

# 과업의 관여 부담 수준과 작업 기억이 고등학교 영어 학습자의 제2언어 어휘 학습에 미치는 영향

이옥희<sup>1</sup> · 이상기<sup>1†</sup>

<sup>1</sup>한국교원대학교

## Effects of Task-Induced Involvement Load and Working Memory on Korean High School Students' Second Language Vocabulary Learning

Ok-hee Lee<sup>1</sup> & Sang-Ki Lee<sup>1†</sup>

<sup>1</sup>Korea National University of Education

---

### ABSTRACT

This study investigated the effects of task-induced involvement load and working memory on Korean high school English learners' vocabulary acquisition and long-term retention. Sixty-six students were divided into three groups and experienced three different tasks that differed in terms of their involvement load: 1) gap-filling, 2) unscrambling sentences, and 3) writing original sentences. Each group took a post-vocabulary test comprising two subtests—form recognition and meaning recognition—immediately after the treatment and one week later. Results showed that the students who performed the task of writing original sentences, which was assumed to have the highest involvement load, gained the highest scores in the immediate meaning recognition test. Further, working memory of the students had a weak but statistically significant correlation with the short-term learning of both form and meaning and the long-term retention of meaning. More detailed findings are presented along with suggestions for pedagogy and future research.

**Keywords:** involvement load hypothesis, working memory, vocabulary learning

---

### 1. 서 론

우연적 학습(incidental learning)은 무언가 다른 과업을 수행한 부산물로서 이뤄진 학습을 의미하며, 특별한 의도를 동반하지 않는다는 측면에서 의도적 학습(intentional learning)과 대조를 이루는 것으로 이해된다(Marsick & Watkins, 2016; Ortega, 2009).

---

<sup>†</sup> Corresponding author: sangkissi@gmail.com



Copyright © 2021 Language Education Institute, Seoul National University.

This is an Open Access article under CC BY-NC License (<http://creativecommons.org/licenses/by-nc/4.0>).

이와 관련하여 제2언어(second language: L2) 어휘 습득 연구 분야에서는 의사소통 중심의 과업을 통한 우연적 어휘 학습의 결과가 어떠한지를 활발하게 살펴왔다(Malone, 2018). 이때 관련 연구 맥락에서는 주로 의미 중심의 읽기 과업이 주어지는 사례들이 많았다(e.g., Cho & Krashen, 1994; Dupuy & Krashen, 1993). 그에 대해 L2 어휘 학습의 효과를 제고하기 위해서는 단순한 읽기 과업 이상의 내용이 제공되어야 할 것이며(Folse, 2006; Horst et al., 1998; Schmitt, 2008), 그 과업이 구체적으로 무엇이나에 따라 과업 수행에 있어서의 정의적, 인지적 측면에서의 부담의 수준이 달라질 것이고, 그 결과 어휘 학습 정도 역시 영향을 받게 될 것이라는 주장이 제기되었다. 이를 관여 부담 가설(involverment load hypothesis)이라고 이른다(Hulstijn & Laufer, 2001).

그간 많은 연구들을 통해 관여 부담 가설의 타당성이 논의되어온 가운데(e.g., Eckerth & Tavakoli, 2012; Folse, 2006; Huang et al., 2012; Keating, 2008; Kim, 2008, 2011; Laufer, 2003; Lu, 2013; Marmol & Sanchez-Lafuente, 2013; Nassaji & Hu, 2012; Soleimani & Rahmanian, 2014; Yang et al., 2017; Zou, 2017), 본 연구 역시 그 연장선에서 채 해결되지 않아 추가적인 고민이 요구되는 대목에 대해 살펴보고자 하였다. 예컨대, 관여 부담 가설의 이론적 측면에서 논의하는 과업의 수준이 실제로 있어서는 예측에서 벗어나는 사례들이 많았고(e.g., Choi, 2015; Folse, 2006; Keating, 2008; Kim, 2015; Marmol & Sanchez-Lafuente, 2013; Park, 2011; Soleimani & Rahmanian, 2014; Yang et al., 2017; Zou, 2017), 연구 결과의 외적 타당도를 확보하는 과정에서 외국어로서 영어를 교수학습하는 국내 교실의 특성을 고려하여 우리나라 학생들에게 적용 가능성이 높은 과업들의 관여 부담 수준이 과연 어느 정도인지를 알아볼 필요가 있었다(cf. Choi, 2017; Sung, 2013).

한편, 본 연구에서는 L2 어휘 습득에 개입하는 또 다른 중요 변인으로 언어 학습 적성(language learning aptitude)의 영향을 함께 고려하고자 하였다. 언어 학습 적성의 여러 구성 요소들 중 특히 작업 기억(working memory)은 복잡한 인지 과제 수행을 위해 정보를 일시적으로 저장하고 처리하는 인지적 체계로서(Baddeley & Hitch, 1974), 어휘 학습과의 연관성이 높은 것으로 알려져 왔다(Atkins & Baddeley, 1998; Linck et al., 2014).

정리하면, 본 연구는 국내 고등학교 영어 학습자들의 어휘 학습 과정에서 과업의 관여 부담 수준이 어떠한 영향을 미치는지, 학습자들의 작업 기억 측면에서의 수준의 차이가 미치는 매개 변인 효과는 어떠한지를 알아보고자 구안되었다. 이러한 맥락에서 제기된 연구문제들은 다음과 같았다.

첫째, 과업의 관여 부담 수준이 고등학교 영어 학습자의 어휘 학습에 미치는 단기 및 장기 효과는 어떠한가?

둘째, 서로 다른 다양한 관여 부담 수준의 과업들을 수행함에 있어 읽기 폭 과제(reading span task)로 측정되는 작업 기억은 고등학교 영어 학습자들의 단기

## 2. 이론적 배경

본 연구는 Hulstijn과 Laufer(2001)에 의해 최초 제기된 관여 부담 가설의 타당성을 실증적으로 입증해 보고자 하였다. 해당 가설은 과업을 수행하는 동안 새로운 어휘를 학습하는 정도가 과업을 수행하는데 필요한 인지적 처리 수준에 따라 달라진다는 가정이다. 이때, 처리 수준은 정보처리의 정교화와 같은 의미로 간주되며, 새로운 어휘 정보를 정교하게 처리할수록 장기기억에 도움이 된다고 본다. 이론에 따르면 관여 부담 수준은 필요(Need, 이하 N), 검색(Search, 이하 S), 평가(Evaluation, 이하 E)의 세 가지 요소에 의해 결정된다.

먼저 필요는 동기적, 비인지적 요소로서 주어진 과업의 수행을 위해 학습자가 목표 어휘를 알아야 할 필요가 있는지의 정도에 따라 N0, N1, N2의 세 수준으로 구분된다. 그리하여 과업을 수행하는 과정에 있어 목표 어휘를 알 필요가 없는 경우(N0), 목표 어휘를 알아야만 하는 외적 필요성이 존재하는 경우(N1), 궁극적으로 학습자의 내재적 동기 측면에서의 자극이 이뤄질 수 있는 경우(N2)로 나뉘게 된다.

검색은 인지적 요인이다. 목표 어휘의 의미를 사전, 교사 등의 자료를 동원해 찾아보고자 하는 학습자의 노력이 동반되는지(S1), 그렇지 않은지(S0)에 따라 두 수준으로 구분된다.

평가 역시 인지적 요인으로서 과업 수행 과정에서 학습자에게 결정을 내리도록 유도하는 정도와 관련되며, E0, E1, E2의 세 수준으로 구분된다. 구체적으로는, 교수학습 과정에서 주어진 맥락 자체를 활용하여 어휘들 간의 의미 차이 혹은 하나의 단어가 가지는 여러 의미들 간의 차이를 인지하도록 유도하는 경우(E1)와 학습자가 스스로 만들어진 맥락에서 목표 어휘를 다른 어휘들과 결합하는 과정에서의 결정을 유도하는 경우(E2)로 나뉘게 된다. 한편, 평가적 요소가 전혀 개입하지 않는 경우는 E0로 이해된다.

이상과 같이 필요, 검색, 평가 측면에서의 내용을 종합하여 과업의 관여 부담 수준이 점수로 산출되게 된다. 이것이 관여 부담 지수(involvement index)이며, 구체적으로는 0점에서 5점 사이에 그 값이 위치하게 된다. 예를 들어, 5점은 목표 어휘를 반드시 알아야 할 강한 필요가 있어 학습자의 내적 동기를 자극하는 경우(N2), 각종의 자료를 동원해 목표 어휘의 의미를 검색해야 하는 경우(S1), 학습자 본인 스스로 창출한 맥락에서 목표 어휘의 타당성을 평가해야 하는 경우(E2)에 주어지는 점수이다. Hulstijn과 Laufer(2001)는 관여 부담 지수가 높은 과업일수록 그를 수행함으로써 어휘 학습이 더욱 촉진될 것이라고 보았다.

Hulstijn과 Laufer(2001)는 EFL 대학생들을 대상으로 주석 달린 지문 읽기(N1, S0, E0=1), 지문 읽고 빈칸 채우기(N1, S0, E1=2), 작문(N1, S0, E2=3)의 세 가지 과업을

적용한 결과, 관여 부담 수준이 높을수록 어휘 평가 점수가 높았으며, 앞의 두 읽기 집단과 작문 집단 간에 유의미한 점수 차가 나타났음을 보고했다. Laufer(2003)는 어휘 주석이 달린 지문을 읽고 이해 문제 풀기(N1, S0, E0=1)와 문장 쓰기(N1, S0, E2=3 [실험 1]), 작문(N1, S0, E2=3 [실험 2])을 각각 비교한 결과, 문장 쓰기와 작문이 읽고 이해 문제 풀기에 비해 즉시와 지연 사후 어휘 평가에서 통계적으로 더 높은 점수로 이어졌음을 밝혔다. 두 연구에 이어 유사한 맥락에서 Keating(2008), Eckert와 Tavakoli(2012), Marmol과 Sanchez-Lafuente(2013), Yang 외 3인(2017) 등도 관여 부담 가설을 대체로 지지하는 연구 결과를 내놓았다. 한편, Zou(2017)는 작문(N1, S0, E2=3), 문장 쓰기(N1, S0, E2=3), 빈칸 채우기(N1, S0, E1=2)의 순으로 어휘 학습이 더 효과적이었음을 보여 가설을 부분적으로 입증하되, 문장 단위 쓰기에 비해 담화 단위 쓰기가 더욱 효과적일 수 있음을 주장하였다.

한편, 관여 부담 가설에 반하는 결과를 내놓은 연구들도 많았다. Folse(2006)는 관여 부담 수준이 서로 다른 빈칸 채우기 활동(N2, S1, E1=4)과 문장 쓰기 활동(N2, S1, E2=5)의 어휘 학습 결과에 상호 간 유의미한 차이가 없었으며, 오히려 세 개의 빈칸 채우기 활동이 하나의 문장 쓰기 활동에 비해 더욱 효과적임을 보여, 어휘 처리의 깊이보다는 요구되는 어휘 인출(retrieval)의 횟수가 보다 중요할 수 있음을 보였다. Lu(2013) 역시 중하위의 영어 능숙도를 가진 학습자에게는 관여 부담 수준이 높은 작문 과업에 비해 빈칸 채우기 과업이 어휘 학습에 더욱 효과적임을 보였다. 이어 Soleimani와 Rahmanian(2014)은 지문의 빈칸을 채우고 이해문제 풀기(N1, S0, E1=2), 읽고 이해문제를 푼 후 문장 쓰기(N1, S0, E2=3)를 각각 수행하게 한 결과, 두 집단의 사후 어휘 평가 점수 간에 차이가 없었음을 보고하였다.

다음으로 국내 연구 사례들을 살펴보면, 특히 빈칸 채우기 과업에 비해 문장 쓰기 과업의 관여도 수준이 높음에도 불구하고 두 과업이 어휘 학습에 미치는 효과와 관련하여서는 일관성 있는 연구 결과를 찾아보기가 어려웠다. Park(2011)은 대학생을 대상으로 읽기(N1, S1, E1=3), 빈칸 채우기(N1, S1, E1=3), 문장 쓰기(N1, S0, E2=3) 등 관여 부담 수준이 동일하게 설정된 세 과업을 통해 어휘 학습의 정도를 살펴본 결과, 문장 쓰기 집단과 읽기 집단 사이에서만 통계적으로 유의미한 차이를 발견하였다. 이에 대해 관여 부담 수준이 동일함에도 수용적 활동보다는 생산적 활동을 할 때(Laufer, 2003), 검색보다는 평가 측면에서 차이가 있을 때 어휘 학습에 더욱 큰 영향을 미치게 된다고 해석하였다. Kim(2015)은 대학생을 대상으로 평가의 정도가 다른 과업-읽고 이해문제 풀기(N1, S0, E0=1), 읽고 이해문제 푼 후 빈칸 채우기(N1, S0, E1=2), 문장 쓰기(N1, S0, E2=3)를 수행하게 한 결과, 영어 능숙도가 높은 학습자들의 경우 문장 쓰기가 빈칸 채우기에 비해 통계적으로 유의미한 수준에서 어휘 학습과 기억에 긍정적인 효과를 가져오는 것을 입증했다. Choi(2015)는 고1 학생들을 대상으로 어휘 주석이 달린 지문 읽기(N1, S0, E0=1), 빈칸 채우기(N1, S0, E1=2), 제한된 문장 쓰기(N1, S0, E2=3)를 수행하게 하였다. 생산적 어휘 지식의 경우 즉시 및 지연 사후 어휘

평가에서 빈칸 채우기와 문장 쓰기 두 집단이 읽기 집단에 비해 통계적으로 유의미하게 우수한 결과를 보인 반면, 수용적 어휘 지식은 집단 간에 차이가 없었다. Sung(2013)은 고1 학생들에게 읽고 이해문제 풀기(N1, S0, E0=1), 읽고 이해문제 풀고 빈칸 채우기(N1, S0, E1=2), 어휘 배열하여 문장 만들기(N1, S0, E2=3)를 수행하게 하고 어휘 학습 정도를 살펴보았다. 그 결과, 빈칸 채우기와 어휘 배열하여 문장 만들기 집단 간에는 유의미한 차이가 없었고, 이 두 집단이 읽기 집단에 비해 더 우수한 어휘 학습결과를 보였다.

한편, Sung(2013)은 한국 고등학생들의 쓰기 수준의 현실을 고려하여 어휘 배열하여 문장 만들기 과업이 쓰기 과업과 동일한 관여 부담 수준을 갖는다고 가정하였고, 이러한 가정은 Choi(2017)의 연구에서도 이어졌다. 그러나 Hulstijn과 Laufer(2001)가 제시한 강한 정도의 평가(E2)는 학습자 스스로가 만들어낸 문장이나 지문에서 목표 어휘와 다른 어휘들의 결합 방식이 어떠한지를 평가하는 것임에 주목해야 한다. 그리고 이는 앞선 연구들에서 가정된 바와는 달리 어휘를 배열하여 문장을 만드는 과업의 경우 관여 부담 지수를 상대적으로 조금 더 낮은 수준의 것으로 이해해야 할 필요가 있음을 시사한다. 이에 본 연구에서는 학습자 스스로 문장을 만들어야 하는 과업과 주어진 어휘를 배열하여 문장을 구성해야 하는 과업이 어휘 학습에 미치는 영향을 직접적으로 비교함으로써 두 과업이 가지는 관여 부담 수준이 과연 어떠한지를 살펴보고자 하였다. 본 연구는 또한 해당 두 가지 과업에 더하여, 빈칸 채우기 과업을 포함함으로써 세 가지의 서로 다른 과업 수준이 어휘 학습에 미치는 영향을 살펴보았다. 세 가지의 과업 모두 외국어로서 영어를 학습하는 국내 고등학교 교실에서 실제 빈번히 활용되는 것들을 고려할 때, 본 연구의 결과를 통해 현장의 어휘 교수 활동 및 평가의 맥락에 유의미한 시사점을 도출할 수 있을 것이라 기대되었다.

또 한편, 본 연구는 과업의 관여 부담 수준이 고등학교 영어 학습자의 어휘 학습에 미치는 영향을 살피는 과정에 있어 개인차 변인으로서의 언어 학습 적성의 역할에 동시에 주목하고자 하였다. 동일한 교수학습 처치를 적용한다고 하더라도 학습자들은 학습의 과정과 그 결과 측면에서 매우 상이한 양상을 보이기 마련이다. 이때 다양한 개인차 요인들 중에서도 언어 학습 적성은 가장 큰 설명력을 가지는 것 중의 하나로 논의되어 왔다(Lee, 2014; Masgoret & Gardner, 2003; Ortega, 2009). 특히 언어 학습 적성 요인으로서의 작업 기억은 어휘 학습과 밀접한 관련성을 맺게 되어, 작업 기억 능력이 뛰어난 학습자일수록 모국어와 L2 어휘 학습 모두에서 더 우수한 수행을 보이고 있음이 여러 연구를 통해 입증되어 온 바 있다(e.g., Atkins & Baddeley, 1998; Chun & Payne, 2004; Linck et al., 2014). 이에 본 연구는 어휘 학습에 미치는 관여 부담 수준의 영향이 연구 참여 학생들의 작업 기억 측면에서의 언어 학습 적성 수준에 따라 유의미하게 달라질 것이라 가정하고 이를 입증하고자 하였다. 이는 교실 학습자의 인지적 측면에서의 다양성을 고려하여 학습자 맞춤형 교수 설계의 가능성을 타진해 보고자 하는 노력의 일환이었다.

### 3. 연구 방법

#### 3.1. 연구 참여자

본 연구는 경기도 시흥시 소재 H 고등학교 2학년 세 개 학급의 87명을 참여자로 하여 진행하였으며, 한 차례라도 검사에서 빠진 학생들을 제외하고 총 66명의 학생들을 최종 분석의 대상으로 삼았다(남학생 34명, 여학생 32명). H 고등학교는 비평준화 지역인 시흥시에서 학구열이 매우 높은 자율형공립고이며, 시흥시 포함 기타 지역(안산, 부천, 광명, 안양 등)에서 자원한 상위권 학생들로 구성된 학교였다. 이를 입증하듯 연구 참여 학생들은 모의고사 성적을 기준으로 학급 평균 2~3등급에 해당하는 높은 성취 수준을 보이고 있었다.

#### 3.2. 읽기 자료 및 목표 어휘 선정

우연적 어휘 학습을 목표로 우선 The Definition and Characteristics of Grit이라는 제목의 읽기 자료를 선정하였다. 해당 자료는 성공하는 사람들의 성격 특성 중 하나로 여겨지는 ‘grit(투지, 기개, 근성)’의 정의와 특징에 관한 내용을 담고 있었다. 글의 구조와 내용이 참여 학생들에게 적절한지를 고려하여 필요한 수정을 행하였으며, 그에 대해서는 영어 원어민의 검토를 거쳤다. 읽기 자료는 29개 문장, 516개의 단어로 이뤄져 있었고, Flesch Kincaid 이독성 지수가 9.4로 나타나 연구 참여 학생들의 수준에 비추어 적절한 것으로 판단하였다(<부록> 참조).

Laufer(1989)는 독자들이 지문의 전체 어휘 중 95%에 대한 지식을 가지고 있을 때 익숙하지 않은 어휘의 의미를 추론할 수 있다고 한 반면, Nation(2001)은 모르는 어휘의 의미를 정확히 추론하기 위해서는 적어도 지문의 98%를 이해할 수 있어야 한다고 보았다. 해당 논의를 고려하여 본 연구에서는 빈칸 채우기 과업의 수행에 있어 학생들이 특별한 어려움을 겪지 않도록 지문의 총 단어 수를 약 500단어로 하고, 그 중 10개의 목표 어휘를 선정하였다(abiding, angst-ridden, dead-end, imbued, indomitable, introspection, perseverance, resilient, revelatory, tippy-top). 이는 지문의 전체 어휘 수 대비 목표 어휘 수의 비율이 대략 2% 수준이었음을 의미한다.

#### 3.3. 과업의 관여 부담 수준 설정

본 연구에서는 고등학교 영어 교실에서 빈번히 활용되고 있는 빈칸 채우기, 어휘 배열 하여 문장 만들기, 문장 쓰기의 세 가지 과업을 연구 참여 학생들로 하여금 수행하도록 하였다. 실제로 본 연구에 참여한 학생들은 평소 수행평가 및 지필평가를 통해 빈칸 채우기와 어휘 배열하여 문장 만들기 과업 유형에 자주 노출되어 왔으며, 문장 단위의

쓰기 활동 역시 수업시간을 통해 자주 접해오고 있었다.

한편, 연구의 타당도를 높이기 위하여 다음과 같은 조치를 취하였다. 먼저, 어휘 주석 제공 방식, 문장길이 및 구조 등을 조정하여 각 과업들 간의 난이도 차이가 더욱 분명할 수 있도록 하였다. 빈칸 채우기 활동에서 어휘 주석을 제공하는 것이 그렇지 않은 경우에 비해 어휘 기억을 촉진시킨다는 기존의 연구들(Hulstijn, 1992; Hulstijn et al., 1996; Watanabe, 1997)에 따라, 빈칸 채우기 과업에서 어휘 주석을 제공하기로 결정하였다. 다만, 빈칸 채우기의 과업 난이도를 다소 낮추기 위하여 목표 어휘의 뜻을 목표 언어인 영어가 아닌 우리말로 제시하였다. 어휘 배열하여 문장 만들기 과업에서는 난이도의 상향 조정을 위해 기존 연구들에 비해 하나의 문장을 구성하는 어휘의 수를 늘리고 다소 복잡한 문장 구조를 적용하였다. 다음으로, 검색 조건을 추가하여 각 과업 간의 난이도 차이가 분명하도록 하였다. 읽기 과업을 수행하는 동안 직접 사전을 찾아본 어휘들이 그렇지 않은 어휘들에 비해, 또한 주석이 제공된 단어들에 비해 더욱 잘 기억된다는 연구 결과(e.g., Laufer, 2000; Watanabe, 1997)를 고려하여 본 연구에서는 빈칸 채우기 과업에서는 주석을 제공하는 반면(S0), 어휘 배열하여 문장 만들기과 문장 쓰기 과업에서는 사전 검색 및 짝과의 정보차 활동이라는 검색 조건을 추가하였다(S1). 이상과 같은 미세 조정 과정을 통해 본 연구에서 적용한 세 가지의 과업 및 각 과업에 대해 가정된 관여 부담 수준은 표 1에 제시된 바와 같았다.

**표 1.** 과업의 종류 및 과업별 관여 부담 수준

집단	과업	관여 부담 수준
1	주석 달린 지문 읽고 빈칸 채우기	N1, S0, E1 = 2
2	어휘 검색 후 어휘 배열하여 문장 만들기	N1, S1, E1 = 3
3	어휘 검색 후 문장 쓰기	N1, S1, E2 = 4

### 3.4. 읽기 폭 과제를 통한 작업 기억 측정

연구 참여 학생들의 작업 기억을 측정하기 위하여 본 연구에서는 Conway 외 5인 (2005)에서 활용한 읽기 폭 과제를 선택하였고, 이를 예비 검사를 통해 고등학교 2학년 학생들의 수준에 맞추어 수정하였다. 이때 PPT 슬라이드 쇼를 활용하여 문제를 제시하고 학습자들이 검사지에 개별적으로 답안을 작성하게 하는 등 집단 검사가 가능하도록 하였다. 학습자들은 화면에 제시되는 영어 문장을 소리 내어 읽고 그에 대한 어법성을 판단하여 검사지에 O 혹은 X를 적은 뒤, 문장의 말미에 제시되는 알파벳을 소리 내어 읽고 그를 후에 기억해 적어야 했다. 처음에는 두 개 문장으로 이루어진 세트가 3회 제시되었고, 이어서 세 개 문장이 한 세트, 네 개 문장이 한 세트, 다섯 개 문장이 한 세트를 이루어 각각 3회씩 제시되는 방식으로 하여 총 12세트, 42개의 문항이 구성되었다.

연습세트

1. He is the boy who  
he likes my sister.

V

연습세트

2. Never trust a man  
whose talks too much.

E

연습세트

“알파벳을 순서대로 적어주세요.”

연습세트 정답

	문장의 어법이 맞으면 O, 틀리면 X	알파벳
연습 1.	X	V
세트 2.	X	E

★ 다음 세트를 준비해주세요.

그림 1. 읽기 폭 과제 문항 예시(두 문장으로 구성된 연습 세트)

본 검사에 앞서 하나의 예시와 하나의 연습 세트를 먼저 제시하여 참여자들의 과업에 대한 이해를 도왔다. 한편, 실험에 활용된 각 문장은 7~12개의 단어로 구성되었으며, 고등학교 2학년이 쉽게 읽고 이해할 수 있는 내용을 담고 있었다. 그림 1은 읽기 폭 과제의 문항 예시이다.

### 3.5. 어휘력 평가

본 연구에서는 Paribakht와 Wesche(1993)가 제안한 VKS(vocabulary knowledge scale) 5점 척도를 활용한 의미인식평가로 연구 참여 학생들의 어휘력을 측정하였다. 그림 2에 제시된 바와 같이 VKS의 다섯 가지 선택지를 통해 학습자가 가지고 있는 어휘에 대한 의미 지식(①~④)과 사용 지식(⑤)을 동시에 확인해볼 수 있었다(cf. Nation, 2001).

① 나는 이 단어를 전에 본 적이 없다.  
 ② 나는 이 단어를 전에 본 적이 있으나, 무슨 뜻인지 모르겠다.  
 ③ 나는 이 단어를 전에 본 적이 있고, 그 의미는 \_\_\_\_\_ (이)라고 생각한다.  
 ④ 나는 이 단어를 알고 있으며, 그 의미는 \_\_\_\_\_ (이)다.  
 ⑤ 나는 이 단어를 문장에서 사용할 수 있다: \_\_\_\_\_

그림 2. VKS(Vocabulary Knowledge Scale)



이와 같이 VKS로 구성된 검사지를 10개의 목표 어휘에 대해 적용하여 사전 검사로 삼았다. 한편, 사후 어휘력 검사를 위해서는 VKS 검사지에 추가하여 형태인식평가를 함께 적용하였다. 형태인식평가는 객관식 10개 문항으로, 각 문항은 가짜 단어를 포함한 유사한 형태의 네 개 어휘 중 학습자가 기억한 목표 어휘를 구분해 내도록 구성되었다. 이때 가짜 단어는 영어의 음운론 및 형태소 규칙에 벗어나지 않도록 제작되었다. 한편, 어휘 형태 추론에 있어 학습자가 스스로의 판단에 대해 확신하는 정도를 5점 척도 상에서 선택하여 답하도록 하였다. 이렇게 제작된 두 개의 검사지는 본 검사 실시 일주일 후 지연 사후 어휘력 검사에서 재차 적용되었다. 이때 문항과 선택지 순서를 조정하여 연습 효과를 최소화하고자 하였다.

### 3.6. 연구 절차

본 실험은 세 학급을 대상으로 정규 수업시간의 세 개 차시를 이용하여 진행되었다. 첫 번째 차시에서는 10개 문항으로 구성된 사전 어휘 검사를 7분 동안, 이어서 작업 기억 검사를 13분 간 실시하였다. 두 번째 차시는 3주 후에 진행되었다. 먼저, 주석 달린 지문 읽고 빈칸 채우기 과업(이하 ‘빈칸 채우기’)을 수행한 첫 번째 집단의 학습자들은 지문을 읽고 우리말 뜻 주석이 달린 10개의 목표 어휘 목록에서 빈칸에 적절한 단어를 선택하여 적어 넣었다. 다음으로, 어휘 검색 후 어휘 배열하여 문장 만들기 과업(이하 ‘어휘 배열하기’)을 진행한 두 번째 집단의 학습자들의 경우에는 먼저 단어 검색 활동을 수행하였다. 짝과 함께 목표 어휘 10개 중 다섯 개씩의 단어를 나누어 맡아 2분 동안 휴대폰으로 그 의미를 검색하였다. 휴대폰을 반납한 후, 다시 2분 동안 짝과 함께 나머지 다섯 개 단어의 의미를 서로 주고받으며 목표 어휘 10개의 의미를 찾아내었다. 이후 개인 활동으로, 검색한 목표 어휘 10개가 하나씩 포함된 10개 문항에서 주어진 어휘 덩어리를 배열하여 우리말 뜻에 맞게 문장을 완성하는 활동을 수행하였다. 마지막으로, 어휘 검색 후 문장 쓰기 과업(이하 ‘문장 쓰기’)을 진행한 세 번째 집단의 학습자들은 두 번째 집단과 동일한 단어 검색 활동을 먼저 수행한 후, 개인 활동으로 목표 어휘를 포함한 10개의 영어 문장을 자유롭게 작성하였다. 각 문장은 목표 어휘를 포함하여 일곱 개 이상의 단어로 구성하도록 하였다. 세 집단 모두 본 과업 직후 형태인식평가(10문항, 3분)와 의미인식평가(10문항, 7분)로 구성된 사후 어휘 평가를 실시하였다. 마지막 세 번째 차시는 일주일 후에 진행되었다. 이때 형태인식평가와 의미인식평가로 구성된 지연 사후 어휘 평가를 실시하여 학습한 바의 지속 가능성 여부를 살펴보았다. 이상과 같은 실험 절차는 표 2에 정리되어 있다.

표 2. 실험 절차

차시	측정 내용 및 도구			시간
1	사전 어휘 평가: 의미인식평가			7분
	작업 기억 검사			13분
3주 간격				
2	집단 1	집단 2	집단 3	25분
	빈칸 채우기	어휘 배열하기	문장 쓰기	
	사후 어휘 평가: 형태 및 의미인식평가			10분
1주 간격				
3	지연 사후 어휘 평가: 형태 및 의미인식평가			10분

3.7. 자료의 분석 및 처리

작업 기억 검사의 결과는 세트와 상관없이 어법성 판단을 정확하게 한 문장에 한해 올바르게 기억해낸 알파벳 하나당 1점을 부여하는 부분점수 방식을 적용하여(Conway et al., 2005) 총 42점 만점으로 채점되었다. VKS 5점 척도를 활용한 어휘 의미인식평가는 총점이 최소 10점에서 최대 60점으로 구성되었다<sup>1)</sup>. 마지막으로 어휘 형태인식평가의 경우에는 10개 문항 각각에 1점씩을 부여하여 총 10점 만점으로 분석이 이뤄졌으며, 그와 함께 5점 척도로 제시된 확신 지수에 대한 응답은 각 문항 당 5점을 만점으로 하여 총 50점 만점으로 계산되었다. 이와 같은 방식으로 산출한 데이터는 자료의 정규성을 확인한 후 SPSS(ver.22.0)를 통해 통계 처리하였고, 모든 결과에 대해서는 유의수준 5%에서 해석하였다. 또한 필요한 경우 Cohen (1988)의 *d*값을 통해 효과 크기를 가늠하여 연구 결과의 해석에 있어 활용하였다.

4. 연구 결과

4.1. 집단 간 동질성 검증 및 사전 어휘 평가 결과

먼저 세 실험 집단의 전반적인 영어 능숙도 수준이 상호 간 동질하였는지를 교내 지필고사<sup>2)</sup> 점수를 통해 알아보았다. 일원분산분석 결과, 세 집단의 영어점수에는 통계

1) VKS의 5개 문항 중 ⑥번의 경우 의미와 어법을 모두 고려하여, 목표어와 관련한 어법이 부적절하나 의미가 적절한 문장이면 5점, 목표 어휘와 관련한 어법과 의미가 모두 적절한 문장이면 6점으로 채점하였다.

적으로 유의미한 차이가 없었음을 확인하였다,  $F(2, 63)=0.19, p=0.83$ . 한편, 목표 어휘에 대한 사전 어휘 지식 평가 결과는 표 3에 제시된 바와 같았다. 각 집단의 사전 검사에서의 평균 점수는 대략 13점으로 나타났는데, 이는 연구 참여 학생들이 목표 어휘를 본 적이 없거나, 본 적이 있더라도 그 뜻을 모르는 수준에서 본 과업에 임하였음을 의미하였다. 목표 어휘에 대한 사전 지식 측면에서의 집단 동질성을 추가적으로 검증한 결과, 세 집단이 서로 간 통계적으로 유의미한 차이가 없었던 동질한 수준이었음이 드러났다,  $F(2, 63)=0.47, p=0.63$ .

**표 3.** 기술통계량: 사전 어휘 평가

집단	<i>n</i>	<i>Mean</i>	<i>Std. Deviation</i>	<i>F</i>	<i>Sig.</i>
1(빈칸 채우기)	21	13.52	1.12		
2(어휘 배열하기)	23	13.00	2.41	.47	.63
3(문장 쓰기)	22	13.00	2.27		

#### 4.2. 과업의 관여 부담 수준이 어휘 학습에 미친 효과

첫 번째 연구문제와 관련하여, 과업의 관여 부담 수준에 따른 어휘 학습의 단기 효과를 먼저 살펴보았다. 표 4에 제시된 바와 같이 사후 어휘 형태인식평가로부터의 각 집단의 평균 점수는 집단 1에서 7.90점, 집단 2에서 8.22점, 집단 3에서 8.64점으로 나타났다. 일원분산분석 결과, 세 집단의 평균 점수 사이에는 통계적으로 유의미한 차이가 없었다,  $F(2, 63)=1.28, p=0.29$ .

**표 4.** 기술통계량: 사후 어휘 형태인식평가

집단	<i>n</i>	<i>Mean</i>	<i>Std. Deviation</i>	<i>d</i>
1(빈칸 채우기)	21	7.90	1.89	
2(어휘 배열하기)	23	8.22	1.31	0.20
3(문장 쓰기)	22	8.64	1.26	0.46

한편, 표 5에 제시된 바와 같이, 사후 어휘 의미인식평가에서는 집단 3의 평균이 32.91점으로 나타나 다른 두 집단에 비해 높았다(집단 1: 27.00점, 집단 2: 26.17점). 집단 1을 기준으로 살펴본 효과 크기 *d*는 집단 3이 0.82로 컸고, 집단 2가 0.13으로 나타나 작은 수준에 채 미치지 못하고 있었다. 기술통계량에 대한 일원분산분석 결과, 세 집단의 사후 어휘 의미인식평가 점수에서 통계적으로 유의미한 차이가 드러났다,

2) 여기서 지필고사는 학기 중에 두 차례 실시되어 그동안 학습한 내용을 바탕으로 영어 능숙도를 평가하는 총괄평가 성격의 중간고사와 기말고사를 의미한다.

$F(2, 63) = 5.60, p < 0.01, \eta^2 = 0.15$ .

표 5. 기술통계량: 사후 어휘 의미인식평가

집단	<i>n</i>	<i>Mean</i>	<i>Std. Deviation</i>	<i>d</i>
1(빈칸 채우기)	21	27.00	5.74	
2(어휘 배열하기)	23	26.17	7.40	0.13
3(문장 쓰기)	22	32.91	8.44	0.82

어휘 학습의 단기 효과와 관련하여 어떠한 집단의 대조에서 통계적으로 유의미한 평균 점수의 차이가 있었는지를 알아보기 위해 Scheffé 사후 검정을 실시하였다. 그 결과, 집단 3과 집단 1, 집단 3과 집단 2 사이에서 통계적으로 유의미한 차이가 나타났다(집단 3 vs. 집단 1:  $p < 0.05$ , 95% CI: 0.32~11.50; 집단 3 vs. 집단 2:  $p < 0.01$ , 95% CI: 1.27~12.20). 그에 비해 집단 1과 집단 2 간의 평균 점수 차이는 통계적으로 유의미하지 못했다( $p = 0.93$ , 95% CI: -4.70~6.35). 요컨대, 과업의 관여 부담 수준이 가장 높았던 문장 쓰기 집단이 다른 두 집단에 비해 단기적인 어휘 의미인식평가에서 통계적으로 유의미한 수준에서 우수한 수행을 보였음을 확인할 수 있었다.

첫 번째 연구문제를 통해서 어휘 학습 정도가 어느 정도 지속되는 것인지에 대해서도 알아보고자 하였다. 그에 관련하여 지연 사후 검사 결과를 살펴보았다. 먼저 표 6에 제시된 바와 같이 세 집단의 지연 사후 어휘 형태인식평가 점수의 평균은 집단 1이 7.71점, 집단 2가 7.09점, 집단 3이 7.86점으로 나타났으나, 그것이 통계적으로 유의미한 수준의 차이는 아니었다,  $F(2, 63) = 0.80, p = 0.45$ .

표 6. 기술통계량: 지연 사후 어휘 형태인식평가

집단	<i>n</i>	<i>Mean</i>	<i>Std. Deviation</i>	<i>d</i>
1(빈칸 채우기)	21	7.71	2.28	
2(어휘 배열하기)	23	7.09	2.64	0.25
3(문장 쓰기)	22	7.86	1.42	0.08

다음으로 지연 사후 어휘 의미인식평가 결과는 표 7에 제시된 바와 같았다. 세 집단의 평균은 집단 1이 21.67점, 집단 2가 22.26점, 집단 3이 23.82점이었다. 일원분산분석 결과, 세 집단 간에는 지연 사후 어휘 의미인식평가 점수에 있어 통계적으로 유의미한 차이가 나타나지 않았다,  $F(2, 63) = 1.08, p = 0.35$ .

표 7. 기술통계량: 지연 사후 어휘 의미인식평가

집단	<i>n</i>	<i>Mean</i>	<i>Std. Deviation</i>	<i>d</i>
1(빈칸 채우기)	21	21.67	3.61	
2(어휘 배열하기)	23	22.26	3.54	0.17
3(문장 쓰기)	22	23.82	3.85	0.58

한편, 본 연구에서는 과업의 관여 부담 수준이 어휘 학습에 미치는 영향과 관련하여 사후 어휘 형태인식 검사를 통해 연구 참여 학생들로 하여금 각 문항에 대해 응답하는 과정에 있어 스스로 확신하는 정도에 대해 5점 척도 상에서 선택하여 답하도록 하였다. 그 결과, 표 8에 제시된 바와 같이, 사후 검사에서의 값이 평균 40.58점( $SD=6.10$ ), 지연 사후 검사에서의 값이 그보다 낮은 평균 35.45점( $SD=9.35$ )으로 나타났다. 응답의 분포를 보다 자세히 살펴보면, 사후 검사에서 매우 확신 49%, 확신 24%, 보통 17%, 불확신 6%, 매우 불확신 2%, 모름 2%로, 73%의 학생들이 형태인식평가 답안에 대해 확신 또는 매우 확신함을 알 수 있었다. 지연 사후 검사에서는 매우 확신 32%, 확신 25%, 보통 23%, 불확신 11%, 매우 불확신 4%, 모름 5%로, 57%의 학생들이 자신들의 답안에 대해 확신 또는 매우 확신하는 것으로 나타나, 사후 평가에 비해 확신 지수가 다소 낮아진 것을 확인할 수 있었다. 대응표본 *t*-검정 결과, 두 검사 단계에서의 확신 정도의 차이가 통계적으로 유의미한 수준으로 드러났다,  $t(65)=6.40$ ,  $p<0.001$ .

표 8. 기술통계량: 어휘 형태 추론 확신 정도

	<i>n</i>	<i>Mean</i>	<i>Std. Deviation</i>	최소값	최대값
사후	66	40.58	6.10	19	50
지연 사후	66	35.45	9.35	5	50

다음으로 형태인식평가로부터의 결과와 확신 정도 간의 관계를 알아보기 위해 피어슨 적률상관관계 분석을 실시하였다. 그 결과, 사후 검사 결과와 확신 정도는  $r=.54$ , 지연 사후 검사 결과와 확신 정도는  $r=.59$ 로 나타났으며, 두 수치 모두 통계적으로 유의미한 수준의 것이었다( $p<0.01$ ). 이는 과업을 수행하면서 새롭게 마주친 단어의 형태를 정확하게 인식하는 정도와 그에 대한 확신 정도가 유의미한 수준에서 상호 간 관련성을 맺고 있음을 의미한다.

#### 4.3. 작업 기억과 어휘 학습 간의 상관관계 분석 결과

관여 부담 수준이 상이한 세 가지 과업을 수행한 각 집단의 작업 기억 검사 결과는

표 9와 같았다. 세 집단의 평균은 42점 만점에 집단 1이 28.67점, 집단 2가 28.17점, 집단 3이 30.55점이었다. 일원분산분석 결과, 세 집단의 작업 기억력에는 통계적으로 유의미한 차이가 없었고, 이를 통해 작업 기억 측면에서 세 집단이 동질하였음을 확인할 수 있었다,  $F(2, 63)=1.07, p=0.35$ .

**표 9.** 기술통계량: 작업 기억 검사

집단	<i>n</i>	<i>Mean</i>	<i>Std. Deviation</i>
1(빈칸 채우기)	21	28.67	7.10
2(어휘 배열하기)	23	28.17	4.68
3(문장 쓰기)	22	30.55	5.20

두 번째 연구문제와 관련하여 작업 기억과 어휘 학습 간의 관계를 알아보고자 피어슨 적률상관계수를 구해본 결과, 작업 기억 점수는 사후 어휘 형태인식평가 결과( $r=.25, p<0.05$ ), 의미인식평가 결과( $r=.27, p<0.05$ ) 및 지연 사후 어휘 의미인식평가 결과( $r=.30, p<0.01$ )와 약한 정도의, 그러나 통계적으로 유의미한 수준의 상관관계를 보였다. 반면, 지연 사후 형태인식평가 결과와는 통계적으로 유의미한 관련성이 없는 것으로 드러났다( $p=0.09$ ). 이를 각 집단별로 살펴보면, 표 10에 제시된 바와 같이, 어휘 배열하여 문장 만들기를 수행한 집단 2의 경우에서만 통계적으로 유의미한 상관관계를 확인할 수 있었다. 특히 집단 2에 속한 학생들의 경우 작업 기억과 어휘 학습 간 상관관계가 중간 정도 크기의 값으로 나타나고 있어 주목할 만하였다( $r=0.40\sim0.69$ ). 보다 구체적으로 보자면, 작업 기억 검사 결과가 사후 어휘 형태인식평가 결과( $r=.58, p<0.01$ ) 및 지연 사후 어휘 형태인식평가 결과( $r=.55, p<0.01$ ), 그리고 지연 사후 의미인식평가 결과와 통계적으로 유의미한 상관관계를 보였다( $r=.69, p<0.01$ ). 또한 사후 어휘 의미인식평가 결과와의 상관관계 역시 유의수준에 다가가고 있음을 확인할 수 있었다( $r=.40, p=0.06$ ).

**표 10.** 작업 기억과 어휘 평가 결과 간 상관관계

			작업 기억		
			집단1 (빈칸 채우기)	집단2 (어휘 배열하기)	집단3 (문장 쓰기)
사후	형태인식	Pearson <i>r</i>	.06	.58	.22
		<i>Sig.</i>	.81	.00	.32
	의미인식	Pearson <i>r</i>	.16	.40	.14
		<i>Sig.</i>	.49	.06	.52
지연 사후	형태인식	Pearson <i>r</i>	.04	.55	-.07
		<i>Sig.</i>	.86	.01	.77
	의미인식	Pearson <i>r</i>	.30	.69	-.09
		<i>Sig.</i>	.19	.00	.70

# 5. 논의 및 결론

본 연구는 고등학교 영어 학습자 총 66명으로 하여금 관여 부담 수준이 서로 다른 세 가지의 과업을 수행하게 하였고, 그로부터 10개의 목표 어휘를 어느 정도 학습하게 되는지를 살펴보았다. 또한 각각의 과업을 수행함에 있어 읽기 폭 과제로 측정해본 학습자의 작업 기억이 그들의 단기 및 장기 어휘 습득과 어떠한 상관관계를 가지게 되는지를 알아보았다. 표 11은 본 연구로부터의 주요 결과의 요약이다.

**표 11.** 주요 연구 결과 요약

		형태인식	집단 1 = 집단 2 = 집단 3
연구문제 1: 관여 부담 수준에 따른 어휘 학습	사후 검사		집단 3 > 집단 1
		의미인식	집단 3 > 집단 2 집단 1 = 집단 2
	지연 사후 검사	형태인식	집단 1 = 집단 2 = 집단 3
		의미인식	집단 1 = 집단 2 = 집단 3
연구문제 2: 작업 기억과 어휘 습득 간 상관관계	<ul style="list-style-type: none"> <li>· 사후 형태인식 &amp; 사후 의미인식에서 <math>p&lt;0.05</math></li> <li>· 지연 사후 의미인식에서 <math>p&lt;0.01</math></li> <li>· 집단 2에서 <math>r=0.40\sim r=0.69^{**}</math></li> </ul>		

사후 어휘 형태인식평가에서는 과업의 관여 부담 수준을 다르게 적용한 세 집단의 어휘 학습 효과에 통계적으로 유의미한 차이가 없었으나, 사후 어휘 의미인식평가에서는 문장 쓰기 집단이 다른 두 집단에 비해 통계적으로 유의미한 수준에서 우수한 수행을 보였다. 또한 빈칸 채우기 집단을 기준으로 하였을 때 문장 쓰기 집단에서  $d=0.82$ 의 큰 효과 크기가 목격되었다. 이처럼 단기적인 어휘 습득 결과는 어휘의 형태에 비해 의미 인식에 있어 집단 간에 큰 차이를 보였는데, 학습자가 목표 어휘를 활용하여 문장을 직접 생산하는 과업이 적어도 단기적인 의미 학습에 있어 더욱 효과적인 교수학습 효과를 이끌게 됨을 시사한다. 과업의 관여 부담 수준이 가장 높은 문장 쓰기 집단이 다른 두 집단에 비해 단기적인 어휘 의미 습득에 있어 우수한 결과를 보인 것은, 문장 쓰기 과업이 빈칸 채우기 과업에 비해 어휘 습득에 더욱 효과적임을 입증한 Zou(2017), Yang 외 3인(2017)의 연구와 맥을 같이 하며, Hulstijn과 Laufer(2001)의 관여 부담 가설을 부분적으로 지지하는 것이라 볼 수 있다. 문장 쓰기 과업은 생산적인 활동을, 빈칸 채우기 활동은 적절한 어휘를 선택하는 수용적인 활동을 요구하는데, 여러 연구에서 더 많은 인지적 노력을 요구하는 생산적인 활동이 수용적인 활동에 비해 어휘 습득에 더욱 효과적임을 입증해왔다(e.g., Huang et al., 2012; Nation, 2001; Yang et al.,

2017). 특히, 쓰기와 같은 생산적인 과업, 그 중에서도 새로운 맥락에서 사용된 어휘가 비생산적인 과업에서 연습된 어휘에 비해 더욱 잘 기억된다는 연구들(Ellis & He, 1999; Hulstijn & Laufer, 2001; Hulstijn & Trompetter, 1998)과 맥을 같이 하는 부분으로, 교사들은 어휘 과업을 계획함에 있어 이 점을 충분히 고려할 필요가 있다. 한국의 교실 환경에서는 주어진 문맥에 적절한 어휘를 채워 넣거나 어휘를 배열하여 문장을 완성시키는 방식의 수동적인 활동을 많이 활용하고 있으나, 어휘 습득 효과의 제고를 위해서는 학습자가 스스로 생각하여 맥락과 문장을 생산해낼 수 있도록 유도하는 활동이 보다 적극적으로 도입될 필요가 있어 보인다.

한편, 본 연구의 결과는 어휘 배열하여 문장 만들기를 문장 쓰기의 대체 과업으로 가정하고 실시한 Sung(2013)과 Choi(2017)의 연구에서 어휘 배열하여 문장 만들기가 빈칸 채우기에 비해 과업의 관여 부담 수준이 높다고 가정하였음에도 불구하고, 어휘 학습 효과 측면에서 통계적으로 유의미한 차이를 발견하지 못했던 이유에 대해 설명을 제공한다. 주어진 어휘를 배열하여 문장을 완성시키는 활동이 학습자 스스로 맥락을 생산하는 문장 쓰기 활동과 과업의 관여 부담 수준 측면에서 동일하다고 볼 수 없으며, 오히려 빈칸 채우기에 상응하는 관여 부담 수준을 가지는 것으로 추론해 볼 수 있을 것이다.

한편, 지연 사후 어휘 형태 및 의미인식평가 모두에서 세 집단 간에 통계적으로 유의미한 차이가 발견되지는 않았다. 어휘 배열하여 문장 만들기와 문장 쓰기 집단은 사전 사용 및 짝 활동으로 어휘 검색을 하며 목표 어휘의 형태와 의미를 연결시킬 기회를 한 차례 가진 다음 새로운 맥락에서 어휘를 배열하여 문장을 완성하거나 스스로 문장을 써 보는 활동을 한 반면, 빈칸 채우기 집단은 지문의 맥락에 비추어 적절한 어휘를 채워 넣기 위해 여러 차례 반복하여 목표 어휘의 의미와 적절성을 평가해 보았을 것이다. 이처럼 빈칸 채우기 활동에서 목표 어휘를 반복적으로 마주치고 맥락에 대해 평가하게 한 것이 목표 어휘를 장기적으로 기억하는 데 도움을 주었고(cf. Yang et al., 2017), 이로 인해 결국 세 집단 간 장기 학습 효과에 있어 도드라진 차이가 목격되지 않았을 것이라 추론된다. 이는 세 개의 빈칸 채우기 활동이 하나의 문장 쓰기 활동에 비해 학습에 더욱 효과적임을 보인 Folse(2006)의 연구에서 어휘 처리의 깊이보다 요구되는 어휘 인출(retrieval)의 횟수가 중요하다고 하였던 결론을 상기시킨다. Hulstijn과 Laufer (2001) 역시 관여 부담 수준이 낮은 여러 개의 과업을 수행하는 것과 관여 부담 수준이 높은 하나의 과업을 수행하는 것 중 어느 것이 어휘 학습에 더욱 효과적일지를 알아보는 후속 연구가 필요하다고 제언한 바 있다. 또 한편, 본 연구에서는 지연 사후 검사를 처치 후 일주일 뒤에 실시하였음을 고려한다면, 후속 연구를 통해 그 시기를 조정하였을 때 교수학습 효과가 어떻게 달라지는지를 탐색해 볼 필요도 있어 보인다.

검색 조건을 달리한 빈칸 채우기(N1, S0, E1=2)와 어휘 배열하여 문장 만들기(N1, S1, E1=3) 과업 상호 간에는 사후 및 지연 사후 검사 모두에서 어휘 학습에 유의미한 차이가 없었던 반면, 검색 조건은 동일하되 평가의 정도를 달리한 어휘 배열하여 문장



만들기(N1, S1, E1=3)와 문장 쓰기(N1, S1, E2=4) 과업 간에는 사후 의미인식평가에서 유의미한 차이가 나타난 것으로 보아, 적어도 단기적인 어휘 의미인식에 있어서는 검색에 비해 평가의 차이가 학습에 더욱 큰 차이를 가져오게 됨을 알 수 있었다. 이는 평가에 비해 검색의 영향력이 보다 더 크다고 한 Laufer(2003)의 실험과는 대조적인 결과이며, 검색에 비해 평가의 차이가 어휘 학습에 더욱 큰 영향을 미친다고 해석한 Park(2011)의 연구와는 일치하는 결과로, 향후 검색 조건에 대한 더욱 세밀한 연구가 후속되어야 할 것으로 보인다. Kim(2011)은 Hulstijn과 Laufer(2001)가 관여 부담 수준의 모든 요소(필요, 검색, 평가)를 동등하게 취급하여 각 요소의 점수의 합을 관여 부담 지수로 제시하였으나, 개별 요소들이 전체 관여 부담 지수에 기여하는 정도가 실제로는 다를 수 있음을 언급한 바 있다. 예를 들어, 단기적인 어휘 습득에 있어서 강한 정도의 평가(E2)가 가장 영향력 있는 요소일 수 있으며, 검색의 차이와 평가의 차이가 갖는 무게가 어휘 습득에 있어 다른 정도로 작용할 수 있는 것이다. 따라서 관여 부담 수준의 세 요소 각각이 전체적인 관여 부담 지수에 기여하는 정도를 균일하게 ‘없음(0),’ ‘중간 정도(1),’ ‘강한 정도(2)’ 등으로 제시할 것이 아니라, 개별 요소가 갖는 무게와 관여 부담 지수의 산출 방식을 더욱 정교하게 다듬을 필요가 있어 보인다.

어휘 추론에 대한 확신 지수의 분포는 사후 형태인식평가 답안에 대해서는 73%가, 지연 사후 형태인식평가 답안에 대해서는 57%가 자신의 추론에 대해 확신 또는 매우 확신하는 것으로 드러났다. 이처럼 상당수의 참여자들이 어휘 형태에 대한 자신의 응답에 대해 확신하는 것으로 보였으나, 대응표본 *t*-검정 결과 사후 및 지연 사후 확신 정도의 차이가 통계적으로 유의미한 것으로 나타났다. 이는 참여자들이 과업 수행 후 어휘 형태를 단기적으로 기억하는 데에는 어느 정도의 자신감을 보였으나 그것이 일주일 후 까지 이어지지는 못했음을 의미한다.

기존의 연구에서 한 걸음 더 나아가 본 연구에서는 개인차 요인으로서의 작업 기억이 서로 다른 관여 부담 수준 하에서 다양한 과업을 수행함으로써 얻게 되는 어휘 학습의 결과와 어떠한 상관관계를 가지게 되는지를 또한 살펴보았다. 그 결과, 작업 기억이 목표 어휘의 형태와 의미 양 측면에서의 단기 기억 및 의미 측면의 장기 기억과 통계적으로 유의미한 수준에서 상관관계를 보이고 있었다. 특히 어휘 배열하여 문장 만들기 집단의 경우에서 작업 기억 검사 결과와 어휘 학습 결과 간의 상관관계가 도드라지게 나타난 점은 특기할만 하였다( $r=0.40\sim0.69$ ). 본 연구가 시사하는 바는 과업을 통한 우연적 어휘 학습 맥락에서 목표 어휘를 기억함에 있어 작업 기억이라는 개인차 변인이 유의미하게 작동할 수 있다는 점이다. 이와 비슷한 맥락에서 Choi(2016)는 작업 기억이 우연적 어휘 학습과 장기 기억에 유의미한 차이를 가져온다고 주장하였고, Lim(2017)은 작업 기억이 문맥을 통해 어휘를 추론하고 기억하는 데 긍정적인 역할을 한다고 설명한 바 있다. 이러한 점을 고려할 때 상관관계 분석을 넘어 작업 기억과 학습 간의 인과관계에 더욱 초점을 맞출 수 있는 연구가 후속되기를 기대한다. 더불어 작업 기억 이외에도 우연적 어휘 학습에 영향을 미칠 수 있는 다양한 개인차 변인에 대한 고려도 계속되어야

할 것이다.

본 연구의 결과는 연구 참여자의 규모나 연구 기간 등 여러 측면에서의 제한적 상황을 고려하여 조심스럽게 해석되어야 옳다. 예를 들면, 본 연구에서 활용한 읽기 폭 과제의 경우, 학교 현장의 시·공간적 제약을 고려하여 집단으로 시행하였다. 그러나 읽기 폭 과제가 작업 기억과 관련한 개인차 변인을 측정하고자 하는 목적을 가지게 됨을 고려할 때 개별적으로 자료를 수집하는 것이 보다 타당한 측면이 있을 수도 있겠다는 점에 유의해야 할 것이다. 그럼에도 불구하고 의미와 의사소통을 중심으로 하는 교수학습의 환경 조성을 강조하는 시대적 흐름에 비추어 우연적 어휘 학습의 가능성을 타진하였다는 점, 어휘 학습에 있어 큰 비중을 점하고 있는 관여 부담 가설과 관련하여 후속 연구를 위한 중요한 이론적 제언을 제공하고 있다는 점에서 본 연구는 의의를 가진다고 하겠다.

## References

- Atkins, P. W. B., & Baddeley, A. D. (1998). Working memory and distributed vocabulary learning. *Applied Psycholinguistics*, 19, 537-552.
- Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), *Recent advances in learning and motivation* (pp. 47-90). New York: Academic Press.
- Cho, K., & Krashen, S. (1994). Acquisition of vocabulary from the Sweet Valley Kids series: Adult ESL acquisition. *Journal of Reading*, 37, 662-667.
- Choi, H. K. (2015). *The impacts of task-induced involvement load on vocabulary acquisition and affective domain of Korean high school students* (Master's thesis) (in Korean). Ewha Womans University, Seoul.
- Choi, M. A. (2017). *The impacts of task-induced involvement load and English proficiency on vocabulary acquisition of Korean high school students* (Master's thesis) (in Korean). Hankuk University of Foreign Studies, Seoul.
- Choi, S. M. (2016). Gloss, working memory capacity, and L2 vocabulary acquisition (in Korean). *(The) Journal of Modern British and American Language and Literature*, 34(3), 247-270.
- Chun, D. M., & Payne, J. S. (2004). What makes students click: Working memory and look-up behavior. *System*, 32, 481-503.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York: Lawrence Erlbaum.
- Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin and Review*, 12, 769-786.
- Dupuy, B., & Krashen, S. (1993). Incidental vocabulary acquisition in French as a foreign language. *Applied Language Learning*, 4, 55-63.

- Eckerth, J., & Tavakoli, P. (2012). The effects of word exposure frequency and elaboration of word processing on incidental L2 vocabulary acquisition through reading. *Language Teaching Research*, 16(2), 227-252.
- Ellis, R., & He, X. (1999). The roles of modified input and output in the incidental acquisition of word meanings. *Studies in Second Language Acquisition*, 21(2), 285-301.
- Folse, K. S. (2006). The effect of type of written exercise on L2 vocabulary retention. *TESOL Quarterly*, 40, 273-293.
- Horst, M., Cobb, T., & Meara, P. (1998). Beyond a clock organ: Acquiring second language vocabulary through reading. *Reading in a Foreign Language*, 11, 207-223.
- Huang, S., Eslami, Z., & Willson, V. (2012). The effects of task involvement load on L2 incidental vocabulary learning: A meta-analytic study. *The Modern Language Journal*, 96(4), 544-557.
- Hulstijn, J. H. (1992). Retention of inferred and given word meanings: Experiments in incidental vocabulary learning. In P. J. Arnaud & H. Béjoint (Eds.), *Vocabulary and applied linguistics* (pp. 113-125). London: Macmillan.
- Hulstijn, J., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning*, 51(3), 539-558.
- Hulstijn, J., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *Modern Language Journal*, 80, 327-339.
- Hulstijn, J. H., & Trompetter, P. (1998). Incidental learning of second language vocabulary in computer-assisted reading and writing tasks. In D. Albrechtsen, B. Henriksen, I. M. Mees, & E. Poulsen (Eds.), *Perspectives on foreign and second language pedagogy* (pp. 191-200). Odense, Denmark: Odense University Press.
- Keating, G. (2008). Task effectiveness and word learning in a second language: The involvement load hypothesis on trial. *Language Teaching Research*, 12, 365-386.
- Kim, J. Y. (2015). The effects of task-induced involvement and learner's English proficiency on English vocabulary learning (in Korean). *English Language Teaching*, 27(3), 125-149.
- Kim, Y. (2008). The role of task-induced involvement and learner proficiency in L2 vocabulary acquisition. *Language Learning*, 58, 285-325.
- Kim, Y. (2011). The role of task-induced involvement and learner proficiency in L2 vocabulary acquisition. *Language Learning*, 61, 100-140.
- Laufer, B. (1989). What percentage of text-lexis is essential for comprehension? In C. Lauren & M. Nordman (Eds.), *Special language: From humans thinking to thinking machines* (pp. 316-323). Clevedon, UK: Multilingual Matters.
- Laufer, B. (2000). Electronic dictionaries and incidental vocabulary acquisition: Does technology make a difference? In U. Heid, S. Evert, E. Lehmann, & C. Rohrer (Eds.), *EURALEX* (pp. 849-854). Stuttgart: Stuttgart University Press.
- Laufer, B. (2003). Vocabulary acquisition in a second language: Do learners really acquire most vocabulary by reading? Some empirical evidence. *Canadian Modern Language Review*,

- Lee, S.-K. (2014). Individual differences in language learning aptitude: issues and possibilities. *Foreign Language Education*, 15, 59-78.
- Linck, J., Osthus, P., Koeth, J., & Bunting, M. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin and Review*, 21(4), 861-883.
- Lim, S. M. (2017). *Effects of text type, vocabulary knowledge, language learning aptitude on incidental vocabulary learning* (Master's thesis) (in Korean). Yonsei University, Seoul.
- Lu, M. (2013). Effects of four vocabulary exercises on facilitating learning vocabulary meaning, form, and use. *TESOL Quarterly*, 47, 167-176.
- Malone, J. (2018). Incidental vocabulary learning in SLA: Effects of frequency, aural enhancement, and working memory. *Studies in Second Language Acquisition*, 40, 651-675.
- Marmol, A., & Sanchez-Lafuente, A. (2013). The involvement load hypothesis: The effect on vocabulary learning in primary education. *RESLA*, 26, 11-24.
- Marsick, V. J., & Watkins, K. (2016). *Informal and incidental learning in the workplace*. New York: Routledge.
- Masgoret, A.-M., & Gardner, R. C. (2003). Attitudes, motivation, and second language learning: A meta-analysis of studies by Gardner and associates. *Language Learning*, 53, 123-163.
- Nassaji, H., & Hu, M. (2012). The relationship between task-induced involvement load and learning new words from context. *IRAL*, 50, 69-86.
- Nation, I. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Ortega, L. (2009). *Understanding second language acquisition*. London: Arnold.
- Paribakht, T. S., & Wesche, M. (1993). The relationship between reading comprehension and second language development in a comprehension-based ESL program. *TESL Canada Journal*, 11, 9-29.
- Park, J. H. (2011). *The role of task-induced involvement load in vocabulary acquisition of Korean college students* (Master's thesis). Ewha Womans University, Seoul.
- Schmitt, N. (2008). Instructed second language learning. *Language Teaching Research*, 12, 329-363.
- Soleimani, H., & Rahmanian, M. (2014). The role of language glossing in a rooted theory: The involvement load hypothesis. *International Journal of Applied Linguistics and English Literature*, 3(4), 6-13.
- Sung, H. M. (2013). *The effects of task-induced involvement load on EFL high school learners' vocabulary learning* (Doctoral dissertation). Sookmyung Women's University, Seoul.
- Watanabe, Y. (1997). Input, intake, and retention: Effects of increased processing on incidental learning of foreign language vocabulary. *Studies in Second Language Acquisition*, 19, 287-307.
- Yang, Y., Shintani, N., Li, S., & Zhang, Y. (2017). The effectiveness of post-reading word-

cused activities and their associations with working memory. *System*, 70, 38-49.

Zou, D. (2017). Vocabulary acquisition through cloze exercises, sentence-writing and composition-writing: Extending the evaluation component of the involvement load hypothesis. *Language Teaching Research*, 21(1), 54-75.

이옥희  
대학원생  
영어교육과  
한국교원대학교  
28173 충북 청주시 흥덕구 강내면 태성탑연로 250  
E-mail: okighost@hanmail.net

이상기  
교수  
영어교육과  
한국교원대학교  
28173 충북 청주시 흥덕구 강내면 태성탑연로 250  
E-mail: slee@knue.ac.kr

접수일자 : 2021. 1. 20  
수정본 접수 : 2021. 4. 8  
게재결정 : 2021. 4. 19

## The Definition and Characteristics of Grit

Before she was a psychologist at the University of Pennsylvania, Angela Duckworth was a middle school math teacher. She found that some of her sharpest students were not performing well, while others who struggled through each lesson were getting A's. "What was (1) revelatory to me was not that effort matters — everybody knows that effort matters," Angela said. "What I got to know was how much it matters. Talent counts, but effort counts twice," insisted Angela.

This revelation led her back to school, this time as a graduate student, to study a character trait she calls grit. According to the Merriam-Webster dictionary, grit is defined as "firmness of character; (2) indomitable spirit." Based on her studies, Angela tweaked this definition to be "(3) perseverance and passion for long-term goals." It is about being (4) resilient and hardworking, so much so that you are willing to continue on in the face of difficulties, obstacles and failures. In her research, she looked at public school students, sales representatives, teachers, and others from various backgrounds. Across all these domains, she found those who did the best were not always the smartest, but they were the grittiest.

Angela believes grit is made up of four key psychological assets: interest, practice, purpose, and hope. Most of us can quickly grasp the importance of patience for success. But that is only half of the equation, according to Angela. The other half is an (5) abiding interest and passion for a singular pursuit. Gritty people wake up thinking about the same questions they go to bed thinking about. Rather than being "discovered" through (6) angst-ridden (7) introspection, psychologists believe interests are developed and deepened over time through continuing engagement with a pursuit.

Psychologists have found that people at the (8) tippy-top of their fields engage with their work differently. You have probably heard of the "10,000 hour rule," the average number of hours it takes an expert to become an expert. That is based on research by K. Anders Ericsson, who found that experts do an intensive-kind of practice called "deliberate practice." It is a purposeful and systematic practice that requires focused attention, while regular practice might include mindless repetitions.

In her interviews with people at the top of their fields, Angela found they describe their work as being (9) imbued with meaning and a desire to help others. She finds

they are driven by a purpose beyond the self.

Gritty people also have hope. They are optimistic about the future and their ability to improve and affect change. To cultivate hope, Angela points to Carol Dweck's work on growth mindset, the belief that intelligence is not fixed but can change over time.

But other research has also pointed to a potential downside to grit. Like stubbornness, too much grit can keep us sticking to goals, ideas, or relationships that should be abandoned. Psychologist Gale Lucas and her colleagues found in one experiment that gritty individuals will persist in trying to solve unsolvable puzzles at a financial cost. Grit does not give you insight into when it will help you prevail and when it will keep you stuck in a (10) dead-end.