

【基盤研究(S)】

人文社会系 (人文学)



研究課題名 多用途型日本手話言語データベース構築に関する研究

工学院大学・情報学部・教授

ながしま ゆうじ
長嶋 祐二

研究課題番号：17H06114 研究者番号：50138137

研究分野：言語学

キーワード：日本手話、手話言語収集、多用途データベース、アノテーション

【研究の背景・目的】

手話は、聴覚障害者のコミュニケーション手段の一つであり、音声言語とは異なる文法体系をもつ独立した対話型の自然言語である。

日本語の音声や言語データは、国立情報学研究所において、音声資源コンソーシアム(SRC: Speech Resources Consortium)が設立され、日本語の研究の発展に寄与している。しかし、音声言語と比較して手話研究は、言語学的にも工学的にも格段に遅れていると過言ではない。原因の一つとして、手話研究において、言語資料の収集方法が定まていないこと、研究者が共通で利用でき議論の土台となる手話言語データベースが存在しないことがある。

本研究では、多用途型日本手話データベースを構築するための方法論の検討、そして、データベースの構築を目的とする。開発を目指す多用途手話データベースは、言語学はもとより、工学、認知科学など学際的な分野での利用を想定した、手話の語彙と文章から構成する。さらに、本研究では、収録した手話のデータの分析を支援するアノテーションシステムの開発も目指す。

【研究の方法】

本研究は、研究者が共通に利用し議論の基礎にすることが可能な、データベースの構築を4つの課題を通して行う。

・課題1:言語資料収集では、言語学、工学などの学際分野で利用可能なデータとするため、収録する手話の語彙、手話文、言語資料提供者などについて検討を行う。

・課題2:言語資料収録では、専門分野ごとの手話の入力形式、収録の時間ならびに空間解像度、保存データ形式、保存方法などについて検討する。想定する入力形式は、3次元動作データ、多視点カメラ画像、距離画像である。一部の収録語彙では、弁別的特徴の分析のため、これら3種類のデータの同期収録を試みる。尚、3次元動作データは、世界最高精度の光学式モーションキャプチャによって収録する。

・課題3:多用途データベース構築では、様々な入力形態の同一データをどのように構造化して、実際の研究に利用可能なデータベースをどのように構築するかを検討する。

・共通課題として、データベースが効率的に機能するためには、各形式のデータが有機的に関連していることが重要である。この課題は、3種類の多用途データ間の新たなアノテーション方式の開発を行う。

図1に、開発中の多用途データのアノテーション支援システムを示す。

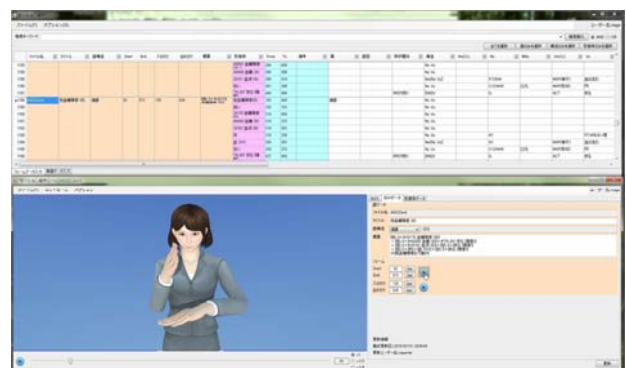


図1 多用途データのアノテーション支援システム

【期待される成果と意義】

本研究では、様々な時間、空間解像度、質の異なる入力媒体からなるデータを同期された多次元的な手話データベースの構築方法の提案と開発である。この方法論が提案できれば、国立学情報研究所の音声資源コンソーシアムへ手話言語資料の登録も可能となる。登録ができれば、様々な領域へ研究者の望む言語資料が提供可能となり、日本の手話研究の発展に大きな貢献となり得ると共に、新たな研究方法の発見など、その有用性は計り知れない。

本研究成果により、手話母語者に対する新たな情報保障の確立などQOLの向上に大きく寄与できる。

【当該研究課題と関連の深い論文・著書】

- ・渡辺桂子, 長嶋祐二: 手話形態素辞書作成のための情報入力支援システム, 電子情報通信学会論文誌 D, Vol.J100-D, No.3, pp.298-309, 2017.
- ・渡辺桂子, 長嶋祐二: 医療用語の手話表現の検討とその単語データベースの構築, 電子情報通信学会論文誌 D, Vol. J99-D No.1, pp.76-89, 2016.
- ・Y. Nagashima, et al.: A Support Tool for Analyzing the 3D Motions of Sign Language and the Construction of a Morpheme Dictionary, Lecture Notes in Computer Science 618, Springer, pp.124-127, 2016.

【研究期間と研究経費】

平成29年度-32年度 109,200千円

【ホームページ等】

<http://www.ns.kogakuin.ac.jp/~wwc1015/>