

감정 음성의 국어 발화 말 경계성조 연구

박 미 영

고려대학교 국어국문학과

Research of Korean utterance-final boundary tones in Emotion speeches

Miyoung Park

Department of Korean language & literature, Korea University

E-mail : pikel@unitel.co.kr

Abstract

The purpose of this paper is to find boundary tone's characteristics in Korean emotion speeches. I mainly focus on investigating patterns and f0 values of boundary tones and f0 values in utterance final phrases.

I. 서론

국어의 억양 연구에서 경계성조에 대한 정의와 음성적 특징은 이전의 많은 연구들에서 체계적으로 진행되어 왔다 [2],[3],[4],[5]. 화자가 전달하고자 하는 문법적, 화용적, 준언어적(화자의 감정 및 태도 등) 의미를 전달하는 기능을 하는 경계성조는 언어학적 해석으로만 정의되는 것이 아니라 음성 합성과 같은 응용 분야에서도 중요한 역할을 하는 음성적 특징이다. 선행 연구 [2]와 [4]에서는 국어 경계성조의 억양 유형 기술뿐만 아니라 그에 대한 음향음성학적인 분석 근거를 들어 국어 억양의 경계성조에 대한 기준안들을 보여줬다. 특히 경계성조를 통해 나타나는 화자의 감정 및 태도 등을 분석함으로써 음성 연구가 최종적으로 나아가야 하는 연구의 방향을 제시했다고 본다.

그러나 음성 발화라는 것이 너무나 복잡적이고 광범위하기 때문에 국어의 억양 연구는 많이 제한될 수밖에 없었다. 그래서 본 연구자는 매우 역동적이고 다양한 음성 발화를 종합적으로 분석할 수는 없지만, 감정을 실현하는 데에서 화자들은 어떠한 음성적 특징, 특히 경계성조를 어떠한 양상으로 실현하는지에 초점을 맞추어 감정 발화의 경계성조 연구를 시작하게 되었다. 이 연구는 자연 발화가 아닌 녹음실에서 텍스트 기반의 음성 자료를 이용하였다는 단점은 있지만, 기존의 감정 음성 연구에서 좀더 나아가 국어 억양 연구, 경계성조에 접근할 수 있는 결과들을 해석해보고

자 진행되었다. 이 연구에서 '기쁨, 슬픔, 화남' 감정과 평상 중립 발화 시에 나타나는 음성을 수집하였다.

II. 실험 및 분석

본 연구를 위해 분석된 음성 자료는 동일한 녹음 텍스트로 감정 연기를 수행한 전문 성우의 발화이다. 성우 전문 녹음실 '보이스위즈(voicewiz)'를 통해 감정 표현의 전문 연기가 가능한 성우들을 섭외하였으며, 6명의 녹음이 진행되었는데, 이번 분석에 이용된 전문 여자 성우 2명은 다음과 같다.

<표 1. 감정 음성 표현 전문 성우 화자 정보>

파일명	화자명	성 별	나 이
F02	ACH	여	31
F03	KIJ	여	27

감정 연구는 실험 설계에서 외부 변인 요소들이 많기 때문에 화자가 실험실에서 직접 감정을 생산하고 실제처럼 얼마나 잘 구현했으며, 어떠한 정밀한 측정을 하였는지가 매우 중요하다[6].¹⁾ 따라서 본 연구자들은 낭독체형 문장이 아닌, 실생활에서 자주 사용될 수 있는 구어체 문장으로 감정 표현이 유발되는 대화 텍스트를 작성하였다.

국어의 경계성조 실현을 보기 위해 종결형을 통일시켰는데, 구어체 대화에서 많이 사용되는 종결 어미 '-

1) 물론 녹음실 환경에서 수집된 감정 음성에서 화자들이 얼마나 감정을 자연스럽게 완벽하게 표현했는지를 판단하는 청각적 확인이 우선시 되어야 하겠지만, 이 연구는 화자들이 감정을 어떻게 표현하는가에 관한 음성의 생성이라는 측면에서 먼저 접근하고자 한다. 실제 자연 발화는 매우 복잡적이고, 외부 변인요소들이 많기 때문에 감정 표현의 공통적인 근거들을 찾기에 어려움이 있다.

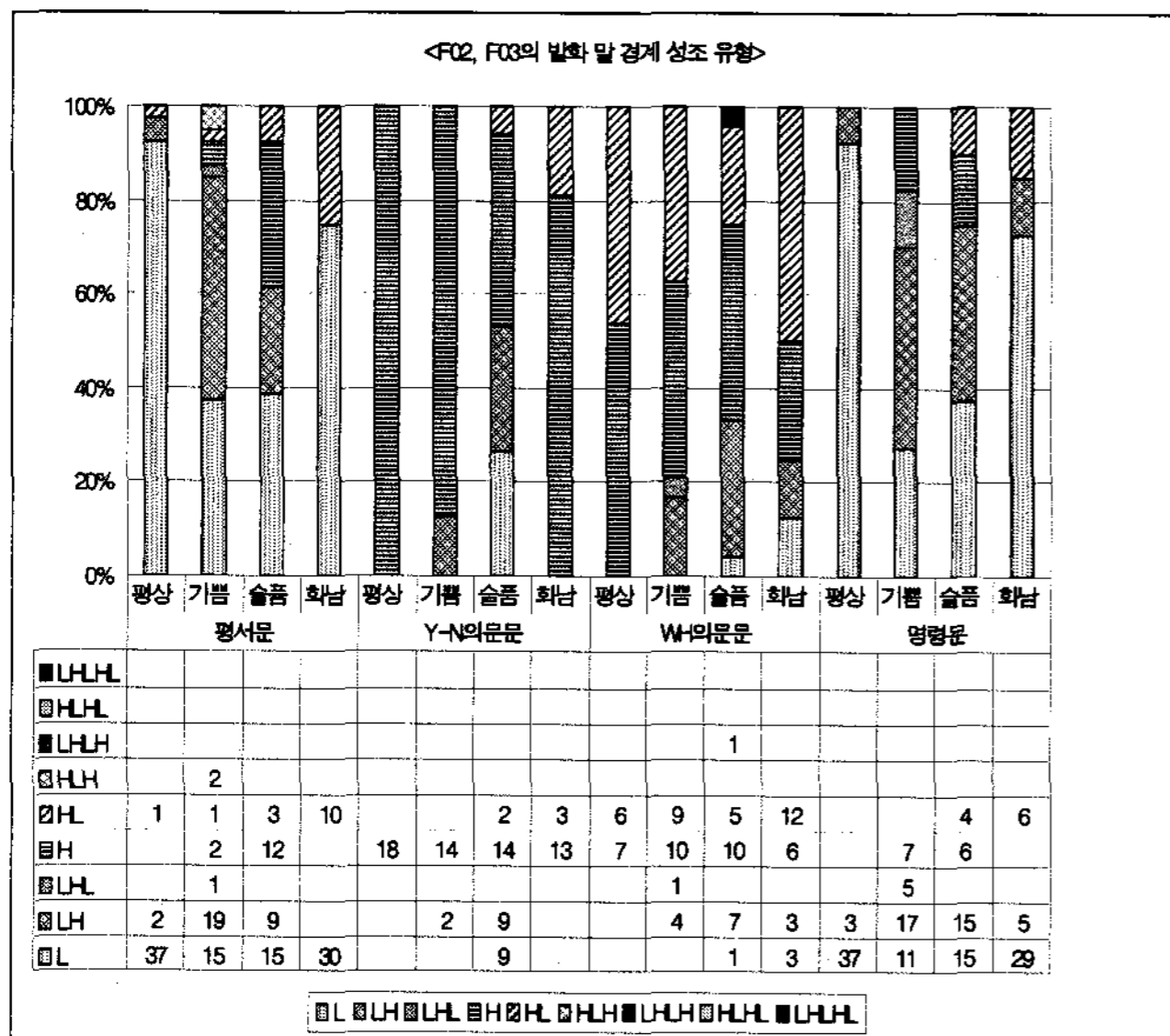
요'로 끝나는 평서문 10개, 의문문 10개(가부 의문문 4개, 의문사 의문문 6개), 명령/청유문 10개의 문장이 들어간 대화 텍스트를 2회씩 반복하여 발화하였다.

녹음에 사용된 장비는 믹서 MACKIE 1604-VLZPRO, 오디오 인터페이스 ESI WaMi Rack 192X, 컴프레서 JOEMEEK VC3Q, 마이크 RODE NT2가 사용되었으며 Cool Edit Pro 2.1에서 22050Hz, 16bit로 디지털화되었다.

분석은 우선 발화 말 억양구의 경계성조 유형 기술은 [5]의 기준을 따랐으며, 발화 문장 전체의 기본주파수와 발화 말 억양구의 경계성조가 실현된 마지막 음절의 H와 L의 기본주파수, 발화 말 강세구의 기본주파수의 최고값(H)과 최저값(L)을 측정 대상으로 하였다. WaveSurfer 1.8.2를 이용하여 기본주파수를 추출하여 경계성조와 발화 말 강세구 기본주파수의 최고값과 최저값을 측정하였는데, 이때 청각적으로 인지되고 억양 곡선에서 안정적인 구간을 중심으로 하였다. 각 발화 문장의 전체 기본주파수값은 pitch_chal 프로그램을 이용하여 최대값, 최소값, 평균 각각의 평균값들을 추출하였다.

III. 결과 및 논의

3.1 발화 말 경계성조 유형



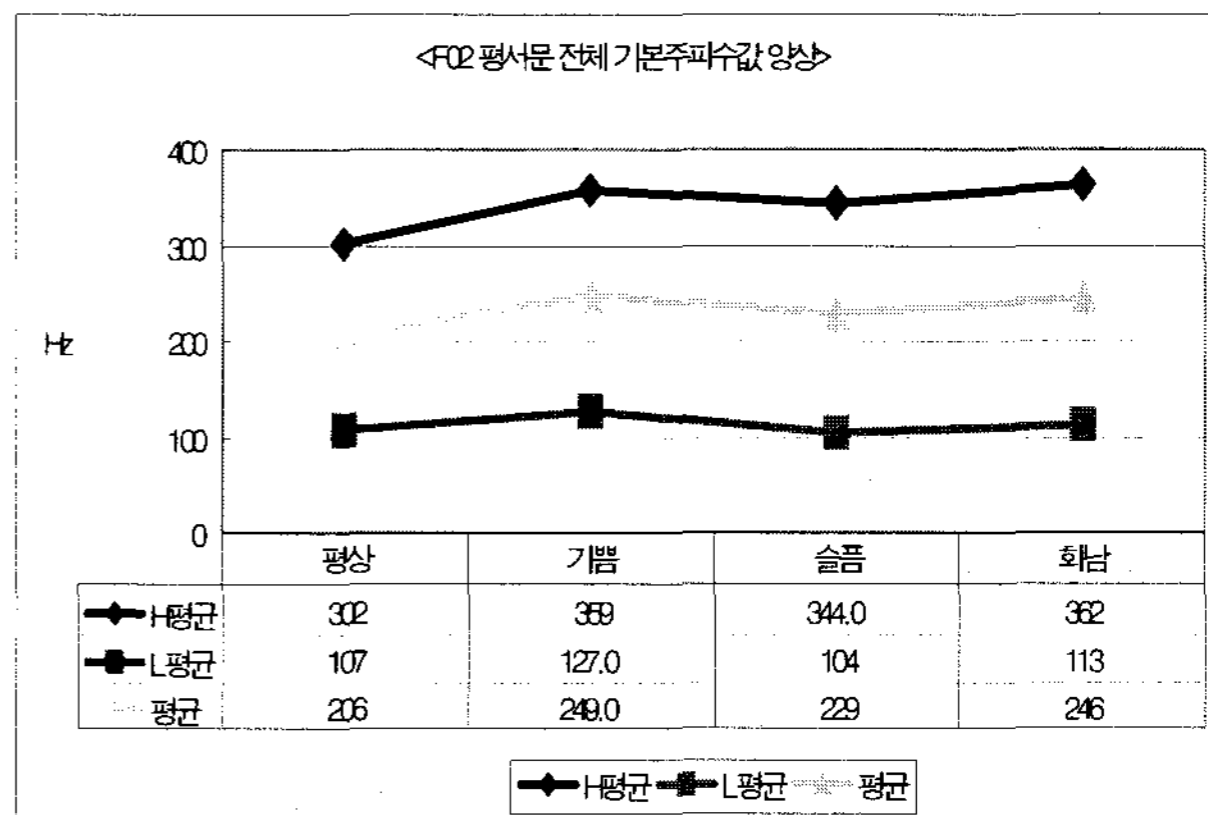
<그림 1. 감정별, 문장 유형별에 따른 발화 말 경계성조 유형>

<그림1>은 화자 2명의 경계성조 유형이다. <그림1>에서 보듯이, 각 문장 유형별로 감정을 표현하지 않는 중립 발화를 하였을 때, 국어 발화 말 경계성조 유형은 선행 연구들과 같은 결과를 보였다. 평서문과 명령/청유문에서는 주로 L%가 실현되었고, 가부의문문에서는 H%, 의문사의문문에서는 H%와 HL%가 나타났다. 그러나 감정을 표현할 때에는 '기쁨' 발화의 평서문

과 명령/청유문에서 LH%가 L%와 비슷하게 나타난다. '슬픔' 발화의 평서문과 명령/청유문에서 H%가 나타났는데, '슬픔' 감정을 표현하기 위해서 발화를 완벽하게 끝내지 못하고 상승조 억양 곡선에서 대체로 발화 마지막 음절 길이가 짧게 나타난 것이 특징이다. 감정 발화에서는 주로 LH%와 HL% 유형이 문장별, 감정별로 고르게 나타났는데, 이때에 H기본주파수값뿐만 아니라 L기본주파수값은 평상 감정에서보다 주파수 대역이 높게 실현되는 것으로 나타났다. [5]에서 제시된 H와 L에 따른 이분법적 억양 유형 기술이 제한적이라는 것과 더불어 [2]에서 논의된 M% 설정 근거를 지지할 수 있는 근거들을 볼 수 있었는데, 다음에서 살펴보겠다.

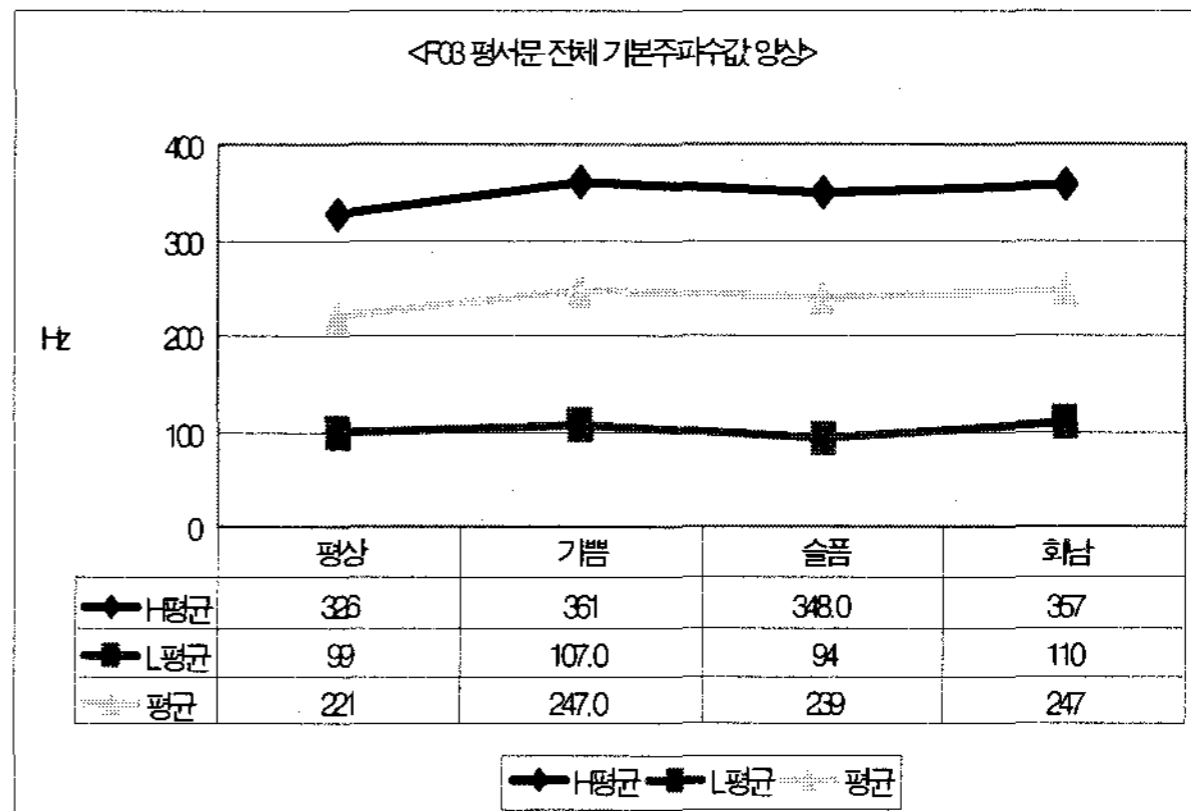
3.2 기본주파수

아래의 <그림2>는 F02가 평서문을 발화하였을 때, 전체 발화문장 기본주파수값의 최대값, 최소값, 평균값의 평균들이다.



<그림 2. F02의 평서문 전체 기본주파수값 양상>

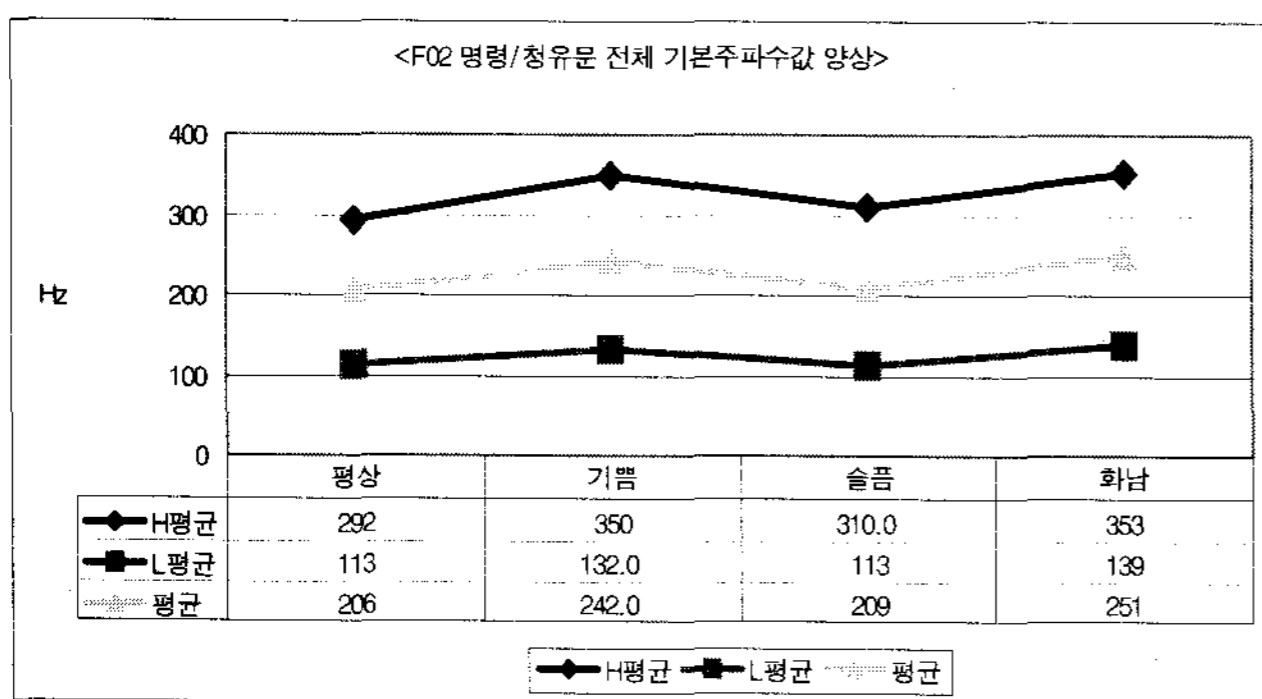
감정 음성들이 평상 발화에서보다 최대값들이 커지는 것을 알 수 있다. 최저값은 크게 차이가 없다. 최고값들의 변화에 따라 평균값이 영향을 받는다.



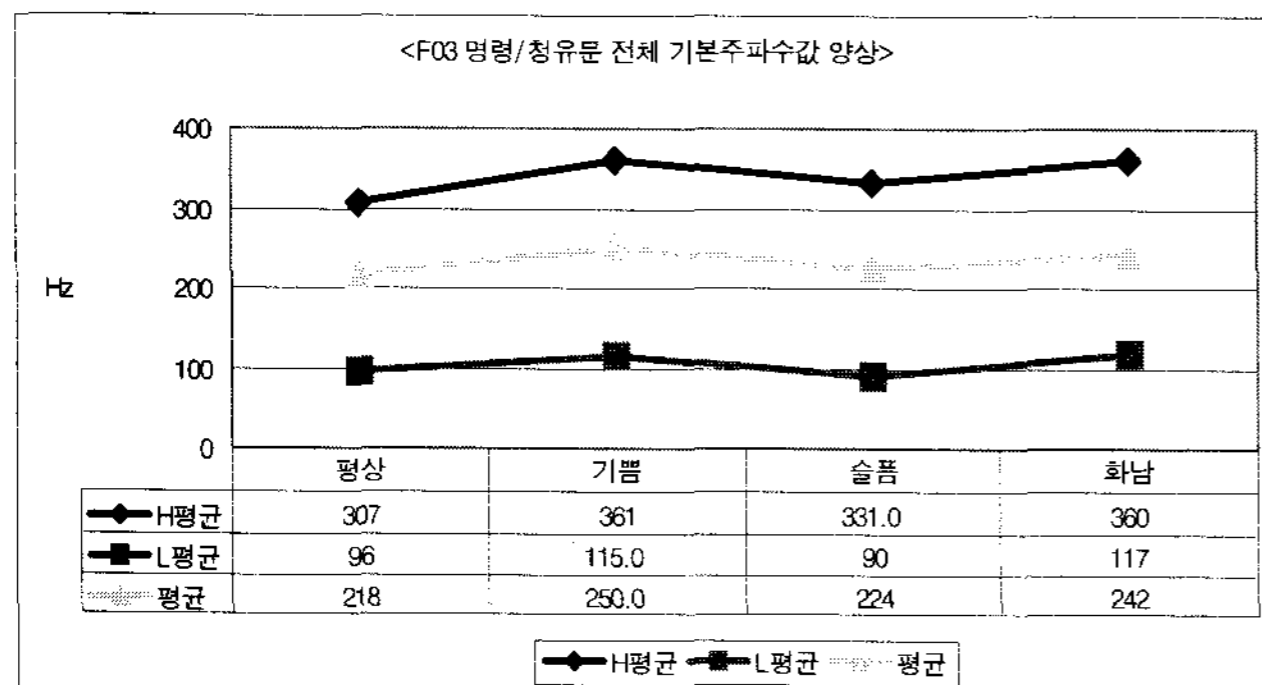
<그림 3. F03 평서문 전체 기본주파수값 양상>

<그림3>에서도 F03의 전체 기본주파수값의 양상이 같은 결과를 보임을 알 수 있다. 두 화자 모두 의문문에서도 최저값들은 평상 발화와 거의 차이가 없고, 최고값들이 더 높은 값을 알 수 있었다.

그러나 명령/청유문의 '슬픔' 감정 발화에서는 최고값의 변화가 '기쁨'이나 '화남' 수준에는 미치지 못하고, 평상 발화와 비슷한 수준을 보이는데, '슬픔'의 감정으로 상대방에게 어떤 행동을 지시하는 명령/청유 의미를 전달한다는 발화 행위가 대화 맥락 상 미흡한 것으로 보인다. <그림4, 5>에서 F02, F03의 명령/청유문 발화 시 전체 기본주파수값들의 실현 양상을 볼 수 있다.



<그림 4. F02 명령/청유문 전체 기본주파수값 양상>

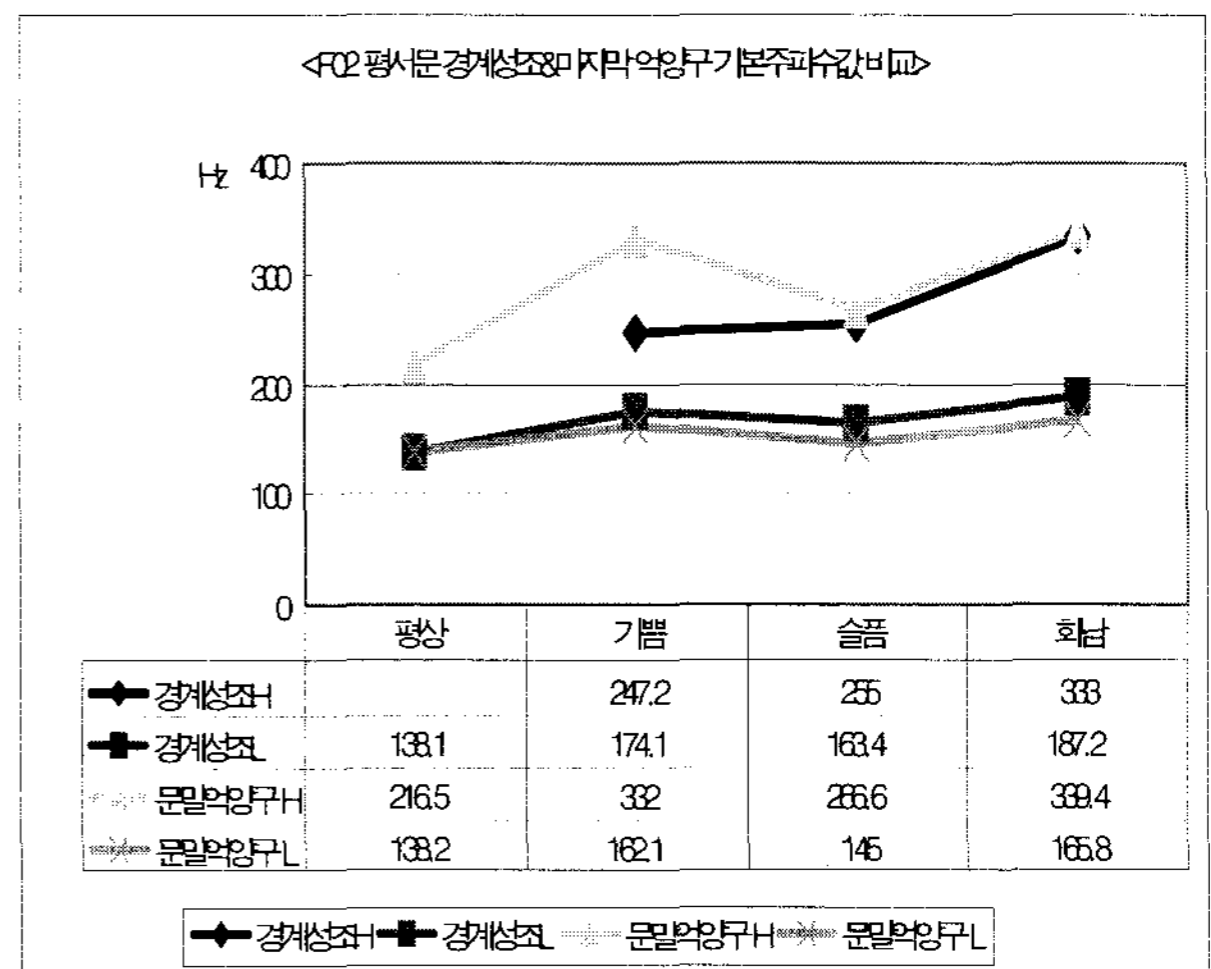


<그림 5. F03 명령/청유문 전체 기본주파수값 양상>

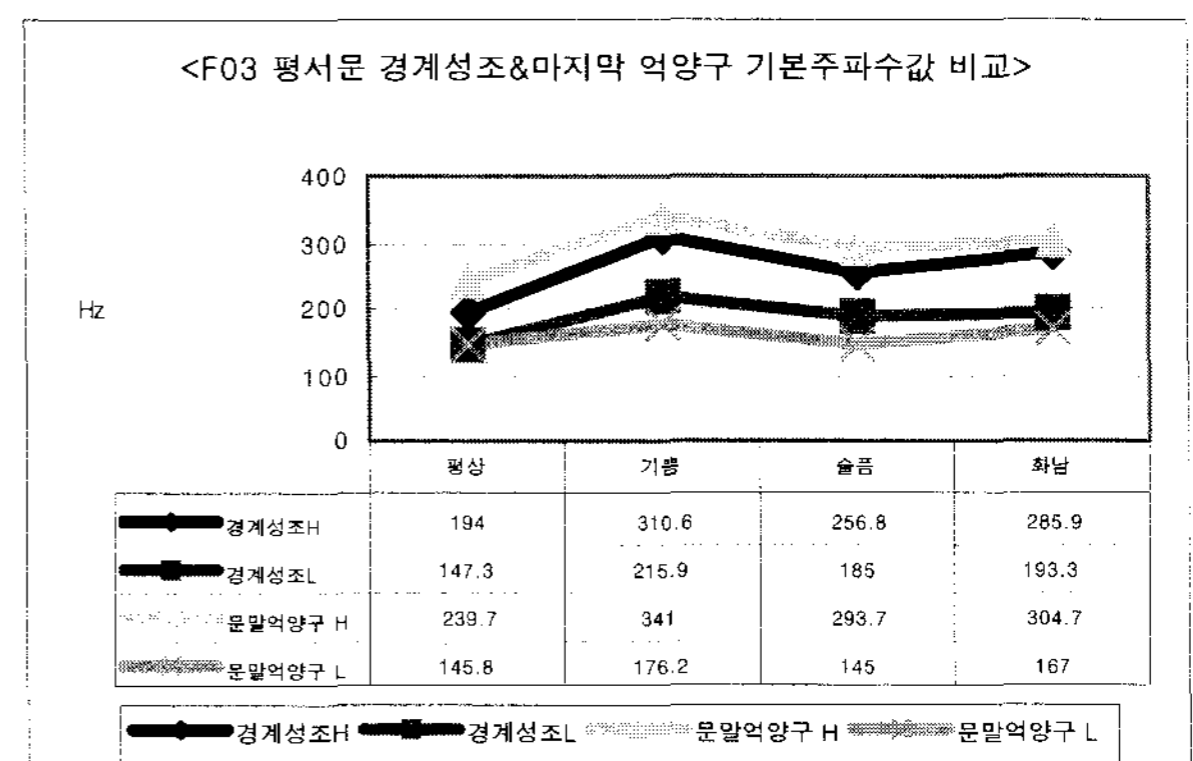
그러나 감정 발화에서 최대값들이 변화하고 최저값들은 평상 발화와 유사하다는 결과는 발화 말 경계성조에서는 조금 다른 양상을 보인다. 다음에서는 F02와 F03의 발화 마지막 억양구 말 강세구 내에서 최고값 H와 최저값 L의 기본주파수와 발화 말 경계성조 유형들의 기본주파수값을 비교하였다.

다음의 <그림6>에서 F02의 평서문의 평상 발화에서 경계성조는 L%만 나타났기 때문에 H값이 표시되지 않았다. F03의 <그림7>과 함께 논의를 해보면, 감정 음성에서 화자들은 모두 마지막 억양구 말 강세구 H가 평상 발화보다 높게 나타난 것을 알 수 있다. 그러나 L값은 '기쁨'과 '화남'뿐만 아니라, '슬픔'에서도 평상 발화보다 높게 실현되었다. 경계성조의 기본주파수

값을 볼 때, 기존 연구들에서 낮은 기본주파수값이 '슬픔' 감정의 음성 특징이라고 하기에는 무리가 있다.



<그림 6. F02 평서문 경계성조와 마지막 강세구 H, L값 비교>



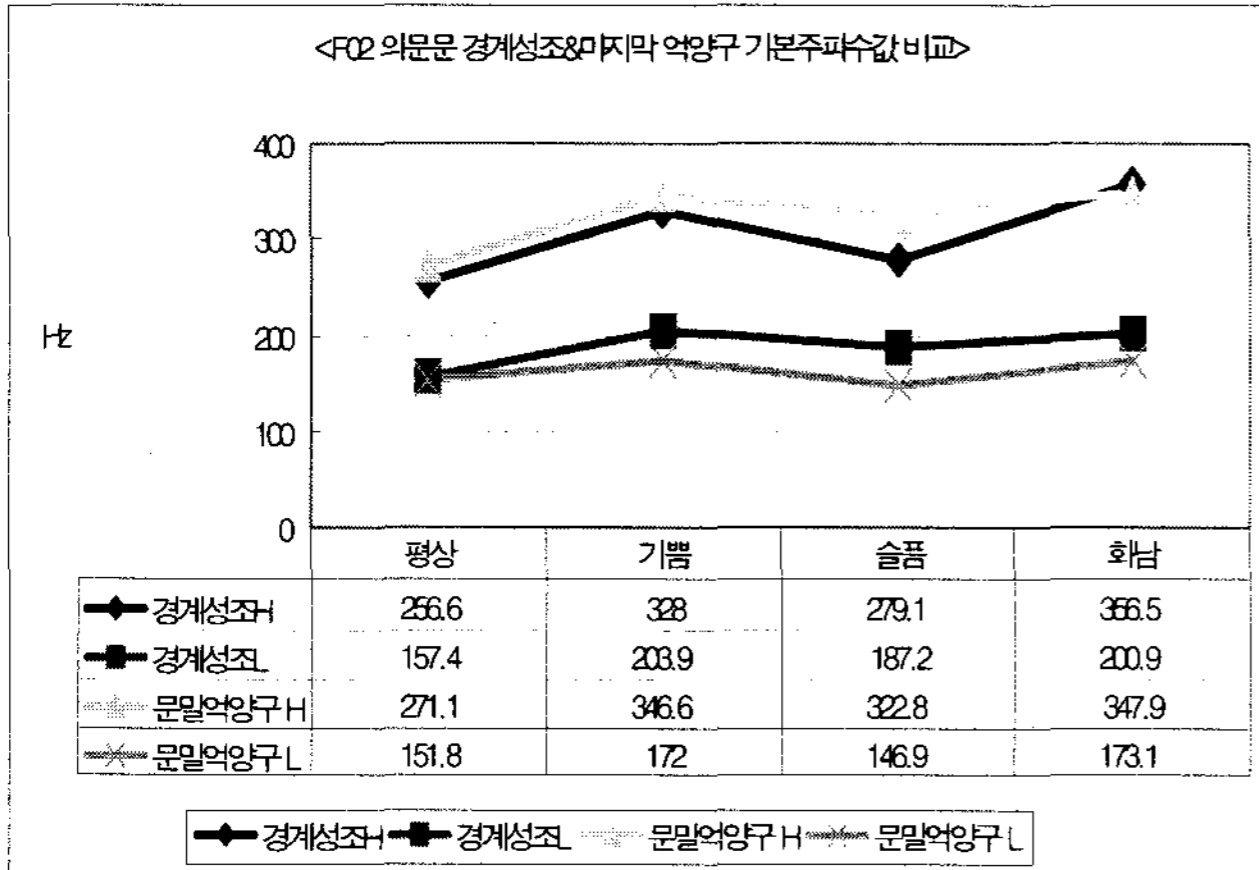
<그림 7. F03 평서문 경계성조와 마지막 강세구 H, L값 비교>

또한 마지막 강세구의 H와 L의 기본주파수값을 전체 발화문장의 기본주파수값 양상과 비교해 보면, L값이 평상 발화에서도 높게 나타난 것을 볼 때, 국어는 발화 말에서 억양 의미가 대부분 실린다는 것을 음향 음성학적으로도 확인할 수 있었다.

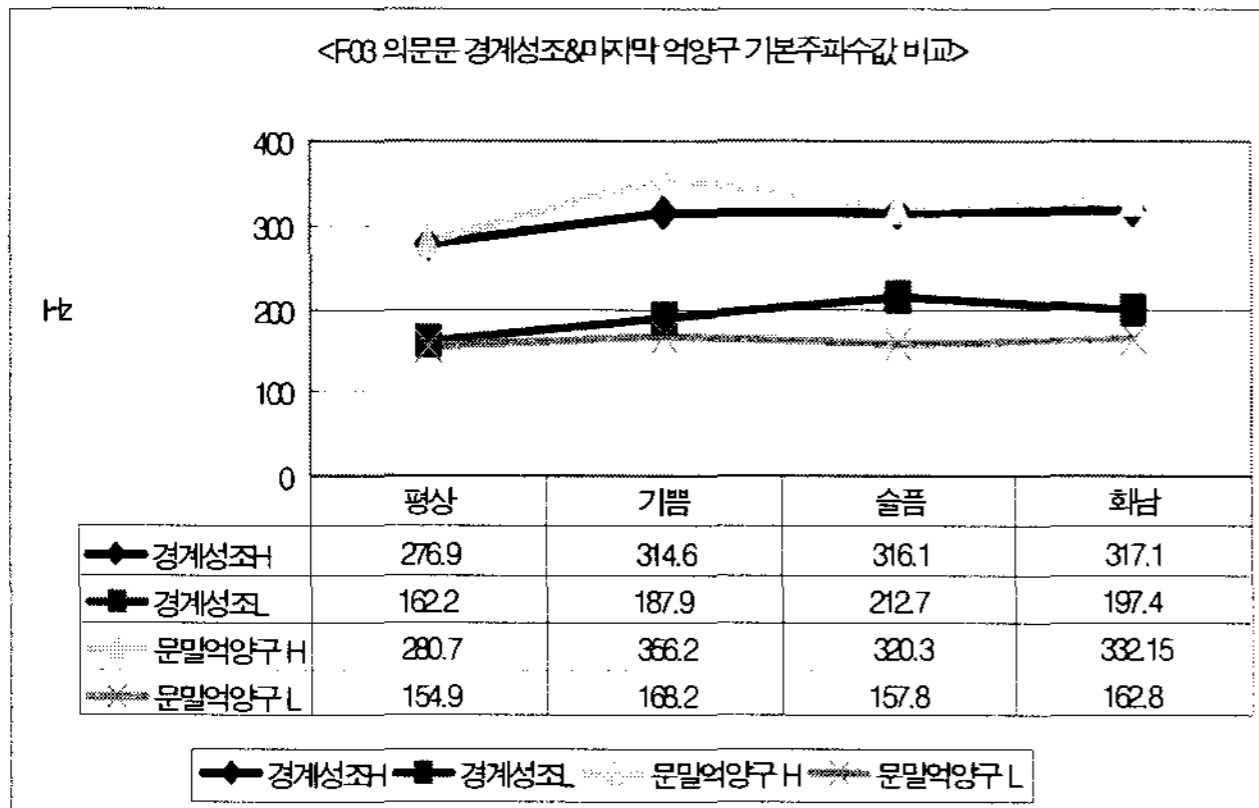
다음의 <그림8, 9>는 의문문에서의 경계성조를 보여 준다. 평서문의 발화에서와 마찬가지로 경계성조의 H와 L의 기본주파수값은 평상 발화보다 모두 상승하였는데, 의문문에서 실현되는 경계성조 H값이 매우 높게 나타나므로 F03의 '기쁨' 음성을 제외하고는 L값들도 평서문에서보다 더 높게 나타나고 있다. 경계성조 분석에서 M톤 설정을 결정하면 논의가 용이할 것이다.

F03은 '슬픔' 감정에서 '기쁨'과 '화남'에서보다도 더 높게 경계성조의 L값이 나타난다. 화자가 발화 마지막 음절을 '슬픔'을 표현하기 위해 안정적으로 발음하지 않고 빨리 끝내면서, HL% 등의 경계성조 유형을 실현

할 때 나타나는 특징으로 보인다.

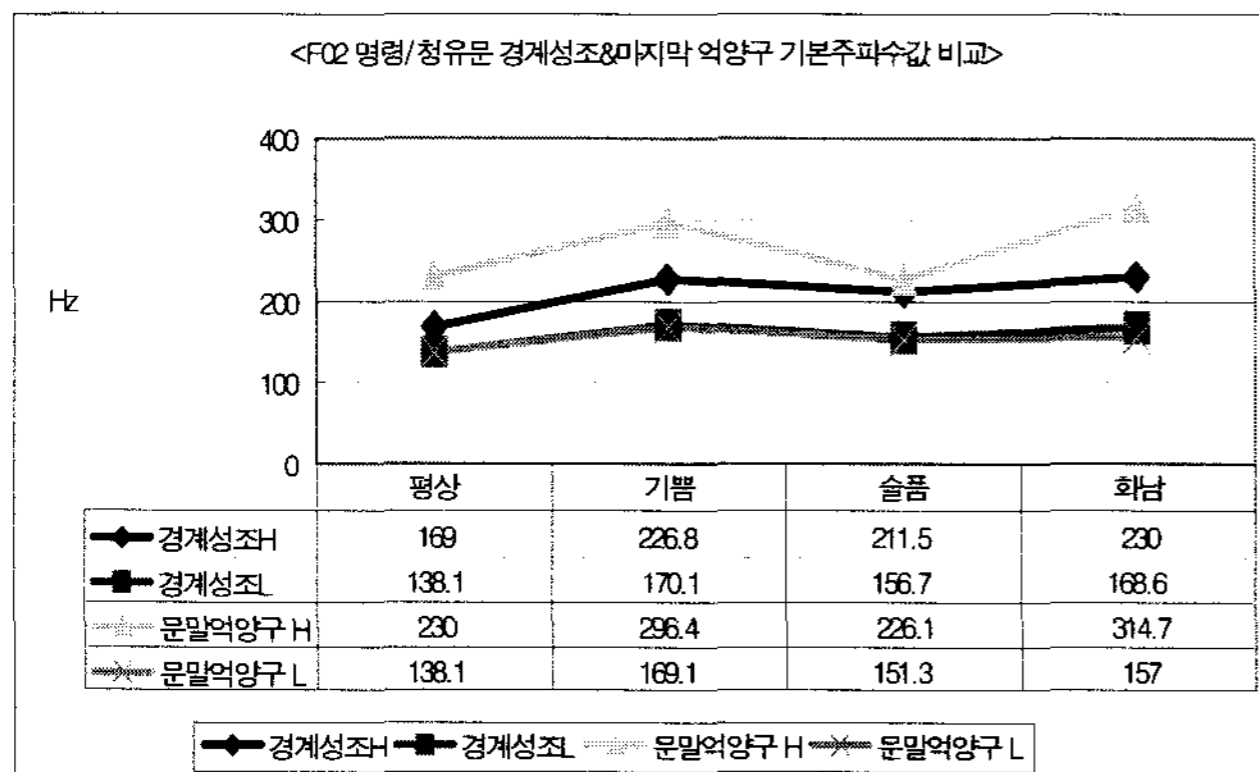


<그림 8. F02 의문문 경계성조와 마지막 강세구 H, L값 비교>

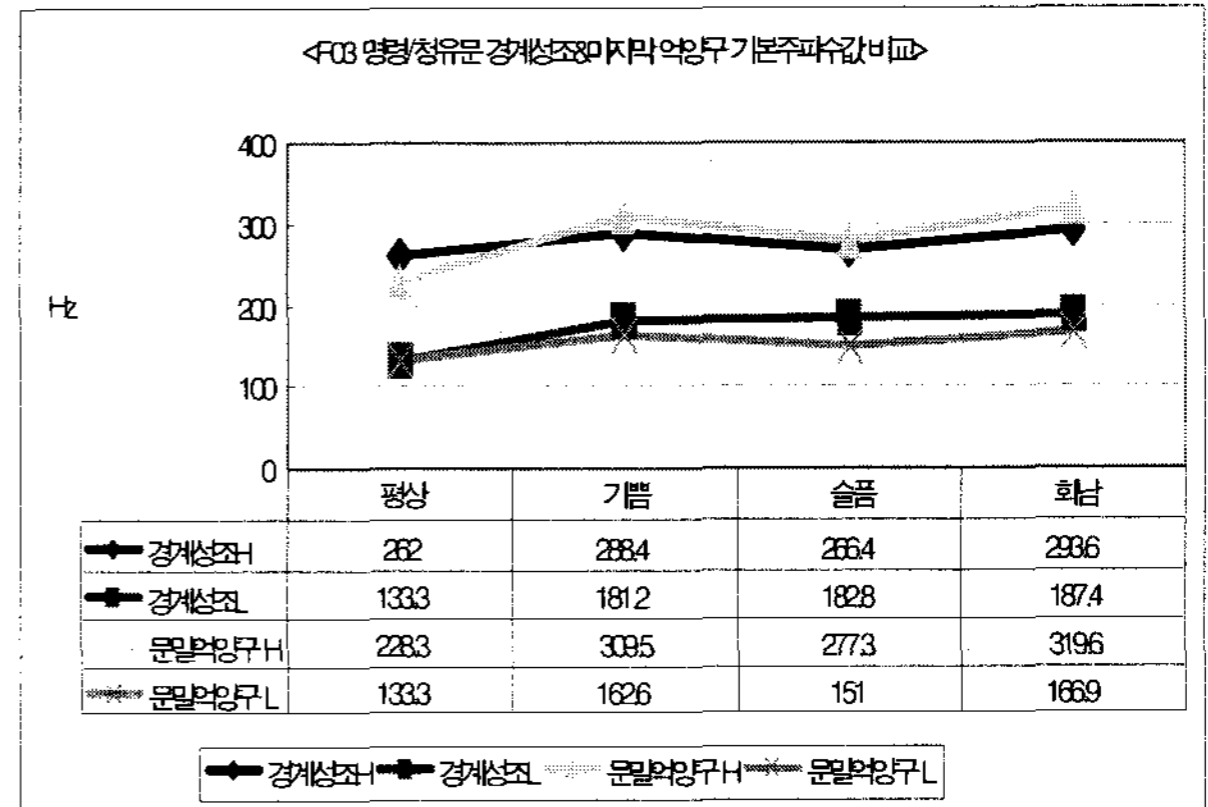


<그림 9. F03 의문문 경계성조와 마지막 강세구 H, L값 비교>

명령/청유문을 발화한 다음의 <그림10, 11>에서는 경계성조의 H값과 L값이 전체 감정 음성에서 유사하게 나타나는 것이 특징이다.



<그림 10. F02 명령/청유문 경계성조와 마지막 강세구 H, L값 비교>



<그림 11. F03 명령/청유문 경계성조와 마지막 강세구 H, L값 비교>

화자들이 감정을 표현하는 데에 명령/청유형 문장 유형에서는 종결어미 '-요'가 전달하는 화용적 의미 특성²⁾의 영향을 받는 것으로 보인다.

IV. 결론

이 연구에서는 분석 자료가 여성 화자 2명으로 부족하지만, 추가 분석 대상인 전문 성우 남녀 화자, 아마추어 남녀 화자에서도 비슷한 경향을 보인다. 객관적인 통계 검증과 실제 자연 발화를 대상으로 분석 결과를 적용해 봐야 하는 과제를 남겨둔다.

참고문헌

- [1] 장인창, 이태승, 박미경, 김태수, 장동식(2005), "발화 내 감정의 정밀한 인식을 위한 한국어 문미억양의 활용." 한국통신학회 논문지 vol.30, 한국통신학회
- [2] 이호영(1998). "국어 핵억양의 음향음성학적 연구." 말소리 제38호. 대한음성학회.
- [3] 한선희, 오미라(1999). "한국어 억양구의 경계톤." 음성과학 제5권 제2호. 한국음성과학회.
- [4] 한선희(2000). "대화체 억양구말 형태소의 경계성조 연구." 음성과학 제7권 제4호. 한국음성과학회.
- [5] Jun, Sun-Ah(2000). "K-ToBI labelling conventions; version 3.1," *UCLA Working Papers in Phonetics*, 99, Journal of Speech Science, Vol. 7, No. 1:143~169.
- [6] Iain R. Murray and John L. Aronott(1993). "Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion." *J. Acoust. Soc.* 93-2. Acoustical Society of America

2) 종결어미 '-요'는 비격식체 청자 높임, 친근함의 의미를 갖는다.