

平成 22 年 5 月 10 日現在

研究種目：基盤研究(C)

研究期間：2007～2010

課題番号：19520434

研究課題名（和文） フレーム意味論に基づく多言語語彙検索ソフトウェアの開発

研究課題名（英文） Development of software tools for multilingual lexicons based on frame semantics

研究代表者

佐藤 弘明 (SATO HIROAKI)

専修大学・商学部・教授

研究者番号：60187226

研究代表者の専門分野：コーパス言語学、自然言語処理

科研費の分科・細目：言語学・英語学

キーワード：フレーム意味論、FrameNet、語彙、コーパス、項構造、日本語、英語、スペイン語

1. 研究計画の概要

(1) **研究の全体構想**：フレーム意味論に基づく英語語彙研究のために自己開発してきたソフトウェアと映画コーパス検索システムをスペイン語、ドイツ語、日本語などの多言語に対応させる。それぞれの言語間での対照研究を可能とする検索機能を開発して、世界の研究者に対してインターネット上で無料公開をして、フレーム意味論関連分野の研究発展に寄与する。

(2) **研究の学術的な特色**：平成 17 年以前の筆者の映画コーパスの研究では、「言語表現の研究には、それが使われる場面にアクセスできるコーパスが有用である」ことを示した。本研究の特色は、話者が言語を用いる際に想起する背景知識を重視する FrameNet の研究に、映画コーパス研究の成果である検索システムとそれが提供する場面情報とを組み合わせる点である。既に多くの研究者に注目されている FrameNet 研究に映画コーパスの研究成果を応用することで、科学研究費補助金を受けて開発を続けてきた映画コーパスの研究成果が海外の多くの研究者に還元することが期待できる。

2. 研究の進捗状況

(1) 本研究は、語彙検索のためのソフトウェア開発を主たる目的としているため、**語彙データ入力**のための大きな予算は計上していない。そのため、語彙データ入力を行っている他の研究組織との協力関係が重要となる。英語、日本語、スペイン語に関しては、それぞれの言語の語彙データを入力している組織がそれぞれ過去 3 年間に大きな予算を獲得してきた。それぞれの組織が新しい語彙

データの入力作業を行った結果、そのデータを本研究で改良した検索ソフトウェアで利用することが可能となり、新しいデータベースとしてインターネット上に無料で公開できた。また、これらの組織の研究者と意見交換を行うために、研究所を訪れ、直接、研究者と会って、ソフトウェアの改良を行った。

映画コーパス用の検索ソフトウェアは、データベース 4th Dimension で開発を続けてきたが、これを FrameNet 用検索ソフトウェアと連携させるために、Web ブラウザーから利用できるように改良した。この結果、1 つの Web ブラウザーから映画コーパスと FrameNet 語彙データとを検索可能になった。

(2) **何をどこまで明らかにしたか**

①英語と他の言語の語彙を対照研究できるように、英語の FrameNet 用検索ソフトウェアの全体のシステムを改良し、それをインターネット上に無料公開した。

②Web ブラウザーから新しく入力された Spanish FrameNet のスペイン語データ検索が行えるシステムを開発して、インターネット上に無料公開した。

③Web ブラウザーから日本語フレームネットの日本語データ検索が行えるシステムを開発して、インターネット上に無料公開した。

3. 現在までの達成度

②おおむね順調に進展している。

(理由)

英語、スペイン語、日本語語彙データベースと、映画コーパスに関しては、新しいデータを検索出来るソフトウェアの開発が成功している。しかし、ドイツ語に関しては、新しい語彙データが利用できる状況にはなっ

いないため、ソフトウェアの機能の改良は行ったが、ドイツ語の新しいデータを本研究で開発したソフトウェアで検索することができていない。

4. 今後の研究の推進方策

ドイツ語語彙データに関しては、新しいデータが公開された場合に、直ぐに本研究で開発する検索ソフトウェアで利用できるようにソフトウェアを改良する。英語、スペイン語、日本語語彙データベースと、映画コーパスに関しては、これまでの研究成果を国際学会等で発表し、検索ソフトウェアを多くの研究者に使わせ、研究成果を広く世界に公開する。

5. 代表的な研究成果

[雑誌論文] (計 3 件)

①Hiroaki Sato, "Browsing Japanese FrameNet with FrameSQL," Proceedings of Pacific Association for Computational Linguistics 2009 (PACLING 2009), 2009 年, CD-ROM収録のためページ番号なし, 査読有り.

②佐藤弘明, 「iPod touchを利用した多言語マルチメディア・パラレル・コーパスの構築」, 『ことばの普遍と変容』, pp. 67-74, 専修大学社会知性開発研究センター, 2009 年, 査読無し.

③Hiroaki Sato, "New Functions of FrameSQL for Multilingual FrameNet," Proceedings of The sixth international conference on Language Resources and Evaluation (LREC 2008), 2008 年, CD-ROM収録のためページ番号なし, 査読有り.

[学会発表] (計 3 件)

①Hiroaki Sato, Browsing Japanese FrameNet with FrameSQL, Pacific Association for Computational Linguistics (PACLING 2009), 2009 年 9 月 3 日, 札幌・北海道大学.

②Hiroaki Sato, Search Functions of FrameSQL for FrameNet, International Conference on Middle and Modern English Corpus Linguistics, 2009 年 7 月 9 日, オーストリア・インスブルック.

③Hiroaki Sato, "New Functions of FrameSQL for Multilingual FrameNet," The sixth international conference on Language Resources and Evaluation, 2008 年 5 月 28 日, モロッコ・マラケシュ.