

구어 자료 활용이 듣기 평가 문항 개발에 미치는 영향 연구

김 지 혜

Abstract

Kim Jihye. 2014. 9. 30. **Research on the Effect of Utilizing Audio Sources in Developing Listening Test Items.** *Bilingual Research* 56, 109-130. This study investigates the effect of audio sources that Korean teachers utilize when developing listening test items.

Before developing listening test items, teachers consider learners' competency and learning targets to set the goal of the test. And after confirming if topics and contents are suitable for the test, teachers collect materials and try to make a meaningful listening test utilizing audio sources as much as possible. However, issues arise at this point: Most of the contents teachers use for a listening test are from written media.

Teachers then change those written contents into spoken ones, such as dialogues or monologues. Listening test items derived from these contents are in fact not different from reading test items, and it causes reading questions appearing in a listening test.

Assuming this kind of test item production somewhat restricts proper evaluation for learners' listening ability, the researcher of this paper looked into how materials used in making listening test affected the test items. In her experiment, well-experienced Korean teachers were divided into two groups: One group was given materials from audio resources, and the other group was given materials from written one. The two groups were then asked to produce listening test items.

In the result, the latter group was turned out to have developed better test items which are more closed to the nature of listening. **(Korea University)**

【Key words】 listening test(듣기 평가), listening comprehension(듣기 능력),

listening text(듣기 텍스트), audio sources(구어 자료),
authentic materials(실자료)

1. 서론

한국어 듣기 평가에서 학습자의 듣기 능력을 타당하게 평가하기 위해서는 평가의 방법과 내용이 의사소통적이어야 하며, 실제 한국어 담화 상황에서 발생할 가능성이 높은 실제적 자료가 활용되어야 한다. 이는 강명순 외(1999)에서 한국어 듣기 평가에 대해 구체적인 논의를 시작한 이래 듣기 교육이나 평가에서 지속적으로 강조되어 온 것이다.

그러나 실제 듣기 평가의 내용을 살펴보면, 출제자가 실자료를 활용한 다면서 구어 자료가 아닌 문어 자료를 가지고 그것을 대화나 독백 형식의 구어체로 바꾸어 듣기 지문을 만들고 그것에 기반해 듣기 문제를 출제하여 읽기 평가에서와 다르지 않은 유형의 문제들이 출제되는 것을 볼 수 있다.¹⁾ 또한 듣기 평가가 이해 영역이라는 점에서 사실 정보의 정확한 이해를 묻는 것이 가장 기본적이고 핵심이 되는 활동이기는 하나 텍스트의 장르나 유형에 상관없이 사실 정보를 묻거나 세부 내용을 파악하는 데에 문항의 유형이 집중되는 것은 재고의 여지가 있다.²⁾

본 연구는 듣기 평가에서 발생하는 이러한 문제가 출제자인 교사가 듣기 문항 개발을 위해 활용하는 최초의 자료로부터 시작되는 것은 아닌지 점검해 보려고 한다. 다시 말해, 문어 자료를 활용해서 듣기 문항을 출제

-
- 1) 임철성(2009:287)은 수능의 듣기 평가도 문어 중심 평가라고 지적하고 있다. 현행 듣기 평가에서는 낭독되는 듣기 지문을 사용하고 있는데, 이 듣기 지문이 문어를 구어로 변환한 수준이어서 낭독 대본을 글 텍스트로 주고 풀게 해도 아무런 문제가 없는 문항들이라고 설명한다.
 - 2) 이준호(2010:337)에서는 한국어능력시험 이해 영역에서 추론하기 기술이 추가되어야 함을 주장하면서 한국어능력시험에서는 사실적·사전적 의미와 정보를 다루는 세부 내용 파악하기 등의 평가 비중이 높다는 문제를 제기하였다.

하는 지금의 방식이 학습자들의 듣기 능력을 타당하게 평가하는 데에 한계가 있다고 보고, 출제자들이 출제를 위한 원시 자료로 활용한 자료의 형태가 문항 유형에 미치는 영향을 살펴보고자 하는 것이다.

2. 듣기 평가에서의 구어 자료 활용

의사소통 중심의 듣기 평가에서는 평가의 유형만큼이나 자료가 중요한 요인이다. 단순히 소리 인지 여부를 확인하는 경우가 아니라면 평가 자료가 실제적이지 못할 경우, 적절한 평가 항목을 추출하기 어렵고 평가의 목적마저 희미해질 수 있기 때문이다. 이에 박민신(2008:3)은 한국어 듣기 교육에서 듣기 텍스트 연구는 진정성(authenticity)이 핵심이 되고 있으며, 실세계 듣기에 기반한 진정성 있는 듣기 텍스트의 구성이 듣기 교육에서 우선되어야 할 것이라고 주장했다.

듣기 교육이나 평가에서 다루어지는 실자료는 대화, 뉴스, 드라마, 라디오, 영화, 다큐 등 구어 자료를 의미한다. 물론 실제 수업이나 평가에서 실자료를 전부 그대로 사용하는 경우는 드물며, 수업의 목적이나 시간에 따라 특정 부분을 편집해서 쓰고 있다. 평가의 경우에는 목적이나 시간뿐만 아니라 자료의 특성에 의해서도 활용이 제한되어, 교사들은 실자료를 참고는 하되 새롭게 듣기 지문을 작성해 그 지문을 기반으로 문제를 출제하는 것이 일반적이다.

강승혜 외(2006)에서는 평가 자료를 기준으로 듣기 평가의 유형을 분류하기도 했는데 담화의 단위, 유형, 주제 등을 기준으로 나누었다. 평가 유형을 분류하기 위해 수집한 자료들은 구어 자료에 한하며, 강승혜 외(2006:144~145)는 평가 자료가 실제성, 자연성, 그리고 명확성을 유지해야 하며 실제 생활에서의 다양한 담화 유형을 반영해야 한다고 주장한다. 그러나 실제 구어 자료를 그대로 듣기 평가의 지문으로 사용할 때, 여러 문제가 발생할 수 있으니 이를 적절히 가공하여 실제 생활에서 쓰

임직한 자료로 만들어야 한다고 설명하고 있다.

이와 같이 듣기 평가에서는 평가의 특성 상 실제 구어 자료를 그대로 사용하는 것이 아니라 그것을 기반으로 해 지문을 작성하고 평가 항목을 추출하는 방식으로 이루어진다.

그러나 현장에서 시행되는 듣기 평가에서 실자료를 활용할 때 발생하는 문제는 평가에 적합하게 쓰일 수 있는 듣기 지문을 얼마나 잘 구성하느냐는 것뿐만 아니라 실자료라는 개념 속에 구어 자료가 아닌 문어 자료도 포함되는 경우가 적지 않다는 것이다. 같은 실자료라고 해도 문어 자료는 듣기 평가에서 실자료에 포함시키기 어렵다. 반복해 읽으면서 혹은 특정 단락이나 문장에 집중해서 정보를 찾을 수 있는 문어 자료와 달리 구어 자료는 제공되는 즉시 소멸해 버리는 이야기 속에서 전략적으로 필요한 내용을 수집하고 의미를 재구성할 수 있어야 하기 때문이다.

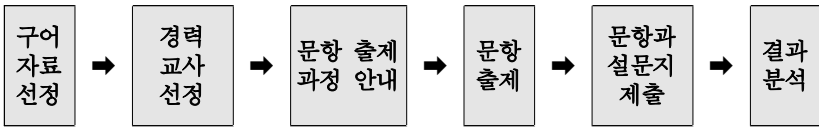
동일한 내용의 자료라고 할지라도 그것이 문어로 구현이 되었을 때와 구어로 구현이 되었을 때 이해하는 방식에는 차이가 발생하며, 이는 평가 문항에도 영향을 미치게 될 것이다. 사람들이 동일한 기사를 신문으로 접했을 때와 동영상 뉴스로 접했을 때, 정보 처리의 방식이나 반응이 상이한 것을 고려하면 충분히 예측이 가능한 부분이다.

본 연구는 듣기 평가에 사용되는 자료가 실자료에 기반해야 한다는 것에 충분한 동의가 이루어져 있는 현 시점에서 듣기 평가에서의 실자료는 문어 자료가 아닌 구어 자료이어야 한다는 점을 다시 한 번 환기시키고 그 중요성을 실험을 통해 밝혀 보고자 한다. 구어 자료로부터 평가 문항을 추출하고 듣기 지문을 작성하는 방식이 아닌 문어 자료를 찾고 그것을 듣기 지문으로 만들어 평가 문항을 추출하는 방식은 분명 학습자의 듣기 능력을 평가하는 문항에 영향을 미칠 것이기 때문이다.

3. 연구 방법

3.1. 연구 절차

연구는 <그림 1>과 같은 절차로 진행되었다. 먼저 평가 문항 개발에 사용될 구어 자료를 선정하였으며, 실험에 참여할 경력 교사 10인을 선정하였다. 다음으로 각 교사들에게 동의를 구하고, 듣기 평가 문항 출제의 과정을 개별적으로 안내하였다. 교사들에게는 연구의 구체적인 목적은 밝히지 않았으며, 한국어 교육 전문가들을 대상으로 듣기 평가 문항 개발의 양상을 살피고자 한다고 설명했다. 교사들은 2주 동안 문항을 출제하였고, 그 내용을 설문지와 함께 제출하였다. 마지막으로 듣기 평가 문항과 설문지의 내용을 분석해 경향을 파악하였다.



<그림 1> 연구 절차

3.2. 연구 자료

출제에 사용된 자료는 동영상 파일 3개이다. A그룹에는 동영상 형태 그대로, B그룹에는 동영상의 내용을 전사해서 문어 자료의 형태로 제시하였다.³⁾ 3개의 동영상 파일은 각각 뉴스, 토론, 드라마로 서로 다른 장르이며 ‘문화재’, ‘게임중독법’, ‘친구와의 갈등’이라는 내용을 담고 있다. 각 자료의 시간은 뉴스가 1분 40초, 토론은 8분 30초, 드라마는 2분 20초

3) 출제자가 듣기 평가를 위해 문항을 개발하기 전에 작성하는 듣기 지문의 원시 자료가 구어인지 문어인지에 따라 문항의 유형이 달라질 것이라는 가정이 본 연구의 출발이다. 따라서 본 연구에서 제시한 두 가지 형태의 자료가 그대로 학습자용 평가를 위한 지문으로 사용되는 것이 아니라 출제자들의 문항 유형 개발의 양상을 살피기 위한 원시 자료로 쓰였음을 밝혀 둔다.

였다. 토론과 드라마의 경우에는 실제 더 긴 시간 동안 진행되는 자료이나 평가에 필요한 부분만을 편집하여 제공하였으며, 드라마는 두 사람의 대화가 진행되는 부분이었다.

<표 1> 연구에 사용된 구어 자료

자료	01	02	03
장르	뉴스	토론	드라마
소요 시간	1분 40초	8분 30초	2분 20초
제목 및 내용	YTN 뉴스 “헐게미가 두려운 목재 문화재”	JTBC 직격토크 “게임중독법 논란” 中	KBS2 드라마특선 “연우의 여름” 中

뉴스와 토론의 경우에는 고급 학습자를 대상으로 한 평가에서 빈번하게 사용되는 텍스트 유형⁴⁾이며, 드라마는 대화 듣기가 가능하기에 선정하여 제시하였다. 듣기 평가에서 뉴스는 보통 사건 사고 뉴스를 많이 사용하는데, 사건 사고 뉴스의 경우 듣기 목적이나 기능이 너무나 명확하여 원시 자료의 형태에 따른 문항 출제의 양상을 살피기에는 적합하지 않을 것이라고 판단하여 제외하고, 정치나 경제 문제가 아닌 일반 사회 문제 중에서 선택하였다. 토론의 경우에는 대부분 비슷한 방식으로 진행되기 때문에 찬반을 분명하게 나눌 수 있거나 이미 익숙한 주제를 피하였고 실제 시험에서는 특정 부분이 편집되어 사용되겠지만, 문항 출제를 위해 실제 시험에 참여하는 교사들에게는 토론 내용을 파악할 수 있도록 주요 쟁점이 담긴 부분을 8분 정도로 편집하여 제시하였다. 마지막으로 드라마의 경우에는 한국 드라마에서 자주 볼 수 있는 갈등 상황을 선택했다. 친구와의 갈등이 최고조에 달해 자신의 입장을 피력하는 두 친구가 말다툼을 하는

4) 김지영 외(2014:91)을 보면, 한국어 숙달도 평가 듣기 문항의 세부 담화 유형은 일상생활 대화, 업무적 대화, 안내 방송, 광고 방송, 일기예보, 뉴스, 교양 정보 프로그램, 강연·강의, 인터뷰·대답, 토론, 식사, 기타 등으로 다양하다.

상황을 드라마에서 골라 편집했다.

구어 자료를 전사한 내용은 아래와 같으며, B그룹의 교사들에게는 연구자가 전사한 텍스트 전문을 제공하여 문어 자료를 읽은 후에 듣기 문항을 출제하도록 하였다.

앵커: 송례문 부실 공사를 계기로 문화재에 대한 관심이 높아지고 있는데요. 목재로 된 문화재가 많아서 흰개미에 의한 피해가 적지 않다고 합니다. 김항우 기자가 취재했습니다.

기자: 지은 지 200년이 넘는 향교입니다. 기둥과 처마, 마루 등 주요 부분이 목재로 만들어져 있습니다. 지은 지 오래 되다 보니 목재 곳곳이 갈라져 있고, 이 틈새로 드나드는 흰개미를 쉽게 볼 수 있습니다. 길이 1cm의 흰개미는 주로 죽은 나무를 갉아먹기 때문에 오래된 목재 건물에는 치명적일 수 있습니다. 게다가 요즘이 흰개미의 번식기라 활동이 가장 왕성합니다. 이 흰개미를 방지하기 위한 작업이 시작됐습니다. 기둥 틈새는 물론 마루와 처마 곳곳에 약을 뿌리고 건물이나 벽 틈새도 막아 줍니다. 흰개미 방제 대상은 경기도 내 160여개 목재 문화재로 이달 말까지 계속됩니다. 국내에는 목재로 된 건축 문화재가 많은 만큼 화재뿐만 아니라 흰개미 등 해충 방제에도 힘을 써야 할 것으로 보입니다.

YTN 뉴스 “흰개미가 두려운 목재 문화재”

사회자: 중독 예방 관리 치료를 위한 일명 중독법 추진을 놓고 온라인이 연일 들썩거리고 있죠. 술과 마약, 도박과 함께 게임을 4대 중독 물질로 규정하고 국가가 통합 관리한다 이것이 현재 법안의 골자인데요. 게임 업체는 우리가 마약 업자나 이렇게 또 반발을 하고 있습니다. 매우 격양된 반응을 보여서 오늘 하루 종일 인터넷은 좀 들썩거렸습니다. 자, 오늘 두 분과 이야기를 나눌 텐데요. 한 분은 직접 법안을 발의한 신의진 새누리당 의원, 그리고 또 한 분은 게임 업계 쪽 입장을 말씀하실 분인데 역시 이분과 함께도 얘기를 좀 나누도록 하겠습니다. 자, 안녕하십니까.

신의진 의원: 네, 안녕하세요.

사회자: 우선 신의진 의원과 얘기를 나누겠습니다. 법안 반대 서명한 네티즌들이 18만 명이 넘어섰다고 하니까, 알고 계시겠지만요. 의원실 홈페이지하고 전화도 마비되었다고 들었습니다.

신의진 의원: 사실 저는 반발하시는 분들의 얘기를 들으면 이 법안에 대한 이해가 부족하고 오해를 많이 하고 있다고 생각을 합니다. 이 법은 게임에 대한 규제가 있는 법이 아니라 중독에 대한 폐해를 예방하고 관리하는 치료를 해 주는 데에 국가가 의무를 갖는 기본법입니다. 그래서 간단히 말씀을 드리면 게임을 지금 하고 있는 사람들을 못 하게 하나, 아닙니다. 그리고 게임을 만드는 업체에 대해서 규제를 하나, 그거 절대 아니고 뭐 범칙금도 없습니다. 그럼 뭐냐, 지금 게임에 중독이 돼서 매일 매일 게임 하느라고 학교도 못 가는 심각한 게임 중독 청소년들이 문제가 생겼을 때 국가가 치료해 주고 또 이런 세상이 안 생기도록 예방을 하는 그런 법안입니다.

(중략)

김종덕 대표: 저허도 게임을 개발하는 입장에서 플레이어들이 일정 시간 동안 게임에 집중하도록 만드는 기본

적인 관점은 유지하고 있습니다. 그렇지만 게임이라는 것은 일정 시간 몰입을 하고 나면 릴렉스하는 시간을 부차적으로 만들도록 되어 있고요. 그게 기본적인 게임 개발의 교본 같은 것입니다. 사실 어떤 특정 학생들이, 중독이나 의원님께서 말씀하시는 그런 부류의 학생들은 사실상 개인들의 문제지 개인이나 가정이나 혹은 학업 스트레스나 이런 것들의 문제에 의해서 게임의 풍선 효과라는 것에 의해서 몰입하는 것이지, 이것이 게임이 잘못 되어서 게임에 중독되는 것은 아니다, 이거고요. 게임 중독이라는 게 의학적으로도 아직 사용하고 있지 않습니다.

(후략)

JTBC 직격토크 “게임중독법 논란” 中

여1: 너 미쳤어? 어디서 알지도 못하면서 함부로 말해?
 여2: 다른 직원들이 그러던데? 지금 니 얼굴 보니까 내가 잘못 들은 게 아닌가 보네.
 여1: 내가 언제 내 얘기 하자지? 윤지원인 척하는 너랑, 너를 윤지원이라고 철썩 같이 믿고 있는 그 남자 애 기하자고.
 여2: 그래서 말하겠잖아, 책임지겠다고.
 여1: 책임? 너 말은 똑바로 해. 난 그냥 얼굴 한번 비춰달라는 거였지. 내 이름 팔아가며 연애 걸라는 얘기는 아니었어. 너, 너 상황 판단이 안 되니? 어리다 어려. 연애 하나 컨트롤 못 하면서 니가 뭘 책임질 수 있는데? 니네 엄마한테 미안하지도 않니? 아직까지 청소부로 니 뒷바라지하시는데 니 나이에다 책임을 좀 지세요.
 여2: 야, 윤지원. 충고 하지 마. 니가 다 안다는 식으로 말해도 될 만큼 나 그렇게 한심한 사람 아니야. 남들이 아는 회사 다니고 차 몰고 좋은 옷에 비싼 화장품에 너 그런 걸로 니가 인생 똑바로 살고 있다고 느끼지? 그래, 그런 니가 부럽고 너처럼 못 사는 내가 못나 보이고 내가 너인 척 그 사람 만난 것도 맞는 데 그래서 부끄러운 건 나니까 니가 말 안 해도 그거 내가 제일 잘 아니까 나한테 함부로 말하지 마.
 여1: 야, 짜증나. 울긴 왜 울어? 지금 울고 싶은 사람이 누군데.
 여3: 근무 시간에 둘이 뭐 하는 거야?

KBS2 드라마특선 “연우의 여름” 中

다음으로 교사들에게는 제공된 설문지 자료는 아래와 같다. 설문지를 통해 교육 기관에서 듣기 문항을 출제할 때 어떤 어려움이 있는지 작성하도록 해, 교사들이 듣기 평가에 대해 어떤 생각을 가지고 있는지 확인했다. 그 후 순서대로 자료를 보거나 읽으면서 듣기 문항을 출제하도록 했다. 교사들에게는 듣기 문제를 출제할 때 어떤 점이 힘들었는지를 우선 질문하고 전체적인 기관 내에서의 듣기 평가에 대한 생각을 물어보았다. 그 다음으로 순서대로 동영상 혹은 듣기 지문을 보면서 적합한 문제

를 출제하고 그 이유를 작성하도록 하였다.

1.	?
2.	가 가 ? ?
3.	

파일명	내용	
01	문항	
	이유	
02	문항	
	이유	
03	문항	
	이유	

3.3. 연구 참여 교사

문항을 제출한 교사들은 한국어 교육 경력이 7년 이상이 된, 다년간의 평가 경험을 가지고 있는 한국어 교사 10명이었다. 전체 교사의 평균 교육 경력은 7.8년이며 모두 여성으로 성별에 따른 차이가 나타나지 않도록 했다. 평가 경력은 한국어 교육 기관에서의 성취도 평가 경험을 의미하며, 실험에 참가한 교사 중 한국어능력시험(TOPIK)과 같은 숙달도 평가에 출제나 채점으로 참여한 경력이 있는 교사는 없었다. 이는 가능한

교육 현장의 일반적인 교사들의 현황을 파악하기 위함이었다. 또한 출제 자이자 교사인 실험의 참여자들이 교육 기관에서의 고급 학습자 대상의 평가 문항을 개발한다는 상황을 가정하고 평가 문항을 작성하는 데에 집중하도록 한 것이다. 교사들의 개별 정보는 <표 2>과 같다.

실험이 진행되는 동안에는 교사들이 서로 누가 실험에 참여하는지 알 수 없도록 연구자가 개별적으로 연락하고 전자 우편으로 자료를 주고받 으며 진행되었다. 교사들은 받은 자료를 약 2주간(2014년 6월 9일~6월 20일) 검토하면서 듣기 문항을 출제했다. A그룹과 B그룹은 각 자료별로 2개에서 3개의 문항을 출제하였으며, 그러한 문항을 출제한 이유를 기술하여 전자 우편으로 제출하였다. 각 그룹의 교사들이 대답한 설문의 답변과 듣기 문항은 다음 장에서 구체적으로 살펴보고자 하겠다.

<표 2> 연구에 참여한 교사 정보

그룹	교유번호	성별	한국어 교육 및 평가 경험
A	가	여	7년
	나	여	8년
	다	여	7.5년
	라	여	8년
	마	여	9년
B	가	여	9년
	나	여	8년
	다	여	7.5년
	라	여	7년
	마	여	7.5년

5) 교사들에게는 어떤 문항을 출제할 수 있는지 그 유형에 집중하도록 했기 때 문에 교사들은 객관식 문항의 답항까지는 기술하지 않았다.

4. 연구 결과 및 논의

4.1. 출제자들의 듣기 평가에 대한 인식

문항의 유형이 듣기 자료가 아닌 교사들의 평가에 대한 목표나 이해의 차이로 달라지지 않도록 먼저 듣기 평가에 대한 교사들의 인식을 살펴보았다. 설문문의 내용을 통해 A그룹과 B그룹의 교사들이 학습자의 듣기 능력을 평가하는 데에 있어 서로 다른 관점을 가지고 있거나, 듣기 평가에서 주요하게 생각하고 있는 항목이 서로 다르지 않음을 확인할 수 있었다. 1번의 질문은 교육 기관에서 듣기 문제를 출제할 때 어떤 점이 어려운지를 물었고, 2번 질문은 현재 교육 기관 내에서 시행되는 듣기 평가의 문제가 뭐라고 생각하는지 물었다. 듣기 평가에 대해서 교사들은 숙달도 평가와 성취도 평가의 괴리, 듣기 평가의 실재성, 듣기 지문의 제작 및 듣기 평가를 위한 자료 수집의 어려움, 목적에 부합하지 않는 평가 문항과 평가 방식에 대해 문제를 제기하고 의견을 제시하였다. 이는 A그룹과 B그룹에서 모두 동일하게 나타난 내용으로 교사들 간의 차이를 발견하기는 어려웠다. 교사들의 구체적인 답변 내용은 다음과 같다.

Q1: 평소 듣기 문제를 출제하실 때, 어떤 점이 힘들었습니다? 그 이유는 무엇입니까?

A(가): 학습한 단어와 표현만을 사용해 듣기 평가를 하다 보니 출제할 수 있는 내용과 유형이 정해져 있다. 해당 급의 숙달도에 적합한 듣기 자료를 찾거나 만드는 것이 어렵고 또한 평가가 학습자의 실생활에서의 듣기 능력을 반영할 수 있는지 의문이 들 때가 많다. 평가를 위한 평가가 되는 경우가 많은데 이러한 이유로 듣기 평가에서 사용되는 자료는 실자료가 아닌 그때그때 상황에 맞게 만들어진 경우가 많다. 녹음을 할 때 정확한 발음

과 억양으로 속도를 조절하여, 그러나 최대한 자연스럽게 연기하는 것이 어렵다. 실생활에서의 듣기 상황과 거리가 있는 모습들이다.

B(나): 급의 수준에 맞춰 학습자가 모를 거라고 추정되는 표현은 되도록 넣지 않고, 답을 찾는 데에 필요한 정보를 중심으로 ‘자연스러운’ 듣기 자료(스크립트)를 만드는 것이 어렵다. 출제자가 직접 원고를 쓰고 녹음해야 하는 과정이 힘들다. 원고를 쓸 때 학습자가 배운 표현과 문법을 사용해서 대화 내용을 구성하되 실제 의사소통 상황과 같이 자연스러워야 하고, 실제 자료를 그대로 사용하지 않고 가공하여 텍스트를 만들어야 하기 때문에 중·고급으로 갈수록 듣기 문제를 출제하는 것이 부담이 된다. 속달도를 지향하는 성취도 평가가 정말 가능한지도 모르겠다.

Q2: 현재 교육 기관 내에서 시행되는 듣기 평가의 문제가 뭐라고 생각하십니까? 그 이유는 무엇입니까?

A(다): 실제로 학습자들이 접하게 되는 듣기 상황은 여러 가지 방해 요인(소음, 발음, 휴지, 모르는 표현 등)이 존재하지만 현재 교육 기관에서 시행되는 듣기 평가들은 학습한 표현과 어휘만을 사용한 정제된 문장으로 이루어진 텍스트이므로 실제적이지 않다. 이 때문에 실제 상황과 같이 전체 맥락에서 중요한 정보만을 선별해 들을 수 있는 능력을 평가하기가 힘들다고 생각한다. 보통 제시된 답항 가운데 알맞은 정보를 찾는 방식으로 듣기 평가가 진행되므로 학습자가 전체 듣기 내용 중 중요하다고 생각되는 정보를 능동적으로 선별해 조직할 수 있는지 확인하기 어렵다. 음성만을 사용하기 때문에 음성이외의 다른 변인들(두 사람이 대화하고 있는 환경이나, 두 사람의 표정, 주변인 등)을 구현하기 힘들어서 듣기 문제 출제를 할 때에는 언어에만 의존하게 된다. 그러나 이로 인해 내용이 실제적이지 않은 경우가 많다.

A(마): 평가 항목을 통해 무엇을 물어보려 하는지에 대한 목표가 불명

확하다고 생각한다. 교사들의 각 급의 성취도 및 숙달도에 대한 이해 부족을 원인으로 들 수 있을 것이다. 듣기 평가 항목들이 얼마나 타당하고 신뢰도가 있는지에 대한 검증이 거의 이루어져 있지 않은 상태에서 평가는 진행 된다. 평가 항목들에 대해 함께 고민하고 분석해 보는 시간이 부족하다. 자료 수집의 어려움, 난이도 등의 이유로 실생활에서의 듣기 상황(유의미한 듣기)을 고려하지 않게 된다.

B(라): 물론 성취도 평가를 위해 급의 수준에 맞게 가공된 듣기 자료를 이용하여 학습한 내용을 잘 알고 있는지, 이를 활용해 들은 내용을 이해할 수 있는지를 파악하는 것도 중요하다고 생각한다. 그러나 평소 수업 시간에는 동료 학습자는 물론 학습자의 한국어 수준에 맞춰 발화하는 교사와의 의사소통에도 어려움을 겪는 학생이 간혹 배운 표현을 열심히 외워서 듣기 시험에서는 좋은 점수를 받는 경우가 있다. 이러한 경우 우리의 듣기 평가 방법이 적절한지 고민을 하게 된다. 대부분의 실제 언어 상황에서는 전화 통화, 라디오 듣기, 안내 방송 듣기 등을 제외하면 청자는 화자의 얼굴 표정이나 몸짓, 주변 환경 등 비언어적인 요소를 통해 화자가 말하는(말하고자 하는) 내용과 의미를 가정할 수 있다. 그런데 현재 우리 교육 기관에서는 음성으로 된 듣기 평가만 시행하고 있다. 특히 중·고급에서는 미디어를 통한 시청각 자료를 활용하여 듣기 수업이 자주 진행되고 있으므로 듣기 평가도 영상과 음성을 포함한 다양한 방식이 요구된다고 본다.

교사들의 의견을 살펴보면, 교사들은 학습자의 숙달도를 고려한 성취도 평가에 있어 평가를 위한 적절한 자료를 찾기가 어렵고, 자료를 교사 스스로 만들어서 녹음하는 과정에서 실제성은 떨어지고 무엇을 평가해야 하는지 평가의 목적도 잃게 되는 경우가 있다고 했다. 또한 기관 내에서 이루어지는 평가에서는 주로 음성 평가가 이루어지기 때문에 그것이 실제로 학습

자들의 듣기 능력을 평가할 수 있는지 회의가 느껴지고, 보다 능동적이고 추론하는 능력이 있는 학습자로 교육하는 것과 그것을 평가하는 것이 어렵다고 했다. 교사들은 중급 이상의 학습자들의 듣기 이해 능력에서 무엇이 중요하고, 어떤 것들이 평가 항목으로 선정되어야 하는지 잘 인지하고 있었고, 그런 평가를 위해 평가 자료와 방법, 내용이 의사소통적이어야 한다고 생각하지만 그런 부분을 해결하는 것이 교사 개인의 문제가 아니라 동료 교사와의 협력 과정을 통해 가능하다고 했다. 이처럼 교사들의 듣기 평가에 대한 인식을 살펴봤을 때, 교사들의 인식의 차이로 인해 문항 개발의 차이가 나타나지는 않을 것이라고 예측할 수 있었다.

4.2. 출제된 듣기 평가 문항 분석

비슷한 경력과 같은 성별, 듣기 평가에 대해 비슷한 관점과 목표를 가지고 있는 A그룹과 B그룹의 교사들이 듣기 문항을 출제했다. 한 그룹은 구어 자료를 원시 자료로 활용했고, 다른 그룹은 문어 자료를 원시 자료로 활용했는데 각 그룹의 문항 유형에는 어떤 차이가 있는지 확인해 보도록 하겠다. 교사들이 출제한 듣기 평가 문항은 다음과 같다.

<표 3> A그룹과 B그룹 교사들의 듣기 문항 출제 내용

A그룹(구어 자료 활용)	자료	B그룹(문어 자료 활용)
<ul style="list-style-type: none"> • 뉴스의 제목으로 알맞은 것을 고르시오. • 흰개미가 문화재가 어떤 피해를 주고 있습니까? • 어떤 방법으로 흰개미 방제 작업을 실시하고 있습니까? • 목조 문화재를 보호하기 위해 어떤 대책을 실시할 계획입니까? 	01	<ul style="list-style-type: none"> • 뉴스의 제목으로 알맞은 것을 고르시오. • 다음 뉴스를 듣고 내용과 같은 것을 고르시오. • 뉴스를 통해 알 수 없는 내용은 무엇입니까? • 목조 문화재를 보호하기 위한 대책이 아닌 것은 무엇입니까? • 흰개미는 주로 어떤 상태에서 많이 생겨납니까?

<ul style="list-style-type: none"> • 논란이 되고 있는 법안은 무엇입니까? • 여자는 이 법안을 반대하는 사람들이 어떤 오해를 하고 있다고 말합니까? • 게임중독의 폐해 및 책임에 대해 두 사람의 의견은 어떻게 다릅니까? • ‘게임 중독법’에 대한 자신의 의견을 밝히십시오. 	<p>02</p>	<ul style="list-style-type: none"> • 토론의 주제는 무엇입니까? • 남자가 반대하는 이유는 무엇입니까? • 남자는 게임 중독이 누구의 문제라고 봅니까? • 여자는 게임 중독법에 필요한 비용을 누가 내야 한다고 했습니까? • 여자의 의견으로 맞는 것을 고르시오.
<ul style="list-style-type: none"> • 두 사람은 어떤 관계입니까? • 여자2가 화를 낸 이유는 무엇입니까? • 두 여자의 갈등의 원인은 무엇입니까? • 여자2는 여자1에 대해 어떤 감정을 느끼고 있습니까? • 이 장면의 앞에는 어떤 내용이 있었을 것 같습니까? 	<p>03</p>	<ul style="list-style-type: none"> • 드라마를 보고 내용과 같은 것을 고르시오. • 두 여자의 갈등의 원인은 뭐라고 생각합니까? • 청소부 여자(여자3)의 생각으로 맞지 않는 것을 고르시오.

반복되는 문항을 제외하고 A그룹과 B그룹의 교사들이 출제한 듣기 문항을 살펴보면, 첫 번째 자료(뉴스)의 경우, A그룹은 뉴스의 제목, 뉴스의 주요 정보인 흰개미로 인한 피해, 방제 방법, 향후 대처법을 확인하고 있음을 알 수 있다. 이는 실제로 한국인 화자들이 뉴스를 보면서 집중하는 내용이기도 하다. B그룹의 경우 같은 내용을 묻는 것으로 보이지만 하나 대책이 아닌 것을 고르라는 문제나 흰개미가 어떤 상태에서 주로 발생하는지와 같은 세부적인 내용을 확인하는 문항이 포함되어 있음을 알 수 있다. 이는 객관식 문항을 만든 후에 선택 답항까지 제시하지는 않았으나 세부 내용을 묻고 있음을 짐작할 수 있다.

두 번째 자료(토론)의 경우, A그룹과 B그룹이 주로 토론의 주제가 무엇이고 여자와 남자가 각각 주장하는 내용이 무엇인지를 묻는 질문이 주를 이루었다. 그러나 B그룹의 경우에 보다 구체적인 질문을 하고 있는 것을 볼 수 있는데, 반대의 이유나 남자가 게임 중독이 누구의 책임이라

고 보는지, 여자가 게임중독법에 필요한 비용을 누가 내야 한다고 했는지를 물어보고 있다. 이는 A그룹에서 두 사람의 의견이 어떻게 다른지 질문하는 것과 게임중독법에 대한 자신의 의견을 밝히려는 문항과는 상당한 차이가 있음을 알 수 있다. 이런 세부적인 정보는 문제를 미리 확인하고 반복해서 들어야 답을 찾을 수 있다. 구어 자료를 보고 문제를 출제하는 경우에는 자료의 특성상 전체적인 맥락이나 그것을 본 후에 청자로서 할 수 있는 행동이나 반응에 대한 문항이 더 많은 것을 볼 수 있다. 문어 자료를 읽고 문제를 출제하는 경우에는 눈으로 확인이 가능한 세부 정보 찾기가 포함되어 있고, 이는 문항 기술에 있어 기존의 틀을 지키고 있다고도 판단할 수 있는 부분이다.

세 번째 자료(드라마)의 경우에는 A그룹에서의 문항의 수가 B그룹에 비해 월등히 많은 것을 알 수 있었다. A그룹은 평균 3개~4개의 문항을 출제했지만 B그룹은 1개만 출제한 교사들도 있었고, 다른 두세 명은 2개 정도의 문제를 출제하였다. A그룹의 경우 두 사람의 관계, 두 여자의 갈등의 원인, 드라마의 앞부분에 어떤 내용이 있었을지 추측해 보는 문항을 출제한 반면, B그룹은 드라마의 내용과 같은 것을 고르라는 문제로 두 사람의 발화에서 볼 수 있는 구체적인 정보를 묻고 있었다. 그러나 구체적인 정보를 묻기에는 대화에서 찾을 수 있는 정보가 적어 문제를 출제하는 것이 쉽지 않았으리라 예측된다.

A그룹과 B그룹의 교사들이 출제한 문항에서 공통적으로 발견할 수 있는 내용은 구어 자료를 기반으로 문제를 출제한 A그룹의 경우에는 정확한 세부 정보를 확인하는 것보다는 전체 흐름이나 주제, 맥락과 상황을 고려한 문제를 주로 출제했다는 것이다. 또한 이러한 질문은 실제 뉴스, 토론, 드라마와 같은 듣기 자료들을 접할 때 주로 집중하게 되는 부분이기도 하다. 그러나 문어 자료로 기반으로 문제를 출제한 B그룹에서는 제목을 묻거나 이유를 묻는 것과 같은 질문이 있기는 하나 기존에 듣기 문제를 출제했던 방식을 따르며 세부 정보를 찾는 질문을 모든 경우

에 포함시키고 있음을 볼 수 있다. 또한 B그룹에서는 들은 후의 활동에 대한 문제가 전혀 출제되지 않은 반면, A 그룹에서는 3명 이상의 교사들에게서 그런 문항이 출제되었다. 이는 교사들이 구어 자료를 보면서 자연스럽게 하게 된 추측 활동을 평가에 반영한 것으로도 볼 수 있겠다.

문항 출제의 이유를 기술하는 부분에서 A그룹의 교사들은 구어 자료를 기반으로 문항을 출제하다 보니 화면의 구성이나 자막 등과 같은 그래픽 정보가 내용을 이해하는 데에 도움을 주었고 맥락이나 흐름을 따라가면서 주요 정보에 관심을 가지게 되었다고 했다.⁶⁾ B그룹의 경우에는 평소 문제를 출제할 때 자주 활용하는 ‘내용으로 맞는 것’이나 ‘맞지 않는 것’을 포함시켜서 내용을 잘 이해했는지 확인하게 했다는 설명을 볼 수 있었다.⁷⁾ 첫 번째 뉴스와 세 번째 드라마에 있어 B그룹의 교사들은 어휘의 난이도가 높아서 학습자들이 어려워할 것이라고 대답했고, 어휘만으로는 어떤 감정인지 예측하기 어렵다는 진술을 추가하여 문어 자료의 한계를 직접적으로 드러내기도 했다.

4.3. 듣기 평가에 대한 교사들의 기대

문항을 모두 출제한 후에는 앞으로 중고급 이상의 학습자들을 대상으로 한 듣기 평가 문항을 개발할 때 어떤 점이 개선되어야 하는지를 물었고, 교사들은 답변은 다음과 같았다.

-
- 6) 교사들이 원시 자료로 삼은 구어 자료에서는 동영상을 시청하면서 이해를 도울 수 있는 다양한 그래픽 자료가 다량 포함되어 있었다. 그러나 실제 듣기 평가에 이것이 포함되는 경우는 드물다. 이 부분은 앞으로 의사소통적 듣기 평가 문항을 개발하는 데에 참고할 수 있는 내용이라고 판단된다.
 - 7) 이해영(2005:290)은 듣기 능력 평가를 위한 문항 유형에서 8회 한국어 능력 시험 듣기 영역을 분석한 결과 세부 내용 이해하기가 전체 1급~6급까지에 고르게 가장 많은 부분을 차지했다. 전체적으로 37%에 달했다.

Q3: 중고급 학습자들을 대상으로 듣기 평가 문항을 개발할 때 어떤 점이 개선되어야 할 거라고 생각합니까? 그 이유는 무엇입니까?

A(나): 학습자들이 일상생활에서 직접 접하게 되는 실자료를 사용해 학습자들의 실제 듣기 능력을 평가할 수 있었으면 좋겠고, 단순히 이해만 하는 듣기가 아니라 들은 내용을 바탕으로 자신의 생각을 말하거나 쓰는 활동이 통합된 평가가 필요하다. 또한 시험을 통한 일회적인 평가보다는 프로젝트와 같은 과제를 통해 듣기 능력을 평가해야 한다.

B(마): 평가에서 실자료를 활용하는 것보다 중요한 것은 어떻게 평가를 하는가 하는 문제이다.⁸⁾ 실제 수업 중에도 듣기의 목적을 분명히 하고, 자신이 원하는 정보를 얻는 혹은 해야 할 행동을 잘 알 수 있는 듣기 활동을 구성하여 실시하고 평가에서도 읽기 평가와 별로 다르지 않은 내용 확인이 아닌 목적이 있는 듣기 평가가 이루어져야 한다.

교사들은 듣기 평가가 현재보다 더 분명한 목적을 가지고 일상생활에서 그러하듯이 듣고 나서 무엇을 할 것인지 고민하게 만들어야 한다고 했다. 앞서 교사들이 출제한 문제를 살펴봤을 때, 교사들은 자신도 인식하지 못하는 동안 구어 자료를 보면서 전체 맥락을 파악하고 각각의 장르에 맞는 듣기 활동을 하고 그 활동에 맞는 평가 문항을 제안한 것을 볼 수 있다. 구어 자료가 아닌 문어 자료를 기반으로 문항을 출제했을 때

8) 박민신(2008:19) 듣기 평가에서 교육적으로 가공된 듣기 텍스트가 갖고 있는 입력의 용이성은 실질적인 문제로서 중요한 의미를 갖는다고 설명한다. 그러나 학습자의 수준에 맞는 입력의 수정이 필수적인 문법이나 어휘 등과는 달리 상호작용적인 측면에서 확인할 수 있는 자연스러운 구어의 특성은 듣기 텍스트를 구성하는 데 있어 진정성을 보장 받아야 하며 학습자들이 정제된 담화 상황이 아니라 실제 담화 상황에서 담화 내용을 얼마나 이해할 수 있는지에 대한 보다 엄밀하고 진정성 있는 평가가 이루어져야 한다고 보았다.

에는 기존의 문제 출제 방식을 그대로 따르게 되었다. 이를 통해 듣기 평가에 적합한 문항을 출제하기 위해 출제의 과정에서 교사가 원시 자료로 삼는 것을 무엇으로 하느냐가 영향을 미친다는 것을 확인할 수 있었다. 박성경·송향근(2008:303-304)은 듣기 평가의 항목을 문맥과 상황 이해의 평가, 사회 문화적 이해의 평가가 이루어져야 한다고 했다. 담화가 이루어지는 맥락과 담화를 전개하는 태도나 방법에 대한 이해력을 평가하고 관용 표현이나 높임법의 체계를 이해하여 주제나 배경, 청자와 화자의 관계 등을 제대로 알아낼 수 있는가를 평가해야 한다는 것이다. 이러한 평가를 위해서는 현재 진행 중인 평가의 원시 자료로부터 다시 살펴 볼 필요가 있다. 본 연구에서 진행된 듣기 문항 출제의 과정에서 구어 자료를 기반으로 문항을 출제한 교사들은 바로 이러한 부분을 평가 문항에 실제로 반영했다고 볼 수 있을 것이다.

5. 결론

한국어 학습자들의 듣기 능력을 평가하기 위한 다양한 방법들이 시도되고 있다. 담화 분석을 통해 듣기 텍스트의 실제성을 높이고 있으며, 듣기 지문을 녹음할 때도 지나치게 정제된 발화가 아닌 실수와 반복, 휴지 등이 있는 자연스러운 상황을 연출하기도 한다. 그러나 평가 문항이 의사소통의 현실을 담아내지 못한다면 학습자들은 결국 형태적인 것들에 집중하게 되고 의사소통적인 듣기 평가는 어려워질 것이다.

우리는 실제 의사소통 상황 속에서 무엇을 언제 들으며, 왜 듣는지를 고민해 봐야 한다. 길을 가다가 음악이 들리거나 옆 사람이 의미 없이 말하는 내용이 들리는 경우를 제외하고는 우리는 반드시 목적을 가지고 무언가를 듣게 되고 그에 따른 반응을 하게 된다. 원론적인 이야기일 수 있겠으나 듣기 평가 문항을 개발하는 데 있어서 개발자는 실제 구어 자료를 충분히 검토해야 할 것이며, 이는 실제성을 확보하는 것뿐만 아니라

개발자로 하여금 무엇에 더 초점을 두어야 하는지 고민하게 만든다. 구어 자료에서 시작된 평가 과정은 듣기 본연의 기능에 보다 집중할 수 있도록 도와주고, 다른 언어 기능과 통합된 듣기 평가를 할 수 있는 가능성을 열어 줄 수 있을 것이다.

강명순 외(1999:52)에서와 같이, 우리가 뭔가를 들을 때 모든 정보에 항상 똑같은 정도의 주의를 기울이지는 않는다. 그것은 인간의 능력으로는 불가능하거나와 그럴 필요도 없기 때문이다. 청자는 듣기 처리 과정에서 청각 기관을 통해 입력된 자료를 자신의 듣기 목적에 따라 선택적으로 처리하면 된다. 따라서 듣기에서 중요한 것은 소리로 입력된 자료를 들었느냐 못 들었느냐가 아니라 의사소통 상황에서 자신에게 필요한 정보를 얼마나 잘 구별해 들을 수 있느냐 하는 것이다. 또한 평가에 이것이 반영되기 위해서는 원시 자료의 선택이 지금과는 달라져야 할 것이며, 현장에서의 문항 개발자인 교사는 듣기 평가의 목적에 적합한 장르와 내용의 실제 듣기 자료를 접하며 문항의 다양성을 꾀해야 할 것이다.

<참고 문헌>

- 강명순 외(1999). 한국어 듣기 능력 평가 방안: 숙달도 평가를 중심으로, <한국어교육> 10-2호, 국제한국어교육학회. 47쪽~94쪽.
- 강승혜 외(2006). 『한국어 평가론』. 서울: 태학사.
- 김지영 외(2014). 한국어 숙달도 평가의 문제은행 도입을 위한 문항 일반 정보 구축 방안 연구, <이중언어학> 54호, 이중언어학회. 75쪽~98쪽.
- 민옥희·김미옥(2012). 동영상 활용한 한국어 듣기 능력 향상에 관한 연구: 드라마와 TV뉴스를 활용하여, <한국언어문화교육학회 제7차 국제학술대회>, 한국언어문화교육학회. 287쪽~301쪽.
- 박미엽(2012). 이주여성의 한국어 듣기 평가에 대한 일고찰: 이주여성을 위한 새로운 한국어 듣기 평가 패러다임, <국어문학> 52호, 국어문화회. 5쪽~29쪽.
- 박민신(2008). 한국어 듣기 평가 텍스트의 진정성 연구: 대화형 텍스트의 상화작

- 용적 측면을 중심으로, <한국어교육> 19-1호, 국제한국어교육학회. 93쪽~112쪽.
- 박성경·송향근(2008). 한국어 듣기·말하기 성취도 평가에 관한 연구: 초급을 중심으로, <우리어문연구> 30집, 우리어문학회. 299쪽~325쪽.
- 안미란(2007). 학문 목적의 한국어 듣기 평가: 대학 지원자의 학업능력평가를 중심으로, <이중언어학> 34호, 이중언어학회. 247쪽~270쪽.
- 양명희·김정남(2011). 『한국어듣기교육론』. 서울: 신구문화사.
- 이준호(2010). ‘한국어능력시험’에서의 추론 기술 평가 연구: 이해 영역을 중심으로, <한국어학> 46호, 한국어학회. 317쪽~351쪽.
- 이혜영(2005). 한국어 이해능력 평가의 원리 및 실제, <한국어교육> 16-3호, 국제한국어교육학회. 275쪽~304쪽.
- 임철성(2009). 듣기 능력과 듣기 평가의 목표: 수능을 중심으로, <화법연구> 14호, 한국화법학회. 279쪽~300쪽.
- 임철성(2011). 국가 수준 학업성취도 평가 개선을 위한 ‘듣기’ 평가 이론의 정립, <국어교육> 134호, 한국어교육학회. 35쪽~58쪽.
- 전은주(2008). 한국어 능력 시험 평가 문항의 내용타당도 분석: 제12회 일반 한국어(S-TOPIK)의 쓰기·듣기·읽기 영역을 중심으로, <국어교육학 연구> 31호, 국어교육학회. 129쪽~165쪽.
- 심윤희(2005). 듣기 문항 분석을 통한 듣기 평가의 지향점, <국어교과교육연구> 3호, 국어교과교육학회. 109쪽~145쪽.
- 차경환·신동일(2001). 『영어청취론』. 서울: 한국문화사.
- 최은규(2006). 유형별로 본 한국어 능력 평가의 실제와 과제: 배치 시험과 성취도 시험을 중심으로, <한국어교육> 17-2호, 국제한국어교육학회. 289쪽~319쪽.
- 황성은·심혜령(2014). 한국어능력시험 중급 듣기 문항 분석 연구: 상황맥락을 중심으로, <언어사실과 관점>, 연세대학교언어정보연구원. 237쪽~259쪽.
- Gary, B.(2010). *Assessing Listening*. Cambridge: Cambridge University Press.
- Shohamy, E., & Inbar, O.(1991). Validation of listening tests: the effects of text and question type. *Language Testing*, 8, 23-40.

김지혜(Kim Jihye)
고려대학교 한국어센터
136-701 서울시 성북구 안암동
전화번호: 010-9229-0729
전자우편: benjoin34@korea.ac.kr

130 이중언어학 제56호(2014)

접수일자: 2014년 8월 11일

심사(수정)일자: 2014년 8월 20일

게재확정: 2014년 9월 17일