Eye movements during spoken word recognition in Russian children

Irina A. Sekerina [*], Patricia J. Brooks

Department of Psychology, College of Staten Island and Graduate Center of the City University of New York, Staten Island, NY 10314, USA

Abstract

This study explores incremental processing in spoken word recognition in Russian 5- and 6-year-olds and adults using free-viewing eye-tracking. Participants viewed scenes containing pictures of four familiar objects and clicked on a target embedded in a spoken instruction. In the cohort condition, two object names shared identical three-phoneme onsets. In the noncohort condition, all object names had unique onsets. Coarse-grain analyses of eye movements indicated that adults produced looks to the competitor on significantly more cohort trials than on noncohort trials, whereas children surprisingly failed to demonstrate cohort competition due to widespread exploratory eye movements across conditions. Fine-grain analyses, in contrast, showed a similar time course of eye movements across children and adults, but with cohort competition lingering more than 1 s longer in children. The dissociation between coarse-grain and fine-grain eye movements indicates a need to consider multiple behavioral measures in making developmental comparisons in language processing.
Published by Elsevier Inc.

Keywords: Spoken word recognition; Cohort effect; Free-viewing eye tracking; 5- and 6-year-olds; Russian; Developmental dissociation

Introduction

Spoken word recognition is fundamental to language use in our everyday lives. Adults perceive speech effortlessly, with an amazing speed of as many as 50 sounds per second in

[*] Corresponding author. Fax: +1 718 982 4114.
  E-mail address: sekerina@mail.csi.cuny.edu (I.A. Sekerina).

their native language (Foulke & Sticht, 1969). This efficiency occurs despite the fact that the speech signal is nearly always ambiguous and often unclear. Current theories of spoken word recognition (e.g., McQueen, 2004) propose that perceptual ambiguity results in words competing with one another during recognition. In the cohort model and its later developments (Gaskell & Marslen-Wilson, 2001; Marslen-Wilson & Tyler, 1980; Marslen-Wilson & Warren, 1994), all words conforming to the incoming sound sequence become active. The competition among lexical candidates occurring when the initial phonemes of spoken words activate other words sharing the same phonemes is known as the cohort effect.

The cohort effect has become an especially fruitful topic for investigation, with advancement of free-viewing eye tracking during the past decade (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995; Tanenhaus & Trueswell, 2005) providing a powerful tool for investigating the time course of spoken word recognition. Allopenna, Magnuson, and Tanenhaus (1998) were the first to investigate spoken word recognition in English-speaking adults using this method. In their Experiment 1, listeners were presented with pictures of four objects on a computer display and were instructed to move them using the mouse. Across trials, the target object (*beaker*) was contrasted with a phonological cohort competitor (*beetle*) or with an unrelated distractor (*carriage*) located in the same position as the cohort competitor in the display. A robust cohort effect was found; starting on average at 200 ms after the onset of the target word in speech, listeners fixated the target and cohort competitor objects more often than the unrelated distractor. Later, at 400 ms after the onset of the target, looks to the target object diverged from the cohort competitor. Thus, the probability of fixating the cohort competitor and distractor objects, as opposed to the target object, depended on the competition between lexical candidates driven by the amount of phonological overlap between the word onsets whose lexical representations were activated. The cohort effect in young adult English-speaking listeners was estimated to be resolved between 200 and 400 ms after the onset of the spoken target. Subsequent eye-tracking studies have shown that listeners' speed of recognition is affected by target frequency (Dahan, Magnuson, & Tanenhaus, 2001), by context in the form of the gender of the preceding article (Dahan, Swingley, Tanenhaus, & Magnuson, 2000), and possibly by vocabulary size. The cohort effect has also been observed in bilinguals involving competition between phonologically similar words from two different languages (Marian & Spivey, 2003; Weber & Cutler, 2004).

Recently, the free-viewing eye-tracking paradigm has been adapted to study the development of spoken language comprehension in preschool children (Snedeker & Trueswell, 2004; Trueswell, Sekerina, Hill, & Logrip, 1999). Fernald and colleagues (Fernald, Perfors, & Marchman, 2006; Fernald, Pinto, Swingley, Weinberg, & McRoberts, 1998; Swingley, Pinto, & Fernald, 1999) have developed a variation of this method known as the listening-while-looking procedure that is suitable to investigate developmental changes in spoken word recognition in infants from 15 to 25 months of age. This procedure reduces task demands, as necessitated by what infants can endure, by presenting very small numbers of highly familiar word stimuli, a small number of trials, a forced-choice between only two objects, and slower speech. In the listening-while-looking task, an infant sitting on a parent's lap views two computer monitors. Pictures of two objects—a target and a distractor—are presented in pairs, one on each of the two monitors. After a silent previewing period of 3–4 s, a spoken stimulus of the type "Where's the [target]?" is played through a loudspeaker. Accuracy and speed of spoken word recognition is assessed in terms of

infants' shifts of gaze to the named target object during the time window of 300–1800 ms after the onset of the target noun. Because it is a forced-choice task, at the onset of the target noun the infant is typically looking at one of the two pictures. Therefore, accuracy is measured separately for distractor-initial trials, where the infant was initially looking at the distractor (with a shift to the target expected), as opposed to target-initial trials, where the infant happens to have initially fixated the target picture (with no shift expected). Speed is measured as the mean latency of the shift from distractor to target on distractor-initial trials. Fernald et al. (1998) observed that 24-month-olds, like adults, shifted their gaze from the distractor to the target before the offset of the target word. Even more impressively, Swingley et al. (1999) showed that infants of the same age not only incrementally processed speech but also exhibited the same cohort effect found in adults. In trials with phonologically unrelated words (e.g., *doggie* vs. *tree*), infants reliably shifted their gaze to the target at approximately 600 ms after the onset of the target noun and prior to its offset (i.e., the average duration of the target nouns was 973 ms). In cohort trials with phonological overlap (e.g., *doggie* vs. *doll*), infants required an additional 272 ms on average to shift their gaze to the target.

Despite the fact that the free-viewing eye-tracking and listening-while-looking paradigms are now firmly established as ideal methods for exploring spoken word recognition in different age groups, there is still a large gap in terms of our understanding of how proficiency in spoken word recognition develops between infancy and adulthood. Garlock, Walley, and Metsala (2001) observed that although spoken word recognition is presumed to underlie important developmental processes such as phonological awareness and reading ability, "less is known about how children represent and process spoken words" (p. 468). Fernald et al. (2006) went further to argue that the development of spoken word recognition in general has a direct impact on both vocabulary growth and the emergence of grammatical abilities. Walley and colleagues (e.g., Metsala & Walley, 1998; Walley, 1993) have taken the position that adult levels of spoken word recognition take a long time to develop, with young children maintaining more holistic word representations than adults. The empirical support for this position comes from studies that employ offline methods (e.g., gating, word repetition, lexical decision) and rely on secondary tasks to measure developmental advances in children's word recognition as opposed to assessing it directly in comprehension. The demands of these tasks have generally restricted participation to first graders and older children (Garlock et al., 2001; Metsala, 1997), leaving early childhood mostly unexplored.

In the current study, we provide new findings on how spoken word recognition develops for an age group between infancy and adulthood that has not been investigated before, namely, kindergarten-age children. The two experiments investigate the speed and efficiency of spoken word recognition in Russian 5- and 6-year-olds and were designed with four goals in mind: (a) to investigate the cohort effect with young children, (b) to explore developmental differences in spoken word recognition by comparing children with adults, (c) to test the feasibility of using the free-viewing eye-tracking method and a rigorously controlled adult-like design to study spoken word recognition in children, and (d) to use more challenging speech materials from Russian, a language with complex phonology, an abundance of multisyllabic words, and unpredictable lexical stress patterns, to extend research on the cohort effect observed in eye-tracking experiments in languages other than Russian. Experiment 1 investigates whether Russian 5- and 6-year-olds are sensitive to phonological similarity in word onsets, that is, whether they experience cohort

competition among lexical competitors. Experiment 2 tested young adults to explore developmental differences in eye movement patterns between children and adults that might reflect a dissociation of controlled eye movements in visual world exploration and automatic eye movements driven by relentless lexical activation (access/competition). In addition, Experiment 2 provides a replication with monolinguals of the within-language cohort effect reported in Russian–English bilingual adults (Marian & Spivey, 2003, Experiment 2).

We motivate our choice of Russian by selecting experimental stimuli with diverse characteristics involving a wide variety of phonological onsets, multisyllabic words, and an abundance of non-initial-stress prosodic templates. All of these factors make Russian strikingly different from English, where spoken word recognition using eye-tracking has relied so far on monosyllabic words (*doll, tree*, and *truck* in Swingley et al., 1999) or short disyllabic words with strong–weak prosodic structure (e.g., *beaker* in Allopenna et al., 1998; *pitcher, penguin*, and *windmill* in Dahan et al., 2001; *baby, birdie*, and *monkey* in Fernald et al., 2006). Mattys and Samuel (2000) showed that non-initial-stress words in English require additional processing as assessed via phoneme and syllable detection tasks. Demonstrating that spoken word recognition proceeds smoothly and swiftly in the face of complex and varied prosodic templates will strengthen the evidence for universal constraints in speech processing.

Experiment 1: Russian 5- and 6-year-olds

The current study modified the methodology of Marian and Spivey (2003, Experiment 2) to allow computerized presentation of stimuli. They tested 14 bilingual English–Russian participants in a monolingual Russian mode; that is, the entire experiment and all interactions with the experimenter were conducted in Russian. Instead of using pictures of objects as in previous English studies, they presented their participants with miniature real objects and asked them to move the objects around on the table. Marian and Spivey conducted only a coarse-grain analysis to assess the difference in the proportion of cohort trials in which the listeners made eye movements to the competitor compared with the proportion of noncohort trials in which the listeners made eye movements to the distractor in the same location. They found that bilingual participants operating in a monolingual Russian mode looked more at the cohort competitor (18% of trials) than at the unrelated distractor (5% of trials). This difference, although significant only in the participant-based analysis, extended to Russian the cohort effect observed in eye-tracking experiments in other languages.

The current study differs from Marian and Spivey (2003) in several ways. First, we conducted both Experiments 1 and 2 with monolingual Russian speakers in Russia, thereby eliminating any possible effects of bilingualism from the results. Second, we presented a larger set of stimulus items in the form of colorized line drawings on a computer (see below for details). Third, to examine the time course of target selection, we conducted a fine-grain analysis of eye movement patterns. To preview the results, the fine-grain analysis provided a view of children's word recognition abilities very different from that of the coarse-grain analysis. Finally, we did not preexpose children or adults to any of the names of the pictures to allow lexical selection to involve the entire lexicon as opposed to a limited response set. Following the main instruction to select the target object, we also added

production trials to ensure that participants—both children and adults—recognized and could name the cohort competitor.

Method

Participants

A total of 36 monolingual Russian-speaking children (21 girls and 15 boys, mean age = 6 years 2 months, range = 5 years 0 months to 6 years 11 months) were recruited and tested at a large preschool and kindergarten center in Moscow. We obtained informed consent from the parents of all participating children using a written consent form (in Russian). Some parents were unobtrusively present during the experiment with their children. Prior to the experiment, all of the children were screened for language disorders by a speech language therapist using existing Russian assessment tests. Because 4 of the children did not complete the entire experiment, the final sample consisted of 32 children.

Design and materials

The materials were 20 experimental and 2 practice sets of picture stimuli (for the complete set of materials, see Appendix A). Each stimulus set consisted of four pictures: a target object (e.g., *banka* 'jar'), a competitor object whose name in Russian overlapped with the name of the target object (e.g., *bant* 'bow') or distractor objects whose names did not overlap with the name of the target object (e.g., *veer* 'fan' or *vilka* 'fork'), and two phonologically unrelated filler objects (e.g., *topor* 'axe' or *korona* 'crown') (see Fig. 1). Each stimulus set was used in four different lists, varying the target (e.g., *banka* 'jar' or *bant* 'bow') and whether the trial was assigned to the cohort or noncohort condition. For example, Fig. 1A was used for a cohort trial with bow or jar as the target, Fig. 1B was used for a noncohort trial with bow as the target, and Fig. 1C was used for a noncohort trial with jar as the target. The initial locations of the target and the cohort competitor/noncohort distractor, as well as their destination locations, were counterbalanced across trials. A total of 8 children were run on each list, comprising 10 trials in the cohort condition and 10 trials in the noncohort condition, with each trial involving a unique stimulus set.

We selected 20 pairs of nouns that referred to concrete objects and were easily rendered in black-and-white line drawings. Of these, 13 sets were created specifically for the experiment and the remaining 7 were a picture version of the miniature object sets of Marian
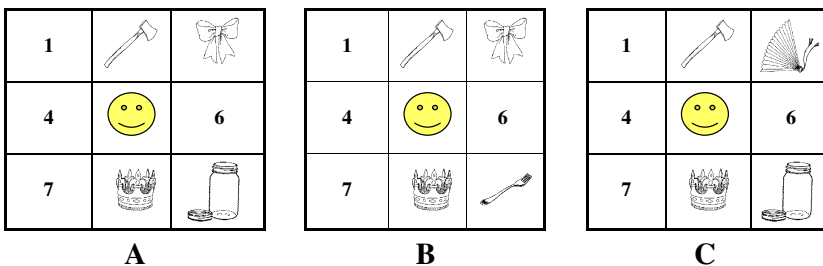

Fig. 1. Stimulus displays for cohort (A) and noncohort (B,C) conditions.

and Spivey (2003, Experiment 2). In one of their sets (e.g., *barkhat* 'velvet'–*baran* 'ram'), *barkhat* 'velvet' was replaced with *baraban* 'drum' because it was impossible to successfully render velvet in a picture format. Inclusion of Marian and Spivey's materials allowed direct comparison of their results with bilingual adults operating in a monolingual Russian mode with the results of our experiments.

The 20 pairs of target and competitor nouns were chosen in such a way that the mean number of overlapping phonemes at onset for the target and competitor was three phonemes (e.g., *banka*–*bant* 'jar–bow'). Two pairs, both from the Marian and Spivey materials, had only two overlapping phonemes (e.g., *šapka*–*šarik* 'hat–balloon'). Together with two distractors and two filler pictures, the entire set of experimental stimuli consisted of 120 line drawings, most of which were selected from Cycowicz, Friedman, Rothstein, and Snodgrass (1997), along with a few additional drawings created specifically for this experiment.

The average word frequencies of the 120 nouns were computed using the new online National Russian Corpus, which has 100 million lexical items (Nacional'nyj korpus russkogo jazyka, 2005). Counterbalancing the assignment of each member of a pair to be target or competitor (e.g., *banka* 'jar' vs. *bant* 'bow') effectively controls for effects of lexical frequency in cohort competition. To determine whether the names of the selected objects were a part of the vocabulary of a typical Russian child, we also examined free association data for the nouns of interest collected from 800 Russian first graders with a mean age of 7 years 6 months (cf. Ovčinnikova, Berseneva, & Dubrovskaya, 2000). These data, shown in Appendix B, represent the mean occurrence of each of the 120 nouns in free association norms for Russian children. To the best of our knowledge, this is the only available source that approximates age of acquisition data for Russian. Only 3 items (i.e., *buton* 'rosebud', *buben* 'tambourine', and *medusa* 'jellyfish') failed to occur in the free association norms; the mean occurrence was 6.1, ranging from 1 to 18. We can conclude that the overwhelming majority of nouns selected for the experiment were well established in the vocabulary of Russian 5- and 6-year-olds.

The spoken instructions presented along with the pictures consisted of three sentences:

(1a) Pokaži, gde zdes' banka.
Show where here jar$_{FEM}$
'Show where the jar is.'
(1b) A teper' pokaži, gde zdes' korona.
And now show where here crown$_{FEM}$
'And now show where the crown is.'
(1c) A teper' nazovi dve drugie kartinki.
And now name two other pictures
'And now name the other two pictures.'

The experimental instruction was always the first one (1a) and contained a carrier phrase and a target noun (e.g., *banka* 'jar'). The second instruction (1b) required a participant to locate one of the two filler objects, and the third one (1c) was intended to elicit the names of the remaining two objects. Crucially, one of the objects to be named in the third instruction for the cohort trials was always the competitor (e.g., *bant* 'bow'), allowing us to assess participants' knowledge of the names for the experimental target–competitor nouns across the four different lists. Thus, the first two instructions assessed spoken word comprehension, whereas the last one assessed word production through picture naming.

Instructions were recorded by a female native speaker of Russian (the first author) using mono-mode sampling at 22,050 Hz. All of the instructions were spoken with normal adult speed, with each sentence recorded individually. No splicing was used so as to preserve naturally occurring coarticulation effects between the last word of the carrier phrase and the target noun. The mean duration of the carrier phrase *Pokaži, gde zdes'* . . . 'Show where here . . .' was 1633 ms, and the mean duration of the target noun was 527 ms.

Procedure

Children were tested individually in a quiet room of their preschool. An undergraduate research assistant brought the child from the classroom and talked to him or her about the study for a few minutes. Prior to conducting the experiment, each child was familiarized with the equipment and the task requirements and was given an opportunity to play with the computer mouse. After obtaining the child's oral consent, the visor was positioned on the child's head and a short eye-tracking calibration was performed.

We used a head-mounted free-viewing eye-tracker (ETL-500, ISCAN) to record children's eye movements during the experiment. Eye movements were sampled at a rate of 30 frames per second and were recorded on a digital Sony DSR-30 videotape recorder. Auditory stimuli were played through the speakers and were recorded simultaneously with eye movements. The eye-tracking portion of the experiment lasted approximately 20 min.

The child was seated at a child-size table facing a laptop computer with a 15-in. LCD screen on which the stimuli were presented (Fig. 2). The first experimenter was seated to the right of the child so that she could monitor the child's behavior during the experiment and provide encouraging feedback. The second experimenter and the eye-tracking equipment were stationed behind the child, as far away as the cable connecting the visor and the eye-tracker would allow. The speakers were located on the child's table on either side of the laptop. Crucially, there were no ''ostensive'' trials prior to the experiment (cf. Swingley et al., 1999); thus, no priming of the words occurred, effectively requiring participants to consider the entire lexicon in word recognition.

Fig. 2. Experimental setup.

The stimulus materials for each trial (four pictures and three spoken instructions) were programmed into an interactive presentation using Macromedia Flash 8 software that allowed for smooth integration of visual and auditory stimuli and was response contingent. Each trial unfolded as in Instruction (1), starting with the presentation of a three-by-three matrix on the laptop monitor containing a bright yellow smiley face in Position 5 (see Fig. 1). Initially, only the smiley face was visible, and it blinked three times at 1-s intervals. We expected that the rhythmical blinking of the smiley face would capture the child's attention and hold his or her gaze until the onset of the first spoken instruction. After 3 s, in addition to the smiley face that remained on the screen, four black-and-white pictures appeared in four different locations in the matrix. The first experimental instruction (i.e., *Pokaži, gde zdes' banka* 'Show where the jar is') was played simultaneously with the appearance of the four pictures. The child was free to scan the display while listening to the carrier phrase prior to the occurrence of the target noun near the end of the first instruction. This provided the child with some time to inspect the pictures.

The child used a child-size mouse attached to the laptop to point and click on the pictures. When the child moved the mouse and the cursor crossed the black line outlining a picture, the white picture with black outline changed from white to a light colorized version. When the child clicked on the picture, the light colorized version was replaced with a full-color version of the picture and the feedback instruction was played (*Pravil'no* 'correct' or *Nepravil'no, podumaj ešče* 'incorrect, think again'). The other three pictures remained black and white. After an incorrect response, the Macromedia Flash program paused until the child attempted to locate the correct picture for the second time. After a correct response, the program proceeded to present the second instruction (i.e., *A teper' pokaži, gde zdes' korona* 'And now show where the crown is'). The procedure was repeated for the second instruction, and the same feedback was provided. Finally, the third instruction requested that the child name the remaining two black-and-white pictures, which also changed into full-color version when clicked on. The child's picture-naming responses were recorded manually by the experimenter along with any self-corrections or comments. At that point, the program revealed a red right-pointing arrow in the corner of the laptop monitor, and the child was trained to click on it to move to the next trial. (The Macromedia Flash presentation of the experiment can be viewed on our website at http://163.238.8.180/~sekerina/research.html.)

Data treatment

Trueswell et al. (1999, p. 102) argued that it is important to analyze the children's performance in eye-tracking experiments using two types of measures: relatively coarse-grain measures of eye movement patterns that give a general idea of how children investigate visual scenes and a more fine-grain record of the moment-by-moment of fixations over time. Following this practice, we analyzed three types of data: (a) accuracy in comprehension (clicking on the target object following the first instruction) and in production (naming the competitor or distractor and one of the two filler pictures following the third instruction), (b) coarse-grain measures of the children's eye movement patterns (number and proportion of trials with eye movements to the target and cohort competitor/noncohort distractor on hearing the target noun), and (c) fine-grain measures of the children's eye movement patterns (latencies of the first look to the target after its onset and moment-by-moment fixations for each of the pictures in the display with the 33-ms

resolution). The reaction times that resulted from clicking on the target were collected but are not summarized here due to the highly variable nature of mouse clicking that depended on each child's fine motor skill development and familiarity with a computer mouse. All of the trials in which the child did not click on the target on the first attempt were coded as errors and were excluded from the coarse-grain and fine-grain eye movement analyses. We also excluded all of the trials in the cohort condition in which the child did not know the correct name of the competitor, thereby rendering these trials equivalent to the noncohort condition. Due to the differing numbers of correct trials per participant and condition, we used proportions of trials as the dependent variable for the coarse-grain eye movement analysis and used average numbers of fixations to the target and to the competitor/distractor objects as dependent variables in the fine-grain eye movement analyses. Eye movement data were subjected to 2×4 factorial analyses of variance (ANOVAs) with trial condition (cohort vs. noncohort) as a within-participants factor and list as a between-participants factor.

To address the main purpose of this study, pinpointing the exact locus of the cohort effect and investigating the time course of how it gets resolved, we conducted additional (pairwise) analyses of fine-grain eye movements. We calculated average numbers of fixations per trial by comparing pairs of objects of interest: (a) the target versus the phonologically related competitor in the cohort trials, (b) the target versus the phonologically unrelated distractor in the noncohort trials, (c) the target in the cohort versus noncohort trials, and (d) the competitor in the cohort trials versus the distractor in the noncohort trials. These fine-grain analyses served two purposes: First, the more detailed analyses of looks to the target versus the competitor and the target versus the distractor should help us to uncover differences in the two conditions whose rapidly changing nature may have gone undetected in the coarse-grain analysis, thereby providing evidence that children are indeed sensitive to cohort competition. It was expected that there would be no differences in looks between these pairs prior to hearing the target noun. However, on arrival of the target in the instructions, when the cohort competition begins, there should be an increased number of eye movements to the target in both cohort and noncohort conditions and to the competitor in the cohort condition as well as a sharp drop in eye movements to the distractor in the noncohort condition. Second, the most interesting comparison turned out to be that between the competitor and the distractor because it revealed the exact locus of the cohort effect and how it got resolved differently for children and adults in the course of target noun recognition.

Results and discussion

Comprehension and production accuracy

The first instruction in the trial (i.e., *Pokaži, gde zdes' banka* 'Show where the jar is') required the child to click on the picture that depicted the target word. The accuracy data were based on 640 data points (20 trials × 32 participants). The item was scored as correct if the child clicked on the picture that depicted the target on the first attempt. Otherwise, the item was scored as an error even if the child provided the correct response on the second attempt. Children's accuracy was almost at ceiling, with 97% correct overall. Comprehension errors occurred with seven target items (*marka* 'stamp' [31% correct]; *čerep* 'skull' [81% correct]; *buben* 'tambourine', *butylka* 'bottle', *plačš* 'raincoat', and *sovok* 'dustpan'

[94% correct each]). We took these data to indicate that all of the targets, except for *marka* 'stamp', were well established in the children's vocabularies. A total of 20 trials (12 cohort and 8 noncohort trials) in which the children made comprehension errors were removed from coarse-grain and fine-grain eye movement analyses reported below.

Next we examined children's accuracy in naming the cohort competitor. Recall that in contrast to the infant studies of Fernald and colleagues, there were no "ostensive" trials in which the stimuli were identified for the child. Our procedure required children to consider their entire lexicon when naming the objects. The third instruction in the trial (i.e., *A teper' nazovi dve drugie kartinki* 'And now name the remaining two pictures') was included to test the children's knowledge of the competitor to ensure that it indeed was a cohort competitor for the target. That is, it was essential to eliminate any cohort trials for which the expected cohort competition did not arise due to the fact that the child did not access the predicted name for the competitor. The accuracy data for naming the competitor were based on 320 data points (10 cohort trials × 32 participants) because in the noncohort trials the children always named a phonologically unrelated distractor.

Children's accuracy in naming the competitor was substantially lower than their comprehension of the target (79% correct). They produced a total of 67 errors in naming the competitor (with only one trial involving errors in both target comprehension and competitor production). Children provided names other than the expected ones at least once for 18 of 20 competitor nouns. Two competitor nouns, *buton* 'rosebud' and *marka* 'stamp', were never named with the expected words ('rosebud' responses were *cvetok* 'flower' and *roza* 'rose'; 'stamp' responses were *kartinka* 'picture' and *kartina* 'painting'). For four other competitor nouns (i.e., *buben* 'tambourine', *čerep* 'skull', *plačš* 'raincoat', and *medusa* 'jellyfish'), children failed to produce any label at all in half of the trials. An additional 12 competitor nouns were given names other than the expected ones at least once. Some of these naming errors are easily explained by the fact that the words were of very low frequency (e.g., 'rosebud', 'tambourine', 'jellyfish'), as can be seen in the child free association norms (Appendix B). Others most likely stemmed from children's difficulties in identifying the objects rendered in picture format (e.g., 'stamp', 'raincoat', 'ram').

Adding together the trials involving errors in comprehension and production, 78 cohort and 8 noncohort trials were eliminated, leaving 242 cohort trials (75.6%) and 312 noncohort trials (97.5%) of the original 640 trials for eye movement analyses.

Coarse-grain eye movement analysis

In the previous spoken word recognition experiments with objects or pictures (Dahan et al., 2000; Marian & Spivey, 2003), the coarse-grain measure of eye movements was operationalized as the proportion of fixations to the phonologically related competitor in the cohort condition compared with the unrelated distractor in the noncohort condition, beginning at approximately 200 ms after the onset of the target noun. This adjustment is necessary to accommodate the time necessary for the programming of an eye movement (conservatively estimated to be 180 ms in Altmann & Kamide, 2004). In previous experiments, this critical window extended through the average duration of the target (400 ms in the Dahan et al. (2000) study) to the moment participants performed the required action (Allopenna et al., 1998; Marian & Spivey, 2003). The cohort effect was observed when there were more fixations to the competitor in the cohort condition than to the distractor in the noncohort condition during this time window. Lexical competition between the

target and the competitor is reflected in alternate fixations between the two competing ref-
erents until the ambiguity is resolved. (In the experiments reported in the literature, it hap-
pened after 600 ms on average.)

In our experiment, all of the eye movements to the competitor or the distractor were
summed during the critical window that started 200 ms from the onset of the target
(e.g., *banka* 'jar') in the spoken instruction and ended when the child clicked on the target.
These data were then used to compute the proportion of trials for cohort and noncohort
conditions in which the competitor or the distractor was fixated at some point during the
critical window. If 5- and 6-year-olds explore the visual world during sentence processing
in a manner similar to adults, on hearing the ambiguous portion of the target noun *ban–*
'jar', we would expect them to look more at the competitor (e.g., *bant* 'bow') in the cohort
condition than at the distractor (e.g., *vilka* 'fork') in the noncohort condition. Surprisingly,
we found that the children considered the competitor and the distractor equally often, in
65 and 69% of the cohort and noncohort trials, respectively ($t < 1$). Thus, these eye move-
ment patterns suggest that our Russian children failed to demonstrate cohort competition
in the coarse-grain analysis.

We then explored a possibility that the cohort and noncohort trials might still be dis-
tinguished in the number of separate eye movements to different objects within each trial,
with more eye movements in the cohort trials than in the noncohort trials. An increased
number of eye movements in the cohort trials might reflect more initial competition
between the target and the competitor. But again we found no difference in the cumulative
number of eye movements per trial between the two conditions, 7.1 in the cohort condition
and 7.4 in the noncohort condition, F_1 and $F_2 < 1$. Thus, the overall pattern of coarse-
grain analysis of eye movements, as reflected in similar proportions of trials with fixations
to the competitor in the cohort condition and to the distractor in the noncohort condition,
did not reveal the cohort effect that was observed in adults in studies testing English
(Allopenna et al., 1998; Dahan et al., 2000) and bilingual Russian–English (Marian & Spi-
vey, 2003) participants.

Fine-grain eye movement analyses

The coarse-grain eye movement data failed to provide any evidence of cohort compe-
tition in spoken word recognition by children. This analysis, however, lacked temporal res-
olution to allow us to investigate lexical competition effects that are potentially too fast
and short-lived. In addition, the cohort effect in the coarse-grain eye movement data could
be masked by the fact that young children in general make more eye movements than do
adults (Nation, Marshall, & Altmann, 2003). We turned to fine-grain eye movement anal-
ysis to explore in detail how lexical competition develops over time.

Our first glimpse of the presence of a cohort effect in children came only when we com-
puted the latencies of the first eye movement to the target using the same critical window as
in the coarse-grain measure, that is, starting at 200 ms from the onset of the target (e.g.,
banka 'jar') in the spoken instruction. Children were indeed sensitive to cohort competition;
they launched the first look to the target significantly faster in the noncohort trials (averag-
ing 627 ms from the onset of the target noun) than in the cohort trials (averaging 823 ms
from the onset of the target noun), $F_1(1, 28) = 6.01$, $p < .02$, and $F_2(1, 19) = 5.83$, $p < .03$.

To explore the time course of lexical activation in a fine-grain temporal way as
the instruction unfolded, we followed Altmann and Kamide's (2004) methodological

suggestions and computed the average number of eye movement toward each object of interest—the target, the competitor (in cohort trials), and the distractor (in noncohort trials)—for each 33-ms frame (i.e., the resolution of the eye-tracking equipment) starting with the onset of the second word in the instructions (i.e., *gde* 'where'). Because the children often produced eye movements to other objects that were not included in the analysis (e.g., the two filler pictures, the smiley face, looks between the pictures, looks away from the display), the average proportions of eye movements to the target and competitor/distractor were not in complementary distribution. In all of the analyses that follow, we directly compared the average number of eye movements to the target with that to the competitor/distractor objects using an ANOVA with participants and items as the random factors (F_1 and F_2, respectively). Trial condition (cohort vs. noncohort) was a within-participants factor (F_1) or a within-items factor (F_2), and list was a between-participants factor (F_1 only).

Two temporal regions of interest were identified. Region 1 was defined as the carrier phrase *Pokaži, gde zdes'* ... 'Show where here is ...', which had an average duration of 1633 ms.

As expected, there were no differences between the average proportions of eye movements to the target and those to the competitor/distractor in Region 1 (F_1 and $F_2 < 1$), nor was there any effect of trial condition (F_1 and $F_2 < 1.3$). The effect of list was not significant, and there were no interactions. Thus, prior to hearing the target noun, children showed similar exploratory movements across cohort and noncohort trials and did not show any preference for the target object versus the competitor/distractor object.

Region 2 was defined as the target noun. This region started right at the end of the carrier phrase at 1634 ms on average; there was a negligible amount of variation for the onset of Region 2 because Region 1, the carrier phrase, was identical for all of the trials. The average duration of the target noun was 527 ms (range = 267–733 ms). Any difference between the cohort and noncohort conditions was anticipated to occur in Region 2 (i.e., starting at 200 ms after its onset due to the time required for the programming of eye movements), but we were uncertain how early and/or short-lived the effect would be. Our main interest was in how the patterns of eye movements would unfold over time, so we plotted the data against a time axis in 33-ms intervals. Fig. 3 shows the average proportions of eye movements launched to the target (e.g., *banka* 'jar') versus the competitor (e.g., *bant* 'bow' in the cohort condition) or the distractor (e.g., *veer* 'fan' in the noncohort condition) for the two trial conditions in Region 1 (e.g., ... *gde zdes'* ... '... where here ...') and Region 2 (the target).

To better localize the time course of the cohort effect, we divided Region 2, the target noun, into three 200-ms bins from the onset of the target noun, corresponding to 1–200, 201–400, and 401–600 ms, respectively. Fig. 4 plots only the Region 2 fragment of Fig. 3, with the *y* axis representing the average numbers of individual eye movements directed toward the target and the competitor/distractor during each of the 200-ms bins of the target noun.

Similar to the results for Region 1, the fine-grain analysis comparing looks to the target versus the competitor/distractor in the first 200-ms bin of Region 2 produced no significant effects (all *F*s < 1.5). Not surprisingly, children did not distinguish between cohort and noncohort trials in their eye movements to the target versus the competitor/distractor at the beginning of the target noun due to the time lag needed for programming of eye movements.
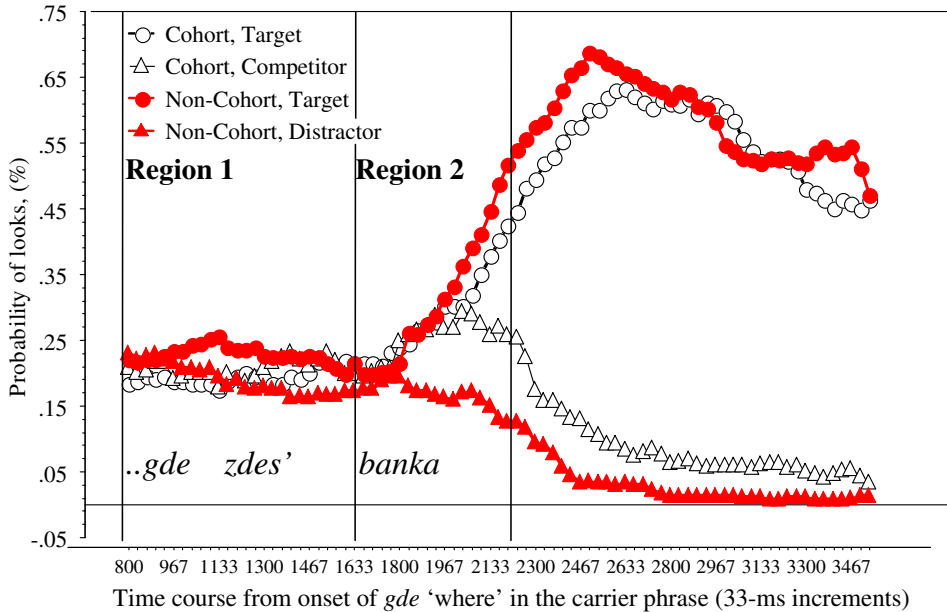
Fig. 3. Experiment 1 (children): Probabilities of eye movements to the target (e.g., *banka* 'jar'), the phonologically related competitor in the cohort condition (e.g., *bant* 'bow"), and the phonologically unrelated distractor in the noncohort condition (e.g., *veer* 'fan') in Region 1 (e.g., ...*gde zdes'* ... '... where here ...') and Region 2 (the target noun) (in 33-ms increments).
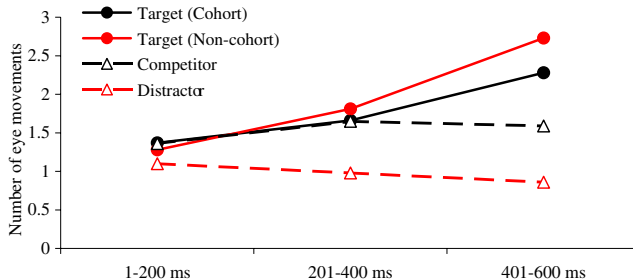

Fig. 4. Experiment 1 (children): Mean numbers of eye movements to the target, the competitor, and the distractor in three 200-ms bins of Region 2, the target noun (e.g., *banka* 'jar').

Lexical competition due to the cohort effect emerged only during the second 200-ms bin of Region 2, where we observed a main effect of object (target vs. competitor/distractor), $F_1(1, 28) = 5.54$, $p < .05$, and $F_2(1, 19) = 4.94$, $p < .05$, mediated by a significant interaction of object and trial conditions, $F_1(1, 28) = 7.02$, $p < .02$, and $F_2(1, 19) = 7.44$, $p < .02$. There was no effect of list or any interaction involving this factor. As shown in Fig. 3, during the second 200-ms bin from the onset of the target noun, children looked equally often at the target objects and competitor objects on cohort trials, F_1 and $F_2 < 1$, but they looked significantly more often at the target than at the distractor on noncohort trials, $F_1(1, 28) = 14.7$, $p < .001$, and $F_2(1, 19) = 7.91$, $p < .02$. Furthermore, there was no

difference in the number of eye movements to the target in cohort and noncohort trials, F_1 and $F_2 < 1$, but children looked significantly more at the competitor in cohort trials than at the distractor in noncohort trials, $F_1(1, 28) = 12.3$, $p < .001$, and $F_2(1, 19) = 9.03$, $p < .01$. Thus, 200 ms after the onset of the target noun, children were more successful in eliminating the distractor object from consideration in noncohort trials than in eliminating the competitor object in cohort trials. The persistence of lexical competition between the target and the competitor on cohort trials provides evidence of a cohort effect in children's spoken word recognition.

The mean offset of the target noun (average duration of 527 ms) occurred during the final 200-ms bin of Region 2, with the offsets of only four target words (10%) extending beyond the endpoint of Region 2. In the final 200 ms of the target noun, we again observed a highly significant main effect of object (target vs. competitor/distractor), $F_1(1, 28) = 46.12$, $p < .001$, and $F_2(1, 19) = 22.71$, $p < .001$, which was again mediated by a significant interaction of object and trial conditions, $F_1(1, 28) = 11.68$, $p < .01$, and $F_2(1, 19) = 15.65$, $p = .001$. However, in contrast to the previous 200-ms bin, children produced significantly more eye movements to the target noun than to the competitor/distractor on both cohort trials, $F_1(1, 28) = 7.63$, $p = .01$, and $F_2(1, 19) = 6.15$, $p < .05$, and noncohort trials, $F_1(1, 28) = 51.80$, $p < .001$, and $F_2(1, 19) = 30.21$, $p < .001$. However, comparing eye movements to the target only, children produced significantly fewer looks to the target on cohort trials than on noncohort trials, $F_1(1, 28) = 4.19$, $p = .05$, and $F_2(1, 19) = 5.01$, $p < .05$, as well as significantly more looks to the competitor on cohort trials than to the distractor on noncohort trials, $F_1(1, 28) = 15.40$, $p = .001$, and $F_2(1, 19) = 9.71$, $p < .01$. In addition to these effects, there was a weak but significant main effect of list, $F_1(3, 28) = 3.39$, $p < .05$, which was due to a larger average number of eye movements overall for the 8 children assigned to List 3 than for the 8 children assigned to List 2. Importantly, the main effect of list did not interact with any of the effects of interest summarized above. In sum, in contrast to the coarse-grain analyses reported above, the fine-grain analyses indicated that children's eye movement patterns differed reliably as a function of trial condition as soon as the programming of eye movements was completed, that is, starting at 200 ms after the onset of the target noun.

Finally, to take full advantage of the 33-ms resolution of the eye tracker, we attempted to locate, for cohort and noncohort trials, the point at which eye movements to the target were significantly more frequent than eye movements to the competitor/distractor. Our goal was to pinpoint, for each trial condition, the average moment when lexical competition is resolved, thereby allowing us to infer the full time course of the cohort effect. For cohort trials, the point of disambiguation was 495 ms from the onset of the target noun (Frame 64, at 2129 ms from the trial onset, with the carrier phrase having an average duration of 1634 ms); at this exact point, the average number of eye movements to the target first exceeded that of eye movements to the competitor, $F_1(1, 28) = 5.59$, $p < .05$, and $F_2(1, 19) = 5.08$, $p < .05$. In contrast, for noncohort trials, disambiguation began much earlier at 198 ms from the onset of the target noun (Frame 55, at 1832 ms from the trial onset), $F_1(1, 28) = 6.37$, $p < .02$, and $F_2(1, 19) = 4.15$, $p < .06$; this was the point at which the average number of eye movements to the target exceeded that of eye movements to the unrelated distractor. That is, it took roughly 300 ms longer for children to identify the target object when it was paired with a cohort competitor.

We finished our examination of the duration of cohort competition by pinpointing the last frame at which the average number of eye movements to the competitor in cohort

trials was significantly greater than that of eye movements to the distractor in the same location in noncohort trials. Children showed significantly more looks to the cohort competitor than to the unrelated distractor until Frame 105 (at 3482 ms from the trial onset), $F_1(1, 28) = 4.11$, $p = .05$, and $F_2(1, 19) = 5.28$, $p < .05$. Thus, children experienced cohort competition for an additional 1353 ms from the point of disambiguation of the noun at Frame 64, indicating lingering activation of the cohort competitor.

Experiment 2: Adults

Method

Participants

A total of 16 monolingual Russian adults (mean age = 21 years, 12 women and 4 men) were recruited from the undergraduate population of the Department of Psychology at Moscow State University. The students were paid an equivalent of $3 (US) in rubles. They were monolingual speakers of Russian and were naive with respect to the nature of the experiment, which took approximately 15 min to complete.

Design and materials

The materials and design were the same as in Experiment 1.

Procedure

The equipment and procedure were the same as Experiment 1 except that the adults used a regular-size mouse.

Results and discussion

Data were analyzed using the same measures as those reported in Experiment 1: comprehension and production accuracy and coarse-grain eye movement analysis, and fine-grain eye movement analysis.

Comprehension and production accuracy

The accuracy data are based on 320 cohort and 160 noncohort trials. Adults were at ceiling for comprehension accuracy (100%). However, they made 9 errors (out of 160) in production, resulting in 94% accuracy in naming the cohort competitor in cohort trials. They provided names other than the expected ones for 3 items. As in the children's production accuracy data, the competitor noun *buton* 'rosebud' was never named with the expected word (all of the responses were 'rose'). Three remaining competitor nouns that resulted in errors were *plačš* 'raincoat' (2 errors), *medusa* 'jellyfish' (2 errors), and *gruzovik* 'truck' (1 error). In addition to the 9 trials with errors, 3 trials (1 in the cohort condition and 2 in the noncohort condition) were lost due to computer error, leaving 150 for cohort trials (93.8%) and 158 for noncohort trials (98.8%) of the original 320 trials for eye movement analyses reported below.

Coarse-grain eye movement analysis

The first difference between child and adult data in our study emerged already at the stage of coarse-grain analysis of eye movements. In contrast to the child data, Russian adults did experience the expected lexical competition as reflected in a larger proportion of trials where there were eye movements to the competitor in the cohort condition in comparison with the proportion of trials where there were eye movements to the distractor in the noncohort condition. As described above in Experiment 1, all of the eye movements to the competitor or the distractor were summed during the critical window that started at 200 ms from the onset of the target (e.g., *banka* 'jar') in the spoken instruction and ended when the participant clicked on the target. These data were then used to compute the proportions of trials for cohort and noncohort conditions in which the competitor or the distractor was fixated during the window's duration. We found a significant effect of the competitor type, $t_1(15) = 1.77$, $p < .05$, one-tailed, and $t_2(39) = 1.84$, $p < .05$, one-tailed, with participants directing eye movements toward the competitor in the cohort trials more often than toward the distractor in the same location in the noncohort trials (47.0 and 37.5%, respectively). Thus, the eye movement patterns of Russian adults in our study revealed the cohort effect that was observed in English speakers (Allopenna et al., 1998; Dahan et al., 2002) and replicated the within-language cohort effect observed in Russian–English bilingual adults (Marian & Spivey, 2003, Experiment 2).

Just as in the child data, we found no difference in the cumulative number of eye movements per trial between the two conditions (4.9 in the cohort condition and 4.8 in the noncohort condition).

Fine-grain eye movement analyses

In addition to the coarse-grain analysis that already revealed a reliable cohort effect for the adults, we also conducted fine-grain analyses of eye movements for the adults to provide a direct comparison with the children. Important differences between the adult and the child moment-by-moment eye movement patterns were found.

The first difference between child and adult data involved the absence of a difference in the latency of the first eye movement to the target across the two trial conditions. The adults in general were faster in launching the first eye movement to the target (425 ms on average after the onset of the target) than were the children (725 ms on average). However, although the children's latencies for the two trial conditions were significantly different, with faster looks to the target in the noncohort condition (627 vs. 823 ms), the adults' latencies were not (430 ms in the noncohort condition vs. 420 ms in the cohort condition), Fs < 1.

Next we compared the average number of eye movements to the target with that to the competitor/distractor across the same Regions 1 and 2 as in the children's data.

Region 1, the carrier phrase *Pokaži, gde zdes'* ... 'where here is ...' (0–1633 ms from the onset of the carrier phrase), unexpectedly showed a main effect of trial condition, $F_1(1, 12) = 4.99$, $p < .05$, and $F_2(1, 19) = 3.98$, $p = .06$. That is, prior to hearing the target noun, adults showed increased exploratory movements to the target and competitor in the cohort trials relative to the target and distractor in the noncohort trials. We suggest that this early significant difference in eye movements for cohort versus noncohort trials reflects an awareness of the phonological similarity of the object names. The adults quickly

became aware of the lexical competition even prior to the arrival of the target in the speech and were getting ready to resolve this competition already from the very beginning of the target. There was also an interaction of object and list, $F_1(3, 12) = 4.23$, $p < .05$, due to spurious variation in numbers of eye movements to the target versus the competitor/distractor across lists.

Region 2, the target noun, comprised the same three separate 200-ms bins as in the analysis of children's eye movements (Fig. 5). Fig. 6 plots just the Region 2 fragment of Fig. 5, with the *y* axis representing the average numbers of individual eye movements directed toward the target and the competitor/distractor during each of the 200-ms bins of the target noun.

In the first 200-ms bin of Region 2, the significant but short-lived main effect of the trial condition found for Region 1 disappeared, $F_1(1, 12) = 3.41$, $p < .10$, and $F_2(1, 19) = 2.93$, $p = .10$. Adults and children alike did not distinguish between cohort and noncohort trials in their eye movements to the target versus the competitor/distractor at the beginning of the target.

Lexical competition due to the cohort effect emerged during the second 200-ms bin of Region 2. There was a significant main effect of object (target vs. competitor/distractor), $F_1(1, 12) = 7.51$, $p < .02$, and $F_2(1, 19) = 4.38$, $p = .05$, as well as a main effect of trial condition, $F_1(1, 12) = 15.53$, $p < .01$, and $F_2(1, 19) = 17.90$, $p < .001$. Adults had significantly more eye movements on cohort trials than on noncohort trials, and they produced more
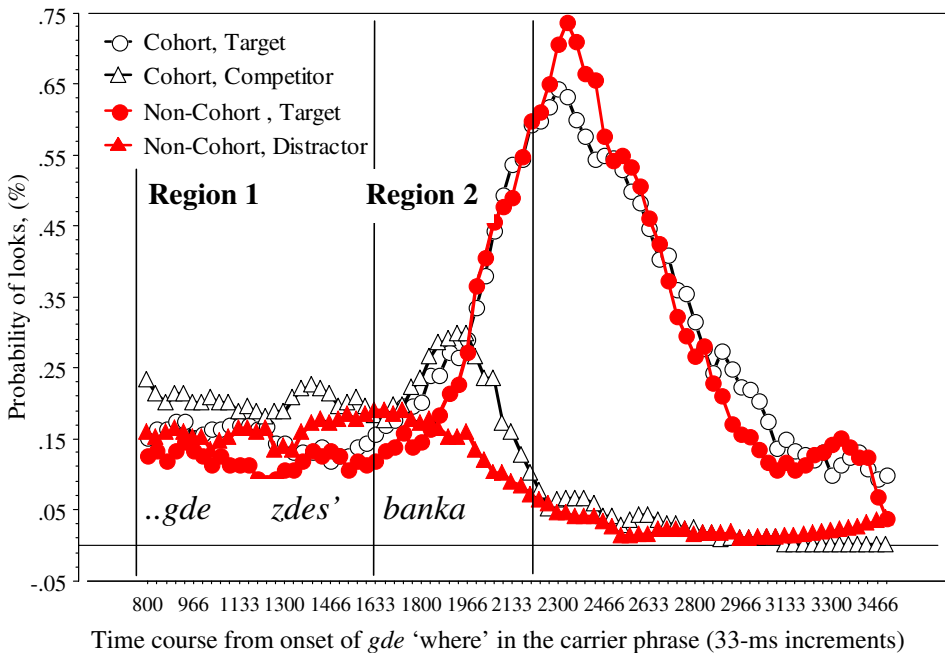

Fig. 5. Experiment 2 (adults): Probabilities of eye movements to the target (e.g., *banka* 'jar'), the phonologically related competitor in the cohort condition (e.g., *bant* 'bow"), and the phonologically unrelated distractor in the noncohort condition (e.g., *veer* 'fan') in Region 1 (e.g., *...gde zdes'* ... '... where here ...') and Region 2 (the target noun) (in 33-ms increments).
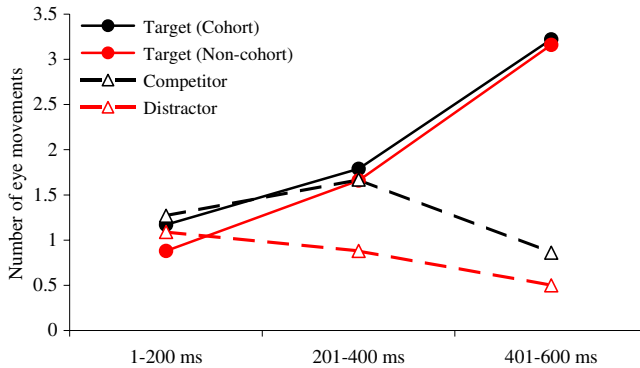
Fig. 6. Experiment 2 (adults): Mean numbers of eye movements to the target, the competitor, and the distractor in three 200-ms bins of Region 2, the target noun (e.g., *banka* 'jar').

eye movements to the target than to the competitor/distractor. There were no other significant main effects or interactions.

In the final 200-ms bin of Region 2, adults showed only a highly significant main effect of object (target vs. competitor/distractor), $F_1(1,12) = 100.56$, $p < .001$, and $F_2(1,19) = 84.92$, $p < .001$, with many more eye movements to the target than to the competitor/distractor across trial conditions.

In sum, the fine-grain analyses indicated that adult eye movement patterns differed reliably as a function of trial condition even before the onset of the target noun. Starting at 200 ms after the onset of the target noun, adults increased looks to the target relative to the competitor/distractor. In contrast to the children, however, the adults failed to show the crucial interaction between trial condition and object (target vs. competitor/distractor) at any point during the trial.

When did the cohort effect get resolved for adults? To answer this question, we compared the moment at which eye movements to the target were significantly more frequent than eye movements to the competitor/distractor for adults and children. Fig. 5 revealed a strikingly similar pattern to that of the children (see Fig. 3). For cohort trials, the point of disambiguation for the adults was at 396 ms from the onset of the target noun (Frame 61, at 2030 ms from the trial onset, with the carrier phrase having an average duration of 1634 ms); at this point, the average number of eye movements to the target first exceeded that of eye movements to the competitor, $F_1(1,12) = 8.90$, $p < .02$, and $F_2(1,19) = 6.43$, $p = .02$. Disambiguation began somewhat earlier for noncohort trials, at 264 ms from the onset of the target noun (Frame 57, at 1898 ms from the trial onset), $F_1(1,12) = 5.56$, $p < .05$, and $F_2(1,19) = 1.69$, $p = .21$; this was the point at which the average number of eye movements to the target first exceeded that of eye movements to the unrelated distractor. Thus, it took adults 132 ms longer on average to select the target when it was paired with a cohort competitor. In comparison, children were slower and required more than twice as much time (297 ms).

Finally, we examined the duration of cohort competition by pinpointing the last frame where the average number of eye movements to the competitor object in cohort trials was significantly greater than the average number of eye movements to the distractor in noncohort trials. Adults made significantly more looks

to the cohort competitor than to the unrelated distractor until Frame 64 (at 2128 ms from the trial onset), $F_1(1, 12) = 5.71$, $p = .05$, and $F_2(1, 19) = 5.06$, $p < .05$. Thus, cohort competition extended for only 99 ms from the point of noun disambiguation at Frame 61. This is in marked contrast to the children who showed lingering activation of the cohort competitor for an additional 1353 ms from the point of noun disambiguation.

General discussion

In two experiments, we investigated the cohort effect in spoken word recognition in monolingual Russian 5- and 6-year-olds (Experiment 1) and compared the results with the control group of young monolingual Russian adults (Experiment 2). Crucially, the task demands and experimental materials were identical in the two studies, allowing for direct comparison of the speed and efficiency of spoken word recognition between the two age groups. Our results show convincingly that the adult-like mechanisms of spoken word recognition are already in place at 5 years of age, with young children not only incrementally processing speech but also exhibiting the same online cohort effect found in adults. This argues against the holistic word representation hypothesis proposed by Walley and colleagues (Garlock et al., 2001; Metsala, 1997; Metsala & Walley, 1998).

The results contribute to three issues concerning patterns of development of processing mechanisms in children's online language comprehension. The first is whether children, like adults, experience and resolve the cohort competition in online spoken word recognition under naturalistic conditions, that is, considering potentially the entire lexicon, a large number of trials, complex lexical items, and conversational speed of speech. The second concerns the behavioral differences between children and adults in online measures, that is, whether children's patterns of eye movements differ from those of adults, reflecting a developmental trajectory in language processing. The third is the importance of methodology in investigating interactions between language processing and cognitive development, with eye tracking becoming a critical tool in uncovering a developmental dissociation between controlled and automatic processes.

First, consider the effect of cohort competition in children. The design of the task created an experimental condition in which the target and the competitor started with the same three phonemes (the cohort condition) that contrasted with a control condition in which the distractor was phonologically unrelated to the target (the noncohort condition). Children's eye movements clearly showed their sensitivity to lexical competition in both latencies and patterns of moment-by-moment eye movements. Children launched the first look to the target an average of 200 ms later in the cohort condition than in the noncohort condition. Fine-grain movement analysis placed the beginning of the cohort effect at 200 ms after the onset of the target, that is, before the noun was completely heard. At that point, children continued to alternate looks between the target and the competitor in the cohort trials but had already stopped looking at the distractor in the noncohort trials. The persistence of lexical competition between the target and the competitor in just the cohort trials provides strong evidence of incremental processing in 5- and 6-year-olds' spoken word recognition.

We also asked whether young children's sensitivity to cohort competition would be influenced by task demands and types of experimental materials. We exposed our child participants to a full set of complex lexical materials identical to the ones used with adults in Experiment 2 without including ostensive trials in a training phase, as has been essential for studies of infant spoken word recognition (Fernald et al., 1998, 2006; Swingley et al., 1999). In addition, in contrast to the English spoken word recognition studies that relied on monosyllabic and short disyllabic words with strong–weak prosodic structure (Allopenna et al., 1998; Dahan et al., 2001; Fernald et al., 2006; Swingley et al., 1999), we used Russian multisyllabic words with non-initial-stress patterns and a wide variety of phonological onsets. Children demonstrated nearly perfect comprehension of the target words (97%) and appropriately fast latencies of eye movements to the target (725 ms after the onset), that is, results that are compatible with the general estimates of accuracy and speed in English spoken word recognition. Thus, our preliminary results for Russian contrast with Mattys and Samuel's (2000) findings for English that non-initial-stress words require additional processing in terms of processing time, accuracy, and/or memory load. Spoken word recognition of longer non-initial-stress Russian words in both children and adults proceeds smoothly and swiftly in the face of complex and varied prosodic templates. A complete testing and verification of the Mattys and Samuel's hypothesis of processing load associated with different stress pattern distributions in Russian remains a matter of future research.

The second issue concerns the developmental differences in spoken word recognition between children and adults. As expected, in terms of latencies of the first eye movements to the target, children were slower than adults (300 ms slower on average). Also, consistent with the findings of Nation et al. (2003), children made more separate eye movements than did adults overall, irrespective of the trial condition. These results conceivably reflect well-established general age-related differences in working memory and allocation of attentional resources. Of greater interest are the specific differences found in coarse-grain and fine-grain analyses of eye movements between the two age groups. The main purpose of this study was to explore how and when children resolve lexical competition in online spoken word recognition in comparison with adults. In general, although both groups demonstrated delayed spoken word recognition for cohort trials, there were a number of systematic differences in eye movement patterns. The first important difference was that adults noticed lexical competition very early, even before the target appeared in the speech stream. This early awareness manifested itself in both fine-grain and coarse-grain eye movements. For the fine-grain eye movements, there were more looks to both the target and the competitor in cohort trials than in noncohort trials in Region 1, during which the adults heard only the end of the carrier phrase, ... *gde zdes'* ... '... where here [is] ...' For the coarse-grain eye movements, the adults considered the competitor in 10% more cohort trials than they considered the distractor in the noncohort trials. This result replicates the magnitude of the cohort effect found in the coarse-grain analysis of eye movements in the English speakers (Allopenna et al., 1998) and Russian–English bilingual speakers (Marian & Spivey, 2003). Children, on the other hand, did not show any differences in eye movements between the cohort and noncohort trials in the coarse-grain analysis, nor did they notice lexical competition until they had actually heard the beginning of the target noun in the speech stream.

The second developmental difference in eye movement patterns involved the time course of cohort competition resolution following the presentation of the target noun. For the adults, the number of looks to the competitor and distractor were indistinguishable from

each other and fell close to zero only 99 ms after the point of target disambiguation. For the children, cohort competition lingered for longer than 1 s (i.e., 1353 ms) after target disambiguation; children continued to look more at the competitor in the cohort trials than at the distractor in the noncohort trials. This result suggests that the children were inefficient at suppressing the activation of the cohort competitor even after they had selected the target. This lingering competition effect resembles the extended time course of interference observed in 5-year-olds in a developmental study of the cross-modal Stroop effect (Hanauer & Brooks, 2003) where participants were instructed to name color patches (e.g., a red square) paired with either a color adjective (e.g., the word *blue*) or a noncolor adjective (e.g., the word *dry*) presented over headphones. Whereas adults and older children showed Stroop-like interference (i.e., longer latencies in naming color patches paired with color adjectives in comparison with noncolor adjectives) only when the auditory distractor and color patch occurred simultaneously, 5-year-olds showed interference even with the distractor preceding the color patch by 500 ms. Thus, as in the current study, the young children were inefficient at resolving the competition between the target noun and the related distractor.

In sum, our results make two important contributions to the literature. First, they provide evidence for a developmental trajectory in spoken word recognition between 5- and 6-year-olds and young adults. Second, they emphasize the importance of using multiple behavioral measures in investigating eye movements in children. We observed a striking behavioral dissociation in children between their general patterns of eye movements (coarse grain) and their moment-by-moment eye movements (fine grain), bringing us to a final issue—the importance of methodology in investigating interactions between language processing and cognitive development. During recent years, eye-tracking has emerged as a critical tool for uncovering developmental dissociations between controlled processes (general patterns of eye movements) and automatic processes (moment-by-moment eye movements), with experimental studies (e.g., Hurewitz, Brown-Schmidt, Thorpe, Gleitman, & Trueswell, 2000; Sekerina, Stromswold, & Hestvik, 2004; Trueswell et al., 1999) providing strong evidence that children parse sentences deterministically; that is, they fail to revise their initially incorrect interpretations of temporary syntactic or referential ambiguities even when disambiguating information becomes available. The current findings also fit well with recent observations of behavioral dissociations in cognitive development. For example, Zelaso, Frye, and Rapus (1996) Ruffman, Garnham, Import, and Connolly (2001) reported differences between preschoolers' awareness of rules and their ability to execute them in dimensional change card sort and false belief tasks, respectively. These dissociations have been attributed to a lack of development of executive functions of the prefrontal cortex responsible for self-regulation, inhibition, planning and modifying behavior, and maintenance of representations in working memory (Davidson, Amso, Anderson, & Diamond, 2006). Taken together with our results, these observations point to a need for further integration of developmental studies of language processing and cognitive development, especially with respect to uncovering how flexibility in attention allocation affects the language processing system.

Acknowledgments

Appendix A

Experimental stimuli (target, competitor, distractor, and fillers) used in Experiments 1 and 2

	Target–competitor	Distractors	Fillers
1	*butylka* 'bottle_{FEM}' *buton* 'rose bud_{MASC}'	*list* 'leaf_{MASC}' *lampočka* 'light bulb_{FEM}'	*kartina* 'picture_{FEM}' *petukh* 'rooster_{MASC}'
2*	*plat'e* 'dress_{NEUT}' *plašč* 'raincoat_{MASC}'	*flag* 'flag_{MASC}' *pero* 'feather_{NEUT}'	*marka* 'stamp_{FEM}' *britva* 'shaver_{FEM}'
3	*banka* 'jar_{FEM}' *bant* 'bow_{MASC}'	*veer* 'fan_{MASC}' *vilka* 'fork_{FEM}'	*korona* 'crown_{FEM}' *topor* 'axe_{MASC}'
4*	*baraban* 'drum_{MASC}' *baran* 'ram_{MASC}'	*jubka* 'skirt_{FEM}' *ruka* 'hand_{FEM}'	*čerepakha* 'turtle_{FEM}' *močalka* 'sponge_{FEM}'
5	*verevka* 'rope_{FEM}' *verbljud* 'camel_{MASC}'	*pingvin* 'penguin_{MASC}' *pugovica* 'button_{FEM}'	*morkov'* 'crown_{FEM}' *pila* 'saw_{FEM}'
6	*krovat'* 'bed_{FEM}' *krokodil* 'crocodile_{MASC}'	*limon* 'lemon_{MASC}' *lošad'* 'horse_{FEM}'	*strelka* 'arrow_{FEM}' *banan* 'banana_{MASC}'
7	*koljaska* 'baby pram_{FEM}' *kolodec* 'well_{MASC}'	*botinok* 'shoe_{MASC}' *mašina* 'car_{FEM}'	*ložka* 'spoon_{FEM}' *sunduk* 'trunk_{MASC}'
8*	*čerepakha* 'turtle_{FEM}' *čerep* 'skull_{MASC}'	*tigr* 'tiger_{MASC}' *šljupa* 'hat_{FEM}'	*spicy* 'knitting needles_{PL}'[a] *lampočka* 'light bulb_{FEM}'
9	*konfeta* 'candy_{FEM}' *konvert* 'envelope_{MASC}'	*platok* 'kerchief_{MASC}' *perčatka* 'glove_{FEM}'	*skamejka* 'bench_{FEM}' *dinosaur* 'dinosaur_{MASC}'
10	*korobka* 'box_{FEM}' *korabl'* 'ship_{MASC}'	*galstuk* 'tie_{MASC}' *gitara* 'guitar_{FEM}'	*moneta* 'coin_{FEM}' *stakan* 'glass_{MASC}'
11*	*busy* 'necklace_{PL}'[a] *buben* 'tambourine_{MASC}'	*ščetka* 'hairbrush_{FEM}' *zvezda* 'star_{FEM}'	*zakolka* 'hairpin_{FEM}' *pomada* 'lipstick_{FEM}'
12	*gruša* 'pear_{FEM}' *gruzovik* 'truck_{MASC}'	*snegovik* 'snowman_{MASC}' *strelka* 'arrow_{FEM}'	*zmeja* 'snake_{FEM}' *divan* 'couch_{MASC}'
13	*medusa* 'jelly fish _{FEM}' *medved'* 'bear_{MASC}'	*kuvšin* 'pitcher_{MASC}' *korova* 'cow_{FEM}'	*skripka* 'violin_{FEM}' *ogurec* 'cucumber_{MASC}'
14	*pal'ma* 'palm tree_{FEM}' *palec* 'finger_{MASC}'	*avtomobil'* 'car_{MASC}' *bočka* 'barrel_{FEM}'	*varežka* 'mitten_{FEM}' *cvetok* 'flower_{MASC}'

Appendix A (*continued*)

	Target–competitor	Distractors	Fillers
15*	*šapka* 'hat$_{FEM}$'	*zamok* 'lock$_{MASC}$'	*ručka* 'pen$_{FEM}$'
	šarik 'balloon$_{MASC}$'	*ključ* 'key$_{MASC}$'	*poezd* 'train$_{MASC}$'
16*	*marka* 'stamp$_{FEM}$'	*pistolet* 'gun$_{MASC}$'	*karta* 'map$_{FEM}$'
	morkov' 'carrot$_{FEM}$'	*nosok* 'sock$_{MASC}$'	*brelok* 'key chain$_{MASC}$'
17*	*karta* 'map$_{FEM}$'	*vinograd* 'grapes$_{MASC}$'	*salfetka* 'napkin$_{FEM}$'
	kartofel' 'potato$_{MASC}$'	*ruka* 'hand$_{FEM}$'	*del'fin* 'dolphin$_{MASC}$'
18	*svin'ja* 'pig$_{FEM}$'	*pauk* 'spider$_{MASC}$'	*ptica* 'bird$_{FEM}$'
	svistok 'whistle$_{MASC}$'	*pila* 'saw$_{FEM}$'	*oslik* 'donkey$_{MASC}$'
19	*sova* 'owl$_{FEM}$'	*kaban* 'boar$_{MASC}$'	*kniga* 'book$_{FEM}$'
	sovok 'dustpan$_{MASC}$'	*koftočka* 'blouse$_{FEM}$'	*luk* 'onion$_{MASC}$'
20	*pomada* 'lipstick$_{FEM}$'	*ananas* 'pineapple$_{MASC}$'	*televizor* 'TV set$_{MASC}$'
	pomidor 'tomato$_{MASC}$'	*akula* 'shark$_{FEM}$'	*zontik* 'umbrella$_{MASC}$'

Note. The asterisk (∗) represents items adapted from the experimental materials used by Marian and Spivey (2003) with bilingual Russian–English adults in their Experiment 2.

[a] Nouns in plural (*pluralia tantum*) do not have a gender.

Appendix B

Corpus frequencies and children's free association data for the target–competitor pairs

	Target–competitor	Gender or number	Simplified phonemic transcription	Translation	Frequencies	
					Child association norms	National Russian corpus
1	**бутылка**	F	/butýlka/	'bottle'	1	663[a]/1000[b]
	бутон	M	/butón/	'rosebud'	0	54/79
2*	**платье**	N	/plát'e/	'dress'	8	1524/7000
	плащ	M	/plášč/	'raincoat'	5	528/1000
3	**банка**	F	/bánka/	'jar'	9	1568/3000
	бант	M	/bánt/	'bow'	5	99/162
4*	**барабан**	M	/barabán/	'drum'	9	314/600
	баран	M	/barán/	'ram'	4	172/313
5	**веревка**	F	/vir'óvka/	'rope'	1	233/381
	верблюд	M	/virbl'út/	'camel'	3	141/250
6	**кровать**	F	/kravát'/	'bed'	5	1062/3000
	крокодил	M	/krakadíl/	'crocodile'	6	184/364
7	**коляска**	F	/kaljáska/	'baby pram'	2	222/525
	колодец	M	/kalódec/	'well'	1	320/800
8*	**черепаха**	F	/čiripákha/	'turtle'	4	98/170
	череп	M	/čérip/	'skull'	6	448/800

Appendix B (*continued*)

	Target–competitor	Gender or number	Simplified phonemic transcription	Translation	Frequencies	
					Child association norms	National Russian corpus
9	**конфета**	F	/kanféta/	'candy'	17	46/57
	конверт	M	/kanvért/	'envelope'	9	516/1000
10	**коробка**	F	/karópka/	'box'	7	464/600
	корабль	M	/karábl'/	'ship'	6	822/3000
11*	**бусы**	Pl	/búsy/	'necklace'	4	151/211
	бубен	M	/búbin/	'tambourine'	0	100/180
12	**груша**	F	/grúša/	'pear'	11	145/246
	грузовик	M	/gruzavík/	'tuck'	1	412/800
13	**медуза**	F	/midúza/	'jellyfish'	0	51/164
	медведь	M	/midvét'/	'bear'	18	541/1000
14	**пальма**	F	/pál'ma/	'palm tree'	2	131/237
	палец	M	/pálec/	'finger'	6	1064/2000
15*	**шапка**	F	/šápka/	'hat'	9	481/800
	шарик	M	/šárik/	'balloon'	7	348/600
16*	**марка**	F	/márka/	'stamp'	3	725/2000
	морковь	F	/markóf'/	'carrot'	3	219/492
17*	**карта**	F	/kárta/	'map'	12	657/1000
	картофель	M	/kartófil'/	'potato'	13	289/576
18	**свинья**	F	/svin'já/	'pig'	8	409/600
	свисток	M	/svistók/	'whistle'	2	193/316
19	**сова**	F	/savá/	'owl'	11	133/332
	совок	M	/savók/	'dustpan'	1	77/100
20	**помада**	F	/pamáda/	'lipstick'	2	103/456
	помидор	M	/pamidór/	'tomato'	5	104/141

Note. The accent (′) is shown on the stressed vowel. An apostrophe (') represents a palatalized consonant.

[a] In this column, the number of texts in which the word was encountered.

[b] In this column, the number of contexts in which the word was encountered.

References

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language, 38*, 419–439.

Altmann, G. T. M., & Kamide, Y. (2004). Now you see it, now you do not: Mediating the mapping between language and the visual world. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 347–386). Hove, UK: Psychology Press.

Cycowicz, Y., Friedman, D., Rothstein, M., & Snodgrass, J. (1997). Picture naming by young children: Norms for name agreement, familiarity, and visual complexity. *Journal of Experimental Child Psychology, 65*, 171–237.

Dahan, D., Magnuson, J. S., & Tanenhaus, M. K. (2001). Time course of frequency effects in spoken-word recognition: Evidence from eye movements. *Cognitive Psychology, 42*, 317–367.

Dahan, D., Swingley, D., Tanenhaus, M. K., & Magnuson, J. S. (2000). Linguistic gender and spoken-word recognition in French. *Journal of Memory and Language, 42*, 465–480.

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia, 44*, 2037–2078.

Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology, 42*, 98–116.

Fernald, A., Pinto, J. P., Swingley, D., Weinberg, A., & McRoberts, G. (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. *Psychological Science, 9*, 72–75.

Foulke, E., & Sticht, T. (1969). Review of research on the intelligibility and comprehension of accelerated speech. *Psychological Bulletin, 72*, 50–62.

Garlock, V. M., Walley, A. C., & Metsala, J. L. (2001). Age-of-acquisition, word frequency, and neighborhood density effects on spoken word recognition by children and adults. *Journal of Memory and Language, 45*, 468–492.

Gaskell, M. G., & Marslen-Wilson, W. D. (2001). Representation and competition in the perception of spoken words. *Cognitive Psychology, 45*, 220–266.

Hanauer, J. B., & Brooks, P. J. (2003). Developmental change in the cross-modal Stroop effect. *Perception and Psychophysics, 65*, 359–366.

Hurewitz, F., Brown-Schmidt, S., Thorpe, K., Gleitman, L. R., & Trueswell, J. C. (2000). One frog, two frog, red frog, blue frog: Factors affecting children's syntactic choices in production and comprehension. *Journal of Psycholinguistic Research, 29*, 597–626.

Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Within- and between-language competition. *Bilingualism: Language and Cognition, 6*, 97–116.

Marslen-Wilson, W. D., & Tyler, L. K. (1980). The temporal structure of spoken word comprehension. *Cognition, 6*, 1–71.

Marslen-Wilson, W. D., & Warren, P. (1994). Levels of perceptual representation and process in lexical access: Words, phonemes, and features. *Psychological Review, 101*, 653–675.

Mattys, S., & Samuel, A. G. (2000). Implications of stress-pattern differences in spoken-word recognition. *Journal of Memory and Language, 42*, 571–596.

McQueen, J. M. (2004). Speech perception. In K. Lamberts & R. Goldstone (Eds.), *The handbook of cognition* (pp. 255–275). London: Sage.

Metsala, J. L. (1997). An examination of word frequency and neighborhood density in the development of spoken word recognition. *Memory and Cognition, 25*, 47–56.

Metsala, J. L., & Walley, A. C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability. In J. L. Metsala & L. E. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 89–120). Mahwah, NJ: Lawrence Erlbaum.

Nacional'nyj korpus russkogo jazyka [National Russian Corpus]. (2005). Retrieved March 30, 2006, from www.ruscorpora.ru/corpora-intro.html.

Nation, K., Marshall, C. M., & Altmann, G. T. M. (2003). Investigating individual differences in children's real-time sentence comprehension using language-mediated eye movements. *Journal of Experimental Child Psychology, 86*, 314–329.

Ovčinnikova, I. G., Berseneva, N. I., Dubrovskaya, L. A., & Penyagina. E. B. (2000). Leksikon mladšego škol'nika (xarakteristika leksičeskogo komponenta jazykovoj kompetencii [Lexicon of the elementary school student (the nature of the lexical component of language competence)]. Perm, Russia.

Ruffman, T., Garnham, W., Import, A., & Connolly, D. (2001). Does eye gaze indicate implicit knowledge of false belief? Charting transitions in knowledge. *Journal of Experimental Child Psychology, 80*, 201–224.

Sekerina, I. A., Stromswold, K., & Hestvik, A. (2004). How do adults and children process referentially ambiguous pronouns? *Journal of Child Language, 31*, 123–152.

Snedeker, J., & Trueswell, J. C. (2004). The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology, 49*, 238–299.

Swingley, D., Pinto, J. P., & Fernald, A. (1999). Continuous processing in word recognition at 24 months. *Cognition, 71*, 73–108.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*, 1632–1634.

Tanenhaus, M. K., & Trueswell, J. C. (2005). Eye movements as a tool for bridging the language-as-product and language-as-action traditions. In M. K. Tanenhaus & J. C. Trueswell (Eds.), *Approaches for studying world-situated language use* (pp. 3–37). Cambridge, MA: MIT Press.

Trueswell, J. C., Sekerina, I. A., Hill, N., & Logrip, M. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition, 73*, 89–134.

Walley, A. (1993). The role of vocabulary development in children's spoken word recognition and segmentation ability. *Developmental Review, 13*, 286–350.

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language, 50*, 1–25.

Zelaso, P. D., Frye, D., & Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cognitive Development, 11*, 37–63.