

Научная статья

УДК 811.112.2

DOI 10.25205/1818-7935-2022-20-3-123-138

Влияние типа ввода на эффективность усвоения коллокаций (корпус vs словарь)

Ольга Юрьевна Савина

Тюменский государственный университет
Тюмень, Россия

osawina@gmail.com, <https://orcid.org/0000-0002-4777-3188>

Аннотация

Настоящее исследование посвящено овладению коллокациями при обучении неродным языкам. Цель исследования – выяснить, зависит ли эффективность усвоения коллокаций от когнитивных усилий обучающихся, приложенных в момент получения информации о коллокации в разных условиях ввода – деконтекстуализированного (в словарях), обогащенного или усиленного (в корпусах). Обогащение представляет собой увеличение частоты целевого объекта, а усиление – более заметное его представление с помощью типографских приемов. Предполагалось, что обработка обогащенных и усиленных корпусных данных потребует больших когнитивных усилий и, следовательно, обеспечит более эффективное произвольное запоминание, чем обработка деконтекстуализированных данных из словарей. Для проверки данной гипотезы был проведен долгосрочный лабораторный эксперимент с повторными измерениями.

Исследование проводилось на материале немецкого языка. Эксперимент включал следующие этапы: (0) срез знаний – определение уровня владения коллокациями; (1) непосредственное тестирование – определение типа ввода информации о коллокациях, обеспечивающего наиболее эффективное понимание; (2) отложенное тестирование – определение типа ввода, обеспечивающего более прочное долговременное произвольное запоминание коллокаций; (3) сравнение результатов непосредственного и отложенного тестирования. Для анализа данных, собранных в ходе эксперимента, использовались статистические методы: непараметрические критерии Краскела — Уоллиса и Манна — Уитни, а также сила эффекта Cohen's d.

Результаты исследования показали, что обработка обогащенных и усиленных корпусных данных, требующая больших когнитивных усилий, обеспечивает более эффективное произвольное запоминание коллокаций, чем обработка деконтекстуализированных данных из словарей, и что этот эффект сохраняется долгосрочно.

Ключевые слова

коллокации, учебная лексикография, корпусная лингвистика, корпусная лингводидактика, обучение на основе данных, немецкий язык

Для цитирования

Савина О. Ю. Влияние типа ввода на эффективность усвоения коллокаций (корпус vs словарь) // Вестник НГУ. Серия: Лингвистика и межкультурная коммуникация. 2022. Т. 20. № 3. С. 123–138. DOI 10.25205/1818-7935-2022-20-3-123-138

Благодарности

Автор выражает признательность своим студентам за участие в эксперименте, Ирине Геннадиевне Горбачевой за ценные консультации по статистическому анализу, а также редакторам и анонимным рецензентам за полезные замечания к более ранним версиям этой статьи.

© Савина О. Ю., 2022

Influence of the Input Type on the Efficiency of Collocations' Retention (Corpus vs Dictionary)

Olga Yu. Savina

University of Tyumen
Tyumen, Russian Federation

osawina@gmail.com, <https://orcid.org/0000-0002-4777-3188>

Abstract

The present research investigates collocation learning strategies in the process of foreign language teaching. The paper aims at analyzing whether the process of effective collocation retention is determined by learners' cognitive effort made at the moment of receiving information about a collocation in various input conditions: decontextualized (in dictionaries), enriched or enhanced (in corpora). Enrichment is viewed as an increase in frequency of the target feature while enhancement consists in a more conspicuous representation of the feature by typographical means.

It is assumed that processing enriched and enhanced corpus data requires higher cognitive effort and, consequently, provides stronger incidental retention compared to analyzing decontextualized dictionary data. To test this hypothesis, we conducted a longitudinal laboratory experiment with repeated measures.

The research focuses on the German language. The experiment included the following stages: (0) a pre-test in order to determine the students' collocation competence; (1) an immediate test targeted at finding out which information about collocations is perceived easier; (2) a post-test to determine the type of input providing stronger long-term involuntary collocation retention; (3) comparison of the immediate test and post-test results. In analyzing experimental data statistical methods have been used: the Kruskal—Wallis and the Mann—Whitney tests for nonparametric data, as well as the Cohen's d effect size test.

The research demonstrates that mental processing of enriched and enhanced corpus data, requiring a higher cognitive effort, provides better incidental collocation retention than processing of decontextualized dictionary data and has a long-term effect.

Keywords

collocation, educational lexicography, corpus linguistics, corpus-based teaching, data-driven learning, german language

For citation

Savina, O. Yu. Influence of the Input Type on the Efficiency of Collocations' Retention (Corpus vs Dictionary). *Vestnik NSU. Series: Linguistics and Intercultural Communication*, 2022, vol. 20, no. 3, pp. 123–138. DOI 10.25205/1818-7935-2022-20-3-123-138

Acknowledgments

The author would like to express her gratitude to the students participating in the experiment, Irina Gorbacheva for valuable advice concerning statistical analysis, the editors and anonymous reviewers for their helpful comments on the earlier versions of this paper.

Введение

Извлечение коллокаций из контекстов или справочных материалов – далеко не тривиальная задача, так как разные источники по-разному предъявляют коллокации, в том числе с точки зрения их визуального представления. Поиск наиболее эффективных путей получения информации о коллокациях и определение факторов, играющих роль в формировании коллокационной компетенции при изучении неродных языков, определяют актуальность настоящего исследования. На необходимость решения проблем в данной области лингводидактики указывают исследователи в различных странах (см. раздел «Результаты эмпирических исследований»).

Цель настоящего исследования – выяснить, зависит ли эффективность сохранения коллокаций в долговременной памяти обучающихся от когнитивных усилий, приложенных в момент получения информации о коллокации в разных условиях ввода: деконтекстуализированного (в словаре), обогащенного или усиленного (в корпусе).

1. Лингводидактический взгляд на коллокации и их представление в словарях и корпусах

Несмотря на разнообразие определений понятия «коллокация», обусловленное различными направлениями исследования, большинство авторов выделяют совместную встречаемость слов как характерную черту коллокаций [Хохлова, 2008, с. 343; Quasthoff, 2011; Павельева, 2016, с. 57; Павловская, Горина, 2017, с. 135; Zaabalawi, Gould, 2017, p. 21]. При этом обычно выделяют так называемую базу, то есть семантически независимое слово (например, *Kaffee, Bett*) и коллокатор – семантически зависимую составляющую (например, *stark, machen*) [Steyer, 2008, p. 187], которая не может быть выбрана, переведена (и выучена) независимо от базы [Hausmann, 2007, p. 218].

Эту семантическую зависимость впервые описал Ф. Й. Хаусманн (1984) [см. Targońska, 2019, p. 181]. Вслед за этим ученые выделили два подхода, ныне общепринятых, к определению коллокаций: количественный, по частотности совместной встречаемости слов, и семантический (или фразеологический), по конвенциональной зависимости коллоката.

Семантический подход фокусируется на установлении семантических отношений между словами (двумя или более) [Gablasova et al., 2017, p. 158]. В то же время, по мнению многих исследователей, коллокации лишь в малой степени проявляют, если проявляют вообще, семантический сдвиг, который является ярким признаком истинных фразеологизмов [Steyer, 2008, p. 187; Reder, 2011, p. 136]. Таким образом, они занимают промежуточную позицию между словом / свободным словосочетанием и идиомой, имеют аналитический характер и благодаря этому относительно легки для запоминания [Hausmann, 2007, p. 229; Reder, 2011, p. 137].

Количественный подход характерен для современной корпусной лингвистики, в рамках которой данные о частотности легко доступны [Reder, 2011, p. 133]. Очевидное достоинство количественного подхода состоит в том, что обучающиеся могут самостоятельно находить коллокации в корпусах [Reder, 2011, p. 133]. Однако, по заключению И. Таргоньской, количественный подход характерен только для дидактики английского как иностранного [Targońska, 2019, p. 182]. В учебниках по немецкому как иностранному чаще всего реализуется семантический подход к коллокациям [Reder, 2011, p. 138]. Первый словарь коллокаций для немецкого языка был составлен У. Квастхоффом в 2011 г. [см. Targońska, 2014, p. 137]. Он включает коллокации для более чем 3 200 частотных существительных, глаголов и прилагательных [Quasthoff, 2011, p. 5].

Считается, что владение коллокациями повышает беглость речи, поскольку если носители языка хранят свой словарный запас в виде более крупных фрагментов, которые можно извлечь из памяти целиком, то это повышает скорость обработки информации по сравнению с его сохранением в виде отдельных слов. Некоторые ученые утверждают, что существует высокая корреляция между общим уровнем владения языком и владением коллокациями [Rahimi, Momeni, 2012, p. 41]. Затраты времени на целенаправленное выучивание коллокаций оправдываются приростом беглости и идиоматичности речи [Nation, 2006, p. 517; Frankenberg-Garcia, 2017, p. 4].

Однако в связи с грамматической и лексической непредсказуемостью таких комбинаций с точки зрения другого языка, являющейся типичной причиной интерференционных ошибок, овладение коллокациями требует особого внимания [Nation, 2006, p. 525; Möhring, 2011, p. 34; Reder, 2011, p. 137]. На начальных этапах освоения иностранного языка обучающиеся плохо осознают необходимость запоминания коллокаций как готовых лексических блоков, которые должны храниться в памяти целиком, и, следовательно, не уделяют достаточного внимания овладению ими [Zaabalawi, Gould, 2017, p. 27]. Довольно часто они комбинируют слова, подспудно руководствуясь правилами родного языка, и могут создать лексически и даже грамматически корректные комбинации, но не соответствующие узусу носителей языка, т. е. неидиоматичные [Boulton, De Cock, 2016, p. 6; Zaabalawi, Gould, 2017, p. 21].

Несмотря на важную роль коллокаций при изучении неродного языка и объективные трудности в овладении ими, они нечасто становятся объектом обучения [Peters, 2012, p. 65; Szudarski, Carter, 2016, p. 246; Targońska, 2019, p. 183]. В лингводидактике немецкого как иностранного коллокации впервые заняли самостоятельное место лишь в публикации 2010 года *Deutsch als Fremdsprache: Ein internationales Handbuch* К. Фандриха и Х.-Ю. Кrumma [см. Targońska, 2014, p. 135].

В то же время некоторые исследователи выделяют специфическую коллокационную компетенцию. И. Таргоньска видит в ее структуре три составляющих: рецептивную, продуктивную и рефлексивную. Рецептивная подразумевает способность понимать значение коллокации, а также идентифицировать неправильные варианты коллокаций; продуктивная – их корректное использование, а также способность исправлять узуально неправильные комбинации слов; рефлексивная – способность видеть неизвестные коллокации в тексте. Соответственно, последняя является базовой для развития первых двух компетенций [Targońska, 2019, p. 184–185].

Деконтекстуализированная форма представления коллокаций в словаре мало способствует их пониманию и запоминанию, так как обычно сводится к перечислению [Sonbul, Schmitt, 2013, p. 128]. Восприятие коллокаций в словаре не поддерживается никакими визуальными приемами, тогда как для развития базовой рефлексивной коллокационной компетенции требуется именно так называемый обогащенный (*enriched*) и усиленный (*enhanced*) ввод [Steyer, 2008, p. 193].

Обогащение входных данных определяется как увеличение частоты предъявления целевого объекта во входном речевом сигнале, а усиление входных данных подразумевает более наглядное представление коллокаций, что достигается с помощью самых различных шрифто-типографских средств (жирный шрифт, подчеркивание, цветная маркировка, облако слов, формат KWIC – в виде фрагментов контекстов, центрированных по искомому слову, и т. д.) [Vyatkina, 2016a, p. 208]. В облаке слов более высокая частотность представлена большим размером шрифта, что привлекает внимание к типичным коллокаторам. Кроме того, форма облака позволяет разместить больше примеров коллокаций для одного слова, чем традиционное представление [Geyken et al., 2008, p. 373].

Широкие возможности для обогащенного и усиленного ввода информации предоставляют корпуса [Roberts et al., 2018, p. 2]. Использование корпусных данных в процессе обучения неродным языкам получило название «обучение на основе данных» (англ. *Data-Driven Learning*, или *DDL*; нем. *datengesteuertes Lernen*). Эта область лингводидактики в настоящее время активно развивается.

2. Результаты эмпирических исследований и вопросы, остающиеся открытыми

Эмпирические исследования последних лет, выполненные учеными разных стран в области «корпусной лингводидактики» [Горина, Царакова, 2021, с. 37], позволяют сделать следующие выводы.

Эксперименты с отложенным тестированием подтверждают, что усиленный ввод более эффективен, чем обогащенный, но не эффективнее деконтекстуализированного. При этом эффект усиленного ввода растет с ростом уровня владения языком [Sonbul, Schmitt, 2013], а комбинация обогащения и усиления ввода дает более высокие результаты по запоминанию и распознаванию коллокаций [Szudarski, Carter, 2016, p. 261–262].

Инструкции учителя влияют на концентрацию внимания обучающихся меньше, чем визуальный акцент в тексте (жирный шрифт или подчеркивание). Более того, усиление ввода побуждает обучающихся обращаться к поиску дополнительной информации о лексических единицах [Peters, 2012].

Несмотря на накопленные результаты, все еще существуют значительные пробелы, требующие внимания. Так, довольно мало исследований посвящено тому, каким путем обучающиеся получают знания о коллокациях [Sonbul, Schmitt, 2013, p. 126]. И. Таргоньска призывает исследовать способы развития коллокационной компетенции и факторы, играющие роль в ее формировании [Targońska, 2014, p. 136, 142].

Малоизученным также остается вопрос об эффективности корпуса по сравнению со словарем. Н. Вяткина предполагает, что использование корпусных данных может быть более эффективным в сравнении со словарем за счет индуктивного обнаружения и обогащения ввода [Vyatkina, 2016b, p. 161]. Также автор отмечает, что достаточно редко изучаются отсроченные эффекты и практически нет исследований, посвященных влиянию когнитивных усилий на эффективность запоминания; по ее мнению, к тому же необходимы исследования таких эффектов, как управляемая индукция, насыщенность и усиление входного сигнала [Vyatkina, 2016a, p. 220].

Большинство проведенных исследований ориентированы на английский язык как иностранный. На необходимость исследований по другим целевым языкам указывают И. Таргоньска [Targońska, 2014, p. 137], Н. Вяткина и А. Бултон [Vyatkina, Boulton, 2017, p. 6].

Кроме того, для более убедительных выводов необходимы репликации – исследования, выполненные в единой методологии, которые позволят задействовать большое количество участников разных уровней [Vyatkina, 2016a, p. 220].

Исследование, о котором сообщается в данной статье, направлено на восполнение указанных пробелов и является вкладом в дискуссию о преимуществах и эффективности использования словарных и корпусных источников информации о коллокациях для немецкого как иностранного.

3. Методология исследования

В рамках исследования подлежала проверке гипотеза о том, что эффективность запоминания коллокаций зависит от объема когнитивных усилий обучающихся, приложенных на этапе ввода новой информации. Обработка обогащенных и усиленных корпусных данных потребует больших когнитивных усилий и, следовательно, обеспечит более прочное произвольное запоминание, чем обработка данных из словарей.

Для проверки данной гипотезы был проведен долгосрочный лабораторный эксперимент с повторными измерениями [Empirische Forschungsmethoden, 2014, p. 239; Döring, Bortz, 2016, p. 209]. Следуя общей тенденции, сформулированной А. Бултоном и С. Де Кок [Boulton, De Cock, 2016, p. 11], для операционализации эффективности овладения коллокациями была использована степень запоминаемости словосочетаний, проверяемая дважды: в непосредственном и отложенном тестировании. Инструментом измерения запоминаемости стали тесты, включающие задания на заполнение пропусков. Подобная методика проведения исследования с указанными задачами уже широко распространена [Boulton, Cobb, 2017].

Исследовательская стратегия включала следующие этапы:

0. Срез знаний для выяснения информированности студентов о коллокациях в изучаемом языке.

1. Непосредственное тестирование для ответа на вопрос, какой тип ввода информации о коллокациях обеспечивает ее наиболее эффективное понимание.

2. Отложенное тестирование для выяснения, какой тип ввода информации о коллокациях обеспечивает их более прочное произвольное запоминание.

3. Сравнение результатов непосредственного и отложенного тестирования для определения эффективности произвольного запоминания в зависимости от типа ввода коллокаций (корпусные и словарные источники).

В качестве источников информации о коллокациях использовались: «Русско-немецкий переводной словарь» (далее – Русс.-нем. сл.) [Großes Schulwörterbuch Russisch, 2008], примеры из корпуса Digitales Wörterbuch der deutschen Sprache (далее – DWDS) [DWDS – Digitales Wörterbuch], толковый словарь DaF от Langenscheidt (далее – Langenscheidt) [Deutsch als Fremdsprache, 2011] и облако слов из корпуса DWDS (далее – Wortwolke) [DWDS – Wortprofil]. Оба словаря – и переводной, и толковый – рассматриваются нами как инструмент деконтекстуализированного ввода коллокаций. Корпусные источники (примеры и облако слов) относятся к обогащенному и усиленному вводу информации.

Тестовыми инструментами на каждом этапе эксперимента были бумажные рабочие листы с десятью пропусками (на первом и втором этапе по каждому из четырех источников информации). В ходе тестирования испытуемые заполняли пропуски в тексте, подбирая правильные, с их точки зрения, немецкоязычные коллокации на основе предложенных бумажных копий информации из соответствующего источника.

Испытуемыми стали 30 студентов, изучающих немецкий как первый или второй иностранный язык и владеющих им на уровне B1–B2. Объем выборки определялся по рекомендациям Н. Дёринг и Ю. Борца [Döring, Bortz, 2016, p. 843]. Данные испытуемых были анонимизированы (работы обозначались индексами)¹. Участники были проинформированы, что эти результаты будут средством контроля их уровня владения коллокациями, но в начале эксперимента они не знали, что этот уровень будет проверен в отложенном тестировании. Это позволило предотвратить целенаправленное повторение материала в период до отложенного тестирования.

Собранные данные первично были обработаны с использованием статистического редактора Excel (версия 2010). Вторичная обработка выполнялась с использованием пакета IBM SPSS Statistics (версия 23) и программы JASP (версия 0.11.1).

Полученные экспериментальные данные были проанализированы по четырем позициям:

0. Срез знаний: сравнение результатов испытуемых.

1. Непосредственное тестирование: сравнение средних баллов по каждому из четырех использованных источников информации.

2. Отложенное тестирование: сравнение средних баллов по каждому из четырех использованных источников информации.

3. Сравнение разности между непосредственным и отложенным тестированием.

В качестве нулевой гипотезы принималось отсутствие различий между группами. В качестве альтернативной гипотезы принималась зависимость сравниваемых величин от использованного источника информации. Для оценки значимости выявленных различий использовались статистические критерии, выбор которых определялся типом данных, характером выборки и типом распределения результатов.

4. Результаты исследования

Описательные статистические данные для результатов разных этапов исследования представлены в табл. 1.

4.0. Срез знаний

Для среза знаний были отобраны коллокации из словаря А. Хэки Бухофер и др. «Устойчивые словосочетания немецкого языка» (например, *den Zug verpassen, ein Medikament einnehmen, einen Flug buchen, dunkles Brot*) [Feste Wortverbindungen des Deutschen, 2014]. Порог эффективности был определен в 60 %, при правильном выполнении более 60 % заданий владение коллокациями оценивалось как удовлетворительное. Срез знаний показал, что студенты недостаточно хорошо владеют коллокациями (см. таблицу 1): среднее значение – 22,5 %, и даже

¹ Для этого эксперимента было получено устное информированное согласие от всех участников предоставить свои результаты тестирования в качестве материалов исследования.

максимальный результат не превысил 55 %. Следовательно, их можно привлечь к тестированию для проверки заявленной гипотезы.

Таблица 1

Описательная статистика*

Table 1

Descriptive Statistics

Источник информации	N	M	Me	SD	min	max
0. Владение коллокациями (срез знаний, балл)						
Без источника	30	2,25	2,0	1,38	0,5	5,50
Понимание и запоминание коллокаций при использовании разных источников информации						
1. Понимание (непосредственное тестирование, балл)						
Русс.-нем. словарь	30	9,25	10,0	1,26	5,0	10,0
DWDS	30	6,38	6,75	2,09	0	9,0
Langenscheidt	30	7,52	8,0	1,85	2,0	10,0
Wortwolke	30	5,75	6,0	1,05	3,5	8,0
2. Запоминание (отложенное тестирование, балл)						
Русс.-нем. словарь	30	1,37	1,0	1,12	0	4,5
DWDS	30	1,0	1,0	0,71	0	3,0
Langenscheidt	30	1,45	1,0	1,01	0	3,0
Wortwolke	30	0,90	1,0	0,76	0	2,5
3. Разность между пониманием и запоминанием коллокаций при использовании разных источников информации						
Русс.-нем. словарь	30	7,88	8,0	1,48	5,0	10,0
DWDS	30	5,38	5,25	2,05	0	9,0
Langenscheidt	30	6,07	6,50	1,75	1,5	9,0
Wortwolke	30	4,85	5,0	1,08	2,5	8,0

* N – размер выборки, M – среднее, Me – медиана, SD – стандартное отклонение, min – минимальное, max – максимальное значение.

Тестирование

Участники могли получить до 10 баллов в каждом тесте (по каждому из источников информации). Обобщенные результаты непосредственного и отложенного тестирования представлены на рис. 1.²

² Примеры коллокаций, использованных в тестовых заданиях: das Gepäck aufgeben, mit bloßem Auge, ein Tor schießen [Großes Schulwörterbuch Russisch, 2008]; starker Kaffee, in Wut geraten, tatkräftig unterstützen [DWDS – Digitales

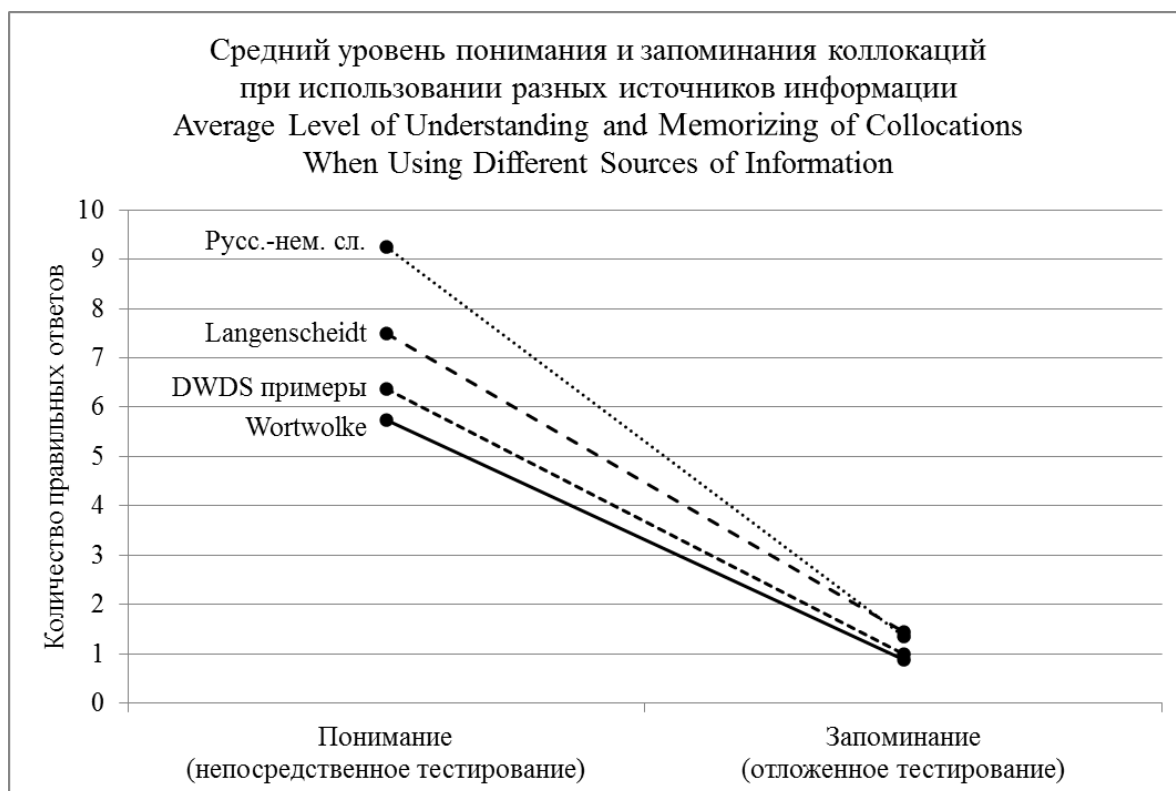


Рис. 1. Результаты непосредственного и отложенного тестирования

Fig. 1. Immediate test and delayed post-test results

График показывает, что:

- 1) количество правильных ответов на непосредственном тестировании различается в зависимости от источника информации;
- 2) количество правильных ответов на отложенном тестировании различается в зависимости от источника информации;
- 3) разность в количестве правильных ответов между двумя тестами (угол падения линий) различается в зависимости от источника информации.

4.1. Непосредственное тестирование

На этапе непосредственного тестирования наибольший средний ранг (самое большое количество правильных ответов) относится к группе Русско-немецкий словарь (97); наименьший средний ранг относится к группе Wortwolke (33). Достоверность различий проверена по критерию Краскела — Уоллиса (см. таблицу 2), который показывает высокую значимость ($p = 0,0000$), что позволяет провести попарное сравнение с использованием U-критерия Манна — Уитни (см. таблицу 3).

Статистически значимая разница между группами на этапе непосредственного тестирования ($p < 0,05$) наблюдается для всех сравниваемых пар. Однако наиболее значимым является отличие показателей любого из трех источников от «Русско-немецкого словаря» ($p = 0,000$), а также отличие Langenscheidt от Wortwolke ($p = 0,000$).

Wörterbuch]; eine vierköpfige Familie, eine Note vergeben, Bett machen, neue Kartoffeln [Deutsch als Fremdsprache, 2011]; verkaufsoffener Sonntag, den Zug erwischen, im schulpflichtigen Alter [DWDS – Wortprofil].

Таблица 2

Результаты критерия Краскела — Уоллиса*

Table 2

Results of the Kruskal-Wallis Test

Источник информации	Rank-sum	Middle rank	H Kruskal–Wallis	<i>p</i>
1. Непосредственное тестирование				
Русс.-нем. сл.	2896	97	56,712	0,0000
DWDS	1428	48		
Langenscheidt	1958	65		
Wortwolke	980	33		
2. Отложенное тестирование				
Русс.-нем. сл.	1981	66	5,223	0,156
DWDS	1649	55		
Langenscheidt	2062	69		
Wortwolke	1569	52		
3. Разность между непосредственным и отложенным тестированием				
Русс.-нем. сл.	2788	93	43,158	0,000
DWDS	1510	50		
Langenscheidt	1863	62		
Wortwolke	1101	37		

* Rank-sum – сумма рангов, Middle rank – средний ранг, H Kruskal–Wallis – критерий Краскела – Уоллиса, p – уровень значимости.

Таким образом, на этапе непосредственного тестирования были выявлены статистически значимые различия по числу правильных ответов в зависимости от использованного источника информации. Эти результаты объясняются тем, что «Русско-немецкий словарь» предоставляет пользователю непосредственный ответ в виде перевода искомого словосочетания, тогда как при использовании других источников информации выбирать коллокации на основе предложенных данных приходится самостоятельно. Значимую разницу между результатами в группах Langenscheidt и Wortwolke, вероятно, можно объяснить особенностями представления информации: в словаре Langenscheidt коллокации приведены как готовые словосочетания (с толкованием значения каждого из них), в облаке же коллокаты представлены, но значение каждого словосочетания никак не комментируется.

Таблица 3

Результаты сравнения групп по критерию Манна – Уитни*

Table 3

Results of the Mann–Whitney Test

Источник информации	Middle rank	U Mann—Whitney	<i>p</i>
1. Непосредственное тестирование			
Русс.-нем. сл.	42,93	77	0,000
DWDS	18,07		
Русс.-нем. сл.	40,38	153,5	0,000
Langenscheidt	20,62		
Русс.-нем. сл.	44,2	39	0,000
Wortwolke	16,8		
DWDS	25,27	293	0,019
Langenscheidt	35,73		
DWDS	35,25	307,5	0,034
Wortwolke	25,75		
Langenscheidt	39,9	168	0,000
Wortwolke	21,1		
3. Разность по группам между непосредственным и отложенным тестированием			
Русс.-нем. сл.	40,97	136	0,000
DWDS	20,03		
Русс.-нем. сл.	39,43	182	0,000
Langenscheidt	21,57		
Русс.-нем. сл.	43,52	59,5	0,000
Wortwolke	17,48		
DWDS	27,35	355,5	0,16
Langenscheidt	33,65		
DWDS	33,93	347	0,124
Wortwolke	27,07		
Langenscheidt	37,87	229	0,001
Wortwolke	23,13		

* Middle rank – средний ранг, U Mann–Whitney – критерий Манна – Уитни, *p* – уровень значимости.

4.2. Отложенное тестирование

Отложенное тестирование было проведено спустя шесть недель, список коллокаций остался прежним. Поскольку установка на запоминание коллокаций испытуемым на первом этапе не давалась, и они не знали о предстоящем отложенном тестировании, на этом этапе проверялось, какой тип ввода информации обеспечивает более прочное произвольное запоминание материала.

Наибольший средний ранг (самое большое количество правильных ответов) относится к группе Langenscheidt (69); наименьший средний ранг относится к группе Wortwolke (52); критерий Краскела – Уоллиса не значим ($p = 0,156 > 0,05$) (см. таблицу 2). Следовательно, на этапе отложенного тестирования и словарная, и корпусная информация сохраняются приблизительно на одном уровне.

4.3. Сравнение разности между непосредственным и отложенным тестированием

Для проверки основной гипотезы исследования (см. раздел «Методология исследования») необходимо сравнить разность между результатами непосредственного и отложенного тестиро-

вания по каждому из источников информации (угол падения линий на рис. 1). Если эта разность минимальна, значит, испытуемый сохранил в своей памяти больше правильных коллокаций.

Медианы разности между пониманием и запоминанием коллокаций отличаются в зависимости от источника информации. Минимальная разность наблюдается в группе Wortwolke – 5, максимальная в группе «Русско-немецкий словарь» – 8 (см. таблицу 1). Для проверки статистической значимости этих различий был использован критерий Краскела — Уоллиса (см. таблицу 2).

Наибольшее ранговое среднее (самая большая разность между правильными ответами в непосредственном и отложенном тестировании) относится к группе «Русско-немецкий словарь» (93), наименьшее – к группе Wortwolke (37). Критерий Краскела — Уоллиса высоко значим ($p = 0,000$). Следовательно, медианы групп значимо отличаются друг от друга, что позволяет провести попарное сравнение групп с помощью критерия Манна – Уитни (см. таблицу 3).

Статистически значимая разница между группами ($p < 0,05$) наблюдается для всех попарных сравнений, кроме пар DWDS vs Langenscheidt ($p = 0,16$) и DWDS vs Wortwolke ($p = 0,124$). Соответственно, в данных комбинациях источник лингвистической информации не влияет существенно на эффективность ее запоминания. Для всех остальных пар принимаем альтернативную гипотезу: различные типы ввода влияют на эффективность запоминания. Запоминание существенно эффективнее при использовании обогащенных и усиленных данных из корпуса (DWDS/Wortwolke) в сравнении с деконтекстуализированным вводом («Русско-немецкий словарь» / Langenscheidt).

Сила эффекта

Для обеспечения методологической строгости исследования использовалась статистика величины эффекта, позволяющая проводить сравнения между исследованиями [Vyatkina, Boulton, 2017, p. 6]. В современной статистике закрепилось мнение, что если сила эффекта невелика, то результат не ценен, даже если есть статистическая значимость [Empirische Forschungsmethoden, 2014, p. 205]. В лингводидактике наиболее распространенной мерой оценки силы эффекта является эффект Козна (Cohen's d) [Boulton, Cobb, 2017]. Результаты расчета эффекта Козна представлены в таблице 4. Сила эффекта была вычислена для парных сравнений на непосредственном и отложенном тестировании, а также для сравнения разности между непосредственным и отложенным тестированием.

Таблица 4

Результаты эффекта Козна

Table 4

Results of Cohen's d

Источник информации	Непосредственное тестирование	Отложенное тестирование	Разность между непосредственным и отложенным тестированием
Русс.-нем. сл. DWDS	1,659	0,391	1,397
Русс.-нем. сл. Langenscheidt	1,094	0,078	1,120
Русс.-нем. сл. Wortwolke	3,013	0,487	2,335
DWDS Langenscheidt	0,574	0,516	0,359
DWDS Wortwolke	0,383	0,136	0,325
Langenscheidt Wortwolke	1,175	0,615	0,836

Для интерпретации полученных показателей обратимся к рекомендациям Л. Плоски и Ф. Освальда [2014] для сферы обучения неродным языкам, в частности, для внутригрупповых дизайнов (слабый эффект – до 0,6, средний – до 1,0, сильный – до 1,4) [Boulton, Cobb, 2017]. Сильный эффект выделен в таблице жирным шрифтом. Кроме того, есть еще три случая очень сильного эффекта ($> 1,4$), в таблице 4 эти ячейки залиты серым.

Сила эффекта, определенная для последнего сравнения, разности между непосредственным и отложенным тестированием, подтверждает статистическую значимость. Деконтекстуализированный ввод существенно уступает в эффективности запоминания обогащенному и усиленному вводу коллокаций.

5. Обсуждение результатов

Полученные результаты сопоставимы с результатами других исследований. Так, в исследовании А. Бултона (2009) по сравнению эффективности словарей и корпусов участвовали студенты французского университета, изучающие английский язык как иностранный [Boulton, 2009]. В рамках эксперимента они на основании корпусных данных (конкордансы в формате KWIC и краткие контексты), а также данных из двуязычных словарей и грамматик выбирали ответ на вопрос о сочетаемости с наречиями. Автор пришел к выводу, что использование корпусных данных более эффективно, чем использование традиционных словарных источников [Boulton, 2009, p. 50]. В нашем исследовании корпусные источники также оказались более эффективными в сравнении со словарными.

Результаты нашего исследования позволяют утверждать, что эффект усиленного и обогащенного ввода гораздо ярче проявляется при восприятии коллокаций (непосредственное тестирование), чем при запоминании (отложенное тестирование). Однако и в отложенном тестировании (т. е. в запоминании) усиленный ввод (облако коллокаций из корпуса) оказался более эффективным, чем деконтекстуализированный (оба словаря), в то время как разница между усиленным и обогащенным вводом (то есть корпусными примерами и облаком слов, соответственно) не существенна.

Таким образом, полученные результаты во многом согласуются с выводами других исследователей, уточняют и подкрепляют их. Представленные нами результаты получены на материале немецкого языка, что расширяет сферу действия выводов. Следует, однако, отметить, что данное исследование было построено на относительно небольшой выборке, поэтому результаты должны интерпретироваться с осторожностью.

Заключение

Представленное исследование позволяет сделать вывод, что когнитивные усилия, прикладываемые обучающимися для запоминания коллокаций в изучаемом иностранном языке, зависят от условий ввода информации. Обогащенный и усиленный ввод способствует более эффективному произвольному запоминанию, чем деконтекстуализированные словарные данные. Корпусные данные по своей природе являются обогащенными и усиленными. Их использование в учебном процессе позволит повысить эффективность усвоения коллокаций.

Список литературы

- Горина О. Г., Царакова Н. С. Корпусные инструменты, маршруты и эксперименты в современной лингводидактике // Вестник НГУ. Лингвистика и межкультурная коммуникация, 2021. № 2. С. 36–53.
- Павельева Т. Ю. Изучение коллокаций на основе лингвистических корпусов текстов // Вестник Тамбовского университета. Серия Гуманитарные науки, 2016. № 3/4 (155/156). С. 56–61.

- Павловская И. Ю., Горина О. Г.** Корпусно-когнитивные методы изучения лексической стороны речи на иностранном языке // Вестник Череповецкого гос. ун-та, 2017. № 1. С. 132–138.
- Хохлова М. В.** Экспериментальная проверка методов выделения коллокаций // *Slavica Helsingiensia* 34. Инструментарий русистики: Корпусные подходы / Под ред. А. Mustajoki. Helsinki: Dep. of Slavonic and Baltic Languages and Literatures, 2008. С. 343–357.
- Boulton A.** Testing the Limits of Data-Driven Learning: Language Proficiency and Training. *ReCALL*, 2009, no. 1, p. 37–54.
- Boulton A., Cobb T.** Corpus Use in Language Learning: A Meta-Analysis // *Language Learning*, 2017. No. 2. Pp. 348–393.
- Boulton A., De Cock S.** Dictionaries as Aids for Language Learning. In: *International Handbook of Modern Lexis and Lexicography* / ed. by P. Hanks, G.-M. de Schryver. New York, Springer, 2016, p. 1–17.
- Deutsch als Fremdsprache.** Duden-Bibliothek. DIGITAL Joanneum Research. Graz. Version 5.1.3.0: Langenscheidt, 2011. CD-Rom.
- Döring N., Bortz J.** Forschungsmethoden und Evaluation. In den Sozial- und Humanwissenschaften. Berlin, Springer, 2016. 1051 S.
- DWDS – Digitales Wörterbuch der deutschen Sprache.** Das Wortauskunftssystem zur deutschen Sprache in Geschichte und Gegenwart. [Электронный ресурс]. URL: <https://www.dwds.de/> (дата обращения: 24.02.2020).
- DWDS – Wortprofil.** Erstellt durch das Digitale Wörterbuch der deutschen Sprache. [Электронный ресурс]. URL: <https://www.dwds.de/wp> (дата обращения: 24.02.2020).
- Empirische Forschungsmethoden für Deutsch als Fremd- und Zweitsprache.** Eine Einführung (UTB Sprachwissenschaft, 8541) / J. Settinieri, S. Demirkaya, A. Feldmeier, N. Gültekin-Karakoç, C. Riemer (Hrsg.). Paderborn : Ferdinand Schöningh, 2014. 1315 S.
- Feste Wortverbindungen des Deutschen.** Kollokationenwörterbuch für den Alltag / A. Häcki Buhofer, M. Dräger, S. Meier, T. Roth, R. Frauchiger (Hrsg.). Tübingen, Francke, 2014. XIX, 1003 S.
- Frankenberg-Garcia A.** Combining User Needs, Lexicographic Data and Digital Writing Environments. TechLing 2017. University of Bologna, 10.11.2017.
- Gablasova D., Brezina V., McEnery T.** Collocations in Corpus-Based Language Learning Research: Identifying, Comparing, and Interpreting the Evidence // *Language Learning*, 2017. No. 67:S1. Pp. 155–179.
- Geyken A., Didakowski J., Siebert A.** Generation of Word Profiles on the Basis of a Large and Balanced German Corpus. In: *Proceedings of the XIII EURALEX International Congress (Barcelona, 15–19 July 2008)* / ed. by E. Bernal, J.A. DeCesaris. Barcelona, Documenta Universitaria; Institut universitari de lingüística aplicada, Universitat Pompeu Fabra, 2008. Pp. 371–383.
- Großes Schulwörterbuch Russisch.** Russisch - deutsch, deutsch - russisch. Berlin : Langenscheidt, 2008. 719 c.
- Hausmann, F. J.** Die Kollokationen im Rahmen der Phraseologie – systematische und historische Darstellung. *Zeitschrift für Anglistik und Amerikanistik*, 2007, Nr 3, S. 217–234.
- Möhrling, J.** Kollokationen im Lernerwörterbuch – Anspruch und Wirklichkeit // *Linguistik online*, 2011. Nr 3. S. 33–53.
- Nation I. S. P.** Learning Vocabulary in Another Language. Cambridge u. a., Cambridge University Press, 2006. 615 p.
- Peters E.** Learning German Formulaic Sequences: The Effect of Two Attention-Drawing Techniques // *The Language Learning Journal*, 2012. No. 1. Pp. 65–79.
- Quasthoff U.** Wörterbuch der Kollokationen im Deutschen. Berlin, De Gruyter, 2011. 551 S.
- Rahimi M., Momeni G.** The Effect of Teaching Collocations on English Language Proficiency // *Procedia—Social and Behavioral Sciences*, 2012. Pp. 37–42.

- Reder A.** Kommen Kollokationen in Mode? Kollokationskonzepte und ihre mögliche Umsetzung in der Didaktik // *Linguistik online*, 2011. Nr 3 (47). S. 131–140.
- Roberts J. C., Frankenberg-Garcia A., Lew R., Rees G., Sharma N.** Visualisation Approaches for Corpus Linguistics: towards Visual Integration of Data- Driven Learning. Paper presented at 3rd Workshop on Visualization for the Digital Humanities. In: *3rd Workshop on Visualization for the Digital Humanities – Vis4DH*, 2018. Pp. 93–104.
- Sonbul S., Schmitt N.** Explicit and Implicit Lexical Knowledge: Acquisition of Collocations under Different Input Conditions // *Language Learning*, 2013. No. 1. Pp. 121–159.
- Steyer K.** Kollokationen in deutschen Wörterbüchern und in der deutschen Wörterbuchforschung. In: *Lexicographica*, 2008. S. 185–207.
- Szudarski P., Carter R.** The role of input flood and input enhancement in EFL learners' acquisition of collocations // *International Journal of Applied Linguistics*, 2016. No. 2. Pp. 245–265.
- Targońska J.** Kollokationen – ein vernachlässigtes Gebiet der DaF-Didaktik? // *Linguistik online*, 2014. Nr 6. S. 127–149.
- Targońska J.** Kollokationskompetenz vs. Sprachfertigkeiten bzw. andere Sprachkompetenzen – ein Forschungsüberblick // *Glottodidactica. An International Journal of Applied Linguistics*, 2019. Nr 1. S. 179–196.
- Vyatkina N.** Data-Driven Learning for Beginners: The Case of German Verb-Preposition Collocations // *ReCALL*, 2016a. No. 2, p. 207–226.
- Vyatkina N.** Data-Driven Learning of Collocations: Learner Performance, Proficiency, and Perceptions // *Language Learning & Technology*, 2016b. No. 20(3). Pp. 159–179.
- Vyatkina N., Boulton A.** Corpora in Language Learning and Teaching // *Language Learning & Technology*, 2017. No. 21(3). Pp. 1–8.
- Zaabalawi R. S., Gould A. M.** English Collocations: A Novel Approach to Teaching the Language's Last Bastion // *Ampersand*, 2017. Pp. 21–29.

References

- Boulton, A.** Testing the Limits of Data-Driven Learning: Language Proficiency and Training. *ReCALL*, 2009, no. 1, p. 37–54.
- Boulton, A., Cobb, T.** Corpus Use in Language Learning: A Meta-Analysis. *Language Learning*, 2017, no. 2, p. 348–393.
- Boulton, A., De Cock, S.** Dictionaries as Aids for Language Learning. In: *International Handbook of Modern Lexis and Lexicography* / Ed. by P. Hanks, G.-M. de Schryver. New York, Springer, 2016, p. 1–17.
- Deutsch als Fremdsprache.** Duden-Bibliothek. DIGITAL Joanneum Research. Graz. Version 5.1.3.0 : Langenscheidt. 2011. CD-Rom.
- Döring, N., Bortz, J.** Forschungsmethoden und Evaluation. In den Sozial- und Humanwissenschaften. Berlin, Springer, 2016. 1051 S.
- DWDS – Digitales Wörterbuch der deutschen Sprache.** Das Wortauskunftssystem zur deutschen Sprache in Geschichte und Gegenwart. [Online]. URL: <https://www.dwds.de/> (zuletzt geprüft am: 24.02.2020).
- DWDS – Wortprofil.** Erstellt durch das Digitale Wörterbuch der deutschen Sprache. [Online]. URL: <https://www.dwds.de/wp> (zuletzt geprüft am: 24.02.2020).
- Empirische Forschungsmethoden für Deutsch als Fremd- und Zweitsprache.** Eine Einführung (UTB Sprachwissenschaft, 8541). J. Settinieri, S. Demirkaya, A. Feldmeier, N. Gültekin-Karakoç, C. Riemer (Hrsg.). Paderborn : Ferdinand Schöningh, 2014. 1315 S.
- Feste Wortverbindungen des Deutschen.** Kollokationenwörterbuch für den Alltag / A. Häcki Buhofer, M. Dräger, S. Meier, T. Roth, R. Frauchiger (Hrsg.). Tübingen, Francke, 2014. XIX, 1003 S.

- Frankenberg-Garcia, A.** Combining User Needs, Lexicographic Data and Digital Writing Environments. TechLing 2017. University of Bologna, 10.11.2017.
- Gablasova, D., Brezina, V., McEnery, T.** Collocations in Corpus-Based Language Learning Research: Identifying, Comparing, and Interpreting the Evidence. *Language Learning*, 2017, no. 67:S1, p. 155–179.
- Geyken, A., Didakowski, J., Siebert, A.** Generation of Word Profiles on the Basis of a Large and Balanced German Corpus. In: *Proceedings of the XIII EURALEX International Congress (Barcelona, 15-19 July 2008)* / ed. by E. Bernal, J.A. DeCesaris. Barcelona : Documenta Universitaria; Institut universitari de lingüística aplicada, Universitat Pompeu Fabra, 2008, p. 371–383.
- Gorina O. G., Tsarakova N. S.** Corpus Routes and Experiments in Language Teaching. *Vestnik NSU. Series: Linguistics and Intercultural Communication*, 2021, vol. 19, no. 2, p. 36–53. (in Russ.)
- Großes Schulwörterbuch Russisch.** Russisch – deutsch, deutsch – russisch. Berlin, Langenscheidt, 2008. 719 c.
- Hausmann, F. J.** Die Kollokationen im Rahmen der Phraseologie – systematische und historische Darstellung. *Zeitschrift für Anglistik und Amerikanistik*, 2007, Nr 3, S. 217–234.
- Khokhlova, M. V.** Evaluation of Methods for Collocation Extraction. *Slavica Helsingiensia 34. Instrumentarij rusistiki: Korpusnye podhody* / ed. by A. Mustajoki. Helsinki : Dep. of Slavonic and Baltic Languages and Literatures, 2008, p. 343–357. (in Russ.)
- Möhrling, J.** Kollokationen im Lernerwörterbuch – Anspruch und Wirklichkeit. *Linguistik online*, 2011, Nr 3, S. 33–53.
- Nation, I. S. P.** Learning vocabulary in another language. Cambridge u. a., Cambridge University Press, 2006. 615 p.
- Paveleva, T. Yu.** Collocations Study based on the Text Corpora. *Tambov University Review. Series: Humanities*, 2016, no. 3/4 (155/156), p. 56–61. (in Russ.)
- Pavlovskaya, I. Yu., Gorina O. G.** Corpus-Based Cognitive Oriented Methods of Lexical Analysis in Foreign Language Studies. *Cherepovets State University Bulletin*, 2017, no. 1, p. 132–138. (in Russ.)
- Peters, E.** Learning German Formulaic Sequences: The Effect of Two Attention-Drawing Techniques. *The Language Learning Journal*, 2012, no. 1, p. 65–79.
- Quasthoff, U.** Wörterbuch der Kollokationen im Deutschen. Berlin : De Gruyter, 2011. 551 S.
- Rahimi, M., Momeni, G.** The Effect of Teaching Collocations on English Language Proficiency. *Procedia – Social and Behavioral Sciences*, 2012, p. 37–42.
- Reder A.** Kommen Kollokationen in Mode? Kollokationskonzepte und ihre mögliche Umsetzung in der Didaktik. *Linguistik online*, 2011, Nr 3 (47), S. 131–140.
- Roberts, J. C., Frankenberg-Garcia, A., Lew, R., Rees, G., Sharma, N.** Visualisation Approaches for Corpus Linguistics: towards Visual Integration of Data- Driven Learning. Paper presented at 3rd Workshop on Visualization for the Digital Humanities. In: *3rd Workshop on Visualization for the Digital Humanities – Vis4DH*, 2018, p. 93–104.
- Sonbul, S., Schmitt, N.** Explicit and Implicit Lexical Knowledge: Acquisition of Collocations under Different Input Conditions. *Language Learning*, 2013, no. 1, p. 121–159.
- Steyer, K.** Kollokationen in deutschen Wörterbüchern und in der deutschen Wörterbuchforschung. In: *Lexicographica*, 2008, S. 185–207.
- Szudarski, P., Carter, R.** The Role of Input Flood and Input Enhancement in EFL Learners' Acquisition of Collocations. *International Journal of Applied Linguistics*, 2016, no. 2, p. 245–265.
- Targońska, J.** Kollokationen – ein vernachlässigtes Gebiet der DaF-Didaktik? *Linguistik online*, 2014, Nr 6, S. 127–149.

- Targońska, J.** Kollokationskompetenz vs. Sprachfertigkeiten bzw. andere Sprachkompetenzen – ein Forschungsüberblick. *Glottodidactica. An International Journal of Applied Linguistics*, 2019, Nr 1, S. 179–196.
- Vyatkina, N.** Data-Driven Learning for Beginners: The Case of German Verb-Preposition Collocations. *ReCALL*, 2016a, no. 2, p. 207–226.
- Vyatkina, N.** Data-Driven Learning of Collocations: Learner Performance, Proficiency, and Perceptions. *Language Learning & Technology*, 2016b, no. 20 (3), p. 159–179.
- Vyatkina, N., Boulton, A.** Corpora in Language Learning and Teaching. *Language Learning & Technology*, 2017, no. 21 (3), p. 1–8.
- Zaabalawi, R. S., Gould, A. M.** English Collocations: A Novel Approach to Teaching the Language's Last Bastion. *Ampersand*, 2017, p. 21–29.

Сведения об авторе

Савина Ольга Юрьевна, кандидат филологических наук, доцент кафедры немецкой филологии Тюменского государственного университета

Information about the Author

Olga Yu. Savina, Cand. of Sci. (Linguistics), Associate Professor of the Department of German Philology at Tyumen University

*Статья поступила в редакцию 01.07.2021-01.06.2022;
одобрена после рецензирования 15.07.2022; принята к публикации 22.07.2022
The article was submitted 01.07.2021-01.06.2022; approved after reviewing 15.07.2022;
accepted for publication 22.07.2022*