

TTA Standard

정보통신 단체표준(국문표준)

TTAK.KO-11.0010/R1

개정일: 2015년 06월 23일

형태소 태깅 말뭉치 작성용

품사 태그 세트

Part-of-Speech Tag Set for
Morphological Annotation of Written Texts



한국정보통신기술협회
Telecommunications Technology Association

형태소 태깅 말뭉치 작성용

품사 태그 세트

Part-of-Speech Tag Set for
Morphological Annotation of Written Texts



본 문서에 대한 저작권은 TTA 에 있으며, TTA 와 사전 협의 없이 이 문서의 전체 또는 일부를 상업적 목적으로 복제 또는 배포해서는 안 됩니다.

Copyright© Telecommunications Technology Association 2015. All Rights Reserved.

서 문

1. 표준의 목적

대규모 말뭉치(corpus)는 자연어 처리 분야에서 가장 기초적인 언어 자원으로 그 중요성이 높아지고 있다. 특히 말뭉치를 분석하여 가공한 형태소 품사 부착 말뭉치(tagged corpus)는 형태소 분석기 개발에 있어서 필수적인 언어 자원으로 그 구축과 관리에 많은 비용이 소요되므로 연구 개발자간에 공유가 가능해야 하고, 기존에 구축되어 있는 것을 재사용할 수 있어야 한다. 본 표준은 품사 부착 말뭉치의 재사용 및 공유를 가능하게 하는 표준 태그 세트를 제공하기 위하여 작성되었다.

2. 주요 내용 요약

1999 년 표준 태그 세트는 일반적으로 많이 사용되는 형태소 품사 위주로 작성되었다. 하지만, 최근에는 구문 분석이나 의미역 인식 등의 상위 수준의 언어 분석에 대한 수요가 늘어나면서 형태소 품사에 대한 세분화 요구가 늘어나고 있다. 그러므로, 본 표준에서는 1999 년 표준 태그 세트보다 조사, 어미, 기호 등을 세분화해서 정리하였다. 본 표준은 14 개 기관에서 사용하는 태그 세트의 비교 과정을 거쳤다[1-10]. 1999 년에 제정된 기존 표준 태그 세트는 27 개 품사로 구성된 반면, 본 표준은 47 개 품사로 세분화되었다.

3. 표준 적용 산업 분야 및 산업에 미치는 영향

언어 분석은 정보 추출, 정보 검색, 질의 응답 등의 자연어 처리 분야에서 기반이 되는 기술이며, 본 표준은 언어 분석 과정 중 형태소 분석에 대한 것이다. 형태소 분석은 언어 분석 과정 중에도 가장 기반이 되는 중요한 기술인데, 정보 추출 등의 응용분야 기술은 형태소 분석의 태그 세트에 매우 의존적이므로 응용 분야의 필요에 부합되도록 형태소 태그 세트를 확장/정제하고 품사 부착 말뭉치 작성용 표준안을 마련할 필요가 있다. 응용 분야의 필요에 적합한 품사 태그 세트와 말뭉치 구축 표준안이 나오면, 표준안에 따르는 형태소 품사 부착 말뭉치가 다양한 기관에서 만들어지고 공유되면서 형태소 분석기 개발 생태계가 자연스럽게 조성되고, 나아가서는 형태소 분석기를 활용하는 응용 분야의 기술 발전에도 기여할 것이다.

4. 참조 표준(권고)

4.1. 국외 표준(권고)

- 해당 사항 없음.

4.2. 국내 표준

- 해당 사항 없음.

5. 참조 표준(권고)과의 비교

5.1. 참조 표준(권고)과의 관련성

- 해당 사항 없음.

5.2. 참조한 표준(권고)과 본 표준의 비교표

- 해당 사항 없음.

6. 지식 재산권 관련 사항

본 표준의 ‘지식 재산권 요약서’ 제출 현황은 TTA 웹사이트에서 확인할 수 있다.

※본 표준을 이용하는 자는 이용함에 있어 지식 재산권이 포함되어 있을 수 있으므로, 확인 후 이용한다.

※본 표준과 관련하여 접수된 요약서 이외에도 지식 재산권이 존재할 수 있다.

7. 시험 인증 관련 사항

7.1. 시험 인증 대상 여부

- 해당 사항 없음.

7.2. 시험 표준 제정 현황

- 해당 사항 없음.

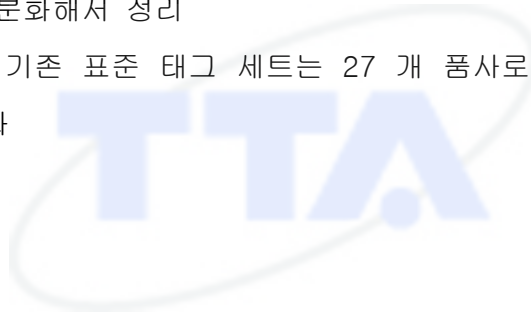
8. 표준의 이력 정보

8.1. 표준의 이력

판수	제정·개정일	제정·개정 내역
제 1 판	1999.12.08.	제정 TTAK.KO-11.0010
제 2 판	2015.06.23.	개정 TTAK.KO-11.0010/R1

8.2. 주요 개정 사항

- 2000 년부터 2014 년까지 발표된 주요 기관의 형태소 태그 세트를 비교 분석
- 최근의 형태소 품사에 대한 세분화 요구에 따라 1999 년 표준 태그 세트보다 조사, 어미, 기호 등을 세분화해서 정리
- 1999 년에 제정된 기존 표준 태그 세트는 27 개 품사로 구성된 반면, 본 표준은 47 개 품사로 세분화



Preface

1. Purpose of the Standard

One of the most essential language resources, large-scale corpora becomes more important in the Natural Language Processing research and application area. Specially tagged corpora obtained by attaching tags to the lexical items in the raw corpora are basic language resources for morphological analysis and require tremendous cost for its development and management. As a result, it is desirable that tagged corpora should be reusable and sharable among the research community. The purpose of this recommendatory standard is to provide commonly usable guidelines for the Part-of-Speech tag set for reuse and sharing of tagged corpus.

2. Summary of Contents

The original version of 1999 mainly defines the part of speech tag set used for general purpose. The needs for more detailed categorization of part of speech tags are increasing recently as the syntactic and semantic analysis are required for a high-level language analysis. Therefore, the proposed standard provides more detailed sub-categorization of auxiliaries, word endings, and symbols by revision of the original version of the part of speech tag set.

To achieve interoperability with the existing tag sets, the proposed tag set was produced by comparing with other 14 different tag sets [1-10].

While the original standard tag set established in 1999 consists of 27 speech tags, the proposed standard tag set in this document extends to 47 tags.

3. Applicable Fields of Industry and its Effect

Language analysis is the essential technology in the Natural language processing areas for information retrieval, information search and Question Answering. The proposed standard is related with the technology of morphological analysis, the core technology among various language analysis technologies. Since the application technologies such as information retrieval are heavily dependent on the POS tag sets, it is required to extend and refine the POS tag set and to establish a standard to construct tagged

corpus satisfying the needs of the applications. The standardized POS tag set and tagged corpus guidelines will contribute to the development of the eco system of the morphological analysis and technology enhancement of various application areas by producing and sharing POS tagged corpus in many different institutes.

4. Reference Standards(Recommendations)

4.1. International Standards(Recommendations)

- None

4.2. Domestic Standards

- None

5. Relationship to Reference Standards(Recommendations)

5.1. Relationship of Reference Standards(Recommendations)

- None

5.2. Differences between Reference Standard(Recommendation) and this Standard

- None

6. Statement of Intellectual Property Rights

IPRs related to the present document may have been declared to TTA. The information pertaining to these IPRs, if any, is available on the TTA Website.

No guarantee can be given as to the existence of other IPRs not referenced on the TTA website.

And, please make sure to check before applying the standard.

7. Statement of Testing and Certification

7.1. Object of Testing and Certification

- None

7.2. Standards of Testing and Certification

- None

8. History of Standard

8.1. Change History

Edition	Issued date	Outline
First Edition	1999.12.08.	Established TTAK.KO-11.0010
Second Edition	2015.06.23.	Revision TTAK.KO-11.0010/R1

8.2. Revisions

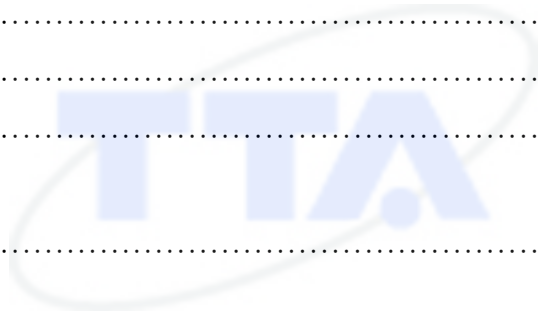
- Comparison of the different tag sets published by major institutes from 2000 to 2014
- Sub-division of the POS tags such as auxiliaries, word endings and symbols according to the recent needs of the sub-division
- The revised tag set consists of 47 POS tags while the existing tag set consists of 27 tags

목 차

1. 개요	1
2. 표준의 구성 및 범위	1
3. 참조 표준(권고)	2
4. 용어 정의 및 약어	3
5. 품사 태그 세트 분류 기준	5
5.1. 태그 세트 분류 체계	5
5.2. 각 태그별 정의 및 분류 기준	7
6. 기타	44
6.1. 준말	44
6.2. 합성어	44
부록 . 참고 문헌	45

Contents

1. Introduction	1
2. Constitution and Scope	1
3. Reference Standards(Recommendations)	2
4. Terms and Definitions	3
5. POS Tag Classification Criteria	5
5.1. POS tag Classification	5
5.2. Definition and Classification Criteria	7
6. Other information	44
6.1. Abbreviation	44
6.2. Compound	44
Appendix . Reference	45



형태소 태깅 말뭉치 작성용 품사 태그 세트

(Part-of-Speech Tag Set for Morphological Annotation of Written Texts)

1. 개요

대규모 말뭉치(corpus)는 자연어 처리 분야에서 가장 기초적인 언어 자원으로 그 중요성이 높아지고 있다. 특히 말뭉치를 분석하여 가공한 형태소 품사 부착 말뭉치(POS tagged corpus)는 형태소 분석 시스템 개발에 있어서 필수적인 언어 자원이다.

품사 부착 말뭉치는 그 중요성이 큰 만큼 구축과 관리에 많은 비용이 소요되므로 연구 개발자 간에 공유가 가능해야 하고, 기존에 구축되어 있는 것을 재사용할 수 있어야 한다. 품사 부착 말뭉치의 공유와 재사용이 가능하기 위해서는 품사 부착 말뭉치에 쓰이는 품사 태그 세트의 표준화가 선행되어야 한다.

표준 태그 세트가 관련 연구 개발자 누구나 수용할 수 있는 체계로 되어 있어서 품사 부착 말뭉치가 표준 태그 세트로 구축된다면 각각이 서로 다른 형태의 품사 부착 말뭉치를 소규모로 구축하던 것을 한 곳에 집중시킴으로써 대규모, 고품질의 언어 자원 확보가 가능해질 것이다.

2. 표준의 구성 및 범위

본 표준은 말뭉치를 분석하여 형태소 단위로 분할할 때, 각 형태소에 부착되는 태그 세트에 대한 정의와 분류 기준을 제공한다. 전체 태그 세트 분류 체계를 제시하고 각 태그에 대한 정의한 분류 기준을 기술한다.

본 표준에서 제시하는 태그 세트는 전산처리적 관점에서 단위와 범주를 설정하여 국어학이나 언어학에서 연구하고 있는 품사 분류 체계와는 그 목적과 용도가 분명히 다르다. 본 표준은 국어학이나 언어학에서 연구한 결과를 가능한 많이 참고하지만, 실제 실행되는 시스템들의 일반적인 환경적 제약과 개발에 드는 노력 및 자원의 제한을 고려한 전산 처리의 효율성을 고려하는 관점에서 분류 체계가 설정되었다. 각 태그별로 제시되는 예는 해당 태그에 속하는 전체 어휘가 아니라 대표적 어휘를 예시한 것이다.

본 표준에서 제시하는 품사 태그 세트의 가장 주된 용도는 자동 품사 태거를 위한 학습자료로 활용되는 품사 부착 말뭉치 작성용이다. 이러한 용도는 의미론적인 어휘 분류나 구문론적 또는 형태론적인 어휘 범주를 연구하는 사람들의 분류 체계와 관점이 다르다는 것을 전제하는 것이다.

언어 분석 과정 중에는 문장 분석 과정으로 형태소 분석이 끝나고 구문 분석을 하게 된다. 구문 분석에서의 애매성 폭발을 줄이기 위하여 품사 태거는 형태소분석 결과인 어절별 분석 결과들을 순위화(ranking)하여 상위 1 개의 분석 결과만 구문 분석의 입력으로 넘겨주기 위한 용도로 자동 품사 태거가 사용됨을 자연언어처리 분야의 연구자들은 주지하고 있을 것이다.

따라서, 태깅용 또는 태그 부착 말뭉치 작성용 품사 태그 세트는 어절 내에서의 다양하고 복잡한 형태소들의 결합 제약을 반영하는 형태론적 분류보다 덜 정밀하며, 문장 내에 멀리 떨어져 있는 어절 간 또는 구절 간의 순서, 공기 제약 관계 등의 현상을 반영하는 구문론적 분류보다도 덜 정밀하다. 그래서, 태깅용 품사 태그 세트는 형태소 분석과 구문 분석 사이의 중간 단계의 처리를 위한 ‘형태-통사분석론’적인 분류를 따른다고 말할 수 있다.

그리고, 자동 품사 태거를 정보 검색에서의 명사 추출의 정확도를 높이기 위해서 사용할 수 있다는 것을 주지하고 있을 것으로 간주하며, 또한 이 권고안을 활용하는 사람들은 통계적 자연 언어 처리의 기본 개념과 제약사항 등을 숙지하고 있음을 가정한다.

끝으로, 태그 세트에서 다루지 않은 세부적인 언어 현상들은 각 시스템의 문법 규칙이나 전자사전의 자질-값에서 반영한다는 것을 가정한다.

3. 참조 표준(권고)

3.1. 국외 표준(권고)

- 해당 사항 없음.

3.2. 국내 표준

- 해당 사항 없음.

4. 용어 정의 및 약어

4.1.1. 언어 자원(language resource)

최근에 유행하는 경험적이고 관찰적인 자연언어 처리 기법에서 귀납적으로 언어 지식을 추출하기 위해 필요한 대량의 다양한 언어 현상을 담고 있는 기초 자료이며, 전자사전, 말뭉치, 품사 부착 말뭉치, 구문구조 부착 말뭉치 등을 말한다. 이 중, 본 표준안과 직접적으로 관련이 있는 ‘말뭉치’, ‘품사 부착 말뭉치’에 대해서만 정의를 내리기로 한다.

4.1.2. 말뭉치(corpus)

신문, 소설, 교과서 등 각 분야의 저서나 출판물을 정해진 저장 양식과 표현 양식에 따라 전자화(machine-readable)한 텍스트 모음을 말한다.

4.1.3. 형태소(morpheme)

우리나라 말에서 의미의 최소단위로 정의되는 단위이다. 예를 들어, ‘나무’의 의미는 ‘나’와 ‘무’의 의미로 더 이상 분해 불가능하므로 ‘나무’는 하나의 의미단위인 형태소가 되고, ‘코웃음’은 ‘코’와 ‘웃음’이 의미가 있으므로 ‘코, 웃음’이 각각 형태소가 된다.

4.1.4. 형태소 분석(morphological analysis)

어절(즉, 한글에서의 띄어쓰기의 기본 단위)을 분석을 위한 기본 단위로 하여, 가능한 모든 문법적인 형태소의 결합 구조와 그 결합 구조에서의 각 형태소가 취하는 품사를 출력하는 것을 말한다. 예를 들면, ‘차는’이라는 입력 어절에 대해 형태소 분석기는 ‘차/명사 + 는/조사’와 ‘차/동사 어간 + 는/어미’를 가능한 결합 구조로 출력할 수 있다.

4.1.5. 품사 태그 세트(POS tag set)

한글을 형태소 단위로 분리했을 때 각 형태소가 형태나 기능적으로 분류되는 품사(POS, Part-of-Speech) 체계를 말한다. 국어학에서는 기본적으로 8 품사를 정하고 있지만, 본 표준에서 말하는 태그 세트는 목적과 용도가 다르므로 전산 처리에 맞는 품사 체계를 말한다. 따라서, 국어학에서 정의하는 품사보다 적용 범위가 좀 더 넓은 것으로써, 형태-통사론적인 관점의 품사 태그 세트라고 볼 수 있다. 그러한 관점에서 전통적인 개념의 품사라는 말보다는 어휘 범주라는 말이 더 적합하지만 일반적으로 많이 쓰이는 용어를 쓰기로 한다.

4.1.6. 태그(tag)

태그 세트를 이루는 구성요소들으로써 태깅의 관점에서 분류한 각각의 품사를 상징하는 암호화된 표식을 말한다.

4.1.7. 태깅(tagging)

말뭉치의 어휘를 형태소 단위로 분석하고, 주변 문맥을 고려해서 가장 적합한 품사 태그를 하나만 부여하는 것이다. 예를 들면, '차'라는 어휘는 동사 어간, 명사, 형용사 어간 등이 될 수 있지만, '나는 차를 마신다'라는 문장에서는 '차/명사+를/조사'의 형태로 하나의 적절한 품사가 부여되고, '나는 공을 찬다'라는 문장에서는 '차/동사어간+는다/어말어미'와 같이 하나의 적절한 품사를 부여하는 것이다.

4.1.8. 품사 부착 말뭉치(POS tagged corpus)

정해진 분리 원칙에 따라 말뭉치의 어절을 형태소 단위로 분리하고, 정해진 표현 양식과 할당 기준 및 문맥에 따라 각 형태소에 1 개의 태그를 부착한 형태의 텍스트 모음을 말한다.

4.1.9. 전산 처리적 관점과 용도

전산 처리적인 관점이라는 것은 좀더 엄밀하게 말하면, 사람들에게 설명을 위한 형태로 언어현상의 특징을 설명하는 형태론, 구문론이 아니다. 형태소 분석기, 품사 태거, 구문 분석기 등과 같은 논리적으로 엄격하고 융통성이 별로 없는 프로그램으로 하여금 주어진 컴퓨팅 환경과 제한된 메모리와 계산 능력을

활용하여 관찰 가능하고 처리 가능한 범위 내의 문장을 입력으로 받아 정해진 형태의 출력을 제한된 시간 내에 출력해야 한다는 제약적 용도를 고려한 관점이다.

5. 품사 태그 세트 분류 기준¹⁾²⁾

5.1. 태그 세트 분류 체계

대분류	소분류	세분류
(1) 체언	명사(NN)	일반명사(NNG)
		고유명사(NNP)
		의존명사(NNB)
	대명사(NP)	대명사(NP)
	수사(NR)	수사(NR)
(2) 용언	동사(VV)	동사(VV)
	형용사(VA)	형용사(VA)
	보조용언(VX)	보조용언(VX)
	지정사(VC)	긍정지정사(VCP)
부정지정사(VCN)		
(3) 수식언	관형사(MM)	성상 관형사(MMA)
		지시 관형사(MMD)
		수 관형사(MMN)
	부사(MA)	일반부사(MAG)
		접속부사(MAJ)
(4) 독립언	감탄사(IC)	감탄사(IC)
(5) 관계언	격조사(JK)	주격조사(JKS)
		보격조사(JKC)
		관형격조사(JKG)

¹⁾ 본 표준의 설정에 있어서 먼저 기존 태그 세트와 호환성을 제고하기 위하여 세종태그셋, 품사 부착 말뭉치 작성용 품사 태그 세트 표준(TTA.KO-11.0010), ETRI, 서울대, KAIST, 포항공대, 울산대, 충북대, MORAN, 연세대, 고려대, MACH, HAM, 표준국어대사전의 14개 기관에서 사용하는 태그 세트의 비교 과정을 거쳤음을 밝힌다.

²⁾ 본 표준 태그세트 분류 기준의 기술 형식은 세종21계획의 결과물인 '형태분석 말뭉치 구축 지침(Ver. 2005-1)' 지침서를 참고했음을 밝힌다.

		목적격조사(JKO)
		부사격조사(JKB)
		호격조사(JKV)
		인용격조사(JKQ)
	보조사(JX)	보조사(JX)
	접속조사(JC)	접속조사(JC)
(6) 의존형태	어미(EM)	선어말어미(EP)
		종결어미(EF)
		연결어미(EC)
		명사형전성어미(ETN)
		관형형전성어미(ETM)
	접두사(XP)	체언접두사(XPN)
	접미사(XS)	명사파생접미사(XSN)
		동사파생접미사(XSV)
		형용사파생접미사(XSA)
	어근(XR)	어근(XR)
(7) 기호	일반기호(ST)	마침표, 물음표, 느낌표(SF)
		쉼표, 가운뎃점, 콜론, 빗금(SP)
		따옴표, 괄호표, 줄표(SS)
		줄임표(SE)
		불임표(물결)(SO)
		기타 기호(SW)
	외국어(SL)	외국어(SL)
	한자(SH)	한자(SH)
	숫자(SN)	숫자(SN)
	분석불능범주(NA)	분석불능범주(NA)

- 품사마다 세 분류된 정도가 다른 것은 품사 태깅의 관점에서 유용성을 검토하여 적합성에 따라 취사선택을 하였기 때문이며, 미학적이거나 구조상의 균형을 고려한 것이 아님을 밝혀둔다. 단지, 분류구조상의 균형성만을 이유로 똑같은 깊이의 품사세분을 하여 품사 태그 부착 말뭉치를 구축하면 태깅의 관점에서 불필요한 낭비와 작업의 어려움을 초래하기 때문이다.

- 태그 세트는 용도에 따라 더 세분되거나 덜 세분된 분류가 제시될 수 있지만, 이 권고안에서 당면 목표로 하는 용도는 자동 품사 태거용 품사 부착 말뭉치 구축에 있음을 다시 한 번 밝혀 둔다.
- 새로운 의미 있는 용도나 언어현상의 출현 또는 발견으로 새 범주의 태그 설정이 필요할 경우에는 이 체계를 유지하면서 계층적으로 새로운 하위분류를 추가하여 점증적인 방식으로 다듬어 나갈 수 있으므로, 이 분류 체계는 개방적이고 수용적인 기본 틀을 제공하는 것을 목적으로 하고 있다.
- 본 권고안은 품사 태그 세트와 품사별 태깅 지침의 표준안 마련에 목적이 있으며, 기본적인 태깅 지침은 세분류를 기준으로 작성하였다. 하지만, 본 권고안의 사용자가 필요에 따라서 소분류와 세분류를 선택해서 사용할 수 있도록 소분류에도 품사 태그를 설정하였다.

5.2. 각 태그별 정의 및 분류 기준

5.2.1. 체언

체언은 명사, 대명사, 수사를 포괄하는 대범주로서, 조사와 결합하거나 그 자체로 다른 체언이나 용언과 어울려 하나의 문장 성분이 될 수 있다.

5.2.1.1. 명사(NN)

명사는 사물의 이름을 나타내는 품사이다. 본 표지에서는 명사를 일반명사, 고유명사, 의존명사로 세분한다.

5.2.1.1.1. 일반명사(NNG)

사물의 이름을 나타내는 단어로써 표준국어대사전에 명사로 등재된 표제어(고유명사와 의존명사를 제외한 모든 명사)와 독립된 음절(한자어), 약어, 고사성어 등 사전 표제어는 아니나 다른 품사로 분석될 수 없는 단위들을 포함한다.

(1) 일반명사로 분석할 수 있는 단어

(가) 표준국어대사전의 명사 표제어

✧ 국어/NNG, 연구/NNG

(나) 1 음절 한자어가 독립된 단위로 사용되는 경우

✧ 서울초등학교 줄 [줄/NNG]

(다) 한자 성어

✧ 백척간두(百尺竿頭) [백척간두/NNG+(/SS+百尺竿頭/SH+)/SS]

(라) 외국어를 음차한 경우

✧ 아이 러브 유 [아이/NNG]

(마) 기타 다른 품사로 분석될 수 없는 단위

✧ 5 관왕 [5/SN+관/NNG+왕/NNG]

방커 C 유 [방커/NNG+C/SL+유/NNG]

(2) 명사 상당어의 분석

(가) 동사의 활용형이 따옴표 없이 문장 속에서 명사처럼 기능하는 경우는 원래 품사대로 분석한다.

✧ 어디 가느냐가 그의 물음이었다. [가/VV+느냐/EC+가/JKS]

(나) 따옴표를 가진 성분이나 요소도 명사처럼 기능할 수 있으나, 원래 품사대로 분석한다.

✧ 그것은 “는”이 아니라 “를”이다. [“/SS+는/JX+”/SS+이/JKC]

(다) 부사 뒤에 격조사가 쓰이는 것도 의미론적인 따옴의 효과에 의하여 부사가 명사적인 용법을 가지는 것이므로 분석은 ‘부사’로 한다.

✧ 가족을 멀리에 보냈다. [멀리/MAG+에/JKB]

5.2.1.1.2. 고유 명사(NNP)

고유 명사는 특정한 사물에 붙여진 이름으로, 기본적으로 최하의어에 속하는 대상을 서로 변별하기 위하여 붙인 이름이며, 원칙적으로 지시 대상만 가질 뿐 의미 내용은 가지지 않는다. 고유명사의 분석 기준은 매우 다양하므로, 본 지침에서는 다음에 제시하는 것만을 고유명사로 인정한다. 또한, 본 지침은 띄어쓰기 단위의 분석을 원칙으로 하고 있으므로, 한 단어 이상으로 구성된 고유명사(‘바람과 함께 사라지다’)와 같은 경우의 분석을 위해 전체를 아우르는 단위를 설정하지는 않는다.

(1) 인명, 종족명

(가) ‘씨(氏), 공(公), 군(君), 양(孃), 웅(翁)’ 등 성 또는 이름 뒤에 같이 쓰이는 호칭어나 직책명은 분리해서 분석한다.

✧ 남수/NNP+군/NNB, 김/NNP+씨/NNB, 최치원/NNP+웅/NNB

✧ 케네디/NNP+씨/NNB, 정/NNP+과장/NNG, 최/NNP+선생/NNG

(나) 성과 이름, 호가 함께 쓰이면 하나의 단위로 분석한다.

✧ 김철수/NNP, 이태백/NNP

(다) ‘씨, 군’ 등과 달리 ‘가(哥)’는 접미사이므로, ‘김가(金哥), 이가(李哥)’는 파생어이다.

✧ 김/NNP+가/XSN

(라) 사람 이름의 뒤에 ‘이’가 붙는 경우는 이름과 함께 하나의 단위로 분석한다.

✧ 진현이/NNP + 가/JKS

(마) 특정한 종족의 이름은 고유명사가 된다.

✧ 알타이족/NNP, 피그미족/NNP, 돌궐족/NNP, 한족/NNP

(2) 지명

(가) 내륙, 바다, 강, 산, 산맥, 호수, 섬, 만, 계곡, 늪, 주 등의 이름

✧ 카스피해/NNP, 템즈강/NNP, 태백산맥/NNP, 미시시피호/NNP, 네바다주/NNP

(나) 도(道), 시(市), 읍(邑), 면(面), 리(里), 군(郡), 구(區), 동(洞), 골, 촌 등의 이름은 그 구역의 종류를 나타내는 말과 함께 전체가 고유명사가 된다.

✧ 서울특별시/NNP, 성북구/NNP, 강진군/NNP, 인창동/NNP, 빨래골/NNP, 해방촌/NNP

(3) 국가명 또는 왕조명

(가) 국가의 명칭, 또는 왕조의 명칭은 고유명사로 분석한다.

✧ 대한민국/NNP, 조선/NNP

(나) 다른 형태가 붙어 국가나 왕조의 존립 기간을 나타내는 경우 일반명사로 분석한다.

✧ 대한제국기/NNG, 조선조/NNG

(다) ‘남한’과 ‘북한’을 의미하는 ‘남, 북, 남북’은 모두 일반명사와 고유명사를 구별한다.

남한을 뜻하는 ‘남’과 북한을 뜻하는 ‘북’을 고유명사로 분석한다.

✧ 남/NNP+과/JKB 북/NNP+의/JKG 의견/NNG 차이/NNG

✧ 남북/NNP 적십자회담/NNG

✧ 북/NNP+미/NNP 회담/NNG

(라) 어떤 국가의 국민을 나타내는 ‘국가+인’은 통합하여 일반명사로 분석한다.

✧ 이집트인/NNG, 아제르바이젠인/NNG, 이스라엘인/NNG, 조선인/NNG

(마) 어떤 국가의 군대를 나타내는 ‘국가+군’은 통합하여 일반명사로 분석한다.

✧ 미군/NNG, 북한군/NNG, 영국군/NNG, 일본군/NNG

(4) 건축물이나 시설물 혹은 구조물의 이름

(가) 도로, 항만, 철도, 전철, 지하철 및 그 명칭과 함께 쓰이는 부대시설은 그 종류를 나타내는 말과 함께 전체가 고유명사가 된다.

✧ 부산항/NNP, 대전역/NNP, 서울지하철/NNP, 테헤란로/NNP

(나) 빌딩, 박물관, 극장 등 건물명은 그 종류를 나타내는 말과 함께 전체가 고유명사가 된다.

✧ 서울역사/NNP, 세종문화회관/NNP, 개나리유치원/NNP, 고려대학교/NNP

✧ 국립중앙박물관/NNP, 국립민속박물관/NNP, 구텐베르크박물관/NNP

✧ 신라호텔/NNP, 미도파백화점/NNP, 동궁예식장/NNP, 명보극장/NNP, 고대병원/NNP

(5) 회사, 학교, 정당, 기관이나 단체의 이름

(가) 특정 회사나 학교, 정당 등의 이름은 고유명사로 분석한다. 단, 특정 회사의 상품명은 고유명사가 아닌 일반명사로 취급한다.

✧ 삼성/NNP, 고려대학교/NNP, 한나라당/NNP, 자유민주주의연합/NNP

✧ 초코하임/NNG, 한메타자교실/NNG

(나) 정부기관의 명칭은 모두 일반명사로 처리한다. 그러나 거기에 인명, 지명 등의 고유명사가 포함된 경우 그 통합형을 고유명사로 처리한다.

✧ 헌법/NNG+재판소/NNG, 대/XPN+법원/NNG, 고등/NNG+법원/NNG, 재정/NNG+경제원/NNG

✧ 서울고등법원/NNP, 서울시경찰서/NNP, 서대문구치소/NNP

(다) 특정 기관이나 단체, 연구소 등의 경우에는 분석하는 것을 원칙으로 한다. 그러나 거기에 인명, 지명 등의 고유명이나 ‘전국’, ‘국제’, ‘세계’ 등이 포함되면 그 통합형을 고유명사로 처리한다.

✧ 대한축구협회/NNP, 전국은행협회/NNP, 한국전자통신연구원/NNP

※ 생활/NNG+체육/NNG+연구소/NNG, 입주자/NNG+대표자/NNG+협의회/NNG

(라) 약어나 준말의 처리

고유명사가 축약된 형태(준말)로 쓰일 경우 본디말과 함께 준말도 인정하여 축약된 형태 그대로를 고유명사로 분석한다. 그리고 일반명사로 분석하는 기관명의 약자는 일반명사로 분석한다.

※ 육사/NNP, 고대/NNP, 자민련/NNP, 서울고법/NNP

※ 정보통신위/NNG

(6) 책, 연극, 영화 등의 창작물의 제목

※ 삼국사기/NNP, 손자병법/NNP, 고래사냥/NNP

(7) 언어명

언어명의 경우 ‘-어’의 형태만을 통합하여 고유명사로 인정한다. ‘한국말’과 같은 경우는 일반명사로 분석한다.

※ 한국어/NNP, 일본어/NNP, 영어/NNP, 알타이어/NNP, 네덜란드어/NNP

5.2.1.1.3. 의존명사(NNB)

의존명사는 자립해서 쓰일 수 없는 명사로, 수식 성분을 반드시 동반해야 한다. 의존명사는 비단위성 의존명사와 단위성 의존명사로 나눌 수 있으나, 본 분석에서는 이를 세분화하지 않는다. 또한 의존명사가 일반명사와 같이 독립적으로 쓰일 때는 일반명사로 분석한다. 의존명사와 일반명사의 구분은 표준국어대사전의 등재 여부에 따른다.

(1) 의존명사이지만, 일반명사처럼 쓰이는 경우

(가) “연대, 연도, 연차”는 “년대, 년도, 년차”와 달리 모두 일반명사로 분석한다.

※ 연도별로 정리된 자료 [연도/NNG]

※ 몇 년도에 일어난 일 [년도/NNB]

(나) “월, 연, 일, 주, 달러, 원” 등은 본래 의존명사이지만, 독립되어 쓰일 경우 모두 일반명사의 자격을 가지므로 일반명사로 분석해야 한다.

※ 나는 월 30 만원을 받는다. [월/NNG]

※ 달러의 가치는 [달러/NNG]

(2) 단위를 나타내는 표현

(가) 길이, 무게, 수효, 시간 따위의 수량을 수치로 나타내는 단위들 중 “미터, 그램, 리터” 등은 의존명사(NNB)로, 외국어로 된 “m, g, l” 등은 기호(SW)로 분석한다.

(나) 일반명사가 단위적인 용법으로 쓰인 경우에는 의존명사가 아니므로 주의한다.

※ 사람, 시간, 그릇 …

- 한 사람이 교실로 들어왔다. [사람/NNG+이/JKS]
- 자장면 한 그릇만 주세요. [그릇/NNG+만/JX]

(3) ‘것’과 구어형 ‘거’의 분석

‘거’의 형태를 그대로 인정하여 분석한다. 그러나 다른 형태와의 결합에서 ‘거’의 형태가 유지되지 않는다면 그 때에는 ‘것’으로 복원하여 분석한다.

- ※ 공부할 거를 준비해 왔니? [거/NNB+를/JKO]
- ※ 공부할 걸 가져왔니? [것/NNB+ㄹ/JKO]
- ※ 연습할 건 있니? [것/NNB+ㄴ/JX]
- ※ 먹을 게 모자라다. [것/NNB+이/JKS]

5.2.1.2. 대명사(NP)

대명사는 그 자체로는 자신의 본유적 지시물을 가지지 않은 채, 다만 사람이나 사물 등 어떤 대상을 간접적으로 지시하는 품사이다. 단, 동일한 대명사가 방언이나 고어의 이형태를 가진 경우에는 이들도 대명사로 같이 분석한다.

(1) 1 인칭 대명사

(가) 1 인칭 대명사

※ 나, 내, 우리, 저, 제, 저희

(나) 2 인칭 대명사

※ 너, 네, 그대, 당신, 댁, 어르신

(다) 기타 대명사

※ 이이, 이분, 그이, 그분, 저이, 저분, 아무, 아무개, 누구, 무엇, 뭐, 어디, 언제, 자기, 개, 재, 애, 이것, 저것, 그것, 이거, 저거, 그거, 여기, 저기, 거기, 이곳, 그곳, 저곳, 어디, 모(某), 모모(某某)

(2) 대명사와 관형사의 두 가지 분석이 가능한 단어

(가) ‘모(某)’는 관형사와 대명사로 분석될 수 있으므로 주의를 요한다.

✧ 모 기업체 [모/MMD]

✧ 김 모씨 [모/NP+씨/NNB]

(나) ‘모모(某某)’도 위와 같이 분석될 수 있다.

✧ 모모가 말했다 [모모/NP+가/JKS]

✧ 모모 기관의 조사를 마쳤다 [모모/MMD]

(3) 대명사의 이형태 분석

(가) ‘이것, 그것, 저것; 이거, 그거, 저거’는 분석하지 않고 대명사로 인정한다.

‘~거’의 경우, 다른 형태와의 결합에서 ‘~거’의 형태가 유지되지 않는다면 ‘~것’으로 복원한다.

✧ 난 저거를 먹을래. [저거/NP+를/JKO]

✧ 나는 여태 그걸 믿어 왔단다. [그것/NP+를/JKO]

(나) 다음과 같이 원형을 밝힐 수 있는 대명사는 원형대로 분석한다.

✧ 내 이제부터는 내 명령을 따라라. [나/NP+의/JKG]

✧ 내게 내게 전자우편으로 알려 다오. [나/NP+에게/JKB]

✧ 네게 어제 네게 보낸 선물이 잘못되었다. [너/NP+에게/JKB]

✧ 제게 문제가 있다면 제게 말씀해 주세요. [저/NP+에게/JKB]

✧ 누가 누가 전화를 하는 지 보고해라. [누구/NP+가/JKS]

✧ 뉘 뉘 집 애기가 이렇게 울고 있는 거야? [누구/NP+의/JKG]

✧ 뭐가 도대체 뭐가 문제라는 거야? [뭐/NP+가/JKS]

(다) ‘뭐’는 ‘무엇’과 대등할 정도로 자주 사용되므로 그 형태 자체를 인정해 준다.

다만, 다음과 같이 조사와 축약되었을 경우에만 원형으로 복원해 준다.

✧ 앞으로 우리가 뭘 하자는 얘기이냐? [무엇/NP+를/JKO]

(라) ‘제’의 경우, ‘제/NP+가/JKS’를 제외하고는 모두 ‘저/NP+의/JKG’로 분석한다.

✧ 제가 갈 것입니다. [제/NP+가/JKS]

✧ 철수는 제 잘못을 안다. [저/NP+의/JKG]

✧ 제 무게를 못 견디다. [저/NP+의/JKG]

✧ 제 값을 하다. [저/NP+의/JKG]

5.2.1.3. 수사(NR)

수사는 사물의 수량이나 차례를 나타내는 품사를 말한다.

(1) 수사의 종류

(가) 양수사

- ✧ 하나, 둘, 셋, 넷, 다섯, 여섯, 일곱, 여덟, 아홉, 열, 스물, 서른, 마흔, 쉰, 예순, 일흔, 여든, 아흔, 백
- ✧ 한둘, 두서넛, 서넛, 너댓, 네다섯, 네댓, 대여섯, 예닐곱, 일여덟, 일고여덟, 열두서넛, 열대여섯,
- ✧ 열일고여덟, 스물두서넛
- ✧ 일, 이, 삼, 사, 오, 육, 칠, 팔 구, 십, 백, 천, 만, 억, 조
- ✧ 기십, 기백, 기천,
- ✧ 수십, 수백, 수천, 수만, 수억, 수십만, 수백만, 수천만

(나) 서수사

- ✧ 첫째, 둘째, 셋째, 넷째, ..., 열째, 열한째, ..., 스물한째, ...,
- ✧ 아흔아홉째, 백째, 백한째, ...

(2) 주의 사항

(가) 복수의 수사가 한 어절 내에 나타날 때에는 전체를 통합해서 분석한다.

- ✧ 백만오천삼십사 [백만오천삼십사/NR]

(나) ‘하나’는 표준국어대사전에 그 품사가 명사와 수사로 되어 있지만 본 지침에서는 수사로 분석한다.

- ✧ 광에 가서 물건 하나만 가져오렴. [하나/NR+만/JX]
- ✧ 우리는 하나로 뭉쳤다. [하나/NR+로/JKB]

(다) 때로 수사와 수관형사의 구별이 애매한 경우가 있다. 임흥빈(1998)의 견해에 따르면, 다음과 같은 특이한 형식을 가진 예만을 수관형사로 취급하고, 그 밖의 것들은 모두 수사로 분석하지만, 본 표준안에서는 조사와 결합한 경우에만 수사로 판별한다.

- ✧ 임흥빈이 제시한 수관형사: 한, 한두, 한두어, 두, 두어, 두세, 두서너, 세, 석, 서, 서너, 네, 너, 녀

(라) ‘제일, 제이’ 등은 접두사 ‘제-’와 수사의 결합으로 분석한다.

- ✧ 제일, 제이, 제삼, 제사, 제오, …, 제구십구, 제백, …
 - [제/XPN+일/NR], [제/XPN+이/NR], …

5.2.2. 용언

용언은 동사, 형용사, 지정사를 가리킨다. 용언 범주에서는 분석 대상이 본용언일 경우에만 동사와 형용사로 구분하여 표시하고, 보조용언의 경우에는 보조동사와 보조형용사를 구분하지 않고 ‘VX’라는 하나의 표지만을 준다. 또한 학교 문법에서 서술격조사로 다루는 ‘이다’는 조사의 범주에 넣지 않고 ‘지정사’라는 용언의 하위 범주에 넣기로 한다. 지정사는 다시 긍정 지정사(VCP)와 부정 지정사(VCN)로 세분된다.

5.2.2.1. 동사(V)

동사는 사물의 움직임이나 작용을 나타내는 용언을 말한다. 동사는 일반적으로 목적어의 필요성 여부에 따라 자동사, 타동사로 나누기도 하지만, 본 분석에서는 그것을 위한 별도의 표지를 세분하지 않고 모두 ‘V’로 표시한다.

5.2.2.2. 형용사(VA)

형용사는 사물의 성질이나 상태를 나타내는 용언을 가리킨다.

5.2.2.3. 보조용언(VX)

이 분석에서는 보조용언을 보조동사와 보조형용사로 하위 구분하지 않는다.

(1) 보조용언 분석 원칙

- (가) 보조용언의 후보는 표준국어대사전에 그 쓰임이 제시되어 있어야 한다.
- (나) 보조용언 앞에는 반드시 다른 용언이 위치해 있어야 한다.
- (다) 보조용언이 동시에 두 개 이상이 연결되어 나타날 수도 있다.

(2) 보조용언의 예와 주의 사항

보조용언의 목록은 다음과 같다. 이 목록은 표준국어대사전을 참고한 것이다.

- ✧ 가다
 - 세월이 흘러 가는 대로 정처없이 떠도는 나그네 [가/VX+는/ETM]
- ✧ 가지다

- 일을 그렇게 해 가지고는 기일을 맞출 수 없다. [가지/VX+고/EC+는/JX]
- ✧ 계시다
 - 손님께서 와 계십니다. [계시/VX+ㅁ니다/EF+./SF]
- ✧ 나가다
 - 정책을 추진해 나가는 과정에서 문제가 생겼다. [나가/VX+는/ETM]
- ✧ 나다
 - 아침에 깨어 나 보니 그가 없어졌다. [나/VX+아/EC]
- ✧ 내다
 - 힘들겠지만 잘 견뎌 내야 한다. [내/VX+아야/EC]
- ✧ 놓다
 - 약속을 잡아 놓고 출장을 가다니 [놓/VX+고/EC]
- ✧ 달다
 - 이번 시험 문제의 정답을 알려 다오. [달/VX+오/EF+./SF]
- ✧ 대다
 - 자꾸 졸라 대는 통에 그만 허락해 주고 말았다. [대/VX+는/ETM]
- ✧ 두다
 - 남겨 둔 돈도 이제 바닥이 났다. [두/VX+ㄴ/ETM]
- ✧ 드리다
 - 영려를 끼쳐 드려 송구하옵니다. [드리/VX+어/EC]
- ✧ 들다
 - 도무지 내 말은 믿으려 들지 않는다. [들/VX+지/EC]
- ✧ 말다
 - 어렵더라도 희망을 잃지 말아야 한다. [말/VX+아야/EC]
- ✧ 먹다
 - 나는 오늘도 수업을 빼 먹었다. [먹/VX+었/EP+다/EF+./SF]
- ✧ 못하다
 - 그 참상을 차마 보지는 못할 것이다. [못하/VX+ㄹ/ETM]
- ✧ 버리다
 - 음식이 다 타 버렸다. [버리/VX+었/EP+다/EF+./SF]
- ✧ 보다
 - 이제는 새벽이 오는가 보다. [보/VX+다/EF+./SF]
- ✧ 빠지다
 - 썩어 빠진 생선을 사오다니 [빠지/VX+ㄴ/ETM]
- ✧ 싶다

- 너를 보고 싶다. [싶/VX+다/EF+./SF]
- ✧ 쌀다
 - 꼬치꼬치 물어 쌀는 통에 정신이 없었다. [쌀/VX+는/ETM]
- ✧ 아니하다
 - 일이 순리대로 풀리지 아니했다. [아니하/VX+았/EP+다/EF+./SF]
- ✧ 않다
 - 시간이 지나도 기차는 오지 않았다. [않/VX+았/EP+다/EF+./SF]
- ✧ 오다
 - 고향을 떠나 온 지 10년이 지났다. [오/VX+ㄴ/ETM]
- ✧ 있다
 - 그녀는 걱정 옷을 입고 있었다. [있/VX+었/EP+다/EF+./SF]
- ✧ 주다
 - 아버지는 아기에게 동화책을 읽어 주었다. [주/VX+었/EP+다/EF+./SF]
- ✧ 지다
 - 한 번 넘어 진 아이는 일어나는 법을 안다. [지/VX+ㄴ/ETM]
- ✧ 치우다
 - 다섯 명이 10인분의 식사를 먹어 치웠다. [치우/VV+었/EP+다/EF+./SF]
- ✧ 터지다
 - 끓인 지 오래 되어서 라면이 불어 터졌다. [터지/VX+었/EP+다/EF+./SF]
- ✧ 하다
 - 나귀를 쉬게 하는 것이 좋겠다. [하/VX+는/ETM]

① 다음과 같은 어절은 보조용언으로 취급되기도 하나, 여기서는 ‘의존명사+접사’로 분석한다. 이들 앞에는 항상 관형어가 온다는 분포적인 특성을 중시한 것이다.

- ✧ 양하다/체하다/척하다/듯하다/법하다/뻥하다 [양/NNB+하/XSA+다/EF+./SF]

② ‘버릇하다’의 경우에는 선행 성분으로 관형형이 오는 것은 아니지만, 일반명사 ‘버릇’과 크게 구별되지 않으므로 ‘버릇’은 명사로 분석한다.

- ✧ 자꾸 물어 버릇하다. [버릇/NNG+하/XSV+다/EF+./SF]

5.2.2.4. 지정사(VC)

지정사는 학교 문법의 서술격 조사에 대응되는 것인데, 용언과 같이 활용한다는 특성을 중시한 술어이다. 여기서는 학교 문법의 ‘이다’를 긍정 지정사로, ‘아니다’를 부정

지정사로 하위 구분한다. 일반적으로 ‘아니다’는 형용사로 다루어지기도 하나, 여기서는 ‘아니다’가 ‘이다’의 부정형이라는 점을 중시하여 ‘부정지정사’로 다룬다.

- ✧ 철수는 매우 우수한 학생이다. [학생/NNG+이/VCP+다/EF+./SF]
- ✧ 철수는 모범적인 학생이 아니다. [아니/VCN+다/EF+./SF]

(1) 지정사 ‘이/VCP’를 복원해야 하는 경우

① 체언에 어미가 직접 연결된 경우

- ✧ 철수는 훌륭한 교사다. [교사/NNG+이/VCP+다/EF+./SF]

② 조사에 어미가 직접 연결된 경우

- ✧ 우리가 그를 본 것은 서울에서다.
 - [서울/NNP+에서/JKB+이/VCP+다/EF+./SF]

③ ‘~였다’

- ✧ 그 당시 나는 아이였다. [아이/NNG+이/VCP+였/EP+다/EF+./SF]

④ 어미 '-라고, -라는, -라도, -라며, -라면서, -라서

- ✧ 나는 그에게 절교라고 말했다. [절교/NNG+이/VCP+라고/EC]
- ✧ 나는 친구라는 말이 좋다. [친구/NNG+이/VCP+라는/ETM]
- ✧ 거지라도 존중해 주어야 한다. [거지/NNG+이/VCP+라도/EC]
- ✧ 그는 최고라며 나를 추켜 주었다. [최고/NNG+이/VCP+라며/EC]
- ✧ 그는 실수라면서 얼버무렸다. [실수/NNG+이/VCP+라면서/EC]
- ✧ 너는 부자라서 우릴 이해하지 못할 것이다. [부자/NNG+이/VCP+라서/EC]

⑤ 인용문 뒤에 오는 “~며”는 지정사를 복원하지 않는다.

- ✧ 얼마나 친절하나?”며 [친절/NNG + 하/XSA + 나/EF + ?/SF + "/SS + 며/EC]

5.2.3. 수식언

5.2.3.1. 관형사(MM)

관형사는 체언 앞에서 그것을 꾸미는 품사를 말한다. 관형사는 지시관형사, 수관형사, 성상관형사로 세분될 수 있다.

5.2.3.1.1. 성상관형사

먼저 체언의 성질이나 상태를 ‘어떠한’의 방식으로 꾸며주는 구실을 하는 관형사를 성상관형사라 한다. 예를 들어 ‘새 목록’에서 ‘새’가 성상관형사인데, 이는 ‘새로운 목록’에서 ‘새로운’이라는 형용사의 관형사형이 하는 역할과 비슷하다.

성상관형사의 예제는 다음과 같다.

✧ 상태

- 온갖, 새, 현, 헛, 윗, 뒷, 온, 못, 외딴, 참, 거짓, ...
- 순(純) 주(主), 정(正), 준(準), 대(大), 소(小), ...

✧ 정도

- 단지(但只), 유독(惟獨), 무려(無慮), 약(若), ...

5.2.3.1.2. 지시관형사

다음 어떤 대상을 가리키는 관형사를 지시관형사라 한다. 지시관형사에는 다음과 같은 것들이 있다.

✧ 이, 그, 저, 요, 고, 조

✧ 이런, 그런, 저런, 다른

✧ 어느, 무슨, 웬

✧ 옛, 올, 현(現), 신(新), 구(舊), 전(前), 후(後), 래(來)

‘이, 그, 저’는 동일한 형태가 지시관형사와 지시대명사로 쓰인다. 예를 들면, “이도 저도 다 싫다.”에서 ‘이, 저’는 대명사이고, “저 여자가 이 거리에서 빵을 판다.”에서는 관형사이다. 이렇게 동일한 형태가 있는 경우, 조사와 결합된 경우에만 대명사로 판별한다.

‘이런, 그런, 저런’은 형용사 ‘이렇다, 그렇다, 저렇다’의 활용형 ‘이러한, 그러한, 저러한’이 줄어든 형식이다. 이런 경우에는 ‘이런, 그런, 저런’의 형태가 표준국어대사전에 관형사로 등재된 경우에만 관형사로 판별한다.

5.2.3.1.3. 수관형사

수량을 나타내는 관형사가 수관형사이다. 양(量)을 나타내는 양수 관형사와 순서(順序)를 나타내는 서수 관형사로 나뉘지만, 본 표준에서는 수관형사로만 구분한다.

수관형사는 대체로 단위를 나타내는 단위성 의존명사 앞에 오지만, 자립 명사 앞에도 온다.

✧ 잣나무 열 그루를 심었다

수관형사와 수사는 조사를 취할 수 있는가에 의하여 구별되는데, 조사를 취하면 수사로 판별한다.

✧ 첫째 분이 나의 형이다. → 첫째: 수관형사

✧ 첫째로 남에게 친절해야 한다. → 첫째: 수사

수관형사의 예는 다음과 같다.

✧ 한, 두, 세(석, 서), 네(넉, 너), 다섯(닷), 엇, 일곱, 여덟, 아홉, 열, 열한, 열두, 열세, 열석, 열서, 스무, 스물한, 스물두, 한두, 두세, 서너, 너댓, ...

✧ 첫, 첫째, 둘째, 셋째, 제일(第一), 제이(第二), ...

✧ 한두째, 주어째, 몇째, 여남은째, 몇몇, 여러, ...

(1) 주의 사항

(가) 관형사는 때로 문맥에 따라 다른 품사로 분석될 가능성이 있으니 문맥을 잘 살펴서 분석해야 한다.

✧ 관형사, 명사 통용

• 올 예산이 다 바닥이 났다. [올/MMD]

• 올 들어 물가가 많이 올랐다. [올/NNG]

✧ 관형사, 부사 통용

• 단 세 명에서 그 일을 꾸렸다. [단/MMD]

• 단, 그 일은 해서는 안 된다. [단/MAJ]

✧ 관형사, 명사, 부사 통용

• 이내 마음을 어찌 알리요. [이내/MMD]

• 아침 들판에 이내가 끼었다. [이내/NNG]

• 그는 이내 떠나갔다. [이내/MAG]

(나) 접미사 ‘-적(的)’이 붙는 경우, 언어학적으로는 ‘-적(的)+체언’의 경우에는 관형사로 분류하고 ‘-적(的)+조사’인 경우에는 명사로 분류하지만, 본 표준안에서는 조사와의 결합여부와 관계없이 모두 명사로 분석한다.

✧ 명사의 부사적인 용법 [부사/NNG+적/XSN+이/VCP+ㄴ/ETM]

✧ 명사의 부사적 용법 [부사/NNG+적/XSN]

5.2.3.2. 부사(MA)

부사는 주로 용언을 꾸며서 그 뜻을 더 세밀하고 분명하게 해 주는 품사를 말한다. 여기서는 부사를 세분하지 않고, 접속부사와 일반부사로만 나누기로 한다.

5.2.3.2.1. 접속부사(MAJ)

접속부사는 문장이나 단어를 이어 주며 뒷말을 꾸미는 부사를 말한다. 접속부사에는 다음과 같은 것들이 있다.

- ✧ 단어 접속: 또는, 그리고, 및, 혹은
- ✧ 문장 접속: 왜냐하면, 즉, 그러므로, 그러나, 하지만, 그리고

(1) 주의 사항

- ① 접속부사는 종종 용언의 활용형으로도 쓰일 수 있으므로 주의한다.
 - ✧ 그래서 마지막에는 조심하라고 했지? [그래서/MAJ]
 - ✧ 영희가 그래서 결석을 했구나. [그렇/VA + 어서/EC]
- ② '그리고나서', '그래도'의 분석
 - ✧ 그리고나서 [그리/MAG+하/XSV+고/EC+나/VX+서/EC]
 - ✧ 그래도 [그러/VV+어도/EC]

5.2.3.2.2. 일반부사(MAG)

접속부사를 제외한 나머지 부사는 모두 일반부사로 판별한다. 일반부사는 문장의 한 성분을 꾸며주는 성분 부사를 말하며, 성상 부사, 지시부사, 부정 부사 등으로 구분할 수 있지만, 본 표준안에서는 일반부사로만 구분한다.

- ✧ 성상 부사 : '어떻게'라는 방식으로 용언을 꾸미는 부사로, '매우, 빨리, 잘' 등이 있다.
- ✧ 지시 부사 : 시간·처소, 또는 특정한 대상을 가리키는 부사로, '이리, 저리, 그때, 내일, 일찍, 이미' 등이 있다.
- ✧ 부정 부사 : 용언의 뜻을 부정하는 부사로, '아니, 잘못, 못' 등이 있다.
- ✧ 의성 부사 : 사물의 소리를 표현하는 부사로, '철썩철썩, 딸랑딸랑, 펑펑' 등이 있다.
- ✧ 의태 부사 : 사물의 모양을 표현하는 부사로, '울긁불긁, 꾸불꾸불' 등이 있다.

(1) 주의 사항

- ① 일반부사는 종종 일반명사와 동일형태를 띠고 있어 구분이 어려운 경우가 있다. 이들은 뒤에 조사가 결합하느냐의 여부와, 문맥에서 후행 명사를 수식하느냐의 여부에 따라 부사와 명사로 분석될 수 있다.
 - ✧ 너의 진짜 속셈이 무엇인지 말해 봐라. [진짜/NNG]
 - ✧ 그 수학 문제는 진짜 어려웠다. [진짜/MAG]
 - ✧ 지금이 공부하기 딱 좋은 때이다. [지금/NNG+이/JKS]
 - ✧ 나는 지금 막 집에 도착했다. [지금/MAG]
- ② 부사적인 용법을 가졌음에도 불구하고 일반부사가 아닌 일반명사로만 표준국어대사전에 등재되어 있는 단어는 오로지 일반명사로만 분석한다.
 - ✧ 구석구석, 무작정, 여기저기, 오랫동안, 이곳저곳, 정작, 좌우간, 처음, 최근, 한때
- ③ 일반부사로 분석하기 쉬운 활용상의 불완전동사인 ‘덩달아, 더불어’는 모두 동사로 옳게 분석해야 함에 주의한다.
 - ✧ 너는 덩달아 왜 난리니? [덩달/VV+아/EC]
 - ✧ 우리 함께 더불어 살아가자. [더불/VV+어/EC]
- ④ ‘명사+없이’는 원칙적으로 ‘일반명사+없이/MAG’로 태깅하지만, 아래와 같이 하나의 단어로 굳어져 사전에 등재된 경우는 ‘없이’ 통합형 자체를 하나의 일반부사로 분석한다.
 - ✧ 관계없이, 그지없이, 꾸밈없이, 꿇임없이, 난데없이, 남김없이 등

5.2.4. 독립언

5.2.4.1. 감탄사(IC)

감탄사는 화자의 부름이나 느낌, 놀람이나 대답을 직접적으로 나타내는 품사를 말한다.

- ✧ 그럼(요), 야호, 어머, 앓, 아, 예, 그래(요), 아니(요), 글썄, 참, 아이구, 와아, 오호, 세상에

(1) 주의 사항

- ① 사람이 입으로 직접 내는 소리를 대상으로 하되, 흥내를 내는 의도가 없는 것과 본능적인 놀람이나 느낌을 나타내는 것을 대상으로 한다. 또한 감탄사와 혼동되는

부사로서 음성상징어류의 부사어가 있는데, 이는 감탄사가 아닌 일반부사로 분석한다.

☆ 야호! 드디어 정상이다. [야호/IC+!/SF]

☆ 쿨럭쿨럭 기침을 했다. [쿨럭쿨럭/MAG]

② 동물의 울음소리 등은 감탄사가 아니라 일반부사로 분석한다.

☆ 검둥이는 멍멍 짖으며 수풀 속으로 뛰어들어갔다. [멍멍/MAG]

③ 욕이나 욕설을 나타내는 말은 전체를 감탄사로 분석한다.

☆ 빌어먹을! [빌어먹을/IC+!/SF]

④ ‘뭐’는 문맥에 따라 대명사와 감탄사의 두 가지 쓰임이 있다.

☆ 원지도 모른 채 [뭐/NP+이/VCP+지/EC+도/JX]

☆ 신문에 뭐 대단한 특종이라도 실렸습니까? [뭐/IC]

⑤ 한 어절이 비정상적으로 늘어나거나 다른 기호가 개입되었을 경우 분석불능 범주(NA)로 분석한다.

☆ 그러어엄/NA, 으~어~이/NA

5.2.5. 관계언

조사는 주로 체언과 결합하여 다른 말과의 문법적 관계를 나타내거나, 특별한 뜻을 더해 주는 품사를 말한다. 조사는 크게 격조사, 접속조사, 보조사로 나눈다. 한글은 조사가 중첩하는 경우가 많은데, 이러한 경우 조사의 결합형은 분리해서 분석함을 원칙으로 한다.

☆ 부산에서도 대형 사고가 있었다. [부산/NNP+에서/JKB+도/JX]

☆ 그녀와의 약속이 갑자기 잡혔다. [그녀/NP+와/JKB+의/JKG]

5.2.5.1. 격조사(JK)

이는 체언과 다른 성분 간의 일정한 문법 관계를 나타내는 조사이다.

5.2.5.1.1. 주격조사(JKS)

선행 체언으로 하여금 주어가 되게 하는 조사이다.

☆ 이/가 책이 보인다. [책/NGG+이/JKS]

☆ 께서 선생님께서 오신다. [선생/NGG+님/XSN+께서/JKS]

☆ (이)서 둘이서 그 일을 꾸렸다고? [둘/NR+이서/JKS]

- ☆ 혼자서 그 일을 꾸몄다고? [혼자/NNG+서/JKS]
- ☆ 께오서 부대장님께오서 [부대장/NNG+님/XSN+께오서/JKS]
- ☆ 께옵서 황제께옵서 드나드신다.[황제/NNG+께옵서/JKS]

5.2.5.1.2. 보격조사(JKC)

선행 체언으로 하여금 서술어 ‘되다, 아니다’의 보어가 되게 하는 조사이다.

- ☆ 이/가 얼음이 물이 되었다. [물/NNG+이/JKC]
- ☆ 철수는 범인이 아니다. [범인/NNG+이/JKC]

5.2.5.1.3. 목적격조사(JKO)

선행 체언으로 하여금 목적어가 되게 하는 조사이다.

- ☆ ㄹ/을/를 너는 바람소리를 들었다. [바람/NNG+소리/NNG+를/JKO]

5.2.5.1.4. 관형격조사(JKG)

선행 체언으로 하여금 관형어가 되게 하는 조사이다.

- ☆ 의 나의 친구는 너 하나뿐이다. [나/NP+의/JKG]

5.2.5.1.5. 부사격조사(JKB)

선행 체언으로 하여금 부사어가 되게 하는 조사이다.

- ☆ (으)로 망치로 못을 박아야지. [망치/NNG+로/JKB]
- ☆ (으)로서 장관으로서 책임을 다해야 한다. [장관/NNG+으로서/JKB]
- ☆ (으)로써 돌로써 지붕을 만든다고? [돌/NNG+로써/JKB]
- ☆ 같이 바보같이 웃고 다닌다. [바보/NNG+같이/JKB]
- ☆ 더러 나더러 이것도 하라고 한다. [나/NP+더러/JKB]
- ☆ 랑 너랑 많이 닮았다. [너/NP+랑/JKB]
- ☆ (으)로부터 TV로부터 받는 영향력이 너무 크다. [TV/SL+로부터/JKB]
- ☆ 마냥 기영이마냥 놀 수만은 없다. [기영이/NNP+마냥/JKB]
- ☆ 마따나 네 말마따나 나도 그래야 한다. [말/NNG+마따나/JKB]
- ☆ 만큼 눈물만큼 콧물도 흐른다니까. [눈물/NNG+만큼/JKB]

- ☆ 보고 영자보고 놀자고 좀 해라. [영자/NNP+보고/JKB]
- ☆ 보다 직관보다는 논리가 동원돼야 한다. [직관/NNG+보다/JKB+는/JX]
- ☆ 에 나는 너에 대해 아무것도 모른다. [너/NP+에/JKB]
- ☆ 에게 너에게 말하기 싫다. [너/NP+에게/JKB]
- ☆ 에게서 나는 철수에게서 그 말을 들었다. [철수/NNP+에게서/JKB]
- ☆ 에서 집에서 학교까지 너무 멀다. [집/NNG+에서/JKB]
- ☆ 에서부터 연구소에서부터 가게까지는 너무 멀다.
[연구소/NNG+에서부터/JKB]
- ☆ 와/과 경미와 함께 다닌다면, [경미/NNP+와/JKB]
- ☆ 처럼 사람처럼 행동하는 동물이 있다. [사람/NNG+처럼/JKB]
- ☆ 하고 그 일하고 관련된 사람은 아무도 없다. [일/NNG+하고/JKB]
- ☆ 한테 그 일은 경비한테 부탁해라 [경비/NNG+한테/JKB]

5.2.5.1.6. 호격조사(JKV)

주로 사람을 가리키는 체언 뒤에 연결되어 그것으로 하여금 부름의 대상이 되게 하는 조사이다.

- ☆ 아 호동아! 이제 그만 일어나거라 [호동/NNP+아/JKV+!/SF]
- ☆ 야 철수야! 밥 먹어라 [철수/NNP+야/JKV+!/SF]
- ☆ 여 주여, 우리에게 힘을 주소서 [주/NNG+여/JKV]
- ☆ (이)시여 신이시여! 우리를 저버리지 마소서
[신/NNG+이시여/JKV+!/SS]

(1) 주의 사항

호격조사와 어말어미는 구분해서 태깅해야 한다.

- ☆ 저기 오는 것이 철수야. [철수/NNP+이/VCP+야/EF+./SF]

5.2.5.1.7. 인용격조사(JKQ)

인용문이나 인용구를, 동사에 대한 부사적 성분으로 도입하는 조사이다.

- ☆ 고 그는 "이제 가도 좋다"고 말했다. [좋/VA+다/EF+/"SS+고/JKQ]

- ☆ (이)라고 문제가 심각하다라고 보고했다.
[심각/XR+하/XSA+다/EF+”/SS+라고/JKQ]
- ☆ 하고 영수는 "이제 가자"하고 말문을 달았다. [가/VV+자/EF+”/SS+하고/JKQ]

(1) 주의 사항

① 인용격조사는 연결어미와 구별하기 어려운 경우가 있으므로 주의한다.

- ☆ 철수는 자기가 학생이라고 말했다.
 - [학생/NNG+이라고/JKQ] (×)
 - [학생/NNG+이/VCP+라고/EC] (○)
- ☆ 철수는 “다음 주에 놀러 가도 좋다”고 말하였다.
 - [좋/VV+다/EF+”/SS+고/JKQ] (○)
 - [좋/VV+다/EF+”/SS+고/EC] (×)

② 인용격조사는 형태만으로 확인할 수 없고 발화 상황까지 고려해야 하는 복잡한 표지이다. 게다가 인용격조사로 인정되는 형태인 ‘라고, 하고’ 등은 원래 용언의 활용형에 불과하다. 하지만 인용격조사를 설정하지 않을 경우에는 인용부호가 들어간 어절의 처리가 어색해진다. 따라서 우리는 인용격조사를 설정하되, 그 쓰임은 인용부호(“, ’,), ■■, },], >, ...)가 있는 경우로만 제한하기로 한다. 물론 인용부호가 빠진 경우에는 어미로 분석한다.

- ☆ 철수는 영희가 좋다고 말했다. [좋/VV+다/고/EC]

5.2.5.2. 접속조사(JC)

두 단어를 같은 자격으로 이어 주는 구실을 하는 조사를 말한다.

- ☆ 와 그 아주머니는 딸기와 사과를 샀다. [딸기/NNG+와/JC]
- ☆ 과 그 기계는 사람과 컴퓨터를 구별하지 못한다. [사람/NNG+과/JC]
- ☆ 나 사과나 배는 모두 몸에 좋은 과일이다. [사과/NNG+나/JC]
- ☆ 량 머루랑 다래랑 먹으며 청산에 살고 싶어라. [머루/NNG+랑/JC]
- ☆ 하고 이번 준비물로 칼하고 연필을 샀다. [칼/NNG+하고/JC]

(1) 주의 사항

① ‘함께 함’의 뜻을 나타내는 접속조사는 부사격조사와 형태상 동일하므로 주의할 필요가 있다.

- ☆ 철수와 영희가 왔다. [철수/NNP+와/JC]
- ☆ 철수와 같이 놀았다. [철수/NNP+와/JKB]

② 표준국어대사전에 조사로 등재(주로 구어체의 경우)된 ‘하며’는 조사로 인정하지 않고 '하/VV+며/EC'로 분석한다.

5.2.5.3. 보조사(JX)

체언이나 부사 또는 용언의 연결 어미나 종결 어미의 뒤에 쓰여 특별한 뜻을 더해 주는 조사를 말한다.

- ☆ 그러 먹습니다그러. [먹/VV+습니다/EC+그러/JX]
- ☆ 까지(꺼정/까장) 너까지 나에게 이럴 줄이야. [너/NP+까지/JX]
- ☆ 깨나 너도 사람깨나 올렸겠구나. [사람/NNG+깨나/JX]
- ☆ (이)나 너나 가라! [너/NP+나/JX]
- ☆ (이)나마 빵이나마 먹어라. [빵/NNG+이나마/JX]
- ☆ ㄴ/은/는 이 종이는 어제 사 온 것이다. [종이/NNG+는/JX]
- ☆ ㄴ커녕/은커녕/는커녕 돈은커녕 먹을 쌀도 없다. [돈/NNG+은커녕/JX]
- ☆ 다 그 물건을 거기다 놓아라. [거기/NP+다/JX]
 - 그 물건을 거기에다 놓아라. [거기/NP+에/JKB+다/JX]
- ☆ 다가 책상을 어디다가 둘까요? [어디/NP+다가/JX]
 - 집에다가 놓아 두어라. [집/NNG+에/JKB+다가/JX]
- ☆ 대로(대루) 너는 너대로 살아라. [너/NP+대로/JX]
- ☆ 따라 오늘따라 택시도 안 잡힌다. [나/NP+더러/JX]
- ☆ 도/두 강아지도 주인은 알아본다. [강아지/NNG+도/JX]
- ☆ (이)란 코알라란 호주에 사는 초식동물이다. [코알라/NNG+란/JX]
- ☆ (이)ㄹ랑 그 일에 대해선 걱정일랑 하지 말아라.
 - [걱정/NNG+ㄹ랑/JX]
- ☆ 마다 사람마다 독특한 개성이 있다. [사람/NNG+마다/JX]
- ☆ 마저 장미마저 시들고 말았다. [장미/NNG+마저/JX]
- ☆ 만 인간은 빵만으로 살 수 없다. [빵/NNG+만/JX+으로/JKB]
- ☆ 밖에 그래 봐야 죽기밖에 더 하랴. [죽/VV+기/ETN+밖에/JX]
- ☆ 부터/부텀 우선 노인부터 태워라. [노인/NNG+부터/JX]
- ☆ 뿐 가진 건 고작 집 한 채뿐. [채/NNB+뿐/JX]

- ☆ 서꺼 국물이나 총각김치서꺼 한 그릇만 주시오. [총각김치/NNG+서꺼/JX]
- ☆ (이)사 내사 그걸 이미 했지. [내/NP+사/JX]
- ☆ (이)야 그가 인간성이야 그만이지. [인간성/NNG+이야/JX]
- ☆ (이)야말로 사과야말로 가을의 과일이다. [사과/NNG+야말로/JX]
- ☆ 요 나는요 그림을요 예쁘게 그립니다. [나/NP+는/JX+요/JX]
- ☆ 조차 이젠 집조차 빼앗기는구나. [집/NNG+조차/JX]
- ☆ 치고 값싼 물건치고 쓸 만하다. [물건/NNG+치고/JX]

(1) 보조사 분석 기준

보조사는 ‘이다’의 활용어미와 구분하기 어려운 경우가 있다. 흔히 보조사로 간주되던 몇몇 형태들은 연결어미와 의미상의 차이가 없으며, 분포상으로도 구별되지 않기 때문에 이런 대상들은 보조사로 분석하지 않는다.

[기준 1] 대상 형태가 용언의 어미로 사용되는가.

[기준 2] 대상 형태가 체언에 후행할 때 서술어의 자격을 가지고 사용되는가.

(가) [기준 1, 2]에 부합하는 다음의 형태들은 모두 ‘연결어미’로 분석한다.

- ☆ (이)ㄴ들, (이)ㄴ즉, (이)든, (이)든지, (이)라도, (이)라서, (이)라야

(나) [기준 1, 2]에 부합하지 않는 다음의 형태들은 ‘보조사’가 된다.

- ☆ (이)나마, (이)야, (이)ㄹ랑, (이)야말로, (이)란

(다) [기준 1]에 부합하지 않으나, [기준 2]에는 부합하는 형태는 ‘중의성’을 가진다.

- ☆ (이)나, (이)요

(라) 다음의 형태는 서술격조사 ‘이다’의 활용형과는 관계가 없으므로 모두 보조사가 된다.

- ☆ 까지, 깨나, 는(은/ㄴ), 대로, 도, 따라, 마다, 마저, 만, 말고, 밖에, 부터, 뿐, 조차, 치고, ㄴ커녕

(2) 주의 사항

(가) 다음의 형태들은 분석 결과에 중의성이 생기므로, 이들을 분석할 때는 특히 주의해야 한다.

- ☆ (이)란 코알라란 동물은 호주에 주로 서식한다.
[코알라/NNG+이/VCP+란/ETM]
- ☆ 코알라란 매우 귀여운 동물이다. [코알라/NNG+란/JX]
- ☆ (이)나 밥이나 빵을 먹도록 해라. [밥/NNG+이나/JC]
- ☆ 그가 비록 열심히 하나 능력은 부족하다. [하/VV+나/EC]
- ☆ 어제 내가 술을 마셨나? [마시/VV+었/EP+나/EF+?/SF]
- ☆ (이)야 철수야 그 일을 할 수 있지? [철수/NNP+야/JX]
- ☆ 내가 좋아하는 것은 철수야.
[철수/NNP+이/VCP+야/EF+./SF]
- ☆ 철수야! 부르는 소리 [철수/NNP+야/JKV]
- ☆ (이)요 밥을 먹다가요 [먹/VV+다가/EC+요/JX]
- ☆ 밥이요 빵이요 [밥/NNG+이/VCP+요/EC]

(나) ‘종결어미+요(보조사)’는 종결어미로 통합하여 분석한다.

- ☆ 마음이 예쁜 사람은 좋아하는 사람이 많으니까요. [많/VA+으니까요/EF+./SF]

(다) ‘비종결어미+요(보조사)’는 통합하지 않고 각각 분석해 준다.

- ☆ 제가 몸이 좀 아파서요 지각을 했어요. [아프/VA+아서/EC+요/JX]
- ☆ 내가요, 왜요 [내/NP+가/JKS+요/JX], [왜/MAG+요/JX]

(라) ‘말고’는 용언 ‘말다’의 활용형으로 처리한다.

- ☆ 돈말고 지혜가 필요하다. [돈/NNG+말/VV+고/EC]

5.2.6. 의존형태

5.2.6.1. 어미 (EM)

5.2.6.1.1. 선어말어미(EP)

용언이 활용할 때, 어간과 어말 어미 사이에 나타나는 것으로 높임법이나 시제, 양태를 나타내는 문법적인 요소이다. 선어말어미의 목록은 연구자에 따라 다를 수 있으나 이 분석에서는 아래의 것만을 선어말어미로 인정한다.

- ✧ -겠- 그 일은 내일 처리하겠다. [처리/NNG+하/XSV+겠/EP+다/EF+./SF]
- ✧ -(으)시-선생님께서 손수 만드신 [만들/VV+시/EP+ㄴ/ETM]
- ✧ -옵- 어머니께 선물을 받치옵고 [받치/VV+옵/EP+고/EC]
- ✧ -았/었- 우리가 먹었던 음식에 문제가 있다. [먹/VV+었/EP+던/ETM]
- ✧ -았었/었었- 거기는 우리가 전에 갔었던 곳이야. [가/VV+았었/EP+던/ETM]

(1) 주의 사항

① 선어말어미가 한 음절로 통합된 경우에는 각각 분리해서 분석한다.

- ✧ -셨- 그 일은 어머니께서 하셨습니다. [하/VV+시/EP+었/EP+다/EF+./SF]

② 다음의 선어말어미는 그 어간이 생략되었을 경우에 어간을 복원해 준다.

- ✧ -겠- 이것은 그대로 두어야겠다. [두/VV+어야/EC+하/VX+겠/EP+다/EF+./SF]

- ✧ -았/었-철수가 그것을 가져오랬다. [가져오/VV+라/EC+하/VV+았/EP+다/EF+./SF]

- ✧ -시- 선생님께서 가자시오. [가/VV+자/EC+하/VV+시/EP+오/EF+./SF]

③ 위의 선어말어미가 포함되지 않은 어미 형태는 그대로 연결어미로 분석한다.

- ✧ -랴- , -대야- , -래야-

5.2.6.1.2. 종결어미(EF)

용언의 어간이나 선어말 어미 뒤에 연결되어 용언의 형식을 완성시키는 어미로서 한 문장을 끝맺는 역할을 하는 어미이다. 그러나 종결어미가 문장의 종결에만 사용되는 것은 아니고, 문맥에 따라서는 연결 어미로 쓰이기도 한다. 본 지침에서는 “SF” 앞에서만 종결어미를 인정한다.

- ✧ -게 그만한 돈이 있으면 좋게. [좋/VA+게/EF+./SF]
- ✧ -구나넌 정말 멋지구나! [멋지/VA+구나/EF+!/SF]
- ✧ -ㄴ가이것이 무엇인가? [무엇/NP+이/VCP+ㄴ가/EF+?/SF]
- ✧ -ㄴ걸이제 시작인걸. [시작/NNG+이/VCP+ㄴ걸/EF+./SF]
- ✧ -ㄴ다이건 말도 안 된다. [되/VV+ㄴ다/EF+./SF]
- ✧ -나 자네 그리로 가나? [가/VV+나/EF+?/SF]
- ✧ -냐 키가 얼마나 크냐? [크/VA+냐/EF+./SF]

- ☆ -네 정말 큰일 났네! [나/VV+았/EP+네/EF+!/SF]
- ☆ -느냐 그것보다 이것이 낫느냐? [낫/VA+느냐/EF+?/SF]
- ☆ -는걸 그는 벌써 갔는걸. [가/VV+았/EP+는걸/EF+./SF]
- ☆ -는구나 앞서 잘 안 보이는구나. [보이/VV+는구나/EF+./SF]
- ☆ -구려 당신도 가시겠구려. [가/VV+시/EP+겠/EP+구려/EF+./SF]
- ☆ -는구려 잘도 먹는구려. [먹/VV+는구려/EF+./SF]
- ☆ -는구먼 공부를 잘하는구먼.
[잘/MAG+하/XSV+는구먼/EF+./SF]
- ☆ -는다 아이가 글을 잘 읽는다. [읽/VV+는다/EF+./SF]
- ☆ -는도다 짐이 조서를 내리는도다. [내리/VV+는도다/EF+./SF]
- ☆ -다 그게 사실이다. [사실/NNG+이/VCP+다/EF+./SF]
- ☆ -ㄹ게 그렇게 할게. [하/VV+ㄹ게/EF+./SF]
- ☆ -ㅁ니까 이제야 옵니까? [오/VV+ㅁ니까/EF+?/SF]
- ☆ -ㅁ니다 이렇게 합니다. [하/VV+ㅁ니다/EF+./SF]
- ☆ -세 제대로 좀 하세. [하/VV+세/EF+./SF]
- ☆ -습니까 그래도 되겠습니까? [되/VV+겠/EP+습니까/EF+?/SF]
- ☆ -습니다 정말 재미있습니다. [재미있/VA+습니다/EF+./SF]
- ☆ -ㅁ시다 다시 만납시다. [만나/VV+ㅁ시다/EF+./SF]
- ☆ -ㅁ시오 서둘러 주십시오. [주/VX+시/EP+ㅁ시오/EF+./SF]
- ☆ -아라/어라 웃기지 말아라. [말/VX+아라/EF+./SF]
- ☆ -으냐물이 얼마나 깊으냐? [깊/VA+으냐/EF+?/SF]
- ☆ -은가그것이 좋은가? [좋/VA+은가/EF+?/SF]
- ☆ -자 밥 먹으러 가자! [가/VV+자/EF+!/SF]
- ☆ -오/으오/소 물이 깨끗하오. [깨끗/XR+하/XSA+오/EF+./SF]
- ☆ -ㅁ디다/습디다 참 좋은 곳입니다.
[곳/NNB+이/VCP+ㅁ디다/EF+./SF]
- ☆ -거든나는 이것이 좋거든! [좋/VA+거든/EF+!/SF]
- ☆ -ㄴ걸/은걸 힘이 꽤 센걸. [세/VA+ㄴ 걸/EF+./SF]
- ☆ -ㄹ걸/을걸 모른다고 할걸. [하/VV+ㄹ걸/EF+./SF]
- ☆ -ㄹ까이제 밥을 할까? [하/VV+ㄹ까/EF+?/SF]
- ☆ -다오그가 가지고 있다오. [있/VX+다오/EF+./SF]
- ☆ -다네일을 망쳤다네 [망치/VV+었/EP+다네/EF+./SF]
- ☆ -다구돈이 많다구? [많/VA+다구/EF+?/SF]
- ☆ -다니까 돈이 없다니까! [없/VA+다니까/EF+!/SF]

- ☆ -냐고/느냐고 그가 누구냐고? [누구/NP+이/VCP+냐고/EF+?/SF]
- ☆ -도다꽃이 아름답도다. [아름답/VA+도다/EF+./SF]
- ☆ -다니무엇을 읽었다니? [읽/VV+었/EP+다니/EF+?/SF]
- ☆ -는가가가 집에 있는가? [있/VV+는가/EF+?/SF]
- ☆ -부디까/습디까 보기에 좋습디까? [좋/VA+습디까/EF+?/SF]
- ☆ -다면서 술은 싫다면서? [싫/VA+다면서/EF+?/SF]
- ☆ -다나 그도 가겠다나. [가/VV+겠/EP+다나/EF+./SF]
- ☆ -렴/으렴 맘대로 해 보렴. [보/VX+렴/EF+./SF]
- ☆ -려무나 책이나 읽으려무나. [읽/VV+으려무나/EF+./SF]
- ☆ -라니까 그 사람이 아니라니까. [아니/VCN+라니까/EF+./SF]
- ☆ -자 잠이나 자자. [자/VV+자/EF+./SF]
- ☆ -세 일이나 하세. [하/VV+세/EF+./SF]
- ☆ -자구나 약속을 좀 미루자구나. [미루/VV+자구나/EF+./SF]
- ☆ -자니까 그만 따지자니까. [따지/VV+자니까/EF+./SF]
- ☆ -지 그가 언제 오지? [오/VV+지/EF+./SF]
- ☆ -아/어/여 밥 먹어! [먹/VV+어/EF+!/SF]
- ☆ -ㄴ세/음세 그날 꼭 오세. [오/VV+ㄴ세/EF+./SF]
- ☆ -단다애들이 다쳤단다. [다치/VV+었/EP+단다/EF+./SF]
- ☆ -니 그게 없니? [없/VA+니/EF+?/SF]

(1) 주의 사항

(가) ‘종결어미+요’는 통합해서 종결어미로 분석한다.

- ☆ 말씀대로 했는걸요. [하/VV+았/EP+는걸요/EF+./SF]

(나) ‘-세요’는 다음과 같이 선어말어미까지 분석한다.

- ☆ 어서 출근하세요. [출근/NNG+하/XSV+시/EP+어요/EF+./SF]

(다) ‘-죠’는 축약형을 그대로 태깅한다.

- ☆ 어서 출근하죠. [출근/NNG+하/XSV+죠/EF+./SF]

5.2.6.1.3. 연결어미(EC)

용언의 어간이나 선어말 어미 뒤에 연결되어 용언의 형식을 완성시키는 어미로서 문장을 종결시키지 못하고 뒤에 오는 절을 연결시켜 주는 어미를 말한다.

- ☆ -거나누가 오거나 알은 체 할 것 없다. [오/VV+거나/EC]
- ☆ -거니비가 오겠거니 생각했다. [오/VV+겠/EP+거니/EC]

- ☆ -건대내가 보건대, 네 말이 옳다. [보/VV+건대/EC]
- ☆ -고 일을 하고 밥을 먹자. [하/VV+고/EC]
- ☆ -곤 숙제한 것도 빌려가곤 한다. [빌리/VV+어/EC+가/VX+곤/EC]
- ☆ -느니앞아서 걱정하느니 나가서 하겠다. [걱정/NNG+하/XSV+느니/EC]
- ☆ -든지외모가 어떠하든지 무슨 상관인가? [어떠하/VA+든지/EC]
- ☆ -며 노래하며 춤을 춘다. [노래/NNG+하/XSV+며/EC]
- ☆ -면서푸르면서 검은 물빛 [푸르/VA+면서/EC]
- ☆ -자마자 오자마자 당했다. [오/VV+자마자/EC]
- ☆ -락 오르락 내리락 [오르/VV+락/EC]
- ☆ -거든가거든 말해라. [가/VV+거든/EC]
- ☆ -거늘이미 늦었거늘 어찌 빨리 가는가? [늦/VV+었/EP+거늘/EC]
- ☆ -건마는 말렸건마는 아직도 축축하다. [말리/VV+었/EP+건마는/EC]
- ☆ -고자병을 낫고자 몸부림쳤다. [낫/VV+고자/EC]
- ☆ -기에실수했기에 용서해 주었다. [실수/NNG+하/XSV+았/EP+기에/EC]
- ☆ -ㄴ다손/다손 입다손 치더라도 구박하지 말자. [입/VA+다손/EC]
- ☆ -ㄴ다기에 앞으로 잘 한다기에 승낙했다. [하/VV+ㄴ다기에/EC]
- ☆ -ㄴ데/는데 예쁘데 미워한다. [예쁘/VA+ㄴ데/EC]
- ☆ -ㄴ들/는들 간다 한들 아주 같까? [하/VV+ㄴ들/EC]
- ☆ -ㄴ즉배가 고프즉 속이 쓰리다. [고프/VA+ㄴ즉/EC]
- ☆ ㄴ지라/는지라 눈이 온지라 길이 미끄럽다. [오/VV+ㄴ지라/EC]
- ☆ -나 눈이 오나 비가 오나 같다. [오/VV+ㄴ나/EC]
- ☆ -나니멀리 보이니니 넓은 들이로다. [보이/VV+나니/EC]
- ☆ -나마맛이 좋지 못하나마 많이 드십시오. [못하/VX+나마/EC]
- ☆ -노니물노니, 포부가 무엇이냐? [물/VV+노니/EC+/,/SP]
- ☆ -는다기에 빵을 먹는다기에 주었다. [먹/VV+는다기에/EC]
- ☆ -니 멍하니 서있다. [멍하/VA+니/EC]
- ☆ -니까웃기니까 좋다. [웃기/VV+니까/EC]
- ☆ -다가자랑하다가 망신당했다. [자랑/NNG+하/XSV+다가/EC]
- ☆ -다기에 그녀가 예쁘다기에 보러 왔소. [예쁘/VA+다기에/EC]
- ☆ -대도시간이 있대도 만나 주질 않는다. [있/VV+대도/EC]
- ☆ -더라도 가더라도 꼭 돌아와라. [가/VV+더라도/EC]
- ☆ -던들진작 알았던들 방법을 취했지. [알/VV+았/EP+던들/EC]
- ☆ -도록미치도록 일했다. [미치/VV+도록/EC]
- ☆ -되 싸우되 꼭 지도록 해라. [싸우/VV+되/EC]

- ☆ -ㄹ뿐더러 비가 올뿐더러 바람도 분다. [오/VV+ㄹ뿐더러/EC]
- ☆ -ㄹ수록 갈수록 태산이다. [가/VV+ㄹ수록/EC]
- ☆ -ㄹ지비가 얼마나 올지 천둥이 다 친다. [오/VV+ㄹ지/EC]
- ☆ -ㄹ지라도 이길지라도 명예롭지는 않다. [이기/VV+ㄹ지라도/EC]
- ☆ -ㄹ지언정 죽을지언정 그 일은 못하겠다. [죽/VA+을지언정/EC]
- ☆ -라고바보라고 생각한다. [바보/NNG+이/VCP+라고/EC]
- ☆ -랍시고 그는 반장이랍시고 행패만 부린다. [반장/NNG+이/VCP+랍시고/EC]
- ☆ -러 청소하러 가자. [청소/NNG+하/XSV+러/EC]
- ☆ -려 웃으려 한다. [웃/VV+으려/EC]
- ☆ -려니와 비용도 문제려니와 일꾼도 문제다. [문제/NNG+이/VCP+려니와/EC]
- ☆ -련마는 보면 반가우련마는 볼 수가 없네. [반갑/VA+련마는/EC]
- ☆ -면 지옥이 존재하면 만원일 것이다. [존재/NNG+하/XSV+면/EC]
- ☆ -므로비가 오므로 가지 않겠다. [오/VV+므로/EC]
- ☆ -아도/어도 암만 봐도 모르겠다. [보/VV+아도/EC]
- ☆ -아서/어서 밧을 놓아서 썩을 잡았다. [놓/VV+아서/EC]
- ☆ -아야이 일은 잘해야 한다. [잘/MAG+하/XSV+아야/EC]
- ☆ -으나밥을 먹으나 마나이다. [먹/VV+으나/EC]
- ☆ -으나마 맛은 없으나마 많이 드세요. [없/VA+으나마/EC]
- ☆ -지마는 비가 오지마는 가야 한다. [오/VV+지마는/EC]
- ☆ -게 개를 굶게 하지 마라. [굶/VV+게/EC]
- ☆ -고 일단 먹고 보자. [먹/VV+고/EC]
- ☆ -아/어 입을 막아 버렸다. [막/VV+아/EC]
- ☆ -지 우기지 못해 버렸다. [우기/VV+지/EC]

(1) 주의 사항

(가) 어미에 따라서는 분석의 중의성이 생길 수 있으므로 문맥 확인을 통해 형태분석을 결정한다.

- ☆ 너는 내가 왔는데 기쁘지도 않니? [오/VV+았/EP+는데/EC]
 - 내가 지금 있는 데가 어디지? [있/VV+는/ETM+데/NNB+가/JKS]
- ☆ 다들 만족하는지 아무런 불평이 없다. [만족/NNG+하/XSV+는지/EC]
 - 너를 만난지도 꽤 오래구나. [만나/VV+ㄴ/ETM+지/NNB+도/JX]

(나) '-음직'은 “음직/EC”로 분석한다. 그러나 ‘바람직, 먹음직’ 등은 그 자체가 하나의 어근이므로 더 이상 분석할 수 없다는 것에 유의한다.

- ☆ 어른답고 믿음직하게 행동해라. [믿/VV+음직/EC+하/VX+게/EC]
- 그것 참 먹음직스럽다. [먹음직/XR+스럽/XSA+다/EF+./SF]
- 그것은 매우 바람직한 일이다. [바람직/XR+하/XSA+ㄴ/ETM]

5.2.6.1.4. 명사형 전성 어미(ETN)

한 문장의 성격을 임시로 바꾸어 다른 문장 속에서 명사적인 역할을 하게 하는 어미를 말한다.

- ☆ -기 그 일은 정말 중요하기 때문이다. [중요/XR+하/XSA+기/ETN]
- ☆ -ㄴ/-음 장사는 신용을 얻음이 제일이다. [얻/VV+음/ETN+이/JKS]

(1) 주의 사항

(가) 불규칙 용언 어간에 명사형 전성 어미가 붙어 있을 경우 ‘-음’이 아닌 ‘-ㄴ’으로 분석한다.

- ☆ 김철수 지음 [짓/VV+ㄴ/ETN]

(나) “음, 기”가 붙은 말이 단순히 명사형이냐 아니면 굳어진 명사이냐 하는 것은 물론 문맥에 따라 결정되어야 하지만 먼저 그것이 “사전”에 등재되어 있느냐의 여부를 살펴보아야 한다.

- ☆ 책을 읽기가 어렵다. [읽/VV+기/ETN+가/JKS]
- ☆ 읽기 교육이 문제가 된다. [읽기/NNG]

5.2.6.1.5. 관형사형 전성 어미(ETM)

용언의 성격을 임시로 바꾸어 다른 문장 속에서 관형사적인 역할을 하게 하는 어미이다.

- ☆ -ㄴ/은 어제 먹은 빵에 이상이 있었다. [먹/VV+은/ETM]
- ☆ -는 잃어버린 물건을 찾는 일은 요즘은 어렵다. [찾/VV+는/ETM]
- ☆ -던 이제까지 미루던 일을 오늘 해치웠다. [미루/VV+던/ETM]

- ☆ -ㄹ/을 나에게는 아직 처리할 일이 있다.
[처리/NNG+하/XSV+ㄹ/ETM]
- ☆ -런 어제런 듯하다. [어제/NNG+이/VCP+런/ETM]

(1) 주의 사항

(가) 불규칙 용언 어간에 관형사형 전성 어미가 있을 경우 ‘-은, -을’이 아닌 ‘-ㄴ, -ㄹ’로 분석한다.

- ☆ 그녀의 고운 얼굴 [곱/VA+ㄴ/ETM]
- ☆ 그녀는 매우 아름다울 것이다. [아름답/VA+ㄹ/ETM]

(나) 종결 어미에 이어서 전성 어미가 올 경우 통합해서 전성어미로 처리한다.

- ☆ 어느 쪽에 더 비중을 두느냐는 것이 [두/VV+느냐는/ETM]

5.2.6.2. 접두사(XP)

접두사는 명사와 수사에 결합하는 접사류를 묶어서 체언접두사만을 설정하기로 한다.

5.2.6.2.1. 체언 접두사(XPN)

명사 접두사에는 한자어계 접두사와 고유어계 접두사가 있는데, 그 목록의 풍부함에 비해 대개가 생산성이 그리 높지 않다. 일단 여기서는 비교적 생산성이 높다고 인정되는 접두사와, 접두사를 분리했을 경우 단일한 표제어로 등재될 수 있는 경우에 한해서 접두사 분석을하기로 한다.

- ☆ 가(假) 가건물
- ☆ 고(高) 고물가
- ☆ 과(過) 과보호
- ☆ 구(舊) 구소련
- ☆ 날 날음식
- ☆ 노(老) 노부부
- ☆ 대(大) 대선배
- ☆ 말 말아들
- ☆ 맨 맨몸
- ☆ 무(無) 무의식
- ☆ 미(未) 미완성
- ☆ 반(反) 반독재

◇ 범(汎))	범세계
◇ 부(不)	부도덕
◇ 불(不)	불합리
◇ 비(非)	비논리
◇ 생(生)	생김치
◇ 소(小)	소강당
◇ 신(新)	신정당
◇ 왕(王)	왕족발
◇ 재(再)	재충전
◇ 저(低)	저임금
◇ 제(第)	제 13 차
◇ 준(準)	준전시
◇ 초(超)	초만원
◇ 최(最)	최고급
◇ 친(親)	친러시아
◇ 탈(脫)	탈냉전시대
◇ 폐(廢)	폐광산
◇ 풋	풋살구
◇ 피(被)	피고소인
◇ 한	한가운데
◇ 헛	헛고생

5.2.6.3. 접미사(XS)

파생 접미사에는 어기의 품사를 바꾸는 것과 그렇지 않은 것이 있는데, 이들을 별도로 구별하여 표지를 부여하지는 않는다.

5.2.6.3.1. 명사파생접미사(XSN)

명사파생접미사는 명사나 다른 어근에 후행하여 그것이 명사의 기능을 수행할 수 있도록 만들어 주는 의존 형태이다. 그러나 명사파생접미사는 연구자에 따라 그 목록이 다르며, 실제로도 구분이 애매한 경우가 많다. 본 분석에서는 접미사의 생산성과 접미사를 제외한 형태의 독립성을 기준으로 다음과 같이 목록을 마련하였다.

- ✧ 가(價) 매매가
- ✧ 가(哥) 김가
- ✧ 경(頃) 두 시경
- ✧ 계(系) 몽고계
- ✧ 계(界) 교육계
- ✧ 광(狂) 메모광
- ✧ 권(圈) 운동권
- ✧ 권(權) 참정권
- ✧ 당(當) 한 사람당
- ✧ 대(臺) 역대
- ✧ 덕(宅) 청주덕
- ✧ 론(論) 비평론
- ✧ 별(別) 가구별
- ✧ 여(餘) 삼십여
- ✧ 류(類) 자연류
- ✧ 률, 율(率) 경쟁률
- ✧ 리(裡) 비밀리
- ✧ 분(分) 분량 일인분
- ✧ 분(分) 3분의
- ✧ 산(産) 중국산
- ✧ 상(上) 역사상
- ✧ 생 1(生)갑자생
- ✧ 생 2(生)견습생
- ✧ 성(性) 인간성
- ✧ 시(視) 영웅시
- ✧ 용(用) 전쟁용
- ✧ 적(的) 사상적
- ✧ 형(型) 기본형
- ✧ 형(形) 도시형
- ✧ 제(制) 봉건제
- ✧ 층(層) 선수층
- ✧ 치(值) 보름치
- ✧ 풍(風) 복고풍
- ✧ 화(化) 도구화



- ☆ 기 기름기
- ☆ 깨 10 분께
- ☆ 꿀 십 원꿀
- ☆ 끼리 전우끼리
- ☆ 끈 노름끈
- ☆ 네 동이네
- ☆ 님 선생님
- ☆ 들 우리들
- ☆ 들이 1ㄹ들이
- ☆ 배기 열 살배기
- ☆ 뺨 조카뺨
- ☆ 씩 만원씩
- ☆ 장이 간판장이
- ☆ 쟁이 심술쟁이
- ☆ 쯤 내일쯤
- ☆ 질 서방질
- ☆ 짜리 백 원짜리
- ☆ 째 1 이틀째
- ☆ 째 2 옹기째
- ☆ 치레 인사치레
- ☆ 투성이 먼지투성이



(1) 주의 사항

(가) 명사파생접미사인 ‘-들’은 그 분포가 매우 다양하여 일부에서는 이를 보조사와 접미사로 나누어 분석하기도 한다. 그러나, 본 분석에서는 이들을 모두 명사파생접미사로 처리한다. ‘먹고들’의 ‘-들’도 선행성분이 어미이긴 하나, 일치하는 대상은 선행하는 명사로 해석할 수도 있기 때문이다.

- ☆ 사람들이 우리 집에 왔다. [사람/NNG+들/XSN]
- ☆ 그들은 밥을 먹고들 싶었다. [먹/VV+고/EC+들/XSN]

(나) ‘-님’은 다음과 같이 세 가지의 분석 중의성을 가지므로 주의해서 분석한다.

- ☆ ‘임’의 의미로 쓰인 경우: 보통명사
 - 님과 이별하다. [님/NNG+과/JKB]

- ✧ 사람의 ‘이름’이나 ‘성’ 뒤에서 쓰인 경우: 의존명사
 - 김철수님께서 오셨습니다. [김철수/NNP+님/NNB+께서/JKS]
- ✧ 그 밖의 경우: 명사파생접미사
 - 과장님이 부르십니다. [과장/NNG+님/XSN+이/JKS]

5.2.6.3.2. 동사파생접미사(XSV)

동사파생접미사는 어기 또는 어근에 붙어서 그것을 동사로 만들어 주는 기능을 갖는 접미사이다. 여기서는 그러한 접미사 중 현재 생산성을 가지고 쓰이는 것만을 인정하여 분석한다.

- ✧ 당하 아군이 공격 당하는 데에는 이유가 있다.
[공격/NNG+당하/XSV+는/ETM]
- ✧ 되 아침식사가 이미 준비되어 있었다. [준비/NNG+되/XSV+어/EC]
- ✧ 시키 오늘 강아지를 운동시키려고 공원에 나갔다.
[운동/NNG+시키/XSV+려고/EC]
- ✧ 하 외국에서 공부하는 일이 쉬운 것은 아니다.
[공부/NNG+하/XSV+는/ETM]

5.2.6.3.3. 형용사파생접미사(XSA)

형용사파생접미사는 어기나 어근에 붙어서 그것을 형용사로 파생시키는 접미사이다. 여기서는 그러한 접미사 중 현재 생산성을 가지고 쓰이는 것만을 인정한다.

- ✧ 답 사람이 사람답게 행동해야 사람이지 [사람/NNG+답/XSA+게/EC]
- ✧ 되 자식된 도리로 어떻게 그런 짓을.. [자식/NNG+되/XSA+ㄴ/ETM]
- ✧ 롬 어려운 일일수록 슬기롭게 대처하라. [슬기/NNG+롬/XSA+게/EC]
- ✧ 스럽 그녀의 사랑스러운 표정을 보거라.
[사랑/NNG+스럽/XSA+ㄴ/ETM]
- ✧ 하 멍청한 표정을 짓지 말아라. [멍청/XR+하/XSA+ㄴ/ETM]

(1) 주의 사항

표준국어대사전에 따라, 기존에 접미사로 분석하던 ‘만하’의 지칭을 변경하여, ‘만’을 보조사로, ‘하’를 동사로 분석한다. ‘만하’는 ‘만/NNB+하/XSA’로 분석되는 경우도 있으므로 주의해야 한다.

- ✧ 그 일을 처리하는 데 철수만한 인재가 없다 [철수/NNP+만/JX+하/VV+ㄴ/ETM]
- ✧ 이 음식은 먹을 만하다. [먹/VV+을/ETM] [만/NNB+하/XSA+다/EF+./SF]

5.2.6.4. 어근(XR)

국어에는 하나의 단어가 조사에 의해 분리되는 현상이 있다. 즉, 파생된 용언에서 보조사 등의 삽입에 의해 어근과 접사가 분리되는 현상이 있다. 지금까지 알려진 형태소 분석 표지에서는 이들에 대해 적절한 표지를 제공하지 못한 채 어근을 명사 등의 범주 속에 포함시켜 왔다. 그러나, 세종 표준안에서는 어근 단위에 대한 표지를 설정하였으므로 이들에 어근이라는 표지를 할당할 수 있다.

- ✧ 따듯도 하다 [따뜻/XR+도/JX] [하/VV+다/EF]

(1) 주의 사항

(가) 본 표준안에서는 용언의 어근인데, 보조사와 결합되거나 독립적인 어절로 사용되는 경우에만 어근으로 분석하고, ‘하다’, ‘거리다’, ‘대다’ 등과 결합해서 용언으로 사용되는 경우는 통합형으로 전체를 용언으로 분석한다.

- ✧ 등직하다 [등직하/VV +다/EF]
- ✧ 취하다 [취하/VV+다/EF]
- ✧ 아른거렸다 [아른거리/VV + 었/EP + 다/EF]
- ✧ 따듯도 하다 [따뜻/XR+도/JX] [하/VV+다/EF]
- ✧ 마포로 정체 극심 [마포/NNP+로/JKB 정체/NNG 극심/XR]

(나) 어근으로 분석되는 대상은 표준국어대사전에 표제어로 등록된 어근만을 대상으로 한다.

5.2.7. 기호

5.2.7.1. 일반기호(ST)

영문이나 한자, 기호 등이 어절 중간에 개입하여 올바른 분석이 불가능한 경우에는 각각의 요소를 분리하여 분석한다. 이 경우 표지를 줄 수 없는 불완전한 형태가 생길 수 있다.

- ✧ 이크로소프트(microsoft)사 [마이크로소프트/NNP+ (/SS + microsoft/SL +)/SS + 사/NNG]
- ✧ 농.수산물 [농/NNG+ ./SP+수산물/NNG]
- ✧ 초.중.고 [초/NNG + ./SP + 중/NNG + ./SP + 고/NNG]

5.2.7.1.1. 마침표, 물음표, 느낌표(SF)

SF 대상은 다음과 같다. 기본적으로 1 바이트 심볼이 대상인데, 유니코드와 그래픽 문자로 존재하는 다른 코드이지만 형태가 동일한 심볼도 대상으로 고려한다.

- ✧ 마침표: .
- ✧ 물음표: ?
- ✧ 느낌표: !

(1) 주의 사항

- (가) ‘.’의 경우에는 문장 종료를 나타내는 경우에만 SF 로 분석하고, 나열형이나 숫자 중간에 사용되는 경우에는 SP 로 분석한다.
- (나) ‘?’와 ‘!’의 경우에는 문장 종료를 나타내는 경우와 이모티콘이나 다른 의미로 문장 중간에서 사용되는 경우에도 모두 SF 로 분석한다.

5.2.7.1.2. 쉼표, 가운뎃점, 콜론, 빗금(SP)

SP 대상은 다음과 같다. 기본적으로 1 바이트 심볼이 대상인데, 유니코드와 그래픽 문자로 존재하는 다른 코드이지만 형태가 동일한 심볼도 대상으로 고려한다.

- ✧ 쉼표: ,
- ✧ 가운뎃점: . .
- ✧ 콜론: : : :
- ✧ 빗금: / /

5.2.7.1.3. 따옴표, 괄호표, 줄표(SS)

SS 대상은 다음과 같다. 기본적으로 1 바이트 심볼이 대상인데, 유니코드와 그래픽 문자로 존재하는 다른 코드이지만 형태가 동일한 심볼도 대상으로 고려한다.

- ✧ 따옴표: “ ‘ ` ~ ”

◇ 괄호표: () { } [] < > < > 《 》 「 」 『 』 【 】 [] < >

◇ 줄표: _ - - - _

5.2.7.1.4. 줄임표(SE)

SE 대상은 다음과 같다. 기본적으로 1 바이트 심볼이 대상인데, 유니코드와 그래픽 문자로 존재하는 다른 코드이지만 형태가 동일한 심볼도 대상으로 고려한다.

◇ 줄임표: .. … …

(1) 주의 사항

(가) 동일한 줄임표 기호가 동일한 어절에 연속적으로 나오는 경우에는 전체를 1 개의 형태소로 분석한다.

◇ 맞춘다.. [맞추/VV + ㄴ다/EF + ../SE]

5.2.7.1.5. 붙임표(물결)(SO)

SO 대상은 다음과 같다. 기본적으로 1 바이트 심볼이 대상인데, 유니코드와 그래픽 문자로 존재하는 다른 코드이지만 형태가 동일한 심볼도 대상으로 고려한다.

◇ 물결: ~

5.2.7.1.6. 기타 기호(SW)

SW 는 일반 기호 하위 분류의 세분화된 기호 외의 나머지 기호를 모두 대상으로 한다. 세부기호에 포함되지 않는 그래픽문자와 이모티콘 등도 모두 기타기호에 해당된다.

5.2.7.2. 외국어(SL)

SL 대상은 한글, 한자를 제외한 모든 외국어가 대상이 된다. 예를 들면, 영어, 일본어 등이다. SL 은 각 외국어의 의미와 상관없이 동일한 어절에 연속적으로 나타난 동일 외국어를 1 개의 형태소로 분석한다.

5.2.7.3. 한자(SH)

SH 는 모든 한자를 대상으로 한다. 한자의 의미와 상관없이 동일한 어절에 연속적으로 나타난 한자를 1 개의 형태소로 분석한다.

5.2.7.4. 숫자(SN)

SN 은 모든 숫자가 대상이 된다. 기본적으로 1 바이트 숫자가 대상인데, 유니코드와 그래픽 문자로 존재하는 다른 코드이지만 형태가 동일한 숫자도 대상으로 고려한다. 동일한 어절에 연속적으로 나타난 숫자를 1 개의 형태소로 분석한다.

5.2.7.5. 분석 불능 범주(NA)

그 자체가 사전에 등재되어 있지도 않으면서, 축약의 정도가 심하거나 분석하기 어려운 방언형의 경우 분석 불능 범주로 처리한다.

✧ 담배가 꼬매턴게 하마 자라서 빼나? [꼬매턴게/NA]

6. 기타

6.1. 준말

준말은, 그것이 본딤말과 대등하게 사용되고 분석 결과가 동일한 어절 단위를 형성할 경우에 한해서만 복원한다. 그러나 다음에서처럼, 본딤말로 복원할 경우 어절 수에 변화가 생길 뿐 아니라 본딤말로 복원하는 정도가 일관성을 띠지 않게 되는 경우는 굳이 복원하지 않는다. 그러나, 이러한 원칙이 모든 경우에 일관적으로 적용될 수 있는 것은 아니다. 결국 준말의 처리는 해당 어절에 따라 임의적일 수 있다.

✧ 라는 [라는/ETM] (○)

✧ [라고/JKQ+하/VV+는/ETM] (×)

6.2. 합성어

합성어는 표준국어대사전에 등재되어 있는 것만을 인정한다.

(1) 주의 사항

(가) 표제어가 사전의 표제어로 등록되어 있는 경우는 그대로 분석한다.

☆ 정치권력 (사전: 정치-권력) [정치권력/NNG]

(나) 합성어로 등재되어 있되 띄어쓰기를 허용한 합성어는 세분하여 분석하는 것을 원칙으로 한다.

☆ 학생운동 (사전표기: 학생^운동) [학생/NNG+운동/NNG]

(다) 합성어로 등록되어 있지 않은 표제어는 분리해서 분석하되, 사전 표제어로 등록되어 있는 최대한 많은 음절수의 단어를 생성하도록 나눈다. 즉 다음 예와 같은 경우 3음절 어휘가 생성되는 첫 번째 분석을 취한다.

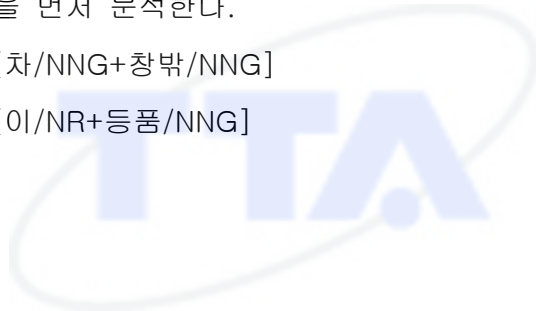
☆ 언어학과 [언어학/NNG+과/NNG] (3음절+1음절)

☆ 언어학과 [언어/NNG+학과/NNG] (2음절+2음절)

(라) 3음절 어휘와 같이 어느 쪽으로 나뉘어도 음절수가 같고, 양쪽 분석이 모두 사전 표제어라면 뒤쪽을 먼저 분석한다.

☆ 차창밖 [차/NNG+창밖/NNG]

• 이등품 [이/NR+등품/NNG]



부 록 |

참고 문헌

- [1] 21세기 세종계획, “형태분석 말뭉치 구축 지침(Ver. 2005-1)”, 2005.
- [2] 한국정보통신기술협회(TTA), “품사 부착 말뭉치 작성용 품사 태그 세트 표준(TTA.KO-11.0010)”, 1999.
- [3] 서울대학교, “꼬꼬마 형태소 분석기”, <http://kkma.snu.ac.kr/>
- [4] 카이스트, “한나눔 한국어 형태소 분석기”, <http://kldp.net/projects/hannanum>
- [5] 권오욱, 정유진, 김미영, 류동원, 이문기, 이종혁, “음절단위 CYK 알고리즘에 기반한 형태소 해석기 및 품사태거”, 제1회 한국어 형태소 분석기 및 품사태거 평가 워크숍 논문집, 76-88 (1999)
- [6] 조용환, “형태소분석기 MORAN”, <http://www.moransoft.co.kr/>
- [7] 윤준태, 이충희, 김선호, 송만석, “연세대 형태소 분석기 morany : 말뭉치로부터 추출한 대량의 어휘 데이터베이스에 기반한 형태소 분석”, 1999년도 제11회 한글 및 한국어 정보처리 학술대회 및 제1회 형태소 분석기 및 품사태거 평가 워크숍, pp.92-98, 1999.
- [8] 김준석, 심준혁, 이근배, “품사 태그 세트의 매핑을 이용한 한국어 품사 태거(POSTAG) 이식”, 1999년도 제11회 한글 및 한국어 정보처리 학술대회 및 제1회 형태소 분석기 및 품사태거 평가 워크숍, pp. 484-490, 1999.
- [9] 심광섭, “초고속 한국어 형태소 분석기 MACH 1.0/2.0”, <http://cs.sungshin.ac.kr/~shim/demo/mach.html>
- [10] 강승식, “HAM: 한국어 형태소 분석 라이브러리”, <http://nlp.kookmin.ac.kr/HAM/kor/ham-intr.html>

표준 작성 공헌자

표준 번호 : TTAK.KO-11.0010/R1

이 표준의 제정·개정 및 발간을 위해 아래와 같이 여러분들이 공헌하셨습니다.

구분	성명	위원회 및 직위	연락처	소속사
표준(과제) 제안	최미란	메타데이터 프로젝트그룹 위원	miranc@etri.re.kr	ETRI
표준 초안 작성자	최미란	메타데이터 프로젝트그룹 위원	miranc@etri.re.kr	ETRI
표준 초안 에디터	최미란	메타데이터 프로젝트그룹 위원	miranc@etri.re.kr	ETRI
표준 초안 검토	서태설	메타데이터 프로젝트그룹 의장	tsseo@kisti.re.kr	KISTI
		외 프로젝트그룹 위원		
표준안 심의	박승민	소프트웨어/콘텐츠 기술위원회 의장	minpark@etri.re.kr	ETRI
		외 기술위원회 위원		
사무국 담당	김영화	부장	ykim@tta.or.kr	TTA
	이상호	선임	sangho.lee@tta.or.kr	TTA



정보통신단체표준(국문표준)

형태소 태깅 말뭉치 작성용 품사 태그 세트
(Part-of-Speech Tag Set for
Morphological Annotation of Written Texts)

발행인 : 한국정보통신기술협회 회장

발행처 : 한국정보통신기술협회

463-824, 경기도 성남시 분당구 분당로 47

Tel : 031-724-0114, Fax : 031-724-0109

발행일 : 2015.06.
