



한국 예비 대학생의 영어 논쟁적 에세이 장르에 대한 다차원 분석

박광현*

명지대학교

황은경

명지대학교

ARTICLE INFO

Received 21 June 2019

Revised 23 July 2019

Accepted 14 August 2019

Examples in: English

Applicable Languages: English

Applicable Levels: Tertiary

KEYWORD

L2 writing/

multi-dimensional analysis/

genre/

영작문/

다차원 분석/

장르

ABSTRACT

Park, Kwanghyun & Hwang, Eunkyung. (2019). A multi-dimensional factor analysis of pre-university students' argumentative essays. *Modern English Education*, 20(3), 129-138.

In L2 writing pedagogy, research on students' essays, holistic evaluation of essays would typically employ qualitative—and manual—techniques, and as a result, findings are not readily scalable to a larger amount of data. Multi-dimensional analysis is a statistical and computational technique to identify underlying linguistic dimensions in a text, which offer a holistic account of the text. Using multi-dimensional analysis, this study examines 881 argumentative essays produced by Korean students as a requirement in a university entrance exam. These essays are compared with essays written by American students. Then, these two sets of essays are compared with a number of other text types, e.g. official letters and newspaper editorials, from two large representative English corpora. The results of this comparison indicate that Korean students' essays show distinct differences from native essays in two linguistic dimensions. Specifically, Korean students' essays are similar to spontaneous speech, telephone conversation, and fiction, while native essays are similar to press editorials, reportage, and academic prose. The findings suggest that Korean students can refer to these target genres in improving their writing and some of these genres can be included in the university writing curriculum.

I. 서론

논쟁적 에세이(argumentative essay)는 영작문 교육과 능숙도 평가의 핵심 장르이다. 이 장르는 IELTS나 TOEFL과 같은 표준화 영어 능력 평가에 포함될 뿐만 아니라, 대학의 기초 영어 글쓰기 교육에서도 필수적으로 다루는 장르이다. 한국 학생들의 에세이에 대한 평가는 일반적으로 2인 혹은

그 이상의 평가자들이 어휘, 문법, 글의 전개(organization), 내용, 어휘와 문법 상의 오류 등과 같은 항목으로 나누어진 평가 기준표를 사용하여 점수를 주고 이를 근거로 총 점수를 산출하는 방식을 사용한다. 경우에 따라, 평가자들이 점수와 더불어 평가 근거를 서술해두는 경우도 있다. 그런데, 평가자의 직관에 의존한 점수와 평가 근거 서술 내용만으로는 평가 기준표의 항목을 넘어서, 학생들의 에세이에 나타

* 제1저자: 박광현(명지대학교, 교수), 교신저자: 황은경(명지대학교, 교수)

박광현

명지대학교 영어영문학과

03674 서울특별시 서대문구 거북골로 34

Tel: (02)300-0539 / Email: kpark@mju.ac.kr

황은경

명지대학교 방목기초교육대학

03674 서울특별시 서대문구 거북골로 34

Tel: (02)300-0830 / Email: hwangek@mju.ac.kr

난 장르적 특성을 이해하기에 부족하다. 학생들이 에세이라는 장르의 성격을 이해하고 이를 바탕으로 적절한 어휘와 문법을 사용하여 글을 쓰도록 교육과정을 준비하기 위해서는 먼저 학생들이 쓴 에세이의 장르적 특성의 이해가 선행되어야 할 것이다.

한국 학생들의 영어 에세이에 나타난 장르적 특성을 이해한다는 것은 텍스트의 어휘 문법적 요소들, 예를 들어, 문장의 길이, 단어의 수와 수동태, 관계절 등 문법 요소의 분석과 함께, 어휘, 문법의 상관 및 복잡도 (E. Hwang, 2013) 또한 고려해야 한다. 아울러, 이들이 이루는 관계를 파악하고 이를 통해서 에세이 장르에서 글쓰기가 이루고자 하는 수사적 목적을 폭넓게 이해해야 한다. 이미 많은 연구를 통해 제 2언어 학습자들의 글쓰기에서 어휘, 문법적 특징과 더불어 수사적 전략과 스타일의 역할이 중요함은 정립된 제안이다 (Hinkel, 2003). 그런데 만약 학생들이 쓴 에세이를 수작업을 통해 담화 수준에서 분석하려 한다면 연구자가 텍스트 분석의 경험과 직관을 가지고 정독해야 하고 이는 많은 시간과 노력이 필요할 것이다.

따라서 본 연구에서는 다차원 분석이라는 통계적 분석 기법을 응용하였다. 다차원 분석은 요인 분석을 통하여 텍스트에 사용된 어휘, 문법적 요소 간의 연관(association)과 군집(cluster) 관계를 밝혀내고, 어떤 요소들이 높은 연관성을 가지고 군집을 이루고, 이 군집이 어떤 언어적 차원을 반영하며, 이 차원들이 어떻게 텍스트의 장르적 특성에 영향을 미치는지 보여주는 통계적 기법이다(Biber, 1988). 다차원 분석 과정의 대부분은 전산적으로 자동화되어 있으므로 결과의 일관성을 확보할 수 있을 뿐 아니라 많은 양의 자료도 처리하고 상호 비교가 가능하다는 장점이 있다. 특히, 본 연구는 한국의 한 대학에 입학하기 위해 영어 에세이 쓰기 시험을 치른 학생들, 즉 예비 대학생들(이후 통칭 한국 학생)의 글을 분석하는 것이 그 목적인 바, 1000명이 넘는 학생들이 쓴 원 자료에서 35만 단어를 상회하는 분량의 에세이의 장르적 특성을 분석하기 위해서 다차원 분석을 적절한 분석 도구로 선택하였다.

본 연구에서는 한국 학생들이 쓴 약 881개의 에세이에 나타난 글의 특성이 장르의 측면에서 어떤 차이를 보이는지 다차원 분석을 통해 살펴볼 것이다. 이를 한국 학생들의 에세이와 원어민의 에세이 사이에 차이가 존재하는지, 그렇다면 어떤 언어적 차원에서 차이가 존재하는지를 알 수 있다. 이 결과는, 그러나, 차이의 유무를 보여줄 뿐, 이 차이가 어느 정도나 큰 것인지를 알 수는 없다. 예를 들어 한 차원에서 1.5 정도의 차이가 난다고 한다면 이 차이가 글쓰기 장르의 측면에서 무슨 의미인지 알기가 어렵다. 따라서 다른 장르와 비교해서 상대적으로 이해해야 하는데, 1.5가 예를 들어 '신문기사와 소설의 차이' 정도라고 한다면, 이미 알려진 장르에 준하여 상대적으로 그 차이를 가늠할 수 있을 것이다. 그러므로 분석 결과를 바탕으로, 한국 학생들과 원어민의

에세이 장르와 원어민이 사용하는 다른 장르 사이에 차이를 비교해보고, 각 자료에 유사한 근접 장르를 통해 그 의미를 논의할 것이다. 비교를 위한 원어민의 장르는 150만 단어 수준의 원어민 영어 참조 코퍼스(reference corpus)를 통해 진행된 기존의 다차원 분석 연구결과를 참조하였다 (Biber, 1988).

한국 학생들과 원어민 에세이에 대한 다차원 분석을 위해 다음과 같이 연구 문제를 설정하고 차원별로 5개의 하위 연구 문제를 설정하였다.

연구 문제: 한국 학생과 원어민의 에세이는 영어 참조 코퍼스의 다른 장르들에 대한 기존 연구에서의 다차원 분석 결과와 어떤 유사성 또는 차이점을 보이는가?

- 1) 상호작용 차원에서 차이와 근접 장르는 무엇인가?
- 2) 내러티브 차원에서 차이와 근접 장르는 무엇인가?
- 3) 명시성 차원에서 차이와 근접 장르는 무엇인가?
- 4) 맥락 의존 차원에서 차이와 근접 장르는 무엇인가?
- 5) 추상성 차원에서 차이와 근접 장르는 무엇인가?

II. 이론적 배경

1. 다차원 요인 분석

다차원 분석(Multidimensional Analysis)은 1990년 후반 경부터 언어 사용 형태를 광범위하게 분석하기 위한 방법론으로써 사용되었다(Biber, 1995, 2009; Biber, Conrad & Reppen, 1998; Biber et al., 2004; Biber, Davies, Jones, & Tracy-Ventura, 2006; Biber & Finegan, 2001; Conrad & Biber, 2001). 문자언어와 음성언어의 차이를 규명하거나, 장르 간의 차이를 많은 언어 자료를 통해 분석하는 공시적 연구는 물론이거니와, 300년에 걸친 영국의 과학 장르의 글을 분석하는 통시적 연구의 방법론으로도 사용되었다(Atkinson, 1996; Biber, Conrad, Reppen, Byrd, & Helt, 2002; Biber et al., 2004). 이와 같이 대규모 연구가 가능해진 것은 언어 코퍼스가 활발하게 구축되고, 이를 이용하기 위한 언어학적 지식, 그리고 전산 기술의 발달이 뒷받침되었기 때문이다. 특히 지난 20년간 학습자의 영어 자료를 모은 학습자 코퍼스 연구 성과가 빠르게 나타남에 따라 학습자 글쓰기에 다차원 분석 기법이 활용될 가능성도 더욱 높아졌다고 하겠다(Granger, 1996, 2002; Granger, Hung, & Petch-Tyson, 2002, 2010; Granger & Rayson, 1998; Granger & Tyson, 1996). 최근에는 다차원 분석 기법을 영어 교육 분야에 활용한 연구들도 발표되고 있고 한국의 영어교육에서도 활용된 바가 있다(Aguado-Jiménez, Pérez-Paredes, & Sánchez, 2012; Asención-Delaney & Collentine, 2011; H. Yoon, 2012). 그러

나 현시점에서 다차원 분석 기법을 한국 영어 학습자들의 에세이에 적용하여 분석한 예는 매우 적다.

초창기의 다차원 분석은 1988년에 발표된 Biber의 연구에서 그 방법론이 정립되었다(Biber, 1988). 이 연구는 크게 두 가지 목적을 가지고 있었는데, 첫째, 한 언어에서 서로 연관관계를 가지는 요소들을 알아내고, 둘째, 언어 요소들 사이의 연관관계를 바탕으로 문어체와 구어체 텍스트의 차이를 규명하는 것이다. 다차원 분석에서 핵심이 되는 개념은 언어 요소 간의 공기(co-occurrence), 그리고 이들이 이루는 연관관계가 반영하는 기저에 존재하는 언어적 차원들(linguistic dimensions)이다. 언어적 차원이라 함은 저자의 스타일 혹은 글쓰기의 목적을 상징하는 추상적인 개념이다. 따라서 차원은 직접 관찰을 할 수 없고, 언어 요소들이 통계적으로 높은 상관관계를 보이며 군집(clustering)하는 형태를 보아서 그 존재를 추론해야 한다(Conrad & Biber, 2001).

차원 추론 과정은 정량적이고 통계적이어서 일견 매우 기계적으로 보이나, 인지적 측면에서 보면 사람이 글을 읽고 장르의 차이 혹은 스타일의 차이를 판단해내는 것과 상당히 유사하다. 사람은 스타일 혹은 글의 종류의 차이를 기계에 비해 높은 수준의 통찰과 직관을 통해서 인식하지만, 어떻게 텍스트가 서로 다른 장르에 속하는지, 혹은 저자들이 서로 다른 스타일을 가지는지를 정량적으로 설명하는 것은 매우 어려운 일일 것이다. 한편 다차원 분석과 같은 기계적인 방법론은 구체적이고 정량적인 언어적 근거를 바탕으로 상이한 스타일과 장르를 추정해낸다. 이 두 가지 방법은 목적을 달성하기 위한 구체적인 기법과 과정이 다르지만, 언어적 요소의 관계를 출발점으로 하여 텍스트의 특성을 판단한다는 지향점은 유사하다. 이러한 점에서 다차원 분석은 사람이 텍스트에서 언어 요소들이 가지는 기능을 객관적으로 이해하도록 도울 수 있는 유용한 테크닉이다.

다차원 분석 방법론의 궁극적 목적은 언어 요소들의 차원이 텍스트에서 수행하는 기능을 밝혀내는 것이다. 차원을 추론하고 그 기능을 규명해내는 데는 크게 두 가지 단계를 거친다. 먼저, 통계적인 기법을 통해 언어 요소의 군집을 찾아내는 정량적 분석 단계인데, 이 단계의 작업은 전산적 방법을 이용하여 상당 부분 자동으로 해낼 수 있다. 다음 단계에서는 군집한 언어 요소들의 개별적 기능을 파악한 후, 왜 특정한 언어 요소들이 서로 공기하는지 정성적으로 분석하는 단계이다. 이 단계는 연구자의 분석을 통해 공기하는 언어 요소들이 어떤 차원을 반영한다고 볼 수 있는지, 만약 그렇다면 그 차원은 텍스트에서 어떤 언어적, 사회적, 심리적 기능을 하는지 추론하는 단계이다. 이 단계에서는 대화 분석, 담화 분석, 텍스트 분석 및 기능 중심 언어학에서 진행된 연구결과가 그 기반이 된다(Halliday, 1994; Martin, 1992).

차원과 언어 요소의 관계에 대해 Conrad와 Biber의 예를 빌리자면, 면대 면 대화에서는 축약형, 약어, 단어나 표현의 반복이 많고, 추상적 표현, 복잡한 명사형 문구, 관계대명사

와 관계사절은 적다(2001). 즉 면대 면 대화라는 특정한 장르에서 공기하는 언어 요소들이 있는 반면, 공통적으로 이 장르를 회피하는 언어 요소들도 있다. 만약 이들 언어 요소는 높은 상관관계로 이 장르에서 공통적으로 나타나거나, 혹은 이 장르에서만 공통적으로 나타나지 않는다면, 이는 이 언어 요소들이 면대 면 대화라는 장르에서 특정한 기능을 하고 있기 때문이다. 이 기능은 예를 들면, 대화를 부드럽게 하고, 상대를 배려하고, 또는 자신의 이야기를 전달하고 상대의 반응을 얻어내는 것일 수 있다. 그렇다면 이런 기능을 수행하는 언어 요소들의 군집은 역동적인 의사소통이라는 기능을 수행하기 위해 '상호작용'이라는 차원을 이루고 있을 수 있다. 이와 같은 차원은 언어를 사용하는 사회생활의 모든 측면에 무수히 존재한다. 배려와 존중을 위한 발화가 있다면 위협과 기만을 위한 발화도 있다. 그 장르적 분류가 배려이든 위협이든 언어적으로 보자면 서로 다른 언어 요소를 조합해서 수행되는 담화적 기능이다. 다차원 분석을 통해 이 두 장르를 분석한다고 상상해보면, 배려는 상대의 의견을 묻는 의문문의 형태와 개인적 의사표현 동사, 그리고 조심스러운 표현을 하는 헤지(hedge)들이 공기할 듯 하고, 위협은 조건문과 짧은 문장, 그리고 명령문들이 공기하고 있을 가능성이 있다라는 분석 결과를 가정해볼 수 있겠다.

언어 요소를 찾아내기 위해서는 전자적으로 처리된 대량의 언어 자료에서 언어 요소들의 빈도를 계산해야 한다. 따라서 언어 자료를 구축해 둔 코퍼스과 더불어, 이 코퍼스를 검색(query) 해낼 소프트웨어 혹은 프로그래밍 기술이 필요하다. Biber의 경우, Lancaster-Oslo-Bergen Corpus(LOB 코퍼스)와 London-Lund Corpus(LL 코퍼스)를 기초자료로 이용하였다(1988). LOB 코퍼스는 1960년대 영국 영어를 수집하여 구축한 100만 단어 규모의 코퍼스이고 LL 코퍼스는 50만 단어 규모의 구어 영어 코퍼스이다. 이 코퍼스들을 구성하는 장르는 본 연구에서 에세이 장르의 특성을 상대적으로 파악하기 위한 기준이 된다. LOB 코퍼스에는 15개 장르가 포함되어있고 LL 코퍼스에는 6개의 구어 장르가 포함되어 있다. 문어 장르의 예는 방송 보도, 신문 사설, 종교적 산문, 취미에 관한 글, 일기, 자서전, 공문서, 학술문, 소설, 유머 등이 있고, 구어 장르는 면대 면 대화, 전화 대화, 공개토론회 인터뷰, 방송, 즉흥적 연설 등이 있다.

장르별 자료가 포함된 코퍼스에서 언어 요소의 기능 분석을 하기 위해서 서로 다른 기능을 하는 문법 항목들의 분류가 필요하다. Biber는 언어를 통해서 표현하는 16가지의 문법/기능 분야를 기술하였는데, 여기에는 시제와 시상, 장소와 시간과 같이 사회적이고 생활과 관련된 분야들이 있는 한편, 명사형, 수동형, 종속절, 전치사구, 조동사 등과 같이 언어구조/형태적 분야와, 어휘의 구체성 등 개념적 분야도 있다(1988). 각 문법/기능 분야를 나눈 후에는 각 언어 요소의 빈도를 파악하고, 이를 이용하여 장르 내에서 어떤 기능을 하는 요소들이 어떻게 군집을 이루는지 조사하여 차원을

추론하게 된다.

2. 차원의 기능적 해석

Biber는 LOB와 LL 코퍼스를 이용한 다차원 분석연구에서는 67개의 언어 요소를 사용하여 총 7개의 차원을 발견하였고 각 차원이 수행하는 기능을 중심으로 분석하였다(1988, p. 102). Conrad와 Biber에 의하면, 7개 차원 중 마지막 6번과 7번 차원은 중요도가 높은 언어 요소의 수가 적고, 따라서 해석이 더 어려워져 다차원 분석을 응용한 대부분의 연구들은 6번과 7번 차원을 제외하고 진행되었다고 밝히고 있고, 같은 이유로 자신들의 연구에도 6번과 7번 차원은 포함하지 않았다(2001, p. 39). 따라서 본 연구에서도 이 두 가지 차원을 제외하고 차원 1부터 5까지 주요 5개 차원을 중심으로 이에 해당하는 53개 언어 요소를 살펴본다.

차원 1은 '상호작용 대 정보전달' 차원이다. 즉 텍스트의 주된 기능과 목적이 정보의 전달에 있느냐 글쓰기가 상대방에게 자신의 입장과 생각을 제시하는데 있느냐의 차이를 보여주는 차원이다. 정보 전달을 중요시하는 장르는 공문서나 공식서한, 뉴스와 신문과 같은 방송매체 등이고, 상호작용을 주된 글쓰기의 요인으로 하는 장르는 소설, 인터뷰, 편지, 대화 등이다. 상호작용이 텍스트 구성의 주된 요인인 경우 높은 상관관계를 보이면서 군집을 이루는 언어 요소는 개인 의견을 피력하는 동사(private verbs), 생각과 추약, 현재형 동사, 의문문 등이다. 이 언어 요소들은 글쓰이의 존재가 글에 드러나게 하고 글쓰이의 생각과 의견, 느낌 등을 전달하기 위한 장치들이다. 반대로 정보전달이 더 중요한 글인 경우 주된 언어 요소들은 명사, 긴 단어, 형용사, 전치사 등이다.

차원 2는 내러티브 차원인데, 글쓰기의 주된 동기가 이야기를 서술하기 위한 것인지 그렇지 않은지를 나타내는 차원이다. 내러티브란 곧 스토리텔링을 의미하는데, 이 차원에 긍정적 기여를 하는 언어 요소들은 과거시제, 완료 시제 등 이야기를 하기 위해 필요한 언어적 장치들이다. 이 차원의 대표적 장르는 대화, 연설, 방송, 인터뷰 등이다. 차원 1과는 달리 차원 2에서는 부정적 영향을 미치는 언어 요소가 없고, 긍정적 언어 요소의 빈도와 군집만 측정한다.

차원 3은 '명시성 대 맥락의존' 차원이다. 텍스트 자체만 가지고 내용을 이해할 수 있는지(명시성) 혹은 텍스트뿐 아니라 외부 정보가 있어야 텍스트를 이해할 수 있는지(맥락의존)를 가늠하는 차원이다. 특정 상황이나 맥락이 없어도 텍스트의 완결성과 이해도가 높은 장르는 학술논문이나 공문서 등이다. 이들 장르에서는 외부 정보, 예를 들어 논문의 저자의 개인적 특성 또는 문서를 작성한 공무원의 개인적 상황과 상관없이 글 자체만으로 내용을 이해할 수 있어야 한다. 반대로 방송이나 개인 편지의 경우는 맥락의존도가 높은 장르이다. 편지를 예로 들자면 독자는 당연히 편지를 쓴 사람에 대해서 알아야 편지의 내용을 잘 이해할 수 있고,

방송, 예를 들어 자막이 나오는 뉴스의 경우 화면이나 소리와 같은 정보가 있는 경우가 그렇지 않은 경우보다 이해도가 높을 것이다. 차원 3에 긍정적으로 기여하는 언어 요소는 관계절이나 전치사와 결합한 관계사 등이고, 부정적 요소는 시간, 장소를 표현하는 부사 혹은 일반 부사이다.

차원 4는 '명시적 설득표현' 차원인데, 자신의 입장을 상대방에게 설득하려는 의도를 분명히 드러낼수록 명시적이다. 대표적인 장르는 언론 사실과 업무용 서한이다. 이들 장르는 권고나 암시보다는 분명하게 설득의 의도를 강하게 밝혀야 하는 장르들이고 이를 위해서 설득표현 동사(demand, insist), 필요표현 조동사(should, need to)와 같은 언어 요소들이 사용된다.

차원 5는 '추상적 정보' 차원이다. 추상성이 높은 대표적인 장르는 학술 텍스트와 공문서이고 추상성이 낮은 장르는 인터뷰와 소설 등이다. 차원 5의 긍정적 언어 요소는 연결어, 수동태, 과거분사 등이다.

3. 제2언어 연구에서 다차원 분석의 활용

다차원 분석은 제2언어 연구와 교육 분야에서 구어와 문어를 구분하여 활용되어왔는데, 자료 수집의 용이함으로 인해 문어 장르의 연구가 더 활발히 진행되어왔다고 하겠다(Biber & Finegan, 1986; Conrad, 2001). 예를 들어, 스페인의 영어 학습자 코퍼스로 다차원 분석을 시행한 연구에서는, 학습자의 작문 스타일을 분류하고, 스페인어의 언어 요소가 영어의 언어 요소에 영향을 주어 독특한 차원을 구성하고 있음을 밝혔다(Asención-Delaney & Collentine, 2011). 이와 같은 발견은 연구자의 직관에 의존하는 연구와 달리 구체적으로 어떤 언어 요소가 어떻게 작문 전체의 스타일과 장르적 특성에 관련을 가지고 영향을 미치는지 보여준다.

교사의 입장에서 본다면 다차원 분석기법은 장르적으로 유사한 학습자료를 준비하고 학생들의 작문에 피드백을 할 때도 참고할 만한 기초자료로 활용할 수 있다. Hardy와 Römer(2013)는 미국 학생들의 글쓰기 과제를 분석하였는데, 미국 학생들의 글쓰기는 학술과 정보, 의견과 사고, 절차와 비절차, 그리고 가능성과 논리성 등 서로 다른 차원이 작용하여 텍스트의 장르적 스타일에 영향을 주고 있음을 발견하였다. 이 발견은 미국 학생들의 글쓰기와 비영어권 학생들의 글쓰기를 비교할 때 활용될 수 있다. 한편, Csomay(2007)는 영어 말하기 수업에서 패턴을 분석하기 위하여 다차원 분석을 활용하였다. 분석 결과를 토대로 수업의 전체적인 흐름과 이때 사용되는 언어 패턴이 어떤 식으로 구성되고 진행되는지 분석하고 말하기 수업의 구조를 체계적으로 보여주었다. 이 연구는 말하기 교육의 경우, 말하기 수업의 담화 구조를 교사의 직관에 의존하기보다는, 데이터를 통해 체계적으로 파악하여 기술하고, 이를 바탕으로 좀 더 효과적인 수업 구성방법을 구상할 수 있다고 제안하

였다. 한국의 영어 말하기 수업 맥락에서는 대학생들의 영어 스피치 발표 원고를 원어민의 연설과 다차원 분석을 통해 비교하여, 그 차이를 기술한 사례가 발표된 바 있다 (Park & Kim, 2016).

다차원 분석은 또한 학습자료의 적절성을 판단하는데 응용할 수 있다. 예를 들어, 교사가 어떤 학습자료가 가르치고자 하는 장르와 유사한 자료인지를 선별하는 데 활용할 수 있다. Al-Surmi(2012)가 주장했듯이 영어 학습자들이 구어 영어를 학습할 때 어떤 학습자료가 자연스러운 구어 발화를 잘 반영하는지 알기 어렵다는 문제점에 착안하여, 어떤 시트콤이나 드라마가 더 자연스러운 영어발화에 가까운 적절한 학습자료가 될 수 있는지 다차원 분석을 통해서 밝혀낸 연구가 그 사례로 볼 수 있다. H. Yoon(2012)은 한국의 상황에서 다차원 분석을 교과서에 적용하여 체계적으로 분석하고 교육적 함의를 밝힌 최초의 연구이다. 이 연구에서는 중학교 영어 교과서의 언어적 차원을 분석한 후 LOB 코퍼스 와 비교하였다. 그리고 분석결과를 바탕으로 교과서에서 격식체와 명시성이 높다는 점을 발견하였고, 이에 따라 교과서의 구어 학습 자료에 대한 제안점을 제시하였다.

언급한 연구들이 다차원 분석의 활용가능성을 탐색하고 유용한 결과를 제시하였으나 기존 연구만으로 한국 학생들의 영어 글쓰기와 영어 교재에 나타난 자료를 분석하기에는 연구도대가 아직 충분하지 않다. 현재까지는 매우 한정된 범위의 학습자 그룹들, 특히 대학생들의 글이 주로 다차원 분석의 대상이 되었을 뿐, 다양한 학습자 그룹을 살펴본 연구들은 매우 적다. 본 연구에서는 연구의 도대를 확장하기 위한 방법으로 연구 대상 그룹을 예비 대학생들, 즉 대학 입학 직전의 학생들로 확대하여 다차원 분석을 적용해본다. 이를 통해, 한국 학생들이 쓴 에세이와 원어민 학생들이 쓴 에세이가 어떤 면(차원)에서 비슷하거나 다른지 알아보고, 더 나아가서, 각 차원에 영향을 미치는 요인을 추적함으로써 한국 학생들이 왜 원어민 학생들과 다른 스타일의 글을 쓰는지 그 원인을 추론할 수 있다. 궁극적으로 본 연구는 학생들의 작문을 개별적인 어휘와 문법 요소를 넘어서서 장르와 담화 수준에서 이해함을 목적으로 한다.

III. 연구 방법

1. 연구 자료

본 연구에서 다차원 분석을 시행할 대상 자료는 한국의 한 대학교에서 예비 대학생들을 대상으로 시행한 영작문 시험에서 수집한 논쟁적 에세이이다. 작문의 주제는 ‘환경보호를 위해 플라스틱을 완전히 폐기해야 하는가’와 같은 논쟁적 주제에 대한 작문을 하도록 선별되었다. 파일 오류 등 기술적 문제가 없는 자료 중 총 1041명의 학생들이 작성한 에

세이는 전자적 처리를 위해 전처리(preprocessing) 과정을 거쳐 전자 파일로 저장되었다. 이 자료 중에서 CEFR(Common European Framework of Reference) 기준으로 최상위 학생(C1, C2 레벨)과 최하위 학생(A1 레벨)의 에세이 160개를 제외하고 A2, B1, B2 3개 레벨에서 총 881개의 에세이를 대상으로 다차원 분석을 시행한다. 에세이는 먼저 대학에서 의뢰한 외부 평가자들이 TOEFL 점수체계에 맞추어서 채점한 후, 연구자들이 TOEFL의 변환기준을 이용하여 CEFR 레벨로 변환하였다 (<https://www.ets.org/toefl>). 이 자료는 총 197,355 단어이고 성별과 연령 정보는 포함되지 않았다. 한국 학생들의 에세이는 이후 NNS Essays라고 칭한다.

한국 학생들의 에세이와 비교하기 위한 원어민 자료는 두 종류를 사용한다. 첫째, 같은 장르 간 비교를 위한 원어민의 에세이는 LOCNESS 코퍼스(the Louvain Corpus of Native English Essays)를 사용한다. LOCNESS 코퍼스는 비원어민의 에세이와 비교하기 위한 목적으로 만들어진 324,304 단어 규모의 원어민 코퍼스이다. 본 연구는 원어민 국적이 변인으로 작용하지 않도록 149,912 단어 규모의 미국 국적의 원어민이 쓴 에세이만 골라서 분석한다. 이 자료는 NS Essays라고 칭한다. 둘째, 한국과 미국 학생들의 에세이에 나타난 장르적 특성을 영어의 다른 장르들과 비교하기 위한 기준이 될 자료는 전문한 바와 같이 LOB와 LL 코퍼스를 사용한다. 이 두 코퍼스에서 학생들의 에세이와 비교할 장르는 뉴스 리포트, 사설, 대중서적, 민간 속설(정치, 역사 문화 등), 순문학, 전기, 에세이, 학술, 과학, 소설, 편지, 대화, 방송 등이다. (<http://clu.uni.no/icame/manuals/LOB/INDEX.HTM>)

2. 분석 방법과 절차

본 연구에서는 요인 분석을 기반으로 한 통계적 분석 기법인 다차원 분석을 방법론으로 이용해서 두 참조 코퍼스를 분석한 결과를 한국 학생과 원어민의 에세이와 비교한다. 먼저 에세이 자료에 나타난 각 언어 요소들의 빈도를 계산하고, 언어 요소의 빈도를 바탕으로 군집을 이루는 요소들이 해당하는 차원에서 가지는 영향력을 수치화하여 참조 코퍼스의 요인분석 결과와 비교한다. 빈도를 계산할 때는 품사정보가 반드시 필요하므로 공개 품사정보 소프트웨어인 Tree tagger(Schmid, 1994)를 사용하여 각 단어에 품사 정보를 부착하였다. Tree tagger는 무료이고 용량이 작으며 사용법이 간편하여 선택하였다. 다음은 품사 부착의 예시이다.

문장 예시: The car is easy to drive

품사 정보 부착 결과:

The_DT car_NP is_VBZ easy_JJ to_TO drive_VB

위의 예에서 약어는 다음과 같은 의미이다: DT: 관사, 한정사(determiner), NP: 명사(구), VBZ: 현재형 동사(단수), VB:

동사 원형

품사 부착이 끝나면 AntConc 소프트웨어(Anthony, 2007)를 사용하여 총 53개 언어 요소의 출현 빈도를 계산한다. 각 언어 요소는 개별적으로 검색해야 하는데 McEnery, Xiao와 Tono(2006)에서 제시한 검색 테크닉을 참조하였다. 예를 들어 현재형(be) 동사를 찾으려 한다면 일일이 수작업으로 검색하지 않고 *_VBZ 검색 조건을 주어 소프트웨어로 검색한다. 검색 조건을 입력하면 다음과 같이 해당 형태의 언어 요소를 검색해서 빈도수를 보여준다.

다음은 *_VBZ를 이용하여 동사의 현재형태를 검색한 결과의 예시이다.

```
abilities_NNS which_WDT is_VBZ difference_NN
One_CD is_VBZ they_PRP know_VBP
Another_DT is_VBZ they_PRP have_VBP artistic_JJ
And_CC the_DT other_JJ is_VBZ they_PRP
I_PRP think_VBP it_PRP is_VBZ the_DT result_NN
```

검색을 통해 빈도를 구한 후에는 비교를 위해 1000단어 당 빈도로 정규화(normalize) 하였다. 그리고 변환된 빈도를 참조 코퍼스의 분석결과를 기준으로 각 요소의 차원 점수(Z-score)를 구할 수 있다. 이 계산 방법은 Biber의 저서에 자세히 기술되어있다(1988, p. 94). 그리고 각 차원별로 소속된 언어 요소의 차원점수를 더하면 해당 차원의 요인 점수, 즉 차원 점수를 산출할 수 있다. 이렇게 한국 학생들의 에세이와 원어민의 에세이에서 구한 차원 점수를 원어민 코퍼스인 LOB과 LL 코퍼스에 있는 장르들의 차원 점수와 비교하면 장르별로 유사성과 차이를 구체적 수치로 분석할 수 있다.

IV. 연구결과 및 논의

1. 한국 학생들의 에세이와 원어민 에세이의 차원점수와 분포

다차원 분석을 통해 한국 학생들의 에세이와 원어민의 에세이의 차원 별 점수를 산출할 수 있었다. 표 1은 차원 별로 언어 요소의 차원 점수이다.

TABLE 1
Dimension Scores of Korean and US Argumentative Essays

Dimension	NNS essays	NS essays
1. Involved vs. informational	12.66	0.47
2. Narrative vs. Non-narrative	-1.75	-1.40
3. Explicit vs. Situation-dependent	2.40	3.74
4. Overt expression of persuasion	0.70	3.39
5. Abstract vs. Non-abstract	2.01	3.06

차원 분석 결과에 의하면 한국 학생들과 원어민의 에세이는 차원 1과 차원 4에서 두드러진 차이를 보이고 차원 2, 3, 5에서는 유사하거나 차이가 크지 않은 것으로 나타났다.

표 1에서 보면 차원 1에서 한국 학생들의 에세이와 원어민 에세이는 가장 뚜렷한 차이를 보인다. 차원 1은 화자가 자신의 존재를 드러내고 상대방에게 자신의 생각, 입장, 감정 등을 좀 더 직접적으로 전달하는 차원이다. 다음으로는 차원 4에서 차이를 보이는데 차원 4는 직접적인 표현을 사용해서 상대를 설득하려는 언어 요소들이 군집하는 차원이다. 두 집단은 차원 3과 차원 5에서도 약간의 차이를 보인다. 이 차이는 근소하지만 두 집단의 글쓰기 성향을 반영해주는 지표 역할을 할 수 있다.

표 2는 두 집단의 차이를 구체적으로 살펴보기 위해서 각 차원에서 언어 요소의 요인 점수를 산출한 것이다.

TABLE 2
Factor Scores of Linguistic Features in Korean and US Argumentative Essays

Dimension	Feature	NNS essays	NS essays	
1 (+)	private verbs	0.05	-0.23	
	THAT deletion	0.32	-0.16	
	contraction	-0.33	-0.55	
	present tense verbs	0.22	-0.30	
	second person pronouns	-0.48	-0.45	
	DO as pro-verb	-0.32	-0.33	
	analytic negation	1.36	0.29	
	demonstrative pronouns	0.04	0.29	
	general emphatics	1.27	0.74	
	first person pronouns	-0.11	-0.44	
	pronoun IT	0.73	0.23	
	BE as main verb	-0.24	-0.73	
	causative subordination	2.33	1.12	
	discourse markers	-0.38	-0.31	
	indefinite pronouns	-0.57	-0.47	
	general hedges	-0.24	-0.28	
	amplifiers	0.06	-0.18	
	sentence relatives	1.17	1.24	
	WH questions	-0.14	0.56	
	possibility modals	3.78	0.68	
1 (-)	non-phrasal coordination	0.18	0.35	
	WH clauses	0.44	0.47	
	final prepositions	-0.41	-0.33	
	other nouns	1.08	1.18	
	word length	0.49	0.39	
	prepositions	-1.20	-0.39	
	type/token ratio	-4.28	-0.51	
	attributive adjectives	0.44	0.07	
	past tense verbs	-0.89	-0.67	
	third person pronouns	-0.14	-0.30	
2	perfect aspect verbs	-1.24	0.04	
	public verbs	0.64	-0.51	
	synthetic negation	-0.27	-0.06	
	present participial clauses	0.30	0.10	
	WH relative clauses	-0.65	-0.67	
	pied-piping constructions	-0.56	-0.07	
	phrasal coordination	1.49	1.68	
	nominalizations	0.09	0.55	
	3 (+)			

3 (-)	time adverbials	-0.54	-0.27
	place adverbials	-0.11	-0.24
	other adverbs	-1.54	-1.74
4	infinitives	0.92	1.38
	prediction modals	0.07	0.52
	suasive verbs	0.37	0.20
	conditional subordination	0.66	0.60
	necessity modals	0.14	0.72
5	split auxiliaries	-1.29	-0.03
	conjuncts	3.53	1.78
	agentless passives	-0.45	0.44
	past participial clauses	-0.07	0.41
	BY-passives	0.12	0.11
	Past participial WHIZ deletion	-0.53	-0.39
	other adverbial subordinators	-0.02	0.71

표 2의 요인 점수를 참조하면 어떤 언어 요소가 한국 학생들과 원어민 사이의 차이를 가져오는지 정량적으로 파악할 수 있다.

그렇다면 각 차원에서 한국 학생들의 에세이는 원어민의 에세이와 구체적으로 얼마나 차이가 나는 것이며, 이 차이는 영어교육의 관점에서 주목할 만한 것인가? 즉, 차원 1에서 두 집단의 에세이는 약 12점의 차원 점수의 차이를 보이는 데 이 차이는 두 집단의 에세이가 매우 다르다는 의미인지, 또는 근소한 차이일 뿐 인지를 알아야 교육적 함의를 찾을 수 있을 것이다. 이에 대한 답을 구하기 위해서는 두 집단의 차이를 비교할 기준이 있어야 하므로, 영어 코퍼스의 다른 장르의 차원 점수와 비교해야 한다.

2. 에세이와 영어 코퍼스의 장르 간 차원 비교

한국 학생들의 에세이와 원어민 에세이의 언어 요소 별 그리고 차원 별 차이를 수치로 정량화했으므로, 다음으로는 이 차이에 의한 장르적 성격을 기존 문헌에서 참조 코퍼스의 여타 장르별 차원 점수를 산출한 내용과 비교한다. 표 3은 두 집단의 에세이와 영어 코퍼스의 장르들의 차원 점수이고 NNS와 NS에세이를 제외한 다른 장르의 차원 점수는 Biber(1988)을 참조한 것이다.

TABLE 3

Dimension Scores of Student Essays and Genres in English

Genre	1	2	3	4	5
NNS Essays	12.66	-1.75	2.40	0.70	2.01
NS Essays	0.47	-1.40	3.74	3.39	3.06
Academic Prose	-14.90	-2.60	4.20	-0.50	5.50
Adventure Fiction	0.00	5.50	-3.80	-1.20	-2.50
Biographies	-12.40	2.10	1.70	-0.70	-0.50
Broadcasts	-4.30	-3.30	-9.00	-4.40	-1.70
Face-to-face Conversations	35.30	-0.60	-3.90	-0.30	-3.20
General Fiction	-0.80	5.90	-3.10	0.90	-2.50
Hobbies	-10.10	-2.90	0.30	1.70	1.20
Humor	-7.80	0.90	-0.80	-0.30	-0.40

Interviews	17.10	-1.10	-0.40	1.00	-2.00
Mystery Fiction	-0.20	6.00	-3.60	-0.70	-2.80
Official Documents	-18.10	-2.90	7.30	-0.20	4.70
Personal Letters	19.50	0.30	-3.60	1.50	-2.80
Popular Lore	-9.30	-0.10	2.30	-0.30	0.10
Press Editorials	-10.00	-0.80	1.90	3.10	0.30
Press Reportage	-15.10	0.40	-0.30	-0.70	0.60
Press Reviews	-13.90	-1.60	4.30	-2.80	0.80
Professional Letters	-3.90	-2.20	6.50	3.50	0.40
Religion	-7.00	-0.70	3.70	0.20	1.40
Romantic Fiction	4.30	7.20	-4.10	1.80	-3.10
Science Fiction	-6.10	5.90	-1.40	-0.70	-2.50
Spontaneous Speeches	18.20	1.30	1.20	0.30	-2.60
Telephone Conversations	37.20	-2.10	-5.20	0.60	-3.70

표 3에 제시된 바와 같이 각 장르는 차원 별로 고유한 차원 점수를 가지고 있다. 이제 차원 점수를 기준으로 하여 한국 학생과 원어민의 에세이를 영어의 여타 장르와 비교하면 어느 장르가 유사하고 어느 장르가 상이한지 알 수 있다. 두 집단의 에세이는 차원 1과 차원 4에서 가장 큰 차이를 보였으므로 이 두 가지 차원을 중심으로 차원 별 차이를 살펴본다.

그림 1은 차원 1에서 에세이와 여타 장르간의 차이를 보여준다.

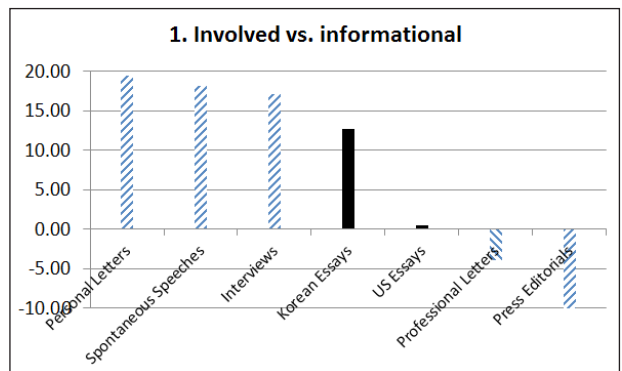


FIGURE 1 Dimension 1 Scores of Student Essays and English Genres

그림 1에서 보면 근접 장르 중 차원 점수가 가장 높은 장르는 개인적 편지이고 그 뒤를 이어서 즉흥적 발화와 이와 유사한 성격의 장르인 인터뷰 등이 있다. 이들은 개인적인 스타일을 강하게 드러내는 장르들인데 한국 학생들의 에세이의 근접장르는 즉흥적 발화와 인터뷰이다. 즉 한국 학생들의 에세이는 마치 즉흥적으로 이야기를 하듯이 또는 인터뷰 질문에 대답하는 듯한 스타일로 구성된다는 의미이다. 한국 학생들은 글에서 독자와 거리를 두어 객관적인 느낌을 주는 글을 쓰기 보다는 자신의 입장과 생각, 그리고 느낌을 직접적으로 드러낸다.

한편 원어민의 에세이는 인터뷰와 같이 개인적인 스타일로 생각, 감정, 입장 전달이 중심이 되는 구어적 장르와 신문 사설과 같이 좀 더 격식을 갖추고 논리적으로 의견을 전달하는 문어적 장르의 중간쯤에서 균형을 갖추고 있다. 유사 장르는 공식 서한인데, 그 목적 상 문어체로 구조와 격식

을 갖추어 내용을 전달하는 것이 중심인 장르이다. 즉 원어민의 에세이는 공식 서한과 같은 스타일을 사용하여 격식을 갖추어 의견을 전하는 것을 목적으로 한다.

다음으로 큰 차이를 보였던 차원 4의 장르 별 차원 점수 분포는 그림 2에서 볼 수 있다.

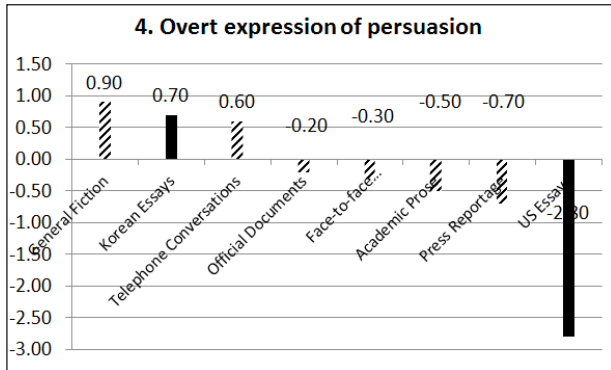


FIGURE 2 Dimension 4 Scores of Student Essays and English Genres

차원 4는 명시적인 설득표현을 반영하는 차원이다. 한국 학생들의 에세이의 유사 장르는 소설과 전화대화이고, 원어민 에세이의 유사 장르는 학술적 산문(대학교과서 지문, 과학 잡지 기사 등)과 신문의 보도기사이다. 명시적인 설득표현은 demand, insist 등과 같이 주장을 나타내는 동사, 그리고 should, must와 같이 당위를 구체적으로 표현하는 조동사를 사용한 표현을 일컫는다. 명시적인 설득표현이 많을수록 좀 더 이해하기 쉽고 분명한 사안을 강조해서 주장하는 맥락이라고 추정할 수 있다. 예를 들어, 전화 대화의 경우 장르 특성상, 학술 산문처럼 복잡한 근거와 논리를 제시하여 주장을 펼치는 것이 어려우므로 좀 더 간단한 형태로 의견을 제시하고 무엇을 주장하는지 분명하게 표현할 필요가 있다. 반대로 신문의 보도기사는 사건의 경위와 사건 전개에의 기승전결 구조와 같이 전체적으로 짜인 구도를 필요로 하여 주어체를 사용하는 대화보다 더 복잡한 구조를 가진다. 학술 산문들은 체계적인 지식을 전달하고 설명해야 하므로 단순하고 분명한 주장만으로 글을 구성하기보다는, 지식 체계에 의한 논리적 전개와 결론이 중요한 장르이다.

차원 1과 차원 4에서 한국 학생들의 에세이와 원어민 에세이를 비교한 결과를 간추려보자면, 한국 학생들의 에세이는 차원 1에서 보듯이 즉흥적이고 대화에 가까운 스타일로 텍스트가 구성되고, 차원 4에서 알 수 있듯이 비교적 간단한 생각과 주장을 제시한 후 주장의 타당성 혹은 중요성을 강조하는 스타일을 취한다. 결과적으로 한국 학생들의 에세이는 전반적으로 마치 의견을 피력하는 인터뷰 장르처럼 자신의 주장을 짧고 분명하게 강조하여 설득하려는 요인이 장르적 특성과 스타일에 영향을 미치고 있다. 이와 비교하여 원어민 에세이는 차원 1에서 나타나듯이 신문 사설이나 공식 서한처럼 주장하는 바가 분명하지만 이 주장은 개인적 입장이 아닌 근거와 논리를 제시하는 언어 요소를 통해 표현되

고, 차원 4에서 나타나듯이 주장을 간략하고 강하게 제시하는 것이 목적이 아니고 학술 산문과 신문 사설처럼 서술 구조를 갖춘 전개를 통해 객관적인 톤으로 제시하여 설득력을 획득하려는 경향이 있다.

V. 결론 및 함의

본 연구는 차원 분석을 통해 한국 학생들의 에세이와 원어민의 에세이의 장르적 차이를 다차원 분석 기법을 이용하여 살펴보았다. 다차원 분석은 어휘 혹은 문법 항목을 개별적으로 살펴보는 분석방식과 달리, 각 언어 요소들이 어우러져 어떻게 글의 전체적인 스타일과 장르를 구성하는 요인으로 작용하는지를 총체적으로 살펴보는 통계 기법이다. 학습자의 글쓰기를 이해하기 위해서는 개별적 분석과 총체적 분석이 모두 유용하지만, 현재 영어교육의 문헌에서는 개별적 분석이 대다수를 이루고, 총체적 분석은 주로 통계기법이 아닌 교사 혹은 평가자의 수작업을 통해 이루어지고 있다. 따라서 통계적, 전산적 기법을 적용해서 교사와 평가자의 총체적 분석을 지원할 방법론적 확장이 필요하다고 하겠다.

본 연구에서 다차원 분석을 적용한 결과, 한국 학생들이 쓴 논쟁적 영어 에세이는 원어민에 비하여 상대적으로 단순한 주장을 강조하는 스타일을 사용하고, 장르로 보자면, 학술적 글쓰기보다는 즉흥적 대화와 유사하였다. 원어민 에세이가 신문 사설과 유사하게 여러 층위의 근거 제시를 통하여 주장에 강한 설득력을 신는다는 점이 그 근본적 차이라는 점을 알 수 있었다. 결과에서 자세히 기술한 차원 별 차이와 유사 장르들은 글쓰기 교육에 함의점으로 제시하기 위해 활용할 수 있다. 본 연구에서 분석한 에세이를 작성한 한국 학생들은 아직 대학생이 아니라, 대학에 입학하기 위하여 작문 시험을 본 학생들이다. 따라서 이 학생들이 대학에 입학하여 수강할 영작문 수업의 교육과정에 바로 그 수업을 수강할 학생들이 취약한 부분을 반영할 수 있을 것이다.

본 연구에서 다차원 분석을 적용해서 얻은 결과는 연구 방법과 자료의 제한으로 인해 후속 연구를 통해 보완할 필요가 있다. 먼저, 연구방법의 측면에서 보자면, 다차원 분석을 위한 언어 요소의 빈도를 계산하는 것은 완전히 자동화되어있지는 않다. 따라서 본 연구에서 사용한 태거와 콩코던서의 검색식의 정확도에 의해 빈도 계산의 결과가 Biber가 참조 코퍼스에서 빈도를 계산한 방식과 달라지므로, 어느 정도 오차가 생기는 한계가 있다. 후속 연구에서는 연구에 사용한 계산 방법을 참조 코퍼스에 직접 적용하여 요인 분석을 행한 후 차원 별 비교를 하는 방식으로 이 한계를 극복할 수 있을 것이다.

연구 과정에서 보다 큰 어려움은 한국 학생들의 에세이와 비교 가능한 원어민 코퍼스 자료의 확보였다. 엄정한 비교를 위해서는 한국 학생들이 작문환경과 유사하게 변인을

통제된 원어민 자료가 필요한데, 본 연구가 진행될 시점에서는 그와 같은 원어민 자료가 가용하지 않아서 대안으로 이미 수집된 학습자 코퍼스의 자료를 사용하였다. 따라서 본 연구의 비교 결과는 변인의 통제가 이루어진 상황에서 원어민 자료를 직접 수집하는 연구를 통해 보완해야 한다. 아울러 본 연구는 학습자와 원어민 간의 차이를 기술하는 기초자료를 통계기법으로 생산하고, 다른 장르와 비교하는 것이 목적이었으므로 생산된 두 자료 간의 차이를 추가 분석을 통해서 통계적으로 비교하지는 않았다. 후속 연구에서는 결과를 통계적 추가분석을 통해 차이를 비교하여 그 차이가 통계적으로 유의미한지 확인할 필요가 있다. 특히 변인이 통제된 상황에서 수집된 원어민 자료와 비교한다면 더 정확한 결과를 얻을 수 있을 것이다.

마지막으로, 본 연구에서는 다차원 분석을 통한 분석 결과를 한국의 영어교육에 유용하게 적용할 방법론을 제시하지 않았다. 이후에는 다차원 분석의 교육적 함의를 밝힐 후속 연구가 필요하다. 예를 들어, 원어민 들의 근접 장르인 신문 사실과 보도 기사를 읽기자료로 분석하고, 한국 학생들의 즉흥적인 스타일과 근거에 기반한 원어민의 글쓰기 스타일과 명시적으로 비교하는 연습도 고려해볼 수 있다. 특히 학생들의 영작문 실력이 수업을 통해 점점 향상됨에 따라, 영작문 교육의 초점이 문법, 어휘의 오류를 피하는 것에서 점차 논쟁적 에세이라는 장르에 적절한 언어사용과 이를 통해 얻는 설득력이 더 중요해지게 될 것으로 생각된다. 그 시점에서는 더욱더 다차원 분석과 같이 장르와 스타일의 비교 분석 방법이 유용하게 활용될 수 있을 것이다.

REFERENCES

- Aguado-Jiménez, P., Pérez-Paredes, P., & Sánchez, P. (2012). Exploring the use of multidimensional analysis of learner language to promote register awareness. *System*, 40(1), 90-103.
- Al-Surmi, M. (2012). Authenticity and TV shows: A multidimensional analysis perspective. *TESOL Quarterly*, 46(4), 671-694.
- Anthony, L. (2007). *AntConc (Version 3.2.1w) [Computer Software]*. Tokyo: Japan.
- Asención-Delaney, Y., & Collentine, J. (2011). A multidimensional analysis of a written L2 Spanish corpus. *Applied Linguistics*, 32(3), 299-322.
- Atkinson, D. (1996). The philosophical transactions of the royal society of London, 1675-1975: A socio-historical discourse analysis. *Language in Society*, 25(3), 333-371. Retrieved from Cambridge Core.
- Biber, D. (1988). *Variation across speech and writing*. Cambridge: Cambridge University Press.
- Biber, D. (1995). *Dimensions of register variation: A cross-linguistic comparison*. Cambridge: Cambridge University Press.
- Biber, D. (2009). A corpus-driven approach to formulaic language in English: Multi-word patterns in speech and writing. *International Journal of Corpus Linguistics*, 14, 275-311.
- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.
- Biber, D., Conrad, S., Reppen, R., Byrd, P., & Helt, M. (2002). Speaking and writing in the university: A multidimensional comparison. *TESOL Quarterly*, 36(1), 9-48.
- Biber, D., Conrad, S., Reppen, R., Byrd, P., Helt, M., Clark, V., Cortes, V., Csomay, E., & Urzua, A. (2004). Representing language use in the university: Analysis of the TOEFL 2000 spoken and written academic language corpus. Retrieved from Educational Testing Service Princeton, NJ website: <https://www.ets.org/Media/Research/pdf/RM-04-03.pdf>
- Biber, D., Davies, M., Jones, J. K., & Tracy-Ventura, N. (2006). Spoken and written register variation in Spanish: A multi-dimensional analysis. *Corpora*, 1(1), 1-37.
- Biber, D., & Finegan, E. (1986). An initial typology of text-types. In J. Aarts & W. Meijs (Eds.), *Corpus linguistics II* (pp. 19-46). Amsterdam: Rodopi.
- Biber, D., & Finegan, E. (2001). Register variation in student and adult speech and writing. In S. Conrad & D. Biber (Eds.), *Variation in English: Multi-dimensional studies* (pp. 157-170). Harlow, England: Pearson.
- Conrad, S. (2001). Variation among disciplinary texts: A comparison of textbooks and journal articles in biology and history. In S. Conrad & D. Biber (Eds.), *Variation in English: Multi-dimensional studies* (pp. 94-107). Harlow, England: Pearson.
- Conrad, S., & Biber, D. (Eds.). (2001). *Variation in English: Multi-dimensional studies*. London: Longman.
- Csomay, E. (2007). Vocabulary-based discourse units in university class sessions. In D. Biber, U. Connor, & T. A. Upton (Eds.), *Discourse on the Move: Using corpus analysis to describe discourse structure* (pp. 213-238). Amsterdam/Philadelphia: John Benjamins Publishing.
- Granger, S. (1996). From CA to CIA and back: An integrated approach to computerized bilingual and learner corpora. In K. Aijmer & B. Altenberg (Eds.), *Languages in contrast* (pp. 37-51). Lund: Lund University Press.
- Granger, S. (2002). A bird's-eye view of learner corpus research. In S. Granger, J. Hung, & S. Petch-Tyson (Eds.), *Computer learner corpora, second language acquisition, and foreign language teaching* (pp. 3-33). Amsterdam: John Benjamins.
- Granger, S., Hung, J., & Petch-Tyson, S. (Eds.). (2002). *Computer learner corpora, second language acquisition, and foreign language teaching*. Amster-

- dam: John Benjamins.
- Granger, S., & Paquot, M. (2010). *ELexicography in the 21st century: New challenges, new applications*. Louvain-la-Neuve: Presses Universitaires de Louvain.
- Granger, S., & Rayson, P. (1998). Automatic profiling of learner texts. *Learner English on Computer*, 119-131.
- Granger, S., & Tyson, S. (1996). Connector usage in the English essay writing of native and non-native EFL speakers of English. *World Englishes*, 15(1), 17-27.
- Halliday, M. A. K. (1994). *An introduction to functional grammar* (2nd ed.). London: Edward Arnold.
- Hardy, J. A., & Römer, U. (2013). Revealing disciplinary variation in student writing: A multi-dimensional analysis of the Michigan Corpus of Upper-level Student Papers (MICUSP). *Corpora*, 8(2), 183-207.
- Hinkel, E. (2003). Simplicity without elegance: Features of sentences in L1 and L2 academic texts. *TESOL Quarterly*, 37(2), 275-301. <https://doi.org/10.2307/3588505>
- Hwang, Eunkyung. (2013). Syntactic complexity of Korean EFL learners' argumentative writing. *Modern English Education*, 14(2), 123-143.
- Martin, J. R. (1992). *English text: System and structure*. Amsterdam: John Benjamins.
- McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-based language studies: An advanced resource book*. London: Routledge.
- Park, Kwanghyun., & Kim, Eunju. (2016). A multi-dimensional analysis of Korean college English learners' speech transcript. *Studies in Foreign Language Education*, 30(3), 25-52.
- Schmid, H. (1994). Probabilistic part-of-speech tagging using decision trees. *Proceedings of international conference on new methods in language processing. UK*, 12, 44-49
- Yoon, Hyunsook. (2012). A multi-dimensional analysis of oral registers in revised middle school English textbooks. *Korean Journal of English Language and Linguistics*, 12(4), 777-799.