



A study on the predictability of acoustic power distribution of English speech for English academic achievement in a Science Academy*

Soon Park · Hyunkee Ahn**

Department of English Language Education, Seoul National University, Seoul, Korea

Abstract

The average acoustic distribution of American English speakers was statistically compared with the English-speaking patterns of gifted students in a Science Academy in Korea. By analyzing speech recordings, the duration time of which is much longer than in previous studies, this research identified the degree of acoustic proximity between the two parties and the predictability of English academic achievement of gifted high school students. Long-term spectral acoustic power distribution vectors were obtained for 2,048 center frequencies in the range of 20 Hz to 20,000 Hz by applying an long-term average speech spectrum (LTASS) MATLAB code. Three more variables were statistically compared to discover additional indices that can predict future English academic achievement: the receptive vocabulary size test, the cumulative vocabulary scores of English formative assessment, and the English Speaking Proficiency Test scores. Linear regression and correlational analyses between the four variables showed that the receptive vocabulary size test and the low-frequency vocabulary formative assessments which require both lexical and domain-specific science background knowledge are relatively more significant variables than a basic suprasegmental level English fluency in the predictability of gifted students' academic achievement.

Keywords: acoustics, long-term average speech spectrum (LTASS), English speech analysis, Science Academy, vocabulary teaching

1. 서론

텐마크를 영어 강국으로 만드는 데 결정적으로 기여한 학자로 인정받는 저명한 영어교육자이자 언어학자 Jespersen(1904)은 “외국어를 가르칠 때 가장 우선적으로 치중해야 하는 대상은

소리(sounds)”라고 지적하면서 외국어 학습자가 모국어와 목표어 사이에 존재하는 차이점에 주목하여 목표어에 근접한 음운 및 음향적 특성을 습득하도록 하는 교수학습법에 초기 교육의 역량을 집중해야 함을 주장하였다. 현존하는 수천 개의 언어 간에는 음소 단계인 분절음 수준(segmental level)에서의 차이뿐만

* This study was supported by Seoul National University research grant in 2020 (No. 700-20200053).

** ahnhk@snu.ac.kr, Corresponding author

Received 17 August 2022; Revised 14 September 2022; Accepted 14 September 2022

© Copyright 2022 Korean Society of Speech Sciences. This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

아니라 리듬, 강세, 지속시간(duration), 주파수 대역(frequential passbands), 진폭(amplitude) 수준 등 초분절적(suprasegmental) 현상에서도 서로 다른 특성을 갖는다(Ladefoged & Johnson, 2014). 실제, 성조언어를 쓰는 중국어 화자는 미국영어 화자에 비해 평균적으로 더 다채롭게 변화하는 기본주파수(fundamental frequency: f_0)값을 활용하여 발성한다는 사실이 실험적으로 검증되었다(Eady, 1982; Keating & Kuo, 2012).

언어의 음향적 특성을 연구한 문헌들은 주로 f_0 및 근접 포먼트 몇 개를 측정하여 정량화하기에 용이한 모음을 대상으로 하는 경우가 대다수인데 이는 언어적 특성을 명백히 규명하기에 충분하지 않다(Behrman, 2022; Johnson, 2011; Mennen et al., 2012; Plack, 2018; Schnupp et al., 2012). 자음에 비해 모음은 대개 더 긴 지속시간을 갖고 있고 음향적 강도(intensity) 측면에서 더 강한 에너지를 가지고 있는 반면, 자음의 지속시간은 10-40 밀리초(ms) 정도에 불과하여 주파수 대역별 스펙트럼 분석에 있어 용이하지 않아, 어음명료도(intelligibility)를 결정하는 데 있어 약 90% 정도의 음향정보를 담고 있는 자음(Amlani et al., 2002; French & Steinberg, 1947)을 기반으로 하여 언어 간 음향 차이를 규명하고자 하는 시도는 대개 음소나 음절 수준의 접근법 위주로 진행되었다(Li et al., 2007). 하지만, Baars & Gage(2010)가 언급한 ‘20-200-2,000’ 현상처럼 음소(phoneme, 20 ms), 음절(syllable, 200 ms) 수준뿐만 아니라 억양(intonation) 등 긴 분석창에서의 운율적 변화(prosody, 2,000 ms)를 모두 포함하고 있는 장시간 분량의 실제적 언어 샘플을 음향적으로 분석하는 방식도 뜻깊은 접근법이라고 판단된다. 관련 분석 툴 및 분석 기기의 비약적 성능 향상을 발판 삼아, 수집 분에서 수집 시간 이상 지속되는 대규모 음성 데이터에 담긴 자음과 모음 정보를 아울러 종합적으로 분석하여 사람이 일상 속에서 실제적으로 체감하는 언어 풍경 속에 담긴 음향적 특성을 포착해낼 수 있는 연구가 가능해졌다(McLoughlin, 2016; Vasilescu et al., 2015).

미국 뉴욕의 벨연구소에서는 20세기 중반부터 유선 전화기 통화 품질 향상을 목적으로 한 바탕정보 확보를 위해 장시간에 걸쳐 지속적으로 발화되는 담화의 음향적 특성을 분석할 수 있는 LTASS(long-term average speech spectrum) 기법을 개발하였다(Amlani et al., 2002; Dunn & White, 1940). 본 기법을 활용하면 순간적인 음향 스펙트럼 특성뿐만 아니라 실제 상황과 동일한 장시간의 관찰 정보에 대한 평균적 음향 스펙트럼을 적분 및 평균하여 음향 스펙트럼 대역별로 시각화할 수 있다.

유선통화 회선의 음질 분석 및 개선을 위해 개발된 LTASS는 이후 청능학(audiology) 분야에서 주로 활용되었는데, 일례로 Cox & Moore(1988)는 2분 분량의 음성 녹음 데이터 60개를 1/3 옥타브 밴드에서 음향 스펙트럼 분석하여 미국 영어와 오스트레일리아 영어 사이의 음향적 차이를 규명하였고, 확보된 수치 정보를 보다 세밀한 보청기 이득값(gain) 설정을 위한 기초자료로 활용하였다.

LTASS를 각 언어별 음향 특성분석을 위해 도입한 국제적 연구도 진행된 바 있으나, 각 언어 화자의 분석 대상 음성 표본의 지속시간은 불과 몇 초에서 몇 십 초 정도였다(Byrne et al.,

1994). Olsen et al.(1987)은 사람들의 일상 속 언어현상에까지 일반화하기 위해서 자음 및 모음을 별도로 분리하거나 단일음절만을 대상으로 음향분석하는 수준을 넘어야만 실용적인 측정 방법으로 인정할 수 있다고 주장했는데, 이는 선행연구들이 진행될 당시의 기술적 수준의 한계 때문이었던 것으로 추정된다.

Noh & Lee(2012)의 연구처럼 LTASS 기법을 활용하여 한국어 화자의 경우 영어 화자에 비해 2,000 Hz 이상 주파수 대역에서 음향 에너지 강도가 유의미한 수준으로 낮은 현상을 규명한 사례는 있어도, 이중언어 화자를 주요 대상으로 진행된 연구는 찾아보기 어려웠음이 사실이다. 그런데, Cho & Munro(2017)는 한국어-영어로 모두 유창한 이중언어 구사자와 한 개 언어만 구사할 수 있는 실험참여자 사이의 LTASS 수치를 Praat을 사용하여 비교한 결과, 양자 사이에 주파수 대역별 음향 스펙트럼 분포에 있어 상당한 차이가 관찰됨을 보고하면서 LTASS 분석기법이 모국어 이외 특정 목표어 유창성 정도를 판별할 수 있는 정량적 기준을 제공할 개연성이 있다고 주장하였다.

이러한 배경하에 본 연구는, 행렬기반 수학/공학 계산 프로그램인 MATLAB[version: 9.12.0.1956245(R2022a); The Math Works, 2022]을 이용하여 담화 수준 이상 장시간 분량의 목표언어 대용량 녹음 데이터를 분석하는 것을 일차적 목표로 실생활의 언어 사용 상황에 근접한 목표어의 음향 스펙트럼 대역별 결과물을 획득하고, 이 작업을 통해 산출한 영어 음향 스펙트럼 곡선과 한국어를 모국어로 하는 영어 학습자의 곡선을 비교하여 영어 원어민과 한국인 영어 화자 간 음향 스펙트럼 분포의 근접도 판별을 시도하고자 한다.

2. 본론

이 연구는 우리나라 전체 고교 재학생 수의 0.2% 정도에 지나지 않는 국내 과학영재학교 재학생 중 연구자가 재직하고 있는 영재학교의 1학년 1, 2학기 영어I과 영어II 수강생 80명을 대상으로 하며, 학생들의 영어발화 녹음을 LTASS 분석하여 얻은 음향 스펙트럼 분포 곡선이 미국영어 화자의 음향 곡선과 얼마나 근접하고 있는지 정도와 영어 학기말고사 학업성취도를 비교하여 두 변인 사이에 얼마나 유의미한 통계적 관계가 존재하는지 살펴볼 것이다. 다각적 고찰을 위해 동일 학생들을 대상으로 하여 실시한 수용성 영어 어휘크기검사(receptive vocabulary size test, rVST), 영어 어휘/문법/독해 형성평가 퀴즈(vocabulary quiz, VQ) 정답 수, 영어 유창성 검사인 공인 영어말하기시험(English Speaking Proficiency Test, ESPT) 성적을 변인으로 추가 분석하였고, 그 결과를 바탕으로 과학영재학교 재학생들에게 중점을 두어야 할 영어 교수학습 항목으로 무엇이 최적의 대상일지에 관해서도 탐색해 볼 것이다.

본 연구가 답변을 모색할 연구 질문(research question, RQ)은 다음과 같다. 과학영재학교 1학년 재학생 발화의 주파수 대역별 음향 스펙트럼 분포는 미국 영어 원어민 화자의 주파수 대역별 음향 스펙트럼 분포와 얼마나 유사한가(RQ1)? 이러한 유사성 정도와 영재학교 영어과목 정기고사 성취도 등 변량과의 통

계적 관계는 얼마나 강한가(RQ2)? 어휘 퀴즈와 영재학교 영어 과목 정기고사 성취도 사이의 관계 강도와 비교해 볼 때 과학영 재학교 신입생들에게 중점을 두어 교수해야 할 최적의 언어 항목은 발음이나 억양 등 초본절적 요소와 어휘 중 무엇인가(RQ3)?

2.1. 연구 데이터 수집

2.1.1. 미국영어 원어민 음성녹음 데이터

미국 원어민의 주파수 대역별 음향 스펙트럼 그래프를 얻기 위해 미국에서 출생하여 성장한 영어 화자 30명(남성 15명, 여성 15명)을 모집하여 녹음 작업을 진행했다. 직접 대면하여 녹음하기에는 비용이나 기술적 측면에서 어려움이 있어, 온라인 프리랜서 구인구직 사이트인 Fiverr의 미국영어 전문 성우녹음(professional voice over) 서비스 그룹 페이지에 참여자 모집 공고를 올리거나 동 사이트 내에 게시된 전문 성우의 프로필을 검색하고 녹음 샘플을 청취한 후 적합하다고 추측되는 미국인 화자에게 직접 메시지를 보내 미국 본토 출생 여부를 확인하고 녹음작업 진행을 요청하였다. 본 실험에 녹음파일을 공급한 미국인 참여자의 출생 및 거주지역은 New Jersey, Virginia, Washington, Pennsylvania, Tennessee, Kansas, Illinois, Michigan, Oregon, Iowa, Utah, California, Texas 등 미국 동서부 전역을 망라했고, 연령대는 20대에서 60대까지 다양했다. 녹음을 위해 사용한 텍스트는 전 세계적으로 널리 번역되어 있고 대부분의 참여자가 유소년 시절 접해본 적이 있어 내용적으로 친숙하며 대본의 흐름상 긍정적 정서 반응을 유도할 확률이 높은 Saint-Exupéry의 *The Little Prince* 20장과 21장을 30명 전부 동일한 방식으로 소리 내어 읽어 녹음하도록 안내하였다. 녹음이 완료된 후, 각 참여자의 전문 경력이나 레코딩 완성도에 따라 녹음파일 1개당 최소 8달러에서 최대 30달러 정도의 참여비를 PayPal을 통해 온라인으로 지급했다.

영화 더빙, 라디오 및 TV 광고 녹음, 오디오북 제작 등을 주업이나 부업으로 하는 전문 성우들이어서 참여자 모두 전문가용 마이크 및 녹음기, 방음 녹음부스(anechoic recording booth) 등 전문가급 녹음 장비를 갖추고 있었으므로 양질의 녹음 데이터를 확보할 가능성이 높았다. 녹음 참여자는 Tascam DR-100MK3, 또는 이에 준하는 전문가용 녹음장비나 PC에 연결되는 오디오 인터페이스 등을 사용하여 MONO, 24 bit(또는 32 bit), WAV 형식으로 음성녹음 파일을 제작했고 샘플링레이트는 44,100 Hz였다. 단, 스테레오로 녹음된 2명의 참여자 파일은 연구자가 오디오 편집 프로그램을 이용하여 MONO로 변환한 후 음향분석을 실시했다.

수집된 30개의 WAV 파일은 세어웨어인 Audacity로 다음과 같이 편집 및 표준화 처리하였다. 첫째, 화자마다 음의 강도(intensity)가 다르므로 분석 대상 음성데이터 전체 중 나타난 최대진폭에서 -1.0 dB 음량으로 각 음성데이터를 노멀라이즈하였다(normalization by removing DC offset, peak amplitude to -1.0 dB). 둘째, 형광등이나 냉난방장치로 인한 저음 노이즈를 삭제

하기 위해 70 Hz를 절삭점으로 하여 옥타브당 12 dB 기울기로 감쇄시키는 설정으로 고역통과필터링하였다(high-pass filtering at the 70 Hz cut-off with the slope of 12 dB/Oct). 셋째, 0.5초 이상 지속되는 침묵 구간은 전부 삭제하여 분석과 관계없는 노이즈의 왜곡 개입 가능성을 회피할 수 있도록 음성발화 사이의 침묵을 삭제하였다(silence truncation at the threshold of -20 dB and 0.5 sec.).

2.1.2. 과학영재학교 재학생 음성녹음 데이터

연구자 중 한 명이 재직 중인 과학영재학교 신입생 87명 중 2021학년도 1학기에 필수과목 중 하나인 영어I을 수강했고, 2021학년도 2학기에 역시 필수과목인 영어II를 수강한 1학년 재학생의 음성녹음 데이터를 수집했는데, 1학년 재학생 중 6명은 본교 학업성취기준에 의거하여 선이수제(placement test)로 필수과목 수강면제를 받았고 1명은 2학기 중반에 휴학하였으므로 총 실험참여 인원은 80명이다. 녹음작업은 1학과 2학기, 2회에 걸쳐 진행되었으며, 1차 녹음의 경우 미국영어 원어민과 마찬가지로 Saint-Exupéry의 *The Little Prince* 20장과 21장을 영어원문 그대로 녹음했고 2차 녹음은 Daniel Keyes의 소설 *Flowers for Algernon* 중 일부를 지정하여 참여자 모두 약 5분에서 10분 분량의 동일한 텍스트를 소리 내어 녹음하도록 안내했다. 이 중 1차 녹음 데이터만 음향분석에 활용하였다.

연구자가 전용하는 강의실에 설치된 전문가용 레코딩 장비를 활용하여 녹음할 것을 학생들에게 권장하였다. 데이터 수집을 위해 투입된 전문 녹음장비 구성은 다음과 같다. 전문가용 Pro 레코더 Tascam DR-100MK3, 지향성 콘덴서마이크 Tascam TM-280, 진동저감형 마이크 거치대 및 스탠드, 전용 팝필터 TM-AG1, Tascam Acoustic Control Filter TM-AR1(벽 등 주변 물체에 반사하여 들어오는 반향음 감쇄 기능). 레코딩 형식은 원어민과 동일하게 MONO, 24 bit, WAV였고 샘플링레이트는 44,100 Hz였다. 원어민의 경우와 동일한 방식으로 MATLAB 분석 전에 전처리 과정을 거쳤다. 전처리된 모든 데이터 파일 중 무작위 추출한 일부를 MATLAB(R2022a; The Math Works, 2022)에 ioSR 툴박스를 인스톨 후 LTAS 코드(LTASS값을 계산하는 코드의 명칭이 LTAS로 명명됨)로 음향분석해서 0-20,000 Hz 사이 2,048개 주파수 지점마다 음향 파워스펙트럼 벡터값을 평균 추출한 후 각각의 주파수 대역별 음향 스펙트럼 분포곡선을 그렸다.

2.1.3. 영어 학기말고사, 수용성 어휘크기검사, 어휘퀴즈, ESPT

음향분석 대상 녹음 자료에 더해, 학업성취도 예측률을 알아보기 위한 추가 지표로 수업 중 확보된 데이터를 정하였다. 2021학년도 2학기 영어II 과목의 정기고사는 서답형, 서술형, 논술형 지필고사를 2회(중간, 기말)에 걸쳐 시행하였고, 수행평가 하위영역은 총 3개로서 Task Portfolio(과제 제출 및 퀴즈 등), Speech Presentation(구술형), Quiz(서답형, 서술형)로 구성되었다. 각각의 반영비율은 지필고사 55%, 수행평가 45%였다. 학기말고사 성적 처리를 위해 모든 영역의 점수를 100점 만점으로

합산하여 A+부터 F까지 평어를 부여하였다.

수용성 어휘크기검사(rVST)는 2021학년도 1학과 2학과 첫 수업시간에 각각 실시된 14,000단어 및 20,000단어 버전의 표준화된 어휘력 평가로서 두 평가의 성취도를 합산한 값을 활용하였으며, 그 당시 해당 학생이 보유하고 있는 어휘력의 크기를 객관적으로 측정할 수 있도록 투입되었다(Coxhead et al., 2015; Nation & Beglar, 2007).

어휘퀴즈(VQ)는 학기 중 총 3회에 걸쳐 실시되었는데, 퀴즈 직전 주까지의 수업 중 학습했던 지문이나 교수학습활동 전반에 걸쳐 등장하는 어휘, 어법, 의미적 빈칸추론, 오류수정 등의 유형으로 구성되었고 각 회당 평균 50문항씩이어서 만점은 총 150점이다.

국내에서 개발된 공인 영어말하기시험인 ESPT(English Speaking Proficiency Test)는 실험참여 과학영재학교의 졸업인증 자격을 취득하기 위해서 1학년 재학생 전원(2, 3학년 탈락자 포함)이 참여하는 평가이다. 2021학년도 참여인원은 총 91명으로서 1학년 전원(87명), 2학년 3명, 3학년 1명이 응시하였으며, 최고점은 674점(2등급), 최저점은 282점(4등급)이었고, 학기 중 영어I과 II를 수강한 80명의 데이터만 분석 대상으로 도입했다. 본 시험의 졸업인증 최소기준(501점) 통과율은 98.9%였다. ESPT는 온라인으로 실시되며 발화자가 영어로 특정 정보를 설명하거나 설정된 상황에 맞게 대응하는 영어 음성언어 활용능력을 측정하는 시험이므로 영어 녹음 데이터의 음향 스펙트럼 분석 결과를 해석하기 위해 보조적으로 투입할 목적으로 변량에 포함시켰다.

아래 표 1에 제시된 세 가지 영어 성취도 지표 중 어휘퀴즈($SD=26.86$)와 ESPT의 표준편차($SD=45.85$)가 지필고사 표준편차($SD=10.59$)나 수용성 어휘크기검사의 표준편차($SD=14.10$)에 비해 월등히 크지만, 수강생별 학기말고사 원점수 및 rVST 점수 차이가 조밀하여도 동점이 거의 없이 차별화되어 있어 변별력 차원에서 네 지표 모두 효과적인 비교 대상으로 활용 가능하였다.

표 1. 정기고사(WT), 수용성 어휘크기검사(rVST), 어휘퀴즈(VQ), ESPT 기초 통계량

Table 1. The descriptive data for WT, rVST, VQ and ESPT

| 변량명 | 자료 수 | 평균 | 표준편차 | 표준오차 |
|---------|------|---------|--------|-------|
| WT | 80 | 84.058 | 10.590 | 1.184 |
| rVST | 80 | 87.106 | 14.097 | 1.576 |
| VQ | 80 | 76.675 | 26.861 | 3.003 |
| ESPT | 80 | 605.625 | 45.853 | 5.126 |
| 결측 수: 0 | | | | |

WT, written test; rVST, receptive vocabulary size test; VQ, vocabulary quiz; ESPT, English Speaking Proficiency Test.

3. 데이터 분석 및 결과

3.1. 음향데이터

미국영어 원어민 남성 참여자 15명과 여성 참여자 15명의 음성녹음 데이터를 각각 MATLAB LTAS 코드로 음향분석한 후 그림 1로 시각화하였는데, 코드에서 음향신호에 대해 설정된

FFT 길이는 4,096이고 hop 사이즈는 2,048이며 각 세그먼트에 Hann 윈도우를 적용했다. 추출된 파워스펙트럼밀도(power spectrum density, PSD)값을 1/6 옥타브 해상도로 가우시안 스무딩(Gaussian-smoothing) 처리하여 음향 그래프 곡선을 부드럽게 조정했다. X축은 대수계산자(logarithmic scale) 방식으로 표시했고 125 Hz부터 8,000 Hz까지 옥타브 단위로 주파수 대역 수치를 명시했다. 미국영어 원어민 데이터와 본교 재학생의 녹음 데이터 모두 70 Hz를 절삭점으로 고역필터링(high-pass filtering)하여 국내 AC 전원의 60 Hz 특성에서 비롯될 가능성이 있는 잡음 및 냉난방기 등 주변에서 유입되는 저음 영역대 소음을 감쇄시키되, 남성의 f_0 값에는 영향을 미치지 않도록 전처리했다. 남성과 여성의 f_0 값의 대략적 중심점이 되도록 그래프의 시작점을 125 Hz로 하였다. 예상할 수 있듯, 모음이 주로 자리잡는 f_0 부터 첫째에서 둘째 포먼트(formant)가 주로 자리잡는 80–1,000 Hz 사이 저음 영역의 음향 파워스펙트럼밀도(PSD)가 높게 나타났고 고음 영역으로 갈수록 평균 누적 에너지 수치가 내려갔으나 본 연구 수준에서는 별도의 청각학적 보정 작업을 실시하지는 않도록 설계했다.

그림 1은 미국식 영어를 모국어로 하는 남성 원어민 화자 15명과 여성 원어민 화자 15명의 음성녹음 데이터를 모두 합하여 확보한 총 4시간 분량에 달하는 녹음데이터를 장기 스펙트럼 평균 계산한 값을 바탕으로 얻은 LTASS 분포 곡선이다.

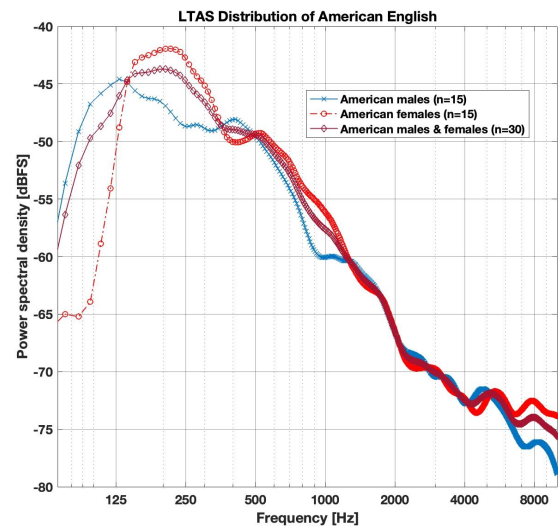


그림 1. 미국영어의 LTASS(long-term average speech spectrum) 분포(n=30)
Figure 1. Long-term average speech spectrum (LTASS) distribution of American English (n=30)

그림 2의 상단은 해당 과학영재학교 1학년 재학생이 본 공인 영어말하기시험(ESPT) 상위, 중위, 하위 30% 각 24명 중 5개씩 총 15건의 무작위추출 샘플 데이터를 분석한 결과이며, 하단 곡선은 영어II 과목 학기말고사 상위, 중위, 하위 30% 각 24명 중 역시 5개씩 총 15개의 샘플을 무작위추출하여 LTASS 음향스펙트럼 분석한 결과이다.

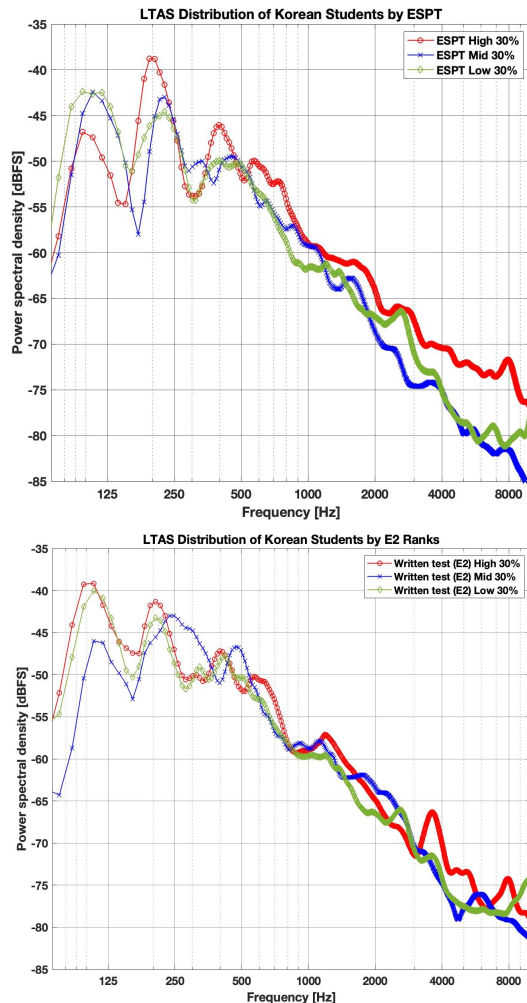


그림 2. 과학영재학교 재학생의 LTASS(long-term average speech spectrum) 분포[위: ESPT(English Speaking Proficiency Test) 상, 중, 하에 따라(n=15), 아래: 지필고사(written test) 상, 중, 하에 따라(n=15)]
Figure 2. Long-term average speech spectrum (LTASS) distribution of Korean students in a Science Academy (Above: according to high, mid and low levels of English Speaking Proficiency Test (ESPT) (n=15); Below: according to written test high, mid and low achievements (n=15))

표 2. 정기고사(WT), 수용성 어휘크기검사(rVST), 어휘퀴즈(VQ), ESPT 상관관계 분석

Table 2. The correlation between WT, rVST, VQ and ESPT

| 변량명 | WT | rVST | VQ | ESPT |
|------|----------------------------------|---------------------------------|----------------------------------|---------------------------------|
| WT | 1 | 0.665 $t=7.872$ $p<.0001$ | 0.758 $t=10.279$ $p<.0001$ | 0.493 $t=5.005$ $p<.0001$ |
| rVST | 0.665 $t=7.872$ $p<.0001$ | 1 | 0.553 $t=5.861$ $p<.0001$ | 0.373 $t=3.552$ $p=.0007$ |
| VQ | 0.758 $t=10.279$ $p<.0001$ | 0.553 $t=5.861$ $p<.0001$ | 1 | 0.362 $t=3.430$ $p=.0010$ |
| ESPT | 0.493 $t=5.005$ $p<.0001$ | 0.373 $t=3.552$ $p=.0007$ | 0.362 $t=3.430$ $p=.0010$ | 1 |

WT, written test; rVST, receptive vocabulary size test; VQ, vocabulary quiz; ESPT, English Speaking Proficiency Test.

각 그래프를 그리기 위해 사용된 데이터의 정규분포 정도를 파악하기 위해 샤피로-윌크 모수 검정(Shapiro-Wilk test)을 실시한 결과, 그림 2 데이터 전부는 비정규 분포로 확인되었으므로, 영어 원어민 PSD와 과학영재학교 재학생 PSD 사이의 통계적 일치성 여부를 알아내기 위해 비모수검정(non-parametric test)인 만-휘트니 U검정(Mann-Whitney U test)을 시행하였다. 첫째, 그림 1에서 남성과 여성을 합하여 산출한 주파수 대역별 미국영어 스펙트럼 곡선과 그림 2 상단의 ESPT 곡선에서 상, 중, 하 수준별 한국 영재학교 학생의 음향 곡선을 대상으로 실시한 검정 결과, 상 수준의 경우($W=2e+06$, $p=.003$), 중 및 하 수준 곡선($W=3e+06$, $p<.000$)에 비해 영어 원어민의 음향 특성 대비 다소 적은 차이만 관측되었지만 상, 중, 하 집단 모두 미국영어와는 통계적으로 유의미한 차이가 존재했다. 둘째, 위와 동일한 미국영어 스펙트럼 곡선과 그림 2 하단의 영어 학기말고사 성적 수준별 곡선 각각을 비교한 결과, 영어 학기말고사 성취도 상위 30%에 속한 학생들의 영어발화 음향곡선 특성이 일반적 유의 수준에서 원어민 곡선에 근접한 수준임을 발견했다($W=2e+06$, $p=.07$). 반면, 영어 학업성취도가 중, 하 수준인 참여자들의 경우에는 영어 원어민 발화와의 음향적 차이가 통계적으로 유의미하게 존재하는 것으로 나타났다($W=3e+06$, $p<.000$). 즉, 본교 재학생들의 수준별 곡선에서 미국영어 원어민의 주파수 대역 분포에 비해 유의미할 정도로 뚜렷한 차이는 여섯 집단 중 다섯에서 관측되지 않았다(RQ1). 이러한 결과로부터 과학영재학교의 수준 높은 영어 자료에 기초하여 서논술형 주관식으로 출제되는 지필고사 중심의 영어 학기말고사 성취도 수준을 예측하는데 있어 주파수 대역별 LTAS 곡선이 최적의 후보 척도로 활용되기에는 적합하지 않다고 해석되었다(RQ2).

3.2. 기타 변량 사이의 통계적 관계 분석

과학영재학교 재학생의 영어 성취도에 대해 영향력 있는 예측 지표로 기능할 수 있는 지표를 추가로 규명하기 위해, 미국영어 원어민과 과학영재학교 1학년 재학생의 수준별 음향 스펙트럼 분포 비교에 이어 2021학년도 2학기 영어II 학기말고사(written test, WT), 수용성 어휘크기검사 점수(rVST, 1학기 초와 2학기 초 2회 실시한 성적 합산), 2학기 중 3회 실시한 어휘퀴즈(VQ, 어휘에 더해 문법과 문맥상 빈칸추론, 오류 수정 문항 등이 종합적으로 들어가 있는 형성평가) 합산 점수, 공인 영어 말하기시험(ESPT) 성적의 네 변량 간 상관관계 및 선형 회귀분석을 실시하였다.

상관계수행렬표(표 2)에 분명히 드러나 있듯이, 음향음성학적 발음의 원어민 근접도가 채점기준에 포함되는 공인 영어 말하기시험(ESPT)과 다른 세 변량 사이의 상관관계는 낮은 편이지만, 수용성 어휘크기를 객관적으로 측정한 점수(rVST)와 학기 중 학습한 내용에 기초하여 시행된 어휘퀴즈(VQ)가 영어 학기말고사 성취도를 예측하는 데 있어서 보다 강력한 예측 척도가 될 수 있다고 판단된다(RQ3).

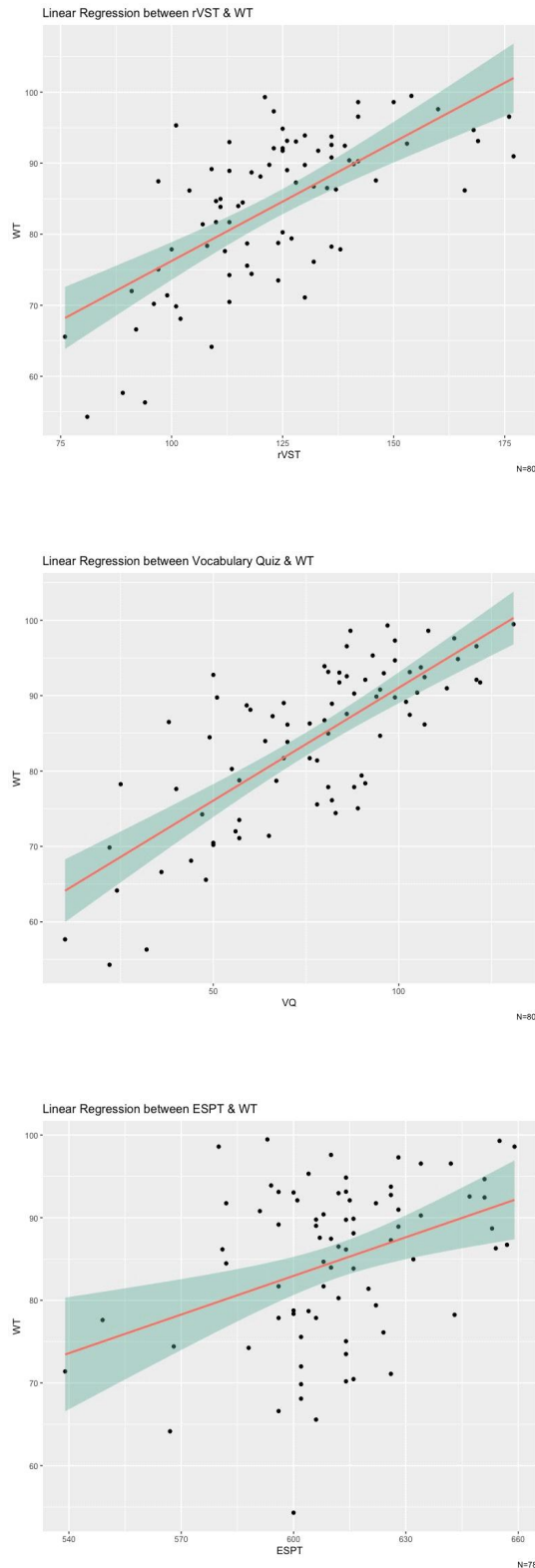


그림 3. 4개 변량[rVST(receptive vocabulary size test), VQ(vocabulary quiz), ESPT(English Speaking Proficiency Test) & WT(written test)] 간 선형회귀분석 산점도

Figure 3. Linear regressions between four variables [receptive vocabulary size test (rVST), vocabulary quiz (VQ), English Speaking Proficiency Test (ESPT) & written test (WT)]

이러한 현상은 그림 3의 선형 회귀분석 그래프에서도 뚜렷이 관측된다. 1학년 1학기 영어I 및 영어II 첫 수업시간에 측정한 수용성 어휘크기검사(rVST)와 2학기 영어II 학기말고사 성취도 (WT) 사이에는 유의미한 관계가 나타났다, $F=62.3$, $df=78$, $p<.01$. 결정계수(efficient of determination) $r^2=0.44$. 이는 강한 수준의 효과가 있다고 간주되는 0.25를 매우 상회하는 수치이다. 2학기 중 수업시간 중 학습한 내용에 근거하여 시행된 형성평가 성격의 어휘퀴즈(VQ)와 영어 학기말고사 성취도 사이에는 더 강한 수준의 유의미한 회귀 관계도가 관찰되었다, $F=106$, $df=78$, $p<.01$. 결정계수 $r^2=0.58$. 반면, 음조나 인토네이션 등 초분절적 음향 특성이 중요한 요인으로 고려되는 공인 영어말하기시험(ESPT)과 영어II 학기말고사 사이에는 매우 미약한 관계성만 존재하는 것으로 나타났다, $F=11.8$, $df=76$ (과도한 왜곡을 초래하는 아웃라이어 값 두 개를 제외하고 계산함), $p<.01$. 결정계수 $r^2=0.13$. 이는 중간 수준의 효과가 있다고 간주되는 수치에 근접한 수준이어서 주파수 대역별 LTASS 분포 곡선의 경우와 마찬가지로 ESPT는 과학영재학교 재학생의 영어 정기고사 성취도를 효과적으로 예측하는 데 기여할 가능성이 현저히 낮은 지표로 밝혀졌다.

4. 결론

지금까지 논의한 데이터 분석 결과를 토대로 추측해 보건대, 과학영재학교 재학생들은 학내 영어 학업성취도 상위권 학생들을 제외하면 외국어로서의 영어발화에 있어 미국 원어민의 주파수 대역별 LTASS 분포 곡선과 통계적인 차이를 보이며, 공인 영어말하기시험(ESPT)이 과학영재학교 영어 성취도를 효과적으로 예측할 수 있는 지표로 채택하기에는 그 강도가 부족함을 알 수 있다. 정기고사 성취도 상위권 학생은 대부분 영어 유창성도 우수하지만, 영어 유창성이 우수하다고 해서 정기고사 성취도가 상위권에 속하지는 않는 것으로 나타났다는 점에도 주목할 필요가 있다. 오히려 이번 연구에 참여했던 학생 중 최상위권의 정기고사 성취도를 보인 학생이 하위권의 영어 발화 유창성을 가졌지만 논리적인 영어 글쓰기 능력, 미국 학부생용 과학 교과서나 뉴욕타임즈 등에서 발췌한 영어 기사 원문으로 수업한 내용에 대해 작성하여 제출한 과업 완성도에서 월등한 실력을 발휘하고 있었다.

이러한 자료로부터 다음과 같이 추리해볼 수 있을 것이다. 첫째, 학년 말 실시되는 공인 영어말하기시험(ESPT)의 합격률이 98%를 상회하는 것으로 보아 과학영재학교 재학생 대부분은 일정 수준의 영어 발화능력과 유창성은 확보하고 있다고 생각된다. 특히 발음 등 초분절적 자질의 구사와 관련된 조음근육(articulatory muscle)의 섬세한 움직임을 통제하는 능력은 민감한 시기(sensitive period)라고도 불리는 시절 대부분 완성되는 것으로 알려져 있는데(Kuhl, 2004; Moon et al., 2013), 본교 재학생 및 학부모와의 교류를 통해 일관되게 얻은 정보로부터 짐작할 수 있듯, 중등학교에 진급하기 이전 유치동 시절에 영어 청취 및 발화 연습에 집중하는 경향성에서 이러한 높은 합격률이

비롯된 것으로 추정해볼 수 있다. 그러나 LTASS 음향 스펙트럼 곡선 간 비교 결과 미국영어 원어민의 음향 특성에 비해 재학생 상당수의 주파수 대역별 음향 특성에 있어 유의미한 분포 차이가 여전히 관측되고 있는 바, 이러한 간극이 영어 유창성 차원에서 어떠한 학술적 의미를 함의하고 있는지에 관해서는 후속 연구를 통해 보다 정교하게 규명해볼 필요가 있다. 둘째, LTASS 음향 스펙트럼 곡선으로 대변되는 운율적 특성(prosody) 지표는 깊이 있는 수과학적 배경지식을 요구하며 과학 분야에 특화된 고난이도 저빈도어휘(low frequency words) 위주로 진행되는 과학영재학교 영어교과의 성취도를 예측하는 지표로 기능하기 어렵다. 실제, 본교 영어교과 성취도 예측에 있어 가장 유력한 후보로 부각된 변량은 측정 당시의 어휘크기를 보여주는 수용성 어휘크기검사(rVST)와 학기 중 학습한 고급 수준의 과학 어휘 및 각 어휘의 라틴어, 그리스어, 프랑스어 등의 어원적 설명(etymology), 문맥에 의거하여 목표 어휘를 추정하도록 고안한 빈칸채우기, 형식적 및 어휘적 오류를 고치는 문항 등으로 구성된 어휘퀴즈(VQ)였으므로, 과학영재학교 영어 수업은 전문적 어휘 교수 위주로 진행할 때 더욱 효과를 발휘할 것이라 추측된다. 해당 과학영재학교의 필수 졸업요건으로 501점 이상을 요구하는 공인 영어말하기시험(ESPT)은 초등 고학년 수준의 실력을 가진 학생이라면 대부분 통과 기준을 충족 가능하므로 과학영재학교 재학생의 영어수준을 적합히 측정하고 재학 중 지속적으로 영어실력 향상을 촉진 및 도모하고자 한다면 ESPT 대신 다른 공인 영어시험에 집중할 수 있도록 제도적으로 보완할 필요가 있다.

1년에 걸쳐 영어 음성 데이터를 수집하여 과학영재학교 재학생의 영어 성취도 예측 지표를 다각도로 분석하고 증거에 기반한 최적의 영어수업 방향을 모색해보고자 설계한 본 연구는 기존의 연구에 비해 양적 측면에서 규모가 큰 음성 데이터를 음향적으로 분석하여 비교하고자 시도하였다는 면에서는 긍정적이거나 질적 연구 수준으로 구체적인 개별화 탐구를 진행하지 못하였다는 데 한계가 있다. 또한, LTASS 분포 곡선을 모두 수치화하여 통계적 유의미성을 주요 대역별로 대조 및 비교 분석하는 기술적 방법을 정교화할 필요가 있으므로 향후 후속연구의 음성음향 스펙트럼 분석에서 ANSI S3.5 표준 등에 근거하여 1/3-1/12 옥타브별 중심주파수 대역별로 세분화된 언어 내 및 언어 간 음향 분석 연구를 실시할 필요가 있다(Bentler et al., 2016).

References

- Amlani, A. M., Punch, J. L., & Ching, T. Y. C. (2002). Methods and applications of the audibility index in hearing aid selection and fitting. *Trends in Amplification*, 6(3), 81-129.
- Baars, B. J., & Gage, N. M. (2010). *Cognition, brain, and consciousness: Introduction to cognitive neuroscience* (2nd ed.). Oxford, UK: Academic Press.
- Behrman, A. (2022). *Speech and voice science* (4th ed.). San Diego, CA: Plural Publishing.
- Bentler, R., Mueller, G., & Ricketts, T. A. (2016). *Modern hearing aids: Verification, outcome measures, and follow-up*. San Diego, CA: Plural Publishing.
- Byrne, D., Dillon, H., Tran, K., Arlinger, S., Wilbraham, K., Cox, R., Hagerman, B., ... Ludvigsen, C. (1994). An international comparison of long-term average speech spectra. *The Journal of the Acoustical Society of America*, 96(4), 2108-2120.
- Cho, S., & Munro, M. J. (2017, September). F0, long-term formants and LTAS in Korean-English bilinguals. *Proceedings of the 31st General Meeting of the Phonetic Society of Japan* (pp. 188-193). Tokyo, Japan.
- Cox, R. M., & Moore, J. N. (1988). Composite speech spectrum for hearing and gain prescriptions. *Journal of Speech, Language, and Hearing Research*, 31(1), 102-107.
- Coxhead, A., Nation, P., & Sim, D. (2015). Measuring the vocabulary size of native speakers of English in New Zealand secondary schools. *New Zealand Journal of Educational Studies*, 50(1), 121-135.
- Dunn, H. K., & White, S. D. (1940). Statistical measurements on conversational speech. *The Journal of the Acoustical Society of America*, 11(3), 278-288.
- Eady, S. J. (1982). Differences in the F0 patterns of speech: Tone language versus stress language. *Language and Speech*, 25(1), 29-42.
- French, N. R., & Steinberg, J. C. (1947). Factors governing the intelligibility of speech sounds. *The Journal of the Acoustical Society of America*, 19(1), 90-119.
- Jespersen, O. (1904). *How to teach a foreign language*. London, UK: George Allen & Unwin.
- Johnson, K. (2011). *Acoustic and auditory phonetics* (3rd ed.). Oxford, UK: Wiley-Blackwell.
- Keating, P., & Kuo, G. (2012). Comparison of speaking fundamental frequency in English and Mandarin. *The Journal of the Acoustical Society of America*, 132(2), 1050-1060.
- Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, 5(11), 831-843.
- Ladefoged, P., & Johnson, K. (2014). *A course in phonetics* (7th ed.). Stamford, CT: Cengage Learning.
- Li, F., Edwards, J., & Beckman, M. (2007, August). Spectral measures for sibilant fricatives of English, Japanese, and Mandarin Chinese. *Proceedings of the XVIth International Congress of Phonetic Sciences* (pp. 917-920). Saarbrücken, Germany.
- McLoughlin, I. V. (2016). *Speech and audio processing: A MATLAB®-based approach*. Cambridge, UK: Cambridge University Press.
- Mennen, I., Schaeffler, F., & Docherty, G. (2012). Cross-language differences in fundamental frequency range: A comparison of English and German. *The Journal of the Acoustical Society of*

- America*, 131(3), 2249-2260.
- Moon, C., Lagercrantz, H., & Kuhl, P. K. (2013). Language experienced *in utero* affects vowel perception after birth: A two-country study. *Acta Paediatrica*, 102(2), 156-160.
- Nation, P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, 31(7), 9-13.
- Noh, H., & Lee, D. H. (2012). Cross-language identification of long-term average speech spectra in Korean and English: Toward a better understanding of the quantitative difference between two languages. *Ear and Hearing*, 33(3), 441-443.
- Olsen, W. O., Hawkins, D. B., & Van Tasell, D. J. (1987). Representations of the long-term spectra of speech. *Ear and Hearing*, 8(5), 100S-108S.
- Plack, C. J. (2018). *The sense of hearing* (3rd ed.). New York, NY: Routledge.
- Schnupp, J., Nelken, I., & King, A. J. (2012). *Auditory neuroscience: Making sense of sound*. Cambridge, UK: MIT Press.
- The Math Works. (2022). MATLAB (version R2022a) [Computer software]. Retrieved from <https://www.mathworks.com/>
- Vasilescu, I., Dutrey, C., & Lamel, L. (2015, October). Large scale data based linguistic investigations using speech technology tools: The case of Romanian. *Proceedings of the 2015 International Conference on Speech Technology and Human-Computer Dialogue (SpeD)* (pp. 1-6). Bucharest, Romania.

• **박순 (Soon Park)**

서울대학교 대학원 영어교육과 박사과정
 서울특별시 관악구 관악로 1
 Tel: 02-880-7670
 Email: drspark@snu.ac.kr
 관심분야: 음향음성학, 영어교육학, 신경언어학, 청각학

• **안현기 (Hyunkee Ahn)** 교신저자

서울대학교 영어교육과 교수
 서울특별시 관악구 관악로 1
 Tel: 02-880-7673
 Email: ahnhk@snu.ac.kr
 관심분야: 영어음성학, 영어교육학, 인공지능, 머신러닝

과학영재학교 재학생 영어발화 주파수 대역별 음향 에너지 분포의 영어 성취도 예측성 연구*

박 순 · 안 현 기

서울대학교 영어교육과

국문초록

본 연구는 미국영어 화자의 평균적 음향 스펙트럼 분포를 확보한 후 과학영재학교 재학생의 영어발화 양상을 비교하여 상대적으로 우수한 지적 역량을 갖춘 우리나라 과학영재들의 초분절적 영어 유창성 양상을 규명하고, 그 근접성 정도가 영재학교 학생의 영어 과목 정기고사 성취도와 어떤 관계성을 갖는지 탐구하고자 진행되었다. 불과 수 초에서 수십 초에 불과한 음성녹음 데이터 위주로 분석을 시행했던 종래의 연구와 달리 총 4시간에 달하는 미국 영어 원어민 화자(남성 15명, 여성 15명)의 음성녹음 자료를 MATLAB(R2022a; The Math Works) 코드로 분석하여 20-20,000 Hz 주파수 범위 내의 대역별로 장기 스펙트럼 음향에너지 분포값을 확보했으며, 이를 기준으로 과학영재학교 1학년 신입생 80명의 녹음데이터 LTASS(long-term average speech spectrum) 분석 수치와 비교한 결과, 영어 과목 학기말고사의 학업성취도 수준이 상위 30% 이내인 학생들의 표본을 제외하고는 미국영어 음향에너지 분포와의 근접성이 통계적으로 유의하지 않다고 밝혀졌다. 영재학교 입학 후 영어 성취도를 예측하기 위한 지표를 발견하기 위해 수용성 어휘크기검사(receptive vocabulary size test), 학기 중 복수 회 실시한 영어 어휘 형성평가 퀴즈 누적 점수, 공인 영어말하기시험(English Speaking Proficiency Test, ESPT) 성취도를 추가 변량으로 하여 정기고사 성취도와의 상관관계 분석 및 각 변량 간 선형 회귀분석을 시행하였는데, 대개 유년시절 완성되는 영어 유창성을 측정하는 ESPT보다는, 1학기 및 2학기 초 실시한 수용성 어휘크기검사 및 수과학 분야 저빈도 어휘 위주 형성평가 점수와의 통계적 유의성이 월등히 높다는 사실이 관찰되었다. 따라서, 본 연구로부터 확보된 이론적 기반을 토대로 국내 영재학교에서는 발음교육보다 과학영재를 주요 대상으로 한 전문적 수준의 저빈도어휘 교육이 보다 효과적인 교수 요목이라 추정할 수 있다.

핵심어: 음향학, LTASS (long-term average speech spectrum), 영어발화 분석, 과학영재학교, 어휘교육

* 본 연구는 서울대학교 2020년도 연구지원을 받아 수행되었음(과제번호: 700-20200053).