DIGITAL CLASSICS ONLINE

Building a Repository of Exercises for Learning Latin

Konstantin Schulz

Abstract: This study introduces quality criteria and a reference implementation of an exercise repository for Latin language exercises, with a special focus on vocabulary. The repository is supposed to be easily accessible to people with no prior knowledge of corpus linguistics or natural language processing. Teachers in high schools can generate exercises themselves, which should be fully customizable and adaptive with regard to the learners.
To facilitate the process of creating new exercises from ancient texts, additional linguistic information is needed. For instance, a Keyword-In-Context analysis enables teachers to investigate usage patterns for single words by looking at visualizations of morphological, syntactic and lexical phenomena.
Besides, exercises need to be findable and accessible. To achieve this, a public repository with an underlying database was created, so exercises can be stored and queried according to their relevant metadata, e.g., vocabulary, textual complexity or interaction type. The repository can be used by teachers to retrieve, modify and try out exercises developed by their fellow pedagogues. In this way, didactic efforts can be shared and built upon not just within the same school, but within the whole country, in many cases even internationally.

Introduction

This study seeks to define quality criteria and develop a reference implementation for a repository of language exercises. The exercises were used for research funded by the German Research Foundation in the CALLIDUS[1] project led by Malte Dreyer, Stefan Kipf and Anke Lüdeling. Consequently, the repository is tightly integrated into the project's software Machina Callida.[2] This application serves as a combined platform for various tools that help teachers of Latin. Previous publications have highlighted other parts of the platform, such as the generation of corpus-based exercises for learning management systems,[3] or using artificial intelligence to support this curation process.[4]

The present study builds on these papers to introduce novel functionality that has not been described previously: a searchable repository with extensive metadata on born-digital, interactive exercises for learning Latin.[5] The aim is to make existing curated materials accessible to a broader public by allowing people to filter them by, e.g., author, work or interaction type. The selection of relevant metadata needs

1 Project number 316618374, accessible at https://hu.berlin/callidus (Last access 11.01.2021).

2 Accessible at https://korpling.org/mc/ (Last access 11.01.2021, public server) and https://scm.cms.hu-berlin.de/callidus/machina-callida (Last access 11.01.2021, source code).

3 Beyer / Schulz (2020).

4 Schulz et al. (2020).

5 The data model was specified using OpenAPI (https://swagger.io/specification/ [Last access 11.01.2021]) and is available at https://scm.cms.hu-berlin.de/callidus/machina-callida/-/blob/master/mc_backend/openapi_models.yaml#L119-188 (Last access 11.01.2021).

to be reflected systematically, which is why the following chapters will make ample use of research literature on corpus linguistics, second language acquisition (esp. Latin pedagogy) and software development. The perspective is therefore interdisciplinary.

Technically, the exercise data is hosted in the cloud services of Humboldt University in Berlin and maintained by the CALLIDUS project. However, access to the data is granted through a public interface, which is represented by a REST API,[6] and by a user-friendly web application.[7] Using the REST API, users can create new exercises and share them with others, thereby contributing to a digital community of Latin teachers. We believe that this combination of reliable technical infrastructure, strict corpus-based approach to language acquisition, and emphasis of collaboration, is a unique characteristic of our learning platform for Latin pedagogy.

In the following, a short analysis of shortcomings in current Latin pedagogy will pave the way for a general estimation of quality criteria for Latin vocabulary exercises. After that, a closer look at evaluation and feedback will reveal important challenges for our system. Finally, we will summarize the major features of our implementation and point to current weaknesses that need to be addressed by future research.

Current State of Learning Latin Vocabulary

In the context of German high schools, various stakeholders are involved in the quality assurance of teaching Latin: students, teachers, researchers, (textbook) publishers and many more. During the last decades, there have been a few efforts to optimize the basic vocabulary for learners at various degrees of proficiency,[8] e.g., by reducing the amount[9] or improving the selection of words to be learned.[10] This discussion about core vocabularies is certainly inspired by the well-known phenomenon of students not being able to understand authentic Latin literature after the first few years of language learning.[11] Such problems arise out of deficient vocabulary training, as can be seen from textbooks focusing on single words instead of meaningful contexts,[12] and repeating important terms too rarely.[13] Therefore, this paper seeks to investigate quality criteria for Latin vocabulary exercises and ways to implement them prototypically in a digital learning environment.

Current vocabularies rely heavily on cross-lingual word equations,[14] which may be suitable for learners at the very beginning,[15] but not at the intermediate and advanced stages because, by this point, students need to develop semantic connections in their mental lexicon.[16] Otherwise, they cannot deal with im-

6 Accessible at https://korpling.org/mc-service/mc/api/v1.0/ui/ (Last access 11.01.2021).

7 Accessible at https://korpling.org/mc/exercise-list (Last access 11.01.2021).

8 Freie und Hansestadt Hamburg, Behörde für Bildung und Sport (2004), 10; Robillard et al. (2014), 2.

9 Schirok (2010), 17.

10 Utz (2000), 146.

11 Schibel (2013), 115.

12 Waiblinger (1998), 13; Siebel (2011), 127.

13 Van de Loo (2016), 136.

14 Van de Loo (2016), 140.

15 Crossley et al. (2010), 56.

16 Crossley et al. (2010), 70.

portant aspects of lexical competence, such as collocations, synonymy or derivation,[17] the latter of which has been identified as particularly important for teaching Latin.[18]

Still, many examples of Latin vocabulary software succumb to the temptation of offering decontextualized equations of cross-lingual form-meaning links: This is true for the ‚Vocabulary Drill' in the Wheelock Latin Exercises,[19] for the ‚Vocabulary Handouts' at The Latin Library,[20] and for flashcard applications like Anki.[21]

Those very same form-meaning links are also a dominant feature in one of the most influential basic vocabularies for Latin in Germany, the so-called Bamberg Vocabulary.[22] However, even if we take the high priority of this lexical representation as granted, we are still facing other problems: The Bamberg Vocabulary has been constructed using a closed-source corpus, i.e., the text collection and the formulae applied to it have not been published. In general, we know the authors and works that it was constructed from, but not the exact text passages or text editions. Besides, it is unclear which words were classified as proper names and removed during preprocessing.[23] Therefore, we have too little information to analyze its quality in detail and we have to extrapolate it from vague hints in the corresponding journal paper. In any case, a well-founded estimation of vocabulary size and selection has to be transparent to meet the requirements of serious educational research, especially in times of the generally acknowledged FAIR data principles.[24]

Upon closer examination of the purpose of vocabulary training, research on the didactics of Latin tends to emphasize translation as the core activity in classes.[25] Translation, however, is a highly complex process that involves more than the memorization of form-meaning links: It can be used to identify cultural peculiarities and to make typological comparisons between different languages.[26] For such use cases, simple equations of single words are rather unsuitable. Instead, they can merely act as a preliminary stage for higher-level tasks, e.g., identifying idiomatic multiword expressions,[27] which need special strategies for adequate translation. One important factor in this regard is contextualization, which also facilitates the identification and linguistic analysis of homonymy, polysemy and several other semantic obstacles.[28] This is in line with the basic assumption of distributional semantics, i.e. the context of a word constitutes its meaning.[29] Additionally, it also necessitates low-level morphosyntactic tasks where learners are supposed to highlight specific phenomena in a text (cf. Fig. 1).

17 González-Fernández / Schmitt (2020), 483.

18 Daum (2016), 15.

19 Accessible at https://web.uvic.ca/hrd/latin/wheelock/ (Last access 11.01.2021).

20 Accessible at http://www.thelatinlibrary.com/101/ (Last access 11.01.2021).

21 Accessible at https://apps.ankiweb.net/ (Last access 11.01.2021). An example for Latin is the DCC Core Latin collection at https://ankiweb.net/shared/info/180623737 (Last access 11.01.2021).

22 Utz (2000).

23 Utz (2000), 152.

24 Wilkinson et al. (2016).

25 Daum (2016), 76; Große (2015), 191.

26 Laviosa (2014), 42.

27 Rayson et al. (2010), 2.

28 Gardner (2007), 251; Webb (2008), 238; Hagiwara et al. (2009), 556; Helm (2009), 97; Gries / Wulff (2013), 348; Herbelot / Ganesalingam (2013), 443.

29 Roller et al. (2014), 1025.

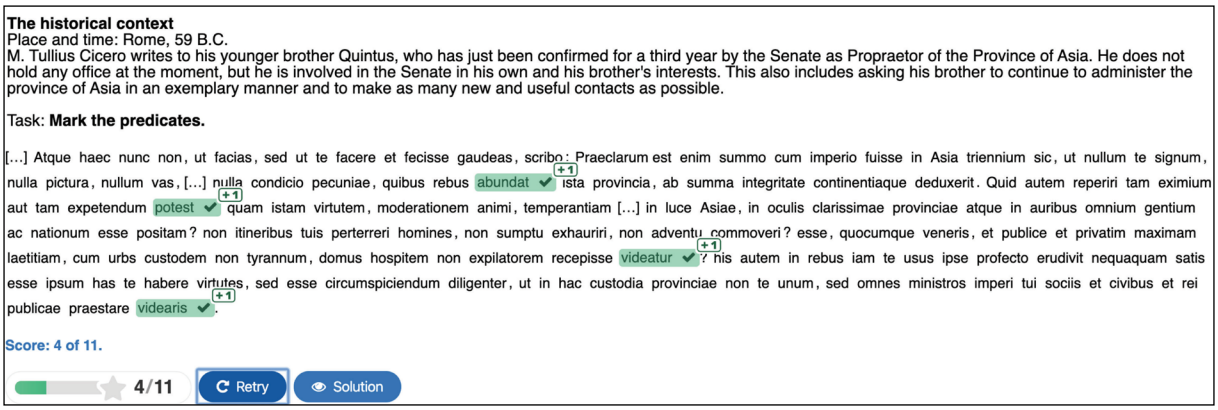

Fig. 1: Text-based exercise for the identification of a morphosyntactic phenomenon, created with H5P.[30]

Designing Digital Exercises for Latin Language Learning

In the following, we will define several quality criteria for exercises, but not in every single detail. Instead, the given example will be explained and analyzed to shed light on the curation process that is linked to the exercise repository. We will particularly highlight the role of contextualization for language acquisition and how this can be implemented in a digital environment. In Fig. 1, students are provided with a heading, a general introduction, a task description, a Latin text, some feedback and control elements. The heading and the introduction are supposed to provide a semantic embedding by describing the historical circumstances of what is portrayed in the Latin text. This is important even for morphosyntactic recognition tasks because previous knowledge of a text's author, topic or background is beneficial for further linguistic analyses.[31] If learners have a basic semantic knowledge (e.g., of the ancient world), they will be able to understand Latin texts more easily,[32] and thus provide a better interpretation or translation. The same correlation also applies to machines when they try to understand natural language.[33] Therefore, the positive influence of prior semantic knowledge on language learning seems to be universal for both humans and machines. This basic interaction is in line with constructivist approaches to learning that emphasize the role of previous knowledge,[34] as well as frame semantics which takes cultural knowledge as the starting point for every language learning process.[35]

The task description should encourage learners to interact with the Latin text. The text is to be presented in a digital format where every word can be clicked on, resulting in its selection for meeting the presented challenge. To assess a learner's performance, the internal digital representation of the text has to be enriched with linguistic annotations, such as part of speech and dependency relation. This way, a learner's selection can be compared to predefined linguistic criteria. In this example, the task is to find predicates, which are defined in Latin grammar books as inflected verb forms that act as the root in the dependency tree of a sentence.[36] However, most of the established linguistic ontologies for describing

30 Joubel (2018). See also https://h5p.org/ (Last access 11.01.2021), where users can create their own digital interactive exercises using a web application. Such exercises may include advanced features like feedback, hints, learning analytics and various types of media.

31 Harrison (2010), 9; Mondahl / Razmerita (2014), 341.

32 Pinkal (1993), 427–428; Fuchs et al. (2015), 211.

33 Bruni et al. (2014), 38; Punyakanok et al. (2008), 266.

34 Mvududu / Thiel-Burgess (2012), 110.

35 Atzler (2011), 61.

36 Menge et al. (2009), 311.

dependency grammar, such as Universal Dependencies,[37] are not entirely congruent with those used in the Classics.[38] Therefore, a domain-specific language has to be employed to provide a mapping between the two perspectives. In this exercise, the linguistic annotations of part of speech and dependency relation are queried (*Look for a word having the part of speech VERB and serving as root in a dependency tree*) in the background. At the same time, learners are asked to mark all predicates in the text. The comparison of results between both descriptions can be used as an indicator of a learner's performance.

The text's form is determined, to some end, by its corresponding edition. In this case, the Perseus Digital Library was used. The underlying critical text editions do not correspond to the state of the art from a philological perspective, but they conform to the principles of FAIR data and offer a uniform access interface through the Canonical Text Services.[39]

The text's content must be meaningful for learners. In textbooks, this is usually addressed by presenting made-up stories about ancient families.[40] When reading authentic Latin literature, though, the focus is shifted towards passages that cover different curricular standards.[41] Therefore, the text in Fig. 1 deals with the administration of Roman provinces, while still somewhat adhering to the family theme by presenting written communication between two brothers. Besides, it allows learners to improve their form recognition skills in a context-based manner: If they encounter a morphologically ambiguous ending, other words from the same paragraph can be used to eliminate at least some of the possible options.

By building every exercise from ancient literature, i.e., authentic texts written by native speakers, an exercise repository can compensate for the lack of native(-like) language input for learners of historical languages. This introduces the side benefit of highly specialized vocabulary training for single texts, authors or genres. Such a major focus on authentic L2 content in didactic materials is relatively unusual for the teaching of historical languages, which often relies heavily on the learners' L1 for communication in class.[42]

Upon completing an exercise, a status bar together with a twofold numerical representation acts as feedback for the learner's performance. Besides, the given answers are marked visually depending on their quality (green for correct, red for incorrect answers). Solutions that have not been found (i.e., false negatives) remain hidden by default, but can be shown on demand. Retries are possible until the exercise has been completed successfully, but can be disabled entirely. All in all, the quality criteria that we may deduce for the repository can be summarized as follows:

- semantic metadata, e.g., in headings and introductions
- authentic texts created by native speakers
- high quality of digital text editions
- linguistic annotations
- domain-specific language for the interface
- keeping track of learners' actions for the assessment of their language proficiency
- feedback
- control elements, e.g., viewing solutions or retrying old exercises

37 Nivre et al. (2017).

38 Kühner / Stegmann (1914); Menge (1914); Menge et al. (2009).

39 Tiepmar et al. (2014).

40 Stratenwerth (2012), 264.

41 Senatsverwaltung für Bildung, Jugend und Sport Berlin (2006), 9–21.

42 Fuhrmann (2003), 10; Große (2015), 191–202; Kuhlmann (2019), 73.

Furthermore, some studies suggest that interactive exercises provide higher motivation for learners.[43] However, this can be due to the novelty effect of technology-based teaching methods,[44] which are still rather uncommon in Latin pedagogy. Besides, even tasks like in Fig. 1 with their rather basic interaction design must not be underestimated because their interlinked view on vocabulary forces learners to tap into a complex combination of available information. Therefore, they have to be trained to study patterns in language use from various perspectives (cf. Fig. 2). The Keyword-In-Context view offers the possibility to study specific phenomena (e.g., usage of pronouns) in a structured and focused manner, integrating morphological, syntactic and semantic patterns.[45] Such visualizations can quickly become very complex,[46] which is why the example in Fig. 2 is restricted to just a few layers of annotation and contains additional formatting for further clarification (alignment of text passages, colouring, regular geometrical shapes).

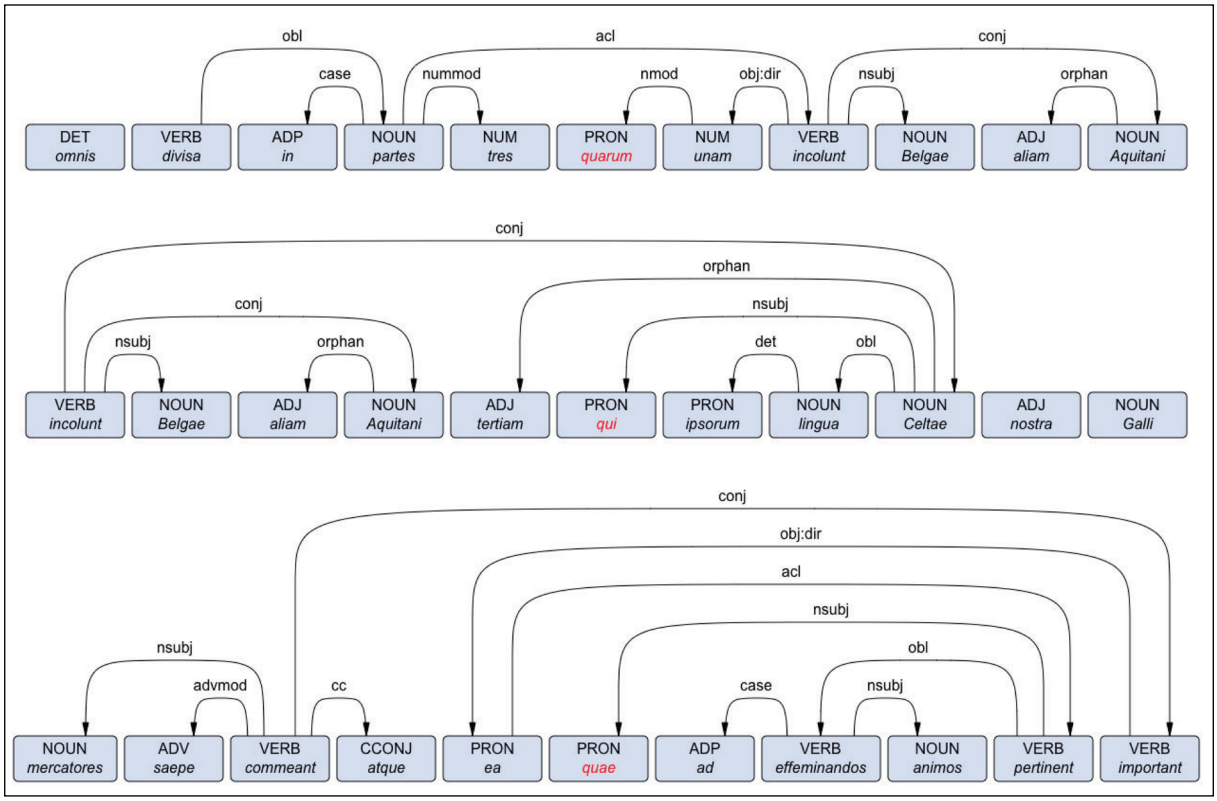

Fig. 2: Keyword-In-Context analysis.[47]

Evaluation

When new lexical knowledge has been acquired, it is notoriously difficult to diagnose that improvement. Traditionally, teachers use lists of word equations, i.e., translations of single words from the foreign to

43 Harecker / Lehner-Wieternik (2011), 5.

44 Merchant et al. (2014), 33.

45 Helm (2009), 97.

46 Fischl / Scharl (2014), 194.

47 The visualization was created using CONLLU Viewer (Kleiweg [2020]). It shows parts of speech and dependency relations for three partly overlapping text passages from Caesar's Gallic War, 1.1.1–1.1.3. Each passage is centred around a pronoun, which is highlighted in red. Arrows indicate the direction of a dependency relation from head to tail. The labels refer to the Universal Dependencies tag set. Parts of speech are given for each word in the upper part of a grey box.

the native language.[48] Instead of *native* language, it is more accurate to refer to it as the language in which the *lessons are conducted* because it is usually not the first language for every single learner in terms of acquisition sequence. This indicates a major issue with teaching Latin (at least in Germany): The heavy focus on the German language is usually depicted as beneficial for native speakers[49] as well as second language learners,[50] but it also systematically discriminates against the latter group by relying on the German language for testing purposes. Thus, at least for the diagnosis of lexical knowledge, we should avoid German elements because the separation of comprehension and translation is important for giving differentiated feedback.

In the oral domain (e.g., in teaching modern foreign languages), such abstractions are already present in existing diagnostic tools like the Toolbox Picture Vocabulary Test (TPVT):[51] Participants listen to tape-recorded words and, after each one of them, choose from 4 possible pictures the one that depicts the word's meaning most accurately. In this manner, the language barrier introduced by translation is eliminated, giving way to a more direct estimation of lexical knowledge. Unfortunately, this is hardly applicable to Latin because historical languages are not learned for the purpose of oral communication.[52] A transfer of the TPVT to the Latin domain would therefore need to provide written stimuli. Moreover, since the perceived distance to the target culture (i.e., the Roman Empire) is comparatively large,[53] some concepts may not be easy to depict and convey in an accurate manner. Therefore, such tools are somewhat limited, but still, their basic principles are valuable for designing new evaluations in the teaching of historical languages. These principles also include the internal differentiation of difficulty levels with respect to a learner's current proficiency, in order to avoid floor and ceiling effects.[54] Moreover, individual items in a test should not be weighted equally, but according to their workload and cognitive complexity: An item that asks me to translate a whole text passage will be more challenging than the morphological analysis of a single word. This distinction becomes even more obvious when the general focus of the test is shifted from form-meaning links to reading comprehension,[55] which suggests itself given the particular emphasis on the reception of literature in teaching Latin.

Another problem is the comparatively small learning input between the two tests. In modern language teaching, students can rely on oral practice for accelerated acquisition since it is much faster than written communication, especially in historical languages like Latin.[56] This lack of repetition and intensity has to be compensated by relying on sophisticated educational designs that integrate psychological concepts like the mental lexicon and spreading activation:[57] Thematically related words should be learned (and possibly tested) together. This also implies the rejection of alphabetical word lists for educational purposes.

Finally, in times of blended learning and e-assessment, digital test tools are becoming increasingly popular. They give the impression of objectivity, (social) justice, reliability and efficiency. However, there are many hidden weaknesses in traditional testing that now become obvious in more formalized, compu-

48 Carter (1997).

49 Große (2015), 202.

50 Siebel (2017), 177–178.

51 Gershon et al. (2013), 54.

52 Siebel (2017), 18.

53 Schauer (2019), 182.

54 Sparrow et al. (2005), 290; Gershon et al. (2013), 56.

55 Schmitt (2014), 950.

56 Schirok (2010), 13; Daum (2016), 76.

57 Bruza et al. (2009), 362/364.

ter-assisted settings. One of them is the lack of a well-defined horizon of expectations for the semantic and morphological parts of translation tasks:[58]

- Which translations (or paraphrases etc.) are appropriate representations of a given target term or concept?
- How do we (consistently) distinguish careless mistakes from a more profound lack of knowledge?
- How do we handle definiteness when translating between language pairs where one part has articles while the other does not?

Recent studies point towards the importance of context and a thorough understanding of its underlying semantics as a prerequisite for adequate translation.[59] This assessment goes beyond the traditional design of vocabulary training, where context was almost entirely eliminated. One approach to reintroduce this complexity in computational settings is a branch of Artificial Intelligence named representation learning, which tries to model each word's semantics by its common textual co-occurrences with other words.[60] However, many special cases are still hard to cover in such frameworks, e.g., multiword expressions.

Feedback

Regardless of whether the evaluation of lexical competence can be automated successfully, we also have to face the challenge of providing high-quality feedback. Usually, this is done in a binary fashion (i.e., correct/incorrect response), with explicit measurements (e.g., a score to be achieved) and a delayed communication of results (e.g., after a few days). For modern languages, there is additional implicit feedback from conversational exercises, e.g., dialogues.[61] Unfortunately, this valuable source of corrective input is mostly unavailable for Latin because of the strong focus on reading. Nevertheless, the same general quality criteria apply: Feedback should be immediate,[62] like a scaffolding,[63] and adapted to a learner's zone of proximal development.[64] This way, students do not just see superficial scores, but a detailed explanation of what they did wrong and what the smallest next step in the right direction might be. Unfortunately, such requirements are hard to meet in both face-to-face and computational settings because they demand a lot of time and/or complex modelling.

A basic, but crucial example consists in the classification of errors: If we do not distinguish between various deviations related to form and meaning,[65] we will fail to give helpful feedback, thus having to fall back to simpler ways of scoring. Besides, a written indication of the locations and types of errors may not be sufficient to encourage corrections. Instead, multimodal feedback (e.g., using videos) may be employed to offer higher incentives for improvement.[66] Furthermore, learners are usually not just interested in their current performance on a single test item, but also on their development over time (cf.

58 Beatty (2013), 209.

59 Hummel (2010), 62.

60 Bengio et al. (2003), 1141; Mikolov et al. (2013), 2; Perez / Cuadros (2017), 51; Wiedemann et al. (2019), 2.

61 Ellis et al. (2006), 340–341.

62 Opitz et al. (2011), 7.

63 Finn / Metcalfe (2010), 959.

64 Shabani et al. (2010), 238.

65 Rudzewitz et al. (2017), 41.

66 Elola / Oskoz (2016), 71.

Fig. 3).[67] This kind of ipsative assessment helps them keep track of their progress and assume responsibility for their learning.

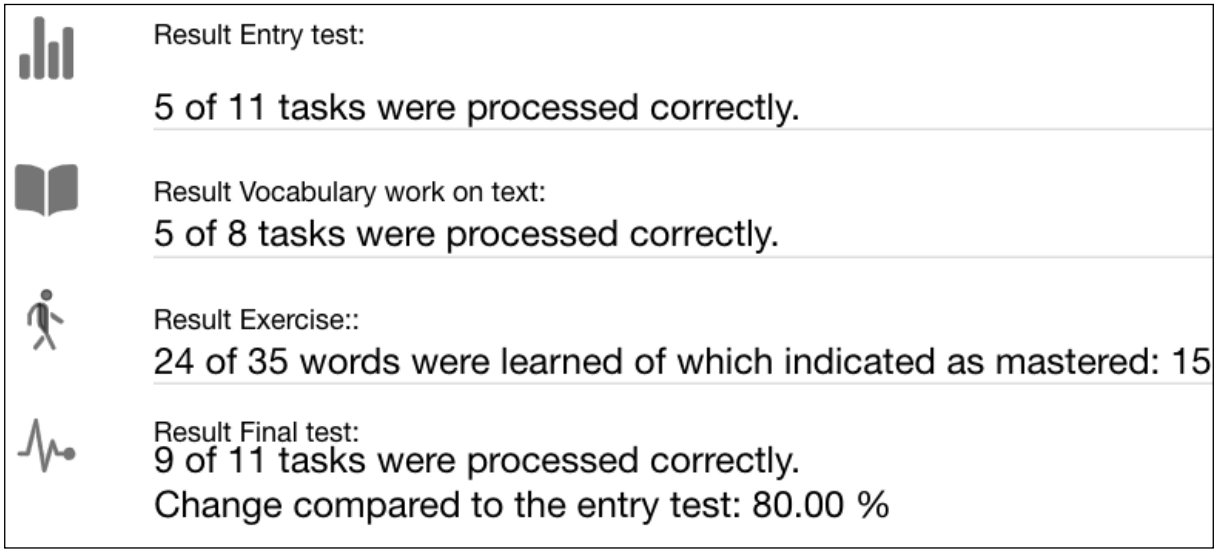

Fig. 3: Explicit summative feedback after a structured learning session, for ipsative assessment.[68]

In addition to this explicit kind of individual feedback, consumers of a well-designed vocabulary framework should also have access to criterion-referenced tests. These may include, among others, measurements of textual complexity and overlap with a target vocabulary.[69] This way, learners can quantify and compare their personal level of competence at various points in time. What is more, teachers may choose the most suitable upcoming exercises on an individual basis by browsing the repository of ready-made materials (cf. Fig. 3). Possible use cases for search in this repository include exercises

- of a preferred interaction type, which may correspond to a certain learning style, e.g., only cloze exercises.[70]
- for a specific author, in order to study the author-specific language use.[71]
- for a specific text passage or subcorpus, in order to prepare for a phase of close reading or exams.
- with a low text complexity for the transition from textbook to authentic literature.[72]
- with as few out-of-vocabulary words as possible, which would have to be supplemented with a gloss in exams or for novice learners.[73]

67 Univio / Pérez (2019), 158.

68 The four parts belong to a curated 45-minute digital session for Latin vocabulary training (accessible at https://korpling.org/mc/test [Last access 11.01.2021]). The Entry and Final tests were identical, thus enabling the learners to see their progress after a specific intervention (Vocabulary work on text and Exercise). Mastery of words in the Exercise part was indicated through self-assessment by the students using a checkbox.

69 Muccigrosso (2004), 422.

70 Schmid (2010), 169.

71 Devine / Stephens (2006), 452; Cordes (2020), 40–41.

72 Schibel (2013), 115.

73 Olimpi (2019), 86.

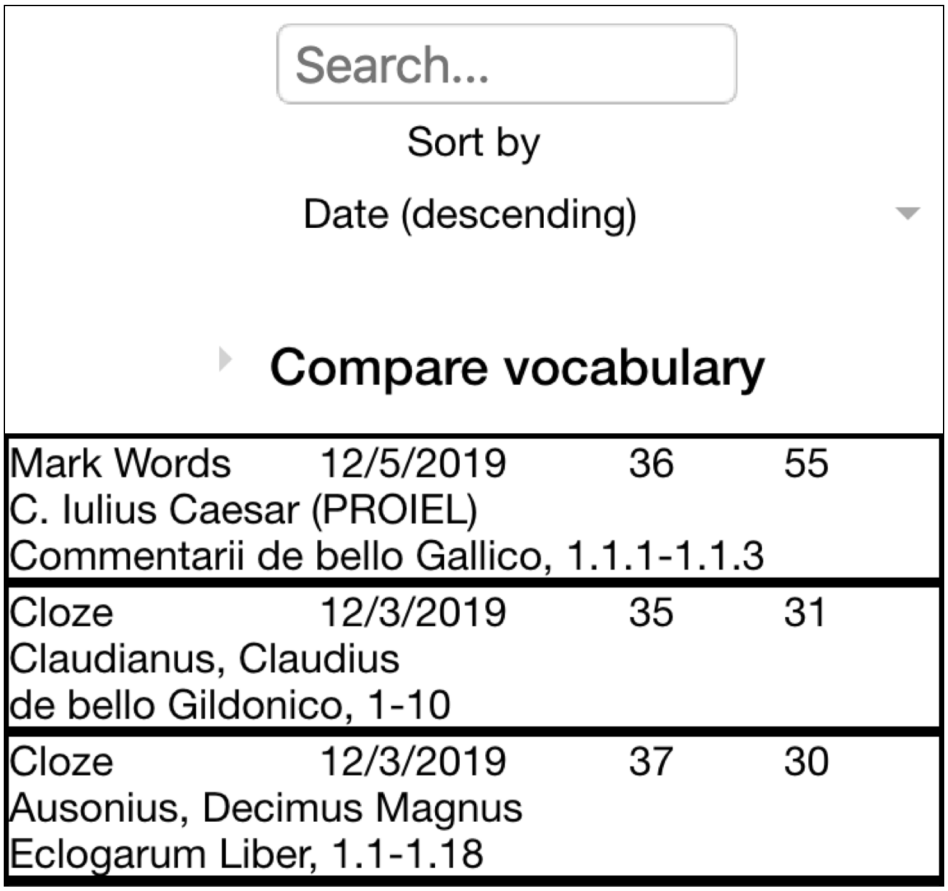

Fig. 4: The exercise repository offers various options for filtering: exercise type, date of last access, author, text passage.[74]

Conclusion

Many of the above-mentioned quality criteria for vocabulary learning and exercise repositories have already been considered in our implementation: Almost all vocabulary exercises involve contextualization, i.e., most words are not presented independently, but as part of a collocation, phrase, sentence or even a whole text passage. Depending on their design, multiple items may be combined into a longer sequence, e.g., with increasing levels of difficulty. Indicators of such difficulty are the familiarity with an exercise's vocabulary or the linguistic features of its base text. These are calculated automatically, so users can sort by them and easily compare various materials.

Moreover, the repository makes use of existing high-quality resources such as text editions, annotated corpora and frameworks for interactive digital exercises. Where linguistic information is missing, it tries to add them automatically. This fallback procedure is error-prone, particularly for complex syntactic annotations, thus decreasing the quality of the curation process and the entire repository.

While exercises also include explicit feedback (either immediately after a single exercise or after a longer period of learning), they do so only in a binary fashion. The correct results are shown and teachers may provide a general explanation, but it is not adaptive and thus not suitable to point learners in the right direction. This shortcoming is probably the most important aspect to consider in the development

74 Every item has measurements of text complexity and the percentage of known words as compared to a reference vocabulary (last two columns).

of future projects. Finally, it seems that the opportunities for individualization offered by a digital learning context seem to ask for an even stronger integration of ipsative assessment. Some of this is already present in the evaluation after the ready-made vocabulary unit, but additional visualizations and a more detailed tracking of learner results are necessary in order to provide a higher overall quality in the curation and reuse of lexical materials.

References

Atzler (2011): J. K. Atzler, Twist in the List: Frame Semantics as Vocabulary Teaching and Learning Tool, PhD Thesis, Austin: University of Texas, 2011, URL: https://repositories.lib.utexas.edu/bitstream/handle/2152/ETD-UT-2011-05-2752/ATZLER-DISSERTATION.pdf?sequence=1&isAllowed=y (Last access 11.01.2021).

Beatty (2013): K. Beatty, Teaching & Researching: Computer-Assisted Language Learning, Routledge, 2013, URL: http://ebook.stkip-pgri-sumbar.ac.id/ebook/bahasa/teaching-researching-computer-assisted-language-learning-pearson-education-esl/download (Last access 11.01.2021).

Bengio et al. (2003): Y. Bengio / R. Ducharme / P. Vincent / C. Jauvin, A Neural Probabilistic Language Model, Journal of machine learning research 3 Feb (2003), 1137–1155.

Beyer / Schulz (2020): A. Beyer / K. Schulz, CALLIDUS – Korpusbasierte, digitale Wortschatzarbeit im Lateinunterricht, in: F. Maier / S. Chronopoulos (eds.), Der Digital Turn in den Altertumswissenschaften, Propylaeum-eBooks, 149–167, 2020, URL: https://doi.org/10.11588/propylaeum.563 (Last access 11.01.2021).

Bruni et al. (2014): E. Bruni / N.-K. Tran / M. Baroni, Multimodal Distributional Semantics, J. Artif. Intell. Res. (JAIR) 49 (2014), 1–47.

Bruza et al. (2009): P. Bruza / K. Kitto / D. Nelson / C. McEvoy, Is There Something Quantum-like about the Human Mental Lexicon?, Journal of Mathematical Psychology 535 (2009), 362–377.

Carter (Aug. 1997): T. G. M. Carter, Latin Vocabulary Acquisition: An Experiment Using Information-Processing Techniques of Chunking and Imagery, English, Dissertation, University of North Texas, Aug. 1997, URL: https://digital.library.unt.edu/ark:/67531/metadc277583/m1/11/ (Last access 11.01.2021).

Cordes (2020): L. Cordes, Wenn Fiktionen Fakten schaffen. Faktuales und fiktionales Erzählen in den spätantiken Panegyrici Latini, Deutsch, in: D. Breitenwischer / H.-M. Häger / J. Menninger (Hrsgg.), Faktuales und fiktionales Erzählen II. Geschichte – Medien – Praktiken, Baden-Baden 2020, 31–56, URL: https://doi.org/10.5771/9783956505126-31 (Last access 11.01.2021).

Crossley et al. (2010): S. A. Crossley / T. Salsbury / D. S. McNamara, The Development of Semantic Relations in Second Language Speakers: A Case for Latent Semantic Analysis, Vigo International Journal of Applied Linguistics 7 (2010), 55–74.

Dascalu et al. (2017): M. A. Dascalu / G. S. Gutu / S. S. Ruseti / I. S. Cristian Paraschiv / P. Dessus / D. A. Mcnamara / S. A. Crossley / S. A. Trausan-Matu, ReaderBench: A Multi-Lingual Framework for Analyzing Text Complexity, in: É. Lavoué / H. Drachsler / K. Verbert / J. Broisin / M. Pérez-Sanagustín (Hrsgg.), Data Driven Approaches in Digital Education, Proc 12th European Conference on Technology Enhanced Learning, EC-TEL 2017, Tallinn, Estonia, September 12–15, 2017, Proceedings, Tallinn, Estonia 2017 606–609, URL: https://hal.archives-ouvertes.fr/hal-01584870 (Last access 11.01.2021).

Daum (2016): M. Daum, Wortschatz und Lehrbuch: Ein Kriterienkatalog für die Wortschatzkonzeption in Lateinlehrwerken, vol. 2, Ars Didactica. Marburger Beiträge zu Studium und Didaktik der Alten Sprachen, Propyläum-eBooks 2016, URL: https://doi.org/10.11588/propylaeum.609 (Last access 13.01.2021).

Devine / Stephens (2006): A. M. Devine / L. D. Stephens, Latin Word Order: Structured Meaning and Information, Oxford University Press, 2006, URL: https://books.google.de/books?hl=en%5C&lr=%5C&id=WY2Nhc3HY3sC (Last access 11.01.2021).

Ellis et al. (June 2006): R. Ellis / S. Loewen / R. Erlam, Implicit and Explicit Corrective Feedback and the Acquisition of L2 Grammar, Studies in Second Language Acquisition 282 (2006), 339–368.

Elola / Oskoz (2016): I. Elola / A. Oskoz, Supporting Second Language Writing Using Multimodal Feedback, Foreign Language Annals 491 (2016), 58–74.

Finn / Metcalfe (2010): B. Finn / J. Metcalfe, Scaffolding Feedback to Maximize Long-Term Error Correction, Memory & Cognition 387 (2010), 951–961.

Fischl / Scharl (2014): D. Fischl / A. Scharl, Metadata Enriched Visualization of Keywords in Context, in: Proceedings of the 2014 ACM SIGCHI Symposium on Engineering Interactive Computing Systems, 193–196, 2014, URL: https://dl.acm.org/doi/pdf/10.1145/2607023.2611451 (Last access 11.01.2021).

Freie und Hansestadt Hamburg, Behörde für Bildung und Sport (2004): Freie und Hansestadt Hamburg, Behörde für Bildung und Sport, Rahmenplan Alte Sprachen: Latein, Griechisch. Bildungsplan achtstufiges Gymnasium Sekundarstufe I, 2004, URL: http://epub.sub.uni-hamburg.de/epub/volltexte/2008/600/pdf/LATGRIE_Gy8.pdf (Last access 11.01.2021).

Fuchs et al. (2015): L. S. Fuchs / D. Fuchs / D. L. Compton / C. L. Hamlett / A. Y. Wang, Is Word-Problem Solving a Form of Text Comprehension?, Scientific Studies of Reading 193 (2015), 204–223.

Fuhrmann (2003): M. Fuhrmann, Bildungsziele im Wandel der Zeiten – und worauf soll es jetzt hinaus? Eine nüchterne Standortbestimmung, auch für Latein und Griechisch, Pegasus-Onlinezeitschrift 32 (2003), 1–11, URL: https://doi.org/10.11588/pegas.2003.2.35716 (Last access 11.01.2021).

Gardner (Apr. 2007): D. Gardner, Validating the Construct of Word in Applied Corpus-Based Vocabulary Research: A Critical Survey, Applied Linguistics 282 (2007), 241–265.

Gershon et al. (2013): R. C. Gershon / J. Slotkin / J. J. Manly / D. L. Blitz / J. L. Beaumont / D. Schnipke / K. Wallner-Allen / R. M. Golinkoff / J. B. Gleason / K. Hirsh-Pasek / M. J. Adams / S. Weintraub, NIH Toolbox Cognition Battery (CB): Measuring Language (Vocabulary Comprehension and Reading Decoding), Monographs of the Society for Research in Child Development 784 (2013), 49–69.

González-Fernández / Schmitt (2020): B. González-Fernández / N. Schmitt, Word Knowledge: Exploring the Relationships and Order of Acquisition of Vocabulary Knowledge Components, Applied Linguistics 41 (2020), 481–505.

Gries / Wulff (2013): S. T. Gries / S. Wulff, The Genitive Alternation in Chinese and German ESL Learners: Towards a Multifactorial Notion of Context in Learner Corpus Research, International Journal of Corpus Linguistics 183 (2013), 327–356.

Große (2015): M. Große, Pons Latinus: Latein als reflexionsbasierte Brückensprache im Rahmen eines sprachsensiblen Lateinunterrichts, in: E. M. F. Ammann / A. Kropp / J. Müller-Lancé (Hrsgg.), Herkunftsbedingte Mehrsprachigkeit im Unterricht der Romanischen Sprachen, vol. 17, Frank & Timme, 185–206, 2015, URL: https://www.peterlang.com/downloadpdf/title/64659 (Last access 18.01.2021).

Hagiwara et al. (2009): M. Hagiwara / Y. Ogawa / K. Toyama, Supervised Synonym Acquisition Using Distributional Features and Syntactic Patterns, Information and Media Technologies 42 (2009), 558–582.

Harecker / Lehner-Wieternik (2011): G. Harecker / A. Lehner-Wieternik, Computer-Based Language Learning with Interactive Web Exercises, ICT for Language Learning (2011), 1–5.

Harrison (2010): R. R. Harrison, Exercises for Developing Prediction Skills in Reading Latin Sentences, Teaching Classical Languages. An Online Journal of the Classical Association of the Middle West and South 21 (2010), 1–30.

Helm (May 2009): F. Helm, Language and Culture in an Online Context: What Can Learner Diaries Tell Us about Intercultural Competence?, Language and Intercultural Communication 92 (2009), 91–104.

Herbelot / Ganesalingam (2013): A. Herbelot / M. Ganesalingam, Measuring Semantic Content in Distributional Vectors, in: Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), 2013, 440–445, URL: https://www.aclweb.org/anthology/P13-2078 (Last access 11.01.2021).

Hummel (2010): K. M. Hummel, Translation and Short-Term L2 Vocabulary Retention: Hindrance or Help?, Language teaching research 141 (2010), 61–74.

Joubel (June 2018): A. Joubel, H5P. Create, Share and Reuse Interactive HTML5 Content in Your Browser, Joubel AS, Tromsø, Norway, URL: https://h5p.org/about-the-project (Last access 11.01.2021), June 2018.

Kleiweg (Sept. 2020): P. Kleiweg, Conllu Viewer. A Web-Based Viewer for Documents in the CoNLL-U Format, Computational Linguistics, University of Groningen, Sept. 2020, URL: https://github.com/rug-compling/conllu-viewer (Last access 11.01.2021).

Kuhlmann (2019): P. Kuhlmann, Sprachausbildung, Aufgabenformate und Übungsdidaktik im Lateinstudium, in: S. Freund / L. Janssen (Hrsgg.), Non ignarus docendi. Impulse zur kohärenten Gestaltung von Fachlichkeit und von Mehrsprachigkeitsdidaktik in der Lateinlehrerbildung, Bad Heilbrunn 2019, 66–78, URL: https://www.pedocs.de/volltexte/2019/16910/pdf/Freund_Janssen_2019_Non_ignarus_docendi.pdf (Last access 11.01.2021).

Kühner / Stegmann (1914): R. Kühner / C. Stegmann, Ausführliche Grammatik der lateinischen Sprache, 2. Teil: Satzlehre, vol. 1, Hannover 1914, URL: https://books.google.de/books?id=PIW7chvuopYC (Last access 11.01.2021).

Laviosa (2014): S. Laviosa, Translation and Language Education: Pedagogic Approaches Explored, London / N.Y. 2014, URL: https://books.google.de/books/?id=cp0QngEACAAJ&redir_esc=y (Last access 13.01.2021).

Menge (1914): H. Menge, Repetitorium der lateinischen Syntax und Stilistik: Ein Lernbuch für Studierende und vorgeschrittene Schüler, zugleich ein praktisches Repertorium für Lehrer, Teile 1–2, Wolfenbüttel 1914, URL: https://books.google.de/books?id=Mr0zAQAAMAAJ (Last access 11.01.2021).

Menge et al. (2009): H. Menge / T. Burkard / M. Schauer, Lehrbuch der lateinischen Syntax und Semantik, Darmstadt 2009.

Merchant et al. (Jan. 2014): Z. Merchant / E. T. Goetz / L. Cifuentes / W. Keeney-Kennicutt / T. J. Davis, Effectiveness of Virtual Reality-Based Instruction on Students' Learning Outcomes in K-12 and Higher Education: A Meta-Analysis, Computers & Education 70 (2014), 29–40.

Mikolov et al. (2013): T. Mikolov / K. Chen / G. Corrado / J. Dean, Efficient Estimation of Word Representations in Vector Space, arXiv preprint arXiv:1301.3781 (2013), 1–12, URL: https://arxiv.org/abs/1301.3781 (Last access 11.01.2021).

Mondahl / Razmerita (2014): M. Mondahl / L. Razmerita, Social Media, Collaboration and Social Learning–A Case-Study of Foreign Language Learning. Electronic Journal of E-learning 124 (2014), 339–352.

Muccigrosso (2004): J. D. Muccigrosso, Frequent Vocabulary in Latin Instruction, The Classical World 974 (2004), 409–433.

Mvududu / Thiel-Burgess (2012): N. Mvududu / J. Thiel-Burgess, Constructivism in Practice: The Case for English Language Learners, International Journal of Education 43 (2012), 108.

Nivre et al. (Nov. 2017): J. Nivre et al., Universal Dependencies 2.1: Morphologically and Syntactically Annotated Corpora of Many Languages, Nov. 2017, URL: https://hal.inria.fr/hal-01682188 (Last access 11.01.2021).

Olimpi (2019): A. Olimpi, Legere Discitur Legendo: Extensive Reading in the Latin Classroom, Journal of Classics Teaching 2039 (2019), 83–89.

Opitz et al. (2011): B. Opitz / N. K. Ferdinand / A. Mecklinger, Timing Matters: The Impact of Immediate and Delayed Feedback on Artificial Language Learning, Frontiers in Human Neuroscience 5 (2011), 1–9.

Perez / Cuadros (2017): N. Perez / M. Cuadros, Multilingual Call Framework for Automatic Language Exercise Generation from Free Text, in: Proceedings of the Software Demonstrations of the 15th Conference of the European Chapter of the Association for Computational Linguistics, 2017, 49–52, URL: https://www.aclweb.org/anthology/E17-3013.pdf (Last access 18.01.2021).

Pinkal (1993): M. Pinkal, Semantik, in: G. Görz (Hrsg.), Einführung in die Künstliche Intelligenz, Bonn 1993, 425–498.

Punyakanok et al. (2008): V. Punyakanok / D. Roth / W.-T. Yih, The Importance of Syntactic Parsing and Inference in Semantic Role Labeling, Computational Linguistics 342 (2008), 257–287.

Rayson et al. (Apr. 2010): P. Rayson / S. Piao / S. Sharoff / S. Evert / B. V. Moirón, Multiword Expressions: Hard Going or Plain Sailing?, Language Resources and Evaluation 441 (2010), 1–5.

Robillard et al. (2014): M. Robillard / C. Mayer-Crittenden / M. Minor-Corriveau / R. Bélanger, Monolingual and Bilingual Children with and without Primary Language Impairment: Core Vocabulary Comparison, Augmentative and alternative communication 303 (2014), 267–278.

Roller et al. (2014): S. Roller / K. Erk / G. Boleda, Inclusive yet Selective: Supervised Distributional Hypernymy Detection, in: Proceedings of COLING 2014 (The 25th International Conference on Computational Linguistics: Technical Papers), 2014, 1025–1036, URL: https://www.aclweb.org/anthology/C14-1097 (Last access 11.01.2021).

Rudzewitz et al. (May 2017): B. Rudzewitz / R. Ziai / K. De Kuthy / D. Meurers, Developing a Web-Based Workbook for English Supporting the Interaction of Students and Teachers, in: Proceedings of the Joint Workshop on NLP for Computer Assisted Language Learning and NLP for Language Acquisition, Gothenburg, Sweden May 2017, 36–46, URL: https://www.aclweb.org/anthology/W17-0305 (Last access 11.01.2021).

Schauer (2019): M. Schauer, Klasse Klassik: Latein im Klassenzimmer, in: K. Beuter / A. Hlukhovych / B. Bauer / K. Lindner / S. Vogt (Hrsgg.), Sprache und kulturelle Bildung: Perspektiven für eine reflexive Lehrerinnen- und Lehrerbildung und einen heterogenitätssensiblen Unterricht, vol. 9, Bamberg 2019, URL: https://fis.uni-bamberg.de/bitstream/uniba/46841/1/FLB9BeuterSpracheopusse_A3a.pdf (Last access 11.01.2021).

Schibel (2013): W. Schibel, Zur Aneignung Lateinischer Literatur und Sprache, Forum Classicum (2013), 113–124.

Schirok (2010): E. Schirok, Wortschatzarbeit, in: T. Doepner / M. Keip (Hrsgg.), Interaktive Fachdidaktik Latein, Göttingen 2010, 13–34, URL: https://static.onleihe.de/content/vandenhoeck/20141125/978-3-647-26411-0/v978-3-647-26411-0.pdf (Last access 11.01.2021).

Schmid (2010): E. C. Schmid, Developing Competencies for Using the Interactive Whiteboard to Implement Communicative Language Teaching in the English as a Foreign Language Classroom, Technology, Pedagogy and Education 192 (2010), 159–172.

Schmitt (2014): N. Schmitt, Size and Depth of Vocabulary Knowledge: What the Research Shows, Language Learning 644 (2014), 913–951.

Schulz et al. (2020): K. Schulz / A. Beyer / M. Dreyer / S. Kipf, A Data-Driven Platform for Creating Educational Content in Language Learning, in: Proceedings of the Conference on Digital Curation Technologies (QURATOR 2020 – Conference on Digital Curation Technologies), Berlin 2020, URL: http://ceur-ws.org/Vol-2535/paper_9.pdf (Last access 21.01.2021).

Senatsverwaltung für Bildung, Jugend und Sport Berlin (2006): Senatsverwaltung für Bildung, Jugend und Sport Berlin, Rahmenlehrplan für die gymnasiale Oberstufe. Gymnasien, Gesamtschulen mit Gymnasialer Oberstufe, Berufliche Gymnasien, Kollegs, Abendgymnasien. Latein, URL: https://www.berlin.de/sen/bildung/unterricht/faecher-rahmenlehrplaene/rahmenlehrplaene/mdb-sen-bildung-unterricht-lehrplaene-sek2_latein.pdf (Last access 11.01.2021), 2006.

Shabani et al. (2010): K. Shabani / M. Khatib / S. Ebadi, Vygotsky's Zone of Proximal Development: Instructional Implications and Teachers' Professional Development, English language teaching 34 (2010), 237–248.

Siebel (2011): K. Siebel, Lateinischer Wortschatz als Brücke zur Mehrsprachigkeit? Eine Durchsicht des Aufgabenspektrums aktueller Lateinlehrwerke, Pegasus-Onlinezeitschrift XII (2011), 102–132, URL: https://doi.org/10.11588/pegas.2011.1.35346 (Last access 11.01.2021)

Siebel (2017): K. Siebel, Mehrsprachigkeit und Lateinunterricht: Überlegungen zum lateinischen Lernwortschatz, vol. 4, Göttingen 2017, URL: https://books.google.de/books?hl=de&lr=&id=YOMsDwAAQBAJ (Last access 11.01.2021).

Sparrow et al. (Jan. 2005): S. S. Sparrow / T. M. Newman / S. I. Pfeiffer, 8 – Assessment of Children Who Are Gifted with the WISC-IV, in: A. Prifitera / L. G. Weiss / D. H. Saklofske (Hrsgg.), WISC-IV Clinical Use and Interpretation (Practical Resources for the Mental Health Professional), Jan. 2005, 281–298, URL: http://www.sciencedirect.com/science/article/pii/B9780125649315500098 (Last access 11.01.2021).

Stratenwerth (2012): D. Stratenwerth, Ziemlich grundsätzliche Überlegungen zur Konzeption von Lateinischen Lehrbüchern. Unter besonderer Berücksichtigung der ersten Lehrbuchtexte und ein paar konkrete Beispiele, Forum Classicum 2012 (2012), 264–270.

Tiepmar et al. (2014): J. Tiepmar / C. Teichmann / G. Heyer / M. Berti / G. Crane, A New Implementation for Canonical Text Services, in: Proceedings of the 8th Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities 2014 (LaTeCH), 1–8, URL: https://www.aclweb.org/anthology/W14-0601 (Last access 11.01.2021).

Univio / Pérez (2019): D. J. Univio / A. d. P. Pérez, Ipsative Assessment of Essay Writing to Foster Reflection and Self-Awareness of Progress, in: E. White / T. Delaney (Hrsg.), Handbook of Research on Assessment Literacy and Teacher-Made Testing in the Language Classroom, Hershey, PA 2019, 157–180, DOI: 10.4018/978-1-5225-6986-2.ch009, URL: https://www.igi-global.com/chapter/ipsative-assessment-of-essay-writing-to-foster-reflection-and-self-awareness-of-progress/217152 (Last access 11.01.2021).

Utz (2000): C. Utz, Mutter Latein und unsere Schüler – Überlegungen zu Umfang und Aufbau des Wortschatzes [BWS], Antike Literatur–Mensch, Sprache, Welt. Dialog Schule und Wissenschaft 34 (2000), 146–172.

Van de Loo (2016): T. Van de Loo, Wortschatzarbeit – Neuere Perspektiven und schulische Praxis, Pegasus-Onlinezeitschrift 16 (2016), 131–151, URL: https://doi.org/10.11588/pegas.2016.0.35254 (Last access 11.01.2021).

Waiblinger (1998): F. P. Waiblinger, Überlegungen zum Konzept des lateinischen Sprachunterrichts. Joachim Gruber zum 60. Geburtstag, Forum Classicum 1998 (1998), 9–19.

Webb (2008): S. Webb, The Effects of Context on Incidental Vocabulary Learning, Reading in a foreign Language 202 (2008), 232–245.

Wiedemann et al. (Oct. 2019): G. Wiedemann / S. Remus / A. Chawla / C. Biemann, Does BERT Make Any Sense? Interpretable Word Sense Disambiguation with Contextualized Embeddings, arXiv:1909.10430 [cs] (2019), 1–10, URL: https://arxiv.org/abs/1909.10430 (Last access 11.01.2021).

Wilkinson et al. (Mar. 2016): M. D. Wilkinson / M. Dumontier / I. J. Aalbersberg / G. Appleton / M. Axton / A. Baak / N. Blomberg / J.-W. Boiten / L. B. da Silva Santos / P. E. Bourne / J. Bouwman / A. J. Brookes / T. Clark / M. Crosas / I. Dillo / O. Dumon / S. Edmunds / C. T. Evelo / R. Finkers / A. Gonzalez-Beltran / A. J. G. Gray / P. Groth / C. Goble / J. S. Grethe / J. Heringa / P. A. C. 't Hoen / R. Hooft / T. Kuhn / R. Kok / J. Kok / S. J. Lusher / M. E. Martone / A. Mons / A. L. Packer / B. Persson / P. Rocca-Serra / M. Roos / R. van Schaik / S.-A. Sansone / E. Schultes / T. Sengstag / T. Slater / G. Strawn / M. A. Swertz / M. Thompson / J. van der Lei / E. van Mulligen / J. Velterop / A. Waagmeester / P. Wittenburg / K. Wolstencroft / J. Zhao / B. Mons, The FAIR Guiding Principles for Scientific Data Management and Stewardship, Scientific Data 3 (2016), 1–9.

Author contact information[75]

Konstantin Schulz

Humboldt-Universität zu Berlin
Unter den Linden 6
10099 Berlin
Tel: 030/2093 9720

E-Mail: schulzkx@hu-berlin.de

75 The rights pertaining to content, text, graphics, and images, unless otherwise noted, are reserved by the author. This contribution is licensed under CC BY 4.0.