

TR-IT-0098

韓国語の形態素処理における予備検討:  
ハングル変換と語彙情報体系

Preliminary Studies for Morphological Processing of Korean:  
A HANGUL conversion and a lexical information system

金 徳奉

Deok-Bong Kim

1995 年 3 月

概要

本報告書は韓国語の形態素処理を行なうにあたり予備的に必要な二つの重要なツールについて述べる。一つは韓国語の形態素分離と単語生成のためのハングル変換プログラムである。もう一つは質の高い韓国語形態素辞書の構築のための韓国語形態素の語彙情報に関する体系である。

エイ・ティ・アール音声翻訳通信研究所

ATR Interpreting Telecommunications Research Laboratories

©(株) エイ・ティ・アール音声翻訳通信研究所 1995

©1995 by ATR Interpreting Telecommunications Research Laboratories

# Contents

1	はじめに	1
2	ハングル変換方法	3
2.1	韓国語アルファベット (ハングル; HANGUL)	3
2.2	ハングルコードの変換方法	5
2.2.1	音節文字からローマ字ハングル子母文字への変換	6
2.2.2	ローマ字ハングル子母文字から音節文字への変換	7
3	韓国語形態素の語彙情報に対する体系	8
3.1	韓国語形態素辞書の構成	9
3.1.1	表層形態 (Surface Form; S-Form)	9
3.1.2	字素形態 (Grapheme Form; G-Form)	9
3.1.3	音素形態 (Phoneme Form; P-Form)	9
3.1.4	正規形態 (Regular Form; R-Form)	10
3.1.5	語彙情報 (Lexical Information; L-Info)	10
3.2	範疇情報 (Morphological Category Information)	10
3.2.1	固有名詞 (Proper Noun; NPROP)	10
3.2.2	動作名詞 (Active Noun; NACTV)	10
3.2.3	形容名詞 (Adjective Noun; NADJV)	11
3.2.4	普通名詞 (Common Noun; NCOMM)	11
3.2.5	代名詞 (Pronoun; NPRON)	11
3.2.6	数詞 (Numeral; NNUMB)	11
3.2.7	動詞語幹 (Verb Stem; VVERB)	11
3.2.8	形容詞語幹 (Adjective Stem; VADJV)	12
3.2.9	補助用言 (Auxiliary Verbal Stem; VAUX1)	12
3.2.10	先語末語尾 (Verbal Prefinal-Ending; VAUX2PEND)	12
3.2.11	転成連結語尾 (Verbal Transformative-Ending; VAUX2TEND)	12
3.2.12	文末語尾 (Verbal Sentential-Ending; VAUX2SEND)	13
3.2.13	冠形詞 (Adnominal Modifier; MADNO)	13
3.2.14	接続詞 (Conjunction; MCONJ)	13
3.2.15	副詞 (Adverb; MADVB)	14
3.2.16	感嘆詞 (Interjection; INTER)	14
3.2.17	主格助詞 (Agentive Particle; PAGNT)	14
3.2.18	冠形格助詞 (Adnominal Particle; PADNO)	14
3.2.19	目的格助詞 (Objective Particle; POBJT)	14
3.2.20	叙述格助詞 (Verbal Particle; PVERB)	15
3.2.21	接続助詞 (Conjunctive Particle; PCONJ)	15
3.2.22	副詞格助詞 (Adverbial Particle; PADVB)	15
3.2.23	主題補助詞 (Subjective Particle; PSUBJ)	15
3.2.24	一般補助詞 (Auxiliary Particle; PAUXI)	15
3.2.25	接頭辞 (Prefix; XPREF)	16

3.2.26	名詞形接尾辞 (Nominal Suffix; XSUFFNOUN)	16
3.2.27	動詞派生接尾辞 (Verb-deriving Suffix; XSUFFVERB)	16
3.2.28	形容詞派生接尾辞 (Adjective-deriving Suffix; XSUFFADJV)	16
3.2.29	副詞派生接尾辞 (Adverb-deriving Suffix; XSUFFADV)	16
3.2.30	記号 (Symbol; SYMBL)	17
3.3	可算情報 (Number Information)	17
3.4	変則情報 (Morphophonological Information)	18
3.5	活用情報 (Conjugation Information)	22
4	おわりに	24

## 1 はじめに

用例に基づく変換主導型翻訳モデルにおいて、形態素解析と形態素生成は翻訳対象の言語の解析と生成の大きな部分を占める。特に韓国語と日本語間の翻訳の場合、形態素処理は極めて重要である。多くの文献から報告されているように韓国語と日本語は語順だけではなく文法的な性質が非常に類似しており、1対1に対応できる単語（形態素）も多いので、複雑な構文処理をしなくても両言語間の翻訳を行なうことができる。しかし、両言語間の翻訳で基本的に要求されることは、両言語が膠着語であるため、効率的に入力文の構成形態素を分割、解析し、形態素を結合することによって出力文を生成しなくてはならないことである。

本報告書は、韓国語の形態素処理を行なうのに予備的に必要である二つの重要なツールについて述べる。一つは韓国語対話の表層表現である音節文字から字素単位の文字に変換したり、逆に字素単位の文字から音節文字を生成するハングル変換プログラムである。このプログラムは次のような韓国語の形態音韻的な特徴のために韓国語形態素処理過程で要求される。

- 韓国語のアルファベット（ハングル）は基本的に24個の子音母音文字（以下、子母文字と略称）で構成されるが、韓国語の表記は子母文字の列ではなく、子母文字の結合（子音+母音、子音+母音+子音）形態からなる数千種類（可能な音節文字の数は11,172であるがコンピュータで処理できるのは韓国語表記によく使用する2,350余の音節文字だけ提供）の音節文字に表現される。しかし、これは音節文字をアルファベットにする日本語の表記と大きく異なる。韓国語の音節文字は韓国語の読み方を容易にするために子母文字を音節単位に集めて書くのに対し、日本語では音節文字自体がアルファベットである。

ㅎ	ㅏ	ㄹ	ㅓ	ㅕ	ㅗ	ㅛ	ㅜ	ㅝ	⇒	한국어	(韓国語)
ㅇ	ㅣ	ㅓ	ㅗ	ㅛ	ㅜ	ㅝ	ㅞ	ㅟ	⇒	일본어	(日本語)
か	ん	こ	く	ど					⇒	かんこくど	(韓国語)
に	ほん	ど							⇒	にほんど	(日本語)

- 韓国語形態素には‘-ㄹ’、‘-ㅓ’、‘-ㅗ’のように子母文字一つで構成される形態素があるだけでなく‘-ㅗ니다’、‘-ㅗ니까’、‘-ㅓ 것이’等のような子母文字から始まる形態素が多い。このような形態素の一番初めの子母文字は、先行する形態素の終音節と結合したり媒介母音‘ㅇ’と結合して新しい音節文字を作る。

가 + ㄴ ⇒ 간 (行った-)  
 가 + ㅂ니다 ⇒ 갑니다 (行きます)  
 먹 + ㄴ ⇒ 먹은 (食べた-)  
 먹 + ㅂ니다 ⇒ 먹음(습)니다 (食べます)

- 韓国語の用言には語幹と語尾が結合して活用する時、音韻縮約か形態音韻変化によって語幹や語尾の形が変わる場合が多い。語幹が変わる多くの場合は語幹の終りの子母文字が脱落したり違う文字に変化したり、または語幹の終りの子母文字が脱落すると同時に語尾の最初の文字が違う子母文字に変わる。さらに、このような形態素結合の際の形態音韻変化によって派生する音節文字の数も、数百種類に及ぶ。

돕 + 어 ⇒ 도와 (助けて)  
 들 + 어 ⇒ 들어 (聞いて)  
 모르 + 어 ⇒ 몰라 (知らなくて)  
 파랗 + 어 ⇒ 파래 (青くて)  
 놓 + 어 ⇒ 놔 (놓아) (置いて)  
 쓰 + 어 ⇒ 써 (書いて)

従って、このような韓国語の形態音韻的特徴のために、韓国語の形態素処理を次の日本語のように音節文字単位とすることはできない。

- 今日予約してたキムヨンヒですけど。

|今日|予約|し|て|た|キム|ヨンヒ|です|けど|。|

오늘 날짜로 예약한 김영희인데요.

|오늘|날짜|로|예약|하|ㄴ|김|영희|이|ㄴ데|요|。|

- いま替えられますか。

|いま|替え|られ|ます|か|。|

지금 바꿔 드릴까요?

|지금|바꾸|어|드리|ㄴ까|요|?|

二章では効率的な韓国語形態素処理のために考案した可読性と効率性を持っているローマ字ハングル字素を提示するとともに、このローマ字ハングル字素に基づくハングル変換手法と音節変換手法について述べる。

韓国語形態素処理を行なう時に必要なもう一つのツールは質の高い韓国語形態素辞書を構築するために必要な韓国語形態素の語彙情報に関する体系である。韓国語は日本語と同じく膠着語である。膠着語は語幹に語尾が付いて活用する特徴を持つ。韓国



- ある位置で与えられた子音が初声子音なのか終声子音なのかを区別するのが難しいので、子母文字列から音節文字を一意に生成できない。これは実際にコンピュータを利用した韓国語処理に大きな負担になる。

- 初声子音と終声子音が同じ文字を使用するために不必要な同音異義語が生じる。

물고 ⇒ ㄹ ㅏ ㄹ ㅓ ㅓ ㅓ (watery)

물꼬 ⇒ ㄹ ㅏ ㄹ ㅓ ㅓ ㅓ (an entry irrigating a rice field)

- 初声子音と終声子音の区別が難しく、また、母音の表現が可変的なために特定文字列の検索に不必要な計算を多く行なうことになる。
- 音韻的には黙音（表記はあるが発音されない音）であるが表記の便宜上、一つの文字として認定されている初声‘ㅇ’の使用と音節母音の可変性のために形態音韻変化を体系的に取り扱うのが難しい。
- コンピュータ内部では一つのハングル子母文字は2バイト（または3バイト）で表現されるのでハングル子母文字そのままの使用はメモリー使用の増加を引き起こす。

上記の問題点を考慮して我々は次のようなローマ字子母文字を考案した。実際にこのローマ字子母文字は、Kim(1994)のものにローマ字の発音とハングル子音母音文字の発音を考慮して若干の修正を加えたものである。

- 初声子音: g(ㄱ), n(ㄴ), d(ㄷ), l(ㄹ), m(ㅁ), b(ㅂ), s(ㅅ), \*\*\*\*\*, j(ㅈ), c(ㅊ), k(ㅋ), t(ㅌ), p(ㅍ), h(ㅎ)
- 半母音: w, y
- 母音: a(ㅏ), E(ㅓ), A(ㅗ), e(ㅓ), o(ㅜ), u(ㅡ), U(ㅡ), i(ㅣ)
- 終声子音: G(ㄱ), N(ㄴ), D(ㄷ), L(ㄹ), M(ㅁ), B(ㅂ), S(ㅅ), Q(ㅇ), J(ㅈ), C(ㅊ), K(ㅋ), T(ㅌ), P(ㅍ), H(ㅎ)

上記の子母文字として書けない音は二つ以上の子母文字を合わせて書く。

- 二重子音 (初声): gg(ㄱㄱ), dd(ㄷㄷ), bb(ㅂㅂ), ss(ㅅㅅ), jj(ㅈㅈ)
- 二重子音 (終声): GG(ㄱㄱ), SS(ㅅㅅ)
- 複合子音 (終声): GS(ㄱㅅ), NJ(ㄴㅈ), NH(ㄴㅎ), LG(ㄹㄱ), LM(ㄹㅁ), LB(ㄹㅂ), LS(ㄹㅅ), LP(ㄹㅍ), LT(ㄹㅌ), LH(ㄹㅎ), BS(ㅂㅅ)

- 複合母音: ya(ㅏ), yA(ㅑ), yo(ㅗ), yu(ㅠ), yE(ㅓ), ye(ㅕ), wa(ㅘ), wE(ㅙ), wi(ㅚ), wA(ㅜ), we(ㅝ), yi(ㅞ), yU(ㅡ)

結果的に韓国語の単語の音節表現をローマ字ハングル子母文字で表現すると次のようになる。

한국어	⇒	h a N g u G A	(韓国語)
일본어	⇒	i L b o N A	(日本語)
쌍방향	⇒	s s a Q b a Q h y a Q	(双方向)
기계번역	⇒	g i g y e b A N y A G	(機械翻訳)

## 2.2 ハングルコードの変換方法

韓国の国家標準に認定されているコンピュータハングルコード体系には完成型コードと組合せ型コード体系 (KSC5601) がある。完成型ハングルコード体系は韓国語表記でよく使用する 2350 余の音節文字に対し辞書の順序に基づく各音節を単純コード化したものである。

それゆえ、一つの音節に対する完成型ハングルコードからその音節の構成子母文字に対する情報を直接的に得ることはできない。これに対して組合せ型ハングルコード体系は音節構成を“初声子音 + 中声母音 + 終声子音”の形態に単純化して一つの音節のコードをそれらの構成子母文字コードから自動的に得られるように考案したものである (図 1 参照)。したがって、一つの音節に対する組合せ型ハングルコードは完成型コードとは異なりそのコードからその音節の構成子母文字情報を容易に得ることができる。

しかし、国家間の通信問題等の理由で大部分のコンピュータは完成型ハングルコード体系を使用している。この場合、ハングル子母文字に対する情報を得るためには、完成型コード一つ一つについての子母文字情報を持つテーブルがなければならない。このようなテーブルは機能的に見る時、完成型ハングルコードに対応する組合せ型ハングルコードを対応させたコード変換テーブル (表 1) と同じである。このようなコード変換テーブルは容易に求められるので完成型コードから子母文字情報を得るために別にテーブルを構成しなくてもよい。実際に我々が韓国語の形態素処理のため作ったハングルコード変換プログラムは韓国科学技術院で作成したコード変換テーブルを利用している。2.2.1 は韓国語形態素解析のために必要な韓国語音節文字 (完成型ハングルコード) からローマ字ハングル子母文字への変換について述べる。2.2.2 は韓国語形態素生成のため必要であるローマ字ハングル子母文字から音節文字 (完成型ハングルコード) への変換について述べる。

1	xxxxx	yyyyy	zzzzz
---	-------	-------	-------

	初声子音 (xxxxx)	中声母音 (yyyyy)	終声子音 (zzzzz)
00000			
00001			
00010		a	G
00011		E	GG
00100			GS
00101		ya	N
00110		yE	NJ
00111		A	NH
01000			D
01001		e	L
01010	g	yA	LG
01011	gg	ye	LM
01100	n		LB
01101	d	o	LS
01110	dd	wa	LP
01111	l	wE	LT
10000	m		LH
10001	b	wi	M
10010	bb	yo	B
10011	s	u	BS
10100	ss		S
10101		wA	SS
10110	j	we	Q
10111	jj	yi	J
11000	c		C
11001	k	yu	K
11010	t	U	T
11011	p	yU	P
11100	h		H
11101		i	
11110			
11111			

図1：2バイト組合せ型ハングルコード体系。

### 2.2.1 音節文字からローマ字ハングル子母文字への変換

完成型ハングルコードで表現された韓国語音節文字は次のような順序によりローマ字ハングル子母文字列に変えられる。

表1：ハングルコード変換テーブル.

ハングル音節文字	完成型コード	組合せ型コード
가	B0A1	A841
카	B0A2	A842
간	B0A3	A845
강	B0A4	A848
...	...	...
힉	C8FE	F3B6

- 与えられた入力各音節に対して次のことを実行する:

段階 1 : ハングルコード変換テーブルを利用して、与えられた音節の完全型コードを対応する組合せ型コードに変換する。

段階 2 : 2バイト組合せ型コードの一番目のビットの次から5ビットずつ三部分に分けて各部分に対して次のような作業を行なう。

- 一番目部分の5ビットをキーにして初声子音文字テーブル(図1)で該当するローマ字ハングル子母文字を出力する。
- 二番目部分の5ビットをキーにして中声母音文字テーブル(図1)で該当するローマ字ハングル子母文字を出力する。
- 三番目部分の5ビットをキーにして終声子音文字テーブル(図1)で該当するローマ字ハングル子母文字を出力する。

### 2.2.2 ローマ字ハングル子母文字から音節文字への変換

ローマ字ハングル子母文字は次のような順序により完成型ハングルコードで表現された音節文字に変換できる。

- 与えられたローマ字ハングル子母文字列の文字を順次的に探索しながら次のことを実行する:

段階 1 : 現在探索された文字の類形にしたがって次を実行する。

- その類形が初声子音であり、そのような類形の出現が初めてならそれを  $C_{i-1}$  に保存して、その出現が初めてでなければ  $C_{i-2}$  に保存する。
- その類形が半母音ならそれを G に保存する。
- その類形が母音ならそれを V に保存する。

- その類形が終声子音であり、そのような類形の出現が初めてならそれを Cf-1 に保存して、その出現が初めてでなければ Cf-2 に保存する。

段階 2 : 表 2 の予備参照テーブルを利用して現在探索された文字が音節の境界を成しているかを検査してそれが音節の境界を形成する文字である場合に限って次を実行する。

- Ci-1 と Ci-2 をキーにして初声子音文字テーブル (図 1) で該当する組合せ型ハングルコードのビット列 (xxxxx) を求める。
- G と V をキーにして中声母音文字テーブル (図 1) で該当する組合せ型ハングルコードのビット列 (yyyyy) を求める。
- Cf-1 と Cf-2 をキーにして終声子音文字テーブル (図 1) で該当する組合せ型ハングルコードのビット列 (zzzzz) を求める。
- 組合せ型ハングルコード “1xxxxxyyyyyzzzzz” をキーにしてハングルコード変換テーブルで該当する完成型ハングルコードを求め、それを出力する。
- Ci-1、Ci-2、G、V、Cf-1、Cf-2 を初期化する。

表 2 : コード変換のための予備参照テーブル.

i \ i+1	初声子音文字	半母音文字	母音文字	終声子音文字	特殊文字
初声子音	□	□	□	▽	▽
半母音	▽	▽	□	▽	▽
母音	○	○	○	□	○
終声子音	○	○	○	□	○
特殊文字	○	○	○	○	○

表 2 はローマ字ハングル子母文字から完成型ハングルコードに表現された音節文字を生成するため予備参照テーブルである。表 2 のテーブルの i は現在探索された子母文字を示し、(i+1) は次に探索する文字を示す。テーブル内にある○は現在探索された文字とつぎに探索する文字の間が音節境界であることを表し、□はその間が音節境界ではないことを表す。▽は子母文字列ができていないことを表す。

### 3 韓国語形態素の語彙情報に対する体系

この章では韓国語と日本語間の対話翻訳のために基本的に必要な韓国語形態素辞書の構成と語彙情報の体系を述べる。一つの言語の語彙情報体系を組み立てることは多

くの時間と努力が必要である。また、それは一人の努力で成し遂げることはできない。今まで韓国語語彙情報の体系に関する研究は多くはなかった。いくつかの研究結果があるものの、それらはたいていの場合、形態素の範疇分類に重点を置いており、分類も複雑すぎて一般的に使用する時は理解しにくい。このような問題点のために韓国語形態素の語彙情報体系に関する研究が必要であった。韓国語形態素の語彙情報体系を組み立てるにあたり重点を置いた部分は次のとおりである。

- 形態素範疇を可能な限り文法に基づいて分類した。言い替えればシステムに依存する形態素範疇に分類することは避けた。
- 語彙情報は形態素処理や語彙の正規化に必要なことだけで構成する。

### 3.1 韓国語形態素辞書の構成

韓国語形態素辞書は次のような形態素情報を持つ語彙レコードの集合として構成される。

#### 3.1.1 表層形態 ( Surface Form; S-Form )

表層形態は一つの形態素の表層表現を表す。一つの形態素の表層表現がいろいろある場合、表層形態は形態音韻的な説明が可能な範囲内でその中の基本形だけを取る。形態音韻的な説明が難しい異形態はすべて一つの独立した表層形態になる。一つの形態素に対する基本形の基準はハングル綴字法 ( 或いは韓国語文法書 ) で規定しているものを原則にする。但し、 “ 으로, 으니까, 으면, … ” や “ 이나, 이면, 이라고, … ” に関しては ‘ 으 ’ や ‘ 이 ’ を除いた残りのものを基本形とみなす。また、音韻規則として説明が難しい “ 가/이, 를/을, 는/은, 와/과 ” のような助詞は二つの異なる表層形態を認める。

#### 3.1.2 字素形態 ( Grapheme Form; G-Form )

字素形態はローマ字ハングル字母文字列で表現された表層形態である。

#### 3.1.3 音素形態 ( Phoneme Form; P-Form )

音素形態は音声認識や音声合成との関係を考慮した音素で表現された表層形態である。しかし、これは音素文字に対する表記が難しく言語翻訳では直接使われないので、保留した。

### 3.1.4 正規形態 ( Regular Form; R-Form )

正規形態は表層形態に対する標準表現を表す。このような正規形態は表層形態が方言や形態音韻規則による取り扱いが難しい縮約語である場合、その標準語、或いは縮約しない言葉にする。また、“이, 을, 은, 과”のような助詞の場合、それに対応する正規形態は“가, 를, 는, 와”にする。

### 3.1.5 語彙情報 ( Lexical Information; L-Info )

語彙情報は形態素処理に必要な表層形態に関する言語情報である。この語彙情報には表層形態の範疇情報、可算情報、変則情報、接続情報、活用情報がある。形態素の範疇情報、可算情報、変則情報、活用情報は3.2節、3.3節、3.4節、3.5節で各々取り扱う。形態素の接続情報は、表層形態に先行することができる形態素の形態範疇に関する情報を表層形態が語尾類とか接尾辞類についてのみ表す。

## 3.2 範疇情報 ( Morphological Category Information )

### 3.2.1 固有名詞 ( Proper Noun; NPROP )

固有名詞は同じ性質の対象の中でどれか一つを他のものと特別に区別する必要があるとき使用される名詞で人名、国名、地名、団体名、職業名、署名、商品名等がある。

例) 영철, 스즈끼, 롯데호텔, 한국, 일본

- 固有名詞は特別な対象に対して付けられた名前なのでそのような属性を持つ対象はたった一つしかない。このような特殊性のために固有名詞は数と関連した言葉と結合しない。
- 固有名詞は長単位分割を基本にする。但し、短単位分割が生成力を持つ場合はそのようにする。例えば、人名の場合は姓と名前に分割する。

### 3.2.2 動作名詞 ( Active Noun; NACTV )

動作名詞は“하다”を付けて動詞を作ることができる名詞でどこに現れても動作性を持つ。

例) 예약, 숙박, 여행, 회의, 문의, 확인, 전화

例) 저는 **전화로 예약**을 한 김영철입니다.

### 3.2.3 形容名詞 ( Adjective Noun; NADJV )

形容名詞は“하다”を付けて形容詞を作ることができる名詞で事物の状態や性質を表す。

例) 만족, 감사, 곤란, 공정, 건강, 간단, 동일

### 3.2.4 普通名詞 ( Common Noun; NCOMM )

普通名詞は、同じ性質を持った対象に対して残らずすべて付けることができる言葉の中で動作名詞と形容名詞を除いたものを言う。

例) 여권, 성함, 주소, 열쇠, 구름, 원, 대학교, 영수증

- 複合名詞は意味の生成力を与える範囲内での短単位分割を原則とする。但し、この時の短単位は国語大辞典（ソウル：民衆書林：94、リヒスン）に対象複合語が一つの見出し語で登録されていない場合に限定する。
- 複合名詞を短単位で分割する場合、そこに含まれている動作名詞や形容名詞はそれをそのまま認定する。

### 3.2.5 代名詞 ( Pronoun; NPRON )

代名詞は事物に名前を付けずにただ示すだけの言葉である。

例) 나, 너, 이, 그, 저, 이것, 그것, 저것, 여기, 저기, 거기

### 3.2.6 数詞 ( Numeral; NNUMB )

数詞は事物の数量や順序を表す言葉で短単位分割を原則とする。

例) 하나, 둘, 셋, 일, 이, 삼, 첫

### 3.2.7 動詞語幹 ( Verb Stem; VVERB )

動詞語幹は事物の動きを過程的に表示する言葉で語尾と結合して多様に活用するのが普通である。

例) 가다, 보다, 잡다, 타다, 먹다, 마시다, 자다

- 動詞は国語大辞典（ソウル：民衆書林：94、リヒスン）では動詞語幹に文末語尾‘다’が付いた形態“動詞語幹+다”と同じ見出し語になっているがここでは動詞語幹部分のみを認定する。

### 3.2.8 形容詞語幹 ( Adjective Stem; VADJV )

形容詞は事物の性質や状態を表示する言葉で動詞と同じように語尾と結合して多様に活用する。

例) 괜찮다, 고맙다, 좋다, 크다, 깨끗하다, 높다, 즐겁다

- 形容詞は動詞と同じように国語大辞典 (ソウル: 民衆書林: 94、リヒスン) では形容詞語幹に文末語尾 ‘다’ が付いた形態 “形容詞語幹+다” と同じ見出し語になっているが、ここでは形容詞語幹部分のみを認定する。

### 3.2.9 補助用言 ( Auxiliary Verbal Stem; VAUX1 )

補助用言は、語幹に付いて使われながらもその言葉に文法的意味を加える自立性がないような、弱い動詞や形容詞を言う。

例) 어 주, 어 드리, 어 보, 어야 하, 고 싶어 하, 지 않으면 안되

### 3.2.10 先語末語尾 ( Verbal Prefinal-Ending; VAUX2PEND )

先語末語尾は、それ自体だけでは単語を完成させることができなくて、常に ‘-다’ のような閉鎖形語尾を必然的に要求する開放形語尾である。この先語末語尾が語幹に与える文法的な情報には主に尊敬と時制がある。

- 尊敬先語末語尾: -(으)시-, -(으)옵-
- 時制先語末語尾: -었-(-았-, -ㅁ-), -겠-, -았었-(-았었-, -ㅁ었-), -ㄴ(-는-), 더

### 3.2.11 転成連結語尾 ( Verbal Transformative-Ending; VAUX2TEND )

転成連結語尾は文章の転成や接続の機能を持った語尾を言う。

- 名詞形転成語尾: 名詞形転成語尾は文章の機能を名詞化する語尾である。

例) -(으)ㅁ, -기

- 冠形形転成語尾: 冠形形転成語尾は文章を冠形詞のように変える語尾である。

例) -ㄴ, -는, -르, -던

- 副詞形転成語尾: 副詞形転成語尾は文章の機能を副詞化する語尾である。

例) -게

- 接続形連結語尾: 接続形連結語尾は前の文章を後ろの文章につなげる機能を持つ語尾である。

例) -고, -면서, -든지 -든지, -(으)며, -지만, -(으)나, -(으)면, -(으)니, -는데, -(으)므로

### 3.2.12 文末語尾 ( Verbal Sentential-Ending; VAUX2SEND )

文末語尾は一つの文章を終らせる機能を持っている語尾である。

- 叙述形文末語尾: -(으, 스)바니다, -(으, 스)바디다, -(으)리다, -데, -(으)리, -더라, -(으)리라, -소, -네, -나이다, -(으)이, -다; -(으)오리다, -(으)리다, -(으)마세, -(으)마
- 疑問形文末語尾: -(으, 스)바니까, -(으, 스)바디까, -(으)리까, -(으)오, -던가, -(으)리라, -더냐, -(으)라, -소, -(으, 느)냐, -오이까, -(으)냐, -나йка, 나, 느냐, 어, 어요, 지, 지요
- 命令形文末語尾: -(으)소서, -(으)시오, -(으)시오, -(으)오, -소, -게, -어라(-아라), -(으)려무나, -(으)렴, -게나, 어, 어요, 지, 지요
- 勸誘形文末語尾: -(으)시오, -(으)세, -자, 어, 어요, 지, 지요
- 感嘆形文末語尾: -(으)바니다그러, -(는)구나, -(는)구려, -(는)구먼, -네그러, -(으)이그러, -요그러, -리세그러, -로구나

### 3.2.13 冠形詞 ( Adnominal Modifier; MADNO )

冠形詞は体言の前についてその体言の意味を限定する自立語である。

例) 어떤, 어느, 아무, 무슨, 이, 그, 한, 몇, 여러

### 3.2.14 接続詞 ( Conjunction; MCONJ )

接続詞は前の文章の意味を後ろの文章に接続させながら、それを修飾する副詞的性格を持つ自立語である。

例) 그리고, 그러나, 하지만, 그런데, 그렇지만, 그렇지만은, 그러면, 그렇다면, 그럼, 따라서, 그러므로, 그래서, 또, 또한, 더구나, 오히려, 곧, 즉, 및, 또, 또한, 또는, 혹은

### 3.2.15 副詞 ( Adverb; MADVB )

副詞は用言やある言葉の前に置かれてその言葉の意味を限定する自立語である。

- 性状副詞: 잘, 높이, 빨리, 매우, 딱, 가장, 바로, 겨우, 아주, …
- 指示副詞: 이리, 저리, 요리, 여기, 거기, 저기, 오늘, 어제, …
- 不正副詞: 못, 안
- 様態副詞: 과연, 물론, 정말, 설마, 아마, 만일, 설령, 비록, …

### 3.2.16 感嘆詞 ( Interjection; INTER )

感嘆詞は話者が自分の感じたことや意志を特別な単語によることなく、直接的に表示する言葉である。従って感嘆詞は単独で話者の感情と意志が表現できるので次に他の言葉が継続されなくても独立した文章のような機能を持つ。

- 感情感嘆詞: 허허, 하하, 에끼, 아이고, 후유, 아, 애고, …
- 意志感嘆詞: 예, 글썽요, 아니오, 여보세요, 고마워요, 죄송합니다, …
- 口ぐせ、問い返し等の感嘆詞: 머, 뭐, 그래, 예, 거시기, 음, 에헴, …

### 3.2.17 主格助詞 ( Agentive Particle; PAGNT )

主格助詞は体言に付いて、その体言に文章の主語資格を持たせる助詞である。

例) 가(이), 께서, 에서, 서

### 3.2.18 冠形格助詞 ( Adnominal Particle; PADNO )

冠形格助詞は先行する体言に付いてその体言が後ろの体言に対して冠形語になるような機能を持つ助詞である。

例) 의

### 3.2.19 目的格助詞 ( Objective Particle; POBJT )

目的格助詞は体言に付いてその体言が後に来る他動詞の目的語になるような機能を持つ助詞である。

例) 를(을)

### 3.2.20 叙述格助詞 ( Verbal Particle; PVERB )

叙述格助詞は体言の後ろに付いて、その体言に叙述する機能を与える。国語大辞典(ソウル:民衆書林:94、リヒスン)や一般文法書で叙述格助詞は指定詞‘이’に語尾が合成された言葉として説明されているが、ここでは指定詞‘이’以後に出てくる語尾が一般用言の語幹の後に出てくる語尾と同じように生成力もあるので指定詞‘이’だけを叙述格助詞として取り扱う。従ってここでは叙述格助詞‘이’は派生接尾辞的役割をするとも言える。

例) 이

### 3.2.21 接続助詞 ( Conjunctive Particle; PCONJ )

接続助詞は二つ以上の体言を同じ資格で接続させる機能を持つ助詞である。

例) 와(과), 하고, (이)며, 에다, (이)랑

### 3.2.22 副詞格助詞 ( Adverbial Particle; PADVB )

副詞格助詞は体言に付いて体言を副詞語になるようにする助詞である。

例) 에, 서, 에서, 에게, 한테, 더러, 께, 에게서, 한테서, (으)로, 에게로, 께로, 한테로, (으)로서, 가(이), 와(과), 하고, 처럼, 같이, 대로, 만큼, 만, 보다, (이)라고, 고, (이)라, 야(아), (이)여, (이)시여

### 3.2.23 主題補助詞 ( Subjective Particle; PSUBJ )

主題補助詞は体言、格助詞、副詞、連結語尾等に付いて先行する言葉に特殊な意味を加える補助詞の中で、その言葉が文章の主題としての役割を持つ助詞である。

例) 는(은, ㄴ), 도, 만

### 3.2.24 一般補助詞 ( Auxiliary Particle; PAUXI )

一般補助詞は格助詞、副詞、連結語尾等に付いて先行する言葉に特殊な意味を加える補助詞の中で主題補助詞を除いた残りの助詞である。

例) 부터, 까지, 마다, (이)야, (이)야말로, (이)든지, (이)라도, (이)닐, 마저, (이)나마, (이)나, 는커녕(은커녕), 커녕, 조차, 서건, 밖에, 그러

### 3.2.25 接頭辞 ( Prefix; XPREF )

接頭辞は後に来る語幹に結合して語幹の意味を制限する接辞である。韓国語の接頭辞は名詞、動詞、形容詞に結合してその意味を制限する機能を持っている。しかし、たいていの韓国語接頭辞は特定語彙とだけ結合してその語彙の意味を制限するので生成力がとても弱い。そのような理由でここでは比較的生成力が高い名詞接続形の漢字接頭辞だけを認定する。

例) 비-, 미-, 부-, 불-

### 3.2.26 名詞形接尾辞 ( Nominal Suffix; XSUFFNOUN )

名詞形接尾辞は先行する体言に付いて、その体言に新たな意味を与える機能を持つ接辞である。

例) -들, -씨, -님, -집, -호, -가

### 3.2.27 動詞派生接尾辞 ( Verb-deriving Suffix; XSUFFVERB )

動詞派生接尾辞<sup>1</sup>は先行する動作名詞に付いて名詞を動詞化する機能を持っている接辞だけを認定する。動詞と同じように動詞派生接尾辞は語尾が付いて多様に活用する。

例) -하, -드리, -되, -시키

### 3.2.28 形容詞派生接尾辞 ( Adjective-deriving Suffix; XSUFFADJV )

形容詞派生接尾辞<sup>2</sup>は先行する形容名詞や普通名詞に付いて名詞を形容詞化する機能を持つ接辞だけを認定する。

例) -하, -답, -스럽

### 3.2.29 副詞派生接尾辞 ( Adverb-deriving Suffix; XSUFFADV )

副詞派生接尾辞<sup>3</sup>は先行する形容名詞に付いて、その言葉を副詞化する機能を持つ接辞“-히”だけを認定する。

---

<sup>1</sup>韓国語の動詞派生接尾辞は先行する名詞、形容詞、副詞に付いて、その言葉を動詞化する接辞を指し示す。しかし、形容詞と副詞に付いてその言葉を動詞化する動詞派生接尾辞は出現するのが稀なので、ここではこれを認定しない。

<sup>2</sup>韓国語の形容詞派生接尾辞は先行する名詞、動詞、副詞、冠形詞に付いて、その言葉を形容詞化する接辞を指し示す。しかし、動詞と副詞と冠形詞に付いて、その言葉を形容詞化する形容詞派生接尾辞はその分布が非常に制約され、また不規則的なので、ここではこれを認定しない。

<sup>3</sup>韓国語の副詞派生接尾辞は先行する名詞、動詞、形容詞に付いて、その言葉を副詞化する接辞を指し示す。しかし、普通名詞と動詞と形容詞に付いて、その言葉を副詞化する副詞派生接尾辞はその分布が非常に制約され、生成力が弱いのでここではこれを認定しない。

### 3.2.30 記号 ( Symbol; SYMBL )

記号は韓国語の文章に現れる文章記号を現わす。ここでは対話文と言う特殊性を考慮して以下の終止符と休止符だけを認定する。

- 終止符: 点 (.) 疑問符 (?) 感嘆符 (!)
- 休止符: コンマ (,)

### 3.3 可算情報 ( Number Information )

可算情報は表題語が普通名詞である場合にだけ現われ、それは普通名詞の前に数量を表す言葉が直接的に来るのかどうかに関する情報を表現する。

- 불가(不可): この普通名詞は文章内でそのすぐ前に数量を表す言葉が直接的に来ることはない。
- 한자수 (漢字数): この普通名詞は文章内でそのすぐ前に数量を指示する言葉が直接的に来ることはできるが、その言葉は漢字数でなければならない。

영 (공)	0	零	zero
일	1	一	one
이	2	二	two
삼	3	三	three
사	4	四	four
오	5	五	five
육	6	六	six
칠	7	七	seven
팔	8	八	eight
구	9	九	nine
십	10	十	ten
백	100	百	a hundred
천	1000	千	a thousand
만	10000	萬	ten thousand
억	100000000	億	a hundred thousand
조	1000000000000000000	兆	a billion

- 한글수 (ハングル数): この普通名詞は文章内でそのすぐ前に数量を指示する言葉が直接的に来ることはできるが、その言葉はハングル数でなければならない。ハングル数は1から99までの数に対して以下のように韓国語固有の数が使われており、それ以上は“백、천、만、억、조”と同じ漢字数が補充される。

하나 (한)	1	一つ	one
둘 (두)	2	二つ	two
셋 (세)	3	三つ	three
넷 (네)	4	四つ	four
다섯	5	五つ	five
여섯	6	六つ	six
일곱	7	七つ	seven
여덟	8	八つ	eight
아홉	9	九つ	nine
열	10	十	ten
스물 (스무)	20	二十	twenty
서른	30	三十	thirty
마흔	40	四十	forty
쉰	50	五十	fifty
예순	60	六十	sixty
일흔	70	七十	seventy
여든	80	八十	eighty
아흔	90	九十	ninety

- 선택 (選択) : この普通名詞は文章内でそのすぐ前に數量を指示する言葉が直接的に来ることはできるが、その言葉は文脈によって漢字数であることもハンゲル数であることもある。

### 3.4 変則情報 ( Morphophonological Information )

変則情報は、用言の語幹に語尾が付いて活用する時に語幹と語尾の形が環境により一定に変化せず、語幹に従って不規則的に変化する形態変異の処理のための形態音韻規則情報である。この変則情報は見出し語が動詞、形容詞、補助用言、動詞派生接尾辞 ‘-하’ と形容詞派生接尾辞 ‘-하’ にだけ現われ、次のような種類がある。

- 규칙 (規則; regul) : 語幹に語尾が付いて活用する時に語幹と語尾の形が全く変化しないなど環境により一定の変化をする。

(1a) 먹다, 먹고, 먹지, 먹더라, 받다, 받고, 받지, 받더라

(1b) 먹어, 먹으니, 먹었다, 받아, 받으니, 받았다

- ㅅ불규칙 (S-不規則; irreS) : ‘ㅅ’ 終声を持った用言の語幹には母音の語尾の前でその ‘ㅅ’ が脱落するものがある。

(2a) 짓다, 짓고, 짓지, 짓더라

(2b) 지어, 지으니, 지었다

(2a') 벗다, 벗고, 벗지, 벗더라

(2b') 벗어, 벗으니, 벗었다

- S- 不規則用言: 짓다, 굶다, 낫다, 잇다, 짓다, 붓다, ...

- S- 規則用言: 벗다, 빗다, 빼앗다, 씻다, 솟다, ...

- ㅂ불규칙 (B- 不規則; irreB) : 'ㅂ' 終聲を持った用言の語幹には母音の語尾の前でその 'ㅂ' が半母音 'ㅓ' や母音 '우' に変わるものがある。

(3a) 굽다(bake), 굽고, 굽지, 굽더라

(3b) 구워, 구우니, 구웠다

(3a') 굽다(curve), 굽고, 굽지, 굽더라

(3b') 굽어, 굽으니, 굽었다

- B- 不規則用言: 굽다(bake), 돕다, 줍다, 곱다, 덩다, ...

- B- 規則用言: 굽다(curve), 뽑다, 입다, 잡다, 줍다, ...

- ㄷ불규칙 (D- 不規則; irreD) : 'ㄷ' 終聲を持った動詞の語幹には母音の語尾の前でその 'ㄷ' が 'ㄹ' に変わるものがある。

(4a) 듣다, 듣고, 듣지, 듣더라

(4b) 들어, 들으니, 들었다

(4a') 달다, 달고, 달지, 달더라

(4b') 달아, 달으니, 달았다

- D- 不規則動詞: 듣다, 건다, 묻다(ask), 일컫다, 깨닫다, ...

- D- 規則動詞: 달다, 돌다, 묻다(bury), 믿다, 얻다, ...

- ㄹ불규칙 (IU- 不規則; irrIU) : 語幹が 'ㄹ' で終る大部分の用言は '어' で始まる語尾と結合する時 'ㄹ' の '으' が脱落すると同時に 'ㄹ' が生じる。この時 'ㄹ' の 'ㄹ' は語幹の終聲音になり、生じた 'ㄹ' は語尾の初聲音になる。

(5a) 흐르다, 흐르고, 흐르지, 흐르더라

(5b) 흘러, 흘르니, 흘렀다

- IU- 不規則用言: 흐르다, 부르다, 오르다, 이르다(inform), 고르다, 다르다, 나르다, 누르다(press), 바르다, 빠르다, ...
  - IU- 規則用言: 따르다
- 러-불규칙 (IA- 不規則; irrlA) : 語幹 ‘르’ で終ったいくつかの用言は子音の語尾が付けば変化しないが、‘어’ で始まる語尾が来ればその ‘어’ を ‘러’ に変えるものがある。
    - (6a) 이르다(arrive), 이르고, 이르지, 이르더라
    - (6b) 이르러, 이르니, 이르렀다
    - IA- 不規則用言: 이르다, 누르다(yellow), 푸르다
- ㅎ-불규칙 (H- 不規則; irreH) : ‘ㅎ’ 終声を持った大部分の形容詞は母音の語尾と結合する時、語幹が語尾と共に変形する。
    - (7a) 파랗다, 파랗고, 파랗지, 파랗더라
    - (7b) 파래, 파라니, 파랴다
    - (7a') 좋다, 좋고, 좋지, 좋더라
    - (7b') 좋아, 좋으니, 좋았다
    - H- 不規則形容詞: 파랗다, 까맣다, 노랗다, 빨갱다, 보얗다, ...
    - H- 規則形容詞: 좋다
- 여-불규칙 (yA- 不規則; irryA) : ‘하’ になった用言やそれが付いてつながった用言に子音語尾が結合すれば、語尾に変化は起こらないが母音語尾が付けば語尾が不規則的に活用する。規則用言に対する母音調和規則は語幹の最後の母音が陽性母音(‘아’ や ‘오’) である場合は ‘아’ 系統の語尾をとるのに対し、yA- 不規則用言は ‘여(yA)’ 系統の語尾をとる。
    - (8a) 하다, 하고, 하니, 하더라
    - (8b) 하여 (해), 하였다 (했다)
    - (8a') 가다, 가고, 가니, 가더라
    - (8b') 가야 (가), 가왔다 (갔다)

- 거라-불규칙 (gAla- 不規則; irrgA) : 最後の母音が ‘아(a)’ である動詞語幹は一般的な語尾が付けば語尾自体には変化はないが、命令形語尾 ‘-아라’ が付けばその ‘-아라’ を ‘-거라’ に変えるものがある。実際に gAla- 不規則動詞は自動詞だけに該当する。

(9a) 가다, 가고, 가니, 가서

(9b) 가거라

(9a') 사다, 사고, 사니, 사서

(9b') 사아라(사라)

— gAla- 不規則動詞: 가다, 자다, 자라다, 일어나다

- 녀라-불규칙 (nAla- 不規則; irrnA) : 最後の母音が ‘오’ である動詞語幹は一般的な語尾が付けば語尾自体には変化はないが、命令形語尾 ‘-아라’ が付けば ‘-녀라’ に変わるものがある。nAla- 不規則動詞は gAla- 不規則動詞と同じく自動詞だけに該当する。

(10a) 오다, 오고, 오니, 와서

(10b) 오녀라

(10a') 보다, 보고, 보니, 봐서

(10b') 보아라

— nAla- 不規則動詞: 오다, 나오다, 들어오다

- 우-불규칙 (u- 不規則; irreu) : ‘우(u)’ で終わっている ‘푸다’ の ‘푸-’ は ‘어’ で始まる語尾の前では ‘우’ が脱落する。

(11a) 푸다, 푸고, 푸니, 푸더라

(11b) 퍼, 뵈다

(11a') 주다, 주고, 주니, 주더라

(11b') 주어, 주었다

— u- 不規則動詞: 푸다

— u- 規則動詞: 주다, 두다, 추다, ...

### 3.5 活用情報 ( Conjugation Information )

活用情報は語幹に語尾が付いて活用する時に語幹に加えられる語尾の文法的な情報を表す。

- 活用情報は接続する語尾の種類に従って時制、相、叙法、様相等の情報を表現する

존경 (尊敬) : -(으)시-, -(으)옵-

현재 (現在) : -ㄴ-, -는-

과거 (過去) : -았(-았-, -ㅁ-)

의지 (意志) : -겠-

대과거 (大過去) : -었었(-았었-, -ㅁ었-)

회상 (回想) : -더-

명사형 (名詞形) : -ㅁ, -기

관형형 (冠形形) : -ㄴ-, -ㄹ-, -는-, -던

부사형 (副詞形) : -게

접속형 (接續形) : -고, -면서, -든지 -든지, -(으)며, -지만, -(으)나, -(으)면, -(으)니, -는데, -(으)므로, ...

서술형 (敘述形) : -(으, 스)ㅁ니다, -(으, 스)ㅁ디다, -(으)리다, -데, -(으)리, -더라, -(으)리라, -소, -네, -나이다, -(으)이, -다, -(으)오리다, -(으)리다, -(으)ㅁ세, -(으)ㅁ마, ...

의문형 (疑問形) : -(으, 스)ㅁ니까, -(으, 스)ㅁ디까, -(으)리까, -(으)오, -던가, -(으)리까, -더냐, -(으)라, -소, -(으, 느)냐, -오이까, -(으)냐, -나이가, 나, 느냐, 어, 어요, 지, 지요, ...

명령형 (命令形) : -(으)소서, -(으)ㅁ시오, -(으)ㅁ쇼, -(으)오, -소, -게, -어라(-아라), -(으)려무나, -(으)렴, -게나, 어, 어요, 지, 지요 ...

청유형 (勸誘形) : -(으)ㅁ시다, -(으)세, -자, 어, 어요, 지, 지요, ...

감탄형 (感嘆形) : -(으)ㅁ니다그러, -(는)구나, -(는)구려, -(는)구먼, -네그러, -(으)이그러, -요그러, -르세그러, -로구나, ...

- 補助用言は本用言に付いてそこに文法的な情報を加える機能を持っているので語尾のように活用情報を持っているものとする。

시도 (試行) : 어 보 (아 보)  
 진행 (進行) : 고 있, 어 오 (아 오), 어 가 (아 가), 어 나가 (아 나가)  
 가능 (可能) : (으)리수 있, (으)리수가 있, (으)리줄 알  
 부정 (否定) : 지 않, 지 못하, 지 말  
 변화 (變化) : 게 되  
 결론 (結論) : (으)ㄴ것이, (으)ㄴ바 있, (으)ㄴ셈이  
 봉사 (奉仕) : 어 주 (아 주), 어 드리 (아 드리)  
 요구 (要求) : 어야 하 (아야 하), 어야 되 (아야 되)  
 완료 (完了) : 어 있 (아 있)  
 기대 (期待) : (으)ㄴ것이, (으)ㄴ터이  
 사동 (使役) : 게 하  
 피동 (受動) : 어 지 (아 지)  
 불가능 (不可能) : (으)리수 없, (으)리수가 없  
 완결 (完結) : 어 버리 (아 버리), 어 내 (아 내), 고 말  
 원인 (原因) : 기 때문이  
 추정 (推定) : (으)ㄴ것 같, (으)ㄴ듯 하, (으)ㄴ가 보, (으)ㄴ모양이  
 유지 (維持) : 어 놓 (아 놓), 어 두 (아 두)  
 희망 (希望) : 고 싶, 기 바라, (으)면 하, 고 싶어 하  
 의도 (意圖) : (으)려고 하, 고자 하  
 강조 (強調) : (으)ㄴ뿐이, (으)ㄴ따름이  
 허가 (許可) : (으)면 되, 어도 되 (아도 되), 어도 좋 (아도 좋), 지 않아도 좋  
 당위 (義務) : (으)ㄴ수 밖에 없, 지 않으면 안되  
 습관 (習慣) : 곤 하  
 경향 (傾向) : (으)ㄴ편이  
 금지 (禁止) : (으)면 안되, 서는 안되  
 과거가능 (過去可能) : (으)ㄴ뻤하  
 시작 (始め) : 어 들 (아 들)  
 가치 (價值) : (으)ㄴ만하, (으)ㄴ직하

허위 (虚偽)	: (으)ㄴ척하, (으)ㄴ체하
위기 (危機)	: (으)ㄴ지경이
상황 (狀況)	: (으)ㄴ터이
과잉 (過剩)	: 어 대 (아 대)
확신 (確信)	: 기 마련이
약속 (約束)	: 기로 하, 기로 되
욕망 (欲望)	: 어 죽겠 (아 죽겠)
제안 (提案)	: (으)면 좋, 는 것이 좋

#### 4 おわりに

韓国語の形態論的処理を行なうのに、なくてはならない重要なふたつのツールについて述べた。すなわち、それは効率的な韓国語の形態素分離と形態素生成のためのハングル変換プログラムと質の高い韓国語形態素辞書を構築するための韓国語形態素の語彙情報に関する体系である。

実際にこのふたつのツールは韓国語と日本語間の双方向機械翻訳システムを構築するのに使用した。ハングル変換プログラムは日-韓機械翻訳では韓国語形態素の結合をするのに利用し、韓-日機械翻訳では韓国語の文章の形態素分離をするのに利用した。また、形態素解析で辞書の見出し語を探索するために必要な辞書の表層形態に関する字素形態を準備するのにも利用した。韓国語形態素の語彙情報の体系は韓-日翻訳システムで韓国語処理規則を作成する基準になり、韓国語対話文のタギング作業と韓国語形態素解析及び生成辞書の情報を記入する指針書として利用した。

前に少し言及したように形態素の語彙情報に関する体系を構築するには多くの時間と努力が必要である。これにはまた、実際に自然言語処理システムで十分な実験と分析を引続き行なってその結果を再度反映させるフィードバックが要求される。韓国語形態素の語彙情報体系についてはまだこのようなフィードバックが一度も実現していない。将来これらに対する作業が実現されなければならない。

#### 参考文献

#### References

- [1] An, D. U., G. C. Kim, and J. H. Lee. 1994. Corpus-Based Modality Generation for Korean Predicates. Literary and Linguistic Computing (to appear).

- [2] Kim, D. B., S. J. Lee, K. S. Choi, and G. C. Kim. 1994. A Two-Level Morphological Analysis of Korean. Proceeding of COLING-94.
- [3] Kim, J. H. and J. H. Seo. 1994. A Korean Part-of-Speech Tag Set for Natural Language Processing. Technical Report. KAIST CS Lab (in Korean).
- [4] Lee, H. S. 1994. The Great Dictionary of Korean. Seoul: Minjung-Seolim.
- [5] Nam, K. S. and Y. G. Ko. 1994. Standard Korean Grammar (revised version). Seoul: Top Publisher (in Korean).

付録：韓国語の機能語リスト

字素形態 (Grapheme Form)	表層形態 (Surface Form)	語彙情報
A	어	문말어미 의문형
A	어	문말어미 명령형
A	어	문말어미 청유형
A bAli	어 버리	보조용언 완결
A bo	어 보	보조용언 시도
A dE	어 대	보조용언 과잉
A dUL	어 들	보조용언 시작
A dULi	어 드리	보조용언 봉사
A du	어 두	보조용언 유지
A iSS	어 있	보조용언 완료
A ji	어 지	보조용언 피동
A ju	어 주	보조용언 봉사
A juGgeSS	어 죽겠	보조용언 욕망
A ga	어 가	보조용언 진행
A nE	어 내	보조용언 완결
A naga	어 나가	보조용언 진행
A noH	어 놓	보조용언 유지
A o	어 오	보조용언 진행
ASS	었	선어말어미 과거
ASSASS	었었	선어말어미 대과거
Ado	어도	전성연결어미 접속형
Ado dwi	어도 되	보조용언 허가
Ado joH	어도 좋	보조용언 허가
Ala	어라	문말어미 명령형
AsA	어서	전성연결어미 접속형
Aya	어야	전성연결어미 접속형
Aya dwi	어야 되	보조용언 요구
Aya ha	어야 하	보조용언 요구
AyamaN	어야만	전성연결어미 접속형
Ayo	어요	문말어미 의문형
Ayo	어요	문말어미 명령형
Ayo	어요	문말어미 청유형
Bdida	ㅂ디다	문말어미 서술형
Bdigga	ㅂ디까	문말어미 의문형
Bnida	ㅂ니다	문말어미 서술형
BnidagUlyA	ㅂ니다그러	문말어미 감탄형
Bnigga	ㅂ니까	문말어미 의문형
Bsidayo	ㅂ시다	문말어미 청유형
Bsyo	ㅂ쇼	문말어미 명령형
BsyogUlyA	ㅂ쇼그러	문말어미 허락형
L	ㄹ	전성연결어미 관형형
L bbANha	ㄹ 뻔하	보조용언 과거가능
L bbuNi	ㄹ 뿐이	보조용언 강조
L ddalUMi	ㄹ 따름이	보조용언 강조
L gAS gaTUmyAN	ㄹ 것 같으면	전성연결어미 접속형
L gASi	ㄹ 것이	보조용언 기대
L jigyaQi	ㄹ 지경이	보조용언 위기

字素形態 (Grapheme Form)	表層形態 (Surface Form)	語彙情報
L juL aL	르줄 알	보조용언 가능
L maNha	르만하	보조용언 가치
L moyaQi	르모양이	보조용언 추정
L su ABS	르수 없	보조용언 불가능
L su baGGe ABS	르수 밖에 없	보조용언 당위
L su iSS	르수 있	보조용언 가능
L suga ABS	르수가 없	보조용언 불가능
L suga iSS	르수가 있	보조용언 가능
L tAi	르타이	보조용언 기대
LbbUNdAIA	르뿐더러	전성연결어미 접속형
Lgga	르까	문말어미 의문형
Lji	르지	전성연결어미 접속형
LjiANjAQ	르지언정	전성연결어미 접속형
LjiNdE	르진대	전성연결어미 접속형
LjiNjA	르진저	문말어미 서술형
Ljilado	르지라도	전성연결어미 접속형
LmaQjAQ	르망정	전성연결어미 접속형
Lse	르세	문말어미 서술형
LsegUlyA	르세그려	문말어미 감탄형
LsuloG	르수룩	전성연결어미 접속형
M	ㅁ	전성연결어미 명사형
M jiGha	ㅁ직하	보조용언 가치
Mma	ㅁ마	문말어미 서술형
Mse	ㅁ세	문말어미 서술형
N	ㄴ	전성연결어미 관형형
N ba iSS	ㄴ바 있	보조용언 결론
N cAGha	ㄴ착하	보조용언 허위
N ceha	ㄴ체하	보조용언 허위
N dUS ha	ㄴ듯 하	보조용언 추정
N gAS gaT	ㄴ것 같	보조용언 추정
N gASi	ㄴ것이	보조용언 결론
N pYANi	ㄴ편이	보조용언 경향
N seMi	ㄴ셈이	보조용언 결론
N tAi	ㄴ타이	보조용언 상황
NUNde	ㄴ는데	전성연결어미 접속형
NUNji	ㄴ는지	전성연결어미 접속형
NdUL	ㄴ들	전성연결어미 접속형
Nda	ㄴ다	문말어미 서술형
Nde	ㄴ데	전성연결어미 접속형
Nga	ㄴ가	문말어미 의문형
Nga bo	ㄴ가 보	보조용언 추정
NjUg	ㄴ죽	전성연결어미 접속형
Nji	ㄴ지	전성연결어미 접속형
SS	ㅍ	선어말어미 과거
SSASS	ㅍ엇	선어말어미 대과거
UBnidagUlyA	읍니다그려	문말어미 감탄형
UBsidayo	읍시다	문말어미 청유형
UBsyo	읍쇼	문말어미 명령형
UBsyogUlyA	읍쇼그려	문말어미 서술형

字素形態 (Grapheme Form)	表層形態 (Surface Form)	語彙情報
UL	을	목적격조사
UL	을	전성연결어미 관형형
UL bbANha	을 뻔하	보조용언 과거가능
UL bbuNi	을 뿐이	보조용언 강조
UL ddalUMi	을 따름이	보조용언 강조
UL gAS gaTUmyAN	을 것 같으면	전성연결어미 접속형
UL gASi	을 것이	보조용언 기대
UL jigyaQi	을 지경이	보조용언 위기
UL juL aL	을 줄 알	보조용언 가능
UL maNha	을 만하	보조용언 가치
UL moyaQi	을 모양이	보조용언 추정
UL su ABS	을 수 없	보조용언 불가능
UL su baGGe ABS	을 수 밖에 없	보조용언 당위
UL su iSS	을 수 있	보조용언 가능
UL suga ABS	을 수가 없	보조용언 불가능
UL suga iSS	을 수가 있	보조용언 가능
UL tAi	을 터이	보조용언 기대
ULbbUNdAlA	을뻔더러	전성연결어미 접속형
ULgga	을까	문말어미 의문형
ULji	을지	전성연결어미 접속형
ULjiANjaQ	을지언정	전성연결어미 접속형
ULjiNdE	을진대	전성연결어미 접속형
ULjiNja	을진저	문말어미 서술형
ULjilado	을지라도	전성연결어미 접속형
ULmaQjaQ	을망정	전성연결어미 접속형
ULsuloG	을수록	전성연결어미 접속형
UM	음	전성연결어미 명사형
UM jiGha	음 직하	보조용언 가치
UMma	음마	문말어미 서술형
UMse	음세	문말어미 서술형
UN	은	전성연결어미 관형형
UN	은	주제보조사
UN ba iSS	은 바 있	보조용언 결론
UN cAGha	은 척하	보조용언 허위
UN ceha	은 체하	보조용언 허위
UN dUS ha	은 듯 하	보조용언 추정
UN gAS gaT	은 것 같	보조용언 추정
UN gASi	은 것이	보조용언 결론
UN pYANi	은 편이	보조용언 경향
UN seMi	은 셈이	보조용언 결론
UN tAi	은 터이	보조용언 상황
UNdUL	은들	전성연결어미 접속형
UNde	은데	전성연결어미 접속형
UNga	은가	문말어미 의문형
UNga bo	은가 보	보조용언 추정
UNjUg	은즉	전성연결어미 접속형
UNji	은지	전성연결어미 접속형
UNkAnyAQ	은커녕	일반보조사

字素形態 (Grapheme Form)	表層形態 (Surface Form)	語彙情報
Udwi	으되	전성연결어미 접속형
Ui	으이	문말어미 서술형
UigUlyA	으이그러	문말어미 감탄형
UIA	으러	전성연결어미 접속형
UlaG	으락	전성연결어미 접속형
Uli	으리	문말어미 서술형
Ulida	으리다	문말어미 서술형
Uligga	으리까	문말어미 의문형
Ulila	으리라	문말어미 서술형
Ulo	으로	부사격조사
UlosA	으로서	부사격조사
UlyA	으려	전성연결어미 접속형
UlyAM	으렴	문말어미 서술형
UlyANmanUN	으런마는	전성연결어미 접속형
UlyAgo ha	으려고 하	보조용언 의도
UlyAmuna	으려무나	문말어미 서술형
UlyAniwa	으려니와	전성연결어미 접속형
Ulya	으랴	문말어미 의문형
UmE	으메	전성연결어미 접속형
UmUlo	으므로	전성연결어미 접속형
UmyA	으며	전성연결어미 접속형
UmyAN	으면	전성연결어미 접속형
UmyAN aNdwi	으면 안되	보조용언 금지
UmyAN dwi	으면 되	보조용언 허가
UmyAN ha	으면 하	보조용언 희망
UmyAN joH	으면 좋	보조용언 제안
UmyANsA	으면서	전성연결어미 접속형
Una	으나	전성연결어미 접속형
Unama	으나마	전성연결어미 접속형
Uni	으니	전성연결어미 접속형
Unigga	으니까	전성연결어미 접속형
Unya	으냐	문말어미 의문형
Uo	으오	문말어미 명령형
Uo	으오	문말어미 의문형
Uolida	으오리다	문말어미 서술형
Uoliida	으오리이다	문말어미 서술형
Usi	으시	존경선어말
UsiBsio	으십시오	문말어미 명령형
Usio	으시오	문말어미 명령형
UsosA	으소서	문말어미 명령형
a	아	부사격조사
a bAli	아 버리	보조용언 완결
a bo	아 보	보조용언 시도
a dE	아 대	보조용언 과잉
a dUL	아 들	보조용언 시작
a dUli	아 드리	보조용언 봉사
a du	아 두	보조용언 유지
a iSS	아 있	보조용언 완료
a ji	아 지	보조용언 피동

字素形態 (Grapheme Form)	表層形態 (Surface Form)	語彙情報
a ju	아 주	보조용언 봉사
a juGgeSS	아 죽겠	보조용언 욕망
a ka	아 가	보조용언 진행
a nE	아 내	보조용언 완결
a naga	아 나가	보조용언 진행
a noH	아 놓	보조용언 유지
a o	아 오	보조용언 진행
aP	앞	명사형접미사
aSS	았	선어말어미 과거
aSSASS	았었	선어말어미 대과거
aSSguna	았구나	문말어미 감탄형
ado	아도	전성연결어미 접속형
ado dwi	아도 되	보조용언 허가
ado joH	아도 좋	보조용언 허가
ala	아라	문말어미 명령형
asA	아서	전성연결어미 접속형
aya	아야	전성연결어미 접속형
aya dwi	아야 되	보조용언 요구
aya ha	아야 하	보조용언 요구
ayamaN	아야만	전성연결어미 접속형
baGGe	밖에	일반보조사
boda	보다	부사격조사
butA	부터	일반보조사
cA	처	명사형접미사
cAlAM	처럼	부사격조사
dAN	던	전성연결어미 관형형
dANdUL	던들	전성연결어미 접속형
dANga	던가	문말어미 의문형
dANji	던지	전성연결어미 접속형
dAlA	더러	부사격조사
dAla	더라	문말어미 서술형
dAlado	더라도	전성연결어미 접속형
dAnya	더냐	문말어미 의문형
dElo	대로	부사격조사
dUL	들	명사형접미사
dUNji	든지	일반보조사
dUNji	든지	전성연결어미 접속형
dUli	드리	동사파생접미사
da	다	문말어미 서술형
da	다	전성연결어미 접속형
daga	다가	전성연결어미 접속형
de	데	문말어미 서술형
di	디	전성연결어미 접속형
do	도	주제보조사
doloG	도록	전성연결어미 접속형
dwi	되	전성연결어미 접속형
dwi	되	동사파생접미사
e	에	접속격조사
ege	에게	부사격조사

字素形態 (Grapheme Form)	表層形態 (Surface Form)	語彙情報
egelo	에게로	부사격조사
egesA	에게서	부사격조사
esA	에서	부사격조사
gANmanUN	건마는	전성연결어미 접속형
gAdUN	거든	전성연결어미 접속형
gAnUL	거늘	전성연결어미 접속형
gAna	거나	전성연결어미 접속형
gAni	거니	전성연결어미 접속형
gAniwa	거니와	전성연결어미 접속형
gUlyA	그러	일반보조사
ga	가	부사격조사
ga	가	주격조사
gaN	간	명사형접미사
gaTi	같이	부사격조사
ge	게	문말어미 명령형
ge	게	전성연결어미 부사형
ge dwi	게 되	보조용언 변화
ge ha	게 하	보조용언 사동
geSS	겠	선어말어미 의지
geSSguna	겠구나	문말어미 감탄형
gena	게나	문말어미 서술형
ggaji	까지	일반보조사
gge	께	부사격조사
ggelo	께로	부사격조사
ggeoBsA	께읍서	주격조사
ggesA	께서	주격조사
gi	기	전성연결어미 명사형
gi bala	기 바라	보조용언 희망
gi ddEmuNi	기 때문이	보조용언 원인
gi malyANi	기 마련이	보조용언 확신
gilo dwi	기로 되	보조용언 약속
gilo ha	기로 하	보조용언 약속
go	고	부사격조사
go	고	전성연결어미 접속형
go	고	접속격조사
go iSS	고 있	보조용언 진행
go mal	고 말	보조용언 완결
go siP	고 싶	보조용언 희망
go siPA ha	고 싶어 하	보조용언 희망
goN ha	곤 하	보조용언 습관
goja	고자	전성연결어미 접속형
goja ha	고자 하	보조용언 의도
gosA	고서	전성연결어미 접속형
gulyA	구려	문말어미 감탄형
gulyA	구려	문말어미 서술형
guna	구나	문말어미 감탄형
gwAN	권	명사형접미사
gwa	과	접속격조사
ha	하	동사파생접미사

字素形態 (Grapheme Form)	表層形態 (Surface Form)	語彙情報
ha	하	형용사파생접미사
hago	하고	부사격조사
hago	하고	접속격조사
hamyA	하며	접속격조사
hante	한테	부사격조사
hantelo	한테로	부사격조사
hantesA	한테서	부사격조사
hi	히	부사파생접미사
ho	호	명사형접미사
i	이	문말어미 서술형
i	이	부사격조사
i	이	서술격조사
i	이	주격조사
iNdUL	인들	일반보조사
idUNji	이든지	일반보조사
igUlyA	이그러	문말어미 감탄형
igo	이고	접속격조사
ila	이라	부사격조사
ilaQ	이랑	접속격조사
ilado	이라도	일반보조사
ilago	이라고	부사격조사
imyA	이며	접속격조사
ina	이나	일반보조사
inama	이나마	일반보조사
isiyA	이시여	부사격조사
iyA	이여	부사격조사
iya	이야	일반보조사
iyamaLlo	이야말로	일반보조사
ja	자	문말어미 청유형
ja	자	전성연결어미 접속형
jamaja	자마자	전성연결어미 접속형
ji	지	문말어미 의문형
ji	지	문말어미 명령형
ji	지	문말어미 청유형
ji aNH	지 않	보조용언 부정
ji aNHUmyAN aNdwi	지 않으면 안되	보조용언 당위
ji aNHado joH	지 않아도 좋	보조용언 허가
ji maL	지 말	보조용언 부정
ji moSha	지 못하	보조용언 부정
jimanUN	지마는	전성연결어미 접속형
jiyo	지요	문말어미 의문형
jiyo	지요	문말어미 명령형
jiyo	지요	문말어미 청유형
jjUM	줍	명사형접미사
joca	조차	일반보조사
kAnyAQ	커녕	일반보조사
kwa	과	부사격조사
lA	러	전성연결어미 접속형
lAni	러니	전성연결어미 접속형

字素形態 (Grapheme Form)	表層形態 (Surface Form)	語彙情報
IUL	룰	목적격조사
la	라	부사격조사
laG	락	전성연결어미 접속형
laQ	랑	접속격조사
lado	라도	일반보조사
lago	라고	부사격조사
laya	라야	전성연결어미 접속형
layamaN	라야만	전성연결어미 접속형
li	리	문말어미 서술형
lida	리다	문말어미 서술형
ligga	리까	문말어미 의문형
lila	리라	문말어미 서술형
lo	로	부사격조사
loguna	로구나	문말어미 감탄형
losA	로서	부사격조사
lyA	려	전성연결어미 접속형
lyAM	렴	문말어미 서술형
lyANmanUN	련마는	전성연결어미 접속형
lyAgo ha	려고 하	보조용언 의도
lyAmuna	려무나	문말어미 서술형
lyAniwa	려니와	전성연결어미 접속형
lya	랴	문말어미 의문형
mE	매	전성연결어미 접속형
mUlo	므로	전성연결어미 접속형
maN	만	접속격조사
maN	만	주제보조사
maNkUM	만큼	부사격조사
mada	마다	일반보조사
majA	마저	일반보조사
manUN	마는	접속격조사
myA	며	전성연결어미 접속형
myA	며	접속격조사
myAN	면	전성연결어미 접속형
myAN aNdwi	면 안되	보조용언 금지
myAN dwi	면 되	보조용언 허가
myAN ha	면 하	보조용언 희망
myAN joH	면 좋	보조용언 제안
myANsA	면서	전성연결어미 접속형
nUN	는	전성연결어미 관형형
nUN	는	주제보조사
nUN gASi joH	는 것이 좋	보조용언 제안
nUNda	는다	문말어미 서술형
nUNga	는가	문말어미 의문형
nUNgulyA	는구려	문말어미 감탄형
nUNguna	는구나	문말어미 감탄형
nUNkAnyAQ	는커녕	일반보조사
nUNya	느냐	문말어미 의문형
na	나	문말어미 의문형
na	나	일반보조사

字素形態 (Grapheme Form)	表層形態 (Surface Form)	語彙情報
na	나	전성연결어미 접속형
naida	나이다	문말어미 서술형
naigga	나이가	문말어미 의문형
nama	나마	일반보조사
nama	나마	전성연결어미 접속형
ne	네	문말어미 서술형
negUlyA	네그려	문말어미 감탄형
ni	니	전성연결어미 접속형
nigga	니까	전성연결어미 접속형
nya	냐	문말어미 의문형
o	오	문말어미 명령형
o	오	문말어미 의문형
oLsida	올시다	문말어미 서술형
oida	오이다	문말어미 서술형
oigga	오이가	문말어미 의문형
olida	오리다	문말어미 서술형
oliida	오리이다	문말어미 서술형
sA	서	부사격조사
sAG	석	명사형접미사
sAN	선	명사형접미사
sAggAN	서건	일반보조사
sAnUN aNdwi	서는 안되	보조용언 금지
sUBdida	습니다	문말어미 서술형
sUBdigga	습디까	문말어미 의문형
sUBnida	습니다	문말어미 서술형
sUBnigga	습니까	문말어미 의문형
se	세	문말어미 청유형
si	시	존경선어말
siL	실	명사형접미사
siki	시키	동사파생접미사
sio	시오	문말어미 명령형
sipi	시피	접속격조사
siyA	시여	부사격조사
so	소	문말어미 서술형
so	소	문말어미 명령형
so	소	문말어미 의문형
sosA	소서	문말어미 명령형
ssi	씨	명사형접미사
wa	와	부사격조사
wa	와	접속격조사
yA	여	부사격조사
yAG	역	명사형접미사
yU	의	관형격조사
ya	야	부사격조사
ya	야	일반보조사
yamaLlo	야말로	일반보조사
yo	요	일반보조사
yo	요	전성연결어미 접속형
yogUlyA	요그려	문말어미 감탄형