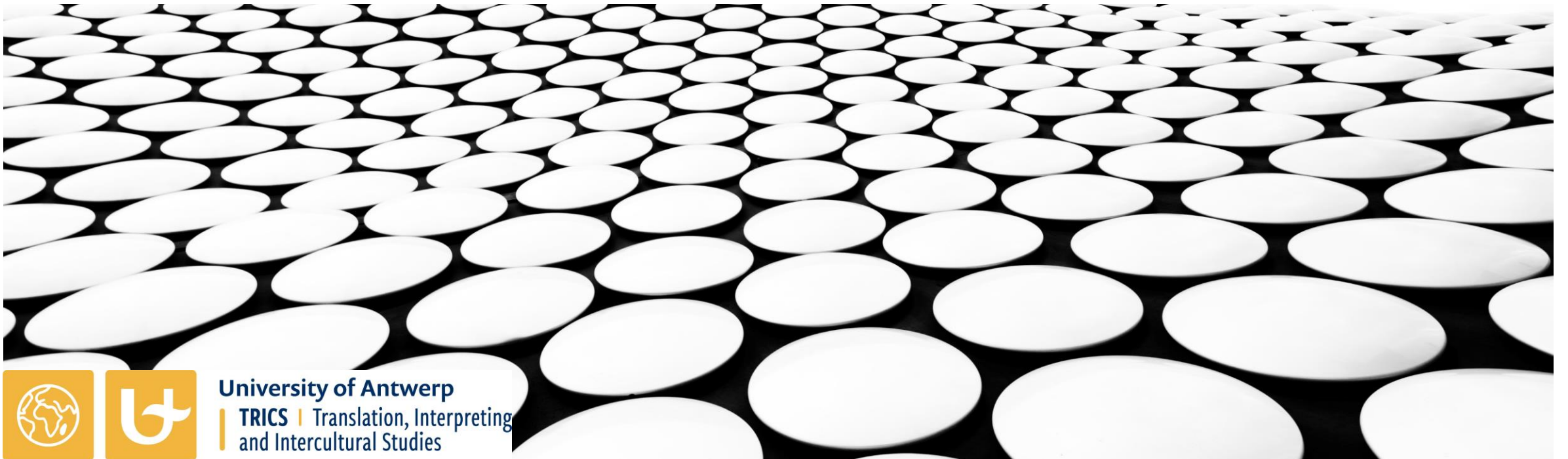

GENERATIVE AI AS A WRITING BUDDY IN THE (L2) WRITING AND TRANSLATION CURRICULUM

WORKSHOP GENERATIVE AI AT FLW – 19 FEB 2024 – CAROLA STROBL



University of Antwerp

TRICS | Translation, Interpreting
and Intercultural Studies

PRELIMINARY THOUGHTS

- (L2) writing and translation pedagogy are **intrinsically intertwined**
- **Translation pedagogy:**

AI-based MT tools trained on LLM since 201x (DeepL, Google Translate, ...)

 - post-editing MT as a (new) task in the translation classroom to promote effective use of MT in translation practice (Balling et al., 2014; Chung, 2020)
- **Writing pedagogy:**

Discussion about integration of digital tools: product- vs. process-oriented approach (Oh, 2022).
Generative AI (GAI) takes writing (and translation) support to a next level (Gayed et al., 2022)

 - Need for new tasks to embed GAI as “writing buddy” into the (L2) writing classroom to promote awareness and effective use (Kasneci et al., 2023)
- **Key question:** WHEN and HOW?

Define necessary skills acquisition level for pedagogically sound integration of MT/GAI

ERASMUS+ PROJECT: **AI WRITE** (2023-2026)

(U HILDESHEIM, ANTWERP, INNSBRUCK, MÄLARDALEN, LIMERICK)

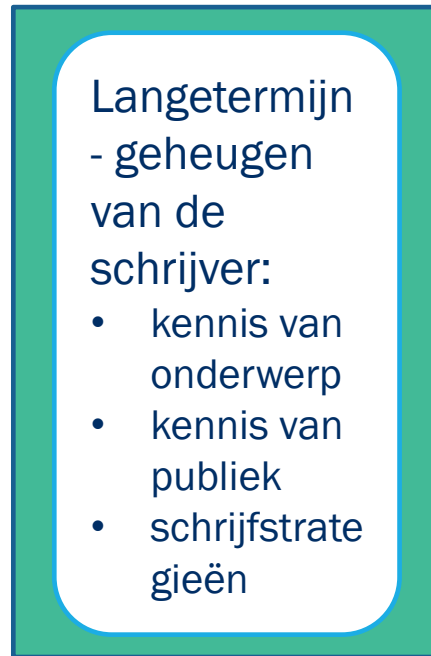


- AI tools for writing: Integration of advanced technology in academic writing within university curricula and student support structures
- Aim: Develop OER for implementation of sound pedagogical practices across European HE institutions
- First Work Package: Literature + tools review to establish state-of-the-art
- **Results:**
 - At this moment, ChatGPT is leading the field (but things may change quickly)
 - GAI tools are mushrooming & increasingly being tailored to specific contexts and needs
 - Translation tools are also being used in HE writing education + vice-versa
 - Expanding body of literature on their integration in HE practice (224 publications)
(experimental studies: writing development + writing evaluation, surveys, practical training)

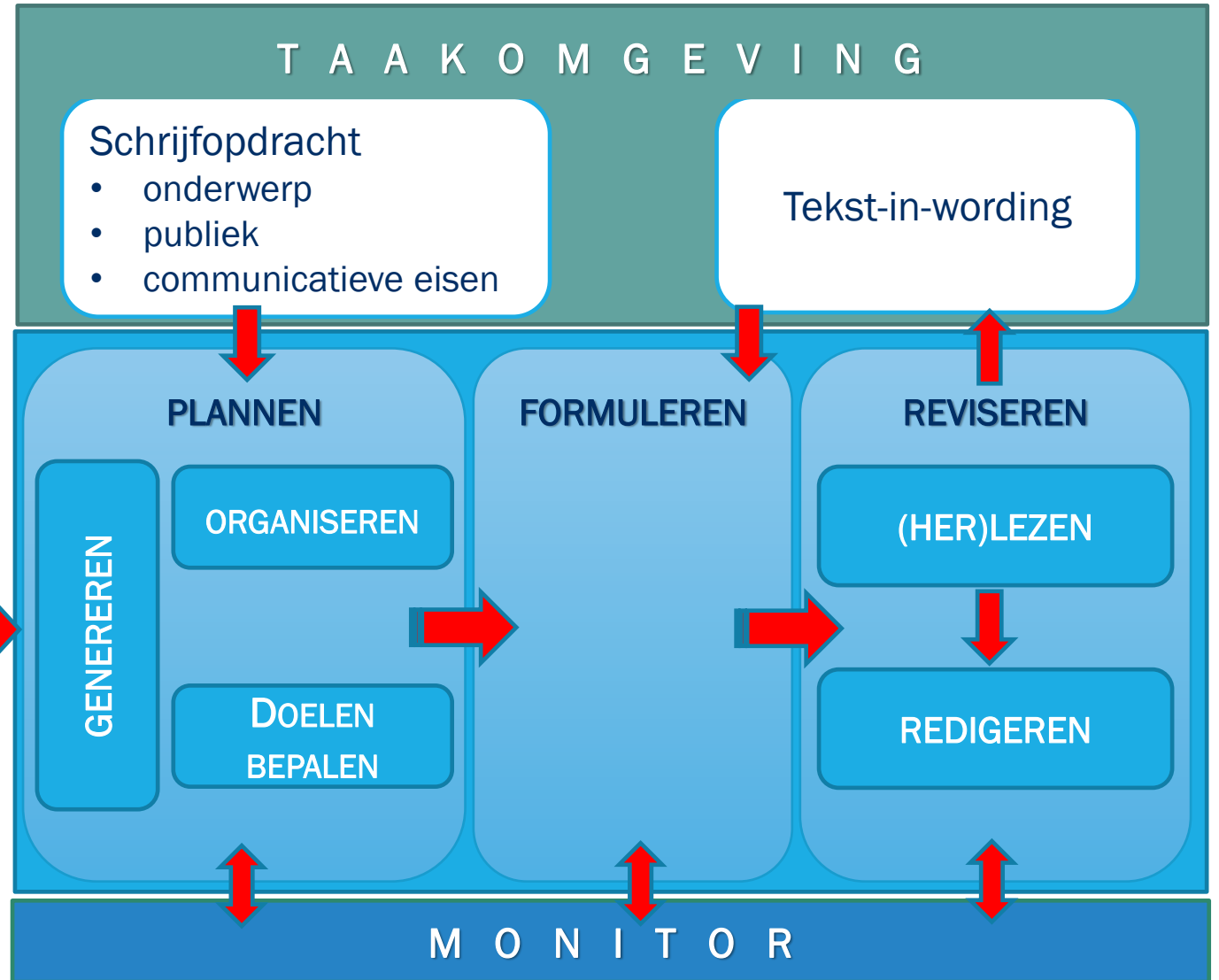
WRITING IS COMPLEX

Several subprocesses & subskills

Cognitively demanding task

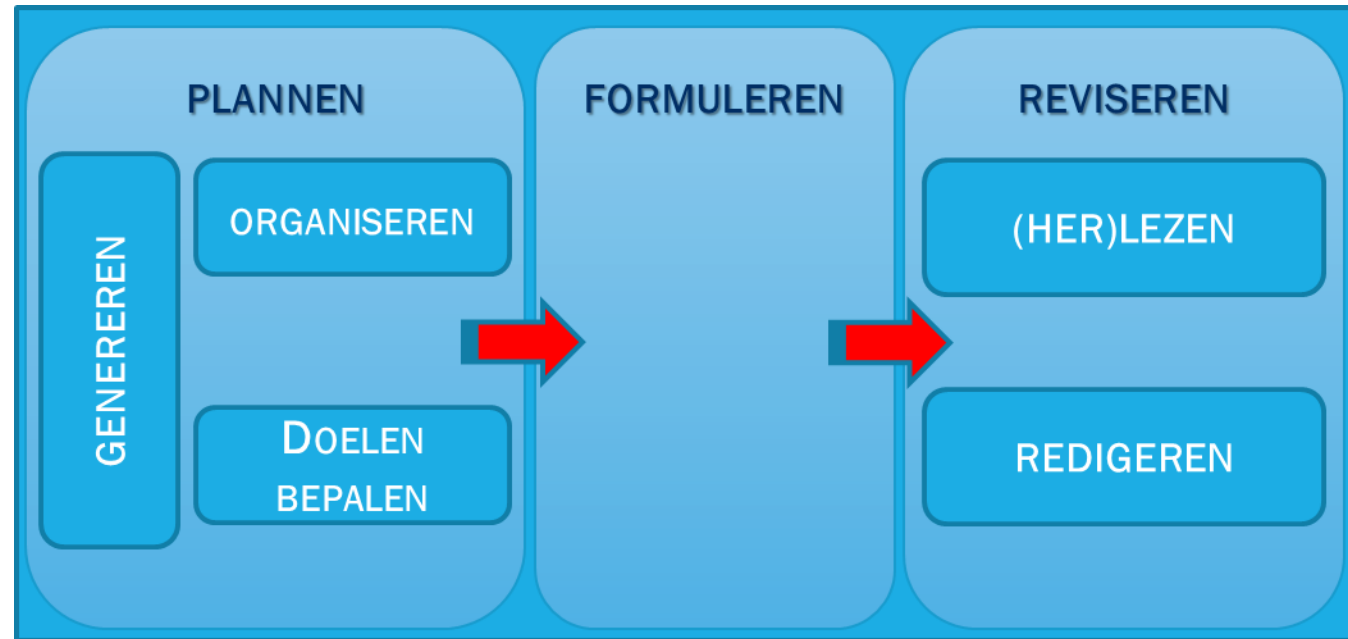


(Hayes & Flower, 1980)



GAI AS BUDDY IN THE WRITING PROCESS

- Support idea generation
- Assist in outline creation
- Evaluate fit of outline with writing goal / audience
- Summarise existing literature
- Rephrasing sources



- Proofreading and editing
- Compare own output with GAI output
- Re-reading and refining output
- Recursive writing

- GAI as dialogical query tool for
 - alternative formulations
 - idiomatic collocations
 - scientific writing style
- Translate ideas from L1 into L2

HOW TO EVALUATE HUMAN & GAI-COLLABORATION?

Process-based evaluation using a (growing) e-portfolio

- Chat history with GAI (prompts, answers): export and organise according to type of query
- All text drafts
- Reflection on changes and edits

Classroom-based feedback on good practices

- Collaboration: let students exchange experiences and good practices / prompts

 **GIGO** (garbage in, garbage out)

CASE STUDY (U ANTWERP & U GRONINGEN)

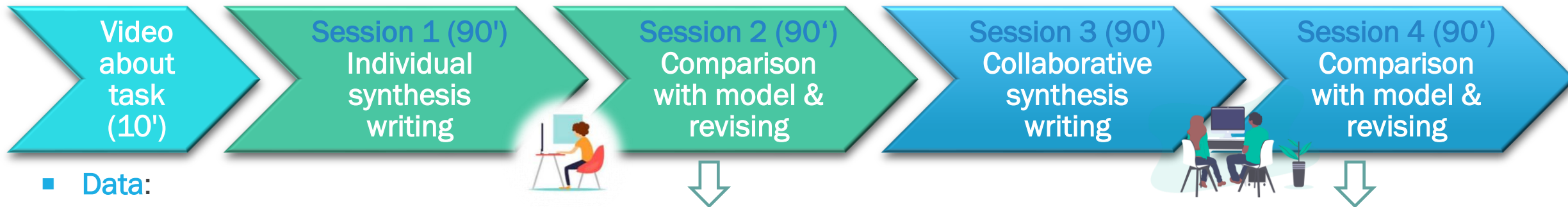
SYNTHESIS WRITING IN L2 GERMAN

Concept: Stimulate 'inner feedback' (Nicol, 2021) through comparison of own text with AI-generated models: **writing > comparing > revising**

- RQ1** What do students **notice** in their own output and in Chat-GPT output based on a guided comparison?
- RQ2** What do students **revise** in their own texts?

METHOD AND DATA

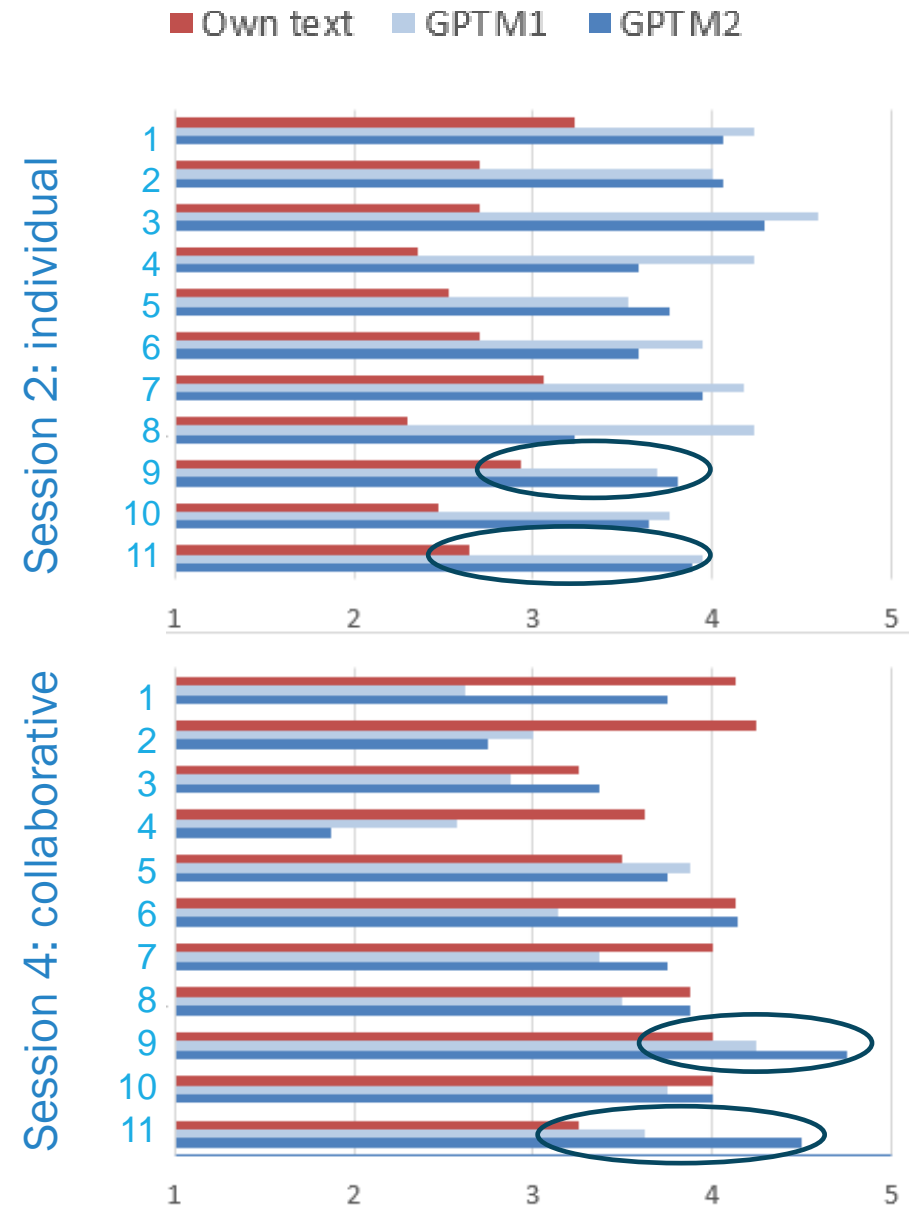
- **Participants:** 22 university students from U of Groningen minoring in L2 German (CEF-levels B2-C1)
- **Task:** Synthesis writing from two popular-scientific source texts on linguistic topics of contemporary German
Environment: Google Docs



- **Data:**
 - **RQ1** “Noticing”: Guided evaluation and comparison of own text with two Chat-GPT models:
 - 11 pre-defined text quality statements (Likert-scale)
 - Free-text comments (three strong + three weak points of the models)
 - **RQ2** “Revision”: Screen-recordings (Screenpresso) and audio-recordings (mobile phones)
9 revision sessions of 6 participants (6 individual and 3 collaborative revisions), 3 coders (Atlas-TI)

RESULTS: GUIDED COMPARISON

1. The synthesis reproduces well the content of both source texts.
2. The synthesis has a clear and logical structure.
3. The introduction summarises the topic of the synthesis.
4. The main body is divided into clear thematic paragraphs.
5. The conclusion clearly rounds off the synthesis.
6. The ideas are clearly linked.
7. The synthesis reads fluidly in one go.
8. The synthesis is reader-oriented: it explains what the reader does not know.
9. The language use overall is correct.
10. The language use overall is varied.
11. The linguistic style is appropriate for an academic synthesis.



RESULTS: FREE COMMENTS ON THE TWO CHAT-GPT MODELS

Strong points

Language use: correct and adequate

In terms of grammar, I would never be able to write such a perfect text containing that many conjunctive and genitive constructions

It is strange that a bot would use humanlike voice, such as “Insgesamt zeigt sich, dass”

Content: good selection

ChatGPT did a much better job than me in selecting the main information of the two source texts

Weak points

Language use: plagiarised from sources, lack of creativity

Given the topic of Kiezdeutsch as a very creative language variety, it is a pity that ChatGPT itself does not use creative language

Content: hallucinations

ChatGPT mentions “die Autorin”, but there is no evidence of the source text being written by a female author.

CONCLUSIONS

RQ1 What do students **notice** in their own output and in ChatGPT output based on a guided comparison?

- Students rated their own output consistently low in terms of linguistic accuracy and appropriate writing style compared with ChatGPT output.
- Students noticed problems with trustworthiness of information of ChatGPT output (Ranalli 2021: “calibrated trust”).
- Overall, students’ confidence with their own text quality compared with ChatGPT output grew during the intervention.

RESULTS: REVISION BEHAVIOUR OF SIX FOCUS PARTICIPANTS

	ALL n=233	Individual mean n=28	Collaborative mean n=20
Revision focus			
• content	30%	32%	28%
• local (word-internal and interpunction)	27%	29%	22%
• lexical choice	14%	12%	20%
• structure	9%	10%	5%
• cohesion	8%	7%	10%
• other (layout, word count)	7%	8%	5%
• grammar (word-external)	6%	4%	13%
Revision necessity	n=222	n=28	n=19
• unnecessary	53%	52%	58%
• necessary	47%	48%	42%
Revision success	n=235	n=27	n=19
• improvement	65%	63%	86%
• neutral	20%	27%	7%
• aggravation	15%	18%	11%

	ALL n=230	Individual n=28	Collaborative n=19
Revision action			
• substitution	38%	46%	19%
• insertion	37%	34%	51%
• deletion	17%	16%	21%
• no action	7%	6%	9%
• move	2%	2%	4%
Revision trigger	n=224	n=27	n=19
• not identifiable	47%	52%	35%
• Google suggestion	29%	35%	16%
• peer discussion	12%	0%	46%
• ChatGPT model	11%	12%	9%
• source texts	0,4%	1%	0%
Information sources	n=231	n=27	n=19
• not identifiable	40%	41%	45%
• Google suggestion	29%	36%	14%
• ChatGPT model	14%	16%	11%
• other online tools	6%	7%	5%
• peer discussion	6%	0%	23%
• Google translate	2%	2%	0%
• other	2%	1%	5%
• Google search	1%	2%	0%

CONCLUSIONS

RQ2 What do students **revise** in their own texts?

- Students revised more in the first (individual) session than in the second (collaborative) session.
- Revision focus is on content (frequently induced by the models) and on local issues (mostly induced by automated Google-suggestions), followed by vocabulary in the third place (↔ literature on model-based revision)
 - Students skillfully draw on their resources for text optimisation.
- More than half of the revisions are unnecessary (“overrevisions”), still often causing text improvement.
- High number of unidentified revision triggers
 - “inner feedback” (Nicol, 2021) was stimulated

MY FAVOURITE QUOTE OF A COLLABORATIVE REVISION SESSION

Ann* [referring to a model]: I like this sentence. Should we just copy-paste it into our text or try to rephrase it?

JOS*: Just copy-paste it! If ChatGPT can do this, we also can.

*pseudonyms



THANK YOU!

Carola Strobl

Associate professor in Applied Linguistics and Translation

University of Antwerp

Department of Applied Linguistics, Translation and Interpreting

Spokesperson research group [TricS – Translation, Interpreting and Intercultural Studies](#) / https://twitter.com/TricS_research
carola.strobl@uantwerpen.be
<https://www.uantwerpen.be/en/staff/carola-strobl/>



Strobl et al. (in press):
Adopting ChatGPT as a writing buddy in the advanced L2 writing class.
Technology in Language Teaching & Learning.



Iryna Menke-Bazhutkina
German language teacher and coordinator



Niklas Abel
German language teacher



Marije Michel
Full Professor -
Chair of Language Learning



University of Antwerp
| TRICS | Translation, Interpreting and Intercultural Studies



university of
 groningen

faculty of arts