
ЭМПИРИЧЕСКИЕ ИССЛЕДОВАНИЯ
EMPIRICAL RESEARCH

Роль полиперспективных репрезентаций в процессах совместного внимания

М.В. Зотов*

Санкт-Петербургский государственный университет, Санкт-Петербург, Россия,
mvozotov@mail.ru

Н.Е. Андрианова**

Санкт-Петербургский государственный университет, Санкт-Петербург, Россия,
natalia-andrianova@mail.ru

А.П. Войт***

Санкт-Петербургский государственный университет, Санкт-Петербург, Россия,
voytalexey@mail.ru

Целью работы являлась проверка гипотезы о зависимости идентификации объектов внимания наблюдаемого человека от понимания контекста коммуникативной ситуации, основанного на сопоставлении точек зрения ее участников. В первом эксперименте испытуемые (N=74) выполняли поиск объектов внимания участников невербальных коммуникативных сцен при наличии/отсутствии информации о коммуникативном контексте, а также при короткой/длительной экспозиции информации для оценки направления взгляда персонажей. При понимании контекста испытуемые обнаружили высокую эффективность идентификации и саккадической детекции объектов внимания персонажей, независимо от доступности информации о направлении их взгляда. Во втором эксперименте после просмотра видеофрагментов коммуникативных ситуаций испытуемые (N=32) выполняли поиск изменений разных типов объектов в кадрах сцен. Лица, понимавшие контекст коммуникации, в первую очередь замечали изменения будущих объектов внимания участников, несмотря на отсутствие «подсказок» в виде ориентации головы и глаз. Результаты обсуждаются с позиции концепций, акцентирующих значение полиперспективных репрезентаций для понимания коммуникативных сигналов.

Ключевые слова: теории референции, совместное внимание, невербальная коммуникация, зрительное восприятие, движения глаз.

Для цитаты:

Зотов М.В., Андрианова Н.Е., Войт А.П. Роль полиперспективных репрезентаций в процессах совместного внимания // Культурно-историческая психология. 2015. Т. 11. № 2. С. 16–27. doi: 10.17759/chp.2015110202.

* *Зотов Михаил Владимирович*, доктор психологических наук, профессор кафедры медицинской психологии и психофизиологии факультета психологии, Санкт-Петербургский государственный университет, Санкт-Петербург, Россия. *mvozotov@mail.ru*

** *Андрианова Наталия Евгеньевна*, ассистент кафедры медицинской психологии и психофизиологии факультета психологии, Санкт-Петербургский государственный университет, Санкт-Петербург, Россия. *natalia-andrianova@mail.ru*

*** *Войт Алексей Петрович*, кандидат физико-математических наук, ведущий инженер физического факультета, Санкт-Петербургский государственный университет, Санкт-Петербург, Россия. *voytalexey@mail.ru*

В настоящее время большое число исследований посвящено проблеме совместного внимания (joint attention), под которым понимают способность человека обращать внимание на тот же объект или событие, на которые обращает внимание другой человек [7]. Во многих работах доминирует взгляд на совместное внимание как на элементарный и даже «рефлекторный» процесс прослеживания взором направления взора другого человека (gaze following) [13; 15; 16; 17]. Утверждается, что в основе совместного внимания лежит «геометрический» механизм («geometric mechanism of joint attention»), предполагающий, что идентификация объекта внимания другого человека осуществляется на основе построения воображаемой линии в пространстве в соответствии с ориентацией его головы и глаз [11; 18]. В рамках данной позиции исследована степень точности, с которой люди способны оценивать угол направления взгляда другого человека [10], а также проанализирован вклад положения тела, поворота головы и ориентации глаз индивида в определение линии его взгляда [18; 27].

Между тем эти работы игнорируют факт, что взгляд другого человека, как и указательный жест, лишь включает отсылку к определенной области пространства, но не содержит достаточной информации для идентификации своего объекта/референта. Воспользуемся примером М. Томаселло [6] и представим человека, взглянувшего на пробегающего мимо кролика. Как мы определим, что именно является объектом его внимания — сам кролик, его лапа, уши, бег и т.д.? Как мы определим, что объектом внимания человека является именно кролик, а не любой другой объект (например, куст), находящийся на линии его взгляда? Между тем в процессе повседневного общения мы способны не только легко идентифицировать, но и предсказывать объекты внимания собеседника, часто даже не прослеживая линию его взгляда. Какие когнитивные механизмы лежат в основе данной способности?

Этот вопрос может быть рассмотрен с позиции теории «кругозора и окружения», разработанной выдающимся отечественным мыслителем М.М. Бахтиным еще в 20-е гг. прошлого века [1; 2]. По мнению М.М. Бахтина, жест, взгляд другого человека или, например, высказывание «так», взятые сами по себе, изолированно, не несут существенной информации и обретают значение лишь при соотнесении их с «внесловесным контекстом» коммуникативной ситуации. Этот «внесловесный контекст» складывается из трех компонентов: 1) общий для собеседников пространственный кругозор («вместе видимое»); 2) общие знания («вместе знаемое»); 3) согласующиеся концептуализации ситуации («согласно оцененное») [1, с. 152]. Наблюдатель может определить значение взглядов, жестов и высказываний лишь тогда, когда он «приобщится» к «общему пространственному и смысловому кругозору» участников коммуникации, т. е. сконструирует и сопоставит соответствующие их перспективе или

точке зрения репрезентации ситуации [там же]. В качестве примера Бахтин приводит игру в разбойников: «Мальчик, играющий атамана разбойников, изнутри переживает свою жизнь разбойника, глазами разбойника смотрит на пробегающего мимо другого мальчика, его кругозор есть кругозор изображаемого им разбойника; то же самое имеет место и для его сотоварищей по игре» [2, с. 67]. Жесты, взгляды и высказывания играющих детей зритель может понять лишь тогда, когда «увидит изнутри их мир так, как они его видят», т. е. разделит с ними знание о разыгрываемом событии («нападение разбойников») и приобщится к их концептуальному видению внешнего окружения («палка — меч, столбы — солдаты»). Одновременно зритель учитывает аспекты окружения, недоступные сознанию участников в данный момент времени, но релевантные их предполагаемым намерениям. М.М. Бахтин пишет: «Близкой к действительной позиции зрителя представляется нам наивная установка того простолюдина, который предупреждал героя пьесы о сделанной против него засаде и готов был броситься ему на помощь во время сделанного на него нападения. Такой установкой наивный зритель занимал устойчивую позицию вне героя, учитывал трансгredientные сознанию самого героя моменты» [там же, с. 71]. Итак, репрезентации, направляющие понимание наблюдателем взглядов, жестов и высказываний участников коммуникации, по своей природе являются полиперспективными, поскольку отражают происходящее одновременно с нескольких перспектив¹.

Вышеописанные положения теории М.М. Бахтина получили подтверждение в исследованиях последних лет, посвященных проблемам социального познания. В работах М. Томаселло и других авторов было показано, что полиперспективные или диалогичные когнитивные репрезентации (perspective or dialogic cognitive representations) являются уникальными для человеческого вида и играют фундаментальную роль в развитии коммуникативных компетенций в детском возрасте [6; 21]. В серии когнитивно-лингвистических исследований было продемонстрировано, что способность человека конструировать когнитивные репрезентации, объединяющие несколько точек зрения (viewpoints) на объекты или события, имеет ключевое значение для понимания и продуцирования различных видов коммуникативных сигналов [23; 26].

В соответствии со сказанным, цель настоящей работы состояла в изучении роли полиперспективных когнитивных репрезентаций в процессах совместного внимания. Проведено два экспериментальных исследования.

Эксперимент 1

Первый эксперимент был направлен на проверку гипотезы о том, что идентификация объектов внимания участников социальной ситуации в большей сте-

¹ Термин «перспектива» используется в широком значении как позиция наблюдателя или точка зрения, которой придерживается наблюдатель, интерпретируя объекты и события [3].

пени зависит от понимания контекста коммуникации (в значении М.М. Бахтина) и в меньшей степени — от визуальной оценки направления их взгляда.

Метод

Участники исследования. В эксперименте приняли участие 74 здоровых испытуемых в возрасте от 19 до 37 лет, которые были разделены на первую (N=18), вторую (N=18), третью (N=17) и четвертую (N=21) экспериментальные группы.

Процедура. Исследование проводилось с помощью системы регистрации движений глаз Tobii X120. Была разработана оригинальная экспериментальная процедура, представляющая собой сочетание задачи саккадической детекции [12] и методики «подсказки взглядом» (gaze-cueing paradigm) [15] (рис. 1).

Суть процедуры состояла в следующем. При отсутствии какой-либо предварительной информации или сразу после просмотра видеофрагмента, позволяющего сформировать представление о контексте коммуникативной ситуации, испытуемые выполняли задачу саккадической детекции объекта внимания одного из персонажей. После фиксации взгляда в центре экрана им предъявлялась динамическая «подсказка», длительность экспозиции которой составляла 200 или 750 мс. В качестве «подсказки» предъявлялась последовательность из пяти кадров, создающая впечатление поворота головы персонажа в ту или иную сторону. Затем в течение 1500 мс. предъявлялось целевое изображение — кадр сцены, ранее не виденный испытуемым. Задача состояла в том, чтобы как можно быстрее перевести взгляд на объект, на который, по мнению испытуемого, смотрит персонаж. Другими словами, испытуемый должен был совершить саккаду от области лица персонажа к предполагаемому объекту его внимания. После окончания пробы он должен был назвать данный объект.

Четыре группы испытуемых выполняли задачу в одном из условий: 1) наличие/отсутствие информации о контексте коммуникации; 2) короткое/длительное время экспозиции информации, предоставляемой для визуальной оценки направления взгляда персонажа. Анализировались вербальные ответы испытуемых, локализация их зрительных фиксаций, угол отклонения первой саккады от линии взгляда

персонажа, оцениваемой по объективным характеристикам изображения.

В качестве стимульного материала использовались пять «немых» видеоизображений социальных сцен из телесериала «Мистер Бин» (Великобритания, 1990—1995), художественного фильма «Новые времена» (США, 1936), художественного фильма «Собачья жизнь» (США, 1918), художественного фильма «Пугало» (США, 1920) длительностью от 20 до 40 с. Приведем содержание сцен 1, 2 и 3.

Сцена 1 (т/с «Мистер Бин»). У кассы супермаркета стоят два покупателя: пожилой мужчина и «мистер Бин». Мужчина отдает продавцу скидочную карту, в то время как «мистер Бин» достает аналогичную карту и демонстрирует ее зрителю. Продавец возвращает мужчине скидочную карту (кладет ее перед ним на прилавок), но последний не замечает этого. «Мистер Бин» кладет свою карту на прилавок рядом с картой мужчины. Не замечая своей карты, покупатель ставит на нее сумку. Видеоклип прерывается в тот момент, когда мужчина делает начальное движение по направлению к прилавку, и зритель ожидает, что вместо своей карты он ошибочно возьмет карту «мистера Бина». Задача: основываясь на оценке поворота головы покупателя, определить объект его внимания.

Сцена 2 (х/ф «Новые времена»). «Маленький бродяга» (Ч. Чаплин) помогает мастеру чинить заводской пресс. Мастер просит его подержать инструмент и пиджак. Не подумав, «бродяга» кладет их под пресс. Когда мастер включает пресс, герой успевает выхватить из-под него инструмент, но забывает про пиджак. Пресс опускается. Видеоклип прерывается в тот момент, когда пресс поднимается, и зритель ожидает, что оба персонажа увидят испорченный пиджак мастера. Задача: основываясь на оценке поворота головы героя Ч. Чаплина, определить объект его внимания.

Сцена 3 (х/ф «Собачья жизнь»). «Маленький бродяга» (Ч. Чаплин) замечает стоящего за забором торговца сосисками. Просунув руку через щель в заборе, «бродяга» ворует сосиску и готовится ее съесть. Проходящий мимо полицейский замечает факт воровства и останавливается, наблюдая через забор за действиями «бродяги». Видеоклип прерывается в тот момент, когда «бродяга» начинает пово-

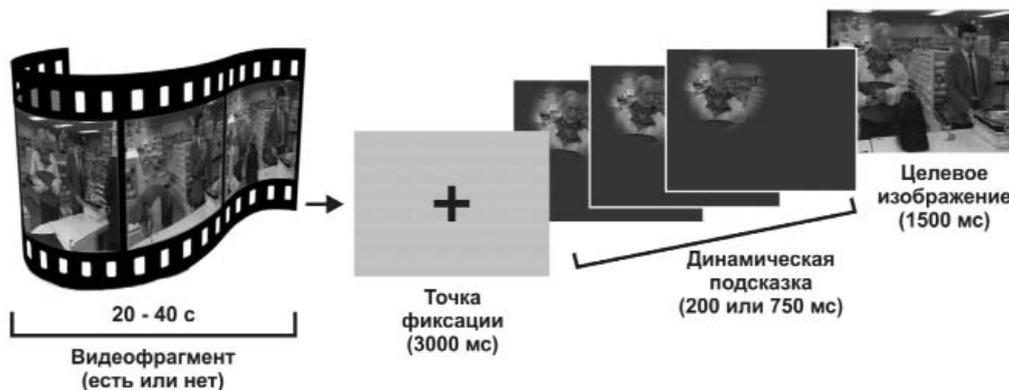


Рис. 1. Задача саккадической детекции объекта внимания персонажа социальной сцены

рачивать лицо к забору, и зритель ожидает, что он заметит полицейского. Задача: основываясь на оценке поворота головы бродяги, определить объект его внимания.

Содержание видеофрагментов и характеристики целевых стимулов исключали использование испытуемыми «восходящей» формы совместного внимания (bottom-up joint attention) [6]. Целевые объекты внимания персонажей обладали небольшим размером (0,2–1,5 угл.град.) и характеризовались низким уровнем «визуальной яркости» (visual saliency), которая оценивалась с помощью инструментария «Saliency Toolbox» [28]. В видеофрагментах отсутствовали признаки движения целевых объектов, могущие привлечь к ним произвольное внимание наблюдателя. Также в видеофрагментах отсутствовали пространственные подсказки, такие как повороты головы и указательные жесты, ориентирующие внимание испытуемых в направлении целевых объектов внимания персонажей. Видеофрагменты всегда прерывались до начала поворота головы персонажа к объекту его интереса. Момент прерывания видеофрагментов являлся неожиданным для испытуемых. В сценах 1 и 2 и 4 целевые объекты внимания персонажа («карта», «пиджак») были связаны с интенциями обоих участников и являлись потенциальными референтами² их предполагаемого взаимодействия. В сценах 3 и 5 целевые объекты были связаны с интенциями центрального персонажа.

Результаты и их обсуждение

Были проанализированы вербальные ответы испытуемых при идентификации объектов внимания персонажей социальных сцен. Распределение ответов испытуемых для сцен 1 и 2 представлено в табл.

Как видно из представленных данных, наличие или отсутствие информации о контексте коммуникативной ситуации оказывало существенное влияние на ответы испытуемых. В сцене 1 48,7% испытуемых, не обладавших информацией о контексте, указали на «корзину» и лишь 8,1% — на «скидочную карту», находящуюся на линии взгляда персонажа. Эти лица демонстрировали большое разнообразие вербальных ответов, обозначая карту как «купюра»,

«что-то на столе» и т.д., в связи с чем их ответы группировались на основе общности референта. В отличие от них, 91,7% испытуемых, обладавших пониманием контекста, указали на «скидочную карту», и лишь 5,5% — на «корзину». Сходные результаты были получены по остальным сценам.

Для показателя отклонения первой саккады от линии взгляда персонажа проводился дисперсионный анализ ANOVA с двумя межгрупповыми факторами *Понимание контекста* (нет/есть) и *Время экспозиции подсказки* (200/750 мс.) и одним внутригрупповым фактором *Номер социальной сцены* (сцены 1, 2, 3, 4 и 5). Выявлено достоверное влияние на данный показатель факторов *Понимание контекста* ($p < 0,001$), *Номер сцены* ($p < 0,001$) и взаимодействия факторов *Понимание контекста*–*Время экспозиции подсказки* ($p < 0,001$) (рис. 2).

Установлено, что увеличение времени экспозиции информации, предоставляемой для «геометрической» оценки изменения направления взгляда персонажа, приводит к достоверному снижению от-

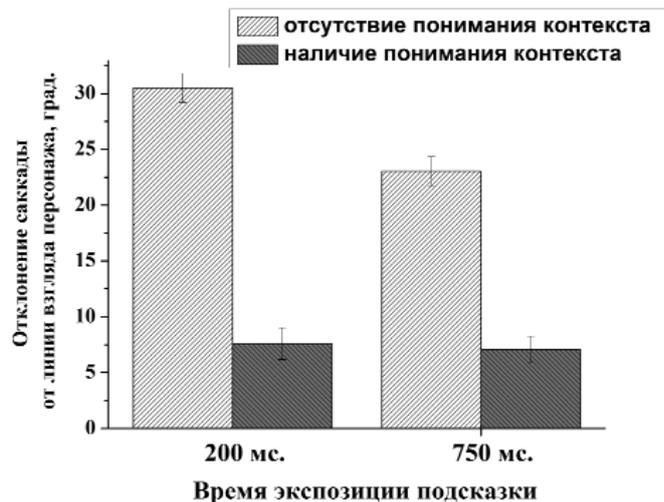


Рис. 2. Влияние понимания коммуникативного контекста и времени экспозиции подсказки на отклонение первой саккады от линии взгляда персонажа.

Примечание. На этом и последующих графиках (рис. 7 и 9) приведены значения среднего и стандартной ошибки для анализируемых показателей

Таблица

Распределение ответов испытуемых при идентификации объектов внимания персонажей сцен 1 и 2

Вид ответа	Сцена 1		Вид ответа	Сцена 2	
	Доля испытуемых, %			Доля испытуемых, %	
	Отсутствие информации о контексте	Наличие информации о контексте		Отсутствие информации о контексте	Наличие информации о контексте
«Скидочная карта»	8,1	91,7	«Пиджак» или «предмет под прессом»	0	94,4
«Корзина»	48,7	5,5	«Что-то на столе»	7,9	2,8
«Руки соседа»	21,6	0	«Прибор на стене»	42,1	0
«Карман соседа»	5,4	0	«Что-то за окном» или «за стойкой»	18,5	0
«Сосед»	5,4	2,8	«Не знаю»	15,8	0
«Не знаю»	10,8	0	Другие ответы	15,7	2,8

² Под термином «референт» понимается объект внеязыковой действительности, являющийся предметом обсуждения и/или взаимодействия участников коммуникативной ситуации [4].

клонения саккад от линии взгляда персонажа у испытуемых, не имевших информации о контексте коммуникации ($p < 0,001$), но не оказывает значимого влияния на эффективность совместного внимания лиц, понимавших контекст ($p > 0,05$).

В целом результаты эксперимента показали, что в случае понимания коммуникативного контекста точная оценка ориентации позы, головы или глаз наблюдаемого человека, постулируемая Дж. Баттеруортом и другими авторами [11; 13; 17; 18], не играет значимой роли в процессах совместного внимания. В условиях дефицита информации о направлении взгляда персонажей испытуемые, обладавшие пониманием коммуникативного контекста, демонстрировали высокую эффективность идентификации объектов их внимания.

Между тем для такой идентификации было недостаточно знания ситуации и понимания намерений ее участников. Необходимо было сконструировать репрезентации, отражающие точки зрения («кругозоры», по М.М. Бахтину) участников, и сопоставить их с информацией, доступной с позиции наблюдателя, т. е. рассмотреть ситуацию одновременно с нескольких перспектив. Чтобы идентифицировать объект внимания покупателя в социальной сцене 1, наблюдатель должен был предварительно не только распознать его намерение взять карту, но и определить ошибку персонажа, вследствие которой он принял карту «мистера Бина» за свою. Для этого было необходимо сопоставить точку зрения покупателя с информацией, доступной с позиции наблюдателя («мужчина не заметил, как положил сумку на свою карту»). Данная социальная сцена сходна с ситуацией задачи «Салли-Энн», используемой для оценки способности к пониманию «ошибочных убеждений» [9]. Чтобы определить объект внимания «бродяги» в сцене 2, наблюдатель предварительно должен был заметить его ошибку, когда «бродяга» оставил пиджак под прессом. Это также достигалось сопоставлением точки зрения персонажа с информацией, доступной с позиции наблюдателя. Наконец, чтобы определить объект внимания персонажа в сцене 3, наблюдатель также должен был реконструировать его точку зрения и сопоставить с информацией, доступной с внешней позиции («бродяга» не заметил полицейского, когда воровал сосиску). Итак, успешная идентификация объектов внимания персонажей предъявленных социальных сцен требовала от испытуемых создания полиперспективных репрезентаций, предполагающих реконструкцию и сопоставление нескольких точек зрения на происходящее.

На рис. 3 и 4 (см. вкладку) показано распределение фиксаций взгляда в группах испытуемых, обладавших и не обладавших пониманием коммуникативного контекста, при поиске объектов внимания участников социальных сцен.

Как видно из представленных данных, понимание коммуникативного контекста оказывает существен-

ное влияние на характер зрительного поиска объектов внимания персонажей сцен. Испытуемые, не просмотревшие видеофрагменты, совершали саккады в направлении линии взгляда персонажей, но испытывали выраженные затруднения в идентификации объектов их внимания, демонстрируя хаотичную и разнонаправленную поисковую активность. Они фиксировали взгляд преимущественно на визуально ярких и крупных объектах, находящихся по направлению взгляда персонажа. Напротив, лица, просмотревшие видеофрагменты, в большинстве случаев решали поставленную задачу путем реализации одной или двух точных саккад от лица персонажа к объекту его внимания. Они демонстрировали высокоэффективный вид совместного внимания, предполагавший быстрое (100–250 мс.) экстрафовеальное опознание объектов взгляда персонажей и выделение их в качестве саккадических целей, несмотря на низкий уровень «визуальной яркости» (visual saliency) этих объектов и наличие конкурирующего предметного окружения.

При объяснении этого факта необходимо учесть два обстоятельства. Во-первых, как отмечалось ранее, испытуемые осуществляли поиск объектов внимания персонажей в ранее не виденных ими кадрах сцены. В предъявленных им видеофрагментах эти объекты были изображены со смещенного ракурса или имели другой вид. Во-вторых, видеофрагменты всегда прерывались неожиданно для испытуемых, и они не могли заранее предугадать, какие именно объекты внимания персонажей им предстоит искать.

Каким образом испытуемым удавалось осуществлять столь точную саккадическую детекцию объектов внимания персонажей в ранее не виденных кадрах сцены в условиях визуальной «малозаметности» (non-saliency) этих объектов и присутствия конкурирующих стимулов? Согласно данным многочисленных исследований [например: 20], столь эффективный зрительный поиск мог основываться только на предварительно известной информации о визуальных признаках и локализации целевых объектов. С учетом факта, что испытуемые заранее не знали, какие объекты им предстоит искать, можно сделать вывод о том, что при восприятии социальной ситуации они удерживали в памяти визуально-пространственные характеристики всех стимулов, рассматриваемых в качестве потенциальных объектов внимания персонажей.

Итак, результаты исследования свидетельствуют о том, что при восприятии социальных сцен наблюдатель в режиме реального времени конструирует и сопоставляет репрезентации, отражающие видение ситуации с точки зрения различных персонажей. Это позволяет ему выделять потенциальные объекты внимания участников ситуации, которые рассматриваются в качестве основы (common ground) их предполагаемого коммуникативного взаимодействия. Информацию о визуальных признаках и локализации этих объектов наблюдатель избирательно

запоминает и удерживает в памяти, что в дальнейшем позволяет ему легко идентифицировать референты взглядов, указательных жестов и высказываний наблюдаемых людей, а также осуществлять практически мгновенный поиск этих объектов/референтов в визуальном пространстве сцены.

Далее мы предположили, что при восприятии коммуникативных ситуаций процесс выделения будущих объектов внимания наблюдаемых участников и запоминания их характеристик происходит автоматически и не зависит от стоящей перед наблюдателем задачи.

Эксперимент 2

Второй эксперимент был направлен на проверку гипотезы о том, что при восприятии социальных сцен люди непроизвольно направляют (со-направляют) свое внимание на предполагаемые будущие объекты внимания участников коммуникативной ситуации без специальной инструкции и «подсказок» в виде ориентации головы и глаз наблюдаемых персонажей.

Метод

Участники исследования. В эксперименте приняли участие 32 здоровых испытуемых в возрасте от 19 до 26 лет, которые случайным образом были разделены на первую (N=16) и вторую (N=16) экспериментальные группы.

Процедура. Процедура эксперимента показана на рис. 5.

В соответствии с инструкцией, испытуемые просматривали видеофрагменты социальных сцен длительностью 20–40 с. После неожиданного прерывания видеофрагмента, испытуемые переводили взгляд на фиксационный крест, после чего выполняли задачу детекции изменений (flicker task) [24]. В течение 20 с. им циклическим образом предъявлялись кадр А и модифицированный кадр А', где были изменены три объекта, в перерыве между ними предъявлялся маскировочный стимул, задача испытуемых состояла в том, чтобы обнаружить все изменения.

Первая группа испытуемых просматривала видеофрагменты «А», позволяющие сформировать

представление о смысловом контексте коммуникации. Вторая группа просматривала «урезанные» видеофрагменты «В», обеспечивающие возможность общей ориентировки в сцене и ее персонажах, но не позволяющие понять коммуникативный контекст. Использовались «немые» видеоизображения трех социальных сцен из телесериала «Мистер Бин». Как и в предыдущем исследовании, видеофрагменты прерывались до начала поворота головы персонажа к объекту его интереса. Первая сцена была идентична вышеописанной сцене 1 первого исследования. Опишем содержание сцены 2.

Сцена 2 (т/с «Мистер Бин»). «Мистер Бин» находится в ресторане, где ему приносят несъедобное блюдо. «Мистер Бин» добивается его замены. Официант приносит новый поднос с блюдом, накрытым крышкой, и ставит его на столик «мистера Бина». Видеокадр прерывается в тот момент, когда «мистер Бин» смотрит на музыканта, играющего на скрипке, и зритель предвосхищает, что официант снимет крышку с блюда и герой увидит, что ему принесли то же самое блюдо.

В задаче детекции изменений предъявлялись видеокадры, отсутствующие в видеофрагментах. По ориентации позы, головы и глаз изображенных на этих кадрах персонажей было невозможно определить объекты их интереса. С помощью инструментария «Saliency Toolbox» [28] был проведен анализ каждого кадра с целью выявления визуально ярких (visual saliency) областей, по результатам которого были созданы модифицированные кадры А', каждый из которых включал 3 типа изменений: 1-й тип — исчезновение визуально «малозаметного» (non-salient) объекта, связанного с будущим вниманием персонажей и их коммуникативным взаимодействием; 2-й тип — исчезновение детали облика персонажа; 3-й тип — исчезновение «визуально яркого» объекта, не несущего смысловой нагрузки. Как и в предыдущем эксперименте, объекты внимания персонажей также являлись потенциальными референтами их будущего коммуникативного взаимодействия (рис. 6, см. вкладку).

Оценивались порядок и время обнаружения испытуемыми всех типов изменений. Также осуществлялась регистрация движений глаз испытуемых в процессе эксперимента.

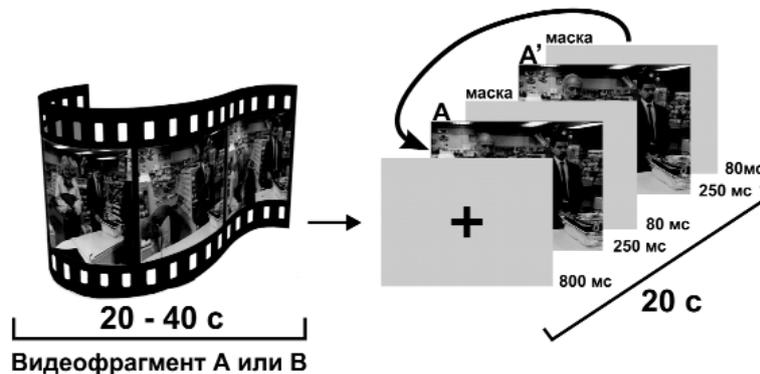


Рис. 5. Задача детекции изменений после просмотра видеофрагмента социальной сцены

Результаты и их обсуждение

Для показателей порядка и времени детекции изменений был проведен дисперсионный анализ ANOVA с одним межгрупповым фактором *Группа* (первая и вторая группы) и двумя внутригрупповыми факторами: *Тип изменения* (1-й, 2-й и 3-й типы) и *Номер сцены* (сцены 1, 2 и 3). Выявлено достоверное влияние на данные показатели фактора *Тип изменения* ($p < 0,001$) и взаимодействия факторов *Группа—Тип изменения* ($p < 0,001$) (рис. 7).

Как показано на рис. 7, испытуемые первой группы, обладавшие пониманием коммуникативного контекста, в первую очередь замечали изменения предметов, рассматриваемых в качестве будущих объектов внимания и коммуникации персонажей. Они демонстрировали минимальное время поиска этих изменений ($1,5 \pm 1,2$ с.) по сравнению с поиском изменений облика персонажа ($11,2 \pm 6,4$ с.) и визуально ярких деталей окружения ($12,2 \pm 4,7$ с.). Напротив, испытуемые второй группы, не обладавшие информацией о смысловом контексте ситуации, более быстро обнаруживали изменения облика персонажей ($6,7 \pm 6,4$ с.) по сравнению с обнаружением изменений объектов внимания персонажей ($9,3 \pm 6,3$ с.) и визуально ярких объектов ($10,9 \pm 5,1$ с.).

На рис. 8 (см. вкладку) показаны «тепловые карты» распределения фиксации взгляда испытуемых обеих групп в течение первой секунды после начала выполнения детекции изменений в сценах 1 и 2.

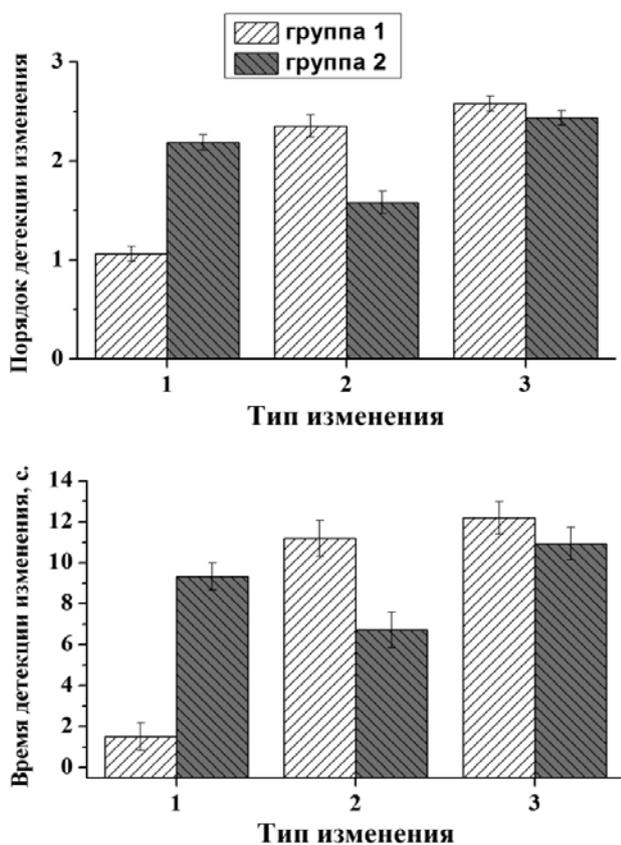


Рис. 7. Порядок и время обнаружения различных типов изменений в кадрах у испытуемых, обладавших (группа 1) и не обладавших (группа 2) пониманием коммуникативного контекста

Видно, что в начальный период выполнения задачи испытуемые, обладавшие информацией о коммуникативном контексте, в отличие от лиц, не обладавших такой информацией, переводят взгляд в область предполагаемых будущих объектов внимания участников сцены, несмотря на отсутствие визуальных указаний на данную область, таких как ориентации позы, головы или глаз персонажей.

Известно, что фиксация взгляда на области изменяющегося элемента не является достаточным условием для обнаружения изменения. Наблюдатели могут локализовать взгляд на одной области, но обрабатывать различные объекты или аспекты объектов [22]. Очевидно, что это явление наиболее выражено в ситуациях, где изменяющиеся объекты представлены в конкурирующем предметном окружении. Для сцен, в которых объекты внимания персонажей находились в конкурирующем окружении, у испытуемых обеих групп было проанализировано количество «неэффективных» зрительных фиксаций, т.е. фиксаций взгляда на области изменяющегося объекта, не приводящих к обнаружению его изменения. Дисперсионный анализ ANOVA с одним межгрупповым фактором *Группа* (первая группа, вторая группа) и двумя внутригрупповыми факторами *Тип изменения* (1-й тип, 2-й тип, 3-й тип) и *Номер сцены* (сцены 2 и 3) показал достоверное влияние фактора *Тип изменения* ($p < 0,001$) и взаимодействия факторов *Группа—Тип изменения* ($p < 0,001$) на количество «неэффективных» фиксаций. Влияние остальных факторов оказалось статистически не достоверным (рис. 9).

Как показано на рис. 9, в отличие от испытуемых второй группы, более одного раза фиксировавших взгляд на области объектов 1-го типа, испытуемые первой группы замечали изменения в этих объектах с первого взгляда, несмотря на то, что данные объекты были представлены в конкурирующем предмет-

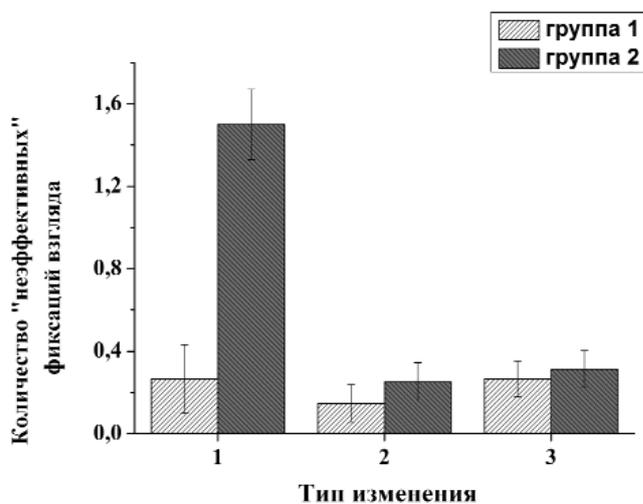


Рис. 9. Количество «неэффективных» фиксаций взгляда на области изменяющихся объектов у испытуемых, обладавших (группа 1) и не обладавших (группа 2) пониманием коммуникативного контекста

ном окружении. Это означает, что они использовали объектно-ориентированную стратегию внимания, предполагающую наличие образа целевого объекта (target template information) в визуальной рабочей памяти [см.: 20].

Между тем, как и в первом эксперименте, испытуемым предъявлялись кадры, в которых целевые объекты были смещены или изображены под другим ракурсом по сравнению с просмотренными видеофрагментами. Видеофрагменты прерывались неожиданно, и испытуемые заранее не знали, поиск каких объектов им предстоит выполнять. Поставленная перед испытуемыми задача — последовательно искать изменения в кадре, которые «бросаются в глаза» — не стимулировала их к осознанному поиску объектов внимания персонажей, поскольку во всех пробах изменения также присутствовали в визуально ярких элементах и деталях облика персонажей. Наконец, в видеофрагментах и кадрах сцен отсутствовали пространственные подсказки, такие как повороты головы и указательные жесты персонажей, ориентирующие внимание испытуемых в направлении объектов 1-го типа.

С учетом сказанного, представленные выше результаты подтверждают выдвинутую гипотезу. При восприятии социальных сцен наблюдатели, вне зависимости от стоящей перед ними задачи, выделяют будущие объекты внимания и/или референты предполагаемого коммуникативного взаимодействия участников и удерживают информацию об этих объектах/референтах в визуальной рабочей памяти. В результате они непреднамеренно направляют внимание на данные объекты и быстро замечают их изменения, несмотря на отсутствие пространственных указаний на эти объекты, таких как ориентация головы и глаз наблюдаемых людей.

Заключение

Способность к идентификации объектов внимания другого человека имеет ключевое значение для процесса коммуникации. При построении высказываний говорящий обычно учитывает, какой объект (событие) или какие аспекты этого объекта (события) находятся в фокусе внимания собеседника [3]. В современной психологической литературе процессы совместного внимания (joint attention) связывают преимущественно со способностью человека проследить взглядом направление взгляда собеседника (gaze following) [13; 15; 16; 17]. Многие работы опираются на концепцию Дж. Баттеруорта, постулирующую, что идентификация объекта внимания другого человека основывается на визуальной оценке ориентации его головы и глаз [11].

Между тем ориентация головы и глаз другого индивида является недостаточным источником информации об объекте его внимания. По направлению взгляда человека невозможно определить, на какой объект из множества расположенных рядом

или на какие аспекты объекта он направил свое внимание. Исходя из этого, выдвинута гипотеза о зависимости идентификации объектов внимания человека от понимания контекста коммуникативной ситуации. С использованием оригинальной экспериментальной процедуры было показано, что лица, обладавшие пониманием коммуникативного контекста, в отличие от лиц, не обладавших таким пониманием, демонстрируют высокую эффективность идентификации объектов внимания наблюдаемых участников коммуникативных ситуаций, независимо от степени доступности информации о направлении их взгляда.

Эти результаты свидетельствуют о несостоятельности концепций, сводящих совместное внимание к элементарному и «рефлекторному» процессу отслеживания направления взгляда человека по ориентации его головы и глаз [11; 18; 27]. В коммуникативных ситуациях идентификация объектов внимания другого человека обеспечивается сложными высокоуровневыми процессами, во многом сходными с процессами идентификации референтов речевых высказываний [5] и указательных жестов [6]. Результаты исследования также противоречат концепции «Mindreading» С. Барона—Коэна [8] и сходным с ней теориям [например: 13], постулирующим большое значение визуальной оценки направления взгляда («eye-direction detector», EDD) для понимания намерений и психического состояния другого человека.

Настоящая работа показала, что понимание контекста коммуникативной ситуации, определяющее успешность идентификации объектов внимания другого человека, предполагает создание и сопоставление когнитивных репрезентаций, отражающих ситуацию с различных точек зрения или перспектив. Этот вывод соответствует положениям теории «кругозора и окружения» М.М. Бахтина [1; 2], а также результатам исследований последних лет, свидетельствующим о ключевом значении полиперспективных репрезентаций (perspectival cognitive representations) для понимания вербальных и невербальных коммуникативных сигналов [21; 23; 26]. Впервые показано, что при восприятии невербальных коммуникативных сцен эти репрезентации обеспечивают выделение будущих объектов внимания и/или референтов предполагаемого взаимодействия наблюдаемых персонажей. Информация о локализации и визуальных признаках этих объектов/референтов избирательно запоминается и удерживается в рабочей памяти, что в дальнейшем позволяет легко идентифицировать референты взглядов и указательных жестов наблюдаемых людей, а также осуществлять практически мгновенный поиск этих объектов/референтов в визуальном пространстве сцены.

С использованием оригинальной модификации методики детекции изменений (flicker task) показано, что при восприятии видеоизображений социальных сцен наблюдатели произвольно направ-

ляют свое внимание на прогнозируемые будущие объекты внимания участников сцены и/или референты их предполагаемого коммуникативного взаимодействия, несмотря на отсутствие осознанной цели поиска данных объектов и визуальных подсказок в виде ориентации головы и глаз наблюдаемых людей. Этот результат позволяет реинтерпретировать феномен «слепоты к социальным изменениям» (social change blindness), описанный в работах зарубежных исследователей Д. Саймонса и Д. Левина [19; 25]. В одном из их экспериментов испытуемые просматривали короткий видеофрагмент следующего содержания. Актер сидит за столом и смотрит на лежащие перед ним документы. Услышав телефонный звонок, актер встает и проходит мимо камеры. В следующих кадрах камера показывает коридор, где уже другой актер подходит к телефону и отвечает на звонок³. После просмотра видеофильма испытуемых спрашивали, заметили ли они что-нибудь необычное. Обнаружено, что около 77% испытуемых не замечали подмену одного человека другим. Этот феномен «слепоты к изменению» авторы объясняют общей неспособностью человека к формированию «детализи-

рованных визуальных репрезентаций окружающего мира» [25, с. 648]. Между тем в свете полученных нами результатов этот феномен можно объяснить иначе. При восприятии вышеописанного видеофрагмента зритель непроизвольно формирует полиперспективную репрезентацию, отражающую сцену с точки зрения наблюдаемого человека. Эта репрезентация направляет его внимание на те объекты, которые релевантны предполагаемым намерениям персонажа («документы», «телефон»). Поскольку детали внешнего облика наблюдаемого человека отсутствуют в такой репрезентации, его подмена не замечается зрителем.

Настоящая работа ставит вопросы о механизмах формирования полиперспективных когнитивных репрезентаций, их структуре, динамике, а также роли, которую они играют в понимании семантики невербальных коммуникативных сцен. Есть основания полагать, что продуктивной теоретической основой при рассмотрении этих вопросов может являться теория концептуальной интеграции Ж. Фоконье и М. Тернера [14], однако требуются дальнейшие экспериментально-психологические и психолингвистические исследования.

Финансирование

Работа выполнена при финансовой поддержке РФФИ (проект №13-06-00616 «Когнитивные механизмы зрительного восприятия ситуаций социального взаимодействия в норме и патологии»).

Литература

1. *Бахтин М.М.* Антропологическая лингвистика: Избранные труды. М.: Лабиринт, 2010. 255 с.
2. *Бахтин М.М.* Эстетика словесного творчества / Сост. С.Г. Бочаров; примеч. С.С. Аверинцева и С.Г. Бочарова. М.: Искусство, 1979. 424 с.
3. *Ирисханова О.К.* Игры фокуса в языке. Семантика, синтаксис и прагматика дефокусирования. М.: Языки славянской культуры, 2014. 320 с.
4. Лингвистический энциклопедический словарь / Под ред. В.Н. Ярцевой; Ин-т языкознания АН СССР. М.: Сов. энцикл., 1990. 682 с.
5. *Падучева Е.В.* Высказывание и его соотношенность с действительностью (референциальные аспекты семантики местоимений). М.: Наука, 1985. 272 с.
6. *Томаселло М.* Истоки человеческого общения. М.: Языки славянских культур, 2011. 328 с.
7. *Фаликман М.В.* Внимание // *Общая психология: в 7 т. Т.4* / Под ред. Б.С. Братуся. М.: Издательский центр «Академия», 2006. 453 с.
8. *Baron-Cohen S.* Mindblindness: An essay on autism and theory of mind. Cambridge, MA: MIT Press, 1995. 183 p.
9. *Baron-Cohen S., Leslie A.M., Frith U.* Does the autistic child have a 'theory of mind'? // *Cognition*. 1985. Vol. 21 (1). P. 37–46.
10. *Bock S.W., Dicke P., Thier P.* How precise is gaze following in humans? // *Vision research*. 2008. Vol. 48. № 7. P. 946–957.

11. *Butterworth G.E., Jarrett N.* What minds have in common is space: Spatial mechanisms serving joint attention in infancy // *British Journal of Developmental Psychology*. 1991. Vol. 9. P. 55–72.
12. *Crouzet S., Kirchner H., Thorpe S.J.* Fast saccades toward faces: face detection in just 100 ms. // *Journal of Vision*. 2010. Vol. 10 (4). P. 1–17.
13. *Emery N.J.* The eyes have it: The neuroethology, function and evolution of social gaze // *Neuroscience and Biobehavioral Reviews*. 2000. Vol. 24. P. 581–604.
14. *Fauconnier G., Turner M.* The Way We Think: Conceptual Blending and the Mind's Hidden Complexities. N. Y.: Basic Books, 2002. 441 p.
15. *Friesen C.K., Kingstone A.* The eyes have it! Reflexive orienting is triggered by nonpredictive gaze // *Psychonomic Bulletin & Review*. 1998. Vol. 5. P. 490–495.
16. *Friesen C.K., Moore C., Kingstone A.* Does gaze direction really trigger a reflexive shift of attention? // *Brain and Cognition*. 2005. Vol. 57 (1). P. 66–69.
17. *Frischen A., Bayliss A.P., Tipper S.P.* Gaze cueing of attention: Visual attention, social cognition, and individual differences // *Psychological Bulletin*. 2007. Vol. 133. P. 694–724.
18. *Langton S., Watt R., Bruce V.* Do the eyes have it? Cues to the direction of social attention // *Trends in cognitive sciences*. 2000. Vol. 4 (2). P. 50–59.
19. *Levin D.T., Simons D.J.* Failure to detect changes to attended objects in motion pictures // *Psychonomic Bulletin and Review*. 1997. Vol. 4 (4). P. 501–506.

³ Видеофрагмент доступен по ссылке: <http://www.youtube.com/watch?v=wBoMjORwA-4>

20. *Malcolm G.L., Henderson J.M.* The effects of target template specificity on visual search in real-world scenes: Evidence from eye movements // *Journal of Vision*. 2009. Vol. 9 (11):8. P. 1–13.

21. *Moll H., Tomasello M.* Co-operation and human cognition: The Vygotskian intelligence hypothesis // *Philosophical Transactions of the Royal Society*. 2007. Vol. 362. P. 639–648.

22. *O'Regan J.K., Deubel H., Clark J.J., Rensink R.A.* Picture changes during blinks: Looking without seeing and seeing without looking // *Visual Cognition*. 2000. Vol. 7. P. 191–212.

23. *Parrill F.* Interactions between discourse status and viewpoint in co-speech gesture // B. Dancygier & E. Sweetser (Eds.) *Viewpoint in Language: A Multimodal Perspective*. Cambridge: Cambridge University Press, 2012. P. 97–112.

24. *Rensink R.A., O'Regan J.K., Clark J.J.* To see or not to see: The need for attention to perceive changes in scenes // *Psychological Science*. 1997. Vol. 8 (5). P. 368–373.

25. *Simons D.J., Levin D.T.* Failure to detect changes to people during a real-world interaction // *Psychonomic Bulletin & Review*. 1998. Vol. 5 (4). P. 644–649.

26. *Sweetser E.* Introduction: viewpoint and perspective in language and gesture, from the Ground down // B. Dancygier & E. Sweetser (Eds.) *Viewpoint in Language: A Multimodal Perspective*. Cambridge: Cambridge University Press, 2012. P. 1–24.

27. *Todorovic D.* Geometrical basis of perception of gaze direction // *Vision research*. 2006. Vol. 46 (21). P. 3549–3562.

28. *Walther D., Koch Ch.* Modeling attention to salient proto-objects // *Neural Networks*. 2006. Vol. 19. P. 1395–1407.

The Role of Polyperspective Representations in Joint Attention Processes

M.V. Zotov*

Saint Petersburg State University, Saint Petersburg, Russia,
mvzotov@mail.ru

N.Ye. Andrianova**

Saint Petersburg State University, Saint Petersburg, Russia,
natalia-andrianova@mail.ru

A.P. Voyt***

Saint Petersburg State University, Saint Petersburg, Russia,
voytalexey@mail.ru

The aim of our work was to test the hypothesis that one's identification of objects of attention in an observed individual depends on his/her understanding of the context of a communicative situation based on the comparison of the participants' points of view. In the first experimental study the subjects (N=74) carried out visual search for objects of attention of participants of non-verbal communication scenes with/without information on the communication context provided, and with short/prolonged display of information for assessing gaze direction of the participants. Understanding the context resulted in a higher efficiency of identification and saccadic detection of the participants' objects of attention in the subjects, regardless of their access to the information on gaze direction. In the second experiment the subjects (N=32), after watching videos of communication episodes, were asked to search for changes of different objects in the scenes. Those subjects who had a clear understanding of the context of communication first of all noticed the changes in the future objects of attention of the participants, in spite of the absence of any 'hints' like head or eye orientation. The outcomes of the research are discussed within the framework of concepts focusing on the significance of polyperspective representations for the understanding of communicative signals.

Keywords: theories of reference, joint attention, non-verbal communication, visual perception, eye movements.

Acknowledgements

The research was conducted with the assistance of the Russian Foundation for Basic Research (project #13-06-00616 "Cognitive Mechanisms in Normal and Abnormal Visual Perception of Social Interactions")

References

1. Bakhtin M.M. *Antropolingvistika: Izbrannye trudy* [Anthropological linguistics: Selected works]. Moscow: Labirint, 2010. 255 p.
2. Bakhtin M. M. *Estetika slovesnogo tvorchestva* [The Aesthetics of Verbal Creativity]. Bocharov S.G. (eds.). Moscow: Iskusstvo, 1979. 424 p.
3. Iriskhanova O.K. *Igry fokusa v yazyke. Semantika, sintaksis i pragmatika defokusirovaniya* [Focus game in language. Semantics, syntax and pragmatics of defocusing]. Moscow: Yazyki slavyanskoi kul'tury, 2014. 320 p.
4. Yartseva V.N.(ed.) *Lingvisticheskii entsiklopedicheskii slovar'* [Linguistic encyclopedic dictionary]. In-t yazykoznaniya AN SSSR. Moscow: Sov. entsikl., 1990. 682 p.
5. Paducheva E.V. *Vyskazyvanie i ego sootnesennost' s deistvitel'nost'yu* (referentsial'nye aspekty semantiki mestoimenii) [The statement and its correlation with reality (referential aspects of the semantics of pronouns)]. Moscow: Nauka, 1985. 272 p.
6. Tomasello M. *Istoki chelovecheskogo obshcheniya* [Origins of human communication]. Moscow: Yazyki slavyanskikh kul'tur, 2011. 328 p.

For citation:

Zotov M.V., Andrianova N.Ye., Voyt A.P. The Role of Polyperspective Representations in Joint Attention Processes. *Kul'turno-istoricheskaya psikhologiya = Cultural-historical psychology*, 2015. Vol. 11, no. 2, pp. 16–27. (In Russ., abstr. in Engl.). doi: 10.17759/chp.2015110202.

* *Zotov Mikhail Vladimirovich*, PhD in Psychology, professor at the Department of Medical Psychology and Psychophysiology, Faculty of Psychology, Saint Petersburg State University, Saint Petersburg, Russia. *mvzotov@mail.ru*

** *Andrianova Natalia Yevgenyevna*, Assistant at the Department of Medical Psychology and Psychophysiology, Faculty of Psychology, Saint Petersburg State University, Saint Petersburg, Russia. *natalia-andrianova@mail.ru*

*** *Voyt Alexei Petrovich*, PhD in Physics and Mathematics, leading engineer at the Faculty of Physics, Saint Petersburg State University, Saint Petersburg, Russia. *voytalexey@mail.ru*

7. Falikman M.V. Vnimanie. T. 4 serii "Obshchaya psikhologiya" [Attention. 4th tom of "General Psychology" series]. Bratus' B.S. (ed.). Moscow: Akademiya, 2006. 453 p.
8. Baron-Cohen S. Mindblindness: An essay on autism and theory of mind. Cambridge, MA: MIT Press, 1995. 183 p.
9. Baron-Cohen S., Leslie A.M., Frith U. Does the autistic child have a 'theory of mind'? *Cognition*, 1985. Vol. 21, no. 1, pp. 37–46.
10. Bock S.W., Dicke P., Thier P. How precise is gaze following in humans? *Vision research*, 2008. Vol. 48, no. 7, pp. 946–957.
11. Butterworth G.E., Jarrett N. What minds have in common is space: Spatial mechanisms serving joint attention in infancy. *British Journal of Developmental Psychology*, 1991. Vol. 9, pp. 55–72.
12. Crouzet S., Kirchner H., Thorpe S.J. Fast saccades toward faces: face detection in just 100 ms. *Journal of Vision*, 2010. Vol. 10, no. 4, pp. 1–17.
13. Emery N.J. The eyes have it: The neuroethology, function and evolution of social gaze. *Neuroscience and Biobehavioral Reviews*, 2000. Vol. 24, pp. 581–604.
14. Fauconnier G., Turner M. The Way We Think: Conceptual Blending and the Mind's Hidden Complexities. New York: Basic Books, 2002. 441 p.
15. Friesen C.K., Kingstone A. The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin & Review*, 1998. Vol. 5, pp. 490–495.
16. Friesen, C.K., Moore, C., Kingstone, A. Does gaze direction really trigger a reflexive shift of attention? *Brain and Cognition*, 2005. Vol. 57, no. 1, pp. 66–69.
17. Frischen A., Bayliss A.P., Tipper S.P. Gaze cueing of attention: Visual attention, social cognition, and individual differences. *Psychological Bulletin*, 2007. Vol. 133, pp. 694–724.
18. Langton S., Watt R., Bruce V. Do the eyes have it? Cues to the direction of social attention. *Trends in cognitive sciences*, 2000. Vol. 4, no. 2, pp. 50–59.
19. Levin D.T., Simons D.J. Failure to detect changes to attended objects in motion pictures. *Psychonomic Bulletin and Review*, 1997. Vol. 4, no. 4, pp. 501–506.
20. Malcolm G.L., Henderson J.M. The effects of target template specificity on visual search in real-world scenes: Evidence from eye movements. *Journal of Vision*, 2009. Vol. 9, no. 11:8, pp. 1–13.
21. Moll H., Tomasello M. Co-operation and human cognition: The Vygotskian intelligence hypothesis. *Philosophical Transactions of the Royal Society*, 2007. Vol. 362, pp. 639–648.
22. O'Regan J.K., Deubel H., Clark J.J., Rensink R.A. Picture changes during blinks: Looking without seeing and seeing without looking. *Visual Cognition*, 2000. Vol. 7, pp. 191–212.
23. Parrill F. Interactions between discourse status and viewpoint in co-speech gesture. In B. Dancygier & E. Sweetser (Eds.) *Viewpoint in Language: A Multimodal Perspective*. Cambridge: Cambridge University Press, 2012. pp. 97–112.
24. Rensink R.A., O'Regan J.K., Clark J.J. To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, 1997. Vol. 8, no. 5, pp. 368–373.
25. Simons D.J., Levin D.T. Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin & Review*, 1998. Vol. 5, no. 4, pp. 644–649.
26. Sweetser E. Introduction: viewpoint and perspective in language and gesture, from the Ground down. In B. Dancygier & E. Sweetser (Eds.) *Viewpoint in Language: A Multimodal Perspective*. Cambridge: Cambridge University Press, 2012. pp. 1–24.
27. Todorovic D. Geometrical basis of perception of gaze direction. *Vision research*, 2006. Vol. 46, Is. 21, pp. 3549–3562.
28. Walther D., Koch Ch. Modeling attention to salient proto-objects. *Neural Networks*, 2006. Vol. 19, pp. 1395–1407.