

「ゆっくり発話」を促す国際映像会話支援システムの提案

小安 宗徳 大島 直樹
東京電機大学情報環境学部情報環境学科

1. はじめに

母語話者(NS)複数名の中に非母語話者(NNS)1名を含むような、母語の異なる話者同士の多人数会話では、NNSが会話に参加しづらく、相互理解が困難である。解決には、NNSへの文字情報提示[1][2]やNSのゆっくりとした発話が有効である[3]。しかし、従来の研究では会話のリアルタイム性が損なわれる恐れがあった。そこで、本研究では、国際映像会話の場に音声認識機能の導入を検討する(図1)。音声認識により、NNSに文字情報が提示され会話内容理解に役立つ。また、音声認識が介入することで、NSのゆっくりとした発話が引き出される。

本稿では、音声認識によってNSの話速の低下が促されるかに着目して検証を行ったケーススタディの結果について報告する。

2. ケーススタディ

多人数会話では問題が複雑化する恐れがあるため、導入として、母語の異なる2者間で映像会話を収録した。収録は、通常国際映像会話を行う収録条件1(10分間)と、NSが音声認識を用いて国際映像会話を行う収録条件2(10分間)の2水準を設けた。収録条件2ではNSに「話した文章が表示され終わったら、次の話しを始めてください」と教示した。各条件終了後に質問紙への回答とインタビューを行った。

NSは標準語を話す大学生、NNSは日本語検定2級の留学生を選出した。これまでに2組(4名)に収録を行った。1組目のNSは22歳男子大学生、NNSはタイ出身の24歳男子留学生で、母国語はタイ語、マレーシア語、日本の滞在期間は5年であった。2組目のNSは23歳大学院生、NNSは台湾出身の22歳短期男子留学生で、母国語は中国語、日本滞在期間は4ヶ月であった。

3. 結果と考察

収録条件2でNSの話速の低下が促されたかを検証する。NSの発話時間と、発音のモーラ数をカウントした結果、1組目、2組目ともに1秒間中の平均モーラ数が

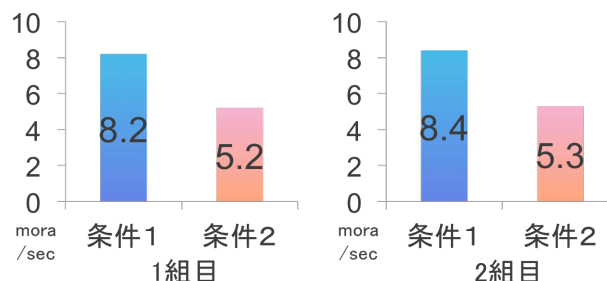


図2 NSの1秒間中の平均モーラ数

減少した(図2)。このことから、手法によりNSの話速が低下したといえる。これはNSが自身の発話を音声認識機能に正しく認識させようと試みた結果、音節1つ1つをはっきりと発音することになったためと考えられる。

ほかにも、収録条件2のNSの発話と発話の間隔が両組とも増加していた。これはNSが音声認識結果の表示完了を待ってから次の話を始めたためと考えられる。

1組目のNNSからは「前の会話を思い出したいときに見ていた」という回答を得た。2組目では「わからない単語があったときに見ていた」という回答を得た。

4. まとめ

本研究では、NSが音声認識システムを使用しながら映像会話を行う手法を提案した。ケーススタディの結果から、音声認識を導入した会話条件においてNSの話速の低下がみられ、また文字情報の提示によってNNSの会話内容理解の支援につながる効果がみられた。

今回の検証ではNSの話速低下に着目してケーススタディを実施したが、今後はNSが発話を終了してから、音声認識完了までの間を待つ心理状態の調査を行う。NSが音声認識させながら会話を行うことによって生じる心理的な負担を軽減する手法を検討することで、NSとNNSが会話内容を相互に理解し、楽しく会話することのできる場の構築を支援するシステム設計を行う。

参考文献

- [1] Clark, Herbert H.; Brennan, Susan E., "Grounding in communication", 1991
- [2] 埴裕美, 宋曉宇, 井上智雄, "母語話者の文字入力による非母語者との会話支援-母語話者による会話中のテキスト入力が音声会話に与える影響-" 2016/5, pp.140
- [3] 小山内一樹, 徳永弘子, 武川直樹, 斎藤博人, "話速変換会話における映像音声の同期・非同期再生条件の比較検討-話者が伝える宛先指定の強さは変化するのか-" 2017/4, pp.1

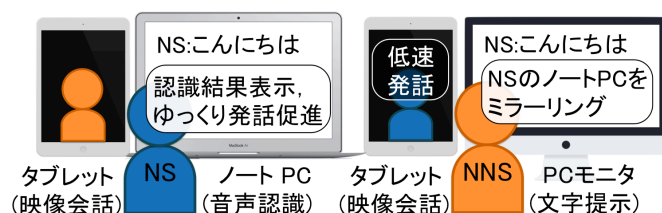


図1 音声認識機能を用いた映像会話