

억제와재인 읽기폭 검사*

Suppression and Recognition Reading Span Test

이 병 택** 이 경 민*** 김 정 오**** 홍 재 성*****
(Byeong-taek Lee) (Kyoung-Min Lee) (Jung-Oh Kim) (Chai-Song Hong)

요약 본 연구의 목적은 회상검사를 기반으로 하는 기존의 읽기폭 검사(Daneman & Carpenter, 1980)를재인검사의 형태로 재구성하여 피검사자의 읽기폭을 측정하는 것이었다.재인 읽기폭 검사의 공존타당도를 확인하기 위해 검사를 구성하는 방해검사단어의 성질을 달리하며 기존의 회상검사와 새로운재인검사간의 상관을 측정하는 일련의 실험이 수행되었다. 검사단어와 무관한 단어들로 방해 검사단어가 구성된 실험 1에서는 기존의 회상검사와재인검사간의 상관이 매우 낮았다. 실험 2에서는 검사단어의 선택을 간섭하도록 구성된 방해검사단어를 포함하였으나 역시 두 검사유형간의 상관이 비교적 낮았다. 실험 3에서는 이전의 시행에서 적절했던 검사단어를 방해검사단어로 사용하였으며, 회상검사와재인검사간의 상관이 유의하게 높았다. 본 연구의 일련의 실험을 통해 얻은 결과는 기존의 읽기폭 검사의 타당도에 관한 시사점을 주는 것 뿐만 아니라, 작업기억을 주제로 한 대규모의 개인차 연구에 필요한 도구를 제공한다는 점에서 의의가 있다.

주제어 읽기폭 검사, 작업기억, 간섭, 회상검사,재인검사, 공존타당도, 구성타당도

Abstract The aim of this study is to make the recall-based reading span test(Daneman & Carpenter, 1980) into the recognition-based test for on-line measuring the capacity of the participant. In order to measure the concurrent validity, a series of experiments is performed with varying features of distractors consisting of the reading span test. In experiment 1, which included irrelevant words as distractor, low correlation was observed. And in experiment 2, including several types of distractors which interfere with the selection of target words, low correlation was observed too. But in experiment 3, including distractors no more relevant in the present trial but relevant in previous trial, high correlation was observed. The results of this study have theoretical implications on the validity of the reading span test and practical implication in that this study provides the tool for the studies on individual differences in working memory capacity.

1. 서론

전통적으로 심리학에서는 사람들의 인지처리의 보편성에 관심을 기울인 반면, 사람들간의 개인차는 측정상

의 오차로 귀인하거나 혹은 최소화하는데 주된 관심을 두었다. 그러나 연구가 더욱 정교해짐에 따라 개인차를 무시하기 어렵게 되었고, 개인차를 통해 보편성을 탐구하려는 시도마저 등장하기에 이르렀다. 한편, 사람들의 인지기능을 설명하기 위해 가장 자주 동원되는 개념이 작업기억이다. Baddeley(1986)에 따르면 작업기억은 처리와 저장을 위한 작업공간이다. 많은 연구자들은 여러 인지처리과정에서의 작업기억의 역할을 밝히고자 하였다. 이런 점에서 오늘날 인지적 처리과정에서의 개인차를 다루는 많은 연구들이 작업기억에서의 개인차를 가정하는 것은 우연이 아니다.

인지기능에서의 개인차를 다루는 대부분의 연구들은 개인차를 측정하는 방법으로 작업기억 용량을 측정하였

** 서울대학교 의학과 신경과학교실
*** 서울대학교 심리학과
**** 서울대학교 불어불문학과
† 서울대학교 인지과학 협동과정
연구세부분야 : 한국어정보처리
키워드: 읽기폭 검사, 작업기억, 간섭, 회상검사,재인검사
주소 : 151-742 서울시 관악구 신림동 산 56-1
서울대학교 인문대학 불어불문학과 홍재성 교수
전화 : 02-880-6123 FAX : 02-887-2493
E-mail : cshong@snu.ac.kr
* 이 연구는 서울대학교 인지과학 연구소를 통하여 서울대학교 간접연구경비 지원에 의해 수행되었음.

다. 작업기억 용량을 측정하는 다양한 방법들이 개발되어 왔으며(Daneman & Carpenter, 1980), 각 측정도구의 타당도와 측정도구들간의 상관에 관한 연구도 많이 이루어졌다(개관을 위해서는 Daneman & Merikle(1996)과 Shah & Miyake(1996)를 보라).

작업기억 용량을 측정하는 가장 대표적인 검사로 Daneman Carpenter(1980)의 읽기폭 검사를 들 수 있다. 읽기폭 검사란 Daneman과 Carpenter에 의해 언어이해측면의 작업기억 용량을 측정할 목적으로 고안된 검사이다. 이 검사에서 피검사자들은 하나씩 제시되는 문장을 소리내어 읽은 후, 각 문장의 마지막 단어를 회상해야 한다. 그러므로 이 검사는 기본적으로 회상검사라고 할 수 있다.

그런데 읽기폭 검사를 이용한 연구들을 비롯한 대부분의 개인차 연구들은 실험실 내에서 소규모로 이루어졌으며, 따라서 대개의 실험실 연구가 갖는 일반화의 문제를 안고 있다고 할 수 있다. 이들 연구들이 실험실 내에서 소규모로 이루어질 수밖에 없었던 것은 검사의 절차와 밀접한 관련이 있다. 읽기폭 검사가 그러하듯이, 대부분의 작업기억 검사가 회상검사이기 때문에 참가자의 작업기억을 검사하기 위해서 반드시 검사자가 함께 참가하여 그 결과를 기록하여야 하며, 이러한 절차의 진행이나 결과의 기록을 위해서 반드시 숙련된 검사자를 필요로 한다. 물론, 회상 검사라 할 지라도 지필 검사인 경우에는 검사자가 반드시 있을 필요는 없지만, 제시된 문장을 소리내어 읽고 마지막 단어를 기억해야 하는 읽기폭 검사의 속성상 검사자의 존재는 필수적이라고 할 수 있다. 이와 같이 읽기폭 검사가 절차상의 특성상 갖는 시간적 공간적인 제약은 비단 읽기폭 검사에만 국한되는 것이 아니고, 대부분의 작업기억 검사에 공통적이라고 할 수 있다. 따라서 기존의 실험실에서 사용되어온 대부분의 작업기억 검사는 대규모의 연구에 적절치 않게끔 구성되어 있다고 할 수 있다.

본 연구의 목적은 기존의 읽기폭 검사가 갖는 절차상의 문제를 극복하여 검사자 없이 피검사자가 단독으로 작업기억 용량을 측정할 수 있는 도구를 개발하는 것이었다. 특히 언어이해능력과 관련된 작업기억 용량을 측정하기 위해 읽기폭 검사를 기본으로 하여 이를 수정하여 개발하고자 하였다. 이와 같은 목적을 위해 크게 두 가지의 문제가 해결되어야 했다.

첫째, 작업기억의 용량을 반영해야 한다는 기존 읽기폭 검사의 목적을 유지하면서, 동시에 기존 검사와 달리 피검사자가 단독으로 검사를 수행할 수 있도록 절차를 수정하는 것이었다. 검사자 없이 피검사자 단독으로

검사를 진행할 수 있도록 절차를 수정하는 과정에서 몇 가지 문제가 해결되어야 했고, 이 문제의 해결이 본 연구에서 핵심적인 과제가 되었다.

즉, 기존의 검사는 검사자가 검사를 진행하기 때문에 피검사자가 문장을 정확히 읽고 있는 지를 확인하는 것이 용이하였다. 그러나 피검사자 단독으로 수행하는 검사에서는 피검사자가 문장을 정확히 읽는지, 그리고 마지막 단어에만 주목하고 있는지 등의 여부를 알기 어렵기 때문에, 피검사자의 수행을 감찰할 새로운 절차가 추가되어야 했다.

둘째, 심리검사라 한다면 갖추어야 하는 요건 중, 타당도의 문제가 매우 중요하다. 기존의 검사는 작업기억 용량을 측정하고 있다는 점에서 어느 정도 그 타당도를 이미 인정받은 검사라고 할 수 있다. 그러나 새로 구성되는 읽기폭 검사는 그 절차가 기존의 읽기폭 검사와 다를수록 타당도를 다시 확보하는 것이 필요하다. 기존의 검사를 수정하여 새로운 검사를 구성할 때 흔히 타당도를 확인하는 방법 가운데 하나가 기존 검사와의 상관을 확인하는 공존타당도이다.

본 연구에서는 새로 구성되는 검사와 기존의 읽기폭 검사의 상관을 측정하여 공존타당도를 확인하고자 하였다. 즉, 기존의 읽기폭 검사와 수정된 읽기폭 검사간의 상관이 높다면, 우리는 수정된 읽기폭 검사 역시 작업기억 용량을 측정하고 있다고 말할 수 있을 것이다. 이를 위해 일련의 실험을 통해 기존의 검사와 높은 상관을 갖는 새로운 검사를 개발하고자 하였다.

2. 전체 방법

본 연구에서는 방해단어의 성질을 달리하면서 세 실험을 실시하였다. 그러나 각 실험에서 새로운 재인 읽기폭 검사의 재료와 절차는 기본적으로 동일하였으며, 다만 방해검사단어의 성질과 수만을 변화시켰다. 또한 새로운 검사의 구체적인 절차는 기본적으로 기존의 읽기폭 검사의 절차와 동일하였다. 즉, 피검사자들은 화면에 하나씩 제시되는 문장을 소리내어 읽고, 읽은 문장의 마지막 어절을 기억하도록 지시받았다. 그 후, 화면에 문장 마지막 어절에 속했던 단어와 방해단어들이 제시되면, 피검사자는 그 중 자신이 기억했던 단어들을 선택해야 했다.

표적단어의 선택을 끝낸 후, 화면에는 다시 하나의 문장이 제시되는데, 피검사자는 이 문장이 그 시행에서 읽었던 문장들의 의미에 비추어 옳은 지의 여부를 판단하여야 했다. 이 절차는 기존의 회상 읽기폭 검사에는

없는 부분이다. 기존의 검사에서는 시행내의 문장에 대한 이해여부를 묻는 절차의 필요성이 적는데, 왜냐하면 피검사자가 문장을 소리내어 읽고 이를 연구자가 함께 참여하고 있기 때문에 확인할 수 있기 때문이다.

그러나 본 연구에서 새로 구성된 재인검사는 기본적으로 검사자의 참여없이 피검사자 단독으로 실시하기 때문에, 피검사자가 문장을 제대로 읽었는 지의 여부를 확인할 수 있어야 하고, 따라서 시행 내에서 읽은 문장의 이해여부를 묻는 절차를 삽입하게 되었다. 검증 문장의 옳고 그름의 여부는 가능한 한 연속되는 두 문장 조건의 열 시행에 걸쳐 절반씩으로 배분되도록 하였다.

마지막으로 검사가 진행되어감에 따라 한 시행에서 제시되는 문장의 수는 2개에서 6개까지 증가하였다. 기존의 읽기폭 검사는 원래 6개 문장 조건까지 구성되었으나, 대개의 경우 5개 문장 조건까지만 실시하는 것이 일반적이다. 그러나 기본적으로 회상의 단서를 명시적으로 제시하는 재인검사는 회상검사에 비해 쉬운 것이 일반적이기 때문에, 6개 문장조건까지 실시하였다.

3. 실험 1

실험 1의 목적은 기존의 읽기폭 검사의 골격을 그대로 유지한 채, 읽기폭 검사가 측정하고자 하는 바를 그대로 측정할 수 있도록 검사를 재인검사로 수정하여, 수정된 검사가 기존의 검사와 높은 상관관을 갖는 지를 확인하는 것이었다. 기존의 읽기폭 검사는 이미 다양한 종류의 언어이해처리과제에서 개인차를 변별함을 보여주어 그 타당성을 인정받은 검사이다(김경중, 1998; 안중찬, 1999; 이병택, 1995, 2002; 이병택, 김경중, 조명환, 1995; 이선화, 2001; Just & Carpenter, 1992; King Just, 1991). 따라서 기존의 회상검사와 새로 구성되는 재인검사간의 상관관이 높다면, 새로운 검사의 공존 타당도가 높다고 할 수 있을 것이다.

3.1 방법

참가자 서울대학교에서 심리학개론을 수강하는 학생 60명이 수업의 일환으로 참가하였다. 이들 가운데 외국인 혹은 이중 언어자는 없었다.

재료 읽기폭 검사는 Daneman과 Carpenter(1980)의 읽기폭 검사를 한국어로 재구성한 이병택(1995)의 검사가 사용되었다. 재인검사는 기본적으로 이병택의 읽기폭 검사를 기본으로 구성되었으나, 검사에 사용되는 문장은 모두 새로 구성되었다. 그리고 선택과제에는 표적단어와 방해검사단어를 포함하였는데, 표적단어는 피검사자

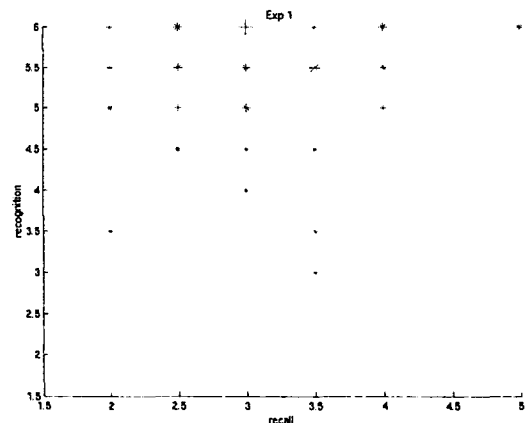
가 그 시행에서 읽은 문장의 마지막 어절의 단어였으며, 표적단어와 동수로 포함된 방해검사단어는 표적단어의 의미와 무관할 뿐만 아니라 그 시행에 포함된 문장에도 속하지 않는 단어들이었다.

예를 들어, 세 문장 조건의 한 시행에서 피검사자들이 표적단어를 선택해야 할 때, 이들은 표적단어 세 개, 그리고 방해검사단어 세 개로, 모두 6개의 단어들을 제시받았다. 검사점수는 세 개의 표적단어를 모두 정확하게 선택한 후, 검증문장에 대해 정확하게 답한 시행만을 옳은 시행으로 간주하였다. 검사점수의 채점 방법은 Daneman과 Carpenter의 규칙을 그대로 유지하였다.

마지막으로 재인 검사를 시작하기 전, 두 문장 조건 2회와 세 문장 조건 2회의 연습시행이 있었다. 연습시행이 끝난 후, 절차에 관한 질문이 없으면 곧 이어 본 시행을 시작하였다. 피검사자들은 기존의 읽기폭 검사와 새로이 재인검사로 구성된 읽기폭 검사를 순차적으로 수행하였다. 두 가지 검사를 실시한 후, 실험에 대한 설명이 이루어졌으며, 모든 실험의 과정을 수행하는데 소요된 시간은 약 25분이었다.

3.2 결과 및 논의

실험 1의 결과는 기존의 검사와의 상관관이 0.066에 불과하여 매우 낮았다($F(1, 58) < 1$). 이러한 결과는 기존 회상검사의 평균점수가 3.03임에 비해, 새로운 재인 검사에서는 5.43에 이르러, 대부분의 피검사자들이 새로운 검사에서 정확히 검사단어들을 선택할 수 있었기 때문으로 보인다. (그림 1)에서도 볼 수 있듯이, 기존 검사상에서 낮은 점수를 받은 피검사자와 높은 점수를 얻은 피검사자가 모두 새로운 검사에서 높은 점수를 얻



(그림 1) 실험 1의 결과 산포도 (*의 크기는 사제수에 비례)

있음을 알 수 있다. 이러한 결과는 새로운 검사의 난이도가 너무 낮았기 때문에, 이로 인해 천장효과(ceiling effect)가 나타난 것으로 해석되어질 수 있다. 이러한 천장효과는 두 검사간의 상관을 낮추는 주요한 요인으로 작용했을 것이다.

4. 실험 2

실험 1에서 노출된 난이도 문제를 해결하기 위해 실험 2가 구성되었다. 실험 1의 재인검사가 난이도가 낮았던 이유는 방해검사단어가 피검사자의 표적단어 선택을 간섭하는데 충분치 못했기 때문이라고 할 수 있다. 본 실험에서는 방해검사단어의 두 가지 측면을 변화시켜서 피검사자들의 적절검사단어 선택에 있어서 간섭을 증가시키고자 하였다. 이를 위해 기억해야 하는 검사단어와 함께 제시되는 방해단어의 수를 증가시키고자하였다.

특히 방해검사단어가 표적단어와 의미적으로 유관한 경우, 피검사자는 작업기억 상에 활성화된 의미를 의미적으로 매우 정교하게 부호화하지 못했다면, 기억하지 않았지만 기억한 것으로 잘못 판단할 가능성이 높을 것이다. 한편 시행 내 문장에 속한 단어가 다시 제시되었을 때도, 문장의 마지막 어절에 대해 적절히 부호화하지 못하고, 제시된 검사단어들을 친숙성에 기초해서 판단한다면 표적단어로 잘못 선택할 가능성이 증가할 것이다. 전체적으로 이러한 조작은 피검사자들의 표적단어 선택을 많이 간섭할 것이고, 따라서 피검사자들의 재인검사상에서의 점수가 낮아져, 천장효과로 인한 낮은 상관을 회피하는 것이 가능해질 것이다.

4.1 방법

참가자 서울대학교에서 심리학개론을 수강하는 학부생 32명이 수업의 일환으로 실험에 참가하였다. 이들 중, 이중언어자 2명은 실험에 제외하였다.

재료 피검사자들이 읽는 문장은 실험 1의 재료를 그대로 사용하였다. 그러나 검사단어의 수가 실험 1에 비해 두 배로 증가하였다. 이것은 방해검사단어의 수가 실험 1에서는 표적단어와 동수였던 것에 반해, 본 실험에서는 표적단어의 세 배였기 때문이다. 방해검사단어 가운데 1/3은 실험 1에서와 같이 표적단어나 문장과 무관한 단어였으며, 다른 1/3은 표적단어와 의미적으로 유관한 단어였고, 나머지 1/3은 시행 내 문장에 속했던 마지막 위치의 단어를 제외한 다른 위치에 속한 단어들이었다.

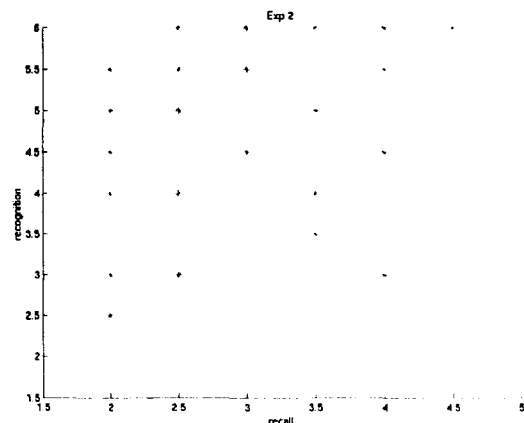
네 문장 조건의 한 시행을 예로 들자면, 피검사자는 표적단어를 선택해야하는 단계에서 표적단어 네 개, 그

리고 방해검사단어 열 두 개를 제시받는다. 열 두 개의 방해검사단어 가운데 네 개는 실험 1에서와 같이 표적단어나 문장의미와 무관한 단어들이었으며, 다른 네 개의 방해검사단어는 표적단어와 의미적으로 유관한 단어를 사용하였다. 그리고 나머지 네 개의 방해검사단어는 읽은 문장 내의 다른 위치에 속한 단어를 사용하였다.

4.2 결과 및 논의

실험 2에서 관찰된 회상검사와 재인검사간의 상관은 0.245 ($F(1, 28) = 1.79$)로, 실험 1에 비해서는 다소 높은 상관을 보였지만, 여전히 비교적 낮은 상관을 보였다. 그러나 기존 회상검사의 평균이 2.97로 실험 1의 결과와 대동소이했음에 비해 재인검사의 평균이 4.63으로 실험 1에 비해 크게 낮아졌다. 이 결과는 본 실험에서 처음에 의도했던 난이도의 조절에는 부분적으로 성공적이었음을 보여준다.

그럼에도 불구하고 기대했던 높은 상관을 관찰하지 못했던 것은 여전히 기존 검사상에서 각 점수 범위의 피검사자들이 새로운 과제에서 비교적 넓은 범위에 걸쳐 골고루 점수를 획득하고 있기 때문으로 보인다(그림 2 참조). 즉, 실험 2의 재인검사에서 사용된 방해검사단어들이 피검사자들로 하여금 표적단어의 선택을 방해하는데는 성공적이었지만, 이것이 회상검사에서 낮은 점수를 받은 집단(저폭집단)과 높은 점수를 받은 집단(고폭집단)을 변별하는데 충분치 못했음을 시사한다.



(그림 2) 실험 2의 결과 산포도 (*크기는 사례수에 비례)

5. 실험 3

실험 3에서는 기존 검사에서 낮은 점수를 받은 집단과 높은 점수를 받은 집단을 잘 변별할 수 있도록 방해 검사단어의 성질을 달리 조작하였다. 새로운 방해 검사단어의 성질은 기존의 작업기억 개념의 정의보다는 근래 새롭게 주창되는 작업기억의 정의에 따르도록 구성된 것이다.

Baddeley(1986)가 저장과 처리를 동시에 강조하며 작업기억 개념을 정리하였지만, 근래 그는 작업기억의 중앙집행부의 중요성을 강조하면서 처리를 더욱 중요시하였다. Baddeley(1993)는 작업기억의 '가장 중요한 성분인 중앙집행부가 저장기 아니라 주의와 협응에 관여하기'(p. 168)이기 때문에 '작업주의(working attention)'라 이름하였어야 옳다고 다시 평가하고 있다. 이러한 논의에서 우리는 최근의 연구들이 저장과 처리를 동시에 강조하는 작업기억 개념에서 벗어나 처리를 강조하고 있음을 알 수 있다.

그렇다면 작업기억의 중앙집행부의 처리기능이란 어떤 것인가? Cantor와 Engle(1993)에 따르면 작업기억의 중앙집행부는 더 이상 적절치 않은 정보를 억제하여 적절한 정보에 더욱 많은 용량을 가용케 하는 것이다. Hasher와 Zacks(1988)에 따르면 한정된 작업기억 공간에 활성화되어 있는 정보들 가운데, 부적절한 정보들을 억제함으로써 더 많은 적절한 정보들이 활성화 상태를 유지할 수 있다. 따라서 작업기억 용량이 큰 사람들은 과거에는 적절했지만 더 이상 적절치 않은 정보를 잘 억제할 수 있는 사람이며, 작업기억 용량이 작은 사람들은 부적절 정보를 잘 억제하지 못하여 적절한 정보의 처리에 용량을 제대로 할당할 수 없는 사람들이다.

만약 본 실험의 재인과제에서 방해검사단어가 이전에는 적절했으나 더 이상 현재 과제 수행에 부적절해진 단어라면, 기존의 회상검사에서 얻어진 점수와 새로운 재인과제에서 얻어진 점수간의 높은 상관을 기대할 수 있을 것이다.

이러한 측면에서 과거에는 과제수행에 적절했으나 더 이상 적절치 않은 정보란 이전 시행의 표적단어가 대표적인 것이다. 즉, 이전 시행에서는 기억해야 했지만, 현재의 시행에서는 억제되거나 혹은 무시되어야 한다. 그렇지 않고 이전 시행에서의 표적단어를 계속 기억하고 있다면 현재 시행에서 새로운 문장의 마지막 단어들을 기억하는데 방해가 될 것이기 때문이다.

Gernsbacher, Varner 및 Faust(1990)의 연구에서 알 듯이, 미숙련 이해집단 혹은 저폭 집단의 특징은 부

적절한 정보를 잘 억제하지 못하는 것이다. 즉, 저폭 집단은 이전 시행에서 적절했던 의미들이 후속 시행에서도 높은 활성화 상태를 유지하고 있을 것이고, 높은 활성화 상태의 표적검사단어들은 표적단어들의 선택을 방해할 것이다. 따라서 저폭은 문장의 수가 증가할수록 방해검사단어들을 더욱 많이 선택하게 될 것이고, 따라서 재인검사 점수가 비교적 낮을 것이다. 그러나 고품 집단은 이전 시행에서는 적절했으나 현시행에서는 더 이상 적절치 않은 단어 의미를 잘 억제할 수 있기 때문에, 방해단어로 이전 시행의 검사단어들이 제시되어도 잘 기억할 수 있을 것이고, 더 많은 문장 조건에서도 많은 표적단어들을 적절히 선택할 수 있을 것이다. 위와 같이 저폭 집단과 고품 집단간의 수행의 차이가 드러난다면 회상검사와 재인검사간의 높은 상관을 관찰할 수 있을 것이다.

5.1 방법

참가자 심리학개론을 수강하는 서울대학교 학생 40명이 수업의 일환으로 실험에 지원하였다. 이들 가운데 3명의 학생은 이중언어자로 실험에 참가하지 않았으며, 37명의 학생이 실험에 참가하였다.

재료 실험 3의 문장은 실험 1과 동일하였다. 또한 실험 2에서와 같이 방해검사단어를 검사단어의 3배수로 그대로 유지하되, 방해검사단어의 1/3은 직전 시행의 표적단어들을, 그리고 다른 1/3은 두 시행 전의 표적단어를 방해검사단어로 사용하였다. 나머지 1/3의 방해검사단어는 실험 1에서와 같이 검사단어나 혹은 그 시행에서 제시된 문장과 어떠한 의미적 관련을 갖지 않는 단어가 사용되었다.

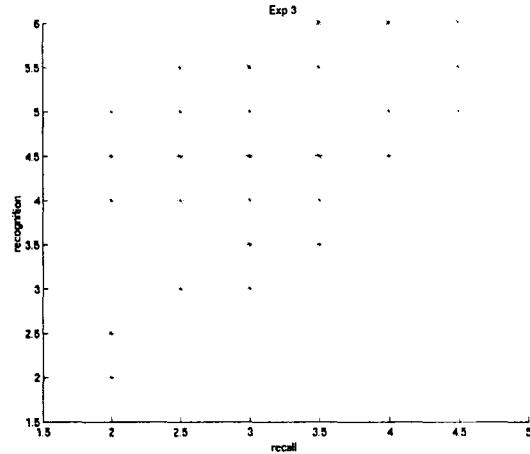
방해검사단어 구성과정의 한 가지 문제는 각 문장 조건의 첫 시행과 둘째 시행에서 방해검사단어를 제대로 구성할 수 없다는 점이었다. 세 문장 조건의 첫 시행을 예로 들면, 직전 시행과 그 이전의 시행은 두 문장 조건으로, 방해검사단어로 각기 두 개씩 밖에 사용할 수가 없었다. 이와 같은 경우에는 부족한 방해검사단어로 표적단어나 문장의 의미와 무관한 단어를 사용하였다. 두 문장 조건의 첫 두 시행에 사용된 방해검사단어는 이전의 연습시행에서 사용된 표적단어들이었다.

5.2 결과

본 실험에서 참가자들의 기존 회상검사의 평균 점수는 3.09였으며, 이 결과는 표 1에서 볼 수 있듯이, 실험 1과 실험 2에서 관찰된 회상검사의 평균과 비슷하다. 그리고 재인 검사의 평균은 4.58로 실험 2의 결과와 대동

소이하였다. 더 중요한 사실은 본 실험에서 실시된 두 검사간의 상관인 0.54로 비교적 높았다는 점이다. 이 결과는 통계적으로도 유의하였다($F(1, 35) = 14.37, p < .001$).

두 검사간의 높은 상관은 실험 1과 실험 2의 산포도인 그림 1이나 그림 2와 비교했을 때, 그림 3에서 보여지는 선형적인 관계를 통해 다시 확인할 수 있다. 즉, 회상검사에서 낮은 점수를 받은 저폭은 재인검사에서도 비교적 낮은 점수를 얻은 것에 반해, 회상검사에서 높은 점수를 받은 고향은 재인검사에서 더욱 높은 점수를 얻은 것으로 보인다. 이러한 결과는 저폭에 비해 고향이 이전 시행에서 적절했던 표적단어를 현재의 시행에서 더욱 잘 억제할 수 있었기 때문이라 해석할 수 있을 것이다.



(그림 3) 실험 3의 결과 산포도 (*의 크기는 사례수에 비해)

6. 전체 논의

본 연구의 목적은 기존의 Daneman과 Carpenter (1980)의 읽기폭 검사를 피검사자 단독으로 검사를 실시할 수 있도록 재인검사로 재구성하는 것이었다. 이를 위해 기존의 읽기폭 검사와 같이 작업기억 용량의 측정이라는 관점에서 타당도를 갖춘 새로운 검사의 개발이 필요했고, 이를 위해 일련의 실험을 통해 새로운 검사와 기존의 회상읽기폭 검사간의 상관을 비교하였다.

실험 1과 실험 2에서 회상검사와 재인검사간의 낮은 상관이 관찰되었지만, 방해검사단어의 성질을 조작한 실험 3에서 회상검사와 재인검사간의 비교적 높은 상관이 관찰되었다. 실험 3에서 관찰된 0.54의 상관이 비교적 높다고 할 수 있는 근거는 원래의 회상검사가 갖는 신뢰도와의 비교이다. 이병택(인쇄중)에 따르면, 기존 회상 읽기폭 검사의 검사-재검사 신뢰도가 약 0.66 정도이며, 동형검사 신뢰도가 약 0.74이다. 또한 원래의 Daneman과 Carpenter의 읽기폭 검사의 검사-재검사 신뢰도는 약 0.41로 알려져있다 (Waters & Caplan, 1996). 뿐만 아니라 작업기억 용량을 측정하는 여러 검사들 가

운데 Turner와 Engle(1989)의 조작폭 검사의 검사-재검사 신뢰도 역시 약 0.67 정도이다(Klein & Fiss, 1999). 검사-재검사 신뢰도가 동일한 검사를 두 번 반복 검사한 결과라는 점을 감안할 때, 서로 다른 과제를 이용한 검사들간의 상관은 비교적 낮기 십상이다.

예컨대, Daneman과 Hannon(2001)에 따르면, 읽기폭사와 Nelson-Denny 이해검사와의 상관인 약 0.53이고, 기폭 검사와 SAT의 수행 상관은 약 0.3에서 0.53을 보였다. 또한 읽기폭 검사와 조작폭 검사간의 상관도 약 0.51 정도로 보고된 바 있다(Engle, Tuholiski, Laughlin, & Conway, 1999).

따라서 개관된 선행연구의 결과들에 비추어 볼 때, 본 연구의 회상 읽기폭 검사와 재인 읽기폭 검사간의 0.54의 상관은 비교적 높다고 할 수 있다. 두 검사간의 상관이 높다는 사실은 새로 구성된 검사의 타당성이라는 측면에서 중요하다. 기존의 읽기폭 검사는 이미 언어이해능력에서의 개인차를 변별하는데 타당성이 인정

<표 1> 각 실험에서의 검사점수 평균과 상관 (괄호 안은 표준편차)

	실험		
	실험 1 (N=60)	실험 2 (N=30)	실험 3 (N=37)
회상검사	3.03 (0.69)	2.97 (0.75)	3.09 (0.73)
재인검사	5.43 (0.69)	4.63 (1.10)	4.58 (1.00)
회상검사와 재인검사간의 r	0.06	0.25	0.54

된 검사이다. 재인 읽기폭 검사가 회상읽기폭 검사와 상관성이 높다는 것은 새로운 재인 읽기폭 검사가 공존타당도를 갖추고 있음을 뜻한다. 즉, 재인 읽기폭 검사 역시 회상 읽기폭 검사와 같이 언어이해능력에서의 개인차를 변별할 수 있음을 시사하는 간접적인 증거로 삼을 수 있기 때문이다.

타당도와 관련하여 실험 3의 방해검사단어의 구성에도 주목할 필요가 있다. Cantor와 Engle(1993)에 따르면 작업기억 용량이란 더 이상 적절치 않은 정보를 억제하는 능력에 다름아니다. 이와 같은 새로운 작업기억 개념의 정의에 비추어 보아, 실험 3에서 사용된 재인 읽기폭 검사는 방해검사단어로 더 이상 적절치 않은 이전 시행의 검사단어를 사용했으며, 이러한 방해검사단어의 선택을 회피하고 적절 검사단어의 선택을 요구하고 있다는 점에서 구성타당도 역시 갖추고 있다고 할 수 있다. 그러나 본 연구에서 개발된 새로운 재인읽기폭 검사가 작업기억 용량의 측정이라는 점에서 완전한 타당도를 갖추고 있다고 할 수는 없다. 앞으로의 연구는 피검사자들이 다양한 언어이해과제를 수행할 때, 그 수행의 결과를 재인읽기폭 검사가 잘 예측하는지를 직접적으로 검토하여 예언 타당도를 갖추어야 할 것이다.

한편 재인검사와 회상검사의 두 과제간의 차이라는 시각에서 우리의 결과를 바라보는 것도 흥미롭다. 회상 과제에는 피검사자들이 검사단어를 회상하기 위해 주어지는 단서가 없다. 그러나 재인과제에는 피검사자들에게 검사단어의 인출에 필요한 단서가 제공된다. 이와 같은 과제간의 차이는 보통 재인의 난이도가 회상에 비해 낮은 것으로 드러난다. 그럼에도 불구하고 Tulving과 Thomson (1973)은 재인실패현상을 통해 재인의 수행 회상에 비해 오히려 저조한 현상을 보고하고 있다. 재인실패현상은 맥락단서의 역할을 통해 설명되어지는 것이 일반적이다.

그런데 본 연구의 실험 3에서 회상검사와 재인검사가 높은 상관성을 갖는 점도 맥락단서의 역할로 잘 설명될 수 있는 것으로 보인다. 즉, 재인회상검사에서 주어진 방해검사가 부적절한 맥락으로 역할을 하여 검사단어의 인출을 방해하였다는 것이다. 이러한 검사단어 인출의 방해를 오늘날에는 간섭이라 달리 부른다고 볼 수 있을 것이다.

본 연구의 일련의 실험은 기본적으로 피검사자들의 언어이해능력에 관한 작업기억 용량을 검사자의 참여 없이 피검사자 단독으로 수행하면서 측정하려는 목적의 일환으로 수행된 것이었다. 이와 같은 상황의 대표적인 예로 웹상에서의 검사를 들 수 있을 것이다. 웹상에서

개인간의 변별이 이루어지고 언어이해능력을 측정하는 다른 종류의 실험이 이루어진다면, 언어이해능력에서의 개인차와 관련된 대규모의 자료축적과 분석이 가능해질 것이다. 또한 이것이 가능하다면, 실험실에서만 이루어져오던 실험의 결과를 일반화시키는데 있어서 갖고 있던 문제를 상당히 극복할 수 있다는 점에서 바람직할 것이고, 이러한 측면에서 본 연구의 의의를 찾을 수 있을 것이다.

7. 참고문헌

- 김경중 (1998). 문장 내의 단어미 선택과제에서 간섭효과와 개인차: 작업기억의 기능과 처리전략. 미발표 서울대학교 석사학위논문.
- 안중찬 (1999). 작업기억 다중성분설에 근거한 덩이글 읽기에서 정합적 연역 추리의 개인차. 미발표 서울대학교 석사학위논문.
- 이병택 (1995). 작업기억 용량에 따른 언어이해처리에서의 개인차. 미발표 서울대학교 석사학위논문.
- 이병택 (2002). 다의어 해소의 어휘과정에서 보이는 개인차: 숙련독자의 문맥의존적인 처리특성. 미발표 서울대학교 박사학위논문
- 이병택 (인쇄중). 읽기폭 검사의 신뢰도 연구. *심리과학*
- 이병택, 김경중, 조명환(1996). 읽기폭에 따르는 언어처리의 개인차: 작업기억과 언어이해. *한국심리학회지: 실험 및 인지*, 8(1), 59-85.
- 이선화 (2001). 덩이글 이해에서 예측 추리의 개인차: 적극적 처리와 소극적 처리. 미발표 서울대학교 석사학위논문
- Baddeley, A. D. (1986). *Working memory*. New York: University Press.
- Baddeley, A. D. (1993). Working memory or working attention. In A. Baddeley & L. Weiskrantz (Eds.), *Attention Selection, awareness, and control* (pp. 152-170 Oxford: Clarendon Press.
- Cantor, J., & Engle, R.W. (1993). Working-memory capacity as long-term memory activation: individual-differences approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 1101-1114.
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal Verbal Learning and Verbal Behavior*, 19, 450-46
- Daneman, M., & Hannon, B. (2001). Using working memory

- to investigate the construct validity of multiple-choice reading comprehension tests such as the SAT. *Journal of Experimental Psychology: General*, 130, (2), 208-223.
- Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychological Bulletin & Review*, 3(4), 422-433.
- Engle, R.W., Tuholski, S.W., Laughlin, J.E., & Conway, A.R.A. (1999). Working memory, short-term memory and general fluid intelligence: a latent-variable approach. *Journal of Experimental Psychology: General*, 128(3), 309-331.
- Gernsbacher, M. A., Varner, K. R., & Faust, M. (1999). Individual differences in general comprehension skill. *Journal of experimental Psychology: Learning, Memory, and Cognition*, 16, 430-445
- Hasher, L., & Zacks, R.T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation*, 22, 193-225. San Diego, CA: Academic Press
- Just, M. A. & Carpenter, P. A. (1992). A Capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, 99, 122-149.
- King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language*, 30, 580-602.
- Klein, K. & Fiss, W.H. (1999). The reliability and stability of the Turner and Engle working memory task. *Behavior Research Methods, Instruments, & Computers*, 31, 429-432.
- Shah, P. & Miyake, A. (1996) The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General*, 125(4), 4-27
- Tulving, E., & Thompson, D. M. (1973). Encoding specific and retrieval processes in episodic memory. *Psychological Review*, 80, 352-373.
- Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, 28, 127-154.
- Waters, G.S., & Caplan, D. (1996). The measurement of working memory capacity and its relation to reading comprehension. *Quarterly Journal of Experimental Psychology*, 49A, 51-74.