

Сочетания местоимений *что-то* и *кто-то* с полными прилагательными в свете интуиции носителей языка

Буров Эдуард Евгеньевич, zhenyabur@mail.ru
Иркутский государственный университет

И. Б. Левонтина отмечает: «Априорно нельзя утверждать, что результаты анализа корпуса (даже если он достаточно большой и претендует на репрезентативность) отражают интуицию носителей языка. Иными словами, языковая интуиция (“patterns of knowledge”) не может быть приравнена к “patterns of use” (как они, в частности, фиксируются в корпусах). <...> Таким образом, может оказаться, что корреляции между корпусными данными и языковой интуицией носителя довольно сложны» [Левонтина, 2022, с. 16].

Представляется интересным провести эксперимент, направленный на исследование связи между корпусными данными и интуицией носителей языка. Хорошую возможность сделать это дают сочетаемостные свойства местоимений-существительных *что-то* и *кто-то*. Как известно, «*что-то* часто употребляется в контексте прилагательного» [Падучева, 2016]. То же, хотя и в меньшей степени, относится к местоимению *кто-то*. Корпусный анализ показывает, что одни сочетания типа «*что-то* / *кто-то* + полное прилагательное» являются статистически устойчивыми¹ (*что-то* знакомое, *что-то* белое, *что-то* грустное, *что-то* вкусное; *кто-то* чужой, *кто-то* огромный, *кто-то* страшный, *кто-то* умный), а другие – нет (*что-то* популярное, *что-то* бесцветное, *что-то* нерадостное, *что-то* сытное; *кто-то* свой, *кто-то* низкий, *кто-то* красивый, *кто-то* глупый).

Эксперимент позволяет проверить, как соотносятся между собой фиксируемая в корпусе устойчивость сочетания типа «*что-то* / *кто-то* + полное прилагательное» и приемлемость этого сочетания для носителей языка.

Гипотеза состоит в том, что между корпусными данными и интуицией носителей языка существует прямая связь, и предложения со статистически устойчивыми сочетаниями значимо более приемлемы, чем предложения с сочетаниями без статистической устойчивости. Подтверждение этой гипотезы означало бы, что носители языка способны не только отличать противоречащее языковым правилам от не противоречащего им (например, воспринимать сочетание *что-то* новое как правильное, а *что-то* новая – как неправильное²), но и определять, что внутри соответствующего языковым правилам более характерно для русской речи, а что – менее.

Было проведено два эксперимента: в первом исследовались сочетания типа «*что-то* + полное прилагательное», во втором – типа «*кто-то* + полное прилагательное».

В ходе первого эксперимента носители языка оценивали предложения, в которых были представлены либо статистически устойчивые сочетания типа «*что-то* + полное прилагательное» (*Он напевал что-то знакомое; У него в руке было что-то белое; Произошло что-то грустное; Мне дали съесть что-то вкусное*), либо статистически неустойчивые сочетания того же типа (*Он напевал что-то популярное; У него в руке было что-то бесцветное; Произошло что-то нерадостное; Мне дали съесть что-то сытное*). В эксперименте, таким образом, исследовался фактор

¹ Ср. понятие коллокации в [Захаров, Хохлова, 2010, с. 137].

² См. о различных ограничениях на сочетаемость лексемы, в том числе морфосинтаксических, в [Апресян, 1974, с. 60-62].

СТАТИСТИЧЕСКАЯ УСТОЙЧИВОСТЬ СОЧЕТАНИЯ, который имел 2 уровня: ВЫСОКАЯ СТЕПЕНЬ СТАТИСТИЧЕСКОЙ УСТОЙЧИВОСТИ / ОТСУТСТВИЕ СТАТИСТИЧЕСКОЙ УСТОЙЧИВОСТИ. Каждое из двух экспериментальных условий было представлено четырьмя лексикализациями.

Участниками эксперимента стали 23 человека – в основном студенты факультета иностранных языков Института филологии, иностранных языков и медиакоммуникации Иркутского государственного университета. Эксперимент был реализован на платформе PCibex [<https://farm.pcibex.net>] с использованием методики оценки приемлемости по шкале Ликерта от 1 (плохое предложение) до 7 (хорошее предложение). Предложения распределялись по экспериментальным листам в соответствии с правилом латинского квадрата (2 листа, в каждом по 4 собственно экспериментальных, или стимульных, предложения).

В экспериментальных листах содержались как собственно экспериментальные предложения, так и филлеры (грамматически приемлемые). В филлерах местоимение *что-то* употреблялось без зависимых прилагательных. Филлеры и стимульные предложения были представлены в соотношении 1 : 1.

Эксперимент распространялся через социальные сети и электронную почту. Анализ данных проводился в среде разработки RStudio с использованием языка программирования R.

В результате эксперимента выяснилось, что средняя нормализованная оценка предложений со статистически устойчивыми сочетаниями типа «*что-то* + полное прилагательное» составляет 0,19, а средняя нормализованная оценка предложений со статистически неустойчивыми сочетаниями того же типа – -0,22. При этом разница между оценками, согласно однофакторному дисперсионному анализу, статистически значима ($df = 1$; $F = 4,413$; $p = 0,0385$).

Второй эксперимент был устроен аналогично первому. В нем принял участие 21 человек. В ходе эксперимента предложения со статистически устойчивыми сочетаниями типа «*кто-то* + полное прилагательное» получили среднюю нормализованную оценку -0,15, а предложения со статистически неустойчивыми сочетаниями того же типа – -0,57. Разница между оценками, как и в случае первого эксперимента, статистически значима (однофакторный дисперсионный анализ: $df = 1$; $F = 4,268$; $p = 0,042$).

Таким образом, исходная гипотеза полностью подтверждается. Носители языка действительно воспринимают как менее приемлемые те сочетания, которые менее характерны для русской речи. Наблюдается прямая связь между корпусными данными и языковой интуицией носителей.

Прямая связь между “patterns of knowledge” и “patterns of use” обнаруживается также и в том, что предложения первого эксперимента получили в среднем более высокие оценки, чем предложения второго. Это коррелирует со следующим фактом: в корпусе большей склонностью употребляться с постпозитивными полными прилагательными обладает местоимение *что-то*.

Литература

- Апресян Ю.Д. Лексическая семантика. Синонимические средства языка. – М.: Наука, 1974. – 368 с.
- Левонтина И.Б. Частицы речи. – М.: Азбуковник, 2022. – 431 с.
- Захаров В.П., Хохлова М.В. Анализ эффективности статистических методов выявления коллокаций в текстах на русском языке // Компьютерная лингвистика и интеллектуальные технологии. – Вып. 9 (16). – М.: Изд-во РГГУ, 2010. – С. 137–143.

Падучева Е.В. Местоимения неизвестности. Материалы для проекта корпусного описания русской грамматики (<http://rusgram.ru>). На правах рукописи. М. 2016.