Hello. We are the Kaldi ASR Research Team. My name is Milan, and in this short presentation, I shall briefly outline a technical overview of the Kaldi ASR Toolkit, and the context and relevance of this project as a whole.

This project utilizes the Kaldi ASR Toolkit, which was developed a decade ago by Johns Hopkins University as a means to develop automatic speech recognition (ASR) models. For this project, an ASR model shall be developed for the application of transcribing live Air Traffic Control (ATC) transmissions to text in real time. The model shall be trained using a 30-hour ATC dataset from the Daytona Beach International Airport. The model converts an inputted audio file into data segments consisting of phonemes, which is the basic unit of spoken sound. The phonemes are then converted into triphones, which are three phonemes in sequence. The triphones are then converted into monophones, which are essentially a simplified sequence of spoken speech. The monophones are then converted into words, and strung together into a sentence and outputted as a text file. Due to the complexity of the required calculations to perform the transcription, a sufficiently powerful computer is a necessity to run the Kaldi ASR Toolkit with ease.

The main objectives of the Kaldi ASR Research project are to create an ASR model that shall transcribe live Air Traffic Control (ATC) transmissions to text in real time, and to create a user manual for the Kaldi ASR Toolkit. A proper understanding of the toolkit is vital for both objectives, and shall require many hundreds of hours worth of research to complete.

The relevance of the Kaldi ASR Research project revolves around the implementation of automatic speech recognition (ASR) for ATC applications, specifically for use by student pilots at Embry-Riddle Aeronautical University. In the future, this team hopes to lay the foundation for ASR implementation in the aerospace industry as a whole.

That is all for now. Thank you for your attention.