

인지과제를 통한 독서이해력의 개인차 연구*

김 영 진[†]

아주대학교 심리학과

최 광 일

아주대학교 사회과학연구소

언어이해 과정 실험연구에서 사용되는 변인과 과제들이, 독서이해력의 개인차를 측정하는 심리검사로 사용될 수 있는가를 알아보기 위해, 어휘처리 과정에 한정된 탐색적 실험을 실시하였다. 단어의 음절 수, 단어빈도, 단어의 자음동화 여부, 단어의 상위 범주 전형성 등의 변인들을 어휘판단과제와 문장검증과제를 사용해 반응시간을 측정하였다. 각 변인의 수준별 차이를 변량분석 하였고 피험자간 변량의 설명량을 분석하였으며, 심리검사 도구로서 사용을 위해 신뢰도를 일반화 가능성 계수로 분석하였고 구성타당도를 위해 확증적 요인분석을 실시하였다. 각 변인의 수준간의 차이는 통계적으로 유의미했고, 일반화 가능성 계수가 .79로 비교적 높게 나와 8개의 측정치가 신뢰로우며 이론적으로 타당한 심리검사가 될 수 있음을 확인할 수 있었다. 본 연구의 한계점과 추후 연구를 위한 시사점을 정리하였다.

주제어 : 인지과제, 독서이해, 개인차, 어휘처리

* 본 연구는 첫 저자의 2003년 한국학술진흥재단의 지원을 받아 수행되었음(과제번호2003-041-H00030).

본 논문에서 사용한 어휘 처리에 관한 실험 분석 자료는 첫째 저자의 지도하에 이루어진 둘째 저자의 박사학위 논문에 포함되었으며, 이는 학위 논문 논리 전개상 필요하다고 판단되어 심사위원장의 허락 하에 이루어 졌음. 논문에 대해 도움말을 주신 심사위원께 감사드립니다.

† 교신저자 : 김영진, 아주대 심리학과, 수원시 영통구 원천동 산5, yjkim@ajou.ac.kr

이미 50여 년 전에 Cronbach(1957)는 심리학이 실험심리학과 상관심리학이라는 두 분야(disciplines)로 나뉘어져 있는 문제를 지적하며, 이를 통합하는 시도가 필요하다고 역설한 바 있다. 어떻게 생각하면 이들 두 분야는 사실 같은 주제에 대해 다른 생각과 접근을 하고 있는 것인지도 모른다. 실험심리학도들은 두 처치 수준간의 평균적인 차이에 보통 관심을 가지며, 실험을 통해 얻은 통계적으로 유의한 차이를 기반으로 이론을 검증하거나 새로운 모형을 만들어 간다. 실험에서 발생하는 개인차는 이들에게는 골치 아픈 문제가 된다. 처치내의 개인차 변량이 이론적으로 중요한 변인의 수준 차이를 보지 못하게 만들기 때문이다. 그래서 여러 수단을 동원해 처치내 피험자들이 가능한 동질적이 되도록 노력한다.

하지만 상관심리학도에게는 개인차가 통제해야 할 오차의 근원이 아니라 오히려 탐구의 대상이라고 할 수 있다. 개인차 혹은 집단차란 생물학적인 혹은 사회적 환경이 일으킨 효과라고 할 수 있다. 모든 유기체들은 환경에 적응을 해야 하며 이 과정에서 모두가 성공적인 적응을 하는 것은 아니다. 그러기에 어떤 측정을 해도 개인차를 관찰할 수 있다. 그렇다면 유기체의 어떤 특성이 적응의 정도와 방식을 결정한 것인가의 의문을 제기할 수 있으면, 여러 특성들 간의 관련성은 풀어야 하는 실타래의 실마리가 되는 것이다.

이 두 가지 입장을, 논문의 처음부터 대비시킨 이유는 바로 본 연구가, 실험심리학 연구와 상관심리학 연구를 합쳐보려는 혹은 통합해보려는 시도이기 때문이다. 언어 이해 과정에 관한 인지심리학 실험에서, 이론적으로

중요한 처치 변인들의 효과를 알아보기 위해 사용되었던 다양한 실험 과제와 방법들이 과연 개인차를 알아보는 일종의 심리 측정, 즉 검사 도구로 사용할 수 있겠는가 연구의 출발점이 되었기 때문이다. 물론 다양한 인지과제를 사용해 개인차를 보려는 시도는 예전부터 있어 왔으며, 대표적인 개인차 변인인 지능 연구에 적용되어 왔다. 점검시간(inspection time)과 같은 여러 실험 과제가, 특별한 이론적 근거 없이, 일종의 지능 검사로 사용되기도 했고, 정보처리모형에 기반을 두고 유추과제(analogy task)가 지능 측정 도구로 사용되기도 했다. 이 지능 측정에 관한 자세한 언급은, 이 주제 자체가 중요하고 흥미로운 것이며 광범위하기에, 본 논문에서는 생략하겠으며 자세한 개관은 이정모 등 (2009)나 Sternberg & Ben-Zeev(2001) 참조하면 될 것이다. 지능 측정에 전통적인 지필 검사 대신에 실험 과제를 사용한 이 접근이 지능, 나아가 일반적인 인지능력의 개인차를 전통적인 심리검사와는 다른 측면에서 이해하는데 도움을 주었다고 할 수 있으며, 반대로 집단 평균에 의존하는 실험 연구를 보완하는 역할도 한다고 할 수 있다. 실제로 김청택(2002)은 개인차 분석법을 통해 실험 결과의 통계적 의미를 확장할 수 있다는 것을 정신회전(mental rotation) 과제 수행 분석을 통해 보여주고 있다.

본 연구에서 적용하려는 이 두 접근은 언어 이해과정 연구, 보다 구체적으로 독서이해과정 연구에서 특히 유용하다고 할 수 있다. 지난 30여 년간 다양한 인지 실험 과제와 방법을 사용해 언어이해과정에 관한 인지심리학적 연구들이 수행되었다. 그리고 이들 연구의 결

과들은 독서이해과정이 어떻게 이루어지는 것인가에 관한, 대부분의 연구자들의 동의를 얻을 수 있는, 모형을 설정할 수 있는 수준에 도달했다고 해도 과언은 아니다. 독서이해과정을 이해하기 위해서는, 지각과정을 통해 단어를 확인하는 체계, 통사분석을 담당하는 체계, 덩이글 표상과 상황모형을 구축하는 체계 등의 하위체계를 상정해야 하며, 이들이 언어적 규칙에 관한 지식과 일반 세상에 관한 지식과 상호작용 하는 것이라는 개념화(예, Kintsch, 1998; Kintsch & Rawson, 2005; Perfetti, 1999; Perfetti, Landi, & Oakhill, 2005)에 이를 제기하는 인지심리학자는 아마 없을 것이다. 아울러 이들 하위 과정 처리에 각각 어떤 변인들이 중요한지를 역시 수많은 연구들이 밝혀내고 있다. 물론 각 하위체계들과 지식이 어떻게 상호작용 할 것인가와 관련하여, 언어 이해를 담당하는 전체적인 인지 건축구조(cognitive architecture)를 어떻게 구체화 할 것인가는 여전히 논쟁거리로 남아 있다.

하지만 이 논쟁과는 중립적으로, 이렇게 잘 정리된 통합 틀, 즉 독서과정에 관한 인지심리학적 이해가 독서이해력의 개인차를 이해하는데 적극적으로 사용될 수 있는 시점이 되었다고 생각할 수 있다. 즉 이들 하위 독서 과정을 포괄하는 독서이해과정 측정치를 곧바로 독서이해력의 개인차 측정치로 사용할 수 있을 것이라는 필자들의 판단이다. 아직 연구자들이 합의할 수 있는 모습을 그리는 것조차 힘든, 지능이니 창의성 같은 개념에 비하면, 독서이해과정은 훨씬 좋은 개인차 측정 연구의 자리를 잡고 있는 셈이다. 인지심리학적 틀에서 독서이해과정의 개인차를 살펴본 여러

연구가 있다(자세한 개관은 최광일, 2007 참조). 몇 가지 대표적인 예만 들면, 초등학교 수준에서 독서의 숙달을 결정하는 개인차를 음성적 부호를 통해 어휘집에 접속하는 능력이 좌우함을 보인 연구(예, Stanovich, 2002)가 있으며, 작업기억 용량과 관련한 통사처리과정의 개인차를 살펴본 연구(예, King & Just, 1991)가 있다. 그리고 덩이글 요약에서도 개인차가 나타남을 보인 연구(예, Hyona, Lorch, & Kaakinen, 2002)가 있으며 Hannon과 Daneman (2001)은 덩이글 기억, 추론, 지식 통합 등의 하위 과정이 독서이해의 개인차에 중요함을 보여주고 있다.

본 연구에서는 위에서 언급한 연구들을 기반으로 하여, 다양한 독서이해과정 측정 과제를 일종의 심리검사 문항처럼 사용할 가능성을 탐구하였다. 구체적인 예를 하나 들어, 어휘접속 과정에 영향을 끼치는 변인의 하나는 단어의 빈도이며, 이는 보통 어휘판단과제의 판단시간의 차이로 확인된다. 그렇다면 “고빈도와 저빈도 단어에 대한 어휘판단 시간을 마치 어휘 능력을 측정하는 지필 심리검사의 문항들처럼 사용할 수 있겠는가” 하는 가능성을 알아보고자 하는 것이다. 혹자는 이러한 시도를 불필요한 것으로 여길 수도 있다. 간단한 지필검사가 있는데, 굳이 컴퓨터로 통제해 시간을 측정해야 하는 과제를 검사 문항처럼 사용할 필요가 있느냐고 의문을 제기할 수 있다. 하지만 기존의 많은 검사들이 전산화 혹은 인터넷 기반 검사로 바뀌고 있다는 점에서 컴퓨터 사용 자체는 문제가 될 수 없다고 보이며, 오히려 지필검사가 제공하지 못하는 어떤 정보를 얻을 수 있느냐가 논의에 초점이 되어야

할 것이다. 소위 말하는 독해력 지필 검사는 개인의 독서이해 능력에 관한 전체적인 점수나 지표를 제공하지만 구체적인 독서 이해 과정의 어떤 하위 체계의 어떤 처리가 문제 되는지는 알려주지 않는 것이 보통이다. 앞서 기술했던 것처럼 독서이해과정의 하위체계들과 중요 변인들을 측정하는 과제를 검사도구로 사용하면, 한 개인이 독서이해의 어떤 처리 과정에서 처리의 어려움 혹은 문제를 일으키는지를 구체적으로 파악할 수 있는 장점을 갖고 있다고 생각된다.

측정 변인과 과제의 선택 개인차를 나타내는 포괄적인 독해력 지수를 만들기 위해서는 어휘, 문장, 덩이글, 모형형성의 네 수준에서 측정이 이뤄져야 한다는 것은 확실하다. 하지만 각 수준에서, 이론적으로 중요한 것으로 밝혀진 변인들 중 어느 것을 선택해야 하는가는 지침이 없다. 그리고 특정 변인의 어떠한 수준을 사용해야하는지도 미리 알 수 없다. 수준의 차이가 잘 나타나는 변인들을 선택해야 하고, 아울러 선택한 변인들이 개인차도 잘 나타내주면 되지만 이 역시 사전에 예측하는 것이 힘들다. 그래서 이를 시도하는 본 연구는 어쩔 수 없이 탐색적인 성격을 가질 수밖에 없다. 그리고 독서이해과정의 모든 하위체계들을 함께 고려하며 그 결과를 분석하고 논의하는 것은, 제한된 지면을 사용할 수밖에 없는 단일 논문으로는 불가능하다. 그래서 본 논문에서는 우선 어휘처리에만 한정하여 결과 분석과 논의를 진행하였다.¹⁾

1) 본 연구는 단어 수준에서 4개의 변인, 문장 수준에서 10개, 덩이글 수준에서 6개, 모형구성에

선택한 어휘 처리의 변인은 표 1에 제시되어 있는 것처럼 음절수(1,2,3,4), 어휘빈도(고, 저), 자음동화 여부, 단어의 상위 범주에서의 전형성 정도 네 가지였다. 음절수 즉 단어길이는 시각적인 단어 처리의 효율성을 알아보기 위한 것이었고, 어휘빈도는 단어접속 실험에서 가장 일관성 있는 결과를 보이는 변인이기에 선택하였다. 자음동화 여부는 음성정보 활용 여부의 개인차를 알아보기 위한 목적이었으며, 단순한 어휘 접속 수준을 넘어서 단어와 단어와의 관계 즉 개념 간의 관계 파악을 알아보기 위해 전형성 정도도 포함시켰다. 이러한 선택에 어떤 이론적 근거가 있는 것은 아니며 임의적인 판단에 근거한 것이다. 선택의 타당성은 이론적 근거보다는 경험적 확인의 문제라고 여겨진다. 처음 세 변인들에 대해서는 어휘판단과제를 사용하였고 마지막 변인은 문장검증과제를 사용하였다. 실험 연구에서는 단어를 무선변인으로 보고 분석이 가능하도록 보통 5-10개의 단어들을 사용하고 적절한 실험설계를 사용하여야 하나(Clark, 1973), 본 연구에서는, 너무 많은 수의 단어를 사용하면, 실시 시간이 길어지기에 2, 4개로 제한하였다. 그래서 어휘 변인을 무선변인으로 보는 F-검증은 하지 않았다. Clark(1973)이 제안한 단일 사례 사용법을 따른 것이다.

요약하면, 본 논문에서는 독서이해과정, 특히 어휘처리 과정에서 이론적 중요성을 갖는

서 4개 등 총 24개의 변인들을 선택하여 실시되었으나 본 논문에서는 이 모두를 포함하여 논문화 하는 것이 불가능하기에 어휘 분석 결과만을 보고하였다. 다른 수준에 대한 분석 결과는 최광일(2007)을 참조 하면 된다.

요인들 네 가지와 두 실험과제들이 독서능력의 개인차를 알아보는 일종의 심리검사 문항처럼 사용될 수 있겠는가를 살펴보고자 하였다.

방 법

참가자 아주 대학교에 재학 중인 학생 73명이 연구에 참여하였고, 반응의 정확도가 낮은 1명을 제외하고 분석에 사용된 참여자는 72명이었다.

실험 재료 사용된 자극은 현재까지 한국의 인지심리학과 언어심리학의 연구들(남기춘, 1997; 이관용, 1991; 이양, 1998)에서 사용되었거나, 두 분야의 표준적인 교재(이정모 외, 2003; 조명한 외, 2003)에 언급된 자료를 기본으로 하였고, 필요에 따라 새로 구성하기도 하였다. 측정 변인으로 사용된 단어특성은 길이(즉 음절의 수), 사용빈도, 지음동화여부, 및 단어의 범주 전형성 정도였다. 단어길이의 효과를 측정하기 위해 1음절에서 4음절 단어 2개씩 총 8단어를 선택하였다. 비단어도 같은 방식으로 8개를 만들었다. 단어길이 효과가 가지고 있는 기본적인 가정은 단어가 길수록 어휘판단과제에서 느린 읽기시간을 보인다는 것이다. 예를 들어 2음절 단어 ‘곡식’은 4음절 단어 ‘소금쟁이’보다 빠르게 어휘접속이 일어날 것이다. 단어 빈도효과를 위해서는 저빈도 단어 4개와 고빈도 단어 4개를 선별하였고, 지음동화 현상의 효과를 측정하기 위해, 시각 단어(지음동화가 일어나지 않는 단어) 4개와 음운단어(지음동화가 일어나는 단어) 4개를 사

표 1. 단어이해과정 측정변인 및 자극 재료

측정 변인	수준(단어 예)
음절수 별 자극 (8개)	1음절 (뱀, 꽃)
	2음절 (곡식, 색깔)
	3음절 (고양이, 나룻배)
	4음절 (국회의원, 소금쟁이)
빈도별 자극 (8개)	고빈도 (눈물, 당신, 생산, 만족)
	저빈도 (판촉, 악상, 상동, 빈축)
자음동화 (8개)	시각 단어 (발명, 감독, 작품, 학습)
	음운 단어 (백년, 입맛, 먹물, 학년)
전형성 단어 (8개)	고전형 (장롱, 버스, 나비, 사과)
	저전형 (찬장, 가마, 거미, 자몽)
	상위범주 (가구, 교통수단, 곤충, 과일)

용하였다. 이 두 종류의 단어는 모든 고빈도로 갖게 만들었다. 그리고 전형성 효과를 측정하기 위해 범주를 대표하는 전형성이 높은 단어 4개와 전형성이 낮은 단어 4개를 선택하였다. 처음 세 변인은 어휘판단과제를 마지막 전형성 효과는 문장검증과제로 구성하였다. 또한 24개의 비단어와 8개의 거짓 문장을 구성하여 판단의 반응 편향을 제거하였다. 본 과제에 사용된 단어자극이 표 1에 제시되어 있다.

절차 참가자들에게 신속하면서도 정확하게 반응하도록 강조하였으며, 과제 순서는 어휘판단 과제를 먼저 실시했고, 연이어 문장판단 과제를 실시하였다. 어휘판단과제로 제시되는 단어들과 문장판단과제를 위한 문장들의 순서는 참가자별로 무선화 하였다.

결과 및 논의

서론에서 언급한 것과 같이 본 연구의 목적은 어휘 수준에서 독서이해력의 개인차를 잘 밝혀낼 수 있는 의미 있는 어휘 변인과 과제를 탐색하는 것이라고 할 수 있다. 그러기에 결과 분석은 두 가지로 이루어졌다. 기존의 실험 연구 결과 분석과 같이, 변인의 수준간의 차이가 나타나기를 알아보는 분석, 즉 실험 분석을 하였으며 아울러 심리 검사의 신뢰도와 타당성을 알아보는 검사 분석을 하였다.

실험 분석 표 2에 실험에 대한 각 변인별 수준간의 변량분석 결과가 제시되어 있다. 단어의 시각적 특징인 음절 변인의 주효과가 통계적으로 유의미 하였으나 ($F(3, 71)=14.53, p<.0001$), 단어 길이와 읽기 시간 사이에 정적인 관계가 관찰 되지 않았다. 즉 1음절 단어

와 2음절 단어를 비교해 보면 2음절 단어가 1음절 단어보다 짧은 읽기시간으로 나타났으며 ($F(1,71)=14.115, p<.000$), 3음절 단어와 4음절 단어의 읽기시간의 차이는 유의하지 않았다. 단어 길이 효과가 가장 두드러지게 나타난 것은 2음절과 3음절 단어의 차이로 2음절 단어의 읽기 시간이 약 110ms 신속했다($F(3, 71)=33.57, p<.0001$).

본 실험에 측정 변인으로 사용된 빈도효과, 자음동화, 및 전형성 효과는 과거 여러 연구 결과와 같이 강력한 읽기 시간의 차이를 보였다. 고빈도 단어가 저빈도 단어보다 약 320ms 짧은 읽기 시간을 보였으며($F(1, 71)= 128.85, p<.000$), 자음동화 현상이 있는 단어가 없는 단어보다 약 200ms 읽기시간이 길었다($F(1, 71) = 35.15, p<.000$). 그리고 전형성이 높은 단어의 문장판단시간이 낮은 단어에 비해 약 340ms 짧았다($F(1, 71)= 7.519, p<.000$).

본 연구의 실험설계에서는 피험자 즉 72명

표 2. 단어이해과정의 하위 변인들의 통계치

변인	수준	평균(표준편차)	통계치(F)	df	유의도
음절수	1음절	669 (132)	전체 14.53 1 vs 2 14.12 2 vs 3 33.57	3, 71	.0001
	2음절	625 (106)			
	3음절	739 (205)			
	4음절	721 (208)			
빈도	고	624 (123)	128.85	71	.000
	저	940 (265)			
자음동화	무	675 (124)	35.15	71	.000
	유	894 (343)			
전형성	고	1225 (319)	45.68	71	.000
	저	1563 (495)			

주) 대부분의 자유도는 두 수준이기 때문에 F(1,71)이다. 자유도가 1이 아닌 경우만 제시하였음.

의 개인들이 네 변인 총 10개의 수준과 교차된다. 그래서 위의 표2에 제시한 것처럼 처치 효과의 변량을 계산하고, 피험자와 처치와의 상호작용을 오차항으로 하여 검증을 할 수 있을 뿐 만 아니라, 개인 간의 차이가 일으킨 변량을 검증하고 그 효과 크기를 산출할 수도 있다. 물론 대부분의 분석에서는 개인 간의 차이는 검증하지 않으며, 검증하면 대부분은 유의하게 나오는 것이 보통이다. 본 결과도 마찬가지로 네 변인들 모두에서 유의한 차이가 나왔다.

개인차가 나타내는 설명량을 오메가스퀘어(ω^2)의 추정치로 계산해보니, 음절(.52), 빈도(.26), 자음동화(.30), 전형성(.69)로 나타났다. 특정한 변인의 수준에 대해 개인들이 보이는 차이의 정도를 반영하는 것으로 이 통계치를 사용할 수 있기에, 이는 흥미로운 결과로 여겨진다. 즉 빈도와 자음동화는 상대적으로 적은 설명량을 보였고, 음절수가 50%, 전형성이 가장 높은 69%의 설명량을 보인 결과는 전자의 두 변인이 상대적으로 적은 개인차 즉 개인간에 일치하는 반응을 보였다는 것이고, 다시 말하면, 전자의 두 변인이 보다는 후자의 두 변인이 개인차를 측정하기에 더 적합한 것이라는 추측을 가능케 한다. 하지만 개인 간의

변량이 크다는 것만으로 좋은 검사 과제 혹은 문항이라고 볼 수 없기에 추후의 연구와 분석이 필요할 것이다.

검사 분석

일반화가능성 분석. 실험 분석에서 유의한 차이로 나타난 2음절과 3음절의 차이, 고빈도와 저빈도의 차이, 자음동화 여부에 따른 차이 즉 시각단어와 음운단어 간의 어휘판단시간 6개와 고전형과 저전형 문장 간의 문장판단시간 2개를 합친 총 8개의 측정치만을 가지고 검사도구로서의 분석을 실시하였다. 우선 일종의 신뢰도라고 할 수 있는 일반가능성 계수(generalizability coefficients)를 계산하였다. 구체적인 계산 절차는 Shavelson과 Webb(1991)을 참조하며, Di Nocera, Ferlazzo, & Borghi(2001)이 사건유발전위(event-related potential) 측정의 신뢰도를 보기위해 사용했던 방법을 함께 고려하였다. 아울러 8개로 측정된 값들이 네 개의 어휘처리 속도(lexical speed) 하위 잠재변인으로 어떻게 구성되는지를 알아보기 위해 확증적 요인분석을 실시하였다.

변량 세 구성성분에 대한 추정치(EVC: Estimates of Variance Components)가 표 3에 제시되어있다. 가장 큰 변이가 검사 단어 때문인 것으로

표 3. 일반화 가능성 분석 결과

	Df	MS	EVC	%
피험자	71	252236.68	24980.29	13
검사단어	7	7973329.41	110012.98	59
피험자X단어(잔차)	497	52394.62	52394.62	28
상대 오차	31529.59	일반화가능성 계수(E_p^2) = .79		
절대 오차	45281.21	의존가능성계수(ϕ) = .55		

나타났으며(59%), 이는 본 연구에서 사용한 자극인 단어의 종류와 두 비교 조건이 읽기시간의 변이에 가장 크게 영향을 끼쳤음을 보여주고 있다. 한편 피험자에 의한 변이도 13% 정도로 단어 처리 속도가 개인에 걸쳐 차이가 있다는 것을 보여 주고 있으나, 다른 변이원에 비해 적어 어휘 처리 속도가 개인차를 크게 보여주지 않고 있다고 생각된다. 피험자와 단어의 상호작용 즉 잔차(residual)는 피험자의 상대적 위치가 본 연구에서 사용한 자극 단어 별로 달라지는 정도를 나타내는 것으로 28%로 나타났다.

앞서 언급했던 것처럼, 일반화 가능성 계수는 기대되는 관찰점수 변량에 대한 보편점수 변량의 비율로, 연구의 목적에 따라 다르게 계산되고 다른 이름으로 불리기도 한다. 상대적인 서열에 초점을 두는 경우(relative decision, 상대적 결정)는 상대 오차(relative error)에 기초하며, 한 개인의 절대적 수준의 지식 혹은 기술 지표 설정이 목적인 경우(absolute decision)에는 절대 오차를 사용하며, 이를 특히 의존가능성 계수(dependability coefficient)라고 부른다.

잔차의 추정 변량을 단어의 수로 나누면 상대오차의 변량 성분이 되며, 피험자의 추정 변량을, 이 값과 상대오차의 변량 성분을 합한 값으로 나누면 상대오차에 기초한 일반화 가능성 계수($E\beta^2$)를 계산할 수 있다. 이 값은 .79로 상대적으로 높게 나왔다. 즉 어휘 처리 속도가 개인차를 보여주는 일관적이며 신뢰로

운 측정치가 될 수 있음을 보여 주는 것이다. 역시 앞서 언급했던 것처럼 일반화가능성 이론에서는 검사 도구의 신뢰도를 올리기 위한 최적의 검사 문항 수를 결정하는데도 사용될 수 있는데, 본 연구에서 사용했던 8개의 단어 자극 수를 배로 늘리면(16개) 일반화 가능성 계수가 .88까지 커지는 것으로 계산되었다. 하지만 본 연구가 함축하는 궁극적인 목적인 독서 이해력의 개인차를 잘 나타내는 단어, 문장, 텍스트 수준의 여러 실험 과제를 찾는 것이기에, 단어 수준에서의 8개의 단어 자극만으로도 .80에 가까운 신뢰도를 보였기에 8개의 단어 자극에 만족할 수 있다.

한편 검사단어의 변량이 포함되는 절대오차 변량 성분에 기초한 의존가능성 계수는 .55로 나와 일반화가능성 계수보다는 낮았다. 즉 개인들의 상대적인 순위는 비교적 신뢰로운 측정치가 되지만 절대적인 등수는 신뢰도가 떨어짐을 보여주는 것이다.

확증적 요인분석. 단어 이해의 개인차를 나타내기 위해 각 하위 요인들이 내재되어 있는 어휘처리 속도(lexical speed) 시스템에 어떤 가중치로 설명될 수 있는가? 이를 위해 어휘 처리 시스템이라는 최상의 가상적인 변인을 설정하고, 각 하위 요인들이 어느 정도 어휘 시스템을 설명할 수 있는가를 밝힐 수 있는 확증적 요인 분석을 실시하였다. 분석에 사용된 측정치들은 일반화 가능성 분석에서 사용

표 4. 확증적 요인분석의 모델 적합도

적합도 지수	자유도	X^2	CFI	TLI	RMSEA	NFI	GFI
확증적 요인분석	18	29.15	.95	.93	.09	.89	.92

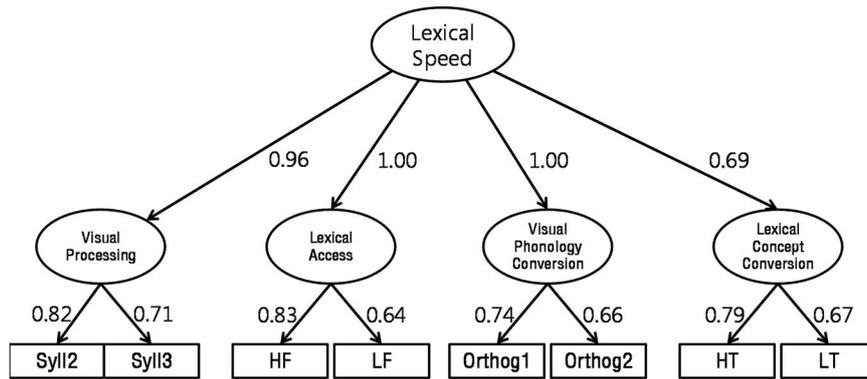


그림 1. 어휘처리 시스템의 요인 모델. Lexical Speed(어휘처리속도), Visual Processing (시각처리), Lexical Access(어휘접속), Visual-Phonology Conversion(시-청각전환), Lexical-Concept Conversion:(어휘-개념 전환). (Syll2,3: 각각 음절수 2,3, HF, LF: 고빈도, 저빈도, Orthog1,2: 자음동화 유, 무, HT, LT: 전형성고, 저)

된 8개의 측정변인 이었고 이들이 하위 요인을 설명하는 강도가 그림 1에 제시되어 있다. 그리고 이 확장적 요인 분석 모델의 적합도 지수가 표 4에 제시되었다.

확증적 요인 분석 모델의 적합도 지수로 유심히 보아야 할 값들은 CFI, TLI와 RMSEA이다 (홍세희, 2000). 본 연구 모형의 지수 값들은 CFI=.95, TLI=.93, RMSEA=.09로 이론모형의 적합도가 좋다고 할 수 있다. RMSEA < .1 이면 보통 적합도(moderate fit) 수준이기에 좋은 모형의 요인을 갖추지 않았을지 모르나, 이러한 경우 홍세희(2000)는 RMSEA의 90% 신뢰구간을 고려함으로써 RMSEA의 신뢰도를 확인할 것을 권고하고 있다. 본 모델에서 RMSEA의 신뢰구간은 .01-.15로 간격이 넓어 추정된 RMSEA의 값에 오차가 많으며 신뢰할 수 없다. 따라서 CFI, TLI를 중심으로 모델의 적합도를 평가해야 하며 본 모델은 매우 좋은 모델에 가깝다고 할 수 있다. 전통적인 적합도

지수인 NFI와 GFI로 평가해 보면 기초부합치 GFI의 값은 .92로 높아 좋은 모델에 속하지만, NFI 값은 .89로 좋은 모델에 포함하기에는 다소 부족하다. 그러나 NFI 값이 표본수에 민감한 지수이기에 다소 적은 표본을 대상으로 실시한 본 연구를 감안하면 충분히 좋은 모델의 여건을 마련했다고 할 수 있다.

확증적 요인 분석 모델을 기초로 어휘처리 시스템을 보면, 단어 길이, 단어 빈도와 자음동화 효과가 어휘처리 확인에 상당한 영향을 미치는 것을 알 수 있다. 전형성 효과는 다른 세 요인들과 비교하여 다소 영향력이 약하지만 그래도 비교적 큰 효과를 보여주고 있다. 또한 주로 쉬운 조건(2음절, 고빈도, 자음동화 없음, 고 전형성)의 측정치들이 어려운 조건보다 추정된 각 요인에 큰 효과를 나타낸다는 것을 보여주고 있다.

요약 및 결론

본 연구는 독서이해 과정 실험 연구에서 사용되고 있는 과제와 변인들이 독서이해력의 개인차를 알아보는 심리 검사로 사용될 수 있는 가능성을 알아보고자 실시하였다. 어휘 판단 과제를 사용해 측정된 단어의 길이, 빈도, 자음동화 여부 등의 6개의 반응시간과, 문장 속에 전형성이 높고, 낮은 단어를 포함시켜 측정된 2개의 문장판단시간은, 개인차를 알아볼 수 있는 측정치가 될 수 있음을 확인시켜 주었다.

실험 분석에서 나타난 것처럼, 네 개 변인 모두 두 수준에서 기존의 실험 결과와 동일하게 유의미한 차이를 보여 본 연구 결과에 신뢰를 부여할 수 있다. 그리고 개인차 즉 피험자간의 변량 계산에서 전형성과 음절 변인의 변량이 빈도, 자음동화 변인보다 크게 나와 전자의 두 변인이 개인차 측정에 상대적으로 더 유용할 가능성을 시사한다.

일반화가능성 계수로 알아본 신뢰도 측정에서도, 비록 8개의 측정치만을 사용했음에도 불구하고 약 .80의 비교적 높은 신뢰도를 보인 것인 상당히 고무적이라고 할 수 있다. 그리고 어휘처리 속도라는 최상의 요인을 설정하고 4개의 하위 요인과 8개의 측정치가 어떤 가중치로 설명될 수 있는가를 분석한 확증적 요인분석 결과는 단어 빈도와 자음동화 효과가 상위 요인을 대체할 수 있음을 시사한다고 해석할 수 있다. 물론 이러한 분석 결과가 다른 두 측정치가 필요 없다는 것을 보여주는 것으로 해석할 수는 없을 것이다. 네 측정치는 개념적으로 서로 다른 인지 과정을 반영하

는 것이기에 각기 다른 측면에 관한 정보를 제공하기 때문이다.

본 논문의 서론에서도 언급한 것처럼, 본 연구는 기본적으로 탐색적일 수밖에 없는 한계가 있다. 첫째로, 본 연구에서는 어휘처리 과정을 단지 두 개의 과제와 네 개의 변인만을 사용해 분석했다. 어휘 처리에 영향을 끼치는 다른 여러 변인들(예, 형태소 처리, 이웃 효과 등)을 추후 연구에서 고려해야 할 것이다.

둘째로, 독서이해과정의 어휘처리 수준 이외의 다른 처리 수준들도 포함시켜야 진정한 의미의 독서이해력의 개인차를 측정하는 심리 검사가 될 수 있기 때문에 이를 다 포함하여 포괄적인 분석을 해야 할 것이다. 상관연구의 장점이 여러 변인들을 동시에 고려하면 분석할 수 있다는 중다변인적 분석이라는 점에서, 다른 수준을 포함하는 중다변인 연구 설계가 필요하다. 아울러 이 연구와 후속해야 할 관련 연구 결과들을 토대로 일종의 “독서 이해력 지수”를 산출해야 할 것이며, 이 점수의 분포 자료를 확보하여 진정한 의미의 독해력의 개인차를 측정하는 심리 검사도구로 사용될 수 있는 기반을 확보해야 할 것이다.

세 번째로 본 연구에서 사용하지 못한 개인별 측정치는 각 변인의 두 수준간의 차이 값이다. 예를 들어 저빈도와 고빈도 단어에 대한 어휘판단 시간의 측정치를 살펴보면, 전체 집단 평균적으로 두 빈도 간에 316ms의 유의한 차이를 보이지만, 개인별로 보면 이 차이 값은 음수 값에서 영을 거쳐 양수 값으로 분포한다. 이 차이 값을 개인차 분석에서 어떻게 사용할 것인가는 앞으로 해결해야 할 문제

이다. 보다 근본적인 측면에서, 반응시간이 과연 개인차 측정치로 적절한가를 다시 돌이켜 볼 필요도 있다고 여겨진다. 한 개인의 검사 상황에 따른 변산성이 크기 때문에, 상대적인 비교를 할 수 있는 기준적인 반응시간 측정(예, 단순 반응시간)을 우선 실시할 수도 있을 것이다.

네 번째로, 세 번째 문제와 어떤 측면에서는 관련이 있는 것으로, 본 연구가 실험적 연구와 상관연구를 모두 고려한다고 했지만, 실험분석에서는 처치간의 변량만을 사용했고, 검사 분석에서는 개인차 변량만을 기초로 하였다. 그러기에 여전히 유기체와 처치 변인들간의 상호작용은 무시한 게 된다는 Cronbach (1957)의 지적은 타당해 보인다. 이러한 분석을 위한 틀이 필요하다고 생각된다.

참고문헌

김청택 (2002). 개인차분석법을 이용한 실험결과의 통계적 분석의 확장. *한국심리학회지: 실험 및 인지*, 14, 107-126.

남기춘 (1997). 한글 단어 재인에서의 단어 길이 효과. *한국심리학회지: 실험 및 인지*, 9, 1-18.

이관용 (1991). 우리말 범주규준 조사. 본보기 산출빈도, 전형성, 그리고 세부특징 조사. *한국심리학회지: 실험 및 인지*, 9, 73-94.

이양 (1998). 한글 단어 지각에서 표음식도와 처리자원의 영향. *한국심리학회지: 실험 및 인지*, 10, 1-16.

이정모 외 (2003). *인지심리학*. 학지사.

조명한 외 (2003). *언어심리학*. 학지사.

최광일 (2007). 독서이해 과정의 개인차: 인지 과제 수행과 안구운동 패턴. *아주대 박사학위논문*.

홍세희 (2000). 구조 방정식 모형의 적합도 지수 선정기준과 그 근거. *한국심리학회지: 임상*, 19, 161-177.

Clark, H. H. (1973). The language as fixed effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, 12, 335-359.

Cronach, I. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, 12, 671-684.

Di Nocera, F., Ferlazzo, F., & Borghi, V. (2001). G theory and the reliability of psychophysiological measures: A Tutorial. *Psychophysiology*, 38, 796-806.

Hannon, B., & Daneman, M. (2001) A new tool for measuring and understanding individual differences in the component processes of reading comprehension. *Journal of Educational Psychology*, 93, 103-128.

Hyöna, J., Lorch, R. F., & Kaakinen, J. K. (2002). Individual differences in reading to summarize expository text: Evidence from eye fixation patterns. *Journal of Educational Psychology*, 94, 44-45.

King, J & Just M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language*, 30, 580-602.

Kintsch, K., & Rawson, K, A.(2005). Comprehension. In Snowling, M, J., &

- Hulme. C. (Eds). *The Science of Reading* (pp. 209-226). Malden, MA; Blacwell.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge, England: Cambridge University Press.
- Perfetti, C. A., & Tan, L. H. (1999). The consistency model of Chinese word identification. In J. Wang, A. Inhoff, A. W. Chen, & H. C. Chen (Eds). *Orthography, Phonology, morphology, and meaning*(pp. 115-134). Mahwah, NJ: Erlbaum.
- Perfetti, C. A., Landi, N., & Oakhill. J. (2005). The Acquisition of Reading Comprehension Skill. In Snowling, M, J., & Hulme. C. (Eds). *The Science of Reading* (pp. 227-247). Malden, MA; Blacwell.
- Shaelson, R. J. & Webb, N. M. (1991). *Generalizability theory: A Primer*. London: Sage.
- Sternberg R. J. & Ben-Zeev, T. (2001). *Complex cognition*. New York: Oxford.

1 차원고접수 : 2010. 6. 3

수정원고접수 : 2010. 6. 18

최종게재결정 : 2010. 6. 25

Investigating Individual differences in Reading Comprehension Ability by using Cognitive Tasks

Youngjin Kim

Kwangill Choi

Department of Psychology, Ajou University

We investigated some language comprehension tasks whether they can be used as psychometric tests(or assessment tools). An exploratory experiment were performed to examine lexical processing times by using lexical decision & sentence verification tasks. Used experimental variables were numbers of syllable, frequency, phonological change, & typicality. We analysed the results by using ANOVA, estimated omega square, generalizability coefficients, and factor analysis in order to examine treatment effects, explained variance, and the reliability and validity of experimental results. All these result suggested the experimental variables and tasks that we used could be good psychometric tests. We discussed some limitations of this study.

Key words : cognitive tasks, reading comprehension, individual differences, lexical processes