

Влияние эмоциональной окраски слов русского языка на их визуальное узнавание в задании на лексическое решение

М.С. Власов, О.А. Сычев

Алтайский государственный гуманитарно-педагогический университет им. В.М. Шукшина

vlasov_mikhailo@mail.ru, osn1@mail.ru

Введение

Эмотивный компонент семантики лексических единиц в настоящее время представляет значительный интерес для экспериментальных исследований визуального узнавания слова. Для этих целей часто используется задание на лексическое решение (lexical decision task) с конкурирующими по эмоциональной окраске стимулами. При этом до недавнего времени лингвистические характеристики слов-стимулов (например, длина слова, частотность слова, количество «орфографических соседей» и др.) строго не контролировались исследователями, в связи с чем результаты оказывались противоречивыми. В поисках объяснения результатов экспериментов, исследователи обращаются к анализу различных лингвистических факторов.

К числу подобных лингвистических факторов, представляющих существенный интерес в контексте визуального узнавания эмоционально-окрашенных слов, относится его «семантическое богатство» (*semantic richness*) (Pexman, 2012). Под этим термином понимается многоаспектный конструкт, отражающий степень вариативности информации, связанной со значениями слова (Pexman, Hargreaves, Siakaluk, Boldner & Pope, 2008). Некоторые авторы определяют это понятие как «богатство семантической репрезентации слова» (Yar et al., 2015). В этот конструкт входит количество семантических характеристик слова, связанных с его референтом; удельный вес «семантических соседей»; количество значений слова-стимула; количество различных «первых ассоциаций», полученных методом свободного ассоциативного эксперимента; образность, то есть степень способности слова порождать ментальные образы; телесно-объектное взаимодействие, то есть степень возможного взаимодействия человеческого тела с референтом слова-стимула; степень активации словом-стимулом сенсорного или перцептивного опыта и эмоциональная валентность, противопоставленная по признакам позитивное, негативное и нейтральное.

Исходя из представлений о семантическом богатстве слова, предполагается, что стимулы с положительной и отрицательной эмоциональной окраской ассоциируются с большим объемом семантической информации, чем нейтральные стимулы, и в большей степени активируют обратную связь на уровне слова. Хотя мнение о сугубо семантической природе процесса визуального узнавания слова в настоящее время не находит поддержки, сами процессы

и механизмы, лежащие в основе этого семантического воздействия, остаются слабоизученными (Balota, Ferrano & Connor, 1991; Pexman, 2012). Кроме того, Yap и Seow (2014) считают, что подобные гипотезы невозможно проверить на уровне значений времени реакции (RTs), так как их анализ не является чувствительным к тонким аспектам принятия лексического решения. При этом в экспериментах, проведенных этими авторами, было показано, что слова с положительной и отрицательной эмоциональной окраской опознавались быстрее, чем слова с нейтральной окраской.

В одной из последних экспериментальных работ по этой проблеме (Kuperman et al., 2014) авторы на большой выборке стимулов (12 658 слов), контролируя максимальное количество лексических и семантических факторов, установили более точные эмоциональные эффекты в процессе визуального опознавания слов: независимый вклад в данный процесс вносят как эмоциональная валентность стимула, так и характер его воздействия на испытуемого («возбуждающий или успокаивающий стимул», *arousal effect*). При этом слова с отрицательной эмоциональной окраской опознаются медленнее, чем слова с положительной окраской, а «возбуждающие» стимулы – медленнее, чем «успокаивающие». Валентность и характер воздействия стимула тесно связаны с его частотностью: эффекты усиливаются для низкочастотных слов. Эти результаты свидетельствуют о том, что существуют сложные эффекты **влияния** эмоциональной окраски слов на их визуальное опознание, однако их исследование требует тщательного контроля лингвистических характеристик предъявляемых слов-стимулов.

Постановка проблемы

Представленный обзор свидетельствует о том, что эмоциональная окраска стимулов как фактор визуального узнавания привлекает значительное внимание зарубежных ученых, однако на материале русского языка подобные исследования являются редкостью. В то же время необходимо отметить, что уже сегодня создается научный задел для перспективных исследований в области обработки эмоционально окрашенной лексики русского языка. Среди таких работ – база данных *ENRuN* Д.В. Люсина и Т.А. Сысоевой (2016). Авторы указывают, что база данных содержит нормативные оценки эмоциональной окраски существительных русского языка. В ходе проведенного опроса для 378 существительных были получены оценки по эмоциональным категориям «радость», «грусть», «злость», «страх» и «отвращение», показана высокая надёжность полученных данных, проанализированы половые различия и связь оценок по различным эмоциональным категориям друг с другом. В данной базе для каждого слова по каждой из пяти эмоциональных категорий приводятся сведения о средней оценке выраженности данной эмоциональной категории, стандартном отклонении, минимальной и максимальной оценках.

В нашем исследовании на основе данных, представленных в этой базе, предпринята попытка исследовать влияние эмоциональной окраски слов на их визуальное узнавание в задании на лексическое решение на материале русского языка. В ходе такого исследования специальное внимание было уделено контролю некоторых основных лингвистических факторов (частотность слова и его длина), а также случайных факторов испытуемого и стимула/слова.

Процедура и методы исследования

Для проверки гипотезы о наличии эффекта эмоциональной окраски слова на их визуальное узнавание было организовано экспериментальное исследование с использованием методики заданий на лексическое решение.

На этапе отбора стимулов – имен существительных, с использованием базы данных *ENRuN*, контролировались следующие лингвистические характеристики слов:

- частотность слова по частотному словарю О.Н. Ляшевской и С.А. Шарова (2009), рассчитываемая как частота лексической единицы на миллион словоформ корпуса, англ. *ipm*, *instances per million*): от 40 до 80 ipm^1 ;

- длина слова в знаках (от 4 до 6) и слогах (2 слога);

- лексико-семантические характеристики слова (эмоциональная окраска): отобраны лексические единицы с максимальными значениями эмоциональной окраски по шкале «радость» и «грусть», а также слова с нейтральной окраской со значением < 2 по обеим шкалам.

В качестве псевдослов использовались сочетания из 5 букв русского языка со слоговой структурой CVCVC (например, *жэтиж*, *бунуб* и т.п.), которые не подвергались дальнейшему анализу. При этом, по возможности, псевдослова подбирались без возможных орфографических соседей из числа кодифицированных слов (то есть, должны быть не похожими ни на одно существующее слово русского языка).

Для эксперимента использовался ноутбук *ASUS Intel Core i3-4030U, 1.9GHz, RAM 4GB* и пробная версия программного обеспечения *Inquisit 5*. Экспериментальный блок состоял из 6 тренировочных проб (3 слова и 3 псевдослова), основной блок из 13 эмоционально окрашенных слов (7 – с положительной и 6 – с отрицательной окраской), 13 нейтральных слов и 26 псевдослов (стимулы и их характеристики представлены в Приложении). Все стимулы предъявлялись рандомизировано, по одному, напечатанные белым шрифтом *Arial 20 pt* по середине экрана, на черном фоне. Перед предъявлением стимула появлялся черный экран на 950 мс, затем символ “*” и сам стимул на 700 мс². После принятия решения и реакции испытуемого предъявлялась следующая проба.

В начале процедуры испытуемым предъявлялась на мониторе следующая инструкция: «В этом задании сначала в центре экрана появится звездочка *, а затем на короткое время

появится последовательность букв. Ваша задача – быстро определить, что перед Вами: настоящее, реальное слово русского языка или несуществующее слово, которого нет в русском языке. Если Вы увидите настоящее слово – нажмите клавишу «Ш», если ненастоящее слово – нажмите клавишу «У». Будьте внимательны. Постарайтесь выполнять задание быстро и точно».

В эксперименте приняли участие 56 испытуемых, 50 женского и 7 мужского пола, средний возраст 26,66 (SD=2,53), студенты факультета иностранных языков АГГПУ им. В.М. Шукшина.

Статистический анализ проводился в среде R с использованием пакетов «lme4», «lmerTest», «pbkrtest» и "MuMIn". В ходе анализа применялись линейные модели со смешанными эффектами, позволяющие учесть влияние не только фиксированных, но и случайных факторов (в задачах лексического решения последние, как правило, представлены в виде факторов испытуемого и стимула/слова) (Baayen, Milin, 2015).

Результаты исследования

Для исключения выбросов использовался способ, ранее применявшийся в целом ряде исследований времени реакции в задачах лексического решения (Yap, Seow, 2014). Из выборки были исключены все значения менее 200 мс, а также значения, отклоняющиеся более чем на 2,5 стандартных отклонения от индивидуальных средних значений ($M \pm 2,5SD$). Неверные реакции, а также реакции на псевдослова в ходе анализа не учитывались. Анализ соответствия распределения данных нормальному с помощью критерия Шапиро-Уилка показал, что у 32 испытуемых распределение времени реакции (RT) существенно отличается от нормального (при $p < 0,05$). С целью коррекции распределения было использовано обратное преобразование ($1/RT$), что позволило существенно приблизить распределение к нормальному: после преобразования, критерий Шапиро-Уилка остался значимым лишь у 13 испытуемых.

К числу типичных вмешивающихся факторов, часто оказывающих влияние на результаты в задачах лексического решения, относятся эффекты номера стимула в последовательности и времени реакции на предыдущий стимул (Baayen, Milin, 2015). Первый из этих эффектов отражает вработываемость и утомляемость испытуемых, а также, в целом, любое постепенное изменение их времени реакции в процессе эксперимента. С ним связан и второй эффект, который, вместе с тем, отражает также менее длительные процессы, объясняющие общие факторы в реакции на относительно короткие последовательности стимулов. Для контроля этих эффектов в предварительную модель были введены факторы номера пробы и времени предыдущей реакции (для первой реакции значение этой переменной было установлено равным среднему для испытуемого). Анализ этой модели показал, что статистически

значимым (при $p < 0,001$) в нашем случае является лишь эффект предыдущей реакции, поэтому данный фактор использовался для контроля во всех последующих моделях. Поскольку длина слова (число букв и слогов) и частотность контролировались при отборе стимулов, разброс по этим показателям был незначительным. Статистически значимого эффекта на время реакции эти переменные не показывали и в дальнейшем анализе не учитывались.

Таким образом, в качестве базовой модели для дальнейшего анализа использовалась модель с одним фиксированным фактором (время предыдущей реакции) и одним случайным фактором (фактор испытуемого). Доля объясняемой дисперсии для этой модели составила $R^2_{\text{общ.}} = 0,407$, при этом доля дисперсии, объясняемая единственным фиксированным фактором, составила $R^2_{\text{фикс.}} = 0,048$ (для оценки долей объясняемой дисперсии использовался метод, предложенный в работе Nakagawa и Schielzeth (2013)). В последующие модели, результаты сравнения которых представлены в табл. 1, были введены пошагово фактор эмоциональности, а также случайный фактор слова и случайные коэффициенты для фактора эмоциональности.

Таблица 1

Сравнение моделей зависимости времени реакции от эмоциональной окраски слова и случайных факторов

№	Модель		Критерий отношения правдоподобия (LRT)		Параметрический бутстреп
	Фиксированные факторы	Случайные факторы	хи-квадрат (df)	p-уровень	p-уровень
1	Время предыдущей реакции	Случайное среднее для испытуемого	-	-	-
2	Время предыдущей реакции + эмоциональность слова	Случайное среднее для испытуемого	13,912 (2)	0,0010	0,0050
3	Время предыдущей реакции + эмоциональность слова	Случайное среднее для испытуемого + случайное среднее для слова	4,349 (1)	0,0370	0,0326
4	Время предыдущей реакции + эмоциональность слова	Случайное среднее для испытуемого + случайные коэффициенты для эмоциональности слова + случайное среднее для слова	1,037 (5)	0,9595	0,7512

Результаты сравнения моделей свидетельствуют о том, что введение случайных коэффициентов для фактора эмоциональности слова (модель 4) не приводит к значимому улучшению модели. Это может означать, что эмоциональность слова показывает довольно

однородный эффект и индивидуальные различия в его проявлении несущественны. В качестве итоговой была выбрана модель 3, включающая в себя два фиксированных и два случайных эффекта. Оценка коэффициентов данной модели и их статистической значимости представлена в таблице 2.

Таблица 2

**Значения коэффициентов фиксированных эффектов для итоговой модели
и оценка их статистической значимости**

	Коэфф.	Стд. ош.	df	t-знач.	p(t)	p(z)	p(Satt)	p(KR)
Константа	1.96e-03	4.56e-05	100,91	42,96	0,0000	0,0000	0,0000	0,0000
Предшествующее ВР	-2.61e-07	2.83e-08	1640,36	-9,21	0,0000	0,0000	0,0000	0,0000
Нейтральная эмоц. слова	-2.06e-06	2.95e-05	23,78	-0,07	0,9448	0,9442	0,9448	0,9444
Позитивная эмоц. слова	7.42e-05	3.31e-05	23,49	2,24	0,0349	0,0251	0,0349	0,0271

Примечание: p(z) – оценка значимости на основе аппроксимации с помощью нормального распределения, p(Satt) - оценка значимости с расчетом степеней свободы с использованием аппроксимации по методу Саттертуэйта, p(KR) - оценка значимости с расчетом степеней свободы с использованием аппроксимации по методу Кенварда-Роджера.

Доля объясняемой дисперсии для итоговой модели в целом составила $R^2_{\text{общ.}} = 0,421$, для фиксированных факторов $R^2_{\text{фикс.}} = 0,056$. Хотя вклад двух факторов, добавленных в итоговую модели по сравнению с базовой, является статистически значимым, прирост показателей объясняемой дисперсии невелик. Этот факт можно интерпретировать как свидетельство довольно слабого влияния эмоциональности слова и случайного фактора слова на фоне весьма существенного вклада факторов испытуемого и предшествующего времени реакции. На основе коэффициентов, представленных в табл. 2 можно также сделать вывод о том, что фактор эмоциональности слова проявляется в том, что время реакции на слова с позитивной эмоциональной нагрузкой значимо меньше, чем на слова с нейтральной или негативной окраской. При этом значимых различий во времени реакции на слова с негативной и нейтральной эмоциональной окраской не выявлено.

Заключение

В результате проведенного экспериментального исследования установлено, что на материале русского языка проявляется эффект эмоциональной окраски слов на их узнавание и время реакции в задаче лексического решения, причем, наименьшее время реакции характерно для слов с позитивной эмоциональной окраской. Этот вывод хорошо согласуется с результатами экспериментальных исследований на материале английского языка (Kuperman et

al., 2014). Ограничения данного исследования связаны с отсутствием контроля некоторых психолингвистических факторов, таких как количество орфографических соседей и возраст усвоения слова-стимула. Уточнение полученных результатов при контроле таких факторов составляет перспективу дальнейших исследований по этой проблеме.

Работа выполнена при финансовой поддержке Российского гуманитарного научного фонда, проект №16-36-01112 «Эмоциональные и лингвистические предикторы времени реакции и точности выполнения экспериментальных заданий на русском языке».

Список литературы:

1. Люсин Д.В. Эмоциональная окраска имён существительных: база данных ENRuN / Д.В. Люсин, Т.А. Сысоева. – 2016 (в печати).
2. Ляшевская О.Н. Частотный словарь современного русского языка (на материалах Национального корпуса русского языка) / О.Н. Ляшевская, С.А. Шаров. – М.: Азбуковник, 2009. – 1087 с.
3. Baayen R.H. Analyzing reaction times / R.H. Baayen, P. Milin // *International Journal of Psychological Research*. – 2015. – V. 3. – N 2. – P. 12–28.
4. Balota D.A. On the early influence of meaning in word recognition: A review of the literature / D.A. Balota, F.R. Ferraro, L.T. Connor // *The psychology of word meanings*. – Hillsdale: Erlbaum. – 1991. – P.187–218.
5. Carreiras M. Effects of orthographic neighborhood in visual word recognition: cross-task comparisons / M. Carreiras, M. Perea, J. Grainger // *Journal of Experimental Psychology: Learning, Memory, and Cognition*. – 1997. – V.23. – N. 4. – P. 857-71
6. Kuperman V. Emotion and language: Valence and arousal affect word recognition / V. Kuperman, Z. Estes, M. Brysbaert, A.B. Warriner // *Journal of Experimental Psychology: General*. – 2014. – V.143. – N. 3. – P.1065-1081.
7. Nakagawa S. A general and simple method for obtaining R² from generalized linear mixed-effects models / S. Nakagawa, H. Schielzeth // *Methods in Ecology and Evolution*. – 2013. – V. 4. – N 2. – P. 133–142.
8. Jiang N. *Conducting Reaction Time Research in Second Language Studies*. / N. Jiang. – NY: Routledge. – 2012. – 282 p.
9. Pexman P.M. Meaning-based influences on visual word recognition / P.M. Pexman // *Visual word recognition: Meaning and context, individuals and development* / J.S. Adelman (Ed.). – Hove, UK: Psychology Press. – 2012. – P. 24–43.

10. Pexman P.M. There are many ways to be rich: Effects of three measures of semantic richness on visual word recognition / P.M. Pexman, I.S. Hargreaves, P.D. Siakaluk, G.E. Bodner, J. Pope // *Psychonomic Bulletin & Review*. 2008 – V. 15. P. 161–167.
11. Rayner K. *Psychology of Reading: 2nd Edition* / Rayner K., Pollatsek A., Ashby J., Clifton Ch.Jr. – NY & London: Psychology Press, Taylor & Francis Group. – 2012. – 496 p.
12. Yap M.J. Semantic richness effects in lexical decision: the role of feedback / M.J. Yap, G.Y. Lim, P.M. Pexman // *Memory & Cognition*. – 2015. – V. 43. – P. 1148-1167.
13. Yap M.J., Seow C.S. The influence of emotion on lexical processing: Insights from RT distributional analysis / M.J. Yap, C.S. Seow // *Psychonomic Bulletin & Review*. – 2013. – V. 21. – N 2. – P. 526–533.

Приложение

Таблица 1. Список слов (n=26), отобранных для эксперимента из базы данных ENRuN

(© Люсин Д.В., Сыроева Т.А.)

Слово	Частотность (ipm)	Кол-во букв	Кол-во слогов	Радость	Грусть	Эмоциональность
Отпуск	44,8	6	2	4,64	0,60	Позитивная
Отдых	57,6	5	2	4,55	0,85	Позитивная
Дружба	59	6	2	4,47	0,77	Позитивная
Талант	62	6	2	4,40	0,36	Позитивная
Мечта	57	5	2	4,34	1,34	Позитивная
Шутка	62,7	5	2	4,23	0,34	Позитивная
Танец	54,1	5	2	3,70	0,70	Позитивная
Горе	48,3	4	2	0,06	4,13	Негативная
Тоска	48,5	5	2	0,09	3,86	Негативная
Гибель	51	6	2	0,09	3,55	Негативная
Тюрьма	75,4	6	2	0,28	2,66	Негативная
Вина	54,6	4	2	0,55	2,38	Негативная
Голод	40,1	5	2	0,09	2,36	Негативная
Бандит	42,8	6	2	0,19	1,25	Нейтральная
Платок	40,3	6	2	1,19	0,91	Нейтральная
Вагон	70,9	5	2	1,04	0,87	Нейтральная
Песок	55,3	5	2	1,85	0,53	Нейтральная
Крыло	67,6	5	2	1,72	0,53	Нейтральная
Доска	67,1	5	2	0,55	0,43	Нейтральная
Стакан	72,8	6	2	1,40	0,40	Нейтральная
Забор	48,3	5	2	0,49	0,40	Нейтральная
Длина	67,7	5	2	0,81	0,34	Нейтральная
Сумка	55,5	5	2	1,68	0,30	Нейтральная
Куртка	41	6	2	1,09	0,30	Нейтральная
Буква	63,5	5	2	1,17	0,25	Нейтральная
Цифра	62,2	5	2	0,64	0,15	Нейтральная

Таблица 2. Список псевдослов, предъявляемых в эксперименте (n=26)

ЪИЪИЪ	бунуб
ЪЭЪЭЪ	бунаб
ЪЫЪЫЪ	бубаб
ЪОЪОЪ	нулаб
ЪАЪАЪ	молон
ЪУЪУЪ	нонон
ЪИЪИЪ	гэтиг
ЪЭЪЭЪ	жэтиж
ЪЫЪЫЪ	гэзэг
мунум	зэгиз
мубам	жэзиж
нунун	зэгиз
нубан	гэтэг

¹ Существуют множество экспериментальных работ, в которых показано, что в заданиях на визуальное узнавание слов высокочастотные стимулы опознаются значимо быстрее низкочастотных, при этом в одной из последних работ (Kuperman et al., 2014) показано, что эффект эмоциональной валентности слова в задании на лексическое решение значим только для слов с низкой частотностью. В связи с этим в наш экспериментальный материал включены слова с относительно низкой частотностью от 40 до 80 ipm.

² В поведенческих исследованиях, по методике LDT, часто встречается следующий диапазон выбросов, исключаемых при анализе времени реакции на слово: до 200 мс и свыше 1500 мс, при этом, как правило, считается, что для опознания слова из 6-10 знаков необходимо менее 1000 мс. (Carreiras et al., 1997; Jiang, 2012.). В работе (Rayner et al., 2012: 52) говорится о том, что в среднем на принятие лексического решения испытуемый тратит 500 мс, а на категоризацию слова – в среднем 700 мс. В связи с тем, что в нашем эксперименте задействован семантический фактор, то есть элементы категоризации лексического значения слова, было решено установить время предъявления стимула 700 мс.