自然言語処理 109－2
（1995．9．14）

連語パターンによる日一韓機械翻訳システムの構築とその評価

朴 哲済+# 文 敬姫# 郭 鍾根* 李 鐘赫#* 李 根培*
+早稲田大学情報科学研究センター
#浦項工科大学情報通信研究所
*浦項工科大学電子計算学科
E-mail: {cjpark, khmoon, kwak}@madonna.postech.ac.kr
{jhlee, gblee}@vision.postech.ac.kr

本稿では，連語パターンによる日一韓機械翻訳システムの実現と評価について述べる．筆者らは，日本語での連語パターンを作成し，変換規則として用いることにより多義語の問題を解決した．述部の生成は様相類意味素テーブルを利用した．日本語解析ではCYK法を改良し，辞書検索回数を少なくした．我々は,このような機械翻訳システムを鉄鋼関係の特許文書翻訳のために構築し，その性能を評価法によって分析した．その結果，翻訳成功率は約93%であるし，長文対応力，翻訳処理速度，解析生成能力において十分有効であることを確認した．

Collocation-Based MT system from Japanese to Korean and Its Evaluation

Chul-Jae Park+# Kyong-Hee Moon# Jong-Geun Kwak* Jong-Hyeok Lee#* GeunBae Lee*
+Centre for Informatics, Waseda University
#Information Research Laboratories, POSTECH
*Computer Science & Engineering, POSTECH
E-mail: {cjpark, khmoon, kwak}@madonna.postech.ac.kr
{jhlee, gblee}@vision.postech.ac.kr

In this paper, we present a collocation-based transfer model in which a source word with multiple meaning may have a set of transfer rules for word-sense disambiguation. To effectively encode the ordering information of modalities and their surface morphemes, we constructed a Modality-Feature-Ordering and Lexicalizing Table called MFOLT. The Japanese morphological analyzer is based on the CYK algorithm which access dictionary less than others. By implementing the above concepts, we constructed a Japanese-Korean machine translation system for iron and steel patent text, and obtained about 93% of accuracy transfer rate. From this experimental result, the system showed great effectiveness.

1．はじめに

日本語と韓国語は言語系統上アルタイ語に属して文法体系が似ている．特に文節単位では両言語の語順がほとんど一致していることから，文節を単位として両言語を１対１に対応して翻訳を行っても相当なレベルの翻訳結果が得られる．従って日一韓翻訳システムの場合，その文法的類似性から実用化のためには，直接翻訳方式を採用している．しかし，日一韓両言語の文節の間には語順が一致しているが，述部の形態素の間には多くの語順の違いがある．また，日本語の否定表現が韓国語では否定的意味を持つ他の一つの用言に対応することがある．また，両言語間１対１に対応出来ないときも多い．従って，直接翻訳方式のシステムでは自然な韓国語述部を生成するのは困難である．

本論文では，直接翻訳方式を利用して構築した日一韓機械翻訳システムを紹介する．特に，意味的制約を持つ連語パターン(collocation pattern)と，対応される変換規則によって日本語形態素の多義性を解決し，妥当な訳語に変換する方法を提案する．また，韓国語の述部に現れる形態素間の部分順序関係を表した述部の様相類意味素テーブルや韓国語接続情報を用いた生成方法を提案する．

日一韓機械翻訳システムが実用化されるためには，両言語間の用言や助動詞の複雑な活用の問題と助詞の使い方の違いなどの解決が問題点として残っている．そして，直接翻訳方式での不合格と判定された文は，多訳性によるものを除き，ほとんどが助詞と述部の扱いに関するものである[8]．このような問題を解決する方法として[8]では，用言テーブルを用いた．用言テーブルとは各用言が持つすべての意味素を組合わせにしたがって韓国語訳語を登録したものを言う．しかし，この方法は用言ごとにテーブルを持ち，テーブルごとに可能な意味素に当たる用言形態を登録することから，辞書が大きくなると共に，辞書作業に膨大な努力が必要である．しかし，韓国語述部において様相類(時制，相，様相，法等)を表現する方法が用言ことに違うものではない．これは補助用言，先語末語尾，語尾によって表現するので,用言ことに表現する必要はない．われわれは日本語述部の様相類に関する意味素を抽出し，これをもとに韓国語述部を生成する手法を具現した．

形態素解析は，CYK法と日本語形態素間の接続可能性の検査によって行う．この方法で可能なすべての解析結果を得たのち，ヒューリスティックスを利用して解析結果を優先順位に従って整列する．

この整列された形態素解析結果は変換部に渡され，いろいろな意味をもつ助詞，語尾，用言の訳語が決められる．変換は日本語形態素の各訳語を選択する連語パターン(collocation pattern)と形態素解析結果をマッチングし，一つの訳語を決める．この連語パターンは例外事例を持つことにより慣用句の処理を簡単にする．このときシソーラスの階層構造(thesaurus hierarchy)によるマッチングも行う．この訳語選択の結果中一番選好度が高いものが生成部に渡される．生成部ではまず，一つに決められた変換結果から述部の意味素を様相類意味素テーブル(MFOLT : Modality Feature Ordering and Lexicalizing Table)に活性化して，否定語，使役語等の処理を行う．次に，MFOLTの順番関係にしたがって韓国語の述部を生成する．語尾等の異形態処理，音韻縮訳および不規則処理はこのとき行う．述部の生成が終ったら，韓国語の接続情報を用いて助詞の異形態処理を行い，全体の韓国語文を生成する．

本論文の構成は次のようになっている．まず2章で，現在日一韓機械翻訳システムにおける問題点と，その解決方法として本システムで提案する変換モデルについて述べる．3章では，類似度スコア計算によって最適訳語を選択する手法について説明する．そして，4章で，述部と助詞の翻訳における問題点とその対策について述べる．5章では，本手法による実験結果を提示し，まとめを行なう．

2 連語パターンによる翻訳処理

2.1 多義語

日一韓翻訳システムでは助詞と用言で多く発生

する多義語の問題を解決することが翻訳の品質を決定すると考える．多義語の問題は日一韓機械翻訳において一番大きな問題点として残っている．次は多義語問題の一例である．

例)　彼が 家／父母 を失う．

例文の「失う」の場合，その意味が「잃다(ilta)」と「여의다(yeoida)」の二つに分けられる．「여의다(yeoida)」の意味として使う場合は，「を」の前に人間の意味が含まれている体言がくる．このように同じ単語が別の意味として使えるのを多義語と呼ぶ．日一韓翻訳においてこの多義語の問題は助詞と用言で多く発生する．助詞と用言の多義語の問題を解決する方法として，われわれは連語パターン(collocation pattern)に基づいた翻訳手法を提案する．助詞の多義性解決モデルでは，助詞の前接形態素が同一品詞の場合は，連語パターンによって意味解決をする．品詞が異なる場合は変換過程ではなく，形態素解析の段階で接続情報を利用して解決する．例えば，入力文字列「見るが」に対して「が」の解析は2通り可能である．一つは韓国語助詞の「이(i) / 가(ga)」，また，「을(eul) / 를(reul)」に当たる．もう一つは韓国語語尾として「지만(jiman)」に当たる．形態素解析において，「が」の前接形態素が「見る」の動詞である．まず，韓国語助詞は接続検査に失敗する．そして，形態素解析の結果から2番目の韓国語語尾「지만(jiman)」が来る．従って，このように先行する形態素の品詞が異なる場合は，形態素解析の接続検査の段階で解決した．

2.2 Collocation-based 変換モデル

変換部では，形態素解析のところから優先順位別にそなえた形態素解析結果（以下パス(path)と呼ぶ）を入力として，各形態素間の意味を決めて該当する韓国語の訳語に変える．変換部は，各形態素の意味を決定するところと，文のパス(path)を選択するところの2段階として構成した．図1にその概要を示す．形態素の意味を決める段階（多義語の問題を解決する段階）では，連語パターン(collocation pattern)を利用して日本語文の各形態素の意味を決定する．変換規則の連語パターンと日本語文をマッチングし，pSIM(pattern similarity)という類似度(similarity)点数を計算する．

連語パターンで意味制約を持つ syntagmatic term と，入力文の形態素間の意味的類似度を表すsSIM(semantic similarity)を計算する．そして，その点数を全部合わせて文全体の連語パターンに対するpSIM(pattern similarity)を計算した．形態素の意味決定段階では，一番高い点数のpSIMを持つ連語パターンが選択され，そのpSIMの値が形態素変換の点数になる．そして選択された連語パターンに例外事例がある場合は，例外事例チェックのため，もう一度マッチングを行う．例外事例のpSIMが変換規則にある連語パターンのpSIM点数より高いときは，例外事例の意味にしたがって訳語が決まる．パス(path)決定段階では，形態素の変換点数を合わせて一番高い点数を選択する．変換部では，求められた最終点数を形態素解析部で求めた点数と一緒に考慮し，一番適切なパスを選択した．

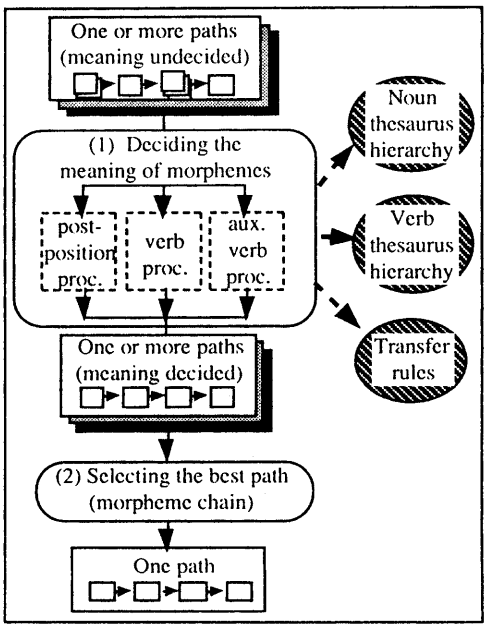

図 1: 変換処理の流れ

2.3 連語パターン

日本語の助詞と用言の意味決定において，一緒に使われった単語は重要な役割をなす．各単語の意味によって前後に使われる単語は構文的，意味的特徴を持つ．われわれはこれを規則化して，多義語の問題を解決する方法として用いた．一般化された連語パターンは文法的，意味的制約を与えるsyntagmatic termと，それらのあいだの関係を表すsyntagmatic operator，そしてbracketで表現する．連語パターンと入力文がマッチングできるとき，意味的制約をもつ名詞と動詞はシソーラス階層構造の中で意味的類似度が計算され，日本語形態素の正確なマッチングが行える．“-”が付いている否定の日本語形態素は，入力文に存在してはならないものを表す．

例えば，日本語用言「組む」の意味の中で
「편성하다(pyeonseonghada) の意味は
目的語を必要とするが，
「한패가 되다(hanpaega doeda)」の意味は目的語が不要である．文の中で目的格が現れたら「한패가 되다(hanpaega doeda)」の意味になる可能性は少ない．連語パターンを表した記号は次のようなものを使った．

◆ Syntagmatic terms
- $ ： それ自身
- N ： 意味的制約を持つ名詞
- V ： 意味，構文的制約を持つ動詞
- 日本語形態素

◆ Syntagmatic operators
- * ： Syntagmatic termの順序関係を表現，隣接に使用
- + ： Syntagmatic termの順序関係を表現，隣接しなくてもよい

◆ Brackets
- [] ： option('/'により区別)
- {} ： set，用言において必要な格を表す (“,”により区別)

変換規則は連語パターン(collocation pattern)，対訳語および，例外事例で構成した．用言と助詞は変換規則を持って意味を決める．名詞の場合は変換規則を持たず，用言と助詞の意味解決段階で名詞の意味制約を利用して意味を決める．例外事例は連語パターンでマッチングしたとき，どちらか一つの規則に対応できるが，マッチングに成功した規則の意味ではないときのためである．例外事例の利用により変換規則を単純化し，rule-based方法の欠点を補うことが出来た．変換規則は変換辞書に登録されている．図2に日本語用言「転がす」の変換規則を示す．

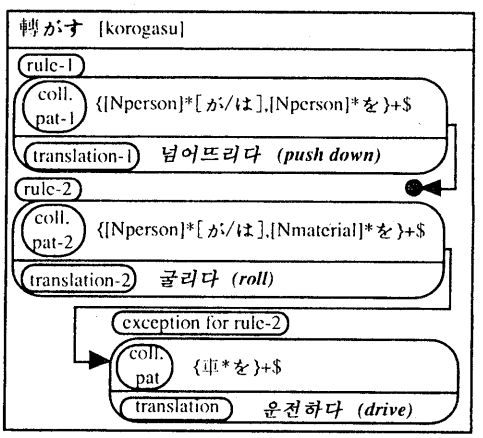

図 2: 「転がす」の変換規則

3. 最適訳語選択手法の提案

3.1 シソーラスの階層構造

シソーラスは入力文の形態素と連語パターンの意味制約間の類似度スコアを計算するときに使用する．動詞シソーラスは構文的，意味的属性によって40種に分類した[10]．名詞のシソーラスは意味によって1000種に区分した[7]．シソーラスは階層構造になっている．名詞シソーラスは4レベールの階層構造に，動詞は2〜5レベールの階層構造である．

3.2 類似度スコア計算

日本語形態素の連語パターンの中で一番高いパターン類似度（図3のpSIM(P,I)）を持つのを選択

することで形態素の意味を決める．選択されたパターン類似度が形態素変換点数（図3のmTS(m)）になる．次にパス内部にあるすべての形態素について変換点数を合わせた後，その値を正規化してパスに関する変換点数を求める（図3のsTS(path)）．

一つのパスに対して変換段階での点数と形態素解析段階の点数を合わせて，最終結果として一つのパスを選択する．このパスは日本語の各形態素に対して意味が決められたもので，韓国語生成部の入力になる．

$$sSIM(P_i,I_j)=\begin{cases}\dfrac{2\times level(MSCA(P_i,I_j))}{level(P_i)+level(I_j)}\times \text{is-a penalty} & \text{(a)}\\ 1\text{(if matched) or }0\text{(if not matched)} & \text{(b)}\\ -1\text{(if matched) or }0\text{(if not matched)} & \text{(c)}\end{cases}$$

(a) P_i is a semantic attribute
(b) P_i is a surface morpheme
(c) P_i is a negative surface morpheme

$$S_{i,j}=\begin{cases}0 & \text{if } i=0 \text{ or } j=0\\ \max\begin{bmatrix}S_{i,j-1}\\ S_{i-1,j}\\ S_{i-1,j-1}+sSIM(P_i,I_j)\end{bmatrix} & \text{otherwise}\end{cases}$$

$$pSIM(P,I)=\frac{\max(0,S_{n,m})}{\sum_i \text{Perfect-Matching-Score-of-}P_i}$$

$$mTS(m)=\max_{P\text{ in }m}(pSIM(P,I))$$

$$sTS(path)=\sum_{m\text{ in path}} mTS(m)\Big/\#\text{ of m with coll. patt.}$$

図 3: 類似度の点数計算

連語パターンPと入力文Iの類似度点数を計算するため，図3のpSIM(P,I)を定義した．pSIM(P,I)は，マッチングの類似度点数として，完全にマッチングされたときの値で割り算をすることにより正規化した．$S_{n,m}$は，動的プログラミング技法を利用することで，パターン内部のP_iが入力文の一番類似な部分とマッチングされた点数として計算する．$S_{n,m}$は，連語パターンの否定的日本語形態素により負数の値を持つ可能性がある．Max(0,$S_{n,m}$)は，マッチング点数が0以下になることを防ぐ．sSIM(P_i,I_j)はパターンのsyntagmatic term P_iと入力文の形態素I_j間の意味的類似度の計算に使用する．

日本語形態素が連語パターンに直接現れるときも類似度計算は正確なマッチングを行う．例えば，sSIM（が，が）は，1になる．sSIM（が，を）は，0になる．否定的な日本語形態素に対するマッチングは，マッチングに成功すると，1の代わりにpenaltyとして-1の値になる．意味的制約をもつ名詞と，用言の意味的類似度はシソーラス階層構造を利用して入力文の該当形態素と比較する．

シソーラス階層構造では，親ノードをたくさん公有すればもっと類似である．[1]では，意味的類似度を計算するため，MSCA(Most Specific Common Abstraction)を用いる方法を提案した．(n+1)階層のシソーラス構造で，階層の下からk番目の階層にあるノードは(k/n)の値を持つ．

例えば，図4でMSCA（会議，滞在）は，aになって意味的長さは1になる．MSCA（会議，提議）は，eになって意味的長さは0になる．しかし，[1]での意味的長さ計算方法は，example-based翻訳で事例と入力を比較するために作成されたことから，入力と事例の意味的属性は必ず一番下の階層に存在する．本論文では，事例たちの一般化(generalized)された形態である連語パターンを用いるため，比較の対象P_iがシソーラス階層構造でもっと上位レベル（more abstract）にある概念になる．例えば，図4のbと「会議」を比較する場合があるので，[1]の方法はそのまま利用できない．

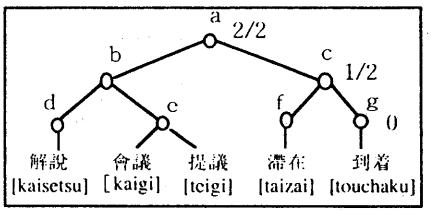

図 4: Thesaurus hierarchy

本論文では，MSCAのパターンと入力の意味的属性レベルも一緒に考慮した．また，入力文の形態素がsyntagmatic termの子孫ノードのときには，もっと類似する．これと他との区別のため“is-a

penalty"を適用した．入力形態素がsyntagmatic termの子孫ノードの場合には，"is-a penalty"が1，それ以外は0.5である（図5参照）．sSIM(P_i,I_j)は，syntagmatic term P_iと入力文の形態素I_jがシソーラス階層構造の一番下レベルのときには[1]の方法と同じになる．

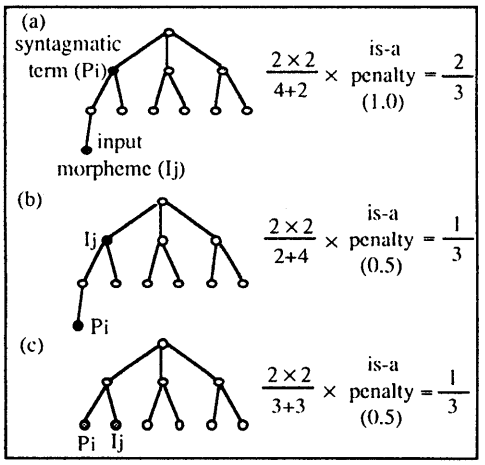

図 5: Is-a penalty

4. 述部と助詞の翻訳における問題点と対策

韓国語生成においては，韓国語の文法規則に従って補助用言を再配置しなければならない．例えば，日本語文「いきません」を韓国語に翻訳すると「 가지 않습니다(gaji anseubnida) 」になる．日本語文は<尊称ー否定>の順であるが，韓国語では<否定ー尊称>の順になる．この点から直接翻訳システムにおいて自然な韓国語の述部を生成するのが困難である．本論文では，韓国語の述部に現れる形態素間の部分順序関係を表現した述部の様相類意味素テーブル(MFOLT)と，韓国語接続情報を用いた韓国語生成手法を提案する．

4.1 辞書の内容と述部の様相類意味素テーブル

辞書には日本語表題語，その表題語ごとの左右接続情報，そして対応する韓国語，韓国語に与えられる接続情報と意味情報が登録される．述部に時制，様相等の意味を与える助動詞や語尾の場合は，韓国語訳語の代わりに意味素を登録した．意味素には，過去，謙譲，推測，受け身，素望，可能，否定等の要素を用いた．用言の場合は，韓国語用言の使役形態，否定語が存在するときの否定語および接続情報が登録される．

日本語では助動詞が用言と結合して述部の意味を表現する．一方，韓国語では，補助用言，先語末語尾，語末語尾を持って用言に意味を付ける．この補助用言たちは，19種類に区別した．これはいろいろなものが同時に使用されるときは，一定の順序を持つ．語尾は先語末語尾と語末語尾に分けた．先語末語尾は語尾の分類に属し，表層文に現れる順序が一定に決められている．先語末語尾は，補助用言，語末語尾と一緒に用言に意味を加える役割をする．その順序関係は次の通りである．

補助用言＞尊称＞時制＞謙譲＞
回想＞推測＞語末語尾

われわれは，様相類意味素テーブル(MFOLT : Modality Feature Ordering and Lexicalizing Table)と呼ぶテーブルを用いて上のような順序関係を表現した．MFOLTは，韓国語の文法的順序によって時制や様相等の意味素を表現し，各エレメントに当る韓国語形態素をその内容としてもつ．図 6にMFOLTの内容を示す．図 6で"CAU1"，"NEG"等は意味素を表す．各エレメントは韓国語述部内部の順序関係によって，左から右に配置されている．各エレメントは韓国語形態素を持っている．例えば，"NEG"エレメントの場合は,韓国語補助用言「 지 않(ji an) 」 持っている．

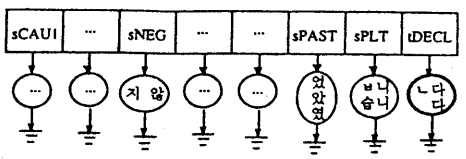

図 6: MFOLT

4.2 述部生成

韓国語生成部では，意味素を用いた述部の生成と，韓国語形態素間の接続可能性によって韓国語助詞の異形態を処理する二つの部分に分けられている.述部の生成は次の5段階の処理によって行う.

(1) 意味素活性化処理 ： まず，用言を探してその用言に意味を付ける助動詞，語尾の意味素を活性化する．一つの形態素が一つ以上の意味素を持つときはその全部をMFOLTに活性化する.

(2) 否定語処理 ： MFOLTに否定の属性が活性化されているのかを確認し，否定語処理を行う．否定の性質が活性化されている場合は，用言が持っている否定語を新しい用言語幹として変えたのち，否定の性質を消す．これにより否定語が存在するとき自然な韓国語らしい文書の翻訳になる.

(3) 使役表現，受け身表現の処理 ： 用言の日ー韓対訳辞書に韓国語訳語と，その用言の使役形態や受け身形態が一緒に登録されている．MFOLTの使役の性質を活性化するとき，辞書に登録された形態の使役，受け身の性質を活性化する．使役の性質が活性化している場合は，用言の使役語の類によって使役表現を作る.

(4) 補助用言，先語末語尾，語尾の異形態処理 ： MFOLTの順序に従って活性化されている性質を持つ韓国語形態素を結合して述部を生成する．この場合，各形態素間の接続検査によって異形態が決める.

(5) 音韻縮訳，不規則処理 ： 用言と補助用言の接続情報に表現されている不規則情報にしたがって不規則処理を行う．次に，音韻条件を検査して縮訳処理を行う.

5．実験結果及び考察

機械翻訳において評価すべき項目は，翻訳言語に依存しない汎用的な技術とそれら言語対に依存するものが多く存在する[11]．ここでは，全体の翻訳文書の質を言語対に依存するものを中心に評価する.

本稿での実験環境は，SPARCStation 10 (90 MIPS)であり，C言語で実装した．これは鉄鋼関係の特許文書を翻訳するためのシステムとして開発した.

定量的評価は，「NHKのニュース」6カ月分の中から49文，「鉄鋼関係の特許文書」約13万件の中から50文を対象とした．まず，日本語が理解できる人が辞書を参照して翻訳した結果と，システムによる翻訳結果を文書別に表1に示す．表1における「加重値」とは，総合評価のために分類別に与えた値である．全体を1と考え，形態素の加重値を0.5にし，他のを0.25にした[6]． スコアとは，各分類内での翻訳成功率に加重値を掛け算した値である．このスコアを合わせてシステムの総合評価点数にした.

次は，項目を細分化して，各文に対する具体的な言語対依存項目を評価した結果を表2に示す.

最後に翻訳文章の質を次のように4段階[5]に分類した.

A ： 訳文の意味が明確である.
B ： 訳文の意味は明確であるが，後編集が必要.
C ： 訳文の意味に少し不明確なところはあるが，推測して理解できる.
F ： 訳文の意味が不明確である.

表3に，翻訳正解率と翻訳の質との相関関係を求めた結果を示す．翻訳正解率は各クラスに属する形態素の数を適合，ほぼ適合，不適合の三つのランクに分けてその比率を求めた.

表 1: 人とシステムの翻訳結果比較

分類	全体個数	加重値	NHKニュース				特許文書			
			システム翻訳		人による翻訳		システム翻訳		人による翻訳	
			成功率(%)	スコア	成功率(%)	スコア	成功率(%)	スコア	成功率(%)	スコア
形態素	2863	0.5	93.39	46.70	98.86	49.43	94.01	47.01	98.63	49.32
単語	1213	0.25	93.36	23.34	96.11	24.03	93.70	23.43	94.60	23.65
熟語	53	0.25	91.39	22.85	99.00	24.75	91.52	22.88	98.54	24.64
総合評価				92.89	-	98.21	-	93.32	-	97.61

表 2: 言語依存評価 (NHKニュース&特許文書)

分類		全体個数	適合		ほぼ適合		不適合	
大分類	小分類		個数	比率(%)	個数	比率(%)	個数	比率(%)
品詞別	名詞	1228	1194	97.23	23	1.87	11	0.90
	動詞	365	310	84.93	43	11.78	12	3.29
	形容詞	34	28	82.35	4	11.76	2	5.88
	形容動詞	18	18	100.00	-	-	-	-
	副詞	25	22	88.00	2	8.00	1	4.00
	連体詞	15	14	93.33	-	-	1	6.67
	接続詞	4	4	100.00	-	-	-	-
	感動詞	0	0	-	-	-	-	-
	助詞	779	747	95.89	20	2.57	12	1.54
	助動詞	66	66	100.00	-	-	-	-
	接辞	61	55	90.16	6	9.84	-	-
	特殊	268	268	100.00	-	-	-	-
単語特徴別	略語	9	9	100.00	-	-	-	-
	地名・人名	71	71	100.00	-	-	-	-
	その他固有名詞	101	100	99.01	-	-	1	0.99
	カタカナ表記	134	134	100.00	-	-	-	-
	2字成語漢字	807	806	99.88	-	-	1	0.12
	数字・アルパベット	102	99	97.06	3	2.94	-	-

表 3: 翻訳正解率と翻訳の質との相関関係

翻訳の質	NHKニュース				特許文書			
	文数	適合 (%)	ほぼ適合 (%)	不適合 (%)	文数	適合 (%)	ほぼ適合 (%)	不適合 (%)
A	37	86.11	13.89	-	39	92.11	7.89	-
B	6	50.00	50.00	-	6	50.00	33.33	16.67
C	3	33.33	33.33	33.33	3	33.33	33.33	33.33
F	3	-	33.33	66.67	2	-	50.00	50.00

5. おわりに

本稿では，大規模で実用的な日一韓機械翻訳システム開発において，変換と生成を中心に述べた．

日本語文での連語パターン(collocation pattern)を変換規則として用いることにより多義語の問題を解決した．パターンと入力文とのマッチングは構文的，意味的に下位区分した名詞と，用言のシソーラス階層構造を利用した．変換規則に例外事例をおくことにより規則をもっと簡単にすることができた．日本語の慣用語が一つの韓国語単語に表現できるときがある．例えば，「手を加える」は韓国語「 손질하다(sonjilhada)」に，「腹が下がる」は「 설사하다(seolssahada)」になって3個の日本語単語が1個の韓国語単語に翻訳される．この場合のため例外事例の処理を改選し，単語を合わせて，一つの単語を作る処理が必要である．

韓国語生成は意味素を利用して直接翻訳システムの問題であった述部の生成における問題を解決した．日本語述部と韓国語述部との形態素間順序の違いは，MFOLTに表現されている順序関係を利用することで解決した．一つの日本語形態素がいろいろな意味を表現している場合は，MFOLTに多くの意味素を活性化して解決した．また，日本語用言に否定の助動詞が使われっているとき，韓国語に否定的訳語が存在する場合は，否定的訳語として生成することによりもっと自然な韓国語表現が得られた．韓国語用言の不規則処理は，韓国語形態素の接続情報に現れた不規則情報を用いて処理した．助詞と語尾の異形態処理は，韓国語接続情報間の接続可能性を調査し処理した．本論文で提案した方法で日本語を韓国語に翻訳した結果，いままで問題点として指摘されてきた文を自然な韓国語に翻訳することが出来た．

本研究で開発した日一韓翻訳システムは現在鉄鋼関係特許文書(約13万件の特許文書)翻訳システムとして稼動し，テストを行っている．われわれは現在このプロジェクトを2年の計画で大規模な辞書構築(鉄鋼関係単語：3万，一般単語：5万)と日本語コーパスを通じた各モジュール別システムの検証を行っている．

参考文献

[1] Eiichiro Sumita : Experiments and Prospects of Example-based Machine Translation, Proc. of 29th ACL, pp. 185-192, 1991.

[2] EunJa Kim, Jong-Hyeok Lee, GeunBae Lee : A Lexical Transfer Model Using Extended Collocational Patterns in COBALT J/K, Proc. of ICCPOL'94, pp. 461-466, 1994.

[3] EunJa Kim, Jong-Hyeok Lee, GeunBae Lee : A Table-Driven Modality Generation in COBAL

T J/K, Proc. PRICAI'94, pp. 759-763, 1994.
[4] Jeannette G. Neal, Elissa L. Feit, and Christine A. Montgomery : Benchmark Investigation/Identification Project, Machine Translation, Vol. 8, No. 1, pp. 77-84, 1993.
[5] Pamela W. Jordan, Bonnie J. Dorr, and John W. Benoit : A First-Pass Approach for Evaluating Machine Translation Systems, Machine Translation, Vol. 8, No. 1, pp. 49-58, 1993.
[6] Yu Shiwen : Automatic Evaluation of Output Quality for Machine Translation Systems, Machine Translation, Vol. 8, No. 1, pp. 117-126, 1993.
[7] 大野 普, 浜西 正人 : 類語新辞典, 角川書店, 東京, 1982.
[8] 金 泰錫, 浦 昭二 : 日韓機械翻訳における意味接続関係を用いた韓国語の生成方法, 情報処理学会論文誌, Vol. 33, No. 12, pp. 1578-1588, 1992.
[9] 下村 秀樹, 並木 美太郎, 中川 正樹, 高橋 延匡 : 最小コストパス探索モデルの形態素解析に基づく日本語誤り検出の一方式, 情報処理学会論文誌, Vol. 33, No. 4, pp. 457-464, 1992.
[10] 寺村 秀夫 : 日本語の構文と意味Ⅰ, 法文社, 1988.
[11] 中岩 浩巳, 森本 康嗣, 松平 正樹, 成田 真澄, 野村 浩郷 : JEIDA 機械翻訳システム評価基準（開発者編）, 情報処理学会自然言語処理研究会, NL-96-10, 1993.