What Eye Movements Can and Cannot Tell us about *Wh*-Movement and Scrambling

Irina A. Sekerina[1], Anna K. Laurinavichyute[2,3], and Olga Dragoy[2]

[1] The College of Staten Island and The Graduate Center, CUNY

[2] National Research University Higher School of Economics, Russian Federation

[3] University of Potsdam, Germany

Abstract Generative grammar postulates a filler-gap dependency in *Wh*-questions. Visual World Paradigm (VWP) studies of this dependency in English have found an increase in fixations to the filler object during and after the verb which was interpreted as filler reactivation (Trace Reactivation Hypothesis) at the gap and explained by the Active Filler Hypothesis. However, it is possible that such fixations are compatible not only with filler-gap processing, but also with a goal-oriented strategy, i.e., the pragmatic computation of an answer to the question. To disentangle these two possible explanations, we conducted two VWP experiments that investigated comprehension of simple Russian *Wh*-questions in which the type of question (subject vs. object) was crossed with scrambling (object-verb vs. verb-subject). For object scrambling, there was no evidence of reactivation of the scrambled filler; for subject scrambling, there was a brief consideration of the scrambled filler, but not at the gap site. Instead, the referent that was the answer to the question was fixated. For object *Wh*-questions, the eye-movement pattern was inconclusive, as it was consistent with both filler-gap and goal-oriented processing. We suggest that the latter strategy of looking for an answer in the visual context may account for eye-movements in all types of *Wh*-movement: when participants answer a question, they prioritize computing the answer (and visually verifying it) over computing filler-gap dependencies.

1 Introduction

There are few phenomena in linguistics that have generated as much excitement as *empty categories (EC)*. ECs (*traces* or *gaps*) are the constituents of a sentence that do not overtly appear in the word string because they lack any phonological or orthographic realization. They generally serve the grammatical function of nominals and occupy various phrasal nodes in the syntactic tree. In order to have

semantic content, an EC must be identifiable, and this is accomplished through association with another referential constituent (i.e., an antecedent, or *filler*) in the tree. The EC must be coindexed (bound) and chain-linked to its moved antecedent, from which it inherits its content.

In sentence processing, sentences that contain traces are referred to as *filler-gap dependencies* and present a special challenge. Comprehenders must project structure onto a string of words, and this is not an easy feat with gaps. How does the human parser identify the position of an inaudible (or invisible) gap and connect the filler with it (for a review see Fodor, 1995)? Frazier and Clifton (1989) proposed the *Active Filler Hypothesis* (*AFH*, later subsumed under the Minimal Chain Principle by De Vincenzi 1991), which says that once an element of a category XP (i.e., a filler) is identified as moved from its argument position, a corresponding empty XP category (i.e., a gap) must be posited as soon as the language allows. Swinney and colleagues (1988) tested the AFH using cross-modal priming in sentences like (1) and found an effect of gap-filling processing in the form of reactivation of a noun such as *girl*, related to the filler (i.e., *boy*), in the *Wh*-trace position. They explained this effect as support for the *Trace Reactivation Hypothesis.*

(1) The policeman saw the boy_i that the crowd at the party accused t_i of the crime.

A strong push in sentence processing research in the 1990s to expand the investigation of filler-gap dependencies other than *Wh*-movement and into languages other than English took place in parallel with developments in syntactic theory regarding word order variation of arguments, known as s*crambling*. Within generative grammar, long-distance scrambling that moves an argument out of its VP and into the sentence-initial position, crossing the clause boundary, was widely believed to be similar to *Wh*-movement, i.e., an instance of A'-movement. In contrast, there was no agreement regarding the short-distance scrambling that had to do with the movement of a direct or indirect object within the same clause. Some syntacticians took it to be an instance of A-movement similar to NP-movement (Bošković & Takahashi, 1998), while others argued for a non-movement, base-generation approach, using German, Dutch, Japanese, and Russian as case studies (see Karimi, 2003).

The Garden-Path theory of sentence processing (Frazier and Fodor, 1978), based on (a) the interaction of grammatical parameters and sentence processing strategies and (b) cross-linguistic validation (Frazier, 2013), aimed to explain how comprehenders deal with word order variation in any language. The theory thus needed to empirically evaluate whether filler-gap processing occurred in scrambling. Cross-modal priming (Swinney et al., 1988) lent itself very well to testing whether scrambled sentences are more like Wh-filler-gap dependencies,

with a trace left behind by a moved object, or are a base-generated phenomenon (see Sekerina, 2003, for a review).

Clahsen and Featherston (1999) found reactivation of a scrambled direct object (DO) in German ditransitive sentences with verb-particle constructions: primes related to the scrambled DO elicited faster lexical decision times than unrelated primes at the position of the gap after the verb, but not at an earlier control position, which favored the Trace Reactivation Hypothesis. Similarly, Nakano, Felser, and Clahsen (2002) demonstrated that in Japanese, gap-filling processing took place with DOs that were scrambled long-distance out of the embedded clause, albeit only for participants with a high working memory capacity. More recently, Marinis and colleagues (Marinis, 2018) found reactivation of the scrambled DO at the gap position as evidence for the base word order IO-DO in Greek. In contrast, van de Koot, Silva, Felser and Sato (2015) did not find an effect of gap-filling processing in similar Dutch sentences and argued for base-generation of different word orders for objects.

Advancement of a more sophisticated online method, namely, *the Visual World eye-tracking Paradigm (VWP*; Trueswell & Tanenhaus, 2004) has brought about a renewed interest in studying filler-gap processing using spoken sentences. In VWP experiments, such sentences are paired with a visual context in which the referents of potential fillers are presented as pictures. Gap-filling is argued to occur when filler reactivation can be observed visually as an increase in looks to the filler at the gap position. Using the VWP, Dickey and colleagues (Dickey, Choy, & Thompson, 2007; Dickey & Thomspon, 2009; see also Sussman & Sedivy, 2003) established reactivation of the filler *who* in spoken *Wh*-questions and object relative clauses (2a-b), but not in passives (2c):

(2) One day a bride and a groom were walking in the mall.
The bride was feeling playful, so the bride tickled the groom.
A clerk was amused.
a. Who$_i$ did the bride tickle t_i today in the mall?
b. Point to who$_i$ the bride was tickling t_i in the mall.
c. Point to who$_i$ was tickled t_i by the bride in the mall.

For *Wh*-movement (2a, Dickey et al., 2007), they found an increase in proportions of fixations on the object (e.g., *groom*) versus the subject (e.g., *bride*) while hearing the verb, which according to the authors, indicated reactivation of the referent filler at the gap position. A similar pattern of eye movements was later found for object relative clauses (2b, Dickey & Thompson, 2009), which the authors argued reflected automatic gap-filling in sentences with different types of *Wh*-movement.

However, one could argue that eye movements in VWP experiments, such as the ones by Dickey and colleagues described above, might not reflect filler

reactivation, but rather pragmatic *goal-oriented processing* (Salverda, Brown, and Tanenhaus, 2011), i.e., when the participants have to explicitly answer a comprehension question, their fixations could reflect a search for a referent that is the answer to the question. Note that in the examples (2a-b), with four referent pictures present (e.g., a groom, bride, cashier, and mall), the answer to the question and the filler *who* refer to the same referent (e.g., *groom*). Dickey and colleagues specifically discarded the idea of a search for the answer to the question/task, arguing that such a strategy should result in an increase of fixations to the gapped object in the passives (2c) as well, which was not observed. But it is possible that the presence of two dependencies, i.e., NP-movement that is embedded in the headless subject relative clause in (2c), can increase the processing load and thus slow down language processing, a necessary prerequisite for the search for the referent. For the English *Wh*-questions and object relatives in (2a-b), the eye-movement patterns are equally compatible with both gap-filling and goal-oriented processing, making it difficult to tease these apart in English.

Russian, with its flexible word order for moved arguments in both *Wh*-questions and scrambled sentences, is much better suited for contrasting the two processing strategies. We had two goals for this study: (1) to obtain cross-linguistic validation of the previously found effect of gap-filling in *Wh*-questions for Russian using the VWP (Dickey et al., 2007; Dickey & Thompson, 2009; Sussman & Sedivy, 2003), and (2) to attempt to separate gap-filling and goal-oriented processing by contrasting Russian simple *Wh*-questions with and without scrambling in which the scrambled phrase is different from the answer to the question. There were several groups of participants in our study, but here we present just the findings from the monolingual Russian-speaking adults who participated in two VWP experiments.

2 Experiment 1

2.1 Method

2.1.1 Participants

The participants were 36 native Russian speakers (M_{Age}= 50; 23 women) residing in Moscow and they did not have any reported neurological disorders. The participants volunteered for the experiment and were tested individually. They all gave written informed consent in accordance with the Declaration of Helsinki.

2.1.2 Design and Materials

Following Dickey et al. (2007), 20 short experimental stories were designed in such a way that each story mentioned three animate protagonists and a location. The stories consisted of three preamble sentences (3a-c) followed by an experimental question in one of two conditions, i.e. a subject *Wh*-question with a scrambled object $girl_{ACC}$ and a moved subject who_{NOM} (4), or an object *Wh*-question with who_{ACC} (5). We will refer to the two conditions as *Subj Wh-question + Obj Scramb*, and *Obj Wh-question*. There is an intricate interaction between information structure and word order in Russian (Bailyn, 2012), but both (4) and (5) are equally natural in the context of the preamble in (3a-c).

(3) a. Однажды девочка и мальчик шли по школе.
One day a girl and a boy were walking around the school.

b. И вдруг мальчик поцеловал девочку.
And suddenly the boy kissed the girl.

c. Учитель очень удивился.
The teacher was very surprised.

(4) SUBJ *WH*-QUESTION + OBJ SCRAMB

Кто$_2$	девочку$_1$	t_2 поцеловал t_1	в школе?[1]
Who$_{NOM}$	girl$_{ACC}$	kissed	at school

'Who kissed the girl at school?'

(5) OBJ *WH*-QUESTION

Кого$_1$	мальчик	поцеловал t_1	в школе?
Who$_{ACC}$	boy$_{NOM}$	kissed	at school

'Who did the boy kiss at school?'

The preamble mentioned a transitive action (3b), two referents involved in the action (*boy* and *girl*), an animate distractor (*teacher*) and a location (*school*). The names of the referents were balanced in length and frequency (Lyashevskaya & Sharov, 2009), and the initial phonemes of the four referents were different in every story. The experimental questions were rotated through two lists in a Latin square design.

The four referents were depicted as black-and-white images. The images were located in the four corners of the screen, with each image occupying 32.5% of the height and width of the visual panel (see Figure 1).

[1] We are showing the subject *wh*-word – t_2 dependency in (4) for clarity, but it is well-established in sentence processing research that it is cost-free.

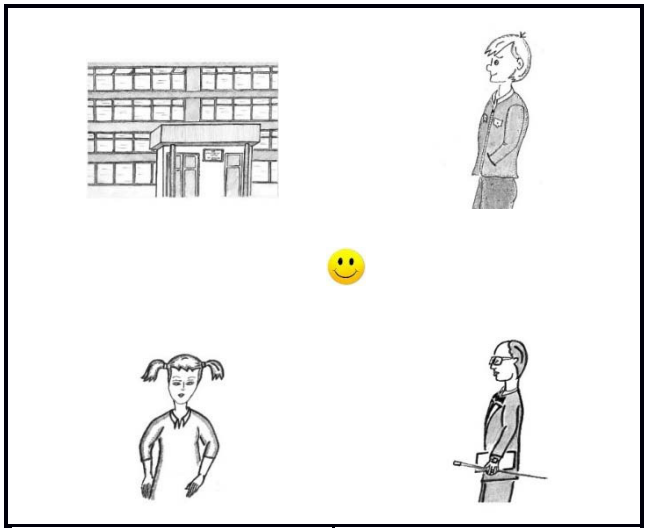

Figure 1. Experiments 1 and 2: Visual context used with the experimental story (3)-(5).

The stories were recorded by a professional announcer, a male native Russian speaker, with a mean speed of three syllables per second. The audio was recorded in .wav format (16 bit mono, 44.1 kHz) and was played through external speakers.

Both 20 experimental and 20 filler stories contained three animate protagonists. The questions in the filler stories probed for where the action had happened. The order was pseudo-randomized such that the experimental stories in two conditions were interspersed evenly with the fillers in a 40-story list.

We predicted that in the Obj *Wh*-questions (5), participants would start looking at the filler object (*girl*) no later than at the gap at the verb, replicating the eye-movement pattern found for English (2a) (Dickey et al., 2007). In the Subj *Wh*-Question + Obj Scramb condition (4), the subject filler is adjacent to its potential gap, and it will attract looks to the subject (*boy*). However, if there is a hypothesized second filler-gap dependency for the scrambled object (*girl*), participants should shift their fixations from the boy to the girl at the verb.

2.1.3 Procedure

The experiment was conducted with the 60 Hz LC Technologies Eyegaze camera (Fairfax, VA, USA) using a chin rest. It began with a 9-point calibration procedure and five practice stories. Each trial started with a central fixation point (a cross in Experiment 1 or a happy face in Experiment 2) presented in the center of the screen for 300 ms; then the audio recording of the story and the visual context

with four referent pictures (counterbalanced across the corners of the screen) were presented simultaneously. The visual context remained on the screen for the duration of the audio recording and for 5 s after the end of the question. Participants were asked to answer the comprehension question by looking at the corresponding image during the 5 s of silence. The experimenter controlled the progression of trials using a second screen, performing a recalibration or offering the participant a short break if needed. The entire experiment lasted for 30-35 minutes, with 10-15 minutes taken by calibration and practice.

2.1.4 Data Treatment and Analysis

In order to track the changes in referent activation over time, four regions of interest (ROIs) were marked in each experimental question (6):

(6)	**ROI 1**	**ROI 2**	**ROI 3**	**ROI 4**
	Wh-word	NP1/NP2	Verb	location
SUBJ WH-QUESTION + OBJ SCRAMB	Kto	devochku		
	who_{NOM}	$girl_{ACC}$	poceloval	v shkole?
OBJ WH-QUESTION	Kogo	mal'chik	kissed	at school
	who_{ACC}	boy_{NOM}		

The beginning and the end of each ROI were marked by two raters based on a oscillogram of the audio recording made using Sound Forge Audio Studio 12. We added 200 ms to each ROI to accommodate saccade planning and execution (Matin, Shao, & Boff, 1993).

Statistical analysis and data visualization were performed using R (R Core Team, 2016) and *ggplot2* (Wickham, 2016). Linear mixed models (LMMs) were estimated with the *lme4* package version 1.1-8 (Bates, Maechler, Bolker, & Walker, 2015). We used the *sjPlot* package (Lüdecke, 2017) to create a table comparing several LMM outcomes.

Each response to a question was coded as correct if the proportion of looks to the correct referent was greater than the proportion of looks to any other referent. We analyzed only the trials for which correct answers were provided (98% of the data) and only those fixations that lasted more than 100 ms. For modeling, empirical logit regression (Barr, 2008) was used; samples were grouped by

participants and by trials and averaged within 50-ms intervals. The dependent variable was the subject advantage, i.e., the difference between the proportion of frames with looks to the subject (*boy*) and the object (*girl*) of the transitive action *kissed*. The independent variables included in the model were Time (in seconds from the start of the ROI), Question Type (coded as '1' for the Subj *Wh*-question + Obj Scramb and '-1' for the Obj *Wh*-question), and their interaction. The model also contained aggregated random intercepts by-participant and by-item as well as random slopes for time.

2.2 Results and Discussion

As the dependent variable was the subject advantage, a positive difference between the proportions of frames with fixations on the subject (*boy*) and the object (*girl*) indicates that the participants looked more to the subject while a negative difference indicates more looks to the object. Figure 2 shows the mean proportions of looks to the two referents in the four ROIs. Statistical analysis (see Table 1 in the Appendix) revealed no difference in the proportions of looks for ROI 1.

Our predictions were confirmed for the Obj *Wh*-questions (5), but not for the Subj *Wh*-Question + Obj Scramb condition (4). For the second half of ROI 2 (*girl*-ACC/*boy*-NOM), there were significantly more looks to the filler object (*girl*) than to the subject (*boy*) in the Obj *Wh*-question (Fig. 2, bottom panel). The opposite pattern characterized the Subj *Wh*-question + Obj Scramb condition, with more looks to the subject than to the object (a main effect of Question Type and interaction with Time). But the hypothesized shift in looks from the subject to the scrambled object never materialized. Even while hearing *girl*-ACC in ROI 2, the participants continued to look at the subject, and this eye-movement pattern carried forward in ROIs 3 and 4, the verb and the location.

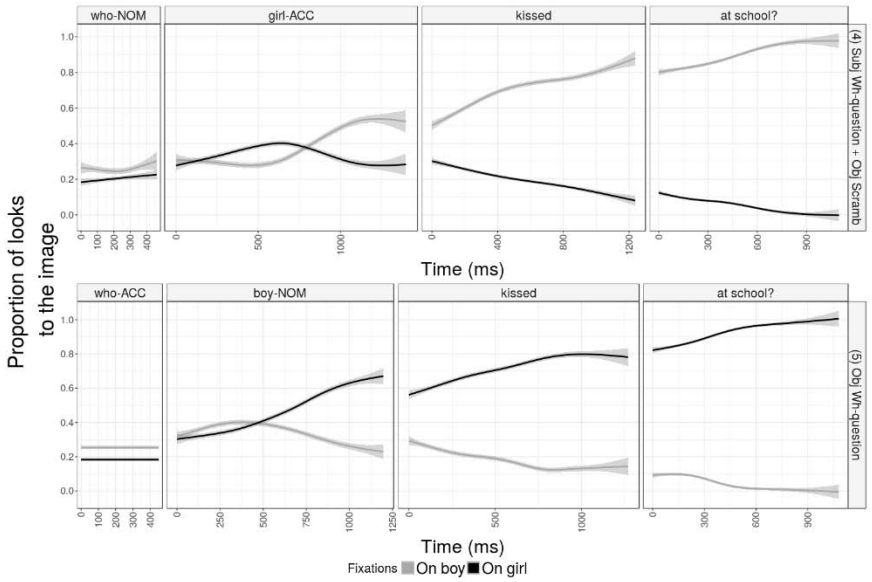

Figure 2. Experiment 1: Proportions of looks to the subject (*boy*) and object (*girl*) across the four ROIs in Subj *Wh*-question + Obj Scramb (4) and Obj *Wh*-questions conditions (5).

Two not mutually exclusive explanations are possible: either object scrambling in Russian does not result in a filler-gap dependency, or the goal-oriented strategy to respond to a question overrode eye movements that might reflect reactivation of the scrambled object at the gap. The first explanation is well-attested in theoretical syntax, where object scrambling is considered an example of *short-distance A-scrambling* (object shift) in Germanic languages and Japanese and is assumed to be base-generated (see Bailyn, 2004, for discussion). Similarly, recent experimental evidence from cross-modal priming in Dutch (van de Koot et al., 2015) has demonstrated a lack of reactivation of the scrambled DO suggesting that there is no syntactic trace in A-scrambling.

If short-distance object scrambling in Russian does not involve movement and thus there is no reactivation of the filler at the gap, what guides eye movements to the subject at and after the verb? Recall that the participants in the present experiment were instructed to silently stare at the picture of the referent that corresponded to the correct answer. The pragmatic goal-oriented processing proposed by Salverda et al. (2011) for VWP studies is an obvious explanation. It argues that a large proportion of eye movements, including anticipatory ones, is more affected by the task at hand than by gap-filling computation.

It might additionally be argued that even in the Obj *Wh*-questions, our participants may have employed this strategy because, just like in the English object *Wh*-questions (2a), the girl is simultaneously the filler object and the correct

answer. Interestingly, this pattern of eye movements was so strong that the looks to the referent that was explicitly mentioned in the sentence (*boy*) could not compete with those to the referent that was the answer but was not mentioned (*girl*), contra the referential priority principle (Knoeferle & Guerra, 2016).

In an attempt to decrease goal-oriented processing, we conducted Experiment 2, in which no explicit fixation of the referent that was the answer was required. We used the same two conditions, Obj-*Wh*-questions and Subj *Wh*-question + Obj Scrambling, but also added two new ones, namely, a scrambling version of the example (5), i.e., an Obj *Wh*-question + Subj Post condition illustrated in the example (8), and a version of the example (4), i.e., a Subj *Wh*-question without scrambling (7). We make the following predictions: First, if goal-oriented processing was artificially boosted by the requirement of fixating the referent in answering the question in Experiment 1, it should be attenuated in all four conditions in Experiment 2. Second, if short-distance scrambling does not involve gap-filling regardless of whether it involves an object or a subject, eye-movement patterns in the pairs of Subj *Wh*- and Obj *Wh*-questions should not differ. Identical eye-movement patterns in processing of all types of *Wh*-questions across both experiments, regardless of the presence or absence of scrambling, would be more consistent with the pragmatic goal-oriented strategy of answering a question rather than the Trace Reactivation hypothesis.

3 Experiment 2

3.1 Method

3.1.1 Participants

Participants were 12 native Russian speakers, 3 of them men (M_{age} = 29.1, range: 18-42) without reported neurological disorders. They were recent immigrants from Russia, Ukraine, Georgia, and Kyrgyzstan who enrolled in the local college to obtain an American college degree. They were also L2 speakers of English of intermediate proficiency. All participants filled out a language history questionnaire and were classified as native Russian speakers, i.e., L2 English learners who continued to speak Russian more than English in their everyday lives. The study was carried out in accordance with the ethical principles of psychologists and was approved by the Institutional Review Board of the College of Staten Island. All participants signed a written informed consent in accordance with the Declaration of Helsinki.

3.1.2 Design and Materials

The same 20 short experimental and 20 filler stories from Experiment 1 were used for Experiment 2. In addition to the two conditions from Experiment 1, two new complementary conditions were designed: Subj *Wh*-question (7), which was parallel to (4), but without scrambling, and Obj *Wh*-question + Subj Post (8), which was parallel to (5), but with a postverbal subject that could be construed as subject scrambling.

(4) SUBJ WH-QUESTION + OBJ SCRAMB

Кто$_2$	девочку$_1$	t_2 поцеловал t_1	в школе?[2]
Who$_{NOM}$	girl$_{ACC}$	kissed	at school

'Who kissed the girl at school?'

(5) OBJ *WH*-QUESTION

Кого$_1$	мальчик	поцеловал t_1	в школе?
Who$_{ACC}$	boy$_{NOM}$	kissed	at school

'Who did the boy kiss at school?'

(7) SUBJ *WH*-QUESTION

Кто$_1$	t_1 поцеловал	девочку	в школе?
Who$_{NOM}$	kissed	girl$_{ACC}$	at school

'Who kissed the girl at school?'

(8) OBJ *WH*-QUESTION + SUBJ POST

Кого$_1$	t_2 поцеловал t_1	мальчик$_2$	в школе?[3]
Who$_{ACC}$	kissed	boy$_{NOM}$	at school

'Who did the boy kiss at school?'

Four lists were created in a Latin square design, with five experimental stories per condition, each based on a spoken sentence ((4)-(5) and (7)-(8)) paired with the corresponding visual context (Figure 1). Three participants were randomly assigned to each list. The auditory stimuli for the two new conditions (7) and (8) were recorded by the same speaker simultaneously with the old (4) and (5). All the remaining details of the design and materials were equivalent to Experiment 1.

3.1.3 Procedure

Participants' eye-movements were recorded by the ISCAN ETL-500 head-mounted eye-tracking system. Each participant underwent a short 5-point

[2] See footnote 1.

[3] We are showing the second potential reverse gap-filler dependency in (8), i.e., t_2 - the postponed subject *мальчик*$_2$, but its existence is debatable and warrants a separate investigation.

calibration procedure prior to the experiment. Eye-movements were sampled at a rate of 30 Hz and were recorded on a SONY DSR-30 digital videotape recorder. Auditory stimuli were played to the participant through speakers.

In contrast to Experiment 1, in which the participants had to silently fixate the referent picture that was the answer to the question, we asked the participants to first answer the comprehension question out loud and then click on the referent picture.

3.1.4 Data Treatment and Analysis

Data treatment and analysis were the same as in Experiment 1. The same four ROIs illustrated in (6) were identified in the two new conditions (7) and (8). However, ROI 2 and ROI 3 were switched: ROI 2 now contained the verb and ROI 3 contained NP1/NP2. Again, we analyzed only the trials for which correct answers were provided (98% of all trials) and only fixations that lasted more than 100 ms. Proportions of looks were averaged by participants and by items within 100-ms intervals. The eye-movement data were coded with 30 frames/sec resolution, so for each 100-ms interval, 3 samples were averaged for each participant. We compared conditions (4) and (5) to see if the goal-oriented strategy to answer the question was attenuated when the explicit requirement of fixating the referent was removed. Conditions (7) and (8) were compared to gain new insight into eye movements in the processing of *Wh*-questions with or without additional scrambling. The independent variables included in the model were Time in ms from the start of the ROI, Question Type (for the first comparison, coded as '1' for the Subj *Wh*-question + Obj Scramb (4) and '-1' for the Obj *Wh*-question (5); for the second comparison, Obj *Wh*-question + Subj Post (8) was coded as '-1' and Subj *Wh*-question (7) as '1'). The model also included aggregated random intercepts as well as random slopes for time, both by participant and by item.

3.2 Results and Discussion

3.2.1 Conditions (4) and (5) (the same as in Experiment 1)

Figure 3 (comparable to Fig. 2 for Experiment 1) shows that in general, the results look similar to Experiment 1, despite Experiment 2 having fewer data points (only 12 participants in Experiment 2 versus 36 in Experiment 1). A statistical comparison (see Table 2 in the Appendix) demonstrated that for ROI 1 (*Wh*-word) and ROI 2 (*girl*-ACC/*boy*-NOM), no difference in proportions of looks across conditions was found. Starting with ROI 3 (*kissed*), there were significantly more looks to the filler object (*girl*) than to the subject (*boy*) in the Obj *Wh*-question

(Fig. 3, bottom panel). The opposite pattern characterized the Subj *Wh*-question + Obj Scramb condition, with more looks to the subject than to the object (a main effect of Question Type and interaction with Time). Just like in Experiment 1, there was no shift in looks from the *wh*-subject to the scrambled object. This preference to fixate the answer referent increased over time; the pattern continued into the following ROI 4 (the location) and became stronger.

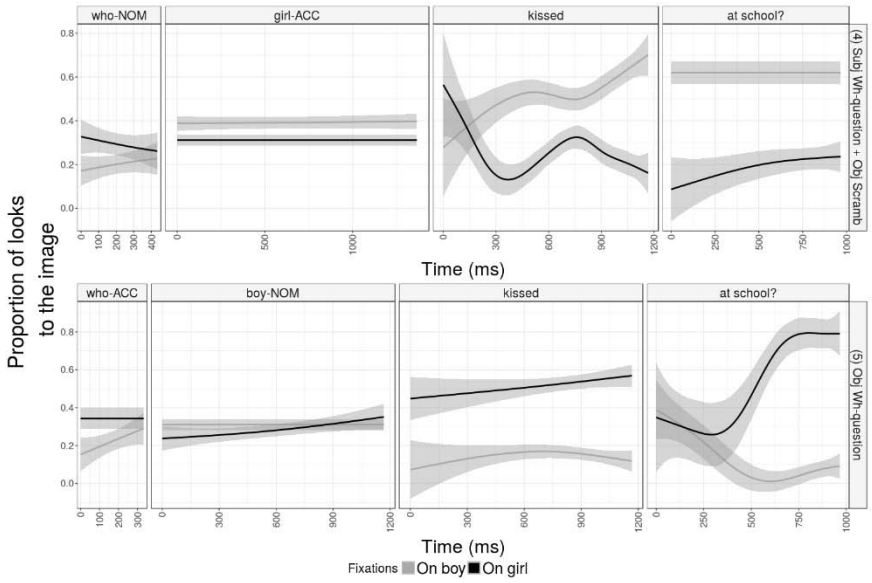

Figure 3. Experiment 2: Proportions of looks to the subject (*boy*) and object (*girl*) across the four ROIs in Subj *Wh*-question + Obj Scramb (4) and Obj *Wh*-questions (5) conditions.

The only difference from Experiment 1 was the locus of divergence in the looks: in Experiment 2, the difference between conditions appeared at the verb, while it manifested in the previous region (*girl*-ACC/*boy*-NOM) in Experiment 1. One possible reason for the delayed effect might be the lower number of observations per condition in Experiment 2. Importantly, the main pattern of eye movements in the later regions of the sentence remained unchanged, which allows us to conclude that the requirement to look at the referent filler for the correct answer in Experiment 1 was not a confound in the observed pattern of fixations.

3.2.2 Conditions (7) and (8)

Figure 4 demonstrates the mean proportions of looks to the two referents in the first four ROIs. The results are less clear here as statistical analysis (see Table 3 in the Appendix) reveals. Similar to Experiment 1, there were more looks to the subject (*boy*) in the Subj *Wh*-questions (7), while there were more looks to the object (*girl*) in the Obj Wh-question + Subj Post condition (8). However, this pattern was observed only in ROI 1 (*who*-ACC/*who*-NOM) and ROI 4 (*at school*). In contrast to Experiment 1, there were no differences in ROI 2 (*kissed*) and ROI 3 (*girl*-ACC/*boy*-NOM). Instead, we found a significant interaction between Question Type and Time in ROI 3, as the participants briefly shifted their looks from the object (*girl*) to the subject (*boy*) in the Obj *Wh*-question + Subj Post (8) (e.g., *Кого$_1$ t_2 поцеловал t_1 мальчик$_2$ в школе?* 'Who$_{ACC}$ did the boy$_{NOM}$ kiss at school?') (Fig. 4, bottom panel), whereas there was no such shift in the Subj *Wh*-questions (7) (top).

This brief shift in condition (8) that lasted approximately 300 ms could be due to one of two reasons. First, it could be driven by a second reverse gap-filler dependency in which the gap preceded the filler subject in postposition, but its implications for processing are unclear. But another explanation is more likely, i.e., that the subject was late and separated from the object by the verb. When it was finally mentioned, participants felt compelled to briefly look at the subject precisely at that point in the sentence. These looks rapidly dissipated leading to a steady increase in looks to the correct answer to the question. In condition (7), where the *Wh*-subject filler is immediately adjacent to its gap, the participants started looking at the answer referent (*boy*) after the verb and never shifted their gaze to the other referent (*girl*), and the immediate adjacency of the subject did not warrant looks.

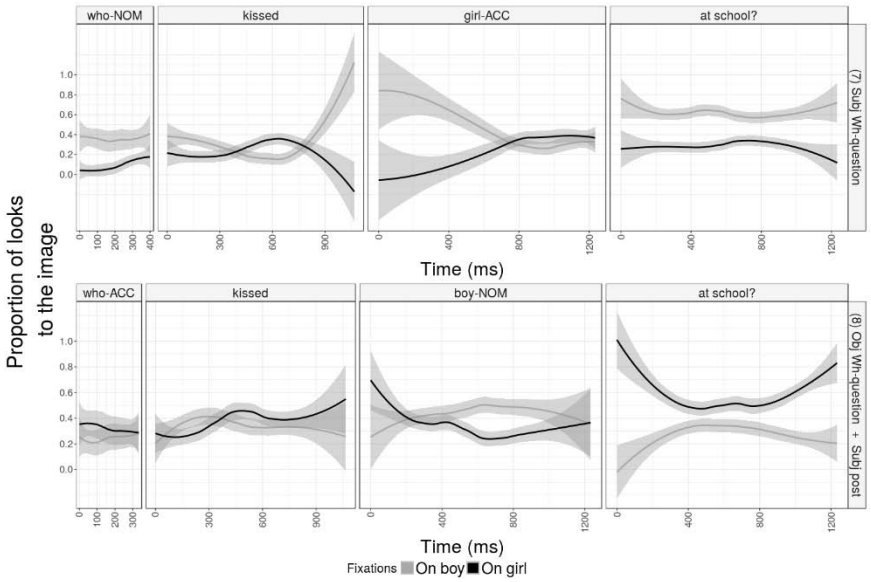

Figure 4. Experiment 2: Proportions of looks to the subject (*boy*) and object (*girl*) across the four ROIs in Subj *Wh*-question (7) and Obj *Wh*-question + Subj Post (8) conditions.

4 General Discussion

The joint results of Experiments 1 and 2 point to two conclusions regarding *Wh*-movement and scrambling in simple Russian *Wh*-questions. First, there was no evidence of reactivation of the scrambled object at the gap at the verb that underwent short-distance scrambling in example (4). The case with the postponed subject in the example (8) is less clear due to a brief shift in looks from the object to the subject. Additional investigation of the reverse gap-filler dependency with postponed subjects is clearly warranted in the future, but for now we ascribe this brief shift to the subject's distance from the object and its direct mention in the sentence. This suggests that scrambling in Russian may not trigger a filler-gap dependency, a conclusion that has some independent support (Bailyn, 2004; Karimi, 2003; van de Koot et al. 2015).

Second, despite the fact that eye movements for scrambling sentences did not reflect reactivation of the scrambled phrase, they nevertheless were quite systematic in that they were most likely guided by the goal-oriented strategy of looking at the answer to the question. It is improbable that the goal-oriented strategy affected only the conditions with scrambling. Moreover, we suggest that the strategy of looking for an answer in the visual context may account for eye movements even in object *Wh*-questions. The increase in fixations on the filler

object at and after the verb in object *Wh*-questions may not necessarily reflect reactivation of the trace, but rather the search for the answer to the question. Importantly, we do not claim that trace reactivation does not happen, we merely suggest that it is not solely reflected in the eye movements in VWP experiments.

It must be noted that the design of our VWP experiments may have encouraged goal-oriented processing: the task was to point/look at the answer to the question, and the case-marked *wh*-word by itself was a sufficiently strong cue that encouraged a search for the correct answer. In particular, the *who-ACC* question always targeted the patient, whereas the *who-NOM* question always targeted the agent of the action described in the story. At this point, it is not possible to ascertain whether participants were involved in strategic processing using *wh*-words as predictive cues or performing normal parsing. To do so, additional filler trials are needed, where *who-ACC* and *who-NOM* target the patient and the agent in one of the preamble sentences different from the experimental question.[4] Finally, to decrease the pressure of goal-oriented processing, it might be advisable to test filler-gap processing in structures other than questions that do not require explicit responses.

We conclude that the eye movements across all four types of simple Russian *Wh*-questions in both experiments are consistent with pragmatic goal-oriented processing: when participants are required to answer a comprehension question, they may prioritize computing the correct answer (and visually verifying the answer they choose) over computing filler-gap dependencies. The goal-oriented hypothesis (Salverda et al. 2011) states that when a referent is incompatible with, or irrelevant to, the current goal, fixations on it will be greatly reduced or even absent, and such absence of fixations on the filler referent speaks neither in favor of nor against successful reference resolution. In conclusion, although we found that processing of *Wh*-questions and scrambling differs in simple Russian sentences, it may be for reasons that do not bear on psycholinguistic theories of filler-gap dependencies. We need to independently verify conclusions regarding the Active Filler and Trace Reactivation Hypotheses for experiments conducted with the Visual World Paradigm that employ questions as linguistic materials.

Admittedly, our evidence in favor of eye movements reflecting goal-oriented processing is circumstantial, rather than direct. Direct evidence in favor of or against the goal-oriented strategy could come from a construction that is uniformly considered to be an instance of syntactic movement, where the moved object does not coincide with the goal of the task (the answer to the question, the object that should be pointed at, etc.). Another, though probably less convincing way, would be to test syntactic constructions that would not require looking for an answer in an experimental task. A more general 'look and listen' task is more

[4] On Monday a boy and girl walked past the teacher. Suddenly, the boy_1 pushed the $girl_2$, which surprised the $teacher_3$. He told both to leave the $school_4$. Nobody realized

(a) who_{ACC2} the boy on Monday pushed t_2. (referent: $girl_{ACC}$)

(b) who_{ACC3} the boy on Monday surprised t_3. (referent: $teacher_{ACC}$).

likely to elicit comprehension-based eye movements that would allow trace reactivation to surface.

Acknowledgments

We would like to thank Anastasia Ulicheva, Maria Ivanova, and Svetlana Kuptsova for their help with planning and conducting the experiment, and Jill Jegerski for providing very useful comments on the chapter. We are especially thankful to our colleagues Janet Dean Fodor, Katy Carlson, and Michael Walsh Dickey whose very thoughtful comments have substantially improved both the content and style of this chapter. The study has been funded by several PSC-CUNY grants to Irina A. Sekerina and by the Center for Language and Brain NRU Higher School of Economics, RF Government grant, ag. № 14.641.31.0004, to Anna Laurinavichyute and Olga Dragoy.

References

Bailyn, J. F. (2004). Generalized inversion. *Natural Language and Linguistic Theory, 22*, 1-49.

Bailyn, J. F. (2012). *The Syntax of Russian.* Cambridge: Cambridge University Press.

Bates, D., Maechler, M., Bolker, B., Walker, S. (2015). Fitting linear mixed-effects models using *lme4. Journal of Statistical Software, 67*(1), 1-48.

Bošković Ž., & Takahashi, D. (1998). Scrambling and last resort. *Linguistic Inquiry, 29*, 347-366.

De Vincenzi, M. (1991) Filler-gap dependencies in a null-subject language: Referential and Non-referential Whs. *Journal of Psycholinguistic Research, 20*(3), 197-213.

Dickey, M. W., Choy, J. J., & Thompson, C. K. (2007). Real-time comprehension of *Wh*-movement in aphasia: Evidence from eyetracking while listening. *Brain and Language, 100*(1), 1-22.

Dickey, M. W., & Thompson, C. K. (2009). Automatic processing of *Wh*- and NP-movement in agrammatic aphasia: Evidence from eyetracking. *Journal of Neurolinguistics, 22*(6), 563-583.

Fodor, J. D (1995). Comprehending sentence structure. In L. R. Gleitman and M. Liberman (Eds.), *An Invitation to Cognitive Science, Volume I: Language*. (pp. 209-246.) Cambridge, MA: The MIT Press.

Frazier, L. (2013). Syntax in sentence processing. In R. P. G. Van Gompel (Ed.), *Sentence Processing.* (pp. 21-50). New York, NY: Psychology Press.

Frazier, L., & Clifton, C., Jr. (1989). Successive cyclicity in the grammar and the parser. *Language and Cognitive Processes, 4*(2), 93-126.

Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition, 6*(4), 291-325.

Karimi, S. (Ed.) (2003). *Word Order and Scrambling*. Malden, MA: Blackwell Publishing.

Knoeferle, P., & Guerra, E. (2016). Visually situated language comprehension. *Language and Linguistics Compass, 10*(2), 66-82.

Lüdecke, D. (2017). *sjPlot: Data Visualization for Statistics in Social Science*. R package version 2.3.3, https://CRAN.R-project.org/package=sjPlot.

Lyashevskaya, O. N., & Sharov, S. A. (2009). *Chastotnyj Slovar' Sovremennogo Russkogo Jazyka (na Materialakh Natsional'nogo Korpusa Russkogo Jazyka*. (in Russian) [*'Frequency Dictionary of Modern Russian* (*based on the materials of the Russian National Corpus)'*]. Moscow: Azbukovnik.

Mack, J. E., Meltzer-Asscher, A., Barbieri, E., & Thompson, C. K. (2013). Neural correlates of processing passive sentences. *Brain Sciences, 3*(3), 1198-1214. doi:10.3390/brainsci3031198

Marinis, T. (2018). Cross-modal priming in bilingual sentence processing. *Bilingualism: Language and Cognition*, FirstView. doi:10.1017/S1366728917000761

Matin, E., Shao, K. C., & Boff, K. R. (1993). Saccadic overhead: Information-processing time with and without saccades. *Attention, Perception, & Psychophysics, 53*(4), 372-380.

R Core Team. (2016). *R: A language and environment for statistical computing [Computer software manual]*. Vienna, Austria. Retrieved from https://www.R-project.org/

Salverda, A. P., Brown, M., & Tanenhaus, M. K. (2011). A goal-based perspective on eye movements in visual world studies. *Acta Psychologica, 137*, 172-180.

Sekerina, I. A. (2003). Scrambling and processing: Dependencies, complexity, and constraints. In S. Karimi (Ed.), *Word Order and Scrambling*. (pp. 301-324). Malden, MA: Blackwell Publishing.

Sussman, R. S., & Sedivy, J. C. (2003). The time-course of processing syntactic dependencies: Evidence from eye movements. *Language and Cognitive Processes, 18*(2), 143-163.

Swinney, D., Ford, M., Frauenfelder, U., & Bresnan, J. (1988). Coreference assignment during sentence processing. In B. Grosz, R. Kaplan, M. Macken, & I. Sag (Eds.), *Language Structure and Processing*. Stanford: CSLI.

Titov, E. (2017). The canonical order of Russian objects. *Linguistic Inquiry, 48*(3), 427-457.

Trueswell, J. C. and Tanenhaus, M. K. (Eds.) (2004). *Approaches to Studying Word-Situated Language Use*. Cambridge, MA: The MIT Press.

van de Koot, H., Silva, R., Felser, C., & Sato, M. (2015). Does Dutch *A*-scrambling involve movement? Evidence from antecedent priming. *The Linguistic Review*, 32(4), 739-776.

Wickham, H. (20016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

Appendix

Table 1. *Exp. 1: Results of linear mixed-effects regression analysis by ROI, conditions (4)-(5).*

	Who$_{NOM/ACC}$			*girl*$_{ACC}$ / *boy*$_{NOM}$			*kissed*			*at school?*		
	Estimate	**SE**	***p***	**Estimate**	**SE**	***p***	**Estimate**	**SE**	***p***	**Estimate**	**SE**	***p***
(Intercept)	0.38	0.21	.069	-0.29	0.20	.136	0.04	0.28	.889	-0.22	0.20	.272
Time (linear)	-0.48	2.17	.824	-4.03	5.75	.483	13.32	5.76	**.021**	3.61	4.85	.457
Time (quad.)	3.22	0.81	**<.001**	7.44	1.59	**<.001**	0.01	1.46	.995	5.89	1.18	**<.001**
Question type	-0.04	0.18	.826	0.46	0.19	**.013**	2.82	0.28	**<.001**	5.37	0.20	**<.001**
Time (linear) x Question	0.38	2.12	.859	27.61	5.48	**<.001**	23.31	5.76	**<.001**	5.06	4.85	.297
Time (quad.) x Question	0.19	0.81	.813	15.17	1.59	**<.001**	-13.70	1.46	**<.001**	-16.42	1.18	**<.001**

Table 2. *Exp. 2: Results of linear mixed-effects regression analysis by ROI, conditions (4)-(5).*

	Who$_{NOM/ACC}$			*girl*$_{ACC}$ / *boy*$_{NOM}$			*kissed*			*at school?*		
	Estimate	***SE***	***p***	***Estimate***	***SE***	***p***	***Estimate***	***SE***	***p***	***Estimate***	***SE***	***p***
(Intercept)	-0.08	0.42	.849	0.32	0.33	.343	0.05	0.28	.848	-0.06	0.82	.938
Time (linear)	2.33	2.34	.319	-1.24	4.18	.766	3.82	3.56	.283	-3.39	6.00	.572
Time (quad.)	-2.48	2.37	.295	0.75	3.17	.812	0.64	2.87	.823	6.34	2.49	**.011**
Question type	-0.26	0.42	.534	0.25	0.33	.460	1.48	0.28	**<.001**	1.99	0.82	**.016**
Time (linear) x Question	0.29	2.34	.902	2.66	4.18	.524	10.01	3.56	**.005**	3.11	6.00	.604
Time (quad.) x Question	-0.89	2.37	.708	1.32	3.17	.677	3.06	2.87	.287	-0.80	2.49	.749

Table 3. *Exp. 2: Results of linear mixed-effects regression analysis by ROI, conditions (7)-(8).*

	Who$_{NOM/ACC}$			kissed			*girl*$_{ACC}$ / *boy*$_{NOM}$			*at school?*		
	Estimate	***SE***	***p***	***Estimate***	***SE***	***p***	***Estimate***	***SE***	***p***	***Estimate***	***SE***	***p***
(Intercept)	0.53	0.37	.150	-0.15	0.41	.714	0.44	0.35	.212	0.26	0.45	.557
Time (linear)	-0.18	2.38	.938	-8.22	4.41	.063	-9.34	4.69	**.046**	-4.67	4.33	.281
Time (quad.)	-0.78	2.34	.737	2.77	2.73	.311	1.65	3.44	.632	-3.17	3.07	.301
Question type	0.59	0.28	**.036**	0.22	0.35	.530	0.11	0.35	.762	0.88	0.44	**.046**
Time (linear) x Question	-4.05	2.37	.088	2.24	3.22	.487	-10.46	4.69	**.026**	5.56	4.14	.179
Time (quad.) x Question	-0.48	2.34	.838	2.68	2.74	.327	5.48	3.44	.111	5.08	3.07	.098