

(続紙 1)

京都大学	博士（情報学）	氏名	曹 珣 (Cao Xun)
論文題目	Understanding and Supporting Listening Comprehension of Non-native Speakers (非母語話者のリスニング能力の理解と支援)		
(論文内容の要旨)			
The goal of this thesis is to understand the listening comprehension of non-native speakers (NNSs) and investigate how to provide them with adaptive support based on their changing needs. Four contributions toward this goal are presented. Chapter 1 outlines the thesis, including the research objective, issues and approaches. Chapter 2 describes the background of this thesis. This chapter begins with a general introduction to the listening comprehension problems of NNSs and technologies that support NNSs in real-time communication. Automatic speech recognition (ASR) technologies are presented in further detail. Chapter 3 investigates 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”, “extant listening comprehension problems”, and “delayed listening comprehension problems”. 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) seem to occur while NNSs are listening, which is not seen in the listening characteristics of native speakers. Chapter 4 reports the advantages and disadvantages of providing a support technology (i.e., ASR transcripts) to NNSs and investigates how NNSs use the technology to solve their problems. 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 does not have enough time to fully exploit the transcripts. For example, NNSs have difficulty shifting between multimodal contents. Chapter 5 examines whether and how different designs of ASR transcripts affect NNSs’ listening experiences. Two display methods of ASR transcripts are examined: accuracy-oriented display which shows transcripts only after the completion of speech input analysis,			

and speed-oriented display which shows the interim analysis results of speech input which are likely to be corrected. 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.

Chapter 6 explores 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 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.

Finally, this thesis is concluded and future directions are suggested in Chapter 7. 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.

(論文審査の結果の要旨)

本論文の目的は、非母語話者のリスニング上の問題を分析し、様々なニーズに基づいて適応的に支援する方法を提案することである。本論文では、以下の成果がまとめられている。

1. 非母語話者のリスニング上の問題の分析

非母語話者のリスニング上の問題を明らかにし、それらが非母語話者に与える負担について分析し、12種類のリスニング上の問題を特定している。さらに、問題の持続時間と、非母語話者が問題を認識するのに要する時間に基づいて、これらの問題を、即時(immediate)問題、残存(extant)問題、および遅延(delayed)問題の三種類に分類している。この分析結果により、問題のパターンによって異なる支援が必要であることを明らかにしている。

2. 非母語話者がリスニング上の問題を解決するための支援技術の利用方法の分析

非母語話者に支援技術(自動字幕生成)を提供することの利点と欠点を明らかにし、問題を解決するための非母語話者の技術の利用方法を分析している。その結果、自動字幕生成技術が、既知の単語の聞き取り失敗などの解決に役立つことを見出している。一方で、不完全な自動字幕が非母語話者の聞き取りを阻害することを明らかにしている。さらに、非母語話者は自動字幕の活用に際して、常に字幕を注視する、必要に応じて字幕を注視するなど、必要に応じて異なる戦略を用いていることを見出している。

3. 支援技術が非母語話者のリスニング体験に与える影響の分析

異なる自動字幕の生成技術が非母語話者のリスニング体験にどのように影響するかを分析している。具体的には、自動字幕の2つの表示方法(精度指向型と速度指向型)を取り上げ、音声を聞くことに集中するほど速度指向の字幕を好む傾向があることなどを明らかにしている。また、ニーズに合わない自動字幕は、非母語話者のリスニング上の理解に悪影響を及ぼすことを明らかにしている。

4. 非母語話者のリスニング上の問題の予測

非母語話者にリスニング上の問題が生じることを自動的に予測して、適切な支援を適応的に提供する方法を探索している。そのために、非母語話者が、音声と自動字幕の両方が提供された場合の視線データを収集し分析している。機械学習を適用することにより、非母語話者の視線データから、遭遇する問題(語彙の問題あるいは文レベルの問題など)が予測可能であることを示している。

以上、本論文は、非母語話者のリスニング上の問題と支援技術の利用方法を分析すると共に、視線データを用いて問題の発生を予測する方法を提案している。この研究成果は、非母語話者の支援技術の開発に資するもので、学術上、實際上寄与するところが少なくない。よって、本論文は博士(情報学)の学位論文として価値あるものと認める。また、平成29年2月15日に実施した論文内容とそれに関連した試問の結果、合格と認めた。

Webでの即日公開を希望しない場合は、以下に公開可能とする日付を記入すること。

要旨公開可能日： 年 月 日以降