

# 平成20年度 情報工学コース卒業研究報告要旨

渡邊 研究室	氏 名	石 黒 雄 佑
卒業研究題目	スライド中の文の構造に基づく変換規則を用いた日英翻訳手法	

現在、研究会や会議など様々な場面で、プレゼンテーションスライドを用いた口頭発表が行われている。海外での発表用のスライドは、英語で作成するのが一般的であるが、英語でスライドを作成するには、慣れていないと多大な労力を要する。そこで、日本語のスライドに機械翻訳を用いて英語のスライドに変換する方法が考えられる。しかし、既存の機械翻訳を用いると、スライド特有の英語表現にすることは難しく、スライドが見にくくなってしまいう問題が発生する。本研究では、この問題の原因が、不要な語句の省略など、スライド中の日本語表現に特有な文構造にあると考えた。そして、このような文構造に対して変換処理を行うことで、スライドに適した英語表現への翻訳を試みた。

機械翻訳に関する問題の原因を調査するために、まずは、スライド中の日本語表現の特徴について調査した。調査の結果スライド中で多用される文の構造は、単文、名詞句、連体節を持つ複文であった。本研究では、これらの中から、スライド特有の英語表現への翻訳が困難な、主語を持たない単文、連体節を持つ名詞句、複文を変換処理の対象とした。処理方法として、まず、主語を持たない単文に対しては、翻訳前に仮主語を補い、翻訳後に主語を除去して主語を持たない英文に変換する。また、連体節を持つ名詞句、複文に対しては、連体節が句となるような換言処理を行い、短く簡潔な英文となるようにする。

さらに、本研究ではスライドを英訳するプロトタイプ・システムを作成した。本システムは、PowerPointで作成されたスライドを翻訳対象とする。本システムの処理の概要は図1に示す通りである。スライド情報抽出処理では、スライド中の文と文の位置、文字の色などの情報を抽出する。次に、文構造解析処理で、抽出された文の種類を判定する。判定された文の種類に応じて、文変換処理を適用し、変換された文を機械翻訳で英訳する。最後に、翻訳文に対して、はじめに抽出した文字の位置、色を反映させてスライドに配置する。以上の手順によって、日本語のスライドから英語のスライドを作成する処理を実現した。図2に生成された英訳スライドの例を示す。

変換処理を行う前と後で翻訳結果を比較し、考察した。変換結果は、主語を持たない単文の処理では簡潔な表現となった文が多かった。しかし、連体節の換言処理では、変換処理の適用された文が少なく、その中で簡潔な表現に変換された文も少なかった。この主な原因は、日本語では連体節の換言ができない場合でも、英訳後の文に除去可能な語句があるためである。

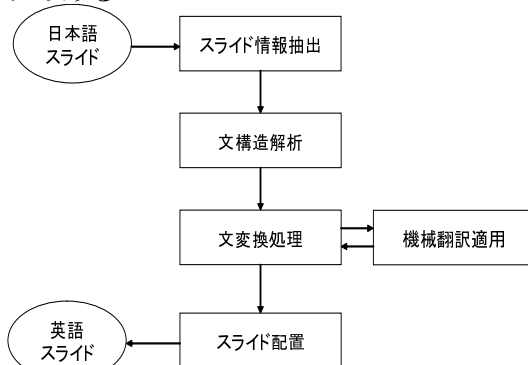


図1 処理概要

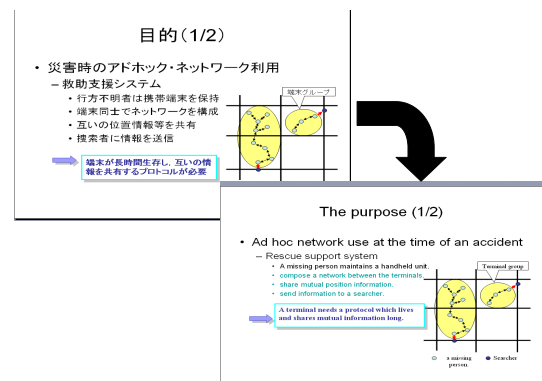


図2 スライド翻訳例