CALLIDUS

A database of exercises for learning Latin

Konstantin Schulz
Corpus Linguistics and Morphology
Institute for German Language and Linguistics
Humboldt University Berlin

Topics

1. Why we need databases and digital exercises

2. Do contexts make us smarter?

3. Evaluation and feedback

4. Tracing learners' development

1. Why we need databases and digital exercises (I)

- Current **situation** of teaching Latin
  - **rarely** specific vocabulary training

Insufficient vocabulary training: Schirok 2010, Van de Loo 2016

1. Why we need databases and digital exercises (I)

- Current **situation** of teaching Latin
  - **rarely** specific vocabulary training
  - Cross-lingual **word equations**
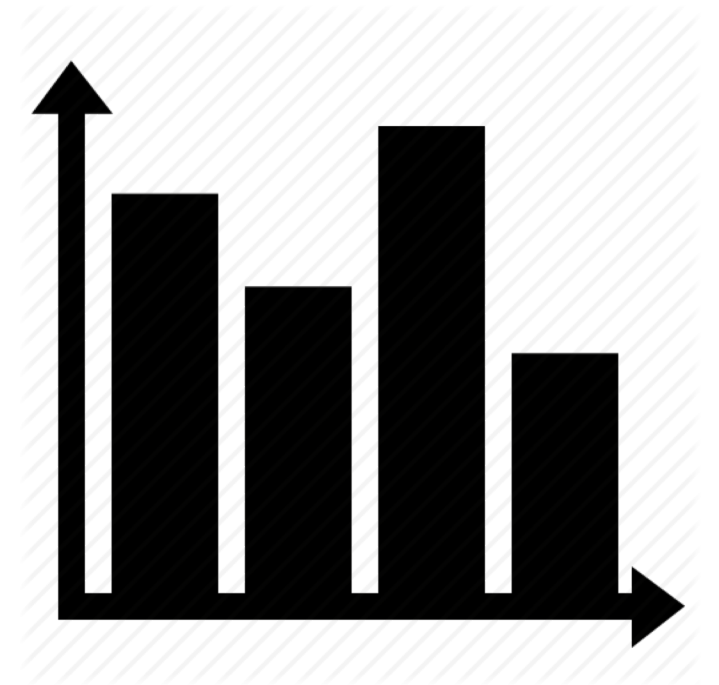
$$x = y$$

1. Why we need databases and digital exercises (I)

- Current **situation** of teaching Latin
  - **rarely** specific vocabulary training
  - Cross-lingual **word equations**
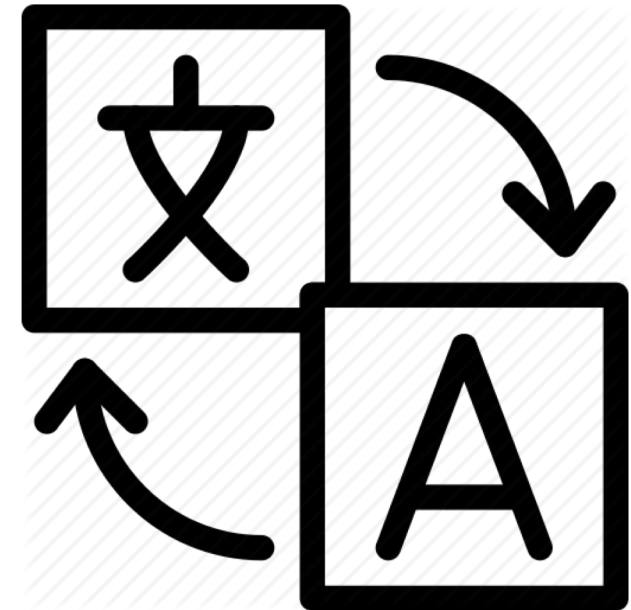   - No common **quality assurance**

1. Why we need databases and digital exercises (I)

- Current **situation** of teaching Latin
  - **rarely** specific vocabulary training
  - Cross-lingual **word equations**
    - No common **quality assurance**
    - often based on **raw frequencies**

raw frequencies: Utz 2000, Daum 2016

1. Why we need databases and digital exercises (I)

- Current **situation** of teaching Latin
  - **rarely** specific vocabulary training
  - Cross-lingual **word equations**
    - No common **quality assurance**
    - often based on **raw frequencies**
    - **translation** = vocabulary knowledge

1. Why we need databases and digital exercises (I)

- Current **situation** of teaching Latin
  - **rarely** specific vocabulary training
  - Cross-lingual **word equations**
    - No common **quality assurance**
    - often based on **raw frequencies**
    - **translation** = vocabulary knowledge
    - **decontextualized** single words
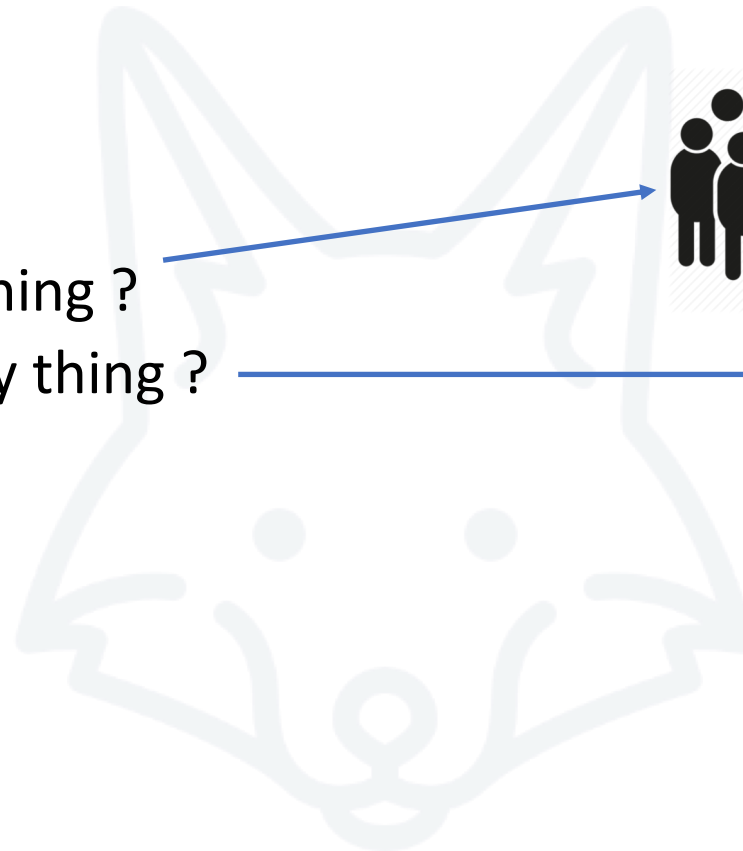
missing contexts: Siebel 2011

1. Why we need databases and digital exercises (II)

- res = thing
  - res **publica** = public thing ?

1. Why we need databases and digital exercises (II)

- res = thing
  - res **publica** = public thing ?
  - res **familiares** = family thing ?

1. Why we need databases and digital exercises (II)

- res = thing
  - res **publica** = public thing ?
  - res **familiares** = family thing ?
  - res **frumentaria** = grain-like thing ?

2. Do contexts make us smarter?

Yes!

- unambiguous context as prerequisite for successful **communication**
- learners' **expectations** on word co-occurrences rely on previous experience
- word **meanings** should be learned in contexts, e.g. a sentence with glosses

communication: Schemann 2009; learner expectations: Robinson & Ellis 2008; contextual vocabulary learning: Webb 2008, Kersten 2010, Nation 2013

3. Evaluation and feedback (I)

evaluation:

Problems	Possible Solutions
lists of lemmata as **representation** for vocabulary competence → Lack of adequate **diagnostic** tools	Applying the **Toolbox Picture Vocabulary Test** to written language Item Response Theory → **individual adaptation** of item difficulty
every lemma weighted **equally**, only total score matters	**variable weighting** of items
decontextualized pairs of **form/meaning**	Strategic reading comprehension: Focus on Form, **derivation, polysemy, collocation**s, based on texts
Lack of language **production**/communication	Vocabulary training needs to be more intensive: → **Spreading Activation** → Stemming, Word Sense Induction, analysis of collocations

Toolbox Picture Vocabulary Test & Item Response Theory: Gershon et al. 2013; reading comprehension instead of form/meaning: Schmitt 2014; Focus on form: Ellis 2009; dimensions of vocabulary competence: González-Fernández & Schmitt 2019; communication in language acquisition: Kersten 2010; Spreading Activation : Bruza et al. 2009

3. Evaluation and feedback (II)

evaluation (continued):

Problems	Possible Solutions
semantics is hard to evaluate automatically	Machine Learning, esp. **Representation learning**, e.g. Word Embeddings

Semantic evaluation: Beatty 2013; Word Embeddings for semantics: Hamilton et al. 2016, Perez & Cuadros 2017

3. Evaluation and feedback (III)

feedback:

- traditional vocabulary tests: **binary** (correct/incorrect), **explicit** (score), **delayed** (after handing in the test)

- historical languages → **rarely implicit** feedback

- goals:
  - immediate, context-dependent **Scaffolding** Feedback
  - **Modeling** learners' development
  - Zone of **proximal** development

implicit vs. explicit feedback: Ellis et al. 2006; zone of proximal development: Shabani et al. 2010; scaffolding feedback: Finn & Metcalfe 2010

English

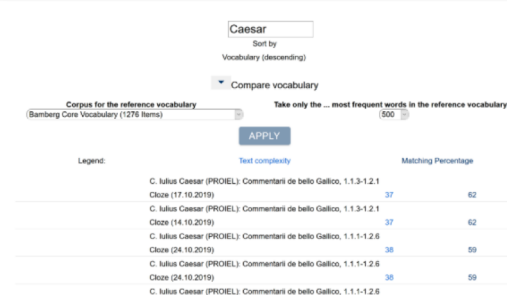# Context matters: Learn to use Latin words smartly!

Here everything has to do with vocabulary exercises to original texts by Cicero, Ovid and Co. Our motto is: No exercise without a reference to the context of the word, as the English linguist John Rupert Firth wrote in 1957: **"You shall know a word by the company it keeps."**

Create exercise

Selection of text
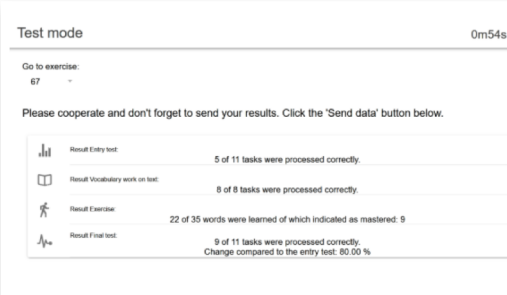Text complexity
Compare vocabulary
Exercise parameters

CONTINUE

Exercise Repository

Exercises created:
Cloze
Mark Words
Matching

CONTINUE

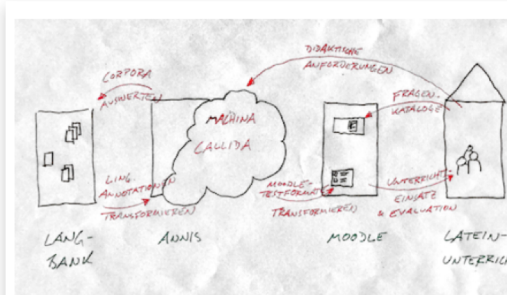## Vocabulary unit

Text work
Exercises
Final test
Evaluation

CONTINUE

Documentation

About the project
Software doc
Exercises doc
Vocabulary unit doc

CONTINUE

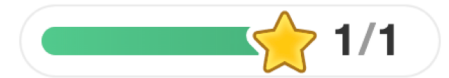|	Result Entry test:	
📊		5 of 11 tasks were processed correctly.
📖	Result Vocabulary work on text:	
		8 of 8 tasks were processed correctly.
🚶	Result Exercise:	
		22 of 35 words were learned of which indicated as mastered: 9
〰️	Result Final test:	
		9 of 11 tasks were processed correctly.
		Change compared to the entry test: 80.00 %

Give a translation.

☐ I already know this.

moderatio animi

serenity ✔ ⓘ

You got 1 of 1 blanks correct.

⭐ 1/1

Before you really get into the exercises, we will start with an overview of the original text.

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, scribo: 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 ✔ [+1] ista provincia, ab summa integritate continentiaque deduxerit. Quid autem reperiri tam eximium aut tam expetendum potest ✔ [+1] quam istam virtutem, moderationem animi, temperantiam […] in luce Asiae, in oculis clarissimae provinciae atque in auribus omnium gentium ac nationum esse positam? non itineribus tuis perterreri homines, non sumptu exhauriri, non adventu commoveri? esse, quocumque veneris, et publice et privatim maximam laetitiam, cum urbs custodem non tyrannum, domus hospitem non expilatorem recepisse videatur ✔ [+1] ? his autem in rebus iam te usus ipse profecto erudivit 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 sociis et civibus et rei publicae praestare videaris ✔ [+1] .

Score: 4 of 11.

⭐ 4/11   ↻ Retry   👁 Solution

His autem in rebus iam te usus ipse profecto erudivit nequaquam satis esse ipsum has te habere virtutes, sed esse circumspiciendum diligenter ut in hac custodia provinciae non te unum sed omnis ministros imperi tui sociis et civibus et rei publicae praestare videaris.

In this regard, the practice has certainly already taught you that it is not enough to possess all these excellent qualities, but that it is necessary to take care that in the provincial administration not only you alone, but all your subordinates seem to be responsible to your allies, to the citizens and to the state.

Give the Latin combination of words behind the translation "in this regard".

[_____]

To whom are Quintus and his subordinates accountable? Give the Latin words in the order of their occurrence.

[_____] [_____] [_____]

Select statements that apply to the content.
If you are not sure what to choose, then read the text again.

[...] And this is not what I am writing to you now to encourage you to such good action, but so that you can rejoice that you have already acted and have acted in this way. It is something very special to have had such unrestricted violence for three years in the province of Asia that no statue, no picture, no vase [...] nor an offer of money has brought you away from your absolute incorruptibility and self-control, although your province offers all these things in abundance. But what can prove to be greater and more desirable than your efficiency, your serenity and your modesty [...] being spread out in the public of the province of Asia, before the eyes of the highly famous province and before the ears of all tribes and peoples? Or that people are not frightened by your journeys through the province, not sucked away by the cost of your reception, not disturbed by your coming? Or that wherever you go, both in public and in private homes, the highest joy breaks out because it turns out that the city has taken in a guardian and not a tyrant, the private home a hospitality and not a robber? In this regard, practice has certainly taught you that it is by no means enough to possess all these excellent qualities yourself, but that it is necessary to take care that in the provincial administration not only you alone, but all your subordinates seem to be responsible to the allies, the citizens and the state.

- [ ] Quintus is popular with everyone in the province.

- [ ] The province of Asia has many things in abundance.

- [ ] Quintus is corruptible and unrestrained.

- [ ] Cicero's brother is governor in the province of Asia.

- [ ] Asia is an unknown province.

- [ ] The provincial administration is in the hands of the Senate of Rome.

esse, quocumque veneris, et publice et privatim ▢▢▢▢▢▢▢ , cum urbs ▢▢▢▢▢▢▢ ⓘnon tyrannum, domus ▢▢▢▢▢▢▢ ⓘnon expilatorem recepisse videatur?

hospitem

maximam laetitiam

custodem

Matching: Assign the matching elements to each other!

Search...

Sort by

Date (descending) ▾

voluntas ▢▢▢▢▢

hostes ▢▢▢▢▢

pueri ▢▢▢▢

diem ▢▢▢▢▢

loco ▢▢▢▢

tertium   tua   tui   nostri   familiari

▸ **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			

4. Tracing learners' development (I)

- Pre- vs. post-test
- No feedback in between

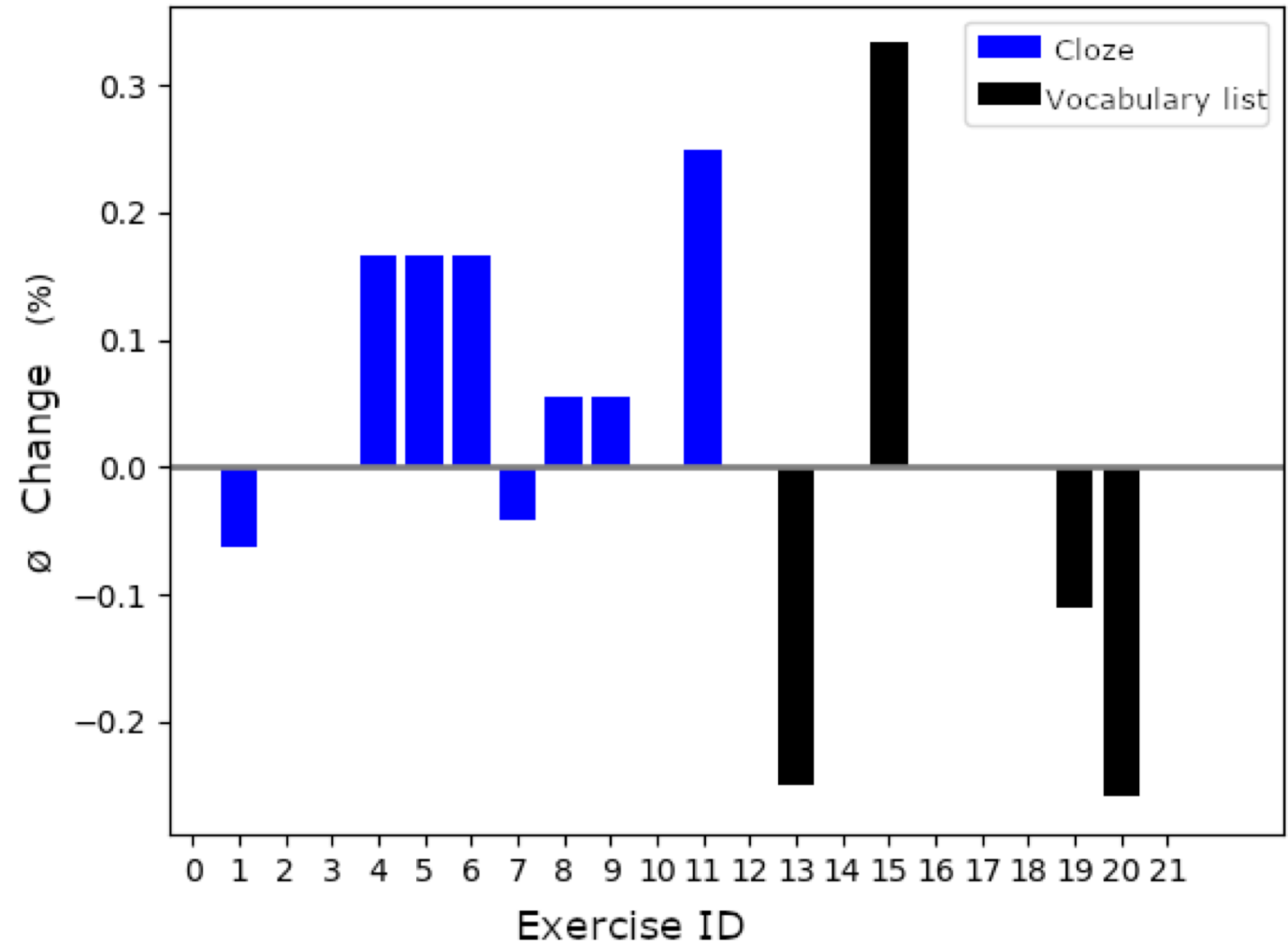Repeated exercises (incl. repetition count)

4. Tracing learners' development (I)

- Pre- vs. post-test
- No feedback in between

Pretest vs. post test, by exercise type

Typical mistakes in free text exercises (I)

dieser, diese, dieses || dieser
oberster || höchster
mit, durch || mit
Geld || das Geld
Weg || der Weg
berühmt, hell, klar || berühmt
Geist, Herz, Seele || das Herz
Tugend, Tapferkeit, Tüchtigkeit || die Tugend
Befehl || der Befehl
Reich || der Befehl
Geld || das Geld
Weg || der Weg
Volk || der Stamm
Mut || das Herz
Weg || der Weg
Geld || das Geld
Beschützer || der Wächter
berühmteste Provinz || die hochberühmte Provinz
der höchste || höchster
und auch || und
Völker und Nationen || die Stämme und Völker
Reich || der Befehl

Nation || das Volk
der Gastgeber || der Gastfreund
mit Oberbefehl || mit dem Oberbefehl
der Begleiter || der Wächter
 || die Gelassenheit
Vermögen || das Geldangebot
die berühmteste Provinz || die hochberühmte Provinz
oberste || höchster
Nation || das Volk
Tüchtigkeit || die Tugend
Staat || die Sache
auich || und
Volk || der Stamm
hochberühmte Provinz || die hochberühmte Provinz
unumschränkte Gewalt || mit dem Oberbefehl
Geldgabe || das Geldangebot
Reich || der Befehl
hell || berühmt
Staatssache || der Staat
Geist || das Herz
Vieh || das Geld
dein || dieser

Typical mistakes in free text exercises (II)

berühmt || hochberühmt
Gewährten || der Wächter
Weg || der Weg
Volk || die Stämme und Völker
Geld || das Geld
Bergleiter || der Bundesgenosse
tapferkeit || die Tugend
in || mit
Fröhlichkeit || die Freude
reich || der Befehl
vermögen || das Geld
dei berühmte provinz || die hochberühmte Provinz
öffentliuch || öffentlich
geist || das Herz
freuen || sich freuen
hin zu || und
stämme und nationen || die Stämme und Völker
der oberste || höchster
hinzugekommen || die Ankunft
weg || der Weg
froh# || hochberühmt
berühmt || hochberühmt

Tüchtigkeit || die Tugend
Reich || der Befehl
Gefährte || der Bundesgenosse
dieser, diese, dieses || dieser
Reise || der Weg
hochberühmte Provinz || die hochberühmte Provinz
die Stämmer und die Völker || die Stämme und Völker
Seele || das Herz
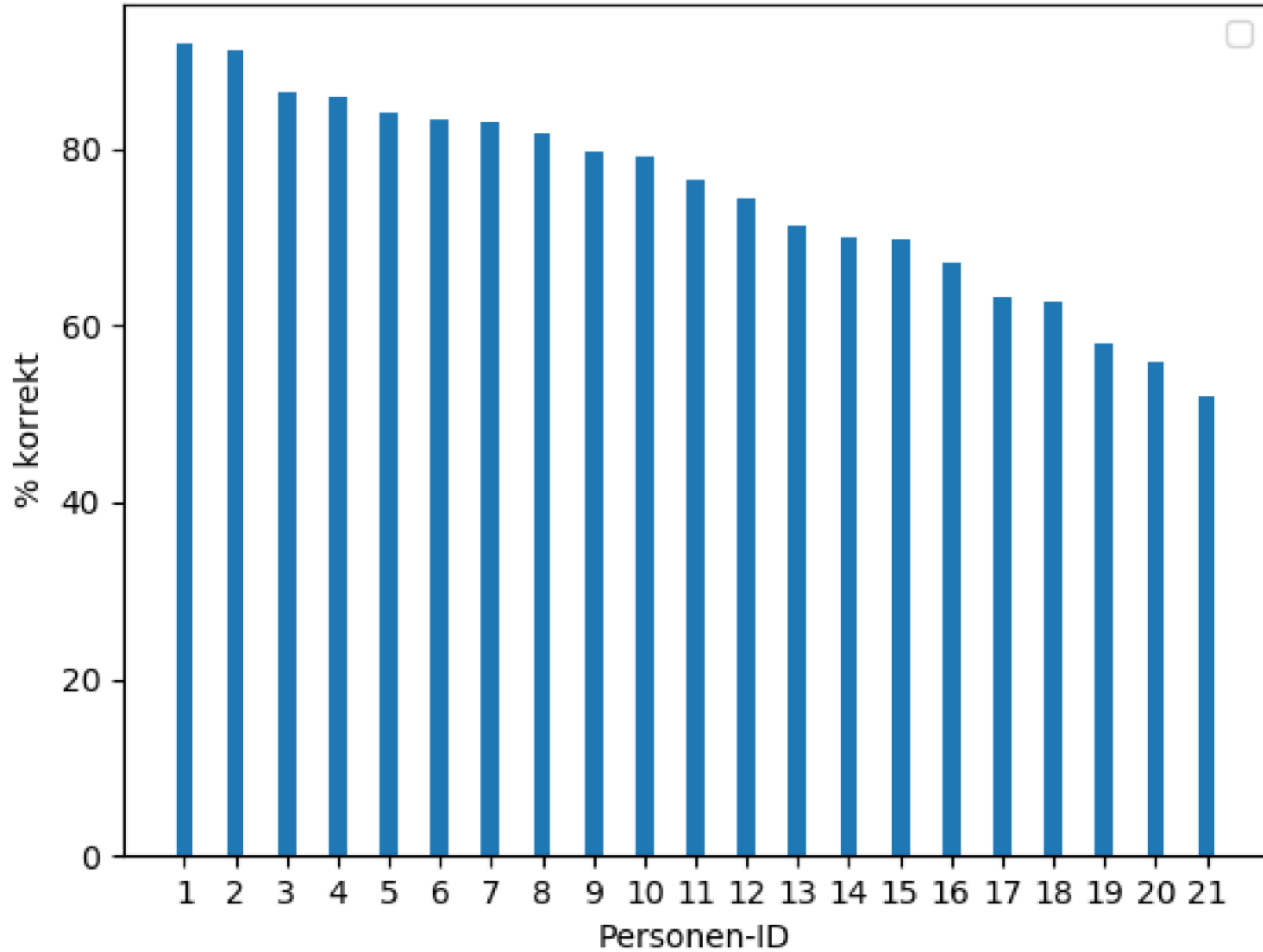Geld || das Geld
der oberste || höchster

4. Tracing learners' development (II): Hypotheses on error sources

- Exercise design:
  - Allow multiple variants, e.g. „this, that" instead of just „this"
  - What is still acceptable? E.g. „absolute power || in supreme command"
  - synonymy, e.g. „supreme || highest"
  - Decontextualized translation / multiple possibilities, e.g. „the most famous province || an illustrious province"
  - determiners, e.g. „in supreme command || in the supreme command"
  - reflexivity, e.g. „freuen || sich freuen" (to rejoice vs. to please)

- learners:
  - Semantically related terms, e.g. „host || guest"
  - powerful collocations / priming, e.g. „state || thing" ($\rightarrow$ res publica)
  - Typographical errors, e.g. „publuic || public" ($\rightarrow$ T9 / autocorrect)
  - Focus on lemma, e.g. „ absolute power || in supreme command"
  - Morphologically similar function words, e.g. „their || this"
  - capitalization, e.g. „weg || der Weg" (away vs. the way)
  - nominalisation, e.g. „arrived || arrival"

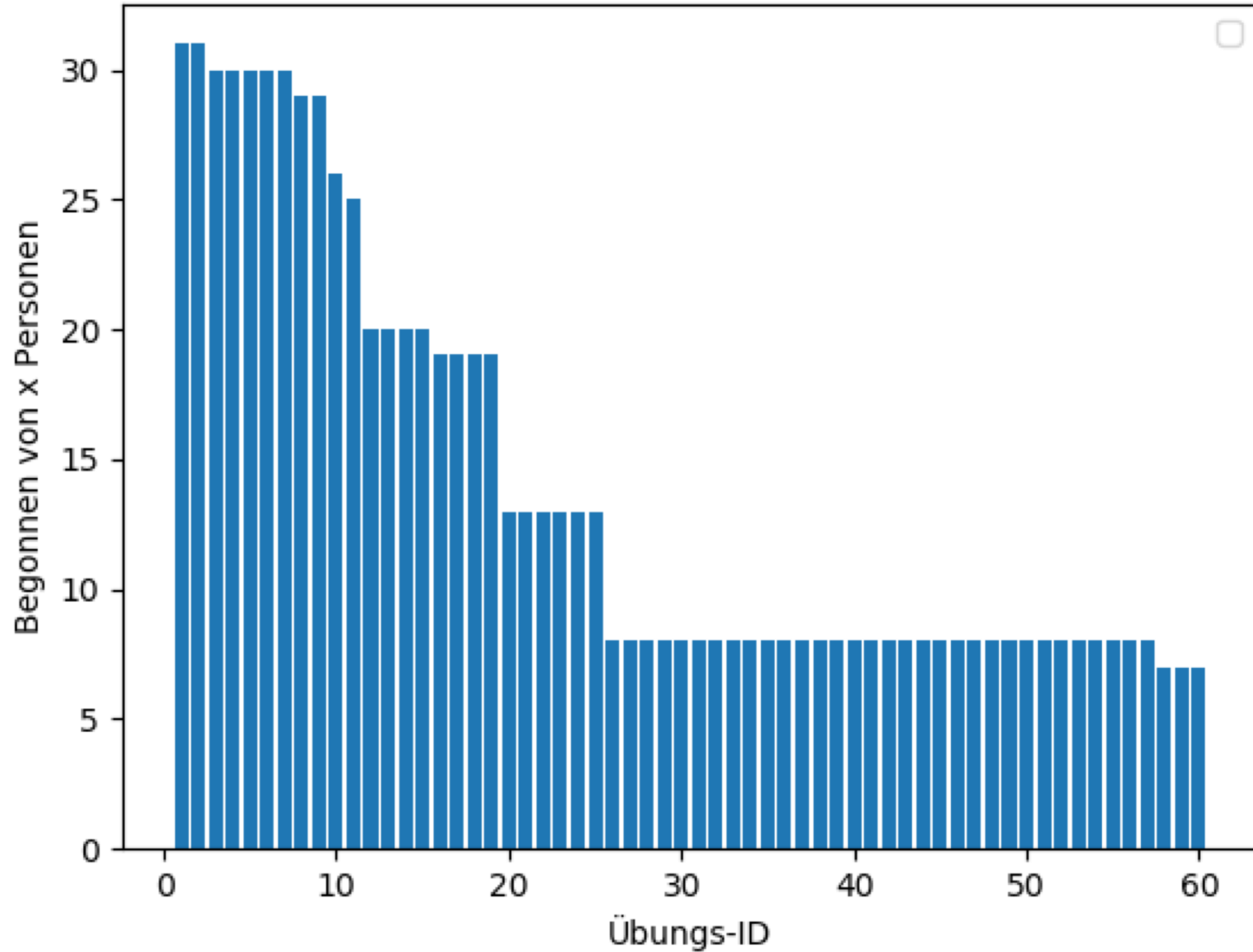Error types in language exercises: Rudzewitz et al. 2017

References

Beatty, K. (2013). *Teaching & researching: Computer-assisted language learning*. Retrieved from http://ebook.stkip-pgri-sumbar.ac.id/ebook/bahasa/teaching-researching-computer-assisted-language-learning-pearson-education-esl/download

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

Daum, M. (2016). *Wortschatz und Lehrbuch: Ein Kriterienkatalog für die Wortschatzkonzeption in Lateinlehrwerken*. Retrieved from https://books.google.de/books?id=CSv8CwAAQBAJ

Ellis, R. (2009). Task-based language teaching: Sorting out the misunderstandings. *International Journal of Applied Linguistics*, *19*(3), 221–246. Retrieved from https://pdfs.semanticscholar.org/2866/fad24dcd59848acb08a5468bacbecc02bf8a.pdf

Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and Explicit Correctiv Feedback and the Acquisition of L2 Grammar. *Studies in Second Language Acquisition*, *28*(2), 339–368. https://doi.org/10.1017/S0272263106060141

Finn, B., & Metcalfe, J. (2010). Scaffolding feedback to maximize long-term error correction. *Memory & Cognition*, *38*(7), 951–961. Retrieved from https://link.springer.com/content/pdf/10.3758/MC.38.7.951

Gershon, R. C., Slotkin, J., Manly, J. J., Blitz, D. L., Beaumont, J. L., Schnipke, D., … Weintraub, S. (2013). 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. Retrieved from https://www.jstor.org/stable/43772790

González-Fernández, B., & Schmitt, N. (2019). Word Knowledge: Exploring the Relationships and Order of Acquisition of Vocabulary Knowledge Components. *Applied Linguistics*. https://doi.org/10.1093/applin/amy057

Hamilton, W. L., Leskovec, J., & Jurafsky, D. (2016). Diachronic word embeddings reveal statistical laws of semantic change. *ArXiv Preprint ArXiv:1605.09096*. Retrieved from https://arxiv.org/pdf/1605.09096.pdf

Kersten, S. (2010). *The mental lexicon and vocabulary learning: Implications for the foreign language classroom*. Retrieved from https://books.google.de/books?hl=de&lr=&id=NEOx13zHKVoC

Nation, I. S. P. (2013). *Learning Vocabulary in Another Language* (2nd ed.). Retrieved from https://books.google.de/books?id=KjYuDwAAQBAJ

Perez, N., & Cuadros, M. (2017). Multilingual call framework for automatic language exercise generation from free text. *Proceedings of the Software Demonstrations of the 15th Conference of the European Chapter of the Association for Computational Linguistics*, 49–52. Retrieved from https://acl-arc.comp.nus.edu.sg//~antho/E/E17/E17-3013.pdf

Robinson, P., & Ellis, N. C. (2008). *Handbook of cognitive linguistics and second language acquisition*. Routledge.

Rudzewitz, B., Ziai, R., De Kuthy, K., & Meurers, D. (2017). Developing a web-based workbook for English supporting the interaction of students and teachers. *Proceedings of the Joint Workshop on NLP for Computer Assisted Language Learning and NLP for Language Acquisition*, 36–46. Retrieved from https://www.aclweb.org/anthology/W17-0305

Schemann, H. (2009). Zur Anlage idiomatischer Wörterbücher. Einige Maximen und Reflexionen. In C. Mellado-Blanco (Ed.), *Theorie und Praxis der idiomatischen Wörterbücher* (pp. 101–117). Retrieved from https://books.google.de/books?hl=de&lr=&id=ReFfAAysHuQC

Schirok, E. (2010). „Wortschatzarbeit ". In T. Doepner & M. Keip (Eds.), *Interaktive Fachdidaktik Latein* (pp. 13–34). Retrieved from https://static.onleihe.de/content/vandenhoeck/20141125/978-3-647-26411-0/v978-3-647-26411-0.pdf

Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning*, *64*(4), 913–951. Retrieved from http://eprints.nottingham.ac.uk/32299/1/Schmitt_final_version_LL.pdf

Shabani, K., Khatib, M., & Ebadi, S. (2010). Vygotsky's Zone of Proximal Development: Instructional Implications and Teachers' Professional Development. *English Language Teaching*, *3*(4), 237–248. Retrieved from https://files.eric.ed.gov/fulltext/EJ1081990.pdf

Siebel, K. (2011). Lateinischer Wortschatz als Brücke zur Mehrsprachigkeit? Eine Durchsicht des Aufgabenspektrums aktueller Lateinlehrwerke. *Pegasus-Onlinezeitschrift*, *XI*(1), 102–132. Retrieved from www.pegasus-onlinezeitschrift.de/2011_1/pegasus_2011-1_siebel.pdf

Utz, C. (2000). Mutter Latein und unsere Schüler—Überlegungen zu Umfang und Aufbau des Wortschatzes. *Antike Literatur–Mensch, Sprache, Welt. Dialog Schule Und Wissenschaft*, *34*, 146–172. Retrieved from https://www.ccbuchner.de/_files_media/mediathek/downloads/267.pdf

Van de Loo, T. (2016). Wortschatzarbeit – neuere Perspektiven und schulische Praxis. *Pegasus-Onlinezeitschrift*, *16*, 131–151. Retrieved from https://journals.ub.uni-heidelberg.de/index.php/pegasus/article/viewFile/35254/28912

Webb, S. (2008). The effects of context on incidental vocabulary learning. *Reading in a Foreign Language*, *20*(2), 232–245. Retrieved from https://files.eric.ed.gov/fulltext/EJ815123.pdf
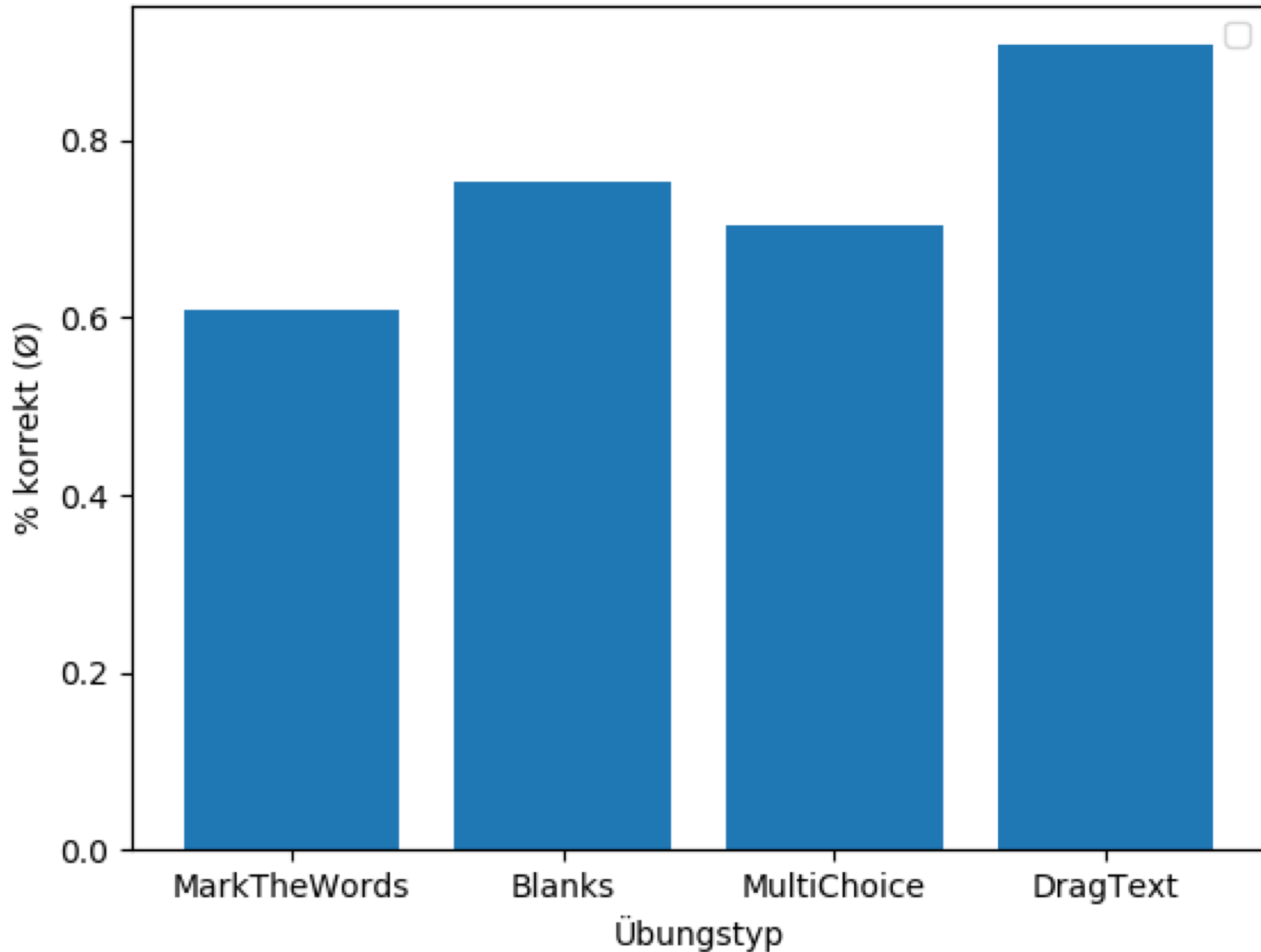
Prozentuale Punktzahl

Hypothese: schwierigere Übungen nötig
(Ceiling-Effekt vermeiden)

Begonnene Übungen

Hypothese: nur für sehr wenige Übungen gibt es verlässliche Daten (d.h. von allen Versuchspersonen)

Durchschnittliche prozentuale Punktzahl je nach Übungstyp

Hypothese 1: MarkWords ist relativ schwer, wenn der Text lang und die maximale Punktzahl hoch ist

Hypothese 2: MultiChoice kann leichter werden, wenn die Aufgabe kleiner ist (z.B. 4 Items inkl. 1 Distraktor)

Hypothese 3: DragText muss schwieriger werden (Ceiling-Effekt vermeiden)

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.