

인공와우 시플레이션에서 나타난 건청인 영어학습자의 영어 말소리 지각

Korean ESL Learners' Perception of English Segments: a Cochlear Implant Simulation Study

임애리¹⁾ · 김다희²⁾ · 이석재³⁾

Yim, Ae-Ri · Kim, Dahee · Rhee, Seok-Chae

ABSTRACT

Although it is well documented that patients with cochlear implant experience hearing difficulties when processing their first language, very little is known whether or not and to what extent cochlear implant patients recognize segments in a second language. This preliminary study examines how Korean learners of English identify English segments in a normal hearing and cochlear implant simulation conditions. Participants heard English vowels and consonants in the following three conditions: normal hearing condition, 12-channel noise vocoding with 0mm spectral shift, and 12-channel noise vocoding with 3mm spectral shift. Results confirmed that nonnative listeners could also retrieve spectral information from vocoded speech signal, as they recognized vowel features fairly accurately despite the vocoding. In contrast, the intelligibility of manner and place features of consonants was significantly decreased by vocoding. In addition, we found that spectral shift affected listeners' vowel recognition, probably because information regarding F1 is diminished by spectral shifting. Results suggest that patients with cochlear implant and normal hearing second language learners would experience different patterns of listening errors when processing their second language(s).

Keywords: cochlear implant, simulation, vocoding, spectral shift, L2 listening

1. 서론

입장 현장에서는 한국어 듣기에 능숙한 인공와우 사용 청각장애인들이 영어 학습, 특히 듣기에서는 어려움을 겪는다는 사실이 널리 알려져 있다(박정란, 이현구, 임정완, 2008). 그러나 인공와우 사용자의 제 2 언어 청취에 대하여 체계적 실험을 통해 연구한 사례는 국내외를 막론하고 흔치 않은 실정이다. 예외적으로 Waltzman, Robbins, Green & Cohen (2003)의 연구에서는 5세 이전에 인공와우 수술을 받은 이중언어 사용자의 제 2 언어 발달을 살펴본 결과, 인공와우 사용자 역시 건청인과 비슷한 수준의 제 2 언어 말하기와 듣기 능력을 갖

출 수 있음을 검증하였다. Thomas, El-Kashlan & Zwolan (2008) 역시 인공와우를 사용하는 이중언어자와 단일언어자의 영어 말하기, 듣기 평가 결과를 비교하여 인공와우 수술이 제 2 언어 습득에 지대한 영향을 미치지 않는다는 것을 주장한 바 있다. 위와 같은 인공와우 사용자의 제 2 언어 습득 연구들이 주로 건청인과 인공와우 사용자 사이의 광범한 제 2 언어 능력 차이에 집중된 반면, 인공와우 사용자와 건청인이 보이는 제 2 언어 발화, 청취간의 세밀한 차이에 대한 실험 연구의 선례는 거의 없다고 해도 과언이 아닐 것이다. 인공와우 사용자들에 특화된 제 2 언어 교육에 대한 제언을 위해서는 인공와우 사용자의 제 2 언어 발화와 청취에 대한 정밀하고 체계적인 연구가 필요하다. 본 연구는 인공와우 사용자들의 영어 듣기 오류 양상을 추정하기 위한 예비 연구⁴⁾로, 건청인과 인

1) 연세대학교, 제1저자, erina21@nate.com

2) 연세대학교, 공동저자, daheekim0726@gmail.com

3) 연세대학교, 교신저자, scrhee@yonsei.ac.kr

접수일자: 2014년 8월 12일

수정일자: 2014년 9월 19일

게재결정: 2014년 9월 20일

4) 말소리 지각 실험이 가능한 정도의 영어교육을 받고 인공와우 시술 과정 및 예후(삽입된 인공와우의 구체적 모델, 시술 시기, 삽입완전도, 수술 후 청각훈련 여부 등)가 흡사한 인공와우 사용자들로 실험 대상을 통제하는 것은 매우 어려운 실정이다. 이와 같은 현실적 어려움을 고려하여, 예

공와우 사용자의 듣기 오류 양상을 추정하기 위하여 건청인에게 원 발화와 인공와우 시뮬레이션을 통과한 발화를 듣게 하고, 인공와우 시뮬레이션에 따른 건청인의 영어 듣기 양상을 살펴보았다.

인공와우 사용자의 듣기 양상을 추정하기 위한 시뮬레이션 연구는 주로 보코딩(vocoding) 여부 및 보코딩 과정에 사용된 채널 수(number of channels), 스펙트럼 이동(spectral shift) 여부 및 스펙트럼 이동의 정도(mm 단위)를 이용하여 수행되었다(Dorman, Loizou, & Rainey, 1997; Dorman, Loizou, Fitzke, & Tu, 1998; Fu & Shannon, 1999; Zhou, Xu, & Lee, 2010). 다양한 보코딩 방법 중 가장 보편적으로 사용되어 온 방법은 채널 보코딩으로, 채널 보코딩은 음성을 주파수 대역에 따라 다양한 구간으로 분리한 후 해당 대역의 에너지 분포를 모방한 백색 소음 등의 불규칙파 또는 사인파 등의 규칙파를 재조합하는 방식의 음성 처리 기법이다(Shannon, Zeng, Kamath, Wygonski & Ekeild, 1995). 이와 같은 보코딩의 원리는 서로 다른 주파역대에 반응하는 여러 개의 전극을 사용하여 와우에 전기적 자극을 주는 인공와우의 원리와 일치하기에, 본 연구에서는 인공와우 사용자들이 듣게 되는 말소리의 정인식률을 추정하기 위하여 보코딩을 사용하였다. 주파수 대역의 수가 많을수록 각 주파수대의 에너지 분포에 대한 정밀한 청각 정보가 제공되며, 그 결과 말지각의 정확성을 증가시키게 된다. 앞선 연구에서는 소음이 없는 조용한 상황에서 모어 듣기를 수행하는 경우, 대개 12개의 대역(인공와우에서는 전극의 수에 대응됨)에서 천정효과(ceiling effect)가 나타난다고 밝힌 바 있다(Dorman et al., 1998; Zhou et al., 2010).

인공와우에서 사용되는 전극의 수에 대응되는 보코딩 시의 주파수 대역 수와는 어느 정도 독립적으로, 인공와우 사용자의 말지각 능력은 인공와우의 삽입 정도에 따라 결정되기도 한다(Dorman et al., 1997; Fu & Shannon, 1999). 건청인의 말소리 듣기에서는 말소리의 특정 주파대에 해당하는 신호가 와우의 해당 부분을 자극하게 되나, 인공와우 사용자의 경우 와우의 삽입 정도에 따라 음성주파수와 와우의 자극 부분 사이에 정확한 일치이 일어나지 못하는 경우가 발견된다(Dorman et al., 1997; Shannon et al., 1995; Fu & Shannon, 1999; Zhou, et al., 2010). 인공와우가 완전히 삽입된 경우(full insertion) 음향 신호에 해당하는 주파수 영역이 와우의 해당 부분에 정확하게 자극을 부여하나, 인공와우가 충분히 삽입되지 못하면 상대적으로 낮은 주파수 영역의 정보가 상대적으로 높은 주파수 영역에 반응하는 와우의 부분을 자극하게 되어 실제 음성 정보와 신호에 의해 자극되는 와우의 부분 사이에 불일치가 야기된다. 요컨대, 임상상의 인공와우의 불완전 삽입(basal

shift)는 청음에 있어 실제 주파수와 와우의 자극 부분 사이의 불일치, 그리고 음향적으로는 스펙트럼 이동(spectral shift)으로 연계되어, 말지각력의 저해 요소가 된다. 주파수대역 수와 스펙트럼 이동을 조작하여 건청인의 말지각력을 연구한 Zhou et al.(2010)은 3mm 이상의 스펙트럼 이동이 일어나게 되면 완전히 삽입된 것과 유의한 차이의 말지각력을 보이게 됨을 밝힌 바 있다.

다만, 위의 연구 결과들이 공히 영어를 모어로 하는 화자들의 모어 듣기를 바탕으로 얻은 결과라는 점을 감안할 때, 한국어를 대상으로 한 연구 또는 외국어 학습자를 대상으로 한 연구는 거의 전무한 실정이다. 본 연구에서는 한국인 영어 학습자가 영어의 자모음을 지각할 때, 보코딩 여부와 스펙트럼 이동 여부에 따라 어떠한 영어 듣기의 어려움이 발생하는지를 살펴보았다. 보코딩과 스펙트럼 이동을 하지 않은 원 자극 듣기의 경우, 한국인 영어 학습자들은 모어와 대상 언어 차이에서 생기는 외국어 듣기의 어려움을 보일 것으로 예상된다(Bohn & Flege, 1990; Best & Tyler, 2007; Kuhl, 1991). 반면, 보코딩과 스펙트럼 이동으로 인해서 부여되는 영어 청취 양상은 인공와우 사용자와 건청인이 어떠한 언어 청취 습득의 어려움을 겪는지에 대한 정보를 제공할 것으로 기대한다.

2. 실험 방법

2.1 지각 자료

지각 자료는 K-SEC(Korean-Spoken English Corpus, 버전 2005. 01, 이석재, 이용주, 이숙향, 강석근, 2005) 말뭉치의 일부로, 노스다코타주(North Dakota) 출신의 영어 원어민 남성 화자 1인이 녹음한 발화를 사용하였다. 화자는 청각 및 언어 병력이 없었으며, 8년 이상의 음성언어평가 경력을 가진 언어 재활사가 화자의 발화를 평가한 결과, 화자에게 음성 및 공명의 문제가 없는 것으로 나타났다.

다음 지각 자료로는 “Say /Cau/, too” 형태의 이끔 문장 내에 위치한, C에 해당하는 단어 초성 위치에 분포하는 자음만으로 변별되는 20 개의 유의미 영어 단어(주: 자음 /dʒ/ 의

- 5) 본 연구 시료에서 성문폐쇄음(glottal stop), 유성치조마찰음(voiced alveolar fricative), 유성치경마찰음(voiced alveopalatal fricative), 연구개비음(velar nasal)이 제외되었는데, 성문폐쇄음은 어휘 초에서 나타날 수는 있으나 [t] 또는 [] 과 빈번히 교체되는 비변별음성인 관계로 청자들이 문자로 표기하기 어려울 것으로 판단하였기 때문이다. 부연하자면, 성문폐쇄음이 들어간 시료 [ʔau]를 듣고 청자가 “I” 로 응답한다면, 반응만으로는 청자가 성문폐쇄음을 듣지 못한 것인지, 청취하였으나 비변별음성으로 간주한 것인지를 판별하기 어려우므로, 성문폐쇄음은 실험 대상 자음으로 사용하지 않았다. 유성치조마찰음 /z/ 는 /zau/ 가 무의미 어휘이며 이와 비슷한 어휘가 없어 사용하지 않았다. 유성치경마찰음 /ʒ/ 과 연구개비음 /ɹ/ 은 영어 음운 규칙에 의해 어

비 연구 단계에서는 실제 음성신호를 인공와우 사용자들이 듣는 것과 흡사한 방식으로 조정된 후 건청인에게 들려주는 연구방법을 택하였다.

경우 /dʒaɪ/ 가 무의미 단어인 관계로, 실제 어휘인 jive, /dʒaɪv/ 로 대체)를 2회씩 발화, 녹음한 40개의 시료가 선정되었다. 자음 수는 20개로, 6개의 폐쇄음 (/p, t, k, b, d, g/), 7개의 마찰음 (/f, v, θ, ð, s, ʃ, h/), 2개의 파찰음 (/tʃ, dʒ/) 및 5개의 비음 및 공명 자음 (/m, n, w, ɹ, l/) 으로 구성되었다.

모음 지각 자료로는 [hVd]의 동일한 음운 환경에서 단모음의 자질만으로 변별되는 10개의 최소대립어휘들(heed, hid, head, had, heard, hod, hawed, hud, hood, who'd)을 2번씩 발화, 녹음한 20개의 시료가 선정되었다.

위 시료들을 프랏 (Praat Version 5.3; Boersma & Weenink, 2014)에서 대역 수를 12개로 고정된 후 소음기반으로 보코딩 (noise vocoding)하여 인공와우 시뮬레이션 자극을 합성하였다. 스펙트럼 이동의 경우, 스펙트럼이 전혀 이동되지 않은 0mm 시뮬레이션과 3mm의 스펙트럼 이동 및 이에 수반한 와우와 해당 주파수간의 간극(tonotopic distance)이 발생한 3mm 시뮬레이션으로 보코딩의 세부 조건을 조작하였다. 이와 같은 시뮬레이션의 결과, 원래 발화(원자극, 총 시료 수 60개)와 두 종류의 인공와우 시뮬레이션을 통과한 자극(각 시뮬레이션 당 시료 수 60개, 총 120개의 시료), 총 180개의 시료가 실험 자극으로 주어졌다. 본 실험 시료 작성을 위해 사용된 보코더의 총괄적 정보를 아래 <표 1>에, 보코더의 채널별 정보 및 각 채널 정보가 대략 어느 정도의 와우 위치에서 처리되는 신호에 대응하는지를 <표 2>에 요약하였다.

표 1. 보코더의 음향적 정보
Table 1. Acoustic specifications of the vocoder

항목	내용
최저 주파수	250Hz
최고 주파수	8700Hz
대역 수	12 개
엔벨로프 필터 주파수	600Hz
합성 필터 모형	고점형 (Peaked)
mm당 합성 필터 톨오프	24
필터 스펙트럼 이동량	(조건에 따라) 0mm 또는 3mm
고점 선정 시간 단위	0.06초
고점 선정 시간 단위의 중첩	0.005초

두에 나타나지 않아 사용하지 않았다.

표 2. 보코더 채널 정보 및 채널 정보에 대응하는 와우 위치
Table 2. Vocoder's channel information and the corresponding cochlear positions

대역	최소 주파수	중앙 주파수	최대 주파수	추정 와우 위치	
				0mm 이동시	3mm 이동시
1	250	305	367	7.25	10.25
2	367	438	518	9.12	12.12
3	518	610	715	11.00	14.00
4	715	834	969	12.87	15.87
5	969	1123	1298	14.75	17.75
6	1298	1498	1725	16.62	19.62
7	1725	1984	2278	18.49	21.49
8	2278	2613	2994	20.37	23.37
9	2994	3428	3922	22.24	25.24
10	3922	4484	5124	24.12	27.12
11	5124	5853	6682	25.99	28.99
12	6682	7626	8700	27.87	30.87

원 자극과 <표 1> 과 <표 2> 에 상술된 과정을 통하여 합성된 음성, “hawed”의 스펙트로그램과 모음의 50% 지점의 스펙트럼은 아래 <그림 1> 과 같다.

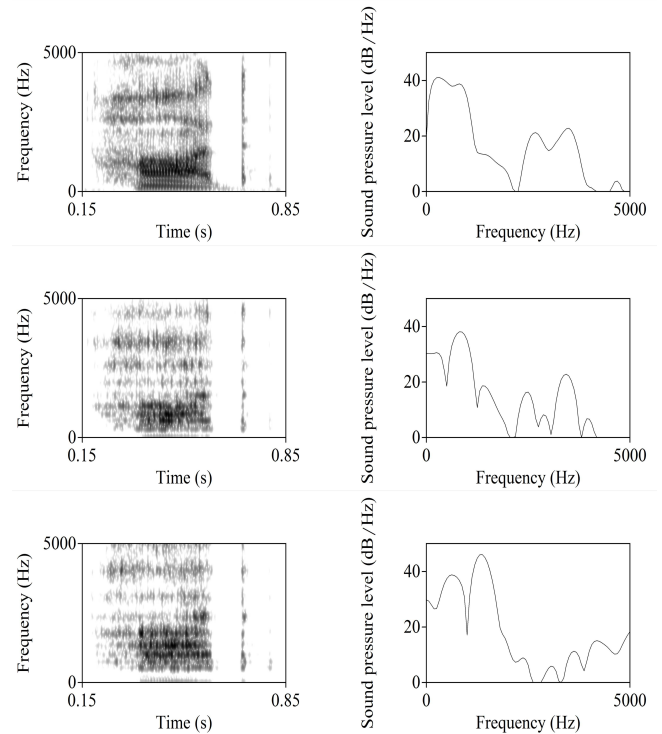


그림 1. 위로부터 “hawed” 에 해당하는 원 자극, 보코딩을 거친 자극, 보코딩과 3mm의 스펙트럼 이동을 거친 자극의 스펙트로그램과 모음 50% 부분의 스펙트럼

Figure 1. Spectrogram and spectrum at the midpoint of the vowel: From top to bottom, original stimulus, stimulus after 12-channel noise vocoding, and stimuli after 12-channel noise vocoding and 3mm spectral shift.

2.2 실험 절차

실험 대상은 말, 언어, 청각 장애의 병력이 없는 성인 10명 이었으며, 전원 영어권 국가 거주 경험이 1년 미만인 한국인 영어 학습자들로 구성되었다.

연구자가 제작한 소프트웨어를 이용, 피험자들에게 자극 단어의 말소리를 무작위로 재생하여 들려주었다. 피험자들은 내장 사운드카드를 이용하는 SMS-M1000 (Rev 1.3) 노트북에서 RP-DJ100 헤드폰을 통하여 자극을 청취하였으며, 실험 시 자극 크기는 70dB로 통일하였고, 재생 시 음량은 피험자 간에 균일하도록 노트북 자체의 음량 30% 수준으로 조절하였다. 실험은 평균 35 dBA의 소음이 있는 조용한 방에서 이루어졌으며, 피험자들은 개별 시료를 들은 후 대상 자음과 모음을 들리는 대로 받아 적도록 하였다.

자극은 3mm 시물레이션, 0mm 시물레이션, 원자극 순서로 피험자들에게 들려주었다. 전체 10인의 피험자 중 절반인 5명은 자음, 모음의 순서로, 나머지 5명은 모음, 자음의 순서로 대상 음소들을 청취하였다. 실험 전 모의 실험이나 연습은 실시하지 않았으나, 이끔 문장과 이끔 단어의 형태, 목표 음소 외의 다른 음소로 이루어진 예시 반응을 실험 전에 문자로 제공하여 실험에 대한 피험자들의 이해를 도왔다. 피험자들에게는 반응을 가능한 한 알파벳으로 기입하도록 권장하였으나 부득이한 경우 한글로 혹은 알파벳과 혼용하여 사용하는 것을 허용하였다.

2.3. 자료 분석

각 음소별로 화자가 의도한 음소와 피험자의 반응을 비교하여 정확도를 측정하였다. 구체적으로 어떠한 오류 양상이 일어나는지에 대한 일반화된 기술을 위해, 의도된 음소와 피험자의 반응의 변별자질(distinctive feature)을 기록하여 각 변별자질별 인식률을 계산하였다. 특히 모음의 경우, 피험자가 한글로 응답한 경우가 많아 한국어 모음의 변별자질과 화자가 발화한 영어 모음의 변별자질을 비교하였다. 모음 긴장성(tenseness)의 경우 피험자가 장음(예: 하아드, 하-드, 하~드)과 단음(예: 핫, 하드)으로 표기한 내용에 따라 장음을 긴장음으로, 단음을 이완음으로 간주6), 발화자가 발화한 음소와의 변별 자질 일치 여부를 비교하였다.

2.4. 통계 분석

청자들의 음소 인식 정확도 및 변별 자질 인식 정확도를 R

6) 장단 또는 긴장-이완성을 표기한 방식에는 청자별 개인차이가 있었다(힐 vs. 히드, 헛 vs. 히이드, 히드 vs. 히:드, 히드 vs. (길게) 히드 등의 표기 사용). 본 실험에 앞서 시행한 예비 실험에 참여한 5명의 피험자 전원 및 보고된 실험의 대상자 10인 전원이 표기를 통해 장단음을 구분한 바, 반응에서 나타난 장단음 여부를 포함하여 결과를 기술하는 것이 유의미할 것으로 간주하였다.

(R Core Team, 2014) 프로그램을 이용, 반복측정 분산분석(repeated measure ANOVA)으로 분석하였다. 종속변인은 음소 인식 또는 변별 자질 인식의 정확도이고, 조작변인은 시물레이션 방법(원자극, 보코딩과 0mm 스펙트럼 이동, 보코딩과 3mm 스펙트럼 이동), 음소의 변별 자질(자음의 경우 조음위치, 방법, 유·무성성, 모음의 경우 전·후설성, 고저성, 원순성 및 긴장성)이었다. 이와 같은 분석을 통해 음소단위의 청취 시 정확도의 차이를 야기하는 변인들을 밝혀내었으며, 나아가 각 변별 자질의 인식에 시물레이션이 어떠한 영향을 미치는지에 대하여 고찰하였다.

3. 연구 결과

3.1. 자음

청자별 자음 인식 정확도의 분포는 아래의 <그림 2>와 같았다.

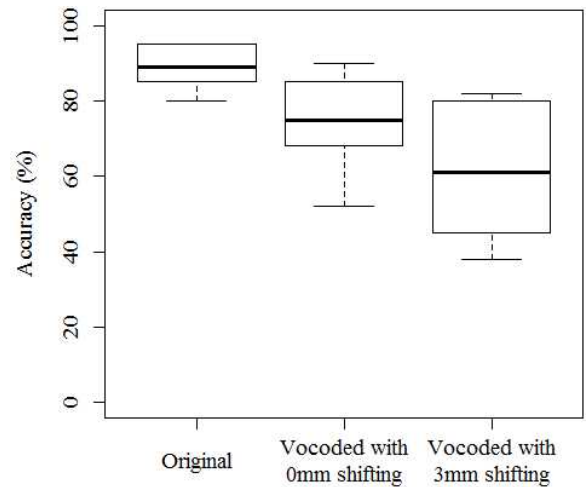


그림 2. 자음 인식 정확도

Figure 2. Accuracy of consonant identification

원 자극 인식 시, 청자들은 평균 88%(청자간 표준편차 6%)의 자음 정인식률을 보였다. 스펙트럼 이동 없이 12채널 보코딩을 거친 자극 인식시의 자음 인식률의 평균은 73%(청자간 표준편차 13%), 3mm 스펙트럼 이동과 12채널 보코딩을 거친 자극 인식률의 평균은 62%였다(청자간 표준편차 16%). 이는 시물레이션의 영향으로 말지각이 어려워질수록 말지각 정확도의 청자 간 변이가 증가함을 시사한다. 자음의 정인식률은 시물레이션에 따라 유의미한 변화를 보였다 ($F(2, 18) = 27.80, p < 0.001$).

원자극과 각 시물레이션 방식 간의 말지각 정확도의 차이를 통계적으로 검증하기 위해 정인식률에 대한 대응표본 t 검정(paired t-test)을 수행하였으며, 복수비교에 인한 오류 가능성을 최소화하기 위하여 유의 수준(alpha-level)을 0.01로 수정,

검증하였다. 원자극과 보코딩을 거친 자극 사이의 차이(정인식률 15% 감소, $t(9) = 5.14, p < 0.001$), 원자극과 보코딩 및 3mm 스펙트럼 이동을 거친 자극 사이의 차이(정인식률 26% 감소, $t(9) = 6.05, p < 0.001$), 그리고 보코딩을 거친 자극과 보코딩 및 3mm 스펙트럼 이동을 거친 자극 사이의 차이(정인식률 11% 감소, $t(9) = 3.53, p < 0.01$) 모두가 통계적으로 유의미한 것으로 나타났다.

시뮬레이션의 효과는 자음의 조음 위치, 조음 방법, 그리고 유무성성에 따라 다르게 나타났다. 자음의 조음 위치 및 시뮬레이션 방법별로 자음 인식 정확도를 요약한 바는 아래 <표 3> 과 같다.

표 3. 자음의 조음 위치별 정인식률
Table 3. Accuracy by place of articulation

	원자극	보코딩을 거친 자극	보코딩과 3mm 스펙트럼 이동을 거친 자극
양순음	98%	76%	61%
순치음	80%	68%	63%
치간음	45%	33%	23%
치경음	96%	87%	73%
구개음	87%	77%	72%
연구개음	100%	58%	38%
성문음	95%	95%	95%

원 자극과 스펙트럼 이동 없이 보코딩만을 거친 자극에 대한 자음 인식 정확도의 차이를 0.01의 유의수준에서 대응 표본 t 검정을 사용, 검증한 결과 평균 22%의 인식률 감소를 보인 양순음 조음 위치와($t(9) = 3.43, p < 0.01$) 평균 42%의 인식률 감소를 보인 연구개 조음위치에서($t(9) = 4.02, p < 0.01$) 통계적으로 유의미한 것으로 나타났다. 양순음의 경우 구강폐쇄음 /p, b/나 접근음 /w/보다 비음 /m/에서의 오류 증가가 두드러졌으며, 청자들은 보코딩을 거친 비음을 유성 폐쇄음 /b/ (3회) 또는 마찰음 /f (3회), v (2회)/로 오인식하는 경향을 보였다. 연구개음의 경우, 원자극은 청자들에 의해 100% 정확하게 인식된 반면, 스펙트럼 이동 없이 보코딩을 거친 자극은 무성 연구개 폐쇄음 /k/의 경우 /f, h, t, θ (각 1회)/ 등으로, 유성 연구개 폐쇄음 /g/의 경우 /b (3회), d (6회), q, s, θ, w (각 1회)/ 등으로 오인식되었다.

원자극과 3mm 스펙트럼 이동 및 보코딩을 거친 자극 사이의 인식률 차이는 양순음(Labial; 평균 37%의 인식률 감소, $t(9) = 4.80, p < 0.001$), 치경(Alveolar; 23%의 인식률 감소, $t(9) = 4.52, p < 0.01$) 그리고 연구개(Velar; 62%의 인식률 감

소, $t(9) = 7.32, p < 0.001$)에서 통계적으로 유의미한 것으로 나타났다. 유성 양순음인 /b, m/의 경우 조음 위치가 비슷한 무성음인 /f 또는 /p/로의 오인식이, 무성양순음 /p/의 경우 치경음인 /t/로의 오인식이 빈번히 나타났다. 치경음의 경우, 폐쇄음 /d/는 /t/로 오인식되었고, 설측음 /l/은 /l, w/ 등으로 오인식되었다. 흥미롭게도 /l/은 원자극 조건에서 역시 /l/로 오인식된 경우가 많았다(총 5회). 반면, /t/에서 /d/는 거의 나타나지 않았으며, /l/을 /l/로 오인식한 경우는 3mm의 스펙트럼 이동과 보코딩 조건에서만 나타났다. 이와 같은 인식의 비대칭성은 건청인과 인공와우 사용자가 영어 듣기에서 보이는 어려움이 다를 것이라는 함의를 갖는다. 나머지 조음 위치(순치음, 치간음, 경구개음 및 성문음)에서는 시뮬레이션의 효과가 통계적으로 무의미한 것으로 나타났다.

스펙트럼 이동 없이 보코딩만 거친 자극과 3mm의 스펙트럼 이동 및 보코딩을 거친 자극 사이의 정확도 차이는 모든 조음 위치에서 통계적으로 유의미하지 않았다⁸⁾.

자음의 조음 방법 및 시뮬레이션 방법별로 자음 인식 정확도를 요약한 바는 아래 <표 4> 와 같다.

표 4. 자음의 조음 방법별 정인식률
Table 4. Accuracy by manner of articulation

	원자극	보코딩을 거친 자극	보코딩과 3mm 스펙트럼 이동을 거친 자극
파열음	98%	77%	56%
파찰음	85%	70%	65%
마찰음	76%	69%	64%
비음	100%	60%	55%
접근음	100%	95%	85%
설측음	75%	70%	45%

조음 방법별 분석 역시 조음 위치별 분석과 흡사한 결과를 보였는데, 두 분석 모두에서 스펙트럼 이동이 자음 정인식률에 미치는 영향은 매우 한정적이었다. 조음 위치별로 분석하였을 경우, 구강폐쇄음에서만 3mm 스펙트럼 이동의 영향이 나타났는데(정인식률 21% 감소, $t(9) = 4.79, p < 0.001$), 이는 주로 /d (오인식 7회 증가), k (오인식 6회 증가), p (오인식 7회 증가)/ 의 오인식률 증가에 기인한 것이었다. 3mm의 스펙트럼 이동은 청자들로 하여금 /d/를 동일 위치에서 조음되는

- 7) 순치음: $t(9) = 1.91; ns$; 치간음: $t(9) = 2.87; ns$
경구개음: $t(9) = 2.59; ns$; 성문음: $t(9) = 0; ns$
- 8) 양순음: $t(9) = 1.91; ns$; 순치음: $t(9) = 0.61; ns$
치간음: $t(9) = 1.50; ns$; 치경음: $t(9) = 2.52; ns$
경구개음: $t(9) = 0.90; ns$; 연구개음: $t(9) = 2.23; ns$
성문음: $t(9) = 0; ns$

무성음인 /t/로 (0mm 스펙트럼 이동시 2회, 3mm 스펙트럼 이동시 6회), /k/와 /p/를 타 위치에서 조음되는 무성음 /t/(0mm 스펙트럼 이동시 1회, 3mm 스펙트럼 이동시 총 12회)로 오인식하게 하였다.

보코딩의 영향은 구강 폐쇄음과 비음의 경우에서 두드러졌는데, 원자극과 보코딩 자극(폐쇄음: 정인식률 21% 감소, $t(9) = 4.63, p < 0.01$; 비음: $t(9) = 4.00, p < 0.01$)은 물론, 원자극과 보코딩 및 3mm 스펙트럼 이동을 거친 자극(폐쇄음: $t(9) = 5.92, p < 0.001$; 비음: $t(9) = 5.01, p < 0.001$) 사이에서도 통계적으로 유의미한 정인식률 차이가 나타났다.

구강 폐쇄음의 경우, 상술한 바와 같이 유성 양순 폐쇄음과 유성 치경 폐쇄음은 동일하거나 흡사한 조음 위치의 무성음 또는 마찰음으로, 유성 연구개 폐쇄음은 /b, d/ 등으로 오인식되었다. 무성음의 경우 /p/는 /t/ (7회) 또는 /k/ (9회)로, /k/는 주로 /t/로 (7회) 오인식된 반면, /t/는 실험오류 또는 /i/로 반응된 3회를 제외하고는 오인식되지 않았다.

비음의 경우 /n/ (총 오인식 10회) 보다는 /m/ (총 오인식 24회)의 오인식 빈도가 높았다. 특히 양순 비음인 /m/의 경우, /b (9회), f (6회), v (3회), p (1회; 총 24회 중 19회에 해당) 등 흡사한 조음 위치의 다른 소리로 인식될 확률이 높았으나, 치경 비음 /n/은 각 1회씩 /b, d, h, j, l, t, dz/, 모음 /i/로 3회 오인식되었다.

마지막으로, 유무성성과 시뮬레이션 방식별 자음 인식률 분석 결과는 자음의 유무성성에 따라 시뮬레이션이 인식률 차이에 미치는 영향이 상이함을 시사하였다. 자음의 인식률을 자음의 유무성성과 시뮬레이션 방법에 따라 분류한 결과는 아래 <표 5> 와 같다.

표 5. 자음의 유무성성별 정인식률

Table 5. Accuracy by voicing

	원자극	보코딩을 거친 자극	보코딩과 3mm 스펙트럼 이동을 거친 자극
무성음	95%	89%	76%
유성음	83%	60%	51%

무성 자음의 인식에 있어서는 원자극과 스펙트럼 이동 없는 보코딩 자극 조건 간의 인식률 차이가 통계적으로 유의미하지 않았으나, 유성 자음 인식에 있어서는 통계적으로 유의미한 것으로 나타났다(무성음 정인식률 6% 감소, $t(9) = 2.02, ns$; 유성음 정인식률 23% 감소: $t(9) = 5.07, p < 0.001$). 자음 인식률은 유, 무성성을 막론하고 3mm 스펙트럼 이동을 동반한 보코딩에 의해 유의미하게 감소하는 양상을 보였다(무성음 정인식률 19% 감소, $t(9) = 5.09, p < 0.001$; 유성음 정인식률

32% 감소, $t(9) = 5.84, p < 0.001$). 스펙트럼 이동 없이 보코딩만 거친 자극과 3mm의 스펙트럼 이동과 보코딩을 거친 자극의 인식률은 무성자음 인식률에만 유의미한 변화를 야기했다(무성음 정인식률 13% 감소, $t(9) = 4.79, p < 0.001$; 유성음 정인식률 9% 감소: $t(9) = 2.02, ns$).

3.2. 모음

청자들은 약 83%의 경우 (600회의 반응 중 496회에 해당) 한글로 어떤 모음을 들었는지를 기입하였다. 영어와 한국어의 모음 체계가 다르다는 점 및 동일한 음성 부호로 기입되는 모음일지라도 언어에 따라 상이한 음성적 특성을 지닌다는 사실 (Johnson, 1997)에 기반할 때, 원자극의 영어 모음과 한글로 기입된 모음 사이의 일대일 대응을 기반으로 말지각의 정오를 판별하기에는 무리가 있다고 간주하여, 변별 자질별(고저설성, 전후설성, 원순성, 긴장성) 정인식률을 산출하였다. 변별 자질별(고저설성, 전후설성, 원순성, 긴장성) 정인식률의 평균을 시뮬레이션 방식별로 요약한 결과는 아래 <표 6> 과 같다.

표 6. 모음의 변별 자질 별 정인식률
Table 6. Accuracy of vowel feature recognition

	원자극	보코딩을 거친 자극	보코딩과 3mm 스펙트럼 이동을 거친 자극
고저설성	71%	71%	61%
전후설성	76%	75%	49%
원순성	82%	78%	68%
긴장성	58%	50%	49%

<표 6> 에서 볼 수 있는 바와 같이, 모음의 변별 자질 인식에 있어서는 보코딩보다 스펙트럼 이동으로 인한 정인식률의 감소가 두드러졌다. 스펙트럼 이동을 수반하지 않은 보코딩 자극이 원자극 대비 평균 3.25%의 정인식률 감소를 보인 반면, 3mm의 스펙트럼 이동은 긴장성을 제외한 모든 변별 자질 지각에 있어 최소 10% 이상의 정인식률 감소를 야기하였다. 스펙트럼 이동 없이 보코딩만을 가한 자극과 3mm의 스펙트럼 이동 및 보코딩을 가한 자극간의 변별 자질 정인식률 차이는 고저설성(정인식률 10% 감소, $t(9) = 3.51, p < 0.01$), 전후설성(정인식률 26% 감소, $t(9) = 3.79, p < 0.01$), 원순성(정인식률 10% 감소, $t(9) = 3.79, p < 0.01$)에서 공히 통계적으로 유의미하였다. 반면, 스펙트럼 이동은 긴장성 지각(정인식률 1% 감소, $t(9) = 0.08, ns$)에 유의미한 정인식률 차이를 가져오지 않았다. 원자극과 스펙트럼 이동 없이 보코딩만을 가한 자극 사이에는 통계적으로 유의미한 정인식률의 차이가 없었다). 마지막으로, 원자극과 3mm 스펙트럼 이동 및 보코딩을

거친 자극 사이의 정인식률 차이는 전후설성(정인식률 27% 감소, $t(9) = 4.16, p < 0.01$) 및 원순성(정인식률 14% 감소, $t(9) = 6.41, p < 0.001$) 지각에서 통계적으로 유의미한 것으로 나타났다.

4. 논의 및 결론

본 연구는 위와 같이 인공와우 사용자가 당면할 영어 듣기 오류 양상을 추정하기 위하여, 건청인이 영어의 자모음을 들 때에 시물레이션을 거치지 않은 발화와 두 가지(스펙트럼 이동이 일어나지 않은 12채널 보코딩, 3mm 스펙트럼 이동이 일어난 12채널 보코딩) 인공와우 시물레이션을 거친 자극을 인식하는 양상에 차이가 있는지를 살펴보았다.

시물레이션의 영향은 대상 말소리가 자음인지 모음인지에 따라 상이하게 나타났다. 자음의 경우 스펙트럼 이동보다는 보코딩의 효과가, 모음의 경우 보코딩보다는 스펙트럼 이동의 효과가 오인식에 더 많은 영향을 미치는 것으로 나타났다. 보코딩에 의해 가장 높은 오인식 증가율을 보인 음소들은 조음 위치로는 순서대로 연구개, 양순, 치경음이었으며 조음 방법으로는 구강폐쇄음과 비음인 것으로 나타났다.

청자들은 원자극과 시물레이션 자극을 청취할 때 서로 다른 오인식 양상을 보였다. 예컨대, 원자극의 인식에 있어서 가장 지각 혼돈이 빈번했던 음소와 오인식된 음소들은 유성 치간 마찰음 /ð/를 /d/로 (총 17회), 파찰음 /dʒ/을 /z/로 (6회), /v/를 /b/로 (6회), /l/을 /r/로 (5회) 오인식한 경우들이었다. 시물레이션 자극 지각에서 역시 원자극 인식에서의 혼돈 양상이 비슷하게 나타난 경우도 있었으나 (/ð/를 /d/로 인식한 경우 총 23회, 파찰음 /dʒ/을 /z/로 인식한 경우 총 16회, /l/을 /r/로 인식한 경우 총 8회), 원자극 인식에서 발생하지 않은 말지각 오류 역시 나타났다. 대표적으로, /g/는 /d/로 (15회), /m/은 /b/로 (9회), /d/와 /p/는 /t/로 (각 8회) 인식된 경우들이 있었다. 특히 비 치경음이 치경음으로 인식된 것은 조음 위치 간 변별이 인공와우 사용자들에게 있어 말지각의 어려움을 초래한다는 기존 연구 결과와 일치한다(이미영, 2005; 안서지, 2005). 유성음을 무성음으로, 비 마찰음을 마찰음으로 인식한 보코딩 상의 오인식은 소음기반 보코딩 과정에서 주파대역의 에너지 집적부분에 백색소음이 삽입되었기 때문일 것으로 추정된다. 또한, 양순음에서 구강폐쇄음이나 접근음보다 비음의 오류 증가가 두드러진 것은 다음과 같이 해석할 수 있다. 구강폐쇄음의 경우 조음 기관의 접근, 폐쇄, 그리고 파열의 분명한 조음 단계가 포먼트 이동, 폐쇄 구간, 짧은 구간동안 나타나는 고주파대역의 에너지라는 음향적 단서로 나타난다. 이러한 시간적 정보가 보코딩에 의해 거의 영향을 받지 않는 반면, 접근음이나

비음 등 조음 단계를 걸쳐 공명이 일어나는 소리의 경우 보코딩에 의해 주파대별 에너지 분포에 대한 정보가 소실될 수 있고, 그 결과 보코딩으로 인한 음인식의 어려움이 한층 증가할 것으로 예상된다. 후속 연구에서는 이와 같은 오류들이 실제 인공와우 사용자들의 영어 청취 시 발견되는지 및 인공와우를 사용하는 영어 원어민의 말지각 오류와 어떠한 대응을 보이는지 등에 대해 알아보고자 한다.

스펙트럼 이동은 자음보다는 모음 인식에 더 많은 영향을 미쳤다. 이 결과는 선행 연구(Zhou et al., 2010) 결과와 일치하는데, Zhou 등(2010)은 스펙트럼 이동이 자음의 조음 방법 인식보다는 모음 등 스펙트럼 신호(spectral cue)에 의존하여 지각되는 말소리 인식에 더욱 많은 영향을 준다는 사실을 밝힌 바 있다. 스펙트럼 이동은 저주파대 에너지를 고주파대로 이동시킴으로서, 자음의 경우 동일 조음 방법과 유-무성성을 갖지만 절대적 에너지 집적 주파대역이 다른 자음과 혼동되어 들리게 하였다(예를 들어, 비음과 구강 폐쇄음의 혼돈 등). 모음의 경우 역시 스펙트럼 이동은 저주파대의 정보, 특히 F1 관련 정보를 왜곡시키게 됨에 따라, 모음의 고저성과 전후설성(F2-F1로 추정 가능) 및 영어의 경우, 전후설성과 밀접한 관련이 있는 원순-비원순성 인식에 어려움을 초래하였다. 또한, 모음 듣기에 관해서는 보코딩보다 스펙트럼 이동의 영향이 말지각에 미치는 영향이 큰 것으로 밝혀졌는데, 이러한 연구 결과는 선행연구(Fu & Shannon, 1999)와 일치한다.

시물레이션을 통해 인공와우 사용자들의 말지각 양상을 알아본 결과는 건청인을 대상으로 한 영어 지각 실험(양병곤, 2005; Kim & Kim, 2003 등)에서 보고한 말지각 오류 양상과 사뭇 다른 점이 많았다. 원자극에 대한 오류는 유음 간 구분 등 한국인 영어 학습자에서 전형적으로 나타나는 오류(양병곤, 2005; Kim & Kim, 2003)가 많은 반면, 시물레이션 자극에 대한 청취 오류는 한국인 학습자의 전형적 오류는 물론, 인공와우 사용자 특유의 조음 위치 및 방법 오류가 더해진 결과로 나타났다. 만일 인공와우 사용자가 시물레이션과 같은 결과를 보인다면, 이들의 영어 듣기 학습 지도와 평가에서 사용되어야 하는 영어 교수 전략은 건청인을 대상으로 한 영어 교수 전략과는 상이한 방식을 사용해야 할 것이다. 영어 모어 청자를 대상으로 한 선행 연구들과 본 연구 간의 차이점도 존재하는데, 예를 들어 Zhou 등(2010)의 연구와 비교하였을 때 본 연구에서 전체적인 정확도가 저하되는 것과 스펙트럼 이동을 거치지 않은 과제에서도 조음 방법 오류가 상당 비율 나타난다는 것이다. 이와 같이 선행 연구와 상이한 결과를 초래한 이유 중 하나는 청자가 영어 학습자였기 때문일 수 있다. 외국어 듣기가 모국어 듣기보다 인지적 노력이 더 많이 든다는 것은 널리 알려진 사실이며, 소음 상황이나 인공와우 시물레이션 등 음향적 명료도가 저하된 상황에서 외국어 학습자가 원어민보다 더 많은 어려움을 겪는다는 것 역시 널리 알려진

9) 고저설성: $t(9) = 0.12; ns$; 전후설성: $t(9) = 0.23; ns$
원순성: $t(9) = 1.59; ns$; 긴장성: $t(9) = 2.38; ns$

사실이다(Ganesh, Vijitha & Subba, 2012; Crandell & Smaldino, 1996; Lecumberri, Cooke & Cutler, 2010). 본 연구에서는 특히 원어인 말소리 듣기 경험이 거의 없는 대상자로 피험자군을 한정하였기 때문에, 제한된 정보를 지닌 외국어 듣기상의 어려움이 강조되었을 것으로 추측한다.

본 연구의 한계점으로는, 반응의 정확성에 대한 문제가 있었다. 대부분의 피험자가 모음 자극을 알파벳으로 적는데 어려움을 느끼고 한글로 반응하였다. 그 결과, 한국어 모음과 영어 모음이 질적으로 매우 상이함(양병곤, 2010)에도 불구하고 각 모음을 변별자질로 구분한 후 자질의 인식률을 평가할 수 밖에 없었다. 후속 연구에서는 이와 같은 방법론적 약점을 보강하기 위하여 객관식 평가나, 사전에 /hVd/ 어휘와 소리 연합을 교육한 후 실험을 수행하는 등의 조치를 취할 수 있을 것이다. 상대적으로 문자-소리간의 불일치가 적게 나타났으나, 자음 지각에서 역시 청자들이 반응한 알파벳 자음이 어떤 음소를 지칭하는 것인지에 대한 판별 문제가 존재하였다. 또한, 시료로 사용된 말소리가 원어인 남성 화자 1인의 발화를 대상으로만 하였다는 점을 감안할 때에, 오류 양상이 특정 화자에 국한된 것인지, 성별과 연령, 지역에 관계없이 모든 미국 영어 모어 화자에 일반화될 수 있는 양상인지에 대한 의문점 역시 뒤따른다. 추후에는 이와 같은 단점을 보완하여, 시뮬레이션을 통한 듣기 실험과 인공와우 사용자의 실제 영어 지각 실험 연구를 제언한다.

감사의 글

실험 소프트웨어를 제작해 주신 이충희님께 감사드립니다.

참고문헌

- Ahn, S. J. (2005). Relationships of speech perception and speech production according to intelligibility in the cochlear implanted. M. S. Thesis, Ewha Womans University.
(안서지. (2005). 인공와우 이식자의 말명료도 정도에 따른 지각과 산출 연구. 이화여자대학교 대학원 언어병리학 협동과정 석사학위논문.)
- Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. *Language experience in second language speech learning: In honor of James Emil Flege*, 13-34.
- Bohn, O. S., & Flege, J. E. (1990). Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics*, 11(3), 303-328.
- Crandell, C. & Smaldino, J. (1996). Speech perception in noise by children for whom English is a second language. *American Journal of Audiology*, 5, 47-51.
- Dorman, M. F., Loizou, P. C. & Rainey, D. (1997). Simulating the effect of cochlear implant electrode insertion depth on speech understating. *Journal of the Acoustical Society of America*, 102, 2993-2996.
- Dorman, M. F., Loizou, P. C., Fitzke, J., & Tu, Z. (1998). The recognition of sentences in noise by normal-hearing listeners using simulations of cochlear-implant signal processors with 6 - 20 channels. *Journal of the Acoustical Society of America*, 104(6), 3583-3585.
- Fu, Q.-J. & Shannon, R. V. (1999). Recognition of spectrally degraded and frequency-shifted vowels in acoustic and electric hearing. *Journal of the Acoustical Society of America*, 105, 1889-1900.
- Ganesh, A. C., Vijitha, S. & Subba, R. T. A. (2012). Objective measurement of listening effort while using first and second language in simulated cochlear implants. *Proceedings of International Symposium on Auditory and Audiological Research 2011*, Nyborg, Denmark.
- Johnson, K. (1997). Speech perception without speaker normalization: an exemplar model. In K. Johnson, & J. W. Mullenix (Eds.), *Talker variability in speech processing* (pp. 145 - 166). San Diego: Academic Press.
- Kim, G. & Kim, S. J. (2003). Korean Native Speakers' Perception of English Sounds According to the Groupings of Phonetic Contrasts. *Speech Sciences*, 10(1), 59-67.
- Kuhl, P. K. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics*, 50(2), 93-107.
- Lecumberri, M., Cooke, M., & Cutler, A. (2010). Non-native speech perception in adverse conditions: A review. *Speech Communication*, 52, 864-886.
- Lee, M. Y. (2005). Open-set Monosyllabic Speech Perception Test for Preschool Children. M. S. Thesis, Yonsei University.
(이미영. (2005). 학령전 아동용 보기 없는 조건의 단음절 말 지각 검사. 연세대학교 대학원 언어병리학 협동과정 석사학위논문.)
- Park, J., Lee, H. G. & Lim, J. (2008). A measurement of the lexical ability of hearing impaired university students. *English Language & Literature Teaching*, 14(4), 247-272.
(박정란, 이현구, 임정완. (2008). 청각장애 대학생의 영어어휘력 실태 조사. *영어어문교육*, 14(4), 247-272.)
- Rhee, S. C., Lee, Y. J., Lee, S. H. & Kang, S. K. (2005). Korean-Spoken English Corpus v.2005-01.

- (이석재, 이용주, 이숙향, 강석근. (2005). Korean-Spoken English Corpus v.2005-01.)
- Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J. & Ekelid, M. (1995). Speech Recognition with Primarily Temporal Cues. *Science*, 270, 303-304.
- Thomas, E., El-Kashlan, H., & Zwolan, T.A. (2008). Children with cochlear implants who live in monolingual and bilingual homes. *Otol Neurotol*, 29(2) 230-4.
- Waltzman, S. B., Robbins, A. M., Green, J. E., & Cohen, N. L. (2003). Second oral language capabilities in children with cochlear implants. *Otology & Neurotology*, 24(5), 757-763.
- Yang, B. (2010). College students' production and perception of English vowels. *English Language Teaching*, 22(4), 165-184.
(양병곤. (2010). 대학생들의 영어모음 발음과 지각. *영어교육연구*, 22(4), 165-184.)
- Yang, B. (2005). A Study of English Consonants Identified by College Students. *Speech Sciences*, 12(3), 139-151.
(양병곤. (2005). 대학생들의 영어자음 인지 연구. *음성과학*, 12(3), 139-151.)
- Zhou, N., Xu, L. & Lee, C.-Y. (2010). The effects of frequency-place shift on consonant confusion in cochlear implant simulations. *Journal of the Acoustical Society of America*, 128(1), 401-409.

• **임애리 (Yim, Ae-ri)** 제1저자
연세대학교 대학원 언어병리학협동과정
서울시 서대문구 연세로 50
Tel: 02-2228-3901
Email: erina21@nate.com
관심분야: 언어병리학, 음성학, 청각학

• **김다희 (Kim, Dahee)** 공동저자
연세대학교
서울시 서대문구 연세로 50
Tel: 02-2123-7517
Email: daheckim0726@gmail.com
관심분야: 음성학, 음운론, 심리언어학

• **이석재 (Rhee, Seok-Chae)** 교신저자
연세대학교 영어영문학
서울시 서대문구 연세로 50
Tel: 02-2123-4483
Email: scrhee@yonsei.ac.kr
관심분야: 음운론, 음성학, 언어이론, 영어교육