Differential effects of constraints in the processing of Russian cataphora

Nina Kazanina

Department of Experimental Psychology, University of Bristol, Bristol, UK

Colin Phillips

Department of Linguistics, Neuroscience and Cognitive Science Program, University of Maryland, College Park, MD, USA

Anaphoric relations between pronouns and their antecedents are subject to a number of different linguistic constraints, which exclude the possibility of coreference in specific syntactic or discourse contexts. Constraints on anaphora may, in principle, impact online sentence processing in a couple of different ways. They may act as constraints on the generation of interpretations, preventing illicit anaphoric relations from ever being considered. Alternatively, they may act as later filters on interpretations, rejecting candidate interpretations after initial consideration. A number of previous studies have sought to determine which of these mechanisms accurately describes the online impact of constraints on anaphora. The current studies present evidence that there is no uniform answer to this question, and that the two mechanisms are both used, for different constraints. Evidence for this is drawn from studies on the processing of two constraints on backwards anaphora or cataphora in Russian that apply in superficially similar contexts but that differ in a number of respects. One self-paced reading study and two judgement studies are reported. The self-paced reading study manipulates the gender congruency between a pronoun and a following name in three pairs of conditions. In conditions where the pronoun–name configuration violates no constraints on anaphora a gender mismatch effect was observed following the name, as in previous studies, suggesting that comprehenders actively search for an antecedent following a cataphoric pronoun. In conditions where the pronoun–name configuration violates Principle C of the classical binding theory no effect of the gender manipulation was observed, suggesting that comprehenders do not even consider the possibility of interpretations that violate this constraint. In conditions where the pronoun–name configuration violates a Russian-specific constraint on cataphora a gender match effect was observed following the name, the reverse of the finding in the no-constraint conditions, suggesting that the constraint applies as a filter on candidate interpretations.

Keywords: Backwards anaphora; Cataphora; Sentence processing; Russian; Principle C.

371

Anaphoric expressions such as pronouns (e.g., *she, him, their*) and reflexives (e.g., *herself, yourself*) are expressions that are dependent for their interpretation on other material in the context where they are used. Anaphora resolution is governed by a variety of different constraints at the levels of syntax, semantics, pragmatics, and prosody (Barss, 2003; Büring, 2005; Huang, 2004), and these constraints vary across languages and across different anaphoric items within individual languages (Cole, Hermon, & Huang, 2000; Kaiser & Trueswell, 2008; Koster & Reuland, 1992). Consequently, anaphora resolution has proven to be a fertile testing ground for theories of language processing (e.g., Garnham, 2001) and in particular for studies of the time course of constraint application in language comprehension.

The current study investigates how grammatical constraints impact the search for antecedents for pronouns, with a focus on *backwards anaphora* or *cataphora*, a type of pronominal dependency in which a pronoun linearly precedes its antecedent. (We use the term "anaphora" here as a neutral term that covers both forward-looking and backward-looking dependencies.) A number of previous studies have investigated whether grammatical constraints on anaphora remove potential antecedents from consideration during online sentence comprehension. Most studies have found that grammatical constraints do impact anaphora resolution, but there are conflicting findings on exactly when and how ungrammatical antecedents are removed from consideration. Here we examine this issue through the lens of a pair of constraints on backwards anaphora in Russian that apply to superficially similar structures yet differ in a number of respects.

Grammatical constraints on anaphora could, in principle, impact the search for antecedents in at least two ways. Consider a sentence like *John thinks that Bill resents him*, in which the pronoun *him* may corefer with the main clause subject *John* but not with its clause-mate subject *Bill*, due to Principle B of the classical binding theory (Chomsky, 1981) or a more recent counterpart (e.g., Reinhart & Reuland, 1993; Reuland,

2001). The impact of the constraint may be to constrain the generation of candidate antecedents, such that the grammatically inaccessible *Bill* is never even considered as an antecedent for the pronoun. Alternatively the constraint may impact a later processing stage, such that both *John* and *Bill* are initially considered as candidate antecedents, and *Bill* is subsequently filtered out by the constraint. Sturt (2003) describes these alternatives as "early" and "late" filters, respectively, but we prefer to restrict the term "filter" to the latter cases only.) Sturt (2003) raises a third possibility, where grammatically inaccessible antecedents are blocked from consideration at an initial automatic stage, but under some circumstances may be reconsidered at a later stage.

Existing studies on the online impact of binding constraints have yielded conflicting results. In a study using a cross-modal lexical decision task Nicol and Swinney argued that Principles A and B of the classical binding theory act as constraints on the generation of candidate antecedents (Nicol & Swinney, 1989). For example, lexical decision times for words presented immediately after the pronoun *him* in sentences like *The boxer told the swimmer that the doctor for the team would blame him for the recent injury* showed facilitation for semantic associates of the two grammatically accessible antecedents (*boxer, swimmer*) but not for associates of the grammatically inaccessible antecedent (*doctor*). Similar conclusions about the role of constraints on anaphora have been drawn from studies using self-paced reading measures (Clifton, Kennison, & Albrecht, 1997; Lee & Williams, 2006), eye tracking during reading (Sturt, 2003), and event-related brain potentials (Xiang, Dillon, & Phillips, 2009). However, another group of studies has found evidence of consideration of grammatically inaccessible antecedents, leading to the conclusion that constraints on anaphora act as late filters on a broader initial set of candidates. For example, Badecker and Straub's (2002) self-paced reading results suggest interference from grammatically inaccessible antecedents in the processing of the pronoun in "multiple match" sentences like *John thought that Bill owed him another chance to solve*

the problem, where the noun phrase (NP) *Bill* is an impossible antecedent due to Principle B. Similar conclusions have been drawn from studies using self-paced reading (Kennison, 2003) and eye tracking in the visual-world paradigm (Runner, Sussman, & Tanenhaus, 2006).

The discrepant results from previous studies may indicate that the parser uses multiple mechanisms to implement constraints on coreference. However, it is also possible that the discrepant results from previous studies may partly be due to methodological differences across studies. Measures like cross-modal lexical decision (Nicol & Swinney, 1989) provide good evidence of the activation of specific candidate antecedents, but offer limited time course information. Word-by-word reading measures such as self-paced reading (Badecker & Straub, 2002) and eye tracking (Sturt, 2003) provide more detailed time course information but less direct measures of antecedent activation. Even measures such as eye tracking in the visual-world paradigm, which promise continuous data and more transparent measures of referent activation, rely on specific assumptions about the link between gaze patterns and anaphoric dependency formation. Another possible source of discrepancy across different studies may be related to how closely the accessible and inaccessible antecedents were matched in discourse prominence, a property that modulates the impact of potential antecedents on processing (see Badecker & Straub, 2002; Sturt, 2003 for discussion). In the current study we test whether multiple mechanisms are required to implement constraints on coreference by using a within-subjects comparison of different constraints and by turning to a different type of anaphoric relation from what has been considered in most previous studies.

Most previous studies on binding constraints in language processing have focused on cases of forwards anaphora, in which comprehenders encounter a pronoun or reflexive after encoding the potential antecedents in memory. Consequently, forwards anaphora resolution is a retrospective process that potentially requires multiple candidate antecedents to be considered simultaneously. Such situations may be particularly sensitive to the contrasting salience of different antecedents, and this may mask the effects of grammatical constraints. The parser's failure to consider a potential antecedent during online interpretation can only be attributed to a grammatical constraint if we are confident that it is not simply ignored due to low discourse prominence. A straightforward way to avoid this concern is to turn to cases of backwards anaphora/cataphora. Cataphora is less frequent than forwards anaphora, yet both natural and common in English—for example, *When she enters the classroom, Zoe sits down at the art table*. Although some instances of apparent cataphora may turn out to be instances of cross-sentential forwards anaphora, there is ample evidence for true backwards anaphora. For example, van Hoek (1997) documents numerous instances of naturally occurring discourse initial backwards anaphora. From the perspective of online interpretation the resolution of backwards anaphora must proceed differently from the resolution of forwards anaphora. After identifying a cataphoric pronoun a comprehender may consider each subsequent noun phrase as a potential antecedent and evaluate it individually at the time of its maximal salience in the discourse (i.e., when it is the current incoming word or phrase).

In an eye-tracking study van Gompel and Liversedge (2003) investigated the time course of reference resolution between cataphoric pronouns and grammatically accessible antecedents. Using sentences like *When she was fed up, the {girl/boy} visited the {boy/girl} very often*, they found that comprehenders attempt to link a cataphoric pronoun (*she*) to a potential antecedent (*the girl/boy*) before they have confirmed that the expression is compatible with the pronoun in gender. Evidence for this comes from a slowdown in reading times when readers encountered a noun phrase that mismatched in gender with a preceding pronoun. Van Gompel and Liversedge argue that this *gender mismatch effect* reflects an unsuccessful attempt to create a referential dependency between the name and the pronoun, and further that this effect could only have arisen if readers attempt to create referential dependencies before

verifying their semantic well-formedness. Kazanina, Lau, Lieberman, Yoshida, and Phillips (2007) replicated the gender mismatch effect using a self-paced reading paradigm and attributed the effect to the parser's active search for an antecedent. They suggest that when comprehenders encounter a cataphoric pronoun in a clause like *While she was working two jobs to pay the bills* they anticipate the upcoming main clause subject position and construct a referential dependency between that position and the pronoun. If the sentence continues with a compatible main clause subject, such as *Kathryn was taking classes full-time*, then comprehension proceeds smoothly. But if the main clause subject mismatches the pronoun, as in *Russell was taking classes full-time*, then the dependency must be revised, and processing disruption ensues. Kazanina and colleagues regard this mechanism as a counterpart of the active dependency formation mechanism that has been extensively documented for filler-gap dependencies (Frazier & Flores D'Arcais, 1989; Garnsey, Tanenhaus, & Chapman, 1989; Kaan, Harris, Gibson, & Holcomb, 2000; Stowe, 1986). We return in the General Discussion to the issue of the specific mechanisms that underlie the gender mismatch effect.

Kazanina and colleagues show further that the gender mismatch effect disappears in syntactic contexts where backwards anaphora is excluded by Principle C of the binding theory. Principle C rules out coreference between a pronoun and any referring expression that it c-commands (Chomsky, 1981; Reinhart, 1983).[1] They found that when comprehenders encounter a nominal in the ccommand domain of a cataphoric pronoun, as in *She was taking classes full-time while {Kathryn/Russell} was working two jobs to pay the bills*, reading times were unaffected by the gender compatibility between the nominal and

the pronoun. This finding was confirmed across several different types of structures that are subject to Principle C. (For related findings in Japanese see Aoshima, Yoshida, & Phillips, 2009.) This led Kazanina and colleagues to argue that Principle C acts as a constraint on generation in online comprehension (for earlier evidence from a naming study see Cowart & Cairns, 1987). It is unlikely that the structurally inaccessible antecedents were ignored in these studies due to insufficient salience, since the critical data came from reading times when those nominals were in the focus of attention.

The finding that comprehenders apparently fail to even consider anaphoric dependencies that violate Principle C raises the question of why this constraint should exert such a powerful effect in online reference resolution, particularly in light of the inconsistent findings from studies on other constraints on anaphora. At least two possibilities suggest themselves. First, there may be a general advantage for constraints on backwards anaphora, due to the possibility for sequential evaluation of potential antecedents. Second, the structural properties of Principle C may be particularly conducive to exclusion of inappropriate antecedents, since the search mechanism may recognize the irrelevance of an entire structural domain—that is, any domain that is c-commanded by the pronoun—in advance of encountering any bottom-up information about potential antecedents in the input. We explore these questions using Russian, which in addition to Principle C exhibits another Russian-specific constraint on backwards anaphora. As described below, the two constraints apply in superficially similar circumstances yet differ in a number of important respects, such as in when and how potential antecedents can be conclusively excluded from an antecedent search.

---

[1] Apparent exceptions to Principle C have been discussed in the linguistic and psycholinguistic literature—for example, the sentence *He was threatening to leave when Billy noticed that the computer had died* in which the pronoun *he* may be understood as coreferential with *Billy* (Harris & Bates, 2002). Although we believe that these cases may not ultimately be counterexamples to the structural Principle C (see Kazanina, 2005, for more discussion), we took care to avoid such potential exceptions in constructing the experimental materials for the present study and verified the acceptability of all contrasts under investigation by offline questionnaires.

Russian, like English, generally allows backwards anaphora. This is illustrated in Example (1) using sentences in which an initial subordinate clause is introduced by the subordinator *posle togo kak* "after" (see Examples below). Also like English, backwards anaphora in Russian is constrained by Principle C, such that a pronoun cannot c-command its antecedent. Hence in Example (2) the pronoun subject of the main clause cannot be interpreted as coreferential with the subject of the following subordinate clause. Intended coreference in the examples is indicated by the use of subscript indices. Since these two constructions in Russian parallel their English counterparts, we expect that manipulation of the gender of the second clause in each case should elicit the same reading-time effects as those previously observed in English.

The second, Russian-specific constraint on backwards anaphora is exemplified by biclausal sentences such as Example (3) that begin with the subordinator *poka* "while" (henceforth, *poka*-sentences). In (3) coreference between the pronoun *on* "he" and the name *Ivan* is disallowed due to an idiosyncratic constraint on backwards anaphora in Russian (note that in the English translation of (3) coreference between *he* and *Ivan* is perfectly acceptable). The constraint, which we refer to as the poka-constraint for the sake of exposition, is cross-linguistically rare and

appears not to apply even in other Slavic languages that are closely related to Russian, such as Polish or Serbo-Croatian. Critically, even in Russian the scope of the poka-constraint is limited: The constraint applies most strongly in contexts in which both the pronoun subject of the poka-clause and the following main subject are agentive, as in (3). Sentences in which one of the subjects is nonagentive are rather more acceptable on the coreference reading, as exemplified by Example (4) with an experiencer main subject (see also Kazanina & Phillips, 2001). Thus, the Russian poka-constraint can be summarized as follows: In poka-sentences an agentive pronoun subject of an embedded clause cannot corefer with an agentive main clause subject. The constraint also applies in sentences containing other subordinators that encode simultaneity, such as *v to vremja kak* "at the time when".

Note that the poka-constraint also does not block coreference in sentences in which the embedded clause expresses a habitual or a generic event. However, in this paper we do not consider such cases, and we sought to avoid them in our experimental materials. More detailed treatments of the contexts where the poka-constraint applies and theoretical accounts of the phenomenon can be found in other works (Antonyuk & Bailyn, 2008; Avrutin &

Examples (1)–(4):

1. Posle togo kak on$_i$ pročital knigu, Ivan$_i$ s"el jabloko.   [no-constraint]
   after   he read   book Ivan ate apple

   "After he read the book, Ivan ate an apple."

2. *On$_i$ čital knigu, poka   Ivan$_i$ el   jabloko.   [Principle C]
   he   read book   while Ivan ate apple

   "He read a book while Ivan ate an apple."

3. *Poka   on$_i$ čital knigu, Ivan$_i$ s"el jabloko.   [*poka*-constraint]
    while he   read book Ivan ate apple

   "While he read the book Ivan ate an apple."

4. Poka on$_i$ čital knigu, Ivan$_i$ vspomnil,   čto   zabyl kupiť xleba.
   while he   read book Ivan remembered   that forgot buy   bread

   "While he read the book Ivan remembered that he forgot to buy some bread."

Reuland, 2004; Kazanina, 2005; Kazanina & Phillips, 2001; Reuland & Avrutin, 2005). For the purposes of the current study it is sufficient to note that there is agreement that the constraint is a part of the grammar of Russian, but that it is not a purely syntactic restriction and that it implicates discourse representations. From the perspective of online language processing the thematic role restriction on the constraint is particularly interesting. Since the thematic role of a subject noun phrase depends upon predicate information, this restriction implies that the applicability of the poka-constraint depends on information that may not be available until after the pronoun–name sequence. This, together with the nonsyntactic nature of the constraint, raises the question of whether the poka-constraint impacts online reference resolution in the same way that Principle C does, or whether it applies as a late filter on candidate antecedents. In the studies that follow we first verify Russian speakers' awareness of Principle C and the poka-constraint and investigate their application during online processing (Experiment 1). We then go on to test whether the results from the online task were confounded by plausibility biases (Experiment 2).

The predictions for the online reading study are straightforward. First, Russian constructions that have the same coreference possibilities as their English counterparts should exhibit the same processing profile as that found in English studies. This means that in constructions where cataphora is freely allowed, Russian speakers should actively construct a cataphoric dependency upon encountering a potential cataphoric pronoun in the first clause. This in turn should lead to processing disruption (mismatch effects) when the anticipated antecedent turns out to be incompatible with the pronoun in its gender. In Russian constructions that are subject to Principle C readers should avoid predicting an antecedent in grammatically illicit positions. The behavioural reflex of this should be a lack of sensitivity to the gender (mis)match between a pronoun and an illicit antecedent—that is, no mismatch effect. Second, predictions for Russian constructions that are subject to the poka-constraint vary, depending

on the way in which the constraint is implemented online. If all constraints on backwards anaphora are implemented similarly, as constraints that block the generation of candidate anaphoric relations, then we might expect to observe a similar processing profile to Principle C sentences—that is, lack of mismatch effects. Alternatively, there are good reasons to expect that the online impact of the poka-constraint should be different. Since the constraint blocks coreference between pairs of agentive nominals, but the thematic role of a subject typically cannot be determined until verb argument structure information is processed, the constraint may not be triggered until after a potential antecedent has already been interpreted. This may lead to an online profile where readers initially construct a cataphoric dependency that must subsequently be rescinded once verb argument structure confirms that the antecedent is an agent. If rescinding a cataphoric dependency leads to processing disruption, then this could lead to greatest processing difficulty in sentences where the pronoun and the potential antecedent match in gender—that is, a gender match effect, the inverse of the reading time profile predicted for constructions where coreference is fully acceptable. This scenario would be particularly interesting, as it would indicate that two constraints on coreference—Principle C and the poka-constraint—that apply in superficially similar configurations are nevertheless implemented differently online. This in turn would imply that it is not possible to give a uniform answer to the question of whether constraints on coreference are implemented online as constraints on the generation of interpretations or as filters that apply after interpretations have been constructed.

EXPERIMENT 1

Experiment 1 had two parts: an offline rating task and an online self-paced reading task. The offline rating task (Experiment 1A) tested the premise of our study that Russian speakers should reject coreference between a cataphoric pronoun and a main clause subject in Principle C or poka-contexts, but should accept coreference in other backwards

anaphora structures that do not violate any grammatical constraints. The online experiment (Experiment 1B) investigated the process of establishing backwards anaphoric dependencies in real-time processing. In particular, we were interested in whether the different types of grammatical information associated with Principle C and the poka-constraint cause grammatically inaccessible antecedents to be removed from consideration in the same fashion.

Materials and design

The critical experimental items in Experiment 1A were based on the experimental items from Experiment 1B: Three of the six conditions from the online task were used in the offline rating task. We first describe the materials for the online experiment (Experiment 1B) and then explain which conditions were used in the offline task (Experiment 1A).

The online experiment had six conditions, consisting of four conditions (the Principle C and poka conditions) organized in a 2×2 factorial design and two additional no-constraint conditions. A sample set of materials for the online experiment is shown in Table 1, and a full list of items is provided in Appendix A. A total of 24 sets of items were constructed according to a

2×2 design with the factors constraint type (Principle C vs. poka-constraint) and gender congruency (gender match vs. gender mismatch) affecting the relation between the pronoun in the first clause and the subject of a second clause. The biclausal target structures (italicized in Table 1) were all embedded inside an enclosing clause, for reasons explained below. The only difference between the gender-congruent and gender-incongruent variants in each pair was the gender of the subject of the second clause (and the following predicate, due to gender agreement on Russian past tense verbs). The gender of the pronoun was counterbalanced across items. The second subject was always a gender-unambiguous proper name, and the number of Cyrillic characters and syllables in the gender-matching and gender-mismatching names was matched. Materials were designed so as to limit the likelihood of a plausibility bias for or against a coreferential interpretation of the name−pronoun pair, independent of constraints on anaphora. In particular, care was taken to use main and embedded clause events that could plausibly be performed either by different agents or by a single agent. Furthermore, pairs of events were chosen such that the plausibility of the disjoint or coreference interpretation was similar for the Principle C and the poka conditions. This was

Table 1. *Sample set of items used in Experiment 1*

Condition	Russian	English translation
Principle C conditions, gender match/gender mismatch	Poskol'ku pered efirom *ona$_i$ prosmatrivala teksty soobščenij, poka* **Marina/Daniil** *grimmirovalas'/grimmirovalsja k načalu s"emok*, Zoja$_i$ pervoj uznala sensacionnuju novost'.	"Since before the broadcast *she looked through the news texts while Marina/Daniel put on make-up for the shoot*, Zoja was the first one to learn about the sensational news."
Poka conditions, gender match/gender mismatch	Poskol'ku pered efirom *poka ona$_i$ prosmatrivala teksty soobščenij,* **Marina/Daniil** *grimmirovalas'/grimmirovalsja k načalu s"emok*, Zoja$_i$ sama opredelila porjadok reportažej v vypuske.	"Since before the broadcast *while she looked through the news texts Marina/Daniel put on make-up for the shoot*, Zoja figured out the order of the reports in the program by herself."
No constraint, gender match/gender mismatch	Xotja *posle togo kak ona$_i$ napisala zakazannuju stat'ju,* **Nataša$_i$/Mixail** *pravila/pravil tekst neskol'ko raz*, Mixail/Nataša$_i$ bol'še vsego gordilsja svoim pervonačal'nym variantom.	"Although *after she wrote the commissioned article Natasha/Michael edited the text several times*, Michael/Natasha was most proud of the original version."

Note: The critical name in the second subject position is shown in bold. The biclausal target structures are shown in italics. Subscript indices indicate the intended licit backward anaphoric dependencies. The acceptability rating task (Experiment 1A) used only sentences in which the second subject matched the gender of the pronoun.

important in order to exclude the possibility that the reading-time profiles in the two pairs of conditions might differ due to a bias in the stimuli rather than to differences in the online application of the two constraints. Experiment 2 further addresses this issue.

Additionally, the experiment contained 12 sentences from the no-constraint conditions. The no-constraint conditions were licit backwards anaphora sentences such as (1) that contained an embedded clause introduced by the subordinators *do togo kak* "before" or *posle togo kak* "after". These embedded clauses linearly preceded the main clause and were therefore structurally parallel to the poka conditions. Similar to the other two conditions, the gender of the second subject NP was varied to obtain gender-congruent and gender-incongruent variants. As in the Principle C and poka-constraint conditions the gender of the pronoun was counterbalanced across items. These structures do not violate any constraints on coreference and were included to provide a baseline measure of how Russian speakers establish licit backwards anaphoric dependencies. In light of the previous findings in English we expected a gender mismatch effect at the second subject position.

The lexical content of the no-constraint conditions differed from the Principle C and poka conditions, because of the differing semantic conditions on the use of the subordinators used. The no-constraint conditions used the subordinators "before" and "after" and thus described events that occurred in succession. The Principle C conditions and the poka-constraint conditions used the subordinator "while" and thus described events that overlapped in time. Consequently, a basic semantic feature of the relation between the events in the sentences had to differ in the no-constraint conditions. Similarly, the subordinators "before" and "after" in Russian require different verbal aspectual forms (perfective) than the configurations that trigger the "poka" constraint in sentences with "while" (imperfective), and hence lexical matching of the verbs was not possible. It was considered more important to maximally control the lexical material and

plausibility between the two pairs of constraint conditions, since the primary goal of the study was to compare the impact on online interpretation of the two constraints on cataphora.

In order to ensure that the cataphoric pronoun could be associated with an acceptable antecedent in every sentence, all target structures were embedded in a further clause introduced by the subordinators *xotja*, "although", or *poskol'ku*, "since". The gender of the third-clause subject was chosen such that each sentence had a unique acceptable antecedent for the pronoun. This follows a design strategy that proved successful in previous studies on English (Kazanina et al., 2007). The lexical content of the third clause was varied between the Principle C and the poka conditions within the same set, in order to maximize the overall plausibility of the sentences. In the Principle C and poka conditions the subject of the third clause matched in gender with the pronoun and served as an acceptable antecedent. In the no-constraint conditions the gender of the third-clause subject mismatched the pronoun in the gender-match condition, due to the availability of coreference between the pronoun and the second-clause subject, but the third-clause subject matched the gender of the pronoun in the gender-mismatch condition.

EXPERIMENT 1A: ACCEPTABILITY RATING TASK (OFFLINE)

The offline rating task tested the prediction that Russian speakers should reject coreference between a cataphoric pronoun and a main clause subject in Principle C or poka-constraint contexts, but should accept coreference in other backwards anaphora structures that do not violate any constraints.

Method

A total of 33 native speakers of Russian from Moscow participated in the experiment. Participants saw three-clause sentences in which the pronoun and the second subject were highlighted in bold face and were asked to judge the

acceptability of a coreferential interpretation of the boldfaced pronoun−name pair, using a scale from 1 (impossible) to 5 (absolutely natural). The specific form of the question (translated from Russian) was: *Can the pronoun in bold and the noun in bold refer to the same person?* Only gender-matching variants of each condition were presented. In addition, in the no-constraint condition the gender of the third subject was changed to match the gender of the pronoun. As a result, in all three conditions the pronoun was followed by two gender-matching nouns, thereby ensuring that any differences across conditions in the rating of the critical second subject would not be confounded by the number of other gender-matching antecedents in the sentence.

The experiment was administered in the form of a pen-and-paper questionnaire using a similar methodology to Gordon and Hendrick (1997). A total of 24 sets of Principle C and poka-constraint conditions were distributed among two lists according to a Latin Square design; each list additionally contained the 12 no-constraint sentences. Each participant saw only a third of all experimental items from one of the lists—that is, 4 sentences from each of three experimental conditions (Principle C, poka-constraint, no-constraint) interspersed with 4 control and 4 filler sentences. The controls were sentences such as (5) with licit backwards anaphora. The filler sentences were structures with licit forwards anaphora such as (6). In total each questionnaire contained 12 experimental items (four per condition) interspersed with 8 fillers.

5. Backwards anaphora (BA) control

    V naznačennyj den' Astaxov, kotoryj obeščal **emu** pomoč' v ustanovke novogo programmnogo obespečenija, tak i ne pojavilsja v novom ofisnom centre **Andreja Isaeva**.

    "On the due date Astakhov, who had promised to help **him** in installing the software still didn't show up at **Andrej Isaev**'s new office.

6. Forwards anaphora (FA) filler

    Nesmotrja na to, čto **Irina** nikogda ne sobiralas' stat' balerinoj, vysokaja ocenka Ekateriny pol'stila **ej** i byla ochen' prijatnoj.

"Although **Irina** was never going to become a ballerina, the high praise from Ekaterina flattered **her** and was very pleasant."

Results and discussion

Fillers with forwards anaphora received an average rating of 4.6/5. Mean rating scores for experimental conditions and the control are given in Table 2. Mean ratings were very low in the Principle C condition and the poka condition and much higher in no-constraint and the backwards anaphora control conditions. A linear mixed-effects model with participants and items as random factors and condition as a fixed factor revealed a strong effect of condition, as witnessed by the fact that the model with the factor *condition* was superior to a simpler model that did not include that factor: $\chi^2(3) = 76.1$, $p < .001$. Planned pairwise comparisons showed that the scores in the no-constraint condition did not differ significantly from those in the backwards anaphora control, $t(262) = -1.1$, $p = .257$, whereas it was rated significantly higher than the Principle C condition, $t(262) = -13.0$, $p < .001$, and the poka condition, $t(262) = -8.8$, $p < .001$. The difference in ratings between the Principle C and the poka condition was significant, $t(262) = -2.8$, $p = .005$, despite the fact that both conditions received consistently low ratings. The slightly higher ratings for the poka-constraint sentences may reflect their structural similarity to the far more acceptable sentences in the no-constraint conditions. (However, the small acceptability difference should not itself

Table 2. *Mean acceptability rating scores by condition in Experiment 1A*

Condition	Score	
	M	*SE*
Principle C	1.5	0.10
Poka-constraint	1.9	0.11
No constraint	3.6	0.12
BA control	3.9	0.11

Note: BA = backwards anaphora.

lead to an expectation for strongly divergent online effects of the two constraints.)

In sum, the acceptability ratings confirm the prediction that Russian speakers should disallow coreference in sentences that are subject to Principle C and the poka-constraint, whereas they accept backwards anaphora in the no-constraint sentences.

EXPERIMENT 1B: SELF-PACED READING EXPERIMENT

Method

Participants

A new group of 48 Russian speakers (ages 18–28 years) was recruited in Moscow. All participants had normal or corrected-to-normal vision and no history of language disorders. They gave informed consent and were paid the equivalent of $10 for their participation.

Procedure

A total of 24 sets of items consisting of a gender-match and a gender-mismatch variant of the Principle C conditions and of the poka-constraint condition (i.e., four conditions per set) were distributed among four presentation lists in a Latin square design. A total of 12 items from the no-constraint conditions were added to each presentation list so that each list contained one variant of each item (yielding six gender-match and six gender-mismatch variants per list). Each list also contained 84 filler sentences that varied in length and complexity and were superficially similar to the target items—for example, they contained proper names or started with a sub-ordinator. Participants were randomly assigned to one of the lists, and the order of the stimuli within a presentation list was randomized for each participant.

Participants were tested using a Windows laptop computer running the Linger software (developed by Doug Rohde, MIT). Sentences were presented in a standard noncumulative word-by-word moving window paradigm (Just,

Carpenter, & Woolley, 1982) using the font Courier New Cyrillic 20. Each sentence initially appeared on the screen masked by dashes that covered all letters and punctuation marks, but left spaces between words unmasked. Each time the participant pressed the spacebar a new word appeared on the screen, and the previous word was remasked by dashes. Care was taken to ensure that all words from the critical second clause appeared on the same line of the display (the first line). A yes/no comprehension question appeared after each sentence in a single display. Comprehension questions were designed so that they were equally likely to question different parts of the sentence (i.e., first, second, or third clause). In order to conceal the main manipulation of the study and reduce the risk of conscious reading strategies the comprehension questions never targeted the pronoun interpretation. Participants were instructed to read sentences at a natural pace and to respond to the comprehension questions as accurately as possible. If the question was answered incorrectly the word "Incorrect" briefly appeared in the centre of the screen. The testing session lasted approximately 40 minutes.

Data treatment

In order to ensure sufficient numbers of data points in each cell of the analysis, participants were excluded from the final analyses if they failed to correctly answer at least half of the comprehension questions in any individual experimental condition, even if they had a high overall accuracy rate. This led to the exclusion of two participants. The final analyses were based on the data from 46 participants. Only sentences for which the comprehension question was answered correctly were included in the analyses of reading times. Furthermore, reading times that exceeded a threshold of 2,500 ms were replaced by the threshold value. This winsorizing procedure affected 0.4% of data points in the Principle C conditions and in the poka-constraint conditions and 0.3% of data points in the no-constraint conditions.

Reading times were analysed in regions that corresponded to a single word, with the exception of the last region in each clause, which represented the average reading time per word for all remaining words in that clause. Reading times were statistically analysed by fitting a linear mixed effect model using the *lmer* function from the *lme4* package in R (Version 2.6.2; CRAN project; R Development Core Team, 2009. Unlike more traditional analyses of variance (ANOVAs), mixed effects models take unaveraged data as input and make it possible to incorporate random effects of both participants and items within a single analysis (for more information on the use of mixed effects models in psycholinguistics see Baayen, 2008). Models were fitted using a restricted maximum likelihood technique. For the data from the poka-constraint and Principle C conditions the model fitting proceeded as follows: Initially a model that only included the random factors (participants and items) was applied. This initial model was next enriched by adding the first fixed factor constraint (Principle C vs. poka-constraint) and subsequently by including the other fixed factor gender congruency (gender match vs. gender mismatch). Finally the interaction Constraint × Gender Congruency was added to the model. Each successive pair of models was evaluated to assess whether the additional factor improved the model fit to the data. The most complex model that significantly improved the fit over the previous model is considered to be the best fitting model, and its estimates are reported below. In all cases where an interaction was significant we report pairwise comparisons and the relevant 95% confidence interval derived by Markov Chain Monte Carlo simulation (Baayen, Davidson, & Bates, 2008). The model-fitting procedure was identical for the no-constraint conditions, with the exception that only one fixed effect (gender congruency) was included. Model fitting was performed for each region in the sentence individually. For the analyses of the binary variable "accuracy (correct/incorrect)" a binomial family was used.

Since exclusion of 2 participants led to an unbalanced number of participants across item lists an additional set of models were tested that included item list as an additional factor. These analyses showed no impact of the list factor, and hence that factor is not considered further. All significant main effects and interactions with $p < .05$ are reported.

Results

Comprehension question accuracy

The mean comprehension question response accuracy was 88.6% in experimental items and 93.2% in filler items. Accuracy rates for individual experimental conditions are given in Table 3.

In the no-constraint conditions there was a marginally significant effect of gender congruency ($z = 1.904$, $p = .057$) due to a higher mean accuracy rate in the gender-match condition than in the gender-mismatch condition. The effect of gender congruency was also significant in the Principle C and poka-constraint conditions ($z = -2.9$, $p = .003$) but the direction of the effect was reversed: A higher mean accuracy rate was found in the gender-mismatch condition than in the gender-match condition (91.3% vs. 86.2%). There also was a marginally significant effect of constraint type ($z = 1.8$, $p = .068$) due to higher accuracy rates in the Principle C conditions than in the poka-constraint conditions (90.2% vs. 87.3%).

Table 3. *Mean comprehension question accuracy rates by condition from Experiment 1B*

Condition	Accuracy	
	M	SE
No constraint		
Gender match	92.8	1.6
Gender mismatch	88.0	2.0
Principle C		
Gender match	88.0	2.0
Gender mismatch	92.4	1.6
Poka		
Gender match	84.4	2.2
Gender mismatch	90.2	1.8

Reading times (RTs)

The results from the no-constraint conditions are reported first, followed by the results from the Principle C and poka conditions. The region means and information on the best fitting model for each region are summarized in Appendix B.

No-constraint conditions

Figure 1 shows word-by-word reading times in the no-constraint conditions. Mean reading-time differences were small throughout the first clause (Regions 1–8), at the critical second subject (Region 9: *Natasha/Michael*) and in the following region. However, in Region 11, two words after the critical second subject NP, the mean reading time in the gender-incongruent sentences was 65 ms longer than that in the gender-congruent sentences (gender congruent, 542 ms; gender incongruent, 607 ms). This difference led to a marginally significant main effect of gender congruency (95% confidence interval, CI $= \pm 50$ ms), $\beta = -48.3$, $t(497) = -1.86$, $p = .057$. The same numerical trend for longer reading times in the gender-incongruent condition continued at the following two regions (Regions 12 and 13), but reached significance only in Region 13 (gender

congruent, 542 ms; gender incongruent, 614 ms; mean difference, −72 ms; 95% CI $= \pm 55$ ms), $\beta = -57.6$, $t(497) = -2.06$, $p = .040$. There were no other significant effects in the remainder of the sentence. Thus, reading times in the no-constraint conditions replicate the gender-mismatch effect observed in studies of cataphora processing in English (Kazanina et al., 2007; van Gompel & Liversedge, 2003).

Principle C and poka-constraint conditions

Figure 2 shows the results from the Principle C conditions, and Figure 3 shows the results from the poka-constraint conditions.

In Region 2 there was an unexpected significant effect of constraint type, $\beta = 30.3$, $t(976) = 2.1$, $p = .030$, and a marginally significant Constraint $\times$ Gender Congruency interaction, $\beta = -36.3$, $t(976) = -1.8$, $p = .069$. This effect must be spurious, since the conditions were identical up to this point, and it should be noted that the effect corresponded to a small difference in mean reading times. At Region 6 of the first clause there was also a marginally significant effect of constraint type, $\beta = 57.5$, $t(976) = 1.94$, $p = .057$, which was most likely due to the
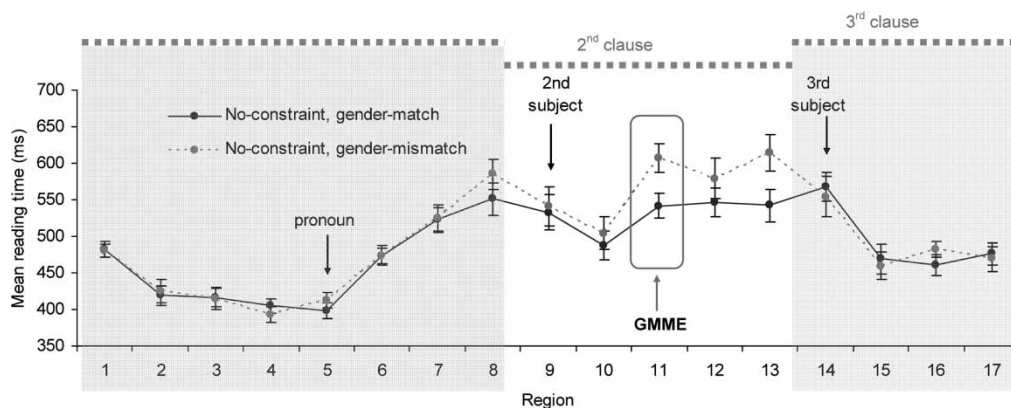
Xotja₁ / posle₂ / togo₃ / kak₄ / ona₅ / napisala₆ / zakazannuju₇ / stat'ju v gazetu,₈ / Nataša (Mixail)₉ / pravila (pravil)₁₀ / tekst₁₁ / neskol'ko₁₂ / raz, ₁₃ / Mixail (Nataša)₁₄ / bol'še₁₅ / vsego₁₆ / gordilsja (gordilas') svoim pervonačal'nym variantom.₁₇

"Although after she₅ wrote the commissioned article Natasha/Michael₉ edited the text several times, Michael/Natasha₁₄ was most proud of the original version."

Figure 1. Mean reading times (in ms) from the no-constraint conditions in Experiment 1B. A gender mismatch effect (GMME) that appeared two words after the critical subject noun phrase (NP) is marked by a box.
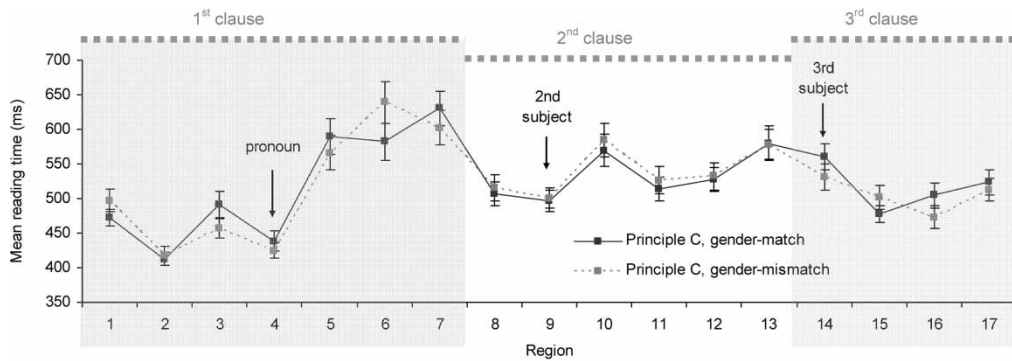
Poskol'ku$_1$ / pered$_2$ / efirom$_3$ / ona$_4$ / prosmatrivala$_5$ / teksty$_6$ / soobščenij,$_7$ / poka$_8$ / Marina (Daniil)$_9$ / grimmirovalas' (grimmirovalsja)$_{10}$ / k$_{11}$ / načalu$_{12}$ / s''emok,$_{13}$ / Zoja$_{14}$ / pervoj$_{15}$ / uznala$_{16}$ / sensacionnuju novost'.$_{17}$

'Since before the broadcast she$_4$ looked through the news texts while Marina (Daniel)$_9$ put on make-up for the shoot, Zoja$_{14}$ was the first one to learn about the sensational news.'

Figure 2. Mean reading times (in ms) from the Principle C conditions in Experiment 1B.

presence of an extra word in the poka-constraint conditions compared to the Principle C conditions. There were no other significant effects or interactions at any region in the first clause.

There was a main effect of constraint type at the critical second subject in Region 9 due to longer mean reading times in the poka conditions than in the Principle C conditions (mean reading times 555 ms vs. 499 ms, respectively), $\beta = -59.4$, $t(978) = -4.04$, $p < .001$. This effect was likely caused by differences in the lexical material in the immediately preceding region. The subordinator while immediately preceded this region in the Principle C conditions and may have provided an effective cue for the appearance of a subject NP in Region 9. There were no significant effects or interactions in Regions 10–11. However, in Region 12, three words after the critical second subject, there was a significant effect of gender congruency, $\beta = 57.8$, $t(960) = 2.48$, $p = .015$, and a marginally significant Constraint × Gender Congruency interaction, $\beta = -61.7$, $t(960) = -1.89$, $p = .057$. Pairwise comparisons within each level of the constraint factor revealed a significant effect of congruency in Region 12 of the poka conditions due to longer mean reading times when the second subject matched in gender with the preceding pronoun (gender congruent, 581 ms; gender incongruent, 523 ms; mean difference = 58 ms;

95% CI = ± 48 ms), $\beta = 58.1$, $t(473) = -2.38$, $p = .018$. This is a gender match effect, the inverse of the effect found in the no-constraint conditions. In the Principle C conditions, on the other hand, reading times were almost identical at both levels of the gender congruency factor (gender congruent, 528 ms; gender incongruent, 533 ms; mean difference = −5 ms; 95% CI = ± 42 ms), $\beta = -2.0$, $t(487) = -0.09$, $p = .957$. There were no significant effects in Region 13.

At the subject of the third clause in Region 14 there was a main effect of gender congruency, $\beta = 45.9$, $t(977) = 2.85$, $p = .006$. In both the Principle C conditions and the poka-constraint conditions mean reading times for the third subject NP were slower in conditions where the earlier second subject NP had matched the gender of the pronoun. Recall that in all four of these conditions the third subject NP provided a grammatically accessible antecedent for the pronoun. In Region 15 there was a main effect of congruency, $\beta = 52.9$, $t(976) = 2.76$, $p = .006$, and a significant Constraint × Gender Congruency interaction, $\beta = -74.3$, $t(976) = -2.76$, $p = .005$. Resolution of this interaction revealed that the effect of congruency was significant in the poka conditions, due to longer mean reading times in the gender-congruent condition than in the gender-incongruent condition (527 vs. 473 ms, respectively;
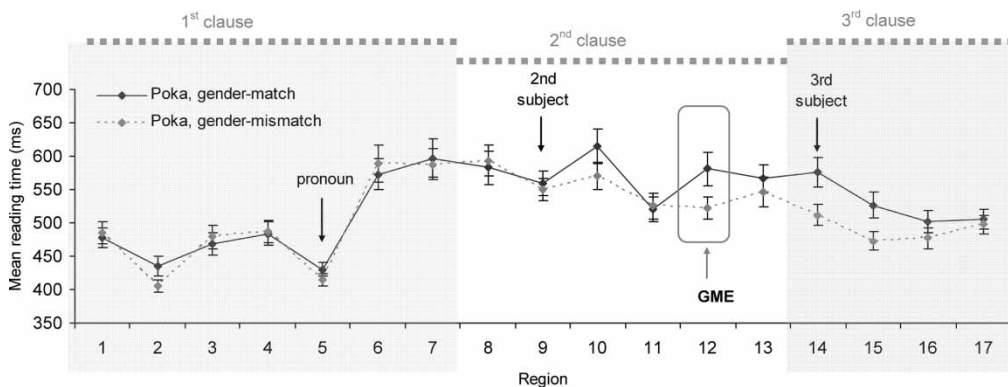
Poskol'ku$_1$ / pered$_2$ / efirom$_3$ / poka$_4$ / ona$_5$ / prosmatrivala$_6$ / teksty$_7$ / soobščenij,$_8$ / Marina (Daniil)$_9$ / grimmirovalas' (grimmirovalsja)$_{10}$ / k$_{11}$ / načalu$_{12}$ / s"emok,$_{13}$ / Zoja$_{14}$ / sama$_{15}$ / opredelila$_{16}$ / porjadok reportažej v vypuske.$_{17}$
'Since before the broadcast while she$_4$ looked through the news texts Marina (Daniel)$_9$ put on make-up for the shoot, Zoja$_{14}$ figured out the order of the reports in the program by herself.'

Figure 3. *Mean reading times (in ms) from the poka conditions in Experiment 1B. A gender match effect (GME) that started three words after the critical subject noun phrase (NP) is marked by a box.*

mean difference = 54 ms; 95% CI = ± 39 ms), $\beta = 51.8$, $t(480) = 2.59$, $p = .013$, but not in the Principle C conditions (gender congruent – 477 ms; gender incongruent – 502 ms; mean difference = –25 ms; 95% CI = ± 33 ms), $\beta = -14.7$, $t(496) = -0.86$, $p = .397$. There were no significant or marginally significant main effects or interactions in the remainder of the sentence (Regions 16–17).

Discussion

In Experiment 1B we found that manipulation of the gender of the subject of the second clause led to a gender mismatch effect in the no-constraint conditions and had no impact on reading times in the Principle C conditions. These results from Russian replicate earlier findings based on similar English sentences (Kazanina et al., 2007; van Gompel & Liversedge, 2003). Additionally, we found a gender match effect at the same region in the poka-constraint conditions, which are subject to a Russian-specific grammatical constraint. This gender match effect is the inverse of the reading-time pattern observed in the no-constraint conditions.

It should be noted that the comprehension accuracy results were largely consistent with the reading-time results. In the no-constraint

conditions the gender-matching condition yielded a higher accuracy rate than its gender-mismatching counterpart, whereas the pattern was reversed in the other two pairs of conditions. This is likely because in the no-constraint condition the gender-matching condition was the easier sentence to process: The parser actively anticipated an antecedent in the second subject position, and this expectation was fulfilled. In the Principle C and the poka conditions, on the other hand, the second subject was easier to rule out in the gender-incongruent condition when it was "doubly" illicit, due to the constraint and to the mismatch in gender with the pronoun.

It is interesting that the Principle C and the poka-constraint conditions yielded different reading-time profiles, despite the fact that both structures are quite similar on the surface and yielded similarly low acceptability ratings for backwards anaphora interpretations. Before interpreting these results, however, it is important to ensure that the differences in the processing of the two pairs of conditions were indeed due to the different constraints on coreference, rather than to artefacts of the stimuli that we used. The sentences in the Principle C and poka conditions differed only in the position of the subordinator *while*. Ideally, this manipulation should have had no other effect

beyond changing the sentence structure so that the sentence invoked either Principle C or the *poka*-constraint. However, this structural change was accompanied by a shift in the figure/ground relation of the two events (E1 and E2) described in the sentence, as illustrated in (7):

7. Principle C: *E1, while E2*
                     [E1 = figure, E2 = ground]

    poka condition: *While E1, E2*
                     [E1 = ground, E2 = figure]

The subordinator *while* establishes a relation in which the embedded event serves as a ground for the main event, which is a figure (also known as fore-grounding/backgrounding, e.g., Matthiessen & Thompson, 1988). Thus, in addition to the structural differences between the conditions in (7), the sentences also differed in which of the events E1 and E2 served as figure or ground. This change of figure/ground relations could be important, because many pairs of events do not have the property that either event can naturally serve as figure while the other takes the role of ground in a *while*-clause. A pair of events that meet this criterion is illustrated in (8), whereas the events in (9) do not.

8. E1 = reading a letter, E2 = eating an apple

    a. Jane read a letter while Bill ate an apple.
                     [E1 figure, E2 ground]

    b. While Jane read a letter, Bill ate an apple.
                     [E1 ground, E2 figure]

9. E1 = breaking a glass, E2 = eating an apple

    a. Jane broke a glass while Bill ate an apple.
                     [E1 figure, E2 ground]

    b. While Jane broke a glass, Bill ate an apple.
                     [E1 ground, E2 figure]

All stimuli used in Experiment 1B were constructed using symmetric event pairs like the pair in (8). However, it is important to also ensure that the pair of events preserves its figure/ground symmetry when the two events are performed by the same agent. This is important, since a semantic bias against a coreferential interpretation of the two subject NPs might lead to spurious evidence for the effects of a linguistic constraint. The examples in

(10) and (11) illustrate an event pair that shows a figure/ground symmetry when the events have different agents, but which is less clearly symmetric when the events are performed by the same agent.

10. E1 = feeling dizzy, E2 = cleaning the floor

    a. Jane was feeling dizzy while Bill was cleaning the floor.
                     [E1 figure, E2 ground]

    b. While Jane was feeling dizzy, Bill was cleaning the floor.
                     [E1 ground, E2 figure]

11. E1 = feeling dizzy, E2 = cleaning the floor

    a. $Jane_i$ was feeling dizzy while she_i was cleaning the floor.
                     [E1 figure, E2 ground]

    b. While $Jane_i$ was feeling dizzy, she_i was cleaning the floor.
                     [E1 ground, E2 figure]

Example 10 shows that the pair of events E1 and E2 is symmetric. When events E1 and E2 are performed by different agents, either of them can serve as a semantically plausible background for the other. Hence, (10a) and (10b) are equally plausible. However, the symmetry disappears if the events are performed by the same person, as in (11). (11a) is quite natural on the coreference reading, and it suggests that E1 and E2 can, in principle, be performed simultaneously by the same agent. However, the coreference reading in (11b) is less plausible, for the reason that *feeling dizzy* is a not an ideal setting for an event of *cleaning the floor* in a setting where these events describe a simultaneous activity by the same person.

The existence of contrasts like the one in (10)–(11) raises the question of whether the materials in Experiment 1 might have contained similar biases. We therefore sought to test whether the pairs of events described in each set were such that switching their figure/ground relation would not introduce a semantic bias towards a coreference or disjoint interpretation, as in (11). If both figure/ground combinations of the events were equally plausible under a coreference interpretation, we could more confidently attribute the processing differences between the Principle C and the *poka* conditions

to the fact that they invoke different constraints on coreference. Experiment 2 addresses this issue.

EXPERIMENT 2

Experiment 2 was a pencil-and-paper questionnaire. Its aim was to test whether the pairs of events used in the materials for the Principle C and poka-constraint conditions in Experiment 1 were balanced and whether inverting their figure/ground configuration would create a bias for a coreferential or disjoint interpretation of the two subject NPs. If there were independent reasons to favour or disfavour a coreferential interpretation of the test sentences, then it would be inappropriate to attribute the reading-time profiles observed in Experiment 1B to effects of grammatical constraints. Additionally, we wanted to ensure that all sentences used in Experiment 1 were amenable to a coreference interpretation in all respects except for the effect of constraints on coreference.

Method

Participants
Experiment 2 was administered to 24 native speakers of Russian residing in Moscow (10 males, 14 females; mean age 30.8 years, range 18–45 years).

Materials and design
All gender-matching variants of the Principle C and poka-constraint conditions from Experiment

1 were used in Experiment 2, yielding 24 sets of two conditions. The sentence-initial subordinators *although* or *since* and the enclosing third clause were discarded. Furthermore, the order of the pronoun and the name was reversed to obtain forwards anaphora, as illustrated in Table 4. Consequently, stimuli from the Principle C conditions were transformed into forwards anaphora sentences in which the main clause preceded the embedded clause ("forwards anaphora, main first"), whereas stimuli from the poka conditions were transformed into sentences with forwards anaphora in which the embedded clause preceded the main clause ("forwards anaphora, embedded first"). Changing the pronoun–name order used in Experiment 1 to forwards anaphora sentences with a name–pronoun order rendered Principle C and the poka-constraint irrelevant and thus made it possible to identify any potential effect of the figure/ground relation of the events on sentence interpretation.

Participants rated the plausibility of each sentence using a scale from 1 (impossible) to 5 (absolutely natural). The critical name and pronoun were highlighted in bold, and participants were explicitly told that they referred to the same person. A total of 24 pairs of target items were distributed among two lists using a Latin square design and were interspersed with 12 control sentences, half of which were highly plausible ("control, highly plausible") or highly implausible ("control, highly implausible") on the coreference reading. These controls were designed to mask the target sentences, and they served as a measure that the participants understood the task correctly.

Table 4. *Sample set of items from Experiment 2*

	Russian	*English translation*
Forwards anaphora, main first	**Marina** prosmatrivala teksty soobščenij, poka **ona** grimmirovalas' k načalu s"emok.	"**Marina** looked through the news texts while **she** put on make up for the shoot."
Forwards anaphora, embedded first	Poka **Marina** prosmatrivala teksty soobščenij, **ona** grimmirovalas' k načalu s"emok.	"While **Marina** looked through the news texts **she** put on make up for the shoot."
Control, highly plausible	Tak kak za poslednie tri goda **ona** ni razu ne brala otpuska, **Olesja** tverdo rešila, čto v etom godu uedet otdyxat' na more.	"Since in the last three years **she** had never taken time off, **Olesja** firmly resolved to go to a seaside resort."
Control, highly implausible	Poka **Inna** naxodilas' po bedro v gipse, **ona** bez truda begala po lestnicam.	"While **Inna's** leg was in a cast, **she** could easily climb stairs."

Note: The participants were explicitly instructed to consider the name and the pronoun in bold as referring to the same person.

Table 5. *Mean plausibility ratings in Experiment 2*

Condition	M	SE
Forwards anaphora, main first	3.2	0.15
Forwards anaphora, embedded first	3.1	0.15
Control, highly plausible	4.5	0.10
Control, highly implausible	1.3	0.10

Results and discussion

The results of the rating test are summarized in Table 5.

As expected, the plausible control sentences received the highest coreference score, and the implausible controls received the lowest score. The forwards-anaphora conditions yielded mean ratings in the middle of the scoring range, suggesting intermediate plausibility of these sentences. Linear mixed-effects modelling revealed a significant effect of condition as a whole, $F(3, 860) = 24$, $p < .0001$. Importantly, a planned comparison of ratings in the two forwards anaphora conditions shows that the ratings for the main-first and the embedded-first forwards anaphora conditions did not significantly differ, $t(574) = 1.45$, $p = .149$. Experiment 2 therefore confirmed that the stimuli used in the Principle C and poka conditions were balanced in terms of the figure/ground relation between the main and embedded events. There was no bias towards higher or lower plausibility of the coreference reading associated with reversal of the figure/ground relation of the events in the two conditions. We can therefore be more confident that the different reading-time profiles observed in Experiment 1B reflect differences in the constraints that apply to each condition.

GENERAL DISCUSSION

We investigated the processing of backwards anaphora in Russian with the goal of understanding whether grammatical constraints impact the online search for antecedents for pronouns in a uniform manner. A self-paced reading experiment (Experiment 1B) investigated the processing of three types of structures containing a cataphoric pronoun in the first clause. In the no-constraint conditions the subject of the second clause was a licit antecedent for the pronoun. In the remaining conditions the second subject was an illicit antecedent due to a constraint on coreference, either Principle C or the Russian-specific poka-constraint. Following a strategy used in previous studies of cataphora processing in English, the gender of the NP in the second subject position was manipulated such that it either matched or mismatched the gender of the cataphoric pronoun. The logic of this design is that in any condition where a cataphoric dependency is attempted we should observe an effect of gender congruency at or shortly after the critical NP. The parser's behaviour in the regions following the critical second subject was different in each of the three condition pairs. In the no-constraint conditions the reading times following the second subject NP were longer when that NP mismatched in gender with the anaphoric pronoun (a gender mismatch effect). There was no effect of gender congruency on the second subject in the Principle C conditions. Finally, in the Russian-specific poka-constraint conditions reading times were longer when the second subject matched in gender with the preceding pronoun (gender match effect). Experiment 2 confirmed that these results are unlikely to be due to a plausibility bias in the pairs of events chosen for Principle C and poka-constraint conditions.

The gender mismatch effect found in the no-constraint condition in Russian replicates earlier findings by van Gompel and Liversedge (2003) and by Kazanina et al. (2007) for similar cases in English, and it suggests that across languages the parser uses a similar mechanism to search for an antecedent for a cataphoric pronoun. The gender-matching no-constraint conditions feature a dependency between a pronoun and a linearly following antecedent. In the absence of a previously mentioned discourse referent the pronoun is temporarily left without a referent, and the parser initiates a search for a licit antecedent in the upcoming input. The gender mismatch effect in the no-constraint conditions reflects the

parser's surprise upon encountering a gender-incongruent nominal in a position where an antecedent was expected. Importantly, as van Gompel and Liversedge point out, the existence of the gender mismatch effect implies that the parser is not hesitant in forming cataphoric dependencies. In order to be disrupted upon recognizing that a noun phrase in the input is gender incongruent, the parser must have initiated dependency formation prior to that moment.

This observation about the earliness of cataphoric dependency formation could reflect either of two mechanisms. Van Gompel and Liversedge (2003) suggest that it is a consequence of an architectural property of the parser that makes word category information about incoming words available before more fine-grained semantic information. In contrast, Kazanina and colleagues regard early formation of cataphoric dependencies as a counterpart of "active" dependency formation found in the processing of filler-gap dependencies (Crain & Fodor, 1985; Frazier & Flores d'Arcais, 1989; Garnsey et al., 1989; Stowe, 1986; Traxler & Pickering, 1996). They propose that dependency formation may be initiated as soon as a cataphoric pronoun is encountered and before a candidate antecedent appears in the input. A mechanism of this kind presupposes the possibility of predictive structure building, but it also avoids the need to assume that an architectural constraint makes syntactic category information available more quickly than other properties of incoming words. Both of these accounts of the gender mismatch effect are compatible with the findings presented here. Moreover, we contend that the two mechanisms are in fact rather similar, since the mechanism proposed by van Gompel and Liversedge must be active to some degree. Even if cataphoric dependency formation is attempted only after a candidate antecedent is encountered in the bottom-up input, the parser must already be in a control state that encodes the fact that it is searching for an antecedent for a pronoun. Without such a mechanism, cataphoric dependency formation could only arise from a massively redundant mechanism that would search backwards for a pronoun every time that it encounters a noun phrase, an

unlikely state of affairs. In addition the control state would need to be sensitive to which structural positions the parser does and does not treat as suitable antecedent positions. This sensitivity is needed to capture the sensitivity of the gender mismatch effect to grammatical constraints on cataphora, as found in this and other studies.

The gender mismatch effect found in the no-constraint conditions also confirms that cataphoric dependencies are acceptable in Russian. If there were a general prohibition against cataphoric dependencies in Russian then we should not have expected to find that Russian speakers actively attempt to construct such dependencies in online processing. Although the gender mismatch effect did not occur immediately upon presentation of the second subject noun, the observed two-word delay is compatible with results from previous studies in English. Van Gompel and Liversedge (2003) found a one-word delay for the mismatch effect in their eye-tracking studies, and our own earlier self-paced reading studies in English showed either an immediate mismatch effect or a one-word delay. Badecker and Straub's self-paced reading studies on forwards anaphora showed a one−two-word delay in the effects of gender congruency (Badecker & Straub, 2002).

Turning to the Principle C conditions, the lack of any effect of the gender manipulation in the second subject in these conditions replicates earlier findings in English (Kazanina et al., 2007). In both languages we take the lack of an effect as evidence that the parser does not consider the second subject as a potential antecedent position for the cataphoric pronoun. This does not mean that the parser fails to actively search for an antecedent in these conditions, only that it does not expect an antecedent in positions that are restricted by Principle C. This apparent blindness to the possibility of coreference in the Principle C conditions cannot easily be dismissed as a consequence of the discourse salience of the second subject NP, since the crucial reading-time measures were taken at a point in time when that NP was in the focus of attention. The Russian findings add cross-linguistic evidence to

earlier conclusions from English and lend further support to the view that Principle C is a cross-linguistically robust constraint on structure generation. In other words, rather than generating cataphoric dependencies that violate Principle C and subsequently rejecting them based on application of a filter, such illicit coreference patterns are never even considered by the parser. More generally, early sensitivity to Principle C can be added to a growing body of evidence that grammatically accurate representations can be built in real time (Crain & Fodor, 1985; Stowe, 1986; Traxler & Pickering, 1996, among others).

Replication of the English findings in the no-constraint conditions and the Principle C conditions provided the background for the most interesting new contribution of the current study. The poka-constraint conditions offered a comparison that has not been possible in previous studies in other languages, as it was possible to compare the reading-time profiles of sentences that are impacted by two different constraints on backwards anaphora. This parallel makes it possible to test whether constraints on coreference apply in a uniform fashion. The Principle C conditions and the poka-constraint conditions had many features in common: The dependent elements were the same pronouns (on, "he", or ona, "she"); the lexical material was identical; the thematic roles of all nominals were the same, and plausibility was matched across conditions; the linear distance between the pronoun and the potential antecedent was almost identical. In addition, the rating study (Experiment 1A) showed that the two constraints led to similarly low acceptability scores. Furthermore, the two constraints were tested using a within-subjects design. Despite these parallels, the reading-time profiles were quite different, with a gender match effect in the poka-constraint conditions and no effect of gender congruency in the Principle C conditions. This implies that the two constraints impact online interpretation in different ways.

A number of previous studies have investigated whether constraints on anaphora impact language processing by blocking the generation of illicit interpretations or by filtering out illicit interpretations after they have been constructed. Most previous studies in this domain have focused on forwards anaphora. Nicol and Swinney argued that constraints on reflexives and pronouns have parallel early impacts on antecedent activation, potentially supporting an interpretation of the constraints as constraints on the generation of interpretations (Nicol & Swinney, 1989). In contrast, Badecker and Straub argued that the same constraints act as later filters on candidate interpretations, although the results from their series of studies were more equivocal (Badecker & Straub, 2002). Runner and colleagues used eye-fixation patterns to argue for differences in the processing of reflexives and pronouns (Runner et al., 2006), but their study focused on special syntactic contexts ("picture noun phrases", such as *Harry's picture of himself*) and ultimately concluded that reflexives in picture noun phrases are exempt from standard grammatical constraints on reflexives. Thus, the current study shows more unequivocally than previous studies that constraints on coreference do not impact language processing in a uniform fashion.

The gender match effect following the critical second subject in the poka conditions implies that participants at least fleetingly considered that subject as a potential antecedent for the preceding pronoun. We interpret the slowdown in the gender-congruent condition as reflecting difficulty that arises when the parser is forced to retract its initial consideration of an analysis in which the name is the antecedent of the preceding pronoun. If the parser were simply blind to the possibility of coreference between the pronoun and the name, then we should have expected to find no gender congruency effect, as found in the Principle C conditions. Thus, it seems hard to avoid the conclusion that the coreference interpretation is at least temporarily considered and subsequently rescinded.

Although the current results provide good evidence that a coreference interpretation is briefly considered and then rescinded in the poka conditions, the results do not provide clear evidence on the question of how actively the cataphoric dependency is constructed in these conditions.

A first possibility is that cataphoric dependency formation proceeded actively in the poka-constraint conditions, just as in the no-constraint conditions, leading to disruption upon encountering a gender-incongruent name (i.e., a gender mismatch effect). If this was the case, then it must be that the gender mismatch effect was overridden in the reading-time profile by the immediately following disruption in the gender-congruent condition, when the poka-constraint is triggered, and the illicit interpretation had to be rescinded. Note that this account entails that the processing disruption caused by retraction of an illicit pronoun–name dependency in the gender-matching poka condition must be stronger than the processing disruption caused by encountering an incongruent pronoun in the gender-mismatching poka condition. This assumption could potentially capture why no gender mismatch effect is observed in these conditions. The assumption is reasonable, in the respect that the gender mismatch effect involves retraction of a prediction, whereas we are claiming that the gender match effect reflects retraction of a fully fledged cataphoric dependency. A second possibility is that the construction of cataphoric dependencies proceeds in a less active fashion when a pronoun is introduced in a poka-clause, and hence that a dependency is only attempted after a gender-matching antecedent has been encountered. The currently available data do not allow us to adequately distinguish between these accounts, but this does not undermine the more important conclusion that a coreference interpretation is fleetingly considered in the poka conditions.

The finding that Principle C acts to block the generation of coreference dependencies, whereas the poka-constraint acts as a filter on dependencies that are initially considered, naturally raises the question of why the two constraints should impact parsing so differently. The current findings are consistent with a number of possible accounts of this difference. It could reflect the fact that Principle C applies to syntactic configurations (Chomsky, 1981) whereas the poka-constraint is a constraint that invokes discourse representations (Avrutin & Reuland, 2004; Kazanina & Phillips,

2001; Reuland & Avrutin, 2005). Alternatively, the difference could reflect the fact that Principle C is a cross-linguistically robust constraint, whereas the poka-constraint is a more idiosyncratic property of Russian. In earlier work we found a developmental dissociation between the two constraints: Russian children have already mastered Principle C at age 3, whereas the poka-constraint does not reliably constrain their interpretations until age 5–6 (Kazanina & Phillips, 2001).

However, we suggest that the most likely cause of the different online effects of the two constraints is the time course of information availability relevant to the two constraints. In the case of Principle C the parser may determine that an entire structural domain is irrelevant to the search for an antecedent, and this can be ascertained in advance of any bottom-up information about potential antecedents. This kind of precomputation is not possible in the case of the poka-constraint. Recall that the poka-constraint applies to pronoun–name sequences in which both nominals have the semantic role of agent. Although agent is by far the most common thematic role for a nominative subject, the thematic role of the subject cannot be confirmed until the predicate of the clause is identified. Therefore, when faced with a sequence $poka\ pronoun_{NOM} \ldots name_{NOM}$ the parser may continue to entertain the possibility that the pronoun and the name may be coreferential, because the thematic role of the name is not yet confirmed. Only when the predicate is reached is it possible to confirm that the name bears an agent role, and at that point the poka-constraint becomes relevant, and the coreference relation must be rescinded, leading to reading-time slowdown. This interpretation is consistent with the time course of the gender match effect, which in Experiment 1B appeared at Region 13, slightly later than the gender mismatch effect in the no-constraint conditions. If this is the appropriate account of why the poka-constraint acts as a filter rather than as a constraint on generation of antecedents then it is interesting to ask why Russian speakers do not immediately assume that a nominative subject is an agent, a statistically likely outcome, and a plausible

consequence of parsing with superficial heuristics (e.g., Towsend & Bever, 2001). However, the likelihood that a nominative subject NP is an agent must be balanced against the likelihood that a *pronoun . . . name* sequence of subject NPs involves a coreference relation. Consequently, more detailed analyses of Russian corpora would be needed in order to determine whether speakers' behaviour in our study genuinely ignores distributional constraints.

CONCLUSION

A number of previous studies have investigated the question of how grammatical constraints impact the search for antecedents for anaphoric expressions in online language processing. Either constraints may act as constraints on the generation of candidate interpretations, which prevent grammatically inaccessible antecedents from ever being considered, or they may act as later filters on interpretations, which serve to remove illicit coreference relations only after they have been initially considered. It is common for discussions of this issue to assume that there is a uniform answer to this question. However, results from our within-subjects comparison of closely matched constraints on backwards anaphora in Russian indicate that a uniform answer may not be possible. Russian speakers appear to implement Principle C as a constraint on generation of interpretations and the poka-constraint as a filter on temporarily considered interpretations. This contrast can be understood in terms of the detailed time course of information that is relevant to application of the two constraints.

REFERENCES

Antonyuk, S., & Bailyn, J. (2008). *Backward pronominalization in Russian: A syntactic account*. Handout from talk presented at the 17th meeting of Formal Approaches to Slavic Linguistics (FASL-17). May, 2008, New Haven, CT, Yale University.

Aoshima, S., Yoshida, M., & Phillips, C. (2009). Incremental processing of coreference and binding in Japanese. *Syntax*, *12*, 93–134.

Avrutin, S., & Reuland, E. (2004). Backward anaphora and tense interpretation. In L. Verbitskaya (Ed.), *Teoreticheskie problemy jazykoznanija* [Theoretical issues in language research]. St. Petersburg, Russia: St. Petersburg University Press.

Baayen, R. H. (2008). *Analyzing linguistic data. A practical introduction to statistics using R*. Cambridge, UK: Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412.

Badecker, W., & Straub, K. (2002). The processing role of structural constraints on the interpretation of pronouns and anaphora. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 748–769.

Barss, A. (2003). *Anaphora: A reference guide*. Malden, MA: Blackwell.

Büring, D. (2005). *Binding theory*. Cambridge, UK: Cambridge University Press.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht, The Netherlands: Foris.

Clifton, C., Kennison, S. M., & Albrecht, J. E. (1997). Reading the words *her*, *his*, *him*: Implications for parsing principles based on frequency and structure. *Journal of Memory and Language*, *36*, 276–292.

Cole, P., Hermon, G., & Huang, C.-T. J. (2000). *Long-distance reflexives*. New York: Academic Press.

Cowart, W., & Cairns, H. S. (1987). Evidence for an anaphoric mechanism within syntactic processing: Some reference relations defy semantic and pragmatic constraints. *Memory & Cognition*, *15*, 318–331.

Crain, S., & Fodor, J. D. (1985). How can grammars help parsers. In D. Dowty, D. Kartunnen, & A. M. Zwicky (Eds.), *Natural language parsing: Psycholinguistic, computational, and theoretical perspectives* (pp. 94–128). Cambridge, UK: Cambridge University Press.

Frazier, L., & Flores d'Arcais, G. B. (1989). Filler-driven parsing: A study of gap filling in Dutch. *Journal of Memory and Language*, *28*, 331–344.

Garnham, A. (2001). *Mental models and the interpretation of anaphora*. Hove, UK: Psychology Press.

Garnsey, S. M., Tanenhaus, M. K., & Chapman, R. M. (1989). Evoked potentials and the study of sentence

comprehension. *Journal of Psycholinguistic Research*, *18*, 51–60.

Gordon, P., & Hendrick, R. (1997). Intuitive knowledge of linguistic co-reference. *Cognition*, *62*, 325–370.

Harris, C. L., & Bates, E. A. (2002). Clausal backgrounding and pronominal reference: A functionalist approach to c-command. *Language and Cognitive Processes*, *17*, 237–269.

Huang, Y. (2004). Anaphora and the pragmatics-syntax interface. In L. R. Horn & G. Ward (Eds.), *The handbook of pragmatics* (pp. 288–314). Oxford, UK: Blackwell.

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General*, *3*, 228–238.

Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, *15*, 159–201.

Kaiser, E., & Trueswell, J. C. (2008). Interpreting pronouns and demonstratives in Finnish: Evidence for a form-specific approach to reference resolution. *Language and Cognitive Processes*, *23*, 709–748.

Kazanina, N. (2005). *The acquisition and processing of backwards anaphora*. Unpublished PhD dissertation, University of Maryland, College Park.

Kazanina, N., Lau, E., Lieberman, M., Yoshida, M., & Phillips, C. (2007). The effect of syntactic constraints on the processing of backwards anaphora. *Journal of Memory and Language*, *56*, 384–409.

Kazanina, N., & Phillips, C. (2001). Coreference in child Russian: Distinguishing syntactic and discourse constraints. In A. Do, L. Dominguez, & A. Johansen (Eds.), *Proceedings of the 25th Annual Boston University Conference on Language Development* (pp. 413–424). Somerville MA: Cascadilla Press.

Kennison, S. M. (2003). Comprehending the pronouns *her*, *him*, and *his*: Implications for theories of referential processing. *Journal of Memory and Language*, *49*, 335–352.

Koster, J., & Reuland, E. (1992). *Long-distance anaphora*. Cambridge, UK: Cambridge University Press.

Lee, M.-W., & Williams, J. N. (2006). *The role of grammatical constraints in intra-sentential pronoun resolution*. Manuscript submitted for publication.

Matthiessen, C., & Thompson, S. A. (1988). The structure of discourse and "subordination". In J. Haiman & S. A. Thompson (Eds.), *Clause combining in grammar and discourse* (pp. 275–329). Amsterdam: John Benjamins.

Nicol, J., & Swinney, D. (1989). The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research*, *18*, 5–19.

R Development Core Team (2009). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Reinhart, T. (1983). *Anaphora and semantic interpretation*. London: Croom Helm.

Reinhart, T., & Reuland, E. (1993). Reflexivity. *Linguistic Inquiry*, *24*, 261–321.

Reuland, E. (2001). Primitives of binding. *Linguistic Inquiry*, *32*, 439–492.

Reuland, E., & Avrutin, S. (2005). Binding and beyond: Issues in backward anaphora. In A. Branco, T. McEnery, & R. Mitkov (Eds.), *Anaphora processing: Linguistic, cognitive and computational modeling* (pp. 139–162). Amsterdam: John Benjamins.

Runner, J. T., Sussman, R. S., & Tanenhaus, M. K. (2006). Processing reflexives and pronouns in picture noun phrases. *Cognitive Science*, *30*, 193–241.

Stowe, L. A. (1986). Parsing WH-constructions: Evidence for on-line gap location. *Language and Cognitive Processes*, *3*, 227–245.

Sturt, P. (2003). The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language*, *48*, 542–562.

Townsend, D. J., & Bever, T. G. (2001). *Sentence comprehension: The integration of habits and rules*. Cambridge, MA: MIT Press.

Traxler, M. J., & Pickering, M. J. (1996). Plausibility and the processing of unbounded dependencies: An eye-tracking study. *Journal of Memory and Language*, *35*, 454–475.

van Gompel, R. P. G., & Liversedge, S. P. (2003). The influence of morphological information on cataphoric pronoun assignment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 128–139.

van Hoek, K. (1997). *Anaphora and conceptual structure*. Chicago: University of Chicago Press.

Xiang, M., Dillon, B., & Phillips, C. (2009). Illusory licensing effects across dependency types: ERP evidence. *Brain and Language*, *108*, 40–55.

APPENDIX A

Full list of stimuli (including comprehension questions) from the self-paced reading task, Experiment 1

English translations of the experimental sentences are provided in order to give readers information about the scenarios tested. Note, however, that the translations do not always preserve the Russian word order and exact phrasing, and no attempt is made to convey the detailed aspectual force of the Russian predicates.

The experiment contained 24 sets the Principle C and poka conditions and 12 sets of no-constraint conditions:

Principle C & poka conditions	*No-constraint conditions*
Condition a: Principle C, gender match	Condition a: no-constraint, gender match
Condition b: Principle C, gender mismatch	Condition b: no-constraint, gender mismatch
Condition c: poka-constraint, gender match	
Condition d: poka-constraint, gender mismatch	

Principle C and poka conditions

1a/b. Поскольку перед Рождеством он сбывал с рук драгоценности, пока Марат/Алина пытался/пыталась получить визы в Аргентину, Стас опасался вызвать подозрение у окружающих.
Пытался/пыталась ли Марат/Алина получить визы в Бразилию?
Since before Christmas he got rid of the jewellery while Marat/Alina attempted(m/f) to secure visas to Argentina, Stas was afraid to raise suspicions with acquaintances.

1c/d. Поскольку перед Рождеством, пока он сбывал с рук драгоценности, Марат/Алина пытался/пыталась получить визы в Аргентину, Стас опасался вызвать подозрение у окружающих.
Пытался/пыталась ли Марат/Алина получить визы в Бразилию?
Since before Christmas while he got rid of the jewellery, Marat/Alina attempted(m/f) to secure visas to Argentina, Stas was afraid to raise suspicions with acquaintances.

2a/b. Поскольку уже в полдень он скрупулезно изучал внутреннюю структуру купола, пока Михаил/Тамара делал/делала карандашные зарисовки постройки, архитектор рассчитывал иметь всю необходимую информацию о храме к концу дня.
Надеялся ли архитектор закончить сбор необходимой информации до конца дня?
Since already at noon he scrupulously studied the inner structure of the dome while Michael/Tamara made (m/f) pencil sketches of the construction, the architect counted on having all of the necessary information about the temple by the end of the day.

2c/d. Поскольку уже в полдень, пока он скрупулезно изучал внутреннюю структуру купола, Михаил/Тамара делал/делала карандашные зарисовки постройки, архитектор рассчитывал иметь всю необходимую информацию о храме к концу дня.
Надеялся ли архитектор закончить сбор необходимой информации до конца дня?
Since already at noon while he scrupulously studied the inner structure of the dome, Michael/Tamara made(m/f) pencil sketches of the construction, the architect counted on having all of the necessary information about the temple by the end of the day.

3a/b. Поскольку перед началом спектакля она разговаривала по рации с охраной, пока Раиса/Сергей рассажи вала/рассаживал по местам высокопоставленных гостей, Инна не видела приключившегося с генералом казуса.
Состоялся ли разговор с охраной после спектакля?
Since before the beginning of the play she spoke with security on the portable radio while Raisa/Sergej sat(m/f) VIP guests in their seats, Inna didn't notice the incident that happened to the general.

3c/d. Поскольку перед началом спектакля, пока она разговаривала по рации с охраной, Раиса/Сергей рассажи вала/рассаживал по местам высокопоставленных гостей, Инна была очень благодарна ей/ему за помощь.
Состоялся ли разговор с охраной после спектакля?
Since before the beginning of the play while she spoke with security on the portable radio, Raisa/Sergej sat(m/f) VIP guests in their seats, Inna was very grateful to her/him for help.

4a/b. Поскольку перед эфиром она просматривала тексты сообщений, пока Марина/Даниил гримировалась/гримировался к началу съемок, Зоя первой узнала сенсационную новость. Узнала ли Зоя сенсационную новость до начала эфира?

Since before the broadcast she looked through the news texts while Marina/Danil put(m/f) on make up for the shoot, Zoja was the first one to learn about the sensational news.

4c/d. Поскольку перед эфиром, пока она просматривала тексты сообщений, Марина/Даниил гримировалась/гримировался к началу съемок, Зоя сама определила порядок репортажей в выпуске. Был ли определен порядок новостей Зоей?

Since before the broadcast while she looked through the news texts, Marina/Danil put(m/f) on make up for the shoot, Zoja figured out the order of the reports in the program by herself.

5a/b. Поскольку вчера днем он подключал измерители давления, пока Артур/Клара настраивал/настраивала специальную вакуумную камеру, Виталий был уверен в полной синхронизации установки. Сомневался ли Виталий в синхронизации установки?

Since yesterday during the day he connected the pressure meter while Arthur/Clara tuned(m/f) the special vacuum camera, Valerij was certain that the equipment was fully synchronized.

5c/d. Поскольку вчера днем, пока он подключал измерители давления, Артур/Клара настраивал/настраивала специальную вакуумную камеру, Валерий надеялся наладить оборудование до наступления сумерек. Потерял ли Валерий надежду настроить оборудование до темноты?

Since yesterday during the day while he connected the pressure meter Arthur/Clara tuned(m/f) the special vacuum camera, Valerij hoped to fix the equipment before dusk.

6a/b. Поскольку всю неделю перед парным финалом он беспечно заигрывал с многочисленными поклонницами, пока Всеволод/Кристина отрабатывал/отрабатывала на корте ключевые удары, Дмитрий отыграл в финальном матче значительно хуже своего партнера/своей партнерши. Играл ли Дмитрий в индивидуальном финале?

Since all week before the doubles final he carelessly flirted with numerous fans, while Vsevolod/Kristina worked(m/f) out key strokes on the court, Dmitri played far worse in the final than his partner.

6c/d. Поскольку всю неделю перед парным финалом, пока он беспечно заигрывал с многочисленными поклонницами, Всеволод/Кристина отрабатывал/отрабатывала на корте ключевые удары, Дмитрий не мог не признать вклад своего партнера/своей партнерши решающим в их победе. Играл ли Дмитрий в индивидуальном финале?

Since all week before the doubles final while he carelessly flirted with numerous fans Vsevolod/Kristina worked(m/f) out key strokes on the court, Dmitri had to acknowledge his partner's contribution to their victory.

7a/b. Поскольку после контрольной она просматривала ученические тетради, пока Ольга/Борис вводила/вводил в компьютер статистику по ошибкам, учительница могла сравнивать результаты между классами. Вводила/вводил ли Ольга/Борис статистику по ошибкам в компьютер?

Since before the exam she looked through students' notebooks while Olga/Boris entered(f/m) data about mistakes into the computer, the teacher was able to compare results between classes.

7c/d. Поскольку после контрольной, пока она просматривала ученические тетради, Ольга/Борис вводила/вводил в компьютер статистику по ошибкам, учительница знала подробные результаты сразу по окончании проверки. Вводила/вводил ли Ольга/Борис статистику по ошибкам в компьютер?

Since before the exam while she looked through students' notebooks Olga/Boris entered(f/m) data about mistakes into the computer, the teacher knew the detailed results immediately after she finished checking the papers.

8a/b. Хотя в прошлом году он впервые тестировал новый вид диеты, пока Ярослав/Варвара готовился/готовилась к ответственному этапу чемпионата мира, Глеб без опасений внедрил свои новшества в схему питания спортсмена/спортсменки. Испытывал ли Глеб опасения относительно новой диеты?

Even though in the past year he tested a new type of diet for the first time while Yaroslav/Varvara prepared(m/f) for the crucial round of the world championship, Gleb fearlessly introduced his own innovations into the competitor's diet.

8c/d. Хотя в прошлом году, пока он впервые тестировал новый вид диеты, Ярослав/Варвара готовился/готовилась к ответственному этапу чемпионата мира, Глеб без опасений внедрил свои новшества в схему питания спортсмена/спортсменки. Побоялся ли Глеб испытать новый вид диеты перед чемпионатом мира?

Even though in the past year while he tested a new type of diet for the first time Yaroslav/Varvara prepared(m/f) for the crucial round of the world championship, Gleb fearlessly introduced his own innovations into the competitor's diet.

9a/b. Хотя в январе она готовилась к квалификационным экзаменам, пока Алла/Иван отдыхала/отдыхал на Канарских островах, Вика не сетовала на судьбу.

Отдыхала/отдыхал ли Алла/Иван на Канарах в августе?

Even though in January she prepared for the qualifying exams while Alla/Ivan relaxed(f/m) in the Canary Islands, Vika did not complain about life.

9c/d. Хотя в январе, пока она готовилась к квалификационным экзаменам, Алла/Иван отдыхала/отдыхал на Канарских островах, Вика не сетовала на судьбу.

Отдыхала/отдыхал ли Алла/Иван на Канарах в августе?

Even though in January while she prepared for the qualifying exams Alla/Ivan relaxed(f/m) in the Canary Islands, Vika did not complain about life.

10a/b. Хотя в воскресенье она учила билеты по физике, пока Варя/Вова слушала/слушал прямой радиорепортаж с 'Евровидения', Рита умудрялась не обращать внимания на происходящее.

Транслировали ли 'Евровидение' по радио в прямом эфире?

Even though on Sunday she studied the physics exam questions while Varja/Vova listened (f/m) to radio coverage of Eurovision, Rita managed not to pay attention to what was going on.

10c/d. Хотя в воскресенье, пока она учила билеты по физике, Варя/Вова слушала/слушал прямой радиорепортаж с

'Евровидения', Рита умудрялась не обращать внимания на происходящее.

Транслировали ли 'Евровидение' по радио в прямом эфире?

Even though on Sunday while she studied the physics exam questions Varja/Vova listened (f/m) to radio coverage of Eurovision, Rita managed not to pay attention to what was going on.

11a/b. Хотя почти все утро он искал представителей редкого вида гусеницы, пока Егор/Маша снимал/снимала на камеру наиболее красивые уголки парка, Аркадий так и не нашел нужный ему экземпляр.

Правда ли, что поиски гусеницы велись вечером?

Even though almost all morning he searched for examples of a rare caterpillar species while Egor/Masha photographed(m/f) the more beautiful corners of the park, Arkadi still did not find the specimen he was looking for.

11c/d. Хотя почти все утро, пока он искал экземпляр редкого вида гусеницы, Егор/Маша снимал/снимала на камеру наиболее красивые уголки парка, Аркадий переживал, что мальчик/девушка скучает.

Правда ли, что поиски гусеницы велись вечером?

Even though almost all morning while he searched for examples of a rare caterpillar species Egor/Masha photographed(m/f) the more beautiful corners of the park, Arkadij was worried that the boy/girl was bored.

12a/b. Хотя каждое воскресенье она пила утренний кофе, пока Оля/Юра решала/решал какойнибудь увлекательный кроссворд, Марина никогда не предлагала помочь.

Правда ли, что Марина никогда не предлагала своей помощи в разгадке кроссвордов?

Even though every Sunday she drank morning coffee while Olja/Jura solved(f/m) some fascinating crossword puzzle, Marina never offered help.

12c/d. Хотя каждое воскресенье, пока она пила утренний кофе, Оля/Юра решала/решал решала какой-нибудь увлекательный кроссворд, Марина никогда не принимала участия в разгадке вопросов.

Правда ли, что Марина никогда не принимала участия в разгадке кроссвордов?

Even though every Sunday while she drank morning coffee Olja/Jura solved(f/m) some fascinating crossword puzzle, Marina never helped out with the crossword.

13a/b. Поскольку в магазине он болтал с симпатичными кондитершами, пока Максим/Наташа пробовал/пробовала разные виды пирожных, Дима получил огромное удовольствие от похода за покупками.

Правда ли, что в магазине работали симпатичные кондитерши?

Since in the store he bantered with the pretty bakers while Maksim/Natasha tasted(m/f) different types of pastries, Dima greatly enjoyed their shopping trips.

13c/d. Поскольку в магазине, пока он болтал с симпатичными кондитершами, Максим/Наташа пробовал/пробовала разные виды пирожных, Дима захотел тоже попробовать что-нибудь из сладкого.

Правда ли, что в магазине можно было попробовать выпечку?

Since in the store while he bantered with pretty bakers Maksim/Natasha tasted(m/f) different types of pastries, Dima also expressed a desire to try some pastry.

14a/b. Хотя уже в полшестого она украшала зеленью салаты, пока Галина/Сережа спешно накрывала/накрывал на стол, хозяйка так и не успела закончить приготовления до прихода гостей.

Украшались ли салаты лимоном?

Even though already at half past five she decorated the salads with greens while Galina/Serezha hastily set(f/m) the table, the hostess did not manage to finish the preparations before the guests arrived.

14c/d.   Хотя уже в полшестого, пока она украшала зеленью салаты, Галина/Сережа спешно накрывала/накрывал на стол, хозяйка не надеялась закончить приготовления до прихода гостей.
Украшались ли салаты лимоном?
Even though already at half past five while she decorated the salads with greens Galina/Serezha hastily set(f/m) the table, the hostess did not hope to finish the preparations before the guests arrived.

15a/b.   Хотя всю субботу она полола разбушевавшиеся сорняки, пока Света/Вадим срезала/срезал сухие ветки с кустов, Валентина, в отличие от подруги/мужа, совершенно не чувствовала усталости.
Утомила ли Валентину работа в саду?
Even though all Saturday she dug up the out-of-control weeds while Sveta/Vadim cut(f/m) the dry branches from the bushes, Valentina, unlike her friend(f)/husband, felt no tiredness.

15c/d.   Хотя всю субботу, пока она полола разбушевавшиеся сорняки, Света/Вадим срезала/срезал сухие ветки с кустов, Валентина считала уборку сада полностью своей заслугой.
Была ли Валентина объективна в своей оценке?
Even though all Saturday while she dug up the out-of-control weeds Sveta/Vadim cut(f/m) the dry branches from the bushes, Valentina considered the garden clean-up to be solely her achievement.

16a/b.   Поскольку весь вечер она составляла квартальный отчет, пока Лида/Вася смотрела/смотрел повтор 'Рождественских встреч', Марина считала себя вправе не готовить ужин.
Смотрела/смотрел ли Лида/Вася повтор 'Рождественских встреч'?
Since all evening she prepared the quarterly report while Lida/Vasja watched(f/m) the rerun of the "Christmas meetings", Marina considered herself justified not to prepare dinner.

16c/d.   Поскольку весь вечер, пока она составляла квартальный отчет, Лида/Вася смотрела/смотрел повтор 'Рождественских встреч', Марина не могла полностью сосредоточиться на работе.
Смотрела/смотрел ли Лида/Вася повтор 'Рождественских встреч'?
Since all evening while she prepared the quarterly report, Lida/Vasja watched(f/m) the rerun of the "Christmas meetings", Marina failed to concentrate fully on her work.

17a/b.   Хотя в ресторане она курила на терассе, пока Алина/Борис обсуждала/обсуждал с официантом прелести японской кухни, Галя слышала все детали их разговора.
Обсуждала/обсуждал ли Алина/Борис с официантом итальянскую кухню?
Even though in the restaurant she smoked on the terrace while Alina/Boris discussed(f/m) the charms of Japanese cuisine with the waiter, Galja caught their conversation in detail.

17c/d.   Хотя в ресторане, пока она курила на терассе, Алина/Борис обсуждала/обсуждал с официантом прелести японской кухни, Галя настояла на заказе исключительно европейских блюд.
Обсуждала/обсуждал ли Алина/Борис с официантом итальянскую кухню?
Even though in the restaurant while she smoked on the terrace Alina/Boris discussed(f/m) the charms of Japanese cuisine with the waiter, Galja insisted on an order consisting exclusively of European dishes.

18a/b.   Поскольку в самолете она расшивала блестками костюм, пока Алена/Гриша заучивала/заучивал наизусть слова финальной песни, Жанна старалась не отвлекать дочь/сына своими просьбами и расспросами.
Правда ли, что Жанна старалась не беспокоить дочь/сына?
Since on the plane she embroidered the suit with sequins while Alena/Gregory learned the words of the final song by heart, Zhanna tried not to distract her daughter/son with her requests and questions.

18c/d.   Поскольку в самолете, пока она расшивала блестками костюм, Алена/Гриша заучивала/заучивал наизусть слова финальной песни, Жанна невольно выучила текст вместе с дочерью/сыном.
Выучила ли Жанна текст текст финальной песни?
Since on the plane while she embroidered the suit with sequins Alena/Gregory learned the words of the final song by heart, Zhanna reluctantly leant the lyrics together with her daughter/son.

19a/b.   Поскольку все утро он пылесосил ковры в лоджии, пока Денис/Ксюша обзванивал/обзванивала друзей относительно предстоящей встречи, Семен ни капли не сожалел о решении встретиться у них дома.
Были ли подвергнуты чистке ковры в лоджии?
Since all morning he vacuumed carpets on the balcony while Denis/Ksjusha telephoned(m/f) friends about the upcoming meeting, Simon did not regret at all the decision to meet at their place.

19c/d.   Поскольку все утро, пока он пылесосил ковры в лоджии, Денис/Ксюша обзванивал/обзванивала друзей относительно предстоящей встречи, Семен плотно закрыл все двери и старался не шуметь.
Были ли подвергнуты чистке ковры в лоджии?
Since all morning while he vacuumed carpets on the balcony Denis/Ksjusha telephoned(m/f) friends about the upcoming meeting, Simon tightly closed all the doors and tried not to make any noise.

20a/b. Хотя весь день он выписывал основные цитаты, пока Илья/Лина писал/писала назначенный на четверг доклад, Кирилл не хотел вникать в суть его/её работы.

Вникал ли Кирилл в суть доклада?

Even though all day he wrote out the main citations while Ilya/Lina wrote(m/f) the report that was due on Thursday, Kirill did not want to grasp the essence of his/her work.

20c/d. Хотя весь день, пока он выписывал основные цитаты, Илья/Лина писал/писала назначенный на четверг доклад, Кирилл не хотел вникать в суть его/её работы.

Вникал ли Кирилл в суть доклада?

Even though all day while he wrote out the main citations Ilya/Lina wrote(m/f) the report that was due on Thursday, Kirill did not want to grasp the essence of his/her work.

21a/b. Хотя два дня назад он обрабатывал собранные ранее данные, пока Игорь/Света проводил/проводила повторный контрольный эксперимент, Геннадий не захотел делиться с ним/ней результатами.

Правда ли, что контрольный эксперимент проводился повторно?

Even though two days ago he processed the data collected previously while Igor/Sveta performed the control experiment, Gennady did not want to share the results with him/her.

21c/d. Хотя два дня назад, пока он обрабатывал собранные ранее данные, Игорь/Света проводил/проводила повторный контрольный эксперимент, Геннадий не захотел делиться с ним/ней результатами.

Правда ли, что контрольный эксперимент проводился повторно?

Even though two days ago while he processed the data collected previously Igor/Sveta performed the control experiment, Gennady did not want to share the results with him/her.

22a/b. Поскольку в прошлую пятницу он проветривал рабочее помещение, пока Яков/Элла распечатывал/расп ечатывала копии праздничных рекламных листовок, Костя винил себя в простуде коллеги.

Правда ли, что Яков/Элла распечатывал/распечатывала копии бухгалтерских форм?

Since last Friday he aired out the work space while Yakov/Ella printed(m/f) copies of the holiday advertisement flyers, Kostja blamed himself for his colleague's illness.

22c/d. Поскольку в прошлую пятницу, пока он проветривал рабочее помещение, Яков/Элла распечатывал/расп ечатывала копии праздничных рекламных листовок, Костя наконец познакомился с новым сотрудником/ новой сотрудницей.

Правда ли, что Яков/Элла распечатывал/распечатывала копии бухгалтерских форм?

Since last Friday while he aired out the work space Yakov/Ella printed(m/f) copies of the holiday advertisement flyers, Kostja finally met his new coworker (m/f).

23a/b. Хотя с утра он делал уборку в квартире, пока Коля/Зина слушал/слушала новый альбом 'Алисы', Иван не стал упрекать его/её ни в чем.

Правда ли, что Иван ни в чем не упрекнул Колю/Зину?

Even though since the morning he cleaned the apartment while Kolja/Zina listened (m/f) to the new album by «Alisa», Ivan did not reproach him/her for anything.

23c/d. Хотя с утра, пока он делал уборку в квартире, Коля/Зина слушал/слушала новый альбом 'Алисы', Иван не стал упрекать его/её ни в чем.

Правда ли, что Иван ни в чем не упрекнул Колю/Зину?

Even though since the morning while he cleaned the apartment Kolja/Zina listened (m/f) to the new album by «Alisa», Ivan did not reproach him/her for anything.

24a/b. Хотя в парке она уныло грызла тыквенные семечки, пока Вера/Витя читала/читал объявления на стенде, Юля не хотела торопить подругу/друга.

Правда ли, что Вера/Витя читала/читал объявления в журнале?

Even though in the park she glumly munched on pumpkin seeds while Vera/Vitja read(f/m) ads on the stand, Julia did not want to hurry her/him up.

24c/d. Хотя в парке, пока она уныло грызла тыквенные семечки, Вера/Витя читала/читал объявления на стенде, Юля не захотела присоединиться к ней/нему.

Правда ли, что Вера/Витя читала/читал объявления в журнале?

Even though in the park while she glumly munched on pumpkin seeds Vera/Vitja read(f/m) ads on the stand, Julia did not want to join her/him.

No-constraint conditions

1a/b. Хотя до того как она открыла собственную фирму, Анна/Дима хорошо зарабатывала/зарабатывал переводами в Лукойле, Дима/Анна не сомневался в правильности её решения/не сомневалась в необх одимости перехода мужа в семейный бизнес.

Была/был ли Анна/Дима недовольна/недоволен зарплатой в Лукойле?

Even though before she started her own company Anna/Dima earned(f/m) good money translating for Lukoil, Dima/Anna did not doubt {the correctness of her decision/the necessity of her husband's move into the family business}.

2a/b. Хотя после того как он закончил генеральную репетицию, Владимир/Кристина был/была полностью доволен спектаклем, Кристина/ Владимир не разделяла/разделял его/её энтузиазма.

Различались ли мнения Кристины и Владимира в отношении спектакля?

Even though after he finished the dress rehearsal Vladimir/Kristina was(m/f) completely satisfied with the play, Kristina/ Vladimir did not share(f/m) her/his enthusiasm.

3a/b. Хотя после того как она написала заказанную статью, Валентина/Александр несколько раз правила/правил текст, Александр все-таки был недоволен ее подачей материала/Валентина больше всего горди лась своим первоначальным вариантом.

Правила/правил ли Валентина/Александр первоначальный текст статьи?

Even though after she wrote the commissioned article Valentina/Aleksandr corrected the text a few times, {Aleksandr nonetheless was dissatisfied with her presentation of the material/Valentina was most proud of her initial version}.

4a/b. Хотя до того как он опубликовал обзор в Литературной Газете, Гавриил/Надежда был/была неизвестен/ неизвестна в писательских кругах, Надежда не сомневалась в появлении интереса к нему после выхода обзора/Гавриил не сомневался в появлении интереса к ней после выхода его обзора.

Был/была ли Гавриил/Надежда широко известен/известна писателям до публикации в Литературной Газете?

Even though before he published the review in Literaturnaja Gazeta Gavriil/Nadezhda was(m/f) unknown(m/f) in literary circles, {Nadezhda did not doubt that there would be much interest in him after the review appeared/Gavriil did not doubt that there would be much interest in her after the review appeared}.

5a/b. Хотя после того как он выиграл зимние соревнования, Павел/Настя недвусмысленно намекнул/нам екнула на участие в Олимпиаде, Настя отказалась от любых комментариев относительно своего подоп ечного/Павел отказался от любых комментариев относительно её высказывания.

Намекнул/намекнула ли Павел/Настя на участие в Олимпиаде?

Even though after he won the winter competitions Pavel/Nastja unambiguously hinted(m/f) at participating in the Olympics, {Nastja refused to make any comments about her protégé/Pavel refused to comment upon her statement}.

6a/b. Хотя до того как она вошла в состав сборной, Наташа/Михаил была/был совершенно неизвестна/ неизвестен специалистам, Михаил хорошо понимал значимость ее успеха/Наташа никогда не сомневалась в его выдающихся тренерских способностях.

Была/был ли Наташа/Михаил знаменитой/знаменитым?

Even though before she entered the team line-up Natasha/Mixail was(f/m) completely unknown(f/m) to aficionados, {Michael well understood the significance of her success/Natasha never doubted his outstanding coaching talent}.

7a/b. Поскольку после того как он переехал работать на север, Юрий/Катя стал/стала получать анонимные угрозы, Катя всеми силами пыталась уговорить его вернуться в деревню/Юрий решил забрать сестру к себе в Норильск.

Пыталась ли Катя уговорить Юрия уехать с севера?/Планировал ли Юрий забрать к себе сестру?

Since after he moved to the north for work Yurij/Katja started(m/f) receiving anonymous threats, {Katja tried everything possible to convince him to return to the village/Yurij decided to move his sister to his place in Norilsk}.

8a/b. Поскольку до того как он не получил печать, Дима/Нина не подписывал/подписывала никаких бумаг, Нина была вынуждена отложить заключение контрактов/Дима был вынужден отложить переговоры о заключении контракта.

Правда ли, что все контракты были подписаны без отлагательств?/Правда ли, что переговоры были проведены в изначально планируемое время?

Since until he received the stamp Dima/Nina did not sign(m/f) any papers, {Nina was forced to postpone the signing of the con-tracts/Dima was forced to postpone the contract-related negotiations}.

9a/b. Поскольку до того как она не разобралась во всех документах, Оксана/Никита не проводила/проводил переговоров с клиентами, Никита не мог сразу оценить ее деловые качества/Оксана не могла рассчиты вать на премию.

Требовалось ли Оксане время на ознакомление с документами?/Верно ли, что Оксана не надеялась на премию?

Since until she understood all the documents Oksana/Nikita did not conduct(f/m) any negotiations with clients, {Nikita could not immediately appreciate her professional qualities/Oksana could not expect a bonus}.

10a/b. Поскольку до того как он устроился в охранное агенство, Николай/Лариса постоянно пересказывал/пересказывала прочитанные детективы, Лариса считала, что новая работа придется ему по душе/Николай надеялся сразить подругу захватывающими случаями из своей практики.

Правда ли, что Лариса устроилась на работу в охранное агенство?

Since before he settled into a job with a security agency Nikolai/Larisa constantly repeated(m/f) the detective stories he/she had read, {Larisa thought that the new job would suit him/Nikolai hoped to impress his girlfriend with fascinating cases from his own work}.

11a/b. Поскольку после того как она перешла на другую работу, Даша/Олег стала/стал забирать детей из садика, Олег/Даша был/была очень рад/рада появившемуся у него/неё свободному времени.

Забирала ли Даша детей из сада?/Была ли Даша огорчена появившемся свободным временем?

Since before she transferred to a new job Dasha/Oleg started(f/m) to pick-up children from kindergarten, Oleg/Dasha was (m/f) very happy(m/f) with the free time available to him/her.

12a/b. Поскольку до того как она начала заниматься латиноамериканскими танцами, Ирина/Антон профессионально занималась/занимался классическим балетом, Антон сразу же по достоинству оценил будущую партнершу/Ирина с большим трудом уговорила его стать ее партнером.

Правда ли, что Ирина/Антон занималась/занимался балетом?

Since before she started to be involved in Latin-American dances Irina/Anton danced(f/m) classical ballet professionally, {Anton immediately recognized his future partner's talents/Irina convinced him with great difficulty to be her partner}.

APPENDIX B

Mean reading times and standard error (in ms) for each region in each condition in the online Experiment 1b

Fixed effect(s) for the best fitting model for each region are shown in the last column with "—" representing the model with no fixed factors other than the intercept. All models had participants and items as random factors.

No-constraint conditions

	No constraint		
Region	Gender match	Gender mismatch	Best fitting model
1	481.8 (13.8)	480.8 (15.8)	—
2	419.1 (13.6)	424.6 (14.1)	—
3	416.6 (10.5)	414.1 (11.3)	—
4	404.6 (11.1)	393.2 (10.2)	—
5	398.8 (10.5)	412.6 (13.9)	—
6	472.9 (16.8)	473.8 (18.4)	—
7	523.3 (22.3)	524.6 (21.1)	—
8	551.1 (23.5)	584.9 (26.6)	—
9 (2nd subject)	533 (19.3)	541.7 (22.6)	—
10	487 (16.9)	504.3 (19.5)	—
11	541.6 (20.1)	607 (27.7)	gender congruency
12	546.6 (21.9)	578.7 (25.3)	—
13	542.1 (20.1)	613.9 (27.6)	gender congruency
14 (3rd subject)	568.2 (20.4)	554 (19.1)	—
15	469.2 (13.8)	459.6 (11.1)	—
16	461.1 (15.7)	481.9 (16.9)	—
17	476 (17.3)	469.6 (15.8)	—

Principle C and poka conditions

	Principle C		Poka		
Region	Gender match	Gender mismatch	Gender match	Gender mismatch	Best fitting model
1	472.4 (11.3)	497.3 (16.7)	478.1 (14.6)	485.1 (16.6)	—
2	412.8 (9.8)	417.8 (13.8)	435.6 (14.9)	405.6 (9.1)	constraint, gender congruency, Constraint × Gender Congruency
3	491.3 (19.6)	456.5 (14.2)	468.5 (17.3)	478.9 (17.4)	—
4	438.1 (14.6)	424.6 (11.6)	483.5 (17.5)	487 (17.2)	—
5	589.5 (26.1)	566.3 (24.2)	430.1 (10.6)	414.7 (9)	—
6	582 (26.4)	639.3 (29.8)	572.5 (23.3)	588.9 (27.2)	constraint, gender congruency, Constraint × Gender Congruency
7	630.3 (24.3)	602.4 (24.9)	597.2 (28.2)	587.7 (23.8)	—
8	506.5 (17.1)	515.3 (18.6)	582.6 (25.4)	593.3 (23.5)	—
9 (2nd subject)	496.6 (16)	500.8 (15)	559 (18.9)	550.1 (17.3)	constraint
10	569.4 (23.1)	584.8 (24.7)	615 (26.6)	570.5 (20.5)	—
11	513.1 (17.3)	526.1 (19.8)	520.7 (18.1)	525.3 (19.5)	—
12	527.9 (16.9)	532.5 (20)	580.8 (24.4)	522.6 (16.6)	constraint, gender congruency, Constraint × Gender Congruency
13	580.1 (25)	578.3 (21.1)	566.8 (20.6)	546.2 (21.9)	—
14 (3rd subject)	560.2 (19.6)	531.1 (18.6)	575.7 (21.8)	511.6 (15.7)	constraint, gender congruency
15	477.6 (12.9)	502.3 (17)	526.7 (19.7)	472.8 (13.9)	constraint, gender congruency, Constraint × Gender Congruency
16	504.5 (18)	473.2 (15.7)	501.4 (16.3)	477.1 (15.7)	—
17	523.8 (18.2)	512.9 (16.8)	506 (14.5)	497.9 (13.7)	—