

Сказал А, говори и Б:
экспериментальное исследование
обработки биномиалов при чтении



Введение

Подходы к восприятию устойчивых выражений

Некомпозиционный
каждое устойчивое словосочетание обрабатывается как единое целое

[Tremblay et al., 2011]

Композиционный
отвергается идея об особом статусе шаблонных выражений в ментальном лексиконе

[McKinnon et al., 2003]

Гибридный
шаблонные выражения - единицы с двойственной природой

[Fan, Wang, 2024]

Цель и гипотезы

Цель работы – выявить особенности обработки биномиалов русского языка как особой единицы ментального лексикона и определить факторы, влияющие на этот процесс.

Гипотезы:

- Высокочастотные биномиалы обрабатываются как целостные языковые единицы;
- Порядок следования компонентов биномиала влияет на процесс его обработки и на успешность восстановления.



Корпусное исследование

- Набор стимулов для исследования производился из Национального корпуса русского языка из раздела Основной корпус (<https://ruscorpora.ru/>).
- Итоговый список состоит из **325 биномиалов** с разной частотой встречаемости выражения в корпусе.
- В экспериментальный список вошло **50 высокочастотных** биномиалов с пороговым значением $\text{ipm} > 0.5$.

Таблица 1. Фразовая частотность биномиала в прямом и обратном порядке, частотность его компонентов

№	Биномиал	Частотность биномиала в НКРЯ		Частотность слова 1 в НКРЯ	Частотность слова 2 в НКРЯ
		ПП	ОП		
1	добра и зла	1 198	43	11 799	11 321
2	жизни и смерти	1 069	32	286 611	70 371
3	мужчина и женщина	585	54	20 110	61 241

- Между порядком слов, частотностью слов и частотностью всего выражения нет прямой зависимости.
- Проведен анализ степени близости компонентов биномиала. Все показатели статистических мер получили положительные оценки, что говорит о неслучайности связи между компонентами.

Меры статистической значимости:

$$MI = \log_2 \frac{f(n,c) \times N}{f(n) \times f(c)}$$

$$\logDice = 14 + \log_2 \frac{2f(n,c)}{f(n) \times f(c)}$$

$$\log\text{-likelihood} = 2 \sum f(n,c) \times \log_2 \frac{f(n,c) \times N}{f(n) \times f(c)}$$

$$t\text{-score} = \frac{f(n,c) - \frac{f(n) \times f(c)}{N}}{\sqrt{\frac{f(n) \times f(c)}{N}}}$$

n — ключевое слово; c — коллокат;
f(n,c) — частота встречаемости ключевого слова n в паре с коллокатом c;
f(n), f(c) — абсолютные частоты ключевого слова n и коллоката c в корпусе;
N — общее число словоформ в корпусе.

Выводы

Механизмы обработки биномиалов

СИНТАКСИЧЕСКИЙ
уровень
Некомпозиционный
подход



ЛЕКСИЧЕСКИЙ
уровень*
Композиционный
подход

различия в скорости прочтения, количестве фиксаций, вероятности регрессий при изменении порядка следования компонентов биномиала

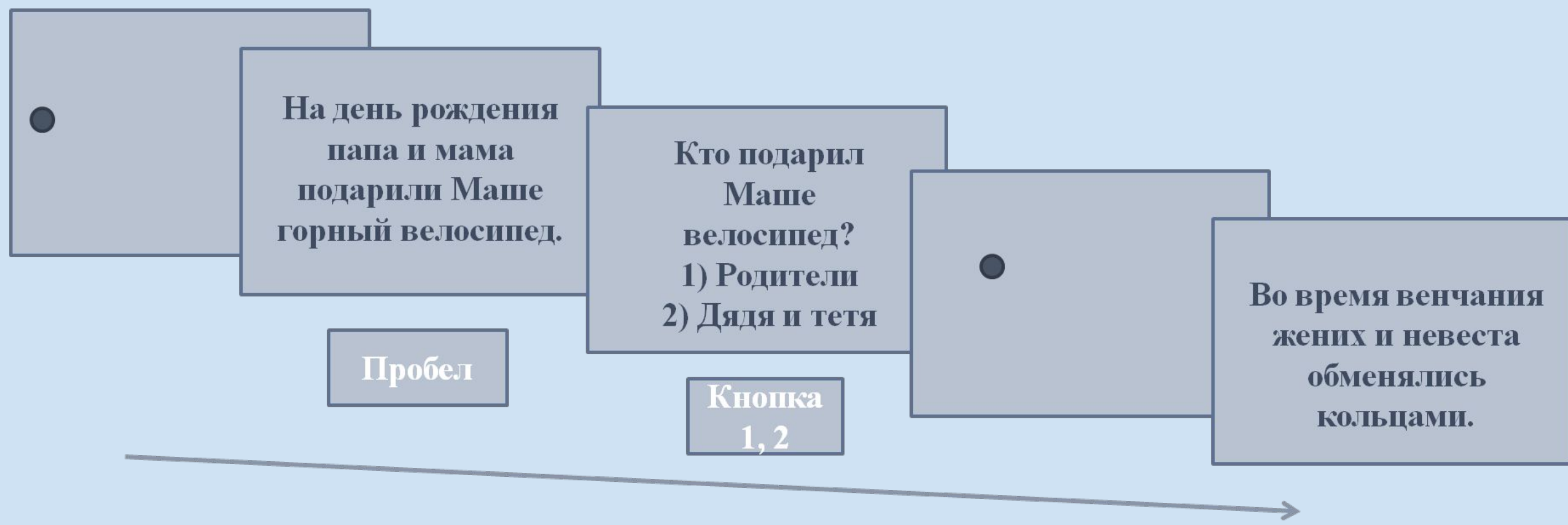
носители языка чувствительны к вариативности компонентов

*** Эксперимент 2**

Перспективой дальнейших исследований является изучение низкочастотных биномиалов и проверка гипотезы о влиянии реверсивности низкочастотного биномиала и частотности каждого из его компонентов на динамику зрительного восприятия.

Эксперимент 1: an eye-tracking study

- Выборка:** В исследовании приняли участие 64 носителя русского языка в возрасте от 18 до 37 лет (M=22,3), 27 мужчин и 37 женщин.
- Стимульный материал:** 50 высокочастотных биномиалов в составе предложения в двух условиях: предпочтительный и обратный порядок следования компонентов.
- В среднем эксперимент занимал 25-35 минут.
- Техническое обеспечение: SR Research Eyelink 1000 Plus.



- На всех этапах обработки **биномиалы с предпочтительным порядком обрабатываются быстрее**, чем биномиалы с обратным порядком следования компонентов.
- Количество фиксаций** на биномиалах с предпочтительным порядком **было меньше**, чем на биномиалах с обратным.
- Вероятность регрессий** к биномиалу в обратном порядке **выше**.
- Когнитивная нагрузка (время прочтения, кол-во фиксаций, вероятность регрессий) **Слова 2** увеличивается. Эффект не обнаружен для Слова 1 и Слова “и”.

Коэффициент обратимости отражает, насколько устойчив данный порядок слов и допустим ли другой. Изменение порядка слов у **необратимых биномиалов** значимо влияет скорость их прочтения: биномиалы с предпочтительным порядком следования компонентов прочтываются быстрее ($p < .05$). Схожий эффект не обнаружен для обратимых биномиалов ($p > .05$).

Таблица 2. Показатели коэффициента в зависимости от типа задачи

Коэффициент обратимости	Тип задачи	Среднее время прочтения биномиала, мс
>90	ПП	665.12
>90	ОП	698.07
<80	ПП	690.17
<80	ОП	717.68

Список литературы

McKinnon R., Allen M., Osterhout L. Morphological decomposition involving non productive morphemes: ERP evidence. 2003. Neuroreport 14, 883–886 10.1097/00001756-200305060-00022

Bobrow S.A., Bell S.M. On catching on to idiomatic expressions // Memory and Cognition. Vol. 1. № 3. 1973. P. 343–346.

Jiang N., Nekrasova T. M. The processing of formulaic sequences by second language speakers. 2007. Mod. Lang. J. 91, 433–445. doi: 10.1111/j.1540-4781.2007.00589.x

Malkiel Y. Studies in irreversible binomials // Lingua 8, – 1959. – P. 113–160.

Mollin S. The (Ir)reversibility of English Binomials: Corpus, constraints, developments // John Benjamins B.V. 2014. P. 265.

Sprenger S. A., Levelt W. J. M., Kempen G. Lexical access during the production of idiomatic phrases. Journal of Memory and Language. 2006. Vol. 54. № 2. P. 161–184.

Tremblay A., Derwing B., Libben G., and Westbury, C. Processing advantages of lexical bundles: evidence from self-paced and sentence recall task. 2011. Lang. Learn. 61, 569–613. doi: 10.1111/j.1467-9922.2010.00622.x

Работа выполнена при поддержке СПбГУ, шифр проекта 124032900009-2