

СРАВНИТЕЛЬНЫЙ АНАЛИЗ НОВОГО СЛОВАРЯ ДЕСКРИПТОРОВ ЛИЧНОСТИ С ТЕЗАУРУСОМ ЛИЧНОСТНЫХ ЧЕРТ А.Г. ШМЕЛЕВА

А.Ю. КАЛУГИН^а, С.А. ЩЕБЕТЕНКО^а, А.М. МИШКЕВИЧ^а,
У.Е. СИРОТИНА^а, Ю.С. ВИТКО^а

^а *Национальный исследовательский университет «Высшая школа экономики», 101000, Россия,
Москва, ул. Мясницкая, д. 20*

A Comparative Analysis of a New Dictionary of Personality Descriptors and the Shmelev's Thesaurus of Personality Traits

A.Yu. Kalugin^a, S.A. Shchebetenko^a, A.M. Mishkevich^a, U.E. Sirotina^a, Yu.S. Vitko^a

^а *HSE University, 20 Myasnitskaya Str., Moscow, 101000, Russian Federation*

Резюме

Статья посвящена сравнению нового русского словаря описателей черт личности, разработанного авторами, с тезаурусом личностных черт, подготовленным А.Г. Шмелевым с коллегами в 1980–1990 гг. В новом словаре слова были отобраны из 140 тыс. прилагательных, существительных, глаголов и наречий. В ходе трехэтапного отбора по «немецкой методологии» были получены «полный» (2384 слова) и «сокращенный» (1253 слова) словари описателей черт личности. Из последнего были удалены однокоренные синонимы и антонимы. На разных этапах подготовки словаря описателей к исследованию подключалось от 6 до 45 экспертов разного возраста, пола, уровня компетентности. Для 624 слов из тезауруса А.Г. Шмелева были найдены полные аналоги в «полном» и «сокращенном»

Abstract

We compared a new Russian dictionary of personality descriptors with the thesaurus of personality traits by Alexander Shmelev and his colleagues developed in the 1980s and 1990s. In the new dictionary, the descriptors were selected from the Open Corpus, a global Russian word collection which had included over 140,000 adjectives, nouns, verbs, and adverbs. We utilized a “German methodology” of the word selection which comprises a three-stage process; in so doing, we consecutively yielded a full (2,384 words) and abridged (1,253 words) dictionaries of the personality descriptors, with cognate synonyms and antonyms being removed from the latter. While developing the full dictionary, we engaged groups of experts who varied in their age, sex, and expertise, in amounts from 6 to 45 experts per group. In the full dictionary, full equivalents to contents of the Shmelev's the-

Работа выполнена при финансовой поддержке РФФ, проект № 21-18-00127.

The research was carried out with the financial support of the RSF, project No 21-18-00127.

словарях. 1353 слова из тезауруса А.Г. Шмелева были отсеяны нами на разных этапах отбора (причины отсева обсуждаются в статье); 109 слов не обнаружили аналогов в первоначальном списке из 140 тыс. слов. 629 слов были в наших словарях, но отсутствовали в тезаурусе А.Г. Шмелева. Словарь А.Г. Шмелева на 26.2% совпадал с «полным» словарем и на 49.8% с «сокращенным» словарем. Выявленные различия в составе словарей не означают, что будут получены разные результаты при изучении структуры личности, так как ключевые характеристики личности могут быть отражены во всех рассмотренных словарях. В этом случае структуры черт личности, полученные на основе сравниваемых словарей, будут схожи. Если же структуры черт личности будут различаться, то это может быть объяснено методологией включения и исключения слов.

Ключевые слова: emic-подход, личность, структура личности, модель «Большая пятерка», описатели черт личности.

Калугин Алексей Юрьевич — старший научный сотрудник, департамент психологии, факультет социальных наук, Национальный исследовательский университет «Высшая школа экономики», кандидат психологических наук, доцент.
Сфера научных интересов: психология личности, психология индивидуальности, анализ данных.
Контакты: kaluginau@yandex.ru

Щебетенко Сергей Александрович — профессор, департамент психологии, факультет социальных наук, Национальный исследовательский университет «Высшая школа экономики», доктор психологических наук, доцент.
Сфера научных интересов: психология личности и индивидуальных различий, социальная психология, методология исследований.
Контакты: sshchebetenko@hse.ru

Мишкевич Арина Михайловна — старший научный сотрудник, департамент психологии, факультет социальных наук, Национальный исследовательский университет «Высшая школа экономики», кандидат психологических наук.
Сфера научных интересов: психология личности, социальные установки.
Контакты: ArinaMishkevich@ya.ru

сaurus were found for 624 words, whereas 1,353 words present in the Shmelev's thesaurus were excluded (the reasons for it are discussed in the paper). From the Shmelev's thesaurus, 109 words were absent in our initial 140,000 list, whereas 629 words in our dictionaries were absent in the Shmelev's thesaurus. Overall, the Shmelev's thesaurus was 26.2% equal to our full dictionary and 49.8% to the abridged dictionary. These differences do not imply different findings to be potentially revealed between the dictionaries, as the content of core personality characteristics can remain the same over the dictionaries. Consequently, estimations of the global Russian personality trait structure using either dictionary could closely fit. Conversely, the methodology of including or excluding descriptors in the particular dictionary could explain why the structures of personality were different.

Keywords: emic-approach; personality; personality structure; Big Five model; personality trait descriptors.

Alexey Yu. Kalugin — Senior Research Fellow, School of Psychology, Faculty of Social Sciences, HSE University, PhD in Psychology, Associate Professor.
Research Area: personality psychology, psychology of individuality, data analysis.
E-mail: kaluginau@yandex.ru

Sergei A. Shchebetenko — Professor, School of Psychology, Faculty of Social Sciences, HSE University, DSc in Psychology.
Research Area: personality psychology and individual differences, social psychology, methodology.
E-mail: sshchebetenko@hse.ru

Arina M. Mishkevich — Senior Research Fellow, School of Psychology, Faculty of Social Sciences, HSE University, PhD in Psychology.
Research Area: personality psychology, social attitudes.
E-mail: ArinaMishkevich@ya.ru

Сиротина Ульяна Евгеньевна — бакалавр, департамент психологии, факультет социальных наук, Национальный исследовательский университет «Высшая школа экономики». Сфера научных интересов: психология личности, эффективность медитативных практик. Контакты: uesiroтина@edu.hse.ru

Ulyana E. Sirotina — Bachelor's Degree in Psychology, School of Psychology, Faculty of Social Sciences, HSE University. Research Area: personality psychology, efficacy of meditation practices. E-mail: uesiroтина@edu.hse.ru

Витко Юлия Станиславовна — ассистент, департамент психологии, факультет социальных наук, Национальный исследовательский университет «Высшая школа экономики». Сфера научных интересов: психология личности, психология духовности, психология здоровья. Контакты: uliyvitko18@gmail.com

Yulia S. Vitko — Assistant, School of Psychology, Faculty of Social Sciences, HSE University. Research Area: personality psychology, psychology of spirituality, health psychology. E-mail: uliyvitko18@gmail.com

Etic- и emic-подходы к определению структуры черт личности

На сегодняшний день одной из самых разработанных и признанных в мировом научном сообществе структур личности является так называемая «Большая пятерка», включающая пять черт: экстраверсию, доброжелательность, добросовестность, нейротизм, интеллект/культуру/открытость опыту¹. Данная модель опирается на «лексическую гипотезу» о том, что основные сведения об индивидуальных различиях исторически нашли отражение в языке (Goldberg, 1981). Существует два основных подхода к изучению структуры личности в рамках данной гипотезы: etic-подход и emic-подход (Cheung et al., 2011; Jiang, 2021). Первый предполагает, что есть некая заданная модель, изначально полученная в некоторой языковой среде; эту модель в дальнейшем «примеряют» к разным культурам. Например, изначально американская пятифакторная модель личности впоследствии нашла подтверждение более чем в 50 странах (McCrae et al., 2005; Schmitt et al., 2007). Такой подход позволяет проверить кросс-культурную устойчивость модели, но практически не учитывает остаточную культурную специфику, отличающую данный культурный контекст от некоего взятого за эталон. Emic-подход, напротив, позволяет учесть культурную специфику, но более трудоемок. Психолексические emic-исследования обычно начинаются с отбора из словарей определенного языка слов, описывающих личность. Когда этот этап пройден и составлен словарь описателей черт личности, его предлагают носителям данного языка, которые оценивают по этим характеристикам себя или другого. Затем полученные результаты обрабатываются с помощью факторного анализа с целью получения эмпирической структуры личности (Thalmayer et al., 2020). Следует отметить, что именно таким путем изначально была получена модель «Большой пятерки» на американской выборке (Allport, Odbert, 1936; Goldberg, 1990; Norman, 1967).

¹ Данная черта является наиболее неустойчивой в своем содержании, что обуславливает комплексность ее названия (Connolly et al., 2014).

Исследования в рамках еміс-подхода показали, что «Большая пятерка» наиболее успешно воспроизводится в германских и славянских языках (Saucier, 2009). Однако даже тогда она имеет культурную специфику. Например, в чешском языке фактор открытости опыту содержательно описывал способности и ловкость (Hřebíčková, 2007), а в голландском и итальянском — социальную непокорность (De Raad et al., 1998).

В других языках «Большая пятерка» воспроизводится хуже, особенно за пределами западных культур (Laajaj et al., 2019; Thalmayer et al., 2020; Wood et al., 2020). Различия наблюдались как в количестве черт, так и в их содержании. Так, исследования на албанском (Shala et al., 2020) и сербском (De Raad et al., 2018b) языках продемонстрировали семь факторов, в то время как на венгерском языке (De Raad et al., 2018a) была получена шестифакторная структура. Как отмечает Дж. Сосье (Saucier, 2018), существуют различия и в содержании черт, выделенных на разных языках, но называемых одинаково.

Всего на данный момент психолексические исследования черт проведены более чем на 30 языках (De Raad et al., 2014; De Raad, Mlačić, 2017). Обнаружена ли универсальная структура личности? По мнению ряда авторов (De Raad, Barelds, 2020; De Raad et al., 2014; и др.), есть три черты, которые более-менее стабильно воспроизводятся в разных языках и культурах (так называемая панкультурная тройка): экстраверсия, доброжелательность и добросовестность. Остальные факторы являются «идиосинкратическими остатками» и могут заметно различаться от культуры к культуре.

Психолексический подход к отбору описателей личности

В мировой практике психолексических еміс-исследований сложились две традиции отбора слов из словарей: голландская (Brokken, 1978; De Raad et al., 1998) и немецкая (Angleitner et al., 1990). Представим их краткое описание (подробнее см.: Мишкевич и др., 2021).

Голландский вариант включает два этапа (Brokken, 1978): 1) из словаря отбираются все прилагательные, которые можно отнести к человеку; 2) отобранные слова разделяются на две категории: «критерий природы» (слово можно подставить в предложение «Этот человек [прилагательное] по своей природе») и «критерий личности» (слово может являться ответом на вопрос «Мистер X, что он за человек?»). Позднее в рамках голландской методологии помимо прилагательных также стали отбирать глаголы и существительные.

Немецкая методология имеет более сложную схему отбора слов (Angleitner et al., 1990). На первом этапе из словаря отбирают слова, с помощью которых можно описать человека или отличить поведение одного человека от поведения другого. Позднее данный этап был расширен за счет дополнительного исключения слов, иррелевантных описанию собственно личности (например, религиозной идентичности или географического происхождения). На втором этапе слова классифицируют по пяти категориям (диспозиции, временные состояния, социальные и репутационные аспекты, наружные характеристики и внешний вид, зависящие от конкретного контекста «термины

ограниченной полезности», метафорические, расплывчатые и т.п.) и 13 подкатегориям. В итоговый список попадают слова только из категории диспозиций.

Ключевое различие этих двух методологий заключается в критериях отбора: в немецкой методологии преимущественно исключаются оценочные признаки, при этом оценочные признаки, относящиеся к описанию темперамента или характера, сохраняются. В целом обе методологии отбора слов-описателей личности дают в итоге схожую структуру черт (Di Blas, Forzi, 1998).

Русскоязычные таксономии черт личности

На русском языке известны единичные попытки определения структуры черт личности на основе слов-описателей (Шмелев и др., 1991; Putilov, 2018).

В исследовании А.А. Путилова (Putilov, 2018) использовались два набора данных. В первый вошли 496 слов, взятых из списков описателей личности зарубежных таксономий. Они были переведены с английского языка на русский, а прилагательные преобразованы в существительные. Далее 414 участников исследования оценивали себя и двух других известных им людей, один из которых им нравился, а другой нет, по трехбалльной шкале: «да», «не знаю», «нет».

Второй список описателей для исследования, включающий 296 существительных, был получен путем объединения 148 существительных, отобранных из ранее полученного списка, с двумя наборами по 74 слова. Каждый из них, в свою очередь, был получен из списков В. Хофсти с соавт. (Hofstee et al., 1992) и А.Г. Шмелева с соавт. (Шмелев и др., 1991). Отбор слов осуществлялся в соответствии с циркумплексной моделью личностных черт, в которой признаки характеризуются их угловыми позициями в двумерном факторном пространстве (Hofstee et al., 1992) по 74 аспектам (фасетам) этой модели. Далее 149 респондентов оценивали себя и двух других известных им людей, один из которых им нравился, а другой нет, по пятибалльной шкале: «да, конечно», «да, наверное», «не знаю», «нет, наверное», «нет, конечно».

Строго говоря, исследование А.А. Путилова не следует емic-парадигме: в качестве источника слов использовался не словарь, а уже сокращенные тезаурусы, полученные на других культурах; при отборе слов автор опирался на априорную модель черт личности (шестифакторная структура, визуализируемая в трехмерном пространстве). Таким образом, речь может идти не о структуре черт личности, основанной на дескрипторах русского языка, а о проверке авторской модели черт личности на русской выборке. В этом отношении исследование Путилова представляет собой комбинацию емic- и еtic-подходов.

В рамках собственно емic-подхода было проведено исследование А.Г. Шмелева с коллегами по составлению «Тезауруса личностных черт» и проверке на его основе факторной структуры личности (Голдберг, Шмелев, 1993; Шмелев, Похилько, 1985; Шмелев и др., 1991; Shmelyov, Pokhil'ko, 1993).

На протяжении десятилетия под руководством А.Г. Шмелева составлялись, расширялись и корректировались словники² личностных черт. Первый словник, включавший 1530 слов, был составлен из личностных конструктов, предложенных студентами, а также из «Алфавита из 1500 черт личности» К.К. Платонова (Платонов, Голубев, 1977) и «Номинальной шкалы личностных качеств» В.Е. Хмелько (Хмелько, 1981). Впоследствии скорректированный и расширенный, дополненный понятиями из толкового «Словаря русского языка» (Ожегов, 1980), словник включил в себя 1650 описателей. Его финальный вариант содержал 2090 слов, характеризующих личность (подробно см.: Голдберг, Шмелев, 1993; Шмелев, 2002; Shmelyov, Pokhil'ko, 1993). Этот вариант (см. в кн.: Шмелев, 2002) был взят нами для сравнения со списком, полученным в ходе нашего собственного исследования русской таксономии личностных черт.

На основе полученного словника путем «интернальных суждений» — оценки связи слов между собой (в противоположность «экстернальным суждениям», суждениям респондентов о себе или другом) — А.Г. Шмелев с коллегами получили 15-факторную структуру личности, затем дополненную 16-м фактором социальной желательности (Шмелев и др., 1991; Шмелев, 2002).

Следует отметить, что подход А.Г. Шмелева с коллегами к отбору слов отличался от принятого за рубежом: ученые основывались более на психосемантической традиции, нежели на психолексической. Как отмечают сами авторы, могла сыграть свою роль ограниченная доступность для них актуальной зарубежной литературы в то время, это приводило к необходимости открывать свои пути для достижения тех же целей, что ставили перед собой зарубежные ученые в изучении структуры личности на основе словарного состава языка (Голдберг, Шмелев, 1993; Shmelyov, Pokhil'ko, 1993). Вероятно, когда зарубежные источники стали доступнее, в многолетнее исследование уже сложно было внести коррективы. Указанное, в конечном счете, породило определенные сложности в проведении кросс-культурных исследований на основе полученной русской таксономии (Saucier, Goldberg, 2001).

Постановка проблемы исследования

Итак, структура личности, полученная А.А. Путиловым на основе описателей личности, не может рассматриваться в качестве примера использования культурно-специфичного *emic*-подхода, так как при ее создании был упущен ключевой элемент — поэтапный процесс отбора слов из словаря конкретного языка. Психосемантический подход А.Г. Шмелева с коллегами, как и психолексический, соотносится с *emic*-подходом. В то же время в международной практике психолексических исследований были выработаны четкие алгоритмы отбора слов, описывающих личность (см., например: Angleitner et al., 1990; De Raad et al., 1998). Единый стандарт позволяет сравнивать структуры

² А.Г. Шмелев (2002) активно использует термин «словник» как синоним «словаря» или «тезауруса» личностных черт. Мы также будем использовать данный термин.

личности, полученные в разных странах, говорящих на разных языках. Однако, как отмечают, например, Дж. Сосье и Л. Голдберг (Saucier, Goldberg, 2001), различия в подходах к отбору слов препятствуют активному использованию русской таксономии А.Г. Шмелева в кросс-культурных исследованиях. При этом стоит учитывать, что различия в списках слов-описателей личности, полученных с помощью психолексического и психосемантического подходов, могут быть минимальны.

Целью данного исследования является определение степени сходства и различия двух русских тезаурусов описателей черт личности: (1) подготовленного под руководством А.Г. Шмелева в традициях психосемантического подхода и (2) подготовленного нами в традициях психолексического подхода³.

Гипотеза исследования: вследствие разницы в подходе А.Г. Шмелева с соавт. и в нашем подходе к отбору слов-описателей черт личности различие сравниваемых тезаурусов будет преобладать над сходством.

Метод

Участники исследования

На разных этапах подготовки словаря описателей к работе подключались разные группы экспертов.

В схеме исследования выделяются этапы, которые, в свою очередь, могут состоять из шагов. На первом этапе в первом шаге исследования участие приняли девять экспертов в возрасте от 21 до 30 лет ($M = 23.8$, $SD = 3.6$); все эксперты были женского пола. Среди них были один кандидат психологических наук, один филолог и семь студентов-психологов старших курсов. Во втором шаге участвовали шесть экспертов в возрасте от 21 до 30 лет ($M = 23.33$, $SD = 3.33$); все эксперты были женского пола. Среди них были один кандидат филологических наук и пять студентов-психологов старших курсов.

Во втором этапе участие приняли шесть экспертов в возрасте от 21 до 30 лет ($M = 24.5$, $SD = 4.32$); все эксперты были женского пола. Среди них были один кандидат филологических наук, один филолог и четыре студента-психолога старших курсов.

В первом шаге третьего этапа участие приняли шесть экспертов в возрасте от 22 до 45 лет ($M = 30.83$, $SD = 8.77$), из них двое мужчин и четыре женщины. Экспертами на этом этапе выступили один доктор психологических наук, два кандидата психологических наук, один кандидат филологических наук и два эксперта с психологическим образованием. Во втором шаге каждое слово оценивали по 12 экспертов (по 6 человек студенческого возраста и по 6 человек старшего возраста), всего 45 человек в возрасте от 19 до 74 лет ($M = 25.93$, $SD = 12.54$), включая 8 мужчин. Эти эксперты не имели психологического образования. Экспертов студенческого возраста было 38 человек (из них 5

³ В представленном сравнении словарей описателей личности не рассматривается словарь А.А. Путилова, так как он был получен путем комбинации еmic- и etic-подходов.

мужчин) в возрасте от 19 до 26 лет ($M = 21.16$, $SD = 1.61$). Семь экспертов (из них трое мужчин) были в возрасте от 30 до 74 лет ($M = 51.14$, $SD = 15.13$).

Подготовка словаря описателей черт личности

Из словаря русского языка (<https://opencorpora.org/dict.php>), полученного на основе свободно распространяемого «Открытого корпуса» (<http://opencorpora.org/>)⁴, были отобраны искомые слова четырех частей речи: прилагательные, существительные, глаголы и наречия. Первоначальный список насчитывал 140 тыс. слов.

Первый этап исследования состоял из двух шагов. В первом шаге первого этапа отбирались слова, описывающие человека, и отсеивались не описывающие. Во втором шаге первого этапа, в соответствии с «немецкой методологией», слова разносились по пяти категориям: 1) слова, связанные с политическим, философским, литературным, религиозным контекстом или искусством; 2) слова, описывающие национальный статус, географическое происхождение или профессию; 3) слова, определяемые как разговорные или вульгарные; 4) архаичные слова; 5) другие слова, с помощью которых можно описать человека или отличить поведение одного человека от поведения другого человека. Если слово попадало в пятую категорию, оно переходило на следующий этап, если в первые четыре — отсеивалось. Поскольку слово могло быть отнесено только к одной категории и была возможность ошибочно удалить многозначные слова, эксперты дополнительно оценивали отсеянные слова на предмет того, насколько они подходят для описания личности, характера и темперамента человека. Оценка проводилась по пятибалльной шкале от 1 («вообще ничего не говорит о личности/характере/темпераменте человека») до 5 («очень много говорит о личности/характере/темпераменте человека»). Слова, в среднем набравшие более 4 баллов, также переходили на следующий этап.

На втором этапе слова классифицировались по 5 категориям (см. раздел «Психолексический подход к отбору описателей личности») и 13 подкатегориям. На следующий этап переходили слова, попавшие в категорию «Диспозиции» (подкатегории: темперамент и характер; ментальные способности и таланты), даже если они одновременно были отнесены и к другим категориям.

На третьем этапе производилась «итоговая чистка» словаря. Число однокоренных синонимов и антонимов оценивала группа из пяти психологов и одного филолога. Понятность смысла слова и частоту его использования применительно к человеку оценивали 12 «наивных носителей языка» на слово. В результате слова либо удалялись, либо сохранялись. Так был получен «полный» словарь (2384 слова), включающий существительные, прилагательные, глаголы и наречия.

Помимо «полного» словаря был подготовлен «сокращенный» словарь (1253 слова), который был очищен от однокоренных синонимов в разных

⁴ Лингвистический корпус — это собрание особым образом подготовленных (в первую очередь, размеченных) машиночитаемых текстов, обладающих максимальной репрезентативностью и предназначенных для решения определенного круга задач (Захаров, 2019, с. 158).

частях речи (например, «честность» и «честный»). Преимущественно сохранялись прилагательные.

Внесем некоторые дополнительные пояснения. Если слово можно было применить к человеку только в качестве метафоры, то такое слово удалялось. Метафоры часто характеризуют не определенное поведение, мысли и чувства (определяя тем самым черты личности), а неочевидную ассоциацию между иррелевантным объектом (событием и т.п.) и чертой. Как следствие, метафоры дают широкий простор для интерпретации и увеличивают дисперсию ошибки при их использовании для диагностики черты личности.

Вульгарные слова на этапе самооценки могут восприниматься респондентами как оскорбительные, что может серьезно повлиять на определение структуры личности. Разговорные слова часто также имеют оскорбительный или метафорический оттенок. Кроме того, они могут иметь сленговый характер, распространенный в определенной субкультуре, и совершенно иначе восприниматься в других возрастных и культурных группах.

Устаревшие слова могут использоваться в художественной литературе, пословицах и поговорках, однако лексическое значение их утрачивается, забывается. Это также может увеличивать дисперсию ошибки, размывать содержание, которое кроется за понятием. Так как оценку слов проводили «наивные носители языка», а не специалисты-филологи, некоторые слова, отнесенные к «архаичным», таковыми не являются с научной точки зрения, хотя и редко используются в речи. Выбор в пользу «наивных носителей языка» был сделан осознанно, так как задача состояла не в том, чтобы научно обоснованно классифицировать материал, но в том, чтобы отразить отношение к словам в популяции. Если какие-то слова воспринимались «экспертами» как не относящиеся к человеку, вульгарные или устаревшие, то эта оценка принималась как характерная для популяции.

Удаление непонятных и редко используемых в отношении человека слов — важная часть работы над словарями описателей, которые будут впоследствии предъявляться носителям языка. Если какое-то слово окажется малознакомым и непонятым, респондент не сможет применить его к своей личности, а значит, ответ будет во многом случайным и внесет «шум» в результаты статистической обработки. Это одна из причин, почему удалялись узкоспециализированные, сленговые и малоизвестные слова.

Отметим, что в тезаурусе личностных черт А.Г. Шмелева (2002) вместо заявленных 2090 слов содержится 2088.

Сопоставление списков описателей личности осуществлялось с помощью среды языка программирования R.

Результаты

Полный список из 140 189 слов был сопоставлен с тезаурусом А.Г. Шмелева (2088 слов); для 1941 слова были найдены полные аналоги. В таблице 1 приведено количество слов, отсеянных нами по тем или иным причинам, но присутствующих в словнике А.Г. Шмелева.

Таблица 1

**Слова, отсеянные на разных этапах исследования, но присутствующие
в тезаурусе А.Г. Шмелева**

Этапы отсева	Количество отсеянных слов из тезауруса А.Г. Шмелева	Примеры исключенных слов
Этап 1. Первичный отбор	428 (6)	
Шаг 1. Исключение слов, которые не относятся к описанию человека (например, «подушка»)	216 (6)	жила, аспид, свинья, продувной
Шаг 2. Исключение слов...	212	
1) связанных с политическим, философским, литературным, религиозным контекстом или искусством	39	анархист, экстремист, догматичный, поэтический
2) описывающих национальный статус, географическое происхождение или профессию	25	звонарь, комедиант, ученый, шут
3) определяемых как разговорные или вульгарные	159	бездарь, дармоед, трепло, чумовой
4) архаичных	18	крамольник, горемычный, окаянный, срамной
Этап 2. Классификация слов	711 (24)	
1. Временные состояния (эмоции, настроения и когниции; физические состояния; наблюдаемая активность)	155 (16)	вдохновенный, грустный, истощенный, сытый
2. Социальные и репутационные аспекты (социальные роли и взаимоотношения; установки и мировоззрения; социальный статус и др.)	589 (13)	брехун*, волонтерист, неподкупный, привилегированный
3. Наружные характеристики и внешний вид (анатомия и телосложение; внешний вид, образ, манеры)	188 (1)	изящный, низкий, опрятный, экстравагантный
4. Термины ограниченной полезности (термины, зависящие от конкретного контекста, или технические термины; метафорические, расплывчатые, устаревшие слова; термины, обозначающие физические/моторные способности)	399 (19)	волокита, параноик, глубокий, плодовый
Этап 3. Итоговая чистка словаря (удаление однокоренных синонимов и антонимов, непонятных и низкочастотных слов)	181 (4)	

Примечание. При сложении отсеянных слов по категориям сумма будет больше, чем указанная на уровне шагов и этапов, т.к. одно и то же слово могло быть отсеяно по разным основаниям; в скобках указано количество удаленных слов из дополнительного списка в 38 близких однокоренных синонимов или антонимов.

* У читателя может возникнуть вопрос: как слово с явно «разговорным и вульгарным» оттенком могло перейти на следующий этап отбора, если соответствующий фильтр был на предыдущем

В списке из 140 тыс. слов не были найдены полные аналоги для 147 слов из тезауруса А.Г. Шмелева. Однако для 38 из них можно найти близкие однокоренные синонимы или антонимы (например, «героичный» — «героический», «заслуженный» — «незаслуженный», «просторечивый» — «просторечный»). В таблице 1 в скобках приведено количество отсеянных слов из этого дополнительного списка.

«Полный» словарь совпадал с тезаурусом А.Г. Шмелева по 622 словам. Из дополнительного списка 34 слова были отсеяны, 2 подобранных однокоренных слова уже входили в указанные 622 слова, 2 слова были новыми и совпадали с нашим списком («героический», «неприспособленный»). Таким образом, 624 слова были одинаковыми в нашем словаре и тезаурусе А.Г. Шмелева. Сходство составляло 29.9% для тезауруса А.Г. Шмелева и 26.2% для «полного» словаря описателей личности. «Сокращенный» словарь при учете однокоренных синонимов из разных частей речи совпадал с тезаурусом А.Г. Шмелева на 49.8%.

Таким образом, 1353 слова из тезауруса А.Г. Шмелева были отсеяны нами на разных этапах отбора. 109 слов тезауруса не нашли соответствий с нашим полным списком. Многие из них можно было бы заменить синонимами (например, «гомонливый» — «шумливый»), но мы не пошли по этому пути, так как синонимы могут иметь свой смысловой оттенок.

Обсуждение

Тестировалась гипотеза о том, что вследствие разницы в подходах отбора слов различие рассматриваемых словарей-описателей личности будет преобладать над сходством. Было обнаружено, что «полный» словарь и тезаурус А.Г. Шмелева содержат 26.2% одинаковых слов (624 из 2384), а «сокращенный» словарь (при учете однокоренных синонимов из разных частей речи) — 49.8% (624 из 1253).

Ряд слов (109) не вошел в изначальный словарь из 140 тыс. слов, но присутствовал в тезаурусе А.Г. Шмелева, что можно объяснить следующим:

- словарь не включал словосочетания («с развитой интуицией», «с развитым воображением», «с чувством долга»);
- в словаре отсутствовали сложные слова, включающие несколько корней, с большой вариативностью в плане сочетания («абстрактномыслящий», «глубокомыслящий», «конкретномыслящий»);

этапе? Поясним. Как отмечалось выше, поскольку эксперты относили слово только к одной из пяти категорий, была опасность отсеять слова, имеющие «личностный контекст», поэтому была проведена дополнительная оценка отсеянных слов. Эксперты оценивали каждое слово по пятибалльной шкале на предмет «описания личности, характера или темперамента человека»: 1 балл — «вообще ничего не говорит о личности/характере/темпераменте человека», 5 баллов — «очень многое говорит о личности/характере/темпераменте человека». Если средняя оценка экспертов была равна 4 или выше, то слово возвращалось в словарь (примером может служить слово «брехун»).

- мы использовали в словаре только существительные, прилагательные, глаголы и наречия; таким образом, в него не вошли причастия: «думающий», «задавленный», «зарвавшийся», «метушийся»;

- тезаурус А.Г. Шмелева включал в себя узкоспециализированные слова, профессиональный сленг («гипертимный», «гипнабельный»);

- тезаурус включал в себя устаревшие либо разговорные слова («духарной», «измочаленный», «стукач», «якающий»).

Также были слова, не вошедшие в тезаурус А.Г. Шмелева, но присутствующие в нашем словаре:

- как уже было указано ранее, сложные слова, состоящие из нескольких корней, имеют большую вариативность в плане сочетания (например, «быстроработающий», «быстрореагирующий», «быстросхватывающий»);

- наиболее частотной частью речи, отсутствующей в тезаурусе А.Г. Шмелева, но вошедшей в наш словарь, были глаголы (314 слов); далее по степени частотности шли существительные-атрибуты⁵ (223 слова), прилагательные (68), наречия (49) и существительные-типы (25)⁶. Как и в случае с причастиями выше, части речи, не включенные в отбор, являются причиной большинства различий в словарях. Тезаурус А.Г. Шмелева включал преимущественно прилагательные и существительные; глаголы и наречия в нем отсутствовали.

«Полный» и «сокращенный» словари в разной степени совпадали с тезаурусом А.Г. Шмелева. Это связано, с одной стороны, с количеством слов в словарях, а с другой — с удалением однокоренных синонимов в «сокращенном» словаре. Например, для слова «авантюрист» из тезауруса А.Г. Шмелева мы находили сопоставимое прилагательное в «сокращенном» словаре — «авантюристичный», так как оно было получено путем обобщения, удаления однокоренных синонимов в разных частях речи («авантюрист», «авантюристичный», «авантюристность»). В «полном» словаре есть слово «авантюрист», и только ему соответствует слово из тезауруса А.Г. Шмелева, «авантюристичный» же и «авантюристность» в тезаурусе отсутствуют. Таким образом, «сокращенный» словарь за счет своей большей обобщенности точнее передает сходство со словником А.Г. Шмелева, чем «полный».

Составление словарей описателей личности — подготовительный этап исследования национальной структуры личности. Совпадение словарей двух российских таксономий лишь на 30–50% говорит о серьезном влиянии используемой методологии на отбор слов. Однако это не обязательно приведет к разным

⁵ В рамках «немецкой методологии» принято различать существительные-типы (например, «агрессор») и существительные-атрибуты (например, «агрессия») (Angleitner et al., 1990).

⁶ Если сложить не вошедшие в словник Шмелева, но присутствующие в нашем словаре слова и 624 указанных выше, сумма окажется больше 1253. Одно слово «сокращенного» словаря обобщало несколько однокоренных слов, имеющих в «полном» словаре (например, «лживый» обобщало слова: «лживый», «лживость», «лгун», «лжец», «лгать»). В тезаурусе Шмелева присутствовали слова «лживый», «лгун» и «лжец», поэтому «лживый» из нашего словаря считался трижды. Подробнее см. далее по тексту. Всего же 629 слов были в «сокращенном» словаре, но отсутствовали в тезаурусе Шмелева.

результатам при изучении структуры личности. Ключевые характеристики личности могут быть отражены как в тезаурусе А.Г. Шмелева, так и в нашем словаре, что в итоге даст схожую структуру черт.

В то же время нельзя утверждать, что таких различий не будет. На наш взгляд, влияние может оказать не только процедура отбора слов (психолексическая или психосемантическая), но и разница в интернальном и экстернальном подходах к получению структуры черт личности. Респондент, примеряющий слова-понятия к своей личности, может отразить в своих ответах всю сложность структуры личности реального человека. На его ответы будут влиять «социальная желательность», «защитные механизмы» (проекция, интроекция, рационализация и т.д.), уровень рефлексивности, самооценка, эрудированность и многое другое. Учет этих моментов отражен и в методологии отбора, когда отсеиваются вульгарные, непонятные, метафоричные или сленговые слова. Эксперт, оценивающий сходство слов, в меньшей степени лично вовлечен, смотрит на ситуацию со стороны. Может показаться, что эксперт более объективен, однако в его случае субъективизм в оценках также будет присутствовать, но он иной: это прежде всего социальные и личностные установки и стереотипы, когнитивные искажения. Более того, эксперт опирается на свой понятийный опыт использования слов как терминов. Напротив, при экстернальном подходе «наивный» респондент оценивает слова в контексте своего жизненного опыта, прибегая к семантической и эпизодической памяти. Таким образом получаемая факторная структура лучше отражает устройство и состав системы личности.

Итак, поставленная цель была достигнута — определена степень сходства и различия двух тезаурусов слов-описателей черт личности, а тестируемая гипотеза получила эмпирическую поддержку.

Заключение

Развернувшиеся в последние десятилетия психолексические исследования в рамках етіс-подхода показали, что, несмотря на достаточно высокую кросс-культурную устойчивость модели «Большая пятерка», в различных языках обнаруживаются и специфические черты, которые могут многое сообщить об особенностях жизни людей в той или иной культуре, нашедших отражение в языке.

На русском языке было проведено, пожалуй, только одно масштабное таксономическое исследование. В 1980–1990-е гг. на основе значительного психосемантического материала группой А.Г. Шмелева был подготовлен тезаурус личностных черт, изучена структура личности, присущая российской (советской) популяции.

Эти изыскания методологически отличались от психолексических исследований, проводимых за рубежом, поэтому их результаты получили ограниченное использование в кросс-культурных исследованиях. Нашей группой была предпринята попытка составить тезаурус слов-описателей черт личности, следуя принятой за рубежом методологии в ее «немецкой» версии. Этот

словарь (в «полном» и «сокращенном» вариантах) сравнивался с тезаурусом А.Г. Шмелева и его коллег. Обнаруженные различия словарей оказались существенными.

Таким образом, гипотеза о том, что вследствие разницы в подходах к отбору слов различие рассматриваемых тезаурусов слов-описателей черт личности будет преобладать над сходством, нашла эмпирическую поддержку. «Полный» словарь отличается от тезауруса А.Г. Шмелева на 73.8%, а «сокращенный» словарь вследствие своей большей обобщенности — на 50.2%.

Ограничения и перспективы исследования

Оценка слов осуществлялась «наивными носителями языка», а не профессиональными филологами и лингвистами. Поэтому слова могли быть не отнесены в филологически корректные категории, а некоторые слова, описывающие личность, могли быть отброшены на каком-либо из этапов. Мы в полной мере осознаем и осознавали данное ограничение, однако участие «наивных носителей языка» — неременное условие получения адекватного списка описателей личности.

На разных этапах отбора количество экспертов, оценивающих одно слово, варьировало от 2 до 12. Малое количество экспертов-оценщиков на первом этапе могло привести к чрезмерному отсеву слов из-за статистической склонности малых выборок к асимметрии и эксцессу, а большее количество на последних этапах могло приводить к избыточному сохранению слов.

А.Г. Шмелев с коллегами помимо составления тезауруса личностных черт также выделили стабильную 15-факторную структуру личности (Шмелев и др., 1991), соотнесли русскую и английскую версии структуры личности (Голдберг, Шмелев, 1993). Повторение этих исследований с использованием полученного нами списка слов может быть перспективным. С одной стороны, это позволит выделить инварианты в структуре личности, с другой — изучить динамику личностных черт. За 35 лет в российском обществе произошли серьезные социально-экономические и культурные изменения, что могло отразиться на структуре черт личности россиянина.

Литература

- Голдберг, Л. Р., Шмелев, А. Г. (1993). Межкультурное исследование лексики личностных черт: «Большая пятерка» факторов в английском и русском языках. *Психологический журнал*, 14(4), 32–39.
- Захаров, В.П. (ред.). (2019). *Моделирование в корпусной лингвистике: специализированные корпуса русского языка*. СПб.: Изд-во Санкт-Петербургского университета.
- Мишкевич, А. М., Щebetенко, С. А., Калугин, А. Ю. (2021). К вопросу об организации психологического исследования русской таксономии черт. *Вестник Пермского университета. Философия. Психология. Социология*, 4, 613–627. <https://doi.org/10.17072/2078-7898/2021-4-613-627>

- Ожегов, С. И. (1980). *Словарь русского языка*. М.: Советская энциклопедия.
- Платонов, К. К., Голубев, Г. В. (1977). *Психология*. М.: Высшая школа.
- Хмелько, В. Е. (1981). Номинальная шкала личностных качеств. *Социологические исследования*, 2, 124–134.
- Шмелев, А. Г. (2002). *Психодиагностика личностных черт*. СПб.: Речь.
- Шмелев, А. Г., Похилько, В. И. (1985). Экспериментальный подход к построению базисного семантического словаря черт личности. *Вестник Московского университета. Серия 14. Психология*, 3, 3–10.
- Шмелев, А. Г., Похилько, В. И., Козловская-Тельнова, А. Ю. (1991). Репрезентативность личностных черт в сознании носителя русского языка. *Психологический журнал*, 12(2), 27–41.

Ссылки на зарубежные источники см. в разделе *References*.

References

- Allport, G. W., & Odbert, H. S. (1936). Trait-names: A psycho-lexical study. *Psychological Monographs*, 47(1), i–171. <https://doi.org/10.1037/h0093360>
- Angleitner, A., Ostendorf, F., & John, O. P. (1990). Towards a taxonomy of personality descriptors in German: A psycho-lexical study. *European Journal of Personality*, 4(2), 89–118. <https://doi.org/10.1002/per.2410040204>
- Brokken, F. B. (1978). *The language of personality* [Unpublished doctoral dissertation]. University of Groningen, Groningen, the Netherlands.
- Cheung, F. M., van de Vijver, F. J. R., & Leong, F. T. L. (2011). Toward a new approach to the study of personality in culture. *American Psychologist*, 66(7), 593–603. <https://doi.org/10.1037/a0022389>
- Connelly, B. S., Ones, D. S., & Chernyshenko, O. S. (2014). Introducing the special section on openness to experience: Review of openness taxonomies, measurement, and nomological net. *Journal of Personality Assessment*, 96(1), 1–16. <https://doi.org/10.1080/00223891.2013.830620>
- De Raad, B., & Barelds, D. P. H. (2020). Models of personality structure. In P. J. Corr & G. Matthews (Eds.), *The Cambridge handbook of personality psychology* (pp. 115–128). Cambridge University Press. <https://doi.org/10.1017/9781108264822.012>
- De Raad, B., Barelds, D. P. H., Timmerman, M. E., De Roover, K., Mlačić, B., & Church, A. T. (2014). Towards a pan-cultural personality structure: Input from 11 psycho-lexical studies. *European Journal of Personality*, 28(5), 497–510. <https://doi.org/10.1002/per.1953>
- De Raad, B., Di Blas, L., & Perugini, M. (1998). Two independently constructed Italian trait taxonomies: Comparisons among Italian and between Italian and Germanic languages. *European Journal of Personality*, 12(1), 19–41. [https://doi.org/10.1002/\(SICI\)1099-0984\(199801/02\)12:1<19::AID-PER290>3.0.CO;2-6](https://doi.org/10.1002/(SICI)1099-0984(199801/02)12:1<19::AID-PER290>3.0.CO;2-6)
- De Raad, B., & Mlačić, B. (2017). The lexical foundation of the Big Five factor model. In T. A. Widiger (Ed.), *The Oxford handbook of the five factor model* (pp. 191–216). New York, NY: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199352487.013.12>
- De Raad, B., Nagy, J., Szirmák, Z., & Barelds, D. P. (2018a). Taxonomy of Hungarian personality traits: Replication and refinement. *International Journal of Personality Psychology*, 4(1), 1–12.
- De Raad, B., Smederevac, S., Čolović, P., & Mitrović, D. (2018b). Personality traits in the Serbian language: Structure and procedural effects. *Journal of Research in Personality*, 73, 93–110. <https://doi.org/10.1016/j.jrp.2017.11.008>

- Di Blas, L., & Forzi, M. (1998). An alternative taxonomic study of personality-descriptive adjectives in the Italian language. *European Journal of Personality*, 12(2), 75–101. [https://doi.org/10.1002/\(sici\)1099-0984\(199803/04\)12:2<75::aid-per288>3.0.co;2-h](https://doi.org/10.1002/(sici)1099-0984(199803/04)12:2<75::aid-per288>3.0.co;2-h)
- Goldberg, L. R. (1981). Language and individual differences: The search for universals in personality lexicons. *Review of Personality and Social Psychology*, 2(1), 141–165.
- Goldberg, L. R. (1990). An alternative “description of personality”: the Big-Five factor structure. *Journal of Personality*, 59(6), 1216–1229. <https://doi.org/10.1037/0022-3514.59.6.1216>
- Goldberg, L. R., & Shmelev, A. G. (1993). Mezhekul’turnoe issledovanie leksiki lichnostnykh chert: “Bol’shaya pyaterka” faktorov v angliiskom i russkom yazykakh [Cross-cultural study of the vocabulary of personality traits: the “Big Five” factors in English and Russian]. *Psikhologicheskii Zhurnal*, 14(4), 32–39.
- Hofstee, W. K., De Raad, B., & Goldberg, L. R. (1992). Integration of the Big Five and circumplex approaches to trait structure. *Journal of Personality and Social Psychology*, 63(1), 146–163. <https://doi.org/10.1037/0022-3514.63.1.146>
- Hřebíčková, M. (2007). The lexical approach to personality description in the Czech context. *Ceskoslovenská Psychologie*, 51, 50–61.
- Jiang, K. (2021). Literature review of measurements of personality traits across cultures. *Journal of Psychological Research*, 3(4), 42–47. <https://doi.org/10.30564/jpr.v3i4.3838>
- Khmel’ko, V. E. (1981). Nominal’naya shkala lichnostnykh kachestv [Nominal scale of personality traits]. *Sotsiologicheskie Issledovaniya [Sociological Research]*, 2, 124–134.
- Laajaj, R., Macours, K., Hernandez, D. A. P., Arias, O., Gosling, S. D., Potter, J., Rubio-Codina, M., & Vakis, R. (2019). Challenges to capture the big five personality traits in non-WEIRD populations. *Science Advances*, 5(7). <https://doi.org/10.1126/sciadv.aaw5226>
- McCrae, R. R., Terracciano, A., & 78 Members of the Personality Profiles of Cultures Project. (2005). Universal features of personality traits from the observer’s perspective: Data from 50 cultures. *Journal of Personality and Social Psychology*, 88(3), 547–561. <https://doi.org/10.1037/0022-3514.88.3.547>
- Mishkevich, A. M., Shchebetenko, S. A., & Kalugin, A. Yu. (2021). On the organization of a psycho-lexical study of the Russian personality trait taxonomy. *Vestnik Permskogo Universiteta. Filosofia. Psihologia. Sociologia [Perm University Herald. Series “Philosophy. Psychology. Sociology”]*, 4, 613–627. <https://doi.org/10.17072/2078-7898/2021-4-613-627> (in Russian)
- Norman, W. T. (1967). *2800 personality trait descriptors: normative operating characteristics for a university population*. Ann Arbor, MI: University of Michigan.
- Ozhegov, S. I. (1980). *Slovar’ russkogo yazyka* [Dictionary of the Russian language]. Moscow: Sovetskaya Entsiklopediya.
- Platonov, K. K., & Golubev, G. V. (1977). *Psikhologiya* [Psychology]. Moscow: Vysshaya shkola.
- Putilov, A. A. (2018). A 3-D look at the Russian personality traits structure. *Current Psychology*, 37(3), 528–542. <https://doi.org/10.1007/s12144-016-9535-y>
- Saucier, G. (2009). Recurrent personality dimensions in inclusive lexical studies: Indications for a Big Six structure. *Journal of Personality*, 77(5), 1577–1614. <https://doi.org/10.1111/j.1467-6494.2009.00593.x>
- Saucier, G. (2018). Culture, morality and individual differences: comparability and incomparability across species. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1744), Article 20170170. <https://doi.org/10.1098/rstb.2017.0170>

- Saucier, G., & Goldberg, L. R. (2001). Lexical studies of indigenous personality factors: Premises, products, and prospects. *Journal of Personality*, 69(6), 847–879. <https://doi.org/10.1111/1467-6494.696167>
- Schmitt, D. P., Allik, J., McCrae, R. R., & Benet-Martínez, V. (2007). The geographic distribution of Big Five personality traits: Patterns and profiles of human self-description across 56 nations. *Journal of Cross-Cultural Psychology*, 38(2), 173–212. <https://doi.org/10.1177/0022022106297299>
- Shala, R. A., De Raad, B., & Arënlju, A. (2020). The structure of Albanian personality descriptive trait-adjectives: A psycho-lexically based taxonomy. *International Journal of Personality Psychology*, 6, 1–11. <https://doi.org/10.21827/ijpp.6.37224>
- Shmelev, A. G. (2002). *Psikhodiagnostika lichnostnykh chert* [Psychodiagnostics of personality traits]. Saint Petersburg: Rech'.
- Shmelev, A. G., & Pokhil'ko, V. I. (1985). Eksperimental'nyi podkhod k postroeniyu bazisnogo semanticheskogo slovarya chert lichnosti [An experimental approach to building a basic semantic dictionary of personality traits]. *Moscow University Psychology Bulletin. Series 14. Psychology*, 3, 3–10.
- Shmelev, A. G., Pokhil'ko, V. I., & Kozlovskaya-Tel'nova, A. Yu. (1991). Rerezentativnost' lichnostnykh chert v soznanii nositelya russkogo yazyka [Representativeness of personality traits in the mind of a Russian language speaker]. *Psikhologicheskii Zhurnal*, 12(2), 27–41.
- Shmelyov, A. G., & Pokhil'ko, V. I. (1993). A taxonomy oriented study of Russian personality trait names. *European Journal of Personality*, 7(1), 1–17. <https://doi.org/10.1002/per.2410070102>
- Thalmayer, A. G., Saucier, G., Ole-Kotikash, L., & Payne, D. (2020). Personality structure in east and west Africa: Lexical studies of personality in Maa and Supyire-Senufo. *Journal of Personality and Social Psychology*, 119(5), 1132–1152. <https://doi.org/10.1037/pspp0000264>
- Wood, J. K., Gurven, M., & Goldberg, L. R. (2020). Ubiquitous personality-trait concepts in 13 diverse and isolated languages: A cluster-classification approach. *European Journal of Personality*, 34(2), 164–179. <https://doi.org/10.1002/per.2246>
- Zakharov, V. P. (Ed.). (2019). *Modelirovaniye v korpusnoy lingvistike: spetsializirovannyye korpusy russkogo yazyka* [Modeling in corpus linguistics: specialized corpora of the Russian language]. Saint Petersburg: Saint Petersburg State University.