

様式 C-7-1

平成19年度科学研究費補助金実績報告書（研究実績報告書）

1. 機関番号 1 4 6 0 3 2. 研究機関名 奈良先端科学技術大学院大学
3. 研究種目名 基盤研究（B） 4. 研究期間 平成18年度～平成20年度
5. 課題番号 1 8 3 0 0 0 5 1
6. 研究課題名 語構成を考慮した多言語の語彙および用語解析システムの研究開発

7. 研究代表者

研究者番号	研究代表者名	所属部局名	職名
1 0 2 1 1 5 7 5	<small>カガナ</small> マツモト, ユウジ 松本, 裕治	情報科学研究科	教授

8. 研究分担者(所属研究機関名については、研究代表者の所属研究機関と異なる場合のみ記入すること。)

研究者番号	研究分担者名	所属研究機関名・部局名	職名
6 0 2 7 2 6 8 9	<small>カガナ</small> イヌイ, ケンタロウ 乾, 健太郎	情報科学研究科	准教授
8 0 3 7 9 5 2 8	<small>カガナ</small> アサハラ, マサユキ 浅原, 正幸	情報科学研究科	助教
5 0 2 7 8 8 1 8	<small>カガナ</small> ハシモト, キョウタ 橋本, 喜代太	大阪府立大学・人間社会学部	准教授
	<small>カガナ</small>		
	<small>カガナ</small>		

9. 研究実績の概要(国立情報学研究所でデータベース化するため、600字～800字で記入。図、グラフ等は記載しないこと。)

下欄には、当該年度に実施した研究の成果について、その具体的内容、意義、重要性等を、交付申請書に記載した「研究の目的」、「研究実施計画」に照らし、600字～800字で、できるだけ分かりやすく記述すること。また、国立情報学研究所でデータベース化するため、図、グラフ等は記載しないこと。

次の3項目に分けて研究を実施した。

- (1) 中国語および日本語の合成語の語構成の整理と分類：中国語と日本語の合成語を複合語と派生語に分類し、語構成に関する分類を行なった。特に、中国語の3文字からなる合成語について、構造の記述を網羅的に行い、自動分類に関する実験を行った。
- (2) 言語解析用辞書の語彙項目の拡充：形態素解析器「茶筌」用の辞書の拡張を目指して、大規模テキストからの未知語の自動抽出に関する検討を行った。また、中国語と日本語の新規の複合語の抽出と登録を行った。中国語について、合計約13万語の辞書を構築した。
- (3) 複合表現・専門用語の自動解析手法の開発：特定分野に頻出する用語の抽出と意味分類を目指し、Web上の大規模なデータからの自動抽出に関する研究を行った。特に、インターネット上で構築が進んでいるWikipediaを対象とし、そのDOM構造を利用して、用語の意味分類を行なう手法を提案し、固有表現辞書の自動構築を行った。

※ 成果の公表を見合わせる必要がある場合は、その理由及び差し控え期間等を記入した調書(A4判縦長横書1枚)を添付すること。

10. キーワード

- | | | |
|------------|-------------|-----------|
| (1) 自然言語処理 | (2) 辞書 | (3) 用語解析 |
| (4) 形態素解析 | (5) 語構成 | (6) 多言語処理 |
| (7) 複合語 | (8) コロケーション | (裏面に続く) |

11. 研究発表（平成19年度の研究成果）

〔雑誌論文〕 計（2）件

著者名	論文標題			
原一夫, 新保仁, 松本裕治	アラインメントと機械学習を応用した並列句解析			
雑誌名	査読の有無	巻	発行年	最初と最後の頁
人工知能学会論文誌	有	22	2007	248-255

著者名	論文標題			
中川哲治, 松本裕治	大域的な情報を用いた未知語の品詞推定			
雑誌名	査読の有無	巻	発行年	最初と最後の頁
情報処理学会論文誌	有	49	2008	1437-1450

著者名	論文標題			
雑誌名	査読の有無	巻	発行年	最初と最後の頁
	有	6	2008	53-60

〔学会発表〕 計（1）件

発表者名	発表標題		
Jia Lu, Masayuki Asahara and Yuji Matsumoto	Analyzing Chinese Synthetic Words with Tree-based Information and a Survey on Chinese Morphologically Derived Words		
学会等名	発表年月日	発表場所	
SIGHAN Workshop on Chinese Language Processing	2008年1月11日	Hyderabad, India	

〔図書〕 計（0）件

著者名	出版社		
書名	発行年	総ページ数	

12. 研究成果による産業財産権の出願・取得状況

〔出願〕 計（0）件

産業財産権の名称	発明者	権利者	産業財産権の種類、番号	出願年月日	国内・外国の別

〔取得〕 計（0）件

産業財産権の名称	発明者	権利者	産業財産権の種類、番号	取得年月日	国内・外国の別

13. 備考

※ 研究者又は所属研究機関が作成した研究内容又は研究成果に関するwebページがある場合は、URLを記載すること。

<http://cl.naist.jp/>