

변형 규칙 기반 한국어 품사 태거의 개선*

임희석[†], 김진동, 임해창

고려대학교 전산과학과 자연어처리 연구실

Improvement of Transformation Rule-Based Korean Part-Of-Speech Tagger

Heui-Seok Lim[†], Jin-Dong Kim, Hae-Chang Rim

Natural Language Processing Lab., Dept. of Computer Science, Korea Univ.

Abstract

변형 규칙 기반 품사 태거는 태깅 규칙을 코퍼스로부터 자동 학습할 수 있고, 견고하며 태깅 결과를 이해하고 분석하기가 쉽다는 장점을 갖는다. 이에 최근 한국어 특성을 고려한 변형 규칙 기반 한국어 품사 태거가 개발되었다. 하지만 이 시스템은 오류 어절의 어휘 정보를 사용하지 않으므로 수정 가능 오류에 대한 변형 규칙이 제대로 학습되지 못하며, 변형 규칙 적용 과정에 새로운 오류를 발생시킨다는 문제점이 있다. 이에 본 논문은 오류 어절의 어휘 정보를 참조할 수 있는 세부 변형 규칙 추출을 이용한 변형 규칙 기반 한국어 품사 태거의 개선 방안을 제안한다. 어휘 정보를 참조할 수 있는 세부 변형 규칙의 형태는 특정 문맥 C에서 어절 W의 어절 태그 α 를 어절 태그 β 로 변형한다와 같다. 제안된 방법은 약 10만 어절 크기의 학습 코퍼스에서 57개의 세부 규칙을 학습하였고, 2만 어절 크기의 실험 코퍼스에 적용한 결과 95.6%의 정확도를 보임으로써 기존의 변형 규칙 기반 품사 태거의 정확도를 약 15.4% 향상시켰다.

1 변형 규칙 기반 한국어 품사 태거

변형 규칙 기반 품사 태거는 태깅 규칙을 코퍼스로부터 자동 학습할 수 있고, 견고하며, 태깅 결과를 이해/분석하기 쉽다는 장점을 갖는다[1, 2, 3]. 이에 최근 한국어 특성을 고려한 변형 규칙 기반 한국어 품사 태깅 시스템이 개발되었는데 [12], 이 시스템은 첨가어적인 한국어

특성을 고려할 수 있는 변형 규칙들과 변형 규칙을 사용하였다[11, 12]. 한국어 특성을 고려할 수 있도록 제안된 변형 규칙은 특정 문맥 C에서 어절 태그 α 를 어절 태그 β 로 변형한다와 같이 어절 태그 변형 규칙이다.

어절 태그 변형 규칙은 어절을 구성하는 형태소 품사 열로 정의되는 어절 태그를 수정함으로써, Brill이 사용한 변형 규칙 형태로 처리할 수 없는 한 어절내 두 가지 이상의 형태소 태깅 오류를 수정할 수 있다. 예를 들면, “내가 먹은¹ 감”에서 ‘먹은’이 먹(명사)+은(조사)로 잘못 태깅 되었을 때, 특정 문맥 C에서 어절 태그, ‘명사+조사’를 어절 태그, ‘동사+어미’로 변형한다와 같은 변형 규칙은 한 어절내 발생한 두가지의 형태소 태깅 오류를 수정할 수 있다.

변형 규칙 기반 한국어 품사 태깅 시스템에서 사용한 초기 태거는 Twoply HMM[6]에서 어절 발생 확률을 계산하기 위하여 사용한 어절내 HMM을 이용하였다. 어절내 HMM은 어절내 형태소 전이 확률과 형태소 발생 확률을 마르코프 모델에 적용하여 어절의 발생 확률을 계산한다. 어절내 HMM을 이용한 초기 태거는 338,016개의 형태소 크기의 학습 코퍼스를 이용하여 학습시켰으며, 실험 코퍼스에서 약 91.6%의 정확도를 보였다.

변형 규칙 생성을 위하여 사용된 규칙들은 크게 어절 태그 문맥, 형태소 품사 문맥, 그리고 어휘 문맥을 고려할 수 있는 형태로 나뉘어진다. 어절 태그 문맥은 주위 어절의 어절 태그가 현재 어절 태그에 미치는 영향을 고려하기 위한 것이며, 형태소 품사 문맥은 이웃 어절의 형식 형태소나 실질 형태소가 특정 어절에 미치는 문

*본 연구는 과학재단의 특정 연구 과제 “한국어 이해에 나타나는 중의성 문제 처리 모델에 관한 연구”의 지원을 받은 것입니다.

¹ 먹(명사)+은(조사)

[표 1] 태깅 오류의 유형

번호	잘못 태깅된 결과	올바른 태그	번호	잘못 태깅된 결과	올바른 태그
1	NNCV	NNCG	11	MAC	MAA
2	VJ+EFD	VX+EFD	12	VX+EP+EFF+SS.	VV+EP+EFF+SS.
3	VX+EFD	VV+EFD	13	VV+EFD	VV+EFD
4	CANNOT	VV+EFC	14	CANNOT	VV+EFD
5	NNB	NNCG	15	MAA	NNCG
6	VX+EFF+SS.	VV+EFF+SS.	16	VV+EFD	NNB+JX
7	VV+EFC	VX+EFC	17	NPP+JCD	MDI
8	NNCG+JCA	NNB+JCA	18	NNB+JCA	NNCG+JCA
9	VX+EFC	VV+EFC	19	VV+EFC	VV+EFC+VX+EFC
10	NNB+JX	NNCG+JX	20	MAC	VV+EFC

정부는 방관자로 나왔어	[있는]:[있(VJ)+는(EFD)]	형편이다.
농약으로 목숨을 잃고	[있는]:[있(VJ)+는(EFD)]	실정이다.
아침이면 집집마다 살아	[있는]:[있(VJ)+는(EFD)]	미이러가 걸어서 혹은 ...
농약 규정에 들어	[있는]:[있(VJ)+는(EFD)]	것만도 27종으로 ...
그나마 서울 가	[있는]:[있(VJ)+는(EFD)]	형들 뭇을 때면 ...
적십자 회원증 가지고	[있는]:[있(VJ)+는(EFD)]	농민은 없다.

[그림 1] 변형 규칙에 의해 수정될 수 없는 오류의 예

법적 영향을 고려하기 위한 규칙 틀이다. 어휘 문맥은 주변 어절의 어절을 참조할 수 있도록 하기 위함이다.

변형 규칙 기반 한국어 품사 태깅 시스템[12]은 57개의 품사 집합을 사용하여 10만 어절 크기의 학습 코퍼스에서 학습한 결과, 101개의 변형 규칙을 추출하였다. 101개의 변형 규칙 중 53개가 형태소 품사 문맥을 고려한 규칙 틀에 의해서 만들어진 것이었으며, 30개의 변형 규칙이 어절 태그 문맥을 이용한 규칙 틀에 의해서 만들어진 것이었다. 나머지 18개의 변형 규칙은 주변 어절의 어휘를 참조할 수 있는 변형 규칙이었다. 101개의 변형 규칙을 약 2만 어절 크기의 실험 코퍼스에 적용한 결과 변형 규칙 기반 한국어 품사 태거[12]는 약 94.8%의 정확도를 보였다.

2 변형 규칙 기반 한국어 품사 태거의 문제점

한국어 특성을 고려한 변형 규칙 기반 한국어 품사 태거는 실험 코퍼스에서 약 94.8%의 비교적 높은 정확도를 보였으나, 태깅 결과를 분석한 결과 많은 수정 가능 오류

와 변형 규칙의 잘못된 적용으로 인한 오류를 포함하고 있었다. 수정 가능 오류란 변형 규칙에 의해서 수정될 것으로 기대되는 오류이지만, 그 오류를 수정할 수 있는 변형 규칙이 학습되지 않아 품사 태깅 이후에도 남아 있는 오류를 의미한다. [표 1]은 변형 규칙 기반 한국어 품사 태거의 상위 20개의 태깅 오류를 나타낸 것이다².

[표 1]에서 ‘잘못된 태깅 결과’ 열은 변형 규칙 기반 한국어 품사 태거가 잘못 태깅한 결과이며, ‘올바른 태그’ 열은 해당 어절의 올바른 어절 태그를 나타낸 것이다. [표 1]의 20가지 오류 유형은 전체 오류 유형 중에서 약 4%에 해당하지만, 이들은 전체 오류의 약 47.8%의 많은 비율을 차지하는 것으로 평균 25회 이상의 오류를 갖는 것이었다.

태깅 오류는 크게 초기 태깅의 오류와 변형 규칙 적용시의 오류로 구분할 수 있다. 초기 태깅의 오류는 변형 규칙에 의해 수정되지 못한 수정 가능 오류를 의미한다. 예를 들면, 연결어미(EFC) 뒤 어절 ‘있는’은 있(VX)+는(EFD)로 태깅되어야 한다. 따라서 초기 태거에 의해 어절, ‘있는’이 있(VJ)+는(EFD)로 잘못

² 품사 집합은 부록을 참조

태깅되었다면, 연결어미 뒤의 어절 태그 $VJ+EFD$ 를 어절 태그 $VX+EFD$ 로 변형한다와 같은 변형 규칙에 의해 코퍼스 내의 연결 어미 뒤에 있는 모든 $VJ+EFD$ 어절 태그가 $VX+EFD$ 로 수정될 것이 기대된다. 하지만 코퍼스 내의 $VJ+EFD$ 를 갖는 모든 어절의 어절 태그를 $VX+EFD$ 로 변형할 경우, 새롭게 발생하는 오류가 오류 감소정도보다 많아지므로 연결어미 뒤의 어절 태그 $VJ+EFD$ 를 어절 태그 $VX+EFD$ 로 변형한다와 같은 변형 규칙은 생성되지 못한다. 따라서 $VJ+EFD$ 로 잘못 태깅된 연결어미(EFC) 뒤의 모든 '있는'은 변형 규칙을 이용한 태깅 후에도 오류가 수정되지 못하고 존재하게 된다. [그림 1]은 위에서 설명한 '있는'과 같은 변형 규칙에 의해서 수정될 수 없는 초기 태깅의 오류의 예를 나타낸 것이다.

변형 규칙 적용시의 오류는 변형 규칙에 의해 잘못 변형된 어절 태그를 수정할 수 있는 변형 규칙의 부재로 인하여 발생하는 것이다. 학습 과정에서 추출된 변형 규칙은 그 규칙이 적용되는 모든 어절에 대해서 항상 올바른 결과를 만드는 것은 아니다.³ 예를 들어, 변형 규칙, 앞 어절의 형식 형태소가 연결 어미(EFC)일 때, 어절 태그 $VV+EFD$ 를 어절 태그 $VX+EFD$ 로 변형한다와 같은 규칙은 연결 어미 뒤의 $VV+EFD$ 로 잘못 태깅된 많은 오류를 $VX+EFD$ 로 수정할 수 있는 규칙이다. 그러나 이 규칙은 “점 1통과 바꾸(VV)+어(EFC) 먹은(VV+EFD) 셈이다.”에도 적용되어 어절 '먹은'의 어절 태그가 $VX+EFD$ 로 잘못된 어절 태그로 변형되게 된다. 그러므로 변형 규칙 적용후 발생한 오류를 수정하기 위해서는 그 오류를 수정할 수 있는 새로운 변형 규칙이 필요하다. 하지만 새로운 변형 규칙이 학습되지 않을 경우 그 오류는 태깅후에도 여전히 남아있게 된다.

3 개선 방안

초기 태깅의 오류와 변형 규칙 적용시의 오류를 가지고 있는 기존의 변형 규칙 기반 한국어 품사 태깅 시스템은 많은 개선의 여지를 가지고 있다. 첫째, 초기 태깅의 오류를 수정할 수 있는 변형 규칙 학습이 가능하도록 하는 것이며, 둘째, 변형 규칙 적용시 발생하는 오류를 제거하기 위한 변형 규칙을 학습하도록 하는 것이다. 본 논문은 오류 어절의 어휘 정보를 사용하는 세부 변형

³ 변형 규칙은 학습 과정에서 규칙을 적용했을 때 발생한 오류 감소 개수와 새롭게 발생한 오류 수의 차이가 특정 임계값 이상인 규칙이다

규칙을 이용하여 초기 태깅의 오류와 변형 규칙 적용시 발생하는 오류를 감소시키고자 한다.

기존 변형 규칙 기반 한국어 품사 태깅이 사용한 변형 규칙은 특정 문맥 C에 속하는 어절 태그 α 를 어절 태그 β 로 변형하므로 변화될 어절의 정보를 전혀 사용하지 않는 매우 일반적인 규칙이었다. 예를 들면, 연결어미 뒤의 어절 태그 $VJ+EFD$ 를 어절 태그 $VX+EFD$ 로 변형한다는 어절 '있는'에만 적용되는 것이 아니라 연결어미 뒤의 $VJ+EFD$ 로 태깅된 모든 어절에 적용되는 일반적인 규칙이다. 따라서, 이 규칙은 연결 어미 뒤의 모든 '있는'에 대해서는 올바른 결과를 만들어 낼 때도 불구하고 다른 어절에서는 오류를 발생시키므로 연결어미 뒤의 '있는'을 수정할 수 있는 변형 규칙으로 학습되지 못한다.

연결 어미 뒤의 '있는'을 수정할 수 있는 변형 규칙을 학습하기 위해서는 어절 '있는'의 어휘 정보를 고려할 수 있는 세부 변형 규칙을 학습할 수 있도록 하여야 한다. 예를 들면, 연결 어미의 뒤에서 $VJ+EFD$ 로 태깅된 '있는'의 어절 태그를 $VX+EFD$ 로 변형한다와 같이 오류 어절의 어휘 정보를 사용하는 세부 규칙을 학습하는 것이다.

오류 어절의 어휘 정보를 사용하면 '있는' 이외에 연결 어미 뒤에서 $VJ+EFD$ 로 태깅된 다른 어절에 영향을 미치지 않는 연결 어미의 뒤에서 $VJ+EFD$ 로 태깅된 '있는'의 어절 태그를 $VX+EFD$ 로 변형한다와 같은 변형 규칙 학습이 가능하다.

오류 어절의 어휘 정보를 고려하면 일반 규칙이 올바르게 적용된 다른 어절에 영향을 미치지 않으면서 변형 규칙 적용시의 오류 어절들만을 수정할 수 있는 변형 규칙 추출이 가능하다.

본 논문에서 어절의 어휘 정보를 고려하기 위하여 사용한 규칙들의 형식은 다음과 같다.

규칙틀 : 특정 문맥 C에서 어절 W의 어절 태그 α 를 어절 태그 β 로 변형한다.

세부 변형 규칙 학습을 위해 사용된 문맥은 [12]에서 사용한 것과 동일한 어절 태그 문맥, 형태소 태그 문맥, 그리고 어휘 문맥을 사용하였다. [그림 2]는 세부 변형 규칙 학습을 위해 사용된 규칙틀의 예를 보인 것이다.

세부 변형 규칙의 학습은 [12]에서 일반 규칙 학습과정에 임계값보다 높은 오류 감소를 보이지 않을 때까지 변화된 10만 어절 크기의 학습 코퍼스에 대하여 수행되었다⁴. 세부 규칙 학습 결과 57개의 세부 변형 규칙이

⁴[그림 2]의 규칙틀을 일반 규칙 학습을 위한 규칙틀과 같이 사

- 1) 현재 어절이 W이고, 앞(뒤) 어절 태그가 z인 경우
 - 2) 현재 어절이 W이고, 앞(뒤) 2번째 어절 태그가 z1이고 앞(뒤) 1번째 어절 태그가 z2인 경우
 - 3) 현재 어절이 W이고, 앞(뒤)의 두 어절 중 한 어절의 어절 태그가 z인 경우
 - 4) 현재 어절이 W이고, 앞 어절의 어절 태그가 z1이고 뒤 어절의 어절 태그가 z2인 경우
- 5) 현재 어절이 W이고, 앞(뒤) 어절의 형식 형태소(실질 형태소)의 품사가 z인 경우
 - 6) 현재 어절이 W이고, 앞 어절의 형식 형태소의 품사가 z1이고 뒤 어절의 실질 형태소 품사가 z2인 경우
- 7) 현재 어절이 W이고, 앞(뒤) 어절이 z인 경우
 - 8) 현재 어절이 W이고, 앞(뒤) 2번째 어절이 z1이고 앞(뒤) 1번째 어절이 z2인 경우

[그림 2] 세부 변형 규칙 학습을 위해 사용된 규칙 틀의 예

[표 2] 학습된 세부 변형 규칙의 예

번호	오류 어절	변경전	변경후	조건
1	있는	VJ+EFD	VX+EFD	[-2 어절 태그] OR [-1 어절 태그] VV+EFC
2	경우	NNB	NNCG	[-1 Tail] JCD
3	바람에	NNCG+JCA	NNB+JCA	[-1 어절 태그] VV+EFD
4	간	NNB	VV+EFD	[1 Head] NNB
5	지난	VV+EFD	MDA	[1 어절 태그] SCD+NNBU
6	하는데	VV+EFC	VX+EFC	[-1 어절 태그] VJ+EFC
7	볼	VX+EFD	VV+EFD	[1 Head] NNB

학습되었으며, [표 2]는 학습된 세부 변형 규칙의 예를 보인 것이다.

세부 변형 규칙 1은 앞에서 첫번째 또는 두번째 어절 태그가 VV+EFC 일때 어절, '있는'의 어절 태그, '형용사(VJ)+연결어미(EFC)'를 어절 태그 '보조용언(VX)+연결어미(EFC)'로 변형하여라를 의미하는 것으로, 기존의 일반 규칙 학습으로는 얻을 수 없는 세부 규칙이다. 2번 규칙은 관형형전성어미(JCD) 뒤에 의존 명사(NNB)로 잘못 태깅된 어절 '경우'의 어절 태그를 보통 명사(NNCG)로 변형하여라를 규칙이다.

4 실험 및 평가

본 논문은 세부 변형 규칙 사용의 타당성을 검증하기 위하여 [12]에서 사용된 코퍼스를 실험 코퍼스로 사용하였다. 실험 코퍼스의 크기는 약 2만 어절이며, 신문 기사, 소설 그리고 수필에서 추출된 어절들로 구성되어 있다. 실험 코퍼스의 약 42% 어절이 2개 이상의 형태소 분석이 가능한 중의성을 갖는 어절이었으며, 어절 당 평균 중의적 분석 개수는 2.4개였다.

용하면 일반 규칙의 학습과 세부 규칙 학습을 동시에 수행할 수도 있다.

실험 결과, 제안된 방법은 어절 단위의 정확도 평가에 의해서 95.6%의 정확도를 보였다. 이는 기존의 변형 규칙 기반 품사 태깅의 정확도를 15.4% 증가시킨 결과이며, 이와 같은 정확도 향상의 원인은 오류 어절의 어휘 정보를 참조하는 세부 규칙을 사용한 결과임을 알 수 있었다.

[표 3]은 [12]에서 학습된 일반 규칙만을 이용한 품사 태깅의 정확도와 일반 규칙과 세부 규칙을 모두 사용한 품사 태깅의 정확도를 보인 것이다.

[표 3] 실험 결과

	규칙 개수	정확도
초기태거(I)	0	91.6%
I+일반 규칙(G)	101	94.8%
I+G+세부규칙	158	95.6%

5 결론

본 논문은 어절의 어휘 정보를 사용하지 않아 수정 가능 오류와 변형 규칙 적용시 오류를 발생시켰던, 기존의 변형 규칙 기반 한국어 품사 태깅의 개선 방안을 제안하였다.

제안된 방법은 오류 어절의 어휘 정보를 참조할 수 있는 세부 변형 규칙 학습이 가능하도록 하고, 이를 품사 태깅시에 사용하는 것이다. 오류 어절의 어휘 정보를 참조하는 세부 변형 규칙은 오류 어절의 특정 문맥에 맞는 변형 규칙 추출이 가능하므로 수정 가능 오류를 줄일 수 있었다. 또한 제안된 방법은 오류 어절들만을 위한 변형 규칙을 학습하므로, 일반 변형 규칙이 올바르게 적용된 어절들에 영향을 미치지 않으면서 코퍼스내 오류를 감소시킬 수 있었다.

제안된 방법은 일반 변형 규칙 학습에 사용된 10만 어절 크기의 학습 코퍼스를 이용하여 학습한 결과, 57개의 세부 변형 규칙을 학습하였다. 일반 변형 규칙 101개와 세부 변형 규칙 57개를 이용하여 2만 어절 크기의 실험 코퍼스에 적용한 결과, 제안된 방법은 일반 변형 규칙만을 사용한 품사 태거와 비교하여 15.4%의 정확도 향상을 보였다.

어휘 정보를 이용한 세부 변형 규칙은 오류 어절 단위의 규칙을 학습하게 된다. 따라서 오류 빈도가 높은 어절에 대한 세부 규칙만을 학습하게 되고, 통사적으로 비슷한 성격을 갖는 어절에 적용될 수 있는 일반성을 갖는 변형 규칙을 학습하지 못한다는 단점이 있다. 따라서 향후에는 일반 규칙과 어절 단위의 세부 변형 규칙의 중간 수준의 일반성을 갖을 수 있는 변형 규칙 추출에 대한 연구가 수행되어야 할 것이다.

References

- [1] Brill, Eric, "A simple rule-based part-of-speech tagger," *Proc. of Third Conf. on Applied Natural Language Processing*, ACL, Trent, Italy, pp.152-155, 1992.
- [2] Brill, Eric, "Some advances in rule-based part-of-speech tagging," *Proc. of Twelfth National Conf. on Artificial Intelligence*, Seattle, WA, pp.722-727, 1994.
- [3] Brill, Eric, "Transformation-based error-driven learning and natural language processing: A case study in part-of-speech tagging," *Computational Linguistics*, Vol.21, No.4, pp.543-565, 1995.
- [4] Klein, S. and Simmons, R.F., "A computational approach to grammatical coding of English words," *JACM*, Vol.10, pp.334-347, 1963.
- [5] Marcus Mitch, "Statistical natural language processing: Current trends and future directions," *Proc. of ATR Int. Workshop on Speech Translation*, 1993.
- [6] 김진동, 임희석, 임해창, "어절 단위의 문맥을 고려한 형태소 단위의 한국어 품사 태깅 모델," *인지과학회 춘계 학술대회 발표 논문집*, pp.97-106, 1996.
- [7] 신상현, 이근배, 이종혁, "TAKTAG: 통계와 규칙에 기반한 2단계 학습을 통한 품사 중의성 해결," *제 7회 한글 및 한국어정보처리 학술대회 발표 논문집*, pp.169-174, 1995.
- [8] 이호, 김진동, 임해창, "언어 정보 획득을 위한 한국어 코퍼스 분석 도구," *제 6회 한글 및 한국어정보처리 학술대회 논문집*, pp.297-304, 1994.
- [9] 임철수, *HMM을 이용한 한국어 품사 태깅 시스템 구현*, 한국과학기술원 전산학과 석사학위 논문, 1994.
- [10] 임해창, 임희석, 윤보현, "자연어처리 연구동향: 통계 기반의 자연어 처리," *한국정보과학회지*, 제 12권 제 9호, pp.20-30, 1994.
- [11] 임해창, 임희석, 이상주, 김진동, "자연어처리를 위한 품사 태깅 시스템의 고찰," *한국정보과학회지*, 제 14권 제 7호, pp.36-57, 1996.
- [12] 임희석, 김진동, 임해창, "한국어 특성에 적합한 변형 규칙 기반 한국어 품사 태깅," *춘계 인공지능연구회 학술발표 논문집*, pp.3-10, 1996.

부록 : 태깅에 사용된 품사 집합

[채언류]

1. NNCG (보통명사)
2. NNCV (동작성 보통명사)
3. NNCJ (상태성 보통명사)
4. NNP (고유명사)
5. NNB (의존명사)
6. NNBU (단위성 의존명사)
7. NPP (인칭 대명사)
8. NPI (지시 대명사)
9. NU (수사)

[조사류]

- 10. JC (격조사)
- 11. JCO (목적격 조사)
- 12. JCD (관형격 조사)
- 13. JCA (부사격 조사)
- 14. JCV (호격 조사)
- 15. JCP (서술격 조사)
- 16. JN (접속 조사)
- 17. JX (보조사)

[용언류]

- 18. VV (동사)
- 19. VJ (형용사)
- 20. VX (보조용언)

[어미류]

- 21. EFC (연결어미)
- 22. EFD (관형사형 전성어미)
- 23. EFN (명사형 전성어미)
- 24. EFA (부사형 전성어미)
- 25. EP (선어말어미)
- 26. EFF (종결어미)

[관형사류]

- 27. MDI (지시 관형사)
- 28. MDU (수 관형사)
- 29. MDA (성상 관형사)

[부사류]

- 30. MAI (지시 부사)
- 31. MAC (접속 부사)
- 32. MAA (성상 부사)
- 33. MAAP (서술 부사)

[독립어류]

- 34. IE (감탄사)

[접사류]

- 35. XPU (수사형 접두사)
- 36. XSN (명사 접미사)
- 37. XSD (관형사 파생 접미사)
- 38. XSV (동사 파생 접미사)
- 39. XSJ (형용사 파생 접미사)
- 40. XSA (부사 파생 접미사)

[기호류]

- 41. SS, (문장기호-)
- 42. SS: (문장기호-:)
- 43. SS; (문장기호-;)
- 44. SS/ (문장기호-/))
- 45. SS. (문장기호-.)
- 46. SS? (문장기호-?)
- 47. SS! (문장기호-!)
- 48. SSA (문장기호-생략)
- 48. SS' (문장기호-')
- 49. SS' (문장기호-')
- 50. SS((문장기호-()
- 51. SS) (문장기호-))
- 52. SS- (문장기호-))
- 53. SSX (문장기호-기타)
- 54. SCD (문자기호-숫자)
- 55. SCU (문자기호-단위)
- 56. SCC (문자기호-화폐단위)
- 57. SCF (문자기호-외국어)
- 58. SCH (문자기호-한자어)