

영어교육 50권 2호 1995년 여름

## 한국의 영어능력 평가의 발전과정: 그 이론과 실제\*

최 용 재

(동국대학교)

Choe, Yongjae. (1995). Development of English proficiency testing in Korea. *English Teaching*, 50(2), 133-150.

English proficiency testing in Korea has made noticeable changes during past three decades in its theories and test formats. Up to 1965, there were very few published articles in the field of TEFL as a whole, let alone the articles on proficiency testing. After three eventful decades, quite a considerable number of articles on English proficiency testing have been published and some of those are really noteworthy in every respect. The test formats have also changed drastically from mainly subjective ones to objective ones. Theorists have claimed that objective test format should be employed to secure the reliability. However, they sometimes contradicted themselves claiming that discrete-point test format should give way to integrative test format to enhance the validity of the test. The real changes that happened in the testing have been done by government run examinations for screening qualified students to enter higher schools. Because of the massive number of examinees, the examinations have to employ objective test format of multiple choice or true-false items to enable the machine grading. The recent development in the English proficiency testing is characterized by regained emphasis on subjective test formats despite their cumbersome grading efforts.

### 1. 서론

지난 30여년에 걸쳐 이루어진 영어능력 평가를 위한 이론의 발전 및 변천과 평가도구의 변화는 자못 괄목할 만하다. 특히 이론적 탐구의 면에서 볼 때, 1960년대 중반까지만 해도 영어학 전반에 대해서도 이론적 연구가 거의 황무지와 같았음은 물론이고 (성찬경 외, 1989), 영어능력 평가에 대한 이론

\* 이 논문은 1993년도 한국학술진흥재단의 공모과제 연구비에 의하여 연구되었음.

적 연구가 겨우 걸음마를 하는 단계에 있었는데 1980년대 후반부터 1990년대에 오면서 상당한 수준에 이르게 되었다. 또한 영어능력 평가 자체를 놓고 볼 때에도, 주관식 위주의 시험에서 객관식 시험으로 다시 주관식 시험으로, 그리고 구두시험에서 지필 시험으로 다시 구두시험으로 바뀌는 경향을 보였다. 이는 평가방법의 변천이 이론의 발전과는 아무런 연관도 없이 시대에 따라서 좌우되는 경향이었다가 이 몇 년 사이에 어느 정도는 이론의 발전과 연관을 갖는 것 같은 느낌이다. 사실 평가의 연구는 그 출발이 이미 실시해 오고 있었던 평가방법 및 형식에 새로운 평가방법과 형식 그리고 도구의 소개로 이루어졌다. 평가에 대한 이 새로운 연구는 순수이론을 개진한 것과 실제 평가도구를 설명한 것들로 나누어 생각할 수가 있다. 여기에서는 이런 순서를 좇아서 지난 30여년 동안 변화하고 발전해 온 평가의 전반에 대하여 서술하려고 한다.

평가에 대한 순수이론은 먼저 평가방법 전반에 걸친 소개로 시작한다. 본격적인 이론의 전개는 최근의 일이고, 외국의 예를 본격적으로 든 것도 1980년대에 들어 와서 일어난 현상이다. 실험연구의 경우에는 주로 언어기술(language skills)별로 이루어졌고, 특히 듣기에 치중했던 것은 듣기의 중요성이 강조되었던 때문인 것 같다. 또한 시험의 형식도 이른 바 객관식 시험으로 분리시험에서 종합적 언어능력의 시험으로 규칙빈칸 메우기 절차(cloze procedure)를 소개하고 연구한 논문들이 많이 쏟아져 나왔다. 그러나 평가의 백미는 뭐니 뭐니 해도 실제 시험에 대한 연구와 비판이다. 고등학교 입학시험에서 출발하여 대학교 입학시험, 대학원 입학시험, 취업의 일종으로 교원임용고사까지를 다루었는가 하면 대학교 교양영어시험에 대한 연구도 있었다. 이런 시험에서 중요한 것은 그 역류현상(washback effect)이어서 이에 대한 언급도 일찍이 있었다. 우리 나라에서 평가도구의 발달은 외국의 평가도구를 본받아 시행한 것이 거의 전부라고 할 수 있는 것이어서 외국의 평가방법이나 도구의 소개 또한 활발했다. 이런 현실 속에서 반성되는 바는 아직도 우리 식의 평가도구를 마련해 본 적이 없다는 것이다.

전체적으로 볼 때, 지난 30여년 사이의 영어능력 평가는 일제시대로부터 전통적으로 실시되어 오던 소규모의 주객관식 방법들에서 벗어나 전국규모의 시험이라는 필요에 따라서 분리시험의 일종인 객관식 선다형 시험이 소개되고 애용되었으나, 이론적 연구와 이해가 쌓이자 분리시험에 대한 반성이 일어났고 그런 현실의 바탕 위에 통합시험의 필요성을 주장하는 목소리가 들리기 시작했다. 더구나 대학입학시험에서 학교마다 주관식 시험의 중요성을 막연하게나마 주장하게 되자 다시 옛날로 되돌아가는 듯한 느낌마저 들게 되

었다. 그러나 우리는 항상 외국의 경우를 들어서 우리를 반성하는 형식을 취하여 오는 과정에서 우리의 실정에 알맞은 평가도구의 개발을 서둘러 왔으나 그 결과는 외국식의 방법을 수정하는 정도에 그쳤다. 지금은 우리의 독자적 형식의 개발을 서두르는 경향을 바라는 시기다. 지난날에 대한 반성은 오늘 날에 우리가 나아갈 길을 제시하고 있는 바, 이제는 성숙된 우리의 역량으로 우리 식의 평가 이론과 방법을 개발할 시기임을 시사한다.

## II. 이론의 발전

평가에 대한 이론적 접근을 맨 처음으로 언급한 시기는 1965년으로 생각되는데, 이 해에 발표된 두 건의 논문에서 그 흔적을 찾아 볼 수가 있다 (박상용, 1965; 서울대학교 어학연구소, 1965). 박상용은 응용언어학이 교실수업에 어떻게 적용되는가를 소개하는 일환으로 당시 미국에서 일고 있던 객관식 영어시험을 소개하는 것이었고, 서울대학교 어학연구소의 연구는 영어학력의 표준화검사를 시도하여 그 검사의 결과를 보고한 당시로서는 획기적이고 선구적 역할을 해냈다. 서울대학교 어학연구소의 보고서는 먼저 표준화검사를 정의하고, 측정도구의 제작과정과 예비시험과 본시험의 실시 그리고 그 결과의 해석 등의 내용을 담고 있어서 시험다운 시험의 모범을 보였다. 표준화 검사의 실험 대상도 31개 고등학교에서 실시하여 실험대상의 신뢰도를 높였을 뿐만 아니라, 결과의 해석에 있어서는 당시로서는 영어교육계에 생소했던 심리측정방법을 동원하여 원점수 (raw score), 백분위 순위(percentile rank), T-점수 (T-score)와 같은 통계적 수치를 나열하여 어느 정도는 제 모습을 갖춘 논문을 내놓았다.

1970년대에 이르러서는 평가방법에 대한 전문적 언급과 더불어 반성도 나오게 되었다. 구체적으로 문법구조를 중심으로 한 영어시험의 시도가 있었고 (이 옥남, 1971), 평가도구의 작성법에 대한 더욱 자세한 안내도 나왔다 (용성해, 1971). 여기서 말하는 평가도구의 유형은 주로 객관식 선다형으로 대규모 시험을 전제로 한 것이다. 이런 대규모 시험을 위하여서는 단시간에 채점을 할 수 있는 방법이 필요한 바, 기계를 이용한 채점이 권장되는 한편 기계를 사용할 경우의 장점이 소개되었다 (이 종찬, 1972). 기계를 사용하여 채점을 할 경우에는 여러 가지 통계를 쉽게 낼 수가 있어서 문항분석, 변별도, 난이도, 표준편차 등등을 처리하는데 편리함을 보여 주었다. 그러나 객관적이라는 선다형 시험이 절대적 지지를 받고 있던 시기에도 개선의 목소리가

들리기 시작했다. 즉 선다형 시험과 단답형 시험을 비교하여 평가방법의 획일적 사용을 지양하고 개선해야 하는 이유를 들었다 (신 홍식, 1976). 뿐만 아니라, 여러 가지 객관식 방법들은 말할 것도 없이 주관식 방법에 이르기까지 평가도구 전반에 걸쳐 부정적 측면을 들어서 더욱 신뢰할 수 있는 평가도구의 개발을 주장하기에 이르렀다 (최 용재, 1974).

1980년대에 들어서서는 영어능력의 측정상의 기술적 문제에 국한하지 않고, 더욱 근본적인 문제에 대한 추구가 시작되었다. 특히 주목할 만한 것은 영어능력의 평가에 있어서 피상적 언어지식이나 사용요령만을 측정하는 이론과 도구의 한계를 지적하고, 방대한 언어자원적 측면에서 더욱 포괄적이고 근본적인 언어능력의 측정을 주장했다 (최 용재, 1982). 물론 이 시기에도 좋은 영어능력 평가도구의 마련을 위한 기술적 고려사항들로서 문항작성시의 주의사항들과 같은 실질적인 노력이 계속되었으며 (권 오량, 1983), 문항분석과 같은 실험적 연구도 진행되었다 (박 원, 1985). 이 때 실시한 문항분석의 내용을 보면 선다형 문제에 그치지 않고 모든 형태의 평가도구 (problem-solving test, pragmatic test, maze test, cloze test, error-correction test, lecturette test, reading comprehension test, listening comprehension test)에 걸쳐서 실시되었다. 이 시기쯤에서는 평가이론의 반성이 있을 법하여, 그런 노력이 보이기는 하지만 아직도 언어능력의 근대적 측정이론을 망라하여 소개하는 수준에서 그치고 있다 (최 인철, 1989).

그러나 1980년대에서 특기할 일은 언어능력의 측정방법에 있어서 해묵은 논쟁이 되살아나는 듯한 경향을 보였다는 것이다. 1970년말 무렵에야 객관식 시험에 대한 논문이 나올 만큼 (강 순경, 1978), 당연하게 받아들여졌던 평가도구의 형태는 객관식 선다형의 분리시험이었다. 그러나 통합적 언어능력의 측정을 내세우는 규칙빈칸매우기 시험 (cloze test)의 소개와 더불어 간단한 실험이 행하여진 이래 (김 순복, 1982, 1983), 이에 대한 해설과 본격적 연구가 속속 발표되어 1990년대에까지 계속되었다 (신 용진과 전 병만, 1983; 임 호순, 1983; 민 대식, 1985; 김 순신, 1994; 최 용재, 1981, 1993a, 1993b). 특히 규칙빈칸매우기 절차 (cloze procedure)에 관한 최초의 본격적인 연구로 들 수 있는 박사학위 논문이 이 시기에 발표되어 이 분야에 대한 적극적 연구에 큰 기여를 했다 (전 병만, 1985).

1990년대에 들어와서는 영어능력 시험과 세계의 언어교육 추세와의 비교가 소개되었고 (홍 성면 1990), 이와 관련하여 의사소통 능력을 기르는 영어교육과 그 결과를 측정하는 의사소통 능력의 측정을 목표로 하는 평가도구에 관한 논의가 새로 나오기 시작했다 (전 병만, 1994b). 물론 영어능력의 측정

나아가서는 언어능력 측정의 가능성도 심도 있게 연구되어 여러 가지 문제가 제기되기도 했다 (최 용재, 1990). 그리고 원론적인 문제의 제기도 끊이지 않아, test와 exam의 차이를 원론적으로 규정하려고 한다거나 이들 사이에 있을 수 있는 서로 다른 문항에 관한 논의도 있었다 (최 용재, 1992). 더욱이 서울대학교 어학연구소가 새롭게 시도한 표준화시험 (Seoul National University Criterion-Referenced English Proficiency Test)의 개발이 보고 되기도 했다 (최 인철, 1994). 이들에서 보이는 경향은 단편적이기는 하지만 1990년대를 어떻게 이끌어갈지 시사하는 바가 크다고 하겠다.

### III. 영어의 기능별 평가

영어능력의 평가에 관한 연구는 전통적인 문법번역식 문제에 대한 연구를 완전히 벗어나서, 주로 영어의 기능 (skills)만을 다루는 논문들이 발표되었다. 특히 음성언어 (spoken language)에 관한 관심과 필요에 따라서 이 부분의 연구가 많았던 것이 특징이라고 하겠다. 그러나 음성언어에 속하는 두 가지 기능 가운데서 여러 가지 제약 때문에 듣기 위주의 연구가 가장 활발했다는 것도 특징 가운데 하나이다. 말하기의 경우는 겨우 한 두엇 정도의 논문이 발표되었을 뿐이다. 읽기와 쓰기도 말하기와 같은 정도로 관심의 밖에 있어 왔다. 음성언어의 평가는 대규모 시험이라는 특수성 때문인지 기계를 통한 듣기가 주종을 이루고, 우리 나라 영어교육의 초창기에 실시했다던 (이 회승, 1984) “담화회” (conversation meeting)같은 실제적인 평가는 보고된 바가 아직은 없다. 특히 중등학교에서 방송을 통한 듣기시험의 실시가 있는 뒤로는 이 분야에 커다란 발전을 가져 왔다고 볼 수 있다.

듣기시험에 관하여서, 1970년대에는 듣기 자체로서 독자적 언급이 거의 없던 듣기시험이 (곽 경수 1977), 1980년대에 들어서면서 한국영어교육학회가 듣기의 중요성을 적극적으로 강조하기 시작한데 힘입어 듣기에 대한 관심이 증폭되었다. 한국영어교육학회는 당시 문교부에 대학입학 예비교사에 청취력을 포함시킬 것을 건의하였다 (1980년 2월 22일자). 이에 대하여 문교부에서는 제한된 시험시간과 별도의 출제를 위한 준비가 부족하다는 이유를 들어서 불가능함을 회신해 왔다 (1980년 4월 30일자). 한국영어교육학회는 이에 굴하지 않고 방송망을 통한 청취력 시험의 가능성을 점검하고 주장하기에 이르렀다 (이 용실, 1982). 한국영어교육학회의 이런 노력이 결실을 맺어, 드디어 1982년 11월에 중등학교 수준에서 방송을 통한 듣기시험이 실시되었

다. 다음 해에는 중등학교에서의 듣기평가에 대한 통계적 연구가 일부지역에 국한된 것이었지만 발표되었다 (권 오량, 1983). 중등학교의 듣기시험이 주로 소리의 구별이라든가 짧은 대화를 듣고 알아듣기에서 그친 반면에, 본격적인 듣고 내용파악하기의 소개는 이중언어교육과 연관하여 이루어졌고 철저하게 고찰되었다 (홍 성면, 1984). 1990년대에 와서도 듣기능력 평가의 연구는 항상 지도방안과 결부하여 실시되고 보고되었다 (박 약우, 1990). 다만 변화가 있었다면 더욱 철저하고 본격적인 연구가 선을 보였다는 것이다. 더욱이 문항의 연구가 청취력 문제에서도 심도 있게 실행되었다 (차 경환, 1993).

말하기 시험은 항상 면담을 전제로 하는 것이지만, 듣고 말하기 시험을 별도로 언급하기 시작한 것이 1980년대 중반이었다 (문 정일, 1985). 말하기 시험의 가장 초보로서 주어진 문장을 구두로 확장하게 하는 방법이 소개되었으나 (Fernandes, 1987), 면담수준의 기술이나 문제에 대한 논문이나 보고는 없었다. 1990년대에 들어와서 발표된 논문들도 단편적인 것들에 지나지 않았다. 성인의 구두발표 능력의 측정과 그것이 그들의 전략적 언어능력과의 상관을 고려한 논문이 나왔다 (오 석봉, 1992). 면담에 대한 연구의 하나로 들 수 있는 것은 면담평가의 신뢰도를 다룬 것이다 (김 영서, 1993). 한 마디로 요약해 보면, 듣고 말하기를 측정하는 데에 있어서 관심의 초점은 시험관의 자질과 훈련 그리고 대규모 시험 등의 여러 가지 여건상 실제적으로 실현할 가능성이 없는 말하기 시험은 관심의 밖으로 친 것 같다. 이런 경향은 문자언어 (written language)에 이르면 더욱 뚜렷해진다.

읽고 쓰기의 평가에 대한 논문을 조사해 보면, 쓰기 시험에 관한 논문은 찾아 볼 수가 없고, 읽기 시험에 관한 논문은 겨우 하나를 발견할 수 있을 뿐이다. 이 논문은 읽고 그 내용을 파악하는 시험 (reading comprehension)의 정의와 문항작성을 소개하면서 질문지의 형식을 단답형으로 할지 선다형으로 할지를 가늠해 보는 입문적 또는 안내서적 내용에서 그치고 있다 (김 용석, 1980). 쓰기 시험에 대한 연구나 언급이 없는 것은 교육현장에서 쓰기를 경시하고 있는 현실의 반영이라고 볼 수도 있고, 대규모 시험에서 실질적으로 채점이 불가능하다는 점도 쓰기가 제외된 이유로 들 수가 있겠다. 사실 TOEFL에서도 쓰기 시험은 흉내만 내다가 아예 제외해버린 것을 볼 때, 이해가 되는 점도 있지만 우리의 실정은 쓰기 교육의 부재에서 말미암았다는 점이 우려되는 바이다. 읽고 내용을 파악하기에 대하여서도 연구가 부진한 이유를 생각해 보면, 지문을 주고 그 지문 안에 있는 사실의 피상적인 인지에 그치고 이른 바 행간을 읽는 식의 문제를 경험하지 못한 실정에서 읽고

내용을 파악하는 능력의 평가는 별로 큰 관심을 일으키지 못한 것 같다.

#### IV. 선발시험

선발시험은 고등학교 입학고사, 대학교 입학고사, 대학원 입학고사, 교사 선발고사 등으로 나누어서 생각할 수가 있겠다. 그러나 많은 학자들이 가장 관심을 갖는 선발고사는 단연 대학 입학 예비고사 또는 대학 수학능력 고사이다. 전국적으로 실시되고, 해마다 70만 여명이 응시하는 시험인 만큼 전문가 비전문가를 막론하고 관심이 지대할 뿐만 아니라 저마다 한 마디씩 소견을 표시하기 마련이다. 또한 대학 입학을 좌우하는 절대적인 척도인 만큼 이 시험을 준비하는 고등학교의 교육이 이 시험의 경향에 따라서 좌우되기 마련이었다. 물론 입학시험의 대비를 주로 삼는 고등학교 교육을 반성하는 목소리도 크게 들리지만, 현실적으로 수험생 자신과 학부모의 요구 그리고 나아가서 사회 전체의 요구가 대학입학시험에서 좋은 점수를 받는 것이므로 어쩔 수 없이 대학 입학시험에 맞추어 수업을 할 수밖에 없는 것이 실정이다. 학교 수업이 대학입학 시험의 출제 경향에 영향을 주는 것보다는 오히려 대학입학 시험의 경향이 학교수업에 영향을 주는 경우가 더 절대적이어서 이른 바 역류현상이 심각한 문제가 되었다.

전국규모의 대학입학 예비고사가 처음으로 실시되던 때와 같은 시기에 이미 역류현상에 대한 우려가 나오고 있었다 (박 상용, 1971). 물론 이 연구는 서울대, 연세대, 고려대, 이화대, 중앙대, 부산대, 전남대, 충남대, 경북대 등 전국 17개 대학의 입학시험문제를 분석하여 문제를 제기하고 있어서 대학입학 예비고사와는 무관한 것처럼 보일 수 있으나, 이들 우수한 대학의 출제경향이 수험생들에게 주는 심리적 부담과 영향을 고려할 때 대학입학 예비고사가 주는 영향도 같은 맥락에서 보아야 한다. 이 연구에서 주목할 만한 것은 음성언어에 대한 문제는 하나도 없이 지필 시험 위주로 발음 맞추기, 문법규칙, 어휘, 철자, 어귀 등과 읽고 답하기와 번역이 전부라는 지적이다. 뒤에 많은 학자들에 의하여 음성언어 능력의 측정으로 잘못 이해되는 발음 맞추기는 발음기호를 통한 물기와 강세 및 쉽자리 등을 묻는 것이 전부인데, 수험생이 정답을 했다고 해서 그가 실제로 올바른 발음을 하거나 들어서 식별할 것으로 믿을 수가 없으므로 한 마디로 말도 되지 않는 문제였다. 이런 실정에서 가뜰이나 고등학교의 영어교육이 입시 위주로 흐르기 마련인데 정히 우려되는 바는 음성언어교육을 소홀히 할 것이라는 개연성이었다. 이런 우려는 10

년 뒤에도 제기되었는 바, 대학입학 예비고사의 문제유형은 문법번역식으로 학습한 학생에게 유리할 뿐이어서 고등학교 교육도 그렇게 될 수밖에 없음을 지적한 논문이 나왔다 (김 덕기, 1980).

더구나 1993년부터서는 대학 수학능력 고사 외에도 이른 바 “대학별 본고사”가 부활되어 학교에 따라서는 자체적으로 마련한 시험문제로 학생을 선발하기 시작하여, 대학입학시험 자체뿐만 아니라 그에 의하여 영향을 받는 고등학교 영어교육에 커다란 변화를 가져오게 되었다. 물론 대학의 입학시험이 고등학교 교육에 영향을 주는 역류현상이 필연적이라고 해도 출제자는 고등학교 교육을 조금만 고려한다면, 그 반대로 고등학교에서의 영어교육이 대학입시에 반영되는 바도 무시할 수가 없을 것이다. 사실 대학에서는

大學入試問題의 傾向이 高等學校 教育에 미치는 影響이 크다고는 생각하면서 그만큼 所重히 出題問題를 다루지 못하는 것도 各大學의 實情이다. 또한 大學 答案紙에 나타난 現象도 斷片的이기는 하지만 高等學校 教育의 一而이 表象化되었다고도 볼 수 있다. (전형국, 1966)

이런 주장은 1980년대 중반에도 있었던 바, 대학입시에서 당시로서는 새롭게 (1) 본고사 (영어)의 부활, (2) 전문가의 활용, (3) 청취시험의 채용, (4) 지식보다는 의사소통 위주의 시험, (5) 고교교육의 올바른 방향을 제시할 수 있도록 고려한 출제 등을 권고하고 있었다 (박 형기, 1986). 사실 모든 수준에 있어서의 입학시험 출제경향과 이를 준비하는 입장에 있는 학교교육과의 상호영향은 입시제도가 있는 한은 영원히 순환적인 상관관계를 가질 수밖에 없을 듯하다.

## 1. 대학입학고사

대학입학 예비고사 (영어)의 문제점과 개선방향에 대하여서는 1978년에 한국영어교육학회 주최로 대토론회가 열려 출제자와 지도교사의 입장을 정리했다. 특히 주목되는 것은 출제자의 입장에 서서 지난 10년간 (1969년부터 1978년까지)의 대학 입학 예비고사 영어문제를 분석하여 제시한 도표이다 (전 상범 1978). 이 분석에 의하면, 읽고 내용을 파악하는 문제가 문법문제와 더불어 가장 많이 출제되었는 바, “내용파악문제 : 문법문제”의 대비는 백분율로 1969년부터 시작하여 1978년에 이르는 동안 43:12, 42:20, 34:24, 12:34, 20:28, 22:44, 22:30, 50:10, 46:22, 54:6 등으로 반비례의 현상을 보인다. 이 밖에도 숙어와 어휘가 다음으로 많은 몫을 차지하고 있음을 보여

주지만, 이보다는 더 적은 몫을 차지하고 있는 기능(?)들이 주목된다. 이들은 철자법과 구두점, 작문, 발음, 회화 등이다. 그러나 발음의 경우에는 발음기호 식별시험에 지나지 않고, 작문은 국문을 영역해 놓은 답지를 고르기이며, 회화의 경우도 간단한 표현을 글로 써 놓고 묻는 것에 지나지 않는 것이었다. 따라서 발음이니 회화니 하는 분류는 건강부회라는 말을 들을 만한 것이어서 고려의 대상이 될 수가 없다.

같은 시기에 나온 다른 연구는 똑같은 자료를 가지고 약간 다른 분석을 보여 주고 있는 바, hearing, speaking, writing, reading, grammar, vocabulary로 분류했다(이 봉삼, 1979). reading, grammar, vocabulary는 앞의 연구와 비슷하였으나, 발음을 hearing으로 speaking을 회화로 본 점이 특이하다. 이런 분류의 구체적 내용을 보면, hearing은 발음기호, 강세위치, 억양 등을 눈으로 알아보는 것이고; speaking은 dialogue, telephone call, tag question 등으로 이루어져 있으며; reading은 'sentence completion, E-K translation, rephrasing, comprehension, inversion 등이다. writing은 punctuation, spelling, K-E translation, rewriting, word-order 등이고; vocabulary는 synonyms, antonyms, word definitions, inflections, affixes, idioms 등이며; 마지막으로 grammar는 전통문법의 내용을 묻는 문제들이다. 여기서도 음성언어를 다룬 문제라고 하는 것들이 음성언어에 대하여 지식으로 알고 있는 바를 묻는 것들이었지 실제로 그 발음을 듣고 말할 수 있는 능력이 있는지의 여부를 묻는 것이 아니었다. 여기서도 앞의 문단에서 지적한 바와 마찬가지로 hearing이니 speaking이니로 부르는 것 자체가 어불성설이다.

고등학교 영어교육의 틀을 좌우해 온 대학입학 예비고사의 객관식 선다형 문제에 대한 사회적 비판이 끊임이 없자, 그 대안으로 나온 것이 이름이 자꾸 바뀌어도 내용은 마찬가지로 대학입학 학력고사에 주관식 문제라고 하여 단답형의 문제를 몇 개 포함시키기에 이르렀다. 1988년의 대학입학 학력고사의 예를 들면 전체 50문항 가운데 8문항이 이른 바 주관식이어서 30%를 차지한 셈이었다. 그러나 아직도 문제점이 많아서 비판이 끊이지 않았던 바, 언어운용능력의 측정을 가미한 시험이 되어야 한다는 주장들이었다(이 영식 1993). 대학입학 학력고사를 개선하기 위하여서는 분리시험을 통합시험으로 바꾸어야 하고, 속도시험(speed test)의 요소를 강조해야 하며, 단순한 언어지식을 묻는 문제를 지양하고 의사소통 능력을 측정하는 방향으로 나아가야 한다는 등의 주장들이었다. 이런 노력과 여론의 반영인지 대학입학 학력고사의 개선책이 발표되어 이제는 대학에 가서 수학할 수 있는 능력의 정도만을 측정하는 시험으로 바뀌었고, 그에 따라 시험형식도 대폭 바뀌어 1990년 12

월에 제1차 실험평가를 시작하여 1992년 11월 제7차 실험평가를 마지막으로 새로운 시험형식을 확정하였고, 1993년 7월 드디어 제1회 “대학 수학능력 고사”를 실시하게 되었다.

1993년부터 새롭게 실시된 대학 수학능력 고사의 특징은 듣기시험을 포함하고 있다는 것이다. 비록 말하기 시험은 포함되지 않았으나, 고등학교 교육에서 음성언어에 관심을 갖게 하는 계기가 되었다. 사실 1993년부터 실시하게 된 새로운 “대학 수학능력 고사”기 있기까지는 그 이전의 학력고사에 많은 문제점이 있었음을 시인하는 용기가 필요했던 것이다. 언어의 기능들을 고르게 측정하는 장치 자체가 미흡했던 것이 사실이었다. 다음의 인용에서 그 내용을 단적으로 보여주고 있다.

‘92년까지 실시되어 온 대학입학 학력고사의 영어시험의 경향은 언어의 네 가지 기능 중 듣기, 말하기, 및 쓰기의 세 기능은 전혀 무시한 채 오로지 ‘읽기’ 기능을 측정하기 위한 독해력과 문법에 편중되어 있었으며, 단편적으로 발음, 어휘, 생활영어의 회화체 문장이나 드물게 영작(통제작문)을 다루고 있다. (전 병만, 1994a)

물론 대학 수학능력 고사에서 이 모든 약점이 보완된 것은 아니지만, 듣기 시험문항의 삽입으로 적어도 음성언어에 대한 관심과 훈련을 유도해왔다는 점이 발전이라고 할 수 있겠다. 듣기문항의 신설은 바로 그런 분야의 연구를 유발하기 시작하면서도 의의가 있다고 하겠다 (차 경환, 1994). 대학입학 시험의 이런 변화는 시험의 성격과 지향점을 제시하려는 노력과 새로운 시험 모형의 실시결과로 오는 반향을 제시하여 차후의 정책에 반영하려는 노력이 나오게 했다 (권 오량, 1992, 1994). 따라서 이런 노력은 시행의 결과를 반성하는 계기를 주었다고 할 수 있다.

## 2. 고등학교에서의 평가 그리고 연합고사

중등학교에서 실시하는 영어능력의 평가는 물론이고, 고등학교 입학 을 위한 “연합고사”마저도 그것이 비록 중학교 영어수업과 직접적인 관련이 있음에도 불구하고 대학과는 직접적인 관계가 없어서인지 대학 교원들의 관심을 끌지 못하여 단 한 편의 논문도 나오지 아니했다. 중등학교에서의 영어능력 평가 및 연합고사에 대한 연구로 지금까지 밝혀진 것들은 교육대학원의 석사 논문이 두 편 있을 뿐이다.

중등학교의 영어능력 평가를 분석한 논문은 서울 시내 중고등학교 20개교

의 영어문제지를 수집하여 영어의 기능과 영어의 요소별로 문제들의 분포와 출제방법을 조사한 것이 하나 발견되었다 (한 성화, 1970). 중학교의 경우에, 기능별 문제의 분포를 보면 말하기 15%, 읽기 32%, 쓰기 54% 등으로 나타났고, 요소별로는 음운 또는 철자 12%, 문법구조 54%, 어휘 34%로 나타났다. 이 조사에서 문제가 되는 부분은 기능별 통계인 바, 말하기는 대화체 문장을 유도하는 문제를 잘못 이해한 것이고 쓰기는 단답형으로 지필로 써넣게 된 문항을 말하는 것이다. 따라서 이 논문에서 올바른 분석은 요소별 분석이다. 여기서도 음운문제는 “다음 밑줄 친 부분이 다른 셋과 발음이 다르게 나는 것은? — 1) Miss 2) Yes 3) has 4) sun”와 같거나, “맞는 ‘accent’의 위치를 골라라 — in-te-res-ting” 정도의 문제이다. 이것들은 이미 앞에서 지적한 바와 같이 올바른 답을 한 학생이 곧바로 올바른 발음을 할 수 있다는 보장이 없고, 학생의 입장에서는 늘 출제되는 문제를 눈으로 외었다가 맞추는 일이 있을 수도 있어서 제대로 된 ‘음운문제라고 할 수는 없을 것이다. 이 논문의 가치는 당시의 중등학교 영어시험의 현실을 반영하고 있는데 있다고 하겠다.

고등학교 입학에 위한 연합고사를 연구한 논문 역시 한 편뿐인 바 (김 진철, 1975), 각시도별로 실시되는 연합고사의 특수성을 감안하여 서울 부산 경기 충청 등의 네 지역에서 문제지를 입수하여 서울에 있는 네 학교에서 표준검사를 실시한 결과를 분석한 것이다. 이 연구에 의하면, 평가문항을 영역별로 나누어 볼 때, 문법 38.3%, 독해력 18.3%, 생활영어 16.5%, 발음 15%, 어휘 11.5% 등으로 나타났다. 이들 가운데 음성언어에 해당하는 발음의 범주에 발음기호, 강세, 쉬어 읽기, 액센트 등을 포함하고 있지만, 평가문항이 모두 선다형으로 되어 있어서 눈으로 보아 짐작하는 것이 될 수밖에 없다. 여기서 주목할 것은 고등학교 입학을 위한 연합고사에 진정한 의미의 음성영어는 전혀 출제되지 않았다는 점이다. 이는 청각구두 방법 (audio-lingual method)을 강조하여 실시하고 있던 당시의 형편으로 보아서는 의외의 사실이다. 특히 문법문제를 지나치게 사용한 것은 아직도 문법번역식 교수방법의 타성에 빠져 영어교육 자체가 영문법 교육에 치중하고 있었다는 증거일 수도 있다고 해석된다.

### 3. 대학원입학 영어시험과 영어교사 임용고시

우리 나라의 대학원 입학시험에 있어서 영어시험이 한 학생의 급락을 크게 좌우하는 요소로 인식되고 있는 실정에서, 평가도구의 타당도와 신뢰도가

높아야 할 것은 필수적이다. 그러나 대다수의 학교에서 영문국역만을 입시문제에 내놓고 있는 바, 이는 독해력을 측정하는 것 같으면서도 번역의 과정에서 오히려 국어의 구사능력을 요구하는 시험이어서 타당도가 낮고 신뢰도도 떨어진다고 볼 수 있다. 이런 점을 감안하여 고려대학교 대학원 입학시험 문제를 분석하여 새로운 대안으로 제시하는 노력을 보인 논문이 한 편이나 발표된 것은 다행이라고 할 수 있다 (김 충배, 1991).

대학원 입학시험과는 맥락이 다른 시험이지만 영어교사 임용고시에 관한 논점이 단 한 편이지만 눈에 뜨인다 (배 양서, 1993). 이 논문이 주장하는 바는 영어교사의 자격요건에 관한 것이지만 그 내용은 영어교사의 자격으로 영어능력과 문화의 이해를 내세우고 있어서 주목된다. 즉 “영어교사 후보자의 전공과목은 그들이 대학 재학시절에 무엇을 공부했던, 영어학이든, 영문학이든, 영어의 네 가지 기능과 영어권의 문화일 수밖에 없다.”는 주장이다. 영어교사를 채용하는 임용고시에서 후보자의 영어능력을 가늠하지는 않고, 엉뚱하게 국사나 교육학이니 하여 영어와는 아무런 관련이 없는 과목을 부과함으로써 영어능력은 썩 좋지 않은 사람이 국사나 교육학의 점수에 힘입어 영어에 능통한 사람을 밀어내고 채용되는 현실을 비판한 것이다. 그러나 구체적으로 영어능력 평가 자체에 대한 분석이나 비판이 없었던 것이 아쉽다.

## V. 대학의 교양영어

대학에서 신입생들에게 실시하는 교양영어대한 평가의 문제는 연세대에서 나온 보고서가 그 효시라고 볼 수 있는 바, 연세대에서는 1960년대 중반에 이미 음성언어시험과 지필 시험을 병행하여 실시하고 있었다 (이 정우, 1965). 그러나 좀더 본격적인 논의는 “제6회 한국 대학어학실험 연구회 보고서”에서 찾아 볼 수 있다. 이 보고서에서는 충남대학의 경우를 예로 들고 교양영어 과정에서 사용하는 평가의 내용을 분석했는데, 그 내용은 “구문, 어휘, 문법 구조, 철자법” 등으로 이루어져 있어서 구두기능과 문화에 대한 배려가 없음을 지적하고 있다 (박 상용, 1970). 여기서 주목할 것은 향후의 교양영어 평가에서 고려할 점을 권고한 내용이다. 그 권고의 내용은 (1) 언어학 이론에 근거한 테스트, (2) 대조분석에 의한 테스트, (3) 학습활동의 병행, (4) 구두 테스트의 첨가, (5) 읽기와 듣기 테스트의 포함 등이다. 이와 맥락을 같이하는 보고서가 바로 다음 해에 나왔는데 (신 상순과 조 명원, 1971), 평가의 내용을 (1) 듣고 내용파악하기, (2) 구문, (3) 읽고 내용파악하기, (4) 쓰기 등으로 잡

아서 말하기를 제외한 모든 분야의 평가를 시범적으로 실시하였다.

대학 교양영어에 관한 논문은 1980년대에도 나왔으나 개선을 주장하는 목소리만 높았지, 현장에서 개선되고 있었다는 조짐은 거의 보이지 않는다. 고려대에서 다양한 문제의 유형으로 평가를 실시하여, “어휘, 속어, 구문, 문장과 절의 의미, 사실적 세부내용, 중심사상과 결론, 추론 및 작가의 태도, 기타 발음” 등으로 나누었으나 실제에 있어서는 선다형 문제로서 구문과 읽기에 치중한 것이었다 (김 충배, 1988). 따라서 대학 교양영어의 개선을 부르짖는 의침은 끊이지 않아, 대학입학 시험을 포함하여 대학에서 실시하는 모든 영어시험에 대한 반성을 촉구하고, 특히 교양영어의 평가에 듣기와 발표력 그리고 쓰기를 첨가해야 한다고 주장한 논문이 발표되었음을 볼 수가 있다 (박형기, 1986).

## VI. 영미국의 표준화고사 소개

영국이나 미국에서 사용하고 있는 표준화시험을 소개하여 우리의 실정을 반성하고 개선의 모형으로 삼으려는 노력이 계속되었다. 맨 처음으로 소개된 시험은 미국 FSI (Foreign Service Institute)의 평가척도 (rating scale)의 소개로, “이런 것도 있다”는 정도의 가벼운 것이었다 (나 건석, 1971). 더욱 본격적인 소개는 대개 우리의 실정과 비교하는 것들이었는 바, 대학입학 예비고사와 TOEFL (Test of English as a Foreign Language)를 비교한 것 (이 봉삼, 1979)을 필두로 영국 UCLES (University of Cambridge Local Examinations Syndicate)에서 실시하는 CPE (Certificate of Proficiency in English)와 우리의 대학입학 예비고사를 비교한 상당히 자세한 소개가 나왔다 (이 봉삼, 1981). 한 발 더 나아가 외국의 시험들을 상호 비교하는 논문도 나오기 시작해서, TOEFL과 MTELP (Michigan Test of English Language Proficiency)를 상호비교하여 비판한 논문이 나왔고 (차 경환, 1988), 유창성 시험의 일환으로서 TOEFL, 영국의 IELTS (International English Language Testing System), 대학 수학능력 고사, 대학별 본고사 등으로 광범한 비교연구도 선을 보이고 있다 (전 병만, 1994). 영미국에서 사용하고 있는 이런 시험들을 소개하고 비교대조한 결과는 우리 학계에 평가의 새로운 계기가 되도록 하려는 의도가 엿보이지만, 실제 그 효과는 아직 확실하지 않다.

## VII. 결론

지난 30년 동안에 한국 영어교육계에 일어난 영어능력 평가의 이론과 실제의 발전과 변화는 실로 괄목할 만한 것이었다. 1960년대 중반까지만 해도 영어교육에 관한 논문이 거의 없다가 되었는데 1990년대에 들어서면서 영어평가에 관하여 발표된 논문들만도 대단한 숫자에 이르고 있다. 또한 논문의 내용과 질에 있어서도 장족의 발전을 거듭해 왔고, 논의의 대상이 다양해졌다. 더욱이 평가의 실제에 있어서도 단순한 이상론으로서의 외침에서 실제적 적용의 단계에 이르렀다. 논문의 내용 면을 보면, 초창기에 개념적이고 해설 또는 설명적이었던 논조에서 벗어나서 이제는 자기의 실험과 이론의 개발에 의한 주장의 개진으로 바뀌었음을 볼 수 있다. 평가도구의 형식도 전통적 형태에서 근대적 형태로 바뀌었다가, 이제는 실험적 차원의 노력을 보이는 경우가 가끔 눈에 뜨인다 (한 예로 1984년에 발표된 최 용재의 “반아쓰기 시험의 한 변형”을 들 수 있다). 더구나 평가의 이론과 실제의 발전이 어느 정도는 보조를 같이하는 단계에 접어들어 바람직한 현상으로 생각된다.

영어능력 평가의 이론은 분리시험과 통합시험 사이의 논쟁이지만, 구체적 시험형태에 이르면 백가쟁명이랄 수 있을 정도로 왕성한 주의주장이 있어 왔다고 볼 수 있다. 1960년대 중반까지만 해도, 영문국역식의 문제가 주류를 이루었던 것이 객관식 선다형 분리시험으로 바뀌었고, 다시 통합식 시험으로 그리고 주관식 시험으로 바뀌어 가는 경향을 보이고 있다. 특히 기능별 평가에 있어서는 읽기에 치중했던 초기의 시험형태에서 벗어나, 듣기가 첨가되고 말하기와 쓰기를 포함시키려는 시도가 두드러지게 나타나고 있다. 특히 영어에 관한 지식을 물어 정확도를 측정하던 시험에서 영어의 사용능력을 물어 유창성을 측정하는 경향을 보인 점은 커다란 변화라고 할 수 있다. 따라서 평가문항도 대화체의 것들이 많이 보여서 생활영어를 반영하는 비율이 높아졌다고 말할 수 있겠다.

영어능력 평가의 쓰임새는 학기마다 실시하는 성취시험으로 가장 많이 쓰이지만, 논문으로 발표될 만큼 관심을 끄는 것은 단연 선발고사이다. 특히 “대학입학 예비고사” 또는 “대학 수학능력 고사”와 “대학별 본고사”는 해당되는 학생과 학부모뿐만 아니라 전체 국민에게 초미의 관심을 갖게 하는 것이어서, 이들 시험의 출제경향이 곧바로 고등학교에서의 영어교육에 반영되어 역류현상이 심각할 정도이다. 대학입학 시험이 문법과 구문에 관한 지식을 묻는 문제 위주에서 영어사용 능력을 측정하려는 문제로 바뀌어짐에 따라 고등학교 영어교육도 그런 방향으로 옮겨 갈 것으로 기대된다. 선발고사 이

외에 논급되는 것으로는 대학교양영어가 있는데, 대학입학시험과는 달리 너무나 구태의연함을 탈피하지 못하고 있어서 지금도 개선을 부르짖는 소리가 높다.

한국 영어교육에 있어서 평가의 장래는 영미국의 이론과 모범을 차용하여 와서 새롭게 우리의 실정에 알맞게 수정보완하여 써 오고 있는 현실에서 우리 스스로 개발하려는 노력이 싹트고 있는 지금이 갈림길이다. 근년에 들어서 실력 있는 소장학자들이 많이 나와서 우수한 논문을 자주 발표하고 있으므로, 지금까지 쌓아 온 선배학자들의 경험을 바탕으로 하여 새로운 도약을 이룰 수 있는 가능성이 있다고 하겠다. 이들에게 주어진 과제는 지난 30년의 짧은 역사에 비추어 볼 때, 새로운 이론을 개발하여 측정분야를 활성화시키는 물론이고 영어의 의사소통능력을 측정하되 언어의 네 기능을 모두 망라할 수 있는 방안과 평가도구를 마련해야 한다. 또한 완전한 표준화검사의 마련으로 교육에 미치는 역류현상이 바람직한 것이 되도록 노력해야 할 것이다.

## 참고 문헌

- 강 순경. (1978). Objective testing of English as a foreign language. 논문집 6 (경기대학교), 251-264.
- 곽 경수. (1977). Test of aural comprehension의 오답분석. *영어교육*, 14, 87-97.
- 권 오량. (1983). 좋은 영어고사의 기술적 고려사항들. *영어교육*, 26, 29-49.
- 권 오량. (1992). 대학수학능력시험 외국어(영어)영역의 성격 규정 및 실험 평가에서 드러난 몇가지 문제점. *이재호 교수 정년기념논문집*. (pp. 221-248). 육군사관학교 영어과.
- 권 오량. (1994). 대학수학능력시험이라는 혁신에 대한 고교 영어 교사의 관심과 요구: 관심에 기초한 채택 모형에 의한 분석. *영어교육*, 48, 265-290.
- 김 덕기. (1980). 대학입시영어의 검토. *영어교육*, 19, 39-51.
- 김 순복. (1982). An experimental cloze test. 논문집 9 (세종대학교), 57-67.
- 김 순복. (1983). The use of cloze procedure for measuring correlation between first and second language proficiency. *영어교육*, 26, 87-106.
- 김 순신. (1994). Cloze기법에 의한 영어 숙달도 측정 —교양영어 교과서의 난이도 분석을 통해서—. *영어교육*, 48, 127-150.

- 김 영서. (1993). 구어면담 평가의 신뢰도. *영어교육*, 46, 137-163.
- 김 용석. (1980). Testing reading comprehension: Multiple-choice v. short-answer items. *육사논문집*, 20, 115-130.
- 김 진철. (1975). 고입연합고사 영어문제의 분석적 연구. *고려대학교 교육대학원 석사논문*.
- 김 충배. (1988). 교양영어 강독시험과 평가의 방향. *영어교육*, 35, 33-82.
- 김 충배. (1991). 대학원 입학 영어시험의 발전 방향. *영어교육*, 41, 51-80.
- 나 건석. (1971). FSI test 소개. 제7회 한국대학어학 실험연구 협의회 보고서. 67-72.
- 대학입시에 대한 건의문. (1980). *영어교육*, 20, 155-156.
- 대학입시에 대한 건의문 (회신). (1980). *영어교육*, 20, 157.
- 문 정일. (1985). Aural/oral tests. *영어교육*, 29-30, 451-468.
- 민 대식. (1985). The cloze procedure as a test of overall EFL proficiency. *영어교육*, 29-30, 395-408.
- 박 상용. (1965). Applied linguistics in the classroom and in English test. *논문집: 인문사회과학편 4 (충남대학교)*. 41-55.
- 박 상용. (1970). Reconsideration of college English testing. 제6회 한국대학어학실험 연구협의회 보고서. 1-13.
- 박 상용. (1971). An analysis of college entrance English examination. 제8회 한국 대학어학실험연구협의회 보고서. 73-93.
- 박 약우. (1990). 영어듣기능력평가의 문항분석과 듣기지도 방안. *영어교육*, 39, 29-75.
- 박 원. (1985). Item analysis of integrative testing in English. *영어교육*, 29-30, 431-450.
- 박 형기. (1986). A critical review of testing English as a foreign language at university level in Korea: Viewed from accepted theories and practices. *영어교육*, 31, 119-134.
- 배 양서. (1993). 영어교사의 자격과 교사 임용고시. *영어교육*, 45, 3-11.
- 서울대학교 어학연구소. (1965). 영어교육측정—서울대학교 어학연구소 작성 영어학력 표준화 검사의 결과보고. *언어교육 1: 1* (서울대학교 어학연구소), 91-110.
- 성 찬경, 장 병길, 김 영일, 최 용재. (1989). 한국영어영문학연구사 1945-1965. *인문과학 19* (성균관대학교), 67-92.
- 신 상순, 조 명원. (1971). 영어교육에 관한 실험평가. 제7회 한국 대학어학

- 실험연구 협의회 보고서, 9-40.
- 신 용진, 전 병만. (1983). Cloze test as a test of integrated skills in Korean university students. *어학 10* (전북대학교 어학연구소), 1-25.
- 신 홍식. (1976). 영어교육 평가방법 개선방안에 관한 연구. *제주간호전문학교 논문집 2*, 7-28.
- 오 석봉. (1992). Test of oral proficiency of adult EFL learners: Focus on implications of adult EFL learner's strategic competence in their oral proficiency. *영어교육*, 43, 173-194.
- 용 성해. (1977). 영어 test에 대한 연구. *논문집 21* (대전공업전문학교), 91-101.
- 이 봉삼. (1979). 한국의 대입예비고사와 미국의 대입자격 영어문제에 관한 비교연구. *Atheneum 2* (단국대학교 영어영문학회), 37-55.
- 이 봉삼. (1981). 영국 대학입학자격 영어시험문제의 분석. *논문집 15* (단국대학교), 185-219.
- 이 영식. (1993). 현행 영어시험에서의 언어운용능력의 도입. *영어교육*, 45, 45-92.
- 이 옥남. (1971). 문법구조를 중심으로 한 영어 테스트. *어학교육 3* (전남대학교 어학연구소), 45-68.
- 이 용실. (1982). 방송망을 통한 청취력 시험의 가능성. *영어교육*, 23, 269-273.
- 이 정우. (1965). 영어성적고사의 실제. *영어수업과 어학실습 1* (연세대학교 교양학부), 73-77.
- 이 종찬. (1972). Language testing에 있어서 기계이용. *영어교육*, 9, 37-40.
- 이 회승. (1984). 구한말의 영어교육에 관한 회고. *영어교육*, 27, 1-5.
- 임 호순. (1983). 언어 테스트: Cloze tests의 신뢰도. *논문집: 인문사회과학편 25* (충북대학교), 25-35.
- 전 병만. (1985). 규칙빈칸매우기 절차에 대한 연구. 전남대학교 박사학위 논문.
- 전 병만. (1994a). 대학입학시험에 있어서 영어시험의 바람직한 방향과 모형 개발에 대한 연구. *영어영문학*, 17 (한국영어영문학회 전북지구) 1-20.
- 전 병만. (1994b). The model of communicative competence and communicative language testing. *영어영문학*, 17 (한국영어영문학회 전북지구), 21-46.
- 전 상범. (1978). 예비고사(영어)의 문제점과 개선방안. *영어교육*, 15, 103-111.
- 전 형국. (1966). 입시에 반영된 영어교육의 단면을 고찰함. *인문과학*, 14-15

- (연세대학교 인문과학연구소), 149-173.
- 차 경환. (1988). Test review: TOEFL vs. MTELP. *영어교육*, 36, 75-83.
- 차 경환. (1994). 대학수학능력시험 영어 듣기문항 분석. *영어교육*, 47, 49-75.
- 차 경환, 이 경민. (1993). Task-based 영어청취력 교재분석을 통한 청취력 문항연구. *영어교육*, 45, 13-43.
- 최 용재. (1974). 언어평가 방법의 부정적 측면. *사대논문집 5* (조선대학교), 119-126.
- 최 용재. (1981). Cloze vs. supply type tests. *Dongguk Review*, 8-9, 155-166.
- 최 용재. (1982). 언어자원이론과 언어평가. *영어교육*, 24, 21-31.
- 최 용재. (1984). 받아쓰기 시험의 한 변형. *영어영문학*, 30(3), 627-639.
- 최 용재. (1990). 언어능력의 측정가능성에 대하여. *논문집*, 29 (동국대학교), 27-58.
- 최 용재. (1992). Test와 Exam, 그리고 그 문항에 관하여. *Dongguk Review*, 20-21, 133-147.
- 최 용재. (1993a). 규칙빈칸메우기 시험의 문항수에 관한 연구. *어학연구*, 29 (서울대학교 어학연구소), 137-155.
- 최 용재. (1993b). 규칙빈칸메우기 시험의 재평가. *논문집*, 32 (동국대학교), 25-50.
- 최 인철. (1989). 언어능력 측정 및 평가이론의 과거, 현재, 미래. *영어교육*, 38, 95-135.
- 최 인철. (1994). Content and construct validation of a criterion-referenced English proficiency test. *영어교육*, 48, 311-348.
- 한 성화. (1970). 언어평가에 관한 연구. 이화여자대학교 교육대학원 석사학위 논문.
- 홍 성면. (1984). Listening skills와 Bilingual education. *영어교육*, 27, 49-90.
- 홍 성면. (1990). 영어 테스트와 세계의 언어교육. *영어교육*, 40, 93-114.
- Fernandes, Mark. (1987). An oral test using expansion drills. *English Teaching*, 33, 167-178.