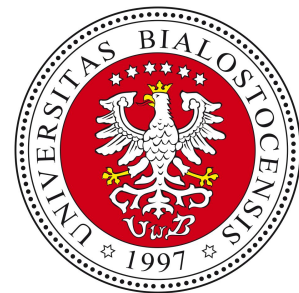


The impact of invisible lexicography on the self-revision of academic English collocations

COLLOcaid

Tomasz Michta



Ana Frankenberg-Garcia



Collocations as a linguistic challenge

For L2 users

- collocations as a challenge for speaking and writing as learners may not be able to piece collocations together from their component parts (Palmer 1933: 14)
- the percentage of deviant collocations does not correlate with years of English in the classroom (Nesselhauf 2004: 235-236)

For L1 users

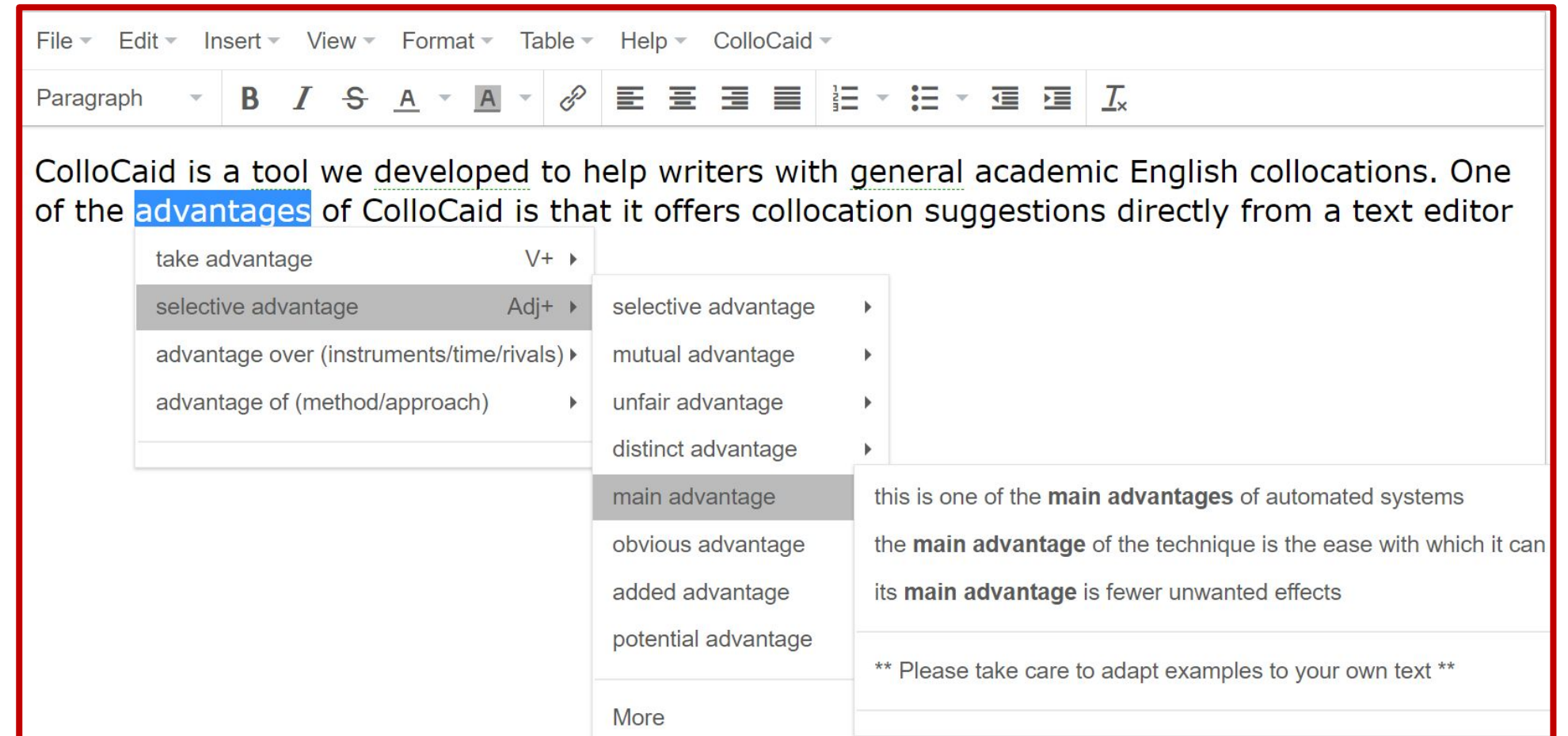
- “L1-English EAP writers tend to be more prone to using general English lexis in academic contexts.” (Frankenberg-Garcia 2018: 103)
- “The main variable affecting the number and variety of collocations available to the participants in the study was their level of academic experience” (Frankenberg-Garcia 2018: 103)

Dictionaries as a solution?

- “the percentage of deviant collocations is exactly the same for those essays where dictionaries were used and for those where they were not” (Nesselhauf 2004: 231)
- “(...) some learners could not find some target collocations in the dictionaries even though they were included there.” (Laufer 2011: 45)
- Learners not always aware of their lexical limitations, so would have no reason to turn to a dictionary (Frankenberg-Garcia 2020)
- Looking up collocations in a dictionary can be particularly disruptive during cognitively demanding tasks, like academic writing. (Frankenberg-Garcia 2020)

ColloCaid: A dictionary goes invisible

- a collocational database that has been experimentally integrated into a text editor (Frankenberg-Garcia et al. 2019, 2021)
- focus on collocations in academic English
- over 32,000 collocation suggestions for 572 of the most frequent nouns, verbs and adjectives plus examples of collocations in context



www.collocaid.uk



Does ColloCaid work?

Frankenberg-Garcia et al. (2019b)

- early versions of the prototype scored between good and excellent on the System Usability Scale (Brooke 2013; Bangor et al. 2009)

Rees (2021)

- ColloCaid found to be less demanding when compared with other collocation tools and dictionaries on the NASA Task Load Index.

Frankenberg-Garcia et.al (2022)

- academic writers and English teachers at Brazilian universities declared they were *very likely* to continue using the tool after a workshop that introduced them to ColloCaid
- one year later, 54% of the writers and 93% of the teachers still used the tool



Our study

Research gap:

- previous studies are based on self-reported perceptions
- no studies so far have looked into the actual revisions by writers using ColloCaid

Aim:

To explore the usefulness of ColloCaid for self-revision of collocations that appear in authentic academic texts written by L2 English students

Research questions

RQ1:

What is the coverage of academic collocations provided by ColloCaid?

RQ2:

What do writers change, and what do they not change?

RQ3:

What motivates writers' decisions to revise (or not) academic collocations?

RQ4:

Do self-revisions with ColloCaid help to improve texts?



Methodology

Participants

- 27 L2 English final-year students of English philology at the University of Białystok
- 18 BA students and 9 MA students
- A range of dissertation topics: linguistics, literature studies and culture studies
- 25 L1 Polish speakers and 2 L1 Ukrainian speakers
- 22 female and 5 male

Procedure

Introduction to ColloCaid

- Video
- Registration

Revision Task

- Approx. 600 word extract from a BA/MA dissertation previously not seen by their supervisor

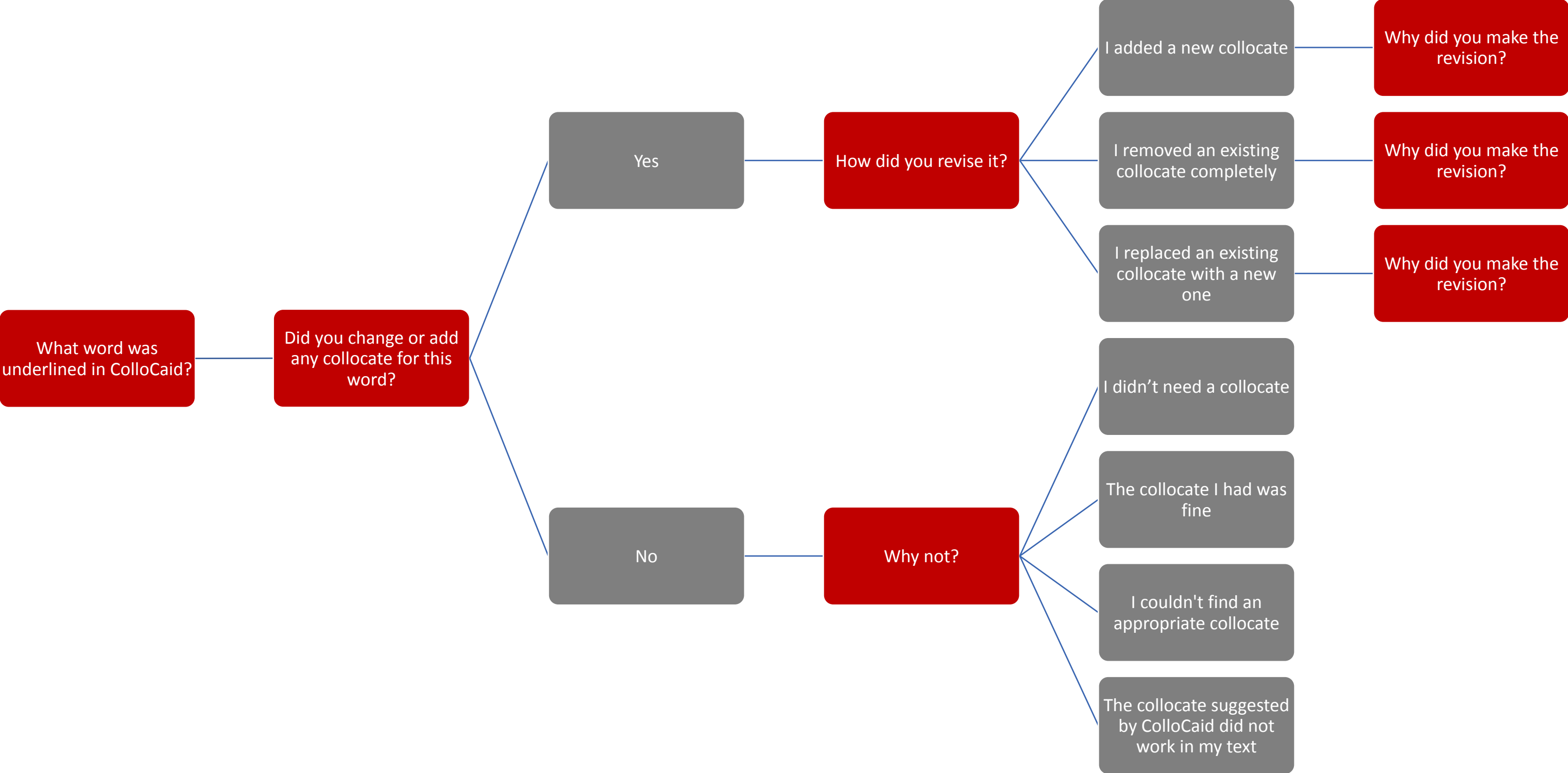
Students self-report on revision

- Survey filled in immediately after the experiment (details in next slide)
- Manual checking of all the revisions reported by the participants

Rating of revisions

- External coders (details coming)

Survey details



External coding

One experienced L1 English proofreader

Second coding forthcoming

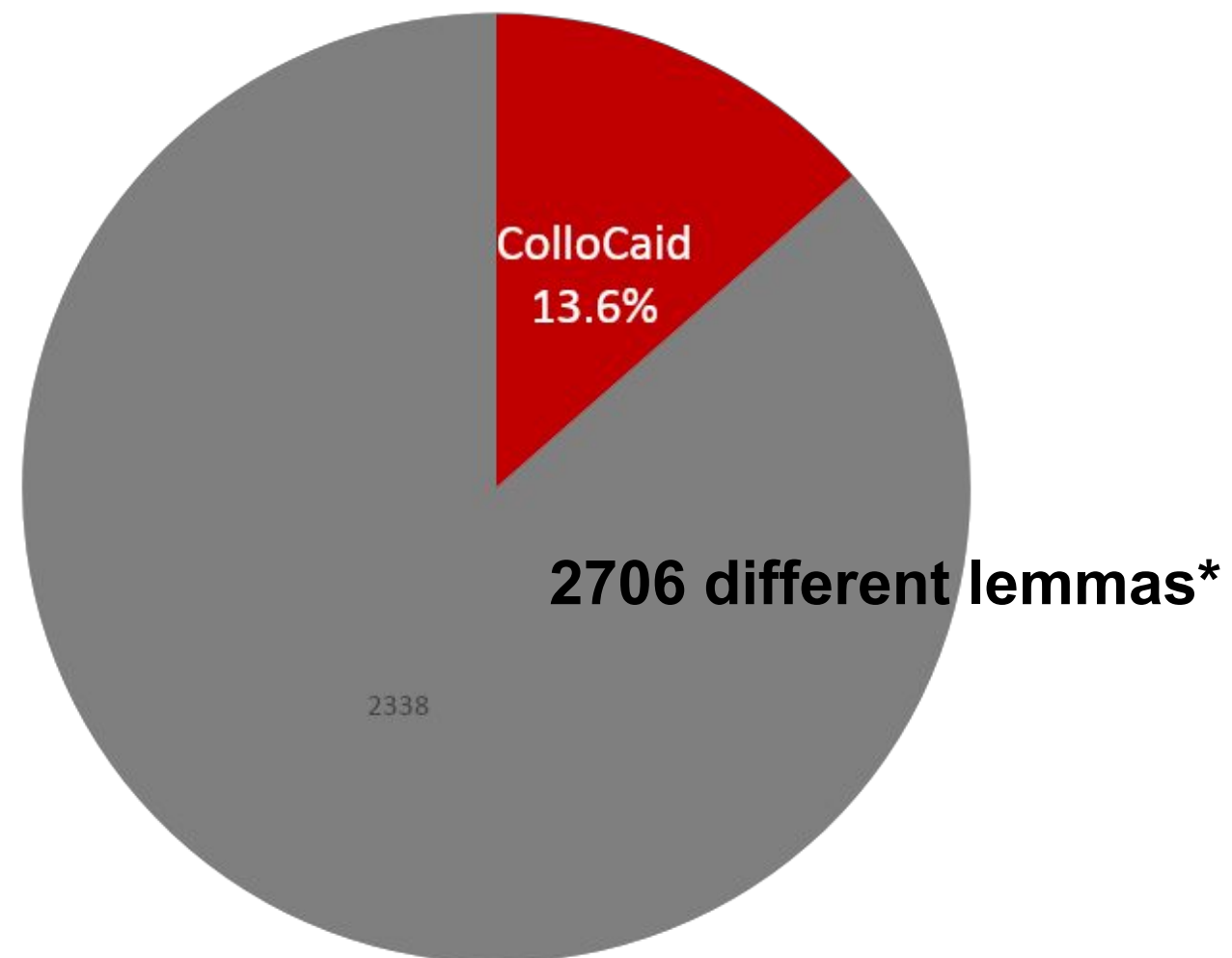
Revision rating taxonomy (Frankenberg-Garcia 1990)

- **positive** (collocation better)
- **negative** (collocation worse)
- **indeterminate** (not enough context to judge)
- **unnecessary** (good collocation replaced with an equally good collocation)
- **ineffective** (bad collocation replaced with an equally bad collocation)
- **consequential** (change because of adjacent changes in text)

Results

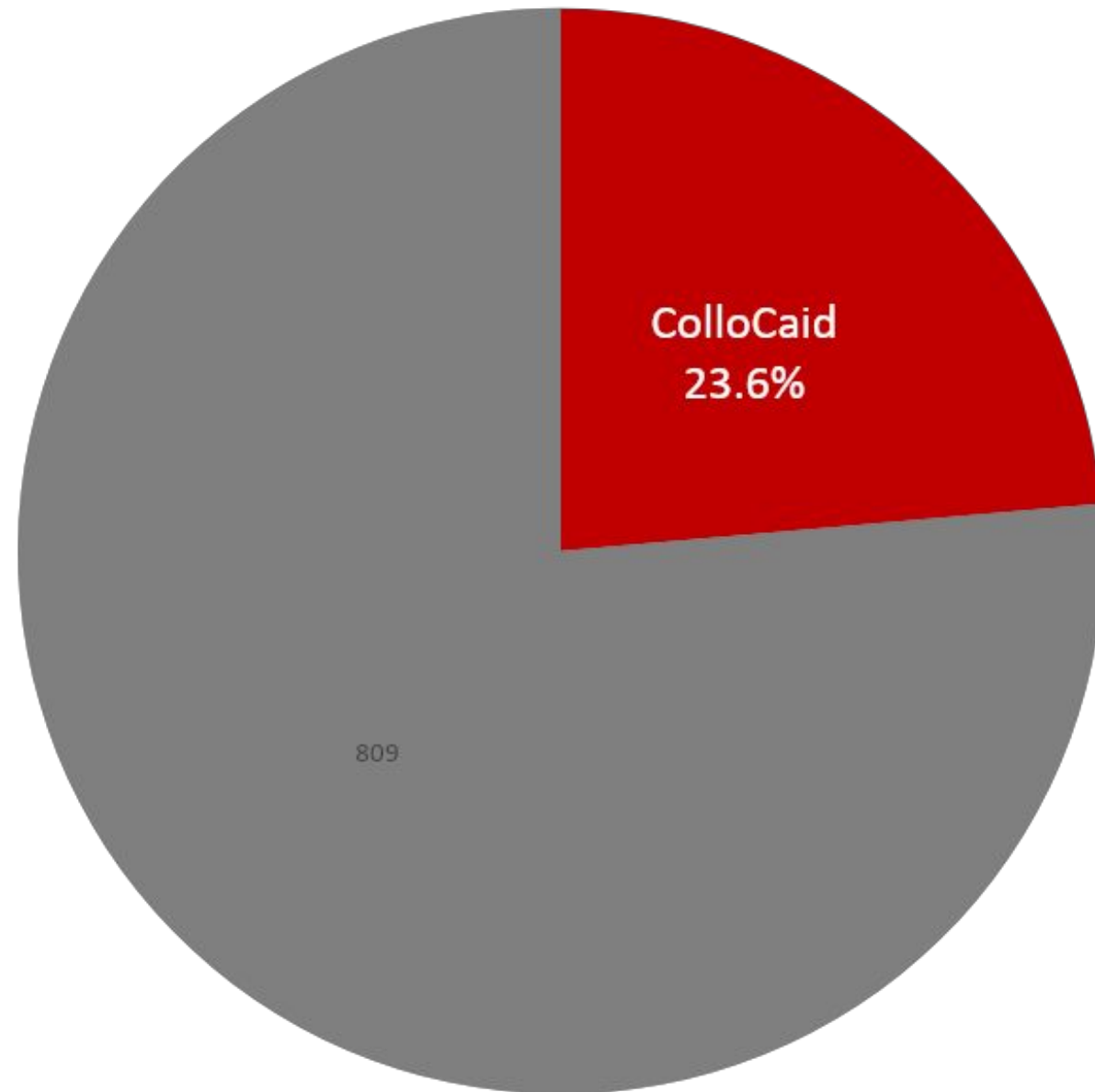
RQ1: ColloCaid Coverage

- Corpus of 16,289 running words on Sketch Engine (Kilgarriff et al. 2004, 2014)
- ColloCaid suggestion every 51.5 words or every 1.8 sentences on average

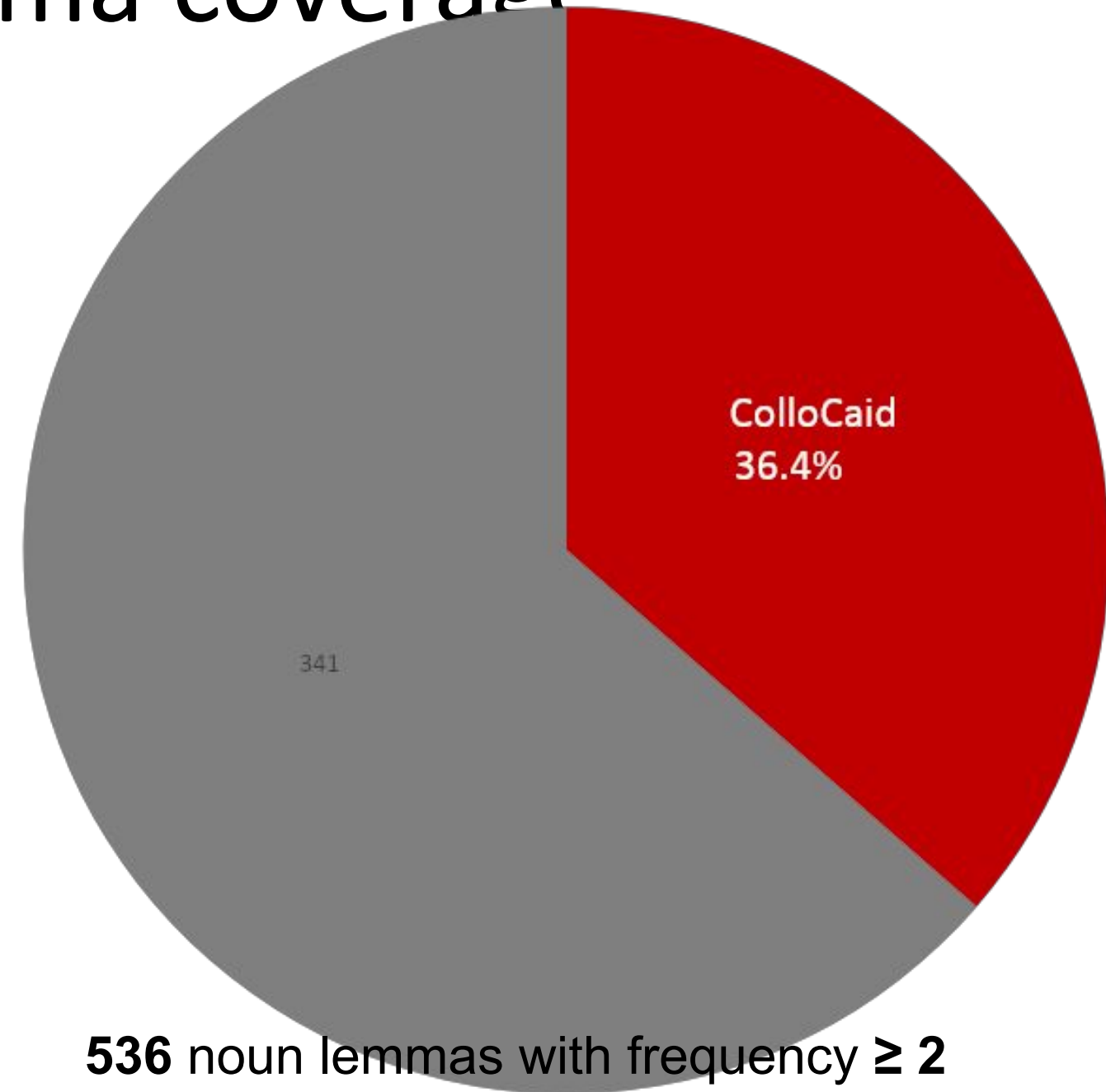


*including articles, prepositions, conjunctions, proper nouns, non-English, non-academic and subject-specific words deliberately not covered in ColloCaid

RQ1: Noun lemma coverage



1059 noun lemmas (excluding proper nouns and non-English words)

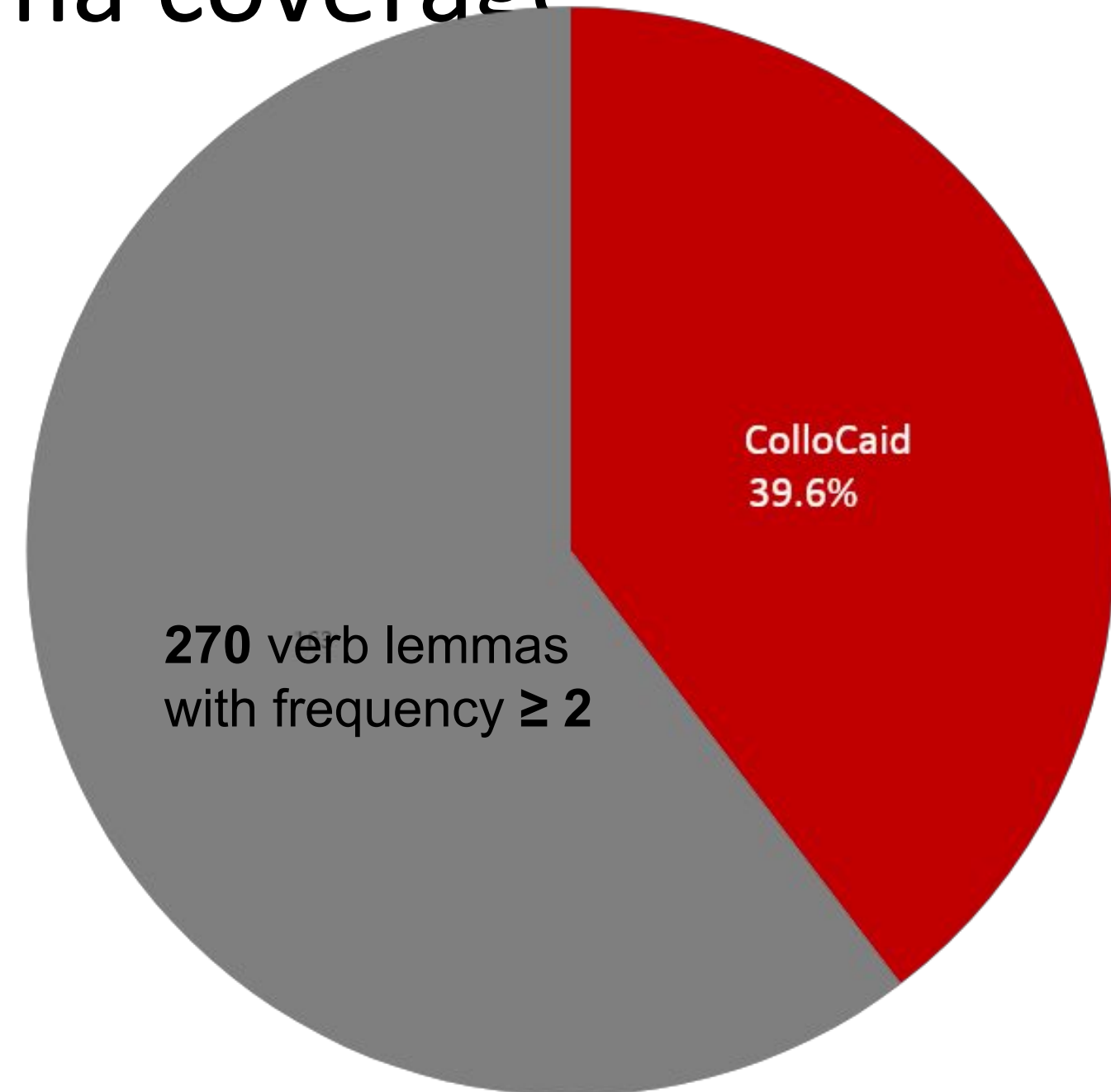
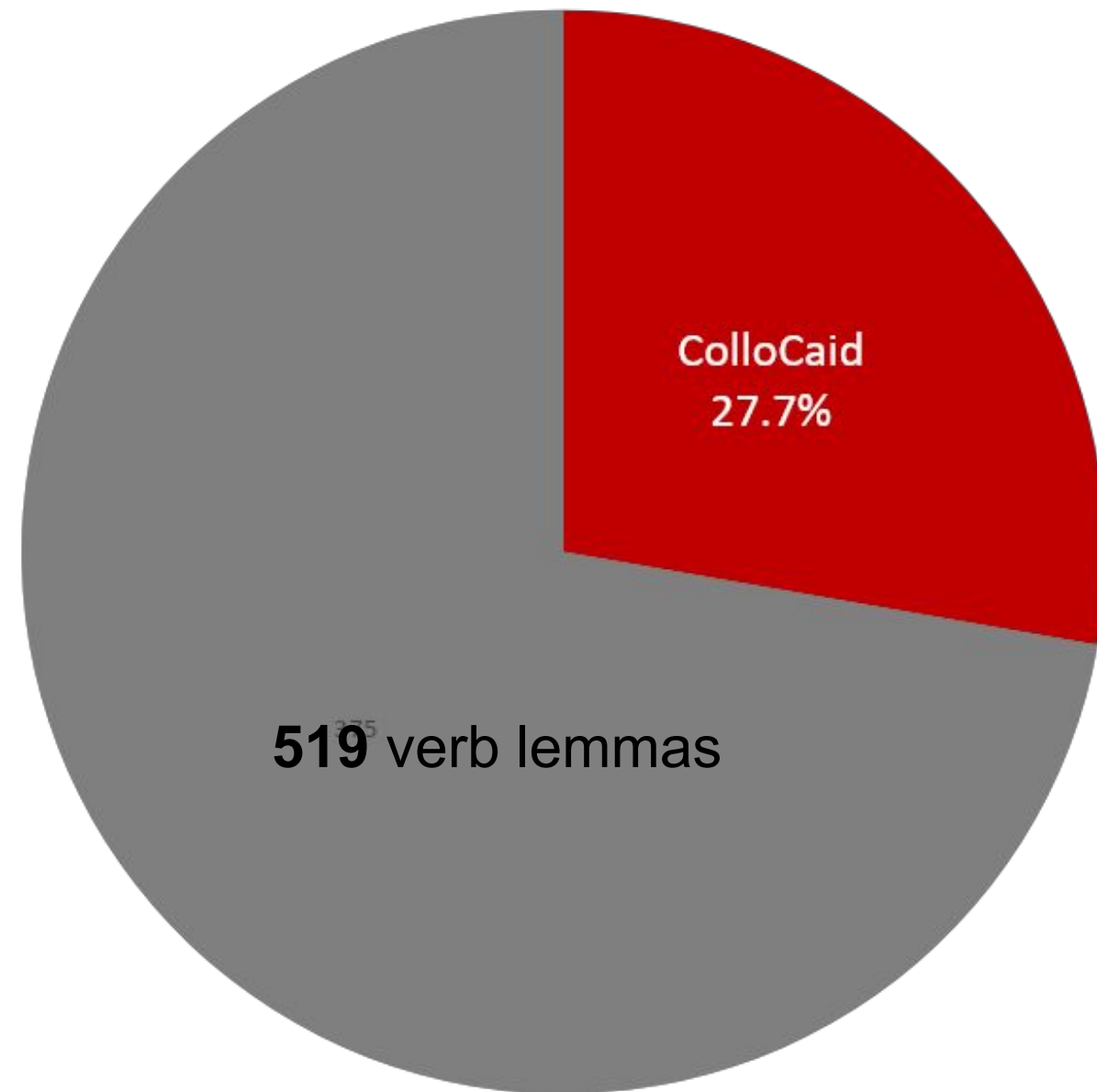


536 noun lemmas with frequency ≥ 2

Types of nouns not covered by ColloCaid

- *vowel, utterance, collocation* (subject-specific)
- *child, artist, people* (general, usually non-academic)
- *user, order, manner* (consider including in future)

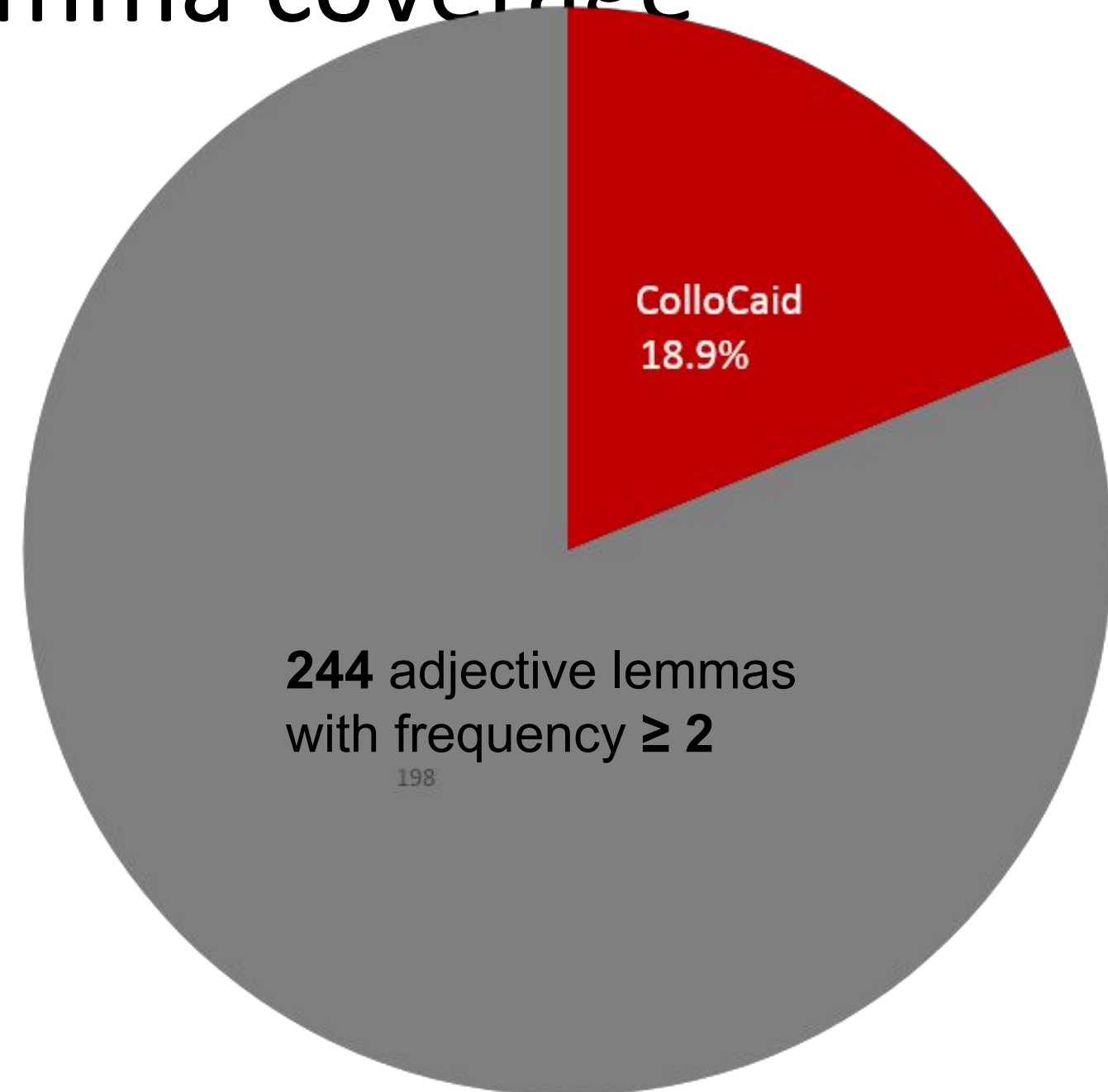
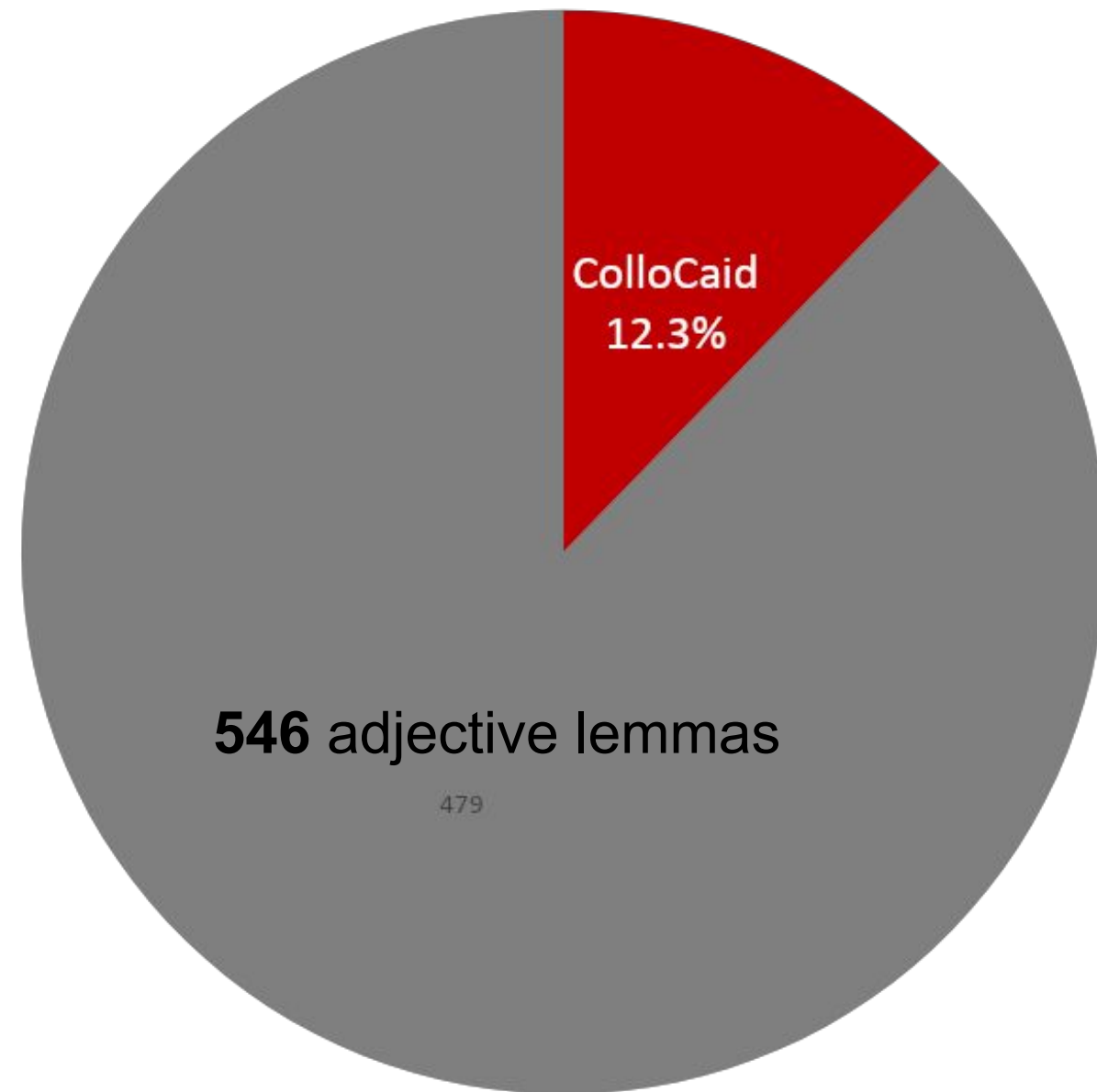
RQ1: Verb lemma coverage



Types of verbs not covered by ColloCaid

- *be, have, do, make, take* (delexicalized verbs)
- *kiss, love, talk, let* (general, usually non-academic)
- *pronounce, collocate, devoice* (subject-specific)
- *convey, perceive, claim* (consider including in future)

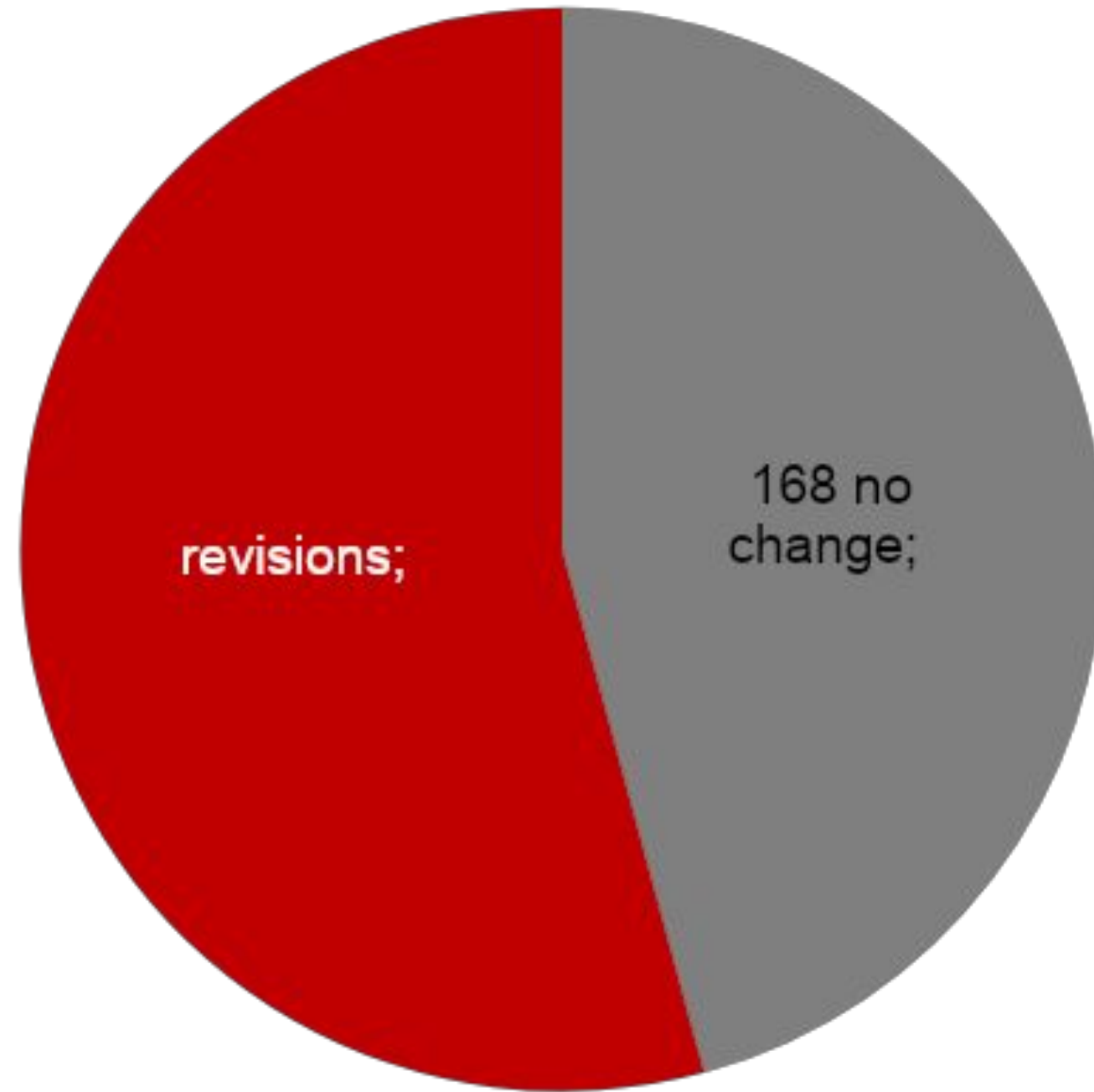
RQ1: Adjective lemma coverage



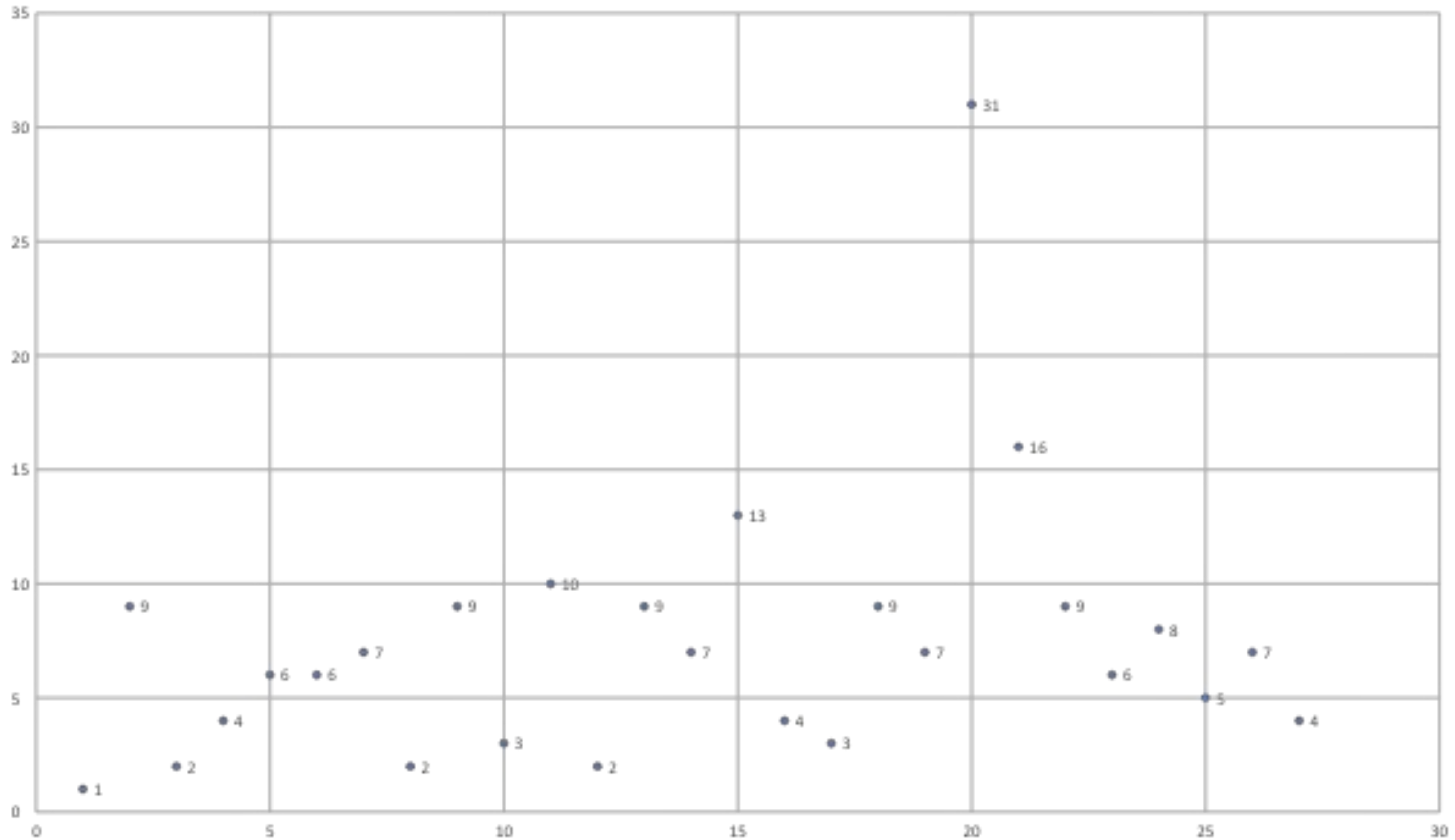
Types of adjectives not covered by ColloCaid

- *main, such, English, female (non-gradable, no collocations)*
- *good, great, young (general, usually non-academic)*
- *lexical, literary, linguistic (subject-specific)*
- *able, large, possible, certain (consider including in future)*

RQ2: What did writers change, and what did they not change?

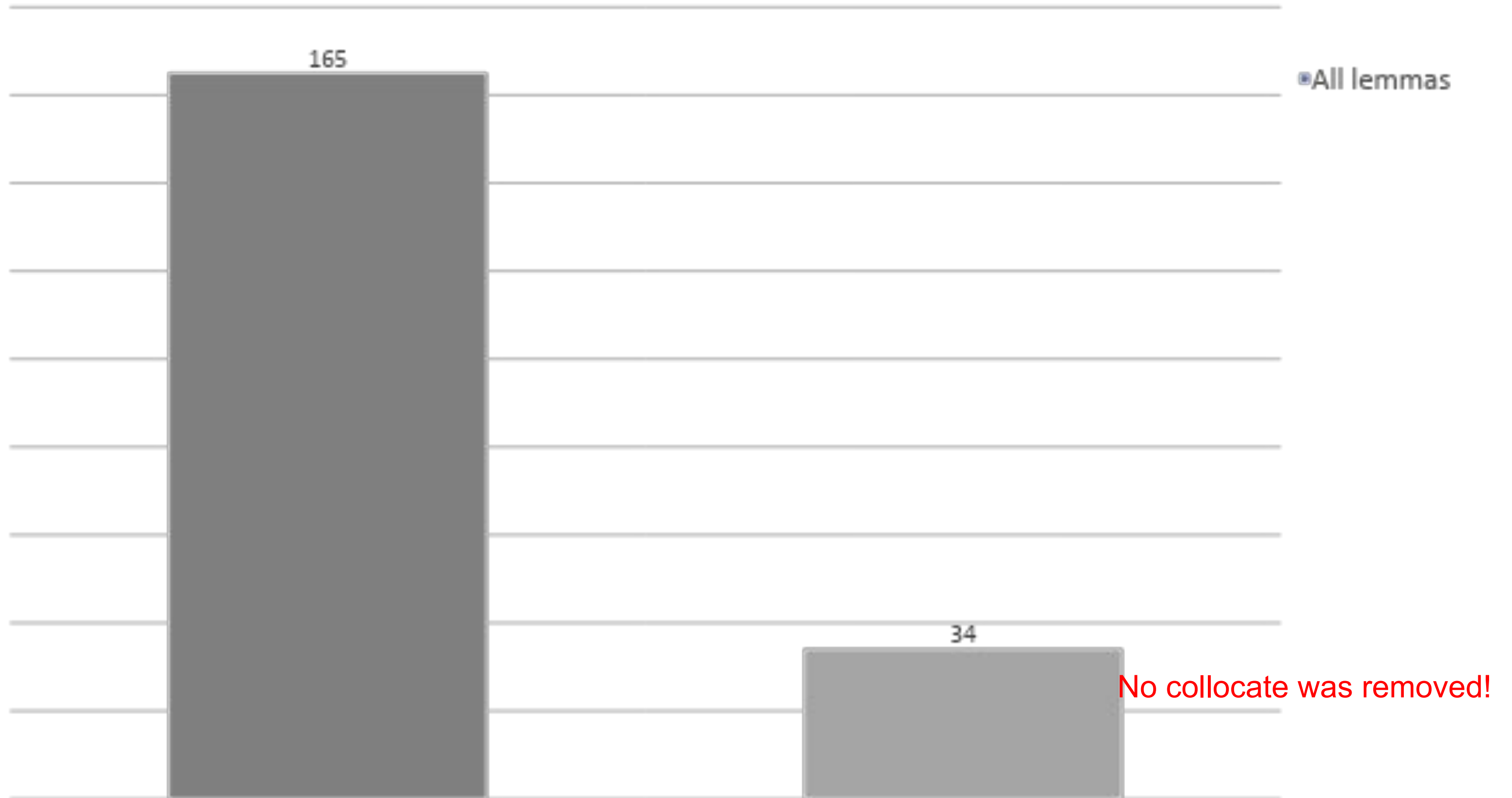


RQ2 - revisions per participant

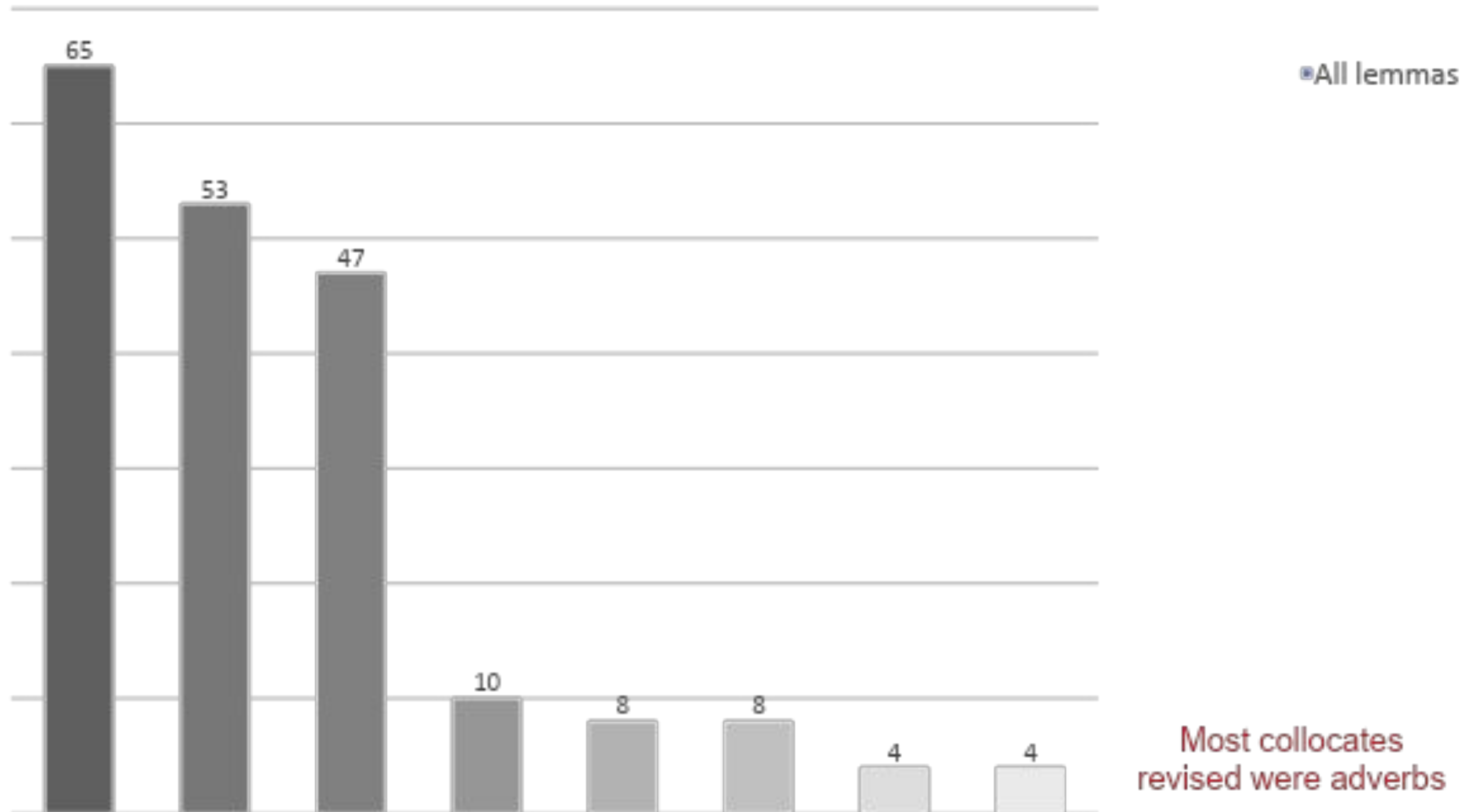


Min = 1
Max = 31
Mean = 7.73
Median = 7
SD = 5.87

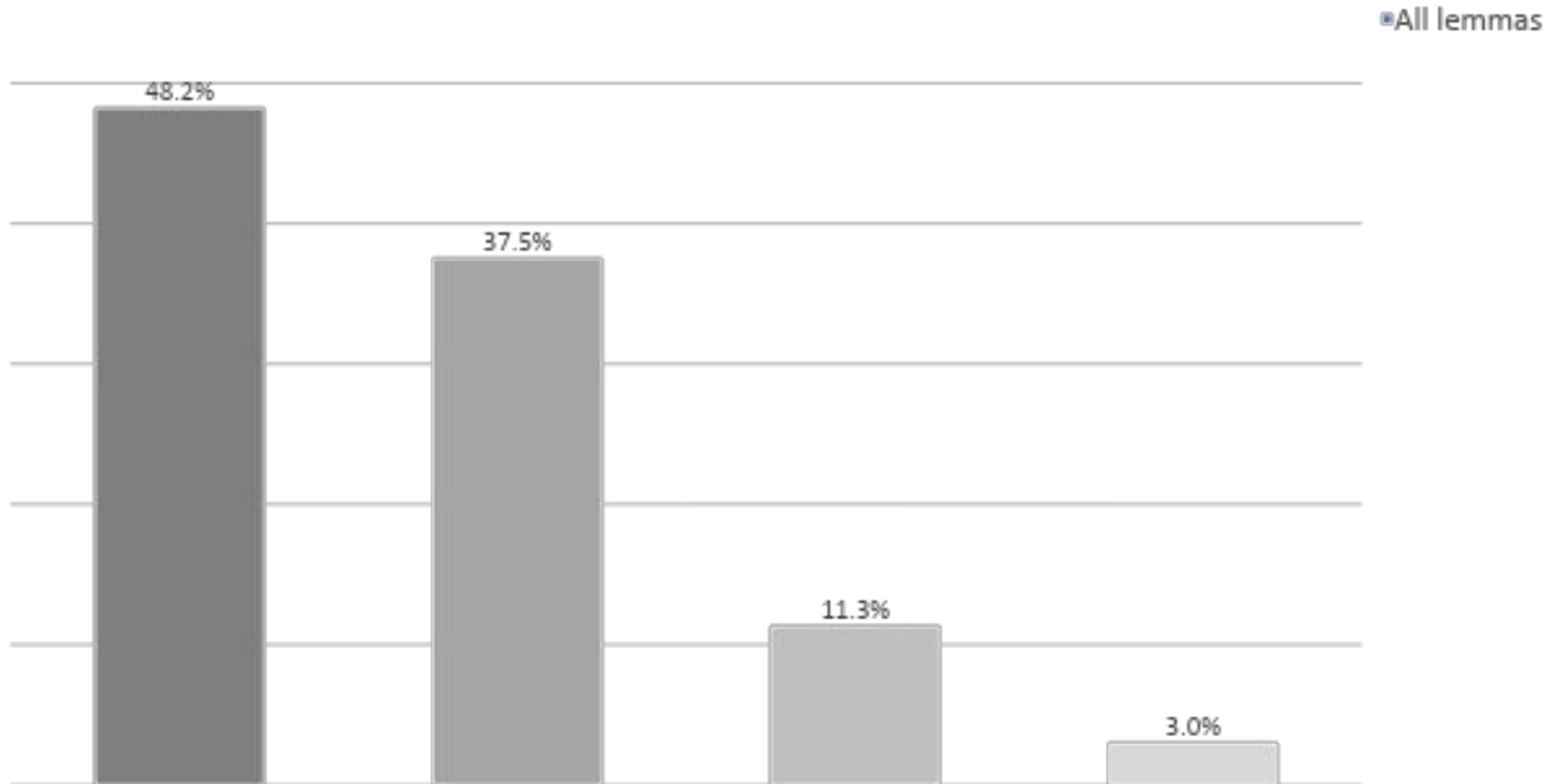
RQ2: How did SS revise collocations?



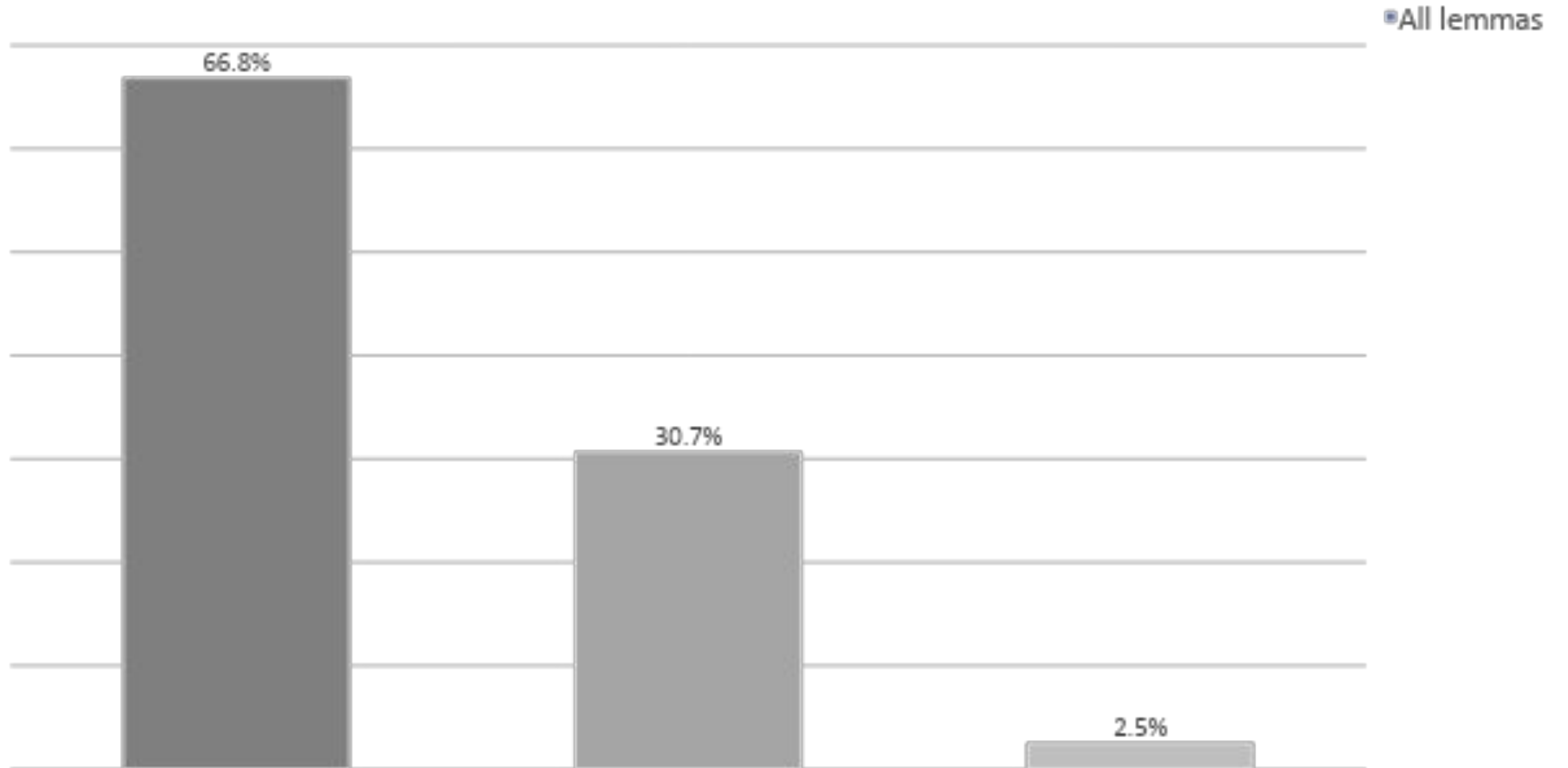
RQ2: Revisions by collocation type



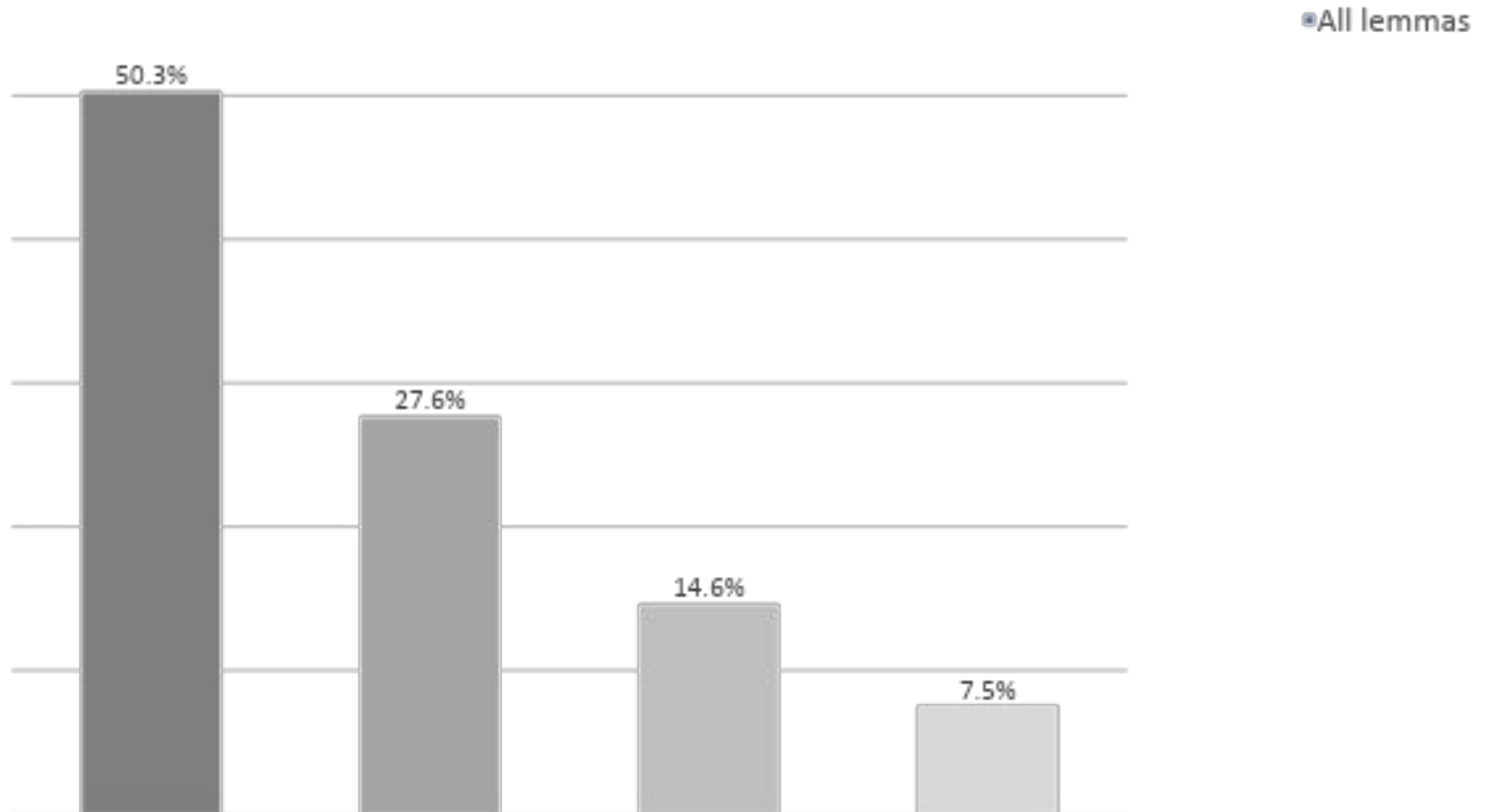
RQ3: motivations for no change (N=168)



RQ3: motivations for revisions (N=199)



RQ4: Revisions by their effectiveness (coder rating)



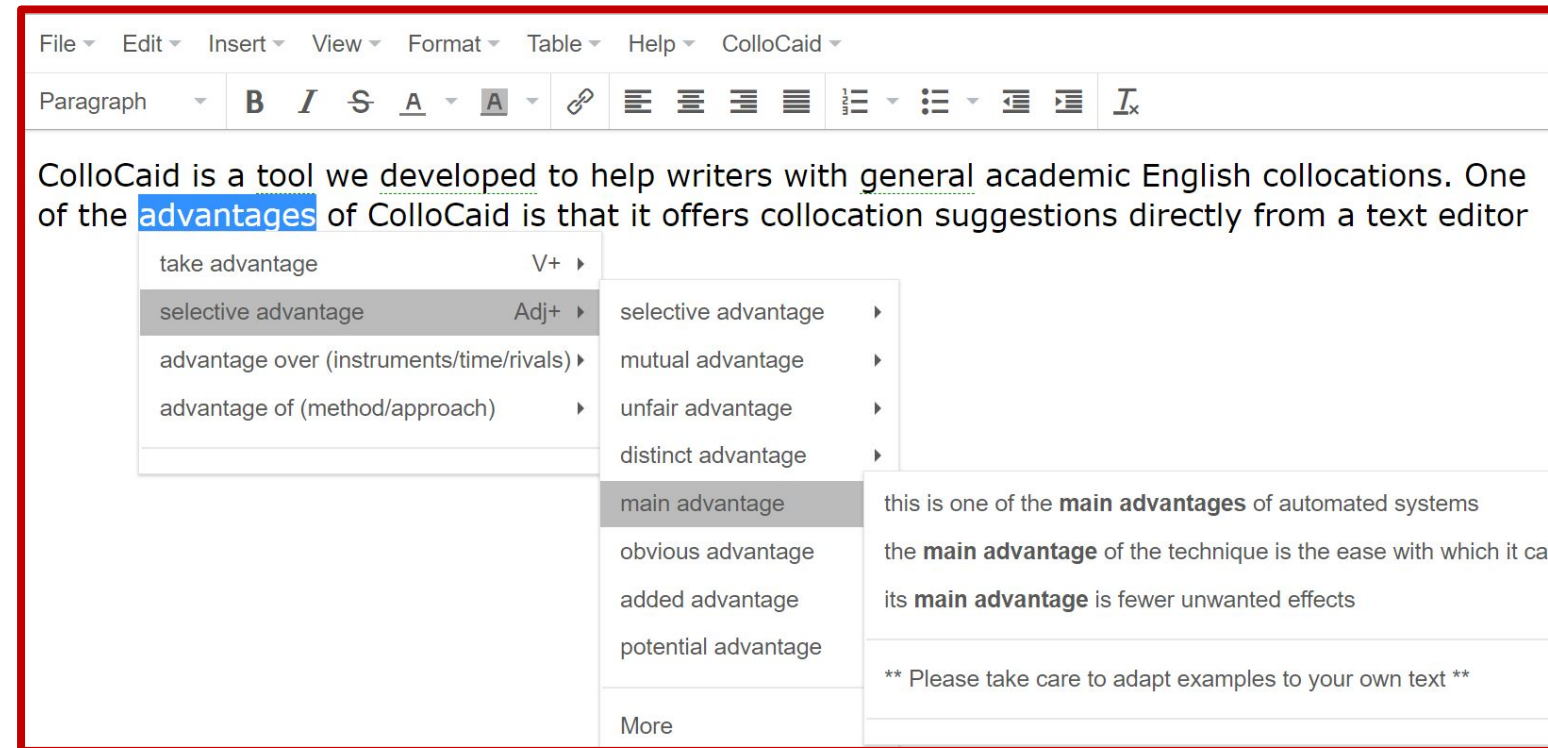
To summarize

- . Findings suggest that ColloCaid aids in improving the quality of student writing
- . ColloCaid offers a good coverage of lemmas (but can still be extended).
- . Participants made on average 7 revisions per 600 words
- . Participants exercised critical judgment in using the tool: did not revise everything underlined in ColloCaid.
- . A majority of revisions:
 - . were motivated by participants' desire to improve their text
 - . involved adding a collocate
 - . especially adverbs (V + Adv collocations)
 - . leading to greater nuance
 - . resulted in improved fluency (according to an external rater)

Further research needed

- . With more and different types of academic writers
 - . other L1 backgrounds (besides Polish/Ukrainian)
 - . diverse levels of academic experience (e.g. PhD students, lecturers and professors)
 - . L1 English users (secondary school students, undergraduates, postgraduate students)
 - . Working in different disciplines (not just Literature, English and Linguistics)
- . To explore how ColloCaid is used when writing academic texts from scratch (rather than revising them).

To conclude



**ColloCaid is an example of invisible lexicography.
We hope to have shown that its effects are clearly visible.**

References

- Frankenberg-Garcia, A. (2018) Investigating the collocations available to EAP writers. *Journal of English for Academic Purposes* 35: pp. 93-104.
- Frankenberg-Garcia, A., Pinto, P.T., Bocorny, A.E. and Sarmiento, S. (2022). Corpus-aided EAP Writing Workshops to Support International Scholarly Publication. *Applied Corpus Linguistics*, 2/3: 1-13. <https://doi.org/10.1016/j.acorp.2022.100029>
- Frankenberg-Garcia, A., Rees, G., Lew, R., Roberts, J., Sharma, N., & Butcher, P. (2019b) ColloCaid: a tool to help academic English writers find the words they need. In Meunier, F., J. Van de Vyver, L. Bradley & S. Thouësny (eds), *CALL and complexity – short papers from EUROCALL 2019* (pp. 1-7). Research-publishing.net.
- Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P. and Suchomel, V. (2014). The Sketch Engine: ten years on. *Lexicography* 1: pp. 7–36.
- Kilgarriff, A., Rychly, P., Smrz, P. and Tugwell, D. (2004). *The Sketch Engine*. In Williams, G. and S. Vessier (eds.), *EURALEX 2004 Proceedings* (pp. 105-115), Lorient.
- Laufer, B. (2011). The Contribution of Dictionary Use to the Production and Retention of Collocations in a Second Language. *International Journal of Lexicography*, 24(1), pp. 29–49.
- Nesselhauf, N. (2004). *Collocations in a Learner Corpus*. Amsterdam/ Philadelphia: John Benjamins Publishing Company.
- Palmer, H.E. (1933). *Second Interim Report on English Collocations*. Tokyo: Kaitakusha.
- Rees, G. P. (2021). Measuring User Workload in e-Lexicography with the NASA Task Load Index. In I. Kosem, M. Cukr (eds.) *Electronic lexicography in the 21st century (eLex 2019): Post-editing lexicography*. Book of abstracts. pp. 54–55.