

## 국내 언론과 KCI 학술 논문에 나타난 '기계번역' 담화 분석

천중성(건국대학교)

이준호(중앙대학교)

### 1. 서론

기계번역은 과거에는 학계가 독점하는 연구 주제 영역이었다면, 이제는 공적인 담론의 주제가 되었다. 기계번역의 빠른 발전, 기계번역의 인간번역 대체 혹은 번역사 일자리의 소멸 등 부정적 전망 혹은 우려를 누구나 일상에서 한 번쯤은 접했을 것이기 때문이다. 이러한 변화 속에서 번역학계에서 기계번역은 최근 몇 년간 급부상한 주제일 수밖에 없었다. 최문선(2019)에 따르면 신경망번역이 출시된 2016년 이후 기계번역 연구의 수가 국내 번역학 학술지에서 급증했다. 하지만 기계번역에 대한 학술적 논의가 촉발된 이유는 기계번역의 발전 혹은 장점에 기인한 것이 아니라 기계번역에 대한 두려움 혹은 경계심 때문이라 보는 것이 합리적일 것이다. 기계번역에 관한 KCI 논문을 분석한 이상빈(2020)의 지적처럼 다수의 연구가 기계번역의 정확성과 오류에 집중해왔음이 이러한 가정을 방증한다.

하지만 기계번역이라는 새로운 기술을 더욱 적극적으로 활용 혹은 기각하기 위해서는 번역학계는 기계번역과 관련된 현실을 그 누구보다 명확하게 진단하고, 일반 대중과 소통하며, 번역학계가 지향해야 할 연구 방향과

활동을 규명해야 한다. 그리고 그 출발점은 기계번역과 관련된 현실을 일반 대중이 어떻게 접하는지 이해하는 것이다.

현실이란 사람들이 살아가면서 자신 주변의 객체와 상황을 인식한 결과물이다. 따라서 현실은 그 자체로 존재한다기보다는, 특정 가치체계 및 관점에 따라 구성된다고 할 수 있다. 대표적인 예시가 새로운 지식이나 사건에 대하여 뉴스가 현실을 구성하는 것이다. 이 과정에서 뉴스는 객관적 현실을 거울처럼 있는 그대로 반영하는 것이 아니라 단편적 사실관계를 텍스트로 조직하여 독자들이 인지할 수 있는 현실로 만들어 낸다. 그리고 이러한 뉴스의 현실 구성 과정을 심층적으로 분석하는 방법의 예시로 프레임(frame) 이론을 들 수 있다. 프레임 이론은 1970년대 비판적 관점에서 뉴스의 이데올로기적 성격을 규명하고 뉴스 전달의 구조적 결정요인을 밝히기 위해 제기되었다(이준웅, 2000, p. 94).

프레임 이론에 따르면 뉴스의 형성 과정에 있어 복잡한 사실관계 중 특정 측면을 선택하고 강조하여, 텍스트에서 그 부분이 더 명확하고 다른 부분보다 더 두드러지게 만들어 내는 작업이 존재한다. 그 결과 뉴스에서 생성된 프레임은 특정 사안에 대한 정의 및 인과관계의 해석, 도덕적 평가를 할 수 있도록 수용자를 인도한다(Entman, 1993, p. 52). 즉, 같은 사안에 대해서도 뉴스 텍스트의 관점과 논리에 따라 특정 사안이 가지는 함의가 달라지고 이에 대한 해석과 의견도 달라질 수 있다는 것이다.

이에 본고는 기계번역과 관련된 뉴스를 분석하여 어떻게 기계번역에 대한 현실이 뉴스를 통해 구성되었는지 조명하며 다음의 연구 질문을 논의한다. 첫째, 최신 기계번역 기술로 평가받는 신경망 기계번역의 출시 시점인 2017년부터 2022년까지의 뉴스를 분석하여 기계번역과 관련된 현실을 구성하는 주요 요소를 파악하고 해당 요소에 대해 뉴스 구성자의 중립적, 부정적, 혹은 긍정적 해석이 어느 정도 개입했는지를 분석한다. 둘째, 이러한 현실에 대응하여 번역학계에서 출간한 논문에서 주요 기계번역 담론은 무엇이었으며, 뉴스가 제시하는 현실과 어떠한 차이가 있는지 알아본다. 셋째, 상기 논의에 기반하여 번역학계의 미래 소통 과제를 제안하고자 한다.

## 2. 선행연구

### 2.1 기계번역에 대한 담론

번역학계 내에서 기계번역에 대한 담론을 직간접적으로 다룬 연구는 다수 존재한다. 하지만 기계번역의 담론에 대한 논의가 번역학계라는 단일 장(場) 내에 국한된 경우가 다수이다. 가장 대표적인 예시 중의 하나가 기계번역에 대한 번역 전문가 및 전공자의 인식에 대한 연구이다(강수정, 2020, 2021; 천중성, 2020; Cadwell et al., 2017; Guerberof, 2013). 또 다른 예시로는 기계번역을 중심으로 이뤄진 연구의 주제 분석과 이를 통해 번역학계의 과제를 제시한 메타 연구, 번역학계 외부에서 이뤄진 특정 담론에 대한 번역학계의 대응을 제안한 연구, 그리고 번역학의 관점에서 바라본 기계번역에 대한 언론 보도의 담론 등으로 구분할 수 있지만, 연구의 수가 많지는 않다.

먼저 번역학계 내에서 기계번역이라는 주제가 어떻게 연구되고 담론이 형성되는지 간접적으로 살펴본 연구로는 다수의 논문을 통시적으로 분석한 메타 연구가 있다(장애리, 2020; 전현주, 2023; Christensen et al., 2017). 본고의 논의 범주인 국내 기계번역 연구에 대해서는 최문선(2019)은 국내 번역학 기계번역 연구 동향을 내용과 키워드 중심으로 분석하였다. 신지선(2020)은 기계번역과 포스트에디팅에 관한 해외 연구 동향을 제시하며 국내에서 자주 논의되고 부족한 연구 영역이 무엇인지 명확하게 제시하고 있다. 여기에서 한 걸음 더 나아가 이상빈(2020)은 KCI 인문학 저널에 나타난 기계번역 연구의 내용과 연구방법을 분석하고, 번역학계의 연구 공백을 지적하고 방법론적 개선을 제안하였다. 상기 세 연구는 메타 연구의 특징 덕분에 전반적 연구 주제 흐름과 기계번역에 대한 학계의 담론 파악에는 도움이 된다. 하지만 연구방법의 특징 때문에 연구자 혹은 번역학계가 기계번역에 대해서 취하는 태도가 명확하게 드러나지는 않으며, 번역학계 외부의 장에서 이뤄지는 기계번역 담론을 충분히 다루기에는 구조적으로 어려움이 있다.

다음으로 기계번역의 등장에 따른 인간번역의 역할 혹은 인간번역사 대체 담론에 대응한 연구가 있다. 박은주와 김동미(2022)는 인간과 기계번역의 공존성을 행위자 네트워크 이론을 사용하여 논의했다. 이들은 기계번역이

또 다른 번역의 행위자로 등장하면서, 번역의 전통적 행위자인 인간의 역할이 전복되지 않기 위한 전제 조건이 무엇인지 고심해야 하며, 번역학계가 행위자 전복을 방지하기 위한 노력을 해야 한다고 지적하였다. 송연석(2018)은 언론에서 제기된 번역사가 필요 없게 된다는 주장에 대응하는 것이 번역학자의 역할임을 강조하며, 번역사라는 직업 성격의 보완 및 강화를 주장하였다. 구체적으로는 번역교육의 방향 설정, 번역평가 및 감수를 통해 잘된 번역 혹은 수용 가능한 번역인지에 대한 정의가 필요함을 제안하였다. 상기 두 가지 연구는 기계번역의 부상에 대응하는 인간번역사의 역할에 대해 훌륭한 논의를 펼쳤다. 하지만 연구의 출발점이라 할 수 있는 인간 역할의 전복 혹은 인간 대체 담론이 어떻게 구성되었고 실재하는지에 대한 논의는 부족하여 아쉬움이 남는다. 따라서 언론 보도 분석을 통해 상기 언급된 담론이 진정으로 우세 담론인지 그리고 많은 대중에게 영향을 주었는지 분석할 수 있다면, 추후 이뤄질 번역학계 연구에서 활용 가능한 논의의 시발점 역할을 할 수 있을 것이다.

세 번째로 기계번역의 언론 보도 프레임에 본격적으로 다룬 연구로는 Vieira(2020)와 천종성과 강수정(2022)의 연구가 있다. Vieira(2020)는 총 284개의 기사 내용에 대해 다수의 양자 선택 질문을 제기하고 그 답변을 정량화하여 기사를 긍정, 부정, 중립의 프레임으로 구분하였다. 상세 내용 분석 결과, 언론 보도는 기계번역의 긍정적인 측면에 초점을 두고 있었다. 또한, 기계번역 제공사의 의견 주장만을 검증 없이 보도한 경우가 많았다고 비판하였다. 천종성과 강수정(2022)은 기계번역이란 주제어를 통해 언론에서 어떠한 담론을 구성하고 있는지 심층적으로 분석하고 있으며, 본고와 가장 유사성이 높은 연구이다. 2017년부터 2022년까지 언론의 헤드라인 데이터를 빅카인즈에서 수집하고, 프레임 이론에 기반하여 뉴스의 의미 네트워크 분석을 시도하였다. 그 결과 신경망번역과 관련된 뉴스가 특정 프레임을 사용했다기보다는 대부분 가치 중립적으로 접근하고 있다고 평가하였다. 하지만 상기 연구는 정교한 분석에도 불구하고 헤드라인만을 분석 대상으로 삼았다는 점은 연구의 한계라 할 수 있다. 따라서 언론 기사 본문에 대한 보다 심층적 연구가 이뤄진다면 연구의 신뢰도가 한층 높아질 것이며, 언론에서 초점을 두고 부각한 내용과 특정 프레임이 존재한다면 이를 식별하는 데 도

움이 될 것이다.

이에 본 연구는 기존의 연구에서 관점을 넓혀 번역학계 외부에서 발생하고 있는 기계번역 담론이 어떻게 기계번역과 관련된 현실을 구성하고 있는지 언론 보도 분석을 시도한다. 또한, 번역학계에서 막연히 받아들이던 부정적 담론이 언론 보도에서 주요한 담론인지를 명확하게 파악할 것이다. 여기에 더해 기존 연구와 달리 기계번역에 대한 학계와 언론의 담론적 격차를 식별하고, 번역학 연구자들이 새롭게 대중과 소통하는 방안을 제시할 것이다.

## 2.2 뉴스분석에 대한 선행 연구

구글과 네이버가 인공지능경망 기반 기계번역 서비스를 출시한 사실과 그와 관련된 담론을 대개의 일반 수용자들은 뉴스를 통해서 접하였다. 이 뉴스들은 기계번역과 관련된 사실적 관계들을 텍스트로 연결하여 의미화된 담론의 형태로 이용자들에게 전달한다. 즉, 기계번역과 관련된 뉴스는 현실을 텍스트로 구성한 의미의 집합체로, 프레임 이론은 뉴스화 과정에서 의미화되는 해석적 집합체를 설명하는 것을 목적으로 한다.

인공지능경망 기반의 새로운 기계번역이 출시되었다는 현실을 뉴스 텍스트로 구성하는 과정에서, 미디어는 이용자들이 그 이슈를 해석하고 평가하는 방법에 영향을 미치는 단어를 선택하여 그 의미를 강조함으로써 독자나 이용자들에게 기계번역 이슈에 대한 해석의 틀을 제시한다. 그리고 독자들은 그 프레임을 통해서 기계번역 뉴스를 수용하게 된다. 이에 대해 이준웅(2005)은 동일한 사건이라도 뉴스 텍스트에서 부각된 단어의 연결에 따라 그 함의가 달라지며, 특히 갈등적 이슈에 대한 여론의 변화가 프레임에 대한 수용자의 해석적 차이에서 비롯된다고 설명한다.

연구자들은 인공지능경망 혹은 기계번역과 같은 과학기술 이슈의 확산과정 역시 뉴스 프레임의 형성과 해석되는 과정을 거친다고 설명한다. 이용자들은 새롭게 부상하는 기술발전 과정에서 소외에 대한 두려움이 있는데 그 교육의 도구로 대중매체를 선택하고, 기술 관계자들도 이용자들의 이해 부족이 기술 운영에 좋지 않은 영향을 끼칠 것에 대비하여 대중매체를 교육 도구화한다. 미디어는 과학기술 보도라는 고품질 정보가 관련 주체들의 관

심을 끌 수 있는 인기 품목으로 간주하고 뉴스를 처리한다(권상희, 2005, p. 37). 이 과정에서 뉴스는 새로운 기술을 긍정적으로 다루어 기술변화와 보급을 촉진하는 가속자의 역할을 할 수도 있고, 부정적 측면을 부각시켜 기술변화를 더디게 하는 제한자의 역할을 할 수도 있다(유승관과 이제영, 2006, p. 174). 결국, 기계번역 기술 역시 어떤 표상이나 단어적 매개를 통해 뉴스 프레임의 구성하는지에 따라 미디어의 기계번역 기술과 서비스에 대한 가속과 제한의 해석적 함의를 유추할 수 있게 된다.

프레임의 분석에 있어 Entman(1991)은 뉴스가 이슈의 특정한 측면을 선택(select)하여 이슈의 도덕적 해법과 처방을 현저(salient)하게 만든다고 설명한다. 미디어와 종사자가 지니는 철학과 이념이 의식적·무의식적으로 뉴스에 반영되는데 그 텍스트는 단어의 선택과 반복을 통해 특정한 방향의 뉴스 메시지를 구성한다. 기계번역과 관련된 뉴스 역시 인공지능과 번역에 대한 수많은 사건과 지식들 사이에 어떠한 측면의 단어나 어휘를 선택하거나 부각하여 기계번역에 대한 미디어의 해석을 현저하게 드러낼 수 있다.

새로운 기술의 확산과정에서 선택성과 현저성을 기반으로 구성된 뉴스 텍스트는 시간에 따라 그 양상이 달라질 수 있다. 기술 확산이 시간 흐름에 따른 정규분포의 모습을 띠고 있어, 기술에 대한 개인적 수용의 크기 역시 시간의 경과에 따라 차이가 나기 때문이다(Rogers et, al., 2014). 기술을 보도하는 뉴스 텍스트에도 기술의 속성을 설명하거나 사실을 구성하는 중심 단어의 패턴이 시간에 따라 달라질 수 있는 것이다. 다시 말해 기계번역의 이용자들이 시간 흐름에 따라 사용 경험이 누적되면서 기계번역과의 상호작용하는 방식이 변화하는데, 그 형태는 미디어에서 제시되는 중심 단어의 변화로 유추할 수 있다.

프레임은 그 형식적 특성에 따라 이슈에 대한 인식이 달라진다는 연구 결과도 있다(Iyengar, 1990). 프레임을 형식적 특성에 따라 주제적 프레임(thematic frame)과 일화적 프레임(episodic frame)으로 구분하는 것이다. 일화적 프레임은 사건의 원인, 영향, 대처에 대한 규명 없이 사건의 발생을 사실 중심으로 기술하는 피상적인 단순보도 프레임이다. 반면 주제적 프레임은 통계적 근거, 전문가 증언 등을 통해서 이슈의 원인이나 배경을 이해하는데 필요한 정보를 제공하며 보다 구조적 관점에서 접근하는 형식의 프레임

이다. 기계번역 뉴스를 전체 사회적 맥락에서 주제화하여 뉴스를 보게되면 그 원인과 대책을 사회적 시스템과 연계하여 고려하게 되고, 일화적 사건 특성에 초점을 맞추게 되면 이슈를 개인화하여 그 사용성에 초점을 맞추게 될 것이다.

뉴스 분석은 뉴스의 담론을 거시적 권력과 경제적 관계에 기반하여 해석할 수 있고, 특정한 텍스트를 미시적으로 분석하여 사회적 상호작용의 의미를 해석할 수도 있다. 본 연구는 프레임이라는 미시적 분석에 기초하여 기계번역 뉴스가 이 문제를 정의하는 방식, 그리고 인과적 해석과 평가의 방법을 파악한다. 또한, 선행연구에 기반하여 뉴스 텍스트에 드러나는 키워드를 통한 현저성, 시간에 따른 프레임의 변화, 프레임의 형식적 특성을 탐구한다.

### 3. 연구 설계

뉴스를 활용한 연구를 진행하기 위해서는 신뢰할 수 있는 데이터의 원천 확보와 연구 목적과 부합하는 데이터 수집 범위를 설정하는 작업이 우선되어야 한다. 또한, 해당 데이터를 분석하기 위한 적절한 도구의 사용과 분석 방법 설정은 연구 신뢰도 향상에 필수적이다.

#### 3.1 데이터 수집

본 연구는 한국언론재단에서 제공하는 뉴스 정보 사이트인 빅카인즈를 사용하였다. 빅카인즈는 뉴스에 대한 빅데이터 분석을 위한 웹 포털이며, 비정형 텍스트를 분석이 가능한 정형화된 데이터로 바꾸어 사회현상을 분석할 수 있는 기초 자료를 제공한다. 데이터의 신뢰성 역시 높아 빅카인즈를 활용한 언론 지형 분석 및 경향성 분석 연구가 다수 이뤄진 바 있다(최수이, 2022; 하동엽, 2023).

연구 목적에 부합하는 데이터 분석 범위의 한정을 위해 대한민국에서 가장 포괄적으로 사용되고 있는 신경망번역 서비스를 제공하는 구글과 네이버의 서비스 출시 시점을 고려하였다. 두 서비스 모두 2016년 하반기에

출시가 이뤄졌고, 본 연구의 데이터 수집은 2023년에 3월에 이뤄졌다. 따라서 2017년부터 2022년까지 전국일간지, 경제일간지, 지역일간지, 방송사, 전문지 등 빅카인즈에서 수집이 가능한 모든 매체에서 ‘기계번역’을 키워드로 검색하여 데이터를 수집하였다. 그 결과 총 569건의 기사에서 325,048 단어 규모의 텍스트를 추출하여 분석을 진행했으며 그 상세 내역은 다음과 같다.

**표 1**  
연도별 기사 건수와 총 단어 수

구분	기사 건수	단어 수
2017	141	74,073
2018	77	39,024
2019	91	62,036
2020	96	44,289
2021	86	48,014
2022	78	57,612
계	569	325,048

본 연구의 또 다른 연구 질문은 기계번역과 관련된 번역학계의 담론이 기사를 통해 나타난 담론과 차이가 있는지를 분석하고, 향후 대응 방안을 모색하는 것이다. 따라서 상기 기사 코퍼스와 비교를 위해 동일한 시간 조건인 2017년부터 2022년까지 기계번역을 키워드로 작성된 논문을 추출하였다. 분석 범위는 KCI에 등재된 다섯 종의 번역학 학술지(번역학연구, 통번역교육연구, 통번역학연구, 통역과 번역, T&I Review)로 한정하였다. 총 82개 국문 논문의 본문을 코퍼스에 포함했으며 총 348,502 단어로 기사 코퍼스와 분량에 있어 큰 차이가 나지 않는다.

### 3.2 실험 설계 및 실행

분석과정에서 코퍼스 분석 도구인 워드스미스(WordSmith) 8.0을 사용하였다. 가장 많이 사용된 기능은 워드리스트를 생성하여 특정 단어의 빈도를 정량화하는 기능, 특정 키워드 주변에 있는 단어를 찾아서 목록화하는 컨코



던스 기능이였다. 여기에 더해 특정 주제나 어휘에 대한 텍스트의 접근 방식 분석에 필요한 경우 빅데이터 도구인 텍스트톰(textom.co.kr)을 사용하여 감성 분석을 진행하였다. 감성 분석이란 감성어 사전과 대상 텍스트를 비교하여 호감, 흥미, 기쁨, 거부, 슬픔 등의 분류에 해당하는 어휘의 비중을 정량화하는 기법으로서 마케팅 연구 혹은 인식 분석 연구 등에 자주 사용된다.

상기 도구를 사용하여 미시적 관찰과 정량적 분석을 통해 가설을 설정 및 초기 검증하고, 관찰 사항을 지지할 수 있는 상세 단어 및 예시를 제시하는 뉴스패키지 분석법을 사용하였다. 먼저 뉴스패키지 분석에서는 가장 많이 사용하는 키워드를 파악하고 그 키워드가 시간의 흐름에 따라 어떻게 달라지는지를 파악하여, 기계번역이 확산되는 과정에서 이슈의 속성이 변화하는 과정을 관찰하였다. 이어서 Vieira(2020)가 사용한 긍정, 부정, 중립의 프레임을 그대로 차용하여 언론 보도의 논조를 긍정, 부정, 중립으로 분류하였다. 또한, 논리 진행의 방법이 합리적 혹은 비합리적인지를 분석했으며, 문장의 서술이 사실, 주장, 주관적 판단, 기대, 전망, 감정인가 여부로 분류하여 특정한 형식적 프레임 형성을 위한 시도가 있었는지 살펴보았다. 마지막으로, 다수의 양자 택일 질문을 활용하여 프레임을 파악하는 기법을 활용하여 언론 보도와 논문에서 기계번역에 대한 담론의 차이가 있는지 거시적 관찰을 도출하고, 해당 관찰 사항을 지지할 수 있는 미시 분석 사례를 제시하였다.

뉴스의 프레임 분석은 연역적 혹은 귀납적으로 실행할 수 있다. 다만 연역적 방법은 기계번역과 같이 새롭게 등장하는 이슈에는 부족한 선행연구로 인해서 적절하지 않을 수 있다. 이에 본 연구에서는 긍·부정의 기준은 Vieira(2020)의 선행연구에 기반한 연역적 방법을 취하고, 기계번역이 뉴스에서 다루어지는 다양한 경향을 살펴보기 위해서 사례를 통한 귀납적 방법을 통해서 프레임을 확인하였다.

## 4. 분석 결과

4장에서는 본 연구의 분석 결과 가장 유의미한 분석 사례를 4.1과 4.2에서 소개한다.

### 4.1 기계번역에 대한 언론 보도 프레임

본 연구는 기계번역에 관한 뉴스 이벤트 구성에 있어 언론이 어떤 의제와 속성을 가지고 보도했는가에 대한 해석적 접근을 목적으로 한다. 뉴스는 이슈를 표현하는 과정에서 현실을 구성하는 방향성에 따라 강조하는 단어인 키워드를 부각시키고 독자들이 뉴스를 해석하도록 영향을 준다. 따라서 이번 절에서는 기사의 총량과 어휘의 총량을 통한 키워드 분석 및 감성 분석을 통해 기계번역 보도에 대한 프레임을 파악한다.

#### 4.1.1 키워드 분석

네이버 파파고의 신경망번역은 2016년 8월 그리고 구글의 신경망번역은 2016년 11월 한국어 서비스를 제공하였다. 새로운 서비스가 출시되면서 이를 홍보하기 위한 기업의 미디어 마케팅 활동이 적극적으로 이뤄졌고, 해당 활동을 언론에서 별다른 필터 없이 그대로 수용했다면 기계번역을 보도하는 기사의 건수가 다른 해 대비 양적으로 많을 것이며, 이로 인해 언론 보도의 수용자는 자연스럽게 기계번역과 관련된 기사에 많이 노출되었을 것으로 추론할 수 있다.

하지만 2016년 기계번역에 관한 총 기사 건수는 118건이었으며, 2017년 오히려 그 수가 141건으로 증가했다. 따라서 신규 서비스 출시 시점에 집중적인 언론 홍보 활동이 이뤄졌다고 주장하기에는 어려움이 있으며, 언론이 기계번역 서비스에 대한 홍보 프레임을 취했다고 주장하기도 어렵다. 다만 2017년은 최근 7년간 기계번역에 대한 언론 보도의 정점이라 볼 수 있고 그 이후 기계번역에 대한 언급은 꾸준히 유지되었다.

**표 2**  
기계번역 관련 기사의 건수와 단어 수

구분	기사 건수	전체 단어
2017	141	74,073
2018	77	39,024
2019	91	62,036
2020	96	44,289
2021	86	49,014
2022	78	57,612

다음으로 실제 대중들이 기계번역과 관련된 정보에 얼마나 노출되었는지 파악할 필요가 있다. 이를 위해 ‘기계번역’과 유사한 의미를 지닌 표현을 식별하는 작업을 먼저 진행하였다. ‘기계번역’은 번역이 자동으로 이뤄진다는 의미를 강조하여 ‘자동번역’이라고도 불린다. 또한, 2016년 새로운 기계번역 서비스가 출시되었을 때는 기존의 기계번역과 다르다는 점을 강조하기 위해 ‘신경망번역’ 혹은 ‘NMT(neural machine translation)’라는 용어를 사용하였다. 이를 고려하여 상기 어휘의 출현 빈도를 연도별로 정량화하였다.

**표 3**  
기계번역, 자동번역, 신경망번역, NMT의 출현 빈도

구분	전체 단어	기계번역	자동번역	신경망	NMT	합계
2017	74,073	266	163	185	164	778
2018	39,024	181	97	111	59	448
2019	62,036	201	131	75	25	432
2020	44,289	194	99	84	6	383
2021	49,014	145	114	82	13	354
2022	57,612	122	117	40	14	293

2017년은 가장 많은 기계번역 관련 기사가 게재된 시기이며 기계번역을 지칭하는 키워드 노출빈도가 가장 높은 시기였다. 또한, 신경망번역이라는

새로운 번역기술에 집중 조명이 이뤄진 시기라 할 수 있다.

<표 3>에 나타난 키워드 중에서 ‘기계번역’과 ‘자동번역’은 새롭게 등장한 기계번역 서비스를 지칭하는 용어라고 할 수 있다. 즉, 구글번역, 파파고와 같은 구체적 서비스를 하나로 묶어주는 용어이다. 반면에서 ‘신경망’과 ‘NMT’는 서비스의 구체적 속성, 기술적 특성을 반영하고 있다. 상기 용어는 기계번역 서비스의 초기 2017년과 2018년에는 서비스 명칭인 ‘자동번역’과 유사한 빈도로 뉴스에 등장했지만, 이후 ‘NMT’는 뉴스에서 거의 사라지는 수준이다. 또한, ‘신경망’ 역시 그 사용 빈도가 줄어들었다.

본 분석 결과는 시간이 지나면서 이용자들에게 ‘기계번역’ 혹은 ‘자동번역’은 번역서비스의 중요한 키워드가 되었고, ‘신경망’과 ‘NMT’와 같은 기술을 설명하는 용어가 뉴스에서 설명이 필요 없는 키워드가 되었음을 보여준다. 그 원인으로서는 짧은 기간에 기계번역 서비스의 사용성이 높아지면서 이용자들에게 새로운 기계번역 서비스가 널리 수용되었거나, 기술에 대해서 이용자들이 이해한다고 뉴스 제공자가 판단했기 때문으로 보인다.

#### 4.1.2 기계번역 뉴스의 감성 분석

조금 더 상세한 분석을 위해 ‘기계번역’, ‘자동번역’, ‘신경망’, ‘NMT’를 둘러싸고 있는 어휘가 지니는 속성을 알아볼 필요가 있다. 이를 위해 언론 보도 코퍼스에서 상기 키워드의 앞과 뒤에 있는 단어 10개를 추출하는 컨코던스 분석을 시행하였다. 이후 컨코던스 결과물을 별도의 파일로 만들어서 빅데이터 분석 도구인 텍스트를 사용하여 감성 분석을 진행하였다.

전체 기사의 ‘기계번역’에 대한 감성 분석에서는 긍정 79.98%와 부정 20.02%, ‘자동번역’에 대해서는 긍정 84.53%와 부정 15.47%, ‘신경망’에 대해서는 긍정 79.87%와 부정 20.13%로 긍정이 압도적으로 높은 비율을 보였다. 다만 ‘NMT’에 대해서는 긍정 69.51% 부정 30.49%로 긍정이 다소 감소한 비율을 보였지만, 긍정의 감성어휘가 전반적으로 더 많이 사용되었음에는 변화가 없다.

**표 4**  
전체 토큰 대비 신경망 및 NMT 토큰의 비중

구분	기계번역	자동번역	신경망	NMT
호감	55.18	66.75	57.56	55.47
흥미	21.07	15.23	19.77	12.41
기쁨	3.21	1.27	1.74	0.73
거부	16.07	12.9	15.41	29.93
슬픔	2.14	2.03	2.6	1.46

긍정과 부정을 구성한 상세 감성은 <표 4>에서 확인할 수 있듯, 호감과 흥미가 가장 큰 비중을 차지했다. 여기에 더해 상기 키큐던스 결과물 내에서 가장 자주 사용된 평가어는 ‘자연스럽다’ 115회, ‘정확하다’ 107회, ‘빠르다’ 66회 등으로 긍정적인 어휘가 다수 사용되었음이 관찰되었다.

<예시 1>

지난 2016년 구글 신경망 기계번역의 기술을 도입했다. 이 기술은 줄임말, 야만정음도 **정확하게** 번역해주는 기능이다. (A 경제지)  
번역 앱의 발전 속도가 무척 **빠르다**. 문장 전체를 통째로 해석하는 <중략> 기술을 적용하면서 **정확도가** 한 차원 높아졌다. (S 경제지)  
인공신경망 방식은 규칙기반 방식에 비해 번역이 **자연스럽다**. (D 일보)  
(이하 모든 예시 강조표시는 저자가 함)

또한 ‘품질’과 ‘성능’을 설명하고 있는 문장 110건을 추출하여, 언론 보도가 기계번역의 품질과 성능에 대해 취한 입장을 분석하였다. 그 결과 중립 및 긍정의 프레임이 주를 이뤘다. 상세하게는 중립적이고 합리적인 사실(43건), 긍정적 사실로부터 합리적으로 도출한 평가(26건)가 주를 이뤘다.

<예시 2>

기계번역의 성능을 지속적으로 개선했고, 번역서비스의 품질을 측정할 수 있는 체계를 마련해 품질 고도화를 위해 노력하고 있다(T 전문지, 중립적이고 합리적인 사실)  
신중희 과파고 리더는 “자체 개발한 품질 평가모델의 우수성을 국제대

회에서 인정받았다” (J 전문지, 긍정적 사실로부터 합리적으로 도출한 평가)

따라서 기계번역에 대한 언론의 보도가 가치 중립적인 접근으로 사회적 수용의 범위를 합의하고 조정하는 과정에 있다는 천종성과 강수정(2022)의 주장을 기사 본문 분석이 일부 지지하고 있다. 하지만 <예시 3>에서 볼 수 있듯 긍정적 평가를 하기 위해 합리적 근거를 제시하기보다는 비합리적이고 주관적인 판단이나 전망에 의존한 경우가 9건 있었다. 그리고 <예시 4>의 사례처럼 본문에는 해당 내용이 전혀 없지만, 제목의 작성 과정에서 편집적 개입이 이뤄진 기사를 일부 발견할 수 있었다.

<예시 3>

파과고 이용자 증가는 통·번역 품질이 개선되면서 나온 결과다. (A 경제지, 긍정적 주관적 평가이지만 비합리적)  
이 솔루션을 이용하면 <중략> 불거진 자막 품질 논란도 해결할 수 있을 전망이다. (B 경제지, 긍정적 전망이지만 비합리적)

<예시 4>

품질 좋아졌어도 ‘번역은 인간의 영역’...문학 취약한 AI (M 경제지, 제목의 편집적 개입)  
한영중일 4개 국어 동시소통 시대 온다 (D 과학지, 제목 편집적 개입)  
구글 넘은 파과고 이제 동시통역 가야죠. (S 경제지, 제목 편집적 개입)

#### 4.1.3 기계번역 뉴스의 프레임

본 장에서 추가적으로 분석한 내용은 2장에서 언급된 기계번역에 대한 인간번역사의 역할 변화 그리고 송연석(2018)이 언급한 ‘기계번역에 의해 인간번역이 사라진다’ 혹은 ‘기계번역에 의해 인간번역이 대체될 수 있다’는 주장이 언론 보도의 프레임에 의해 시작되었는가 하는 점이다. 이를 판단하기 위해 ‘대체’라는 단어가 들어간 모든 문장을 추출하여 상세 분석하였다. 총 111개의 문장 중 번역과 직접적인 관련이 없는 경우가 90회였으며, 인공지능에 의해 여타 직업군이 대체될 것이라는 논의가 주류를 이뤘다. 실제로 번역과 관련된 경우는 총 20회였으며 ‘사실+주장’이 9회, ‘주관적 판단’이 11회로 나타났다. 논조에서는 인간번역 대체 불가를 주장하는 내용이 16회

였고, 인간번역 대체 가능성을 주장하는 내용은 불과 4회에 불과했으며, 합리적 논리 전개라 보기도 어려웠다.

<예시 5>

-대체 불가 논조

AI라고 해서 바로 전문 번역가들을 대체할 수 있는 것은 아니다. (J 전문지, 주관적 판단)

AI 번역이 사람을 대체할만 한가. “그렇지 않다” (H 일보, 사실+주장)

-대체 가능 논조

지난해 11월 개선된 구글 번역은 주요 언어들 간 전문 통번역사의 기능을 대체할 수 있는 수준이다. (H 일보, 주관적 판단)

기존에 영어를 잘하는 직원들이 떠났던 단순 번역 업무를 AI가 대체하며 직원들은 실질적 커뮤니케이션에만 집중할 수 있게 됐다. (M 경제지, 사실+주장)

다음으로 ‘사라\*’ 라는 단어는 전체 기사에서 총 73회 등장했다. 이중 번역과 직접적 관계가 없는 경우는 56회였으며, 언어장벽 등이 사라진다는 가정을 포함하면 번역과 관련된 경우는 총 17회, 번역사와 직접적 관련된 경우는 8회에 불과했다. 번역사의 미래와 현재에 대한 서술 방법을 코딩한 결과 흥미롭게도 전망이 1회, 가정이 2회, 주관적 판단 2회, 사실+주장 3회이며, 객관적이고 전문적 정보에 입각한 언급이라 보기는 어려웠다. 또한, ‘사라지지 않는다’로 요약할 수 있는 언급이 7회, ‘사라진다’로 요약될 수 있는 언급은 1회에 불과했다.

<예시 6>

-사라진다

잡코리아가 최근 발표한 자료에 따르면 통번역가는 AI, 로봇에 의해 사라질 직업 1위에 꼽혔다. (다수의 언론사, 사실+주장)

-사라지지 않는다

언론이 인간 번역가의 위기 프레임을 조장했기 때문입니다. 그러나 인간 번역가는 사라지지 않을 것이고 <후략> (S 일보, 주관적 판단)

기사의 세부 내용을 보면 표면적으로는 ‘사라진다’ 혹은 ‘대체된다’는 단어가 나오지만, 기사 전체의 논조는 그렇지 않음을 주장하기 위해 가정이나 조건을 언급한 경우가 다수였다.

결론적으로 ‘대체’와 ‘사라진다’와 관련된 뉴스는 모두 사실적 관계에 기반하지만, 대체 여부의 문제는 주관적 주장에 기초했다는 것을 알 수 있다. 즉, 단순 인력의 대체 여부를 주관적으로 판단했지, 이것이 번역산업의 구조적 특성을 변화시키거나 사람들이 기계번역을 사용하는 사회적 구조의 변화를 설명하고 있지 않다. 이는 Iyengar(1990)가 주장한 일화적 프레임에 해당한다. 기계번역의 일자리 대체나 번역사의 사라짐에 대해 폭넓은 맥락에서 통계자료와 증언을 통한 구조적 현상을 파악하기보다 특정한 이벤트나 이슈를 설명하면서 주관적 판단으로 근거를 보강하려 했기 때문이다.

결국 일화적 프레임은 사건의 대안적 고민을 제시하지 않고 기존의 고정관념의 프레임에 맞추어 사건을 해석하게 된다(이재승과 박영숙, 2013). 즉, 기술발전이 인간의 일자리를 감소시킬 것이라는 고정관념이 산업혁명 당시와 마찬가지로 기계번역 도입에도 적용된 것이다. 기계번역이라는 새로운 기술의 등장에 적합한 새로운 관점을 통계와 증언, 그리고 맥락적 분석에서 대안적 방향을 제시하기보다 신기술의 등장을 기존의 고정관념의 프레임에 맞추어 이슈의 해석을 주장한 것이다. 따라서 기계번역의 도입으로 번역사가 대체되거나 사라진다는 접근은 주관적 고정관념에 기반한 일화적 프레임이다.

또한, 언론 보도에서도 포스트에디팅을 비롯한 통번역에 대한 다양한 통합적 접근으로 산업적 변화와 그에 따른 일자리 대체나 변화가 일부 논의되는 것을 고려하면, 이러한 고정관념의 일화적 프레임은 단순 흥미 위주의 접근을 통해 온라인에서 반복 재생을 목적으로 하고 있음을 유추할 수 있다. 이처럼 관심을 끌 수 있는 기사를 반복 재생시키는 미디어의 관행이 일반 대중에게 기계번역의 인간 대체 담론을 형성한 것으로 보인다.

4.1의 논의를 요약하자면 언론에서 적극적으로 의도된 프레임을 사용하여 기계번역에 대한 특정 메시지를 전달하려 노력했는지 판단하기는 쉽지 않다. 하지만 새로운 서비스가 출시된 다음 해인 2017년은 기계번역과 관련된 정보가 가장 많이 노출된 기간이었으며, 언론은 새로운 번역 기술을 소



개하기 위해서 ‘신경망’과 ‘NMT’ 같은 전문용어를 통해서 기술적 속성을 설명해 왔지만, 이 용어들은 빠르게 뉴스에서 사라졌다. 이는 기계번역을 이용하는 이용자가 그 속성을 파악하는 데 많은 시간이 걸리지 않았음을 시사한다.

또한, 기계번역과 관련된 보도는 중립성 혹은 객관성 확보를 위해 노력하고 있지만, 감성어휘 분석이나 기계번역을 둘러싼 어휘의 상세 분석을 보면, 보도의 내용이 기계번역의 긍정적 측면에 치우쳐 있다고 볼 수 있다. 따라서 언론의 이러한 논조가 일반 대중이 기계번역에 대한 긍정적인 인식을 형성하는 데 이바지했을 가능성이 크다.

일부 뉴스에서 객관적 증거나 통계 혹은 증언에 입각하지 않고 주관적, 나아가 비합리적 주장을 하는 사례도 있었다. 본 연구에서는 이러한 접근을 고정관념에 기반한 일화적 프레임으로 분류하였다. 기계번역의 도입으로 발생하는 산업과 사회의 변화를 구조적인 주제로 다룬 것이 아니라, 기존의 고정관념에 기반하여 해석의 방향을 일화적으로 제시한 것이다. 더욱 심각하게는 제목이 원문의 내용을 곡해하는 경우도 있었다. 비록 이러한 자극적 정보는 소수이지만, 일반 대중에게 강한 인상을 전달할 수 있고 뉴스를 클릭할 수 있는 동인이 되기도 한다. 결국, 언론이 만들어준 기계번역에 대한 긍정적 인식의 토대 위에 특정 프레임을 가진 소수의 자극적 정보가 일반 대중들이 기계번역에 대한 고정관념을 형성하는 데 기여했을 가능성이 크다 하겠다.

#### 4.2 언론 보도와 학계 논문의 담론 차이

하나의 대상에 대한 담론이 존재하기 위해서는 그 담론을 구성하는 주체와 주변의 요소들이 필요하다. 예를 들어 기계번역 자체는 하나의 기술일 뿐 스스로 행위성을 가지고 있지 않다. 미디어가 기계번역에 대한 뉴스를 구성하기 위해서는 행위성을 가지는 기술의 개발자, 사용자, 그리고 이 기술로 인해 영향을 주거나 받는 주체 및 객체들의 이야기는 필연적이며, 그 주체들과 연관된 핵심 단어가 담론을 구성한다고 할 수 있다. 이번 절에서는 언론 보도와 학계의 논문에 나타나는 기계번역 담론을 주도한 주체와 키워드를 분석하여 담론의 차이가 있는지를 살펴보고 이러한 차이가 특정 프

레이프로 이어지는지 살펴보고자 한다.

주체에 대한 상세 분석을 위해 워드스미스를 사용하여 전체 단어를 빈도수 별로 나열하는 워드리스트를 생성하였다. 이후, 주체로 인식할 수 있으며 사용 빈도가 높은 명사를 선별하는 과정을 거쳤다. 그 결과 ‘네이버’, ‘구글’, ‘카카오’, ‘파파고’, ‘번역사’, ‘번역가’, ‘번역자’, ‘교수’를 초기 분석 대상으로 설정할 수 있었다.

표 5  
언론 보도와 논문에 나타난 주체의 빈도

	언론	논문
파파고	409	437
네이버	1,076	325
구글	1,159	836
카카오	887	65
번역사	93	823
번역가	262	259
번역자	38	44
교수	647	96

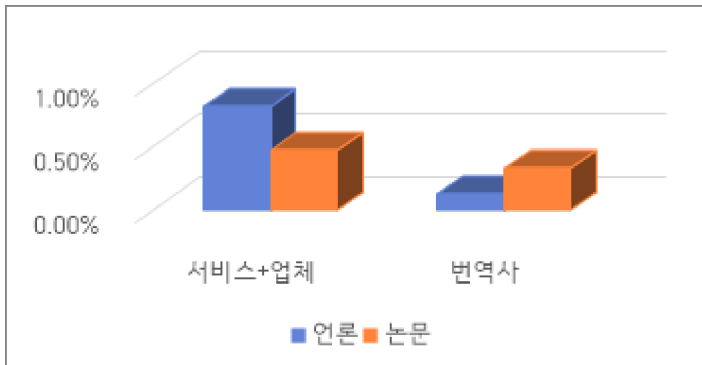
<표 5>에서 볼 수 있듯 언론 보도에서 기계번역의 담론은 서비스 제공사와 그 서비스가 주도하고 있었다. 특히 ‘네이버’와 ‘파파고’는 함께 1,485회 언급이 이뤄졌으며, ‘구글’은 1,159회나 언급이 되었다. 이는 ‘기계번역’이 1,109회 언급된 것보다 더 많다. 상기 데이터는 언론 보도는 기계번역 제공사에 대한 정보가 주를 이룬다는 Vieira(2020)의 연구 결과와 맥을 같이한다.

반면에 번역을 수행하는 사람을 통칭하는 ‘번역사’, ‘번역가’, ‘번역자’의 합산 언급 빈도가 총 393회밖에 되지 않는다는 점은 언론 보도에서 번역 수행의 주체인 인간은 담론의 주변부에 있음을 시사한다. 오히려 번역 수행의 주체보다 ‘교수’의 비중이 647회나 된다. 이는 학술적 전문성이 있는 교수의 언급을 인용, 객관적 정보 전달을 위해 언론에서 노력한 결과로 풀이된다. 하지만 번역학과 관련된 교수를 인용한 문구는 74회에 불과하며, 이는 이공

계 교수 중 가장 많이 인용된 교수 1인에 대한 언급(150회)에도 한참 미치지 못하는 수치이다. 특히 이공계 및 기술 개발과 관련된 교수들의 언급이 압도적으로 많았다는 점에서, 언론은 번역 수행 주체인 인간보다는 기술에 더욱 집중하는 양상을 보였다.

반면 논문에서는 '파파고', '네이버', '구글'의 언급은 언론 보도 대비 적으며, '번역사'의 언급은 언론 대비 절대량에서 많다. 물론 논문 코퍼스가 기사 코퍼스보다 2만 토큰 이상 크기 때문에, 가중평균을 적용해 보면 기계번역 서비스 제공업체에 대한 언급은 언론 보도 대비 논문 코퍼스에서 40% 정도 적으며, 번역 수행의 주체인 '번역사'와 '번역가'의 빈도는 100% 이상 많았다.

그림 1  
언론과 논문 코퍼스의 구성 요소 비교



즉, 언론은 서비스와 해당 서비스를 제공하는 업체가 논의의 중심이다. 하지만 학술 논문은 상대적으로 서비스 및 서비스 제공업체에 대한 논의는 적으며, 번역 수행의 주체인 인간번역사가 논의의 중심인 것으로 관찰된다. 이러한 논의 중심의 차이는 필연적으로 코퍼스 전체에서 사용되는 체언 및 용언의 차이로 이어질 수밖에 없다. 따라서 언론 보도와 논문 코퍼스에서 생성한 워드 리스트에서 상위 100위 안에 있는 단어 중 명사만을 선별하여 비교를 진행했으며 그 결과는 <표 6>과 같다.

표 6  
KCI 논문과 언론 보도 최빈출 명사

언론 보도	빈도	KCI 논문	빈도
AI	3,775	기계번역	4,464
기술	2,717	연구	3,591
서비스	2,597	포스트에디팅	2,275
인공지능	2,162	평가	2,087
기계번역	1,125	오류	1,852
데이터	1,106	분석	1,717
기능	763	텍스트	1,613
자동번역	721	문장	1,600
이하 교차 비교			
문장	545	기술	1,012
평가	462	데이터	624
분석	358	서비스	336
텍스트	166	AI	260
오류	112	인공지능	220

논문 코퍼스는 번역학 학회지에서 기계번역을 키워드로 사용한 논문만 추출할 것이기에 ‘기계번역’과 ‘연구’가 가장 많이 언급된 의미어인 것은 당연한 결과이다. 하지만 ‘기계번역’에 대한 언급은 총 4,464회로 언론 보도의 1,125회 대비 3.98배 밀도가 높다. 여기에 더해 논문은 ‘텍스트’ 및 ‘문장’ 등 기계번역의 결과물에 대한 ‘분석’이 주를 이룬 것을 볼 수 있다. 반면 언론 보도는 ‘인공지능’, ‘AI’, ‘기술’, ‘서비스’ 등 번역과 텍스트보다는 기술에 관한 서술이 주를 이뤘다.

또 한 가지 흥미로운 점은 언론 보도와 논문의 기계번역에 대한 입장 차이가 명확하게 관찰되었다는 점이다. 논문 코퍼스에서 ‘오류’가 1,852회나 언급되었고 명사 상위 5위 안에 든다는 것은 이상빈(2020)의 보고처럼 오류 중심의 기계번역 연구가 이뤄졌다는 점을 확인할 수 있는 데이터이다. 이는 4.1에서 논의한 대로 기계번역을 긍정적으로 바라보는 언론의 접근과는 사뭇 대조적이다. 일례로 ‘오류’, ‘한계’, ‘단점’ 등 기계번역의 부정적 측면을

대변하는 단어를 두 코퍼스 간에 비교하면 언론 보도와 논문 간에는 현격한 차이가 존재한다.

표 7

기사와 논문 코퍼스의 핵심 부정 명사 빈도 비교

단어	오류	한계	단점
기사종합	112	109	22
논문종합	1,852	244	138

이러한 기계번역에 대한 입장 차이를 다시 확인하기 위해, ‘기계번역’과 관련된 키큐던스를 각 코퍼스에서 추출하여 부정적 감성어휘 분석을 텍스트롬에서 추가로 시행하였다. 텍스트롬은 부정적 감성을 거부감, 슬픔, 놀람, 분노, 두려움, 통증 등으로 자체 분류하고 있으며, 분석 결과 가장 큰 부정적 감정은 거부감으로 드러났다.

추가 분석을 위해 기계번역에 대한 거부감을 나타내는 상세 어휘를 두 코퍼스에서 추출하여 출현 빈도와 종류를 비교해 보았다. <표 8>에서 확인할 수 있는 바와 마찬가지로, 논문 코퍼스에서 기사 코퍼스 대비 훨씬 더 다양한 종류의 부정적 감성어휘가 높은 빈도로 나타났음을 확인할 수 있었다.

표 8

기사와 논문 코퍼스의 거부감 어휘 종류 및 빈도 비교

언론 보도	빈도	KCI 논문	빈도
난해하다	48	어렵다	120
권위적	12	부족하다	40
복잡하다	8	별로	20
부담스럽다	6	복잡하다	17
어색하다	5	심하다	13
부족하다	5	나쁘다	13
어렵다	3	어색하다	11
	이하 1회 미만	심각하다	10

		추상적	6
		부담스럽다	6
		난해하다	6
		불만족하다	4
		번거롭다	4

따라서 언론 보도에서는 기계번역에 대해 긍정적인 언급이 상대적으로 많지만, 논문에서는 기계번역에 대해 부정적 언급이 상대적으로 많다고 분석할 수 있다.

하지만 번역학계의 논문이 단순히 기계번역의 단점을 부각하기 위한 프레임에 취한 것이 아니라, 이를 극복하기 위한 기계번역의 활용방안을 제시하고 있다는 사실을 ‘포스트에디팅’의 높은 언급 빈도에서 알 수 있다. 그리고 기계번역 결과물의 수정 기법 및 도구와 관련된 어휘인 ‘포스트에디팅’, ‘MTPE’, ‘CAT’의 빈도를 비교하면 언론과 논문의 기계번역에 대한 접근법 차이는 더욱 극명해진다.

표 9

기사와 논문 코퍼스의 포스트에디팅 관련 단어 빈도 비교

단어	포스트에디팅	MTPE	CAT
논문종합	1,218	201	71
기사종합	18	10	2

이상의 논의를 요약하자면 기계번역에 대한 학계의 논의와 언론 보도에는 큰 격차가 존재한다는 것이다. 학계는 논의의 방점이 번역과 번역사에게 있으며, 기계번역의 장점보다는 단점과 한계에 집중하는 경향을 보였다. 그리고 그 결과 자연스럽게 기계번역의 단점을 보완하는 방안에 집중하고 있다. 따라서 프레임 이론으로 보자면 일화적 프레임보다는 주제적 프레임에 가까운 접근을 취하고 있다. 하지만 언론 보도는 기계번역이란 기술의 긍정적인 측면에 더욱 집중하고, 개인의 사용성을 중심으로 부각하고 있다. 따라서 기술 변화에 따른 사회구조적 변화를 고찰한 것이 아니라, 기계번역

신기술의 등장이라는 사건에만 초점을 맞춘 일화적 프레임을 취하고 있음을 다시 한번 확인할 수 있다.

상기 논의가 의미하는 바는 번역학계가 이러한 담론의 격차를 명확하게 이해하고 언론 보도와 학계 담론의 중간점을 발굴하고 제안할 수 있어야 한다는 것이다. 그리고 바로 그 시작은 기계번역의 장점과 단점 사이에 명확한 균형을 잡을 수 있는 정보와 통찰의 제공일 것이다. 그 다음 단계로는 기계번역의 장점을 가장 효율적으로 활용하고 단점은 보완할 수 있는 최적의 번역 결과물 산출 방식을 제안해야 할 것이다.

## 5. 결론

본고는 기계번역을 둘러싼 현실에 대해 세 가지 연구 질문을 제기하였고, 해당 연구 질문에 답하기 위해 코퍼스의 양적 분석과 텍스트 수준의 뉴스패키지 분석과 감성 분석을 실행하였다. 먼저 기계번역을 둘러싼 현실이 언론 보도를 통해 어떻게 구성되었는지 분석하기 위해 2017년부터 2022년 사이에 출간된 기계번역과 관련된 언론 기사를 수집하여 325,048 단어의 코퍼스를 구축하고 분석을 시도했다. 둘째, 번역학계의 기계번역에 대한 논의가 어떻게 구성되었는지 탐구하기 위해 348,502 단어의 논문 코퍼스를 구성하여 상세 분석을 진행하고 기사 코퍼스와 담론의 차이가 있는지 분석을 진행하였다. 셋째, 기사 코퍼스와 논문 코퍼스의 담론 차이에 기반하여 번역학의 미래 과제를 제안하기 위해 논의를 펼쳤다.

첫 번째 연구 질문인 뉴스 구성자의 중립적, 부정적, 혹은 긍정적 해석이 어느 정도 개입했는지 여부에 대해서는, 특정 의도를 가진 프레임을 가지고 언론 보도가 이뤄졌다고 보기는 어려웠다. 구체적으로는 기계번역의 홍보를 위해서 기사의 양이 폭발적으로 많았다거나, 기사의 내용이 비합리적으로 기계번역의 낙관적 측면만을 강조했다고 볼 수는 없었다. 다만 정량적 분석과 감성 분석의 결과를 볼 때 전반적으로 기계번역의 기술적인 측면과 긍정적인 측면에 기사가 편중된 것으로 나타났다. 또한, 프레임의 형식으로 보자면 기계번역이라는 이슈에 대해 신기술 확산의 영향력이나 구조적

변화를 포괄적으로 고려했다기보다는 신경망 번역이라는 신기술에 대해 현상적이고 고정관념에 기반한 특정 일화를 부각하는 프레임을 사용한 기사가 주류를 이뤘다. 만약, 기계번역에 대한 주제적 프레임으로 보도가 이뤄졌다면 기계번역에 관한 결과, 조건, 통계 증거, 전문가 증언 등이 더욱 보강되어야 했을 것이다. 달리 말해 기계번역 결과물의 실질적 사용을 연구하는 번역학계의 통계나 전문가 증언 등이 더 많이 사용되었어야 했는데, 언론 보도에서는 기계번역의 생산자인 이공계 전문가들의 주장이 주로 다뤄졌다는 점은 아쉬움을 남긴다.

여기에 더해 일부 기사에서는 본문에 없는 내용을 제목에 의도적으로 편집하여 독자의 관심을 끌기 위해 노력한 사례가 있었다. 또한, 인간번역사 대체에 관한 기사는 극소수였으나 이를 확대 및 재생산하는 언론의 관행이 일반 대중의 기계번역에 대한 잘못된 인식에 영향을 주는 것으로 풀이된다.

두 번째 연구 질문인 언론 보도와 번역학 논문에서의 기계번역에 대한 관점 차가 있는가에 대해서는, 큰 격차가 존재하는 것으로 나타났다. 언론 보도는 기계번역에 대해 일화적 프레임으로 기술과 장점에 집중하였다. 또한, 번역의 주체인 번역사보다는 서비스 제공사에 더욱 집중하여 기사를 작성하였다. 반면, 번역학 논문에서는 번역의 주체를 사람에게 두고 있다. 물론, 언론 보도 대비 기계번역에 대해 주제적 프레임으로 접근하려 노력하였으나 다수의 논문에서 기계번역의 결함과 단점을 일화적으로 지적하는데 그친 예도 있었다. 하지만 학술 논문에서는 기계번역의 단점만을 지적한 것이 아니라 기계번역을 합리적으로 활용하기 위한 포스트에디팅 등의 연구가 다수 이뤄졌다는 점, 그리고 송연석(2018)의 연구처럼 언론의 기계번역에 대한 막연한 긍정론에 균형을 잡기 위해 노력했다는 점은 기계번역에 대해 한층 더 포괄적이고 객관적인 현실 구현에 기여했다는 점에서 의미가 있다.

세 번째 연구 질문인 미래 번역학계의 역할에 대해서는, 언론과 연구 사이의 기계번역에 관한 담론이 매우 큰 격차를 보이는 시점에서 번역학계의 역할을 고민해야 한다. 먼저 소통의 형태에 대한 성찰이 필요하다. 이미 언론 보도는 일반 대중들에게 기계번역에 대한 장점과 편리성에 대한 정보를 전파했음을 인정할 필요가 있다. 그리고 학술 논문만을 통해서만 일반 대중



의 의식 변화를 도출하기에는 한계가 있음을 받아들여야 한다. 따라서 일반 대중과 소통하기 위한 언론 기고, 더 많은 사용자 교육을 위한 산학협력, 일반인 의식 개선을 위한 강연회 등이 더 많이 이뤄질 필요가 있다. 다음으로 소통하는 내용에 있어 조정이 필요하다. 기계번역에 단점이 있는 것은 사실이지만 기계번역의 단점만을 부각해서는 대중들이 받아들일 수 있는 현실을 구성하기 어려울 것이다. 따라서 기계번역의 단점과 함께 올바른 기능적 사용에 대한 논의가 더욱 많이 이뤄질 필요가 있다. 대표적으로 포스트에디팅, 프리에디팅, 그리고 이 작업을 효율적으로 수행하기 위한 도구에 대해서 더 많은 연구와 대중 소통이 필요할 것이다.

본 연구는 기계번역에 대한 언론 보도 및 학술 논문의 담론을 프레임 이론을 사용하여 비교하고 있다. 이를 통해 기계번역에 대한 현실이 어떻게 구성되고 있으며 학계는 그러한 현실에 어떻게 대응해왔는지를 고찰하고 미래 과제까지 제안했다는 점에서는 의미가 있다. 하지만, 2023년 기사와 논문이 분석에 포함되지 않았기에 본 고에서 제시한 내용은 시간이 지나면서 다르게 해석될 가능성이 있다. 또한, 동일 기사의 반복 재생성에 대해서 충분히 다루지 못하고 있으며, 독자의 반응이 본 연구의 분석 범위 밖이라는 점은 본 연구의 한계라 할 수 있다. 여기에 더해 지면의 한계 때문에 더 상세한 정성적 분석을 시행하지 못한 점 역시 아쉬움으로 남는다. 마지막으로 본고에서 제기된 기사 제목의 편집자 개입에 대한 문제에 관해서는 미래 연구 과제로 남겨두고자 한다.

## 참고문헌

- 강수정. (2020). 전문번역사들의 NMT에 대한 인식과 수용에 관한 연구 — 심층인터뷰를 중심으로. 번역학연구, 21(3), 9-35.
- 강수정. (2021). 통번역대학원생들의 기계번역 사용현황과 인식에 관한 연구. 번역학연구, 22(3), 41-63.
- 곽은주, 김동미. (2022). 인간과 기계번역의 공존을 위한 담론: 행위자 네트워크 이론을 중심으로. 통번역교육연구, 20(3), 5-32.

- 권상희. (2005). 인터넷 뉴스프레임: 인터넷 미디어발달의 장기적인 뉴스보도 경향연구. 한국언론정보학보, 30, 35-87.
- 송연석. (2018). 기계번역 담론에 대한 비판적 고찰. 번역학연구, 19(1), 119-145.
- 신지선. (2020). 기계번역 포스트에디팅에 관한 해외 연구 동향. 번역학연구, 21(4), 87-114.
- 유승관, 이제영. (2006). 유비쿼터스 환경에 대한 언론보도와 수용자의 인식 유형에 관한 연구. 한국언론정보학보, 32, 169-207.
- 이상빈. (2020). 기계번역에 관한 KCI 연구논문 리뷰: 인문학 저널 논문(2011~2020년 초)의 논의내용과 연구방법을 중심으로. 통역과 번역, 22(2), 75-104.
- 이재승, 박경숙. (2013). 지역 텔레비전의 다문화 관련 뉴스 프레임 분석: KBS제주, 제주MBC, JIBS, KCTV JEJU를 중심으로. 언론과학연구, 13(1), 326-358.
- 이준용. (2000). 프레임, 해석 그리고 커뮤니케이션 효과. 언론과 사회, 29, 85-153.
- 이준용. (2005). 갈등적 사안에 대한 여론 변화를 설명하기 위한 프레임링 모형 검증 연구: 정부의 통일 정책에 대한 뉴스 프레임의 형성과 해석적 프레임의 구성을 중심으로. 한국언론학보, 49(1), 133-162.
- 장애리. (2020). 중국 번역학계 ‘기계 번역’ 연구 동향 분석. *T&I REVIEW*, 10(2), 111-133.
- 전현주. (2023). ‘기계번역’ 관련 국내 번역학 연구동향 분석. 융합영어영문학. 8(3), 85-121.
- 천중성. (2020). 전문번역사들의 기계번역 수용에 관한 연구. 한국융합학회논문지, 11(6), 281-288.
- 천중성, 강수정. (2022). ‘인공신경망 기계번역’ 관련 기사의 의미구성에 대한 연구: 헤드라인에 대한 의미네트워크분석을 중심으로. 번역학연구, 23(4), 37-65.
- 최수이. (2022). 빅카인즈 시스템을 활용한 게더타운 뉴스 빅데이터 분석. 한국컴퓨터정보학회논문지, 27(3), 53-61.

- 최문선. (2019). 국내 번역학 기계번역 연구 동향:내용 분석과 키워드 분석을 중심으로. *언어학연구*, 24(1), 275-297.
- 하동엽. (2023). 빅인즈를 활용한 교권침해 관련 국내 언론보도 변화 분석. *교육행정학연구*, 41(4), 317-341.
- Cadwell, P., Sharon O., & Carlos S. C. T. (2017). Resistance and accommodation: Factors for the (non-) adoption of machine translation among professional translators. *Perspectives*, 26(3): 1-21.
- Christensen, T. P., Flanagan, M., & Schjoldager, A. (2017). Mapping translation technology research in translation studies. An introduction to the thematic section. *Hermes-Journal of Language and Communication in Business*, (56), 7-20.
- Entman, R. M. (1991). Framing US coverage of international. *Journal of Communication*, 41(4), 6-27.
- Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of Communication*, 43(4), 51-58.
- Guerberof, A. (2013). What do professional translators think about post-editing? *JosTrans*, 19, 75-95.
- Iyengar, S. (1990). Framing responsibility for political issues: The case of poverty. *Political Behavior*, 12(1), 19-40.
- Rogers, E. M., Singhal, A., & Quinlan, M. M. (2014). Diffusion of innovations. In D. W. Stacks, M. B. Salwen, & K. C. Eichhorn (Eds.), *Integrated approach to communication theory and research* (pp. 432-448). Routledge.
- Vieira, L. N. (2020). Machine translation in the news: A framing analysis of the written press. *Translation Spaces*, 9(1), 98-122.

## **Analysis of discourses on machine translation in media reports and scholarly articles**

**Jong-sung Chun**

Department of Culture and Media, Konkuk University

**Jun-ho Lee** (corresponding author, brandon4tni@cau.ac.kr)

Korean-English Advanced Interpretation and Translation Program, Chung-Ang University

### **Abstract**

This study addresses three questions regarding the reality of machine translation (MT). First, it explores how media reports have shaped perceptions of MT. To this end, the researchers conducted a mixed-method analysis of media reports published from 2017 to 2022, incorporating a quantitative corpus analysis, a text-level news package analysis, and a sentiment analysis. The findings indicate no significant framing bias; however, the articles tend to emphasize MT's technical and positive aspects, often framing it as a technological novelty rather than providing comprehensive discussions. Some articles present subjective and illogical arguments, potentially generating misconceptions about MT among the general public. Second, the study compares media reports and academic papers on MT. Media reports focus on the benefits of the technology and the perspectives of service providers, while academic papers mainly address human translators and the practical use of MT, including post-editing. This comparison highlights a significant discourse gap between the two realms. Third, building on the previous findings, the study discusses the future role of translation studies, emphasizing the need for improved public communication through media contributions, industry-academic collaborations, and public lectures. Despite limitations such as the omission of the most recent data and the lack of qualitative information, the study provides valuable insights into the framing of MT in media and academia, proposing future directions for the field of translation studies.

**Keywords:** Machine translation; human translation; media; sentiment analysis;

discourse analysis

**키워드:** 기계번역, 인간번역, 미디어, 감성 분석, 담화 분석

천종성(<https://orcid.org/0000-0003-0988-0035>)

건국대학교 문화미디어학과 조교수

[jongsungk@konkuk.ac.kr](mailto:jongsungk@konkuk.ac.kr)

이준호(교신저자, <https://orcid.org/0000-0003-0397-6829>)

중앙대학교 국제대학원 전문통번역학과 조교수

[brandon4tni@cau.ac.kr](mailto:brandon4tni@cau.ac.kr)

논문 투고: 2024년 5월 14일

1차 심사 완료: 2024년 5월 30일

2차 심사 완료: 2024년 6월 9일

게재 확정: 2024년 6월 15일