

한국인 영어 학습자의 관계대명사 사용에 관한 코퍼스 기반 연구*

이정은 (동국대학교)
신정아 (동국대학교)

Lee, Kyung Eun and Shin, Jeong-Ah. 2017. A corpus-based study on the use of relative pronouns by Korean learners of English. *Korean Journal of English Language and Linguistics* 17-4, 817-837. This study examined the use of relative pronouns, *who* and *which* (subdivided into restrictive and non-restrictive clauses), by high school learners and college freshman learners of English in Korea in terms of their frequency and errors displayed in their writings. Then, the frequency rates found were compared with those of high school English textbooks. The corpus analyses showed that in writing, high school learners tend to overuse the *which*-clauses with many grammatical errors in comparison to college freshman learners. However, corpora from both learner groups revealed a significant overuse of restrictive relative clauses compared to the textbook corpus, while their use of non-restrictive relative clauses was rarely shown. These results were discussed in terms of Keenan and Comrie's (1977) Noun Phrase Accessibility Hierarchy and learners' avoidance (Kleinmann, 1977).

Keywords: L2 learner corpus, frequency, errors, relative pronouns, restrictive, non-restrictive

1. 서론

최근에 코퍼스와 관련하여 학습자 코퍼스(learner corpus)를 다 양한 영어 교육 분야에서 활용하려는 시도가 이루어지고 있다(이석 재와 정채관 2014). 지금까지의 한국인 영어 학습자 코퍼스를 활 용한 연구는 부사구, 정형화된 연결어구, 연어 등을 살펴보는 영어 어휘 연구가 대부분이었다(e.g., 이은주 2009, 2010, 이정은과 김동

* 이 논문은 제1저자의 석사학위논문의 자료를 바탕으로 수정 보완하여 작성한 것임. 제1저자: 이정은, 교신저자: 신정아

규 2013, 한나래와 이수화 2009). 학습자 코퍼스를 이용하여 학습자의 영어 문법 구문의 사용을 살펴본 연구는 매우 제한적이었다 (e.g., Shin 2011).

본 연구는 학습자들이 어렵게 생각하는 문법 구문 중 하나인 영어 관계대명사 구문(Celce-Murcia, Larsen-Freeman and Williams 1983)을 한국인 고등학생과 대학신입생 학습자 코퍼스를 이용하여 비교하였다. 기존 연구 중에는 관계대명사에 있어서 습득에 관한 연구와 교과서를 분석한 연구, 고등학생 집단과 대학생 집단을 각기 따로 연구 대상으로 삼은 연구들은 있었으나, 두 개의 학습자 집단과 교과서를 비교하는 선행연구는 없었다. 따라서 본 연구에서는 고등학생과 대학생 신입생 집단으로 두 학습자 코퍼스와 교과서 코퍼스와 비교한 결과 값이 어떤 시사점을 나타내고 있는지 알아보려고 한다.

특히 본 연구에서는 고등학교 1,2학년들의 영작문에 관계대명사 *who*와 *which*의 사용 빈도가 어떻게 나타나는지 분석하고, 어떠한 형태의 오류들이 나타나는지 알아보려고 한다. 그리고 고등학교를 졸업하고 대학에 입학하는 신입생들의 영작문에서의 빈도와 오류는 고등학생 1,2학년과 어떻게 다른지 살펴보고자 한다. 마지막으로 현행 고등학교 교과서에 제시된 관계대명사 *who*와 *which*의 쓰임이 앞의 두 집단에 어느 정도 반영이 되어 있는가를 보기 위해서, 관계대명사 제한적 용법뿐만 아니라 계속적 용법으로 쓰인 빈도는 어떻게 나타나는지, 그 사용의 차이가 얼마큼 나타나는지 비교하는 것도 본 연구의 목적이다. 본 연구의 구체적 질문은 다음 두 가지와 같다.

첫째, 고등학교 1,2학년과 대학 신입생 영작문에서 한정적 용법과 계속적 용법 관계대명사 *who*와 *which*의 사용 빈도와 오류의 유형이 차이가 나는가?

둘째, 고등학교 영어 교과서 본문을 코퍼스로 하여 한정적 용법과 계속적 용법의 관계대명사의 *who*와 *which*의 빈도가 앞서 살펴본 두 학습자 영작문 집단에서의 사용 빈도와 비교하여 어떠한 특징을 가지고 있는가?

2. 선행 연구 및 이론적 배경

2.1 L2 학습자의 관계대명사 습득 가설

L2 학습자들을 대상으로 한 관계대명사 습득에 관한 연구들에서는 관계대명사의 습득 난이도에 따라 학습자들이 어떠한 형태의 관계대명사를 어려워하는지 나타나있다. 우선, Sheldon(1974)이 제안한 평행 기능(parallel function) 가설은 관계대명사를 습득할 때 명사구인 선행사의 문법적인 기능과 관계절 안에서의 관계대명사 문법적 기능이 일치할 때 습득이 쉽다는 것이다. 예를 들어, “The man [who needed a job] helped the woman”의 문장은 Subject Head Noun, Subject target of relativization의 첫 글자를 이용하여 [SS]유형으로 분류하였고, 마찬가지로 “The dog [that the woman owned] bit the cat”와 같은 문장은 Subject Head Noun, Object target of relativization인 [SO] 유형으로, “The saw the boy [who entered the room]”은 [OS] 유형으로, “A man bought the clock [that the woman wanted]”은 [OO] 유형으로 나누어 앞서 제시한 두 역할이 일치 하지 않을 경우 관계대명사 습득이 어렵다는 주장이다.

한편, Kuno(1974)가 제시한 인지 난이도 가설(Perceptual Difficulty Hypothesis, PDH)은 인간의 인지처리능력을 근거로, 학습자의 문장 이해에 있어서 작용하는 인간의 기억체계와 문장구조 사이의 관계에 중점을 두었다. 인간의 작업기억 능력의 한계로 주절의 중간에 한정적 용법으로 쓰인 관계절이 삽입되었을 때, 인지 처리 과정이 방해를 받게 되어 습득이 어렵다는 것이다. 이 가설에서의 주요한 논지는 관계절의 위치이고 그에 따라 관계절이 삽입된 문장의 이해와 습득이 영향을 받음을 의미한다. 따라서 인지 난이도 가설에 따르면, 습득 위계는 OS / OO > SS / SO라고 할 수 있다.

또한, Keenan과 Comrie(1977)는 여러 언어의 유형적 대조(typological comparison)를 바탕으로 관계대명사 습득의 위계를 정리하여 명사구 접근성 위계 가설(Noun Phrase Accessibility Hierarchy, NPAH)을 제안하였다. 이 가설은 50여개의 언어들의 문

법적인 구조를 비교하고, 관계절이 가지는 보편적인 속성과 습득 순서를 다루었다. 즉, 이 가설에서는 관계절 그 자체 내에서 관계대명사의 기능에 따라 Subj > Direct Obj > Indirect Obj > Oblique > Genitive > Object of Comparison로 위계를 나누었다.

그리고 Hamilton(1994)이 주장한 주어-목적어 위계 가설(SO Hierarchy Hypothesis, SOHH)에서는 인지 난이도 가설(PDG)과 명사구 접근성 위계 가설(NPAH)을 합친 혼합형 가설로 O'Grady(1987, 1999)의 처리 비연속성(processing discontinuity)의 개념을 근간으로 하고 있다. 처리 비연속성은 선행사와 관계절 내의 이동의 흔적(trace) 사이의 절 또는 구의 수가 많으면 구조적 거리도 증가하게 되어 관계절의 습득 어려움 또한 증가한다고 설명한다. 그래서 여기서의 위계를 도식화 하면 OS > OO / SS > SO와 같다.

위 살펴본 위계 가설과 관련한 학습자의 관계대명사 습득 선행연구들을 살펴보면, 3.8세부터 5.5세의 영어를 모국어로 사용하는 어린이들을 대상으로 습득 양상을 알아본 Sheldon(1974)의 연구는 4가지의 유형 중 관계대명사와 선행사의 기능이 같은 유형을 쉽게 습득한다고 가정했다. Kuno(1974)의 연구에서는 관계절의 위치에 따른 모국어의 습득 어려움을 주장하였지만, Doughty(1991)는 ESL 학습자를 대상으로 실험한 연구에서는 모국어 습득과 다른 양상을 보인다고 하였다.

특히, 한국인 영어 학습자들을 대상으로 한 연구의 결과는 다양하였다. 고미숙(2007)의 연구에서는 고등학생이 문법성 판단과제에서 인지 난이도 위계(PDH)를 따르는 것으로 나타났지만 영작 과제에 있어서는 접근성 위계(AHH)를 따르는 것으로 나타났다. 김태희(2009)의 연구에서는 고등학생들이 접근성 위계(AHH)를 따른다는 결과를 얻었다. 이권창(2014)은 고등학생들이 인지 난이도 위계(PDH)를 따르는 결과를 얻기도 했다. 김남국(2014)은 대학생 55명을 대상으로 산출 과제와 이해 과제 두 과제를 제시한 결과, 한국 대학생들의 관계대명사 습득 순서 예측 가설로서 인지 난이도 가설(PDH)과 주어-목적어 위계 가설(SOHH)을 따르는 것으로 확인되었으며, 과제 유형에 따라 습득 순서가 다르다는 결과를 얻었다. 이렇게 선행연구에서 관계대명사 습득 난이도 위계 가설은 각

연구의 대상자와 언어 습득 환경이나 각자의 실험 과제에 따라 상이한 결과를 나타냈다.

2.2 관계사의 코퍼스 연구

관계사 관련한 코퍼스 연구로는 교과서를 코퍼스로 사용한 연구와 그 빈도를 원어인 코퍼스와 비교하는 연구가 대부분이었다. 예를 들어, 김효은(2007)은 고등학교 영어 교과서에 관계대명사의 구문이 얼마나 적절하게 제시되고 있는지를 분석하기 위해, Longman Grammar of spoken and written English(LSWE Corpus)와 비교하여 얼마나 실제 영어를 반영하고 있는지를 알아보았다. 이 연구는 7차 교육과정에 해당하는 고등학교 1학년 영어교과서 15권의 본문 내용을 코퍼스로 구축하여 *who*, *which*, *that*을 찾고, 관계대명사 아닌 경우의 수를 수작업으로 제거하였다. 그 결과, 관계대명사 *that*은 LSWE Corpus와 교과서 모두에서 높은 빈도를 나타냈으며, 교과서에서 관계대명사 *who*는 사람을 선행 명사로 가지며 주격으로 사용되는 경우가 지배적이었으며, LSWE Corpus에서도 높은 빈도로 *who*가 주격으로 사용되었다. 관계대명사 *which*의 경우 교과서에의 사용 빈도는 가장 낮았으나 LSWE Corpus에서는 사용 영역에 따라 *which*가 *who*만큼 자주 사용되었다. 마지막으로 교과서에서의 관계대명사 *that* 사용 시, 선행사가 사람을 나타내는 경우는 드물지만 LSWE Corpus에서는 *who*만큼 *that*이 사람을 가리키는 선행명사를 가지는 경우가 많았다는 결론을 얻을 수 있었다.

장진임(2011)의 연구도 앞선 연구와 비슷하게 2007 교육과정에 따른 고등학교 3학년 영어 교과서에 관계대명사의 구문이 적절하게 또 얼마나 실제 사용을 반영하고 있는지는 알아보았다. 교과서의 읽기자료와 고등학교 3학년 영어 교과서 5종을 코퍼스로 구축하였고 사용 현황을 유형별, 용법별, 격별 순으로 분석하였다. 그 결과 관계대명사 *that*의 사용빈도는 교과서와 교과서의 읽기 자료사이에서 그 차이가 없음을 보였고 나머지 관계대명사 *who*와 *which*에 있어서는 유형별, 용법별, 격별로 다른 결과를 나타내었다. 이 중 용법별 분석에 있어서 주격 관계대명사 *who*와 *which*를 분석했을 때 *who*는 제한절로 사용되는 빈도가 높게 나타났으며, *which*는 비제

한절로 사용되는 빈도가 높게 나타났다는 결과를 얻었다.

이제까지 중등 영어 교과서를 코퍼스로 구축하여 빈도를 조사한 연구는 있었으나, 학습자 코퍼스로 관계대명사 쓰임을 비교한 논문은 없었다. 그래서 본 연구는 L2 학습자들의 영작문 코퍼스를 이용하여 학습자들의 관계대명사의 습득의 특징을 규명하고, 더불어 L2 학습자들이 기본적으로 노출되는 기본 교재인 현행 학교 교과서의 본문과 비교하여 그 반영 정도 또한 비교해 보고자 한다. 그리고 학습자 코퍼스에 나타난 오류의 패턴을 분석 후 어느 위계 가설을 지지하는 결과를 나타내는지 알아보고자 한다.

3. 연구 방법

3.1 연구 대상

본 연구에서는 총 3가지 종류의 텍스트를 코퍼스로 사용하였다. 먼저, 서울 소재의 고등학교에서 수준이 다양한 1, 2학년 학생 111명의 영작문으로 코퍼스를 구축하였다. 111명의 학생들은 모두 여학생으로 이루어져 있으며, 영작문과 말하기의 주제는 각각 이웃과의 단절을 극복 할 수 있는 방법, 내가 소개하고 싶은 관광지로 자유롭게 제한된 A4 내외의 분량으로 수업 시간 내에 기술하였다. 영작문 작성은 전자기기를 제외한 다른 종류의 사전검색을 허용하는 환경에서 이루어졌다. 대학생 영작문으로 사용한 코퍼스는 연구자가 연구용으로 사용 허가를 받은 연세 영어 학습자 코퍼스 YELC이다. YELC는 2011년 연세대학교에 입학 예정인 총 3,286명을 대상으로 치러진 영작문 시험에서 100자 이내로 작성하는 영작문과 200자 이내로 작성하는 영작문을 모아서 구축한 한국인 영어 학습자 코퍼스 이다(이석재와 정채관 2014). 본 연구에서 사용 된 세 번째 코퍼스는 고등학교 영어 교과서로, 이는 2009 개정 교육과정에 따른 것으로, 현재 학교에서 학습자들이 접할 수 있는 교재이다. 총 7개의 출판사의 교재를 사용 하였으며, 교재는 기본 교재인 고등학교 영어 I, 영어 II와 실용 영어에 해당하는 실용영어 I, 실용영어 II, 실용영어 독해와 작문, 독해와 작문의 221개의 본문을 사용

하였다. 구축된 코퍼스에는 본문 이외의 듣기, 말하기 지문들은 포함하지 않았다. 대다수의 학교현장에서 접하는 것이 본문이라 여겨 본문만 코퍼스 구축에 사용하였다. 본문의 형태가 교과서 마다, 단원마다 다르게 나타나 간혹 구어체나 대화체의 형식으로 나타난 본문이 존재하였지만 이는 본문의 형태라고 간주하고 그대로 포함하였다.

3.2 연구 도구

본 연구에서 고등학생과 대학생의 영작문을 코퍼스로 분석하고, 현 영어 교과서를 분석함에 있어서 사용한 코퍼스 프로그램은 개인, 학교 또는 대학에 적합하게 만들어졌으며 윈도우기반 시스템에서 실행 가능한 무료 애플리케이션인 AntConc3.4.4w 프로그램이다 (Anthony 2004). AntConc는 사용하기 쉬운 직관적인 그래픽 사용자 인터페이스를 포함하고 있으며, 어휘 분석(concordance)과, 단어 및 키워드 빈도 생성기, 클러스터 및 어휘 묶음 분석 및 단어 분포도를 나타내는 도구를 제공한다. 본 연구에서는 어휘 분석 도구(concordance tool)를 주된 분석 도구로 이용하였다.

관계대명사 *who*와 *which*를 앞서 언급한 AntConc 프로그램을 통해 분류한 뒤, 이들의 사용 빈도를 비교하기 위해서 사용한 것은 Log-likelihood 값이다. 비교하고자 할 코퍼스를 대상으로 고려해야 할 쟁점들 또한 존재하는데, Rayson와 Garside(2000)는 대표성(representativeness), 코퍼스 내의 균질성(homogeneity within the corpora), 코퍼스의 비교가능성(comparability of the corpora), 통계 검정의 신뢰성(reliability of statistical tests)으로 4가지의 고려해야 할 요소들을 나열 하였다. 먼저 대표성은 모든 주요 텍스트 유형의 표본을 포함해야 한다고 한다(Leech 1993). 다음으로 코퍼스 내의 균질성이란 특성은 다른 코퍼스로부터의 구별되는 특징을 발견하는데 목표를 둔다(Rayson and Garside 2000). 코퍼스의 비교가능성은 비교 할 코퍼스의 대상들이 같은 방식으로 표본이 추출되어야 한다는 설명으로 표본 추출의 선택적 방법이라고도 설명한다. 마지막으로 통계 검정의 신뢰성은 코퍼스 규모에 관련하여 통계의 신뢰성이 바탕이 되어야 함을 말하는 것이라 설명한다. 이러

한 요인들을 고려하여, Dunning(1993)은 Log-likelihood ratio가 과거 자주 빈도 비교에서 사용 되던 Pearson의 chi-squared검사를 대체하는 검사라고 제안했다. Log-likelihood(LL)는 다음 표 1에 나오는 분할표의 구성에 의해 계산이 되었다(Rayson and Garside 2000).

표 1. Log-likelihood 계산에 필요한 빈도

	코퍼스 1	코퍼스 2	합계
목표 단어 빈도	a	b	a+b
다른 단어 빈도	c-a	d-b	c+d-a-b
합계	c	d	c+d

Rayson와 Garside(2000)에 따른 해석으로 <표 1>에 언급된 값들을 해석하자면, ‘c’ 값은 코퍼스 1의 단어들의 수, ‘d’값은 코퍼스 2의 단어들의 수와 일치하며 ‘c’, ‘d’ 값은 N의 값이라 한다. ‘a’와 ‘b’의 값은 측정값이라고 부르며 이 값들은 O의 값이라 한다. 이를 바탕으로 기대되는 기대값(E)을 다음 그림 1에 나타난 공식에 따라 계산한다.

$$E_i = \frac{N_i \sum_j O_j}{\sum_j N_j}$$

그림 1. 기대값(E) 구하는 공식

위 표와 공식에 따르면 $N_1=c$ 가 되고 $N_2=d$ 가 되기에, $E_1 = c * (a+b) / (c+d)$ 가 되고, $E_2 = d * (a+b) / (c+d)$ 가 된다(Rayson and Garside 2000). 기대값은 두 개의 코퍼스의 규모를 고려하기 때문에 공식에 적용하기 전, 해당하는 수치들을 표준화할 필요는 없다. 다음 그림 2는 Log-likelihood(LL)의 값을 구하는 공식이다.

$$-2 \ln \lambda = 2 \sum_i O_i \ln \left(\frac{O_i}{E_i} \right)$$

그림 2. Log-likelihood(LL)값 구하는 공식

본 연구에서는 LL의 값을 구하기 위해 해당하는 수치를 직접 대입하여 구하지 않아도 만들어진 수식을 온라인으로 이용 할 수 있는 Log-likelihood calculator를 이용하였다. Log-likelihood calculator를 통해 해당하는 Corpus1, 2에 고등학생과 대학생 코퍼스의 전체 코퍼스 규모(corpus size)를 입력하고 관계대명사 *who*와 *which*에 해당하는 값을 넣어주게 되면 두 집단의 사용 빈도 결과를 비교해서 볼 수 있다. 결과 값의 유의미성 판단은 Log-likelihood 값을 통해할 수 있다. 다음 표 2는 Rayson과 Garside(2000)에 의한 Log-likelihood 값에 따른 *p*값의 범위이다.

표 2. Log-likelihood값에 따른 유의도 검사

Log-likelihood값	<i>p</i> 값
3.84 이상	$p < 0.05$
6.63 이상	$p < 0.01$
10.83 이상	$p < 0.001$

표 2에서 볼 수 있듯이, Log-likelihood값이 3.84이상이 나올 경우 두 집단의 사용 빈도 차이가 유의미하다는 것을 알 수 있다.

3.3 연구 과정 및 방법

본 연구과정의 순서는 다음과 같다. 고등학생의 111개의 영작문을 텍스트(text) 파일 형식으로 저장하였다. 학생들이 작성한 영작문에서 나타난 반점, 대·소문자 사용, 기타 문법적인 오류, 철자상의 오류 등은 있는 그대로 수정하지 않고 입력하였다. 다음으로 사용한 대학생 영어 학습자 코퍼스인 YELC는 연세대학교 영어영문학연구소 영어코퍼스 연구실 홈페이지에서 코퍼스 자료를 요청하여 받을 수 있었다. 또 다른 코퍼스 구축은 고등학교 영어 교과서 본문 파일이다. 본문에 해당하는 내용만 남겨서, 오타와 오류를 직접 수

정한 후 텍스트(text) 파일 형식으로 저장하였다. 위 세 가지의 텍스트를 AntConc3.4.4w 프로그램에 가동시켜 파일을 불러 온 다음, 전체적인 타입(type)과 토큰(token)을 먼저 살펴본 후, 목표 구문인 관계대명사를 검색하였다. 여기서 타입(type)은 서로 다른 단어의 수를 가리키는데 모든 중복된 경우를 하나로 계산 한 수이고, 토큰(token)은 해당 코퍼스에 사용된 모든 단어의 전체 출현 수를 말한다(고광윤과 백주현 2010). 이때 *who* 나 *which*를 검색하게 되면 해당되는 *who* 나 *which* 앞, 뒤로 어떠한 단어의 형태가 와있는지를 살펴 볼 수 있다. *who* 와 *which*를 그 자체로 검색하게 되면, 관계대명사 이외의 의문사로 사용된 경우의 수도 함께 검색이 되기에 관계대명사 *who*, *which*를 찾기 위해 태깅(tagging)을 하여 관계대명사로 사용하지 않은 경우는 제외하였다. 관계대명사 *who*, *which*를 찾은 후, 그 중 계속적 용법으로 사용된 문장도 분류하였다. 세 집단의 코퍼스 유의미성을 살펴보기 위해 online Log-likelihood calculator (<http://ucrel.lancs.ac.uk/llwizard.html>)를 사용하였다. Log-likelihood calculator에서 해당하는 Corpus 1, 2에 비교하고 싶은 집단의 전체 코퍼스 크기(Corpus size)와 보고자 하는 값을 Frequency of word에 넣어 나온 빈도 결과 값을 비교하였다.

4. 결과 및 논의

4.1 학습자 코퍼스에서 나타난 관계대명사 빈도

고등학교 1, 2학년의 코퍼스의 전체 단어 타입(word types)은 3,640개, 단어 토큰(word tokens)은 36,549개였다. 콘코던스 프로그램으로 *who*를 찾았을 때 총 116개의 값이 나왔고, 여기서 관계대명사 *who*가 아닌 것들을 제거하고 나면 89개의 관계대명사 *who*의 빈도 값을 알 수 있었다. 이 중 계속적 용법으로 사용된 *who*는 3번이었다. 그리고 *which*를 입력 하였을 때 전체 *which*의 빈도는 87개이었으며, 그 중 관계대명사에 해당하는 값은 79개로 추출할 수 있었다. 관계대명사 *which* 사용 가운데 계속적 용법으로 쓰인

*which*의 빈도는 27개로 나타났다.

한편, 연세 학습자 코퍼스 YELC에 대한 결과 값을 살펴보면, 전체 단어 타입은 21,219개, 단어 토큰은 1,099,478로 나타났다. 그 중 태깅 작업을 거친 결과, 관계대명사 *who*의 개수는 3,797개였고 계속적 용법인 *who*는 55개였다. *which*도 같은 방법으로 분류 하였을 때, 전체 *which*의 개수 1,834개에서 26개를 제외한 1,808개의 관계대명사 *which*의 개수를 얻을 수 있었고 그 중 계속적 용법의 *which*는 222개였다. 다음 표 3은 고등학생 영작문 코퍼스와 대학생 영작문 코퍼스(YELC)에 나타난 코퍼스의 크기와 한정적 용법과 계속적 용법의 관계대명사의 *who*와 *which* 빈도수를 나타낸 것이다.

표 3. L2 학습자 코퍼스에 나타난 관계대명사 빈도

코퍼스	Tokens	Types	한정적 <i>who</i>	한정적 <i>which</i>	계속적 <i>who</i>	계속적 <i>which</i>
고등학생	36,549	3,640	89 (.244%)	79 (.216%)	3 (.008%)	27 (.074%)
대학생	1,099,478	21,219	3,797 (.345%)	1,808 (.164%)	55 (.005%)	222 (.020%)

표 3에 따르면 고등학생 영작문 코퍼스에 가장 많이 사용된 관계대명사는 *who*이다. *which*는 10번의 빈도 차이로 *who* 다음으로 많이 사용되었다. 대학생 영작문 코퍼스에서도 역시 *who*가 *which*보다 많이 사용되었다. 또한 관계대명사의 문법적 기능에 의해 한정적 용법과 계속적 용법으로 분류하여 볼 때, 두 집단 모두 계속적 용법으로 사용된 횟수 보다 한정적 용법으로 사용된 횟수가 더 많음을 알 수 있다. 그리고 계속적 용법으로 사용된 경우에는 *which*가 *who*보다 많이 사용되었다. 단순빈도만으로 볼 때, 두 학습자 집단은 사용 빈도에 있어서 같은 형태를 띠고 있음을 알 수 있다.

고등학교 영어 교과서 중 7개의 출판사의 총 214개의 본문을 텍스트 파일로 변환하여 AntConc3.4.4w프로그램에 넣었을 때, 전체 단어 타입은 10,938개, 단어 토큰은 169,514개라는 결과 값이 나왔다. 그 중 전체 *who*의 개수는 290개, 전체 *which*의 개수는 260개이었으며, 태깅 작업을 거친 후 관계대명사로 사용된 *who*는 258개이고 그 중 계속적 용법의 *who*는 44개이었으며, 관계대명사

*which*는 223개이고 계속적 용법의 *which*는 122개라는 결과 값을 얻을 수 있었다. 단순 빈도만 보았을 때, 한정적 용법으로 쓰인 관계대명사가 계속적 용법으로 쓰인 관계대명사보다 많음을 알 수 있다. 다음 표 4는 고등학교 영어 교과서 코퍼스에 나타난 코퍼스 크기인 전체 토큰과 타입, 한정적 용법과 계속적 용법의 관계대명사의 *who*와 *which* 빈도수를 나타낸 것이다.

표 4. 고등학교 영어 교과서에 관계대명사 빈도

코퍼스	Tokens	Types	한정적 who	한정적 which	계속적 who	계속적 which
영어 교과서	169,514	10,938	258 (.152%)	223 (.132%)	44 (.026%)	122 (.072%)

표 4에 따르면 고등학교 영어 교과서 코퍼스에 가장 많이 사용된 관계대명사는 *who*이며 계속적 용법으로 사용된 관계대명사도 *which*가 계속적 용법 *who*보다 많이 사용됨을 알 수 있다. 단순 빈도의 결과만 보면 앞의 두 학습자 코퍼스 집단과도 같은 패턴을 보이지만, 계속적 용법으로 쓰인 *who*의 빈도가 학습자 코퍼스에 비해 상대적으로 높음을 알 수 있다.

4.2 두 학습자 코퍼스와 고등학교 영어 교과서의 통계 비교

고등학생 영작문 코퍼스와 연세 학습자 코퍼스(YELC) 두 집단의 관계대명사 빈도가 통계적으로 차이가 있는지 비교하기 위해서 online Log-likelihood calculator (<http://ucrel.lancs.ac.uk/llwizard.html>)를 사용하였다. Log-likelihood calculator 결과는 코퍼스 1에서 목표단어 빈도와 상대 빈도, 코퍼스 2에서 목표 단어 빈도와 상대빈도를 나타내고, 상대빈도 차이를 보여준다. 코퍼스 2의 상대빈도에 나타나있는 + 은 코퍼스 2에 비해 코퍼스 1이 상대적으로 더 많이 사용(overuse)했다는 것을 의미하며, -은 코퍼스 2에 비해 코퍼스 1이 상대적으로 덜 사용(underuse)했다는 의미한다. 이러한 calculator 해석을 바탕으로 정리한 산출된 관계대명사 *who*와 *which*의 한정적 용법과 계속적 용법으로 두 학습자 코퍼스를 비교한 값을 표로 정리하면 다

음 표 5와 같다.

표 5. 두 학습자 코퍼스의 Log-likelihood 결과 값

	한정적	한정적	계속적	계속적
	who	which	who	which
고등학생 코퍼스	89	3	79	27
고등학생 코퍼스의 상대빈도	0.24	0.01	0.22	0.07
YELC 코퍼스	3,797	55	1,808	222
YELC 코퍼스의 상대빈도	0.35-	0.01+	0.16+	0.02+
Log-likelihood (LL) 값	11.89**	0.60	5.21*	29.17***

Note. +고등학생 집단이 YELC집단 보다 더 사용함. * $p < 0.05$,

*** $p < 0.001$

위의 결과를 보면, 관계대명사 *who*는 대학생 영작문에 비해 고등학생 영작문에서 통계적으로 유의미하게 덜 사용 되었다는 것을 알 수 있으며 *which*는 제한적 용법과 계속적 용법 둘 다 대학생 영작 코퍼스보다 고등학교 영작문에서 통계적으로 유의미하게 더 많이 사용한 것을 알 수 있었다.

고등학생 영작문 코퍼스와 고등학교 영어 교과서 코퍼스의 두 집단의 비교는 다음 표 6과 같다.

표 6. 영어 교과서와 고등학생 코퍼스의 Log-likelihood 결과 값

	한정적	한정적	계속적	계속적
	who	which	who	which
고등학생 코퍼스	89	3	79	27
고등학생 코퍼스의 상대빈도	0.24	0.01	0.22	0.07
영어 교과서 코퍼스	258	44	223	122
영어 교과서 코퍼스의 상대빈도	0.15+	0.03-	0.13+	0.07+
Log-likelihood (LL) 값	13.48**	5.24*	13.22**	0.02

Note. +고등학생 집단이 교과서에 비해 더 사용함. * $p < 0.05$, *** $p < 0.001$

위의 결과는 관계대명사 계속적 용법 *who*는 고등학교 영어 교과서 보다 고등학생 영작문에서 유의미하게 덜 사용 되었다는 것을 보여주며, 반면 계속적 용법의 *which*는 더 사용 되었으나 그 값은 통계적으로 유의미하지 않았다. 나머지 제한적 용법의 관계대명사 *who*와 *which* 역시 교과서에 나타난 빈도 보다 고등학생 영작문에서 통계적으로 유의미하게 더 많이 사용이 되었음을 알 수 있다.

마지막으로 연세 영어 학습자 코퍼스(YELC)인 대학생 영작문 코퍼스와 고등학교 영어 교과서 코퍼스 간의 비교이다. 연세 영어 학습자 코퍼스의 대상은 고등학교를 졸업하고 대학에 입학한 신입생 이었고, 대학 신입생 작문 코퍼스와 고등학교 교과서에 나타난 빈도와 비교하였을 때 얼마만큼의 빈도 차이를 보이는지의 결과는 다음 표 7과 같다.

표 7. 영어교과서와 대학생 학습자 코퍼스의 Log-likelihood 결과 값

	한정적	한정적	계속적	계속적
	who	which	who	which
YELC 코퍼스	3797	55	1808	222
YELC 코퍼스의 상대빈도	0.35	0.01	0.16	0.02
영어 교과서 코퍼스	258	44	223	122
영어 교과서 코퍼스의 상대빈도	0.15+	0.03-	0.13+	0.07+
Log-likelihood (LL) 값	206.94**	56.90**	10.48**	107.45**
	*	*		*

Note. +YELC가 교과서에 비해 더 사용함. ** $p < 0.01$, *** $p < 0.001$

표 7에 따르면 대학생 영작문 코퍼스와 고등학교 교과서 코퍼스 사이의 관계대명사 비교는 고등학생 영작문 코퍼스와 교과서 코퍼스 사이의 비교와 비슷한 양상을 보였다. 계속적 용법의 *who*에서만 고등학교 교과서 코퍼스보다 대학생 영작문에서 유의미하게 덜 사용되었고, 한정적 용법의 *who*, 계속적 용법 및 한정적 용법의 *which*는 모두 고등학교 교과서 보다 대학생 영작문에서 유의미하게 더 많이 사용되었음을 알 수 있다.

4.3 관계대명사 *who*와 *which* 사용 오류

본 연구에서는 고등학생 영작문 코퍼스에서 나타난 한정적, 계속적 용법의 관계대명사 *who*와 *which*의 오류 종류를 크게 5가지의 범주로 나누었다. 선행사와의 동사 수일치 오류, 선행사의 부재나 선행사 자체의 잘못 쓰임, 마지막으로 주절의 동사가 누락된 경우가 오류의 5가지의 범주이다. 고등학생 영작문에 나타난 오류의 종류에 따른 빈도를 정리하면 다음 표 8과 같다.

표 8. 고등학생 코퍼스에서의 나타난 오류의 종류와 빈도

오류의 종류	빈도
수일치 오류	12
선행사 부재	1
선행사 오류	1
관계절안 동사 오류	0
불필요한 관계대명사 사용 오류	0

표 8에 나타난 것처럼 고등학생 영작문 코퍼스에 나타난 오류는 총 14개로 나타났다. 그 중 선행사와의 동사 수일치에 해당하는 오류가 12번으로 가장 빈번하게 나타났고, 관계대명사와 선행사와의 거리가 멀지 않음에도 수일치에 있어서 나타났는데, 특히 복수 동사 사용의 오류가 빈번함을 알 수 있었다. 다음으로 본 연구에서 사용한 대학생 영작문 코퍼스인 연세 영어 학습자 코퍼스(YELC)에 나타난 오류의 종류와 각각 발생한 빈도는 다음 표 9와 같다.

표 9. YELC에서 나타난 오류의 종류와 빈도

오류의 종류	빈도
수일치 오류	193
선행사 부재	3
선행사 오류	12
관계절 내 동사 오류	15
불필요한 관계대명사 사용 오류	2

표 9에서 보여주듯 대학생 영작문 코퍼스에서 나타난 오류의 종류는 선행사와의 동사 수일치 오류, 선행사의 부재에도 불구하고 관계대명사를 사용한 경우, 선행사 자체의 오류, 관계대명사 사용 후 주절의 동사의 누락으로 분류할 수 있었으며, 오류의 총 빈도는 225번이었다. 오류의 빈도로 볼 때 선행사와의 동사 수일치 오류가 다른 오류에 비해 큰 차이로 빈번하게 나타났음을 알 수 있다. 이는 앞서 살펴본 고등학생 영작문 코퍼스에서도 가장 빈번한 오류였다. 선행사와의 동사 수일치 오류 문장 중, 특히 관계대명사 *who*를 사용 하는 문장에서 선행사가 *person*, *people*일 경우, 혹은 *someone*, *one*과 같은 대명사를 사용할 경우 오류의 빈도가 높아짐을 알 수 있었다. 선행사들에 따라 동사 수일치의 오류가 발생한 경우는 총 193번으로 그중 관계대명사 *who*를 사용할 때 발생한 오류는 151번이다. 이때 *person*, *people*이 선행사로 쓰인 경우 오류는 67번, *someone*, *one*이 선행사로 쓰인 경우 발생하는 오류의 빈도는 22번 이었다. 이는 모두 관계대명사 *who*를 사용하는 경우에

발생한 오류이다.

4.4 두 학습자 코퍼스의 관계대명사 사용 오류 빈도 비교

지금까지 고등학생 영작문 코퍼스와 대학생 영작문 코퍼스(YELC)에서 나타난 오류의 종류와 그에 따른 빈도를 알아보았다. 두 집단의 전체 코퍼스 규모를 고려한 Log-likelihood의 결과는 표 10과 같다.

표 10. 두 학습자 코퍼스에서의 전체 오류 빈도 비교

	전체 오류 빈도
고등학생 코퍼스	12
고등학생 코퍼스의 상대빈도	0.03
YELC 코퍼스	193
YELC 코퍼스의 상대빈도	0.02+
Log-likelihood (LL) 값	5.59*

Note. +고등학생 집단이 YELC보다 더 오류가 많음. * $p < 0.05$

고등학생 영작문 코퍼스와 대학생 영작문 코퍼스(YELC)에 발생한 오류의 빈도를 비교 하였을 때 나타난 +를 통해 고등학생 영작문 코퍼스에서 YELC보다 더 많은 오류의 빈도를 보였고, 두 값을 비교한 Log-likelihood의 값은 통계적으로 유의미하였다. 특히, 두 집단의 오류 분류 중에서 가장 높은 빈도를 나타내었던 선행사와의 동사 수일치 오류에 대하여, 고등학생 영작문 코퍼스에서 나타난 선행사와의 동사 수일치 오류는 총 12번 발생하였고, 대학생 영작문 코퍼스(YELC)에서는 총 193번 발생하였다. 이들의 빈도 비교한 통계값은 표 11과 같다.

표 11. 두 학습자 코퍼스에서의 수일치 오류 빈도 비교

	수일치 오류 빈도
고등학생 코퍼스	15
고등학생 코퍼스의 상대빈도	0.04
YELC 코퍼스	225
YELC 코퍼스의 상대빈도	0.02+
Log-likelihood (LL) 값	3.70

Note. +고등학생 집단이 YELC에 비해 더 사용(overuse)함

위의 표에서 나타난 Log-likelihood값에 의하면 3.8을 넘지 못하는 3.70이기에 두 빈도 비교에 있어서는 유의미성이 입증되지 않았

다. 따라서 두 집단을 비교하였을 때 선행사와의 동사 수일치 오류 빈도 비교는 통계적으로 차이가 없다는 결과를 얻을 수 있었다.

즉, 두 학습자 코퍼스 집단에서 나타난 오류의 종류는 어느 정도 일관성을 나타내었다. 선행사와의 동사 수일치 오류가 두 집단 모두에게서 가장 빈번하게 나타났다. 오류의 형태를 자세히 살펴보면, 선행사와 관계절의 동사 거리가 2절점 이상을 넘지 않는 가까운 거리임에도 불구하고 오류의 수는 높게 나타났다. 또한 관계대명사 *who*를 사용하는 경우 발생하는 오류가 *which*를 사용하는 경우에서는 한 번도 발생하지 않는 특징을 볼 수 있었다. 이는 관계대명사의 선행사의 사람 및 사물에 제한된 특성과도 연관될 수 있을 것이다(Izumi, 2003). 또한, 계속적 용법의 관계대명사를 사용하는 문장에 있어서 오류는 발생되지 않았다는 점도 특이점이라 볼 수 있다. 이는 사용 빈도가 애초에 낮다는 점과 사용을 꺼리는 L2학습자의 습득 과정에서 보이는 회피(avoidance)현상으로 설명 될 수 있다(Kleinmann 1977).

관계대명사 습득 위계 가설과 연관하여, 선행연구에서 고등학생 집단에 있어서는 인지 난이도 위계 가설(PDH)과 명사구 접근성 위계(NPAH) 가설을 뒷받침하는 결과들이 많았다. 습득 가설 연구는 보통 학습자들에게 다양한 과제를 제시하고 결과를 얻는 경우가 많지만, 본 연구 대상인 학습자 영작문 코퍼스와 습득 위계를 연관지어 비교해보면, 전반적으로 고등학생들과 대학 신입생의 학습자 코퍼스에서 다양한 종류의 관계대명사가 발견되지 않았다. 또한, 두 코퍼스에서 보인 오류의 빈도는 SS, OS에서 Subject 주어위치가 대부분이었다. 따라서 두 학습자 코퍼스에서 학습자들은 상대적으로 사용하기 쉬운 관계대명사의 용법을 사용하고 있고, 그 단계에서의 오류 발생도 많이 발생하는 것으로 보아, 관계사의 기능에 따라 나눈 명사구 접근성 위계 가설(NPAH)을 따르는 것으로 파악할 수 있다.

5. 결론

본 연구는 L2 학습자가 학습하는 여러 문법적인 내용들 중 관계

대명사가 가지고 있는 자체적 특성으로 인해 다른 문법 내용들 보다 습득하고 사용하는데 어려움이 있다는 것을 고려하여, 고등학생과 대학 신입생의 작문 학습자 코퍼스에서 관계대명사 *who*와 *which*의 빈도와 오류를 알아보았다. 그리고 그 결과를 고등학교 교과서 코퍼스와 비교하였다.

첫 번째 연구 질문은 두 L2 학습자 집단인 고등학교 1,2 학년과 대학 신입생 영작문에서 한정적 용법과 계속적 용법 관계대명사 *who*와 *which*의 사용 빈도와 오류의 유형이 차이가 나는지에 대한 질문이었다. 빈도 비교 결과는 고등학생 코퍼스에서 대학생 코퍼스보다 관계대명사 *which*의 제한적 용법 및 계속적 용법 둘 다 더 많이 사용되었으며, 관계대명사 *who*는 덜 사용되었다. 오류 빈도 비교 결과, 고등학생 코퍼스에서 대학생 코퍼스보다 통계적으로 유의미하게 더 많은 오류가 나타났다. 하지만, 고등학생 코퍼스에서 나타나지 않았던 관계절의 동사 오류와 불필요한 관계대명사 사용 오류는 대학생 코퍼스에서 발견되었다. 종합해보면, 고등학생 학습자들이 *which*를 더 많이 사용(*overuse*)하는 경향과 더 많은 오류의 빈도를 보였다.

두 번째 연구 질문은 고등학교 영어 교과서 본문을 코퍼스로 하여 한정적 용법과 계속적 용법의 관계대명사의 *who*와 *which*의 빈도가 앞서 살펴본 두 학습자 영작문 집단에서의 사용 빈도와 비교하여 어떠한 특징을 가지고 있는가에 대한 질문이었다. 두 학습자 집단 모두 고등학교 영어 교과서보다 한정적 용법의 *who*와 *which*를 더 많이 사용하였다. 이는 두 집단의 계속적 용법에 대한 회피 현상으로 파악된다.

종합적으로, 관계대명사 *who*와 *which*는 두 학습자 코퍼스 집단에서 사용 빈도와 오류 패턴에 있어 크게 두드러지는 차이는 보이지 않았다. 두 코퍼스에서 다양한 종류의 관계대명사가 발견되지 않았고, 두 코퍼스에서 보인 오류의 빈도는 SS, OS에서 Subject 주어위치가 대부분인 점을 보면, 습득 위계에 있어서 명사구 접근성 위계 가설(NPAH)을 뒷받침하는 결과라고 할 수 있다.

하지만, 고등학교 영어 교과서와 비교하였을 때 두 학습자 코퍼스 집단 모두 한정적 용법이 계속적 용법 보다 더 많이 사용했음을 알 수 있다. 교과서에서 계속적 용법에 대한 예문이 상당수 존재함

에도 불구하고, 이는 학습자의 회피 현상을 반영하는 결과라고 할 수 있다. 하지만, 관계대명사 계속적 용법의 회피 현상에 대한 구체적인 원인 및 학습·교육적 효과 등에 대한 연구가 적어, 이 부분에 대한 심도 깊은 연구가 필요하다.

참고문헌

- 고광윤·백주현(Goh, G.-Y. and J.-H. Baek). 2010. 한국 대입영어시험 어휘의 코퍼스 분석(A corpus-based analysis of the vocabulary used in the CSAT English exam and two analogous tests). 《영어학연구》(*English Language and Linguistics*) 16, 1-26.
- 고미숙(Ko, M.-S.). 2007. 한국 고등학교 학생들의 관계절 습득에 관한 연구(The study on Korean EFL high school students' acquisition of relative clauses). 《교과교육학연구》(*Journal of Research in Curriculum Instruction*) 11, 57-75.
- 김남국(Kim, N.-G.). 2014. 한국 대학생들의 영어 관계대명사 습득 양상(A study on the acquisition of the English relative pronoun by Korean college students). 《영어영문학》(*The Journal of Mirae English Language & Literature*) 19-4, 181-207.
- 김태희(Kim, T.). 2009. 한국 고등학교 학생들의 영어관계절 습득(*The acquisition of English relative clauses by Korean high school students*). 이화여자대학교 대학원 영어영문학과 석사학위논문(Unpublished master's thesis, Ewha Womans University, Seoul, Korea).
- 김효은(Kim, H.-E.). 2007. 코퍼스에 언어학적 접근법에 의한 고등학교 영어교과서의 관계대명사 분석(*A corpus-based analysis of uses of relative pronouns in high school English textbooks*). 동아대학교 교육대학원 석사학위논문(Unpublished master's thesis, Dong-a University, Busan, Korea).
- 이권창(Lee, K.-C.). 2014. 한국 고등학교 학습자의 영어 관계절 구문 습득 양상 연구: 유형별 난이도 위계와 오류 분석(*A study on the acquisition of English relative clause constructions by Korean high school learners: The difficulty order of the relative clause types and error analysis*). 한국교원대학교 대학원 석사학위논문(Unpublished master's thesis, Korea National University of Education, Cheongju, Korea).
- 이석재·정채관(Lee, S.-C. and C. K. Chung). 2014. 한국 예비 대학생의 영어 사용 특성 파악을 위한 대규모 공개 영어 학습자 코퍼스 구축 및 분석(Compilation of the Yonsei English Learner Corpus (YELC) 2011 and Its Use for Understanding Current Usage of English by Korean Pre-university Students). 《한국콘텐츠학회논문지》(*The Journal of the Korea Contents Association*) 14-11,

- 1019-1029.
- 이은주(Lee, E.-J.). 2009. 한국인 영어 학습자 코퍼스를 활용한 정형화된 연결어구 분석(A corpus-based study of the Korean EFL learners' use of formulaic sequences). 《외국어교육》(*Foreign Languages Education*) 16-2, 321-340.
- 이은주(Lee, E.-J.). 2010. 한국 대학생 영어 학습자 코퍼스에 나타난 견지 부사구 (Stance Adverbials)의 사용(The use of stance adverbials in the Korean EFL college students' writing corpus). 《외국어교육》(*Foreign Languages Education*) 17-3, 347-366.
- 이정은·김동규(Lee, J.-E. and D. Kim). 2013. 영어 어휘 연구 최근 동향 분석: 2000 년 이후 학술지 논문을 바탕으로(A synthesis of current research on English vocabulary in Korea: A content analysis of journal articles from 2000 to 2011). 《외국어교육》(*Foreign Languages Education*) 20-3, 113-140.
- 장진입(Jang, J. I.). 2011. 고등학교 영어교과서에 나타나는 관계절에 대한 연구(A corpus-based analysis of the uses of relative pronouns in high school English textbooks). 부경대학교 교육대학원 석사학위논문(Unpublished master's thesis, Pukyong National University of Education, Busan, Korea)..
- 한나래·이수화(Han, N.-R. and S.-H. Lee). 2009. 학습자 코퍼스를 이용한 영어 전치사 오류 교정 모델 개발(Developing a model for English preposition errors using a learner corpus). 《언어학》(*Journal of the Linguistic Society of Korea*) 53, 163-185.
- Anthony, L. 2004. AntConc: A learner and classroom friendly, multi-platform corpus analysis toolkit. *Proceedings of IWLeL*, 7-13.
- Celce-Murcia, M., D. Larsen-Freeman and H. A. Williams. 1983. *The Grammar Book: An ESL/EFL Teacher's Course*. Boston, MA: Heinle & Heinle.
- Doughty, C. 1991. Second language instruction does make a difference. *Studies in Second Language Acquisition* 13(4), 431-469.
- Dunning, T. 1993. Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics* 19(1), 61-74.
- Hamilton, R. L. 1994. Is implicational generalization unidirectional and maximal? Evidence from relativization instruction in a second language. *Language Learning* 44(1), 123-157.
- Izumi, S. 2003. Processing difficulty in comprehension and production of relative clauses by learners of English as a second language. *Language Learning* 53(2), 285-323.
- Keenan, E. L. and B. Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8(1), 63-99.
- Kleinmann, H. H. 1977. Avoidance behavior in adult second language acquisition. *Language Learning* 27(1), 93-107.
- Kuno, S. 1974. The position of relative clauses and conjunctions.

- Linguistic Inquiry* 5(1), 117-136.
- Leech, G. 1993. 100 million words of English. *English Today* 9(1), 9-15.
- O'Grady, W. (1987). *Principles of Grammar and Learning*. Chicago: University of Chicago Press.
- O'Grady, W. (1999). Toward a new nativism. *Studies in Second Language Acquisition* 21, 621-633.
- Rayson, P. and R. Garside. 2000. Comparing corpora using frequency profiling. *Proceedings of the Workshop on Comparing Corpora*, 1-6.
- Sheldon, A. 1974. The role of parallel function in the acquisition of relative clauses in English. *Journal of Verbal Learning and Verbal Behavior* 13(3), 272-281.
- Shin, J.-A. 2011. Overpassivization errors in Korean college students' English writings. *Korean Journal of Applied Linguistics* 27(3), 255-273.

예시 언어(Examples in): 영어(English)

적용가능 언어(Applicable Languages): 영어(English)

적용가능 수준(Applicable Level): 중고등학교(Secondary)

이경은

동국대학교 교육대학원 영어교육전공
04620 서울특별시 중구 필동로 1길 30
Tel: 010) 9577-3284
E-mail: 10nara26@hanmail.net

신정아

동국대학교 영어영문학부
04620 서울특별시 중구 필동로 1길 30
Tel: 02) 2260-3167
E-mail: jashin@dongguk.edu

논문접수: 2017년 10월

논문수정: 2017년 11월

게재결정: 2017년 12월