

УДК 81'322

DOI: 10.18413/2313-8912-2024-10-1-0-5

Литвинова Т. А.¹ 
Паничева П. В.² 

**Индивидуальные различия в ассоциативном значении
слова сквозь призму языковой модели и семантического
дифференциала**

¹ Воронежский государственный педагогический университет
ул. Ленина, 86, Воронеж, 394043, Россия
E-mail: centr_rus_yaz@mail.ru

² Воронежский государственный педагогический университет
ул. Ленина, 86, Воронеж, 394043, Россия
E-mail: ppolin86@gmail.com

*Статья поступила 01 февраля 2024 г.; принята 14 марта 2024 г.;
опубликована 30 марта 2024 г.*

Информация об источниках финансирования или грантах. Исследование выполнено в Воронежском государственном педагогическом университете при поддержке гранта РФФИ № 21-78-10148 «Моделирование значения слова в индивидуальном языковом сознании на основе дистрибутивной семантики».

Аннотация. Существование индивидуальных различий в семантике слова признается многими исследователями. Однако установление и описание подобных различий представляет собой сложную научную задачу, связанную с необходимостью проведения трудоемкой семантической разметки и неизбежным субъективизмом исследователя. В настоящей работе нами предлагается метод выявления различий в индивидуальной семантике слов на основе автоматически рассчитанных оценок ассоциативного значения слова по шкалам семантического дифференциала. Используя дистрибутивную семантическую модель word2vec, обученную на многомиллионном корпусе текстов, и метод Concept Mover's Distance, мы получили для каждого ассоциативного ряда оценки по 18 шкалам семантического дифференциала. В нашей работе впервые данный метод, активно используемый в новейших работах, рассматривающих текст как данные (преимущественно выполненных в русле computational social science), применяется по отношению к такому объекту анализа, как ассоциативный ряд, с целью описания индивидуальных различий в семантике слов. В качестве материала для исследования мы использовали специально созданный датасет, содержащий ассоциативные реакции к словам-стимулам, важным для русского языкового сознания, данные о психологических характеристиках респондентов (черты «Большой пятерки») и их эмоциональном состоянии в момент прохождения тестирования. Применяя комплекс методов анализа многомерных данных (метод главных компонент, факторный анализ, иерархическая кластеризация на главных компонентах), мы разделили слова-стимулы на группы в зависимости от выраженности индивидуальных различий в их семантике. Нами также была установлена связь эмоционально-психологических характеристик респондентов и автоматически рассчитанных оценок ассоциативного значения слов-стимулов по шкалам семантического дифференциала. Описанная методика анализа может применяться для получения оценок ассоциа-

тивных рядов (а также контекстов употребления слов в текстах) по любым семантическим оппозициям и предлагается как дополнение к традиционным методам выявления психологически реального значения слова. Используемый в работе датасет и код для воспроизведения полученных результатов на языке R доступны для исследователей.

Ключевые слова: Значение слова; Семантический дифференциал; Дистрибутивная семантика; Дистрибутивные семантические модели; Ассоциативный эксперимент

Информация для цитирования: Литвинова Т. А., Паничева П. В. Индивидуальные различия в ассоциативном значении слова сквозь призму языковой модели и семантического дифференциала // Научный результат. Вопросы теоретической и прикладной лингвистики. 2024. Т. 10. № 1. С. 61-93. DOI: 10.18413/2313-8912-2024-10-1-0-5

**Индивидуальные различия в ассоциативном значении слова
сквозь призму языковой модели и семантического дифференциала**

Почему нам так сложно понимать друг друга?	Решение	Результаты
<p>Различия в индивидуальной семантике слова</p> 	<p>Датасет ассоциаций и данных о респондентах + получение оценок ассоциативного ряда по бинарным шкалам (семантический дифференциал) с использованием языковой модели + методы многомерного анализа (PCA, HCPC)</p> 	<p>Выделены кластеры значений высокочастотных слов русского языка</p> <p>Определено влияние психологических характеристик и эмоциональных состояний респондентов на особенности коннотативного значения слов</p> <p>Самое «однозначное» слово - <i>обман</i>, <i>русский</i> и <i>мама</i> расположены близко в многомерном пространстве, а <i>думать</i> имеет противоположные значения для разных групп респондентов</p> <p>Код и датасет доступны на гитхаб</p>
<p>Как выявлять и описывать такие различия?</p> 	<p>Исследование выполнено при поддержке Российского научного фонда</p> 	<p>СЕТЕВОЙ НАУЧНЫЙ ЖУРНАЛ НАУЧНЫЙ РЕЗУЛЬТАТ RESEARCH RESULT</p> 

UDC 81'322

DOI: 10.18413/2313-8912-2024-10-1-0-5

Tatiana A. Litvinova¹ 
Polina V. Panicheva² 

**Individual differences in the associative meaning of a word
through the lens of the language model and semantic
differential**

¹ Voronezh State Pedagogical University
86 Lenin St., Voronezh, 394043, Russia
E-mail: centr_rus_yaz@mail.ru

² Voronezh State Pedagogical University
86 Lenin St., Voronezh, 394043, Russia
E-mail: ppolin86@gmail.com

Received 01 February 2024; accepted 14 March 2024; published 30 March 2024

Acknowledgements. The research was carried out at Voronezh State Pedagogical University with the support of the Russian Science Foundation, Grant No. 21-78-10148 “Modeling the Meaning of a Word in Individual Linguistic Consciousness Based on Distributive Semantics”.

Abstract. The fact that there are individual differences in word semantics is recognized by lots of scholars. However, establishing and describing such differences is a complex scientific task involving labor-intensive semantic marking and the researcher’s inevitable impartiality. In this study, we set forth a method for identifying differences in the individual semantics of words based on automatically calculated estimates of the associative meaning of a word on semantic differential scales. Using the word2vec distributive semantic model, trained on a multimillion corpus of texts as well as the Concept Mover's Distance method, for each associative series we obtained (associative series consisting of the same number of elements – 5 words – had been selected) scores on 18 semantic differential scales. In our study, for the first time, this method, which has been widely used in the latest studies that consider text as data (mostly performed in the mainstream computational social science), has been applied to such an analysis object as an associative series for describing individual differences in the semantics of words. As the material for the study, we employed a specially designed dataset containing associative reactions to stimuli – high-frequency words in the Russian language, data on the respondents' psychological characteristics (Big Five traits) and their emotional state at the time of testing. Using a set of methods for analyzing multidimensional data (principal component method, factor analysis, hierarchical clustering on principal components), we divided the stimulus words into groups depending on the degree of individual differences in their semantics. We also established a connection between the respondents' psychological characteristics and automatically calculated the estimates of the associative meaning of the stimulus words on semantic differential scales. The described analysis technique can be used in order to obtain the estimates of associative series (as well as contexts of word use in texts) for any semantic oppositions and is set forth as a supplement to the traditional methods of identifying the psychologically real meaning of a word. The dataset used in the study and the code for reproducing the results obtained in the R language are available to a wider research community.

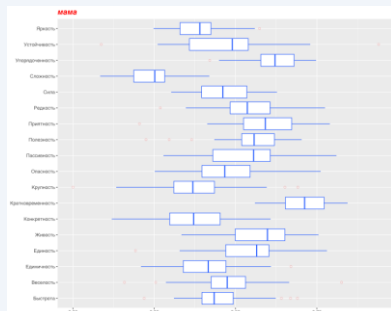
Keywords: Word meaning; Semantic differential; Distributive semantics; Distributive semantic models; Associative experiment

How to cite: Litvinova, T. A. and Panicheva, P. V. (2024). Individual differences in the associative meaning of a word through the lens of the language model and semantic differential, *Research Result. Theoretical and Applied Linguistics*, 10 (1), 61-93. DOI: 10.18413/2313-8912-2024-10-1-0-5

INDIVIDUAL DIFFERENCES IN THE ASSOCIATIVE MEANING OF A WORD THROUGH THE LENS OF THE LANGUAGE MODEL AND SEMANTIC DIFFERENTIAL

Why is it so hard to understand each other?

Differences in individual meaning of a word

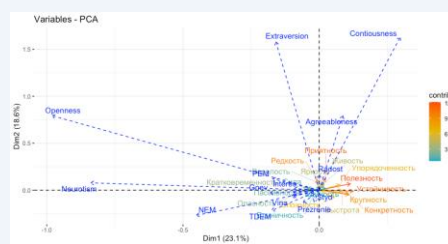


How to reveal and describe these differences?



Solution

Dataset consisting of word associations and respondent metadata + automatic evaluation of association series over binary scales (semantic differential) using distributional semantic model and concept mover distance + multivariate methods (PCA, HPCP)



The research is supported by Russian Science Foundation



Results

Clusters of similar word meanings were obtained for highly frequent Russian words

The effect of psychological traits and emotional states of respondents on word meaning was revealed and thoroughly described

Deception (обман) has the fewest semantic differences among respondents, *Russian (русский)* and *mom (мама)* are close to each other in multidimensional space, while *think (думать)* has controversial meanings for different groups of respondents

Code and dataset are available on github



Введение

Исследование индивидуальных различий в значении слова относится к одной из фундаментальных задач психолингвистики и психосемантики. Указанная проблема требует привлечения экспериментальных данных и комплекса исследовательских методов, ведущими из которых являются ассоциативный эксперимент в различных его видах и методы семантического шкалирования. Несмотря на большое число исследований в данной области и значительные теоретические достижения, существует ряд проблем методологического характера, решение которых, в том числе с использованием современных методов автоматической обработки текстов (NLP) и информационных технологий в целом, способствовало бы дальнейшему прогрессу исследований в области изучения особенностей индивидуального значения слова.

Во-первых, общедоступные базы данных ассоциаций, как правило (см. подробнее: Литвинова и др., 2022), не дают возможности исследовать значение слова на индивидуальном уровне, поскольку

содержат агрегированные данные об ассоциативном значении слова (агрегация проводится на уровне слова-стимула либо – реже – на уровне групповых характеристик респондентов). Между тем на необходимость создания индивидуального словаря ассоциаций указывал еще Ю. Н. Караулов в работе 1993 года (Караулов, Коробова, 1993). Е. И. Горошко в работе 2001 года отмечает, что индивидуальное ассоциативное поведение очень редко выступает предметом исследовательского внимания (Горошко, 2001). В работе 2022 года (Wulff et al., 2022) содержится описание датасета (проект My Small World of Words) индивидуальных ассоциативных реакций 8 респондентов, при этом указывается на уникальность такого датасета; очевидно, создание подобных наборов данных является актуальной, но крайне трудоемкой исследовательской задачей (см. подробнее об этом: Литвинова, Любова, 2023).

Кроме того, имеющиеся базы данных ассоциаций, как правило, не позволяют оценивать влияние эмоционального состояния, психологических особенностей

человека и т.д. на характер продуцируемых им ассоциаций. Как представляется, создание актуальных (то есть периодически обновляемых) баз данных ассоциаций в рамках концепции гражданской науки, в том числе с использованием мобильных приложений с элементами геймификации, является перспективным направлением дальнейших исследований в данной области, учитывая наблюдающийся интерес интернет-пользователей к словесным онлайн-играм (отметим, что название браузерной игры-головоломки Wordle стало самым популярным запросом в Google в 2022 году¹).

Следует отметить, что исследование вербальных ассоциаций осложняется и отсутствием общепринятой классификации ассоциативных реакций (см. подробнее об этом: Степыкин и др., 2023). Кроме того, предложенные методики анализа ассоциативных данных трудоемки и связаны с «субъективной интерпретации зафиксированных наборов ассоциаций самим исследователем» (Горошко, 2001), а также со значительными временными затратами.

В связи с указанными обстоятельствами в работах последних лет (преимущественно англоязычных) для исследования вербальных ассоциаций все чаще привлекаются ресурсы и инструментарий компьютерной лингвистики (см. обзор: Litvinova et al., 2023), в частности данных дистрибутивных семантических моделей разных типов и методов извлечения информации из них.

Дистрибутивные семантические модели (эмбединги слов, Word embeddings, далее также – ДСМ, языковая модель) опираются на связи между словом и его языковым окружением (Ellis, 2019; Lenci, 2018). ДСМ, как правило, обучаются на больших данных – многомиллионных

корпусах текстов. В процессе обучения модели накапливают данные о взаимоотношениях слов. Даже если слова не употребляются вместе, но используются в близких контекстах, считается, что они имеют схожие значения. ДСМ воспроизводят репрезентации значения слов в ментальном лексиконе, поскольку они реконструируют их на основе данных об их реальном употреблении в дискурсах разных типов. Отмечается, что ДСМ являются ценным материалом для исследования человеческого мышления и социолингвистических аспектов использования языка (см. подробнее: Arseniev-Koehler, Foster, 2022).

Учеными предпринимались многочисленные попытки параметризованного описания семантики слова на основе ДСМ. Для настоящего исследования особенно ценной является работа Hollis и Westbury (2016), в которой было показано, что из ДСМ возможно извлечь близкие к «человеческим» оценки слов по аффективным параметрам значения, традиционно выделяемым в англоязычных работах по семасиологии: оценка (valence или pleasantness), возбуждение (arousal; интенсивность эмоции, вызванной стимулом) и доминантность (dominance; степень контроля, проявляемая стимулом), что позволяет «автоматизировать» процесс семантического шкалирования и таким образом получить оценки большого числа слов (в отличие от традиционных психолингвистических экспериментов). Указанные измерения в целом соответствуют факторам, выделенным Ч. Осгудом – создателем метода семантического дифференциала. Однако названными измерениями не ограничивается состав семантических признаков, которые можно извлечь из ДСМ, как правило, содержащих несколько сот измерений.

Разработанные к настоящему времени методы компьютерной семантики позволяют извлекать из ДСМ, обученных на многомиллионных корпусах текстов, ценную информацию, с помощью которой

¹ Самым популярным запросом в Google в 2022 году стало слово «Wordle» // <https://www.kommersant.ru/doc/5708068> (дата обращения: 27.02.2024)

возможно реконструировать представления человека о тех или иных объектах и понятиях (Grand et al., 2022), в том числе в динамике (Daenekindt, Schaap, 2022), обнаруживать неосознаваемые и явные ассоциации и представления (в том числе стереотипы), которые ранее возможно было выявлять только с использованием имплицитного ассоциативного теста (Garg et al., 2018; Kozłowski et al., 2019), и устанавливать степень семантической связи текста с тем или иным концептом, даже если он напрямую не называется в тексте (Stoltz, Taylor, 2019).

Вследствие сказанного методы компьютерной семантики стали активно использоваться в социологических и культурологических работах, рассматривающих текст как источник данных (концепция *texts as data*, см. обзор: Matsui, Ferrara, 2022). Однако применительно к результатам ассоциативного эксперимента, их описанию и интерпретации подобные методы, несмотря на их широкое распространение в других областях, до настоящего времени, насколько нам известно, не применялись.

Целью настоящей работы является исследование индивидуальных различий в значении слов русского языка, важных для русского языкового сознания, с использованием методов компьютерной семантики, позволяющих получать оценки слов по шкалам семантического дифференциала, и анализ влияния психологических характеристик и эмоционального состояния респондентов на особенности ассоциативного значения слов, представленного в виде таких оценок.

Вслед за В. П. Серкиным под значением слова мы понимаем образующую сознания, которая является «субъективной формой отражения индивидуального и общественно-исторического опыта, приобретенного в процессе личной истории индивидуальных и совместных деятельностей» (Серкин, 2008: 98), что указывает на наличие в значении слова как уникальной, индивидуальной, составляющей, так и связанной с групповыми

характеристиками респондентов компоненты.

Новизна исследования определяется следующими обстоятельствами. Во-первых, нами предложен новый метод исследования индивидуальной семантики, основанный на объединении двух ведущих традиционных психолингвистических методов исследования индивидуального значения слова - ассоциативного эксперимента и семантического шкалирования - и методов компьютерной семантики. Во-вторых, впервые предложен автоматический метод получения оценок ассоциативного значения слова по шкалам семантического дифференциала. В-третьих, впервые с использованием методов многомерного анализа данных (в том числе продвинутых методов анализа, таких как кластерный анализ на главных компонентах) оценена степень выраженности индивидуальных различий в семантике слов, важных для русского языкового сознания, в зависимости от индивидуально-психологических характеристик респондентов.

В работе используется созданный авторским коллективом датасет, специально предназначенный для исследования индивидуальных различий в семантике слов: ассоциативные реакции представлены в нем в неагрегированном формате и сопровождаются обезличенными данными о респонденте (пол, возраст, результаты психологического тестирования).

Наша работа следует принципам открытой науки: датасет и код для воспроизведения результатов открыты для исследователей². Все расчеты выполнены на языке R.

1. Обзор литературы

1.1. Семантический дифференциал как метод исследования психологически реального значения слова в психосемантике и других дисциплинах

Одним из ведущих методов исследования значения слова в

² <https://github.com/Litvinova1984/semantic-differential-assoc-DSM/tree/main>

психосемантике, которая представляет собой совокупность «методологических подходов, моделей, методов, методик и результатов исследований психологического содержания (индивидуальных и групповых) значений и систем значений» (Серкин, 2008: 11–12), является метод семантического дифференциала (далее также – СД), предложенный Ч. Осгудом (Osgood et al., 1957).

Метод СД позволяет изучать значение стимула (не только слова, но и стимулов других модальностей) во множестве измерений и представляет собой комбинацию ассоциативного эксперимента и процедур семантического шкалирования тех или иных стимулов с помощью шкал-антонимов естественного языка (Серкин, 2008). Как и ассоциативный эксперимент, СД исследует коннотативную часть значения, понимаемую как «комплекс контекстуальных ассоциаций» (там же: 54). Именно условия деятельности и опыт, составляющие в совокупности индивидуальный деятельностный контекст значения, и определяют индивидуальность ассоциаций и оценок (там же).

В работах Ч. Осгуда была доказана перспективность рассмотрения семантики слова как многомерного пространства, в котором семантические различия слов преимущественно определяются различиями по аффективным измерениям. Ключевым выводом, полученным с использованием данного метода в работах Ч. Осгуда, является вывод о том, что большая часть вариации оценок стимулов по шкалам СД может быть объяснена тремя ортогональными (то есть независимыми друг от друга) факторами: оценки, силы и активности, что согласуется с теорией управления эмоциями (*affect control theory*), согласно которой люди интерпретируют события и ведут себя с учетом культурно обусловленных смыслов, определяемых такими измерениями, как *evaluation*,

potency и *activity* (EPA) (см. также: Schröder et al., 2016).

Работа Ч. Осгуда и соавторов (Osgood et al., 1957) неоднократно цитировалась в качестве доказательства низкой размерности значения слова (то есть того, что все многообразие значения слова может быть описано тремя измерениями), однако подобное утверждение не совсем верно. В более поздней работе (Osgood, 1969) исследователь подчеркивал, что СД отражает именно аффективные компоненты значения, которым, как известно, значение слова не ограничивается, – денотативные компоненты значения описываются другими параметрами. Подобное утверждение согласуется с полученными в компьютерной семантике результатами, согласно которым ДСМ кодируют компоненты значения, соответствующие современным представлениям о нейробиологически обусловленной структуре значения (Utsumi, 2018).

Результаты, полученные с использованием СД, дают возможность формализованного описания индивидуальных различий в значении слов, поскольку слово-стимул представляется как вектор в многомерном семантическом пространстве. Лингвисты, применяющие методику СД, «открывают» ранее не известные аспекты (коннотативного) значения, которое может быть выявлено и описано не только у эмоционально окрашенных слов, но и у традиционно относимых к нейтральным: «По-видимому, это связано с тем, что человек оценивает все явления, с которыми он сталкивается, отнюдь "не бесстрастно", поэтому эмоциональную окраску получают все слова, "пропущенные" через его сознание и опыт» (Глухов, 2005: 301).

Как отмечают исследователи, метод СД, «несмотря на критику исходных теоретических положений Ч. Осгуда, широко используется в практике современных лингвистических, психолингвистических и психологических

исследований» (Новиков, Новикова, 2011: 70), в социологических работах (Сикевич, 2016) и др. Так, в БД Scopus поисковый запрос *semantic differential* в поле «Название статьи, реферат ключевые слова» выдает 5324 документа, при этом наблюдается устойчивый интерес ученых к данному методу (за последние 10 лет в год публикуется около 150 статей, соответствующих указанному поисковому запросу). Большая часть изданий, публикующих данные работы (1708), относится к предметной области «Психология», на втором месте – «Общественные науки» (1501), на третьем – «Медицина» (1431). Издания, публикующие работы по разделам гуманитарных наук (куда входит и лингвистика), выпустили 640 работ, использующих методику СД.

Таким образом, методика СД является активно используемым инструментом современных исследований, применяющимся в разных областях знаний, и одним из ведущих методов исследования психологически реального значения слова.

Как и любой метод работы с респондентами, СД связан с определенными ограничениями, связанными прежде всего с ограничением по числу респондентов, оцениваемых стимулов и шкал. Кроме того, при использовании данного метода, как правило, производится сравнение ответов респондентов в рамках эксперимента, тогда как продуктивным было бы также сравнение ответов респондентов с некоторым базовым, общим уровнем. В качестве такого уровня могут выступать данные языковых моделей, обученных на многомиллионных корпусах текстов, которые могут служить в качестве универсальной базы знаний, отражающей структуру ментального лексикона носителя того или иного языка. Сравнение оценок респондентов по шкалам СД с оценками, полученными на основе ДСМ, может существенно расширить возможности СД как метода.

1.2. Семантическое шкалирование с использованием методов компьютерной семантики

В настоящее время развитие методов компьютерной семантики позволяет получить оценки отдельных слов и текстов, отражающие их положение на любых антонимических шкалах (при условии, что слова-антонимы присутствуют в ДСМ), то есть автоматизировать процедуру семантического шкалирования.

Одной из первых работ, в которых на основе ДСМ были получены оценки текстов по культурно значимым бинарным оппозициям, стала работа «Geometry of Culture» (Kozłowski et al., 2019) (на данный момент имеет более 200 цитирований в БД Scopus). Ее авторы предложили новый подход, связанный с использованием ДСМ в социологическом анализе культуры и позволяющий конструировать культурные измерения, задаваемые путем операции вычитания векторов слов в ДСМ («бедный – богатый», «красивый – уродливый», «старый – новый» и т.д.). Под культурными измерениями (cultural dimensions) понимаются параметры, которые люди используют в повседневной жизни для классификации акторов и объектов в мире (Там же: 911). Авторы цитируемой работы указывают, что проекции слов на подобные измерения отражают их частотные ассоциации (как неосознаваемые, так и регистрируемые в ходе соцопросов), присущие представителям той или иной культуры. Например, проецируя на измерение «благополучие» названия профессий, можно обнаружить, что традиционно хорошо оплачиваемые профессии, такие как банкир и юрист, располагаются на одном конце шкалы, тогда как плохо оплачиваемые (*няня*, *плотник* – все примеры взяты из цитируемой статьи) – на другом. Авторы статьи объясняют подобное обстоятельство тем, что контексты, в которых используется слово «банкир», часто содержат слова с семой «богатый», а *няня* – слова с семой «бедный».

Проанализировав тексты миллионов книг, опубликованных более чем за 100-летний период, с использованием описанного метода, Kozłowski с соавторами сделали важные выводы об изменениях в коллективных представлениях о социальных классах в США, происходивших на протяжении XX века.

В другой известной работе, связанной с выявлением стереотипов в текстах (Caliskan et al., 2017), было показано, что позиция вектора слова на шкалах «Гендер» и «Раса» тесно связана с ассоциациями, выявленными с использованием имплицитного ассоциативного теста (Greenwald et al., 1998), и отражает негативные расовые и гендерные стереотипы, присутствующие в текстах в неявном виде. Названная работа убедительно показала перспективность использования ДСМ для исследования культурных коннотаций концептов.

М. А. Taylor и D. S. Stoltz (2021) модернизировали метод Concept Movers Distance (Stoltz, Taylor, 2019), с помощью которого возможно получать количественные оценки выраженности в тексте того или иного понятия (даже если самих слов, выражающих исследуемые понятия, в тексте нет) на основе ДСМ, специально для случая бинарных шкал и реализовали его в библиотеке R text2map (Stoltz, Taylor, 2022).

Используя разработанный метод, М. А. Taylor и D. S. Stoltz (2021) продемонстрировали, что число смертей в пьесах Шекспира имело положительную статистически значимую связь с оценкой пьес по бинарной шкале «Смерть – жизнь», но подобной связи не было установлено для шкал, определенных только одним словом данной антонимической пары.

Благодаря тому, что предложенный авторами метод реализован в библиотеке на языке R, другие исследователи также стали использовать его для решения разнообразных задач, связанных с анализом текстов. Так, в работе Voyer et al. (2022) на материале исторического корпуса

текстов, описывающих социальные нормы, были исследованы изменения оценок по шкале «Strange – Normal», вычисленные для фрагментов текстов, в которых говорится об основных группах мигрантов в США, в зависимости от года создания текста, и выявлены временные характеристики и траектории подобных изменений для каждой из исследуемых групп.

В работе S. Daenekindt и J. Schaap (2022) данный метод использовался для исследования семантических изменений в дискурсе о популярной музыке по шкале «Гендер» за 20-летний период. Был проанализирован корпус из более чем 20000 текстов музыкальных обзоров, и установлено смещение оценок с течением времени в сторону «женского» полюса шкалы.

Таким образом, в компьютерной семантике разработаны и успешно апробированы методы, позволяющие извлекать из текстов оценки их семантической близости к тому или иному полюсу бинарной шкалы (даже в случае отсутствия в текстах соответствующих слов-антонимов) на основе информации, закодированной в ДСМ. Насколько нам известно, подобные методы ранее не применялись по отношению к ассоциативным реакциям, за исключением нашего пилотного исследования (Литвинова и др., 2023), в котором мы рассматривали три шкалы семантического дифференциала и анализировали связь оценок ассоциативного значения слова по данным шкалам, рассчитанным автоматически, с психологическими особенностями респондентов.

1.3. Влияние психологических характеристик и эмоционального состояния индивида на ассоциативное значение слова

Влияние эмоций на результаты ассоциативного эксперимента рассматривалось в «громадном количестве исследований» (Горошко, 2001), а сам ассоциативный эксперимент может быть использован как

психодиагностическая методика. Наиболее изучены особенности протекания процесса ассоциирования в состоянии эмоциональной напряженности, стресса, аффекта. В работах, выполненных на русскоязычном материале (см. их обзор, а также результаты собственных исследований автора в монографии (Горошко, 2001)), отмечается смысловая связь ассоциативных реакций с контекстом ситуации стресса, их повышенная оценочность, увеличение стереотипности реакций, количества отрицательно окрашенных реакций и отказов от реагирования вообще. В работе (Schmolling, 1978) говорится о противоречивости данных о влиянии стресса на характер ассоциаций, полученных на англоязычном материале, что связывается с методологическими различиями в исследованиях, затрудняющими сопоставление результатов. В самой названной работе было показано, что в ситуации стресса (в сравнении с нормой) снижается число частотных реакций, однако подобный эффект не проявлялся для слов, одинаковые ассоциаты к которым были даны более чем 60% респондентов (то есть стимулов с сильной ассоциативной связью с реакциями). Был сделан вывод о том, что уменьшение числа частотных ассоциативных реакций у слов-стимулов, имеющих слабую либо умеренную связь с ассоциатами, является типичной реакцией на стресс.

Работа Isen et al. (1985) стала одним из первых исследований, посвященных анализу влияния позитивного аффекта на частотность ассоциативных реакций (на данный момент статья имеет более 600 цитирований в Scopus и 1500 цитирований в Google Scholar). В исследовании было показано, что респонденты в состоянии позитивного аффекта давали больше необычных ассоциативных реакций на эмоционально нейтральные слова, чем в обычном состоянии. Отметим, что в данной работе, наряду с аффектом, рассматривалось влияние эмоциональной оценочности стимула на характер ассоциативных реакций, и было показано,

что реакции на положительно окрашенные слова более разнообразны и менее стереотипны, чем реакции на стимулы с нейтральной и отрицательной эмоциональной окраской.

В работе (Lukavsky, 2004) описаны результаты эксперимента, в ходе которого респондентам предлагалось не только продуцировать ассоциации, но и оценить валентность слова-стимула. Была установлена тенденция респондентов оценивать нейтральные слова как позитивные. Автор указанной работы связывает данное наблюдение с контрастом между негативными и нейтральными словами, который воспринимается как более значительный, нежели контраст между позитивными и нейтральными словами. Участники эксперимента, страдающие маниями и зависимостями, воспринимали нейтральные слова как еще более позитивные, чем респонденты в контрольной выборке. Напротив, страдающие тревожными и депрессивными расстройствами, а также респонденты с расстройствами личности воспринимали нейтральные слова как более негативные. Восприятие лиц, страдающих шизофренией, отличалось от всех остальных групп, однако общего тренда установить не удалось. Цитируемая работа ценна для нас тем, что в ней указывается на важность дополнения ассоциативного эксперимента данными о субъективных оценках эмоциональной окрашенности слов-стимулов.

Таким образом, как показал анализ соответствующих работ, при исследовании влияния эмоционального состояния индивида на характер продуцируемых им ассоциативных реакций необходимо учитывать множество факторов, в том числе характеристики самого слова-стимула (его эмоциональную окрашенность, силу связи с ассоциатами) и психологические особенности респондентов.

Кроме того, как отмечается самими исследователями, работающими в данной

предметной области, сравнение описанных в различных работах результатов затруднено вследствие сложности самой проблемы: существуют различные подходы к пониманию термина «эмоция», отсутствуют общепринятые классификации эмоций. Продуктивной представляется комплексная оценка эмоционального состояния респондента по набору теоретически обоснованных шкал. Е. И. Горошко предлагает использовать для подобных исследований «в качестве некоторой обобщающей концепции по теории эмоций» (Горошко, 2001) концепцию теории дифференциальных эмоций Томкинса, основываясь на которой К. Изард выделила 10 базовых эмоций, которые определяют компонентный состав разнообразных проявлений эмоциональной жизни (интерес, радость, удивление, горе, гнев, отвращение, презрение, страх, стыд, вина). В рамках данной теории эмоции рассматриваются как элементы системы, которые могут противопоставляться друг либо же находиться в иерархических отношениях (см. подробнее о данной классификации с описанием каждой эмоции в работе (Зуева, 2006)).

Мало исследованной является проблема влияния устойчивых личностных характеристик информантов на результаты ассоциативного эксперимента. В англоязычных работах большое внимание уделяется анализу влияния такой характеристики, как креативность, на характер ассоциативных реакций (см., например, работу Merseal et al. (2023), выполненную на основе данных, собранных в ходе масштабного проекта Big-C Project, посвященного исследованию креативности, в том числе путем сравнительного анализа ассоциативной продукции известных представителей творческих профессий, ученых и рядовых носителей языка), при этом преимущественно исследуется уровень стереотипности/необычности реакций (в работах последних лет, в том числе в

процитированной выше, – с использованием данных ДСМ).

В ряде работ указывается на связь отдельных психологических характеристик информантов и особенностей ассоциирования. Так, на связь оценок респондентов по шкалам «Экстраверсия/Интроверсия», «Торможение/Импульсивность» со временем реакции, числом частотных реакций указывается в работе (Eberhard & Owens, 1975). Т. Merten отмечает положительную корреляцию между выраженностью психотизма и числом уникальных ассоциативных реакций в различных условиях проведения ассоциативного эксперимента (Merten, 1993). В названных работах, а также в исследовании (Innes, 1972) отмечается, что сравнение полученных результатов затруднено вследствие использования разных инструментов оценки психологических характеристик респондентов и опоре на разные теоретические концепции.

Как и в случае с эмоциями, продуктивным представляется исследование комплекса личностных характеристик, составляющих систему. Одной из наиболее влиятельных и широко используемых теоретических моделей личности человека является «Большая Пятерка» («The Big Five»). Данная модель состоит из следующих черт личности: экстраверсия, доброжелательность, добросовестность, нейротизм (другое название – негативная эмоциональность) и открытость опыту. Валидность модели была подтверждена и на российской популяции, разработаны инструменты для оценки выраженности черт большой пятерки у русскоязычных респондентов (см. подробнее о модели и русскоязычных тестах: Калугин и др., 2021).

Данная модель является наиболее часто используемой в современных работах (преимущественно англоязычных), посвященных психологическому профилированию автора текста и – шире – проблеме отражения в тексте психологических характеристик его создателя (см.

обзор: Azucar et al., 2018). Нам, однако, не известно о работах, в которых бы названная модель применялась по отношению к анализу ассоциативных реакций в их связи с психологическими характеристиками респондентов (отметим, что работы по влиянию психологических характеристик информантов на результаты ассоциативного эксперимента датируются в основном 1960-1990-ми годами).

По нашему мнению, для дальнейшего прогресса в области изучения влияния эмоционально-психологических характеристик индивидов на особенности ассоциативного (то есть психологически реального) значения слова необходимо использование валидных тестов, рассматривающих эмоциональное состояние и психологические характеристики респондентов как систему, учет эмотивности слов-стимулов и использование методов анализа ассоциативных реакций, предполагающих возможность автоматизации и воспроизводимости результатов и, следовательно, снижающих уровень субъективности анализа и временные затраты исследователя на его проведение. Подобные методики не отменяют, безусловно, качественный («ручной») анализ ассоциативных данных, но могут служить его эффективным дополнением.

2. Материал и методы

2.1. Материал

2.1.1. Описание эксперимента

Материалом данного исследования выступил собранный нами датасет ассоциативных реакций к 50 словам, важным для русского языкового сознания (Литвинова и др., 2022). Слова были отобраны нами из списка стимулов Русского сопоставительного ассоциативного словаря (РСпАС) (Черкасова, 2008), содержащего 253 стимула, отобранных по материалам трех ассоциативных опросов, проведенных с интервалами 10–20 лет с носителями русского языка. В данный словарь вошли русские слова-стимулы, которые повторялись в трех или двух массовых ассоциативных опросах, по материалам которых опубли-

кованы важнейшие ассоциативные словари русского языка. Слова-стимулы тщательно отбирались авторами словарей как вербализующие наиболее важные для русского языкового сознания объекты, понятия, свойства, действия и т.д.

Для создания датасета из 253 стимулов были отобраны 47 на основании частоты встречаемости в текстах разных жанров, входящих в БД RusIdiolect (Litvinova, 2021) (поскольку создание датасета являлось одним из этапов сравнительного исследования особенностей индивидуально-значения слова в ассоциативном эксперименте и дискурсе): *бабушка, бог, богатый, враг, встретить, глаз, глупый, добро, дом, дорогой, друг, думать, душа, жизнь, история, лес, любить, маленький, мама* (в РСпАС отсутствует, но есть слово «мама»), но в материале БД RusIdiolect частотность слова «мама» значительно выше, чем у слова «мать», в связи с чем в итоговый список стимулов мы внесли именно эту лексему), *место, много, настоящий, начало* (в РСпАС отсутствует, но есть *начать*), *ненавидеть, новый, обещать, обман, огонь, плохо, помощь, простой, путь, радость, результат, родной, русский, свет, связь, семья, смерть, смысл, счастье, труд, увидеть, учиться, хотеть*. К этому списку были добавлены три слова, частотных в текстах БД RusIdiolect, но отсутствующих в РСпАС, - *дерево, мир* (оба слова входят, согласно данным (Уфимцева, 2009), в ядро русского языкового сознания), а также слово *Россия*.

Опросная анкета³ содержала 50 слов-стимулов, при этом 10 из них были представлены два раза в случайном порядке для проверки устойчивости ассоциативного значения слова в языковом сознании (*добро, дом, друг, жизнь, мир, настоящий, путь, семья, счастье, хотеть*), 17 стимулов были продублированы изображениями. Таким образом, каждый участник эксперимента должен был продуцировать

³ <https://github.com/Litvinova1984/semantic-differential-assoc-DSM/blob/main/Questionnaire.pdf>

77 (50 вербальных стимулов + 10 повторяющихся вербальных стимулов + 17 изображений) ассоциативных рядов.

В эксперименте приняли участие 49 человек (из них 42 женщины, возраст участников – 17-25 лет) – студентов российских вузов (Воронежа, Перми, Санкт-Петербурга), для которых русский язык является родным. Информация об эксперименте распространялась в соцсетях; желающие принять участие в эксперименте получали ссылку на анкету и выполняли задание на личных электронных устройствах. Были также получены логи файлов, дающие возможность исследовать клавиатурное поведение респондентов во время продуцирования ассоциативных реакций (Литвинова и др., 2022).

Перед началом эксперимента был проведен онлайн-семинар, в ходе которого участники получили дополнительные разъяснения по процедуре эксперимента. Участники получали инструкцию в письменном виде дать не менее 5 ассоциативных реакций на каждый стимул (вербальный либо визуальный). Считается, что подобный способ проведения эксперимента (продолженный ассоциативный эксперимент) «обеспечивает доступ не к вербальным сетям, а к глубинным структурам знания (долговременной памяти человека) и запускает процессы естественного семиозиса (смыслопорождения) в соответствии с нормами и правилами национальной культуры» (Курганова, 2019: 27), а письменный способ фиксации реакций – в нашем случае подкрепленный также записями логов – «еще больше усиливает определенную направленность сознания участников экспериментов на актуализацию того или иного фрагмента сознания» (там же).

В итоге был получен 3771 ассоциативный ряд от 49 респондентов (в двух анкетах отсутствовали по 1 ассоциативному ряду).

Большинство респондентов выполнили задания в соответствии с инструкцией, однако некоторые респонденты указа-

ли меньшее число ассоциатов, чем того требовала инструкция.

Помимо выполнения заданий по продуцированию ассоциаций все респонденты до выполнения задания проходили тест на определение эмоционального состояния, а также – в любое удобное для них время – тест «Большая пятерка» и заносили результаты обоих тестов в анкету. Конкретные варианты тестов были выбраны в ходе консультации со специалистом-психологом. Тестирование проводилось онлайн, по предложенным респондентам ссылкам. Для получения оценок по 5 шкалам «Большой пятерки» использовалась одна из последних методик тестирования, свободно доступная для некоммерческого использования (Калугин и др., 2021). Для исследования эмоционального состояния респондентов использовался тест «Шкала дифференциальных эмоций», в результате прохождения которого респонденты получали оценки по 10 шкалам, отражающим уровень выраженности эмоций, выделяемых в рамках теории дифференциальных эмоций, а также по 3-м комбинированным шкалам: шкале позитивных эмоций (ПЭМ) и шкале острых негативных эмоций (НЭМ), показывающих соответственно степень позитивного либо общий уровень негативного эмоционального отношения к ситуации, а также по шкале тревожно-депрессивных эмоций (ТДЭМ), которая показывает уровень относительно устойчивых индивидуальных переживаний тревожно-депрессивного комплекса эмоций, опосредующих субъективное отношение к ситуации.

Данные анкетирования были сведены в единый табличный файл, где строки соответствуют индивидуальным ассоциативным рядам, колонки – метаданным (ID респондента, тип стимула, баллы по тестам)⁴.

⁴ https://github.com/Litvinova1984/semantic-differential-assoc-DSM/blob/main/raw_dataset.xlsx

2.1.2. Описание датасета

В данном исследовании анализировались только ассоциативные реакции на вербальные стимулы. Далее был проведен контроль выполнения задания путем анализа логов: были сверены текстовые файлы и файлы логов, после чего из дальнейшего анализа были удалены ответы респондента 14 (для одного стимула пользователь подобрал 1 ассоциацию, что не соответствует исходному заданию), респондента 31 (в логе у многих стимулов записан не тот порядок следования, что в итоговом файле, а также разнятся ассоциаты), респондента 33 (приведен либо один ассоциат, либо они не приведены вообще).

Далее были удалены ряды, содержащие множественные (более двух реакций) ассоциаты в виде словосочетаний. Если ряд содержал 1-2 реакции-словосочетания, они сокращались до одного слова, несущего основную смысловую нагрузку, путем независимой разметки тремя дипломированными лингвистами (например, *быть популярным* было сокращено до *популярный* и т.д.); в случае разночтений реакция удалялась.

Далее реакции были обработаны при помощи библиотеки *udpipe* (определена лемма и часть речи), после чего удалены ассоциаты, отсутствующие в ДСМ (как правило, это имена собственные, а также служебные слова и местоимения).

Для дальнейшего анализа мы оставили только ассоциативные ряды, состоящие (после обработки) из одинакового числа ассоциативных реакций ($n = 5$) (были удалены как ряды, состоящие из меньшего числа слов, так и ряды с большим числом слов). Всего для экспериментов, результаты которых представлены в настоящей статье, был отобран 1981 ассоциативный ряд (число респондентов – 46). Данные о распределении ассоциативных реакций по

авторам и стимулам в виде графиков в дополнительных материалах к статье⁵.

Представленная схема обработки материала, безусловно, не является единственно возможной, однако для данного эксперимента, который является пилотным, мы выбрали подобный «строгий» подход к отбору ассоциативных рядов с целью их максимальной сопоставимости по типу реакций (однословные, многословные) и их числу. В дальнейших исследованиях нами будет проведено сопоставление полученных результатов с данными, полученными на материале без обработки многословных ассоциаций и удаления рядов, содержащих разное число ассоциаций.

Датасет исследования размещен в дополнительных материалах к статье⁶.

2.2. Методы анализа материала

2.2.1. Списки шкал семантического дифференциала

Список шкал семантического дифференциала был составлен на основе материалов, представленных в работе В. Ф. Петренко «Основы психосемантики» (Петренко, 2005: 94–97). Ниже приведены названия факторов и соответствующие антонимические пары, выделенные В. Ф. Петренко; названия шкал даны нами условно для удобства обозначения на графиках и понимания по тексту:

1) фактор «Оценка»: *приятный* – *неприятный* (шкала «**Приятность**»), *полезный* – *бесполезный* (шкала «**Полезность**»), *веселый* – *грустный* (шкала «**Веселость**»), *живой* – *безжизненный* (шкала «**Живость**»);

2) фактор «Активность»: *быстрый* – *медленный* (шкала «**Быстрота**»), *яркий* – *тусклый* (шкала «**Яркость**»), *пассивный* – *активный* (шкала «**Пассивность**»);

3) фактор «Упорядоченность» (или «Стабильность»): *упорядоченный* – *хаотичный* (шкала «**Упорядоченность**»),

⁵ <https://github.com/Litvinova1984/semantic-differential-assoc-DSM/tree/main/Figures>

⁶ https://github.com/Litvinova1984/semantic-differential-assoc-DSM/blob/main/dataforanalysis_rus.xlsx

устойчивый – изменчивый (шкала «Устойчивость»), кратковременный – длительный (шкала «Кратковременность»), единый – дискретный (шкала «Единность»);

4) фактор «Сложность»: конкретный – абстрактный (шкала «Конкретность»), сложный – простой (шкала «Сложность»);

5) фактор «Сила»: большой – маленький (шкала «Крупность»), сильный – слабый (шкала «Сила»);

6) фактор «Обычность»: редкий – частый (шкала «Редкость»), единственный – повторяющийся (шкала «Единичность»);

7) фактор «Комфортность»: опасный – безопасный (шкала «Опасность»).

Указанные шкалы выступают в нашем исследовании как бинарные оппозиции, которые будут сконструированы на основе ДСМ. Мы не объединяли шкалы по факторам; отдельным этапом нашего исследования являлся факторный анализ оценок ассоциативных рядов по данным шкалам, вычисленным на основе ДСМ, результаты которого мы сравнили с выделенными В. Ф. Петренко факторами.

2.2.2. Языковая модель

Предобученные языковые модели рекомендуются использовать для анализа текста в широком культурном и лингвистическом контексте (Taylor, Stoltz, 2021), что и является одной из задач нашего исследования – получение оценок ассоциативного ряда как единого целого по шкалам семантического дифференциала на основе статистики совместной встречаемости слов в больших корпусах текстов (насчитывающих миллионы слов и тысячи контекстов).

Для настоящей работы нами была использована предобученная модель word2vec. Выбор данной модели обусловлен тем, что, как показали исследования, она является одной из наиболее эффективных в задаче предсказания оценок близости слов, даваемых респондентами (Jordan et al., 2022; Pereira et al., 2016), и одной из наиболее часто используемых в

исследованиях, применяющих те или иные методы семантической проекции.

Для вычисления оценок по шкалам семантического дифференциала была использована предобученная на корпусах НКРЯ и Википедии за ноябрь 2021 года модель

ruwikiruscorporga_upos_cbow_300_10_2021⁷.

Размер корпуса, на котором обучалась модель, – 1,2 миллиарда слов, объем словаря – 249 333 слов, размер окна – 10. Модель, выпущенная в декабре 2021 года, является наиболее «свежей» из доступных для скачивания и наиболее близка к дате создания нашего датасета (октябрь – декабрь 2022 года). Как отмечается в литературе (Taylor, Stoltz, 2021), важным является выбор языковой модели, максимально близкой по временному периоду ко времени создания анализируемых данных. Кроме того, поскольку языковая модель лежит в основе работы функции, извлекающей значение СД, важным является выбор «надежной» модели, натренированной на большом объеме текстов.

2.2.3. Получение значений оценок по шкалам семантического дифференциала

Для получения основанных на языковой модели оценок слов, входящих в ассоциативный ряд, по антонимическим шкалам мы использовали функцию CMDist из библиотеки text2map на языке R (Stoltz, Taylor, 2022). Как отмечалось в разделе 1.2, данная функция вычисляет для каждого документа (в нашей работе мы рассматриваем каждый ассоциативный ряд как «документ») значение, отражающее степень близости всех составляющих документа слов к полюсам заданной антонимической пары (то есть представляет собой оценку по шкале СД). Функция get_direction позволяет получать оценки по шкале несколькими способами, в том числе использовать для каждого полюса шкалы несколько слов и усреднять их векторы, однако в нашей работе мы стремились максимально приблизиться к

⁷ <https://rusvectors.org/ru/models/>

методике семантического дифференциала, поэтому каждый полюс шкалы соответствовал одному слову, в связи с чем мы использовали дефолтную опцию 'paired' функции `get_direction`: для каждой антонимической пары из каждого слова вычитается строго одно слово. Таким образом, каждый ассоциативный ряд (который обрабатывался как текст) получал оценку по каждой шкале СД.

Теоретически разброс значений оценок лежит в пределах от -3 до 3 , однако для сравнения оценок по словам данные необходимо нормировать путем приведения их к одному диапазону. На следующем этапе анализа мы нормализовали оценки, приведя их значение в диапазон от 0 до 1 с использованием пакета `Wmisc` (опция нормализации – `range`, первое значение в диапазоне $0-1$ представляет собой заменное минимальное значение переменной, второе – максимальное), так что значения СД ближе к 1 указывали на близость оценки к полюсу СД, выраженному первым словом оппозиции (*приятный, веселый*) (или степень выраженности обозначенного в названии шкалы понятия – приятности, веселости), значения, близкие к 0 , – ко второму члену оппозиции (*неприятный, грустный*), значения, близкие к $0,5$, указывали на слабую связь ассоциативного значения слова-стимула с данной шкалой СД.

2.2.4. Методы анализа данных

При подборе методов исследования мы ориентировались на необходимость получения интерпретируемых результатов, а также на возможность как анализа явной (прямой) структуры данных, так и выявления имплицитной (скрытой) структуры данных – оценок по шкалам СД: «Сопоставление результатов, полученных при помощи двух методов, является наиболее информативным» (Серкин, 2008: 276). В связи с этим мы использовали совокупность методов анализа данных: 1) эксплораторные методы, описывающие явную структуру данных (вычисление коэффициента вариации оценок по шкалам

СД, составление и визуализация профилей слов); 2) методы, направленные на поиск скрытой структуры в данных, – методы сокращения размерности (РСА), факторный анализ, кластеризация на главных компонентах с последующей интерпретацией различий между кластерами.

3. Результаты и обсуждение

3.1. Анализ явной структуры данных

На первом этапе анализа мы вычислили коэффициент вариации оценок по каждой шкале СД.

Коэффициент вариации (*coefficient of variation*, *cv*, далее также – *КВ*) количественно выражает вариативность исследуемого признака и вычисляется путем деления стандартного отклонения на среднее и умножения на 100 , что позволяет проводить сравнение вариативности разных признаков.

Высокие значения *КВ* указывают на более высокую вариативность признаков, то есть более широкий разброс их значений относительно среднего, и наоборот.

Значения *КВ* для исследуемых шкал СД были рассчитаны двумя способами для исследования разных аспектов вариативности оценок. Нами был вычислен разброс значений *КВ* оценок слов-стимулов по шкалам СД, сгруппированным по словам-стимулам (например, был вычислен коэффициент вариации оценок по шкале «**Опасность**» для слова *бабушка*, данных всеми респондентами, то же самое для слова *бог* и т.д.; далее на основе полученных значений построены диаграммы размаха, на которых «ящик» отражает 50% значений *КВ*, выбросы обозначены красными точками). Также нами была построена диаграмма размаха, отражающая дисперсию *КВ* оценок всех стимулов по шкалам СД, сгруппированных по респондентам (рисунки размещены в папке⁸).

Мы сравнили значения *КВ* оценок, сгруппированных по словам-стимулам и по респондентам, с использованием

⁸ <https://github.com/Litvinova1984/semantic-differential-assoc-DSM/tree/main/Figures/>

t-критерия Уэлча и установили, что для всех шкал СД средние значения коэффициента вариации оценок, сгруппированных по респондентам, больше, чем значения КВ оценок, сгруппированных по стимулам.

Наиболее широкий разброс оценок наблюдается для шкал «Яркость», «Сложность», «Конкретность», «Единичность». Более «согласованны» оценки по шкалам «Кратковременность», «Полезность», «Редкость» (сказанное касается сгруппированных как по респондентам, так и по стимулам КВ оценок).

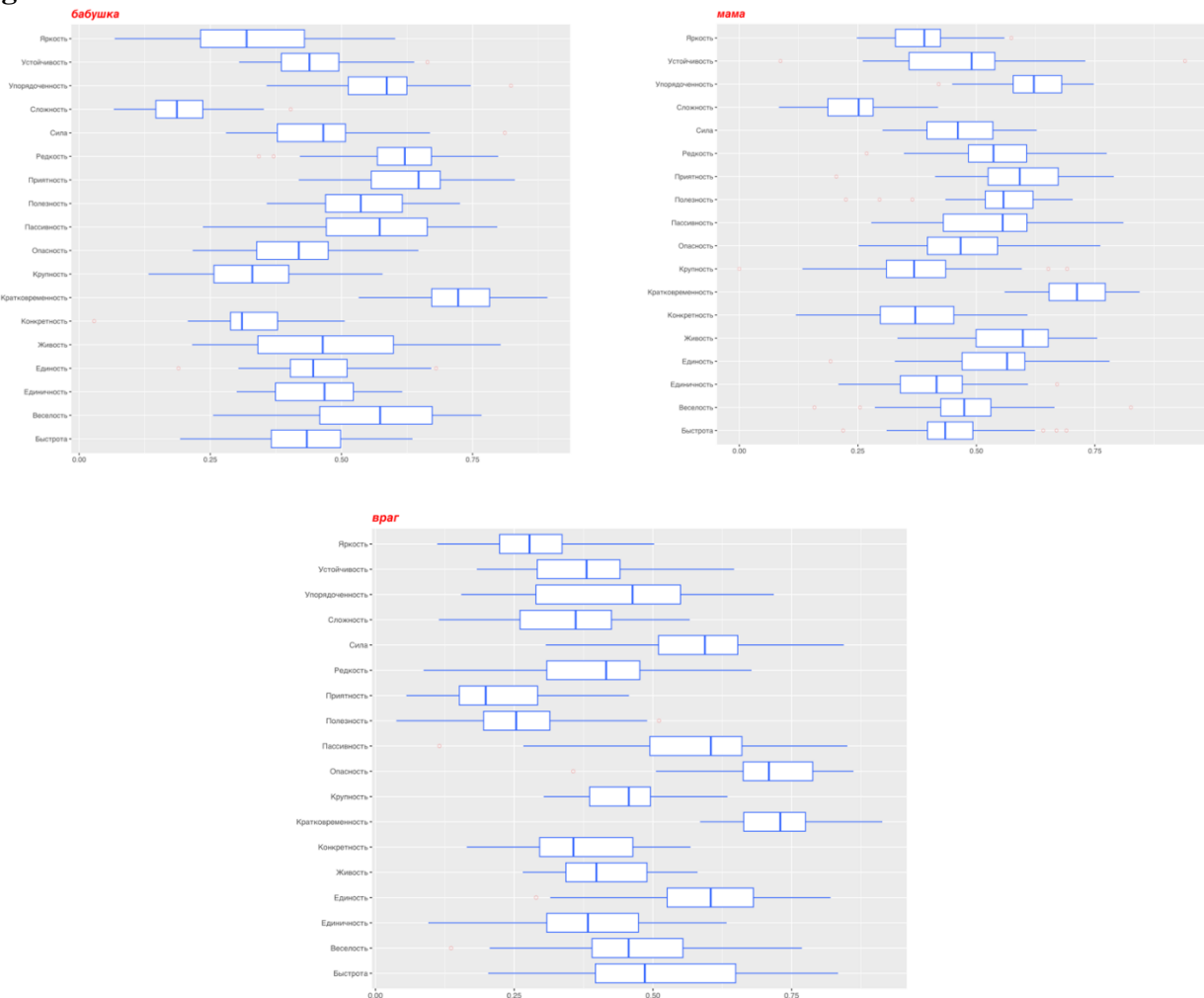
Далее для каждого слова-стимула нами был построен профиль оценок по всем шкалам СД. Рассмотрим построенные профили на примерах слов-стимулов *бабушка*, *мама* и *враг* (рис. 1). Подобная визуализация («ящики с усами») позволяет

увидеть различия в профилях слов, отражающие не только медианные значения, но и размах вариации.

Мы видим, что профили стимулов *мама* и *бабушка* более похожи друг на друга, чем на профиль стимула *враг* (что ожидаемо), однако и между ними есть различия. Стимул *бабушка* имеет более высокие средние оценки по шкале «Веселость», «Пассивность», «Приятность», стимул *мама* – более высокие оценки по шкале «Единичность», «Живость», «Опасность», «Крупность», «Конкретность», «Сложность».

Стимул *бабушка* имеет большую вариацию оценок полезности, живости, яркости, веселости, быстроты, *мама* – устойчивости, конкретности.

Рисунок 1. Профили отдельных стимулов
Figure 1. Profile of individual stimuli



Враг в сравнении с приведенными выше стимулами менее «яркий», «упорядоченный» (хотя оценка по данному признаку имеет широкий размах), более «сложный», «сильный», «опасный», менее «редкий», «приятный», «полезный», «живой», менее веселый (но только в сравнении с *бабушкой*), более быстрый.

Построение подобных профилей (еще больше примеров визуализаций приведено в папке⁹ с использованием «ящика с усами» является эффективным методом сравнения коннотативного значения слов у отдельных респондентов, групп респондентов и т.д. и позволяет выявить ранее не известные компоненты значения.

3.2. Анализ скрытой структуры данных

Исследования с использованием коэффициента вариации и визуализация профилей слова с использованием «ящика с усами» позволяют анализировать явную (Серкин, 2008) структуру в данных. Для выявления скрытой структуры в датасете – наборе ассоциативных рядов с метаданными – будут использованы методы снижения размерности (в частности, метод главных компонент, PCA) с последующей кластеризацией на главных компонентах (НСРС) и интерпретацией различий в кластерах.

Первым этапом выявления скрытой структуры в данных является анализ корреляций между параметрами. Корреляционный анализ оценок AP по шкалам СД (рис. 2; на рис. также обозначены уровни корреляционной связи). Построенная при помощи пакета *corrplot* коррелограмма позволяет выявить группы связанных признаков и направления корреляции. Среди

наиболее значимых положительных корреляций – корреляции оценок по шкалам «Крупность» и «Устойчивость», «Крупность» и «Сложность», «Устойчивость» и «Конкретность», «Полезность» и «Устойчивость», «Яркость» и «Живость», «Приятность» и «Веселость». Среди наиболее значимых отрицательных корреляций – корреляции оценок по шкалам «Конкретность» и «Сила», «Сложность» и «Кратковременность», «Крупность» и «Кратковременность», «Устойчивость» и «Кратковременность», «Устойчивость» и «Опасность», «Упорядоченность» и «Опасность». В целом, как показывает критерий Kaiser-Meyer-Olkin (КМО), данные коррелируют друг с другом умеренно (КМО = 0.634, Bartlett test 23212.97, df(153), p.value < 0.001), что означает их пригодность для дальнейшего факторного анализа.

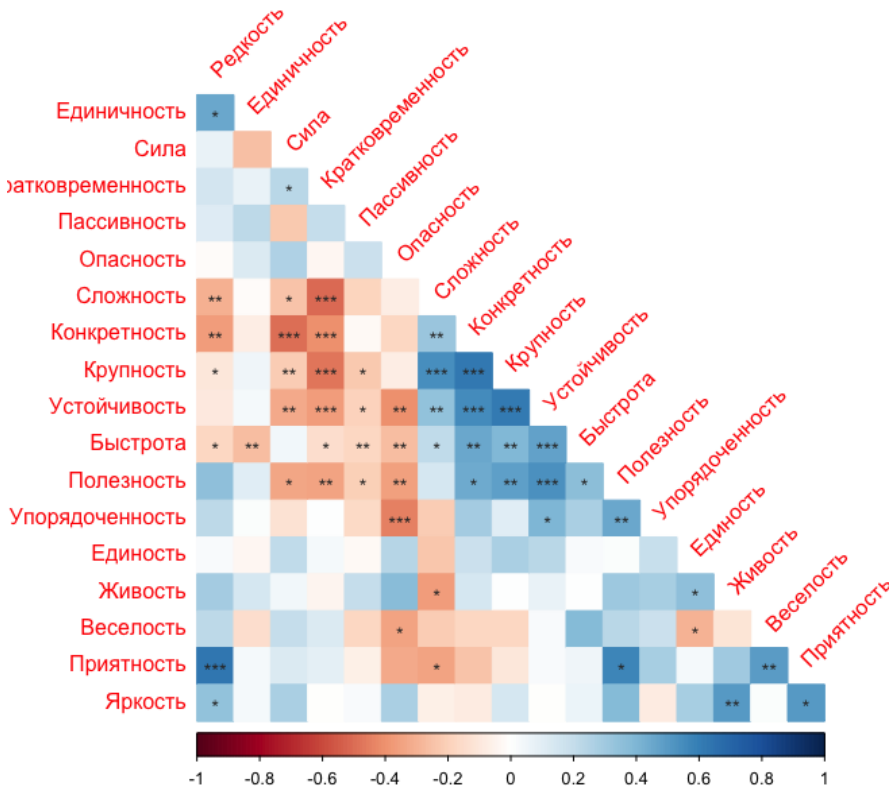
Коррелограмма, построенная по результатам психологического тестирования информантов (рис. 3), показывает наличие связей между исследуемыми характеристиками эмоционального состояния (прежде всего – между различными негативными эмоциями, а также между радостью, интересом и интегральным показателем положительных эмоций РЕМ). Черты «Большой пятерки» в среднем коррелируют умеренно (отметим отрицательную корреляцию между негативной эмоциональностью и экстраверсией).

Эмоциональные состояния и психологические характеристики коррелируют на довольно низком уровне (за исключением отрицательной корреляции стыда и экстраверсии).

⁹ <https://github.com/Litvinova1984/semantic-differential-assoc-DSM/tree/main/ПРОФИЛИ%20СЛОВ>

Рисунок 2. Корреляционная матрица оценок по шкалам семантического дифференциала (***) – $p < 0.001$, ** – $p < 0.01$, * – $p < 0.05$)

Figure 2. Correlation matrix of ratings according to semantic differential scales (***) – $p < 0.001$, ** – $p < 0.01$, * – $p < 0.05$)



Для проведения факторного анализа (метод PCA, вращение varimax) мы оставили первые 7 компонент (по числу выделенных в работе В. Ф. Петренко (2005) факторов). Кроме того, выбор данного значения соответствует правилу локтя (elbow rule). В общей сложности 7 компонент объясняют 82,37% вариации, что является хорошим показателем (рекомендуется для

дальнейшего анализа оставлять компоненты, объясняющие не менее 70% вариации) (Abdi and Williams, 2010).

В таблице 1 представлены факторы в порядке убывания собственных значений и объясненной вариации и наибольшие (свыше 0,5 по модулю) нагрузки шкал по выделенным факторам.

Рисунок 3. Корреляционная матрица данных психологического анкетирования респондентов (***) – $p < 0.001$, ** – $p < 0.01$, * – $p < 0.05$)

Figure 3. Correlation matrix of values of psychological characteristics of respondents (***) – $p < 0.001$, ** – $p < 0.01$, * – $p < 0.05$)

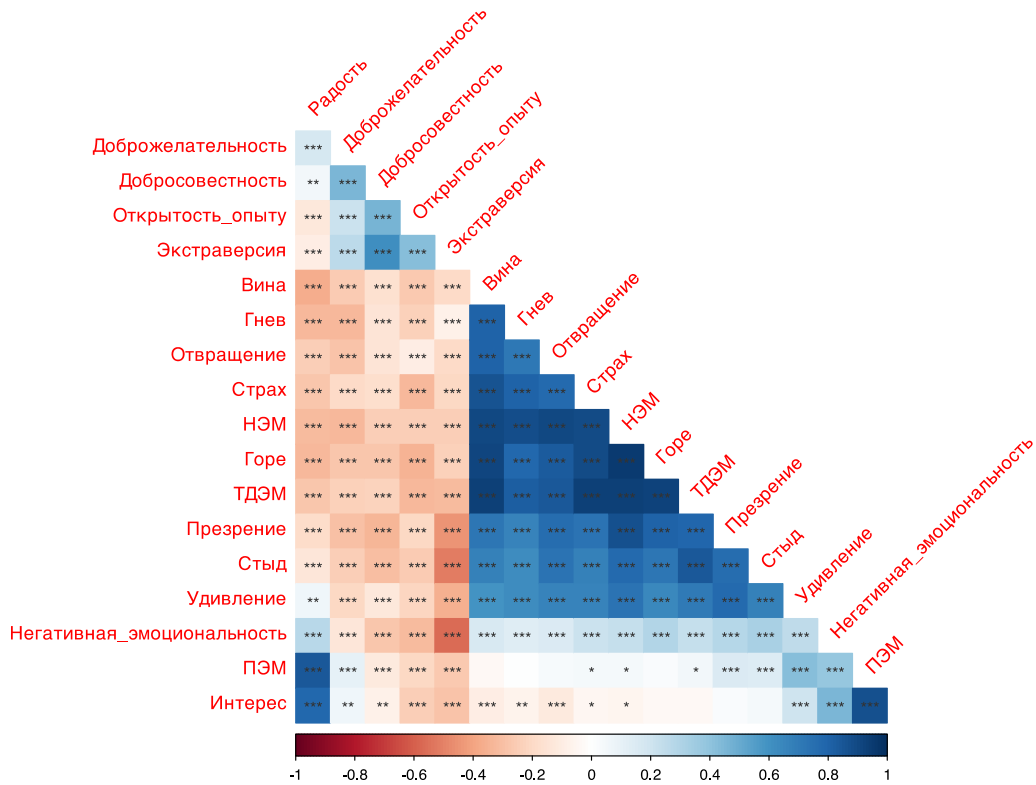


Таблица 1. Выделенные факторы и нагрузки шкал по факторам
Table 1. The selected factors and scales loadings

Шкала СД /Фактор и объясненная вариация, %	1 (RC1), 16,97	2 (RC2), 16,08	3 (RC5), 11,63	4 (RC3), 11,48	5 (RC7), 10,09	6 (RC4), 8,56	7 (RC6), 7,57
Приятность		0,88					
Полезность							
Веселость				-0,51			
Живость				0,65			
Быстрота	0,75						
Яркость		0,71					
Пассивность							0,85
Упорядоченность			0,82				
Устойчивость	0,74						
Кратковременность					0,78		
Единость				0,85			
Конкретность	0,66						
Сложность							

Шкала СД /Фактор и объясненная вариация, %	1 (RC1), 16,97	2 (RC2), 16,08	3 (RC5), 11,63	4 (RC3), 11,48	5 (RC7), 10,09	6 (RC4), 8,56	7 (RC6), 7,57
Крупность	0,82						
Сила					0,53		-0,47
Редкость		0,69				0,53	
Единичность						0,87	
Опасность			-0,69				

Наибольший процент вариации объясняют первые два фактора, один из которых (фактор 2) может быть назван фактором оценки. Как отмечает В. Ф. Петренко на основании проведенных им экспериментальных исследований, данный фактор является ведущим в экспериментах с вербальными стимулами. Максимальную нагрузку по данному фактору в наших экспериментах имеет шкала «Приятность» (отметим, что нагрузки данной шкалы по фактору оценки на уровне 0,8 отмечены и в цитируемой работе В. Ф. Петренко). Также высокие нагрузки по данному фактору имеют шкалы «Яркость», «Полезность», «Редкость», причем во всех этих шкалах названный фактор является доминирующим.

Фактор 1, объясняющий, наряду с фактором оценки, наибольший процент вариации в наших данных (наибольшие нагрузки по данному фактору имеют шкалы «Быстрота», «Крупность» «Устойчивость», «Конкретность»), может быть условно назван фактором «Устойчивость». Он не имеет прямых аналогов среди выделенных В. Ф. Петренко факторов.

В нашем материале, как и в экспериментах В. Ф. Петренко, выделяется фактор «Упорядоченность» (фактор 3). Отметим отрицательную корреляцию шкалы «Опасность» с данным фактором («то, что упорядочено, безопасно»).

В нашем материале, в отличие от экспериментов В. Ф. Петренко, выделяется отдельный фактор «Единость» (фактор 4), наибольшую нагрузку по которому имеет шкала с соответствующим названием (в экспериментах В. Ф. Петренко данная шкала входит в фактор «Упорядочен-

ность»). Также высокую нагрузку по данному фактору имеет шкала «Живость».

Осгудовский фактор **силы** (фактор 5) представлен двумя основными шкалами. Помимо силы, наибольшую нагрузку по данному фактору имеет шкала «Кратковременность», что указывает на положительную связь соответствующих понятий.

Отметим полное совпадение состава шкал, представляющих фактор «**Обычность**» (фактор 6), в наших данных и экспериментах В. Ф. Петренко. Несмотря на то что данный фактор представлен всего двумя шкалами, как и в экспериментах исследователя, в нашем материале он также имеет четкое содержание и ясную семантику (иное название фактора – «частота встречаемости»).

Осгудовский фактор «Активность» в нашем материале объяснял небольшой процент вариации (фактор 7). Отметим отрицательную корреляцию с данным фактором шкалы «Сила», что указывает на связь пассивности и бессилия.

Таким образом, в целом в нашем материале выделяются основные шкалы, выделенные В. Ф. Петренко, однако конкретный состав шкал, представляющих факторы, может отличаться. Отметим ведущую роль фактора оценки, отмеченную в экспериментах Ч. Осгуда и В. Ф. Петренко, а также полное совпадение структуры выделенного В. Ф. Петренко фактора «Обычность».

Продолжая исследовать скрытую структуру в данных, перейдем к анализу ассоциативного значения изучаемых слов, выраженного в оценках по шкалам СД, с использованием метода главных компонент. Сочетание пакетов factextra (Kas-

смерть, ненавидеть, обман, встретит, плохо.

По третьей компоненте противопоставляются стимулы *смысл, обман, настоящий², ненавидеть, плохо*, с одной стороны, *место, дом1, богатый, огонь, дом2, лес* – с другой. Положительно коррелируют с данной компонентой оценки по шкалам СД «**Опасность**», «**Живость**», «**Единность**», «**Яркость**»; отрицательно – оценки по шкалам «**Веселость**», «**Быстрота**», «**Приятность**».

Таким образом, РСА позволяет установить основные источники различий в семантике слов, учитывая при этом связь признаков (оценок по шкалам СД).

Дальнейшим шагом нашего исследования является проведение кластерного анализа с целью обнаружения групп ассоциативных рядов, близких друг другу исходя из оценок по исследуемым шкалам СД. Существует множество методов кластерного анализа. В нашей работе мы используем впервые в семантических исследованиях метод кластеризации данных, основанный не на «сырых» значениях, а на выделенных на первом этапе компонентах (Hierarchical Clustering on Principal Components, НСРС). Данный метод реализован в небольшом числе библиотек. В нашем исследовании мы использовали его реализацию в библиотеке FactoMineR. Данный метод комбинирует три основных метода анализа многомерных данных: РСА, иерархическую кластеризацию и алгоритм кластеризации K-means. Проведение кластеризации не на сырых данных, а на результатах РСА позволяет получить более устойчивое кластерное решение (Husson, 2010).

Выбор числа кластеров в данном методе основан на критерии loss of inertia. Наименьшее значение данного критерия наблюдается при переходе от

5-кластерного решения к 4-кластерному; следовательно, как оптимальное было выбрано 4-кластерное решение. Результат кластеризации представлен на рис. 5.

Следующим этапом анализа стала интерпретация различий между кластерами. Для проверки существования различий в оценках слов по шкалам СД проводился однофакторный дисперсионный анализ (*one-way ANOVA*) с *последующим* множественным анализом (использовался статистический тест Tukey) для выявления парных различий.

Тест *one-way ANOVA* подтвердил наличие различий по кластерам в средних значениях оценок по всем шкалам СД (значения *p-values* ниже 0,05). Следовательно, все переменные были использованы для дальнейшей интерпретации.

С использованием библиотеки autoEDA был осуществлен отбор самых информативных признаков (шкал) СД для разделения кластеров (таблица 2) (применялся критерий Kolmogorov Smirnov distance measure).

Для оценки различий по характеристикам респондентов в кластерах мы использовали непараметрический аналог однофакторного дисперсионного анализа – критерий *Крускала – Уоллиса с последующим множественным анализом*. Статистически значимые различия между кластерами были обнаружены для следующих устойчивых характеристик респондентов (в порядке убывания значения *p*): доброжелательность, открытость опыту ($p < 0,0005$), добросовестность ($p < 0,005$), экстраверсия ($p < 0,05$) и для следующих эмоциональных состояний: радость ($p < 0,0005$), презрение, интерес, ПЭМ ($p < 0,005$), стыд ($p < 0,05$).

Далее слова-стимулы на основании частоты встречаемости в кластерах были условно распределены нами на три группы.

Рисунок 5. Визуализация результатов анализа с использованием метода иерархической кластеризации на главных компонентах (НСРС)

Figure 5. Visualization of the analysis results using the hierarchical clustering method on the main components (НСРС)

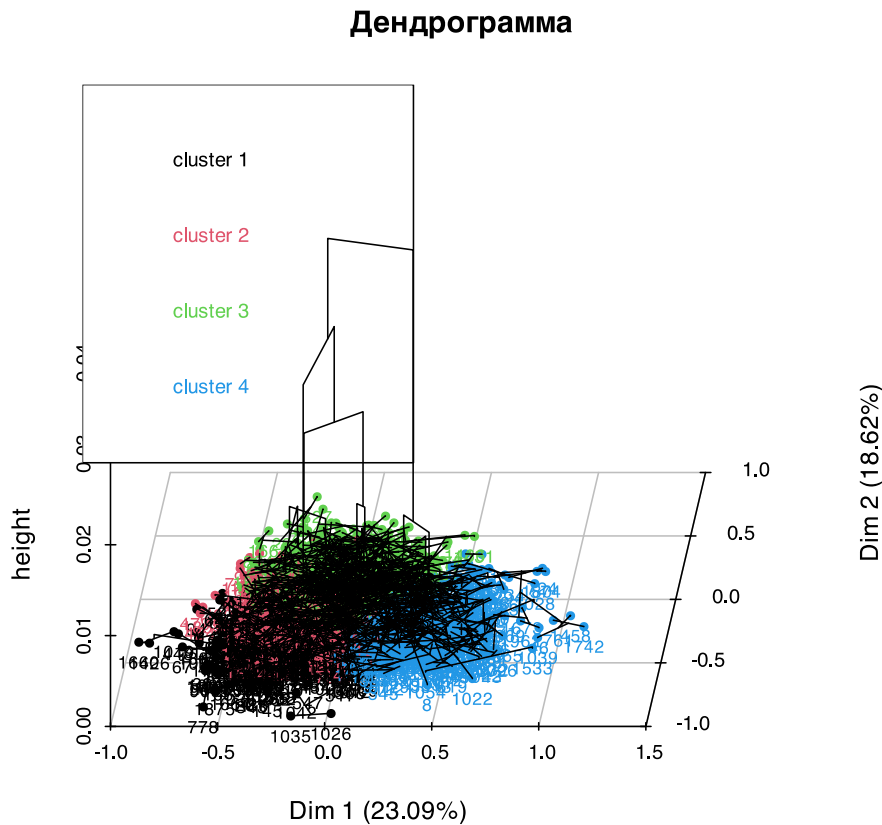


Таблица 2. Вклад шкал в кластерное решение
Table 2. Contribution of SD scales to the cluster solution

Шкала СД	Вклад шкалы в разделение на кластеры, в процентах	Автоматическая оценка вклада шкалы в разделение на кластеры
Приятность	86	Высокий
Полезность	76	Высокий
Редкость	71	Высокий
Упорядоченность	62	Высокий
Живость	56	Высокий
Яркость	51	Высокий
Веселость	47	Средний
Опасность	41	Средний
Сложность	38	Средний

1. Стимулы с одним основным значением, выраженным в оценках по шкалам СД, то есть преимущественно входящие в один из кластеров (относительная частота встречаемости в одном из класте-

ров выше **0,5**, то есть более 50% всех ассоциативных рядов данного стимула входят в один из 4-х выделенных нами кластеров, в других кластерах – ниже **0,3**): *бог, враг, глупый, делать, добро1, душа, любить,*

мама, место, настоящий1, настоящий2, ненавидеть, обман, огонь, плохо, простой, путь1, результат, свет, связь, смерть, счастье1, счастье2, труд, учиться.

2. Стимулы с двумя значениями были разделены нами на две подгруппы: а) стимулы с двумя основными значениями (частота встречаемости в одном из кластеров не ниже **0,3**, но не более **0,6**, разница между частотами встречаемости в двух доминирующих кластерах менее **0,25**): *бабушка, богатый, встретить, глаз, добро2, дом1, дорогой, друг1, друг2, думать, лес, мир2, много, обещать, путь2, семья2, смысл*; б) стимулы с двумя значениями, одно из которых является доминирующим (частота встречаемости стимула в каждом из кластеров не ниже **0,25**, но не выше **0,6**; разница между ними более **0,25**): *семья1, радость, дом2, маленький*.

3. Стимулы с несколькими значениями/диффузным значением разделены нами на две подгруппы: а) с тремя основными значениями (частота встречаемости стимула в каждом из кластеров не ниже **0,2**): *дерево, жизнь1, мир1, новый, помощь, увидеть, хотеть1, хотеть2*; б) стимулы с диффузным значением (максимальная частота встречаемости в каждом кластере не превышает **0,4**, другие значения не превышают **0,25**): *жизнь2, история*.

Отметим, что стимулы, ассоциации на которые респонденты давали два раза, оказываются, как правило, в одной группе (*настоящий, счастье, друг, хотеть*) либо в подгруппах одной группы (*жизнь, семья, дом*), что свидетельствует об устойчивости их ассоциативного значения, несмотря на различное лексическое наполнение рядов. Стимулы *добро, путь, мир* оказались в разных группах, что, возможно, обусловлено многозначностью указанных стимулов.

Анализ распределения респондентов по кластерам показал, только 3 респондента заметно «тяготеют» к одному из кластеров (свыше 50% их ассоциативных реакций входит в один кластер). 19 респондентов имеют тенденцию к предпо-

чтению одного из кластеров (частотность встречаемости их ассоциативных реакций в одном из кластеров от 0,4 до 0,5).

Рассмотрим характеристики каждого из кластеров. **Первый кластер** характеризуется **низкими** значениями по шкалам «**Приятность**», «**Полезность**», «**Веселость**», «**Живость**», «**Быстрота**», «**Упорядоченность**», «**Устойчивость**», «**Редкость**», «**Яркость**», «**Конкретность**», **высокими** значениями – «**Пассивность**» и «**Опасность**», «**Сила**». Среди респондентов, чьи ответы отнесены в данный кластер, преобладают респонденты с самыми низкими оценками по *доброжелательности*, экстраверсии, *добросовестности*, интересу, радости, самые высокие оценки по шкалам «**Стыд**», «**Презрение**». Мы условно назвали данный кластер **негативным**.

Как в основной входят в данный кластер лексемы *враг, глупый, ненавидеть, обман, плохо, смерть*. В качестве одного из частотных вхождение в данный кластер зафиксировано для стимулов *думать, начало, обещать, увидеть*. От 20 до 30% от общего числа рядов к стимулам *жизнь2, история, огонь, смысл* входят в данный кластер, а также от 10 до 20% ассоциативных реакций к стимулам *бог, встретить, душа, любить, учиться, хотеть1*.

Для респондентов, ассоциативные ряды которых составляют **второй кластер**, характерны самые высокие значения *доброжелательности*, экстраверсии, *добросовестности*, открытости опыту, интереса, радости и интегрального показателя позитивных эмоций, низкие оценки по шкалам «**Стыд**» и «**Презрение**». Таким образом, данный кластер можно охарактеризовать как противоположный первому по характеристикам респондентов.

Ассоциативные ряды, входящие во **второй кластер**, характеризуются самыми высокими значениями по шкале «**Веселость**», «**Кратковременность**», «**Сила**», умеренно высокими оценками по шкале «**Приятность**», низкими значениями по шкале «**Яркость**»,

«**Опасность**», самыми низкими значениями по шкале «**Живость**», «**Единость**», «**Конкретность**», «**Крупность**».

Как в основной входят в данный кластер стимулы *место, огонь, путь1, свет*. В качестве доминирующего (при наличии других частотных значений) данный кластер выступает для следующих стимулов: *бабушка, встретить, глаз, дом1, дом2, лес, маленький, хотеть2*, является одним частотных для стимулов *богатый, дерево, жизнь2, мир1, новый, путь2, радость, увидеть, учиться, хотеть2*. Данный кластер можно с условностью назвать *позитивная сила*.

Для респондентов, ассоциативные ряды которых составляют **третий** кластер, характерны умеренно высокие значения по шкалам доброжелательности, экстраверсии, добросовестности, открытость опыту, средние значения интереса, радости, низкие оценки по шкале «Презрение».

Ассоциативные ряды, входящие в **третий** кластер, характеризуются самыми высокими значениями по шкалам «**Приятность**», «**Живость**», «**Яркость**», «**Упорядоченность**», «**Единость**», «**Редкость**», «**Единичность**», самыми низкими значениями по шкале «**Сложность**». Как в основной (без иных высокочастотных кластеров) в данный кластер входят стимулы *бог, добро1, душа, любить, мама, настоящий1, настоящий2, русский, счастье1, счастье2*. В качестве доминирующего (при наличии других частотных) данный кластер выступает для следующих стимулов: *добро2, дорогой, друг1, история, мир2, помощь, радость, семья1, семья2, смысл, увидеть*, в качестве одного из частотных – для стимулов *встретить, глаз, дерево, дом1, друг2, жизнь1, лес, маленький, мир1, много, хотеть1, хотеть2*. Данный кластер можно назвать *приятная витальность*.

Для респондентов, ассоциативные ряды которых составляют **четвертый** кластер, характерны умеренно низкие значения по шкале доброжелательности, экстра-

версии, добросовестности, низкие значения открытости опыту, средние значения интереса, радости, презрения, умеренно низкие оценки по интегральному показателю позитивных эмоций.

Ассоциативные ряды, входящие в **четвертый** кластер, характеризуются невысокими значениями по шкале «**Приятность**», «**Живость**», самыми высокими – по шкале «**Полезность**», «**Быстрота**», «**Упорядоченность**», «**Устойчивость**», «**Конкретность**», «**Сложность**», «**Крупность**», низкими – по шкале «**Кратковременность**», «**Сила**», «**Опасность**». Назовем условно данный кластер *полезная упорядоченность*.

Как в основной (без иных высокочастотных кластеров) в данный кластер входят стимулы *богатый, делать, простой, результат, связь, труд, учиться*. В качестве доминирующего (при наличии других частотных) данный кластер выступает для следующих стимулов: *друг2, думать* (интересно, что второй по частотности вхождения кластер этого стимула – первый), *жизнь1, жизнь2, мир1, много, новый, путь2, хотеть1*, как один из частотных – для стимулов *дерево, добро2, дом2, дорогой, друг1, мир2, начало* (второй по частотности вхождения кластер этого стимула – первый), *обещать, помощь, семья2, смысл, хотеть2*.

Таким образом, как показал проведенный анализ, около половины слов-стимулов имеют общее – по оценкам семантического дифференциала – для большинства респондентов значение (то есть ассоциативные ряды к данному слову-стимулу входят преимущественно в один кластер). К таким словам относятся преимущественно стимулы из первого, *негативного*, кластера: *враг* (83% от общего числа всех реакций к стимулу входят в кластер 1), *глупый* (60% от общего числа всех реакций к данному стимулу входят в кластер 1, хотя более 20% ассоциативных рядов данного стимула входят в кластер 3 *приятная витальность*, что может свидетельствовать о связи индивидуальной

семантики этого слова и психологических характеристик респондентов), *ненавидеть* (84%), *обман* (93% всех ассоциативных рядов к данному стимулу входит в один кластер, что говорит о минимальных индивидуальных различиях в семантике данного стимула), *плохо* (76%), *смерть* (78%). Таким образом, слова из негативной семантикой, как правило, характеризуются невысокой степенью выраженности индивидуальных различий (по крайней мере применительно к анализируемому нами датасету), хотя для ряда слов из данного кластера отмечено наличие подобных различий.

Среди стимулов, входящих в другие кластеры, одни из самых «однозначных» – *свет* (80% ассоциативных рядов данного стимула входит в кластер 2 *позитивная сила, настоящий*^{1,2} (свыше 70% ассоциативных рядов данного стимула входит в кластер 3 *приятная витальность*), *бог, душа* (70 и 76% вхождений в кластер 3 *приятная витальность* соответственно), *простой, результат, труд* (69, 86, 83% ассоциативных рядов данных стимулов соответственно входят в кластер 4 *полезная упорядоченность*).

Представляется весьма примечательным тот факт, что ассоциативные ряды некоторых слов-стимулов входят с близкой частотностью в противоположные по семантике кластеры (1 и 4, 1 и 3), например, ассоциативные ряды стимулов *думать, история, начало, обещать, увидеть*, что свидетельствует о значительных различиях в индивидуальной семантике указанных слов в языковом сознании наших респондентов, обусловленных в том числе различиями в психологических характеристиках и эмоциональных состояниях последних.

Выводы

Несмотря на то что существование различий в индивидуальной семантике слова признается многими учеными, исследование подобных различий представляет собой непростую задачу. Комплексность и сложность данной задачи обусловлена несколькими факторами, и прежде

всего трудностями объективации критериев для описания семантических параметров (компонентов значения). В данной работе нами предложен метод получения оценок ассоциативного значения слова по шкалам семантического дифференциала с использованием языковой модели на основании специально созданного датасета – набора ассоциативных реакций на слова-стимулы, важные для русского языкового сознания. Указанный метод позволяет квантифицировать и описывать различия в индивидуальной семантике слов. В совокупности с рядом продвинутых методов анализа данных, таких как кластерный анализ на главных компонентах, данный метод позволил нам разделить исследуемые слова-стимулы на группы в зависимости от выраженности индивидуальных различий в их семантике.

Нами также обнаружена связь индивидуального значения слова, измеренного по шкалам семантического дифференциала, с психологическими характеристиками респондентов: выделенные нами кластеры ассоциативных рядов различаются как по психологическим характеристикам респондентов, так и по оценкам по шкалам семантического дифференциала. При этом было установлено, что для разных стимулов степень влияния психологических характеристик и состояний респондента на особенности их семантического представления в языковом сознании различается.

Метод оценивания семантики языковых единиц по отношению к бинарным оппозициям, основанный на использовании данных дистрибутивных семантических моделей, ранее применялся в основном по отношению к текстам. В данной работе мы впервые применили данный метод по отношению к ассоциативным рядам, что позволило нам по-новому подойти к проблеме изучения ассоциативного значения слова в индивидуальном языковом сознании. Предложенная методика ни в коей мере не заменяет традиционные методы исследования психологически реального значения слова (ассоциативный экс-

перимент, семантический дифференциал), но может служить их эффективным дополнением.

Предложенный метод может быть использован с различными вариантами семантического дифференциала и – шире – с различными бинарными оппозициями, а также с разными языковыми моделями.

Кроме ассоциативных рядов, предложенная методика может быть применена к текстам для исследования индивидуальной семантики слов, воплощенной в дискурсах разных типов.

Как и любое другое, наше исследование имеет **ограничения**. Во-первых, полученные результаты непосредственно связаны с используемой языковой моделью, ее настройками, составом корпуса, на котором она обучалась, и, соответственно, статистикой совместной встречаемости слов в корпусе. В данной работе была использована только одна языковая модель, и хотя ее выбор был нами обоснован, применение других языковых моделей и сравнение полученных результатов представляется одним из перспективных направлений дальнейших исследований. Также ограничения текущего исследования связаны со сравнительно небольшим объемом датасета, на котором оно проводилось. Расширение датасета – как по линии увеличения числа респондентов, в том числе за счет расширения состав социальных групп, участвующих в опросах, так и по линии увеличения числа слов-стимулов – позволит в дальнейшем уточнить полученные нами результаты и усовершенствовать предложенную методику анализа.

Список литературы

Глухов В. П. Основы психолингвистики: учебное пособие для студентов педвузов. М.: АСТ: Астрель, 2005. 351 с.

Горошко Е. И. Интегративная модель свободного ассоциативного эксперимента. М.: ИЯ РАН; Харьков: РА-Каравелла, 2001. 320 с.

Зуева Е. А. Эмоции как объект лингвистических исследований // Иностранные языки в профессиональном образовании: лингвометодический контекст: материалы межвуз.

науч.-практ. конф., Белгород, 17-18 мая 2006 г. / Белгор. ун-т потреб. кооперации. Белгород, 2006. С. 148-154.

Калугин А. Ю. Психометрика русскоязычной версии Big Five Inventory-2 / Калугин А. Ю., Щebetенко С. А., Мишкевич А. М., Сото К. Д., Джон О. // Психология. Журнал Высшей школы экономики. 2021. № 18 (1). С. 7-33.

Караулов Ю. Н., Коробова М. М. Индивидуальный ассоциативный словарь // Вопросы языкознания. 1993. № 5. С. 5–15.

Курганова Н. И. Ассоциативный эксперимент как метод исследования значения живого слова // Вопросы психолингвистики. 2019. № 3 (41). С. 24-37.

Литвинова Т. А., Заварзина В. А., Любова С. Г. База данных ассоциативных реакций, содержащая информацию о клавиатурном поведении респондентов // Известия Воронежского государственного педагогического университета. 2022. № 4 (297). С. 240–249.

Литвинова Т. А. Исследование значения слова в индивидуальном языковом сознании с использованием метода семантической проекции / Литвинова Т. А., Котлярова Е. С., Любова С. Г., Паничева П. В. // Russian Linguistic Bulletin. 2023. № 12 (48). URL: <https://rulb.org/archive/12-48-2023-december/10.18454/RULB.2023.48.50> (дата обращения: 29.01.2024). <https://doi.org/10.18454/RULB.2023.48.50>

Литвинова Т. А., Любова С. Г. Разметка ассоциативных реакций по типам отношений «стимул-ассоциат» как этап создания аннотированного многокомпонентного корпуса ассоциативных реакций индивида // Проблемы изучения живого русского слова на рубеже тысячелетий: материалы XI Междунар. науч. конф. (г. Воронеж, 27–28 октября 2023 г.). Воронеж: Воронежский государственный педагогический университет, 2023. С. 49–58.

Новиков А. Л., Новикова И. А. Метод семантического дифференциала: теоретические основы и практика применения в лингвистических и психологических исследованиях // Вестник РУДН. Серия: Теория языка. Семиотика. Семантика. 2011. № 3. С. 63–71.

Петренко В. Ф. Основы психосемантики. 2-е изд., доп. СПб.: Питер, 2005. 480 с.

Серкин В. П. Методы психологии субъективной семантики и психосемантики: учебное пособие для вузов. М.: ПЧЕЛА, 2008. 382 с.

Сикевич З. В. Метод семантического дифференциала в социологическом исследовании (опыт применения) // Вестник СПбГУ. Серия 12. Социология. 2016. Вып. 3. С. 118–128.

Степыкин Н. И. Исследование динамики ментального лексикона по данным свободного ассоциативного эксперимента / Степыкин Н. И., Багана Ж., Слободова К. Н., Фуникова С. В. // Научный результат. Вопросы теоретической и прикладной лингвистики. 2023. № 2. С. 19–33. DOI: 10.18413/2313-8912-2023-9-2-0-2

Уфимцева Н. В. Образ мира русских: системность и содержание // Язык и культура. 2009. № 4 (8). С. 98–110.

Черкасова Г. А. Русский сопоставительный ассоциативный словарь. М.: ИЯ РАН, 2008. 194 с.

Abdi H., Williams L. J. Principal component analysis // Wiley Interdiscip. Rev. 2010. № 2. P. 433–459.

Arseniev-Koehler A., Foster J. G. Sociolinguistic Properties of Word Embeddings // Dehghani M., Boyd R. L. Handbook of Language Analysis in Psychology. New York: Guilford Press, 2022. P. 464–477.

Azucar D., Marengo D., Settanni M. Predicting the Big 5 personality traits from digital footprints on social media: A meta-analysis // Personality and Individual Differences. 2018. Vol. 124. P. 150–159.

Caliskan A., Bryson J. J., Narayanan A. Semantics derived automatically from language corpora contain human-like biases // Science. 2017. Vol. 356 (6334). P. 183–186.

Daenekindt S., Schaap J. Using word embedding models to capture changing media discourses: a study on the role of legitimacy, gender and genre in 24,000 music reviews, 1999–2021 // Journal of Computational Social Science. 2022. № 5. P. 1615–1636.

Eberhard C., Owens W. A. Word Association as a Function of Biodata Subgrouping // Developmental Psychology. 1975. Vol. 11 (2). P. 159–164.

Ellis N. C. Essentials of a Theory of Language Cognition // The Modern Language Journal. 2019. Vol. 103. P. 39–60.

Garg N., Schiebinger L., Jurafsky D., Zou J. Word embeddings quantify 100 years of gender and ethnic stereotypes // Proceedings of the National Academy of sciences of the United States of America. 2018. Vol. 115, iss. 16. P. E3635–E3644.

Grand G. et al. Semantic Projection Recovers Rich Human Knowledge of Multiple Object Features from Word Embeddings // Nat Hum Behav. 2022. Vol. 6 (7). P. 975–987.

Greenwald A. G., McGhee D. E., Schwartz J. L. K. Measuring Individual Differences in Implicit Cognition. The Implicit Association Test // Journal of Personality and Social Psychology. 1998. № 74. P. 1464–1480.

Hollis G., Westbury C. The principals of meaning: Extracting semantic dimensions from co-occurrence models of semantics // Psychonomic Bulletin & Review. 2016. Vol. 23. P. 1744–1756.

Husson F., Josse J., Pages J. Principal component methods-hierarchical clustering-partitional clustering: Why would we need to choose for visualizing data // Appl. Math. Dep. 2010. Vol. 17. P. 1–17.

Innes J. M. The relationship of word-association commonality response set to cognitive and personality variables // Br J Psychol. 1972. Vol. 63 (3). P. 421–428.

Jordan M. C., Giallanza T., Ellis C. T. et al. Context Matters: Recovering Human Semantic Structure from Machine Learning Analysis of Large-Scale Text Corpora // Cognitive science. 2022. № 46 (2), Article e13085.

Isen A. M. The influence of positive affect on the unusualness of word associations / Isen A. M., Johnson M. M., Mertz E., Robinson G. F. // J Pers Soc Psychol. 1985. № 48 (6). P. 1413–1426.

Kassambara A., Mundt F. Factoextra: Extract and visualize the results of multivariate data analyses // R Package Version. 2017. № 1. P. 337–354.

Kozlowski A. C., Taddy M., Evans J. A. The Geometry of Culture: Analyzing the Meanings of Class through Word Embeddings // American Sociological Review. 2019. № 84 (5). P. 905–949.

Lê S., Josse J., Husson F. FactoMineR: An R package for multivariate analysis // J. Stat. Softw. 2008. № 25. P. 1–18.

Lenci A. Distributional Models of Word Meaning // Annual Review of Linguistics. 2018. № 4. P. 151–71.

Litvinova T. A. Mapping the field of word association research using text mining approach / Litvinova T. A., Zavarzina V. A., Kotlyarova E. S., Lyubova S. G. // 5th International Conference on Information Technology and Computer Communications (ITCC 2023), June 15–17, 2023, Tianjin, China. ACM, New York, NY, USA. P. 90–98.

Litvinova T. RusIdiolect: A New Resource for Authorship Studies // Lecture Notes in Networks and Systems. 2021. Vol. 186. P. 14-23.

Lukavsky J. Subjective valence of the test words an enhancement of the word association test // Ceskoslovenska Psychologie. 2004. № 48. P. 203-214.

Matsui A., Ferrara E. Word Embedding for Social Sciences: An Interdisciplinary Survey. URL: <https://arxiv.org/pdf/2207.03086.pdf> (accessed 29.01.2024).

Merseal H. M. Free association ability distinguishes highly creative artists from scientists: Findings from the Big-C Project / Merseal H. M., Luchini S., Kenett Y. N., Knudsen K., Bilder R. M., Beaty R. E. // Psychology of Aesthetics, Creativity, and the Arts. 2023. Advance online publication. <https://doi.org/10.1037/aca0000545>

Merten T. Word association responses and psychoticism // Personality and Individual Differences. 1993. Vol. 14 (6). P. 837–839.

Mikolov T. Distributed Representations of Words and Phrases and Their Compositionality / Mikolov T., Sutskever I., Chen K., Corrado G. S., Dean J. // Proceedings of the 26th International Conference on Neural Information Processing Systems, vol. 2, Curran Associates Inc., NY: USA, 2013. P. 3111–3119.

Osgood C. E., Suci G. J., Tannenbaum P. H. The Measurement of Meaning. Chicago: University of Illinois Press, 1957. 342 p.

Pereira F., Gershman S., Ritter S., Botvinick M. A comparative evaluation of off-the-shelf distributed semantic representations for modelling behavioural data // Cognitive Neuropsychology. 2016. № 33. P. 175–190.

Schmolling P. Cognitive Responses under Stress: The Effect of Examination Stress on Commonality of Verbal Associations // Psychological Reports. 1978. № 43 (3_suppl). P. 1131-1138.

Schröder T., Hoey J., Rogers K. B. Modeling Dynamic Identities and Uncertainty in Social Interactions: Bayesian Affect Control Theory //

American Sociological Review. 2016. № 81 (4). P. 828–855.

Stoltz D. S., Taylor M. A. Concept Mover's Distance: measuring concept engagement via word embeddings in texts // Journal of Computational Social Science. 2019. Vol. 2. P. 293–313.

Stoltz D. S., Taylor M. A. text2map: R tools for text matrices // Journal of Open Source Software. 2022. № 7 (72). P. 3741.

Taylor M. A., Stoltz D. S. Integrating Semantic Directions with Concept Mover's Distance to Measure Binary Concept Engagement // Journal of Computational Social Science. 2021. Vol. 4. P. 231–242.

Utsumi A. A Neurobiologically Motivated Analysis of Distributional Semantic Models // Proceedings of the 40th Annual Conference of the Cognitive Science Society (CogSci2018). Madison, WI, USA. P. 1147-1152.

Voyer A. From Strange to Normal: Computational Approaches to Examining Immigrant Incorporation Through Shifts in the Mainstream / Voyer A., Kline Z. D., Danton M., Volkova T. // Sociological Methods & Research. 2022. № 51 (4). P. 1540–1579.

Wulff D. U. Data From the MySWOW Proof-of-Concept Study: Linking Individual Semantic Networks and Cognitive Performance / Wulff D. U., Aeschbach S., De Deyne S., Mata R. // Journal of Open Psychology Data. 2022. № 10 (5). P. 1–8.

References

Glukhov, V. P. (2005). *Osnovy psikholingvistiki: uchebnoe posobie dlya studentov pedvuzov* [Fundamentals of Psycholinguistics: a Textbook for Students of Pedagogical Universities], ACT Astrel' Publishing, Moscow, Russia. (In Russian)

Goroshko, E. I. (2001). *Integrativnaja model' svobodnogo assotsiativnogo eksperimenta* [Integrative Model of Free Associative Experiment], Institute of Linguistics RAS, Moscow, Russia; Izdatelskaia gruppa "RA-Karavella", Kharkov, Ukraine. (In Russian)

Zueva, E. A. (2006). Emotions as an object of linguistic research, *Proc. of the Interuniversity scientific and practical conference "Foreign Languages in Vocational Education: Linguistic and Methodological Context"*, Belgorod University of Cooperation, Economics & Law, Belgorod, Russia, 148-154. (In Russian)

Kalugin, A. Yu., Shchebetenko, S. A., Mishkevich, A. M., Soto, Ch. D. and John, O. (2021). Psychometric Properties of the Russian

Version of the Big Five Inventory–2, *Psikhologiya. Zhurnal Vysshej shkoly ekonomiki*, 18 (1), 7–33. <https://doi.org/10.17323/1813-8918-2021-1-7-33> (In Russian)

Karaulov, Yu. N. and Korobova, M. M. (1993). Individual associative dictionary, *Voprosy yazykoznanija*, 5, 5–15. (In Russian)

Kurganova, N. I. (2019). Association experiment as a method for studying the meaning of a living word, *Voprosy psiholingvistiki*, 3 (41), 24–37. <https://doi.org/10.30982/2077-5911-2019-41-3-24-37> (In Russian)

Litvinova, T. A., Zavarzina, V. A. and Lyubova, S. G. (2022). Database of associative reactions containing information about the keyboard behavior of respondents, *Izvestiya Voronezhskogo gosudarstvennogo pedagogicheskogo universiteta*, 4 (297), 240–249. https://doi.org/10.47438/2309-7078_2022_4_240 (In Russian)

Litvinova, T. A., Kotlyarova, E. S., Lyubova, S. G. and Panicheva, P. V. (2023). The study of the meaning of a word in an individual linguistic consciousness using the method of semantic projection, *Russian Linguistic Bulletin*, [Electronic], 12 (48), available at: <https://rulb.org/en/archive/12-48-2023-december/10.18454/RULB.2023.48.50> (Accessed 8 March 2024). <https://doi.org/10.18454/RULB.2023.48.50> (In Russian)

Litvinova, T. A. and Lyubova, S. G. (2023). Marking up associative reactions by types of "stimulus-associate" relationships as a stage in creating an annotated multicomponent corpus of associative reactions of an individual, *Proc. of the XI International Scientific Conference "Problems of studying the living Russian word at the turn of the millennium"*, Voronezh State Pedagogical University, Voronezh, Russia, 49–58. (In Russian)

Novikov, A. L. and Novikova, I. A. (2011). The method of semantic differential: theoretical foundations and practice of application in linguistic and psychological research, *Vestnik RUDN. Seriya: Teoriya yazyka. Semiotika. Semantika*, 3, 63–71. (In Russian)

Petrenko, V. F. (2005). *Osnovy psikhosemantiki* [Fundamentals of psychosemantics], Piter, St. Petersburg, Russia. (In Russian)

Serkin, V. P. (2008). *Metody psikhologii subyektivnoi semantiki i psikhosemantiki* [Methods of psychology of subjective semantics and

psychosemantics], PCHELA, Moscow, Russia. (In Russian)

Sikevich, Z. V. (2016). The method of semantic differential in sociological research (application experience), *Vestnik SPbGU. Seriya 12. Sotsiologiya*, 3, 118–128. <https://doi.org/10.21638/11701/spbu12.2016.309> (In Russian)

Stepykin, N. I., Bagana, Zh., Slobodova, K. N. and Funikova, S. V. (2023). Investigation of the dynamics of the mental lexicon according to the data of a free associative experiment, *Nauchny rezultat. Voprosy teoreticheskoi i prikladnoi lingvistiki*, 2, 19–33. <https://doi.org/10.18413/2313-8912-2023-9-2-0-2> (In Russian)

Ufimtseva, N. V. (2009). Russian image of the world: consistency and content, *Yazyk i kultura*, 4 (8), 98–110. (In Russian)

Cherkasova, G. A. (2008). *Russkiy sopostavitelnyy associativnyy slovar* [Russian comparative associative dictionary], Institute of Linguistics RAS, Moscow, Russia. (In Russian)

Abdi, H. and Williams, L. J. (2010). Principal component analysis, *Wiley Interdiscip. Rev.*, 2 (4), 433–459. <https://doi.org/10.1002/wics.101> (In English)

Arseniev-Koehler, A. and Foster, J. G. (2022). Sociolinguistic Properties of Word Embeddings, in Dehghani, M. and Boyd, R. L. (eds.), *Handbook of Language Analysis in Psychology*, Guilford Press, New York, USA, 464–477. (In English)

Azucar, D., Marengo, D. and Settanni, M. (2018). Predicting the Big 5 personality traits from digital footprints on social media: A meta-analysis, *Personality and Individual Differences*, 124, 150–159. <https://doi.org/10.1016/j.paid.2017.12.018> (In English)

Caliskan, A., Bryson, J. J. and Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases, *Science*, 356 (6334), 183–186. <https://doi.org/10.1126/science.aal4230> (In English)

Daenekindt, S. and Schaap, J. (2022). Using word embedding models to capture changing media discourses: a study on the role of legitimacy, gender and genre in 24,000 music reviews, 1999–2021, *Journal of Computational Social Science*, 5, 1615–1636.

<https://doi.org/10.1007/s42001-022-00182-8>

(In English)

Eberhard, C. and Owens, W. A. (1975). Word Association as a Function of Biodata Subgrouping, *Developmental Psychology*, 11 (2), 159–164. (In English)

Ellis, N. C. (2019). Essentials of a Theory of Language Cognition, *The Modern Language Journal*, 103, 39–60.

<https://doi.org/10.1111/modl.12532> (In English)

Garg, N., Schiebinger, L., Jurafsky, D. and Zou, J. (2018). Word embeddings quantify 100 years of gender and ethnic stereotypes, *Proceedings of the National Academy of sciences of the United States of America*, 115 (16), E3635–E3644. <https://doi.org/10.1073/pnas.1720347115> (In English)

Grand, G., Blank, I. A., Pereira, F. and Fedorenko, E. (2022). Semantic Projection Recovers Rich Human Knowledge of Multiple Object Features from Word Embeddings, *Nat Hum Behav*, 6 (7), 975–987. <https://doi.org/10.1038/s41562-022-01316-8> (In English)

Greenwald, A. G., McGhee, D. E. and Schwartz, J. L. K. (1998). Measuring Individual Differences in Implicit Cognition. The Implicit Association Test, *Journal of Personality and Social Psychology*, 74 (6), 1464–1480. <https://doi.org/10.1037/0022-3514.74.6.1464> (In English)

Hollis, G. and Westbury, C. (2016). The principals of meaning: Extracting semantic dimensions from co-occurrence models of semantics, *Psychonomic Bulletin & Review*, 23, 1744–1756. <https://doi.org/10.3758/s13423-016-1053-2> (In English)

Husson, F., Josse, J. and Pages, J. (2010). Principal component methods-hierarchical clustering-partitional clustering: Why would we need to choose for visualizing data, *Appl. Math. Dep.*, 17, 1–17. (In English)

Innes, J. M. (1972). The relationship of word-association commonality response set to cognitive and personality variables, *Br J Psychol*, 63 (3), 421–428. (In English)

Jordan, M. C., Giallanza, T., Ellis, C. T., Beckage, N. M. and Cohen, J. D. (2022). Context Matters: Recovering Human Semantic Structure from Machine Learning Analysis of Large-Scale Text Corpora, *Cognitive science*, 46 (2), e13085. <https://doi.org/10.1111/cogs.13085> (In English)

Isen, A. M., Johnson, M. M., Mertz, E. and Robinson, G. F. (1985). The influence of positive affect on the unusualness of word associations, *J Pers Soc Psychol*, 48 (6), 1413–1426. (In English)

Kassambara, A. and Mundt, F. (2017). Factoextra: Extract and visualize the results of multivariate data analyses, *R Package Version*, 1, 337–354. (In English)

Kozlowski, A. C., Taddy, M. and Evans, J. A. (2019). The Geometry of Culture: Analyzing the Meanings of Class through Word Embeddings, *American Sociological Review*, 84 (5), 905–949.

<https://doi.org/10.1177/0003122419877135>

(In English)

Lê, S., Josse, J. and Husson, F. (2008). FactoMineR: An R package for multivariate analysis, *J. Stat. Softw.*, 25, 1–18. <https://doi.org/10.18637/jss.v025.i01> (In English)

Lenci, A. (2018). Distributional Models of Word Meaning, *Annual Review of Linguistics*, 4, 151–71. <https://doi.org/10.1146/annurev-linguistics-030514-125254> (In English)

Litvinova, T. A., Zavarzina, V. A., Kotlyarova, E. S. and Lyubova, S. G. (2023). Mapping the field of word association research using text mining approach, *Proceedings of 5th International Conference on Information Technology and Computer Communications (ITCC 2023), Tianjin, China*, 90–98.

<https://doi.org/10.1145/3606843.3606858>

(In English)

Litvinova, T. A. (2021). RusIdiolect: A New Resource for Authorship Studies, *Lecture Notes in Networks and Systems*, 186, 14–23. (In English)

Lukavsky, J. (2004). Subjective valence of the test words an enhancement of the word association test, *Ceskoslovenska Psychologie*, 48, 203–214. (In Czech)

Matsui, A., Ferrara, E. (2018). Word Embedding for Social Sciences: An Interdisciplinary Survey, available at: <https://arxiv.org/pdf/2207.03086.pdf> (Accessed 8 March 2024). (In English)

Merseal, H. M., Luchini, S., Kenett, Y. N., Knudsen, K., Bilder, R. M. and Beaty, R. E. (2023). Free association ability distinguishes highly creative artists from scientists: Findings from the Big-C Project, *Psychology of Aesthetics, Creativity, and the Arts*. Advance online publication, available at:

<https://psycnet.apa.org/doi/10.1037/aca0000545>

(Accessed 8 March 2024). (In English)

Merten, T. (1993). Word association responses and psychoticism, *Personality and Individual Differences*, 14 (6), 837–839. (In English)

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S. and Dean, J. (2013). Distributed Representations of Words and Phrases and Their Compositionality, *Proceedings of the 26th International Conference on Neural Information Processing Systems*, vol. 2, Nevada, USA, 3111–3119. <https://doi.org/10.5555/2999792.2999959> (In English)

Osgood, C. E., Suci, G. J. and Tannenbaum, P. H. (1957). *The Measurement of Meaning*, University of Illinois Press, Chicago, USA. (In English)

Pereira, F., Gershman, S., Ritter, S. and Botvinick, M. (2016). A comparative evaluation of off-the-shelf distributed semantic representations for modelling behavioural data, *Cognitive Neuropsychology*, 33, 175–190. <https://doi.org/10.1080/02643294.2016.1176907> (In English)

Schmolling, P. (1978). Cognitive Responses under Stress: The Effect of Examination Stress on Commonality of Verbal Associations, *Psychological Reports*, 43 (3_suppl), 1131–1138. (In English)

Schröder, T., Hoey, J. and Rogers, K. B. (2016). Modeling Dynamic Identities and Uncertainty in Social Interactions: Bayesian Affect Control Theory, *American Sociological Review*, 81 (4), 828–855. <https://doi.org/10.1177/0003122416650> (In English)

Stoltz, D. S. and Taylor, M. A. (2019). Concept Mover's Distance: measuring concept engagement via word embeddings in texts, *Journal of Computational Social Science*, 2 (2), 293–313. <https://doi.org/10.1007/s42001-019-00048-6> (In English)

Stoltz, D. S. and Taylor, M. A. (2022). text2map: R tools for text matrices, *Journal of Open Source Software*, 7 (72), 3741. <https://doi.org/10.21105/joss.03741> (In English)

Taylor, M. A. and Stoltz, D. S. (2021). Integrating Semantic Directions with Concept Mover's Distance to Measure Binary Concept Engagement, *Journal of Computational Social Science*, 4 (1), 231–242. <https://doi.org/10.1007/s42001-020-00075-8> (In English)

Utsumi, A. (2018). A Neurobiologically Motivated Analysis of Distributional Semantic Models, *Proceedings of the 40th Annual Conference of the Cognitive Science Society (CogSci2018)*, Madison, WI, USA, 1147–1152. (In English)

Voyer, A., Kline, Z. D., Danton, M. and Volkova, T. (2022). From Strange to Normal: Computational Approaches to Examining Immigrant Incorporation Through Shifts in the Mainstream, *Sociological Methods & Research*, 51 (4), 1540–1579.

<https://doi.org/10.1177/00491241221122596> (In English)

Wulff, D. U., Aeschbach, S., De Deyne, S. and Mata, R. (2022). Data From the MySWOW Proof-of-Concept Study: Linking Individual Semantic Networks and Cognitive Performance, *Journal of Open Psychology Data*, 10 (5), 1–8. <https://doi.org/10.5334/jopd.55> (In English)

Все авторы прочитали и одобрили окончательный вариант рукописи.

All authors have read and approved the final manuscript.

Конфликты интересов: у авторов нет конфликтов интересов для декларации.

Conflicts of interests: the authors have no conflicts of interest to declare.

Татьяна Александровна Литвинова, доктор филологических наук, ведущий научный сотрудник, Воронежский государственный педагогический университет, научно-исследовательская лаборатория компьютерной семасиологии, Воронеж, Россия.

Tatiana A. Litvinova, Doctor of Philological Sciences (D.Sc. in Philology), Leading Researcher, Voronezh State Pedagogical University, Computer Semasiology Laboratory, Voronezh, Russia.

Полина Вадимовна Паничева, кандидат филологических наук, заведующий научно-исследовательской лабораторией, Воронежский государственный педагогический университет, научно-исследовательская лаборатория компьютерной семасиологии, Воронеж, Россия.

Polina V. Panicheva, Candidate of Philological Sciences (Ph.D. in Philology), Head of Laboratory, Voronezh State Pedagogical University, Computer Semasiology Laboratory, Voronezh, Russia.