

한국어 원거리 음성의 운율적 특성

Prosodic Characteristics of Korean Distant Speech

이 숙 향*, 김 선 희**, 김 중 진***

(Sook-hyang Lee*, Sunhee Kim**, Jong-Jin Kim***)

*원광대학교 영중어문학부, **서울대학교 인문정보연구소, ***한국전자통신연구원

(접수일자: 2006년 1월 20일; 채택일자: 2006년 2월 16일)

본 논문의 목적은 한국어 원거리 음성의 운율적 특성을 규명하는 것으로, 36개의 2음절어를 4명의 화자 (여성 화자 2명, 남성 화자 2명)가 원거리 환경과 일반 환경에서 발화한 총 288개의 2음절어를 분석대상으로 하였다. 실험 결과 지속시간과 에너지의 경우는 일반 음성에 비하여 원거리 음성의 첫음절에 대한 둘째음절의 비율이 유의미하게 큰 것으로 나타났다. F0 대역폭의 경우에도 원거리 음성에서의 대역폭이 평이 음성에 비해 큰 값을 보였다. 억양 패턴에 있어서는 원거리 음성의 경우에 둘째음절에 'HL%'의 복합 경계성조가 실현되거나 첫음절에 'L +H' 성조가 실현되기도 하였으며 이 두 가지가 한 단어에 모두 실현되는 경우도 있었다.

핵심용어: 운율적 특성, 원거리 음성, F0 대역폭, 에너지, 억양 패턴

투고분야: 말소리 생성 및 인지 분야 (12.1)

The aim of this paper is to investigate the prosodic characteristics of Korean distant speech. Four speakers (2 males and 2 females) produced 36 2-syllable words in both distant-talking and normal environments, totaling 288 spoken 2-syllable words. The results showed that ratios of second syllable to first syllable in vowel duration and vowel energy were significantly larger in the distant-talking environment compared to the normal environment and f0 range also bigger in the distant-talking environment. In addition, 'HL%' contour boundary tone in the second syllable and/or 'L +H' contour tone in the first syllable were used in the distant-talking environment.

Keywords: Prosodic Characteristics, Distant Speech, F0 Range, Energy, Intonation Pattern

ASK subject classification: Speech Production and Perception (12.1)

I. 서론

본 연구에서 원거리 음성이란 일정한 거리에 있는 음성 시스템을 조작하기 위하여 화자가 음성을 보다 명료하게 전달하고자 발화한 음성을 의미한다. 본 연구는 이와 같이 원거리 음성에서 관찰되는 한국어의 분절음의 운율적 변화 중 운율적 특징을 지속시간, 에너지, F0와 억양 패턴을 중심으로 평이한 음성과 비교하여 살펴보고자 한다. 이러한 연구는 원거리 음성의 보상 문제를 위한 기초연구로서, 궁극적으로 음성인식 시스템 성능 개선에 기여하기를 기대한다.

상대방에게 명료하게 들리도록 하기 위하여 목소리를 높

이는 경우에 발생하는 효과는 여러 환경에서 관찰될 수 있는데, 잡음환경에서 의사소통을 시도하는 경우 [1-2-3] 나 청각 장애인과 대화하는 경우[4-5], 거리가 떨어진 곳에 있는 사람과 대화하는 경우[6], 잘 알아듣지 못하는 외국인에게 말하듯이 말하는 경우[7], 일정한 거리에 있는 PC를 음성으로 조작하기 위한 경우[8], 또는 언어적 강세 혹은 포커스를 실현하기 위하여 명료하게 발음하는 [9-10] 등 다양하다.

기존의 연구들에 의하면 각각의 경우에 그 대상 언어 및 언어 수집 환경이나 수집 방법의 상이성으로 인하여 엄밀한 의미에서의 비교는 용이하지는 않지만, 대체로 (i) F0와 에너지 및 모음 지속시간의 증가, (ii) 에너지 분포의 이동, (iii) F1의 증가 등이 공통적인 음향학적 특징으로 보고 되었다.

그러나 이러한 특징은 분절음 차원에서 그치고 운율적

특성에 관한 연구는 본격적으로 이루어지지 않았으며 한국어의 원거리 음성에 관한 연구 또한 분절음의 음향적 특성을 다룬 연구는 소수 있지만[11-12] 운율적 특성을 다룬 연구는 없다.

II. 분석 대상 및 분석 방법

본 연구에서 사용하는 음성데이터는 PC 명령어를 녹음한 것으로서 이전 연구[1-2]에서 이미 사용했던 데이터 가운데, 남, 녀 각 2명의 데이터를 이용하였으며 1-4음절어 50개 단어 중 36개의 2음절어만을 선택하여 총 288개 음성파일 (36개 단어 x 4 화자 x 2 환경)을 분석 대상으로 하였다. 피험자는 모두 서울 경기지역에서 성장하였으며 남녀 각 2명으로, 여성 화자는 30대, 남성 화자는 20대와 40대였다. 음향분석은 Praat 4.2.21을 이용하여 지속시간과 에너지, F0 및 억양패턴에 대한 분석이 이루어졌다. 지속시간과 에너지 분석은 각 음절의 모음을 대상으로 자동측정 프로그램을 이용하여 측정이 이루어졌으며, F0 분석은 수동으로 최고점과 최저점을 추출하였다. 억양 패턴 분석은 K-ToBI 분석틀을 이용하였다. 측정결과에 대한 통계 분석은 SPSS 10.0을 이용하였으며 유의수준은 $P < 0.05$ 이다.

III. 분석 결과 및 논의

3.1. 지속시간

음절 간 모음 지속시간을 비교한 결과 표 1에서 볼 수 있듯이 모음 지속시간은 평이 음성과 원거리 음성 모두에서 첫음절에 비해 둘째음절에서 길며 M48의 평이 음성을 제외하고 모두 음절 간 유의한 차이를 보이고 있다. 발화상황에 따라 음절 간 모음 지속시간 비율이 달라지는지 보기 위하여 둘째음절 모음의 지속시간 대 첫음절 모음의 지속시간의 백분율을 구하였다 ((둘째음절/첫음절)*100). 표 2에서 볼 수 있는 바와 같이 음절 간 모음의 지속시간 비율은 대체적으로 평이 음성에 비해 원거리 음성에서 큰 값을 보이고 있다. 발화상황 간 차이를 화자별로 보면 화자 F26와 M48만이 유의한 차이를 보이고 있으며 이는 뒤에서 논의될 발화상황 간 경계성조 억양 패턴 차이에 기인하는 것 같다. 즉, 이 두 화자는 다른 두 화자와는 달리 대부분의 단어에서 둘째음절의 경

표 1. 음절별 발화상황별 모음 지속시간 평균, 표준편차 및 음절 간 차에 대한 Paired Samples t-test 결과

Table 1. Results of paired samples t-test for vowel duration differences between syllable positions in a word.

화자	발화상황	평균(표준편차)		샘플수	t-test 유의도
		첫음절 (ms.)	둘째음절 (ms.)		
전체	원거리	120.36 (68.70)	222.15 (87.44)	139	.000
	평이	112.05 (61.10)	153.48 (69.52)	139	.000
M1	원거리	97.24 (43.19)	171.67 (68.42)	34	.000
	평이	73.38 (32.76)	122.44 (52.23)	34	.000
F26	원거리	102.71 (54.16)	244.09 (103.33)	35	.000
	평이	130.54 (55.95)	176.17 (71.05)	35	.007
M48	원거리	105.43 (57.61)	215.80 (69.41)	35	.000
	평이	114.11 (67.35)	139.69 (60.37)	35	.100
F79	원거리	175.40 (83.01)	255.60 (82.20)	35	.000
	평이	129.06 (65.52)	174.74 (78.24)	35	.018

표 2. 발화상황별 둘째음절 대 첫음절의 모음 지속시간 비율 평균 및 표준편차와 발화상황 간 차이에 대한 Paired Samples t-test 결과

Table 2. Results of paired samples t-test for vowel duration ratio differences between speaking conditions.

화자	평균(표준편차)		샘플수	t-test 유의도
	원거리(%)	평이(%)		
전체	253.48 (189.94)	181.33 (128.49)	139	.000
M1	220.51 (140.94)	209.30 (143.71)	34	.478
F26	313.9562 (204.84)	162.02 (99.34)	35	.000
M48	284.92 (236.30)	170.73 (127.05)	35	.000
F79	193.60 (140.69)	184.06 (140.24)	35	.677

계성조를 평이 음성에서는 단순 경계성조 (simple tone) 인 'L%'에서 원거리 음성에서 복합성조 (contour tone) 중 하나인 'HL%'로 변화시킴으로써 원거리 음성의 둘째음절이 평이음성의 둘째음절에 비해 상당히 길어지는 결과를 초래하였으며 따라서 원거리 음성의 둘째음절 대 첫음절의 모음 지속시간 비율에 큰 변화를 가져왔다고 할 수 있겠다.

둘째음절 대 첫음절의 모음지속시간 비율의 화자 간 차이를 보기 위하여 각 발화상황별로 일원분산분석을 실시한 결과, 평이 음성에서는 화자 간 유의한 차이를 보이

1) F79는 원거리 음성에서 둘째음절 즉 경계성조에서 복합성조가 나타나는 것이 아니라 첫음절에서도 복합성조가 나타나는 특이한 패턴을 보였음.

표 3. 음절별 발화상황별 모음 에너지 평균, 표준편차 및 음절 간 차에 대한 Paired Samples t-test 결과

Table 3. Results of paired samples t-test for vowel energy differences between syllable positions in a word.

화자	발화상황	평균(표준편차)		샘플수	t-test 유의도
		첫음절 (dB)	둘째음절 (dB.)		
전체	원거리	62.92 (6.65)	65.50 (5.59)	139	.000
	평이	54.32 (9.50)	52.64 (9.31)	139	.000
M1	원거리	58.24 (3.90)	60.94 (2.03)	34	.001
	평이	47.68 (2.98)	48.76 (2.19)	34	.118
F26	원거리	61.49 (7.20)	66.54 (4.37)	35	.002
	평이	66.94 (6.37)	64.63 (5.89)	35	.030
M48	원거리	69.71 (3.75)	72.46 (2.15)	35	.002
	평이	56.80 (2.70)	54.80 (4.04)	35	.019
F79	원거리	62.11 (5.13)	61.91 (3.74)	35	.838
	평이	45.66 (4.19)	42.26 (4.19)	35	.000

표 4. 발화상황별 둘째음절 대 첫음절의 모음 에너지 비율 평균 및 표준편차와 발화상황 간 차이에 대한 Paired Samples t-test 결과

Table 4. Results of paired samples t-test for vowel energy ratio differences between speaking conditions.

화자	평균(표준편차)		샘플수	t-test 유의도
	원거리(%)	평이(%)		
전체	104.98 (12.38)	97.35 (10.43)	139	.000
M1	105.20 (9.47)	102.81 (9.99)	34	.010
F26	110.03 (17.56)	97.09 (10.98)	35	.000
M48	104.30 (7.78)	96.72 (9.05)	35	.000
F79	100.37 (10.86)	92.95 (9.59)	35	.000

지 않았으나 원거리 음성에서는 화자 간 유의한 차이를 보였으며[F(3,135)=3.143, p=0.027] Tukey 사후검정 결과 F79, M1, M48 < M1, M48, F26의 순으로서 F79과 F26 간에 유의한 차이를 보였다.

3.2. 에너지

음절 간 모음 에너지를 비교한 결과 표 3에서 볼 수 있는 바와 같이 발화상황에 따라 다른 양상을 보이고 있다. 대체적으로 원거리 음성에서는 첫음절에 비해 둘째음절에서 큰 값을 보이며 반대로 평이 음성에서는 둘째음절에 비해 첫음절이 큰 값을 보이고 있다. 화자별로 유의도를 살펴보면 평이 음성의 화자 M1과 원거리 음성의 F79을 제외하고 모두 유의한 차이를 보이고 있다.

표 5. F0대역폭과 최고값, 최저값의 성별 간 차이. '/' 왼쪽은 남성화자, 오른쪽은 여성화자의 값을 나타냄

Table 5. Gender differences in F0 range, F0 maximum and F0 minimum. Values to the left of "/" are for males while values to the right of "/" are for females.

분석대상	원거리 음성		평이 음성	
	평균(표준편차)	t-test 유의도	평균(표준편차)	t-test 유의도
대역폭	98.21/148.53 (43.87/57.65)	.000	46.61/86.40 (18.24/35.14)	.000
최고값	224.48/349.24 (32.32/31.11)	.000	131.49/238.44 (17.07/37.33)	.000
최저값	126.27/200.71 (33.115/46.32)	.000	84.89/152.04 (10.85/11.67)	.000

표 6. 화자별 F0대역 및 발화상황 간 대역폭 차에 대한 t-test결과

Table 6. F0 range for individual speakers and results of t-test for F0 range differences between speaking conditions.

화자	측정 대상	원거리	평이	평균 원거리/평이 (Hz)	대역폭 원거리/ 평이 (Hz)	t- test 유의 도
		최저/최고 (Hz)	최저/최고 (Hz)			
M1	최고점	180/222	102/177	200.17/125.91 (10.6/18.75)	71.66/ 39.88	.000
	최저점	84/182	73/140	128.51/86.03 (30.50/14.08)		
F26	최고점	288/390	186/294	333.64/227.03 (30.84/34.17)	129.64 /81.59	.000
	최저점	143/302	105/159	204/145.44 (46.65/10.00)		
M48	최고점	183/290	115/168	248.11/136.22 (28.55/13.41)	124.03 /52.24	.000
	최저점	84/214	76/105	124.08/83.78 (35.77/6.35)		
F79	최고점	300/400	201/353	364.83/249.86 (22.62/37.29)	167.66 /91.22	.000
	최저점	159/339	135/187	197.17/158.64 (46.37/37.29)		

발화상황에 따라 음절 간 에너지 비율이 달라지는지 보기 위하여 둘째음절 모음의 에너지 대 첫음절 모음의 에너지의 백분율을 구하였다 ((둘째음절/첫음절)*100). 표 4에서 볼 수 있듯이 음절 간 모음의 에너지 비율은 화자 4명 모두 발화상황 간에 유의한 차이를 보이고 있으며 지속시간에서 유의한 차이를 보이지 않은 화자 M1과 F79도 에너지에서는 유의한 차이를 보이고 있다. 이는 둘째음절, 즉 마지막 음절의 에너지를 크게 발음하는 것이 원거리 음성의 전달의 효율성을 높이는 기본 전략의 하나일 수 있지 않을까 한다.

둘째음절 대 첫음절의 모음에너지 비율의 화자 간 차이를 보기 위하여 각 발화상황별로 일원분산분석을 실시한 결과, 두 발화상황 모두에서 화자 간 유의한 차이를 보였다 (평이 음성 [F(3, 135)= 5.773, P=0.001], 원거리 음성 [F(3, 135)= 3.810, P=0.012]). Tukey 사후검정

표 7. 평이 음성의 화자별 억양 패턴. 음영은 'H L%'의 높은 사용빈도를 보기 쉽게 하기 위하여 사용하였음. '?'은 분석이 어려운 경우임

Table 7. Intonation pattern for individual speakers under normal speaking conditions. Shade is used for high frequency of 'H L%' and '?' is for intonation patterns difficult to analyze.

단어	F26	M48	F79	M1
검색	L HL%	HL%	HL%	HL%
날짜	HL%	L+H L%	HL%	L HL%
다음	HL%	L HL%	HL%	HL%
달기	HL%	L HL%	HL%	HL%
뒤로	HL%	HL%	HL%	HL%
메뉴	L+H L%	HL%	HL%	L HL%
메일	HL%	HL%	HL%	L HL%
반복	HL%	HL%	HL%	H L%
보기	HL%	HL%	HL%	L HL%
불름	HL%	HL%	HL%	HL%
설정	HL%	HL%	HL%	HL%
스톱	HL%	HL%	HL%	HL%
시작	HL%	HL%	HL%	HL%
암호	HL%	L HL%	HL%	L HL%
연결	HL%	HL%	HL%	L+H L%
열기	HL%	HL%	HL%	L+H L%
오전	HL%	HL%	HL%	L+H L%
오프	HL%	HL%	HL%	L HL%
오후	HL%	HL%	HL%	L+H L%
이전	HL%	L+H L%	HL%	HL%
일정	HL%	HL%	HL%	L+H L%
잠금	HL%	L+H L%	HL%	L+H L%
재생	HL%	L+H L%	HL%	HL%
저장	HL%	L+H L%	HL%	HL%
전송	HL%	?	HL%	L+H L%
정지	L+H L%	L+H L%	HL%	HL%
종료	HL%	?	HL%	HL%
주소	HL%	L+H L%	HL%	HL%
증지	HL%	HL%	HL%	?
처음	HL%	HL%	HL%	HL%
추가	HL%	HL%	HL%	HL%
축소	HL%	HL%	HL%	HL%
통화	HL%	HL%	HL%	HL%
해제	HL%	HL%	HL%	HL%
확대	HL%	HL%	HL%	HL%
확인	HL%	HL%	HL%	HL%

결과 평이 음성에서는 F79, M48, F26 < M48, F26, M1의 순으로 그룹이 지어졌으며 원거리 음성에서는 F79, M48, M1 < M48, M1, F26의 순으로 그룹이 지어졌다.

3.3. F0

성별 간 F0의 최고값, 최저값, 그리고 대역폭을 비교한 결과 표 5에서 볼 수 있듯이 두 발화상황 모두에서 성별 간 유의한 차이를 보이고 있으며 여성이 남성에 비해 큰 값을 보였다.

원거리 음성에서 F0의 변화 양상을 보기 위하여 단어 내에 F0의 최고점과 최저점을 측정하였다. F0 대역폭의

표 8. 원거리 음성의 화자별 억양 패턴. 음영은 'L HL%'의 높은 사용빈도 (특히 화자 M48과 F26)를 보기 쉽게 하기 위하여 사용하였음

Table 8. Intonation pattern for individual speakers under distant-talking conditions. Shade is used for high frequency of 'L HL%'.

단어	F26	M48	F79	M1
검색	L HL%	L+H L%	L HL%	L+H L%
날짜	L HL%	L HL%	L+H L%	L HL%
다음	L HL%	L HL%	H HL%	L+H L%
달기	L HL%	L HL%	L HL%	L HL%
뒤로	L+H L%	L+H L%	HL HL%	L+H L%
메뉴	L+H L%	L HL%	L HL%	L HL%
메일	L HL%	L HL%	L HL%	L HL%
반복	L HL%	L HL%	L HL%	L+H L%
보기	L HL%	L HL%	HL HL%	L HL%
불름	L HL%	L HL%	L HL%	L HL%
설정	H HL%	L HL%	HL HL%	H L%
스톱	H HL%	H HL%	H L%	H L%
시작	?	L HL%	H HL%	H L%
암호	L HL%	L HL%	L+H HL%	L HL%
연결	L HL%	L HL%	L+H HL%	L+H L%
열기	L HL%	L HL%	L+H HL%	L+H L%
오전	L HL%	L HL%	L+H HL%	L+H L%
오프	L HL%	L HL%	L+H HL%	L HL%
오후	L HL%	L HL%	L+H HL%	L HL%
이전	L HL%	L HL%	L+H L%	L HL%
일정	L HL%	L HL%	L+H HL%	L+H L%
잠금	L HL%	L HL%	L+H L+H L%	L+H L%
재생	L HL%	L HL%	L HL%	L+H L%
저장	L HL%	L HL%	H L+HL%	L+H L%
전송	L HL%	L HL%	L HL%	L+H L%
정지	L HL%	L HL%	L+H L+H L%	L+H L%
종료	L HL%	L HL%	H HL%	L+H L%
주소	L+H L%	L+H HL%	L HL%	L+H L%
증지	L HL%	L HL%	L+H L+H L%	L+H L%
처음	H L%	H HL%	H L+HL%	H L%
추가	H L%	L HL%	H L%	H L%
축소	H L%	L HL%	H HL%	H L%
통화	H L%	L HL%	H L%	H L%
해제	H L%	L HL%	H HL%	L+H L%
확대	H L%	L+H HL%	H HL%	H L%
확인	H L%	L+H HL%	H L+HL%	L HL%

원거리 음성과 평이 음성 간 차이를 보기 위하여 t-test를 실시한 결과 표 6에서와 같이 모든 화자가 두 발화상황 간에 유의한 차이를 보이고 있으며 원거리 음성에서의 대역폭이 평이 음성에 비해 큰 대역폭을 보이고 있다. 이러한 원거리 음성에서의 큰 대역폭의 구체적인 실현 양상을 보기 위하여 F0 최고점과 최저점의 평균값을 비교하였으며 이 두 지점의 최저값과 최고값 비교 또한 수행하였다. 평이 음성에 비해 상대적으로 큰 F0 대역폭

실현 방법은 적어도 두 가지가 있을 수 있겠다. F0 최고점과 최저점을 동시에 상승시키되 최고점 상승폭에 비해 최저점의 상승폭을 적게 하거나, 최고점은 상승시키되 최저점은 하락시키는 방법이 있을 수 있겠다. 표 6에서 보듯이 화자 4명 모두의 F0 최고점과 최저점의 평균값이 평이 음성에 비해 원거리 음성에서 큰 값을 보일 뿐더러 이 두 지점의 최저값과 최고값 비교에서도 비교 대상의 모든 값이 평이 음성에 비해 원거리 음성에서 큰 값을 보임으로써 F0 최고점과 최저점의 동반상승에 의한 F0 대역폭 확대가 이루어졌음을 알 수 있다.

F0 대역폭의 화자 간 차이를 보기 위하여 각 발화상황별로 일원분산분석을 실시한 결과, 두 발화상황 모두에서 화자 간 유의한 차이를 보였다 (평이 음성 [F(3, 139)=25.944, P=0.000], 원거리 음성 [F(3, 139)=26.799, P=0.000]). Tukey 사후검정 결과 평이 음성에서는 M1 < M48, F26 < F79의 순으로 그룹이 지어졌으며 원거리 음성에서는 M1, M48 < F26, F79의 순으로 그룹이 지어졌다. 화자 간 그룹이 어떻게 지어지는가에서는 두 발화상황 간에 차이를 보이고 있으나 화자 간 순서에는 변화가 없음을 볼 수 있으며 여성 화자의 대역폭이 남성 화자의 것에 비해 큰 것으로 나타났다. 평이 음성에서는 남성화자 M1의 대역폭이 가장 작고 여성화자 F79의 대역폭이 가장 크며 나머지 두 화자의 대역폭은 이 두 화자의 중간값을 보인 반면, 원거리 음성에서는 성별에 따라 그룹을 달리 함을 볼 수 있다.

3.4. 억양 패턴

평이 음성과 원거리 음성 간에 억양 패턴의 차이를 보이기 위하여 평이 음성에서의 억양 분석을 수행하였다. 표 7에서 보듯이 평이 음성에서의 억양 패턴은 화자에 따라 정도는 다르지만 'H L%'가 주를 이루고 있는 반면, 원거리 음성에서는 화자 간에 차이를 보이고 있다 (표 8 참조). 화자 M48과 F26은 'L HL%' 패턴이 주를 이루고 있는 반면, 화자 M1은 'L HL%' 패턴보다는 'L+H L%' 패턴을 보다 빈번히 사용하고 있으며 화자 F79은 'L HL%' 패턴 외에 'L+H HL%' 패턴 또한 많이 사용하고 있다. 흥미로운 것은 첫음절의 시작성조와 둘째음절의 경계성조 패턴에 있어서 화자 간에 그룹핑을 달리 하고 있다는 점이다. 예를 들어, 경계성조의 경우, 화자 M48, F26, F79는 공통적으로 복합 경계성조 'HL%'를 사용하는 반면 화자 M1은 단순 경계성조 'L%'를 주로 사용하고 있다. 첫음절 성조의 경우는 화자 M48과 F26은 'L' 또는 'H' 중 한 개

의 성조가 실현되는 반면, 화자 F79과 M1의 경우에는 'L' 과 '+H' 두 개의 성조가 다 실현되고 있다. 즉, 원거리 음성 전달 전략으로 M48과 F26은 복합 경계성조를 사용하여 마지막 음절을 더 들리리게 하는 반면, M1은 첫음절에 두 개의 성조를 다 사용함으로써 첫음절을 보다 더 들리리게 하고 있으며 F79은 이 두 가지 전략을 모두 사용하여 두 음절 모두 들리도록 하는 전략을 사용하고 있다고 볼 수 있다.

한 가지 주목할 것은 화자 M48과 F26이 위에 논의한 억양 패턴에서는 같은 양상을 보였으나 강세구 시작 성조에서 두 화자 간에 차이를 관찰 할 수 있었다. 두 화자 모두 평이 음성에서는 첫음절이 어떤 분절음으로 시작하든 대부분 'H'로 시작하는 데 반해 원거리 음성에서는 화자 F26은 첫음절의 분절음 유형에 상관없이 'L'로 시작하는 반면 화자 M48은 강자음으로 시작하는 경우에는 'H'로 시작하고 그 외에는 'L'로 시작하고 있다(3).

또 한 가지 특기할 만한 것은 see-saw효과[14]로서 본 연구에서와 같이 하나의 강세구 그 자체가 억양구를 이룰 때 강세구 내의 음절수가 적은 경우에 억양구의 경계성조 유형에 따라 강세구 내의 성조가 영향을 받을 수 있음을 보여주고 있다는 것이다. 그 예로 평이 음성에서 경계성조가 'L%'인 경우, 강세구가 강자음으로 시작하지 않더라도 첫 성조가 'L'로 실현되지 못하고 'H'로 실현되고 있다. 이와는 반대로 원거리 음성에서 경계성조가 'HL%'인 경우, 강세구가 강자음으로 시작하더라도 첫음절 성조 'H'가 'L'로 바뀌는 경우도 관찰되었다. (화자 M48).

IV. 결 론

본 연구는 2음절어 고립어를 이용하여 원거리 음성의 운율적 특성을 살펴보았다. 화자에 따라 정도를 달리 하지만 대체적으로 모음 지속시간은 발화상황에 상관없이 첫음절에 비해 둘째음절에서 길며 음절 간 모음의 지속시간 비율 ((둘째음절/첫음절)*100)은 평이 음성에 비해 원거리 음성에서 큰 값을 보였다. 에너지는 원거리 음성에서 첫음절에 비해 둘째음절에서 큰 값을 보이며 평이 음성에서는 그 반대 양상을 보였다. 음절 간 모음의 에너지 비율 ((둘째음절/첫음절)*100)은 평이 음성에 비해 원거리 음성에서 통계적으로 유의하게 큰 값을 보였다.

F0 대역폭 또한 두 발화상황 간에 유의한 차이를 보였으며 원거리 음성에서 F0 최고점과 최저점의 동반상승에 의한 F0 대역폭의 확대를 볼 수 있었다. 억양 패턴은 평이 음성의 단순 경계성조 'L%'가 원거리 음성에서 복합 경계성조 'HL%'로 바뀌거나 평이 음성 첫음절에 'L' 또는 'H' 시작 성조 하나만이 실현되는 대신 원거리 음성에서는 보통 첫음절과 둘째음절에 실현되는 'L'과 '+H' 두 개 성조가 모두 실현되기도 하였다.

이와 같이 네 명의 화자 모두 공통적으로 운율적인 측면에서 에너지, F0, 억양패턴이라는 운율적 특성 변화로 원거리 음성의 효율적인 전달을 꾀하는 것으로 나타났으나, 구체적인 변화 양상에서 화자 간 차이를 보였다. 원거리 음성에서 첫음절의 에너지에 비해 둘째음절의 에너지 증가율이 평이 음성에 비해 높다는 원거리 음성에서 F0 최고점과 최저점의 동반상승에 의한 F0 대역폭의 증가 등은 화자 모두 비슷한 경향을 보였으나 원거리 음성에서의 억양 패턴은 화자에 따라 다른 양상을 보였다. 평이 음성에서의 억양 패턴은 화자에 따라 정도는 다르지만 'H L%'가 주를 이루고 있는 반면, 원거리 음성의 경우는 화자 간 차이를 보였다. 화자 M48과 F26은 복합 경계성조 'HL%'를 사용하여 마지막 음절을 더 돌돌리게 하는 반면, 화자 M1은 보통의 경우 첫음절과 둘째음절에 실현되는 'L'과 '+H' 성조를 첫음절에 다 실현시키고 경계성조는 단순 성조인 'L%'를 실현시킴으로써 첫음절을 보다 더 돌돌리게 하였으며, 화자 F79는 이 두 가지 전략, 즉 두 음절 모두에 복합성조를 사용하여 두 음절 모두 돌돌리도록 하는 전략을 사용하는 것으로 나타났다. 요약하자면, 일반적으로 둘째음절의 지속시간과 에너지를 첫음절에 비해 길게 또는 크게 함으로써, f0 대역폭을 평이 음성에 비해 원거리 음성에서 더 크게 함으로써, 그리고 단순 성조대신 복합 성조 사용으로 억양 패턴을 달리 함으로써 원거리 음성의 전달 효과를 높이려는 것으로 나타났다.

감사의 글

본 논문에서 사용된 음성 데이터는 한국전자통신연구원(ETRI)에서 제작되었습니다. 또한 본 연구는 제1저자의 원광대학교 2005학년도 교내 일반과제 연구비 지원으로 수행되었습니다.

참고 문헌

1. J.-C. Junqua, "The Influence of Acoustics on Speech Production: A Noise-Induced Stress Phenomenon as the Lombard Reflex," *Speech Communication* 20, 13-22, 1996.
2. J. Hansen, "Analysis and compensation of speech under stress and noise for environmental robustness in speech recognition," *Speech Communication* 20, 151-173, 1996.
3. E. Lombard, "Le signe de l'ivation de la voix", *Annales des maladies de l'oreille, du larynx, du nez et du pharynx*, 37, 101-119, 1911.
4. S.H. Ferguson, and D. Kewley-Port, "Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners," *J. Acoust. Soc. Am.* 112(1), 259-271, 2002.
5. S. Kóster, "Acoustic Characteristics of Hyperarticulated speech for different speaking style," *Proc. ICASSP*, 2, 873-876, 2001.
6. J.-S. Linéard, and M.-G. Di Benedetto, "Effect of vocal effort on spectral properties of vowels," *J. Acoust. Soc. Am.* 106(1), 411-422, 1999.
7. S.J. Moon & B. Lindblom, "Interaction between duration, context, and speaking style in English stressed vowels," *J. Acoust. Soc. Am.* 96(1), 40-55, 1994.
8. 우수영, 롭바드 효과 보상 필터를 이용한 강인한 특징 추출 방법. 한국과학기술원 석사논문, 2003.
9. A.M.C. Sluijter and V.J. van Heuven, "Spectral balance as an acoustic correlate of linguistic stress," *J. Acoust. Soc. Am.* 100(4), 2471-2485, 1996.
10. A.M.C. Sluijter, V.J. van Heuven, and J.J.A. Pacilly, "Spectral balance as a in the perception of linguistic stress," *J. Acoust. Soc. Am.* 101(1), 503-513, 1997.
11. 김선희, "한국어 롭바드 음성의 지속시간," *말소리* 54호, 1-14, 2005.
12. 이숙향, 김선희, "한국어 원거리 음성의 모음의 음향적 특성," *말소리* 55호, 61-76, 2005.
13. Sun-Ah Jun, "K-ToBI (Korean ToBI) labelling conventions: Version 3," *UCLA Working Papers in Phonetics* 99, 149-173, 2000.
14. Sun-Ah Jun, "Influence of microprosody on macroprosody: a case of phrase initial strengthening," *UCLA Working Papers in Phonetics* 92, .97-116, 1996.

저자 약력

• 이 숙 향 (Sook-hyang Lee)



1980년 2월: 서울대학교 사범대학 영어교육과 (학사)
 1984년 8월: 서울대학교 대학원 언어학과 (석사)
 1994년 3월: Ohio State University 언어학과 (박사)
 1995년 3월~1996년 3월: 원광대학교 시청각교육관
 전임강사
 1996년 4월~1997년 3월: 원광대학교 영어영문학과
 전임강사
 1997년 4월~2001년 3월: 원광대학교 영어영문학과
 조교수
 2001년 4월~2006년 3월: 원광대학교 영중어문화부 부교수
 2006년 4월~현재: 원광대학교 영중어문화부 교수

• 김 선 회 (Sunhee Kim)



1963년 3월 9일생.
1985년 2월: 연세대학교 불어불문학과 문학사
1986년 10월: 파리7대학 (프랑스)언어학석사
1990년 10월: 고등사회과학대학원 (프랑스) 언어학박사
1991년 3월~2002년 2월: 연세대학교 문과대학 시간강사
2002년 3월~2004년 2월: 광운대학교 연구교수
2004년 9월~2005년 8월: KAIST BK21 연구교수
2005년 9월~현재: 서울대학교 인문정보연구소 선임연구원

• 김 종 진 (Jong-Jim Kim)



1969년 8월 24일생
1995년 2월: 원광대학교 컴퓨터공학과 (학사)
1997년 2월: 원광대학교 컴퓨터공학과 (석사)
2000년 2월: 원광대학교 컴퓨터공학과 (박사수료)
2000년 7월~현재 : 한국전자통신연구원 (ETRI)
음성언어정보연구부 선임연구원