

## The accuracy of AI in comprehending disfluent L2 Chinese speech

**Abstract:** Second language (L2) learners often encounter situations in which AI fails to understand their speech when attempting to use AI for language learning. This study examines the speech recognition accuracy of the latest AI system (OpenAI Whisper) when processing disfluent L2 speech. Sixteen L2 Chinese learners at the Intermediate-High to Advanced-Low proficiency levels and fourteen Chinese native speakers participated in the study. Participants read aloud eight Chinese texts with an average length of 1,651 characters, totaling 13,208 characters. During the read-aloud tasks, pronunciation errors included initial errors (0.8%), final errors (6.9%), and tone errors (92.3%), as well as non-fluency phenomena such as frequent pauses, inappropriate pauses, and repetitions (mean number of silent pauses = .64 per second; mean silent pause duration = .49 seconds; mean repetitions = .016 per second). The results show that the AI system was able to ignore non-fluency phenomena such as frequent or inappropriate pauses and repetitions without compromising comprehension. However, learners' errors in initials, finals, and tones negatively affected AI comprehension. Although the AI could infer part of the speech from context, an average of 7.9% of pronunciation errors still led to recognition errors. It is found that AI achieved a recognition accuracy of 99.8% for native read-aloud speech, whereas the recognition accuracy for non-fluent L2 read-aloud speech at the Intermediate-High to Advanced-Low levels was 92.1%.

**Keywords:** AI, L2 Chinese speech, comprehension, fluency, accuracy