

**Синтаксическое поведение идиом и неидиоматических выражений:
корпусное исследование (на материале немецкого языка)¹**

Е. Б. Кротова (Институт языкознания РАН)

Аннотация

В данной статье сравнивается синтаксическое поведение немецких идиом с глагольным компонентом *gehen* с синтаксическим поведением самого глагольного компонента и с их парафразами. Также рассматривается ряд неидиоматических выражений, содержащих тот же глагольный компонент. Проверяется значимость различий между синтаксическим поведением идиом и неидиоматических выражений.

Ключевые слова

Идиомы, синтаксис, корпусная лингвистика, фразеография, обработка естественного языка.

Идиомы² с точки зрения синтаксиса интересны тем, что обладают дефектной парадигмой: модификации, допустимые для свободных словосочетаний, не всегда допустимы для идиом. Верно и обратное: так, к примеру, в [Иомдин 2003, 217] говорится о возможности фразем нарушать синтаксические нормы и стандарты, существующие вне сферы идиоматики (напр., правила согласования и управления). Есть отдельные идиомы, для которых утверждение о дефектной парадигме кажется не требующим доказательств. Например, *abwarten und Tee trinken* 'поживём — увидим', которая употребляется только в этой форме. Однако для большинства идиом наличие дефектности парадигмы не очевидно и требует проверки на больших корпусных данных.

¹ Исследование выполнено при поддержке гранта РФФИ № 18-012-00335А «Ориентационные метафоры во фразеологии: особенности употребления и лексико-грамматических изменений фразеологизмов в славянских, германских и кельтских языках».

² В данной статье под идиомами понимаются «сверхсловные образования, которым свойственна высокая степень идиоматичности и устойчивости». Подробнее в [Баранов, Добровольский 2008, 57].

В статье анализируется синтаксическое поведение 18 употребительных идиом немецкого языка, содержащих глагольный компонент *gehen* ‘идти’. Рассматривается употребление идиом в пассиве, в сослагательном наклонении, в инфинитивных конструкциях с *zu*, с модальными глаголами и с отрицанием. Также большое внимание уделяется употреблению идиом в различных временных формах: презенс (Präsens), претерит (Präteritum), перфект (Perfekt), футур I (Futur I). Сравнение будет проводиться как с самим глаголом *gehen*, так и с парафразами к идиомам и с 12 частотными свободными словосочетаниями с *gehen* (*ins Schwimmbad gehen* ‘идти в бассейн’, *einkaufen gehen* ‘идти за покупками’ и т.д.). Предполагается, что, с одной стороны, на синтаксическое поведение идиомы будет накладываться ограничения глагольный компонент, с другой стороны, идиома будет вести себя сходно со своим парафразом. Также проверяется, будет ли синтаксическое поведение неидиоматических выражений похоже на поведение идиом.

Предварительно нами уже была собрана информация по указанным выше модификациям для 100 идиом немецкого языка, содержащих различные глагольные компоненты (подробнее в [Krotova 2018; Krotova в печати]), что позволит сопоставить данные по глагольному компоненту *gehen* с частотностями, характерными для идиом.

Целью статьи является проверка гипотезы о несхожести синтаксического поведения рассматриваемых идиом и неидиоматических выражений.

Текстовый материал был взят из архива корпусов Института немецкого языка в г. Мангейме (W — Archiv der geschriebenen Sprache) [DeReKo]. На его основе был создан виртуальный корпус, содержащий только немецкие источники, начиная с 1990 г. (австрийские и швейцарские источники были исключены из рассмотрения). Корпус состоит в основном из публицистики, процент художественной литературы в нем крайне мал, устной речи не представлено.

DeReKo является одним из самых больших корпусов немецкого языка. Большой объем корпуса позволяет изучать даже те идиомы, которые нечасто встречаются в письменной речи. Следует отметить, что идиомы характерны скорее для устной речи, поэтому

даже те идиомы, которые по опросам носителей языка употребительны, могут быть слабо представлены в корпусе. К тому же Мангеймский архив в основном состоит из газетных корпусов, а некоторые идиомы довольно сложно представить себе в публицистическом тексте (например, если они не относятся к нейтральному или разговорному стилю, а являются сниженными).

Тексты в корпусе по большей части не являются синтаксически аннотированными, а для целей исследования требовалось наличие грамматической информации. Поэтому для каждой рассматриваемой идиомы или неидиоматического выражения из корпуса скачивалась выдача, которая затем отдельно аннотировалась. По каждому поисковому запросу корпус позволяет скачать до 10 тысяч примеров.

Запросы для идиом формулировались максимально широкие. Например, для идиомы *an die Nieren gehen* ‘брать [хватать] за живое (кого-л.); сильно подействовать (на кого-л.)’ запрос выглядел следующим образом: (an /+w2 &Niere) /s0 &gehen. То есть, между предлогом *an* и существительным *Niere* может быть один токен. *Niere* является леммой, а не словоформой. В одном предложении с этой синтагмой должна встретиться лемма *gehen*. Таким образом, запрос является максимально широким, чтобы не исключать возможное варьирование.

Получившийся корпус (до 10 тысяч примеров на одну идиому / словосочетание) аннотировался с помощью программы GATE (General Architecture for Text Engineering) — системы обработки естественного языка [GATE].

При обработке использовались имеющиеся в GATE средства для обработки немецкого языка (German Tokeniser, German Sentence Splitter, German Stanford POS Tagger, German POS-mapping Feature Gazeteer, German gazeteer, German grammar). Также мною были созданы словарь для глаголов и правила для обработки контекстов.

Данные аннотируются следующим образом. Искомое словосочетание помещается в тэг *Idiom*. Внутри тэга помещается грамматическая информация о глаголе (в каком вре-

мени встретился, в каком лице и числе), а также может указываться наличие адвербиальных или атрибутивных модификаций.

Аннотированные данные можно скачать из GATE в формате xml и далее обрабатывать их. Пример разметки в формате xml из GATE для идиомы *an die Nieren gehen* ‘брать [хватать] за живое (кого-л.); сильно подействовать (на кого-л.)’:

Nichts ist mehr zu spüren von der Bedrängnis, die ihn geplagt haben muß. <Idiom gate:gateId="26998" number="sg" adv="sehr" perfekt="sein" tense="perfekt" person="3">Wie sehr ihm diese Zeit *an die Nieren gegangen ist*</Idiom>, zeigt die letzte Szene, eine Dokumentaraufnahme. (Berliner Zeitung, 03.06.1998)

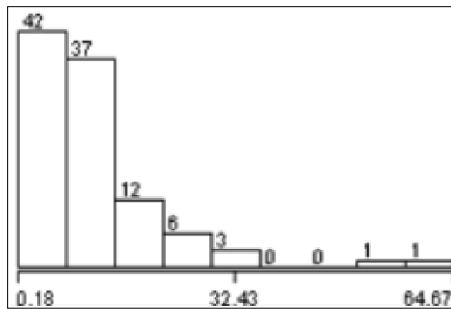
В первую очередь посмотрим на результаты анализа для глагола *gehen*. В общей сложности для него было проанализировано 10 138 контекстов:

Презенс	62,83%
Претерит	26,24%
Перфект	4,78%
Футур I	1,38%
Пассив	0,12%
Конъюнктив I	0,77%
Конъюнктив II	1,58%
Повелительное наклонение	0,29%
Инфинитивные конструкции с <i>zu</i>	1,81%
Модальные глаголы	5,93%
Употребление с отрицанием <i>nicht</i>	7,3%
Употребление с отрицанием <i>kein</i>	0,5%
Употребление с отрицанием <i>nie</i>	1,63%

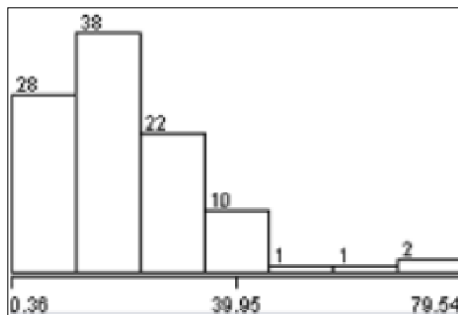
Как упоминалось ранее, мною были собраны данные по указанным выше параметрам для 100 идиом немецкого языка. Статистическая информация была проанализирована с помощью программы [Weka]. Данная программа позволяет, в том числе, по-

смотреть распределение частотностей по различным параметрам (модификациям и временным формам), а также кластеризовать идиомы по схожести их синтаксического поведения. Результаты описаны в [Krotova в печати]. Здесь же посмотрим только на некоторые из них, а именно на данные по употреблению идиом в трех временных формах — презенсе, претерите, перфекте:

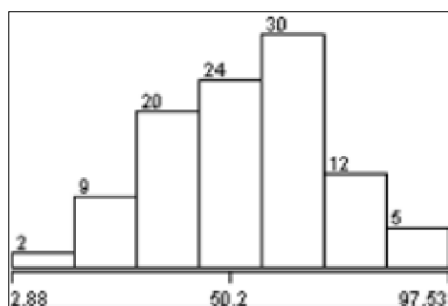
Perfekt



Präteritum



Präsens



Если сопоставить табличные данные по глагольному компоненту *gehen* с частотностями, характерными для идиом, можно увидеть, что ощутимой разницы нет. Глагол встречается в презенсе в 63% случаев, примерно эта частотность характерна и для боль-

шинства идиом. То же самое касается претерита (27%) и перфекта (5%). Частотность употребления в перфекте на удивление низкая. Посмотрим на 100 первых аннотированных примеров с глаголом *gehen* и выясним, чем объясняются такие цифры.

Среди первых 100 примеров в 24% случаев глагол *gehen* встречается в конструкции *es geht um* ‘речь идет (о ком-л., чём-л.)’ (15 из 24 случаев — презенс, 7 — претерит). В 25% случаях встретился другой глагол с отделяемой приставкой (*ausgehen, zurückgehen* и т.п.). Пока такие случаи отделить не получается, т.к. многие отделяемые приставки могут совпадать по форме с предлогами. Таких ложных контекстов меньше для словосочетаний, т.к. второе слово накладывает свои ограничения на семантику глагола. В остальном встречаются как неидиоматические употребления (*jmd. ging ins Internet* ‘кто-л. зашел в Интернет’), так и идиомы (*von Bord gehen* ‘уволиться, уйти в отставку’) и коллокации (*in Erfüllung gehen* ‘исполниться’). Употребление в перфекте относительно редко, в основном встречается презенс и реже претерит. В качестве глагола движения в чистом виде *gehen* употребляется в корпусе не так часто. Возможно, схожесть статистики для глагола со статистикой для идиом объясняется как раз тем, что в корпусе *gehen* в основном употребляется в составе устойчивых словосочетаний.

На редкое употребление глаголов с широкой семантикой в своем первом значении указывается, в частности, в статье Дж. Синклэра [Sinclair 2004]. Так называемые *common verbs* (основные глаголы) часто употребляются в корпусе не в своем прямом значении [Sinclair 2004: 42]. Например, глаголу *to take* носителями языка приписывается значение *to reach for and hold* ‘протянуть руку к какому-либо объекту и взять его’. Тем не менее, анализ корпуса показывает, что в таком значении глагол употребляется относительно редко. Связано это с тем, что основные глаголы могут выступать в качестве вспомогательных глаголов, в качестве *phrasal verbs* (фразовых глаголов), а также часто используются в коллокациях и иных идиоматических сочетаниях, например, *to take time* ‘требовать времени’.

Все проанализированные идиомы с глагольным компонентом *gehen* были разбиты на три группы. Основное сравнение идет по употреблению в презенсе и претерите, т.к. по многим другим параметрам результаты слишком малы и могут быть статистической погрешностью. Идиомы первой группы по сравнению с *gehen* употребляются реже в презенсе и чаще или сопоставимо в претерите. Идиомы второй группы встречаются реже и в презенсе, и в претерите. Идиомы третьей группы чаще употребляются в презенсе и реже или сопоставимо в претерите. Рядом с идиомой дается описание синтаксического поведения ее парафразы, если таковой анализировался. Парафраз сравнивается также с *gehen* ради однообразия. Понятно, что речь идет о разных глаголах и их поведение будет отличаться. Также в конце каждой группы приводится список неидиоматических выражений с *gehen*, которые синтаксически сходно себя ведут.

I. В первую группу вошли те идиомы, которые по сравнению с глаголом *gehen* реже встречаются в презенсе, но чаще или сопоставимо в претерите. Сравним их синтаксическое поведение между собой, а также с их парафразами.

Идиома *bergauf gehen* 'идти в гору' чаще встречается в претерите по сравнению с глаголом *gehen* (на 10%³), несколько реже в презенсе (на 5%) и перфекте (на 2%), также она употребляется реже с отрицанием (на 3,5%).

Идиома *abwärts gehen* 'идти под откос' чаще встречается в претерите (на 22%), реже в презенсе (на 15%) и перфекте (на 2%), реже с отрицанием (на 3,5%).

Идиома *bergab gehen* 'идти под откос' чаще в претерите (на 22%), реже в презенсе (на 17%) и перфекте (на 3%), реже с отрицанием (на 5%).

При этом синонимичная идиоме *bergab gehen* идиома *aufwärts gehen* 'идти в гору' ведет себя синтаксически иначе: несколько чаще глагола *gehen* эта идиома употребляется в презенсе (на 4%), сопоставимо в претерите, чаще с модальными глаголами (на 4%).

3 Здесь и далее приводятся округленные числа.

Среди этих 4 идиом наблюдается следующая тенденция: имеющие положительную семантику *bergauf gehen* и *aufwärts gehen* реже употребляются в претерите и чаще в презенсе по сравнению с имеющими негативную семантику *bergab gehen* и *abwärts gehen*.

Сравним синтаксическое поведение этих идиом с их парафразами *sich verbessern* ‘улучшаться’ и *sich verschlechtern* ‘ухудшаться’.

Глагол *sich verbessern* намного реже глагола *gehen* встречается в презенсе (на 48,5%), примерно также часто в претерите (чаще на 3%), намного чаще в перфекте (на 39%), чаще с модальными глаголами (на 4%), немного реже с отрицанием.

Глагол *sich verschlechtern* также намного реже употребляется в презенсе (на 53,5%), реже в претерите (на 6%), намного чаще в перфекте (на 54%), примерно также часто с модальными глаголами и немного реже с отрицанием.

Получается, что парафразы ведут себя по-другому: они намного чаще рассматриваемых идиом употребляются в перфекте, чем в претерите, и редко встречаются в презенсе.

Рассмотрим другие идиомы, попавшие в эту группу.

Идиома *durch den Kopf gehen* ‘(у кого-л.) в голове пронеслась [промелькнула] мысль (о чём-л.); (кому-л.) (в голову) пришла мысль (о чём-л.); (что-л.) вертится в голове (у кого-л.)’ реже употребляется в презенсе (на 9%), немного чаще в претерите (на 3%), чаще в перфекте (на 5%), реже с отрицанием (на 6%).

Сравним с глаголами, способными выступать парафразами (при изменении диатезы): *sich überlegen* и *nachdenken* ‘раздумывать, размышлять’.

Глагол *sich überlegen* реже употребляется в презенсе (на 28%), реже в претерите (на 7%), чаще в перфекте (на 23%), чаще с инфинитивом с *zu* (на 4%) и чаще с модальными глаголами (на 24%).

Глагол *nachdenken* также реже употребляется в презенсе (на 13%), реже в претерите (на 16%), чаще в перфекте (на 14%), чаще с инфинитивом с *zu* (на 10%) и чаще с модальными глаголами (на 12%).

Получается, что парафразы ведут себя схоже, но от идиомы отличаются. Парафразы чаще встречаются в инфинитиве с *zu* и с модальными глаголами. Для идиомы это нехарактерно.

В первую группу попала еще одна идиома — *in die Hose gehen* ‘провалиться, закончиться неудачей [крахом, фиаско]’. Она встречается реже в презенсе (на 24%), чаще в претерите (на 9%), чаще в перфекте (на 15%), чаще с модальными глаголами (на 7%).

Сравним ее с парафразом *scheitern* ‘провалиться, потерпеть фиаско, рухнуть (о планах)’. Глагол реже встречается в презенсе (на 34%), чаще в претерите (на 7%), чаще в перфекте (на 22%).

В этом случае парафраз ведет себя схоже с идиомой, хотя числа несколько отличаются.

Также анализировалось синтаксическое поведение неидиоматических словосочетаний. В них входят не только свободные словосочетания, но и частотные коллокации с *gehen*.

Среди проанализированных 12 сочетаний пять попали в первую группу, т.к. реже встречаются в презенсе и сопоставимо часто в претерите:

zu Ende gehen ‘закончиться, подойти к концу’ реже в презенсе (на 17%), немного чаще в претерите (на 3%), чаще в перфекте (на 12%), реже с отрицанием (на 5%);

zur Schule gehen ‘идти в школу’ реже в презенсе (на 11%), примерно также часто в претерите (реже на 1%), чаще в перфекте (на 5%), чаще в инфинитивных конструкциях (на 3%), чаще с модальными глаголами (на 8%);

zu Fuß gehen ‘идти пешком’ реже в презенсе (на 8%), примерно также часто в претерите (чаще на 1%), чаще в инфинитивных конструкциях (на 5%), чаще с модальными глаголами (на 16%);

nach Hause gehen ‘идти домой’ реже в презенсе (на 13%), сопоставимо в претерите (чаще на 3%), чаще в инфинитивных конструкциях (на 4%), чаще с модальными глаголами (на 14%);

ins Kloster gehen ‘уйти в монастырь’ реже в презенсе (на 23%), сопоставимо в претерите (чаще на 2%), намного чаще в инфинитивных конструкциях (на 15%), чаще с модальными глаголами (на 3%).

Для первого словосочетания, *zu Ende gehen* ‘закончиться, подойти к концу’, видится естественным редкое употребление в презенсе и частое в перфекте, т.к. речь идет о завершении некоего процесса. Более частое употребление фраз *zur Schule gehen* ‘идти в школу’, *zu Fuß gehen* ‘идти пешком’ и *nach Hause gehen* ‘идти домой’ с модальными глаголами, возможно, частично объясняется особенностями корпуса: в основном корпус содержит публицистику, и можно предположить, что если некая повседневная активность упоминается в новостном тексте, то часто она будет иметь модальность (например, *пришлось* идти пешком, т.к. сломался автобус). Фраза *ins Kloster gehen* ‘уйти в монастырь’ значительно чаще других используется в инфинитивных конструкциях, что также может объясняться узусом (упоминается не просто уход в монастырь, а принятие соответствующего решения, вынужденность и т.п.).

II. Во вторую группу вошли те идиомы, которые по сравнению с глаголом *gehen* реже встречаются и в презенсе, и в претерите. Сравним их синтаксическое поведение между собой, а также с их парафразами.

Рассмотрим идиому *auf den Grund gehen* ‘разобраться (в чём-л.); вникнуть в суть (чего-л.)’. По сравнению с глаголом *gehen* она реже встречается в презенсе (на 18%), реже в претерите (на 14%), сопоставимо в перфекте, чаще с инфинитивом с *zu* (на 27%), чаще с модальными глаголами (на 11%), реже с отрицанием (на 5%).

Сравним с парафразом *eine [die] Ursache finden* ‘найти причину’. Он реже встречается в презенсе (на 31%), реже в претерите (на 7%), чаще в перфекте (на 21%), употребляется в пассиве (15%), чаще с инфинитивом с *zu* (на 16%), чаще с отрицанием (на 19%). Таки образом, и парафраз, и идиома чаще встречаются с инфинитивом с *zu*, но при этом идиома чаще встречается с модальными глаголами, а парафраз — в перфекте, а также с отрицанием и в пассиве.

Идиома *durch die Lappen gehen* ‘(кто-л.) убежал (от кого-л.), (кто-л.) улизнул (от кого-л.), (кто-л.) удрал (от кого-л.); (что-л.) ускользнуло (от кого-л.); (кто-л.) теряет (что-л.)’ встречается реже в презенсе (на 18%), реже в претерите (на 7%), чаще в перфекте (на 17%), чаще в конъюнктиве II (на 5%).

Сравним с парафразом *entkommen* ‘убежать, ускользнуть’. Он также реже употребляется в презенсе (на 45%), но при этом чаще встречается в претерите (на 8%), чаще в перфекте (на 34%), чаще с модальными глаголами (на 16%). Таким образом, синтаксическое поведение идиомы значительно отличается от поведения парафразы.

Идиома *auf den Leim gehen* ‘попасться на удочку (кому-л.)’ встречается значительно реже в презенсе (на 27%), реже в претерите (на 11%), намного чаще в перфекте (на 23%), чаще с инфинитивом с *zu* (на 8%) и с отрицанием (на 11%). Для нее парафразы не рассматривались, т.к. они были в пассиве.

Идиома *hausieren gehen* ‘навязываться всем (с разговорами, идеями); раструбить, раззвонить (что-л.), кричать (о чём-л.) на каждом углу’ употребляется реже в презенсе (на 2%), реже в претерите (на 6%), чаще с инфинитивом (на 5%) и чаще с модальными глаголами (на 5%), чаще с отрицанием (на 8%).

Сравним с парафразом *überall erzählen* ‘везде рассказывать’. Он употребляется значительно реже в презенсе (на 20%), реже в претерите (на 9%), чаще в перфекте (на 29%), часто в пассиве (17%), чаще в инфинитиве с *zu* (на 2%), реже с отрицанием (на 3%). За счет употребления в пассиве и частого употребления в перфекте синтаксическое поведение идиомы и парафразы расходятся.

Идиома *durchs Feuer gehen* ‘пройти сквозь огонь и воду (ради кого-л.)’ сопоставимо часто употребляется в презенсе, реже встречается в претерите (на 17%), чаще в перфекте (на 5%), чаще в конъюнктиве II (на 7%), чаще в инфинитив (на 2,5%).

Сравним с парафразом *Schlimmes erleben* ‘пережить ужасное’. Он реже встречается в презенсе (на 56%), реже в претерите (на 22%), чаще в перфекте (на 51%), реже с от-

рицанием (на 5%). Синтаксическое поведение парафразы значительно отличается от идиомы.

Также в эту группу попала идиома *seinen Weg gehen* ‘идти своим (собственным) путём; найти свой собственный путь (в жизни)’. Она реже встречается в презенсе (на 8%) и реже в претерите (на 14%), чаще в перфекте (на 6%), чаще с инфинитивом с *zu* (на 12%), чаще с модальными глаголами (на 10%).

Сравним с парафразом *vorankommen* ‘продвигаться вперёд; преуспеть в жизни’. Он также реже употребляется в презенсе (на 20%), реже в претерите (на 21%), чаще в перфекте (на 3%), чаще с инфинитивом с *zu* (на 28%), чаще с модальными глаголами (на 8%). Синтаксическое поведение парафразы в целом схоже с идиомой, хотя парафраз значительно чаще встречается с инфинитивом с *zu*.

В эту группу попали следующие неидиоматические словосочетания:

ins Ausland gehen ‘отправиться за границу’ — реже в презенсе (на 12%) и реже в претерите (на 9%), чаще с инфинитивом (на 14%), и модальными глаголами (на 6%);

zum Arzt gehen ‘пойти к врачу’ — реже в презенсе (на 10%) и реже в претерите (на 15%), чаще с инфинитивом (на 11%), и модальными глаголами (на 19%), чаще с отрицанием (на 8%);

schlafen gehen ‘идти спать’ встречается реже в презенсе (на 5%), реже в претерите (на 11%), чаще в перфекте (на 5%), чаще в инфинитиве (на 4%), чаще с модальными глаголами (на 8%).

Как и у неидиоматических выражений из группы 1, данные фразы чаще употребляются в инфинитиве и с модальными глаголами, что, как указано выше, может объясняться публицистическим характером корпуса.

III. В третью группу вошли те идиомы, которые по сравнению с глаголом *gehen* чаще встречаются в презенсе и реже или сопоставимо в претерите.

В эту группу вошли несколько идиом со схожей семантикой:

Идиома *auf den Keks gehen* ‘действовать на нервы (кому-л.); доставать (кого-л.)’ чаще встречается в презенсе (на 17%), реже в претерите (на 14%).

Идиома *auf den Geist gehen* ‘действовать на нервы (кому-л.)’ также чаще в презенсе (на 14%), реже в претерите (на 13%).

Идиома *auf die Nerven gehen* ‘действовать на нервы (кому-л. чем-л.)’ также чаще в презенсе (на 8%), реже в претерите (на 9%).

Сравним с парафразом *nervös machen* ‘заставлять нервничать’. Он встречается чаще в презенсе (на 4%), реже в претерите (на 14%), чаще в перфекте (на 4%), чаще в инфинитиве (на 3%), чаще с отрицанием (на 9%). В целом синтаксическое поведение парафраза сходно с поведением идиом.

Рассмотрим идиому *an die Nieren gehen* ‘брать [хватать] за живое (кого-л.); сильно подействовать (на кого-л.)’. Она чаще встречается в презенсе (на 6%), несколько реже в претерите (на 8%), чаще в перфекте (на 5%), реже с отрицанием (на 3%).

Сравним с двумя парафразами. Фраза *schwer zu schaffen machen* ‘создавать кому-л. трудности’ сопоставимо часто употребляется в презенсе, реже в претерите (на 3%), чаще в перфекте (на 3%), часто в пассиве (15%), реже с отрицанием (на 5%). За исключением употребления в пассиве, невозможного для данной идиомы, синтаксическое поведение фразы схоже с поведением идиомы.

Сравним с глаголом *belasten* ‘обременять, наносить вред’. Он реже встречается в презенсе (на 27%), реже в претерите (на 9%), чаще в перфекте (на 30%), часто в пассиве (15%), чаще в инфинитиве с *zu* (на 3%), чаще с отрицанием (на 4%). Видно, что его синтаксическое поведение отличается от идиомы.

К этой же группе относится упомянутая ранее *aufwärts gehen* ‘идти в гору’: чаще глагола *gehen* употребляется в презенсе (на 4%) и чаще с модальными глаголами (на 4%).

Рассмотрим идиому *(nicht) aus dem Sinn gehen* ‘не выходить из головы (у кого-л.)’. Она чаще употребляется в презенсе (на 6%), немного реже в претерите (на 3%), реже с

модальными глаголами (на 4%). Т.к. идиома является эксплицитно негативной, употребляется она в основном с отрицанием (чаще на 84%).

В качестве парафразы к идиоме можно рассматривать фразу *ständig daran denken müssen* 'постоянно думать о чем-л.'. Она не анализировалась, т.к. редко встречается в корпусе.

Можно сравнить синтаксическое поведение этой идиомы с синонимичной (*nicht aus dem Kopf gehen* 'не выходить [идти] из головы (у кого-л.); крутиться в голове (у кого-л.)'). Она сопоставимо часто употребляется в презенсе и претерите, несколько чаще в перфекте (на 3%), реже с модальными глаголами (на 2,5%), в основном с отрицанием (чаще на 85%). В целом поведение этих синонимичных идиом схоже.

Из неидиоматических выражений в группу попали фразы:

vor sich gehen 'происходить'. Она чаще встречается в презенсе (на 9%) и реже в претерите (на 6%), а также реже с отрицанием (на 5%);

einkaufen gehen 'идти за покупками' чаще встречается в презенсе (на 3%), реже в претерите (на 13%), чаще в инфинитиве (на 8%), чаще с модальными глаголами (на 12%);

ins Schwimmbad gehen 'идти в бассейн' сопоставимо в презенсе (чаще на 1%), реже в претерите (на 10%), сопоставимо в перфекте (чаще на 1%), чаще в инфинитиве с *zu* (на 7%), чаще с модальными глаголами (на 6%).

Наиболее близкой к исходному глаголу оказалась фраза *in [durch] den / im Wald gehen* 'ходить по лесу'. Она встречается сопоставимо в презенсе, сопоставимо в претерите (реже на 2%), сопоставимо в перфекте (чаще на 1%), чаще в инфинитиве с *zu* (на 2%), сопоставимо с модальными глаголами.

Глагол *gehen* накладывает на идиомы определенные синтаксические ограничения. Так, он практически не употребляется в пассиве, т.к. не является переходным.

Большинство рассматриваемых идиом и неидиоматических выражений также используется в пассиве меньше чем в 1% случаев, за исключением *seinen Weg gehen* — 1,15%, *auf den Grund gehen* — 2,69%.

Среди проанализированных парафразов большинство также не являются переходными глаголами (или не включают таковые в свой состав) или являются возвратными. Они практически не встречаются в пассиве (от 0 до 1,6%): *vorankommen*, *nachdenken*, *lästig werden*, *scheitern*, *sich verschlechtern*, *sich verbessern*, *sich überlegen*, *entkommen*. Есть несколько конструкций, включающих переходный глагол, однако они также практически не употребляются в пассиве (от 0,15 до 1,03%): *nervös machen*, *schwer zu schaffen machen*, *Schlimmes erleben*. Видимо, в этих случаях ограничения на пассивизацию накладывает семантика фразы.

В трех случаях пассив встречается в парафразах относительно часто (от 14,5 до 16,5%): *Ursache finden*, *belasten*, *überall erzählen*.

Как показал статистический анализ большого количества корпусных данных для рассматриваемых идиом, их синтаксическое поведение в целом отличается как от поведения глагольного компонента, так и от их парафразов. Синонимичные идиомы с одинаковым глагольным компонентом могут употребляться схожим образом, но в целом сходным синтаксическим поведением могут обладать разные по значению идиомы. Для некоторых идиом могут наблюдаться корреляции между их синтаксическим поведением и поведением парафразы (*in die Hose gehen* — *scheitern*), но на все идиомы это не распространяется. Между идиомами и парафразами остаются семантические различия и различия в узусе, которые не дают вести им себя синтаксически одинаково. Глагольный компонент идиомы также может накладывать ограничения, например, невозможность пассивизации (*auf den Grund gehen* — *Ursache finden*).

Рассмотренные неидиоматические словосочетания (*einkaufen gehen*, *ins Ausland gehen*, *ins Schwimmbad gehen*) синтаксически ведут себя схоже с идиомами. Их можно поделить на те же группы, и отличия в частотности отдельных параметров от частотностей *gehen* в них также могут быть существенны. Можно было бы предположить, что их поведение будет более однородным. То, что их частотности не совпали с *gehen*, неудивительно. Как было показано, в корпусе *gehen* часто встречается в устойчивых словосочета-

ниях. Однако можно было бы ожидать, что между собой у них будет меньше отличий. Возникает вопрос, в какой мере можно говорить о том, что синтаксическое поведение идиом в принципе отлично от поведения частотных словосочетаний. Чтобы подтвердить или опровергнуть эту гипотезу, требуется сопоставимый анализ для идиом и неидиоматичных выражений с другими глагольными компонентами. Также изучение поведения свободных словосочетаний сделает более наглядным, в каких случаях действительно можно говорить о дефектности парадигмы идиомы.

В заключение хотелось бы указать на лексикографическую и методическую ценность подобных статистических сведений об идиомах. Понимание того, какие модификации для идиомы частотны и употребление в какой временной форме для нее характерно, может помочь при составлении комментариев к словарным статьям и при выборе иллюстративного материала. С методической точки зрения важно знать прототипические контексты употребления идиомы. Например, большинство рассмотренных идиом с *gehen* редко употребляются в перфекте, хотя такое употребление вполне возможно. Если изучающий язык не будет владеть этой информацией, он с высокой вероятностью допустит ошибку при употреблении этой идиомы: при изучении немецкого языка претериту как письменной форме уделяется меньшее внимание, чем перфекту как разговорной форме.

Источники

DeReKo: Deutsches Referenzkorpus // <https://cosmas2.ids-mannheim.de/cosmas2-web>.

Программное обеспечение

GATE: General Architecture for Text Engineering // <https://gate.ac.uk>.

Weka: Waikato Environment for Knowledge Analysis // <https://www.cs.waikato.ac.nz/~ml/weka>.

Литература

Баранов А. Н., Добровольский Д. О. Аспекты теории фразеологии. М.: Знак, 2008.

Иомдин Л. Л. Большие проблемы малого синтаксиса // Труды международной конференции по компьютерной лингвистике и интеллектуальным технологиям Диалог'2003. Протвино, 2003. С. 216—222.

Dobrovol'skij D. Idiom-Modifikationen aus kognitiver Perspektive // Kamper, H., Eichinger, L.M. (Hrsg.) Sprache — Kognition — Kultur. Sprache zwischen mentaler Struktur und kultureller Pragung. Berlin / New York: de Gruyter, 2008. S. 302—322.

Krotova E. Syntaktische Modifikationen der Phraseme im Deutschen // Carmen Mella-do/Herbert Holzinger/Nely Iglesias/Ana Mansilla (eds.): Produktive Muster in der Phraseologie. Reihe Studia Phraseologica et Paroemiologica. Verlag Dr. Kovač (Hamburg) (*в печати*).

Krotova E. Computer-aided Analysis of Idiom Modifications in German // Proceedings of the XVIII EURALEX International Congress: Lexicography in Global Contexts. Edited by: Jaka Čibej, Vojko Gorjanc, Iztok Kosem and Simon Krek. Ljubljana, 2018, pp. 523—531// <https://e-knjige.ff.uni-lj.si/znanstvena-zalozba/catalog/book/118>.

Sinclair J. Intuition and annotation — the discussion continues // Advances in Corpus Linguistics / Papers from the 23rd International Conference on English Language Research on Computerized Corpora (ICAME 23) Göteborg 22-26 May 2002. Amsterdam/New York, NY, 2004, IX. P. 39—59.

**Syntactic behavior of idioms and non-idiomatic expressions:
a corpus study (based on the material of the German language)**

**E. B. Krotova (Institute of Linguistics,
Russian Academy of Sciences)**

This article compares the syntactic behavior of German idioms with the verb component *gehen* with the syntactic behavior of their verb component and their paraphrases. The syntactic behavior of a number of non-idiomatic expressions containing the same verb component is also analyzed. The article examines how significant the differences are between the syntactic behavior of idioms and non-idiomatic expressions.

Keywords: idioms, syntax, corpus linguistics.