

운율경계강도 예측을 위한 품사셋 비교 연구

엄 기완*, 김 진영*, 김 선미**, 성 평모***, 이 현복****

* 전남대학교 전자공학과 ** 서울대학교 뉴미디어 연구소
*** 서울대학교 전기공학부 ****서울대학교 언어학과

Comparison of Three POS Sets in Prosody Break Index Estimation

Ki-Wan Eom*, Jin-Young Kim*, Sun-Mi Kim**, Koeng-Mo Sung***, Hyun-Bok Lee****

* Dept. of Electronics Eng. , Chonnam Nat'l Univ.

** New Media Research Center, Seoul Nat'l Univ.

*** School of Electrical Eng. ,Seoul Nat'l Univ.

****Dept. of Linguistics, Seoul Nat'l Univ.

eom@dsp.chonnam.ac.kr, kimjin@dsp.chonnam.ac.kr, sunmi@acoustics.snu.ac.kr, kmsung@ee.snu.ac.kr

요 약

본 논문에서는 문장의 문법 구조로부터 운율 경계 강도를 효율적으로 예측하기 위해서, 문법 정보의 세밀함에 따라 품사셋을 3 단계로 설정하였다. 그리고 운율 경계 강도를 예측하는데 있어서 어떠한 품사셋이 최적인가를 알아보기 위해 150문장의 코퍼스를 구축하였으며, 세 종류의 품사셋에 대해 코퍼스를 수작업으로 품사분석을 하였다. 청취실험으로 결정한 운율 경계 강도를 바탕으로 확률론적인 모델링 방법을 사용하여 예측하는 실험을 하였다. 이러한 예측결과를 평가 비교하여 최적의 품사셋을 정하였다.

1. 서 론

사람은 화자의 음성을 청취할 때 음의 경계 위치를 지각한 후 화자의 의도를 해석하는데, 그 정보를 이용한다. 이처럼 청취자로 하여금 발화의 끝감을 지각시키는 곳을 운율구 경계라 한다. 이러한 운율구 경계 위치를 찾는 문제는 합성기의 자연성을 향상시키는 중요한 변수이다.^[1-2] 또한, 이는 연속음 음성인식에 있어서, 발화된 문장을 작은 단위로 인식하게 되므로 인식의 복잡성을 줄이 수 있게된다.^[3-5] 일반적으로 운율구 형성은 문장의 품사정보와 구문정보에 의해서 이루어진다고 알려

져 있다.

한국어의 문장을 정밀하게 분석하기 위해서는 한국어의 품사(Part of Speech, POS)를 가능한 한 세분화고 이를 바탕으로 태깅을 수행하는 것이다. 그러나, 한국어 음성합성기라는 시스템을 고려하여 볼 때, 수 백개의 품사를 사용하는 것은 불가능한 것이다. 왜냐하면, 한국어의 억양규칙을 구현함에 있어서 bi-gram, tri-gram 또는 penta-gram의 복잡한 통계학적 문법을 사용하게 되는데 품사의 종류가 수백 개에 이르러서는 실제로 문법을 구현하기가 불가능하기 때문이다. 따라서, 제한된 개수의 품사를 가지고 한국어의 억양을 적절한 오차를 가지고서 분석할 수 있는 품사셋이 필요하다.

본 논문에서는 품사셋을 설정하기 위해 다음과 같은 기준을 두었다. 문장의 문법 구조로부터 운율 경계 강도를 효율적으로 예측하기 위해서는 운율 구조 형성에 영향을 줄 수 있는 문법 정보는 다 주고 운율 구조 형성에 영향을 주지 않는 정보는 주지 않으면 된다. 그러나 어떤 문법적 특징이 운율 구조에 영향을 주는지는 문법적 특징의 운율적 실현에 대한 포괄적인 연구가 행해진 후에야 답할 수 있는 문제이다. 이에 이 연구에서는 문법 정보의 세밀함에 따라 품사셋을 3 단계로 설정하고 각 품사셋에 대한 운율 경계 강도 예측 성능을 평가 비교하여 최적의 품사셋을 정하고자 한다.

2. 품사셋 정의

2.1. 1단계 품사셋

1단계는 김선미(1997)의 품사셋으로 그 이론적 기반을 허웅(1983, 1995)에 두고 남기섭·고영근(1985), 장석진(1993), Lee, Hyun-Bok(1989), 이익섭·임홍빈(1984)를 참조하여 고안한 것이며, 태깅셋 고안의 기본원칙은 운율적인 행동이 유사한 문법적 성분들은 한가지 기호로 통합하며, 문법적으로는 유사한 기능을 갖지만 운율적 기능이 다른 것은 별개의 기호를 준 것이다. 즉, 운율구 형성에 영향을 줄 수 있다고 생각되는 통사 형태적 정보를 최대한 자세히 준 것이다.¹⁶⁻¹²⁾

이렇다면, 체언이 들어간 어절의 경우, 품사는 명사, 의존명사, 대명사, 수사로 구분하고, 격조사는 주격, 속격, 목적격, 위치격, 방편격, 대비격으로 구분하였다. 격조사와 보조사의 연결 등 조사가 겹쳐서 나타날 때는 이를 순서대로 다 표시하여, 예를 들어 '학교(N)-에서(L)-는(M)'은 'NLM'으로 표시하였다. 용언이 들어간 어절의 경우, 품사는 동사, 형용사, 지정사로 나누고 용언의 품사와 어말어미의 정보를 주었다. 용언의 어간에 전성어미나 종결어미가 붙고 여기에 다시 조사가 붙은 경우 이렇다 표시하였다. 예를 들어 '아름답(A)-기(N)-는(M)'은 'ANM'로, '보내(V)-느냐(I)-에(L)'는 'VIL'로 표기하였다.

이 품사셋은 문법 기호 하나로도 그 앞뒤의 통사 구조를 어느 정도 예측할 수 있으며, 한국어의 모든 가능한 어절을 체계적으로 태깅할 수 있다는 장점을 가진 반면, 품사셋이 커서 대용량의 코퍼스를 필요로 한다는 단점이 있다. 1단계의 품사셋에 정의된 품사의 종류는 총 232개이다.

2.2. 2단계 품사셋

1단계의 품사셋이 코퍼스가 적을 경우 통사 구조와 운율 구조와의 관계에 대한 전체적인 시각을 방해하는 면이 없지 않으므로 품사셋을 다음과 같이 축소하였다.

체언의 경우, 대명사와 수사를 명사류에 포함시켜 체언의 품사를 명사류(N)와 의존명사(B)의 두 가지로 설정하였다. 조사에 있어서는 위치격, 방편격, 대비격의 격조사를 부사격 하나로 묶어, '학교-에서'와 '학교-로서'를 모두 'NV'로 표시하였다. 또한 격조사와 보조사가 겹쳐서 나타날 때에는 보조사만 표시하여, '학교에서는'과 '학교로서는'을 둘 다 'NM'으로 표시하였다. 용언의 경우는 동사와 형용사를 동사류로 묶어 용언의 품사를 동사류(V)와 지정사(C)의 둘로 설정하였다. 따라서 '달리-고'과 '붙잡-고' 같은 어절을 모두 'VC'로 표시하였다. 2단계의 품사셋에 정의된 품사의 종류는 총 110개이다.

2.3. 3단계 품사셋

품사셋을 더 축소하여, 체언이나 용언의 품사를 표시하지 않고, 조사나 어말어미만을 표시하되 2단계 품사셋에 준하였다. 이 품사셋은 해당 어절의 통사적 기능만을 표시한 것으로 1단계 품사셋이 어절의 통사적 기능뿐 아니라 형태론적 특징을 표현하고 2단계 품사셋이 축소된 형태소 정보를 가진 것과 비교된다. 1단계의 품사셋에 정의된 품사의 종류는 총 18개이다.

본 연구에서 사용된 코퍼스 중 한 문장을 위의 세 가지 품사셋의 문법 기호로 적용한 품사분석한 예는 아래와 같다.

<품사분석 예>

·1단계 품사셋

생일이(NS) 한(D) 생명이(NS) 태어난(VA) 것을(BO) 기념하고(VC) 축하하는(VA) 날이라는(CDQA) 것은(BS) 동서고금을(NO) 막론하고(VC) 모두(E) 같다(AD).

·2단계 품사셋

생일이(NS) 한(D) 생명이(NS) 태어난(VA) 것을(BO) 기념하고(VC) 축하하는(VA) 날이라는(CA) 것은(BM) 동서고금을(NO) 막론하고(VC) 모두(E) 같다(VD).

·3단계 품사셋

생일이(NS) 한(D) 생명이(NS) 태어난(VA) 것을(NO) 기념하고(VC) 축하하는(VA) 날이라는(VA) 것은(NM) 동서고금을(NO) 막론하고(VC) 모두(E) 같다(VD).

3. 음성코퍼스

본 연구에서 사용된 실험 문장은 중학교 국어 교과서와, <한국어의 표준발음>의 '음성독본'에서 무작위로 발췌한 150개의 문장으로 문체(style)는 낭독체이다. 문장의 길이는 대개 12어절에서 16어절 정도이다. 품사분석은 수작업을 통해 위의 3 단계 품사셋에 대해 분석을 하였다. 그리고 서울태생의 남성화자가 발췌된 150개의 문장을 보통속도로 말성한 것을 녹음하였으며, 녹음기는 SONY의 DAT(Digital Audio Tape-Recorder) TCD-D8로 하고 마이크로폰은 Shure사의 SM58S를 사용하였다. 이렇게 녹음된 음성 데이터는 운율 경계 강도 설정을 위한 정취실험에 사용하였다. 정취실험은 총 10명이 참가하였으며, 이들은 발화된 음성을 듣고 경계강도에 따라 표 1과 같이 4 가지 중 하나를 선택하였다. 그리고 이들의 정취실험 결과들 중에서 최빈치(mode)를 정취실험에 의한 운율 경계 강도로 결정하였다.

표 1. 운율경계강도 분류

운율경계 강도(BI)	설 명
BI_0	어절간에 경계가 전혀 느껴지지 않는다.
BI_1	어절간에 약한 경계가 느껴진다.
BI_2	어절간에 강한 경계가 느껴진다.
BI_3	어절간에 매우 강한 경계가 느껴진다.

다음 표 2에는 150문장에 대한 품사분석 결과 문장에서 발생한 bi-gram, tri-gram 수, 그리고 각각의 운율 경계 강도별 발생횟수를 나타내었다.

표 2. 150문장의 품사분석 결과 단위: 개

품사셋	품사개수	bi-gram	tri-gram	BI_0	BI_1	BI_2	BI_3
1단계	97	597	1295	1043	359	213	144
2단계	44	314	974	1024	358	213	144
3단계	18	202	820	1024	358	213	144

위의 결과에서 보면, 본 연구에서 사용된 150문장 코퍼스에서 나타난 품사 개수는 각 단계의 품사셋마다 97, 44, 18개로 각각에 정의된 품사셋의 크기에 비해서 발생한 개수가 매우 적다고 할 수도 있다. 그러나 여기에 사용된 문장들이 랜덤하게 추출한 것이므로, 이는 다른 텍스트를 사용했을 경우에도 유사한 결과가 나오리라 여겨진다.

4. Markov 모델에 의한 운율경계강도 추정

4.1 예측실험

보다 정확한 운율 경계 강도를 예측하기 위해서는 텍스트의 구문정보뿐만 아니라 음성정보도 필요하나 본 연구가 운율 경계 강도를 예측하는데 있어 어떠한 문법 기호체계가 유효한가에 대한 것이므로, 여기에서는 품사 정보만을 이용하여 운율 경계 강도를 예측하였으며, 품사정보로는 bi-gram과 tri-gram을 사용하였다. 그리고 운율 경계 강도의 추정방법에는 Markov 방식을 사용하였으며, 문장전체에 대한 운율 경계 강도를 Viterbi 탐색 방법을 이용하여, 최적의 운율 경계 강도열을 선택하였다.

bi-gram과 tri-gram을 사용한 확률값은 식 1과 2와 같다.

$$P(D_i, D_{i+1} | BI_i) \quad (1)$$

$$P(D_{i-1}, D_i, D_{i+1} | BI_i) \quad (2)$$

위 식에서 D_i 는 i 번째 어절에 해당하는 품사이고, BI_i 는 i 번째 어절 뒤에 나타나는 운율 경계 강도를 나타낸다.

운율 경계 강도 유형의 천이확률(transition probability) 값으로 N-gram을 사용하게 되면, 천이확률은 다음 식과 같다.

$$P(BI_i) = P(BI_i | BI_{i-1}, BI_{i-2}, \dots, BI_{i-N+1}) \quad (3)$$

$$= P(BI_i | BI_{i-1})$$

운율 경계 강도 유형별 발생확률은 Bayes' rule에 의해 다음과 같이 정리될 수 있다.

$$P(BI | D) = \frac{P(D | BI) P(BI)}{P(D)} \quad (4)$$

위의 확률을 구하는 방법으로 Viterbi search 알고리즘을 사용하였다.

다음 표3, 4 그리고 5는 세 가지 품사셋에 대한 운율 경계 강도 예측 결과에 대한 혼돈표(confusion table)을 나타낸 것이다.

표 3. 운율경계강도 예측결과(1단계 품사셋) 단위: %

예측값 \ 실제값	bi-gram				tri-gram			
	BI_0	BI_1	BI_2	BI_3	BI_0	BI_1	BI_2	BI_3
BI_0	94.2	4.8	0.9	0.1	98.0	1.8	0.2	0.0
BI_1	56.7	32.2	7.1	4.1	20.1	77.4	1.4	1.1
BI_2	34.2	16.0	39.9	16	15.5	8.5	71.4	4.7
BI_3	23.5	15.7	12.4	48.4	6.3	0.7	2.8	90.3

표 4. 운율경계강도 예측결과(2단계 품사셋) 단위: %

예측값 \ 실제값	bi-gram				tri-gram			
	BI_0	BI_1	BI_2	BI_3	BI_0	BI_1	BI_2	BI_3
BI_0	94.9	3.8	0.9	0.4	97.5	2.1	0.2	0.2
BI_1	53.9	34.9	7.8	3.3	36.9	57.8	2.8	2.5
BI_2	24.3	14.4	45.7	15.6	16.9	12.2	62.0	8.9
BI_3	16.3	3.9	24.2	55.6	7.6	3.3	10.4	78.5

표 5. 운율경계강도 예측결과(3단계 품사셋) 단위: %

예측값 \ 실제값	bi-gram				tri-gram			
	BI_0	BI_1	BI_2	BI_3	BI_0	BI_1	BI_2	BI_3
BI_0	95.5	3.8	0.5	0.2	98.3	1.4	0.2	0.1
BI_1	50.4	38.2	7.3	4.1	33.8	61.7	2.2	2.2
BI_2	21.8	14.8	46.5	16.9	16.4	10.8	63.8	8.9
BI_3	13.7	3.9	19.0	63.4	6.3	2.8	4.2	86.8

4.2 결과분석

예측결과를 나타내는 위의 혼돈표에서 그림 1과 같은 가중치를 주어 각 품사셋에 대한 COST를 계산하였다.

즉 실제 BI가 0인데 1로 예측한 경우에는 그것이 발생한 횟수에 가중치 1을 곱하고, 2나 3으로 예측한 경우에는 가중치 2를 곱하게 된다. 이처럼 모든 경우에 대해 계산하고, 발생된 총 횟수로 나누어 각 품사셋에 대한 COST를 계산하였다. 그러므로 이 값은 작을수록 더 좋은 성능을 나타낸다.

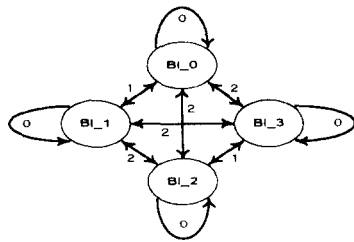


그림 1. 예측혼돈에 따른 가중치

표 6. 각 품사셋에 대한 COST 비교

품사셋	COST(bi-gram)	COST(tri-gram)
1단계	0.42	0.14
2단계	0.34	0.20
3단계	0.37	0.22

그리고 다음 표 7에는 실험결과를 각 품사셋에 대해 전체 인식률로 나타내었다. 여기서 도수1의 의미는 bi-gram 이나 tri-gram에서 운율 경계 강도(0-3)중에서 어느 하나만 한번 나타난 경우이다. 이러한 경우는 예측이 100%가 되므로 예측성능을 비교하는데 의미가 없다. 그러므로 도수가 1인 경우를 제외한 인식률(도수2이상)을 아래 표에 도수 1을 포함한 경우와 함께 비교하여 나타내었다.

표 7. 전체 인식률 비교 단위: %

품사셋	bi-gram		tri-gram	
	도수1이상	도수2이상	도수1이상	도수2이상
1단계	70.8	65.7	89.9	75.8
2단계	74.6	73.0	85.6	77.6
3단계	72.8	72.3	83.4	77.5

V. 결론

본 논문에서는 문법 정보의 세밀함에 따라 품사셋을 3 단계로 설정하였다. 그리고 품사셋 중 운율 경계 강도

를 예측하는데 있어서 어떠한 품사셋이 적절한가를 정의하기 위해 bi-gram과 tri-gram을 사용한 확률론적 모델링 방법에 의해 그 성능을 평가 비교하였다. 지금까지 살펴본 바와 같이 위의 3가지 품사셋들은 운율 경계 강도 예측에 있어서 비슷한 성능을 나타내지만, 그 중 2단계 품사셋이 좀더 유효함을 알 수 있다. 그러나 품사셋의 크기에 따른 복잡성을 감안하면, 1단계 품사셋을 사용하여도 무방하리라 여겨진다. 본 연구결과를 토대로 향후 대용량의 코퍼스를 구축하여, 위의 결과를 검증하고자 한다.

* 본 논문은 '97 학제간연구지원사업 지원에 의해 이루어진 연구 결과물 중 하나입니다

참고문헌

- [1] J. Allen, M. S. Hunnicutt & D. Klatt, From Text to Speech : The MITalk System, Cambridge University Press, 1987.
- [2] Ostendorf, Parse Scoring with Prosodic Information: An Analysis and Synthesis Approach, in Computer Speech and Language, July 1993.
- [3] A. J. Hunt, A generalized Model for Utilising Prosodic Information in Continuous Speech Recognition, Speech Technology Research Group, University of Sydney, 1995.
- [4] A. J. Hunt, Models of Prosody and Syntax and their Application to Automatic Speech Recognition, Ph.D thesis. University of Sydney, 1995.
- [5] Y. J. Kim, S. H. Lee, Y. H. Oh, Relationship Between Prosodic features and Dependency Relation, ICSP'97, 1997.
- [6] 김선미, 한국어의 리듬 단위와 문법 구조 -음성합성에서 리듬 구현의 자연성 향상을 위한 음성언어학적 연구-, 서울대학교 언어학과 박사학위논문, 1997.
- [7] 남기심·고영근, 표준 국어 문법론, 탑출판사, 1985.
- [8] 이익섭·임홍빈, 국어문법론, 학연사, 1984.
- [9] 장석진, 정보기반 한국어 문법, 언어와 정보, 1993.
- [10] 허웅, 국어학 - 우리말의 오늘, 어제 -, 샘문화사, 1983.
- [11] 허웅, 20세기 우리말의 형태론, 샘문화사, 1995.
- [12] Lee, Hyun-Bok, Korean Grammar, Oxford University Press, New York, 1989.