$ in Czech questions, normalized for sound duration across conditions (IP boundaries also normalized) with geom_smooth.

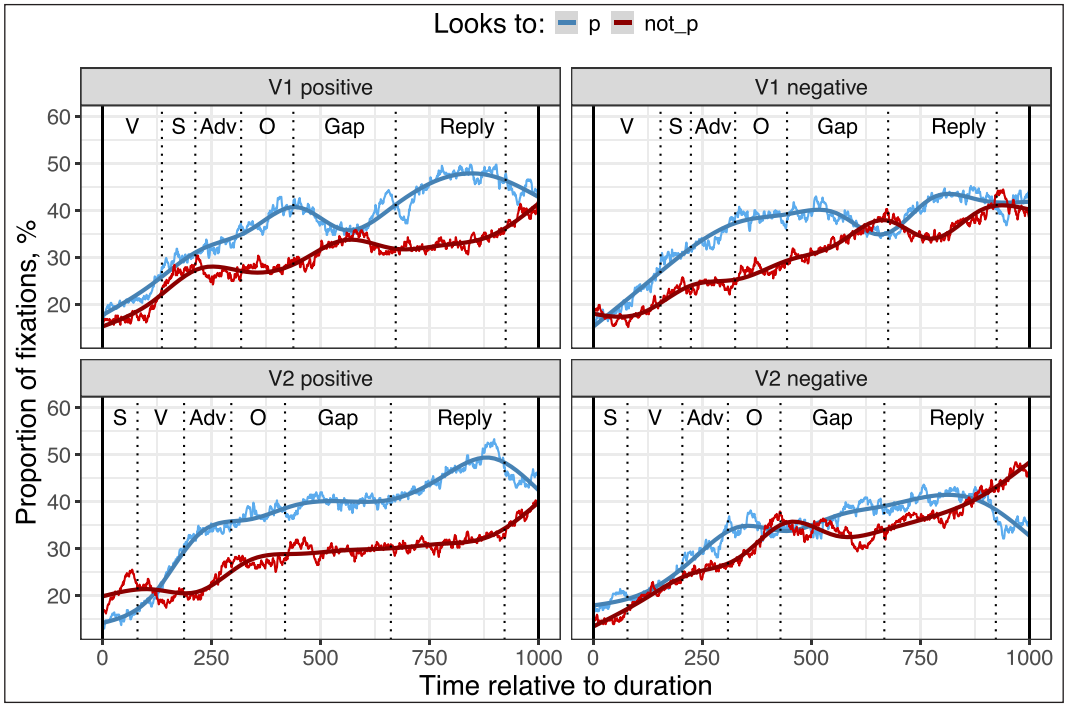

Figure 8: Proportions of looks to p and $\overline{p}$ in Russian questions, normalized for sound duration across conditions (IP boundaries also normalized) with geom_smooth.

subsets for V1 and V2 questions that included data from both languages.[16] As each participant was exposed to stimuli only from one language, the random slope of language by participant was not included. Concerning items, they were parallel in Russian and Czech. Therefore, it was assumed that effects of both predictors as well as their interaction could vary by item. In both analyses, random effects structure was gradually simplified, as described in 3.4, if confronted with conversion and singularity issues.

5.4 Results

Figure 9 displays the progression of mean LOG-RATIO values for positive and negative PQs with verb-initial (V1) and verb-medial (V2) word orders in Czech and Russian. Fixations on the *p* picture dominate across the board, most of the values being above 0 (except for the first and the last interest periods in some conditions). This overall trend seems to be consistent across

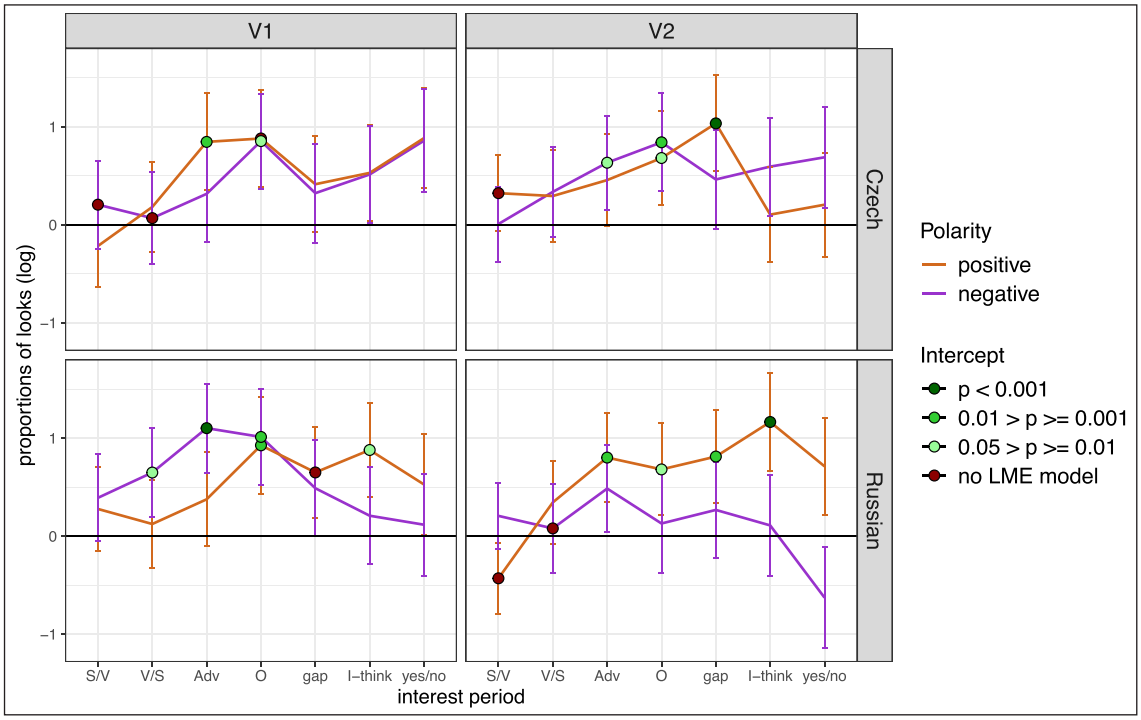

Figure 9: Average progression of LOG-RATIO of proportions of fixations on p over $\overline{p}$ by interest period in Czech and Russian questions. Error bars indicate 95% confidence intervals. Green dots show where LOG-RATIO was significantly different from 0. Red dots show where no LME model could be fitted.

[16] Model formula: log_ratio_transf~polarity_contr∗language_contr + (1 + polarity_contr∗language_contr||item) + (1 + polarity_contr||participant).

all conditions, apart from negative Russian V2 questions. However, an indirect comparison of statistical intercept-based models (see 3.4) reveals differences in temporal emergence of a significant preference of the p picture over the $\overline{p}$ picture.[17]

5.4.1 Czech

In positive V1 questions, the preference for the p picture over the $\overline{p}$ picture is statistically significant in the adverbial period. In negative V1 questions, on the other hand, the effect emerges only in the object. Starting from the object for positive V1 questions and the gap for negative V1 questions, no significant preference was found. Czech V2 questions constitute a somewhat reversed pattern concerning POLARITY. In this case, it is the negative V2 condition in which the significant difference is reached already in the adverbial (and continues in the object), whereas in positive V2 questions, it is present only in the object and the gap. In spite of these temporal differences, the statistical analyses which tested the influence of the fixed predictors did not reveal any significant effects of VERB POSITION, POLARITY, or an interaction thereof for any interest period in the Czech data.

5.4.2 Russian

In the positive V1 condition, a significant preference for the p picture was found in the object region and in the first part of the response; no reliable estimation was obtained for the gap period. Negative V1 questions exhibit the earliest significant preference for p throughout the experiment, namely starting from the subject and continuing until the object. In positive V2 questions, fixations on p significantly prevail from the adverbial to the first part of the response on. The negative V2 condition in Russian is the only one for which no significant preference for the p picture was found in any interest period.

Based on the statistical models that included VERB POSITION and POLARITY as predictors, significant effects were obtained in several interest periods for Russian. These are illustrated in **Figure 10**, which shows predicted values of the dependent variable across conditions and relevant interest periods.[18] In the adverbial period, an interaction of POLARITY and VERB POSITION

[17] It must be noted that for certain combinations of interest period and condition, no LME model could be fitted, due to convergence issues. Such cases are marked by red dots in the plot. It would be unreasonable to compare a random effects model with a linear regression model which only includes fixed predictors. Therefore, we remain agnostic as to whether the dependent variable was significantly different from 0 in such cases.

[18] The analysis also detected a significant main effect of VERB POSITION in the first interest period in Russian, which was the verb in the V1 and the subject in the V2 conditions ($t = 2.2$, $p = 0.032$). The LOG-RATIO was estimated to be higher in V1 than in the V2 questions. Since no information on the target object and, therefore, on the pictures corresponding to p and $\overline{p}$ was presented during the subject period in the V2 condition, one might suggest that the locus of fixations was due to chance in V2 questions, whereas the information from the verb in V1 questions could trigger some looks to p.

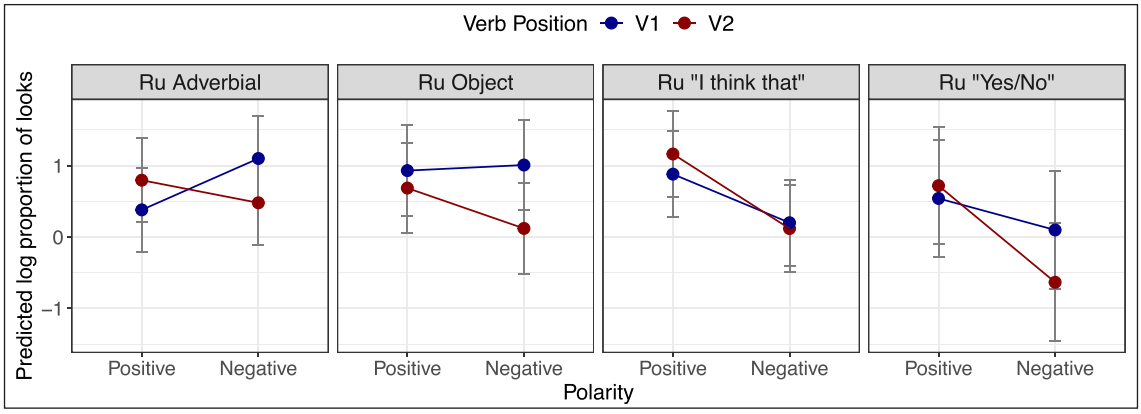

Figure 10: LOG-RATIO predicted by LME models in selected interest periods of Russian questions.

($t = -2.173$, $p = 0.035$) was found. In the V2 condition, positive questions were estimated to have more fixations on p during this period than negative questions. For V1 questions, the model predictions are reversed, with negative questions showing a stronger preference for p. In the object period, the model identified a main effect of VERB POSITION ($t = 2.074$, $p = 0.0497$), V1 questions being characterized by more looks to p than V2. Auxiliary LME models fitted to the data subsets divided by POLARITY reveal that the main effect is due to the simple effect of VERB POSITION in the negative condition ($t = 2.320$, $p = 0.029$); the corresponding simple effect in the positive condition is absent ($t = 0.573$, $p = 0.57$). Finally, a main effect of POLARITY was discovered in both parts of the response ($t = 2.997$, $p < 0.01$ for 'I think that' and $t = 3.547$, $p < 0.001$ for 'yes/no').

5.4.3 Language comparison

We found an interaction of POLARITY and LANGUAGE in the adverbial period of V1 questions ($t = -2.184$, $p = 0.037$). The corresponding model predictions are shown in **Figure 11**. The statistical inference indicates different directions of the POLARITY effect in the two languages. While in Czech, positive questions are characterized by more looks to p than negative questions, in Russian it is the negative ones that instigate more fixations on p. No significant effects were found in the periods of V2 questions that were investigated.

5.5 Discussion

The results of Experiment 2 largely corroborate previous findings of Tian et al. (2021) for English and French: PQs generally trigger looks to the positive picture, even if they have negative polarity. We see this in most of our experimental conditions (see **Figure 9**): in Czech, the positive picture received more fixations than the negative one at the latest in the adverbial or the object period, independently of polarity and verb position; in Russian, the same holds, with the exception of

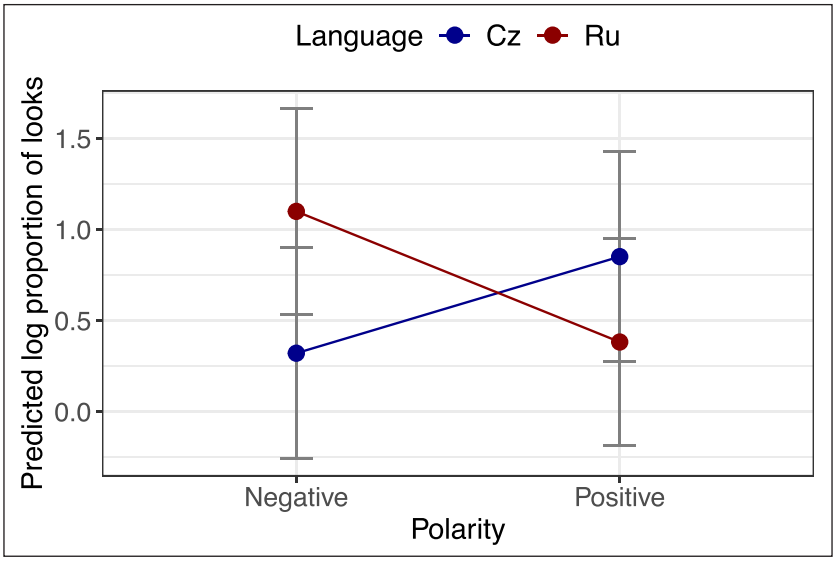

Figure 11: LOG-RATIO predicted by the LME model for the adverbial period of Czech and Russian V1 questions.

V2 negative questions, where neither fixations on the p picture nor on the $\overline{p}$ picture dominate significantly. We interpret this result as evidence for the salience of positive epistemic bias in PQs.[19] That positive epistemic bias can be present in both positive and negative PQs, although in somewhat different flavors, is illustrated by the following pair of questions, which are formulated in English for simplicity; an analogous generalization applies to Russian and Czech (see also Section 2 for discussion).[20]

(14) a. Did Mary clean her shoes?
⇝ The speaker expects her to have done so and checks if the expectation is met [positive epistemic bias, based, e.g., on prior agreement].

b. Didn't Mary clean her shoes?
⇝ The speaker has until recently believed that she did [positive epistemic bias, based on prior belief]; possibly (not necessarily; see Goodhue, 2022), this belief is in conflict with evidence (direct or reported) that she didn't [negative evidential bias].

If our reasoning is right, then epistemic bias was more salient for our participants than evidential bias. One reason why this might be so could be that while epistemic bias is private to the speaker

[19] An anonymous reviewer wonders why the participants, who merely observe a conversation (not being directly involved in it), should be attending to speaker-anchored bias. We assume that speaker-related (epistemic) bias is an implication conveyed by the PQ (see Section 2). If a participant processes all the implications conveyed by the stimulus (regardless of who the addressee is), it is expected that s/he also processes the bias.

[20] This is not to say that PQs exhibit exactly the same bias properties in English as they do in Russian or Czech. There are important differences pertaining to the modal force and flavor of the biases (see Šimík to appear), but none with clear consequences for our experimental results.

and is being conveyed conventionally – possibly in the form of a conventional implicature and, thus, potentially as a new piece of information – evidential bias is based on the common ground of both discourse participants and, thus, has the status of a presupposition.[21] While our design did not particularly favor the salience of one type of bias over another, it is likely that evidential bias relies on explicit contextual or situational cues to become salient. The absence of such cues in our design might have led to the increased activation of epistemic bias, which arguably does not require any contextual triggering.

We further hypothesized that Russian V1 questions – due to the presence of *li* – would align with the expectations drawn from Hamblin's (1973) semantics. The results are consistent with this hypothesis, at least in the early periods in the positive V1 condition, where a significantly increased fixation on the positive picture (over the negative picture) steps in rather late, namely, in the object period. This is later than in Russian positive V2 questions as well as in Czech V1 questions, where the positive picture is fixated on already in the adverbial period.[22] This is consistent with the expectation that *li* forces the participants to consider both alternatives for a longer time. The competing hypothesis that the preference for the positive picture, apparently characteristic of PQ processing, would show itself earlier in the Russian V1 condition could, therefore, not be confirmed. A similar delay is not attested in Russian negative V1 questions (which also involve *li*), where positive fixations step in very early – already in the subject period (persisting until the object period). This can be ascribed to a salient positive epistemic bias triggered by negation, which – by hypothesis – overrides the assumed effect of *li*. An anonymous reviewer rightly wonders why the effect of positive epistemic bias should be more pronounced in negative V1 questions than in their positive counterparts. We hypothesize that this is due to the different flavor of the epistemic bias in positive and negative questions: while the former induce what we called a "default" bias, corresponding to the speaker's expectations or hopes, the latter induce a stronger type of bias, based on the speaker's prior belief that the positive prejacent was true. Czech V1 questions exhibit a different pattern. Both positive and negative questions exhibit relatively early positive fixations (considering that the utterance is reliably disambiguated as a question only in the subject period), slightly earlier in positive questions (adverbial) than in negative questions (object). We do not know why the positive bias steps in slightly later in negative questions than in positive ones.

If we factor in the assumption that PQ processing corresponds to positive fixations (and set aside the interfering effect of *li* in Russian positive V1 questions), we see that the results are only partly consistent with our expectations about the timing of question interpretation. Russian

[21] For an account of both types of biases consistent with this view, see Korotkova (2023).

[22] An auxiliary statistical model applied to the subset of positive V1 questions has not yielded a significant simple effect of LANGUAGE in the adverbial period in V1 questions; we only see a related interaction between POLARITY and LANGUAGE; see **Figure 11**.

negative V1 questions align well with the expectation: positive fixations step in immediately after processing the negative verb + *li*. Czech positive V1 questions are recognized as questions (i.e., instigate positive picture fixations) as soon as the VS combination has been uttered, as expected. This process appears to be somewhat delayed in negative V1 questions, for unclear reasons. Russian positive V2 questions also follow the expected pattern: positive fixations occur right after the pitch-accented verb has been processed. In Czech V2 questions, question interpretation was expected to become apparent only on the object, which realizes the final rising tone – a distinctive feature of Czech PQs. It is, therefore, unexpected that positive fixations occur already on the object in positive questions and on the adverbial in negative questions (in the latter case, even earlier than in V1 questions). We hypothesize that the intonation of declarative (V2) PQs in Czech is more complex than previously assumed (Romportl, 1951; Veroňková, 2002). In particular, it is likely that the question-related intonation contour is not confined just to the last/nuclear prosodic phrase of the utterance (here, the object), but spreads over a larger portion of the utterance, so that language users recognize an utterance as a question relatively early (see Baltazani et al., 2015, for evidence to this effect from Greek). Verifying this post-hoc hypothesis would require a proper acoustic investigation of Czech PQs (in comparison to assertions) – not just their final prosodic phrase (as in Veroňková, 2002), but the whole utterance.

It may seem implausible to claim a positive epistemic bias also for Czech positive V2 questions, since they superficially resemble English rising declaratives, which (in the positive variant) typically involve a positive evidential bias and a negative speaker expectation (Gunlogson, 2002).[23] However, Czech V2 questions differ from English rising declaratives in an important respect: They can contain not only inner, but also outer, negation, compatible with positive polarity items (Staňková, 2023). Moreover, Malamud & Stephenson (2015) argue that even English rising declaratives involve tentative commitments put up by the speaker and included in his expectations about further discourse (projected). Thus, epistemic biases of the default type envisaged here may well be at stake even in that subtype of Czech V2 questions which is closer to rising declaratives.

We now return to the negative V2 condition in Russian, which was the only one in which preference for fixation on the p picture (or $\overline{p}$ picture, as a matter of fact) did not reach significance. The relatively balanced fixation profile in this condition might suggest a Hamblin-style interpretation of the question processed. In light of all our other results and the results of Tian et al. (2021), we consider this interpretation highly unlikely. One possibility is that the observed pattern is brought about by the expected meaning of the responses.[24] As is clear from

[23] We are grateful to an anonymous reviewer for insisting on this point. We agree that this is puzzling and cannot be fully resolved here, since we did not embed questions in biased contexts, as in, e.g., Domaneschi et al. (2017).

[24] Another possibility could be that some of the negative PQs were actually mistaken by hearers as negative assertions, thus counter-balancing the amount of fixations on the p picture. Meyer & Mleinek (2006) found evidence that

Figure 12, most negative V2 questions yielded the selection of the negative picture. Therefore, the participant's expectation of choosing the $\overline{p}$ picture, even before the response polarity was revealed, could have affected the fixations.

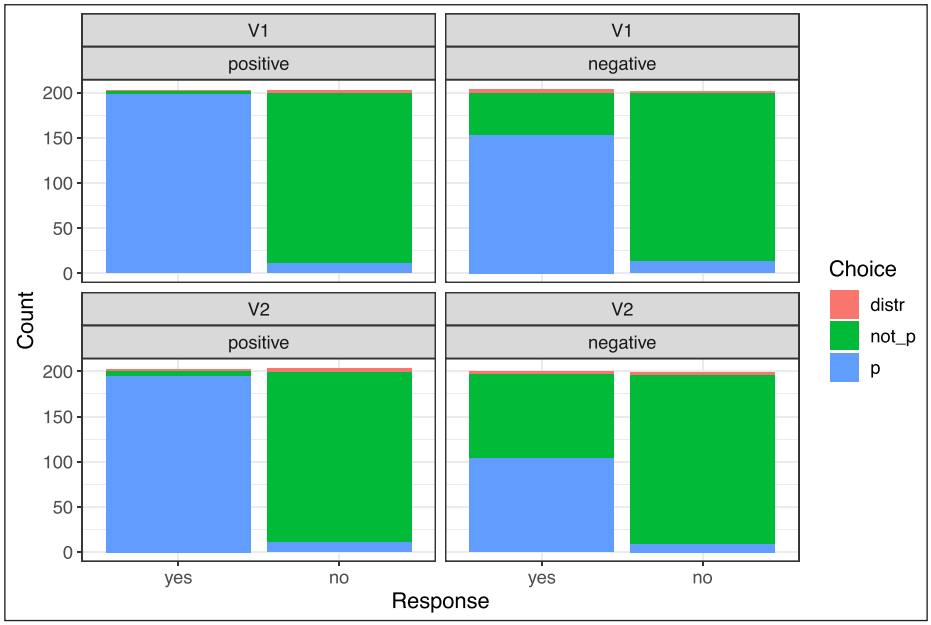

Figure 12: Picture choices in Russian questions.

The close-to-equal proportions of picture choices after the positive response can be connected to the interpretation of negation. Russian negative V2 PQs are generally ambiguous between a low-negation and a high-negation reading. According to González-Fuente et al. (2015) and Esipova (2021), the Russian response particle *da* 'yes' as a reaction to a negated proposition is mainly understood as agreeing with it, not as conveying the positive alternative.[25] This is consistent with low negation which is interpreted at the propositional level. In this case, *da* would confirm the negative proposition, leading to the selection of the negative picture. High negation, on the other hand, whether analyzed as operating at the level of speech acts (Krifka, 2015; Romero & Han, 2004, a.m.o.) or as a separate semantic operator (Šimík, to appear, a.m.o.; Repp & Geist, 2025), establishes no negative propositional antecedent (Krifka, 2013). In this case, *da* conveys the positive alternative (Geist & Repp, 2023), triggering the choice of the positive picture.

Russian positive V2 assertions produced with a contrastive focus on V were erroneously categorized as PQs in 50% of the cases, when presented out of context. While they did not find any miscategorizations in the opposite direction, concluding that the PQ accent is strongly distinctive for sentence mood in Russian, it is conceivable that negation could blur this clarity.

[25] Geist & Repp (2023) present experimental results that potentially contradict this claim, but acknowledge that their items might have had a semantic interpretation different from the one assumed in the literature. We leave this discussion for future research.

The reasoning about the ambiguity between low and high negation is comparable to that concerning French PQs in the Tian et al. (2021) study. Although a marginally significant shift of attention to the *p* picture was also found for negative PQs in French, it was much reduced in comparison to the clear-cut preference for the *p* picture in positive PQs. The authors hypothesize that the presence vs. absence of a positive epistemic bias may be responsible for the pattern observed in French PQs vs. English low-negation PQs. More research is needed to uncover the bias profile of Russian negated V2 questions and see whether similar considerations apply here.

6. Comparison of assertions and questions

In both Czech and Russian, V2 PQs do not differ from the corresponding assertions in word order. The only way to distinguish them is through intonation, specifically, by a dedicated pitch accent on the verb in Russian and by a rising tone at the end of the utterance in Czech (see Section 1 for a brief description). To investigate how illocutionary force expressed by intonation might impact the processing of an utterance, a cross-experimental comparison between assertions and V2-questions was conducted. We remind the reader that both experiments – Experiment 1 (assertions) and Experiment 2 (questions) – were part of a single experimental arrangement, which makes the comparison valid; effectively, it amounts to the manipulation of FORCE (assertion vs. question) between items (and within subjects).

If our hypothesis that the processing of PQs is affected by the positive bias they convey is correct, we expect questions to trigger more looks to the positive picture when compared to assertions. This difference is expected to be stronger for negative utterances, as negative assertions lead to more looks to $\overline{p}$, while negative questions still convey positive epistemic bias.

6.1 Descriptive findings

Figures 13 and **14** illustrate how the proportions of fixations on *p* and $\overline{p}$ evolved over time for assertions and V2 questions in Czech and Russian, respectively. The visualization suggests that assertions might exhibit a different pattern in comparison to questions. Whereas the picture preference while processing assertions seems to correspond to their polarity, questions show an overall tendency to fixations on *p*. An exception is formed by negative Russian V2 questions, already discussed in 5.5. Notably, questions in all the other conditions apparently display even a stronger preference for *p* than the positive assertions.

6.2 Models

Two categorical predictors were included: POLARITY and FORCE. Positive POLARITY and question (FORCE) were coded as 1, negative POLARITY and assertion (FORCE) as –1.

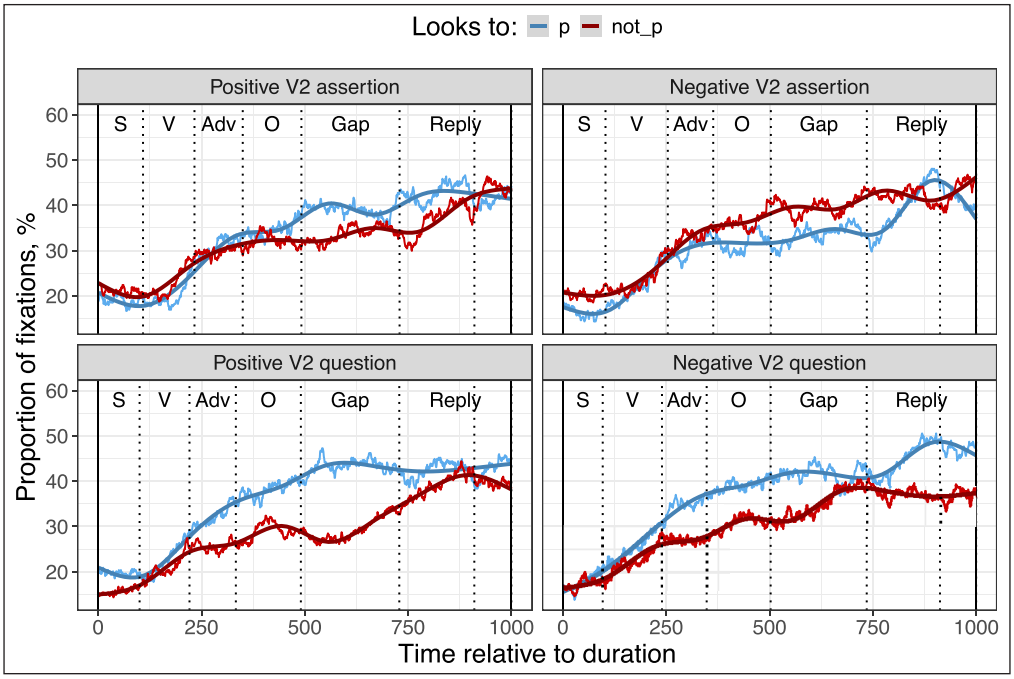

Figure 13: Proportions of looks to p and $\overline{p}$ in Czech assertions and V2 questions, normalized for sound duration across conditions (IP boundaries also normalized) with geom_smooth.

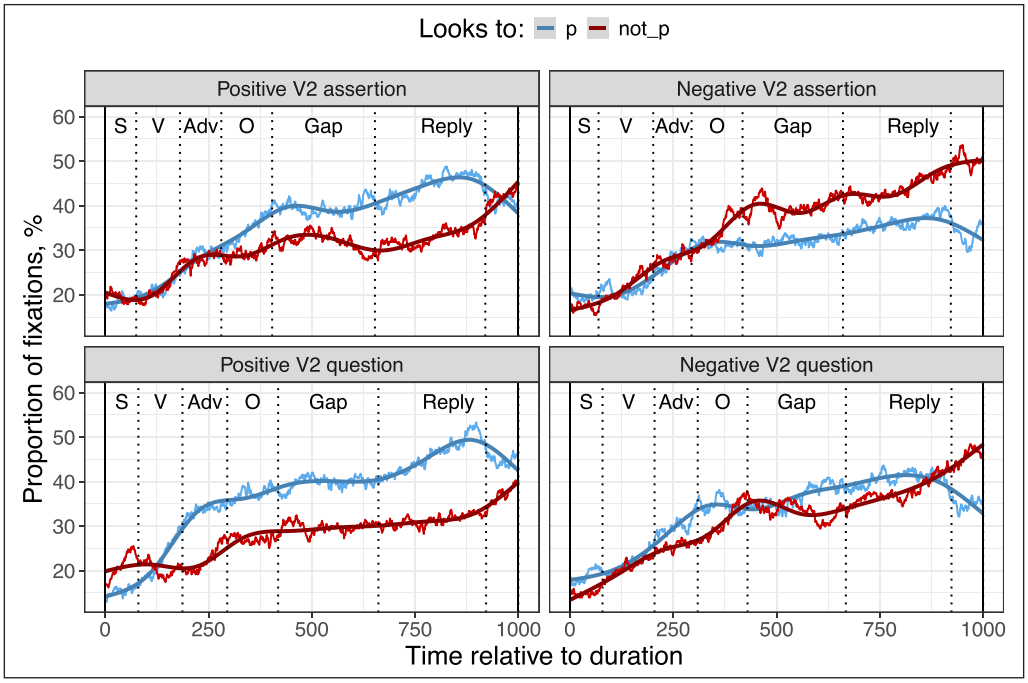

Figure 14: Proportions of looks to p and $\overline{p}$ in Russian assertions and V2 questions, normalized for sound duration across conditions (IP boundaries also normalized) with geom_smooth.

In this analysis as well, the starting point was the maximal reduced model.[26] Crucially, no random slope for FORCE by item was included. This term was not supported by the data, as FORCE was manipulated between items; all the items from Experiment 1 were assertions, and all items from Experiment 2 were questions.

6.3 Results

The mean LOG-RATIO values plotted across interest periods for different conditions and languages are shown in **Figure 15**. These representations technically repeat those given in **Figures 3** and **9** for assertions and V2 questions, respectively. However, in these visualizations, the conditions are combined differently, which allows for a comparison from another perspective. It is evident that assertions generally have lower LOG-RATIO values compared to V2 questions.

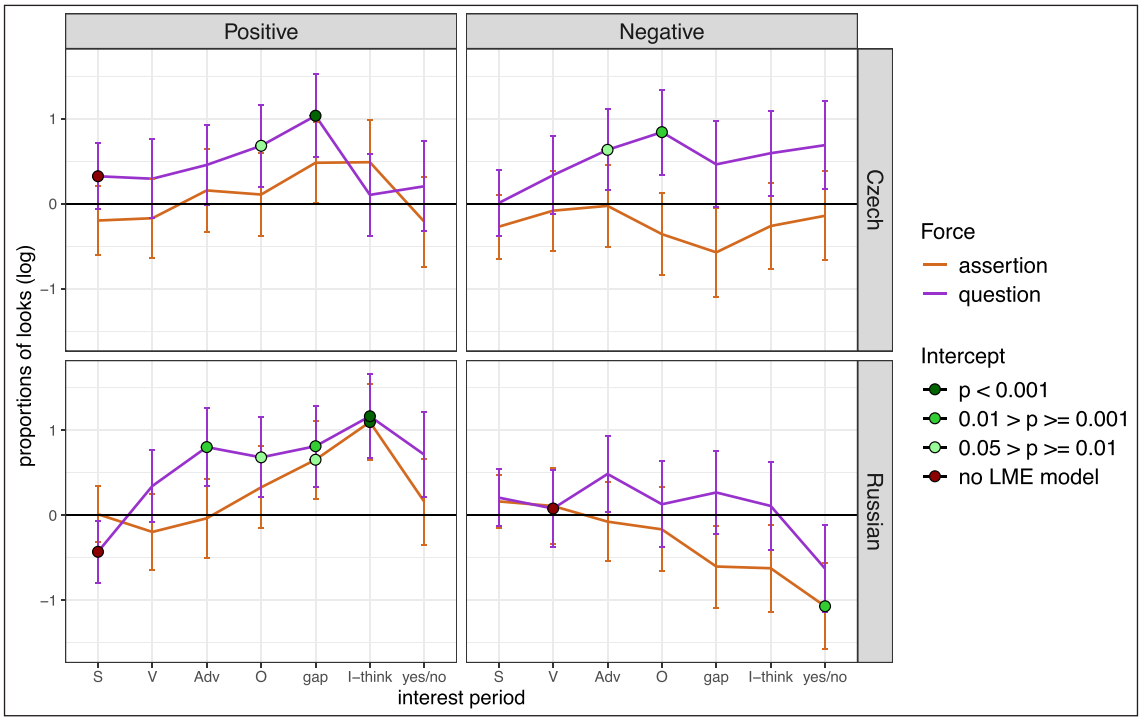

Figure 15: Average progression of LOG-RATIO of proportions of fixations on p over $\overline{p}$ by interest period in Czech and Russian assertions and V2 questions. Error bars indicate 95% confidence intervals. Green dots show where LOG-RATIO was significantly different from 0. Red dots show where no LME model could be fitted.

[26] Model formula: log_ratio_transf~polarity_contr * force_contr + (1 + polarity_contr||item) + (1 + polarity_contr * force_contr||participant).

6.3.1 Czech

Both positive and negative V2 questions reach a peak where the preference for p becomes significant. For assertions, on the other hand, no significant inclination could be found either for p or for $\overline{p}$. This distinction is supported by the LME models that contained FORCE and POLARITY as independent variables. A main effect of FORCE was discovered in the object ($t = 2.587$, $p = 0.014$) and in the gap ($t = 2.130$, $p = 0.04$), higher values being predicted for questions than assertions. Moreover, an interaction of FORCE and POLARITY was found in the first part of the response ($t = -2.362$, $p = 0.023$). While positive assertions instigated fixations on p and negative assertions on $\overline{p}$, predicted values for both positive and negative questions are above 0; interestingly, negative questions had a stronger preference for p. Lastly, a main effect of POLARITY manifested itself in the gap ($t = 3.070$, $p < 0.01$). These effects are visualized by the model prediction plots in the three leftmost panels of **Figure 16**.

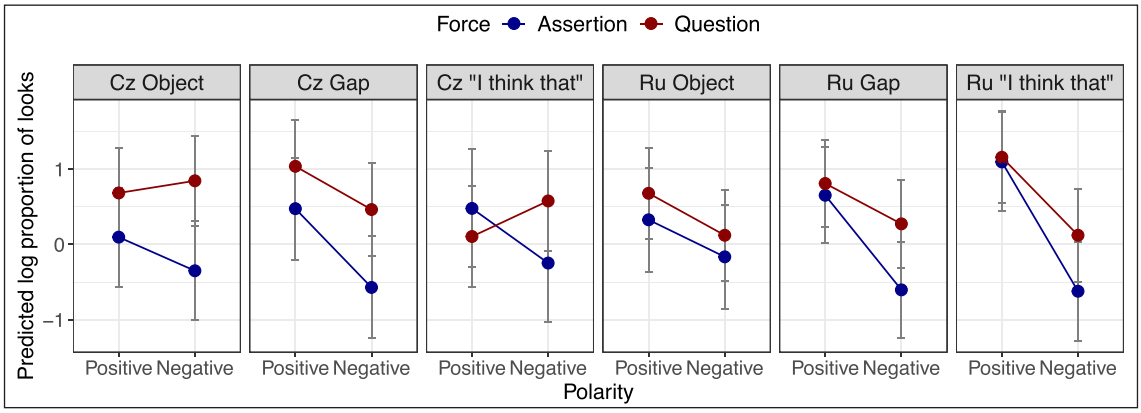

Figure 16: LOG-RATIO predicted by LME models in selected interest periods of Czech and Russian assertions and V2 questions.

6.3.2 Russian

The situation in Russian is less clear-cut. Although the tendency of questions to have higher LOG-RATIO values than assertions can be observed descriptively, no statistical confirmation could be obtained. In the positive condition, both questions and assertions reach a significant prevalence of p fixations. However, in questions it is obtained earlier (from the adverbial on) than in assertions (from the gap on). As for the negative utterances, both questions and assertions show a rather balanced pattern, with a downward tendency in the later periods. The preference for $\overline{p}$ becomes significant in the second part of the response to negative assertions (see 4.5 for discussion). The statistical analysis with FORCE and POLARITY as fixed predictors, therefore, could capture only the difference between positive and negative utterances, but not between questions and assertions. A main effect of POLARITY was discovered in the object ($t = 2.134$, $p = 0.04$), the gap

($t = 3.114, p < 0.01$), the first ($t = 4.143, p < 0.001$) and the second ($t = 3.594, p < 0.001$) parts of the response.[27]

6.4 Discussion

The comparison of V2 questions and assertions supports our interpretation of the results of the question experiment (see 5.5). We remind the reader that any differences between questions and assertions we observe here are due to their formal (prosodic) properties and not to their different discourse use: both types of utterances were presented in the very same discourse setting, namely, as lead-in utterances to be responded to by the addressee's (dis)agreement. In both languages, we see that questions instigate more positive fixations than the corresponding assertions; in Czech, this difference is statistically significant, but in Russian, only a tendency.[28] This is particularly striking in Czech, where, judging by the descriptive findings, negative questions appear to lead to stronger positive fixations than the corresponding positive assertions.

These results underline what appears to be an inherent "positivity" of PQs. That PQs – whether positive or negative – convey positive speaker-related bias has been discussed repeatedly in the theoretical literature. Our results, together with those of Tian et al. (2021), suggest that this meaning component is particularly salient in the processing of PQs.

Finally, we would like to point out the relatively small degree of preference of looks to p over $\overline{p}$ (or to $\overline{p}$ over p) in assertions (with the exception of the negative V2 condition in Russian; see 5.5 for some discussion). This preference rarely reached significance in assertions, while it did so relatively often in questions. In a sense, and very informally expressed, assertions behave more Hamblin-like than questions in our design: they instigate more balanced looks to p vs. $\overline{p}$ than the corresponding questions. One possibility why this may be so is that assertions (as well as questions) in our design were always followed by 'I think that yes/no'. The repeated exposure to this type of dialogue may have prompted the participants to expect a response and, therefore, interpret assertions as relatively neutral invitations to update the common ground, open to being challenged by the addressee's perspective. Compared to that, PQ force is apparently not just a move that sets two alternatives side-by-side, but rather one that conventionally conveys positive bias.

7. Experiment 3: Particles

We have attributed the "positivity" of PQs – both positive and negative – to the salience of the positive epistemic bias they convey. If our interpretation is right, we expect the fixations to be

[27] A main effect of POLARITY was also found in the subject in Russian ($t = -2.230, p = 0.035$). As no information about either the object discussed in the dialogue or the polarity of the utterance would be available at that point, this effect cannot be reliably interpreted based on our hypotheses.

[28] Arguably, the lack of the expected effect in Russian is at least partly caused by the balanced looks to p and $\overline{p}$ in the negative V2 question condition; see 5.5 for discussion.

affected by particles which convey or strengthen question bias. Experiment 3 addresses this hypothesis by investigating the processing of Russian questions with the particle *razve* and Czech questions with the particle *snad*.

7.1 Design and hypotheses

The Russian particle *razve*, licensed only in PQs, is dedicated to conveying/strengthening both epistemic and evidential bias, and, more specifically, the fact that they are in conflict. *Razve* occupies the clause-initial position and is followed by other constituents in a non-interrogative order (SVO, in our case); it is incompatible with V1 and *li*. Also for this reason, we only manipulate POLARITY in this experiment. It has been observed and argued (Korotkova, 2023; Restan, 1969) that a positive question with *razve*, such as (15a), conveys a positive evidential bias and negative epistemic bias. Negation reverses the biases, as in (15b).

(15) a. Razve Kostja narval na pole cvety?
RAZVE Kostja picked on field flowers
'Did Kostja pick flowers on the field? (It appears he did [evidential bias]; I thought he didn't [epistemic bias])'

b. Razve Kostja ne narval na pole cvety?
RAZVE Kostja NEG picked on field flowers
'Did Kostja not pick flowers on the field?' (It appears he didn't [evidential bias]; I thought he did [epistemic bias])'

Turning to Czech, although it possesses various particles which optionally occur in PQs, it does not have one that would correspond to the Russian *razve* straightforwardly. The question-initial particle *copak* (or *cožpak*) has been argued to have an analogous semantic effect (Nekula, 1996), but it belongs to a literary register, so we did not deem it suitable for our experimental setting. We instead decided to test the particle *snad*, which has been called "dubitative" in the Czech linguistic literature (Běličová, 1989; Štícha, 1995a, 1995b). Similarly to *razve*, it usually conveys a conflict of biases (Chodounská et al., 2025). It can occur in both interrogative (V1) and declarative questions, although the former is more frequent, which is why we use the verb-initial form in our design; see (16). Similar to *razve* and distinct from English *really* in PQs, Czech *snad* is compatible with both outer-negation and inner-negation readings.

(16) a. Nařezal snad Oldřich odpoledne dřevo?
cut SNAD Oldřich afternoon wood
'Did Oldřich cut the wood in the afternoon? (You seem to suggest he did [evidential bias], but I doubt it [epistemic bias].)'

b. Nenařezal snad Oldřich odpoledne dřevo?
NEG.cut SNAD Oldřich afternoon wood
'Didn't Oldřich cut the wood in the afternoon? (You seem to suggest he didn't [evidential bias], I thought he did [epistemic bias].)'

We hypothesize that the conflict of biases, made salient by *razve* and *snad*, will manifest itself in the processing of PQs in that participants will pay attention to both pictures. We assume that the particles encode not only epistemic but also evidential bias conventionally, in contrast to the PQs without *razve*, where evidential bias is more context-dependent (see 5.5). It is, therefore, to be expected that in this experiment, we will not observe such a strong preference for the positive picture, since both contradicting biases will be reflected in the processing. In the absence of more detailed knowledge of how the epistemic and evidential biases are processed (thus far, we only seem to have observed the effect of epistemic bias), we do not attempt to formulate more specific hypotheses about the direction and timing of fixations. The experiment, thus, has an exploratory component.

7.2 Russian *razve*: Results

As can be observed in **Figure 17**, questions with *razve* exhibit a characteristic pattern, especially in the positive condition. Although, in general, looks to p prevail, there are intervals during which the proportion of looks to $\overline{p}$ appears to be higher, based on the descriptive findings. Those intervals are a substantial part of the gap period as well as the second part of the reply. In the negation condition, the pattern appears more balanced, with a slight preference for p in the object period and parts of the gap and the first part of the reply. Towards the second part of the reply, a mild tendency towards $\overline{p}$ seems to emerge. There is also an interval of apparently balanced looks in the middle of the gap.

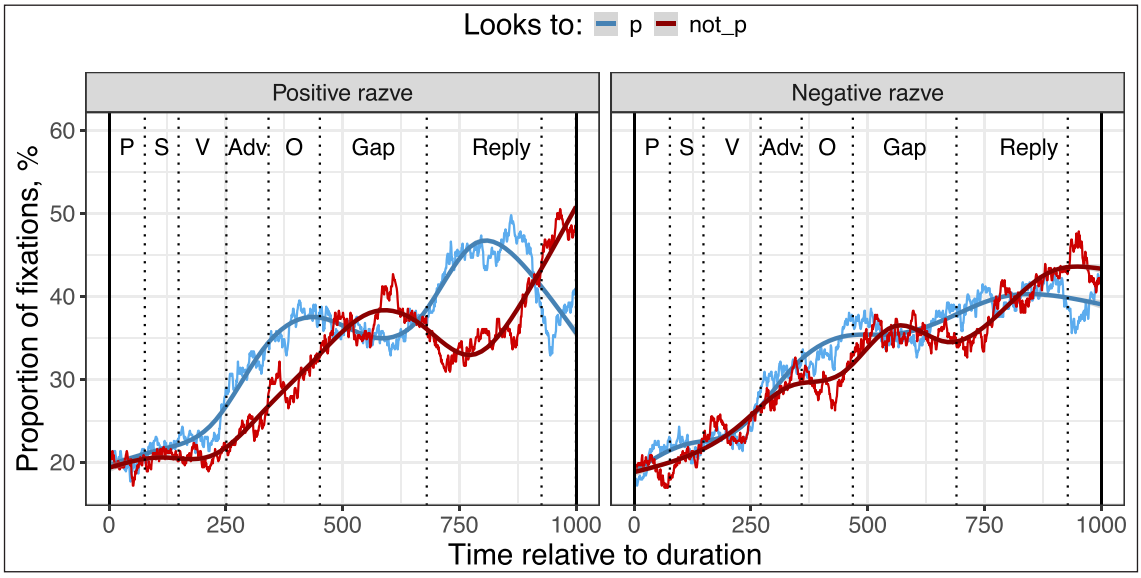

Figure 17: Proportions of looks to p and $\overline{p}$ in Russian questions with the particle *razve*, normalized for sound duration across conditions (IP boundaries also normalized) with geom_smooth.

The visualization of the mean LOG-RATIO progression across interest periods in **Figure 18** suggests that the differences between the proportions of fixations to p and $\overline{p}$ in *razve* questions are rather moderate. A significant preference for p is only present in the first part of the response in the positive condition. The analysis comparing positive *razve* questions to negative ones revealed no significant effect of POLARITY in any interest period.

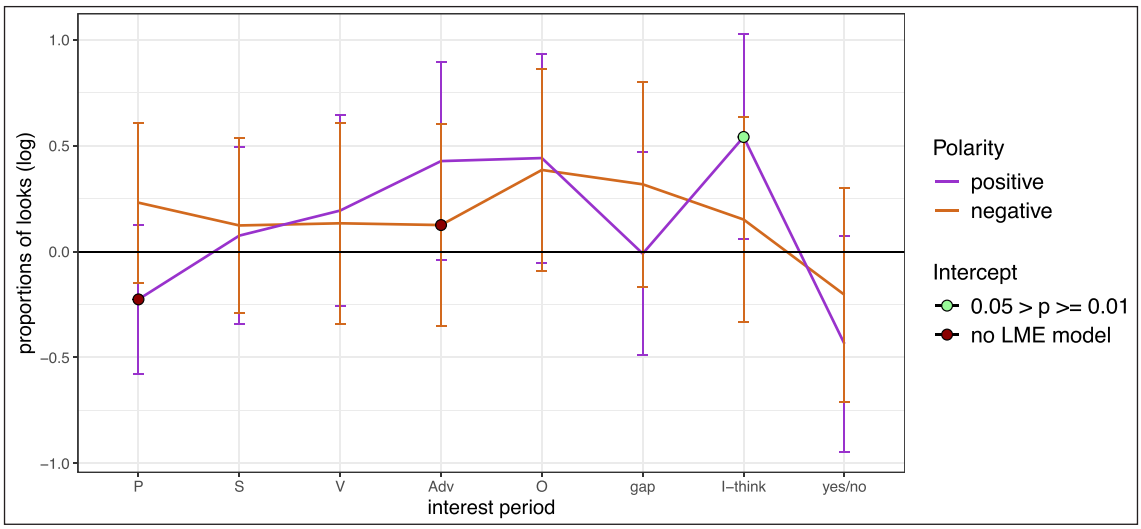

Figure 18: Average progression of log-ratio of proportions of fixations on p over $\overline{p}$ by interest period in Russian V2 questions with particle *razve*. Error bars indicate 95% confidence intervals. Green dots show where log-ratio was significantly different from 0. Red dots show where no LME model could be fitted.

To test if *razve* questions differ reliably from simple V2 questions, a cross-experimental analysis was conducted on the relevant data set. LME models with POLARITY and PARTICLE as independent variables were fitted to each interest period, starting from the subject.[29] Positive POLARITY and *razve* PARTICLE were coded as 1, negative POLARITY and no PARTICLE as –1. The particle period in *razve* questions was removed from the analysis, since no corresponding interval was present in the simple V2 condition and no information on the target object has been revealed up to this period. The statistical analysis did not reveal significant effects for interest periods during the question and the gap.

Moreover, a post-hoc test was conducted to investigate the temporary preference for $\overline{p}$ in the gap after the positive *razve* questions, as noticed above. A LME model was fitted to compare this interest period in questions with and without *razve* inside the positive condition. *Razve* PARTICLE was coded as 1, no PARTICLE as –1. A significant effect of PARTICLE was found ($t = -2.045$, $p = 0.0498$). It might indicate a reliable distinction between simple positive V2 questions, which

[29] Model formula: log_ratio_transf~polarity_contr∗particle_contr + (1 + polarity_contr||item) + (1 + polarity_contr∗particle_contr||participant).

show a clear preference for p during the gap (see, e.g., **Figure 8**) and positive *razve* questions, where an inclination towards $\overline{p}$ emerges in the same period.

7.3 Czech *snad*: Results

The data visualization in **Figure 19** indicates patterns somewhat similar to those observed for Russian *razve* questions (see **Figure 17**). Also in Czech *snad* questions, there seem to be fluctuations of fixation proportions, especially in the negative condition. Short intervals of a more balanced distribution of looks can be seen at the beginning of the gap and in some parts of the reply. In positive *snad* questions, there might also be a slight tendency towards a similar proportion of looks to p and $\overline{p}$ at the beginning of the gap. However, no preference for $\overline{p}$ comparable to the one found for Russian positive *razve* questions can be identified. A difference between the positive and the negative conditions seems to emerge during the response. In positive *snad* questions, the preference for p apparently grows towards the end of the dialogue, while in the negative condition, the fixation proportions become more balanced.

The results of the dependent variable analysis, depicted in **Figure 20**, are in accordance with these observations. The majority of the mean LOG-RATIO values are above 0, indicating more fixations on p than on $\overline{p}$. However, in no combination of interest period and condition was this preference estimated as significant by the corresponding LME model. A statistical analysis comparing positive and negative *snad* questions revealed a main effect of POLARITY in the second part of the response ($t = 2.194$, $p = 0.044$). In the positive condition, the dependent variable was estimated to be significantly higher than in the negative one.

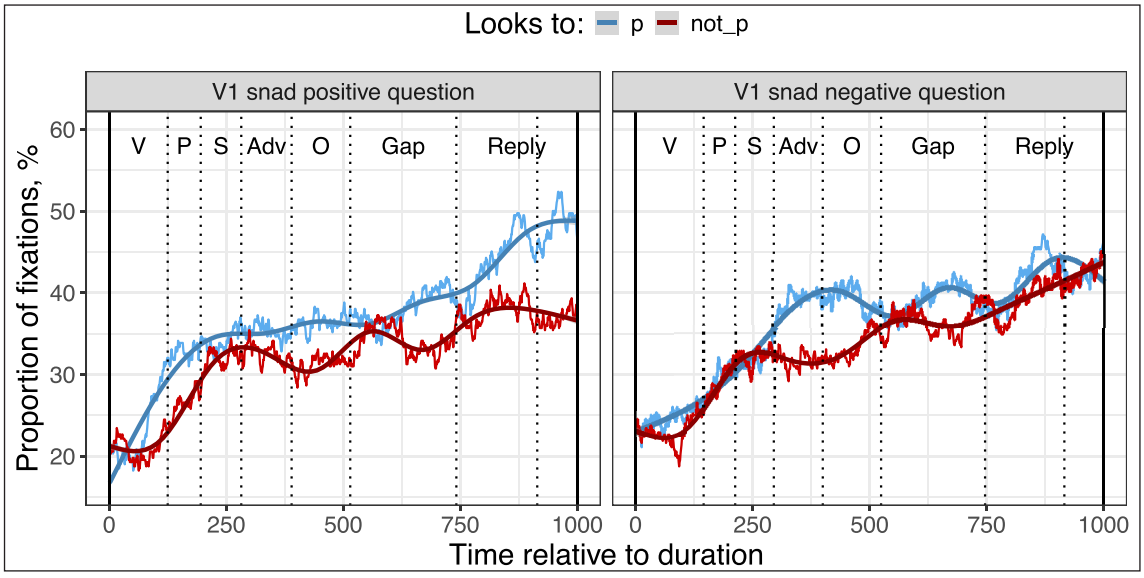

Figure 19: Proportions of looks to p and $\overline{p}$ in Czech questions with the particle *snad*, normalized for sound duration across conditions (IP boundaries also normalized) with geom_smooth.

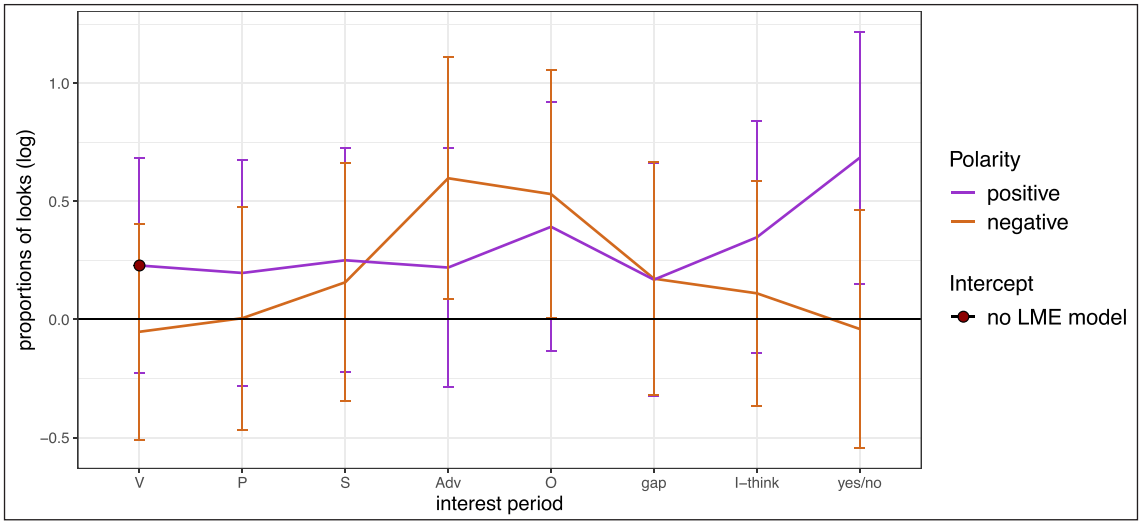

Figure 20: Average progression of log-ratio of proportions of fixations on p over $\overline{p}$ by interest period in Czech V1 questions with particle *snad*. Error bars indicate 95% confidence intervals. Green dots show where log-ratio was significantly different from 0. Red dots show where no LME model could be fitted.

7.4 Discussion

Despite a lack of decisive statistical support (arguably due to the relatively low statistical power of the experiment), the pattern we observe in the processing of *razve* questions – particularly in the positive condition – can be regarded as distinct from those we found for other PQs. Although certain conditions without *razve* also exhibit short intervals of apparently balanced looks or even preference of $\overline{p}$ (see **Figure 8**), we argue that this tendency is more pronounced at least in positive *razve* questions. Potential evidence for this is the effect of PARTICLE in the post-hoc comparison of positive V2 PQs with and without *razve*. For ease of exposition, we will use the ad hoc label *seesaw pattern* to describe the observed alternations in picture preference.

We believe that the seesaw pattern can be considered a preliminary validation of our hypothesis that bias plays a central role in the processing of PQs. The manner in which the participants' attention alternates between the positive and negative pictures is clearly systematic. If interpreted against the background of the corresponding questions without *razve*, we see significantly increased fixations on the negative picture in the positive condition of *razve* questions (and, correspondingly, decreased fixations on the positive picture). This can be understood as a reflection of the processing of the negative epistemic bias conveyed by the particle. This effect is preceded and followed by looks to the positive picture, suggesting an alternating consideration of the two conflicting biases.

Although we also observe a slight tendency toward the seesaw pattern in negative *razve* questions, the pattern there is much less prominent and less systematic. It is likely that the combination of *razve* and negation has led to a processing overload and, subsequently, to failure to yield a systematically alternating fixation pattern.

Albeit somewhat inconclusive, we consider the results of the *razve* experiment promising. Besides providing circumstantial support to our general claim that bias is essential for PQ processing, they also pave the way to a better understanding of how epistemic and evidential bias could be represented in the semantics of PQs and in an interlocutor's mind.

The results of the *snad* experiment are qualitatively similar to those of the *razve* experiment: we found a rather balanced pattern of fixations. We also observe a tendency towards the seesaw pattern, which we assume reflects the processing of both types of biases – epistemic (contrasting with the polarity of the question) and evidential (in accord with the question polarity). The results of the *snad* experiment are not statistically robust, but, together with the *razve* experiment results, provide a promising point of departure for a deeper investigation of the processing of question bias.

8. General discussion and conclusion

Our visual world eye-tracking study of Russian and Czech has contributed new findings which we hope will promote a better understanding of the semantics, pragmatics, and processing of polar questions, constructions of considerable semantic and pragmatic complexity. Counter to what one can reasonably consider to be the predictions of Hamblin's question semantics, and in accordance with previous experimental findings for English and French (Tian et al., 2021), the processing of a polar question generally does not lead to balanced fixations on both pictures representing the propositions in the question's denotation. What happens instead is that participants pay more attention to the picture corresponding to the positive alternative – largely independently of the question's polarity. Since we have also tested the processing of positive and negative assertions, we are confident that the observed pattern is not just residue of the experimental paradigm employed, but, indeed, the effect of prosodically or syntactically conveyed question mood. Following Tian et al. (2021), we have interpreted the "positivity" of polar questions as the salience of positive epistemic bias. The exact source of the bias could be different in positive and negative questions and will likely be modulated by contextual factors, which we have not controlled for in our experiment. But the fact that positive bias in both positive and negative polar questions has been identified in previous studies (AnderBois, 2019; Büring & Gunlogson, 2000; Gärtner & Gyuris, 2017; Šimík, to appear; among others; see Romero, 2024, for a recent survey) lends some credibility to our interpretation of the results.

The idea that the processing of polar questions is strongly affected by question bias has further been supported by the results of our experiments with particles. The particles *razve* in Russian and *snad* in Czech have both been argued to convey bias conflict: the mutual contradiction of epistemic and evidential bias. Despite the relatively low statistical power, the findings point towards a pattern different from that of ordinary polar questions, since we did not observe such a

strong preference for the positive picture. Instead, participants seem to pay attention to pictures representing both propositional alternatives. It is notable that the pattern of fixations is not stable: in the course of processing, positive and negative fixations appear to alternate repeatedly – what we have dubbed the seesaw pattern. We believe that this fixation pattern points to a promising avenue for future research, where the semantic/pragmatic nature of question biases (i.e., whether a particular type of bias is a presupposition, a conventional implicature, or a conversational implicature; cf. Goodhue, 2022; Korotkova, 2023; Romero, 2015; Romero & Han, 2004; among others) could be tested by investigating the timing and strength of positive and negative fixations.[30]

One question that remains open is why a similar seesaw pattern has not been observed in the processing of questions without particles, which have likewise been argued to be compatible with conflicting biases (Büring & Gunlogson, 2000; Domaneschi et al., 2017; Romero & Han, 2004; Šimík, to appear; a.o.). What we see instead are predominant positive fixations. One possible answer is that polar questions without particles (and, more generally, without other explicit strategies) are merely *compatible* with conflicting scenarios, but *do not convey* the conflict of biases. What seems compatible with this answer is Domaneschi et al.'s (2017) finding that, when faced with a conflicting scenario and when given the explicit choice, participants prefer to convey the conflict by using the adverb *really*, as in *Really?! Is there a train in the early morning?*[31]

While fixation timing was not a conclusive measure in our experiments with particles, it proved to be useful in interpreting the results of the question experiment. In accordance with our hypothesis that polar question processing is affected by positive epistemic bias, we see that positive fixations tend to increase soon after it becomes clear to the participant that what s/he is listening to is a polar question.

Our expectations about the cross-linguistic differences between Russian and Czech have only partially been borne out. One of the differences concerns the use of *li* in V1 questions in Russian. Due to its morphological and arguably semantic kinship to *ili* 'or' (Metslang et al., 2017), we expected this particle to activate both polar alternatives more strongly than in Czech, where no such particle is used. This expectation was borne out in positive questions, where the onset of the predominantly positive fixations is delayed, compared to the Czech V1 condition. No such

[30] Psycholinguistic work on presuppositions and implicatures is extensive but sometimes conflicting. Still, most studies agree that they both integrate immediately and pose little difficulty (e.g., Degen et al., 2021; Hirsch et al., 2019; Schwarz, 2015). Presuppositions and scalar implicatures sometimes show similar response times, but prosody and entailment tests reveal clear differences and support the distinction between them (Bill et al., 2018). With more research, it will become clearer what type of inference question bias is.

[31] Domaneschi et al. (2017) only offered the "really" option in a scenario with positive evidential and negative epistemic bias (correlating with positive *really* questions), not in the opposite pattern, so it remains unclear how *really* would be used in negative questions.

effect is observed in negative V1 + *li* questions, in which the positive fixations step in very early, arguably due to the positive epistemic bias associated with high negation. Another difference concerns the use of V2 questions, which are characterized as neutral in Russian (Esipova, 2025), but as evidentially biased toward the proffered alternative in Czech (Staňková & Šimík, 2025). We observe no effects which would be consistent with these expectations. We attribute this to the fact that evidential bias is not made salient enough in our paradigm to affect the fixation patterns.

Our study of polar question online processing in Czech and Russian is among the first of its kind. It has corroborated previous findings (Tian et al., 2021) on French and English, while also yielding qualitatively new insights. At the same time, it leaves many issues open and suggests new research directions. Having seen that bias plays a pivotal role in the processing of polar questions, it remains to be determined why the dominant bias is positive, and arguably epistemic; negative epistemic bias was only reliably elicited in Russian positive questions with *razve*. This might have to do with the general primacy of the positive prejacent in the processing of negation (Tian et al., 2016). However, the nature of epistemic bias could also play a role. If epistemic bias is a conventional implicature, for instance (cf. Korotkova, 2023; Romero, 2015), potentially conveying information that is new to the addressee, it might be perceived as more salient than evidential bias, which is arguably a presupposition (Korotkova, 2023), and, hence, backgrounded and less salient. The salience of bias in the processing of polar questions calls for experimental designs with contextual manipulations. Controlling the context could help to make the negative alternative and/or evidential bias more salient and, thus, detectable in processing measures. Another way of bringing out biases, their semantic nature, their pragmatic source, and the timing of their processing is by investigating specific interrogative constructions or particles, as we have attempted in Experiment 3. Despite the relatively low statistical power, we saw that questions with particles indeed exhibit a pattern consistent with the processing of a conflict between epistemic and evidential bias.

Data accessibility statement

All data, scripts and notebooks are available on OSF; see Onoeva et al. (2025): osf.io/w9rsc.

Ethics and consent

The Research Ethics Committee of the Faculty of Arts, Charles University has found no inconsistencies or shortcomings with respect to the requirements arising from the legislation of the Czech Republic, internal and other regulations (Ref.No.: UKFF/299333/2023, in Prague 19 May 2023). The Ethics Committee of the German Linguistic Society (DGfS) has found no ethical concerns regarding the conduct of the study (Ethikvotum No. 2020-18-210112, in Frankfurt/ Main 12 January 2021).

Funding

This study was conducted within the project “Modelling the question–statement opposition in Slavic languages (QueSlav)”, funded by the Czech Science Foundation (GAČR), project No. 21-31488J, and by the German Science Foundation (DFG), project No. ME4125-5/1.

Acknowledgments

The authors thank Pia Knoeferle, Filip Smolík, Jan Volín, Daniela Palleschi, and Anna Laurinavichyute for helpful feedback and methodological advice. We are grateful to Georg Lohfink for assistance with data collection, experimental programming, and software troubleshooting. We also thank Michaela Chodounská and Adéla Koštejnová for data collection in Prague, and Hana Dufková for annotation in Praat. We thank audiences at the Slavic Linguistics Colloquium, the ERCEL Colloquium at Charles University, the workshop “Polar Question Meaning[s] Across Languages” at the University of Amsterdam, the Psycholinguistik-Kolloquium at HU Berlin, 11th Novi Sad Workshop on Psycholinguistic, Neurolinguistic and Clinical Linguistic Research (PNCLR11), and Psycholinguistics of Slavic Languages (PsychoSlav 2024) at the University of Wrocław for valuable feedback on earlier versions of this work. The experimental stimuli and visuals were kindly provided by Ye Tian and colleagues, for which we are very grateful. We are grateful to the editor, the proofreaders and the anonymous reviewers for their insightful comments, which improved the manuscript.

Competing interests

The authors have no competing interests to declare.

Authors' contributions

- Mariia Razguliaeva contributed to (i) research questions and hypotheses formulation, (ii) preparation of experimental stimuli and data collection in Berlin, (iii) data analysis and statistics, (iv) writing sections 3.3, 3.4, and the descriptive findings, models, and results subsections for the subexperiments, as well as (v) revising and proofreading the manuscript.
- Maria Onoeva contributed to (i) research questions and hypotheses formulation, (ii) experimental design and preparation of experimental stimuli, (iii) data analysis and statistics, (iv) writing sections 3.1, 3.2, as well as (v) revising and proofreading the manuscript.
- Radek Šimík initiated the research and contributed to (i) research question and hypothesis formulation, (ii) experimental design, (iii) writing of sections 1, 2, and the discussion parts, and (iv) revising and proofreading the manuscript.
- Roland Meyer co-initiated the research and contributed to (i) research question and hypothesis formulation, (ii) writing of sections 3 (intro), 5.5, 6 and 8, and (iii) revising and proofreading the manuscript.
- Kateřina Hrdinková adapted Tian and colleagues' visual stimuli and translated the linguistic stimuli to Czech.

ORCiD IDs

Mariia Razguliaeva: 0009-0006-6791-0844

Maria Onoeva: 0000-0001-8115-5183

Radek Šimík: 0000-0002-4736-195X

Roland Meyer: 0000-0003-2023-0527

References

Abels, K. (2007). Deriving selectional properties of 'exclamative' predicates. In A. Späth (Ed.), *Interfaces and interface conditions* (pp. 115–140). De Gruyter. https://doi.org/10.1515/9783110926002.115

Alonso-Ovalle, L. (2006). *Disjunction in alternative semantics* [Doctoral Dissertation]. University of Massachusetts. https://www.semanticsarchive.net/Archive/TVkY2ZlM/alonso-ovalle2006.pdf

AnderBois, S. (2019). Negation, alternatives, and negative polar questions in American English. In K. von Heusinger, E. Onea, & M. Zimmermann (Eds.), *Questions in discourse – Volume 1: Semantics* (pp. 118–171). Brill. https://doi.org/10.1163/9789004378308_004

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge University Press. https://doi.org/10.1017/CBO9780511801686

Baltazani, M., Kainada, E., Nicolaidis, K., & Lengeris, A. (2015). The prenuclear field matters: Questions and statements in standard modern Greek. *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS 18)*. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0941.pdf

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2018). Parsimonious mixed models. https://arxiv.org/abs/1506.04967

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Běličová, H. (1989). Zjišťovací otázky kladné a záporné v slovanských jazycích. *Slavia: Časopis pro slovanskou filologii*, *58*, 97–105.

Biezma, M., & Rawlins, K. (2012). Responding to alternative and polar questions. *Linguistics and Philosophy*, *35*(5), 361–406. https://doi.org/10.1007/s10988-012-9123-z

Bill, C., & Koev, T. (2022). Really: Ambiguity and question bias. *Proceedings of Sinn und Bedeutung*, *26*, 130–148. https://doi.org/10.18148/SUB/2022.V26I0.992

Bill, C., Romoli, J., & Schwarz, F. (2018). Processing presuppositions and implicatures: Similarities and differences. *Frontiers in Communication*, *3*. https://doi.org/10.3389/fcomm.2018.00044

Boersma, P., & Weenink, D. (2009). Praat: Doing phonetics by computer (version 5.1.13). http://www.praat.org

Bolinger, D. (1978). Yes–no questions are not alternative questions. In H. Hiż (Ed.), *Questions* (pp. 87–105). Reidel. https://doi.org/10.1007/978-94-009-9509-3_3

Bosker, H. R., Quené, H., Sanders, T., & De Jong, N. H. (2014). Native 'um's elicit prediction of low-frequency referents, but non-native 'um's do not. *Journal of Memory and Language*, *75*, 104–116. https://doi.org/10.1016/j.jml.2014.05.004

Brown, V. A. (2021). An introduction to linear mixed-effects modeling in R. *Advances in Methods and Practices in Psychological Science*, *4*(1). https://doi.org/10.1177/2515245920960351

Büring, D., & Gunlogson, C. (2000). Aren't positive and negative polar questions the same? *Ms. UCSC/UCLA*. https://semanticsarchive.net/Archive/mYwOGNhO/polar_questions.pdf

Chodounská, M., Koštejnová, A., & Šimík, R. (2025). Částice *náhodou* a *snad* ve zjišťovacích otázkách. *Slovo a slovesnost*, *86*(3), 203–235. https://doi.org/10.58756/s6238632

Ciardelli, I., Groenendijk, J., & Roelofsen, F. (2019). *Inquisitive semantics*. Oxford University Press. https://doi.org/10.1093/oso/9780198814788.001.0001

Cross, C., & Roelofsen, F. (2014). Questions. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/sum2015/entries/questions/

Davidson, D. (1967). Truth and meaning. *Synthese*, *17*(1), 304–323. https://doi.org/10.1007/BF00485035

Dayal, V. (2016). *Questions*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199281268.001.0001

Degen, J., Kursat, A., & Leigh, J. (2021). Seeing is believing: Testing an explicit linking assumption for visual world eye-tracking in psycholinguistics. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *43*. https://escholarship.org/uc/item/6182t9jb

Domaneschi, F., Romero, M., & Braun, B. (2017). Bias in polar questions: Evidence from English and German production experiments. *Glossa: A Journal of General Linguistics*, *2*(1), 26. 1–28. https://doi.org/10.5334/gjgl.27

Esipova, M. (2021). Polar responses in Russian across modalities and across interfaces. *Journal of Slavic Linguistics*, *29*(FASL 28 extra issue), 1–11. https://ojs.ung.si/index.php/JSL/article/view/151

Esipova, M. (2025). Prosody across sentence types. *Semantics and Linguistic Theory*, *34*, 68–87. https://doi.org/10.3765/pe3dtd58

Gärtner, H.-M., & Gyuris, B. (2017). On delimiting the space of bias profiles for polar interrogatives. *Linguistische Berichte*, *251*, 26–49. https://doi.org/10.46771/2366077500251_2

Geist, L., & Repp, S. (2023). Responding to negative biased questions in Russian. In P. Biskup, M. Börner, O. Mueller-Reichau, & I. Shcherbina (Eds.), *Advances in formal Slavic linguistics 2021* (pp. 145–177). Language Science Press. https://langsci-press.org/catalog/book/411

González-Fuente, S., Tubau, S., Espinal, M. T., & Prieto, P. (2015). Is there a universal answering strategy for rejecting negative propositions? Typological evidence on the use of prosody and gesture. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00899

Goodhue, D. (2022). Isn't there more than one way to bias a polar question? *Natural Language Semantics*, *30*(4), 379–413. https://doi.org/10.1007/s11050-022-09198-2

Groenendijk, J. A. G., & Stokhof, M. J. B. (1984). *Studies on the semantics of questions and the pragmatics of answers* [Doctoral Dissertation]. University of Amsterdam. https://stokhof.org/wp-content/uploads/2020/09/groenendijk-stokhof_ssqpa.pdf

Gunlogson, C. (2002). Declarative questions. *Semantics and Linguistic Theory*, *12*, 124–143. https://doi.org/10.3765/salt.v12i0.2860

Gunlogson, C. (2008). A question of commitment. *Belgian Journal of Linguistics*, *22*, 101–136. https://doi.org/10.1075/bjl.22.06gun

Gutzmann, D. (2015). *Use-conditional meaning: Studies in multidimensional semantics*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198723820.001.0001

Hallett, P. E. (1986). Eye movements. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (pp. 78–101, Vol. 1). Wiley.

Hamblin, C. L. (1973). Questions in Montague English. *Foundations of Language*, *10*(1), 41–53. https://www.jstor.org/stable/25000703

Hirsch, A., Zehr, J., & Schwarz, F. (2019). Presupposition projection from disjunction in online processing. *Proceedings of Sinn und Bedeutung, 21*(1), 547–566. https://ojs.ub.uni-konstanz.de/sub/index.php/sub/article/view/154

Hrdinková, K., & Šimík, R. (2025). The meaning of Czech response particles. In B. Gehrke, D. Lenertová, R. Meyer, D. Seres, L. Szucsich, & J. Zaleska (Eds.), *Advances in formal Slavic linguistics 2022* (pp. 225–254). Language Science Press. https://doi.org/10.5281/zenodo.15394176

Ito, A., & Knoeferle, P. (2022). Analysing data from the psycholinguistic visual-world paradigm: Comparison of different analysis methods. *Behavior Research Methods, 7*(55), 3461–3493. https://doi.org/10.3758/s13428-022-01969-3

Karttunen, L. (1977). Syntax and semantics of questions. *Linguistics and Philosophy*, *1*(1), 3–44. https://www.jstor.org/stable/25000027

Kisler, T., Reichel, U., & Schiel, F. (2017). Multilingual processing of speech via web services. *Computer Speech & Language*, *45*, 326–347. https://doi.org/10.1016/j.csl.2017.01.005

Korotkova, N. (2023). Conversational dynamics of Russian questions with *razve*. In M. Onoeva, A. Staňková, & R. Šimík (Eds.), *Proceedings of Sinn und Bedeutung 27* (pp. 328–346). Institute of Czech Language and Linguistic Theory, Faculty of Arts, Charles University. https://doi.org/10.18148/sub/2023.v27.1073

Korotkova, N. (submitted). A new perspective on negative bias in polar questions: The view from Russian. In R. Eckardt, G. Walkden, & N. Dehé (Eds.), *The Oxford handbook of non-canonical questions*. https://lingbuzz.net/lingbuzz/008228

Krifka, M. (2001). For a structured meaning account of questions and answers. In C. Féry & W. Sternefeld (Eds.), *Audiatur vox sapientiae: A festschrift for Arnim von Stechow* (pp. 287–319). Akademie-Verlag. https://doi.org/10.1515/9783050080116.287

Krifka, M. (2006). Association with focus phrases. In V. Molnár & S. Winkler (Eds.), *The architecture with focus phrases* (pp. 105–136). de Gruyter. https://doi.org/10.1515/9783110922011.105

Krifka, M. (2011). Questions. In K. von Heusinger, C. Maienborn, & P. Portner (Eds.), *Semantics: An international handbook of natural language meaning* (pp. 1742–1785, Vol. 2). de Gruyter. https://doi.org/10.1515/9783110255072.1742

Krifka, M. (2013). Response particles as propositional anaphors. In T. Snider (Ed.), *SALT 23: Proceedings from the 23rd conference on Semantics and Linguistic Theory* (pp. 1–18). CLC Publications. https://doi.org/10.3765/salt.v23i0.2676

Krifka, M. (2015). Bias in commitment space semantics: Declarative questions, negated questions, and question tags. In S. D'Antonio, M. Moroney, & C.-R. Little (Eds.), *SALT 25: Proceedings from the 25th conference on Semantics and Linguistic Theory* (pp. 328–345). CLC Publications. https://doi.org/10.3765/salt.v25i0.3078

Krifka, M., & Trinh, T. (2025). Introduction. In T. Trinh, A. Benz, D. Goodhue, K. Yatsushiro, & M. Krifka (Eds.), *Biased questions: Experimental results and theoretical modelling* (pp. v–xxiv). Language Science Press. https://doi.org/10.5281/zenodo.17158174

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Lüdecke, D. (2018). Ggeffects: Tidy data frames of marginal effects from regression models. *Journal of Open Source Software*, *3*(26), 772. https://doi.org/10.21105/joss.00772

Magnuson, J. S. (2019). Fixations in the visual world paradigm: Where, when, why? *Journal of Cultural Cognitive Science*, *3*(2), 113–139. https://doi.org/10.1007/s41809-019-00035-3

Malamud, S. A., & Stephenson, T. (2015). Three ways to avoid commitments: Declarative force modifiers in the conversational scoreboard. *Journal of Semantics*, *32*(2), 275–311. https://doi.org/10.1093/jos/ffu002

Metslang, H., Habicht, K., & Pajusalu, K. (2017). Where do polar question markers come from? *STUF – Language Typology and Universals*, *70*(3), 489–521. https://doi.org/10.1515/stuf-2017-0022

Meyer, R., & Mleinek, I. (2006). How prosody signals force and focus – A study of pitch accents in Russian yes–no questions. *Journal of Pragmatics*, *38*(10), 1615–1635. https://doi.org/10.1016/j.pragma.2005.05.011

Müller, K. (2020). *Here: A simpler way to find your files* [R package version 1.0.1]. https://CRAN.R-project.org/package=here

Munteanu, A., & Kiss, A. (2025). Form-meaning relations in Russian confirmative and surprise declarative questions. *Language and Speech*. https://doi.org/10.1177/00238309251314862

Nekula, M. (1996). *System der Partikeln im Deutschen und Tschechischen: Unter besonderer Berücksichtigung der Abtönungspartikeln*. Niemeyer. https://doi.org/10.1515/9783110930924

Onoeva, M., & Staňková, A. (2025). Polar questions in Russian and Czech: An exploratory spoken corpus investigation. In B. Gehrke, D. Lenertová, D. Seres, R. Meyer, L. Szucsich, & J. Zaleska (Eds.), *Advances in formal Slavic linguistics 2022* (pp. 429–446). Language Science Press. https://doi.org/10.5281/zenodo.15394189

Onoeva, M., Razguliaeva, M., Šimík, R., & Meyer, R. (2025). Processing of polar questions in Czech and Russian: Evidence for the effect of question bias [Open Science Framework, paper data]. osf.io/w9rsc

OpenAI. (2025). *ChatGPT* (Version 4o). https://chat.openai.com/chat

Pastukhov, A. (2024). *eyelinkReader: Import gaze data for EyeLink eye tracker* [R package version 1.0.2. https://alexander-pastukhov.github.io/eyelinkReader/]. https://github.com/alexander-pastukhov/eyelinkReader/

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Repp, S., & Geist, L. (2025). Negative polar questions in Russian: Question bias and question concern. In T. Trinh, A. Benz, D. Goodhue, K. Yatsushiro, & M. Krifka (Eds.), *Biased questions: Experimental results and theoretical modelling* (pp. 101–139). Language Science Press. https://doi.org/10.5281/zenodo.17158182

Restan, P. (1969). *Sintaksis voprositel'nogo predloženija. Obščij vopros.* Universitetsforlaget.

Roberts, C. (1996). Information structure in discourse: Towards an integrated formal theory of pragmatics. In J. H. Yoon & A. Kathol (Eds.), *Ohio State University Working Papers in Linguistics 49: Papers in Semantics* (pp. 91–136).

Roberts, C. (2012). Information structure: Towards an integrated formal theory of pragmatics. *Semantics & Pragmatics*, *5*(6), 1–69. https://doi.org/10.3765/sp.5.6

Roelofsen, F. (2019). Semantic theories of questions. In M. Aronoff (Ed.), *Oxford research encyclopedia of linguistics*. Oxford University Press. https://doi.org/10.1093/acrefore/9780199384655.013.504

Roelofsen, F., & Farkas, D. F. (2015). Polarity particle responses as a window onto the interpretation of questions and assertions. *Language*, *91*(2), 359–414. https://www.jstor.org/stable/24672234

Roelofsen, F., & van Gool, S. (2010). Disjunctive questions, intonation, and highlighting. In M. Aloni, H. Bastiaanse, T. Jager, & K. Schulz (Eds.), *Logic, language, and meaning: 17th Amsterdam Colloquium, Amsterdam, December 2009, Revised selected papers* (pp. 384–394). Springer. https://doi.org/10.1007/978-3-642-14287-1_39

Roelofsen, F., Venhuizen, N., & W. Sassoon, G. (2012). Positive and negative polar questions in discourse. *Proceedings of Sinn und Bedeutung*, *17*, 455–472. https://ojs.ub.uni-konstanz.de/sub/index.php/sub/article/view/355

Romero, M. (2015). High negation in subjunctive conditionals and polar questions. In E. Csipak & H. Zeijlstra (Eds.), *Proceedings of Sinn und Bedeutung 19* (pp. 519–536). University of Göttingen. https://doi.org/10.18148/sub/2015.v19i0.247

Romero, M. (2024). Biased polar questions. *Annual Review of Linguistics*, *10*, 279–302. https://doi.org/10.1146/annurev-linguistics-022421-064837

Romero, M., & Han, C.-H. (2004). On negative yes/no questions. *Linguistics and Philosophy*, *27*(5), 609–658. https://doi.org/10.1023/b:ling.0000033850.15705.94

Romportl, M. (1951). Melodie (tónový průběh) otázky zjišťovací v hovorové střední češtině. *Listy filologické / Folia philologica*, *75*(6), 268–274. https://www.jstor.org/stable/23459298

Rudis, B., Bolker, B., & Schulz, J. (2017). *Ggalt: Extra coordinate systems, 'geoms', statistical transformations, scales and fonts for 'ggplot2'* [R package version 0.4.0]. https://CRAN.R-project.org/package=ggalt

Schwarz, F. (2015). Presuppositions are fast, whether hard or soft – Evidence from the visual world. *Semantics and Linguistic Theory*, *24*, 1–22. https://doi.org/10.3765/salt.v24i0.2414

Signorell, A. (2024). *Desctools: Tools for descriptive statistics* [R package version 0.99.58]. https://CRAN.R-project.org/package=DescTools

Šimík, R. (to appear). Polar question semantics and bias: Lessons from Czech/Slavic. In B. Gehrke & R. Šimík (Eds.), *Topics in the semantics of Slavic languages*. Language Science Press. https://lingbuzz.net/lingbuzz/007552

Šimík, R., Razguliaeva, M., Onoeva, M., & Meyer, R. (2023). Processing of polar questions in Czech and Russian: An eye-tracking study [Open Science Framework, preregistration]. https://doi.org/10.17605/OSF.IO/NWKA9

SR Research Ltd. (2020). *SR Research Experiment Builder 2.3.1* [computer software]. Mississauga, Ontario, Canada. https://www.sr-research.com/experiment-builder/

Staňková, A. (2023). *The expression of speaker's bias in Czech polar questions* [Master's thesis]. Charles University. https://hdl.handle.net/20.500.11956/182788

Staňková, A., & Šimík, R. (2025). Negation in Czech polar questions. *Journal of Slavic Linguistics*, *33*(FASL issue), 1–16. https://ojs.ung.si/index.php/JSL/article/view/472

Štícha, F. (1995a). Otázky predikátové: Inference, implicitnost a explicitní výrazy ilokučních funkcí I. *Slovo a slovesnost*, *56*(2), 98–109. http://sas.ujc.cas.cz/archiv.php?art=3625

Štícha, F. (1995b). Otázky predikátové: Inference, implicitnost a explicitní výrazy ilokučních funkcí II. *Slovo a slovesnost*, *56*(3), 204–217. http://sas.ujc.cas.cz/archiv.php?art=3638

Sudo, Y. (2013). Biased polar questions in English and Japanese. In D. Gutzmann & H.-M. Gärtner (Eds.), *Beyond expressives: Explorations in use-conditional meaning* (pp. 275–295). Brill. https://doi.org/10.1163/9789004183988_009

Szabolcsi, A. (1981). Compositionality in focus. *Folia Linguistica*, *15*(1–2), 141–162. https://doi.org/10.1515/flin.1981.15.1-2.141

Tian, Y., Ferguson, H., & Breheny, R. (2016). Processing negation without context – Why and when we represent the positive argument. *Language, Cognition and Neuroscience*, *31*(5), 683–698. https://doi.org/10.1080/23273798.2016.1140214

Tian, Y., van Tiel, B., Clin, É., & Breheny, R. (2021). Representing polar questions. *Journal of Psycholinguistic Research*, *50*(6), 1535–1555. https://doi.org/10.1007/s10936-021-09814-y

van den Brand, T. (2024). *Ggh4x: Hacks for 'ggplot2'* [R package version 0.3.0]. https://CRAN.R-project.org/package=ggh4x

van Rooy, R., & Šafářová, M. (2003). On polar questions. *Semantics and Linguistic Theory*, *13*, 292–309. https://doi.org/10.3765/salt.v13i0.2887

Veroňková, J. (2002). *Rozdíly v průběhu základního tónu relevantní pro percepční rozlišování melodémů v češtině* [Doctoral dissertation]. Charles University. https://hdl.handle.net/20.500.11956/7717

Wickham, H. (2011). The split-apply-combine strategy for data analysis. *Journal of Statistical Software*, *40*(1), 1–29. https://www.jstatsoft.org/v40/i01/

Wickham, H. (2016). *Ggplot2: Elegant graphics for data analysis*. Springer-Verlag. https://ggplot2.tidyverse.org

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wickham, H., Henry, L., & Vaughan, D. (2023). *Vctrs: Vector helpers* [R package version 0.6.5]. https://CRAN.R-project.org/package=vctrs