

Abstract

Automatic speech recognition (ASR) holds promise for supporting simultaneous interpreting (SI), yet its real-world performance remains inadequately mapped. This two-part study addresses that gap. The first part reports a scoping review of 46 SI studies, revealing inconsistent ASR performance across tools and mixed evidence on whether ASR assistance benefits or disrupts SI product and process. Guided by these findings, the second part benchmarks four ASR tools — Zoom, TerpMate, iFLYTEK, and InterpretBank — across speech conditions varying in difficulty and pace. Within this study, Zoom and TerpMate outperformed iFLYTEK and InterpretBank on accuracy and stability, while InterpretBank and Zoom outperformed the other two on latency. Performance on problem triggers — specialised terminology and numbers — declined across all tools as difficulty increased. The study found that ASR dependability should not be assumed, particularly under challenging conditions, and advocates for multi-dimensional evaluation frameworks in future research.