
Статьи

НА ПОВЫШЕННЫХ ТОНАХ: РОЛЬ ПРОСТРАНСТВЕННОГО ПОЗНАНИЯ В КРОСС-МОДАЛЬНОМ ВЗАИМОДЕЙСТВИИ ЭМОЦИОНАЛЬНОЙ СЕМАНТИКИ И АУДИАЛЬНОГО ВОСПРИЯТИЯ

Е.А. АНДРЮЩЕНКО^а, Е.Н. БЛИНОВА^а, Ю.Ю. ШТЫРОВ^б,
К.Г. МИРОШНИК^а, В.В. ТИМОХОВ^а, А. ДЖАНИЯН^с,
О.В. ЩЕРБАКОВА^а

^а Санкт-Петербургский государственный университет, 199034, Россия, Санкт-Петербург, Университетская наб., д. 7/9

^б Орхуский университет, 8000 Дания, Орхус, ул. Университетсбюэн, д. 3 стр. 1710

^с Новый Болгарский университет, 1618, Болгария, София, б-р. Монтевидео, д. 21

The Impact of Spatial Cognition on Cross-Modal Interaction between Emotional Semantics and Auditory Perception

Е.А. Andriushchenko^а, Е.Н. Blinova^а, Y. Shtyrov^б, К.Г. Miroshnik^а, V.V. Timokhov^а,
А. Janyan^с, О. V. Shcherbakova^а

^а Saint Petersburg State University, 7/9 Universitetskaya emb., Saint Petersburg, 199034, Russian Federation

^б Aarhus University, 3 bld. 1710 Universitetsbyen, Aarhus C, 8000, Denmark

^с New Bulgarian University, 21 Montevideo Str., Sofia, 1618, Bulgaria

Резюме

Теория воплощенного познания предполагает укоренение абстрактных, в том числе эмоциональных, концептов в сенсомоторном взаимодействии с физической средой. Так, верхняя область пространства ассоциируется с положительными

Abstract

According to the embodied cognition view, abstract concepts, including emotional ones, are grounded in our sensorimotor experience of the physical environment. For example, emotionally positive words

Исследование проведено при финансовой поддержке РНФ, проект № 22-28-01020 («Роль эмоциональной регуляции в мультисенсорной интеграции вербальной и невербальной информации: психологические и психофизиологические аспекты»).

The research was supported by RSCE, project N 22-28-01020 («The role of affective regulation in multisensory integration of verbal and non-verbal information: psychological and psychophysiological investigations»)

переживаниями, а нижняя — с отрицательными. Взаимодействие между различными доменами репрезентаций выражается в том числе в эффекте кросс-модального соответствия — фасилитации восприятия в одной модальности конгруэнтными аспектами стимулов в другой. Нами была впервые предпринята попытка вызвать данный эффект с помощью аудиальных стимулов различной высоты и вербальных стимулов, обозначающих эмоциональные состояния и имеющих соответствующие пространственные коннотации. В основной части исследования приняли участие 36 добровольцев, которым в ходе эксперимента одновременно предъявлялись аудиальные (тональные послышки 1000 и 2000 Гц) и вербальные (слова, различающиеся по эмоциональной валентности и соотносимые с различными частями пространственного поля) стимулы. Задача респондентов заключалась в идентификации высоты (низкая/высокая) предъявленного тона. Анализ различий во времени реакции с помощью линейных смешанных моделей показал наличие статистически значимых различий между конгруэнтным (например, эмоционально положительное слово — высокий тон) и неконгруэнтным, а также конгруэнтным и контрольным условиями. Помимо этого было оценено наличие эффекта кросс-модального соответствия для каждого стимульного слова и для обоих звуков по отдельности, что показало его наибольшую выраженность для высокого тона и для части вербальных стимулов. Таким образом, был обнаружен эффект кросс-модального соответствия, возникающий при когнитивной обработке стимулов разной модальности, осуществляемой за счет механизмов разного уровня сложности: от перцептивного анализа аудиальных сигналов до выделения семантики лексических единиц, обозначающих абстрактные идеи. Отдельным результатом работы стало создание базы данных вербальных стимулов с заданными эмоциональными и пространственными параметрами, которые могут быть использованы в широком круге психо- и нейролингвистических исследований.

Ключевые слова: кросс-модальное соответствие, ориентационная метафора, пространственная коннотация, эмоциональное восприятие, воплощенное познание.

(happiness) are associated with the upper part of vertical space, whereas the negative ones (desperation) are linked with the lower parts of physical space. Interactions between different representational domains are expressed, among others, in the so-called cross-modal correspondence effect — facilitation of stimulus processing in one modality by congruent aspects of information presented in another modality (e.g., high-pitch sound — high vertical location). The present study attempted, for the first time, to induce this effect using auditory stimuli of varied pitch and emotional words with defined spatial connotations. Thirty-six volunteers (26 women, 18–34 years old) were simultaneously presented with 1000 or 2000 Hz tones and words of varied emotional valence that had been rated for their spatial associations in vertically oriented space. The participants' task was to identify the pitch of the presented tone as high or low. The analysis of reaction times using a linear mixed-effects model demonstrated statistically significant differences between congruent (e.g., emotionally positive word — high tone) and noncongruent conditions, as well as between congruent and neutral conditions. The effect of cross-modal correspondence was also evaluated for each stimulus word as well as for high and low tones separately, showing its specificity to high-pitch sounds and a subset of words. In sum, the results showed a cross-modal correspondence effect that spans across different levels of cognitive processing: from perceptual analysis of auditory signals to verbal semantic processing of complex abstract ideas. Importantly, we also present a set of Russian-language nouns rated for their spatial and emotional properties that can be used in future psycho- and neurolinguistic studies on embodied semantics.

Keywords: cross-modal correspondence, orientational metaphor, spatial representation, emotional perception, embodied cognition.

Андриющенко Екатерина Александровна — инженер-исследователь, факультет психологии, Санкт-Петербургский государственный университет.
Сфера научных интересов: психология эмоций, психолингвистика.
Контакты: kateand625@gmail.com

Блинова Екатерина Николаевна — инженер-исследователь, факультет психологии, Санкт-Петербургский государственный университет.
Сфера научных интересов: психолингвистика, психология чтения.
Контакты: blinova_e.n@mail.ru

Штыров Юрий Юрьевич — профессор, ведущий научный сотрудник, Центр функционально-интегративной нейронауки, Орхусский университет (Дания), PhD in Psychology.
Сфера научных интересов: нейробиология языка и речи.
Контакты: yury@cfin.au.dk

Мирошник Кирилл Геннадьевич — аспирант, факультет психологии, Санкт-Петербургский государственный университет.
Сфера научных интересов: психология креативности.
Контакты: cyril.miroshnik@gmail.com

Тимохов Виктор Викторович — студент магистратуры, Санкт-Петербургский государственный университет.
Сфера научных интересов: нейроэкономика, когнитивная нейронаука.
Контакты: viktor-timohov@mail.ru

Джаниян Армина — доцент, Исследовательский центр когнитивных наук, Новый Болгарский университет, PhD in Psychology.
Сфера научных интересов: психолингвистика.
Контакты: ajanyan@cogs.nbu.bg

Щербакова Ольга Владимировна — доцент, факультет психологии, кафедра общей психологии, Санкт-Петербургский государственный университет, кандидат психологических наук, доцент.
Сфера научных интересов: когнитивная психология, когнитивная нейронаука, психология интеллекта и креативности, психолингвистика.
Контакты: o.shcherbakova@spbu.ru

Ekaterina A. Andriushchenko — Junior Research Fellow, Faculty of Psychology, Saint Petersburg State University.
Research Area: psychology of emotion, psycholinguistics.
E-mail: kateand625@gmail.com

Ekaterina N. Blinova — Junior Research Fellow, Faculty of Psychology, Saint Petersburg State University.
Research Area: psycholinguistics, psychology of reading.
E-mail: blinova_e.n@mail.ru

Yury Shtyrov — Professor, Principal Investigator, Center for Functionally Integrative Neuroscience, Aarhus University, PhD in Psychology.
Research Area: neurobiology of speech and language.
E-mail: yury@cfin.au.dk

Kirill G. Miroshnik — Ph.D. student, Faculty of Psychology, Saint Petersburg State University.
Research Area: psychology of creativity.
E-mail: cyril.miroshnik@gmail.com

Viktor V. Timokhov — Master's student, Saint Petersburg State University.
Research Area: neuroeconomics, cognitive neuroscience.
E-mail: viktor-timohov@mail.ru

Armina Janyan — Assistant Professor, Research Center for Cognitive Science, New Bulgarian University, PhD in Psychology.
Research Area: psycholinguistics.
E-mail: ajanyan@cogs.nbu.bg

Olga V. Shcherbakova — Associate Professor, Faculty of Psychology, Department of General Psychology, Saint Petersburg State University, PhD in Psychology, Associate Professor.
Research Area: cognitive psychology, cognitive neuroscience, intelligence and creativity, psycholinguistics.
E-mail: o.shcherbakova@spbu.ru

В рамках концепции воплощенного познания (embodied cognition) принято считать, что устройство концептуальной системы человека базируется на его сенсомоторном опыте (Barsalou, 2008). Предполагается, что концепты представляют собой обобщенные ментальные реконструкции реальных свойств объектов, которые они репрезентируют. Например, к числу таких свойств может быть отнесено расположение референтного концепту объекта в естественной среде. При этом укоренение новых концептуальных структур в опыте может происходить и косвенно, за счет активации в памяти уже известных слов, имеющих сенсомоторные коннотации, что обеспечивает концептуальной системе возможность выхода за рамки непосредственно воспринимаемого (Günther et al., 2020).

Одной из наиболее влиятельных теорий, развивающих эту идею, является теория концептуальной метафоры, согласно которой в основе процессов метафоризации лежат процедуры обработки таких структур знаний, как фреймы и сценарии (Лакофф, Джонсон, 2004). Свойства этих структур, с одной стороны, обусловлены генетически детерминированной сенсомоторной организацией человека, а с другой — индивидуальными особенностями его отношений с физической средой. Другими словами, упорядочивание и структурирование опыта осуществляется как за счет работы биологически обусловленных механизмов, так и за счет разнообразных средовых воздействий, — например, культуры (Lakoff, Johnson, 1980).

Одним из основных видов метафор являются ориентационные метафоры, т.е. основанные на восприятии расположения объектов в пространстве. Метафоры такого типа образуются произвольно, связаны с пространственными отношениями (например, «верх — низ») и, как предполагают авторы, возникают в силу того, что человеку свойственно определенным способом взаимодействовать с внешней средой на телесном уровне. Так, все объекты и ситуации, ориентированные вверх, часто наделяются положительным эмоциональным значением, а объекты и ситуации, ориентированные вниз, — отрицательным. Физической основой такого отождествления, вероятно, является то, что склоненная поза человека обычно отражает печаль и депрессивное состояние, а прямая поза — позитивный эмоциональный настрой (Лакофф, Джонсон, 2004). В языке ориентационные метафоры проявляются в таких фразах, как, например, «я чувствую себя на вершине блаженства», «быть на седьмом небе от счастья» или, напротив, — «настроение ниже плинтуса», «провалиться сквозь землю». Как показывают эти примеры, эмоциональная валентность метафорических высказываний напрямую связана с когнитивной схемой организации ментального пространства: положительно эмоционально окрашенные слова ассоциированы с верхней частью пространственного поля, а негативно эмоционально окрашенные — с нижней. Следовательно, когнитивная обработка слов, обозначающих эмоциональное состояние, вероятно, сопровождается актуализацией ментальной схемы пространства, ориентированного относительно вертикали. Данное предположение подтверждается результатами исследований, продемонстрировавших, что время реакции в ответ на конгруэнтное (соответствующее предполагаемой

пространственной ассоциации) предъявление вербальных стимулов, обладающих эмоциональной коннотацией и расположенных на разной высоте, оказывается значительно меньше времени реакции при неконгруэнтном (несоответствующем) предъявлении стимулов, — например, идентификация положительно окрашенных слов происходит быстрее, если они расположены в верхней части зрительного поля (Gozli et al., 2013; Woodin, Winter, 2018).

Интересно, что описанный эффект также проявляется на уровне анализа фонологических характеристик стимулов. Так, в одном из исследований (Auracher, 2017) участникам было необходимо соотнести псевдослова с различным артикуляционным местом гласных (которое зависит от степени и места подъема языка и того, принимают ли участие губы в произнесении звука) с изображениями животных и иллюстрациями поз тела, выражающих эмоции. Авторы предположили, что место артикуляции гласных будет влиять на соотнесение содержащих их псевдослов с изображениями различных животных или поз тела, имеющими эмоциональную окраску (например, физическое или социальное доминирование). Так, предполагалось, что угрожающая поза будет соотноситься с псевдословами, в которых присутствуют «задние гласные» (язык при произнесении сдвинут назад, как в псевдослове «котопу»). Результаты показали, что на соотнесение артикуляционно-акустических характеристик фонем с изображениями влияют семантические признаки представленных на картинках объектов. Полученные данные позволили заключить, что абстрактные семантические понятия могут функционировать как интерфейс между различными сенсорными системами, облегчая возникновение соответствия между признаками разной модальности — так называемого кросс-модального соответствия (Ibid.). Можно предположить, что описанные авторами результаты обусловлены актуализацией моторных схем, опосредующих артикуляторные процессы. Однако альтернативное объяснение заключается в возникновении у человека целостных пространственных репрезентаций, актуализирующихся при выполнении когнитивной обработки различного уровня сложности: от идентификации простых перцептивных характеристик стимулов до выделения семантики абстрактных понятий. Поскольку формирование таких репрезентаций может представлять собой один из ключевых познавательных механизмов, перспективным представляется изучение вопроса о том, возможна ли модуляция эффекта кросс-модального соответствия с помощью перцептивных аудиальных и вербальных (семантически нагруженных) стимулов, ассоциированных с разными частями пространственного поля.

Кросс-модальное соответствие определяют как эффект совместимости/сопоставимости свойств стимулов различных сенсорных модальностей, — например, звуков высокой частоты и объектов, расположенных в верхней части зрительного поля (Spence, 2011). В частности, результаты решения задач, содержащих конгруэнтные (соответствующие друг другу по определенной характеристике) стимулы различной модальности, значительно отличаются от результатов, полученных при работе с неконгруэнтными (не соответствующими друг другу) стимулами. Такой эффект возникает у многих

людей и, возможно, даже является универсальным человеческим свойством, проявляющимся достаточно рано: известно, что уже к двум годам дети способны соотнести громкие звуки с крупными геометрическими формами (Smith, Sera, 1992), а в пятилетнем возрасте надежно сопоставляют яркость визуального стимула с громкостью звука (Bond, Stevens, 1969).

Описано большое количество типов кросс-модального соответствия (Spence, 2011), среди которых особый интерес представляет аудиовизуальное. В ряде исследований изучалась взаимосвязь восприятия частоты звука и различных характеристик визуального стимула. Так, в одной из работ было обнаружено соответствие между высотой воспринимаемого звука и размером предъявляемого одновременно с ним изображения, а также его вертикальным расположением в пространстве (Evans, Treisman, 2010). При рассмотрении вопроса о том, как яркость, насыщенность, размер и вертикальное положение визуального стимула могут влиять на определение частоты аудиального стимула, была обнаружена общая закономерность суммирования частных кросс-модальных соответствий: наличие ассоциаций высокого тона звука с высокой яркостью, высокой насыщенностью, небольшим размером и высоким расположением визуального стимула в пространстве (Jonas et al., 2017). Авторы другого исследования также выявили эффект аудиовизуального соответствия между высотой тона и размером визуального стимула, отметив, что он является относительным, т.е. зависит от контекста (другими словами, идентификация объекта как большого или маленького зависит от стимулов, которые предъявлялись до этого), и достаточно гибким, так как интерпретация одного и того же стимула способна меняться на противоположную (Brunetti et al., 2018).

На данный момент не существует единого мнения о том, как именно влияет соответствие между слуховыми и зрительными стимулами на качество итогового перцептивного образа. Некоторые эмпирические данные свидетельствуют о положительном эффекте такого соответствия между характеристиками аудиальных и визуальных стимулов: например, в одном из экспериментов было показано, что установление связей между конгруэнтной информацией, поступающей по разным сенсорным каналам, повышает общее качество восприятия (Störmer, 2019). Также в ряде поведенческих исследований было продемонстрировано, что визуально-пространственная и аудиальная информация взаимодействуют между собой и в случае, когда эти стимулы оказываются конгруэнтны друг другу, их обработка ускоряется (например: Evans, 2020). Однако в серии психофизических экспериментов было обнаружено негативное влияние кросс-модального соответствия на качество восприятия: участникам было значительно труднее различать расположение или оценивать время предъявления аудиального стимула при одновременном предъявлении конгруэнтного ему визуального стимула по сравнению с неконгруэнтным (Parise, Spence, 2009). Наличие таких противоречивых результатов позволяет предположить, что конгруэнтность стимулов оказывает избирательное влияние на возникновение эффекта кросс-модального соответствия при выполнении задач, характеризующихся определенным уровнем когнитивной сложности.

Возникновение эффекта кросс-модального соответствия между частотой звука и высотой расположения объектов в пространстве объясняют по крайней мере двумя разными способами. С одной стороны, предполагается, что психика может усваивать лежащие в основе этого эффекта закономерности окружающей среды за счет научения (Ibid.). Однако более полное объяснение того обстоятельства, что для обозначения как параметров аудиальных сигналов (высокой и низкой частоты звука), так и расположения визуальных стимулов в пространстве воспринимаемого или воображаемого зрительного поля мы используем одинаковый набор характеристик, возможно в рамках гипотезы лингвистического опосредования (Martino, Marks, 1999; Spence, 2011). Так, слова «высокий» и «низкий» описывают и частоту звука, и пространственную высоту, воспринимаемую зрительно (Spence, 2020). Другими словами, судя по полученным ранее данным, эффект аудиовизуального соответствия нестабилен и, возможно, связан с устройством речевого опыта человека (Spence, 2011; Puigserver et al., 2019). В пользу такого вывода говорят и данные о том, что носители разных языков различаются по своим пространственно-временным ассоциациям: например, в голландском языке, как и в русском, используют метафору высоты тона, а в турецком — метафору его толщины («тонким» называют высокий тон звука, а «толстым» — низкий). В эксперименте (Dolscheid et al., 2020), в ходе которого участникам предъявляли звук того или иного тона и просили выбрать либо толстую линию, расположенную высоко, либо тонкую линию, расположенную низко, было показано, что носители турецкого и голландского языков не только используют разные метафоры пространственного тона, но и по-разному ассоциируют высоту тона с пространственным расположением объекта: в то время как носители турецкого языка для сопоставления с высоким звуком преимущественно выбирали тонкую линию, расположенную внизу экрана, носители голландского языка отдавали предпочтение толстой линии, расположеннойверху экрана. При этом, как отмечают авторы, такого рода ассоциации оказываются в разной степени восприимчивы к языковым факторам: например, соответствие между тоном звука и пространственной локализацией является более гибким и в большей степени подверженным влиянию конкретного языка, по сравнению с другими типами кросс-модального соответствия (Ibid.). В другом исследовании было показано, что аудиовизуальное соответствие возникает как у носителей английского языка, в котором для обозначения тона звука и расположения объекта в пространстве используются одинаковые слова, так и у носителей испанского и каталанского языков, в которых для обозначения этих категорий используются различные прилагательные. Однако у англоязычных участников был обнаружен больший эффект аудиовизуального соответствия при определении высоты звука, что также говорит о влиянии языка на проявление данной разновидности кросс-модального соответствия. При этом в случае с громкостью звука никаких значимых различий между группами обнаружено не было: носители обоих языков используют идентичные слова для описания громкости и пространственной локализации (Fernandez-Prieto et al., 2017). Можно предполагать, что аудиовизуальное соответствие возникает уже на

уровне простой перцептивной обработки воспринимаемой человеком информации, однако использование того или иного языка может усиливать возникающие ассоциации.

Таким образом, имеющиеся на данный момент эмпирические данные являются достаточно разнородными. Во-первых, не существует единого мнения относительно того, как именно — положительно (повышая качество итоговой перцептивной обработки и уменьшая время реакции) или отрицательно (снижая качество восприятия и увеличивая время ответа) — влияют друг на друга конгруэнтные характеристики стимулов различной модальности. Во-вторых, неясно, в какой мере языковые средства могут опосредовать возникновение эффекта аудиовизуального соответствия. В частности, если предполагать, что кросс-модальные соответствия могут возникать не только на перцептивном, но и на более высоком уровне семантической обработки, то необходимо ответить на вопрос о том, как влияет активация общего лингвистического или метафорического кода на связь между пространственным расположением визуального объекта и характеристиками звука, такими как высота тона или громкость. Например, известно, что слова «высокий» и «низкий» в английском языке, как и в русском, активируют и слуховые, и пространственные компоненты концептуальных схем (*Ibid.*). Это означает, что когда для определения звукового тона используется слово «высокий», одновременно может быть активировано мысленное представление об объекте, имеющем высокое пространственное расположение. Однако до сих пор неясно, наблюдается ли аналогичная тенденция при обработке вербальных стимулов более комплексного характера — например, языковых метафор, в которых пространственная коннотация служит для передачи сложной абстрактной идеи (переживаемой человеком эмоции). Поэтому в рамках настоящей работы нами была предпринята одна из первых попыток ответить на данный вопрос. Целью нашего исследования стало изучение пространственных аспектов семантики вербальных стимулов, описывающих эмоциональные состояния, путем создания эффекта кросс-модального соответствия при решении задачи на определение высоты звука.

Стимульный материал исследования

Для проведения экспериментальной части исследования необходимо было создать стимульный материал, который включал бы визуальные (семантически нагруженные, вербальные) и аудиальные (невербальные) стимулы для проверки гипотезы о возникновении эффекта кросс-модального соответствия.

Вербальные стимулы

На начальном этапе работы мы создали базу психолингвистических стимулов, характеризующихся пространственными коннотациями. Эта задача была реализована в две последовательные ступени. На первой ступени двумя экспертами

(первым и вторым авторами данной статьи — бакалавром и магистром психологии, специализирующимися в области психолингвистики и когнитивных наук и имеющими опыт исследовательской работы в этой сфере) были составлены наборы стимульных слов, обозначающих 1) объекты или явления физической среды и части человеческого тела (60 слов, например, «штиль», «ключица», см. таблицу 1), из которых в дальнейшем после их независимой оценки отбирались контрольные нейтральные стимулы для основного эксперимента, и 2) эмоциональные состояния (60 слов, например, «блаженство», «гнев», см. таблицу 2), из которых отбирались целевые стимулы для создания эффекта кросс-модального соответствия. Стимульные слова первой группы были отобраны путем анализа словарей; слова второй группы были получены при работе со специализированной базой данных (Bradley, Lang, 1999), а затем переведены на русский язык.

Таблица 1

**Оценка пространственной локализации природных объектов и частей тела
по семибальной шкале**

<i>Природные объекты</i>								
	М	SD		М	SD		М	SD
Солнце	6.75	0.55	Даль	4.44	1.68	Трава	1.72	0.70
Луна	6.61	0.73	Лес	4.03	1.18	Газон	1.61	0.60
Небо	6.39	0.69	Штиль	3.50	1.40	Песок	1.58	0.69
Облако	5.97	0.74	Огонь	2.47	1.03	Лужа	1.50	0.65
Радуга	5.89	0.78	Вода	1.89	1.06	Почва	1.42	0.60
Туча	5.89	0.78	Тропа	1.75	0.77	Пропасьт	1.28	0.78
Молния	5.61	1.13	Камень	1.81	0.62	Яма	1.22	0.48
<i>Части тела</i>								
	М	SD		М	SD		М	SD
Макушка	4.50	1.48	Глаз	4.14	1.05	Локоть	3.47	0.97
Лоб	4.31	1.35	Ухо	4.14	1.25	Спина	3.33	0.83
Веко	4.28	1.32	Скула	4.11	1.21	Живот	3.17	0.77
Висок	4.28	1.32	Борода	4.11	1.21	Талия	3.14	0.76
Ресницы	4.28	1.30	Щека	4.06	1.17	Пупок	3.00	0.79
Голова	4.25	1.25	Усы	4.06	1.22	Таз	2.72	0.70
Бровь	4.25	1.25	Рот	4.03	1.18	Бедро	2.89	0.78
Лицо	4.22	1.27	Язык	3.97	1.30	Ягодицы	2.83	0.85
Чуб	4.22	1.42	Шея	3.78	1.15	Колено	2.53	0.61
Челка	4.19	1.35	Лопатка	3.64	1.02	Голен	2.25	0.65
Челюсть	4.19	1.28	Плечо	3.56	0.91	Пятка	1.67	0.79
Зубы	4.19	1.31	Ключица	3.47	0.94	Стопа	1.61	0.73
Нос	4.17	1.25	Грудь	3.47	0.94	Носок	1.61	0.69

Таблица 2
Оценка характеристик стимулов, обозначающих эмоциональные состояния и ассоциированных с разной пространственной локализацией

Слово	Валентность			Интенсивность			Контроль			Локализация		
	M	SD	Min	Max	M	SD	Min	Max	M	SD	Min	Max
Подавленность	2.30	0.77	1.00	4.00	3.73	1.64	1.00	7.00	3.36	1.34	2.00	6.00
Разочарование	2.03	0.85	1.00	4.00	4.33	1.41	2.00	7.00	3.79	1.47	1.00	7.00
Мука	1.73	1.10	1.00	6.00	4.94	1.90	1.00	7.00	2.58	1.32	1.00	6.00
Грусть	2.79	0.96	1.00	5.00	3.39	1.25	1.00	6.00	4.27	1.31	2.00	7.00
Тоска	2.82	1.10	1.00	6.00	4.36	1.69	2.00	7.00	3.73	1.51	1.00	7.00
Презрение	2.12	0.89	1.00	4.00	3.88	1.76	1.00	7.00	4.64	1.62	1.00	7.00
Огорчение	2.76	0.61	1.00	4.00	3.79	1.19	1.00	6.00	4.09	1.57	1.00	7.00
Благоговение	5.39	1.09	3.00	7.00	4.30	1.78	1.00	7.00	5.12	1.56	1.00	7.00
Страсть	5.00	0.97	3.00	7.00	5.58	1.44	1.00	7.00	4.21	1.80	1.00	7.00
Удовольствие	5.55	0.75	4.00	7.00	4.03	1.29	2.00	7.00	4.94	1.66	1.00	7.00
Веселье	5.73	0.76	4.00	7.00	5.15	1.00	3.00	7.00	5.06	1.56	1.00	7.00
Уверенность	5.85	0.91	4.00	7.00	4.00	1.71	1.00	7.00	4.36	1.41	1.00	7.00
Радость	6.06	0.66	5.00	7.00	4.82	1.33	2.00	7.00	5.03	1.59	1.00	7.00
Блаженство	6.06	1.03	4.00	7.00	4.55	1.87	1.00	7.00	4.70	1.69	1.00	7.00
Торжествование	5.61	0.97	3.00	7.00	5.24	1.50	1.00	7.00	5.06	1.60	1.00	7.00
Восхищение	5.82	0.98	3.00	7.00	5.15	1.33	3.00	7.00	5.06	1.46	2.00	7.00
Ликование	5.94	0.97	4.00	7.00	5.67	1.45	1.00	7.00	4.67	1.63	1.00	7.00
Вдохновение	6.30	0.88	3.00	7.00	5.00	1.25	2.00	7.00	3.30	1.93	1.00	7.00
Любовь	6.24	1.00	4.00	7.00	4.91	1.61	1.00	7.00	3.88	1.98	1.00	7.00
Свобода	6.09	1.10	4.00	7.00	4.00	1.98	1.00	7.00	4.27	1.57	1.00	7.00
Счастье	6.58	0.61	5.00	7.00	5.18	1.55	2.00	7.00	4.09	1.84	1.00	7.00
Восторг	6.39	0.61	5.00	7.00	5.61	1.14	2.00	7.00	4.39	1.58	1.00	7.00

Отобранные стимулы, обозначающие объекты физической среды и части тела, были протестированы на независимой выборке ($n = 36$; впоследствии никто из участников данного этапа не вошел в основную выборку эксперимента) с помощью онлайн-опроса через Google Forms. Респондентам было необходимо указать с помощью семибальной шкалы Ликерта, где обычно располагается данный объект относительно вертикальной оси (1 — «очень низкое положение», 7 — «очень высокое положение»). Были подготовлены две параллельные формы для контрбалансировки порядка предъявления стимулов.

Далее аналогичная процедура была проведена со стимулами, семантика которых связана с эмоциональным состоянием. В данном случае участникам ($n = 32$; никто из респондентов этого этапа не принимал участие в предшествующем опросе) помимо оценки пространственной локализации было предложено оценить слова по еще трем семибальным шкалам: 1) шкале эмоциональной валентности (1 — «очень негативное состояние», 7 — «очень позитивное состояние»), 2) шкале эмоциональной силы/интенсивности (1 — «очень низкая интенсивность», 7 — «очень высокая интенсивность») и 3) шкале, отражающей способность человека контролировать обозначенное словом эмоциональное состояние (1 — «эмоция совершенно не поддается контролю», 7 — «эмоция легко поддается контролю»). Было подготовлено четыре параллельные формы опроса, где названные показатели предлагались для оценки в различном порядке.

По результатам проведения статистического анализа данных (учитывались средние значения оценок, позволяющие оценить субъективное восприятие пространственной локализации понятия участниками, и стандартные отклонения оценок, служащие индикатором согласованности вынесенных респондентами суждений), собранных с помощью этих двух опросов, нами были отобраны вербальные стимулы для основного этапа исследования (см. таблицу 3). В их перечень вошли слова, вызвавшие у участников наиболее однозначные пространственные ассоциации, в том числе слова, обозначающие эмоциональные состояния, отнесенные большинством участников к верхней («восторг») или нижней («тоска») части вертикальной пространственной оси, а также нейтральные слова — природные объекты и части тела («даль», «живот»), отнесенные респондентами к средней части вертикальной оси и используемые в эксперименте в качестве контрольных. Для каждой группы слов (ассоциированных с верхней позицией в пространстве — позитивных, с нижней позицией — негативных, а также со средней — нейтральных) было отобрано пять наиболее отвечающих требованиям стимулов. Из-за относительных различий в представленности слов, отражающих эмоциональные состояния (по сравнению с такими категориями, как «предметы» или «объекты окружающей среды»), нам не удалось стандартизовать стимулы по всем психолингвистическим характеристикам (в частности, длине и частотности словоупотребления). Однако мы проконтролировали наиболее релевантные для нашего исследования семантические параметры слов — валентность и интенсивность ассоциированных с ними эмоциональных состояний, а также особенности восприятия их пространственных коннотаций.

Таблица 3

Стимуляные слова, использованные при проведении основного эксперимента

Слово	Оценка пространственной локализации				Частотность, ipm
	M	SD	Min	Max	
Позитивные стимулы					
Восторг	6.12	0.64	4.00	7.00	56.2
Счастье	6.09	0.72	4.00	7.00	149.2
Вдохновение	5.79	0.90	4.00	7.00	12.8
Ликование	5.73	0.88	2.00	7.00	5.6
Восхищение	5.67	0.95	3.00	7.00	18.3
Негативные стимулы					
Подавленность	2.39	0.67	1.00	5.00	1.2
Разочарование	2.55	0.93	1.00	5.00	16.6
Тоска	2.94	0.75	1.00	5.00	48.5
Презрение	2.97	1.07	1.00	7.00	14.5
Огорчение	3.00	0.55	2.00	5.00	7.1
Нейтральные стимулы					
Живот	3.17	0.77	2.00	5.00	65.6
Талия	3.14	0.76	2.00	4.00	10.6
Пупок	3.00	0.79	1.00	4.00	4.1
Даль	4.44	1.68	1.00	7.00	18
Лес	4.03	1.18	1.00	7.00	211.5

Аудиальные стимулы

В качестве невербальных аудиальных стимулов были использованы тональные послыки разной частоты — звуковые сигналы продолжительностью 100 мс с высотой тона 1000 и 2000 Гц. Аудиальные стимулы были подготовлены с использованием звукового редактора Audacity v. 2.3.0 ®¹.

Выборка исследования

Выборка основного этапа исследования состояла из 36 добровольцев (26 женщин, 18–34 года, M = 23.36, SD = 3.93). В их число вошли люди с высшим или неоконченным высшим образованием разного профиля. С помощью предварительного онлайн-анкетирования были отобраны участники, соответствовавшие следующим требованиям: ведущая рука — правая; русский является единственным родным языком; нормальные или скорректированные до

¹ Программное обеспечение Audacity® защищено авторским правом (© 1999 2021 Audacity Team. URL: <https://audacityteam.org/>) и распространяется на условиях GNU General Public License. Audacity® является зарегистрированной торговой маркой.

нормальных зрения и слух; отсутствие неврологических и психиатрических заболеваний. Все респонденты до начала исследования подписали информированное согласие, одобренное Этическим комитетом Санкт-Петербургского психологического общества.

Экспериментальная парадигма исследования

Для разработки экспериментальной парадигмы использовалась программная среда NBS Presentation v.21.0 (Neurobehavioral Systems, Беркли, Калифорния, США).

В ходе эксперимента участники располагались на расстоянии 60–65 см от экрана компьютера. Аудиальные стимулы предъявлялись через наушники с интенсивностью 75 дБ, а вербальные стимулы были представлены визуально на мониторе компьютера (ASUS ROG PG278Q, частота обновления: 144 Гц, диагональ: 27"; разрешение экрана: 2560 × 1440 пикселей) и предъявлялись в центре экрана на сером фоне (130, 130, 130 в цветовом пространстве RGB).

Полная структура данного этапа исследования может быть представлена совокупностью следующих процедур:

1. Обучение. Респондентам для ознакомления последовательно предъявлялись два звука разной частоты (1000 и 2000 Гц), параллельно на экране в письменной форме указывалась релевантная характеристика звука (является ли он высоким или низким; 8 проб).

2. Тренировочная серия. Респондентам предоставлялась инструкция и предлагалось потренироваться в выполнении экспериментальной задачи, описанной далее (использовались вербальные стимулы, не задействованные в основной части эксперимента; 10 проб).

3. Основная серия. Участникам предъявлялся целевой звук (1000 Гц или 2000 Гц) вместе с указанием нажать на клавиатуре кнопку, которая соответствует верной характеристике этого звука (высокий или низкий). Звуки подавались с помощью наушников одновременно со словами, которые предъявлялись на экране монитора в течение 100 мс. Вербальные стимулы были разделены на три группы в зависимости от их эмоциональной валентности и локализации в пространстве (положительный/отрицательный/нейтральный — для вербальных стимулов, отражающих эмоциональные состояния; высокий/низкий/нейтральный — для вербальных стимулов, обозначающих предметы физической среды). Последовательность предъявления в ходе эксперимента была следующей: фиксационный крест (500 мс), стимулы (звук и слово; 100 мс), окно для ответа (2000 мс), межстимульный интервал (1000 мс) (см. рисунок 1). Всего каждому респонденту было предъявлено 180 пар стимулов (15 слов × 2 звука × 6 повторений) в псевдорандомизированном порядке. Пары стимулов были распределены на три группы: 1) конгруэнтные стимулы (слова, обозначающие положительные эмоции, и высокий звук; слова, обозначающие отрицательные эмоции, и низкий звук); 2) неконгруэнтные стимулы (слова, обозначающие положительные эмоции, и низкий звук; слова, обозначающие отрицательные эмоции, и высокий звук); 3) контрольные стимулы

**Процедура предъявления стимулов в ходе выполнения участниками
экспериментального задания**



(нейтральные слова и высокий/низкий звук). Ожидалось, что эффект кросс-модального соответствия проявится в меньшем времени реакции на конгруэнтные пары стимулов в сравнении с неконгруэнтными парами. Общее время экспериментальной серии составило примерно 15 минут.

Для выполнения участниками экспериментального задания были использованы первая и вторая кнопки на пульте Cedrus RB-740 (Cedrus Corp., Сан-Педро, Калифорния, США; расстояние между кнопками составляло 30 мм), позволяющем наиболее точно фиксировать время реакции и правильность ответа, которые выступали критериями для оценки проявления эффекта аудиовизуального соответствия. Для половины респондентов кнопка «1» обозначала высокий звук, а кнопка «2» — низкий; для второй половины респондентов — наоборот. Необходимость смотреть на экран неоднократно подчеркивалась во время инструктажа, а соблюдение этого требования контролировалось с помощью камеры, находящейся справа от участника и позволяющей отслеживать направление его взгляда.

Математико-статистические методы обработки данных

Анализ данных осуществлялся с помощью линейных смешанных моделей в программной среде RStudio 2021.9.2.382 с применением пакета *lme4* (Bates et al., 2015). В качестве случайного эффекта во всех моделях выступал фактор респондента, а в качестве фиксированного эффекта — фактор конгруэнтности. Так как общая экспериментальная процедура включала показ видеороликов разной эмоциональной валентности, влияние которых не представляет интереса

для настоящей работы², было принято решение учитывать данный фактор как контрольную переменную, которая вводилась в качестве фиксированного эффекта. Для оценки параметров модели применялся метод ограниченного максимального правдоподобия. Вклад фиксированных эффектов определялся с помощью t -критерия Вальда (с аппроксимацией Кенварда — Роджера), доверительных интервалов и величины эффекта по d_z Коэна для зависимых выборок, которая интерпретировалась в свете «классических» пороговых значений (Cohen, 1988). Сырые данные и код анализа доступны по ссылке: <https://osf.io/mn5cf/>.

Результаты

В качестве зависимой переменной во всех видах анализа выступало время реакции для правильных ответов. В ходе предварительного анализа были отсеяны пробы, для которых значение времени реакции либо было меньше 100 мс, либо отклонялось более чем на 2.5 SD от среднего внутри отдельных групп по фактору конгруэнтности. Такой критерий отсева выбросов можно считать оптимальным, так как он обеспечивает баланс между необходимостью исключения экстремальных значений, в особенности характерных для правосторонних хвостов экс-Гауссовских распределений, к которым относится время реакции, и сохранением как можно большего числа валидных наблюдений. Для всех моделей была проведена проверка допущений о нормальности распределения (по Q-Q графикам) и гомоскедастичности остатков (по графику рассеяния остатков и предсказанных значений), и соответствующие допущения не были отклонены.

Результаты анализа для смешанной модели представлены в таблице 4. В неконгруэнтном условии респонденты в среднем отвечали на 16 мс позднее, чем в конгруэнтном условии, что в стандартизированном виде соответствует умеренной величине эффекта ($d_z = 0.53$, 95% CI [0.18, 0.89]). Сходные различия были обнаружены между конгруэнтным и нейтральным условиями ($d_z = -0.53$, 95% CI [-0.89, -0.18]), тогда как различия между неконгруэнтным и нейтральным условиями были незначительны — пересечение доверительных интервалов больше 50% (Cumming, 2009).

² Наше исследование является частью более крупного исследовательского проекта, направленного на изучение роли эмоциональной регуляции в мультисенсорной интеграции вербальной и невербальной информации. В его рамках проводилось несколько экспериментальных воздействий, включающих варьирование условий конгруэнтности и предъявление видеороликов разной эмоциональной валентности для модуляции эмоционального состояния испытуемых (Панкратова, Люсин, 2018), а также измерение выраженности различных стратегий регуляции эмоций и некоторых психофизиологических показателей. В данной работе мы оценивали влияние пространственных аспектов семантики вербальных стимулов на возникновение эффекта кросс-модального соответствия. Все дополнительные переменные, рассмотрение которых не входило в цели настоящей работы, но которые потенциально могли повлиять на ее результаты, были полностью проконтролированы на этапе статистического анализа.

Поскольку выявленный эффект конгруэнтности оказался на границе между слабой и умеренной величиной эффекта, было принято решение проанализировать наличие эффекта кросс-модального соответствия для каждого из стимульных слов по отдельности, так как для разных стимулов эффект кросс-модального соответствия мог проявиться по-разному. Этот эффект считался проявившимся, если точечная оценка величины эффекта для сравнения неконгруэнтного и конгруэнтного условий попадала в диапазон малой величины эффекта. В итоге выраженный эффект соответствия был обнаружен для части слов («счастье», «подавленность» и «презрение»; см. таблицу 5); для остальных стимулов различия между неконгруэнтным и конгруэнтным условиями оказались незначительными.

Таблица 4

Результаты линейной смешанной модели для фактора конгруэнтности

	Предикторы (эффекты)	Оценка параметра	95% CI	<i>t</i>	df	<i>p</i>
Фиксированные эффекты	Свободный член	444.37	—	18.32	34	< 0.001
	Конгруэнтность: неконгруэнтное условие	15.94	6.82–25.06	3.42	70	0.001
	Конгруэнтность: нейтральное условие	13.65	4.53–22.77	2.93	70	0.005
	Видеоролики: позитивная валентность	–65.36	—	–1.92	33	0.064
	Видеоролики: негативная валентность	–109.22	—	–3.20	33	0.003
Случайные эффекты	τ00 (респондент)	6843.10				
	σ ² (остаток)	390.20				

Таблица 5

Эффект конгруэнтности по отдельным словам

Слово	<i>d</i> _z Коэна [95% CI]	Интерпретация
Восхищение	0.15 [–0.18–0.49]	Незначительный эффект
Ликование	0.18 [–0.16–0.51]	Незначительный эффект
Вдохновение	0.01 [–0.34–0.32]	Незначительный эффект
Счастье	0.36 [0.02–0.71]	Малый эффект
Восторг	–0.05 [–0.38–0.28]	Незначительный эффект
Подавленность	0.37 [0.03–0.72]	Малый эффект
Разочарование	0.01 [–0.32–0.34]	Незначительный эффект
Тоска	0.13 [–0.20–0.47]	Незначительный эффект
Презрение	0.27 [–0.06–0.61]	Малый эффект
Огорчение	0.19 [–0.14–0.52]	Незначительный эффект

В качестве дальнейшего эксплораторного анализа было произведено сравнение времени реакции для конгруэнтных и неконгруэнтных условий в зависимости от высоты звука. Для этого в качестве фиксированных эффектов в уравнение смешанной модели были введены взаимодействие факторов конгруэнтности и высоты звука и их главные эффекты (см. таблицу 6). Результаты показали, что при переходе от неконгруэнтных проб к конгруэнтным ускорение ответов наблюдается для высокого звука ($d_z = 0.51$, 95% CI [0.17, 0.87]), но не для низкого ($d_z = 0.14$, 95% CI [-0.19, 0.47]).

Обсуждение

Анализ влияния конгруэнтности разномодальных характеристик стимулов на время реакции позволил выявить предполагаемый эффект кросс-модального соответствия — более легкую идентификацию перцептивных характеристик аудиального стимула при их конгруэнтности пространственной коннотации вербального стимула. Время реакции при предъявлении конгруэнтных стимулов было меньше, чем в случае неконгруэнтных и нейтральных, при этом значимых различий между неконгруэнтным и нейтральным условием обнаружено не было. Вероятно, соответствие характеристик разномодальных стимулов друг другу позволяло респондентам сформировать общую ментальную репрезентацию, основанную на пространственных ассоциациях, актуализированных разными стимулами. Как следствие, это целостное представление способствовало более легкому (а потому и быстрому) выполнению задачи. Полученные нами результаты существенно расширяют представления о закономерностях, ранее обнаруженных другими исследователями при использовании иного стимульного материала (Gozli et al., 2013;

Таблица 6
Результаты линейной смешанной модели для выявления взаимодействия факторов конгруэнтности и высоты звука

	Предикторы (эффекты)	Оценка параметра	<i>t</i>	df	<i>p</i>
Фиксированные эффекты	Свободный член	440.95	18.37	36	< 0.001
	Конгруэнтность: неконгруэнтное условие	26.65	3.21	105	0.002
	Низкий звук	5.38	0.65	105	0.519
	Неконгруэнтное условие × Низкий звук	21.42	1.82	105	0.071
	Видеоролики: позитивная валентность	64.45	1.92	33	0.061
	Видеоролики: негативная валентность	107.63	3.24	33	0.003
Случайные эффекты	τ00 (респондент)	6295.00			
	σ ² (остаток)	1242.00			

Woodin, Winter, 2018). В частности, для реализации нашего исследования были отобраны стимулы двух типов — аудиальные и вербальные, — требующие когнитивной обработки разного уровня сложности (от простой перцептивной идентификации до семантического анализа), но в обоих случаях лишь опосредованно (за счет устройства вербального опыта носителей русского языка) ассоциированные с определенным расположением в пространстве.

Так, выявленное нами уменьшение времени реакции при согласовании характеристик аудиальных и вербальных стимулов, ассоциированных с разными частями пространственного поля, может указывать на возникновение эффекта кросс-модального соответствия, проявляющегося практически моментально при обработке стимулов, не требующих пространственного анализа. Данный результат подтверждает предположение о том, что актуализация ментальной схемы пространства, ориентированного относительно вертикальной оси, является базовым когнитивным механизмом, опосредующим понимание семантики слов, обозначающих эмоциональные состояния (Лакофф, Джонсон, 2004).

Также для более детального рассмотрения полученного эффекта мы сравнили условия, в которых предъявлялись высокие и низкие звуки. Было обнаружено, что ускорение выполнения экспериментальной задачи при конгруэнтном предъявлении стимулов оказывается значительно более выраженным при использовании высокого звукового сигнала в сравнении с низким. Интересно, что схожие результаты обнаруживаются при предъявлении более простых визуальных стимулов в верхней пространственной позиции, что объясняется асимметрией восприятия и внимания (Jóhannesson et al., 2018), так как верхняя часть зрительного поля имеет преимущество перед нижней: длительность фиксаций, совершенных в верхней части поля зрения, ниже, чем длительность фиксаций в нижней; также отмечается более четкое представительство верхнего поля зрения в подкорковых и корковых зрительных центрах (Greene et al., 2019). Схожие результаты были получены в наших предыдущих исследованиях, выполненных с использованием более простого стимульного материала — круга, предъявлявшегося в различных пространственных позициях (Janyan et al., 2022).

Помимо этого, был проанализирован эффект аудиовизуального соответствия для каждого из использованных вербальных стимулов, обозначающих эмоциональные состояния. Этот — более подробный — анализ показал, что эффект соответствия явно обнаруживается лишь для некоторых использованных стимульных слов. Такие результаты позволяют предположить, что, несмотря на достаточно тщательную процедуру подбора стимульного материала, итоговый набор слов оказался не полностью сбалансированным по тем параметрам, которые в принципе не могли быть проконтролированы (длина и частотности словоупотребления). Первым возможным объяснением полученных результатов являются различия в частотности и длине вербальных стимулов. Проведенная оценка частотности использованных нами слов позволила заключить, что ожидаемый эффект соответствия возникал при работе респондентов с наиболее частотным позитивным стимулом («счастье»), в то

время как для негативных стимулов тенденция была обратной — данный эффект проявился в большей степени для наименее частотных слов («презрение» и «подавленность») (см. таблицу 3). Однако важно отметить, что систематическое изучение данного фактора не входило в задачи настоящего исследования, из-за чего выявленные нами закономерности носят описательный характер, а для их более полной содержательной интерпретации необходимо проведение дополнительных экспериментов с использованием расширенного количества стимулов, сбалансированных по частотности употребления. Второе объяснение может заключаться в недостаточно богатом эмоциональном глоссарии респондентов: согласно данным одного из исследований, проведенного на русскоязычной выборке, у 26.2% респондентов в возрасте 10–14 лет и 22.1% респондентов в возрасте 15–18 лет обнаружена алекситимия, проявляющаяся в сложностях осознания и выражения собственных эмоций. При этом треть подростков в каждой из возрастных категорий была отнесена к группе риска по данному параметру (Юткина, 2017). С учетом времени, прошедшего со момента публикации этой работы, можно предполагать, что описанные эмоциональные особенности характерны для участников нашей выборки как представителей того же поколения. Вероятно, сопряженная с высоким уровнем алекситимии бедность эмоционального словаря не позволяет респондентам быстро распознавать использованные в исследовании стимулы и сопоставлять их с собственным телесным опытом, что необходимо для формирования эффекта кросс-модального соответствия при использовании ориентационной метафоры.

В связи с этим в будущих исследованиях следует обратить дополнительное внимание на контроль частотности и длины вербальных стимулов, а также объема словарного запаса участников, связанного с характеристикой эмоциональных состояний. Отдельный интерес может представлять учет данных об особенностях функционирования эмоциональной сферы респондентов.

Отдельным результатом нашей работы стало создание набора вербальных стимулов с независимой оценкой их эмоциональных параметров и пространственных коннотаций. Данные слова могут в будущем послужить основой для разнообразных экспериментальных работ, направленных на выяснение механизмов кросс-модальных взаимодействий в рамках теории пространственного воплощения эмоций.

Литература

- Лакофф, Д., Джонсон, М. (2004). *Метафоры, которыми мы живем*. М.: Едиториал УРСС.
- Панкратова, А. А., Люсин, Д. В. (2018). Видеоролики для индукции эмоций в лабораторных условиях: нормативные данные и кросс-культурный анализ. *Экспериментальная психология*, 11(2), 5–15.
- Юткина, О. С. (2017). Изучение уровня алекситимии у детей школьного возраста. *Современные проблемы науки и образования*, 2, 68. <https://www.elibrary.ru/item.asp?id=29036132>

References

- Auracher, J. (2017). Sound iconicity of abstract concepts: Place of articulation is implicitly associated with abstract concepts of size and social dominance. *PLoS ONE*, 12(11), Article e0187196. <https://doi.org/gchcmq>
- Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, 59, 617–645. <https://doi.org/10.1146/annurev.psych.59.103006.093639>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.186737/jss.v067.i01>
- Bond, B., & Stevens, S. S. (1969). Cross-modality matching of brightness to loudness by 5-year-olds. *Perception & Psychophysics*, 6(6), 337–339. <https://doi.org/10.3758/BF03212787>
- Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW): Instruction manual and affective ratings*. Technical Report C-1, The Center for Research in Psychophysiology, University of Florida.
- Brunetti, R., Indraco, A., Del Gatto, C., Spence, C., & Santangelo, V. (2018). Are crossmodal correspondences relative or absolute? Sequential effects on speeded classification. *Attention, Perception, & Psychophysics*, 80(2), 527–534. <https://doi.org/10.3758/s13414-017-1445-z>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.
- Cumming, G. (2009). Inference by eye: Reading the overlap of independent confidence intervals. *Statistics in Medicine*, 28(2), 205–220. <https://doi.org/10.1002/sim.3471>
- Dolscheid, S., Çelik, S., Erkan, H., Küntay, A., & Majid, A. (2020). Space-pitch associations differ in their susceptibility to language. *Cognition*, 196, Article 104073. <https://doi.org/10.1016/j.cognition.2019.104073>
- Evans, K. K. (2020). The role of selective attention in cross-modal interactions between auditory and visual features. *Cognition*, 196, Article 104119. <https://doi.org/10.1016/j.cognition.2019.104119>
- Evans, K. K., & Treisman, A. (2010). Natural cross-modal mappings between visual and auditory features. *Journal of Vision*, 10(1), Article 6. <https://doi.org/10.1167/10.1.6>
- Fernandez-Prieto, I., Spence, C., Pons, F., & Navarra, J. (2017). Does language influence the vertical representation of auditory pitch and loudness? *i-Perception*, 8(3). <https://doi.org/10.1177/2041669517716183>
- Gozli, D. G., Chow, A., Chasteen, A. L., & Pratt, J. (2013). Valence and vertical space: Saccade trajectory deviations reveal metaphorical spatial activation. *Visual Cognition*, 21(5), 628–646. <https://doi.org/10.1080/13506285.2013.815680>
- Greene, H. H., Brown, J. M., & Strauss, G. P. (2019). Shorter fixation durations for up-directed saccades during saccadic exploration: A meta-analysis. *Journal of Eye Movement Research*, 12(8), Article 5. <https://doi.org/10.16910/jemr.12.8.5>
- Günther, F., Nguyen, T., Chen, L., Dudschig, C., Kaup, B., & Glenberg, A. M. (2020). Immediate sensorimotor grounding of novel concepts learned from language alone. *Journal of Memory and Language*, 115, Article 104172. <https://doi.org/10.1016/j.jml.2020.104172>
- Janyan, A., Shtyrov, Y., Andriushchenko, E., Blinova, E., & Shcherbakova, O. (2022). Look and ye shall hear: Selective auditory attention modulates the audiovisual correspondence effect. *i-Perception*, 13(3), 1–10. <https://doi.org/10.1177/20416695221095884>
- Jóhannesson, Ó. I., Tagu, J., & Kristjánsson, Á. (2018). Asymmetries of the visual system and their influence on visual performance and oculomotor dynamics. *European Journal of Neuroscience*, 48(11), 3426–3445. <https://doi.org/10.1111/ejn.14225>

- Jonas, C., Spiller, M. J., & Hibbard, P. (2017). Summation of visual attributes in auditory–visual cross-modal correspondences. *Psychonomic Bulletin & Review*, 24(4), 1104–1112. <https://doi.org/10.3758/s13423-016-1215-2>
- Lakoff, G., & Johnson, M. (1980). The metaphorical structure of the human conceptual system. *Cognitive Science*, 4(2), 195–208. <https://doi.org/cxpmcn>
- Lakoff, G., & Johnson, M. (2004). *Metafori, kotorymi my zhivem* [Metaphors we live by]. Moscow: Editorial URSS. (Original work published 1980)
- Martino, G., & Marks, L. E. (1999). Perceptual and linguistic interactions in speeded classification: Tests of the semantic coding hypothesis. *Perception*, 28(7), 903–923. <https://doi.org/10.1068/p2866>
- Pankratova A. A., Lyusin D. V. (2018). Videoroliki dlya izucheniya emotsiy v laboratornykh usloviyakh: normativnyye dannyye i kross-kul'turnyy analiz [Videos for the study of emotion in the lab: Normative evidence and cross-cultural analysis]. *Experimental Psychology*, 11(2), 5–15.
- Parise, C. V., & Spence, C. (2009). ‘When birds of a feather flock together’: synesthetic correspondences modulate audiovisual integration in non-synesthetes. *PLoS ONE*, 4(5), Article e5664. <https://doi.org/10.1371/journal.pone.0005664>
- Puigcerver, L., Rodríguez-Cuadrado, S., Gómez-Tapia, V., & Navarra, J. (2019). Vertical mapping of auditory loudness: Loud is high, but quiet is not always low. *Psicológica Journal*, 40(2), 85–104. <https://doi.org/10.2478/psicolj-2019-0006>
- Smith, L. B., & Sera, M. D. (1992). A developmental analysis of the polar structure of dimensions. *Cognitive Psychology*, 24(1), 99–142. [https://doi.org/10.1016/0010-0285\(92\)90004-L](https://doi.org/10.1016/0010-0285(92)90004-L)
- Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, 73(4), 971–995. <https://doi.org/10.3758/s13414-010-0073-7>
- Spence, C. (2020). Simple and complex crossmodal correspondences involving audition. *Acoustical Science and Technology*, 41(1), 6–12. <https://doi.org/10.1250/ast.41.6>
- Störmer, V. S. (2019). Orienting spatial attention to sounds enhances visual processing. *Current Opinion in Psychology*, 29, 193–198. <https://doi.org/10.1016/j.copsyc.2019.03.010>
- Woodin, G., & Winter, B. (2018). Placing abstract concepts in space: quantity, time and emotional valence. *Frontiers in Psychology*, 9, Article 2169. <https://doi.org/10.3389/fpsyg.2018.02169>
- Yutkina, O. S. (2017). Izuchenie urovnya aleksitimii u detei shkol'nogo vozrasta [The inspection of standard of alexithymia among schoolchildren]. *Modern Problems of Science and Education*, 2, 68. <http://shorturl.at/dmrBT>