

한국어 문화소에 관한 기계번역과 인간번역의 비교 분석 연구

노금송* · 왕원**

〈차 례〉

- 서론
- 연구대상과 연구방법
- 기계번역과 인간번역의 비교 분석 결과
- 결론

[국문초록]

본고는 한국어 문화소를 연구대상으로 총 92개의 문화소 관련 문장을 기계번역 시스템인 유다오와 파파고를 사용하여 얻은 번역 결과물을 인간번역과 비교 분석하였다. 통계 결과 기계번역의 오류량은 무려 인간번역의 5.5배에 달하였다. 형태적 오류에서 인간번역이 기계번역보다 어휘 대체 및 누락 오류가 더 큰 비중을 차지하였으며 통사적 오류에서 기계번역이 인간번역보다 오류가 더 큰 비중을 차지하였다. 기계번역의 오류는 형태, 통사 두 영역에서 두루두루 다양하게 나타났다. 번역 품질을 평가한 결과 10점 만점에서 인간번역은 9.03점, 유다오번역은 6.5점, 파파고번역은 6.44점으로 인간번역이 기계번역에 비해 월등히 높은 점수를 받았다. 기계번역은 특정문화어, 관용구, 민속풍습, 역사문화 등 문화소와 관련된 어휘와 문장 번역에서 정확성이 두드러지게 떨어졌다. 복문의 경우 분절, 어순 배열 등에 치질이 생겨 거의 모든 문장에서 오류가 발생하여 가독성이 아주 낮았다. 이를 기반으로 기계번역의 오류 요인을 한중 어휘 쌍 데이터의 부족, 문장표현 및 문법 지식의 결여, 번역전략 및 번역방법의 결여 등으로 꼽았다.

[주제어] 기계번역, 인간번역, 문화소, 오류, 비교 분석

* 베이징제2외국어대학 한국어학과 교수

** 베이징제2외국어대학 한국어학과 석사과정생

1. 서론

인공지능 기술의 발달과 함께 신경망 기반 기계 번역기는 통번역 현장에 서나 한국어 교육 현장 및 해외 관광과 같은 일상생활에까지 널리 활용되고 있다. 카카오톡(Kakao talk)의 경우 외국어로 보낸 문자를 ‘복사하기’하여 메시지 입력창에 ‘#’를 입력한 후 ‘붙이기’ 하고 검색하면 자동으로 카카오톡 AI 번역 페이지가 제공된다. 중국의 대표적인 메신저 앱 위챗(Wechat, 微信)에서는 외국어로 된 텍스트나 그림을 길게 누르면 여러 항목이 나오는데 거기에서 번역을 클릭하면 자동으로 중국어로 번역해준다. 그만큼 규칙기반, 통계기반, 신경망 기반의 발전을 거친 네이버 파파고, 구글, 유다오(有道), 바이두(百度) 등 기계번역의 정확도가 점점 높아지고 있음을 설명해주고 있다. 한국 최대 포털 사이트에서 운영하는 네이버의 파파고와 카카오톡이 번역기 같은 경우, 풍부한 한국어 데이터를 확보하고 있기에 번역 결과에 있어 비교적 우수하다는 평가를 받고 있다. (이성화 · 김세현, 2018; 황요한 · 이제영 · 신동진, 2020; 곽중철 · 한승희, 2018; 이주리아, 2018)

기계번역의 정확성과 우수성이 평가되면서 일부 매체들의 조사 보도에 따르면 인공지능이 대체할 직업 1순위로 번역가, 통역사로 꼽기도 하였다.¹⁾ 업계 내에서 번역가, 통역사의 인공지능 대체설에 대한 찬반 논쟁이 이어지고 있고 학술계에서도 신경망 기계번역에 관한 연구가 한국과 중국에서 모두 활발히 이루어지고 있다. 그간의 기계번역에 대한 논의는 통번역학 및 컴퓨터공학을 중심으로 이루어졌고 본 연구와 관련이 있는 통번역학계에서는 오류 유형과 분석, 기계번역 평가 및 정확성 분석 등을 중심으로 논의가 이루어졌다.

기계번역 오류 유형과 분석에 관한 연구로 김윤정(2018), 곽중철 · 한승희(2018), 기유미(2018), 이서이 · 한용수(2022), 임소영 · 임소정(2023)

1) 시사IN: 과학책 번역가가 생성 AI 시대를 바라보는 방식(<https://v.daum.net/v/20230519075148484>), 조선일보: AI는 번역가를 대체할까? “속도 빨라 기능” “문맥까진 안돼”(<https://v.daum.net/v/20230417030150675>) 등 참조.

등이 있는데 기유미(2018)의 연구를 보면 문형별로 분류한 중국어 문장을 각각 네이버 파파고와 Google 번역기에 입력한 뒤, 결과물을 비교 분석하고 오류 유형을 다음과 같이 세분화하였다. 먼저, SVO형식 문형, 把자문, 被자문, 过자문의 번역은 비교적 원활했으나, 형용사술어문, 동사술어문, 의문문 등에서 오류가 비교적 많이 나타났다. 다음으로 오류 유형의 경우 서술어 선택 및 누락의 문제가 가장 두드러졌으며 불필요한 단어의 부가, 어순오류 등으로 정리할 수 있었다. 끝으로 단어의 중의성 문제를 제기하였는데 간단한 문장임에도 불구하고 기계는 사람과 같이 맥락을 읽어내지 못해 부정확한 의미로 번역하는 경우가 있다고 하였다. 기계번역 평가 및 정확성에 관한 연구로 강병규·이지은(2018), 한승희(2021), 이지용(2022) 등이 있는데 강병규·이지은(2018)의 연구를 보면 신경망 기계번역 서비스를 제공하는 기업 가운데 구글, 네이버, 바이두 3사의 번역기 작동 원리를 설명하고, 중한 번역의 실례를 통해 번역의 정확률을 제시하였다. 연구대상은 중국어 의사소통 기본표현 170문장, 실용문 200문장으로 단어 번역, 구 번역, 문장 번역의 측면으로 구분하였고 10점 척도로 평가하였다. 연구 결과, 의사소통 기본표현의 경우 네이버 신경망 기계번역의 정확률이 가장 높았고, 실용문 번역에 있어서 구글 신경망 기계번역과 네이버 신경망 기계번역의 정확률이 높은 것으로 나타났다. 인간번역과의 품질 비교에 관한 연구로 김가희·조영래(2021)가 있는데 수필 문학 작품을 대상으로 인간번역과 기계번역의 결과물을 비교 분석하여 번역 품질 면에서 인간번역과 기계번역은 큰 차이를 보이지 않았고 가독성도 우수하다고 제시하였다. 이 밖에도 중국인 한국어 학습자를 대상으로 한 기계번역 사용과 활용에 관한 연구로 전미화(2022a, b), 기계번역 포스트에디팅에 관한 연구로 周童·朱明爰(2022) 등이 있다.

이상에서 살펴본 바와 같이 기계번역에 관한 학계에서의 연구는 언어 쌍의 번역 방향이 대부분 중한이 주를 이루었고, 중국의 한국어학계에서 한중 기계번역에 관한 연구는 영어-중국어, 일본어-중국어 등 기타 언어 쌍에 비해 극히 드물게 이루어졌음이 중국 학술데이터베이스 CNKI에서 검색되었

다. 이에 따라 본고는 신문이나 기본적인 의사소통 표현 등 기계번역이 비교적 쉽게 처리가 가능한 텍스트가 아닌 한국 소설에 나타난 문화소 관련 문장을 연구대상으로 기계번역과 인간번역의 결과물을 비교 분석하고자 한다. 인간번역과의 비교를 통해 기계번역의 오류 유형, 오류 분포 등 구체적 정보를 보다 정확하게 분석한 기초 상에서 오류가 발생하게 된 요인을 살펴봄으로써 기계번역 시스템의 개발과 수정에 도움을 제공하고자 한다.

2. 연구대상과 연구방법

1) 연구대상

본고는 1992년 한국일보 창작문학상을 수상한 『녹천에는 똥이 많다』에 나타난 한국어 문화소를 연구 범위로 잡았다. 본 작품은 1980년대 한국 사회의 모습을 생생하게 그려낸 작품으로 억압적인 정치 체제와 불평등한 사회 구조 속에서 살아가는 사람들의 고통과 절망을 리얼하게 표현하였다. 한국 현대문학의 중요한 작품으로 80년대 한국의 사회생활의 일면을 그대로 보여주었기 때문에 이 작품에 나타난 문화소가 포함된 문장을 연구대상으로 잡았다. 문화소의 개념에 대해 이근희(2005)는 ‘원천 언어를 사용하는 사회 공동체의 역사, 경제, 정치, 언어 관습을 둘러싼 고유의 특정 문화에서 비롯되는 어휘’로 규정하였고, 권인경(2013:39)은 ‘문화적인 특수한 어휘뿐만 아니라 한 사회에서 지니고 있는 사회적, 문화적 배경에서 사용하는 언어적 관습과 표현’으로 정의하고 있다. 본고에서도 학계에서 정의한 개념을 받아들여 ‘원천 문화권의 특성이 반영된 고유한 표현’을 연구대상으로 삼았다.

이근희(2003: 6-7)는 문화소를 고유명사, 특정 문화와 관련된 어휘, 특정 사건과 관련된 어휘, 관용구, 도량형 단위, 언어사용 관습에서 비롯되는 표현으로 분류하고 있는데 여기에서 고유명사, 도량형 단위의 번역은 기계번역이 상대적으로 쉽게 처리가 가능한 부분이라고 생각돼 본고의 연구대상

에서 제외하기로 하였다. 나머지 부분을 문장 단위로 추출한 결과 모두 92개가 선정되었다. 이 92개의 문장을 세분하면 의식주행(설렁탕, 자취방 등), 사회역사(백골단, 달동네 등), 정치(재야 단체, 적화 통일 등), 문화관습(사주 관상, 토정비결), 특정사건(6월항쟁, 80년 광산 폭동 사건 등), 관용구(마른하늘에 날벼락 떨어지다, 뒤로 자빠져도 코가 깨지다 등)로 나눌 수 있는데 이를 본고의 연구대상으로 삼았다.

2) 연구방법

본고에서는 먼저 앞에서 선정된 문화소 관련 한국어 원문을 한국과 중국에서 상대적으로 많이 활용되고 있는 파파고와 유다오 두 번역기를 사용하여 번역 결과물을 도출해낼 것이다. 다음 역자 春喜(2021)에 의해 번역된 『鹿川有许多粪』과 비교하여 기계번역의 오류 유형과 번역 품질에 대해 분석할 것이다. 이 책은 중국의 정규 출판사인 우한대학출판사(武汉大学出版社)에서 출간한 역서로 이를 참조물로 기계번역에 존재하는 문제점을 살펴보고 오류 요인을 분석함으로써 번역시스템 개발과 수정을 위한 근거를 제공하고자 한다.

김윤정(2018), 서보현·김순영(2018), 이상민(2020) 등에서 신경망 기계번역 오류의 유형을 세분화하여 제시하고 있는데 기존연구들은 분류 기준이 명확하지 않아 혼란스러운 느낌을 주고 있으므로 본고는 오류 분류 기준을 문장의 구성단위에 두어 형태와 통사 크게 두 개 영역으로 나누어 분류를 시도해 보았다.

[표 1] 번역 오류 유형

상위 분류	하위 분류
형태적 오류	어휘 대체
	어휘 추가
	어휘 누락
통사적 오류	수식(중심어에 대한 수식에 오류가 발생한 경우)
	문장 어순(문장 어순 배열 오류)
	의연화(관용구나 절의 의미를 풀어서 옮김)
	성분 누락
	단어 배합(단어와 단어의 결합에 생긴 오류)
	절 누락
	논리적 관계(절과 절의 논리적 관계 오류)
	분절
	시제
	의문법
	부정법
	피동법
	사동법

본고는 오류 유형에 대한 분석과 통계를 진행한 후 번역의 결과물에 대해 품질평가를 진행할 것이다. 기계번역의 품질은 컴퓨터가 자동으로 평가하거나 전문가가 주관적으로 평가할 수 있다. 후자는 학계에서 적극적으로 활용되고 있으며 이 과정에서 번역 성능의 정량화를 위해 번역 결과에 점수를 부여하기도 한다. 본고에서도 번역의 결과물에 대한 평가를 20여 년간의 한중, 중한 통번역 강의 및 실천 경력을 가지고 있는 제1저자와 한국어능력시험 6급을 패스한 중국어 모국어 학습자인 제2저자의 주관적 판단에 의존하기로 하였다.

번역 결과물에 대한 객관적인 평가를 위해서는 평가 기준을 먼저 세워야 할 것이다. 신경망 기계번역 품질평가에 대해 기유미(2018), 정해연(2018), 권도경 외(2019), 황지연·양슬아(2020) 등 선행연구에서 인간평가모형을 제시하였는데 이를 종합하여 본고는 번역 품질평가의 척도를 10점으로 정하고 해당 항목에 따라 오류가 발생한 부분을 감점하는 방식을 취한다. 총

92개의 문장을 하나하나 채점하여 득점한 평균치를 최종 평가 결과로 채택할 예정이다. 본고에서 채택한 평가모델을 표로 나타내면 아래와 같다.

[표 2] 번역 품질평가 모델

상위 항목	하위 항목	비중
정확성	단어와 문장 의미의 정확성, 문법 표현의 정확성	4
완결성	단어, 절의 누락이 없는 역문의 완성도	2
일관성	절과 절의 논리성	1
가독성	문장의 유창성, 도착 언어의 모국어 독자 수용성	2
형식성	오타자, 인용 및 문장부호 등	1
합계		10

위의 번역 품질평가 모델에 의거해 평가 과정을 예를 들어 구체적으로 설명하면 아래와 같다.

(1) ST: “짐승도 죽을 때면 고향을 찾는다고 했는데, 역시 그 말이 틀린 게 아니야.”

TT1(인간): “都说落叶归根, 这句话一点也没错。”

TT2(유다오): “俗话说, 动物死后也要回到故乡, 这话果然没错。”

TT3(파파고): “听说野兽死的时候也会回故乡, 果然那句话没有错。”²⁾

예문 (1)은 한국어 문화소인 속담 “짐승도 죽을 때면 고향을 찾는다”가 들어있는 문장인데 인간번역에서 속담을 제대로 이해하여 생략할 것은 생략하고 그에 대응하는 중국어인 ‘落叶归根’이라는 완전히 대응되는 사자성어를 사용해 등가번역을 구현하였기에 10점 만점으로 평가하였다. 유다오 번역은 한국어 속담을 문자 그대로 축자역을 하였는데 이와 같은 번역은 원문의 정보를 제대로 전달하지 못하였을 뿐만 아니라 동물이 왜 고향으로 돌아가야 하는지 오리무중에 빠지게 하므로 정확성 항목에서 2점 감점, 가독성

2) 아래의 예문에서도 인간번역을 TT1로, 유다오번역을 TT2로, 파파고번역을 TT3으로 대체하기로 하였다.

항목에서 2점 감점으로 총 6점으로 평가, 파파고에서도 속담을 글자 그대로 번역하고 속담 뒤에 오는 절에서도 성분 오류가 발생한 관계로 정확성 항목에서 3점 감점, 일관성 항목에서 1점 감점, 가독성 항목에서 2점 감점으로 총 4점으로 평가하였다.

3. 기계번역과 인간번역의 비교 분석 결과

본 장에서는 2장에서 선정된 오류 분류 및 품질평가 기준을 적용하여 인간번역과 기계번역의 오류를 비교 분석하고, 다음 번역의 품질평가를 통해 오류가 생기게 된 요인에 대해 살펴보고자 한다.

1) 문화소 번역 오류 유형

인간번역과 유다오번역, 파파고번역 결과물을 오류 유형별로 살펴본 결과 아래의 [표 3]과 같이 나타났다.

[표 3] 문화소 번역 오류 유형 분류표

상위 분류	하위 분류	인간 (개수)	오류율 (%)	유다오 (개수)	오류율 (%)	파파고 (개수)	오류율 (%)
형태적 오류	어휘 대체	19	59.3	86	50	87	52.1
	어휘 추가			3	1.7	2	1.2
	어휘 누락	5	15.6	7	4.1	3	1.8
동사적 오류	수식	1	3.1	19	11	11	6.6
	문장 어순			9	5.2	15	8.9
	외연화	3	9.3	16	9.3	16	9.6
	성분 누락			7	4.1	6	3.6
	단어 배합			5	2.9	5	3
	절 누락	2	6.2	2	1.2	4	2.4
	논리적 관계			5	2.9	7	4.2

	분절			6	3.5	4	2.4
	시제	1	3.1	1	0.6		
	의문법			1	0.6	1	0.6
	부정법	1	3.1	3	1.7	4	2.4
	피동법			1	0.6	1	0.6
	사동법			1	0.6	1	0.6
합계	32	100	172	100	167	100	

첫째, 번역의 오류량을 보면, 인간번역은 32개, 기계번역인 유다오는 172개, 파파고는 167개로, 기계번역의 오류량이 무려 인간번역의 5.5배에 달하였다. 물론 기계번역이 인간번역에 비해 번역의 속도는 훨씬 빠르지만 정확도는 현저히 떨어진다는 것이 입증되는 부분이기도 하다.

둘째, 형태적 오류를 보면, 어휘 대체 오류가 인간번역에서는 59.3%, 유다오에서는 50%, 파파고에서는 52.1%로 나타나 인간번역에서나 기계번역에서 어휘 대체 오류가 전체 오류에서 절반 이상을 차지하였다. 여기서 주목해야 할 부분은 인간번역이 기계번역에 비해 어휘 대체 오류가 7.2%-9.3%나 많은 비중을 차지한다는 것이다. 예를 들어보면 아래와 같다.

(2) ST: 전쟁이 끝나자 갑자기 그 용천뱅이들이 부쩍 늘어났다. (P58)

TT1(인간): 战争结束后, 龙川白突然多了起来。(P53)

TT2(유다오): 战争一结束, 龙川人就突然增加。

TT3(파파고): 战争结束后, 龙川银鱼突然增多。

예문(2)의 세 번역문은 얼핏 보기에 모두 정문으로 보이나 자세히 들여다 보면 밑줄 친 단어는 중국인 독자들이 이해하는 데 어려움이 따른다. ‘龙川白’, ‘龙川人’, ‘龙川银鱼’가 물건인지, 사람인지, 뱀어인지 알 수가 없다. ‘용천뱅이’라는 말은 전라, 충청 지역의 방언으로 ‘문둥이’라는 의미를 뜻한다. ‘용천’이라는 말은 ‘문둥’이라는 의미이고 ‘뱅이’는 ‘병자’, ‘환자’라는 의미이다. 물론 이 문장의 앞부분에서 이미 ‘문둥이’라는 단어가 나왔기 때문에 이

를 ‘麻风病人’으로 중복 처리하는 데는 무리가 있을 것이다. 그렇다고 ‘龙川’ 뒤에 ‘白’, ‘人’, ‘银鱼’를 붙여 새로운 단어를 만들어 그 의미를 대체시키는 방법은 취할 바가 아니라 생각되며 ‘龙川’의 뒤에 ‘病人’ 혹은 ‘麻风病人’의 방언으로 그 의미를 증가시키는 것이 더 이상적인 번역이라는 생각이 든다.

어휘 대체 오류에 이어 어휘 누락 오류를 보면, 인간번역은 15.6%, 유다오는 4.1%, 과파고는 1.8%로 인간번역의 어휘 누락 오류가 기계번역에 비해 무려 11.5%-13.8%나 높은 비중으로 나타났다. 그만큼 기계번역은 출발언어 문장 내의 어휘를 하나하나 찾아내 누락 없이 번역하였음을 알 수 있고 이에 반해 인간번역은 역자의 실수로 인해 누락되었을 것으로 추측해 볼 수 있다. 다음의 예를 보자.

(3) ST: 남자는 반질반질한 대머리에다 기름때가 묻은 당꼬바지를 입고 있었

는데, 우리를 훑어보는 눈초리가 왠지 사납게 느껴졌어요. (P67)

TT1: 男人是一个()光头, 穿着一条满是油垢的收脚裤。他打量着我们, 眼神里莫名有种凶狠的感觉。(P61)

TT2: 男子光着身子的秃子, 穿着沾满油垢的唐裤, 不知为什么打量我们的眼神让人觉得很凶。

TT3: 男人光秃秃的, 穿着沾满油垢的糖衣裤, 不知怎么的, 打量我们的眼神觉得凶狠。

예문(3)에서 인간번역은 중심어인 ‘대머리’의 수식어 ‘반질반질한’을 생략하였고 유다오번역에서는 원문의 의미와 완전히 다른 수식어로 대체하여 오류가 생겼으며, 과파고에서는 아예 중심어인 ‘대머리’를 생략하여 성분 누락 오류가 발생하였다. 물론 인간번역에서처럼 수식어 ‘반질반질한(发亮的)’을 누락해도 문장의 의미를 전달하는 데는 큰 무리가 되지 않지만 우스꽝스러운 주인공의 이미지 전달이 100% 이루어졌다고 보기는 어렵다.

셋째, 통사적 오류를 보면, 인간번역에서는 절 누락, 외연화 오류가 각각 9.3%, 6.2%로, 전체 통사적 오류에서 1위, 2위를 차지하고 있다. 이는 역자

의 실수, 혹은 번역의 스킬과 관련되는 오류로 추측해 볼 수 있다. 다음의 예를 보자.

(4) ST: 때마침 설렁탕이 나오자 그는 일언이폐지하고 숟가락을 들어 먹기 시작했다. (P16)

TT1: 刚好牛骨炖汤上桌了, 他()抓起勺子开始吃饭。(P10)

TT2: 正好牛杂碎汤出来了, 他一口废除, 拿起勺子吃了起来。

TT3: 恰巧牛杂碎汤一出, 他一言废止, 拿起勺子吃了起来。

(5) ST: “맛나고 싱싱한 오징어나 김밥 사이소오! 둘이 목다가 하나 죽어도 모르는 오징어나 김밥.”(P162)

TT1: “来来来, 快来买好吃又新鲜的鱼糕和紫菜包饭! 新鲜的鱼糕和紫菜包饭!”(P156)

TT2: “好吃新鲜的矣东或紫菜包饭! 两个人吃着吃着死也不知道的矣东或紫菜包饭。”

TT3: “美味新鲜的鱼饼或紫菜包饭csio! 两个人一起住, 死都不知道的鱼饼或者紫菜包饭。”

예문(4)의 ‘일언이폐지하고’를 인간번역에서는 누락하였고, 유다오와 파파고에서는 한국어 관용구를 직역 처리하였는데 누락한 경우 역문에는 오류가 생기지 않았지만 원문에 대한 충실성이 떨어진 완전하지 않은 번역이 되었고 기계번역의 외연화 처리는 우격다짐으로 번역한 느낌이 짙어 중국인 수용자들이 이해하는 데 어려움이 따른다. 여기서 ‘일언이폐지하고’는 중국어 성어인 ‘二话不说’으로 번역하는 게 더 정확한 것으로 여겨진다. 예문(5)의 ‘둘이 목다가 하나 죽어도 모르는’도 한국어 속담인데 중국어에 대응되는 표현이 없는 관계로 인간번역에서나 기계번역 모두에서 외연화 전략을 취하였다. 인간번역은 비문은 아니지만 원문의 의미를 축소 처리했다는 느낌을 주고 유다오는 직역 처리, 파파고는 방언 ‘목다가’를 ‘투숙하다’의 의

미로 대역해 큰 오류를 범하였다. 도착언어에 완전히 대응되는 관용구 표현이 없을 때 그와 유사한 관용구로 대체시켜 “香噴噴、让人回味无穷”과 같이 번역하는 것이 더 적합하다는 생각이 든다.

기계번역에서의 통사적 오류를 볼 때 수식, 외연화, 어순, 성분 등 유형에서 다양하게 오류가 발생하였는데 유다오번역과 파과고번역은 수식 오류에서 각각 11%, 6.6%, 외연화 오류에서 각각 9.3%, 9.6%, 문장 어순 오류에서 각각 5.2%, 8.9%, 성분 누락 오류에서 각각 4.1%, 3.6%로 나타났다. 외연화 오류에 관해서는 예문(4)에서 이미 설명했기에 생략하기로 하고 아래에서는 기타 오류 유형의 예를 들어보도록 하겠다.

(6) ST: 왜냐하면, 장병만씨와 내가 처음 만난 것이 바로 그해 6월 어느날, 소위 닭장차라고 불리는 경찰 호송버스 안에서였기 때문이다. (P7)

TT1: 因为正是在那年六月的某一天, 在名为“鸡笼车”的警用押运大巴上, 我第一次见到了张丙万。(P1)

TT2: 因为张炳万和我第一次见面就是在当年6月的某一天, 在一辆被称为“鸡窝车”的警察护送车上。

TT3: 因为张秉万和我第一次见面是在当年6月的某一天, 在被称为“鸡场车”的警察护送巴士上。

(7) ST: 떨어진 옷 하나로 몇 달을 버티고, 하루 세끼의 끼니를 삼백 원짜리가 락국수나 라면으로만 때우면서도 나는 악착같이 돈을 모았어요. (P71)

TT1: 我执着地攒钱, 一件破衣裳撑几个月, 一日三餐只靠三百韩元的粗面条或者方便面凑合。(P64)

TT2: 用一件掉下来的衣服坚持了几个月, 一天三顿饭都只吃300韩元的面条或方便面, 但我还是拼命攒钱。

TT3: 光靠一件掉下来的衣服撑了几个月, 一天三顿饭只吃三百块的面条或拉面, 我拼命攒钱。

(8) ST: 사람들은 몸과 몸이 밀착되어 서로 밀고 밀리면서도 끊임없이 군중이
 불어나길 원했고 그래서 자주 목소리를 합해 “애국시민 동참하라 홀라
 홀라”를 노래하기도 했다. (P17)

TT1: 大家的身体彼此紧贴, 挤来挤去, 却依然渴望人数的增加, 因此不
 断地齐声歌唱“爱国市民一起来吧, Hula Hula”。(P11)

TT2: 人们希望身体和身体紧贴在一起, 互相推挤, 但群众不断增加, 所
 以经常合着声音唱“爱国市民共同参与吧, 吧, 吧, 吧”

TT3: 人们希望身体和身体紧密相连, 互相推挤, 但群众不断增多, 因此
 经常合声演唱“爱国市民参与吧, 呼啦呼啦”

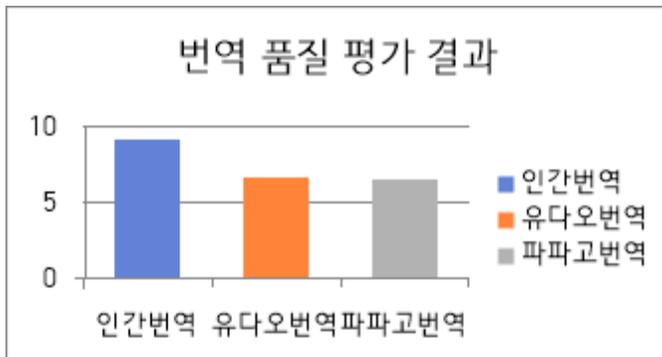
예문(6)의 “소위 닭장차라고 불리는”이 ‘경찰 호송버스’를 수식하고 있는 데 인간번역에서나 기계번역에서 모두 직역 처리하여 비문으로 나타났다. 세 역문에서는 ‘닭장차’라는 이름을 가진 ‘경찰 호송버스’로 대역시켰다. 사실 여기서 쇠창살이 설치된 경찰 호송버스는 닭장과 비슷하다고 하여 비유의 의미로 쓰인 것이지, ‘닭장차’라는 이름을 단 경찰 호송버스를 지칭하는 것이 아니다. 따라서 이 절은 “在比喻成鸡笼的囚车上”으로 번역하는 것이 더 적합한 것으로 보인다. 예문(7)을 보면 인간번역과 달리 기계번역은 축자역의 방식을 취해 원문의 한 구절 한 구절을 그 글귀 그대로 본래의 뜻에 충실하게 번역하였는데 이는 중국어에서 자주 보이는 문장 표현 형식이 아니다. 세 개 절로 구성된 이 문장의 내용으로 볼 때 앞의 두 절과 후행절은 전체와 부분의 관계로 볼 수 있는데 이런 경우 중국어에서는 위의 인간번역과 같이 어순의 조정을 통해 전체를 선행절에 두고 부분을 후행절에 배열한다. 예문(8)의 기계번역을 보면 ‘希望’의 위치가 잘못 배정되어 발생한 오류로 마치 사람들은 몸과 몸이 밀착되어 서로 밀고 밀리기를 원하는 것처럼 보인다. 사실 사람들이 원하는 것은 군중이 불어나는 것임으로 인간번역과 같이 ‘希望’의 의미가 ‘人数的增加’의 앞에 와야 더 적합한 번역이 된다.

전반적으로 볼 때 형태적 오류에서 인간번역은 74.9%, 유다오번역은 55.8%, 파과고번역은 55.1%로, 인간번역이 기계번역에 비해 어휘 대체 및

누락 오류가 절대적 비중을 차지한다는 것을 알 수 있고 통사적 오류에서 인간번역은 25.1%, 유다오번역은 44.2%, 파파고번역은 44.9%로, 기계번역이 인간번역에 비해 형태, 통사 영역에서의 오류가 더욱 다양한 형태로 나타난다는 것을 알 수 있다.

2) 번역 품질평가 및 오류 요인 분석

2장에서 제시한 번역 품질평가 모델에 의거해 인간번역과 기계번역의 결과를 평가한 결과 아래와 같은 점수를 얻게 되었다.



10점 만점에서 인간번역은 9.03점, 유다오번역은 6.5점, 파파고번역은 6.44점으로 인간번역이 기계번역에 비해 월등히 높은 점수를 받았다. 인간번역은 실수에 의한 원문 누락을 피한다면 더 높은 점수를 받을 수 있을 것이다. 중국의 대표적인 번역시스템 유다오와 한국의 대표적인 번역시스템 파파고는 특점이 거의 비슷하여 번역 수준의 우위를 가릴 수 없을 정도이다. 다음의 표는 구체적인 번역 평가 항목에서의 득점 상황이다.

[표 4] 각 평가 항목에서의 득점 상황

평가 항목(점수)	인간번역	유다오번역	파파고번역
정확성(3)	2.75	1.4	1.4
완결성(2)	1.6	1.7	1.68
일관성(2)	1.9	1.45	1.47
가독성(2)	1.88	1.15	1.19
형식성(1)	0.9	0.7	0.7
합계(10)	9.03	6.5	6.44

표에서 보듯이 인간번역은 정확성, 일관성, 가독성 항목에서 각각 2.75 점, 1.9점, 1.88점을 득점하여 기계번역보다 월등히 높은 점수를 받아 번역의 질이 훨씬 우위에 있음이 입증되었다. 완결성 항목에서 인간의 실수와 능력의 한계로 인해 기계번역에 비해 점수가 낮기는 하지만 전반적으로 만족할 만한 결과라고 볼 수 있겠다. 기계번역의 완결성은 인간번역을 능가하지만 문장 하나하나를 살펴보면 정확성, 일관성, 가독성 면에서 인간번역보다 못하다. 특히 특정문화어, 관용구, 민속풍습, 역사문화 등 문화소와 관련된 어휘와 문장 번역의 정확도는 아주 낮으며 심지어 복문의 경우 분절, 어순 배열에 차질이 생겨 거의 모든 역문에서 오류가 발생하여 가독성이 매우 낮다. 기계번역에 이런 오류가 생기게 된 요인은 무엇일까? 위의 분석을 기반으로 아래와 같이 정리해 보도록 하겠다.

첫째, 가장 손꼽히는 부분은 번역시스템에 한중 어휘 쌍의 부족, 또 혹은 대응되는 어휘가 시스템에 있다고 하더라도 여러 의미항 중에서 해당 문장 담화 맥락과 의미에 맞게 어울리는 단어를 선택해야 하는 부분에서 의미를 정확하게 인식하지 못하는 경우가 있었다. 예컨대 자취방, 오텍, 양반, 동기, 국민학생, 전세금, 운동권, 대공과 등 문화특정어는 물론, 당꼬바지, 아버지, 용천뱅이 등 방언과 오리발 내밀다, 고무신 거꾸로 신다, 못된 땅아지 빨나다 등 관용구는 의미를 이해하지 못해 직역으로 처리하였다.

둘째, 대화 중 행위자 표현에서 한국어는 주어 생략하는 무주어문을 많이 사용하는 편이지만 중국어는 주어문이 대부분이다. 기계번역은 이런 문

장표현의 특징을 학습하지 못한 데다가 문장 담화 맥락을 인식하지 못하는 관계로 한국어의 무주어문을 그대로 중국어의 무주어문으로 번역함으로써 오류가 발생한다. 예를 들면, “그저 그때가 육이오 무렵이었으니까, 전쟁통에 부모를 잃어버린 게 아닌가 짐작할 뿐이었지요.”와 같은 경우 주어가 생략되었는데 유다오와 파파고에서 주어들을 모두 생략한 “当时正值六二五之际, 只能推测是不是因为战争失去了父母。”, “只是因为当时是六零五左右, 所以推测是不是因为战争失去了父母。”로 번역하여 오류가 발생하였다. 이에 반해 인간번역은 “我只是猜测, 那时正值六二五战争时期, 可能是在战乱中失去了父母吧。”로 번역하여 주어를 명확히 하였다.

셋째, 부정법, 사동법, 피동법, 의문법, 시제 등 문법 지식의 부족으로 인해 오류가 발생하였다. 예를 들면, “하여튼 이번에 지서장님 수고가 많았어요. 오래간만에 월척 하나 낚으신 거나 아닌지 모르겠네.”와 같은 문장을 보면 유다오와 파파고는 각각 “不管怎样, 这次分行长辛苦了。不知道是不是好久没钓到一条大鱼了。”, “总之, 这次支署长辛苦了。不知道是不是好久没钓到一条大鱼了。”으로 번역하였는데 번역문의 문말을 보면 이상하지 않을 수 없다. 원문에서는 부정을 통한 추측의 의미를 표현한 것인데 기계번역은 제대로 이해하지 못해 모두 반문으로 표현하였다. 추측의 의미인 “总之, 支署长这次辛苦了。不知这次算不算钓了一条久违的大鱼。”로 번역하는 게 더 정확할 것이다.

넷째, 번역은 출발언어를 읽고 이해하고 그에 대응되는 어휘를 찾아 순서를 배열하고 도착언어로 표현한 다음, 그 역문을 원문과 대조하여 교정 및 감수하는 등 일련의 과정을 거친다. 어느 한 일환에서 차질이 생긴다면 오류가 발생하기 쉽다. 신경망 기반 기계번역은 훈련 데이터, 딥러닝, 세밀한 알고리즘에 의해 역문이 도출된다. 물론 속도는 빠르지만 도착언어의 표현 습관에 맞게 적절하게 표현하는 데는 아직 부족한 듯하다. 인간이 번역할 때 사용되는 이국화, 자국화 번역전략이나 문맥 상황에 맞게 정보를 추가하고 생략하고 또 긴 문장을 분절하고 성분의 앞뒤 순서를 조정하는 등 번역 방법에는 능숙하지 않다. 예를 들면, “이렇게 병약한 노인들까지 간첩 조직

에 이용하는 것은 적화 통일을 위해서는 수단과 방법을 가리지 않는 북괴의 악랄성을 다시 한번 입증하는 것인데, 이들은 십여 년 전부터 지령에 따라 중요 정보를 수집 보고해왔고, 검거 당시에 난수표, 공작금, 단파라디오 등 의심의 여지가 없는 각종 증거품까지 압수되었다.”와 같은 문장은 절이 서로 뒤엉켜진 아주 복잡한 복문이다. 유다오와 파파고는 각각 “这样体弱多病的老人到应用于间谍组织为实现赤化统一, 手段和方法尹锡宪不分的恶劣省再次证明, 他们在十年前就按照指令一直在收集重要信息报告、检举当时暗号票、元氏、短波收音机等产生怀疑的余地的各种证据到扣押了。”, “连体弱多病的老人也利用在间谍组织上, 再次证明了为了赤化统一不择手段的‘北傀’的恶毒性, 他们从十多年前开始根据指令收集并报告重要情报, 被检举当时没收了乱收票、工作金、短波收音机等各种毫无疑问的证据。”로 번역하였다. 중심어에 대한 수식, 성분의 순서, 분절 등 여러 항목에서 오류가 발생하였다. 하지만 인간번역을 보면 “为了赤化统一, 把老弱病残也利用到间谍组织, 这种不择手段的做法, 再次证明了北部傀儡集团的恶毒性。 他们从十几年前开始接受指令并收集重要情报, 拘捕时的随机数表、活动经费、短波无线电等各种确凿的证物已被全部没收。”로 이런 문제점을 모두 해결하였다.

4. 결론

본고는 『녹천에는 똥이 많다』에 나타난 한국어 문화소 관련 문장을 연구 범위로, 총 92개의 문화소 관련 문장을 추출하여 중국의 대표적 번역시스템인 유다오와 한국의 대표적 번역시스템인 파파고를 각각 사용하여 얻은 번역 결과물을 인간번역 결과물인 『鹿川有许多粪』과 비교 분석하였다. 먼저 번역 결과물의 오류 유형을 형태적, 통사적 측면으로 나누어 살펴보았는데 기계번역의 오류량은 무려 인간번역의 5.5배에 달하였다. 형태적 오류의 비중을 보면 인간번역은 74.9%, 유다오번역은 55.8%, 파파고번역은 55.1%

로, 인간번역은 어휘 대체 및 누락 오류가 절대적 비중을 차지하였다. 통사적 오류의 비중을 보면, 인간번역은 25.1%, 유다오번역은 44.2%, 파파고번역은 44.9%로, 기계번역의 오류는 형태, 통사 두 영역에서 두루두루 다양하게 나타났다.

다음, 번역 품질평가 모델에 의거해 비교 분석한 결과 10점 만점에서 인간번역은 9.03점, 유다오번역은 6.5점, 파파고번역은 6.44점으로 인간번역이 기계번역에 비해 월등히 높은 점수를 받았다. 기계번역은 정확성, 일관성, 가독성 등 세 개 항목에서 득점이 낮고 특히 특정문화어, 관용구, 민속풍습, 역사문화 등 문화소와 관련된 어휘와 문장 번역에서 더욱 두드러지게 나타났다. 복문의 경우 분절과 어순 배열에 차질이 생겨 거의 모든 문장에서 오류가 발생하여 가독성이 아주 낮았다. 이를 기반으로 기계번역의 오류 원인을 한중 어휘 쌍 데이터의 부족, 문장표현 및 문법 지식의 결여, 번역전략 및 번역방법의 결여 등으로 꼽았다.

기계번역의 활용이 폭넓은 범위에서 이루어지고 있는 배경하에서 본 연구는 기계번역의 발전 수준 및 품질 상황을 살펴봄으로써 번역시스템의 개발과 수정 등에 시사점이 있을 것이라 기대된다. 또한 기계번역이라는 기술적인 수단을 이용한 통번역 수업의 도입에도 현실적인 의의가 있을 것이라 여겨진다. 하지만 문화소 관련 연구대상의 추출로부터 오류 유형의 분류, 번역 품질의 평가 등에 이르기까지 모두 객관적인 기술 수단이 아닌 주관적 판단에 의해 산출된 것이라 신빙성이 다소 떨어진다. 이런 부족 부분은 향후 연구에서 개선, 보완되어야 할 것이다.

■ 참고문헌

- 곽중철 · 한승희, 「포스트에디팅 측정지표를 통한 기계번역 오류 유형화 연구」, 『통번역학연구』 22-1, 한국외국어대학교 통번역연구소, 2018, 1~25쪽.
- 권인경, 「문화소 번역 방법 연구 - 제주도 관광안내텍스트를 중심으로 -」, 『동화와 번역』 25, 건국대학교 동화와 번역연구소, 2013, 37~59쪽.
- 기유미, 「한중 기계번역 오류의 문형별 비교분석: 네이버 파파고 번역기와 구글번역기를 중심으로」, 『중국어연구』 74, 한국외국어대학교 중국연구소, 2018, 3~32쪽.
- 김가희 · 조영래, 「수필번역에서 인간번역과 인공지능번역의 비교분석 - 바스넷의 『번역의 성찰』을 중심으로」, 『인공지능인문학연구』 제8권, 중앙대학교 인문콘텐츠연구소, 2021, 52~70쪽.
- 김윤정, 「구글 번역」 기반 텍스트 유형별 중한 기계번역 오류 현황 분석」, 『中國語文學誌』 第65輯, 중국어문학회, 2018, 321~355쪽.
- 이근희, 「문화와 밀접한 상관어의 번역전략」, 『번역학연구』 14(2), 한국번역학회, 2003, 5~27쪽.
- _____, 『번역의 이론과 실제』, 한국문화사, 2005.
- 이서이 · 한용수, 「중국어 색채어 '黑' 신경망 기계번역의미 오류 분석 - '黑'류 형용사가 쓰인 문장의 중한번역을 중심으로 -」, 『비교문화연구』 제66집, 경희대학교 글로벌인문학술원, 2022, 69~99쪽.
- 이성화 · 김세현, 「영-한 및 한-영 기계번역 품질향상을 위한 프리에디팅 기법 제안」, 『번역학연구』 19-5, 한국번역학회, 2018, 121~154쪽.
- 이주리애, 「인공신경망 기계 번역의 한일/일한 번역 품질에 대한 예비연구: 품질향상 검토와 교열코드 적용」, 『통역과 번역』 20-1, 한국통역번역학회, 2018, 43~71쪽.
- 이지용, 「한국어 관용구의 기계 번역 양상 연구 - {마음+서술어}형태를 중심으로 -」, 『한국어 의미학』 78, 한국어어미학회, 2022, 59~88쪽.
- 이창동, 『녹천에는 똥이 많다』, 문학과지성사, 1992.
- 임소영 · 임소정, 「구어체 한-중 시번역의 오류 양상 연구 - 넷플릭스 「솔로지옥」 발화를 중심으로 -」, 『中國語文學』 第92輯, 영남중국어문학회, 2023, 299~330쪽.
- 전미화, 「중국 대학 한국어 학습자의 기계번역 사용 경험 및 인식 연구」, 『인문사회21』 제13권 6호, 2022a, 953~962쪽.
- _____, 「중국 대학 한국어 교육에서의 기계번역 활용 방안 연구」, 『인문사회21』 제13권 6호, 2022b, 1207~1216쪽.
- 한승희, 「기계번역 품질에 대한 요인분석」, 『통번역학연구』 제25권 2호, 한국외국어대학교 통번역연구소, 2021, 147~170쪽.
- 황요한 · 이재영 · 신동진, 「온라인 기계 번역기의 과거, 현재 그리고 미래: 오류 분석을 중심으로」, 『언어연구』 36-2, 한국현대언어학회, 2020, 191~210쪽.
- 황지연 · 양슬아, 「한국어-중국어 신경망 기계 번역에 대한 수동평가」, 『통번역학연구』 제24권 4호, 한국외국어대학교 통번역연구소, 2020, 175~206쪽.
- 春喜, 『鹿川有许多粪』, 武汉大学出版社, 2021.
- 周童 · 朱明爱, 「韩汉机器翻译的译后编辑研究 - 以Naver Papago为例」, 『韩国语教学与研究』 第四期, 2022, 152~160쪽.

A comparative analysis between human translation and machine translation of Korean culture-bound terms

Lu Jinsong* · Wang Yuan**

A comparative analysis between human translation and machine translation of Korean culture-bound terms

This paper took the Korean culture-bound terms as the research object. 92 sentences were extracted and input into Youdao and Papago translation system respectively, and the results were compared with the human translation. Statistics showed that the number of errors in machine translation (MT) was up to 5.5 times that of human translation (HT). The proportion of lexical substitution and omission in morphological errors in HT was larger than that in MT, while MT accounted for a larger proportion of syntactic errors. In the quality evaluation results out of 10 points, HT scored 9.03 points, which was significantly higher, with Youdao 6.5, and Papago 6.44. The accuracy of MT decreased significantly when translating idioms and words related to folk customs, history and society. Errors in segmentation and word order arrangement of complex sentences appeared in almost all MT results, leading to low readability. The error factors of MT included the lack of bilingual vocabulary data, correct grammar knowledge, and proper translation strategies.

Key words: machine translation, human translation, culture-bound terms, errors, comparative analysis

논문투고일: 2024년 3월 15일 || 심사완료일: 2024년 5월 13일 || 게재확정일: 2024년 5월 21일

* Professor, Department of Korean, Beijing International Studies University

** Graduate student, Department of Korean, Beijing International Studies University