

論文 / 著書情報
Article / Book Information

題目(和文)	コーパスに基づく辞書記述の精緻化の研究
Title(English)	
著者(和文)	柏野和佳子
Author(English)	Wakako Kashino
出典(和文)	学位:博士(学術), 学位授与機関:東京工業大学, 報告番号:甲第10284号, 授与年月日:2016年7月31日, 学位の種類:課程博士, 審査員:高村 大也,新田 克己,寺野 隆雄,三宅 美博,奥村 学
Citation(English)	Degree:Doctor (Academic), Conferring organization: Tokyo Institute of Technology, Report number:甲第10284号, Conferred date:2016/7/31, Degree Type:Course doctor, Examiner:,,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	審査の要旨
Type(English)	Exam Summary

論文審査の要旨及び審査員

(2000字程度)

報告番号	乙 第 号	学位申請者	柏野 和佳子	
論文審査員	氏 名	職 名	氏 名	職 名
	主査 高村 大也	准教授	三宅 美博	教授
	奥村 学	教授		
	寺野 隆雄	教授		
	新田 克己	教授		

汎用の基本的辞書の整備は、言語学上も言語処理の応用上も重要である。しかしながら従来の辞書は、その構築過程において編集者の知識や内省への依存度が高いことなどにより、これまで以上に精度の高い処理を実現するための情報源としては不十分であった。本論文は、大規模な言語コーパスに基づく客観的・分析的手法を導入することを柱として、これまでよりも精緻な汎用辞書の記述方法を提案するものであり、「コーパスに基づく辞書記述の精緻化の研究」と題し、全6章より構成されている。

第1章「序論」では、本研究の背景として、言語資源としての辞書の構築の重要性、およびコーパスを活用する辞書記述の重要性を述べている。また、本研究で着目する辞書情報として、辞書構築の骨格となる見出し語選定、語義区分、多義構造の記述、および解析・生成処理の精度向上が特に期待できるコロケーション(共起情報)の記述、使用域の記述を挙げている。

第2章「従来の辞書とその問題点」では、まず、従来の(人間用の)国語辞典について検討し、構築作業の主観依存性、非網羅性、暗黙の前提による情報の省略、を問題点として述べている。次に、従来の計算機用日本語辞書について検討し、非構造化、生成処理用の情報不足、などの問題点を指摘している。

第3章「コーパスを用いた辞書記述」では、前章で指摘した問題点のコーパスによる解決可能性を検討している。まず、近年構築され利用が可能になった『現代日本語書き言葉均衡コーパス(BCCWJ)』を中心に現在利用できる日本語コーパスの概要を述べている。次に、コーパスに基づく主な利点として、(1)頻度情報や実際の用例に基づき辞書情報の抽出や記述が客観的・明示的に行えること、(2)実際の用例を網羅的に収集できること、(3)多様な資料から得る用例により使用域を具体的に記述できること、を述べている。それらを踏まえ辞書記述方法の基本設計を行い、提案法の新規性と期待される効果について述べている。

第4章「コーパスに基づく見出し語選定と語義区分の精緻化」では、まず、現代語と古語の間に位置する「古風な語」を例に、コーパスの頻度情報を使った見出し語選定方法を提案している。次に、用例や用法に着目する12通りの「下位区分の判断基準」を設け、説明可能な一貫した区分を行える語義の区分の方法を提案している。最後に、多義構造を比喻による意味の拡張という観点から捉え、多義構造の記述を明示的に行う方法を提案している。

第5章「コーパスに基づく辞書記述内容の精緻化」では、まず、従来の多義性解消の試みにおいては主に個々の語の多義性のみが着目されてきたところ、複数の語の結び付きであるコロケーションの多義性という観点からの検討が重要であることを指摘し、その解決案として、文型情報に加え、名詞句内の先行句に関する辞書情報を記述するというコロケーションの記述方法を提案している。次に、文体情報を使用域の記述として取り込むために、文体的特徴を表す分類指標を設計し、コーパスに文体情報を付与するアノテーション方法を提案している。さらに、「古風な語」を例に、コーパスから得られる出典情報を活用して辞書記述を行う方法を提案している。最後に、「外来語」を例に、新聞記事データベースの頻度情報より時間的使用推移の分類型を得て、それを辞書に記述する方法を提案し、その効果を説明している。

第6章「結論」では、本研究で達成した事項について述べ、その後、今後の課題について述べている。

以上を要するに、本論文は、従来の辞書記述において、様々な制約によって十分な記述が行われていなかった側面、および編集者の主観的判断に強く依存していた側面のうち、応用上あるいは言語学上の観点から特に重要と考えられる点に着目し、コーパスに基づく辞書記述の精緻化手順を具体的に提案し、客観的・明

示的・網羅的・具体的な記述の実現性を示している。このことは、学術上貢献するところが大きい。よって、博士(学術)の学位を授与するに十分な価値をもつものと認められる。

注意：「論文審査の要旨及び審査員」は、東工大リサーチポジトリ(T2R2)にてインターネット公表されますので、公表可能な範囲の内容で作成してください。