

## 審査の結果の要旨

氏名 宮田 玲

近年、多言語での情報提供の必要性はますます高まっているが、予算の問題などから、機械翻訳（MT）に頼っている組織も多い。とりわけ自治体のような公共組織では正確な情報提供が求められるが、近年の機械翻訳精度の向上にもかかわらず、日本語を起点言語とする MT は実用に求められる水準に達しているとは言いがたい。

本論文は、その問題を解決するために、文書の構造化、制限言語（CL）、用語統制を導入することで、起点言語の可読性を改善しつつ MT による目標言語の質を実用精度にすることを目的とする。第 1 章では、こうした問題背景を整理するとともに、文単位の MT の限界を指摘し、制限オーサリングという解決を提案する。また、社会的必要性も考慮し、対象を自治体文書に、言語方向を日英に絞り込む。第 2 章では関連研究を調査し、CL と文書構造の統合という枠組みが、提案はあるものの具現化されていないことを示す。

第 3 章から第 5 章は、要素規則と技術の開発にあてられる。第 3 章では文書構造を分析し、構成要素が Darwin Information Typing Architecture（DITA）のタスク文書の枠組みで定義されることを示すとともに実際に自治体文書向け要素を確定している。第 4 章では CL 規則策定のプロトコルを開発し、合計で 60 の CL 規則を定め、規則に基づく MT と統計的 MT それぞれ 2 つを対象にその有効性を起点言語の質の維持改善と目標言語の質の改善から評価し確認している。第 5 章では、自治体関連用語の統制可能性を検討し、それにより実際にカバレッジと言語間対応の双方が改善されることを実証的に示している。

第 6 章では、文書構造要素に依存した CL を定義する。これにより、文単位という MT の制約を乗り越え、適切な訳し分けが可能になることを示している。

第 7 章から第 9 章では、以上の規則を組み込んだ文書執筆支援（オーサリング）システム MuTUAL の実装と評価にあてられる。第 7 章でシステムの説明がなされる。第 8 章では CL チェッカの有効性を、また第 9 章ではシステムの利用可能性を、それぞれ研究協力者の参加に基づき評価し、有効性と実用可能性を示している。

本研究は、情報アクセスという図書館情報学の中核的課題の一環として多言語情報提供を捉え、潜在的に無限の文を処理するという言語処理の視点に対して有限の知識を編成するという図書館情報学の理論的視点を導入することで、今後の人間＝機械協調による多言語情報提供の有望なモデルを構築するとともに、それを実現するための技術的な詳細を詰めることで、理論的枠組みの有効性を示している。さらに、文書執筆支援システムを構築し実用可能性を実証的に評価することで、提案した枠組みと技術が現実に応用可能であることを示している。以上の点で、図書館情報学、多言語流通モデリング、テクニカルライティング、MT の研究に対する貢献は極めて大きい。よって、本研究は博士（教育学）の学位を授与するにふさわしい水準にあるものと判断された。