

개방 파열의 스펙트럼상의 특성

박 한 상

대구한의대학교 국제어문학부

Spectral Characteristics of Release Bursts

Hansang Park

Faculty of Foreign Language and Literature, Daegu Haany University

E-mail: phans@dhu.ac.kr

Abstract

This study investigates Korean phonation types in terms of spectral characteristics of release bursts. Particularly, this study compares intensity of the speech signal in the release burst, the center of gravity and skewness of the spectra of the release bursts across phonation types for the Korean alveolar plosives. The results showed that there was no significant difference in intensity, center of gravity, or skewness across phonation types but a significant difference across speakers.

I. 서론

파열음은 폐쇄구간(occlusion), 짧은 전이구간(transient), 마찰소음구간(fricative segment), 기식구간(aspirative segment), 그리고 포먼트 전이를 보이는 모음의 앞부분(the initial part of a following voiced sound)으로 이루어져 있다.[1] 즉, CV 연쇄체에서 C가 파열음일 경우 개방전의 조음과정은 폐쇄구간과 개방파열(release burst)로 이루어져 있으며, 개방후의 조음과정은 마찰소음, 기식, 그리고 모음의 앞부분으로 구성되어 있다.[1] 전이구간, 즉 개방파열은 불규칙한 공기의 움직임이 미치는 영향을 제외하면 성도의 압력개방에 대한 반응이다.[1] 개방파열시 전이구간의 길이는 약 2-30 ms이며 일반적으로 10 ms 미만이다.[1] 전이구간의 포먼트 구조는 시간에 따라 매우 심하게 변하고 뒤이어오는 마찰소음과 같은 zero 함수에 영향을 받기 때문에 분명하게 드러나지 않을 수도 있다.[1]

전이구간은 성도의 압력개방에 대한 반응이기 때문에 파열음의 개방시 나타나는 짧은 전이구간의 스펙트럼상의 특성은 조음위치를 연구하는 데 이용되어 왔다.[2] 양순음의 스펙트럼은 평평하거나 떨어지는 기울기를 보이고, 치조음은 증가하는 기울기를 보이며, 연구개음은 증가 후 하강을 보인다고 하였다.[2] 한편 마찰소음에 대하여 스펙트럼의 평균(mean), 무게중심(center of gravity), 왜도(skewness), 첨도(kurtosis) 등의 기술통계적인 자료를 이용하여 어두에 나타나는 장애음의 조음위치를 기술한 연구도 있다.[3] 한편 조음위치의 연구와는 별개로 발성유형의 차이를 연구하는 데에 파열개방의 스펙트럼상의 특성이 이용된 바 있다.[4][5] 특히 한국어의 격음과 경음이 연음보다 개방파열의 에너지가 더 크다고 한 바 있다.[4][5] 본 연구는 조음위치가 같은 한국어 격음, 연음, 경음의 개방파열의 강도(intensity)가 다르며 스펙트럼의 에너지 분포에서 차이를 보인다는 가설을 세우고 문두에 나타나는 한국어 치조파열음의 조음시 나타나는 짧은 전이구간(release burst)에 대하여 음성파형의 강도와 스펙트럼의 무게중심과 왜도를 연구하여 한국어 삼중대립의 특성을 밝히고자 한다.

II. 연구방법

2.1 실험자료

본 연구에서 사용된 음성자료는 한국어 치조파열음 [t^h, t*, t] 및 마찰음 [s^h, s*] 그리고 경구개 파찰음 [tj^h, tj*, tj] 다음에 따라오는 모음 /a/를 포함하는 2음절의 동사원형이다. 각각의 동사원형은 "_____가 아니라 _____다"라는 틀 문장(frame sentence)에 넣어서 발음하였다. 틀 문장의 두 번째 자리에는 동사원형의 두

음절을 서로 바꾸어 넣었다. 본 연구에서는 파열음만을 대상으로 한다.

2.2 녹음

앞 절에서 본 연구에서는 녹음자료 중 일부만 취하여 연구한다고 밝힌 바 있다. 원녹음자료는 동사원형을 포함한 8개의 실험문장이었다. 실험 문장 각각을 10회 반복하여 얻은 240개의 실험문장을 무작위로 순서를 배열한 리스트를 작성하여 3명의 화자에게 규칙적인 시간간격을 두고 읽게 하였다. 화자는 모두 서울 출신의 20대 남성들이었다. 녹음은 UT Austin의 언어학과에 있는 음성실험실에서 이루어 졌다. Marantz SuperScope를 이용하여 아날로그 테이프에 녹음을 한 뒤 SoundScope를 이용하여 디지털 신호로 옮겼다. 표본추출률(sampling rate)은 22,050 Hz였다.

2.3 측정

개방파열시 전이구간의 길이가 2-30 ms 이며 일반적으로 10 ms 이하라는 지적이 있었고[1] 전이구간의 스펙트럼을 얻기 위하여 6.4 ms의 음성샘플을 채취하여 연구한 적이 있다.[6] 실제음성자료를 살펴본 결과 경음의 경우 개방파열의 전이구간이 끝난 뒤 곧바로 성대진동이 시작되는 것을 관찰할 수 있었다. 본 연구에서는 성대진동부분이 포함되는 것을 피하기 위하여 5 ms의 음성신호를 추출하였다. 개방파열의 시작점은 스펙트로그램을 참조하여 파형에서 심한 변화를 보이기 시작하는 점으로 정하였고 모든 음성자료에 대하여 일괄되게 5 ms의 음성신호를 얻었다. 그림 1, 2, 3에 음성신호 추출의 예가 제시되어 있다.

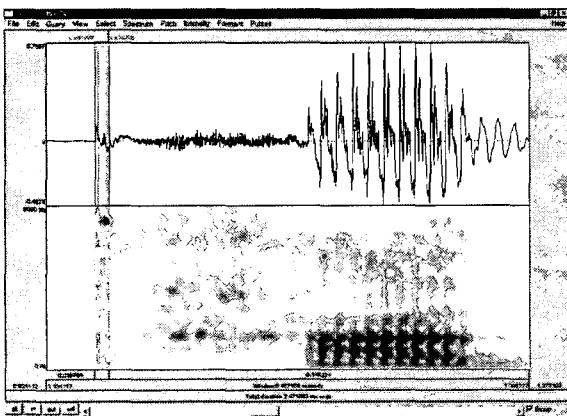


그림 1 [ʰada]의 개방파열 구간

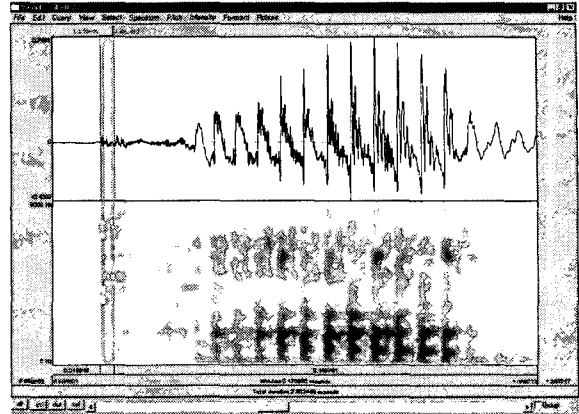


그림 2 [tada]의 개방파열 구간

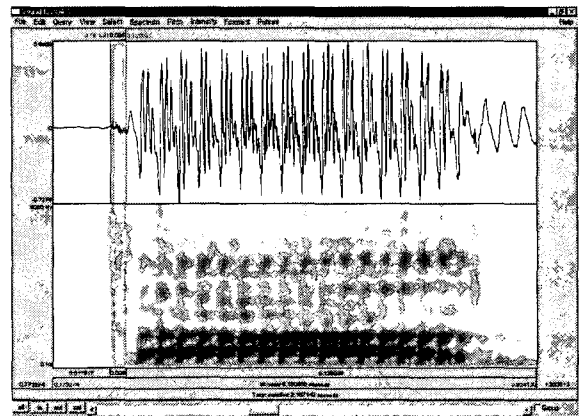


그림 3 [t'ada]의 개방파열 구간

음성신호의 추출 및 측정에 Praat 4.1.5를 이용하였다. 추출된 각각의 음성신호에 대하여 먼저 강도를 측정하였고 다음으로 그 음성신호의 스펙트럼에서 무게 중심과 왜도를 측정하였다. 추출된 음성신호에 대한 파형과 스펙트럼이 그림 4와 그림 5에 각각 제시되어 있다.)

- 1) 그림 5에서는 앞에서 언급한 스펙트럼의 기울기 모양, 치조음은 증가하는 기울기를 보이지 않는다. 이것은 preemphasis를 하지 않았기 때문이다. 본연구에서는 preemphasis를 하지 않은 자료를 바탕으로 먼저 연구한다. preemphasis를 한 자료는 다음의 연구로 미룬다.

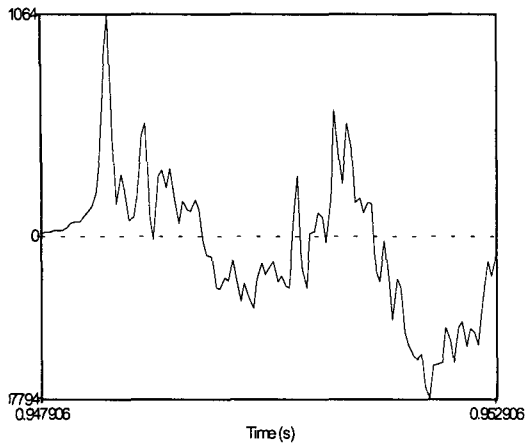


그림 4 [t^{na}]의 개방파열 구간의 파형

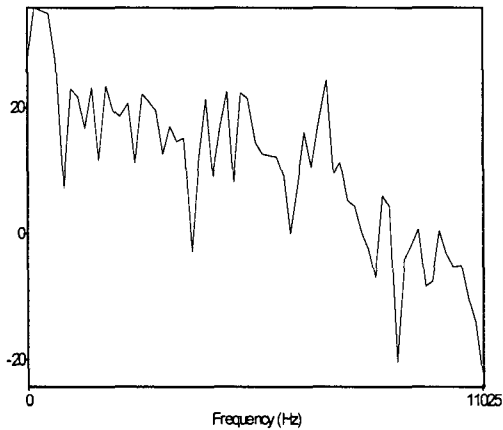


그림 5 [t^{na}]의 개방파열 구간의 스펙트럼

강도(intensity)는 식 (1)과 같이 계산된다.

$$10 \log_{10} \left(\frac{1}{nP_0} \right) \sum_{i=1}^n x_i^2 \quad (\text{식 } 1)$$

식1에서 n 은 표본의 수이고, P_0 는 인간의 청취 문턱치 압력(auditory threshold pressure), 즉 2×10^{-5} Pa이며, x 는 각 표본의 진폭(amplitude)을 나타낸다. 무게중심은 식 (2)에 의해 계산되었다.

$$\frac{\int_0^{\infty} f S(f) |^p df}{\int_0^{\infty} |S(f)|^p df} \quad (\text{식 } 2)$$

식 (1)에서 f 는 주파수이고 $S(f)$ 는 스펙트럼이며 p 는 파워이다. 위의 식 (1)에서 분모는 스펙트럼의 에너지

를 나타내며 분자에서 p 값이 1일 경우 절대 스펙트럼, 즉 복소수로 표현되는 FFT의 결과가 되며 p 값이 2일 경우 파워스펙트럼이 된다. 본 연구에서 파워는 2로 하였다. 따라서 무게중심은 전 주파수 영역에 대하여 파워스펙트럼에 의해 가중치가 주어진 주파수의 평균값이다. 스펙트럼의 무게중심은 에너지 분포에 대하여 주파수의 평균값이 얼마나 높은가를 측정함으로써 어떤 음성신호가 가지는 스펙트럼상의 특성을 나타내는 것이다. 그리고 왜도는 에너지 분포가 얼마나 한쪽으로 치우쳐져 있는가를 재는 것이다. 그림 5에서 제시된 스펙트럼의 무게중심은 약 988 Hz이다.

왜도는 식 (3)과 같이 계산된다.

$$\frac{\int_0^{\infty} (f-f_c)^3 |S(f)|^p df}{\int_0^{\infty} |S(f)|^p df} \quad (\text{식 } 3)$$

위의 식 (3)에서 f_c 는 스펙트럼의 무게중심을 나타낸다. 본 연구에서는 파워를 2로 하였다. 그림 5에서 제시된 스펙트럼의 왜도는 약 2.6이다.

III. 연구결과

실험결과 나타난 강도(Intensity), 무게 중심(Center of Gravity), 왜도(Skewness)에 대하여 화자(Speaker)간 그리고 발성유형(Phonation Type)간 차이를 보여주는 평균과 표준편차 값이 표 1에 제시되어 있다.

		Phonation Type					
		Aspirated		Larynx		Fornis	
		평균	표준편차	평균	표준편차	평균	표준편차
Speaker LGS	Intensity	58.42	2.77	56.56	4.58	57.30	3.19
	Center of Gravity	1161	424	1107	470	1400	524
	Skewness	2.24	.96	2.42	.94	1.77	.76
LSS	Intensity	61.32	2.37	60.97	2.49	62.72	2.95
	Center of Gravity	1224	334	1231	584	1109	400
	Skewness	2.04	.85	2.50	.98	2.38	.71
SHJ	Intensity	58.07	2.38	58.84	2.72	59.05	2.10
	Center of Gravity	1094	370	705	391	830	478
	Skewness	2.27	.91	3.49	1.36	3.32	1.33

표 1 강도, 무게중심, 왜도의 평균과 표준편차

표 1에 드러난 강도, 무게중심, 왜도에 대한 평균값의 차이가 화자간 그리고 발성유형간에 통계적으로 유의미한지 검증하기 위하여 강도, 무게중심, 왜도 등 각각의 종속변수에 대하여 화자와 발성유형을 독립변수로 하는 일변량이원분산분석(two-way ANOVA)을 실시하였다. 먼저 강도에 대한 분석 결과 화자와 발성유형 사이에 유의미한 상호작용효과(interaction effect)가 없었다. ($F(4,81) = 0.814, p > 0.05$) 발성유형에 대해서는 집단간 유의미한 차이가 없었다. ($F(2,81) = 0.719, p >$

0.05) 반면에 화자에 대해서는 집단간 유의미한 차이가 있었다.(F(2,81) = 16.776, p < 0.0001) Tukey's HSD를 이용한 사후검증 결과 LSS가 SHJ나 LGS보다 유의미하게 컸다. 다음으로 무게중심에 대한 분석 결과 화자와 발성유형 사이에 유의미한 상호작용효과가 없었다.(F(4,81) = 1.294, p > 0.05) 발성유형에 대해서는 집단간 유의미한 차이가 없었다.(F(2,81) = 0.820, p > 0.05) 반면에 화자에 대해서는 집단간 유의미한 차이가 있었다.(F(2,81) = 5.447, p < 0.01) Tukey's HSD를 이용한 사후검증 결과 SHJ가 LSS나 LGS보다 유의미하게 작았다. 마지막으로 왜도에 대한 분석 결과 화자와 발성유형 사이에 유의미한 상호작용효과가 없었다.(F(4,81) = 1.639, p > 0.05) 발성유형에 대해서는 집단간 유의미한 차이가 없었다.(F(2,81) = 3.028, p > 0.05) 반면에 화자에 대해서는 집단간 유의미한 차이가 있었다.(F(2,81) = 6.882, p < 0.01) Tukey's HSD를 이용한 사후검증 결과 SHJ가 LSS나 LGS보다 유의미하게 컸다.

요약하면, 강도, 무게중심, 왜도 모두에 대하여 발성유형 사이에는 유의미한 차이가 없었으나 화자 사이에는 유의미한 차이가 있었다. 이는 발성유형이 다르더라도 개방파열의 스펙트럼에는 의미있는 차이가 없다는 것을 의미한다. 이러한 결과는 기존의 연구와 다르다.[4][5] 다른 이유는 기존의 논문에 자세한 방법이 나와 있지 않아 알 수 없다.

IV. 결론

본 연구에서는 치조음(격음, 연음, 경음)은 조음위치가 같으므로 무게중심과 왜도에서는 유의미한 차이가 없으며 강도에서는 유의미한 차이가 있다는 가설을 세우고 실험을 통해 살펴보았다. 실험결과 강도, 무게중심, 왜도 모두에 대하여 발성유형 사이에는 유의미한 차이가 없었으나 화자 사이에는 유의미한 차이가 있었다. 이 결과를 바탕으로 무게중심과 왜도에 대해서는 가설을 받아들이고 강도에 대해서는 가설은 기각한다. 본 연구를 통해 발성유형이 다르더라도 조음위치가 같으면 개방파열의 스펙트럼에는 큰 차이가 없음을 확인하였다.

참고문헌

- [1] G. Fant, Stops in CV-Syllables, Speech Sounds and Features, pp. 110-139, 1973.
- [2] S. Blumstein, K. Stevens, Perceptual invariance and onset spectra for stop consonants in different vowel and environments, JASA 67, pp. 648-62, 1980.
- [3] K. Forrest, G. Weismer, P. Milenkovic, R. Dougall (1988) Statistical Analysis of Word-initial Voiceless Obstruents: Preliminary Data, JASA 84, pp. 115-123, 1988.
- [4] C.-W. Kim, On the autonomy of the tensity feature in stop classification, Word 21, pp. 339-359, 1965.
- [5] T. Cho, S.-A. Jun, P. Ladefoged, Acoustic and aerodynamic correlates to Korean stops and fricatives, Journal of Phonetics 30(2), pp. 193-228, 2002.
- [6] K. Stevens, S. Manuel, M. Matthies, Revisiting place of articulation measures for stop consonants: Implications for models of consonant production, Proceedings of the 14th ICPhS, pp. 1117-1120, 1999.