

ЯЗЫКОЗНАНИЕ

Научная статья

УДК 811.133.1'373

doi: 10.17223/19996195/58/1

КОРПУСНОЕ ИССЛЕДОВАНИЕ СИНТАГМАТИЧЕСКОЙ СОЧЕТАЕМОСТИ (НА ПРИМЕРЕ КОЛЛОКАЦИЙ С КОМПОНЕНТОМ *PROFESSEUR*)

Юлия Валерьевна Богоявленская^{1,2}, Анна Сергеевна Жесткова³

¹ Уральский федеральный университет им. первого Президента России
Б.Н. Ельцина, Екатеринбург, Россия

^{2,3} Петрозаводский государственный университет, Петрозаводск, Россия
^{1,2} jvbog@yandex.ru

³ zhest1806anna@gmail.com

Аннотация. Анализируется синтагматический (коллокационный) потенциал французского слова *professeur*. Под коллокацией понимается комбинация двух или более лексических единиц, имеющих тенденцию к совместной встречаемости. В основе этой комбинации лежит семантико-грамматическая взаимообусловленность ее составных элементов. Коллокации характеризуются устойчивостью, рекуррентностью и воспроизводимостью. Работа выполнена в русле корпусной лингвистики, ее методологическая база совмещает применение общенаучных методов и приемов анализа (наблюдение, описание, классификация, сравнительно-сопоставительный анализ), метода выявления коллокатов с использованием грамматики лексико-синтаксических шаблонов, элементов методики унификации словарных дефиниций и метода комплексного комбинаторно-семантического анализа. Актуальность исследования обусловлена, прежде всего, значимостью коллокаций: они свойственны всем языкам и имеют яркую национально-языковую специфику.

Разработана методика семантико-синтагматического анализа сочетаемости лексических единиц, согласно которой на первом этапе были проанализированы значения слова *professeur* в нескольких словарях, сформулировано унифицированное лексическое значение, а также выявлены коллокаты слова, иллюстрирующие словарные статьи. На втором этапе состоялось обращение к корпусу frTenTen2017, размещенному на сетевой платформе Sketch Engine. Методом автоматического извлечения получен 713 641 случай употребления слова, распределенных по скетчам – лексико-синтаксическим шаблонам, наполнение которых можно ранжировать по показателям частоты взаимной встречаемости коллокации и ее типичности (уникальности). На третьем этапе сопоставлены данные словарей и выборки корпуса. Определено, что словарные коллокаты лишь частично подтверждаются в скетчах, построенных на материале многомиллиардного корпуса. При этом установлен ряд устойчивых коллокаций, не нашедших места в словарях. Все это свидетельствует о недостаточном отражении словарями реального функционирования слова в речи. Выявлены логико-понятийные, семантические, сочетаемостные, стилистические ограничения коллокаций с компонентом *professeur*. Тестирование сервиса Sketch Engine показало ряд его

недостатков, которые частично можно скорректировать и которые необходимо учитывать в работе со скетчами. Практическая ценность исследования заключается также в возможности применения полученных результатов в лексикографической практике и преподавании ряда теоретических и практических лингвистических дисциплин.

Ключевые слова: коллокация, коллокат, синтагматическая сочетаемость, Sketch Engine, корпусная лингвистика, методика семантико-синтагматического анализа

Для цитирования: Богоявленская Ю.В., Жесткова А.С. Корпусное исследование синтагматической сочетаемости (на примере коллокаций с компонентом *professeur*) // Язык и культура. 2022. № 58. С. 8–28. doi: 10.17223/19996195/58/1

Original article

doi: 10.17223/19996195/58/1

CORPUS STUDY OF SYNTAGMATIC CO-OCCURRENCE (BASED ON THE EXAMPLE OF COLLOCATIONS FORMED BY THE COMPONENT *PROFESSEUR*)

Julia V. Bogoyavlenskaya^{1, 2}, Anna S. Zhestkova³

¹ Ural Federal University named after the first President of Russia B.N. Yeltsin,
Yekaterinburg, Russia

^{2, 3} Petrozavodsk State University, Petrozavodsk, Russia

^{1, 2} jvbog@yandex.ru

³ zhest1806anna@gmail.com

Abstract. The syntagmatic (collocation) potential of the French word *professeur* is analysed. Collocation refers to a combination of two or more lexical units that tend to occur together. This combination is based on the semantic and grammatical interdependence of its constituent elements. Collocations are characterized by stability, recurrence and reproducibility. This paper is realized in course of the corpus linguistics, its methodological base combines the usage of general scientific research methods and techniques (observation, description, classification, comparative analysis), of the method of identifying the collocates using the grammar of lexical and syntactic patterns, the parts of the method of unification of dictionary definitions and the method of complex combinatorial and semantic analysis. The relevance of the study is mainly due to the significance of collocations: they are characteristics of all languages and they have a bright national linguistic specificity. In the paper a method of semantic and syntagmatic analysis of co-occurrence of lexical units is designed. According to it, firstly, the word definitions in several dictionaries were analysed, a unified lexical meaning was formulated, and besides, the collocates of the word were identified from the dictionary entries. Secondly, the usage of the corpus frTenTen17 from the web service Sketch Engine took place. By the method of automatic extraction 713 641 cases of usage of the word were found. They were classified by sketches – lexical and syntactic patterns, that can be ranked according to the indicators of the frequency of mutual occurrence of collocation and of the score (uniqueness). Thirdly, the data from dictionaries and from corpus were compared. It was found that dictionary collocations are only partially confirmed in sketches from the multibillion corpus. At the same time, a set of stable collocations that had not been found in the dictionaries was determined. All this indicates that dictionaries do not sufficiently reflect the real usage of

the word in speech. Logical, semantic, co-occurring, stylistic limitations of collocations of the word *professeur* were revealed. Testing of the service Sketch Engine has displayed a number of shortcomings, which can partly be corrected and which need to be taken into consideration while working with sketches. The practical value of the research also consists in the possibility of applying the results in lexicographic practice and teaching a number of theoretical and practical linguistic disciplines.

Keywords: collocation, collocates, syntagmatic co-occurrence, Sketch Engine, corpus linguistics, methods of semantic and syntagmatic analysis

For citation: Bogoyavlenskaya Yu.V., Zhestkova A.S. Corpus study of syntagmatic co-occurrence (based on the example of collocations formed by the component *professeur*). *Language and Culture*. 2022;58: 8-28. doi: 10.17223/19996195/58/1

Введение

Коллокация – один из наиболее часто используемых терминов корпусной лингвистики. В широком смысле под коллокациями понимается последовательность двух и более лексических единиц, имеющих тенденцию к совместной встречаемости. Специалистами отмечается, что в основе этого явления лежит семантико-грамматическая обусловленность элементов словосочетания [1]. В корпусной лингвистике это понятие переосмысливается и уточняется. Поскольку «во главу угла ставится частота совместной встречаемости, то коллокации определяются как “статистически устойчивые словосочетания”, как фразеологизированные, так и свободные» [2. С. 138].

Вхождение в научный обиход термина «коллокация» состоялось благодаря трудам представителей Лондонской лингвистической школы в 1930–1950-х гг. Г. Палмер предложил называть коллокациями последовательности слов, которые лучше изучать как целые слова, а Дж.Р. Фёрс издал книгу, в которой ввел термин «значение через “коллокацию”» для обозначения регулярного сочетания слов с другими словами. Лингвист предложил классификацию коллокаций, выделив привычные, общие и узуальные, а также технические, уникальные, персональные и нормальные типы коллокаций. Его работы дали толчок к более активному изучению коллокаций его учениками и последователями, среди которых выделялись М.А.К. Халлидей, Р.Г. Робинс, Дж. Синклер. Одни ученые, например Т.Ф. Митчелл, признавали необходимость как лексического, так и грамматического анализа коллокаций. Он полагал, что коллокации следует изучать в рамках грамматических матриц, выстраивающихся на основе коллокационных связей. Другие исследователи не поддерживали этот подход, исключая участие грамматики. П. Мэттью считал, что под коллокацией следует понимать «отношение» в границах одной синтаксической единицы между отдельными лексическими элементами, в которых слова привычно употребляются вместе (более подробно см. работу [3. С. 169–178]).

Интерес к синтагматической сочетаемости слов в России оформился в работах еще в XIX в. Н.В. Крушевский писал, что некоторые слова «возбуждают в нашем уме» другие связанные с ними слова, что основано на «психическом законе ассоциации по смежности» (цит. по: [3. С. 179]). Сочетаемость слов освещалась в трудах Л.В. Щербы, В.В. Виноградова, А.И. Смирницкого и многих других исследователей. Впервые термин «коллокация» появился в «Словаре лингвистических терминов» О.С. Ахмановой, где определяется как «лексико-фразеологически обусловленная сочетаемость слов в речи как реализация их полисемии» [4. С. 199].

Первым серьезным трудом, посвященным изучению понятия «коллокация», является монография Е.Г. Борисовой «Коллокации: что это такое и как их изучать», где автор определил коллокации как неидеоматические словосочетания, в которых свободный компонент не меняет своего значения, при этом значение несвободного отличается от того, что реализуется в свободных словосочетаниях [5]. В монографии рассматриваются синтаксические, семантические, морфологические и стилистические свойства коллокатов, а также утверждается необходимость изучения коллокаций как самостоятельного объекта.

В настоящее время коллокации привлекают активный интерес исследователей в связи с изучением возможностей и методов их автоматического выявления в корпусах, их семантико-синтаксических, аксиологических и иных функций, роли в обучении иностранным языкам, рассматриваются различные виды коллокаций: колоративные, политические и т.д. Одним из наиболее спорных вопросов является соотношение понятий «коллокация» и «фразеологизм». Коллокация может рассматриваться как родовое понятие по отношению к идиомам [6], как одна из разновидностей устойчивых сочетаний, как разновидность фразеологизмов (см., например, [1, 5, 7]), как единицы, имеющие признаки синтаксически и семантически целостной единицы, выходящие за пределы фразеологического фонда языка (см., например, работу [8]), как единицы «серой зоны» между лексикой и фразеологией [9].

В нашей работе мы опираемся на понимание коллокации как комбинации двух или более лексических единиц, имеющих тенденцию к совместной встречаемости. В основе этой комбинации лежит семантико-грамматическая взаимообусловленность элементов словосочетания. Уточним, что термин «словосочетание» мы считаем родовым по отношению к коллокации. Словосочетания могут быть как свободными, так и не свободными, коллокация же обладает определенной устойчивостью, рекуррентностью и воспроизводимостью, т.е. употребляется в готовом виде, «по традиции».

Целью предпринимаемого исследования является анализ синтагматической сочетаемости французского слова *professeur*. Поставленная

цель предполагает решение следующих задач: 1) разработка методики семантико-синтагматического анализа, применяемого для выявления особенностей синтагматической сочетаемости конкретной лексики; 2) выявление унифицированного лексикографического значения слова *professeur* и репертуара его коллокатов в современных словарях французского языка; 3) инвентаризация и анализ коллокатов слова *professeur* по лексико-синтаксическим моделям на материале корпуса frTenTen, размещенного на платформе Sketch Engine; 4) сравнительно-сопоставительный анализ полученных результатов. Дополнительная задача состоит в тестировании сервиса с точки зрения эффективности изучения коллокатов той или иной лексической единицы.

Актуальность исследования коллокаций обусловлена, прежде всего, их значимостью: они свойственны всем языкам и имеют яркую национально-языковую специфику. Для выявления синтагматической сочетаемости слова необходима инвентаризация коллокатов слова на материале объемного репрезентативного корпуса. Новизну исследования мы видим в выявлении комбинаторного потенциала социально-значимого слова *professeur*, прежде не попадавшем в фокус внимания исследователей. Практическая ценность усматривается нами в разработке методики детального анализа синтагматической сочетаемости лексических единиц, возможности применения полученных результатов в лексикографической практике и преподавании ряда теоретических и практических лингвистических дисциплин.

Методология исследования

Как отмечают исследователи коллокаций, достоверные данные о синтагматической сочетаемости слов и их частотности можно получить только на материале миллиардных корпусов [10]. На данный момент уже существует немало сервисов, обеспечивающих выявление коллокаций: сервис поиска в корпусах, созданных в Университете Лидса [11], корпусах Aranea Университета им. А. Коменского [12], корпус книг Google Books Ngram Viewer [13], сервис поиска устойчивых сочетаний в «АОТ» (Автоматическая обработка текстов) [14] и др. Одной из наиболее качественных и востребованных платформ для автоматического извлечения и изучения коллокатов является сетевая платформа Sketch Engine [15], созданная в 2003 г. командой разработчиков во главе с А. Килгарифом для анализа текстовых данных и управления корпусами текстов. Одним из ее преимуществ считается успешное сочетание статистических критериев и методов, основывающихся на собственно лингвистических моделях [2. С. 141].

Охарактеризуем данный сервис и особенности работы с его лингвистическим инструментарием, позволяющим извлечь необходимые

для исследователя данные. Sketch Engine располагает более 500 корпусами на более чем 100 языках. Для работы на платформе необходима регистрация, после которой пользователь выбирает интересующий язык или корпус. Функция поиска коллокаций на данной платформе позволяет выявить сочетаемостный потенциал слова, обусловленный особенностями, накладываемыми синтаксисом того или иного языка, семантикой и языковым узусом. Данный поиск реализуется во вкладке Word Sketch, где пользователь может получить так называемые скетчи изучаемого слова.

Под скетчами понимается «описание речевого поведения слова, полученное автоматическим путем обобщения информации о всех контекстах, в которых исследуемое слово встретилось в корпусе» [16. С. 110]. Результаты разбиваются на столбцы – списки коллокатов, построенных по определенным лексико-синтаксическим шаблонам, обработанным анализатором сервиса. Эти шаблоны (модели) отражают частотные грамматические взаимодействия изучаемого слова с другими словами. При работе с коллокатами есть возможность просмотреть такие показатели, как frequency (показатель частоты совместной встречаемости) и score (показатель крепости связи слов: чем он выше, тем чаще эти коллокаты встречаются именно друг с другом; низкий балл означает, что слова часто сочетаются и с другими словами). Пользователь может сортировать получившиеся результаты по данным критериям. Нажатием на число рядом со словом в списке можно просмотреть контексты из корпуса со ссылками на источники. Серьезным преимуществом сервиса является тот факт, что пользователь работает с коллокациями, построенными по определенным моделям, в выборки не попадают бессмысленные комбинации, например существительное с пунктуационным знаком.

Следующий вопрос касается математического аппарата для установления силы синтагматической связи между элементами словосочетаний. Этой цели служат меры ассоциации (меры ассоциативной связанности, уникальности коллокации), учитывающие как линейную близость и частоту совместной встречаемости, так и частоту в данном корпусе и некоторые другие факторы. В Sketch Engine до 2006 г. использовалась мера MI, затем этот метод был заменен на AssociationScore (Ascore). В настоящее время в системе применяется мера logDice, основанная на коэффициенте Дайса и особой гибридной мере. LogDice выражает степень типичности коллокации. Она рассчитывается с учетом как частотности ключевого слова, так и его коллоката. Для нас принципиальное значение имеет преимущество, которое обеспечивает данная мера, а именно: размер корпуса не оказывает влияния на ее значение, что позволяет получать объективные результаты на материале многомиллиардного корпуса из семейства TenTen, которым мы планируем воспользоваться.

ся для исследования. Целевой размер корпусов составлял 10 млрд слов на каждый язык, что и дало название семейству. По факту объем французского корпуса составил 6 845 630 573 токена.

Эмпирическим материалом исследования послужили скетчи коллокатов французского слова *professeur*. В выборке представлены как формы мужского, так и женского рода, поскольку последняя может встречаться как в том же виде, так и приобретать показатель женского рода (буква *e* в конце слова – *professeure*). Обе формы лишь недавно получили академическое одобрение [17]. Выбор этого материала обусловлен констатацией того факта, что данное слово, имеющее несколько соответствий в русском языке (учитель, преподаватель, профессор), получает недостаточное описание в словарях с точки зрения его сочетаемостных свойств и частоты употребления. Наш исследовательский интерес обусловлен двумя факторами: 1) частотностью самого слова; 2) социальной значимостью профессии, востребованной и уважаемой во всех национальных культурах.

Логика исследования выстраивается в соответствии с поставленными выше задачами: от обобщения значений и представленности сочетаемостного потенциала слова *professeur* в ведущих словарях французского языка (Le Petit Robert [18], Dictionnaire Larousse [19], Dictionnaire de l'Académie française en ligne, 9^{ème} édition [20], Dictionnaire Ortolang [21]) к анализу коллокатов, извлеченных в корпусе frTenTen2017 системы Sketch Engine и сопоставлению полученных данных.

В основе нашей методологии лежит комбинация общенаучных методов и приемов анализа и методов корпусной лингвистики. К первым относятся наблюдение, описание, классификация, сравнительно-сопоставительный анализ. Ко вторым – метод выявления коллокатов с использованием грамматики лексико-синтаксических шаблонов при помощи анализатора Sketch Engine. В разработке методики анализа, применяемой в данной работе, мы также использовали элементы методики унификации словарных дефиниций [22] и элементы метода комплексного комбинаторно-семантического анализа, описанного в работе [23].

Исследование и результаты

Согласно разработанной нами методике семантико-синтагматического анализа, на первом этапе обратимся к анализу толкований, а затем к анализу коллокатов изучаемого слова, представленных во французских словарях.

Словарь Le Petit Robert ограничивается кратким толкованием и двумя синонимами: *Personne (diplômée) qui enseigne une discipline. → enseignant. Professeur des écoles → instituteur* [18. P. 2880].

Словарь Larousse приводит два значения и указывает на существование аббревиатуры, используемой в разговорной речи, и формы

женского рода существительного: «Professeur, e. 1. Personne qui enseigne une matière, une discipline: *Professeur d'équitation, de guitare, de français. C'est un excellent professeur* (syn. enseignant); 2. Membre de l'enseignement (abrév.fam. prof): *Un professeur des écoles, d'université*. REM. Au féminin, on rencontre aussi *une professeur* [19. P. 2048]. Уточним, что использование формы женского рода данного слова приобрело официальный характер с 2019 г. в рамках феминизации названий профессий. В докладе Французской академии указывается, что женский род может быть маркирован артиклем *une professeur* или артиклем и окончанием женского рода *une professeure* [17].

Словарь Французской академии дает следующее определение: Celui, celle qui transmet un savoir, qui dispense la connaissance d'une discipline, d'une science, d'un art. *Un professeur et ses élèves. Professeur de grec, de mathématiques, de danse, d'art dramatique. Elle est professeur de piano*. Ценным для нашей работы является то, что словарь выделяет также специальное значение слова, сопровождая его примерами употребления с коллокатами, и его переносное значение: Spécialement. Titulaire d'un diplôme d'enseignement. Fig. Personne qui prône une idée, un point de vue, une doctrine, etc., qui cherche à l'imposer, à l'inculquer aux autres. *Ce penseur est un professeur de scepticisme, de vertu* [20].

Наиболее полное толкование представлено в словаре Ortolang. Поскольку словарная статья довольно объемна, мы не приводим ее здесь полностью. В ней уточняется, что в национальной системе образования преподаватель может работать в средней и высшей школе, иметь научное звание или должность по конкурсу: Enseignant(e) du second degré ou du supérieur, titulaire d'un grade universitaire ou d'un titre acquis par concours lui conférant ce titre. Возможность употребления в женском роде приписывается только разговорному сокращению *prof*. Словарь существенно расширяет представление о семантике слова и его стилистических особенностях: указывается пейоративное значение (Personne qui cherche à propager, à imposer certains principes, certaines idées ou doctrines. Synon. *maître. Professeur d'athéisme, de vertu, de sagesse*), отмечается возможность иронической коннотации, употребления с неодушевленными существительными (*[Le talent de bien dire] est le professeur de l'espèce humaine*). Упоминается, что слово может употребляться в арго в следующем значении: *Vieille courtisane qui excelle à faire l'éducation des petits jeunes gens*. Ценными данными, представленными в словаре, являются многочисленные коллокаты, иллюстрирующие комбинаторный потенциал слова. Словарная статья снабжена примерами его употребления во французской литературе, датированной XIX–XX вв. [21].

Анализ словарных дефиниций показывает, что количество значений в словарях не совпадает, как и не совпадают пометы. Так, неодно-

значно расценивается словарями значение «человек, пропагандирующий некоторые идеи, принципы, доктрины, точку зрения, навязывающий, внушающий их другим»: словарь Ortolang ставит помету «пейоративное», а словарь Французской академии трактует его как переносное. Обобщая данные лексикографических источников, сформулируем унифицированное значение слова:

1. *Personne (diplômée) qui enseigne une discipline, une science, une technique, un art. Être professeur; bon, éminent professeur; être nommé, devenir professeur; un poste de professeur.*

2. [Dans l'Éducation Nationale] Enseignant(e) du second degré ou du supérieur, titulaire d'un grade universitaire ou d'un titre acquis par concours lui conférant ce titre. *Professeur adjoint, associé; conseil des professeurs; professeur de l'enseignement secondaire, supérieur.*

3. *Personne qui transmet un savoir-faire, des connaissances dans un domaine quelconque. C'est grand-père qui a été mon professeur* (GYP. *Souv. pte fille*. 1927. P. 59).

4. Ce qui apprend quelque chose à quelqu'un. [*Le talent de bien dire*] *est le professeur de l'espèce humaine et le secrétaire de l'esprit humain* (TAINE. *Nouv. Essais crit. et hist.* 1865. P. 110).

5. *fig., péj.* *Personne qui cherche à propager, à imposer aux autres certains principes, certaines idées ou doctrines, un point de vue, etc. Professeur d'athéisme, de vertu, de sagesse.*

6. *arg.* *Vieille courtisane qui excelle à faire l'éducation des petits jeunes gens.*

Переходим к анализу коллокатов, представленных в вышеперечисленных словарях. Эти данные мы распределили по группам в соответствии с теми лексико-синтаксическими шаблонами, которые предлагает Sketch Engine, с некоторой адаптацией к традиционным обозначениям. В частности, модель «verbs with X as object» соответствует в нашей работе модели V + O (professeur), «verbs with X as subject» – S (professeur) + V, «modifiers of X» – ADJ + N (professeur), «X de» – N (professeur) + DE + N, «prepositions preceding noun/nouns after preposition» (de, à, hors, etc.) – N + DE + N (professeur), N (professeur) + + À + N, N(professeur) + HORS + N. Такой подход позволит нам в дальнейшем оптимизировать анализ данных, полученных в этом сервисе, и их сопоставление. В словарях были выявлены коллокаты, построенные только по пяти шаблонам (моделям). Они получают следующее наполнение:

ADJ + N (professeur): *bon, mauvais, jeune, vieux, éminent, certifié* (Ortolang)

certifié, agrégé, émérite, associé (Dictionnaire de l'Académie française en ligne)

excellent (Larousse 2021)

V + O (professeur): être professeur, être nommé professeur, devenir professeur (Ortolang)

N + DE + N (professeur): chaire de professeur, poste de professeur (Ortolang)

chaire de professeur, salle des professeurs, conseil des professeurs, robe d'un professeur, épitoge d'un professeur (Dictionnaire de l'Académie française en ligne)

N (professeur) + DE + N: professeur de français, de mathématiques, de chimie, d'histoire, de géographie, de philosophie, de langues, de médecine, de théologie, de droit, de lettres, de gymnastique, de danse, de musique, professeur de judo, de ski, de tennis; professeur de collège, de lycée, d'université, de faculté; professeur de sixième, de terminale; d'athéisme, de vertu, de sagesse (Ortolang)

professeur de grec, de mathématiques, de danse, d'art dramatique, de piano, professeur de littérature comparée; professeur d'université, des universités, des écoles; professeur de scepticisme, de vertu (Dictionnaire de l'Académie française en ligne)

professeur d'équitation, de guitare, de français, des écoles, d'université (Larousse)

professeur des écoles (Le Petit Robert)

N (professeur) + À + N: professeur au petit (grand) Séminaire, à la Sorbonne, au Collège de France, au conservatoire (Dictionnaire Ortolang)

N (professeur) + HORS + N: professeur hors classe (Dictionnaire de l'Académie française en ligne).

Анализ словарных коллокатов позволяет сделать следующие выводы:

1. Наиболее репрезентативную картину сочетаемостного потенциала слова дает Dictionnaire Ortolang (42 коллоката), далее по убыванию: Dictionnaire de l'Académie française en ligne (20), Dictionnaire Larousse (8), Le Petit Robert (1 коллокат).

2. Не все лексико-синтаксические модели представлены в словарях: их количество варьируется от 1 до 5.

3. Коллокаты, приведенные в словарях, практически не совпадают, логики в отборе коллокатов для иллюстрации словарных статей не выявляется.

4. Частотность коллокатов не отражена в словарях, порядок их презентации в словарных статьях не ясен.

5. По данным проанализированных словарей, наибольшее количество коллокатов строится по модели **N (professeur) + DE + N**, далее по убыванию **ADJ + N (professeur)**, **N + DE + N (professeur)**, **V + O (professeur)**, **N (professeur) + À + N**.

6. Модель **N (professeur) + HORS + N** получает только один случай наполнения: *professeur hors classe* в Dictionnaire de l'Académie française en ligne. Несколько коллокатов по модели **N (professeur) + À +**

N и V + O (**professeur**) были приведены только в словарной статье Dictionnaire Ortolang.

Переходим ко второму этапу – анализу данных, автоматически извлеченных из корпуса frTenTen. Корпус включает тексты, автоматически извлеченные с сайтов lefigaro.fr, fr.wikipedia.org, convention-entreprise.fr, forum.doctissimo.fr, docplayer.fr, cartesfrance.fr и других (более подробно см. [15]). В данном корпусе имеется 713 641 случай употребления слова *professeur*, его род в скетчах не маркирован.

В сервисе Sketch Engine предлагаются скетчи по 10 моделям (рис. 1, а, б).

WORD SKETCH French Web 2017 (frTenTen7)

professeur as noun 713,641

verbs with "professeur" as object	verbs with "professeur" as subject	modifiers of "professeur"	pronominal possessors of "professeur"	prepositional phrases
agréer 7,224 10.57 ... professeur agréé	assister 753 8.54 ... professeur assistant à	émérité 11,064 11.84 ... professeur émérite	leur 10,152 4.89 ... leurs professeurs	"professeur" de 229,152 32.11% ...
inviter 4,642 8.07 ... professeur invité à	enseigner 566 0.12 ... professeurs enseignant	titulaire 3,061 9.80 ... professeur titulaire de	notre 4,120 4.19 ... nos professeurs	"professeur" à 73,160 10.25% ...
nommer 7,875 8.4 ... il est nommé professeur	suppléer 191 7.42 ... professeur suppléant de	honoraire 2,462 9.67 ... professeur honoraire	mon 4,176 4.14 ... mon professeur	"professeur" on 10,676 1.5% ...
certifier 7,317 8.36 ... professeurs certifiés	garantir 84 0.37 ... Professeurs Garanties	stimuler 3,350 8.70 ... Les professeurs stimuler	son 14,791 3.7 ... son professeur	"professeur" dans 3,782 0.53% ...
associer 5,661 8.34 ... professeur associé à	dispenser 80 6.16 ... professeurs dispensant	émient 1,025 8.24 ... émient professeur	tes 461 3.09 ... tes professeurs	"professeur" pour 2,223 0.31% ...
diplômer 1,299 7.92 ... un professeur diplômé	encadrer 84 6.03 ... professeurs encadrant	référer 676 7.83 ... professeur référer	voire 2,750 2.95 ... voire professeur	"professeur" sur 940 0.13% ...
devenir 13,128 7.86 ... découvert professeur	corriger 71 5.82 ... professeur corrige	chercheur 500 7.36 ... professeurs et chercheurs	ton 418 0.06% ... ton professeur	"professeur" par 565 0.08% ...
expérimenter 1,101 7.32 ... par un professeur expérimenté	distribuer 79 5.76 ... le professeur distribue	assistant 526 7.29 ... professeur assistant	vous 118 0.02% ... vous professeur	"professeur" sans 312 0.04% ...
retailer 462 6.71 ... professeur retailé	demande 518 6.75 ... professeur demande	ordinaire 830 7.14 ... professeur ordinaire à		"professeur" pendant 252 0.04% ...
affilier 445 6.47 ... professeur affilié à	visiter 82 5.68 ... professeur visitant à	ancien 5,276 6.74 ... ancien professeur de		"professeur" depuis 123 0.02% ...
chevronner 301 6.05 ... avec un professeur chevronné	expliquer 296 5.66 ... professeur explique	extraordinaire 558 6.47 ... professeur extraordinaire		"professeur" contre 118 0.02% ...
être 49,648 5.96 ... est professeur	veiller 79 5.62 ... Le professeur veille	formateur 241 6.41 ... professeurs formateurs		

а

"professeur" and/or ...

"professeur" and/or ...	prepositions preceding nouns after preposition	noun modifiers of "professeur"	nouns modified by noun "professeur"	usage patterns
élève 7,813 0.28 ... élève et professeur	par 27,206 4.8 ... par le professeur	documentaliste 1,231 12.84 ... le professeur documentaliste	mère 231 8.03 ... et d'une mère professeur	with_definite_article 156,226 21.80 ...
étudiant 3,900 0.36 ... étudiants et professeurs	parmi 774 4.72 ... parmi les professeurs	stagiaire 412 10.95 ... les professeurs stagiaires	colleague 105 0.75 ... colleagues professeurs	with_indefinite_article 42,236 5.92 ...
université 3,142 7.71 ... Professeur Université	avec 13,652 4.53 ... avec le professeur	tuteur 56 8.65 ... un professeur tuteur	suite 147 8.47 ... devant par la suite professeur	
instituteur 1,331 7.4 ... instituteurs et professeurs	chez 1,442 4.49 ... chez le professeur	accompagnateur 63 8.44 ... les professeurs accompagnateurs	et d'un père professeur 178 8.35 ...	
parent 2,145 7.34 ... parents et professeurs	selon 1,813 4.22 ... Selon le professeur	universitaire 40 0.20 ... professeur universitaire	apprenti 50 7.6 ... étudiants apprentis professeurs	
chercheur 1,509 7.19 ... professeur et chercheur	entre 1,029 3.04 ... entre les professeurs	langage 19 7.64 ... professeur langage	temps 141 8.07 ... en même temps professeur	
écrivain 1,363 6.85 ... écrivain et professeur	entre 1,943 3.7 ... entre professeurs	histoire 35 7.83 ... professeur histoire	femme 196 6.87 ... la première femme professeur	
compositeur 1,069 6.77 ... compositeur et professeur	pour 7,974 3.25 ... pour les professeurs	encadrent 28 7.61 ... les professeurs encadrants	tour 62 6.79 ... à son tour professeur de	
médecin 1,375 6.72 ... médecin et professeur	de 134,083 3.19 ... des professeurs	relais 67 7.46 ... les professeurs relais	fiche 65 6.74 ... la fiche professeur	
philosophe 917 6.64 ... philosophes professeurs	devant 241 2.64 ... devant les professeurs	remplaçant 21 7.43 ... les professeurs remplaçants	époque 35 6.27 ... à l'époque professeur	
sociologue 946 6.64 ... sociologue professeur	à 19,080 2.38 ... aux professeurs	vocaliste 22 7.4 ... professeurs vocalistes	poste 25 6.05 ... le poste professeur	
directeur 1,473 6.61 ... professeur et directeur	contre 77 2.29 ... en outre professeur	agréé 19 7.23 ... les professeurs agréés	élève 25 6.04 ... les élèves professeurs	

б

Рис. 1. Скетчи коллокаций с компонентом *professeur*

В рамках статьи невозможно описать все лексико-семантические модели, мы обратимся к тем шести, которые были представлены в словарях, а также проанализируем скетч с коллокациями, построенными по модели S (**professeur**) + V.

Поскольку сервис предлагает большие списки коллокатов, мы приняли решение ограничиться значением score 5,0 для иллюстрации

хода анализа в статье. Мы также будем обращаться к показателю частотности и коллокациям с меньшим значением score для сравнения и проверки релевантности выделенных групп. Прежде всего рассмотрим те модели коллокатов, которые представлены в словарях.

Модель **ADJ + N (professeur)** включает модификаторы адъективного типа: *émérite, titulaire, honoraire, similaire, éminent, référent, chercheur, assistant, ordinaire, ancien, extraordinaire, formateur, principal, universitaire, étudiant, spécialiste, particulier, futur, responsable, excellent, contractuel, savant, débutant, québécois, absent, célèbre, compétent, illustre, auxiliaire, coordonnateur, brillant, jeune, permanent, pascal, volontaire, intervenant, anglophone, auguste, vieux*.

Анализ позволяет выделить группы право- и левосторонних адъективных модификаторов существительного *professeur*: прилагательные, уточняющие его профессиональный статус, должность, квалификацию (*émérite, titulaire, honoraire, référent, chercheur, assistant, débutant*), социальный статус, опыт, признание в обществе или профессиональных кругах (*éminent, excellent, célèbre, illustre, compétent, illustre*), а также фокусируют внимание на чертах характера, возрасте (*responsable, jeune, vieux*). Эти модификаторы имеют ярко выраженную позитивную окраску: *éminent, responsable, excellent, célèbre, compétent, illustre*, реже нейтральную окраску: *chercheur, référent, futur*. В тех случаях, когда коннотация модификатора зависит от контекста (например, *jeune, vieux*), мы обращаемся к анализу контекстов. Анализ коннотаций модификаторов *jeune* и *vieux*, в частности, показывает, что эти прилагательные употребляются либо с нейтральной окраской, сообщая о возрасте преподавателя (*Jules Isaac est un jeune professeur d'histoire; elle rencontre à la gare de Bruxelles un vieux professeur d'art italien*), либо с позитивной окраской (*le jeune professeur: une lumière brille dans ses yeux quand il explique; ce bon vieux professeur Sergent Pepper eut la merveilleuse idée de faire un cours spécial cinéma*). Небольшая часть прилагательных характеризует национальность или язык, на котором говорит преподаватель (*québécois, anglophone*). Сказанное в полной мере относится и к прилагательным, оказавшимся ниже значения 5,0.

Анализ скетча показал, что в него ошибочно были отфильтрованы три имени собственных: *Pascal* (122/5,34), *Auguste* (89/5,02), *Julien* (44/3,87). В сервисе эти имена фигурируют со строчной буквой в начале слова. Также ошибочно в данную группу коллокатов был включен модификатор *joyeux* (122/4,93), примеры использования которого иллюстрирует преимущественно имя собственное (*Je trouve abusif par le professeur Joyeux de tout mélanger avec une allusion aux vaccinations*).

Словарные коллокаты *jeune, vieux, éminent, titulaire, principal, émérite* были зафиксированы и в выборке сервиса. Однако сочетаемость изучаемого существительного с прилагательными *bon, mauvais, certifié*,

agrégé, associé в данной выборке не подтверждена. Очевидно, что такие коллокации возможны, и возникает вопрос, почему они не обнаруживаются системой. Дело в том, что список коллокатов по показателю score заканчивается на значении 4,41, а сочетаемость с *bon* и *mauvais* обладает, как выяснилось, низкой уникальностью и была включена в выборку только по показателю frequency (1 267/3,34 и 91/2,15). Это очень невысокие показатели, что ярко проявляется при сравнении. Например, комбинация *professeur d'université* имеет показатель frequency 16 481 / score 9,8; *professeur de français* – 6 911/8,98, *professeur émérite* – 11 064/11,84. Интересный факт: коллокация *bon professeur* встречается в 14 раз чаще, чем *mauvais professeur*, что свидетельствует о положительной оценке преподавателя во французском обществе.

Что касается модификаторов *certifié, agrégé, associé*, то они имеют отглагольное происхождение, поэтому в данном сервисе их следует искать в выборке по модели **V + O (verbs with X as object)**. Ее анализ приводит к выводу, что наиболее частотными отглагольными формами являются *agrégé, invité, certifié, associé, diplômé, expérimenté, retraité, affilié, chevronné, qualifié, passionné, renommé, réputé*. Все перечисленные коллокаты распределяются по группам, описанным выше.

Проанализируем использование слова *professeur* в качестве дополнения к глаголу или именной части сказуемого (за исключением отглагольных форм, перечисленных выше). Модель **V + O (professeur)** включает: *nommer, devenir, être, recruter, expliquer, contacter, remercier, rencontrer, souligner, naître, suivre, déclarer, assister, aider, chercher*. Эта группа является самой неоднозначной и требует тщательного анализа. Часть глаголов (*nommer, recruter, suivre*) использована в контекстах как в личной форме, так и в отглагольной, в том числе входящей в причастные обороты. Преимущественно в личной форме используются глаголы *devenir, être, remercier, rencontrer, assister, aider, chercher*. Сопоставление со словарными коллокатами, приводимыми словарем Ortolang относительно глаголов *être, être nommé, devenir* (в других словарях они даже не были упомянуты), позволяет заключить, что эти глаголы относятся к списку наиболее типичных коллокатов существительного *professeur* и располагаются в верхних строках выборки.

Анализ коллокатов, построенных по данной модели, показал некоторые недостатки системы. Во-первых, сервис не учитывает порядок слов во французском предложении, что привело к включению глаголов *expliquer, souligner* и *déclarer* в данную выборку. Эти глаголы оформляют прямую речь, что требует инверсионного порядка слов: *En 2015, dans le monde, un décès sur dix était lié au tabagisme, explique le professeur Guy Joos; «Nous avons 60 % d'élèves boursiers», souligne un professeur d'Averroès; «On pourrait établir un parallèle avec l'adaptation biologique», déclare le professeur Otto*. Среди глагольных коллокатов со

значением менее 5,00 также находим подобные случаи: *préciser, dire, conclure, insister, commenter, noter* и др. Все эти глаголы относятся к семантической группе глаголов речевой деятельности и должны были наполнить модель **S (professeur) + V**, которую мы рассмотрим позже. Во-вторых, казус с глаголом *renaître*, который произошел в форме причастия прошедшего времени *né*, который совпадает с именем собственным *René*. Все попавшие в выборку случаи иллюстрируют именно эту случайную коллокацию: *les idées du Professeur René Dumont; un rapport du professeur René Piedelièvre* и др. В-третьих, сервис не смог различить совпадающие в настоящем времени формы глаголов *être* и *suivre* (*je suis*) и включил их все в выборку по глаголу *suivre*. Произошла путаница глаголов *étayer* и *être* (в скетче представлен первый глагол, в конкордансе наблюдаем исключительно форму *étaient*). По непонятной причине в данный скетч был внесен коллокат *émérite* (59/3,81), который также возглавляет список коллокатов – модификаторов, но уже с показателями 11 064/11,84. В выборке также присутствует коллокат *fiche* (119/4,53), который, видимо, был распознан системой как форма глагола *ficher*. Анализ конкорданса показал, что все приведенные примеры демонстрируют употребление этого слова в качестве существительного (*Ressources en ligne accessibles avec tablettes. Fiche professeur, fiche élève*). Это значит, что данный коллокат должен был быть включен в скетч, построенный по модели «**nouns modified by noun “professeur”**». Оказалось, что такая коллокация перечислена в данном скетче, но с другими показателями: 65/6,74.

Следующую, очень объемную группу составляют коллокаты, построенные по модели **N (professeur) + DE + N**: *université, philosophie, mathématique, école, universités, français, anglais, littérature, histoire, sociologie, eps, lycée, science, physique, lettre, histoire-géographie, théologie, piano, médecine, économie, psychologie, droit, musique, chant, chimie, collège, yoga, langue, danse, dessin, géographie, lettres, biologie, cour, allemand, éducation, faculté, espagnol, art, psychiatrie, rhétorique, sciences, guitare, enseignement, géologie, botanique, violon, composition, sport, linguistique, anatomie, anthropologie, conservatoire, secondaire, informatique, discipline, archéologie, maths, technologie, mathématiques, latin, physiologie, astronomie, classe, grec, écoles, chaire, physique-chimie, zoologie, degré, clinique, gymnastique, marketing, fle, académie, arabe, philologie, théâtre, orgue, mécanique, chirurgie, finance, management, neurologie, italien, génétique, russe, psychopathologie, département, religion, universités-praticien, arts, chinois*.

Данная модель дает представление о дисциплинах, преподаваемых преподавателями (*philosophie, mathématique, français, anglais, littérature, histoire*), о месте работы преподавателя, его принадлежности к среднему или высшему звену системы образования (*université, école,*

lycée, conservatoire). Поскольку сервис дает возможность вывода в скетч только коллокатов, представляющих собой исключительно биграммы, то коллокаты-триграммы не дифференцируются. Этот недостаток частично компенсируется расшифровкой, расположенной под выделенной биграммой. Так, можно понять, что коллокация с *art* является триграммой: *professeur d'arts plastiques*, а коллокация с *degré* включает *second*: *professeur du second degré* и т.д. На модель **N (professeur) + DE + N** приходится 229 152 случая употребления, что представляет 32,11% от общего количества употреблений с предлогами. Но сопоставление со словарями показывает, что коллокаты *de scepticisme, de vertu* (Dictionnaire de l'Académie Française en ligne), *de vertu, de sagesse* (Dictionnaire Ortolangue), их синонимы в данном корпусе отсутствуют. Таким образом, значение Fig., péj. *Personne qui prône une idée, un point de vue, une doctrine, etc., qui cherche à l'imposer, à l'inculquer aux autres* в этом корпусе не подтверждается.

Модель **N + DE + N (professeur)** представлена в словарях коллокациями с компонентами *chaire, salle, conseil, robe, épitoge*. Сочетаемость с компонентами *chaire, poste, salle, conseil* находит подтверждение в корпусе, с *robe, épitoge* – нет. Сервис приводит следующие коллокаты, не отраженные в словарях: *site, recrutement, direction, concours, fiche, métier, Association, postes, assistant, formation, diplôme, carrière, équipe, laboratoire, supervision, fonctions, grade, houlette, nomination*. Это самая неопределенная группа, скетч которой представлен в системе: в нем собраны все случаи употребления слова *professeur* после предлога *de*. Они никак не дифференцированы и потому не снабжены показателями частотности и типичности.

Модель **N (professeur) + A + N** заполняется следующими коллокатами: *université, faculté, collège, département, institut, conservatoire, école, lycée, sorbonne, académie, musée, séminaire, sciences, université, harvard, l, hec, cnam, escp, sciences-po, upmc, ens, iep, oxford, London, laboratoire, retraite, esse, stanford, chaire, uqam, iufm, beaux-arts, inalco, university Massachusetts, polytechnique, columbia, princeton, centre, temps, école, domicile, mines*. В данном скетче содержатся названия образовательных организаций и их структурно-административных подразделений, в которых работают преподаватели (*université, faculté, collège, département*), конкретных учреждений, расположенные как во Франции, так и в Великобритании и США (*Sorbonne, Harvard, UPMC, ENS, IEP, Oxford*), а также указания на социально-профессиональный статус преподавателя (*à la retraite, à temps plein / partiel / complet, à domicile*). Анализ несколько затруднен из-за замен заглавных букв на строчные, а также невозможности просмотра триграмм в отдельном списке с указанием частоты и степени типичности (выявить триграммы возможно лишь при просмотре конкорданса) и полной формы

коллокации с артиклями и предлогами. Сопоставление со словарными коллокатами показывает, что все они находят подтверждение в данном скетче.

Модель **N (professeur) + HORS + N** представлена небольшим количеством комбинаций, частотность которых чрезвычайно мала, но высока степень типичности: *professeur hors classe* (8/7,04), *professeur hors contrat* (4/6,15), *professeur hors pair* (19/5,9), *professeur hors champ* (2/5,48). Следовательно, сочетаемость *professeur hors classe*, предлагаемая в Dictionnaire de l'Académie française, подтверждается. Однако ввиду ее крайне редкой встречаемости, на наш взгляд, можно было бы ее и опустить, но привести другие, более частотные и актуальные для современного французского языка.

Модель глагольных коллокаций, где опорный компонент выступает в качестве субъекта **S (professeur) + V**, не нашла отражения ни в одном словаре. Это довольно объемная группа, объединяющая глаголы, дающие представление о типичных действиях преподавателя: *assister, enseigner, suppléer, gamberger, dispenser, encadrer, corriger, distribuer, demander, visiter, expliquer, veiller, remplacer, conseiller, encourager, adapter, remarquer, noter, transmettre, adjoindre, choisir, donner, apprendre, intervenir, guider, évaluer, régir*.

Как перечисленные выше, так и глагольные коллокаты со значением score ниже 5,0 объединяются в следующие семантические группы: глаголы речевой деятельности (*expliquer, déclarer, préciser, corriger, demander, conseiller, intervenir*), глаголы социального воздействия (*assister, enseigner, veiller, apprendre, guider, évaluer*), глаголы интеллектуальной деятельности (*choisir, savoir, penser*), замены (*suppléer, remplacer*), передачи объекта (*transmettre, donner, distribuer*), модальные глаголы (*pouvoir, devoir*), глаголы качественного состояния (*travailler, participer*) и некоторые другие. Все эти глаголы имеют нейтральную стилистическую окраску.

Как и в предыдущих случаях, были выявлены недостатки сервиса. В выборке оказался глагол *gamberger* (84/6,37), использующийся в разговорной речи. Однако анализ его конкорданса показывает, что все приведенные примеры содержат фамилию преподавателя (*Le professeur Gamberge est un savant un peu fou, toujours présent pour répondre aux questions des enfants...*). То же касается глагола *régir* (46/5): в конкордансе перечислены примеры с именем собственным (*Comme l'a expliqué le professeur Régis Aubry lors de son audition...*). Как видим, эти комбинации не столь уж многочисленны, но получают довольно высокую оценку степени уникальности. Снова фиксируем путаницу глаголов *étayer* и *être*: в скетче фигурирует *étayer*, а в конкордансе наблюдается *étaient*. Как происходило распределение *étaient* по скетчам с разными моделями — остается непонятным. Коллокат *foldingue* (20/4,3) также по

ошибке попал в данную выборку (*Il ressemblait de plus en plus à Doc, le professeur foldinguede Retour vers le futur*). Любопытно, что в скетче отсутствуют глаголы физического и душевного состояния, глаголы с эмоциональным значением. Можно предположить, что подобная сочетаемость неуместна с лексемой, использующейся для номинации профессии.

Итак, подведем итоги. Разработанная методика семантико-синтагматического анализа позволила на первом этапе выявить унифицированное лексикографическое значение слова *professeur*. Дальнейший анализ корпуса подтвердил актуальность первых трех значений. Остальные три, являющиеся результатом семантического переноса, требуют подтверждения и более тщательного поиска. Мы полагаем, что они окказиональны, не соответствуют общепринятому употреблению, обусловлены специфическими контекстами. В проанализированных контекстах слово *professeur* используется в номинативном (прямом) значении, его коллокаты тоже. С точки зрения нормы, узуса данные коллокации узуальны. Синтагматический анализ показал наличие большого количества лексем, с которыми изучаемое слово входит в довольно устойчивые коллокации, выстраивающиеся по ряду лексико-синтаксических шаблонов: компоненты этих коллокаций отражают постоянную связь понятий и с различной регулярностью воспроизводятся вместе. По логико-понятийному критерию изученные коллокации относятся к неустойчивым, среди них отсутствуют комбинации, имеющие фразеологически связанные, идиоматические значения. Слово *professeur* является стилистически нейтральным, и, как показал анализ, его коллокаты обладают преимущественно той же характеристикой. Коллокаты, имеющие положительную коннотацию, встречаются довольно редко (за исключением модели с адъективными модификаторами), негативную – в единичных случаях. С точки зрения стилистической окраски коллокаты слова *professeur* относятся к общеупотребительной лексике.

Разумеется, полученные данные отражают специфику используемого источника – веб-корпуса, сформированного из текстов новостных лент, газетных статей, научных, учебных, официально-деловых текстов, бесед в форумах, блогов, Википедии и других сайтов. В корпусе не представлена современная художественная литература и другие типы текстов, на которые распространяется авторское право. Поэтому нельзя считать, что данный корпус в полной мере отражает все особенности функционирования лексем и их коллокатов во французском языке на данном этапе его эволюции. Но используемый многомиллиардный корпус, тем не менее, охватывает огромную часть узуса, дает достаточно четкое представление о специфике употребления слов в речи, является надежным источником всех остальных типов текстов и, как показывает сопоставление со словарями, проливает свет на синтагма-

тическую сочетаемость слов и на ограничения этой сочетаемости значительно лучше. Полагаем, что подобные исследования могут дополняться сопоставлением с данными, полученными на материале других корпусов, для подтверждения или корректировки выводов.

Тестирование сервиса, которое мы запланировали как дополнительную задачу, дало положительные результаты. Сервис позволяет эффективно выявлять коллокации с учетом лексико-синтаксических шаблонов, ранжировать их по частотности или степени типичности, наблюдать конкретные реализации коллокаций, анализировать их ограничения. Объем анализируемого корпуса и качество обработки информации, базирующееся на применяемой методикой установления силы синтагматической связи, представляют для исследователя ценный и богатый материал. В ходе анализа были выявлены незначительные недостатки, которые следует учитывать при работе с сервисом и которые, как мы надеемся, разработчики смогут исправить. К этим недостаткам отнесем ошибки в распределении по лексико-синтаксическим шаблонам, связанные с 1) неучетом особенностей порядка слов во французском предложении (инверсия при оформлении прямой речи); 2) именами собственными, которые не идентифицируются системой и некорректно распределяются по разным моделям; 3) с трудностями идентификации начальной формы (лемматизации) слова в тех случаях, когда она совпадает с другой лексемой. Неудобным является отсутствие возможности работы с триграммами и более крупными коллокациями. Как полагаем, в работе с сервисом могут выявиться и иные ошибки и трудности. Все они, однако, обнаруживаются и корректируются при переходе пользователя к анализу конкретных контекстов, но вносят неточности в статистическую оценку частотности и степени уникальности коллокации. Повторимся, что при сравнении с другими сервисами они незначительны, не оказывают чувствительного влияния на общие результаты и частично могут быть скорректированы в ходе работы.

Заключение

Коллокация выполняет важную роль в речи: она уточняет и конкретизирует значение слов, которые ее образуют. Коллокациям свойственна структурно-семантическая целостность, их состав неслучаен и имеет целый ряд ограничений: логико-понятийных, семантических, сочетаемостных, стилистических. Эта информация, безусловно, должна быть отражена в словарях. Однако данные о синтагматической сочетаемости слова *professeur*, приведенные во французских словарях, подтверждаются в нашем исследовании реального языкового материала лишь частично. Причина, как мы полагаем, кроется, в том числе, в материалах, к которым обращаются лексикографы. Словари иллюстрируют особен-

ности функционирования лексем в речи, опираясь либо на литературу, датированную XVI–XX вв., базы данных ATILF, BDLP, словаря литературы XIX–XX вв. TLFi и другие источники (словарь *Ortolang*), либо на собрание литературных произведений проекта Gutenberg и корпуса *Europarl*, включающего тексты протоколов Европарламента за 1996–2011 гг. (*Le Petit Robert*), либо на примеры, составленные самими редакторами словаря (*Larousse, Dictionnaire de l'Académie française*), либо привлекают иные ресурсы.

Объем и количество коллокаций, вводимых в словарь, их источники, безусловно, зависят от лексикографических установок разработчиков словаря. Но именно узус, на наш взгляд, должен являться одним из определяющих факторов включения слов в словари и их лексикографического описания с учетом изменений постоянно эволюционирующей языковой системы. Не можем не согласиться и поддерживаем вывод В.П. Захарова о том, что корпусно-статистический подход позволяет создавать невозможные ранее описания языка [10. С. 83]. Данный подход, как показало исследование, эффективен и открывает широкие перспективы для изучения реального функционирования слов в речи.

Список источников

1. *Иорданская Л.Н., Мельчук И.А.* Смысл и сочетаемость в словаре. М. : Языки славянских культур, 2007. 672 с.
2. *Захаров В.П., Хохлова М.В.* Анализ эффективности статистических методов выявления коллокаций в текстах на русском языке // Компьютерная лингвистика и интеллектуальные технологии. 2010. № 9 (16). С. 137–143.
3. *Онал И.О.* Коллокация как объект изучения в отечественной и зарубежной науке: диахронический аспект // Вопросы языкознания в мультикультурном пространстве современного мира. Новосибирск : Изд-во ЦРНС, 2018. С. 169–191.
4. *Ахманова О.С.* Словарь лингвистических терминов. М. : Советская энциклопедия, 1966.
5. *Борисова Е.Г.* Коллокации: что это такое и как их изучать. М. : Филология, 1995. 49 с.
6. *Manning C., Shutze H.* Collocations // Foundations of Statistical Natural Language Processing. Cambridge, Massachusetts : The MIT Press, 2002. P. 151–189.
7. *Баранов А.Н., Добровольский Д.О.* Классификация фразеологизмов: новый взгляд на классическую проблему // Слово, фразеологизм, текст в литературном языке и говорах : сб. ст. Орёл : ОГУ, 2010. С. 80–87.
8. *Кунин А.В.* Курс фразеологии современного английского языка. Москва : Высш. шк. ; Дубна : Феникс, 1996. 381 с.
9. *Копотев М.В., Стексова Т.И.* Исключение как правило. Переходные единицы в грамматике и словаре. М. : Языки славянских культур, 2016. 170 с.
10. *Захаров В.П.* Лингвистика больших корпусов // Компьютерная лингвистика и вычислительные онтологии : сб. науч. ст. Труды XVIII объедин. конф. «Интернет и современное общество» (IMS-2015) (СПб., 23–25 июня 2015 г.). СПб. : Университет ИТМО, 2015. С. 82–93.
11. *Корпус* Университета Лидса. URL: <http://corpus.leeds.ac.uk/it/>
12. *Корпуса* Aranea. URL: http://ucts.uniba.sk/aranea_about/
13. *Корпус книг* Google Books Ngram Viewer. URL: <https://books.google.com/ngrams>

14. **АОТ** (Автоматическая обработка текстов). URL: http://aot.ru/online_demo.html
15. **Сервис Sketch Engine**. URL: <https://www.sketchengine.eu/>
16. **Кротова Е.Б.** Sketch Engine для лингвистических исследований // Германистика сегодня : материалы Междунар. науч.-практ. конф. (Казань, 16–17 октября 2018 г.) / под ред. М.А. Кулковой. Казань : Изд-во Казан. ун-та, 2019. С. 107–112.
17. **Феминизация** названий профессий во французском языке : доклад Французской академии. URL: http://www.academie-francaise.fr/sites/academie-francaise.fr/files/rapport_feminisation_noms_de_metier_et_de_fonction.pdf
18. **Dictionnaire** Le Petit Robert de la langue française 2021 / éd. A. Rey, J. Rey-Debove, P. Robert. Paris : Le Robert, 2020. 2836 p.
19. **Dictionnaire** Le Petit Larousse illustré 2020. Paris : Librairie Larousse, 2020. 2048 p.
20. **Dictionnaire** de l'Académie Française en ligne. 9^{ème} éd. URL: <https://www.dictionnaire-academie.fr/article/A9P4464>
21. **Dictionnaire** Ortolang. URL: <https://www.cnrtl.fr/definition/professeur>
22. **Виноградова О.Е.** Унифицированное лексикографическое значение как тип значения // Вестник Воронежского государственного университета. Серия: Филология. Журналистика. 2012. № 2. С. 32–34.
23. **Влавацкая М.В.** Комбинаторно-семантический анализ коллокаций как метод лингвистического исследования (на примере колоративных коллокаций, образованных по адъективному типу) // Научный диалог. 2020. № 4. С. 40–57.

References

1. Iordanskaja L.N., Mel'chuk I.A. (2007) Smysl i sochetaemost' v slovare [Meaning and compatibility in the dictionary.]. M.: Jazyki slavjanskih kul'tur. 672 p.
2. Zakharov V.P., Khokhlova M.V. (2010) Analiz effektivnosti statisticheskikh metodov vyjavenija kollokacij v tekstakh na russkom jazyke [Analysis of the effectiveness of statistical methods for detecting collocations in texts in Russian] // Komp'juternaja lingvistika i intellektual'nye tehnologii. 9 (16). pp. 137-143.
3. Onal I.O. (2018) Kollokacija kak objekt izuchenija v otechestvennoj i zarubezhnoj nauke: diahronicheskij aspekt [Collocation as an object of study in domestic and foreign science: a diachronic aspect] // Voprosy jazykoznanija v multikul'turnom prostranstve sovremenogo mira: monografija. Novosibirsk: Izdatel'stvo CRNS. pp. 169-191.
4. Akhmanova O.S. (1966) Slovar' lingvisticheskikh terminov [Dictionary of linguistic terms]. M.: Izd-vo Sovetskaja enciklopedija.
5. Borisova E.G. (1995) Kollokacii: chto eto takoe i kak ikh izuchat' [Collocations: what they are and how to study them]. M.: Filologija. 49 p.
6. Manning C., Shutze H. Collocations (2002) Foundations of Statistical Natural Language Processing. pp. 151-189.
7. Baranov A.N., Dobrovol'skij D.O. (2010) Klassifikacija frazeologizmov: novyj vzgljad na klassicheskiju problemu [Classification of phraseological units: a new look at the classic problem] // Slovo, frazeologizm, tekst v literaturnom jazyke i govorah: sbornik statei. Orel: OGU. pp. 80-87.
8. Kunin A.V. (1996) Kurs frazeologii sovremenennogo anglijskogo jazyka [Course of phraseology of the modern English language]. M.: Vysshaja shkola; Dubna: Feniks. 381 p.
9. Kopotev M.V., Steksova T.I. (2016) Iskljuchenie kak pravilo. Perekhodnye edinicy v grammatike i slovare [Exception as a rule. Transitional units in grammar and vocabulary]. M.: Jazyki slavjanskih kul'tur. 170 p.
10. Zakharov V.P. (2015) Lingvistika bol'shih korpusov [Linguistics of the large corpora] // Komp'juternaja lingvistika i vychislitel'nye ontologii: sb. nauch. statei. Trudy XVIII objedin. konf. "Internet i sovremennoe obshchestvo" (IMS-2015). SPb: Universitet ITMO. pp. 82-93.
11. Korpus Universiteta Lidsa [Leeds University corpus]. URL: <http://corpus.leeds.ac.uk/it/>
12. Korpusa Aranea [Aranea corpus]. URL: http://ucts.uniba.sk/aranea_about/

13. Korpus knig Google Books Ngram Viewer [Corpus of books of Google Books Ngram Viewer]. URL: <https://books.google.com/ngrams>
14. AOT (Avtomaticeskaja obrabotka tekstov) [Automatic word processing]. URL: <http://aot.ru/onlinedemo.html>
15. Servis Sketch Engine [Sketch Engine service]. URL: <https://www.sketchengine.eu/>
16. Krotova E.B. (2019) Sketch Engine dlja lingvisticheskikh issledovanij [Sketch Engine for linguistic research] // Germanistika segodnya. pp. 107-112.
17. Feminizacija nazvanij professij vo francuzskom yazyke. Doklad Francuzskoj Akademii [Feminization of occupational names in French: Report of the French Academy]. URL: http://www.academie-francaise.fr/sites/academie-francaise.fr/files/rapport_feminisation_noms_de_metier_et_de_fonction.pdf
18. Rey A., Rey-Debove J., Robert P. (2020) Dictionnaire Le Petit Robert de la langue française 2021. Paris: Le Robert. 2836 p.
19. Dictionnaire Le Petit Larousse illustré 2020 (2020). Paris: Librairie Larousse. 2048 p.
20. Dictionnaire de l'Académie Française en ligne. 9ème éd. URL: <https://www.dictionnaire-academie.fr/article/A9P4464>
21. Dictionnaire Ortolang. URL: <https://www.cnrtl.fr/definition/professeur>.
22. Vinogradova O.E. (2012) Unificirovanoe leksikograficheskoe znachenie kak tip znachenija [Unified lexicographic value as a type of value] // Vestnik Voronezhskogo gosudarstvennogo universiteta. Seriya: Filologija. Zhurnalistika, 2. pp. 32-34.
23. Vlavackaja M.V. (2020) Kombinatorno-semanticheskij analiz kollokacij kak metod lingvisticheskogo issledovanija (na primere kolorativnykh kollokacij, obrazovannykh po adjektivnomu tipu) [Combinatorial-semantic analysis of collocations as a method of linguistic research (on the example of colorative collocations formed according to the adjective type)] // Nauchnyj dialog, 4. pp. 40-57.

Информация об авторах:

Богоявленская Юлия Валерьевна – доктор филологических наук, доцент, профессор кафедры лингвистики и профессиональной коммуникации на иностранных языках, Уральский федеральный университет им. первого Президента России Б.Н. Ельцина (ул. Мира, 19, Екатеринбург, Россия, 620002); профессор кафедры немецкого и французского языков, Петрозаводский государственный университет (пр. Ленина, 33, Петрозаводск, Россия, 185035). E-mail: jvbog@yandex.ru.

Жесткова Анна Сергеевна – студент, Петрозаводский государственный университет (пр. Ленина, 33, Петрозаводск, Россия, 185035). E-mail: zhest1806anna@gmail.com

Вклад авторов: все авторы сделали эквивалентный вклад в подготовку публикации. Авторы заявляют об отсутствии конфликта интересов.

Information about the authors:

Bogoyavlenskaya Yu.V., D.Sc. (Philology), Associate Professor, Professor of the Chair of Linguistics and Professional Communication in Foreign Languages, Ural Federal University named after the first President of Russia B.N. Yeltsin (Mira str., 19, Yekaterinburg, Russia, 620002); Petrozavodsk State University (Lenin Ave., 33, Petrozavodsk, Russia, 185035). E-mail: jvbog@yandex.ru

Zhestkova A.S., student, Petrozavodsk State University (Lenin Ave., 33, Petrozavodsk, Russia, 185035). E-mail: zhest1806anna@gmail.com

Contribution of the authors: the authors contributed equally to this article. The authors declare no conflicts of interests.

Поступила в редакцию 12.04.2021; принята к публикации 04.05.2022

Received 12.04.2021; accepted for publication 04.05.2022