

문학텍스트의 한-아 기계 번역 오류 연구 - 『82년생 김지영』을 중심으로*

곽 순 례
(한국외대)

1. 서론

이 연구는 한국사회에 젠더(gender) 갈등이 심각한 상황에서 사회적으로 그리고 정치적으로 큰 관심을 끌고 있는 『82년생 김지영』의 아랍어 인간 번역문과 기계 번역문을 언어와 문화 어휘 범주에서 비교·분석하는 것을 목적으로 한다.

『82년생 김지영』은 ‘여성’이라는 젠더적 기준으로 선별된 에피소드로 구성되어 사회에서 뜨거운 논란을 일으켰던 작품이다. 이 소설은 1982년 한국에서 태어난 여성 김지영의 생애를 중심으로 실제로 여성에 대한 제도적 차별은 사라졌지만 가정, 학교, 직장 등 우리 사회에서 크고 작은 방식으로 잔존하는 보이지 않는 성차별적 관념들을 집중적으로 조명했다(연희선 2022.7.21.).

아랍 국가에서 젠더 측면에서 여성의 위치를 단편으로 보여주는 것은 이혼

* 이 논문은 2020년 대한민국 교육부와 한국연구재단의 인문사회분야 중견연구자지원사업의 지원을 받아 수행된 연구임(NRF-2020S1A5A2A01047483). 이 연구는 2022학년도 한국외국어대학교 교내학술연구비의 지원에 의하여 이루어진 것임.

권이다. 가부장 중심 가족제도에 기반을 둔 중동 이슬람 문화권에서 전통적으로 결혼과 이혼에 대한 결정은 남성에게 부여된 고유 권한으로 인식되어왔으며, 여성은 이 과정에서 공식적으로 배제되어왔다. 특히 일부다처가 허용되고 이혼권은 주로 남자에 의해 행사되어왔다(엄익란 2007: 51). 심지어 일부 아랍 국가에서는 간통을 하거나 강간을 당한 여성을 죽임으로써 가문의 수치를 씻어내는 ‘ 명예 살인’이 일어날 정도이다(엄익란 2015: 13). 아랍어판 번역가는 동아일보(임현석 2021.7.21.)와 가진 인터뷰에서 “한국 여성 김지영의 이야기는 이집트 여성의 이야기이기도 합니다. 여성에 대한 차별은 보이지 않는 곳에서 더 많이 일어나고 억압적인데 그 점이 많이 비슷합니다. 그래서 이집트와 아랍 여성들이 이 소설에 공감할 수 있다고 생각해 번역에 나섰습니다”라고 밝혔다.

조남주의 소설 『82년생 김지영』은 2016년 발간 이후 국내에서만 130만 부 넘게 판매되었다(공감언론 뉴시스 2022.7.11.). 한국문학번역원에 따르면 이 소설은 2016년부터 2020년까지 해외에서 30만 부 이상 판매되어 가장 많이 팔린 한국 문학작품이다. 이 기간 동안 한국문학번역원의 지원으로 해외에서 출간된 한국 문학은 2017년 127종에서 2021년에는 186종까지 늘어났고, 민간 재단인 대산문화재단 지원분까지 합하면 한해 200종 이상 해외에서 출간되었다. 지금까지 총 37개 언어권에서 658종의 한국 문학작품이 출간되었다(강종훈 2022.1.18., 정영현 2022.1.18.).

2021년 3월에 이집트에서 출간된 아랍어판은 사우디아라비아와 쿠웨이트 등에서도 출간되었으며 4개월 만에 초판이 다 팔려 2021년 7월 재인쇄에 들어갔다(임현석 2021.7.21.).

본 연구에서는 아랍어로 번역 출간된 소설 『82년생 김지영』의 인간번역과 구글번역기의 번역을 비교하여 오류 유형을 분류하여 범주별로 정량적으로 분석하고, 사례별로 정성적으로 분석한다. 이를 통해 문학텍스트의 한국어-아랍어 기계번역의 가능성과 한계를 가늠해 보고자 한다.

2. 선행연구

2.1 한국어-아랍어 언어쌍 기계번역

구글사가 2016년 11월 인공지능(AI)을 활용한 신경망 기계번역(Neural Machine Translation, NMT)을 적용한 번역 웹서비스를 제공하기 시작하면서 한국어-아랍어 언어쌍의 기계번역 품질이 획기적으로 개선되었다. 기계번역은 자연언어 처리에서도 가장 대표적인 응용분야인데, 현재 언어학, 번역학, 교육학 분야에서 기계번역 관련 연구가 활발히 이루어지고 있으나 한국어-아랍어 언어쌍에서는 연구가 제한적으로 이루어지고 있다.

본 연구는 국내에서 최초로 진행되는 문학텍스트의 한국어-아랍어 방향 기계번역 연구이다. 아랍어-한국어 방향 기계번역과 관련하여 진소영(2019)은 1988년 노벨 문학상을 수상한 나집 마흐푸즈(1911~2006) 작품 『أولاد حارتنا』의 한국어 번역 『우리 동네 아이들』의 아드함장에 나타난 ‘보다’ 동사를 중심으로 기계번역과 인간번역의 결과물을 비교 분석하였고, 기계번역 오류를 줄이기 위한 해결방안을 제안하였다. 아랍어의 구어체는 표준어와 많은 차이가 있으며 권역별로 다르게 나타난다. 이러한 어려운 아랍의 언어환경을 고려하여 바니아타(Baniata 2019)는 신경망 기계번역과 멀티태스크 방식으로 학습하는 아랍어 방언 기계번역 모델인 「A Multitask-Based Neural Machine Translation Model for Arabic Dialects(아랍어 방언을 위한 멀티 태스크 기반 신경망 기계 번역)」를 연구하였으며, 광순례(2021)는 비문학텍스트의 아랍어-한국어 기계번역 고찰을 통해, 문맥에 맞는 다의어의 어휘 선택 등과 같이 다른 여러 언어 쌍에서 발생할 수 있는 차이가 있고, 동철이의어, 동철이음이의어의 문맥에 맞는 의미 선택, 동사의 대명사 주어와 동사의 대명사 목적어의 명사화, 형용사와 부사 간 품사 전환, 부정성 주어의 처리 등과 같이 아랍어-한국어 언어 쌍에서만 발생할 수 있는 차이를 확인하였다. 이와 더불어 프리에디팅과 포스트에디팅을 통해 기계번역 품질을 개선시킬 수 있는 가능성을 실례를 통해 확인하였다. 이계연(2021)은 기계번역 후에 번역의 품질 개선을 위해 행해지는 포스트에디팅 기법을 학부 아랍어 번역교육에 적용하였고, 교육적 효과를 확인하였다.

2.2 소설 『82년생 김지영』의 번역과 젠더 문제

소설 『82년생 김지영』의 번역연구와 관련하여, 코다야 마유미(2019)는 소설 속에 나타난 문화 관련 어휘의 한일 번역 양상을 일본 문화에 없거나 익숙하지 않은 한국 문화 관련 어휘는 각주를 달아 설명하였고, 신조어와 생활문화 관련 어휘는 앞에 부연설명을 추가하여 이해도를 높였으며, 젠더 갈등을 일으켰던 작품인 점을 감안해 여성주의 관점을 반영한 어휘는 의역을 하거나 부연설명을 하였고, 의성어·의태어는 일본어에 대응되는 표현이 많아 그대로 적용하였다고 분석하였다. 또한 일본에서 20만 부 이상 팔린 원인이 한국문화 관련 어휘를 일본인이 쉽게 이해할 수 있도록 번역한 가독성과 충실성이라고 보고 있다.

이유아(2020)는 여성문말표현을 중심으로 소설 속에 나타난 젠더 표현 양상 연구에서 ‘여성어’와 ‘여성문말표현’을 정의하고 여성문말표현마다 사용된 빈도수를 조사하고 유년기, 취업기, 임신육아기를 지나면서 여성문말표현 사용이 증가하는 것을 자아성장과 여성차별에 대한 경험이 증가한 결과로 추정하고 있다.

최단(2022)은 소설의 원본과 중국어 번역판을 중국인 학습자들이 동시 접근하여 텍스트를 읽는 과정에서 차이점을 체감하면서 문화의 충돌을 대리 경험하게 되고 다중언어문화의 소양을 쌓아가는 것이 가능하다고 피력하고 있다.

소설 『82년생 김지영』에 나타난 젠더 갈등과 관련하여, 소설은 여성혐오 저항담론의 대중성을 획득해내고, 여성주의적 시각에서 정의한 ‘여성혐오(misogyny)’ 개념의 대중화를 시도하였고, 자기중심적 성격을 지닌 여성혐오가 여성의 삶에 미친 영향과 타자화된 여성의 문제점을 보여줌으로써 현실에서 문제 제기가 가능한 담론 환경을 제공하였으며, 신자유주의와 여성혐오의 유착관계를 밝혀 ‘여성과 남성은 이미 평등하다’는 전제에 균열을 가했다(김현, 손병우 2020). 이 외에도 페미니스트 번역전략 연구(최지영, 이정순 2022), 토픽 모델링을 활용한 페미니즘 텍스트 읽기에 나타난 독자 담론 분석(최숙기 2019), 소설을 원작으로 한 영화 리뷰에 나타난 페미니즘의 프레임에 대해 연구(차현주, 한연추, 권상희 2021) 등이 있다.

2.3 문화 어휘 번역

문화 어휘 번역은 인간번역에서조차 어렵다. 최근 K팝의 BTS, 영화 「기생

춤, 「미나리」, K드라마 「이상한 변호사 우영우」, 웹드라마 「오징어게임」 등의 인기에 힘입어 다양한 한국 문화콘텐츠가 세계 각국에 수출되고 있고, 전 세계의 많은 사람들이 K문화에 빠져 있다. 이제 한국은 문화계의 거물이 되었다(최수호 2021.11.4.). 2021년 9월 영국 옥스퍼드 영어사전(OED)은 한류(hallyu) 등 한국 문화 관련 단어 26개를 새로 실었다(이기욱 2021.10.13.). 대표적인 한국의 문화 어휘인 김치, 김밥, 불고기, 삼계탕과 같은 음식 이름부터 한복, 한글과 같은 어휘까지 아랍인들에게 많이 알려져 의미로 번역되지 않고 한글 발음을 그대로 표기한 음역 형태로 번역된다. 이러한 점을 고려하여 본 연구에서는 문화 어휘가 기계번역에서 어느 정도 반영되어 의미 번역이나 음역형태로 번역되는지 확인한다.

문학번역의 궁극적인 목적은 원작이 가지고 있는 미학적 효과와 감동을 다른 언어로 구체적으로 재현해냄으로써 원작을 읽은 독자가 받은 문학적 감동이 가장 유사한 형태로 다른 언어권 독자들에게 전달되도록 하는데 있다. 그러므로 원작을 읽은 독자와 번역된 작품을 읽은 독자가 가장 근접한 범주 내에서 작품에 대한 이해와 감동을 공유할 수 있을 때에 문학번역은 그 의도하는 목표를 성공적으로 완수했다고 평가 받을 수 있다(이형진 2011: 322). 이러한 목표를 실현하기 위한 문학번역에서 가장 큰 난제는 문화 어휘 번역이라 할 수 있다.

김효중(2004: 37)에 따르면 문화 어휘란 어느 특정의 상황에서 특정의 이해 그리고 이에 상응하는 행위를 유발하는 어느 언어권의 모든 문화적, 사회·경제적 소여성을 말하며, 언어적 요인뿐만 아니라 국가마다 다르게 정의된 비언어적 현상, 제도 등이 여기에 속한다.

이근희(2008: 248-249)는 번역자가 직면하는 문화 관련 어휘의 범주를 (1) 고유명사: 특정 인물, 건물, 조직, 단체, 저서 등의 이름, (2) 특정 문화와 관련된 어휘: 의식주, 지역, 사회, 관습 등과 관련된 어휘, (3) 특정 사건이나 인물과 관련된 어휘, (4) 관용어: 관습적으로 많이 쓰이는 속담, 유행어, 은어, 관용구, (5) 도량형의 단위: 화폐, 거리, 높이, 중량 등의 어휘 5가지 범주로 나누었다.

김재희(2018)는 한국문학작품에 나타난 문화소의 한-영, 한-아 번역방법을 분석하고, 호칭어는 대부분 영어와 아랍어번역에서 상위어전략을 사용하였으며 의미를 살리기 위해 추가적인 설명이 제시되었다고 밝힌다.

곽순례(2017)는 한국문학작품이 번역되는 언어권에서 대중적인 성공을 거두기 위해서는 자국화 번역 전략이 필요함에도 불구하고 대치할 목표어가 없는 경우 음역된 어휘가 상당량 나타났다고 밝히고 있다.

소설 『82년생 김지영』이 평범한 여성의 학교생활, 취업 후 사회생활, 결혼 후 가정생활 등 삶의 환경과 주변 인물들과의 관계 등을 다루고 있어, 이 소설에 생활 관련 어휘, 신조어와 줄임말, 호칭어, 한국적 표현, 음식명 등의 문화 어휘가 많이 나타났다.

3. 연구 내용과 방법

3.1 연구 내용

본 연구는 아랍어로 번역 출판된 한국어 문학작품의 인간번역과 기계번역을 실시한 결과물을 언어와 문화 범주에서 비교·분석한다. 본 연구는 미국 뉴욕 타임스 2020년 꼭 읽어야 할 책 100선에 선정된 조남주 작가의 소설 『82년생 김지영』을 내용으로 한다.

『82년생 김지영』의 아랍어 번역판은 한국 문학작품을 꾸준히 출판하고 있는 이집트 기자(الجميزة)에 위치한 출판사 دار مصفاة للنشر والتوزيع والدراسات(dār ṣafāf lilmnashr wa-altawzīf wa-aldirāsāt)에서 2021년 출간되었고, 번역은 아인샴스대학교 알알순대학 한국어과를 졸업하고 (주)이집트 한국문화원에서 문화 행사 기획 및 조직 담당자로 근무하고 있는 منار أحمد الديناري(Manar Ahmed Eldenary)에 의해 이루어졌다. منار(Manar)는 해외문화홍보원의 코리아넷뉴스(알라 압둘 아띠 2021.3.19.)와 가진 대담 기사에서 “가부장적 성차별 현실을 고발해 한국 남성들에게 강한 반감을 일으켰다는 기사를 읽었다. 이 기사를 읽고 호기심이 생겨 소설을 읽게 되었고, 소설의 여자 주인공과 아랍 여성들 간에 공통점이 많다는 것을 느끼고 번역할 결심을 하게 되었다.”라고 밝히고 있다.

이와 같이 한국과 아랍세계의 젠더와 관련된 사회적 공통점과 국내외의 출판시장에서 소설 『82년생 김지영』의 위상과 아랍어 초판의 완판을 고려하고, 한국어를 전공한 아랍인이 번역한 점을 확인할 수 있어서 소설 『82년생 김지영』

을 연구 내용으로 선정하였다. 연구 내용의 한국어 텍스트는 공백 포함 93,075 자이며, 아랍어 텍스트의 어절¹⁾ 수는 총 30,794개다.

3.2 연구 방법

인공지능을 활용한 국내에서 가장 대표적인 NMT 기반 기계번역기인 네이버의 파파고에서는 아랍어 번역 서비스를 제공하지 않아, 세계적으로 가장 대표적인 NMT 기반 기계번역기인 Google Translate를 사용하여 한국어-아랍어 기계번역을 진행하였다.

Google Translate를 이용하여 한국어 텍스트를 번역할 때 기계번역의 품질을 낮게 만드는 한국어 문장 성분의 모호성을 제거하지 않았으며, 문장 중 단어에 오타가 있거나 단어 순서가 어긋난 것과 같은 문장의 부적절한 요소인 노이즈를 제거하거나 내용 일부를 변경하는 가공 작업을 하지 않았다.

문장 단위로 의미에 지장을 주지 않는 범위에서 한 번에 한국어 약 500(공백 포함)글자 분량씩 2022년 4월 30일~5월 6일 동안 기계번역을 일괄 진행하였고, 번역 후 한국어 원문, 출판된 인간번역 결과물, 기계번역 결과물을 엑셀 파일에 병렬코퍼스로 만들었다.

곽순례(2022)는 한국어-아랍어 비문학텍스트 기계번역 연구에서 언어 범주의 오류를 어휘, 통사, 화용, 문화·정보의 4개 항목으로 분류하였다. 이 표를 기준으로 언어 범주를 수정 보완하여 분류하였고, 본 연구의 분석 대상이 문화어휘가 많이 포함된 문학텍스트인 점을 고려하여 문화·정보 범주를 문화 범주로 수정하고 항목을 세분화하였다.

문화어휘 범주는 이근희(2008: 248-249)의 문화어휘 범주 분류를 고려한 고유명사, 한국적 표현, 관용어 항목에, 『82년생 김지영』에 많이 나타나는 문화어휘의 특성을 고려하여 생활 관련 어휘, 호칭어, 음식명 항목을 분류 항목에 추가하였다. 이근희의 분류 중 ‘특정 사건이나 인물과 관련된 어휘’ 항목은 ‘IMF’ 1건만 나타났고, 도량형과 종교 항목의 문화어휘도 빈도수가 작아 별도 항목으로 분류하지 않았다. 신조어와 줄임말은 소설 『82년생 김지영』의 아랍어 번역가 *متر أحمد الديناري*가 “한국어와 영어를 합쳐 만든 신조어들이 번역하기 어려웠

1) 문장에서 띄어쓰기가 되어 있는 말의 덩어리들을 의미

어요 ‘된장녀’와 같이 대응되는 아랍어가 없는 어휘들의 의미를 아랍 독자들에게 전달하는 것이 가장 어려웠어요.”(알라 압둘 아띠 2021.3.19.)라고 밝히고 있어 별도 항목으로 추가하였다.

수정된 오류 범주는 다음과 같다.

오류범주	오류 내용
어휘	① 용어와 어휘의 의미 ② 대명사(인칭대명사, 지시대명사, 동사의 대명사 주어)의 문장 내 응결성과 응집성 강화를 위한 명시화 ③ 명사류(고유명사, 명사, 대명사, 수사, 대명사, 동사의 대명사 주어 등)의 성과 수 ④ 어휘 누락 ⑤ 부정을 나타내는 접두어, 부정사
통사	① 번역문에 맞게 품사 전환 ② 능동태와 수동태 ③ 부정성 주어 ④ 부사의 위치 ⑤ 문장의 완성
화용	① 문맥을 고려한 다의어의 적절한 의미 ② 문체(대화체, 경어체, 서간문체 등) ③ 문장 부호 ④ 띄어쓰기 ⑤ 통일된 어휘 ⑥ 자연스러운 표현 ⑦ 가독성, 비문 ⑧ 접속사 ⑨ 문장 종결의 유형(평서문, 의문문, 명령문, 청유문, 감탄문 등)
문화	① 고유명사 ② 한국적 표현 ③ 관용어 ④ 생활관련 어휘 ⑤ 신조어와 줄임말 ⑥ 호칭어 ⑦ 음식명

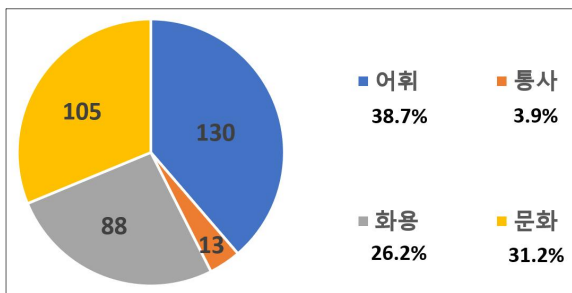
한국어 원문, 인간번역과 기계번역 결과물을 비교하여 오류를 분석한 후 위와 같은 범주로 분류하여 정량적으로 분석하였다.

그 후 항목별로 사례를 정성적으로 분석하고, 이러한 분석과정을 통해 오류의 원인을 파악하고, 기계번역의 한계와 가능성을 확인하였다.

이와 함께 K콘텐츠의 세계적인 인기에 힘입어 아랍세계에서도 한국의 문화 어휘가 많이 알려져서 기계번역에서도 반영되고 있는지 확인하였다.

4. 분석 사례 및 결과

본 장에서는 인간번역과 기계번역 오류의 정량적 분석과 정성적 분석을 사례를 제시하여 살펴본다. 백분율은 소수점 아래 2자리를 반올림하였다.



4.1 정량적 분석

기계번역을 인간번역과 비교하여 나타난 오류를 범주별로 분류한 결과는 다음과 같다.

범주	내용	계
어휘	①용어, 어휘 의미	66(19.6%)
	②대명사	16(4.8%)
	③성과 수	41(12.2%)
	④어휘 누락	3(0.9%)
	⑤부정어	4(1.2%)
	소계	130(38.7%)

통사	①번역문에 맞게 품사 전환	10(3.0%)
	②능동태와 수동태	1(0.3%)
	③무정성 주어	1(0.3%)
	④부사의 위치	0(0.0%)
	⑤문장의 완성	1(0.3%)
	소계	13(3.9%)
화용	①문맥을 고려한 다의어의 적절한 의미	12(3.6%)
	②문체	0(0.0%)
	③문장 부호	0(0.0%)
	④띄어쓰기	1(0.3%)
	⑤통일된 어휘	10(3.0%)
	⑥자연스러운 표현	27(8.0%)
	⑦가독성, 비문	35(10.4%)
	⑧접속사	3(0.9%)
	⑨문장 종결의 유형	0(0.0%)
	소계	88(26.2%)
소계		231(68.8%)
문화	①고유명사	1(0.3%)
	②한국적 표현	40(11.9%)
	③관용어	7(2.1%)
	④생활관련 어휘	36(10.6%)
	⑤신조어와 줄임말	10(3.0%)
	⑥호칭어	7(2.1%)
	⑦음식명	4(1.2%)
	소계	105(31.2%)
총계		336(100.0%)

분석된 총 오류는 336건이며, 그중 언어 범주 오류는 231건(68.8%), 문화 범주 오류는 105건(31.2%)이다. 언어 범주에서 어휘 관련 오류가 130건으로 38.7%를 차지하고, 통사와 관련된 오류가 13건으로 3.9%, 화용과 관련된 오류가 88건으로 26.2%를 차지하였다.

어휘 범주에서 어휘의 단순 의미 오류가 66건(19.6%), 성과 수가 41건(12.2%), 대명사가 16건(4.8%)이 나타났다. 아랍어는 한국어와 다르게 사람뿐만 아니라 사물도 성과 수가 구별되어 있고, 사람의 경우 1인칭, 2인칭, 3인칭으로 분류되어 있어 성과 수 범주에 오류가 많이 나타났으며, 이러한 어휘를 지칭하는 대명사에서도 오류가 많이 나타났다.

통사 범주가 가장 오류가 적게 나타났는데 그 중 품사와 관련된 오류가 10

건(3.0%) 나타났다. 통사 범주는 언어의 일반 규칙이 적용되는 분야라 어휘와 화용 범주에 비교해 볼 때 현저하게 오류가 적게 나타났다.

화용 범주에서 오류는 가독성, 비문이 35건(10.4%), 자연스러운 표현이 27건(8.0%), 다의어가 12건(3.6%)으로 빈도수가 많이 나타났다.

문화 범주에서 오류는 한국적 표현이 40건(11.9%), 생활 관련 어휘가 36건(10.6%), 신조어와 줄임말이 10건(3.0%)으로 나타났다.

4.2 어휘 범주 정성적 분석

4.2.1 어휘 의미 오류 사례

한국어	정대현 씨는 밤 12시 가 다 되어 퇴근하고,
인간번역	عادة ما يعمل جونج داي هيون حتى منتصف الليل
기계번역	يترك جونج داي هيون العمل في الساعة 12:00 ظهرا

‘밤 12시’가 기계번역에서 ‘낮 12시’로 번역되었다. ‘밤’이 다의어 중의 하나의 의미가 아닌 전혀 예상치 못한 상대 개념인 ‘낮’으로 번역되었다. 반면 인간번역에서 ‘밤 12시’를 ‘자정’으로 번역하였고 글자 그대로의 직역이 아닌 동일한 의미의 다른 어휘로 번역하여 자연스러움이 더해졌다.

한국어	김지영 씨가 정색하며 손을 탁 쳐 냈다.
인간번역	ولكنها أبعدت يده بوجه تعلوه نظرة صارمة
기계번역	لكن جي يونغ كيم صنفق بيدها بابتسامه مكلفة

‘정색하며’가 기계번역에서 ‘위선적인 미소 또는 씩소’로 번역되었고, 인간번역에서는 ‘엄격한 눈초리(표정)’로 번역되었다. 앞의 문장 ‘정대현 씨는 귀엽다는 듯 김지영 씨의 볼을 살짝 잡았는데,’를 고려해 보았을 때 기계번역에는 문맥에 맞는 어휘는 아니지만 부정적인 의미가 반영되어있다. ‘손을 탁 쳐 냈다’가 기계번역에서 ‘손뼉을 치다’로 번역되었다. ‘손을 탁 쳐 냈다’의 의미에 거부감이나 싫어하는 감정이 들어 있는데 이러한 의미가 표현되지 않았다.

한국어	계집질 안 하고, 마누라 때리지 않은 게 어디냐고,
인간번역	بل كانت تؤمن حقا بأنه زوج صالح كونه لا يضربها أو يهونها
기계번역	اعتقدت بصدق أنه ليس عاهرة ولم يضرب زوجته

‘계집질 안 하고’가 기계번역에서 ‘창녀가 아니다’로 번역되었고, 인간번역에서는 ‘배우자를 배반하지 않다’로 번역되었다. 기계번역에서 ‘계집질’ 어휘가 부인이 아닌 여성과의 성관계를 의미하는 표현으로 잘 번역되었으나 문장 번역에서는 ‘창녀가 아니다’라고 번역되어 다른 의미로 표현되었다.

한국어	아들이 못해도 넷은 있어야 되는 법이야.
인간번역	وحتى إذا لم يستطع أي منهم ذلك يجب أن يكون لديك أربعة أبناء
기계번역	حتى لو لم يستطع الابن ، يجب أن يكون هناك أربعة

이 문장은 기계번역에서 ‘아들이 할 수 없어도 넷은 있어야만 한다’로 의미 번역이 아닌 글자 그대로 직역되어 정확한 의미가 전달되지 않았다. ‘아들이’에서 주격조사가 나타나 ‘아들’이 주어로 번역되었고 ‘적어도’의 의미를 가진 ‘못해도’가 ‘할 수 없다’로 번역되었다. 원문 ‘넷은 있어야 되는 법이야’에 주어가 생략되어 있어, 문맥에 따라 선행 문장의 주어인 ‘아들’을 주어로 여겨 남성 3인칭 단수 대명사 주어 동사로 번역되었고 수사 ‘넷’ 뒤에 피계수 명사(아들의 복수 소유격 형태)가 누락되었다. 인간번역에서도 ‘적어도’의 의미를 가진 ‘못해도’가 ‘할 수 없다’로 번역되었다. 그러나 수사 ‘넷’ 뒤에 피계수 명사가 나타났고, 생략된 주어에 모음부호가 나타나지 않아 남성인지 여성인지 알 수 없으나 ‘2인칭 단수’로 명시화되었다.

한국어	혼자 이 집안 떠메고 있는 것처럼 앓는 소리 좀 하지 마.
인간번역	فلا داعي لأن تتحدث عن الأمر وكأنك تتحمل مسؤولية هذا المنزل وحدك
기계번역	لذلك لا تتحدث عن مرضك كما لو كنت تحمل هذا المنزل بمفردك

‘이 집안 떠메고 있는 것처럼’이 기계번역에서 ‘이 집을 감당하다’로 번역되어, 원문에 의미인 ‘집에 대한 책임을 떠맡다’의 ‘책임’ 의미가 누락되었고, ‘앓는 소리’가 ‘병, 질환’으로 글자 그대로 직역되었다. 인간번역에서 ‘앓는 소리’

가 선행 문장에 기술된 내용을 의미하는 ‘그 일’로 번역되어 의미가 축소되었다. ‘않는 소리’를 ‘الشكاوى’로 번역하면 어느 정도 의미가 표현될 것으로 보인다.

4.2.2 대명사 오류 사례

한국어에는 대명사에 성의 구별이 없지만 아랍어에서는 1인칭을 제외한 2인칭과 3인칭에서 남성과 여성으로 구별된다. 또한 1인칭은 단수와 복수로 구별되고, 2인칭과 3인칭은 단수, 쌍수, 복수로 구별된다. 그리고 모든 대명사는 주격, 소유격, 목적격으로 구별되고, 대명사는 단지 사람이나 사물을 지칭하는 것 이외에도 한정사 대신 명사 뒤에 연결형의 제2요소로 쓰여 한정의 의미를 나타내고 뜻을 명시화하기도 한다. 이와 같이 한국어 대명사와 아랍어 대명사 간에 여러 가지 다른 특성들이 있어 기계번역에서 이를 반영하기 어려워 대명사와 관련된 오류가 많이 발생하였다.

한국어	추석이 되어 시댁에 갔을 때 일이 터졌다.
인간번역	بحلول عيد التشوسوك الذي اعتادت فيه الأسرة زيارة أهل الزوج
기계번역	حدث شيء ما عندما ذهبت إلى منزل أهل زوجي في تشوسوك

‘추석’ 명절은 기계번역에서 음역되었다. ‘시댁’은 기계번역에서 ‘내 남편의 가족들’로 번역되었다. 이 부분은 1인칭 시점이 아닌 3인칭 시점으로 쓰여져 ‘그녀 남편의 가족들’로 번역되어야 한다. 즉 1인칭 단수가 아닌 3인칭 여성 단수 대명사로 번역되어야 한다. 한국어에는 자주 대명사가 생략되며, 원문에서 ‘시댁’이 지시하는 바는 ‘그녀의 시댁’이다. 한국어에서 아랍어로 번역할 때 생략된 대명사를 찾아 명시화하여 의미를 명확하게 해야 한다. 그런데 인간번역에서 ‘시댁’이 ‘남편의 가족들’로 번역되어 누구의 남편인지 알 수 없게 대명사가 한국어처럼 같이 생략되었다. 이는 한국어 원문의 영향 때문으로 보인다.

한국어	“언니는 분유 맛없어?”
인간번역	أختي، هل تكرهين طعم هذا الحليب
기계번역	أختك لا تحب مسحوق الحليب

‘언니’가 기계번역에서 ‘당신의 언니’로 번역되었다. 원문에는 대명사가 생략되어 있으나 대화문이므로 ‘언니’는 ‘나의 언니’로 번역되어야 한다. 즉 2인칭 여성 단수가 아닌 1인칭 단수 대명사로 번역되어야 한다. 또한 대화문에서 ‘언니’가 2인칭 여성 단수이기 때문에 ‘맛없어’의 동사 주어도 동일하게 2인칭 여성 단수로 번역되어야 하는데 기계번역에서 대화문이 아닌 서술문에 쓰여진 ‘언니’처럼 3인칭 여성 단수로 번역되었다. 이와 같이 원문에 생략된 대명사가 명시화되어 번역되어야 하고, 문맥에 맞게 성, 수, 격에 맞는 대명사형으로 번역되어야 한다.

4.2.3 부정어 오류 사례

한국어	김지영 씨가 더 이상 분유 먹을 나이가 아니라 거나
인간번역	كبر سنهما على تناول هذا الحليب
기계번역	كيم جي يونغ لم تعد كبيرة بما يكفي لتناول الحليب الصناعي

부정형 술어가 들어간 이 문장은 기계번역에서 ‘김지영씨가 분유를 먹을 정도로 더 이상 나이가 많지 않다’로 반대의 의미로 번역되었다. ‘더 이상 나이가 많지 않다’가 아닌 ‘더 이상 어리지 않다(لم تعد صغيرة)’로 번역되어야 한다. 인간번역에서 나이는 ‘이 우유를 마시기에는 너무 커버렸다’로 부정형 술어를 사용하지 않고 의미 번역이 되었다. 단, 김지영씨 한 명을 지칭하는데 3인칭 여성 단수 대명사를 사용하지 않고 3인칭 쌍수로 대명사를 사용하여 오역되었다. 번역사가 문맥을 잘 이해하지 못해 김지영씨와 김지영씨의 언니 두 사람을 지칭하는 것으로 이해한 것으로 보인다.

한국어	일곱 살 막둥이는 절대, 절대, 엄마와 잘 거라고
인간번역	أصر الطفل ذو السنوات السبع أنه لن يتام أبداً دون أمه
기계번역	أصر الصغرى البالغة من العمر سبع سنوات على أنها لن تتام أبداً مع والدتها

‘절대, 절대, 엄마와 잘 거라고’는 기계번역에서 ‘엄마와 절대 자지 않을 거라고’로 반대의 의미로 번역되었다. ‘절대’는 ‘어떠한 경우에도 반드시’의 의미로 긍정어뿐만 아니라 부정어와도 같이 쓰이지만 부정어와 호응하여 쓰이는 경우가 많다. 기계번역은 학습데이터셋트를 학습하면서 성능이 높아지는데(지인

영, 김희동 2020: 204) 학습데이터세트에 ‘절대’가 부정어와 많이 쓰여서 부정의 의미로 번역된 것으로 보인다.

한국어	애 보고..... 손목을 안 쓸 수가 없어요
인간번역	على أن أعتني بالطفلة... لن أستطيع أن أريح رسغى
기계번역	لا يمكننى استخدام معصمى لمشاهدة الأطفال

‘손목을 안 쓸 수가 없어요’는 이중 부정형 문장이다. 기계번역에서 ‘손목을 쓸 수가 없어요’로 반대의 의미로 번역되었다. 인간번역에서는 ‘안 쓰다’가 ‘편하게 하다’로 번역되어 ‘손목을 편하게 둘 수가 없다’의 의미로 이중 부정형 문장이 부정형 문장으로 의미 번역되었다.

한국어	“절반 좋아하네. 못해도 7대 3이거든? 내가 7, 당신이 3.”
인간번역	النصف! عزيزي النسبة سبعين إلى ثلاثين في المئة على الأقل، أنا سبعون وأنت ثلاثون
기계번역	أنا أحب نصفها. على الأقل 7 إلى 3؟ أنا 7 ، أنت 3.

‘절반 좋아하네’는 긍정형 문장이지만 반어법 표현이다. 기계번역에서 ‘절반을 좋아한다’로 글자 그대로 직역되어 반대의 의미가 되었다. 반어법이야말로 인간이 아닌 기계가 그 의미를 파악해서 번역하기 어렵다고 할 수 있다.

4.2.4 다의어 오류 사례

한국어	“몇 번 타세요?”
인간번역	أي حافلة تركيبى
기계번역	كم مرة تركيبى

‘번’의 다양한 의미 중 자주 쓰이는 의미는 ‘일의 차례’, ‘일의 횟수’, ‘사물의 차례’, 번호를 의미하는 ‘숫자’이다. 문맥에 따라 ‘번’은 ‘숫자’를 의미하는데 기계번역에서 ‘횟수’에 해당하는 의미로 번역되었다. 동사 ‘타다’는 문맥에 따라 남학생이 주인공 김지영에게 묻는 질문이므로 여성 2인칭 단수 미완료 동사로 번역되어야 하는데 기계번역에서 남성 2인칭 단수 미완료 동사로 번역되었

다. 인간번역에서 번호 대신 의미를 살려 ‘어느 버스를 타세요’로 번역되었다. 원문 단어의 다양한 의미 중 문맥에 적절한 의미를 선택하는 과정이 기계번역에서 어려우며, 이러한 점이 기계번역의 한계를 보여준다.

한국어	너 항상 내 앞자리에 앉잖아. 프린트도 존나 웃으면서 주잖아
인간번역	تجلسين أمامي دائما بالفصل، ودائما ما تمررين لي أوراق الدرس بإتسامة عريضة
기계번역	أنت تجلس دائما أمامي. تعطي البصمات بإتسامة لعنة

원문에서 ‘프린트’는 학원 수업 자료를 의미한다. 기계번역에서 ‘프린트’가 ‘발자국’으로 번역되었다. 이는 ‘프린트’가 영어를 음역한 어휘라서 영어 명사 ‘print’의 다양한 의미 중 ‘발자국’으로 번역된 것으로 보인다. ‘존나 웃으면서’가 기계번역에서 ‘빌어먹게 웃음지으며’로 번역되었고, 비속어 ‘존나’가 부정적 의미로 번역된 것으로 보이며 문맥에서 ‘존나’는 ‘활짝’의 의미로 번역되었다.

한국어	두 명의 간호사와 한 명의 의사가 번갈아 내진 을 했다.
인간번역	طبيب وممرضتين يتناوبون عليها لإجراء فحوصات عنق الرحم
기계번역	تناوبت ممرضتان وطبيب على هن الزوال

산부인과에서 진행되는 검사인 ‘내진’을 기계번역에서 지진과 관련된 어휘로 파악했으나 ‘지진을 견디어 났’의 의미가 아닌 ‘지진의 흔들림’으로 번역되었다. ‘내진’의 여러 의미에서 문맥에 맞는 적절한 어휘 선택이 이루어지지 않았다. 문장 내에 주어가 ‘두 명의 간호사와 한 명의 의사’인데 지진과 관련 있는 어휘로 번역되었다는 점은 기계번역 과정에서 연어가 고려되지 않았으며, 기계번역의 한계를 보여준다.

한국어	우리 피부톤 비슷해서 잘 받는 색도 비슷하잖아.
인간번역	لدينا لون بشرة متقارب لذا سيبدو جيدا عليك أيضا
기계번역	وبالتالي فإن الألوان التي تحصل عليها هي نفسها

‘잘 어울리는’의 의미인 ‘잘 받는’을 기계번역에서 ‘그녀가 얻은(받은)’으로

번역되었다. 동사 ‘받다’의 여러 의미 중 문맥상 ‘색깔이나 모양이 어떤 것에 어울리다’의 의미를 나타내는데 ‘다른 사람이 주거나 보내오는 물건 따위를 가지다’의 의미로 번역되었다. ‘잘 받는다’가 수식하는 어휘 ‘색’이 고려되지 않았으며 기계번역의 한계를 보여준다.

4.2.5 자연스러운 표현 오류 사례

한국어	손바닥만 한 그늘도 없는 운동장 한가운데였다.
인간번역	الملعب الرياضي بالجامعة حتى لم تترك للظل طرف أمانة
기계번역	وكانت في منتصف الملعب بلا ظل بحجم كف اليد

‘손바닥만 한 그늘도 없는’이 기계번역에서 글자 그대로 ‘손바닥 크기의 그늘도 없는’으로 직역되었다. 이와 같이 번역을 하면 무슨 뜻을 전달하기 위한 것인지 짐작해서 이해는 하겠지만 아랍에서 쓰는 표현이 아니라 자연스럽지 않은 번역투의 표현이다.

한국어	애, 너 힘들었니?
인간번역	أكان ذلك مرهقا لك؟
기계번역	مرحبا ، هل تواجه صعوبة في ذلك؟

이 문장은 명절 음식을 준비하느라 힘들었던 김지영에게 시어머니가 한 말이다. ‘힘들다’에 ‘어렵거나 곤란하다’의 의미도 있으나 이 문장에서 ‘힘들다’는 ‘가사노동으로 육체적으로 힘이 들다’의 의미로 번역해야 한다. 기계번역에서 ‘그 일에서 어려움에 직면했니?’의 의미로 번역하여 ‘힘들다’의 의미가 부자연스럽게 번역되었다. 또한 기계번역에서 ‘애’가 ‘안녕’으로 번역되었고, 시제도 현재시제로 번역되었으며, 동사의 주어도 2인칭 남성 단수로 번역되었다. 동사의 주어는 김지영을 의미하므로 2인칭 여성 단수가 되어야 한다.

한국어	그리고 등산 동아리에서 첫 번째 남자 친구도 사귀었다.
인간번역	والتقت بحبيبها الأول بنادي الحركة بالجامعة
기계번역	كما صنعت أول صديق لي في ناد للتسلق

‘첫 번째 남자 친구도 사귀었다’가 기계번역에서 ‘첫 번째 남자 사람 친구 (남사친)를 만들었다’로 번역되었다. 명사 ‘صديق’와 동사 ‘صنعت’는 연어로 같이 쓰이지 않는 어휘로 부자연스러운 번역투의 문장으로 나타났다. ‘등산’은 기계번역에서 ‘التسلق’로 번역되었다. ‘التسلق’는 암벽등반을 지칭하며 신체활동이나 운동을 위해 산을 오르는 등산은 ‘الحركة’이므로 오역되었다.

한국어	은영 아빠는 어쩔 거야?”
인간번역	ماذا ستفعل يا أبا أون يونج
기계번역	فماذا سيفعل والد أون يونج

이 대화는 남편과 부인 간에 이루어지고 있다. 그런데 기계번역에서 호격으로 부르는 2인칭 남성 단수 대화 상대방이 3인칭 남성 단수로 번역되어 대화의 상대가 아닌 제3자의 이야기를 대화하는 내용으로 바뀌는 오류가 발생되었다. 한국 사회에서 남편이나 부인을 부를 때 자녀 이름을 사용하여 ‘~의 아빠, ~의 엄마’로 부른다. 아랍사회에서도 연세 드신 분을 이름 대신 ‘~아빠, ~엄마’로 부르며 이를 ‘الكنية’라고 한다. 따라서 위의 문장을 달리 표현한다면 ‘은영 아빠, 당신은 어쩔 거야?’와 같은 의미이다. 인간번역에서는 은영 아빠가 호격사와 함께 번역되었다. 이는 문맥을 고려하지 못하는 기계번역의 한계를 보여준다. 반면 한국어의 ‘어쩔 거야(어떻게 할거야)’는 아랍어로 번역될 때 의문사 ‘كيف(어떻게)’가 아닌 의문사 ‘ماذا(무엇)’가 사용되는데 기계번역과 인간번역 모두 정확하게 번역되었다. 이는 기계번역의 가용성을 보여준다.

한국어	짜꿍은 이제 엄마한테 죽었다며 어깨를 축 늘어뜨린 채 집에 갔고, 선생님은 김지영 씨를 교실에 남겨 했다.
인간번역	رجع الولد الى بيته مطأطأ الرأس مقتنعا أن أمه ستقتله لما فعل، بينما أبقّت المعلمة جي يونج بالفصل بعد انتهاء اليوم الدراسي
기계번역	عاد شريكى إلى المنزل وكتفيه متدلّيتين قائلاً إن والدته ماتت ، وغادر المعلم كيم جي يونج في الفصل

‘짜꿍은 이제 엄마한테 죽었다며’에서 동사 죽었다는 미래시제를 의미한다. 기계번역에서 ‘والدته ماتت’(엄마가 그를 죽였다)로 과거시제로 번역되었다. 인간번

역에서는 ‘أمه ستقله’(엄마가 그를 죽일 것이다)로 미래시제로 번역되었다.

위의 오류 사례들을 종합해보면, 기계번역에서 각 어휘의 의미는 어휘만 별도로 번역했을 때 거의 정확하게 번역되는데, 앞뒤 문장 구성 요소들과 결합되는 경우 문맥과 다른 의미로 번역되는 경우가 많이 나타났다. 특히 부정어가 섞여 있으면 문장의 의미를 파악하지 못해 반대의 의미로 번역된 경우가 많이 나타났다. 문맥에 따라 적절한 의미를 선택해야 하는 다의어 번역은 아직 부족한 점이 많았고, 문맥에 전혀 맞지 않는 의미로 번역된 경우가 많았다. 대명사는 예측한 바처럼 기계번역에서 많은 오류가 발생하였다. 특히 대화문에서 1인칭이나 2인칭을 호칭어로 지칭한 경우에 주어가 정확히 파악되지 않았다.

4.3 문화 범주 정성적 분석

문화 관련 어휘는 인간번역에서도 쉽지 않은 범주이다. 이러한 점을 고려해 볼 때 기계번역 결과를 인간번역과 비교하는 것에 큰 의미를 두기 어렵다. 따라서 본 연구에서 문화 분야 오류 중 전통적인 문화 분야를 제외한 부분을 분석하였다.

4.3.1 줄임말 오류 사례

한국어	어린 여공들은 직장 생활이 원래 그런 건 줄 알고
인간번역	ظلت الفتيات أن هذه هي ظروف الحياة العملية
기계번역	عرفت الأميرات الشابات أن الحياة العملية كانت هكذا
파파고 번역	تعتقد الفتيات الصغيرات أن هذه هي الطريقة التي يتم بها العمل

‘어린 여공들’은 기계번역에서 ‘젊은 공주들’로 번역되었다. ‘공’을 ‘공주’로 인식해서 번역된 듯하다. 원인을 분석하기 위해 구글 번역기에서 원 문장을 영어로 번역한 결과 ‘Young female princesses know that work life is like that’으로 나타났으며, 이것이 아랍어 번역에 영향을 미친 것으로 판단된다. 이 문장을 네이버 파파고에서 번역하면 ‘Young girls think that’s how work is done’으로 번역된다. 이 문장을 구글 번역기에서 번역하면 ‘어린 여성들’로 번역되었다.

한국어	연말에 치러진 공채 에 합격했다
인간번역	وتعيّنت بالشركة من خلال فتح باب التوظيف للعامة بنهاية السنة
기계번역	واجتاز التوظيف المفتوح الذي عقد في نهاية العام

‘공채’는 ‘공개 채용’의 줄임말이다. 예측과 달리 기계번역에서 자연스러운 표현은 아니지만 ‘التوظيف’로 의미가 정확하게 전달되었다. 아랍어로 ‘공개 채용하다’는 ‘يفتح باب التوظيف’와 같이 숙어처럼 쓰이는 틀에 박힌 표현이 있다. 인간번역에서는 정확하게 이와 같은 표현으로 번역되었다. 다만 ‘공채’가 줄임말이라 풀어서 써서 원문보다 길게 번역되었다.

한국어	우리가 칼퇴 하는 회사도 아닌데 뭐. 그냥 30분 날로 먹는 거지
인간번역	كما لو أن أحدا في هذه الشركة يغادر المكتب في موعده! هي فقط تاهو لـ 30 دقيقة مجانية
기계번역	إنها ليست حتى الشركة التي نتركها. فقط أكله نينا لمدة 30 دقيقة

‘칼퇴’는 ‘칼 퇴근’의 줄임말로 ‘퇴근 시간이 되자마자 조금도 지체 없이 바로 퇴근함’을 뜻한다. 기계번역에서 ‘퇴근하다’의 의미로 전달된 것으로 보아 ‘퇴근’의 의미는 학습된 것으로 보인다. ‘날로 먹는 거지’의 의미는 ‘힘을 들이지 아니하고 일을 해내거나 어떤 것을 차지하는 것을 낫잡아 이르는 말’인데 기계번역에서 ‘익히지 않은 상태로 먹기만 하였다’로 직역되었다.

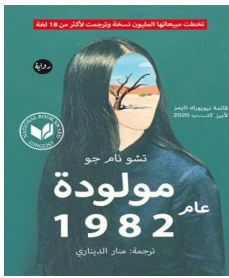
한국어	외근을 나가거나, 반차 를 내고 병원에 갈 때는
인간번역	كانت عندما تخرج لمهمة خارج الشركة أو تأخذ نصف يوم إجازة للذهاب إلى الطبيب
기계번역	عندما كنت أذهب إلى العمل أو دفعت نصف سيارة للذهاب إلى المستشفى

‘반차’는 ‘오전이나 오후 동안 주어지는 휴가’의 의미로 기계번역에서 ‘반(1/2)대의 자동차’로 번역되었다. ‘반차’의 의미가 학습되어 있지 않아 ‘반차를 내다’에서 동사 ‘내다’도 기계번역에서 ‘지불하다, 내다’로 번역되었다.

한국어	수업료에 시터 비용까지 더하니 금액이 너무 커져 부담스럽기도 했다.
인간번역	كما أن الرسوم الدراسية مع راتب جلسة الأطفال كان عبئًا كبيرًا
기계번역	عندما أضفت رسوم المعهد إلى الرسوم الدراسية ، كان المبلغ كبيرًا جدًا وكان مرهقًا

‘시터’는 ‘베이비시터’를 줄여서 표현한 것으로 기계번역에서 ‘좌석’으로 번역되었다. 인간번역에서는 ‘베이비시터’로 줄이지 않고 번역되었다.

4.3.2 호칭 오류 사례



『82년생 김지영』의 아랍어 제목은 왼쪽 표지 이미지에서 볼 수 있듯이 ‘김지영’ 이름이 빠진 『82년생 여성』으로 번역되었다. 이름이 자신의 정체성을 드러내는데 제목에서 이름이 빠진 채 번역되었다는 것이 젠더 측면에서 시사하는 바가 크다고 본다.

한국어	당신 뭐야. 꼭 장모님 같아
인간번역	ما خطبكِ تحدثين كوالدك
기계번역	من أنتِ تبدو تمامًا مثل حماتك

이는 남편이 아내에게 하는 대화이므로 ‘장모님’은 ‘아내의 어머니’를 의미한다. 아랍어에서는 어휘 ‘حم’가 문맥에 따라 시아버지와 장인을 의미하고, 또 어휘 ‘حمات’가 시어머니와 장모를 의미한다. 기계번역에서 ‘장모님’이 ‘당신(아내)의 장모’ 또는 ‘당신(아내)의 시어머니’로 번역되었다. 장인과 장모 호칭은 사위(남성)와의 가족관계 호칭이므로 아내의 장모가 있을 수 없으며, 아내의 시어머니로 번역되면 의미가 달라진다. 한국의 가족관계 호칭어가 기계번역에 학습되었다면 일어날 수 없는 오류가 발생했다.

한국어	처가에는 엉덩이 한 번 붙였다 그냥 가고
인간번역	وتذكرني بزيارة سريعة لا تكاد تجلس فيها
기계번역	لكني وضعت مؤخرتي على أهل زوجي مرة واحدة وانهب

‘처가’는 ‘아내의 본가’를 의미한다. 기계번역에서 ‘처가’가 ‘시가’로 번역되었다. 한국의 가족관계 호칭이 학습되었다면 ‘처가’라는 어휘는 남편이 아내의 본가를 지칭할 때 사용하므로 ‘나(남성)의 남편의 집’이라고 번역될 수 없다.

4.3.3 음식 이름 오류 사례

한국어	저녁부터는 사골을 우리고, 갈비를 재고
인간번역	وأَمْضُوا مَسَاعِمَ فِي سَلْقِ عِظَامِ الدَّيْجَةِ وَتَبِيلِ ضُلُوعِهَا
기계번역	في المساء ، نطهو عظام اللحم البقري ، ونخزن الضلوع الاحتياطية

기계번역에서 ‘사골을 우리다’가 ‘우리가 소의 뼈를 요리하다’로 번역되었고, ‘갈비를 재다’는 ‘비축용 갈비를 저장하다’로 번역되었다. ‘갈비’는 아랍인들에게도 많이 알려진 음식 이름인데 아직까지 기계번역에서 학습되지 않은 것으로 보인다. 인간번역에서 ‘사골을 우리다’가 ‘동물 사체의 뼈를 삶다’로 소 대신 상위어로 번역되었고, ‘갈비를 재다’가 ‘갈비뼈를 양념하다’로 번역되었다.

한국어	갈비를 찌고, 송편을 빚고
인간번역	سلق اللحم، وتحضير كعك الأرز
기계번역	أعدت أضلاع على البخار وصنعت Songpyeon

‘갈비’는 바로 앞의 사례 다음 단락에 바로 오는데 ‘갈비를 찌고’가 기계번역에서 ‘갈비를 스팀으로 요리했다’로 번역되었고, ‘송편을 빚다’는 ‘Songpyeon을 만들었다’로 번역되었다. 송편이 음식인지 아닌지도 학습이 안 된 듯하다. 인간번역에서 ‘갈비를 찌고’는 ‘고기를 삶다’로 번역되었고, ‘송편을 빚다’는 ‘떡을 만들다’로 상위어로 번역되었다. 여기에서 ‘찌다’라는 동사가 음식과 관련 있는 어휘로 기계번역에서 정확히 학습된 것으로 보인다.

한국어	김치도 없이 고구마를 먹던 겨울밤
인간번역	فيتناولون البطاطا الحلوة المطهية على البخار من دون طبق من الكيمشي
기계번역	وتناولوا البطاطا الحلوة بدون الكيمشي

‘김치’는 기계번역이나 인간번역에서 모두 ‘김치’로 음역되었고 표기도 김치의 종류가 많아 정관사를 붙여 한정 상태로 표현되었다. 이는 ‘김치’는 이미 한국의 음식 이름으로 아랍사람들에게 잘 알려져 있다는 의미이다.

위의 사례 분석을 볼 때, 줄임말 ‘여공’은 네이버의 파파고를 사용해 영어로 번역하고, 구글 번역기를 사용하여 영어-아랍어 번역을 진행하면 ‘어린 소녀들’로 번역된다. ‘공장에서 일하는’은 누락되었지만 파파고가 구글 번역기의 번역보다 의미가 더 명확하게 번역되었다. ‘공채’, ‘칼퇴’의 의미도 어느 정도 의미가 반영되어 번역되었으며 ‘반차’도 휴가라는 의미로 번역되었다.

가족 호칭어는 가족관계가 파악되지 않아 사위와 관련된 어휘가 아내와 관련되게 번역되었고, 아내와 관련된 어휘가 남편과 관련되게 번역되었다. 음식 이름 중에서는 김치만 음역되었고 사골과 갈비는 음역은 아니지만 음식의 종류로 번역되었고, 송편은 음역도 아닌 영어로 번역되었다.

5. 논의 및 결론

앞서 살펴본 오류 사례들의 분석 결과를 기계번역 관점에서 보면 언어 범주에서 각 어휘의 의미는 대부분 정확하게 번역되었지만, 문장 구성 요소들과 결합되면 다른 의미로 번역되는 경우가 많았다. 특히 문장내 부정어가 포함되거나 이중부정, 반어법인 경우, 문장 전체의 의미가 반대로 번역된 경우가 많이 나타났고, 다의어인 경우 문맥과 전혀 상관없는 의미로 번역된 경우가 많았다. 대명사는 예측대로 기계번역에서 많은 오류가 발생하였다. 특히 대화문에서 1인칭이나 2인칭을 호칭으로 지칭한 경우, 기계번역에서 주어가 정확히 파악되지 못하였다. 어휘의 의미 번역에서는 기계번역의 가용성이 확인되었으며, 문맥에 맞는 의미 선택에서는 기계번역의 한계가 확인되었다.

문화 범주에서 줄임말은 대체로 정확하게 번역되었으며, 가족 호칭어는 한

국인들에게도 어려운 용어라 기계번역에서 해결하기 어려운 문화 어휘로 판단되었다. 특히 대화문에서 가족 호칭어가 나타날 때 호칭어가 영똥한 사람을 지칭하는 번역으로 나타났다. 음식 이름 번역에서 ‘김치’는 음역으로 나타났고, 다른 음식 이름은 음역되지 않았지만 음식으로 인식되어 번역되었다. 예상과 달리 줄임말 번역에서 기계번역의 가용성을 확인하였고, 가족 호칭어에서 기계번역의 한계가 확인되었다. 음식 이름의 음역은 아랍세계에서 한국문화콘텐츠의 전파에 달려있다고 본다.

『82년생 김지영』의 한글텍스트 총 93,075자, 아랍어 번역텍스트 총 30,794개 아랍어 어절 분석은 적지 않은 분량이지만 한 작품에 국한되어 분석 자료를 객관화하기는 어렵다. 그렇지만 국내에서 최초로 문학텍스트의 한국어-아랍어 기계번역을 시도하고 인간번역과 비교를 통해 기계번역의 가용성과 한계를 확인한 점에서 의미가 있다. 향후 본 연구 결과를 기반으로 프리에디팅과 포스트에디팅을 활용한 교육 분야, 비문학텍스트와 문학텍스트의 기계번역 비교 연구 등 후속 연구를 진행할 예정이다.

참고문헌

- 곽순례 (2017) 「아랍어 번역에 나타난 문화 어휘에 관한 고찰—신경숙의 『엄마를 부탁해』를 중심으로—」, 『아랍어와 아랍문학』 21(3): 55-87.
- 곽순례 (2021) 「아랍어-한국어 기계번역에 관한 고찰—비문학텍스트를 중심으로—」, 『아랍어와 아랍문학』 25(1): 1-27.
- 김재희 (2018) 「한국문학작품에 나타난 문화서 번역 방법: 한·영, 한·아 번역을 중심으로」, 『글로벌문화연구』 9(2): 27-48.
- 김현, 손병우 (2020) 「여성혐오 담론의 경합과 공존: 소설 <82년생 김지영> 현상을 중심으로」, 『한국언론정보학보』 103: 127-146.
- 김효중 (2004) 『새로운 번역을 위한 패러다임』, 서울: 푸른사상사.
- 노진서 (2007) 「한·영 시에 활용된 문화소의 번역 전략」, 『이중언어학』 35: 127-146.
- 엄익란 (2007) 『이슬람의 결혼문화와 젠더』, 서울: 도서출판 한울.

- 엄익란 (2015) 『금기, 무슬림 여성을 엿보다』, 서울: 도서출판 한울.
- 이계연 (2021) 「포스트에디팅 기법을 도입한 아랍어 번역교육」, 『통번역교육연구』 19(1): 81-106.
- 이근희 (2005) 『번역산책』, 서울: 한국문화사.
- 이유아 (2020) 「소설 『82년생 김지영』의 번역에서 나타난 젠더 표현 양상에 대한 고찰—여성문말표현을 중심으로—」, 『일본어문학』 88: 149-178.
- 이형진 (2011) 「신경숙의 Please Look After Mom의 영어 서평에 나타난 문학 번역 평가의 관점」, 『세계문학비교연구』 37: 303-328.
- 지인영, 김희동 (2020) 「신경망기계번역 기술 진화와 번역품질 분석」, 『통번역학연구』 24(3): 191-223.
- 진소영 (2019) 「아랍어-한국어 기계번역의 오류 유형 및 해결방안」, 『아랍어와 아랍문학』 23(4): 53-69.
- 차현주, 한연추, 권상희 (2021) 「영화 <82년생 김지영>의 한·중 리뷰에 나타난 페미니즘 프레임 연구」, 『영화연구』 89: 5-46.
- 최단 (2022) 「문화 부등성에 따른 중국인 학습자의 비판적 수용 연구—『82년생 김지영』의 원본 및 중국어 번역본 비교를 중심으로—」, 『한국어문교육』 39: 483-510.
- 최지영, 이정순 (2022) 「페미니스트 번역전략 ‘협력(co-creation)’의 한중번역 적용을 위한 소고—여성주의 기반작품에서 ‘창조적 변형’을 위한 실천 방법을 중심으로—」, 『한중인문학연구』 74: 23-48.
- 코다야 마유미 (2019) 「『82년생 김지영』의 문화 관련 어휘에 대한 한일 번역 비교」, 『다문화사회와 교육연구』 4: 207-233.
- Baniata, Laith Hasan Okleh (2019) 『A Multitask-Based Neural Machine Translation Model for Arabic Dialects』, 박사학위논문, 경북대학교 대학원.

<인터넷 자료>

- 강종훈 (2022.1.18.) 「‘82년생 김지영’, 최근 해외서 가장 많이 팔린 한국 문학」, 『연합뉴스』, <https://www.yna.co.kr/view/AKR20220118041700005>, 2022년 5월 19일 검색.
- 공감언론 뉴시스 (2022.7.11.) 「연극 ‘82년생 김지영’ 9월 백암아트홀 무대에」,

- 『연합뉴스』, https://newsis.com/view/?id=NISX20220711_0001939148&cID=10701&pID=10700(2022, 2022년 7월 20일 검색).
- 연희선 (2022.7.21.) 「“페미니스트 Yes or No” 소유진 연극 ‘82년생 김지영’ 두고 찬반 논쟁」, 『OSEN』, <http://osen.mt.co.kr/article/G1111901054>, 2022년 7월 25일 검색.
- 이기욱 (2021.10.13.) 「옥스퍼드사전에 우리말 올해 26개 등재 ... 45년치보다 많은 이유」, 『동아일보』, <https://www.donga.com/news/article/all/20211013/109691254/1>, 2022년 6월 30일 검색.
- 임현석 (2021.7.21.) 「‘82년생 김지영’은 이집트 여성들과 비슷해 공감」, 『동아일보』, <https://www.donga.com/news/article/all/20210721/108061812/1>, 2022년 5월 19일 검색.
- 정영현 (2022.1.18.) 「해외에서 가장 많이 팔린 한국 소설은 ‘82년생 김지영」, 『서울경제』, <https://www.sedaily.com/NewsView/260WVHDOAD>, 2022년 5월 19일 검색.
- 최수호 (2021.11.04.) 「‘한국은 어떻게 문화 거물 됐나’...NYT 한국문화콘텐츠 조명」, 『연합뉴스』, <https://www.yna.co.kr/view/AKR20211104170100009>, 2022년 5월 19일 검색.
- حوار مع منار الديناري مترجمة رواية 'مولودة' (2021.3.19.) 'آلاء عبد العاطي' 『1982 عام』, *KOREA.net*, <https://arabic.korea.net/NewsFocus/HonoraryReporters/view?articleId=196119>, 2022년 5월 31일 검색.

<분석 대상>

조남주 (2016) 『1982년생 김지영』, 서울: 민음사.

تشو، نام جو (2021) *مولودة عام 1982*، الجيزة: دار صفصافة للنشر والتوزيع والدراسات

[Abstract]

**A Study on the Errors of Korean-Arabic Machine Translation in
Literary Text—Focusing on *Kim Ji-young, Born in 1982***

Soonlei Gwag

(Hankuk University of Foreign Studies)

This study aims to identify the availability and limitations of neural machine translation (NMT) through a comparative analysis of the Korean-Arabic human translation and NMT of a literary text. To this end, the NMT of Cho Nam-Joo's novel *Kim Ji-young, Born 1982* is analyzed compared to its human translation published in Arabic. Errors in the NMT output are classified into linguistic and cultural categories for quantitative analysis, and exemplary errors are qualitatively examined. According to the findings for the linguistic category, the meaning of words is almost accurately translated by NMT, but many errors involve polysemous words that must be translated according to their context. In particular, when negative words are included, the whole sentence is often translated into the opposite meaning. Many errors occur in the translation of pronouns as predicted. In the cultural category, abbreviations and food names are translated correctly in general, but the terms describing family relationships are translated into different meanings in many cases. Although this study uses limited data for analysis, it is meaningful in that it confirms the availability and limitations of Korean-Arabic NMT.

Keywords: Human Translation, Neural Machine Translation, Linguistic Category,
Cultural Category, Literary Text, Korean-Arabic Translation

주제어: 인간번역, 신경망 기계번역, 언어 범주, 문화 범주, 문학 텍스트, 한-아 번역

곽순례(<https://orcid.org/0000-0001-9196-6465>)

한국외국어대학교 통번역대학원 한아과 부교수

gwag@hufs.ac.kr

관심분야: 통번역교육, 통번역평가, 아랍 및 이슬람 문화 교육, 기계번역, 의료
통역

논문투고: 2022년 8월 7일

1차심사 완료: 2022년 8월 26일

2차심사 완료: 2022년 9월 9일

게재 확정: 2022년 9월 19일