CALLIDUS

Computer-Aided Language Learning:
Lexikonerwerb im Lateinunterricht durch korpusgestützte Methoden

Humboldt-Universität zu Berlin

Computer and Media Service, Corpus Linguistics, Pedagogy of Classics

Andrea Beyer, Malte Dreyer, Stefan Kipf, Milena Kühnast, Anke Lüdeling, Konstantin Schulz, Andreas Vollmer

Research Questions

Question 1:   Are corpus-based methods more supportive in vocabulary acquisition than other methods used in teaching foreign languages?

Question 2:   Do the differences between historical languages and modern languages affect vocabulary acquisition? How?

Question 3:   What implications do corpus-based methods have for the development of tasks?

Question 4:   What does vocabulary competence mean? How can vocabulary competence be acquired and expanded?

Background: Instruction in Latin in German schools

➤ 3rd most important foreign language: almost 700.000 learners (ca. 7%) in 2015, but there is a huge decline in numbers after the textbook phase

➤ Latin starts between the 5th and 11th grade, depending on school types; average span of learning Latin ca. 4 years

➤ Latin is controversial, therefore it has to justify its place in the canon of subjects. → Key question: What does a student gain by learning Latin? (except some (useless?) knowledge of a „dead language")

➤ Some responses: knowledge about language(s), literature, ancient history and European culture, transferable skills like Translations-, Interpretations- und Texterschließungskompetenz → "Bildung"

Specific background: Vocabulary acquisition in Latin classes

➤ A known problem is that students don't learn their vocabulary (well) enough.

➤ The suggested solutions are: reducing the amount of vocabulary and/or focussing on learning techniques.

➤ Vocabulary is currently presented in word equations like
        ratio | ionis, f. | Vernunft, Rechnung, Methode | (maybe +) FW. rationell .

➤ The core vocabulary is questionable because of its selection of words, the given (German) word meaning and the design.

➤ Even in textbooks, vocabulary work is inconsistent and exercises focus mostly on a low cognitive level (focus on repetition and forms).

Research focus: Vocabulary acquisition in a historical language

➤ How many words do you need to learn to understand and/or translate a Latin text? (size)

➤ Which words should be learnt to understand Latin literature in school? (selection)

➤ How should the (core) vocabulary be displayed, repeated and trained regarding the following considerations:
    – What is a word? (word as a lexical entity, e.g. a 1-/2-/3-gram)
    – How do humans store (language) information? (mental lexicon)
    – How should vocabulary be taught? (model of vocabulary acquisition; methods)

❖ What are the implications of corpus-based vocabulary learning for exercises, tasks and tests? (context vs. word equations)

➤ What is specific about learning a historical language? (language acquisition)

Research focus: Using Corpora for teaching

➤ How do we work with a corpus (any collection of written or spoken texts)?
    – association measures (significance vs. effect size)
    – modeling distributional semantics as vector space
    – retrieving text metadata from external sources (e.g. biographical information)
    – dynamic linguistic analysis of texts (e.g. text complexity)

➤ How do we use the advantages of corpora for language teaching (authenticity, context, frequency)?

➤ Is it "equations versus collocations" or rather "equations and collocations"?

➤ How do we find collocations in a text?

➤ How do we break down the linguistic theory with regard to the use in a school context?

Intervention study (Ovid): less complex tasks

What you should know about the intervention study

➤ Study design: pretest, intervention (test groups) with context-based vocabulary tasks (Ovid: Pyramus and Thisbe) and introduction to vocabulary learning, posttest, subsequent vocabulary test

➤ 60 Participants: 2 schools (Berlin), Latin as 2nd/3rd language, 3 groups: 2 x 10th grade (test + control), 1 x 13th grade (test)

➤ Time period: October to December 2018 / January 2019, depending on the schools

Ovid-Study: Average results
(10_t_pre 40%, 10_t_post 44%, 10_k_pre 59%, 10_k_post 57%, GK_t_pre 52%, GK_t_post 55%)

Definition of less complex tasks

Five out of ten tasks are considered to be less complex because they take up either a author-related element (Ovid: hexameter, hyperbata) or some language-related elements well known to students (morphological congruency, lexicon entry, target language phrasing).

Formal vocabulary knowledge tasks

Proposition 1: In a historical language, pronunciation must be explicitly developed and consciously learned so that a word or a group of words can also be anchored in the mental lexicon via the phonological dimension of perception.
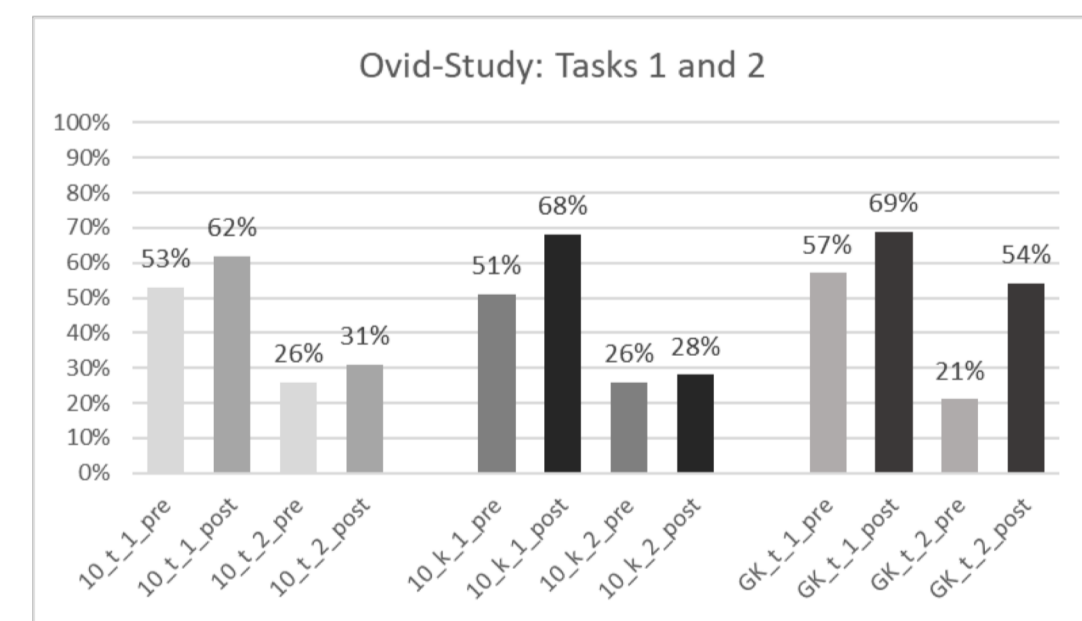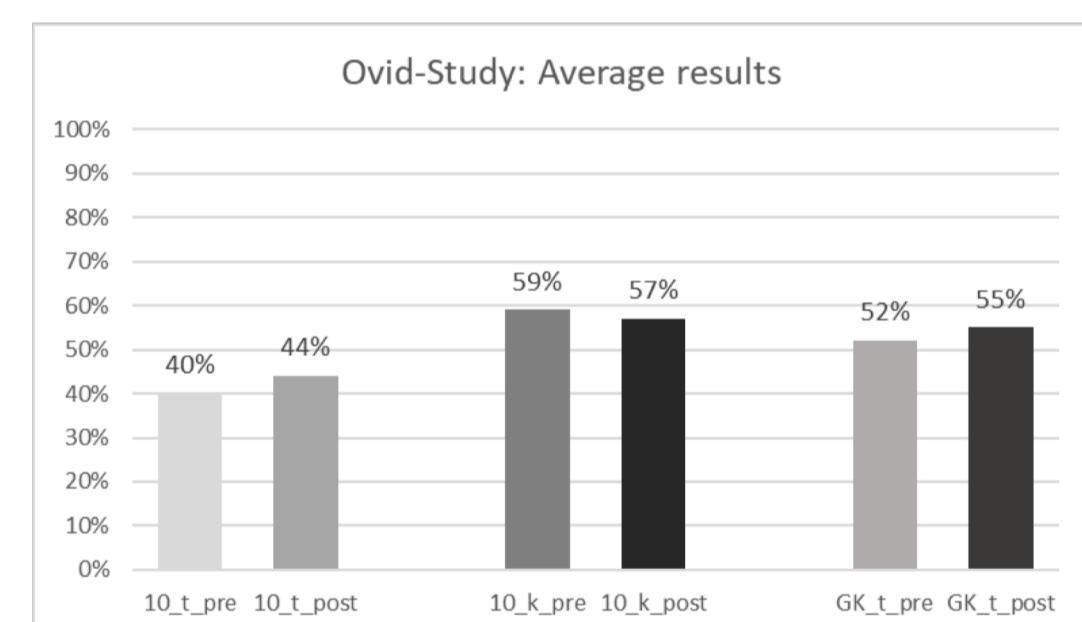

Ovid-Study: Tasks 1 and 2

Proposition 2: In a historical language, a word with its morphological facets must be learned explicitly and reflection-based, because the aim of language acquisition is not language production in the historical language, but detailed text comprehension, which begins at the lexeme or possibly morpheme level. To this end, mastery of the metalanguage is indispensable in order, among other things, to support the rapid retrieval of a mental lexicon entry.
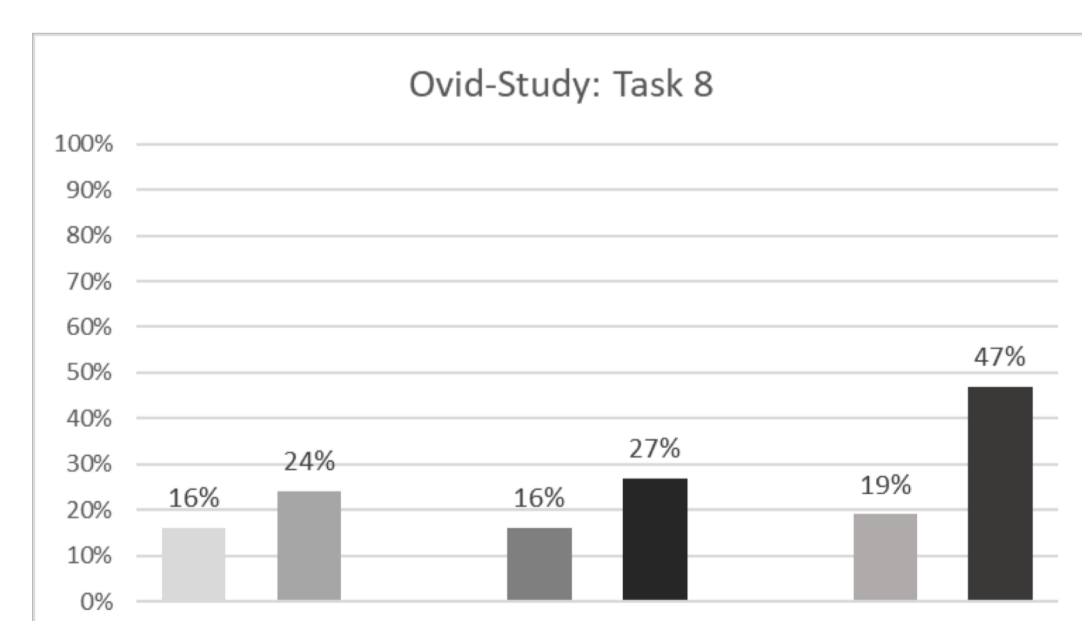

Ovid-Study: Task 8
(10_t_pre 16%, 10_t_post 24%, 10_k_pre 16%, 10_k_post 27%, GK_t_pre 19%, GK_t_post 47%)

Intervention study (Ovid): complex tasks

Functional vocabulary knowledge tasks

Proposition 3: In a historical language, word concepts must be understood as comprehensively as possible and anchored in the mental lexicon. Besides, the meaning of a word or a group of words must be worked out in various contexts in order to finally enable an adequate interlingual representation (translation, mediation, paraphrase etc.).

Proposition 4: In a historical language, syntactic knowledge must be consciously used to structure linguistic information, since often only limited intuitive comprehension is possible due to the great spatial and temporal distance, the otherness of language structures and the receptively oriented acquisition of language.

Ovid-Study: Task 3
(61%, 68%, 71%, 78%, 69%, 67%)

Strategic vocabulary knowledge tasks

Proposition 5: In a historical language like Latin, whose vocabulary has influenced Indo-European languages to varying degrees, the derivation of meanings must be systematically practiced and demanded in order not only to strengthen the networks in the mental lexicon, but also to expand the (rather small) learning vocabulary by a potential vocabulary.
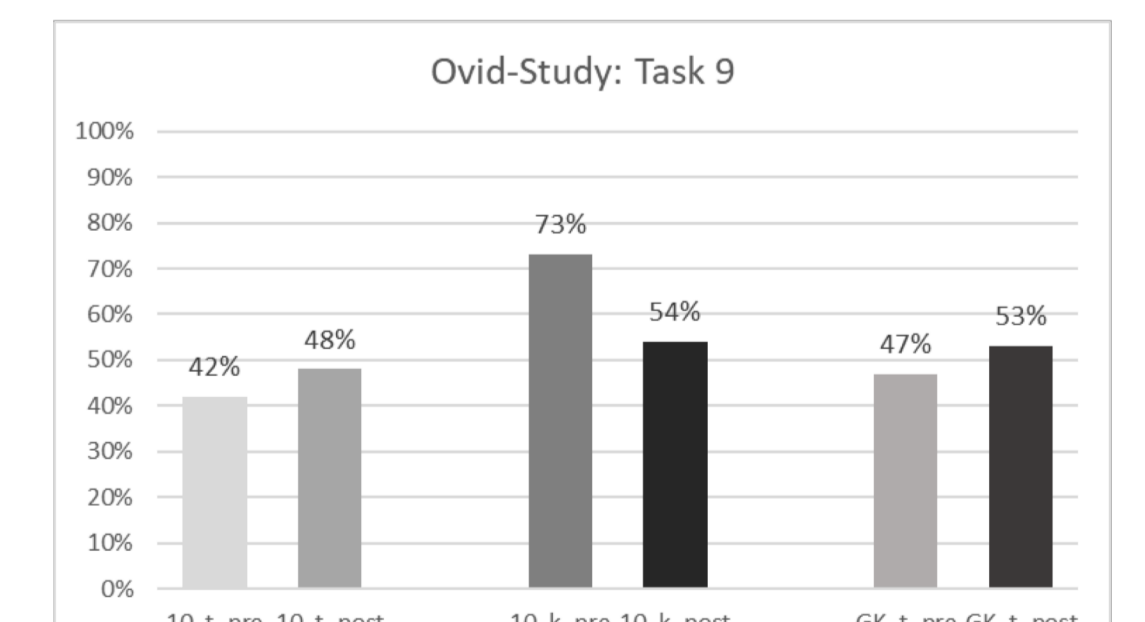

Ovid-Study: Task 9
(10_t_pre 42%, 10_t_post 48%, 10_k_pre 73%, 10_k_post 54%, GK_t_pre 47%, GK_t_post 53%)

Proposition 6: In a historical language, it is particularly important to learn strategies for understanding texts (topic - comment, systematic use of the information provided, etc.), since the text is the only source of information.

First conclusions

➤ The automation of retrieval processes from the mental lexicon, which is necessary for language acquisition, is mostly not successful. (→ task 5: Hyperbata; task 8: Grundform).

➤ The metacognition skills also required for language acquisition are only available to a small extent and are obviously further limited by a low level of language competence in the German (technical or) standard language. (→ task 5: Hyperbata; tasks which require explanations).

➤ The (corpus-based) lexicon acquisition theory on which the research design is based cannot be implemented by a mere contextualization of vocabulary tasks. The results of the additional tasks in pre- and post-tests suggest that the context must be addressed explicitly: none of the study participants was able to fully answer the comprehension questions on the overall context (e.g. end of the myth).

Further Questions

➤ Variety of task types: Which task formats are indispensable for vocabulary acquisition and for what reasons?

➤ Metacognitive Elements: How can linguistic problems be distinguished from cognitive or content-related problems?

➤ How high is the dependency between the success of the concept (data-driven language learning) and the level of language competence?

➤ What should be the goals of Latin classes in the 21st century?

➤ How can necessary psycholinguistic automation processes be integrated into Latin classes in a varied and purposeful way?