

КЕМПИ
Арина Анатольевна

бакалавриат, Петрозаводский государственный
университет (Петрозаводск, Россия),
rina.kempi@mail.ru

КОЛИЧЕСТВЕННАЯ И КАЧЕСТВЕННАЯ ДИНАМИКА ОРФОГРАФИЧЕСКИХ ОШИБОК В НЕМЕЦКОЯЗЫЧНЫХ ТЕКСТАХ НА РАЗНЫХ ЭТАПАХ ОБУЧЕНИЯ

Научный руководитель:
Котюрова Ирина Аврамовна

Рецензент:
Токко Наталия Ивановна
Статья поступила: 12.02.2024;
Принята к публикации: 28.06.2024;
Размещена в сети: 30.06.2024.

Аннотация. Орфографические ошибки – распространенное явление в процессе изучения иностранных языков. В статье представлен анализ и выявлена количественная и качественная динамика орфографических ошибок в текстах на немецком языке, взятых из ученического корпуса ПАКТ у студентов на разных этапах обучения – начальном, среднем и завершающем. Делается вывод о том, что с начального до завершающего этапа обучения количество орфографических ошибок сокращается. Основное сокращение приходится на начальный этап обучения при дальнейшем снижении динамики.

Ключевые слова: ученический корпус, корпус студенческих текстов, ПАКТ, образовательная аналитика, корпусная лингвистика, орфографические ошибки

Для цитирования: Кемпи А. А. Количественная и качественная динамика орфографических ошибок в немецкоязычных текстах на разных этапах обучения // StudArctic Forum. 2024. Т. 9, № 2. С. 111–118.

Введение. Возникновение ошибок у студентов, изучающих иностранные языки, – явление неизбежное, как при любом обучении. Исследование этих ошибок – предмет многочисленных работ как лингвистов, так и педагогов [Иванова: 250], [Соснина: 40], [Котюрова, Токко: 68], [Дрога: 103], [Кондрашова: 28]. В последнее время на помощь им приходит корпусная лингвистика – относительно новое научное направление, которое является мощным инструментом для языковых исследований, в том числе и лингводидактических [Майорова: 42]. В связи с этим не вызывает сомнений актуальность данного исследования, выполненного с применением корпусных технологий.

Одним из ключевых преимуществ корпусной лингвистики в образовании является возможность анализа реальных языковых данных. Ученические языковые корпуса, как Петрозаводский аннотированный корпус текстов (ПАКТ), в котором содержатся тексты на двух языках – немецком и французском, в которых проведена разметка ошибок согласно классификации из 90 пунктов [Котюрова, Щеголева: 326], позволяют преподавателям эффективнее оценивать результаты обучения, выявлять слабые стороны студентов, отслеживать динамику сокращения количества ошибок к последнему курсу обучения, а учащиеся имеют возможность просматривать ошибки в своих работах для их предотвращения в дальнейшем. Предметом для анализа в данной статье стали орфографические ошибки в текстах студентов.

Целью данного исследования является выявление динамики орфографических ошибок на разных этапах обучения. Для этого было необходимо решить две основных задачи: провести анализ количественной статистики таких ошибок в лонгитюдном срезе, а также интерпретировать эти ошибки в содержательном аспекте, то есть провести анализ качественных изменений. Таким образом, предметом исследования стали орфографические ошибки в текстах студентов первого, третьего и пятого года обучения.

Для данного исследования из корпуса студенческих текстов ПАКТ были отобраны все примеры – предложения, которые содержали в себе орфографические ошибки, допущенные в работах студентов 1, 3 и 5 курсов. Все орфографические ошибки, встречающиеся в данных примерах, были сгруппированы по нескольким признакам, и на основе этих признаков была создана классификация орфографических ошибок. Опираясь на эту классификацию, мы провели статистический анализ, наглядно показывающий количественную и качественную динамику орфографических ошибок у студентов с начала и до завершения обучения. По состоянию на ноябрь 2023 г. в ПАКТ аннотированы немецкоязычные тексты 1, 3 и 5 курсов объемом 24569, 30856 и 12136 токенов соответственно.

Результаты исследования. Примеров, содержащих в себе орфографические ошибки, в текстах студентов первого курса оказалось 315, в текстах у третьего курса – 171, в текстах студентов на пятом курсе обучения – 63. Для выявления количественной динамики в орфографических ошибках студентов следовало посчитать и сравнить количество данных ошибок в текстах на каждые 100 токенов. Таким образом, удалось выявить (см. табл. 4), что на каждые 100 токенов на первом курсе обучения приходится примерно 1,28 ошибки, на третьем курсе – 0,55 ошибки, и на пятом – 0,52 ошибки.

Содержательный анализ всех этих примеров позволил распределить их на следующие группы в зависимости от типа ошибки:

- неверное написание слова с заглавной / строчной буквы;
- лишняя буква;
- отсутствие нужной буквы;
- замена одной буквы на другую;
- англоязычное написание слова вместо немецкоязычного;
- перестановка букв внутри слова;
- раздельное / слитное написание слова;
- ошибки в написании умлаута;
- ошибки в написании имен собственных;
- ошибки в написании ß;
- ошибки, связанные с наличием или отсутствием знака при дефисном написании.

Проанализировав частотность типов встречающихся орфографических ошибок в текстах первокурсников, мы можем сказать о том, что наиболее частотными являются ошибки, связанные с написанием слов со строчной буквы вместо прописной. Например: *Die Natur und **umwelt** in Schweden sind sauber und hübsch*. Примеров, содержащих в себе такой тип ошибки, насчитывается 63, что составляет 20 % от общего числа орфографических ошибок на 1 курсе.

Второй по частотности у начинающих изучать язык студентов является орфографическая ошибка, когда одна буква заменяется на другую: *Jedoch, ich schenke große **Audmerksamkelt** auf Sport noch, und trainiere dreimal pro Woche*. Данный тип ошибки составляет 18 % (57 примеров).

Ошибки, связанные с отсутствием нужной буквы, занимают третье место по частотности и составляют 14,6 %. Например: *Er baut seine Sportlaufbahn, nimmt an **Wettbewrben** teil und natürlich gewinnt*.

Четвертой по распространенности ошибкой в текстах первокурсников является написание слова с прописной буквы вместо строчной (13 %): *Es gibt viele verschiedene Kleidungsstile und **I**ch verstehe nicht immer was in Mode ist.*

На пятом месте – лишняя буква в слове. Например: *Es hat drei Stockwerke: ein Keller, ein **E**rdgeschloss und ersten Stock.* Данный тип ошибки составляет 9,8 % от общего числа орфографических ошибок в текстах студентов, обучающихся на первом курсе.

Далее по количеству следуют ошибки в написании умлаута (6,3 %). Среди 20 примеров встретилось 19, в которых умлаут отсутствует, и лишь один, в котором умлаут написан ошибочно: *Hinzu kommt noch, das Training ist verpflichtend und der Sportler ist immer unter **D**rück: er muss immer gewinnen.*

Небольшая доля ошибок в работах первокурсников связана с неверным написанием имен собственных или названий (5 %): *Der Artikel stammt aus der Zeitschrift «**V**itamin **D**»* (правильное наименование журнала – Vitamin De).

4,1 % от общего числа ошибок, связанных с орфографией, приходится на ошибки в написании ß. Например: *Er macht viel **S**pazz und nimmt nicht so viel Zeit.*

В текстах первокурсников встретилось 10 примеров (3,2 %), связанных с отдельным написанием слова вместо слитного: ***D**esigner **K**leidung gibt ein selbstbewussteres Gefühl und das ist sehr wichtig, weil es attraktiv macht.*

Далее по убыванию идут ошибки, связанные с перестановкой букв внутри слова (2,2 %): *Die Sonne **s**chneit nicht so oft, und ich fühle mich depressiv.*

Англоязычное написание слов вместо немецкоязычного встретилось в 5 примерах (1,6 %): *Das sind meine **H**obbies und mir gefallen sie sehr.*

На предпоследнем месте стоит ошибка – отсутствие знака при дефисном написании (1,3 %). Например: *Der Arzt hat gesagt, dass ich tief **e**in und **a**usatmen soll.*

И, наконец, на последнем месте по частотности совершения находится слитное написание слова там, где нужно писать слово отдельно (0,6 %): *Es hängt alles von meiner Laune ab, **n**aja, ein bisschen vom Wetter.*

Таблица 1

Статистика орфографических ошибок на 1 курсе (начальный этап обучения)

Тип ошибки	Количество примеров	Процент от общего числа орфографических ошибок
Написание слова со строчной буквы	63	20 %
Замена одной буквы на другую	57	18 %
Отсутствие нужной буквы	46	14,6 %
Написание слова с прописной буквы	41	13 %
Лишняя буква	31	9,8 %
Ошибки в написании умлаута	20	6,3 %
Ошибки в написании имен собственных	16	5 %
Ошибки в написании ß	13	4,1 %
Раздельное написание слова	10	3,2 %
Перестановка букв внутри слова	7	2,2 %
Англоязычное написание	5	1,6 %
Знак при дефисном написании	4	1,3 %
Слитное написание слова	2	0,6 %

В таблице 1 отражена сводная картина орфографических ошибок на первом курсе, когда студенты только приступили к изучению немецкого языка с нуля.

Далее перейдем к ошибкам в текстах студентов, обучающихся на третьем курсе, т. е. примерно в середине общего курса изучения немецкого языка, который рассчитан на 5 лет обучения. Примеров, содержащих в себе орфографические ошибки, здесь уже значительно меньше – 171 (0,55 ошибки на каждые 100 токенов). Связано это, безусловно, с тем, что к третьему курсу обучения студенты значительно улучшают свои знания немецкого языка благодаря долгой практике, лучшему изучению правил и др. Тем не менее ошибки в работах третьекурсников все же присутствуют.

Характерно, что на третьем курсе сохраняются те же самые типы ошибок, что и на первом курсе обучения (см. типы в таблице 1), однако распределение частотности этих ошибок меняется.

Наиболее частотной на третьем курсе, как и на первом, является ошибка «написание слова с прописной буквы вместо строчной» (17,5 %): *Außerdem kann **I**ch meine Fassung verlieren, wenn jemand meiner Meinung nicht zustimmt.*

На втором месте – ошибки, связанные с отсутствием нужной буквы/букв в слове (16,9 %), например: *Im Artikel handelt es sich um **E**rrinerungen und Erfahrungen von iranischen Frauen, die heute in Tübingen leben,* и с ошибочным написанием слова со строчной буквы вместо прописной (16,9 %): *Und seit ein paar Wochen habe sie schlaflose **n**ächte.*

Следующей из наиболее частотных ошибок является замена буквы в слове (12,3 %). Например: *Er hat das Buch langweilich gefunden, so er ist **e**ingeschlaben.*

Сравнивая частотность орфографических ошибок на первых трех местах у студентов первого и третьего курсов, можно заметить, что в текстах студентов преобладают практически одни и те же типы ошибок: первокурсники чаще всего пишут слова со строчной буквы вместо прописной, а третьекурсники – наоборот. Также на обоих курсах учащиеся довольно часто допускают опечатки, заменяя или не дописывая нужные буквы.

У третьекурсников также часто возникают ошибки в написании имен собственных и названий, хотя и наблюдается динамика снижения подобных ошибок – 13 % на первом курсе против 9,4 % на третьем: *Die Insel **K**iji ist eine der berühmtesten Wahrzeichen der Republik Karelien.*

Одной из распространенных ошибок является написание лишней буквы в слове, однако и здесь наблюдается снижение с 9,8 % до 8,7 %: *Ende Mai findet ein **R**ock**B**estival «Tunnel» statt.*

Следом идут слитное написание слова вместо отдельного (4,7 %) и ошибки в написании умлаута (4,7 %). Примерами могут послужить предложения: *Deshalb gehen sie dorthin, wo die Menschen zu Hause sind, wo sie Sport treiben und ihre Familien und **F**reundesind,* а также *Diese Gebäude **g**ehort zum Postmoderne.*

Студенты третьего курса также допускают и ошибки в отдельном написании слова в случае, когда требуется писать его слитно (4 %): *Diese Kathedrale ist ein sehr **m**ajestätisch es und bedeutendes Gebäude auf der ganzen Welt.*

Одной из наименее частотных является ошибка, связанная с перестановкой букв в слове (1,7 %), например: *Dies liegt vielleicht daran, dass die vorderen Spitzboegen unvollendet **a**susehen.*

Англоязычное написание слова вместо немецкоязычного (1,2 %) и ошибки, связанные с отсутствием знака при дефисном написании (1,2 %), находятся на предпоследнем месте по частотности совершения. Примеры: *Es sieht aus wie ein **A**pfel**g**arden* и *Psychologen verwenden das **B**ig **F**ive-Modell, um einen Charakter zu beschreiben.*

На последнем месте – ошибки в написании ß. Данный тип составляет 0,6 % и

насчитывает лишь один пример: *Der Mann hat eine ernste Miene, raue Haut, ein schmales Gesicht, einen dichten Bart und eine **Grosse**, gerade Nase.*

Сравнивая таблицу 2, где указана статистика для работ третьего курса с таблицей 1, можно увидеть, что типы ошибок и их очередность по убыванию в целом схожи. Отклонения наблюдаются в написании слова со строчной буквы, на третьем курсе этот тип ошибки стал наиболее частотным. Студенты третьего курса также стали чаще совершать ошибки в слитном написании слова вместо отдельного.

Таблица 2

Статистика орфографических ошибок на 3 курсе (средний этап обучения)

Тип ошибки	Количество примеров	Процент от общего числа орфографических ошибок
Написание слова с прописной буквы	30	17,5 %
Написание слова со строчной буквы	29	16,9 %
Отсутствие нужной буквы	29	16,9 %
Замена одной буквы на другую	21	12,3 %
Ошибки в написании имен собственных	16	9,4 %
Лишняя буква	15	8,7 %
Слитное написание слова	8	4,7 %
Ошибки в написании умлаута	8	4,7 %
Раздельное написание слова	7	4 %
Перестановка букв внутри слова	3	1,7 %
Англоязычное написание	2	1,2 %
Знак при дефисном написании	2	1,2 %
Ошибки в написании ß	1	0,6 %

Для того чтобы отследить динамику орфографических ошибок у студентов на трех ступенях обучения, приведем статистику и с пятого курса. Здесь типов орфографических ошибок выделяется значительно меньше, чем на 1 курсе, и незначительно меньше, чем на 3 (0,52 ошибки на каждые 100 токенов). То есть можно сделать вывод, что усвоение орфографических правил происходит в основном на начальном этапе обучения.

На первом месте, как и в предыдущем периоде, идут два противоположных друг другу типа – написание слова со строчной буквы вместо прописной и наоборот (26,9 %). Примерами из текстов, написанных пятикурсниками, могут послужить предложения *Ich habe viel gesehen **filme**, in denen Bösewichte zum Polizisten kamen «...»* и *Es ist bemerkbar, dass die **Erneubaren** Energien vor 10 Jahren überhaupt nicht genutzt wurden.*

Следом идут ошибки в написании имен собственных (23,8 %), в частности ошибки, связанные с неверным написанием имени героини романа Г. Гессе «Степной волк» Гермины: *Im Gespräch mit Harry sagt **Hermi**na: es ist viel flacher, Harry, wenn du für etwas Gutes und Ideales kämpfst und nun meinst, du müsstest es auch erreichen.*

Третьим по распространенности является отсутствие нужной буквы в слове (11,1 %): *Das Buch hat mir gefallen aber **man**mal meine Gefühle gemischt waren.*

Далее следуют ошибки, связанные с заменой одной буквы на другую (4,7 %): *Einige Leute vertrauen den **Massens**edien, und deswegen hassen einender.*

Примеров на лишнюю букву в слове (3,2 %) встречается всего два, и одним из них

является предложение: *Ich war damals schockiert, obwohl ich selbst, ohne es zu wissen, keine Kosmetika mehr benutzte.*

Тип ошибки, связанный со слитным написанием слова вместо отдельного составляет 1,6 % от общего числа орфографических ошибок: *Aber es war immer nur Sympathie und Freundlichkeit, was erfand, man lud ihn ein, man beschenkte ihn, schrieb ihm nette Briefe, aber nahe an ihn heran kam niemand, Bindung entstand nirgends, sein Leben zu teilen war niemand gewillt und fähig.*

Наравне с предыдущим типом идет ошибка, связанная с отсутствием знака при дефисном написании (1,6 %): *Sie schickte mir Fotos, und ich mochte seine Siedlung in **Frankfurt am Main** sehr.*

Таблица 3

Статистика орфографических ошибок на 5 курсе (завершающий этап обучения)

Тип ошибки	Количество примеров	Процент от общего числа орфографических ошибок
Написание слова со строчной буквы	17	26,9 %
Написание слова с прописной буквы	17	26,9 %
Ошибки в написании имен собственных	15	23,8 %
Отсутствие нужной буквы	7	11,1 %
Замена одной буквы на другую	3	4,7 %
Лишняя буква	2	3,2 %
Слитное написание слова	1	1,6 %
Знак при дефисном написании	1	1,6 %

Анализируя статистику орфографических ошибок на завершающем этапе обучения, можно сказать, что к пятому курсу у студентов совсем уходят такие ошибки, как англоязычное написание вместо немецкоязычного, отдельное написание слова вместо слитного, ошибки в написании умлаута, ошибки в написании ß и перестановка букв внутри слова.

Таблица 4

Количественная динамика орфографических ошибок на разных этапах обучения

Курс	1 курс	3 курс	5 курс
Общее количество орфографических ошибок	315	171	63
Количество токенов в проверенных текстах	24569	30856	12136
Количество ошибок на каждые 100 токенов	1,28	0,55	0,52

Выводы. Анализ примеров, содержащих в себе ошибки на орфографию на трех курсах – трех этапах обучения, а также данные в таблице 4, позволяют сделать вывод, что с первого года обучения (начального этапа) к пятому (завершающему этапу) наблюдается сокращение орфографических ошибок в студенческих текстах, при этом стремительное уменьшение количества орфографических ошибок в текстах учащихся происходит с первого

по третий курс, а разница в частотности орфографических ошибок на 3 и 5 курсе – незначительная. Связано это, безусловно, с тем, что на первом курсе немецкий язык студенты изучают с нуля и не знают большинства нюансов немецкой орфографии, а уже к третьему курсу благодаря углублению в немецкий язык, тщательному изучению правил и долгой практике показывают гораздо лучшие результаты. Тем не менее следует отметить, что на среднем этапе обучения у студентов присутствуют абсолютно те же самые типы орфографических ошибок, что и на начальном этапе.

Следует отметить, что количественная динамика от третьего курса к пятому меняется незначительно, но также в лучшую сторону. Тем не менее к завершающему этапу обучения классификация орфографических ошибок в текстах студентов становится уже. К пятому курсу учащиеся перестают совершать ошибки, вызванные интерференцией с английским языком (англоязычное написание вместо немецкоязычного), а также ошибки в написании ß или умлаутов, что свидетельствует о положительной качественной динамике.

Список литературы

Дрога М.А. Типичные орфографические ошибки студентов-иностранцев и пути их преодоления // Проблемы преподавания филологических дисциплин иностранным учащимся. 2014. С. 103-107.

Иванова Н.К. Орфографические ошибки в английском языке: опыт классификации на основе интернет-материалов // Вестник КГУ. 2012. № 1. С. 250-253.

Кондрашова Н.В. Прогнозирование и исправление студенческих ошибок при обучении иностранным языкам // Научный диалог. 2015. № 7(43). С. 27-47.

Котюрова И.А. Анализ типичных ошибок русскоязычных студентов в текстах на немецком языке с помощью ученического корпуса ПАКТ / И.А. Котюрова, Н.И. Токко // Непрерывное образование: XXI век. 2022. № 4(40). С. 66-76.

Котюрова И.А. Корпус студенческих текстов на немецком языке как источник данных для образования и науки / И.А. Котюрова, Л.В. Щеголева // Вопросы образования. 2022. № 4. С. 322-349.

Майорова А.Д. Корпусная лингвистика: исторический и лингводидактический аспекты // Международный научно-исследовательский журнал. 2017. № 5-2(59). С. 42-46.

Соснина Е.П. К анализу лингвистических ошибок учащихся на базе учебного корпуса текстов // Известия Самарского научного центра Российской академии наук. Социальные, гуманитарные, медико-биологические науки. 2017. № 5. С. 39-44.

Arina A. KEMPI

bachelor's degree, Petrozavodsk State University
(Petrozavodsk, Russia),
rina.kempi@mail.ru

QUANTITATIVE AND QUALITATIVE DYNAMICS OF SPELLING ERRORS IN GERMAN TEXTS ACROSS THE STAGES OF LEARNING

Scientific adviser:

Irina A. Kotiurova

Reviewer:

Natalia Tokko

Paper submitted on: 02/12/2024;

Accepted on: 06/28/2024;

Published online on: 06/30/2024.

Abstract. Spelling errors are a frequent occurrence during the process of learning foreign languages. This study analyzes and discusses the quantitative and qualitative changes in spelling errors in German texts sourced from the PACT learner corpus of texts written by students across their initial, intermediate, and final learning stages. The findings indicate a decrease in spelling errors from the initial to final stages of learning, with the most significant reduction observed in the initial stage, followed by a gradual decrease in errors.

Keywords: learner corpus, corpus of student texts, PACT, educational analytics, corpus linguistics, spelling errors

For citation: Kempf, A. A. Quantitative and Qualitative Dynamics of Spelling Errors in German Texts Across the Stages of Learning. *StudArctic Forum*. 2024, 9 (2): 111–118.

References

Droga M.A. Typical spelling mistakes of foreign students and ways to overcome them. *Problems of teaching philological disciplines to foreign students*, 2014, pp. 103-107. (In Russ.)

Ivanova N.K. Spelling mistakes in the English language: experience of classifying on the basis of the Internet-sources. *Vestnik of Kostroma State University*, 2012, No. 1, pp. 250-253. (In Russ.)

Kondrashova N.V. Prediction and correction of students' mistakes when teaching of foreign languages. *Scientific Dialogue*, 2015, No. 7(43), pp. 27-47. (In Russ.)

Kotiurova I.A., Tokko N.I. Typical errors analysis by Russian-speaking students in German texts in the learner corpus PACT. *Lifelong Education: The 21st Century*, 2022, No. 4(40), pp. 66-76. (In Russ.)

Kotiurova I.A., Shchegoleva L.V. Learner corpus in German as a data source for education and science. *Educational Studies*, 2022, No. 4, pp. 322-349. (In Russ.)

Maiorova A.D. Corpus linguistics: historical and linguistic diagnostic aspects. *International Research Journal*, 2017, No. 5-2(59), pp. 42-46. (In Russ.)

Sosnina E.P. Research of pupils' linguistic errors on the base of learner corpus. *Izvestiya of the Samara Russian Academy of Sciences Scientific Center. Social, Humanitarian, Medicobiological Sciences*, 2017, No. 5, pp. 39-44. (In Russ.)