

## 한국어 단모음의 지각적 모음공간과 심적 표상\*

Perceptual Vowel Space and Mental Representation of Korean Monophthongs

최 양 규\*\*  
Yang-Gyu Choi

### ABSTRACT

The purpose of this study was to examine whether the same vowel sounds are perceived differently by the two local dialect speakers, Seoul dialect speakers (SDS) and Kyungnam dialect speakers (KDS), whose vowel systems differ each other. In the first experiment SDS and KDS heard vowels synthesized in vowel space with F1 by F2 and categorized them into one of 10 Korean monophthongs. The results showed that SDS and KDS perceived the synthesized vowels differently. For example, /ɛ/ versus /e/ contrast, /ø/, and /y/ are differentiated by SDS, whereas they are perceptually confused by KDS. We also observed that /i/ could not be perceived unless the vowel synthesis included F3 or higher formants.

In the second experiment SDS and KDS performed the similarity rating task of 10 synthesized Korean monophthongs. Two-dimensional MDS solution based on the similarity rating scores was obtained for each dialect group. The first dimension can be named 'vowel advancement' and the second 'vowel height'. The comparison of the two MDS solutions showed that the overall psychological distances among the vowels are shorter in KDS than SDS and that especially the distance between /ʌ/ and /i/ is shorter in KDS than SDS. The result suggested that perception or mental representation of vowels depends on the vowel system of the listener's dialect or language. Further research problems were discussed in the final section.

**Keywords:** Vowel Perception, Speech Perception, Mental Representation

### 1. 서 론

본 연구는 첫째, 모음의 질(quality)을 결정한다고 알려져 있는 물리적인 특성인 포먼트(formant) 주파수 값(특히 F1과 F2)에 따라서 그 모음에 대한 지각이 어떻게 달라지는지를 체계적으로 조사하여, 지각에서의 모음공간을 확인하려는 것이며, 둘째, 각 단모음들간의 유사성 자료(similarity data)를 수집하여 다차원척도법(MDS: multi-dimensional scaling)을 통해 우리말 단모음들이 심리적으로는 어떻게 표상되는지를 규명하려는 것이다. 아울러 동일한

\* 본 연구는 한국과학재단 목적기초연구(R05-2003-000-12327-0) 지원으로 수행되었음.

\*\* 춘해대학 유아특수치료교육과

모음들에 대한 지각과 심적 표상(mental representation)이 화자의 방언에 따라서 서로 어떻게 다른지를 알아보고자 하였다.

포만트는 성도(vocal tract)의 공명구조 잘 반영하고 있으며, 성도의 길이가 약 17 cm 정도 되는 남성화자에게서 생성될 수 있는 포만트는 5 개인데, 이 중에서 주파수가 낮은 순서대로 보았을 때 제1 포만트(F1)와 제2 포만트(F2)가 모음의 질을 결정하는 것으로 알려져 있다 (Jenkins, 1987; Peterson & Barney, 1952).<sup>1)</sup>

Peterson & Barney(1952)는 성인 남녀를 포함하여 아동들까지 모두 76 명의 화자가 산출한 모음들을 스펙트럼 분석해 본 결과, F1과 F2가 다른 포만트들에 비해, 그 강도가 강하고 각 모음마다 현저한 차이가 있음을 확인하였다. 그들은 F1과 F2 값을 x, y 좌표값으로 하는 2 차원 평면상에 모음들을 배치하여 도식화하여 각 모음 범주들이 모음공간에서 고유의 영역을 차지하고 있음을 제시한 바 있다. 따라서 그들은 첫 두 포만트가 모음의 특질을 결정한다는 결론에 도달하게 되었다. 만일 F1과 F2가 모음의 특질을 결정한다면 산출된 다른 모음들 간의 F1, F2값 차이를 보여주는 것 뿐 아니라 그 역으로 다른 조합의 F1과 F2값으로 합성된 모음들이 서로 다른 범주의 모음들로 지각됨을 보여줄 수 있어야 할 것이다.

한편, 말소리의 지각과정은 산출과정에 근거하여 이루어진다는 운동이론(motor theory)에 따르면 산출할 수 있는 모음에 따라서 지각할 수 있는 모음이 결정된다고 한다(Liberman & Mattingly, 1985). 만일 그렇다면, 언어나 방언마다 사용되는 자음과 모음 체계가 각기 다르므로 어떤 언어나 방언을 사용해 왔느냐에 따라 지각되는 자음 혹은 모음이 다를 가능성이 있다. 예를 들어, 영어에는 사용되지 않으나 스웨덴어에서는 사용되는 /y/와 /ɥ/가 영어를 사용하는 미국인들에게는 식별(identification)되지 않는다고 한다(Stevens, Liberman, Studdert-Kennedy & Öhman, 1969). 이러한 예는 모음지각은 듣는 사람의 언어 경험에 좌우될 가능성을 시사하고 있다.

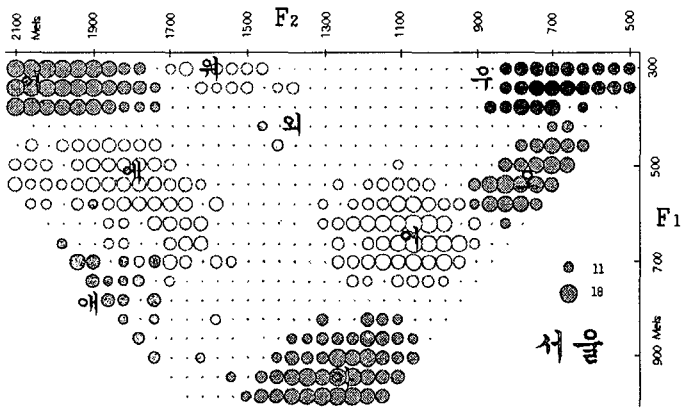
우리말에도 서울방언에서는 사용되나 경남방언에서는 사용되지 않거나 구분이 분명하지 않은 모음들이 있다. 예를 들어, 경남방언에서는 /애/와 /에/, /어/와 /오/ 간의 구분이 모호하며, /외/와 /위/는 단모음으로 발음되지 않는다고 한다(김영송, 1981; 이기문, 김진우, 이상억, 1991). 국어학자에 따라 약간 차이가 있으나 일반적으로 서울방언에서는 /아/, /애/, /어/, /에/, /오/, /외/, /우/, /위/, /으/, /이/가 사용되는 10 모음체계 혹은 /외/와 /위/를 제외한 8 모음체계로, 경남방언은 /아/, /애/, /어/, /오/, /우/, /이/가 사용되는 6 모음체계로 발음된다고 한다(김영송, 1981). 모음의 산출과 지각간에 상호 밀접한 연결관계가 존재한다면, 서로 다른 모음 체계를 가진 서울방언 사용자와 경남방언 사용자들은 F1과 F2를 합성하여 만든 동일한 모음 자극들을 서로 다르게 범주화할 가능성이 있다.

따라서 이러한 여러 가지 사실들에 근거하여 최양규, 신현정 그리고 권오식(1997)은 F1과 F2가 모음지각에 어떠한 역할을 하는지를 구체적으로 규명하기 위하여 F1과 F2, 이 두 포만트만으로 달리 조합되어 합성된 소리가 어떻게 여러 가지 다른 모음들로 지각되는지를 알아보고, 아울러 다른 모음체계를 가진 서울방언과 경남방언 사용자들 간에 지각에서의 차이를 알아보았다.

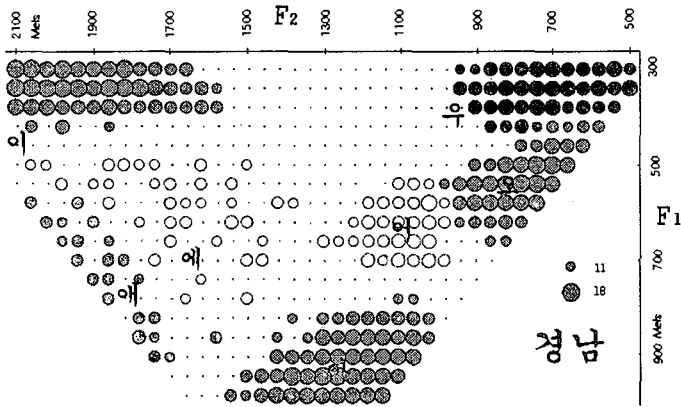
그 결과인 그림 1을 보면 F1과 F2에 의한 모음공간에서 각 모음들이 고유 영역을 차지하

1) 이후부터 제1, 제2, 제3포만트나 그 주파수를 각각 F1, F2, F3으로 지칭할 것이다.

고 있으며 각 모음들의 상대적 위치가 산출된 모음의 분석 결과와 동일하였다. 이 결과는 말소리 지각에서도 산출된 말소리에서처럼 F1과 F2가 중요한 역할을 한다는 것을 시사하였다. 서울방언 사용자의 경우 /에/와 /애/ 영역과 /외/와 /위/의 영역이 좁기는 하지만 구분되어 나타나고 있으나 경남방언 사용자의 경우 이 모두가 나타나고 있지 않다. 특기할 만한 사항은 서울방언 사용자에게서는 /어/와 /으/ 영역이 구분되어 나타나고 경남방언 사용자에게서는 구분되어 나타나지 않을 것으로 예상하였으나, 두 방언 모두에서 모음 /으/에 해당하는 영역이 나타나지 않았다는 점이다. 이것은 F1과 F2만으로는 /으/ 모음으로 지각되는 소리가 제대로 합성되지 않음을 시사하고 있다. 다시 말해서 /으/ 모음으로 지각되기 위해서는 F3 이후의 포만트가 필요하다는 것으로 해석할 수 있다. 또한 합성음의 지각에서 얻어진 F1과 F2가 산출



(A)



(B)

그림 1. 2 개의 포만트(F1, F2)만으로 합성된 모음에 대한 서울방언 사용자(A)와 경남방언 사용자(B)의 지각적 모음공간(최양규 등, 1997). (9 명의 실험참가자로부터 얻은 각 자극에 대한 총 18 회 반응 중 11 회 이상 일치 반응을 보인 자극들을 일치빈도에 따라 원의 크기로 표시하였다)

된 말소리의 F1, F2에 비해 전반적으로 큰 값을 보이고 있는데, 이것은 합성음에서 F3 이후의 포만트가 생략됨으로써 특정 모음으로 식별되기 위해 필요한 합성음의 두 포만트 값이 모음공간의 바깥쪽으로 이행한 것이라고 추정되었다.

따라서 본 연구의 실험 1에서는 F3 이후의 포만트를 추가하여 합성하였을 때, 모음공간에서 패턴이 어떻게 달라질 것인지, 즉 모음의 지각적인 특질에 어떤 변화가 일어나는지를 알아보려고 하였다. 아울러 F1, F2만으로 합성된 모음에 대한 최양규 등(1997)의 연구와 비교해서, F3 이후의 포만트가 포함되어 합성된 모음에 대해서는 지각에서 서울방언과 경남방언 사용자간의 차이가 어떤 양상으로 나타나는지를 알아보려고 하였다.

## 2. 실험 1: 한국어 단모음의 지각적 모음공간

본 연구의 실험 1에서는 첫째, 다섯 개의 포만트 모두를 사용하되 F3, F4, F5는 적절하게 고정시키고,<sup>2)</sup> F1과 F2만을 달리 조합하여 합성된 모음을 제시했을 때, 어떻게 우리말에서 사용되는 여러 가지 단모음들로 지각되는지를 확인함으로써 F1과 F2의 정보가 모음지각의 필요충분조건인지를 살펴보고자 하였다. 아울러 지각에서의 모음공간이 산출과 비교해서 볼 때 어떠한 패턴으로 나타나는지를 알아보았다. 둘째, 다른 모음체계를 가진 서울방언과 경남방언 사용자들의 모음지각에서 어떤 차이가 있는지를 알아보았다. 즉 지각에서 나타나는 두 방언의 모음공간의 패턴을 비교해 보고자 하였다.

### 2.1 방법

#### 2.1.1 실험참가자

태어나서 줄곧 20년 이상 서울 또는 부산 및 경남지역에 거주하였던 20-40대 서울방언 사용자 11명과 경남방언 사용자 11명이 실험에 참가하였다.<sup>3)</sup>

#### 2.1.2 자극

F1과 F2의 주파수의 범위를 각각 380-860 Mel, 580-1780 Mel로 하여 Klatt (1980)의 직/병렬 포만트 합성기(cascade/parallel formant synthesizer)로 합성한 소리를 자극으로 사용하였다. F1과 F2 각각에서 간격을 40 Mel로 하여, 범위 내에서 조합하여 합성가능한 음들 중에서 포만트들간의 간격이 매우 좁아서 소리가 왜곡되는 경우, F1이 F2보다 높은 값을 가지는 경우, 그리고 F1, F2 모두가 매우 높은 값을 가지는 몇몇 경우를 제외시킨 총 247개의 조합을 만들어 소리를 합성하였다. 음성합성에 사용한 소프트웨어는 Klatt (1980)의 프로그램이 연구의 용도에 맞게 수정한 것이다. 자극을 합성하기 위해 프레임당 40개의 파라메타를

- 
- 2) 산출된 모음을 분석한 결과들을 보면 F1과 F2는 모음에 따라 큰 차이를 보이지만 F3, F4, F5는 거의 비슷한 값을 보였다. 따라서 본 연구에서 F3은 F2와 충돌이 일어나지 않을 만큼의 약간의 변화만을 주고 나머지 F4, F5는 모두 동일한 값을 부여하였다.
- 3) 사전연구에서 지각반응에 성별에 따른 차이가 관찰되지 않았기 때문에 실험에 참가한 서울방언 사용자(남7명, 여4명)와 경남방언 사용자(남5명, 여6명)간의 성별 비율을 통제하지는 않았다.

사용했는데, 이 중 12 개의 파라메타(AV, F0, F1, F2, F3, F4, F5, B1, B2, B3, B4, B5)의 값이 자극합성을 위해 조정되었다. F2가 2,160 Hz (1660 mel) 이상인 10 개의 자극은 예외로 F3을 3,040 Hz로 고정시키고 나머지는 모두 2,600 Hz로 하였다. F1의 대역폭(B1)과 F2의 대역폭(B2)은 각각 60 Hz와 90 Hz로 고정시켰다(Klatt, 1980). F4, F5, B1, B2, B3, B4, B5는 각각 3,300, 3,850, 60, 90, 200, 200, 150 Hz로 모든 자극에 공통적으로 적용되었다. 나머지는 Klatt (1980)의 기본값을 그대로 사용했다. 소리의 진폭(AV)은 최초 120 ms 동안 10 dB에서 점차로 65 dB로 커지고 그 다음 100 ms 동안 지속하다가 마지막 200 ms 동안 서서히 30 dB로 낮아지도록 하였다. 음고를 결정하는 기본주파수(F0)는 첫 20 ms 동안 140 Hz로 하고 나머지 300 ms 동안은 140 Hz에서 점차로 113 Hz로 낮아지도록 했다. 프레임당 10 ms 분량의 음파가 생성되도록 했고, 한 자극을 합성하는 데 모두 42 프레임의 파라메타를 사용하여 자극의 지속시간이 420 ms가 되도록 하였다. 표본속도는 10,000 Hz로 하였다. 자연스런 모음을 합성하기 위해 직렬(cascade)방식으로 합성하였다. 이렇게 한 것은 포먼트들간의 관계 속에서 포먼트들의 진폭값(A1, A2, A3, A4, A5)이 자동적으로 결정되기 때문이다(Klatt, 1980). 자극은 실시간으로 합성되어 Sony MDR-CD1000 헤드폰을 통해 실험참가자에게 제시되었다.

### 2.1.3 도구

사운드 카드가 장착되어 있고 운영체제로 Linux가 설치되어 있는 펜티엄 PC가 자극합성과 제시 및 실험통제에 사용되었다. 자극청취에는 Sony MDR-CD1000 헤드폰이 사용되었다. 자극합성 모듈은 C언어로 프로그램 되었고 자극제시 및 통제는 X-Window 상에서 Tcl/Tk로 프로그램 되었다.

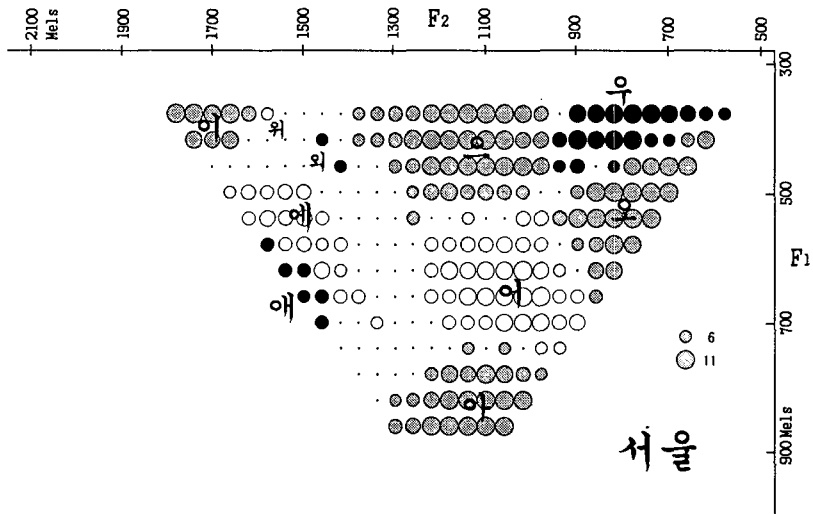
### 2.1.4 절차

실험은 춘해대학 상담심리과 실험실에서 한 번에 한 사람씩 개별적으로 실시되었다. 실험참가자는 실험실에 설치된 컴퓨터 모니터 앞에 앉아 컴퓨터로 합성되는 자극소리를 하나씩 듣고서 그에 해당하는 키보드의 모음 글쇠를 누르도록 하였다. 다음 시행으로 넘어갈 때는 마우스에 연결된 확장 버튼을 사용하도록 하였다. 확장 버튼 위에는 '다음'이라고 표시해 두었으며, 자판(keyboard) 가운데 11 개의 글쇠(key)에는 '아', '어', '오', '우', '으', '이', '애', '에', '외', '위', 'X' 등의 글자가 적힌 스티커를 붙여두어 실험참가자가 반응해야 할 해당 글쇠를 쉽게 찾도록 했다. 제시된 자극에 해당하는 글자가 적힌 글쇠를 누르되, 애매하거나 잘 모르는 자극의 경우에는 'X'를 누르도록 하였다. 실험참가자는 헤드폰을 쓰고 먼저 20 회의 연습시행을 한 뒤 동일한 절차로 진행되는 본 실험에 참가하였다. 실험은 두 세션으로 나누어 진행되었다. 한 세션 당 247 시행이었고, 한 세션 내에서 모든 자극음들이 실험참가자마다 무선적인 순서로 제시되었다. 세션간에는 약 5 분 정도의 휴식기간이 있었다.

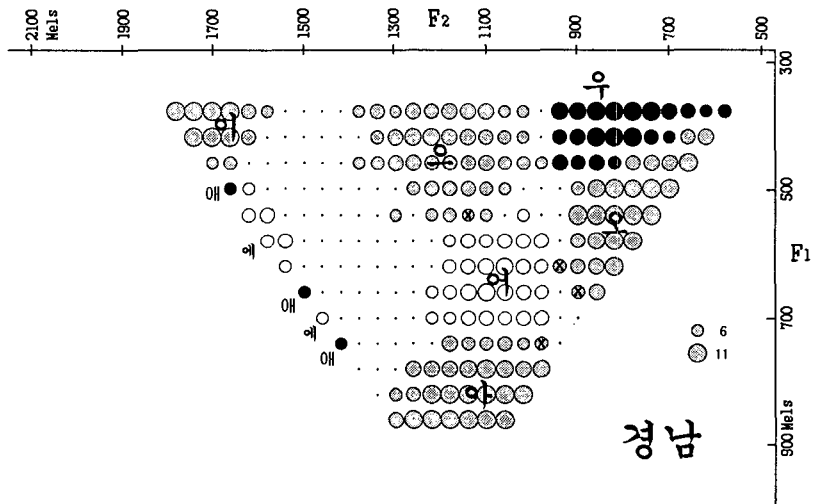
## 2.2 결과 및 논의

각 자극음에 대한 각 실험참가자의 반응은 세 번이었는데(3 세션 동안) 이 중 적어도 두 번 이상 일치하는 경우만 분석에 포함시켰다. 두 집단 각각에서 11 명 중 6 명 이상의 실험참가자들간에 일치를 보이는 자극의 위치에는 일치 빈도에 따라 원크기를 달리하여 각 방언사

용자 집단별로 반응 결과를 표시한 것이 그림 2이다. 그림 2를 보면 F1과 F2 차원에서 만들어진 2 차원 공간상에 각 모음들이 고유의 영역을 차지하고 있는 것을 볼 수 있다. 또한 각 모음들간의 상대적 위치는 모음들을 조음위치에 따라 도식화한 모음 사각도(혹은 삼각도)와 거의 일치하고 있음을 볼 수 있다(그림 3 참조). 그리고 F1과 F2, 이 두 포만트만으로 합성된 모음에 대한 결과인 그림 1과, F3 이후의 포만트가 추가되어 5 개의 포만트로 합성된 모음에 대한 결과인 그림 2를 서로 비교해서 살펴보면 다음과 같은 사실을 추론해 볼 수 있다. F3 이후의 포만트가 없이 F1과 F2만으로 특정 모음으로 지각되기 위해서는 F1과 F2가 높은 값을



(A)



(B)

그림 2. F1과 F2를 달리하여 5 개의 포만트로 합성한 모음에 대한 서울방언 사용자(A)와 경남 방언 사용자(B)의 모음공간(X 표시가 된 자극은 서울방언 사용자 /어/로 지각한 자극임)

가져야한다. 따라서 모음공간이 그림 1의 경우 그림 2에 비해 전반적으로 넓게 나타났던 것이라 생각할 수 있다. 그림 1에서는 나타나지 않았던 모음 /으/ 영역이 넓게 나타났는데 이것은 한국어 /으/ 모음의 지각에 F3 이후의 포만트가 중요한 역할을 한다는 것을 시사하고 있다. 다시 말해서 다른 모음과는 달리 한국어 /으/ 모음 지각을 위해서는 F1과 F2만으로는 불충분하고 F3 이후의 포만트가 필요하다는 것이다.

서울방언 사용자와 경남방언 사용자간의 몇 가지 차이점을 발견할 수 있다. 가장 대표적인 차이는 /에/와 /애/, /외/와 /위/ 간의 식별에서의 차이이다. 서울방언 사용자들의 경우는 /에/와 /애/ 그리고 /외/와 /위/의 고유영역이 어느 정도 나타나는 반면, 경남방언 사용자의 경우에는 /에/와 /애/를 서로 혼동하고 있고 /외/와 /위/에 대한 반응도 없었다. 이것을 볼 때 서울방언 사용자는 /에/와 /애/, /외/와 /위/를 어느 정도 구분하여 지각하는 반면, 경남방언 사용자들의 경우는 /에/와 /애/, /외/와 /위/ 간의 분명한 구분을 하지 못함을 시사 받을 수 있다. 그림 2(B)를 보면 경남방언에서도 서울방언과 마찬가지로 각각 /어/와 /으/로 들리는 고유의 영역이 나타나는 것을 볼 수 있는데 이러한 결과는 예상과는 다르게 경남방언 사용자들이 /으/와 /어/를 다르게 지각할 수 있음을 보여준다.

그런데 서울방언 사용자들의 /위/와 /외/ 그리고 /애/와 /에/는 경남방언과는 달리 모음공간에서 고유의 영역을 차지하기는 하나 그 영역이 다른 모음들에 비해 좁다. 이것은 서울방언의 모음체계에서 /위/, /외/가 점차 사라져가고 있고 /에/와 /애/간의 대비도 점차 약해지는 추세임을 보여준다. 한편 경남방언 사용자들은 /으/와 /어/를 지각적으로 변별하는 것으로 나타났는데 이것은 경남방언의 모음체계가 점차 /으/와 /어/의 대비(contrast)를 수용해가고 있음을 보여준다. 다시 말해서 서울방언은 10 혹은 8 모음체계에서 7 모음체계로, 경남방언은 6 모음체계에서 7 모음체계로의 통시적 변천과정(diachronic transition)에 있음을 시사하고 있다(이호영, 1996).

합성음의 지각 결과에서 얻어진 본 실험 1의 결과를 산출된 모음의 분석 결과들과 비교해 보면 모음들 간의 상대적 위치가 대단히 유사함을 발견할 수 있다. 이것은 모음의 산출과정과 마찬가지로 지각과정에서도 F1과 F2, 이 두 개의 포만트가 중요한 역할을 하고 있음을 말해준다. 지각에서의 모음공간과 산출된 모음에서 나타나는 모음공간 간에 거의 유사한 패턴을 보이고 있다는 사실은 모음의 산출과 지각간의 관련성을 시사해 주는 것으로 생각된다. 보다 구체적으로 말해서 첫 두 포만트가 모음지각을 위한 상당한 정보를 제공하므로 지각에서도 거의 구분되는 모음영역을 보였다고 생각할 수 있다. 모음 산출 시에 성도의 구조에 의해 공명된 주파수대역에 따른 독특한 포만트 주파수가 결정되므로 포만트는 조음기관의 구조 혹은 운동에 관한 정보를 갖고 있다고 할 수 있으며, 말소리 지각 시스템이 이 정보를 이용하여 조음기관의 운동을 추론하여 화자가 의도한 음소를 지각할 가능성이 있다. 이러한 생각은 운동 이론으로 잘 설명될 수 있다.

그림 2에서 모음영역의 중심에서 벗어날수록 일치빈도가 점차 작아짐을 볼 수 있는데 이것은 자음지각과는 달리 모음지각은 범주적이라기보다 등위적(graded)이라는 사실을 시사한다. 그리고 그림 2를 보다 자세히 살펴보면 각 모음 범주의 경계가 두 방언에서 다르게 나타남을 찾아볼 수 있다. 예컨대 경남방언 사용자의 /어/ 영역(27 개 자극)은 서울방언(37 개 자극)에 비해 좁게 나타났다. 따라서 동일한 모음자극에 대해서 두 방언 사용자들이 다르게 지

각하는 경우가 발생했다. 예컨대 F1과 F2가 각각 620과 940 Mels, 660과 900 Mels, 740과 980 Mels, 540과 1140 Mels인 네 개의 자극들(그림 2B에서 X로 표시된 자극)을 서울방언 사용자는 모두 /어/로 지각하였으나 경남방언 사용자는 이 중에서 첫 두 자극을 /오/로 지각하였고 나머지 두 자극을 각각 /아/와 /으/로 지각하였다. 이 결과는 모음공간상의 각 모음범주 경계가 언어적 환경에 좌우됨을 시사하고 있으며 이 사실은 사용하는 방언이 다를 때, 즉 언어학습환경이 다르면 지각도 달라짐을 보여준다.

### 3. 실험 2: 우리말 단모음의 심적 표상

본 실험 2에서는 각 단모음들간의 심리적 거리를 보다 자세하게 규명하기 위해 F1과 F2로 표현되는 모음공간이 심리적으로는 어떻게 표상되는지를 알아보고자 하였다. 서로 분명하게 구분되는 모음일수록 심리적으로 더 멀리 표상되고, 보다 유사하게 지각되는 모음일수록 심리적으로 더 가까이 표상될 것이다.

모음은 조음음성학적으로 볼 때 전후성(advancement), 고저성(height) 그리고 원순성(roundness) 차원에서 구분된다. 전후성은 모음 조음과정에서 혀의 최고점의 전후위치를, 그리고 고저성은 혀의 최고점의 고저위치를 말한다. 원순성은 입술 모양의 둥근 정도이다. 구체적인 예를 들면, 전후성에 따라 전설모음(이, 에), 중설모음(으, 어), 후설모음(우, 오)으로 구분되고, 고저성에 따라 고모음(이, 으, 우), 중모음(에, 어, 오), 저모음(아)으로 구분된다(김무림, 1993; 김영송, 1981; 이호영, 1996). 예컨대 /이/는 대표적인 전설모음이며, /우/는 후설모음에 속한다. /오/, /우/는 원순음이며, /이/는 비원순음에 속한다. 이러한 모음분류는 조음음성학적 연구에 근거한 것으로 그림 3과 같이 모음사각도로 도식화하여 나타낼 수 있다.

	전설	중설	후설
고	이 위	으	우
중	에 외	어	오
저	애	아	

그림 3. 국어 모음 사각도

모든 단모음은 F1과 F2의 두 차원으로 이루어진 이차원 공간에 배열될 수 있다. 실험 1에서 우리말의 경우에도 F1과 F2를 달리하여 합성하면 거의 모든 단모음이 생성될 수 있음이 입증되었다(그림 2 참조). 그렇다면 모음의 질을 결정하는 이러한 물리적 차원에 대응되는 심리적 차원은 무엇이며 이 차원에 근거한 심리적 공간에서 단모음들이 어떻게 표상되는가?

다차원척도법(MDS)은 모음지각 연구에서 약 30년 이상 사용되어 온 대표적인 수리적 도



구로서 모음의 심리적 표상을 알아내거나 심리적 차원을 찾아내는 데 현재로서는 가장 유용한 방법으로 알려져 있다(Flege, Munro, & Fox, 1994; Fox, 1982, 1983; Fox, Flege, & Munro, 1995; Fox & Trudeau, 1988; Rakerd & Verbrugge, 1985; Shepard, 1972; Singh & Woods, 1971; Terbeek, 1977). 본 실험에서는 10 개의 우리말 단모음들(아, 에, 어, 예, 오, 외, 우, 위, 으, 이) 간의 유사성 평정자료(similarity rating data)를 다차원척도법으로 분석함으로써 그 속에 내재하는 심리적 차원과 심적 표상을 밝혀보려 하였다. 아울러 이러한 단모음의 심리적 차원이 방언에 따라 다르게 나타나는지 그리고 각 심리적 차원상에 나타나는 모음의 심적 표상에서 서울과 경남방언 사용자간에 어떤 차이가 있는지를 알아보았다. 다시 말해서 모음체계가 다른 방언에 따라 모음들이 심적 공간에서 어떻게 다르게 표상되는 지를 알아보고자 하였다.

### 3.1 방법

#### 3.1.1 실험참가자

부산·경남방언 사용자로서 춘해대학의 상담심리과 재학생 11 명이 그리고 서울방언 사용자로서 고려대학교 심리학과 대학원생 8 명이 실험에 참가하였다. 이들의 나이는 20-30 세이었다.

#### 3.1.2 자극

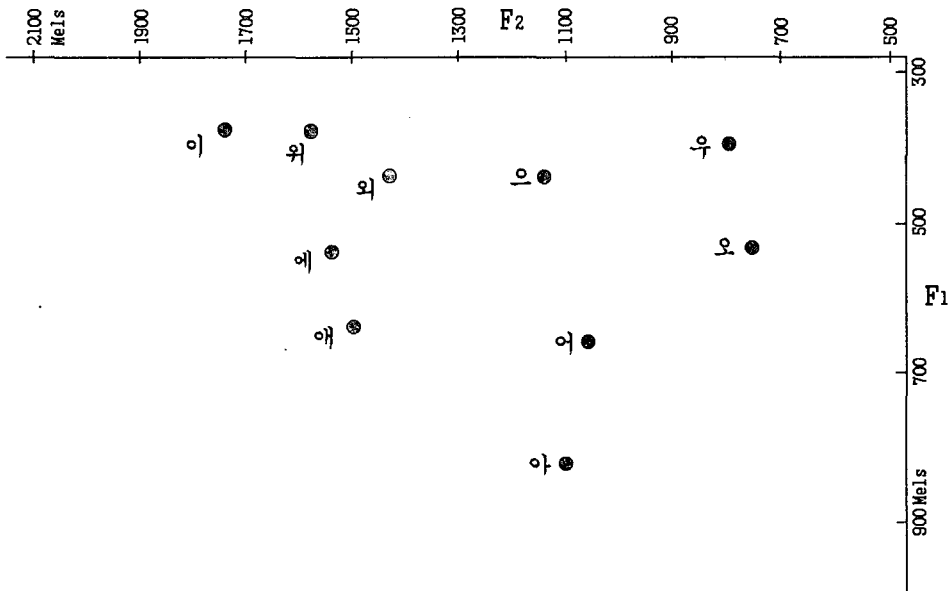


그림 4. 그림 2A(서울)에서 각 모음영역의 중심에 해당하는 위치

표 1. 모음의 분류와 모음자극 합성에 사용된 F1과 F2 [단위: Hz (Mel)]

혀의 전후위치	혀의 높이	입술 모양	모음	F1	F2
전설	고	평순	이	301(380)	2338(1739)
		원순	위	301(380)	1990(1580)
	중	평순	애	454(540)	1910(1541)
		원순	외	376(460)	1627(1393)
	저	평순	애	560(641)	1750(1459)
	중설	고	평순	으	370(454)
중		평순	어	580(660)	1082(1058)
저		평순	아	765(819)	1144(1100)
후설	고	원순	우	310(390)	730(791)
	중	원순	오	454(540)	710(774)

Klatt (1980)의 직/병렬 포만트 합성기를 이용하여 F1과 F2만으로 합성된 10 개의 우리말 단모음(아, 애, 어, 에, 오, 외, 우, 위, 으, 이)이 자극으로 사용되었다. 각 자극은 실험 1에서 우리말 각 모음의 F1과 F2 공간에 속하면서 동시에 가장 높은 빈도로 해당 모음으로 판단된 자극들이었다(그림 4 참조). 각 모음을 합성하는 데 사용된 F1과 F2 수치가 표 1에 제시되어 있다. 나머지는 실험 1과 동일하였다.

### 3.1.3 도구

실험 1과 동일하였다.

### 3.1.4 절차

실험은 춘해대 상담심리과 실험실과 고려대 심리학과 실험실에서 한번에 한 사람씩 개별적으로 실시되었다. 10 개의 모음으로 가능한 45 쌍의 모음 쌍에 대한 유사성 자료를 얻기 위해 9 점 척도상에서 모음쌍의 유사성 비교 평정을 실시하였다. 실험참가자들에게는 두 음이 유사할수록 큰 점수를 주도록 지시하였다. 아울러 가능한 한 '1'부터 '9'까지 다양하게 반응하도록 요구하였다. 한 쌍의 모음이 500 ms 간격을 두고 헤드폰을 통해 차례로 제시되면 실험 참가자는 '1'에서 '9'까지의 스티커가 붙여진 키보드에서 해당 글쇠를 눌러 반응하였다. 반응한 후 스페이스바를 누르면 다음 시행으로 넘어가도록 하였다.

실험은 두 세션으로 구성되었고 각 세션은 다시 두 블록으로 나뉘어졌다. 각 블록에서 45 개의 모든 모음쌍들이 무선적 순서로 제시되었다. 모음쌍 내의 두 모음의 제시순서 효과를 가능한 한 배제시키기 위하여 각 세션의 블록 2에서는 모음쌍 내의 자극제시순서를 블록 1에서의 모음쌍 내의 자극제시순서와 반대가 되도록 하였다. 세션 1이 끝나면 5분 정도 휴식이 있는 후 두 번째 세션이 동일한 절차로 실시되었다. 본 시행 전에 20 회의 연습시행이 실시되었다. 결과적으로 실험참가자당 총 200 시행이 실시되었고, 각 모음쌍은 네 차례 평정되었다.

### 3.2 결과 및 논의

45 개의 각 모음쌍별로 유사성 평정치의 중앙치(median)를 구하여 구성된 행렬을 SPSS 10.0의 ALSCAL을 이용하여 각 방언별로 MDS해(解)를 구하였다. 스트레스값과 각 차원에 대한 해석의 용이성을 고려하여 최종적으로 이차원의 해를 구하였다. 각 차원과 자극의 물리적 속성간의 관계를 보다 명확하게 파악하기 위해서 각 자극의 F1, F2와 차원값, D1, D2간에 최대의 상관성을 보이도록 두 차원의 축을 회전시켜 각 자극의 차원값을 다시 구하였다. 각 단모음 자극의 변경된 각 차원상의 값과 단모음 분류를 나타낸 것이 표 2이고, 변경된 차원 1과 차원 2의 이차원 평면에 각 단모음 자극을 각 지역방언별로 도식화한 것이 그림 5이다. 그림 5를 보면, 차원 1은 모음의 전후성을 반영하고 있으며, 차원 2는 모음의 고저성을 반영하고 있음을 알 수 있다. 또한 실험 1과 마찬가지로 서울방언 사용자나 경남방언 사용자 모두에게서 /애/와 /에/, /외/와 /위/가 심리적 공간에서 서로 인접해 있다. 이것은 실험 1에서 언급한 것처럼 두 모음대비(vowel contrast)가 사라져간다는 사실을 거듭 시사하고 있다. 이 네 개의 모음을 제외한 다른 모음들은 이차원 공간에서 서로 적당한 거리를 두고 있는데 이것은 서로 지각적으로 변별이 용이함을 시사해 주고 있다.

두 방언의 모음에 대한 심리적 표상을 비교해 보면 전체적인 패턴은 서로 유사하나 가장 현저하게 두드러진 차이는 /어/와 /오/간의 심리적인 거리이다. 경남방언 사용자의 심리적 공간에서 /어/와 /오/는 서울방언 사용자보다 훨씬 더 가까이 위치해 있다. 이것은 경남방언 사용자에게는 /어/와 /오/가 서울방언 사용자보다 심리적으로 또는 지각적으로 훨씬 더 유사하다는 것을 보여준다. 서울과 경남방언 사용자의 /어/와 /오/간의 평균 유사성 평정점수는 각각 3.1과 5로서 그 차이는 통계적으로도 유의했다( $p < .05$ ).

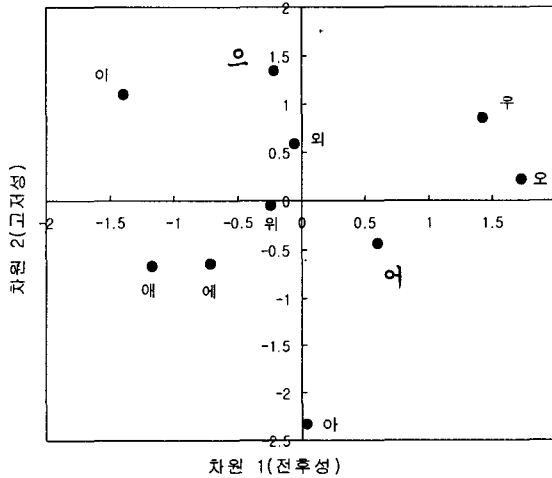
그리고 서울방언 사용자와 비교해 볼 때 경남방언 사용자는 단모음 /외/, /위/, /오/, 이 3개의 단모음을 심리적으로 매우 가깝게 표상하고 있으며, /오/를 /우/와는 더 멀리 /어/와는 보다 가깝게 표상하고 있음을 알 수 있다. 그 반면 /애/와 /에/는 두 방언 사용자간에 큰 차이가 없다.

그림 5에 나타난 모음들의 심리적 공간에서의 배열패턴을 보면 실험에 사용된 모음 자극들의 F1과 F2에 의한 모음공간 상의 패턴(그림 4)과 매우 유사하다는 사실을 볼 수 있다. 이것은 다차원최도법으로 찾아진 두 지각차원이 모음의 물리적 차원이라 할 수 있는 F1과 F2와 상응함을 보여주는 결과라 하겠다. 이러한 상응관계를 확인하기 위해 각 방언별로 모음들의 각 심리적 차원 상의 좌표값과 F1, F2값들간의 상관분석을 해본 결과, 차원 1은 F2와, 그리고 차원 2는 F1과 서울방언에서 각각  $-0.895$ ,  $-0.878$ , 경남방언에서 각각  $-0.925$ ,  $-0.914$ 의 높은 상관을 보였다( $p < .001$ ). 따라서 차원 1은 모음의 전후성과, 그리고 차원 2는 모음의 고저성과 대응되는 차원이라고 해석할 수 있다. 이러한 해석은 영어권에서 수행된 연구결과들과 일치한다(Fox, 1982, 1983; Rakerd & Verbrugge, 1985).

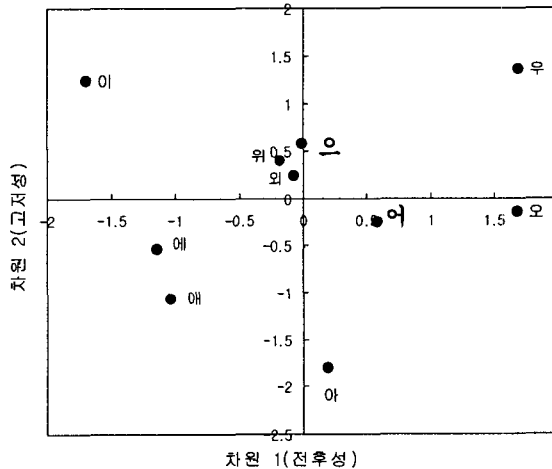
실험 2의 결과를 요약하면 첫째, 경남방언에서 /어/와 /오/는 서울방언에서보다 심리적으로 보다 유사하여 변별이 용이하지 않다. 둘째, 우리말 모음의 경우 2 개의 심리적 차원이 물리적·음향적 차원인 F1과 F2와 밀접히 연관되어 있음을 알 수 있다.

표 2. MDS 분석 결과로 얻어진 차원 1과 차원 2에서의 각 모음자극의 차원 값과 F1, F2값 (Mel)

혀의 전후위치	혀의 높낮이	모음	F1	F2	서울방언		경남방언	
					차원1	차원2	차원1	차원2
전설	고	이	301	2338	-1.394	1.104	-1.693	1.242
		위	301	1990	-.241	-.049	-.188	.391
	중	에	454	1910	-.717	-.644	-1.143	-.541
		외	376	1627	-.051	.598	-.074	.244
저	애	560	1750	-1.177	-.667	-1.037	-1.070	
중설	고	으	370	1202	-.211	1.349	-.008	.582
	중	어	580	1082	.605	-.435	.588	-.245
	저	아	765	1144	.045	-2.338	.192	-1.800
후설	고	우	310	730	1.417	.860	1.689	1.350
	중	오	454	710	1.724	.222	1.674	-.153



(A) 서울방언 사용자의 심적 표상



(B) 경남방언 사용자의 심적 표상

그림 5. 우리말 단모음에 대한 다차원 척도 분석 결과

#### 4. 종합 논의

본 연구에서는 두 실험을 통해 서울방언과 경남방언에서 산출되는 모음뿐만 아니라 지각되는 모음도 다르다는 사실을 보여주었다. 이는 언어적 경험을 통한 학습이 모음의 범주를 결정한다는 것을 시사한다. 다시 말해서 사용하는 방언이 다를 때, 특히 쓰여지는 모음이 다를 때 모음의 지각도 다르다는 것이다. 이 사실은 언어학습 환경이 다르면 지각도 달라질 수 있음을 시사한다. 다시 말해서 말소리의 산출과 지각은 밀접하게 관련됨을 시사하고 있다.

실험 1에서 경남방언 사용자들의 경우 /어/와 /으/가 모음공간에서 서로 혼동되어 있을 것이라고 기대되었으나 결과는 각각 고유의 영역을 차지하여 /어/와 /으/를 변별하여 지각할 수 있음을 보여주었다. 이것은 시대가 지남에 따라, 그리고 특히 학교에서 행해지는 국어교육으로 인해 30세 전의 비교적 젊은 세대들의 경우,<sup>4)</sup> 경남방언에서도 /어/와 /으/간의 대조가 점차 커지는 경향을 반영해 주는 것으로 해석되었다.

그리고 모음공간에서 각 단모음 범주의 경계가 방언마다 다르게 나타났다. 따라서 동일한 모음자극에 대해서도 두 방언 사용자들이 다르게 지각하는 경우가 발생했다. 예컨대 서울방언 사용자들이 /어/로 지각하는 자극들을 경남방언 사용자들이 /으/, /아/나 /으/로 지각하는 경우가 있었다.<sup>5)</sup> 이것은 모음범주의 경계가 언어적 환경에 좌우됨을 시사한다.

실험 2의 결과를 보면 서울방언에 비해 경남방언에서는 아직도 /어/와 /으/가 심리적으로 매우 가깝고 유사한 상태로 표상되고 있다는 사실을 알 수 있다. 이러한 사실 역시 모음지각이 산출되는 모음체계와 모종의 관계가 있음을 시사하고 있다. 말소리의 지각과 산출간의 어떻게 연계가 되는지가 밝혀진다면 말소리의 지각 및 발음의 교정 혹은 외국어 청취 및 발음 교육을 위한 보다 효과적인 훈련방법을 모색할 수 있을 것이다.

예컨대, 지역방언을 사용하는 사람들 중에서 모음지각검사를 통해 /애/와 /에/ 혹은 /어/와 /으/ 간의 차이를 지각하지 못하는 사람들을 대상으로 지각변별훈련을 실시한 다음에 지각변별훈련 전과 비교해서 훈련 후에 발음이 어느 정도 개선되는지를 모음 산출평가를 통하여 확인해 볼 수 있을 것이다. 아울러 변별되어 지각되지 않는 모음 쌍(예컨대 /애/와 /에/)들에 관한 산출훈련(혹은 발음훈련)이 지각에 미치는 영향을 알아볼 수 있다. 지각이 산출에 미치는 영향과 산출이 지각에 미치는 영향을 서로 비교해보는 연구를 통해서 말의 지각과 산출간의 연계를 확인할 수 있을 것이다. 더 나아가 지각 개선 혹은 산출 개선을 위해 어떤 방법이 가장 효과적인지를 시사 받을 수 있으며, 효과적인 외국어 훈련 프로그램이나 언어치료 프로그램 개발을 위한 차후 연구의 기초자료를 제공받을 수 있을 것이다.

최근에 널리 유행하는 상업적인 영어청취학습 프로그램들은 과거와는 달리 특히 언어산출을 강조하는 경향이 있다. 만일 산출을 강조하는 이러한 프로그램들이 효과적이라면 이것은 산출이 지각 또는 청취를 촉진시켜주는 효과 때문이라고 추정해 볼 수 있겠다. 이에 관한 보다 정교한 연구가 차후에 계속 진행된다면 효과적인 언어교정과 외국어 교육 방법 모색에 크게 기여할 수 있을 것이다.

4) 본 연구의 실험 1에 참여한 실험참가자들이 바로 이 연령 대에 속한다.

5) 실생활에서, 경남방언 사용자들 중에는 '떡 사세요'라고 하는 어떤 서울방언 사용자의 말소리가 '뚝 사세요'로 들리는 이도 있다.

## 참 고 문 헌

- 김무림. 1993. *국어음운론*. 한신문화사.
- 김영송. 1981. *우리말 소리의 연구*. 과학사.
- 이기문, 김진우, 이상억. 1991. *국어음운론*. 학연사.
- 이호영. 1996. *국어음성학*. 태학사.
- 최양규, 신현정, 권오식. 1997. "F1과 F2 모음공간에서 합성된 한국어 모음 지각." *음성과학*, 1, 201-211.
- Flege, J. E., M. Munro & R. A. Fox. 1994. "Auditory and categorical effects on cross-language vowel perception." *Journal of Acoustical Society of America*, 95(6), 3623-3640.
- Fox, R. A. 1982. "Individual variation in the perception of vowels: Implications for a perception-production link." *Phonetica*, 39, 1-22.
- \_\_\_\_\_. 1983. "Perceptual structure of monophthongs and diphthongs in English." *Language and Speech*, 26, 21-59.
- Fox, R. A., J. E. Flege & M. J. Munro. 1995. "The perception of English and Spanish vowels by native English and Spanish listeners: A multidimensional scaling analysis." *Journal of Acoustical Society of America*, 97(4), 2540-2551.
- Fox, R. A. & M. D. Trudeau. 1988. "A multidimensional scaling study of esophageal vowels." *Phonetica*, 45, 30-42.
- Jenkins, J. 1987. "A selective history of issues in vowel perception." *Journal of Memory and Language*, 26, 542-549.
- Klatt, D. H. 1980. "Software for a cascade/parallel formant synthesizer." *Journal of Acoustical Society of America*, 67, 971-995.
- Liberman, A. M., & I. G. Mattingly. 1985. "The motor theory of speech perception revised." *Cognition*, 21, 1-36.
- Peterson, G. E. & H. L. Barney. 1952. "Control methods used in a study of the vowels." *Journal of Acoustical Society of America*, 24, 175-184.
- Rakerd, B. & R. R. Verbrugge. 1985. "Linguistic and acoustic correlates of the perceptual structure found on an individual differences scaling study of vowel." *Journal of Acoustical Society of America*, 77(1), 296-301.
- Shepard, R. N. 1972. "Psychological representation of speech sounds." In E. D. David & D. P. Dence (eds.), *Human Communication: A Unified View*, McGraw Hill: New York, 67-113.
- Singh, S. & G. R. Woods. 1971. "Perceptual structure of 12 American English vowels." *Journal of Acoustical Society of America*, 49, 1861-1866.
- Stevens, K. N., A. M. Liberman, M. Studdert-Kennedy & S. E. G. Öhman. 1969. "Cross-language study of vowel perception." *Language and Speech*, 12, 1-23.
- Terbeek, D. 1977. "A cross-language multidimensional scaling study of vowel perception." *UCLA Working Papers in Phonetics*, No. 37.

▲ 최양규

울산시 울주군 웅촌면 곡천리 산 72-10 (우: 689-870)

춘해대학 유아특수치료교육과

Tel: +82-52-270-0221, 0220 Fax: +82-52-270-0209

E-mail: ygchoi@choonhae.ac.kr