

형태소 해석을 이용한 합성음성의 음운 및 운율처리

金尙龍, 金正壽
三星綜合技術院 基盤技術研究所

I. 서론

인간의 의사를 조음기관을 통하여 다른 이에 전달할 때 음절과 음절의 조음결합, 문장의 운율 뿐 아니라 그 문장의 구조에 알맞는 띄어 읽기가 필요하며 이는 음질의 자연성 향상에 크게 기여하고 있다.^[1]

현재까지 한국어 합성에 관한 많은 연구가 이루어져 왔고^[2], 그중에는 기능어(어미, 조사)를 이용한 문서음성 변환장치(Text to Speech)를 소개한 것도 있으나 인간의 자연적인 발성상태의 음성과는 그 성능에서 차이점을 드러내고 있다.^{[3] [4]}

본 논문에서는 한국어가 가지는 독특한 특징점을 고려한 음소와 음소의 천이구간처리, 문장의 운율 및 pause 처리 등을 기술하려 한다. 처리를 위해 우리는 한국어의 자연어처리 기술을 도입하고 있으며, 구문해석(syntactic analysis)시 야기되는 복잡성을 줄이기 위해 단순화한 처리방법을 사용한다.

언어처리의 결과로 한국어 문서(한글, 영어, 숫자, 특수기호가 혼용된 한글 주류의 문서)에 관한 언어코드(음운기호와 운율기호)를 얻을 수 있으며, 이를 한국어 규칙합성기에 적용시키고 있다.

Ⅱ장에서는 한국어 특성과 구조에 관하여 간단히 고찰하고, Ⅲ장에서는 한국어 문서의 형태소 및 구문해석 과정에 관하여 기술코자한다. 본 처리과정을 통해 얻어진 언어코드를 이용한 문서의 규칙합성 결과를 Ⅳ장에 기술하며, Ⅴ장에서는 결론과 향후 과제에 관하여 언급한다.

Ⅱ. 한국어 특성

한국어는 형태상 분류로는 첨가어, 계통상 분류로는 알타이 어족에 포함되며, 본 논문에 관련된 특성을 형태적, 구문적, 음운적으로 나누어 간략히 정리하였다.^[5]

1. 형태적 특징

한국어는 뜻을 나타내는 실질 형태소에 문법적 관계를 표시하는 형식 형태소가 붙는 첨가적 성질을 가진다. 체언에 조사가 연결되어 격변화하거나, 어간에 어미가 연결되어 문법적 관계를 나타내는 예를 볼 수 있으며 형태소의 범주는 그림1과 같다.

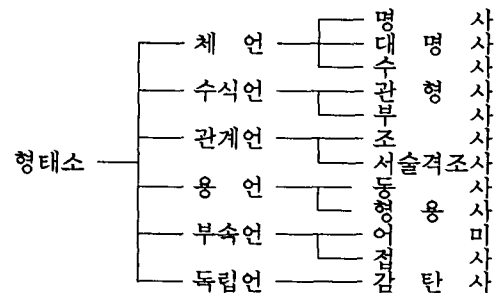


그림 1. 형태소 범주 분류

2. 구문적 특징

정상적인 문장의 구조는 '주어+목적어/보어+서술어'의 어순을 가진다. 그러나 대화체의 말에서는 주어 생략되는 경우가 많고, 수식어는 보통 수식을 받는 말 앞에 놓이며, 서술어는 동사, 형용사, '체언+이다'의 형태를 취한다.

이러한 구문적 특성을 규칙화 하기 위해 한국어의

구문요소(문장성분)를 그림 2와 같이 분류한다.

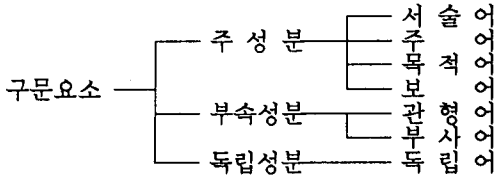


그림 2. 구문요소 범주 분류

3. 음운적 특징

한국어 음운은 그림 3과 같이 분류할 수 있으며 음운변화는 자음과 자음이 연결되는 부분에서 새로운 천이음을 생성하는데, 본 논문에서는 자음동화, 구개음화, 음운첨가, 경음화 및 음의 장단화를 고려하고 있다.

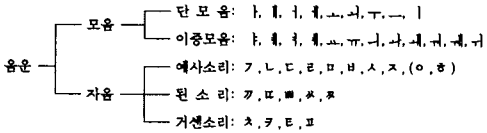


그림 3. 한국어 음운

문서음성합성시 자음의 'ㅇ' 음은 모음부에 흡수된 형태로 해석하며, 'ㅎ' 음은 문맥에 따라 그 형태가 7가지 이상으로 구별되나, 경우에 따라 'ㅇ' 음과 같이 처리되기도 하고 예사소리로 처리되기도 한다. 기타 자음의 경우 원칙적으로 초성에서는 무성음화 되어 있으나 발음되는 시점의 앞뒤 연결음에 따라 유성화의 정도가 크게 다르다. 보통의 경우 'ㄴ', 'ㄹ', 'ㄹ' 등은 다른 자음과 달리 유성자음으로 분류하고 있다.

III. 한국어 문서의 해석

1. 개요

한국어 문서해석은 그림 4와 같이 전처리, 형태소 해석, 구문해석, 언어코드 생성의 4부분으로 구성하며, 입력문서는 한글·한자코드^[6]와 ASCII코드로 구성된 한국어 문서로 가정하였다.

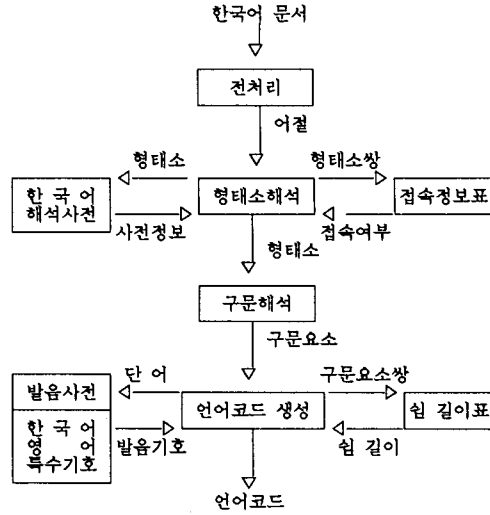


그림 4. 문서 해석부

문서해석시 필요한 한국어 해석사전은 형태소 표제어, 예외발음 여부 및 장모음 위치, 형태소 범주 코드 등의 정보를 지니고 있고, 형태소 해석사 전 정보를 제공하며, 접속정보는 최장일치법에서 사용한다. 그림4에 나타난 각각의 block은 이어지는 절에서 자세히 설명코자 한다.

2. 전처리

전처리는 한국어 문서를 입력으로 받아 어절 단위로 분리하는 일을 수행한다. 여기서 어절이란 띄어쓰기가 이루어지는 기본단위를 말한다. 전처리에는 한글·한자 변환, 어절 분리, 그리고 문장 분리 등의 처리가 포함된다. 한자 코드를 한글 코드로 변환하는 방법은 한글에 해당하는 한자의 첫번째 코드와 연속되는 갯수가 들어 있는 한자·한글 변환표를 이용하는 것이다.

어절 분리는 문장을 형태소해석 단위인 어절 단위로 분리하는 것을 말하며, 구성 문자에 따라 한글(K), 영어(E), 숫자(N), 그리고 특수기호(S) 어절로 분류한다. 어절 분리는 Finite State Automata에 의해 처리된다.

문장 분리는 문서를 문장 단위로 분리하는 것이다. 문장을 분리하는 단서는 소수점이 아닌 마침표('.'), 물음표('?'), 느낌표('!') 등을 문장의 끝으로 보는 것이다. 또한 공란(space line)도 문장의 끝으로 보는데, 이것은 제목이나 표제 다음에는 일반적으로 공

란을 두기 때문이다. 국민교육헌장을 입력할 경우의 전처리 결과의 예가 그림 5에 있다.

어 절	Class
우리는	K
민족중	K
의	K
역사적	K
사명을	K
피고	K
이	K
땅에	K
태어났다	K
.	S

그림 5. 전처리의 예

3. 형태소해석

문장을 분석하여 끊어 읽기 단위로 재구성하는 일은 형태소 해석부터 시작된다. 한국어 형태소 해석이란 띄어쓰기의 단위인 어절의 구성요소(형태소)를 밝히는 것이다. 본 논문에 구현된 형태소 해석부를 설명하기 위해, 한국어 형태소 분류, 해석사전 구조, 형태소해석의 중심인 형태소 조합찾기와 최장일치를 위한 접속정보표에 관하여 기술한다.

1) 형태소 분류

형태소는 일정한 음성에 일정한 뜻이 결합되어 있는 말의 가장 작은 단위, 즉 최소의 유의적 단위로 정의되며, 한국어의 경우 그림 1과 같이 대분류된다. 처리 관점에 따라 각기 다르나, 본 논문에서는 대분류된 12가지의 품사를 기초하여 60여가지로 부록 A와 같이 상세분류하였다.

2) 해석사전 구조

형태소 분석은 빈번한 사전 검색이 요구된다. 이를 위해 사전은 검색시간이 최소화 되도록 구성되는 것이 바람직하다. 이를 위해 해석사전을 글자수별로 구성하며, 각각을 Trie로 구성하여 탐색시간을 최소화 하였다.^[7]

본 논문에서 구축된 해석사전의 규모는 약 88,000 단어이며, 음운기호 생성 단계에서 사용되는 한국어 발음사전은 약 5,000 단어, 영어 발음사전은 약 5,500 단어로 구성되었다.

3) 형태소 조합 찾기

어절은 문장을 구성하는 도막도막의 마디로 의미적 구성 단위이며, 문장 성분의 최소단위이다. 어절은

형태소가 모여 이루어지며 본 논문에서는 형태소 해석의 정확도를 높이기 위해, 그 어절을 구성할 수 있는 모든 가능한 형태소 조합을 찾는다. 형태소 조합 찾기의 결과는 트리형태가 된다. 예를 들어, 어절 '우리는'에 대한 가능한 형태소 조합은 그림 6과 같다.

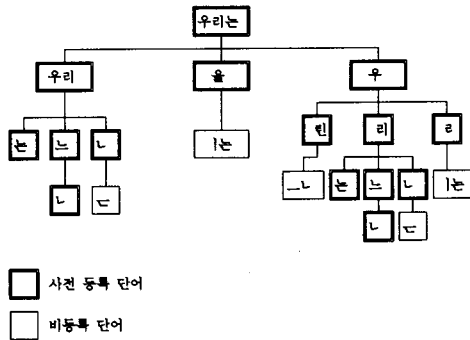


그림 6. 형태소 분석트리

형태소 분석 트리를 구성하기 위해서는 주어진 어절의 첫 글자에서 시작하여 사전을 탐색하며 등록된 형태소가 발견되면 형태소 분석트리의 노드를 추가하는 것을 반복적으로 행한다

형태소 분석트리로부터 올바른 형태소 리스트를 찾기 위해 형태소 분석트리의 루트(root) 노드로부터 하위 노드와의 접속가능성을 접속정보표를 이용하여 조사하면서 단말 노드까지 찾아가면 된다. 단말 노드가 미등록인 경우는 형태소 리스트에서 제외하며, 모든 단말노드가 미등록인 경우, 즉 형태소해석에 실패하는 경우 복합명사로 간주한다.

4) 최장일치를 위한 접속정보표

Depth first 탐색 알고리즘과 유사한 최장일치법은 한국어 형태소해석시 분석 오류 확률이 적은 것으로 알려져 있다.^[8] 특히 음성합성시 요구되는 음운 및 운율정보 추출에는 접속정보표를 이용하는 최장일치법이 실험결과 만족한 결과를 보였으며, 해석 문장의 형태소 리스트의 단순화가 가능하였다. 알고리즘 구현시 탐색의 순서는 단말 노드로부터 수행하는 것이 수행시간상 이점이 있음을 알 수 있었다.

최장일치법까지 마친 형태소 해석의 결과가 입력문장이 국민교육헌장의 첫 문장인 경우를 예로 그림 7에 주어져 있다.

형태소	범주
우리 는 민족중흥 의 역사적 사명을 띠고 이땅에 태어나 다	인칭대명사 동용보조사 보통명사 관형사 상관형사 목적격조사 동사 대동사 지시관형사 보통명사 부사 동사 선어말어미 평서형종결어미

그림 7. 형태소 해석 예

4. 구문해석

인간이 긴 문장을 적절히 끊어 읽는 것은 문장의 구문구조를 분석해 낼 수 있기 때문이다. 구문해석에서는 형태소 해석의 결과인 문장성분의 가장 작은 단위인 어절이 어떠한 구문요소로 쓰여졌는가를 토대로 구문구조를 분석한다.

본 절에서는 구문해석 방법을 구문요소의 분류, 홀문장화, 격해석, 독립어처리의 순으로 기술한다.

1) 구문요소의 상세 분류

어절은 문장성분의 가장 작은 단위로서 문장에서의 쓰임새에 따라 그림 2와 같이 분류된다. 본 논문에서는 통용보조사나 접속조사로 끝나는 어절인 경우 각각 통용어, 접속어로 분류하였고, 어절의 구성내역에 따라 상세분류하여 19가지로 나누었다. 구문요소의 상세분류는 부록B에 보였다.

2) 홀문장화

한국어 문장은 크게 홀문장과 겹문장으로 나뉘어진다. 홀문장은 주어와 서술어의 관계가 한번 이루어져 있는 문장이고, 겹문장은 홀문장이 서로 이어지거나 여러 겹으로 된 문장이다. 홀문장화는 홀문장이 서로 이어져 있는 겹문장을 홀문장으로 분리하는 것을 말한다.

3) 격해석

격해석에서는 각 어절에 대하여 표층격을 결정한 다. 각 어절의 표층격은 마지막 형태소의 범주에 따라 결정된다. 예를 들어 어절 '우리는'의 경우, 마지막 형태소 '는'의 형태소 범주가 통용보조사이므로

통용격으로 결정한다.

4) 독립어 처리

독립어는 주성분이나 부속성분과 직접적인 관계가 없이 운율이 주어지며, 문장 전체에 영향을 주는 성분이다. 본 논문에서는 감탄사, 문장부사, 호격조사로 끝나는 어절 및 순서를 표시하는 글자 등을 독립어로 간주한다.

독립어 처리까지 마친 구문해석의 결과가 입력문장이 그림7과 같은 경우, 그림8과 같이 주어진다.

어절	범주
우리는	통용어
민족중흥의	관형어
역사적	관형어
사명을	목적어
띠고	서술어
이	관형어
땅에	부사어
태어났다	서술어

그림 8. 구문해석 예

5. 음운기호 생성

언어코드 생성부에서는 구문해석에서 얻어진 구문요소에 대하여 음운기호 생성을 실시한다. 음운기호 생성에서는 한글, 영어, 숫자, 그리고 특수기호에 대하여 각각의 규칙에 따라 음운기호를 생성한다.

1) 한글 읽기

한글 읽기는 표기된 기호와 발생되는 기호의 상이 함으로부터 기인되는 발음법의 문제이다. 즉 단어를 구성하는 자소는 그들의 상호관계에 의하여 영향을 받아 인간이 발생하기에 적합한 음으로 바뀌는 성질이 있다. 자소간의 관계에 의하여 발생하는 음운규칙은 한국어 표준발음법에 제정되어 있다.^[6]

음운변동은 규칙적으로 처리되는 경우가 대부분이지만 약 6 ~ 7%의 불규칙 변동이 일어남을 알 수 있었다. 규칙 변동인 경우는 발음 변환표와 표준발음법 규칙에 따라 처리하여 음운기호를 생성하고, 불규칙인 경우는 한국어 발음사전에서 탐색하여 음운기호를 생성한다. 특히 동일철자의 단어 중 서로 다르게 발음되는 경우는 발견적 해결 방법 (heuristic approach)을 사용하였다.

2) 영어 읽기

영어의 경우에 있어서, 사전에 존재하는 영어 단어

는 등록되어 있는 발음대로 읽고, 그렇지 않은 경우는 각각의 철자를 한자씩 읽는다.

3) 숫자 읽기

숫자는 문장 중에 '숫자 + 단위성 의존명사', '숫자.숫자', '숫자.….숫자'와 같은 형태로 나타나며 그 외에 '8.15' 등과 같은 관용적인 표현을 쓸 때도 있다. 각각의 경우에 읽는 예를 표 1에 나타내었다.

표 1. 숫자 읽기의 예

숫자 + 의존명사	3명 (세명), 3인 (삼인), 3남 (서남)
숫자.숫자	7.638 (칠백육삼팔)
숫자	3429 (삼천사백이십구), 3.429 (삼천사백이십구)
숫자.….숫자	36, 26, 726, 12 (삼십육 이십육 칠백이십육 십이)
관용	8.15 (팔일오), 6.25 (육이오), 5대 (오대 또는 다섯대), 12:12 (열두시 십이분)

4) 특수기호 읽기

사전에 존재하는 특수기호는 등록되어 있는 발음대로, 그렇지 않은 경우는 읽지 않는다. 일본어 문자의 경우도 특수기호로 취급하여 사전에 등록된 대로 읽는다.

6. 운율기호 생성

운율기호 생성은 구문요소간의 연결에 따라 쉽 길이와 억양패턴을 생성한다.

1) 쉽 길이를 결정

문장의 의미를 잘 전달하기 위해서는 적절히 끊어 읽어야 한다. 이것에 해당하는 것이 쉽 길이를 결정이다. 본 시스템에서는 표 2와 같이 6 단계의 경우로 나누어 쉽 길이를 결정하였다.

표 2. 쉽 길이 및 의미

쉽 길이	의미
5	문장간 문장간 독립어 뒤 숨을 쉬기 위한 쉽 의미 전달을 위한 쉽 쉽 없음
4	
3	
2	
1	
0	

쉽 길이를 결정하는 과정은 다음과 같다. 먼저, 구문요소 사이의 구문적 결합도를 이용하여 구문요소간 쉽 길이를 결정한다. 이 단계에서 쉽 길이 5, 4, 3이

결정된다. 다음으로, 끊어 읽기 단위가 12개 이상의 모음을 갖고 있는 관형어이고 앞뒤의 단위가 3개 이상의 모음을 갖고 있으면, 쉽 길이를 2로 조정한다. 또한, 끊어 읽기 단위가 12개 이상의 모음을 갖고 있는 명사이고 앞뒤의 단위가 3개 이상의 모음을 갖고 있으면, 쉽 길이를 1로 조정한다.

2) 운율기호 생성

본 논문에서는 운율패턴을 발화 단위의 유성음 수와 연관관계에 따라 정의하며, 이러한 결정 규칙을 바탕으로 운율기호 생성표를 작성하여 어절 단위로 처리하였다. 표 3은 한국어 음성의 운율 통계를 기초로 작성된 것이며 유성음수가 4개까지를 정리하였다.

표 3. 운율 패턴

유성음수 \ 유형	1	2	3	4
TYPE 0				
TYPE 1				
TYPE 2				
TYPE 3				

IV. 실험 및 고찰

본 논문에서 제안한 알고리즘의 성능을 평가하기 위하여 '국민교육헌장'과 논설문 등을 입력문서로 하여 각 단계별 정확도를 측정하였다. 이 결과를 표 4에 정리하였다. 측정 단위는 어절로 하였고 전체 어절수에 대한 실패 어절수의 비를 기록하였다.

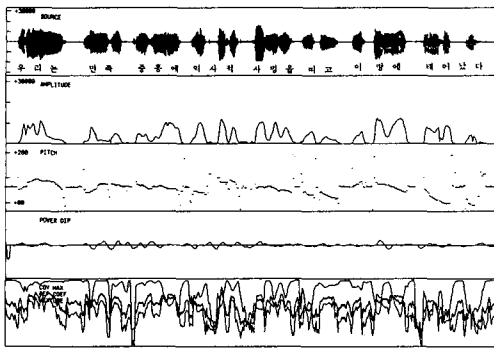
표 4. 시스템 성능 평가

입력문서	형태소해석	구문해석	음운기호	쉽 길이	운율패턴
국민교육헌장	1/145	3/145	0/145	8/145	7/145
국어교과서	11/557	26/557	0/557	36/557	28/557
국사교과서	13/632	29/632	0/632	38/632	35/632
논설	15/598	34/598	2/598	39/598	35/598
신문기사	12/348	22/348	2/348	36/348	33/348
성능	97.7%	95.0%	99.8%	93.1%	94.0%

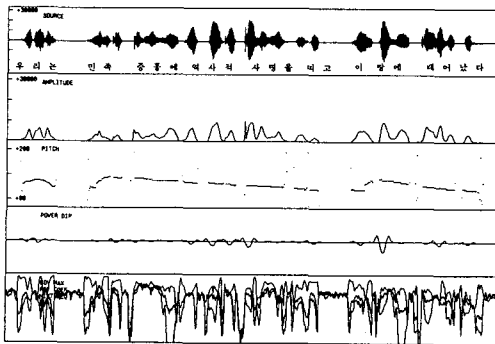
총 172 문장, 2280 어절에 대한 형태소해석과 구문해석의 성공률은 각각 97.7%, 95.0%로 나타났다. 오류의 주된 원인은 하나의 어절이 여러가지로

분석되는, 즉 모호성에 인한 것이다. 이러한 모호성을 줄이기 위해서는 문맥정보를 이용하는 해석법을 추가하는 것이 바람직하리라 본다.

음운기호 생성의 성공률이 99.8%로 나타났다. 실패의 원인은 숫자의 읽기에서 한자어로 읽느냐 고유어로 읽느냐에 실패한 경우이다. 이러한 실패를 없애기 위해서는 숫자 읽기에 휴리스틱 정보가 더 추가되어야 할 것이다.



(a) 원음



(b) 합성음

그림 9. 원음과 합성음의 비교

음운기호 즉 쉼 길이와 운율 패턴 생성의 성공률은 각각 93.1%, 94.0%로 나타났다. 그 기준은 사람이 낭독한 문장을 앞장에서 정의한 쉼 길이와 운율패턴으로 분별한 다음 본 시스템이 생성한 결과와 비교한 것이다. 부정확한 쉼 길이는 음성의 자연성에 치명적

인 영향을 미치므로, 정확도를 높이기 위해서는 인간의 발성 습관에 대한 휴리스틱도 추가되어야 할 것이다. 그림9는 형태소해석과 구문해석이 완료된 후 그로부터 운율패턴과 쉼 길이 정보를 가지고 만들어진 음성 파형과 원음의 차이를 보인 것이다.

V. 결론

본 연구는 무제한 음성합성에 필요한 기술 중 문서 해석에 관련된 형태소해석, 구문해석, 그리고 언어코드 생성에 관한 문제를 다루었다. 본 논문은 형태소해석과 구문해석을 통하여 문장의 구조를 파악하고, 적절한 끊어 읽기와 자연스러운 운율 정보의 생성이 가능하도록 하였다.

이 과정에서 88,000 단어의 한국어 해석사전을 구축하였다. 언어 코드중 음운기호를 정확히 생성하기 위해 한국어, 영어, 특수기호 발음사전을 구축하였다. 또한, 인간의 발성과 유사한 음운기호를 얻기 위해 독창적으로 설계된 쉼 길이 결정 알고리즘과 운율패턴 생성 알고리즘은 IV장에서 보인 바와 같이 우수한 것으로 평가된다.

향후 과제로는 구문해석부를 강화하여 문장의 계층구조, 단어간 의미 결합구조 파악을 통한 운율기호 생성으로 인간의 발성과 좀더 유사하도록 하는 것이다. 또한, 발음사전내에 단어의 액센트 정보를 이용할 수 있도록 사전정보를 보완하여야 할 것이다.

參考文獻

- [1] Y.Sagisaka, "Speech Synthesis from Text," *IEEE Communication Magazine*, pp. 35-41, Jan. 1990.
- [2] 구준모 외5인, "한국어 무제한 음성합성 시스템: 가라사대," 음성통신 및 신호처리 워크샵 논문집, p.201, 1992.
- [3] 최운천 외4인, "고품질의 한국어 문장음성변환 시스템: 글소리II," 음성통신 및 신호처리 워크샵 논문집, pp.193-196, 1992.
- [4] 김형욱 외3인, "한국어 문장-음성 변환기의 운율 제어용 구문분석기," 음성통신 및 신호처리

워크샵 논문집, pp.218-223, 1992.
 [5] 조규빈, 하이라이트 교교문법, 지학사, 서울, pp.16-18, 1990.
 [6] 한국공업진흥협회, 한국공업규격 KSC5601, 1987.
 [7] E.Horowitz and S.Shani, Fundamentals of Data Structures in Pascal,

Computer Science Press, pp.512-520, 1983.
 [8] 김효준 외2인, "이식성이 있는 한국어 형태소 해석기 개발 연구," 87 특정연구과제 결과 발표 논문집, 1988
 [9] 문교부, "표준발음법," 표준어 규정, 삼성이데아, pp.64-70, 1988. ㉸

부록 A. 형태소 범주표

품 사		코드	보 기	
명	고 유 명 사	1	김유신, 두만강, 대한민국	
	보 통 명 사	동작성	2	독립, 희망, 입학, 탐류
		비동작성	3	사람, 철수, 나무, 하늘
	의 존 명 사	보편성	4	분, 것, 데, 바
		주어성	5	지, 수, 나휘, 리
		서술성	6	따름, 터
		부사성	7	대도, 채, 만, 채, 만큼, 줄
	단위성 명 사	한자어	8	원, 미터
		고유어	9	채, 개, 분, 마리, 섬, 자루
	(중략)			
접두사	관형사성	54	것-, 만-	
	부사성	55	및-, 및-, 절-, 새-, 첫-	
	접미사	한정적	56	-꾼, -꾸러기, -님, -디달-
		명사화접미사	57	-(으)마, -이, -기, -에
		동사화접미사	58	-거리, -이-, -히-
		형용사화접미사	59	-업/업-, -브-, -브-
부사화접미사	60	-껏-, -내-, -로-, -이-, -히		

부록 B. 구분요소 범주표

구 분 범 주		코드	보 기
서술어	체언 + 조사	1	시합이다. 꽃이로구나
	용언 + 어미	2	했다. 사랑한다. 아름답다
주 어	체언 + 조사	3	내가, 사람이
	용언 + 어미 + 조사	4	하기가
목적어	체언 + 조사	5	너를, 받을
	용언 + 어미 + 조사	6	함을
보 어	체언 + 조사	7	사람이
	용언 + 어미 + 조사	8	아름다움이
종용어	체언 + 조사	9	우리는
	용언 + 어미 + 조사	10	사랑함은
접속어	체언 + 조사	11	한과
	용언 + 어미	12	사랑하고
관형어	관형사	13	이, 새, 한
	체언 + (조사)	14	나의, 삼의
	용언 + 어미	15	사랑함의
부사어	부사	16	참, 내일
	체언 + 조사	17	은행에, 국민으로서
독립어	용언 + 어미	18	사랑함으로써
	감탄사, 문장부사, 접속부사	19	그러나, 아이고, 그래

筆者紹介



金 尙 龍

1957年 8月 24日生

1980年 2月 한국항공대학 전자과 졸업(학사)

1982年 2月 한국과학기술원 전기 및 전자과 졸업(석사)

1989年 2月 한국과학기술원 전기 및 전자과 졸업(박사)

1980年 1月 ~ 1993年 1月 삼성전자 근무

1993年 2月 ~ 현재 삼성종합기술원 근무

주관심분야: 음성처리(인식, 합성), 기계번역, DSP, Queueing Network analysis, ISDN



金 正 壽

1964年 5月 18日生

1988年 2月 연세대학교 전산학과(학사)

1990年 2月 한국과학기술원 전산학과(석사)

1990年 1月 ~ 1993年 1月 삼성전자 근무

1993年 2月 ~ 현재 삼성종합기술원 근무

주관심분야: 자연어 처리, 음성합성 및 인식, 컴퓨터 그래픽스