

Секция «Фонетика и фонология в пространстве лингвистических исследований»

**Устойчивость графической репрезентации слова в ментальном лексиконе:
экспериментальное исследование на материале французского языка**

Коваленко Елизавета Сергеевна

Студент (бакалавр)

Национальный исследовательский университет «Высшая школа экономики» -
Санкт-Петербург, Санкт-Петербургская школа социальных и гуманитарных наук,
Санкт-Петербург, Россия

E-mail: eskovalenko@edu.hse.ru

Орфографические ошибки в большинстве случаев рассматриваются как следствие слабого орфографического представления в ментальном лексиконе [Perfetti, Hart 2002: 189]. В настоящей статье будет рассмотрено, как происходит распознавание слова и лексический доступ у носителя французского языка, в каких словах он чаще всего допускает ошибки при декодировании написанного, сколько времени требуется на идентификацию слова, написанного как правильно, так и с орфографической ошибкой.

В 2017 году Садаф Рахманьян и Виктор Куперман с помощью метода регистрации движения глаз и задачи лексического решения провели два экспериментальных исследования, цель которых состояла в том, чтобы проверить, как орфографические ошибки, которые часто встречаются в текстах, влияют на структуру ментального лексикона носителя языка [Rahmanian, Kuperman 2017: 7]. На материалах английского языка и иврита было доказано, что частотные альтернативные варианты написания слов могут влиять на визуальное декодирование правильно написанных единиц, поскольку они увеличивают энтропию в ментальном лексиконе и размывают орфографические представления. В ментальном лексиконе фиксируется более одного варианта написания слова: правильный и с ошибкой, что приводит к затруднению при идентификации слова [Kuperman 2017, 2021]

В 2021 году Виктор Куперман совместно с коллегами провел еще одно исследование, которое основывалось на изучении правильных и неправильных вариантов написания слов, представляющих собой альтернативные факторы по значению и произношению (например, 'tomorrow' и 'tommorrow') [Kuperman 2021: 11]. Так, было выявлено, что высокая энтропия приводит к более медленному распознаванию слов, даже если они написаны правильно, что было показано на материалах английского, китайского, греческого, финского языков и иврита. Важно также отметить, что и высокочастотные, и низкочастотные слова распознавались медленно при наличии частотных вариантов ошибочного написания у этих слов.

Обоснование исследования на материале французского языка

На материале французского языка мы провели аналогичное исследование, для которого характерны иные типы орфографических ошибок, например, ошибки, связанные с диакритическими знаками, с дифтонгами или трифтонгами, с «немыми» гласными или согласными в конце слов. Однако, прежде всего, стоит отметить, что во французском языке существует множество слов, орфограммы которых соответствуют определенным орфографическим подходам, в результате чего часто возникают сложности с восприятием слов или предложений на слух и их дальнейшей графической репрезентацией. Так, можно отметить, что наиболее частотные орфографические ошибки встречаются в словах с диакритическими знаками ('évènement' / 'avènement'), с диграммами (<eu> может соответствовать [ø] - 'deux', [œ] - 'jeûne', [y] - (il) 'eut') или триграммами (<œu> может соответствовать [œ] - 'cœur' или [ø] - 'vœu'), с удвоенными согласными ('pâte' / 'patte').

Следует добавить, что носители французского языка также допускают орфографические ошибки из-за лишних добавлений букв, ('cauchemar(d)', 'cat(h)échisme'), неправильного порядка букв ('infarctus' / 'infractus') или отсутствия окончания 'gagueure' / 'gagueur').

Важно отметить, что частотность орфографических ошибок в последнее десятилетие увеличивается из-за развития различных средств массовой информации. Это связано с тем, что раньше число редактируемых текстов (книги, журналы, газеты) превалировало над неотредактируемыми текстами (личные письма), так как последних было намного меньше. К XXI в. появился Интернет, различные социальные сети, блоги, и к настоящему времени существенно выросла доля текстов, не подвергающейся профессиональной редакции и корректуре. Стоит отметить, что подобная разница в воздействии печатных и электронных изданий может быть важным источником индивидуальных различий в когнитивных способностях человека. Кит Станович указал, что печатные издания содержат дескриптивные знания, которые часто недоступны в электронных средствах массовой информации или устных источниках [Stanovich 1993: 146]. Печатные издания являются более богатым источником для расширения словарного запаса, чем устная речь или телевидение [Hayes, Ahrens 1988: 399]. Так, на примере нашего эксперимента мы рассмотрим, как это отражается на работе ментального лексикона.

Материал

Для начала эксперимента было отобрано 30 слов (список слова представлен в таблице 1), в которых часто допускают орфографические ошибки. Далее, для каждого слова была подобрана 'пара', в которой допущена орфографическая ошибка. Для нахождения частотности употребления и подбора 'пары' использовалась лингвистическая база данных для слов французского языка - Sketch Engine frTenTen: Corpus of the French Web [URL: <http://www.sketchengine.eu/frtente-french-corpus/>]. Так, было составлено два протокола: в первом протоколе половина слов была написана правильно, половина - с ошибкой; во втором - половина верных, которые в первом протоколе были неправильными, половина с ошибкой, которые были правильными в первом случае. Стоит добавить, что основными орфографическими ошибками, представленными в стимульном материале, были пропуск букв, вставка лишней буквы, замена одной буквы на другую и ошибка в надстрочном знаке - диакритике. Исследование проводилось с помощью методики орфографического решения (spelling decision task).

Участники и процедура экспериментальной части

Для проведения экспериментальной части была выбрана платформа IBEX FARM. [URL: <https://spellout.net/ibexfarm>]. В исследовании приняли участие 36 человек (19 женщин и 17 мужчин). Все испытуемые являлись носителями французского языка. Их возраст варьировался от 16 до 49 лет. Важным пунктом прохождения эксперимента было то, чтобы испытуемые профессионально не работали с текстами. Во время прохождения эксперимента испытуемым предъявлялись отдельные слова, и они должны были решить, правильно ли написано слово, или в нем допущена орфографическая ошибка. Программа регистрировала ответ и время реакции.

Результаты и обсуждение

Было выявлено, что вероятность дать правильный ответ зависит от частотности слова ($Z = 2,903$, $p\text{-value} = 0,0037$).

Что касается слов без орфографических ошибок, к которым были даны правильные ответы, то на скорость их распознавания значимо влияет взаимодействие двух факторов: частотности слова и энтропии написаний, то есть соотношения верных и неверных

написаний слова в корпусе: $b = -63,264$; $SE = 28,429$; $t\text{-value} = -2,225$; $p\text{-value} = 0,0267$. Наличие большого количества неверных написаний слова в корпусе текстов размывает его графический облик в ментальном лексиконе носителей языка, особенно неустойчивы орфографические репрезентации слов с низкой частотностью.

В заключение мы решили проверить, какой тип орфографической ошибки наиболее выражен в неправильно написанных словах: было выявлено, что реже всего участники распознают ошибки в словах с заменой или добавлением букв, а чаще всего - ошибки, связанные с диакритическими знаками.

Источники и литература

- 1) Hayes D. P., Ahrens M. Vocabulary simplification for children: A special case of motherese? In: Journal of Child Language, 1988, Vol. 15. Pp. 395–410.
- 2) Kuperman V., Kyröläinen A-J. Prevalence of spelling errors affects reading behavior across languages. In: Journal of Experimental Psychology, 2021, Vol. 23. Pp. 1–52.
- 3) Perfetti C. A., Hart L. The lexical quality hypothesis. In: Precursors of functional literacy, 2002, Vol. 11. Pp. 67–86.
- 4) Rahmanian S., Kuperman V. Spelling errors impede recognition of correctly spelled word forms. In: Scientific Studies of Reading, 2017, Vol. 23(1). Pp. 24–36.
- 5) Stanovich K. E. Does reading make you smarter? Literacy and the development of verbal intelligence. In: Advances in child development and behavior, 1993, Vol. 24. Pp. 133–180.

Иллюстрации

Таблица 1

Слово	Количество букв	Частотность, ipm	«Пара»	Частотность пары, ipm
Recueillir	10	61525	Receuilir	304
Siphon	6	9037	Syphon	506
Obédience	9	9079	Obédiance	56
Accueillir	10	245758	Acceuilir	1992
Gueuler	7	4141	Geuler	95
Sapharmaüm	10	2756	Capharmaüm	6
Déménager	9	21608	Déménager	85
Bouilloire	10	9926	Bouilloire	31
Enthousiasme	12	83256	Entousiasme	186
Rébellion	9	40999	Rébelion	13
Déshonorer	10	1330	Deshonorer	23
Catéchisme	9	13731	Cathechisme	17
Cauchemar	9	37015	Cauchemard	1039
Dégingandé	10	400	Déguingandé	19
Gageure	7	5679	Gageur	64
Infarctus	9	15013	Infractus	83
Abasourdi	9	2896	Abassourdi	28
Hypnotisé	9	1922	Hynoptisé	6
Aborigène	9	4026	Arborigène	29
Effervescence	13	16765	Effervescence	8
Dilemme	7	19031	Dileme	53
Métonymie	9	3557	Métonimie	5
Fatigant	9	5591	Fatigant	4303
Extravagant	11	5122	Extravaguant	97
Convaincant	11	16822	Convainquant	3330
Agressif	8	30892	Aggressif	589
Nourrir	7	107171	Nourir	747
Omettre	7	7330	Ommettre	44
Absence	7	474113	Absense	68
Évènement	9	344548	Èvènement	135303

Рис. 1. Стимульный материал