

音声認識技術を用いた日本語聴解支援システムの検討

Study of Japanese listening comprehension support system using ASR

余 暁俊

Xiaojun Yu

岡山大学 阿部研究室

Abe Laboratory, Okayama University

概要 本研究では、日本語に関する語彙、文法、読解に優れた留学生でも聴解が困難である課題を解決するための、音声認識技術を用いた日本語聴解支援システムの構築を目指す。本報告では、留学生の学習レベルと「日本語教育語彙表」[1]を用いた未知語推定により、聴解支援を提案した。未知語の情報提示方式に関する評価実験をおこない、提案した未知語推定方式及び情報提示方式について検討をおこなった。

1 はじめに

日本における留学生数が年々増加している。しかし、第二言語の理解には限界があり、日本語に関する語彙、文法、読解に優れた留学生でも、教員や日本人学生の話し言葉が聞き取れないというケースが知られている。読解と比べても、聴解いわゆる聞き取りの方が留学生にとってより困難であると考えられる。

音声認識は、話し言葉を文字列に変換する技術である。音声認識技術を活用した講義・講演字幕付けシステム [2]、観光客向け音声翻訳アプリ VoiceTra[3] などがある。しかし、留学生のための聴解支援システムはあまり研究されてない。

本研究では、音声認識技術を活用して、日本語聴解を支援するシステムの構築を目標とする。本報告では、音声入力全体が正しくテキストに変換できることを前提とし、「日本語教育語彙表」と留学生の日本語学習レベルを用いた未知語推定により、異なる未知語情報提示方式の有効性を調べた。また、評価実験から得られた未知語データの分析をおこない、提案した未知語推定方式について検討する。

2 日本語聴解支援の検討

2.1 聴解に影響する要因

聴解は、聴き手が持っているすべての知識を総動員して、内容を予測しながら聴き、意味を構築していく活動である。日本語教育の場合では、水田は、独話における中国人上級学習者の聞き取り上の問題を調査し、学習者の聞き取り上の問題の約半数以上は、「分節化ができない」、「辞書的意味がわからない」であることを報告している [4]。

2.2 未知語推定による聴解支援

2.1 節で取り上げた「分節化ができない」と「辞書的意味がわからない」という二つの聞き取り上の問題に

ついて考察をおこなった。

本研究では、音声認識技術の活用を提案する。留学生が分節化できない音声であっても、十分な学習データから学習された音響モデルと言語モデルを持つ音声認識器の方が分節化できる可能性がある。また、「辞書的意味がわからない」の問題に関しては、単語の意味、用例、同義語、対訳などの情報を提示することによって解決できると考えられる。

しかし、Rubin[5]により整理された第二言語における聴解に与える要因から見ると、留学生の言語能力レベルや、社会的・文化的な背景知識などの違いによって、聴解に与える影響が異なるため、個人ごとの未知語推定は難しい課題である。本研究では、日本語能力試験 [6] を基準に、学習レベルごとの留学生の未知語推定を目標とする。「日本語教育語彙表」の語彙レベルを用いて、留学生の学習レベル以上の単語を未知語とし、未知語の情報を適切な形式で提示することで、聴解の補助になることを目指す。

3 未知語の情報提示に関する評価実験

3.1 実験条件

評価実験には、日本語話し言葉コーパス [7] の一つの模擬講演データの音声ファイルと書き起こしテキストを使用した。長さ約 13 分の音声ファイルを聞いて、下記の点を考慮して、人手で 11 分割した。分割した音声には、1 から 11 まで番号を付けた。

- 一つの発話が終わる所で分ける
- 分割した一つの音声の長さはほぼ均等である

上級日本語学習者レベルである中国人の大学院生 3 名により、評価実験をおこなった。11 個の音声をもとに、書き起こしテキストも同様に分割した。そして、最も意味理解の影響を与える名詞、動詞、形容詞を抽出した。抽出には、形態素解析器 MeCab (ipa-dic 辞書) を用いた。

抽出された名詞、動詞、形容詞に対して、未知語の推定をおこなった。実験説明で使った 1 番音声を除いて、残り 10 個音声に含まれる総単語数は 1986 個である。そのうち、名詞、動詞、形容詞の数は 1087 個である。提案手法により上級学習レベルの留学生にとっての未知語数は 324 個である。

実験で用いる五つの情報提示方式及び音声ごとの提示方式を表 1 に示す。音声のみの場合では、全て聞き取ることは難しいと考えられる。一方で、音声の書き起こしを提示すれば、「音声を聞く」という行動より、「文書を読む」という行動をしていると考えられ、音声の内容を十分に理解できると予測している。音から言葉を思い出せば、聴解の練習に有益であると考え、未知語の発音を表す音素列を提示する方式を提案した。単語の辞書的意味が分からない場合では、母語対訳の提示により解決できることが期待できる。また、今回の実験参加者は 3 名とも中国人であり、日本語と同じく漢字を使用する。未知語の表記をそのまま提示する方式が有効であると考えられる。

実験手順を以下に示す。

1. 音声を流し始めると、画面上に情報が提示される。提示した情報を見ながら順番に音声を聞く。
2. 一つの音声が終わったら、アンケートに記入する。アンケートの質問内容を表 2 に示す。Q1 と Q2 の答えは 1 (0~19%) から 5 (80~100%) の五段階評価であり、Q3 は 1 (とても反対) から 5 (とても賛成) の五段階評価である。
3. 聞いた音声の名詞、動詞、形容詞のみのテキストから聞き取れなかった単語に○をつける。

3.2 実験結果

アンケートの結果を図 1 に示す。音声のみの場合では、出現した単語を全て聞き取ることは難しい。一方で、書き起こしを提示する場合の評価が最もよい。読解が優れた留学生でも、聴解が困難であることが明らかとなったが、この方式は本研究の聴解の補助になる目標と離れている。結果から、中国人の留学生において、未知語の表記を提示する方式はより有効的であることが確認できる。また、音素列を提示しても、聞き取りやすさはあまり変化がなかった。事後の調査により、理由は音素列が分かりにくいである。音素列の代わりに、単語のローマ字表記や、ひらがな表記など単語の発音情報提示方式を検討する。

実験参加者 3 名は実際に聞き取れなかった単語数、26, 45, 31 のうち、提案手法により推定できた単語数は 13, 32, 22 である。推定精度を向上するために、未知語の推定方式に関する検討をおこなう必要がある。

また、音声認識器 Julius の話し言葉モデルキット [8] を用いて、2 から 11 番の音声を認識させた。単語正解精度 (Word Accuracy) は 69.4 % である。実験参加者が実際に聞き取れなかった単語のうち、正しく認識できた単語数は、23, 39, 20 である。全体の約 8 割が認識できた。認識誤りを含む場合でも、留学生にとっての有益な情報を提供できると考えられる。

表 1: 情報提示形式

情報提示形式	音声の番号	情報提示形式	音声の番号
情報なし	1, 2, 7	未知語の表記	5, 10
未知語の音素列	3, 8	音声の書き起こし	6, 11
未知語の対訳	4, 9		

表 2: アンケートの質問内容

Q1	出現した単語をどれくらい聞き取れましたか
Q2	話の内容をどれくらい理解できましたか
Q3	音声のみの場合と比べて、今回の場合は聞き取りやすかったと思いますか (音声の番号が 1,2,7 の時、質問内容は：今回の場合は聞き取りやすかったと思いますか)

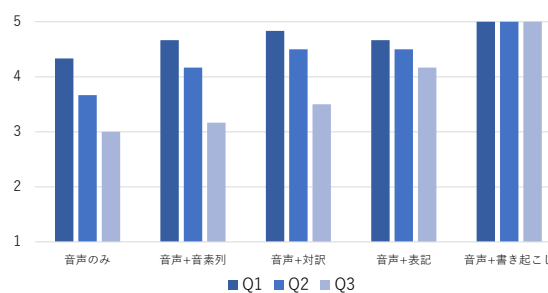


図 1: アンケートの結果

4 まとめ

本報告では、日本語聴解を支援するために、未知語の推定及び情報提示形式について検討し、評価実験をおこなった。実験の結果、上級中国人留学生の場合、漢字表記の未知語をそのまま提示することが、最も聴解の補助になることが分かった。今後は、追加実験により、異なる学習レベルの留学生に対する情報提示方式に関するデータ集計をおこなう。

参考文献

- [1] Yuriko Sunakawa, Jae-ho Lee, Mari Takahara, “The construction of a database to support the compilation of Japanese learners’ dictionaries,” *Acta Linguistica Asiatica*, pp. 97–115, 2012.
- [2] 河原 達也, “ICT 音声認識の活用による講演講義の字幕付与,” *情報処理*, Vol . 56, No . 6, pp. 543–546, 2015.
- [3] Matsuda Shigeki, “Multilingual speech-to-speech translation system: VoiceTra,” 2013 IEEE 14th International Conference on Mobile Data Management, 2013.
- [4] 水田 澄子, “独話聞き取りにみられる問題処理のストラテジー,” *世界の日本語教育*, pp. 49–64, May 1996.
- [5] J. Rubin, “A review of second language listening comprehension research,” *Modern Language Journal*, Vol . 78 , No . 6, pp. 199–221, 1994.
- [6] “日本語能力試験,” <http://www.jlpt.jp/>, accessed Jan. 2018.
- [7] K. Maekawa, “Corpus of Spontaneous Japanese : its design and evaluation,” In *Proc. ISCA & IEEE Workshop on Spontaneous Speech Processing and Recognition (SSPR)*, pp. 7-12.
- [8] “話し言葉モデルキット,” <http://julius.osdn.jp/index.php?q=dictation-kit.html>, accessed Jul. 2018.