

氏名	しば やま まもる 柴 山 守
学位の種類	工 学 博 士
学位記番号	論 工 博 第 2480 号
学位授与の日付	平 成 3 年 5 月 23 日
学位授与の要件	学 位 規 則 第 5 条 第 2 項 該 当
学位論文題目	Studies on Thai Language Processing with Emphasis on Input/Output Techniques (入出力方式を重点にしたタイ語の言語処理に関する研究)
論文調査委員	(主 査) 教 授 星 野 聰 教 授 長 尾 真 教 授 池 田 克 夫

論 文 内 容 の 要 旨

本論文は、表音文字で非分割的特徴をもつタイ語の計算機端末、音節認識機構、出力システムから構成されるタイ語入出力システムの実現を目的として、タイ語入力方式におけるローマ字・タイ字変換のための改良型翻字方式、タイ語の音節構造や正書法の言語学的分析からの音節形成規則の導出と定義による音節認識モデル、高品質なタイ語の印刷制御方式について、理論的および実験的研究の成果をまとめたもので7章より成る。

第1章は緒論であり、本研究の背景、目的ならびに研究内容の概要について述べている。

第2章では、本研究の対象であるタイ語の言語処理に関して、タイ語古典テキストと辞書、タイ語正書法、ローマ字・タイ字変換のための翻字方式と学習曲線、タイ語印刷制御方式の特徴的性質を説明し、タイ語入出力システムの実現の必要性和本論文の理論的基礎について述べている。

第3章では、計算機端末におけるタイ語テキスト入力時のタイピストのキー打鍵負荷の軽減と入力速度の向上を計るタイ語入力方式を実現するために、従来のローマ字・タイ字変換のための翻字方式を改良した自動連続変換機能を備える改良型翻字方式を提案し、その変換テーブルとアルゴリズムを提示している。またこの方式は、タイ語電子タイプライタの入力方式であるタイ字直接入力方式に比較して鍵盤上の配列キー数を減少できること、タイ語古典テキスト(約100万字)の入力作業において従来の翻字方式に比較してキー打鍵数を低減できることを示している。

第4章では、直接入力方式、改良型翻字方式の2つのタイ語入力方式の比較研究を行うために新たにタイ語テキスト・エディタを開発し、タイ-タイ語辞書の見出し語(約3万語)における各字母の生起頻度回数に基づく2つの入力方式のキー打鍵数の比較、実際の入力作業におけるタイピストの記憶負担量、指・手の使用率や移動距離、学習効果の特性について述べている。また、改良型翻字方式で必要となるローマ字・タイ字変換テーブルの記憶負担量の軽減のため、新たに単純型翻字方式を提案している。直接入力方式に比較して、改良型翻字方式におけるタイピストの記憶負担量や指の移動距離が短縮されること、入力作業における学習曲線の測定では文字入力速度が速く有効的であることを確認している。

第5章ではタイ語古典テキストに対する形態素解析の手法と実験・評価およびタイ語音節認識モデルと音節認識機構の実現と評価について論じている。形態素解析ではタイ語の音節構造や正書法の言語学的分析から音節形成規則を導出、定義して分かち書きの実験を行い、新たにバックトラック機能を備えた音節最長一致法を提案し、従来の最長一致法に比較して有効性を示している。また、文の音節単位への分割のために有限オートマンによるタイ語音節認識モデルを提案し、タイ語音節認識機構を実現している。本認識機構による自動分かち書きの実験結果から、発見的手法による音節認識モデルと機構の改良により、形態素解析辞書を使用せずに約94%の分かち書き率が得られることを確認している。

第6章では、計算機の表示装置やレーザービーム式印刷装置を対象にしたタイ語出力方式について述べている。タイ語の各字母の表示装置上における合成表示方式と印刷装置上における高品質な印刷制御方式の実現方法について提示し、文字幅と母音、声調字母の印字位置制御のための文字列依存型最適化テーブルの導入とアルゴリズムについて提案している。また、各行の音節区切り位置での改行や右揃え印刷を制御するために音節形成規則を用いたアルゴリズムを提示し、タイ語古典テキストのデータベース化や計算機用例索引の出版で実用化された結果について示している。

第7章は結論であり、本論文の成果と今後の課題を要約し、タイ語の言語処理に関する研究への寄与について述べている。

論文審査の結果の要旨

表音文字で非分割的特徴をもち表記法が複雑なタイ語の計算機による言語処理を行うためには、効率的な入出力方式の実現が必要である。しかし、これらの技術や解析手法に関する研究は殆どなく、従来からタイ語テキストの計算機処理では膨大な時間と人手が必要とされてきた。本論文はタイ語の音節構造や正書法の言語学的分析に基づいて、これらの問題を扱い、タイ語古典テキストのデータベース化や計算機用例索引の出版で有効性を確認したもので、得られた主な成果は以下のとおりである。

1. ローマ字・タイ字変換のための自動連続変換機能を備えた改良型翻字方式を提案し、鍵盤上の配列キー数がタイ文字直接入力方式に比べて約2分の1に減少できること、タイ語の入力時のキー打鍵数が従来の翻字方式より約1割低減できることを示した。

2. 直接入力方式と改良型翻字方式による入力法について、タイピストの記憶負担量、手の使用率や移動距離、学習効果の特性について測定し、提案した方式ではタイピストの記憶負担量が約6割軽減できること、指の移動距離が約2割短縮できること、入力速度が約1割向上することを確認した。

3. タイ語の音節構造や正書法の言語学的分析から音節形成規則を導出して、バックトラック機能を備えた音節最長一致法を提案して、約98%の分かち書き率を得て有効であることを示した。また有限オートマンによるタイ語音節認識モデルを提案し、辞書を必要としない発見的手法によるタイ語音節認識機構を実現して、約94%の分かち書き率が得られることを確認した。

4. タイ語音節形成規則に基づく文脈依存型の印刷制御方式を提案して、この方式によるタイ語出力システムを実現し、タイ語古典テキストのデータベース化や計算機用例索引の出版で実用化した。

以上を要するに、本研究は計算機へのタイ語入力の性能向上を計り、タイ文字の分かち書きを自動的に

認識する印刷制御方式を考案し，実現したものであって，学術上また實際上寄与することが少なくない。
よって本論文は工学博士の論文として価値あるものと認める。

また平成3年3月7日，論文内容とそれに関連した事項について試問を行った結果，合格と認めた。