

# К вопросу о базовой лексике и принципах ее описания в двуязычном словаре сочетаемости для начинающих

Irina Makarova Tominec

UP FHŠ, Oddelek za slovenistiko, Oddelek za uporabno jezikoslovje, Inštitut za medkulturne študije

irina.mt@fhs.upr.si

---

## Введение

В практике преподавания иностранных языков проблема определения и корректного описания так называемой базовой лексики не теряет своей актуальности. Усвоение иностранного языка проходит через определенные этапы, от простого к сложному, и уже на самом начальном этапе важно дать учащемуся те слова, которые являются самыми употребительными и в качестве таковых дадут ему необходимый минимум языковых средств для общения в рамках ограниченного набора тем и грамматических конструкций. Успешность обучения таким образом во многом зависит как от определения самого состава «самых употребительных» слов, так и от их корректного представления в имеющихся дидактических материалах.

При ближайшем рассмотрении оказывается, что как вопрос определения состава базовой лексики, так и вопрос ее адекватного представления представляет собой отдельную лингводидактическую проблему. В нашей статье мы представим принципы формирования лексических минимумов, проведем сопоставление двух действующих лексических минимумов для РКИ (элементарный и базовый) с новым частотным словарем русского языка. Основное внимание будет уделено представлению принципов описания базовой лексики в

рамках нового словаря сочетаемости русского языка для словенцев, работа над которым не прекращается и в настоящее время.<sup>1</sup>

## 1. Принципы выделения базовой лексики

Проблема создания списка слов для различных уровней языковой компетенции, на который можно было бы ориентироваться в практике преподавания, имеет свою историю. Первый список так называемой базовой лексики английского языка, известный как *General Service List of English Words*, был опубликован Майклом Уэстом в 1953 году, в результате длительной работы коллектива специалистов в течение как минимум двух десятилетий. Список содержал приблизительно 2 000 лексических единиц, выбранных в результате анализа частотности английских слов в письменных текстах. Как выяснилось, владение ограниченным объемом базовой лексики обеспечивает достаточно высокий процент понимания текстов на соответствующем языке (West 1953).

В 90-ые годы список базовой лексики был уточнен и дополнен в работах Пола Нэшена, преподавателя-англиста из Новозеландского университета, который с помощью методов корпусной лингвистики обновил списки базовой лексики для различных уровней владения английским языком (Nation 2004, Nation 2016). В 2000 году был опубликован список так называемой академической лексики, обеспечивающей понимание текстов научного характера (Coxhead 2000). В последнее десятилетие количество списков базовой английской лексики для учебных целей еще более возросло.<sup>2</sup> Создание списков, как правило, основывается на статистическом анализе репрезентативных корпусов национального языка, в состав которых включены неспециализированные тексты разной жанровой, стилистической принадлежности и максимально разнообразной тематической направленности. При этом по возможности учитываются как письменные, так и устные тексты. При отборе лексических единиц важен не только показатель абсолютной частотности (*frequency rank*), но и коэ-

1 Рабочая версия в свободном доступе, см. Makarova Tominec in dr. 2019.

2 В 2013 году были опубликованы обновленные списки базовой и академической лексики (Browne et al. 2013a, Browne et al. 2013b). Свой список базовых английских слов внесен в учебные планы обучения английскому языку в Израиле и рекомендован для использования на международном уровне (Revised English Curriculum 2018).

фициент их распределения (*dispersion rank*) по текстам самых разных жанров и стилей. Для создания неспециализированных лексических списков отбирается высокочастотная лексика, которая регулярно используется в максимально широком наборе письменных и устных текстов, что определяется различными методами статистического (корпусного) анализа. В качестве дополнительных критериев в литературе (Nation 2004, Nation 2016) приводятся также стилистическая нейтральность, словообразовательная активность и методическая обоснованность (*teacheability*). В частности, принципиальной особенностью списков английской лексики является указание не отдельного слова, а словесной семьи (*word family*), в которой наряду со словом указывается набор частотных аффиксов, участвующих как в слово-, так и в формообразовании (Bauer & Nation 1993). Таким образом, современный подход к формированию списков базовой лексики основывается на статистической обработке репрезентативного корпуса национального языка.

Задачей определения базового списка русской лексики с 70-ых годов прошлого столетия активно занимался ряд авторов, в частности Петр Никитич Денисов и особенно Валерий Вениаминович Морковкин. Последний руководил авторским коллективом «Лексических минимумов русского языка», в котором термин лексический минимум получил свою подробную методологическую разработку. Лексические минимумы были определены как «лексические перечни нарастающего объема, охватывающие лексическое ядро современного русского языка», предназначенные для усвоения учащимися в течение определенного промежутка времени (Морковкин 1985, 10). Указанные перечни слов были предложены в качестве рекомендуемой основы для создания учебных материалов для различных уровней языковой компетенции. В первой версии «Лексических минимумов» (Морковкин 1985) речь идет о семи градуальных списках, содержащих от 500 до 3 500 полнзначных слов, каждый предыдущий список целиком включался в последующий и одновременно наращивался на 500 дополнительных лексических единиц. В следующей версии (Богачева и др. 2003)<sup>3</sup> набор списков достиг количества

3 В 2003 году коллективом авторов Государственного института РЯ им. Пушкина был опубликован труд «Система лексических минимумов русского языка», который также содержит перечень структурных слов и 10 лексических списков, включающих 5 000 самых важных русских слов. Концептуальный подход к выделению базовой лексики остался прежним.

десяти, в состав которых в общей сложности вошло приблизительно 5 000 слов. Состав лексики был определен на базе сравнительного анализа восьми частотных словарей русского языка, данные которых сведены к единому знаменателю по специально выработанной методике. Статистическая основа представлена в отдельном разделе, в виде подробных таблиц с данными по частотности лексических единиц. Тематическая принадлежность включенных лексических единиц определялась на базе анализа программ, учебников и учебных пособий по РКИ и представлена в отдельном разделе «Тематический словарь-минимум современного русского языка».

Знаменательной особенностью первых ЛМ является их преимущественная ориентация на пассивные виды речевой деятельности (чтение, слушание и перевод с русского языка на родной). Второй особенностью русских ЛМ по версии 1985 года было выделение отдельного, так называемого «нулевого» списка основных структурных лексических единиц, которые «обеспечивают функционирование полнозначных (знаменательных) слов и в конечном счете формируют синтаксическую и модальную структуру текста» (Морковкин 1985, 7). В этот список были включены слова различной частеречной принадлежности: существительные с количественным значением, существительные, соотносимые с фазовыми глаголами, десемантизированные глаголы в составе глагольно-именных образований (*носить характер, вести войну*), фазовые глаголы, количественные, порядковые и собирательные числительные, ряд раздельнооформленных наречий, местоимений и вводных слов (*до упора, друг друга, одним словом*), а также группа неоднословных частиц речи, предлогов, союзов, союзных слов, частиц и междометий (*в духе, да и, куда бы ни, не что иное как, вот это да*). При этом усвоение функциональных слов из нулевого списка было отнесено на более поздние этапы обучения.

В 2000-е годы лексические минимумы были частично обновлены и переосмыслены. Связано это с тем, что они стали «важнейшим компонентом» российской государственной системы тестирования, которая была включена в европейскую структуру языкового тестирования ALTE и приведена в соответствие с общеевропейскими стандартами. Новые лексические минимумы сформированы с целью определить словарный запас иностранцев, изучающих русский язык, на различных уровнях языковой компетенции. Система новых лексических минимумов русского языка включает 6 уровней: элемен-

тарный (A1), базовый (A2), первый сертификационный (B1), второй сертификационный (B2), третий сертификационный (C1) и четвертый сертификационный уровень (C2). Каждый уровень владения русским языком базируется на соответствующем лексическом минимуме. На элементарном уровне список слов максимально минимизирован и составляет 780 единиц, на базовом – около 1300 слов, на первом сертификационном – 2300 слов, на втором около 5 000 слов, на третьем – около 10 000. Для последнего уровня набор слов не определяется.

Сопоставление новых лексических минимумов со старыми показывает, что, с одной стороны, они продолжают традицию лексических минимумов Морковкина, в плане разграничения слов, предназначенных для активного и пассивного усвоения. Новые лексические минимумы так же содержат спецификацию тем, предназначенных для активного усвоения на разных уровнях обучения, и включают соответствующий раздел со словами разной частеречной принадлежности, сгруппированными на тематической основе (тематический словарь-минимум). С другой стороны, новые лексические минимумы предлагают альтернативное решение, в частности, в рамках описания служебной лексики: последние включены в основной список слов, их усвоение таким образом методически распределено по различным уровням обучения. Как отмечают составители лексических минимумов, в качестве критерия отбора базовой лексики были использованы следующие показатели: 1) стилистическая нейтральность, 2) способность входить в словосочетания, 3) семантическая значимость (способность слова обозначать часто встречающиеся предметы и явления), 4) способность слова к словообразованию (слова, имеющие высокую словообразовательную способность, служат для создания потенциального словаря), 5) частотность (при этом учитывались показатели частотности по частотным словарям, по использованию в учебниках РКИ, «тематическая» частотность).

Если для списков базовой английской лексики, созданных на основе корпусного анализа, для каждого уровня декларируется способность понимания определенного процента текстов общей направленности (*coverage*), российские ЛМ этой информации не содержат. В них декларируется способность овладевшим ими учащимся понимать и создавать тексты определенного объема. Данные о статистической зависимости между частотностью русских слов и их встречаемостью в текстах были представлены в «Русском частотном словаре» (Шаров

2001), построенном на основе предварительного корпуса СРЯ. В нем приведены следующие данные:

- 1000 наиболее частотных лемм покрывает 64.07% текста,
- 2000 наиболее частотных лемм покрывают 71.95% текста,
- 3000 наиболее частотных лемм покрывают 76.51% текста,
- 5000 наиболее частотных лемм покрывают 82.06% текста.

Приведенные данные впрочем сложно напрямую сопоставить с действующими лексическими минимумами, предложенными для иностранных учащихся на различных уровнях языковой компетенции. В связи с этим в рамках настоящего исследования было решено соотнести ЛМ базового уровня (1300 лексических единиц) с данными Нового частотного словаря РЯ (Ляшевская, Шаров 2009), составленного на базе Национального корпуса русского языка. Сопоставление было произведено с помощью программы корпусного анализа *AntConc*, которая позволяет статистически вычислить процент совпадения соотносимых списков. Как оказалось, в числе первых 1300 наиболее частотных лемм можно обнаружить только 45% слов, включенных в ЛМ базового уровня (см. рис.1).

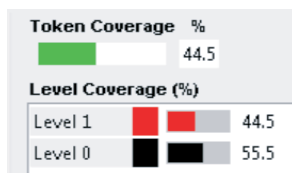


Рис. 1. Совпадение первых 1300 самых частотных лемм с ЛМ базового уровня

Данный показатель свидетельствует о том, что частотность слов действительно не является приоритетным критерием включения лексических единиц в действующие лексические минимумы. Поскольку они были разработаны для ограниченного тематического набора, прописанного для каждого уровня, мы решили проверить процент лексического совпадения с текстами из интернет-источника «Корпуса звучащей речи», тематика которых совпадает с заявленным тематическим диапазоном ЛМ. Для сравнения были использованы 20 неспециализированных (устных) текстов небольшого объема (до 250 слов) из сферы повседневного общения (о семье). Как показал анализ,

положительная корреляция подтвердилась, процент совпадения достиг 70% (см. ниже).

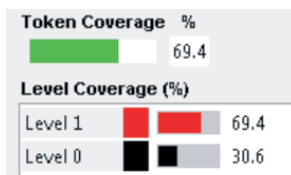


Рис. 2. Совпадение лексического состава избранных устных текстов на тему «семья» с ЛМ базового уровня (A<sub>2</sub>)

Разработка новых дидактических материалов для усвоения лексики в словенской аудитории, как нам кажется, должна основываться на действующих лексических минимумах, являющихся «важнейшим компонентом» российской системой тестирования, в связи с чем в целях создания относительно небольшого в числовом отношении нового двуязычного словаря (в настоящий момент около 1100 словарных статей) был использован лексический минимум базового уровня (A<sub>2</sub>), с частичной заменой и наращением списка лексических единиц (см.ниже).

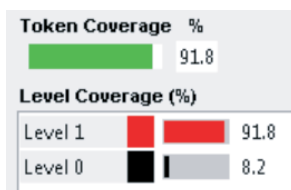


Рис. 3. Совпадение словаря нового двуязычного словаря для начинающих с ЛМ базового уровня (A<sub>2</sub>)

Разница между словарем нового двуязычного словаря и ЛМ связана с исключением ряда, по нашему мнению, малоактуальных в словенской аудитории или устаревших слов (*аспирант, аспирантка, аспирантура, булочная, магнитофон, марка, милиционер, милиция, молодежный, молодежь, народ, народный, правительство*), с добавлением новой предметной лексики из области цифровых технологий и соцсетей (*смартфон, гаджет, приложение, программа, соцсеть, пост*), а также с включением некоторых слов из ЛМ следующего уровня (*бывший, бухгалтер*). В целом, согласно рекомендациям

действующих ЛМ, создаваемые на их основе дидактические материалы могут отступать от рекомендованного списка на 10%.

## 2. Базовая лексика: лексикографический аспект

Наряду с критериями отбора базовой лексики, не менее важным является вопрос способа описания и представления лексического ядра: что собственно значит «знать слово», какая информация о слове должна быть предоставлена иностранцу, чтобы обеспечить не только адекватное понимание/распознавание определенного количества русских слов, но и их правильное употребление.

Если целью является предоставление списка слов для активного владения, принципиально важным признается предоставление разнотрановой информации о слове, включающей в себя как собственно-лексические, так и грамматические аспекты (Nation 2004, Nation 2016). Если говорить о русской лексике, наряду со значением слова иностранцу будет полезна информация о частотности слова, о его частеречной отнесенности, приведение набора основных грамматических форм, указание возможных особенностей и исключений, о месте слова в системе парадигматических отношений (гипероним, антоним, синоним), сведения о лексической и синтаксической сочетаемости, примеры словосочетаний и предложений. Полезной информацией для усвоения русской лексики являются словообразовательные связи слова, поскольку «словообразование органически связано с пополнением лексического запаса учащихся» (Иванова 1977, 73).

Существенным элементом, обеспечивающим возможность активного употребления слова, особенно на начальном этапе, является его перевод на родной язык учащихся. Как в свое время писал В.В. Морковкин, если определять словарь активного типа как «такой, который обеспечивает переход от мысли к языковому выражению, и помнить, что мысль у иноязычных учащихся формируется на родном языке, становится очевидным, что между мыслью и требуемым языковым выражением (русским!) появляется средний член, именно иноязычное слово или словосочетание, через посредство которого только можно прийти к русскому слову или словосочетанию» (Морковкин 1977, 91).

Таким образом сложился концепт двуязычного словаря для начинающих, который строится по принципу представления сочетаемости слова в зависимости от его частеречной принадлежности, на



основе базовых грамматических (синтаксических) связей. Раскрытие сочетаемости на основе синтаксических связей, с одной стороны, способствует раскрытию лексического значения слова, с другой стороны, иллюстрирует способы его вхождения в предложения.

Далее подробнее приводим структуру словарной статьи.

### 3. Структура словарной статьи

#### 3.1 Заголовочное слово

Заголовочное слово приводится с обозначенным местом ударения, с указанием части речи и аудиозаписью произношения. Обычно в заглавной же строчке указывается перевод слова на словенский язык. Глаголы совершенного и несовершенного вида, которые в традиционной русской лексикографии считаются одной лексемой (Ожегов, Шведова 1992), в словаре для начинающих с учетом их различных грамматических свойств (набора временных форм) рассматриваются по отдельности, в разных словарных статьях. В этом случае отсылка к соотносимому слову (глаголу СВ или НСВ) приводится сразу в заглавной строчке (*говорить – сказать*).

В качестве средства семантизации наряду с переводом также используется визуальное изображение (картинка).

В связи с тем, что в настоящее время словарь содержит относительно небольшое количество (чуть более 1100) заголовочных слов, которые на 90% совпадают с действующим лексическим минимумом ТРКИ базового уровня (A2), в текущей версии словаря специальные отметки с указанием частотности заголовочного слова отсутствуют.

#### 3.2 Место в лексической системе

Сразу под заголовочным словом указываются возможные гиперонимы, синонимы, антонимы. Особенно важным кажется приведение гиперонима, слова с вышестоящим значением, так как оно открывает возможности к парафразированию как к одной из частотных компенсирующих стратегий речевого поведения говорящих на начальных уровнях языковой компетенции.

Важно отметить, что развернутость системы лексических отношений заголовочного слова тесно связано с его значением, так, например, наличие антонимов/синонимов достаточно трудно установить для предметной лексики. Прослеживается определенная

зависимость лексических связей и от частеречной принадлежности заголовочного слова: для качественных прилагательных (*новый, интересный*) бывает трудно подобрать прилагательное-гипероним со значением, более конкретным, чем «свойство/качество».

### 3.3 Сочетаемость лексическая и синтаксическая

Модель сочетаемости заголовочных слов в предлагаемом концепте двуязычного словаря строится в зависимости от их частеречной принадлежности, т.е. на базовых лексикограмматических характеристиках.

Сочетаемость существительного «бабушка» в словаре представлена следующим образом. На первом месте показано несколько вариантов сочетаний заголовочного слова с согласованным прилагательным (*какая бабушка?*). Далее следует набор частотных предикатов, сочетающихся с заголовочным словом (*что делает бабушка?*). Ряд эксплицируемой сочетаемости замыкают частотные варианты, связанные с комбинаторными возможностями различных предложно-падежных форм заголовочного слова (*бабушку навестить/слушаться*), (*бабушке помогать/позвонить*), (*к бабушке зайти/заехать*), (*у бабушки спросить/жить*), (*о бабушке заботиться*).

Информация о глаголе эксплицирует его лексическую валентность, модель управления: сочетаемость представлена в наборе трех-пяти вариантов лексического заполнения синтаксических позиций, возможных при заголовочном глаголе (*бежать куда? откуда? за кем? через что? где? как?*), включая стандартную для глагола позицию обстоятельства (*как?*), а также сочетания с фазовыми (*начать/перестать/броситься бежать*) и модальными глаголами (*любить/уметь готовить*), если они при данном глаголе частотны и употребительны.

Сочетаемостные возможности заголовочного прилагательного представлены максимально наглядно и последовательно, сначала в сочетании с существительными единственного числа мужского, женского и среднего рода, затем с существительными множественного числа. При этом существительные не повторяются (*бедный человек, бедная женщина, бедное население, бедные студенты*). Систематически приводятся сочетания заголовочного прилагательного с интенсификатором (*очень /довольно/абсолютно/совсем не*), если такие сочетания употребительны. Систематически приводятся

сочетания прилагательных с глаголами (*быть/стать/казаться/оказаться/считать+ каким*).

Если в позиции заголовочного слова указано наречие, что происходит в случае особой семантической значимости данного наречия (*бесплатно*), приводятся частотные сочетания с глаголами (*бесплатно отдать, получить*), а также с интенсификаторами (*абсолютно, совершенно*).

Указание словосочетаемости показало себя результативным в случае экспликации значения предлогов, так как разные лексико-семантические варианты, как правило, связаны с различными предложно-падежными конструкциями, следующими за предлогом в том или ином его значении (*на чём/на ком? ≠ на кого/что?*).

В случае экспликации значения вопросительных местоимений (*где, что, когда, какой, который, чем*), которые могут выполнять функцию союзных слов, обе функции приводятся в одной словарной статье, но разделены по значениям. В одном случае они указаны как вопросительные местоимения, в другом – как союзные слова, вводящие придаточные предложения.

### 3.4 Грамматические формы слова

Учащиеся на начальном этапе обучения часто нуждаются в эксплицитном указании *основных* грамматических форм русских слов, согласно программе изучения языка на базовом уровне (Vozelj, Urbas 2008). С учетом более широких возможностей цифровых словарей, для существительных приводится полный набор предложно-падежных форм, для глаголов эксплицировано спряжение во всех временных формах (настоящее, прошедшее и будущее время), формы императива в единственном и множественном числе, а также формы сослагательного наклонения. Прилагательные приводятся в сочетании с существительными женского, мужского и среднего рода во всех падежных формах в единственном и множественном числе. Отдельно указываются краткие формы (если они употребительны). Для заголовочных прилагательных и наречий приводятся также формы сравнительной и превосходной степени, иллюстративные предложения содержат примеры развернутых сравнительных конструкций.

Важно отметить, что в отдельных случаях наличие определенных грамматических форм связано только с определенным значе-

нием заголовочного слова. Если рассмотреть многозначное существительное «дело», широкий набор форм глаголов страдательного наклонения чаще встречаются с одним его лексико-семантическом вариантом, а именно «юридическая процедура» (*дело возбуждено, открыто, закрыто, рассмотрено, передано в суд*). Поскольку в тематический диапазон начинающих данная сфера не входит, в двуязычном словаре это значение не указывается, как и связанные с ним формы страдательного наклонения.<sup>4</sup>

### 3.5 Предупреждение типичной ошибки

Дифференциальный характер<sup>5</sup> двуязычного словаря мы понимаем как необходимость учета родного языка учащихся в описании лексических единиц. В практике составления словарных статей этот принцип означает, что в случае, если при усвоении заголовочного слова в словенской аудитории часто возникают ошибки, словарная статья должна содержать специальную рубрику, предупреждающую их возникновение. В нашем случае это рубрика *Posebnosti*, которая обращает внимание учащихся на возможные сложности и несовпадения в значении либо функционировании заголовочного слова, что может привести к его ошибочному употреблению.

Типологию лексических ошибок, подлежащих рассмотрению в двуязычном учебном словаре для начинающих, невозможно разработать в отрыве от грамматики. Как нам представляется, именно для начального уровня языковой компетенции взаимозависимость лексики и грамматики весьма велика, чем мы и обосновываем синергический принцип выделения возможных сложностей, допускающий включение грамматических по своей природе сложностей в зону словарной статьи, предназначенную для предупреждения типичных ошибок.

К числу частотных ошибок, наблюдаемых в русскоязычных текстах, написанных словенскими учащимися на начальном уровне языковой компетенции, мы относим явление расширенной паронимизации, которое понимаем как ошибочное словоупотребление в результате (полного или частичного) наложения значения и функции

4 В качестве альтернативного решения возможно включение в учебный словарь соответствующего значения и связанных с ним форм, но с экспликацией более высокого уровня языковой компетенции (Capel 2010).

5 Термин заимствован из работ В.В. Морковкина (1977).

внешне похожих слов (по звучанию или по написанию). Имеются в виду частотные случаи смешения слов *delati* и *делать*, *iti* и *идти*, *znati* и *знати*, *familija* и *фамилия*, предлогов *na* и *на*, словоформ *boljši* и *больше*, в результате интерференции родного языка.

Отдельная сложность – предупреждение калькирования моделей грамматического управления, в том случае, когда они в русском и словенском языке не совпадают. Именно на это в словаре сочетаемости обращается особое внимание: *позвонить кому-то* (при словенском эквиваленте с зависимым словом в винительном падеже *roklicati koga*), *поздравлять кого-то* (при словенском эквиваленте с зависимым словом в дательном падеже *čestitati komu*), *поблагодарить кого-то* (при словенском эквиваленте с зависимым словом в дательном падеже *zahvaliti se komu*).

Особую сложность для усвоения представляет ограниченная лексическая сочетаемость, особенно если она не совпадает с родным языком учащихся. Например, русское прилагательное *хороший* имеет иные комбинаторные возможности, чем его словенский эквивалент *dober*: *хорошая погода – lepo vreme*, *хорошее дело – pametna zadeva*, *хорошие слова – lepa beseda*.

### 3.6 Иллюстративные предложения

Иллюстративные предложения содержат наиболее типичные контексты использования заголовочного слова, в словаре для начинающих они, как правило, связаны с устными формами речи. Одновременно с этим иллюстративные предложения должны показать примеры функционирования грамматических форм и конструкций, которые могут представлять сложность для словенских учащихся.

*Как у вас дела?* Первое по списку иллюстративное предложение для заголовочного существительного *дело* иллюстрирует клишированную форму речевого этикета.

*Людмила Улицкая – популярный автор.* Данный пример выбран для иллюстрации типичного для русского языка использования существительных мужского рода из тематической группы профессии/занятий для наименования женщин, что совсем не свойственно словенскому языку.

В отдельных случаях при составлении иллюстративных предложений использовались примеры из Национального корпуса русско-

го языка, структура которых в случае необходимости упрощалась (грамматически и лексически).

### 3.7 Словообразование

Одним из важных аспектов описания базовой лексики мы считаем включение частотных словообразовательных связей заголовочного слова. Именно усвоение словообразовательных связей позволит учащемуся расширять свой (как минимум пассивный) словарный запас.

В текущей версии словаря состав данной рубрики разнороден и намеренно минимизирован. Как правильно, приводятся частотные феминитивы (*писатель-писательница, француз-француженка*), деминутивы (*рука – ручка*), а также однокоренные с заголовочным словом лексические единицы различной частеречной отнесенности (*спорт – спортивный, турист – туристический*).

### 3.8 Связанные словосочетания

Идиоматический аспект при работе с иноязычной лексикой является особенно проблематичным, так как в большинстве случаев речь идет о многословных сочетаниях, значение которых не выводимо из значения составляющих эти сочетания лексических единиц. По этой причине указанные сочетания представляют особую сложность для усвоения и требуют отдельной проработки.

В текущей версии двуязычного словаря сочетаемости для начинающих данная рубрика довольно разнородна: в ней представлены клишированные выражения, свойственные разговорной речи (*Доброе утро!*), контекстуально ограниченные примеры использования, а также сочетания, идиоматизированные в различной степени (*не в этом дело*). В отдельных случаях в текущей версии словаря с целью минимизировать количество значений многозначных слов в этой рубрике представлены контекстуально ограниченные примеры использования заголовочных слов в переносном значении (*идти – время/фильм/речь идет; бежать – время бежит*).

## 4. Развивающая направленность словаря

Двуязычный словарь сочетаемости для начинающих предусматривает широкие возможности развития словарного запаса

пользователей словаря, так как 1) предлагает перевод использованных словосочетаний на словенский язык, 2) набор сочетаемости включает помимо прототипических, самых частотных, еще и набор дополнительных, синтаксически однотипных, вариантов, 3) представляет слово в контексте его лексических связей (гиперонимы, синонимы, антонимы), 4) представляет словообразовательный потенциал заголовочного слова, 5) представляет идиоматически ограниченные формы, 6) предупреждает типичные ошибки словоупотребления с учетом родного языка учащихся.

В текущей версии словарь включает 1 176 заголовочных русских слов, что на 91,8% совпадает со списком слов, входящих в лексический минимум базового уровня владения русским языком (уровню А2, согласно Общеревропейскому языковому портфолио), что, по нашей оценке, может соответствовать словарному запасу словенских учащихся после первого года обучения русскому языку (напр., в гимназии).

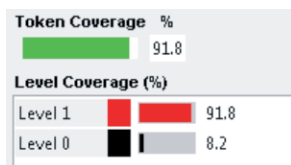


Рис.4. Процент совпадения словаря двуязычного словаря с ЛМ (А2)

Поскольку словарь сочетаемости показывает заголовочные слова в более широком контексте их функционирования, в составе широкого набора словосочетаний или иллюстративных предложений, словарные статьи в целом включают значительно большее количество русских слов. В этом состоит обучающая направленность словаря, задачей которого является расширение активного словарного запаса учащихся.

## 5. Визуализация

Развитие цифровых словарей значительно расширило возможности лексиграфического описания. В качестве существенного преимущества может быть рассмотрен, в частности, возросший размер словарной статьи по сравнению с печатной версией, а также

дополнительные технические возможности - цветные иллюстрации, аудиозапись, ссылки.

При концептуализации словаря значительное внимание было уделено также визуальному облику словарной статьи, одним из существенных элементов которой стали 1) параллельный перевод на словенский язык всех слов и словосочетаний, 2) наличие картинки, 3) возможность поступенчатого представления различных вариантов сочетаемости, в отличие от традиционного перечисления вариантов сочетаемости через запятую, принятого в печатных словарях, 4) приведение основных грамматических форм заголовочного слова в полном объеме, отказавшись от специальных обозначений классов, рядов, типов и так далее. Приоритетным принципом в данном случае является максимальная доступность и понятность словаря, упрощение его использования в целевой аудитории.

В качестве иллюстративного примера рассмотрим словарную статью для заголовочного слова «автор».

Рис. 5. Пример словарной статьи

### **Áвтор** samostalnik, moški spol **avtor, avtorica**



Vir: pixabay

►... автóр avtor, avtorica  
совреме́нный sodoben  
молодо́й mlajši  
изве́стный znan  
зарубе́жный tuj  
начаи́ающий na začetku kariete  
талантливы́й nadarjen  
популярный priljubljen  
(мой/наш) любимы́й moj/naš najljubši

►áвтор ... avtor, avtorica  
книги knjige



статья članka  
прое́кта projekta  
ромáна romana  
стихотворéния pesmi  
произведéния dela  
интернет-поста́ posta  
сообщéния sporočila

►... áвтором  
статья postati avtor(-rica)  
быть biti avtor(-ica)  
являться biti avtor(-ica)  
счита́ть šteti koga za avtorja(-ico)

►áвтор ... avtor, avtorica  
пишет piše  
счита́ет meni  
говори́т pravi  
предлага́ет predlaga  
описывае́т opisuje  
критикуе́т kritizira

►... áвтора avtorja, avtorico  
зна́ть poznati  
назва́ть imenovati  
указа́ть navesti, omeniti  
вспо́мнить / забы́ть (имя) áвтора spomniti se / pozabiti (ime) avtorja,  
avtorice  
цити́ровать citirati  
хва́лить / крити́ковáть hvaliti / kritizirati

►... на áвтора na avtorja, avtorico  
де́лать ссылку sklicevati se  
ссылáться sklicevati se

►... с áвтором  
согласи́ться soglašati  
встрéтиться dobiti se z avtorjem/avtorico

познакомиться spoznati avtorja, avtorico  
 встреча srečanje z avtorjem, avtorico  
 интервью intervju z avtorjem, avtorico  
 договор pogodba z avtorjem/avtorico

►... автору avtorju, avtorici  
 позволить omogočiti  
 принадлежать pripadati  
 ответить odgovoriti  
 верить verjeti  
 принести успех, признание, гонорар prinesiti uspeh, priznanje,  
 honorar

		Ednina	Množina
Imenovalnik	Это - кто?	автор	авторы
Rodilnik	Нет когó?	автора	авторов
Dajalnik	Дать книгу комú?	автору	авторам
Tožilnik	Любить когó?	автора	авторов
Orodnik	Говорить с кем?	автором	авторами
Mestnik	Думать о ком?	об авторе	об авторах

Людмила Улицкая - популярный автор. Ljudmila Ulickaja je priljubljena avtorica.

А вы не знаете, кто автор (этого произведения)? Ali veste, kdo je avtor (tega dela)?

Автором картины является Владимир Малевич. Avtor slike je Vladimir Maljevič.

Автором проекта считается Андрей Петров. Avtor projekta naj bi bil Andrej Petrov.

авторский (экземпляр) avtorski (izvod)

## 6. Отбор сочетаний в словарь: результаты и дискуссия

Существенным моментом, влияющим на полноту и актуальность представляемой информации о сочетаемости слова, является вопрос о способах/критериях отбора подходящих словосочетаний. В качестве основного инструмента мы использовали платформу статистического анализа языковых корпусов *Sketch Engine*, которую

можно применить в том числе и для анализа ряда русскоязычных корпусов.<sup>6</sup>

Для отбора частотных предикатов при субъекте «автор» на платформе *Sketch Engine* была получен следующий список частотных глаголов, ранжированных, с одной стороны, по устойчивости образуемых сочетаний (индекс из 3 цифр), с другой стороны, по частоте (индекс из 5 цифр) (см. таблицу ниже).

Словосочетания ранжированные по устойчивости				Словосочетания ранжированные по частотности			
автор as noun 4,637,774×				автор as noun 4,637,774× Sorted by freq			
subject_of				subject_of			
писать	32,684	8.08	...	писать	32,684	8.08	...
пишет автор				пишет автор			
утверждать	14,505	7.82	...	считать	29,582	7.62	...
считать	29,582	7.62	...	быть	20,351	2.74	...
отмечать	18,781	7.27	...	предлагать	19,315	6.9	...
подчеркивать	8,127	7.21	...	отмечать	18,781	7.27	...
рассказывать	12,922	7.16	...	говорить	17,071	5.86	...
автор рассказывает				говорит автор			

Рис.6. Результаты анализа словосочетаний с помощью платформы *Sketch Engine*

Как видим, часть глаголов из левого и правого столбца совпадает (*писать, считать, отмечать*), а часть нет. С учетом того, что целевая аудитория словаря – начинающие изучать русский язык, набор предоставляемых сочетаний количественно ограничен (около 5 позиций). При окончательном отборе предпочтение дается совпадающим коллокациям, т.е. тем из них, которые показали высокую устойчивость (8.08, 7.62 и 7.27) и высокую частотность (32,684, 29,582 и 18,781 на миллион словупотреблений) одновременно.

6 Для наших целей мы использовали корпус Russian Web 2011 sample (ruTen-Ten11), самой большой из русскоязычных корпусов на *Sketch Engine*, насчитывающий 998 миллионов словупотреблений.

Другой пример – выбор словосочетаний на позицию левого согласуемого прилагательного к слову «автор» (какой автор?).

Словосочетания ранжированные по устойчивости				Словосочетания ранжированные по частности			
<b>a_modifier</b>				<b>a_modifier</b>			
<b>античный</b>	5,906	8.54	...	<b>разный</b>	14,390	6.85	...
античных авторов				разных авторов			
<b>неизвестный</b>	7,884	8.46	...	<b>современный</b>	13,849	6	...
неизвестного автора				современных авторов			
<b>зарубежный</b>	11,109	8.07	...	<b>молодой</b>	12,311	6.55	...
и зарубежных авторов				молодых авторов			
<b>начинающий</b>	4,826	8.01	...	<b>известный</b>	11,153	6.41	...
начинающих авторов				известных авторов			
<b>анонимный</b>	2,845	7.8	...	<b>зарубежный</b>	11,109	8.07	...
анонимный автор							

Рис. 7. Результаты анализа коллокаций с помощью платформы *Sketch Engine*

Как видим, и в этом случае самые устойчивые (слева) и самые частотные словосочетания (справа) совпадают не полностью. Дополнительный анализ контекстов употребления показал, что прилагательное *разный*, указанное в качестве самого частотного определения перед существительным *автор*, обычно используется в конструкциях во множественном числе (*разных авторов*, *разными авторами*), т.е. функционально ограничено, в связи с чем было исключено из итогового набора, который выглядит следующим образом: 1) *современный*, 2) *молодой*, 3) *известный*, 4) *зарубежный*.

Отдельного комментария требует возможность включения сочетаний существительных с качественными прилагательными типа *хороший/плохой*, «малоинформативных», с точки зрения традиционной лексикографии (см. Морковкин 1977). Преимущественным критерием, влияющим на включение словосочетания в двуязычный словарь, мы считаем дифференциальную направленность двуязычных словарей, что означает обязательный учет родного языка целевой аудитории, в связи с чем словосочетания, структура которых в родном и целевом языке не совпадает (*хорошая погода* ≠ *lepo vreme*), отбира-

ются в первую очередь, в связи с высокой вероятностью возможной ошибки.

Как показал наш опыт словарной работы, в качестве дополнительного инструмента контрастивного анализа лексики можно использовать русский ворднет (RuWordNet). Речь идет о лексической базе данных (11,5 тысяч слов и выражений), разработанной для различных задач процессирования речи искусственным интеллектом (например, для целей машинного перевода), в которых русские существительные, глаголы и прилагательные организованы в концептуальные иерархические списки.

Ворднет как инструмент может быть весьма полезен при указании места заголовочного слова в лексической системе языка, поскольку воспроизводит лексические отношения системного характера, по принципу: вышестоящее/нижестоящее слово, антонимические отношения, синонимические отношения, отношения часть – целое, отношение частеречной синонимии. Рассмотрим пример качественного прилагательного *хороший*. Русский ворднет включает его вместе с синонимом (*положительный*) в состав вышестоящего концепта *качество, свойство*, предоставляет широкий набор потенциальных синонимов, относящихся к восемнадцати концептуальным полям, а также широкий набор антонимов.

Использование ресурса русского ворднета оказалось полезным при семантизации слова *бабушка*. В русском языке слово *бабушка* используется не только для обозначения члена семьи, но и для обозначения любой пожилой женщины. Словенскому лексическому эквиваленту *babica* такой перенос значения не свойственен, для второго значения будет использовано сочетание *starejša gospa*. Использование электронной онтологии концептуальных полей русского языка, таким образом, может облегчить и ускорить работу по составлению словаря.

Использование этого источника имеет свои ограничения, в чем можно убедиться на следующем примере. Русский ворднет включает слово *актер* в набор синонимов (синсет) *актер, актриса, артист, артистка, лицедей, лицедейка*, формирующих семантическое поле (домен) *притворщик*. Словенский перевод *igravec* в таких контекстах встречается реже. Синонимический словарь словенского языка приводит слово *glumač*, корпус *gigafida* содержит некоторое количество примеров со словом *glumec*, которое встречается и в словаре литера-

турного словенского языка. В связи с тем, что в русском языке второе, переносное значение (актер = притворщик) реализуется в ограниченном наборе контекстов, а именно когда существительное *актер* занимает позицию предикатива, чаще в устных высказываниях оценочного характера (*ну ты актер!*), в двуязычном словаре для начинающих этими примерами, как кажется, можно пренебречь. Ограниченное функционирование переносного значения подтверждает и дифференциальный анализ слов (актер: притворщик) с помощью статистического анализа контекстов употребления *Sketch Engine*, который показал, что набор дифференциальных контекстов очень невелик, переносное значение актуализуется за счет предстоящих прилагательных с отчетливой негативной коннотацией (*лживый, бессовестный, лицемерный*) или же потенциально возможно в сочетании с ограниченным набором глаголов (*сбежать, убежать, влюбиться*). В связи с тем, что прилагательные, в контексте которых это значение фиксируется, не входят в ЛМ, а при глаголах это значение лишь потенциально возможно (не обязательно), считаем, что включение данного значения в двуязычный словарь сочетаемости для начинающих было бы излишним.<sup>7</sup>

## Выводы

Таким образом, описание русской лексики в двуязычном словаре сочетаемости для начинающих базируется на следующих соображениях: 1) опора на ядерную лексику языка, позволяющую при сравнительно небольшом количестве единиц обеспечить понимание половины слов в текстах общего характера, 2) представление слова в рамках конструкции – словосочетания или предложения. Обоснование подобной точки зрения обнаруживаем на разных уровнях: во-первых, включение слова в конструкцию эксплицирует возможности его включения в речь и таким образом облегчает его использование. Во-вторых, привязка заголовочного слова к частотным конструкциям, контекстам его использования набирает популярность и на теоретическом уровне, в рамках конструктивной грамматики, согласно которой конструкции, в которые включаются лексические единицы, далеко не случайны, часто идиоматичны/идиосинкретичны и содержат много важной семантической информации о конкретном

7 Альтернативным решением могло бы быть включение данных контекстов с переводом в рубрику иллюстративных примеров или связанных сочетаний.

слове (Goldberg 2006). Об особом значении конструкций писали и лексикографы-практики (Морковкин 1985), которые подчеркивали важность экспликации в словарях сочетаемости как регулярной грамматической сочетаемости, так и «собственной», индивидуально-ограниченной сочетаемости описываемого слова.

Далее, с учетом дидактической направленности предлагаемого двуязычного словаря, важнейшим, если не преимущественным принципом становится дифференциальный характер представления лексического материала. Это означает, что в зону преимущественного внимания попадают точки несовпадения лексических единиц двух языков, поскольку они с высокой долей вероятности становятся причиной ошибки. Это в равной степени относится как к различиям на уровне значения (или концепта), так и к различиям на уровне употребления (грамматических форм, синтагматики). Представляется, что предупреждение лексической интерференции – одна из важнейших задач двуязычных дидактических словарей. В этом плане эффективным инструментом качественной доработки такого словаря можно считать учет типичных ошибок в русскоязычных текстах, создаваемых словенскими учащимися.<sup>8</sup>

Отдельной задачей является, естественно, актуализация лексического состава словаря путем включения слов из области соцсетей, цифровых технологий, а также лексики, связанной со словенскими реалиями. Параллельно важна и работа по постепенному расширению набора лексических единиц.

В качестве существенного ограничения при создании двуязычных дидактических пособий необходимо отметить тот факт, что не вся информация, подлежащая экспликации в словаре, на данный момент поддается статистическому анализу, что не позволяет полностью объективизировать выбор лексических единиц. Думается, что может оказаться весьма полезным и параллельный корпус, работа над которым, по нашим сведениям, ведется в Институте русского языка им. В.В. Виноградова и на кафедре русистики Филологического факультета Люблянского университета.<sup>9</sup>

8 Работа по расширению корпуса таких текстов особенно перспективна для создания двуязычных пособий, в настоящее время ведется на базе Факультета гуманитарных наук Приморского университета.

9 Выражаем благодарность др.Младену Ухлику за предоставленные сведения.

В качестве полезного дополнения к двуязычным дидактическим словарям для начинающих мог бы оказаться комплекс упражнений, сгруппированных вокруг самых частотных лексических ошибок (паронимия, несовпадение грамматического управления, различия в лексической сочетаемости, идиоматические словосочетания).

Все вышесказанное предполагает лексикографическую работу по сопоставлению русского и словенского языка в нескольких перспективных, на наш взгляд, направлениях: учет типичных ошибок на базе корпуса русскоязычных текстов, написанных словенскими учащимися, включение переводных эквивалентов из параллельного корпуса текстов, учет результатов тестирования рабочей версии словаря в словенской аудитории, с последующей доработкой и расширением представляемых сведений о русском слове. В связи с этим возникнет и необходимость периодического обновления двуязычного словаря сочетаемости, который, как мы надеемся, сможет найти своего пользователя в словенской аудитории.

## Источники и литература

### Источники

- Andryushina N.P. in Kozlova, T.V. 2015. *Lexical minimum of Russian as a foreign language. Level A1. Common language*. 5th edition. Sankt-Petersburg: Zlatoust
- Andryushina N.P. in Kozlova, T.V. 2015. *Lexical minimum of Russian as a foreign language. Level A2. Common language*. 5th edition. Sankt-Petersburg: Zlatoust.
- Andryushina N.P. et al. 2015. *Lexical minimum of Russian as a foreign language. Level B1. Common language*. 7th edition. Sankt-Petersburg: Zlatoust.
- Богачева, Г.Ф., Луцкая Н.М., Попова З.П., Морковкин В.В. 2003. Система лексических минимумов современного русского языка. 10 лексических списков. От 500 до 5000 самых важных русских слов. Москва, АСТ Астрель. [Bogačeva, G.F., Luckaja N.M., Popova Z.P., Morkovkin V.V. 2003. *Sistema leksičeskikh minimumov sovremennogo ruskogo jazyka. 10 leksičeskikh spisov. Od 500 do 5000 samyh važnyh russkih slov*. Moskva: AST Astrel'].
- Денисов П.Н. и др. 1983. Словарь сочетаемости слов русского языка: ок. 2500 словарных статей. Под ред. П.Н. Денисова и В.В. Морковкина. Москва: Институт русского языка им. Пушкина: Издательство



Русский язык. 2-ое издание. [Denisov, P.N. idr. 1983. *Slovar sočetaemosti slov ruskogo jazyka: ok.2500 slovarnyh statej*. Pod red. P.N. Denisova i V.V. Morkovkina. Moskva: Institut ruskogo jazyka im.Puškina: Izdatel'stvo Russkiy jazyk. 2-oe izdanie].

Морковкин, В.В. и др. 1985. Лексические минимумы русского языка. Под ред. В.В. Морковкина. Москва: Русский язык. [Morkovkin, V.V. idr. 1985. *Leksičeskie minimumy ruskogo jazyka*. Pod red. V.V. Morkovkina. Moskva: Russkiy jazyk].

Ожегов С.И., Шведова, Н.Ю. 1992. Толковый словарь русского языка. Москва: Издательство Азъ. [Ožegov, S.I. in Štedova N.Ju. 1992. *Tolkovyyj slovar' ruskogo jazyka*. Moskva: Izdatel'stvo Az.]

Шанский Н.М. и др. 1981. 4000 наиболее употребительных слов русского языка. Москва: Издательство Русский язык. [Šanskiy, N.M. idr. 1981. *4000 naibolee upotrebitel'nyh slov ruskogo jazyka*. Moskva: Izdatel'stvo Russkiy jazyk.]

Revised English Curriculum Including Band III Lexics. Principles and Standards of Learning English as an International Language for All Grades. State of Israel. Ministry of Education. Pedagogical Secretariat. Language Department. English Language Education. 2018.

West M., 1953. *A General Service List of English Words*. London: Longman, Green & Co.

### Литература

Иванова, Т.Ф. 1977. «Частичное гнездование слов в учебном толковом словаре для иностранцев.» Проблемы учебной лексикографии. Под ред. П.Н. Денисова и В.В. Морковкина. Москва: Издательство Московского университета. 72–84. [Ivanova, T.F. 1977. »Častičnoe gnezdovanie slov v učebnom tolkovom slovare dlja inostrancev.« *Problemy učebnoj leksikografii*. Pod red. P.N. Denisova i V.V. Morkovkina. Moskva: Izdatel'stvo Moskovskogo universiteta. 72–84.]

Морковкин В.В. 1977. «Русско-иноязычный словарь сочетаемости, его лингвистические основы и структура». Проблемы учебной лексикографии. Под ред. П.Н. Денисова и В.В. Морковкина. Москва: Издательство Московского университета. 84–110. [Morkovkin V.V. 1977: «Russko-inojazyčnyj slovar' sočetaemosti, ego lingvističeskie osnovy i struktura». *Problemy učebnoj leksikografii*. Pod red. P.N. Denisova i V.V. Morkovkina. Moskva: Izdatel'stvo Moskovskogo universiteta. 84–110.]

- Bauer, L. & Nation, I. S. P. 1993. »Word families.« *International Journal of Lexicography*, 6 (4): 253–279.
- Capel, Annette. 2010. »A1–B2 Vocabulary: Insights and Issues Arising from the English Profile Wordlists Project.« *English Profile Journal 1*. Published online by Cambridge University Press. <https://doi.org/10.1017/S2041536210000048>;
- Goldberg, Adele E. 2006. *Constructions at Work: The Nature of Generalization in Language*. Oxford Scholarship Online. E-book.
- Nation, I.S.P. 2001. «Using small corpora to investigate learner needs.» *Small Corpus Studies and ELT. Theory and practice*. John Benjamins Publishing Company. Amsterdam/Philadelphia, 31–45.
- Nation, I.S.P., 2004. «A study of the most frequent word families in the British National Corpus.» V P. Bogaards & B. Laufer (Eds.). *Vocabulary in a Second Language Acquisition: Selection, Acquisition and Testing* (pp.3-13). Amsterdam: John Benjamins.
- Nation, I.S.P. 2016. *Making and Using Word lists for Language Learning and Testing*. Amsterdam: John Benjamins.

#### *Интернет-источники*

- Browne, Charles, Culligan, Brent, Phillips, Joseph. 2013a. *New General Service List: the most important word for second language learners of English*. <http://www.newgeneralservicelist.org/>
- Browne, Charles, Culligan, Brent, Phillips, Joseph. 2013b. *New Academic Word List*. <http://www.newgeneralservicelist.org/nawl-new-academic-word-list>
- Coxhead, Averil. 2000: *A new academic word list*. *Tesol Quarterly* 34 (2): 213–238. <https://doi.org/10.2307/3587951>
- Кибрик А.А. и др. *Корпуса звучащей речи*. [Kibrik A.A. et al. *Korpora zvučašćej reči* <http://spokencorpora.ru/>]
- Ляшевская, О.Н., Шаров С.А., 2009. Частотный словарь современного русского языка (на материалах Национального корпуса русского языка). Москва: Азбуковник. [Ljaševskaja, O.N. in Šarov, S.A. 2009. *Častotnyj slovar' sovremennogo ruskogo jazyka (na materialah Nacional'nogo korpusa ruskogo jazyka)*. Moskva: Azbukovnik <http://dict.ruslang.ru/freq.php>]
- Makarova Tominec, Irina idr. 2019. *Rusko-slovenski kolokacijski slovar za začetnike / Russko-slovenskij slovar sočetaemosti dlja načinajuščih*. [https://www.termmania.net/slovarji/243/Rusko\\_slovenski\\_slovar\\_za\\_zacetnike](https://www.termmania.net/slovarji/243/Rusko_slovenski_slovar_za_zacetnike)

Nation, I.S.P. 2017. *The BNC/COCA Level 3 partial word family lists* (Versions 1.0.0) [Data file]. <http://www.victoria.ac.nz/lals/staff/paul-nation.aspx>

Шаров С. А. 2001: Частотный словарь русского языка [Šarov, S.A. 2001: *Častotnyj slovar' russkogo jazyka*. <http://www.slovorod.ru/freq-sharov/index>]

*Thesaurus of Russian language RuWordNet*. <https://ruwordnet.ru/en>

Vozelj, Ines in Urbas, Janja. 2008. *Učni načrt. Ruščina. Gimnazija: splošna, klasična, strokovna gimnazija: obvezni predmet in matura (420 ur), izbirni predmet (140 ur)* [http://www.mss.gov.si/fileadmin/mss.gov.si/pageuploads/podrocje/ss/programi/2008/Gimnazije/UN\\_RUSCINA\\_gimm.pdf](http://www.mss.gov.si/fileadmin/mss.gov.si/pageuploads/podrocje/ss/programi/2008/Gimnazije/UN_RUSCINA_gimm.pdf)

## Povzetek

*O jedrnem besedišču in pristopih k leksikografskem opisu le-tega v dvojezičnem kolokacijskem slovarju za začetnike*

Članek se osredotoča na načela oblikovanja seznamov osnovnega, t.i. jedrnega besedišča (*core lexics*) tujega jezika, prirejenih za didaktične namene, in sicer za usvajanje na različnih ravneh jezikovne kompetence. Posebej se obravnava pojem *leksikalnega minimuma*, izdelan v okviru didaktike ruskega jezika. Relevantnost veljavnih leksikalnih minimumov ruskega jezika se utemeljuje na podlagi primerjave le-teh s podatki Novega frekvenčnega slovarja ruske leksike, ki je sestavljen z upoštevanjem statistične analize besedišča Nacionalnega korpusa ruskega jezika. Ugotovljeno je, da pri izboru temeljnega besedišča poleg frekvenčnosti veljajo merila stilistične nevtralnosti, semantične vrednosti, kompatibilnosti in besedotvorne aktivnosti posameznih besed.

V drugem delu članka se podrobneje predstavijo izhodišča za leksikografski opis osnovnega besedišča v okviru predlaganega koncepta dvojezičnega kolokacijskega slovarja ruščine. V članku je poudarjen diferencialni pristop k predstavitvi besedišča, ki izpostavi tipične napake v primeru rusko-slovenskega stika tako na ravni pomena, kot tudi na ravni slovnic. Podrobno se predstavi struktura slovarskega članka. Kot prednost digitalne oblike predlaganega slovarskega priročnika se izpostavi možnost sprotne dodelave opisa, z upoštevanjem rezultatov nastajajočega korpusa pisnih esejev v ruščini, ki so jih sestavili slovenski učenci.

## Summary

*On core lexis wordlists and lexicographic approaches used for bilingual learner dictionary of collocations (beginner level)*

The article highlights the principles of compiling wordlists of core lexis, tailored for didactic use at different levels of foreign language competence. Special attention is dedicated to the term lexical *minimum*, developed within the framework of didactics of Russian (Morkovkin). The validity of actual lexical minimums of Russian is analysed by comparing them with a wordlist of the New frequency dictionary of Russian lexis. It is found that selection criteria for Russian lexical minimums, in addition to frequency, also apply the criteria of stylistic neutrality, semantic value, compatibility, and word-formation activity.

The second part of the article presents in more detail the lexicographic approaches within the proposed concept of a bilingual collocation dictionary of Russian. The article emphasizes the contrastive approach to the bilingual vocabulary that highlights the typical errors at both the level of meaning and the level of grammar. The structure of the dictionary-article is presented in more detail. The possibility of updating the vocabulary description, using data of an emerging learner corpus of Russian compiled by Slovene students, is highlighted as an advantage of the digital format of the proposed dictionary.