Improved German Spelling Acquisition through Learning Analytics

3 AUTHORS:

Konstanze Edtstadler
Kirchliche Pädagogische Hochschule Wien/…
3 PUBLICATIONS   1 CITATION

SEE PROFILE

Markus Ebner
Graz University of Technology
5 PUBLICATIONS   1 CITATION

SEE PROFILE

Martin Ebner
Graz University of Technology
338 PUBLICATIONS   1,537 CITATIONS

SEE PROFILE

Improved German Spelling Acquisition through Learning Analytics

Authors

Konstanze, Edtstadler
konstanze.edtstadler@kphvie.ac.at
University College of Teacher Education Vienna/Krems, Researcher
Vienna / Krems, Austria

Markus, Ebner
markus.ebner@tugraz.at
Institute of Information Systems and Computer Media, Graz University of Technology, Researcher
Graz, Austria

Martin, Ebner
martin.ebner@tugraz.at
Department Educational Technology, Graz University of Technology, Head of Department
Graz, Austria

Many pupils struggle with the acquisition of the German orthography. In order to meet this struggle a web based platform for German speaking countries is currently developed. This platform aims to motivate pupils aged 8 to 12 to improve their writing and spelling competences. In this platform pupils can write texts in the form of blog entries concerning everyday events or special topics. Since the core of this platform consists of an intelligent dictionary focussing on different categories of misspellings, students can improve their own spelling skills by trying to correct their mistakes according to the feedback of the system. Teachers are informed about specific orthographic problems of a particular student by getting a qualitative analysis of the misspellings from this intelligent dictionary. The article focuses on the development of the intelligent dictionary, details concerning the requirements, the categorization and the used wordlist. Further, necessary information on German orthography, spelling competence in general and the platform itself is given. By implementing methods of learning analytics it is expected to gain deeper insight into the process of spelling acquisition and thus serves as a basis to develop better materials in the long run.

Tags

German orthography, learning analytics, qualitative analysis of misspellings, categorization, Technology Enhanced Learning

1. Introduction

This article is concerned with a learning analytics approach in the field of German orthography. Due to the increasing internet usage in the field of education, the amount of data that is produced is rising daily. This data is shared between various kinds of institutions around the globe (Piety, 2013). Furthermore, the heavy use of the Internet generates enormous data about learners' behavior and leaves traces of every interaction (Duval, 2010). Thus, interaction between students and a learning platform can be captured and used for later analysis in order to gain an insight into a learners' learning process (Khalil & Ebner, 2015). This can then be used for early detection of learning issues and enables teachers to actively intervene accordingly in order to solve such issues effectively (Siemens et al., 2011; Greller & Drachsler, 2012).

German orthography is known to be quite difficult to master. People from different social classes, of different ages and with varying degrees of education, struggle with spelling words correctly. However, the importance of correct spelling for social acceptance is quite high. It affects primary-school pupils' as well as a university students' everyday life inside and outside

schooling. Still, instructions in German orthography are often unsystematic and not particularly attractive for children.

The development of the IDERBLOG-Platform[1] aims to solve such problems by combining technology enhanced learning and learning analytics with the acquisition of German orthography (Ebner et al., 2015a). The platform should serve as an attractive and motivating innovation for children to acquire German orthography appropriately and more easily. It has also advantages for teachers and researchers, as the application of learning analytics supports them in their decision making process by providing them with an overview of possible educational interventions (Ebner et al., 2015b).

Outline

The next section gives a short overview of the German orthography as well as orthographic competence and learning analytics. The following section is concerned with the development of the information system of the platform, its interface design process and the planned workflow. The two succeeding sections focus on to the intelligent dictionary and the feedback system. The article aims to give an overview of the categories, the requirements and the process of categorization of the intelligent dictionary.

2. Related work

German Orthography

German orthography uses an alphabetic writing system. Alphabetic writing systems are characterised by mirroring the phonemic structure of the spoken language to the written language, which leads to the assumption that words are spelled as they sound (cf. Katz & Frost, 1992, p. 149).

This phonological principle is applied to a varying degree of consistency in different languages. It leads to a continuum of orthographies ranging from transparent to opaque ones with a huge impact on spelling instruction and acquisition. In transparent orthographies like Serbian, Turkish or Italian each phoneme (notated consecutively with / /) is represented by one letter – or more precisely – grapheme (notated consecutively with < >). Therefore, the assumption to spell a word as it is heard is quite true in these orthographies. In opaque orthographies like English or French "there is a pronounced discrepancy

between pronunciation and orthography" (Klees, 1989, p. 137). Consequently, learners are confronted with unreliable correspondences since – in the case of English for example – "the alphabet contains just 26 letters yet these correspond to 44 phonemes associated with 102 functional spelling units." (Snowling, 1989, p. 1). The German orthography can be found in the middle between transparent and opaque orthographies. Following Nerius (2007) it consists of two basic principles, the phonological and the semantic principle. Part of the phonological principle are the phoneme-grapheme-correspondences (PGC), which are mostly not in a 1:1 relationship, e.g. /a:/ can be <a> in <Wal> *whale*, <aa> in <Saal> *hall* or <ah> in <kahl> *bald* (cf. Meinhold & Stock, 2007, p. 122). Part of the semantic principle is, according to Nerius (2007, p. 89 ff), the morphological principle - among the lexical, the syntactic and the textual one. This principle is responsible for spelling one morpheme in the same way in all words in which it occurs. This often leads to a conflict with the phoneme-grapheme-correspondences: e.g. spelling the word *dog*, in German pronounced as /hunt/, following the PGC would lead to the misspelling (usually indicated with an asterisk) <*Hunt>. It has to be spelled <Hund> because of the plural form /hundə/ *dogs*. The spelling <Hund> with a <d> is kept the same in all words, like <Hündin> *female dog* or <Hundeschlitten> *dog sled*. The spelling of the orthographically challenging 'Umlaut' (= vowel mutation spelled as ä/äu) in morphologically complex words is also due to the morphological principle (e.g. <Hände> *hands*, not <*Hende>).

These principles and their value for German orthography highly affect the didactic approach of teaching – especially in higher classes and additional trainings. In general, the spelling instruction at the beginning of literacy acquisition is clearly phoneme based (cf. Landerl & Thaler, 2006). This is the reason why children are able to write (new) words relying on their knowledge of pronunciation soon after getting to know the PGC. Words written in this way can also be read, but they are often not orthographically correct (e.g. <*falipt> for <verliebt> *in love*). Especially for children who are not speaking the standard German variety, the influence of the spoken language is evident in their spelling. Due to other sub-principles of the semantic basic principle further orthographic challenges are – for example:

- Nouns must be spelled with capital letters – a feature that can only be found in the German orthography (cf. Valtin, 1989, p. 119). It leads to many mistakes – even in texts of well-educated adults.

- Homophone words are sometimes, but not always, spelled differently (e.g. /liːt/ as <Lied> song or <Lid> eyelid, but /notə/ as <Note> for mark and note) (cf. Nerius, 2007, 167).

- Compounds are usually spelled compound (e.g. <Teetasse> tea cup). Depending on the meaning of a combination of words, it must be spelled separately or compound (e.g. <Schweinebraten> roast pork or <Schweine braten> – to fry pork) (cf. Fuhrhop, 2011, p. 107).

Spelling Competence

Especially in the German speaking world correct spelling is considered very prestigious, but students consider spelling instructions often as boring and formal (cf. Küttel, 2007, p. 418f). Additionally, teachers often do not pay attention to the systematic principles that stand behind the spelling of certain words. This often leads to the assumption, that it is not possible to teach German orthography systematically (cf. Fröhler, 2002). In contrast to other areas of language learning, there is hardly space to argue about the correct or incorrect spelling of a word. This orthographical stiffness can probably serve as an explanation for its importance.

It is important to understand that the spelling competence of a person does not only include the knowledge of the correct spelling of a given word and knowing the rules of orthography. It also includes being sensitive to misspelled words, knowing how to correct them, using spelling aids and applying strategies (cf. Sommer Stumpenhorst, 2012; Naumann, 2008). Concerning instruction, it is not enough to simply offer different online or offline exercises. "Children or student's need purposeful reading and writing in a broad range of situations, in an environment that values risk-taking. They will develop spelling competence as they implement their knowledge of the spelling system, receive feedback and refine their hypotheses." (Government of South Australia, 2011, p. 6). Furthermore, children should be encouraged to think about and reflect language in order to become aware of the structure of words (cf. e.g. Tsesmeli & Seymour, 2006). Due to the different principles of German Orthography, metalinguistic awareness must be established beyond phonological awareness (cf. e.g. Naumann, 2008). For example, children must be encouraged to see the morphological link between singular and plural form (e.g. <Hälse> because of <Hals> *necks*, <Rind> because of <Rinder> *cows*).

Learning Analytics

The field of Learning Analytics tries to consider the learning process as a whole in its full complexity. According to Baker et al. (2012) and Neuhold (2013) it is important to keep feedback and its visual representation as simple as possible to avoid confusion and unreasonable interpretation on the side of the stakeholders. Campbell et al. (2007) provide a model for the analysis process in five steps: capture, report, predict, act and refine. Clow (2012) used these five steps as a basis for his learning analytics cycle. This iterative process consists of four main components: learners, data, metrics/analytics and intervention (Clow, 2012). To get an overview about the whole process Khalil & Ebner (2015) added stakeholders to the cycle. Nevertheless, the main idea of Learning Analytics is to provide and process a learners' data in an appropriate way in order to facilitate teachers to react and (if necessary) to intervene. For instance, Taraghi et al. (2015a) introduced an analytical approach to model a learner's profile according to their answering behavior. Moreover, the analysis of different error types can lead to findings that help to enhance the learning process as a whole (Taraghi et al., 2015b).

3. Information system

The platform (information system) for the project is currently under development and yet not available for public presentation. Nevertheless, this section will provide basic design ideas to ensure good age-appropriate interface design and usability (Ebner et al., 2015a). In the second section the planned workflow of the analysis will be outlined.

Writing by using the Computer

Since developing writing skills and acquiring orthographic competence is important and writing with computers is attractive for children, the IDERBLOG-Platform combines these components. The aim is not to replace handwriting by typing on keyboards, but to take advantage of the digital age. "For some people with major handwriting problems, personal computers are a boon." (Høien & Lundberg, 2000, p. 68)

A further advantage of writing on a computer is, to train the ability to correct texts. Since corrections are made within a digital text, corrections do not leave traces in contrast to a handwritten text. Consequently, a text can be edited several times until it becomes publishable. Furthermore, the IDERBLOG-Platform is "providing relevant reasons and audiences for writing"

(Government of South Australia, 2011, p. 8) as children can publish their texts. Therefore, it is expected that the motivation to formulate a text and to revise it several times is possibly higher in contrast to typical essay writing in a classroom where the addressee is almost only the teacher.

Concerning the training of orthographic skills, the IDERBLOG-Platform offers an intelligent dictionary, which does not only count the number of mistakes in a text, but also categorizes the mistakes in different orthographic areas. In contrast to the work of Thelen (2010), that analyses misspellings in German orthography, we do not only focus on beginning spellers but also on more advanced learners. One of the most important features of the intelligent dictionary is that it offers feedback and hints for the correction of a mistake. Additionally, the platform offers a number of exercises that are connected and categorised according to spelling mistakes and therefore meet the need of practice in a specific area of spelling.

Interface Design

The platform is generally designed for children the primary school (age 8 to 12) with the focus on a graphically appealing and age-appropriate web interface (Liebal et al., 2011). For this purpose, a graphic designer created drafts that have been examined and rated by students from different schools. The designs that were favoured by the majority were then, in a second step, developed further and afterwards integrated into the platform.

Another important part of the platform is usability. We had to ensure that the students can reach the most important parts of the platform in less than five clicks. This convenient accessibility in combination with attractive figures should ensure high motivation in fulfilling the task of writing texts. In ongoing usability tests (Holzinger et al., 2005) we continue to improve the concept step by step.

Workflow of the Platform

The students, as shown in Figure 1, can write their texts on the provided platform. First the text will be analysed orthographically by the intelligent dictionary (which will be described in the next section). Proper feedback will be provided to the student, based on error type and category. The student has the choice to either try to correct the wrong words or to hand-in the text directly to the teacher. This intermediate step encourages the

expertise of independent correction (Bartnitzky et al., 2010). After the submission, the teacher should inspect the text for further corrections and/or improvements. Notes can be made and delivered with the final correction to the student. After this step the text can be published in the class blog of the school (if appropriate).
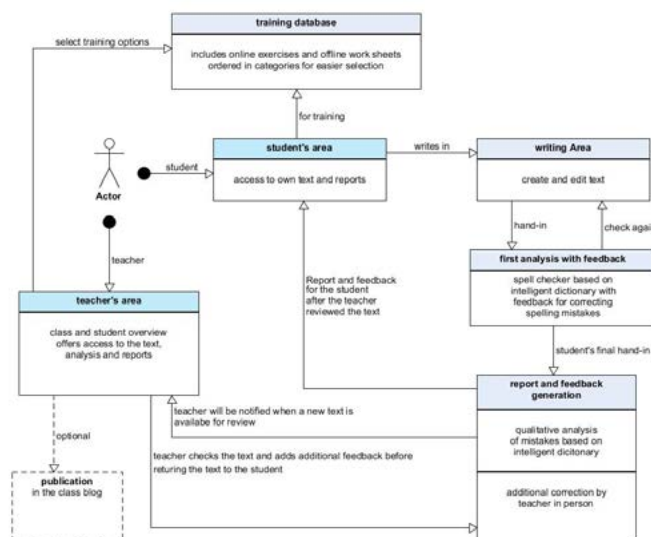

Figure 1: Workflow of the platform

The methods of learning analytics will be used for further analysis of the texts (Siemens, 2012). The results will be provided to students, teachers and parents in an appropriate way. Further, an overview of the frequency of mistakes and possible systematically made errors is offered. In the long run changes in a students' performance will be measured (Schön et al., 2012).

Training database

The platform will include an additional training database, as shown in Figure 1, with selected online exercises and offline work sheets. This database will aid teachers and students to find appropriate exercises to improve the performance in problematic areas (as a consequence of the learning analytics analyses). The exercises and work sheets are congruently ordered in categories and sub-categories for easier selection.

4. The intelligent dictionary

Categorization of mistakes

A word can be either orthographically right – in case the spelling of a given word exists in a list of correctly spelled word, called

dictionary, or wrong – in case it does not exist. In science, mistakes are analyzed in different categories depending on the purpose of the study: e.g. for the English language Broc et al. (2013) categorise spelling errors of people with specific language impairments in phonologically acceptable vs. unacceptable. Flor & Futagi (2012) focus on non-word misspellings in the context of spell checker. In school often a quantitative approach is applied, which means counting the number of misspelled words. In addition to the correct-wrong dichotomy there are some other ways in the categorisation of incorrectly spelled words that lead to a greater insight into the orthographic competence.

One way is to count the number of correctly written graphemes of a given word: This helps to analyse the progress of extremely weak or very young spellers (cf. May, 2010). For example, the spelling for /V/e/r/k/äu/f/e/r seller in *F/e/r/k/eu/f/a contains only 4 out of 8 correct graphemes in contrast to */V/e/r/k/eu/f/e/r with 7 out of 8 correct graphemes. Although both writing attempts are wrong, the second one is much better. This grapheme-based approach is a rather detailed and time consuming way of correcting. Therefore, it is generally only applied in a standardized spelling test called "Hamburger Schreibprobe" (May, 2010) which provides pre-defined templates for the quite small amount of words used in the test.

Another way of categorizing incorrectly spelled words is to define the type of mistake(s) and to collect the various frequencies for the given categories in order to identify the orthographic areas that need to be worked on (cf. e.g. Naumann, 2008, p. 139; Thomé & Thomé, 2014). The determination and the assessment of these categories vary and are highly depending on the purpose. The applied systems range from unpublished templates developed by teachers to published and buyable ones. For example, scientifically based and evaluated templates for texts can be found in the "Oldenburger Fehleranalyse" (OLFA) (Thomé & Thomé, 2014) and those specifically meant for the qualitative analysis of standardized tests can be found in the "Aachener Förderdiagnostische Rechtschreibfehleranalyse" (AFRA) (Herné & Naumann, 2002). When using the OLFA (Thomé & Thomé 2014) the teacher has to collect texts containing a certain amount of mistakes. Each mistake of a word has to be analysed and categorized in one of the 35 categories (e.g. lower case letter instead of upper case letter, upper case letter instead of lower case letter, omission of a vowel, addition of a vowel, etc.) that are described in detail in an accompanying manual. Since the different categories are related to the stages of spelling development the teacher gets to know the level of spelling competence of a student.

In all described cases above, the time consuming analysis of misspellings must be done by the teacher personally. This requires effort to get familiar with the theory of German orthography and the (applied) way of analysing the mistakes. From our experiences, a detailed analysis is made only by highly specialised people in rare cases. Since a clear qualitative analysis of misspellings is the basis for a good and target oriented intervention, the IDERBLOG-Project aims to conduct the analysis in large part automatically in order to support teachers and consequently foster the spelling acquisition process for children.

Requirements for analysis

The categories of the qualitative analysis for the intelligent dictionary need to fulfil some requirements on scientific, technical and practical basis. In order to fulfil all these requirements, the system of categories is established on different hierarchical levels from fine to coarse grained. This has the advantage, that the system stays flexible as each level is mainly dedicated to a specific purpose. We had to take into account that many different words belong to one category of mistakes. In order to provide a detailed analysis, we split a category into specific phenomena (see table 1). Based on those we have a proper fine-grained level for the application of learning analytics. However, those phenomena on such a detailed level are not suitable for a general feedback. Therefore, the phenomena of this specific level are merged in order to retrieve a qualitative analysis for the teacher with a manageable amount of categories and in order to be linked to the database containing appropriate orthographic exercises. It also needs to be taken into account that the naming of the categories that are visible for the teachers and/or children, are scientifically correct but still easy to understand and consistent with the established terms used in school environment, which are not always consistent with the scientific terminology.

Method

In order to establish the different categories, a literature survey was conducted and well-known approaches for qualitative analysis for misspellings within the German orthography were evaluated (cf. Edtstadler, in press). At the same time, 55 short texts of 3rd grade students from Germany and a limited

number of longer texts of 5th and 7th graders from Austria were collected. In a next step a draft of categories was developed based on the findings of our research of relevant literature as well as by using, as a starting point, established categories of existing analysis methods such as OLFA and AFRA. This draft was rearranged, modified and extended in order to meet the requirements of the intelligent dictionary. It was especially challenging to construct the categories in a way that the description of the phenomena fits in with the possibilities of programming misspellings as well as with the categories for the teachers and the database with the exercises. Additionally, specific misspellings due to the existence of different German varieties are considered.

The usability and suitability of the draft's categories were further checked by assigning one mistake encountered in the above mentioned texts from different regions of the German speaking area to a phenomenon, for which a feedback can be given. Also, phenomena of mistakes that were not found in the quite small amount of texts' (mentioned above), but are theoretically possible and/or by experience encountered in students' texts were added: For example, in the texts an inflected form shows that the 'Umlaut' is substituted by <e> (<*fengt> he/she/it catches instead of <fängt> because of <fangen> to catch),

therefore, the substitution in plural forms (<Apfel - Äpfel> apple - apples), derivations (<Glanz – glänzend> shine - shiny), and comparisons (<warm - wärmer> warm -warmer) were also added.

General Description of the Categories

The categories are established on a linguistic and orthographic basis, also by regarding previous findings of the theory of German Orthography (e.g. Nerius, 2007). Consequently, the system (see table 1) is divided in two parts: On the one hand, the system contains the parts that are invisible for the user where the scientific and theoretical basis can be found. This is necessary for the description of the phenomena and consequently for programming the possibly misspelled words. The visible parts, on the other hand, appear in the qualitative analysis for the teacher, serve for the selection of exercises from the training database and appear in the feedback the writer gets in case a word is not spelled correctly.

Since this system is quite complex, the described system is shown in table 1. The orthographic area of 'Umlaut' serves as an example.

Linguistic level (not visible)	Ortho-graphic area (not visible)	description/ rule based phenomenon (not visible)	Category for the teacher (visible)	Category of spelling exercise (visible)	Sub-category of spelling exercise (visible)	Example of a misspelled word	Feedback for the writer
Morpho-logical level	Um-laut	Inflection of nouns: e/eu for ä/äu	Umlaut derivable	Morpho-logical hints	Derivation	apples: not <*Epfel> but <Äpfel> because of <Apfel> apple	Think, if there exists a base form with a.
Morpho-logical level	Um-laut	Inflection of verbs: e/eu for ä/äu	Umlaut derivable	Morpho-logical hints	Derivation	he/she/it catches er/sie/es not <*fengt> but <fängt> because of <fang-en> to catch	Think, if there exists a base form with a.
Morpho-logical level	Um-laut	Comparison of adjectives: e/eu for ä/äu	Umlaut derivable	Morpho-logical hints	Derivation	warmer: not <*wermer> but <wärmer> because of <warm> warm	Think, if there exists a base form with a.
Morpho-logical level	Um-laut	Word formation / derivation: e/eu for ä/äu	Umlaut derivable	Morpho-logical hints	Derivation	shiny: not <*glenzend> but <glänzend> because of <Glanz> shine	Think, if there exists a base form with a.

Table 1: Example of the system of the intelligent dictionary in the orthographic area of 'Umlaut'

Starting from four linguistic levels – phonological, morphological, lexical and syntactic – the categories are further divided into orthographic areas. The phonological, the lexical and the syntactic level consist of three orthographic areas. The morphological level that is used for giving insight to the complex system (cf. partly Nerius, 2007, p. 158ff) contains five orthographic areas, including 'Umlaut' (see table 1), which are essential for spelling words correctly:

1. Morpheme constancy: lexical and grammatical morphemes are spelled the same way in compounding, derivation and inflection, even when the sound cannot be heard (e.g., <Ohrring> earring, <Weihnachten> Christmas because of <weihen> hallow, <stehen> not <*stehn> because of the suffix –en, <Verkäufer> seller because of the prefix ver-).

2. Morphological hints for using capital letters: the use of capital letters for nouns is quite difficult and depending on the syntax, but because of certain suffixes, derivations can easily be identified as nouns that must be written with capital letters, e.g. <*belohnung> gratification because of the suffix –ung.

3. Morphological hints for not using capital letters: there are also some suffixes that indicate that a given word is not written with a capital letter, although morpho-syntax can change the word class, e.g. <*Furchtbar> horrible is correctly spelled <furchtbar> because of the suffix –bar, but it is spelled with capital letters in the phrase <etwas Furchtbares> something horrible whereas the use of the same word form as an adjective requires the use of lower case, e.g. <ein furchtbares Gewitter> a horrible thunderstorm.

4. 'Umlaut': Because of phoneme-grapheme correspondences, especially in the area of Austria, the 'Umlaut' is often incorrectly written as <e>, e.g. <*glenzend> instead of <glänzend> shiny, since the 'Umlaut' needs to be applied because of the base morpheme <Glanz> shine (for details, see table 1).

5. Terminal devoicing: In German a word is pronounced with a devoiced obstruent at the end of the word, but spelled with the voiced variant of the phoneme-grapheme correspondences (e.g. /hunt/, but spelled as <Hund> dog because of the wordform /hunde/ whereas <Brot> bread is spelled as <Brot> because of /bro:t – bro:tə/).

Each orthographic area is associated with a wide range of phenomena. These phenomena are formed in a way that they can function as a rule for programming the possible mistakes (see table 1). The number of phenomena is depending on the given orthographic area and can be expanded and reduced, based on evidence. The following example will help to show the variety in the number and the characteristics of phenomena: In the orthographic area of morpheme constancy the category derivational suffixes (for the analysis for the teacher) summarises the phenomenon of misspelling suffixes such as –ig (e.g. <lustig> funny) (written in different ways depending on the spoken German variety as –ich <*lustich>, -isch <*lustisch>, -ik <*lustik>), and the phenomenon of spelling the suffix –lich as <*-lig> as well as further phenomena describing the misspelling of other derivational suffixes.

As mentioned above, it is important to work with a manageable amount of categories when offering the qualitative analysis for the teacher. Therefore, the currently 110 phenomena are linked with 34 categories of the qualitative analysis for the teacher. In the example in table 1, the four defined phenomena for misspelling the 'Umlaut' are summed up in one category that tells the teacher that within a certain amount of mistakes the 'Umlaut' was derivable, but incorrectly spelled with the wrong grapheme.

The categories of the qualitative analysis are then connected with and/or mirrored in the labels of the orthographic exercises available on the platform. For an easier orientation they are divided in categories and sub-categories of exercises. However, the labelling of the exercises is in some cases more coarsely grained than the category of the qualitative analysis itself (e.g. the category upper case instead of lower case and the category lower case instead of upper case are labelled as upper and lower case exercises, since a lot of exercises practice both at the same time). This is due to the fact, that in the first step only already existing exercises are available on the platform, but in the progress of the project specific exercises will be developed.

The phenomena also function as the starting point for the feedback of the intelligent dictionary, which will be described more in detail below.

5. Feedback from the Intelligent Dictionary

All of the categories and phenomena form the basis for the analysis and application of the intelligent dictionary, which is the core of the platform. The idea is that a child, who misspells

a word, does not only get the feedback that the word is spelled incorrectly, but also gets a hint for correcting it. The given feedback is connected with the phenomenon. In order to keep a straightforward number of feedbacks the same feedback will be given – whenever possible - for more than one phenomenon within an orthographic area. It is formulated in a way that forces the child to think about the spelling and further encourages the development and application of spelling strategies. Therefore, no direct commands (e.g. "use <ä> instead of <e>") for correcting the word are included in the feedbacks. The correction will only be successful if the child reflects on the hint in combination with the misspelled word. This approach stands in contrast to the usual word-correction process where either the misspelled word is marked or the correct word needs to be selected from a variety of offered words. In both cases the correction will probably not lead to a deeper understanding of correct orthography.

Klicpera et al. (2003, p. 255) mention that in order to acquire correct spelling, it is important to offer exercises that allow the autonomous correction in a motivating context. Experienced teachers and trainers for dyslexic children know that poor spellers have problems in identifying their mistakes in a text. But as soon as a hint for correcting the word is given, they often know how to spell it correctly. This is a successful, but a time and energy consuming way of improving orthographic competence. In order to avoid exhaustion, the intelligent dictionary gives this feedback instead of a teacher, which also has the advantage that the intelligent dictionary can and will repeat the feedback several times. In case a child spells <*Epfel> instead of <*Äpfel> apples the intelligent dictionary provides the feedback "Think if there exists a base form with "a"?", the same feedback will be given in case the child writes <*fengt> instead of <fängt> to catch or <*glenzend> instead of <glänzend> shiny. There is no strict application of a syntactic pattern for the feedback since the wording of the feedback is chosen rather on a didactic than on a formal basis.

Wordlist of the intelligent dictionary

Since this intelligent dictionary so far, is designated to function as a first prototype, only a selection of words functions as the basis for programming the dictionary. For the first prototype we had to choose around 1000 words. Generally such a selection of words would be based on the frequency of the CELEX (1995) database – although this would propose some problems (cf. Brysbaert et al., 2011). Selecting words only based on frequency in general without considering the frequency of words in children's language is especially problematic for the development of an application aimed at children. Also the fact that the selected words should be prone to be misspelled had to be considered (for a discussion see Risel, 2008).

In order to meet these requirements, the word list for the prototype of the intelligent dictionary is based on the basic vocabulary of three German Federal States (Bavaria, Hamburg, Berlin-Brandenburg). In the next step it was checked, whether in these basic vocabularies the 100 most frequently misspelled words of 4th graders (compiled and made available by Tacke, 2008) are included. Words that do not appear in any form in one of these three basic vocabularies were included (e.g. <kommt> comes was not included since <kommen> to come is a word of the basic vocabulary, but <ziemlich> quite was included since it does not appear in one of the basic vocabularies). At the end, the word list for the prototype of the intelligent dictionary ended up containing around 1100 words.

Since German has a rich morphology and texts are not merely made of words that are listed in a dictionary, it is necessary to list all the possible word forms of a given word in order to construct all possible misspellings for the intelligent dictionary in a next step. The collection of all possible word forms of a given word (or to be precise of a lemma) is based on the CELEX (1995) database. This incorporation of all word forms enlarged the wordlist to over 7500 orthographically correct words. In German the variation in the number of word forms for a given lemma is quite high as is proven by the following examples:

- For the adjective <ähnlich> similar the CELEX (1995) database has 17 word forms (<ähnlich, ähnliche, ähnlichen, ähnlicher, ähnlichem, ähnliches, ähnlichst, ähnlichste, ähnlichsten, ähnlichster, ähnlichstem, ähnlichstes, ähnlichere, ähnlicheren, ähnlicherer, ähnlicherem, ähnlicheres>) including inflection for singular, plural, the different cases and comparison.

- For the regular verb <arbeiten> *to work* 10 word forms can be found in the CELEX (1995) database in summary, whereas for the irregular verb <beginnen> to begin exactly 24.

- For the noun <Beispiel> *example* exist only four word forms.

- For prepositions no other word form can be found since they cannot be modified.

For these 7500 word forms of this wordlist all possible mistakes for a given category are constructed and connected with the feedback. For instance, for the orthographic area of 'Umlaut' all words containing an <ä> in the wordlist must be searched for. To consider the different phenomena where substituting <ä> for <e> is a mistake due to the morphological principle, the search is done separately for verbs (e.g. <*fengt> instead of <fängt> he/she/it catches), nouns (e.g. <Epfel> instead of <Äpfel> apples), adjectives (e.g. <*kelter> instead of <kälter> colder) and derivations (e.g. <glenzend> instead of <glänzend> shiny). But, since the 'Umlaut' in the word <ähnlich> and its substitution with the incorrect <e> does not qualify for the morphological level it must not be included in this category but has to be added in another appropriate category with according feedback.

6. Conclusion

In this article we introduced a platform that aims to motivate children to improve their spelling skills by writing and publishing texts. In this platform an intelligent dictionary is integrated and based on the presented system of categories, the intelligent dictionary gives feedback in order to enable children to correct mistakes with the help of this feedback. The platform also provides a qualitative analysis for teachers, who can use the results in order to help pupils with the improvement of word spelling. Concerning learning analytics, the occurred misspellings can also be used for an in depth analysis.

The development of the platform and the intelligent dictionary is still under construction and changes are still possible. There are still issues such as the identification of several mistakes in one word that will most certainly lead to further discussion in the future. However, we are positive that this combination and the interdisciplinary work of the IDERBLOG-Project will in future motivate more children from grade 3 on to write texts and to improve their spelling competence. Further, we can support teachers by providing analysis and material for the improvement of spelling. The active application of the methods of learning analytics in this area of language learning will help us to understand the process of spelling acquisition in more detail. It is expected that this unique combination in one platform has a positive impact on didactic approaches, education and science.

[2]   LPM Saarland, Beethovenstraße 26, 66125 Saarbrücken, Germany
[3]   University of Teacher Education Vienna, IBS/DiZeTIK, Grenzackerstraße 18, 1100 Vienna, Austria - Europe
[4]   Albert-Weisgerber School St. Ingbert, Robert-Koch-Straße 4, 66386 St. Ingbert, Germany
[5]   School of Raeren, Hauptstraße 45, 4730 Raeren, Belgium
[6]   University College of Teacher Education Vienna/Krems, Mayerweckstraße 1, 1210 Vienna, Austria - Europe
[7]   Graz University of Technology, Department Educational Technology, Münzgrabenstraße 35a, 8010 Graz, Austria - Europe

References

Baker, R. S. J. D., Duval, E., Stamper, J., Wiley, D. & Buckingham Shum, S. (2012). Panel: educational data mining meets learning analytics. In Buckingham Shum, S., Gasevic, D. & Ferguson, R. (Eds.) Proceedings of 2nd International Conference on Learning Analytics & Knowledge (LAK'12), New York, USA, 20.

Bartnitzky, H. (2010). Individuell fördern–Kompetenzen stärken. Grundschule aktuell (9), 6-11.

Broc, L., Bernicot, J., Olive, T., Favart, M., Reilly, J., Quémart, P. & Uzé, J. (2013). Lexical spelling in children and adolescents with specific language impairment: Variations with the writing situation. Research in Developmental Disabilities 34, 3253–3266.

Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A.M., Bölte, J. & Böhl, A. (2011). The word frequency effect: a review of recent developments and implications for the choice of frequency estimates in German. In: Experimental Psychology 58(5), 412-424.

Campbell, J. P., DeBlois, P. B. & Oblinger, D. G. (2007). Academic analytics: A new tool for a new era. EDUCAUSE review, 42(4), 40.

CELEX (Center for Lexical Information) - Linguistic Data Consortium (1995). The CELEX Lexical Database, Release 2, German, Version 2.5.

Clow, D. (2012). The learning analytics cycle: closing the loop effectively. In Buckingham Shum, S., Gasevic, D. & Ferguson, R. (Eds.) Proceedings of 2nd International Conference on Learning Analytics & Knowledge (LAK'12), New York, USA, 134-138.

Duval, E. (2010). Attention Please! Learning Analytics for Visualization and Recommendation. In Proceedings of LAK11: 1st International Conference on Learning Analytics and Knowledge 2011, Retrieved October 8, 2015 from https://lirias.kuleuven.be/bitstream/123456789/315113/1/la2.pdf

Ebner, M., Taraghi, B., Ebner, M., Aspalter, C., Biermeier, S., Edtstadler, K., Gabriel, S., Goor, G., Gros, M., Huppertz, A., Martich, S., Steinhauer, N., Ullmann, M. & Ziegler, K. (2015a). Design für eine Plattform zum Schreibenlernen im Grundschulalter. In Rathmayer, S. & Pongratz, H. (Eds.) Proceedings of DeLFI Workshops 2015 co-located with 13th e-Learning Conference of the German Computer Society (DeLFI 2015) Munich, Germany, September 1st, 2015, 118-122.

Ebner, M., Taraghi, B., Saranti, A. & Schön, S. (2015b). Seven features of smart learning analytics-lessons learned from four years of research with learning analytics. eLearning papers 40, 51–55.

Edtstadler, K. (in press). Qualitative Fehleranalyse im Schriftspracherwerb. Kritik und Kriterien. In Lindner D., Gabriel S. & Beer R. (Eds.) Dialog Forschung – Tag der Forschung 2015, Lit-Verlag.

Fuhrhop, N. (2011). System der Getrennt- und Zusammenschreibung. In Bredel, U. & Reißig, T. (Eds.) Weiterführender Orthographieerwerb. Schneider Verlag Hohengehren, 107-128.

Flor, M. & Futagi, Y. (2012). On using context for automatic correction of non-word misspellings in student essays. In Proceedings of the Seventh Workshop on Innovative Use of NLP for Building Educational Applications, Association for Computational Linguistics, 105-115.

Fröhler, H. (2002). Fernitzer Grundwortschatz 2002. Wortliste für die Grundschule. Ausgabe 2002. Wien: Eigenverlag.

Greller, W. & Drachsler, H. (2012). Translating learning into numbers: A generic framework for learning analytics. Journal of Educational Technology & Society, 15(3), 42-57.

Government of South Australia. Department for Education and Child Development, issuing body. (2011). Spelling: from beginnings to proficiency: a spelling resource for planning, teaching, assessing and reporting on progress, [Adelaide, South Australia]: Department for Education and Child Development. Retrieved October 8, 2015 from http://www.decd.sa.gov.au/literacy/files/pages/Programs%20and%20 Resources/Spelling_resource_FINAL.pdf

Herné, K.-L. & Naumann, C.L. (2002). Aachener Förderdiagnostische Rechtschreibfehler-Analyse, Alfa-Zentaurus.

Holzinger, A., Errath, M., Searle, G., Thurnher, B. & Slany, W. (2005). From extreme programming and usability engineering to extreme usability in software engineering education (XP+ UE→ XU). In Computer Software and Applications Conference, 2005. COMPSAC 2005. 29th Annual International. (2), IEEE, 169-172.

Høien, T. & Lundberg, I. (2000). Dyslexia: From Theory to Intervention. Kluwer.

Katz, L. & Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. Haskins Laboratories Status Report on Speech Research, 147-160. Retrieved October 8, 2015 from http://www.haskins.yale.edu/sr/SR111/SR111_11.pdf

Khalil, M. & Ebner, M. (2015). Learning Analytics: Principles and Constraints. In Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications, EdMedia 2015. AACE Waynesville, NC, USA, 1326-1336.

Klees, M. (1989). Developmental Dyslexia in French. In Aaron, P.G. & Malatesha Joshi, R. (Eds.) Reading and Writing Disorders in Different Orthographic Systems. Kluwer Academic Publishers, 137-142.

Klicpera, Ch., Schabmann A. & Gasteiger-Klicpera, B. (2003). Legasthenie. Reinhardt.

Küttel, H. (2007). Probleme des Erwerbs der Orthographie. In Nerius, D. (Ed.) Deutsche Orthographie. Georg Olms, 417-451.

Landerl, K. & Thaler, V. (2006). Reading and spelling acquisition and dyslexia in German. In Joshi, R.M. & Aaron, P.G. (Eds.) Handbook of Orthography and Literacy, Lawrence Erlbaum, 121-134.

Liebal, J. & Exner, M. (2011). Usability für Kids. Springer Fachmedien Wiesbaden.

May, P. (2010). Hamburger Schreib-Probe 1-9. Manual. Klett.

Meinhold, G. & Stock, E. (2007). Phonem-Graphem-Beziehungen in deutschen Wörtern. In Nerius, D. (Ed.) Deutsche Orthographie. Georg Olms, 108-128.

Naumann, C. L. (2008). Zur Rechtschreibkompetenz und ihrer Entwicklung. In Bremerich-Vos, A., Granzer, D. & Köller, O. (Eds.) Lernstandbestimmung im Fach Deutsch, Beltz, 134-159.

Nerius, D. (2007). Deutsche Orthographie, Georg Olms.

Neuhold, B. (2013). Learning Analytics-Mathematik Lernen neu gedacht, BoD–Books on Demand.

Piety, P. J. (2013). Assessing the educational data movement, Teachers College Press.

Risel, H. (2008). Arbeitsbuch Rechtschreibdidaktik, Schneider Verlag Hohengehren.

Schön, M., Ebner, M. & Kothmeier, G. (2012). It's just about learning the multiplication table. In Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, ACM, 73-81.

Siemens, G. (2012). Learning analytics: envisioning a research discipline and a domain of practice. In Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, ACM, 4-8.

Siemens, G. & Long, P. (2011). Penetrating the Fog: Analytics in Learning and Education. EDUCAUSE review, 46(5), 30.

Snowling, M. J. (1989). Developmental Dyslexia: A Cognitive Developmental Perspective. In Aaron, P.G. & Joshi, R.M. (Eds.) Reading and Writing Disorders in Different Orthographic Systems, Kluwer Academic Publishers, 1-23.

Sommer Stumpenhorst, N. (2012). Rechtschreibkompetenz entwickeln. Retrieved October 9, 2015 from http://www.rechtschreib-werkstatt.de/rsl/gk/html/kompetenz.html

Tacke, G. (2008). Rechtschreibüben als Projektunterricht für die Klassenstufen 4 bis 8. Retrieved October 9, 2015 from http://www.cyberkidz.de/pdf/Projektunterricht-worter-5-7.pdf

Thelen, T. 2010. Automatische Analyse orthographischer Fehler von Schreibanfängern. PhD-thesis: University of Osnabrück. Retrieved November 8, 2015 from https://repositorium.uni-osnabrueck.de/bitstream/urn:nbn:de:gbv:700-201006096307/1/thesis_thelen.pdf

Thomé, G. & Thomé, D. (2014). Oldenburger Fehleranalyse für die Klassen 3-9, Instrument und Handbuch, Institut für sprachliche Bildung.

Taraghi, B., Saranti, A., Ebner, M., Müller, V. & Großmann, A. (2015a). Towards a Learning-Aware Application Guided by Hierarchical Classification of Learner Profiles, Journal of Universal Computer Science, 21(1), 93-109.

Taraghi, B., Frey, M., Saranti, A., Ebner, M., Müller, V. & Großmann, A. (2015b). Determining the Causing Factors of Errors for Multiplication Problems. In Ebner, M., Erenli, K., Malaka, R., Pirker, J., Walsh, A. (Eds.) Immersive Education, Springer, 27-38.

Tsesmeli, S. N. & Seymour, P.H.K. (2006). Derivational morphology and spelling in dyslexia. Reading and Writing, 19(6), 587–625.

Valtin, R. (1989). Dyslexia in the German Language. In Aaron, P.G. & Joshi, R.M. (Eds.) Reading and Writing Disorders in Different Orthographic Systems, Kluwer Academic Publishers, 119-135.