

# 대화형 음성인터페이스 시스템을 위한 국어 문말 경계역양의 분류<sup>1)</sup>

## 1. 머리말

최근 들어 지능형 로봇이나 텔레매틱스 등의 응용분야를 위한 대화형 음성인터페이스 시스템의 개발이 활발해지고 있다. 대화형 음성인터페이스 시스템이란 사용자와 시스템 간의 의사소통이 음성에 의한 대화 형태로 이루어지는 시스템이다. 인간인 사용자가 음성으로 시스템에 명령을 하면 이 음성을 시스템이 인식하고 그 의미와 사용자의 의도를 파악하여 대화체의 합성음으로 응답을 해주는 것이다.

대화체 음성은 단순한 정보 전달을 목적으로 하는 낭독체 음성에 비해 억양이나 지속시간 등의 운율적 요소가 다양하게 나타나는데, 이 중에서 억양은 단지 발화의 자연성을 높이기 위해서만이 아니라 화자의 감정이나 태도를 전달하는 의사소통적인 목적으로 사용된다.<sup>2)</sup> 억양이란 문장에 얹히는 높이 곡선을 말하는데, 억양 중에서 문장의 끝 음절에 얹히는 억양 패턴인 문말 경계억양은 한국어의 경우 발화 내용에 대한 화자의 태도를 나타내므로 발화의 의미와도 관련이 있다. 특히 대화에서는 생략문, 비정형문, 한 단어로 된 문장 등 문장의 형식을 갖추지 않은 소형문(minor sentence)이 많이 나타나는데, 이 때 그 발화의 의미 해석을 문말 억양에 의존해야 할 경우가 많다.

인간과 컴퓨터 간의 의사소통을 수행하는 대화시스템에 있어서도 마찬가지로 사용자의 발화의 의미와 그 의도를 해석하기 위해서 문말 억양이 가지는 의사소통적 정보의 추출이 필요하다. 또한 인간인 사용자에게 시스템의 의도를 정확하게 전달하기 위해서는 합성음의 생성에 있어 발화 의도에 맞는 적절한 문말 억양의 구현이 요구된다. 이렇게 문말 억양이 가지는 의사소통적 정보를 대화형 음성인터페이스 기술에 적극적으로 활용하기 위해서는 문말 억양의 의미와 기능을 화용론적 측면에서 분석하고, 다양한 억양 패턴 중에서 의사소통적 정보를 변별하게 하는 요소가 무엇인지를 추출하는 작업이 선행되어야 할 것이다.

본 연구의 목적은 대화체 음성 인식과 음성 합성을 포함하는 대화형 음성인터페이스 시스템을 위한 한국어 문말 경계억양 패턴을 분류하고 그 목록을 작성하는 것이다.<sup>3)</sup> 이제까지 경계억양의 분류에 관한 연구로는 한국어 운율의 주석 체계(annotation system)로 널리 알려진 K-ToBI(Korean Tone and Break Indices)<sup>4)</sup>가 있는데, 여기에서의 분류는 경계억양의 기능이나 의미적인 측면보다는 억양의 형태에 기반해 이루어진 분류이다. 한편, 이호영(1990, 1991, 1996, 1999)<sup>5)</sup>에서는 같은 대상에 대해 핵억양이라는 명칭 하에, 억양의 형태만이 아니라 억양 패턴 간의 기능적인 대립 관계를 고려하여 분류하고 있다. 문말 경계억

1) 이 연구는 한국 정보통신부의 지원을 받아 이루어졌다.

2) Bolinger(1989)의 표현을 빌리면 억양은 “발화시에 어떻게 느끼는가, 또는 발화 내용에 대해 어떻게 느끼는가(how we feel about what we say, or how we feel when we say)”를 나타낸다.

3) 분류된 문말 억양의 목록은 인식이나 합성을 위해 구축된 음성 데이터베이스의 수동 운율 레이블링(prosodic labeling)과 자동 운율 레이블러(labeler)의 훈련, 그리고 운율모델의 생성 등에 적용될 수 있다.

4) Jun(2000) 참조.

5) Lee(1990)와 이호영(1991, 1996, 1999)에서 보이는 핵억양의 분류는 용어만 다소 차이가 있을 뿐 분류 내용 자체는 같으므로 이후의 언급에서는 이호영(1996)으로 지칭한다.

양이 나타내는 의사소통적 정보라는 측면을 대화체의 음성 기술의 개발에 적극 활용하기 위해서, 본 연구에서는 억양의 의미와 기능적 측면을 고려한 이호영(1996)과 같은 접근 방식으로 분류를 시도하고자 한다. 그러면서, 여기서 좀 더 세분화될 필요가 있는 억양 유형을 다시 검토해보고 이에 대한 새로운 분류를 하고자 한다. 본 연구에서는 발화의 의사소통적인 정보에 보다 직접적으로 관여하는 것이 문말 위치에 나타나는 경계억양이라고 보아 일단 분류의 범위를 문말 경계억양으로 국한시키고자 한다.

본 연구에서는 분석을 위해 다양한 대화체 음성이 녹음된 음성 데이터를 사용하였는데, 2장에서 그 구축 과정과 내용에 대해 설명한다. 3장에서는 문말 경계억양의 기능에 대해서 간단히 언급하고, 4장에서는 앞선 연구의 분류 체계들을 비교해 가면서, 대화형 음성인터페이스 기술 개발을 위한 문말 경계억양의 분류를 시도하고자 한다. 마지막으로 5장에서는 분류된 문말 경계억양 목록을 기반으로 태깅된 음성 데이터를 대상으로 문말 경계억양 유형의 분포를 조사해본다.

## 2. 음성 데이터의 구축

문말 경계억양의 분석과 분류, 그리고 그 분포 조사를 위해서 본 연구에서는 두 종류의 음성 데이터를 사용하였다. 이 두 음성 데이터는 대화체 합성시스템을 개발하기 위해 두 차례에 걸쳐서 수집된 음성 데이터베이스의 일부들이다. 이 두 데이터의 차이는 발성목록의 구성에 있는데, 하나는 대화체 코퍼스로부터 추출된 개별 문장으로 이루어진 발성목록을 전문 성우로 하여금 발성하게 하여 녹음한 데이터이고, 다른 하나는 대화 형식을 유지하여 추출된 발성목록을 역시 같은 성우가 발성하여 녹음한 데이터라는 점이다.<sup>6)</sup> 따라서 두 번째 음성 데이터는 성우가 문장의 내용만이 아닌 대화의 맥락까지 파악하여 발성하게 함으로써 보다 자연스러운 대화체 운율을 표현할 수 있도록 만들어진 것이다. 본 연구에서는 다양한 종결어미 분포를 고려하여 구축된 4853 문장의 음성 데이터 I을 기본 분석 데이터로 하고, 대화식으로 구성된 1000문장의 음성 데이터 II를 추가적으로 분석에 사용하기로 한다. 데이터 구축의 상세 내용을 다음에 소개한다.

### 2.1. 음성 데이터 I

음성 데이터 I을 위한 문장 세트의 추출에 사용된 코퍼스는 약 97만 문장의 대화체 텍스트 코퍼스로서 콜센터 상담 대화, 여행자 영역 대화, TV 오락프로그램, 영화 대본 등의 다양한 영역의 대화체 문장을 수집한 것이다.<sup>7)</sup> 텍스트의 작성 방법 상 두 가지의 유형을 포함하는데, TV 오락프로그램을 녹음하여 이를 전사한 전사 텍스트가 있는가 하면, 영화 대본 텍스트와 같이 발화를 상정하여 작성된 텍스트도 있다.<sup>8)</sup> 콜센터 상담 대화 텍스트는 일

6) 분석 데이터로 여러 명의 화자의 음성을 이용하는 것이 바람직하다고 할 수 있겠으나, 화자가 전문 성우이므로 보편적인 표준 한국어의 발음과 억양을 구사하고 있고, 같은 유형, 혹은 같은 어미의 문장이 다른 내용으로 다수 포함되어 있으므로 한 화자의 음성으로도 억양 패턴을 분석하기에 충분하다고 판단하였다.

7) 이 가운데에서 상담대화와 여행자 영역 대화는 ETRI에서 구축한 것이고, TV 오락프로그램 전사 텍스트는 21세기 세종계획에서 구축한 것이다.

8) 구현정·전영옥(2002:21)에서는 구어의 문자화의 종류를 발화의 전사, 발화 상정, 발화 편집으로

반인으로 하여금 상담 주제와 질문 키워드만을 제시하고 상담원과 즉흥적인 대화를 나누도록 하고 이를 녹음한 데이터를 전사자가 수동 전사하는 방법으로 작성된 것이고, 여행자 영역의 대화 문장은 영어 구문집 문장을 자연스러운 한국어의 대화체 문장으로 의역하여 작성된 텍스트이다. 이 모든 영역의 텍스트의 공통점은 전사된 텍스트이건 발화 상정으로 작성된 텍스트이건 모두 한국어의 자연스러운 대화체 문장들이고, 내용면에서는 전문적인 내용이나 정치 대담 등을 포함하지 않은 일상 대화 수준의 발화라는 것이다. 이 코퍼스는 문장 단위로 되어 있고 대화의 순서가 유지되지 않은 코퍼스이다. 또한 화자명이 표기되지 않았다.

발성목록은 종결어미 형태의 분포를 고려하여 추출하였다. 이렇게 추출한 이유는 한국어의 종결어미가 문장 내용에 대한 화자의 태도를 나타내면서, 억양 패턴과도 밀접한 관련성을 보이기 때문이다. 종결어미의 형태에 따라 특정 억양 패턴이 선호되어 쓰이는 어미가 있는가 하면, 모든 억양 패턴이 다 나타나는 어미들이 있다. 따라서 다양한 형태의 억양 패턴을 포함하는 데이터를 구축하기 위해서는 다양한 형태의 종결어미를 포함하는 문장 세트를 추출할 필요가 있었다.

종결어미의 분포를 고려한 문장 세트를 추출하기 위해서는 다음의 작업을 수행하였다. 우선, 대화체 코퍼스를 형태소분석기를 이용하여 자동으로 형태소 태깅하고<sup>9)</sup>, 그 태깅 결과를 바탕으로 문장의 종결어미를 분리하여 코퍼스 전체에 나타나는 종결어미의 리스트를 추출하고, 그 어미 형태의 분포를 고려하여 문장 세트를 추출하였다. 결과적으로 코퍼스에 포함된 모든 종결어미 형태가 적어도 한 번씩 출현하되, 그 출현 빈도에 비례하여 문장 세트에도 포함되도록 하였다. 그 결과, 365개의 종결어미 형태를 포함하는 4853개의 문장이 추출되었다.

이렇게 작성된 발성목록을 전문 여성 성우로 하여금 자연스러운 대화체의 억양으로 발성하도록 하여 녹음하였다.<sup>10)</sup> 녹음은 방음장치가 된 전문 스튜디오에서 이루어졌고 16bit, 16KHz로 저장되었다.

## 2.2. 음성 데이터 II

두 번째 음성 데이터를 위한 문장 세트의 추출에 사용된 코퍼스는 TV나 라디오 드라마 대본으로만 구성된 것으로 약 9만 발화의 대화체 텍스트 코퍼스이다. 이 코퍼스는 첫 번째 코퍼스와는 달리 화자가 명시되고 대화의 순서가 유지된 것이다.

음성 데이터 II는 TV와 라디오 드라마의 대본으로 구성된 약 9만 발화의 대화체 텍스트 코퍼스를 일단 대화에서의 말차례(turn)를 최소 4번 이상 포함하는 대화 단위로 나누고,<sup>11)</sup> 각 대화 단위 내의 발화에 포함된 음소열, 문형, 화행, 종결어미 형태를 구하고, 이러한 요소들이 다양하게 포함되도록 대화 세트를 추출하였다. 다음은 그 발성목록의 구성 대화문의

구분한 바 있다.

- 9) 형태소 분석기는 ETRI의 음성/언어정보연구부에서 자체 개발한 프로그램을 이용하였다. 대화체 텍스트의 형태소 분석 성능은 현재 약 95% 정도이다.
- 10) 코퍼스가 여러 영역의 텍스트로 구성된 것임에도 불구하고, 문장 추출 과정에서 각 영역의 문장들이 뒤섞이고 개별 문장 단위로 녹음이 되었기 때문에 영역에 따른 발화 스타일의 차이는 나타나지 않았다.
- 11) 대화 단위를 나눌 때에는 발성의 자연스러움을 고려하여, 대화 구조상 동의나 인지, 이해 등 선행 발화에 직접 연결되어 나타나는 발화 행위가 대화 단위 첫 발화로 나오지 않도록 하였다.

한 예이다.

<음성 데이터 II의 대화문의 예>

재범 : 누워있어서 움직이기 그런데... 문 좀 열어주실래요?  
여1 : 네.  
: 나가요.  
민이 엄마: 어머, 손님 계셨네?  
: 재범 엄... 재범 엄마, 지금 뭐해요?  
재범 엄마: 나 지금 얼굴 맞사지 받고 있었어요.  
여1 : 말하지 마세요.  
: 주름살이 더 생겨요.  
: 근데 이쪽 어머니!  
: 피부 너무 안 좋으시다.  
: 관리 좀 받으셔야겠어요.  
민이 엄마: 나가면 피부 종단 소리 많이 듣는데...

이렇게 작성된 발성목록은 음성 데이터 I과 같은 전문 여성 성우로 하여금 같은 환경에서 발성하도록 하고, 역시 같은 방식으로 녹음과 저장이 이루어졌다.<sup>12)</sup> 발성시에는 대화의 맥락을 살려 자연스럽게 발성하도록 하였다.

이렇게 녹음된 데이터 가운데 1000문장을 본 연구의 두 번째 분석 자료로 사용하였다. 음성 데이터 I에 비해 데이터 II는 문장 수가 적으며, 종결어미의 유형도 음성 데이터 I만큼 다양하지 못하다. 따라서 음성 데이터 I을 기본 분석 데이터로 삼고 음성 데이터 II는 억양 패턴의 전반적인 분포 결과를 비교하기 위한 데이터로 이용하였다.

### 3. 문말 경계억양의 기능

언어행위에서 하나의 독립적인 의사는 보통 문장이라는 단위를 통해 전달되는데, 문장의 의미는 명제적 의미와 명제에 대한 화자의 태도로 구성된다. 한국어에서 화자의 태도는 보통 종결 어미를 통해 표현되지만 일부는 문말 경계억양으로(글말에서는 문장부호로) 표현된다. 그럼에도 불구하고 종결 어미나 화자의 태도를 표현하는 다른 요소들과 같이 구체적인 의미가 아닌 다소 추상적인 의미를 띠기 때문에 문말 경계억양의 의미를 규정하기는 매우 어렵다.<sup>13)</sup> 그럼에도 불구하고 문말 경계억양은 한국어에서 다양한 패턴으로 나타나면서 문장 차원만이 아니라 담화 차원에서도 기능하는 것을 볼 수 있는데, 다음에 그 기능들에 대해 간단히 언급하기로 한다.

12) 분석 데이터로 여러 명의 화자의 음성을 이용하는 것이 바람직하다고 할 수 있겠으나, 화자가 전문 성우이므로 보편적인 표준 한국어의 발음과 억양을 구사하고 있고, 같은 유형, 혹은 같은 어미의 문장이 다른 내용으로 다수 포함되어 있으므로 한 화자의 음성으로도 억양 패턴을 분석하기에 충분하다고 판단하였다.

13) Hirst(2005:335)에서는 억양의 의미에 관해 여러 방식에서의 다각적인 접근이 있었음에도 불구하고 아직 억양이 어떤 방식으로 발화의 의미에 기여하는지에 대한 의견의 일치가 이루어지지 않고 있음을 지적하였다. 그 밖에 억양과 의미 혹은 억양과 기능에 대한 논의로는 Bänziger and Scherer(2005:253-254), Hirschberg(2002:31-43), Bolinger(1989) 등이 있다.

### 3.1. 문장 차원에서의 기능

문장 차원에서 문말 경계억양은 명제 내용에 대한 화자의 태도를 표현하는 기능을 한다. 종결어미가 나타내는 서법에 호응하여 문장 유형을 나타내는 데에 보조적인 역할을 하기도 하지만, 종결 어미가 서법에 있어 무표적(unmarked)인 경우에는 문말 경계억양이 문장 유형을 결정짓는 기능을 대신하기도 한다. 종결어미 ‘-아/어(요)’로 구성되는 문장이 높은 수평조<sup>14)</sup>(K-ToBI에서의 H%)의 문말 경계억양에 의해 가부관정의문문(yes/no-question)이 되는 것은 문말 경계억양이 서법과 같은 기능을 하는 예이다.<sup>15)</sup>

또한 같은 질문이라도 문말 경계억양의 형태에 따라 다른 태도를 나타낼 수 있다.

- (1) 가: 대청봉이 어디예요?  
나1: 대청봉이요? ↗  
나2: 대청봉이요? ↘

(1)의 예는 <가>의 질문에 대해 <나>가 되묻는 식의 대화인데, <나1>은 문장의 끝음절에서 높이 올라가는 억양 패턴을 사용하고, <나2>는 높이 올라갔다 내려오는 억양 패턴을 사용하고 있음을 나타낸다.<sup>16)</sup> 내용적으로는 같은 문장이나 이 두 발화에 나타나는 화자의 태도는 다르게 들린다. <나1>의 발화는 화자가 ‘대청봉’이라는 정보에 대해 갖는 어떠한 주관적인 전제가 없이 그대로 되묻는 질문인 데 비해, <나2>의 발화에는 ‘대청봉’이라는 정보 내용이 화자가 이미 알고 있는 정보라는 전제적 의미가 함축되어 있음을 느낄 수 있다. 이 같은 패턴의 억양은 문장의 정보 내용을 확인하려는 태도를 나타내며 이러한 의미로 인해, ‘인식’이나 ‘인지’의 양태적 의미(modality)를 가지는 종결어미와 호응하여 쓰이기도 한다.

- (2) 남자 친구 생겼지? ↘  
(3) 너 많이 혼났구나? ↘  
(4) 여기서 며칠 더 묵는다면서? ↘

(2), (3), (4)의 예문에서 ‘-지’<sup>17)</sup>, ‘-구나’, ‘-다면서’ 등의 종결어미가 가부관정 의문문을 구성하는 경우에 <나 1>의 억양보다는 <나 2>에 보이는 억양이 선호되는 것은 확인의 태도를 나타내는 이 억양의 의미가 ‘-지’라는 종결어미가 갖는 ‘이미 앎’이라는 의미, 그리고 ‘-구나’가 가지는 ‘지금 앎’의 의미<sup>18)</sup>, ‘-다면서’가 가지는 ‘전해들음(hearsay)’의 의미와 상호관련성을 보이기 때문이다. 화자가 인식한 사실을 청자에게 확인 질문하는 발화에서는 위의 종결어미에 공통으로 내포된 ‘인식’의 의미에 호응하는 확인의 태도를 나타내는 억양이 사용되는 것이 자연스럽다고 할 수 있다.

14) 여기서의 억양 유형의 명칭은 이호영(1996)을 따른다.

15) 그러나 문장 유형이 반드시 억양 패턴과 일치하지 않는다는 사실에 대해서는 국내외의 연구에서 이미 논의된 바 있다. Bolinger(1989:2), Gussenhoven(2002:47), Verdugo(2005:2), Huang(2001:758), 이호영(1999:34), Jun & Oh(1994:323-326) 참조.

16) 구체적인 억양 패턴의 분류나 기술은 4장에서 하기로 한다.

17) ‘-지’가 가부관정 의문문에 사용될 경우 오름조에 비해 내림조의 억양이 우세하게 쓰인다는 사실은 이미 이숙향(1985:50)에서 논의된 바 있다.

18) ‘-지’와 ‘-구나’의 양태 의미에 대해서는 장경희(1995) 참조.

한편, 하나의 형태로 여러 의미를 가지는 종결어미가 그 양태적 의미를 구별하기 위해 서로 다른 문말 경계억양을 취하는 경우도 있다. ‘-르걸’과 같은 종결어미는 추측을 나타내는 경우에는 (5)의 발화와 같이 상승하는 억양이, 후회를 나타내는 경우에는 (6)에서와 같이 하강하는 억양이 사용된다.

- (5) 나라두 그럴걸? ↗
- (6) 알았으면 차를 갖고 올걸. ↘

이렇게 문말 경계억양은 문장 차원에서 발화 내용에 대한 화자의 태도를 나타내는데, 서법의 종결어미를 대신하여 문장 유형을 결정하는 기능을 하기도 하고, 문장의 양태적 의미와 호응하여 사용되기도 한다. 그리고 다의적인 종결어미의 경우는 문말 경계억양 패턴이 양태 의미를 변별하게 하는 역할을 하기도 한다.

### 3.2. 담화 차원에서의 기능

이밖에 문말 경계억양은 담화 차원에서 특정 화행(speech act)를 특징짓는 데에 보조적인 역할을 담당하기도 한다. 대화 혹은 담화의 마감(closing)에서 많이 나타나는 길게 이어지는 으뜸조의 억양이 그 중의 하나이다.



<그림 1> 대화의 마감에 쓰인 발화의 문말 경계억양의 예.  
“즐거운 하루 되세요.”

담화의 마감(closing) 행위에 쓰이는 발화는 여러 가지 유형이 있는데, “안녕히 가세요.,” “잘 가.”와 같이 헤어질 때 쓰는 관습적인 표현만이 아니라 “감사합니다”, “좋은 시간 되세요.”, “상담원 ~이었습니다.” 등 다양한 내용의 문장이 실제 대화의 마감 행위에 쓰인다. 심지어는 단순히 “네/예~.,” “그래~.” 등과 같이 인지 표시의 단어로 마감 행위를 대신하는 경우도 많다. 이러한 발화에서 공통적으로 나타나는 낮은 으뜸조의 억양은 일반 문장에서의 낮은 으뜸조와는 달리, 억양의 시작 부분이 길게 수평조로 이어지다가 으뜸이 나타난다. 이러한 억양에서 우리는 화자가 대화를 끝내려 하는 의도를 가지고 있음을 파악할 수 있다. 그 밖에도 같은 형태의 ‘네/예’라는 응답어가 그 억양 패턴에 따라 상대 발화에 대한 인지

(acknowledgement)의 표시도 되고, 긍정의 대답이 되기도 하고, 단순한 부름에 대한 응답이 되기도 하는 것은 문말 경계역양이 담화 차원에서 담당하는 기능으로 해석할 수 있다.

### 3.3. 청자에 대한 태도 표시 기능

이 밖에 문말 경계역양은 화자의 듣는 이에 대한 태도를 나타내는 데에도 기능한다. 같은 문장이라도 상대방에 대해 친밀한 태도로 발화할 수 있고, 반대로 무뚝뚝한 태도로 발화할 수 있다. 예컨대, 콜센터의 상담 대화 같은 경우에 상담원의 발화를 들어보면 대개 진술이나 행위 요청의 발화의 경우에도 역양의 끝을 올리는 경우가 많다. 이는 듣는 이 중심의 태도로 발화하고 있음을 표시하기 위한 것으로 해석된다. 끝이 올라가는 문말 경계역양에는 수평이나 내려가는 역양에 비해 청자지향적인(listener-oriented) 태도가 나타난다. 낭독체 합성기와 달리 지능형 로봇이나 텔레매틱스와 같이 사용자에게 대한 친밀감을 표시할 필요가 있는 대화체 음성합성기에서는 문말 경계역양이 갖는 이러한 특성을 이용하여 발화 스타일을 조절할 수 있을 것이다.

## 4. 문말 경계역양의 분류

국어의 역양 패턴의 분류에 대해서는 많은 연구가 이루어져 왔는데, 그 가운데 역양의 전사에 흔히 이용되고 있는 전사 체계(transcription system)로 Jun(2000)의 K-ToBI가 있다.<sup>19)</sup> 여기에서는 역양구 경계 역양(Intonation phrase boundary tone)의 범위 안에서 문말 경계역양을 분류하고 있다. 역양구란 문장과 강세구의 중간 단위로서 역양구의 마지막 음절에 부가되는 역양 패턴이 역양구 경계 역양이다. Jun(2000)에서는 역양구 경계 역양의 유형을 피치 곡선의 형태(shape)에 기반하여 분류하고, L과 H, 그리고 그 조합으로 이루어진 이분 체계적 기호로 표현하고 있다.

반면, 이호영(1996)에서는 문장을 말마디(K-ToBI의 역양구에 해당)와 말토막(K-ToBI의 강세구에 해당)이라는 단위로 나누고, 말마디의 끝 음절에 부가되는 역양을 핵역양(nuclear tone)이라 하였는데, 핵역양의 분류에 있어 형태만이 아니라 음역에서의 높이도 고려하여 L과 H 사이에 M을 설정했다. 음역에서의 절대적인 높이를 분류의 기준으로 두는 것은 문말 경계역양의 기능적인 측면을 고려한 것이다. 따라서 K-ToBI의 경계 역양 중 H%를 H%와 M%으로, HL%은 HL%와 ML%으로, LH%는 LH%와 LM%으로 나누었다.<sup>20)</sup>

19) K-ToBI(Korean Tones and Break Indices) system은 영어의 ToBI system (Pierrehumbert 1980)을 한국어에 적용한 것으로 한국어의 경계강도(break)와 톤(tone)에 대한 주석 체계(annotation system)이다. K-ToBI system에서는 한국어의 문장을 역양구(IP, intonational phrase)와 이를 구성하는 강세구(AP, accentual phrase)로 나누고, 강세구와 역양구에 나타나는 톤 목록을 제시하고 있다.

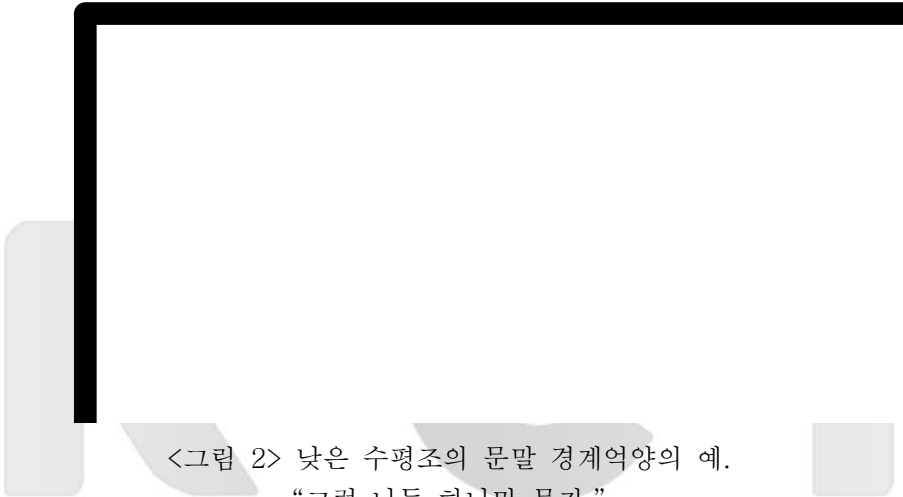
20) 이호영(1999:37)에서는 핵역양과 K-ToBI의 경계역양의 차이를 다음의 표로 제시하였다.

핵역양	K-ToBI의 경계역양	이호영(1999)의 수정안
높은수평조	H%	H%
가운데수평조		M%
낮은수평조	L%	L%

본 연구에서는 이호영(1996)에서 보인 분류 기준을 기반으로 하되 머리말에서 밝혔듯이 분석의 대상을 일단 문말 경계역양으로만 국한하는 것으로 한다.<sup>21)</sup>

그리고 3장에서 언급한 담화적 기능을 하는 특수한 억양 형태들은 이번 논의의 대상에서 제외하기로 한다. 피치 곡선의 관찰과 분석을 위해서는 음성분석 프로그램인 Praat(Boersma & Weenink 2005)을 이용하였다.

우선, K-ToBI에서의 L%은 다양한 음성 데이터를 관찰하다 보면 형태 면에서 차이가 나는 두 가지 유형을 발견할 수 있다. 하나는 K-ToBI의 L%에 대한 정의에서와 같이, ‘한 억양구의 마지막 강세구에서 나타나는 기본주파수(F0)의 최고점에서부터 그 구의 전체에 걸쳐 완만하게 내려가거나 수평이 되는 성조’이고, 다른 하나는 마지막 음절의 기본 주파수가 시작 부분부터 끝부분에 이르기까지 급격히 떨어지는 성조이다.



<그림 2>와 <그림 3>은 각각의 예이다. <그림 2>의 문말 경계역양은 L%에 대한 기존의 설명과 같이 완만하게 내려가는 낮은 수평조인데 비해, <그림 3>의 문말 경계역양은 마지막 음절의 기본주파수가 급격히 떨어지는 억양이며, 음의 세기도 약해진다. 또한 청각적으로 인지되는 화자의 태도에도 차이가 있는데, <그림 2>의 발화에서는 화자의 확신 있는, 혹은 단정적인 태도가 나타나는 데에 비해 <그림 3>의 발화에서는 그러한 태도가 나타나지 않는다. 이러한 문말 경계역양이 나타나는 발화의 유형을 보면, 대부분 화자의 확신적 태도를 나타낼 필요가 없는 독백의 발화라든지, 문장의 형식을 갖추지 않고 접속절이나 명사구

높내림조	HL%	HL%
낮내림조		ML%
온오름조	LH%	LH%
낮오름조		LM%
오르내림조	LHL%	LHL%
내리오름조	HLH%	HLH%
	LHLH%	
	HLHL%	
	LHLHL%	

21) Jun(2002)에서의 경계역양과 이호영(1996)에서의 핵역양은 같은 대상을 지칭하는 명칭으로 본 논문에서는 경계역양이라 통일하여 사용한다.

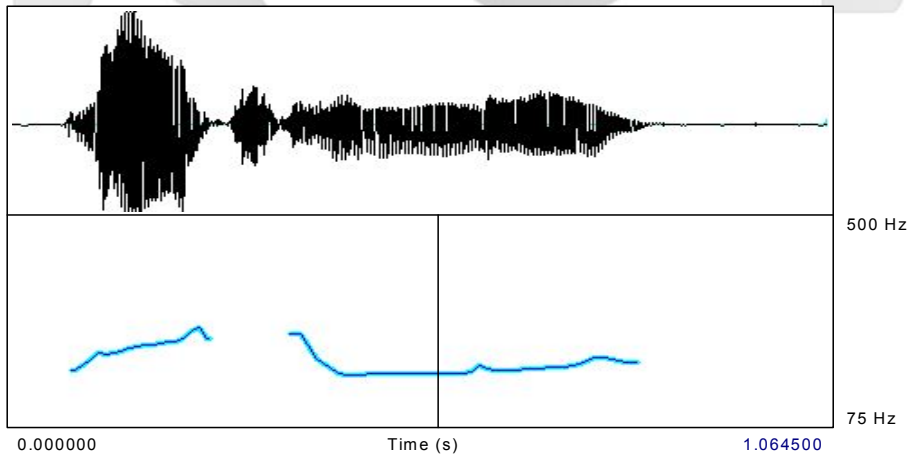


로 끝나는 발화 등이다. 이러한 억양은 화자의 단정적 태도를 나타내는 것으로 인식되는 일반적인 낮은 수평조의 억양과는 달리, 화자의 확신성 여부의 표명에 있어서 ‘무표적 (unmarked)’이라 할 수 있다. 본 연구에서는 이러한 억양을 일반적인 낮은수평조의 억양과 구분하여 문말 경계억양의 한 유형으로 분류하고, 임의의 기호인 LL%로 표기하기로 한다.



<그림 3> 급격히 떨어지는 문말 경계억양의 예.  
“훨씬 더 재미있을텐데.”

다음, K-ToBI에서의 H%는 M%과 H%로 구분한다. 이 구분은 이호영(1996)에서 기술한 바와 같이 가운데수평조와 높은수평조가 기능적으로 대립하는 경우가 있기 때문이다. 그렇게 분류하는 근거는 다음의 <그림 4>와 <그림 5>의 예에서 확인할 수 있다.

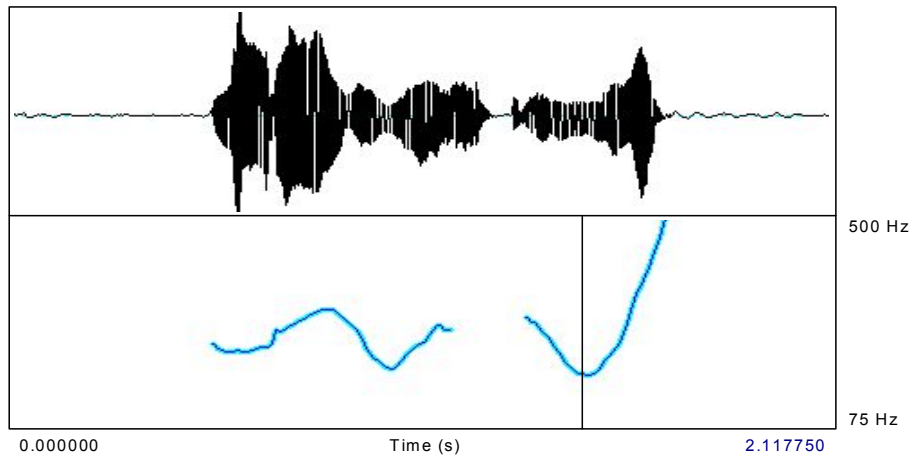


<그림 4> 가운데수평조(M%)의 문말 경계억양의 예.  
“나 잘래.”

<그림 4>의 “나 잘래.”는 마지막 음절의 피치가 앞 음절에 비해 약간 올라간 경우이다. 위의 그림만으로는 기본주파수의 오름이 매우 완만하여 거의 앞 음절과 거의 같은 높이로 이어지는 것처럼 보이나 기본주파수 값은 차이가 있으며,<sup>22)</sup> 청각적으로도 앞 음절보다 기본

22) 앞 음절의 최저점의 기본주파수가 180.8Hz이고, 마지막 음절의 끝지점의 기본주파수가 207.3Hz

주파수가 높아지는 것이 인지된다.



<그림 5> 높은 수평조(H%)의 문말 경계역양의 예.  
“나랑 외식할래?”

이에 비해 <그림 5>의 발화 “외식할래”는 앞 음절에 비해 마지막 음절에서 기본 주파수의 급격한 상승이 보인다. 같은 ‘-르래’라는 종결어미에 얹힌 두 억양은 모두 끝음절에 기본 주파수의 상승이 있으나 그 상승 피치의 최고점의 차이에 의해 하나는 진술이 되고, 다른 하나는 질문이 되는 기능적인 대립관계를 보인다. 즉, 낮게 올라가는 억양은 듣는 이에 대한 응답의 요청의 의미가 안 나타나는 데에 비해, 기본 주파수가 화자의 음역의 끝까지 올라가는 억양은 듣는 이에 대한 응답 요청의 의미가 나타나는 것이다.

‘-아/어’나 ‘-르래’와 같이 둘 이상의 문장 유형에 통용되는 종결어미로 구성된 발화에서는 이렇게 기본주파수의 최고점의 높이의 차이가 문장 유형의 구분에 변별적인 기능을 할 수 있다. 이러한 경우, 기능상으로 볼 때 가운데수평조 M%은 H%보다는 오히려 L%과 같은 부류에 속한다고 할 수 있다.

기본주파수의 음역에서의 최고점의 높이라는 기준은 오름조(LH%)나 내림조(HL%)에서도 마찬가지로 적용된다. 다음은 모두 마지막 음절이 시작되고 난 후에 기본 주파수의 오름이 시작되는 오름조의 예들이다.

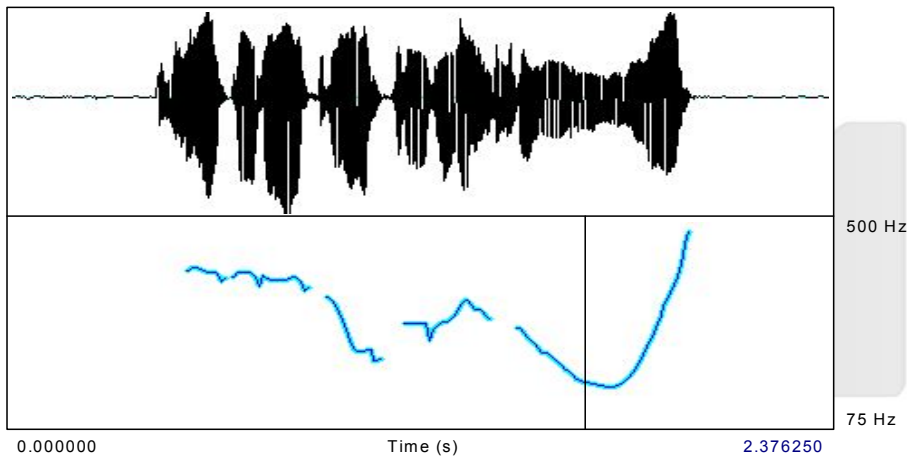
<그림 6>과 <그림 7>에서 기본 주파수의 오름의 최고점의 높이가 다름을 보이는데 <그림 6>의 발화는 상승의 높이가 음역의 중간 이하로 낮은 데에 비해 <그림 7>의 발화에서는 화자의 음역의 최고 높이까지 도달함을 볼 수 있다. 같은 어미인 ‘-네요’가 사용된 발화에서 <그림 6>의 예는 듣는 이의 반응을 구하는 발화가 아니나 <그림 7>의 예는 문장 내용에 대해 화자가 단정적인 태도를 보이면서 동시에 듣는 이의 반응을 구하는 발화가 된다. 음성에서 나타나는 기본 주파수의 상승의 높이의 차이는 텍스트 상에서는 위의 예들처럼 문장부호의 차이로 나타난다.

낮은 올림조의 억양 중에는 간혹 끝부분이 약간 내려가는 패턴이 관찰된다. <그림 8>의 예가 이를 보여준다. <그림 8>의 문말 경계역양 형태는 시각적으로 볼 때 끝부분이 약간 내려감을 볼 수 있는데, 이것이 청각적으로도 느껴지는 경우도 있으나 그 차이가 일반적인 낮오름조와 비교할 때 매우 미미하여 그 구분이 어려운 경우가 많다.<sup>23)</sup> 또한 기능적으로도

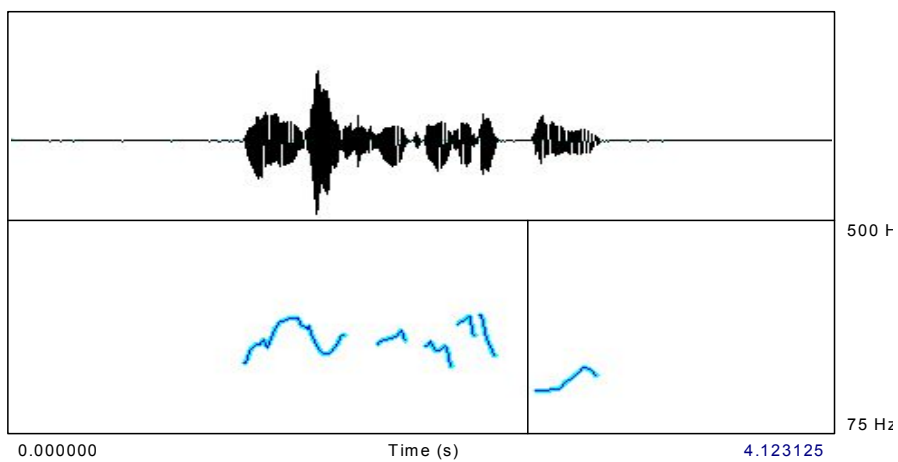
이다.



<그림 6> 낮오름조(LM%)의 문말 경계역양의 예.  
 “끓주렸겠네요.”



<그림 7> 온오름조(LH%)의 문말 경계역양의 예.  
 “카드로 결제를 하셨네요?”



<그림 8> 낮오름조(LM%) 문말 경계역양의 다른 예.  
 “오늘 날씨 되게 좋다.”

별 차이를 보이지 않으므로 이러한 유형의 억양은 낮은 올림조(LM%)에 포함시키기로 한다.

다음은 내림조(K-ToBI의 HL%)의 억양 패턴을 살펴보기로 한다. 이호영(1996)에서는 내림조를 그 최고점의 높이에 따라 낮내림조와 높내림조로 구분하였다. 낮내림조의 예는 <그림 9>의 발화에 나타난다.



<그림 9> 낮내림조(ML%) 문말 경계억양의 예.  
“국세청에서 조사 나온 거지.”



<그림 10> 높내림조(HL%) 문말 경계억양의 예 1.  
“내 연구실로 좀 오지.”

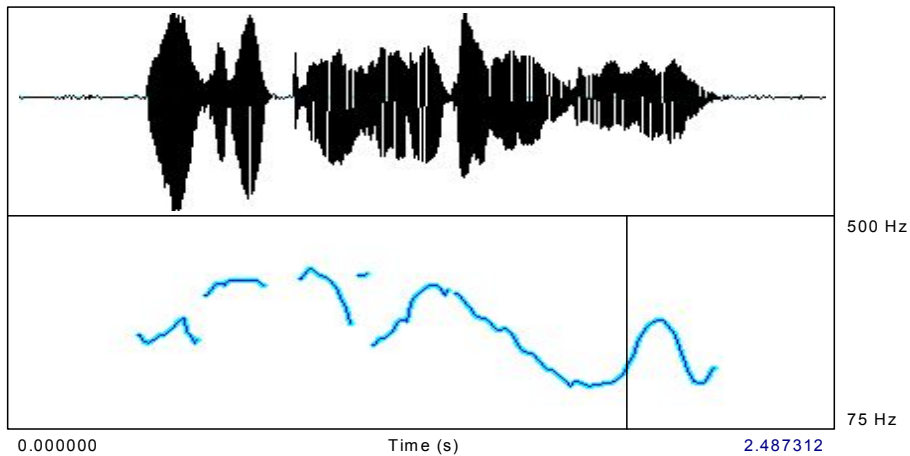
<그림 9>의 문말 경계억양은 상승의 최고점의 기본 주파수 값이 화자의 음역에서 가운데보다 낮다. 여기서 끝부분에 보이는 기본주파수의 상승은 청각적으로는 인지되지 않는 부분으로 무시한다. <그림 9>의 발화에서는 종결어미 ‘-지’와 낮내림조의 문말 경계억양으로 인해 화자에 대해 자신이 이미 알고 있는 바를 다짐하는 태도가 나타난다.

<그림 10> ~ <그림 12>의 발화는 내림조이면서 기본주파수의 상승의 최고치가 화자의 음역의 중간보다 높은 경우의 예들이다.

23) Hirst(2005:337)에서는 억양의 전사에 있어서 듣는 이의 청각적인 판단과 해석이 우선시되어야 함을 강조했다.

<그림 10>에서 기본주파수의 최고치는 화자의 음역에서 가운데보다 위에 위치하며, 낮 내림조의 최고점보다 높다. 그러나 음역의 끝까지 올라가지는 않는다. <그림 9>의 낮내림조에 비해 <그림 10>에서는 화자가 말하려는 내용을 듣는 이에게 강하게 인식시키려는 태도가 느껴진다.

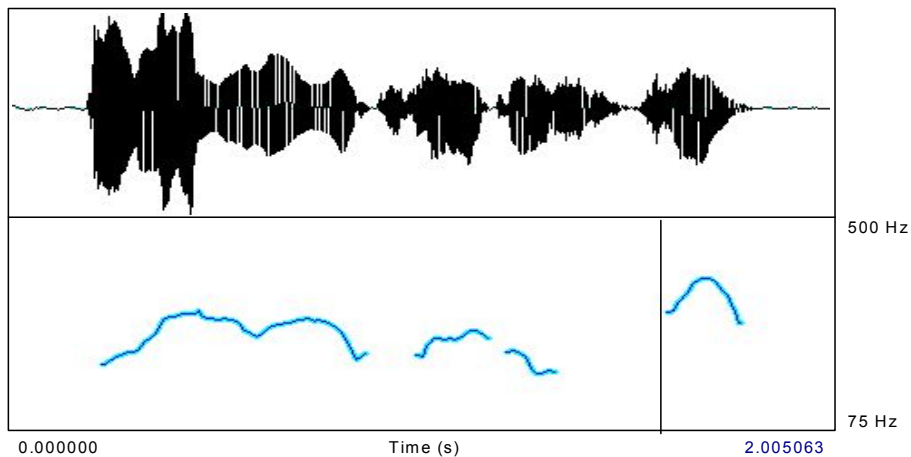
같은 유형으로 다른 발화의 예를 하나 더 보기로 한다.



<그림 11> 높내림조(HL%) 문말 경계역양의 예2.  
“왜 그렇게 오래 걸리는 거야?”

<그림 11>의 발화에서도 문말 경계역양에 의해 화자가 듣는 이에 대해 너무 오래 걸린다는 사실에 항의하는 태도가 느껴진다.

이번에는 기본주파수의 최고치가 화자의 음역 끝까지 올라가는 경우를 보기로 한다.



<그림 12> 높내림조(HL%) 문말 경계역양의 예 3.  
“어머니는 잘 계시지?”

<그림 12>에서는 높내림조이면서 기본주파수의 최고치가 화자의 음역 끝까지 올라가는 데, 이러한 억양 패턴에서는 듣는 이에 대해 응답을 요구하는 태도가 느껴진다. <그림 10>, <그림 11>의 문말 경계역양은 문장 내용을 듣는 이에게 강하게 인식시키려는 태도를 나타

낼 뿐, 대답을 요청하는 태도는 나타나지 않는 데에 비해, <그림 12>의 음역 끝에서 내려오는 억양은 듣는 이의 응답을 요청하는 태도를 나타냄으로써 발화를 질문으로 만드는 데에 결정적인 역할을 한다.

본 연구에서는 이 두 높내림조가 같은 조건에서 이러한 태도적 차이를 보인다는 점에서 이 두 패턴을 구분하고, <그림 12>과 같이 기본주파수의 최고치가 화자의 음역 끝까지 올라가는 문말 경계억양의 패턴을 임의의 기호인 TL%<sup>24)</sup>로 표기한다. 정리해 보면, 위의 세 억양 패턴을 비교해 볼 때, HL%은 듣는이에 대한 응답의 요청의 태도가 나타나지 않는다는 점에서 기능상으로는 TL%보다는 ML%에 더 가까운 것이고, 의미적인 차원에서 ML%과 구분이 되는 억양 패턴이라 하겠다.<sup>25)</sup>

그렇다면 이같은 억양 패턴이 듣는 이에게 응답 요청의 태도를 나타낸다고 해서 과연 내림조가 나타나면서 질문의 기능을 하는 모든 발화가 모두 TL%의 억양 패턴을 가지는가 하는 의문이 제기될 수 있다. 이 문제는 3.1.에서 이미 언급한 바와 같이 억양과 문형의 연관성에 대한 문제로 이어지며, 이들 연구에서의 공통적인 결론은 특정 문형과 특정 억양 사이에는 필연성이 없다는 것이다.

국어의 경우, 텍스트 상에서 서술, 의문, 청유, 명령, 감탄문 등의 문장 서법을 결정짓는 요소는 종결어미이고, 여기에 문장 부호가 보조적인 역할을 한다고 할 수 있다. 반면, 음성을 매개로 하는 구어에서는 억양이라는 운율적 요소가 문장 부호가 가지는 역할을 대신한다고 할 수 있다. 듣는 이의 입장에서 여기에는 문장의 내용과 화용적인 요소인 맥락 정보도 화자의 의도를 해석하는 데에 이용된다. 그런데 종결어미가 무표적인 경우, 듣는 이는 문말 경계억양과 문장의 내용, 그리고 화용적인 요소인 맥락 정보에 의존하여 화자의 의도를 해석하게 된다. 이들 가운데 어느 요소가 화자의 의도 해석에 더 결정적인 역할을 하게 되는가 하는 문제는 차치하고, 문말 경계억양이 기능을 갖는다는 것은 다른 요소들이 문형 결정에 단서로서 기여할 수 없는 상태에서 문말 경계억양은 문형 파악에 결정적인 역할을 할 수 있다는 것이다.

예를 들면 <그림 12>의 “어머니는 잘 계시지”라는 발화가 HL%로 발성되었을 때도 이를 질문으로 해석할 수 있다. 이는 문장의 내용과 듣는 이가 가지고 있는 배경 지식이 발화를 질문으로 해석할 가능성을 높이기 때문이다. 만일 화자와 듣는 이 사이에 대화 맥락에 대한 어떠한 배경 지식이 없는 상태에서, 내용상 청유문인지 의문문인지가 불분명한 “가시지요”라는 문장을 HL%의 억양 패턴으로 발화했을 경우에는 TL%의 억양 패턴으로 발성했을 경우에 비해 질문으로 해석할 가능성이 훨씬 낮아지고 가자는 뜻의 청유문으로 해석할 가능성이 커진다. 이러한 경우 듣는 이가 갈 건지를 확인하는 질문으로 받아들일도록 하기 위해서 화자는 의식적으로 TL%의 억양 패턴을 사용하려 할 것이다.

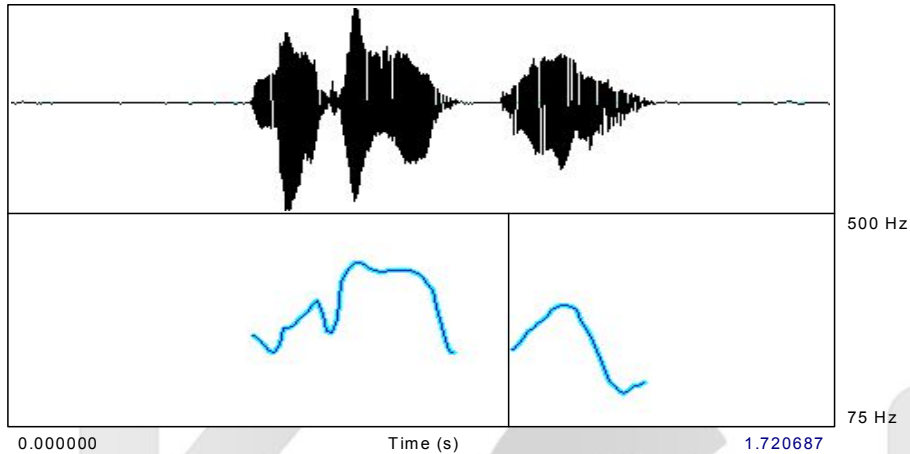
따라서 TL% 경계성조는 문장을 의문문으로 판단하게 하는 문법적, 어휘적 혹은 화용적 요소가 드러나지 않는 상황에서 문장을 질문으로 만드는 잠재적인 기능을 가진다는 점에서 다른 내림조 패턴과 구분될 타당성을 가진다고 할 수 있다.

종결어미 중에는 그 의미 특성상 HL%이 쓰일 수는 있으나 TL%이 쓰이면 부자연스러워지는 어미가 있다. 약속을 나타내는 종결 어미인 ‘-크게’가 쓰인 문장에서 문말 경계억양으로 HL%이 쓰일 경우가 있는데 이 때 기본 주파수의 최고값이 화자의 음역의 가운데 이상

24) TL에서의 T는 Top을 의미한다.

25) 이호영(1996)에서 HL%로 분류한 것은 사실상 본 연구의 HL%가 아닌 TL%을 지칭하는 것이라 할 수 있다.

올라가기는 하나 끝까지 올라가지는 않는다. <그림 13>의 발화에서 ‘-르게’에 나타나는 문말 경계억양은 듣는 이로 하여금 문장 내용을 강하게 인식시키려는 태도를 나타낸다. 앞의 ‘-지’와는 달리 ‘-르게’는 응답을 요구하는 의문문으로는 쓰이지 않기 때문에 문말 경계억양의 기본주파수의 최고값은 화자의 음역 끝까지 올라가지는 않는다. 만일 합성시스템에서 <그림 13>의 문장에 TL%의 억양이 구현된다면 부자연스러운 합성음으로 들릴 것이다. 이렇게 문장의 의미나 종결어미의 유형에 따라 이 두 억양을 구분하여 구현할 필요가 있다는 점에서도 두 억양 패턴의 구분은 필요하다고 하겠다.



<그림 13> 높내림조 문말 경계억양의 예 3.  
“내가 할게.”

본 연구에서는 오르내림조(LHL%)도 그 정점의 높이에 따라 낮은 오르내림조(LML%)와 높은 오르내림조(LHL%)로 구분한다. LML%과 LHL%의 구분은 문형을 결정하는 데에 관계하지는 않으나 화자의 태도를 나타내는 데에 있어 의미적인 차이를 갖는다. LHL%은 LML%에 비해 듣는 이에 대해 자신의 의견을 더 강요하는 태도를 나타내어 강한 감탄이나 따짐의 태도를 나타내는 데에 쓰이는 것을 다음의 <그림 14>와 <그림 15>예에서 볼 수 있다. <그림 14>의 문말 경계억양에 나타난 태도는 화자가 정보를 인지했음을 강조하는 의미인데, <그림 15>에서의 문말 경계억양에서는 화자가 인식한 사실을 다짐함과 동시에 듣는 이에게까지 인식시키려는 태도를 엿볼 수 있다. 이렇게 기존 분류체계에서의 LHL%은 기본주파수의 최고점이 높고 낮음에 따라 청각적으로 강조의 태도가 느껴지고 낮은 경우보다 강조의 태도가 느껴진다는 점에서 그 의미적인 차이를 발견할 수 있다. 대화체 음성과 같이 화자의 감정이나 강조의 태도를 운율로 표현하기 위해서는 이러한 의미적 차이를 구분할 필요가 있다. 실제 음성 데이터를 분석해 보았을 때 비슷한 내용의 문장이라도 감탄부호가 붙은 경우와 그렇지 않은 경우의 억양의 차이가 기본주파수의 최고점의 높낮이에 의해 표현되는 경우가 많았다.<sup>26)</sup> 다양한 감정과 태도가 나타날 수 있는 대화체 합성음의 구현에 있어서 입력텍스트 상의 감탄부호의 유무의 차이를 음성을 표현해야 할 경우를 위해서도 이 두 유형의 억양 패턴을 구분할 필요가 있다고 본다.

26) 음성 데이터 I에서 LML%로 발생된 문장 가운데 15.8%가 감탄부호가 사용된 문장인 데에 비해 LHL%로 발생된 문장 가운데에는 42.3%가 감탄부호가 사용되었음을 확인하였다.



<그림 14> 낮오르내림조(LML%)의 예.  
“그렇군요.”



<그림 15> 높오르내림조(LHL%)의 예.  
“오, 그림 걸어갈 수 있겠군요!”

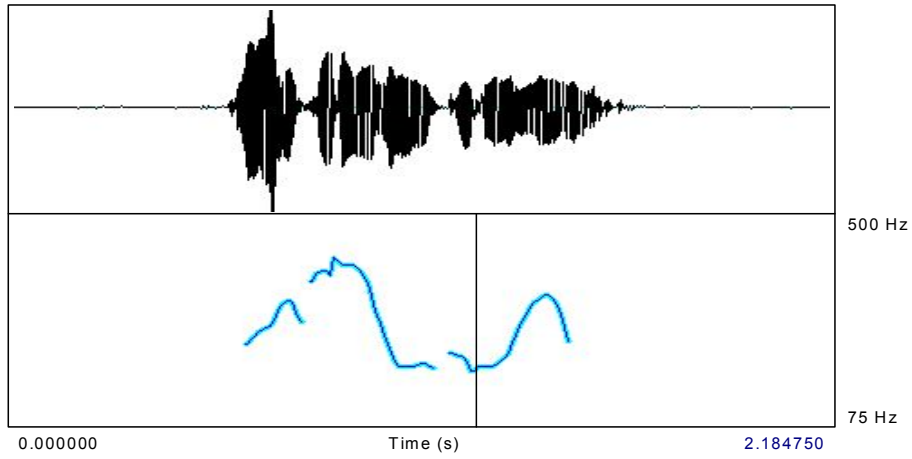
LHL%은 <그림 16>의 발화와 같이 따짐이나 설득의 태도를 강조할 때도 쓰인다. <그림 16>에 보이는 억양 곡선은 <그림 15>의 것과 비교할 때 약간의 차이를 보인다. <그림 15>의 오르내림 곡선은 피치가 음역의 밑까지 완전히 내려갔다가 약간 상승하는 형태인 데 반해, <그림 16>은 피치가 음역의 밑까지 내려가지 않고 중간에 끝나 버린다. <그림 15>의 억양 패턴을 LHLH%로 분류하지 않고 LHL%로 분류하는 이유는 청각적으로 느껴지는 화자의 태도가 <그림 16>의 억양 패턴과 유사하기 때문이다.

내리오름조(HLH%)는 데이터에서 최고점의 차이에서 의미나 기능적 대립을 보이는 예를 발견하지 못해 더 분류하지 않고 하나의 패턴으로 설정한다.<sup>27)</sup> 내리오름조의 예는 <그림 17>에서 볼 수 있다.

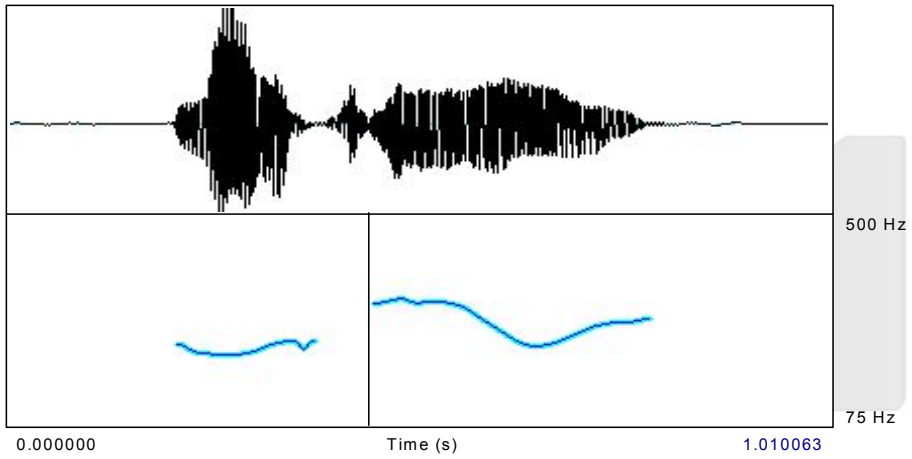
이 밖에 음성 데이터 II에서는 HLHL% 억양 패턴도 나타났다. <그림 18>이 그 예이다.

27) 본 연구에서 제시하는 문말 경계억양의 목록은 추후의 데이터 분석에 의해 새로운 유형이 관찰되면 추가될 수 있음을 밝힌다.

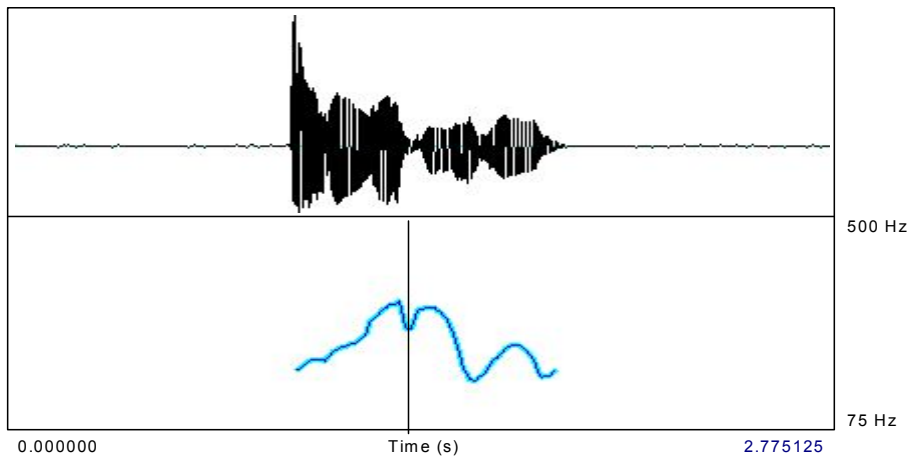




<그림 16> 높오르내림조(LHL%) 문말 경계억양의 다른 예.  
“얼굴이 왜 그래?”



<그림 17> 내리오름조(HLH%) 문말 경계억양의 예.  
“맞아.”



<그림 18> 내리내림조(HLHL%) 문말 경계억양의 예.  
“안 나긴!”

HLHL% 억양 패턴에서도 최고점의 차이에 의한 대립을 이루는 예들은 보이지 않아 하나의 패턴으로 설정한다.

이상의 문말 경계억양의 분류 결과를 기존 연구의 분류와 비교하여 표로 보이던 다음과 같다.

K-ToBI	이호영(1999)	필자의 문말 경계억양
L%	L%	LL%
		L%
H%	M%	M%
	H%	H%
LH%	LM%	LM%
	LH%	LH%
HL%	ML%	ML%
	HL%	HL%
		TL%
LHL%	LHL%	LML%
		LHL%
HLH%	HLH%	HLH%
LHLH%	-	-
HLHL%	-	HLHL%
HLHLH%	-	-

<표 1> 기존 연구의 문말 경계억양 분류와 필자의 분류.

## 5. 문말 경계억양의 분포

위에서 분류된 문말 경계억양 목록을 가지고 대화체 음성 데이터 I(4853문장)과 음성 데이터 II(1000문장)의 문말 경계억양 태깅을 수행하였다. 억양 태깅은 청취에 의한 시각과 음성분석 프로그램인 Praat(Boersma & Weenink 2005)을 이용한 피치 곡선의 확인을 통해서 이루어졌다. 억양 태깅은 훈련된 언어학자 한 사람(필자)이 담당을 하였는데, 이는 여러 사람의 작업에 의한 일관성 결여의 문제를 피하기 위해서였다. 문말 경계억양의 태깅은 분석 프로그램의 시각적인 관찰을 기반으로 하되 청각적인 판단을 중시하였다.

또한 문형별 억양 패턴의 분포를 알아보기 위해 문형 태깅도 수행하였다. 문형은 일반적인 서법의 유형 구분에 따라 평서문, 의문문, 명령문, 청유문, 감탄문으로 구분하되 의문문은 가부 판정 의문문(yes/no-question)과 의문사 의문문(WH-question), 선택 의문문 등으로 세분하여 태깅하였다. 문형 태깅은 거의 수동으로 이루어졌다. 일반적으로는 종결어미의

유형, 또는 의문사 포함 여부에 따라 문형이 구분이 될 수 있으나, 대화체에는 ‘-아/어’와 같이 서법상으로 무표적인 종결어미로 구성된 문장들이 많으므로 문장부호나 문장 내용을 참조해야만 되기 때문이다.

문말 경계역양이 태깅된 두 음성 데이터를 가지고 각 역양 패턴별 빈도를 구해본 결과는 다음과 같다. 우선 <표 2>는 음성 데이터 I에서의 문말 경계역양 패턴의 분포를 보인 것이다.

역양유형	H	HL	L	LH	LHL	LL	LM	LML	M	ML	HLH	TL	계
빈도	729	303	55	341	138	211	1569	557	76	263	2	613	4853
%	14.9	6.2	1.1	7.0	2.8	4.3	32.3	11.5	1.6	5.4	0.0	12.6	100

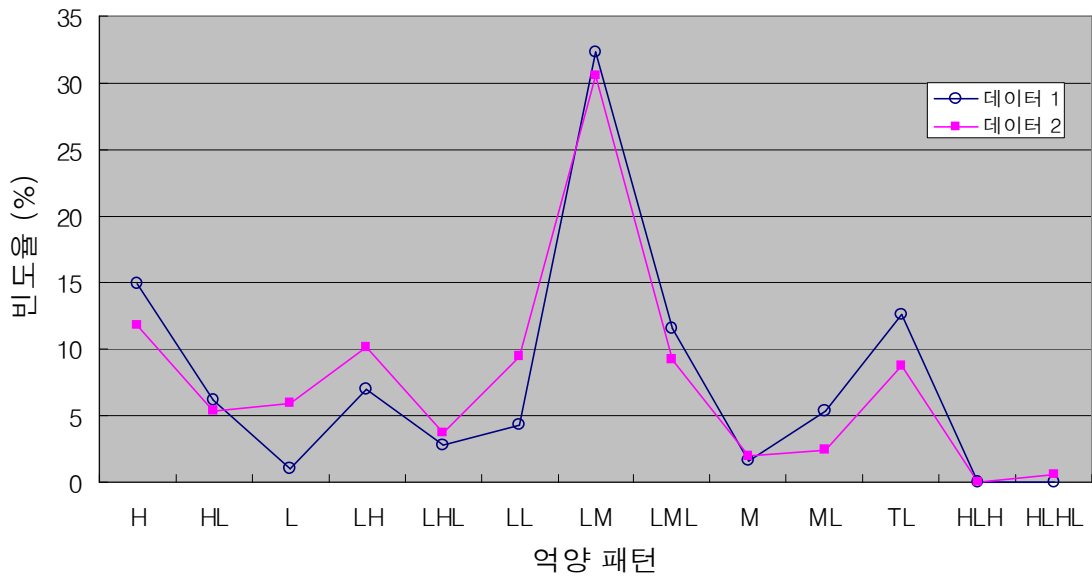
<표 2> 음성 데이터 I에서의 문말 경계역양의 사용 빈도와 빈도율.

한편, 음성 데이터 II에서의 문말 경계역양의 사용 빈도와 빈도율은 <표 3>과 같이 나타났다.

역양유형	H	HL	L	LH	LHL	LL	LM	LML	M	ML	HLHL	TL	계
빈도	118	54	60	102	37	94	306	92	20	24	6	87	1000
%	11.8	5.4	6.0	10.2	3.7	9.4	30.6	9.2	2.0	2.4	0.6	8.7	100

<표 3> 음성 데이터 II에서의 문말 경계역양의 사용 빈도와 빈도율.

<표 2>와 <표 3>에 나타난 문말 경계역양 유형의 사용 빈도율을 비교해 보면 <그림 19>와 같다.



<그림 19> 두 음성 데이터에서의 문말 경계억양의 빈도율 비교

<그림 19>에서 보듯이 두 데이터 상에서 나타난 억양 패턴의 빈도율은 대체로 비슷한 분포를 보이고 있다. 특히, LM%이 두 데이터에서 모두 높은 빈도율을 차지하고 있음을 볼 수 있다.

다음은 문형별로 문말 경계억양의 분포를 조사한 결과이다. 각 문형에 대한 문장수가 균등하지 않기 때문에 문형별로 빈도와 빈도율을 조사하여 각각 표로 보이기로 한다.<sup>28)</sup>

우선 평서문에서의 억양 패턴의 분포를 보면 LM%이 가장 많이 쓰이는 것으로 나타났다.<sup>29)</sup>

억양유형	H	HL	L	LH	LHL	LL	LM	LML	M	ML	HLH	TL
빈도	123	174	42	119	98	173	1311	458	68	248	2	154
%	4.1	5.9	1.4	4.0	3.3	5.8	44.1	15.4	2.3	8.4	0.1	5.2

<표 3> 평서문에서의 문말 경계억양의 빈도.

의문문은 의문사 의문문(WH-question)과 가부판정 의문문(yes/no-question), 선택의문문(alternative question)으로 나누어 빈도를 조사했다. 다음은 그 결과이다.

28) 문형별 억양 패턴의 분포는 보다 다양한 종결어미 형태를 포함하고 있는 음성 데이터 I을 분석한 결과만 보이기로 한다.

29) 낭독체에서는 평서문(declarative)에서 L%이 가장 일반적으로 나타난다고 Jun(2000)에서 언급한 바 있다.

억양유형	H	HL	L	LH	LHL	LL	LM	LML	M	ML	HLH	TL
빈도	150	78	2	130	9	3	11	9	3	0	0	98
%	30.4	15.8	0.4	26.4	1.8	0.6	2.2	1.8	0.6	0	0	19.9

<표 4> 의문사 의문문에서의 문말 경계억양의 빈도.

<표 4>에서 보면, 의문사 의문문에서는 H%가 가장 많은 빈도를 차지하고, 그 다음으로 LH%가 가장 많이 나타났음을 볼 수 있다. 일반적으로 H%는 가부관정 의문문에서 전형적으로 사용되는 것이고, 의문사 의문문에는 HL%이나 LH%, 혹은 낮은 H%가 주로 사용되는 것으로 인식되어 왔다. 이는 이제까지의 의문사 의문문의 연구가 주로 어느 한 어미 형태(주로 ‘-아/어(요)’)로 구성된 문장들만을 대상으로 이루어져 왔기 때문이다.<sup>30)</sup> 실제로 발성된 의문사 의문문의 음성 데이터를 살펴 보면, 억양 패턴의 사용 분포는 종결어미의 유형에 따라서 차이가 남을 볼 수 있다.

본 연구에서는 따라서 의문사 의문문을 구성하는 주요 어미별 억양 유형의 빈도를 조사해 보기로 하였다. 그 결과는 <표 6>과 같았다.

<표 6><sup>31)</sup>을 보면, 모든 어미들이 두 가지 이상의 억양 패턴과 같이 쓰임을 볼 수 있는 동시에, HL%이나 TL%이 선호되는 어미 부류가 있는가 하면 H%나 LH%가 선호되어 사용되는 어미들이 있음을 알 수 있다.

어미	억양유형										
	HL	TL	H	LH	LML	LHL	L	M	LL	LM	
-버니까/습니까	21	9	2	2	0	0	0	0	0	3	
-아/어(요)	14	7	21	32	1	0	0	0	0	0	
-지(요)	0	0	56	4	0	0	0	0	0	0	
-ㄴ대/대/래/내/재(요)	2	0	8	11	0	0	0	0	0	0	
-다고/라고/자고/냐고(요)	0	0	10	0	3	4	0	0	1	0	
-는데/(으)ㄴ데(요)	0	3	23	10	0	2	0	0	0	0	
-니	0	23	1	2	0	0	0	0	0	0	
-고/구(요)	0	1	2	5	1	0	0	0	0	0	
-게(요)	0	0	1	3	0	1	0	0	0	0	
-(으)려고(요)	0	1	4	10	0	1	0	0	0	0	
-ㄴ지/는지/(으)ㄴ지(요)	4	0	1	10	0	0	2	2	0	3	
-(으)ㄴ까(요)	12	12	2	10	0	0	0	0	0	0	
-르래/을래(요)	1	3	1	0	0	0	0	0	0	0	
-(으)ㄴ가/나요	15	4	4	5	0	0	0	0	0	0	
-(으)ㄴ가/나	3	5	0	7	0	0	0	1	1	0	

30) 의문사 의문문의 운율 관련 연구로는 Jun and Oh(1994), 김미란 외(2000) 등이 있다.

31) <표 6>에서 각 어미 유형에 가장 높은 빈도는 어두운 배경색으로 표시하였다. 어미 유형은 이 형태와 ‘-요’가 붙은 준대형을 포함하는데, 단 ‘-(으)ㄴ가요/나요’와 ‘-(으)ㄴ가/나’는 구분하였다. 이는 ‘-(으)ㄴ가/나’가 독백 의문문을 구성하므로 ‘-요’가 붙은 의문문과는 다른 태도를 나타내기 때문이다.

-더라	0	0	1	3	0	0	0	0	0	0	0
-냐니까	1	0	1	3	3	1	0	0	0	0	

<표 6> 음성 데이터 I에서의 어미 유형별 억양 패턴의 사용 빈도.

다음은 가부관정 의문문과 억양 패턴간의 분포를 조사한 결과이다.

억양유형	H	HL	L	LH	LHL	LL	LM	LML	M	ML	HLH	TL
빈도	398	25	0	67	4	2	17	14	0	2	0	345
%	45.5	2.9	0	7.7	0.5	0.2	1.9	1.6	0	0.2	0	39.5

<표 5> 가부관정 의문문에서의 문말 경계억양의 빈도.

가부관정 의문문에서는 H%과 TL%이 압도적으로 많이 사용되었음을 볼 수 있다.

또한 특정 종결어미 형태와 특정 억양 패턴이 관련을 보이는 현상도 관찰할 수 있는데, 앞에서도 언급했듯이 ‘-지(요)/-죠’, ‘-다면서요’, ‘-다며’, ‘-다고(요)’, ‘-ㄴ대(요)’와 같은 종결어미들이 확인을 요청하는 뜻의 의문문에서 TL%이 고빈도로 사용되었다.

선택 의문문은 모두 4문장의 예 중에 LH%가 3번, LM%이 1번 나타났다.

<표 6>, <표 7>은 명령문과 청유문에서의 억양 패턴의 분포를 보인 것이다. 명령문과 청유문에서도 서술문과 마찬가지로 LM%이 가장 많은 빈도로 나타났음을 볼 수 있다.

억양유형	H	HL	L	LH	LHL	LL	LM	LML	M	ML	HLH	TL
빈도	19	17	9	9	20	26	149	50	5	7	0	3
%	6.1	5.4	2.9	2.9	6.4	8.3	47.5	15.9	1.6	2.2	0	1.0

<표 6> 명령문에서의 문말 경계억양의 분포.

억양유형	H	HL	L	LH	LHL	LL	LM	LML	M	ML	HLH	TL
빈도	2	6	1	2	3	4	57	19	0	6	0	0
%	2.0	6.0	1.0	2.0	3.0	4.0	57.0	19.0	0	6.0	0	0

<표 7> 청유문에서의 문말 경계억양의 분포.

감탄문에서의 억양 패턴의 빈도 분포는 다음의 <표 8>에 보인다.

억양유형	H	HL	L	LH	LHL	LL	LM	LML	M	ML	HLH	TL
빈도	0	3	1	0	4	3	23	7	0	0	0	10
%	0	5.9	2.0	0	7.8	5.9	45.1	13.7	0	0	0	19.6

<표 8> 감탄문에서의 문말 경계억양의 분포.

감탄문에서도 LM%이 가장 많은 빈도를 나타냈고, 그 다음으로 TL%, LML% 순으로 많이 쓰였음을 볼 수 있다.

이상의 분석 결과를 보면 서술문과 청유, 명령, 감탄문 등에서는 모두 LM%이 가장 높은 빈도로 쓰였음을 알 수 있다. 특기할 것은 낭독체 음성에서 서술문의 전형적인 억양 패턴으로 인식되던 L%은 극히 낮은 빈도를 보이고 그 대신 LM%이 가장 높은 빈도를 보였다는 것이다. 이는 성우의 발성상의 특성이라고도 할 수 있으나, 단정적이고 무뚝뚝한 느낌의 L%보다 듣는 이에 대해 친밀한 느낌을 주는 LM%을 주로 사용한 전반적인 발화 태도를 반영하는 것으로 해석할 수도 있을 것이다.

## 6. 맺음말

이상으로 필자는 대화형 음성인터페이스 기술 개발에 있어 음성인식과 음성합성의 운용 처리에 기반이 되는 문말 경계억양의 분류에 대해 논의하였다. 분류에 앞서 문말 경계억양이 가지는 기능을 문장 차원과, 담화 차원, 그리고 청자에 대한 발화 태도라는 관점에서 논의하였다. 문장 차원에서 문말 경계억양은 명제 내용에 대한 화자의 태도를 표시하는 기능을 한다. 따라서 국어에서 이러한 기능을 담당하는 종결어미의 양태적 의미와 가장 밀접한 관련성을 보인다. 문말 경계억양은 문장 유형을 표현하는 종결어미에 호응하여 사용되기도 하고, 종결어미가 서법적으로 무표적인 경우, 혹은 문장 내용이나 발화의 맥락 정보만으로 문형이 변별되지 않을 경우 문말 경계억양은 화자의 의도를 표현하거나 해석하는 데에 결정적인 기능을 하기도 한다. 따라서 문말 경계억양은 화자의 태도를 나타내는 데에 독립적으로 기능을 하는 것이 아니라 형태, 운율, 화용상의 다른 요소들과 유기적으로 관련을 맺으며, 문장 유형의 결정에 잠재적인 기능을 가진다고 할 수 있다.

또한 문말 경계억양은 담화 차원에서 특정한 화행을 표시하는 것을 보조하는 기능을 한다. 대화의 마감(closing)에서 전형적으로 나타나는 억양은 화자가 자신의 발언권을 넘겨주면서 대화를 마감하려 하는 태도를 나타내어 대화를 자연스럽게 끝맺는 데에 기여한다. 그 밖의 문말 경계억양의 기능으로 화자의 듣는 이에 대한 발화 태도를 표시하는 기능이 있음을 논의했다. 서술이나 행동 요청 등의 듣는 이의 응답이나 의견을 필요로 하지 않는 유형의 발화에서 나타나는 상승 억양은 화자가 청자에 대해 청자지향적인 태도로 발화하고 있음을 보임으로써 듣는 이에 대해 친밀한 느낌을 주도록 하려는 의도의 표시라고 할 수 있다. 문말 경계억양이 갖는 이러한 다양한 측면에서의 기능과 효과를 대화형 음성인터페이스 시스템의 음성 인식 과 합성 기술에 이용한다면 사용자와 시스템 간에 보다 자연스러운 대화를 수행하도록 하는 효과를 거둘 수 있을 것이다.

본 논문에서는 대화체 음성을 대상으로 하는 음성인식이나 음성합성에서 대화체 음성의 이러한 특성을 적극적으로 이용하기 위한 기초 작업으로 의미와 기능에 기반한 문말 경계억양의 분류를 시도하였다. 즉, 두 가지의 억양 패턴이 문형이나 화행을 결정하는 데에 기능적으로 대립 관계를 보일 수 있으면 이 둘을 분류하는 것으로 하고, 기능만이 아니라 의미상으로도 청각적으로 구분되는 차이를 보이면 역시 이를 분류하는 것으로 하였다. 억양구 경계 억양을 형태에만 기반하여 기술하던 Jun(2000)의 분류와 달리, 기능적인 기준을 도입하여 이에 수정안을 제시한 이호영(1996)과 같은 접근 방식으로 분류를 시도하면서 이 분류에서 세분화가 필요하다고 여겨지는 억양 패턴들을 다시 세분화하여 분류하였다.

Jun(2000)과 이호영(1996)이 공통으로 설정한 L%의 억양 패턴을 본고에서는 둘로 나누

어 완만하게 하강하거나 수평인 전형적인 L%과 마지막 음절의 시작에서 끝까지 급경사를 이루며 떨어지는 톤으로 구분하고, 화자의 단정적인 태도가 느껴지는 일반적인 L%과 달리 화자의 단정적인 태도가 나타나지 않는 이 억양 패턴에 대해 임의의 기호 LL%을 부여했다.

Jun(2000)의 H%나 LH%는 이호영(1996)에서의 분류대로 화자의 음역에서의 기본주파수의 최고점이 도달하는 높이에 따라 M%과 H%, LM%과 LH%으로 분류하였다. 그러나 HL%은 억양 곡선의 최고점이 화자의 음역에서 어떤 높이까지 미치느냐에 따라 다시 청자에 대한 태도가 달라짐을 관찰하였다. 기본주파수의 최고점이 화자의 음역의 꼭대기까지 올라가는 경우는 듣는 이에게 응답을 요구하는 태도가 나타나고 그렇지 않은 경우는 단지 발화 내용을 강하게 인식시키려는 태도만 보일 뿐 응답의 요구가 나타나지 않는다는 사실을 근거로 HL%은 다시 HL%와 TL%로 분류하였다. LHL%도 기본주파수의 최고점이 높고 낮음에 따라 높은 쪽이 낮은 쪽보다 화자의 확신을 듣는이에게 강하는 인식시키려는 태도가 느껴진다는 점에서 LML%과 LHL%로 분류하였다. 실제로 발성목록과 화자가 발성한 음성을 비교해 보았을 때도 텍스트상에서 느낌표가 있고 없고의 차이를 성우가 기본주파수의 최고점의 높고 낮음으로 구별하여 표현한다는 사실을 확인하고, 합성음에서 이러한 텍스트상의 차이를 음성적으로 구현하기 위해서도 LML%과 LHL%로 분류할 필요가 있음을 논하였다. 음성 데이터에서 발견된 HLH%와 HLHL%는 그러한 대립 관계를 발견하지 못해 더 구분하지 않았다.

이상에서 나온 문말 경계억양의 목록을 가지고 대화체 음성 데이터를 태깅하여 문말 경계억양의 분포를 조사하고, 다시 문형별로 억양의 빈도를 조사하였다. 그 결과, 음성 데이터 I 과 II 모두에서 LM%이 가장 높은 빈도로 나타났다. 억양 패턴의 분포는 어미의 형태에 따라서도 달리 나타나는데, 특히 의문문의 경우에 그 차이가 크게 남을 알 수 있었다.

서술과 명령, 청유, 감탄문에서는 전반적으로 LM%이 가장 높은 빈도로 사용된 것으로 나타났다. 일반적으로 서술문에 전형적으로 사용되는 것으로 여겨지는 L%보다 LM%이 높은 빈도를 보이는 것은 단정적인 느낌의 L%보다는 듣는 이에 대해 친밀한 느낌을 주는 LM%으로 발성한 화자의 발화 태도를 반영하는 것으로 해석할 수 있다.

본 연구에서 미처 다루지 못한 부분은 일반 국어 화자들이 문말 경계억양을 문장에서의 기능, 의미와 관련해 어떻게 지각하고 구분하는지 하는 문제를 청취실험을 통해 다루지 못했다는 점이다. 이에 대해서는 후고로 미루기로 한다. 또한 본 연구에서는 문장 차원에서 기능하는 문말 경계억양 패턴을 분류의 대상으로 삼았는데, 차후의 연구에서는 담화 차원에서 기능하는 문말 경계억양에 대한 논의가 이어져야 하리라고 보며, 앞으로 의미, 화용론적 입장에서의 국어의 억양 연구가 보다 활성화되어야 할 것으로 본다.



## 참고 문헌

- 구현정 · 전영옥(2002), 「구어와 구어 전사 말뭉치」, 『한국어 구어 연구(1)』, 말뭉치 기반 국어 연구 총서9, 서울: 한국문화사, 9-71.
- 김미란, 신동현, 최재웅, 김기호(2000), 「초점과 관련된 의문문 억양 패턴 실험」, 『음성과학』 제7권 제4호, 203-217.
- 이숙향(1985), 「한국어 문미억양에 관한 연구」, 『말소리』 제9-10호, 대한음성학회, 28-90.
- 이숙향, 김종진(2005), 「대용량 데이터베이스를 이용한 한국어 운율 특성에 관한 연구」, 한국음향학회지 제24권 2호, 한국음향학회, 117-126.
- 이호영(1991), 「한국어의 억양체계」, 『언어학』 제13호, 한국언어학회, 25-39.
- \_\_\_\_\_(1996), 『국어 음성학』, 서울: 태학사.
- \_\_\_\_\_(1999), 「국어 핵억양의 음향음성학적 연구」, 『말소리』 38, 대한음성학회, 25-39.
- 장경희(1985), 『현대 국어의 양태범주 연구』, 서울: 탑출판사.
- Bänziger T. and Scherer K. R.(2005), “The role of intonation in emotional expressions”, *Speech Communication* 46, 253-254,
- Boersma, P. & Weenink, D.(2005), Praat: doing phonetics by computer (Version 4.3.19) [Computer program]. Retrieved July 20, 2005, from <http://www.praat.org/>.
- Bolinger D.(1989), *Intonation and Its Uses: Melody in Grammar and Discourse*, Stanford: Stanford University Press.
- Gussenhoven, C.(2002), "Intonation and Interpretation: Phonetics and Phonology", *Proceedings of Speech Prosody 2002*, Aix-en-Provence, 47-57.
- Hirst D. J.(2005), “Form and function in the representation of speech prosody”, *Speech Communication* 46, 334-347.
- Hirschberg J.(2002), “Communication and prosody: Functional aspects of prosody”, *Speech Communication* 36, 31-43,
- Huang X., Acero A., Hon H.-W.(2001), *Spoken Language Processing*, Prentice Hall PTR, New Jersey.
- Jun, S. A.(2000), “K-ToBI (Korean ToBI) Labelling Conventions, ver. 3.1”, <http://www.linguistics.ucla.edu/people/jun/ktobi/K-tobi.html>, 2000.
- Jun, S. A. and Oh, M.(1994), “A prosodic analysis of three sentence types with “WH” words in Korean”, *Proc. ICSLP'94*, Yokohama, Japan, 323-326.
- Lee, H. Y.(1990), *The Structure of Korean Prosody*, Ph.D. thesis, University College London, Published by Hanshin Publishing Co.
- Pierrehumbert, J.(1980), *The Phonology and Phonetics of English Intonation*, Ph.D. thesis, MIT.
- Verdugo, D. R.(2005), "The nature and patterning of native and non-native intonation in the expression of certainty and uncertainty: Pragmatic effects", *Journal of Pragmatics* vol. 37, Issue 12, 2086-2115.

Keyword: 대화형 음성인터페이스 시스템(dialog system), 문말 경계억양(sentence-final boundary tone), 대화체 음성(conversational-style speech)

#### ABSTRACT

Classification of sentence-final boundary tones for the development of Korean dialog systems

Oh, Seung-Shin

The purpose of this research is to provide a classification of sentence-final boundary tones in conversational-style speech for the development of Korean dialog systems. While IP boundary tones in the K-ToBI system are classified based on their shapes, our classification is based on their meaning and function as proposed by Lee(1990, 1996). We propose a more refined classification based on the analysis of speech data containing a wide range of sentence ending types. We then investigate the global distribution of boundary tone types in the speech data as well as their distribution per sentence type.

