

Advancing Verbal Communication Skills Through
Transcript-Based Feedback :
Designing an AI-Assisted Language Learning Application,
“Stepping Stones,” to Support ESL Students in
Academic and Professional Contexts

Jeein Iris Bae

A thesis
submitted in partial fulfillment of the
requirements for the degree of
Master of Design

University of Washington
2025

Committee:
Axel Roesler
Karen Cheng

Program Authorized to Offer Degree:
School Of Art, Art History, Design

©Copyright 2025

Jeein Iris Bae

University of Washington

Abstract

Advancing Verbal Communication Skills Through Transcript-Based Feedback :
Designing an AI-Assisted Language Learning Application, “Stepping Stones,” to Support
ESL Students in Academic and Professional Contexts

Jeein Iris Bae

Chair of the Supervisory Committee:

Axel Roesler

Department of Design

This thesis investigates communication challenges faced by English as a Second Language (ESL) learners from high-context language backgrounds, such as Korean, Japanese, and Chinese, in online academic and professional settings. Cultural norms around indirectness, shared context, and cautious self-expression often influence how learners communicate in English, leading to subtle but persistent barriers to clarity, participation, and confidence. Adopting a research-through-design approach, the study combines interviews, diary studies, and online meeting transcript analysis to surface recurring pain points in learner speech. These insights informed the development of *Stepping Stones*, an AI-powered learning application that provides context-aware feedback based on learners’ actual spoken interactions. While many conventional language learning tools often emphasize grammar instruction

and scripted practice, this system draws on learners' real conversational data to provide personalized feedback and lesson flows. By analyzing naturally occurring speech rather than pre-written responses, it addresses individual communication patterns shaped by cultural context and supports growth in expressive clarity, pragmatic fluency, and habit awareness. This thesis presents a design-driven exploration of how culturally rooted communication barriers might be addressed through personalized feedback grounded in real learner speech, aiming to raise new possibilities for how interaction design can respond to the needs of high-context ESL learners.

1. Introduction

Learners from high-context language backgrounds, such as Korean, Japanese, and Chinese, often face challenges adapting to English-speaking academic environments, where verbal clarity, linear reasoning, and direct articulation are more explicitly valued. These expectations, particularly prevalent in American English communication norms, contrast with high-context communication styles that tend to be more indirect, ambiguous, and reliant on the listener's ability to interpret unspoken cues and shared assumptions (Gudykunst, Ting-Toomey, & Chua, 1988).

This distinction between high-context communication style to low-context communication style has been widely examined in intercultural communication studies. For example, Joynt and Warner (1996) found that in high-context cultures, communication depends heavily on nonverbal cues, shared experiences, and implicit understanding. In contrast, low-context communication, such as that common in the U.S., relies on direct, explicit language where meaning is carried primarily through words (Gudykunst & Nishida, 2001). These cultural-linguistic differences often manifest in spoken English as hesitation, indirect phrasing, or translation-like logic. Speakers from high-context cultures tend to rely on shared assumptions rather than direct verbal expression, which can lead to misunderstandings and miscommunication in low-context settings like American classrooms or professional meetings (Scollon & Scollon, 2001).

As of the 2023–2024 academic year, the United States welcomed approximately 1.13 million international students, with over 25% coming from high-context language regions such as South Korea, Japan, Taiwan, and China (Institute of International Education, 2024). Despite strong academic qualifications, many of these students report difficulty participating confidently in discussions

or team-based projects, citing unclear feedback, performance anxiety, and unfamiliar communication norms as major barriers.

To address these challenges, I designed an AI-powered language learning application, *Stepping Stones*. Unlike traditional grammar-focused English education applications, *Stepping Stones* draws on real Zoom meeting transcripts: turning learners' actual speech into customized feedback and coaching. This can enable language learning closer to real-life interaction, rather than relying on artificial prompts or pre-written dialogue. Rather than correcting errors in isolation, *Stepping Stones* helps learners strengthen clarity, logical flow, and expressive confidence over time. While *Stepping Stones* does not prescribe a single "correct" way of speaking, it helps learners notice and revise unclear patterns shaped by cultural habits, supporting clearer, more intentional expression over time.

The objective of this thesis is to present the design of *Stepping Stones*, a transcript-based, AI-assisted learning tool that offers ESL learners personalized feedback on real meeting conversations. While its core function is language support, this product is intentionally designed with learners in design or communication-intensive disciplines in mind, where speaking clearly in collaborative settings is essential.

This paper explores the design rationale, interaction model, and pedagogical framework that guide *Stepping Stones*. The "Solution Strategy" section outlines what exactly has been designed: a set of UI applications, feedback flows, and learning cycles aimed at helping learners revise unclear or indirect speech patterns. Rather than offering theoretical models, this work demonstrates how key design decisions were translated into an actual, functioning product informed by real learner behaviors.

2. Background / Literature Review

A significant proportion of international students in the U.S. (over 30% as of 2023 and 2024) come from high-context cultures such as Korea, Japan, Taiwan, and China (Institute of International Education, 2024). For these learners, challenges with English extend beyond vocabulary or grammar gaps. They stem from deeper differences in communication norms. High-context languages rely heavily on shared cultural knowledge and implicit cues, making communication context-dependent rather than explicitly verbal. In contrast, English, particularly in academic and professional settings, favors clarity, directness, and explicit articulation (Hall, 1976).

Language anxiety is another well-documented barrier. Learners experiencing anxiety often avoid speaking tasks, over-script their responses, or rely on literal translation strategies (Horwitz et al., 1986; MacIntyre & Gardner, 1994). This thesis reaffirms the continuing relevance of this issue through interviews and diary studies, showing that performance pressure and lack of meaningful feedback persist as major obstacles.

Recent scholarship emphasizes the value of timely, structured, and context-aware feedback in fostering long-term language development (Li, 2020; Hyland & Hyland, 2006). Yet existing EdTech platforms, (e.g., Duolingo, ELSA Speak, and Speak) tend to focus on correctness, pronunciation, or gamified completion, often neglecting cultural nuance and expressive tone. They rarely address how emotional confidence or communication habits affect language use.

Many learners now turn to AI-based tools like ChatGPT or Grammarly for efficient language output. While helpful, these tools are often perceived as task-oriented, offering immediate correction rather than fostering long-term growth. Educational apps such as Duolingo, ELSA Speak, and Speak integrate AI for grammar and speaking practice. While ELSA and Speak accommodate varied proficiency levels better than Duolingo, they rely on pre-scripted exercises rather than user-generated speech, and do not track cultural-linguistic patterns over time. Their feedback remains generic, lacking the contextual depth or reflection-based support needed for complex academic communication. These limitations reinforce the need for applications that offer personalized, context-aware feedback grounded in learners' actual usage and growth trajectory.

In contrast, *Stepping Stones* supports learners by addressing not only grammatical accuracy, but also expression style shaped by cultural background and performance anxiety. It aims to promote both clarity and mutual understanding in multilingual academic and professional settings.

3. Solution Strategy / Design Approach

3.1 Design Philosophy & Strategic Framework

This thesis adopts a Research-through-Design (RtD) strategy, wherein iterative prototyping serves as both a method of problem exploration and a tool for testing potential solutions. Rather than approaching the project with a fixed and predetermined definition of the problem, the design process uncovered communication barriers through cycles of refinement. These barriers (e.g., vague phrasing, indirect speech, or topic delay) are often rooted in the cultural and linguistic patterns of high-context language speakers.

The RtD approach enabled the identification of nuanced, context-specific challenges that are typically overlooked by conventional design methodologies. In response to these findings, the project integrates three interrelated philosophical frameworks:

First, a learner-centered design perspective ensures that the application prioritizes the real-world needs of English learners, particularly in meeting and classroom communication scenarios.

Second, a culturally responsive design lens guides the application away from rigid models of correction and instead encourages alignment with the pragmatic and rhetorical norms of high-context learners.

Third, the design incorporates a low-pressure learning environment that emphasizes reflection and gradual improvement. While emotional resonance was not a primary design goal, it is intentionally embedded to support learner confidence and long-term engagement.

3.2 Iterative Design Process and Co-Design

The design process unfolded across five iterative cycles, each shaped by direct feedback from ESL students with backgrounds in design. Although formal usability testing was not conducted, these participants engaged actively throughout the development process as co-designers. Their contributions influenced interface structure, interaction flow, feedback language, and the overall tone of the application.

Three methodological approaches supported this participatory design process. Autobiographical design played a foundational role by drawing on the researcher's personal experience as a high-context ESL learner. This perspective helped surface subtle cognitive and emotional needs that often go unaddressed in grammar-focused tools. In parallel, user-centered design informed interface decisions, ensuring that the lesson structures were predictable, cognitively manageable, and emotionally safe for learners. Participatory design practices also enabled iterative feedback on mid-fidelity prototypes. Participants shaped a range of application elements, from tone and lesson length to visual layouts and the gamified reward mechanic, such as the pebble stacking feature. Together, these methods contributed to a design process that was responsive, contextually grounded, and co-developed with the learner community.

3.3 Technology & Pedagogical Choices

AI was selected not for novelty, but because of its ability to deliver analysis grounded in real, user-generated data, drawn from learners' actual meetings and conversations. This ensures that feedback is not hypothetical, but based on each user's real-life communication context, making the application both practical and personally relevant.

To address the multifaceted challenges faced by high-context ESL learners, the category system organizes feedback into four pedagogical categories: Flow & Thought Process, Accuracy, Nuance & Lexical Precision, and Habit Awareness. This framework was not pre-defined, but rather emerged through iterative engagement with learner transcripts, reflective journaling, and co-design input. Each category was deliberately constructed to capture a distinct yet interrelated dimension of verbal expression, moving beyond conventional grammar correction models.

AI enables consistent tracking of these four dimensions across real learner data, transforming abstract challenges (e.g., indirect logic or patterned filler use) into observable and coachable behaviors.

Flow & Thought Process helps learners improve logical sequencing, linear articulation, and coherence. This reduces frustration and misunderstandings that stem from culturally different information structures.

e.g., A learner may delay the main point until the end of a sentence, reflecting topic-last logic from their first language.

Accuracy targets grammar and sentence structure to improve clarity. It builds linguistic correctness and reduces fear of being judged.

e.g., Misusing verb tense or article use may obscure otherwise clear ideas.

Nuance & Lexical Precision supports the use of professional, natural expressions with appropriate tone and register. This enables expressive precision and builds confidence.

e.g., Saying “I will explain some bad effects” instead of “I’ll outline the key concerns” may sound overly casual or vague.

Habit Awareness surfaces repeated filler words, translation patterns, and phrasing habits that reduce clarity over time. This develops self-awareness and normalizes mistakes as part of a learnable process.

e.g., Overusing “actually” or translating idioms literally like “catch mind” instead of “understand.”

These categories were also chosen based on communication challenges frequently observed among learners from high-context language cultures. In such learners, sentence-level accuracy is often not the primary problem. Instead, their difficulties stem from indirect expression, topic delay, and hesitation rooted in cultural logic. Studies in intercultural communication (e.g., Hall, 1976; Scollon & Scollon, 2001; Maynard, 1997) confirm that high-context speakers tend to structure ideas differently than expected in low-context English conversations. For example, they may withhold main ideas until the end of a sentence, rely heavily on hedging, or translate politeness norms literally.

Thus, the Flow & Thought Process category was added to support logical sequencing and prioritization. Nuance & Lexical Precision addresses culturally encoded tone and formality mismatches. Habit Awareness surfaces hidden translation patterns that would otherwise become fossilized. Together, these categories reflect the types of breakdowns that occur when culturally influenced communication patterns meet low-context expectations.

Unlike many language learning applications that focus solely on grammar correction or pronunciation scoring, this application was designed to address the actual sources of communication breakdown in real-life conversations. These four categories were also chosen based on recurring challenges observed among learners from high-context language cultures, where indirect expression, topic delay, and formality mismatches—rather than grammatical errors—often lead to misunderstanding. Studies in intercultural discourse (Scollon & Scollon, 2001; Maynard, 1997) support that such learners tend to organize ideas in ways that conflict with low-context English norms. To address this, Flow & Thought Process was added to support linear articulation, while Nuance & Lexical Precision helps recalibrate tone and formality expectations. Habit Awareness targets subtle fossilized patterns that often go unnoticed. Together, these categories reflect a spectrum of challenges that go beyond correctness to capture deeper issues of coherence, pragmatics, and learner behavior—aligning closely with established second language acquisition models that classify learner errors as grammatical, pragmatic, discourse-level, or habit-driven (Ellis, 1997; Corder, 1981; Kasper & Blum-Kulka, 1993).

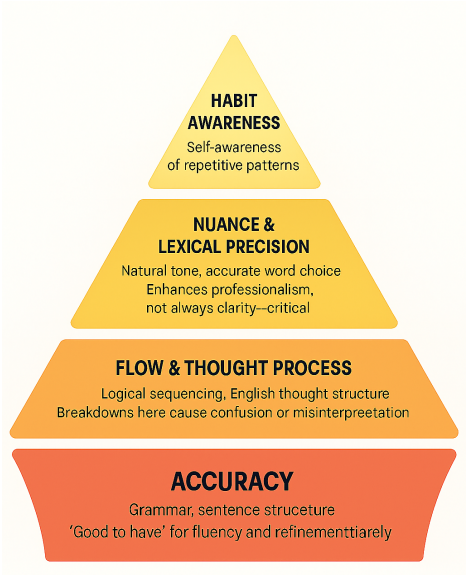


Figure 3.3.1 This table maps established second language error types—ranging from grammatical to pragmatic and habitual issues—to the four feedback categories used in this application. It illustrates how the framework captures a full spectrum of verbal performance challenges.

These categories were developed through iterative synthesis of user interviews, literature research, performance observation, and autobiographical reflection. They were not borrowed wholesale from linguistics but instead designed to be teachable, observable, and relevant to learners' actual behavior.

These categories defined the product architecture:

- Micro Feedback Cards: Cards deliver bite-sized lessons, grouped by issue type with accessible labels
- Coaching Priority Logic: Lessons are sequenced based on communication impact, not grammar rules
- Lesson Plan Structure: Lessons follow a Reflect → Refine → Reinforce progression within each category

This structure supports a reflection-reinforcement learning loop, helping learners gradually notice and resolve recurring communication issues at their own pace, without needing to process all feedback at once.

4. From Insight to Interface

4.1 Translating Research into Design

The application's core features are grounded in qualitative research findings and co-design feedback, particularly those related to common communication patterns observed among high-context learners. These learners frequently delay the main idea in conversation, rely heavily on literal translation, and often use vague or hedging expressions. Rather than treating these tendencies as errors, the design reframes them as culturally influenced patterns that can be addressed through structured reflection and gradual adaptation.

This reorientation informed both the content and form of the lesson modules. Each lesson is built around real ESL conversation data and organized into bite-sized learning tasks that guide learners through a reflective revision process. The learning sequence typically includes a brief diagnosis of the language pattern, a focused micro-task for correction, and a contextual reapplication of the improved sentence in a simulated real-world scenario.

These instructional structures are further reinforced by interface choices that prioritize low cognitive load, personalized feedback, and a tone of support. For example, card-based feedback interactions mimic the natural revision loop that occurs in peer-based learning, while interface language avoids evaluative phrasing that may increase learner anxiety. Together, these design decisions translate abstract user insights into tangible interaction mechanisms that directly target common ESL speaking challenges in professional and academic settings.

4.2 Design Objectives

The following objectives emerged directly from qualitative research insights and guided the application's interaction strategies, content scaffolding, and feedback logic. These objectives reflect the educational philosophy and intended user experience behind the application. They aim to reframe communication improvement not as rule acquisition, but as an emotionally and culturally sustainable learning process.

Objective 1: Language Growth through Pattern Awareness.

Encourage learners to recognize and revise recurring speech patterns, particularly those influenced by direct translation or high-context communication logic. These patterns include vague modifiers, topic delay, and overuse of hedging expressions. The application is designed to help learners identify these habits and internalize more effective, context-appropriate alternatives.

Objective 2: Cultural Sensitivity in Feedback. Incorporate feedback that highlights culturally influenced phrasing and offers clearer alternatives while preserving the speaker's intent. Rather than flagging language as “awkward,” the application explains how expressions may reflect indirectness common in the learner's native language and provides more culturally appropriate English phrasing.

Objective 3: Emotional Safety and Reduced Performance

Anxiety. Reduce learner hesitation and pressure by minimizing evaluative tone, using supportive feedback language, and allowing flexibility in lesson timing and length. Emotional clarity is encouraged by helping users reflect on their patterns in a nonjudgmental environment, reducing fear of mistakes and promoting long-term confidence.

Objective 4: Long-Term Growth with Immediate Usefulness.

Blend short-term corrective feedback with mechanisms for longitudinal tracking and reapplication. Lessons are drawn from users' actual meetings, making the content personally relevant and directly applicable. This enables users to revisit language in authentic contexts and see improvement over time.

5. Design Intervention

5.1 Implementing Language Pattern Awareness

Lesson modules are structured to surface recurring issues in learner speech—such as vague modifiers, indirect phrasing, or backloaded topic delivery. Feedback cards highlight these patterns, explain their communicative impact, and guide users to revise them in more contextually effective ways.

Beyond merely correcting sentences, each lesson provides a rationale that helps learners understand *why* certain expressions might cause confusion, misalignment, or delays in collaborative settings. For example, when learners repeatedly use vague hedges like “kind of” or “maybe,” the application flags this habit, explains how it may weaken the clarity of their intent, and encourages a firmer phrasing alternative.

The habit-tracking dashboard also visualizes repeated phrasing tendencies across sessions, helping users monitor progress and develop greater expressive control over time.

5.2 Delivering Culturally Sensitive Feedback

Each feedback item avoids generic judgments such as “awkward” or “unnatural.” Instead, the application contextualizes unclear expressions by explaining how certain phrasings may reflect translation logic or culturally implicit speech norms.

For instance, a sentence like *“I was just thinking it might be good, if possible...”* might be flagged with the following explanation: *“This phrasing reflects polite deferral common in Korean speech, but in English, it may feel unclear or overly hesitant. Consider saying: ‘I recommend we try...’ to express the same idea more directly.”*

This contrastive approach helps learners internalize culturally specific rhetorical norms, enabling them to choose language more intentionally rather than defaulting to literal translation patterns.

5.3 Reducing Performance Anxiety Experience

To reduce psychological pressure and foster a sense of confidence during use, Stepping Stones deliberately avoids evaluative or corrective tones. Instead, it adopts a supportive voice throughout its interface and feedback system. Rather than presenting mistakes as failures, the system reframes them as recurring patterns that can be refined over time. Furthermore, the application maintains a consistent instructional structure—Reflect → Refine → Reinforce—which provides users with a sense of orientation and control. By minimizing performance anxiety and offering a low-stakes, judgment-free environment, Stepping Stones aims to support sustained engagement that could hypothetically lead to long-term learning growth.

5.4 Supporting Long-Term, Transferable Growth

All lesson content is derived from user-submitted transcripts of real online meetings, ensuring that revision practice remains deeply relevant to the learner's academic or professional domain. This allows users to revisit the same contextual material across multiple sessions and iterate improvements over time.

The application also identifies shifts in tone, expression, or structural habits over repeated encounters with similar contexts—thereby promoting durable learning beyond single-task correction. Micro-feedback is stored in a personal dashboard that tracks improvement trends, surface expression tendencies, and celebrates consistent progress.

Together, these interventions transform abstract goals into concrete, culturally aware, and emotionally sustainable features, supporting not just English fluency, but the learner's sense of voice, confidence, and adaptability.

6. Findings / Insights (from Primary Research)

6.1 Key Insights from Qualitative Interviews

To uncover nuanced language struggles and digital tool use among ESL learners, this study conducted one-on-one interviews and diary studies with seven East Asian female university students in the U.S. Participants came from Korean, Japanese, Taiwanese, and Chinese linguistic backgrounds and ranged in their time living in the U.S. from under six months to over seven years. This context grounds the following key insights that emerged from the qualitative data.

1. **Performance anxiety and risk-avoidant speech behavior**

Most participants reported experiencing high levels of anxiety when speaking in English, particularly in academic discussions, classroom presentations, and group projects. Some also noted similar discomfort in casual or social situations, but academic contexts were described as the most stress-inducing. To avoid making mistakes, learners tended to refrain from participating unless their opinions strongly aligned with what had already been said. This led to reduced engagement and minimal verbal contribution in discussions. One participant noted, “Unless my thought is very different from others’, I often choose to stay quiet.”

2. **Automatic avoidance responses triggered by communication anxiety**

When unsure of what to say or feeling overwhelmed, many learners resorted to vague expressions or off-topic, “safe” comments. These were not deliberate coping strategies, but spontaneous responses to anxiety or mental overload in

high-pressure situations. Participants described moments of mental freeze—not due to lack of grammar knowledge, but due to the emotional strain of real-time conversation.

This pattern highlights the importance of creating emotionally safe environments for language learners. Instead of reinforcing perfection, design strategies should support learners in staying calm, confident, and engaged under pressure.

3. **Cultural-linguistic friction from high-context language backgrounds**

“I think in Japanese first and then try to make it sound English.” One participant, who had lived in the U.S. for over seven years, described her persistent struggle to express ideas formed in her native language. She found that thoughts formed in her first language did not translate clearly into English, often resulting in vague or disorganized expressions. For more complex writing tasks, she consistently relied on ChatGPT—not for grammar correction, but to reframe her ideas into “English-sounding” structures. ***This reflects a deeper challenge not of vocabulary, but of how ideas are cognitively framed, shaped, and delivered.***

This experience is emblematic of a broader pattern often observed among learners from high-context communication cultures. Learners from Korean and Japanese backgrounds frequently face difficulty aligning their implicit, layered communication style with the more explicit and direct expectations of English. Addressing this may require helping learners reframe their ideas earlier in the planning or pre-speaking stage, using tools that guide structural clarity while preserving their intended nuance. This insight later informed design explorations that considered when and how learners might be scaffolded—whether through pre-conversation outlining, contrastive examples, or soft prompts guiding structure.

4. **Heavy reliance on AI tools, but limited learning impact**

As part of the research, participants were also asked about the AI tools they currently use to manage language-related challenges

in their academic life. While the focus of this study is on AI-powered educational tools, it was essential to first understand the broader ecosystem of AI usage among learners. It was necessary to examine how learners currently interact with general-purpose AI tools in academic settings. Understanding their usage patterns, intentions, and unmet needs helped reveal design opportunities for more intentional, learning-oriented AI interactions. Interestingly, ChatGPT, DeepL, and occasionally Grammarly emerged as the dominant tools used for writing support.

All seven participants actively used AI tools like ChatGPT, DeepL, and Grammarly to revise their writing—both their own and their peers’—and mentioned that such use was common among their ESL networks. However, learners overwhelmingly expressed that while these tools helped them complete academic tasks, they did not support long-term language development. Several key limitations emerged:

a. **Lack of retention and reflection:** Learners reported that even after refining English with AI, they could not recall or reuse the improved expressions. One participant shared, “I don’t feel like I’m improving. I just fix and forget.” Another stated, “I think it decreases my skills because I rely on it too much. Now I feel that I am losing those skills such as writing effectively.”

b. **Task-focused use, limited reflective learning:** While learners used AI tools frequently—often on a daily basis for academic writing—they predominantly used them to complete tasks rather than as tools for language learning. ChatGPT, DeepL, and occasionally QuillBot or Grammarly were the most commonly mentioned. Six out of seven participants described using these platforms to finish assignments quickly, often without reflecting on the revised output. One participant shared that she primarily uses AI to check grammar and polish structure, but does not think about why certain changes are suggested. This suggests that the limited learning impact stems not from the AI tools themselves, but from users’ utilitarian approach—prioritizing immediate task completion over language development.

c. Incidental learning through repeated AI interaction:

Interestingly, one participant described a different experience: she began to notice learning patterns when repeatedly regenerating responses on ChatGPT until it gave her a sentence that felt right. In that process, she realized that certain expressions she intended could only be captured in specific English forms, which taught her nuances in structure and tone. She explained that AI-generated sentences often reflected more natural U.S.-style phrasing and helped her reverse-learn patterns she had misunderstood or over-applied from textbook grammar. However, even in this case, her learning came not from intentional study, but as a byproduct of persistent interaction and a personal tendency to question and refine outputs. This insight reinforced the idea that effective learning via AI does not happen by default and informed the decision to design features that promote reflective iteration, like micro-revision prompts and contrastive examples that help learners internalize structure over time.

d. Mismatch with learner voice and tone: Several participants expressed discomfort with the overly formal or generic tone of AI-generated text. One learner said, “What AI gives is too professional—I feel shy to use them. It doesn’t feel like my voice.” Another shared that she had to regenerate many times to find a sentence that included phrasing she liked.

e. Lack of personalized insights: Despite interacting with these tools frequently, learners were unable to get a holistic understanding of their recurring issues or progress over time. Even when prompted, tools like ChatGPT were unable to synthesize past inputs or identify patterns across sessions.

These limitations exposed both behavioral and emotional gaps in current AI-assisted learning practices, which the following design strategies aimed to address.

5. Mixed feelings toward AI as a learning companion

Participants expressed ambivalence about the ethical and emotional implications of using AI. One learner shared, “It feels

like cheating—it's not mine. I keep checking it again and again. It's not really my writing, so I feel uneasy submitting it." Another remarked, "Using AI embarrasses me. It feels like it's not what students should do." In contrast, one participant explained, "All the ideas are mine, so I don't feel guilty. I think it actually helps grow my critical thinking in the early stages of design."

These perspectives reflect varied assumptions about authorship, ownership, and responsibility in the context of AI-augmented learning. For some, the use of AI raised ethical discomfort or a sense of diminished authorship. For others, it was seen as an opportunity to develop ideas and structure them better. This divergence highlighted the need for a product that supports student agency, frames AI as a collaborative assistant rather than a generator, and encourages growth through guided input rather than one-click output. This insight ultimately shaped the application's pedagogical positioning: to promote AI as a scaffold for self-driven learning, not a shortcut to completion.

6.2 Competitor Analysis

Despite the growing availability of AI-powered English learning tools, many fail to address the nuanced challenges experienced by high-context language learners. These platforms (e.g., such as Duolingo, ELSA Speak, Grammarly, and Speak) typically prioritize surface-level correctness, focusing on grammar, vocabulary, and pronunciation. While such tools may occasionally flag sentences as "unnatural" or "unclear," they rarely engage with the deeper cultural and linguistic reasoning behind these expressions. This limitation is especially significant for learners from East Asian language backgrounds, whose discourse patterns often reflect different expectations around clarity, directness, and interpersonal tone.

One critical issue lies in the lack of cultural-linguistic sensitivity in the feedback these applications provide. Learners from high-context backgrounds may construct sentences that are grammatically correct but subtly misaligned with English communication norms due to indirect phrasing or

translation-based logic. For example, a sentence like “I’m not sure, but maybe it could be considered...” might be flagged as vague or ineffective, despite the fact that such hedging is a culturally normative speech strategy in Korean or Japanese. These systems rarely acknowledge this nuance, leading to a form of correction that subtly invalidates the learner’s communication style. As a result, learners may internalize the belief that their natural patterns are simply “wrong,” rather than contextually different. Moreover, these systems miss the opportunity to guide learners toward culturally adaptive strategies that build on their existing strengths, rather than replacing them.

A second limitation concerns the lack of transparency in how corrections are delivered. Tools such as Grammarly, DeepL, and ELSA Speak frequently revise or rewrite learner input without providing explicit reasoning behind those changes. Learners may receive alternative phrasing that is labeled as “more natural,” but are often left uncertain about which part of their original expression was problematic, or why the revision improves clarity or tone. In speaking-focused platforms like Speak, feedback is often reduced to short, generic responses—such as “Try again” or “Great job”—in order to preserve conversational flow. While this keeps the user experience fluid, it offers little opportunity for reflection or deeper learning. Several interviewees expressed concern that although they appreciated the speed of corrections, they felt uncertain whether they were truly improving or simply becoming dependent on suggestions. This concern is echoed in user discussions across online forums, where learners frequently question whether such tools lead to real speaking improvement.

Third, many platforms lack continuity across sessions, making it difficult for learners to build on past progress. Speaking tools like Speak often operate on an open-topic basis but do not retain contextual memory between conversations. Each session begins as a blank slate, without knowledge of the learner’s prior inputs, goals, or topic progression. This fragmented experience prevents learners from iteratively refining their expressions across similar scenarios. In contrast, effective language learning often relies on returning to familiar themes and gradually improving how ideas are structured and expressed. The absence of continuity in these

tools limits learners' ability to develop fluency through cumulative practice.

Finally, most tools offer minimal support for long-term habit awareness. Platforms such as Duolingo, ELSA Speak, Grammarly, and DeepL typically evaluate inputs on a sentence-by-sentence basis, without tracking recurring patterns in phrasing, tone, or expression. As a result, learners are rarely made aware of habitual behaviors (e.g., filler word overuse, repetitive grammar structures, or indirect framing) that may hinder their clarity over time. In contrast to human tutors, who often recognize and address such patterns during ongoing instruction, AI tools tend to treat each interaction as isolated. This lack of longitudinal insight limits opportunities for learners to reflect on and improve their speech habits in a structured, progressive way.

7. Design Outcome / Output

This section outlines how Stepping Stones implemented the design objectives introduced in Section 4.2. The design outcomes include both visual and interaction-level solutions aligned with the educational and emotional goals of the project. The focus is not on demonstrating a polished, finished product, but on showing how iterative design decisions progressively embodied Stepping Stones's intended function, tone, and pedagogical structure.

7.1 Design Iteration

Design decisions evolved over four core versions, each integrating key feedback from users and pushing the concept closer to the intended goals of emotional safety, cultural sensitivity, and pattern-based learning.

Version 1: Introducing Modular Feedback Categories.

The first version introduced a dashboard structure divided into four feedback categories: Accuracy, Flow, Nuance, and Habit. The interface was designed to highlight progress and reflection trends within each category. However, internal feedback raised concerns about whether these visualizations were actually meaningful or actionable. Static metrics and generalized phrasing were perceived as emotionally flat, limiting user engagement.

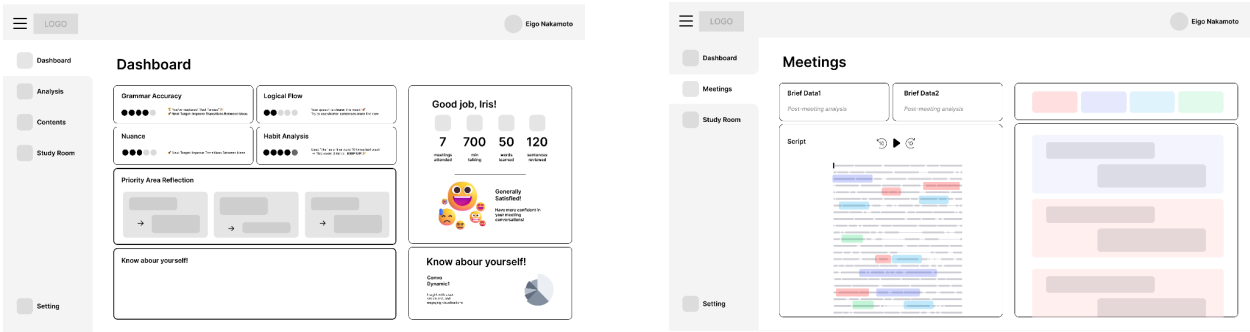


Fig. 7.1.1

Version 2: Emphasis on Data Visualization.

To address these issues, the second version introduced trend cards and category-specific feedback overviews to improve clarity and visual appeal. A paper prototype was tested for speed and comprehension. While these additions made Stepping Stones more dynamic, participants reported feeling overwhelmed by the volume of content. They also found the color-coded categories either confusing or emotionally charged, highlighting the need to reduce cognitive load and improve tone clarity.

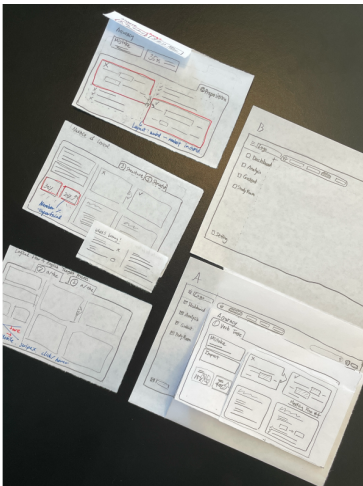
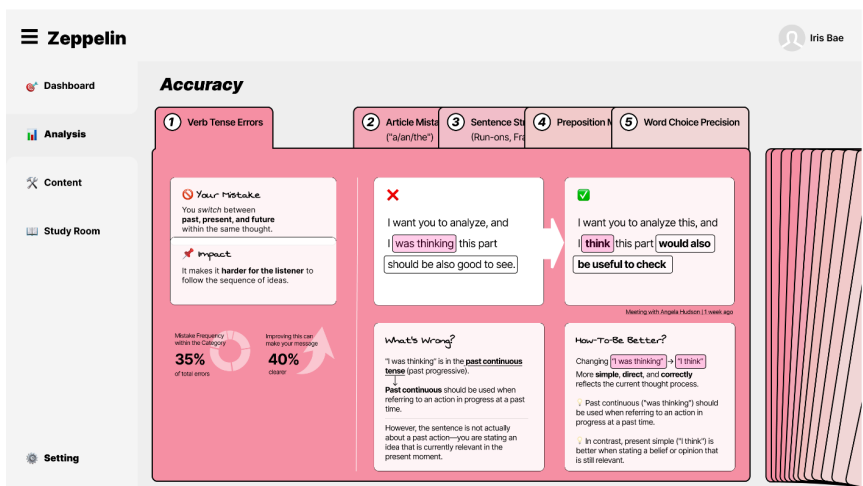


Fig 7.1.2

Version 3: Redesigning for Emotional Clarity and Confidence

The third version prioritized emotional clarity and user confidence by integrating features like Actual vs. Perceived Fluency tracking and personalized Communication Reports. Feedback tone became more supportive, and insights from diary studies were used to reflect learners' own progress. Despite these improvements, some users remained unsure about the purpose or value of dashboard metrics, which still appeared infrequently updated and too abstract.

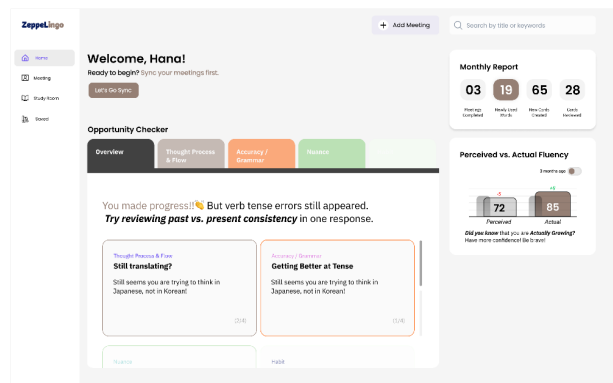
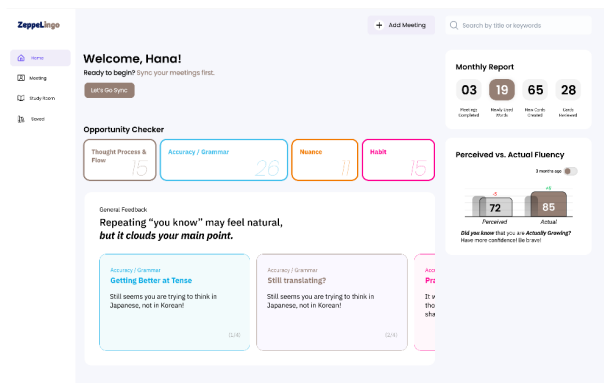
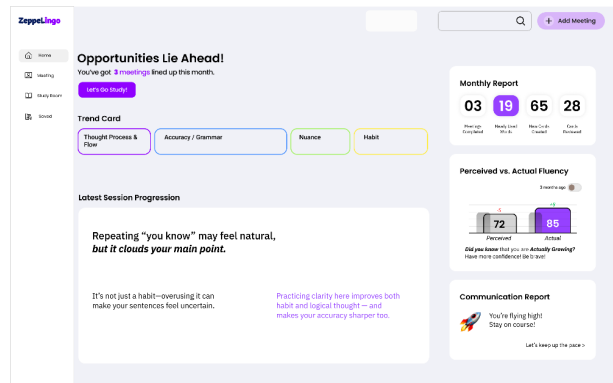
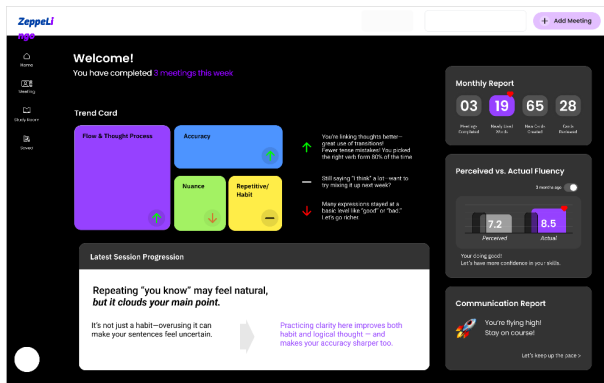


Fig 7.1.3

Version 4: Context-Based Learning and AI Integration. The fourth and final iteration marked a shift toward personalized, real-life application by integrating an AI-based lesson generator linked to user-submitted transcripts. Learners could now choose the scope and duration of lessons, easing performance pressure. Feedback and reflection were tied directly to real conversation excerpts, and chat-like micro-feedback made the experience feel more interactive and contextual.

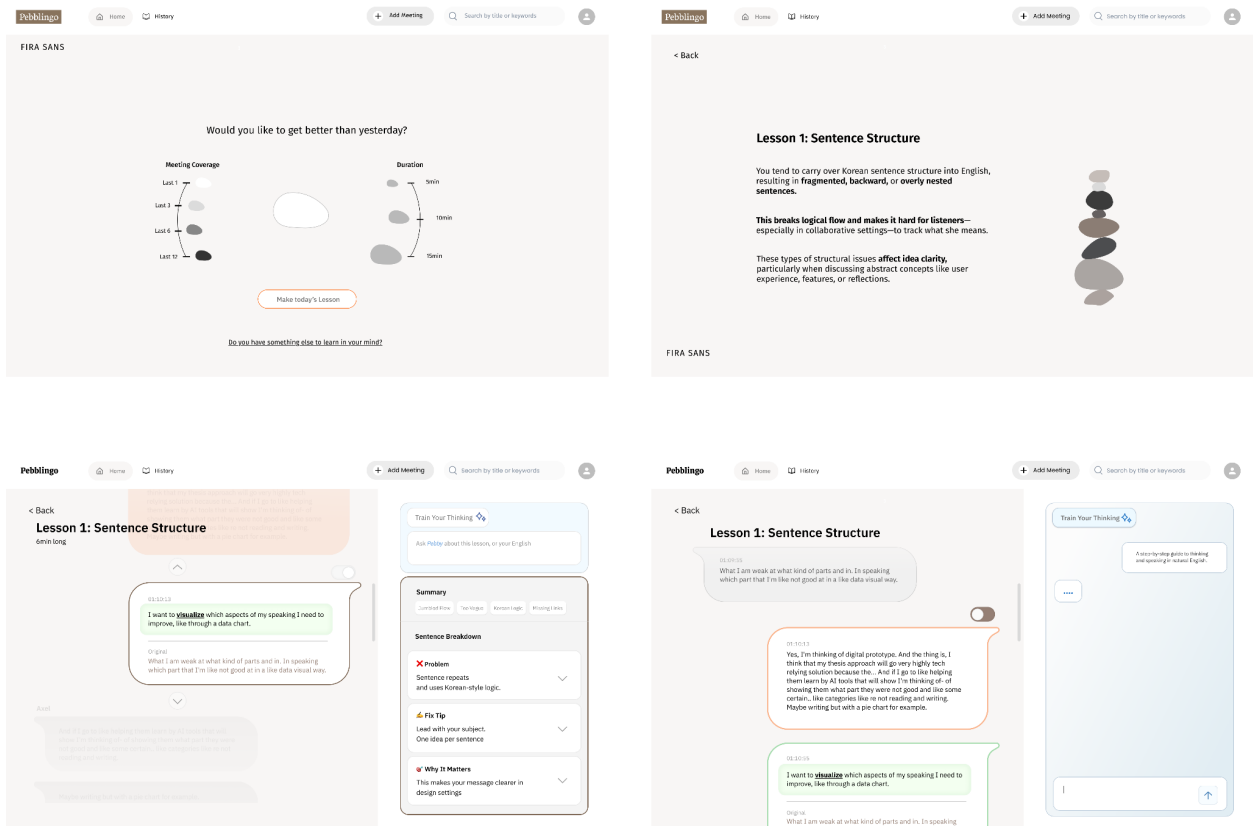


Fig. 7.1.4

Across all four iterations, key product qualities were refined and clarified. Emotional safety was supported through a friendly tone, flexible pacing, and consistent structure. Cultural sensitivity became embedded in Stepping Stones through contrastive, AI-informed feedback strategies. Pattern awareness was reinforced through categorized habit tracking and lesson breakdowns—laying the groundwork for a more intuitive and culturally responsive learning experience.

7.2 Final Output

The final design of *Stepping Stones* comprises a set of interrelated features intended to support structured language development, with particular sensitivity to the needs of learners from high-context cultural backgrounds. The interface, feedback mechanisms, and learning sequence are designed to help learners identify, reflect on, and incrementally improve unclear or habitual speech patterns, while minimizing cognitive overload. Each element simultaneously facilitates instructional engagement and sustained learner motivation.

Smart Lesson Generator: Learners begin by selecting a recorded meeting, which Stepping Stones analyzes to generate personalized lessons based on actual spoken transcripts. These lessons are not templated; instead, they are constructed dynamically through AI-driven analysis that identifies communication-critical issues, either due to high frequency or substantial impact on meaning. This ensures that each lesson addresses areas of genuine communicative relevance for the learner.

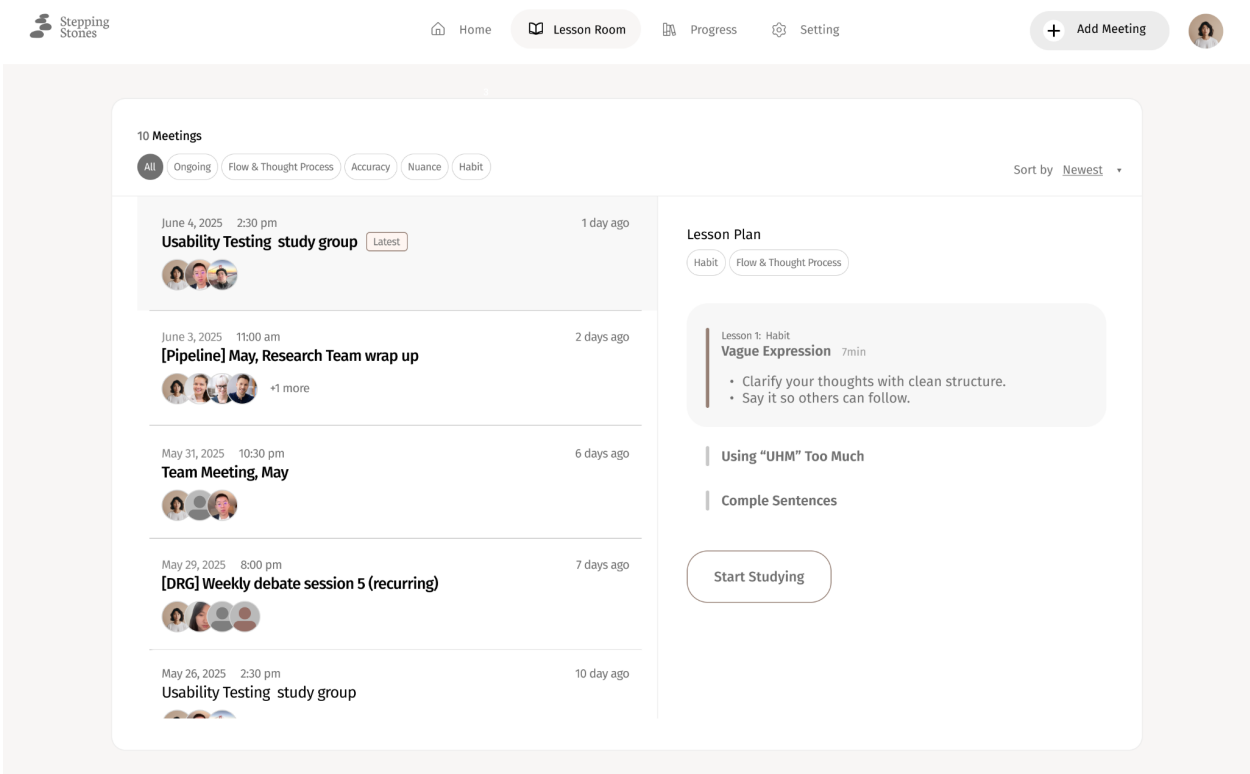


Fig 7.2.1

Bite-Sized Feedback Cards. Rather than overwhelming learners with extended instructional text, Stepping Stones provides feedback through compact, interactive cards that appear immediately after an answer is submitted. These bite-sized feedback cards are designed to deliver focused insights regardless of whether the learner’s response is correct or incorrect. Each card presents a specific type of information—such as an explanation for an incorrect answer, identification of recurring linguistic patterns, clarification of the question’s intent, or a brief conceptual summary related to the lesson objective. By distributing feedback across multiple discrete cards, Stepping Stones encourages manageable reflection, reduces cognitive load, and facilitates sustained engagement with the learning process.

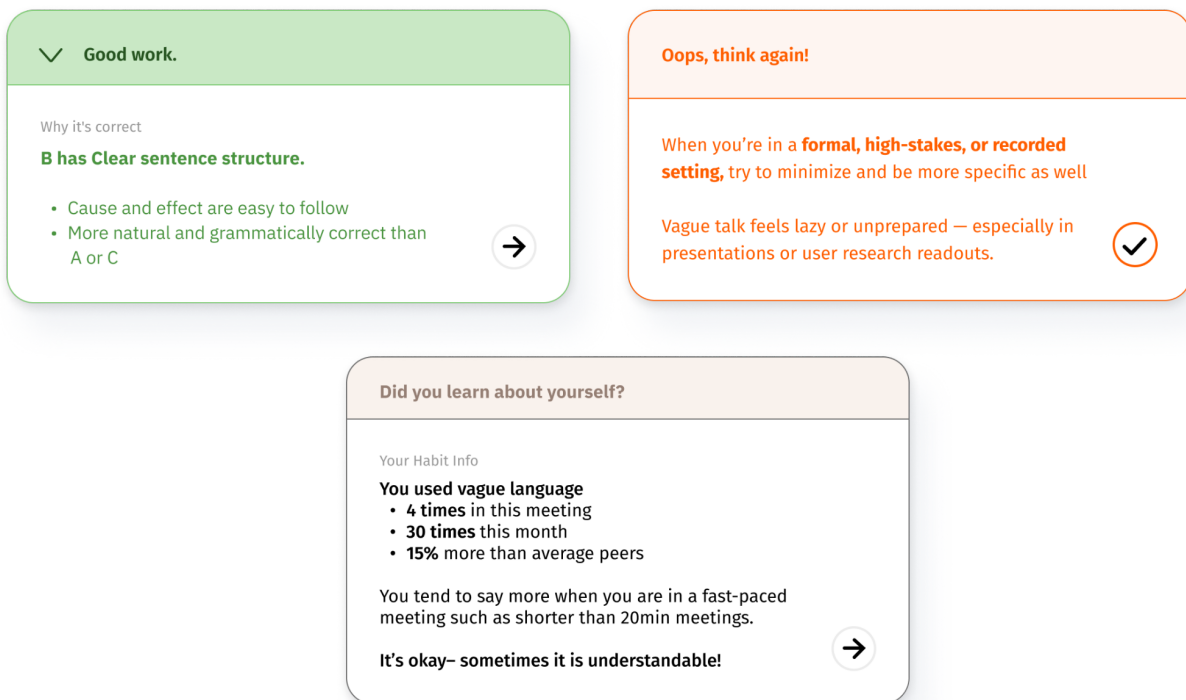


Fig. 7.2.2

Four Feedback Categories: Each lesson in Stepping Stones is categorized into one of four feedback areas—Accuracy, Flow and Thought Process, Nuance and Lexical Precision, and Habit Awareness—each corresponding to a distinct type of communication challenge. These categories not only guide the pedagogical focus of the lesson but also help learners monitor their development across key dimensions, as the assigned category is visibly displayed in both the lesson summary and the progress bar. Building on this classification, each lesson follows a structured instructional cycle composed of three stages: Reflect, Refine, and Reinforce. In the Reflect stage, learners examine their original expression alongside a clearer alternative to identify and understand the underlying issue. During the Refine stage, they engage in targeted exercises designed to practice and internalize more effective linguistic patterns. Finally, the Reinforce stage prompts learners to record themselves applying the revised expression, supporting consolidation through active use. This cyclical model fosters incremental, sustainable improvement and encourages greater communicative confidence over time.

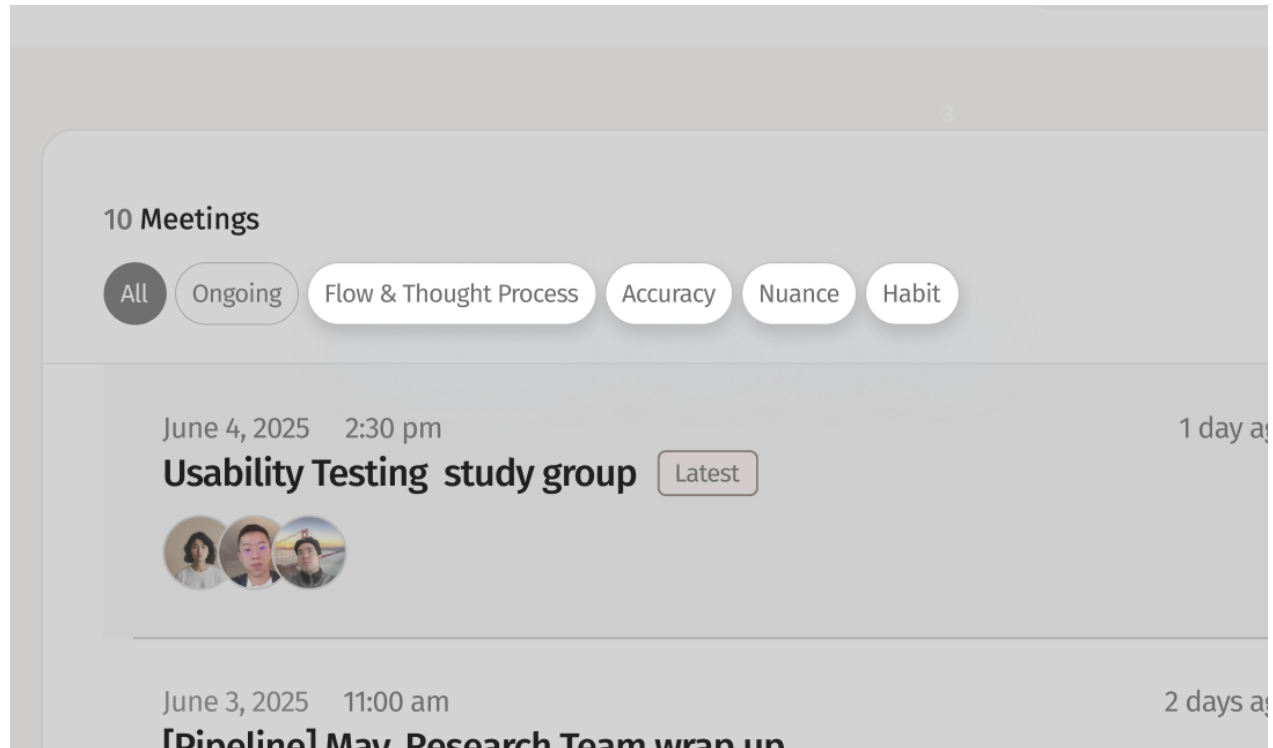


Fig 7.2.3 Category filter on 'Lesson Room' screen

Exercises for High-Context Learners: The Flow and Thought Process category is especially relevant for learners from high-context language backgrounds, whose speech may reflect indirect logic, ambiguous phrasing, or translation-based sentence structures. A representative lesson in this category typically begins by asking the learner to identify the intended meaning behind a vague or unclear sentence they previously used, fostering awareness of how their original phrasing may hinder clarity. In the Refine stage, the lesson unpacks overlapping issues—such as syntactic ambiguity, unnatural word order, or indirect structuring—through guided, step-by-step practice. Finally, in the Reinforce stage, learners are prompted to reconstruct and vocalize the improved sentence with minimal scaffolding, encouraging autonomous restructuring and more direct expression. This progression supports both cognitive and behavioral shifts in language production, helping learners move toward clearer, audience-centered communication.

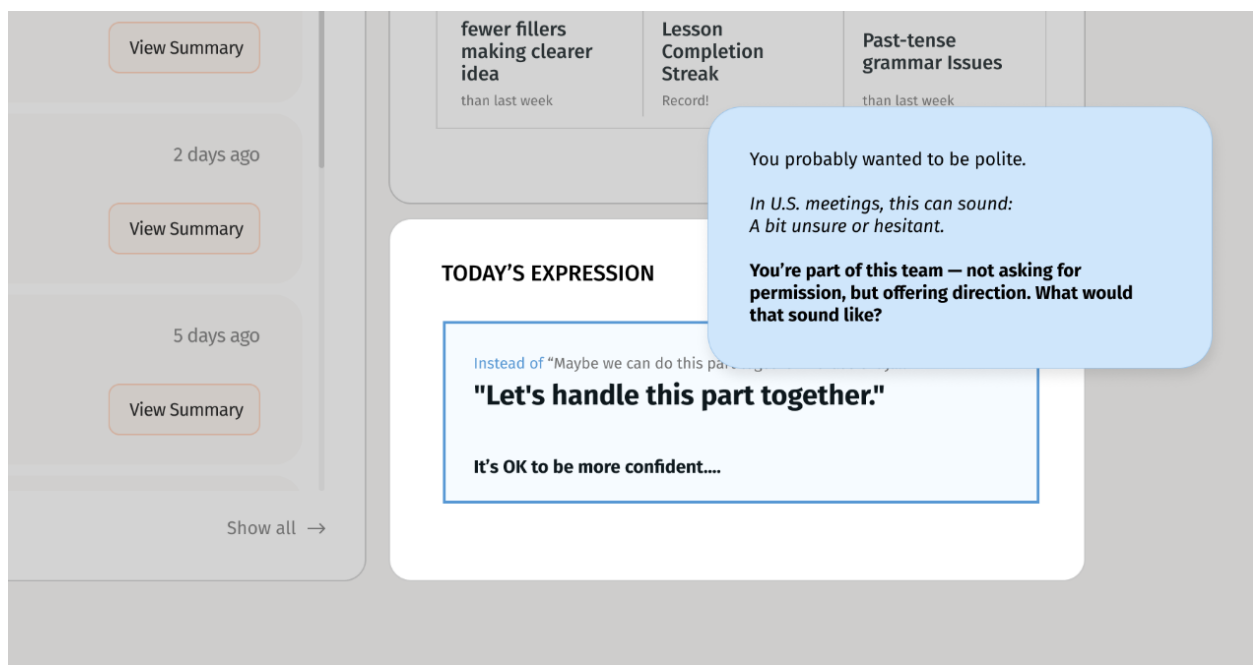


Fig.7.2.4 *Today's Expression* panel high-context learner-focused suggestion on 'Home' screen, with tooltip pop up showed up by mouse hover on the panel

Progress Dashboard and Motivation: Learner progress is represented through an interactive spider chart that tracks performance across the four core feedback categories: Accuracy, Flow and Thought Process, Nuance and Lexical Precision, and Habit Awareness. As learners complete lessons, the corresponding segments of the chart are dynamically filled in, offering a visual summary of growth and highlighting areas that may need further attention. Lessons that have been generated but not yet completed are clearly marked, allowing learners to manage their progress intentionally. To maintain motivation without increasing pressure, the interface incorporates subtle gamification elements—such as animated transitions and soft color changes—that celebrate progress while preserving a calm and focused learning environment.

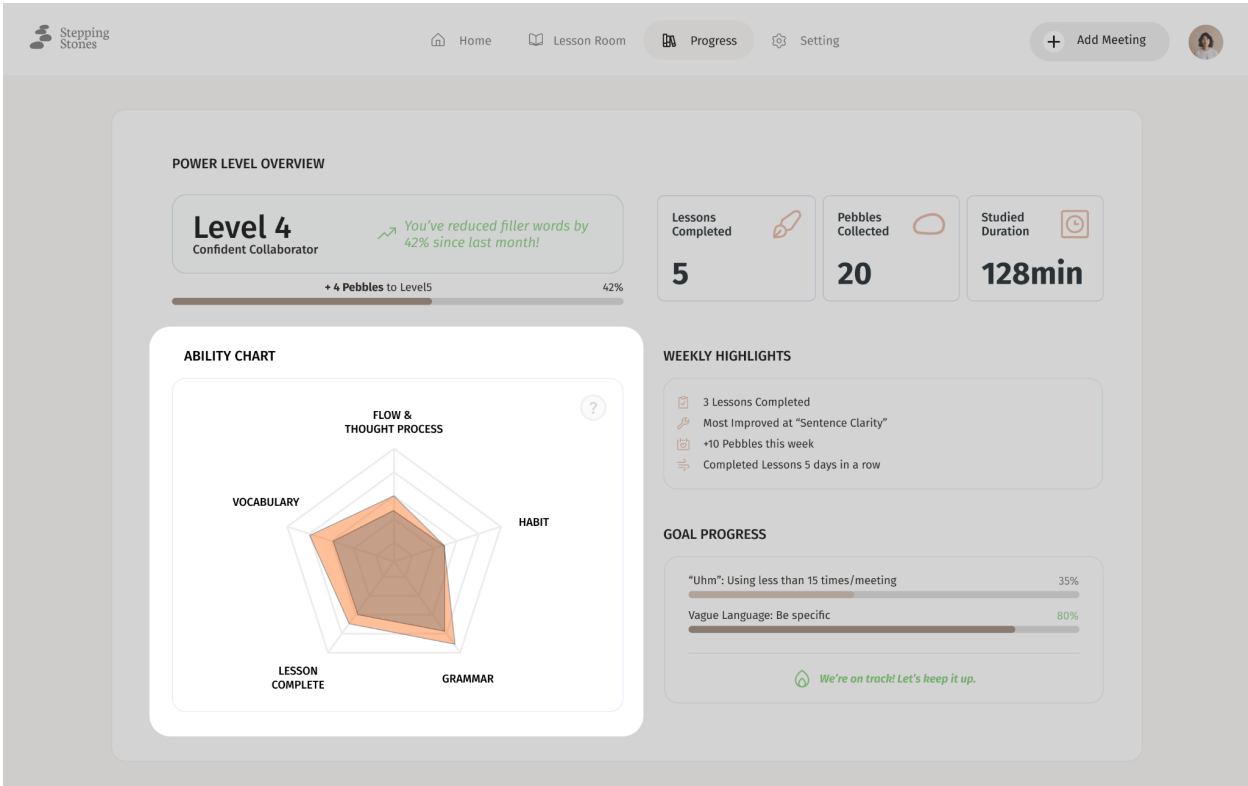


Fig.7.2.5 Spider chart on 'Progress' screen

Visual Tone and Brand Language: The product interface is designed to evoke a warm, grounded, and professional atmosphere. It features a muted, nature-inspired color palette alongside metaphoric visual elements, such as stacked pebbles, to symbolize gradual progress and steady growth. This visual language is complemented by UX writing that is intentionally supportive and reassuring, aiming to ease performance anxiety and foster a sense of calm focus. Together, these design choices create an emotionally considerate environment that encourages learners to stay engaged and confident throughout their learning journey.

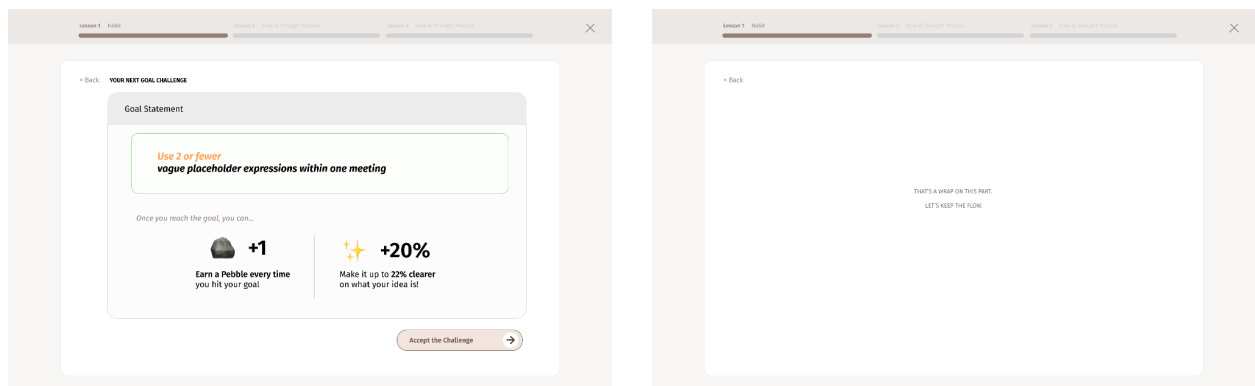


Fig.7.2.6

Together, these visual and interaction elements form a cohesive product experience that reflects Stepping Stones's core values: cultural sensitivity, emotional sustainability, and personalized learning. By combining intentional visual tone, flexible feedback architecture, and real-life speech data, *Stepping Stones* offers more than just error correction. It provides a reflective space where high-context learners can grow at their own pace, feel seen in their communication struggles, and gain the confidence to express themselves more clearly and authentically in academic and professional contexts.

8. Implications / Future Work / Limitation

This thesis proposes a lesson structure[lets expand: The thesis proposes an overall digital product that employs as lesson structure.] designed to support communication clarity for high-context learners. While the current version offers meaningful scaffolding [how? Back up.] through feedback cards and structured lessons, it is not a complete solution. Several limitations remain, and future work is needed to improve both precision and adaptability.

8.1 Limitations of the Current application.

While the current application offers valuable post-meeting feedback, it faces three key limitations: delayed feedback timing, structural inflexibility, and privacy concerns. First, because feedback is generated only after meetings, learners receive no real-time support during high-pressure moments like presentations or interviews. Second, the fixed three-step lesson format (Reflect → Refine → Reinforce), though helpful for habit formation, lacks adaptability for advanced users or those with specific goals. Lastly, reliance on user-submitted Zoom recordings raises privacy and social concerns, particularly in group settings where participants may feel uneasy about being recorded. To address this, future versions may adopt alternatives like transcript-only uploads or anonymized live transcription to reduce friction while maintaining learner autonomy.

To mitigate this concern, future versions may explore alternative methods such as transcript-only uploads, blurred identity extraction, or live transcription tools that do not require full video/audio capture. These approaches could promote ethical transparency while preserving learner comfort and autonomy.

8.2. Realistic Areas for Future Improvement

Future iterations of *Stepping Stones* could benefit from increased adaptivity based on user behavior. For example, lesson sequencing might be dynamically prioritized using data on which lessons are skipped, repeated, or completed—allowing the application to surface the most relevant communication issues for each learner. Feedback content could also be refined by detecting patterns in user behavior, such as when a

learner repeatedly corrects an issue but fails to apply it in future sessions. Additionally, rather than grouping users by generalized proficiency levels, segmenting them by communication contexts—such as team discussions versus mentor feedback sessions—could enable more precise adjustments in lesson tone and scope. While this work does not attempt to present a universal solution, it offers a focused and emotionally considerate feedback model that can evolve through adaptive logic and broader application in diverse real-world settings.

9. Conclusion / Reflection

This thesis addressed a subtle yet critical issue: ESL students from high-context language backgrounds often produce grammatically correct English that still feels vague, overly cautious, or indirect—leading to reduced participation, misunderstandings, and communication anxiety in academic settings. To respond to this, I adopted a research-through-design approach that prioritized real learner behavior over prescriptive language rules. ESL design students participated as co-designers, and their feedback directly shaped each iteration of the product.

The resulting tool, *Stepping Stones*, analyzes Zoom transcripts and generates personalized lessons across four key categories: Accuracy, Flow & Thought Process, Nuance & Lexical Precision, and Habit Awareness. Rather than focusing solely on correctness, the application supports learners in understanding why certain expressions may be unclear and guides them toward clearer, more confident communication. By reframing errors as growth opportunities and grounding lessons in real conversations, this design promotes culturally sensitive, habit-aware language development. Ultimately, it aims to equip learners with lasting tools for participating more fully in collaborative, multilingual environments—inside and beyond the classroom.

10. Bibliography

Corder, S. P. (1981). *Error Analysis and Interlanguage*. Oxford University Press.

Ellis, R. (1997). *Second Language Acquisition*. Oxford University Press.

Gudykunst, W. B., Ting-Toomey, S., & Chua, E. (1988). *Culture and Interpersonal Communication*. SAGE Publications.

Gudykunst, W. B., & Nishida, T. (2001). Bridging Japanese/North American Differences. *Intercultural Communication Series*. SAGE Publications.

Hall, E. T. (1976). *Beyond Culture*. Anchor Books.

Horwitz, E. K., Horwitz, M. B., & Cope, J. A. (1986). Foreign Language Classroom Anxiety. *The Modern Language Journal*, 70(2), 125–132.

Hyland, K., & Hyland, F. (2006). Feedback on Second Language Students' Writing. *Language Teaching*, 39(2), 83–101.

Institute of International Education. (2024). *Open Doors Report on International Educational Exchange*.
<https://opendoorsdata.org/>

Joynt, P., & Warner, M. (1996). *Managing Across Cultures: Issues and Perspectives*. International Thomson Business Press.

Kasper, G., & Blum-Kulka, S. (1993). *Interlanguage Pragmatics*. Oxford University Press.

Li, S. (2020). The Effectiveness of Written Corrective Feedback: A Meta-analysis. *The Modern Language Journal*, 104(1), 79–101.

MacIntyre, P. D., & Gardner, R. C. (1994). The Subtle Effects of Language Anxiety on Cognitive Processing in the Second Language. *Language Learning*, 44(2), 283–305.

Maynard, S. K. (1997). Japanese Communication: Language and Thought in Context. *University of Hawaii Press*.

Scollon, R., & Scollon, S. W. (2001). *Intercultural Communication: A Discourse Approach* (2nd ed.). Blackwell Publishing.

van Lier, L. (2004). *The Ecology and Semiotics of Language Learning: A Sociocultural Perspective*. Kluwer Academic Publishers.

Warschauer, M., & Healey, D. (1998). Computers and Language Learning: An Overview. *Language Teaching*, 31(2), 57–71.

Zheng, B., & Warschauer, M. (2015). Participation, Interaction, and Academic Achievement in an Online Discussion Environment. *Computers & Education*, 84, 78–89.

Special thanks to

Aidan Moeck

Tiger Zhao

William Zhou

Soyun Moon

Dahae Cheon

Li-yuan Chiou

*for their insights, encouragement,
and support during the design process.*