



Методы сетевого моделирования структуры семантической памяти у изучающих иностранный язык

А. В. Бармин¹, Б. Б. Величковский²

^{1,2}Московский государственный лингвистический университет, Москва, Россия

²Московский государственный университет имени М. В. Ломоносова, Москва, Россия

¹art.barmin@mail.ru, ²velitchk@mail.ru

Аннотация. Цель настоящей статьи – определение методов сетевого моделирования структуры семантической памяти у изучающих иностранный язык. Авторами статьи был проведен теоретический анализ отечественных и зарубежных литературных источников, посвященных данной проблеме. Результаты теоретического обзора показывают, какие методы сетевого моделирования существуют сегодня и какие из них могут эффективно использоваться в целях исследования структуры семантической памяти у изучающих иностранный язык.

Ключевые слова: семантическая память, изучение иностранного языка, сетевое моделирование, семантическая сеть, теория графов

Благодарности: выражаем благодарность студенту 3-го курса бакалавриата Пиастро Роману Алексеевичу за помощь в поиске и систематизации литературы, а также в подготовке рукописи к отправлению в редакцию.

Для цитирования: Бармин А. В., Величковский Б. Б. Методы сетевого моделирования структуры семантической памяти у изучающих иностранный язык // Вестник Московского государственного лингвистического университета. Образование и педагогические науки. 2024. Вып. 1 (850). С. 105–110.

Review article

Methods for Network Modeling the Structure of Semantic Memory of Foreign Language Learners

Artem V. Barmin¹, Boris B. Velichkovsky²

^{1,2}Moscow State Linguistic University, Moscow, Russia

²Lomonosov Moscow State University, Moscow, Russia

¹art.barmin@mail.ru, ²velitchk@mail.ru

Abstract. This article is devoted to defining methods for network modeling of the structure of semantic memory of foreign language learners. The authors of the article conducted a theoretical analysis of domestic and foreign literary sources devoted to the problem under consideration. The results of the theoretical review show what network modeling methods exist today and which of them can be effectively used to study the structure of semantic memory of foreign language learners.

Keywords: semantic memory, learning a foreign language, network modeling, semantic network, graph theory

Acknowledgments: we would like to express our gratitude to Piastro Roman Alekseevich for his assistance in searching and systematizing literature, as well as in preparing the manuscript for sending to the editor.

For citation: Barmin, A. B., Velichkovsky, B. B. (2024). Methods for network modeling the structure of semantic memory of foreign language learners. Vestnik of Moscow State Linguistic University. Education and Teaching, 1(850), 105–110.

ВВЕДЕНИЕ

Семантическая память (СП) представляет собой отдел долговременной памяти, отвечающий за хранение общей информации о мире (например, Земля – планета Солнечной системы) [Tulving, 1972]. В настоящее время приобретают актуальность исследования СП у изучающих иностранный язык. Во-первых, это связано с разработкой новых интеллектуальных систем, моделирующих СП человека для оптимизации изучения иностранного языка [Runge, Hovy, 2020; A platform for image recommendation in foreign word learning, 2019]. Во-вторых, в настоящее время до сих пор сохраняется потребность в поиске эффективных средств оценки качества усвоения иностранного языка [Дубинина, Каменская, 2018]. Предполагается, что разработка этих средств может быть основана на моделировании и анализе структуры СП человека.

В современной психологии и когнитивной науке существует несколько подходов к моделированию структуры СП. Среди этих подходов наиболее популярными считаются сетевой и дистрибутивный. В сетевой парадигме, структура СП представляют собой множество узлов, объединенных ребрами, тогда как в дистрибутивной парадигме, она может быть представлена в виде системы статистических показателей распределения слов в языковых корпусах. Хотя дистрибутивные модели весьма эффективны в объяснении возможных механизмов, лежащих в основе работы СП человека, и сегодня находятся на пике достижений в развитии искусственного интеллекта (например, ChatGPT от OpenAI), они могут упускать ряд психологических факторов (например, аффективных, перцептивных, моторных), которые могут быть доступны в поведенческих экспериментах. В этом смысле, дистрибутивные модели сильно уступают сетевым моделям, основанным на ответах реальных испытуемых в экспериментальных условиях. Действительно, результаты сравнения двух типов моделей показывают, что дистрибутивные модели эффективнее сетевых в решении таких задач, как, например, установление сходства объектов [De Deyne et al., 2016] или в словесных играх [Kumar, Steyvers, Balota, 2021]. Так, существуют веские основания полагать, что сетевой подход в определенном смысле превосходит дистрибутивный в плане своей «психологичности». В связи с этим далее будут рассмотрены известные способы сетевого моделирования структуры СП и определены возможности их применения в исследованиях структуры СП у изучающих иностранный язык.

Целью настоящей статьи является определение методов сетевого моделирования структуры СП у изучающих иностранный язык.

Задачи исследования:

- рассмотреть известные на сегодняшний день методы сетевого моделирования структуры СП;
- определить возможности применения этих методов в исследованиях структуры СП у изучающих иностранный язык.

МЕТОДЫ СЕТЕВОГО МОДЕЛИРОВАНИЯ СТРУКТУРЫ СП

Сетевое моделирование представляет собой метод построения и анализа сетевых моделей, отражающих организацию тех или иных изучаемых явлений в виде множества узлов и множества ребер, объединяющих некоторые пары этих узлов [Морозова, 2017]. В настоящее время в психологии и когнитивной науке существует несколько методов построения сетевой структуры СП, не считая метода построения сетей на основе норм свободных ассоциаций и тезаурусов. Данный метод был продемонстрирован в работе M. Steyvers и J. Tenenbaum [Steyvers, Tenenbaum, 2005]. Так, одним из известных методов является задача вербальной беглости (verbal fluency task), в которой испытуемым предлагается назвать как можно больше объектов из заданной категории (например, «животные») за фиксированный промежуток времени. Результатом выполнения данной задачи является суммарное количество верно названных объектов, принадлежащих к той или иной категории, которое может сравниваться со стандартными показателями выполнения задачи. Однако наиболее важно то, что результаты выполнения данной задачи могут использоваться для построения сетевых моделей структуры СП. Например, G. Zhang (и коллеги) показали, что графы, построенные на основе результатов прохождения задачи вербальной беглости здоровыми испытуемыми, представляют собой «цепи» (тип графа, в которых для всех ребер, кроме первого и последнего, одна из вершин является общей с предыдущим ребром, а вторая с последующим), тогда как сети пациентов, страдающих болезнью Альцгеймера и Паркинсона, не являются «цепями» и имеют ряд отличительных признаков, например, «петли» (ребра, соединяющие узел N сети с самими собой) [Zhang et al., 2022]. Однако, главная заслуга G. Zhang (и коллег) заключается в том, что им удалось и выявить различия между сетями трех групп по макроуровневым сетевым характеристикам, таким как «средний кратчайший путь», «плотность» и «диаметр».

Другим методом является задача (парадигма) «снежного кома» (snowball sampling paradigm)

[Morais, Olsson, Schooler, 2013]. Согласно этой задаче испытуемым предлагается придумывать ассоциации к заранее отобранным словам (напр., взятым из психолингвистических баз данных), называемых «семенами» (seeds) в течение фиксированного промежутка времени. При этом задача испытуемого заключается не в придумывании как можно большего количества ассоциаций, а в отражении самого подходящего, что придет им в голову в ответ на поочередно предъявляющиеся стимулы. По завершении первой итерации, испытуемым предлагается придумать ассоциации к собственным словам-ассоциациям, данным к словам – «семенам». Так, данная задача подразумевает бесчисленное количество итераций, поэтому семантическая сеть растет словно снежный ком. Следует отметить, что данная задача является весьма трудоемкой – в оригинальном исследовании использовалось всего шесть итераций, но и их выполнение длилось около шести недель почасовых сессий. По этой причине D. Wulff и R. Mata разработали сокращенную версию задачи «снежного кома» (mini-snowball), состоящую всего из двух итераций [Wulff, Mata, 2022]. Сокращенная версия задачи «снежного кома» была успешно использована учеными в их исследовании семантической репрезентации понятия риска.

Последним методом, к которому следует обратиться в настоящей работе, является задача семантической связанности (semantic relatedness task), разработанная Y. Kenett (и коллегами) [Kenett et al., 2017]. Согласно данной задаче испытуемым предлагается оценить связанность заданных пар слов по шкале с полюсами «не связаны» и «сильно связаны». На основании результатов прохождения данной задачи выстраивается «матрица смежности» графа (adjacency matrix) (матрица, в которой столбцы и строки являются узлами, а показатели на их пересечении принимают одно из двух значений: 0 – узлы не связаны или 1 – узлы связаны). Например, в исследовании связи сетевой структуры СП и креативности M. Benedek (и коллеги) использовали 28 слов, разбив их на 378 уникальных пар для предъявления в задаче семантической связанности испытуемым двух групп, различающихся по уровню креативности [Benedek, et al., 2017]. По результатам оценок связанности пар слов выстраивались матрицы смежности 28x28, которые были использованы учеными для моделирования сетевой структуры СП испытуемых двух групп с целью выявления различий между ними по сетевым характеристикам.

ВОЗМОЖНОСТИ ПРИМЕНЕНИЯ МЕТОДОВ СЕТЕВОГО МОДЕЛИРОВАНИЯ В ИССЛЕДОВАНИИ СЕМАНТИЧЕСКОЙ ПАМЯТИ У ИЗУЧАЮЩИХ ИНОСТРАННЫЙ ЯЗЫК

Одной из основных методологических особенностей исследований структуры СП с использованием метода сетевого моделирования является сравнение семантических сетей испытуемых разных групп (например, здоровых и больных, молодых и пожилых, креативных и некреативных и т. д.). В связи с этим в качестве выборки для исследований структуры СП у изучающих иностранный язык могут выступать испытуемые с разным уровнем владения иностранным языком. В таких исследованиях могут использоваться как связанные, так и несвязанные выборки, т. е. речь идет о лонгитюдных исследованиях и исследованиях с использованием метода поперечных срезов. Кроме того, необходимым условием в таких исследованиях является использование в качестве стимулов иноязычных лексических единиц.

Так, в задаче вербальной беглости испытуемым двух групп, различающихся по уровню владения иностранным языком, может быть предложено перечислить на иностранном языке как можно больше объектов из определенной категории за фиксированный промежуток времени. Поскольку предполагается, что в подобных исследованиях будут участвовать испытуемые без нейродегенеративных заболеваний и психических расстройств, представляется целесообразным построение и анализ агрегированных, а не индивидуальных сетей, как это было сделано в исследовании M. P. Agustin-Llach [Agustin-Llach, 2022]. Действительно, анализ индивидуальных сетей представляется нецелесообразным по причине возможной однородности их структурных особенностей (все сети могут представлять собой «цепи»).

Согласно задаче снежного кома испытуемым с разным уровнем владения иностранным языком может быть предложено придумывать ассоциации к исходным словам на изучаемом иностранном языке. При этом в качестве «семян» могут использоваться слова, отобранные из специальных психолингвистических баз данных, позволяющих отфильтровать их по возрасту усвоения. Например, среди англоязычных баз данных можно выделить MRC и нормы усвоения слов Купермана [Coltheart, 1981; Kuperman, Stadthagen-Gonzalez, Brysbaert, 2012]. Например, электронный словарь Cambridge Dictionary¹ обеспечивает такую возможность бла-

¹URL: <https://dictionary.cambridge.org/>

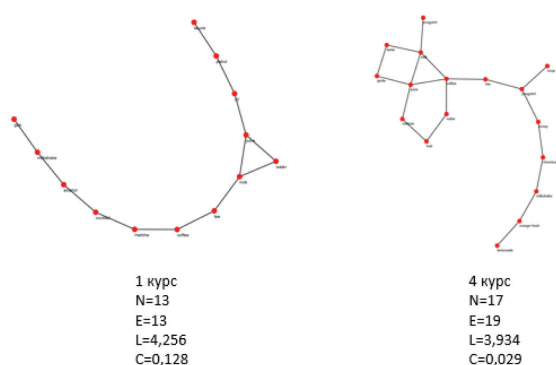


Рис. 1. Агрегированные семантические сети двух студентов 1-го курса МГЛУ и двух студентов 4-го курса МГЛУ, обучающихся по направлению «Языкознание и литературоведение», построенная при помощи метода задачи вербальной беглости; N – количество узлов, E – количество ребер, L – показатель средней длины кратчайшего пути, C – уровень кластеризации

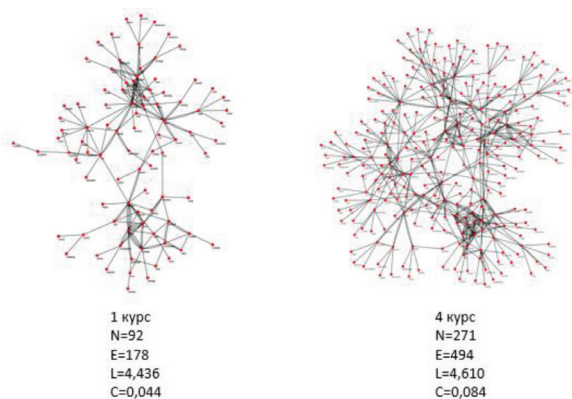


Рис. 2. Семантические сети студента 1-го курса МГЛУ и студента 4-го курса МГЛУ, обучающихся по направлению «Языкознание и литературоведение», построенная при помощи метода задачи снежного кома; N – количество узлов, E – количество ребер, L – показатель средней длины кратчайшего пути, C – уровень кластеризации

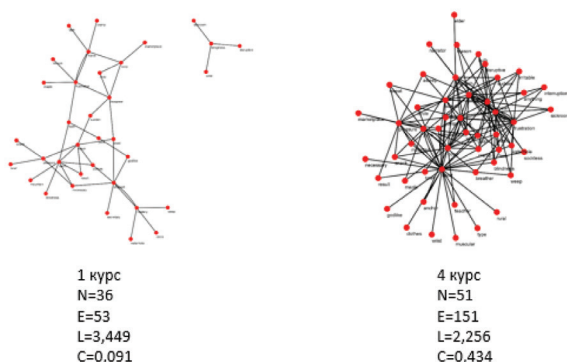


Рис. 3. Семантические сети студента 1-го курса МГЛУ и студента 4-го курса МГЛУ, обучающихся по направлению «Языкознание и литературоведение», построенная при помощи метода задачи семантических суждений; N – количество узлов, E – количество ребер, L – показатель средней длины кратчайшего пути, C – уровень кластеризации

годаря специальной маркировке слов, позволяющей определять соответствие слов тому или иному уровню владения иностранным языком по системе CERF.

Наконец, в задаче семантических суждений разным группам испытуемых может быть предложено определить семантическое сходство некоторых пар иностранных слов, отобранных либо из психолингвистических баз данных либо из баз свободных ассоциаций [Nelson, McEvoy, Schreiber, 2004]. Они могут быть отфильтрованы по силе связанности. Существенное преимущество данного метода заключается в том, что помимо оценок семантических суждений, он также подразумевает оценку времени реакции. Так, например, чем ближе два слова находятся в сети СП, тем меньше будет время реакции при определении их семантического сходства [Collins, Loftus, 1975].

В завершение описания возможностей использования методов сетевого моделирования в изучении структуры СП у изучающих иностранный язык, нами были приведены образцы построения и анализа сетевых моделей, основанных на выполнении трех вышеупомянутых задач: задачи вербальной беглости (см. рис. 1), задаче снежного кома (см. рис. 2) и задаче семантических суждений (см. рис. 3). Визуализация и анализ семантических сетей были реализованы в Cytoscape 3.10.1. В задаче вербальной беглости испытуемым предлагалось назвать на английском языке как можно больше объектов из категории «fluids» за одну минуту. Согласно задаче снежного кома в качестве «семян» были отобраны случайные слова из базы данных Купермана [Kuperman, Stadthagen-Gonzalez, Brysbaert, 2012], отфильтрованные по возрасту усвоения до девяти лет. Испытуемым предлагалось придумывать ассоциации на английском языке к каждому отобранному слову, а затем к собственным словам на протяжении одной минуты для каждого предъявленного стимула. В задаче вербальной беглости использовалось 56 случайно отобранных английских слов из базы данных Купермана [Kuperman, Stadthagen-Gonzalez, Brysbaert, 2012], отфильтрованные по возрасту усвоения до девяти лет. Из 56 слов были составлены случайные пары, которые предъявлялись испытуемым. Испытуемые должны были оценить степень семантической связанности пар слов по шкале от одного – слова не связаны до пяти – слова сильно связаны. По результатам выполнения данной задачи выстраивались матрицы смежности 56x56, на основе которых строились индивидуальные семантические сети. Полученные сети были отфильтрованы по оценкам связанности пар слов от трех до пяти (пары слов, связанность которых была оценена до двух были исключены из анализа).

Полученные результаты не являются репрезентативными в полной мере, поскольку они были получены на крайне небольшой выборке испытуемых, поэтому следует рассматривать их только как образцы выполнения рассматриваемых задач.

ЗАКЛЮЧЕНИЕ

В настоящем теоретическом обзоре были представлены существующие методы сетевого моделирования структуры СП и определены

возможности их применения в исследовании структуры СП у изучающих иностранный язык. Результаты обзора показывают, какие методы сетевого моделирования существуют сегодня и какие из них могут использоваться в целях исследования структуры семантической памяти у изучающих иностранный язык. Результаты проведенного исследования могут быть использованы в разработке систем автоматизации изучения иностранного языка, в создании эффективных средств оценивания уровня владения иностранным языком.

СПИСОК ИСТОЧНИКОВ

1. Tulving E. Episodic and semantic memory // Organization of memory / Eds. E. Tulving, W. Donaldson. New York: Academic Press, 1972. P. 381–402.
2. Runge A., Hovy E. Exploring neural entity representations for semantic information // Proceedings of the Third BlackboxNLP Workshop on Analyzing and Interpreting Neural Networks for NLP. 2020. P. 204–216. DOI: <https://doi.org/10.48550/arXiv.2011.08951>
3. A platform for image recommendation in foreign word learning / M. N. Hasnine et al. // Companion Proceeding of the 9th International Conference on Learning Analytics & Knowledge (LAK'19). March 4–8, 2019, Tempe, Arizona, USA. 2019. P. 187–188.
4. Дубинина Г. А., Каменская Н. В. Поиск новых форм контроля и оценки качества подготовки по иностранному языку в неязыковом вузе // Вестник Московского государственного лингвистического университета. Образование и педагогические науки. 2018. Вып. 2 (796). С. 202–211.
5. De Deyne S. et al. Structure at every scale: A semantic network account of the similarities between unrelated concepts // Journal of Experimental Psychology: General. 2016. Vol. 145 (9). P. 1228–1254.
6. Kumar A. A., Steyvers M., Balota D. A. Semantic memory search and retrieval in a novel cooperative word game: A comparison of associative and distributional semantic models // Cognitive Science. 2021. Vol. 45 (10). P. 1–33.
7. Морозова О. Структурное сетевое моделирование в когнитивной науке // Психологические исследования. 2017. Вып. 10 (55). С. 1–13.
8. Steyvers M., Tenenbaum J. B. The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth // Cognitive science. 2005. Vol. 29 (1). P. 41–78.
9. Zhang G. et al. Graph theoretical analysis of semantic fluency in patients with Parkinson's disease // Behavioural Neurology. 2022. P. 1–7. DOI: 10.1155/2022/6935263.
10. Morais A. S., Olsson H., Schooler L. J. Mapping the structure of semantic memory // Cognitive science. 2013. Vol. 37 (1). P. 125–145.
11. Wulff D. U., Mata R. On the semantic representation of risk // Science advances. 2022. Vol. 8 (27). P. 1–13.
12. Kenett Y. N. et al. The semantic distance task: Quantifying semantic distance with semantic network path length // Journal of Experimental Psychology: Learning, Memory, and Cognition. 2017. Vol. 43 (9). P. 1–21.
13. Benedek M. et al. How semantic memory structure and intelligence contribute to creative thought: A network science approach // Thinking & Reasoning. 2017. Vol. 23 (2). P. 158–183.
14. Agustin-Llach M. P. How age and L2 proficiency affect the L2 lexicon // System. 2022. Vol. 104. DOI: <https://doi.org/10.1016/j.system.2021.102697>.
15. Coltheart M. The MRC psycholinguistic database // The Quarterly Journal of Experimental Psychology. 1981. Vol. 33 (4). P. 497–505.
16. Kuperman V., Stadthagen-Gonzalez H., Brysbaert M. Age-of-acquisition ratings for 30,000 English words // Behavior research methods. 2012. Vol. 44. P. 978–990.
17. Nelson D. L., McEvoy C. L., Schreiber T. A. The University of South Florida free association, rhyme, and word fragment norms // Behavior Research Methods, Instruments, & Computers. 2004. Vol. 36 (3). P. 402–407.
18. Collins A. M., Loftus E. F. A spreading-activation theory of semantic processing // Psychological Review. 1975. Vol. 82 (6). P. 407–428.

REFERENCES

1. Tulving, E. (1972). Episodic and semantic memory. In Tulving, E., Donaldson, W. (Eds.), Organization of memory (pp. 381–402). New York: Academic Press.
2. Runge, A., Hovy, E. (2020). Exploring neural entity representations for semantic information. In Proceedings of the Third BlackboxNLP Workshop on Analyzing and Interpreting Neural Networks for NLP (pp. 204–216). <https://doi.org/10.48550/arXiv.2011.08951>.

3. Hasnine, M. N. et al. (2019). A platform for image recommendation in foreign word learning. In Companion Proceeding of the 9th International Conference on Learning Analytics & Knowledge (LAK'19) (pp. 187–188). Tempe, 2019.
4. Dubinina, G. A., Kamenskaya, N. V. (2018). Looking for new forms of assessment in foreign language training at non-linguistics universities. Vestnik of Moscow State Linguistic University. Education and Teaching, 2(796), 202–211. (In Russ.)
5. De Deyne, S. et al. (2016). Structure at every scale: A semantic network account of the similarities between unrelated concepts. Journal of Experimental Psychology: General, 145(9), 1228–1254.
6. Kumar, A. A., Steyvers, M., Balota, D. A. (2021). Semantic memory search and retrieval in a novel cooperative word game: A comparison of associative and distributional semantic models. Cognitive Science, 45(10), 1–33.
7. Morozova, O. (2017). Structural network modelling in cognitive science. Psychological studies, 10(55), 1–13. (In Russ.)
8. Steyvers, M., Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. Cognitive science, 29(1), 41–78.
9. Zhang, G. et al. (2022). Graph theoretical analysis of semantic fluency in patients with Parkinson's disease. Behavioural Neurology, 1–7. 10.1155/2022/6935263.
10. Morais, A. S., Olsson, H., Schooler, L. J. (2013). Mapping the structure of semantic memory. Cognitive science, 37(1), 125–145.
11. Wulff, D. U., Mata, R. (2022). On the semantic representation of risk. Science advances, 8(27), 1–13.
12. Kenett, Y. N. et al. (2017). The semantic distance task: Quantifying semantic distance with semantic network path length. Journal of Experimental Psychology: Learning, Memory, and Cognition, 43(9), 1–21.
13. Benedek, M. et al. (2017). How semantic memory structure and intelligence contribute to creative thought: A network science approach. Thinking & Reasoning, 23(2), 158–183.
14. Agustin-Llach, M. P. (2022). How age and L2 proficiency affect the L2 lexicon. System, 104. <https://doi.org/10.1016/j.system.2021.102697>.
15. Coltheart, M. (1981). The MRC psycholinguistic database. The Quarterly Journal of Experimental Psychology, 33(4), 497–505.
16. Kuperman, V., Stadthagen-Gonzalez, H., Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. Behavior research methods, 44, 978–990.
17. Nelson, D. L., McEvoy, C. L., Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. Behavior Research Methods, Instruments, & Computers, 36(3), 402–407.
18. Collins, A. M., Loftus, E. F. (1975). A spreading-activation theory of semantic processing. Psychological Review, 82(6), 407–428.

ИНФОРМАЦИЯ ОБ АВТОРАХ

Бармин Артем Вячеславович

аспирант кафедры психологии и педагогической антропологии
Института гуманитарных и прикладных наук Московского государственного лингвистического университета
младший научный сотрудник лаборатории когнитивных исследований основ коммуникации
Московского государственного лингвистического университета

Величковский Борис Борисович

доктор психологических наук, профессор кафедры методологии психологии факультета психологии МГУ
имени М. В. Ломоносова, директор Лаборатории когнитивных исследований основ коммуникации
Московского государственного лингвистического университета

INFORMATION ABOUT THE AUTHORS

Barmin Artem Vyacheslavovich

Post-graduate Student of the Department of Psychology and Pedagogical Anthropology
of the Institute of Humanities and Applied Sciences of Moscow State Linguistic University
Junior Researcher of the Laboratory for Cognitive Studies of Communication, Moscow State Linguistic University

Velichkovsky Boris Borisovich

Doctor of Psychology, Professor at the Department of Methodology of Psychology
Faculty of Psychology, Lomonosov Moscow State University,
Head of the Laboratory for Cognitive Studies of Communication, Moscow State Linguistic University

Статья поступила в редакцию
одобрена после рецензирования
принята к публикации

03.11.2023
21.11.2023
16.01.2024

The article was submitted
approved after reviewing
accepted for publication