

MRI에 의한 모음의 성도 단면적 측정 및  
면적 변이에 따른 합성 연구\*

Measurement of the vocal tract area of vowels  
By MRI and their synthesis by area variation

양 병 곤\*\*  
(Byunggon Yang)

ABSTRACT

The author collected and compared midsagittal, coronal, coronal oblique, and transversal images of Korean monophthongs /a, i, e, o, u, ɨ, ʌ/ produced by a healthy male speaker using 1.5 T MR, VISION. Area was measured by computer software after tracing the cross-section at different points along the tract. Results showed that the width of the oral and pharyngeal cavities varied compensatorily from each other on the midsagittal dimension. Formant frequency values estimated from the area functions of the seven vowels showed a strong correlation ( $r=0.978$ ) with those analyzed from the spoken vowels. Moreover, almost all of 35 students who listened to the synthesized vowels from area data perceived the synthesized vowels as equivalent to the spoken ones. Movement of constriction points of vowel /u/ with wider lip opening sounded /ɨ/ and led to slight changes in vowel quality. Jaw and tongue movement led to major volume variation with an anatomical limitation. Each corner vowel varied systematically from a somewhat constant volume of the average area. Thus, the author proposed that any simulation studies related to vocal tract area variation should reflect its constant volume. The results may be helpful to verify exact measurement of the vocal tract area through vowel synthesis and a simulation study before having any operation of the vocal tract.

**Keywords:** vowel, MRI, vocal tract, area function, formant synthesis

---

\* 이 논문은 1997년도 한국학술진흥재단의 자유공모과제 연구비에 의하여 연구되었음.

\*\* 동의대학교 인문대학 영어영문학과

## 1. 머리말

사람은 성대를 일정한 속도로 진동하여 생긴 음원을, 조음기관인 혀, 턱, 입술을 움직여, 인강, 구강, 비강으로 구성된 성도 공간을 변형시켜 원하는 모음을 발성한다. 턱의 움직임은 구강의 공간을 넓히거나 좁히는 역할을 하고 혀의 움직임은 구강의 공간에서도 앞 뒤 중간의 좁힘점을 이동시킴으로써 그 위치에 따라 전설, 중설, 후설모음으로 구분한다. 이 때 체적이 일정한 혀의 움직임은 인강과 구강의 면적을 다양하게 바꾸어 준다. 입술 모양에 따라 모음은 앞으로 등글게 하여 내밀면 원순모음이 되고, 평평하게 하면 평순모음, 그리고 모음 /아/와 같은 비원순 모음이 있다. 이렇게 성도의 공간을 적절히 조절함으로써 거기서 발생하는 공명주파수의 조합이 바뀌어지고, 이에 따른 모음의 청각적 인상도 달라지게 된다. 결국, 성도의 단면 공간에 따라 모음이 달라지기 때문에 성도 단면적에 대한 수량적인 연구를 통해 모음의 특징과 기능을 설명하거나 비정상적인 또는 수술 전·후에 구강과 인강의 단면적을 보상하여 보다 자연스러운 발성을 하는데 필요한 정보를 구할 수 있다.

본 연구의 목적은 첫째, 최근에 개발된 MRI 장비를 이용하여 건강한 화자가 국어 단모음을 발성할 때의 성도의 정중단면의 모양을 구하여 각 모음의 특징을 관찰해 보고, 둘째, 성문으로부터 입술에 이르는 성도의 각 지점의 단면적 값을 측정하여 비교하고, 셋째, 단면적 값에서 음향적 공명주파수를 추정해내어 실제 음성에서 추출한 포먼트와의 상관도를 구하고, 또한 Klatt의 포먼트 음성합성기인 SenSyn을 이용하여 추정한 포먼트 값을 입력하여 합성했을 때 청각적으로 어떠한 소리로 들리는지를 실험해 본다. 마지막으로, 모음 사각도의 모퉁이에 해당하는 /아, 우, 이/의 단면적 값의 평균을 구하고 이 기준선에서 각각의 모음이 어떻게 구강과 인강의 공간을 좁히거나 늘려서 원하는 모음을 생성하는가를 조사해본다. 이러한 연구는 국어발음의 조음학적인 특성에 대한 자료를 수집할 수 있는 중요한 연구이며 동시에 성도 공간에서 구강이나 인강의 면적 변화가 어떠한 음향적인 변화를 가져오는지 미리 예측하여 임상적으로 활용하는 데도 도움이 될 수 있을 것이다.

## 2. 모음의 성도 단면적 연구

지금까지 모음을 발성할 때의 성도 단면적에 대한 다양한 연구가 있었지만, 입체적인 모양에 대한 연구는 최근의 자기공명영상 (Magnetic Resonance Image: MRI) 기법에 의해 보다 완전한 형태를 지향할 수 있게 되었다. X선에 의한 단층촬영(Computerized Tomography: CT) 영상은 구강과 인강의 정확한 모양을 나타낼 수 있음에도 불구하고 구강과 인강의 경계지점인 인두강 지역의 촬영에 어려움이 많았다. 왜냐하면, CT는 횡단면만을 얻을 수 있고, 이 정보를 적절히 재처리하여

성도의 수평 또는 수직 단면 영상을 구할 수는 있지만, 인두강 지점을 따라 비스듬한 각도의 영상을 얻는데는 한계가 있기 때문이다. 덧붙여, CT는 X선의 투과율이 많으면 검게 나타나고 뼈 조직과 같이 밀도가 높은 경우는 투과율이 적어 희게 나타나지만, MRI는 신체의 조직내의 수소원자와 양성자의 수량에 따라 반응이 나타나므로 성도와 같은 연한 조직내의 경계를 살펴보는 데 편리하다. 보통 성도의 단면을 X선 촬영으로 할 때 조영제나 혀에 가는 쇠사슬을 드리워서 연한 조직의 경계선을 찾기도 하지만 구개부나 인두강 지역은 여러 구조가 한꺼번에 겹쳐져 나타나기 때문에 선명한 영상을 얻기가 어려웠다. 또한 피험자를 방사능에 자주 노출시키면 신체에 여러 가지 장애를 초래할 위험의 소지도 있다. 이에 비해 MRI는 다양한 각도에서 촬영이 가능하며 또한 방사능에 의한 위험성도 없기 때문에 최근의 성도 단면적 측정에 자주 이용되고 있다(Baers 외 1987, 1991; Foldvick 외 1991; Matsumura and Sugira 1990; Moore, 1992; Yang and Kasuya 1994, 1996). 단점으로 아직까지 해상도가 낮고 연속적인 동작을 나타낼 만큼 개선이 되지는 않았지만, 앞으로 조음 동작을 실시간으로 측정할 수 있는 기기도 나오게 될 것이다.

모음을 발성할 때 성도 단면적에 대한 연구로는 스웨덴어를 대상으로 한 Fant(1960)의 성도 단면적 측정과 단면 모양에 따른 공명주파수의 관계를 규명한 연구가 있다. 이 연구에서는 좁힘점 이론으로 요약되는 다양한 결과를 보여주고 있다. 성도의 좁힘점 이론은 성도의 어느 부분이 좁혀지는가에 따라 공명하는 포먼트 주파수가 달라지는 규칙을 나타내 준다. Wood(1979)는 문헌에서 수집한 X선 촬영을 한 38세트의 모음 조음에서 성도 면적을 구했다. 또한, 새로운 X선 활동사진을 영어, 아라비아어에서 구했는데, 경구개, 연구개, 인강 윗부분, 인강 아래 부분의 네 지점이 모음의 자질을 정의하는데 적절함을 관찰했다. 특히, 이런 모음 조음이 좁힘점의 위치에 따라 선형적으로 변하다가도 비선형적으로 비약하는 특징이 있음을 관찰했고 주어진 음소환경이 달라지더라도 구강에서의 혀좁힘 위치가 일정함을 지적했다. 또한 성도 가운데서도 Stevens(1972)가 지적한 비약적인 영역(quantal region)을 피하여 네 개의 좁힘점이 나타나고, 구강부위에서 언어마다 전반부 또는 후반부를 선호하는 경향이 있다고 했다. Baer 외(1991)에서는 MRI를 이용하여 음성 컴퓨터 모델을 정의하기 위한 기본 자료를 수집했는데, 두 명의 남성화자에게 네 개의 모음 /i,a,u,x/를 발성하게 하고 이를 MRI 이미지로 단면적을 측정하고 그 값에 상응하는 포먼트 값을 구한 뒤 이를 발성한 모음 포먼트 분석값과 비교했다. 포먼트 값에 차이가 일부 보였으나 지각실험에서는 거의 동일한 모음으로 들었다는 결과가 나왔다. Demolin 외(1996)에서는 남자 2명, 여자 2명의 건강한 피험자가 발성한 불어 구강모음 /i,e,ε,a,y,ø,œ,ɔ,o,u/에 대한 성도 단면적을 MRI로 측정했는데 남녀 성도의 길이가 차이가 나기 때문에 4명 사이의 규칙성을 볼 수는 없었지만 구강과 인강의 경계점인 목젓지점에서 후설모음의 좁힘점이 생기고, 정중위치에서 혀의 함몰의 크기가 전설모음에는 중요하나 후설모음에는 중요하지 않다는 사실을

지적하고 있다. 또한 인강 폭이나 입술 영역에 따라 모음을 구분했는데 이를 제대로 나타내기 어렵다는 점과 피부와 공기의 경계선을 찾는 데 어려움이 있다고 했다. Yang과 Kasuya(1996)는 일본어 화자의 성도를 연구했는데 세 명의 남성과 세 명의 여성화자가 발성한 일본어 모음 다섯 개의 3차원 성도를 MRI로 재구성했다. 이어서, 이들을 구강, 인강, 후강 지역으로 나누어 남녀의 성도 길이를 균일, 비균일, 축척 방식을 통해 정규화(normalization)시켜서 포먼트 값을 비교했다. 그 결과는 첫째, 남녀 성도의 신체적 차원이 연속적으로 분포되어 있고, 비균일 척도는 통계적으로 남녀 차이가 유의하지 않고 인강/구강비와 성도 길이가 개인 차이를 구분하는 중요한 요소임을 지적했다. 이들은 컴퓨터 프로그램을 이용하여 성도 공간과 피부의 경계선을 자동 추정할 때 치아가 나타나지 않기 때문에 구강부분의 분을 떠서 보정했으며 치아를 보정한 단면적에서 구한 포먼트 값과 보정하지 않는 경우의 값의 차이는 크지 않았음을 밝혔다. Gracco 외(1994)는 정상인과 비정상인의 성도 영상을 3차원으로 분석했다. 성문 상부에 있는 설근, 후두개, 가성대를 제거하고 후두만을 남겨둔 환자의 경우를 조사했다. 수술 전의 성도의 3차원 영상에서 구한 공명주파수와 수술 후의 모양에서 구한 값이 실제 발성에서 구한 음성의 공명주파수와 매우 높은 상관도를 보였기 때문에 수술 전에 성도의 모양을 임의로 변형했을 때 음질의 변화를 예측할 수 있음을 보였다. 국어 모음 성도의 단면 자료로는 김영송(1981)이 있는데, 성도 정중단면을 X선 촬영으로 상세히 밝혀놓았으나, 성도 단면적 값을 제시하지 않아서 단면적에 의한 입체적인 좁힘점 위치를 잡는데 어려움이 있었다. 지금까지 살펴본 연구에서 성도의 단면적은 포먼트 값과 밀접한 관계를 가지고 있으며 화자는 성도의 단면적을 혀와 턱을 이용해 변화시켜 원하는 모음을 발성함을 알 수 있다.

### 3. 연구 방법

이 연구의 대상으로는 동일한 모양으로 모음을 발음할 수 있도록 훈련을 하고, 보철에 의한 영상 신호의 간섭이 전혀 없는 건강한 남성화자 1명을 대상으로 했다. 자기공명영상은 1.5 T MR, VISION(Siemens Co.)을 이용하여 피험자가 누운 자세에서 국어 단모음 /아, 어, 에, 오, 우, 으, 이/를 발성하고 있을 때 촬영하여 구했다. 발성은 여러 번 연습을 통해 촬영할 때 입모양이 일정하게 유지되도록 했다. 촬영할 때의 펄스 시퀀스는 FLASH (fast low angle shot) 2D, TR (repetition time) 100 ms, TE (echo time) 5 ms, Flip angle 30 Matrix 98X128 pixels, FOV (field of view) 130 mm, NEX (number of excitement) 2, Slice Thickness 10 mm이었고, 촬영시간은 20초였다. 촬영시간을 짧게 한 것은 피험자가 모음을 발성할 때 성도 단면 모양을 일정하게 유지할 수 있는 시간이기 때문이다. 촬영시간을 늘리면 영상이 매우 정밀하게 나타나지만 조음 동작을 동일하게 유지하기가 어려워지게 되는 단점이 있다. 정중단면은 이비인후과 의사가 직접 영상을 보면서 성도의 정중

양으로 맞추었고 먼저 낮은 해상도의 고속 촬영을 통해 성도의 윤곽을 확인한 뒤 단면적을 구할 부분을 좁힘의 변화가 많은 부분에 하나씩 모두 10개 전후의 직선 모양으로 성도 중심선에 수직이 되도록 하여 지정했다. 이때 단면적을 구하기 위해 설정한 선이 인접한 선과 겹쳐질 때 자기장의 영향으로 영상이 중복되면서 흐려지기 때문에 가능하면 겹쳐지지 않도록 이동시켰다. 성문 부위는 경추 4번과 5번 관절 사이의 연골부 중심과 감상연골의 성도가 가장 앞쪽으로 나온 지점을 연결한 지점으로 했고, 입술부위는 끝부분이 얼굴 전면에 나와 있어서 면적값을 구하기 힘들기 때문에, 양 입술의 중앙을 선택했다. 그림 1은 모음 /아/의 정중단면과 성도 단면적 촬영위치

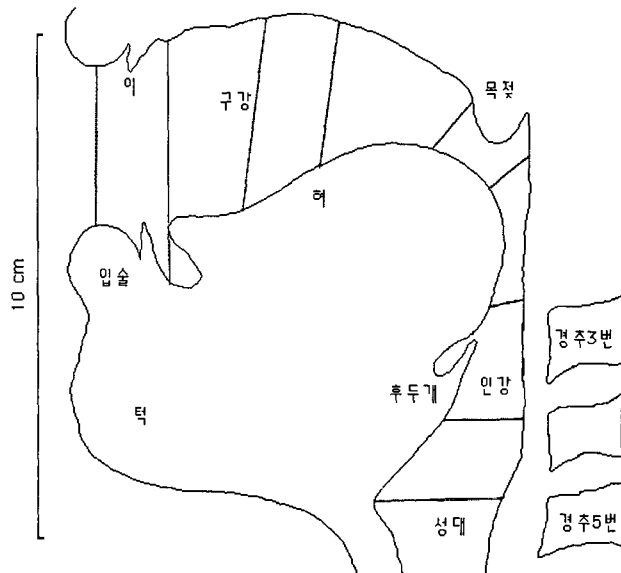


그림 1. 모음 /아/의 정중단면과 성도 단면적 촬영위치

MRI 필름에서는 치아부분이 활성 수소원자가 없으므로 검게 나타나기 때문에 관상단면 CT를 촬영하여 구강에서의 치아의 단면을 확인하고 비디오 영상을 찍어 입술과 치아의 높이를 추정하여 보정하는 방법을 사용했다. 성도 단면적은 성도의 해부학적 구조를 세심히 관찰하여 알고 있는 저자가 MRI 영상 필름을 유리창에 고정시키고 투명용지를 덧붙인 뒤 투영된 성도의 윤곽을 따라 연필로 그렸으며, 의심나는 부분은 이비인후과 의사의 조언을 구했다. 그려진 정중단면과 성도 단면적을 기준 눈금을 이용하여 실측치의 비율로 확대하여 PowerPC에 U-MAX Scanner로 입력했다. 이어서 실측에 가까운 성도의 단면 공간을 Photoshop3.0에서 매직봉 기능을 이용하여 선택한 뒤 검게 채워 복사하고 이를 Area Properties(v.3.2)라는 소프트웨어에 입력하여 면적을 구했다. Area Properties는 화면의 픽셀 수를 세어

단면적을 구하는 프로그램으로써 임의의 사각형의 실제 면적을 측정하여 확인했는데 오차가 거의 없었다. 구해진 단면적 값을 이용해 성도 모양을 정하고 여기서 발생하는 포먼트 주파수를 구하는 프로그램은 Liljencrants(1986)의 Formfrek을 이용했다. 이 프로그램은 시스템 방사 위상각에서 영점을 찾는 원리로 포먼트 값을 구하고 입술 끝점의 방사효과를 수정한다. 포먼트 값은 맨 아래부터 다섯 개의 음향적 공명주파수( $F_1, F_2, F_3, F_4, F_5$ ) 값을 구했다.

실제 발생된 음성의 포먼트 값은 피험자가 MRI 촬영 때와 같이 누운 자세에서 발생한 7개의 단모음을 PC에 20,000 Hz의 표본속도로 입력하고 이를 SFS(Speech Filing System) 음성분석소프트웨어를 이용하여 스펙트로그램 가운데 가장 안정된 모음 구간의 포먼트 값을 구했다. 모음의 발생기간 중모음 포먼트 값은 계속 변하기 때문에 전체 모음구간에서 약 1/3지점에 해당하는 스펙트럼을 기준으로 삼았다.

음성합성은 보다 정밀하게 음성 파라미터를 조정할 수 있는 포먼트 합성 방식의 SenSyn1.0(Sensimetrics)을 이용했다. 단모음 합성에 필요한 진폭, 지속시간, 피치 변화는 양병곤(1995)의 단모음 합성연구의 파라미터를 사용하고 포먼트 값은 MRI의 단면적 값에서 Formfrek을 이용해 추정된 값을 입력했다. 청취실험은 합성된 음성을 녹음테이프에 저장하여 조용한 교실에서 학생들에게 일정한 순서 없이 적당한 음량으로 들려주고, 응답지의 해당 번호에 제시된 7개모음 가운데 하나에 동그라미를 치게 했다.

#### 4. 분석 및 토론

##### 4.1 모음의 정중단면 비교

먼저 정중단면에서 성도의 길이를 비교해보았다. 측정 방법은 Hypertalk로 프로그램을 하여 그림 자료에서 성문에서 입술 끝지점까지 성도의 중심선을 따라 마우스를 움직여 가면서 점을 찍었고 이 점까지의 거리를 구해서 계속 더해 간 뒤, 인강과 구강을 지나 입술까지의 성도의 중심선의 길이를 측정했다(양병곤 외 1998 참고). 성도 단면적을 입술 끝에서는 측정하기 어렵기 때문에 입술의 중양을 선택해서 단면적을 구했으나, 성도의 실제 측정 길이는 양 입술에서 약간 더 나아가 끝부분을 서로 연결한 부분까지였다. 덧붙여, 이 과정에서 후두개와 목젓부위, 원순모음에서 혀가 뒤로 당겨져 생긴 혀끝부위와 아랫니 사이에 생기는 홈 등에서 성도 중심선이 급작스레 변해야 하지만, 그 차이는 크지 않기 때문에 자연스럽게 연결되는 중심선 모양을 그려 성도 길이를 재었다. 이렇게 측정한 길이를 비교해보면, 원순모음 /우/가 입술을 둥글게 내밀어 발음하기 때문에 18 cm으로 가장 길고, 모음 /에/와 /으/가 16.2 cm로 짧게 나타났다(표 1, 표 2 참조). 7개 모음의 평균값은 16.9 cm이다. 따라서 모음 발음시 성도의 중심선을 따라 측정한 길이는 원순모음 /우/를 제외하고는 큰 차이가 없음을 알 수 있다.

이어서 모음 정중단면을 각각 조음 위치에 따라 나눠서 비교해 보았다. 그림 2

는 국어모음 삼각도에서 모퉁이에 해당하는 모음 /아, 우, 이/를 동시에 나타냈다. 이 때 그림 1에서 보이는 경추 3~5번과 인강벽을 일치시켰다. MRI로 촬영할 때 피험자의 두경부를 고정시켜서 상하로만 움직일 수 있게 했고, 정중단면은 동일한 처리과정을 통해 나온 결과이기 때문에 상대적인 비교는 의미가 있을 것이다.



그림 2. 모음 /아, 우, 이/의 정중단면 비교

전설모음 /이/는 구강부분에서 혀와 입천장이 매우 가까이 근접했고, 일정한 체적을 가진 혀의 일부가 이동했기 때문에 반대쪽인 인강 부분은 아주 넓은 공간이 생겼다. 모음 /아/는 이와는 반대 모양을 하고 있는데 혀가 인강 쪽으로 많이 이동하여 인강이 좁아지고 구강이 넓다. 턱을 내려 입안의 공간을 확보하기 위해 성문의 위치도 약 0.5 cm 가량 아래로 내려가 있다. 아래 입술은 약 2.8 cm 정도 이동을 했는데 반해, 인강 중심부의 폭은 2 cm의 차이를 보여 약 0.8 cm 정도의 차이가 난다. 후설모음 /우/는 전설모음 /이/에 비해 입술 부위가 1 cm 더 앞쪽으로 밀려졌으며 아래 입술은 이보다 약 0.6 cm로 약간 적게 앞쪽으로 튀어나왔다. 전체 성도 길이로서는 후설 부분이 목젖에 가깝게 접근하여 더 길어져 평순모음 /으/에 비해 1.8 cm 이상 길어진 결과를 가져왔다. /우/의 혀 모양은 혀끝을 뒤로 밀어서 타원 모양으로 만들어 모음 /이/에 비해 구강 앞부분이 넓어진 공간을 확보하고 목젖과 턱을 이은 선에서 겹쳐보면, 인강 부분은 거의 비슷한 모양의 구조로 되어있다. /아/와 /우/를 비교해 보면 피험자는 턱을 아래로 내림과 동시에 구개의 윗부분도 울려서 발음하고 있다. 윗입술의 이동은 약 1.8 cm이고 아랫입술의 이동거리는 1.5 cm이다. 혀의 높이 차이는 1.2 cm, 후설부분의 차이는 0.7 cm가 된다. 전설모

음과 후설모음을 발음할 때 화자는 윗입술과 혀의 앞쪽 부분을 더 활용하여 공간을 확보하고 있음을 알 수 있다.

나머지 모음의 정중단면은 지면상 제시하지 못하나 다음과 같은 연구 결과가 나타났다. 모음 /으/는 /이/와 비슷한 입술과 턱 모양을 하고 있지만 혀의 전반부가 약 1.5 cm 좁아지고 그만큼 혀가 뒤로 이동하여 인강 상부가 좁아진 모양으로 발생된다. 모음 /으/와 /우/의 인강 부분은 거의 동일하나 구강 전반부인 혀의 끝부분의 동작에서 차이가 있다. 즉, 모음 /우/는 뒤로 최대한 수축한데 반해 모음 /으/는 혀끝을 얇게 펼쳐서 아랫니 쪽으로 당겨낸 점이 다르다. 원순모음 /우/와 /오/의 발음 모양을 비교해 보면 입술이 약 0.5 cm 정도 더 나온 /우/가 성도의 길이는 길어졌으나, 상대적으로 입술을 내밀기 위해 혀의 뒷부분을 목젖 쪽으로 당긴 동작에서 혀끝부분이 아랫니 쪽에서 멀어져서 구강 전면부가 넓게 나타난다. 모음 /어/와 /아/를 비교해 보면 인강 부위는 비슷하나 구강 전면 부분에서 턱의 벌림 정도에 차이가 나는데 양 입술의 벌림 정도는 약 2 cm가 되며, 이에 따라 혀의 중앙부분의 상하 폭이 가장 크다. /오/와 /에/의 차이는 혀 앞부분과 중앙부분이 내려가고 아랫입술이 약 0.7 cm정도 내리고 발음하고 있으며 인강 부위는 거의 차이가 없음을 알 수 있다. /어/와 /에/를 비교해보면 혀 중앙부분이 2 cm 내려가고 구강 중앙부분이 넓어진 반면 인강 상부는 약 1.5 cm 더 좁아지게 되었다. 모음 /으/는 후설 부위를 2 cm 들어올림으로써 목젖부위에 가까운 구강부분이 좁아진 특징을 보여주고 이에 따라 인강 부분이 0.7 cm정도 더 넓어지게 되었다.

지금까지 면적이 아닌 정중단면에서 모음의 구강 및 인강, 입술 등 성도 각 지점의 폭을 비교해 보았다. 원순모음 /우/는 입술을 더 길게 내밀어 성도 길이가 평순모음보다 길어졌으며 모음 /오/도 약간 내밀었으나 턱을 내렸기 때문에 혀와 입천장의 공간이 넓어져 성도 길이에서는 차이가 별로 없었다. 모음의 조음 동작은 주로 혀의 전후 이동에 의해 이뤄지고 있다. 특히, 혀의 체적은 일정하기 때문에 상하로 움직이는 부위가 좁아지거나 넓어짐에 따라 구강과 인강의 폭이 상호 반비례하는 관계를 보여주었다. 턱의 움직임은 성문 부위를 향해 비스듬하게 당겨지거나 밀어내기 때문에 이 동작에 따른 구강이나 인강의 변화는 성문에서 멀어질수록 변화가 크고 가까울수록 변화 폭이 적어짐을 알 수 있다. 양 입술사이의 폭을 측정해 본 결과는 모음 /아/가 3.1 cm로 가장 길고 /어/와 /아/가 1 cm 정도, /이/와 /으/가 0.4 cm, /오/가 0.5 cm, 우가 0.2 cm로 가장 좁게 벌렸다. 실측치는 MRI 필름을 확대하여 윤곽선을 부드럽게 하기 위해 처리되는 과정에서 다소 줄어들었지만, 동일한 방법으로 처리되었기 때문에 상대적인 비교가 가능하다.

#### 4.2 성도 단면적 비교

실제 음성의 특징은 성도의 길이에 영향을 받지만, 앞서 성도의 길이 비교에서 살펴봤듯이 사람의 성도의 길이는 입술을 이용해서 늘리는 방법 외에는 큰 변화를



기대할 수 없다. 성도 길이에서 1 cm 차이가 제 3포먼트 주파수로는 약 200 Hz 정도 차이가 나고(Yang 1996: 252), 단모음 합성지각실험에서 이러한 차이는 청각적으로 별로 큰 차이를 못 느낀다는 연구 결과(양병곤 1995)를 고려해 볼 때 성도의 길이에 의한 차이는 별로 큰 영향을 미치지 못한다고 할 수 있다. 그러나 조음기관인 혀와 턱을 움직여 입안의 단면적을 자유롭게 변화시킬 수 있기 때문에 다양한 음질의 모음을 구사할 수 있다. 피험자가 국어 단모음을 조음할 때의 성도 단면적 값은 표 1, 2와 같다.

표 1. 모음 /아, 에, 이/의 성문으로부터의 거리(cm)와 단면적(cm<sup>2</sup>)

아		에		이	
거리	단면적	거리	단면적	거리	단면적
0	2.10	0	2.88	0	2.60
1.75	1.68	0.87	1.59	3.01	2.71
4.12	2.75	3.19	5.85	4.56	7.19
6.66	1.87	4.80	6.09	7.27	10.04
8.23	2.19	7.33	3.89	8.77	4.70
10.90	7.67	8.86	4.04	10.80	2.56
12.40	10.98	10.60	2.80	12.30	0.84
14.20	12.29	12.10	2.52	13.90	0.71
16.70	8.99	13.80	4.05	15.10	0.75
		16.20	8.30	17.10	4.30

표 2. 모음 /오, 우, 어, 으/의 성문으로부터의 거리(cm)와 단면적(cm<sup>2</sup>)

오		우		어		으	
거리	단면적	거리	단면적	거리	단면적	거리	단면적
0	2.53	0	2.99	0	2.78	0	2.16
1.86	3.91	1.65	8.03	2.34	6.59	0.96	2.48
3.87	3.69	3.82	8.51	4.87	2.67	2.96	5.29
5.60	1.80	5.45	5.22	7.25	2.95	4.71	5.87
7.72	1.51	7.42	1.71	8.48	2.24	7.62	1.40
9.37	1.35	9.24	1.49	10.90	6.01	9.18	1.45
11.52	6.91	11.70	3.96	12.20	8.28	11.10	2.39
13.28	10.04	13.60	13.04	13.80	6.82	12.60	2.97
15.01	7.06	15.20	6.10	16.50	7.66	13.90	4.44
17.38	2.94	18.00	0.74			16.20	3.27

성도의 각 지점의 단면적 값의 차이를 모음별로 살펴보면, /이/와 /에/의 차이는 성문에서 약 7.5 cm 부근에서 6 cm<sup>2</sup>, 구강쪽 14 cm 부분에서는 약 3.3 cm<sup>2</sup>의 차이가 있다. 구강 부위의 폭 차이에 비해 인강쪽의 변화가 거의 두 배 정도 나타나고 있다. 모음 /이/와 /으/를 비교해 보면 면적 차이가 나는 곳은 성문으로부터 약 7 cm 지점으로 약 8 cm<sup>2</sup>의 차이가 있다. 이것은 앞서 폭의 차이가 약 1.5 cm임을 생

각해 보면 인강 부위의 폭과 길이의 차이가 매우 큼을 알 수 있다. 성도 단면적에 따라 포먼트 주파수 값이 달라짐을 생각해 볼 때, 정중단면에 의한 모음 비교는 완전하지 못함을 알 수 있다(김영송 1981).

후설모음 /어, 오, 우/는 성문에서 8~9 cm 전후의 위치에 좁힘점이 형성되었고 모음 /아/는 턱을 내림으로써 그보다 약 2 cm 뒤쪽에 좁힘이 일어났다. 이에 반해 전설모음 /이/는 약 14 cm 지점인 구강 전반부에 좁힘이 일어났고, /에/는 12 cm 지점인 구강 중간부분에 좁힘이 있다. 모음 /으/는 8 cm 지점에 좁힘이 일어나 후설모음 /우/와 가깝기 때문에 중설모음으로 분류하기보다는 후설모음의 모양을 하고 있다. 모음 /어/와 /아/는 성문에서 약 14 cm 지점인 혀 끝부분에  $5 \text{ cm}^2$ 의 면적 차이가 난다. /어/는 /아/에 비해 구강 부분과 인강 부분이 거의 동일한 값의 차이를 보인다. 영어 모음에서 제시하는 균일관의 모양과는 다른 독특한 발음이라고 할 수 있다. 모음 /어/와 /우/의 면적 차이는 성문에서 8 cm 떨어진 목젖 부위에서 약  $1 \text{ cm}^2$ 의 차이를 보였는데 이렇게 후설 부분을 내린 결과 구강과 인강 부위가 각각  $3 \text{ cm}^2$  넓어지게 되는 결과를 보였다. 즉, 혀의 일부가 이동하여 넓어지면 그만큼의 면적이 다른 공간에 더해진다.

그러면 성도의 체적은 항상 일정하게 유지되는가? 성문에서 입술까지의 성도의 단면적 값을 모두 합친 체적을 구해 보았다. 측정 과정은 각 모음의 단면적을 나타내는 그래프를 실측치의 크기로 확대하여 단면적 부분을 Area Properties에 입력하여 체적을 구했다. 그림 3은 이와 같이 구한 성도 공간 단면적의 총합인 체적을 크기 순으로 나타내 준다.

이 값을 보면 성도 공간의 체적이  $78.8 \text{ cm}^3$ 에서  $103.4 \text{ cm}^3$ 까지 다양한데 턱을 크게 벌리지 않고 발음하는 고모음 /오, 우, 이/가 체적이 작고 저모음은 크다. 그러나 대체로 체적의 변화를 가져오는 원인은 턱과 혀의 움직임이 조합되어 나타남을 알 수 있다. 후설 저모음인 /아/와 전설중모음인 /에/는 거의 비슷한 체적을 나타내고 있으며, 모음 /으/가 가장 큰 체적을 차지하고 있다.

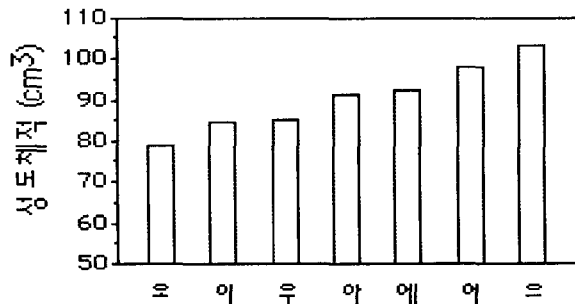


그림 3. 각 모음의 성도 체적

앞서 모음 /아/와 /이/의 성도 인강과 구강의 정중단면의 모양이 서로 반대로 되어 있었는데 성문에서 8 cm 거리를 기준으로 인강과 구강의 체적을 구해 본 결과 모음 /아/에서는 인강부가 17 cm<sup>3</sup>이고 구강부는 이것의 약 4배인 74 cm<sup>3</sup>이었는데 반해, 모음 /이/에서는 인강부가 58 cm<sup>3</sup>이고 이것의 반인 28 cm<sup>3</sup>가 구강부를 차지하고 있다. 두 모음은 체적에서 5 cm<sup>3</sup>의 차이를 보이지만 혀의 움직임은 성도에서도 구강과 인강의 체적비를 변화시키되 전체 체적은 큰 변화가 없음을 증명해 준다. 모음 /우/에서 좁힘지점인 성문에서 9 cm 떨어진 지점을 중심으로 구강부와 인강부를 나누어 보면 양쪽의 체적이 43 cm<sup>3</sup>으로 거의 동일한 체적을 보였다. 혀의 뒷부분을 목젖에 가까이 접근시킴으로써 구강과 인강의 공간이 거의 동일하게 나누어 졌는데, 그만큼 혀의 운동이 체적 조절에 큰 영향을 주고 있음을 알 수 있다.

#### 4.3 단면적 값에 의한 포먼트 값 추정과 면적 변이에 따른 음성 합성

실제 음성에서 SFS를 이용해 측정한 포먼트 값과 단면적 값을 이용해 Formrek으로 추정한 포먼트 주파수 값은 표 3과 같다.

표 3. 음성 분석에 의한 포먼트 값과 단면적에서 추정한 값의 비교

모음	음성 분석에 의한 포먼트 값			단면적에서 추정한 포먼트 값		
	F <sub>1</sub>	F <sub>2</sub>	F <sub>3</sub>	F <sub>1</sub>	F <sub>2</sub>	F <sub>3</sub>
아	744	1,167	2,816	762	1,243	2,691
어	526	1,120	2,818	545	1,229	2,482
에	544	2,027	2,672	492	1,726	2,487
오	412	738	2,620	473	928	2,727
우	296	706	2,344	244	729	2,802
으	311	1,387	2,445	430	1,387	2,803
이	284	2,377	3,093	264	2,141	2,679

여기서 나온 값을 포먼트 합성기인 SenSyn을 이용해 합성한 음성을 쓰니 미니 디스크에 담아 조용한 강의실에서 음성학 수업을 듣는 동의대학교 영문과 3학년 학생들에게 임의의 순서로 3번씩 들려주어 해당되는 모음에 표시하게 했다. 응답지에는 7개의 모음이 해당번호 옆에 인쇄되어 있고 학생들은 이것을 듣고 표시했고 분석은 9번째부터 마지막까지 두 번 되풀이 된 경우를 헤아렸다. 실험에 참가한 50명의 학생 중 /오/와 /어/를 구분하지 못하는 15명의 학생을 제외한 남자 13명, 여자 22명의 응답 빈도수는 표 4와 같다.

표 4. 합성모음 청취실험 결과표

	아	어	에	오	우	으	이
아	560						
어		560					
에			560				
오		5		554		1	
우					560		
으				1		559	
이							560

이 표에서 살펴보면 모음 /오/를 모음 /어/로 표시한 경우가 다섯 번이고 /으/로 표시한 경우가 한 번이고 /으/를 /오/로 표시한 경우가 한 번이다. 따라서, 합성된 음을 거의 모든 피험자가 실제 발음으로 들었음을 나타내준다. 이 청취 실험 결과는 성도 단면적의 포먼트 주파수에서 구한 값이 실제 발음의 값과 일치함을 나타낸다. 통계적으로도 두 포먼트 값을 통계 소프트웨어인 StatViewSE로 처리했을 때 상관관계수가 0.978로 매우 높은 상관관계를 보였다. 물론 F<sub>3</sub>값에서는 차이가 많이 나는 경우도 있으나 단모음 합성 실험에서 800 Hz 이상의 변화도 동일하게 지각하는 사람의 청각능력을 고려해보면 동일한 모음으로 들을 확률이 높다고 할 수 있다. 이는 앞서 Baer 외(1991)나, 러시아 모음의 단면적 값에서 Formfreq를 이용해 구한 포먼트 값을 SenSyn으로 합성한 실험(양병곤 1997)에서도 뒷받침된다.

만일 성도의 좁힘점을 이동시키면 어떠한 포먼트 주파수가 나오고 이에 따른 합성음의 청각적 인상은 어떠한 것인가? 이러한 연구는 성도에 이상이 생긴 환자에게 수술 전에 성도의 모의 변형 실험을 통해 사전에 조음상의 문제점을 예상하는데 도움이 될 것이다(Gracco 외 1994). 앞서 저자는 Fant (1970)의 러시아 단모음의 일부를 선택해 후강과 구강, 입술의 단면적을 변형했을 때 나타난 포먼트 값의 변이가 좁힘점 이론과 일치하고 청각적 인상도 면적의 변화에 따라 달라짐을 밝힌 바 있다(양병곤 1997). 실제 성도의 단면적의 변화는 매우 다양하게 나타나기 때문에 무작정 어느 부위를 넓히거나 좁히는 과정은 중요한 결과를 얻는데 어려움이 많다. 특히 성도의 지점마다 비선형적으로 포먼트 값이 변하는 영역이 있고 이런 비약적인 지역(quantal region)에서는 갑작스러운 음질의 변화를 가져오지만 그 부분 전후에서는 상당한 면적 변화에서도 전혀 음질의 변화가 없다(Stevens 1972; Wood 1979). 예를 들어, 모음 /우/보다는 약간 넓은 입술 면적은 그대로 둔 채 그림 4와 같이 목젓주위의 좁힘점을 1 cm 간격으로 옮겨가도 음질의 차이를 가져올 정도의 포먼트 값의 변화가 없다. 저자의 청각으로는 /으/와 /어/음으로 전이하는 과정에 있는 발음이 들렸으나 여전히 입술이 약간 벌어진 상태의 발음이어서 /으/의 음질을 가지고 있으나 좁힘점의 이동에 따른 음질의 변화는 거의 없었다.

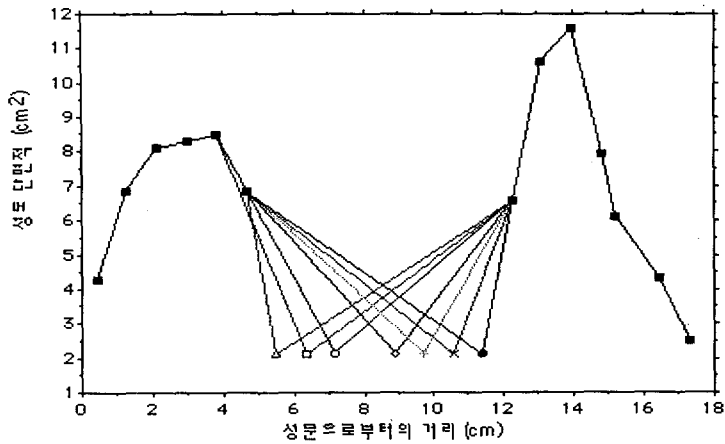


그림 4. 모음 /우/의 좁힘점 변화 실험

따라서, 어떤 면적 변화의 일관성을 찾기 위해서는 기준이 되는 지점을 찾아야 한다. 이 논문에서는 국어의 모음 삼각도에서 모통이에 해당하는 모음인 /아, 우, 이/ 세 모음의 단면적 값 그래프를 실측치로 인쇄한 뒤 이를 성문에서 1 cm 간격으로 자로 재어서 입력하고 동일한 위치에서의 평균값을 그림 5와 같이 나타내 보았다. 모음 /우/의 입술부분을 제외한 성문에서 17 cm까지를 나타냈다.

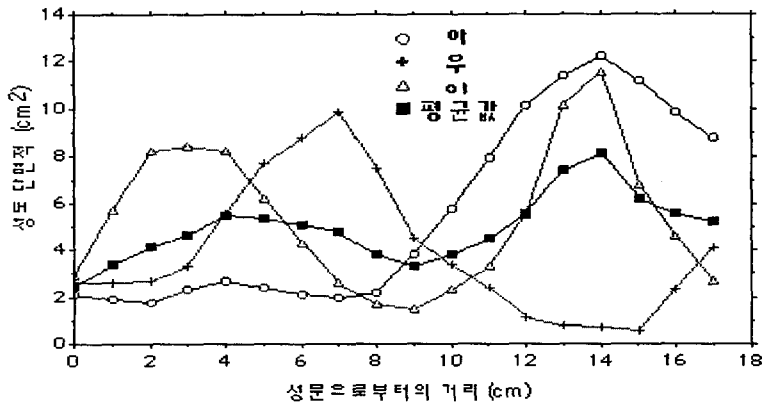


그림 5. 모음 /아, 우, 이/의 단면적 값의 평균

그림 5에서 보면 평균값의 좁힘점은 9 cm 지점에 나타났고 중모음 /어/에 가까운 모양을 하고 있다. 이 단면적 값을 Formfreq에 입력하여 포먼트 주파수를 구한 뒤 합성하였는데 모음 /어/로 들렸다. 앞서의 청각실험에 이 음을 넣어서 피험자들이 표시한 결과로는 총 560번 가운데 549번은 /어/로 들었다. 그 가운데 17번은

/오/로 들었고, 두 번은 /오/로 들었으나 절대 다수가 중모음 /어/로 들었다. 이 평균값을 중심으로 세 모음을 비교해 보면, 성도의 모양이 일정한 면적으로 변형되고 있음을 알 수 있다. 예를 들어, 모음 /아/는 인강 부분이 줄어든 반면 구강 부분이 그만큼 넓어졌다. 모음 /우/는 혀의 중앙 부분이 목젓에 접근함으로써 그만큼의 공간이 인강과 구강으로 나뉘어져 넓어졌다. 모음 /이/는 혀가 입천장에 가까이 다가가면서 구강의 면적이 좁아진데 비해, 그만큼의 체적이 인강에 더해졌다. 덧붙여, 앞서 체적 측정에서도 나타났듯이, 성도의 체적은 턱을 내리거나 올리는 동작에 의해 약 20 cm<sup>3</sup>의 차이를 오게 하지만 대체로 전체 체적의 변화는 크지 않았다. 이러한 관계에서 살펴보면 성도 단면적의 변이는 단순하게 입술만을 넓히거나 좁히기 보다는 평균값을 나타낸 면적 값을 중심으로 한쪽 부분을 넓힌 만큼 다른 쪽을 줄여 가는 정교한 합성 실험이 되어야 할 것이다. 앞으로 이러한 자연스런 성도변형에 의한 합성 실험을 수행할 계획이다.

## 5. 맺음말

이 연구에서는 건강한 남성화자가 모음 7개를 발성할 때의 성도 정중단면과 성도의 각 지점의 단면적을 측정하여 모음의 특징을 살펴보았다. 이어서, 단면적에서 추정된 포먼트 주파수를 이용해 합성한 음성의 청취실험과 모음 삼각도의 모퉁이에 해당하는 모음의 단면적의 평균을 구했다. 연구 결과로는 성도 정중단면에서 화자는 모음발화를 위해 혀나 턱을 이용하여 구강과 인강의 폭을 변화시켜서 원하는 모음을 발성했다. 성도의 폭은 구강과 인강의 폭이 상호 반비례하는 관계를 보였다. 턱의 움직임은 수직으로 내려가지 않고 비껴서 움직이기 때문에 성도의 공간은 성문에서 멀어질수록 변화가 크고, 가까울수록 적어졌다. 이어서 성도 단면적을 비교했는데 정중단면에서 본 좁힘점과 단면적의 좁힘점이 달랐다. 성도의 면적값을 합한 체적에서는 개방모음과 폐쇄모음으로 구분되지는 않았다. 면적 값에 의한 포먼트 값 추정은 발성음에서 추정한 값과 상관도가 매우 높았고, 청취실험에서도 동일한 모음으로 들은 비율이 높았다. 세 개의 모퉁이에 위치한 모음의 단면적 평균값에서 추정한 포먼트 값을 합성했을 때 많은 피험자들이 /어/로 들었다. 성도의 면적 변화를 관찰해 본 결과, 성도의 면적변화 실험은 체적을 일정하게 유지하는 경향을 보여주었다. 앞으로 구강과 비강의 면적을 동시에 조절하는 보다 정교한 단면적 변화에 따른 합성 실험을 통해 인접 모음의 지각적 경계를 밝혀보고자 한다.

이 연구 결과의 기대효과로는 MRI를 이용하여, 모음발성 시의 상세한 조음 구조를 밝히는 과정을 통해 사람의 조음 체계를 깊이 있게 이해할 수 있고, 성도 단면적 실측값에서 나오는 포먼트 값을 합성을 통해 들어봄으로써 의학적으로 어떠한 인강 모양이 조음에 도움이 되는지를 사전 모의 실험을 통해 예측하는 데 도움이 될 것이다. 아울러, 이러한 새로운 분석방법을 보다 더 확대하여 동적인 조음기관의 움직임을 포착할 수 있게 되면 이를 응용하여 외국어를 학습할 때 문제가 되

는 발음을 어떻게 습득할 수 있는지를 구체적으로 밝혀서, 옳은 발음법을 익힐 수 있는 연구를 위한 기초를 제공할 것이다.

### 참 고 문 헌

- 김영송. 1981. *우리말 소리의 연구*. 서울: 과학사.
- 양병곤. 1995. 합성한 한국어 단모음의 지각실험 연구. *언어* 20, 3, 127-146.
- 양병곤. 1996. 합성한 한국어 이중모음의 지각실험 연구. *언어* 21, 3, 829-843.
- 양병곤. 1996. 라링고그래프에 의한 한국인의 성문 파형 분석 및 합성모음의 청각 실험. *언어* 21, 4, 1025-1040.
- 양병곤. 1998. 성도변형에 따른 모음 포먼트의 변화 고찰. *음성과학* 제 3권, 83-92.
- 양병곤, 황수진, 박병래. 1998. MRI에 의한 성도 단면적 측정. *한국음성과학회 제 5 회 학술발표회 논문집*. 27-30.
- Baers, T., J. C. Core, L. C. Gracco, and P. W. Nye. 1991. Analysis of vocal tract shape and dimensions using magnetic resonance imaging: vowels. *J. Acoust. Soc. Am.*, 90, 2, 799-828.
- Baer, T., J.C. Gore, S. Boycew, and P. W. Nye. 1987. Application of MRI to the analysis of speech production. *Magnetic Resonance Imaging*, 5, 1-7.
- Demolin, D., T. Metens, and A. Soquet. 1994. Three-dimensional measurement of the vocal tract by MRI. *ICSLP 96 Proceedings*. Vol. 1, 272-275.
- Fant, G. 1970. *Acoustic Theory of Speech Production*. The Hague: Mouton.
- Foldvick, A. K., O. Husby, J. Kvaerness, I. C. Norli, and P. A. Rinck. 1991. MRI for filming articulatory movements. *Proceedings of the 12th ICPhS*. Aix-en-Provence, 34-36.
- Gay, T. 1992. Acoustic and perceptual effects of changes in vocal tract constrictions for vowels. *J. Acoust. Soc. Am.*, 92, 3, 1301-1309.
- Gracco, C., C. T. Sasaki, R. McGowan, E. Tierney, and J. Gore. 1994. Magnetic resonance imaging(MRI) in vocal tract research: clinical application. *ASA 127th Meeting Abstract* IpSP35.
- Klatt, D. H. and L. C. Klatt. 1990. Analysis, synthesis, and perception of voice quality variations among female and male talkers. *J. Acoust. Soc. Am.*, 87,2, 820-857.
- Matsumura, M. and A. Sugira. 1990. Modeling of 3-dimensional vocal tract shapes obtained by magnetic resonance imaging for speech synthesis. *Proceedings ICSP 90*. 425-428.
- Moore, C. A. 1992. The correspondence of vocal tract resonance with volumes obtained from magnetic resonance imaging. *Journal of Speech and Hearing Research*, 35., 1009-1023.
- Sondhi, M. M. 1974. Model for wave propagation in a lossy vocal tract. *J. Acoust. Soc. Am.*, 55, 5, 1070-1075.
- Sondhi, M. M. 1986. Resonance of a bent vocal tract. *J. Acoust. Soc. Am.* 79, 4, 1113-1116.

- Sondhi, M. M. and J. Schroeter: 1987. A hybrid time-frequency domain articulatory speech synthesizer. *IEEE Trans. ASSP*, 35, 7, 955-967.
- Stevens, K. N. 1972. The quantal nature of speech: evidence from articulatory-acoustic data. In David, E.E. & Denes, P.B., eds. *Human Communication, a Unified View*, 51-66.
- Wood, S. 1979. A radiographic analysis of constriction locations for vowels. *J. of Phonetics* 7, 25-43.
- Yang, B. 1992. An acoustical study of Korean monophthongs. *J. Acoust. Soc. Am.* 91, 4, 2280-83.
- Yang, B. 1996. A comparative study of American English and Korean vowels produced by male and female speakers. *J. of Phonetics* 24, 1, 245-261.
- Yang, C. S and H. Kasuya. 1994. Accurate Measurement of vocal tract shapes from magnetic resonance images of child, female and male subjects. *Proceedings ICSLP 94*, 623-626.
- Yang, C. S. and H. Kasuya. 1996. Speaker individualities of vocal tract shapes of Japanese vowels measured by magnetic resonance images. *Proceedings ICSLP 96*, 949-952.

접수일자 :1998. 9. 10.

게재결정: 1998. 11. 10.

- ▲ 부산광역시 부산진구 가야동 산 24  
 동의대학교 인문대학 영어영문학과(우: 614-714)  
 Tel.: (051) 890-1227 (O), 896-7636 (H)  
 Fax.: (051) 890-1209  
 e-mail: bgyang@hyomin.dongueui.ac.kr  
<http://hyomin.dongueui.ac.kr/~bgyang/index.html>