

국어 낭독체 발화의 운율경계 예측*

김효숙(언어과학), 김정원(언어과학), 김선주(언어과학),
김선철(언어과학), 김삼진(언어과학), 권철홍(대전대)

<차례>

- | | |
|---------------|--------------------|
| 1. 서론 | 3. 운율경계 예측 |
| 2. 데이터 및 CART | 3.1. nonIP와 IP 예측율 |
| 2.1. 데이터 | 3.2. zero와 AP 예측율 |
| 2.2. CART | 4. 결론 |

<Abstract>

Prediction of Break Indices in Korean Read Speech

Hyo Sook Kim, Chung Won Kim, Sun Ju Kim,
Seoncheol Kim, Sam Jin Kim, Chul Hong Kwon

This study aims to model Korean prosodic phrasing using CART(classification and regression tree) method. Our data are limited to Korean read speech. We used 400 sentences made up of editorials, essays, novels and news scripts. Professional radio actress read 400sentences for about two hours. We used K-ToBI transcription system. For technical reason, original break indices 1,2 are merged into AP. Differ from original K-ToBI, we have three break index Zero, AP and IP.

Linguistic information selected for this study is as follows: the number of syllables in 'Eojeol', the location of 'Eojeol' in sentence and part-of-speech(POS) of adjacent 'Eojeol's.

We trained CART tree using above information as variables. Average accuracy of predicting NonIP(Zero and AP) and IP was 90.4% in training data and 88.5% in test data. Average prediction accuracy of Zero and AP was 79.7% in training data and 78.7% in test data.

* 주제어: 낭독체 발화, 운율경계, CART 방법

1. 서론

국어의 운율경계로 ‘말토막’과 ‘말마디’가 있다[6,7]. ‘말토막’ 경계는 쉽이 없거나, 있어도 임시적이고 불완전한 아주 짧은 쉽을 나타낸다. 말토막 경계에 나타나는 휴지는 경계 앞 음절을 장음화시키고 앞 말토막과 뒤 말토막의 억양과 리듬의 끊김을 초래한다. ‘말마디’는 숨쉬기의 단위이다. 말마디의 경계에도 휴지가 부과되는데, 말마디 경계에 부과되는 휴지는 종종 숨쉬기에 의한 에너지의 공백을 수반하며 경계 앞 음절의 장음화를 초래한다. 말마디의 경계에서 나타나는 휴지는 말토막 경계에서 나타나는 휴지보다 더 길게 실현된다. K-ToBI에서도 이와 유사한 단위를 정의하고 있는데 ‘말토막’은 ‘Accentual phrase(AP)’에 ‘말마디’는 ‘Intonation phrase(IP)’에 대응한다. 단, AP에는 경계 앞 장음화현상은 없다는 점에서 ‘말토막’ 단위와 차이가 있다. 정의의 차이는 있으나 국어의 운율경계를 크게 두 가지로 보는 입장은 다른 연구에서도 나타나고 있다[2-4]. 그리고 음성합성기를 위한 대용량의 데이터를 자동으로 처리하기 위해서 휴지 길이를 기준으로 경계강도를 매기기도 한다[1].

본 논문에서는 K-ToBI의 운율전사체계를 이용하였다[11]. 운율경계 분석에 이용한 음향적 특징은 경계 앞 장음화, 음높이의 변동, 묵음구간, 음절 변동 등이다[8]. 경계태깅을 할 때 분석자의 판단에 일관성을 기하기 위해서 경계 앞 장음화현상이 있으나 묵음구간이 짧은 경우는 IP가 아닌 AP에 포함시켰다. 그리하여 K-ToBI의 Break Index 0은 본 논문의 zero에, break index 1,2는 본 논문의 AP에, break Index 3은 본 논문의 IP에 각각 대응한다.

2. 데이터 및 CART

2.1. 데이터

논설, 수필, 소설 및 방송문장으로 구성된 400문장(5620어절)을 전문 성우(여성)에게 보통 빠르기로 읽혔다. 400문장은 모두 평서형 종결어미로 끝나는 문장이다. 운율경계가 생기는 최소의 단위를 어절로 보았는데 합성어 구성을 이루며 한 어절의 음절수가 길어 어절안에서 경계가 생길 때에는 텍스트를 수정하여 어절 경계를 주었다. 예를 들어 ‘노량진수산주식회사’와 같은 어절을 성우가 읽을 때 ‘노량진`수산`주식회사’와 같이 읽었다면 ‘노량진 수산 주식회사’처럼 띄어쓰기를 수정하였다.

문장부호(특히 쉼표와 마침표)가 운율경계 결정에 중요한 역할을 하므로 별도의 형태소로 취급하여 어절 경계를 주었다. 따라서 5620개 어절에 문장부호로 인해

어절이 추가되어 6207개 어절이 되었다. 수작업 태깅 결과 zero는 987개 어절, AP 경계는 3427개 어절, IP 경계는 1793개 어절이었다.

운율경계는 형태-통사적으로 긴밀한 관계로 연결된 어절보다는 그렇지 않은 어절에 올 경우가 많다[9]. 낭독체발화에서는 문장 전체의 구조를 파악하기 보다는 텍스트에서 제시되는 형태순으로 순차적으로 처리되어 발화될 수 있다[8]. 따라서 문장을 분석할 때 형태위주의 문법분석을 통해 운율경계를 시도한 경우가 있다 [3,4,8,9].

본 논문에서는 형태소 정보만을 이용하여 운율경계를 예측해보았다. 그리고 어절의 구성에서 허사의 정보가 운율경계 결정에 중요한 역할을 하므로 어절의 품사를 실사부가 놓이는 왼쪽품사와 허사부가 놓이는 오른쪽 품사로 나누어서 분석하였다[4,9]. 허사부가 없는 어절의 경우는 왼쪽의 품사를 오른쪽 품사로 그대로 이용하였다. 기존의 연구들을 참고하여 다음과 같은 23개의 품사목록을 작성하였다[1-4].

<표 1> 형태소분석에 이용된 품사셋(23개)

1: 관형격조사	13: 부호
2: 관형사	14: 수사
3: 관형어미	15: 쉽표
4: 단위명사	16: 의존명사
5: 동사	17: 접속부사
6: 명사	18: 접속조사
7: 명사형어미	19: 종결부호
8: 목적격조사	20: 종결어미
9: 보조사	21: 종속 연결어미
10: 보조연결어미	22: 주격조사
11: 부사	23: 주체격조사
12: 부사격조사	

2.2. CART

본 연구에서는 CART를 이용하여 운율경계를 결정하는 결정 트리를 생성하였다. 관측어절 앞뒤로 최대 2개 어절까지의 품사를 고려하였다[5]. 최적 트리를 구할 때는 Ose(0 standard error)법을 이용하고 크기가 최소인 트리를 선택할 경우에는 1se rule을 이용하였다[4]. 형태소 정보 이외에도 관측어절의 음절수와 관측어절 다음어절의 음절수와 관측어절의 문장내에서의 위치를 고려하였다. 품사정보는 범주형

변수이며 나머지는 실수형 변수이다.

<표 2> 운율경계 결정트리에 사용된 설명변수목록(13개)

- <syl> 관측 어절의 음절 수
- <n_syl> 관측 어절 다음 어절의 음절 수
- <loc> 문장의 모든 어절에 일련번호를 매긴 후 관측어절의 일련번호를 문장 전체의 어절 수로 나눈 수
- <ppL> 관측 어절 앞 앞 어절의 왼쪽 품사
- <ppR> 관측 어절 앞 앞 어절의 오른쪽 품사
- <pL> 관측 어절 앞 어절의 왼쪽 품사
- <pR> 관측 어절 앞 어절의 오른쪽 품사
- <L> 관측 어절의 왼쪽 품사
- <R> 관측 어절의 오른쪽 품사
- <nL> 관측 어절 다음 어절의 왼쪽 품사
- <nR> 관측 어절 다음 어절의 오른쪽 품사
- <nnL> 관측 어절 다음 다음 어절의 왼쪽 품사
- <nnR> 관측 어절 다음 다음 어절의 오른쪽 품사

운율경계 결정 트리는 세 가지 운율경계를 목적변수로 주었을 때 예측율이 훈련문장(4188어절)에서 77.5%, 테스트문장(2019어절)에서 75.2%로 나타났다. 본 연구가 음성합성시스템을 목적으로 하고 있으므로 IP의 예측율을 높이기 위해 트리를 NonIP와 IP를 결정하는 트리과 zero와 AP를 결정하는 트리로 분리하였다. 6207어절을 가운데 4188(약 67%)어절을 훈련자료로 하고 2019(약 33%)어절을 테스트자료로 하여 nonIP(zero, AP)와 IP를 결정하는 트리를 만들었다. 그리고 IP로 결정된 어절을 제외한 나머지 4414어절 가운데 2991(약 67%)어절을 훈련자료로 하고 1423(약 33%)어절을 테스트자료로 하여 zero와 AP를 결정하는 트리를 만들었다.

3. 운율경계 예측

3.1. nonIP 와 IP 예측율

훈련자료에서 90.4%의 예측율을 보였으며 테스트자료에서는 88.5%의 예측율을 나타냈다.(Standard Tree Use rule 기준)

<표 3> nonIP와 IP 예측율 (테스트문장, 2019어절)

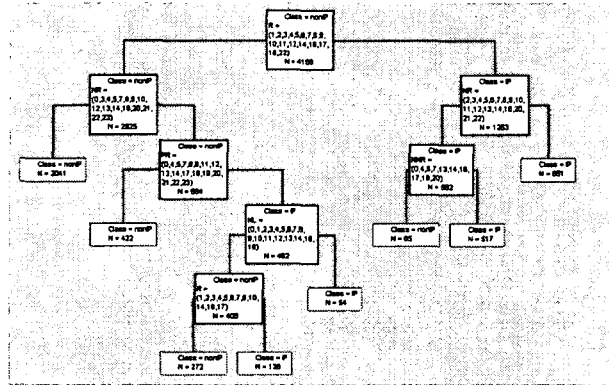
Actual Class	Predicted NonIP	Class IP	Actual total
NonIP	1283.00	157.00	1440.00
IP	76.00	503.00	579.00
Pred. Tot	1359.00	660.00	2019.00
Correct	0.891	0.869	
Tot. Correct	0.885		

NonIP와 IP의 결정에서 <R>과 <nR>, <nL> <nnR>, <nnL>이 기여도가 높은 것으로 나타났다.

<표 4> nonIP와 IP 결정에서의 변수 기여도

변수	기여도
R	100.000
nR	77.317
nL	71.470
nnR	68.841
nnL	66.316
L	38.006
pR	10.187
n_syl	7.084
pL	5.561
ppR	5.321
loc	3.217
ppL	3.041
syl	1.310

트리 가지치기를 더 많이 하는 결과를 가져오는 1se rule을 적용했을 때는 훈련 자료에서 87.9%의 예측율을 그리고 테스트자료에서는 88.8%의 예측율을 나타냈다. 다음은 1se rule을 적용한 트리이다.



<그림 1> nonIP와 IP의 classification tree(Ise rule)

트리의 왼쪽가지는 분류기준이 Yes일 경우이며 오른쪽 가지는 분류기준이 No인 경우이다. 예를 들어 가장 오른쪽에 있는 최종 노드의 자료는 다음과 같은 과정을 통해 IP로 결정된다. 오른쪽 품사가 쉼표, 종결부호, 종결어미, 종속연결어미, 주격조사인 경우(즉 분류기준이 No인 경우) 오른쪽 가지로 분류되고 다음 어절의 오른쪽 품사가 관형격조사, 접속부사, 접속조사, 종결부호인 경우(역시 분류기준이 No인 경우)에 다시 오른쪽 가지로 분류되고 이 때 최종 노드가 되어 IP로 결정된다.

3.2. zero와 AP 예측을

훈련자료에서는 79.7%의 예측율을 보였으며 테스트자료에서는 78.7%의 예측율을 나타냈다(Standard Tree Ise rule). Ise rule 을 적용했을 때 훈련자료에서 77.2%, 테스트자료에서 78.6%의 예측율을 나타냈다.

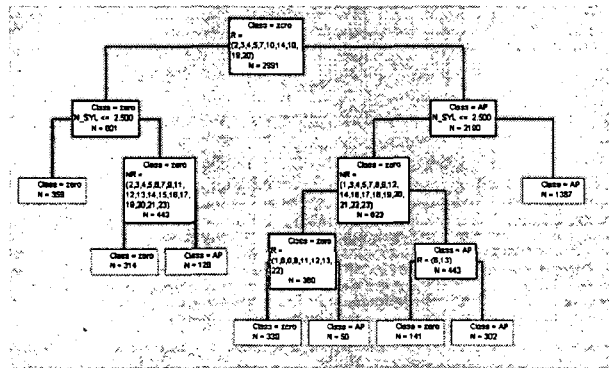
<표 5> zero와 AP 예측율(테스트문장, 1423어절, Ose)

Actual Class	Predicted Zero	Class AP	Actual total
Zero	259.00	70.00	329.00
AP	233.00	861.00	1094.00
Pred. Tot.	492.00	931.00	1423.00
Correct	0.787	0.787	
Tot. Correct	0.787		

Zero와 AP의 결정에서는<R>, <L>, <n_syl>이 기여도가 높은 것으로 나타났다.

<표 6> zero와 AP 결정에서의 변수 기여도

변수	기여도
R	100.000
L	57.112
n_syl	54.539
nR	45.252
nL	37.302
syl	35.689
pR	27.183
nnR	13.456
nnL	10.004
ppR	4.210
loc	2.010
ppL	1.328
pL	0.370



<그림 2> zero와 AP의 classification tree(1se rule)

4. 결론

운율경계 예측율에서 nonIP와 IP의 예측율은 테스트자료에서 88.5% 그리고 zero와 AP는 79%였다. 이 가운데 특히 zero와 AP의 예측율이 낮는데 이는 수작업 태

강 시 IP는 뚜렷한 음향적 특징으로 인해 비교적 일관성있게 뚜렷하게 구분할 수 있었던 반면 일부 AP는 zero와 구분하기가 어려웠던 데에서 비롯되었다고 판단된다. 추후에는 AP와 zero의 특징을 보다 세분화하여 분석의 일관성을 높이고 방법을 향상시켜zero와 AP 경계에측율을 높이도록 할 것이다.

참 고 문 헌

- [1] 김상훈, 박준, 이영직(2001), 코퍼스에 기반한 한국어 문장/음성변환 시스템(Corpus-based Korean Text-to-speech Conversion System), 「한국음향학회지」 제20권 제3호, pp.24~33.
- [2] 김진영, 박상언, 엄기완, 최승호(2000), 코퍼스 기반 한국어 합성기의 억양 구현 방안, 「음성과학」 제7권 2호, pp.193~208.
- [3] 엄기완, 김진영, 김선미, 이현복(1998), 품사셋에 의한 운율경계강도의 예측, 「말소리」 제 35-36호, pp.145~156.
- [4] 이상호, 오영환(1998), CART를 이용한 운율구 추출 및 휴지 기간 모델링, 제15회 음성통신 및 신호처리 워크샵, pp.81~86.
- [5] 이용주, 이숙향, 김종진, 고현주, 김영일, 김상훈, 이정철(1998), K-ToBI 기호에 준한 F0 곡선 생성 알고리즘, 「말소리」 제35-36호, pp.131~143.
- [6] 이현복(1993), 「한국어의 표준발음」, 교육과학사 3rd.
- [7] 이호영(1996), 「국어음성학」, 태학사.
- [8] 성철재(1995), 한국어 리듬의 실험음성학적 연구-시간 구조와 관련하여-, 서울대학교 언어학과 박사학위논문
- [9] 정희선(1996), 한국어 리듬단위 생성기 연구-낭독체 중심으로-, 제2회 음성학 학술대회 자료집, pp.217~224.
- [10] Jun, Sun-Ah (2000), K-ToBI Labelling Conventions, <http://www.linguistics.ucla.edu/people/jun/ktobi/k-tobi.html>.

접수일자: 2002년 5월 3일

게재결정: 2002년 5월 24일

▶김효숙(Hyo Sook Kim)

주소: 151-716 서울특별시 관악구 봉천4동 882-5 2F

소속: (주)언어과학부설 음성공학연구소

전화: 02) 881-8125

Fax: 02) 887-8127

E-mail: soogy35@dreamwiz.com

▶김정원(Chung Won Kim)

주소: 151-716 서울특별시 관악구 봉천4동 882-5 2F

소속: (주)언어과학부설 음성공학연구소

전화: 02) 887-8125

E-mail: kimpaq@intizen.com

▶김선주(Sun Ju Kim)

주소: 151-716 서울특별시 관악구 봉천4동 882-5 2F

소속: (주)언어과학부설 음성공학연구소

전화: 02) 887-8125

E-mail: repose@intizen.com

▶김선철(Seoncheol Kim)

주소: 151-716 서울특별시 관악구 봉천4동 882-5 2F

소속: (주)언어과학부설 음성공학연구소

전화: 02) 887-8125

E-mail: dacos@eoneo.co.kr

▶김삼진(Sam Jin Kim)

주소: 151-716 서울특별시 관악구 봉천4동 882-5 2F

소속: (주)언어과학부설 음성공학연구소

전화: 02) 887-8125

E-mail: sjmonk@eoneo.co.kr

▶권철홍(Chul Hong Kwon)

주소: 300-716 대전광역시 동구 용운동 96-3 대전대학교

소속: 대전대학교 정보통신공학과

전화: 042) 280-2555

E-mail: chkwon@dju.ac.kr