

인간번역과 기계번역의 여성어 · 남성어 비교 연구: 한일 번역본 『82년생 김지영』의 문말(文末)표현을 중심으로*

조승연

(이화여자대학교)

Cho, Seungyeon. (2023). Comparative study of feminine and masculine languages of human translation and machine translation: Focused on the sentence-end expressions in the Japanese translations of the Korean novel 『Kim Jiyoung, Born 1982』. *The Linguistic Association of Korea Journal*, 31(1), 45-63. This study is a case study of machine translation in literature, which is performed by applying two types of machine translations (Google and Papago) to the novel 『Kim Jiyoung, Born 1982』. The purpose of this study is to examine the applicability of machine translation in literature. To this end, we have looked at the usage patterns of the sentence-end expressions in female and male characters of the human-translated Japanese novel and how those sentence-end expressions appeared in machine translations. This study is not for evaluating mistranslations and performance in machine translation in literature. As a research method, we constructed the texts of the Korean novel and the Japanese translation as a corpus, and performed a cross-analysis (chi-square test) with the analysis data. Also, we contrasted human and machine translations by analyzing the sentence-end expressions of the characters according to 有泉 (2016)'s seven classification methods of sentence-end expressions. The study results show that both types of machine translations translated the female characters' sentence-end expressions to a level close to human translation.

주제어(Key Words): 문학기계번역(Literature translation), 여성어(female language), 남성어(male language), 코퍼스 분석(corpus analysis)

* 본 연구를 위해 학술적 글쓰기를 지도해 주신 이화여자대학교 통역번역대학원 이지은 교수님과 통계 자문을 주신 충남대학교 언어학과 이용훈 교수님께 깊은 감사의 인사를 올립니다.

1. 서론

2016년 11월, 구글이 새로운 ‘Google 번역’ 서비스를 선보이며 심층학습 기반의 ‘신경망 기계번역(Neural Machine Translation, NMT)’이 등장하였다. 이후 NMT에 대한 기대감이 고조되었고 NMT 관련 연구는 오류 분석을 시작으로 기계번역의 효율성과 경제성을 인정하고 기계번역의 도입을 위해 프리에디팅, 포스트에디팅 분야로 발전하였다(이해미, 2022, p. 129). 이러한 연구들이 진행되면서 부족한 빅데이터 정보량과 품질을 이유로(안미영, 2020, p. 103) 기계번역에 대한 부정적인 의견이 제시되었지만, 한편으로 NMT가 자연스럽게 정확한 번역을 제공하는 것으로 알려지면서 곧 인간 번역사를 능가할 것이라는 의견도 과거보다 설득력을 얻게 되었다(김슬기, 2018, p. 3). NMT는 기계학습을 통해 눈부신 속도로 진화하였고 기계번역의 품질이 높아지면서 문학기계번역의 가능성을 타진하는 연구도 등장하였다. 해외 언어 산업 전문매체인 SLATOR¹⁾는 2018년 1월 19일 문학작품을 영어-카탈란어 언어 쌍으로 기계번역 연구를 진행하고 NMT 번역물 중 17~34%가 전문 번역가 품질에 버금가는 수준이었다고 보도했다.

문학기계번역의 실현 가능성을 모색해보기 위해서는 먼저 기계번역 연구가 공학과 인문학이 만나는 다학제적 연구라는 점을 전제로 해야 한다(이준호, 2019, p. 146). 기계번역을 인문학에 초점을 맞춘다면 문화적 전환(cultural turn), 다시쓰기(rewriting), 문화소(culture specific item) 등을 언급하며 문학기계번역은 불가하다는 주장을 펼칠 수 있지만, 공학적 관점에서 기계번역 엔진을 개발한 연구자들은 기계학습을 통해 더 많은 데이터를 구축한다면 기계 또한 인간에 가까운 번역을 할 수 있다고 결론을 내릴 것이다(이준호, 2019, p. 146). 따라서 두 학문이 만나는 영역에서 논의를 시작하는 것이 바람직하다. 또한 기계와 번역의 조합은 공학과 인문학의 조합이라고도 할 수 있기에 기계번역이 넘어야 할 산이라고 할 수 있는 문학번역에서 질적인 한계점을 극복하고 인문학의 장(field)에서 먼저 그 해결책을 내놓아야 한다(이해미, 2022, p. 126).

본 연구의 목적은 인간번역과 기계번역이 머지않은 미래에 공존할 수 있는 협력적 관계가 될 것이라는 전제하에 인간번역과 기계번역의 문학번역 사례를 살펴보고, 이를 통해 인공지능 번역기 시대에 기계번역의 현재를 객관적으로 설명함으로써 기계번역의 확장 가능성에 대해 고찰하는 것이다. 이에 본 연구는 기계번역의 현주소를 인간번역을 대체할 수 없는 분야 가운데 하나로 꼽히는 문학번역이라는 한정된 영역에서 살펴보고자 한다. 이를 위해 소설 『82년생 김지영』한일 번역본에 나타난 모든 등장인물 대화문에 먼저 초점을 맞추고, 일본어에서 여성성과 남성성을 잘 드러내는 문말(文末)어미를 살펴보기로 한다. 이를 바탕으로 2중(구글, 파파고)의 기계번역이 원문에 나타난 여성문말표현과 남성문말표현을 어떻게 처리하였는지 분석한 후 인간번역과 2중(구글, 파파고)의 기계번역 결과물을 비교하였다.

1) <https://slator.com/technology/machine-translates-literature-and-about-25-was-flawless-research-claims/>

2. 선행연구

2.1. 문학기계번역에 관한 연구

국내 문학기계번역의 가능성을 논의한 연구 가운데 한국어와 일본어 언어 쌍에 관한 연구는 적은 데 비해 한국어와 서구어 언어 쌍의 연구는 비교적 많이 진행되어 있다. 가령 한국어와 영어 언어 쌍의 연구를 진행한 최윤영(1996)은 문학번역에서 기계번역이 어려운 이유로 인간의 창조적이고 예술적인 능력을 기계가 가질 수 없다고 주장하였다. 이상빈(2016)은 창의적 글쓰기가 요구되는 문학작품과 같은 분야에서는 기계번역과 인간번역의 차이는 훨씬 크다는 의견을 제시하였다. 마승혜(2017)는 번역 장르 중에서도 창의력과 다면적 해석, 선택 능력 등이 요구되는 문학번역을 기계가 인간을 대체할 수 있을지에 대해 의문을 제기하면서 문학작품의 번역 능력과 기계화되기 어려운 인간 능력의 비교 분석을 통해 문학번역의 기계화는 어려울 것이라는 결론을 내렸다. 이처럼 문학기계번역에 관한 국내 연구는 창의성과 예술성을 중심으로 한 이론적 접근에 대부분 치우쳐져 있다. 또한 국내 연구의 경우 소수의 사례 연구만이 존재하며 연구에서 제시하는 데이터가 해외 연구와 비교했을 때 한정적이거나 객관성이 다소 떨어진다.

반면 해외 연구의 경우 문학기계번역의 한계를 지적하는 견해가 대부분인 국내 연구에 비해 그 가능성을 타진하는 연구 사례도 보인다. 보이트와 주라프스키(Voigt & Jurafsky, 2012)는 문학작품의 문체와 텍스트적인 문체를 기계번역이 극복할 수 없다고 지적하면서도, 문학작품의 결속성(cohesion)을 재현하는 데 있어 기계번역이 인간번역에 비해 70% 정도의 능력이 있음을 밝혔다. 베사시어와 슈바르츠(Besacier & Schwartz, 2015)는 문학작품을 포스트에디팅으로 번역한 후 번역 엔진의 반복 학습을 통해 포스트에디팅의 생산성이 향상될 수 있다고 주장한 바 있다. 토랄과 웨이(Toral & Way, 2015)는 프랑스어와 영어, 프랑스어와 이탈리아어의 언어 쌍에서 일부 문장은 인간번역과 기계번역의 결과가 동일했으며, 기계번역이 인간번역과 유사한 경우에는 포스트에디팅을 통해 문학번역의 생산성이 향상될 수 있다는 의견을 제시하였다. 이들은 2015년과 2018년에 NMT를 문학번역에 적용한 연구를 진행하였다. 먼저 2015년에 실행한 연구에서는 스페인어와 카탈란어 처럼 유사한 언어 사이에 적용해 본 결과 텍스트의 응집성(cohesion), 구문 구조, 비유적 언어 번역 등에서 많은 오류가 발생 하였다고 평가하였다(Toral & Way, 2015, p. 25). 이후 2018년도에 진행된 영어와 카탈란어를 대상으로 한 연구에서 NMT 번역이 기존 구(phrase) 단위의 기계번역(PBMT, Rule-based Machine Translation)²⁾에 비해 BLEU 자동 평가에서 11퍼센트 향상된 결과가 제시되었다(Toral & Way, 2018, p. 12).

이처럼 문학기계번역을 둘러싼 국내외의 선행연구들을 살펴보면 그 가능성과 한계를

2) 기계번역 초반에 많이 사용했던 기술로, 개발자가 직접 정의한 규칙에 따라 번역(김슬기, 2018, p. 3).

지적하는 의견이 분분하지만 한 가지 공통점은 문학번역 일부에서 기계번역의 가능성을 모색하는 초기 단계라 할 수 있다는 점이다. 따라서 본 연구는 젠더 표시 기능을 하고 낙양스에 영향을 미치는 일본어의 문말표현에 국한하여 문학기계번역의 현재를 살펴보는 것에 연구 의의를 두고자 한다.

2.2. 일본어 문말표현

일본어를 성별을 중심으로 살펴보면 여성이 일반적으로 사용하는 말과 남성이 일반적으로 사용하는 말로 구분할 수 있다(奥田, 2016, p. 41). 일본어는 언어사용에 있어 남녀차가 타 언어보다 뚜렷하다고 평가되는데 실제 언어생활에서 남성어와 여성어의 차이가 어휘, 문법 등에 나타난다. 특히 여성이 사용하는 어휘, 표현, 말투 등에서 더 많은 제약이 관찰된다. 예를 들어 남성들이 사용하는 「腹減った(Hara hetta 배 고파)」, 「水をくれ(Mizuwo kure 물 줘)」와 같이 접두사가 누락된 표현이나 명령형 표현 등을 여성이 사용할 경우 '품위가 없다', '상스럽다' 와 같은 인상을 타인에게 줄 수 있다. 이는 동일한 표현을 어떤 성별이 발화하느냐에 타인에게 주는 인상이 달라질 수 있다는 것을 의미한다. 이렇듯 일본어는 여성과 남성이 주로 사용하는 말투가 대체적으로 정해져 있는 현상을 통해 사회적으로 성별에 따라 요구되는 말투가 구분된다는 것을 알 수 있다.

문말표현이란 문말어미에 해당하는 표현으로 일본어에서 주로 술어 부분을 가리키는데 대체로 술어의 넓은 의미로 사용된다(탁성숙, 2002, p. 14). 일본어의 문말표현은 여러 가지 기능이 있으며³⁾(황수영, 2015, pp. 1-2) 다양한 문말표현과 액센트 등으로 섬세하고 미묘한 뉘앙스를 부여할 수 있다. 일본어에서 남녀차를 드러내는 요소로는 ①중조사(문말표현), ②호칭(1인칭, 2인칭, 3인칭), ③음변화(音変化), ④억양(intonation), ⑤어휘(부사어, 형용사), ⑥문법, ⑦경어, ⑧파라언어, ⑨청중의 성별 등을 꼽을 수 있는데(Ogawa, 2004, pp. 26-27) 문말표현은 이처럼 일본어에서 남녀차를 명확하게 드러내는 술어부에 해당되는 형식이다.

(1) 문말표현

- a. 韓国人だわ(Kankokujinn-dawa) 한국인이네.
- b. 韓国人だぞ(Kankokujinn-dazo) 한국인이야.

3) 첫째로 그 자체를 묘사한다. 둘째로 그 사실(사물·사건)을 긍정하는가, 부정하는가, 이미 실현했는가, 아직 실현하지 않았는가라고 하는 판단을 나타낸다. 셋째로 사실(사물·사건)에 대한 화자의 태도, 즉, 확실히 단정하는가, 의문의 여지가 있는가, 의지가 있는 표현인가 등을 나타낸다. 넷째로 화자의 청중에 대한 태도의 표현을 포함할 수 있다. 어떠한 행동을 하는 것을 요구하는가, 질문하는가, 단순히 지적하는 것인가 등을 포함할 수 있다. 그리고 화자 측과 관계된 것인가, 그렇지 않은가에 따라 나타나는 대우표현 또한 문말표현의 중요한 요소가 된다(황수영, 2015, pp. 1-2).

위와 같이 일본어는 가장 문말에 있는 자립어가 문장의 중심 요소가 되어 이 단어와 연결되는 조동사나 조사를 포함한 부분이 문장 전체의 의미에 깊이 관여하기 때문에 성별을 짐작할 수 있다⁴⁾. 그러나 이는 문법적인 구분이 아닌 일반적으로 ‘남성들이 쓰는 언어’, ‘여성들이 쓰는 언어’와 같이 막연하게 인식되며 그 특징은 인칭대명사나, 감탄사, 문말어미 등에 나타난다. 이 가운데 대화문 문말에 주로 보이는 문말표현은 주로 소설이나 영화, 드라마상에서 등장인물의 남성성과 여성성을 표현하기 위해 사용되기는 하나, 시대의 변화와 함께 실제로 명확한 젠더적 차이는 점차 희미해지고 있다. 이는 언어 사용과 관련하여 젠더 규범의 영향은 있으나 상황에 따라 어떠한 표현을 사용할지는 화자의 선택이며, 젊은 세대일수록 점차 ‘남성어’와 ‘여성어’의 구분 없이 언어 사용이 ‘중성화’ 되어 가는 것을 의미한다(水本, 2005, p. 25). 그럼에도 불구하고 여전히 일본소설에서 ‘남성어’와 ‘여성어’가 많이 사용되는 것은 실제 사용 유무와는 상관없이 사람들에게 ‘남성어’와 ‘여성어’에 관한 지식이 공유되어 있기 때문으로 보인다.

구체적으로 일본 국내의 남성과 여성의 실제 문말표현 사용과 관련된 연구를 살펴보면 일본 직장에서의 여성 언어 사용 실태에 관한 조사가 있다. 尾崎(1997)는 20대~50대의 여성 19명을 대상으로 분석한 결과, 일본어의 대표적인 여성문말표현 가운데 하나인 「わ(wa)」의 사용 빈도가 80년대에 비해 현격히 줄었다고 밝혔다. 尾崎(1997)가 수집한 데이터베이스를 토대로 中島(1997)는 의문문 형식에서 사용되는 「わね(wane) (상승형)」와 「わよね(wayone) (상승형)」의 사용 빈도 또한 감소하였으며 대표적인 남성문말표현인 「かな(kana)」, 「かね(kane)」, 「だよね(dayone)」 등을 여성 또한 사용하는 경향이 있다고 분석하였다.

청년층이라는 특정 세대를 연구 대상으로 언어사용 실태 조사를 한 小川(1997)는 일본인 대학생(여성 65명, 남성 64명)을 대상으로 이들의 대화를 문자 데이터로 추출해 남녀별로 문말표현의 사용 현황을 조사하였다. 그 결과 현대의 젊은 여성들이 즐겨 쓰는 문말표현은 「のよ(noyo)」와 「の(no)+과거형」이었으며 대표적인 여성문말표현으로 꼽혔었던 「わ(wa)」와 「체언+よ(yo)」는 잘 사용하지 않았다고 분석했다. 한편 대표적인 남성문말표현인 「だよ(dayo)」, 「だよね(dayone)」를 여성이 사용하는 경우가 많았으며, 여성과 남성이 모두 쓰는 문말표현인 「よ(yo)」와 「の(no) (의문형)」의 사용빈도가 높아지는 등 문말표현 사용에 있어 남녀간의 차이가 점점 없어지고 있다고 주장하였다.

2000년대에 들어서면서 문말표현과 관련한 일본 국내 연구는 점차 남녀간의 차이가 없어지고 있는 일본의 사회적 분위기를 반영하기 시작하였다. 有泉(2013, p. 62)는 문말표현 사용에 있어 기존의 이분법적인 구분 방식에서 벗어나 새로운 문말표현 방식을 제시하였다. 일본의 조사대상자인 20대 학생 15명으로부터 얻은 약 40종류의 문말표현을 바탕으로 젠더 식별 경향(문말표현과 젠더의 관련성 지표)과 젠더 특이성(남성성·여성성의 강도 지표)을

4) a.는 여성이 쓰는 문말표현, b.는 남성이 쓰는 문말표현이다.

측정한 후 이를 사용빈도에 따라 7가지 카테고리로 나누어 문말표현 방식을 제시하였다. 남녀별 사용빈도를 강도에 따라 배타적(최고 강도) 형식, 고특이형식, 저특이형식으로 구분하고 여성과 남성 모두 사용하는 문말표현을 중성형식으로 분류하였다. 이를 정리하면 표 1과 같다.

표 1. 有泉(2013)의 문말표현 분류

분류카테고리	문말표현
배타적 남성형식	だぜ
고특이 남성형식	だろ、んだろ、んだぜ、ぞ、だい、かい、よな
저특이 남성형식	だ、んだ、んだよ、だな、だよ、(の)か、な、だろう
배타적 여성형식	わ、わよ
고특이 여성형식	なの、ね、よ(명사)、かしら、よね、のね、のよ、なのよ
저특이 여성형식	ね、でしようね、でしよ、んでしよ、でしよう、の、の?、んだもん
중성형식	かな、から、わけ、よ

이처럼 여성문말표현에 관한 최근의 일본 국내 연구는 여성과 남성의 명확한 구분이 점점 없어지는 이른바 ‘규범의 일탈’ 현상이 나타나고 있다고 정리할 수 있다(小林, 2009, p. 3).

한편 TV 드라마나 소설에 나타나는 남성과 여성의 문말표현 사용 실태는 尾崎(1997), 小川(1997)의 연구와는 다소 차이를 보인다. 젊은 세대일수록 점차 ‘남성어’와 ‘여성어’의 구분 없이 언어 사용이 ‘중성화’ 되어 간다고 주장한 水本(2005)는 젊은이들 사이에서 사라져 가고 있는 여성문말표현이 드라마 안에서는 여전히 많이 사용되고 있다고 지적하며 극본 작가를 대상으로 설문조사를 실시하였다. 그 결과 등장인물들 사이에서 빈번하게 사용되고 있는 여성문말표현이 극본 작가들의 스테레오 타입⁵⁾에 의존하는 바가 크다고 주장하였다. 특히 항의나 반론 등 감정을 표출하는 자기주장이 강하게 드러나는 장면에서 빈번하게 사용되고 있다고 언급하면서 이러한 경향이 TV 드라마나 소설 속에서 여성등장인물의 캐릭터를 극대화시킨다고 분석하였다.

그 밖에 일본어 문말표현에 관한 국내 연구로는 곽은심(2014, 2015)과 이유아(2020) 등이 있다. 곽은심(2014, 2015)은 두 논문에서 한국어 소설을 일본어로 번역할 때 나타나는 여성문말표현이 어떻게 번역되고 이때 어떤 기능을 담당하고 있는지에 대해서 비교 분석하였다. 나아가 여성문말표현의 사용을 중심으로 소설에서 여성 인물상이 번역 과정에서 어떻게 표현되었는지 살펴보았다. 이유아(2020, p. 155)는 『82년생 김지영』에서 제시된 젠더 의식과 관련하여 여성문말표현을 정리하였다. 텍스트에 나타난 주요 여성등장인물 5명의 대사가 문말

5) 어떤 특정한 대상이나 집단에 대하여 많은 사람이 공통으로 가지는 비교적 고정된 견해와 사고(두산백과 참조).

표현을 통해 어떻게 드러나고 있는지 살펴보았다. 본 연구는 이유아(2020)의 소설 『82년생 김지영』에 나타난 여성문말표현 연구를 차용하여 『82년생 김지영』의 한일 번역본의 문말표현 양상이 기계번역에서 어떻게 나타났는지 코퍼스 분석과 교차분석을 통해 살펴보았다.

3. 연구방법

3.1. 텍스트 선정

연구 결과를 일반화하고 나아가 타당성을 담보하기 위해 다음과 같은 두 가지 사항을 핵심적으로 고려하였다. 첫째, 텍스트 선정에 있어 폭넓은 독자층이 읽고 일반 대중에게 잘 알려진 대표적인 현대 한국문학 작품을 선정해야 한다. 조남주 작가의 『82년생 김지영』은 2016년 10월 출판된 이후 2021년 기준 누적 판매부수 130만 부를 돌파하였다. 한국소설로는 신경숙 작가의 『엄마를 부탁해』 이후 9년 만이다. 이러한 인기에 힘입어 2018년 12월 8일 일본에서 번역 출간되었고 2019년 7월 기준 누계 발행 부수 13만 부를 기록하는 아시아 문학으로서 이례적인 큰 인기를 끌고 있다.

둘째, 본고는 인간번역과 기계번역을 비교하는 접근 방식을 취하고 있기에 인간 번역본이 있어야 한다. 번역자는 제1회 일본번역대상을 수상한 바 있고 한국어 시집을 출간한 적이 있는 사이토 마리코(齋藤真理子)이다. 위와 같은 두 가지 조건에 따라 『82년생 김지영』의 한일 번역본 텍스트를 연구 대상으로 선정하였고 본 연구의 타당성을 인정받기에 적합한 작품이라 판단하였다.

본고에서는 한국소설로는 일본에서 유례없는 공감과 인기를 얻고 있는 『82년생 김지영』의 한일 번역본에 나타난 여성과 남성의 문말표현 양상을 먼저 살펴보고, 원작소설 『82년생 김지영』에 나타난 여성과 남성의 문말표현이 2종(구글, 파파고)의 기계번역에서 어떻게 드러났는지 고찰해 보고자 한다.

3.2. 연구 절차

본 연구는 소설 『82년생 김지영』한일 번역본에 등장하는 모든 등장 인물 대화문의 문말어미에 초점을 맞추어 여성문말표현과 남성문말표현을 젠더와의 관련성 및 여성성·남성성의 강도에 따라 7가지 카테고리로 구분한 有泉(2013)의 구분 방식을 토대로 살펴보고, 또한 원문을 기준으로 기계번역이 소설에 등장하는 대화문(ST)⁶⁾의 문말표현을 어떻게 처리했

6) ST는 'Source Text'의 준말로 '원문'을 가리킨다. 본 연구를 위해 살펴본 MT는 2021년 11월 29일, 2022년 1월 18일, 2022년 2월 4일에 도출한 것이며, 본고에서 인용하는 번역결과는 가장 마지막의 번역결과

는지를 구글 엔진과 파파고 엔진을 선정하여 분석하였다. 기계번역의 엔진 선택에 있어서는 다음 사항들을 고려하였다. 본 연구의 목적은 NMT의 활용 가능성을 타진하는 것이라는 점을 감안하여 일반 대중과 학계에서 가장 많이 언급하는 범용 엔진 두 종류를 사용하였다. 실험 데이터 구축 및 사용자 인터페이스가 용이해야 한다는 조건을 충족시키기 위해 구글 번역과 파파고 번역을 선정하여 연구를 진행하였고 비교 분석 시 파일명은 사이트 마리오의 번역본을 HT로 하고 구글 번역은 GMT로, 파파고 번역을 PMT로 표기하였다.

연구 대상이 된 텍스트는 소설속 등장인물의 모든 대화문이며, 소설 전체를 기계번역한 후 모든 등장인물의 대화문만을 추출하였다. HT의 총 단어수는 13,300 개로 이 가운데 여성등장인물의 총 단어수는 8,180 개, 남성등장인물의 총 단어수는 5120 개이며, GMT 9,736 단어, PMT 10,730 단어이다(표 2참조).

분석을 위해 인간번역과 기계번역 결과물을 텍스트 파일(*.txt)로 저장 후 3개의 코퍼스를 구축하였고 코퍼스 분석 도구인 안트콩크(Antconc) 3.2.4버전으로 키워드 리스트를 추출한 뒤 문말표현을 중심으로 분석하였다.

표 2. 인간번역과 기계번역의 단어수 비교

	번역자	파일명	총 단어수
인간번역	Saito Mariko	HT	13,300
기계번역	Papago(2022)	GMT	9,736
	Google(2022)	PMT	10,730

4. 사례 분석

4.1. 문말표현 사용 횟수

문말표현 가운데 『82년생 김지영』 번역본에 등장한 남성등장인물과 여성등장인물의 문말표현을 선행연구에서 살펴본 有泉(2013)의 7가지 분류 방식으로 코퍼스 분석을 한 결과는 표 3과 같다.

를 기준으로 한다.

표 3. 남성등장인물과 여성등장인물의 문말표현 사용횟수(HT)

분류카테고리		문말표현	남성등장인물 사용횟수	여성등장인물 사용횟수
남성 문말 표현	배타적 남성형식	だぜ	-	-
		だろ	3	-
	고특이 남성형식	んだろ	-	-
		んだぜ	-	-
		ぞ	3	-
		だい	1	-
		かい	8	2
		よな	-	2
	저특이 남성형식	だ	1	-
		んだ	5	6
		んだよ	-	-
		だな	2	-
		だよ	11	18
		(の)か	1	-
な		6	-	
だろう	1	-		
여성 문말 표현	배타적 여성형식	わ	-	2
		わよ	-	1
	고특이 여성형식	なの	1	2
		ね	9	36
		よ(명사)	-	-
		かしら	-	-
		よね	1	8
		のね	-	1
		のよ	-	-
	なのよ	-	1	
	저특이 여성형식	ね	-	-
		でしょうね	1	1
		でしょ	-	6
		んでしょ	-	3
でしょう		-	10	
の		-	7	
の?		3	23	
んだもん	-	1		
중성형식	かな	-	2	
	から	-	2	
	わけ	-	1	
	よ	16	87	

문말표현의 사용횟수를 살펴보면 인간번역 HT의 문말표현 가운데 남성등장인물과 여성등장인물이 가장 많이 사용한 것은 중성형식 문말표현인 「よ」였다. 여성과 남성의 사용 빈도를 각각 살펴보면 여성이 87회, 남성이 16회이다. 그중에서도 「よ」는 반론, 비난, 항의 등의 감정을 표현하는 수단으로 상대방의 언행을 시정하기 위해 자신의 생각을 강도있게 주장할 때 사용된다(水本, 2007, p. 15). 여성등장인물 중에서도 주인공인 김지영이 13회, 오미숙(어머니)은 20회로 다른 인물들보다 「よ」의 사용 빈도가 높게 조사되었다. 주인공 김지영은 유년기보다는 성인이 되어 취업과 임신, 육아 등을 거치면서 「よ」를 많이 사용하였다는 점을 알 수 있다(이유아, 2020, p. 175). 이는 김지영이 성장하면서 자아가 성숙하고 여성차별에 대한 경험이 늘어난 결과라고 추정해 볼 수 있다. 김지영의 어머니 오미숙의 경우는 평범한 가정주부이지만 생활력이 강한 어머니로서 묘사되고 있으며 이러한 그녀의 「よ」사용은 경제력과 함께 자신감을 느끼게 되면서 더욱 두드러졌다고 할 수 있다.

일반적으로 「よ」는 「だ+よ」의 형태로 남성이 많이 사용하지만 현대 사회를 살아가는 여성에게도 적지 않게 나타나고 있다(水本, 2005, p. 31). 본 연구의 HT 분석 결과를 보더라도 여성등장인물이 사용한 문말표현 가운데 「だ+よ」의 사용횟수는 18회로 나타나 남성의 사용 빈도가 높다고 여겨졌던 문말표현의 여성 사용 빈도 또한 관측되어 「よ」가 더이상 남녀를 명확하게 구분한다고는 볼 수 없다. 구체적으로 다음 예시를 살펴보면 이러한 경향은 더욱 명확하게 드러난다. 예시문 2)는 면접을 보러 가는 주인공 김지영을 태운 남자 택시기사의 대사이며 예시문 3)은 면접시험을 보러 간 김지영의 대사이다.

(2)

ST (P.100)	“나 원래 첫 손님으로 여자 안 태우는데, 딱 보니까 면접가는거 같아서 태워준 거야.”
HT (P.93)	「ふだんは最初の客に女は乗せないんだけどね、ぱっと見て面接だなと思ったから、乗せてやったんだよ」 (“평상시에는 첫 손님으로 여자는 안 태우는데, 얼핏 보니까 면접같아서 태워준 거야”)

(3)

ST (P.103)	“저런 개자식은 손모가지를 부러뜨려 봐야지.”
HT (P.96)	「そんな下司野郎は、手、折っちゃえばいいんだよ」 (“저런 개자식은 손을 부러뜨려 봐야돼”)

(2)와 (3)의 예시문을 통해서도 알 수 있듯이 중성형식으로 구분된 문말표현「よ」가 남녀 구분 없이 사용되고 있어 현대 일본 사회에서 나타나고 있는 남성문말표현과 여성문말표현의 변화를 소설에서도 엿볼 수 있다.

다음은 구글 엔진과 파파고 엔진으로 분석한 남성등장인물과 여성등장인물의 문말표현 사용횟수 결과이다. 분석 결과는 표 4, 표 5와 같다.

표 4. 남성등장인물과 여성등장인물의 구글 엔진 문말표현 사용횟수(GMT)

분류 카테고리	문말표현	남성등장인물 사용횟수	여성등장인물 사용횟수	
남성 문말 표현	배타적 남성형식	だぜ	-	
	고특이 남성형식	だろ	-	
		んだろ	-	
		んだぜ	-	
		ぞ	1	
		だい	-	
		かい	-	
		よな	-	
	저특이 남성형식	だ	5	19
		んだ	5	11
んだよ		2	-	
だな		1	-	
だよ		-	3	
(の)か		2	-	
な		-	-	
だろう	-	-		
여성 문말 표현	배타적 여성형식	わ	-	
	わよ	-	-	
	고특이 여성형식	なの	3	3
		ね	12	40
		よ(명사)	3	8
		かしら	-	-
		よね	1	-
		のね	-	-
		のよ	-	-
	なのよ	-	1	
	저특이 여성형식	ね	-	1
		でしょうね	-	-
		でしょ	-	6
んでしょ		-	3	
でしょう		1	10	
の		1	7	
の?		6	23	
んだもん	-	1		

분류 카테고리	문말표현	남성등장인물 사용횟수	여성등장인물 사용횟수
중성형식	かな	-	-
	から	1	5
	わけ	-	-
	よ	3	8

표 5. 남성등장인물과 여성등장인물의 파파고 엔진 문말표현 사용횟수(PMT)

분류 카테고리	문말표현	남성등장인물 사용횟수	여성등장인물 사용횟수	
남성 문말 표현	배타적 남성형식	だぜ	-	
	고특이 남성형식	だろ	-	1
		んだろ	-	-
		んだぜ	-	-
		ぞ	1	-
		だい	-	-
		かい	-	1
	저특이 남성형식	よな	-	-
		だ	-	6
		んだ	5	20
		んだよ	2	-
		だな	-	1
だよ		-	19	
(の)か		2	3	
な	-	1		
だろう	-	6		
여성 문말 표현	배타적 여성형식	わ	-	
	わよ	-	1	
	고특이 여성형식	なの	3	1
		ね	4	16
		よ(명사)	-	-
		かしら	-	-
		よね	-	1
		のね	-	-
		のよ	-	-
	なのよ	-	1	
	저특이 여성형식	ね	-	-
		でしょうね	-	-
		でしょ	-	-
		んでしょ	-	-
		でしょう	1	1
の		1	-	
の?		6	14	
んだもん		-	-	

분류 카테고리	문말표현	남성등장인물 사용횟수	여성등장인물 사용횟수
중성형식	かな	1	-
	から	-	-
	わけ	-	-
	よ	4	23

구글 엔진과 파파고 엔진으로 기계번역을 한 GMT와 PMT 모두 사용 빈도가 가장 높았던 문말표현은 HT의 결과와 상이했다. 구글 엔진의 경우 가장 많이 사용된 문말표현은「ね」였으며, 사용 빈도는 여성이 36회 남성이 9회이다.

「ね」는 대표적인 여성문말표현 가운데 하나이다. 「ね」는 ①동의, 확인, ②주의 환기와 다짐, ③가벼운 감동, ④의문 등의 용법으로 사용된다 7). 이중 여성에게 특징적인 것은 「写真ね Shsinne(사진이네)」, 「一番ね Ichibanne(최고네)」와 같이 '명사+ね'형으로 사용되는 경우 그리고 「作ってくれるしね Tukuttekerurusine(만들어 주기도 하고)」와 같이 접속사 말미에 붙어 사용되는 경우가 특징적이라 할 수 있다.

파파고 엔진의 경우 여성문말표현인 「の?」가 가장 많이 사용되었으며 사용횟수는 여성이 28회 남성이 6회이다. 「の?」는 의문문에서 주로 사용되는데 이를 사용함으로써 문장의 주체가 여성임을 드러낼 수 있다.

4.2. 교차분석(chi-squared test)

코퍼스 분석을 통해 얻은 데이터를 활용하여 남녀 두 집단 간에 인간번역과 기계번역(구글, 파파고)의 문말표현 사용 비율 차이를 파악하기 위해 교차분석을 시행하였다.

먼저 남성등장인물의 분석 결과를 살펴보면 다음과 같다.

표 6. 남성등장인물의 남성문말표현 사용 비율 카이제곱검증 결과

남성문말표현	んだ	だよ	기타	합계	χ^2	<i>p</i>
HT	5	11	26	42	8.594	0.072
GMT	5	2	9	16		
PMT	5	2	3	10		
합계	15	15	38	68		

인간번역과 기계번역에 출현하는 문말표현의 빈도수에 대한 χ^2 -검정결과 남성등장인물의 남성문말표현 사용에 대한 분포는 통계적으로 유의미하지 않았다((4,N=68)=8.594,

7) 네이버 백과사전 참조

$p=0.072$). 이는 남성등장인물이 HT와 GMT, PMT에서 남성문말표현을 취하는 방식이 동일하다고 할 수 없다는 것을 의미한다. 「んだ」, 「だよ」 그리고 기타 남성문말표현을 골고루 사용하고 있어 기계번역의 경우 남성등장인물의 남성문말표현을 제대로 번역하지 못한 채 임의로 문말표현을 사용하였다고 할 수 있다.

남성등장인물의 여성문말표현 사용에 대한 분포는 통계적으로 유의미하지 않았다 ((4,N=56)=3.809, $p=0.432$). 즉 HT와 GMT, PMT의 여성문말표현 빈도수는 동일한 모집단에서 임의로 추출되었다고 할 수 있다. 이는 남성등장인물이 여성문말표현을 취하는 방식이 HT와 GMT, PMT에서 동일하다고 할 수 없다는 것을 의미한다. 기계번역의 경우 「ね」, 「の?」 그리고 기타 여성문말표현을 골고루 사용하고 있어 남성등장인물들의 여성문말표현을 제대로 번역하지 못한 채 임의로 여성문말표현을 추출하였다고 할 수 있다.

표 7. 남성등장인물의 여성문말표현 사용 비율 카이제곱검증 결과

여성문말표현	ね	の?	기타	합계	χ^2	p
HT	9	3	3	15	3.809	0.432
GMT	12	6	8	26		
PMT	4	6	5	15		
합계	25	15	16	56		

남성등장인물의 중성형식 문말표현 사용에 대한 분포도 통계적으로 유의미하지 않았다. ((N=25) =3.940, $p=0.139$). 이는 남성등장인물이 중성형식 문말표현을 취하는 방식이 HT와 GMT, PMT에서 동일하다고 할 수 없다는 것을 의미한다. 기계번역 시 남성등장인물들의 중성형식 문말표현을 제대로 번역하지 못한 채 임의로 중성형식 문말표현을 번역해 사용하였다고 할 수 있다.

표 8. 남성등장인물의 중성형식 문말표현 사용 비율 카이제곱검증 결과

중성문말표현	よ	기타	합계	χ^2	p
HT	16	0	16	3.940	0.139
GMT	3	1	4		
PMT	4	1	5		
합계	23	2	25		

다음은 여성등장인물의 남성문말표현 사용 비율을 분석한 결과이다. 인간번역과 기계번역에 출연하는 문말표현의 빈도수에 대한 χ^2 -검정결과, 여성등장인물의 남성문말표현 사용

에 대한 분포는 통계적으로 유의미하였다((6,N=122)=45.850, p<0.001). 이는 특정 기계번역(구글, 파파고)에서 여성등장인물이 남성문말표현 「だ」, 「んだ」, 「だよ」를 더 많이 사용한다고 할 수 있으며 기계번역은 여성등장인물의 남성문말표현을 임의로 추출하지 않고 번역하였음을 시사한다.

표 9. 여성등장인물의 남성문말표현 사용 비율 카이제곱검증 결과

남성문말표현	だ	んだ	だよ	기타	합계	χ^2	p
HT	0	6	18	4	28	45.850	<0.001
GMT	19	11	3	3	36		
PMT	6	20	19	13	58		
합계	25	37	40	20	122		

남성문말표현으로 구분한 「だよ」를 여성등장인물이 사용한 예문을 살펴보면 다음과 같다. 예문은 김지영의 어머니인 오미숙의 대사로서 HT와 PMT가 “진짜야”의 ST번역을 남성문말표현인 「だよ」로 처리했음을 알 수 있다.

ST (p.36)	“진짜야. 초등학교 때는 오남매 중에서 엄마가 제일 공부 잘했다. 큰외삼촌보다 더 잘했어.”
HT (p.31)	「ほんとだよ。国民学校のときは五人きょうだいの中で私がいちばん勉強ができたんだから。上の伯父さんよりできたんだよ」 (“진짜야. 국민학교 시절엔 오남매 중에 내가 공부를 제일 잘 했어. 큰외삼촌보다 더 잘했다니까”)
GMT	「本当だ。国民学校の時はお兄さんの中でお母さんが一番勉強上手だった。大外叔父よりも上手だった」 (“진짜다. 국민학교 시절엔 오빠들 중에 내가 공부를 제일 잘 했어. 큰외삼촌보다 더 잘했다”)
PMT	「本当だよ。国民学校の時は5人兄弟の中でお母さんが一番勉強ができた。叔父さんよりもっとよくやった」 (“진짜야. 국민학교 시절엔 오남매 중에 엄마가 공부를 제일 잘 했다. 큰외삼촌보다 더 잘했다”)

이는 일반적으로 「だよ」는 남성들이 사용하는 경우가 많다고는 하나 여성에게도 적지 않게 나타나고 있기 때문에 「だよ」의 사용 유무가 남녀를 구분한다고 볼 수 없다는 이유아(2020, p. 158)의 연구 결과와도 일치한다. 인간번역에서는 오미숙의 “진짜야”의 문말어미를 「だよ」로 처리함으로써 오미숙의 젠더의식을 보여주려는 번역자의 의도가 반영된 것으로 보인다.

다음은 여성등장인물의 여성문말표현 사용에 대한 분포 결과이다. χ^2 -검정결과, 여성등장인물의 여성문말표현 사용에 대한 분포는 통계적으로 유의미하였다($4, N=220$)= 33.981 , $p<0.001$). 즉 특정 기계번역(구글, 파파고)에서 여성등장인물은 여성문말표현 「ね」, 「の?」를 더 많이 사용한다고 할 수 있다. 이는 기계번역이 여성등장인물의 여성문말표현을 잘 살려서 번역함으로써 소설 속 여성등장인물의 여성성을 잘 드러내었다고 할 수 있다.

표 10. 여성등장인물의 여성문말표현 사용 비율 카이제곱검증 결과

여성문말표현	ね	の?	기타	합계	χ^2	p
HT	36	23	43	102	33.981	<.001
GMT	40	14	12	66		
PMT	16	28	8	52		
합계	92	65	63	220		

마지막으로 여성등장인물의 중성형식 문말표현 사용에 대한 분포이다. χ^2 -검정결과, 여성등장인물의 중성형식 문말표현 사용에 대한 분포는 통계적으로 유의미하였다($2, N=128$)= 19.627 , $p<0.001$). 즉 특정 기계번역(구글, 파파고)에서 여성등장인물은 특정한 중성형식 문말표현 「よ」를 더 많이 사용한다고 할 수 있다. 이는 기계번역이 여성등장인물의 중성형식 문말표현을 잘 살려서 번역하였음을 알 수 있다.

표 11. 여성등장인물의 중성형식 문말표현 사용 비율 카이제곱검증 결과

중성문말표현	よ	기타	합계	χ^2	p
HT	87	5	92	19.627	<.001
GMT	8	5	13		
PMT	23	0	23		
합계	118	10	128		

이상 카이제곱검증 결과를 통해 기계번역에서 남성등장인물과 여성등장인물 간 문말표현 사용의 차이가 명확하게 드러났음을 알 수 있다. 남성 등장인물의 문말표현 사용의 경우 기계번역이 임의로 번역을 한 반면 여성등장인물의 문말표현은 구글 엔진과 파파고 엔진 모두 HT 수준의 번역물을 도출하였다.

5. 결론

본 연구는 두 종류(구글, 파파고)의 기계번역을 소설 『82년생 김지영』에 적용한 문학기계번역의 사례 연구이다. 본 연구의 목적은 단순히 문학 기계번역의 오역 및 결과물 중심의 평가에서 벗어나 한일 번역본에 나타난 여성등장인물과 남성등장인물의 문말표현 사용 양상을 살펴보고 이러한 문말표현이 기계번역에서 어떻게 나타났는지 고찰하는 데 있다. 이를 위해 소설 원작과 한일 번역본의 텍스트를 코퍼스로 구축하고 분석데이터로 교차분석(chi-square test)을 실행하였고 有泉(2013)의 7가지 문말표현 분류 방식에 따라 등장인물의 문말표현을 분석한 후 인간번역과 기계번역을 대조하였다.

언어 구조가 유사한 한국어와 일본어 언어 쌍에서 인간번역과 기계번역을 비교 분석한 결과 두 종류의 기계번역 모두 여성등장인물의 문말표현을 HT에 가까운 수준으로 번역하였다는 사실을 확인하였다.

이는 문학기계번역의 한계점을 쏟아 내는 국내 연구 속에서 기계번역이 특정 문학 작품의 여성문말표현 사용에 있어 등장인물들의 여성성을 잘 드러내었다는 연구 결과로 우리에게 시사하는 바가 크다.

본 연구의 목적은 단순히 문학기계번역의 오역 및 결과물 중심의 논의에서 탈피하여 소설 『82년생 김지영』의 한일 번역본에 나타난 여성등장인물과 남성등장인물의 문말표현 사용 양상을 살펴보고 이러한 문말표현이 기계번역에서 어떻게 나타났는지도 함께 고찰하는데 있다.

NMT의 등장으로 기계번역이 인간번역을 대체하게 되는 것은 아닌지, 번역가라는 직업이 결국 사라지게 되는 것은 아닌지, 앞으로 번역 교육이 과연 필요할 지에 대한 논의가 활발한 지금 인공지능 기술의 발전 속도를 고려했을 때 머지않아 다양한 종류의 문학작품까지도 기계번역이 가능한 시대가 올 수 있다.

다만 본 연구는 소규모 데이터와 두 종류의 범용 엔진을 통한 분석 결과만을 제시한 것이기에 추후 분석 대상을 확장하고 다른 장르와의 비교를 통한 분석도 이루어져야 할 것이다. 또한 기계번역의 결과 해석에 있어 연구자의 주관적인 해석이 반영되었다는 점도 본 연구의 한계라 할 수 있다. 그럼에도 불구하고 본 연구는 문학기계번역 관련 사례 연구가 부족한 국내의 연구 현실 속에서 장르를 구체화하고 한국어-일본어 언어 쌍의 문학기계번역 적용 사례를 분석하여 그 가능성을 모색하고자 했다는 점과 인간 번역과의 비교를 통해 기계번역의 어떠한 점을 보완해야 할지 생각해 볼 수 있었다는 점에서 의미가 있다고 생각한다. 본 연구를 바탕으로 향후 연구에서는 문학 장르를 조금 더 세분화하고 장르별로 차별화된 후속 연구가 이루어지기를 희망한다.

참고문헌

- 곽은심. (2014). 韓国小説の日本語翻訳に見られる女性登場人物の言葉遣いに関する一考察-女性ジェンダー-標示形式の使用を中心に-. *일본문학학보*, 64, 5-28.
- 곽은심. (2015). 韓日幡譯授業における文末表現の指導に關する研究 -性差を表す文末詞の使用を中心に-. *일본어교육연구*, 7-28.
- 김슬기. (2018). 인공지능으로 날개 단 자동통번역 서비스. Digieco. KT경제경영연구소. Retrieved from <http://digieco.co.kr/ktfront>
- 마승혜. (2017). 기계화되기 어려운 인간 능력과 문학번역 능력 비교 고찰 및 논의. *통번역학연구*, 21(3), 55-77.
- 마승혜. (2018). 문학작품 기계번역의 한계에 대한 상세 고찰. *통번역학연구*, 22(3), 65-88.
- 안미영. (2020). 한국어와 영어의 언어적 차이에 따른 기계번역의 문제점과 그에 대한 포스트 에디팅 방향 제안. *영어영문학*, 25(1), 103-130.
- 이상빈. (2016). 트랜스크리에이션, 기계번역, 번역교육의 미래. *통역과 번역*, 18(2), 129-152.
- 이용훈. (2016). *R을 활용한 코퍼스언어학과 통계학*. 한국문화사.
- 이유아. (2020). 소설 『82년생 김지영』의 번역에서 나타난 젠더 표현 양상에 대한 고찰. *일본어문학*, 88(1), 149-178.
- 이준호. (2019). 문학번역 적용을 위한 기계번역의 현주소. *통번역학연구*, 23(1): 143-167.
- 이창수. (2021). 기계학습 알고리즘을 활용한 문학번역에서의 기계 번역과 인간 번역 결과물 분류 연구. *번역학연구*, 22(1), 199-217.
- 이해미. (2022). 국내 언어학계와 일본어학계의 기계번역 관련 연구 동향 분석. *일본어학연구*, 71, 125-143.
- 탁성숙. (2002). *문말표현에 관한 연구*. 보고서, 13-21.
- 최윤영. (1996). 기계번역과 문학번역. *번역연구*, 4, 51-61.
- 황수영. (2015). 한국인 일본어 학습자에 있어서의 문말표현에 관한 양상 연구 : 종조사와 중도종료표현, 추량표현을 중심으로. *일본어교육*, 72-81
- 有泉優里. (2013). 会話文末における「男ことば」と「女ことば」の分類-ジェンダー識別傾向とジェンダー-特異性を指標として-. *日本語とジェンダー*, 13, 63-72.
- 奥田亜未. (2016). 채팅에서의 일본인 여성 젠더표현 사용경향에 관한 연구 -문말표현을 중심으로-. *일본어교육연구*, 139-156.
- 小川早百合. (2004). 話し言葉男女差-定義・意識・実態. *日本語とジェンダー*, 4, 25-39.
- 尾崎喜光. (1997). 「女性専用形式のいま」現代日本語研究会編 『女性のことば-職場編

—」 ひつじ書房.

中島悦子 (1999). *女性の言葉・職場編* ひつじ書房.

林明子. (2005). 젠더による日本語表現のヴァリエーション—若年層の用いる発話末の表現形式と音声的特徴—. *EAJIS (European Association for Japanese Studies)*, 43.

水本光美. (2005). テレビドラマにおける女性言葉とジェンダーフィルター. *日本語ジェンダー学会学会誌*, 5, 23-46.

水本光美. (2007). 主張どの強い場面における女性文末詞使用-実際の会話とドラマとの比較—. *国際論文第五号*, 北九州市立大学, 13-22.

水本光美. (2008). 드라마に使われる女性文末詞-脚本家の意識調査より. *日本語とジェンダー*-8, 73-94.

Hutchins, J. (2005). Current commercial machine translation systems and computer-based translation tools: system types and their uses. *International Journal of Translation*, 17(1-2), 5-38.

Toral, A., & Way, A. (2015) 'Machine-assisted Translation of Literary Text: A Case Study', *Translation Spaces*, 4, 241-268.

Toral, A., & Way, A. (2018) What level of quality can neural machine translation attain on literary text?. *arXiv preprint arXiv*, 1801, 04962.

Voigt, R., & Jurafsky, D. (2012). Towards a literary machine translation: The role of referential cohesion. In *Proceedings of the NAACL-HLT 2012 Workshop on Computational Linguistics for Literature*, 18-25.

조승연

03760 서울특별시 서대문구 이화여대길 52 이화여자대학교

통역번역대학원 한일통역학과 강사 및 동대학원 박사

이메일: hearty226@hanmail.net

Received on November 22, 2022

Revised version received on January 11, 2023.

Accepted on March 31, 2023