CALLIDUS

Computer-**A**ided **L**anguage **L**earning: Lexikonerwerb im Late**i**nunterricht **d**urch korp**us**gestützte Methoden

(Vocabulary Acquisition in Latin using Corpus-based Methods)

at the Humboldt-Universität zu Berlin

Contents

- Who are we?
- Background: Instruction in Latin in German schools
- Research purpose and focus
  - Questions
  - Corpora
  - Vocabulary learning
  - Task-based learning
- Methodology
  - Prolog: What is an intervention study?
  - DBR: A specific intervention study
  - The reuse and developing of software tools
- Status quo
  - The intervention studies
  - The software
  - Problems and open questions

Who are we? – A colourful mix!

Corpus linguistics:
Anke Lüdeling
Konstantin Schulz
Thomas Krause

Didactics of Latin:
Stefan Kipf
Andrea Beyer

Computer and Media Service:
Malte Dreyer
Konstantin Schulz
Andreas Vollmer

+ advisors
+ schools

Background:
Instruction in Latin in German schools

- 3rd most important foreign language: almost 700.000 learners (ca. 7%) in 2015, but 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 (see above).

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

Discussions about vocabulary learning have been ongoing for many years in Latin didactics (e.g. Steinthal 1971, Thurow 1981, Siebel 2011, Daum 2016, Kuhlmann 2016), and often reference the psychology and methods of learning in particular (e.g. Steinhilber 1978, Esser 1999, Nickel 2005, Bösch 2012, Sass 2015, van de Loo 2016). Nevertheless no one really challenges the traditional way of learning in word equations.

Contents

Research Question(s)

Main Question:

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

Some of the subordinated questions:

- What implications do corpus-based methods have for the development of tasks?
- How can we systematically obtain lexical information about (problematic) texts, esp. in order to use it for language acquisition?
- How can we reuse available corpora and search tools for use in schools?
- What are the criteria of a user-friendly software in a teaching/learning environment?
- …

Corpora

- What is a corpus? → Practical definition: any collection of written or spoken texts

- How do we work with a corpus ?
  - association measures (significance vs. effect size)
  - modeling distributional semantics as vector space

- Advantages of corpora for language teaching: authenticity, context, frequency

---

On collocations and covarying collexems, see also Stefanowitsch & Gries 2005 and Gries 2013. For corpora in general, see Lüdeling & Kytö 2008, V. For the specific advantages of corpus-based language teaching, see Gilquin & Granger 2010, 359.

Equations vs. Collocations

- Previously: res = „thing"

➤ res publica = public thing ?

➤ res familiares = family-related thing ?

➤ res frumentaria = grain-related thing ?

- Now: „A collocation is a word combination whose semantic and/or syntactic properties cannot be fully predicted from those of its components […]."

➤ How do we find collocations in a text?

For collocations, see Evert 2005, 17.

Perseus Corpus:
Top 10 Noun + Adjective, Lemma:

- 127: res publicus (t^{fisher} = 1.00)

- 42: populus Romanus (t^{fisher} = 1.00)

- 18: vox magnus (t^{fisher} = 1.00)

- 16: nomen meus (t^{fisher} = 1.00)

- 14: pars tertius (t^{fisher} = 1.00)

- 13: res novus (t^{fisher} = 1.00)

- 13: eques Romanus (t^{fisher} = 1.00)

- 11: res omnis (t^{fisher} = 0.82)

- 10: deus immortalis (t^{fisher} = 1.00)

- 10: aes alienus (t^{fisher} = 1.00)

➜ t-value = 1 for 1237/3851 collocations | Lemmata sum: 9715

Metadata: Sentence Count: 4936 | Token Count: 79697 | Lemma Count: 9715 | Sentence Length Mean: 16.15 | Sentence Length Min: 1 |Sentence Length Max: 91 | Punctuation Count: 6724 | Punctuation Mean per Sentence: 1.36

Perseus Corpus:
Top 10 Verb + Dative Noun, Lemma:

- 7: dico ecclesia (t^{fisher} = 1.00)
- 6: scribo angelus (t^{fisher} = 1.00)
- 3: sum usus (t^{fisher} = 1.00)
- 3: do letum (t^{fisher} = 1.00)
- 3: do populus (t^{fisher} = 0.98)
- 2: sum praesidium (t^{fisher} = 1.00)
- 2: facio res (t^{fisher} = 1.00)
- 2: obsto scelus (t^{fisher} = 1.00)
- 2: sum mens (t^{fisher} = 1.00)
- 2: parco fama (t^{fisher} = 1.00)

➔ t-value = 1 for 363/621 collocations | Lemmata sum: 9715

Various Measures, Top 10, Lemma:

Fisher	Dice	PMI	Chi-Squared (~Phi)	Jaccard
1.00: populus Romanus	1.00: culter venatorius	12.24: funditor Baliaris	4846.00: culter venatorius	1.00: culter venatorius
1.00: deus immortalis	1.00: sacrificium anniversarius	12.24: cupiditas furiosus	4846.00: sacrificium anniversarius	1.00: sacrificium anniversarius
1.00: deus omnipotens	1.00: funditor Baliaris	12.24: focus magicus	4846.00: funditor Baliaris	1.00: funditor Baliaris
1.00: ager Picenus	1.00: cupiditas furiosus	12.24: murra Oronteus	4846.00: cupiditas furiosus	1.00: cupiditas furiosus
1.00: majores noster	1.00: focus magicus	12.24: Phoebe Leucippis	4846.00: focus magicus	1.00: focus magicus
1.00: Augustus divus	1.00: murra Oronteus	12.24: maritus Phrygius	4846.00: murra Oronteus	1.00: murra Oronteus
1.00: bellum Philippensis	1.00: Phoebe Leucippis	12.24: Hermione Spartanus	4846.00: Phoebe Leucippis	1.00: Phoebe Leucippis
1.00: Babylon magnus	1.00: maritus Phrygius	12.24: Thebe Cadmeus	4846.00: maritus Phrygius	1.00: maritus Phrygius
1.00: pontifex magnus	1.00: Hermione Spartanus	12.24: tigris Armenius	4846.00: Hermione Spartanus	1.00: Hermione Spartanus
1.00: miles gregarius	1.00: Thebe Cadmeus	12.24: Eridanus venetus		1.00: Thebe Cadmeus

Various Measures, Top 10, Lemma:

Poisson-Stirling	Student's t-Test	Likelihood Ratio	MI-like exponent 3	MI-like exponent 5
402.53: res publicus	10.64: res publicus	735.00: res publicus	59.89: res publicus	966034.19: res publicus
190.48: populus Romanus	6.34: populus Romanus	310.90: populus Romanus	16.88: populus Romanus	29777.00: populus Romanus
66.96: pars tertius	3.73: vox magnus	99.99: pars tertius	3.11: aes alienus	436.55: pars tertius
62.34: aes alienus	3.68: nomen meus	95.42: aes alienus	2.50: Augustus divus	310.56: aes alienus
60.79: eques Romanus	3.67: pars tertius	94.99: deus immortalis	2.23: pars tertius	301.37: eques Romanus
56.88: deus immortalis	3.54: eques Romanus	94.63: eques Romanus	2.13: deus immortalis	212.77: deus immortalis
50.57: bellum civilis	3.14: aes alienus	77.11: bellum civilis	2.12: ager Picenus	178.68: vox magnus
45.50: deus omnipotens	3.13: deus immortalis	75.61: deus omnipotens	2.06: potestas tribunicius	167.40: nomen meus
42.93: ager Picenus	3.11: bellum civilis	70.25: ager Picenus	2.00: culter venatorius	137.36: bellum civilis
42.68: potestas tribunicius	2.80: deus noster	66.39: potestas tribunicius	2.00: sacrificium anniversarius	87.15: deus omnipotens

Raw Frequencies vs. Meta-Analysis (Top 10, Lemma)

Raw

- 127: res publicus (t^{fisher} = 1.00)
- 42: populus Romanus (t^{fisher} = 1.00)
- 18: vox magnus (t^{fisher} = 1.00)
- 16: nomen meus (t^{fisher} = 1.00)
- 14: pars tertius (t^{fisher} = 1.00)
- 13: res novus (t^{fisher} = 1.00)
- 13: eques Romanus (t^{fisher} = 1.00)
- 11: res omnis (t^{fisher} = 0.82)
- 10: deus immortalis (t^{fisher} = 1.00)
- 10: aes alienus (t^{fisher} = 1.00)

Meta

- 0.00: res publicus
- 4.00: populus Romanus
- 12.00: pars tertius
- 14.00: aes alienus
- 21.00: deus immortalis
- 25.00: eques Romanus
- 34.00: bellum civilis
- 39.00: deus omnipotens
- 46.00: ager Picenus
- 50.00: potestas tribunicius

Corpora as vector spaces

In principio erat verbum et verbum erat apud deum et deus erat verbum.

1. In: 1, principio: 2, erat: 3, verbum: 4, et: 5, apud: 6, deum: 7, deus: 8

2. In -> [1,2,3], principio -> [1,2,3,4], …

3. $erat_1$ -> [1,2,3,4,5], $erat_2$ -> [5,4,3,6,7], $erat_3$ -> [5,8,3,4]

4. erat -> $erat_1$ + $erat_2$ + $erat_3$ … (Magic happens here. Not really addition, though.)

5. „What is a typical context for ‚erat'?" → „verbum"

For distributional semantics and vector spaces, see Mikolov et al. 2013, 2.

Top10 Ranking similar words for „deus" in the Perseus Corpus

with Vulgata	without Vulgata
spiritus	verbum
audio	quaero
oculus	o
vox	levis
caelum	nec
iesus	fleo
dominus	amo
habito	puella
ecclesia	modo
quoniam	durus

Contextual search:
1. All segments for the corpus
2. Only the ones containing the target word
3. Makeshift ranking according to relevance for hits in the Top10: $r = \text{(hit count)}^2 / \text{segment length}$
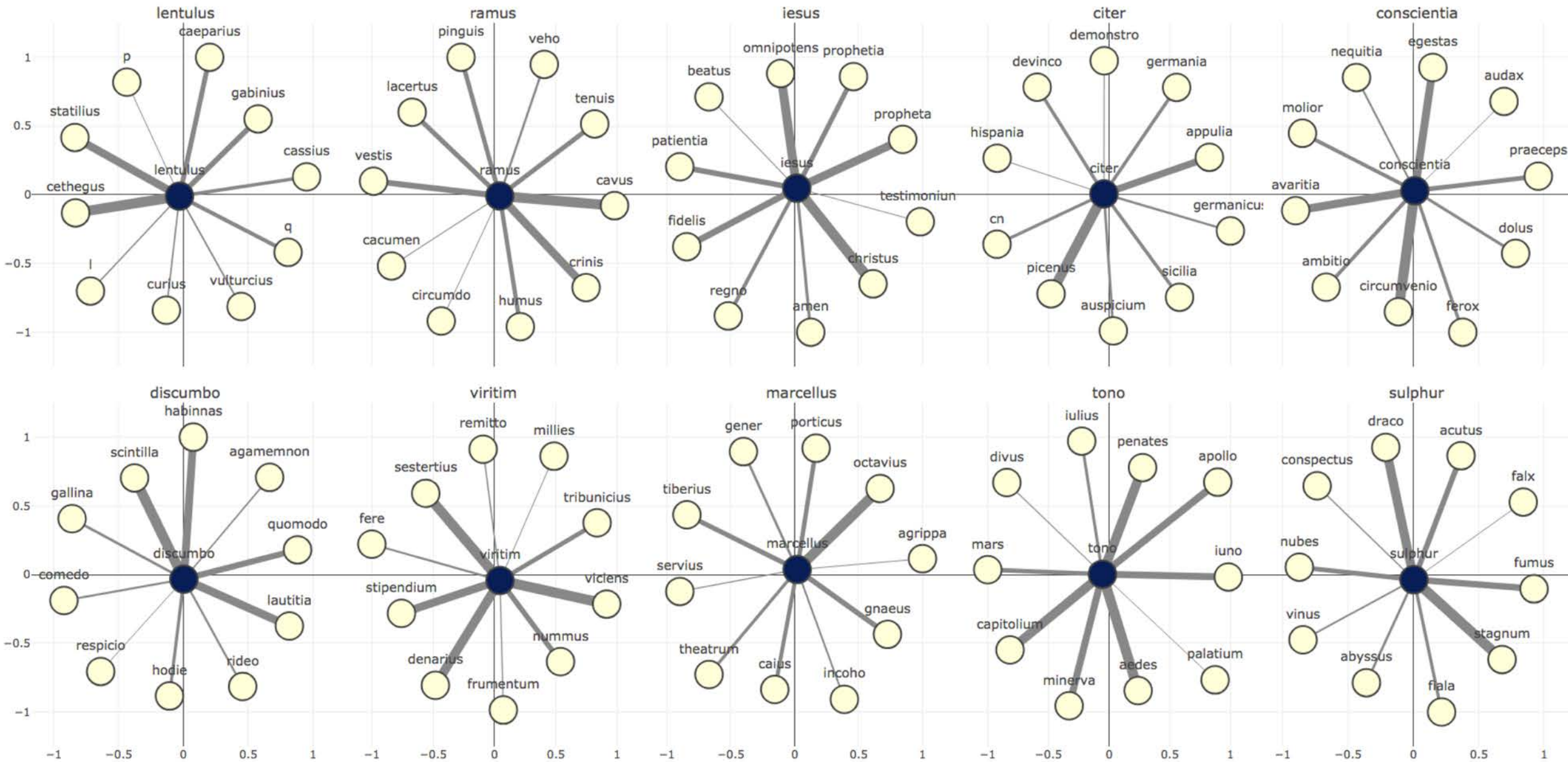
Top10 best contexts for „deus"

with Vulgata	without Vulgata
deus ecce deus	O **di immortales**
alleluia quoniam regnavit **Dominus** Deus noster omnipotens	deus ecce deus
et sextus angelus tuba cecinit et audivi vocem unum ex cornibus altaris aurei quod est ante oculos Dei	desine iam revocare tuis periuria verbis **Cynthia** et oblitos parce movere deos
qui habet aurem audiat quid **Spiritus** dicat ecclesiis vincenti dabo ei edere de ligno vitae quod est in paradiso Dei mei	Tunc contra deus
et sustulit me in spiritu in montem magnum et altum et ostendit mihi civitatem sanctam **Hierusalem** descendentem de caelo a Deo habentem claritatem Dei	Pauci quos aequus amavit **Iuppiter** aut ardens evexit ad aethera virtus dis geniti potuere
et scribam super eum nomen Dei mei et nomen civitatis Dei mei novae **Hierusalem** quae descendit de caelo a Deo meo et nomen meum novum	Non illo melior quisquam c ne **amantior** aequi vir fuit aut illa metuentior ulla deorum
exulta super eam caelum et sancti et **apostoli** et prophetae quoniam iudicavit Deus iudicium vestrum de illa	et quamvis duplici correptum ardore iuberent hac Amor hac Liber durus **uterque deus** subiecto leviter positam temptare lacerto oscula -que admota sumere et arma manu non tamen ausus eram dominae turbare quietem expertae metuens iurgia saevitiae
et habitabit cum eis et ipsi populus eius erunt et ipse Deus cum eis erit eorum Deus	non me deus obruit
magna et mirabilia opera tua Domine Deus **omnipotens** Deum adora	Homines **sumus** non **dei**
	Anchisa generate deum certissima proles Cocyti stagna alta vides Stygiam -que paludem di cuius iurare timent et fallere numen

Perseus Corpus: Author, Work, Texts

- **Petronius**, Cena Trimalchionis: 14177 tokens (17.79%)
- **Sallustius**, Catilina: 13191 tokens (16.55%)
- **Jerome**, Vulgate Bible / Apocalypsis: 9309 tokens (11.68%)
- **Cicero**, In Catilinam, 1.1-2.11: 6652 tokens (8.35%)
- **Tacitus**, Annales: 3531 tokens (4.43%)
- **Augustus**, Res Gestae: 3035 tokens (3.81%)
- **Caesar**, De bello Gallico: 1556 tokens (1.95%)

Top10 cooccurrence fields (by similarity)

Vocabulary Learning

- 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, 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?

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

- How should the vocabulary „lists" be designed?

It appears that the existing core vocabulary (Utz 2000) is no longer discussed, although it was not compiled in a scientific way (Beyer & Schulz 2018). – To get an overview of the mental lexicon, we recommend Aitchison 2012 and for the mental lexicon and vocabulary learning Kersten 2010. See Bruza et al. 2009, 362 for further reading. For the difference between orthographical words and lexemes, see Jackson 2014, 8. For the characteristics of corpus-based language learning, see Campoy et al. 2010, 20.
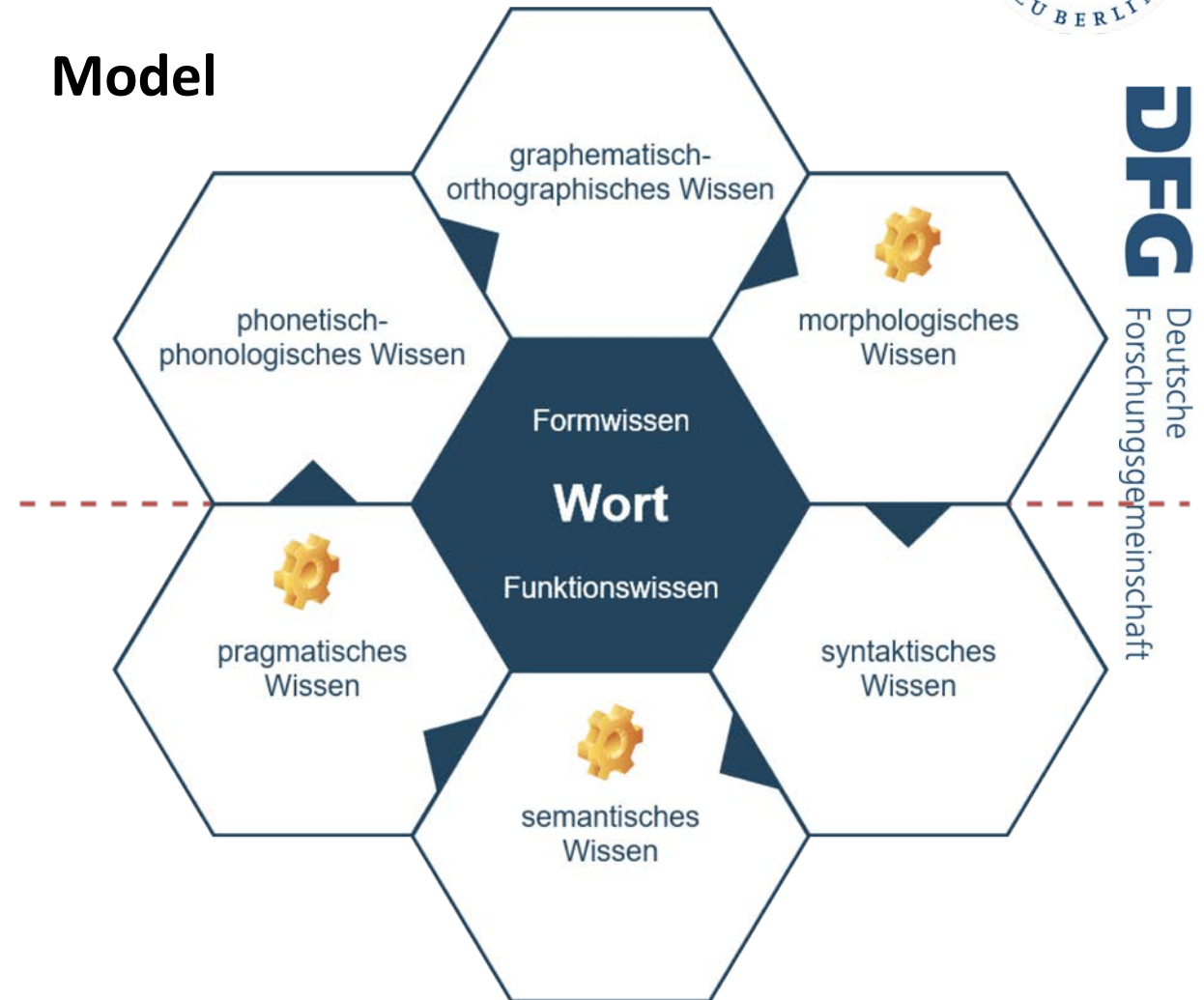
Task-based learning

- „Academic tasks, […], are defined by the answers students are required to produce and the routes that can be used to obtain these answers." (Doyle 1983)

- Ellis (2003) defines tasks more closely: they are supposed to be meaning-focused whereas exercises are supposed to be form-focused.

- The distinction between cognitively complex (task) and cognitively simple (exercise) is clearer: a task might contain exercises but not vice versa.

- Therefore task-based (language) learning involves complex, meaning-focused tasks through which an objective can be attained (Long 1985).

As often, the Anglo-American discussion about tasks and task-based learning „arrived" later in Germany. But esp. after the first PISA results (2001) and the following discussion on „Kompetenzen" (skills) the distinction between task and exercise respectively task-based learning („Aufgabenorientierung") became almost ubiquitous – except in the didactics of Latin. They still don't differentiate between tasks and exercise and are mostly not aware of task-based learning (Kipf 2012, Beyer 2018).

Corpus-based vocabulary tasks

Principles

- context-based: → intralingual, less equations, nevertheless translations of chunks

- focusing on different aspects of vocabulary (→ model)

- explicit and reflective

- connecting to previous knowledge

- productive (in parts)

Model

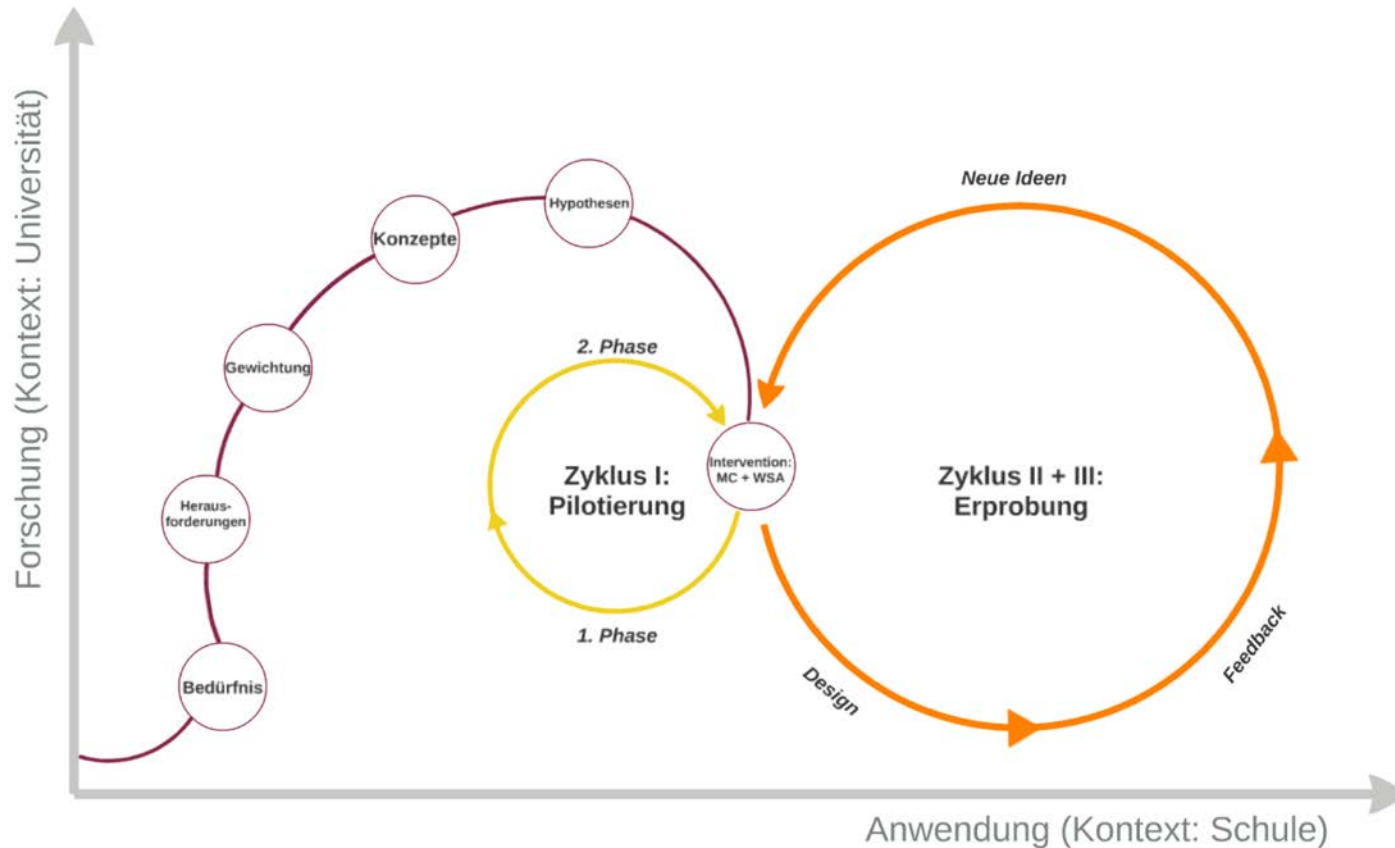
Cf. Steinhoff (2009, 22)

Contents

Prolog: What is an intervention study?

- In contrast to medical intervention trials, an (educational) intervention study is not an experiment in a closed setting.
- This kind of study has a characteristic structure:

  pre-test – intervention (or not) – post-test
- There are normally two groups (test and control). These groups must be comparable, e.g. two classes at least of the same grade in the same school.
- In the pre-test the level of a certain kind of knowledge (and/or skills) is to be evaluated, e.g. how well the students perfom in vocabulary tasks.
- In the intervention one of the groups gets the intervention material, e.g. special tasks and exercises for vocabulary learning. The other group gets nothing in particular. But apart from that both groups work with the same material.
- The post-test is identical to the pre-test. The level of knowledge is evaluated once more and compared to the results of the pre-test. The comparison might focus on the individual development and/or that of the group compared to the control.

Design-Based Research
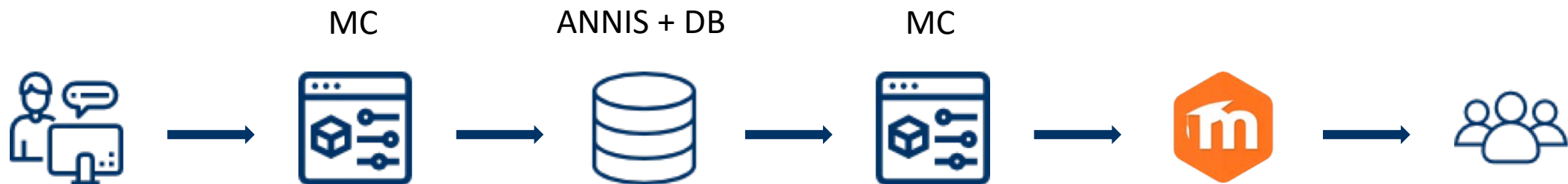# as a specific intervention study

„Design Research is not defined by methodology. All sorts of methods may be employed. What defines design research is its purpose: sustained innovative development." (Bereiter 2002, 330)

→ iterative, interventionist and theory-oriented

→ purpose: do real work in practical educational contexts

Design-based research is a highly recommended research design in an educational context. Therefore this methodology is also called Educational Design Research (EDR) (McKenney & Reeves 2012). For further reading see Bereiter 2002, Cobb et al. 2003, Anderson & Shattuck 2012, Euler & Sloane 2014, Reinmann 2014, Bakker & van Eerde 2015.

The reuse and developing of software tools

- Reuse: ANNIS, PROIEL, PERSEUS, PERSEIDS, Moodle, GitLab, …

- Use of POS-Taggers, Stemming, Lemmatization, Tokenization

- Developing a software = **m**achina **c**allida → (intended) workflow:

The teacher decides on a corpus and vocabulary exercises as the first step in using the machina callida (MC). MC uses the search tool ANNIS (Krause & Zeldes 2016) for the chosen corpus and then builds the answers with context for the chosen task, e.g. a fill-in-the-gap exercise. Afterwards MC exports this to Moodle where the user(s) work(s) on the exercise.

Contents

2 Types of Intervention

experienced learners

- Corpus: Cicero, Quint.1, Ovid, Pyramus et Thisbe

- 9th/10th grade (~ 15 y), 5-6 y Latin; 12th grade (~17y), 4-5y Latin (June 2018, autumn 2018)

- Setting: 12/18 lessons, 2 classes
  - basis = same texts with vocabulary support,
  - add-on = vocabulary learning material/ lesson
  - tests (Cicero) = tasks e.g. on morphology, on collocations, on strategies for deriving meaning, finally a fill-in-the-gap task (3 level)
  - tests (Ovid) = tasks on vocabulary skills „live", i.e. not isolated, but coherent tasks with a context
  - pre-=post-test: performance development A → B
  - previous performance evaluated by the marks from an earlier vocabulary test

beginners

- Corpus: textbook

- 5th grade (~10y), (August 2018)

- Setting: 1 year, 2 x 2 classes, pre-/ post-test (demanding text: reading comprehension, working on the vocabulary, all in German), tests in between

- different approach to vocabulary learning (context vs. word equations)

Examples of vocabulary tasks (productive)

(First, give an introduction and German and Latin examples for collocations, then …)

Exercise:

Give three meaningful collocations by taking one word from each box (left to right) and translating them.

Box 1:
alicui	summo	in
operam	bono	in
alicui	equestri	

Box 2:
initium	loco	animo
loco	salutem	fidem
imperio	damus	

Box 3:
ortus	faciunt	est
ut	affert	es
natus	datis	

Beispiel: *operam damus ut – wir geben uns Mühe, dass …*

Examples of vocabulary tasks (reflective)

a) *proelium committit*   *facinus committit*

b) *deos colimus*   *agros colimus*

c) *labor itineris*   *labores perficiunt*

d) *gravis cura*   *gravis homo*

e) *causa magna*   *pecuniae causa*

f) *a praetore videtur*   *mihi videtur*

g) *rationem reddo*   *nulla ratione aliquid facio*

h) *imperium romanum*   *imperium praetoris*

Exercise:
The underlined words differ in meaning in their given context. They are polysemic. Decide which German meaning might be suitable and translate the phrases.

Example of (test) tasks (context-based)

302 Nympha colit, sed ____ venatibus apta, ____ arcus
flectere quae soleat nec quae contendere cursu,
solaque naiadum celeri non nota Dianae.

305 Saepe suas illi fama est dixisse sorores:
"Salmaci, vel iaculum vel pictas sume pharetras,
____ tua cum duris venatibus otia misce."
____ iaculum sumit ____ pictas illa pharetras,
____ sua cum duris venatibus otia miscet,

310 ____ modo fonte suo formosos perluit artus,
saepe Cytoriaco deducit pectine crines
____, quid se deceat, spectatas consulit undas;
nunc perlucenti circumdata corpus amictu
mollibus ____ foliis ____ mollibus incubat herbis;

315 saepe legit flores.

Dort wohnt eine Nymphe, aber sie ist <u>weder</u> eine Jägerin, <u>noch</u> ist sie es gewohnt,
den Bogen zu spannen oder um die Wette zu laufen;
als einzige der Naiaden ist sie der schnellen Diana nicht bekannt.

Es geht das Gerücht, ihre Schwestern hätten oft zu ihr gesagt:
„Salmacis, nimm einen Wurfspieß oder einen bunten Köcher
<u>und</u> vermische deine Muße mit der harten Jagd.“
Aber jene nimmt keinen Wurfspieß und auch keinen bunten Köcher
noch vermischt sie ihre Muße mit der harten Jagd,
sondern badet nur ihre schönen Glieder in ihrer Quelle.

Oft kämmt sie ihre Haare mit einem Kamm aus Buchsbaumholz und befragt die
bewunderten Wellen, was ihr stehe.
Nun legt sie sich, in ein durchscheinendes Gewand gehüllt, entweder auf weiches
Laub oder ins weiche Gras.
Oftmals sammelt sie Blumen.

3. In den Lücken fehlen die **Konjunktionen** *nec, sed, et,* und *aut*. **Setze** sie (ggf. mehrfach) **passend ein**.   **10 / __**

3. Fill in the gaps with the conjunction *nec, sed, et* or *aut*. Use each as often as necessary.

Example of (test) tasks (semantic, morphology)

340 … At ille,

scilicet ut vacuis et inobservatus in herbis,

huc it et hinc illuc, et in adludentibus undis

summa pedum taloque tenus vestigia tingit;

nec mora, temperie blandarum <u>captus</u> aquarum

345 <u>mollia</u> *de* tenero *velamina corpore ponit.*

Tum vero placuit, nudaeque <u>cupidine</u> formae

Salmacis exarsit: …

Zusammenfassung (Paraphrase):

Aber Hermaphroditus glaubt sich nun unbeobachtet, geht ein wenig hin und her und testet zuerst mit den Füßen das Wasser. Weil es angenehm warm ist, zieht er sich aus, um im Teich baden zu gehen. Da gefiel er Salmacis erst recht und sie wurde noch begieriger nach seiner nackten Gestalt: …

7. Übersetze die zwei kursiv gesetzten **Wortverbindungen zielsprachlich**, d.h. finde einen Ausdruck bzw. eine Redewendung, die den Inhalt der Aussage trifft. **7. Translate** the collocations in italics. 2 / __

8. Gib die Grundform (Infinitiv, Nominativ) der unterstrichenen Wörter **an** und **erkläre**, wie Du Dir diese **Form herleiten** kannst. 6 / __

Beispiel: *blandarum* → *blandarum* und *aquarum* passen zusammen, *aqua* ist ein Substantiv, *blandarum* ist dann wohl ein Adjektiv, dann ist –arum eine Endung und zwar für den Gen.Pl.f., also bleibt bland-, an das man –us für den Nom. anhängt, weil es ein Adj. der o-/a-Dekl. ist = *blandus*

8. Give the lemma of the underlined words and **explain** how you derived the lemma.

Example of vocabulary entries (textbook)

Kontext (lat.)	Grundform und gramm. Informationen	Kontext (dt.)
* Sextus Selicius Comis **dominus** est.	**dominus**, domini, m. (Substantiv, weibl. Form: domina)	Sextus Selicius Comis ist _____ .
* Aurelia domina **est**.	**esse**, sum (Hilfsverb, Kopula)	Aurelia _____ die Hausherrin.
* Marcus **filius** est.	**filius**, filii, m. (Substantiv, weibl. Form: filia)	Marcus ist _____ .
* Gaia **et** Paulla filiae sunt.	**et** (Konjunktion)	Gaia _____ Paulla sind ihre Töchter.
* Sextus Selicius Comis **pater** liberorum est.	**pater**, patris, m. (Substantiv)	Sextus Selicius Comis ist _____ der Kinder.
* Aurelia **mater** liberorum est.	**mater**, matris, f. (Substantiv)	Aurelia ist _____ der Kinder.
* Sextus et Aurelia et liberi et servi **familia** sunt.	**familia**, familiae, f. (Substantiv)	Sextus, Aurelia, die Kinder und die Sklaven bilden _____ .
* **Etiam** Gallus servus hic habitat.	**etiam** (Konjunktion)	_____ der Sklave Gallus wohnt hier.
* Sextus Selicius in domicilio **est**.	**esse** (Verb)	Sextus Selicius _____ im Haus.
** **Silentium** Sexto Selicio placet.	**silentium**, silentii, n. (Substantiv)	_____ gefällt Sextus Selicius.
* **Ubi** liberi sunt?	**ubi** (Interrogativpronomen)	_____ sind die Kinder?

Example of vocabulary work (concept learning)

What do you think about if someone says „family"?

What might a Roman have meant, if he spoke about *familia*?

avus (+ ava)   *(filia +) mater + pater*   *filia + filius*   *servus*   *serva*

Testing vocabulary in context (productive)

1. Ergänze die Sätze sinnvoll.

Die Familie aus dem Lehrbuch wird vorgestellt:

Sextus Selicius Comis _____ est.

Aurelia _____ est.

Marcus filius _____ .

Gaia _____ Paulla filiae sunt.

Marcus et Gaia et Paulla _____ sunt.

Wer fehlt hier und gehört auch zur _familia_? _____

Bonus: Bilde einen lateinischen Satz zur fehlenden Person.

_____

1. Find matching Latin words.
The family of the textbook is introduced:

Who is missing who also belongs to this family?
Bonus: Write a Latin sentence about this person.
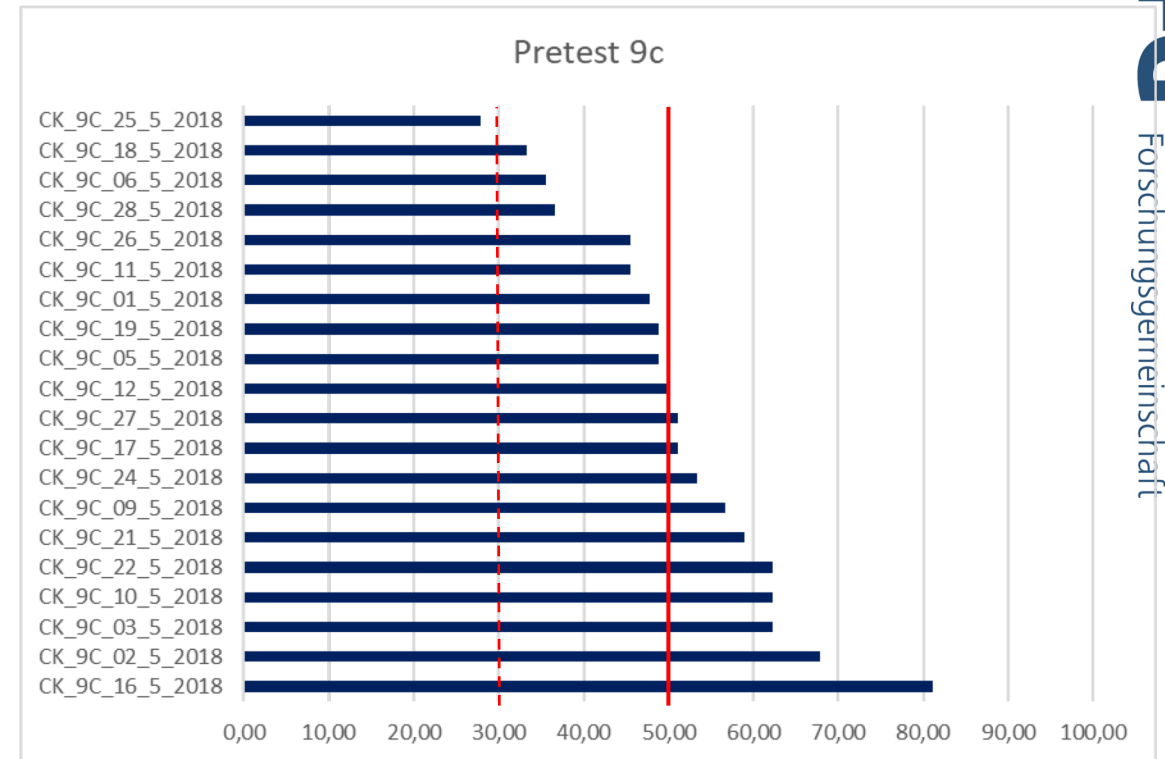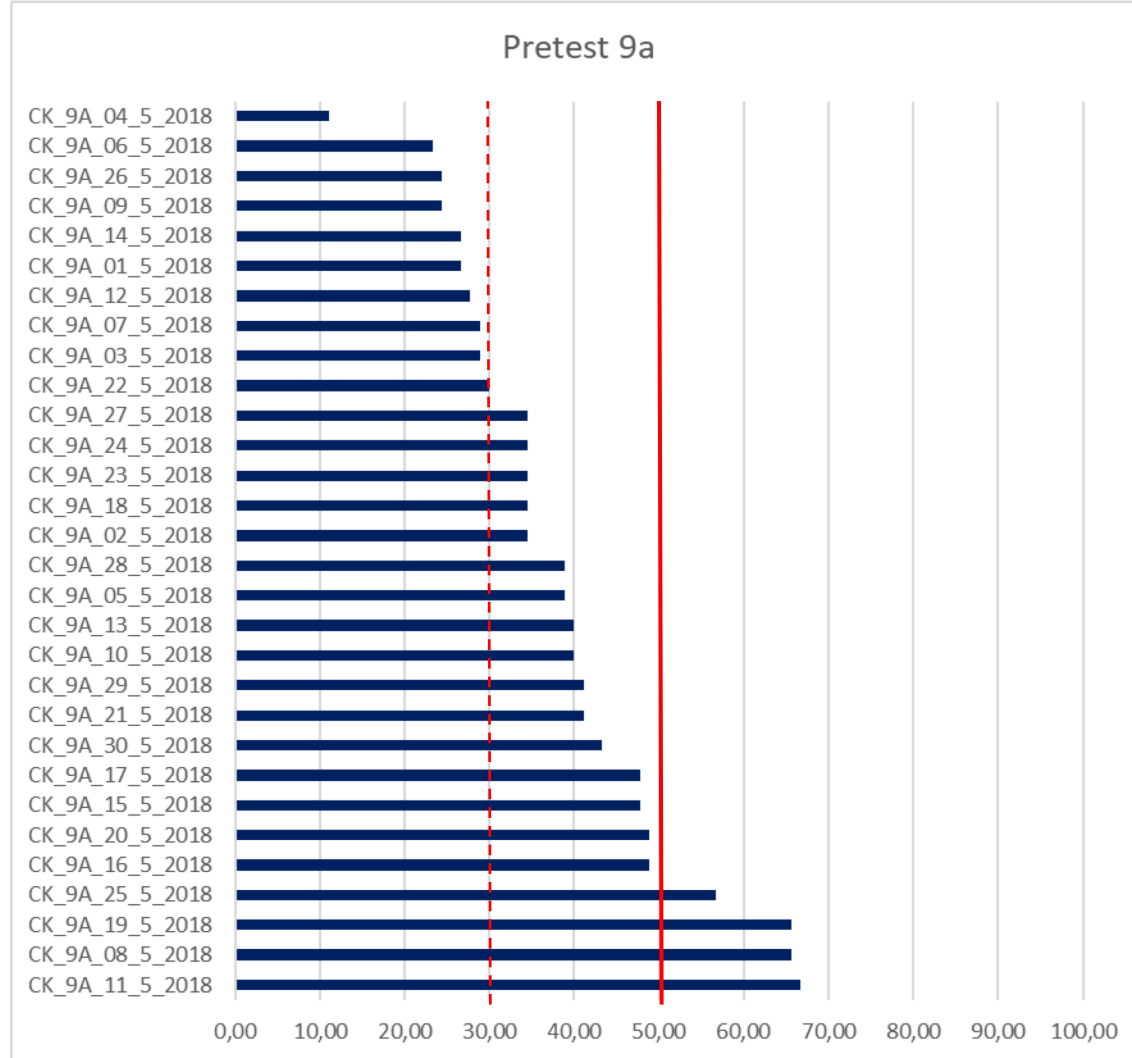
Testing vocabulary in context (pragmatic)

4. Fill in the gaps with „small" Latin words. Pay attention to the context.

4. Setze im deutschen Lückentext die lateinischen „kleinen Wörter" sinnvoll ein.

Sextus Selicius Comis _____ Aurelia warten auf ihre Kinder. Sie sind nämlich noch

immer _____ da. _____ geht die Tür auf _____ Marcus tritt

ein. Hinter ihm kommen _____ seine Schwestern Gaia _____ Paulla.

_____ ist es mit der Stille vorbei. Aber einer fehlt immer noch. Wo steckt nur …?

S.S.C. and Aurelia expect their children. They are still not home. Suddenly the door opens and Marcus comes in. Gaia and Paulla, his sisters, come in after him. It becomes very noisy. But someone is still missing. Where might he be …?

First results of the pretests, 9th grade (June 2018)

Some interesting findings

- There are large differences between classes at the same grade. On average, one class performed much worse, esp. in more abstract and complex tasks.
- Some results concerning vocabulary:
  - crosslingual interferences in language production, e.g.
    litteras scribere **habeo** (instead of: scripsit) – ich **habe** Briefe geschrieben;
    ab-scribere (nonexistent in Latin) – **ab**-schreiben
  - phonological interferences in translation, e.g. **frater** – **Vater** (instead of: Bruder); facili**ter** – leich**ter** (instead of: leicht)
  - overgeneralising in language production, e.g. *vide-tor* (nonexistent in Latin) – Seher, Beobachter (analogical to *ora-tor*)
  - morphological uncertainties in language production, e.g. *amicitia magnum est* (instead of: *amicitia magna est*)
  - lexical uncertainties, e.g. *nos = noster*
  - syntactical (and/or morphological) uncertainties, e.g. *litteras scribere debet* – Briefe müssen geschrieben werden (instead of: er/sie muss einen Brief schreiben)

Machina Callida and Export to Moodle

```
<quiz>
    <question type="category">
        <category>
            <text>$course$/tst_drgNdrp</text>
        </category>
    </question>
    <question type="ddwtos">
        <name>
            <text>Drag and Drop: Cic. Q.Fr. 1,1</text>
        </name>
        <questiontext format="html">
            <text>
                <![CDATA[<p></p><p>Ordne den Lücken die richtigen Konjunktionen zu!</p><p>[[1]] non
dubitabam [[2]] hanc epistulam multi nuntii, fama denique esset ipsa sua celeritate superatura tu [[3]] ante ab
aliis auditurus esses annum tertium accessisse desiderio nostro [[4]] labori tuo, tamen existimavi a me
quoque tibi huius molestiae nuntium perferri oportere.</p><br><p></p>]]></text>
        </questiontext>
        <generalfeedback format="html">
            <text></text>
        </generalfeedback>
        <defaultgrade>1.0000000</defaultgrade>
        <penalty>0.3333333</penalty>
        <hidden>0</hidden>
        <shuffleanswers>0</shuffleanswers>
        <correctfeedback format="html">
            <text>Your answer is correct.</text>
        </correctfeedback>
        <partiallycorrectfeedback format="html">
            <text>Your answer is partially correct.</text>
        </partiallycorrectfeedback>
        <incorrectfeedback format="html">
            <text>Your answer is incorrect.</text>
        </incorrectfeedback>
        <shownumcorrect/>
        <dragbox>
            <text>etsi</text>
            <group>1</group>
```

Machina Callida:

    select author/text

    choose task type and parameters

    export data to Moodle-XML

← Moodle-XML

Problems and open questions

- Corpus data is missing: LangBank
  - Alternatives:
    - repositories (e.g. LAUDATIO) are not yet ready for use
    - treebanks are scarce and cover a very limited amount of texts

- Topics are heavily influenced by teachers, textbooks and a small core of Latin authors

- How do we use the testing effect for vocabulary tasks?

- What are the differences between learning vocabulary in the first/second language and learning it in foreign languages?

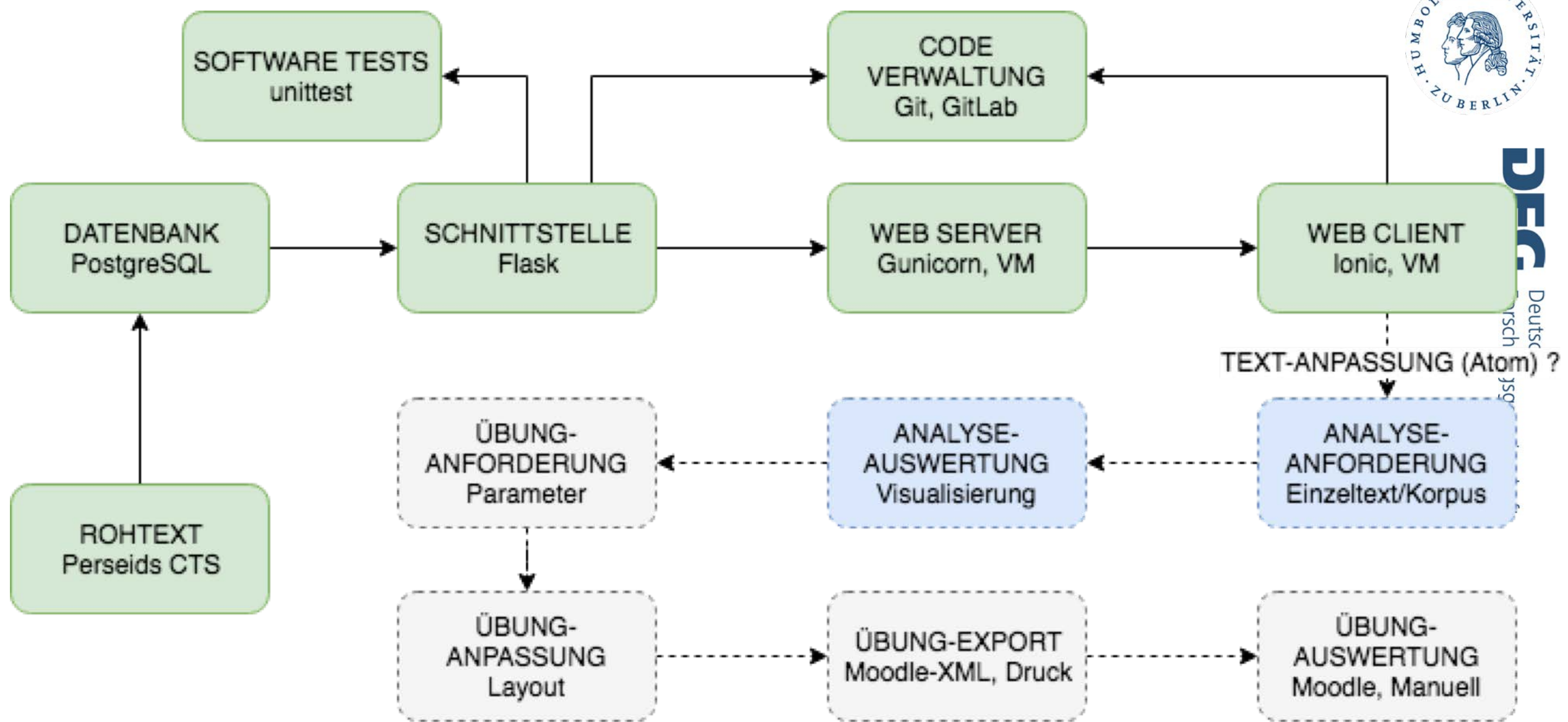- How do we exploit psycholinguistic concepts (e.g. lexical concepts) for Latin vocabulary learning?

For the effect of integrating tests into your practicing routines, see Einstein et al. 2012, 190. For (second) language acquisition in general, see Saville-Troike & Barto 2016, 2. For lexical concepts, see Evans 2009, 2. For the difference between first/second language and foreign language, see Schmitz 2012, 9.

Thank you for paying attention

Lemma	Stamm-formen	Nom. Adj.	Gen. Subst. /Adj.	Genus	Son-der-for-men	Wortart	Syntax Konstr.	Lat. Bsp.	Dt. Bed.	Dt. Bsp.	FS		dt. FW	Zu-sätze	Tags (#)
# ā/ab							Abl.		von, von … her						
abdūcere	abdūcō, abdūxī, abductum								wegführen						
# abesse	absum, āfuī						ā + Abl.		abwesend sein, fehlen				absent, Absentenheft		
# abīre	abeō, abiī, abitum								weggehen				Abitur, Abiturient		
# ac/atque									1 und, und auch, 2 wie, als						
# accēdere	accēdō, accessī, accessum						ad + Akk.		herbeikommen, hinzukommen		e. access f. accéder i. acceder s. accèdere				
accendere	accendō, accendī, accēnsum								anzünden, anfeuern		i. accèndere				
# accidere	accidō, accidī								geschehen, sich ereignen		e. / f. accident i. /s. accidente				
# accipere	accipiō, accēpī, acceptum								erhalten, erfahren		e. to accept f. accepter		akzeptieren		
accūsāre	accūsō								anklagen, beschuldigen		e. to accuse f. accuser		Akkusativ		
ācer		ācris, ācre							energisch, scharf, heftig		e. eager f. aigr i. agro s. agrio				
acerbus		acerba, acerbum							bitter, grausam, rücksichtslos		i. /s. acerbo				
# aciēs			aciēī	f.					Schlachtordnung						

Perseus Corpus: Author, Work, Texts

- **Petronius** Arbiter, Petronius, Petronius: 14177 tokens (17.79%)

- C. **Sallusti** Crispi, Catilina, Iugurtha, Orationes Et Epistulae Excerptae De Historiis, **Catilina**: 13191 tokens (16.55%)

- Jerome, **Vulgate** Bible, **Apocalypsis**: 9309 tokens (11.68%)

- C. **Suetonius** Tranquillus, Divus **Augustus**, 1.1-154.4: 8313 tokens (10.43%)

- Marcus Tullius **Cicero**, M. Tulli Ciceronis Orationes: Recognovit brevique adnotatione critica instruxit Albertus Curtis Clark Collegii Reginae Socius, **In Catilinam, 1.1-2.11**: 6652 tokens (8.35%)

- **Phaedrus**, Augusti libertus**, Fabulae Aesopiae, libri I-III**: 6588 tokens (8.27%)

- **Propertius**, **Elegies, 1.1-1.22**: 5297 tokens (6.65%)

- **Ovid**, Metamorphoses, **Metamorphoses**: 5209 tokens (6.54%)

- Cornelius **Tacitus**, Annales: 3531 tokens (4.43%)

- None, **Res Gestae**: 3035 tokens (3.81%)

- Publius **Vergilius** Maro, Bucolics, Aeneid, and Georgics Of Vergil, **Aeneis 6, 1-295**: 2839 tokens (3.56%)

- Gaius Iulius **Caesar**, C. Iuli Commentarii Rerum in Gallia Gestarum VII A. Hirti Commentarius VII, **De bello Gallico, 2.1-2.3; 2.5; 2.7; 2.9; 2.14-2.18; 2.32-2.33**: 1556 tokens (1.95%)

Perseus Corpus: Metadata

- Sentence Count: 4936
- Token Count: 79697
- Lemma Count: 9715
- Sentence Length Mean: 16.15
- Sentence Length Min: 1
- Sentence Length Max: 91
- Punctuation Count: 6724
- Punctuation Mean per Sentence: 1.36

Bibliography

Aitchison, J. (2012): Words in the mind - An introduction to the mental lexicon. 4. Aufl. Chichester, West Sussex, Malden, MA: Wiley-Blackwell.

Anderson, T.; Shattuck, J. (2012): Design-Based Research: A Decade of Progress in Education Research? In: *Educational Researcher* 41 (1), S. 16–25. DOI: 10.3102/0013189X11428813.

Bakker, A.; van Eerde, D. (2015): An introduction to design-based research with an example from statistics education. In: A. Bikner-Ahsbahs, C. Knipping und N. C. Presmeg (Hg.): Approaches to qualitative research in mathematics education. Examples of methodology and methods (Advances in mathematics education), S. 429–466.

Beard, M. (2012): Do the Classics Have a Future?, http://www.nybooks.com/articles/2012/01/12/do-classics-have-future/

Bereiter, C. (2002): Design Research for Sustained Innovation. In: *Cognitive Studies, Bulletin of the Japanese Cognitive Science Society* 9 (3), S. 321–327, zuletzt geprüft am 30.12.2014.

Beyer, A. (2018): Das Lateinlehrbuch aus fachdidaktischer Perspektive. Theorie - Analyse – Konzeption. Universitätsverlag Winter GmbH, Heidelberg. i.Vorb.

Beyer, A., Kipf, S., Liebsch, A.-C. & Zimmermann, S. (2017). Zwischen Aktualität und historischer Forschung: Entwicklungstendenzen in der Fachdidaktik Latein. Seminar 23 (4), 5–18.

Beyer, A.; Schulz, K. (2018): CALLIDUS – Korpusbasierte, digitale Wortschatzarbeit im Lateinunterricht. i.Vorb.

Bösch, F. (2012): Methodische Überlegungen zur Wortschatzarbeit im Lateinunterricht. Saarbrücken: AV Akademikerverlag.

Bruza, P., Kitto, K., Nelson, D., & McEvoy, C. (2009). Is there something quantum-like about the human mental lexicon? *Journal of Mathematical Psychology*, *53*(5), 362–377. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2834425/pdf/nihms160244.pdf

Campoy, M. C., Cubillo, M. C. C., Belles-Fortuno, B., & Gea-Valor, M. L. (2010). *Corpus-Based Approaches to English Language Teaching*. A&C Black. Retrieved from https://books.google.de/books?hl=en&lr=&id=L6zzqz8JPUgC

Cobb, P.; Confrey, J.; diSessa, A.; Lehrer, R.; Schauble, L. (2003): Design Experiments in Educational Research. In: *Educational Researcher* 32 (1), S. 9–13, zuletzt geprüft am 30.12.2014.

Daum, M. (2016): Wortschatz und Lehrbuch. Ein Kriterienkatalog für die Wortschatzkonzeption in Lateinlehrwerken. Speyer.

Doyle, W. (1983). Academic Work. *Review of Educational Research 53* (2), 159–199.

Einstein, G. O., Mullet, H. G., & Harrison, T. L. (2012). The testing effect: Illustrating a fundamental concept and changing study strategies. *Teaching of Psychology*, *39*(3), 190–193. Retrieved from https://pdfs.semanticscholar.org/2668/bee20e4fd7b2175d2e1340de3373b77d37b0.pdf

Esser, D. (1999): Innovative Übungsformen für die Wortschatzarbeit. In: *Der Altsprachliche Unterricht* 42 (4), S. 44–48.

Euler, D.; Sloane, P. F.E (Hg.) (2014): Design-Based Research. 1. Aufl. Stuttgart: Steiner, Franz (Zeitschrift für Berufs- und Wirtschaftspädagogik, 27).

Evans, V. (2009). *How words mean: lexical concepts, cognitive models, and meaning construction*. Oxford University Press on Demand. Retrieved from https://s3.amazonaws.com/academia.edu.documents/34321936/ve_lexical_concepts.pdf

Evert, S. (2005). *The statistics of word cooccurrences: word pairs and collocations* (Dissertation). Universität Stuttgart. Retrieved from https://elib.uni-stuttgart.de/bitstream/11682/2573/1/Evert2005phd.pdf

Gilquin, G., & Granger, S. (2010). How can data-driven learning be used in language teaching. In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (Vol. 359370, pp. 359–370). Retrieved from https://sites.uclouvain.be/cecl/archives/Gilquin_Granger_2010_How_can_DDL_be_used_in_language_teaching.pdf

Gries, S. T. (2013): 50-something years of work on collocations. In: *International Journal of Corpus Linguistics* 18 (1), S. 137–166.

Hensel, A. (2017). Überprüfen von Textverständnis - neue Wege der schriftlichen Leistungsfeststellung im altsprachlichen Unterricht. *Der Altsprachliche Unterricht 60* (4+5), 2–11

Jackson, H. (2014). *Words and their Meaning*. Routledge. Retrieved from https://books.google.de/books?id=JGm4AwAAQBAJ

Kersten, S. (2010): The mental lexicon and vocabulary learning. Implications for the foreign language classroom.

Kipf, S. (2012). Kompetenzen im Lateinunterricht?! *LGBB 56* (4), 63–77.

Korn, M. (2015). Das Handlungsfeld Sprachunterricht. In S. Kipf & P. Kuhlmann (Hrsg.), Perspektiven für den Lateinunterricht I (Studienbücher Latein, S. 27–33). Bamberg: Buchner

Korn & Behrendt, Schülerzahlen im Fach Latein und Entwicklungsperspektiven der Fachdidaktik, FC 3/2016, 156-57

Korn, M., Kuhlmann, P. & Scholz, I. (2017). Das EPA-Grundproblem. In S. Kipf, P. Kuhlmann & M. Schauer (Hrsg.), *Studienbücher Latein/Perspektiven für den Lateinunterricht II. Praxis des altsprachlichen Unterrichts/Ergebnisse der Dresdner Tagung vom 19./20.11.2015* (Studienbücher Latein, Bd. 10, S. 7–12). Bamberg: Buchner, C.C.

Krause, T., & Zeldes, A. (2016). ANNIS3: A new architecture for generic corpus query and visualization. *Literary and Linguistic Computing*, *31*(1), 118–139. https://doi.org/10.1093/llc/fqu057

Kuhlmann, P. (2016): Wortschatzlernen im Lateinunterricht - Mythen und Fakten. In: *Forum Schule* 63, S. 40–56.

Long 1985, 89, in: Long, M. H. (1985): A role for instruction in second language acquisition: task-based language teaching, in: Hyltenstam, K. und Pienemann, M. (Hrsg.):   Modelling and assessing second language development. Clevedon, 77-99.

Lüdeling, A., & Kytö, M. (2008). *Corpus linguistics: An international handbook*. Citeseer. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.455.9749&rep=rep1&type=pdf

McKenney, S. E.; Reeves, T. C. (2012): Conducting educational design research. New York: Routledge.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *ArXiv Preprint ArXiv:1301.3781*. Retrieved from https://arxiv.org/pdf/1301.3781.pdf

Nickel, R. (2005): Wortschatzarbeit: Wort – Satz – Text. In: *Der Altsprachliche Unterricht* 48 (6), S. 4–11.

Reinmann, G. (2014): Design-based Research: Auftakt für eine methodologische Diskussion entwicklungsorientierter Bildungsforschung. Online verfügbar unter http://gabi-reinmann.de/wp-content/uploads/2013/05/Reader_Entwicklungsforschung_Jan2015.pdf, zuletzt geprüft am 28.03.2015.

Robillard, M., Mayer-Crittenden, C., Minor-Corriveau, M., & Bélanger, R. (2014). Monolingual and bilingual children with and without primary language impairment: core vocabulary comparison. *Augmentative and Alternative Communication*, *30*(3), 267–278. Retrieved from 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

Sass, A. (2015): Sprachenübergreifendes Vokabellernen. Eine qualitativ-interpretative Studie zur Vernetzung der Fächer Englisch und Latein. Dissertation. Staats-und Universitätsbibliothek Bremen.

Saville-Troike, M., & Barto, K. (2016). *Introducing second language acquisition*. Cambridge University Press. Retrieved from https://s3.amazonaws.com/academia.edu.documents/37176416/Introducing_and_Foundations_of_Second_Language_Acquistion_Muriel_Saville-Troike.pdf

Schmitz, S. (2012). *Singen statt Vokabeln pauken: Fremdsprachenlernen mit musikalischer Unterst tzung*. Diplomica Verlag. Retrieved from https://books.google.com/books?hl=en&lr=&id=tKx79tlWllsC

Siebel, K. (2011): Lateinischer Wortschatz als Brücke zur Mehrsprachigkeit? Eine Durchsicht des Aufgabenspektrums aktueller Lateinlehrwerke. In: *Pegasus-Onlinezeitschrift* XI (1), S. 102–132.

Stefanowitsch, A.; Gries, S. T. (2005): Covarying collexems. In: *Corpus Linguistics and Linguistic Theory* 1 (1), S. 1–43.

Steinhilber, J. (1978): Wortschatzübungen im lateinischen Anfangsunterricht. In: *Der Altsprachliche Unterricht* 21 (3), S. 88–92.

Steinhoff, T. (2009): Wortschatz – eine Schaltstelle für den schulischen Spracherwerb?, in: Feilke, H.; Kappest, K.-P.; Knobloch, C. (Hrsg.): SPASS. Schriftenreihe der Universität Siegen, H. 17.

Steinthal, H. (1971): Zum Aufbau des Wortschatzes im Lateinunterricht. In: *Der Altsprachliche Unterricht* 14 (2), S. 20–50.

Thurow, R. (1981): Psychologische Analyse des Wortschatzerwerbs im Lateinunterricht. In: *Der Altsprachliche Unterricht* 24 (4), S. 5–23.

Utz, C. (2000): Mutter Latein und unsere Schüler - Überlegungen zu Umfang und Aufbau des Wortschatzes. In: P. Neukam (Hg.): Antike Literatur - Mensch, Sprache, Welt, Bd. 34. München (Dialog - Klassische Sprachen und Literaturen), S. 146–172.

van de Loo, T. (2016): Wortschatzarbeit – neuere Perspektiven und schulische Praxis. In: *Pegasus-Onlinezeitschrift* 16, S. 131–151.