京都大学	博士（情報学）	氏名	Xun Cao
論文題目	Understanding and Supporting Listening Comprehension of Non-native Speakers		

（学位論文の要約）

Non-native speakers (NNSs) often face comprehension difficulties when listening to native speakers' (NSs') speech. Due to the need to process continuous streams of speech during listening, even when NNSs encounter a comprehension problem, they cannot dedicate the time and resources to resolve such problems. As a result, they often get left behind and sometimes even miss the key points of the speech. Previous studies have suggested many technologies to support NNS comprehension in real-time communication. However, such technologies may impose an extra burden on NNSs if they do not match their current needs. The goal of this thesis is to understand the listening comprehension of NNSs and investigate how to provide them with adaptive support based on their changing needs. Four contributions toward this goal are presented.

- **Understanding Listening Comprehension Problems of NNSs**

  We investigate NNSs' listening comprehension problems and how they burden NNSs. Through an exploratory analysis of the data collected in a laboratory experiment, twelve types of listening comprehension problems are identified. These problems are further classified into three patterns based on their persistence and the time taken to perceive them: "immediate listening comprehension problems" (74% of all problem occurrences), "extant listening comprehension problems" (25%), and "delayed listening comprehension problems" (1%). The findings suggest that different patterns of problems require different support. For example, most of the immediate listening comprehension problems are related to words or phrases and are perceived instantly. For these problems, it would be best to provide support that could immediately solve their problems (e.g., bilingual dictionaries). In addition, a snowball effect of listening comprehension problems (i.e., one problem triggers another problem) seems to occur while NNSs are listening, which is not seen in the listening characteristics of NSs.

- **Investigating NNSs' Use of Support Technology for Solving Problems**

  We report the advantages and disadvantages of providing a support technology to NNSs and investigate how NNSs use the technology to solve their problems. Automatic speech recognition (ASR) transcripts are examined as a representative technology. Results show that ASR transcripts benefit NNSs by helping them solve certain problems (e.g., "do not recognize words they know"), but imperfect ASR transcripts (e.g., errors and no punctuation) sometimes confuse the NNSs and even generate new problems. To make better use of the ASR transcripts, NNSs adopt different strategies of listening and reading the transcripts; some follow them throughout the listening; some only check them when necessary. Furthermore, post-task interviews and gaze analysis of the participants reveal

that NNSs do not have enough time to fully exploit the transcripts. For example, NNSs have difficulty shifting between multimodal contents.

- **Exploring the Impact of Different Designs of Support Technology on NNSs' Listening Comprehension**

  We examine whether and how different designs of ASR transcripts affect NNSs' listening experiences. Two display methods of ASR transcripts are examined: speed-oriented display which shows the interim analysis results of speech input which are likely to be corrected, and accuracy-oriented display which shows transcripts only after the completion of speech input analysis. In the experiment, the NNSs engaged in two listening tasks with ASR transcripts presented via the two display methods. Analysis of the results shows that the more the NNSs pay attention to listening to the audio, the more they tend to prefer speed-oriented transcripts, and vice versa. Mismatched transcripts have negative effects on NNSs' listening comprehension.

- **Predicting Listening Comprehension Problems of NNSs**

  We explore whether it is possible to detect NNSs' problems. While ASR transcripts can improve the listening comprehension of NNSs, certain issues such as vocabulary problems or sentence-level problems remain unsolved. If a system can automatically detect such problems and provide appropriate support (e.g., showing a translated word for a vocabulary problem), the listening comprehension of NNSs might be further improved. The potential of using eye-tracking data for such detection is explored. Gaze data from NNSs who engaged in listening tasks (ASR transcripts were provided) were collected. By applying machine learning techniques, the study shows that it is possible to make reasonably accurate predictions (83.8%) about the types of problems encountered by NNSs using eye-tracking data. In addition, vocabulary problems are characterized by having significantly larger maximum fixation durations, shorter average/maximum saccade lengths, shorter variance of saccade lengths, and more regressive saccades than sentence-level problems. Furthermore, the optimal length of gaze data to predict the listening comprehension problems of NNSs is found to be one second.

In summary, this thesis presents efforts to design an adaptive system for NNSs, which automatically changes the type of support based on their changing needs so that it does not impose additional burden on them. To this end, the work investigates NNSs' comprehension problems, how NNSs use a support technology to solve their problems, how the technology affects their listening experiences, and whether it is possible to detect their problems to provide appropriate support.