

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
Филологический факультет

АКТУАЛЬНЫЕ ПРОБЛЕМЫ ЛИНГВИСТИКИ И ЛИТЕРАТУРОВЕДЕНИЯ

Сборник материалов IX (XXIII)
Международной научно-практической
конференции молодых учёных
(14–16 апреля 2022 г.)

Выпуск 23

Томск
Издательство Томского государственного университета
2022

7. Méndez-Bértolo C. Word frequency modulates the processing of emotional words: convergent behavioral and electrophysiological data // *Neurosci. Lett.* 2011. Vol. 494. P. 250–254.

doi: 10.17223/978-5-907572-04-1-2022-45

**СОЗДАНИЕ БАЗЫ ДАННЫХ АБСТРАКТНЫХ
СУЩЕСТВИТЕЛЬНЫХ: ВОПРОС О МЕТОДОЛОГИИ
ПРОВЕДЕНИЯ ОПРОСА**

Вольская Ю.А.

Казанский федеральный университет, аспирант

**CREATING DATABASE OF ABSTRACT NOUNS:
ON THE QUESTION OF SURVEY METHODOLOGY**

Volskaya Yu.A.

Kazan Federal University, postgraduate student

В статье исследуется проблема методологии проведения психосемантического опроса с целью создания базы данных рейтингов конкретности/абстрактности. В центре внимания – вопрос об использовании подробных инструкций, которые предлагаются респондентам перед прохождением опроса.

Ключевые слова: абстрактные существительные, база данных, психосемантический опрос

This article is devoted to the study of methodology for conducting a psychosemantic survey, the main purpose of which is creating a database of «concreteness» / «abstractness» ranking. It focuses on the question of using the detailed instructions that are offered to respondents before taking the survey.

Key words: abstract nouns, database, psychosemantic survey

Научный руководитель: В.Д. Соловьев, д-р физ.-мат. наук, профессор КФУ.

Анализ современной литературы показывает, что абстрактность – это значимое понятие для многих исследований, например, в области нейролингвистики, в исследованиях памяти, в области изучения познания [1]. Для реализации подобных исследований необходимы точные данные о том, какие слова являются конкретными, а какие абстрактами.

На основе различных языков создаются базы данных (БД) абстрактных и конкретных слов с указанием степени проявления

признаков абстрактности/конкретности. Такие рейтинги были созданы для английского [2], голландского [3] и китайского языков [4]. На данный момент для русского языка создается база данных с рейтингами абстрактности/конкретности [5, 6]. Для проекта уже были собраны рейтинги для 1500 слов, однако словник планируется расширить.

При этом необходимо обратить внимание на методологию создания словаря абстрактных существительных, оценки которого собирались в ходе психосемантических опросов.

В ходе опроса респонденты получали списки из 50 существительных, которые нужно было оценить на степень конкретности/абстрактности по шкале Лайкерта от 1 до 5, где крайние значения шкалы, 1 и 5, указывают на конкретность и абстрактность, соответственно; показатель 3 указывает на то, что слово может расцениваться и как конкретное и как абстрактное. В списки слов были включены наиболее частотные существительные русского языка [7].

Первая тысяча слов оценивалась русскоязычными студентами Казанского федерального университета [5] и Белорусского государственного педагогического университета [8]. По каждому слову были получены оценки 40 носителей языка.

Оценки для 500 слов собирались путем проведения опросов на площадке Яндекс.Толока. На каждые 50 слов опрашивалось 60 человек разной возрастной категории: 30 человек в возрасте от 18 до 30, 30 человек – от 31 до 55. Также при анализе результатов учитывался возраст, пол, уровень образования (среднее, среднее специальное, неоконченное высшее, высшее, высшее филологическое). Для проверки качества ответов в каждый список было включено 10 контрольных слов. Не принимались ответы с единообразием оценок, а также те ответы, которые были даны прежде того, как истечет минимальное время.

При проведении опросов на основе 1000 слов для создания подобной базы данных русского языка авторы ограничились короткими инструкциями с указанием примеров абстрактных и конкретных слов [5]. Точные критерии для определения разряда

существительных указаны не были, так как считалось, что это может значительно повлиять на ответы респондентов. Таким образом, возможно снизить влияние субъективного опыта. При проведении психосемантических опросов для сбора рейтингов английских слов [2] использовались достаточно развернутые инструкции. Подобные английским инструкциям использовались и инструкции для создания базы данных голландского [3], китайского языков [4].

В инструкциях к опросам при создании базы данных для указанных языков можно выделить следующие основные повторяющиеся пункты: 1. Определение конкретности и абстрактности с опорой на сенсорное восприятие; 2. Примеры очень конкретных и очень абстрактных слов; 3. Указание на то, что в данном опросе не может быть «верных» или «неверных» ответов; просьба отвечать в зависимости от собственного субъективного опыта.

Для проверки степени влияния инструкций на ответы респондентов был проведен эксперимент, в рамках которого проводилось два опроса на определение степени абстрактности 100 слов среди двух групп респондентов. При проведении опроса респонденты первой группы получили подробные инструкции с указанием критериев разграничения абстрактных и конкретных существительных; респонденты второй группы – краткие пояснения с некоторыми примерами очень конкретных и очень абстрактных слов. Всего в эксперименте участвовало 60 человек (все – студенты Казанского федерального университета, обучающиеся на старших курсах филологического направления).

В список слов для опроса были включены те единицы, для которых уже были собраны рейтинги абстрактности [5]. Для описываемого эксперимента были взяты слова, которые демонстрируют весь диапазон степеней конкретности/абстрактности (от самых конкретных до самых абстрактных) и которые являются наиболее частотными для русского языка [7].

Респондентам были отправлены электронные формы опроса. Первой группе – с подробными инструкциями, второй группе – с краткими пояснениями и с примерами. Списки со словами были

разделены на списки по 50 слов, поэтому эксперимент проходил в два этапа. В первый день студенты получили первые списки, через три дня студентам были отправлены новые списки со следующими словами.

Между показателями опроса среди респондентов первой группы и среди респондентов второй группы был рассчитан коэффициент корреляции по Спирмену. По первому списку слов коэффициент равен 0,917, по второму списку слов – 0,907. Как видно, корреляционная связь достаточно сильная, что указывает на незначительную разницу в ответах респондентов.

Однако отметим, что наибольшая разница в ответах была обнаружена именно при указании степени абстрактных слов.

Коэффициент корреляции для данных слов равен 0,577, что указывает на более слабую корреляционную связь в сравнении с общими показателями. При этом во всех случаях респонденты первой группы оценивали указанные слова как более абстрактные, т.е. чаще выбирали показатель 4 или 5. Респонденты второй группы чаще выбирали показатель 2 или 3, что отразилось на общих результатах: ни одно слово не получило среднее значение больше или равное 4,5, всего три слова получили среднюю оценку больше 4 (*сила, любовь, душа*). Это указывает на то, что подробные инструкции несколько упрощают задачу респондентов при идентификации абстрактных единиц.

Корреляционная связь между оценками конкретных слов несколько выше: 0,728. При этом ни одно слово не было оценено респондентами разных группы с разницей больше или равной 1 или же с разницей меньше или равной –1.

Приведённые формальные критерии в подробных инструкциях позволили несколько уменьшить влияние субъективного языкового опыта респондентов, что, в первую очередь, повлияло именно на оценки абстрактности слов. Данные результаты интересны в рамках концепции «mode of acquisition» (MoA) [9].

Итак, при оценке конкретности респондентами двух групп не было обнаружено значительной разницы, поскольку конкретные слова в связи с перцептивным опытом наиболее понятны и просты

для восприятия. Для восприятия абстрактных слов респондентам необходимо иметь определённые языковые знания о данных лексических единицах. В данном случае подробные инструкции упрощают процесс оценки абстрактных единиц.

Список источников

1. Solovyev V. Concreteness/Abstractness Concept: State of the Art // *Advances in Intelligent Systems and Computing*. 2021. 1358 AIST. P. 275–283.
2. Brysbaert M., Warriner A.B., Kuperman V. Concreteness ratings for 40 thousand generally known English word lemmas // *Behavior research methods*. 2014. № 46 (3). P. 904–911.
3. Brysbaert M., Stevens M., De Deyne S., Wouter V., Gert S. Norms of age of acquisition and concreteness for 30,000 Dutch words // *Acta psychologica*. 2014. № 150. P. 80–84.
4. Xu X., Li J. Concreteness/abstractness ratings for two-character Chinese words in MELD-SCH // *PLoS ONE*. 2020. № 15 (6). URL: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0232133> (дата обращения: 03.03.2022).
5. Solovyev V.D., Ivanov V.V., Akhtiamov R.B. Dictionary of Abstract and Concrete Words of the Russian Language: A Methodology for Creation and Application // *Journal of research in applied linguistics*. 2019. № 10. P. 215–227.
6. Volskaya Y.A., Zhuravkina I.S., Lobanov A.P. Dictionary of abstract the words of the Russian language: Nouns with high numerical measure of abstractness // *International Journal of Criminology and Sociology*. 2020. № 9. P. 219–230.
7. Ляшевская О.Н., Шаров С.А. Частотный словарь современного русского языка : на материалах Национального корпуса русского языка. 2-е изд., испр. и доп. М. : Словари.ру, 2015. 1087, [1] с.
8. Zhuravkina I., Soloviev V., Lobanov A., Danilov A. Comparative Analysis of Concreteness/Abstractness of Russian Words // *Conference of Open Innovation Association, FRUCT*. 2020. Vol. 2020. April, Is. P. 464–470.
9. Della Rosa Pasquale A., Catricalà E., Vigliocco G., Cappa S.F. Behavior Research Methods Beyond the abstract-concrete dichotomy: Mode of acquisition, concreteness, imageability, familiarity, age of acquisition, context availability, and abstractness norms for a set of 417 Italian words // *Behavior Research Methods*. 2010. № 4. P. 1042–1048.