

시, 도교육청 주관 표준화된 영어 듣기평가가 영어 교수·학습에 미치는 환류효과: 광주광역시 중, 고등학교를 중심으로*

정행

(호남대학교)

Jung, Haeng. 2006. The Washback Effects of Standardized English Listening Test Conducted by Municipal or Provincial Offices of Education: Focusing on the Secondary Schools in Gwangju Metropolitan City. *The Linguistic Association of Korea Journal*, 14(3), 177-194. This study aims to investigate washback effects of the nationwide standardized English listening test on English teaching and learning in Korean EFL secondary schools. The following research questions were raised for this study. 1) What were the test's purposes or expected effects on learning or teaching? 2) What perceptions and attitudes did the students and teachers have about the listening test? 3) What influences did the test have on students' English learning? 4) What influences did the tests have on teachers' English teaching? To answer the above questions, the data were collected through interviews and 2 different questionnaires (both the students' questionnaire and the teachers' questionnaire). Interviews were carried out to 12 test designers, developers and administrators. Questionnaires were conducted to a total of 234 students and a total of 87 teachers. Descriptive statistical analysis of the collected data indicated the following results. 1) The listening test was originally developed in order to increase the awareness of the importance of the students' English oral communicative abilities. 2) Students were much interested in the test (47.0%) and they perceived the test was beneficial to their development of listening abilities (35.9%). Teachers responded that the test positively influenced the students' listening abilities (4.04). 3) The listening test had an impact on the students' learning contents and led to the study on the colloquial expressions and the communicative functions. 4) The listening test had an impact on the teachers' teaching contents and techniques.

Key Words: Testing, English Language Testing, Listening Test, Language

* 이 논문은 2004년도 한국학술진흥재단의 지원에 의하여 연구되었음
(KRF-2004-002-A00156)

1. 서론

1980년대에 국제화, 세계화의 흐름과 함께 우리나라에서는 영어구사력을 갖춘 인재양성이 필요하게 되었고, 이에 따라 영어교육계에서는 영어교육의 목표를 주로 학생들의 영어 의사소통 능력을 키우는 데에 두었다. 그리하여 의사소통 중심의 영어교육이 도입되었고, 중·고등학교 교실현장에서 의사소통 중심의 영어교육이 실현될 수 있도록 하기위한 방안으로 교육부의 주도하에 교육목표와 함께 교육과정, 교재, 교육방법, 교육평가에서의 여러 가지 변화가 시도되었다.

이에 발맞추어 각 시, 도교육청에서는 지필고사 위주의 영어평가 방식에 의존해왔던 것에서 벗어나 관할 모든 중·고등학교에 학년별 영어 듣기평가를 주관하여 실시함으로써 평가를 통한 의사소통 중심의 영어교육으로의 변화를 주도하고자 하였다. 이러한 영어 듣기평가는 전국규모로 확대되어 갔으며 1985년부터는 국영방송인 한국방송공사(KBS)의 방송망을 통해 연 2회 실시되었고, 1991년부터는 교육방송(EBS)망을 통해 전국적으로 동시에 실시되었다. 또한 이것은 1998년부터는 연 4회로 횟수를 늘려 시행되기 시작했고, 서울을 제외한 15개 시, 도교육청이 공동으로 주관하는 전국적인 규모의 표준화된 평가로서 지금까지 시행되어오고 있다.

이러한 영어 듣기평가의 도입과 시행은 평가를 통해 교수, 학습에 큰 변화를 꾀하고자 하는 시도로서 이루어졌으며, 듣기평가에서 좋은 점수를 받기 위해서는 교실에서 교사와 학생은 평가에서 중시하는 언어기능과 유형에 맞추어 가르치고 배워야 하므로 궁극적으로 영어교수, 학습의 내용과 방법에 있어서 변화를 초래할 수밖에 없다는 가정에서 출발하였다(Alderson, 1986). 이것을 Alderson과 Wall(1993)은 평가의 환류효과(washback effect)라고 정의하였으며, 평가가 교수, 학습의 마지막 단계에서 학생들의 학습목표 도달 정도를 측정하기 위한 도구로서 역할을 할뿐만 아니라 이와 같이 오히려 평가가 교수, 학습에 영향을 주고 변화를 가져오는 효과를 준다고 주장하였다.

우리나라에서 중, 고등학생들을 대상으로 표준화된 듣기평가가 시행된지 약 20년 이상이 되었다. 그동안 듣기평가의 시행은 평가를 준비하는 학생이나 교사를 포함한 개인, 그리고 교육제도나 사회제도 전반에 걸쳐서 영

향을 미쳤다. 그러나 듣기평가 자체에 대한 평가 즉 평가의 구인요소나 내용, 유형이 학생들의 듣기능력을 측정하는 데에 타당한가(validity)에 대한 연구나 평가결과에 대한 신뢰도(reliability) 연구(Alderson et al., 1995; Heaton, 1988, Hughes, 1989; McNamara, 2001; Weir, 1993)는 거의 이루어지지 않고 있다. 또한 듣기평가가 학생들의 의사소통능력 향상을 이끌기 위한 효과를 기대하고 출발했다면 과연 평가가 그동안 교수, 학습에 어떠한 영향을 미쳤는가에 대한 구체적이고 실증적인 연구가 필요할 것이다(Andrews, 1994a, 1994b). 나아가 듣기평가가 원래 기대했던 의사소통 중심의 교수, 학습에 도움이 되는 환류효과(beneficial washback effect)를 주도록 하게 하기 위해서는 평가시행기관이나 평가개발자들은 듣기평가의 환류효과를 검토해서 평가의 목적에 장애가 되는 요인이나 문제점들을 찾아내고 평가를 개선해 나가야 할 것이다.

본 연구는 현재 15개 시, 도교육청이 주관하여 우리나라 중, 고등학교 학생들을 대상으로 연 4회 실시하는 표준화된 영어 듣기평가에 대한 평가로서, 듣기 평가가 영어 교수, 학습에 미치는 환류효과를 실증적으로 검토하고, 나아가 평가실시의 원래 목적이나 기대효과와 관련하여 영어 듣기 평가가 학생들의 영어 의사소통능력 향상에 도움이 되는 환류효과를 주기 위해서는 어떤 점들을 고려해야 할 것인가를 제안하는 데에 그 목적이 있다. 또한 본 연구는 연구 범위를 광주광역시 중등학교로 정하고 광주광역시 중, 고등학생들에게 시행되는 영어 듣기평가의 문제점과 실태를 파악함으로써 광주광역시 지역 중등학교들의 영어 교수, 학습의 질과 학생들의 영어 의사소통능력을 향상시키는 데 기여할 수 있도록 평가를 개선하는 데에 그 목적을 둘 것이다.

2. 이론적 배경

평가가 교수, 학습에 미치는 영향에 대한 연구는 평가에 대한 평가의 필요성이 강조되면서 여러 평가개발자와 연구자들에 의해서 시행되었다. Alderson(1986)은 언어 교육과정에서의 혁신은 언어평가에 있어서의 혁신에 의해서 이루어진다고 주장하였으며, Popham(1987)은 평가는 교사의 관심을 평가항목의 내용에 두게 함으로써 측정에 의해 이끌어진 교육(measurement-driven instruction)을 초래한다고 하였다. 또한 Morrow(1986)는 평가의 타당도는 그 평가가 교육에 좋은 영향을 미치는 정도에 의해서 측정되어야 한다고 보고 이를 환류타당도(washback

validity)라고 정의했다. Frederiksen과 Collins(1989)는 평가는 평가가 측정하고자하는 인지적인 기능을 조장하거나 향상시키는 방향으로 교육제도에 변화를 유도할 때 타당성을 가진다고 주장하고 이러한 타당도를 체제적타당도(systemic validity)라고 정의했다.

한편 Alderson과 wall(1993)은 평가가 긍정적으로, 혹은 부정적으로 교수. 학습에 막대한 영향을 줄 수 있으며, 평가가 좋은 영향을 주지 않는다고 Morrow(1986)나 Frederiksen과 Collins(1989)의 주장처럼 타당도가 부족하다고 말할 수는 없다고 주장하였다. 그들은 평가가 교육에 좋은 영향을 주지 못한 데에는 평가 자체의 요인뿐만 아니라 평가와 관련된 사회, 교육, 학교 등 다른 힘이 영향을 미칠 수 있다고 주장하고 있으며, 평가가 교수. 학습에 미치는 영향을 단순히 평가의 타당도와외의 관계에서만 설명하려는 것에 대해서 경고하였다. 또한 그들은 평가가 긍정적으로나 부정적으로 교수. 학습에 미치는 모든 영향 즉 Baker(1991)나 Bachman과 Palmer(1996)가 평가영향(test impact)이라고 불렀던 현상을 통틀어 환류효과(washback effect)라고 칭하였다. 나아가 그들은 환류효과는 한 가지가 아닌 복합적인 요인들과 관계되어지며, 정확히 교수. 학습에 어떤 영향을 주는가를 알아보기 위해서는 이론적이 아닌 교실 안으로 들어가서 자료를 수집하여 분석하는 실증적인 연구를 해야 한다고 주장하였다.

Alderson과 Wall(1993)은 환류효과에 대한 실증적인 연구방법을 도입하여 스리랑카 학생들의 영어 의사소통 능력을 측정하기 위해 혁신적으로 개발된 영어능력평가가 교수. 학습내용, 방법, 그리고 이들의 평가에 대한 태도에 미치는 연구를 실시하였으며 그 결과 평가가 교수. 학습내용에는 영향을 주지만 방법에는 영향을 주지 않았다고 보고했다. 또한 Alderson과 Hamp-Lyons(1996)는 교사와 학생들에 대한 설문지와 면담, 수업관찰을 통해 외국어로서의 영어능력 시험(TOEFL)인 토플 평가의 환류효과를 실증적으로 연구하고, 그 결과 평가가 교수. 학습 내용뿐만 아니라 방법에도 영향을 준다는 것을 보여주었다. 그러나 연구결과는 또한 교육방법은 평가에 의해서보다는 교사 개인적인 성격과 관련된 교수유형(teaching style)에 의해 훨씬 영향을 받는다는 것을 제시했다. Cheng(2000)은 평가가 교수. 학습에 미치는 영향은 평가결과가 개인들의 신상에 결정적인 영향을 미치는 중대한 시험(high-stakes test)일수록 크다는 것을 강조했다.

이와 같이 평가가 교수. 학습에 미치는 영향에 대한 이론적, 실증적인 연구들은 Alderson(1991)의 주장에서처럼 평가의 환류효과에 대한 연구가 평가의 타당도나 신뢰도 연구와 마찬가지로 평가를 평가하기 위한 연구의 한

분야로서 자리를 잡고 연구되어야 할 필요와 가치가 있음을 보여준다. 그러나 평가의 환류효과에 대한 지금까지의 연구들은 Viphavee(2001)가 지적한 바와 같이 개별적인 평가에 따른 사례연구(case study)와 유사하며, 평가의 환류효과에 대한 일반적인 결과를 도출해 내기에는 양적으로 아직 미흡하다. 따라서 지금 현재 시행되고 있는 여러 평가들에 대한 환류효과를 실증적으로 연구하여 평가가 교수, 학습에 미치는 영향을 구체적으로 검증하고 환류효과에 대한 일반적인 결과를 도출하는 데에 기여할 많은 제안이 이루어져야 할 것이다.

3. 연구 내용 및 방법

본 연구의 내용은 듣기평가 실시의 목적과 기대효과에 대한 검토, 듣기평가에 대한 교사, 학생들의 인식과 태도 분석, 듣기평가가 의사소통 중심의 교수, 학습내용과 방법에 미치는 영향 분석, 듣기평가의 원래 목적이나 기대 효과에 장애가 되는 요인을 분석하고 듣기평가가 교수, 학습에 도움이 되는 환류효과를 갖게 하기 위한 제안을 하는 것으로 이루어진다. 연구자료 수집을 위해서는 장학사나 영어듣기평가 관련자와의 면담 및 학생용 설문지와 교사용 설문지와 같은 도구를 개발하여 활용하였다. 각 각의 설문지에는 응답자에 대한 기본 정보, 듣기평가에 대한 인식 및 태도, 교수, 학습내용, 교수, 학습 방법과 관련된 문항이 포함되었다. 구체적인 연구 방법 및 절차는 다음과 같다.

3.1. 연구 질문

본 연구의 목적은 시, 도교육청 주관 표준화된 영어 듣기능력 평가가 영어교실에서 중, 고등학생들의 교수, 학습에 미치는 환류효과(washback effect)를 검토하는 데에 있다. 이를 위해 본 연구는 다음 네 가지 연구 질문들에 대한 답을 구한다.

질문 1: 듣기평가 실시의 목적과 기대효과는 무엇인가?

질문 2: 듣기평가에 대해 학생과 교사는 어떠한 인식과 태도를 가지고 있는가?

질문 3: 듣기평가가 학생들의 영어 학습내용과 방법에 어떠한 영향을 미치는가?

질문 4: 듣기평가가 교사들의 영어 수업내용과 방법에 어떠한 영향을 미치는가?

3.2. 연구 대상

본 연구는 연구의 대상을 광주광역시 중, 고등학교로 한정하고 면담은 광주시 교육청 영어과 장학사 및 듣기평가 개발과 시행에 참여한 교사들 12명을 대상으로 실시하였으며, 설문은 학생용 설문과 교사용 설문을 따로 개발하여 실시하였다. 학생용 설문은 시내 중학교 4개교와 고등학교 2개교를 임의로 선택하여 한 학교에 한 학년의 특정 반을 대상으로 실시하였으며, 중학생 158명 고등학생 76명 총 234명이 참여했다. 교사용 설문은 중학교 영어교사 56명과 고등학교 영어교사 31명 총 87명을 대상으로 실시하였다.

3.3. 연구 도구

3.3.1. 면담

면담은 연구질문 1에 대한 답을 구하기 위해 이루어졌으며 평가관리자 및 개발자를 대상으로 실시하였다. 면담은 주로 평가의 목적과 기대효과에 관해 질문하였지만 교사용 설문을 보충하기 위해 듣기평가의 긍정적인 영향, 평가결과에 대한 피드백 유형, 출제내용이나 유형을 수업에 반영하는 방법, 현 듣기평가의 개선점 등에 대해서도 질문하였다.

3.3.2. 학생용 설문지

학생용 설문지에는 ①응답자에 대한 기본 정보 (성별, 학년, 영어듣기능력 수준), ②듣기평가에 대한 인식 및 태도 (평가가 듣기 능력에 도움이 되는가, 평가가 전반적인 영어 구술능력에 도움이 되는가, 듣기평가가 동기유발에 미치는 정도, 듣기평가 문제에 대한 관심도, 듣기평가에 대한 만족도) ③듣기평가와 학습내용 (듣기 학습내용, 듣기 학습주제, 듣기 학습을 위한 의사소통 기능, 듣기를 위한 발음 학습) ④듣기평가와 학습방법 및 전략 (교수법/학습법, 수업/학습매체, 수업/학습활동, 수업/학습전략) ⑤기타 (듣기 학습장소, 듣기 학습교재)의 범주에 관해 문항들을 작성하여 포함시키

고, 각 문항들에 대해 응답자는 아니다(1), 잘모름(2), 그렇다(3)와 같이 세 가지로 응답하도록 하였다. 또한 설문지 문항의 타당도를 높이기 위해 사전 시험(pilot test)을 실시하여 문항이 조사 범주에 부적절하거나 뜻이 애매한 것은 제외시켰다.

3.3.3. 교사용 설문지

교사용 설문지에는 ①응답자에 대한 기본 정보 (교육대상 학년, 교육 경력), ②듣기평가에 대한 인식 및 태도 (듣기평가와 의사소통능력과의 관계, 듣기평가와 듣기능력과의 관계, 듣기평가가 듣기학습에 주는 긍정적인 영향, 듣기평가가 듣기학습에 대한 동기유발 정도, 듣기평가가 교수·학습 내용에 미치는 영향 정도, 듣기평가가 교수·학습 유형 및 전략에 미치는 정도, 듣기평가 내용 타당도 정도, 듣기평가 유형 타당도 정도, 듣기평가 신뢰도, 듣기평가의 지속에 대한 바람 정도) ③듣기평가와 수업내용 (듣기수업을 위한 어휘나 표현, 듣기 수업의 대화나 지문 주제, 의사소통기능, 대화나 지문의 문장구조, 발음이나 음운규칙) ④듣기평가와 학습방법 (수업활동 유형, 수업자료, 수업매체, 듣기전략, 평가유형에 대한 친숙도) ⑤평가 준비를 위한 수업활동 (듣고 빈칸 채우기, 듣고 답 고르기, 듣고 답 말하기, 듣고 그림 연결하기, 듣고 표 완성하기, 듣고 지시사항 따르기, 듣고 따라 말하기) ⑥듣기평가의 긍정적 효과를 저해하는 요인의 범주에 관한 문항들을 주로 Likert의 5점 척도를 사용하여 응답하도록 하였고, ⑦기타 개방형 질문(평가의 목적과 기대효과, 듣기평가의 긍정적인 영향, 평가결과에 대한 피드백 유형, 출제내용이나 유형을 수업에 반영하는 방법, 현 듣기평가의 개선점 등의 타당성, 신뢰성, 인식 및 태도, 동기유발 정도, 평가결과 성적 반영도)을 포함시켰다.

3.4. 연구 자료 수집 및 분석

본 연구는 신뢰할 만한 자료 수집을 위해 Brown(2001)이 제시한 삼각형 측량 (triangulation) 방법을 받아들여 연구 대상을 모두 세 가지로 하였다. 즉 연구 대상은 한 집단의 의견만 반영되지 않도록 평가관리자 및 개발자, 교사, 학생의 세 집단을 활용하였다.

면담을 통해 수집한 자료들은 주로 개방형 질문에 대한 답이기 때문에 설문지와는 달리 정량적인 분석(quantitative analysis)이 아닌 정성적인 분

석(qualitative analysis)을 하였다. 자료 분석에 대한 신뢰를 높이기 위해서는 TESOL(2000b)에서 제시한 정성적인 연구 지침서 (TESOL Quarterly Qualitative Guidelines)에 따라 수집한 자료를 범주별로 정리하고 분석하였다. 설문지를 통해 수집한 자료들은 설문 문항들이 3점 척도나 Likert의 5점 척도를 활용하여 정도를 표시한 것들이므로 SPSS 통계 프로그램 (version 12.0)을 활용 기술적인 통계(Descriptive Statistics) 분석을 하였다. 각 문항에 대해 빈도수와 백분율, 평균, 표준편차를 내서 응답의 정도 차이를 분석하고, 각 범주들 간의 차이를 비교하였다.

4. 연구결과

4.1. 듣기평가의 목적과 기대효과

듣기평가는 원래 어떠한 목적을 가지고 실시되었으며 듣기평가 시행을 통해 어떤 결과를 기대하였는가를 알아내기 위해 광주광역시 교육청 영어과 장학사 및 듣기평가 개발과 시행에 참여한 교사들 12명을 대상으로 면담을 실시하였다. 면담결과를 정리해 보면 아래 표 1과 같이 요약된다.

표 1. 듣기평가의 목적과 기대효과에 관한 면담결과

면담결과	빈도수
영어 교수. 학습에 있어서 구두 의사소통 능력의 중요성에 대한 인식을 높임	5명
문법 훈련 위주의 영어 교수. 학습 활동에서 의사소통 연습 중심으로 교수. 학습활동의 변화를 꾀함.	3명
듣기 기능을 강화하기 위한 교수. 학습에 동기를 부여하고 증진.	2명
듣기평가의 내용, 유형, 형태를 통해 듣기 교수. 학습에 대한 방향 유도.	1명
듣기 교수. 학습을 위한 자료 개발	1명

표 1의 면담결과에서 보여주듯이 듣기평가는 영어교육 현장에서 읽기 쓰기와 같은 문자언어를 통한 소통 외에 말하기 듣기와 같은 음성언어를 통한 구두 의사소통 능력의 중요성을 인식시키기 위한 목적으로 실시되었으며, 교사와 학생들은 듣기평가에서 좋은 점수를 받기 위해 구두 의사소통 중심의 표현에 중점을 둘 것이고 궁극적으로 듣기평가는 영어 교수학습 내

용과 방법, 활동, 자료에 있어서 큰 변화를 가져오리라 기대하고 실시되었음을 알 수 있다.

4.2. 듣기평가에 대한 학생과 교사의 인식과 태도

현재 1년에 4회 실시되고 있는 듣기평가에 대해 학생과 교사는 어떠한 인식과 태도를 가지고 있는가를 알아보기 위해 설문조사를 실시한 결과 학생들은 표 2와 같이 응답하였고, 교사들은 표 3과 같이 응답하였다.

표 2. 듣기평가에 대한 학생들의 인식과 태도

설문내용	응답자 반응		
	부정적 %(응답자 수)	중립 %(응답자 수)	긍정적 %(응답자 수)
1. 듣기 기능 향상에 도움이 됨	18.4% (n=43)	45.7%(n=107)	35.9% (n=84)
2. 말하기 기능 향상에 도움이 됨	35.9% (n=84)	48.7%(n=114)	15.4% (n=36)
3. 듣기평가에서 좋은 점수를 받기 위해 열심히 공부함	32.9% (n=77)	25.2% (n=59)	41.9% (n=98)
4. 듣기평가의 내용, 유형, 형식에 관심이 큼	16.7% (n=39)	36.3% (n=85)	47.0%(n=110)
5. 현재의 듣기평가에 만족함	34.6% (n=81)	29.5% (n=69)	35.9% (n=84)

(N=234, 1=아니다, 2=잘 모르겠다. 3=그렇다)

위의 표 2는 학생용 설문결과를 정리한 것이다. 설문에 대해 응답한 총 학생 수는 중학교 4개교 158명과 고등학교 2개교 총 234명이며, 이들은 각 항목에 대해 숫자로 (1=아니다, 2=잘 모르겠다. 3=그렇다) 부정적, 중립, 긍정적임을 표시하였다.

표 2에서 학생들은 듣기평가의 내용, 유형, 형식에 대해 관심이 크다고 답하였고(47.0%), 듣기평가에서 좋은 점수를 받기 위해서 열심히 공부한다는 데에 동의(41.9%)하였다. 또한 듣기평가가 듣기 기능 향상에 도움이 된

다는 데에는 긍정적으로 답(35.9%)하였지만 말하기 기능 향상에 도움이 된다는 데에는 부정적인 견해(15.4%)를 보였다. 현재의 듣기평가에 대한 만족도에 있어서는 긍정적으로 보는 학생의 수가 약간 더 많지만 긍정적인 입장과 부정적인 입장이 거의 비슷한 숫자(35.9%, 34.6%)로서 평가에 대한 학생들의 만족도를 높이기 위해서는 현재의 듣기평가가 상당히 개선되어야 함을 보여주고 있다.

한편 평가에 대한 교사들의 인식과 태도를 조사하기 위한 교사용 설문결과는 표 3과 같다. 설문에 응답한 교사들은 중학교 영어교사 56명과 고등학교 영어교사 31명 총 87명으로서 이들은 각 항목에 대해 동의하는 정도를 5점척도 (1=매우 동의하지 않음, 5= 매우 동의함)를 사용해서 숫자로 표시하였다.

표 3. 듣기평가에 대한 교사들의 인식과 태도

설문내용	응답자 반응	
	평균	표준편차
1. 학생들의 의사소통 능력 향상에 도움이 됨	3.13	.38
2. 학생들의 듣기 기능 향상에 도움이 됨	4.25	.24
3. 학생들의 듣기 학습에 긍정적으로 영향을 줌.	4.04	.32
4. 듣기에 대한 학생들의 관심과 필요성을 꾀함.	4.32	.15
5. 평가항목의 내용이 타당함.	3.85	.26
6. 평가항목의 유형이 타당함.	3.07	.45
7. 평가결과는 신뢰할 만 함.	3.26	.36
8. 현재의 평가가 지속되어야 함.	3.58	.34

(N=87, 1=매우 동의하지 않음, 5= 매우 동의함)

표 3에 따르면 교사들은 듣기평가가 듣기에 대한 학생들의 관심과 필요성을 갖게 하는 데에 가장 큰 도움이 된다고 보고 있으며(4.25), 학생들의 의사소통 능력 향상에는 크게 도움이 되지 않으나(3.13) 듣기기능 향상에 도움이 되고(4.25), 듣기 학습에도 긍정적으로 영향을 준다(4.04)고 보고 있다. 평가의 타당도에 대해서는 타당하다고 보는 편(3.85, 3.07)이어서 타당도를 더 높일 필요가 있지만 내용이 유형보다 더 타당하다고 여기고 있어

평가의 유형에 있어서 보 다 큰 개선이 필요함을 보여 주었다. 평가결과 점수에 대한 신뢰도에 있어서도 타당도와 비슷한 수준으로 신뢰할 만한 정도(3.26)에 그치고 있어 신뢰도를 높일 수 있는 방안이 필요하다고 본다. 현재의 평가가 지속되어야 하는가에 대해서는 긍정적인 편(3.58)이나 평가에 대한 만족도를 높이기 위해서는 평가에 대한 개선이 따라야 함을 알 수 있다.

표 2와 표 3이 각 각 보여주는 학생 설문결과와 교사 설문결과를 비교해 보면 학생과 교사는 모두 듣기평가에 대해 비슷한 인식과 태도를 가지고 있음을 알 수 있다. 학생과 교사 모두 듣기평가가 듣기에 대한 관심을 불러일으키고 듣기 기능의 향상과 듣기학습에 매우 긍정적인 영향을 준다고 보고 있어 이는 현재 듣기평가가 원래 시행 목적과 기대효과에 부합됨을 알 수 있다. 그러나 평가의 타당도나 신뢰도, 만족도에 있어서는 학생과 교사 모두 긍정적이긴 하지만 숫자가 높지 않아 이를 높이기 위해서는 현재의 평가를 개선해야 함을 보여주었다.

4.3. 듣기평가가 학생들의 영어 학습내용과 방법에 미치는 영향

듣기평가가 학생들의 영어 학습내용과 방법에 미치는 영향을 조사하기 위한 설문은 크게 듣기 학습내용, 듣기 학습방법 및 전략, 듣기 학습장소, 듣기 학습자료의 네 범주로 구분하여 이루어 졌다.

표 4. 학생들의 듣기 학습내용

설문내용	응답자 반응		
	부정적 %(응답자 수)	중립 %(응답자 수)	긍정적 %(응답자 수)
1. 회화체의 어휘나 표현 중심	23.5% (n=55)	29.9% (n=70)	46.6%(n=109)
2. 일상생활과 관련된 주제 중심	31.2% (n=73)	33.8% (n=79)	35.0% (n=82)
3. 실생활 관련 의사소통 기능 중심	24.5% (n=57)	40.8% (n=95)	34.8% (n=81)
4. 문장구조나 문법 위주	48.3%(n=113)	31.2% (n=73)	20.5% (n=48)
5. 발음상의 특징, 억양, 강세 중심	28.2% (n=66)	34.2% (n=80)	37.6% (n=88)

(N=234, 1=아니다, 2=잘 모르겠다. 3=그렇다)

표 4는 듣기 학습내용에 관한 설문 결과이다. 표 4에서 학생들은 듣기를 위해 회화체의 어휘나 표현을 중심으로 학습한다는 데에 매우 긍정적으로 답(46.6%) 하였고, 발음학습을 많이 하고(37.6%), 일상생활과 관련된 주제(35%), 실생활과 관련된 의사소통 기능 중심의 학습을 위주(34.8%)로 함에는 동의하는 편이었지만 듣기를 위해 문장구조나 문법위주의 학습을 한다는 데에 대해서는 매우 부정적으로(48.3%) 답하였다. 이는 듣기평가가 학습내용에 영향을 미쳤으며, 문장구조나 문법위주의 학습에서 회화체의 어휘나 표현, 의사소통 기능 중심의 학습으로 내용이 바뀌었음을 보여준다.

표 5. 학생들의 듣기 학습방법 및 전략

설문내용	응답자 반응		
	부정적 %(응답자 수)	부정적 %(응답자 수)	부정적 %(응답자 수)
1. 대화 참여자들의 관계	14.1% (n=33)	29.1% (n=68)	56.8%(n=133)
2. 대화의 주제 파악	13.2% (n=31)	24.4% (n=57)	62.4%(n=146)
3. 대화의 장소와 상황	7.3% (n=17)	25.2% (n=59)	67.5%(n=158)
4. 말하는 사람의 의도 파악	11.5% (n=27)	25.6% (n=60)	62.8%(n=147)
5. 모르는 단어나 표현은 맥락으로 짐작	8.6% (n=20)	18.5% (n=43)	73.0%(n=170)

(N=234, 1=아니다, 2=잘 모르겠다. 3=그렇다)

표 5는 학생들의 듣기 학습방법 및 전략에 관한 설문결과를 보여준다. 학생들은 영어 듣기를 하는 과정에서 모르는 단어나 표현을 맥락을 통해 짐작하고(73.0%), 대화의 장소나 상황을 통해 세부사항을 이해하려고 하며(67.5%), 말하는 사람의 의도를 파악하려고 하고(62.8%), 대화의 주제를 생각하며 듣고(62.4%), 대화에 참여한 사람들의 관계를 파악하며 듣는(56.8%) 듣기전략을 주로 활용한다고 답하였다. 이는 듣기평가가 학생들의 듣기 학습전략에 영향을 주었음을 알 수 있고 과거의 부분에서 전체로 가는 상향식 전략 위주에서 전체를 통해 부분을 이해하려는 하향식 전략위주로 바뀌었음을 보여준다.

듣기 학습장소에 대한 설문조사결과에서 학생들은 학교(35.2%), 학원(28.4%), 집(22.5%), 기타(13.9%) 순으로 답하였다. 학생들의 듣기 학습은 학원이나 집보다는 학교에서 가장 많이 이루어지고 있으며 주로 아침자습이나 점심시간 자율학습시간에 학교에서 방송수업을 통해 다 함께 학습한다고 답하였다. 이것은 듣기학습이 공교육에서 자리 잡아 감을 보여주는

것으로 해석되어 진다.

마지막으로 듣기 학습자료에 관한 설문조사에서 학생들은 상업용 듣기교재(36.4%), 학교 교과서(30.5%), 팟송(0.8%), 드라마(0.2%), 영화(0.1%) 순으로 답하였다. 조사결과에 따르면 학생들은 학습자료로서 듣기평가 대비용 모의평가가 들어있는 상업용 듣기교재를 가장 많이 활용하고 있으며, 다음으로는 교과서의 듣기내용을 활용하고 있었다. 이는 학생들이 듣기평가와 직접 관련된 자료만을 선호하고 있음을 보여주고 있으며, 실제 언어생활에서 듣기와 관련이 있는 팟송, 드라마, 영화와 같은 실제자료(authentic material)는 거의 활용하지 않고 있음을 보여주고 있다. 따라서 앞으로의 듣기평가에서는 교과서의 내용 외에 실제자료가 보다 더 많이 활용될 수 있도록 개선되어야 할 것이고 이를 통해 학생들이 보다 더 실제 자료를 듣기 학습에 활용하도록 유도해야 할 것이다.

4.4. 듣기평가가 교사들의 영어 수업내용과 방법에 미치는 영향

듣기평가가 교사들의 영어 학습내용과 방법에 미치는 영향을 조사하기 위한 설문은 크게 듣기 수업내용, 듣기 수업기법, 듣기평가 준비를 위한 수업활동, 듣기평가의 긍정적인 영향에 방해가 되는 요인의 네 범주로 구분하여 이루어 졌다.

표 6. 교사들의 듣기 수업내용

설문내용	응답자 반응	
	평균	표준편차
1. 평가에서 활용된 어휘나 표현	3.82	.25
2. 평가에서 활용된 대화나 지문의 주제	3.54	.37
3. 평가에서 활용된 의사소통 기능	3.63	.43
4. 평가에서 활용된 대화나 지문의 문장구조	1.85	.21
5. 평가에서 활용된 발음, 음운규칙	2.47	.13

(N=87, 1=매우 동의하지 않음, 5= 매우 동의함)

위의 표 6은 영어 듣기평가가 교사들의 수업내용에 미치는 영향에 대한 설문조사 결과를 보여준다. 표 6에 따르면 교사들은 평가에서 활용된 어휘나 표현(3.82), 의사소통 기능(3.63), 대화나 지문의 주제(3.54)가 수업내용에

상당히 영향을 준 것으로 답하고 있다. 반면 발음이나 음운규칙(2.47), 문장 구조(1.85)에 대해서는 큰 영향을 주지 않는다고 답하고 있다.

표 7. 교사들의 듣기 수업기법

설문내용	응답자 반응	
	평균	표준편차
1. 평가가 수업활동 유형에 영향	3.74	.26
2. 평가가 수업자료 활용에 영향	3.92	.21
3. 평가가 수업매체 활용에 영향 (녹음기, 비디오, CD, 멀티미디어)	3.48	.45
4. 평가가 듣기전략 사용에 영향	4.13	.28
5. 평가가 평가유형에 친숙해지는 수업에 영향	4.58	.24

(N=87, 1=매우 동의하지 않음, 5= 매우 동의함)

영어 듣기평가가 교사들의 수업기법에 미치는 영향을 설문조사한 결과는 표 7과 같다. 교사들은 듣기평가가 평가유형에 친숙해지도록 하는 데에 가장 큰 영향(4.58)을 미치고, 이어서 듣기전략 사용(4.13), 수업자료 활용(3.92), 수업활동 유형(3.74), 수업매체 활용(3.48) 순으로 상당히 영향을 준다고 응답하였다.

듣기평가 준비를 위한 수업활동으로 교사들이 교실에서 선호하는 수업활동에 대한 조사에서 교사들은 듣고 정답 고르기(72.3%), 듣고 빈칸 채우기(42.3%), 듣고 따라 말하기(32.4%), 듣고 답 말하기(8.5%), 듣고 그림 연결하기(5.8%), 듣고 표 완성하기(4.5%), 듣고 행동하기(3.8%) 순으로 응답하였다. 이러한 결과는 평가에서 가장 많이 활용되는 듣고 정답 고르기 유형이 교사들의 수업에 큰 영향을 주고 있음을 보여주는 것으로 해석되며, 평가의 유형이 듣고 정답 고르기 유형에 만 편중되지 말고 보다 다양해져서 교사들이 다양한 수업활동을 쓰도록 유도하여야 함을 보여준다.

마지막으로 듣기평가의 긍정적인 영향에 방해가 되는 요인에 대한 조사에서 교사들은 정해진 유형의 평가항목(38.5%), 평가에 대한 교사들의 관심부족(27.3%), 평가에 대한 학생들의 관심부족(22.7%), 평가결과가 성적에 반영되는 비중이 낮음(15.4%), 평가의 신뢰도 부족(9.6%), 평가의 타당도 낮음(7.5%)을 이유로 들었다.

5. 결론 및 제언

지금까지 본 연구에서는 시, 도교육청 주관 표준화된 영어 듣기평가가 영어 교수·학습에 미치는 환류효과를 조사 분석하였다. 장학사 및 듣기평가 개발과 시행에 참여한 교사들에 대한 면담, 학생과 교사를 대상으로 실시한 설문조사결과 듣기평가는 영어교수·학습의 내용과 방법에 모두 큰 영향을 주었음을 알 수 있었다.

먼저 듣기평가가 학생들의 영어 듣기 학습내용과 방법에 미치는 영향을 분석한 결과 다음과 같은 결론을 얻어냈다. 듣기평가는 학생들의 학습내용에 있어서 문장구조나 문법위주의 학습으로부터 회화체의 어휘나 표현, 발음, 의사소통 기능 중심의 학습으로 변화를 가져다주었다. 또한 학습방법에 있어서는 듣기평가가 학생들의 듣기 학습전략에 큰 영향을 주었음을 알 수 있었다. 학생들은 과거 평가시행 이전이나 초기에는 단어를 먼저 이해하고 단어의 결합을 통해 점 점 큰 구나 문장의미를 이해하려는 부분에서 전체로 가는 상향식 전략을 주로 사용하여왔으나 평가시행 20년이 지난 지금에는 세부 단어를 모르더라도 전체 상황이나 맥락을 통해 주제를 이해하고 모르는 단어의 의미를 짐작해 가는 전체를 통해 부분을 이해하려는 하향식 전략을 주로 활용하고 있음을 보여주었다.

다음 듣기평가가 교사들의 영어듣기 수업내용과 방법에 미치는 영향을 분석한 결과 얻어낸 결론은 다음과 같다. 수업내용에 있어서 교사들은 평가에서 활용된 어휘나 표현, 의사소통 기능, 대화나 지문의 주제를 포함하여 다루고 있었다. 이것은 교사들도 학생들과 마찬가지로 평가를 의식하고 수업내용에 있어서 문법위주에서 벗어나 회화체의 표현이나 의사소통 기능 중심으로 변화를 시도해왔음을 보여준 것으로 해석되어진다. 수업기법에 있어서 교사들은 학생들을 평가의 유형에 익숙해지게 하기 위해 노력하고 있었고, 평가에서 가장 많이 활용된 기법인 듣고 정답 고르기를 주요 수업 활동으로 활용하였다.

이와 같이 영어 듣기평가는 1985년 전국규모로 시행된 이래로 거의 20년 동안 꾸준히 시행되어 오면서 우리나라의 중등 영어교수·학습의 내용과 방법에 큰 영향을 미쳤다. 또한 평가의 목적이나 기대효과에서 예상했던 바와 같이 교사와 학생 모두에게 구두의사소통능력의 중요성에 대한 인식을 높였고 큰 변화를 가져다주었다.

그러나 현행 표준화된 영어 듣기평가가 학생들의 구두의사소통능력을 키우는 데에 기여할 수 있는 보다 큰 유익한 환류효과를 가져다 주기위해서

는 다음과 같은 몇 가지 점에서 평가가 개선되어야 할 필요가 있다고 본다.

첫째, 평가의 유형에 있어서 다양해질 필요가 있다. 지금까지 주로 평가에서 활용된 유형은 듣고 내용을 이해했는가를 확인하기 위해 듣고 내용에 일치하는 답을 고르기였다. 따라서 교사와 학생 모두 듣기 기능을 위한 교수. 학습 방법으로 듣고 고르기에 너무 의존하는 경향이 있다. 그러나 실생활에서의 듣기 기능은 듣고 고르기 외에 듣고 응답하기, 듣고 행동하기, 듣고 양식에 기입해 넣기 등 여러 가지 다양한 형태로 반응을 보여야 하므로 평가가 실생활에서의 듣기 기능을 강화하기 위한 학습으로 변화를 가져다 주도록 하기 위해서는 듣기평가의 유형을 다양화해야 할 필요가 있다.

둘째, 듣기평가가 학생들의 구술능력 및 전반적인 의사소통 능력을 키우는 데에 기여하도록 하기 위해서는 현재와 같이 단순히 듣기능력만을 평가하는 분리형 평가(discrete test) 보다는 듣고 말하기, 듣고 요지 쓰기 등과 같이 듣기 기능을 다른 언어 기능과 통합하여 평가하는 통합형 평가(integrative test)로 발전시켜야 한다고 본다. 그리하여 듣기평가가 듣기 기능에 관한 평가일 뿐만 아니라 전달되는 의미를 이해하고 그에 대한 반응으로 의미를 전달하는 의사소통 중심의 평가(communicative testing)가 되도록 개선되어야 할 것이다.

셋째, 듣기평가에서 활용된 듣기자료가 보다 실생활에서 접할 수 있는 실용자료(authentic materials)로 확대되어져야 한다. 현재 듣기평가에서 활용되는 언어자료는 주로 두 사람 혹은 몇 사람 간의 대화를 듣고 답하도록 되어있어 대화가 주 자료가 되고 있지만 실제 언어생활에서의 듣기자료는 대화뿐만 아니라 안내 방송, 상품광고, 뉴스, 드라마, 영화에 이르기 까지 다양하다. 따라서 듣기평가가 단순히 교실에서의 평가라는 한계를 벗어나 학생들의 실생활에서의 듣기능력을 평가하는 평가로서 타당도가 높은 평가로 발전되기 위해서는 다양한 실생활에서 접할 수 있는 자료들이 평가자료로 활용되어야 할 것이다.

이상과 같은 몇 가지 점을 수정 보완하여 현행 듣기평가가 개선되어진다면 듣기평가에 대한 타당도와 신뢰도가 높아질 것이고, 듣기평가는 중등 영어 교수 학습에 더 큰 유익한 환류효과를 가져다 줄 것으로 기대되어진다.

참고문헌

- Alderson, J. C. (1986). Innovations in language testing. In M. Portal (ed.), *Innovations in testing: Proceedings of the IUS/NFER conference* (pp. 93-105). London: NFER/Nelson.
- Alderson, J. C. (1991). *Language testing in the 1990s: How far have we come? How much further have we to go?* Retrieved June 2, 2003, from ERIC EDRS, ED 365145.
- Alderson, J. C., Clapham, C., & Wall, D. (1995). *Language testing construction and evaluation*. Cambridge: Cambridge University Press.
- Alderson, J. C., & Hamp-Lyons, L. (1996). TOEFL preparation courses: A study of washback. *Language Testing*, 13(3), 280-297.
- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14(2), 115-129.
- Andrews, S. (1994a). The washback effect of examinations: Its impact upon curriculum innovation in English language teaching. *Curriculum Forum*, 4(1), 44-58.
- Andrews, S. (1994b). Washback or washout? The relationship between examination reform and curriculum innovation. In D. Nunan, V. Berry, & R. Berry (Eds.), *Bringing about change in language education* (pp. 67-81). Hong Kong: Dept of Curriculum Studies, University of Hong Kong.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Baker, E. (1991). *Alternative assessment and national policy*. Paper presented at the National Research Symposium on limited English Proficient students' issues: Focus on evaluation and measurement, Washington, DC.
- Brown, J. D. (2001). *Using surveys in language programs*. Cambridge: Cambridge University Press.
- Cheng, L. (2000). *Washback or backwash: A review of the impact of testing on teaching and learning*. Retrieved June 2, 2003, from ERIC EDRS, ED 442280.

- Frederiksen, J. R., & Collins, A. (1989). *A system approach to educational testing*. *Educational Researcher*, 18(9), 27-32.
- Heaton, J. B. (1988). *Writing English language tests*. London: Longman.
- Hughes, A. (1989). *Testing for language teachers*. Cambridge: Cambridge University Press.
- McNamara, T. (2001). *Language testing*. Oxford: Oxford University Press.
- Morrow, K. (1986). The evaluation of tests of communicative performance. *Innovations in language testing: Proceedings of the IUS/NFER conference* (pp. 1-13). London: NFER/Nelson.
- Popham, W. J. (1987). The merits of measurement-driven instruction. *Phi Delta Kappa*, 68, 679-682.
- TESOL. (2000b). TESOL Quarterly qualitative guidelines. *TESOL Quarterly* 31(2), 387-388.
- Viphavee, V. (2001). An interview with J. Charles Alderson. *Issues in Applied Linguistics*, 12(1), 91-109.
- Weir, C. J. (1993). *Understanding and developing language tests*. New York: Prentice Hall.

Haeng Jung

Department of English Language and Literature

Honam University

59-1, Seobong-dong, Gwangsan-gu

Gwangju, 506-704

Tel: 82-62-940-5523

Email: hjung@honam.ac.kr

Received: 30 Jul, 2006

Accepted: 5 Sep, 2006