

Building a Database of Exercises for Learning Latin^{*}

Konstantin Schulz
konstantin.schulz@hu-berlin.de

Humboldt-Universität zu Berlin, Germany

Abstract. CALLIDUS is an interdisciplinary project with researchers from the Classics, corpus linguistics and the IT department who are jointly trying to answer the following question: Can corpus-based acquisition of vocabulary add significant value to the way we learn Latin, compared to traditional methods? For our studies, we use approaches like data-driven and computer-assisted language learning. This enables us to dynamically derive meaningful exercises from text corpora, so vocabulary can be acquired in authentic contexts using our public software (<https://korpling.org/mc/>). Its exercise generator is designed to be easily accessible to people without a previous knowledge of corpus linguistics or natural language processing. Thus, teachers in high schools are able to produce exercises independently and according to their own needs.

However, we also want the exercises to be findable and reusable. To achieve this, we have designed a public repository with an underlying PostgreSQL database, 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 (or students) 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.

To facilitate the process of creating new exercises or enhancing existing ones, we also provide various kinds of additional information about texts, so people can decide more easily what to focus on. This is helpful not just for teaching, but also for research: The Keyword In Context analysis allows the investigation of usage patterns for single words by looking at visualizations of morphological, syntactic and lexical phenomena, much in the same way as a dynamic enhanced concordance.

Keywords: Language exercise · Language learning · Learning software · Exercise repository

^{*} This research is funded by the German Research Foundation (project number 316618374) in a project lead by Malte Dreyer, Stefan Kipf and Anke Lüdeling.

1 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 infrequent efforts to optimize the basic vocabulary [13, p. 10; 31, p. 2] for learners at various degrees of proficiency, e.g. by reducing the amount [37, p.17] or improving the selection of words to be learned [46, p.146]. This discussion 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 [36, p.115]. Such problems arise out of deficient vocabulary training, as can be seen from textbooks focusing on single words instead of meaningful contexts [41, p.127] and repeating important terms too infrequently [47, p.136]. 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 [47, p.140], which may be suitable for learners at the very beginning [6, p.56], but not at the intermediate and advanced stages because, by this point, students need to develop semantic connections [6, p.70] in their mental lexicon. Otherwise, they cannot deal with important aspects of lexical competence, such as collocations, synonymy or derivation [16, p.3], the latter of which has been identified as particularly important for teaching Latin [7, p.15].

However, even if we take the high priority of form-meaning links as granted, one of the most influential basic vocabularies for Latin, the so-called Bamberg Vocabulary [46], has been constructed using a closed-source corpus, i.e. the text collection and the formulae applied to it have not been published. Therefore, we cannot realistically analyze the quality and 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 [49].

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 [7, p.76; 18, p.191]. 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, p.55]. 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 [30, p.2] 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 [14, p.251; 17, p.348; 20, p.556; 22, p.97; 23, p.443; 48, p.238]. This is in line with the basic assumption of distributional semantics, i.e. that the context of a word constitutes its meaning [32, p.1025]. Additionally, it also supports low-level morphosyntactic tasks where learners are supposed to highlight specific characteristics in a text

(cf. Fig. 1). However, even such tasks with their 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 beforehand to study patterns in language use from various perspectives (cf. Fig. 2).

The historical context
Place and time: Rome, 59 B.C.
M. Tullius Cicero writes to his younger brother Quintus, who has just been confirmed for a third year by the Senate as Propraetor of the Province of Asia. He does not hold any office at the moment, but he is involved in the Senate in his own and his brother's interests. This also includes asking his brother to continue to administer the province of Asia in an exemplary manner and to make as many new and useful contacts as possible.

Task: Mark the predicates.

[...] Atque haec nunc non, ut facias, sed ut te facere et fecisse gaudeas, scribe. Praeclarum est enim summo cum imperio fuisse in Asia triennium sic, ut nullum te signum, nulla pictura, nullum vas, [...] nulla condicio pecuniae, quibus rebus abundat ista provincia, ab summa integritate continentiae deduxerit. Quid autem reperiri tam eximium aut tam expetendum potest quam istam virtutem, moderationem animi, temperantiam [...] in luce Asiae, in oculis clarissimae provinciae atque in aurbus omnium gentium ac nationum esse positam? non itineribus tuis perterri homines, non sumptu exhausti, non adventu commoveri? esse, quocumque veneris, et publice et privatim maximam laetitiam, cum urbs custodem non tyrannum, domus hospitem non expilatorem recepisse videatur? his autem in rebus iam te usus ipse profecto erudit nequaquam satis esse ipsum has te habere virtutes, sed esse circumspiciendum diligenter, ut in hac custodia provinciae non te unum, sed omnes ministros imperi tui socios et civibus et rei publicae praestare videaris.

Score: 4 of 11.

4/11 [Retry](#) [Solution](#)

Fig. 1. Text-based exercise for the identification of linguistic phenomena

2 Collocations and Context

One exemplary operationalization for the acquisition of such context-based lexical knowledge consists in collocations. They can be understood as "word combination[s] whose semantic and/or syntactic properties cannot be fully predicted from those of its components" [11, p.17], i.e. we *need* the surrounding words to fully understand the meaning of a specific target word. Since frequent repetition and variation are desirable factors in high-quality vocabulary training, we have to rely on systematic language resources [1, p.3], e.g. text corpora. However, not every corpus is suitable for this purpose: In language learning, teachers tend to focus on a traditional, rather small canon of authors to be read [19, p.21; 4, p.13]. Such a limited corpus may not provide enough opportunities to study the meaning of certain words in the form of collocations [2, p.275]. Thus, our challenge is to cope with the task of teaching vocabulary in a highly efficient manner, exploiting the few textual resources and, where necessary, extending them. This is particularly reasonable in cases where additional linguistic evidence is needed and the exact preservation of the original context does not take the top priority. One such use case lies in the analysis of polysemy, where networks of collocations can act as representations of a word's multifaceted meaning [21, p.8; 14, p.251].

Abstracting from this to a higher level, we may argue that the focus on authentic contexts is a prerequisite for successful communication [35, p.111], which in our case manifests itself in the form of consuming ancient literature and is treated as a crucial part of teaching Latin [39, p.9]. Besides, modern theories of language learning assume that learners' linguistic experiences, i.e.

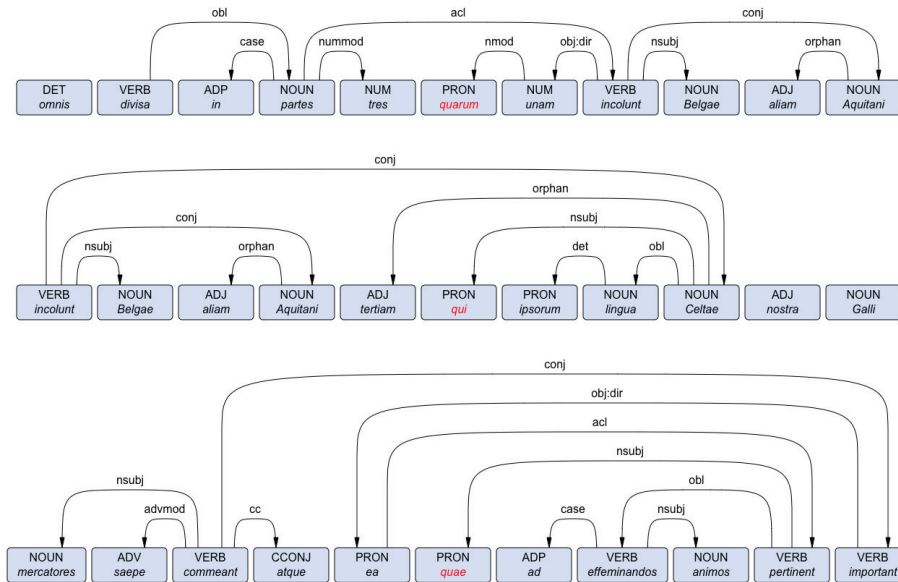


Fig. 2. A *Keyword In Context* view offers the possibility to study specific phenomena (e.g. usage of pronouns) in a structured and focused manner, integrating morphological, syntactic and semantic patterns.

which contexts they have seen frequently, shape their expectations and thus predetermine their ability to understand new communicative input [8, p.374]. Therefore, if we want them to read complex pieces of literature, we have to gradually introduce certain aspects of it right from the start, which is a major shortcoming of current textbooks that oversimplify the original language [44, p.509] to provide a less frustrating learning experience. On top of that, it is not enough to just sort the vocabulary by degree of difficulty. In fact, each lexical item has to be acquired in multiple different contexts, e.g. within text passages or accompanied by glosses and other explanatory material [25, p.142; 27, p.351/353].

3 Evaluation

Even 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 the native language. 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 [18, p.202] as well as second

language learners [42, p.177f.], but it also systematically discriminates against the latter group by relying on the German language for testing purposes. Thus, at least for diagnosis, 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 [15, p.54]): Participants listen to tape-recorded words and, after each one of them, choose from 4 possible pictures that which 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 [42, p.18]. 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 long [34, p.182], some concepts may not be easy to depict accurately. 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 [43, p.290; 15, p.56]. Moreover, individual items in a test should not be weighted equally, but according to their extent and complexity. This becomes obvious when the general focus of the test is shifted from form-meaning links to reading comprehension [38, p.950], which suggests itself given the particular emphasis on the reception of literature in teaching Latin.

Another problem lies in the comparatively small learning input between two tests. In modern language teaching, students can rely on oral practice for accelerated acquisition since it is much faster than written communication. This lack of repetition and intensity has to be compensated in historical languages by relying on sophisticated educational designs that integrate psychological concepts like the mental lexicon and spreading activation [5, p.362/364]: 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, computer-assisted settings. One of them is the lack of a well-defined horizon of expectations for the semantic and morphological parts of lexical evaluation [3, p.209]:

- Which translations (or paraphrases etc.) are appropriate representations of a given target term or concept?
- How do we (consistently) distinguish typographical errors from a 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 [24, p.62]. 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 [29, p.51]. However, many special cases are still hard to cover in this framework, e.g. multiword expressions.

4 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, cf. Fig. 3) and a delayed communication of results (e.g. after a few days). For modern languages, there is additional implicit feedback [9, p.340f.] from conversational exercises, e.g. dialogues. 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 [28, p.7], scaffolding [12, p.959] and adapted to a learner's zone of proximal development [40, p.238]. 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 modeling.

Give a translation.

I already know this.

moderatio animi

serenity ✓ ⓘ

You got 1 of 1 blanks correct.

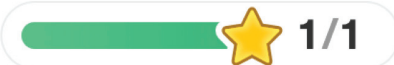
 1/1

Fig. 3. Immediate explicit feedback in a traditional setting for vocabulary learning

A basic, but crucial example consists in the classification of errors: If we do not distinguish between various deviations related to form and meaning [33, p.41], we will fail to give helpful feedback, thus having to fall back to mere scoring. Besides, a simple 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 [10, p.71]. Furthermore, learners are usually not just interested in their current performance on a single test item, but also on their development over time ([45, p.158]; cf. Fig. 4). This kind of ipsative assessment helps them keep track of their own progress and assume responsibility for their learning.

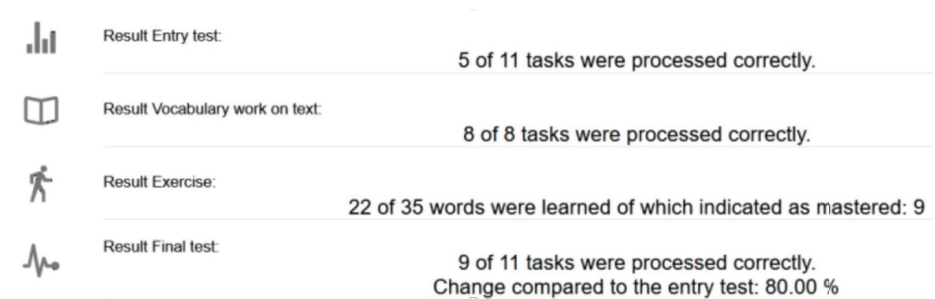



Fig. 4. Explicit summative feedback after a structured learning session, for the purposes of ipsative assessment

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 (cf. Fig. 5) and of overlap with a target vocabulary. 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. 6).

5 Conclusion

Many of the above-mentioned quality criteria for vocabulary learning have already been considered and implemented in the CALLIDUS project's software: 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 to 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.

Text complexity 

Overall complexity: 39.99
 Word count: 303
 Sentence count: 13
 Words per sentence (Ø): 23.31
 Word length (Ø): 5.81
 Number of different word forms: 215
 Number of different parts of speech: 13
 Lexical density: 0.62
 Punctuation mark count: 13
 Main clause count: 13
 Subclause count: 23
 Infinitive count: 2
 Participle count: 4
 Gerund count: 2
 Number of Ablativi Absoluti: 1

The overall measure for text complexity takes into account all the single values presented here in equal proportions. Word and sentence counts are used with predefined ranges of 9 steps (0 to 10, 10 to 50 etc.). Lexical density is measured by comparing the number of unique content words to the overall length of the text.

Fig. 5. Immediate explicit feedback in a traditional setting for vocabulary learning

Search...

Sort by
Date (descending) ▼

► Compare vocabulary

Mark Words	12/5/2019	36	55
C. Iulius Caesar (PROIEL) Commentarii de bello Gallico, 1.1.1-1.1.3			
Cloze	12/3/2019	35	31
Claudianus, Claudius de bello Gildonico, 1-10			
Cloze	12/3/2019	37	30
Ausonius, Decimus Magnus Eclogarum Liber, 1.1-1.18			

Fig. 6. The exercise repository offers various functions for filtering. Every item has measurements of text complexity and percentage of known vocabulary (last two columns).

By building every exercise from Latin literature, i.e. authentic texts written by native speakers, the project tried to 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 novel for the teaching of historical languages, which usually relies heavily on the learners' L1 for communication in class.

While the software also gives explicit feedback (either immediately after an exercise or after a longer period of learning), it does 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 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 pre-structured vocabulary unit, but additional visualizations and a more detailed tracking of learner results seem to be necessary in order to provide a higher overall quality in the curation of lexical materials.

Bibliography

- [1] Aijmer, K.: Corpora and Language Teaching, vol. 33. John Benjamins Publishing (2009), <https://books.google.de/books?hl=de&lr=&id=I1-hYaUJLu0C>
- [2] Baker, P., Gabrielatos, C., Khosravinik, M., Krzyżanowski, M., McEnery, T., Wodak, R.: A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. *Discourse & Society* **19**(3), 273–306 (2008), https://www.researchgate.net/profile/Ruth_Wodak/publication/240705984_A_Useful_Methodological_Synergy_Combining_Critical_Discourse_Analysis_and_Corpus_Linguistics_to_Examine_Discourses_of_Refugees_and_Asylum_Seekers_in_the_UK_Press/links/0c960533c26f1ba1f3000000.pdf
- [3] Beatty, K.: Teaching & Researching: Computer-Assisted Language Learning. Routledge (2013), <http://ebook.stkip-pgri-sumbar.ac.id/ebook/bahasa/teaching-researching-computer-assisted-language-learning-pearson-education-esl/download>
- [4] Beyer, A.: Das Lateinlehrbuch Aus Fachdidaktischer Perspektive: Theorie - Analyse - Konzeption. Universitätsverlag Winter GmbH, Heidelberg (2018)
- [5] Bruza, P., Kitto, K., Nelson, D., McEvoy, C.: Is there something quantum-like about the human mental lexicon? *Journal of Mathematical Psychology* **53**(5), 362–377 (2009), <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2834425/pdf/nihms160244.pdf>
- [6] Crossley, S.A., Salsbury, T., McNamara, D.S.: The development of semantic relations in second language speakers: A case for latent semantic analysis. *Vigo International Journal of Applied Linguistics* **7**, 55–74 (2010), <https://pdfs.semanticscholar.org/5f05/7cfe49a7496b453489f03029440e44d804ef.pdf>
- [7] Daum, M.: Wortschatz Und Lehrbuch: Ein Kriterienkatalog Für Die Wortschatzkonzeption in Lateinlehrwerken, *Ars Didactica. Marburger Beiträge Zu Studium Und Didaktik Der Alten Sprachen*, vol. 2. BoD–Books on Demand (2016), <https://books.google.de/books?id=CSv8CwAAQBAJ>
- [8] Ellis, N.C.: Usage-based and form-focused language acquisition: The associative learning of constructions, learned attention, and the limited L2 endstate. In: Robinson, P., Ellis, N.C. (eds.) *Handbook of Cognitive Linguistics and Second Language Acquisition*. Routledge (2008), https://www.academia.edu/6633691/HANDBOOK_OF_COGNITIVE_LINGUISTICS_AND_SECOND_LANGUAGE_ACQUISITION
- [9] Ellis, R., Loewen, S., Erlam, R.: Implicit and Explicit Corrective Feedback and the Acquisition of L2 Grammar. *Studies in Second Language Acquisition* **28**(2), 339–368 (Jun 2006). <https://doi.org/10.1017/S0272263106060141>, <https://www.cambridge.org/core/journals/studies-in-second-language-acquisition/article/implicit-and->

- explicit-corrective-feedback-and-the-acquisition-of-l2-grammar/CDE67D4A4E286921DA4BE9C40BAD9FE6/core-reader
- [10] Elola, I., Oskoz, A.: Supporting Second Language Writing Using Multimodal Feedback. *Foreign Language Annals* **49**(1), 58–74 (2016). <https://doi.org/10.1111/flan.12183>, <https://onlinelibrary.wiley.com/doi/abs/10.1111/flan.12183>
- [11] Evert, S.: The Statistics of Word Cooccurrences: Word Pairs and Collocations. Dissertation, Universität Stuttgart (2005), <https://elib.uni-stuttgart.de/bitstream/11682/2573/1/Evert2005phd.pdf>
- [12] Finn, B., Metcalfe, J.: Scaffolding feedback to maximize long-term error correction. *Memory & Cognition* **38**(7), 951–961 (2010), <https://link.springer.com/content/pdf/10.3758/MC.38.7.951>
- [13] Freie und Hansestadt Hamburg, Behörde für Bildung und Sport: Rahmenplan alte Sprachen: Latein, Griechisch. Bildungsplan achtstufiges Gymnasium Sekundarstufe I (2004), http://epub.sub.uni-hamburg.de/epub/volltexte/2008/600/pdf/LATGRIE_Gy8.pdf
- [14] Gardner, D.: Validating the Construct of Word in Applied Corpus-based Vocabulary Research: A Critical Survey. *Applied Linguistics* **28**(2), 241–265 (Apr 2007). <https://doi.org/10.1093/applin/amm010>, <https://pdfs.semanticscholar.org/b578/612101f506d0270b002e8a051d5424233478.pdf>
- [15] Gershon, R.C., Slotkin, J., Manly, J.J., Blitz, D.L., Beaumont, J.L., Schnipke, D., Wallner-Allen, K., Golinkoff, R.M., Gleason, J.B., Hirsh-Pasek, K., Adams, M.J., Weintraub, S.: NIH Toolbox Cognition Battery (CB): Measuring Language (Vocabulary Comprehension and Reading Decoding). *Monographs of the Society for Research in Child Development* **78**(4), 49–69 (2013), <https://www.jstor.org/stable/43772790>
- [16] González-Fernández, B., Schmitt, N.: Word Knowledge: Exploring the Relationships and Order of Acquisition of Vocabulary Knowledge Components. *Applied Linguistics* (2019). <https://doi.org/10.1093/applin/amy057>, <https://academic.oup.com/applij/advance-article-pdf/doi/10.1093/applin/amy057/27364938/amy057.pdf>
- [17] Gries, S.T., Wulff, S.: The genitive alternation in Chinese and German ESL learners: Towards a multifactorial notion of context in learner corpus research. *International Journal of Corpus Linguistics* **18**(3), 327–356 (2013), https://www.researchgate.net/profile/Stefan_Gries/publication/263252961_The_genitive_alternation_in_Chinese_and_German_ESL_learners_Towards_a_multifactorial_notion_of_context_in_learner_corpus_research/links/54407e850cf2fd72f99ddae6/The-genitive-alternation-in-Chinese-and-German-ESL-learners-Towards-a-multifactorial-notion-of-context-in-learner-corpus-research.pdf
- [18] Große, M.: Pons Latinus: Latein als reflexionsbasierte Brückensprache im Rahmen eines sprachsensiblen Lateinunterrichts. In: Ammann, E.M.F., Kropp, A., Müller-Lancé, J. (eds.) *Herkunftsbedingte Mehrsprachigkeit Im Unterricht Der Romanischen Sprachen*, vol. 17, pp. 185–206. Frank

- & Timme (2015), <https://books.google.de/books?hl=de&lr=&id=OEHWBgAAQBAJ>
- [19] Guillory, J.: Cultural Capital: The Problem of Literary Canon Formation. University of Chicago Press (Sep 2013), <https://books.google.de/books?hl=en&lr=&id=QhKMIIiwkgC>
- [20] Hagiwara, M., Ogawa, Y., Toyama, K.: Supervised synonym acquisition using distributional features and syntactic patterns. *Information and Media Technologies* 4(2), 558–582 (2009), https://www.jstage.jst.go.jp/article/imt/4/2/4_2_558/_pdf
- [21] Hamilton, W.L., Leskovec, J., Jurafsky, D.: Diachronic word embeddings reveal statistical laws of semantic change. arXiv preprint arXiv:1605.09096 (2016), <https://arxiv.org/pdf/1605.09096.pdf>
- [22] Helm, F.: Language and culture in an online context: What can learner diaries tell us about intercultural competence? *Language and Intercultural Communication* 9(2), 91–104 (May 2009). <https://doi.org/10.1080/14708470802140260>, <https://doi.org/10.1080/14708470802140260>
- [23] Herbelot, A., Ganesalingam, M.: Measuring semantic content in distributional vectors. In: Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers). vol. 2, pp. 440–445 (2013), <https://www.aclweb.org/anthology/P13-2078>
- [24] Hummel, K.M.: Translation and short-term L2 vocabulary retention: Hindrance or help? *Language teaching research* 14(1), 61–74 (2010), <https://journals.sagepub.com/doi/pdf/10.1177/1362168809346497>
- [25] Kersten, S.: The Mental Lexicon and Vocabulary Learning: Implications for the Foreign Language Classroom. Narr Francke Attempto (2010), <https://books.google.de/books?hl=de&lr=&id=NE0x13zHKVoC>
- [26] Laviosa, S.: Translation and Language Education: Pedagogic Approaches Explored. Routledge (2014), <http://translationindustry.ir/Uploads/Pdf/Translation%20and%20Language%20Education.pdf>
- [27] Nation, I.S.: Learning Vocabulary in Another Language. Cambridge University Press, 14 edn. (2012), <https://www.cambridge.org/core/books/learning-vocabulary-in-another-language/491314AA1B451AD04F353600F1C9F0D#>
- [28] Opitz, B., Ferdinand, N.K., Mecklinger, A.: Timing matters: The impact of immediate and delayed feedback on artificial language learning. *Frontiers in Human Neuroscience* 5, 1–9 (2011). <https://doi.org/10.3389/fnhum.2011.00008>, <https://www.frontiersin.org/articles/10.3389/fnhum.2011.00008/full>
- [29] Perez, N., Cuadros, M.: 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. pp. 49–52 (2017), <https://acl-arc.comp.nus.edu.sg/~antho/E/E17/E17-3013.pdf>

- [30] Rayson, P., Piao, S., Sharoff, S., Evert, S., Moirón, B.V.: Multiword expressions: Hard going or plain sailing? *Language Resources and Evaluation* **44**(1), 1–5 (Apr 2010). <https://doi.org/10.1007/s10579-009-9105-0>, <https://link.springer.com/content/pdf/10.1007/s10579-009-9105-0.pdf>
- [31] Robillard, M., Mayer-Crittenden, C., Minor-Corriveau, M., Bélanger, R.: Monolingual and bilingual children with and without primary language impairment: Core vocabulary comparison. *Augmentative and alternative communication* **30**(3), 267–278 (2014), https://www.researchgate.net/profile/Manon_Robillard/publication/263096012_Monolingual_and_Bilingual_Children_With_and_Without_Primary_Language_Impairment_Core_Vocabulary_Comparison/links/5662505308ae15e7462f04ca.pdf
- [32] Roller, S., Erk, K., Boleda, G.: Inclusive yet selective: Supervised distributional hypernymy detection. In: *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*. pp. 1025–1036 (2014), <https://www.aclweb.org/anthology/C14-1097>
- [33] Rudzewitz, B., Ziai, R., De Kuthy, K., Meurers, D.: 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*. pp. 36–46. LiU Electronic Press, Gothenburg, Sweden (May 2017), <https://www.aclweb.org/anthology/W17-0305>
- [34] Schauer, M.: Klasse Klassik: Latein im Klassenzimmer. In: Beuter, K., Hlukhovich, A., Bauer, B., Lindner, K., Vogt, S. (eds.) *Sprache Und Kulturelle Bildung: Perspektiven Für Eine Reflexive Lehrerinnen- Und Lehrerbildung Und Einen Heterogenitätssensiblen Unterricht*, vol. 9, p. 169. University of Bamberg Press (2019), https://fis.uni-bamberg.de/bitstream/uniba/46841/1/FLB9BeuterSpracheopusse_A3a.pdf
- [35] Schemann, H.: Zur Anlage idiomatischer Wörterbücher. Einige Maximen und Reflexionen. In: Mellado-Blanco, C. (ed.) *Theorie Und Praxis Der Idiomatischen Wörterbücher*, pp. 101–117. Max Niemeyer (2009), <https://books.google.de/books?hl=de&lr=&id=ReFfAAysHuQC>
- [36] Schibel, W.: Zur Aneignung lateinischer Literatur und Sprache. *Forum Classicum* pp. 113–124 (2013), <https://journals.ub.uni-heidelberg.de/index.php/fc/article/download/39366/33034>
- [37] Schirok, E.: „Wortschatzarbeit“. In: Doepner, T., Keip, M. (eds.) *Interaktive Fachdidaktik Latein*, pp. 13–34. Vandenhoeck & Ruprecht, Göttingen (2010), <https://static.onleihe.de/content/vandenhoeck/20141125/978-3-647-26411-0/v978-3-647-26411-0.pdf>
- [38] Schmitt, N.: Size and depth of vocabulary knowledge: What the research shows. *Language Learning* **64**(4), 913–951 (2014), http://eprints.nottingham.ac.uk/32299/1/Schmitt_final_version_LL.pdf
- [39] Senatsverwaltung für Bildung, Jugend und Sport Berlin: Rahmenlehrplan für die gymnasiale Oberstufe. Gymnasien, Gesamtschulen mit gymnasialer

- Oberstufe, Berufliche Gymnasien, Kollegs, Abendgymnasien. Latein (2006), https://www.berlin.de/sen/bildung/unterricht/faecher-rahmenlehrplaene/rahmenlehrplaene/mdb-sen-bildung-unterricht-lehrplaene-sek2_latein.pdf
- [40] Shabani, K., Khatib, M., Ebadi, S.: Vygotsky's Zone of Proximal Development: Instructional Implications and Teachers' Professional Development. *English language teaching* **3**(4), 237–248 (2010), <https://files.eric.ed.gov/fulltext/EJ1081990.pdf>
- [41] Siebel, K.: Lateinischer Wortschatz als Brücke zur Mehrsprachigkeit? Eine Durchsicht des Aufgabenspektrums aktueller Lateinlehrwerke. *Pegasus-Onlinezeitschrift* **XI**(1), 102–132 (2011), www.pegasus-onlinezeitschrift.de/2011_1/pegasus_2011-1_siebel.pdf
- [42] Siebel, K.: Mehrsprachigkeit Und Lateinunterricht: Überlegungen Zum Lateinischen Lernwortschatz, vol. 4. Vandenhoeck & Ruprecht (2017), <https://books.google.de/books?hl=de&lr=&id=YOMsDwAAQBAJ>
- [43] Sparrow, S.S., Newman, T.M., Pfeiffer, S.I.: 8 - Assessment of Children Who Are Gifted with the WISC-IV. In: Prifitera, A., Weiss, L.G., Saklofske, D.H. (eds.) *WISC-IV Clinical Use and Interpretation*, pp. 281–298. *Practical Resources for the Mental Health Professional*, Academic Press, Burlington (Jan 2005). <https://doi.org/10.1016/B978-012564931-5/50009-8>, <http://www.sciencedirect.com/science/article/pii/B9780125649315500098>
- [44] Tok, H.: TEFL textbook evaluation: From teachers' perspectives. *Educational Research and Reviews* **5**(9), 508–517 (2010), <https://pdfs.semanticscholar.org/7745/e55e22bbec42afdb11ce25d519d56ff038db.pdf>
- [45] Univio, D.J., Pérez, A.d.P.: Ipsative Assessment of Essay Writing to Foster Reflection and Self-Awareness of Progress. *Handbook of Research on Assessment Literacy and Teacher-Made Testing in the Language Classroom* pp. 157–180 (2019). <https://doi.org/10.4018/978-1-5225-6986-2.ch009>, <https://www.igi-global.com/chapter/ipsative-assessment-of-essay-writing-to-foster-reflection-and-self-awareness-of-progress/217152>
- [46] Utz, C.: Mutter Latein und unsere Schüler — Überlegungen zu Umfang und Aufbau des Wortschatzes [BWS]. *Antike Literatur—Mensch, Sprache, Welt. Dialog Schule und Wissenschaft* **34**, 146–172 (2000), https://www.ccbuchner.de/_files_media/mediathek/downloads/267.pdf
- [47] Van de Loo, T.: Wortschatzarbeit – neuere Perspektiven und schulische Praxis. *Pegasus-Onlinezeitschrift* **16**, 131–151 (2016), <https://journals.ub.uni-heidelberg.de/index.php/pegasus/article/viewFile/35254/28912>
- [48] Webb, S.: The effects of context on incidental vocabulary learning. *Reading in a foreign Language* **20**(2), 232–245 (2008), <https://files.eric.ed.gov/fulltext/EJ815123.pdf>
- [49] Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.W., da Silva Santos, L.B., Bourne, P.E., Bouwman, J., Brookes, A.J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C.T., Finkers, R., Gonzalez-Beltran, A., Gray,

A.J.G., Groth, P., Goble, C., Grethe, J.S., Heringa, J., 't Hoen, P.A.C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S.J., Martone, M.E., Mons, A., Packer, A.L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M.A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., Mons, B.: The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* **3**, 160018 (Mar 2016). <https://doi.org/10.1038/sdata.2016.18>, <https://www.nature.com/articles/sdata201618>