

What you apply is not what you learn! Examining students' strategies in German capitalization tasks

Nathalie Rzepka
Hochschule für Technik und Wirtschaft
Berlin
nathalie.rzepka@htw-berlin.de

Hans-Georg Müller
Universität Potsdam
Potsdam
hgmuelle@uni-potsdam.de

Katharina Simbeck
Hochschule für Technik und Wirtschaft
Berlin
katharina.simbeck@htw-berlin.de

ABSTRACT

The ability to spell correctly is a fundamental skill for participating in society and engaging in professional work. In the German language, the capitalization of nouns and proper names presents major difficulties for both native and nonnative learners, since the definition of what is a noun varies according to one's linguistic perspective. In this paper, we hypothesize that learners use different cognitive strategies to identify nouns. To this end, we examine capitalization exercises from more than 30,000 users of an online spelling training platform. The cognitive strategies identified are syntactic, semantic, pragmatic, and morphological approaches. The strategies used by learners overlap widely but differ by individual and evolve with grade level. The results show that even though the pragmatic strategy is not taught systematically in schools, it is the most widespread and most successful strategy used by learners. We therefore suggest that highly granular learning process data can not only provide insights into learners' capabilities and enable the creation of individualized learning content but also inform curriculum development.

Keywords

Student strategies, Learning type, Online learning, German language, Spelling, Learning analytics

1. INTRODUCTION

The German language is known to be difficult to learn not only for nonnative speakers but also for native speakers who struggle with orthography [26]. However, a high degree of orthographic competence is crucial for successful communication with authorities and for professional success, as studies on employers and personnel selection show [21, 27].

One of the many peculiarities in the German language is capitalization. While nouns and proper names are generally capitalized, there are different linguistic perspectives on which words are considered nouns. Subsequently, learners can apply various redundant strategies to identify nouns.

Nathalie Rzepka, Hans-Georg Müller and Katharina Simbeck "What you apply is not what you learn! Examining students' strategies in German capitalization tasks". 2021. In: Proceedings of The 14th International Conference on Educational Data Mining (EDM21). International Educational Data Mining Society, 566-572. <https://educationaldatamining.org/edm2021/> EDM '21 June 29 - July 02 2021, Paris, France

Previous research has further indicated that these cognitive strategies for capitalization result in different patterns of errors that can be distinguished from each other [18]. While some learners consider the entire phrase when deciding whether to capitalize a word, others focus on only the word itself, especially the word ending, as an indication of the correct capitalization. Other learners use the words' meaning or take a pragmatic approach.

This paper aims to contribute to a better understanding of learners' cognitive strategies while processing capitalization tasks in German spelling courses. To this end, we use anonymized learning data on capitalization from the online platform orthografietrainer.net. The dataset consists of 9,647,385 single exercises completed by 30,658 users.

Identifying learners' cognitive strategies for capitalization tasks can enable educators and learning platforms to offer individualized help. Moreover, it can improve learning success by informing the implementation of personalized adaptive learning environments. Furthermore, comparing the predominant cognitive strategies in our large dataset to widely taught strategies in school can help inform future curriculum development. Previous studies of textbooks show that the set of rules taught in school contains semantic, morphological, and syntactic properties but almost completely lacks pragmatic strategy instruction [20]. Nonetheless, we found strong evidence that the pragmatic perspective is the major approach used by students of German.

In summary, we study the following research questions:

- RQ 1: Which cognitive strategies for capitalization are used by learners in grades 5 to 9?
- RQ 2: How does the use of capitalization strategies differ by grade level and gender?
- RQ 3: How do the predominant capitalization strategies used by learners compare to the strategies taught in school?

To answer the research questions, we proceeded as follows: The words used in the capitalization exercises on the online learning platform were manually one-hot encoded with 18 grammatical features associated with the four cognitive strategies for capitalization. In the next step, the four cognitive strategies for solving capitalization tasks were modeled as decision trees. Subsequently, the results of the four decision tree models were compared word by word with the solutions of more than 30,000 users.

2. RELATED WORK

2.1 Grammatical and cognitive approaches to German noun capitalization

The German orthographic system is complex and difficult to master. In contrast to other European writing systems, the difficulties relate less to spelling and more to the indication of grammatical structures. This can be illustrated by the capitalization of nouns, a

peculiarity of the German spelling system. Unlike in many other languages, in German, all nouns are capitalized. The ostensibly simple spelling rule that “nouns have to be capitalized” forces the speller to define precisely what is considered a noun and what is not. On closer examination, this question has a variety of very different possible answers.

On the one hand, there are many obvious nouns, such as people, places, things, and proper nouns. However, beyond that, every part of speech in German can be formally or functionally transformed into a noun. This is sometimes recognizable by a change in suffixes (cf. Ex. 1). In other cases, it can only be inferred from the syntactic context, for example, when articles or prepositions are added (cf. Ex. 2).

Ex. 1: fahren (V) → der Fahrer (N)
to drive → the driver

Ex. 2: fahren (V) → das Fahren (N)
to drive → the driving

The situation is further complicated by idiomatic expressions that formally contain a noun but that, from a pragmatic point of view, have lost their nominal characteristics (cf. Ex. 3). For instance, the supposed noun in Example 3 can still be formally complemented by an adjective, but this otherwise typical procedure for nouns is contrary to what a native German speaker would say. For this reason, the capitalization of such phrases is highly controversial in orthographic theory [5] and is a common source of error among students.

Ex. 3: im Allgemeinen → but not: im *häufigen Allgemeinen
in general → in *common general

Consequently, all of the nouns in the first sentence of Jane Austen’s “Pride and Prejudice” can be identified as nouns across several linguistic levels and work in both English and German:

Ex. 4: “It is a truth universally acknowledged that a single man in possession of a good fortune must be in want of a wife.”

Ex. 5: “Es ist eine allgemein anerkannte Wahrheit, dass ein Junggeselle im Besitz eines schönen Vermögens sich nichts mehr wünschen muss als eine Frau.”

Four of the nouns in the sentence occur with articles and in a typical syntactic environment for nouns. In addition, “man” and “wife” are identifiable as nouns by their concrete semantics. The words “truth” and “possession” are also marked morphologically since they were derived from the adjective “true” and the verb “to possess” with the help of a derivative ending.

The difficulties of coherent noun definition thus lie in the fact that a term may have a different extension depending on the linguistic perspective, although the semantic, morphological, syntactic and pragmatic perspectives agree in regard to a broad core of words. On the periphery, however, different perspectives lead to different conceptual boundaries and, consequently, to different orthographic decisions. The ground truth for what constitutes correct writing is therefore a mix between these different perspectives and defined by the Council of German Orthography [15].

The teaching of these different perspectives has been shown by an analysis of different textbooks [18]. The author found that capitalization is practically always introduced semantically. With the beginning of grammatical education in later primary school classes, morphological properties of nouns are added (especially the property numerus and some typical derivational endings), and articles as typical noun companions are introduced. At the secondary level, this knowledge is supplemented more systematically by further

morphological and, especially, syntactic properties of the noun group (e.g., gender, case, other determiners besides the article). However, in all courses, noun identification is based exclusively on formal-grammatical grounds. Only one of the textbooks examined also refers to pragmatic properties of the nouns [20].

Müller [20] demonstrated that errors in capitalization correlate strongly with different linguistic perspectives. Thus, some learners are apparently guided more by semantic aspects and others more by morphological, syntactic or pragmatic factors. These findings provide a starting point for our study, in which we attempt to model the different perspectives on noun capitalization using learning analytics methods to test whether different learning types can be distinguished.

2.2 Cognitive strategies of capitalization

Very little literature exists on the differentiation of orthographic strategies. Theoretical models [16, 24] distinguish between lexical and syntactic approaches, which roughly correspond to semantic and morphological strategies on the one hand and syntactic and pragmatic strategies on the other. Studies on the success of both approaches [28] have been limited to very small corpora and have produced partly contradictory results. The proposal of a division into four individual strategies was made by [19], who also found initial indications of different error profiles on the basis of an empirical study.

According to our linguistic considerations, the investigation is based on four theoretically distinguishable capitalization strategies:

The semantic strategy capitalizes words that have a concrete meaning. This strategy is primarily taught in early elementary first grade: “Things that can be touched have to be capitalized.”

Katze, Hand → but not: *nacht, *meinung,
cat, hand → night, opinion

The morphological strategy is to capitalize words that are classified as nouns because of the type of word and the word ending (word derivation):

Läufer → but not: (das) *laufen
Runner → (the) running

The syntactic strategy is to capitalize words that occur in a typical nominal syntactic environment, preferably in combination with attributes, articles or other determiners.

Die (totale) Dunkelheit → but not: *dunkelheit
ängstigt mich.
The (total) darkness → Darkness frightens me.

The pragmatic strategy is to capitalize words that are used in the current discourse like a nominal unit, which does not apply to all nouns. Pragmatically proper nouns can be supplemented with attributes or substituted with pronouns, which is often not possible with nouns in fixed phrases.

Der Grund → but not: im *grunde
the ground → in the ground
(saying for: “basically”)

Typically, the use of several strategies leads to success. Furthermore, there are many words with which capitalization errors are made only very rarely, for example, articles, prepositions, and pronouns. There are only a few words where using only one strategy leads to the correct result, and these words are not representative in the German language. Nevertheless, learners apply the strategies to different degrees and thus arrive at different results.

2.3 Spelling error analysis

We identify the learners' use of the previously introduced cognitive strategies through the analysis of error patterns. The analysis of spelling errors can help in understanding students' cognitive approaches to assignments [2]. Many studies use spelling error analyses to gain knowledge about second language learners; for example, studies [3, 4] analyzed spelling errors of native Arabic speakers in English courses or programs. Others investigate special subpopulations, as the authors of [2, 23] did with dyslexic learners. In addition to differences between native language and foreign language learning and between subpopulations, there are different classification schemes for spelling errors. Some authors have used Cook's classification from 1999 [9], which differentiates between omission, substitution, addition, transposition and sound-based errors [3, 4]. There are major differences between writing systems, and Abu-rabia [1] showed that these differences also affect spelling errors. For the German language, Landerl and Wimmer [14] used the "phoneme distance score" as a scoring method for spelling errors. Defior and Serrano [10] divided Spanish spelling errors into seven different classes of errors, which consist of substitutive spelling, partial spelling, random letters and nonorthographic spelling. Czech spelling errors were divided into phonological errors on the one hand and orthographic, morphological, grammatical, and lexical errors on the other hand [7]. The information gained about the learners can later be used in adaptive environments for different educational approaches to best address each student's abilities [11].

2.4 Learner-Level Adaptation

Adaptive learning environments aim to improve learning success by building personalized models of each student's knowledge, preferences and difficulties [6, 12]. The goal of such an adaptation is to individually optimize the learning path for each student [17]. This can lead to higher motivation, less overload and frustration, and, thus, better results [17]. Personalized adaptation to the student's needs can appear in a variety of forms, including task sequencing, intelligent solution analyses and problem-solving support [6]. The adaptations and the subsequent assessment of adaptive learning environments use a range of different data [8]. The parameters used most often in learner-level adaptation are parameters that refer to the user him or herself and his or her profile as a learner to optimize content [17, 22]. The learner profile consists, among other components, of the learner's behavioral pattern, learner preferences, cognitive traits or learning style as well as performance data [8, 17]. The learner's behavioral pattern can be analyzed by tracing his or her activities on an online platform. Learner preferences thus basically describe learners' preferred materials [22]. Another approach is to adapt a system based on learners' cognitive traits. These traits are their cognitive abilities, for example, their working memory capacity, abstraction ability or analysis ability [22]. There are various definitions of learning style. However, they all agree that there are different ways that learners experience learning [13]. Fang et al. [11] also differentiated between features of a learner's interaction with a system and individual differences between learners in terms of, for example, skill and knowledge.

All this information can be used by teachers to gain a better understanding of their students, leading to opportunities to adapt their teaching, materials or tests [13]. In addition, learners can be provided with appropriate materials and tasks that meet their needs. Finally, learning styles differ in terms of the sequencing of tasks [13]. The relationship between learning styles and the structure of the learning material has been investigated by, for example, the authors of [25], who found that students whose learning styles and

multimedia preferences match the material in their online course have higher scores.

In the context of this article, we suggest using the information gained about cognitive strategies for capitalization to display matching tips on online spelling platforms and to evaluate the difficulty of an exercise task in terms of which strategy is used.

3. DATASET

3.1 Orthografietrainer.net platform

The learning platform [orthografietrainer.net](https://www.orthografietrainer.net) offers online exercises for improving German spelling skills, including exercises on capitalization, punctuation, and spelling. The platform provides immediate and extensive individual feedback, which is impossible in a classroom setting. The training platform is built based on the pedagogical assumption that spelling requires not only knowledge but also skills. Thus, the focus is not on the regular learning of rules but on repeated practice [18].

The platform offers material for three different user groups: teachers, students and guests. Teachers register themselves and their entire class. They assign appropriate tasks to their students, who work on the tasks. Teachers can view their students' results via a dashboard. Additionally, any interested person can log in as a guest and complete tasks and tests.

A special exercise form on the platform is the competence test, which determines competence levels in capitalization, punctuation and separated or combined spelling. Any identified knowledge gaps are visualized, and appropriate exercises are suggested. A pre-test, an intermediate test and a posttest are available and show improvements made over time. For this study, we use only data from competence tests on capitalization, not regular training data, as the test's standardized structure allows for better comparison. Moreover, in competence tests, all sentences are new to users.

3.2 Description of the dataset

For this paper, anonymized, event-level competence test data from [orthografietrainer.net](https://www.orthografietrainer.net) from April 1, 2020 to November 17, 2020 are used. Each answer to a sentence corresponds to one record in our dataset. During the analyzed time period, schools in Germany, Austria and Switzerland were closed for several weeks due to the COVID-19 pandemic. In this period, 46,356 users visited the online platform and completed a total of 65,645 capitalization task sessions. When processing the tasks, nearly 50% of the sentences were answered incorrectly, which means that the answers each contained at least one mistake.

The platform was heavily used during the first wave of the COVID-19 pandemic. During the German school holidays in July and August, there was less practice; online training activity increased again in autumn.

The dataset contains information about the class level and gender of users. The German school system includes grades 1 to 13, with 1 being the youngest children and 13 being the oldest (Figure 1).

We decided to exclude all users in grades 1 to 4, as those learners are not well represented in the data set and the difficulty of the capitalization exercises is not adjusted for them. Students in grade 10 and above are also excluded. Older students who are still assigned capitalization exercises are well behind the average learning path and thus represent a marginal group that would bias the data. Most of the users are in grade 7. Our dataset contains slightly more girls (51%) than boys (49%).

Age	School	Class level
18	Secondary School (Second Phase)	13
17		12
16		11
15	Secondary School (First Phase)	10
14		9
13		8
12		7
11		6
10		5
9	Primary School (Sometimes extended to Grade 6)	4
8		3
7		2
6		1

Figure 1. German School System (simplified)

3.3 Using decision trees to replicate different cognitive strategies

The capitalization of German words depends on various grammatical categories, such as the beginning of sentences, word types and clauses. The 180 sentences in the competency test that deal with capitalization contain 2679 words that begin with either lowercase or uppercase letters. These 2679 words were manually categorized into 18 grammatical categories. After one-hot encoding of the labels, we obtained a data frame with dimensions of 2679 x 58.

It cannot be assumed that people use only one strategy; instead, it is likely that each person uses different manifestations of a variety of strategies. To be able to analyze the students' adoption rates of the strategies in regard to capitalization, we first needed to gain insight into how a student would process a word if he or she were only to use one strategy and had only one preferred learning type.

For this purpose, decision trees were used to replicate the four cognitive strategies by attributing to them only the grammatical features corresponding to the given strategy. Afterwards, the sentences from the competence tests were predicted by the decision trees and then validated to determine whether the user classified the words correctly or incorrectly in terms of capitalization. This provided us with the error profiles that would result if only one of the four strategies were used to decide on proper capitalizations. Table 1 shows the strategies and their grammatical features.

Table 1. Strategies with grammatical features

Strategy	Grammatical features
Syntactic	Clause, Article, 2nd person, Determinator, Is prefix, Attribute, Complement of a prepositional phrase, Beginning of sentence, Core nominal pronoun
Semantic	Concrete, Polite form, Semantic word type
Pragmatic	Theme-Rheme, Attributable, Proper name, as an attribute not separable from noun sequence
Morphological	Word type, Noun ending

In the decision trees, 77 % of the words were processed correctly by all four strategies. These are mostly words, where users make only a few mistakes, such as articles, pronouns, prepositions, and conjunctions. They are not interesting for further analyses, as they do not provide insights into differences between the strategies. The beginnings of sentences are also filtered out because they are a special case and cause bias in the data: in the structure of the exercises

on the online platform, the beginnings of sentences are in upper case letters per default. Students rarely click on such words to change the letter to lower case. However, as the beginning of a sentence is a syntactical feature, only the syntactic strategy processes these words correctly. Keeping the sentence beginnings part in the dataset would lead to bias, as most users would have a high adoption rate of the syntactic strategy precisely because the sentence beginnings are correct by default.

Table 2. Percentage of correct words per strategy

Syntactic	Semantic	Morphological	Pragmatic
74,11%	32,14%	62,05%	79,69%

The distribution of the remaining 448 words shows that strategies have different success rates (Table 2). The semantic strategy only processes approximately 30% of the words correctly, while the syntactic and pragmatic approaches are much better. This is not surprising, as the meaning of a word is less informative for the determination of capitalization than its grammatical use in a sentence.

Table 3. Sample of a merged data set

Word ID	User ID	Success	Syntactic	Semantic	Morphological	Pragmatic
255	452	1	1	1	1	0
256	128	0	1	0	0	1
257	427	1	0	0	0	1

In the next step, user data and the results from the decision trees are merged. The resulting data frame contains a word processed by a user in each row. For each word, there is information on whether the user capitalized the word correctly and how the decision tree models processed the item. Table 3 shows a sample of the resulting data set. In total, there are 1,355,641 records from more than 30,000 users.

4. RESULTS

To answer the first research question "Which cognitive strategies for capitalization are used by learners in grades 5 to 9", we compare users' error profiles with the error profiles of the decision tree classifiers. That for, we calculated the percentage of answers that matched. The adoption rate was calculated by dividing the sum of matching responses by the sum of processed words for each cognitive strategy. In this calculation, we did not consider whether the word was capitalized correctly. Instead, the result expresses only whether the words were processed in the same way by a user and by one of the four models.

Table 4 presents the average adoption rate per strategy in percentages. The models implementing the syntactic, morphological and pragmatic strategies were in alignment with the users' answers for, on average, 65% to 72% of the words. However, the result for the semantic strategy matched only approximately 40% of users' answers.

Table 4. Adoption rate by strategy

Syntactic	Semantic	Morphological	Pragmatic
65,33%	39,92%	66,32%	72,27%

When interpreting the results, it must be considered that several strategies can be used simultaneously when answering a task. This is always the case if the word cannot be answered exclusively by one strategy. Thus, overall, the adoption rate is over 100%.

4.1 Success rates

Thus far, we have primarily discussed the adoption of the four capitalization strategies. Now, we will examine the successful application of strategies for determining correct capitalization. Figure 2 shows the correlation between the adoption rates and the success rates per strategy.

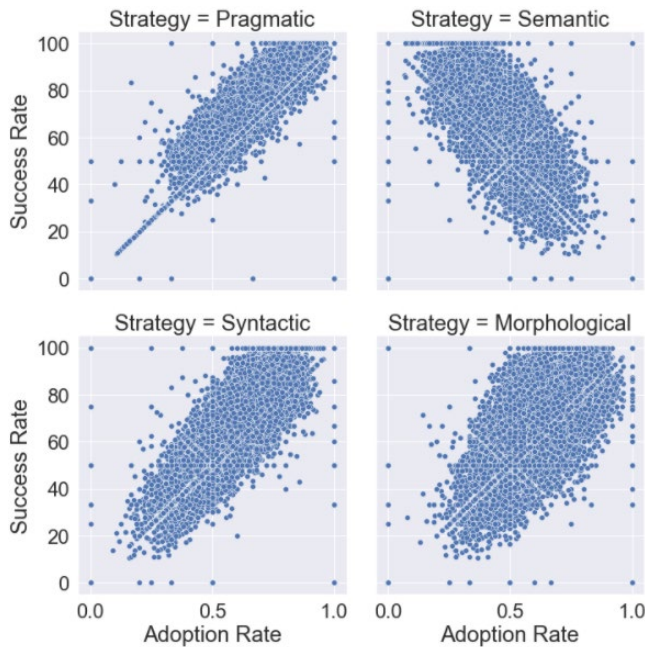


Figure 2. Correlation of success rate and adoption rate by strategy

The adoption of pragmatic, syntactic and morphological strategies led to increased success rates. The correlation is strongest for the pragmatic strategy. In contrast, the higher the share of words solved in agreement with the semantic strategy, the lower the success rate was. These correlations also exist when grade levels are considered in isolation. The success rates of the different strategies are also similar across grade levels.

The success distributed by class level and gender shows that students in higher grades tended to have lower success rates (Figure 3). While grades 5 to 7 had similar success rates, these declined from grade 8 onwards. The lowest success rates were found in grade 9. There is a very small difference between male and female success rates; however, in grades 7 to 9, male students correctly capitalized fewer words than female students did. It is possible that the data in these years reflect cognitive strategy shifts and corresponding temporary uncertainties.

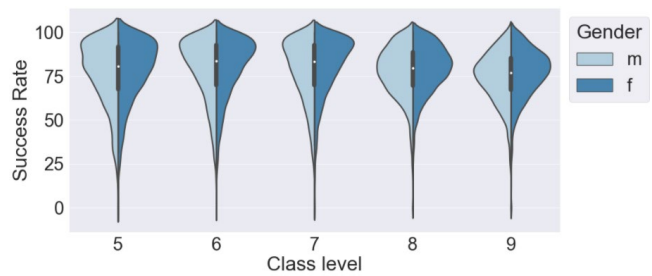


Figure 3. Success rates by class level and gender

4.2 Distribution by class level and gender

The second research question “How does the use of capitalization strategies differ by grade level or gender” is addressed by Figure 4. Looking at the distribution of the average adoption rate by strategy and grade level, we see that preferred strategies evolve over time and shift according to gender.

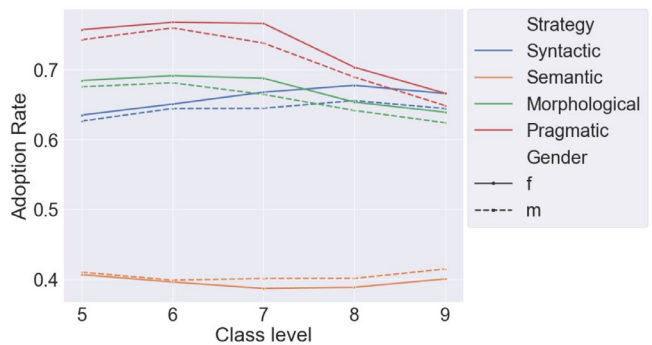


Figure 4. Distribution of adoption rates by class level and gender

The rate of adoption of the pragmatic strategy is very high from the beginning until it decreases sharply after grade 7 for girls and after grade 6 for boys. This is interesting, as the pragmatic strategy is the only strategy that is not explicitly taught in school even though it is very useful for determining correct capitalization (Figure 2). The pragmatic strategy is only surpassed in frequency by the syntactic strategy in grade 9, and the latter increases in use with every grade. Although the use of the syntactic strategy increases more for girls than for boys, in both cases, it ends up being on par with the pragmatic strategy.

Apparently, this reflects stronger grammatical skills among older students. Learners often start a second foreign language in grade 7 (usually Spanish or French), which increases the need for understanding grammatical concepts that are less explicit in their first foreign language, English. At the same time, usage of the morphological strategy also decreases from grade 7 onwards (as early as grade 6 for boys). These findings fit the students’ learning biography, as grammatical instruction progresses from morphological to syntactic issues, and therefore orthographic instruction focuses on morphological strategies first. The adoption rate of the semantic strategy decreases until grade 7 but then increases again. This fits with the results regarding the success rate of the semantic strategy, which shows a weakening of knowledge from grade 8 onwards. The increase in semantic strategy use thus goes hand in hand with the students’ lower success rates.

Looking at the differences in gender, we have already seen in Figure 3 that boys in grades 7 to 9 answer fewer words correctly than girls. If we now look at the use of the strategies by boys and girls

in Figure 4, we see that boys, especially in grades 7 to 9, use the semantic strategy more frequently, which is the least successful strategy and whose use correlates negatively with the success rate. Girls, on the other hand, use the other three strategies more frequently during this period, which correlate positively with the success rate.

In summary, we can again identify a difference between the semantic strategy and the other strategies. Even though the semantic approach is taught first, most learners do not adopt it for subsequent learning. The adoption rate of the pragmatic and morphological strategy decreases, while the syntactic strategy adoption rate increases. However, the pragmatic approach, which is rarely taught, is applied most frequently.

5. DISCUSSION

We used event-level learning data from an online spelling trainer to analyze cognitive strategies used by students for processing German language capitalization tasks. We built four decision tree classifiers to model capitalization strategies that use only syntactic, semantic, morphological or pragmatic features. As expected, as grammatical information in language is redundant, models often produce overlapping results. We compared the models' output to user error profiles. We found that the strategies are adopted to different degrees and that strong correlations—both positive and negative—between the adoption rates of strategies exist.

Furthermore, the distribution of adoption rates by grade level shows that strategies are represented among older and younger teenagers to different degrees. This variation by grade level is particularly interesting when compared to the rules taught at school, which answers the third research question “How do the predominant capitalization strategies used by learners compare to the strategies taught in school?”. The first capitalization strategy taught at school is the semantic strategy: things that can be touched have to be capitalized. Even though this is taught first, students follow it only partly—and rightly so, as the semantic strategy is the least successful in determining correct capitalization. The pragmatic strategy (capitalizing a word if it occurs in a typical textual context for nouns), however, is the only one that is not taught explicitly in school. Nevertheless, this is the strategy with the highest adoption rate and with the highest success rate in our research. The syntactic strategy presupposes a deeper understanding of grammar than the semantic and pragmatic strategies and thus increases with grade level. Although the syntactic strategy and the grammatical knowledge required for employing it begin to be introduced in grade 5, it is only later that students apply it. This may be because students' actual understanding of German grammar increases when they begin learning a second foreign language in grade 7. Since many grammatical concepts are not present in English, a deeper engagement with grammar might only begin when students begin learning a second foreign language. This could lead to a different way of looking at spelling, which is then reflected in the use of the syntactic strategy. The use of the morphological strategy decreases over time as the use of the syntactic strategy increases.

When considering the success rates in combination with the adoption rates, it is particularly interesting that the semantic strategy adoption rate correlates negatively with success rate. This again shows that the teaching of the semantic strategy as the basic rule does not lead to success. The strongest positive correlation with the success rate is the pragmatic strategy adoption rate.

6. CONCLUSION

In this paper, we have contributed to three aspects of learning analytics. We have identified cognitive strategies of learners using error analyses, compared adoption rates and drawn conclusions for curriculum development from the results.

First, we were able to model cognitive strategies for solving German language capitalization tasks. The four strategies (syntactic, semantic, morphological and pragmatic) do partially overlap. We have shown that the different learning strategy adoption rates can be observed in user error profiles (RQ1). This opens up opportunities for individualized training and therefore for higher motivation and learning success for students.

Second, we found that learners prefer different strategies depending on their grade level and gender (RQ2). This information can be used to adapt the online platform orthografietrainer.net to various learner levels. For example, based on this information, the difficulty of the words can be calculated more specifically for each user, and task sequencing can be adjusted to be neither too difficult nor too easy. This reduces the potential for frustration caused by tasks that are too difficult and also increases motivation. Furthermore, with tasks that represent typical sources of error for a user, the platform could display appropriate tips and hints. If the error analysis results are made available to the teacher on the dashboard of the online platform, he or she can see which rules have not yet been observed by the students and can adapt lessons accordingly. Further research could include the implementation and subsequent validation through A/B testing of such improvements.

Finally, our findings lead to a better understanding of how capitalization is learned and taught (RQ3). Our research shows that there is a great discrepancy between which strategies are taught in class and which strategies are used by students. We therefore suggest that highly granular learning process data can not only provide insights into learners' abilities and enable individualized learning content but also inform curriculum development.

Other future analyses could investigate whether the learning strategies can be applied to other grammatical areas, such as separated and combined spelling.

7. ACKNOWLEDGEMENTS

This research was funded by the Federal Ministry of Education and Research of Germany in the framework “Digitalisierung im Bildungsbereich” (project number 01JD1812A).

REFERENCES

- [1] Abu-rabia, S. 2002. Reading in a root-based-morphology language: the case of Arabic. *Journal of Research in Reading* 25, 3, 299–309. DOI= <https://doi.org/10.1111/1467-9817.00177>.
- [2] Abu-rabia, S. and Taha, H. 2004. Reading and spelling error analysis of native Arabic dyslexic readers. *Reading and Writing: An Interdisciplinary Journal* 17, 7-8, 651–690. DOI= <https://doi.org/10.1007/s11145-004-2657-x>.
- [3] Alhaisoni, E. M., Al-Zuoud, K. M., and Gaudel, D. R. 2015. Analysis of Spelling Errors of Saudi Beginner Learners of English Enrolled in an Intensive English Language Program. *English Language Teaching* 8, 3, 185–192.
- [4] Al-Oudat, A. 2017. Spelling Errors in English Writing Committed by English-Major Students at BAU. *Journal of Literature, Languages and Linguistics* 32.

- [5] Bredel, U., Müller, A., and Hinney, G. 2010. Schriftsystem und Schriffterwerb. *Linguistisch - didaktisch - empirisch* 289.
- [6] Brusilovsky, P. and Peylo, C. 2003. Adaptive and Intelligent Web-based Educational Systems. *International Journal of Artificial Intelligence in Education* 13, 159–172.
- [7] Caravolas, M. and Volín, J. 2001. Phonological spelling errors among dyslexic children learning a transparent orthography: the case of Czech. *Dyslexia* 7, 4, 229–245. DOI= <https://doi.org/10.1002/dys.206>.
- [8] Chen, W., Joe-Wong, C., Brinton, C. G., Zheng, L., and Cao, D. 2018. Principles for Assessing Adaptive Online Courses. *International Educational Data Mining Society*.
- [9] Cook, V. 1999. Teaching spelling.
- [10] Defior, S. and Serrano, F. 2005. The initial development of spelling in Spanish: From global to analytical. *Reading and Writing* 18, 1, 81–98. DOI= <https://doi.org/10.1007/s11145-004-5893-1>.
- [11] Fang, Y., Shubeck, K., Lippert, A., Chen, Q., Shi, G., Feng, S., Gatewood, J., Chen, S., Cai, Z., Pavlik, P., Frijters, J., Greenberg, D., and Graesser, A. 2018. Clustering the Learning Patterns of Adults with Low Literacy Skills Interacting with an Intelligent Tutoring System. *Proceedings of the 11th International Conference on Educational Data Mining* 11.
- [12] Gaudioso, E. and Boticario, J. G. 2003. Towards web-based adaptive learning communities. *Proceedings of the 11th International Conference on Artificial Intelligence in Education* 11.
- [13] Klačnja-Milićević, A., Vesin, B., Ivanović, M., and Budimac, Z. 2011. E-Learning personalization based on hybrid recommendation strategy and learning style identification. *Computers & Education* 56, 3, 885–899. DOI= <https://doi.org/10.1016/j.compedu.2010.11.001>.
- [14] Landerl, K. and Wimmer, H. 2000. Deficits in phoneme segmentation are not the core problem of dyslexia: Evidence from German and English children. *Applied Psycholinguistics* 21, 2, 243–262. DOI= <https://doi.org/10.1017/s0142716400002058>.
- [15] Lange, J. *Über den Rat*. Retrieved April 26, 2021 from <https://www.rechtschreibrat.com/ueber-den-rat/>.
- [16] Maas, U. 1992. *Grundzüge der deutschen Orthographie*. Niemayer, Tübingen.
- [17] Muhammad, A., Zhou, Q., Beydoun, G., Xu, D., and Shen, J. 2016. Learning path adaptation in online learning systems. In *2016 IEEE 20th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*. IEEE, 421–426. DOI= <https://doi.org/10.1109/CSCWD.2016.7566026>.
- [18] Müller, H.-G. 2008. Das Onlineportal Orthografietrainer.net - Lernpsychologische und didaktische Überlegungen zu einem automatisierten digitalen Übungskonzept.
- [19] Müller, H.-G. 2014. Zur textpragmatischen Funktion der Groß- und Kleinschreibung des Deutschen. *ZGL Zeitschrift für germanistische Linguistik* 42, 1, 1–25.
- [20] Müller, H.-G. 2016. Der Majuskelgebrauch im Deutschen. Groß- und Kleinschreibung theoretisch, empirisch, ontogenetisch. *Reihe Germanistische Linguistik* 305, Habermann, M. and Hausendorf, H., Eds. Berlin: de Gruyter.
- [21] Nimz, K. and Möhlmann, H. 2021. Zum Einfluss orthographischer Fehler in Bewerbungsschreiben. Eine experimentelle Untersuchung mit Personalverantwortlichen. In *Neue Wege des Orthografieerwerbs. Forschung-Vermittlung-Reflexion*, Kepser, M., Schallenberg, S. and Müller, H.-G. Eds. Lemberger, Wien.
- [22] Prematha, K. R. and Geetha, T. V. 2015. Learning content design and learner adaptation for adaptive e-learning environment: a survey. *Artificial Intelligence Review* 44, 4, 443–465. DOI= <https://doi.org/10.1007/s10462-015-9432-z>.
- [23] Protopapas, A., Fakou, A., Drakopoulou, S., Skaloumbakas, C., and Mouzaki, A. 2013. What do spelling errors tell us? Classification and analysis of errors made by Greek schoolchildren with and without dyslexia. *Reading and Writing* 26, 5, 615–646. DOI= <https://doi.org/10.1007/s11145-012-9378-3>.
- [24] Röber-Siekmeyer, C. 2003. *Ein anderer Weg zur Groß- und Kleinschreibung*. Klett-Grundsulverlag, Leipzig.
- [25] Surjono, H. D. 2015. The Effects of Multimedia and Learning Style on Student Achievement in Online Electronics Course. *Turkish Online Journal of Educational Technology - TOJET* 14, 1, 116–122.
- [26] Thomé, G. and Eichler, W. 2008. Rechtschreiben Deutsch. In *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie*, Klieme, E. Ed. Beltz Pädagogik. Beltz, Weinheim, Basel, 104–111.
- [27] Varnhagen, C. K. 2000. SHOOT THE MESSENGER AND DISREGARD THE MESSAGE? CHILDRENS ATTITUDES TOWARD SPELLING. *Reading Psychology* 21, 2, 115–128. DOI= <https://doi.org/10.1080/02702710050084446>.
- [28] Wahl, S., Rautenberg, I., and Helms, S. 2017. Evaluation einer syntaxbasierten Didaktik zur satzinternen Großschreibung. *Didaktik Deutsch: Halbjahresschrift für die Didaktik der deutschen Sprache und Literatur* 22, 42, 32–52.