

분과 **3**

디지털 환경에서의 한국어문화교육

사회 : 오지혜(세명대학교)

말하기 수업에서의 챗봇 빌더 활용 방안

엄정윤(영남대학교)

1. 서론

한국어 교육자가 말하기 수업을 설계할 때 중요하게 생각하는 것은 학습자들의 성공적인 대화 경험이다. 그래서 대부분 말하기 수업을 위한 교재에는 배운 문법을 활용한 대화문을 제시하고 그 대화문을 동료 학습자와 연습하는 과정이 들어가게 된다.

한국어 학습자들 역시 말하기 수업에서 성공적인 대화 경험을 중요하게 생각하지만 한편으로 자신의 말하기 오류 수정에 대한 기대도 있다. 김영주(2011)는 학습자는 말하기를 자신의 목표 언어 능력을 점검하고 수정하는 주요한 목표어 습득 활동으로 보며 특히 학습자들은 원어민 교사의 명시적인 오류 수정을 선호한다고 하였다.

그러나 대화 경험은 대화 상대자의 목표어 수준에 영향을 받을 수 있다. 일반적인 교육 현장에서는 교실 내에서 유일한 원어민인 교사가 모든 학습자와 개별적으로 대화 연습을 하기는 어렵고 동료 학습자와 서로 짝을 지어 연습하는데 이때 상대 동료 학습자의 목표어 숙달 수준에 따라 대화 경험의 만족도는 달라질 수 있다. 또한 동료 학습자와의 대화 연습을 교사가 처음부터 끝까지 관찰하며 오류 수정을 해 주기가 어렵고 한 명의 교사가 모든 그룹을 돌아보며 관찰하는 것이 쉽지 않기 때문에 명시적 피드백에도 한계가 있다.

팬데믹 시대가 길어지면서 이러한 한계는 뚜렷하게 나타난다. 온라인 및 비대면 수업에서 학습자들이 소그룹으로 나누어 대화 연습을 하더라도 교사는 학습자들의 연습 장면을 계속 볼 수 없기 때문에 적절한 피드백을 주기 어렵고 한국어 교육기관을 통하지 않고 혼자 공부하는 학습자에게는 대화상대와 피드백을 해 주는 동료 학습자 및 교사가 없으므로 말하기 연습은 머릿속으로 이루어지거나 혼자 글을 읽어 내는 것 외에는 별다른 방법이 없다.

본 연구는 이렇게 변화하는 학습 상황에서 더 분명한 기존 교실 상황의 한계를 보완해 줄 교육 방법이나 도구가 필요하다고 보고 교육 도구로서 챗봇의 활용 방안을 모색해 보고자 한다.

챗봇은 채팅을 하는 로봇이라는 뜻으로 인간의 음성이나 문자가 입력이 되면 미리 설계한 대답 및 대화가 가능한 프로그램이다. 이미 영어 교육에서는 실제 수업 현장에서 챗봇을 사용해 개인 맞춤형 교육의 가능성을 증명하고 학습에도 효과적이라는 연구 결과들이 많다¹⁾. 한국에서도 챗봇 활용 방안에 대한 연구가 진행되고 있으나 아직까지 영어 교육에 비해 관련 논의가 많이 이루어지지 않고 있다.

본 연구에서는 말하기 수업의 단계인 대화 연습을 도와 줄 학습 도구로서의 챗봇을 만들어 학습자들이 참여하는 양상을 보고자 한다. 그리고 챗봇 활용이 효과적인 대화 연습의 기능과 교사의 피드백을 보완할 수 있는지를 살펴보고 대화 연습과 피드백 기능을 중점으로 말하기 수업에서 챗봇 활용의 교육적 효과를 논의하고자 한다.

1) 한다운(2020), 유영진(2021), 윤광진(2021), 이은비(2021), 유지연·유훈식(2021)

2. 말하기 수업과 챗봇

2.1. 말하기 수업 활동의 변화

외국어 교육의 말하기는 목표어가 가지고 있는 사회·문화적 배경에 대한 이해를 전제한다. 또 유창성과 정확성, 대화의 맥락을 이해하는 담화 능력도 요구된다. 이를 위하여 외국어 교육 학계와 현장에서는 말하기를 효과적으로 교육하고자 하는 시도가 계속되어 왔다.

한국어 교육에서도 말하기 기능 교육의 중요성을 인식하며 많은 연구가 진행되어 왔다. 특히 한국어 기능 연구사를 살핀 김중섭(2015)²⁾에 따르면 142편의 기능 교육 연구 논문에서 듣기, 읽기, 쓰기 관련 논문을 모두 합쳐 53%인데 비해 말하기 관련 논문이 47%를 차지해 다른 기능에 비해 활발히 연구되었다는 것을 알 수 있다.

한국어 교육에서 말하기교육 연구는 길지 않은 역사 속에서 끊임없이 생동하며 발전해 왔다. 이민경(2010)은 말하기교육 연구의 시대적 흐름에 따라 1980년대를 ‘말하기교육의 초기 연구’의 시대, 1990년대를 ‘다양한 주제의 말하기교육 연구의 출발’의 시대, 2000년 이후를 ‘말하기교육 연구의 양적·질적인 증가’의 시대로 나누었다.

특히 2010년대 전후는 그동안 쌓아온 연구들을 자양분 삼아 실질적인 교수방법과 모형들의 구체적 적용 사례들이 많이 등장하였다. 학습자 중심의 수업 활동이 활발히 연구되었고 토론, 짝활동, 스토리텔링, 플립드 러닝, 과정극 등으로 상호작용과 의미 중심의 교수법 연구도 활발히 이루어졌다. 이와 함께 교육 자료로서 매체를 활용한 말하기 활동에 대한 연구도 크게 증가했는데 드라마, 광고, 비디오 같은 영상 매체와 동화나 동요, 모바일을 활용한 말하기 활동 등이 있다³⁾. 이러한 교실 활동들은 학습자의 말하기 부담감을 크게 감소시켰고 학습자가 말하기 수업에 능동적으로 참여할 수 있어 학습 효과에도 긍정적인 변화를 가져왔다. 또 말하기의 학습 목표가 단순한 의사소통에서 상호작용을 거쳐 의미협상으로 변화했다. 즉, 이제 한국어 교육에서 말하기는 학습자가 적극적으로 주체적으로 상호작용을 하며 다양한 상황 속에서 유의미한 활동을 통해 의미협상을 하는 훌륭한 매개로 여겨지고 있다.

그러나 2020년대는 교육계의 내·외부적인 상황에 의해 급격한 변화를 겪고 있어 교육 현장과 학계에서는 학습 효과와 학습 지속성을 위해 더욱 과감하고 다양한 연구와 논의가 진행되고 있다. 한국어 교육이 겪고 있는 변화는 다음과 같다.

첫 번째, 학습자의 변화이다. 글보다는 영상이, 주어진 콘텐츠를 소비하기 보다는 참여하는 것이 더 익숙하고 편한 학습자가 증가하고 있다. 또 한국어 학습의 동기와 목적에서도 취미나 여가 등이 차지하는 비중이 커지고 있다. 말하기 수업에서 교재의 활자를 통해 지문을 읽고 연습하는 것 이상의 흥미 요소가 필요한 이유이다.

두 번째 변화는 수업 형태의 변화이다. 온라인 및 비대면 수업, 가상 공간에서의 수업, 개별 학습자의 증가 등으로 교실이라는 전통적이고 물리적인 환경에서 벗어나 점점 다양하고 복잡한 환경에서 교육이 이루어지고 있다. 이러한 수업 형태의 변화는 시대적 필요에 의한 변화이기는 하지만 교실 말하기 수업에 비해 또래 집단과의 상호작용을 어렵게 하고 목표어의 노출에도 소홀해질 수밖에 없다.

한상미(2011)는 말하기 교육에 최고의 방법을 판단하는 대신 학습자의 요구나 학습 목표를 바탕으로 현장의 교사나 교육 기관에서 학습자의 모국어, 연령, 학습 동기, 숙달도 수준 등의 다양한 변인을 고려하여 최선의 방법을 선택할 필요

2) 국제한국어교육학회의 학술지 <한국어 교육>의 1호(1989)부터 26-1호(2015)까지 게재된 142편의 기능 교육 연구 논문을 분류하였는데 그 중 말하기 관련 연구는 67편이었다.

3) 김정은(2014), 권세희(2010), 김재희(2017), 신정아(2017), 이경·윤영(2017), 정진(2018), 이정화·박혜옥(2018), 서수진(2019), 최유정·김대희(2019), 김지현(2020), 신윤아(2021), 이은비(2021), 유영진(2021), 윤광진(2021), 한다운(2020)

가 있다는 점을 강조했다. 기존의 연구와 협력할 수 있고, 학습자들의 학습 동기를 고취시키면서 달라진 상황과 조건에 맞는 최선의 교육 방식과 매개에 관심을 가지고 그 교육적 효과를 논의하고 검증하는 과정이 필요하다.

2.2 말하기 교육 도구로서의 챗봇

비대면 수업이나 온라인 수업에서 기존의 교실 말하기 활동의 상호작용 효과를 대신하기 위해 여러 가지 연구를 시도하고 있다. 모바일을 활용한 말하기 과제 수행에 대한 연구(김지현·성인경, 2021), 음성합성기술인 TTS를 활용한 발표하기 사례연구(박진철, 2021), 가상현실(VR) 매체를 활용한 여성결혼이민자 대상 한국어 교육 방안(심은지·유훈식, 2019), 학습자 인식을 중심으로 한 챗봇을 통한 스마트러닝 한국어교육의 학습 효과(유지연·유훈식, 2021) 등의 연구가 있다. 여러 도구를 활용한 교육 중 특히 챗봇은 다수의 장점을 가지고 새로운 교육 방법으로 각광받고 있다. 챗봇의 장점은 다음과 같다.

첫째, 입력과 출력이 모두 가능하고 시간과 장소의 제약을 받지 않는다.

둘째, 학습자 개인이 컴퓨터나 스마트폰으로 쉽게 접근할 수 있다.

셋째, 교실 수업 외에서도 교사의 1:1 피드백이 용이하다.

넷째, 학습자들의 말하기 불안을 감소시킬 수 있다.

한국어 교육에서도 챗봇의 교육적 효과를 입증하는 연구가 있는데 박정아·이향(2020)은 챗봇이 한국어 교육에서 비대면 수업 및 차세대 온라인 수업의 수업 도구로 활용될 수 있다고 보고, 한국어 학습에서의 구체적인 챗봇 활용 방안과 코딩이나 프로그래밍에 관한 지식이 없는 교사가 챗봇을 제작할 수 있는 대표적인 챗봇 빌더들의 기능과 특성을 함께 제시하였다.

챗봇은 충분히 수업 설계의 변인으로 자리잡을 수 있으며 학습의 개인화가 지향되는 상황에서 새로운 대안으로 모색해 볼 필요가 있다. 그러나 아직은 챗봇의 장점과 소개에 대한 연구에 그쳐 실제 교육 효과를 검증하는 연구는 드물다. 챗봇의 교육적 효과를 입증하기 위해서는 다양한 환경과 방법으로 연구할 필요가 있다.

본 연구에서는 학습자들의 챗봇 활용이 교실 말하기 수업의 대안으로 활용될 수 있는지를 알아보기 위해 말하기 수업 내용 중 대화 연습을 챗봇으로 만들어 학습자들에게 과제로 제시하고 그 교육적 효과를 논의하고자 한다.

3. 연구 방법

3.1 연구 참여자

말하기 수업에서 챗봇 활용 방안을 보기 위해 본연구 시작 전에 교사가 챗봇 빌더를 공부하고 여러 차례에 걸쳐 연습하는 준비 단계를 가졌다. 본연구에서는 교사가 챗봇 빌더를 활용해 연구에 쓰일 대화문을 설계하는 것과 학습자들의 실제 챗봇 대화 연습이 있었다. 한국어 교육 기관 3군데에서 각각 두 명씩 자원하여 총 6명이 참가했다. 다양한 환경에서 챗봇이 사용되는 양상을 보고 싶어서 세 집단을 선정했다. 교육 기관에 따라 연구 참여자 집단을 A집단, B집단, C집단으로 구분하였다. A 집단은 대면 수업이고 B, C 집단은 비대면 수업이다. 또 A, B집단은 학문 목적인 반면 C집단의 수업은 일반 목적이다.

A집단의 학생 두 명(미얀마 1명, 일본 1명)은 학문 목적의 한국어 학습자로 D기관에서 현재 『서울대 한국어』 4B로 공

부하고 있다. 대면 수업 중 연구자에게 직접 챗봇 설명을 듣고 귀가 후 각자의 공간에서 참여했다. 이 집단의 학생들은 서울대 한국어 4단계를 이수하자마자 대학원 진학할 예정이다. B집단의 학생 두 명은 Y기관에서 『서울대 한국어』 3B로 한국어를 공부하고 있으며(베트남, 2명) 현재 ZOOM수업 중이어서 교사에게 비대면으로 설명을 듣고 수업 시간에 실험에 참여했다. 이 학생들은 3단계를 이수 후에도 기관에서 다음 단계를 수강할 예정이다. C집단(인도네시아, 2명)은 일반 목적의 학습자들이다. 현재 공무원 연수 프로그램을 통해 한국어를 공부하고 있으며 교재는 『사랑해요 한국어2』를 사용하고 있다. 인도네시아에 거주 중이며 ZOOM으로 한국어 수업을 듣고 있다. 수업 시간에 담당 교사에게 설명을 듣고 참여 했다. 이를 정리하면 다음과 같다.

〈표 1〉 연구 참여자 정보

	성별	연령	국적	거주국가	한국어 학습 단계	수업형태	학습목적
A1	여	20대	미얀마	한국	서울대한국어4B	대면수업	학문목적
A2	여	20대	일본	한국	서울대한국어4B	대면수업	학문목적
B1	여	20대	베트남	한국	서울대한국어3B	비대면수업(ZOOM)	학문 및 취업목적
B2	여	20대	중국	한국	서울대한국어3B	비대면수업(ZOOM)	학문목적
C1	여	30대	인도네시아	인도네시아	사랑해요 한국어2	비대면수업(ZOOM)	일반목적
C2	여	30대	인도네시아	인도네시아	사랑해요 한국어2	비대면수업(ZOOM)	일반목적

3.2 연구 도구

본 연구는 챗봇을 활용한 말하기 수업 방안에 대한 것으로 연구의 시작을 챗봇 설계 과정부터 두었다. 코딩이나 프로그래밍 등에 전문적이지 않은 교사가 직접 챗봇을 설계하기 위해서는 챗봇 빌더가 필요하다. 챗봇 빌더란 챗봇 시스템을 구성하는 복잡한 기능들은 제공하고 이용자는 활용 목적과 활용할 영역에 맞춰 간단한 기능이나 명령의 추가만하면 챗봇을 생성할 수 있도록 플랫폼을 말한다. 챗봇 빌더는 종류에 따라 제공하는 기본 템플릿이나 형태 등이 다양하다. 본 연구에서 챗봇 빌더로 구글에서 개발한 다이얼로그 플로우(Dialogflow ES)를 선정했으며 선정 기준은 다음과 같다.

첫째, 교사의 설계 용이성을 위해 한국어를 지원해야 하며 생성 작업이 직관적이고 쉬워야 한다

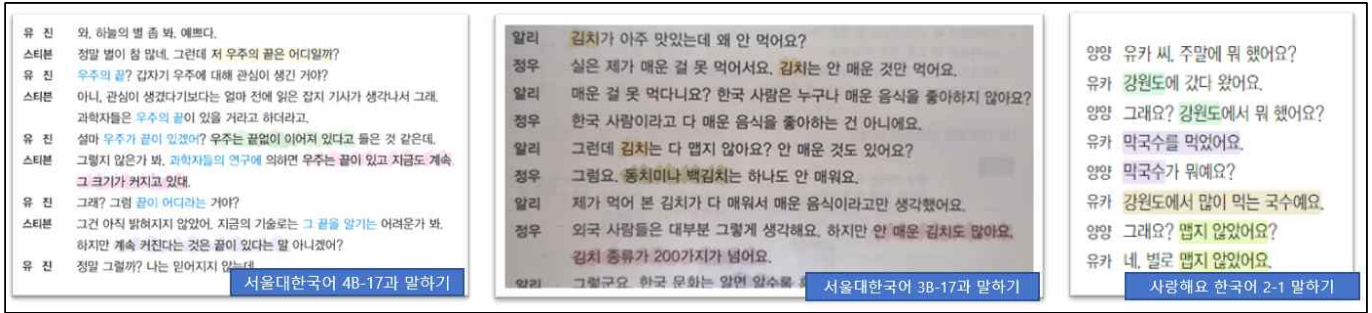
둘째, 학생의 사용 편리성을 국내 뿐 아니라 국외에서도 접근 가능성을 고려해야 한다.

셋째, 대화 연습의 기능을 수행하여야 하므로 음성 입력과 출력이 되어야 한다.

마지막으로 피드백 기능을 위해 학습자가 자신의 발화를 눈으로 확인할 수 있고 교사도 학생의 발화를 사후 점검할 수 있어야 한다.

박정아·이향(2021)에 따르면 다이얼로그 플로우는 텍스트 봇과 보이스봇을 위한 단일 통합 플랫폼으로 개발자가 예상 대화 시나리오를 입력해 놓으면 챗봇은 사용자에게 텍스트, 음성, 시각적인 응답을 할 수 있다고 하였다. 특히 본 연구자는 다양한 메시지 플랫폼으로 연동할 수 있다는 점과 google에서 개발하고 무료라서 접근이 용이하다는 점, 교사가 비교적 설계하기 쉽다는 점에 주목하여 본 연구를 위한 챗봇 빌더로 선정하였다.

집단별 대화 연습문은 학습자들의 진도에 맞춰 교재의 대화 연습문으로 생성했으며 A집단은 『서울대 한국어』 4B의 16과 말하기, B집단은 『서울대 한국어』 3B 17과 말하기, C집단은 『사랑해요 한국어』 2, 3과의 대화 연습문을 사용했다.



〈그림 1〉 집단별 사용한 교재 대화 연습문

대화 연습은 거의 모든 한국어교재에 있을 정도로 중요한 기능 연습 중 하나이다. 이민아 외(2015)는 대화 연습은 학습자들에게 언어 기능을 수행하는 활동인 동시에 문화적 요소가 반영된 것이므로 학습자가 실제 상황과 그 담화적 의미를 이해하고 생성할 수 있기 때문에 중요한 활동이라고 보았다. 또 이미향·이윤주(2018)는 교재의 대화문은 그 사회의 대표성을 부여받으므로 언어 교재에서 대화 연구에 집중하여야 한다고 보았다.

대화 연습은 문법과 어휘처럼 혼자 공부할 수 없고 쓰거나 읽기와는 달리 장소에도 영향을 받는다. 함께 대화를 주고받는 동료나 피드백을 제공해 줄 교사가 없으면 효과를 알기 어렵기 때문에 교실 수업에 의존적인 활동이라고 할 수 있다. 그렇기 때문에 비대면과 개별학습에서 챗봇이 대화 연습의 기능을 수행할 수 있는지와 적절한 피드백을 제공할 수 있는지를 알아 볼 것이다. 본 연구에서는 챗봇 설계도 연구의 한 부분으로 보고 있기 때문에 챗봇 설계 과정은 4장에서 자세히 소개하려고 한다.

3.3 연구 절차

먼저 예비연구 단계에서 교사가 챗봇 빌더를 선정하고 교재의 대화문을 챗봇으로 구현하기 위해 연구자가 직접 설계하는 것을 연습하였다. 설계 후 연구자가 직접 사용해 보는 과정을 반복하며 설계 오류를 수정한 후 연구 참여자 집단을 정하였다. 연구자는 챗봇 빌더로 연구 참여자들의 학습 진도에 맞는 대화 연습용 챗봇을 만든 뒤 학습자들이 참여하도록 했으며 챗봇 빌더에 자동으로 기록된 결과를 통하여 학생들의 사용 양상을 분석해 보았다.

연구 후 참여자들을 대상으로 챗봇 대화 연습의 참여 소감에 대해 간단한 설문을 실행했다. 이를 정리하면 다음과 같다.

1. 예비 연구

- 챗봇 빌더 선정
- 챗봇 설계 연습

2. 연구 참여자 선정

3. 본 연구

- 챗봇 대화문 설계
- 챗봇 사용의 실제
- 설문 조사

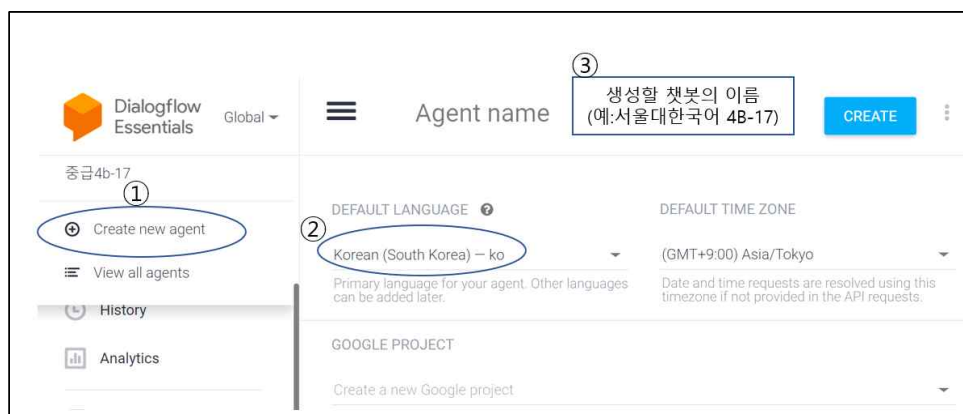
4. 자료 분석

〈그림 2〉 연구 절차

4. 연구 결과

4.1 챗봇 설계 및 활용의 실제

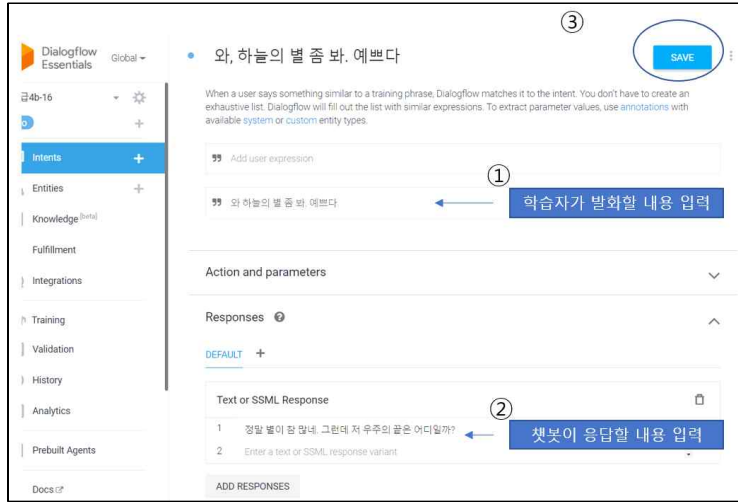
본 연구에서 사용한 챗봇 빌더인 Dialogflow Essentials는 구글 로그인으로 무료 사용이 가능하다. 대화문 생성을 위해 먼저 create new agent를 선택해 새로운 Agent이름으로 연습할 교재의 단원명이나 제목을 입력하고 언어 설정을 한국어로 바꿔 준다.



〈그림 3〉 Creat new agent(agent 생성하기)

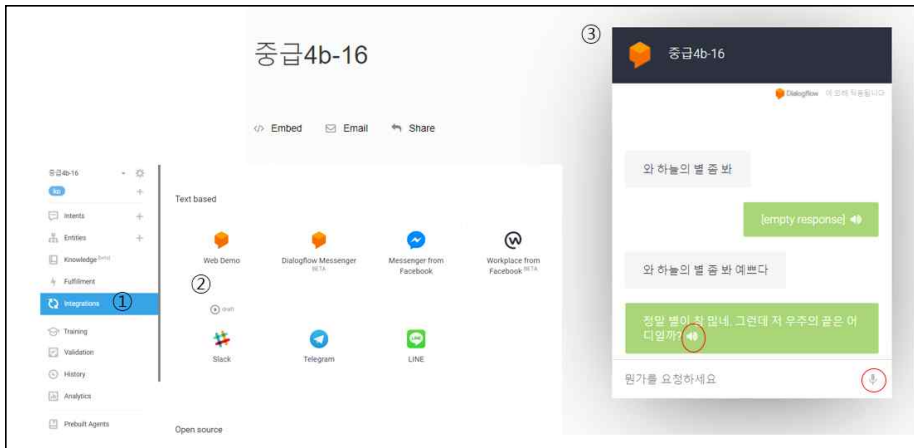
새 agent를 생성한 후에는 좌측의 intent를 추가하여 하나의 intents에 한 쌍의 대화문을 입력하면 된다. 대화문을 입력하는 방법은 intent안의 Training Phrases에 학생이 발화해 주기를 바라는 문장을 입력해 주고 Responses에는 챗봇이 학생의 발화에 반응해 주기 바라는 문장을 입력하면 된다. intents는 추가 생성할 수 있으며 이렇게 모든 대화쌍을 각각의

intents에 입력하면 대화문 전체를 연습하기 위한 챗봇이 완성된다.



〈그림 4〉 intents에 대화쌍 입력

모든 Intents의 입력이 끝나면 좌측의 메뉴 중 integrations에 들어가서 text based의 web demo를 선택하면 대화를 시작할 수 있다. web demo이외에도 선택지가 많으나 대부분 beta 중이거나 음성 출력이 안 되는 경우가 많아 연구자는 web demo를 선택했다. web demo로 생성된 url에 들어가면 챗봇 창이 나타난다. 챗봇 창 우측 하단의 마이크를 활성화 시키고 첫 문장을 말하면 발화된 문장이 그대로 전사될 뿐 아니라 챗봇이 입력해 놓은 대화쌍으로 반응을 하는데 이때 마이크를 눌러 챗봇의 발화를 들어볼 수 있다.

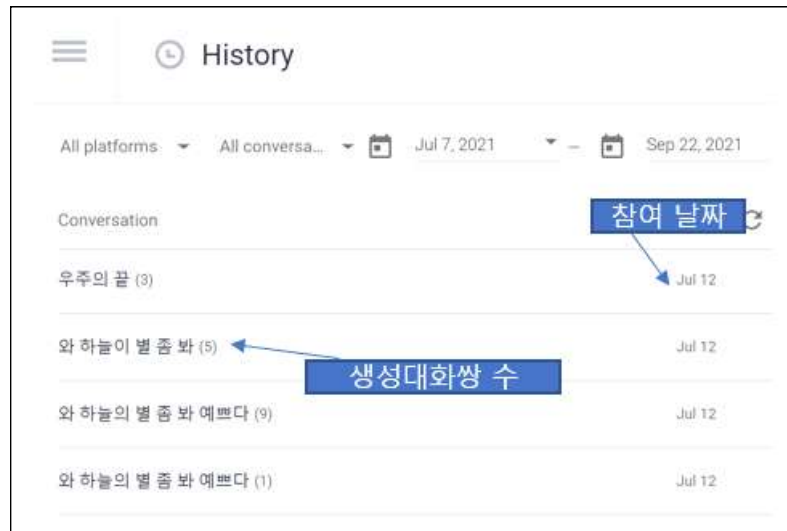


〈그림 5〉 챗봇 실행 장면

챗봇이 완성되면 교사는 수업 시간에 학생들에게 챗봇의 사용방법과 화면에 대해서 설명해 주고 챗봇 URL을 학생들에게 전달하면 된다. 학생은 구글에 가입할 필요가 없이 바로 받은 링크를 실행할 수 있다. 챗봇을 실행하면 〈그림 5〉와 같은 연습 창이 뜨며 마이크가 활성화되는지를 확인하고 바로 주어진 지문으로 연습을 시작할 수 있다.

학생들에게 각자 편한 공간과 시간을 선택해 휴대폰이나 컴퓨터 등의 단말기로 연습한 후 연습이 끝나면 한 부분을 캡처해서 교사에게 보내게 했다. 학생들은 기숙사나 집, 교실, 아르바이트 중에 사용을 했으며 한 대화문으로 여러번 반복

하여 연습할 수 있다. 자신의 발화 실수가 텍스트화 된 장면을 캡처해서 보낸 학생도 있었고 몇 번의 연습으로 실수가 없는 문장을 발화한 뒤 그 부분을 캡처해서 보내기도 했다.



〈그림 6〉 HISTORY에 자동 전사된 학습자들의 발화

〈그림 6〉을 보면 Dialogflow의 메뉴 중 history에 해당 agent의 대화 참여 기간과 참여자와 챗봇간의 모든 대화들이 자동으로 텍스트화 되어 있는 것을 볼 수 있다. 교사는 학생들의 종료 문자를 받은 후 history에서 대화 내용을 확인할 수 있다.

4.2 교육적 효과

학습자가 말하기 수업을 통해 기대하는 것은 대화 연습할 수 있는 기회와 동료 및 교사에게 제공받을 수 있는 피드백이다. 챗봇이 대화 연습의 기능과 교사의 피드백을 보완할 수 있는지를 살펴보기 위해 연구 결과와 연구 종료 후 실시한 간단한 설문조사를 분석하여 다음과 같은 효과를 확인할 수가 있었다.

첫째, 교실 수업의 대화 연습은 대화 상대가 존재하는 교실 안, 수업 중이라는 제약이 있었지만 챗봇과의 대화 연습은 시간과 공간의 제약 없이 자유롭게 대화를 시작할 수 있다는 장점이 있었다.

연구에 참여한 학생은 이동 중, 기숙사 방, 아르바이트 중에 과제를 수행했다고 말했다. 〈그림 7〉의 C1 참가자의 history를 보면 이를 유추할 수 있다.

요 카시 뭐 했어요 주말에 뭐 했어요 (5)	Aug 17
아니요 (9)	Aug 17
주말에 뭐 했어요 (25+)	Aug 17
좋아 (5)	Aug 17
육화 씨 주말에 뭐 했어요 (25+)	Aug 14
아리랑을 배웠어요 (8)	Aug 14
전주에 갔다 왔어요 (2)	Aug 14
안녕하세요 (4)	Aug 12

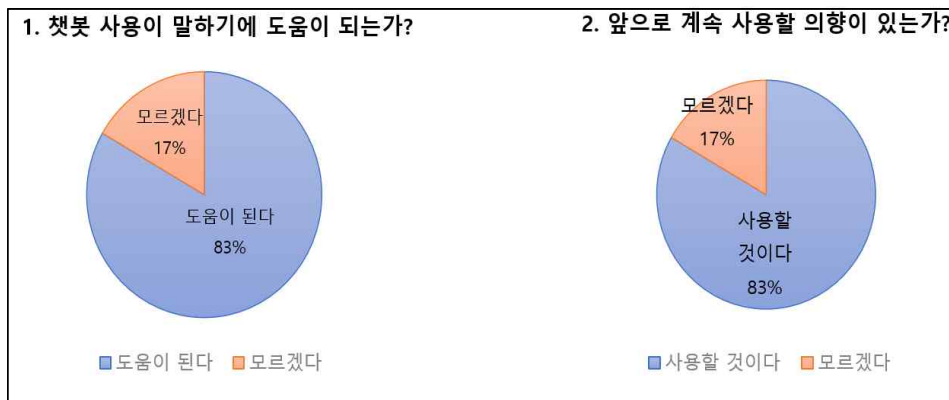
과제 제출일 →

Aug 14
Aug 14
Aug 12

〈그림 7〉 C1 참가자의 history

〈그림 7〉을 보면 학생들은 과제를 제출한 후에도 원하는 시간에 대화를 시도한 것을 볼 수 있다. 즉 학생들은 챗봇을 활용하여 시간과 장소를 선택해 대화 연습을 할 수 있고 의지를 가지고 주체적으로 대화 연습에 참여할 수 있다.

둘째, 연구 종료 후 참가자에게 3개 문항의 간단한 설문 조사를 실시하였는데 대부분 학습자들은 챗봇과의 대화 연습을 긍정적으로 인식하고 있다는 것을 알 수 있었다.



〈그림 8〉 설문 1번과 2번 문항과 응답

설문 결과를 보면 챗봇 사용이 말하기에 도움이 되는냐는 질문에 1명을 제외한 5명이 도움이 된다고 답했다. 또 대다수 학생들은 챗봇을 앞으로도 말하기 대화 연습으로 사용하고 싶다고 답했다. 설문 결과는 학생들이 챗봇을 대화 연습의 훌륭한 상대로 인식하고 있는 것이라고 볼 수 있다.

설문 3번은 챗봇 수업에 대한 의견을 자유롭게 써달라고 요청한 것으로 학습자들의 응답을 분류하면 5개 의견으로 나눌 수 있다.

〈표 2〉 설문 3번 문항과 응답

3. 챗봇 수업에 대한 의견	응답 수
"제 발음을 눈으로 볼 수 있어서 좋았습니다."	5
"챗봇 사용이 재미있었습니다."	4
"폰에서 마이크가 안 켜져서 노트북에서 했습니다."	1
"역할을 바꿔서도 연습할 수 있으면 좋겠습니다."	1
"내 발음이 틀린 게 맞아요?"	1

학습자는 자신의 발음을 볼 수 있다는 것과 함께 챗봇 사용 자체에 흥미를 가지고 참여했다는 것을 알 수 있다. 이러한 응답이 나온 이유는 학생들이 수업에 흥미를 가지고 능동적·주도적으로 학습에 참여했기 때문인 것으로 풀이된다.

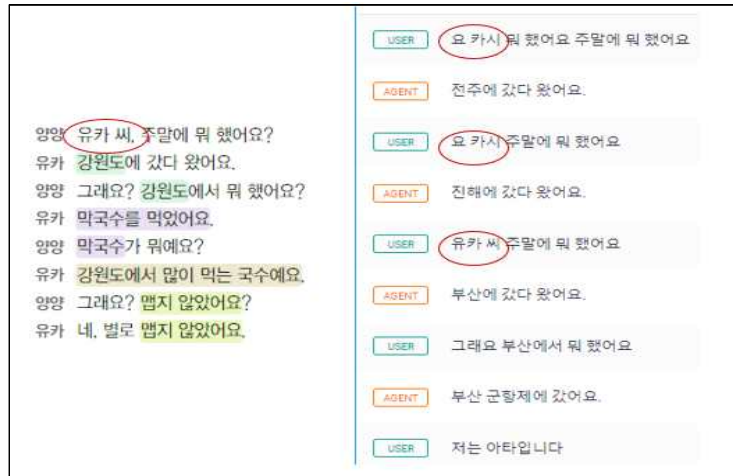
셋째, 교실 수업에서의 동료 학습자 및 교사 피드백의 한계점을 보완할 수 있었다. 학습자가 대화를 시도하는 동안 챗봇은 교사에 의해 설계된 대답을 하게 된다. 반복적으로 이러한 모범 발화에 노출되면 동료 학습자나 교사에게 피드백을 받는 것과 같이 즉각적이지는 않더라도 피드백의 효과를 줄 수 있다.

〈그림 9〉는 C2학생의 history에 전사된 내용이다. 앞에서 충분히 챗봇과 대화를 진행하다가 어느 순간 챗봇의 대답을 따라하는데 이는 학생이 원어인 교사의 발화를 따라하는 것과 같아 보인다.

USER	주말에 뭐 했어요
AGENT	진해에 갔다 왔어요.
USER	휴가시 주말에 뭐 했어요
AGENT	강원도에 갔다 왔어요.
USER	강원도 갔다 왔어요
AGENT	그래요? 강원도에서 뭐 했어요?
USER	그래요 강원도에서 뭐 했어요

〈그림 9〉 C2학생의 history

피드백과 관련하여 또 주목할 점은 자가 피드백이 원활히 일어났다는 것이다. 〈그림 10〉을 보면 학생의 발화를 챗봇이 이해하고 대답을 했음에도 자동 전사된 글을 본 학습자는 자신의 발음이 틀렸다는 것을 인지하고 계속 시도하는 것을 볼 수 있다. 이렇게 가시적인 피드백을 보고 스스로 수정을 위해 노력하는 것은 챗봇이 피드백의 기능을 충실히 하고 있다는 것을 보여 준다.



〈그림 10〉 자가 피드백의 사례

5. 결론 및 한계

본 연구는 교육 도구로서의 챗봇 설계와 대화 연습을 도와줄 학습 도구로서의 챗봇 활용의 실재를 연구하고 챗봇 대화 연습의 교육적 효과를 논의해 보았다. 연구로 얻은 결론은 다음과 같다.

첫째, 학습자들은 자신들이 편한 시간과 장소를 선택해서 챗봇과의 대화 연습을 할 수 있었다. 즉 챗봇을 활용한 연습은 대면·비대면, 국내외 학습자 뿐 아니라 개별 학습자에게도 좋은 학습 도구가 될 수 있음을 확인했다.

둘째, 학습자들은 챗봇과의 대화 연습에 흥미를 느꼈으며 자신들의 말하기 실력에 도움이 될 것이라 기대했다. 즉, 학습자들은 흥미와 기대를 가지고 주체적·능동적으로 수업 활동에 참여할 수 있었다.

셋째, 챗봇의 피드백 기능을 확인할 수 있었다. 챗봇을 활용한 대화 연습을 통해 학습자가 챗봇의 모범 문장을 모방할 수 있다. 또 학습자들이 자신의 발화 오류를 시각적으로 인지할 수 있고 능동적으로 수정할 수 있다. 교사 역시 history에 자동 기록된 자료들을 활용하여 학습자에게 피드백을 줄 수도 있다.

한편 이번 연구 과정에서 발견된 한계는 다음과 같다.

첫째는 챗봇을 활용한 수업은 시간과 장소의 제약을 받지 않는 편이라서 비대면과 대면, 국내외 학생에게 모두 가능했지만 연구 설계 단계에서 특정 국가나 특정 국가의 단말기로는 접속이 불가하다는 것을 알게 되었다.

둘째, 설문조사와 history에 전사된 대화 횟수 등으로 학습자들이 챗봇과의 대화 연습을 긍정적으로 인식한다는 것을 알 수 있었다. 그러나 참여자 수가 너무 적고 학습자들이 자발적으로 연구에 참여한 점을 봤을 때 결과를 일반화하기는 어렵다.

셋째, 학생들이 마이크를 켜고 발화하는 모든 대화가 history에 자동적으로 전사가 되지만 학습자들이 그들의 이름을 말하지 않으면 전사된 대화문이 누구의 것인지 알 수가 없었다. 또 언어적 표현에 대한 피드백 기능은 있었으나 학습자의 비언어적인 표현에 챗봇이 적절히 반응할 수 없거나 기록을 하지는 못한다는 한계가 있었다.

이 연구는 교실 수업의 물리적 한계를 극복하고 학습자들에게 대화 연습의 기회와 피드백을 제공할 수 있는 수업 도구를 제안하는 데 의의가 있다. 이 연구가 한국어 수업에서 실제적인 도움이 되기를 바라며 앞으로도 한국어 교육에서 챗봇 활용 방안에 대한 모색이 계속 되어야 할 것이다.

■ 참고문헌 ■

- 서울대학교 언어교육원(2017), 『서울대 한국어』 3B, 서울: (주)투판즈
- 서울대학교 언어교육원(2017), 『서울대 한국어』 4B, 서울: (주)투판즈
- 서울대학교 언어교육원(2019), 『사랑해요 한국어』 2, 서울: 서울대학교출판문화원
- 권세희(2010), 과제 중심 말하기 활동이 초급 한국어 학습자의 학습 불안과 흥미에 미치는 영향, 한남대학교 대학원 석사학위 논문.
- 김시연(2020), zoom을 활용한 중국 대학 온라인 한국어 말하기 교육 사례-상해 S 대학 한국어 말하기 수업을 중심으로, 이중언어학 제80권, 이중언어학회, 1-29쪽, A Case Study of Online Korean Speaking Education at University in China Using ZOOM -Focused on Speaking Class at Shanghai S University-학회이름, pp. 1-29
- 김영주(2011), 주제를론 : 한국어 교실에서의 말하기 교육 방안, 국제한국어교육학회 추계학술발표논문집 Vol.2011 No.- [2011], 국제한국어교육학회, 41-51쪽.
- 김재희(2017), 플립러닝을 활용한 한국어 글쓰기 교수학습 설계·모형 개발 연구 - 문제중심학습(PBL)을 기반으로, 고려대학교 대학원 박사학위논문.
- 김지현(2020), 모바일 러닝을 활용한 한국어 말하기 활동 방안 연구, 한국콘텐츠학회 논문집 제20권 3호, 한국콘텐츠학회, 440-451쪽.
- 김지현·조희영(2021), 피드백을 활용한 비대면 한국어 말하기 수업 사례 연구, 교사교육연구 제60권 1호, 교사교육연구학회, 21-36쪽.
- 김정은(2014), 역할극 활동을 통한 한국어 말하기 교육 방안 연구, 한양대학교 대학원 석사학위논문.
- 김중섭(2015), 기획 논문 : 한국어 기능 교육: 말하기, 듣기, 읽기, 쓰기 교육 연구의 성과와 과제. 한국어 교육 Vol.26 No.4 [2015], 국제한국어교육학회, 359-388쪽
- 박정아·이향(2021), 한국어 교육에서의 챗봇 빌더 활용 방안, 국제한국어교육학회 추계학술발표논문집 Vol.2021 No.- [2021], 국제한국어교육학회, 101-124쪽
- 서수진(2019), 한국 현대문학을 활용한 토론 토의 수업 모형 연구-조남주의 『82년생 김지영』을 중심으로, 한국언어문화학 제 16권 1호, 국제한국어문화학회, 87-211쪽
- 신윤아(2021), 한국어 교육에서 유튜브 활용 블렌디드 교육 방안 연구-관용 표현을 중심으로, 상명대학교 대학원 석사학위논문.
- 신정아(2017), 스토리텔링을 활용한 한국어 교육 방안-동시 작품을 중심으로, 어문총론 제73권 0호, 한국문화언어학회, 185-209쪽.
- 유영진(2021), 인공지능 챗봇 활용 학습 환경에 대한 설계기반연구 :초등 영어 말하기 수업 중심으로, 한국교원대학교 대학원 석사학위논문.
- 유지연·유훈식(2021), 챗봇을 통한 스마트러닝 한국어교육의 학습 효과: 학습자 인식을 중심으로, 한국멀티미디어언어교육학회 24권 2호, 연속간행물, 82-105쪽.
- 윤광진(2021), AI를 활용한 PBL 기반의 영어교수학습 모형 개발, 아주대학교 대학원 교육학박사학위논문.
- 윤화영(2020), 인공지능 시대에 다중지능 이론을 활용한 역량중심 한국어 수업 설계 방안, 인하대학교 교육학박사학위논문.
- 이경·윤영(2017), 한국어 말하기 능력 신장을 위한 플립드 러닝(flipped learning)기반 문법 교육 방안 연구, 교육문화연구 23권 4호, 인하대학교 교육연구소, 333-361쪽.
- 이미향·이운주(2018), 상담 원리로 본 한국어 교재 대화문의 전형성 연구-사회통합프로그램 한국어 1~4 교재를 중심으로, 국제 한국어교육학회 국제학술발표논문집 2018권 0호, 국제한국어교육학회, 687-688쪽.
- 이민경(2010) 한국어 말하기교육 연구의 내용과 방법에 관한 고찰 한국언어문화교육 제4권2호 87~105
- 이민아·윤보은·홍성지·정주화·한명화·이동은(2015), 한국어교재 대화의 화용적 분석-실제성을 중심으로, 국제한국어교육학회 추계 학술발표논문집 Vol.2015, 국제한국어교육학회, 200-221쪽.
- 이은비(2021), AI스피커를 통한 영어회화 학습유형에 대한 사용자 경험 연구, 홍익대학교 대학원 석사학위논문.
- 이정희·박혜옥(2018), 한국어 말하기 능력 향상을 위한 과정극(process drama)활용에 대한 고찰, 한국어 교육 제29권 1호, 국제 한국어교육학회, 199-222쪽.
- 정진(2018), 공익광고를 활용한 한국어 교수 방안, 새국어교육 제2권 114호, 한국국어교육학회, 165-191쪽.

- 최유정·김대희(2019), 토론 프로젝트를 통한 KFL 학습자의 말하기 능력 신장에 관한 연구, 국어문학 제72권, 국어국문학회, 689-711쪽.
- 한다운(2020), AI 챗봇 활용이 EFL 학습자의 말하기 능력 및 정의적 영역에 미치는 영향. 전남대학교 대학원 교육학박사학위논문.
- 한상미(2011), 주제토론 : 한국어 교실에서의 말하기 교육 방안, 국제한국어교육학회 추계학술발표논문집, Vol.2011 No.-[2011], 57-62(6쪽)

〈토론〉 “말하기 수업에서의 챗봇 빌더 활용 방안”에 대한 토론문

신윤경(가천대학교)

이 연구는 요즘과 같은 비대면 수업에서 말하기 연습이 충분히 이루어지지 않을 때 보완할 수 있는 방안에 대한 것이다. 바로 ‘챗봇 빌더’를 사용하는 것이다. 사실 비대면 수업에서 의사소통의 간극을 메꾸기 가장 어려운 영역이 ‘말하기’이기도 하다. ‘읽기’나 ‘쓰기’는 문자 언어의 특성을 활용한 시각 자료를 장소에 구애받지 않고 활용할 수 있으나, ‘말하기’는 같은 학습 언어를 구사하는 사람과 언어 환경이 뒷받침되어야 하기 때문이다. 따라서 이런 시간과 장소에 구애받지 않고 말하기 연습을 할 수 있는 ‘챗봇 빌더’ 활용은 의의가 있다. 따라서 그 의의를 더욱 살릴 수 있도록 몇 가지 질문과 제안을 함으로써 연구의 완성도를 높이는 데 일조하고자 한다.

1. 먼저, ‘챗봇 빌더’를 적용할 수 있는 학습자 수준이 너무 한정되어 있는 것 같다. 대화쌍에서 교사가 예상 답안을 입력하고 그 답을 해야만 정답으로 인정되기 때문에 초급 수준의 학습자들에게는 적합하지만, 중·고급 수준의 학습자들에게는 한정될 수밖에 없다. 중·고급 학습자들은 정답은 아니지만 맥락에 맞는 다른 문법과 어휘를 사용해서 답을 해도 허용된다는 것을 알고 있기 때문이다. 따라서 대화의 양상이 너무 단순해지고 획일화된다면, 맥락과 화용에 맞게 유창한 대화를 구사할 수도 있어야 하는 중·고급 학습자들에게는 오히려 역효과가 될 수도 있다. 그런데 연구에 참여한 학습자들 중에 4단계를 마친 학습자들이 있었는데, 그들에게는 어떤 양상과 한계가 있었는지 궁금하다.

2. 연구에 참여한 학습자들을 3집단으로 나누면서 한 집단은 대면, 다른 두 집단은 비대면 상황이라고 구분했다. 또한, 학문목적과 일반목적 학습자로 구분하기도 했다. 물론 참여자 숫자가 너무 적어 일반화해서 말할 수는 없지만 학습자들을 구분했던 상황마다 다른 결과가 나왔는지 궁금하다. 특히, ‘대면’, ‘비대면’ 상황의 구분은 연구의 목적이 비대면 상황에서의 한계를 극복하기 위한 것이기 때문에 심층 인터뷰를 통해서라도 결과와 효과를 비교해 보는 것을 제안한다.

3. 마지막으로, 이번 연구의 문제점을 보완하여 2차 연구를 진행할 계획이라고 했는데 결론에 제시된 한계는 대부분 ‘챗봇 빌더’ 설계와 실현의 과정에서 생겨난 물리적 문제점들이다. 물론 이런 물리적 문제점을 해결하는 것도 필요하지만, 대부분의 교수자들은 ‘챗봇 빌더’를 활용할 때 어떤 상황과 수준의 학습자들에게 어떤 형식의 대화 내용을 제시하는 것이 효과적인지 궁금할 것 같다. 따라서 원래 설계하셨던 2차 연구가 무엇인지도 궁금하고, 이런 내용들이 포함된 2차 연구를 진행해보는 것도 의미가 있을 것 같다고 제안해 본다.

한국 현대시 텍스트의 문화 어휘 분석 연구

- 김소월 시 코퍼스를 중심으로 -

황혜란(홍익대학교)

1. 서론

최근 문학작품을 통한 한국어교육은 문학작품이 가지는 효용성으로 활발하게 전개되었다. 문학의 장르 중 시는 사회 문화적인 맥락을 민감하게 포착하여 풍성하게 담아내고 있으며, 인간의 보편적인 정서를 다루고 있어 학습자가 자국 문학 학습의 경험을 통해 시 작품을 향유함으로써 적극적으로 학습에 참여할 수 있다는 장점이 있다. 또한, 텍스트의 길이가 짧아 전문을 학습하기에도 용이하다.

현대시를 활용한 최근의 연구들은 학습자들의 배경지식을 활성화하거나 자국 문화와 비교를 통한 상호문화주의¹⁾의 관점에서 많은 연구가 이루어지고 있다.²⁾ 현대시 작품 중에서도 가장 많은 연구가 이루어지는 시는 단연 김소월의 작품이다. 김소월의 시는 한국전통시의 전형이자 한국인에게 가장 익숙하고 사랑받는 시이기 때문이다. 노금숙(2010)은 김소월의 시와 마야의 '진달래꽃'을 중심으로 노래와 시의 상호텍스트성을 활용하여 다문화사회의 한국어 수업 방안을 제시하였고, 김진호(2012)는 김소월의 작품을 대상으로 한국문화교육 방안을 설계하였다. 조수진(2013)은 「진달래꽃」, 「초혼」, 「접동새」를 중심으로 김소월 시에 나타난 이별의 정서 교육 방안을 제시하고 있다. 박서영(2019)는 한국어문학교육을 위한 김소월 시의 중국어 번역에 관한 연구를 진행하였다.

현대시를 교육할 때 상호문화적 관점의 접근은 작품들 간의 상호텍스트성으로 연결된다. 상호텍스트에 대해 크리스테바(J.Kristeva, Toril Moi(ed)1986:36-37)는 “하나의 텍스트는 작가와 그의 잠재적인 시간의 관계인 수평적 축과, 텍스트 그 자체와 다른 텍스트들이 형성되는 수직의 틀, 이 둘로 이루어지는 좌표에 놓인다.”고 하였다. 즉 상호텍스트성은 텍스트와 텍스트의 관계라고 할 수 있는데, 넓은 의미에서 텍스트는 인간의 언어로 이루어진 모든 것을 말한다. 상호텍스트의 관점에서 독자적으로 형성된 텍스트는 없으며, 시 작품의 경우도 개별적인 텍스트가 아닌 과거의 텍스트의 영향을 받으며

1) 상호문화주의(interculturalism)란 사회 내에 존재하는 다문화적 사실에 대한 인정을 넘어 서로 다른 문화들 사이(inter)의 대화와 상호 작용을 강조하는 교류의 철학이다. 상호문화주의는 다른 문화를 일방적으로 배우는 것이 아닌 다른 문화들 간의 차이를 인정하고 쌍방의 상호작용을 중요시하는데 서로 다른 문화들의 복수적 공존을 주장하는 다문화주의와는 구별된다. 상호문화적 접근 방법을 통한 교육은 상호문화능력을 함양하는 것을 목표로 하는데 상호문화능력이란 타문화에 대한 지식을 바탕으로 해당 문화를 존중하고 이해하면서 자신의 문화를 상대화시킬 수 있는 능력을 말한다. 김창근(2015). 「상호문화주의의 원리와 과제: 다문화주의의 대체인가 보완인가?」, 『윤리연구』 103, 한국윤리학회, 188쪽 참조.

2) 양민애(2008)는 이문화 교육의 필요성 방향을 제시하였고, 조수진(2009)은 정의적 영역의 접근으로 간문화 교육, 전홍(2012)은 문화 이해를 중심으로 한국어 학습자의 상호작용적 시 읽기 양상을 살폈고, 김혜민(2015)는 한국어 상호문화 교수전략을 제시, 김예리나(2018)은 학습자의 경험을 통한 상호문화교육의 구체적인 방안을 설계하였다. 박제은(2019)는 중급 학습자를 대상으로, 김유리(2019)는 중국인 학습자를 대상으로 현대시에 내재된 문화 지식 교육 방안을 마련하였다.

미래의 텍스트에 영향을 미친다. 이러한 맥락에서 오지혜(2015: 87-99)는 한국의 현대시 텍스트를 공부할 때 학습자들은 작품 속에 내재된 문화 요소들을 학습하고 이를 자국문화와 비교할 수 있으며, 비슷한 주제의 다른 작품들과 비교하기도 하고, 번역을 통해 새로운 형태의 텍스트를 생산하기도 한다고 하였다. 학습자들은 시를 통해 한국의 문화에 접근할 때 텍스트를 이해하고 문화 지식을 구성하며 자국 문화와의 비교를 통해서 적극적으로 의미를 구성해 나가는 것이다.

시인은 느낌이나 생각을 표현하기 위해 일상어를 선택하여 공을 들여 다듬는다. 정련의 과정을 거친 시어들은 함축적이고 내포적인 의미를 지니고 있어 독자의 상상력을 불러일으킨다. 그러나 시어의 특징은 실제 교육 현장에서 활용되기보다는 이론적 접근에만 머무르게 하는 요인이 되기도 한다. 시어는 그 함축성으로 인해 일상어의 의미를 기준으로 판단하기에는 다중적 의미를 이해하기가 어렵기 때문이다.³⁾ 이러한 문제를 해결하기 위하여 주석을 달아 어려운 어휘에 대한 문화적 맥락을 밝히거나 가치나 정서 등의 개별 문화 항목에 집중하여 교육 방안을 설계하기도 한다. 그러나 구체적인 교육 방안을 설계할 때 어느 어휘에 주석을 달 것인지, 어느 문화 항목을 교육 내용으로 삼을 것인지에 대한 일관된 기준과 기초 자료는 거의 없다고 할 수 있다. 따라서 세계의 다양한 문화와의 상호문화적 관점에서 한국 현대시 작품을 비교 분석하고 작품과 작품 사이를 연계할 수 있는 일관된 기준과 기초 자료를 제공하는 것이 필요하다.

현대시 기초 자료와 관련된 연구로 김병선(2004)은 1990년대 초반부터 약 10여년 간 『한국현대시데이터베이스(KPD)』를 구축하였다. 그 중 『한국현대시 코퍼스(Korean Modern Poetry Corpus)』는 현대시어 용례사전 편찬을 위한 범위(1923~1950에 출판된 창작시집)에 포함되어 있는 현대 창작시 9,893편을 수록하였으며 본 연구의 기반인 「김소월 코퍼스」도 『한국현대시 코퍼스(KoPoCo)』에 구축되어 있다. 김병선(2004)에서는 『한국현대시 데이터베이스(KPD)』의 구성에 대해 상세히 설명하고 있으며, 김병선(2015)은 텍스트의 계량으로부터 시작해 어휘 고유성과 문체, 통사 구조 계량, 통시적 탐구와 문체 등 한국『한국현대시 코퍼스(KoPoCo4)』에 수록된 현대시 8000여 편의 모든 언어적 단위를 기술통계학으로 접근·분석하였다. 문화 어휘 기초 자료와 관련한 연구로 황혜란(2021)은 오정희의 「중국인 거리」를 대상으로 소설 텍스트의 어휘에 문화 분류 코드를 부여하는 과정을 구체적으로 밝히었다.

본 연구에서는 『한국현대시 코퍼스(KoPoCo)』를 기반으로 황혜란(2021)의 방법론을 적용하여 시 텍스트의 문화 항목을 체계적으로 분류하고자 한다. 대상 작품은 『한국현대시 코퍼스(KoPoCo)』 중 「김소월 코퍼스」 대상으로 하고 분류 기준은, HRAF 문화 코드를 기준으로 한다. 문화 분류 코드는 주제, 제목, 어휘의 순서로 부여할 것이다.

2. 분석의 기준

2.1. 분석 대상 텍스트

현대시 텍스트의 문화 항목을 추출하기 위해서는 먼저 분석의 대상이 되는 텍스트가 필요하다. 본 연구는 분석 대상 텍스트로 『한국현대시 코퍼스(KoPoCo)』를 기반으로 하였다. 『한국현대시 코퍼스(KoPoCo)』는 한국학중앙연구원의 프로젝트로 구축된 데이터베이스로 1923년부터 1950년 사이에 출판된 창작시집의 작품이 수록되었다. 코퍼스의 규모는 344명 시

3) 소설의 경우도 일상 언어의 어휘가 쉽다고 하더라도 작품 속에 내재된 문화적 간극으로 인하여 난이도 차이가 발생한다. 소설 어휘의 난이도는 ‘황혜란(2017), 『한국문학교육용 텍스트의 이독성 연구』, 한국학중앙연구원 박사학위 논문’, 소설 어휘의 문화적 간극은 ‘이다니엘(2019), 「외국인 학습자 위한 한국 소설 이독성 측정 모형 개발 연구-문화적 공백 개념을 중심으로」, 경희대학교 석사학위 논문’ 참조.

4) KoPoCo는 ‘Korean Modern Poetry Corpus’의 약칭이다.

인의 8,201 편의 작품에서 추출한 모두 60만 개 이상의 어휘(token)에 42,485 종의 어종(type)으로 구성되어 있다. 방대한 양의 현대시 코퍼스 중에서 본 연구에서는 한국의 대표 시인이라 할 수 있는 김소월의 시 작품들을 분석하고자 한다. 『김소월 시 코퍼스』는 『진달래꽃』 수록 작품 126편, 『소월시초』에 수록 작품 19편으로 총 145편, 시어 수는 8,203개이다.

김소월의 시 작품을 분석 대상 텍스트로 선정한 이유는 김소월이 한국인에게 가장 익숙한 시인이자 가장 한국적인 색깔을 지닌 시인이기 때문이다. 김소월의 시는 내용과 형식 면에서 한국시의 전형이며, 전통적이라는 평가가 일반적이다. 그의 시는 수많은 가곡으로 작곡되고 영감을 주기도 하였으며, 현재까지도 한국인들이 가장 많이 애송하는 시로 서정적인 공감대를 폭넓게 지니고 있다. 또한 개인의 서정 너머에는 식민지 상황에서의 한국 민족의 고통과 슬픔을 보여주고 있다. 김소월 시의 시어들은 평범한 일상적 언어뿐만 아니라 관서지방의 토착 어휘들이 풍부하여 개인과 민중의 정서를 보여주고 있다. 따라서 김소월의 시에 내재된 역사-문화적 배경이나 전통적인 문화 요소들을 추출해 내는 작업은 상호문화적 관점에서의 시교육을 위한 유의미한 자료가 될 것이다.

김소월 코퍼스는 동음이의어 및 다의어분석 말뭉치의 형태로 구성되어 있으며 어휘 용례 색인이 제시되어 있고, DB로도 구축된 상태이다.

〈그림 1〉 「김소월 코퍼스」의 구성 예시

1	활용형	한글형	기본형	소분류	중분류	앞둠	키워드	뒤문맥	제목
2	싸을	까을	가01ng	장소-위치	1 명사:장소	\$나는 오히려 못물	싸을	싸고 싸든다./!@	가을 저녁에
3	싸엿	가엿	가01ng	장소-위치	1 명사:장소	<술잔디>/ 잔디, 잔디./ 深深山川에	싸엿	금잔디./!@	술잔디
4	싸에	까에	가01ng	장소-위치	1 명사:장소	(나의 집)/ 들	싸에	떠러져 나가 안즌 메시슬의/ 넓은 바다의 물따 뒤	나의 집
5	싸에	까에	가01ng	장소-위치	1 명사:장소	\$다니다 옛 동무 무덤	싸에/	正月 대보름날 달마지!//!@	달맞이
6	싸의	까의	가01ng	장소-위치	1 명사:장소	보습 대일 땅이 있었더면)/ 나는 꿈꾸었노라,	싸의	하로 일을 다 맞추고/ 夕陽에 마을로 도라오는 꿈	바라건대는 우리에게 우리의
7	싸에는	까에는	가01ng	장소-위치	1 명사:장소	\$원모래 모래 빗긴 船艙	싸에는/	한가한 배노래가 멀리 자즈며/ 날 젊을고 안개는 깃	山 위에
8	가의	가의	가01ng	장소-위치	1 명사:장소	\$新開地에도 봄은 와서 가냘픈 밭줄은/ 독	가의	어슴푸레한 개버들 어린 엄도 추기고./ 난별에 파릇	爽快한 아침
9	싸에	까에	가01ng	장소-위치	1 명사:장소	<새벽>/ 落葉이 밭이 숨는 못물	싸에/	웃둑웃둑한 나무 그림자/ 물빛조차 여섯프러히 써	새벽
10	싸의	까의	가01ng	장소-위치	1 명사:장소	새운 일도 업지 안치만/ 아직도 깨마다는 당	싸의	꿈은 잊지만!//!@	임에게
11	싸의	까의	가01ng	장소-위치	1 명사:장소	당신을 생각하면 지금이라도/ 비 오는 모래밭	싸의	꿈은 잊지만/ 당신은 니저 바린 서름이외다!//!@	임에게
12	싸에	까에	가01ng	장소-위치	1 명사:장소	\$津頭江 가림	싸에	살든 누나는/ 津頭江 암마을에/ 와서 읊니다!//!@	겹동새
13	싸에	까에	가01ng	장소-위치	1 명사:장소	@옛날, 우리 나라/ 먼 뒤쪽의/ 津頭江 가림	싸에	살든 누나는/ 이북어미 식생에 죽었습시다!//!@	겹동새
14	싸의	까의	가01ng	장소-위치	1 명사:장소	\$오오 가슴이어, 모닥불 피어오르는/ 내 한세	싸의	가을도 갖서라./!@	찬 저녁
15	싸	까	가01ng	장소-위치	1 명사:장소	스뚱내 고개를 넘고 넘어/ 짐 싣고 댕든 말도	싸,	夕陽 손에/ 고요히 조오는 한때는 다 있나니./!@	追悔

〈그림 1〉은 「김소월 코퍼스」 구성 예시이다. 「김소월 코퍼스」는 원본에서 한글형으로 바꾼 후 기본형을 추출한 한 별도의 어휘 분류 체계를 통해 어휘를 분류하였다. 기본형 정보는 어휘의 기본형, 원어 표기, 동음이의어 첨자, 다의어 구분용 부호, 품사 기호로 이루어진다. 기본형의 추출은 김병선(2015)의 방식에 따라 일반명사(ng), 고유명사(nm), 의존명사(nb), 대명사(np), 수사(nr), 동사(vv), 형용사(va), 관형사(mm), 부사(ma), 감탄사(ic)로 표기되어 있다.(김병선, 2015:349-386)

각 어휘는 의미 분류 단계에서 대분류-중분류-소분류 세 수준의 단계로 구성되어 있다. 대분류는 품사, 중분류는 넓은 의미, 소분류는 세분화된 의미이다. 예를 들어 가01ng은 경계에 가까운 바깥쪽 부분을 나타내는 어휘로 01은 표준국어대사전 첨자, ng는 명사를 나타내는 품사 기호, 의미는 대분류는 명사, 중분류는 '장소', 소분류는 '위치'이다. 장소의 중분류는 다시 소분류 체계로 장소-방위, 장소-영역, 장소-위치, 장소-기타의 5종류로 나누어진다. 중분류와 소분류 체계의 어휘 수는 〈표 1〉과 같다.

〈표 1〉 「김소월 코퍼스」의 어휘 분류 체계

중분류	소분류 항목 ⁵⁾	중분류 어휘수	중분류	소분류 항목	중분류 어휘
I 고유:국명	1	1	I 수사:숫자	2	14
I 고유:기타	1	3	I 의존:단위	5	44
I 고유:시설	1	1	I 의존:의존	5	139
I 고유:인명	1	12	II 동사:동작	6	1625
I 고유:지명	1	69	II 동사:사태	1	336
I 대명:시물	3	28	II 동사:정신	9	219
I 대명:시기	1	2	II 동사:동작	6	1625
I 대명:인간	6	355	II 동사:사태	1	336
I 대명:일반	1	45	II 동사:정신	9	219
I 대명:처소	4	49	II 어근:어근	2	4
I 명사:개념	5	273	II 조동:행동	10	141
I 명사:고체	3	8	II 조형:형용	4	30
I 명사:기체	1	5	II 형용:감각	19	278
I 명사:동물	7	110	II 형용:상태	15	282
I 명사:사회	5	36	II 형용:정서	10	171
I 명사:산천	4	182	III 관형:성상	4	42
I 명사:시간	11	510	III 관형:수효	1	84
I 명사:식료	3	8	III 관형:지시	4	154
I 명사:식물	7	174	III 부사:공간	3	22
I 명사:신체	10	209	III 부사:부정	2	59
I 명사:액체	2	35	III 부사:성분	2	36
I 명사:예술	5	48	III 부사:성상	9	318
I 명사:오감	7	159	III 부사:시간	7	51
I 명사:의류	3	36	III 부사:양태	7	26
I 명사:이동	4	66	III 부사:의문	1	15
I 명사:인간	17	649	III 부사:접속	2	127
I 명사:인공	4	111	IV 감탄:감탄	5	57
I 명사:자연	6	325	V 조사:조사	1	1
I 명사:장소	4	333	VI 접사:접미	1	1
I 명사:주거	4	54	총합계	272	8203
I 명사:해양	3	31			

2.2. 문화 분류 체계

김소월 시 텍스트의 문화 항목을 추출하고 분석하기 위해서 문화를 범주화하고 항목화하는 기준이 필요하다. 본 연구에서는 문화 인류학의 관점을 적용하여 HRAF 문화 분류 체계를 활용하고자 한다.

HRAF(Human Relation Area Files)⁶⁾는 1930년대 미국 예일대학의 인간관계연구소(Institute of Human Relation)에서 세계의 모든 문화자료를 수집하고 그 자료를 주제 분류하여 항목별로 찾을 수 있도록 정리한 문서이다. HRAF 파일은

5) 소분류 항목의 수는 총 272종으로 지면의 한계상 소분류의 개수만 표시하였다.

6) HRAF 에 관한 자세한 정보는 'https://hraf.yale.edu'에서 확인할 수 있다.

특정 문화에 대한 정보를 담고 있는 모든 텍스트들이 모아지고 지리적으로 나누어지고 주제로 분류된다. 이 파일에는 두 개의 분류 체계가 있는데 OWC(The Outline of World Cultures, 6th ed. 1983)⁷⁾와 OCM(The Outline of Cultural Materials)이다. 본고에서는 OCM 코드를 활용하고자 한다. HRAF의 OCM 분류 체계는 대부분의 분류 코드가 주제 관련 내용으로 구성되어 있다.

표 2. HRAF OCM 코드 대분류 체계

코드	대분류	코드	대분류
100	ORIENTATION	500	WATER, AIR and SPACE TRANSPORTATION
110	BIBLIOGRAPHY	510	LIVING STANDARDS AND ROUTINES
120	METHODOLOGY	520	RECREATION
130	GEOGRAPHY	530	ARTS
140	HUMAN BIOLOGY	540	COMMERCIALIZED ENTERTAINMENT
150	BEHAVIOR PROCESSES & PERSONALITY	550	INDIVIDUATION AND MOBILITY
160	DEMOGRAPHY	560	SOCIAL STRATIFICATION
170	HISTORY AND CULTURE CHANGE	570	INTERPERSONAL RELATIONS
180	TOTAL CULTURE	580	MARRIAGE
190	LANGUAGE	590	FAMILY
200	COMMUNICATION	600	KINSHIP
210	RECORDS	610	KIN GROUPS
220	FOOD QUEST	620	COMMUNITY
230	ANIMAL HUSBANDRY	630	TERRITORIAL ORGANIZATION
240	AGRICULTURE	640	STATE
250	FOOD PROCESSING	650	GOVERNMENT ACTIVITIES
260	FOOD CONSUMPTION	660	POLITICAL BEHAVIOR
270	DRINK, DRUGS, AND INDULGENCE	670	LAW
280	LEATHER TEXTILES, AND FABRICS	680	OFFENSES AND SANCTIONS
290	CLOTHING	690	JUSTICE
300	ADORNMENT	700	ARMED FORCES
310	EXPLOITATIVE ACTIVITIES	710	MILITARY TECHNOLOGY
320	PROCESSING OF BASIC MATERIALS	720	WAR
330	BUILDING AND CONSTRUCTION	730	SOCIAL PROBLEMS
340	STRUCTURES	740	HEALTH AND WELFARE
350	EQUIPMENT AND MAINTENANCE OF BUILDINGS	750	SICKNESS
360	SETTLEMENTS	760	DEATH
370	ENERGY AND POWER	770	RELIGIOUS BELIEFS
380	CHEMICAL INDUSTRIES	780	RELIGIOUS PRACTICES
390	CAPITAL GOODS INDUSTRIES	790	ECCLESIASTICAL ORGANIZATION

7) OWC는 지리적으로 세계를 8개 영역으로 나누어 구분하고 아시아는 A, 유럽은 E, 아프리카는 F, 중동 지역은 M, 북미지역은 N, 오세아니아는 O, 러시아 지역은 R, 그리고 남미지역은 S 등으로 알파벳코드를 부여한다. 그리고 8개의 작은 영역에서 더 구체적인 지역은 그 다음의 알파벳을 부여하여 하위 지역으로 구분한다. 그리고 더 세부적인 하위 지역의 단위는 수자가 부여된다. 예를 들면 한국은 AA01이다.

400	MACHINES	800	NUMBERS AND MEASURES
410	TOOLS AND APPLIANCES	810	SCIENCES AND HUMANITIES
420	PROPERTY	820	IDEAS ABOUT NATURE AND PEOPLE
430	EXCHANGE	830	SEX
440	MARKETING	840	REPRODUCTION
450	FINANCE	850	INFANCY AND CHILDHOOD
460	LABOR	860	SOCIALIZATION
470	BUSINESS AND INDUSTRIAL ORGANIZATION	870	EDUCATION
480	TRAVEL AND TRANSPORTATION	880	ADOLESCENCE, ADULTHOOD AND OLD AGE
490	LAND TRANSPORT	890	GENDER ROLES AND ISSUES
		900	TEXTS
		910	ARCHAEOLOGICAL ANALYSES

〈표 2〉는 HRAF OCM 코드 대분류 체계이다. OCM 분류 체계는 0~999까지 대·중·소 단위로 분류하고 있는데, 코드 000, 100, 110, 120, 900, 910 등은 정보원, 연구 방법, 텍스트, 분석법 등을 다루고 있으며, 코드 130부터 890까지는 문화 관련 주제로 구성되어 있어 인류문화와 관련한 대부분의 내용을 대상으로 하고 있다.

OCM은 세계의 문화 특성에 따라 주제의 범주를 정하고 세 자리 숫자 코드를 부여한다. 처음 두 자리는 대분류이고, 마지막 세 번째 자리 숫자는 대분류의 주제를 세분화하는 중분류이다. 예를 들어 590(family)은 591(residence), 592(house hold), 593(family relationships), 594(nuclear family) 등과 같은 방식이다. 또한 중분류 체계 아래에서는 소분류로 세부적인 색인어가 있다. 색인어는 웹 사이트 검색을 통해 확인할 수 있다. 방대한 양의 문화 내용을 정리해 낼 수 있는 만큼 HRAF의 문화 분류코드를 활용한다면 문학텍스트에 쓰인 어휘의 의미를 항목별로 분석할 수 있으며, 일상어의 사용이 시 작품 속에서 문화적으로 어떻게 함축되어 있는지를 파악할 수 있을 것이다.(황혜란, 2021:269)⁸⁾

HRAF 분류 체계를 기준으로 삼은 이유는 상호텍스트성을 위한 범위의 확장 문제와 디지털화의 가능성을 찾고자 하기 때문이다. HRAF는 한국의 문화뿐만 아니라 세계의 문화 특성에 따라 주제의 범주를 정하고 있어 문학텍스트에 나타나는 이문화적인 요소들을 추출하는 데 활용할 수 있다. 주제 항목도 80개 주제 아래 9개의 세밀한 중분류 체계가 있어 다룰 수 있는 문화 항목의 범위가 넓으며 중분류 아래에 색인어는 A-Z까지 10,000개 이상의 키워드를 검색할 수 있다.⁹⁾ 문학 텍스트에서 사용되는 소재들은 한국의 문화뿐만 아니라 이문화적인 요소까지 전 세계의 문화 항목들이 담길 수 있기 때문에 분류 기준이 세밀한 HRAF가 누락되는 어휘가 적으며, 작품 간 비교, 작가 간 비교, 외국인 학생의 경우 자국 문학과 비교 등 그 활용의 범위가 넓을 것이라 판단된다.(황혜란, 2021:269)

교수 현장에서 문학 작품을 문화 교수의 도구나 교재 제작 시에도 교수자가 참고할 수 있는 자료가 충분하지 않다. 문학작품을 교수 현장에서 손쉽게 활용하기 위해서 자료를 디지털화하여 구축해야 한다. 이 과정에서 각 어휘에 태그를 달아 주어야 하는데 숫자 코드로 구성된 HRAF 코드는 이러한 관점에서 활용성이 높다. 어휘에 분류 코드를 부여하는 작업은 이미 구축된 코퍼스를 기반으로 하지만 개별 어휘의 의미를 하나씩 확인하고 분류하는 작업은 연구자의 품이 들어

8) 한국어교육에서의 문화 항목은 대체로 교육해야 할 문화 교수 항목 중심으로 이루어져 있다. 대표적으로 국립국어원의 『국제통용 한국어교육 표준모형』에서는 한국어교육에서 교수해야 할 문화 영역에 대해 목표와 내용을 등급별로 정리하여 기술하였다 그러나 『국제통용 한국어교육 표준모형』의 문화 항목은 등급별 교수 내용을 중심으로 하고 있기 때문에 다양하고 광범위한 세계의 문화 항목을 포괄하기에는 범위의 한계가 있다.

9) 필자가 공개된 웹사이트 검색을 통한 확인한 색인어는 10,427이다. 별도의 공개된 색인어 목록이 없으므로 색인어의 수는 다소 오차가 있을 수 있다.

갈 수밖에 없다. 이러한 작업은 효율적인 분석 시스템을 만들기 위한 준비 과정이라 할 수 있으며, 이 과정을 통해 체계적인 분류 체계와 충분한 양의 기초 자료가 확보될 수 있을 것이다.

2.3. 분석의 절차

현대시 텍스트의 문화 항목을 분석하기 위해서 김소월 시 텍스트를 주제, 제목, 어휘 단위로 분석하였다. 코드를 부여하는 과정은 다음과 같다. 첫째 HRAF의 코드 번호를 「김소월 코퍼스」의 주제에 부여하였다. 전체 작품의 주제에 해당하는 문화 분류 코드를 적용함으로써 개별 작품 전체의 문화 요소를 파악할 수 있다. 둘째 시 텍스트의 제목에 HRAF 분류 코드를 부여하였다. 제목에 분류 코드를 부여했을 때 일상어로 된 시의 제목이 주제에서는 다른 문화 분류코드가 적용될 수도 있다. 이러한 경우는 일상적인 시가 어떻게 시에서 문화적 요소를 함축하는지 살펴볼 수 있는 계기가 될 수 있다. 셋째 텍스트의 어휘 단위에 HRAF 문화 분류 코드를 부여하였다.

「김소월 코퍼스」는 어절 단위로 기본형을 추출하여 품사별로 분류해 놓았다. 이 기본형 어휘에 HRAF 분류 코드를 부여하여 문화적인 요소들 살펴보았다. 문화 분류 코드는 「김소월 코퍼스」 기본형을 대상으로 부여하였다. 「김소월 코퍼스」의 대분류는 품사를 기준으로 하였고, 중분류와 소분류의 의미 체계는 연역적 방식이 아닌 귀납적이고 기술적인 방식으로 이루어져 있다. 분석의 용이성을 위해 일정한 의미군으로 묶은 소분류 체계를 기준으로 우선 분류한 후 소분류 체계로 분석되지 않는 경우 좀 더 넓은 의미의 중분류 체계를 분석하였다. 또한 중분류와 소분류 체계의 어휘 중에서도 HRAF로 분석하는 과정에서 기본형의 의미와 작품 속에서의 쓰임에 다른 코드를 부여하기도 하였다. 즉 귀납적 방법으로 작성된 「김소월 코퍼스」의 어휘 분류 체계를 수렴하여 연역적 분류 체계인 HRAF 문화 분류 코드를 재차 부여하는 방식이다.

기본형의 품사 중에서는 명사에 문화 코드를 부여하였다. 동사와 형용사는 동작과 상태를 나타내며 부사와 관형사는 수식의 역할을 한다. 문화 코드를 부여할 수 실질적 의미를 지니는 것은 의미의 범주화가 이루어진 명사에 해당하므로 명사에 한정하여 어휘를 분석하였다. 시 텍스트의 특성상 사물 그 자체를 나타내는 것보다 사물이나 사태의 상태나 동작을 나타내는 동사와 형용사의 표현이 많고, 꾸밈 표현인 부사나 관형형 용언이 많다. 동사가 정적/동적인 문화의 특징, 행동 중심적이거나 사유 중심적인 문화적 특징을 나타내기도 하고, 형용사는 시적 정서를 표현하는 중요한 장치이다. 그럼에도 불구하고 의미의 범주화를 통해 분류 코드를 부여하는 본 연구 작업의 특성상 실질적 의미를 지니고 있는 명사를 분석 대상으로 한정하였다. 분석된 어휘는 총 8,203개(token) 중 3,448개(token)이다.

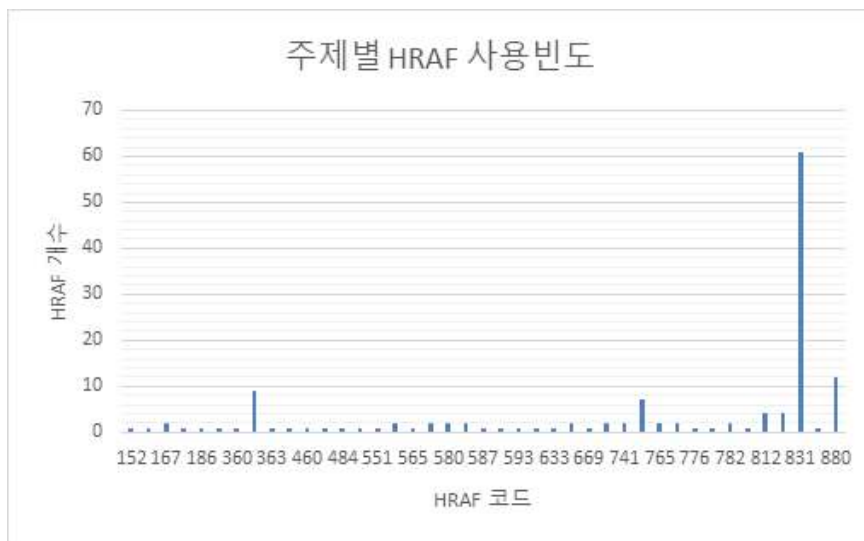
3. 분석의 결과

김소월의 시는 우리 민족 고유의 정서와 민요적 율격을 바탕으로 하고 있다. 그의 시는 그리움, 슬픔, 외로움 등 비극적인 한의 정서를 표현하면서 이면에는 존재에 대한 깊은 성찰을 보여 주고 있다. 그의 시에 드러나는 정서와 성찰은 시의 구조와 리듬, 시어 등을 통해 형상화되고 있다. 김소월의 시 전반을 이루는 정조와 전제적인 구조와 율격 등을 제외하고, 다양한 해석을 가능하게 하는 함축적 의미의 시어를 하나의 범주로 규정하는 것은 온당하지 않을 수 있다. 그러나 분류 코드를 통한 분석은 해석 차원의 아닌 기초 자료를 마련하는 차원에서 이루어진다. 필요한 자료를 손쉽게 찾고 비교하고 해석하며, 교수자나 연구자의 필요에 따라 또 다른 텍스트로 가공할 수 있다. 따라서 본고에서의 의미 부여 기준은 가장 일반적인 기준으로 부여되었으며, 하나의 어휘에 하나의 코드를 부여하였다.

3.1. 주제별 문화 코드

김소월 시를 전반적으로 관통하는 주제는 ‘이별의 정한’으로 그리움과 슬픔 등의 정서를 표현하고 있다. 코퍼스의 구성이 어절 단위로 되어 있고 분류 코드도 어절 단위로 부여하고 있기 때문에 이별의 슬픔, 임에 대한 그리움 등은 각각의 어휘에 코드를 부여해야 한다. 그러나 서정시라는 장르의 특성상 정서나 감정과 관련된 주제를 별도로 코드를 부여할 필요는 없다고 판단하였기 때문에 핵심적인 내용에 대해서만 코드를 부여하였다. 예를 들어 ‘임에 대한 그리움’이라는 주제에 코드를 부여할 때 그리움의 정서는 감정(152)는 별도로 부여하지 않고 임(831)에 대한 코드를 부여하였다. 다만 주제가 명확하게 어휘로 명확하게 드러나는 경우는 설움(152)과 같이 감정(152)의 코드를 부여하였다. 주제를 부여한 작품은 모두 145편으로 분류 코드를 부여한 결과는 다음과 같다.

〈그림 2〉 주제별 HRAF 사용의 빈도 양상



〈그림 2〉는 주제별 HRAF 코드의 사용 빈도를 나타낸 것이다. 주제별 분류에서 가장 많이 사용된 주제는 800~890 범위에 있다. 위의 그래프를 보면 800번대 주제 중 가장 많은 코드는 831 남녀간의 사랑에 해당하는 코드로 임, 이별 등의 주제어에 831번을 부여하였다. 830번대의 대분류는 성(SEX)이고, 중분류 831번은 성적 특질(Sexuality)이며 김소월 시에 적용한 소분류 색인어는 사랑(로맨스)이다. 소분류 색인어는 별도의 번호 코드가 없으므로 중분류의 번호 코드가 부여되었다. 단순히 그래프를 살펴보아도 김소월 시를 관통하는 주제가 ‘이별의 정한’에 해당함을 확인할 수 있다.

〈표 3〉 주제별 HRAF 사용의 빈도 양상

코드	개수	코드(문화코드: 개수)
100 ~ 199	7	132(기후:1), 152(감정:1), 153(태도:1), 167(이주:2), 170(혁명:1), 186(민족성:1)
200 ~ 299	1	226(뉴스:1)
300 ~ 399	10	361(고향:10)
400 ~ 499	4	428(계승:1), 460(노동:1), 461(노동과 여가:1), 484(여행:1)
500 ~ 599	15	513(수면:1) 538(문학:1), 551(개인이름:1), 563(유랑:2), 565(농민:1), 577(윤리:2), 580(결혼:2) 581(부부사랑:2), 587(재혼:1), 590(가족:1), 593(부모:1)

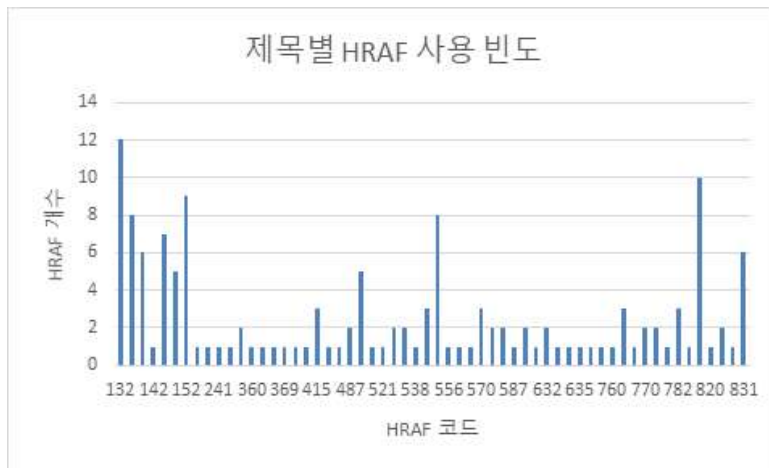
600 ~ 699	5	632(마을:1), 633(도시:1), 636(식민지:2), 669(혁명:1)
700 ~ 799	19	735(가난:2), 741(신뢰:2), 761(삶과죽음:7), 765(사별:2), 770(불교:2), 776(영혼:1), 778(신성한장소:1), 782(기도:2)
800 ~ 899	83	805(시간:1), 812(존재:4), 820(자연:4), 831(사랑:61), 839(기타성행위:1), 880(인생:12)

〈표 3〉에서 주제별 문화 코드의 작품 수를 구체적으로 확인할 수 있다. 김소월 시의 주제 중 많은 빈도를 차지하는 것을 살펴보면 831(사랑:61), 880(인생:12), 361(고향:10), 761(삶과죽음:7)의 빈도수가 높고, 사랑, 인생, 고향, 삶과 죽음 등의 주제가 많음을 코드 분류를 통해 확인할 수 있다. 예를 들어 김소월의 시 작품 중 가장 널리 알려진 〈진달래꽃〉은 831(사랑)의 시로 분류되며, 〈초혼〉은 765(사별), 〈길〉은 563(유랑)이 부여된다.

3.2. 제목별 문화 코드

주제와 제목 간의 관계를 살피기 위해 시의 제목은 각 어휘의 일반 의미에 해당하는 코드를 부여하였다. 코퍼스의 구성이 어절 단위로 되어 있고 분류 코드도 어절 단위로 부여하고 있기 때문에 여러 어절로 이루어진 제목은 1차적으로 핵심어에 코드를 부여하고, 2차, 3차 순으로 어휘에 코드를 부여하였다. 그러나 분류 체계의 일관성을 유지하기 위해서 제목의 핵심어 코드만을 반영하였다. 김소월 시의 제목에 코드를 부여한 결과는 다음과 같다.

〈그림 3〉 제목별 HRAF 사용의 빈도 양상



〈그림 3〉은 제목별 HRAF 문화 코드의 사용 빈도를 나타낸 것이다. 제목별 분류에서 가장 많은 코드는 100번대와 500번대의 대범주이다. 100번대 주제 중 가장 많은 코드는 132(기후:12)로 계절과 날씨 등이 해당한다. 그 다음으로 많이 사용된 것은 805(시간:8), 552(동물·사물이름:8), 등이다. 132번에는 「가을 아침에」, 「가을 저녁에」, 「구름」, 「눈」, 「눈 오는 저녁」, 「바람과 봄」, 「봄밤」, 「봄비」, 「비단 안개」, 「여름의 달밤」, 「오는 봄」, 「오시는 눈」이 있다. 805번에는 「면後日」, 「밤」, 「새벽」, 「예전엔 미처 몰랐어요」, 「자나 깨나 앓으나 서나」, 「저녁때」, 「해 넘어가기 前 한참은」, 「해가 山마루에 저물어도」가 있다. 552번은 「개미」, 「귀뚜라미」, 「닭 소리」, 「닭은 꼬꾸요」, 「부엉새」, 「접동새」, 「제비①」, 「제비②」가 있다.

〈표 4〉 제목별 HRAF 사용의 빈도 양상

코드	개수	중분류 코드(개수)
100 ~ 199	49	132(기후:12), 133(지형과 지질:8), 137(식물:5), 142(신체:2), 150(행동과 성격:7), 151(감각과 지각:5), 152(동인과 감정:9), 157(성격특성:1)
200 ~ 299	3	226(낚시:1), 241(경작:1), 276(기분전환 및 비치료약물:1)
300 ~ 399	7	342(주거:2), 360(정착자:1), 363(거리와교통), 365(공공시설:1), 369(도시와농촌생활), 373(빛:1)
400 ~ 499	8	410(도구 및 기구:1), 415(기구:3), 428(상속:1), 436(교환매체:1), 487(경로:2)
500 ~ 599	36	513(수면:5), 516(자세:1), 521(대화:1), 524(게암:2), 530(예술:2), 538(문학:1), 551(개인이름:3), 552(동물과 사물 이름:8), 556(부의 축적), 560(사회계층화:1), 562(성별:1), 570(대인관계:3), 577(윤리:2), 580(결혼:2), 587(재혼:1), 593(가족관계:2)
600 ~ 699	7	602(친족관계:1), 632(마을:2), 633(도시:1), 634(1), 635(지구:1), 675(계약:1)
700 ~ 799	14	740(건강과복지:1), 760(죽음:1), 761(삶과 죽음:3), 764(장례:2), 770(종교:2), 777(운과 기회:2), 782(기도와 희생:3)
800 ~ 899	20	802(숫자:1), 805(시간:8), 820(자연과사람:1), 821(민족기상학:2), 824(민족식물학:2), 831(상:6)

〈표 4〉에서 제목별 문화 코드의 작품 수를 구체적으로 확인할 수 있다. 김소월 시의 제목 중 많은 빈도를 차지하는 것을 살펴보면 831(사랑:61), 552(동물과 사물 이름:8), 133(지형과 지질:8), 137(식물:7)이 높은 빈도를 차지하는 것을 알 수 있다. 사랑, 인생, 고향, 삶과 죽음 등의 주제가 동물과 식물 등에 투영되었음을 코드 분류를 통해 확인할 수 있다. 예를 들어 〈진달래꽃〉은 의 주제는 831(사랑), 제목은 824(민족식물학)이며, 〈초혼〉은 주제는 765(사별), 제목은 764(장례), 〈길〉은 주제는 563(유랑), 제목은 487(경로)가 부여된다.

3.3. 어휘별 문화 코드

주제와 제목의 어휘에 문화 코드를 부여한 후 어휘에 문화 코드를 부여하였다. 시어는 일상적이고 지시적인 어휘에 다양한 비유적 의미를 함축한다. 이러한 이유로 시어에 문화 코드를 부여할 때 문제가 되었다. 어휘에 분류 코드를 부여하는 과정에서 1차적으로는 일상적이고 지시적인 기본 의미에 문화 코드를 부여하고 동일한 어휘에 비유적이고 상징적 의미를 지니는 의미를 2차적으로 부여하였다. 그러나 본 연구의 목적은 작품의 해석이 아닌 어휘의 분류에 있고, 분류의 일관성을 유지하기 위해서 어휘의 기본 의미에 문화 코드를 부여한 것만을 대상으로 하였다. 김소월 시 어휘에 코드를 부여한 결과는 다음과 같다.

〈그림 4〉 어휘별 HRAF 사용의 빈도 양상



〈그림 4〉는 어휘별 HRAF 사용의 빈도를 나타낸 것이다. 어휘별 분류에서 가장 많은 비중을 차지하는 것은 800번대 100번대이다. 800번대에서 가장 많은 코드는 805(시간:457회)이다. 805(시간)의 하위 범주에는 개념, 계기, 계절, 기간, 단위, 요일, 작금, 전후, 절기, 주야, 지시가 있다. 이 중에서 가장 많은 것은 주야(805:124회)로 가을밤00ng, 저녁00ng 등이 있고, 때01ng, 세월02ng, 계절을 나타내는 봄날00ng 등이 있다. 다음으로 많이 사용된 것은 133번(지형과지질:209)이다. 133번의 하위 범주에는 산야, 하천, 호소, 기타, 영역, 위치, 지역명칭, 도서, 바다, 해변이 있다. 이 중에서도 가장 많이 사용된 어휘는 산야(133:114회)로 강산01ng, 난별02ng, 땅01ng, 산01ng, 만산편야00ng 등이 있다.

〈표 5〉 어휘별 HRAF 사용의 빈도 양상

코드	개수	중분류 코드(개수)
100 ~ 199	1672	103(지명:8), 130(지리:39), 131(위치:311), 132(기후:211), 133(지형과 지질:219), 134(휴:19), 135(광물자원:6), 136(동물:28), 137(식물:149), 140(인간:57), 141(인체측장:3), 142(신체설명:180), 147(생리학적데이터:28), 150(광물자원:104), 151(감각과 지각:106), 152(동인과 감정:102), 153(행동수장:6), 157(성격특성:2), 190(언어:94),
200 ~ 299	99	201(제스처 및 표지:5), 203(뉴스 및 정보:1), 206(전화 및 통신:1), 210(기록:1), 226(뉴스:4), 241(경작:13), 260(식품소비:4), 265(식품서비스산업:1), 271(물과 갈증:29), 273(알코올 음료:2), 276(기분전환 및 비치료약물:3), 290(의류:15), 291(일반복장:15), 292(특수복:2), 293(도구:3)
300 ~ 399	126	312(물공급:2), 316(채광 및 채석), 340(구조물:20), 342(주거:37), 344(공공구조물:2), 356(개실관리:2), 363(거리와 교통:18), 364(쓰레기처리 및 위생시설:2), 365(공공시설:5), 368(기타 시설:2), 370(에너지 및 전력:2), 372(불:11), 373(빛:21), 376(수력:1)
400 ~ 499	174	410(도구 및 기구:8), 411(무기:3), 412(일반도구:16), 415(기구:57), 432(사고 팔기:1), 435(가격과 가치:5), 436(교환매체:2), 437(교환거래:1), 454(저축과 투자:1), 461(노동과 여가:1), 463(직업전문화:8), 464(노동공급과 고용:1), 482(짐 운반:13), 484(여행:7), 487(경로:44), 491(고속도로와 교량:6)
500 ~ 599	436	501(보트:16), 504(항마시설:1), 512(일상:5), 513(수면:61), 515(개인 위생:1), 516(자세:14), 524(게임:6), 526(스포츠:4), 530(예술:2), 532(대표예술:2), 533(음악:30), 534(악기:4), 535(댄스:4), 537(운반:2), 538(문학:10), 550(개별화 및 이동상:1), 551(개인이름:12), 552(동물사물의 이름:83), 553(이름지정:10), 556(부의축적:1), 560(사회계층:35), 561(연령:2), 562(성별:24), 570(대인관계:44), 571(사회관계:3), 572(우정:16), 574(방문·환대), 577(윤리:10), 578(적대:1), 580(결혼:6), 585(결혼식:1), 587(재혼:2), 590(가족:4), 593(가족관계:19)

600 ~ 699	135	601(친족용어:1), 602(친족관계:7), 607(형제자매:13), 618(문중:1), 630(영토조작:9), 632(마을:27), 633(도시:16), 635(지방:53), 672(책임:2), 675(계약:4), 690(정의:2)
700 ~ 799	80	713(병가:4), 729(전쟁참전 용사:1), 752(신체부상:2), 761(삶과 죽음:12), 763(사망:2), 764(장례:10), 765(애도:1), 770(종교:4), 771(종교일반:1), 776(영과 산:16), 778(신성한 물건과 장소:3), 780(종교관습:1), 781(종교체험:2), 782(기도와 희생:2), 787(계사와 점술:1), 788(의식:11), 792(예언자와 수행자:1)
800 ~ 899	726	800(숫자와 측정:48), 802(숫자:5), 805(시간:457), 812(철학:4), 820(자연과 사람:1), 821(민족식물학:121), 822(민족물리학:35), 824(민족식물학:27) 831(성적 특질:4), 850(영·유아:4), 856(발달과 성숙:1), 870(교육:1), 880(인생:18), 883(청소년 활동:1), 886(노년:6)

〈표 5〉에서 어휘별 문화 코드의 수를 구체적으로 확인할 수 있다. 어휘별 HRAF 사용을 대분류와 중분류 체계로 분류했을 때 다음과 같다. 김소월 시의 어휘 중 많은 빈도를 차지하는 것을 살펴보면 131(위치:311), 132(기후:211), 133(지형과 지질:219)이 높은 빈도를 차지하는 것을 알 수 있다. 예를 들어 「진달래꽃」은 의 주제는 831(사랑), 제목은 824(민족식물학)이며, 사용된 어휘는 영변(133:지형과지질), 약산(133:지형과지질), 진달래꽃(824:민족생물학) 등이다. 「초혼」 주제는 765(사별), 제목은 780(종교관습), 어휘는 이름(553:이름지정), 설움(152:감정), 사람(140:인간) 등이다. 〈길〉의 주제는 563(유랑), 제목은 487(경로), 어휘는 정주(635:지방), 관산(635:지방), 길(487:경로) 등이 부여된다.

4. 어휘 분석의 실제

김소월의 시는 한국의 전통 민요의 울조와 토속적인 언어를 통해 한국적인 서정시의 전통을 확립하고 있다. 그의 시세계 전반에 나타나는 주제는 그리움, 외로움, 쓸쓸함 등의 이별의 정한의 감정이며 시적 화자 이러한 감정은 자연물을 대상으로 투영한다. 또한 그의 시의 그리움은 단순히 임에 대한 그리움으로 그치는 것이 아니라 고향에 대한 그리움으로 이어지며 당대 현실에 대한 민족의 마음을 대변하기도 한다. 아래에서는 김소월의 작품 중 교육텍스트로 가장 많이 다루어지는 작품들을 예로 들어 부여된 문화 코드의 양상을 구체적으로 살펴보도록 하겠다.

4.1. 이별의 정한 「진달래꽃」

「진달래꽃」은 1922년 「개벽」에 발표한 김소월의 작품으로 ‘진달래꽃’이라는 향토적 소재와 민요조의 리듬으로 서정적 효과를 자아낸다. 여성적인 경어체와 비애의 정서를 고조시키는 이 작품은 이별의 정한을 형상화시킨 김소월의 대표적이며 한국에서 가장 사랑받는 시 중 하나이다. 「진달래꽃」 텍스트에 부여된 문화 코드의 양상은 다음과 같다.

주제: 이별의 정한(831:사랑-이별)

제목: 진달래꽃(824:식물_화종)

나 보기가 역겨워

가실 때(805:시간-개념)에는

말없이 고이 보내 드리오리다.

영변(寧邊)(133:지역-명칭)에 약산(藥山)(133:지역-명칭)

진달래꽃(137:식물-화종)

아름(800:개념-관념) 따다 가실 길에 뿌리오리다.

가시는 걸음걸음(512:인간-활동)
 놓인 그 꽃(824:식물-화종)을
 사뿐히 즈려밟고 가시옵소서.

나 보기가 역겨워
 가실 때(805:시간-개념)에는
 죽어도 아니 눈물(147:신체-분비) 흘리오리다.

문화 코드가 부여된 양상을 살펴보면 「진달래꽃」의 주제는 이별(831), 제목은 진달래꽃(824)의 코드가 부여되었다. 이별의 감정을 투영한 진달래꽃은 식물 중에서는 토속적인 식물에 해당하므로 식물과는 별도의 영역인 824(민족식물학)으로 분류하였다. 이별의 정한을 드러내는 어휘들은 진달래꽃01ng, 눈물01ng, 이별의 과정은 길01ng, 걸음걸음00ng, 아픔01ng 등에서 나타난다고 할 수 있다. 특히 영변02nm, 약산04nm은 평안북도 위치하는 지명과 산으로 진달래꽃과 함께 토속적인 정서를 더욱 짙게 나타낸다.

진달래꽃은 사랑과 희생, 눈물은 슬픔을 상징하고 있는데 이러한 비유적인 의미는 2차 작업을 통해 이루어진다. 꽃을 뿌리는 행위는 불교의 전통의식인 이별에 대한 방어적 표현임과 동시에 불교에서의 전통의식인 산화공덕(散花功德)을 연상시킬 수 있다. 부처님이 지나가시는 길에 꽃을 뿌려 그 발길을 영화롭게 한다는 축복의 의미로 (770:불교)의 문화 코드를 추가로 부여할 수도 있다.

4.2. 갈 곳 없는 떠돌이의 정서 「길」

「길」은 정처 없이 떠도는 시적 화자의 서글픔 심정과 고독을 형상화하는 작품이다. 자문자답의 형식은 화자의 고독한 처지를 더욱 강조한다. 할아버지의 사업실패로 가세가 기울어 고향을 떠날 수밖에 없었던 소월의 개인사적인 차원에서도 식민지 치하의 우리 민족의 차원에서도 「길」은 삶의 터전인 고향을 상실하고 떠도는 유랑의 정서를 효과적으로 형상화하고 있다. 「길」 텍스트에 부여된 문화 코드의 양상은 다음과 같다.

주제: 유랑(563)
 제목: 길(487:이동-통로)

어제(805:시간-작금)도 하로밤(805:시간-기간)
 나그네(484:인간) 집(342:주거-가옥)에
 가마귀(552:동물-조류) 가왯가왯 울며 새었소.
 오늘(805:시간-작금)은
 또 몇 십 리(十里)
 어디로 갈까.
 산(山)(133:산천-산야)으로 올라갈까.
 들(133:산천-산야)로 갈까.
 오라는 곳이 없어 나는 못 가오.
 말(190:개념-언어) 마소, 내 집(342:주거-가옥)도
 정주(635:지역-명칭) 광산(定州郭山)(635:지역-명칭)
 차(車)(501:이동-수단) 가고 배(501:이동-수단) 가는 곳이라도.

여보소, 공중(131:장소-영역)에
 저 기러기(552:동물-조류)
 공중(131:장소-영역)엔 길(487:이동-통로) 있어서 잘 가는가?
 여보소, 공중(131:장소-영역)에
 저 기러기(552:동물-조류)
 열십자(十字)(802:개념-숫자) 복판(131:장소-위치)에 내가 썼소.
 갈래갈래 갈린 길(487:이동-통로)
 길(487:이동-통로)이라도
 내게 바이 갈 길(487:이동-통로)은 하나(802:개념-숫자) 없소.

문화 코드가 부여된 양상을 살펴보면 「길」의 주제는 유랑(563), 제목은 길(487:이동-통로)의 코드가 부여되었다. 정치 없이 떠도는 유랑이 감정을 갈 곳 없는 길에 비유하고 있다. 화자의 암담한 마음은 가마귀(552:동물-조류), 고향에 대한 그리움은 기러기(552:동물-조류), 방향성을 잃은 모습은 열십자(十字)(802:개념-숫자) 복판(131:장소-위치)으로 나타내고 있다. 삶의 터전인 고향의 모습을 나타내는 어휘는 집(342:주거-가옥), 정주(635:지역-명칭) 곽산(定州郭山)(635:지역-명칭)이라는 구체적 지명을 더해 향토적인 정서를 불러일으키고 그리움의 정서를 구체화한다.

4.3. 생과 사, 죽음 「초혼」

「초혼」은 입을 잃은 비탄의 절정을 보여주는 작품으로 죽은 자의 혼령을 불러내는 고복의식을 모티브로 하고 있다. 불행한 결혼 생활을 하다가 죽음을 맞이한 소월의 연인 오순을 위해 지은 시로 알려져 있는데 삶과 죽음의 거리를 뛰어넘지 못하는 비통함과 공허함이 느껴진다. 「초혼」에 부여된 문화 코드의 양상은 다음과 같다.

주제: 사별-비탄(765)
 제목: 초혼(764: 장례)

산산이 부서진 이름(553:개념-기능)이여!
 허공(821:자연-천체)중에 헤어진 이름(553:개념-기능)이여!
 불러도 주인(570:인간-기타)없는 이름(553:개념-기능)이여!
 부르다가 내가 죽을 이름(553:개념-기능)이여!

심중(150:인간-정신)에 남아있는 말(190:개념-언어) 한마디(190:개념-언어)는
 끝끝내 마저하지 못하였구나.
 사랑하던 그 사람(140:인간)이여!
 사랑하던 그 사람(140:인간)이여!

붉은 해(821:자연-천체)는 서산마루(131:장소-위치)에 걸리었다.
 사슴(552:동물-짐승)의 무리(800:개념-기능)도 슬피운다.
 떨어져 나가 앓은 산(133:산천-산야) 위(131:장소-위치)에서
 나는 그대의 이름(553:개념-기능)을 부르노라.

설움(152:감정)에 겹도록 부르노라.
 설움(152:감정)에 겹도록 부르노라.

부르는 소리는 비켜가지만 하늘(821:자연-천체)과 땅(133:산천-산야) 사이(131:장소-위치)가 너무 넓구나.

선재로 이 자리에 돌(134:산천-암석)이 되어도
부르다가 내가 죽을 이름(553:개념-기능)이여!
사랑하던 그 사람(140:인간)이여!
사랑하던 그 사람(140:인간)이여!

문화 코드가 부여된 양상을 살펴보면 「초혼」의 주제는 사별-비탄(765), 제목은 초혼(764:장례)의 코드가 부여되었다. 사랑하는 사람을 잃은 비통한 심정을 '초혼'이라는 상례를 빌어 표현하고 있다. 초혼은 사람이 죽었을 때 망자가 생전에 입던 옷을 들고 지붕에 올라가서 망자의 이름을 부르는 의식이다. 이는 남은 자의 처절한 미련의 몸부림이라 할 수 있다. 화자의 공허하고 비통한 마음은 허공(821:자연-천체), 반복되는 이름(553:개념-기능), 사람(140:인간)에서 보여진다. 말(190:개념-언어) 한마디(190:개념-언어) 못해 설움(152:감정)이 가득한 심중(150:인간-정신)은 허공과 대비적이다. 죽은 이에게 더 가까이 하고자 산(133:산천-산야) 위(131:장소-위치)에서 이름을 불러 보지만 하늘(821:자연-천체)과 땅(133:산천-산야) 사이(131:장소-위치)는 닿을 수 없는 거리다. 해(821:자연-천체)는 서산마루(131:장소-위치)있고 화자는 돌(134:산천-암석)이 되어 기다리겠다고 한다. 돌은 일반적인 의미의 134(암석)의 코드를 부여하였다. 그러나 여기에서의 돌은 남편을 기다리다가 돌이 되었다는 망부석 설화를 배경으로 있어 538(문학-설화)의 코드를 추가로 부여할 수 있다.

이밖에도 김소월의 시에서 생과 사, 그리고 죽음이 많이 다루어지고 있는데 「무덤」(761), 「사노라면 사람은 죽는 것을」(761), 「생과 돈과 死」(761), 「생과 死」(761), 「하다못해 죽어 달 내가 올라」(761)이 있다.

5. 결론

지금까지 시 텍스트의 어휘에 문화 코드를 부여하는 방안을 구체적으로 살펴보았다. 분석 대상은 『한국현대시 코퍼스』(KoPoCo)에 수록된 「김소월 코퍼스」를 대상을 하였다. 「김소월 코퍼스」는 『진달래꽃』 수록 작품 126편, 『소월시초』에 수록 작품 19편으로 총 145편, 시어 수는 8,203개로 이 중 명사는 3,448개이다.

「김소월 코퍼스」에 문화 코드를 부여하는 과정은 다음과 같이 진행되었다. 먼저 HRAF의 코드를 김소월 시 텍스트의 주제에 부여하였다. 다음으로 시 텍스트의 제목에 HRAF 분류 코드를 부여하였다. 마지막으로 시 텍스트의 어휘 단위에 HRAF 문화 분류 코드를 부여하였다. 그리고 「진달래꽃」, 「길」, 「초혼」을 예로 들어 문화 코드가 부여된 양상을 살펴보았다.

본 연구는 현대시 텍스트에 나타나는 문화 항목들을 세계적으로 활용되고 있는 문화 분류 코드를 활용하여 분석함으로써 한국문학교재나 학습자료의 기초 자료를 제공하는 방안을 마련하고자 하였다. 일상어의 의미와 시 작품 속에서의 의미 차이에서 학습자가 느끼는 문화적 공백감을 최소화하여 상호문화적인 관점에서 시교육을 통한 교육적 효과를 높일 수 있을 것이다. 또한 텍스트의 수정하거나 각주를 제공, 번역에 대한 참고 자료로 활용하는 등 현대시 텍스트의 2차 가공을 용이하게 할 것이다. 시어는 그 함축성으로 인하여 특히 어려움이 있는데 문화 어휘를 분석하여 함축된 의미를 제시한다면 오역을 그만큼 줄일 수 있을 것이며, 상호문화 비교를 위한 지점을 제시하여 문화 비교의 일종의 네비게이션 역할을 할 수 있을 것이다.

그럼에도 분석 대상을 명사로 한정하고 어휘의 기본 의미에만 문화 코드를 반영하여 함축적인 시어의 특징을 충분히 반영하지 못하였다. 시어의 특징에 맞는 2차, 3차 의미에 추가로 문화 코드를 반영하여 시 텍스트에 적합한 문화 분류 체계를 발전시켜 나가는 것은 후 연구과제로 남긴다.

■ 참고문헌 ■

- 김병선, 『한국현대시코퍼스(Korean Modern Poetry Corpus)』, 한국학중앙연구원.
Human Relations Area Files(HRAF) [online][cited 2021.06.03.], <https://hraf.yale.edu/>
- 김병선 외(2007), 「한국 현대시어 빈도 사진」, 한국문화사.
- 김병선(2004), 「한국 현대시 데이터베이스의 구성과 그 활용방안」, 『한국언어문학』 53, 한국언어문학회.
- 김병선(2015), 『현대시와 문학 통계학』, 한국학중앙연구원 출판부. pp.349-386.
- 김예리나(2018), 「학습자 경험을 활용한 한국어 상호문화교육 연구」, 서울대학교 석사학위 논문.
- 김유리(2019), 「한국현대시를 활용한 상호문화 지식 교육 연구-중국인 중·고급 학습자의 비교를 중심으로」, 서울대학교 석사학위 논문.
- 김중섭(2011), 『국제통용 한국어교육 표준 모형 2단계』, 국립국어원.
- 김진호(2012), 「외국어로서의 한국어교육에서 시문학을 활용한 한국문화 교육 연구: 김소월의 작품을 중심으로」, 단국대 박사학위 논문.
- 김창근(2015), 「상호문화주의의 원리와 과제: 다문화주의의 대체인가 보완인가?」, 『윤리연구』 103, 한국윤리학회. p.188.
- 김해옥·김하수(2006), 「다문화시대의 한국문학을 통한 언어문화 콘텐츠 구축에 관한 연구」, 『현대문학의 연구』 28, 한국문학연구학회.
- 김혜민(2015), 「한국어 상호문화교육을 위한 교수 전략 연구」, 서울대학교 석사학위 논문.
- 노금숙(2010), 「다문화사회의 한국어수업: 노래와 시의 상호텍스트성을 활용한 한국어교육-김소월 시와 마야의 “진달래꽃”을 중심으로-」, 국어교육학연구 38, 국어교육학회.
- 박서영(2019), 「한국어문학 교육을 위한 김소월 시 중국어 번역 연구」, 전남대학교 박사학위 논문.
- 박수환(2004), 「구례 류씨가 생활일기 『기어(紀語)』의 전산화 방안 연구-HRAF를 이용한 생활사 연구의 기초연구자료 확보를 목적으로-」, 『고문서연구』 25, 한국고문서학회.
- 박제은(2019), 「상호문화적 접근에 의한 현대시 활용 한국어교육 연구-중급 한국어 학습자를 중심으로」, 경인교육대학교 석사학위 논문.
- 신윤경(2014), 「한국문학 읽기를 위한 어휘 주석 연구-사회·문화적 배경지식 및 맥락 이해 추가 설명을 중심으로」, 『한국언어문화학』 11-1, 국제한국언어문화학회.
- 양민애 (2008), 「한국어 교사 교육에서의 이문화(異文化) 교육의 필요성과 방향」, 『이중언어학』 38, 이중언어학회.
- 오지혜(2015), 「한국어 학습자의 번역텍스트 분석- 시 텍스트를 대상으로-」, 『언어와 문화』 11-11, 한국언어문화교육학회, pp.87-89.
- 오지혜(2018), 「재외동포 수필에 나타난 문화 어휘 연구」, 『교육문화연구』 24-1, 인하대학교 교육연구소.
- 윤여탁(2010), 『외국어로서의 한국문학교육』, 한국문화사.
- 이다니엘(2019), 「외국인 학습자 위한 한국 소설 이독성 측정 모형 개발 연구-문화적 공백 개념을 중심으로」, 경희대학교 석사학위 논문.
- 이자옥(2002), 「김소월 시에 나타난 자연 이미지 연구」, 연세대학교 석사학위 논문.
- 장노현(2014), 『소월이 지금 나에게로 왔다』, 글누림.
- 전미순·이병운(2011), 「한국어 문화어휘에 관한 일고찰: 문화어휘 분석을 중심으로」, 『언어와 문화』 7-1, 한국언어문화교육학회.
- 전홍(2012), 「한국어 학습자의 상호작용적 시 읽기 양상 연구 -문화 이해를 중심으로」, 『국어교육연구』 29, 서울대학교 국어교육연구소.
- 조수진(2009), 「한국어교육의 간문화 교육 연구-정의적 영역에서의 접근」, 『한국언어문화학』 6-2, 국제한국언어문화학회.
- 조수진(2013), 「한국어교육에서 김소월 시에 나타난 이별의 “정서” 교육 방안 -〈진달래꽃, 초혼, 접동새〉를 중심으로-」, 한국어교육 24-2. 국제한국어교육학회.
- 조옥라(1993), 「한국인류학에서의 비교연구와 HRAF」, 『비교문화연구』 1, 일신사.
- 황혜란(2017), 『한국문학교육용 텍스트의 이독성 연구』, 한국학중앙연구원 박사학위 논문.

황혜란(2021), 「소설 텍스트의 문화 항목 분류 체계 연구 -HRAF 분류 코드 적용을 중심으로-」, 한국언어문학 116집, 한국언어
문학회. p.269

J.Kristev, Toril Moi(ed)(1986), The Kristeva Reader, New York: Columbia Up. pp.36-37.

〈토론〉 “한국 현대시 텍스트의 문화 어휘 분석 연구
-김소월 시 코퍼스를 중심으로-”에 대한 토론문

김금숙(상지대학교)

이 연구는 『한국현대시 코퍼스(KoPoCo)』를 기반으로 김소월 시 텍스트의 어휘에 나타나는 문화 항목을 체계적으로 분류하고자 한 연구입니다. 논문에서는 김소월 시의 어휘 중 명사에 대해 HRAF 문화 코드를 기준으로 주제, 제목, 어휘의 순서로 문화 분류 코드를 부여하고 있습니다. 시 텍스트의 어휘에 문화 코드를 부여한다는 것이 다소 생소하지만 이러한 작업이 전반적으로 이루어지면 현대시를 활용한 한국 문화 교육이나 현대시 번역 작업 등을 위한 기초 자료로서 유용하게 쓰일 수 있을 것 같습니다. 그러한 점에서 연구자께서 수행하신 작업이 매우 의미 있는 것이라고 생각하면서 연구자의 논지에 대부분 동의하면서 발표문 읽었습니다. 다만 토론자의 소임을 다하고자 발표문을 읽으면서 들었던 몇 가지 궁금증을 중심으로 토론문을 작성해 보고자 합니다.

1. 제목에서, 그리고 본문 기술에서 한 번 ‘문화 어휘’라는 용어를 사용하셨는데, 이것을 어떤 의미로 사용하신 것인지 여쭙보고 싶습니다. 문화적인 의미를 담고 있는 문화 어휘가 따로 있다고 보신 것인지, 아니면 시에 쓰인 모든 어휘를 문화 어휘로 지칭한 것인지 궁금합니다. 이에 대한 기술이 논문이나 주석에 포함되면 좋을 것 같습니다.
2. 논문에서 현대시 텍스트에 나타나는 문화 항목들을 문화 분류 코드를 활용하여 분석함으로써 한국문학 교재나 학습 자료의 기초 자료를 제공하는 방안을 마련하고자 한다고 밝히고 있는데요. 한국문학 교재나 학습 자료를 만드는 데 본 연구 결과가 어떻게 활용될 수 있는지에 대해 구체적인 예시를 들어 주시면 좋겠습니다.
3. 논문을 읽으면서 흥미롭기는 했지만, 분석 대상이 ‘시’이고 분류 코드를 부여하는 것이 ‘시어’라는 점이 조금 간과된 것은 아닌가 하는 의문이 들었습니다. 연구자가 말씀하신 대로 HRAF의 문화 분류 코드를 활용하면 문학 텍스트에 쓰인 어휘의 의미를 항목별로 분석할 수 있다는 것에는 전적으로 동의합니다. 시의 주제와 제목에 각각 문화 분류 코드를 부여하여 제목과 그 속에 함축되어 있는 문화적 요소를 살펴볼 수 있다는 것에도 동의하고요. 그러나 “일상어의 사용이 시 작품 속에서 문화적으로 어떻게 함축되어 있는지를 파악할 수 있을 것”이라는 부분은 잘 해결되지 않은 듯합니다. 시의 제목뿐만 아니라 시에 쓰인 어휘들은 모두 일상어와는 달리 함축성을 지니며, 시어를 통해 구축되는 이미지들 속에는 시인의 생각이나 정서가 다양하게 형상화되어 있습니다. 그런데 3.3에서는 “어휘에 분류 코드를 부여하는 과정에서 1차적으로는 일상적이고 지시적인 기본 의미에 문화 코드를 부여하고 동일한 어휘에 비유적이고 상징적 의미를 지니는 의미를 2차적으로 부여하였다. 그러나 본 연구의 목적은 작품의 해석이 아닌 어휘의 분류에 있고, 분류의 일관성을 유지하기 위해서 어휘의 기본 의미에 문화 코드를 부여한 것만을 대상으로 하였다.”

라고 기술하고 있습니다. 그리고 실제 연구 결과들을 살펴보아도 분석 대상 어휘가 나름의 특수성을 지니는 ‘시어’라는 것을 크게 고려하지 않은 것처럼 보입니다.

4. 연구자께서 “세계의 다양한 문화와의 상호문화적 관점에서 한국 현대시 작품을 비교 분석하고 작품과 작품 사이를 연계할 수 있는 일관된 기준과 기초 자료를 제공하는 것이 필요하다.”고 하셨는데요. 어휘에 코드를 부여하는 것은 일종의 태그를 다는 것이므로 방대한 어휘를 정리하여 기초 자료를 제공하는 것은 의미 있다고 생각합니다. 그러나 이러한 분류 작업이 주석을 달아서 어휘에 대한 문화적 맥락을 설명하는 것을 대신할 수 있을지는 잘 모르겠습니다. 또 이러한 분류가 문화 항목 선정의 일관된 기준을 제공하는 데 어떻게 기여할 수 있을지도 잘 그려지지 않습니다.