



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

제 138 회 석사학위논문

지도교수 김 선 회

미국 아동 애니메이션 영화 대본에 나타난
영어 어휘들의 빈도 분석

**An Analysis of English Word Frequency in the American
Children's Animation Movie Scripts**

중앙대학교 대학원

영어영문학과 영어학전공

김 진 하

2023년 2월

미국 아동 애니메이션 영화 대본에 나타난
영어 어휘들의 빈도 분석

**An Analysis of English Word Frequency in the American
Children's Animation Movie Scripts**

이 논문을 석사학위논문으로 제출함

2023년 2월

중앙대학교 대학원
영어영문학과 영어학전공
김 진 하

김진하의 석사학위논문으로 인정함

심사위원장 최 영 진 (인)

심 사 위 원 이 성 호 (인)

심 사 위 원 김 선 회 (인)

중앙대학교 대학원

2023년 2월

목 차

제1장 서 론	1
1.1 연구 목적과 필요성	1
1.2 연구 자료의 범위	2
1.3 본고의 구성	7
제2장 이론적 배경	8
2.1 코퍼스 연구	8
2.2 지프의 법칙	10
2.3 애니메이션 영화	11
2.4 코퍼스와 애니메이션을 활용한 선행 연구	12
제3장 연구 방법	15
3.1 분석 자료	15
3.2 분석 방법 및 절차	17
제4장 분석 결과 및 논의	21
4.1 영어 어린이 애니메이션 영화 어휘 분석 결과	21
4.2 애니메이션 영화 대본 코퍼스 어휘 빈도수 그래프 분석	23
4.3 애니메이션 영화 <i>UP</i> 대본 코퍼스 분석	25
4.4 애니메이션 영화 <i>Inside Out</i> 대본 코퍼스 분석	29
4.5 애니메이션 영화 <i>Soul</i> 대본 코퍼스 분석	32
4.6 명사 사용 빈도 비교분석	34
4.7 동사 사용 빈도 비교분석	37

4.8 형용사 사용 빈도 비교분석	39
제5장 결론	41
참고문헌	44
국문초록	47
<i>Abstract</i>	49

표 목 차

<표 1> <i>UP, Inside Out, Soul</i> 의 타입과 토큰	21
<표 2> <i>Type-Token Ratio</i> 비교	22
<표 3> 영화 <i>UP</i> 대본 코퍼스 명사 빈도순 20 이내	26
<표 4> 영화 <i>UP</i> 대본 코퍼스 동사 빈도순 20 이내	27
<표 5> 영화 <i>UP</i> 대본 코퍼스 형용사 빈도순 20 이내	28
<표 6> 영화 <i>Inside Out</i> 대본 코퍼스 명사 빈도순 20 이내	29
<표 7> 영화 <i>Inside Out</i> 대본 코퍼스 동사 빈도순 20 이내	30
<표 8> 영화 <i>Inside Out</i> 대본 코퍼스 형용사 빈도순 20 이내	31
<표 9> 영화 <i>Soul</i> 대본 코퍼스 명사 빈도순 20 이내	32
<표 10> 영화 <i>Soul</i> 대본 코퍼스 동사 빈도순 20 이내	33
<표 11> 영화 <i>Soul</i> 대본 코퍼스 형용사 빈도순 20 이내	34
<표 12> <i>UP, Inside Out, Soul</i> 대본 코퍼스 명사 빈도순 10 이내	35
<표 13> <i>UP, Inside Out, Soul</i> 대본 코퍼스 동사 빈도순 10 이내	37
<표 14> <i>UP, Inside Out, Soul</i> 대본 코퍼스 형용사 빈도순 10 이내	39

그림 목 차

<그림 1> 레마(<i>lemma</i>) 리스트 예시	18
<그림 2> 영화 <i>UP</i> 어휘 빈도 분포	23
<그림 3> 영화 <i>Inside Out</i> 어휘 빈도 분포	24
<그림 4> 영화 <i>Soul</i> 어휘 빈도 분포	24

제1장 서론

1.1 연구 목적과 필요성

대규모 데이터와 통계 소프트웨어의 이용이 어려웠던 시기부터 이러한 것들의 이용이 용이하게 된 최근에 이르기까지 여러 학자들은 다양한 말뭉치에 포함된 어휘들의 사용에 대한 분석을 시도하였다. 언어학적 관점에서는 어휘들의 사용 빈도를 조사하고 해당 언어의 어휘 사용 특성을 논의하기도 하고, 사용된 어휘들의 빈도가 어떠한 분포를 보이는지 확률적으로 분석하기도 하였다. 텍스트 분석의 관점에서는 어휘 사용에서 나타나는 텍스트의 특성에 대해 논의하기도 하고 여러 텍스트의 어휘 빈도에서 관찰되는 인간의 언어 행위의 특성을 분석하기도 하였다. 최근에는 문학 텍스트의 분석에서도 어휘의 사용 양상이 활용되기도 한다. 어휘 사용에 관한 비교적 초기 연구들에서는 주로 문학 텍스트에 나타난 어휘들의 빈도 분포를 분석하였다.

그리고 Zipf(1949)는 James Joyce의 소설 *Ulysses*에 포함된 어휘를 분석하여 어휘 빈도순에 따른 어휘 빈도는 $1, 1/2, 1/3, \dots$ 로 분포한다는 지프의 법칙(Zipf's law)를 제안하였고, Newman(2005)와 Clauset et al.(2009)은 Herman Melville의 소설 *Moby Dick*에 포함된 어휘들의 빈도에 대한 분석을 통해, 지프의 법칙이 모든 텍스트의 어휘 빈도에 적용될 수 있음을 보여 주었다. 그러나 문학 텍스트에 포함된 어휘들의 빈도에 대한 연구들 중 아동들을 대상으로 하는 텍스트를 분석한 연구는 상대적으로 드물다. 그 이유 중 하나는 어휘 빈도와 관련된 연구가 주로 언어학적 관점에서 이루어진 반면, 아동을 대상으로 한 텍스트에 나타난 어휘들이 해당 언어의 어휘 빈도를 충분히 반영하고 있지 못하기 때문이다. 즉, 아동

을 대상으로 한 텍스트는 어휘 빈도에 나타난 보편적인 언어학적 특성을 분석하고자 하는 연구들을 위해서는 적절한 말뭉치라고 볼 수 없다. 그러나 만약 아동을 대상으로 한 텍스트에 나타난 어휘들의 빈도가 언어 습득 과정 중에 있는 아동 또는 성인이 되기 전의 아동들의 해당 언어의 어휘 빈도를 반영한다고 가정한다면, 이러한 텍스트에 대한 어휘 빈도 분석 연구가 그들의 어휘 사용 특성을 보여줄 수 있다는 점에서 학문적 의의가 있다고 볼 수 있다. 또한, 텍스트에 포함된 어휘 빈도의 특성 분석을 통해 충분하지는 않지만, 텍스트의 특성을 파악할 수 있다면 텍스트 분석의 관점에서도 그 의의를 찾을 수 있을 것이다.

이러한 관점 하에, 본 연구는 아동을 대상으로 한 영어 텍스트 가운데 애니메이션 영화 대본 세 편 *UP*, *Inside Out*, *Soul*을 택해, 해당 텍스트에 포함된 어휘들의 빈도를 분석하고 그 분석을 통해 드러난 어휘 사용의 특성들에 대해 논의하고자 한다.

1.2 연구 자료의 범위

본 연구의 분석 자료는 *UP*, *Inside Out*, *Soul*로 모두 미국 영화사 ‘Fixar’의 제작으로 상영된 아동 애니메이션 영화이다. 이 애니메이션을 선택한 이유는 다음과 같다. 첫째, 동일한 주제에 따라 구성된 텍스트 일 경우 어휘 빈도의 비교가 의미를 가질 수 있다는 판단에서다. 다른 주제를 다룬 텍스트를 비교할 경우 출현 어휘들의 종류와 빈도는 다를 수밖에 없다. 따라서 어휘 빈도의 비교가 무의미할 수 있다. 이런 판단 하에, ‘가족’과 ‘성장’ 그리고 ‘모험’을 주제로 다룬 텍스트를 연구 자료로 선정하였다. 또한, 본 연구가 아동들의 어휘 사용에 초점을 맞추고 있

으므로, 직접 아동들이 발화한 텍스트는 아닐지라도, 아동을 대상으로 애니메이션 자료의 범위를 국한하였다. 애니메이션 영화의 특성상 그 내용이 아동을 대상으로 하고 있을지라도 모든 연령대가 이해하고 공감하는 내용으로 구성되어 있어, 모든 연령대가 시청할 수는 있다. 그러나 주요 시청 대상으로 아동을 포함하고 있고 그 내용도 아동을 중심으로 이루어지므로, 사용되는 어휘들의 범위와 수준은 아동들에 맞추어져 있는 것이 일반적이다.

둘째, 본 연구가 영어 어휘 빈도를 비교하고 분석하는 데 있으므로, 영어를 사용하는 애니메이션 영화로 국한하였다. 제작된 애니메이션의 언어적 환경 및 배경은 모두 미국이다. 마지막으로, 동일한 배경 속에서 가족이라는 주제가 아동들에게 전달하고 있는 메시지의 중요성 역시 고려 대상이었다. 따라서 본 연구가 분석 대상으로 택한 애니메이션들은 비슷한 연령층의 아동들에게 가족을 떠올릴 만한 배경 묘사나 특정한 어휘들의 사용 빈도가 높을 것으로 예상할 수 있다.

본 연구는 코퍼스 언어학의 관점에서 텍스트를 분석하므로, 코퍼스 언어학에서 행해진 분석 방법을 연구의 방법으로 택한다. Biber(2015, p.15)에 따르면, 코퍼스 기반 분석의 본질적 특징 가운데 하나는 대규모로 일정한 원칙에 따라 수집된 “코퍼스”라는 자연 텍스트를 분석의 기초로 이용한다. 그러므로 정량적 분석 기법과 정성적 분석 기법을 모두 활용하는 것이다. 연구 방법과 관련한 본 연구의 의의 가운데 하나는 이미 선행 연구에서 수집하고 정제 과정을 거친 말뭉치가 아니라 본 연구가 텍스트를 직접 선정하고 어휘를 추출하여 정제한 말뭉치를 분석 대상으로 삼았다는 점이다.

애니메이션 영화에 포함된 어휘들을 영어 교육적 관점에서 분석한 연구들이 존재하기는 하지만, 본 연구는 언어학적 관점 텍스트 분석의 관

점에만 초점을 맞추고자 한다. 언어 연구는 크게 구조 연구와 용법 연구로 나눌 수 있다. 구조 연구가 언어 현상의 분석을 통해 언어 구조에 내재된 특성을 밝히고자 하는 연구라면, 용법 연구는 언어 사용자의 사용양상을 분석하여 언어 사용에서 나타나는 특성을 밝히고자 하는 연구라고 할 수 있다, 즉, 화자나 저자가 어떻게 언어 자원을 최대한 사용하여 그들의 의도를 청자에게 전달하고자 하는 방식의 특성을 살펴볼 수 있다 (Biber 외, 2015, p.11). 이런 측면에서 볼 때, 본 연구는 구조 연구라기 보다는 용법 연구에 더 가깝다고 볼 수 있다.

본 연구가 채택한 코퍼스 기반 연구 방법의 장점 중 한 가지는 정량적 분석을 위해 다양한 컴퓨터 소프트웨어를 활용할 수 있다는 것이다. 컴퓨터 소프트웨어를 활용하면 수작업으로는 달기 불가능한 방대한 규모의 자연 언어의 데이터베이스를 용이하게 구축하고 분석할 수 있기 때문에, 복잡한 언어 용법 유형들을 확인하고 분석하는 것이 가능하다. 그리고 일관되고 신뢰도 높은 분석 결과를 제공할 수 있다. 또한, 컴퓨터 소프트웨어가 입력된 자료를 정리하고 분석하는 동안, 연구자는 어려운 의사 결정과 언어적 판단을 내릴 수 있다(Biber 외, 2015, p.15).

앞에서 언급했듯이, 이러한 코퍼스 기반 연구 방법의 특성에 비추어 볼 때, 본 연구는 기존에 이미 존재하는 정제된 말뭉치를 대상으로 분석하기보다는 정제된 말뭉치를 구축하여 그것에 대해 분석한다는 점에서 의의가 있다고 할 수 있다. 또한, 영어 어휘를 분석하는 데 있어 영어권 문학·문화적 배경을 가진 애니메이션 자료를 활용하는 것도 의미가 있다고 할 수 있다. 특히, 영어 습득 과정 중에 있는 아동에 초점을 맞춘 텍스트들이라는 점에 주목할 수 있다. 그러므로 영어 습득 과정 중에 출현하는 어휘 빈도의 특성을 분석한다고 볼 수 있다는 점에서 본 연구의 결과는 영어 습득 분야에도 흥미로운 시사점을 줄 수 있을 것으로 기대한

다. 무엇보다도, 지금까지 본 연구와 유사한 선행 연구들을 거의 찾을 수 없다는 점에서 본 연구는 이 분야 후속 연구를 위한 기초 자료를 제공할 수도 있을 것이다.

앞에서도 언급했지만, 코퍼스 분석 기반으로 애니메이션 영화를 활용한 선행 연구들은 그리 많지 않은데, 몇몇 관련 연구들은 영어 교육과의 관련성 속에서 이 주제를 다루었다. 김은현(2010)은 아동 판타지 영화의 코퍼스 기반 어휘 분석에서 초등 영어교육 과정의 어휘와 비교 분석하여 어휘 사용의 유사점과 차이점을 밝혔다. 그러나 연구 분석에 사용된 판타지 영화와 동화에 비교 대상인 영어 교과서의 문장 수와 어휘 수가 매우 적으며 어휘의 의미, 해석 외에는 보충 설명이 충분하지 않다.

김은아(2013)는 영어 애니메이션 영화의 어휘 수준에 관한 코퍼스 기반 분석에서, 초·중등 교육과정의 권장 어휘 목록과 비교하여 애니메이션 영화에 사용된 어휘와 비교하였다. 1940년부터 2013년까지 국내에서 상영된 영어 애니메이션 64편의 대본 코퍼스를 구축하여 BNC코퍼스 기반으로 애니메이션 영화의 이해에 필요한 객관적인 어휘 수준을 비교한 결과, 애니메이션 영화의 어휘 규모가 초등영어 교육과정에서 배우는 어휘 유형의 수를 훨씬 초과한다는 것을 보여 주었다. 그러나 사용된 영어 애니메이션을 선정하게 된 명확한 기준이 필요하며, 다량의 코퍼스를 통한 양적 분석에 그치고 더 나아가 다양한 논점이 없음에 다소 아쉬운 점이 있다.

이가람(2012)은 시리즈 영어 애니메이션 6종을 코퍼스로 구축하여 각각의 애니메이션 어휘의 양적·질적 특성과 수준을 분석하였다. 그래서 한국 초등학생 영어 학습자에게 활용하는 것을 다루었다. 그 결과, 시리즈 영어 애니메이션의 어휘는 초등영어 교과서 어휘보다 다양했다. 그리고 어휘의 반복 빈도수가 초등학교 교과서에서의 반복 빈도수보다 높게

나왔다. 그러므로 시리즈 영어 애니메이션은 필수 기본 어휘가 많이 반영되어 초등영어 교육 자료로 활용하기에 적합하다고 판단한 것이다. 애니메이션 대본에서 어휘들의 반복됨에 대한 빈도수를 측정한 점은 연구 목적에 다양하게 활용될 수 있었다. 위의 선행 연구들을 통해 보았을 때, 전체적으로 한국의 영어 교육안에서 활용되는 방안을 제안한 애니메이션 영화 코퍼스 기반 분석 연구가 진행되는 경향이 있음을 알 수 있다.

본 연구는 앞에서 언급한 선행 연구들과는 달리 영어 교육과의 관련성보다는 언어학적 관점에서 어휘 사용의 빈도에 대한 분석에 초점을 맞춘다. 본 연구에서 코퍼스 기반 양적 분석을 통해 다루게 될 연구 문제들은 다음과 같다.

첫째, 세 편의 애니메이션 영화 대본 텍스트에 포함된 어휘들을 총 출현 횟수를 나타내는 토큰(token)과 출현한 어휘 유형의 수를 나타내는 타입(type)으로 구분하여 그 차이가 어떠한지 조사한다. 토큰은 문어 혹은 구어 텍스트 안의 어형을 모두 세는 것으로 띄어쓰기 기준 텍스트를 구성하는 총 어휘의 개수를 의미하고, 타입은 텍스트 내의 어휘의 종류를 파악할 수 있다(김윤서, 2016, p.32). 그리고 토큰 대비 타입의 비율 (Type-Token Ratio)을 조사하여 각 애니메이션에서 어휘 사용의 다양성 정도를 비교한다. 이러한 조사를 통해 어휘 사용 측면에서 각 애니메이션 텍스트의 규모와 어휘 사용 방식을 알 수 있을 것이다 (김은현, 2010, p.33).

둘째, 각 애니메이션 영화 대본 텍스트에 등장하는 어휘들이 주는 정보들은 무엇이며, 미국 아동 애니메이션 영화 대본 코퍼스에 대해 살펴볼 것이다. 본 연구 분석에서 사용될 애니메이션 영화에 공통된 주제는 ‘가족’ 과 ‘성장물’ 이다. 이러한 애니메이션 영화에서 볼 수 있는 영어 어휘들을 나열하였을 때 각 어휘의 분포를 비교하여 주제와 관련된 정

보들에 대해 논의할 수 있다.

셋째, 총 세 개의 미국 애니메이션 영화 간 어휘 빈도수를 분석하여, 미국 애니메이션 영화에서 높은 빈도로 사용된 어휘들을 알아볼 것이다. 예를 들면, 명사, 동사, 형용사 품사별로 나눈 어휘 리스트를 구축한다. 그리고 각 대본 코퍼스에서 출현한 어휘 빈도가 높은 순으로 알아본다. 본 논문에서는 미국 아동 애니메이션 영화 대본 텍스트에서 출현 빈도가 높은 어휘를 제시하고 이를 통해 미국 애니메이션 영화의 어휘 사용 특성을 분석하고 논의할 것이다.

1.3 본고의 구성

본고의 구성은 다음과 같다. 2장에서는 이론적 배경을 설명한다. 코퍼스 언어학과 애니메이션에 대한 정의를 설명한다. 나아가 코퍼스 언어학과 애니메이션을 활용한 선행 연구들을 살펴본다. 그리고 3장에서는 본 연구의 분석 자료들을 밝히고 총 세 가지 영어 애니메이션 영화들의 배경과 특징들을 설명한다. 또한, 연구의 절차 및 방법을 소개한다. 4장에서는 컴퓨터 프로그램을 활용하여 코퍼스를 분석한 결과를 중점적으로 다룬다. 고빈도 어휘 리스트나 공통된 어휘 사용을 비교하여 분석한 결과를 설명한다. 그리고 빈도와 토큰(token) 및 타입(type)의 결과를 놓고 이야기한다. 4장에서는 미국 아동 애니메이션 영화 *UP*, *Inside Out*, *Soul* 총 세 편의 코퍼스 데이터를 활용하여 분석하였다. 각 대본 코퍼스의 분석 결과를 명사, 동사, 형용사 각 20개의 고빈도 어휘들로 나누었다. 그런 다음, 품사마다 비교한 결과를 공통점 및 차이점을 논한다. 마지막으로, 5장에서는 본 연구의 코퍼스 기반 분석 결과들에 대해 결론 및 한계점에 대해 논의한다.

제2장 이론적 배경

2.1 코퍼스 연구

권혁승(2008, p.3)은 코퍼스 언어학이란, 언어 연구의 여러 분야 중 음운, 형태, 구조, 의미와 같이 특정한 분야나 특정한 이론을 연구하는 것이 아니라, 언어 연구의 제 영역에서 다양한 언어 현상에 대한 답을 찾기 위해 코퍼스를 이용하여 언어학적 분석, 기술, 연구를 수행하는 연구 방법론이라고 정의했다. 그리고 코퍼스 언어학은 음운론, 통사론 등 언어학의 특정 분야를 독립적으로 형성하지 않고, 음성, 음운, 형태소, 어휘, 구, 절, 문장, 담화 등 모든 언어적 단위에 적용된다. 이는 코퍼스 언어학의 독립된 학문 체계를 갖기보다는 각 연구 단위에 응용되는 도구 학문적 성격을 갖는다는 것을 보여 준다(김진해, 2006, p.77).

특히, 코퍼스 연구에서 어휘 분석을 할 때, 텍스트의 상당 부분을 구성하는 고빈도 어휘들에 주목할 필요가 있다. 또한, 해당 코퍼스의 텍스트 장르나 크기 또한 코퍼스 연구에 결정적인 요소가 된다. 앞에서 언급한 바 있듯이, 타입과 토큰 비율은 어휘적 다양성 정도를 살펴보는 데 유용한 정보를 제공한다.

이론 언어학은 인간이 선천적으로 타고난 언어 능력의 본질에 대한 규명에 초점을 맞추고 있으나, 코퍼스 언어학은 인간의 언어 사용에서 나타나는 특성에 주목한다. 따라서 코퍼스 언어학은 이론 언어학과 언어 연구의 대상에 있어서 상반되는 개념을 갖고 있으나 (권혁승, 2008, p3), 이론 언어학에서 제기된 문제 또는 새롭게 밝혀진 여러 현상에 대한 계량적 증거를 제시하기도 하고 코퍼스 언어학적 결과가 이론 언어학적 증거

들에 의해 뒷받침되기도 하므로 이론 언어학과 코퍼스 언어학은 상호 보완적 관계에 있다고도 볼 수 있다. 코퍼스를 활용할 수 있는 분야는 언어 연구의 거의 모든 분야에 걸쳐서 가능하다. 즉 어휘, 문법, 의미, 담화, 화용, 문체, 언어교육, 언어역사, 방언, 심리언어, 언어문화, 사회언어 등의 연구에서 코퍼스 자료를 활용하여 보다 정확하고 실증적 자료를 얻을 수 있다 (김성식, 2007, p.293). 게다가 이론 언어학과 달리 코퍼스 언어학은 연역적 연구 방법이 아닌 실제 사용한 언어를 관찰하고 분석한 통계적 결과를 이론으로 발전시키는 귀납적 연구 방법을 취하고 있다 (권혁승, 2008, p.4). 그리고 코퍼스 연구는 구어와 문어 등 각 영역에서 사용되는 패턴이나 어휘들을 확인하고, 텍스트에 나타난 언어 사용의 화용 정보까지 파악할 수 있어 언어의 진정성 연구에도 도움이 된다 (김은현, 2010, p.119). 또한, 산출된 언어 자료를 이용하여 언어를 기술하는 코퍼스 언어학의 연구 방법의 한계를 지적하는 견해들도 있다. 그러나 대규모 언어 자료를 수집할 수 있고 이것을 과학적 분석 도구를 사용하여 분석할 수 있다면, 이를 통해 도출된 결과는 인간의 언어 능력을 반영한 것으로 볼 수 있을 것이다.

따라서 코퍼스 언어학의 연구 방법에서 중요한 것은 일관되고 오류 없이 코퍼스를 수집하고 분석할 수 있는 수집 및 분석 도구를 선택하는 것이다. 이와 관련된 컴퓨터 프로그램 및 소프트웨어의 종류는 매우 다양하다. 어휘 분석을 위한 것들에 국한하여 살펴본다면 다음과 같은 것들을 수 있다. 어휘 목록 및 색인 목록을 보여주는 KWIC(Key Word in Context) 기능을 가진 Concordance Generator(Smith, 1966), CLOC(Reed, 1978) 등의 초창기 프로그램에서 오늘날의 WordSmith Tools, MonoConc Pro, AntConc 같은 개인용 프로그램이나, Corpus.byu.edu, SketchEngine, Wmatrix 같은 웹기반 프로그램 등이 존재한다. 이 분석도구들은 개발자

가 다르기 때문에, 자료 처리 방식이 다르거나 호환이 안 되는 경우가 있고, 웹기반 프로그램은 사용자 등록을 하고 정기적으로 사용권을 구매해야 하는 어려움이 있다 (김성식, 2018, p.53). 또한, 연구 결과의 유의미성을 확인하기 위해 코퍼스 언어학의 연구 방법에서 반드시 수행되는 것은 계량적 분석을 위한 통계 분석이다.

2.2 지프의 법칙

이탈리아 경제학자 빌프레도 파레토(Vilfredo Pareto)는 사회와 사람들의 돈과 영향력 면에서 상위 20%인 ‘vital few’와 하위 80%인 ‘trivial many’로 자연스럽게 나뉘는 것을 발견했다. 당시 이탈리아 인구의 20%가 국가 전체 부(富)의 80%를 보유하고 있음을 발견한 것이었다. 이 법칙은 흔히 80:20 법칙 또는 "파레토 원리"라고 학자 이름을 따서도 불린다. 이처럼, 언어학 쪽에서도 이러한 패턴들이 존재한다. 이것을 발견한 사람은 하버드 대학 언어학 교수 George Kingsley Zipf이다. 지프는 사람들이 쓰는 어휘들을 사용 횟수별로 순위를 매겨 어휘 사용 빈도를 기준으로 조사하였다. 그 결과, 극히 많이 쓰이는 소수의 어휘와 자주 사용되지 않는 다수의 어휘로 나누어지는 것을 발견하게 된다. 예를 들어, *the* 나 *and* 같은 어휘는 매우 많이 쓰이는데, 이러한 고빈도 어휘들은 몇 가지 개수에 불과하다는 것이다. 이것을 지프의 법칙(Zipf's law)라고 한다. 지프의 법칙의 핵심은 대량의 텍스트에 사용된 어휘에서 빈출 순위와 빈도를 집계하면, 빈출 순위가 k 번째 빈도는 빈출 순위 첫 번째 빈도를 $1/k$ 한 값이 되는 것이다. 빈출 순위가 두 번째인 어휘 빈도는 빈출 순위가 첫 번째인 어휘 빈도 약 $1/2$ 이고 빈출 빈도가 세 번째인 어휘 빈도는 빈출 순위가

첫 번째인 어휘 빈도의 약 1/3이다. 지프는 어느 텍스트로부터 영어 어휘마다 빈도의 높은 순서로 나열하면, 빈출 순위(r)와 빈도(f)를 곱하면 거의 같은 값이 되는 것을 발견했다.

2.3 애니메이션 영화

애니메이션(animation)이란, 라틴어로 ‘*anima*’로 ‘생명, 영혼, 정신’을 가리킨다. 즉, 움직이는 사물에 생명을 불어넣어 움직임을 준다는 의미를 가지고 있는데, 다른 말로는 동화(動畵)라고도 한다(김재원, 2008, p.97). 애니메이션은 하나의 문화 텍스트이고 텍스트로서 애니메이션은 그 안에 일정한 구조를 가지게 되고, 애니메이션의 서사 구조를 구성하는 스토리, 캐릭터, 애니메이션 움직임 등은 유기적인 연관성을 지니게 된다. 애니메이션의 시공간 속에는 문화가 존재하며, 그 문화는 애니메이션 제작자가 제공한 기반 안에서 관객들과 소통하며 새로운 문화를 만들어 내기도 한다(김지수, 2012, 4p). 영화와 애니메이션은 그 역사도 비슷하고 영사기와 촬영기 같은 영화 장치를 사용하여 완성한다는 점에서 매우 유사한 점을 갖는다.

컴퓨터의 등장으로 영화 장치들은 디지털화 되었지만 여전히 그 역할들은 유지되고 있다(정준호, 2017, p.17). 그리고 Gorčević(2021, p.177)은 애니메이션 영화 안에서 색깔이나 감정들이 어휘에 영향을 끼칠 만큼 중요하다고 주장한다. 그러므로 긍정적이거나 부정적인 의미를 가진 어휘들이 곧 사람의 감정을 유발한다고 볼 수 있다. 이러한 감정적인 표현들이 곧 언어적 기능 중 하나임을 말하고 있다. 또한, 애니메이션 영화 ‘도리를 찾아서(*Finding Dory*)’를 활용하여 초등 영어 교수·학습 개발

에 적용하여 보여준 강경아 외(2018)는 교과서와는 달리, 문화, 삶, 생활, 희로애락 등 인간의 다반사를 보여주는 매체라고 애니메이션 영화를 소개하였다. 그리고 아동을 대상으로 제작된 애니메이션은 대부분 폭력적인 내용이나 선정적인 내용을 포함하지 않으며, 언어의 적절성뿐만 아니라 정의적 요인에도 긍정적인 영향을 주기 때문에 교육적 자료로서 적합한 경우가 많다 (강경아 외, 2018, p.199). 더불어, 외국인들의 실제 문화를 알게 된 것, 언어의 표현이 실감나게 느껴진 것 등을 애니메이션 활용 수업의 장점으로 결론에서 제시하였다.

2.4 코퍼스와 애니메이션을 활용한 선행 연구

본 연구를 위해 코퍼스 분석 기반으로 애니메이션 영화를 활용한 선행 연구들을 살펴보았다. 김은현(2010)은 어린이 판타지 영화 코퍼스를 구축하고, 영어 동화 코퍼스와 초등 영어 교과서 코퍼스 어휘를 비교 및 분석하여 두 코퍼스 간에 어휘 사용의 유사점과 차이점을 밝혔다. 그 결과, 판타지 영화 코퍼스는 고빈도 어휘가 품사별 워드리스트에 일치하는 경우가 대부분이다. 반면에 영어 교과서 코퍼스는 워드리스트와 고빈도 어휘가 일치하는 개수가 적었다. 그러나 연구 분석에 사용된 판타지 영화와 동화에 비교 대상인 영어 교과서의 문장 수와 어휘 수가 상대적으로 매우 적으며 어휘의 의미, 해석 외에는 보충 설명이 충분하지 않다.

두 번째로, 김은아(2013)는 영어 애니메이션 영화의 어휘 수준에 관한 코퍼스 기반 분석에서, 초·중등 교육과정의 권장 어휘 목록을 애니메이션 영화에 사용된 어휘와 비교하였다. 그리고 코퍼스 분석을 위해 코퍼스 프로그램인 ‘WordSmith Tools 6.0’을 사용하였다. 1940년부터 20

13년까지 국내에서 상영된 영어 애니메이션 64편의 대본 코퍼스를 구축하여 BNC코퍼스 기반으로 애니메이션 영화의 이해에 필요한 객관적인 어휘 수준을 비교한 결과, 애니메이션 영화의 어휘 규모가 초등 영어 교육 과정에서 배우는 어휘 타입의 수를 훨씬 초과한다는 점을 알 수 있다. 그러므로 영어 애니메이션 영화를 영어 학습용으로 사용하는 경우엔 학습자의 수준을 고려해야 하며, 각별한 주의가 요구됨으로 결론을 내렸다. 그러나 사용된 영어 애니메이션을 선정하게 된 명확한 기준이 필요하며, 다량의 코퍼스를 통한 양적 분석에 그치고 더 나아가 다양한 논점이 없음에 다소 아쉬운 점이 있다.

세 번째로, 이가람(2012)은 아동 영어 학습자료로 사용되고 있는 시리즈 영어 애니메이션의 코퍼스를 구축하여 각각의 애니메이션 어휘의 수준을 분석하였다. 한국 초등학교 영어 학습자를 대상으로 어휘를 비교하였다. 그리고 시리즈 영어 애니메이션 어휘의 특성과 수준을 밝혔다. 연구의 목적은 영어 교사와 학습자에게 자료 선택에 필요한 지표를 제공하는데 의의를 두었다. 위의 선행 연구들을 통해 전체적으로 한국의 영어 교육에서 활용되는 방안을 제안한 애니메이션 영화 코퍼스 기반 분석 연구가 진행되는 경향이 있음을 알 수 있다.

네 번째로, 성민창, 김기택(2021)은 영어 애니메이션 영화를 대상으로 우리나라 초등학교 영어 교육에서 사용되는 어휘를 비교하여 총 25편의 영어 애니메이션 영화에 나오는 어휘를 분석하였다. 그래서 영어 애니메이션 영화에서 나오는 어휘 규모와 종류들을 보여주었다. 결론적으로, 타입-토큰(TTR) 분석을 통해 애니메이션 영화에서 고빈도 어휘가 반복적으로 사용됨을 알 수 있다. 그러나 통사, 형태, 음운 등 다양한 문법적 지식과 능력이 있는데 단순히 어휘 측면에만 국한시켰다는 한계를 결론에서 언급하였다.

마지막으로, 박서영(2019)은 애니메이션 영화에서 더빙 텍스트 코퍼스를 활용하였다. 연구 대상으로 구어 담화의 구성 요소인 담화표지를 분석하였다. 정량분석과 정성분석을 통해 연구방법에 적용하였다. 영상번역 연구와 코퍼스 연구방법을 선택하여 연구되었다. 그 결과, 영어 담화표지 번역을 위한 한국어 담화표지가 다양하다는 결과가 나왔다. 그리고 담화표지를 정리, 분석한 결과물들을 코퍼스 구축을 통해 연구에 실시시켰다는 것에 의의를 두었다. 다만, 영상 번역 분야에서 코퍼스 연구를 통해 접근한 방식은 찾아보기 어렵다. 그러므로 앞으로 영상 텍스트와 담화 그리고 코퍼스 연구방법을 통한 연구는 유의미한 결과로 이어 나갈 수 있을 것으로 기대해 볼 수 있다.

제3장 연구 방법

3.1 분석 자료

본 연구를 위해 선정한 세 편의 아동 애니메이션 영화 *UP*, *Inside Out*, *Soul*은 모두 픽사와 디즈니 영화사의 제작으로 상영된 것이다. 픽사(Pixar)와 디즈니(Disney)는 2006년에 합병되었으며, 현재 픽사는 디즈니의 자회사가 되었다. 분석 자료의 선정 조건 중 하나는 개봉 날짜가 너무 오래되지 않은 것이어야 한다는 것이었고 인지도 및 흥행 여도 고려 대상이었다. 다음은 분석 자료인 세 편의 애니메이션 영화에 대한 설명이다. 전부 동일한 주제인 ‘가족’, ‘성장’ 그리고 ‘모험’으로 분석 자료로 선정하였다. 첫 번째로, 애니메이션 영화 *UP*은 2009년에 개봉한 작품이다. 줄거리는 일찍이 아내를 먼저 잃은 남성 주인공 78세 노인 ‘칼(Carl)’의 모험 이야기가 전개된다. 이 애니메이션 영화 포스터에서도 볼 수 있듯이, 수 천 개의 풍선으로 자신의 집을 하늘로 들어 올린다. 하늘로 들어 올린 집으로 아내와 약속했던 곳을 향해 떠난다. 그리고 8세 소년 ‘러셀(Russell)’과 함께 여행을 시작한다. 노년 생활을 흥미진진한 모험으로 보낸 Carl은 마침내 과거에 아내와 함께했던 모든 날이 모험이었음을 깨닫게 된다. 게다가 애니메이션 영화가 끝날 때까지도 주인공 78세의 Carl의 여정은 끝나지 않는다. 그는 부인의 기억을 통해 매 순간 함께 한다.

두 번째로, *Inside Out*은 2015년에 개봉한 작품이다. 이 애니메이션 영화감독은 자신의 딸이 성장하면서 변화하는 성격을 보고 이야기를 구성

했다고 한다. 애니메이션 영화 속 주인공으로 소녀 ‘라일리(Riley)’와 그녀의 감정들인 ‘기쁨이(Joy)’, ‘슬픔이(Sadness)’, ‘소심이(Fear)’, ‘까칠이(Disgust)’ 그리고 ‘버럭이(Anger)’가 등장한다. 이러한 감정들은 ‘섬’으로 만들어져 있다. 그리고 애니메이션 영화 내용의 처음부터 끝까지 라일리의 탄생부터 모든 성장 과정들을 감정들을 의인화를 시켜서 보여준다. 한 가정의 미네소타에서 샌프란시스코로 이사를 하게 되면서 혼란을 겪게 되는 어린 소녀 라일리의 적응과정을 보여준다. 그래서 새로운 도시, 집, 학교에서 변화되는 감정들의 이야기들을 알 수 있다. 그리고 라일리의 일생에 중요한 날에만 만들어지는 ‘핵심기억(Core Memory)’이 만들어지는 과정을 보여준다. 그런데 주인공 소녀 라일리가 슬픈 기억을 할 때마다, 슬픔이는 핵심기억의 자리를 건들게 된다. 한 소녀의 성장을 통해 감정들과 함께 보여주면서 인생에 기쁨과 슬픔 그리고 좌절까지 생각해 볼 수 있는 교훈을 전달해준다.

이 애니메이션 영화 내용에서 주목할 점은, 애니메이션 영화에서 기쁨이(Joy)의 시선에서 라일리의 성장 과정을 보여준다는 점이다. 사건 전개를 슬픔이(Sadness), 소심이(Fear), 까칠이(Disgust) 그리고 버럭이(Anger)가 아니라, 기쁨이(Joy)로 설정한 것은 아이들에게 긍정적인 교훈을 주기 위한 장치라고 볼 수 있다.

마지막으로, *Soul*은 2020년에 개봉한 작품이다. 개봉 당시에는, 주인공 ‘조(Joe)’가 흑인이라는 점에서 개봉과 동시에 사람들의 이슈를 끌었다. 주인공뿐만 아니라, 다수의 등장인물이 흑인이 등장하여 흥미롭고 흔하지 않은 어린이용 애니메이션 영화 속 캐릭터들이다. *Soul* 영화 속의 주인공을 통해서, 인생에 있어서 풀리지 않는 문제에 답을 찾기 위한 모험을 보여준다.

즐거리는 중학교 교사이면서 재즈 뮤지션인 46세 Joe의 갑작스런

교통사고로 이야기가 시작된다. 그래서 주인공 Joe의 삶을 마감할 수도 있었을 법한 신기한 경험들을 보여준다. 우주의 영역인 ‘You Seminar’을 발견하고 들어가게 된다. 즉, 영혼들의 세계에 가게 되고 영혼들의 멘토 역할을 하게 된다. 결국, ‘영혼’과 ‘죽음의 세상’의 이야기를 통해 우리 각자의 삶들은 소중한 것이며 아름답다는 깨달음과 큰 울림을 주는 이야기이다.

본 연구의 준비를 위해 대본 코퍼스 구축 과정에 있어서, 영어 애니메이션 영화 대본들은 인터넷에 오픈소스(Open-Source)를 통해서 무료로 구할 수 있었다. 더구나 픽사(Pixar), 디즈니(Disney)의 애니메이션 영화들은 워낙 유명하여 인터넷 사이트에서 대본들을 쉽게 볼 수 있다. 본 논문에서는 애니메이션 영화 대본에 있는 내레이션(Narration) 대본들까지 포함하였다. 내레이션의 역할은 아동 애니메이션 영화에서 각 장면의 정보 제공 및 해석에 도움을 주어 전반적인 대사의 흐름에 많은 영향을 끼치기 때문에, 내용의 이해를 돕는 수많은 내레이션의 대사들은 본 연구의 분석에서 제외하지 않았다.

3.2 분석 방법 및 절차

본 연구에서는 도쿄 와세다 대학교의 Laurence Anthony 교수가 개발한 ‘AntConc3.5.8’ 프로그램을 사용하였다. 앞으로 4장 분석에서 어휘 리스트와 고빈도 수 등을 확인하기 위하여 사용하였다. 본 연구의 분석에서 사용된 AntConc3.5.8 프로그램은 Laurence Anthony의 홈페이지(<https://www.laurenceanthony.net/software/antconc/>)에서 누구나 무료로 이용할 수 있다. 본 연구에서 사용된 AntConc3.5.8은 장점이 많은 프

로그래밍이다. 프로그램의 크기가 매우 작아서 대부분의 컴퓨터 환경에서 큰 문제 없이 활용할 수 있다는 장점이 있다. 또한, 주요 컴퓨터 운영체제(Windows, Mac OS X, Linux)를 모두 지원하여 다양한 컴퓨팅 플랫폼에서 활용할 수 있다 (고광윤, 2019, p.20).

게다가 AntConc3.5.8은 코퍼스의 고빈도 어휘 리스트를 구축할 수 있으며 텍스트 파일의 어휘 규모와 어휘 다양성 등을 확인할 수 있다. AntConc3.5.8은 파일뷰(File View), 클러스터(Cluster), 콜로케이트(Collocate), 워드리스트(Word List) 등의 기능이 있다. 위 기능 중에 본 연구에서는 워드리스트(Word List) 기능을 사용하여 고빈도 어휘를 정렬하였다. 그리고 클러스터(Cluster) 기능을 사용하여 용례를 정리하였다. 또한, 동명사 또는 복수형 명사들의 본래 모습 어휘인 ‘레마(lemma)’를 설정하였다. 고빈도 어휘 리스트를 분석할 때 어휘의 원형을 비교할 때 도움을 받을 수 있었다. 이러한 레마 처리를 위해, AntConc 홈페이지에서 레마 파일을 다운로드할 수 있다.

<그림 1> 레마(lemma) 리스트 예시

Corpus Files

UP.txt

Concordance

Concordance Plot

File View

Clusters/N-Grams

Collocates

Word List

Keyword List

Lemma Types: 1889

Lemma Tokens: 13939

Search Hits: 0

Rank	Freq	Lemma	Lemma Word Form(s)
71	33	one	one 32 ones 1
72	33	through	through 33
73	32	balloon	balloon 18 balloons 14
74	32	see	saw 2 see 17 seen 5 sees 8
75	32	sit	sit 9 sits 23
76	31	explorer	explorer 25 explorers 6
77	31	stop	stop 20 stopped 1 stops 10
78	30	behind	behind 30
79	30	window	window 27 windows 3
80	29	close	close 9 closer 8 closes 10 closing 2
81	29	rock	rock 16 rocked 1 rocks 12
82	29	want	want 23 wanted 3 wants 3
83	28	call	call 12 called 3 calling 3 calls 10

프로그램 AntConc3.5.8의 시작 화면에서 바로 볼 수 있는 몇 가지 기능들이 있다. 고광윤(2019)의 「AntConc」에 따르면, AntConc의 기능들은 다음과 같다. 첫 번째로, Concordance 기능이다. 저장한 텍스트 파일에 찾고자 하는 어휘를 검색을 할 수 있다. 그리고 특정 어휘를 검색함으로써 같이 사용된 어휘와 문맥을 볼 수 있다. 두 번째로, File View의 기능은 Concordance에서 실행된 검색 내용들을 보여주는 창이다. 즉, File View는 코퍼스를 저장한 텍스트 파일을 자세히 보여주는 도구이다. 세 번째로, Clusters, N-Grams는 둘 다 텍스트 파일 안에 있는 어휘 덩어리들을 뽑아내는 도구로, 어휘들의 연속적인 패턴을 보여준다.

특히, Clusters는 검색어를 포함한 일정 길이의 어휘 덩어리를 보여준다. 빈출 어구를 뽑아내는 도구라고 할 수 있다. ‘어휘 덩어리’는 인간이 언어를 습득할 때 중요한 역할을 한다. 하나의 어휘가 아니라 같이 사용되는 어휘 덩어리로 익히게 되면 훨씬 기억하기 쉽다. 네 번째로, Collocates는 특정 문장 안에 연어(連語)를 검색할 수 있는 도구이다. 어떤 어휘는 다른 어떤 어휘와 함께 나타나는 경향이 있다. 이렇게 어휘들이 함께 나타나는 현상을 연어(連語)라고 한다. 즉, 두 어휘가 연어 관계가 없는 경우에 기대되는 것보다 더 많이 함께 나타나는 현상을 의미한다고 할 수 있다 (강범모, 2011, p.112).

다섯 번째로, Word List의 기능은 설정한 코퍼스 텍스트 파일에 사용된 어휘들의 목록을 만들어 주는 것이다. 그래서 텍스트 파일의 전체 토큰과 타입 수를 알 수 있다. 그뿐만 아니라, 파일 전체의 어휘 리스트를 만들어 한눈에 볼 수 있다. 여섯 번째로, Keyword List는 키워드를 산출해준다. 코퍼스 텍스트 파일에 있는 어휘 가운데 빈도수가 높은 어휘 또는 낮은 어휘를 키워드로 보여주는 기능이다.

본 연구 방법과 같이 Antconc와 엑셀 프로그램을 이용한 선행 연

구를 살펴보았다. 정소우(2014)는 영어 교육 논문 코퍼스에 나타난 학문 어휘 분석을 하였다. 연구 방법으로 Antconc와 엑셀 프로그램을 선택하였다. 분석 결과, 자료의 어휘량이 커짐에 따라, 자료에 나타나는 학문 어휘군과 학문 어휘 수는 증가하지만, 자료의 총 어휘 수 대비 학문 어휘 비율은 영향을 거의 받지 않는다는 것을 알 수 있었다. 그러나 연구 분야나 연구 주제는 학문 어휘 비율에 영향을 미친다는 것은 알 수 없었다.

본 연구의 분석의 절차는 다음과 같다. 첫째, 본 연구의 분석 대상인 총 세 편의 애니메이션 영화 작품의 대본을 수집하였다. 영어 애니메이션 영화 대본은 인터넷 사이트를 통해 무료로 다운을 받아볼 수 있다. 그다음, 대본들을 텍스트 파일(UTF-8)로 입력하고 저장하여 코퍼스 파일을 구축하였다. 둘째, 애니메이션 영화 대본 코퍼스를 각각 AntConc3.5.8 어휘 목록 기능을 사용하여 이용하여 고빈도 순으로 나열하였다. 셋째, 텍스트 파일로 입력 및 수정하여 구축한 대본 코퍼스를 추가적으로 수정하였다. 예를 들면, 한 장면당 부가적인 설명이 되어있는 문장이나 애니메이션 영화 대본 안에 있는 모든 고유명사와 의성어 그리고 문장부호를 제거하였다. 넷째, 수정된 코퍼스 텍스트 파일을 바탕으로 코퍼스 총 어휘 개수 그리고 타입과 토큰을 알아본다. 여섯째, 각 대본 코퍼스 데이터마다 어휘 사용 규모를 알아보기 위해 TTR의 비율을 알아본다.

일곱 번째, 통계 프로그램 R을 이용하였다. R 3.5.8을 활용하여 데이터를 비교 분석하였다. 그러기에 앞서, 제일 먼저 데이터를 불러오기 위해, 기존의 원본 데이터 파일들을 <.csv> 확장자로 표시되는 파일방식으로 새로 저장해준다. 그렇게 해야만 R 프로그램에서 파일을 읽기 위해 불러올 수 있기 때문이다. R 3.5.8을 활용하여 데이터를 비교 분석하였다. 그러기에 앞서, 제일 먼저 데이터를 불러오기 위해, 기존의 원본 데

이더 파일들을 <.csv> 확장자로 표시되는 파일방식으로 새로 저장해준다. 그렇게 해야만 R 프로그램에서 파일을 읽기 위해 불러올 수 있기 때문이다. 다음 제4장에서는 시각적으로 빈도의 형태를 알아보기 위해 빈도 그래프를 만들고 설명하였다. 또한, 각 대본 코퍼스마다 엑셀 프로그램을 이용하여 명사, 동사, 형용사별로 나누어 어휘 사용과 빈도를 비교한 결과에 대해 자세히 살펴보겠다.

제4장 분석 결과 및 논의

4.1 영어 어린이 애니메이션 영화 분석 결과

본 4장은 각각의 애니메이션 영화 대본 코퍼스의 분석 결과를 다루어 이야기한다. 앞에서 본 2장에서는 코퍼스 언어학과 애니메이션 영화에 대하여 설명하였다. 그리고 2.3장에서는 어린이 애니메이션 영화와 코퍼스 언어학을 활용한 선행 연구들을 살펴보았다. 다음 아래의 표는 본 연구의 분석 대상인 총 세 가지 어린이 애니메이션 영화 *UP*, *Inside Out*, *Soul*의 타입과 토큰의 개수를 의미한다.

<표 1> *UP*, *Inside Out*, *Soul*의 타입과 토큰

	<i>UP</i>	<i>Inside Out</i>	<i>Soul</i>
Type	2394	2962	2093
Token	13939	16753	19120

타입과 토큰의 개수는 코퍼스 데이터 비교할 때 가장 기본적이며 분석하기 유용하다. 타입이란, 어휘의 형태를 말한다. 그리고 토큰이란, 코퍼스 데이터인 문장들에서 띄어쓰기가 된 어휘의 전체 개수를 말한다. 즉, 코퍼스 어휘들의 전체규모로 볼 수 있다.

위의 <표 1>에서는 총 세 가지 애니메이션 영화 대본 코퍼스 정리 작업을 한 후의 타입과 토큰의 개수를 비교하였다. 세 편의 대본 코퍼스 분석의 첫 번째로, 서로 어휘의 양 크기 차이는 많이 나지 않은가 형평성을 위해 비교하여 알아본 것이다.

첫 번째로, 애니메이션 영화 *UP*에서 타입의 수는 2394이며, 토큰의 개수는 13939이다. 두 번째로, *Inside Out*의 타입의 수는 2962이며, 토큰의 수는 16753이다. 마지막으로, *Soul*의 타입의 수는 2093이며, 토큰의 수가 19120이 나왔다. 세 가지 어린이 애니메이션 영화 대본 코퍼스의 타입과 토큰의 개수가 큰 차이를 보이진 않았다. 다음은, 총 세 편의 애니메이션 영화 대본 코퍼스에서 토큰과 타입의 개수를 나누어 서로 비교해 보았다.

< Type-Token Ratio 계산 방법 >

$$TTR = \frac{Type}{Token} \times 100$$

<표 2> Type-Token Ratio 비교

	<i>UP</i>	<i>Inside Out</i>	<i>Soul</i>
TTR	17.2	17.7	10.9

텍스트의 크기가 작으면 일반적으로 타입/토큰 비율이 높으며, 텍스트의 크기가 클수록 그 비율이 낮아지는 경향이 있다 (강범모, 2011,

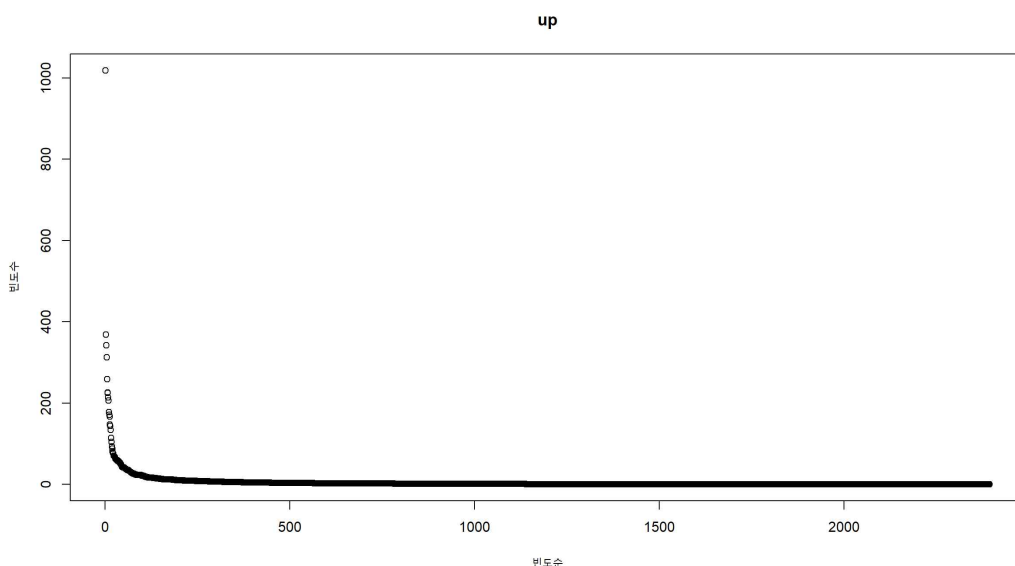
p.71). 그러므로 코퍼스 언어학에서 어휘를 분석할 때 토큰과 타입 그리고 Type-Token Ratio(TTR)은 중요한 분석지표가 된다. 토큰이란 어휘 계산(word count)을 뜻하며, ‘실행된 어휘’의 의미를 나타내는 ‘Running Word’로 사용되기도 한다. 이러한 TTR를 계산한 분석지표는 코퍼스의 규모를 나타내는 것이다 (김은현, 2010, p.33).

본 연구 대상인 세 편의 대본 코퍼스의 TTR 계산 결과, 텍스트의 크기가 클수록 비율이 낮아지므로, 세 가지 애니메이션 영화 대본 코퍼스의 규모는 *Soul* > *UP* > *Inside Out* 순서로 크다.

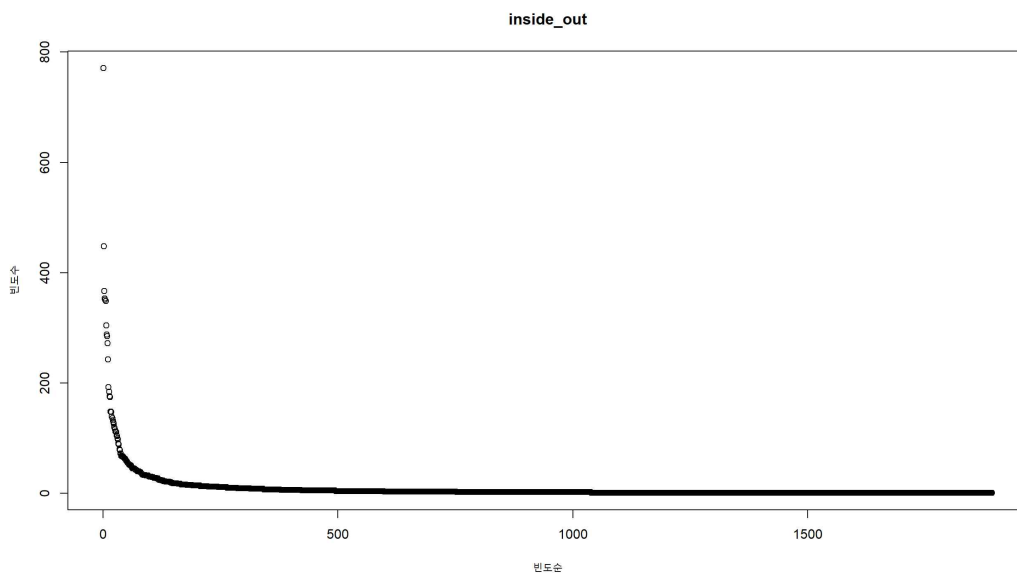
4.2 애니메이션 영화 대본 코퍼스 어휘 빈도수 그래프 분석

다음 아래의 그래프는 R studio를 이용하여 각 애니메이션 영화 대본 코퍼스의 총 어휘 빈도수를 나타낸 것이다. X축은 ‘빈도수’를 나타내며, Y축은 ‘빈도수’를 나타낸다.

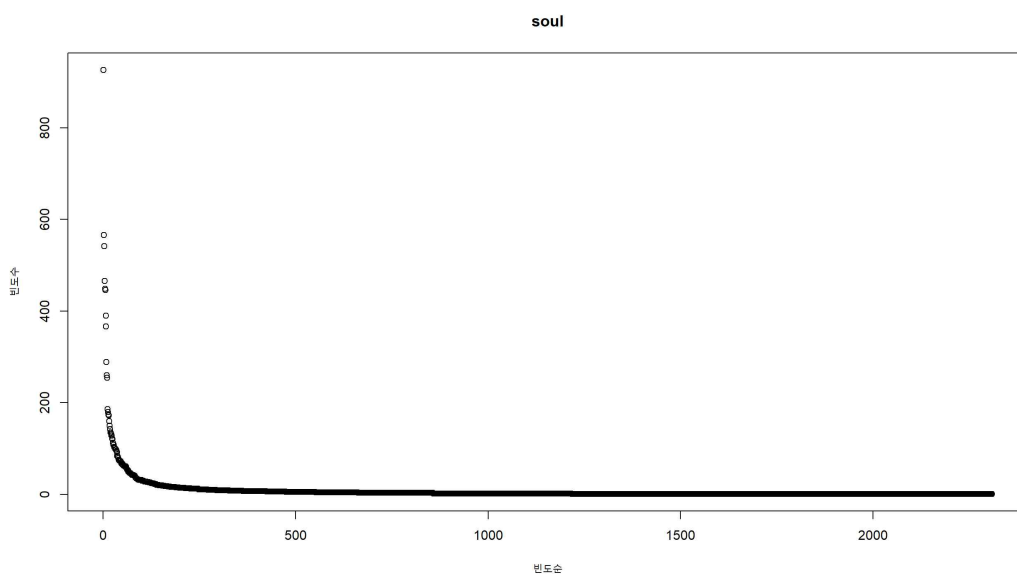
<그림 2> 영화 *UP* 어휘 빈도 분포



<그림 3> 영화 *Inside Out* 어휘 빈도 분포



<그림 4> 영화 *Soul* 어휘 빈도 분포



위의 그래프들은, 세 개의 코퍼스 데이터 어휘 빈도수가 모두 비슷한 패턴의 그래프임을 보여준다. 즉, 위의 그래프처럼 애니메이션 영화의 대본 코퍼스 데이터에서도, 소수의 고빈도 어휘와 다수의 저빈도 어휘들의 분포가 마치 알파벳 L자 형태로 눈에 띄게 나누어지는 결과가 나온다. 그러므로 지프의 법칙이 적용될 것이라는 가설에 유의미한 결과가 나타났다. 그리고 그래프들 안에 작은 점 하나가 동떨어져 있는 모습을 볼 수 있다. 좌측 끝 모서리에 모여있는 지점은 소수 몇 개의 어휘들로 해석할 수 있다. 또한, 그 소수의 어휘는 가장 빈도수가 높은 Y축 맨 위쪽에 위치한다는 것을 알 수 있다. 즉, 대본 코퍼스의 어휘의 빈도수는 단 몇 가지 어휘들만 빈도수가 가장 높고, 나머지 다수의 어휘는 빈도수가 비슷하거나 대부분 낮다. 분석 결과, 세 개의 코퍼스 데이터에서 사용된 어휘 빈도 특성이 지프의 법칙으로 유사하며, 서로 데이터의 크기가 차이점이 크게 없음을 알려준다.

4.3 애니메이션 영화 UP 대본 코퍼스 분석

다음은 어린이 애니메이션 영화 *UP* 대본 코퍼스 데이터를 활용한 어휘 분석하였다. 앞으로 다루게 될 애니메이션 영화에서는 아동을 대상으로 제작된 것이므로 고난도 어휘는 상당히 적은 편에 속한다. 본래 대본 코퍼스 데이터에서 가공 단계를 거친 ‘텍스트 파일’을 이용하였다. 고빈도의 어휘를 산출하기 위해 AntConc 프로그램의 Key Word 기능을 사용하였다. 각 애니메이션 대본 코퍼스에서 어휘를 비교 및 분석을 위해, AntConc 프로그램을 이용하여 어휘 리스트 결과를 먼저 엑셀 파일로 옮겼다. 그런 다음, 엑셀 파일의 어휘 리스트를 명사, 동사 그리고 형용

사 각각 따로 분리하였다.

엑셀 파일에서 작업할 때 앞의 <그림 1>에서처럼, 레마 리스트를 참고하여 명사와 동사 그리고 형용사를 나누었다. 레마란 어휘의 원형을 말한다. 예를 들어, *is, are, were* 어휘들의 원형은 *be*가 된다. 따라서, 어휘 *be*는 *is, are, were* 들의 레마 형태가 되는 것이다. 다음은 애니메이션 영화 *UP* 대본 코퍼스의 데이터 결과에서 명사만 뽑아서 고빈도 순으로 20개씩 어휘를 나열하였다. 그 결과는 다음 <표 3>과 같다.

<표 3> 영화 *UP* 대본 코퍼스 명사 빈도순 20 이내

빈도순	명사	빈도수	빈도순	명사	빈도수
1	house	135	11	adventure	24
2	bird	117	12	chair	24
3	dog	98	13	hose	24
4	door	43	14	paradise	24
5	alpha	35	15	cane	23
6	balloon	32	16	room	23
7	explorer	31	17	chocolate	22
8	window	30	18	time	22
9	rock	29	19	wilderness	21
10	master	28	20	baby	19

그 결과, *UP* 애니메이션 영화 이야기 속 상황이 그려지는 명사들이 눈에 띄는 것을 알 수 있다. 주인공 Carl이 엄청난 양의 풍선을 지붕 위에 매달고 집을 이동한다. 그리고 여행을 떠나는 장면이 영화 속 장면과 포스터에서도 나온다. 그 결과, 대본 코퍼스의 명사 어휘 고빈도 리스트에서 *house*와 *balloon*이 포함돼 있음을 알 수 있다. 애니메이션 영화 *UP*의 포스터 그림 이미지와 맞는 어휘를 자주 등장시킨 것으로 볼 수 있다. 그리고 명사 *house* 다음으로 많이 출현한 어휘는 *bird*와 *dog*이다. 즉, 사람 캐릭터가 주인공인 애니메이션 영화이지만, 사람 또는 사물과 관련된

어휘가 아닌 동물이 가장 많이 언급되었다. 그 까닭은 아동 대상 영상물이란 이유로, 사람 캐릭터뿐만 아니라 재미를 주고 친근한 ‘동물’ 캐릭터 등장이가 많은 것으로 해석될 수 있다. 그런데 반대로, 빈도수가 1인 명사 어휘 중 하나가 *zoology*라는 것을 발견할 수 있었다. 그리고 다음 <표 4>는 20개의 동사를 고빈도 순으로 나열한 것이다.

<표 4> 영화 *UP* 대본 코퍼스 동사 빈도순 20 이내

빈도순	동사	빈도수	빈도순	동사	빈도수
1	be	300	11	walk	38
2	look	90	12	like	37
3	get	89	13	pull	35
4	have	88	14	run	35
5	go	84	15	come	34
6	do	59	16	see	32
7	fall	59	17	sit	32
8	will	47	18	stop	31
9	can	40	19	close	29
10	take	38	20	want	29

동사의 경우에는 *be* 동사 다음으로 *look*, *get* 그리고 *have* 순으로 어휘가 자주 출현하였다. 그리고 주제와 내용 안에서 ‘모험’과 연관된다는 점을 비추어보면 다음과 같이 고빈도를 해석할 수 있다. 바로, 고빈도 리스트에서 *walk*, *pull*, *run*, *come*, *sit* 등 유독 행동에 비중을 둔 동사가 자주 등장한다는 점을 알 수 있다. 다음 아래의 <표 5>는 20개의 형용사를 고빈도 순으로 나열한 것이다.

<표 5> 영화 *UP* 대본 코퍼스 형용사 빈도순 20 이내

빈도순	형용사	빈도수	빈도순	형용사	빈도수
1	dirigible	33	11	small	9
2	good	17	12	real	8
3	old	17	13	sorry	7
4	past	17	14	clear	7
5	light	13	15	dead	7
6	south	11	16	elderly	7
7	alone	11	17	large	7
8	long	10	18	little	7
9	big	9	19	massive	7
10	high	9	20	red	7

형용사 고빈도 리스트 경우에는 *dirigible*이 가장 많이 출현하였다. 앞에서 명사 분석에서도 언급하였듯이, 주제와 내용 안에서 모험과 연관이 된 까닭으로 해석될 수 있다. 연이어서, *good*, *old*, *past* 그리고 *light* 순서대로 많이 출현하였다. 그리고, 죽음과 관련된 *dead*라는 형용사와 나이가 연상되는 *old*와 *elderly*가 고빈도 어휘로 등장한다. 이러한 고빈도 형용사들은 주인공 78세 노인 Carl의 캐릭터상 특징과 젊은 날의 과거를 회상하는 이야기로 인한 영향으로 볼 수 있다. 반면에 흥미롭게도, 빈도수가 1인 형용사 어휘들 가운데 *young*이 등장함을 알 수 있었다.

4.4 애니메이션 영화 *Inside Out* 대본 코퍼스 분석

다음은 어린이 애니메이션 영화 *Inside Out* 대본 코퍼스 데이터를 활용한 어휘 분석 결과이다. 위에 4.3 애니메이션 영화 *UP* 대본 코퍼스 분석에서의 절차와 동일하다. 다음은 대본 코퍼스 데이터 결과에서 고빈도 순으로 명사 어휘들만 각각 20개씩 어휘를 나열하였다. 그 결과는 다음 <표 6>과 같다.

<표 6> 영화 *Inside Out* 대본 코퍼스 명사 빈도순 20 이내

빈도순	명사	빈도수	빈도순	명사	빈도수
1	joy	351	11	dream	34
2	sadness	192	12	way	33
3	memory	119	13	family	30
4	island	68	14	day	29
5	anger	67	15	console	28
6	fear	66	16	hockey	27
7	headquarter	65	17	stage	27
8	core	56	18	time	27
9	mom	55	19	guard	24
10	train	39	20	friend	23

그 결과, *Inside Out* 애니메이션 영화 이야기 속 상황들이 그려지는 명사 어휘들을 볼 수 있다. 주인공 소녀 라일리의 감정 중 joy라는 명사가 고빈도 어휘 첫 번째 순위이다. 감정을 나타내는 어휘 중에서도 긍정적인 어휘일수록 많이 사용되었다. 반면에, joy의 정 반대 의미인 sadness가 두 번째로 빈도가 높은 명사이다. 애니메이션 영화 내용에서도 마찬가지로 기쁨(Joy)과 슬픔(Sadness)이 항상 같이 다니는 모습을 보여준다. 즉, 인생에 ‘기쁨’과 ‘슬픔’은 늘

공존한다는 교훈처럼, 명사의 고빈도 리스트에서도 연달아 순서대로 나타나고 있다.

이뿐만 아니라, *anger, fear* 등 감정 상태를 표현하는 어휘가 상당히 많이 등장한다. 앞서 이야기한 것과 마찬가지로, 감정과 관련된 명사 어휘들이 빈도가 높은 결과로 나왔다. 또한, 주제와 내용 부분에서 ‘성장’과 ‘가족’과 관련되어 있고, 한 소녀의 성장 과정을 통해 그려진 이야기이므로 *memory, mom, family, time, friend*라는 명사 어휘들이 고빈도 순위에 포함된다는 점을 알 수 있다. 다음은 코퍼스 데이터 결과에서 동사 어휘들만 고빈도 순으로 각각 20개씩 어휘를 나열하였다. 그 결과는 다음 <표 7>과 같다.

<표 7> 영화 *Inside Out* 대본 코퍼스 동사 빈도순 20 이내

빈도순	동사	빈도수	빈도순	동사	빈도수
1	be	448	11	make	50
2	have	139	12	know	48
3	go	120	13	come	47
4	do	112	14	let	41
5	get	111	15	walk	40
6	look	100	16	wait	36
7	like	65	17	play	34
8	can	64	18	turn	34
9	run	63	19	see	32
10	will	58	20	try	31

고빈도 어휘 동사 순은 *be* 동사 다음으로 *have, go, do* 그리고 *get* 순서대로 많이 출현하였다. 그리고 *look, like, can* 등의 동사 어휘들 볼 수 있다. 동사의 고빈도 어휘들이 전체적으로 긍정적인 의미의 어휘들이 많이 사용되었다. 왜냐하면, 애니메이션 영화 내용에서 소녀 라일리의 관점으로 보여주는 것이 아니라, 라일리의 감정 중 하나인 기쁨이(Joy)의

관점으로 소녀의 이야기를 대신 전달하여 보여주기 때문이라고 해석될 수 있다. 반대로, 빈도수가 1인 명사 어휘 가운데 부정적인 의미로 *disbelief*, *disappointment* 어휘들이 출현하였다. 다음은 코퍼스 데이터 결과에서 고빈도 순으로 형용사 어휘들을 각각 20개씩 어휘를 나열하였다. 그 결과는 다음 <표 8>과 같다.

<표 8> 영화 *Inside Out* 대본 코퍼스 형용사 빈도순 20 이내

빈도순	형용사	빈도수	빈도순	형용사	빈도수
1	new	40	11	bad	14
2	good	33	12	nice	13
3	happy	32	13	close	12
4	long	31	14	dark	12
5	great	30	15	open	12
6	blue	28	16	cool	10
7	subconscious	25	17	alone	8
8	smile	17	18	big	8
9	sure	17	19	past	8
10	sorry	16	20	old	7

형용사의 경우, *new*, *good*, *happy* 순서로 형용사 어휘들이 많이 출현하였다. 한 소녀의 감정을 다룬 애니메이션 영화 내용인데, 고빈도 어휘 리스트를 통해 긍정적인 형용사 어휘들을 훨씬 더 많이 사용하였음을 알 수 있다. 그러므로, 긍정적인 형용사를 더 많이 활용하여 영화 속에서 감정들을 그려냈음을 알 수 있다. 이뿐만 아니라, *long*, *great*, *blue*, *subconscious*, *smile* 순서대로 형용사 어휘들 출현이 많다는 것을 보여준다. 반대로, 빈도수가 1인 형용사 어휘 가운데 부정적인 의미의 *disrespectful*, *dismissive*가 출현하였다.

4.5 애니메이션 영화 *Soul* 대본 코퍼스 분석

마지막으로, 어린이 애니메이션 영화 *Soul* 대본 코퍼스 데이터를 활용한 어휘 분석 결과이다. 위의 애니메이션 영화 *UP*, *Inside Out* 대본 코퍼스 분석에서의 절차와 동일하다. 그리고 코퍼스 데이터 결과에서 명사 어휘들을 고빈도 순으로 각각 20개씩 어휘를 나열하였다. 그 결과는 아래의 <표 9>와 같다.

<표 9> 영화 *Soul* 대본 코퍼스 명사 빈도순 20 이내

빈도순	명사	빈도수	빈도순	명사	빈도수
1	soul	135	11	music	31
2	earth	74	12	mentor	30
3	body	68	13	eye	24
4	cat	46	14	door	23
5	life	42	15	hall	22
6	time	42	16	mom	20
7	counselor	41	17	club	20
8	suit	35	18	gardner	20
9	portal	32	19	piano	19
10	jazz	31	20	seminar	19

그 결과, *Soul* 애니메이션 영화 제목과 똑같은 어휘인 *soul*이 가장 많이 출현한 명사로 나왔다. 명사 *soul*이 고빈도인 이유는, 현재의 삶과 영혼의 삶을 그린 애니메이션 영화 속 이야기의 영향이 크다. 그리고 현재의 삶과 영혼(죽음)의 삶을 다녀온 이야기로부터 알 수 있듯이, 결과에서도 *earth*가 두 번째, *body*가 세 번째로 많이 출현한 고빈도 명사임을 보여준다. 게다가, 명사 *life*와 *time* 어휘를 자주 사용함으로써, 애니메이션 영화 *Soul* 이야기를 통해 인간의 ‘삶’에 대한 문제를 해결해 나가는 교훈이 담겼다는 점을 짐작해 볼 수도 있다. 그리고 빈도수가 1인 명사

어휘 중에서도 *psychology, psychologist*가 출현하였다. 삶은 다른 영화 안에 관련된 고빈도 명사들과 더불어, 심리학과 관련된 어휘도 등장하였음을 알 수가 있다. 다음은 코퍼스 데이터 결과에서 고빈도 순으로 동사 어휘를 각각 20개씩 어휘를 나열하였다. 그 결과는 다음 <표10>과 같다.

<표 10> 영화 *Soul* 대본 코퍼스 동사 빈도순 20 이내

빈도순	동사	빈도수	빈도순	동사	빈도수
1	be	541	11	know	63
2	get	134	12	will	61
3	look	132	13	walk	60
4	have	129	14	try	53
5	go	108	15	run	47
6	do	95	16	take	46
7	can	92	17	play	41
8	like	72	18	lose	35
9	come	67	19	pass	34
10	see	64	20	let	32

동사의 경우에는 *be* 동사 다음으로 *get, look, have* 순서대로 높은 빈도로 출현하였다. 특히, *can, like, will* 동사들이 고빈도 리스트에서 볼 수 있다. 왜냐하면, 애니메이션 영화 내용에서 주인공 Joe가 현실과 죽음의 세상을 경험하게 되면서 Joe가 삶의 의미를 찾아가는 교훈을 주는 것으로 연관을 지을 수 있다. 즉, 이러한 고빈도 어휘들을 통해, 삶에 대한 의지와 가능성을 자주 표출시켰다고 볼 수 있다. 물론, 긍정적인 동사 어휘들이 자주 사용되었다. 다음은 코퍼스 데이터 결과에서 고빈도 순으로 형용사 어휘를 각각 20개씩 어휘를 나열하였다. 그 결과는 다음 <표 11>과 같다.

<표 11> 영화 *Soul* 대본 코퍼스 형용사 빈도순 20 이내

빈도순	형용사	빈도수	빈도순	형용사	빈도수
1	new	59	11	different	11
2	right	49	12	empty	10
3	good	40	13	angry	9
4	great	39	14	full	9
5	astral	16	15	young	9
6	little	14	16	long	8
7	sure	14	17	nice	8
8	bad	12	18	fine	7
9	dark	12	19	happy	7
10	old	12	20	wrong	7

형용사의 경우에는 *new*, *right*, *good* 등 순서로 형용사 어휘들이 많이 출현하였다. 애니메이션 영화 속에서 주로 삶과 관련된 내용의 영향으로 인해, *old*와 *young*처럼 서로 대조적인 어휘가 고빈도 어휘 리스트에 속해 있음을 알 수 있다. 이점은 앞에서 언급했던, *Inside out*의 고빈도 리스트와 비슷한 맥락으로 해석될 수 있다. 그리고 고빈도 어휘 리스트에 대부분 긍정적인 의미의 형용사 어휘가 많이 사용되었음을 알 수 있다.

4.6 명사 사용 빈도 비교분석

다음은 총 세 가지 어린이 애니메이션 영화 *UP*, *Inside Out*, *Soul* 애니메이션 영화 대본 코퍼스 안에서 명사 어휘만을 고빈도 비교분석을 이야기한다. 고빈도 순으로 명사 어휘만 각각 10개씩 나란히 나열하였다. 그 결과는 다음 <표 12>과 같다.

<표 12> *UP, Inside Out, Soul* 대본 코퍼스 명사 빈도순 10 이내

빈도순	<i>UP</i> (빈도수)	<i>Inside Out</i> (빈도수)	<i>Soul</i> (빈도수)
1	house (135)	joy (351)	soul (135)
2	bird (117)	sadness (192)	earth (74)
3	dog (98)	memory (119)	body (68)
4	door (43)	island (68)	cat (46)
5	alpha (35)	anger (67)	life (42)
6	balloon (32)	fear (66)	time (42)
7	explorer (31)	headquarter (65)	counselor (41)
8	window (30)	core (56)	suit (35)
9	rock (29)	mom (55)	portal (32)
10	master (28)	train (39)	jazz (31)

세 가지 애니메이션 영화 대본 코퍼스 데이터를 고빈도 순으로 명사를 나열하였다. 그 결과, 첫 번째로 가장 출현이 많은 명사 어휘들에서 공통점을 찾을 수 있었다. 먼저, 애니메이션 영화 *UP*의 내용에서 고빈도 어휘인 *house*와 *balloon*이 나온다. 그리고 애니메이션 영화 *Inside Out*에서 가장 많이 등장하는 감정이자 주인공인 *joy*라는 어휘가 가장 빈도가 높다. 마지막으로 애니메이션 영화 *Soul*에서 가장 많이 출현한다.

이러한 첫 번째 순위의 고빈도 명사 어휘는 애니메이션 영화의 제목과 동일하거나, 이야기를 떠올리며 관련 이미지를 그려낼 수 있는 어휘에 속하는 명사들이다. 전부 애니메이션 영화 포스터나 내용 안에서 반드시 주인공 또는 교훈과 직결되는 명사 어휘들이 고빈도 순위에서 첫 번째라는 점이다. 그러므로 부정적인 의미의 명사 어휘보다 긍정적인 명사 어휘 사용이 고빈도 리스트에 많이 찾아볼 수 있다.

4.7 동사 사용 빈도 비교분석

다음은 본 연구의 분석 대상인 총 세 가지 어린이 애니메이션 영화 *UP*, *Inside Out*, *Soul*의 고빈도 순으로 동사 어휘만 각 10개씩 나열하였다. 그 결과, 아래의 <표 13>과 같다.

<표 13> *UP*, *Inside Out*, *Soul* 대본 코퍼스 동사 빈도순 10 이내

빈도순	<i>UP</i> (빈도수)	<i>Inside Out</i> (빈도수)	<i>Soul</i> (빈도수)
1	<i>be</i> (300)	<i>be</i> (448)	<i>be</i> (541)
2	<i>look</i> (90)	<i>have</i> (139)	<i>get</i> (134)
3	<i>get</i> (89)	<i>go</i> (120)	<i>look</i> (132)
4	<i>have</i> (88)	<i>do</i> (112)	<i>have</i> (129)
5	<i>go</i> (84)	<i>get</i> (111)	<i>go</i> (108)
6	<i>do</i> (59)	<i>look</i> (100)	<i>do</i> (95)
7	<i>fall</i> (59)	<i>like</i> (65)	<i>can</i> (92)
8	<i>will</i> (47)	<i>can</i> (64)	<i>like</i> (72)
9	<i>can</i> (40)	<i>run</i> (63)	<i>come</i> (67)
10	<i>take</i> (38)	<i>will</i> (58)	<i>see</i> (64)

고빈도 순으로 동사를 나열하였다. 그 결과, 첫 번째로 가장 많은 출현을 한 동사는 *be* 동사이다. 첫 번째, 애니메이션 영화 *UP*의 고빈도 동사 어휘 순서는 *look, get, have* 그리고 *go* 순서대로 출현한다. 두 번째, *Inside Out*에서 *have, go, do* 그리고 *get* 순서대로 출현한다. 마지막으로 *Soul*에서 *get, look, have* 그리고 *go* 순서대로 출현한다. 세 가지 애니메이션 영화의 동사 사용 빈도 공통점은, 동일한 동사 어휘들이 출현한다는 점이다. *have, look, get, go, do*가 10위 안에 반드시 존재한다. 즉, 고빈도 동사 어휘들은 세 편의 애니메이션 영화 대본에서 차이점이 크게 없다고 볼 수 있다. 또한, 긍정적인 의미의 동사 어휘들이 빈도 높게 출현한다는 점을 알 수 있다. 세 가지 애니메이션 영화에서 어려운 동사가 아닌 기본적인 난이도가 쉬운 동사들이 고빈도로 출현되었음을 알 수 있다.

4.8 형용사 사용 빈도 비교분석

다음은 본 연구의 분석 대상인 총 세 가지 어린이 애니메이션 영화 *UP*, *Inside Out*, *Soul*의 고빈도 순으로 형용사 어휘만 각 10개씩 나열하였다. 그 결과, 아래의 <표 14>와 같다.

<표 14> *UP*, *Inside Out*, *Soul* 대본 코퍼스 형용사 빈도순 10 이내

빈도순	<i>UP</i> (빈도수)	<i>Inside Out</i> (빈도수)	<i>Soul</i> (빈도수)
1	<i>dirigible</i> (33)	<i>new</i> (40)	<i>new</i> (59)
2	<i>good</i> (17)	<i>good</i> (33)	<i>right</i> (49)
3	<i>old</i> (17)	<i>happy</i> (32)	<i>good</i> (40)
4	<i>past</i> (17)	<i>long</i> (31)	<i>great</i> (39)
5	<i>light</i> (13)	<i>great</i> (30)	<i>astral</i> (16)
6	<i>south</i> (11)	<i>blue</i> (28)	<i>little</i> (14)
7	<i>alone</i> (11)	<i>subconscious</i> (25)	<i>sure</i> (14)
8	<i>along</i> (10)	<i>smile</i> (17)	<i>bad</i> (12)
9	<i>big</i> (9)	<i>sure</i> (17)	<i>dark</i> (12)
10	<i>high</i> (9)	<i>sorry</i> (16)	<i>old</i> (12)

세 편의 애니메이션 영화의 대본 코퍼스를 고빈도 순으로 형용사를 나열하였다. 그 결과, 첫 번째로 가장 많은 출현이 있는 형용사 어휘에 공통점을 찾을 수 있었다. 앞에서 애니메이션 영화 *UP*의 내용에서 중요한 역할을 하는 명사 어휘 *house*가 고빈도 리스트에 나온다. 모험을 통해 이동과 여행을 그리는 내용에 걸맞게 형용사 어휘 *dirigible*이 고빈도 리스트의 첫 번째 순서이다. 두 번째로, *Inside Out*에서 가장 많이 등장하는 고빈도 어휘는 *new* 그리고 *good*이다. 즉, 주인공 소녀 라일리의 성장 과정을 보여주는 이야기이므로 항상 새로운 것을 접하게 되는 것이기 때문이다. 마지막으로, *Soul*에서 많이 출현하는 형용사 어휘는 *new, right, good, great* 순이다. 이 어휘들 또한 마치, 위의 라일리처럼, Joe도 역시 새로운 모험에 접하게 된다. 현재의 삶이 아니라 영혼의 삶을 경험하게 되는 이야기이므로 모든 것이 새로운 것이기 때문이다. 세 편의 애니메이션 영화 대본 코퍼스에서 형용사 고빈도를 통해 알 수 있는 공통점은 각각 애니메이션 영화 대본 코퍼스에서 고빈도 명사와 관련된 형용사가 사용되었다는 점이다. 또한, 형용사 고빈도 리스트에서 부정적인 의미보다 긍정적인 의미의 형용사가 더 많이 포함되어 있음을 볼 수 있다.

제5장 결론

본 연구에서는 총 세 편의 미국 아동 애니메이션 영화 *UP*, *Inside Out*, *Soul*의 대본 코퍼스 데이터를 활용하여 어휘 빈도를 분석하였다. 본 논문의 제2장에서는 애니메이션에 코퍼스 연구를 활용한 선행 연구들을 살펴보았다. 제3장에서는 본 연구에서 분석 자료인 세 편의 애니메이션 영화에 대한 설명과 줄거리를 이야기하고, 코퍼스 데이터 분석을 위해 사용한 AntConc 3.5.8과 R 프로그램을 설명하였다. 마지막으로 제4장은 코퍼스 데이터 분석 결과에 대해 다루었다. 각 대본 코퍼스 데이터를 고빈도 순서 20개씩 명사, 동사, 형용사로 나누어 고빈도 어휘들에 대해 분석하였다. 또한, 지프의 법칙에 적용되는지 빈도수-빈도순 그래프를 통해 결과를 알아보았다. 먼저, 애니메이션 영화 대본 코퍼스마다 타입과 토큰부터 비교하여 결과를 제시하였다. 반대로, 빈도수가 1회인 저빈도 어휘도 같이 살펴보았다. 앞서 설명한 바와 같이, 타입의 빈도가 높을수록 다양한 어휘가 사용되었음을 뜻한다. 토큰의 빈도는 어휘의 개수이므로 토큰 빈도가 높을수록 텍스트의 양이 많은 것을 뜻한다. 그 결과는 아래와 같다.

첫째, 애니메이션 영화 *UP*에서 타입의 수는 2394이며, 토큰의 개수는 13939이다. 두 번째로, *Inside Out*의 타입의 수는 2962이며, 토큰의 수는 16753이다. 마지막으로, *Soul*의 타입의 수는 2093이며, 토큰의 수가 19120이다. 그리고 각 애니메이션 영화 대본 코퍼스의 TTR 계산 결과 순서는 다음과 같다. 텍스트의 크기가 클수록 비율이 낮아진다. 그러므로 본 연구에서 활용된 세 편의 애니메이션 영화 대본 코퍼스의 규모 크기는 *Soul* > *UP* > *Inside Out* 순서로 컸다. 그러나 실질적인 코퍼스 데이터 분량의 크기 차이는 많은 편이 아니므로 어휘 비교 대상으로 적절하였다.

둘째, 전부 애니메이션 영화 포스터나 내용 안에서 반드시 주인공 또는 교훈과 직결되는 명사 어휘들이 고빈도 순위에서 첫 번째라는 점을 발견하였다. 전부 동일한 주제인 가족, 성장 그리고 모험으로 정하였다. 이러한 첫 번째 순위의 고빈도 명사 어휘는 애니메이션 영화의 제목과 동일하거나, 이야기를 떠올리며 관련 이미지를 그려낼 수 있는 어휘에 속하는 명사들이다. 또한, 고빈도 명사들은 긍정적 의미의 어휘가 다수인 반면에, 빈도수가 1인 만큼 부정적 의미의 명사 어휘들은 출현이 낮음을 알 수 있었다.

셋째, 세 가지 애니메이션 영화의 동사 사용 빈도 공통점은, 동일한 동사 어휘들이 출현한다는 점이다. 예를 들어, *have, look, get, go, do*의 동사 어휘들이 10위 안에 반드시 존재한다. 즉, 고빈도 ‘동사’ 어휘들은 세 편의 애니메이션 영화 대본에서 차이점이 크게 없다고 볼 수 있다. 또한, 긍정적인 의미의 동사 어휘들이 빈도 높게 출현한다는 점을 알 수 있다. 세 가지 애니메이션 영화에서 어려운 동사가 아닌 기본적인 난이도가 쉬운 동사들이 고빈도로 출현 되었음을 알 수 있다.

넷째, 본 연구를 통해 흥미로웠던 점 가운데 하나는 아동을 대상으로 제작된 애니메이션 영화 대본 코퍼스는 세 편의 애니메이션에 공통 단어의 고빈도가 서로 겹친다는 점이다. 그러므로 자주 반복되고, 난이도가 쉬운 어휘들로 구성되었다는 점이다. 그리고 아동 애니메이션 영화를 굳이 미리 시청하지 않아도, 고빈도 어휘 ‘명사’ 어휘들은 쉽게, 유추가 가능하다는 점이다. 이러한 면을 통해 영어교육에서 활용될 수 있는 방안이 충분히 많을 것으로 기대해 볼 수 있다.

다섯째, 고빈도 형용사 어휘들은 고빈도 명사 어휘들을 통해 영향을 받는다는 점이다. 고빈도 명사 어휘 리스트만 보고, 관련된 형용사들을 떠올려 볼 수 있다. 고빈도 동사 어휘들은 서로 다른 어휘들이 아니

다. 즉, 복잡하고 긴 어휘가 아닌, 기본적인 동사 어휘들을 활용하여 애니메이션 대본들이 작성되었다. 이러한 면 또한, 영어교육에 충분히 교육 자료로써 활용될 수 있다고 생각된다.

마지막으로, 영어학 분야의 양적 연구 논문들 가운데 ‘애니메이션 영화’ 연구 분야는 활발한 편이 아니다. 그러나 향후 연구들에 조금이나마 보탬이 될 수 있을 것이라는 생각을 해 볼 수 있다. 반대로, 코퍼스 언어학과 애니메이션 영화 둘 다 해당하는 선행 연구들을 찾기에 풍족하지 않아서 아쉬운 점도 있었다. 그 결과, 본 연구에서는 분석 대상의 분량이 적은 편이다. 그러므로 단순히 아동 애니메이션 영화의 어휘 빈도 분석의 대표적 특징이라고 단정 짓기에는 매우 부족하다. 그러나 이렇게 애니메이션 영화 대본 코퍼스 데이터를 활용한 통계처리는 물론이며, 코퍼스 데이터 연구 분야에 도전한 것은 다음 연구를 위해 기여할 수 있다고 생각한다.

참고문헌

- 강경아, 이제영, 황치복. (2018). 애니메이션 영화를 활용한 초등영어 교수·학습 모형의 개발 및 적용. 영어영문학연구, 44(3), 197-219.
- 강범모. (2011). 「언어 컴퓨터 코퍼스 언어학」. 서울: 고려대학교출판부
- 고광윤. (2019). 「The AntConc Book 코퍼스 분석과 활용 완벽 가이드」. 경기도: LARRABEE LEARNING
- 권혁승. (2008). 코퍼스 언어학의 실제 및 응용. 응용언어학, 24(3), 1-30.
- 김성식. (2018). 코퍼스 언어학의 과제와 전망. 영어교육연구, 30(1), 47-71.
- 김윤서. (2016). 영어 그림책의 코퍼스 기반 어휘수준 분석. 연세대학교 교육대학원 석사학위 논문.
- 김은아. (2013). 영어 애니메이션 영화의 어휘수준에 관한 코퍼스 기반 분석. 연세대학교 교육대학원 석사학위 논문.
- 김은현. (2010). 코퍼스 기반 어린이 판타지영화에 나타난 어휘 분석. 청주교육대학교 교육대학원 석사학위 논문.
- 김재원. (2008). 애니메이션과 초현실주의 회화에서 연계된 메타모포시스에 관한 연구. 한국디자인포럼, 19(0), 93-105.
- 김지수. (2012). 동. 서양 사상과 문화에 따른 애니메이션 특징 분석. 한국컴퓨터게임학회논문지, 25(4), 63-72.
- 김진해. (2006). 코퍼스언어학적 관점에서 본 의미의 본질. 한국어 의미학, 21, 75-104.

- 박서영. (2019). 영한 더빙텍스트에 나타난 담화표지 번역 연구: 애니메이션 코퍼스 분석을 중심으로. 한국외국어대학교 대학원 석사학위 논문.
- 성민창, 김기택. (2021). 2015 개정 영어과 교육과정 초등 수준 어휘 목록을 활용한 영어 애니메이션 영화의 어휘 사용 분석. *영어학*, 21, 87-109.
- 이가람. (2012). 시리즈 영어 애니메이션 어휘 수준에 관한 코퍼스 기반 분석. 연세대학교 교육대학원 석사학위 논문.
- 정소우. (2014). 영어 교육 논문 코퍼스에 나타난 학문 어휘 분석. *언어*, 39(4), 997-1019.
- 정준호. (2017). 애니메이션이란 무엇인가?-애니메이션의 매체적 특성과 정의에 관한 연구. *PREVIEW: 디지털영상학술지*, 14(1), 63-86.
- Biber, D., Conrad, S., and Reppen, R. (2015). 바이버의 코퍼스 언어학 (Corpus linguistics). (유석훈, 김유영 공역). 고려대학교 출판문화원. (원저 1998 출간).
- Clauset, A., Shalizi, C. R., and Newman, M. E. (2009). Power-law distributions in empirical data. *SIAM Review*, 51(4), 661-703.
- Gorčević, A. R., Dazdarević, S. N., and Zoranić, A. L. (2021). Dysphemisms in animated films. *Journal of Language and Literary Studies*, 12(35), 175-194.
- Newman, M. E. (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics*, 46(5), 323-351.
- Zipf, G. (1949). *Human Behavior and the Principle of Least Effort*. New York: Addison-Wesley.

<참고 웹사이트>

영어 애니메이션 영화 대본 : <https://www.kci.go/>

픽사(Pixar) 공식 홈페이지 : <https://www.pixar.com/>

Antconc : <https://www.laurenceanthony.net/software/antconc/>

R download : <https://posit.co/download/rstudio-desktop/>

국 문 초 록

미국 아동 애니메이션 영화 대본에 나타난 영어 어휘들의 빈도 분석

김진하

영어영문학과 영어학전공

중앙대학교 대학원

본 연구의 목적은 총 세 편의 미국 아동 애니메이션 영화 *UP*, *Inside Out*, *Soul*의 대본 코퍼스 데이터를 활용하여 어휘 빈도를 활용하여 비교 및 분석하는 것이다. 본 연구에서는 대본 코퍼스 분석을 위해 AntConc 3.5.8와 R 프로그램을 사용하였다. 먼저, 각 대본 코퍼스 데이터를 고빈도 순서 20개씩 명사, 동사, 형용사로 나누어 고빈도 어휘 목록을 알아보았다. 또한, 지프의 법칙에 적용되는지 빈도수-빈도순 그래프를 통해 결과를 알아보았다.

그 결과는 다음과 같다. 첫째, 애니메이션 영화 대본 코퍼스 데이터가 지프의 법칙에도 적용되었다. 둘째, 애니메이션 영화의 주제와 내용에 의해 고빈도 어휘에도 영향을 준다는 것을 발견하였다. 셋째, 아동을 대상으로 제작된 애니메이션 영화 대본 코퍼스는 공통 단어의 고빈도가 서로 겹친다는 점을 발견하였다. 넷째, 애니메이션 영화를 굳이 미리 시청하지 않아도, 고빈도 어휘 명사 어휘들은 유추가 쉽게 가능하다는 점을 알 수 있었다. 이러한 면을 통해 영어교육에서 활용될 수 있는 방안이 충분히 많을 것으로 기대해 볼 수 있다.

마지막으로, 본 연구는 영어학과 언어학 분야에 있어서 애니메이션 영화 대본 코퍼스 데이터를 활용하였다. 그 점에서 통계 처리는 물론이며, 애니메이션 분야를 코퍼스 데이터 연구 분야에 도전한 점은 다음 연구를 위해서 기여할 수 있을 것으로 본다.

핵심어: 코퍼스, 지프의 법칙, 타입-토큰 비율, 영어학, 어휘 빈도,
아동 애니메이션 영화 대본

ABSTRACT

An Analysis of English Word Frequency in the American Children's Animation Movie Scripts

Jinha Kim

Major in English Linguistics

Department of English Language and Literature

The Graduate School

Chung-Ang University

The purpose of this study is to compare and analyze using word frequencies using script corpus data from three American children's animation films *UP*, *Inside Out*, and *Soul*. In this study, AntConc 3.5.8 and R programs were used for script corpus analysis. First, each script corpus data was divided into 20 high-frequency orders into nouns, verbs, and adjectives to find out the high-frequency word list. In addition, Zipf's law was applied. Therefore the results were examined through frequency-frequency ordered graphs.

The summary of the results is as follows. First, the animation film script corpus data was also applied to Zipf's law. Second, it was found that the high frequency affects the word by the theme and content of the animation film. Third, it was found that the frequency of common words overlapped with each other in the animated film corpus produced for children. Fourth, it was found that high-frequency word noun word is easy to infer without having to watch animated movies in advance. Though this

aspect, it can be expected that there will be enough methods that can be used in English education.

Finally, this study used the animation film script corpus data in the fields of English and linguistics. In that respect, I believe that challenging not only statistical processing but also the field of animation to the field of corpus data research can contribute to the next research.

Keywords: corpus, Zipf's law, type-token ratio, English linguistics, word frequency, Children's Animation Movie Script

감사의 글

2년이라는 시간이 무색할 만큼 이제야 조금 알 것 같은데, 벌써 졸업을 앞두고 되었습니다. 가장 먼저, 석사 논문을 완성하기까지 항상 함께해주신 지도교수님께 감사의 말씀을 전해드리고 싶습니다. 매주 논문지도를 통해 많이 부족한 저를 끝까지 이끌어주신 김선희 교수님께 감사드립니다. 이뿐만 아니라, 항상 저희를 생각해주시고 아껴주신 지도교수님의 은혜를 잊지 않겠습니다. 그리고 논문 심사를 통해서 꼼꼼한 피드백과 아낌없는 조언을 해주신 최영진 교수님과 이성호 교수님께 감사의 말씀을 전해드리고 싶습니다.

그동안 중앙대 영어영문학과 교수님들께 교육 조교로 근무하면서 들었던 존경심과 감사한 마음을 한 분 한 분께 전해드리고 싶습니다. 그리고 학과 행사에 힘쓰시고 조교들을 항상 챙겨주셨던 학과장님이신 장성진 교수님께 감사의 말씀 전해드리고 싶습니다. 또한, 대학원 생활 동안 저에게 가장 힘이 되어주고, 학업에 힘든 순간이 있을 때마다 함께 이겨내며 서로 의지를 할 수 있었던 동기 이민영 선생님에게 감사한 마음을 전합니다.

마지막으로, 조금이나마 힘들지 말라고 직접 운전해서 학교에 태워주시고 항상 지지해 주시는 어머님께 존경하고 사랑한다고 전하고 싶습니다. 감사합니다.

2022년 12월

302관 서라벌홀

영어영문학과 사무실에서

김진하 드림