

한국인 영어 발음의 좋음과 나쁨 인지 평가에 영향을 미치는 초분절 매개변수 연구

A study on the Suprasegmental Parameters Exerting an Effect on the Judgment of Goodness or Badness on Korean-spoken English

강석한¹⁾ · 이석재²⁾

Kang, Seokhan · Rhee, Seok-Chae

ABSTRACT

This study investigates the role of suprasegmental features with respect to the intelligibility of Korean-spoken English judged by Korean and English raters as being good or bad. It has been hypothesized that Korean raters would have different evaluations from English native raters and that the effect may vary depending on the types of suprasegmental factors. Four Korean and four English native raters, respectively, took part in the evaluation of 14 Korean subjects' English speaking. The subjects read a given paragraph. The results show that the evaluation for 'intelligibility' is different for the two groups and that the difference comes from their perception of L2 English suprasegmentals.

Key words: pronunciation evaluation, English education, prosody, F0, duration, pause

1. 서론

영어 발음 평가에 대한 기존 연구들은 영어 원어민과 한국인 평가자 사이에 높은 상관관계를 나타낸다고 주장하였다(이석재, 박전규, 2003; 정현성 외 4인, 2008; 윤원희, 2009; Lee C-H, 2010). 이석재, 박전규(2003)는 음성학 훈련을 받은 한국인 전문가 그룹, 비전문가 그룹, 원어민 그룹 간의 평가와 자동 인식기를 기반으로 한 기계식 평가 사이의 상관관계를 연구하였다. 연구 결과는, 한국인 평가자 사이의 상관관계가 가장 높고 ($r=0.98$), 전문가/비전문가 한국인 평가자와 원어민 평가자와의 상관관계 ($r=0.92$)도 높았으며, 인간 평가와 기계 평가 사이에도 유의한 수준의 상관관계가 있었다($r=0.72$). 정현성 외 4인(2008)에서도 영어 음성학 박사급 인력 5명과 원어민 3명이 참여한 발음 평가에서 상당히 높은 수준의 상관관계 유의미도를 보여

주었다($p<0.0001$). 윤원희(2009)에서도 비록 일부 문장에서는 한국인과 원어민 사이에 약간의 평가가 달라지고 있지만, 대체적으로 두 집단 사이의 평가는 상당히 유의미한 일치도를 지니는 결과를 보고하고 있다.

이러한 선행연구에서 한국인과 원어민 평가자의 평가 결과가 매우 유사하다는 결과가 있음에도 불구하고 일부 항목에서는 통계적으로 유의미할 정도의 차이를 보이는 것으로 보고되고 있다. 윤원희(2009)에서는 숙련도(proficiency), 리듬(rhythm), 억양(intonation), 이해도(intelligibility)로 나누어진 평가 영역에서 이해도 측면에서 한국인과 원어민의 차이를 보이는 것으로 보고되고 있다. Lee C-H(2010)은 13명의 영어 원어민평가자와 9명의 한국인 평가자들이 3명의 한국인 영어 발화를, 듣기, 문장완성, 대담, 문법 및 문장 구조, 어휘, 발음, 능숙도, 응답 속도 등 8개의 항목을 연구하였다. 결과는 문법 영역과 발음 영역에서 상당한 정도의 두 집단 간 차이를 보이는 것으로 보고하고 있다.

그렇지만 평가 영역에서 발생하는 비원어민 발음에 대한 원어민과 비원어민의 평가 결과 차이가 비원어민 발화 요소 중 어떤 음성신호에 기인하는 것인지는 상대적으로 연구가 적었다. 제2언어 학습분야에서, 비원어민의 L2 발화의 외국인투의 발음 및 억양에 영향을 미치는 요소로서 주로 연령(Tahta,

1) 서울대학교, kang45@snu.ac.kr, 제1저자

2) 연세대학교, scrhee@yonsei.ac.kr, 교신저자.

접수일자: 2011년 5월 31일

수정일자: 2011년 6월 19일

게재결정: 2011년 6월 20일

Wood, Loewenthal, 1981; Guion, Flege, Liu, Yeni-Komshina, Grace, 2000), 경험(Flege, Liu, 2001; Mennen, 2004, Trofimovich, Baker, 2006; Kang, Guion, Ahn, Rhee, 2010), 모국어(Archibald, 1995, 1998; Davis, Kelly, 1997; Guion, Harada, Clark, 2004), 동기(Conrad, 1991; Moyer, 1999)와 같은 여러 요소들이 제2언어 학습자들의 이해 가능한 발화 및 인지에 영향을 미친다고 주장하고 있다.

이러한 요소들과 관련된 음성 분석은 대체적으로 분절적 접근과 초분절적 접근 두 가지로 이루어지고 있다. 일반적으로 목표어로서의 능숙한 발화는 분절적 요소(Jenkins, 2000; Munro, Derwing, 2008)와 초분절적 요소(Anderson-Hsieh, Koehler, 1988; Tajima, Port, Dalby, 1997)가 제2언어 학습자의 발화 이해도에 영향을 미치는 것으로 보고하고 있다. 그러나 상대적으로 많은 연구들은 초분절 요소가 원어민과 유사한 발음 구현에 더 큰 영향력을 발휘한다고 주장하고(James, 1976; Anderson-Hsieh, Johnson, 1992; Bradlow, Toretta, Pisoni, 1996), 다양한 언어 환경에서 영어 원어민은 의미 있는 F0 곡선을 사용하는 반면에, 제2언어로서의 영어 학습자들은 이런 곡선을 거의 사용하지 못한다고 보고하고 있다. Kang, Guion, Rhee, Ahn(2010)에서는, 한국어 영어 학습자의 경우 F0 곡선의 이용 부족은 단조로운 억양 형태를 유발하고, 필연적으로 매우 완만한 억양 기울기를 도출하지만, 이에 반해 영어권 체류 경험자들은 다양한 억양 형태를 사용함으로써 영어 원어민과 유사한 발화를 구현한다고 보고하고 있다.

일반적으로 초분절 요소는 음성학적으로 F0, 발화 속도, 휴지 빈도 및 구간, 강도(dB)등의 신호로 구현되며, 일반적으로 L2 학습자들이 목표어에 능숙함에 따라 이런 초분절 요소가 목표어에 접근한다는 사실은 많이 알려져 있다(Guion et al., 2000). 한국어인 영어 학습자들의 경우 목표어인 영어에 능숙해짐에 따라, F0 범위가 늘어나고(이석재, 조철현, 문선영, 2003; Kang, Guion, Ahn, Rhee, 2010), 발화 속도가 빨라지며(Guion et al., 2000), 문장 말 음절 구간이 길어지고(Kang, Ahn, 2010), 휴지 빈도가 줄어들며 구간 길이가 짧아진다고(Trofimovich, Baker, 2006) 보고하고 있다. 제2언어 학습 분야의 연구를 종합하면, 고도의 발음 이해도를 지닌 L2 발화자는 목표어인 영어 발화에서 폭넓은 F0 간격, 상대적으로 빠른 발화 속도, 그리고, 짧은 휴지 구간을 지닌다고 볼 수 있다.

한편 지금까지 진행되어온 영어 교육 분야의 발음 평가 연구들은 주로 한국어인 평가자들의 평가가 신뢰성을 가질 수 있는지에 대하여 주로 연구가 이루어져 왔다(이석재, 박진규, 2003; 정현성의 4인, 2008; 윤원희, 2009; Lee C-H, 2010). 본 연구는 여기서 한 걸음 더 나아가 만약 외국인과 한국어인 평가자들 사이에 발음 영역 평가에 차이가 있다면, 모국어 변수에 의하여 어떤 음성학적인 매개 변수가 평가에 영향을 미치는지에 대하여 밝히려고 한다. 이 연구는 한국어인 평가자를 대상으로 한 영어

평가가 발음 영역에서 어느 정도 신뢰성을 확보할 수 있는지에 대한 연구와 관계가 있으며, 앞으로의 영어 말하기 평가의 보완점 및 대책을 음성학적인 측면에서 접근하려고 한다.

본 논문의 구조는 2장에서 한국어인 영어 학습자의 발화와 이에 대한 영어 원어민과 한국어인 평가자들의 인지평가, 두 가지로 이루어진 실험 방법을 다룰 것이고, 3장에서 그 결과를 제시할 것이다. 4장 토의 부분에서는 한국어인 평가자와 원어민 평가자 사이에 차이를 보이는 발음 이해도(intelligibility)에 대하여 비교 연구를 한 후, 이러한 차이를 여러 초분절 매개 변수 위주로 토론을 할 것이다. 그리고 5장에서는 결론이 주어질 것이다.

2. 실험

2.1. 발음 평가

인지 평가 자료로 사용될 음성 자료의 녹음을 위하여 14명의 한국어인 발화자(남성 9명, 여성 5명)가 주어진 영어 단락을 읽었다. 이 발화자들은 주로 서울에 위치한 한 대학교 1학년 학생들이고, 전공은 다양하며, 대학입시 영어 수학능력 시험 성적이 1-2등급에 속하는 우수한 학생들이다. 이들 중 어떤 피험자도 6개월 이상의 영어 사용권 국가로의 체류 경험이 없다. 그들의 발음은 조용한 교실에서 14대의 컴퓨터를 통하여 화면에 읽기 자료가 제시되었고, 피험자가 이를 읽으면 컴퓨터에 그들의 발화가 녹음되었다. 이 녹음된 읽기 자료는, 약 1주일 후 4명의 내국인 평가자와 5명의 외국인 평가자³⁾가 발음 영역에서 개별적으로 평가를 하였다. 평가자들은 대부분 영어 영문학 분야의 박사 학위 소지자들이며, 대학에서 교양영어를 담당하고, 연령이 33세-46세 정도 되는 강의 경력이 5-12년 정도 되는 중견 교/강사들이다. 평가자들에게는 소정의 평가료가 주어졌다. 발화자 및 평가자들의 구체적인 정보는 다음과 같다.

표 1. 발화자 정보

ID	성별	연령
f1	여	19
m1	남	19
f2	여	19
f3	여	20
m2	남	22
m3	남	21
m4	남	20
f4	여	19
m5	남	19
m6	남	19
m7	남	20
m8	남	20

3) 수정 논문에서는 한명의 외국인 평가자를 교체했다. 이유는 다른 평가자에 비하여 경력이 너무 차이가 났고(30년), 평가가 다른 평가자에 비하여 확연히 다르게 나타났기 때문이다.

m9	남	21
f5	여	19

표 2. 평가자 정보

ID	한국인 평가자				원어민 평가자			
	R1	R2	R3	R4	R5	R6	R7	R8
연령(세)	46	41	38	34	33	38	31	43
강의경험(년)	12	4	5	9	7	5	6	9
성별	남	남	여	여	여	남	남	여

평가는 5 단계 평가 방법을 이용하였다. 영어 말하기 평가 분야에서는 대규모 국제 공인 시험인 경우 대부분 9-12단계 평가방법을 채용하는 경향이 있다. 예를 들어, ACTFL-OPI는 10등급, ISLPR은 12등급, MATE 12등급을 이용한다. 그러나, 일부 시험에서는 5-8단계 등급을 이용하는데 이는 평가기관의 규모 및 시험 성격에 따라 달라진다 (참고: G-TELP 5등급, ICAO 6등급, FCE 6등급). 본 평가는 연구용으로 제작되었기 때문에 5 단계 방법을 채용했다. 평가 방법은 OPIc(2010)을 참고하였다. 점수는 0에서 4점까지 5단계를 제시하였다. 다음은 구체적인 평가 루브릭이다.

(1) 평가 분류

- 4: 외국인 어투의 액센트가 거의 없고, 발화 실수가 존재하지 않는다. 발화자는 정확한 의미 전달을 위하여 다양한 억양을 구사하고, 문장 강세를 적절하게 위치시킬 수 있다.
- 3: 가끔씩 비-원어민 발화가 나타나지만, 발화는 항상 이해도를 지닌다. 명확하고 자연스런 발화와 억양을 습득하였다. 문장 강세를 적절히 위치시키지만, 이해도에 영향을 미치지 않는 미소한 오류를 나타낸다.
- 2: 외국인 어투의 억양이 산발적으로 명확히 나타나고, 발화 오류가 생기지만, 발음은 어느 정도 이해 가능하다. 억양과 문장 강세는 적절하지만, 원어민 이해도에 어느 정도의 혼란을 불러일으킨다.
- 1: 상당한 정도의 외국인 액센트에도 불구하고 일반적으로 발화는 이해할만하다. 그러나 분절음의 실수와 외국인 어투의 억양 및 강세가 빈번하여 이해도를 상당히 떨어뜨린다.
- 0: 발화자의 발음은 완전히 이해 불가하다.

평가자들은 1회(확실한 경우)에서 3회 정도(불확실한 경우) 듣고 개별적으로 평가하도록 하였다. 평가 시에 헤드셋 착용은 자율로 결정하도록 하였다.

2.2. 발화 분석

14명의 발화자는 컴퓨터 화면에 떠오른 다음 문단을 읽도록 하였다. 읽기 전에 45초의 준비시간이 주어졌다.

Several years ago, the scientists of SNJ developed ways of investigating the nature of the atmosphere of the past by studying air cut in the ice around the North or South Pole. According to their theory, when snow falls, air is trapped between the snow flakes. The snow turns to ice with the air still inside. Over the years more snow falls on top, making you layers of ice. But the trapped air these scientists believe remains exactly as it was when the snow originally felt. To find what air is like 300 years ago, you use a drill in the shape of the hollow tub to cut deep into the layers of ice.

분석에 사용된 매개 변수는 F0 범위, 발화 속도, 휴지 구간의 세 영역을 설정하였다. 이 매개 변수들은 발화가 능숙함에 따라 통계적으로 가장 유의미하게 달라지고, 발화 평가에 결정적인 영향을 미친다고 알려진 것이다(Trofimovich, Baker, 2006; Kang, Guion, Ahn, Rhee, 2010).

가. F0 범위: F0 범위는 영어 능숙도를 나타내는 중요한 인자이다(Backman, 1979; Willems, 1982; 이석재, 조철현, 문선영, 2003). 이 연구에서는 발화자의 전체 단락에서 최저 수치의 F0와 최고 수치의 F0를 측정하였다. 기계 상의 오류를 제거하기 위하여 자동 측정 대신에 수작업을 통하여 직접 눈으로 확인하여, 최소 5ms 정도의 기본 주파수가 지속되고, 이전 발화 F0 곡선에 이어지는 경우에만(즉, 일시적으로 나타나는 단속적인 F0를 제거하여) 측정치로 삼았다. 기본적인 측정 방법은 Ladefoged (2003)을 따랐다.

나. 발화 속도: 발화 속도(speech rate)는 제2언어 학습자의 능숙도를 측정하는 중요한 인자이다(Derwing & Munro, 1997; Guion et al., 2000). 본 연구에서는 1초에 몇 개의 음절이 발화하는지를 측정하였다. 측정 방법은 웨이브 폼과 스펙트로그램을 통하여 각 음절의 시작점과 끝점을 측정하였고, 이 측정에는 자음인 경우 웨이브 폼의 움직임, 자음의 과일, 마찰구간을 참고했고, 모음인 경우 F0 움직임, 제2 포먼트를 주요 측정 지침으로 삼았다. 측정 방법은 Ladefoged(2003)을 따랐다.

다. 휴지 길이: 휴지 길이는 제2외국어 화자의 발화 이해도를 결정짓는 중요한 요소이다(Cenoz, 2000). 이 휴지 길이는 휴지 빈도와 더불어 모국어 배경과는 관계없이, L2 학습자들에게 명백히 나타나는 범언어적인 현상이다(Trofimovich, Baker, 2006). 본 연구에서는 발화자가 읽은 전체 단락 중 발음된 부분을 제외한 전체 휴지 구간을 측정하였고, 이를 60초당 휴지 길이로 계산하였다.

3. 결과

3.1. 발음 평가

한국인 4인, 외국인 4인이 발음 평가에 참여하였고, 그 결과는 다음과 같다.

표 3. 두 집단 각 평가자의 평가 평균 점수 및 표준 편차

ID	한국인 평가자					외국인 평가자				
	R1	R2	R3	R4	소계	R5	R6	R7	R8	소계
평균	2.29	2.00	2.21	2.79	2.32	1.14	1.93	1.71	1.77	1.77
표준 편차	0.61	0.68	0.70	0.58	0.69	0.36	0.73	0.72	0.96	0.71

발음 평가의 결과를 놓고 보면, 한국인 평가자들은 비교적 후한 점수를 주며, 그 편차는 작지만, 원어민은 상대적으로 낮은 점수를 부여하고, 그 편차는 상대적으로 크다. 각 집단 평가자들의 평가들 사이의 신뢰도를(inter-rater reliability) 측정을 위하여, 평가자들의 평가점수에 대한 선형 관계를 나타내는 Pearson correlation 지수를 구하였다. 그 결과는 $r=0.478$ ($p<0.01$)로 상당히 유의미하게 나타났다. 이는 평가자들의 평가가 집단 요인에 의하여 상당한 영향을 받고 있음을 보여주고 있다. 한편 한 평가자에 대하여(intra-rater) 얼마나 체계적으로 평가를 내리는지를 살펴보기 위하여 Spearman correlation 지수를 구하였다. 지수값은 $r=0.409$ ($p<0.01$)로서 상당히 유의미하게 나타났다. 이는 각각의 평가자들의 14명의 피험자 발음에 대한 평가들이 비교적 일관성 있게 진행되고 있음을 보여주고 있다.

평가자들의 평균 분포를 도표로 나타내면 다음과 같다.

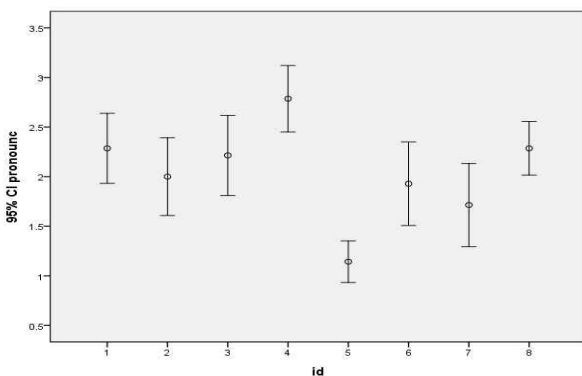


그림 1. 한국인 평가자(1-4번)와 원어민 평가자(5-8번)들의 평가 분포.

한국인 평가자들은 0-4점 분포에서 평균 1.5에서 3점 사이에서 균일하게 평가를 하고 있다. 반면에 영어 원어민들은 평균 0.7점에서 2.5점 사이에서 평가를 하고 있다. 이는 원어민들의 발음 평가가 상당히 엄격하다는 기존의 연구 결과를 뒷받침하

고 있다(Lee C-H, 2010). 또한 원어민들의 평가 분포는 한국인들에 비하여 비교적 비균질적이다. 이는 그들의 평가가 상당히 자유롭다는 것을 의미한다.

각 집단이 부여하는 평균 점수는 매우 다르게 나타났다. 한국인 평가자들은 평균 2.32이 부여된 반면에, 영어 원어민들은 평균 1.77점을 부여하고 있었다. 다음 표(표 4)는 각 발화자에 대한 한국인 및 원어민의 평균 평가 점수와 이를 좋음(A)/보통(B)/나쁨(B)으로 구분 판정한 도표이다. 각 세단계의 판정은 각 집단별로 상위 30%는 좋음(A), 40%는 보통(B), 그리고 30%는 나쁨(C)으로 판정을 하였다. 이는 각 기관에서 이루어지는 영어 말하기 점수가 최종적으로는 상대적 단계 평가방식으로 연결되기 때문에, 개별적 점수를 집단별 판정으로 변환하여 비교 연구하였다.

표 4. 각 평가자들이 평가한 점수 및 판정표

발화자	한국인 평가		원어민 평가	
	평균 점수	판정	평균점수	판정
f1	2.25	보통(B)	1.50	나쁨(C)
m1	2.75	좋음(A)	1.75	보통(B)
f2	2.0	보통(B)	1.75	보통(B)
f3	2.50	보통(B)	2.75	좋음(A)
m2	1.50	나쁨(C)	1.25	나쁨(C)
m3	2.50	보통(B)	1.75	보통(B)
m4	2.50	보통(B)	1.25	나쁨(C)
f4	3.0	좋음(A)	3.25	좋음(A)
m5	2.75	좋음(A)	2.0	좋음(A)
m6	2.75	좋음(A)	2.25	좋음(A)
m7	1.75	나쁨(C)	1.25	나쁨(C)
m8	2.25	보통(B)	2.0	보통(B)
m9	2.0	나쁨(C)	1.75	보통(B)
f5	2.0	나쁨(C)	1.75	보통(B)

한국인 평가자들은 좋은 발음인 경우 평균 2.81, 보통 발음인 경우 2.33, 나쁜 발음인 경우 평균 1.81점을 부여하고 있는 반면, 외국인 평가자들은 좋은 발음인 경우 평균 2.56, 보통 발음인 경우 1.79, 나쁜 발음인 경우 1.31점을 부여하고 있다. 두 집단의 각 단계별 평균 점수를 비교해 보면, 보통 단계 (B) 평균 점수는 두 집단이 비슷하지만, 영어 원어민의 좋음(A)과 나쁨(C)의 평가는 한국인 평가자에 비하여 점수가 하락하고 있다. 이는 원어민 평가가 한국인 평가에 비하여 엄격하다는 의미이다. 두 집단사이에 동일한 평가를 받은 피험자는 8명이며, 나머지 6명은 평가 판정이 각기 달랐다.

3.2. 발화 분석

아래 표 5는 두 집단 간의 좋은/나쁜 발음이라고 평가한 각 집단의 상위층 30% 및 하위층 30%, 계 60%의 발화자들에 대

하여 F0 범위, 휴지 길이, 발화 속도의 평균과 표준 편차를 측정하는 것이다.

각 측정된 음향 분석에서 원어민 평가자들은 한국인 평가자에 비하여 F0 범위와 발화 속도에서 좋은 발음과 나쁜 발음 사이에 뚜렷한 차이를 보이고 있다. 그러나, 휴지 길이에 대해서는 한국인 평가자들의 차이가 원어민보다 더 컸다.

표 5. 각 집단 별 좋은 발음/나쁜 발음 간의 F0 범위, 휴지 길이, 발화 속도의 평균값과 표준편차

Measure	원어민 평가자		한국인 평가자	
	나쁨	좋음	나쁨	좋음
F0 범위	129 (54)	139 (32)	118 (47)	127 (43)
휴지 길이	0.21 (0.05)	0.20 (0.04)	0.25 (0.05)	0.21 (0.06)
발화 속도	2.99 (0.15)	3.28 (0.43)	3.22 (0.14)	3.24 (0.42)

가. F0 범위

F0 범위는 L2의 배경언어와 밀접한 관계를 가진다(Van Bezooijen, 1995; Scherer, 2000). 따라서 영어 원어민과 유사한 한국인 영어 학습자의 F0 범위 발화 및 인지는 영어 발화 이해도와 밀접한 관계를 맺는다.

한국인 평가자들의 좋은 발음은 F0 범위가 평균 110 Hz 정도이고, 나쁜 발음의 평균은 평균 107 Hz 정도였다. 반면에 원어민 평가자들을 살펴보면, 좋은 발음과 나쁜 발음사이의 격차는 더 커진다. 그들이 평가하는 좋은 발음은 평균 139 Hz 정도가 되며, 나쁜 발음의 평균도 129 Hz였다. 도표로 나타난 결과는 다음과 같다.

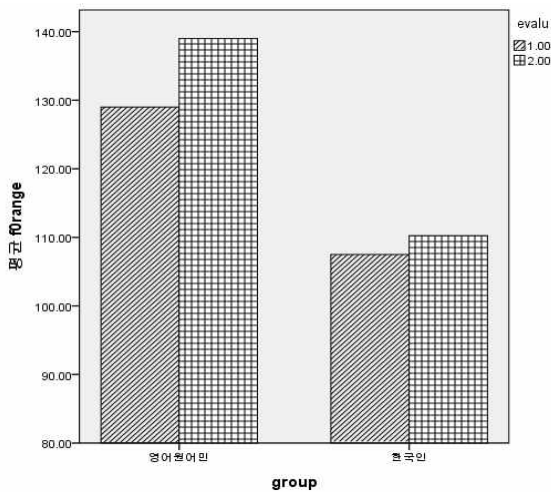


그림 2. 두 집단에 의하여 분류 평가된 F0 범위 (1: 나쁜 발음, 2: 좋은 발음)

위 그림에서 보듯이 원어민 판정 발음이 더 명확한 F0 차이

를 가지고 온다. 원어민 집단에서는 좋은/나쁜 발음 사이의 F0 차이가 평균 27 Hz 정도이지만, 한국인 집단에서는 그 차이가 불과 9 Hz이다. 이는 영어 원어민이 선정한 F0 범위가 큰 좋은 발음이 한국인 경우 나쁜 발음으로 선정되어, 한국인 평가의 F0 범위의 격차가 좁아지는 경우가 있기 때문으로 분석된다.

나. 휴지 길이

휴지 길이는 목표어 발화의 어려움을 나타내며, 이는 L2 발화의 정보 해석과 기억과 관련이 있는 것으로 추정된다 (Schachter, Christenfeld, Ravina, Bilous, 1991). 흥미로운 사실은 휴지 구간은 배경 언어나 문화권과는 관계없이 L2 초보자에게 더 흔하게 나타나는 범언어적인 현상이라는 것이다(Holmes, 1995).

한국인 평가자들의 판정한 나쁜 발화는 휴지 길이가 평균 초당 0.25초로, 좋은 발화는 평균 초당 0.21초로 판정하고 있다. 이는 영어 원어민과 매우 유사하다. 원어민인 경우 나쁜 발화는 초당 0.21초, 좋은 발화는 초당 평균 0.20초이었다. 이는 휴지 길이가 갖는 한국인들의 인지 능력이 원어민들보다 더 민감하다는 것을 의미한다. 확실히 한국인 평가자들은 F0 범위보다 발화 및 휴지 구간에 대하여 더 민감하다. Kang, Ahn (2010)의 한국인과 영어 원어민의 F0 범위 및 구간에 대한 발화와 인지 연구에서는, 한국인 영어 학습자들이 F0 보다 휴지 구간에 더 민감하게 반응하는 것으로 보고하고 있다.

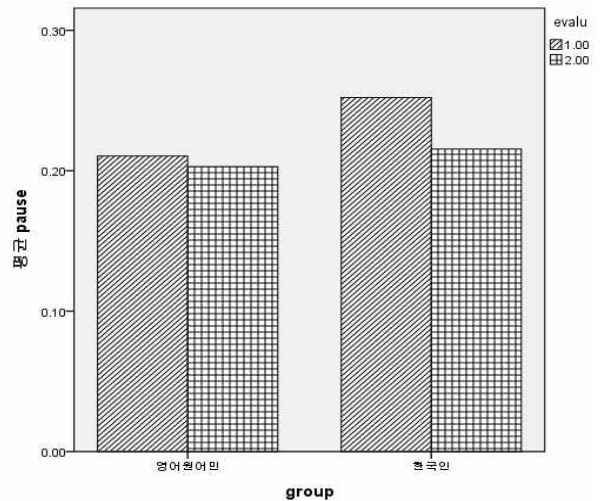


그림 3. 두 집단에 의하여 분류 평가된 휴지 길이 (1: 나쁜 발음, 2: 좋은 발음)

휴지 길이가 짧을수록 단어 사이의 연결이 자연스럽게 이어진다는 점을 고려한다면, 두 집단 모두 지속적인 발화를 좋은 발음의 한 요소로 여긴다.

다. 발화 속도

L2 초보자들은 L2 경험자나 L1 원어민보다 발화 속도가 느린 것으로 보고되고 있는데, 이는 음운론적인 정보 해석의 어려움, 조음 문제, 심리적인 요소등과 관련이 있다(Munro, Derwing, 1995, 2001). 명백히 느린 발화는 외국인 억양의 신호(Anderson-Hsieh, Koehler, 1988)로 여기며, 발음 평가 상에도 영향을 미친다(Zuengler, 1988).

한국인 평가자들의 좋은 발음은 초당 음절수가 평균 3.24개이며, 나쁜 발음은 초당 평균 3.22개이다. 반면에 원어민 평가자들의 좋은 발음은 초당 음절수가 평균 3.28이며, 나쁜 발음은 평균 2.99개이다. 초당 음절수가 많을수록 빠른 발화임을 감안하면, 두 집단 모두 빠른 발화에 대하여 좋은 평가를 주고 있지만, 한국인 평가자의 좋음/나쁜 발음 차이는 이 변수에서는 거의 차이를 보여주지 못하고 있다.

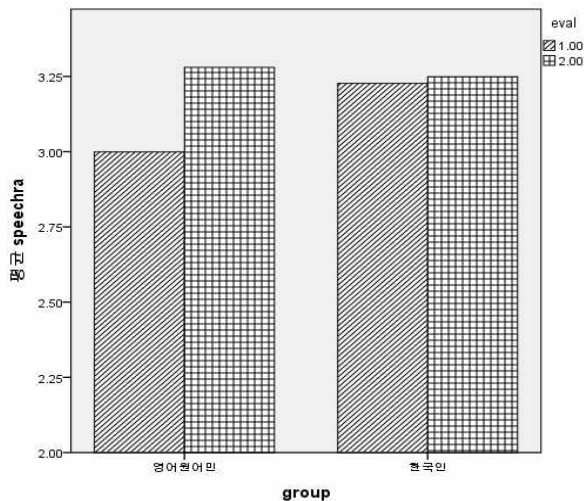


그림 4. 두 집단에 의하여 분류 평가된 발화 속도
(1: 나쁜 발음, 2: 좋은 발음)

위 그림에서 보듯이 원어민 판정 발음이 더 명확한 발화 속도 차이를 가지고 온다. 원어민 집단에서는 좋은/나쁜 발음 사이의 발화 속도 차이가 초당 평균 0.30개의 음절이지만, 한국인 평가자들의 판정 차이는 초당 0.02개에 불과하다. 이는 영어 원어민의 판정이 뚜렷한 차이를 보이는 반면, 한국인들의 판정은 발화 속도의 차이가 원어민 보다 떨어진다.

4. 토의

4.1. 발음 평가

전술한 바대로 집단 요소가 각 평가자들의 평가에 상당한 영향을 미친다($r=0.478$, $p<.001$). 이는 한국인 집단과 원어민 집단의 평가경향이 서로 다르게 나타나고 있음을 보여준다. 한국인 평가자들의 발음 평가에 대한 경향은 평가기준의 이해와 심리

적인 확실성의 문제로 요약할 수 있다. 발음 평가 지침에 기술된 평가 방법에 대하여 대부분의 한국인 평가자들은 기준 적용에 혼란을 겪고 있었다. 실제적으로 이런 기준 적용 문제는 일부 연구에서 언급된 것이며(Lumley, McNamara, 1995; 신동일, 2006), 다만 발음 영역 평가인 경우 외국어로서 영어를 사용하는 한국인 평가자에게는 그 혼란의 정도가 더 클 수 있다는 점이다. 기준 적용에 대한 문제는 자연히 한국인 평가자들 간의 평가 일치도를 증가시키고 있다. 이는 기준 적용에 대한 미숙함이 중위 점수대인 2점(4점 만점) 주위로 몰리는 현상을 가속화시킨다. 평가자들은 상위 점수인 만점을 주는 것과 하위 점수인 0점 혹은 1점을 주는 것을 극도로 회피한다. 이런 경향은 상위권 및 중상위권 평가 참여자들에게는 손해로 작용하며, 역으로 중하위권 및 하위권 학생들에게는 이익으로 작용한다.

통계 집단인 영어 원어민 평가자들도 이러한 평가 상의 문제에서 자유로울 수 없다. 일정부분 이 집단에서도 기준 적용에 대한 혼란현상을 볼 수 있으며, 중위권 집중 현상도 나타난다. 무엇보다 평가자에 의하여 비교적 균질한 요소를 보이는 한국인 평가자와 달리, 집단 내에서의 평가자별 분포가 다양하게 벌어지는 것을 발견할 수 있는데, 이는 평가자의 권한이 더 크게 작용함을 의미한다.

4.2. 음향 분석

보편적으로 한국인들의 평가에 대한, 피험자 음성의 초분절 요소 분석은 대체적으로 외국인 평가와 그 추세를 같이한다. 즉, 좋은 발음은 폭넓은 F0 범위, 짧은 휴지 구간, 상대적으로 빠른 발화를 좋은 발음이라고 평가한다. 그러나 한국인 평가자들은 그 범위가 원어민과 상당히 차이를 보인다. 대표적으로 F0 범위 영역에서 한국인들이 판정한 좋은 발음과 나쁜 발음 사이의 차이가 원어민에 비하여 상당히 떨어지는 것으로 나타났다. 한국인들의 좋은 발화와 나쁜 발화 사이의 F0 범위 차이는 불과 9 Hz로, 원어민 평가의 27Hz에 비하여 상당히 낮다. 이는 모국어의 영향으로 보이는데, 강석한, 이석재(2011) 연구에 의하면 한국인이 발화한 한국어의 특징은 영어에 비하여 상대적으로 높은 F0 평균 수치와 낮은 F0 범위를 나타낸다고 보고하고 있다. 그리고, F0 범위는 문장 초보다는 문장 말에서 더 벌어진다.

F0 범위 차이에서 보듯이, 발화 속도에서도 비교적 원어민 평가에서 더 뚜렷하다. 역시 발화 속도에 있어서도, 한국인 발화자의 평가는 초당 음절수가 약 0.02개의 미소한 차이를 보이는 반면, 원어민 평가자들은 약 0.30개의 명확한 차이를 보여주고 있다. 그러나 휴지 구간은 거의 두 집단이 동일하다. 한국인 평가는 좋은 발화와 나쁜 발화 사이의 휴지 구간이 원어민과 거의 유사하다.

한국인 판정 발음이 초분절 요소에서 경계가 불분명한 것은 원어민 판정과 다른 발음들 때문으로 보인다. 14명의 발화를 판

정한 이 연구에서는, 14개 파일 중 6개 파일인 40%에 대하여 두 집단은 각기 다른 발음 판정을 내리고 있다. 이 각기 다르게 판정된 발음으로 인하여 두 집단사이의 초분절 요소에 대한 음성 기준이 달라지고 있다.

5. 결론

전반적으로 영어 발음 평가에서 한국인 평가자들은 원어민 평가자와 유사한 경향을 보일지는 모르지만, 초분절 요소에 대한 음성학적 분석에서는 확실한 차이가 있는 것으로 보인다. 이 부분에서 원어민은 한국인 평가자에 비하여 좀 더 명확한 차이를 보여주고 있었다. 이는 앞으로 한국인을 활용한 영어 발음 평가에서 새로운 접근법이 필요하다는 것을 의미한다. 앞으로 한국인에 의한 한국인 영어 발화 평가는 음성 교육을 통한 평가자 교육에 역점을 두는 것이 필요하다. 이 음성 교육 지도에는 제2언어 학습 분야의 연구 결과에서 도출된 중요한 신호들을 기반으로 하는 실제 음성 발화로 교육에 임하는 것이 좋을 것이다.

참고문헌

- Kang, S.H., & Rhee, S-C. (2011). "The study on prosody change for native English-spoken Korean", ms.
(강석한, 이석재 (2011). "영어 원어민 한국어 초분절 요소의 변화 연구", 미출판 논문.)
- Shin, D.I. (2006). *English evaluation in Korea*. Seoul: Hankuk Munhwasa.
(신동일 (2006). "한국의 영어 평가학". 서울: 한국문화사.)
- Rhee, S-C., & Park J-K. (2003). "The correlation study on evaluation of Korean-spoken English for Korean/native English/ILT", Conference Proceedings of Korean Phonetics Society, Autumn 2003, pp. 83-87.
(이석재, 박전규 (2003) "한국인의 영어 문장 발음에 대한 한국인/원어민/ILT 평가 점수 사이의 상관관계," 대한음성학회 2003 가을 학술대회 발표집, 83-87.)
- Rhee, S-C., Cho C-H, Moon S-Y. (2003). "F0 difference of English between Korean and native English speakers and its role in speaking evaluation", *Speech Sciences*, 10-4, 93-104.
(이석재, 조철현, 문선영 (2003). "한국인과 원어민 영어 발화의 F0 고저 범위 차이와 발음 평가에 있어서의 그 역할", *음성과학*, 10-4, 93-104.)
- Yoon, W-H. (2009). "The study on Korean-spoken English between Koreans and native English speakers", *Language Science Study*, 48, 201-217.
(윤원희 (2009). "한국인 영어 학습자 발음 평가에 대한 한국인 과 원어민의 차이", *언어과학연구*, 48, 201-217.)
- Jeong, H-S., Jang T-Y., Yoon, W-H., Yoon, I-S., & Sa, J-J. (2008). "The study on automatic evaluation on Korean English learners", *Korean Journal of Linguistics*, 42, 165-196.
(정현성, 장태엽, 윤원희, 윤일승, 사재진 (2008) "한국인 영어 학습자의 자동 측정 방법에 대한 연구", *언어학*, 42집, 165-196.)
- ACTFL-OPIC Korea (2010). *OPIC Evaluation criteria for English speaking*. Seoul: OPIC Korea:
- Anderson-Hsieh, J., & Koehler, K. (1988). "The effect of foreign accent and speaking rate on native speaker comprehension". *Language learning*, 38, 561-613.
- Anderson-Hsieh, J., & Koehler, K. (1998). "The effect of foreign accent and speaking rate on native speaker comprehension". *Language learning*, 38, 561-613.
- Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). "The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure", *Language learning*, 42, 529-555.
- Archibald, J. (1995). *Phonological acquisition and phonological theory*. Hillsdale, NJ: Lawrence Erlbaum Associate.
- Archibald, J. (1998). *Second language phonology*. Amsterdam: John Benjamins.
- Backman, N. E. (1979). "Intonation errors in second language pronunciation of eight Spanish speaking adults learning English", *Interlanguage Studies Bulletin*, 4-2, 239-266.
- Bradlow, A.R., Torretta, G.M., & Pisoni, D. B. (1996). "Intelligibility of normal speech: Global and fine-grained acoustic-phonetic talker characteristics". *Speech Communication*, 20, 255-272.
- Cenoz, J. (2000). "Pauses and hesitation phenomena in second language production". *ILT Review of Applied Linguistics*, 127-128.
- Conrad, B. K. (1991). *The relationship between empathy and pronunciation ability: A study of elementary-level college students of German*. Unpublished doctoral dissertation, University of Texas at Austin.
- Davis, S. M., & Kelly, M. H. (1997). "Knowledge of the English noun - verb stress difference by native and nonnative speakers". *Journal of Memory & Language*, 36, 445-460.
- Derwing, T. M., & Munro, M. J. (1997). "Accent, intelligibility, and comprehensibility". *Studies in Second Language Acquisition*, 19, 1-16.
- Flege, J.E. & Liu, S. (2001). "The effect of experience on adults' acquisition of a second language". *Studies in Second Language*

- Acquisition*, 23-4, 527-552.
- Guion, S. G., Flege, J. E., Liu, S. H., Yeni-Komshian, & Grace, H. (2000). "Age of learning effects on the duration of sentences produced in a second language". *Applied Psycholinguistics*, 21, 205-228.
- Guion, S. G., Flege, J.E., & Loftin, J.D. (2000). "The effect of L1 use on pronunciation in Quichua-Spanish bilinguals". *Journal of Phonetics*, 28, 27-42.
- Guion, S.G., Harada, T., & Clark, J. J. (2004). "Early and late Spanish-English bilinguals' acquisition of English word stress patterns". *Bilingualism: Language and Cognition*, 7, 207-228.
- Holmes, M. (1995). "A crosslinguistic comparison of the production of utterances in discourse", *Cognition*, 54, 169-207.
- James, E. (1976). "The acquisition of prosodic features of speech using a speech visualizer", *IRAL*, 14-3, 227-243.
- Jenkins, J. (2000). *The phonology of English as an international language*. Oxford: Oxford University Press.
- Kang, S., & Ahn, H. (2010). "The automatic measurement of prosody in English pronunciation test". *Korean Journal of Applied Linguistics*, 26(4), 121-150.
- Kang, S., Guion, S., Rhee, S., & Ahn, H. (2010). "The effect of language immersion on second language intonation". *Proceedings of Korean society of speech sciences*, 135-136.
- Ladefoged, P. (2003). *Phonetic data analysis*. Oxford: Blackwell Publishers.
- Lee C-H. (2010). "Improving inter-rater reliability in oral proficiency test at college level", *Modern English Language and Literature*, 54-1, 367-385.
- Lumley, T., & McNamara, T. (1995). "Rater characteristics and rater bias: Implication for training", *Language Testing*, 12, 54-71.
- Mennen, I. (2004). "Bi-directional interference in the intonation of Dutch speakers of Greek". *Journal of Phonetics*, 32, 543-563.
- Moyer, A. (1999). "Ultimate attainment in L2 phonology: The critical factors of age motivation and instruction", *Studies in Second Language Acquisition*, 21, 81-108.
- Munro, M. J., & Derwing, T. M. (1995). "Foreign accent, comprehensibility, and intelligibility in the speech of second language learners", *Language Learning*, 45-1, 73-97.
- Munro, M. J., & Derwing, T. M. (2001). "Modelling perceptions of the accentedness and comprehensibility of L2 speech: The role of speaking rate", *Studies in Second Language Acquisition*, 23, 451-468.
- Munro, M., & Derwing, T. (2008). "Segmental acquisition in adult ESL learners: A longitudinal study of vowel production", *Language Learning*, 58, 479-502.
- Schachter, S., Christenfeld, N., Ravina, B., Bilous, F. (1991). "Speech dysfluency and the structure of knowledge", *Journal of Personality and Social Psychology*, 60, 362-367.
- Scherer, K. (2000). "A cross-cultural investigation of emotional inferences from voice and speech: implications for speech technology". In Lewis & Haviland, *The Handbook of Emotion*. New York: Guilford.
- Tajima, K., Port, R., & Dalby, J. (1997). "Effects of speech timing on intelligibility of foreign accented English", *Journal of Phonetics*, 25, 1-24.
- Tahta, S., Wood, M., & Loewenthal, K. (1981). "Foreign accents: factors relating to transfer of accent from the first language to a second language", *Language and Speech*, 24-3, 265-272.
- Trofimovich, P. & Baker, W. (2006). "Learning second language suprasegmentals: effect of L2 experience on prosody and fluency characteristics of L2 speech", *Studies in Second Language Acquisition*, 28, 1-30.
- van Bezooijen, R. (1995). "Sociocultural aspects of pitch differences between Japanese and Dutch women", *Language and speech*, 38-3, 253-265.
- Willems, N.J. (1982). *English intonation from a Dutch point of view*. Dordrecht: Foris Publication.
- Zuengler, J. (1988). "Identity markers and L2 pronunciation", *Studies in Second Language Acquisition*, 10, 33-49.

• 강석한 (Kang, Seokhan), 제1저자

서울대학교 외국어교육연구소

서울시 관악구 관악로 599

010-9120-2433

kangs45@snu.ac.kr

관심분야: 음성학, 영어발음평가, 음성정보학

• 이석재 (Rhee, Seok-Chae), 교신저자

연세대학교 문과대학 영어영문학과

서울시 서대문구 연세로 50

02-2123-4483

scrhee@yonsei.ac.kr

관심분야: 음성학, 음운론, 영어발음평가, 음성코퍼스