

氏 名（本 籍）	Nguyen My Chau（ベトナム）
学 位 の 種 類	博 士（工学）
学 位 授 与 番 号	甲第 268 号
学 位 授 与 日 付	平成 17 年 9 月 14 日
専 攻	電子情報システム工学専攻
学 位 論 文 題 目	日本語－ベトナム語機械翻訳に関する研究 ―連体修飾構造を中心に― (A Study on Japanese-Vietnamese Machine Translation ― Focusing on Adnominal Embedding Structure ―)
学位論文審査委員	(主査) 教 授 池 田 尚 志 (副査) 教 授 伊 藤 昭 教 授 速 水 悟 教 授 寺 島 隆 吉

論文内容の要旨

本論文は日本語からベトナム語への機械翻訳の方法について言語学的に分析し、特に両言語で違いの大きい連体修飾構造の日本語－ベトナム語対照に関して翻訳規則を提案している。また当該研究室で開発している機械翻訳システム jaw/Vietnamese 上に、提案した翻訳規則を実装し、翻訳結果を評価・分析している。このほかにも多くの例文について jaw/Vietnamese 上で翻訳実験を行い、翻訳システムおよび翻訳規則の開発に多くの提言を行っている。

ベトナム語は語形変化がない孤立語であり文要素の並ぶ順序は基本的に SVO である。ベトナム語では語順は文の正確な意味を表すために極めて重要である。一方日本語は膠着語に属し、基本的語順は SOV であるが、語順規則は比較的ゆるやかである。このようにベトナム語と日本語は構造的に大きな違いがあるが、中国の漢字文化を介して共通の面もみられる。機械翻訳の研究では、この両言語の対照言語学的研究がまず最初に重要である。日本語－ベトナム語機械翻訳に関する研究は他にまだほとんどなく、本論文はその開拓の第一歩と位置づけられる。

本論文は序論と結論を含めて 7 章で構成されている。

第 1 章は、研究の背景と目的および概要を述べている。

第 2 章では、機械翻訳研究の歴史と現状について概観している。また、ベトナム語に関する自然言語処理・機械翻訳の状況について述べている。

第 3 章では、ベトナム語の文法構造の概要について述べている。

第 4 章では、日本語・ベトナム語機械翻訳システム jaw/Vietnamese について述べ、100 例文に対する翻訳実験について述べている。jaw/Vietnamese システムは当該研究室で開発している、日本語をベトナム語に翻訳する機械翻訳システムである。システムを検証するために、ある英作文例文集中の 1102 例文からランダムに 100 例文を抽出し、

jaw/Vietnamese で翻訳実験を行って (1047 パターンの翻訳規則を記述している)、翻訳規則の記述方式の設計についての検討、翻訳実験結果の評価分析を行っている。それに基づいて翻訳規則およびシステムの修正を行い、改良した結果、100 例文に対して約 90% の正解率を得ている。

第 5 章では、日本語の連体埋め込み修飾表現のベトナム語への機械翻訳方法を述べている。日本語では埋め込み文による修飾は「連体形の動詞/形容詞+名詞」という構造的には単一の形をとるが、ベトナム語では関係節構造や名詞節構造などいくつかの構造をとるため、その構造分類規則が必要になる。さらに、先行詞や名詞化のための形式名詞などについての訳語選択規則が必要になる。本章では、これらのことについて分析し、連体修飾構造に対する機械翻訳規則を提案している。提案した規則により 714 日本語例文について人手で評価しており、被修飾名詞が一般名詞の場合 (262 文) の構造分類規則の正解率 88%、被修飾名詞が形式名詞「の」の場合 (374 文) の構造分類規則と訳語選択規則の正解率としてそれぞれ 87% と 95%、また被修飾名詞が形式名詞「こと」の場合 (78 文) の訳語選択規則の正解率として 88% を得ている。

第 6 章では、機械翻訳システム jaw/Vietnamese における日本語の「 N_1 の N_2 」という構造の翻訳処理について述べている。

日本語の名詞が名詞を修飾する場合は必ず「の」を介して「 N_1 の N_2 」という形を取るが、「の」によって結びつけられた 2 つの名詞の意味関係は、連用補語の連体化、述語名詞の連体修飾語化、所有/所属/全体・部分の関係など多様である。ベトナム語ではこの「 N_1 の N_2 」は、 N_1 と N_2 の意味関係に依って様々な前置詞を使い分けたり、いくつかの異なった語順の多様な形で表現される。本章では両言語の対応関係を分析し、意味解析の結果を単語の共起情報から記述した上で用例パターンを作成している。具体的には、名詞句の前後の表層的な情報を解析し、日本語語彙大系の意味属性を用いて、用例パターン規則を用意する手法である。「 N_1 の N_2 」を含む 279 例文について、提案した翻訳規則を jaw/Vietnamese で実験・評価して、概略 80% の正解率を得ている。

第 7 章では、研究結果のまとめおよび今後の課題について述べている。

論文審査結果の要旨

申請論文は、日本語 - ベトナム語機械翻訳のための翻訳規則について、両言語の言語的知識を駆使して分析し、特に日本語 - ベトナム語間で違いの大きい連体修飾表現に関する翻訳規則について、計算機で取り扱える規則として翻訳手法と共に提案している。また連体修飾構造に限らず、比較的単純なさまざまな例文について、当該研究室で開発中の機械翻訳システム jaw/Vietnamese によって翻訳実験を行い、翻訳結果を分析して、翻訳規則の記述法について多くの有用な考察を行っている。

機械翻訳の研究では、対象とする両言語に関する対照言語学的研究がまず最初に重要であるが、前記の研究内容はそのような意味で大変貴重である。日本語とベトナム語との間の機械翻訳に関する研究は他に未だほとんどなく、本論文はその開拓の第一歩として高く評価できる。特に名詞句構造の機械翻訳に関する本格的な研究は他の言語対でも少なく、

本研究の内容には大きな価値がある。

日本語とベトナム語の間では、ベトナム語は語形変化がない孤立語である；文要素の並び順序は基本的にSVOである；語順は文の正確な意味を表すために極めて重要である、などの特徴があるが、一方日本語は膠着語に属する；基本的語順はSOVである；語順規則は比較的ゆるやかである、など構造的に大きな違いがある。また一方では、中国の漢字文化を介して共通の面もみられる。このように同じアジアの言語として、両言語の対照研究は大変興味深い内容がある。日本とベトナムの関係は近年大きく進展しており、文化的にも経済的にも日本とベトナムの交流は大きくなってきている。日本語とベトナム語との間の機械翻訳に関する研究は、日本とベトナムの経済的文化的交流に大きな貢献をするであろうことは間違いなく、本研究はその嚆矢として高く評価できる。

日本語では埋め込み文による修飾は「連体形の動詞/形容詞+名詞」という構造的には単一の形をとるが、ベトナム語では関係節構造や名詞節構造などいくつかの構造をとるため、その構造分類規則が必要になる。さらに、先行詞や名詞化のための形式名詞などについての訳語選択規則が必要になる。本論文では、これらのことについて分析し、連体修飾構造に対する機械翻訳規則を提案している。提案した規則により714日本語例文について人手で評価しており、被修飾名詞が一般名詞の場合(262文)の構造分類規則の正解率88%、被修飾名詞が形式名詞「の」の場合(374文)の構造分類規則と訳語選択規則の正解率としてそれぞれ87%と95%、また被修飾名詞が形式名詞「こと」の場合(78文)の訳語選択規則の正解率として88%を得ている。日本語の連体埋め込み構造とベトナム語の表現との複雑な対応を整理・規則化したことは、機械翻訳に大きく貢献するのはもちろんであるが、言語学的にもその意義は大きい。

また、名詞が名詞を修飾する場合は、日本語では必ず「の」を介して「N₁のN₂」という形を取るが、「の」によって結びつけられた2つの名詞の意味関係は、連用補語の連体化、述語名詞の連体修飾語化、所有/所属/全体・部分の関係など多様である。ベトナム語ではこの「N₁のN₂」は、N₁とN₂の意味関係に依って様々な前置詞を使い分けたり、いくつかの異なった語順の多様な形で表現される。本論文では両言語の対応関係を分析し、意味解析の結果を単語の共起情報から記述した上で用例パターンを作成している。具体的には、名詞句の前後の表層的な情報を解析し、日本語語彙大系の意味属性を用いて、用例パターン規則を用意する手法である。「N₁のN₂」を含む279例文について、提案した翻訳規則をjaw/Vietnameseで実験・評価して、概略80%の正解率を得ている。「N₁のN₂」については日英翻訳の研究でもこれまでにたびたび取り上げられてきたが、日英と日越ではその対応は同じではなく、この領域の研究を多言語の場に広げた意義は大きい。

これらの結果の主要部分は、この分野の代表的な論文誌および国際学会で評価され学術論文2編として発表されている。

以上のことから、本論文は博士(工学)の学位論文として完成された内容を有し、価値あるものと確認した。

最 終 試 験 結 果 の 要 旨

最終試験（公聴会）を平成 17 年 8 月 3 日に開催し、口頭試問を行い審査した。審査委員会での審議の結果、合格と判定した。