

한국어 화자의 영어 양순음 /b/와 순치음 /v/ 식별에서 시각 단서의 효과

The Effect of Visual Cues in the Identification of the English Consonants /b/ and /v/ by Native Korean Speakers

김 윤 현¹⁾ · 고 성 룡²⁾ · Valerie Hazan³⁾

Kim, Yoon hyun · Koh, Sungryong · Hazan, Valerie

ABSTRACT

This study investigated whether native Korean listeners could use visual cues for the identification of the English consonants /b/ and /v/. Both auditory and audiovisual tokens of word minimal pairs in which the target phonemes were located in word-initial or word-medial position were used. Participants were instructed to decide which consonant they heard in 2 x 2 conditions: cue (audio-only, audiovisual) and location (word-initial, word-medial). Mean identification scores were significantly higher for audiovisual than audio-only condition and for word-initial than word-medial condition. Also, according to signal detection theory, sensitivity, d' , and response bias, c were calculated based on both hit rates and false alarm rates. The measures showed that the higher identification rate in the audiovisual condition was related with an increase in sensitivity. There were no significant differences in response bias measures across conditions. This result suggests that native Korean speakers can use visual cues while identifying confusing non-native phonemic contrasts. Visual cues can enhance non-native speech perception.

Keywords: non-native speech perception, visual cue, bilabial, labio-dental

1. 서론

유성음인 양순 파열음 /b/와 순치 마찰음 /v/는 한국어에 없는 음소이다. 한국어를 모국어로 하는 화자(이하 한국어 화자)들은 두 음소를 한국어 연음 /p/로 지각하는 비율이 높아 잘 변별하지 못한다(예, Schmidt, 1996). 이 음소들은 조음 과정에서 눈에 띄는 정도로 입 모양 움직임에 차이가 있다. 파열음인 /b/는 다물었던 두 입술을 터뜨려 나오는 말소리이고, /v/는 마찰음으로 윗니를 아래 입술에 살짝 댄 채로 공기를 그 좁은 틈으로 새나오게 해서 만드는 말소리이다. 외국어 말소리를 들을 때 발음하고 있는 입 모양도 볼 수 있다면 말소리만 들

을 때에 비해 어떤 말소리인지 더 잘 식별할 수 있을까? 이 연구는 한국어 화자에게 영어 음소 /b/와 /v/를 식별하도록 할 때, 발음하고 있는 사람의 얼굴을 함께 보여 주는 것이 식별에 효과적인지, 그리고 시각 단서가 식별에 미치는 영향이 낱말에서 음소의 위치(단어 머리 vs. 단어 내)에 따라 차이가 있는지를 조사하였다.

영아는 생후 4, 5 개월이 되면 말소리에 맞는 얼굴의 움직임을 옳게 탐지해낼 수 있다(Kuhl & Meltzoff, 1982). 연구자들은 /a/ 또는 /i/ 모음을 들려주면서 모니터에 각 모음을 조음하는 두 얼굴을 제시한 다음, 영아들이 어느 쪽을 보는지 조사하였다. 영아들은 들려진 모음에 맞게 조음하는 얼굴 쪽을 더 오래 응시하였다. 이는 말소리 습득 과정에서 청각 단서뿐 아니라 얼굴, 입의 움직임 같은 시각 단서도 함께 학습됨을 시사한다. 이렇게 학습된 시각 단서들은 청각 단서만으로 말소리 식별이 어려울 때 도움이 된다. 예를 들어, 말소리가 주변 소음 때문에 잘 들리지 않을 때 시각 단서들이 말소리 식별을 향상시킨다(Sumby & Pollack, 1954). 청각에 손상이 있는

1) 서울대학교, toithaka@gmail.com

2) 서울대학교, koh@snu.ac.kr, 교신저자

3) University College London, v.hazan@ucl.ac.uk

접수일자: 2012년 7월 25일

수정일자: 2012년 9월 3일

게재결정: 2012년 9월 4일

청자들에게 청각 단서와 조음 관련 시각 단서를 함께 제시하면 청각 단서만 제시할 때보다 말소리 식별을 잘한다(Grant & Seitz, 1998). 혼동이 되는 외국어 말소리의 경우에도 시각 단서와 소리 단서를 함께 제시하면 말소리만 제시한 경우에 비해 더 잘 식별한다(Hazan, Sennema, Faulkner, Ortega-Llebaria, Iba, & Chung, 2006). Hazan 등(2006)은 스페인어 화자(스페인어가 모국어인 화자)와 일본어 화자(일본어가 모국어인 화자)가 외국어 음소인 영어 양순음(/b/, /p/)과 순치음(/v/)을 식별할 때 시각 단서에 담긴 정보에 얼마나 민감한지 조사하였다. 이 외국어 화자들은 말소리만 들려 줄 때보다 조음 중인 얼굴을 함께 보여줄 때 영어 음소들을 더 잘 식별하였다.

맥거크 효과(McGurk & MacDonald, 1976)는 청자들이 말소리를 식별할 때 이런 시각 단서를 자동적으로 사용한다는 것을 보여 준다. 청각 단서와 시각 단서가 서로 다른 말소리에 대한 정보를 제공한 경우에, 청자는 말소리를 청각 정보에 따라서만 식별하지를 못한다. 예를 들어, /ga/를 조음할 때의 입 모양(시각 단서)과 /ba/의 말소리(청각 단서)를 함께 제시하면, 청자는 /da/로 식별하는 비율이 높다. 그런데 맥거크 효과의 크기는 청자가 사용하는 모국어에 따라 차이가 있다(예, 최양규, 2004; Sekiyama, 1997; Sekiyama & Tohkura, 1991). 일본어, 중국어, 한국어 화자들은 영어 화자들에 비해 맥거크 효과가 작았다. 연구자들은 이들 언어를 모국어로 하는 화자들이 영어 화자들에 비해 언어적, 또는 문화적 이유로 언어 습득 과정에서 조음 관련 시각 단서를 무시하도록 훈련되었기 때문이라고 주장하였다. 맥거크 효과 실험 패러다임을 이용한 연구(Sekiyama & Burnham, 2008)에 따르면, 영어 화자들의 경우에 6세에서 11세에 이르기까지 시각 단서의 영향이 점점 커졌지만, 일본어 화자들은 이런 발달적 변화를 보이지 않았다. 청자가 시각 단서를 무시한 것은, 모국어 말소리의 시각 단서가 충분한 정보를 제공하지 못하거나 대화 중 얼굴을 잘 보지 않아서 조음 관련 시각 단서를 잘 처리하지 못하게 되기 때문일 수 있다(Sekiyama, 1997). 예를 들어, 중국어의 경우 성조에 따라 의미가 달라지는데, 시각 단서는 이에 대한 정보를 제공해주지 못한다. 이런 점 때문에 중국어 화자들은 언어 습득 과정에서 시각 단서보다 청각 단서에 더 주의하도록 훈련되었을 가능성이 높다. 한편, 동양 문화권에서는 자신보다 지위가 높은 사람의 얼굴을 뺨히 보는 것을 무례한 것으로 여기기 때문에, 이의 영향으로 사람의 얼굴을 정면으로 보는 것을 피한다. 이 때문에 말소리를 식별할 때 조음 관련 시각 정보와 청각 정보를 결합시켜 처리하는 능력을 발달시키지 못했을 수도 있다(Sekiyama & Tohkura, 1991).

모국어에 없는 말소리 대립의 지각에서 시각 단서의 활용 정도가 모국어에 따라 다를까? 스페인어 화자와 일본어 화자를 대상으로 영어 /b/, /v/ 음소 식별, 그리고 일본어 화자와 한국어 화자를 대상으로 영어 /l/, /r/ 자음 식별을 조사한 연구

(Hazan 등, 2006)에 따르면, 맥거크 효과의 경우처럼 언어에 따라 시각 단서의 활용에 차이가 있다. /b/-/v/ 음소 대립의 경우에 두 언어 화자들은 앞서 말했듯이 시각 단서가 같이 제시되었을 때 그렇지 않을 때보다 잘 식별하였지만, 스페인어 화자의 식별이 일본어 화자보다 전반적으로 나았을 뿐 아니라 식별할 때 시각 단서의 효과가 컸다. 자극으로 시각 단서만 제시하였을 때 스페인어 화자는 음소 식별률이 82.3%였던 반면에 일본어 화자는 우연 수준인 50.7%였다. 실험 참여자들의 모국어 말소리와 영어 말소리의 관련성 때문에 이러한 차이가 생겼을 수 있다. 영어 음소 /b/의 경우에 스페인어와 일본어에 음성이나 조음 동작이 비슷한 음소 또는 이음이 있다. 하지만 /v/의 경우는 다르다. 두 언어에 영어 /v/와 일치하는 음소는 없지만, 스페인어에는 조음 동작이 영어 /v/와 비슷한 무성음 /f/가 있다. 반면에 일본어에는 조음 동작이 비슷한 음소가 없다.

Hazan 등(2006)은 모국어 말소리와 영어 말소리의 관련성이 영향을 미친다면 영어 /r/과 /l/을 식별할 때에 일본어 화자보다 한국어 화자에게 시각 단서가 더 효과적일 것이라고 예측하였다. 그 근거로, 일본어에는 주로 탄설음 [r]로 실현되는 유음이 하나 있을 뿐이어서 영어의 유음들을 받음할 때와 같은 조음 동작을 요구하는 말소리가 없지만, 한국어에는 음성이 [r], [l], [n]로 실현되는 유음들이 있어 한국어 화자들이 설측음 [l]을 별개의 음소로 사용하지 않아도 영어 /l/을 받음할 때와 같은 조음 동작을 경험한다는 점을 들었다. 그런데 그들의 실험 결과에 따르면, 시각 단서의 효과는 두 언어 사용자 모두에게서 유의하지 않았다. 한국어 화자들이 일본어 화자에 비해 전반적으로 높은 식별률을 보여 주었지만, 두 언어 화자 모두 시각 단서를 음성 단서와 함께 제시하여도 음성 단서만 제시한 경우에 비해 식별률이 유의하게 향상되지 않았다. 이 결과는 모국어와의 관련성 외에 시각 단서 활용에 영향을 미치는 다른 요인이 있을 가능성을 시사한다. 조음하는 동안 입이나 혀의 모양 등의 시각 단서가 잘 보이는지 여부가 /r/과 /l/의 식별에 영향 주었을 수 있다.

이 연구의 목적은 한국어 화자가 모국어에 없는 외국어 말소리를 식별할 때 시각 단서를 활용할 수 있는지 살펴보는 것이다. 한국어 화자들이 영어 유음 /r/과 /l/을 식별할 때 시각 단서는 별로 도움이 되지 않았다(예, Hazan 등, 2006). 이 연구에서는 /r/, /l/보다 조음 동작의 차이가 시각적으로 더 분명하게 드러나는 /b/와 /v/를 이용하여 한국어 화자들이 이 두 음소를 식별할 때 시각 단서를 활용할 수 있는지 조사하였다. 시각 단서의 활용에 식별하려는 외국어 말소리와 모국어 말소리의 음성 또는 조음 자질들의 관계가 중요하다면, 한국어는 유성음인 양순음이 없다는 점에서 일본어에 비해 영어와 더 다르기 때문에 한국어 화자의 경우에는 일본어 화자에 비해 시각 단서 효과가 없거나 작아야 할 것이다. 이와 달리, 시각

단서의 활용에 모국어와의 관계보다 시각 단서의 두드러짐(salience)이 중요하다면, 음성과 함께 시각 단서를 제시하였을 때 식별이 더 잘 될 것이다.

또, 식별해야 할 음소가 단어의 머리에 나올 때와 단어의 중간에 나올 때 시각 단서의 활용 정도가 같은지 다른지 검토하였다. 단어 내 위치처럼 음소가 놓인 환경에 따라 같은 음소라도 식별률이 다르다. Logan, Lively, Pisoni(1991)가 일본어 화자들에게 단어를 들려 주고 영어 /t/과 /l/을 식별하도록 한 결과에 따르면 두 음소가 단어의 끝에 단독으로 나올 때(tower - towel) 가장 잘 식별하였고, 단어의 앞에서 다른 자음에 연이어 나올 때(breed - bleed) 식별을 가장 못하였다. 단어의 맨 앞에 단독으로 음소가 나올 때에도(rock - lock) 단어의 끝에 단독으로 나오는 경우보다 식별이 좋지 않았다. 소리만 듣고 어떤 음소인지 식별해야 할 때 그 음소가 단어 내 어디에 위치하는가에 따라 식별률이 다른 것처럼, 시각 단서의 제시에 따른 식별의 향상 여부도 표적 음소가 단어의 어디에 있는가에 따라 다를 수 있다.

이 연구에서는 시각 단서의 효과를 측정하기 위해 식별률 뿐 아니라, 신호탐지 이론(Macmillan & Creelman, 2005)에 따라 적중률(hit rate)과 오경보율(false alarm rate)을 근거로 민감도(sensitivity)와 반응 편중(response bias) 값을 구하였다. /b/ 음소가 탐지되어야 할 신호라고 가정하면, 자극 /b/를 /b/로 옳게 식별하는 비율이 적중률, 자극 /b/를 /v/로 틀리게 식별하는 비율이 실패율(miss rate), 자극 /v/를 /b/로 잘못 식별하는 비율이 오경보율, 자극 /v/를 /b/가 아닌 것으로(즉, /v/로) 옳게 식별한 비율이 정확 기각률(correct rejection rate)이다. 일반적으로 식별률은 표적 자극에 대해 옳게 반응한 비율(적중률)만을 말한다. 이 수치가 같더라도 표적 자극이 아닌 것을 표적 자극으로 잘못 식별하는 비율은 다를 수 있다. 이를 오경보율이라 하는데, 식별률이 같더라도 오경보율이 다르면 신호에 대한 탐지 능력이 비슷하다고 볼 수 없다. 적중률이 높거나 오경보율이 낮을수록 신호에 대한 민감도 d' 은 커진다. 그리고, 반응 편중 값 c 가 양(+)이면 오경보율이 낮은 대신 적중률도 낮음을 의미한다. 이를 통해 신호탐지 능력과 별개로 탐지자의 반응 경향을 확인할 수 있다.⁴⁾

2. 방법

2.1 참여자

서울대학교에서 심리학 개론을 수강하는 18명(남자 11명, 여자 7명, 18세-25세)의 대학생이 과목 이수를 위한 의무로 실

험에 참가하였다. 이들은 모두 한국어가 모국어였다. 영어권 국가에 체류한 경험이 있는 학생이 두 명이었는데, 둘 모두 한 달 이상 체류하지는 않았다. 영어 능력 검사(TEPS)에서 취득한 점수는 530점에서 914점까지 다양하였다.

2.2 기구

방음 설비된 실험실에서 실험하였다. 외장형 사운드카드(Sound Blaster Audigy 2 NX)를 장착한 데스크톱 컴퓨터로 자극을 제시하고 반응을 수집하였다. 실험 프로그램은 DMDX 소프트웨어로 작성되었다(Forster, 2003). 참여자들은 사운드카드에 연결된 헤드폰(Sennheiser HD212Pro)을 통해 음성 자극을 들었고, 모니터를 통해 해당 자극을 발음하고 있는 화자의 얼굴 영상을 보았다. 반응 수집을 위해 키보드의 두 키('v', 'b')가 사용되었다. 영상 자극은 영국 UCL Speech, Hearing and Phonetic Sciences 학과의 녹음실에서 Canon XL-1 DV 캠코더로 녹화되었으며, 자극 음성은 Bruel and Kjaer type 4165 마이크를 이용해 캠코더로 녹음되었다.

2.3 자극

/b/-/v/ 최소 대립쌍 단어 16개가 사용되었다. 두 음소가 단어의 앞에 위치하는 대립쌍이 8개였고(ballet-valet, bat-vat, bend-vend, best-vest, bigger-vigour, boat-vote, bowel-vowel, burble-verbal), 단어의 중간에 위치하는 대립쌍이 8개였다(cupboard-covered, dribble-drivel, fibre-fiver, hobble-hovel, knobble-novel marble-marvel, rebel-revel, sabre-savour). 자극들은 이전 연구(Hazan, Sennema, Iba, & Faulkner, 2005)를 위해 다음과 같이 만들어졌다. 4명(남자 2명, 여자 2명)의 영어(British English) 화자들이 각각의 단어를 발음하였다. 단어별로 조음 동작 관련 시각 정보가 담긴 비디오 파일을 만들었으며, 각 파일의 시작과 끝 장면은 중립적인 얼굴 표정이 표시되도록 편집하였다. 각 자극의 영상은 250*300 픽셀, 25 프레임/초, 그리고 음성은 표본추출률 22,050 Hz로 만들었다. 음성만 제시되는 자극은 영상과 음성이 동시에 제시되는 자극에서 음성만 추출한 것으로, 이 자극이 제시될 때 모니터에는 검은 색 화면이 표시되었다.

2.4 절차

얼굴 영상 제시 여부(A 조건-음성만 제시, AV 조건-얼굴과 음성을 함께 제시)와 표적 음소 위치(i 조건-단어의 첫 음소가 표적 자극, m 조건-단어의 두 번째 음절 첫 음소가 표적 자극)의 효과를 조사하였다. A 조건과 AV 조건의 제시 순서는 참여자 간 상쇄되도록 하였고, A 조건(또는 AV 조건) 내에서 i 조건, m 조건 순으로 자극을 제시하였다. 참여자는 빈 영상+음성 또는 얼굴 영상+음성으로 제시되는 단어의 처음 또는 중간의 음소가 /b/인지 /v/인지 판단하여 해당키를 누르라는 지

4) 민감도 $d' = z(f) - z(h)$,

$$\text{반응편중 } c = \frac{z(h) + z(f)}{2}$$

(z: 누적 정규분포의 역함수, h: 적중률, f: 오경보율)

시를 받았다. 응답을 하지 않은 경우에는 자동적으로 다음 자극이 제시되었다. 각 조건 내에서 자극들은 무선(random)으로 제시되었다. 조건 별로 세 번의 연습 시행 후 본 시행이 시작되었다. 실험은 약 20 분 걸렸다.

3. 결과

실험 결과, 참여자들의 A 조건 평균 식별률은 78.4%(표준편차 17.2)였고, AV 조건의 평균 식별률은 83.1%(표준편차 20.3)이었다. 참여자들은 조음 동작 관련 시각 단서를 확인할 수 있는 경우 그렇지 않은 경우에 비해 /b/, /v/ 음소를 더 잘 식별하였다. 모든 실험 조건에서 개인차가 크긴 했지만, 시각 단서가 제공되고 표적 음소가 어두에 위치할 때 평균 식별률이 가장 높았으며(평균 86%, 표준편차 20.4), 음성만 제공되고 음소가 어중에 위치할 때 평균 식별률이 가장 낮았다(평균 75.3%, 표준편차 18.2) <그림 1>.

반복측정 이원변량분석으로 시각 단서 제시 여부와 단어 내 음소 위치가 /b/와 /v/의 식별에 미친 효과를 검토한 결과, 시각 단서 주효과와 음소 위치 주효과가 있었다 [시각 단서: $F(1, 17) = 4.9, p < .05$, 음소 위치: $F(1, 17) = 13.6, p < .05$]. 두 변인 간 상호작용 효과는 유의하지 않았다 [$F(1, 17) = .15, n.s.$]. 이 분석으로 조음 동작 단서가 참여자들의 음소 식별에 도움을 주고, 단어 중간보다는 초두에 나올 때 음소가 더 잘 식별됨을 알 수 있었다. 또, 표적 음소를 식별할 때 음소가 단어의 어디에 있는가에 따라 시각 단서가 더 영향을 주거나 하지는 않음을 확인하였다.

시각 단서 유무에 따른 식별률의 차이가 음소에 대한 민감도 차이와 관련 있는지 혹은 참여자들 반응 기준의 편중과 관련 있는지 살펴보았다. 민감도와 반응 편중의 계산에는 베이 지안 분석 방법을 썼다(Dennis, Lee, & Kinnell, 2008; Lee, 2008).⁵⁾ 그 결과, 평균 민감도는 음소가 어두에 나오고 청각 단서와 시각 단서가 함께 제시되었을 때 가장 높았고, 음소가 어중에 나오고 말소리만 제시된 조건에서 가장 낮았다<표 1>. 민감도와 반응 편중에 대하여 시각 단서 유무(음성만 제시, 음성과 얼굴 영상 동시 제시), 음소 위치(단어의 초두, 단어의 중간)가 변인인 반복측정 이원변량분석을 실시하였다. 민감도에서 시각 단서 유무와 음소 위치의 주효과가 각각 유의하였다 [시각 단서: $F(1, 17) = 7.14, p < .05$, 음소 위치: $F(1, 17) = 11.39, p < .01$]. 시각 단서가 제시될 때(AV 조건) 또는 표적 음소가 어두에 나올 때(i 조건) 민감도가 높아졌다. 민감도에 대한 시각 단서와 음소 위치 간 상호작용 효과는 유의하지 않

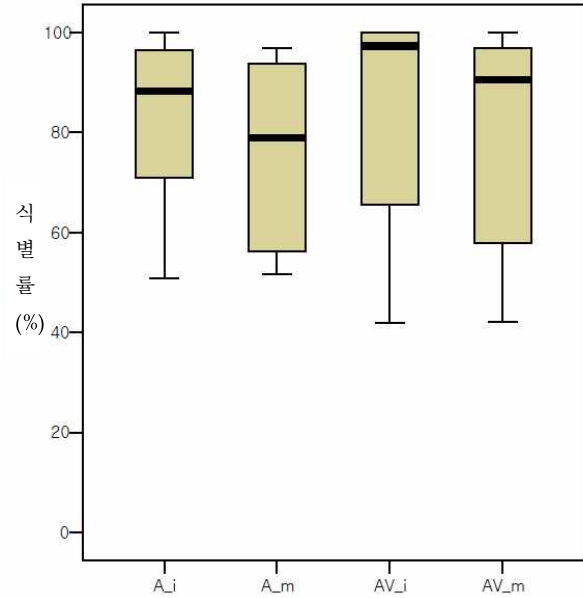


그림 1. 조음 관련 시각 단서 유무와 표적 음소의 위치에 따른 한국어 화자의 영어 음소 /b/, /v/의 식별률 (A: 음성 단서만, AV: 음성 단서+시각 단서, i: 단어 초두, m: 단어 중간)

Figure 1. Mean identification scores in A_i (audio-only, word-initial), A_m (audio-only, word-medial), AV_i (audio-visual, word-initial), and AV_m (audio-visual, word-medial) conditions

았다 [$F(1, 17) = .12, n.s.$]. 식별률의 차이가 음소들에 대한 민감도 이외에 참여자들의 반응 결정 기준 때문에 생긴 것인지 보기 위해 반응 편중 값을 분석하였다. 두 요인의 주효과는 모두 유의하지 않았다 [시각 단서: $F(1, 17) = .65, n.s.$, 음소 위치: $F(1, 17) = .65, n.s.$]. 유의한 상호작용 효과도 발견되지 않았다 [$F(1, 17) = 1.59, n.s.$]. 이 결과는, 앞서 밝힌 시각 단서 유무나 음소 위치에 따른 식별률 차이가 반응 편중이 아니라 참여자들의 음소에 대한 민감도 차이와 관련 있음을 시사한다.

표 1. 민감도와 반응편중의 평균(괄호 안, 표준편차)
Table 1. Mean sensitivity d' and response bias c (SD in parenthesis)

	A		AV	
	i	m	i	m
민감도	1.60 (.89)	1.29 (.97)	1.89 (1.08)	1.60 (1.11)
반응편중	.05 (.22)	.15 (.29)	.14 (.21)	.12 (.23)

5) Lee(2008)에서 한 대로 민감도와 반응편중의 사전(prior) 분포로 정규분포를 가정했고, 표집분포는 이항분포를 가정했다. 이 수치로 피험자 별 민감도와 반응 편중의 분포를 추정했고 이 분포의 평균을 대표치로 삼아 통계처리를 했다.

4. 결론 및 논의

한국어 화자들은 영어 음소 /b/, /v/를 혼동한다. 한국어 화자인 참여자들은 이 혼동하기 쉬운 음소들이 단어 초두에 있을 때 단어 중간에 있을 때보다 잘 식별하였다. 또, 음소가 어떤 위치에 있는 조음 관련 시각 단서가 제공되면 음소의 식별이 청각 단서만으로 식별할 때보다 나왔다. 하지만, 시각 단서가 음소의 위치에 따라 식별에 도움이 더 되거나 하지는 않았다. 즉, 청각 단서의 활용이 상대적으로 쉬운 음소 환경(어두에 위치할 때)과 어려운 음소 환경(어중에 위치할 때)에서 시각 단서를 제공했을 때 평균 식별률의 향상 정도는 별로 다르지 않았다(4.4% 대 5%). 그리고 식별률의 향상이 참여자들의 반응 편중 때문이 아니라 음소 위치와 시각 단서 유무에 따른 자극에 대한 민감도 변화와 관련이 있음을 확인하였다. 정리하면, 영어 음소 /b/와 /v/를 식별할 때 한국어 화자들도 시각 단서를 활용할 수 있다. 주의를 기울이기 힘든 단어 중간에 음소가 나올 때에도 시각 단서에서 말소리 정보를 얻을 수 있다.

이 결과는 한국어 화자들이 영어 음소 /r/과 /l/을 식별할 때 시각 단서를 활용하지 못한 것과 대비된다(예, Hazan 등, 2006). 음소에 따른 이러한 차이는 말소리 식별에서 시각 단서의 효과를 설명하기에 문화 요인설(Sekiyama & Tohkura, 1991)이 충분하지 않음을 시사한다. 언어 습득 과정에서 문화적 요인 때문에 시각 단서에서 얻은 정보를 청각 단서에서 얻은 정보와 결합시키는 능력에 결함이 생겼다면, /r/-/l/ 음소 대립의 경우와 마찬가지로 모국어에 없는 /b/와 /v/를 식별할 때에도 시청각 정보들을 결합하는 데 실패할 가능성이 높기 때문이다.

또, 모국어와 외국어 말소리의 관계가 시각 단서의 활용에 결정적으로 영향을 주는 것 같지도 않다. 일본어에 비해 한국어 양순음이 영어 양순음과 비슷한 점이 적음에도 불구하고, 음성과 함께 시각 단서를 제시하면 일본어 화자처럼 식별률이 유의하게 향상되었다. 반대로 유음의 경우에는 일본어와 달리 영어 /l/과 비슷하게 조음되는 이음이 한국어에 있음에도 불구하고, 두 언어 화자의 경우 모두 /r/과 /l/의 식별에서 시각 단서의 효과가 나오지 않았다(예, Hazan 등, 2006).

모국어와 외국어의 관련성 못지않게 시각 단서의 두드러짐 정도의 차이가 시각 단서의 활용 여부를 결정하는 중요한 요인일 수 있다. /r/과 /l/을 식별할 때 한국어 화자들이나 일본어 화자들은 입과 혀의 움직임에 주의를 기울였지만 그 차이를 변별하지 못했을지 모른다. 즉, 말소리를 식별할 때 시각 단서에 주의를 기울이지만 단서가 뚜렷이 드러나지 않아 정보를 얻는데 실패하였을 가능성이 있다. 이와 달리 /b/와 /v/는 조음 동작이 시각적으로 분명하게 드러나서 시각 단서의 차이를 알아채기 쉬웠을 수 있다.

차이가 두드러지지 않은 시각 단서도 훈련으로 그 차이를 변별할 수 있게 만들 수 있다. Hardison(2003)은 한국어가 모국어인 화자들을 대상으로 비디오를 이용하여 음성과 함께 조음하는 얼굴 영상을 보여 주어 /r/과 /l/을 식별하도록 훈련시켰다. 이러한 훈련으로, 시각 단서가 제공될 때 참여자들의 식별률이 더 높아졌다(다른 예, Hazan 등, 2005). 말소리 지각 훈련뿐 아니라 시각 단서에 대한 훈련을 할 필요가 있다. 음성 지각을 위한 훈련과 함께 조음 관련 시각 단서의 변별을 촉진하는 훈련을 같이 병행한다면, 외국어 화자와 얼굴을 맞대고 대화하는 상황에서 청각 단서로 식별이 쉽지 않을 때 시각 단서로 보완할 수 있을 것이다.

감사의 글

좋은 지적을 해 주신 심사위원님들께 깊이 감사드립니다. 이 논문은 2008년 정부(교육과학기술부)의 재원으로 한국학술진흥재단의 지원을 받아 수행된 연구입니다. [인문사회 분야 : KRF-2008-356-H00004]

참고문헌

- Choi, Yang-Gyu (2004). The influences of visual information in place of articulation on speech perception: The McGurk effect in Korean subjects. *Journal of Speech & Hearing Disorders*, 13, 117-128.
- (최양규 (2004). 조음위치에 관한 시각정보가 말소리 지각에 미치는 영향: 한국인의 McGurk 효과. *言語治療研究*, 13, 117-128.)
- Dennis, S., Lee, M. D. & Kinnell, A. (2008). Bayesian analysis of recognition memory: The case of the list-length effect. *Journal of Memory and Language*, 59, 361-376.
- Forster, K. I. & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods*, 35, 116-124.
- Grant, K. W. & Seitz, P. F. (1998). Measures of auditory-visual integration in nonsense syllables and sentences. *The Journal of the Acoustical Society of America*, 104, 2438-2450.
- Hardison, D. M. (2003). Acquisition of second-language speech: Effects of visual cues, context, and talker variability. *Applied Psycholinguistics*, 24, 495-522.
- Hazan, V., Sennema, A., Iba, M. & Faulkner, A. (2005). Effect of audiovisual perceptual training on the perception and production of consonants by Japanese learners of English. *Speech Communication*, 47, 360-378.
- Hazan, V., Sennema, A., Faulkner, A., Ortega-Llebaria, M., Iba,

- M. & Chung, H. (2006). The use of visual cues in the perception of non-native consonant contrasts. *The Journal of the Acoustical Society of America*, 119, 1740-1751.
- Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A. & Siebert, C. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition*, 87, 47-57.
- Kuhl, P. K. & Meltzoff, A. N. (1982). The bimodal perception of speech in infancy. *Science*, 218, 1138-1141.
- Lee, M. D. (2008). BayesSDT: Software for Bayesian inference with signal detection theory. *Behavior Research Methods*, 40, 450-456.
- Logan, J., Lively, S. & Pisoni, D. B. (1991). Training Japanese listeners to identify English /r/ and /l/: A first report. *The Journal of the Acoustical Society of America*, 89, 874-886.
- Macmillan, N. & Creelman, C. (2005). *Detection theory: A user's guide*. Mahwah, NJ: Lawrence Erlbaum Assoc. Inc.
- McGurk, H. & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264, 746-748.
- Schmidt, A. M. (1996). Cross-language identification of consonants. Part 1. Korean perception of English. *The Journal of the Acoustical Society of America*, 99, 3201-3211.
- Sekiyama, K. (1997). Cultural and linguistic factors in audiovisual speech processing: the McGurk effect in Chinese subjects. *Perception & Psychophysics*, 59, 73-80.
- Sekiyama, K. & Burnham, D. (2008). Impact of language on development of auditory visual speech perception. *Developmental Science*, 11, 306-320.
- Sekiyama, K. & Tohkura, Y. (1991). McGurk effect in non-English listeners: Few visual effects for Japanese subjects hearing Japanese syllables of high auditory intelligibility. *The Journal of the Acoustical Society of America*, 90, 1797-1805.
- Sumby, W. H. & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *The Journal of the Acoustical Society of America*, 26, 212-215.

• **김윤현(Kim, Yoon hyun)**

서울대학교 심리학과
서울시 관악구 관악로 1
Email: toithaka@gmail.com
관심 분야: L2 말소리 지각

• **고성룡(Koh, Sungryong) 교신저자**

서울대학교 심리학과
서울시 관악구 관악로 1
Tel: 02-880-9107

Email: koh@snu.ac.kr

관심 분야: 언어심리학, 읽기 과정, 안구운동

• **Valerie Hazan**

UCL Speech Hearing and Phonetic Sciences
Chandler House, 2 Wakefield Street London WC1N 1PF
Tel: + 44 20 7679 4076 Fax: + 44 (0) 207 679 4010
email: v.hazan@ucl.ac.uk