

# 구어체 한-중 AI번역의 오류 양상 연구

- 넷플릭스 「슬로지옥」 발화를 중심으로 -

임소영\* · 임소정\*\*

## <目次>

I. 서론	III. 구어체 한-중 AI번역의 오류 양상
1. 문제 제기	1. 오류 유형의 분류 기준
2. 선행 연구	2. 오류 유형별 사례 분석
II. 연구 대상 및 방법	IV. 결론
1. 연구 대상	
2. 연구 방법	

## I. 서론

### 1. 문제 제기

IT 기술이 비약적으로 발전함에 따라 현대인의 일상은 인터넷, 컴퓨터 그리고 스마트 기기들과 불가분해졌다. 중국어를 공부할 때도 모르는 단어가 나오면 온라인으로 검색하고 생소한 중국어 자료를 읽어야 할 경우에도 기계번역의 도움을 받는 등 모든 과정이 빠르고 간편해졌다. 인터넷이 대중화되기 이전에는 모르는 한자를 검색하려면 종이 사전을 찾아야 했는데, 그러려면 해당 한자의 부수(部首)를 먼저 식별한 다음, 부수를 제외한 나

\* 한국방송통신대학교 실용중국어학과 겸임교수: 제1저자

\*\* 한밭대학교 중국어과 부교수: 교신저자

머지 필획의 총 획수를 계산하는 등 적잖은 시간과 수고를 필요로 하였다.

이러한 아날로그 방식도 2000년대 초반 전자사전이 등장하면서 빠르게 사라지기 시작하였다. 성인의 손바닥만 한 전자사전 안에 종이책 수백, 수천 권의 내용이 데이터베이스화 되고, 원어민 음성 지원 등의 부가 기능도 탑재되니 외국어 학습이 보다 편리해졌다. 인터넷 기술이 보편화된 이후로는 전자사전 역시 웹 기반의 온라인 어학사전, 어플리케이션 기반의 모바일 사전 등에 의해 대체되었다. 뿐만 아니라 대량의 언어 코퍼스를 기계학습(machine learning)하여 텍스트 번역 서비스를 제공하는 대형 플랫폼들이 생겨나 전 세계의 사용자들에게 편의를 제공하고 있다. 대표적인 예가 구글(Google), 네이버(Naver), 바이두(百度) 등 각국의 빅테크 기업들의 번역 서비스인 구글번역(Google Translate), 네이버 파파고(Naver Papago), 바이두번역(百度翻译) 등이다.

물론, 기계번역의 초기 발전 단계에는 번역 결과물의 품질이 그리 좋지 않았고, 실소를 유발하는 오역들도 종종 발생해 서비스 신뢰도를 크게 떨어뜨리고는 했다. 2016년 구글을 필두로 딥러닝에 기반한 신경망 기계번역(NMT: Neural Machine Translation) 기술이 새롭게 적용되면서, 기계번역의 전반적인 품질에 혁신적인 개선이 있었고, 기계번역은 편리하고 효율적인 보조 수단으로 자리매김하고 있다. 김윤정(2018)이 중국어 교과목을 수강하는 대학생 47명을 대상으로 설문 조사한 결과, 중국어를 학습한 경력과는 무관하게 수강생 전원(100%)이 중한번역을 할 때 기계번역 프로그램을 사용한다고 응답하였다. 강수정(2021)에 의하면, 통번역대학원 전공생 205명 중에서 기계번역을 학습에 적용하여 사용하고 있는 비율이 88.8%(182명)으로 10명 중 9명이 사용하고 있었고, 주로 사용하는 플랫폼은 네이버 파파고 73.1%(133명)와 구글번역 22.0%(40명)이었다. 굳이 외국어 전공자가 아니더라도 기계번역은 의외로 일상의 다양한 상황에서 폭넓게 활용되고 있다. 2022년 기준, 영국의 다양한 연령, 성별, 인종으로 구성된 1200명의 참가자를 대상으로 조사한 결과, 기계번역을 사용한 경험이 있는 비율이 75.9%(911명)이었고, 사용 목적은 외국어로 된 텍스트 읽

기, 인터넷 서핑, 소셜미디어 활동 등 일상적 용도(80.1%)가 가장 높았으며 업무(27.8%), 학습(22.8%) 목적은 그 후순위였다. (Vieira et al., 2022)

이렇듯 기계번역의 성능과 품질이 계속 발전하면서 사용 범위와 활용도 역시 확대되고 있지만, 인간의 고유한 감성과 창의력이 들어간 텍스트를 기계가 번역하기에는 아직 역부족이라는 것이 정설이다. 예컨대, 형식면에서 문법을 정확하게 따르고 구조면에서 논리 관계가 명확하며 내용면에서는 객관적 사실을 건조하게 기술한 기사문, 보고서, 제품 기술 설명서 등 정형화되어있고 비문학적이며 정보성이 강한 텍스트는 기계번역이 대체로 양호하게 처리한다. 반면, 인간의 상상력과 창의력을 기반으로 한 문학 텍스트, 그리고 사회적 경험의 총체인 상식(common sense)이 전제되어야 비로소 해석이 가능한 농담이나 언어유희 등은 기계번역이 상대적으로 미진한 영역이다. 이에 본 연구는 대한민국의 20~30대 젊은이들이 일상생활에서 실제로 사용하는 구어(spoken language)를 기계번역이 어떻게 처리할까 하는 문제의식에서 출발한다.

본고는 넷플릭스의 연애 리얼리티 프로그램 <솔로지옥> 시즌2 출연진과 방송 패널들의 발화(utterance) 데이터를 연구 대상으로 하여, 한국과 중국을 대표하는 신경망 기계번역 플랫폼 5종 (네이버 파파고, 카카오번역, 百度翻译, 腾讯翻译, 有道翻译)의 한-중 기계번역 결과를 고찰한다. 그 가운데 다수의 기계번역에서 공통적으로 나타나는 오류 유형을 집중적으로 귀납하여 분석하고자 한다.

## 2. 선행 연구

최문선(2019)에 의하면 신경망 기계번역 관련 학술 연구는 해당 기술이 본격적으로 상용화된 이후인 2017년을 기점으로 급속도로 증가하였다. 이상빈(2020)은 기계번역을 다룬 인문학 논문 121편을 분석한 결과, 연구 주제 면에서는 기계번역 정확성/오류(63%), 인간번역과의 비교(31%), 기계 간 성능 비교(19%), 교육 도구로서의 활용성(23%), 프리/포스트 에디

팅(30%) 등이 있었고, 프리에디팅은 포스트에디팅에 비해 주제 다양성이 상대적으로 부족하다고 평가하였다. 분석에 활용된 번역기는 구글 번역(73편)과 네이버 파파고(36편)가 가장 많았다.

한-중 기계번역으로 범위를 좁혀 들어가면, 강병규·이지은(2018)은 신경망 기계 번역 모델의 작동 원리를 설명하고 초급 수준의 ‘중국어 의사소통 기본표현’ 170개 문장, 중고급 수준의 중국어 실용문 200개 문장을 구글, 네이버, 百度 기계번역으로 정확률을 평가하였다. 초급 기본표현의 경우, 네이버(9.31점), 百度(9.21점), 구글(7.05점)로 구글이 최하위였는데, 구글이 중국어-영어, 영어-한국어 등 병렬코퍼스를 학습하는 과정에서 영어식 표현에 영향을 받았을 가능성을 언급하였다. 중고급 실용문의 경우, 구글(7.24점), 네이버(7.22점), 百度(6.69점) 순으로 난이도가 높아지면서 번역품질도 저하되었으며 구글이 영어 위주의 다양한 웹 문서를 학습하기 때문에 상대적으로 문어체에 강하다는 해석을 하였다. 권도경 외(2018)는 구글 번역과 네이버 파파고로 중한번역 추이를 분석한 결과, 시간의 흐름에 따라 번역 결과물이 발전하는 방향성은 발견되지 않았다고 언급하며, 데이터의 양적 축적보다는 고품질의 병렬 코퍼스의 확보가 번역 품질 개선에 더 중요하다고 보았다. 기유미(2018)는 신HSK 3~4급 수준의 한중 작문 연습문제 580개 문장을 네이버, 구글 번역에 입력하여 한중 기계번역의 오류 유형을 점수 척도로 측정하였다. 김윤정(2018)은 성어 100개, 초급 회화 문장 127개, 뉴스 기사 2편을 구글로 중한 번역한 결과를 유형별로 귀납하였다. 김혜림(2020)은 신문사설의 한중 기계번역과 인간번역의 문체 차이를 다루었으며, 기계번역은 메타담화의 변이 없이 문자 그대로 번역되는 특징을 보이는 데 반해 인간번역은 추가와 강조가 나타났다고 하였다.

이들 연구는 중-한 또는 한-중 신경망 기계에 대한 일정한 평가 항목을 제시하고 오류 유형을 분석하며 나아가 교육현장에 시사점을 제시하였다는 의미가 있다. 언어쌍의 번역 방향은 대부분 중국어에서 한국어로의 번역이 주를 이루었고, 기계번역에 대한 인간의 사전적·사후적 개입의 관

점에서는 모국어인 한국어 사후교정을 수업 설계에 활용하는 방안에 대해 언급하였다. 이에 비해 본 연구는 신문 기사나 보고서 등 기계번역이 비교적 처리를 쉽게 하는 텍스트가 아니라 기계가 가장 어려워하는 구어체를 연구 분석의 대상으로 삼고자 한다. 또한, 기존 연구에서 많이 사용되었던 구글 번역의 경우, 2022년 10월 기준으로 중국 시장에서 사실상 철수하였고 한국어-중국어 언어쌍에서는 경쟁력 있는 성능을 보여주지 않았던 점을 감안하여 분석 도구에서 제외하였다. 대신 국내 플랫폼인 네이버 파파고와 카카오톡번역을 선정하고, 구글 번역이 철수한 시점에서 중국의 가장 선도적인 기계번역 플랫폼인 百度, 腾讯, 有道를 활용하여 한-중 번역의 오류 양상을 관찰하고자 한다.

## II. 연구 대상 및 방법

### 1. 연구 대상

본고는 한국인들이 일상생활에서 실제로 사용하는 구어체를 AI번역이 어떻게 중국어로 번역하는지 파악하기 위해, 전 세계적으로 인기를 끌었던 넷플릭스 연애 리얼리티 프로그램 <솔로지옥2>의 발화 데이터를 연구 대상으로 삼았다. <솔로지옥2>의 총 10개의 에피소드 중에서, 1화 영상의 한국어/중국어 자막을 선택 활용하였으며, 자막 단위로 나누어진 총 문장 개수는 940개, 영상 분량은 약 75분이다.

다양한 영상 콘텐츠 중에서 넷플릭스(Netflix)를 선택한 이유는 넷플릭스는 전 세계 1위 온라인동영상서비스(OTT) 기업으로서의 시장 지배력과 글로벌 파급력 때문이다.

넷플릭스가 자체적으로 제작한 오리지널 작품인 <솔로지옥>은 20대, 30대 매력적인 싱글 남녀가 출연하는 리얼리티 프로그램으로, 매칭에 성공하

여 커플이 되어야만 나갈 수 있는 외딴 섬인 ‘지옥도’에서 공동생활하면서 솔직한 애정을 쌓아가는 과정을 중계한다. 연애 리얼리티는 전 세계 어디에서나 통용되는 남녀 간의 사랑 이야기와 섬세한 감정 서사를 꾸밈없이 보여주기에 이번 솔로지옥 출연진과 방송 패널들의 발화를 분석하면서 대한민국의 청춘 감성과 살아있는 언어 데이터로는 최적이라는 판단으로 본 연구의 분석 대상으로 선정하였다.

## 2. 연구 방법

넷플릭스의 공식 한국어/중국어 자막은 별도의 확장 프로그램을 이용하여 추출하였으며, 한국어 자막 텍스트를 네이버 파파고·카카오·번역·百度翻译·腾讯翻译·有道翻译 등 5개 플랫폼에서 중국어로 번역한 결과를 분석하였다. AI번역의 특성상 시기에 따라 번역 결과가 유동적으로 변화할 가능성이 있으므로 해당 번역 작업은 2023년 1월 한 달 간 진행되었음을 명시한다. 본 연구에서 한국 플랫폼 2개, 중국 플랫폼 3개 등 5개의 AI번역 플랫폼을 선택한 이유는 AI번역의 출력 결과가 해당 기계가 학습한 언어 데이터의 가용성 및 품질과 긴밀히 연관되어 있고, 입력어와 출력어 자체의 특성이 번역에 영향을 미치는 만큼, 양국의 대표적인 플랫폼을 다룰 필요가 있다고 보기 때문이다. 또한, 한 두 개의 플랫폼 분석만으로는 기계번역의 공통적인 한계점을 논하기에는 지나친 일반화가 될 수 있기에 가급적 여러 개의 AI번역을 사용하기로 하였다.

네이버파파고와 카카오·번역을 선정한 기준은 다음과 같다. ‘2022년 연령대별 인기 앱’ 통계에 따르면, 10대에서 60대까지 전 연령대에서 가장 많이 사용한 앱이 카카오톡(1위), 유튜브(2위), 네이버(3위) 순이었다.<sup>1)</sup> 이처럼 한국인의 일상을 지배한다고 해도 과언이 아닌 앱 중에서 기계번역을 지원하는 곳이 카카오와 네이버다. 특히 네이버 파파고는 10대 이하에

1) 출처: <10대 파파고, 20대 토스, 30~60대는...> 조선일보 2022.12.30.

선 9위를 차지했는데, 이는 동 서비스가 초등학생부터 중·고등학생들에게 이미 친숙한 외국어 학습 수단임을 의미한다. 기계번역의 성능은 해당 플랫폼의 유저들이 생성하는 언어 데이터를 자동적으로 학습하며 발전하기에, 사용자의 수와 사용시간이 많은 네이버 파파고는 1순위 분석 대상이었다. 실제 번역학 연구에서 네이버 파파고는 한국어와 중국어의 일상적 표현과 구어체에 강하며 번역의 정확도가 한중 언어 쌍에서 압도적으로 높음이 확인되기도 하였다(장애리 2017). 카카오번역은 비록 네이버에 비하면 후발주자로 출발하였지만 카카오톡의 범용성과 실시간으로 생성되는 풍부한 채팅 대화 자원을 보유하고기에 자연스러운 한국어 구사력 측면에서 특히 기대가 되었으며, 과일럿 테스트 결과 한국산 플랫폼답게 전반적으로 자연스러운 한국어와 때로는 네이버 파파고를 능가하는 재치 넘치는 표현을 출력하는 등 앞으로의 잠재력이 충분하다고 보였다.

중국 플랫폼의 경우, 중국 정부가 지정한 3대 AI 선도기업인 BAT (BAIDU, ALIBABA, TENCENT)에 착안했으며, 百度, 阿里巴巴(알리바바), 腾讯(텐센트) 모두 기계번역 서비스를 지원하지만, 阿里巴巴의 경우 2023년 2월 말을 기점으로 내부 업무 조정을 이유로 번역서비스를 중단한다고 밝힌 바 있다. 나머지 百度翻译과 腾讯을 분석 도구로 선정했으며, 기계번역에 전통이 있는 Netease의 자회사 YOUDAO 번역 플랫폼을 추가하여 다양성을 추구하였다.

이렇게 양국의 번역 플랫폼 5로 출력한 한중 번역은 정확성과 유창성, 자연스러움을 기준으로 하여 전문가의 주관적 평가를 실시하였으며, 번역 품질에 대한 2차 검증은 한국어를 전공한 중국어 원어민 그룹의 검수를 거쳤으며, 추가적으로 중국 온라인 커뮤니티에서 실제 사용되는 용례를 검색하여 최종 판단하였다.

### Ⅲ. 구어체 한-중 AI번역의 오류 양상

#### 1. 오류 유형의 분류 기준

본 장에서는 국내의 네이버파파고, 카카오i 번역과 중국의 百度, 有道, 騰訊 번역 등 AI 번역 모델 5종의 주요 오류 양상을 살펴보고 여러 번역에서 공통적으로 나타나는 오류에 대해 집중적으로 논의해보고자 한다. 본 연구는 번역기 간의 성능 비교에 초점을 둔 것이 아니라 현 단계 구어체 번역의 전반적인 오류 유형을 파악하는 데 목표를 두고 있다. 입력문에 대해 동일한 어휘 또는 문법 항목에서 5개의 번역기 가운데 과반수인 3개 이상의 번역기에서 오류가 나타날 경우 분석 대상으로 삼았다. <솔로지옥 2> 1화에 나오는 940개의 문장을 고찰한 결과, 3개 이상의 기계번역에서 오류가 나타나는 문장이 총 92개, 비율로 계산하면 약 10%인 것으로 드러났다.

기계 번역의 오류 유형에 대한 기존 연구에서는 오류 유형을 다음과 같이 나누었다. 서보현·김순영(2018)은 누락, 단어/절 배열, 어휘 오류, 맥락, 미완성, 오타자 등으로, 이준호(2019)는 단순 오류<sup>2)</sup>, 단어의 오역, 추가<sup>3)</sup>, 누락<sup>4)</sup>, 문법 등 다섯 가지로, 김윤정(2018)은 필수성분 누락, 과도한 유추, 발음표기, 어휘의미오용, 어휘의미나열, 고유명사오용, 관용표현미인식, 문화소오용, 구문왜곡, 논리오용, 기사형식부적응, 기타 등 12가지로 분류하고 있다. 기존 연구의 대부분은 번역의 결과물에서 드러나는 표면적인 현상에 초점을 맞추어 오류 양상을 분류하고 있다. 예를 들어, ‘누락’은 출발어에 있던 성분이 도착어로 번역되면서 없어진 것을 말하고, ‘추가’는 출

2) 맞고 틀림을 이분법적으로 확실하게 구분할 수 있는 단순 번역 실수.

3) 원문에 없던 의미 등이 번역에서 추가됨.

4) 번역문에서 단어 등이 임의로 누락되어 원문의 의미가 충분히 전달되지 않은 경우.

발어에 없던 성분이 도착어에서 추가된 경우를 뜻하며, ‘발음표기’는 출발어의 특정 어휘를 도착어에서 발음 그대로 표기한 경우를 일컫는다. 즉, 이러한 오류 유형은 모두 출발어가 도착어로 번역된 이후의 결과적 현상에 대한 논의이다.

본고는 번역 오류를 유발한 1차적인 원인을 출발어의 텍스트에 착안하여 언어 요소를 유형화하는 분류 방식을 취한다. 5개의 AI번역이 특정 번역 오류를 범하였을 때 오류의 구체적인 양상들이 AI번역마다 다르다 하더라도 오류의 근본적인 원인은 하나로 귀납되는 경우가 많다. 이에 본고는 출발어의 언어현상 파악에 중점을 두고 오류 유형을 분류할 필요가 있다고 보았다. 예를 들어 한국어 어휘 ‘배’를 A, B, C라는 3개의 AI번역이 각각 ‘肚子’, ‘船’, ‘梨’라고 제 각기 다르게 번역하게 된 근본 원인은 출발어의 어휘가 동음이의어라는 점에 있다. 그래서 출발어의 언어 현상 측면에서 오류 유형을 조명해보는 것도 의미가 있다고 보았다. 또한, 한국어와 중국어의 AI번역과 관련한 연구는 기유미(2018), 김혜림(2020) 등을 제외하면 대부분 번역 방향이 중→한이었고, 이와 더불어 한중 번역 연구에서 일상적인 구어 대화체를 소재로 하여 출발어인 한국어의 특징을 다룬 논문도 거의 없다. 번역의 오류를 제대로 분석하기 위해서는 양 언어에 대한 균형적인 이해가 필요하다. 그동안 한중 혹은 중한 번역 연구에서는 주로 중국어에 포커스가 맞추어져 있었다. 이에 본 연구에서는 출발어인 한국어 텍스트에 초점을 맞추어 오류 분석 과정에서 한국어 대화체의 고유한 특징에 주목해보고자 한다.

우선 AI번역 오류의 유형을 크게 어휘, 통사, 화용 등 3개의 층위로 나누고 AI번역에서 빈번하게 나타나는 오류를 고찰하여 출발어의 언어현상을 기준으로 각 층위 별로 유형을 세분화하기로 한다. 어휘 층위에서는 단어의 사전적 의미 혹은 지시적 의미를 중심으로 다룬다. 대부분의 단어는 여러 의미가 공존한다. 본래 지니고 있는 기본의미에서부터 의미가 파생되고 지시적 의미에서 비지시적인 의미로 확장되기도 한다. Leech(1981:23)는 단어의 의미를 지시적이고 고정적인 개념 의미와 그 외 연상적 의미

등 비개념적 의미로 구분하였지만 의미들 간의 경계가 뚜렷하지 않고 고정적이지 않다고 하였다. 이처럼 의미항 간의 경계가 모호하고 의미의 유동적 측면이 있기는 하지만 사전에 등록된 개념적 의미와 상황 맥락 속에서 얻게 되는 상황적 의미 혹은 비개념적 의미는 어느 정도 구분이 가능하다. 이에 본고는 단어의 사전적 의미, 지시적 의미를 어휘 층위에서 다루고, 상황 맥락에 의존한 화용적 의미는 화용 층위에서 다루기로 한다. 어휘, 통사, 화용 층위 내부에서도 각각 세부적으로 여러 오류 양상들이 나타났지만, 본고는 AI번역이 범하는 오류 가운데 발생 빈도가 높은 오류 유형을 위주로 논의하고자 한다. 어휘 층위에서 오류 유형은 주로 어휘의 중의성과 관용적 표현 등 2개의 영역으로 나누어 볼 수 있고, 통사 층위의 오류 유형은 성분 생략과 복잡한 구조가 주를 이루었으며, 화용 층위에서는 담화표지, 문체화법, 축약형 문장 등 3개 영역이 주된 오류 양상으로 나타났다.<sup>5)</sup>

## 2. 오류 유형별 사례 분석

### 1) 어휘 층위

어휘란 일정한 범위 안에서 사용되는 단어의 집합으로서, 어휘에는 실질적인 의미를 지니는 내용어 뿐 아니라 접사나 어미 결합형, 조사 결합형 등 기능어가 포함되며, 연어구나 관용구와 같이 단어를 넘어선 단위도 포함될 수 있다.<sup>6)</sup> 의미는 추상적인 기호체계와 실존 세계의 대상과의 관계를 바탕으로 성립된다. 일반적으로 어휘 번역은 ‘사과-蘋果’, ‘컴퓨터-電腦’, ‘뛰다-跑’, ‘먹다-吃’와 같이 출발어와 도착어의 어휘의 기본 의미가 상당 부분 일치하여 기계적으로 쉽게 매칭된다. 그렇다고 하여 어휘 번역

5) 3.2절에서 오류 유형을 설명하면서 예시로 제시하는 문장 안에 여러 개의 오류 유형이 동시에 나타나는 경우도 있을 것이다. 특정 오류 유형에 대해 설명을 하기 위한 예시에서 다른 유형의 오류는 언급하지 않을 예정이다.

6) 김은혜(2017:15-16) 인용.

이 늘 이와 같이 단순하게 처리되는 것은 아니다. 한 단어를 여러 의미로 해석할 수 있는 어휘적 중의성(ambiguity)으로 인해 번역 오류가 빈번하게 나타난다. 또한 특정 사회나 커뮤니티에서 사회 문화적으로 관습화되어 사용되는 관용어 등도 문화적 차이로 인해 번역 오류가 적지 않게 나타난다. 본 절에서는 어휘 층위에서의 중의적 표현과 문화적 특수어휘인 관용구를 중점적으로 다루기로 한다.

어휘의 다의성과 관련하여 Taylor(1995)는 한 범주의 여러 의미들이 서로 간의 유사성을 매개로 연결되어 있으며 모든 의미는 어떤 측면에서는 화용적이라고 볼 수도 있음을 주장하였다. 의미항은 고정적이지 않고 유동적이며, 의미항들 간의 경계도 모호하다. 그렇기 때문에 다의어 등을 번역할 때에도 분명 문맥이나 상황적 요소에 대한 고려는 필수이다. 어휘 사용자는 문맥에서 의미 관계를 따져보고 추론 과정을 거쳐 하나의 의미를 선택하게 된다. 하지만 그렇다고 하여 어휘적 의미와 화용적 의미를 동일시할 수는 없다. 여러 의미 가운데 구체적으로 어떤 특정 의미가 쓰였는지가 맥락에 의해 결정된다고 하여 전부 ‘화용’ 의미라고 통칭하지는 않는다. 어휘 영역의 표현에 화용적 요소가 개입된다는 것은 문맥과 상황에 부합하는 하나의 어휘적 ‘의미’로 구체화되는 것일 뿐, 그 의미 자체가 진리조건적으로 영향을 미치지 않는 ‘화용’ 층위의 용법이라거나 비지시적인 의미라는 것은 아니다. 또한 본고는 사회 문화적 요소가 다분히 개입되었다고 볼 수 있는 관용구도 어휘 영역에서 다루기로 하는데, 그 이유는 관용구가 거시적인 사회 문화적 배경에 의해 발생한 의미를 지닌다면, 화용 의미는 발화 당시의 미시적인 상황 맥락에 의해 얻는다는 점에서 관용구와는 구별되기 때문이다.<sup>7)</sup>

7) 화용 의미는 사전적 의미나 지시적 의미와는 달리 말로 풀어서 정확하게 설명하기 어려우며 상황적인 요소들이 복합적으로 작용하여 만들어진 상황 의존적 특성을 지닌다.

(1) 어휘의 중의성

본 절은 출발어의 특정 어휘가 지니고 있는 어휘적 중의성으로 인해 도착어로의 번역 과정에서 발생하는 오류를 다룬다. 광의의 어휘적 중의성은 크게 다의어와 동음이의어로 구분할 수 있다.

언어의 경제성 원칙에 따라 한 단어가 여러 세부 의미항을 지니는 것은 세계 여러 언어의 공통된 현상이다. 단어의 다의 관계까지 완벽하게 일치하는 외국어 단어를 찾기란 쉽지 않다. 이러한 미스매칭 때문에 번역 과정에서 어휘 오역이 생길 가능성이 높다. 5의 기계 번역에서 공통적으로 발생한 92개의 오류 사례를 분석한 결과, 출발어의 다의어 오류가 20개로 발생 빈도수가 가장 높았다.<sup>8)</sup> 구체적인 오류 사례는 다음과 같다.

〈표 1〉 다의어 ‘시즌’에 대한 중국어 번역

입력문		잠깐 봤는데도 지난 시즌보다 더 뜨겁고 좀 더 자극적이고 강해진 게 느껴지지 않나요?
출력문	네이버	看了一会儿, 是不是感觉比上赛季更火热、更刺激、更强了?
	카카오	看了一会儿, 不觉得比上一季更火热、更刺激、更强大了吗?
	百度	看了一会儿, 也感觉到比上一季更火热更刺激更强了吧?
	腾讯	看了一会, 你不觉得比上赛季更热, 更刺激, 更强壮吗?
	有道	看了一会儿, 不觉得比上个赛季更火热、更刺激、更强了吗?
	Netflix 자막	开场片段很简短, 但看起来是不是比上一季更辣、更刺激?

‘시즌’은 영어 ‘season’에서 온 외래어로서, 《고려대 한국어대사전》에 따르면 크게 두 가지 의미로 사용되고 있다. 첫째, ‘어떤 활동이 가장 활

8) 본고의 원문 텍스트에서 다의어가 동음이의어에 비해 개수 자체가 많아 번역 오류의 발생 빈도도 높게 나타났다.

발히 이루어지는 시기', 둘째, '인기 있는 텔레비전 프로그램을 계속적으로 제작하여 방송할 때 제목의 이름으로 덧붙여 쓰는 말'이다. 위의 입력문은 〈솔로지옥 시즌2〉 패널의 오프닝 멘트로 후자의 의미로 사용된 것이다. 이에 적합한 중국어 번역문은 '季'이다. 그런데 해당 어휘가 포함된 문장을 5개의 번역기에 입력해보니, 네이버 파파고, 腾讯, 有道 등 3개 번역기에서 '시즌'을 첫째 의미에 해당하는 '赛季'로 번역하였다. 문장 하나만 놓고 본다면 '赛季'로 번역한 문장도 자체적으로는 문법이나 의미적으로 아무 문제가 없다. 하지만 상기 입력문에 대한 출력문으로서는 완전히 잘못된 문장이라고 할 수 있다. 이 문장이 사용된 상황과 전후 맥락에 부합하지 않기 때문이다.

이외에 한국어 어휘 '몸'을 '身材'로 번역해야 하는 상황에서 5개의 AI번역이 모두 '身体'로 잘못 번역한 사례도 있고, '회전목마가 돌아가다'의 '돌아가다'를 '回去'로 번역하여 '旋转木马回去了'와 같은 황당한 번역이 나오기도 하였다. 어휘의 다의성에서 유발된 오류는 어느 정도 맥락과 연관이 있다. 두 개 이상의 의미 또는 용법을 지니는 단어는 그것이 쓰인 상황적 맥락에 의해 그 중의성이 해소되기 때문이다. 중의성을 지닌 단어가 텍스트에서 구체적으로 어떤 의미로 쓰였는지는 발화자의 의도나 상황적 맥락만 잘 파악하고 있다면 충분히 알아낼 수 있다. AI번역 과정에서 다의어가 엉뚱한 의미로 오역이 되었다는 것은 기계가 화자의 의도나 상황적 맥락을 제대로 파악하지 못하였다고 볼 수 있다.

다음으로 출발어에 동음이의어가 존재하는 관계로 도착어로의 번역 과정에서 해당 어휘와 음만 동일한 완전히 다른 단어로 오역된 경우도 있다. 한국어 '배'를 중국어로 옮길 때 '船', '肚子', '梨' 중 어떤 것으로 번역할 것인가 하는 문제가 있다.<sup>9)</sup> 음운의 조합으로 만들 수 있는 음절 조합이 한정적인 만큼 동음이의어의 발생은 필연적이다. 본 연구에서는 동음이의

9) 엄격하게 말하면 신체 부위로서의 '배'와 교통수단으로서의 '배', 과일 종류인 '배' 간에는 장단음 등의 음성적 구별이 있으나 그 차이가 크지 않기 때문에 일반적으로 이들을 동음이의어로 간주한다.

어 오류가 2건 발견되었다.<sup>10)</sup> 다음은 그 중 한 예이다.

〈표 2〉 동음이의어 ‘무’에 대한 중국어 번역

입력문		1인1무	
출력문	기계번역	네이버 파파고	一人一亩
		카카오i	一人一份
		百度	一人一亩
		腾讯	一人一亩
		有道	一人一平
	Netflix 자막	一人一根萝卜	

‘무’는 동음이의어가 다수 존재하는 단어이다. 텍스트에서는 채소를 뜻하는 단어로 사용된 것인데 AI 번역에서는 이와 발음이 같은 ‘이랑 무(亩)’로 번역하였다. 위의 입력문은 〈솔로지옥 시즌2〉에서 먹는 무가 산더미처럼 쌓여있는 것을 보고 출연진 중 한 명이 인당 무 한 개씩 먹을 수 있을 만큼 무의 양이 많음을 ‘1인1X’의 형식을 사용하여 재치 있게 표현한 것이다. 해당 어휘가 포함된 문장을 5개의 번역기에 입력해보니, 네이버 파파고, 腾讯, 有道 등 3개 기계번역에서 ‘무’를 첫 번째 의미에 해당하는 ‘이랑 무(亩)’로 번역하였다. 이렇듯 화자가 표현을 참신하고 기발하게 할수록 기계번역은 그 언어활동의 다양성이나 변화 속도를 따라가지 못하는 것으로 판단된다.

10) 번역문이 단순히 하나의 원인 때문에 오류가 발생하였다고 보지는 않는다. 예를 들어, ‘1인1무’의 번역 오류도 화자의 독특한 표현법이 영향을 주어 기계번역이 동음이의어 가운데 선택 오류가 생겼을 수 있다. 하지만 ‘무’가 동음이의어이기 때문에 여러 ‘무’ 가운데 하나를 선택해야 하는 상황을 만든 것은 분명한 사실이다. 즉, ‘무’의 번역에 오류가 생긴 가장 근본적인 원인은 해당 단어가 동음이의어이기 때문이라고 본다.

(2) 문화적 특수어휘

본 절은 사회 문화적으로 관습화된 의미로 사용되는 관용어나 신조어 등이 도착어로의 번역 과정에서 액면 그대로 번역하여 오역이 발생한 경우가 다수 발견되었고, 또 의미가 어느 정도 파악이 되어도 중국어에는 그에 완벽하게 부합하는 어휘 자체가 존재하지 않아 번역상의 정확도와 자연스러움이 떨어지는 경우도 있었다. 본 연구에서 확인된 문화적 특수 어휘 오류는 14개가 있었다. 구체적인 예는 다음과 같다.

〈표 3〉 관용어 ‘면접 프리패스상’의 번역 오류

입력문		되게 이분 면접 프리패스상 아니에요?	
출력문	기	네이버 파파고	这位不是面试合格相吗?
	계	카카오i	这不是面试自由通行证奖吗?
	번	百度	这位不是面试自由通行证吗?
	역	腾讯	真的是这位面试自由通行证奖吗?
		有道	你不觉得他很容易就能通过面试吗?
		Netflix 자막	他看起来能在所有工作面试中都得到好成绩

중국에는 ‘면접 프리패스상’에 대응하는 적절한 어휘가 존재하지 않는다. 〈표 3〉에서 보면 카카오i, 百度, 腾讯 등 3개의 기계번역이 입력어를 제대로 출력하지 못하고 있다. ‘면접 프리패스상’에 딱 들어맞는 어휘가 없으므로 네이버 파파고의 경우 자체적으로 ‘面试合格相’이라는 표현을 고안해내었다. 하지만 중국인들은 해당 표현이 이해는 가능하지만 중국에서는 쓰지 않는 어색한 표현이라고 하였다. 사회가 변화하면서 새로운 언어 표현이 양산되는 것은 자연스러운 현상이다. 취업난이 심각한 근래 한국에서는 ‘취준생’, ‘공시족’, ‘일루’, ‘소확행’ 등 많은 줄임말 형태의 신조어가 생겨나고 있다. 본 연구에서는 ‘답이 오다’, ‘대시하다’, ‘작업멘트’, ‘산으로 가다’, ‘무쌍’, ‘대학 커뮤니티’, ‘승부 보다’, ‘돌쇠’ 등 우리나라 특유의 관용적 표현이 AI번역에서 높은 오류율을 보였다.

정수현(2018)에서는 134개의 한국문화 특수어휘를 고찰하고 있는데 이런 어휘를 ‘번역할 수 없는 말’이라고 칭하고 있다. 그만큼 어휘에 대한 사회 문화적 맥락을 모르고 사전적 의미만 가지고는 단순하게 번역을 할 수 없다는 뜻이다. 한국 사람들 사이에서는 흔하게 쓰이는 말이지만 그에 대응하는 외국어 어휘나 문장을 찾지 못하여 난감할 때가 많다. 언어는 한 나라의 문화를 반영하기도 하는데 그 속에는 사람들의 생각, 관습, 사고방식과 의식구조 등 문화적 요소가 내포되어 있다. 이러한 문화적 배경 지식이 없으면 표면적으로 드러나는 언어 표현만 가지고는 정확한 의미를 파악할 수 없다. 출발어 의미에 대한 파악조차 되지 않으면 도착어로의 정확한 번역은 당연히 기대하기 어렵다.

## 2) 통사 층위

### (1) 성분 생략

출발어의 통사 구조적인 문제로 인하여 도착어로 번역하는 과정에서 잘 못된 도착어로 출력되는 양상이 나타나기도 하였고, 92개의 오류 사례를 분석한 결과 출발어의 구조적인 문제로 인해 발생하는 오류 사례가 20개로 오류 발생 수가 세 번째로 많았다. 통사 구조적인 측면에서 주로 두 가지의 오류 유형이 보이는데 그중 첫 번째가 바로 출발어에서 주어, 목적어, 관형어 등의 문장 성분이 생략된 경우이다.

목정수(2010)에 따르면 우리말 구어는 논항이 실현되지 않거나 논항이 실현되더라도 조사가 없는 경우가 문어에 비해 훨씬 많다. ‘논항-술어’의 관계 중 ‘주어-서술어’ 관계나 ‘목적어-서술어’ 관계로 포착되는 명시적 명사 논항과 더불어 문맥이나 어미를 통해 복구할 수 있는 비명시적 논항도 있다. 노대규(1996)에 의하면 구어는 음성으로 실현된 언어이며 즉흥적, 직접적, 즉각적인 특성을 지니는 데 비해 문어는 문자를 매개로 실현되며 사유적, 간접적, 계획적인 특성을 지닌다.

이렇게 구어와 문어는 기본 성격이 다른데 언어의 문법이라고 하면 일반적으로 문어를 바탕으로 형성된 것을 말한다. 사람들이 평소 일상적으로

사용하는 대화체의 구어는 역동적이고 간결함을 추구하는 구어의 특성상 주어나 목적어 등 문장 성분의 생략이 많고 격 표지나 조사 등의 사용도 최소화하는 경향이 있다. 이는 문어를 바탕으로 이루어진 일반 문법의 기준에서는 비문으로 처리되기도 하는 양상들이다. 하지만 이런 식의 과감한 생략이 대화 속에서는 오히려 훨씬 자연스러울 때가 많다.

〈술로지옥〉도 젊은 남녀들이 일상에서 주고받는 살아있는 대화체이다 보니 문장 내 성분이 생략된 형태가 다수 보였다. 목정수(2010)는 구어의 문법 체계는 기존 문어 중심의 문법체계와는 다르게 파악할 필요가 있다고 주장하며 구어의 규칙과 질서를 순수 문법적인 차원에서 규명하고자 시도하였다. 본고 역시 대화체의 성분 생략 현상을 언어 사용자의 단순 실수나 비문으로 간주할 것이 아니라 이를 대화체 특유의 특성으로 이해하고 비중 있게 다룰 필요가 있다고 본다. 다음 예시를 보자.

〈표 4〉 주어 생략으로 인한 번역 오류

입력문		( )빠졌나? 잠깐만. 어, ( )빠졌나? ( )빠진 거 같기도 하고.
출 력 문	기계번역	네이버 파파고 掉了嗎? 好像又瘦了。 好像掉色了。
		카카오i 啊, 掉了嗎? 好像掉了。 好像掉了。
		百度 嗯掉了嗎? 好像掉进去了好像掉进去了
		腾讯 啊, 掉了嗎? 好像沒了。 好像沒了。
		有道 他掉进去了嗎? 好像漏掉了。 好像漏掉了。
		Netflix 자막 出来了嗎? 等等。 出来了嗎? 感觉好像出来了

〈표 4〉의 입력문은 참가자가 풀밭을 건다가 밭에 가시가 박혔다가 빠진 상황에서 나온 발화인데, 구체적으로 ‘무엇이’ 빠졌는지, 주어가 누락되어 있다. 넷플릭스 자막은 생략된 성분에 대해 파악하여 원문의 의미를 무난하게 번역하였으나, 기계번역은 모두 오역을 하고 있다. 흥미로운 점은 네이버 파파고는 ‘빠지다’의 대상 파악이 되지 않자, 대상을 ‘물건’, ‘살(또는

체중), ‘색’ 순으로 임의로 바뀌가며 번역하고 있다. 사실 해당 입력문이 사용된 정황을 보면 앞의 맥락을 통해 ‘빠진’ 대상이 ‘가시’라는 사실을 쉽게 파악할 수 있다.

- ① 세정: 아, 따가워.
- 용재: 괜찮으세요?
- 세정: 가시 박힌 거 같은데. 아, 가시 박혔다.
- ... (중략)
- 용재: 괜찮으시면 제가 해 드릴까요?
- 세정: 아, 네.
- 세정: 빠졌나? 잠깐만. 어, 빠졌나? 빠진 거 같기도 하고.

위와 같이 직전 대화에서 ‘가시’가 박힌 상황임을 알려줬음에도 불구하고 기계는 그 앞의 텍스트까지 연결시키지 못하였고 동사 ‘빠지다’가 직접적으로 사용된 문장에만 국한하여 그 동작의 주어를 매칭해보다가 실패한 것으로 보인다.

다음은 문장 내 주어와 목적어가 생략된 예시이다.

〈표 5〉 주어와 목적어 생략으로 인한 번역 오류

입력문		제가 ( ) 좋다고 하면 ( ) 다 ( ) 좋다고 하던데요	
출력문	기	네이버 파파고	只要我说好, 大家都说好。
	계	카카오i	我说好就都说好。
	번	百度	我说喜欢就都喜欢啊?
	역	腾讯	听说我好的话都好呢?
		有道	我说喜欢就说喜欢
		Netflix 자막	当我说我喜欢他们时,他们都说他们也喜欢我

〈표 5〉의 입력문은 여자 출연진 본인이 ‘누구를(목적어)’ 좋다고 하면

‘상대방(주어)’도 ‘자신을(목적어)’ 좋다고 한다는 인터뷰 내용이다. 이때 출발어에서 주어와 목적어가 생략되어 표현되었는데, 5개의 AI 모델은 전부 번역이 불완전하거나 아예 잘못되었다. 중국 원어민 감수자는 위와 같은 배경 상황이 주어지더라도 중국어의 경우 한국어처럼 주요 문장 성분을 생략하면 어색하다고 답하였다.

다음은 문장 내 관형어가 생략된 예시이다.

〈표 6〉 관형어 생략으로 인한 번역 오류

입력문		아, 미소가 너무 예쁘시네	
출력문	기계번역	네이버 파파고	啊, 微笑太漂亮了。
		카카오i	啊, 笑得真漂亮。
		百度	啊, <u>你的</u> 微笑真漂亮。
		腾讯	啊, <u>你的</u> 微笑真漂亮。
		有道	啊, <u>你的</u> 微笑好美啊。
	Netflix 자막	<u>他的</u> 笑容很好看	

〈표 6〉의 입력문은 한 패널이 남자 출연진의 인터뷰 영상을 보면서 미소가 아름답다고 칭찬을 하는 상황이다. 패널은 ‘누구의’ 미소인지, 문법적으로는 ‘미소’를 수식하는 관형어 성분을 생략하고 표현하였다. 맥락상 ‘他的’가 들어가야 올바른 번역이 되는데 검토한 5개의 기계번역 가운데 중국 AI번역 3개는 모두 ‘你的’라는 관형어를 임의로 추가하였다. 이는 명백한 오역이다. 넷플릭스 자막팀은 상황적 맥락을 파악하고 ‘他的’를 추가하여 올바르게 잘 번역하였고, 네이버 파파고와 카카오i는 입력문 구조 그대로인 관형어가 누락된 상태로 보수적으로 번역하여 명백한 오역은 피할 수 있었다.

다음은 문장 내 서술어가 생략된 예시이다.

〈표 7〉 서술어 생략으로 인한 번역 오류

입력문		와, 소름( ).	
출력문	기계번역	네이버 파파고	哇, 鸡皮疙瘩
		카카오	哇, 好诡异
		百度	哇, 鸡皮疙瘩
		腾讯	哇, 鸡皮疙瘩
		有道	哇, 起鸡皮疙瘩
	Netflix 자막	我起鸡皮疙瘩了	

위에서 ‘소름’이라고 한 것은 놀랍거나 소름 끼치는 상황에서 ‘소름이 돋는다’, 즉 몸에서 일어나는 반응을 표현하는 방식으로 본인의 감정을 드러낸 것이다. 최근 젊은 사람들의 대화체에서 자주 발견되는 언어 현상은 바로 줄임말이다. 인터넷 채팅 문화가 발달하며 줄임말을 사용하는 것은 전 세계적인 추세라고도 할 수 있다. 젊은 사람들 사이에서 줄임말 신조어가 많이 탄생함으로써 세대 간에 의사소통이 어려워지는 모습도 보이고 있다. 5개 기계번역 중 네이버, 百度, 腾讯은 ‘소름’을 한국식 그대로 ‘鸡皮疙瘩’로 번역하였는데 중국 원어민들의 감수 결과, 주어진 상황에서 ‘鸡皮疙瘩’는 어색하고 동사 ‘起’를 추가하여야 자연스러워지는 것으로 드러났다. 중국에서도 젊은 세대 중심으로 줄임말을 많이 사용하고는 있지만 한국어만큼 줄이는 정도가 심하지는 않는 편이다. 이러한 차이는 한국어의 구어가 본래부터 생략형 언어였다는 점을 감안하면 언어 자체의 특성에 기인한 것으로 볼 수 있다.

(2) 복잡한 통사구조

AI번역이 입력문의 복잡한 통사구조를 잘 인식하지 못하여 오역이 발생한 경우도 있었다. 다음은 수식어 내부 구조가 복잡하고 길어 기계가 문장 구조를 잘못 인식하고 오류를 범한 사례이다. 이런 오류는 총 5개로 통사 오류 유형 가운데 2위이다. 아래 예문을 보자.

〈표 8〉 복잡한 통사 구조로 인한 번역 오류

입력문		저는 제가 좀 진지하게 확실하게 좋아해서 만난 사람이 없었어서	
출력문	기계번역	네이버 파파고	因为我比较认真地、确实地喜欢, 因为没有见过的人,
		카카오	我是真的很喜欢, 所以没有见过任何人,
		百度	我是因为我非常认真的喜欢, 所以没有遇到过这样的人,
		腾讯	我比较认真的喜欢, 所以没有见过的人,
		有道	我自己比较喜欢认真切实, 所以我没有遇到的人,
	Netflix 자막	我从来没跟我真正喜欢的人约过会, 也没有达到确认喜欢某人的程度	

위의 문장은 남자 출연진이 본인이 이 연애 프로그램에 출연하게 된 이유를 설명하는 장면에서 나온 말이다. 이 문장은 ‘제가 좀 진지하게 확실하게 좋아해서 만난’이 관형어로서 뒤에 나오는 명사 ‘사람’을 수식하고 있다. 관형어의 내부적인 구조를 들여다보면, 기본적으로 연동 구조로 이루어져있기도 하고 두 개의 부사어가 동시에 하나의 서술어를 수식하고 있는 형태라 상당히 복잡한 구조이다. 이에 모든 기계 번역은 관형어 내부의 연동 구조를 분리시켜 번역하는 등 문장 구조 파악에 실패하였다.

### 3) 화용 층위

담화 맥락 속에서 얻게 되는 화용적 의미를 제대로 파악하지 못함으로써 오류가 발생하기도 한다. ‘지시적 의미’(denotation)란 텍스트를 구성하는 단어의 글자 그대로의 의미를 뜻하는데, 단어를 축자적으로 번역하여도 정확하게 맞아떨어지는 경우도 물론 있지만 앞뒤 맥락이나 발화 상황에 대한 배경 정보 등을 동원해야만 정확한 번역이 이루어지는 경우가 많다. 김은혜(2007:76-77)에 따르면 사람들은 단어를 사전에 등록된 의미에만 국한하여 사용하지 않고 사용자의 배경 지식, 사전 경험, 문화적 지식, 발

화 맥락에 맞게 의미를 조정하여 사용한다고 한다. 또한, 단어는 그 자체가 지닌 의미와는 별개로 거시적인 맥락에서 특정한 상황과 결부하여 사용되기도 하는데 이런 사용 문맥에 의한 의미는 언어 사용자들이 역동적으로 의미를 재구성하는 과정에서 발생하는 것이기 때문에 맥락 의존적이고 경우에 따라서는 맥락을 벗어나면 그 의미가 사라질 수도 있음을 지적하고 있다. 본 절은 의미적으로나 통사적으로 설명하기 어려운 화용 층위의 표현에 대해 살펴보기로 한다. 세부적으로는 특정 어휘가 담화표지로 기능하는 사례, 화자의 심적 태도나 개인적 성향이 내포된 화법의 문제, 축약형 문장 등 세 항목으로 나누어 논의하고자 한다.

AI 기계번역은 입력된 단어를 따로 식별하는 방식이 아니라 입력된 전체 내용을 컴퓨터가 학습을 통해 가장 적합한 언어로 조합하여 출력하는 방식이다. 그만큼 신경망 기계번역은 기존의 규칙 기반이나 통계 기반의 기계번역보다 맥락을 종합적으로 판단한다는 점에서 보다 정확한 번역이 가능해졌다. 그럼에도 불구하고 본고의 고찰 결과, 지시적 의미 외에 함축 의미(connotative meaning)를 나타내거나 화자의 태도, 감정, 상태 등 화용론적 요소에 대한 이해가 많이 요구되는 문장에서 적지 않은 오류가 발견되었다. 5개의 기계번역에서 공통적으로 발생한 92개의 오류 사례 가운데 총 18개 사례가 이러한 화용 의미의 인지 오류와 관련이 있었으며, 모든 오류 유형 가운데 발생 빈도수는 3위이다.

### (1) 담화표지(Discourse Marker)

담화표지란 주로 화자의 담화 전략에서 기인하며 특별한 어휘적 의미가 없어 문장의 진리조건적 의미에는 영향을 미치지 않는 화용적 층위에서 기능을 하는 표현을 일컫는다. 담화표지는 대개 특정 어휘가 기본의미에서 문법화 과정을 거쳐 화용론적 문맥에서 화용상의 기능을 수행하게 된 것이다. 담화표지가 주로 구어체에서 많이 나타나는 만큼, 본 연구에서 상당히 자주 출현하고 있음을 확인할 수 있다. 우리는 평상시에 대화를 할 때 무의식적으로 담화표지를 자주 사용한다. 담화표지란 기존의 명시적 의미

에서 벗어나 담화 속에서 새로운 화용적 기능을 수행하는 언어 표현을 뜻한다. 의사소통 과정에서 담화표지를 적당히 잘 쓰면 담화가 자연스럽게 연결되고 의사소통의 효율을 높일 수 있다.

본고가 일상생활 속에서 사람들이 실제로 주고받은 대화체를 연구대상으로 삼고 있는 만큼, 텍스트에 담화표지가 많이 출현하고 있다. 담화표지는 사전적 의미를 나타내는 것이 아니기 때문에 표현만 놓고 기계적으로 단순하게 매칭하면 오역이 되기 쉽다. <솔로지옥2> 패널들의 대화에서는 담화표지 ‘아니’가 종종 발견된다. 한국어에서 ‘아니’는 본래 부정(否定)을 나타내는 표현으로서 구어에서 주로 부사나 감탄사로 사용되는데, 이와 동시에 다양한 담화 기능도 가진다. 담화표지로서의 ‘아니’에 대한 연구로 김미숙(1997), 이한규(2012) 등이 있다. 김미숙(1997)은 담화표지 ‘아니’는 수용불가의 태도를 나타내거나 주의를 집중시키거나 실수를 교정하는 기능이 있다고 하였고, 이한규(2012)는 담화표지 ‘아니’에 맞장구치기, 시간 끌기, 자기 방어, 화제 전환, 다른 의견 제시, 관심 끌기, 말차례 가져오기, 자기 수정 등의 기능이 있다고 주장한다. 아래 예를 보자.

<표 9> 담화표지 ‘아니’에 대한 중국어 번역

입력문		아니, 회전목마가 돌아가더라고.	
출 력 문	기 계 번 역	네이버 파파고	不是 旋转木马在转
		카카오i	不是是旋转木马
		百度	丕, 旋转木马在转。
		腾讯	丕, 旋转木马回去了。
		有道	旋转木马在转
	Netflix 자막	有个旋转木马	

‘아니 회전목마가 돌아가더라고’라는 발화는 한 패널이 이전 시즌보다 천국도의 부대시설이 화려해지고 좋아졌다고 하자 다른 패널이 회전목마도 있다고 덧붙인 것이다. 즉, 앞서 패널이 언급한 내용에 다른 패널이 맞

장구를 치면서 해당 내용에 부연 설명을 하는 상황이다. 이때 ‘아니’는 상대방의 말을 부정하는 의미가 아니라 상대방의 말에 덧붙여서 부연 설명하는 담화기능을 지닌다고 볼 수 있다.

중국어에는 ‘아니’의 대응 표현으로 ‘不是’가 있다. ‘不是’도 사전적 의미 외에 화용 기능도 어느 정도 존재한다. 马永娜(2018)은 반문어기부사 ‘不是’가 문법화되어 화용 의미가 생겨났다고 하나 그 문법화 정도가 아직 미숙한 단계라고 주장한다. 马永娜(2018)의 주장에 따르면 ‘不是’가 발언권 획득, 주의 환기, 보충 설명 등 화용 기능을 하는 상황이더라도 담화 참여자 간에 기본적으로 의견 대립이나 모순이 분명 존재하며 이는 ‘不是’의 기본 의미인 부정 의미가 살아있기 때문인 것으로 보았다. 실제로 <솔로지옥2> 담화 텍스트에 출현한 ‘아니’에 대한 번역 결과만 놓고 보더라도, ‘아니’에 비해 ‘不是’의 문법화 정도가 미숙함을 알 수 있다. <표 4>에서 ‘아니’를 대부분의 기계번역은 ‘不是’, ‘不’로 번역하고 있는데 본 연구의 어감 테스트에 참여한 중국인들은 해당 번역문을 모두 어색하다고 보았다.

구중남(2015:76)은 발화된 문장을 제대로 이해하려면 담화표지에 대한 이해가 필수적인데, 담화표지는 어휘적 의미가 없거나 모호하여 화자와 청자를 고려해야 하고 그것의 기능이 복합적이고 문맥에 따라 달라지기 때문에 그 기능을 명확하게 파악하기 위해서는 면밀한 주의가 필요하다고 지적한다. 그만큼 담화표지에 대한 번역은 난이도가 아주 높은 언어 요소 중의 하나이다. 번역자의 능력 부족 혹은 도착어 안에 출발어에 부합하는 적절한 표현이 존재하지 않는 등 여러 이유로 담화표지를 정확하게 번역하기가 어려운 상황이라면, 어휘적 의미로 번역하여 문장을 어색하게 만드는 것보다는 차라리 담화표지를 번역하지 않는 보수적인 방법을 취하여 진리조건적 의미라도 보존하는 방법을 쓰는 것을 제안한다. 하지만 이런 번역을 결코 완벽하고 좋은 번역이라고 부를 수는 없을 것이다.

(2) 화법

본고는 화자가 발화 상황이나 문맥에 따라 전략적으로 특정 언어 형식을 선택하여 사용하는 언어 변종을 ‘화법’으로 칭한다. 같은 상황에 놓였더라도 사람마다 자신이 취하는 전략이 다르고 또 화자 자신의 개성과 성격에 따라 화법이 결정된다. 그러다보니 개인의 주관적이고 개성 있는 말투까지 AI번역이 충실하게 담아내기란 현실적으로 쉽지 않다. 다음 예시들을 보자.

〈표 10〉 ‘아이’에 대한 중국어 번역

입력문		그래도 이 아이는 제가 책임감 있게 한번 해내 보고	
출력문	기계번역	네이버 과과고	这孩子我还是负责地做一次吧。
		카카오i	这个孩子我会负责地做一次。
		百度	等一下我会负责地完成这件事
		腾讯	这个孩子还是让我负责地做一次吧。
		有道	这个孩子我也要负责地做一次。
	Netflix 자막	我应该负起责任有始有终	

한국어에서 화자의 개인적 성격이나 담화 전략 등에 따라 어떤 상황에서 특정 사물을 애정 어린 감정으로 ‘아이’로 표현하는 경우가 있다. ‘아이’는 기본적으로 사랑스럽고 보호해주고 싶은 이미지를 연상시키며, 이러한 표현을 사용하는 화자에게도 귀엽고 부드러운 이미지를 주는 효과를 가져다준다. 위의 발화문은 여성 출연자인 서은이 식재료를 다듬는 과정에서 자기가 심혈을 기울여 다듬고 있던 ‘무’를 ‘아이’로 본인의 감정과 태도를 주입하여 의인화한 것이다. 그런데 AI번역 5개 모두가 ‘아이’를 그대로 ‘孩子’로 직역하고 있다. 중국어 ‘孩子’는 한국어 ‘아이’와 같이 의인화되어 쓰이는 용례가 드물다. 이는 어휘를 사용한 화자의 발화 의도나 발화 당시의 상황적 맥락, 문화적 배경 지식에 대한 이해 없이 단순히 어휘적 측면

에 국한하여 접근한 결과라고 볼 수 있다.

다음 예시는 여성 출연진이 전화를 받을 때 애교 섞인 말투로 ‘여보세용?’이라고 한 것을 AI번역 5개 중 3개가 오역하였다.

〈표 11〉 ‘여보세용’에 대한 중국어 번역

		입력문	여보세용?
출 력 문	기 계 번 역	네이버 파파고	喂 ?
		카카오i	喂 ?
		百度	亲爱的 ?
		腾讯	亲爱的世用?
		有道	老公世勇?
	Netfix 자막	喂 ?	

개인적인 말투 자체가 애교가 많은 사람일 수도 있고 담화 전략적으로 애교 섞인 귀여운 말투로 말을 하는 경우 한국에서는 어미에 ‘-오’를 붙여서 말한다. 네이버와 카카오는 이러한 뉘앙스까지는 전달은 못했지만 전화를 받을 때 응대하는 명제진리는 잘 전달을 하였다. 그에 반해 중국 AI번역은 모두 어미에 붙는 ‘-오’의 기능 파악에 실패하여 기본적인 ‘여보세용?’로 번역하거나 ‘여보’와 ‘세용’을 분리하여 번역하기도 하였다. 엄밀히 말해 ‘여보세용?’은 사전에 정식으로 등록된 표준어가 아니며, 개인 취향이나 말하는 습관에 따라 붙이는 표현이기 때문에 이 또한 해당 발화자에 대한 이해가 있어야 정확하고 자연스러운 번역이 나올 수 있다.

(3) 내용 축약

본고의 연구 대상은 기본적으로 대화체인데다가 줄임말을 좋아하는 젊은 세대들의 대화이다 보니 내용을 압축하거나 과감하게 생략하는 문장들이 자주 등장한다. 이러한 종류의 발화는 이해하는 데 있어 맥락 의존도가 상당히 높은 경우에 해당한다. 본 절에서 말하는 ‘축약’은 앞서 3.2.2절 통

사 층위에서 다루었던 ‘성분 생략’과는 구분된다. ‘축약’은 여러 개의 언어 요소들을 한 데 압축하여 나타내는 형태이고, ‘생략’은 주어, 목적어, 관형어 등 본래 있어야 할 위치의 특정 문장 성분이 누락된 경우다. 이하의 축약형 예시를 보자.

- ② 서은: 근데 아까 보셨어요? 음식 하는 데?  
 나딘: 무 30개?  
 서은: [웃으며] 그 와중에 30개였어요?  
 나딘: 그 정도 될걸요?

②는 부엌에 엄청난 양의 무가 쌓여 있는 것에 대해 서은과 나딘이 나눈 대화 내용이다. 서은이가 ‘그 와중에 30개였어요?’라고 한 것은 ‘그 짧은 시간 안에 (그 많은 무를 세서 무가) 30개였(다는 것을 캐치했)어요?’라는 의미로 말한 것이다. 이러한 사례는 단순히 통사 층위에서 특정 문장 성분이 생략되었다고 보기에는 무리가 있다고 본다. ‘그 와중에’는 기간을 나타내는 부사어로서 그 ‘기간’ 안에 ‘어떤 동작을 하다’라는 서술어와 상응하는데, 구체적으로는 ‘무의 개수를 세었다’ 또는 ‘무가 많다는 사실을 포착하였다’ 등의 서술어와 호응해야 자연스럽다. 하지만 ‘그 와중에’와 ‘30개였다’는 논리적으로 연결이 어색하다. ‘그와중에 30개였다’는 곧 여러 문장이 복합적으로 혼재되어 하나로 표현된 축약형 문장이라고 볼 수 있다. 형식과 논리가 어긋난 이 문장을 제대로 파악하고 번역하려면 우선 짧은 시간 안에 무의 개수를 파악하는 게 무리일 정도로 무가 많았다는 사실에 공감대가 형성이 되어 있어야 한다. 그러나 기계는 이러한 대화 참여자들의 인식과 맥락까지 정밀하게 파악하기는 힘들었을 것으로 본다. 네이버 파파고를 제외한 나머지 4개의 기계번역 모두 완전히 오역을 하였고, 네이버 파파고 역시 어감을 제대로 살려서 번역하지는 못하였다.

〈표 12〉 내용 축약으로 인한 번역 오류

입력문		그 와중에 30개였어요?	
출력문	기	네이버	是30个吗 ?
	계	카카오	中间是30个吗?
	번	百度	那时候是30个吗 ?
	역	腾讯	其中30个吗 ?
	기	有道	其中有30个?
		Netflix 자막	有30个 ?

이상에서 한국어 구어체를 중국어로 번역할 때, 기계번역 플랫폼 5종에서 공통적으로 나타나는 주요 오류 유형을 살펴보았다. 오류 유형을 표로 정리하면 아래와 같다.

〈표 13〉 한-중 기계번역 오류 양상

오류 유형	세부 유형	개수	항목별 비율
어휘	중의성	22	24%
	특수 어휘	14	15%
	기타(고유 명사 등)	5	5%
	소계	41	45%
통사	성분 생략	20	22%
	복잡한 구조	5	5%
	기타(화제문 등)	3	3%
	소계	28	30%
화용	담화 표지	8	9%
	화법	6	7%
	축약형	4	4%
	소계	18	20%
기타	특수 외래어	5	5%
합계		92	100%

#### IV. 결론

이상으로 본 연구에서는 구어체 대화문으로 이루어진 리얼리티 프로그램 <솔로지옥>의 출연진 발화 텍스트를 연구 대상으로 삼아 네이버 파파고, 카카오i, 百度翻译, 腾讯翻译, 有道翻译 등 기계번역의 구어체 오류 유형을 검토해보았다. 본 연구에서는 번역 결과의 오류 여부를 판단한 후에 유형을 귀납하는 방식을 취하였다. <솔로지옥2> 1화의 940개의 문장을 5개의 기계번역으로 출력해보았는데, 그중 3개 이상의 기계번역에서 오류를 범한 문장이 총 92개였다.

본 연구는 번역 오류가 유발된 원인을 출발어 텍스트와 출발어 특성에서 파악하여 오류 유형을 크게 어휘 층위, 통사 층위, 화용 층위로 나누어 논의하였다. 이상 세 가지 영역을 비교해보면 어휘, 통사, 화용 순으로 오류 빈도수가 높은 것으로 드러났다. 어휘 층위에서는 주로 다의어와 특수 어휘에서 오류가 빈발하였고, 다의어로 인한 오류 비율이 상대적으로 높았다. 통사 층위에서는 통사 성분의 생략으로 인한 오류와 통사 구조의 복잡함으로 야기된 오류가 주를 이루었는데, 이 둘 중에 통사 성분의 생략으로 인한 오류가 압도적으로 많았다. 화용 층위에서는 담화표지, 화법, 내용 축약 순으로 오류가 빈번하게 발생하였다.

이처럼 AI 기계번역의 한국어 구어체 번역 오류를 층위별로 분석하였는데 개인 유튜브를 비롯한 미디어 콘텐츠가 폭발적으로 성장하는 이 시대에 고품질의 구어체 번역에 대한 수요는 갈수록 높아질 것이다. 본 연구의 구어체 오류 양상을 바탕으로 하여 차기 연구에서는 한국인 사용자들이 구어체 한국어를 중국어로 번역할 때, 기계번역 플랫폼들을 보다 능동적이고 효율적으로 활용할 수 있도록 모국어 텍스트를 사전 교정하는 일련의 규칙을 연구해볼 수 있을 것이다.

## &lt; 참고문헌 &gt;

- 강병규·이지은, <신경망 기계번역의 작동 원리와 번역의 정확률>, 《중어 중문학》 제73집, 2018.
- 강수정, <통번역대학원생들의 기계번역 사용현황과 인식에 관한 연구> 《번역학연구》 22(3), 2021.
- 고려대학교 민족문화연구원, 《고려대 한국어대사전》, 고려대학교 출판부.
- 구종남, 《국어의 담화표지》, 경진, 2015.
- 권도경, 문수정, 김정수, 허윤정, <2018년 인공지능망 기계번역의 중한 번역 현황 - 구글 번역과 네이버 파파고를 중심으로>, 《중국문학》 98권, 2019.
- 기유미, <한중 기계번역 오류의 문형별 비교분석: 네이버 파파고 번역기와 구글 번역기를 중심으로>, 《중국연구》 74권, 2018.
- 김미숙, <대화구조로 본 ‘아니’의 기능>, 《담화와 인지》 4권, 22호, 1997.
- 김윤정, <‘구글 번역’ 기반 텍스트 유형별 중한 기계번역 오류 현황 분석>, 《중국어문학지》 65집, 2018.
- 김은혜, 《외국어로서의 한국어 어휘 교육론》, 서울: 역락, 2017.
- 김혜림, <기계번역과 인간번역의 문체 비교: 신문사설 한중번역을 중심으로>. 《통번역학연구》 24(1), 2020.
- 노대규, 《한국어의 입말과 글말》, 국학자료원, 1996.
- 목정수, <구어 한국어를 위한 문법 모형>, 《한국어학》 46, 2010.
- 서보현·김순영, <기계번역 결과물의 오류유형 고찰>, 《번역학연구》 19(1), 2018.
- 이상빈, <기계번역에 관한 KCI 연구논문 리뷰: 인문학 저널 논문(2011~2020년 초)의 논의내용과 연구방법을 중심으로>, 《통역과 번역》 22(2), 2020.
- 이준호, <신경망기계번역의 객관적 평가를 위한 예비연구: 자동평가와 수

- 동평가의 균형점), 《통번역학연구》 23(3), 2019.
- 이한규, <한국어 담화표지어 ‘아니’의 의미>. 《현대문법연구》 67, 2012.
- 장애리, <국내 기계 통번역의 발전 현황 분석: 한중언어쌍을 중심으로>, 《번역학연구》 18(2), 2017.
- 정수현, 《한국문화 특수어휘집》, 서울: J&J Culture, 2018.
- 최문선, <국내 번역학 기계번역 연구 동향: 내용 분석과 키워드 분석>을 중심으로, 《언어학연구》 24(1), 2019.
- 马永娜, <“不是”作为话语标记的语用功能分析>, 《课程教育研究·学法教法研究》 第26期, 2018.
- Leech, G.N., *Semantics*, Penguin Books, 2018.
- Taylor, J., *Linguistic Categorization*, Oxford University Press, 1995. 조명원·나익주 옮김, 《인지언어학이란 무엇인가-언어학과 원형이론》, 한국문화사, 1997.
- Vieira et al. *Machine translation in society: insights from UK users*, Language Resources and Evaluation, 2022
- <10대 파파고, 20대 토스, 30~60대는...> 조선일보 2022.12.30.

### <Abstract>

In this study, natural colloquialisms that young people in their 20s and 30s actually use in daily life in Korea were studied, and five machine translation platforms, such as Naver Papago and Kakaoi in Korea and ‘百度’, ‘腾讯’ and ‘有道’ in China, were used as subjects of study. This study examined the error aspects of colloquial machine translation between Korea and China. This study does not focus on the evaluation and comparison of machine translation platform performance,

but focuses more on the inherent characteristics of Korean spoken language, looking at how AI machine translation handles subtle situational nuances or linguistic and cultural contexts, and what they have in common. Efforts were made to reveal the error patterns. While previous studies on Korean–Chinese machine translation mainly dealt with written language such as newspaper articles, editorials, speeches, and reports, and even if they dealt with colloquial language, they focused on Chinese conversation or refined dialogues appearing in HSK textbooks. It is differentiated in that machine translation was studied by taking live dialogues actually used in daily life by young Koreans in their 20s and 30s who appeared in season 2 of the Netflix original reality program 〈Single's Inferno〉.

Key Words : 한-중 기계번역(Korean–Chinese machine translation), 오류 양상(error pattern), 구어체(spoken language), 문맥 인식(contextual awareness), 문법(grammar)