

Based on the provided prototype screenshots and user research documents, here is the expert HCI and Human-AI interaction evaluation.

EVALUATION FORM

A) Session info

- **Expert:** Gemini
- **Expertise:** HCI / UX / AI Interaction Design

B) Task outcomes (quick status)

Task	Status	Severity (0–4)	Notes
T1: Onboarding	Issues	2	No visible onboarding/framing of AI capabilities. Language direction is clear, but user settings (proficiency, goals) are hidden or assumed.
T2: Highlighting	OK	1	Color coding effectively links source and target. P1 benefits from this visual clarity for speed.
T3: AI Cues	OK	1	"Language Insight" panel provides good contextual anchoring. The visual diagram for "UP" aids immediate recognition.
T4: Expansions	Issues	2	Expansions like "UP: similar uses" create a very long vertical scroll. P1 might find this "long technical explanation" frustrating.
T5: Feedback	OK	1	Granular "Useful/Not Useful" buttons are present, supporting trust calibration (G15).

C) Nielsen heuristics checklist

- **[X] H1 Visibility of system status:** The system clearly highlights which text segment is currently being analyzed in the side panel.
- **[X] H3 User control and freedom:** Users can easily close the Language Insight panel using the "X". However, there is no way to "undo" or minimize specific expansion sections.
- **[X] H4 Consistency and standards:** The interface follows standard MT layouts (Left Source / Right Target).
- **[X] H7 Flexibility and efficiency of use:** The "more examples" expansion supports **P2**'s need for ready-made sentences , but may be too much info for **P1**.
- **[X] H8 Aesthetic and minimalist design:** The Language Insight panel becomes cluttered when multiple expansions are open.

Evidence: The expansion of "UP: similar uses" results in a dense gallery of images and text that may overwhelm users seeking quick, actionable info.

D) Human–AI checklist (Amershi et al.)

- **G1 Make clear what the system can do: Partial.** While translation is obvious, the pedagogical/insight purpose is only revealed upon clicking highlights.
- **G4 Show contextually relevant info: Yes.** The insight is directly anchored to the selected phrasal verb.
- **G8 Support efficient dismissal: Yes.** Clear "X" to close the panel.
- **G11 Make actions/outcomes understandable: Yes.** It explicitly states *why* "grow up" was used (numerical growth of populations).
- **G15 Encourage granular feedback: Yes.** Useful/Not Useful buttons allow for local AI correction.

E) Persona impact tagging & Walkthrough summary

Issue	Severity	Persona Impacted	Rationale
Scroll Fatigue	2	P1	Anna needs "quick, short info" ; long expansions of "UP" meanings slow her down.
Technical Depth	1	P2, P3	Carlos and Marta value terminology accuracy and benefit from the "similar uses" nuance.
Language Profile	2	All	The system assumes B1 proficiency ; no visible toggle to adjust insight complexity for a C1 user like Marta.

Light walkthrough summary:

- **P1 (Anna):** Focuses on T2/T3. Highlighting helps her cross-reference quickly. She likely ignores deep expansions (T4) due to time pressure and a dislike for technical complexity.
 - **P2 (Carlos):** Uses T4 extensively. He checks "more examples" to ensure the phrasal verb fits the formal tone required for his technical reports.
 - **P3 (Marta):** Engages with "similar uses" (T4) to find alternative phrasal verbs. She uses the "Useful" button (T5) to help the AI learn her preference for non-literal translations.
-