

Battuguldur Tuyatsetseg

Below is an opponent-style review and point-based evaluation of *Tuyatsetseg.pdf* using the criteria and grading scale shown in *opponensi_lap_en.pdf*.

0) Use of AI & suspected plagiarism (separate, non-point criterion)

What the thesis does well (AI transparency)

- The thesis contains a dedicated annex on **LLM use**, stating that a model (named as **Anthropic Claude**) was used mainly for **academic English phrasing**, discussion of OAM interpretation, and Flask-structure advice, and that **LLM text was not incorporated verbatim**; the author also includes a **responsibility statement**.

Red flags / inconsistencies (raise in defense)

- In the **Hungarian abstract**, the prototype link appears with a tracking parameter: `?utm_source=chatgpt.com`, which is a typical artifact of copying a ChatGPT-produced hyperlink. This is **in tension** with the annex narrative that only Claude was used and that no verbatim LLM output was pasted. It is not proof of misconduct, but it **warrants clarification**.
- There are **internal inconsistencies** that often arise in AI-assisted or heavily copy-edited writing:
 - **Direction identifier conflict** for A2 (Reading speed): earlier it is treated as **minimise** (Dir=1), later a dictionary table lists it as **maximise** (Dir=0).
 - **Scale mismatch**: many UX attributes are presented as **1–10** in the main evaluation tables, but later the attribute dictionary describes **Likert 1–5** and mental load **1–20**.
- Citations in the body are sometimes formatted as bracketed references with empty links (e.g., ``), which may be an export/conversion issue, but in final submission it counts as **formal referencing weakness**.

Plagiarism likelihood (what can and cannot be concluded here)

- Based on document inspection alone, I **cannot** confirm plagiarism. No definitive “copied-from-source” matches are established in this review (a proper check requires similarity software and/or systematic source comparison).
- However, the **style uniformity**, **template-like phrasing**, and the **ChatGPT tracking parameter** justify:
 - a. running an official similarity check (Turnitin, etc.), and
 - b. asking the student to **show drafts / prompt logs / commit history** to evidence authorship of key parts (implementation + evaluation).

Bottom line: AI use is partly documented, but the thesis contains artifacts and inconsistencies that should be explicitly addressed at defense.

1) Topic & Objectives (max 40)

Strengths

- The topic is **clearly defined and relevant** to informatics and HCI: designing and evaluating an **RSVP speed-reading web application**.
- The thesis provides a **clear aim** and a well-structured list of **specific objectives** and **tasks** aligned with development + evaluation.
- Target groups and expected utility are explicitly stated, with reasonable scope limitations (e.g., no clinical claims).

Weaknesses / deductions

- The novelty is **incremental**: RSVP tools exist generally; the distinctive element here is mainly the **particular prototype + OAM-based evaluation framing**, not a new RSVP theory or controlled experimental contribution.

Score: 34 / 40

2) Review of the Literature (max 40)

Strengths

- Literature chapter is **broad and logically segmented**: testing concepts, usability foundations, RSVP theory, digital reading cognition, and OAM/COCO methodology.
- Uses **recognizable core sources** (e.g., RSVP/speed-reading debates, usability heuristics, digital vs print reading).
- The thesis explicitly connects literature to the chosen **evaluation KPIs** and design decisions.

Weaknesses / deductions

- Reference base is **heavily pre-2020** (the thesis itself reports ~73% "old" sources), which is understandable in classic cognitive literature but still suggests **limited coverage of newer HCI/reading-interface work**.
- Some **formal and internal inconsistencies** reduce credibility:
 - Reference-count statements contradict themselves ("41 references" vs a later mention of "same 31 references").
 - In-text citations appear with empty hyperlink placeholders, which in a final thesis is a **formatting/referencing defect**.

Score: 30 / 40

3) Presentation of the Author's Own Work (max 60)

Strengths

- A working system is described with **clear architecture**: Flask backend for parsing TXT/PDF/DOCX, frontend in HTML/CSS/JS, client-side RSVP playback logic, and JSON token transfer.
- Implementation details are sufficiently concrete (WPM → interval formula; ORP-like highlighting logic; controls; progress estimation).
- Evaluation is **structured** (12 participants × 12 attributes) and includes OAM tables, COCO estimation scores, ranking, and a discussion of limitations (sample size, no baseline, self-selected speed, etc.).
- Security considerations are at least acknowledged (file validation, no persistent storage claim, HTTPS via hosting).

Weaknesses / deductions

- Methodology rigor is **prototype-appropriate but not research-strong**:
 - no control/baseline (normal reading),
 - self-selection confound for WPM,
 - small sample size,
 - comprehension operationalized as short post-test (surface level).
- The evaluation specification suffers from **consistency errors** (direction IDs and measurement scales), which weakens the "objectivity" claim of the OAM pipeline unless corrected.

Score: 46 / 60

4) Structure, Style, and Formal Aspects (max 40)

Strengths

- Structure is clear: intro → literature → development/evaluation → discussion → conclusions → future work → annexes.
- Language is generally fluent, academic, and readable (especially in English sections).

Weaknesses / deductions

- **Formal issues** appear repeatedly:
 - inconsistent scales and direction IDs across tables,
 - citation formatting issues (empty links),
 - some annex/table formatting is messy in the PDF rendering, and images appear as placeholders in the extracted text.
- The **ChatGPT tracking parameter** in the abstract link is an avoidable formal/professional mistake.

Score: 28 / 40

5) Overall Impression (max 20)

Overall strengths

- Solid bachelor-level outcome: **implemented prototype + structured evaluation + honest limitations.**

Overall concerns

- The thesis is weakened mainly by **formal inconsistencies** and **AI-transparency artifacts** that need clarification.

Score: 14 / 20

Point Summary & Grade (per opponensi_lap_en scale)

	☰ Criterion	☰ Max	☰ Awarded
1	Topic & Objectives	40	34
2	Review of the Literature	40	30
3	Presentation of Own Work	60	46
4	Structure, Style, Formal Aspects	40	28
5	Overall Impression	20	14
6	TOTAL	200	152

Final grade band: Average (140–159 points) according to the form's grading scale.

Recommended questions for the defense (focused on risk areas)

1. **AI transparency:** Why does the abstract contain a `utm_source=chatgpt.com` link? What tools were used exactly, and where is this documented?
2. **Evaluation validity:** Please reconcile the A2 direction ID contradiction and explain which specification was actually used in

the COCO computation.

3. **Measurement scales:** Are the UX ratings 1–10 or 1–5? Why does the dictionary differ from the OAM table?
 4. **Reproducibility:** Can you show the OAM workbook and demonstrate how the COCO ranks were generated from it?
 5. **Prototype verification:** Walk through ingestion of PDF/DOCX and show how tokenization avoids losing content (headers/footers, hyphenation, etc.).
-

Quick, practical improvement list (if revisions are allowed)

- Remove tracking parameters from URLs; standardize prototype URL(s).
 - Fix the A2 direction inconsistency and harmonize **all attribute scales/units** across tables and annex dictionary.
 - Ensure citation formatting is consistent and complete in the final PDF (no empty-link placeholders).
 - Add a short appendix note: “Where AI was used” + “exact prompts and outputs” (even a minimal log) to fully align with transparency.
-

If you want, I can also **fill the opponensi form textually** (as a ready-to-paste “General summary” paragraph under each heading), matching the style of your departmental evaluations.