

База русских идиом (БРИ) с нормированными психолингвистическими параметрами

Гриднева Екатерина Михайловна¹, Здорова Нина Станиславовна^{1,2}, Иваненко Анастасия Андреевна¹, Макарова Полина Сергеевна¹, Грабовская Мария Андреевна¹

¹НИУ ВШЭ, ²Институт языкознания РАН

Цель данного исследования – создание Базы русских идиом (БРИ) с нормированными психолингвистическими параметрами. Наиболее разработанные классификации русских идиом основываются на описании степени семантической слитности компонентов идиомы. Данных по этому параметру недостаточно для проведения психолингвистических исследований о языковой обработке идиом. Заполнение лакуны на предмет распространенного описания психолингвистических параметров идиом определяет актуальность настоящей работы. В данной работе описаны результаты создания Базы из 376 русских идиом, содержащей в себе объективные (подсчитанные на основании корпусов) и субъективные (полученные в ходе опроса носителей) психолингвистические параметры. В опросе приняли участие 485 носителей русского языка в возрасте от 18 до 76 лет ($M = 35,3$, $SD = 13,4$). Были удалены результаты участников, родным языком которых русский не является, и тех, кто указал более одного родного языка. Всего было исключено 29 респондентов. Так, анализ проводился по результатам финальной выборки из 456 респондентов (354 женщины, 93 мужчины, 9 предпочли не указывать свой гендер). К объективным параметрам относятся: частотность употребления идиом (по корпусам НКРЯ и SketchEngine, подкорпус ruTenTen11), длина (в словах и символах) и тип синтаксической структуры, а к субъективным – встречаемость, знакомость, буквальность, предсказуемость и место распознавания идиомы. Параметр встречаемости отражает, как часто носители сталкиваются с идиомой в коммуникации (Gernsbacher, 1984). По параметру знакомости можно установить, насколько хорошо носитель языка осведомлён о значении идиомы (Nusbaum, Pisoni, & Davis, 1984; Libben & Titone 2008). Встречаемость и знакомость оказывают влияние на распознавание слов (Connine, Mullbx, Shemoff, & Yelen, 1990; Gernsbacher, 1984), на скорость чтения (Schweigert & Moates, 1988) и восприятие новых метафор (Blasko & Connine, 1993), которые, как и идиомы, являются частью фигуративного языка. Параметр буквальности означает возможность дословной интерпретации идиомы. Предсказуемость идиом определяется как вероятность правильного завершения неполной идиомы. Под местом распознавания идиомы в нашей работе вслед за (Beck & Weber, 2016) подразумевается место за словом, после которого не менее 70% участников верно продолжили выражение как контролируруемую идиому. Влияние данных характеристик на процессы обработки и порождение идиом было описано в работах по экспериментальной психолингвистике на материале английского и немецкого языков (Cacciari & Corradini 2015; Debra & Connine 1994; Beck & Weber 2016; 2020). Отличительной чертой БРИ является описание лексической вариативности идиом (Таблица 1.), а также анализ влияния и корреляции объективных и субъективных параметров между собой (Таблица 2.), которые не приводились в аналогичных ресурсах на других языках.

Таблица 1. Лексическая вариативность в продолжении идиом по первым словам

Идиома	Предсказуемость	Место распознавания	Лексическая вариативность	Лексическая вариативность: процент выбора иной формы
выйти из ума	1%	выйти из ума	выйти из_себя, выйти из_тени, выйти из_тупика, выйти_в свет, выйти_из себя, выйти_на свет	38, 3, 2, 3, 12, 1
волей-неволей	92%	волей- неволей	отсутствует	отсутствует

Исходя из примеров в Таблице 1, можно предположить, что место распознавания идиомы связано с лексической вариативностью: если идиома правильно завершается только при указании всех её компонентов, то это повышает ее лексическую вариативность. Однако если идиома легко определяется по ее началу, это уменьшает вероятность лексической вариативности. Более предсказуемые идиомы обладают меньшей лексической вариативностью. Эту гипотезу мы также планируем проверить в будущих исследованиях, однако она косвенно подтверждается работами о параметрах предсказуемости (Cassiaji & Tabossi, 1988) и месте распознавания идиомы (Beck & Weber, 2016).

Таблица 2. Результаты корреляционного анализа субъективных психолингвистических параметров

	Предсказуемость	Встречаемость	Знакомость
Встречаемость	0,39****		
Знакомость	0,31****	0,79****	
Буквальность	-0,03	0,12*	0,04

Примечание: Уровни значимости: **** $p < .0001$, ** $p < .01$, * $p < .05$

Из Таблицы 2 видно, что самая высокая положительная корреляция обнаруживается для параметров знакомости и встречаемости (0,79). Слабая положительная корреляция наблюдается между предсказуемостью и встречаемостью, а также между предсказуемостью и знакомостью. Не наблюдается корреляции и наблюдается слабая корреляция между буквальностью и другими субъективными параметрами.

Полные результаты анализа с серией линейных моделей на каждый субъективный параметр как зависимую переменную, а также описательная статистика по

субъективным психолингвистическим параметрам (среднее, стандартное отклонение, диапазон) будут приведены далее.

БРИ может быть использована для подготовки психолингвистических экспериментов, исследований в области корпусной и компьютерной лингвистики, а также в практике преподавания русского языка как иностранного школьникам и иностранным студентам.

Библиография

Баранов А. Н., Добровольский Д. О. Аспекты теории фразеологии. — М. : Знак, 2008. — 656 с.

Кустова Г. И. Словарь русской идиоматики. Сочетания слов со значением высокой степени. М., 2008. URL: <http://dict.ruslang.ru/magn.php?act=list&list=>

Arnon, T., Lavidor, M. Cognitive control in processing ambiguous idioms: evidence from a self-paced reading study. *J Psycholinguist Res* 52, 261–281 (2023). <https://doi.org/10.1007/s10936-022-09861-z>

Bates D, Mächler M, Bolker B, Walker S (2015). “Fitting Linear Mixed-Effects Models Using lme4.” *Journal of Statistical Software*, 67(1), 1–48. doi:10.18637/jss.v067.i01.

Beck, S.D., Weber, A. Context and Literality in Idiom Processing: Evidence from Self-Paced Reading. *J Psycholinguist Res* 49, 837–863 (2020). <https://doi.org/10.1007/s10936-020-09719-2>

Cristina Cacciari & Paola Corradini (2015) Literal analysis and idiom retrieval in ambiguous idioms processing: A reading-time study, *Journal of Cognitive Psychology*, 27:7, 797-811, DOI: 10.1080/20445911.2015.1049178

Cieślicka AB. Idiom Acquisition and Processing by Second/Foreign Language Learners. In: Heredia RR, Cieślicka AB, eds. *Bilingual Figurative Language Processing*. Cambridge University Press; 2015:208-244.

Holsinger 2013 <https://journals.sagepub.com/doi/10.1177/0023830913484899>

Kessler, R., Weber, A., & Friedrich, C. K. (2021). Activation of Literal Word Meanings in Idioms: Evidence from Eye-tracking and ERP Experiments. *Language and Speech*, 64(3), 594–624. <https://doi.org/10.1177/0023830920943625>

Libben, M.R., Titone, D.A. The multidetermined nature of idiom processing. *Memory & Cognition* 36, 1103–1121 (2008). <https://doi.org/10.3758/MC.36.6.1103>

Mancuso, A., Elia, A., Laudanna, A., & Vietri, S. (2020). The role of syntactic variability and literal inter-pretation plausibility in idiom comprehension. *Journal of Psycholinguistic Research*, 49(1), 99–124

McGlone, M. S., Glucksberg, S., & Cacciari, C. (1994). Semantic productivity and idiom comprehension. *Discourse processes*, 17(2), 167–190

R Core Team (2023) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna. <https://www.R-project.org/>

Rommers, J., Dijkstra, T., & Bastiaansen, M. (2013). Context-dependent Semantic Processing in the Human Brain: Evidence from Idiom Comprehension. *Journal of Cognitive Neuroscience*, 25(5), 762–776. doi:https://doi.org/10.1162/jocn_a_00337

(4) (PDF) Cognitive control in processing ambiguous idioms: evidence from a self-paced reading study. Available from: https://www.researchgate.net/publication/359409636_Cognitive_control_in_processing_ambiguous_idioms_evidence_from_a_self-paced_reading_study [accessed Jan 28 2024].

Debra A. Titone & Cynthia M. Connine (1994) Descriptive Norms for 171 Idiomatic Expressions: Familiarity, Compositionality, Predictability, and Literality, *Metaphor and Symbol*, 9:4, 247-270, DOI: 10.1207/s15327868ms0904_1

Titone, D., Lovseth, K., Kasparian, K., & Tiv, M. (2019). Are figurative interpretations of idioms directly retrieved, compositionally built, or both? Evidence from eye movement measures of reading. *Canadian Journal of Experimental Psychology / Revue canadienne de psychologie expérimentale*, 73(4), 216–230. <https://doi.org/10.1037/cep0000175>