This paper describes a practical speech-to-speech translation system architecture for operation in face-to-face situations, from a system point of view. The current achievements in speech recognition and language translation technologies are still insufficient to build a real working speech-to-speech translation system, even in a limited domain of application. To build such a working speech-to-speech translation system, it would be necessary to consider a system point of view that includes system design concept, usability and human correction capability. As a possible solution to meet this requirement, we suggest a speech-to-speech translation system architecture that maximizes the keyword transfer rate, a performance measure which we introduce in this paper.

Full Paper