

# 라링고그래프에 의한 한국인의 성문파형 분석 및 합성모음의 청각실험\*

양 병 곤  
(동의대학교)

1

Yang, Byunggon. 1996. An Analysis of Korean Glottal Waves by the Laryngograph and a Perceptual Study of Synthesized Vowels. *Korean Journal of Linguistics*, 21-4, 1025-1040. Glottal waveforms of vowels, words and sentences produced by a healthy Korean male and a female speakers were collected by the laryngograph. These waves were compared with those of sound pressure. Also, source parameters of vowel [i] and [ja] were partially modified by a formant synthesis software program and presented to 20 male and female listeners. Results are as follows: First, no predictable pattern of the glottal waves of vowels was observed. Second, intonation and tone variation did not result in different glottal waves. Third, the VOT position of glottal waves synchronized with the initial vowel pattern of the sound pressure waves, thus the voicing of intervocalic stop could be better identified by the glottal waves. Fourth, no noticeable difference was perceived when the source parameters of Oq and Sq were varied but overall auditory impression of the synthesized vowels with modified TL sounded like low-pass filtered signals. This study will be helpful in developing automatic speech synthesizers, especially modelling the source. (Dong-eui University)

## 1. 머리말

음성은 원음을 생성하는 성대와 이를 필요한 형태로 걸러내는 성도의 조합으로 생성된다(Fant 1960). 성대의 진동은 기본주파수로 나타나며, 화자의 감정이나 특정한 방언에 따라 독특한 유형을 보여준다. 또한, 발성방식에 따라 미세한 음질의 변화도 보인다. 따라서, 성대의 진동은 음성분석이나 합성 및 외국어 학습에 필요한 기본적인 자료를 제공해주므로 연구할 가치가 있다.

본 논문에서는 Fourcin이 개발한 라링고그래프로 정상적인 건강한 한국인 남녀화자가 발성한 모음, 단어 및 문장의 성문파를 측정하여 음압파와 비교 분석을 통해 연구하고 또한 음원의 파라미터를 변화시켜 합성한 모음

---

\*이 연구는 1995년도 한국학술진흥재단 자유공모과제 연구조성비에 의하여 연구되었음.

의 청각 실험을 한다. 구체적으로 다음과 같은 연구문항에 대한 해답을 구하여 한국어 음성 분석 및 합성에 대한 음원의 기본자질을 깊이 있게 이해하는데 도움을 주고자 한다.

- (1) 라링고그래프에서는 성대의 여달음이 어떻게 나타나고 모음에 따른 변화가 있는가?
- (2) 성조와 억양에 따라 성문과와 성대의 열림비율은 어떻게 변하는가?
- (3) 한국어 파열음의 성문과형의 특징은 무엇인가?
- (4) 모음의 음원파라미터를 변화시켜 합성했을 때 청각적인 인상은 어떤가?

## 2. 성대의 작용과 음원 특성

성대는 보통 전방에서 후방으로 세등분하여 나뉘었을 때, 전방의 두 부분은 막질로 되어있고, 후방의 한 부분은 연골질로 되어 있다. 이 후방의 1/3 부분은 실제 발성시 움직이지 않으며 대체로 전방의 두 부분이 주로 움직인다(Childers 외 1986). 일반적으로 수용되고 있는 성대의 작용은 공기역학적 근육탄성이론(aerodynamic myoelastic theory)으로 설명한다. 성대의 진동은 성대의 질량과 길이에 반비례하고 긴장도에 비례하기 때문에 보다 두껍고 긴 성대는 낮은 주파수를 갖게 되고, 가늘고 짧은 성대는 높은 주파수를 갖게 된다. 덧붙여, 사람마다 감정을 나타낼 때 턱을 내밀어 올리면서 환상연골을 이용하여 성대를 당겨서 팽팽히 만들 수 있는데 이렇게 되면 진동수가 빨라지므로 기본주파수가 높아지게 된다. 이런 변화를 통해 사람의 감정을 표현하는 독특한 억양을 가지게 된다. 따라서, 성인 남성의 기본주파수는 70~200Hz까지 변하며, 여성은 140~400Hz의 범위를, 아이들에서는 180~500Hz로 변한다(Boothroyd 1986).

이러한 성대의 작용에 대해 음성병리학자들은 발성체제에 관심을 갖고 환자의 성대 근접 촬영을 통해 성대의 임상적인 반응을 분석하여 악성종양의 크기나 위치를 알아내거나 장애자들의 발성구조에 대한 근본적인 문제점을 진단하고 치료할 수 있는 방안을 모색해왔다. 이들이 사용했던 대표적인 두 가지 방식으로는, PGG(Photoglottography)와 EGG (Electroglottography)가 있다. PGG를 측정하기 위해서는 광원을 입안이나 인두부에 두고, 성대 밑에 성문을 통과한 빛의 양을 측정하는 광탐지기를 두어 발성할 때 성문이 열린 단면적에 비례하는 빛의 양을 측정한다. 반면, EGG는 한쌍의 전극을 갑상연골의 양쪽 피부에 부착시켜 전류가 한쪽 전극에서 성대 조직을 통과하여 다른 쪽 전극으로 횡단하면서 조직 통로의 접촉면적과 길

이에 따라 저항 값이 달라짐을 이용한다. 즉, 성대가 분리되면 전류가 지나갈 통로가 적어지고, 직선거리가 아닌 후두의 전후방지역을 돌아가기 때문에 저항이 증가한다. 반면, 성대가 접촉하면 보다 많은 전도통로가 생기고 저항 값도 낮아진다. 즉, 조직저항은 성대의 접촉면적에 반비례한다(Childers 외 1986). Fourcin(1981)은 성문의 공간이 아니라 성대의 접촉면적을 나타내기 때문에 laryngogram이라 불렀다. Titze/Talkin(1981)은 더 정밀한 관찰을 통해, 후두에서 측정된 값은 중앙표면 전체의 접촉에 대한 통합된 측정치이며 전선가닥과 같이 좁은 지향성을 보이기보다는 성문주위로 전달되는 모든 요소까지도 반영한다고 지적했다. EGG는 성대의 단함을 다루고 PGG는 성대의 열림을 다루기 때문에 서로 상보적인 측정 방법이라고 할 수 있다(Baer 외 1982).

성문에서 생성되는 음원은 두 가지 중요한 음향적 특징을 갖고 있다. 첫째, 기본주파수 또는 배음 주파수 간격과 들쭉, 주파수에 따른 진폭유형이다. 기본주파수 간격은 성문파의 진동주기가 빠르면 성기계 나타나고 느리면 촘촘하게 나타난다. 보통 대화를 할 때의 발성은 성문파형의 모퉁이 부분이 보다 더 둥글게 그려진다. 따라서 성문스펙트럼은 고주파로 갈수록 옥타브당 약 15dB씩 더 날카롭게 떨어진다. 반면에 아주 강하게 발생할 때는 성문진동이 보다 날카로운 모퉁이 부분이 생겨서 약 9dB 씩 떨어진다(Fant 1973; Rothenberg 1973). 실제 발성 시에는 얼굴표면에 입이 있어서 머리에 의한 방사특성이 실려지게 되므로, 입술밖에 나오는 음성파형은 한 옥타브당 약 6dB씩 떨어지는(-12dB + 6dB = -6dB) 모양으로 나타나고, 후강의 소리흡수 때문에 600~1000Hz 근처지역에 함몰된 부분이 생기기도 한다(Pickett 1987: 67-70).

말소리의 날카로움과 둔탁한 정도를 나타내는 음색은 성대 주름의 진동 방식에 따라 다르다. 다양한 진동 방식의 음원 특성을 반영하여 음성합성을 할 때는 다음과 같은 파라미터를 조정한다. 첫째, 성문파의 진폭(AV), 기본주파수값의 변화( $f_0$ ), 성문이 한번 열렸다 닫히는 한 사이클의 지속시간에서 열린 구간이 차지하는 비율값(opening quotient: Oq), 성문의 열린 부분에서 상승하는 부분의 지속시간과 하강하는 비율(speed quotient: Sq), 성대 전체길이를 따라 동시에 닫히지 않으면 모퉁이 부분이 둥글게 되어 생기는 성문스펙트럼의 기울기(Tilt: Tl) 등이 있다. 이들 값을 변화시켰을 때 Oq는 첫 번째 배음의 상대적인 강도를 결정하는데, 기본주파수값이 높으면 성대의 떨림이 빠르게 일어나므로 빠르게 닫혔다 열리는데 단힘이 보다 더 지속되기 때문에 열림 구간 비율은 낮게 된다. 정상적인 발화는 이 비율이 50% 정도인데 소리를 크게 하면 성대주름이 더 깊게 닫혀있기 때문에 그 보다 낮은 수치 값을 갖게 된다. 즉, 성문하 기압(Psg) 이 높아져서 더 떨

리 열렸다가 더 힘있게 당겨주기 때문이라고 한다(Clark/Yallop 1992). 그 결과 고주파 부분의 진폭이 높아지게 되어 청각적으로는 음성이 강하고 더 뚜렷한 인상을 받게되며, 이와는 반대로 작은 목소리에서는 50% 이상의 Oq값을 갖게된다. 즉, 목청에서 발성 노력이 적어서 성대 주름이 완전히 닫히지는 않고 지속적으로 공기가 흐르기 때문에 숨쉬는 소리로 들린다. 또한, Sq는 스펙트럼 상에 영점으로 나타나고, Tl 변수는 고주파배음을 감폭시킨다. 음성합성 소프트웨어인 SenSyn 1.0는 Lilyencrants/Fant (LF) model로 성문파의 파라미터를 각각 변화시킬 수 있도록 되어 있다.

### 3. 성문파 분석 및 합성실험

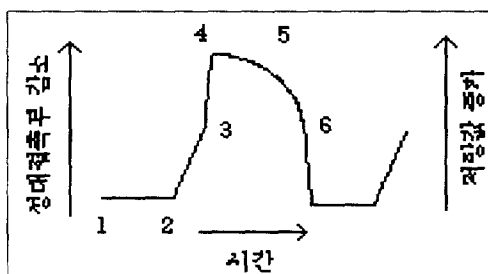
본 연구에서는 경상도 방언을 사용하는 건강하고 정상적인 청력을 가진 37세의 남녀화자 2명이 보통 속도로 발성한 음성과 음원의 특징을 분석했다. 음성 수집과 음원 분석은 IBM PC의 SPG 분석프로그램을 실행하여, 음성을 10kHz의 표본속도로 계수화했다. 동시에 감상연골의 양 측면에 동전 크기 모양의 전극을 부착하여 성대의 여달음에 따른 전류의 흐름을 라링고그래프를 사용하여 수집했다. Fourcin의 라링고그래프에는 분석보드가 있어서 입력된 음성, 음원의 파형, f0와 Oq의 비율이 그래프로 나타나며 이들 파형을 두개씩 하나의 파일로 저장하도록 되어있다. IBM PC의 분석 환경이 고정되어 있어서 섬세한 파형 변화를 볼 수 없기 때문에, 이들 파일을 텍스트로 변환 저장하여 Macintosh LC630의 Signalyze 2.45 음성분석 소프트웨어로 읽어들이 분석했다. 두 컴퓨터 사이의 자료 변환에서 특수문자때문에 파일을 상호 교환하여 처리하는데 어려움이 있고, 또한 모든 파일을 옮기는데 많은 시간이 소요되므로 교내 인터넷을 활용하여 자료를 전달하고 받는 방법을 취했다. 덧붙여, PC의 프로그램에서 분석된 성문파형(Lx)과 음압곡선(Sp)의 저장방식이 한 파일에 번갈아 저장되어있기 때문에 이를 짝수와 홀수 번째로 나누어 따로 파일 명을 부여해서 분리 저장하는 방법을 사용했다. 특히, 파일의 앞부분에 음파의 특성에 대한 문자 자료가 부착되어 있어서 Signalyze 그래프의 성문 파형이 작은 진폭의 신호로 보이므로 이를 제거하여 보다 확대된 파형을 눈문에 실었음을 밝혀둔다.

#### 실험 1

성문파의 정확한 해석은 결과 파형을 해석하는데 매우 중요한 출발점이 된다. 필자는 성문파에 관한 여러 가지 연구논문을 조사하는 과정에서 다양한 성문파의 모양을 볼 수 있었다. 특히, 정반대의 파형을 두고 거기에

맞게 설명을 하다보면 잘못을 범하게 된다. 예를 들어, Childers 외(1986)의 연구에서는 그림 1과 같이 성대의 저항 값의 크기를 나타내는 성문과형을 보이고 있다. 1-2에서는 성대가 최대한 접촉되어 있으며, 2지점에서 성대의 뒤쪽에서 앞으로 지퍼처럼 분리되어 3지점에서 성대의 윗부분이 떨어지고 4지점에서 성대가 완전히 열린다. 5지점에서 성대 닫힘이 시작되어 5-6-1로 아래에서 위로 닫힌다. 즉, 4-6까지가 성대의 닫힘을 나타내는 구간으로 표시되고 나머지 부분은 성대의 열림을 나타내는 것으로 기술했다.

그림 1. 성문과의 저항변화를 보이는 과형



반면, Fourcin의 문헌에서는 성대의 전류의 흐름 정도를 그래프로 그려내는 형태로 기술되어 있다. 따라서, 위의 그림과는 정반대의 모양이 된다. 연구문항 (1)의 해답을 구하기 위해, 숨을 멈출 때는 성대가 닫혀있다는 사실을 이용하여 Fourcin의 라딩고그래프에서 다음과 같은 실험을 해보았다. 화자에게 숨을 깊이 들이쉬운 뒤, 어떤 물건을 들 때처럼 의도적으로 숨을 갑자기 멈추어 허파에 공기를 가두어 두듯이 했다가 곧이어 숨을 [하--] 하면서 내쉬게 했다.

그림 2. 성대의 여닫음에 따른 음압과 성문과의 변화

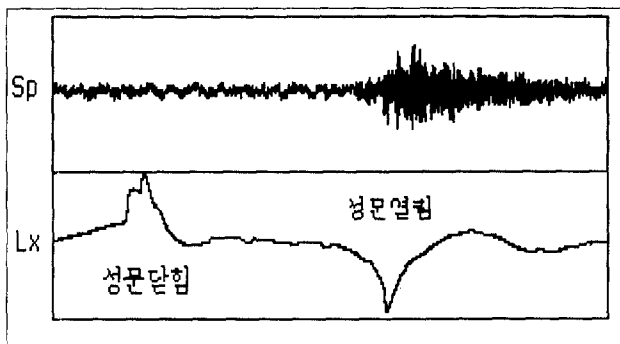
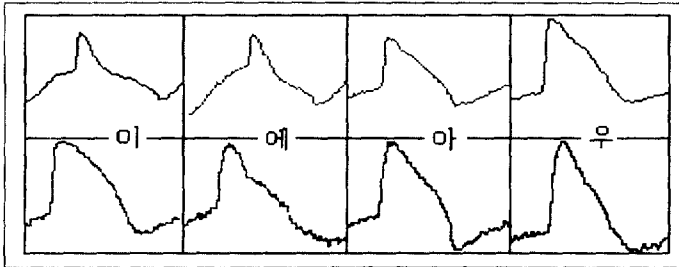


그림 2에서 위는 시각적으로 구분하기 쉽게 다소 증폭한 음압파형(Sp)을 보이고 있고 아래는 성대의 떨림을 나타내는 성문파형(Lx)이다. 처음의 꼭지점은 성대가 서로 붙어서 Fourcin의 라랑고그래프가 성대의 전류가 흐르는 정도를 나타내고 있음을 확인할 수 있다. 성대가 서로 맞붙어 있으면 전류가 짧은 거리에서 바로 전달되고 또한 많은 가닥의 통로가 생성되어 높은 전류가 흐르게 된다. 반면, 성대가 서로 떨어져 있으면 전류가 통하는 거리도 멀고 또한 양도 줄어들게 된다. 본 논문에서는 이러한 기준점을 염두에 두고 여러 가지 실험에 임하고자 한다.

다음으로는 모음에 따라 성문파는 달리 나타나는가? 의 질문에 대한 답을 구하기 위해 남녀 화자가 발성한 모음 [이 에 아 우]의 총모음지속시간 중 1/3에 해당하는 시간점의 성문파형을 각각 하나씩 그림 3에 나타내 보았다. 턱의 움직임이나 혀의 전후 움직임에 따른 성대의 길이나 동작방법에 약간의 변화가 있을 것으로 추정되나 실제 파형에서는 뚜렷한 차이가 없는 모양을 보이고 있다.

그림 3. 남녀화자의 모음의 성문파 모양



남성화자에서 [이 에]의 단히는 부분이 날카로운 삼각형을 그리고 있으나 이는 강하게 발음한 것이고 후반부에서는 [아 우]와 비슷한 모양을 하고 있었다. 여성의 파형은 열림 부분에서 약간의 소음이 나타나며 파형 전체에서 오르내림이 매우 심한 모습을 보이고 있다. 이러한 경향은 본 실험과 별도로 5세 남자 어린이의 음원파형에서도 소음과 같은 파형을 보였고 특히, 성대의 여단음이 분명히 나타나지 않는 일반 정현파의 모양을 보였다.  $f_0$ 가 남성화자에서는 평균 127Hz(Oq 평균 44%)이고 여성화자에서는 평균 217Hz(Oq 평균 36%)였다. 그림 3에서 보면 남녀화자에 따른 차이가 약간 있으나 전체적으로는 비슷한 모양을 하고 있으며 전설 또는 후설성에 따른 규칙적인 변화유형을 보이지 않는다.

실험 2

경상도 방언에서 [아이]는 high low(HL)의 성조로 흔히 발음되고, [여우]는 LH의 성조로 발생된다. 이러한 성조의 차이가 성문파에서는 어떤 모양으로 나타나는지 실험해 보았다. 그림 3은 남성화자가 두 단어를 발생했을 때 나타나는 성문파에서 전반부와 후반부를 같은 비율로 확대하여 나타낸 그림이다.

그림 4. 성조에 따른 성문파의 변화

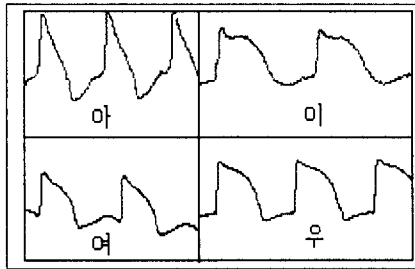
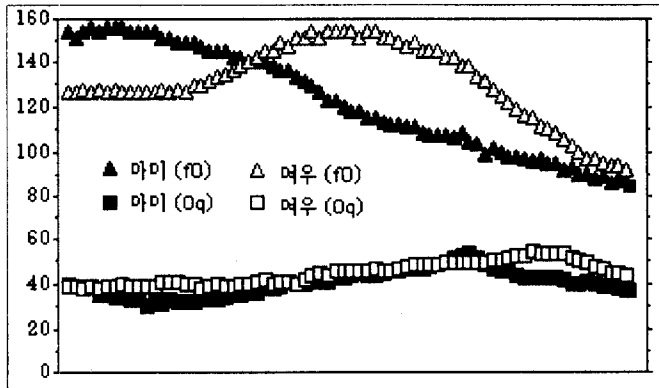


그림 4에서 살펴보면 높은 음조의 [아]와 [우]의 경우에는 성대의 떨림이 빠름을 알 수 있다. 또한 낮은 음조의 [미]와 [여]에서는 성대의 떨림이 상대적으로 적다는 것을 나타내준다. 특히, 두 파형의 모양에서 전반부는 [아]의 경우에 모서리 부분이 날카롭게 만들어지는데 반해 후반부의 [미]와 [우]의 경우에는 보다 부드러운 모서리를 형성함을 보여준다. 이들 단어의  $f_0$ 와  $Oq$ 의 특징을 그래프로 나타내보면 그림 5와 같다.

그림 5. [아이] 와 [여우]의  $f_0$ 와  $Oq$ 의 변화(y축은 주파수(Hz)를 나타냄)



첫 음절의 성조가 높은 [아이]의 경우에는  $f_0$ 가 153에서 100Hz 이하로 변하고 성대의 열림비율은 약 33%에서 46%로 변했다. 두 번째 음절에 강세가 주어진 [여우]에서는  $f_0$ 가 120에서 150Hz로 변하고  $O_q$ 는 약 36%에서 46%로 변했다. 따라서, 성대의 떨림 수를 갖게 하여 강한 음조를 만들고 성조의 차이를 보인다고 할 수 있다. 반면,  $O_q$ 에서는 두 단어간의 차이가 별로 없어 보인다. 즉 전반부에서는 약간 내려갔다가 후반부에 약간 올라갔다가 내려오는 비슷한 패턴을 보이고 있다.

만약 음조의 변화가 많을 때는 어떠한 성문파형을 나타낼 것인가? 이번에는 /다다다다/의 네음절을 음악의 4음계인 도미솔도의 음높이로 발성했을 때 어떠한 성문파의 변화가 있는가를 조사해 보았다. 남녀화자가 발성한 음성파형 중에서 다섯 번째 주기의 성대 여담음에서부터 음압파와 성문파 세 개씩 그림 6에 나타내 보았다.

그림 6. 남녀의 음압파와 성문파의 변화

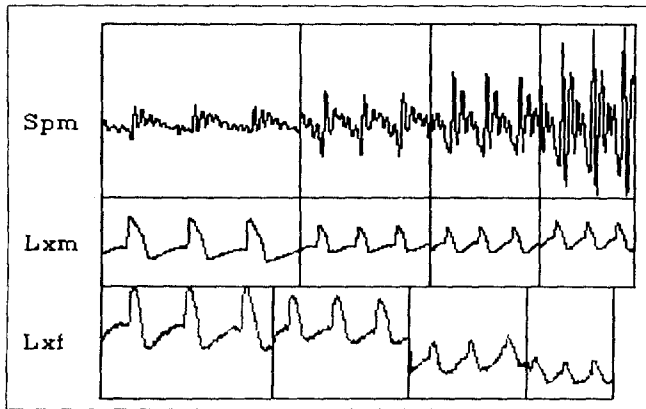


그림 6에서 살펴보면 남성화자(Lxm)의 낮은 음 도에서는 성문파가 드물게 배치되어 있으나 높은 음 도에서는 보다 잦은 파형을 보이고 있다. 즉, 기본주파수가 점점 더 높아가고 있음을 알 수 있다. 실제, 낮은 도음에서는 기본주파수값이 평균 110Hz( $O_q$ 의 평균은 32%)이고, 미음에서 높은 도음까지, 145Hz(31%), 175Hz(33%), 220Hz(31%)로 변화되어 가고 있다. 반면 여성화자(Lxf)의 경우에는 성문파가 점점 더 잦아지는데 음파의 단헨 부분에 소음처럼 떨림이 많이 있고 또 진류의 흐름도 낮은 음에서는 높낮이가 크나 높은 음에서는 좁은 진폭을 보여 성대의 떨림이 매우 빠르게 일어나고 있음을 보여준다.  $f_0$ 와  $O_q$ 의 변화는 각각 평균 193Hz(29%), 247Hz(27%), 291Hz(26%), 383Hz(29%)이다.  $O_q$ 에서는 특별한 변화를 보이지 않는다. 남



녀 화자 모두에서 한 옥타브 높은 도음은 낮은 도음의 두 배에 해당하는 주파수 값을 갖고 있다. 남성화자의 음압곡선(Spm)에서는 진폭이 점점 더 증가하고 있었으나 여성화자에서는 약간의 증가추세만 보였다. 결국, 강세의 차이를 보이는 것은 f0와 진폭 모두 변화를 보이고 있음을 알 수 있다. 앞으로 이 두 가지 파라미터 중 어느 것이 청각에 더 중요한 영향을 미치는가에 대한 합성실험이 필요하다. 성문파형의 모양에서는 낮은 도음보다 높은 도음에서 모퉁이가 더 날카롭게 변화하고 있음을 보이고 있으나 실제 상대 열림비율을 조사해 보면 앞서 제시한 이음절 단어의 파형과 거의 차이가 없는 약간 상승하다가 내려오는 유형을 남녀 파형 모두에서 발견할 수 있었다.

다음으로 억양에 따른 성문파형의 특징을 알아보기로 한다. “이것은 책입니다”와 “이것은 책입니까?”의 두 문장을 발성했을 때 성문파형을 살펴보기로 하자. 먼저 전체적인 기본주파수와 억양곡선을 살펴보면 그림 7과 같다.

그림 7. 서술문과 의문문에서의 성문파형의 변화  
(y축은 주파수(Hz)를 나타냄)

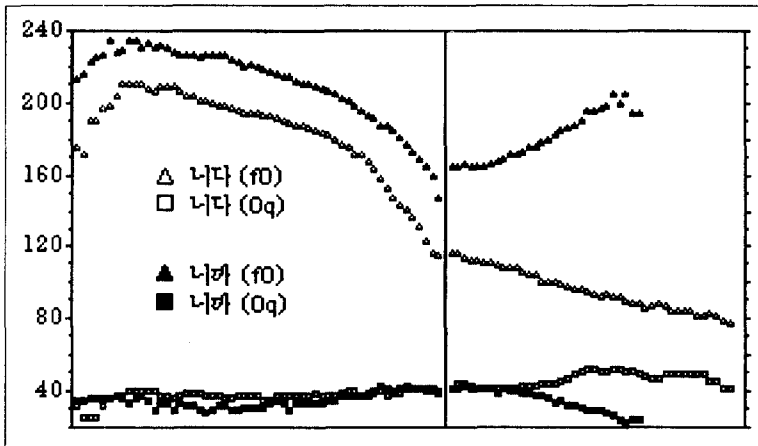


그림 7에서 전반부는 [니]에 해당하고 후반부는 각각 [다]와 [까]에 해당한다. [니] 부분의 상대적인 기본주파수 패턴은 동일하다고 볼 수 있다. 반면 [다]와 [까]는 기본주파수에서 상호 반대되는 모양을 하고 있다. 특히, 성문파형의 Q0에서는 의문문에서는 하강하는 모양을 보인 반면에 서술문에서는 약간 상승하는 모양을 보이고 있다. 즉, 성대의 노력과 열림구간의 비율이 마지막 음절에서는 상호 반비례하는 경향을 보이고 있다. [다]부분

에서의 기본주파수의 평균은 180Hz(표준편차[s.d.]는 14Hz)이고 성대열림구간의 비율은 35%(s.d. 7%)이다. 반면, [까] 부분에서의 기본주파수의 평균은 96Hz(s.d. 11Hz)이고 성대열림구간의 비율은 46%(s.d. 4%)이다. 그러나 마지막에서 두 번째 음절에서는 앞서의 음계실험에서도 보았듯이 이러한 반비례관계가 뚜렷이 나타나지 않는다. 심지어  $f_0$ 가 355Hz 이상이 될 때는  $O_q$  측정값에 문제가 있음을 발견했다. 그러면, 이러한 경향이 성문과형에서는 어떤 모양으로 나타날까? 그림 8은 이 문장의 마지막 두 음절의 파형을 나타내준다.

그림 8. [니다]와 [니까]의 성문과형

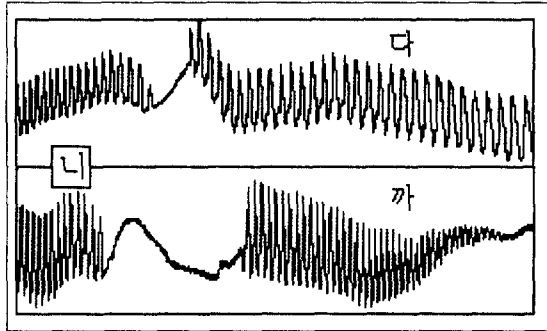


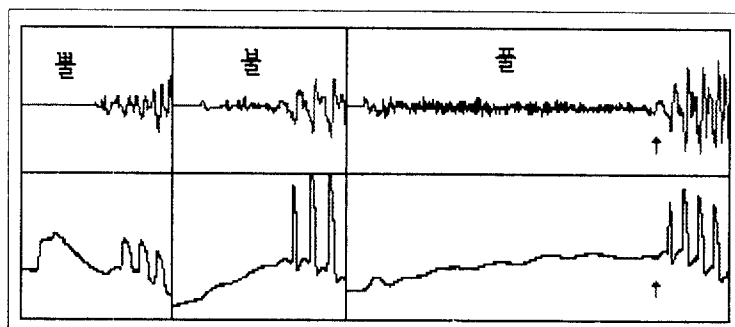
그림 8에서 앞부분은 거의 같은 모양이나 마지막 음절에서 성문과형의 차이를 보이고 있다. 서술문에서는 성대의 떨림이 느리고 모음이 부분이 날카롭지 않으나 의문문에서는 성대의 떨림이 잦고 또한 급격히 진동 폭이 줄어들고, [까]의 전반부에서 아주 날카롭게 모양을 이루다가 마지막에는 둥근 정현파 모양을 하고 있다. 이는 강하게 빨리 발음하고 곧 떨림이 멈추기 때문일 것이다. 덧붙여, [까]로 시작되는 된소리에서는 [니]를 발음하고 성대를 약간 붙였다가 떨어지게 하는 과형에서 보듯이, 성문아래 기압을 높였다가 발성하는 특징을 보이고 있다(그림 2 참조).

### 실험 3

다음은 성문과형과 성대진동개시시간의 관계에 대해 알아보기로 한다. 성대진동개시시간(voice onset time)은 파열음에서 성대가 열린 부분부터 모음의 진동구간이 막 시작되는 시점까지의 지속시간을 말한다(MacKay 1987). Lisker/Abramson(1964)의 연구에서 단어의 첫 음절에 오는 11개 언어의 파열음의 범주를 VOT로 구분할 수 있음을 보였다. 일반적으로 성대

진동의 시작을 음압 곡선에서는 쉽게 구분할 수 없기 때문에 주로 그 범위를 정할 때 성대가 막 열리는 순간부터 규칙적으로 성대가 떨리는 파형을 보이는 위치의 바로 앞 영점을 교차하는 지점까지를 잡고 있다(VOT에 대한 더 상세한 내용은 Yang(1993)을 참고하기 바람). 성대 진동 개시점은 성문파형에서 정확히 포착할 수 있으므로 규칙적으로 성대가 여닫기 시작하는 정확한 시점을 찾아보기 위해 /불 뿔 풀/의 시험음절을 발음하여 음압파형과 비교하여 보기로 한다.

그림 9. 파열음의 종류에 따른 음압파와 성문파



[뿔]에서는 보다 더 많은 압력이 필요하기 때문에 성대의 규칙적인 진동이 개시되기 직전에 닫혔다 열리는 부분이 분명히 나타나있다. [풀]에서는 성대가 터지는 부분에 약간의 융기된 모양이 보인다. [뿔]에서는 VOT의 길이가 매우 짧고 [풀]에서는 아주 길다(각각, 5ms, 110ms). [불]의 발음에서는 아주 부드럽게 시작하는 성문파형을 나타내고 있다. 성대진동개시시간이라고 정의했을 때 성문파형의 정확한 시작점은 위 그림의 작은 화살표(↑)에서 찾을 수 있다. 음압파형에서도 소음이 끝나고 규칙적인 모음의 진동이 시작되는 지점이 성문파형의 시작을 알리는 위치와 일치함을 알 수 있다. 따라서, 성대열림의 시작부는 모음 시작부의 패턴을 추적하여 그 모양이 막 시작되는 지점으로 정의할 수 있을 것이다. 여성화자의 경우에는 뿔의 경우에는 성문이 닫혀서 성문하기압을 높인 뒤 발성하는 모양이 보이거나 나머지 두 종류의 파형에서는 급작스런 폭모양의 뚜렷한 전류의 흐름이 보이지 않았다.

덧붙여, 무성파열음이 모음사이에 오는 경우에는 어떤 성문파를 보이는가? 그림 10은 [아다]를 발음하여 측정된 음압파(Sp)와 성문파(Lx)를 보여준다.

그림 10. [아다]의 음압과 성문과

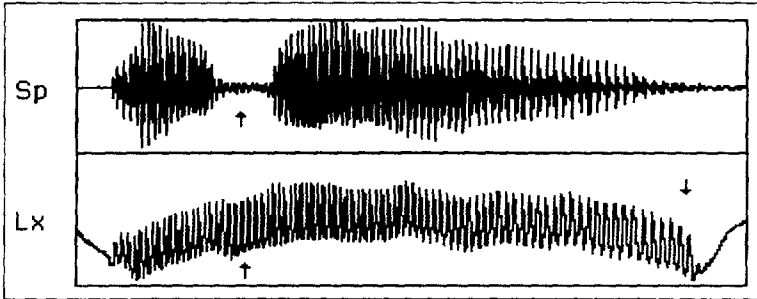


그림 10에서 살펴보면 [아다]의 연속된 빠른 발음에서는 뒤에 오는 무성파열음 [ㄷ]이 유성음화되어서 성문과형에서는 거의 같은 크기로 진동하고 있음을 보이고 있다. 또한 마지막 음절의 끝부분에서는 비록 진폭은 줄어들었지만 성대의 진동은 같은 모양으로 그대로 진행되고 있음을 알 수 있다. 전체적인 윤곽에서는 각 음절의 시작부분이 아주 강하게 발음되었기 때문에 촘촘하게 그려지고, 성문과형도 이와 마찬가지로 질게 나타났다. 실제 음파에서는 추적하기 어려운 유무성 구간에 대한 특징을 성문과형에서 쉽게 나타낼 수 있다는 장점을 이용하면 자모음의 유성성에 대한 중요한 잣대로 사용할 수 있다. 덧붙여, 이런 발성을 천천히 할 때는 성문과형의 진동이 없었으나 약간 빠른 속도에서는 연속적인 성문과형이 나타났다.

실제 그림 9의 [니다] 부분은 성대떨림이 연속되지 않음을 보이고 있다. 따라서, 모음간의 유성음화 규칙은 조건적이 되어야 할 것이다. 즉, 빠른 속도로 발성할 때는 유성음화가 일어나지만 느린 속도에서는 일어나지 않음을 보았다. 앞으로 여러 가지 다른 속도의 발성실험을 통해 유성화의 경계선에 대한 연구가 필요하다.

#### 실험 4

실험 2에서 일반적인 모음의  $Oq$ 는 서서히 상승하다가 끝에서 약간 하강하는 모양을 보았다. 만약  $Oq$ 의 변화가 음질에 영향을 준다면 지금까지의 음성합성 방식에서 50%로 균일하게 고정시켜 합성했던 방식에 수정이 필요할 것이다. 따라서, 본 실험에서는 단모음 [이]와 이중모음 [야]의 음원 파라미터를 여러 가지로 변환하여 청각 실험을 했다. 단모음 [이]와 [야]의 합성용 파라미터는 양병곤(1995, 1996)에 실려있으며, 그 중 음원파라미터만 새로운 Hypertalk 프로그램을 작성하여 자동 변환하여 SenSyn 1.0을 이용하여 합성음을 만들었다. 동시에 Natural Model로 설정된 환경을 LF

Model로 선정하여 합성음을 만들었다. 참고로 LF 모델은 음원의 특성을 미세하게 조정하기 위한 모델임을 밝혀둔다. 청각 실험은 앞서의 두 논문에서 제시한 자동 청각실험방식을 이용했다. 이중모음 [야]의 변형된 파라미터 목록을 예로 들어보면 다음과 같다.

표 1. [야] 합성용 파라미터의 설정표

```

Synthesis specification for file: 'ya.Oq.par'
SenSyn Version 1.0
Total number of waveform samples = 8000
CURRENT CONFIGURATION:
60 parameters
SYM V/C MIN VAL MAX DESCRIPTION
-----
SS C 1 3 3 Source switch
      (1=impulse, 2=natural, 3=LF model)
(중간부분 생략)
Oq V 10 50 99
      Open quotient (voicing open-time/period), in %
(중간부분 생략)
varied Parameters:
time F0 AV Oq F1 F2 F3
  0 1300 26 79 263 2398 3136
  5 1300 27 79 287 2380 3105
(중간부분 생략)
 390 945 19 50 739 1272 3017
 395 934 16 50 739 1272 3021
    
```

표 1에서 살펴보면 Oq가 79에서 50으로 하강했음을 보여주고 있다. 청각 실험을 하기 위해 자연스럽게 들리는 Oq 50%의 합성음을 기준음으로 하고, 최대값과 최소값 사이를 아래 위 각각 이등분하여 하강 및 상승하는 모양으로 자동으로 변하도록 Hypertalk 프로그램으로 만든 파일을 합성음과 기준음을 상호 비교하는 방식을 취해 청각실험을 실시했다.

청각실험은 조용한 연구실에서 경상도 출신의 동의대 영문과 남녀 학생 20명을 선정하여 들려주고 청각적으로 차이가 나는 경우에는 손을 들게 하고, 청각적 인상을 개별적으로 물어 기록했다. 결과는 상당한 Oq와 Sq의 변화에도 특별히 다르다는 청각적 인상을 받지 않았으나 T1의 변화에서는 전체적으로 고주파부분이 줄어들어서 저주파통과 여과기를 거친 듯한 음색을

보였다. 따라서, 변화 폭이 별로 감지되지 않았기 때문에 이번에는 조절할 수 있는 최대한의 변화를 주어 음을 합성했다. 즉 Oq는 처음 99%에서 10%까지 하강(fOq: falling Oq)하거나 그 반대로 상승하는 파일(rOq: rising Oq)을 합성하고, Sq는 500에서 100으로 하강(fSq)하거나 그 반대로 상승하는 파일(rSq)을 합성했다. 덧붙여 T1값은 41에서 0으로 하강(fT1)하거나 그 반대로 상승하는 파일(rT1)을 합성하여 들려주었다. 이중모음과 단모음을 함께 들려주었을 때, 청각적 실험에 참여한 학생들의 반응을 요약하면 다음과 같다. fOq는 뒷부분에서 목갈리는 듯한 소리가 났다. rOq는 별다른 차이가 없었다. fSq는 진폭이 서서히 상승했다 하강하는 음파를 만들고 실제 소리가 커지며 길게 끄는 듯한 느낌을 주었다. rSq는 강하게 시작하여 빨리 끝나는 듯한 인상을 주고 음파의 진폭도 빨리 떨어졌다. 처음 부분이 매우 카랑카랑한 목소리로 들렸다. T1의 변화에서 fT1은 아주 심한 코먹은 소리처럼 들리며 스펙트로그램에서는 F<sub>1</sub>이 강하게 나타나나 F<sub>3</sub>는 거의 나타나지 않을 정도로 진폭이 줄었다. rT1은 약간 코먹은 소리가 되며 진이음 부분이 약간 강하게 들렸다. 덧붙여, 성문파의 Oq를 10% 간격으로 균일하게 변형하여 합성했을 때 Oq가 60%일 때 부드럽고 자연스러운 인상을 주었고, 보다 낮은 값에서는 굵고 탁한 목소리로 들렸고, 70%에서는 약간 명량한 목소리로 들렸다. 100%에서는 자연스럽게 들리지 못한 인상을 주었다. 앞으로 보다 많은 파라미터를 조합하여 변환하는 실험을 할 필요가 있으며 또한 청각적 인상을 정확히 말로 표현하기가 힘들어서 청각실험에 대한 몇 가지 기준을 마련할 필요가 있다.

#### 4. 맺음말

본 연구에서는 라링고그래프를 이용하여 건강한 남녀화자가 발성한 모음, 단어 및 문장의 성문파를 측정하여 음압파와 비교 분석하고, 모음의 음원 파라미터를 변화시켜 합성한 모음의 청각실험을 했다. 연구 결과로는 첫째, 성문파에서 성대의 여담음을 추적할 수 있었고, 모음에서는 성문파의 모양에 어떤 규칙적인 패턴을 보이지 않았다. 둘째, 음조 또는 억양의 변화에 따라 f0와 음압의 변화는 보였으나 Oq에서는 뚜렷한 반비례관계를 보이지 않았다. 셋째, 한국어 과열음에서는 성대를 닫아서 성문하기압을 올린 흔적이 된소리에서 뚜렷이 발견되었고, 성대진동개시점의 위치는 음압파형과 비교했을 때 소음이 끝나고 규칙적인 모음의 패턴을 보이는 지점이 성문파와 일치하며, 음압파에서는 측정하기 힘든 유성화 현상을 성문파를 통해 쉽게 확인할 수 있었다. 아울러, 모음사이의 유성화현상도 발생속도에 따라 조건적임을 관찰했다. 넷째, 음원의 특성을 변화시켜 합성한 모음의

청각실험에서 Oq와 Sq의 변화에는 그다지 뚜렷한 청각적 인상을 보이지 않았고, Tl의 변화는 전반적으로 저주파 여과기를 거친 듯한 인상을 보였다.

본 연구에서는 단일 파라미터의 변화의 비교에 치중하였으나 앞으로 음원과 성도 파라미터의 조합에 의한 청각실험이 필요하다. 또한, 보다 많은 화자의 음원을 수집하여 성문과형에 대한 평균스펙트럼을 조사하여 성대노력에 따른 여러 가지 다른 스펙트럼의 특징을 분석하여 이들에 대한 일반적인 원칙을 조사할 필요가 있다. 또한 개인별 발화 특성을 추출하고 이들의 발음이 여러가지 다른 음색을 띠고 있다는 인상을 주는 이유에 대한 상관관계에 대한 연구가 있어야 할 것이다. 본 연구의 결과가 남녀의 음원에 대한 보다 다각적인 이해에 도움을 주고, 음원에 대한 기본 자료를 제공하여 음성합성과 분석에 널리 활용될 수 있기를 기대한다.

## 참고문헌

- 양병곤. 1995. "합성한 한국어 단모음의 지각실험 연구," 언어 20-3, 127-146.
- 양병곤. 1996. "합성한 한국어 이중모음의 지각실험 연구," 언어 21, 829-843.
- Baer, T., A. Loefqvist and N. S. McGarr. 1982. "Laryngeal Vibrations: A Comparison between High-speed Filming and Glottographic Techniques," *JASA* 73, 1304-1380.
- Boothroyd, A. 1986. *Speech Acoustics and Perception*. Austin, Texas: Pro-Ed.
- Childers, D. G., D. M. Hicks, G. P. Moore and Y. A. Alsaka. 1986. "A Model for Vocal Fold Vibratory Motion, Contact Area, and the Electroglottogram," *JASA* 80, 1309-1320.
- Fant, G. 1960. *Acoustic Theory of Speech Production*. The Hague: Mouton.
- Fant, G. 1973. *Speech Sounds and Features*. Cambridge, Mass: MIT Press.
- Fourcin, A. J. 1981. "Laryngographic Assessment of Phonatory Function," in C. L. Ludlow and M. O. Hart, eds., *Proceedings of the Conference on the Assessment of Vocal Pathology, ASHA Rep. No. 11*, 116-127. Washington, D.C.: American Speech and Hearing Association.
- Lisker, L. and A. S. Abramson. 1964. "A Cross-language Study of Voicing in Initial Stops: Acoustical Measurements," *Word* 20, 384-422.
- MacKay, I. R. A. 1987. *Phonetics and the Science of Speech Production*. Boston: Little, Brown and Co.
- Pickett, J. M. 1987. *The Sounds of Speech Communication*. Austin, Texas: Pro-Ed.
- Rothenberg, M. 1973. "A New Inverse-filtering Technique for Deriving the

- Glottal Air Waveform during Voicing," *JASA* 53, 1632-1645.
- Titze, I. R. and D. Talkin. 1981. "Simulation and Interpretation of Glottographic Waveforms," in E. L. Ludlow and M. O. Hart, eds., *Proceedings of the Conference on the Assessment of Vocal Pathology, ASHA Rep. No. 11*, 48-55. Washington, D.C.: American Speech and Hearing Association.
- Yang, B. 1993. "A Voice Onset Time Comparison of English and Korean Stop Consonants," *Dongueui Nonjip* 20, 41-59.

부산광역시 진구 가야동 산 24번지  
동의대학교 인문대학 영어영문학과  
614-714  
E-mail: bgyang@turtle.dongueui.ac.kr  
Fax: +82-51-890-1209

접수일자: 1996. 10. 22.

게재결정: 1996. 12. 4.