

# 日韓音声翻訳システムの設計

4K-6

鈴木 雅実 井ノ上 直己 野垣内 出  
KDD研究所

## 1 はじめに

音声翻訳システムの実現には多くの課題があり、これまでパイオニア的な研究成果が幾つか発表されているものの、システム全体としての明確な設計方針に基づいたものはまだ少ないと言える。本稿では、ホテル予約をタスクとする日韓音声翻訳システムの構築について、構成要素である音声認識と言語翻訳処理の特性やヒューマンファクタ等を考慮した設計方針と問題点について報告する。特に、リアルタイム処理等の基本性能と認識率の向上、および翻訳の精度と頑健性の追求に関して、現時点でのシステム構成上の工夫について述べる。

## 2 音声翻訳システムの問題点

### 2.1 要素技術の統合の観点から

音声翻訳システムを構成する要素技術である、音声認識と言語翻訳はそれぞれ性格の異なる処理内容を含んでおり、これらを統合する際には種々の問題が生じることが指摘されているが、具体的な試みについての報告はまだ少ない。

音声認識処理においては、一般に短いシラブルからなる語（日本語の助詞等）の正確な認識は難しいとされている。また翻訳処理においても、形態素レベルの短い処理単位毎に翻訳規則を適用することは、適切な結果の導出に対して処理の複雑化が問題となる。

一方、多量のテキストデータから、頻出する表現の連鎖を求め、共起（コロケーション）として分析した結果が報告されている（[1]ほか）。このような広い意味での連語は、認識や翻訳の単位として一括した取り扱いをすることが望ましい。

そこで、我々の日韓音声翻訳システムでは処理の効率を重視し、かつ認識・翻訳結果の精度を向上させることをねらった設計を行なうこととし、それを実践に移すことを目的の一つとしている。

### 2.2 関連研究

日本語／英語（ドイツ語）間の音声翻訳システムについては、ATRをはじめ国内外の幾つかの研究機関において基礎的な研究がなされている。[2]しかし、従来の研究の進め方が、どちらかという要素技術（特に音声認識）のレベルアップを主体としていたため、音声認識部と言語翻訳部の連携が必ずしも効率的とは言えない状況にあった。一方、種々の言語的な制約を一律な形式で宣言的に記述し、それらを充足しながら最適解を求める理論的な枠組みも提案されているが、具体的なタスクでの実験結果等はまだ報告されていない。

比較的近い語族に属すると思われる日韓両言語間の機械翻訳に関して、近年幾つかの研究が報告されているが、その中で金 [3] は、形態素レベルでの語彙トランスファーを基本とする、日韓機械翻訳方式を提案した。その対象は技術的な内容のテキストであるが、この手法を話言葉に適用する試みも報告されている [4]。さらに、井ノ上 [5] では、主として日韓の言語の類似性／相違点に着目して、話言葉の翻訳上の問題点について考察している。

[3][4]では、日本語文の形態素解析結果を用いて語彙交換を行なうことを前提としているが、話し言葉のように特定の表現形式が繰り返し使用されるような認識対象については、前に述べたようなコロケーションを単位とする処理が効率的である。

## 3 日韓音声翻訳システムの設計方針

以上述べたような背景の下に、現在取り組んでいる日韓音声翻訳システムの試作において、高性能なシステムの実現と今後の研究方向の明確化を図るために重視している項目について説明する。各項目は、要素技術の統合と利用者への親和性の両面を意識している。

### 3.1 コンパクトな構成とリアルタイム処理

筆者らが目指しているのは、可能な限り小規模な構成の一纏まりの音声翻訳システムを具現化することであり、それによってシステムを使用した実験の容易性や、可搬性を高めようとするものである。具体的には、

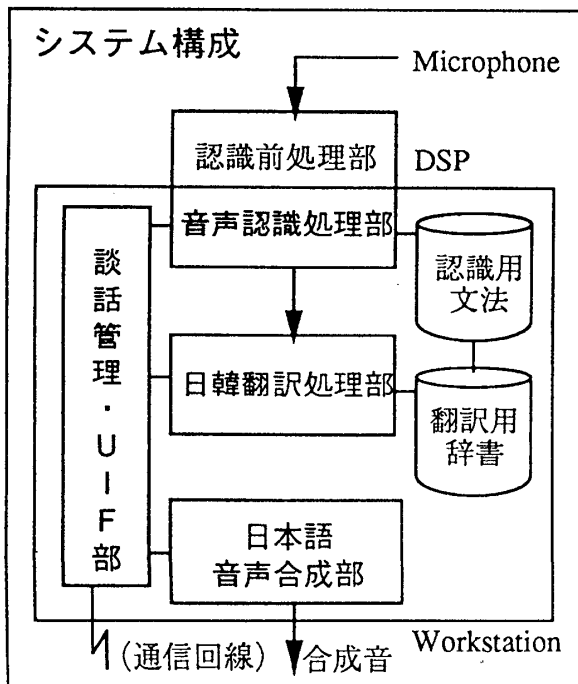
一台のワークステーションと音声認識用の数枚のDSP、および通常の周辺機器の範囲内で音声翻訳システム（日本語音声合成を含む）を構築する（下の構成図を参照）。

上のようなハードウェア構成の下に作成したシステムの基本性能として、処理の精度はもちろんであるが、処理速度の点でもリアルタイムに近い動作を目標とする。音声翻訳に限らず、対話システムの利用者にとっての待ち時間を極力短縮し、自然な対話の流れを損なわないようにすることは、システム設計上の必須の条件である。このため、次項でも述べるが、日韓両言語の相違点を十分検討した上で [5]、必要十分な処理量で期待する性能の実現を図っている。

### 3.2 認識 / 翻訳用の言語知識の連携

発声 / 認識方法は連続発声 / 連続音声認識とし、文末の長いポーズで終話を検出する。音声認識手法は離散分布型の HMM に基づき LR 文法による単語予測型の駆動方式を採用している。利用者にとっての許容範囲の中間目標として、上位 5 位以内の文理解率として 90% 以上を目指す。翻訳方式は金 [3] を参考にして、意味関係を含む接続制約を主とする逐次翻訳方式を採用している（このため十分な頑健性をもつ）が、前章で述べたように連語を重視して処理単位を設定しているため、話し言葉への対応が柔軟である。

さらに、翻訳辞書をシステム全体のマスター辞書と位置付け、音声認識文法における認識単位（語彙）も、これと連携させて記述することにより、



1) 音声認識 / 翻訳の両面での精度向上 2) 言語知識の管理の容易性に貢献している。なお、タスク内の単語数は現在約 1000~1500 を想定している。

## 4 今後の課題

[3] [5] にも述べられている通り、日韓両言語は語順等が近い関係にあることから、文書の翻訳上の問題は英語への翻訳等と比較して少ないと考えられるが、話し言葉を対象とした場合は予想以上に難しい現象に遭遇している。とりわけ、文脈に依存した多義性の解消や、文化的な背景にも関わる、言語運用上の相違点（敬語その他）等も存在する。これらの問題解決には当面は局所的な談話知識によるヒューリスティクスを試行する。また、ホテル予約のタスクでは会話内容がかなり限定されており（原則としてホテル側が会話を主導）、文のタイプの遷移や同タイプ内の語彙的な特徴に強い傾向が見られることから、これらを利用した認識候補の絞り込みを図る。

このほか、規則化することが困難な現象についての用例に基づく翻訳手法も検討に値すると思われる。典型的な例では、日本語の助詞「の」に対して韓国語にも、ほぼ同様な助詞が存在するが、対応関係は一様ではなく、韓国側では  $\rho$  となることが多い。この種の現象への接近法の一つとして類似例の検索が考えられる。

また本稿では触れなかったが、マルチモーダルインタフェースについても検討中であり、この音声翻訳システム上での種々の試行を基に、将来的に望ましい HCI の実現を図る方向で漸進的に取り組む予定である。

## 5 おわりに

以上、日韓音声翻訳システムの試作に当たって、設計上の重視項目と工夫点について述べた。両言語の相違点を考慮しつつ、音声認識と翻訳処理用の言語知識を連携させて、コンパクトなシステム構成によるリアルタイムに近い音声言語翻訳処理を実現しつつある。なお、今後の予備評価を含めた詳細は別途報告する。

## 参考文献

- [1] 北 研二, 他: “仕事量基準を用いたコーパスからの定型表現の自動抽出”, 情報処理学会論文誌, Vol.34 No.9, pp.1937-1943, 1993.
- [2] 竹沢 寿幸: “ATR 音声言語翻訳実験システム ASURA”, 情報処理学会第 46 回全国大会, 3-161, 1993.
- [3] 金 泰錫, 浦 昭二: “意味接続関係に基づく翻訳テーブルを用いた日韓機械翻訳システム”, 情報処理学会第 43 回全国大会, 3-203, 1991.
- [4] 金 政仁, 他: “日韓機械翻訳における話し言葉の翻訳処理”, 情報処理学会第 47 回全国大会, 3-179, 1993.
- [5] 井ノ上 直己, 他: “日韓音声翻訳システムの考察”, 電子情報通信学会 93 年秋季全国大会, D-65, pp.6-67, 1993.