

운율레이블링 기준의 검토 및 제안

이숙향, 신지영, 김봉완, 이용주
원광대, 고려대, 원광대SiTEC, 원광대
shlee@wonkwang.ac.kr

Prosody Labelling Convention

Sook-hyang Lee*, Jiyoung Shin**, Bong-Wan Kim**, Yong-Ju Lee*
Wonkwang Univ.*, Korea Univ.**, SiTEC, Wonkwang Univ.*
E-mail : shlee@wonkwang.ac.kr, shinjy@korea.ac.kr, bwkim@sitec.or.kr,
yjlee@wonkwang.ac.kr

Abstract

The purpose of this paper is to review one of the prosody labelling conventions for Korean, K-ToBI convention(ver. 3.1) and to propose a couple of modifications and suggestions.

I. 서론

음성 코퍼스 구축 및 활용에 있어 음운 레이블링과 함께 운율 레이블링이 필수적인 작업 과정의 하나이기 때문에, 여러 해에 걸쳐 음성학자들과 음성공학자들이 공동으로 레이블링 기준안 제작과 함께 이에 대한 검토, 수정 보완작업을 병행해 왔다. 그 한 예가 최근에 이루어진 원광대학교 SiTEC의 음운, 운율 레이블링 작업이다. 그 중 운율 레이블링 작업은 전국적 망을 가지고 있는 SiTEC 음성학 분야 7개 협력사이트¹⁾가 2년여에 걸쳐 공동 작업으로 이루어졌다. 남, 녀 성우 각각 한 명이 4,392개의 문장을 읽어 총 8,784개의 녹음된 문장을 레이블링하였다. 본 논문은 한국어 운율 구조에 대해 논하고 이러한 공동 운율레이블링 작업의 근간이 된 K-ToBI(Korean Tone and Break Indices) 기준안을 간단히 소개한다. 그리고 이 기준안에 제시

되어 있지 않은 여러 다양한 레이블링 환경에 대한 몇 가지 규정과 보다 체계적이고 객관적인 레이블링을 위하여 운율 요소들에 대한 정량화의 시도를 제안한다.

II. 한국어 운율구조

그림 1에서와 같이 한국어 발화문장의 운율구조는 계층구조로 이루어져 있으며 하단의 음절층에서 시작하여, 단어/어절층, 악센트구(Accentual Phrase)층, 그리고 최상단에 억양구(Intonation Phrase)층으로 구성된다. 악센트구는 하나 이상의 어절로 이루어지며 억양구는 악센트구의 상위층으로서 하나 이상의 악센트구로 이루어진다.

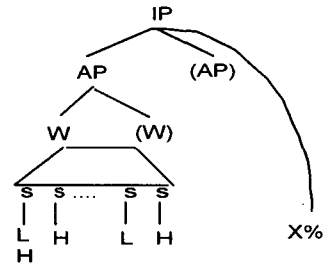


그림 1. 서울 방언의 운율구조

IP(Intonation Phrase): 억양구
AP(Accentual Phrase): 악센트구
W(Prosodic Word): 운율 단어
s(syllable): 음절
X%: 억양구 경계성조

1) 7개 협력사이트는 경북대(이은영), 고려대(신지영), 부천대(이순향), 서울대(이호영), 아주대(문승재), 연세대(이석재), 원광대(이숙향)임.

억양구는 구 마지막에 경계성조로 구분되어지고 구 말 장음화 현상(phrase final lengthening)을 수반하며 그 하위층에 속하는 악센트구간의 downstep이 일어나는 최대구간이 된다. 반면에 악센트구는 기본형으로 [L +H ... L+ Ha] 성조 패턴을 가지며 구 말 장음화 현상은 보이지 않는다[8][2]. 그리고 경음화와 모음 장음화(전남 방언), 그리고 유성음화가 적용되는 최대 구간이기도 하다[8]. 일반적으로 악센트구초 [L +H]는 L이 첫 음절에, +H가 둘째 음절에 없으며 악센트구말 [L+ Ha]는 구 말 끝에서 둘째 음절에 L+가 그리고 마지막 음절에 Ha가 없힌다. 또한 기본형 [L +H ... L+ Ha]는 여러 다양한 요인에 의하여 다양한 형태로 실현되어 나타나기도 한다. 첫째, 악센트구의 시작자음이 강자음 (/s/와 /h/를 포함한 유기음과 경음)이면 [H +H L+ Ha]의 성조 유형이 되며 둘째, 악센트구의 음절수가 4개보다 적을 때는 악센트구초 둘째 성조 +H 또는 악센트구말 L+가 실현되지 않거나 두 성조 모두가 미실현되기도 한다. 셋째, 악센트구말의 Ha 대신 La로 끝나는 경우도 간혹 관찰된다[1][9].

III. K-ToBI 레이블링 기준안

한국어 운율구조 연구는 그 동안 음성, 음운학자들에 의해, 여러 가지 접근방법으로 이루어져 왔다 [4][5][7][10]. Beckman과 Jun은 이러한 연구 결과들을 Pierrehumbert의 운율 이론적 토대[12]에 접목시켜 한국어 운율구조 기술 체계의 하나인 K-ToBI 기술 체계를 고안하였다[6]. 그 후 1996년에 1차 수정을 거치고(2차 버전), 2000년도에 다시 한 번 수정을 거듭하여 3차 버전이 나왔다[9].

K-ToBI 레이블링 기준안은 그림 2와 같이 5 개의 기술 층을 둔다²⁾.

1. 단어층 a word tier
2. 음운[기저]성조층 a phonological tone tier
3. 음성[표면]성조층 a phonetic tone tier
4. 연결층 a break-index tier
5. 기타층 a miscellaneous tier

단어층에서는 한글 맞춤법에 따라 띄어쓰면 일단 하나의 단어로 가정한다.

성조층은 악센트구와 억양구의 존재 또는 경계를 나

2) 기준안 3차 버전이 과연 한국어의 운율구조를 기술하기에 적절한 체계인지를 보기 위한 연구의 일환으로 수행한 레이블링 실험에서 실험에 참여한 20명의 레이블러간의 일치도가 비교적 높게 나타났다[10].

타내는 기저성조층과 실제 다양한 형태로 나타나는 두

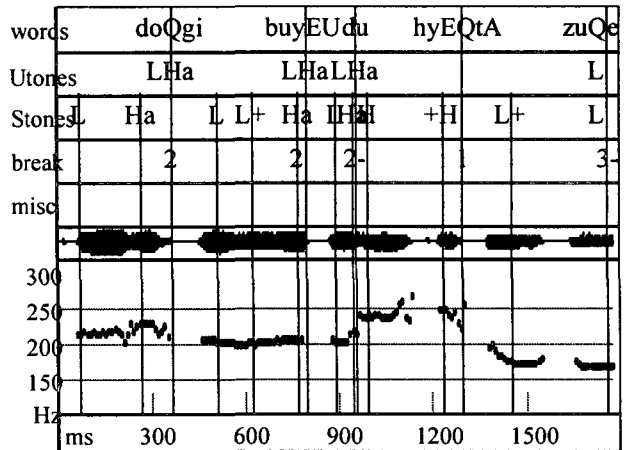


그림 2. K-ToBI 레이블링 기준안의 운율레이블링의 한 예.

개 구의 억양패턴을 기술하는 표면성조층으로 나뉜다.

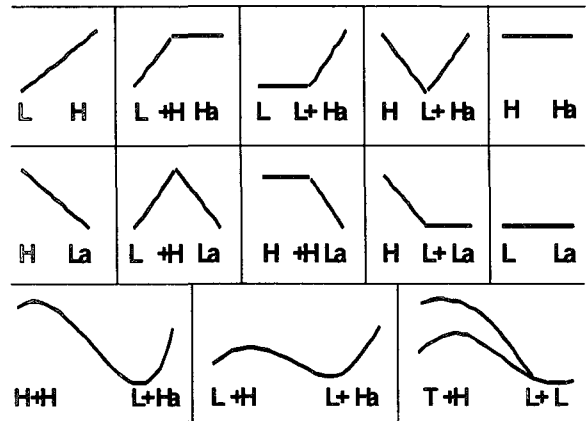


그림 3. 한국어 악센트구의 14가지 유형의 표면형 성조의 도식적 모양.

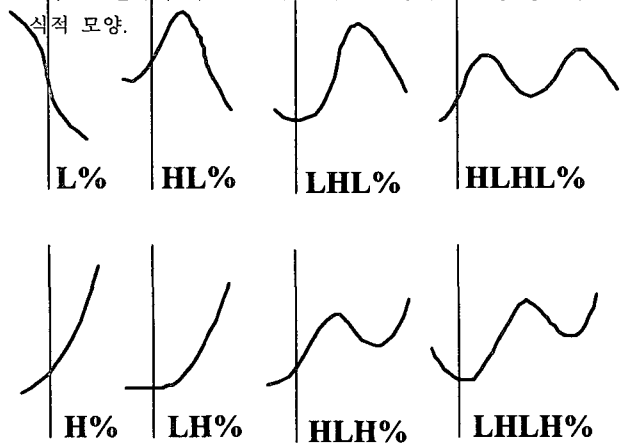


그림 4. 한국어의 다양한 억양구 유형.

K-ToBI 레이블링 기준안에서 가정하고 있는 한국어 악센트구는 14개의 유형이며(그림 3), 억양구는 8개의 유형(그림 4)으로서 이들 중 하나가 표면성조층에 각각 억양구와 악센트구의 표면 유형으로 기술된다.

연접층은 발화문장의 단어와 단어간, 그리고 문장의 마지막 단어와 문장 끝에 오는 휴지간의 연접의 정도를 나타낸다. K-ToBI 레이블링 기준안은 '0'에서 '3'까지 4 단계의 연접 또는 끊김의 정도를 가정한다. 판단은 레이블러의 청각적 판단에 의하며 '0'은 음성적으로 cliticization이 일어나거나, 일음절 불완전 명사에 해당하며, '1'은 단어간의 연접 정도, '2'는 pause의 강한(주관적) 느낌이 없는 약한 구경계에 해당하는 끊김으로서 전형적으로 악센트구 경계에서 느끼는 끊김 정도를 말한다. '3'은 pause가 실제 waveform이나 spectrogram 상에서 관찰되는 무음구간이든 또는 무음구간은 없지만 구말장음화 현상에 의해 느껴지는 것이든간에 pause의 강한 느낌을 주는 강한 구경계에서 느껴지는 끊김의 정도를 말한다. 전형적으로 억양구의 경계성조와 함께 느껴지는 끊김이 이에 해당한다. 실제 성조층과 연접층간의 불일치가 일어나는 경우를 위해 불일치를 나타내는 'm'을 더해준다 (2m, 3m). 연접의 정도 판단에 확신이 서지 않는 경우에는 '-'를 더해준다 (1-, 2-, 3-).

기타층에서는 성조와 연접 정도 외에 레이블러의 판단에 따라 필요하다면 뭐든 기술해줄 수 있다. 예를 들어, 화자의 문장 발화 중에 기침소리가 났다든지, 말을 더듬는다든지 또는 말하는 중간에 다른 화자가 끼어들었다든지 등의 것을 기술해줄 수 있다.

IV. 검토 및 제안

K-ToBI 레이블링 기준안 자체에 대한 검토 및 제안과 SiTEC 레이블링 공동작업 시 협력사이트와의 논의 결과 협의되었던 수정, 보완사항을 제시한다. 그리고 보다 객관적인 기술을 위해 운율요소들의 음성학적 특성에 대한 몇 가지 정량화 시도를 제안한다.

1. K-ToBI 레이블링 기준안 3.1버전에서 억양구는 기저성조층이나 표면성조층 모두에서 실제 표면형 8개 유형 중 하나를 기술해주고 있다. 반면에, 악센트구는 기저성조층에서는 실제 표면형을 기술하지 않고 단순히 기저형만을 기술하고 (LHa), 표면성조층에서 14개의 다양한 성조 중 하나를 기술해주고 있다. 따라서, 악센트구의 기술과 일관되면서 또한 하나이었던 성조층을 기준안 3차 버전에서 둘로 나눈 취지에 합당하

게, 억양구도 기저성조층에서 8개의 억양구 경계성조를 대표하는 또는 커버하는 성조를 하나 선정하여 그것으로 억양구의 경계성조를 나타내는 것이 바람직하다고 생각한다. 가장 빈도수가 높은 L%를 기저 억양구 경계성조로 제안한다.

2. 실제로 SiTEC 공동 레이블링 작업에서는 기저성조층은 기술하지 않기로 하였다. 이 층은 언어학적으로는 중요한 의미를 가질 수 있으나, 표면성조층으로부터 예측 가능하므로 필요한 경우, 이를 표면성조층에서 도출해내는 것으로 하여 기술에서 제외하였다

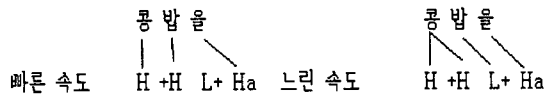
3. 억양구 경계에서 에너지가 약하여 억양곡선이 억양구 마지막까지 이어지지 못하고 끊기는 경우가 종종 발생하는데 이 경우, 레이블러의 청각적 판단에 의거하여 경계성조 유형을 기술하되, 기호는 보통의 억양구 경계성조 기술과 마찬가지로 단어층의 경계와 일치하는 곳에 둔다. 또한 pitch 분석 프로그램의 오류에 의해 억양곡선상에 잡음이 나타날 수 있는데 이를 무시하고 청각적 판단을 우선으로 한다.

4. 모음무성음화에 의해 악센트구 중간에 억양곡선이 끊기는 경우가 발생하는 경우, 특히 유기음과 고모음의 결합에서 빈번히 발생하는데 이 경우 또한 운율 판단은 레이블러의 청각적 판단에 의해 기술하며 모음 중간 지점에 성조기호를 둔다.

5. 보다 객관적이고 체계적인 기술과 자동 레이블링을 위해서는 운율적 요소들에 대한 정량화가 필요하다. 예를 들어, 악센트구의 유형별 빈도수나 음절수와 유형간의 관계 등에 관한 여러 연구가 있긴 하나, 화자 수나 출신 성분, 발화 스타일, 발화 속도, 발화 환경, 그리고 자료의 크기에서 볼 때, 아직 미약하다. 다양한 그리고 동시에 대규모의 음성코퍼스 구축과 분석을 필요로 한다. 운율레이블링 작업 시 많은 어려움을 호소하는 경우가 +H, Ha, H%간의 구분이다. 또한 L+, La, L% 간의 구분 또한 어려움이 많다. 이와 관련한 연구 결과들을 보면, 서울 방언에서 Ha 유형 악센트구 내의 내림(+H L+)의 경사도는 음절수에 따라 경사도가 달라지는 것으로 나타났다(4음절: -0.08 Hz/ms , 6음절: -0.20 Hz/ms)[8]. 그러나 악센트구 경계에 나타나는 내림(La H)은 악센트구 내에 나타나는 내림보다 경사가 급하게 나타났다 ($-0.5 \sim -0.7 \text{ Hz/ms}$). 같은 연구에서 서울 방언에서와는 달리 {L +H La} 또는 {H +H La} 패턴을 보이는 전남 방언 악센트구에서는 음절수가 내림(+H La)의 경사도에 큰 영향을 미치지 않는 것으로 나타났다 ($-0.15 \sim -0.18 \text{ Hz/ms}$). 다른 연구 결과

를 보면, 서울 방언의 La 유형 악센트구의 내림의 경사도는 전남 방언의 내림보다 일반적으로 경사가 급하게 나타났으며 서울 방언에서 Ha 유형의 악센트구 내림은 음절수가 많아질수록 경사가 급하게 나타난 반면 La 유형의 악센트구 내림은 음절수가 많아질수록 경사가 완만하게 나타났다[3]. 이 외에 악센트구의 오름(L+H, L+Ha)의 경사도에 대한 연구 결과들도 있다.

또한 발화속도와 운율실현과의 관계에 대한 연구로는 악센트구의 성조 실현은 음절수의 영향을 받는다기보다는 악센트구의 지속시간의 영향을 받는 결과를 보여 주는 연구 결과가 있었다[3]. 예를 들어, 보통 발화속도에서 첫째 음절과 둘째 음절에 실현되는 고음 [H+H]가 느린 발화속도에서는 첫째 음절에만 실현되는 경우들이 종종 관찰되었다. 예를 들어 /콩밥을/을 빠른 속도나 보통 속도로 읽었을 때에는 /콩/과 /밥/에 각각 고음 H와 +H가 실현되었으나 느린 속도에서는 /콩/에만 고음이 실현되었다. /밥을/은 [L+Ha]로 실현되었다. 즉, 빠른 속도에서는 시간이 부족하여 구말 L+가 실현되지 못함으로써 {H Ha}로 실현되었으며 느린 속도에서는 첫째 음절 /콩/이 충분히 길어서 악센트구 초 H와 +H 모두 첫째 음절에 연결되고 오히려 둘째 음절 /밥을/은 L+에 연결되어 저음으로 실현되는 것으로 볼 수 있겠다.



위와 같은 운율 요소들에 관한 음성학적 특성들에 대한 정량화 또는 운율요소들에 미치는 여러 다양한 요인들에 대한 연구가 본격적으로 시작되면 보다 객관적이고 따라서 용이한 기술이 이루어질 것이며 또한 보다 정확한 운율 자동 레이블링화에 한 걸음 더 다가갈 수 있게 될 것이다.

V. 결론

본 논문은 최근 음성공학과 음성학 분야간의 학제간 연구로서 이루어진 운율레이블링 공동 작업에서 사용한 운율레이블링 기준안인 K-ToBI 레이블링 기준안에 대하여 검토하고 작업 시 협력사이트들간의 논의 결과 결정된 수정, 보완 사항을 제시하였다. 그리고 보다 체계적이고 객관적인 운율레이블링을 위한 음성학적 특성들에 대한 정량화 또한 제시하였다.

참고문헌

- [1] 이숙향, *한국어 ToBI 기준안 및 Segmentation 기준안*, 미출판 원고, 한국전자통신연구소 용역 보고서, 1997.
- [2] 이숙향, "한국어 운율구조와 관련한 모음 및 음절 길이," *말소리*, 35-36, pp.13-24.. 1998.
- [3] 이숙향, "한국어 운율구조 기술 체계에 대한 연구: K-ToBI 기술 체계를 중심으로," *언어학* 10-2, pp.1-18, 2002.
- [4] 이현복, (1974). 서울말의 리듬과 억양. *어학연구*, 10(2), pp.15-25, 1974.
- [5] 이호영, *국어음성학*, 태학사, 1996.
- [6] Beckman, M. & S.-A. Jun, "K-ToBI annotation convention." Unpublished manuscript. Ohio State Universti and UCLA, 1995
- [7] S.-A Jun, *The phonetics and phonology of Korean prosody: intonational phonology and prosodic structure*, Doctoral dissertation, Ohio State University, 1993.
- [8] S.-A. Jun, The accentual phrase in the Korean prosodic hierarchy. *Phonology*, 15, 189-226, 1998.
- [9] S.-A. Jun, *K-ToBI(Korean ToBI) Labelling conventions, Version 3.0*. Unpublished manuscript, UCLA, 2000.
- [10] S.-A. Jun, S.-h. Lee, K.H. Kim, & Y.J. Lee, "Labeler Agreement in transcribing Korean Intonation with K-ToBI," *Proceedings of the 2000 International Conference on Spoken Language Processing*, 3, pp.211-214, 2000.
- [11] S.-h. Lee, "Intonational domains of the Seoul dialect of Korean" a paper presented at the 117th meeting of the Acoustical Society of America, Syracuse, N.Y.; An abstract in *Journal of the Acoustical Society of America*, 85, suppl. 1, S99, 1989.
- [12] J.B. Pierrehumbert, *The Phonology and Phonetics of English Intonation*. Doctoral dissertation. MIT, 1980.