

韓日機械翻訳のための音韻表現形式による用言の活用処理†

李 秀 炫†† 小 沢 慎 治††

本論文では韓日機械翻訳における用言の活用処理システムについて述べる。日本語と韓国語は構文的類似性があるが、文の中心的役割を果たす用言の活用文法がそれぞれ違い、活用構造が複雑に発達している。このような活用構造を簡潔に処理する方法の開発が、両言語間の機械翻訳のために重要な問題であると考えられる。したがって、本論文では両言語とも用言の活用を音韻表現形式によって簡単に処理する方法を提案する。音韻表現形式とは語幹と語尾活用間の母音調和の関係によって活用形式を表したもので、以下のような方法を考える。韓国語について、(1)用言の語幹と語尾間の変格活用と音韻調和の現象を分析し、(2)母音調和を伴う文法素を用いて語幹タイプと語尾活用タイプを作成する。(3)種々の変形を起こす文法素と母音調和の属性、文法規則、接続条件などの処理に必要な知識を吸収する組合せパターンを作成し、用言の活用解析に用いる。日本語について、(4)上の概念を拡張し、用言の活用処理のための語幹タイプと活用タイプを表現し、(5)活用タイプを付属語との接触条件およびポイントとして用いることによって音便処理を含む用言の活用生成処理が柔軟に、しかも簡潔に行われることを実験例を用いて示す。韓日両言語の各々音韻を基に辞書を作成し、翻訳システムを構成して韓国語用言の活用解析から日本語用言の活用生成を行っている。本システムは両言語とも従来のように膨大な活用規則をほとんど必要としない特徴を持っている。

1. はじめに

近来、機械翻訳に関する研究が盛んに行われ、2言語間の翻訳や多言語間の翻訳も活発に進められている¹⁾²⁾。一方、韓日両言語間の機械翻訳についても盛んに進められている^{3)~9)}。しかし、韓国語用言の活用は種々の変格活用と音韻調和を伴うため、処理システムが大変複雑になる。一方、日本語の5段動詞の活用は従来、行ごとに違う規則や助動詞と他の助動詞との連接による規則が膨大になる。かつ、韓日機械翻訳の場合、両言語の述語の表現差による述部間の構造変換が難しい問題になっている。したがって、両言語の活用解析と生成のための簡潔な処理システムの開発が必要とされている。日本語用言の活用文法に関しては多くの研究がある^{9)~12)}。その中で、音素表記の言語学的研究としてB. Bloch⁹⁾は構造言語学に基づいて日本語の活用とその派生形に関する研究を行ったことがあり、また、最近になって戸田¹⁰⁾が日本語用言における活用概念を用いずに、それを構成する語を述語複合体として再構成した。安西ら¹¹⁾は文献10)の考え方に基づいて述語複合体を処理するための知識表現形式を

開発し、計算機上での実現を行った。

本論文では、韓日機械翻訳システムにおける用言の活用処理のために文献9)~12)とは違い、かつ、上記の諸問題の解消を与える方法の1つとして音韻表現形式を提案する。この方法は以下のような考えに基づき、韓国語については①~③、日本語については④~⑥に分けて説明する。①韓国語の変格活用と音韻調和の現象とは実際どのようなもので、どう融合するかをまず調べることにする。②辞書には語幹の意味属性のほか、語幹と語尾(日本語の付属語に対応する)にそれぞれ音韻機能を持つ23種の語幹タイプと7つの語尾活用タイプで表現し、語幹タイプは語幹の機能と共に後向情報として用いる。語尾活用タイプは同じ機能意味を持つ語尾選択の曖昧性を解消するために機能意味を表す語尾集合をさらに7つの音韻カテゴリーに分けて、それぞれ前向情報として用いる。③変形文法の属性、音韻調和の属性および活用文法、接続条件など、解析と生成に必要な知識を組合せパターンに吸収し、語幹タイプ、語尾活用タイプを合わせて3つのパターンからなる音韻表現形式を表している。④音韻表現概念を拡張し、日本語用言の活用語尾をローマ字の音素によって分離し、子音は14種の語幹タイプに、母音は5つの活用タイプにまとめる。これらの語幹タイプと活用タイプはそれぞれ5段動詞の後向情報および前向情報として用いると共に、音素に用いる。⑤活用タイプはさらに、付属語の前向情報やポイントとして用いるこ

† Inflecting Process of Verbs by Phonological Form for Machine Translation System between Korean-Japanese by SOOHYUN LEE and SHINJI OZAWA (Department of Electrical Engineering, Faculty of Science and Technology, Keio University).

†† 慶應義塾大学理工学部電気工学科

とによって「5段動詞」,「上一段・下一段」,「サ変・カ変」,「形容詞」の4つの類と付属語(助動詞など)との接続を一貫して規則的に活用語尾の変化ルールが適用できるようにする。⑥音便を含めて6つの活用タイプの枠組を作成し,付属語の先頭ノードに位置付ける。ただし,この場合,活用タイプは語幹,付属語どちらにも属しないようにする。

本論文では,以上の音韻表現を用いて辞書および処理システムを構成し,韓国語用言の音韻調和および変格活用を組合せパターンを介して解析し,辞書と対照し,原形語幹と語尾を抽出して対応変換をした後,日本語用言の生成処理を行っている。その結果,韓日両言語の複雑な活用処理が簡単になり,特に,5段動詞は5母音の活用によって従来のように行ごとに異なる活用規則を用意する必要がなくなり,助動詞と他の助動詞との接続などが簡潔に,しかも,自然的音接が可能になることを例を用いて示す。本システムはMu-Lispで記述している。

2. 韓日両言語の活用形式

2.1 用言の基本活用形式

韓国語用言の活用形式は語幹と語尾を直接つなぐ場合と,語幹と語尾間に補助語幹が媒介される場合がある。補助語幹はテンス(現在,過去,未来,回想),尊称,母音調和を伴う文法素,使役と受動の接辞機能を持っているが,活用はしない。活用するのは語尾である。また,補助語幹は語幹と語尾間に入れても,入れなくても語尾がほとんど変わらない特徴を持っている。

一方,日本語用言の活用形式は語幹と活用語尾が結合し,さらに助動詞が結合しながらまた,活用し,他の助動詞に対する語幹として作用するため,両言語は活用構造の違いが大きいと言える。両言語の基本活用形式を(a),(b)式に示す。

$$VP_k = STM_k + [Aux_STM] + F_k \quad (a)$$

$$VP_j = STM_j + Fin_f + Aux_VB_{inf} \quad (b)$$

VP: 用言句, STM: 語幹, F_k: K語尾,

Aux_STM: 補助語幹, j, k: 日本語, 韓国語,

Aux_VB_{inf}: 付属語, Fin_f: J活用語尾

韓国語は形容詞の語尾がほとんど動詞と共通なので,両者を同じルーチンで処理できる。一方,日本語の動詞と形容詞はその活用の形にわずかの違いがあるので両言語の語尾辞書は1:2に対応付ける。また,韓国語には形容動詞の規定がない。しかし,「ha/da

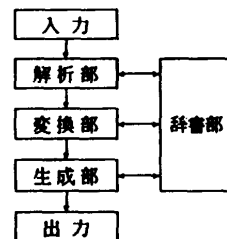


図1 システムの構成

Fig. 1 Block diagram of system.

(하/다)動詞の変形が意味的には日本語の形容動詞に対応できる。われわれは現在,両言語各々約「1100個」の単語を調べ,同一システム上で双方向に解析と生成を行うシステムを作成して用いている¹⁾。ここでは説明を省略する。

本システムの処理手順は図1のように大きく分けて解析部,変換部,生成部,辞書部の4つの部分からなる。解析部では組合せパターンを用いて韓国語の音韻調和と変格活用を解析し,辞書と対照して原形語幹と語尾および処理に必要な情報を抽出して,変換部に渡す。変換部では語順変換および述語の統語構造を調整し,日本語の音韻表現の助述部を辞書から抽出して生成部に渡す。生成部では音韻文法機能をチェックして日本語用言の生成を行い,その結果を出力する。

2.2 音素表記の言語学的活用表現

音素表記の日本語用言の活用における言語学的研究の中でB. Bloch⁹⁾,戸田¹⁰⁾と田中¹²⁾の主な部分を取り上げる。

[1] B. Blochは構造的に日本語の活用と活用語派生の分析を行った。その主な部分は①否定形容詞として単純形と派生形の2つの語幹を定義した。例えば,母音動詞 tabe/ru (食べる)の語幹 tabe-と,否定 naを含める tabena-を派生語幹と決め,②子音動詞 mat/u (待つ)の場合も同様に否定形容詞 matana-と mat-の2つの語幹を決めている。③音便の場合は基本形 mat-外にまた,異形態として語幹 ma·t̄-を立てる(t̄:音便)ことによって1つの動詞が多数の語幹を持つことになる。④接辞 ba(ば)の場合,子音動詞には eba,母音動詞には rebaを規定している。

[2] 戸田は動詞,助動詞などの活用概念を用いなくて,それを構成する述部を述語複合体として再構成している。この複合体は動詞と副動詞に分け,動詞(語幹)から始まってスペース,付属語などの独立語として区切られるまでとしている。

これらを述語複合体構成ユニットとして頭音,尾音

が子音か、母音かによって接続条件を識別し、つなぎ音を用いて連結している。例えば、語幹「起 ki」とVV型副動詞「eba」を接続する場合、母音連続禁止条件によってつなぎ子音「r」を用い、CC型副動詞「mas」と接続する場合はつなぎ母音「i」を用いる。

[3] 田中は動詞「飛ぶ、取る、問う」などの語根「TOB, TOR, TOW」から共通部分「TO」を見出し語とし、語根から欠けた「B, R, W」を(TO VHEAD (-B -R -W)0)のように辞書に書いて日本語の解析を行った。

3. 韓国語用言の活用処理

3.1 用言の音韻調和と変格活用

ここでは、音韻調和とは実際どんなもので、変格活用とどう融合するのかわかを調べ、それを伴う特定文法素を抽出する。

[1] 音韻調和

音韻調和には母音調和、媒介母音の挿入/削除、母音縮略および母音省略の4種がある。①母音調和：語幹末の子音の有無条件および語幹と補助語幹を含む語尾(以下語尾と呼ぶ)がそれぞれ陽性対陽性、陰性対陰性に母音の同質性が存在する時、語幹と語尾間で母音調和を伴う。②媒介母音の挿入/削除：語幹末の子音の有無条件と子音の連続禁止条件によって語幹と語尾間に媒介母音を挿入/削除する。③母音縮略：用言の語幹末の母音と語尾の先頭母音に文字の連音条件が成立する時、1つの母音に縮略される。④母音省略：語幹末の子音がない時、同一母音の連続禁止条件によって語尾の先頭母音を省略する。

[2] 変格活用

代表的な例文を用いて文法素の変化や音韻調和の現象を調べる。

例1)

- (a) ggaedad + _myeon → ggaedar/eumyeon :
悟れば
 개달+ _면 개달/으면
- (b) deud + assda → deur/eosssda : 聞いた
 들+았다 들/었다
- (c) ga + assda → gass/ da : 行った
 가+았다 갔/ 다
 ↑ ↑
 原形 変形

例1 a)の「ggaedad(개달) + _myeon(_면)」は「d(ㄷ)変則」動詞の語幹と語尾を示している。活用の場

表1 韓国語の規則用言の語幹タイプ構成
Table 1 Stem types of Korean regular verbs.

Type	規 則		例 文		
	属性(V1)	条件(C2)			
VA1	ㅌ, ㅍ, (ㅊ, ㅌ) a. ae. (eo, e)	0	가/다 ga/da	行く	(1)
VA2	ㅂ o	0	보/다 bo/da	見る	(2)
VB3	ㅣ, ㅡ, ㅍ, (ㅌ) i. eu. u. (oi)	0	주/다 ju/da	与える	(3)
VA4	ㅌ, ㅊ a. o	1	받/다 bad/da	受ける	(4)
VB5	(ㅌ, ㅊ)* (a. o)*	1	먹/다 meog/da	食べる	(5)

V1: 語幹末の母音, C2: 語幹末の子音の有無
() *内の母音以外のすべての母音

合、原形語幹「ggaedad(개달)」末の音素「d(ㄷ)」が「r(ㄹ)」に変わって変形語幹「ggaedar(개달)」になると共に、原形語幹と語尾「myeon(면)」間には子音連続禁止条件によって媒介母音「eu(으)」が挿入され、語尾活用は「eumyeon(으면)」に変わる。例1 b)の「d(ㄷ)変則」動詞の原形語幹と語尾が「deud(들) + assda(았다)」の場合、語幹は「deur(들)」に、語尾は「eosssda(었다)」に変わって変形語幹と語尾間では母音調和を伴い、表層上に変形が現れる。このように同じ変則語幹であっても語尾によって活用形態が異なって来る。例1 c)は原形語幹と語尾間の母音連続禁止条件によって補助語幹「ass(았)」の「a(아)」だけが語幹と同音省略によって消去され、音素「ss(ㅍ)」は語幹にシフトされて「gassda(갔다)」になる。

3.2 用言の語幹タイプと語尾活用タイプ

[1] 規則用言の語幹タイプ

ここでは音韻表現形式の一環として規則用言の語幹タイプの決め方について述べる。規則用言は語幹末の母音属性(V1)と語幹末の子音(C2)の有無条件によって語幹タイプを決める。例えば、基本形「ga/da(가/다)」の場合、語幹「ga(가)」は母音「a(ㅏ)」を持っており、語幹末の子音(C2)の有無条件は「0」である。語尾「da(다)」と語幹が母音調和を伴う。この場合、語幹タイプをVA1とし、表1に示す。他の語幹タイプも同様にしてVA2, VB3, VA4, VB5などの5種に決める。VAは陽性母音、VBは陰性母音を示す。これらの語幹タイプは母音を基準としているので動詞と形容詞が同類のタイプになる。ただしVA2は形容詞にはない。

[2] 変格活用の語幹タイプへ拡張

前に述べた結果から変格活用は語幹の部分変化や

様々な音韻調和が複合されるので、共通性を見出すことが難しくなる。そこで、規則用言の語幹は活用しても不変に保たれる点に留意し、表1の規則用言の語幹タイプに基づいて変格活用の語幹タイプへ拡張することが考えられる。すなわち、陽性と陽性、陰性と陰性がそれぞれの同質性によって母音調和が簡単に取れるので、すべての用言の語幹タイプは表1に基づいて総23種の語幹タイプに拡張することができる。例2)に示した変格活用および語幹と語尾間で種々の音韻調和に用いる「音韻カテゴリー」は活用音に制限しており、例1)の例文および表1を含めて語幹タイプの付け方について述べる。

例2)

- (a) $ggaedad(VA6) + assda \rightarrow ggaedar/assda$: 悟った
 깨단 았다 깨달/았다
- (b) $geod(VB7) + assda \rightarrow geor/eosdda$: 歩いた
 걸 았다 걸/았다
- (c) $ar(VA8) + assda \rightarrow ar/assda$: 分かった
 알 았다 알/았다
 ↑ ↑
 原形 変形

表1に基づいて例2)の変則用言は原形語幹に語幹タイプ(VA6, VB7, VA8)を与える。すなわち、例2a)の原形語幹「 $ggaedad$ (깨단)」末の母音「a」が陽性であり、語幹末に子音「d」が介していることから表1の規則(4)と一致する。さらに原形語幹「 $ggaedad$ 」末の母音「a」は語尾「 $assda$ (았다)」の母音「a」と同音に一致するため語幹と語尾間で母音調和が取れ、語幹タイプを「VA6」とする。例2b)は原形語幹「 $geod$ (걸)」の母音「eo」が陰性であるが、原形語尾「 $assda$ 」の母音が陽性「a」であるので「eo」と「a」は母音調和が取れない。したがって、原形語幹「 $geod$ (걸)」は「d(ㄷ)」変則で、変形語幹「 $geor$ (걸)」に変わると共に、語尾は陰性母音「 $eosdda$ (었다)」に変わり、表層上では母音調和が取れるので語幹タイプ「VB7」を付ける。例2)(c)が(a), (b)と違う点は語幹末の音素が活用によって変わらないということである。(c)も同じように語幹タイプ VA8 を付ける。

[3] 語尾活用タイプ

言語学的分類では主に機能意味に重点をおいているのが一般である¹⁷⁾。しかし、生成システムでは同じ意味を持つ語尾であっても語幹によって異なる形態の語尾が用いられる。

- 例3) ga | $seo \rightarrow gaseo$: 行って
 가 | 서 가서
 $bad + aseo \rightarrow badaseo$: 受けて
 받 | 아서 받아서
 $meog$ | $eoseo \rightarrow meogeoseo$: 食べて
 먹 | 어서 먹어서
 ↑
 (F6)

例えば、例3)のように異形態の語幹「 ga (가)」、「 bad (받)」、「 $meog$ (먹)」などに機能上同じ副詞形語尾(F6)が用いられるので活用の際、どの形態の語尾(F6)を選択するか曖昧さが生じ、意味だけでは不十分になり、その種類も多い。

そこで、本論文では、同じ機能意味の類の中で代表的な形態を語尾辞書に登録し、先行語幹タイプと音韻的に対応できる語尾活用タイプを決める。これは先行語幹タイプと「ペア」になるもので、表2のように7つの音韻的語尾活用タイプを立てることによって選択の曖昧さを解消することができる。「E1, E2, E3, …, E7」は韓国語の語尾活用タイプを音韻に表現したもので、これらの「E1, E2, …」は表2のようにそれぞれ機能意味「F1, F2, F3, …」などの集合からなる。すなわち、 $E1 = (F1, F4, F6, \dots)$, $E2 = (F1, F5, F6, \dots)$, $E3 = (F2, F6, F9, \dots)$, …などがそれぞれである。したがって、E1, E2, E3, …は母音調和の機能のみでなく、F1, F2, F3, …の機能意味も内包しているので選択の曖昧性を吸収することができる。表2の語尾活用タイプは以下のように定義する。

① E1タイプ：すべての韓国語用言の語幹タイプの中で、どの語幹タイプと結合しても不変の語尾をE1とする。

② E2タイプ：母音「a」から始まる語尾をE2とする。

③ E3タイプ：子音「n」から始まる語尾で、先行語幹タイプがVA8, VB9の場合だけ語幹末の子音「r」が消去され、他の語幹タイプと結合する時は、変形が生じない語尾とする。

④ E4タイプ：語尾の先頭が子音「m」, 「r」で始まり、先行語幹タイプがVA6, VB7の時、語幹末の子音「d」を「r」に置換し、媒介母音「eu」を挿入する。また、VA8, VB9と結合する時は語幹に「r」を添加する。

⑤ E5タイプ：VA6, VB7と結合する時はE4と同様に子音「d」を「r」に置換し、媒介母音「eu」を挿

表2 韓国語の音韻語尾活用タイプ
Table 2 Phonological inflecting type for Korean endings.

Type	語尾の機能 (Fi)
E1	da(F1), ja(F4), ge(F6), go(F6), gi(F8),...
E2	assda(F1), ari(F5), aseof(F6), aya(F6),...
E3	neunga(F2), neunji(F6), neun(F9),...
E4	ryeogo(F7), myeon(F7), myeonseo(F8),...
E5	rji(F6), rma0jeo0(F7), rjeonjeo0(F7),...
E6	nda(F1), ndanda(F1), ndaneun(F9),...
E7	bnida(F1), bsida(F4),...

- GY → (GY_M0(#\$)) (a)
 - RB → (RB(#(PM α))) (b)
 - DB → (DB(#(LR β γ))) (c)
 - IN1 → (IN1(#(OI c))) (d)
 - IN2 → (IN2(#(OI m))) (e)
 - HA1 → (HA1(#(LR v v')))) (f)
 - HA2 → (HA2(#(PM u))) (g)
 - CH1 → (CH1(#(#(LR a v v')))) (h)
 - CH2 → (CH2(#(#(LR w i v')))) (i)
 - KE → (KE(#/#(ES l))) (j)
- 但し, consonants : α,β,γ
vowels : c,v,v',u,n,i,w,m,l

図2 活用文法ルールと基本パターン形式

Fig. 2 Examples of conjugation rules and basic pattern forms.

- ```
(KVT '(
 (KVT00(
 (GY_M0(#$))) (a)
 (KVT01(
 ((NR(#))(HA1(#(LR eo a)))) (b)
 ((NR(#))(HA1(#(LR yeo a)))) (c)
 ...))
 (KVT10(
 ((RB(#(PM r)))(NR(#))) (d)
 ((DB(#(LR r d)))(NR(#))) (e)
 ...))
 (KVT11(
 ((EU(#(PM eu)))(HA1(#(LR eo a)))) (f)
 ((BB(#(PM b)))(IN1(#(OI u)))) (g)
 ((KE(#/#(ES b)))) (h)
 ((DB(#(LR r d)))(HA1(#(LR eo a)))) (i)
 ((REU(#(PM eu)))(HA1(#(LR ra a)))) (j)
 ((DB(#(LR r d)))(IN1(#(OI eu)))) (k)
 ...))
))
```

図3 組合せパターン解析機 (一部)

Fig. 3 Examples of combination patterns.

入するが, VA13, VB14 と結合する時は語幹末の子音「b」を削除し, 媒介母音「u」を挿入する。

⑥ E6 タイプ: 表2のように語尾が「nda」で始まり, 語幹タイプが VA4, VB5, VA6, VB7 と VA13, VB14, VA15, VB16 の場合は媒介母音「neu」を挿入する。

⑦ E7 タイプ: 語尾が子音「b」で始まり, 語幹タイプが E6 と同様に VA4, VB5, VA6, VB7 と VA13, VB14, VA15, VB16 の場合は媒介母音「eu」を挿入する。

### 3.3 組合せパターン解析機

3.1 節で述べた変格活用の分析結果を①様々な語幹末の変形音素の属性, ②母音調和の属性, ③語幹と語尾間の媒介母音の削除/挿入条件 (以下, 斜線/の前は解析, 後は生成の場合), ④文法規則, ⑤接続条件にまとめて, まず活用処理に必要な基本パターンを作成し, 次に組合せパターン解析機を構成した。以下これらについて述べる。

#### [1] 基本ルールとパターン

基本ルールをパターン化した形式を図2に示す。音韻調和の現象による変格活用を起す特定音素の数は21種がある。これらの文法素を組み合わせるための基本ルールは次のとおりである。

(a) GY\_M0: 原形語幹と語尾の直接結合による規則的活用ルール。

(b) RB: 子音 α の添加/削除の変換ルール。

(c) DB: 子音 β と γ の置換ルール。

(d) IN1: 媒介母音 c の削除/挿入の変換ルール。

(e) IN2: 語幹と語尾活用間の媒介母音 m の削除/挿入の変換ルール。

(f) HA1: 語幹と語尾活用間の母音調和ルール。

(g) HA2: 母音 u の添加/削除による母音省略ルール。

(h) CH1: 語と語間の母音縮略ルール。

(i) CH2: 先行語の母音 i と後行語の母音縮略ルール。

(j) KE: 音素シフトルール。

記号「#」は語幹に対する処理を, 「\$」は語尾に対する処理を意味し, 「#/\$」は語幹と語尾が同時に処理されることを意味する。「#/\$」は語幹と語尾の音素移動を意味する。「PM」は語幹の部分音素の「添加/削除」を表す。「LR」は解析の際, Left-handの音素をRight-handの音素に置換し, 生成の際はその逆に適用する。「OI」は母音の「削除/挿入」を意味する。

#### [2] 組合せパターン解析機

組合せパターンは基本パターンを活用形態に従って, 2つずつ組み合わせて拡張したものである。原形語幹と語尾が直接結合できるものと変格活用の形態に大別し, 30種類の Lisp Form からなる音韻表現の組合せパターンに解析機を構成して図3に示す。この解

析機は語幹や語尾の活用形態によって4つのカテゴリー (KVT00, 01, 10, 11) から成り立ち、句単位の処理ができるようにパターンを組み合わせている。呼び出し① KVT00は原形語幹と語尾が直接結合できる規則的なもの、② KVT01は語幹が原形(NR(#))で、語尾が変わるもの、③ KVT10は語幹が変形で、語尾が不変(NR(\$))のもの、④ KVT11は語幹と語尾とも変わることを示す。変格活用の動詞は活用しても変わらない部分を語幹の見出し語とする。例えば、動詞「deud/da(듣/다):聞く」の場合、「deu·d」は「d変則」の原形語幹として辞書に登録する。これに対する変格活用は「deur/eosdda(듣/었다):聞いた」のように語幹が「deu·r」になる。したがって、語幹の共通部分「deu·」だけを見出し語とし、残りの「d」か「r」かによって原形と変格活用が識別され、判別条件として用いられる。語尾の変形も同様にして条件が求められる。

#### [解析の場合]

組合せパターン解析機がどのように適用されるかを例1bの入力「deureosdda:聞いた」を用いて図4(a)の解析過程を説明する。まず、語幹「deu·r」と語尾「eosdda」に分割して原形語幹と語尾を搜索する。語幹と語尾両方の条件によって組合せパターンのKVT11に入り、図3(i)の「DB組合せパターン」の適用条件「r」と「eo」を図4(a)の入力と対照し、変形語幹「deu·r」末の音素「r」と語尾「eosdda」の先頭音「eo」に一致するかを調べる。一致すれば、入力

の変形語幹「deu·r」の音素「r」を「d」に、語尾「eosdda」の音素「eo」を「a」に置換し、原形語幹「deu·d」と語尾「assda」に変換する。この原形語幹と語尾を辞書と対照し、必要な情報を辞書から取り出し、解析結果「(deu·d VB7)(assda E2 F1)」を変換部に渡す。

本論文では、韓国語を解析して日本語を生成する問題に限定するので、ここで、韓国語生成の問題は議論外である。しかし、解析に用いる組合せパターンが解析と生成に共用できることを図4(a)(b)に示す。

#### [生成の場合]

図4(b)の入力は原形語幹 (deu·d VB7) と語尾 (assda E2 F1) であり、原形語幹末の「d」と語尾の先頭音「a」によって「DB変則」を判別し、直接「DB組合せパターン」が選択される。図4(b)は図4(a)とは逆方向へ「DB組合せパターン」のルールをRight-handからLeft-handに適用する。この際、「DB組合せパターン」の適用条件は「d」と「a」になり、この条件に従い、入力の原形語幹の音素「d」を「r」に、語尾の音素「a」を「eo」に置換することによって原形語幹と語尾が変格活用「deu·r」と「eosdda」に変換される。この際、語幹の母音「eu」と語尾の母音「eo」が母音調和を満たし、図4(b)の「deureosdda」を出力する。

## 4. 日本語用言の活用処理

### 4.1 動詞の語幹タイプと活用タイプ

日本語用言の活用処理のための活用語尾<sup>14)</sup>をローマ字音素に分離し、音韻表現形式によって生成する方法を試みる。

①「5段動詞」、「上一段・下一段」、「サ変・カ変」、「形容詞」など、4つの類が一貫して5段動詞のように規則的な語尾変化のルールが適用できるように類別の語幹タイプを決める。②5段動詞の活用母音が語幹タイプの種類と無関係になる共通性を用いて、どのように活用構造を表現し、処理すべきかということが問題であると考え、5つの活用母音を活用タイプに決める。③4つの類の語幹タイプに従う活用音素を構成要素とするパターンを作成し、日本語用言の活用生成を構造的に処理する。

#### [1] 5段動詞の語幹タイプ

5段動詞は活用語尾を音素表記によって音韻的共通性を取り出し、動詞の語幹タイプと活用タイプに分ける。すなわち、「カ行」の「KA, KI, KU, KE, KO」の子音「K」は一定である。したがって、活用しても変

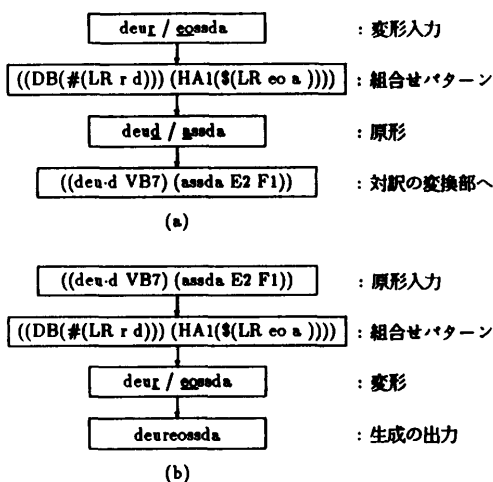


図4 韓国語の活用解析と生成例

Fig. 4 Examples of inflecting word analysis and generation for Korean.

わからない語幹本来の定義を満たす「K」などは動詞の語幹タイプに定義する。他の行も同様に語幹タイプを「W,K,S,T,N,M,B,R,G」の9つにまとめる。これらの語幹タイプは各々語幹の部分として語幹に加えると共に後向情報として用いる（行くは例外として扱う）。以下、例を見やすくするために語幹タイプと活用タイプは大文字で表す。

例) yom/u (読む) → (yo•M)

kak/u (書く) → (ka•K)

また、活用語尾の母音「A,I,U,E,O」は5段動詞のすべての語幹タイプに共通であるので音便を加えて「A,I,U,E,O,T」の6つの活用タイプに定義する。これらの活用タイプ（音便 T' 外）は助動詞を含める付属語の活用カテゴリの媒体として用い、語幹タイプとの自動的音接を可能にしている。そのためには、6つの活用タイプに従い、付属語を整理し、活用タイプは付属語と離して語幹や付属語のどちらにも含めない文法素として位置付けることが重要である。

[2] 上一段・下一段, サ変・カ変, 形容詞の語幹タイプ

「上一段・下一段」動詞の語幹タイプは「VS」に単一化し、この「VS」と活用タイプ「A,I,...」を表3に適用して活用パターン「φ,φ,ru,re,ro,φ」を取り出す。この場合、語幹タイプ「VS」は活用パターンを指定し、活用タイプはポイントとして用い、付属語との接続を行う。「サ変・カ変」動詞と「形容詞」も同様に語幹タイプを「SU, KU, AJ」とし、活用タイプはポイントとして表3に適用し、活用処理を行う。5段動詞以外の語幹タイプの例を次に示す。

例) oki/ru (起きる) → (oki•VS)

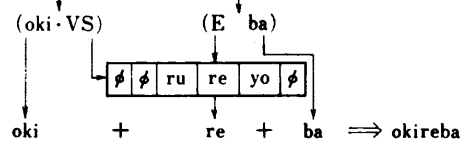
su/ru (する) → (φ•SU)

ooki/i (大きい) → (ooki•AJ)

次の例4)では、入力解析結果から対応する(oki•VS)と(E ba)を抽出し、表3の「上一段・下一段」動詞の活用パターンを介し、活用タイプ(E)をポイントとして用いることによって活用音「re」と接辞「ba」

と共に抽出する。この場合、語幹タイプは情報として用いるので空白となり、語幹「oki」、活用音「re」、接辞「ba」を合わせて「okireba」が生成される。

例4) (ireona•φ VA1) (myeon E4 F7)



4.2 助動詞の接続と態助動詞

[1] 助動詞の接続

助動詞「seru, reru, tai, nai,...」などは付属語でありながら活用するため次の例のように活用タイプと語幹タイプを前後に持つのが主動詞と違う点である。動詞型活用の助動詞は語幹タイプを「VS」とし、形容詞型活用の助動詞は語幹タイプを「AJ」とする。

例) re/ru (れる) → (A re•VS)

na/i (ない) → (A na•AJ)

ta (た) → (T'ta)

ba (ば) → (E ba)

助動詞「masu」は特殊活用し、他の助動詞との一括処理は困難であり、「masen, masita」などは活用しないため、そのまま韓国語の叙述語尾と1:1に対応させる。生成の際、語幹と助動詞および必要な情報を辞書から抽出し、次のような規則によって接続処理を行う。

[規則1] 5段動詞は語幹タイプと活用タイプをそれぞれ A<sub>0</sub>, B<sub>0</sub> としたとき、A<sub>0</sub> は語幹の後向情報を示し、B<sub>0</sub> は付属語の前向情報を示す。これらの A<sub>0</sub>, B<sub>0</sub> は接続条件として用いると共に5段動詞の場合は音素になり、音韻カテゴリー「A<sub>0</sub>, B<sub>0</sub>」は次のとおりである。

A<sub>0</sub>: 「W,K,S,T,N,M,B,R,G」

B<sub>0</sub>: 「A,I,U,E,O,T」(ただし T' は音便タイプ)

[規則2] 語幹タイプ D<sub>0</sub> は後向情報として用い、この場合に限って B<sub>0</sub> は接続ポイントとして用いる。音素としては用いない。

D<sub>0</sub>: 「VS,SU,KU,AJ」

例5)は入力文を解析した結果を用いて、辞書から抽出した日本語の動詞と助動詞との接続関係を示し、規則1, 2によるチェック過程を示している。特に、助動詞は前後に情報を持っている。

例5) (a) (yo•M) (A re•VS) (u φ)

→ yom+a+re+ru → yomareru

(b) (ka•K) (A se•VS) (I masen)

→ kak+a+se+φ+masen

表3 上一段・下一段とサ変・カ変, 形容詞の活用パターン  
Table 3 Conjugation pattern for kamiitidan-simoitidan, sahen•kahan and adjective.

| 活用    | 語幹タイプ | 活用タイプ |      |      |      |      |     |
|-------|-------|-------|------|------|------|------|-----|
|       |       | A     | I    | U    | E    | O    | T'  |
| 上一・下一 | VS    | φ     | φ    | ru   | re   | yo   | φ   |
| サ変    | SU    | si    | si   | suru | sure | siro | si  |
| カ変    | KU    | ko    | ki   | kuru | kure | koi  | ki  |
| 形容詞   | AJ    | ku    | karo | i    | kere | φ    | kat |

## → kakasemasen

以上は助動詞の音接関係を示したが、態(voice)の扱いは次に述べる。

## [2] 態助動詞

韓日両言語の動詞が使役、受動、可能の形態で用いられる場合、文の深層格が変わる<sup>15),18)</sup>。韓国語は使役、受動を作る要素の種類が多く、語幹によって接辞(i, hi, gi, ri, …), 補助動詞による統辞構造の(-ge hada), (-eo jida)が選択され、日本語の使役と受動動詞より制限性と多様性がある。可能動詞はその逆で韓国語は単一動詞である。これ以外の格変更は両言語ともほぼ同じである。ここでは両言語の使役、受動、可能に関する態の対応関係について述べる。

(1) 韓国語の使役表現は①形態的使役、②統辞的使役、③語彙的使役などに分ける。③は極めて制限されているのでここでは①、②について述べる。①語幹と形態的使役を作る接辞(i, hi, gi, ri, …)は使役と受動に共用するが、制約がある。②統辞的使役動詞は単一の補助動詞が用いられる。

例6) (a) irghida(읽히다)

(b) irgge hada(읽게 하다)

(c) yomaseru(読ませる)

例6c)の日本語の使役動詞の1つに韓国語の使役動詞は(a), (b)のように多様性がある。例6a)の「irghida(읽히다)」は接辞「hi, i, ri, gi, …」の中で「hi」を用いた使役の場合で、語幹によって極めて制限され、使用範囲も狭く、規則性がないから語幹に付けて辞書に登録する。例6b)の「irgge hada(읽게 하다)」は先行副詞形語尾「ge(게)」と後行「hada(하다)」動詞からなる統辞的使役である。すべての動詞語幹に適用される代表的な使役動詞である。しかし、(b)は2つの動詞によって形成されるため、日本語の助動詞「せる」に直接対訳はできなく、構造変換が必要になる。変換方法は、先行副詞形語尾と後行の補助動詞にスペースを挿入して離し、「ge hada」の枠組に書き換えて使役を表現し、日本語の使役助動詞「せる」に対応させる。処理例は図9に示す。

(2) 受動も使役と同様に①形態的受動、②統辞的受動、③語彙的受動などに分ける。韓国語の受動における接辞および統辞受動については文献7)で述べているが、使役、可能は扱っていない。韓国語動詞「meogda(먹다)」に対する受動は例7a, b)に示す。

例7) (a) meoghida(먹히다)

(b) meogeo jida(먹어 지다)

## (c) taberareru(食べられる)

例7a)は使役と同じ接辞「hi」が用いられて受動になるが、語幹によって制限され、使役と同様に語幹に付けて登録する。また、例7a, b)共に日本語の受動「れる」に対応される。例7b)は統辞的受動で、主動詞の語幹に副詞形語尾「eo」を付けて活用し、補助動詞「jida」によって受動になる。これはすべての動詞に用いられる代表的な受動動詞である。処理例は例11)(6)に示す。

(3) 韓国語の可能は例8a)のようにすべての動詞語幹に冠形形語尾「r, eur, reur」を付けて活用し、後に形式名詞「su」と肯定(あるいは否定)を示す単一の補助動詞「issda」からなり、語幹に付く活用形「r su issda」の3つの成分が直接「られる」に対訳されるので簡単に可能動詞の処理ができる。

例8) (a) mideur su issda(믿을 수 있다)

(b) sinjirareru(信じられる)

しかし、対応する日本語の可能を表す態の取り方は複数の形式になるため、多義的処理を必要とする。

以上、韓国語の使役、受動を作る接辞および補助動詞の種類が多い(使役10種、受動8種、可能1種)。すべての動詞に適用される代表的な使役、受動については図9の例文および翻訳結果の例11)(3), (6)に処理例を示したが、深層格変換はここでは扱っていない。

## 4.3 音便処理

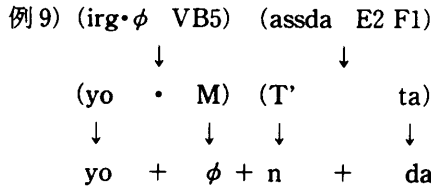
5段動詞に「te, ta, tari, …」などの付属語が接続される場合、音便を起す活用語尾は語幹タイプに従う一定のパターンがある。例えば、「kaita(書いた)」は連用形「kakita」の「ki」が音便「i」に変わったもので、語幹の基本形は(ka・K)となり、付属語「ta」は活用タイプ「T」によって(T'ta)になる。したがって、音便「kaita」の原形形式は(ka・K)(T'ta)に表現される。同様に「isoida」の場合は(iso・G)(T'ta)に表現され、「G:T'ta」の共起によって付属語「ta」を「da」に書き換える。同様にして語幹タイプが「M, B, N」の場合は撥音「n」に変わり、語幹タイプ「T, R, W」は促音「t」に変わる。

このような語幹タイプと活用タイプ(T')との対応関係によって音便と付属語の変化に対する音便ルールをまとめて図5に示す。4種の音便形「it, id, tt, nd」を用いて生成を行う。



(K : T') → it (a)  
 (G : T') → id (b)  
 (K, T, R, W : T') → tt (c)  
 (B, M, N : T') → nd (d)

図5 音便ルール  
 Fig. 5 Forms of onnbinn-rule.



例9)は基本形「yomu(読む)」が「yonda(読んだ)」に活用する場合で、入力解析結果から対応する日本語の(yo·M)と(T'ta)を抽出し、図5(d)の音便ルール(M:T')の共起によって「T'」を撥音「n」に書き換えると共に付属語「ta」を「da」に変換する。次に、原形語幹(yo·M)の語幹タイプ「M」を空白処理して、語幹「yo」、撥音「n」、付属語「da」を合わせることで「yonda」が生成される。

5. 処理システム

5.1 辞書

用言辞書の一般形式は図6のように成り立っている。図6(a)は語幹(stem)形式を表し、語幹タイプVAi/VBjは機能判別や語尾との音接情報として用い、Vは動詞、AiとBjはそれぞれ陽性と陰性の情報を示す。ViWj(1 ≤ i ≤ n, 1 ≤ j ≤ m)は動詞の格パターン情報を表している。このVi, Wjによって格パターンを抽出し、組合せDPN(dynamic Pattern Net)を構成して構文・意味解析に用いた<sup>16)</sup>。Viは必須格、Wjは任意格の選択に、また、Viは多義語の選択に有用である。しかし、本論文では用言の活用およびその派生語の処理に問題を限定するので、多義語処理はしていない。両言語の語幹の辞書はポインタPtによってつなぎ、それぞれの「キー」を持ち、双方向への検索ができるようにしている。両言語の語幹辞書の実例を図7(a), (b)に示す。

語尾辞書の形式は図6(b)のように韓国語の叙述および機能語尾ENDkに対応する日本語の動詞と形容詞の語尾を分けて1つのフレーム内に対応付ける。Eiは韓国語語尾の語尾活用タイプを、inf1, inf2は日本語の付属語の活用タイプを示す。Ri, Ti, Fiはそれぞれ韓国語の尊敬、テンスなどの語尾機能を表し、図8は実際の両言語語尾の対応辞書を示している。特に、

(stem VAi/VBj (ViWj) Pt)  
 (a)  
 (ENDk Ei ((inf1 Fi-V) (inf2 Fj-A))  
 Ri Ti Fi)  
 (b)

図6 辞書の形式  
 Fig. 6 Format of dictionary.

(KVTBL '(  
 (meog·φ VB5 (V30W9 1))  
 (irg·φ VB5 (V4W9 2))  
 (no·r VA8 (V1W9 3))  
 (as·eu VB20 ((V2Q4 4)  
 (V3S1 5)  
 (V5W9 6)  
 ... )  
 (mandeu·r VB9 (V27R8 7))  
 (nu·b VA14 (V15W8 8))  
 (ju·φ VB3 (V27Y2 9))  
 (bo·φ VA2 (V4W8 10))  
 (ggaeda·d VA6 (V5Q7 11))  
 (geo·d VB7 (V37Y5 12))  
 ... ))  
 (a)

(JVTBL '(  
 (tabe· VS (V30W9 1))  
 (yo· M (V4W9 2))  
 (aso· B (V1W9 3))  
 (ka· K (V2Q4 4))  
 (kabu· R (V3S1 5))  
 (tuka· W (V5W9 6))  
 (tuka· R (V27R8 7))  
 (yokotawa· R (V15W8 8))  
 (kure· VS (V27Y2 9))  
 (mi· VS (V4W8 10))  
 (sato· R (V5Q7 11))  
 (aru· K (V37Y5 12))  
 ... ))  
 (b)

図7 韓国語と日本語の動詞語幹の辞書例  
 Fig. 7 Examples of dictionary for Korean and Japanese stems.

日本語の動詞と形容詞の語尾は図8, 表3のように「ます, です」外は活用の形にわずかの違いが見える。

5.2 翻訳処理

韓国語の用言を入力として日本語の用言を翻訳する過程例10)を図9に示す。韓国語の助述表現は主動詞と補助動詞が連続される。対応する日本語の助述部は助動詞が用いられ、構造的相異が生じ、構造変換が必要になる。

例10)の入力、韓国語述部「norge hayeosseubnida (놀게 하였습니다)」を「asobasemasita(遊ばせまし

```
(ENDTAB '(
 (asseubnida E2 ((I masita) (U desita)) R2 T2 F1)
 (eubnida E1 ((I masu) (U desu)) R2 T1 F1)
 (bnida E7 ((I masu) (U desu)) R2 T1 F1)
 (assda E2 ((T' ta) (T' ta)) R1 T2 F1)
 (go E1 ((T' te) (A te)) R1 T1 F6)
 (myeonseo E4 ((I nagara) (U nagara)) R1 T1 F8)
 (ryeogo E4 ((O uto) (U to)) R1 T1 F7)
 (gido E1 ((T' tari) (T' tari)) R1 T1 F8)
 (myeon E4 ((E ba) (E ba)) R1 T1 F7)
 (a E2 ((T' te) (A te)) R1 T1 F6)
 (ge ha E1 (A se VS))
 ...))
```

図8 韓国語と日本語の語尾の辞書  
Fig. 8 Examples of dictionary for Korean and Japanese endings.

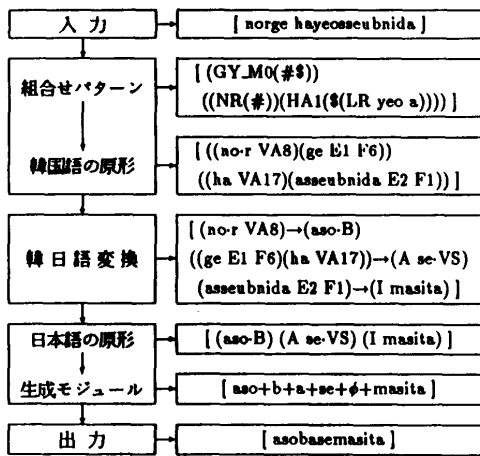


図9 韓国語と日本語の翻訳過程例10  
Fig. 9 Example of process for translation between Korean and Japanese.

た]に翻訳する例を用いて説明する。入力文の主動詞「norge」と補助動詞「hayeosseubnida」をそれぞれ図3(a)と(c)2つの組合せパターンを介して解析して辞書と参照し、主動詞の原形語幹と語尾((no-r VA8)(ge E1 F6))と補助動詞の原形語幹と語尾((ha-φ VA17)(asseubnida E2 F1))を抽出して変換部に渡す。変換部では、図9のように①韓国語の主動詞の語幹(no-r VA8)と日本語の主動詞の語幹(aso-B)を対応させる。②残りの副詞形語尾(ge E1 F6)と後続の補助動詞の語幹(ha φ VA17)はそれぞれ直訳できない。そこで、先行主動詞の副詞形語尾(ge E1 F6)と補助動詞の語幹(ha φ VA17)を組み、構造変換して使役を表し、日本語の助動詞(A se-VS)に対応させる。③次に、過去、尊敬、叙述を表す補助動詞の語尾(asseubnida E2 F1)を日本語の過去、丁寧を表す(I

masita)に対応させる。④辞書から抽出した日本語の(aso-B), (A se-VS), (I masita)を生成部に渡し、語幹タイプと活用タイプとの関係(B:A, VS:I)を規則1, 2によってチェックし、各項目を合わせて訳語[asobasemasita]を出力する。

[翻訳結果例]

例 11)

(1) (a) georeosdda(걸었다) <歩いた>  
→ ((geo-d(걸) VB7) (assda(았다) E2 F1))

↓

(b) ((aru-K) (T' ta)) → aruita

(2) (a) nuumyeon(누우면) <横たわれば>  
→ ((nu-b(눕) VA14) (myeon(면) E4 F7))

↓

(b) ((yokotawa-R) (E ba))

→ yokotawareba

(3) (a) norge hago sipji anhasdda  
(놀게 하고 싶지 않았다)  
<遊ばせたくなかった>  
→ ①((no-r(놀) VA8) (ge(게) E1 F6)) +  
②((ha-φ(하) VA17) (go(고) E1 F6)) +  
③((sip-φ(싶) AJ5) (ji(지) E1 F6)) +  
④((anh-φ(안) AJ5) (assda(았다) E2 F1))

↓

(b) ((aso-B) (A se-VS) (I ta-AJ) (A na-AJ) (T' ta))

→ asob+a+se+phi+ta+ku+na+kat+ta

→ asobasetakunakatta

(4) (a) irgeo boa jugesseubnida(읽어 보아 주겠습니다)  
<読んでみてください>

→ ((irg-φ(읽) VB5) (a(아) E2 F6)) +  
((bo-φ(보) VA2) (a(아) E2 F6)) +  
((ju-φ(주) VB3) (gesseubnida(겠습니다) E1 F1))

↓

(b) ((yo-M) (T' te)) + ((mi-VS) (T' te)) + ((kure-VS) (I masu))

- yo+φ+n+de+mi+φ+te+kure+  
φ+masu
- yondemitekuremasu
- (5)(a) irggo sipji anheubnida(읽고 싶지  
않습니다)  
〈読みたくないです〉
- ((irg•φ(읽)VB5)(go(고)E1 F6))+  
((sip•φ(싶)AJ5)(ji(지)E1 F6))+  
((anh•φ(않)AJ4)(eubnida(읍니다)  
E1 F1))
- ↓
- (b) ((yo•M)(I ta•AJ)(A na•AJ)(U  
desu))
- yom+i+ta+ku+na+i+desu
- yomitakunaidesu
- (6)(a) meogeo jida(먹어 지다)〈食べられる〉
- ①((meog•φ(먹)VB5)(a(아)E2  
F6))+  
②((ji•φ(지)VB3)(da(다)E1 F1))
- ↓
- (b) ((tabe•VS)(A rare•VS)(U φ))
- tabe+φ+rare+ru+φ
- taberareru

翻訳結果は例11)(a)韓国語活用の原形抽出と(b)対応する日本語の活用生成に分けて主な部分だけ簡単に説明する。例11)(1)では、韓国語の動詞「georeosdda〈歩いた〉」を図3(i)の組合せパターンに適用し、変格活用と音韻調和の現象を解析し、辞書から原形語幹(geo•d VB7)と語尾(assda E2 F1)を抽出する。対応する日本語の原形語幹(aru•K)〈歩く〉と音便の活用形式(T' ta)を辞書から抽出し、(aru•K)(T' ta)に並べ、語幹タイプ「K」と活用タイプ「T'」の接続条件(K:T')によって「T'」を「i」に置き換え、[aruita](iの前でKが消失)を出力する。例11)(2a)では、「b変則」と媒介母音「u」の削除が生じるが、図3(g)の組合せパターンを介して解析し、辞書から原形語幹と語尾を取り出す。対応する日本語の5段動詞の語幹と接続助詞を抽出して(2b)のように並べ、語幹タイプ「R」と活用タイプ「E」の接続条件を規則1によってチェックするだけで合成出力される。例11)(3a)の韓国語入力「norge hago sipji anhasdda〈遊ばせたくなかった〉」の述部は4つの用言句から成り立っている。この述部は主

動詞に3つの補助用言が接続されている。この例は用言の句単位に図3(a)の組合せパターンを反復適用して解析し、次々に辞書と対照して原形語幹と語尾①, ②, ③, ④を抽出して変換部に渡す。変換部ではまず, ①の主動詞語幹(no•r VA8)に対応する日本語の主動詞の語幹(aso•B)を抽出する。次に、韓国語①の副詞形語尾(ge E1 F6)と②の補助動詞の語幹(ha•φ VA17)とを組んで、使役に表現して構造変換し、日本語の使役助動詞(A se•VS)に対応させる。同様に、次々と反復して先行語尾と後行補助用言の語幹とを組んで、日本語の助動詞に対応させることによって否定を含む3つの助動詞が(3b)のように正確に抽出される。さらに、(I ta•AJ)と(A na•AJ)は表3を介し、形容詞パターンの活用語尾「ku」, 「kat」を処理した後、各タイプの接続条件(B:A, VS:I, AJ:A, AJ:T')と音韻的關係を規則1, 2によってチェックすることだけで、使役を含む訳語が合成される。例11)(4a)の韓国語の入力は叙述語を含む3つの動詞が連続しており、対応する日本語は助動詞を含めて4つの動詞が接続されている。入力の主動詞は図3(b)の組合せパターンを介して解析した結果を日本語の原形語幹に対応させ、規則1, 2によって各タイプの接続条件をチェックすることによって合成し、出力する。例11)(5)入力の述部「irggo sipji anheubnida〈読みたくないです〉」は主動詞と補助動詞が例11)(3)の一部と類似性があるので解析に関する説明は省略する。ところで、入力の述部を翻訳した日本語は「読みたくないです」と「読みたくありません」の2つに訳される可能性がある。どれを選択するかは話者、聴者によって異なるもので、「masu, desu」のニュアンスによる選択の規則は本システムには設けていない。例11)(6a)は受動の場合である。図3(a), (b)の組合せパターンを介して、入力の主動詞と補助動詞を解析してそれぞれ辞書と対照し、原形語幹と語尾を抽出して①, ②に示す。まず, ①の主動詞の語幹(meog•φ VB5)に対応する日本語の主動詞の語幹(tabe•VS)を抽出する。次に, ①の副詞形語尾(a E2 F1)と②の語幹(ji VB3)を組んで、受動を表現し、対応する日本語の助動詞(A rare•VS)を取り出す。続いて②の語尾(da E1 F1)を(U φ)に対応させ、(6b)日本語の受動形式を取る。さらに、助動詞の語幹タイプ「VS」を表3に対応させて「ru」を抽出すると共に各タイプの接続条件をチェックして訳語を出力する。以上、韓日両言語

の助述部は非常に微妙な変化をするので構造変換が困難であるが、本システムでは柔軟に処理できた。

## 6. おわりに

本論文では音韻表現形式を開発し、韓国語用言の活用を解析して、日本語用言の活用を生成するシステムについて説明した。本システムは韓国語用言の変格活用における種々の変化をする素性、語幹と語尾活用間で様々な母音調和を伴う素性、文法規則、接続条件など、活用処理に必要な知識を30種の組合せパターンに圧縮することができた。すなわち、語幹タイプ、語尾活用タイプ、組合せパターンの3つのパターンを連動することによって、同一システム上で、解析と生成を双方向に処理することができるのみでなく、種々の活用規則や接続情報をほとんど組合せパターンに吸収することによって簡潔に処理することができた。

これに対応する日本語の用言は活用語尾を音素に分離し、5段動詞の語幹タイプの種類に無関係になる活用母音の共通性を取り出し、「5段動詞」、「上一段・下一段」、「サ変・カ変」、「形容詞」を一貫して規則的に活用語尾の変化ルールが適用できるように14種の語幹タイプと6つの活用タイプとの枠組の音韻表現形式を立てた。その結果、韓国語はすべての活用規則と条件が組合せパターンに吸収され、処理が簡単に、両言語の述部間の構造変換が容易になった。日本語の5段動詞の行ごとに違う活用規則や接続条件を用意する必要がなくなり、かつ助動詞と他の助動詞との接続もタイプをチェックすることだけで簡単に生成できることを例を用いて示した。韓国語用言の活用解析と生成に用いる文法規則、すなわち、組合せパターンの容量は約1.5 Kbytesであり、プログラムの大きさは約63ステップに圧縮した。一方、日本語用言の活用生成のためのプログラムの大きさは35ステップに圧縮され、プログラムの大きさが非常に小容量である。本システムは韓国語の場合、1つの組合せパターン枠組で4-6個の文法素が同時に解析される。5つの用言句からなる述部を解析する場合は、24-30個の文法素がチェックされることになる。したがって、一般の文法規則を並べた場合に比べて、10倍以上に処理能力が向上することになる。日本語の場合は不規則が2つしかないことによって韓国語の活用より処理能力が上がる。本システムの有効性を確かめるために高校国語の参考書な

ど<sup>17)-19)</sup>から述語の構成要素である「助動詞、付属語」対応語が1つから5つの用言からなる代表的な述部を切り取って実験を行った。解析用の用言辞書にはKIUV-II<sup>20)</sup>を採用した。平均翻訳時間は1.5 sec以内に収まった(80386 CPUマシン)。韓国語の用言の見出し語は約6千語であり、実験した代表的見出し語は約2千語である。

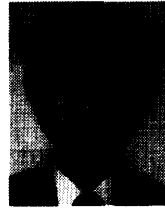
## 参考文献

- 1) 野村浩郷, 田中穂積(編): 機械翻訳, bit 別冊, pp. 143-215, 共立出版(1988.9).
- 2) 山本利文, 辻井潤一, 長尾 真: Mu-プロジェクトの英日機械翻訳システムにおける多品詞語の解消, 情報処理学会自然言語処理研究会, 53-7 (1986.1).
- 3) Lee, J. K. and Han, S. K.: Syntactic Analysis of Korean Sentence for Machine Translation System, *KITE*, Vol. 18, No. 5, pp. 15-21 (1981).
- 4) 韓国科学技術院システムセンター: 韓日・日韓機械翻訳システム開発研究(II), 研究報告書(1986.10).
- 5) Lee, S. H., Ozawa, S. and Lee, J. K.: A Study on Extraction of Aspect and Modality Information in Korean, *KJCS*, Vol. 1, No. 2, pp. 255-277 (1989).
- 6) Lee, S. H., Ozawa, S. and Lee, J. K.: Morphological Analysis of the Korean Language, *KITE*, Vol. 26, No. 4, pp. 526-534 (1989).
- 7) Choi, K. S.ほか: 日本語翻訳システムの開発環境下での韓国語翻訳システムの開発の為の一考察, 情報処理学会自然言語処理研究会, 68-4 (1988).
- 8) 李 義東, 中嶋正之, 安居院猛: 語と語の関係を用いた意味解析による日韓単文機械翻訳システム, 電子情報通信学会論文誌, Vol. J72-D-II, No. 10, pp. 1689-1695 (1989).
- 9) Bloch, B.: ブロック日本語論考, 林栄一(監訳), pp. 1-24, 研究社(1985).
- 10) 戸田正直: 日本語口語文法の再構成の試み, 昭和61年度科学研究費補助金(特別推進研究(2))研究成果報告書, 6章(1987).
- 11) 神岡太郎, 土屋孝文, 安西祐一郎: 述語複合体の生成とその表現, 情報処理学会論文誌, Vol. 30, No. 4, pp. 457-466 (1989).
- 12) 田中穂積: 計算機による自然言語の意味処理に関する研究, 電子技術総合研究所研究報告第797号, pp. 58-63 (1979).
- 13) Lee, S. H., 小沢慎治: 韓日機械翻訳システムにおける日本語の形容動詞の処理, *SITA '89*, pp. 839-844 (1989).
- 14) 西尾 実, 岩淵悦太郎, 水谷静夫(編): 岩波国語辞典第四版, p. 1217, 岩波書店(1988).
- 15) 柴谷方良: 日本語の分析, pp. 120-155, 大修館書

- 店(1988).
- 16) Lee, S. H., Han, K. R. and Lee, J. K.: A Bidirectional MT System between Korean and Japanese Based on a Pattern Net, *SITA '87*, pp. 405-410 (1987).
- 17) 安田吉実, 孫 洛範: ESSENCE 韓日辞典, pp. 2109-2131, 民衆書林(1987).
- 18) Seo, W. I.: 使動法記述試案, pp. 187-207, 文研 1 (1974).
- 19) 金 均一: 日本語文法研究, pp. 60-191, 学文社 (1983).
- 20) 韓国仁荷大学: 計算機処理のための韓国語基本用言 KIUV-II (1983).

(平成4年3月26日受付)

(平成4年9月10日採録)



**李 秀炫 (正会員)**

昭和34年生。昭和58年韓国仁荷大学校工科大学電子工学科卒業。昭和60年仁荷大学校大学院修士課程修了。現在、慶應義塾大学大学院博士課程在学中。自然言語処理、機械翻訳などに興味がある。大韓電子工学会会員。



**小沢 慎治 (正会員)**

昭和42年慶應義塾大学工学部電気卒業。昭和47年同大学院博士課程修了。昭和45年慶應義塾大学工学部電気助手。同大専任講師・助教授を経て現在同大電気工学科教授。その間、昭和59年米メリーランド大学訪問助教授。デジタル通信および画像・音声のデジタル信号処理の研究に従事。著書「デジタル信号処理」, 「基礎通信工学」(実教出版)。工学博士。電気学会, 信学会, IEEE各会員。