한국어 장애음 지각에서의 VOT와 F0의 상관 관계

김미담
서울대학교

The Correlation of VOT and F0 In the Perception of Korean Obstruents

Midam Kim
Seoul National University

midam@dreamwiz.com

Abstract

The present thesis examines the correlation of VOT and F0 in the three-way distinction of Korean obstruents, conducting production and perception tests. In the production test, one female native speaker of Korean with a Seoul dialect (the author) recorded 15 repetitions of a monosyllabic word list including /ka, kha, k*a, pa, pha, p*a, ta, tha, t*a, ca, cha, c*a/ in random order. VOT and F0 of the following vowels were measured, and the result was significant for the three-way distinction with a strong correlation between VOT and F0, and also in the VOT-F0 plot, no overlapping among the domains was observed. As for the perception test, I manipulated the data recorded in the production test, heightening or lowering their F0 values. In all, 14 subjects (seven males and seven females) participated in the identification test. The result was as follows: the fortis stimuli were not influenced by F0 changes, and the VOT and F0 values at the lenis-aspirated boundary were negatively correlated. From these results I concluded the following: 1) VOT and F0 can distinguish the three domains of Korean obstruents without overlapping; 2) the fortis perception does not need F0 as its acoustic cue; and 3) VOT and F0 in the distinction between the lenis and aspirated are in the phonetic trading relation[2].

Ⅰ. 서론

그 동안 한국어 장애음의 삼중 대립에 대한 연구에서 가장 많이 다루어진 음향 단서는 VOT와 후행모음의 F0였다. VOT는 경음을 평음과 격음으로부터 분리하며, F0는 평음을 경음과 격음으로부터 분리한다고 알려져 있지만, VOT와 F0, 두 음향 단서에 있어서 모두 영역간 중첩 현상이 종종 보고 되어 왔다. 즉, VOT도 F0도, 단독으로는 한국어의 삼중 대립을 만들 수 없는 것으로 보인다. 여기서 우리는 두 단서 사이의 상관 관계에 의한 삼중 대립의 가능성에 대해 생각해 볼 수 있겠다.

그간 한국어 삼중 대립에 대한 연구는 대개 산출 실험이었고, 지각 실험은 많지 않았다[1,3,4,5,8]. VOT든 F0든 어떤 요소가 한국어 자음 지각에 정말로 "기여" 하는지 확증하려면, 산출 실험 결과로는 부족하며, 반드시 변인 통제된 지각 실험 결과가 있어야 한다. 산출 실험 결과는 발화에 어떤 요소가 얼만큼 들어 있는지 보여 주는 것일 뿐, 그 요소가 실제 지각에 미치는 영향을 보여 줄 수는 없기 때문이다.

삼중 대립을 이루는 자음에 대한 지각 실험에 비해 후행 모음에 대한 지각 실험은 비교적 최근에 시작되었다.[3,4,5,8] 그러나 그 중에는 후행 모음의 주파수 조작이라는 기술적 문제 때문에 변인 통제된 실험[3,5]이 거의 없었다. 연구하기로 한 음향 단서만을 조작하고, 나머지 변수를 통제하는 과정 없이 얻은 결과로는 과연 어떤 변수가 결과에 영향을 미쳤는지 판단할 수 없게 되며, 그러한 상황에서 내려진 결론은 근거가 미약할 수 밖에 없다[4,8]. 그러므로 지각 실험에서 합당한

결론을 도출하려면 자극에 대한 변인 조작, 통제 실험이 꼭 필요하다.

이 논문에서는 모음의 F0 값을 조작하고 산출 실험 결과 얻어지는 다양한 VOT 값을 이용하여, VOT와 F0의 상관 작용에 의해 한국어 장애음의 삼중 대립이 어떻게 지각되는지 알아 보았다. 먼저, 지각 실험의 예비 실험으로서 산출 실험을 하고, 그 결과를 가지고 지각 실험을 진행하였다.

Ⅱ. 산출 실험

1. 실험 방법

20대의 서울 여성 화자(저자)가 단위 문장 없이 무순으로 구성된 한국어 단음절 단어 목록 가 /ka, kha, k*a, pa, pha, p*a, ta, tha, t*a, ca cha, c*a/를 일정한 음높이로 15번 발화하였다.

발화는 AKG C1000S 마이크를 통해 조용한 방에서 SB Platinum Ex 사운드 카드가 장착된 컴퓨터에 직접 녹음 되었다. 표본 추출률은 22050 Hz였고, 마이크와 발화자 사이의 거리는 5 cm로 고정되었다. 파일 저장은 wave로 하였다.

CoolEdit2000을 이용하여 전체 발화 내용에서 각 단어를 잘라내었으며 Praat 4.0.53을 이용하여 VOT와 후행 모음 시작 부분의 F0를 측정하였다. VOT는 파형상 파열이 시작하는 부분부터 스펙트로그램상 모음의 F1 시작점에서 가장 가까운 영교차점까지로 측정하였다. F0는 VOT가 끝나는 지점, 즉 성대 진동이 시작하는 지점에서 측정하였다.

2. 측정 결과

그림 2.1.을 보면, VOT는 경음이 가장 짧고, 격음이 가장 길며, 평음은 중간값을 보이고, F0는 평음이 가장 낮고, 격음이 가장 높으며, 경음이 중간값을 보이는 것으로 나타났다. 세 영역을 VOT와 F0 중 한 가지 척도로만 나누자면 중첩 현상이 일어나지만, 두 가지 척도를 가지고 나누면 세 영역이 중첩 없이 나뉘는 것을 볼 수 있다.

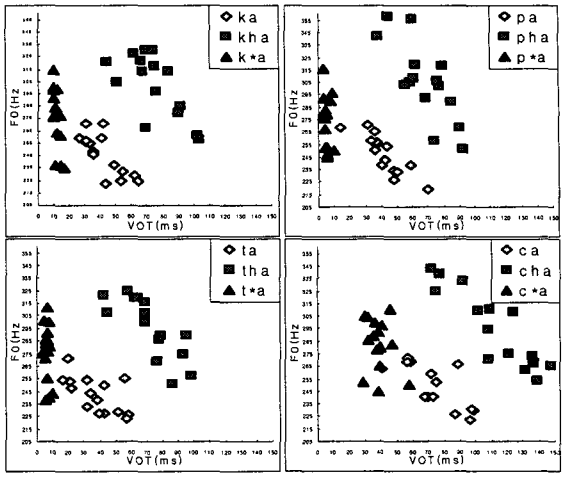

그림 2.1. 조음 위치별 평음/격음/경음의 VOT-F0 산포도

표 2.1. VOT와 F0의 평균, 표준 편차와 범위

		연구개 파열음		양순 파열음		치조 파열음		경구개 파찰음	
		VOT	F0	VOT	F0	VOT	F0	VOT	F0
평음	평균	41.8	246	41.7	247	37.4	242	72.9	250
	표준편차	11.75	16.7	12.87	15.5	13.96	14.4	17.44	17.6
	최소값	26.2	222	13.6	219	16.2	223	41.1	222
	최대값	63.6	271	69.8	271	58.2	271	98.2	272
격음	평균	75.0	302	68.1	305	72.4	296	112.2	295
	표준편차	16.96	25.6	15.92	31.6	16.99	24.2	24.82	30.5
	최소값	44.1	258	38.1	251	42.8	251	71.3	253
	최대값	103.4	331	92.2	358	98.9	325	147.6	343
경음	평균	11.7	277	5.3	274	6.3	276	38.8	282
	표준편차	2.26	25.2	1.82	21.1	1.64	22.8	7.37	21.0
	최소값	9.6	235	3.0	245	3.7	238	29.4	245
	최대값	16.7	315	9.5	316	9.9	312	57.8	311

평음, 격음, 경음의 VOT와 F0 값이 각각 영역별로 구별되는지 알아 보기 위해, 유의 수준 0.05의 ANOVA 분석과 Tukey's HSD 사후 분석을 실시하였다(표 2.2.). 그 결과, 세 영역에서의 VOT와 F0값이 cha와 c*a의 영역 간 F0 값을 제외하고 모두 유의미하게 구별되는 것으로 나타났다.

표 2.2. VOT와 F0의 ANOVA 결과

	VOT		F0	
velar	F(2,42)=104.685	p<0.0005	F(2,42)=22.478	p<0.0005
	ka-k^{h}a	p<0.0005	ka-k^{h}a	p<0.0005
	kha-k*a	p<0.0005	k^{h}a-k*a	p=0.013
	k*a-ka	p<0.0005	k*a-ka	p=0.002
labial	F(2,42)=105.876	p<0.0005	F(2,42)=22.379	p<0.0005
	pa-p^{h}a	p<0.0005	pa-p^{h}a	p<0.0005
	pha-p*a	p<0.0005	p^{h}a-p*a	p=0.002
	p*a-pa	p<0.0005	p*a-pa	p=0.01
alveolar	F(2,42)=101.202	p<0.0005	F(2,42)=25.093	p<0.0005
	ta-t^{h}a	p<0.0005	ta-t^{h}a	p<0.0005
	t^{h}a-t*a	p<0.0005	t^{h}a-t*a	p=0.037
	t*a-ta	p<0.0005	t*a-ta	p<0.0005
palatal affricate	F(2,42)=62.284	p<0.0005	F(2,42)=14.502	p<0.0005
	ca-c^{h}a	p<0.0005	ca-c^{h}a	p<0.0005
	c^{h}a-c*a	p<0.0005	c^{h}a-c*a	p=0.326
	c*a-ca	p<0.0005	c*a-ca	p=0.001

이번에는 VOT와 F0값의 상관이 세영역 사이에서 구별되는지 알아 보기 위해 유의 수준 0.05의 MANOVA 분석을 실시하였다(표 2.3.). 그 결과, 모든 조음 위치에 있어서 세 영역에서의 VOT와 F0의 상관이 유의미하게 구별되는 것으로 나타났다.

표 2.3. VOT와 F0의 MANOVA 결과

다변량 검정(Wilks의 람다)			
효과	값	F	유의확률
ka/kha/k*a	0.053	F(4,82)=68.246	0.000
pa/pha/p*a	0.045	F(4,82)=76.453	0.000
ta/tha/t*a	0.053	F(4,82)=68.967	0.000
ca/cha/c*a	0.069	F(4,82)=57.507	0.000

Ⅲ. 지각 실험

1. 실험 방법

산출 실험에서 녹음된 자료의 F0 값을 Praat의 manipulation 기능으로 조작하여 지각 실험용 자극을 만들었다. F0의 조작 범위는 각 조음 위치별로 평음은 격음의 최고점까지, 격음은 평음의 최저점까지, 경음은 격음의 최고점과 평음의 최저점까지로 하여, 각각 10 Hz씩 변화시켰다. 한 음절 내의 F0 변화 곡선은 또 다른 음향 단서가 될 수 있으므로[7], 모든 자극은 단 하나의 F0 값만을 갖도록 했다. 조작은 산출 실험 녹음된 모든 자료에 대해 행해졌으므로 결과적으로 표 2.1.에 제시한 VOT 범위에 아래 표 3.1.과 그림 3.1.에 제시한 F0 범위를 가진 자극 2014개를 만들게 되었다.

표 3.1. 자극의 개수와 F0 범위

원본	개수	F0 조작 범위	원본	개수	F0 조작 범위	원본	개수	F0 조작 범위	원본	개수	F0 조작 범위
ka	150	222	pa	185	219	ta	147	223	ca	164	222
k^{h}a	140	to	p^{h}a	149	to	t^{h}a	131	to	c^{h}a	132	to
k*a	193	331	p*a	233	358	t*a	180	325	c*a	210	343
합계	483		sum	567		sum	458		sum	506	

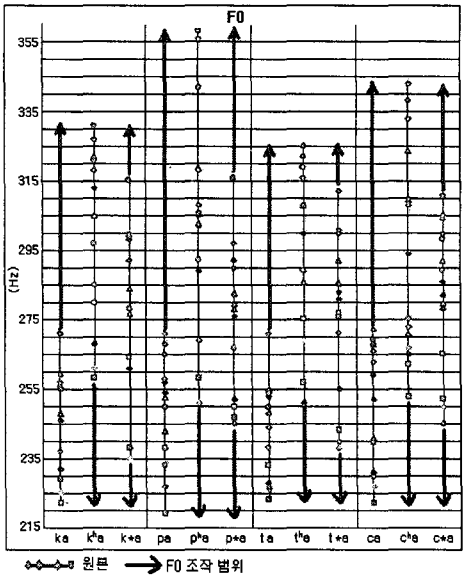

그림 3.1. F0 조작 범위

피실험자는 남녀 각 7명, 총 14명의 20대 서울말 화자로 하였으며, 모두 청력에 이상이 없었다.

실험은 서울대학교 언어학과의 방음실에서 진행되었으며, 지각 실험용 프로그램으로는 Inquisit 1.33이 이용되었다. 피실험자는 컴퓨터 앞에 앉아 Sennheiser eH 1430 헤드폰을 쓰고서 모니터 앞에 나타난 세 개의 한글 글자 중에서 들리는 소리에 해당하는 글자를 마우스로 찍도록 하였다.(identification test) 조음 위치별로 4 개의 실험을 구성하였으며 피실험자는 나흘에 걸쳐서 각각 다른 날에 4 개의 실험을 마쳤다. 각 실험은 10개의 블록으로 구성되었고, 각 블록을 실행하는 데는 2분 가량이 소요되었으며, 블록간 쉬는 시간을 포함하여 하나의 실험에는 30분 정도가 소요되었다.

2. 실험 결과

지각 실험 결과, 경음의 조작본에 대한 반응은 70% 이상 '경음'이었으므로, 경음의 지각은 F0 변화에 영향을 받지 않는다는 결론을 내릴 수 있었다.

평음과 격음의 조작본에 대한 반응은 F0 값에 따라 범주적 지각을 보였다. 즉, 원본이 평음이었든, 격음이

었든 간에, F0 값이 높아질수록 격음의 반응률이 높아지고, F0 값이 낮아질수록 평음의 반응률이 높아졌다.

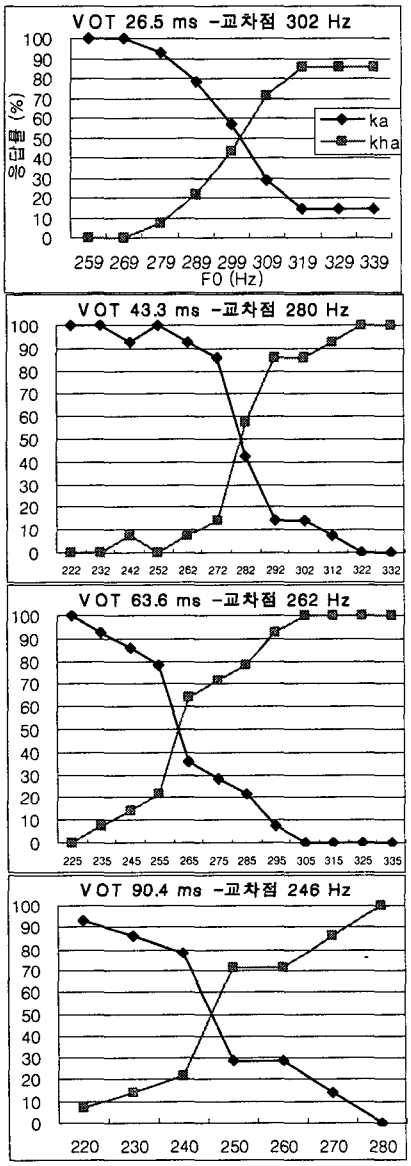

그림 3.2. VOT가 26.5, 43.3, 63.6, 90.4 ms인 자극에서의 F0에 따른 /ka,-kha/의 반응률

이 현상과 해당 VOT 값을 연관 지어 살펴 보면, VOT 값이 높아질수록 평음 반응과 격음 반응의 교차점의 F0 값이 낮아지는 것을 관찰할 수 있었다(그림 3.2.).

이에 따라, 평음 반응과 격음 반응의 교차점, 즉, 평음 지각과 격음 지각의 경계점의 VOT와 F0의 상관계수를 구해 보니 모든 조음 위치에서 -0.9를 넘는 높은 음의 상관을 보였다(표 3.2.).

표 3.2. 평음-격음 경계점의 VOT-F0 상관 계수경계점의 VOT-F0 상관 계수

경계	경계점의 VOT-F0 상관 계수
ka-k^ha	-0.923
pa-p^ha	-0.956
ta-t^ha	-0.954
ca-c^ha	-0.942

이 경계점들을VOT-F0 평면에 산포도로 나타내면 그림 3.3.와 같다. 그림에 보인 직선은 각 조음 위치별 점들의 추세선이다.

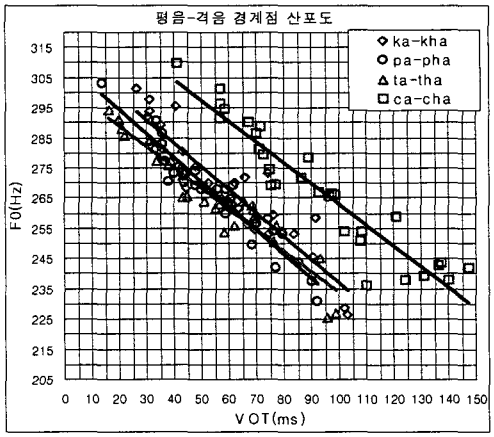

그림 3.3. 평음-격음의 경계점 산포도와 추세선

그림 3.3에서 조음 위치별로 보인 각 추세선의 아랫부분은 50% 이상 평음으로 지각된 영역이며, 그 윗부분은 50% 이상 격음으로 지각된 영역이다.

격음 지각을 중심으로 보면, VOT가 짧으면 F0가 높아야 격음으로 인식되고, VOT가 길면 F0가 낮아도 격음으로 인식된다. 평음 지각을 중심으로 다시 생각해 보면, F0가 높으면 VOT가 짧아야 평음으로 인식되고, F0가 낮으면 VOT가 길어도 평음으로 인식된다고 해석할 수 있다.

그러므로 한국어 자음 지각에 있어서 VOT가 아주 짧으면 그 F0 값과 관계 없이 경음으로 인식되며, VOT가 그보다 긴 경우에는 VOT와 F0가 큰 쪽이 격음으로, VOT와 F0가 작은 쪽이 평음으로 인식되되, 그 때의 평음-격음 지각의 경계에서는 VOT와 F0가 음의 상관 관계에 있게 된다고 결론 내릴 수 있다.

Ⅳ. 토의

이상의 결과에서 보면, 한국어 자음의 삼중 대립에서, 경음은 홀로 '아주 짧은 VOT'라는 조건만 가지고서 분리되며, 격음과 평음은 상당 부분의 VOT 값과 상당 부분의 F0 값을 공유하되 VOT와 F0의 상관 관계를 이용하여 분리됨을 알 수 있다. 이 연구가 지금까지의 연구들과 가장 다른 점이 여기에 있다고 하겠다. 즉, 격음과 평음의 구분은 VOT만으로도, F0만으로도 불충분하며, VOT와 F0의 상호 작용이 있어야 가능하다는 것이다. 이를 통해, 지금껏 보고되었던 격음-평음 간 VOT, F0 범위의 겹침 현상에 대한 의문이 해소될 수 있을 것이다.

이러한 지각 단서 통합 현상은 영어에 대한 연구에서는 많이 밝혀진 바 있다.[2] 한국어의 지각 단서로서 그 동안 밝혀진 여러 음향 단서들에 대해서도 이러한 연구가 이어져야 할 것으로 생각한다. 하나의 소리를 지각할 때는 그 관련 지각 단서들이 각각 따로 작용하는 것이 아니라, 하나의 통합된 단서로서 작용하기 때문이다.[6]

Ⅴ. 결론 및 향후 과제

이 논문에서는 VOT와 F0에 의한 한국어 장애음의 삼중 대립 지각을 알아 보기 위해 산출 실험과 지각 실험을 실행하였다. 산출 실험 결과,VOT와 F0는 각각 평음, 격음, 경음의 세 영역에서 유의미한 구별을 보이며 분포하였으며, 상관된 VOT와 F0 또한 세 영역에서 유의미하게 구별되어 분포함을 알 수 있었다. 지각 실험 결과는, 아주 짧은 VOT는 경음을 평음과 격음으로부터 분리하며, VOT가 경음보다 긴 범주에서는 VOT와 F0가 상보적으로 기여하여 평음과 격음을 분리하는 것으로 나타났다(그림 3.4).

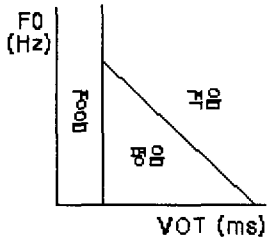

그림 3.4. VOT-F0 평면에서의 한국어 삼중대립 지각도

앞으로는 같은 자극에 대한 discrimination test를 통하여 정확한 격음-평음 경계 값을 얻어 볼 계획이며, 이번 연구에서 배제했던 지각 단서인 F0 변화 곡선에 대한 지각 실험, 화자에 따른 지각 특성, 특히 성조 언어로 알려진 경상 방언에서의 삼중 대립 실현 양상에 대한 연구, VOT와 F0 이외의 지각 단서들의 상호 작용에 대한 연구, 단음절 무의미 단어가 아닌 유의미 단어를 자극으로 하는 지각 실험, F0 뿐 아니라 VOT의 길이에 대한 조작 실험 또한 뒷따라야 할 것으로 생각한다.

참고문헌

[1] 이경희, 정명숙, "한국어 파열음의 음향적 특성과 지각 단서", *음성과학*, 7, 2, 2000.

[2] Repp, B.H., "Phonetic trading relations and context effects: New experimental evidence for a sppech mode of perception" *Psychological Bulletin*, 92, 81-110, 1982.

[3] Mi-Ran Kim, Charles Read, Keith Kluender, & Andrew Lotto, "Production and perception of work-initial stops by Korean adults", *JASA*, 94, 1993.

[4] T. Cho, "Vowel Correlates to Consonant Phonation: An Acoustic-Perceptual Study of Korean Obstruents", M.A. Thesis, University of Texas at Arlington, 1996.

[5] Jeoungim Han, "The Phonetics and Phonology of "Tense" and "Plain" Consonants in Korean", Ph. D dissertation, Cornell University, 1996.

[6] Pickett, J. M., *The Acoustics of Speech Communication: fundamentals, speech perception theory, and technology*, Allyn & Bacon, 1999.

[7] Mi-Ryoung Kim, "The nature and magnitude of segmental effects of F0 in Korean", *Eoneo*, 26, 3, 2001.

[8] Mi-Ryoung Kim, Patrice Speeter Beddor, & Julie Horrocks "The contribution of consonantal and vocalic information to the perception of Korean initial stops", *Journal of Phonetics*, 30, 2001.