

1세대와 2세대 한국계 미국인 작가의 한국전쟁소설에 나타난 감정 어휘의 비교: R을 이용한 감성분석

김현욱

(사우스 캐롤라이나 주립대학교)

백승현*

(전남대학교)

Kim, Hyunwook, & Baek, Seunghyun. (2023). Comparison of emotional vocabulary usage in Korean War novels between first- and second-generation Korean American authors: A sentiment analysis using R. *Studies in Foreign Language Education*, 37(4), 67-82.

The purpose of this study was to compare the patterns of emotional vocabulary usage in Korean War novels by first- and second-generation Korean American authors and to explore the expected ratio of sentiment vocabulary usage by each generation. Based on sentiment analysis and Chi-squared test, the results showed that (1) in terms of frequency, both first- and second-generation use more positive sentiment vocabulary than negative one. More importantly, the first generation employs more positive sentiment vocabulary than the second one, and (2) with regards to the expected ratio of sentiment vocabulary usage across generations, the first-generation writers are 1.076 times more likely to use positive sentiment vocabulary than their counterparts. The current findings derived from a distant reading framework are not consistent with Hirsch (2012)'s post-memory research grounded in a close reading account. In keeping with Hirsch's expectation that the first-generation's traumatic memories would be passed on to the second-generation, both first- and second-generation should use more negative sentiment vocabulary than positive one. The possible reasons for the difference in the sentiment vocabulary usage by each generation are discussed and pedagogical implications are further addressed in detail.

* 제1저자: 김현욱, 교신저자: 백승현

I. 서론

기억 연구(memory study)의 선구자인 Hirsch(1997, 2001, 2012)는 홀로코스트(Holocaust) 생존자의 트라우마적 기억(trumatic memory)¹⁾이 시공간을 넘어 다음 세대에 전승될 수 있다고 주장하며 포스트 메모리(post-memory)라는 개념을 제시했다. 포스트 메모리는 “트라우마적 경험을 갖고 있는 생존자 자녀들의 문화적, 집단적 트라우마”(cultural or collective trauma to the experiences of their parents)를 일컫는데 생존자 부모의 구술, 편지, 또는 사진 등의 2차적 매개체에 의해 형성된다(Hirsch, 2001, p. 9). 그녀의 연구는 과거의 역사적 상처를 문화적 또는 문학적으로 복원할 수 있는 새로운 방향성을 제시했다는 점에서 의의가 있지만 생존자의 어떤 구체적인 감정이 다음 세대에 어떻게 영향을 미치는지에 대한 명확한 답을 제시하지는 않는다. 이는 신비평(new criticism) 이후의 지배적인 읽기 방식인 꼼꼼히 읽기(close reading)²⁾의 한계점으로 보인다. 즉, 주목하고 싶은 특정 부분에 대한 발췌(excerpt) 부분을 읽는 것과 한 개인이 인식할 수 있는 범위의 한계로 인해 전체적인 감정 패턴의 변화나 연결 관계를 파악하지 못한 결과로 보인다. 이러한 맥락의 일환으로, 꼼꼼히 읽기가 특정한 장르의 문학작품의 전체적인 내용전개와 구조를 파악하는데 한계점을 갖는다고 지적하며 멀리서 읽기(distant reading)³⁾의 방법으로 문학의 본질을 더 잘 이해할 수 있다고 주장한다(Moretti, 2000). 컴퓨터 프로그램을 이용하는 멀리서 읽기는 대규모 자연어(natural language)⁴⁾ 텍스트 데이터를 양적(quantitative)으로 분석하여 문학 작품의 패턴, 경향성, 문학작품들 간의 관계를 파악하는데 유리하다. 이러한 배경에서, 자연어에는 인간의 감정을 담는 어휘나 문장들이 많기 때문에 자연어에 담긴 주관적 감정 상태를 체계적으로 정량화하는 감성분석(sentiment analysis)⁵⁾은 자연어 처리(natural language processing)의 주요 관심 분야의 하나이다. 이러한 감성분석은 Hirsch (1997, 2001, 2012)의 연구가 해결하지 못 했던 트라우마적 기억을 가진 생존자와 다음 세대의 감성의 전이와 추이에 대한 방법론적 한계점을 극복하는 대안으로써 새로운 이론적 시각을 제시할 수 있을 것이다.

이러한 이론적·방법론적 배경에서, 본 연구의 목적은 자연어 처리 기술 중 하나인 감성분석을 통해 1세대와 2세대 한국계 미국작가들⁶⁾이 저술한 각 세대별 한국전쟁 소설 2권⁷⁾에 나타난 감정 및 감성 어휘

1) Hirsch (2001)의 연구에서 말하는 생존자의 트라우마적 기억은 홀로코스트와 같은 정신적, 신체적 또는 생명의 위협을 직면했던 상황에서 갖게 된 극도로 고통스러운 기억을 말한다. 이는 현재에도 끊임없이 생각나거나 후유증을 초래할 정도의 영향력을 가지는 기억을 말한다.

2) 꼼꼼히 읽기란 문학 작품을 내용 위주로 간단히 살펴보는 것이 아니라 단어 선택, 이미지리(imagery), 문장 구조, 상징, 미묘한 의미 등 세밀한 부분에 주목하여 작품을 깊이 있게 이해하는 방식을 말한다(Smith, 2016).

3) 멀리서 읽기란 꼼꼼히 읽기와 대비되는 방식으로, 개별 작품을 꼼꼼하게 읽는 것이 아니라 대규모의 텍스트 데이터를 컴퓨터 분석 기법을 통해 처리하고 문학작품들의 경향성이나 추세를 파악하는 읽기 방식이다(Smith, 2016).

4) 자연어(natural language)란 인간이 의사소통을 위해 사용하는 언어를 가리키며, 컴퓨터가 사용하는 0과 1의 이진코드로 이루어진 언어인 기계어(machine language)와 대비되는 개념이다(Eisenstein, 2019). 예를 들어, 영어 'Hello'는 자연어에 해당하고, '01001000(for 'H') 01100101(for 'e') 01101100(for 'l') 01101100(for 'o') 01101111(for 'o')'은 자연어 'Hello'를 의미하는 기계어이다.

5) 감성분석은 텍스트 데이터에 나타난 감정 관련 어휘를 구별하고 분석하는 자연어 처리 기술이다(Pozzi, 2017). 본 논문에서는 National Research Council of Canada (NRC) 감정 어휘사전(Mohammad, 2022)의 분류에 따라, 기쁨, 분노, 혐오 등을 나타내는 감정(emotion)과 긍정과 부정을 나타내는 감성(sentiment)으로 구분하여 기술한다.

6) 본 논문에서 1세대 한국계 미국인 작가는 한국에서 태어나고 자란 후 한국전쟁을 겪은 후 미국으로 이민한 작가를 말하고, 2세대 한국계 미국인 작가는 1세대 이민자의 자녀로 미국에서 태어났거나 미국 문화 속에서 자란 작가를 의미한다.

7) 1세대 한국계 미국작가 중 Kim(1964)의 *순교자(The Martyred)*와 Pak (1983)의 *죄의 대가(Guilt Payment)*를 분석에 사용하였고 2세대 한국계 미국작가 중에서는 Choi (1998)의 *외국인 학생(The Foreign Student)*과 Lee (2010)의 *향복한*

를 분석하여 비교하는 것이다. 앞서 언급한 감성분석은 특정 대상 또는 주제에 관한 감정 관련 어휘를 인식하고 분류하여 감정 어휘의 빈도수를 기반으로 기본적인 감정 및 감정 변화의 패턴을 탐구하게 한다. 그러나 단순히 빈도수를 통한 패턴 결과는 감성의 흐름만 제시하기에 세대별 작가가 사용한 어휘의 정확한 감정 비교가 어려우므로 후속 분석으로 카이스퀘어 검정(Chi-squared test)을 사용하여 세대별 작가와 감정 어휘의 사용 연관성과 세대별 작가 대비 감정 어휘의 사용 비율을 파악하고자 한다. 본 연구는 하나의 소설 또는 그 일부를 분석하는 꼼꼼히 읽기에 기반한 미시적 접근을 넘어서 멀리서 읽기에 근거한 거시적 접근과 추론적 통계(inferential statistics)분석을 병합하여 소설 내부의 감성의 흐름과 세대별 작가의 감정 어휘 사용의 관계에 대한 새로운 통찰력을 제공할 것이다. 나아가서 L2-영어 소설을 읽을 때, 학습자가 작가의 감정 어휘 사용 패턴을 파악하고 감성의 추이를 이해하여 궁극적으로는 소설작품의 서사 구조를 파악하는데 도움이 될 것이다.

II. 이론적 배경

기억 연구 분야에서 생존자의 트라우마적 기억이 문학적으로 어떻게 표상되는지 대해서 오랫동안 논의되어 왔다. 기억연구에서 홀로코스트 생존자와 그 자녀들의 트라우마에 대한 Hirsch (1997, 2001, 2012)의 포스트 메모리 연구가 선구적인데, 최근 그녀의 연구는 다양한 분야에서 적극적으로 차용되고 있다. 특히, 최근 다수의 아시아계 미국문학 연구자들이 Hirsch의 포스트 메모리 개념을 적용해 베트남전쟁과 한국전쟁 문학을 분석해왔다(Chu, 2011; Goodwin, 2015; Hong, 2014; Kim, 2009). 기존의 포스트 메모리 연구들은 문학 작품 연구의 ‘지배적인 방식’(primary methodology)인 꼼꼼히 읽기를 사용해왔다(Jockers, 2013). 하지만, 디지털 인문학(digital humanities) 연구자들은 꼼꼼히 읽기가 작고 미시적인 몇 가지 부분에만 집중해서 장리적 특징이나 반복되는 패턴과 같은 거시적인 윤곽을 파악하기 어렵다고 주장한다(Smith, 2016). 한편, Moretti (2000, p. 57)는 세계문학에 대한 논문에서 연구자가 무수히 많은 문학 작품을 모두 읽어내는 것은 불가능하므로 컴퓨터 소프트웨어를 활용해 문학 작품의 텍스트 데이터를 결합하고 분석함으로써 “장치, 주제, 비유—또는 장르 및 시스템”(devices, themes, tropes—or genres and systems)을 파악할 수 있다는 주장을 하였다. 나아가서 그는 그래프, 지도, 수형도 등을 통하여 멀리서 읽기가 “전체적인 상호 연결성에 대한 더 예리한 감각, 모양, 그리고 관계 및 구조”(a sharper sense of their overall interconnection. Shapes, relations, structures) 등을 생산해 내는 읽기를 할 수 있다는 것을 제안하였다(Moretti, 2005, p. 1).

하지만 멀리서 읽기를 활용한 문학 작품 분석은 아직은 생소한 개념이기 때문에 컴퓨터 과학 기법을 활용한 양적 연구 방식에 대해서 관심이 미흡하고 이러한 방법론의 연구가 활발하게 이루어지지 않고 있는 실정이다. 이러한 맥락에서, 멀리서 읽기 방법론에 대한 회의론적 시각 중에 하나는 멀리서 읽기는 기존의 꼼꼼히 읽기 연구로 인한 성과를 재확인하는 것에 지나지 않는다는 것이다. 일례로, 유희석(2005, pp. 152-153)은 Moretti (2000)의 멀리서 읽기를 “2차적 읽기”라 칭하며, 그의 주장이 그저 지역 문학 작품 연구를 종합하는 것에 불과하며, “인간과 그 역사현실을 단순히 기능적인 변수로 치환”하는 것에 불과하다는 비판적인 입장을 취한다. 하지만, 김용수(2019)는 한국 학계가 과학주의에 대한 부정적 인식으로 인해, 멀리서 읽기와 같은 실험적인 방법에 관심을 기울이기보다는 세계 문학론에 집중해 Moretti

사람들(The Surrendered)을 분석에 사용하였다.

(2000, 2005)의 연구의 한계와 맹점을 지적하는 비판적인 접근으로 일관하는 것은 생산적인 방향이 아니라고 지적한다. 오히려 컴퓨터 기법을 활용한 새로운 문학 작품 읽기 방식은 단순한 문학 작품 연구의 종합화와 시각화를 넘어서, 100년을 넘어서는 긴 시간 단위 속의 문학사적 흐름에 대한 통찰력을 제시하고 기존의 꼼꼼히 읽기와는 다른 결과를 제시할 수 있을 것이다.

이러한 새로운 읽기 방식의 접근에 기반하여 문학작품을 분석하려는 일련의 연구들은 감성분석을 사용하였는데(Gao et al., 2016; Mohammad, 2011; Ptaszynski et al., 2013), 이 분석은 전산 언어학(computational linguistics)⁸⁾기법을 사용하여 텍스트로부터 감정과 관련된 긍정, 부정, 중립⁹⁾의 어휘의 정보를 추출하는 자연어 처리 기법이다(Pozzi, 2017). 이러한 관점에서, Medhat et al. (2014)은 감성분석을 활용한 기존 연구에 대한 종합적인 개요를 제시하고 있는데 구체적으로 2011년부터 2014년까지 출판된 54개의 연구에 사용한 주된 데이터는 고객들의 제품리뷰이고 세부적으로는 제품 리뷰, 뉴스 기사, 온라인 기사, 블로그 포스팅 등으로 분류 되는데 이 중 오직 두 개의 연구만이 문학 텍스트를 데이터로 사용하였다.

하지만 이 두 개의 연구도 문학 작품을 직접적으로 분석한 연구라기보다는 문학 텍스트 데이터를 사용한 컴퓨터 기술(technique)의 개선에 관한 연구(Mohammad, 2011; Ptaszynski et al., 2013)에 가깝다. 구체적으로 살펴보면, Mohammad (2011, p. 106)의 연구는 ‘감정 단어 밀도’(emotion word density)라는 개념을 활용하여 더 나은 텍스트 검색 결과를 얻기 위해 문학 작품을 데이터베이스에 어떻게 분류해야 하는지를 제시하고 있다. 또한, Ptaszynski et al. (2013, p. 168)은 아동 문학 텍스트의 네러티브를 활용하여, 이야기 구조에서 어떤 인물과 어떤 감정이 연관되는지를 인식하는 ‘감정 인식’(recognition of emotions)기술을 발전시킨 것이다. 즉, 위에 언급한 사전 연구들은 감정 관련 어휘를 이해하기 위한 컴퓨터 모델의 성능 향상을 목표로 하기 때문에, 문학 작품 분석과 이해의 관점에서 출발한 이론적 배경과 시사점을 제공하고 있지는 않다.

문학 작품 분석에서 본격적으로 감성분석을 사용한 연구는 ‘감성 시계열’(sentiment time series) 개념을 사용한 것인데 이 개념은 시간에 따른 텍스트의 감성 어휘들을 분석하여 서사 구조의 발전에 따른 감성 변화의 흐름을 추적하기 위한 기법이다(Giachanou & Crestani, 2016). 이러한 감성 시계열 분석을 사용한 Gao et al. (2016)의 연구는 소설 작품의 감성 흐름의 파악을 통해 새로운 문학적 해석을 제시한 다기보다는 소설 14권에 관한 웹 리뷰 자료에 나타난 대중 의견의 장기적 의존성¹⁰⁾을 예측한 연구에 가깝다.

이러한 한계점을 극복하기 위하여, 본 연구는 1세대와 2세대 한국계 미국 작가가 쓴 한국전쟁 미국소설 각각 2편을 대상으로 감성분석을 시행하여 감성 변화 추이와 더불어 세대별 작가의 감성 어휘 사용 양상을 규명하려고 한다. 이를 위해, 다음과 같은 연구문제를 제시한다.

(1) 각 세대별 한국전쟁 소설에 나타난 감정 어휘의 빈도수 비교와 감성 변화 추이는 어떠한가?

8) 전산 언어학은 인간의 언어 이해하고 생성하도록 하는 컴퓨터 시스템을 연구하는 분야로 기계번역(machine translation), 음성인식(speech recognition) 등 다양한 분야에서 활용된다(Grishman, 1986).

9) NRC 감정 어휘사전은 긍정 단어의 예시로는 aspiring(열망 있는), comfort(위안), 부정 단어의 예시로는 abandoned(버려진), hypocrite(위선자), 중립 단어의 예시로는 administrative(행정상의), account(계좌) 등이 있다.

10) 장기적 의존성이란 시계열 데이터에서 과거의 값이 현재나 미래에 값에 영향을 주는 경향성을 의미하는 것으로 장기적 의존성을 갖는다는 것은 특정한 경향성이 앞으로도 지속되어 독자의 높은 관심도를 얻은 소설이라는 것을 의미한다.

(2) 한국전쟁 소설에 나타난 세대별 작가와 감정 어휘 사용의 연관성과 세대별 작가 대비 감정 어휘의 사용 비율은 어떠한가?

이 연구 문제를 해결함으로써 본 연구는 L2-영어 소설의 감정 변화 추이를 시각적으로 제시하여 갈등이나 내용전개에 대한 전체적인 서사 구조의 이해를 도울 수 있을 것이다. 나아가서, 감정 어휘 사용 비교를 통해 L2-영어 학습자들이 세대별 작가의 성향 및 특징을 파악하여 궁극적으로는 L2-영어 소설의 심리적 장벽을 낮추고 소설 구조에 대한 이해도를 증진시킬 수 있을 것이다.

III. 연구방법

1. 데이터 분석

한국전쟁 소설은 전쟁 중이거나 끝난 후부터 지금까지 총 100여권 정도가 출간되었다(정연선, 2019). 대부분의 소설은 참전군인들에 의해 저술되었고, 이 중 12권이 1세대와 2세대 한국계 미국 작가들에 의해 쓰여 졌다(Choi, 1991, 1993, 1994; Choi, 1998; Fenkl, 1996; Keller, 2002; Kim, 1964; Kim, 2018; Lee, 2004; Lee, 2010; Pak, 1983; Yoon, 2013). 본 연구의 목적이 한국전쟁의 직접 경험의 유무에 따른 1세대와 2세대 한국계 미국작가의 한국전쟁소설에 나타난 감정 변화추이와 세대별 작가와 감정 어휘 간의 관계 분석이기 때문에, 한국전쟁을 직접 경험한 1세대 작가들의 소설과 한국전쟁을 직접 경험하지 않았지만, 그 전쟁을 경험한 부모를 둔 2세대 작가들의 소설을 데이터 분석 자료로 정하였다. 이 기준에 따라 1세대 한국계 미국작가 중 Kim (1964)의 *순교자(The Martyred)*와 Pak (1983)의 *죄의 대가(Guilt Payment)*를 선택하였고, 2세대 한국계 미국작가 중에서는 Choi (1998)의 *외국인 학생(The Foreign Student)*과 Lee (2010)의 *항복한 사람들(The Surrendered)*을 분석에 사용하였다. 한편, Choi (1991, 1993, 1994) 역시 한국전쟁을 직접 경험한 1세대 작가에 속하지만, 그녀의 소설은 3부작으로 구성되어 있어 불균형한 데이터 양으로 인한 통계분석에 편향을 가져올 수 있으므로 분석에서 제외하였다.

소설 텍스트에 나타난 어휘의 감성분석을 위해 소설 텍스트 데이터 전 처리 과정을 두 단계로 수행하였다. 첫 번째 전 처리 단계에서, 소설의 하나의 챕터를 한 개의 분석 대상 단위로 설정하여 각 소설의 하나의 챕터를 기준으로 텍스트 파일을 생성하였다. 대상 소설 네 권의 모든 챕터의 단어 수 기준으로 박스 플롯(box plot)으로 분석을 하면, 각 챕터의 단어 수 기준, 하위 25퍼센트가 1,533단어이다. 하위 25퍼센트보다 작은 단어 수를 가진 챕터는 통계분석의 타당성을 확보하기 위하여, 단어 수가 1,500 미만인 챕터는 분석에서 제외하였다. 예를 들면, Kim (1964)의 *순교자*의 경우 41 챕터로 구성되어 있지만 하나의 완성된 장으로 보기에 짧은 챕터, 즉 단어 수가 1,500 미만인 챕터는 분석에서 제외하였다. 결과적으로, Kim (1964)의 *순교자*의 경우 21개의 챕터로 재구성하였고 Pak (1983)의 *죄의 대가*는 13개의 챕터로 구성되고 모든 챕터가 1,500 단어 이상이었기에 1세대 작가군의 소설은 총 34개의 챕터로 구성되었다. 한편, 2세대 한국계 미국작가의 소설의 경우 Choi (1998)의 *외국인 학생*이 14개의 챕터 그리고 Lee (2010)의 *항복한 사람들*이 19개의 챕터로 구성되어 있고 모든 챕터가 1,500 단어 이상이었기에 총 33개의 챕터로 2세대 작가군의 소설을 구성되었다.

두 번째 전 처리 단계에서는, 각 세대별 소설군의 전체 어휘 수 차이에 따른 통계분석의 오류를 제거하기 위해 단어 수 보정작업을 수행하였다. 1세대 작가군의 두 소설의 총 단어 수는 123,018단어이고, 2세대 작가군의 두 소설의 총 단어 수는 265,229단어이다. 이 차이는 통계적 분석 결과를 왜곡시키기에 충

분하다고 판단하여, 단어의 데이터 값을 다음과 같이 보정하였다. 1세대 한국계 미국작가와 2세대 한국계 미국작가 소설 총 단어 수의 비율이 2.16이므로, 1세대 한국계 미국작가의 데이터 값에 2.16을 곱하여 보정된 데이터 값을 산출하였다.

2. 감성분석

본 연구에서는 National Research Council of Canada (NRC) 감정 어휘사전¹¹⁾(Mohammad, 2022)을 기반으로 한 감성분석을 시행하여 1세대와 2세대 한국계 미국작가의 한국전쟁소설 4권의 감정 및 감성 어휘를 분석하였다. NRC 감정 어휘사전은 감정 어휘의 집합인데 이는 기대(anticipation), 기쁨(joy), 놀람(surprise), 신뢰(trust), 분노(anger), 혐오(disgust), 두려움(fear), 슬픔(sadness) 8개의 감정(emotion)으로 단어를 분류¹²⁾하고, 다시 기대, 기쁨, 놀람, 신뢰를 긍정(positive)으로 분노, 혐오, 두려움, 슬픔은 부정(negative)의 감성(sentiment)으로 분류한다. 이렇게 분류된 NRC 감정 어휘사전의 단어와 소설에 출현한 단어를 비교하기 위해 tidytext R 패키지(2023)를 사용하였고 이 과정에서 텍스트를 토큰(token)으로 분류하여 비정형 데이터(unstructured)를 표의 형태로 구조화하는 텍스트 전처리 과정을 거쳤다.

전처리 과정을 거친 후, R 프로그램을 사용한 감성분석은 기본적으로 대상 데이터를 NRC 감정 어휘사전 속 어휘와의 비교를 통하여 각 해당 감정과 연관이 없으면(not associated) 0으로, 연관이 있으면(associated) 1로 코딩하여 특정 감정 어휘군의 빈도를 계산한다. 이러한 분석 과정에서 하나의 단어는 특정 감정 어휘군에만 속할 수도 있지만 여러 개의 감정 어휘군에 속할 수도 있다. 예를 들어, NRC 감정 어휘사전에 ‘abandoned’라는 단어는 감성 연관 어휘(Word-Sentiment Associations)에서는 부정으로 분류되고, 감정 연관 어휘(Word-Emotion Associations)에는 분노, 두려움, 슬픔에 세 가지 감정에 각각 해당된다. 즉, 감성분석 기술은 감성 사전에서 ‘abandoned’이라는 단어가 존재하기 때문에 소설 속에 ‘abandoned’이라는 단어가 출현한다면 분노, 두려움, 슬픔 감정 어휘군과 부정 감성에 각각 1의 값을 부여한다. 이러한 분석 절차를 통하여 최종적으로 감성분석의 결과 값은 전체 긍정 감성 단어의 빈도수에서 부정 감성 단어의 빈도수를 뺀 수치로 양(+)의 값이면 긍정의 감성이고, 음(-)의 값이면 부정의 감성을 나타내는 것이다.

3. 카이스퀘어 분석

두 범주형 변수들 특히, 명목척도(nominal scale) 사이에 연관성(association)이 있는지 또는 독립적

11) 감성 분석을 위해 다양한 감정어휘사전이 사용되며 Bing, AFINN, NRC 감정어휘사전이 대표적이다(Silge & Robinson, 2017). Bing 감정어휘사전은 단어를 긍정과 부정 감성의 범주로 이진 분류하고 AFINN 감정어휘사전의 경우에는 감정단어에 -5에서 5까지의 점수를 할당해 양(+)의 값은 긍정 감성을 나타내고 음(-)의 값은 부정 감성을 나타낸다. 반면, NRC 감정어휘사전은 긍정과 부정 감성 이외에도 기대, 기쁨, 놀람, 신뢰 그리고 분노, 혐오, 두려움, 슬픔과 같은 다양한 감정을 기준으로 어휘를 분류한다. 본 연구에서는 세대별 한국전쟁 소설에 나타난 감정 어휘의 빈도수 비교와 세대별 작가의 다양한 감정 어휘 사용 추이를 파악하기 위해 NRC 감정사전을 사용하였다.

12) 각 감정 어휘군의 예시는 다음과 같다. 기대: aspiring(열망하는), 기쁨: delightful(정말 기분 좋은), 놀람: wonderful(아주 멋진), 신뢰: admirable(존경스러운), 분노: annoying(짜증스러운), 혐오: awful(끔찍한), 두려움: anxious(불안해하는), 슬픔: depressed(우울한).

인지를 파악하는 카이스퀘어 분석은 관찰빈도(observed counts)와 기대빈도(expected counts) 사이의 차이를 계산하고 검정 통계량을 산출하여 두 변수의 연관성의 유무를 판단하는 통계 분석이다(Baek, 2019). 본 연구에서는 세대별 작가와 감정 어휘 유형이 명목척도로서, 2(세대별 작가: 1세대 vs. 2세대) x 2(감정 어휘 유형: 긍정 vs. 부정) 카이스퀘어 분석을 하여 세대별 작가와 감정 어휘 유형의 연관성과 세대별 작가 대비 감정 어휘의 사용 비율을 규명할 것이다. 즉, 세대별 작가의 긍정 또는 부정 감정 어휘 사용에 통계적으로 유의미한 차이가 있는지의 여부를 파악하고, 만약 차이가 있다면 세대별 작가 대비 감정 어휘의 사용 비율의 기대치를 밝혀내고 감정 어휘 대비 어느 세대의 작가일 가능성을 제시할 수 있을 것이다.

IV. 연구결과

1. 세대별 감정 어휘의 빈도수와 감성 변화의 추이

첫 번째 연구 문제인 ‘각 세대별 한국전쟁 소설에 나타난 감정 어휘의 빈도수 비교와 감성 변화 추이’에 대한 결과를 살펴보고자 한다. 먼저 아래의 표 1은 보정 값을 반영하여 1세대, 2세대 작가군의 감정과 감성의 빈도를 비교한 것이다.

표 1. 1세대와 2세대 작가군의 감정과 감성의 빈도

	기대	기쁨	놀람	신뢰	긍정	감성 수치 (긍정 - 부정)
1 세대	5612	4689	3078	9139	12465	1750
	분노	혐오	두려움	슬픔	부정	
	4409	3352	6307	5381	10716	
2 세대	기대	기쁨	놀람	신뢰	긍정	717
	4900	4092	2632	5675	9614	
	분노	혐오	두려움	슬픔	부정	
	3238	2560	4246	4459	8897	

Note. 긍정=기대+기쁨+놀람+신뢰; 부정=분노+혐오+두려움+슬픔; 감성 수치는 긍정 빈도에서 부정 빈도를 차감한 값. 양(+)의 값이면 긍정, 음(-)의 값이면 부정의 감성.

전반적으로 감성 유형에 상관없이 1세대가 2세대보다 감정 어휘 사용 빈도가 높다. 또한 긍정과 부정 감정 어휘의 차이는 세대와 상관없이(1세대: 1750; 2세대: 717) 양의 값(+) 즉, 긍정적 감성이 부정적 감정보다 우위를 점하고 있으며 1세대 감성 수치가 2세대 감성 수치보다 더 큰 것을 알 수 있다. 결론적으로, 빈도 수 측면에서 1세대가 2세대보다 긍정적인 어휘를 더 많이 사용한다는 것을 알 수 있다. 구체적으로는 세대별 비교에서 긍정 감성 유형에 속하는 신뢰 연관 어휘와 부정 감성 유형에서는 두려움 연관 어휘에서 빈도수의 차이가 가장 크게 나타났고, 세대간의 신뢰의 차이(3464)가 두려움의 차이(2061)보다 더 크게 나타났다. 이 결과를 통해 앞서 언급한 1세대가 2세대보다 긍정적 어휘를 더 많이 사용한다는 것을 재확인할 수 있다.

그림 1은 세대별 작가의 각 챕터에서 출현한 긍정 감성 어휘(기대, 기쁨, 놀람, 신뢰)와 부정 감성 어휘(분노, 혐오, 두려움, 슬픔) 각각의 최댓값, 최솟값, 제 1사분위수(25%), 중앙값(median: 50%), 그리고

제 3사분위수(75%)를 나타낸 박스플롯이다. 1세대 작가군의 긍정 감성 어휘에 대한 사분범위(interquartile range: 75% - 25%)는 247이고 2세대 작가군의 긍정 감성 어휘에 대한 사분범위는 108로, 1세대 작가군의 긍정 감성 어휘가 2세대에 비해 중앙값을 기준으로 더 넓게 분포되어 있음을 알 수 있다. 마찬가지로, 1세대 작가군의 부정 감성 어휘에 대한 사분범위는 253.25이고 2세대 작가군의 부정 감성 어휘에 대한 사분범위는 79로, 1세대 작가군의 부정 감성 어휘도 2세대에 비해 더 넓게 분포되어 있다. 즉, 이 결과는 1세대가 2세대보다 긍정이든 부정이든 감성 유형에 상관없이 더 많은 감성 어휘를 사용하고 있다는 것을 의미한다. 한편, 각 세대 내에서의 긍정과 부정 감성을 비교해보면 1세대 긍정, 부정 감성의 평균은 각각 366.56과 315.09이고 2세대의 경우에는 291.33과 269.61으로 관찰되어 세대에 상관없이 감성 유형에서 긍정 감성 어휘를 부정 감성 어휘보다 더 많이 사용하였다. 중앙값을 통해 중심 경향성을 살펴보면 1세대 긍정, 부정 감성의 중앙값이 각각 273과 196.5이고 2세대 긍정, 부정 감성의 중앙값이 각각 264와 252로 세대에 상관없이 긍정 감성 어휘의 중앙값이 높는데, 이는 평균 비교의 결과와 마찬가지로 세대에 상관없이 긍정 감성 어휘에서 더 높은 중심 경향성을 보인다는 것을 의미한다.

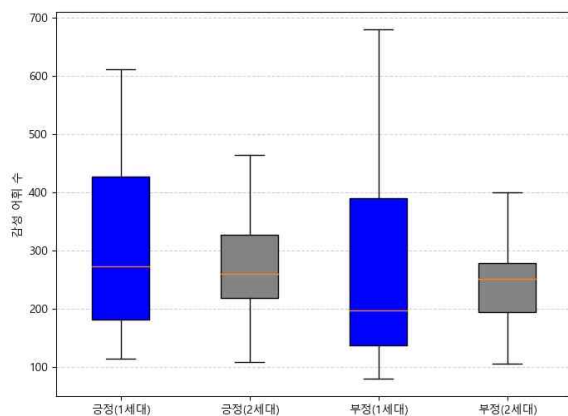


그림 1. 세대 별 긍정, 부정 감성 어휘 사용 수 비교

다음으로는 각 소설의 감성 추이를 세대별로 살펴보면서 비교 검토한 결과이다. 그림 2의 챕터 1부터 21까지는 1세대 작가 Kim (1964)의 *순교자* 21개 챕터의 감성 수치(긍정 - 부정)를 그래프로 나타낸 것이고 챕터 22부터 34까지는 1세대 두 번째 작가인 Pak (1983)의 *죄의 대가* 13개 챕터의 감성 수치를 그래프로 나타낸 것이다. Kim (1964)의 *순교자*의 경우, 챕터 1과 20을 제외한 19개 챕터들은 대체적으로 양(+)의 값을 보이며 최댓값은 186, 최솟값은 -35로 나타나기에 전반적으로 긍정적인 감성을 강하게 표출하고 있다는 것을 알 수 있다. 한편, 1세대 두 번째 작가인 Pak (1983)의 *죄의 대가*의 경우에는 총 13개 중 절반 정도인 6개 챕터가 부정적인 감성을 나타내며 최댓값은 272이고, 최솟값은 -147이다. 이 작품의 경우, 후반부 챕터로 갈수록 긍정적인 경향성을 보이는 것을 알 수 있다.

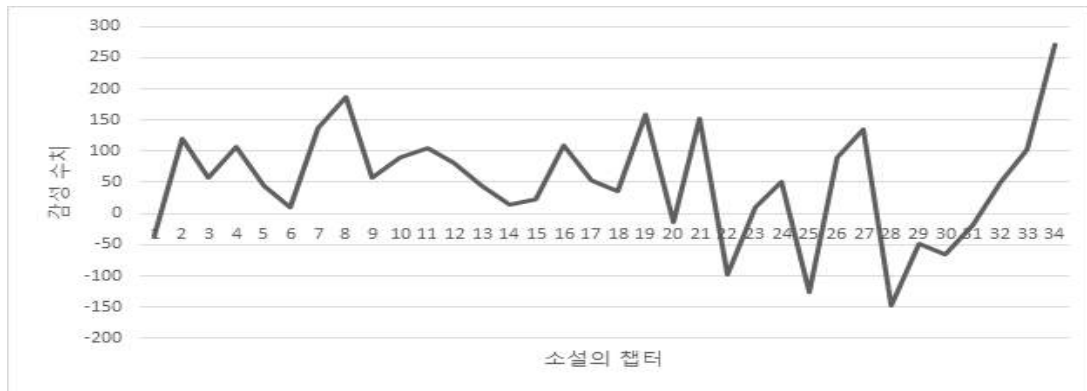


그림 2. 1세대 소설 감성 추이: Kim (1964)의 *순교자*; Pak (1983)의 *죄의 대가*

종합적으로 살펴보면, 1세대 작가의 경우 34개의 챕터 중 8개의 챕터(1, 20, 22, 25, 28, 29, 30, 31)를 제외한 26개의 챕터들은 양(+)의 감성 값을 보이고 있다. 즉, 그림 2에서 전반적으로 긍정적인 감성이 지배적인 것을 알 수 있다. 이 결과는 표 1에서 제시된 1세대 작가군의 감성 수치가 1,750으로 긍정적 감성 관련 어휘를 우세하게 사용하고 있는 것을 재확인할 수 있다. 1세대에서 주목할 만한 특징은 전반적인 감성이 안정적으로 긍정적이며 긍정과 부정 감성의 변동 주기가 적은 편이라는 것이다.

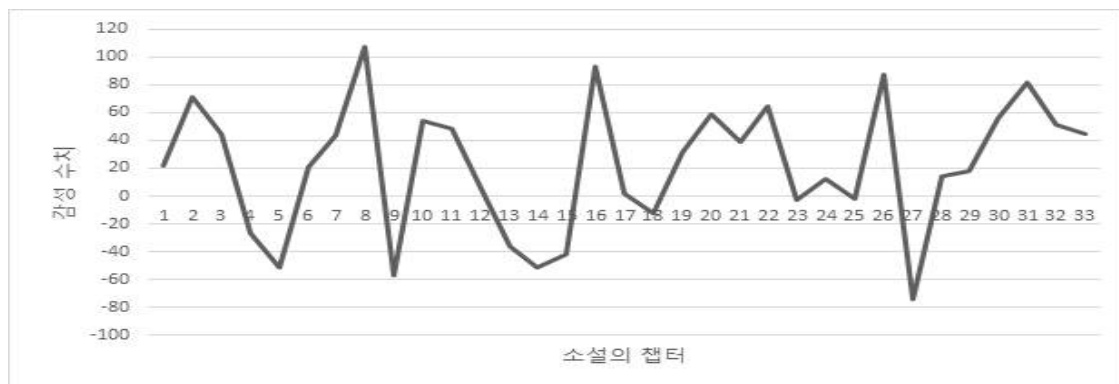


그림 3. 2세대 소설 챕터별 감성 추이: Choi (1998)의 *외국인 학생*; Lee (2010)의 *행복한 사람들*

1세대 작가의 그래프와 마찬가지로 그림 3의 챕터 1부터 14까지는 2세대 작가 Choi (1998)의 *외국인 학생* 14개 챕터의 감성 수치(긍정 - 부정)를 그래프로 나타낸 것이고 챕터 15부터 33까지는 2세대 두 번째 작가 Lee (2010)의 *행복한 사람들* 19개 챕터의 감성 수치를 그래프로 나타낸 것이다. Choi (1998)의 *외국인 학생*의 경우에는 1세대 작가의 작품처럼 전반적으로 양(+)의 값을 보이며 최댓값이 107, 최솟값이 -57이다. 또한, 2세대 두 번째 작가 Lee (2010)의 *행복한 사람들*의 경우 역시 전반적으로는 긍정적인 감성이 우세한 것으로 보이며 최댓값이 93이고, 최솟값이 -74로 나타났다.

종합적으로 살펴보면, 2세대 작가의 경우 33개의 챕터 중 10개의 챕터(4, 5, 9, 13, 14, 15, 18, 23, 25, 27)를 제외한 23개의 챕터들은 전반적으로 양(+)의 감성 값을 보이며 2세대 작가군 역시 긍정 감성 어휘가 우위를 점하고 있다. 이 결과는 표 1에서 제시된 2세대 작가군의 감성 수치가 717으로 긍정적 감성 관련 어휘를 지배적으로 사용하고 있다는 것을 다시 한번 확인할 수 있다. 2세대에서 특이한 점은 긍정과 부정 감성의 변동 주기가 1세대에 비해 더욱 빈번하게 나타나고 있다는 것이다.

그림 2와 3에서 나타난 두 작가 세대를 비교해볼 때 1세대의 순교자의 최댓값은 186, 최솟값은 -35, 좌의 대가의 최댓값이 272, 최솟값이 -147인 반면 2세대의 외국인 학생의 최댓값이 107, 최솟값이 -57, 항복한 사람들의 최댓값이 93이고, 최솟값이 -74로, 감성 차이의 폭은 1세대 작가군이 2세대 작가군에 비해 더 큰 것을 알 수 있다. 비록 두 세대 모두 양(+)의 값을 갖는 긍정 감성의 챗터들이 우세하지만 1세대 그래프에 나타난 양(+)의 값, 즉 긍정 감성이 2세대 그래프에 나타난 양(+)의 값보다 더 폭이 크고 더 많은 챗터에서 관찰되었다. 따라서, 1세대 작가가 2세대 작가에 비해 긍정 감성 어휘를 더 많이 사용한다는 것을 알 수 있고 이러한 결과는 표1에 나타난 1세대, 2세대 작가군의 감성 수치가 각각 1750, 717인 것을 재확인 시켜주고 있다.

2. 세대별 작가와 감성 어휘 유형의 연관성

다음은 두 번째 연구 문제인 ‘한국전쟁 소설에 나타난 세대별 작가와 감성 어휘 유형의 연관성과 세대별 작가 대비 감성 어휘의 사용 비율’을 살펴보고자 한다.

표 2. 감성*그룹 교차표(cross-tabulation)

		그룹		총합	
		1 세대	2 세대		
감성	긍정	관찰빈도	12,465	9,614	22,079
		기대빈도	12,276.1	9,802.9	22,079
		감성 내	56.5%	43.5%	100.0%
		그룹 내	53.8%	51.9%	53.0%
	부정	관찰빈도	10,716	8,897	19,613
		기대빈도	10,904.9	8,708.1	19,613
		감성 내	54.6%	45.4%	100.0%
		그룹 내	46.2%	48.1%	47.0%
총합	관찰빈도	23,181	18,511	41,692	
	기대빈도	23,181	18,511	41,692	
	감성 내	55.6%	44.4%	100.0%	
	그룹 내	100.0%	100.0%	100.0%	

위의 표 2에 나타난 바와 같이 세대별 그룹 내의 감성 어휘 비율은 1세대의 경우 긍정 감성이 53.8%(12,465), 부정 감성이 46.2%(10,716)이고, 2세대의 경우 긍정이 51.9%(9,614), 부정이 48.1%(8,897)이다. 감성 유형 내의 그룹 비율은 긍정의 경우 1세대는 56.5%, 2세대는 43.5%이고, 부정의 경우 1세대는 54.6%, 2세대는 45.4%로 나타나고 있다. 표 2에 나타난 관찰빈도와 기대빈도의 차이에 관한 통계적 유의성, 즉 세대별 작가와 감성 유형과의 연관성을 확인하기 위하여 카이스퀘어 분석을 한 결과는 표 3에 제시되었다.

표 3. 카이제곱 검정

	값	자유도	Asymp. Sig. (양측 검정)	Exact Sig. (양측 검정)	Exact Sig. (단측 검정)
Pearson 카이제곱	13.923*	1	.000		
연속성 보정**	13.850	1	.000		
Fisher의 정확도 검정				.000	.000
유효 사례 수	41692				

*0퍼센트 셀(0.0%)은 기대빈도가 5 이하이다. 최소 기대빈도는 8708.06이다.

**2x2 표로 계산되었다.

위의 표 3에서 제시된 바와 같이 기대빈도 5 이하의 셀이 0%이므로 Pearson 카이제곱 값($\chi^2 = 13.923$, $df = 1$)을 확인하여 통계적으로 유의미하였으므로($p = .000$), 세대별 작가의 감성 어휘 유형 간의 연관성이 있음을 확인할 수 있다. 또한 이 연관성의 효과의 크기(effect size)를 나타내는 Cramer의 V 값은 0.018($p = .000$)로 나타났다. 이 결과는 세대별 작가의 긍정 감성 어휘와 부정 감성 어휘 사용에 빈도의 차이가 있다는 것을 의미한다. 다음으로 연관성의 강도 즉, 감성 어휘의 사용 비율의 기대에 측력을 알아보기 위하여 표 4의 위험추정(risk estimate) 값을 확인해 보도록 하자.

표 4에서 볼 수 있듯이 95% 신뢰구간에 1이 포함되어 있지 않기 때문에 감성에 대한 승산비(odd ratio)¹³⁾, 1.076이 통계적으로 유의미하다는 것을 알 수 있다. 다시 말하면, 이 결과는 세대별 작가 대비 감성 어휘의 사용 비율의 기대치를 제시하는 것으로 1세대 한국계 미국 작가군은 2세대 작가군에 비해서 긍정 감성 어휘를 사용할 가능성이 1.076배 높다는 것을 의미한다.

표 4. 위험 추정(risk estimate)

	값	95% 신뢰구간	
		하위	상위
감성에 대한 승산비(긍정/부정)	1.076	1.036	1.119
코호트 집단 = 1 세대	1.033	1.016	1.051
코호트 집단 = 2 세대	.960	.939	.981
유효 사례 수	41692		

한편, 감성 어휘 비율 대비 어느 세대의 작가군에 속할지에 대한 가능성을 제시하기 위한 상대 위험도(relative risk) 분석 결과에 따르면 긍정 감성 어휘를 사용하면 부정 감성 어휘 대비 1세대가 될 확률이 1.033배가 높고 다시 말하면, 2세대가 될 가능성이 0.96배 낮다는 것을 의미한다. 결론적으로 세대별 감성 어휘 사용에 관한 연관성은 1세대일수록 2세대보다 긍정 감성 어휘를 더 많이 사용한다는 것을 카이스퀘어 분석을 통해 재확인할 수 있었다.

V. 결론과 논의

13) 승산비는 기대 예측력을 의미하고 1보다 클 경우 긍정적인 기대치(+), 1보다 작을 경우 부정적인 기대치(-)를 의미한다.

본 연구의 목적은 세대별 감정 어휘의 빈도수 비교와 감정 변화의 추이 그리고 세대별 작가와 감정 어휘 유형의 연관성을 파악하는 것이었다. 첫 번째로, 감정 어휘의 빈도수에서 나타난 감정 수치가 1세대는 1750이고 2세대는 717로 나타났다. 또한 R의 감성분석에 기반하여 감정 추이에 대한 결과로 두 작가군 모두 양(+)의 값, 즉 긍정의 감성을 보여주지만 1세대의 챗터별 최댓값, 최솟값이 각각 272, -147이고 2세대의 경우, 각각 107, -74로 감정 차이의 폭은 1세대가 더 크고 양(+)의 값을 갖는 챗터들이 2세대에 비해 1세대에서 더 많이 관찰된다는 것을 시각적으로 확인할 수 있었다. 한편, 감정 추이 그래프를 살펴보면 1세대는 긍정과 부정 감성의 변동 주기가 적은 편이고 긍정 감성이 안정적으로 나타났지만 이에 반해 2세대의 경우 변동 주기가 훨씬 더 빈번하여 긍정 감성이 1세대에 비해 상대적으로 불안정하다는 것을 알 수 있었다.

두 번째로 카이제곱 분석을 통하여 세대별 작가의 감정 어휘 유형 간의 연관성($X^2 = 13.923$, $df = 1$, $p = .000$)이 있음을 확인할 수 있었고 이 결과는 세대별 작가에 따라 감정 어휘 유형(긍정 vs. 부정)의 사용 빈도에 차이가 있다는 것을 의미하였다. 나아가서, 위험추정 값을 통해 그 연관성의 강도를 확인하든 1세대 작가군이 2세대 작가군에 비해 긍정 감정 어휘를 사용할 가능성이 1.076배 더 높고 긍정 감정 어휘를 사용하면 부정 감정 어휘 대비 1세대 작가군이 될 가능성이 1.033배 높다는 것을 확인할 수 있었다.

결론적으로 본 연구는 단순한 데이터의 종합화와 시각화를 넘어서서 세대별 작가와 감정 어휘 사용의 관련성에 대한 통계적 데이터 분석을 통하여 세대별 작가군의 감정 어휘 유형 간의 빈도의 차이가 있고 1세대가 2세대에 비해 긍정적 어휘를 더 많이 사용하는 경향과 기대 가능성이 있다는 결과를 도출하였다. 이러한 세대별 작가의 감정 어휘 사용 양상에 관한 연구 결과는 Hirsch(1997, 2001, 2012)가 생존자의 트라우마적 기억이 자녀 세대로 전이된다는 포스트 메모리의 주장과는 상반된다. 구체적으로 살펴보면, 꼼꼼히 읽기에 기반한 Hirsch의 포스트 메모리 주장에 의하면 1세대의 작품은 전쟁에 의한 트라우마적 기억으로 인해 부정 감정 어휘가 지배적이어야 하고 이에 영향을 받은 2세대의 작품 역시 비슷하게 부정 감정 어휘가 우세해야 한다. 하지만, Hirsch의 주장과는 다르게 본 연구에서는 1세대와 2세대 작가군 모두가 부정보다는 긍정 감정 어휘의 사용이 더 많았고 1세대가 2세대에 비해 긍정 감정 어휘를 상대적으로 더 많이 사용하였다. Hirsch의 연구가 본 연구의 통계 분석 결과와 상반된 결과를 가져온 이유를 꼼꼼히 읽기 방식과 일정 부분의 발췌독에서 비롯된 방법론적 문제에서 찾아볼 수 있다. 즉, 꼼꼼히 읽기 방식으로 다소 지엽적인 부분에 집중하면서 전체적인 감정의 흐름이나 패턴을 파악하는 것을 간과한 것으로 추정할 수 있다.

반면, 본 연구는 멀리서 읽기 방식을 기반으로 한 감성분석의 방법론으로 작품 전체 챗터의 종합적 검토를 통해 세대별 작가군의 감정 어휘의 빈도와 감정 추이에 관한 경향을 총체적으로 파악할 수 있었다. 나아가서, 세대 간의 좀 더 면밀한 비교와 분석을 하기 위하여 감성분석의 토대 위에 통계적 분석과 해석을 추가하여 각 세대별 작가의 감정 어휘 사용간의 연관성과 세대별 작가 대비 감정 어휘의 사용 비율에 대한 통찰력을 제공할 수 있었다. 본 연구 결과의 잠재적 원인을 이해하기 위해서는 세대별 작가의 성장 배경과 미국 사회와 문화에 반응하는 방식을 살펴보는 것이 필요할 것이다.

1세대 한국계 미국 작가들은 한국전쟁을 경험하여 트라우마적 기억을 안고 낮은 환경으로 이민하여 사회, 경제, 문화적 적응 과정에서 어려움과 고통을 겪었을 것이라는 것은 부인할 수 없는 사실이다. 하지만, 그들에게 가장 중요한 것은 새로운 나라 미국에서의 생존이었을 것이고 생과사를 오가는 전쟁 경험과 비교하면 미국 정착에서 오는 고난은 그렇게 심각한 문제는 아니었을 것이다(Kim, 2008). 오히려, 그

들은 미국이라는 새로운 터전을 마지막 기회로 보았을 가능성이 있다. 이 결과는 각 작가의 문화적 성장 배경이 글쓰기에 영향을 미칠 수 있음을 보여준다. 1세대 한국계 미국 작가의 경우에는 2세대 작가보다 한국의 집단주의적 성장 배경을 더 가지고 집단주의적 정체성을 형성 후에 미국으로 이민하였을 것이다. 또한 그들은 새로운 터전인 미국에서의 생존을 위해서 한인 공동체의 구성원에 대한 신뢰를 바탕으로 한 협력이 필요했을 것이다. 이러한 세계관이 1세대 작가의 글에서 더 긍정적인 단어를 사용하고 더 희망적인 주제를 표현하도록 이끌었을 수 있다.

이에 반해, 2세대 한국계 미국작가들은 미국에서 유년기와 학창시절을 보내면서 소수인종이기 때문에 직면해야 했던 차별과 상처에 민감하였을 것이고 차별적인 환경에 적응하는데 고군분투했을 것이다(Kim, 2008). 그 민감성은 2세대 작가의 글에서 미묘하게 비판적인 관점을 가질 수 있도록 하는 원동력이 되어서 2세대가 1세대에 비해 긍정적인 어휘를 더 적게 사용하는 양상을 보인 것이라고 해석할 수 있다. 더구나 2세대 한국계 미국 작가들은 보다 개인주의적인 문화적 배경에서 성장하여 개인의 자립에 보다 초점을 맞추며 타인에 대한 신뢰라는 가치에는 상대적으로 작은 관심을 갖게 됐을 것이다. 이러한 서로 다른 성장배경에 기인하여 1세대가 2세대가 좀 더 긍정적인 어휘를 사용한 것으로 보인다.

한편, L2-영어 교육적인 측면에서 감정 어휘사전의 분류 유형(긍정: 기대, 기쁨, 놀람, 신뢰, 부정: 분노, 혐오, 두려움, 슬픔)은 작품의 구조와 감성의 흐름을 파악하는데 활용될 수 있을 것이다. 예를 들어, L2-영어 소설 읽기 수업에서 감정 어휘사전에 제시된 긍정 감성의 기대, 기쁨, 놀람, 신뢰와 부정 감성의 분노, 혐오, 두려움, 슬픔의 총 8개 감정 카테고리라 그 카테고리에 해당되는 어휘를 학습자에게 제공한 후, 수업을 위해 발췌한 소설의 하나의 챕터에서 감정 유형과 그 유형에 해당하는 어휘를 연결시키게 하고 그 의미를 학습 시킨다. 학습된 각 감정 어휘를 양(+의 긍정 감성과 음(-)의 부정 감성으로 분류하게 하고 긍정과 부정 감성의 빈도수를 파악하게 하여 목표 챕터의 최종 감성을 결정하게 한다. 동일한 방법으로 각 챕터의 감성을 결정한 후 소설 전체의 감성 변화 추이선을 그려보게 하는 활동을 할 수 있을 것이다.

이러한 일련의 과정들이 L2-영어 소설을 읽는 학습자들에게 감성의 빈도수와 같은 수치와 감성 추이선과 같은 시각적 자료를 목표 작품의 문맥을 이해하는데 도움을 줄 수 있는 정보로 제공함으로써 목표 소설에 대한 ‘이해 증진 보조’(Comprehension-Aiding Supplementation)¹⁴⁾의 역할을 할 수 있을 것이다(Krashen et al., 2018, p. 1). 또한 이러한 학습 활동은 소설의 전반적인 감성의 흐름을 이해하고 구체적인 감정을 파악하여 L2-영어 소설을 읽는 심리적 장벽을 낮출 수 있을 것이다. 나아가서 이러한 학습 과정을 통하여 L2-영어 학습자는 소설 전반의 감성 변화 추이에 대한 민감도를 높이고 궁극적으로는 소설의 등장인물로 확장하여 감정 어휘를 기반으로 한 등장인물의 성격파악도 가능할 수 있을 것이다.

본 연구에서 선택한 세대별 작가 2명의 소설 총 4권에 나타난 감정 어휘 사용 양상만으로 각 세대 작가의 특징을 규정하는 일반화의 어려움을 극복하기 위해 다수의 세대별 작가를 확보하고 소설뿐만 아니라 시, 드라마 등 다양한 장르의 작품을 분석하는 후속 연구가 필요해 보인다.

참고문헌

14) 최근의 연구들(윤택남 외, 2019; 윤택남, 김병선, 2021)에서 이해 증진 보조의 일종인 워드클라우드 활용 영어 수업이 L2-영어 학습자들의 어휘력 향상에 도움이 되며 자신감과 학습 동기를 향상 시킨다는 결과를 도출하였다.

- 김용수. (2019). 세계문학과 디지털인문학 방법론: 한국 학계의 모레티 연구. *비평과 이론*, 24(3), 59-78.
- 유희석. (2005). 세계문학에 관한 단상: 프랑코 모레티의 발상을 중심으로. *안과밖*, 18, 134-153.
- 윤택남, 석경오, 최가영. (2019). 워드클라우드 활동이 초등영어학습자들의 어휘능력과 정의적 태도 변화에 미치는 영향. *외국어교육연구*, 33(4), 429-447.
- 윤택남, 김병선. (2021). 워드클라우드 활용이 예비초등영어교사들의 교과 내용 지식과 정의적 영역에 미치는 효과. *외국어교육연구*, 35(3), 79-94.
- 정연선. (2019). *잊혀진 전쟁의 기억: 미국소설로 읽는 한국전쟁*. 문예출판사.
- Baek, S. (2019). *A guide to running SPSS package for second language acquisition and applied linguistics research*. Chonnam National University Press.
- Choi, S. (1991). *Years of impossible goodbyes*. Houghton Mifflin.
- Choi, S. (1993). *Echoes of the white giraffe*. Houghton Mifflin.
- Choi, S. (1994). *Gathering pearls*. Houghton Mifflin.
- Choi, S. (1998). *The foreign student*. Harper Collins Publisher.
- Chu, S. (2011). *Do metaphors dream of literal sleep?: A science-fictional theory of representation*. Harvard University Press.
- Eisenstein, J. (2019). *Introduction to natural language processing* (Illustrated ed.). The MIT Press.
- Fenkl, H. I. (1996). *Memories of my ghost brother*. Dutton.
- Gao, J., Jockers, M. L., Laudun, J., & Tangherlini, T. (2016). A multiscale theory for the dynamical evolution of sentiment in novels. *Proceedings of 2016 International Conference on Behavioral, Economic and Socio-Cultural Computing (BESC)* (pp. 1-4). Durham, NC: IEEE.
- Giachanou, A., & Crestani, F. (2016). Tracking sentiment by time series analysis. *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 1037-1040). New York, NY: Association for Computing Machinery.
- Goodwin, M. A. (2015). Mapping memory in Tran's *Vietnamerica*. *CLCWeb: Comparative Literature and Culture*, 17(3), 1-7.
- Grishman, R. (1986). *Computational linguistics: An introduction* (1st ed.). Cambridge University Press.
- Hirsch, M. (1997). *Family frames: photography, narrative, and postmemory*. Harvard University Press.
- Hirsch, M. (2001). Surviving images: Holocaust photographs and the work of postmemory. *Yale Journal of Criticism*, 14(1), 5-37.
- Hirsch, M. (2012). *The generation of postmemory: Writing and visual culture after the Holocaust*. Columbia University Press.
- Hong, C. K. (2014). Disorienting the Vietnam War: G.B. Tran's *Vietnamerica* as transnational and transhistorical graphic memoir. *Asian American Literature: Discourses and Pedagogies*, 5, 11-22.

- Jockers, M. L. (2013). *Macroanalysis digital methods and literary history*. University of Illinois Press.
- Keller, N. O. (2002). *Fox girl*. Viking.
- Kim, C. H. (2018). *If you leave me*. William Morrow.
- Kim, D. Y. (2009). "Bled in, letter by letter": Translation, postmemory, and the subject of Korean War: History in Susan Choi's *the foreign student*. *American Literary History*, 21(3), 550-583.
- Kim, N. Y. (2008). *Imperial citizens: Koreans and race from Seoul to LA*. Stanford University Press.
- Kim, R. E. (1964). *The martyred*. George Braziller.
- Krashen, S., Mason, B., & Smith, K. (2018). Some new terminology: Comprehension-aiding supplementation and form-focusing supplementation. *Journal of English Language Teaching*, 60(6), 12-13.
- Lee, C. (2010). *The surrendered*. Riverhead Books.
- Lee, E. J. (2004). *The good man*. Bridgeworks.
- Maynard, D., & Funk, A. (2012). Automatic detection of political opinions in tweets. In R. García-Castro, R. Fensel, & D. Antoniou (Eds.), *Proceedings of the 8th International Conference on The Semantic Web: ESWC 2011 Workshops* (pp. 88-99). Berlin, Germany: Springer.
- Medhat, W., Yousef, A. H., & Mohamad, H. K. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5(4), 1093-1113.
- Mohammad, S. (2011). From once upon a time to happily ever after: Tracking emotions in novels and fairy tales. *Proceedings of the 5th ACL-HLT Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities* (pp. 105-114). Portland, OR: Association for Computational Linguistics.
- Moretti, F. (2000). Conjectures on world literature. *New Left Review*, 1, 54-68.
- Moretti, F. (2005). *Graphs, maps, trees: Abstract models for literary history*. Verso.
- Mohammad, S. (2022, August). *NRC Word-Emotion Association Lexicon*. Saif M. Mohammad's Research Website. <https://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm/>
- Pak, T. (1983). *Guilt payment*. Bamboo Ridge Press.
- Pozzi, F. (2017). *Sentiment analysis in social networks* (1st ed.). Morgan Kaufmann.
- Ptaszynski, M., Dokoshi, H., Oyama, S., Rzepka, R., Kurihara, M., Araki, K., & Momouchi, Y. (2013). Affect analysis in context of characters in narratives. *Expert Systems with Applications: An International Journal*, 40(1), 168-176.
- Silge, J., & Robinson, D. (2017). *Text Mining with R: A Tidy Approach*. O'Reilly Media.
- Smith, B. H. (2016). What was "close reading"? A century of method in literary studies. *Minnesota Review*, 87(1), 57-75.
- Yoon, P. (2013). *Snow hunters*. Simon & Schuster.

〈Korean Abstract〉

김현욱, 백승현. (2023). 1세대와 2세대 한국계 미국인 작가의 한국전쟁소설에 나타난 감정 어휘의 비교: R을 이용한 감성분석. *외국어교육연구*, 37(4), 67-82.

본 연구의 목적은 1세대와 2세대 한국계 미국 작가가 집필한 한국전쟁소설에 나타난 감정 어휘의 사용 양상을 검토하고, 각 세대별 감정 및 감정 어휘 사용 빈도를 비교하여 각 세대별 감정 어휘 사용의 기대치를 파악하고자 하였다. 감성분석과 카이제곱 검정을 기반으로 다음과 같은 결과를 도출하였다: (1) 빈도 측면에서 1세대, 2세대 모두가 긍정 감성 어휘를 부정 감성 어휘보다 더 많이 사용하고 1세대가 2세대에 비해 더 많은 긍정 감성 어휘를 사용하였다. (2) 세대 간 감정 어휘 사용 예상 비율의 측면에서 1세대가 2세대에 비해 긍정 감성 어휘를 사용할 확률이 1.076배 더 높은 것으로 나타났다. 멀리서 읽기 접근법을 활용한 본 연구는 꼼꼼히 읽기에 기반을 둔 Hirsch (2012)의 포스트 메모리 연구의 주장과 상반된 결과를 도출하였다. 즉, 1세대의 트라우마적 기억이 2세대에게 전이될 것이라는 Hirsch의 주장에 따르면, 세대와 상관없이 두 세대 작가군 모두가 부정 감성 어휘를 긍정 감성 어휘보다 더 많이 사용했어야 한다. 세대별 감정 어휘 사용 빈도와 비율의 차이에 대한 잠재적 원인과 교육적 시사점을 논의하였다.

Key words: Chi-squared test, Close reading, Digital humanities, Distant reading, Sentiment analysis/ 카이스퀘어 검정, 꼼꼼히 읽기, 디지털 인문학, 멀리서 읽기, 감성분석

Examples in: Korean, English
Applicable Languages: English
Applicable Levels: University

Kim, Hyunwook
Doctoral Student
Department of English, University of South Carolina
1400 Greene St, Columbia, SC 29225, USA
TEL: (803) 777-4203
E-MAIL: hyunwook@email.sc.edu

Baek, Seunghyun
Professor
Department of English Language and Literature, Chonnam National University
77 Yongbong-ro, Buk-gu, Gwangju, Korea 61186
TEL: (062) 530-3169
E-MAIL: sh3940@jnu.ac.kr

Received in October 09, 2023
Revised version received in November 03, 2023
Revised version accepted in November 05, 2023