

研究種目：若手研究 (B)

研究期間：2007～2010

課題番号：19700627

研究課題名 (和文) 聴覚障害者のためのリアルタイム字幕呈示システムにおける最適情報呈示に関する研究

研究課題名 (英文) An optimum method of displaying speakers' face images and captions for a real-time speech-to-caption system for the deaf

研究代表者

黒木 速人 (KUROKI HAYATO)

筑波技術大学・産業技術学部・准教授

研究者番号：00345159

研究分野：人間情報工学

科研費の分科・細目：科学教育・教育工学・教育工学

キーワード：ヒューマン・インタフェース, 認知科学, 障害者支援, 高等教育支援, 教育工学

1. 研究計画の概要

音声認識を用いたリアルタイム字幕呈示システムにおいて、現状では精度100%の字幕の生成は困難であり、字幕は誤認識を含む不完全な文章になる。しかしながら、重要なのは最終的な伝達内容に対する理解を100%に近づけることであり、そのためには不完全部分を補うための情報を内容理解の阻害なしに最適な状態で呈示することが重要となる。

そこで本課題では、現在までに行ってきた字幕生成過程における精度向上のみならず、字幕呈示過程においてもヒトが阻害なく情報補完・統合を行うために、誤認識を含む不完全文と発話時のノンバーバル情報を呈示させる際に、どのような顔の部位や情報間の呈示タイミングなどと言った、情報呈示に関する最適な方法を探る研究を行う。

2. 研究の進捗状況

検討すべき内容を以下の小項目4つに分類し、現在までに3つの小項目に関して検討を行った。

(1) 不完全な文に対し話者の顔などのノンバーバル情報の活用による効果とその最適表示の方法

①ノンバーバル情報として顔のどの領域を呈示するのが良いか

②ノンバーバル情報と字幕との呈示タイミングをどう制御するのが良いか

③ノンバーバル情報と字幕を同期させて呈示する際、どの程度の精度が要求されるか (ヒトの時間ずれの検知限の把握)

④ノンバーバル情報と字幕をある時間間隔で呈示する際、2つの方法をどこまで統合した情報として保持できるか (ヒトの情報

に対する時間的統合限界)

①では、不完全文からなる字幕のみの呈示(字幕のみ)を基準として、不完全文に対し話者の発話時の顔全体映像を付加した場合(字幕+顔)、同様に口元近接映像を付加した場合(字幕+口元)の3種で比較した。完全文(正解文)に対する回答文の正答率を文理解の向上の指標として健聴者と聴覚障害者に対し実験を行った結果、いずれの被験者群においても正答率は、「字幕のみ」<「字幕+顔」<「字幕+口元」となる傾向を示した。このことからノンバーバル情報を付加することで文理解が向上する傾向が統計的有意差として確認された。

②では、ノンバーバル情報として話者の発話時の顔全体映像を用い、ノンバーバル情報に対して不完全文の呈示に時差(-5秒～+5秒)を設けた試料に対し、いずれの時差が文理解の向上に貢献するかを比較した。①の実験と同様に正答率をそれぞれの被験者群で求めた結果、正答率は時差なし(±0秒)ではあまり向上が認められず、字幕先行側に1秒ずらした場合(-1秒(字幕先行1秒))が高くなる傾向を示した。また字幕先行側に外れるほど正答率が単調増加する傾向を示した。これらの傾向は、何れの被験者群においても統計的有意差として確認された。

③では、話者の顔情報と不完全字幕の呈示時差(-5秒～+5秒)を意図的に設け、その呈示時差を被験者が検知できるかどうかを定量的に計測することを試みた。その結果、被験者により異なる傾向を示したが、概ね呈示時差0～+2秒の間に許容限があると推察される結果を得た。しかしながら、この定性実験の結果、検知限には個人差が大きいことがわかった。今後の定量実験の実施方法と解析方

法を再考する必要がある。

3. 現在までの達成度

③やや遅れている

(理由)

研究項目として挙げている項目のうち、項目①～②の呈示部位・呈示時差に関しては研究をほぼ完了した。これらに関しては、現在、論文化の作業を行っている段階である。項目③の同期呈示の検知限においては、定性実験がほぼ終了した段階である。定性実験の結果、検知限の範囲は概ね把握することができたが、この検知限の定性実験結果を見る限り、他の実験結果と比較して非常に個人差が大きいことがわかった。また被験者群に関しても同様で、聴覚障害者群・健聴者群で群ごとの傾向は認められない結果となっている。そのため、定量実験の実施方法と解析方法を当初の通りではなく再考する必要が生じた。これが現在までの達成度の自己評価を「③やや遅れている」とした理由である。項目④の複数情報の時間的統合限に関しては、本年度予定通り取り組む予定である。

4. 今後の研究の推進方策

先述通り、項目①～②の呈示部位・呈示時差実験に関しては、ほぼ実験実施・結果解析が終わり、得られた結果を順次原著論文としてまとめている段階である。項目③の同期呈示の検知限の実験に関しては、定性実験の結果により、検知限の範囲は概ね把握することができたが、個人差が非常に大きいため、今後の定量実験の手続きと解析方法を早急にまとめて実験に取り組む必要がある。項目④の時間的統合限は予定通り進めていく予定である。

5. 代表的な研究成果

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計5件)

- ① 黒木速人, 三好茂樹, 白澤麻弓, 河野純大, 萩原彩子, 磯田恭子, 田中美希, 金澤貴之, 大倉孝昭: アメリカにおける聴覚障害者のための文字による情報保障技術に関する報告, 筑波技術大学テクノレポート, Vol.15, pp.181-185, 2008, 査読無
- ② 中野聡子, 金澤貴之, 牧原功, 黒木速人, 上田一貴, 井野秀一, 伊福部達: 聴覚障害者向け音声同時字幕システムの読みやすさに関する研究(1)-改行効果に焦点を当てて-, ヒューマンインタフェース学会論文

誌, Vol.10, pp.435-444, 2008, 査読有

- ③ 中野聡子, 金澤貴之, 牧原功, 黒木速人, 上田一貴, 井野秀一, 伊福部達: 音声認識技術を利用した字幕呈示システムの活用に関する研究-聴覚障害者のニーズに即した呈示方法-, メディア教育研究, Vol.5, pp.63-72, 2008, 査読有
- ④ 中野聡子, 牧原功, 金澤貴之, 中野泰志, 新井哲也, 黒木速人, 井野秀一, 伊福部達: 音声認識技術を用いた聴覚障害者向け字幕呈示システムの課題-話し言葉の性質が字幕の読みに与える影響-, 電子情報通信学会論文誌 D, Vol.J90-D, No.3, pp.808-814, 2007, 査読有
- ⑤ 黒木速人, 井野秀一, 中野聡子, 堀耕太郎, 伊福部達: 音声同時字幕システムにおける内容理解の向上を目的とした話者の顔情報の呈示方法, ヒューマンインタフェース学会論文誌, Vol.9, pp.205-212, 2007, 査読有

[学会発表] (計1件)

- ① H. Kuroki, et al., Effective Timing of Displaying Speakers' Face and Captions to Achieve Higher Comprehension, 23rd Annual International Technology and Persons with Disabilities Conference (CSUN2008), 2008.3.14, Los Angeles, USA.

[その他]

ホームページ

筑波技術大学・機関リポジトリ

<http://www.a.tsukuba-tech.ac.jp/repo/dspace/>