

Evaluator: MT

The evaluator is an american-born heritage speaker who grew up in a predominantly Chinese-speaking family. She has experience speaking some Mandarin, though states that proficiency has decreased since growing up and going to college.

User Experience:

During the user experience, the evaluator picked the Chinese option and began practicing. When she arrived at the choice of conversation or scenario roleplay, she did express some confusion about what each option meant, wishing for a short description of each. Additionally, she notes that it would be nice to have a short description of the scenario during the roleplay as a reminder. Since she chose Chinese as the language to learn, the evaluator also noted that being able to observe the pinyin of the words, in addition to the characters, would also be of use to help her with things like pronunciation.

The user also liked the suggestion box and the dictionary options to help her continue the conversation and to look up meanings and translations, though she did note that the dictionary definitions appeared to be too technical at times (for example, when searching up the word 'chicken,' in english, it provided a technical definition with the species name in addition to everything else). Finally, she expressed a wish to be able to edit existing/saved scenarios as well, and noted that combining the login/language selection screen with conversation/roleplay scenario choices could condense the amount of navigation needed to start practicing.

Assessment Portion:

Overall, the evaluator scored both conversation/roleplay experience and language learning helpfulness as 3 out of 5. She cited that sometimes, the responses by the AI felt a bit too long to be realistic for a 'casual' conversation, noting it feels like practice from a textbook. Additionally, she feels that the tool itself, with its suggestion and dictionary options, were good for practicing language aspects like grammar and vocabulary, but wanted more 'realism' in conversations to be useful or real-world scenarios like travel. Finally, one other feature she was interested in being implemented was the ability to highlight or hover over certain words in the chatbox to get the definition and translation, rather than copy and pasting into the dictionary.

Evaluator: LY:

The evaluator was someone who was originally from China before moving to the US. As such, she has experience in both English and Mandarin, but is more limited in her English Proficiency.

User Experience:

During the experience phase, the user first decided to choose the English option to practice. Upon reaching the conversation/roleplay, the user expressed initial confusion between the two options, stating that conversation could emulate certain scenario roleplays if requested. It was only after I showcased the features for custom scenarios (customizing both AI role and overall scenario) that she understood the differences. In the end, she chose scenario roleplay for the user section.

During the roleplay, the user expressed that she liked the dictionary and suggestion options, citing that it helped with knowing what to say to continue the conversation and to also look up certain words that she did not know. The user also intentionally misspelled certain words, used poor grammar, and responded in different languages to see how the AI would react, and liked that it was still able to discern her meaning and continued to use the same language set in the option. She cited this was useful, as she sometimes wanted to express something in her native language, but did not know how to in the target language. Finally, the user expressed enjoyment of the lesson page, particularly its inclusion of both a critique and three lessons to work on.

Assessment Portion:

Overall, the user gave the conversation and roleplay experience of the system a score of 4 out of 5, citing that she did enjoy the conversations, but felt like some of the responses were too long/the AI talked too much, and that in a casual conversation that would be useful for practicing, such responses would be shorter. For language learning effectiveness, she gave the system a score of 5 out of 5, citing her enjoyment of the lessons, suggestions, and dictionary features. For additional feedback, the user also felt like having a way for the system to detect and notify the user of incorrect spelling/grammar immediately would be useful, rather than getting an evaluation at the end with the lesson page. She cites that this immediate feedback would help her as she conversed, keeping the lessons and suggestions fresh in her mind.

Evaluator: WT:

This evaluator was originally born and lived in China before moving to the US, and therefore has high proficiency in his native Chinese language, but lower proficiency in English.

User Experience:

The user first wished to test the Chinese language capabilities of our system, selecting custom scenarios in scenario roleplay as well in the form of a doctor's visit. During his user session, he noted liking the inclusion of the dictionary and suggestion features, but also wanted options to directly highlight words and phrases/hover over them to look them up, rather than relying on typing or copy-and-pasting them into the dictionary. In later scenario sessions, he also wondered if there could be options to choose both a user's native language and target language and a button that can switch text between the two, similar to Google Chrome's ability to translate pages from one language to another, saying it could help people learn overall meanings/translations of the word. The user also expressed a wish to have reactive corrections/highlighting when using the system, as well as suggestions that improve upon one's intended input. For example, when he typed in 'i want to order salmon,' he expressed wanting the system to suggest better ways to phrase the sentence as a response. Finally, in the lessons page, he stated that he wishes there was also lessons that can assist in pronunciation improvement, since the lessons mainly consisted of improving spelling/grammar due to input being text (or audio transcribed into text).

Assessment Portion:

The user gave a score of 4 out of 5 for both the conversation experience and for helpfulness in language learning. More specifically, the user states that some of the responses were too long to seem like it was for a casual conversation. Additionally, he felt that for scenario roleplay, having additional features such as descriptions of the environment or images could help set a better context of the scene and enhance the ability to practice. He also suggested adding in more options to learn how to improve speaking, feeling that it would be useful given the system being designed to emulate casual conversations with others.