

코메트릭스를 활용한 한국 대학생들의 영어 논설문과 설명문 요약문 비교 분석*

전문기(건국대학교)
최윤희(충신대학교)**

Jeon, Moongee and Choe, Yoonhee. 2019. A Coh–Metrix analysis of Korean EFL learners' summary writings in the English argumentative and expository texts. *Korean Journal of English Language and Linguistics* 19–3, 539–559. This study aims to investigate to what extent Korean EFL learners' summary writings of the English argumentative and expository texts are different in terms of lexical, semantic, syntactic, discourse features and readability levels. The participants were 26 Korean college students majoring in English language education in a university located in Seoul. Their summary writings of English argumentative and expository texts were collected and analyzed using a computational assessment tool, Coh–Metrix. Besides, the results of Coh–Metrix analysis were compared using ANOVA. The results indicate that the participants' summary writings in the two different text types were statistically different in terms of their lexical, semantic, syntactic, discourse features and the readability levels. In particular, the participants used more words, longer sentences, more low–frequency and concrete words and adjectives in their argumentative summary. In contrast, the participants used more nouns and verbs in their expository summary. For syntactic and discourse features, more additional conjunctions, syntactic complexity and more co–referential cohesion were shown in their argumentative summary. This study provides some pedagogical implications for teaching Korean EFL learners' summary writings of different English text types and raising text–specific linguistic features.

Keywords: Coh–Metrix, English summary writing, argumentative text, expository text, linguistic feature, computational linguistics

1. 서론

최근에 영어의 다양한 담화 유형 및 장르의 특징과 차이점들에 대한 인식의 중요성이 대두되면서 교육에도 이러한 차이점에 대한 인식을 적용하기 시작하였다(곽혜영 2018, Bailey

* 이 논문은 2018년도 건국대학교 KU학술연구비 지원에 의한 논문임.

건국대학교 KU학술연구비는 교신저자와 무관함.

** 제1저자: 전문기, 교신저자: 최윤희

2015, Park and Lee 2018). 특히, 학문적 목적의 영어교육에서 다양한 담화의 유형과 글의 장르에 대한 이해가 중요함에 따라, 다양한 담화 유형 및 장르에 따라 서로 다른 유형의 글들의 수사학적, 통사적, 어휘적 구조가 유형 간에 어떠한 차이가 있는 지에 대한 비교 연구가 관심을 받고 이루어지고 있다(곽혜영 2018, 안수진 2018, McGee 2014, Park and Lee 2018, Yasuda 2011). 이러한 연구들은 학문적 목적의 영어교육에서 빈번하게 다뤄지는 설명문(expository writing) 또는 논설문(argumentative writing)과 일상적인 글에서 많이 접할 수 있는 기술문(description)과 서사문(narration)에 대한 비교 분석을 주로 수행하였다(배식한 2019, 안수진 2018). 이와 더불어 최근에는 주어진 영어 텍스트를 읽고 요약하는 요약문(summary)을 작성하는 과정에 대한 관심도 높아졌다. 이에 요약문 작성이 학문적 목적의 영어 평가 시험인 TOEFL 시험에서 영어 쓰기 영역의 한 과업으로 개발되었고, 한국 교육과정평가원 주관으로 개발되는 중등 영어과 교원임용경쟁시험에서도 한 문항으로 개발되어 타당하고 신뢰도 있는 교사 선발 문항으로 사용되고 있다.

다양한 영어 글의 담화 유형 및 장르에 관련한 선행연구들을 살펴보면, 각 개별 담화 유형이나 장르에 따라 영어 원어인 저자의 글과 영어를 제2언어 또는 외국어로 배우는 학습자들이 작성한 글이 수사학, 통사적, 구성적, 어휘적으로 어떻게 다른지 비교 분석한 연구들이 주를 이루고 있다(안수진 2018, Bae and Min 2017, Kim 2016, McGee 2014, Min, Paek and Kang 2019). 또한 다양한 장르의 주어진 글에 대한 요약문 연구들은 대체로 질적 양적으로 분석한 연구들(McGee 2014, Park and Lee 2018, Yasuda 2011)이 대부분이었다. 그러나 다양한 담화 특징을 가지고 있는 장르의 글을 읽은 후 요약한 글을 수사학적인 특징 뿐 아니라 자세한 언어적 특징을 종합적으로 양적 비교 분석한 연구는 드문 편인 것으로 보인다. 특히, 최근 컴퓨터 언어학과 코퍼스 언어학의 급진적인 발전으로 개발되어 다양한 텍스트 분석에 사용되고 있는 Coh-Metrix와 같은 객관적이고 신뢰할 만한 분석 도구를 사용하여 다양한 장르의 글을 요약한 글을 비교 분석한 연구는 매우 드물다.

Coh-Metrix 시스템의 신뢰성은 국내에서 다양한 연구들을 통해 확인되었다. 예를 들면, Coh-Metrix를 활용하여 중·고등학교에서 사용하는 교과서의 어휘의 난이도 및 이독성을 비교 분석한 연구(전문기 2011, 전문기, 임인재 2010) 및 수능 영어 영역의 어휘와 수능연계 EBS 교재의 어휘의 난이도를 Coh-Metrix를 이용하여 비교 분석한 연구(김미혜, 이성원 2018, 김제우, 이동주 2017, 성일호 2014, 황영민, 이경민 2014) 등이 있다.

Coh-Metrix를 이용한 수능의 읽기 지문을 분석한 최근 연구로 고나은과 신정아(2017)는 EBS-수능 연계 정책의 전과 후에 대학수학능력시험의 읽기 영역 지문의 난이도를 Coh-Metrix를 이용하여 비교 분석하였다. 난이도 분석을 위해 어휘, 통사, 정합성 측면과 이독성 측면에서 EBS-수능 연계 전후의 영어 지문의 난이도에 유의미한 차이가 있는지를 살펴보았다. 어휘적 측면에서는 단어 수, 문장 길이, 텍스트 내에서 얼마나 다양한 어휘들이 사용되는지를 측정하는 내용어와 모든 단어의 타입-토큰 비율(type-token ratio), 내용어의 CELEX 어휘빈도 평균, 모든 단어의 CELEX 로그 어휘빈도 평균을 객관적인 수치

로 비교하였다. CELEX 데이터베이스는 네덜란드 어휘정보센터(Dutch Centre for Lexical Information)에서 만들어진 것으로 어휘빈도 정보를 담고 있다. 이 데이터베이스는 텍스트 내에 포함된 어휘들의 빈도를 산출하는데 활용된다. 통사적인 측면에서는 명사구 밀도(noun density) 점수와 본동사 앞 단어 평균 어휘수로 비교하였다. 정합성 측면에서는 인접한 문장 사이의 명사 중복(noun overlap) 문장 평균 수, 인접한 문장 사이의 논항 중복(argument overlap) 문장 평균 수, 인접한 문장 사이의 어간 중복(stem overlap) 평균 수와 인접 문장 간의 의미적 정합성(semantic cohesion)을 비교 분석하였다. 이독성 측면에서는 표준 읽기가독성 측정치(Flesch Reading Ease 점수, Flesch-Kincaid Grade Level 점수)로 비교 분석하였다. Flesch Reading Ease 점수와 Flesch-Kincaid Grade Level 점수는 문장길이와 단어길이에 의해 계산되는데 텍스트의 이해도 및 읽기난이도를 측정하는데 이용된다. 분석 결과, EBS-수능 연계 정책의 전과 후에 대학수학능력시험의 읽기 지문의 난이도에 차이가 있음이 확인되었다.

또한, 김재은과 최인철(2015)은 고등학교 영어교과서, EBS 수능 연계 교재와 대학수학능력시험의 난이도를 Coh-Metrix를 이용하여 비교분석하였다. 그들은 세 종류의 자료에서 듣기 스크립트와 독해의 선택지는 제외하고 읽기 지문만을 대상으로 코퍼스를 구축해서 TTR(type/token ratio)과 STTR(standardized TTR)을 비교분석하였다. 또한 이독성 지수 평균을 비교하기 위해 세 자료 간 일원분산분석(ANOVA)을 실시하였다. 연구 결과, EBS 수능 연계 교재와 대학수학능력시험 간의 난이도에 차이가 있음이 발견되었다. 이상에서 살펴본 바와 같이 Coh-Metrix는 국내에서도 다양한 텍스트의 특징을 객관적으로 분석하는데 신뢰롭게 적용될 수 있는 양적인 방법으로 많이 사용되고 있다.

본 연구에서는 Coh-Metrix 시스템을 활용하여 한국인 대학생들이 영어 설명문과 논설문을 읽고 요약한 요약문의 언어적 특성을 비교분석하였다. 구체적으로 과제(설명문 요약과 논설문 요약) 즉, 담화 유형의 정보가 대학생들이 작성한 요약문에 미치는 효과를 분석하고자 하였다. 이러한 담화 유형의 차이가 요약문에 미치는 효과를 분석하기 위해 본 연구에서는 Coh-Metrix 시스템의 기본텍스트 측정치(평균 어휘 수, 평균 문장 수, 평균 단락 수, 평균 문장 길이, 평균 단어 길이), 어휘빈도수 측정치(내용어에 대한 어휘빈도수), 어휘속성 측정치(심상성, 구체성, 습득나이), 어휘다양성 측정치(타입-토큰 비율), 주요 품사 측정치(명사, 동사, 형용사, 부사), 인칭대명사 측정치(1인칭, 2인칭, 3인칭 대명사), 접속사 측정치(인과, 부가, 시간 접속사), 표준 읽기가독성 측정치(Fleach Reading Ease 점수, Flesch-Kincaid Grade Level 점수), 통사구조 복잡성 측정치(명사구 밀도, 본동사 앞 어휘 수), 참조정합성 측정치(인접된 문장 간 논항중복 비율, 모든 문장 간 논항중복 비율), 의미적 유사성 측정치를 분석하였다. 본 연구의 주된 목적을 명세화하여 제시하면 다음과 같다.

첫째, 본 연구는 서로 다른 담화 유형의 과제(설명문 요약 과제 vs. 논설문 요약 과제)가 EFL 대학생들이 작성한 요약문에 어떠한 영향을 미치는 지를 Coh-Metrix 시스

템의 다양한 측정치에 기반해 통계적으로 분석하고자 하였다. 둘째, 이러한 결과를 바탕으로 학생들이 수업 현장에서 글쓰기 과업을 수행할 때 효과적으로 적용할 수 있는 교육 방법적 시사점 및 함의점을 제시하고자 하였다.

2. 이론적 배경 및 선행 연구

2.1 영어의 다양한 담화 유형에 따른 비교 연구

영어 작문 관련 연구에서는 영어 원어민 저자들이 작성한 글과 영어를 제2언어 또는 외국어로 배우는 학습자들이 작성한 글을 언어의 여러 특질에 따라 양적 또는 질적으로 비교 (linguistic comparison)한 연구들이 주를 이루어 왔다(Yasuda 2011). 최근에는 다양한 글의 담화 유형 및 과업의 유형에 따라 언어적 특징이 어떻게 다른 지를 비교한 연구가 관심을 받고 있다. 이러한 연구들 중 논설문이 가장 많은 관심을 받으며 연구가 이루어진 것으로 보인다(McGee 2014, Min, Paek and Kang 2019). 또한 논설문은 글의 다른 유형인 이야기 글 즉 서사문이나 설명문과도 비교하여 연구되고 있다(Bae and Min 2017, Hwang 2012). Hwang(2012)은 논설문과 이야기 글을 통사 구조의 복잡성, 정확성과 유창성 (complexity, accuracy, and fluency)을 바탕으로 분석하였다. 또한, 최윤희(2013)는 국가 영어능력평가시험 2급의 쓰기 과업인 개인 의견쓰기(personal opinion writing)와 논설문에서 고등학생들의 쓰기 수행이 어떠한 차이가 있는 지를 내용(content), 구성(organization)과 언어사용(language use) 측면에서 비교 분석하였다. 위 두 연구들은 논설문과 일상에서 많이 접할 수 있는 글인 이야기 형식의 서사문 또는 개인 의견 쓰기 글을 통사 구조와 내용적 측면으로 비교해 보았다는 점에서 의의가 있다.

이와는 달리 학문적 목적의 글로 많이 사용되는 설명문과 논설문을 비교한 연구도 이루어져 왔다. 학문적 목적의 영어쓰기 교육에서 논설문 뿐 아니라 원인과 결과(cause and effect) 및 비교와 대조(comparison and contrast)등의 유형은 설명문(expository)의 대표적인 중요한 글의 유형 중 일부이다(배식한 2019). 안수진(2018)은 한국인 대학생들이 작성한 영어 논설문과 설명문을 영어 원어민들이 작성한 글의 코퍼스와 비교하여 Coh-Metrix로 비교분석하였다. 안수진(2018)의 연구는 두 가지 서로 다른 담화 유형의 글을 원어민의 코퍼스와 비교하여 기존의 연구 방법과는 달리 Coh-Metrix로 분석하였고 통계적으로 유의미한 차이까지 살펴보았다는 점에서 연구 방법적으로 많은 의의를 제시하였다고 보인다. 이 연구 결과에 따르면, 설명문과 논설문 모두에서 원어민이 한국인 영어 학습자들보다 통계적으로 유의미하게 문장 수와 평균 문장의 길이가 유의미하게 길었다고 한다. 그러나 안수진(2018)의 연구는 한국인 대학생과 영어 원어민 대학생이 작성한 글의 언어적 특징을 설명문과 논설문에서 각각 개별적으로 비교하였고, 설명문과 논설문에 나타난 측정

치의 값을 서로 비교하지는 않았기 때문에 글의 담화 유형에 따른 차이를 비교하지는 못하였다는 한계가 있어 보인다.

Bae와 Min(2017)은 한국인 대학생 학습자들의 영어 글과 영어 원어민 학생들이 쓴 글과 이미 출판된 학문적 글의 코퍼스를 기반으로 글의 다양한 담화 유형(이야기 글, 원인/결과, 비교/대조 및 논설문)에 따라 접속사의 사용에서 어떠한 차이가 있는지를 비교 분석하였다. 그 결과 글의 담화 유형에 따라 다양한 접속사를 사용하는 것으로 나타났다. 원인과 결과 글에서는 원인과 결과를 나타내는 접속인 *because, due to, therefore* 등이 많이 사용되었고, 비교와 대조하는 글에서는 *in comparison, by contrast* 등의 유사한 점과 다른 점을 나타내는 접속사들이 많이 사용된 것으로 나타났다. 이들의 연구는 네 개의 서로 다른 글의 담화 유형에 따른 연결사들의 사용 빈도를 양적으로 비교하였다는 점에서 의의가 있어 보인다. 그러나 네 개의 담화 유형의 차이에 따라 접속사의 사용을 통계적으로 유의미한 차이가 있는지를 좀 더 객관적으로 비교 분석하지 않았다는 점에서 한계점이 있다.

이상의 연구들에서 살펴보았듯이, 영어의 다양한 담화 유형에 따른 어휘, 의미, 담화 및 통사적 특징이 한국인 영어 학습자가 작성한 다양한 영어 글의 유형과 언어적으로 어떠한 차이가 있는지를 Coh-Metrix 등의 객관적이고 신뢰할 만한 도구를 사용하여 종합적으로 비교 분석한 연구는 부족한 것으로 보인다.

2.2 영어로 요약하기 관련 국내 연구

주어진 글을 읽고 요약문 쓰기는 주어진 글의 주제를 파악하여 적절하게 환원(paraphrasing)하는 능력을 갖추고 있어야 가능하다. 이러한 영어로 주어진 글을 영어로 환원하는 역량은 학문적 목적의 글쓰기에서 많이 요구되는 글쓰기 역량이다. 그러나 한국인 영어 학습자들은 영어 요약문을 작성할 때 많은 어려움을 겪고 있다고 한다(김정덕 2019, Choi 2012). 최근에 이루어진 요약문 쓰기에 관한 국내 연구들을 살펴보면 다음과 같다. Kim(2016)은 한국 고등학생들의 영어 요약하기 수행 양상을 분석하였고, 그 결과 한국 고등학교 학생들은 영어 지문의 주제와 사소한 내용을 구분할 수는 있으나, 아이디어들을 연결하는 연결사 사용에 어려움이 있는 것으로 나타났다고 보고하였다. 또한, 영어로 요약할 때에 환원하기 또는 변형하기(transformation) 등의 고급 전략은 매우 드물게 사용하는 것으로 나타나 한국 고등학생들의 요약하는 역량은 매우 초보적인 수준이라고 보고하였다. 이러한 결과를 바탕으로 Kim(2016)은 고등학생들에게 요약할 기회를 많이 제공하며 체계적인 교육이 이루어져야 한다고 주장하였다. 또 다른 국내 연구로 김정덕(2019)은 한국 대학생 9명을 대상으로 영어로 요약하는 글쓰기의 어려움을 질적 연구 방법인 근거 이론적 grounded theory 접근방법으로 분석하였다. 그 결과 연구 참가자들은 중·고등학교 시절부터 암기 위주의 내신 및 대학수학능력시험 대비 준비로 분주하여 제대로 된 영어 요약하기 과업을 해 본적이 없어서 어려워하는 것으로 나타났다고 한다. 또한 영어로 환원하거나 부

분 글을 요약하는 것 뿐 아니라 전체 글을 영어로 요약해 본 경험이 거의 없어 주어진 영어 지문을 요약하는 데 어려움이 있다고 토로한 것으로 나타났다 보고하였다.

Hwang(2010)은 한국인 대학생 30명을 대상으로 이야기 글인 서사문과 설명문에 대한 요약문을 비교분석하였다. 이와 더불어 연구 참가자들이 요약한 글과 본래 주어진 읽기 문항의 독해력 테스트 점수와의 상관관계를 분석하였다. 그 결과 연구 참가자들은 이야기 글의 구조가 상대적으로 더 단순하고 이해하기 쉬워서 요약을 더 쉽게 잘 했다고 인식하고 있었다고 한다. 반면, 설명문은 글의 논리적인 흐름을 이해해야 하므로, 글의 응집장치들을 잘 이해해야 가능하다고 보고하였다고 한다. Hwang(2010)의 연구는 위에서 살펴본 Kim(2016)과 김정덕(2019)의 연구에서 살펴보지 않았던 글의 유형 즉 설명문과 이야기 글에 따라 요약문의 수행 결과가 어떻게 차이가 나는지를 연구하였다는 점에서 주목할 만한 연구로 보인다.

좀 더 최근의 연구인 곽혜영(2018)은 한국 대학생들의 영어로 요약한 글에서 환원하는 양상이 의미적, 어휘적, 구문적 측면에서 질적으로 어떠한 지를 살펴보았다. 그 결과 연구 참가자들은 주어진 글을 영어로 환원할 때 어휘와 구문을 통합하는 전략을 가장 많이 사용하였고, 어휘를 유사한 의미의 다른 어휘로 변경하는 전략도 많이 사용하였다고 한다. 또한 의미적인 측면에서는 새로운 정보가 추가되거나 원문의 내용이 일부 삭제되어서 불완전하게 의미가 전달되는 경우가 많았다고 한다. 이와 더불어 한국인 학생들은 요약하거나 환원할 때에 문맥상 적절하지 않은 동의어로 대체하여 영어 원문의 의미가 변질되어 유사하지 않은 의미가 되기도 하였다고 한다. 곽혜영(2018)의 연구는 우리나라 영어교육에서 적절하고 효율적으로 환원하는 방법에 대한 교육의 필요성을 시사했다는 점에서 큰 의의가 있다고 보인다.

Park과 Lee(2018)는 영어 작문 과업의 복잡성이 한국인 영어 학습자의 요약하기 수행에 미치는 영향에 대해 분석하였다. 이들은 Robinson(2001)의 과업의 인지적 복잡성 모델을 활용하여 네 개의 서로 다른 복잡성을 가진 영어 지문을 요약하는 과업을 비교 분석하였다. 그들이 비교한 글의 담화는 논설문과 설명문이었으며, 복잡성의 차이는 요약문의 길이에 대한 제한 여부 즉, 70단어 또는 단어 길이에 제한이 없는 요약문 작성이었다. 위의 연구들은 글의 다양한 담화 유형에 따라 그 글을 읽는 독자들의 인지 과정이 다르므로 요약하는 과정도 다를 것이고 이는 학습자들이 작성한 요약문에서도 언어적, 의미적, 구조적, 통사적, 어휘적 측면에서 특징이 다르게 나타날 수 있다는 점을 시사하고 있다.

그 동안의 선행연구들을 검토한 결과 국내외 해외 연구들의 대부분은 영어 원어인 저자들이 작성한 작문 코퍼스와 L2 화자의 영어 작문 코퍼스에서 글의 수사적 또는 통사적 특징들을 비교 분석한 연구들이 대부분이었다. 영어 글의 다양한 담화 유형에 따른 수사적, 통사적, 어휘적 차이를 살펴보거나, 또는 다양한 담화 유형의 글에 대한 영어 요약문의 차이를 분석하는 것은 학문적 목적을 위해서도 매우 중요하다. 그럼에도 불구하고 한국인 EFL 영어 학습자들이 다양한 영어 담화 유형의 글을 요약한 요약문에 어떠한 언어적, 어휘적,

통사적, 담화적 차이가 있는 지를 종합적으로 비교 분석한 연구는 매우 드물었다. 따라서 본 연구에서는 한국인 대학생을 대상으로 논설문과 설명문을 제시하여 읽게 한 후, 요약문을 작성하게 하고 이를 Coh-Metrix 시스템을 이용하여 비교 분석하였다.

3. 연구방법

3.1 연구대상(연구 참가자 및 연구자료) 코퍼스

본 연구는 서울에 위치한 한 4년제 대학의 영어 교육을 전공하는 학생 26명을 대상으로 수행되었다. 이들은 각각 논설문과 설명문(원인과 결과에 대한 글)을 읽고 요약문을 작성하는 과업을 수행하였다. 연구 대상과 수집 자료의 내용은 표 1과 같다. 연구 참가자들은 ‘Will A.I. replace human teachers in the future?’란 주제의 논설문과 ‘Causes and effects of melatonin’에 관한 글을 읽고 한 개의 주제문(topic sentence), 두 세 개의 뒷받침하는 근거들 그리고 한 개의 결론문으로 요약하는 과업을 수행하였다. 각 주제 당 30분의 시간이 허용되었다.

표 1. 연구 참가자 및 수집 자료

성별		학년		요약문	
남	10명	2학년	2명	설명문(원인결과)	26 개
여	16명	3학년	18명	논설문	26 개
		4학년	6명		
합계	26명	합계	26명		

연구 참가자들은 주어진 글을 자세히 읽고 주제문을 작성해야 하고 그 주제문을 뒷받침하는 근거들을 찾고 두세 개의 근거 문장들과 한 개의 결론 문장을 작성하였다. 그러나 주어진 글에서 연속하는 단어 4개 이상을 그대로 베껴서 쓰는 것은 허용되지 않았고 적절하게 패러프레이징 즉 환원하는 능력이 요구되었다.

3.2 코퍼스 구축과 분석 도구

본 연구에서는 English as a Foreign Language(EFL) 학생들이 수업 시간을 통해 작성한 두 유형의 영어에세이 코퍼스 자료(논설문과 원인결과의 설명문)를 Coh-Metrix 컴퓨터 시스템을 통해 분석하였다. Coh-Metrix 시스템은 미국 멤피스대학교(The University of Memphis)에 소재한 지능형시스템연구소(Institute for Intelligent Systems)에 의해서 개발

된 텍스트 분석 소프트웨어로서 다양한 유형의 영어 코퍼스 자료를 분석하는데 광범위하게 활용되고 있다(Graesser, Jeon, Cai and McNamara 2008, Graesser, Jeon, Yan and Cai 2007, Graesser, McNamara, Louwerse and Cai 2004). 좀더 구체적으로 설명하면, 본 연구에서는 개인 PC용 Coh-Metrix 시스템이 사용되었다. 개인 PC용 Coh-Metrix 시스템은 txt 파일 확장자 형태의 자료만 처리할 수 있기 때문에 본 연구에서 분석된 영어에세이 코퍼스 자료는 모두 txt 파일 형태로 구축되었다. 영어에세이 코퍼스는 26명의 EFL 학생들에 의해 작성된 총 52개(논설문 26개, 설명문(원인/결과) 26개)의 파일로 구성되었다.

본 연구에서 두 유형의 에세이 자료를 분석하는데 사용된 측정치는 총 28개로 이러한 측정치는 기본텍스트 측정치(평균 어휘 수, 평균 문장 수, 평균 단락 수, 평균 문장 길이, 평균 단어 길이), 어휘빈도수 측정치(내용어에 대한 어휘빈도수), 어휘속성 측정치(심상성, 구체성, 습득나이), 어휘다양성 측정치(타입-토큰 비율), 주요 품사 측정치(명사, 동사, 형용사, 부사), 인칭대명사 측정치(1인칭, 2인칭, 3인칭 대명사), 접속사 측정치(인과, 부가, 시간 접속사), 표준 읽기가독성 측정치(Fleach Reading Ease 점수, Flesch-Kincaid Grade Level 점수), 통사구조 복잡성 측정치(명사구 밀도, 본동사 앞 어휘수), 참조정합성 측정치(인접된 문장 간 논항중복 비율, 모든 문장 간 논항중복 비율), 의미적 유사성 측정치(인접 문장 간 LSA 코사인, 모든 문장 간 LSA 코사인)를 포함한다.

4. 연구결과 및 논의

본 연구에서는 논설문 요약문과 설명문 요약문 코퍼스의 차이를 통계적으로 검증하기 위해 일원반복측정설계(one-way repeated measures design)가 사용되었고 유의수준은 5%로 설정하였다.

4.1 기본텍스트 측정치 분석 결과

표 2에 제시된 것처럼, 논설문 요약문의 평균 어휘 수 값은 설명문 요약문의 평균 어휘 수 값보다 통계적으로 유의미하게 높았다. 또한 요약문 코퍼스에 사용된 문장의 길이 측정치 값도 논설문 요약문이 더 높았다.

표 2. 기본텍스트 측정치의 평균(표준편차)

기본텍스트 측정치	논설문 요약문(n=26)	설명문 요약문(n=26)	F	p
평균 어휘 수	79.2(16.0)	66.8(14.5)	16.836	.000
평균 문장 수	4.04(.77)	3.85(.46)	1.499	.232
평균 단락 수	1.08(.27)	1.00(.00)	2.083	.161
평균 문장 길이	19.9(4.0)	17.6(4.1)	5.879	.023
평균 어휘 길이	1.735(.130)	1.738(.135)	.006	.939

반면에 요약문에 사용된 문장 수, 단락 수, 평균 어휘 길이는 두 유형 간에 통계적으로 유의미한 차이를 보이지 않았다. 이러한 결과는 요약문 유형의 차이가 학생의 문체에 영향을 미쳤음을 시사한다. 다시 말하면, 본 연구에 참여한 EFL 학생들은 논설문 요약문을 작성할 때 더 많은 어휘와 더 긴 문장을 사용하였다. 이러한 결과는 주어진 원 자료들의 어휘 수, 문장 수, 문장 길이와 어휘 길이를 측정하고 원 자료의 결과가 이러한 결과에 영향을 미쳤는지도 살펴보아야 할 것이다.

4.2 어휘빈도수 분석 결과

어휘빈도수 분석 결과는 표 3에 제시되었다. Coh-Metrix 소프트웨어가 제공하는 어휘 빈도 수 측정치는 대수변환(logarithmic transformation)을 통해 산출된 값이다. 이러한 이유는 대수변환을 하면 어휘빈도 분포가 정규분포에 더 가까워지기 때문이다(Graesser et al. 2004).

표 3에 제시된 어휘빈도수 분석 결과를 살펴보면, 논설문 요약문에 사용된 어휘 빈도수 값은 설명문 요약문에 사용된 어휘 빈도수 값보다 통계적으로 유의미하게 높다. 일반적으로 텍스트 내에 고빈도(high-frequency) 어휘가 많이 사용되면 그 텍스트의 이해 난이도는 감소하는 경향이 있다. 요약하면, 이러한 결과는 논설문 요약문에 상대적으로 고빈도 어휘가 더 많이 사용되었음을 나타낸다. 다시 말하면, 설명문 요약문에 좀 더 어려운 어휘가 사용되었다고 볼 수 있다. 이러한 결과는 요약문 장르에 의한 차이가 학생들이 사용하는 어휘 빈도에 영향을 주었음을 나타낸다.

표 3. 어휘빈도수 측정치의 평균(표준편차)

어휘빈도 수	논설문 요약문(n=26)	설명문 요약문(n=26)	F	p
어휘빈도 수(내용어)	2.34(.13)	2.24(.21)	4.558	.043

이 결과는 안수진(2018)의 연구에서 설명문과 논설문 모두에서 영어 원어민들이 한국인 영어 학습자들보다 단어 빈도에서 낮은 평균을 보여 출현 빈도가 낮은 어휘 즉 난이도가 더

높은 어휘를 구사했으나, 한국인은 논설문에서 빈도가 높은 어휘 즉 “일상생활에서 빈번하게 접하는 어휘들을 사용하여 글을 쉽게 쓰고”(안수진 2018, pp. 188-189), 설명문에서는 다의성이 높은 어휘들을 많이 사용해서 글이 모호해 지는 경향이 있다는 결과와 맥을 같이 하는 것으로 보인다.

4.3 어휘속성 분석 결과

본 연구에서 분석된 어휘속성 측정치 결과는 표 4에 제시되었다. Coh-Metrix 프로그램은 MRC 어휘 데이터베이스에 등록된 어휘 속성 정보에 기초해서 어휘속성 측정치를 산출한다(Graesser et al. 2004).

표 4. 어휘속성 측정치의 평균(표준편차)

어휘속성	논설문 요약문(n=26)	설명문 요약문(n=26)	F	p
심상성	414.7(17.4)	386.8(29.2)	17.313	.000
구체성	387.1(16.7)	360.3(26.8)	19.325	.000
습득나이	361.9(24.9)	366.3(31.4)	.286	.597

심상성(imageability)은 어휘가 얼마나 잘 형상화되는지를 나타내고 구체성(concreteness)은 어휘가 얼마나 구체적인지를 제시한다. 습득나이 측정치는 어휘가 학습되는 연령 관련 정보를 제시한다. 이러한 측정치의 값은 100에서 700사이로 표시된다. 표 4에 제시된 것처럼, 습득나이 측정치에 대해서는 두 요약문 유형 간에 통계적으로 유의미한 차이가 없었지만 심상성과 구체성 측정치에 관해서는 유의미한 차이가 발견되었다. 이러한 결과는 학생들이 논설문 요약문을 작성할 때 더 구체적이고 형상화가 잘 되는 어휘들을 사용했음을 제시한다.

4.4 어휘다양성 분석 결과

어휘다양성 측정치는 표 5에 제시된 것처럼 타입-토큰 비율(type-token ratio) 값을 나타낸다. 타입-토큰 비율 값은 텍스트 내에 얼마나 다양한 어휘들이 사용되는지를 측정하는데 활용된다. 타입은 텍스트 내에 제시된 각각의 어휘의 수를 나타내고 토큰은 그 어휘가 반복되어 제시되는 수를 제시한다. 따라서 토큰 값은 항상 1이 된다. 구체적인 예를 제시하면, 만약 영어 에세이 내에 ‘topic’이라는 어휘가 5번 사용되면 ‘topic’의 타입은 1이고 토큰은 5가되기 때문에 ‘topic’의 타입-토큰 비율 값은 $0.2(1/5 = 0.2)$ 가 된다.

표 5. 어휘다양성 측정치의 평균(표준편차)

어휘다양성	논설문 요약문(n=26)	설명문 요약문(n=26)	F	p
타입-토크 비율(내용어)	.77(.06)	.91(.06)	68.729	.000

표 5를 살펴보면, 논설문 요약문과 설명문 요약문 두 유형 간에 유의미한 차이가 발견되었다. 이러한 결과는 요약문 유형에 따라 학생들의 문체가 변화되었음을 시사한다. 다시 말하면, 학생들은 설명문 요약문을 작성할 때 더 다양한 어휘들을 사용하는 경향이 있었다.

4.5 주요 품사 분석 결과

표 6은 주요 품사에 대한 분석 결과는 제시한다. 본 연구에서는 내용어(명사, 동사, 형용사, 부사)에 대한 결과를 제시한다.

표 6 주요 품사 측정치의 평균(표준편차)

주요 품사	논설문 요약문(n=26)	설명문 요약문(n=26)	F	p
명사 수	255.2(32.3)	279.9(48.9)	6.529	.017
동사 수	125.7(21.9)	165.9(31.3)	31.580	.000
형용사 수	122.9(21.1)	95.3(34.2)	12.065	.002
부사 수	72.0(30.0)	66.3(27.1)	.588	.450

Coh-Metrix 시스템에서 주요 품사에 대한 측정치의 값은 텍스트 길이에 의해 설명되는 분산(variance)을 통제하기 위해서 발생점수(incidence score)로 변환된다(Graesser et al. 2004). 어떤 어휘에 대한 발생점수는 1,000개의 어휘 당 그 어휘가 얼마나 반복되어 나타나는지를 지칭한다. 표 6에 제시된 결과를 살펴보면, 부사에 대해서는 두 유형 간에 유의미한 차이가 발견되지 않았다. 그러나 학생들은 논설문 요약문을 작성할 때 더 많은 형용사를 사용하였다. 반면에 그들은 설명문 요약문을 작성할 때 더 많은 명사와 동사를 사용하였다. 이러한 결과는 장르의 차이가 학생들이 형용사, 명사, 동사를 사용하는데 영향을 주었음을 시사한다. 부연하면, 주요 품사에 대한 분석 결과는 학생들이 설명문 요약문을 작성할 때 명사와 동사를 더 많이 사용함으로써 원인과 결과의 관계를 좀더 명확하게 설명하려고 시도하는 경향이 있음을 나타낸다. 반면에 논설문 요약문을 작성할 때는 형용사를 더 많이 사용하여 자신의 개인적인 견해, 의견과 주장을 좀 더 자세하고 명확하게 전달하려고 시도하는 경향이 있음을 제시한다. 또한 이러한 결과는 안수진(2018)의 연구에서 설명문에서는 한국인이 부사와 대명사를 원어민 저자보다 유의미하게 더 많이 사용하였고 논설문에서는 한국인 영어 학습자들이 영어 원어민 저자보다 형용사와 동사의 사용빈도가 높았다는 결과와 어느 정도 같은 선상에 있어 보인다. 논설문 요약문에서는 한국인 영어 학습자들이 자신의 개

인적인 견해를 형용사를 사용하여 더 잘 표현하고자 하는 경향을 보여준다고 보인다.

4.6 인칭대명사 분석 결과

인칭대명사 분석 결과는 표 7에 제시되었다. 인칭대명사 측정치도 발생점수로 제시되었다. 표 6에 제시된 것처럼, 학생들은 논설문 요약문 및 설명문 요약문을 작성할 때 2인칭 대명사를 전혀 사용하지 않았다.

1인칭 대명사를 사용함에 있어서는 두 요약문 유형 간에 통계적으로 유의미한 차이가 발견되지 않았다. 그러나 학생들은 논설문 요약문을 작성할 때 더 많은 3인칭 대명사를 사용하였다. 이러한 결과는 장르의 차이가 학생들이 3인칭 대명사를 사용하는 방식에 영향을 미쳤음을 나타낸다. 또한 이러한 결과를 주요 품사의 분석 결과와 연결해 설명할 수도 있다.

표 7. 대명사 측정치의 평균(표준편차)

대명사	논설문 요약문(n=26)	설명문 요약문(n=26)	F	p
1인칭 대명사 수	3.34(6.4)	3.53(7.2)	.013	.911
2인칭 대명사 수	None	None	None	None
3인칭 대명사 수	8.53(8.79)	.47(1.66)	20.643	.000

주요 품사의 분석 결과 설명문 요약문 조건의 학생들은 명사를 더 많이 사용하는 경향이 있었다. 따라서 명사의 사용빈도가 증가하면서 자연스럽게 명사를 지칭하는 대명사의 사용빈도가 감소했을 가능성이 있다고 할 수 있겠다.

4.7 접속사 분석 결과

요약문에 사용된 접속사 분석 결과는 표 8에 제시되었다. 접속사 측정치도 발생점수로 제시되었다. 표 8에 제시된 것처럼, 학생들은 시간 접속사(until, till 등)를 사용할 때는 두 유형 간에 차이를 보이지 않았다. 그러나 설명문 요약문을 작성할 때는 예측할 수 있는 것처럼 원인과 결과의 관계를 설명하기 위해 인과접속사(therefore, because 등)를 수치적으로는 좀 더 많이 사용하였다. 그러나 인과접속사 사용에 의한 차이 역시 통계적으로 유의미한 차이를 나타내지는 않았다.

표 8. 접속사 측정치의 평균(표준편차)

접속사	논설문 요약문(n=26)	설명문 요약문(n=26)	F	p
인과 접속사 수	42.1(17.7)	54.6(28.2)	3.514	.073
부가 접속사 수	19.4(12.1)	9.7(10.5)	9.870	.004
시간 접속사 수	7.8(10.1)	10.3(13.0)	.672	.420

반면에 학생들은 논설문 요약문을 작성할 때 더 많은 부가 접속사(*and, moreover* 등)를 사용하였다. 이러한 결과는 논설문 요약문을 작성할 때 학생들이 논증적 증거들을 제시 및 연결하기 위해 부가적 접속사를 많이 사용했을 가능성이 있다고 해석할 수 있겠다. 또한 이러한 결과는 Bae와 Min(2017)의 연구에서 네 개의 서로 다른 글의 유형 즉, 이야기 즉 서사 글, 원인과 결과글, 비교와 대조글과 논설문 중 한국 대학생들은 논설문에서 가장 많은 연결사(connector)들을 사용하며, 이러한 다양한 연결사 중에서도 부가접속사(enumeration/addition)을 38.6%로 가장 빈번하게 과잉 사용(Bae and Min 2017, p. 25)한다는 결과와 같은 선상에 있는 결과로 보인다. 비록 Bae와 Min(2017)의 결과는 한국인 대학생들이 작성한 글에 대한 조사 결과라 본 연구에서 살펴본 요약문과는 다르지만, 주어진 텍스트를 영어로 요약한 글에서도 한국인 영어 학습자들은 영어 작문을 할 때와 같은 양상을 보이는 것을 알 수 있다. 이러한 결과는 한국 영어 학습자들에게 담화 유형에 따라 다양한 연결사를 사용할 수 있도록 교수 방법에 변화가 있어야 함을 시사한다.

4.8 표준 읽기가독성 지표 분석 결과

표준 읽기가독성 지표 측정치와 관련하여, Coh-Metrix 도구는 교육학 분야에서 텍스트의 읽기 및 이해 난이도를 평가하기 위해서 많이 사용되는 Flesch-Reading Ease(FRE) 점수와 Flesch-Kincaid Grade Level(FKGL) 점수를 제공한다(Graesser et al. 2004).

표 9. 표준 읽기가독성 지표 측정치의 평균(표준편차)

표준 읽기가독성 지표	논설문 요약문(n=26)	설명문 요약문(n=26)	F	p
Flesch-Kincaid Grade Level	11.7(.56)	11.2(.85)	7.128	.013
Flesch Reading Ease score	39.8(9.4)	42.0(9.8)	.723	.403

FRE 점수는 0과 100 사이의 점수로 계산되는데 값이 클수록 읽기 및 이해난이도가 낮음을 의미한다. 반면에 FKGL 점수는 1과 12사이의 점수로 표현되는데 값이 클수록 읽기 및 이해난이도가 높음을 뜻한다. 특히, FKGL 점수는 미국 초중고 학년등급(1학년에서 12학년)에 대응되기 때문에 영어텍스트 및 교재의 난이도를 평가하는데 광범위하게 활용되고 있다. 표 8을 살펴보면, 두 유형의 텍스트는 FRE 점수에 대해서는 통계적으로 의미 있는 차

이를 보이지 않았다. 반면에 논설문 요약문의 FKGL 측정치는 설명문 요약문보다 유의미하게 높았다. 이러한 결과는 논설문 요약문의 읽기 및 이해난이도가 설명문 요약문보다 높음을 시사하며 구체적으로 어떻게 높은지 원 자료에 대한 Coh-Metrix를 통한 정밀하고 신뢰할 만한 분석도 이루어져야 할 필요가 있음을 시사한다. FRE 점수와 FKGL 점수는 사실상 단어길이와 문장길이에 의해 크게 영향을 받기 때문에 이러한 결과는 표 2에 제시된 평균 문장길이의 분석 결과와 일치한다.

그럼에도 불구하고 두 에세이의 FKGL 점수는 각각 11.7(논설문 요약문) 및 11.2(설명문 요약문)로 11학년 이상의 수준을 보여주고 있다. 따라서 본 연구에 참여한 EFL 학생들이 작성한 영어에세이의 읽기 및 이해난이도가 일반적으로 높은 편이라고 할 수 있겠다.

4.9 통사구조 복잡성 분석 결과

문장의 통사 구조의 복잡성과 관련하여 Coh-Metrix 프로그램은 표 10에 제시된 측정치를 제공한다. 명사구 밀도(noun density)는 통사적 복잡성을 분석하는데 사용되는 대표적인 측정치 중 하나로써 명사구 구조의 복잡성을 나타낸다. 다시 말하면, 이 측정치는 명사구의 표제어(headword)를 수식하는 수식어의 개수에 기반하여 계산된다. 구체적으로 명사구 내의 수식어의 수를 명사구 전체의 어휘 수로 나눈 값이 명사구 밀도 점수이다. 따라서 이 값을 클수록 명사구의 길이가 길어지면서 문장의 길이도 증가하기 때문에 전체적으로 문장의 통사적 복잡성이 증가한다. 본동사 앞 어휘 수는 본동사 앞에 있는 평균 어휘 수를 지칭한다. 따라서 명사구 밀도 점수와 유사하게 이 값이 클수록 전체적으로 문장의 통사구조의 복잡성이 증가한다(Graesser et al. 2004).

표 10. 문장 통사 구조 측정치의 평균(표준편차)

통사구조 복잡성	논설문 요약문(n=26)	설명문 요약문(n=26)	F	p
명사구 밀도	.69(.12)	.77(.24)	2.719	.112
본동사 앞 어휘 수	5.8(2.2)	3.7((1.8)	29.687	.000

표 10에 제시된 결과를 보면, 명사구 밀도 점수에 대해 논설문 요약문과 설명문 요약문은 유의미한 통계적인 차이를 보이지 않았다. 그러나 논설문 요약문의 본동사 앞 어휘 수 측정치 값은 설명문 요약문보다 높았다. 이러한 결과는 논설문 요약문의 통사적 복잡성이 설명문 요약문보다 높음을 나타내며 아울러 장르의 변화가 학생들의 문체에 영향을 주었음을 시사한다. 또한 이러한 결과는 표 2에 제시된 평균문장길이 분석 결과 및 표 9에 제시된 FKGL 점수 분석 결과와도 일치한다. 안수진(2018)의 연구에서 설명문과 논설문 모두에서 영어 원어인 저자들이 전치사구, 수동태 밀도가 높았고, 한국인 영어 학습자들은 동사구, 부정사구를 더 자주 사용하는 것으로 나타났다고 한다. 즉 영어 원어인 저자들은 글의 답화

유형에 상관없이 한국인보다 통사적으로 더 복잡한 문장을 많이 작성하였고, “한국인은 일반적으로 짧은 문장을 선호하고 복잡한 구조는 회피”(안수진 2018, p. 195)하고 쉽고 단순하게 쓰려는 경향이 있다고 하였다. 본 연구에서는 명사구 밀도와 본동사 앞 어휘 수만을 조사하여 논설문 요약문에서 한국인 영어 학습자들이 좀 더 통사적으로 복잡한 문장을 작성하는 경향이 있음을 알게 되었다. 이러한 글의 유형에 따라 통사 구조의 복잡성을 연구하기 위해 좀 더 다양한 측정치를 분석해 볼 필요가 있어 보인다.

4.10 참조정합성 분석

참조정합성(co-referential cohesion)은 언어학 및 언어심리학 분야에서 텍스트 이해(text comprehension)를 연구하는데 이용되는 중요한 개념으로써 일반적으로 서로 이웃한 문장 사이에 명사, 대명사, 명사구가 중복되어 제시될 때 형성된다(Graesser et al. 2007). 참조정합성이 형성되면 텍스트의 이해도가 낮아지기 때문에 결과적으로 참조정합성이 높은 텍스트는 이해하기 쉬운 경향이 있다.

본 연구에서는 표 11에 제시된 것처럼, 참조정합성을 분석하기 위해 서로 이웃한 문장 간의 논항 중복과 모든 문장 간의 논항 중복 측정치를 이용하였다. 텍스트 내에서 서로 이웃한 문장 간의 논항 중복은 지엽적 정합성(local cohesion)을 의미하고 특정 문장과 텍스트 내의 다른 모든 문장 사이의 논항 중복은 전반적 정합성(global cohesion)을 제시한다. Coh-Metrix 프로그램의 서로 이웃한 문장 간의 논항 중복 측정치는 논항이 중복된 서로 이웃한 문장 쌍의 수를 텍스트 내의 전체 이웃한 문장 쌍들의 수로 나눈 값이고, 모든 문장 사이의 논항 중복 측정치는 논항이 중복되어 사용된 문장 쌍의 수를 계산하여 이 값을 전체 문장 쌍의 수를 계산하여 나눈 값을 나타낸다(Graesser et al. 2004).

표 11. 참조정합성 측정치의 평균(표준편차)

참조정합성 측정치	논설문 요약문(n=26)	설명문 요약문(n=26)	F	p
논항 중복(인접 문장 간)	.87(.21)	.63(.35)	10.195	.004
논항 중복(모든 문장 간)	.89(.19)	.61(.34)	15.804	.001

참조정합성 분석 결과를 살펴보면, 논설문 요약문의 참조정합성 측정치 값이 모두 설명문 요약문의 값보다 높다. 이러한 결과는 논설문 요약문의 이해난이도가 설명문 요약문보다 낮다고 볼 수 있고 아울러 요약문 장르의 변화가 요약문의 참조정합성에도 영향을 미쳤음을 제시한다. 안수진(2018)은 영어 원어민들의 설명문이 지시적 응집성과 의미적 응집성 모든 지표에서 한국인 영어 학습자들의 설명문보다 높아 응집성이 높은 문장을 사용한다고 주장하였다. 즉 설명하는 글에서 응집성이 높다는 것은 더 명확하게 독자들에게 의미를 설명하고 이해해 도움을 줄 수 있다는 것을 의미한다. 또한, 이는 한국인 영어 학습자들이 응집

장치를 사용하는 데에 다소 부족함이 있다는 것을 의미하는 것으로 보인다(안수진 2018, Bae and Min 2017). 반면에 논설문에서는 한국인이 영어 원어인 저자들에 비해 중복 수치가 높아 난이도가 낮은 쉬운 텍스트를 구사하였다고 보고하였다. 표 6의 주요 품사결과에 제시된 것처럼, 학생들은 논설문 요약문을 작성할 때 더 많은 대명사(3인칭 대명사)를 사용하는 경향이 있었다. 따라서 이러한 경향성으로 인해 텍스트 내에 대명사 사용이 증가하면서 논항(말하자면, 대명사) 중복 측정치 값도 증가했다고 볼 수 있겠다. 이러한 결과들은 한국인 영어 학습자들이 담화 목적 즉 설명하는 글 또는 자신의 의견을 타당하게 주장하는 글에 따라 응집성 있는 글로 요약하는 역량이 다소 부족함을 시사한다.

4.11 의미적 유사성 분석 결과

Coh-Metrix 프로그램의 문장 간의 의미적 유사성(semantic similarity) 측정치는 일종의 수학적 모형인 Latent Semantic Analysis(LSA)에 의해 산출된다. LSA의 기본 원리는 방대한 크기의 코퍼스 자료에 기반해 생성된 일종의 대형 어휘-문장 행렬(matrix)을 수학 공식인 SVD(singular value decomposition)에 의해 축소된 행렬(말하자면, 의미적 공간)로 재생성하여 이를 텍스트 구성 요소(어휘, 문장, 단락) 간의 의미적 유사성을 분석하는데 응용하는 것이다(Landauer, Foltz and Laham 1998). 이 의미적 공간(축소된 행렬) 안에서 어휘 간, 문장 간, 단락 간의 의미적 유사성(말하자면, LSA 코사인 값)이 계산된다.

표 12. 의미적 유사성 측정치의 평균(표준편차)

참조정합성 측정치	논설문 요약문(n=26)	설명문 요약문(n=26)	F	p
LSA 코사인(인접 문장 간)	.40(.12)	.17(.06)	52.956	.000
LSA 코사인(모든 문장 간)	.38(.11)	.16(.06)	49.629	.000

본 연구에서는 표 12에 제시된 것처럼, 두 개의 LSA 값을 분석에 이용하였다. 말하자면, 참조정합성 분석에서처럼 본 연구에서는 서로 이웃한 문장 간의 LSA 측정치와 전체 문장 간의 LSA 측정치 값을 분석하였다. 표 11의 분석결과를 살펴보면, 참조정합성 분석 결과와 유사하게 논설문 요약문의 LSA 측정치 값이 모두 설명문 요약문의 값보다 높았다. 이러한 결과는 참조정합성의 분석 결과와 맥락을 같이하는 것으로 장르의 변화가 학생들의 글을 쓰는 방식을 변화시켰다는 점을 제시한다. 즉 한국인 영어 학습자들은 독자에게 이해가 쉽도록 잘 전달해야 하는 설명문에서 의미적 유사성 수치가 논설문 보다 낮으므로, 설명문을 요약할 때 의미적 유사성이 높아지는 교육을 받을 필요가 있음을 시사한다.

5. 결론

이 연구는 한국인 대학생들에게 영어 논설문과 원인과 결과의 설명문을 읽게 하고 요약하게 한 후 그 요약문을 Coh-Metrix로 분석하여 한국 영어 학습자들의 영어 요약문의 언어적 어휘, 담화, 이독성 측면에서 어떠한 차이가 있는지를 살펴보고자 한 연구이다. 참가자들의 주어진 논설문과 설명문에 대한 요약문을 Coh-Metrix로 분석한 결과 중 통계적으로 유의미한 차이가 있는 것을 요약 정리하면 다음과 같다. 연구 참가자들이 논설문을 요약한 글이 평균 어휘 값과 문장의 길이의 측정치가 통계적으로 유의미하게 설명문을 요약한 것보다 높게 나타났다. 또한, 구체적이고 심상성 높은 어휘와 고빈도 어휘의 사용도 논설문을 요약한 글에서 더 높게 나타났다. 논설문을 요약한 글에서 더 많은 형용사를 사용한 것으로 나타났고, 통사적 복잡성도 높게 나타났으며 부가 접속사의 사용 빈도도 통계적으로 유의미하게 높았다. 반면에 연구 참가자들이 주어진 원인과 결과의 설명문 글을 요약한 요약문에서는 논설문을 요약한 것보다 동사를 더 많이 사용하였고, 더 많은 저빈도 어휘를 사용하였으며, 타입-토크 비율에 따르면 더 다양한 어휘를 사용한 것으로 나타났다. 두 요약문 모두 FKGL지수는 11이상으로 미국 11학년의 어휘 수준으로 나타났으나, 설명문의 어휘수준(어휘빈도)이 논설문의 어휘수준(어휘빈도)보다 더 높은 것으로 나타났다.

본 연구는 기존의 연구들이 간과했던 영어 텍스트의 담화 유형 즉 설명문 또는 논설문에 따라 한국 영어 학습자들의 요약문의 특징이 어휘, 의미, 통사, 담화 그리고 읽기 수준에 따라 어떻게 다른지를 Coh-Metrix라는 객관적이고 신뢰할 만한 도구를 사용하여 양적으로 분석하였다는 점에서 연구 방법적 의의가 있다고 보인다. 또한 본 연구의 결과를 통해 한국인 영어 학습자들은 논설문을 요약할 때 대체로 설명문 요약보다는 통사적으로 복잡하고 단어 및 문장의 길이가 긴 문장을 많이 사용한다는 것을 알게 되었다. 반면에 설명문에서는 저빈도 어휘를 많이 사용하고 정합성이 다소 떨어져서 독자들에게 설명을 해야 하는 담화의 목적과 특성을 잘 반영하지 못하는 요약문을 작성하고 있다는 점도 유추할 수 있다. 이러한 결과를 바탕으로 본 연구는 한국인 영어 학습자들에게 영어 담화의 특성과 목적에 대한 인식을 높이고 목적에 적합한 요약문을 작성할 수 있도록 효과적인 영어 작문 교수 학습 방법이 개발되어야 함을 시사하므로 교육적 의의도 제공한다.

본 연구는 다음과 같은 한계도 지니고 있다. 첫째, 연구 참가자들이 서울에 위치한 영어 교육을 전공하는 예비교사들이므로 본 연구의 결과를 한국 대학생 전체로 일반화하기에는 무리가 있다. 또한, 설명문에 해당하는 원 자료의 글이 원인과 결과에 대한 글이므로 비교와 대조의 담화 유형을 지닌 설명문도 이러한 결과라고 보기에는 무리가 있다. 이에 후속 연구들은 설명문 요약문의 언어적 특징을 살펴볼 때에 원인과 결과 글뿐만 아니라 비교와 대조의 설명문도 포함하여 분석하면 좀 더 일반화 할 수 있는 결과를 도출할 수 있을 것으로 보인다. 이와 더불어, 원 자료 글인 설명문과 논설문의 주제에 대한 연구 참가자들의 친숙도(topic familiarity)에 따라 요약문의 특징이 다를 수 있으므로 후속 연구들은 다양한 주

제의 논설문과 설명문을 요약하게 한 후 그 요약문을 분석해 볼 필요가 있다. 본 연구에서는 원 자료 텍스트의 어휘, 의미, 통사, 담화와 읽기 수준을 살펴보지는 않았으나, 원 자료의 언어학적 특징들도 학습자들이 작성한 요약문에 영향을 미칠 수도 있으므로 후속 연구들은 원 자료의 이러한 언어적 특징이 요약문에 어떠한 영향을 미치는 지에 관해서도 조사할 필요가 있다. 추가적으로 본 연구는 한국인 영어 학습자들이 영어의 담화 유형에 따라 요약문 작성에서 차이가 있었다는 것을 밝히는 것에 그쳤으므로, 이러한 본 연구의 결과를 바탕으로 한국인 영어 학습자들이 부족한 부분을 중점적으로 영어 읽고 쓰기 교육에 적용한 후에는 어떠한 차이가 있는 지도 후속 연구에서 이루어진다면 영어 쓰기 교육에 주요한 시사점을 제공할 것으로 기대된다.

참고문헌

- 곽혜영(Kwak, H.). 2018. 한국인 대학생 영어 학습자의 요약문에 나타난 패러프레이즈 사용에 관한 연구(A study on the paraphrasing usage of Korean university students in English summary writing). 《언어과학연구》(*Journal of Linguistic Science*) 85-1, 1-23.1
- 김미혜·이성원(Kim, M. and S. Kim). 2018. EBS 자료와 CSAT 및 영어교과서 어휘 분석(Vocabulary analysis on EBS materials, CSAT, and English textbooks). 《예술인문사회융합멀티미디어논문지》(*Asia-pacific Journal of Multimedia Services Convergent with Art, Humanities, and Sociology*) 8-4, 417-427.
- 김정덕(Kim, J. D.). 2019. 한국 대학생들의 요약하기 능력 부족에 관한 질적 연구(A qualitative study on Korean college students' lack of summarizing ability). 《영어교과교육》(*Journal of the Korea English Education Society*) 18-1, 1-23.
- 김제우·이동주(Kim, J. and D. J. Lee). 2017. 코퍼스기반 고등학교 어휘 사용 연구(A corpus-based study on the use of vocabulary in high school English I·II textbooks, college scholastic ability tests, and EBS materials). 《언어과학연구》(*The Journal of Linguistic Science*) 80, 51-74.
- 배식한(Bae, S. H.). 2019. 담론의 네 양식을 이용한 대학 글쓰기 교육- 묘사, 서사, 설명, 논증의 필요성과 활용 방안(Academic writing education using four modes of discourse: The need for description, narration, exposition, and argumentation and a way to use them). 《리터러시연구》(*The Korean Journal of Literacy Research*) 10-1, 239-268.
- 성일호(Sung, I.). 2014. 이독성 공식과 Coh-Metrix를 활용한 우리나라 고등학교 영어교과서 이독성 분석(An analysis of Korean high school English textbooks through

- readability formulae and Coh-Metrix). 《영어영문학연구》(*Studies in English Language & Literature*) 40-4, 299-320.
- 안수진(Ahn, S.). 2008. 한국인 대학생과 영어 원어민 대학생의 영어 설명문, 논설문의 코메트릭스 분석을 통한 비교(An analysis of Coh-Metrix on the differences in English expository and argumentative writing of Korean and native English university student). 《새한영어영문학》(*New Korean Journal of English Language and Literature*) 60-3, 177-205.
- 전문기·임인재(Jeon, M-G and Lim, In- Jae). 2010. 고등학교 영어 교과서 읽기 자료의 코퍼스 언어학적 비교 분석(A corpus-based linguistic analysis of high school English textbooks). 《외국어 교육》(*Foreign Languages Education*) 17-1, 209-233.
- 전문기(Jeon, M-G). 2011. Coh-Metrix를 이용한 중학교 1학년과 2학년 개정 영어교과서 읽기자료의 코퍼스 언어학적 연계성 분석(A corpus-based analysis of the continuity of the reading materials in middle school English 1 and 2 textbooks with Coh-Metrix). 《언어과학연구》(*The Journal of Linguistic Science*) 56, 201-218.
- 최윤희(Choe, Y.). 2013. 국가영어능력평가시험 2급 쓰기 과업 유형이 한국 영어 학습자의 영어 쓰기 수행에 미치는 영향(The effects of writing task types in the National English Ability Test level 2 on Korean EFL learners' English writing performance). 《영어학》(*Korean Journal of English Language and Linguistics*) 13-1, 161-186.
- 황영민·이경남(Hwang, Y-M. and K-N, Lee). 2014. 고등학교 영어 교과서와 대학수학능력 영어시험의 어휘 비교. 《외국어교육》(*Foreign Languages Education*) 21-1, 191-214.
- Bae, Jiyoung and Sujung Min 2017. A genre-based study of connectors in L2 students' academic writing. *Secondary English Education* 10-1, 3-30.
- Bailey, S. 2015. *Academic Writing: A Handbook for International Students*. New York: Routledge.
- Choi, Y. H. 2012. Paraphrase practices for using sources in L2 academic writing. *English Teaching* 67-2, 51-79.
- Graesser, A. C., M. Jeon, Z. Cai and D. S. McNamara. 2008. Automatic analyses of language, discourse, and situation models. In J. Auracher and W. van Peer eds., *New beginnings in literary studies*, 72-88. Cambridge: Cambridge Scholars Publishing.
- Graesser, A. C., M. Jeon, Y. Yan and Z. Cai. 2007. Discourse cohesion in text and

- tutorial dialogue. *Information Design Journal* 15-3, 199-213.
- Graesser, A. C., D. S. McNamara, M. M. Louwerse and Z. Cai. 2004. Coh-Metrix: Analysis of text on cohesion and language. *Behavioral Research Methods, Instruments, and Computers* 36-2, 193-202.
- Hwang, E. 2012. *Korean EFL College Learners' Linguistic Features in Narrative and Argumentative Writings in terms of CAF*. Unpublished doctoral dissertation, Sookmyung Women's University, Seoul.
- Hwang, H. J. 2010. A comparison of summarizations of different text genres and their relationship to reading ability. *Korean Journal of Applied Linguistics* 26-3, 69-99.
- Kim, N. 2015. The use of paraphrase in L2 writing with two factors: L2 proficiency and source text genre. *Modern English Education* 16-3, 71-97.
- Kim, S-A. 2016. Characteristics of Korean high school students' summary writing of English reading texts. *Korean Journal of Applied Linguistics* 32-4, 89-118.
- Kim, Y. 2013. Students' learning processes and outcomes through text summary as a reading strategy. *Korean Journal of English Language and Linguistics* 13-1, 17-49.
- Landauer, T. K., P. W. Foltz and D. Laham. 1998. Introduction to latent semantic analysis. *Discourse Processes* 25, 259-284.
- McGee, I. 2014. The pragmatics of paragraphing English argumentative text. *Journal of Pragmatics* 68, 40-72.
- Min, S., J. K. Paek, and Y. Kang. 2019. Exploring the use of hedges and stance devices in relation to Korean EFL learners' argumentative writing qualities. *English Teaching* 74-1, 3-23.
- Park, H. and H. Lee. 2018. The effects of task complexity on Korean adult EFL learners' summary writing. *Modern English Education* 19-1, 62-75.
- Park, S-K. 2013. Korean EFL learners' use of cohesive devices in narrative and argumentative essays. *Studies in English Education* 18-2, 51-81.
- Robinson, P. 2001. Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics* 22-1, 27-57.
- Yasuda, S. 2011. Genre-based tasks in foreign language writing: Developing writers' genre awareness, linguistic knowledge, and writing competence. *Journal of Second Language Writing* 20, 111-133

예시언어(Examples in): 영어(English)
적용가능 언어(Applicable Languages): 영어(English)
적용가능 수준(Applicable Level): 모든 수준(All)

전문기(Jeon, Moongee)
교수(Professor), 영어영문학과(Dept. of English)
건국대학교(Konkuk University)
서울시 광진구 능동로 120(120 Neungdong-ro, Gwangjin-gu, Seoul 05029, Korea)
Tel: 02) 450-3356
E-mail: mjeon1@konkuk.ac.kr

최윤희(Choe, Yoonhee)
교수(Professor), 영어교육과(Dept. of English Education)
충신대학교(Chongshin University)
서울시 동작구 사당로 143(143 Sadang-ro, Dongjak-gu, Seoul 06988, Korea)
Tel: 02) 3479-0382
E-mail: yhchoi@chongshin.ac.kr

논문 접수(Received): 2019년 8월 10일 (August 10, 2019)
논문 수정(Revised): 2019년 9월 10일 (September 10, 2019)
게재 확정(Accepted): 2019년 9월 19일 (September, 19, 2019)