

AI 시대 통번역 담론의 미디어 프레이밍 분석: 한국 뉴스 빅데이터를 중심으로*

이지민(경상국립대학교)

1. 서론

인공지능(AI) 기술의 급속한 발전으로 인해 언어 서비스 산업 전반에 걸쳐 근본적인 변화가 발생하고 있다. 특히 기계번역, 실시간 통역, 음성 인식 기술의 정확도가 비약적으로 향상되면서, AI가 인간 번역가와 통역사를 대체할 수 있는가에 대한 논의가 학계와 업계를 넘어 일반 사회에서도 활발하게 제기되고 있다. 이러한 상황에서 대중의 인식과 담론 형성에 핵심적인 역할을 담당하는 언론 미디어가 AI 시대의 통번역 이슈를 어떤 방식으로 프레임화하고 있는지를 분석하는 것은 통번역학적·미디어학적으로 중요한 의미가 있다.

그간 국내 언론은 AI 번역 도구의 상용화, AI 동시통역 서비스의 확대, 통번역사 직업군의 위기론 등 다양한 측면에서 AI와 통번역의 관계를 보도해 왔다. 그러나 개별 보도의 관점과 논조를 넘어 언론 전체가 AI 시대의 통번역을 어떤 의제로 구성하고 있는지, 즉 어떤 주제를 중심에 놓고, 이러한 주제에 어떤 식으로 접근하는지에 대한 체계적인 분석은 아직 충분히 이

* 이 연구는 2026 경상국립대학교 신입교원 연구기반 조성 연구비에 의하여 수행되었음, GNU-NRFSG-0001

루어지지 않았다.

이에 본 연구는 텍스트마이닝 기법을 활용해 최근 1년간의 한국 뉴스 기사를 대상으로 AI 시대 통번역에 관한 미디어 담론의 주요 의제와 구조를 분석하고자 한다. 구체적인 연구 질문은 다음과 같다.

- 1) 한국 미디어는 AI 발전과 통번역을 어떠한 주제로 다루고 있는가?
- 2) 각 주제는 어떠한 가치 프레임으로 제시되는가?
- 3) 한국 미디어의 AI 시대 통번역 담론이 갖는 통번역학적 시사점은 무엇인가?

이를 위해 2장에서는 이론적 배경을 제시하고, 3장에서 연구 방법을 설명한 후 4장에서 연구질문 1과 2에 대한 분석 결과를 제시하며, 5장에서 이 결과가 지닌 통번역학적 시사점(연구질문 3)을 논의한다.

2. 이론적 배경

2.1 언론 담론과 미디어 프레임링 연구

미디어 프레임링(framing)은 언론이 현실을 있는 그대로 반영·전달하는 것이 아니라 특정한 방식으로 선택하고 구성한다는 문제의식에서 출발한다. 사회학자 고프만(Goffman, 1974)이 경험을 조직하는 해석의 틀로서 ‘프레임’ 개념을 제시한 이래, 엔트만(Entman, 1993)은 이를 커뮤니케이션 연구의 핵심 개념으로 정착시켰다. 그는 프레임링을 ‘선택(selection)’과 ‘부각(salience)’의 과정으로 규정하고, 텍스트의 일부 측면을 두드러지게 만들어 특정한 문제 정의, 원인 해석, 도덕적 평가, 해결책 제시를 촉진하는 것으로 정의하였다. 즉 프레임링은 담론이 ‘무엇을 말하는가’를 넘어 그것을 ‘어떻게 규정하고 평가하는가’의 문제로서 특정 사안을 수용자가 특정 관점에서 보고 생각하도록 유도하는 것이라 할 수 있다.

프레임링은 이후 미디어 효과를 설명하는 이론으로 발전했다(Scheufele,

1999). 갠슨과 모딜리아니(Gamson & Modigliani, 1989)는 미디어 담론이 사회적으로 구성된 ‘해석 패키지(interpretive package)’를 통해 여론과 상호작용한다고 보았다. 프레임은 사안을 개별 사건으로 다루는 일화적(episodic) 프레임과 맥락·구조 속에서 다루는 주제적(thematic) 프레임으로 나눌 수 있으며(Iyengar, 1991; Druckman, 2001), 이 프레임에 따라 수용자들이 책임을 지우는 주체가 개인 또는 정부·구조적 요인 등으로 달라지는 것으로 나타났다(Iyengar, 1991). 특히 사안을 위험과 기회라는 상반된 평가 방향으로 배치하는 가치 프레임은 EU 확대에 대한 유럽 언론 보도가 여론에 미치는 영향(Schuck & de Vreese, 2006)이나 인터넷 도입 초기 독일 언론의 보도 태도(Rossler, 2001)처럼, 새로운 이슈의 평가적 성격을 포착하는 데 활용되고 있다.

국내 커뮤니케이션학계에서도 프레이밍을 연구에 적용하고 있다. 차현주와 권상희(2020)는 2010년에서 2019년까지 국내 주요 커뮤니케이션 학술지 6개에 게재된 프레이밍 연구 98편을 메타분석했는데, 연구가 양적 연구와 이득손실 프레임을 독립 변인으로 삼는 실험설계에 편중되어 있고, 신문·텔레비전 등 전통 미디어를 다룬 연구가 다수이며, 소셜미디어 연구가 빠르게 부상하고 있다고 설명했다.

이외에 뉴스 담론을 계량적으로 분석한 연구들은 프레이밍을 통해 특정 대상이 어떻게 재현되는지를 분석했다. 최신인과 최은정(2015)은 신조어 관련 신문기사 51건을 프레임 분석과 언어 네트워크 분석으로 검토해, ‘언어 파괴’, ‘소통 장애’와 같은 부정적 프레임이 집중적으로 나타나고 있으며 규범적 집단만을 정보원으로 삼아 청소년을 ‘우리말 파괴의 주범’으로 재현하는 경향이 있다고 밝혔다. 이은별(2022)은 30여 년간의 아프리카 관련 국제 뉴스를 LDA 토픽모델링으로 분석하여, 한국 언론이 강대국의 시각을 무비판적으로 이식하고 테러, 내전, 빈곤 등 일탈적 의제만을 부각함으로써 대륙 전체를 타자화하는 ‘아프리카성’을 강화한다고 논증한다. 유소연과 임규건(2021)은 코로나19 감정 관련 뉴스 헤드라인을 의미 네트워크 분석으로 시각화하여 ‘코로나-블루’를 중심으로 한 담론 구조를 규명하였다.

인공지능이라는 기술 의제 자체를 다룬 연구로 박종화 등(2022)은 국내 뉴스 빅데이터 서비스(빅카인즈)와 미국 기사 데이터베이스(LexisNexis

Academic)를 통해 수집한 한·미 인공지능 관련 뉴스 17,295건을 대상으로 주제어 빈도분석, LDA 토픽모델링, 언어 네트워크 분석을 결합해 양국 언론이 인공지능을 어떻게 다루는지를 비교하였다. 그 결과 한국 언론은 챗봇, 스마트스피커 등 서비스 출시를 중심으로 인공지능을 미시적 수준에서 다루는 반면, 미국 언론은 직업의 미래, 프라이버시, 기술 패권 전쟁 등 거시적이고 비판적인 의제를 상대적으로 더 비중 있게 다룬다는 점을 밝혔다. 이들 연구는 토픽모델링, 네트워크 분석과 같은 텍스트마이닝 기법이 뉴스 담론의 프레임과 의미 구조를 대규모로 포착하는 데 유효함을 증명한다.

2.2 통번역 담론 프레임

통번역에 관한 미디어 프레임링을 분석한 연구는 국내외를 막론하고 매우 드문데, 비에이라(Vieira, 2020)는 기계번역 관련 영어 신문기사 284건을 대상으로 긍정·부정 프레임을 각 3개 문항(제목의 정서적 표현 여부, 기사 전반의 강조점(이점 대 단점), MT의 결과나 사용 결과 묘사)을 대상으로 이진식으로 코딩해 기사별 점수를 산출하는 가치 프레임링 분석을 수행하였다. 그 결과 전체의 57%가 긍정 프레임을 보였고 부정은 29%에 그쳤으며, 이러한 긍정 편향이 2016년 신경망 기계번역 출시 시기와 무관하다는 점에서 언론이 신기술을 애초에 ‘경이로운 대상’으로 규정하는 경향이 있다고 해석했다. 비에이라(2020)를 제외하면 통번역과 미디어의 관계를 다룬 기존 연구는 주로 통역사와 관련된 연구로 한정되는데, 주로 개별 사건이나 특정 직군을 대상으로 통역사라는 직업인의 이미지가 언론에서 어떻게 재현되는가를 다루었다. 일례로 두와 왕(Du & Wang, 2021)은 최근 10년간 중국 언론에 보도된 통역 관련 기사 60건을 대상으로 헤드라인, 본문의 주제, 수사적 분석과 사진의 다중모드 분석을 결합해 회의통역사의 이미지를 다차원적으로 분석하였다. 그 결과 통역사는 대형 행사나 유명 인사와 결부된 ‘스타’ 내지 공적 셀럽으로 재현되는 동시에 권력에 결속되고 대중과는 거리가 있는 존재로 그려지고 있는 것으로 파악되었다. 디리커(Diriker, 2022)는 2017년 6월부터 2021년 3월까지 터키에서 발행부수가 가장 많은 신문 3종의 온라인판을 분석하여 무엇이 통역사에 관한 미디어 담론을 촉발하며 통역사가 어떻게 재현되는지를 규명하였는데, 법정통역 관련 언급이 가장 많이 나

타남에도 대부분 스쳐 지나가는 방식으로만 언급되어 해당 영역 통역사에 게는 가시성과 발언권이 거의 주어지지 않았다고 설명한다. 친과 리(Qin & Li, 2026)는 2021년 미중 알래스카 회담에 관한 중국 소셜미디어 뉴스 보도를 다중모드로 분석하여, 통역사가 이례적으로 ‘국민 스타’라는 중심적 이미지(미모의 스타, 몰모델, 국가 대표)로 재현된 사례와 이를 구성하는 프레이밍 전략을 규명하였다.

국내에서는 통번역사에 대한 언론 프레이밍 중심의 연구는 찾아보기 힘들며, 대중의 인식에 대한 연구가 주를 이룬다. 먼저 권상미(2020)는 봉준호 감독의 통역사로 화제를 모은 샤론 최에 대한 대중의 인식을 살펴보았는데, ‘봉준호 수상소감’을 검색어로 유튜브 조회수 상위 10개 영상을 선정하고 셀레니움을 이용하여 해당 영상의 댓글 총 4,192건을 수집하였다. 수집된 텍스트는 키워드 분석, 연관어 분석, LDA 토픽모델링 기법으로 분석했는데, 그 결과 댓글에서 가장 높은 빈도로 등장한 키워드는 ‘통역’(통역이라는 행위와 통역사라는 사람을 동시에 지칭)이었다. 분석 결과, 수용자들은 통역사의 전달력과 이해력을 높이 평가하는 한편, 외모, 나이, 표정과 같은 외형적 요소나 어학 실력 자체에도 상당한 관심을 보이는 것으로 나타나, 미디어 속 통역사가 여전히 피상적 이미지 위주로 소비되고 있음을 시사하였다. 검색어가 ‘봉준호 수상소감’이었음에도 정작 최고 빈도 키워드가 ‘봉준호’나 ‘수상’이 아닌 ‘통역’이었다는 점은 통역사라는 매개자에 대한 대중적 관심이 감독에 대한 관심 못지 않게 크다는 것을 시사하며, 이는 본 연구가 주목하는 ‘통번역사 담론의 가시성’ 문제와도 맞닿아 있다.

장애리(2025)도 통역사 샤론 최 관련 유튜브 영상의 댓글을 분석해 대중 평가에 나타난 인간 통역사의 차별적 강점을 파악했다. 댓글 3,762개를 대상으로 KoNLPy(Okt)를 활용한 텍스트마이닝, 공출현 분석, 의미 네트워크 중심성 분석과 함께 LDA 토픽모델링을 수행했다. 이를 통해 대중 평가에 나타난 인간 통역사의 핵심역량은 언어 능력, 이해 능력, 커뮤니케이션 능력이 라고 결론지었다. 장보윤과 허지운(2026)은 이른바 ‘쿠광 청문회’ 통역 논란과 관련한 유튜브 댓글 10,897개를 셀레니움으로 수집하고, LDA 토픽모델링(1단계 K = 6, 2단계 K = 5)과 대표 댓글에 대한 정성 분석을 결합하는 한편, 전문 통역사 6명을 면접하여 대중과 통역사의 인식 간극을 규명하였

다. 대중은 통역사에게 자의적 요약을 기대하거나 통역사와 발원 주체를 동일시하는 등 통역 규범에 대한 인식 결여를 드러낸 반면, 통역사들은 공적 회의에서의 ‘통역 프로토콜’ 제정을 촉구하는 등의 차이가 있는 것으로 나타났다.

이처럼 국내에서의 통번역사에 관한 담론은 주로 댓글을 대상으로 해 대중의 인식을 파악하는 것에 집중했는데, 대규모 텍스트로 분석한 연구는 토픽모델링을 적극적으로 활용하는 것으로 나타났다. 토픽모델링은 대량의 문서 집합으로부터 잠재 주제(토픽)를 추출하는 텍스트 마이닝 기법으로, 그 중 잠재 디리클레 할당(Latent Dirichlet Allocation: LDA)이 가장 간단한 형태다(Blei, 2012, p. 78). 방대한 온라인 담론에서 대중 인식의 잠재 구조를 포착하는 데 효과적이며, 본 연구가 뉴스 기사에 LDA를 적용하는 것과 방법론적으로 연결된다.

요컨대 기존 통번역사 인식 연구는 특정 사건에 대한 대중 반응(댓글)에 초점을 두어, 통번역 직업이 AI를 도구이자 위협으로 양가적으로 인식하며 대중은 통역 규범을 종종 오해한다는 점을 밝혔다. 또한 앞서 살펴본 통역사 미디어 재현 연구(Diriker, 2022; Du & Wang, 2021; Qin & Li, 2026)는 통역사의 직업적 이미지가 특정 사건이나 국가적 맥락에서 어떻게 구성되는지를 규명했으나, 그 분석 단위는 통역사라는 행위자의 이미지에 머물러 있어 AI라는 기술적 의제가 통번역 전반에 걸쳐 언론에서 어떤 프레임으로 구성되는지는 다루지 않는다. 아울러 이러한 인식이 형성되는 담론적 환경, 즉 언론이 AI 시대의 통번역을 어떤 의제와 관점으로 프레임화하는지를 대규모로 분석한 연구는 아직 드물다. 따라서 개별 사건이나 특정 집단이 아니라 최근 1년간의 통번역 관련 뉴스 담론 전체를 대상으로 미디어 프레임링을 분석함으로써 통번역사를 포함한 통번역에 관한 인식 연구를 미디어 담론의 층위에서 보완할 필요성이 발생한다.

3. 연구 방법

3.1 연구 대상 및 자료 수집

본 연구는 한국언론진흥재단이 운영하는 뉴스 빅데이터 분석 시스템 빅카인즈(BigKinds)를 활용해 자료를 수집했다. 빅카인즈는 국내 54개 언론사의 뉴스를 표준화된 메타데이터(일자, 언론사, 제목, 본문 등)와 함께 제공하는 국내 유일의 뉴스 빅데이터 분석 시스템으로, 개별 언론사 홈페이지를 일일이 크롤링하는 방식에 비해 표본의 대표성과 수집의 재현 가능성을 확보할 수 있다는 장점이 있다. 수집 기간은 2025년 7월 1일부터 2026년 6월 30일까지 최근 1년으로 설정했는데, AI 생성형 기술이 통번역 산업에 본격적으로 도입 및 확산된 시점 이후의 담론을 포착함으로써 연구 질문 1과 2(현재 시점의 주제와 프레임 구조)에 답하기 위한 것이다. 키워드는 AI 시대 통번역 담론을 포괄적으로 수집하기 위해 설계한, “AI AND (통역 OR 번역 OR 통역사 OR 번역사 OR 통번역 OR 통번역사)”로 정했다. 이 검색어는 기사 제목을 대상으로 적용하였다. 제목 단위 검색은 본문 전체를 대상으로 한 검색에 비해 AI와 통번역이 기사의 부차적 소재로만 스쳐 지나가는 사례(예: 통번역과 무관한 기사에 AI 관련 문구가 본문 일부에 등장하는 경우)를 원천적으로 배제한다. 기사를 전수 검토한 사후 관련성 필터링 작업을 통해서도 표본의 주제 적합성이 확보되었음을 확인했다.

수집된 기사는 총 489건이었다. 이 중 동일 제목의 기사가 동일 일자 동일 언론사에 중복 수록된 경우를 확인해 제거했다. KBS 5건, 매일경제 1건, 파이낸셜뉴스 1건, 헤럴드경제 1건 등 총 8건을 제거해 최종 481건을 분석 대상으로 확정했다. 분석 대상 언론사는 조선일보, 동아일보 등의 종합 일간지 외에도 전자신문, 헤럴드경제, 파이낸셜뉴스, 이데일리, 머니투데이 등 경제산업 전문지가 혼재했다. 수집된 기사에는 제목, 본문, 언론사명, 보도일자, URL 등의 정보가 포함되어 있다. 빅카인즈에서는 본문을 첫 200자로 한정해 제공하므로 구체적인 담론 파악을 위해서는 전체 본문이 필요해 빅카인즈에서 확보한 URL 목록을 대상으로 생성형 AI 어시스턴트(Claude)의 웹 페이지 콘텐츠 조회 기능을 이용해 URL 단위로 본문을 개별 확인·추출했다.

접근이 차단된 77건은 연구자가 수작업으로 본문을 수집했다. AI를 통해 수집한 기사 중 50건을 연구자가 무작위 표집해 해당 URL의 원문과 대조해 확인한 결과 모두 정확하게 수집이 되었음을 확인했다.

본 연구의 표본 규모(481건)는 한·미 인공지능 뉴스 담론을 비교한 박종화 등(2022, 17,295건)에 비해 작다. 그러나 박종화 등(2022)은 ‘인공지능’이라는 포괄적 기술 의제를 국가 간 비교를 위해 대량 수집한 반면, 본 연구는 ‘AI×통번역’이라는 특정 산업·직업 의제로 검색어를 좁혀 제목 단위로 수집했으므로 모집단 자체가 481건으로 한정된다. 본 연구는 제목 단위 검색으로 표본의 주제 적합성을 우선한 결과 481건으로 표본이 한정되었고, 문서당 평균 토큰 수(166.7개)도 역시 크지 않다. 단, LDA는 문서 수가 적더라도 토픽 수(K)를 낮게 설정하고 응집도 검증을 거치면 안정적인 결과를 얻을 수 있다(Griffiths & Steyvers, 2004).

3.2 분석 방법

3.2.1 전처리

수집된 기사 본문을 분석에 적합한 형태로 가공하기 위해 다음 절차를 순차적으로 수행했다. 첫째, URL, 기자명, 저작권 고지 등 분석과 무관한 정보를 정규표현식으로 제거했다. 둘째, Python KoNLPy의 Okt 형태소 분석기를 사용해 본문에서 명사를 추출했다. 셋째, Okt가 복합명사를 분리하는 한계를 보완하기 위해 ‘인공지능’, ‘동시통역’, ‘생성형 AI’, ‘기계번역’ 등 통번역 분야의 주요 복합용어는 원본 텍스트에서 정규표현식으로 직접 추출해 토큰 목록에 추가하는 이중 추출 방식을 적용했다. 넷째, ‘기자’, ‘저작권’, ‘무단재배포’ 등 뉴스 특유의 불용어와 한 글자 단어를 제거했다. 다섯째, 분석의 일관성을 위해 ‘인공 지능’과 ‘인공지능’을 ‘AI’로, ‘동시 통역’을 ‘동시통역’으로 통일하는 용어정규화를 수행했다. 전처리 결과, 토큰 총 80,192개, 고유어휘 6,252개가 추출되었으며, 문서당 평균 토큰 수는 166.7개였다.

추출 시 명사만을 추출한 이유는 LDA와 언어네트워크 분석과 같은 토픽·의미 구조 분석에서 명사가 담론의 소재를 가장 직접적으로 반영하기 때문이며, 이는 한국어 텍스트마이닝 연구에서 통용되는 방식이기도 하다(예: 이은별, 2022). 단, 이 과정에서 동사나 형용사에 담긴 평가 정보(예: ‘우려하

다’, ‘기대된다’)가 소실될 수는 있으나, 이는 원문 전체를 대상으로 하는 가치 프레이밍 코딩을 통해 상당 부분 보완된다.

3.2.2 주제어 빈도 분석 및 언어네트워크 분석

주제어 빈도 분석과 언어네트워크 분석은 기사 본문만을 대상으로 수행했다. 본고의 표본은 제목에 AI 및 통번역 관련 키워드가 동시에 출현하는 기사로 표집했으므로, 제목을 빈도 집계에 포함할 경우 표집 조건이 된 검색어(AI, 번역, 통역 등)가 설계상 자동으로 과다계상되는 순환적 편향 문제가 발생한다. 이를 방지하기 위해 제목은 빈도 집계 대상에서 제외했다. 아울러 라틴문자로 표기되는 ‘AI’는 형태소 분석 과정에서 단독어와 복합어(예: AI번역, 생성형AI, AIT통역)으로 나뉘므로 빈도 보고 시 AI단독어와 복합어를 구분해 제시한다.

3.2.3 토픽모델링과 가치 프레이밍

본 연구는 담론의 ‘무엇(주제)’과 ‘어떻게(가치 방향)’를 개념적으로 구분한다. 토픽은 담론의 소재적 구성을 가리키고, 가치 프레임은 그 소재에 부여된 긍정·부정의 평가 방향을 가리킨다. 본고에서는 LDA로 각 기사의 대표 토픽(T)을 확정하고, 가치 프레이밍(긍정적 프레이밍은 P, 부정적 프레이밍은 N) 코딩을 각 기사에 독립적으로 적용한 다음 토픽 × 가치 교차분석을 수행한다.

먼저 주제 구조 추출에는 LDA(Blei, 2012; Griffiths & Steyvers, 2004)를 사용했다. 토픽 수 결정을 위해 $K = 2$ 부터 8까지 c_v 응집도 점수(Röder et al., 2015)를 3회 반복 산출한 결과, $K = 3\sim 4$ 구간($M = .33$)에서 정체되어 있던 응집도가 $K = 5$ ($M = .40$)에서 뚜렷이 상승했으며 이후 상승폭이 완만해지는 경향을 보였기에 $K = 5$ 를 최종 토픽 수로 채택하고 각 토픽의 확률가중치 상위 20개 단어를 질적으로 검토해 해석 가능성을 확인했다.

가치 분류를 위해서는 언론 보도를 대상으로 위험과 기회 유형의 가치 프레임을 코딩해 보도의 긍정 및 부정 비율을 측정한 비에이라(2020)의 가치 프레이밍을 대규모 뉴스 코퍼스에 적용 가능한 형태로 확장했다. 비에이라(2020)는 긍정·부정 각 3개 문항(총 6개)을 기준으로 사용한 반면(2.2절 참

조), 본 연구는 인간-AI 협업, 공공서비스 편익 등 비에이라(2020)의 6개 문항으로는 포착되지 않는 담론적 요소가 반복적으로 관찰됨에 따라 프레임 요소를 8개(긍정 4, 부정 4)로 확장했다. 긍정측은 P1 성능기술 진보, P2 편익·유용성(언어장벽 해소, 업무 효율 등), P3 기회·긍정 전망, P4 인간-AI 협업 보조이고, 부정측은 N1 오류·품질 한계, N2 직업·생계 위협, N3 위험·우려(프라이버시 포함), N4 과장·비판이다. 작성한 코딩북(부록 참조)에 따라 전체 기사에 대해 8개 프레임 요소가 있음(1)/없음(0)으로 코딩했다. 긍정 점수가 부정 점수보다 크면 ‘긍정’, 그 반대면 ‘부정’으로 분류했고, 두 점수가 같으면 중립으로 분류했다.

아울러 부정 기사 내에서 부정성의 귀속 대상을 세분화하기 위해, N1(오류·품질 한계) 또는 N3(위험·우려)가 코딩된 기사는 ‘AI 기술 자체에 대한 비판’으로, N2(직업·생계 위협)만 단독으로 코딩되고 N1, N3는 코딩되지 않은 기사는 ‘AI 옹호형’(AI의 성능이나 효용 자체는 문제 삼지 않으면서 인간 전문직의 대체 가능성만을 위협 요인으로 제시하는 기사)으로 구분했다(관련 내용은 4.3절에서 논의).

이어 제2코더가 코드북을 따라 전체의 약 10%인 48건(전체 코퍼스에서 낱자 오름차순 정렬 후 매 10번째 기사)과 T2 79건 전수를 대상으로 독립적으로 코딩했다. T2 79건을 전수 검증 대상에 포함한 것은 예비 검토에서 전체 부정 기사 중 다수가 이 토픽에 집중될 것으로 관찰되었기 때문이다. 부정 사례처럼 발생 빈도가 낮은 범주는 무작위 표집만으로는 검증 표본에 충분히 포함되지 않을 위험이 있다. 이중 코딩 대상은 48건과 T2 79건을 합쳐 총 127건으로, 이는 전체 481건의 26.4%에 해당한다. 이는 비에이라(2020)가 채택한 약 10% 검증 표본보다 넓은 범위이기도 하다. 이후 Cohen’s k 값을 산출했는데, 그 값은 $k = .91$ 로 나타났다. 이는 랜디스와 코크(Landis & Koch, 1977)의 기준상 ‘거의 완벽한(almost perfect)’ 일치 수준($k = .81 \sim 1.00$)에 해당한다. 해당 코딩이 8개 프레임 요소의 있음/없음이라는 비교적 명확한 조작적 정의를 따르는 이진 코딩이며, 8개 항목의 점수 비교로 최종 3분류(긍정, 부정, 중립)를 최종적으로 도출하는 절차의 특성상 충분히 나올 수 있는 수치다.

토픽과 가치의 연관성 파악을 위해서는 토픽×가치 교차표를 구성하였다.

교차표에서 기대빈도가 5미만인 셀이 있어 독립성 검정은 피셔의 정확검정(Fisher's test)을 적용했다.

3.2.4 언어 네트워크 분석

언어 네트워크 분석은 텍스트에서 단어들의 공출현 빈도를 바탕으로 단어 간 관계를 네트워크 형태로 표현하는 기법이다. 주제어 빈도분석이 개별 단어의 중요도를 파악하고, 토픽모델링이 단어들의 잠재적 주제군집을 추출하는 것이라면, 언어 네트워크 분석은 담론 내에서 단어들이 어떤 맥락으로 연결되어 있는지의 구조, 즉 의미 구조를 시각화한다는 점에서 세 방법이 상호보완적으로 기능한다.

전체 코퍼스 중 기사 본문을 분석 대상으로 해서 상위 빈도 단어 100개를 노드로 설정하고, 두 단어가 같은 기사 안에 함께 출현한 빈도가 80회 이상인 단어 쌍을 네트워크의 엣지로 연결했다. 이는 문장 단위의 공출현을 적용한 최신훈과 최은정(2015)과 달리 문서(기사) 단위의 공출현을 채택한 것으로, 문서 단위 공출현은 연결이 상대적으로 조밀해지는 경향이 있어 상위 100개 단어와 공출현 80회 이상이라는 기준으로 네트워크의 밀도를 조정했다. 네트워크 구성과 지표 산출에는 Python의 NetworkX 라이브러리를 사용해 연결 중심성(degree centrality)을 산출했다. 연결 중심성은 한 노드가 네트워크 내 다른 노드들과 얼마나 많이 연결되어 있는가를 나타내는 지표로, 값이 높을수록 담론 내에서 중심적 역할을 하는 단어임을 의미한다.

4. 분석 결과

4.1 주제어 빈도분석

주제어 빈도분석 결과, 481건 기사 전체에서 가장 높은 빈도를 보인 단어는 ‘번역’(1,930회)으로 나타났으며, 이어 ‘언어’(1,055회), ‘서비스’(1,029회), ‘AI’([단독] 789회), ‘기술’(710회), ‘실시간’(701회), ‘통역’(692회), ‘기능’(657회), ‘외국인’(560회) 순으로 나타났다.

AI를 필수 키워드로 해 검색한 기사임을 고려할 때 AI 빈도가 상대적으로 낮은 것이 아닌가 하는 의문을 가질 수 있으나, 위의 빈도는 단독으로 등장한 경우(789회)만을 나타낸 것이며, 다른 어휘와 결합한 복합어로 ‘AI번역’(239회), ‘생성형AI’(157회), ‘AI통역’(105회)이 별도로 출현해 AI복합어는 총 501회에 이른다. 단독과 복합을 합산하면 AI계열은 1,290회로 ‘언어’, ‘서비스’를 상회하는 최상위 빈도를 형성한다. 아래 <표 1>은 본문 기준 상위 30개 토큰 빈도를 나타낸다.

표 1

본문 주제어 빈도 상위 30(제목 제외, N=481)

순위	토큰	빈도	순위	토큰	빈도
1	번역	1,930	16	동시통역	392
2	언어	1,055	17	건설	365
3	서비스	1,029	18	다국어	355
4	AI(단독)	789	19	정보	339
5	기술	710	20	시스템	333
6	실시간	701	21	솔루션	329
7	통역	692	22	개발	327
8	기능	657	23	환경	313
9	현장	588	24	고객	302
10	지원	569	25	데이터	297
11	외국인	560	26	플랫폼	290
12	기업	489	27	확대	284
13	글로벌	434	28	업무	269
14	음성	418	29	영어	264
15	도입	400	30	운영	263

아래 <표 2>는 통역 또는 번역을 포함한 토큰만을 집계한 결과다. AI 기술과 관련된 어휘 중에서는 ‘동시통역’(392회), ‘실시간번역’(134회), ‘실시간통역’(109회), ‘기계번역’(55회)이 자주 등장했다. 주목할 점은 ‘번역’(1,930회)이 ‘통역’(692회)의 약 2.8배에 달한다는 것으로, AI 기술이 텍스트 번역 분

야에서 더 빠르게 상용화되는 현실과 함께 미디어가 통역보다 번역에 더 크게 주목하고 있음을 반영한다.

표 2
번역/통역 포함 개별 토큰 빈도

순위	토큰	빈도	순위	토큰	빈도
1	번역	1930	11	번역기	72
2	통역	692	12	다국어번역	57
3	동시통역	392	13	기계번역	55
4	AI번역	239	14	자동번역	50
5	번역가	166	15	순차통역	11
6	통번역	136	16	통번역사	10
7	실시간번역	134	17	번역본	9
8	실시간통역	109	18	번역자	8
9	AIT통역	105	19	번역사	7
10	통역사	89	20	번역학	3
			21	번역인력	3

직업인으로서의 통번역사를 살펴보면, ‘번역가’(‘번역자’, ‘번역사’ 포함)는 181회, ‘통역사’(‘통번역사’ 10회 포함)는 99회 등장해 1,029회 나타난 ‘서비스’의 10~18% 수준에 그쳤다.

4.2 LDA 토픽 모델링

LDA 토픽 모델링을 통해 5개의 토픽(T)을 도출했는데, 각각은 핵심어에 기반해 다음과 같이 명명했다.

- T1 공공·금융 통역 서비스
- T2 AI 시대 통번역사 역할 담론
- T3 AI 웨어러블 번역기기
- T4 기업용 통번역 솔루션

T5 건설현장 소통·미디어 콘텐츠 번역

T1 ‘공공·금융 통역서비스’는 총 96건으로 ‘외국인’, ‘통역’, ‘다국어’, ‘고객’, ‘상담’, ‘보험’, ‘행정’, ‘금융’, ‘민원’ 등이 핵심어로 나타났다. 공공행정 및 금융기관에서 외국인을 위한 AI 통역 서비스를 도입하는 내용을 주로 다루고 있다.

T2 ‘AI 시대 통번역사 역할 담론’은 ‘번역가’, ‘인간’, ‘한국’, ‘통역사’, ‘교수’, ‘문화’, ‘문학’, ‘교육’, ‘영어’ 등이 핵심어로, AI 시대 통번역사의 역할과 직업적 정체성에 관한 논의를 담고 있다. 총 79건이 이에 해당한다.

T3 ‘AI 웨어러블 번역기’는 70건이 포함되는데, ‘안경’, ‘구글’, ‘스마트’, ‘메타’, ‘글라스’, ‘학습’, ‘사용자’, ‘음성’ 등이 핵심어로, 구글과 메타 등 글로벌 빅테크 기업의 AI 스마트 안경에 탑재된 실시간 번역 기능에 관한 보도로 구성된다.

토픽4 ‘기업용 통번역 솔루션’은 ‘기업’, ‘글로벌’, ‘통역’, ‘솔루션’, ‘동시 통역’, ‘업무’, ‘데이터’, ‘혁신’, ‘회의’ 등이 핵심어로 나타났고, 기업 환경 내의 AI 통번역 솔루션 도입과 확산을 다루는 기사가 가장 많이 포함되어 있다. 기사 수도 160건으로 가장 많다.

토픽5 ‘건설현장 소통 및 미디어 콘텐츠 번역’은 76건으로 ‘현장’, ‘건설’, ‘음성’, ‘안전’, ‘근로자’, ‘호반’, ‘콘텐츠’, ‘보이스’, ‘텍스트’, ‘더빙’ 등이 핵심어로, 건설현장 외국인 근로자와의 소통 지원 및 미디어 콘텐츠의 AI 더빙 및 번역을 포함한다.

4.3 토픽별 가치 분포

토픽별 가치 분포 분석 결과는 아래 <표 3>과 같다.

표 3
토픽별 가치 프레임 분포(N=481)

토픽	긍정	부정	중립	계 (비중)
T1 공공·금융 통역 서비스	96 (100%)	0 (0.0%)	0 (0.0%)	96 (20.0%)
T2 AI 시대 통번역사 역할 담론	52 (65.8%)	18 (22.8%)	9 (11.4%)	79 (16.4%)
T3 AI 웨어러블 번역 기기	62 (88.6%)	6 (8.6%)	2 (2.9%)	70 (14.6%)
T4 기업용 통번역 솔루션	156 (97.5%)	1 (0.6%)	3 (1.9%)	160 (33.3%)
T5 건설현장 소통·미디어 콘텐츠 번역	75 (98.7%)	0 (0.0%)	1 (1.3%)	76 (15.8%)
총계	441 (91.7%)	25 (5.2%)	15 (3.1%)	481 (100%)

토픽과 가치는 통계적으로 유의한 연관성이 있는 것으로 나타났다(Fisher's exact test, $p < .001$, Cramér's $V = .312$). 서비스와 솔루션 성격의 토픽(T1, T4, T5)은 사실상 거의 대부분이 긍정으로 분류되었으며, 특히 공공·금융 통역 서비스(T1)는 100%가 긍정적인 프레임으로 나타났다.

본 분석에서 가장 두드러지는 점은 부정 프레임이 특정 토픽에 편중되어 있다는 것이다. 전체 부정 비율은 25건(5.2%)에 불과하지만, 이 가운데 72%(18건)가 T2(AI 시대 통번역사 역할 담론)에 집중되어, T2만 놓고 보면 부정 비율이 22.8%에 달한다. T2의 기사의 부정적인 프레임은 통번역학과·외국어학과 폐지, 번역가 일감 급감(‘업계 추정 약 80%’), 인간 번역의 고유 가치 옹호, AI 번역의 낡아스 및 문화 맥락 한계 등 ‘직업 정체성·생계 위협’과 ‘AI 한계’ 담론으로 구성된다. 그런데 이 부정적 기사 18건 가운데 N1(오류·품질 한계) 또는 N3(위협·우려)가 코딩되어 AI 기술 자체에 비판적인 기사)는 11건(해당 토픽의 13.9%)에 불과했다. 반면 N2(직업·생계 위협)

1) AI 기술의 한계나 위험을 직접 지적한다. 예컨대 "AI가 소설도 번역도 대신할까 서울국제도서전의 답은"(전남일보, 2026.06.25.)에서는 AI가 시대적 맥락을 완전히

만 단독으로 코딩되어 AI의 성능은 문제 삼지 않으면서 통번역사의 일자리 위협만을 다룬 ‘AI 옹호형’ 기사²⁾는 7건(토픽 대비 8.9%)으로 나타나, 부정 프레임 내에서도 AI 기술 자체에 대한 비판보다는 인간 노동의 대체 가능성에 초점을 맞춘 보도가 두드러진 것으로 나타났다. 이는 여전히 AI에 대한 긍정적 프레임이 담론 전반을 지배하고 있음을 시사한다.

반면 T3(AI 웨어러블 번역 기기)의 부정적 기사 6건(8.6%)은 전부 동일한 사건, 즉, KBS의 아르테미스호 발사의 실시간 AI 자막 오역 사고에 대한 보도이다. 2026년 4월 2일(머니투데이)에서 4월 3일(서울신문, YTN, 한국일보, 매일경제, 파이낸셜뉴스)에 집중되어 있으며 동일한 사건에 대한 보도로 한정된다. 이는 기술 자체에 대한 지속적 비판이 아니라 단일 사건이 여러 매체에 중복 보도된 결과다. 따라서 해석 시 이를 단일 사건 효과로 분리해 다루는 것이 타당할 것이다.

4.4 언어 네트워크 분석 결과 분포

언어 네트워크 분석 결과, 상위 100개 단어로 구성된 백본 네트워크(공출현 80회 이상)는 75개 노드와 493개 엣지로 구성되고(밀도 0.178), ‘생성형AI’가 형태소 분석 과정에서 분절되어 형성된 2개 노드가 고립된 것을 제외하면 73개 노드가 하나의 연결된 네트워크(거대 성분)를 형성했다.

담론의 중심에는 기술·서비스 관련 어휘가 자리한다. 강도(가중 연결도, 해당 연결이 얼마나 자주 발생했는가를 나타냄) 기준 상위 어휘는 ‘언어’(10,658), ‘번역’(10,033), ‘실시간’(9,470), ‘AI’(9,321), ‘서비스’(8,869), ‘기술’(8,339), ‘지원’(8,221)의 순이었다. 매개 중심성(서로 다른 집단을 잇는

전달하지 못한다는 우려가 제기되었으며, "AI는 평면적 번역 인간만큼 다양한 해석 못해"(매일경제, 2025.10.12.)에서도 AI 번역의 해석적 한계가 지적되었다.

2) AI의 성능이나 결함을 언급하지 않은 채 대체 결과만을 서술한다. 일례로 "AI의 일자리 위협, ‘번역가’에게는 이미 현실...번역가協 ‘번역 일 80% 줄었다’"(아주경제, 2025.12.14.)는 번역 일감이 약 80% 감소했다고 전했고 "AI에 발맞춰 전공 대수술 中 대학, 번역·촬영 등 16개 전공 폐지"(노컷뉴스, 2026.03.10.) 역시 AI의 한계에 대한 언급 없이 학과 구조조정 사실만을 전했다.

다리 역할을 얼마나 하는가를 나타냄) 역시 ‘언어’(0.233)와 ‘번역’(0.231)이 가장 높아, 이 두 어휘가 서로 다른 하위 영역을 잇는 교량 역할을 함을 보여준다. 가장 강하게 연결된 단어 쌍은 ‘번역-언어’(281회), ‘언어-실시간’(274회), ‘언어-AI’(270회), ‘번역-AI’(267회)로, 담론의 축이 ‘AI, 실시간, 언어, 번역’이라는 기술 중심 어휘로 짜여 있음을 드러낸다. 아래 표4는 강도 기준 상위 10개 단어의 중심성 지표를 나타낸다.

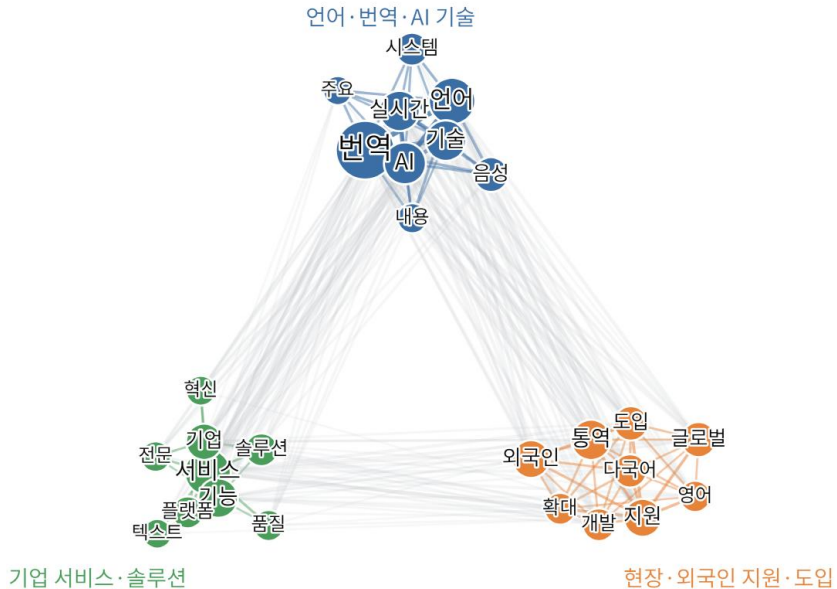
표 4

언어 네트워크 중심성 상위 10개 단어(백분, 공출현 $\geq$ 80)

순위	단어	빈도	강도	연결중심성	매개중심성	군집
1	언어	1,055	10,658	0.905	0.233	군집1
2	번역	1,930	10,033	0.905	0.231	군집1
3	실시간	701	9,470	0.716	0.074	군집1
4	AI	789	9,321	0.770	0.087	군집1
5	서비스	1,029	8,869	0.676	0.058	군집3
6	기술	710	8,339	0.676	0.052	군집1
7	지원	569	8,221	0.568	0.025	군집2
8	기능	657	7,153	0.473	0.022	군집3
9	기업	489	6,618	0.365	0.006	군집3
10	도입	400	6,356	0.338	0.003	군집2

군집 분석 결과 크게 세 개의 의미 영역이 도출되었다. 첫째, 가장 큰 군집(41개 어휘)은 ‘언어, 번역, AI, 기술, 실시간, 음성, 동시통역’ 등으로 구성되어 ‘AI 번역·통역 기술’의 핵심 영역을 이룬다. 둘째, ‘통역, 현장, 지원, 외국인, 글로벌, 도입, 다국어’로 구성된 군집(17개)은 산업 현장과 외국인을 대상으로 한 통역 지원 및 도입 담론을 나타낸다. 셋째, ‘서비스, 기능, 기업, 솔루션, 고객, 플랫폼, 품질’로 구성된 군집(15개)은 기업의 번역·통역 서비스 및 솔루션 담론을 형성한다. 아래 그림 1은 이 군집 구조를 시각화한 것이다.

그림 1
AI 통번역 담론 언어네트워크



이러한 네트워크의 의미 구조는 앞선 토픽모델링 및 가치 프레이밍 분석과 정합적이다. 담론이 ‘AI 번역·통역 기술’, ‘현장·외국인 지원 및 도입’, ‘기업 서비스와 솔루션’의 세 축으로 조직되는 양상을 통해 기술·산업 솔루션 프레임이 압도한다는 가치 분석 결과(전체 긍정 91.7%)와 토픽 구성(공공·금융 통역 서비스, 기업용 통번역 솔루션, 건설현장 소통 등)이 단어 관계의 층위에서도 지지됨을 확인했다.

5. 논의

5.1 기술 및 산업 솔루션으로서의 AI 통번역 프레임 지배

본 연구에서 가장 두드러진 사항은 한국 미디어의 AI 통번역 담론이 압

도적으로 기술·산업 솔루션 프레임 중심으로 구성되어 있다는 점이다. 토픽 모델링에서 가장 많은 기사(160건, 33.3%)가 포함된 T4(기업용 통번역 솔루션)를 위시하여, T1(공공·금융 통역 서비스, 96건)과 T5(건설현장 소통 및 미디어 콘텐츠 번역, 76건)까지 포함하면 전체의 69.0%(332건)가 AI 통번역을 기술 도구 또는 산업 서비스의 관점에서 다루고 있다. 이는 상위 빈도어에서 ‘서비스’(1,029회), ‘기술’(710회), ‘기업’(489회), ‘솔루션’(329회)이 높은 순위를 차지한다는 사실과도 일치한다. 이는 언어 네트워크 분석 결과와도 정합적이다. <표 4>의 강도(가중 연결도) 지표에서 상위권을 차지한 어휘 역시 ‘서비스’(8,869), ‘기술’(8,339), ‘지원’(8,221)으로 상위 빈도어와 대체로 일치하며, 특히 군집 분석에서 ‘서비스, 기능, 기업, 솔루션, 고객, 플랫폼, 품질’이 하나의 독립된 의미 군집을 형성한다는 점은 기업 서비스·솔루션 담론이 토픽모델링의 T4(기업용 통번역 솔루션)와 별개의 층위에서도 응집된 하나의 의미 영역을 이루고 있음을 뒷받침한다.

이러한 긍정 편향의 정도는 선행연구와 대비할 때 더욱 두드러진다. 긍정 프레임이 57%, 부정 프레임이 29%로 나타난 영어권 기계번역 보도 연구(Vieira, 2020)의 연구와 대비해, 본 연구의 한국 뉴스 표본에서는 긍정이 91.7%이고 부정은 5.2%에 그쳤다. 두 연구의 표본 시기, 언어권, 매체 구성이 상이해 직접적인 인과 비교는 어렵지만, 이러한 격차는 한국 언론 보도에 기업·서비스 출시를 알리는 기사(T1, 3, 5)의 비중이 상대적으로 크다는 표본 구성상의 특징과 맞물려 있을 가능성이 있으며, 이는 4.3절에서 지적인 T1의 100% 긍정 문제와도 함께 해석할 필요가 있다.

이러한 서비스·제품 중심의 미시적 프레임은 인공지능 전반에 대한 국내 언론 보도 경향과도 궤를 같이한다. 전술했듯이 박종화 등(2022)은 한국 언론의 인공지능 기술과 서비스 출시라는 미시적 수준에서 이루어지는 반면, 미국 언론은 직업의 미래, 프라이버시, 기술 패권 전쟁 등 거시적이고 비판적인 의제를 상대적으로 더 폭넓게 다룬다고 보고한 바 있다. 본 연구의 AI 통번역 담론에서도 기업용 솔루션과 공공서비스 출시 기사(T1, T4, T5)가 압도적 비중을 차지하고, 통번역사라는 행위자나 그 직업적 함의를 다루는 담론(T2)은 상대적으로 주변화되어 있다는 점에서, 이는 통번역이라는 특정 영역에 국한되지 않는 한국 언론의 인공지능 보도 전반의 구조적 경향의 연

장선으로 해석할 수 있다.

이러한 패턴을 통해 알 수 있는 점은 한국 미디어가 AI통번역을 주로 경제적 가치와 산업적 효율성의 프레임을 통해 보도하고 있다는 것이다. 미디어 프레임 이론의 관점에서 볼 때, 이러한 기사는 독자들이 AI통번역을 기술 혁신의 성과로 인식하도록 하지만 통번역의 문화적·윤리적·직업적 차원의 복잡성은 상대적으로 주변화하는 효과를 가져올 수 있다.

5.2 통번역사 담론의 주변화와 부정적 관점의 편재

총 481건 기사 중 AI시대 통번역사의 역할과 정체성을 언급하는 기사(T2)는 79건(16.4%)에 불과하다. 빈도분석에서도 번역가(번역자, 번역사 포함)는 181회, 통역사(통번역사 10회 포함)는 99회 등장해 1,029회 나타난 서비스의 10~18% 수준에 그쳐, 직업인으로서의 통번역사에 대한 담론이 기술 서비스 담론에 비해 현저히 주변화되어 있음을 알 수 있다.

언어 네트워크 분석에서도 ‘번역사(가)’와 ‘통역사’는 연결 중심성 상위권에 포함되지 않았다. 특히 통번역 직업인을 지칭하는 ‘번역가’, ‘통역사’ 등의 어휘는 강한 공출현 연결을 형성하지 못해 백본 네트워크의 주변부에 머물거나 진입하지 못하는데, 이는 부정·직업위협 담론이 전문직 정체성이라는 특정 영역(T2)에 국소적으로 편재하고 있다는 앞선 발견과 맞물린다.

이는 통번역학적으로 중요한 함의가 있다. AI가 통번역 시장에 미치는 영향에 대한 전문직 담론이 학계와 업계에서 활발히 전개되고 있음에도 불구하고 한국 미디어는 통번역사 당사자의 관점과 직업적 강약점을 충분히 조명하지 않고 있다. 이러한 패턴은 한국 미디어가 AI 통번역을 주로 경제적 가치와 산업적 효율성의 프레임에서 보도하고 있음을 다시 한 번 확인한다. 특히 긍·부정 평가에서도 드러났듯이 한국 미디어에서 ‘위협, 우려, 한계’ 등의 프레임이 AI 통번역 기술 전반에 고르게 분포하는 것이 아니라 전문직 정체성 담론에 집중되어 있다. 즉, 기술과 서비스는 긍정 일색으로 재현되는 반면, 부정적 평가는 ‘사람(전문직)’의 문제로 프레임되는 구조가 관찰된다. 이는 대중의 통번역사 직업에 대한 인식 형성에 부정적인 영향을 미칠 수 있으며, 통번역 전문 교육의 사회적 가치를 약화시킬 수 있다.

5.3 이주 노동자 및 외국인 소통 도구로서의 프레임

T1과 T5를 합산한 172건(35.8%)에서 ‘외국인’(560회), ‘근로자’(201회) 등의 단어가 핵심어로 등장하는 점을 볼 때, 한국 미디어가 AI 통번역을 이주 노동자 및 외국인의 언어 장벽 해소 수단으로 빈번히 보도하고 있음을 알 수 있다. 해당 기사에서 ‘행정’은 99회(T1 93회, T5 6회), ‘민원’은 107회(T1 107회, T5 0회) 출현한다. 특히 ‘민원’은 전체 코퍼스(112회) 중 대부분이, ‘행정’도 전체 131회 중 상당수가 T1(금융·통역서비스)에 집중되어, 두 어휘가 T5(건설 현장)보다도 공공행정 및 민원창구 맥락에서 훨씬 더 뚜렷하게 나타나는 담론임을 알 수 있다. 이는 한국 사회의 외국인 노동자 급증이라는 사회적 맥락, 특히 건설·제조업 현장의 언어 소통 문제가 언론 보도에 적극 반영된 결과로 해석된다.

5.4 실시간성의 강조와 번역·통역 속성 재정의

‘실시간’이 전체 6위(701회)로 높은 빈도를 보이며, 공출현 분석에서도 AI-실시간, 번역-실시간이 최상위 쌍에 포함된다는 사실을 볼 때 한국 미디어가 AI 통번역의 핵심 가치로 ‘속도’와 ‘즉시성’을 가장 강조하고 있음을 알 수 있다. 이는 T4의 ‘동시통역’(392회)과 T5의 ‘음성’(418회) 등 음성 기반 번역 기술에 대한 관심과도 맥락을 같이한다.

이러한 미디어 프레임은 번역·통역의 가치를 정확성과 문화적 적합성보다 속도와 접근성으로 재정의하는 경향을 내포한다. 통번역학의 관점에서 볼 때 번역의 품질에 대한 사회적 기대를 변화시킬 수 있는 담론적 효과가 있다.

5.5 번역과 통역 담론의 불균형

‘번역’(1,930회)이 ‘통역’(692회)과 비교해 빈도가 약 2.8배에 높은 불균형을 볼 때 AI기술이 텍스트 번역 분야에서 더 빠르고 상용화되고 있는 현실을 반영함과 동시에, 미디어가 통역보다는 번역에 더 크게 주목하고 있음을 알 수 있다. 동시통역을 포함한 구술 통역 분야에서도 AI의 영향이 확대되고 있음에도 불구하고(서은경, 2026) 이에 대한 담론적 관심은 상대적으로

제한적이다. 이는 통역 분야의 직업적 변화와 교육적 대응에 관한 사회적 논의가 아직 충분히 활성화되지 않았음을 시사한다.

6. 결론

6.1 요약

본 연구는 2025년 7월 1일부터 2026년 6월 30일까지 최근 1년 간 한국 언론에 보도된 AI 시대 통번역 관련 기사 481건을 대상으로, 주제어 빈도분석, LDA 토픽모델링, 가치 프레임링 코딩, 언어 네트워크 분석을 결합해 미디어 담론의 구조를 규명했다. 분석 결과는 다음 네 가지로 요약된다.

첫째, 한국 미디어의 AI 통번역 담론은 기업용 통번역 솔루션, 공공·금융 통역 서비스, 건설현장 소통 및 미디어콘텐츠 번역을 중심으로 하는 기술산업 솔루션 프레임이 압도적으로 나타났으며, 전체 기사의 91.7%가 긍정 프레임으로 분류되어 비에이라(2020)가 보고한 영어권 보도의 긍정 비율(57%)을 크게 상회했다.

둘째, 통번역사라는 직업인의 역할 및 정체성을 다루는 담론은 전체의 16.4%에 그쳤고, 전체 부정 기사(25건)의 72.0%(18건)가 이 토픽에 집중되어 (T2 내부 부정 비율 22.8%), 기술서비스는 긍정 일색으로, 부정적 평가는 ‘사람(전문직)’의 문제로 프레임되는 비대칭 구조가 확인되었다.

셋째, 이주 노동자·외국인 관련 담론에서 ‘행정’과 ‘민원’ 빈도가 두드러져, AI 통번역이 공공행정·민원 창구의 언어 장벽 해소 수단으로 주로 소환되고 있음을 보여주었다.

넷째, ‘실시간’이 상위 빈도어로 나타나고 ‘번역’이 ‘통역’보다 2.8 배 더 많이 언급되어, 미디어가 통번역의 가치를 정확성보다 속도접근성 중심으로, 그리고 통역보다 번역 중심으로 재구성하고 있음을 시사했다.

6.2 의의와 한계

본 연구의 통번역학적 의의는 다음과 같다. 첫째, 개별 사건이나 특정 직군에 국한되었던 기존의 통역사 미디어 재현 연구(Diriker, 2022; Du & Wang, 2021; Qin & Li, 2026)와 달리, AI시대 통번역이라는 기술산업적 의제 전반을 대규모 뉴스 코퍼스로 체계적으로 분석함으로써 통번역사 인식 연구를 미디어 담론의 층위에서 보완했다는 의의가 있다. 특히 481건을 질적으로 통독하는 방식으로는 포착하기 어려운 대규모 패턴 — 예컨대 부정 프레임의 72%가 단일 토픽(T2)에 집중되는 구조적 편향 — 은 토픽×가치 교차분석을 통해서만 통계적으로 확인 가능하다는 점에서, 본 연구 결과는 질적 내용분석과 구별된다.

둘째, 한국어 뉴스 코퍼스에 가치 프레이밍 코딩을 적용하고 토픽모델링과 결합함으로써, 담론의 ‘무엇’과 ‘어떻게’를 개념적으로 분리해 분석하는 방법론적 틀을 제시했다. 아울러 주제어 빈도분석, LDA 토픽모델링, 언어 네트워크 분석을 결합하는 본 연구의 분석 절차에는 박종화 등(2022)이 한미 인공지능 뉴스 담론을 비교하는 데 적용한 텍스트마이닝 접근을 통번역이라는 특정 산업·직업 영역으로 확장·적용했다는 방법론적 의의도 있다. 물론 표본 규모(481건)는 한미 인공지능 뉴스 담론을 비교한 박종화 등(2022, 17,295건)에 비해 작다고 할 수 있으나 두 연구는 모집단 정의 자체가 다르다. 박종화 등(2022)은 ‘인공지능’이라는 포괄적 기술 의제를 국가 간 비교를 위해 대량 수집한 반면, 본 연구는 ‘AI x 통번역’이라는 특정 산업 및 직업으로 검색어를 좁혀 제목 단위로 수집했으므로 모집단이 상대적으로 작을 수밖에 없다.

아울러 미디어가 통번역사라는 행위자보다 기술서비스를 압도적으로 긍정적인 프레임으로 재현하는 동시에, 부정적 평가는 전문직 정체성 담론에 국한시키는 비대칭 구조를 실증함으로써, 통번역 전문 교육의 사회적 가치와 통번역사 직업 정체성에 대한 대중 인식 형성에 미디어가 미칠 수 있는 영향을 구체적으로 제시했다.

그러나 본 연구는 다음과 같은 한계를 지닌다. 먼저, 기술적 측면에서 Cohen’s $\kappa = .91$ 은 8개 프레임 요소 각각이 아니라 최종 3 분류 라벨에 대한 일치도이므로, 항목별 신뢰도가 이보다 낮을 가능성을 후속 연구에서 별도

로 검증할 필요가 있다.

또한 본 연구에서 확장한 8개 프레임 요소는 비에이라(2020)의 6개 문항을 토대로 연구자가 코퍼스 검토 과정에서 판단해 도출한 것으로, 전문가 패널이나 별도의 예비조사를 통한 항목 타당도 검증을 거치지 않았다는 한계가 있다. 후속 연구에서는 이를 보완할 필요가 있다.

그리고 본 연구는 대상 데이터 기간이 최근 1년에 불과해 해당 담론이 지금까지 지속되어 온 담론이었는지 또는 변화하고 있는지의 여부, 또는 예외적인 보도 행태였는지의 여부에 대한 검증이 불가능했다는 것이다. 특히 AI 도입의 이전 또는 초기와 현재의 보도에 차이가 있을 것으로 보이나 현 연구에서는 연구 설계상 해당 분석을 수행할 수 없었다는 약점이 존재한다. 또한 한국어 뉴스로 표본을 한정했으므로, 타 언어권·타 시기 연구와의 비교는 시사적 수준에 머문다는 한계가 있다. 후속 연구에서는 시계열 비교, 언론사 유형별(IT/경제지 vs 종합지) 프레임 차이, 통번역사 개인 인터뷰를 통한 담론-인식 간극에 대한 심층 분석 등으로 보완할 필요가 있다.

참고문헌

- 권상미. (2020). 유튜브 댓글 토픽모델링 분석을 통한 봉준호의 통역사 샤론 최(Sharon Choi)에 대한 수용자의 인식 연구. *통역과 번역*, 22(2), 1-28.
- 박종화, 임민성, 김정환. (2022). 언론은 인공지능(AI)을 어떻게 다루는가?: 뉴스 빅데이터를 통한 한국과 미국의 보도 경향 분석. *정보시스템연구*, 31(1), 175-195.
- 서은경. (2026. 7. 10.). AI와 통번역: 공존을 위한 숙련 캠프 [세미나 발표]. 한국통번역사협회, 서울, 대한민국.
- 유소연, 임규건. (2021). 텍스트 마이닝과 의미 네트워크 분석을 활용한 뉴스 의제 분석: 코로나 19 감정을 중심으로. *지능정보연구*, 27(1), 47-64.
- 이은별. (2022). 한국 국제뉴스의 아프리카 보도 지형: 빅카인즈를 이용한 토픽모델링을 중심으로. *한국아프리카학회지*, 67, 157-194.
- 장보윤, 허지운. (2026). 전문 통역에 대한 통역사와 대중의 인식 연구 쿠팡

- 청문회 통역 사례를 중심으로. *T&I REVIEW*, 16(1), 7-46.
- 장애리. (2025). AI 시대 인간 통역사의 강점 극대화를 위한 역량 탐색: 사론
최 통역 유튜브 동영상 댓글 분석을 중심으로. *번역학연구*, 26(3),
349-393.
- 차현주, 권상희. (2020). 프레임(Framing) 이론 메타분석: 2010-2019 주요
학술지에 게재된 연구를 중심으로. *한국소통학보*, 19(4), 119-148.
- 최신인, 최은정. (2015). 신문기사의 신조어 재현 양상 연구. *새국어교육*,
102, 215-243.
- Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*,
55(4), 77-84. <https://doi.org/10.1145/2133806.2133826>
- Diriker, E. (2022). The (re)presentation of interpreters in the Turkish media: A
diachronic and cross-setting analysis. *The Translator*, 28(2), 196-214.
<https://doi.org/10.1080/13556509.2022.2105677>
- Druckman, J. N. (2001). On the limits of framing effects: Who can frame?
Journal of Politics, 63(4), 1041-1066.
<https://doi.org/10.1111/0022-3816.00100>
- Du, Y., & Wang, B. (2021). A multi-dimensional analysis of the representation
of conference interpreters in the Chinese media. *Translation Spaces*, 10(2),
306-328. <https://doi.org/10.1075/ts.18011.du>
- Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm.
Journal of Communication, 43(4), 51-58.
<https://doi.org/10.1111/j.1460-2466.1993.tb01304.x>
- Gamson, W. A., & Modigliani, A. (1989). Media discourse and public opinion
on nuclear power: A constructionist approach. *American Journal of
Sociology*, 95(1), 1-37. <https://doi.org/10.1086/229213>
- Goffman, E. (1974). *Frame analysis: An essay on the organization of
experience*. Harvard University Press.
- Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of
the National Academy of Sciences*, 101(Suppl 1), 5228-5235.
<https://doi.org/10.1073/pnas.0307752101>

- Iyengar, S. (1991). *Is anyone responsible? How television frames political issues*. University of Chicago Press.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159-174.
- Qin, K., & Li, X. (2026). The interpreter as a “national star”: A multimodal study of Chinese social media news narratives on the diplomatic interpreters in the 2021 ChinaUS Alaska talks. *Interpreting*, 28(1), 91-121. <https://doi.org/10.1075/intp.00126.qin>
- Rössler, P. (2001). Between online heaven and cyberhell: The framing of “the internet” by traditional media coverage in Germany. *New Media & Society*, 3(1), 49-66. <https://doi.org/10.1177/14614440122225985>
- Röder, M., Both, A., & Hinneburg, A. (2015). Exploring the space of topic coherence measures. *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining (WSDM '15)*, 399-408. <https://doi.org/10.1145/2684822.2685324>.
- Scheufele, D. A. (1999). Framing as a theory of media effects. *Journal of Communication*, 49(1), 103-122. <https://doi.org/10.1111/j.1460-2466.1999.tb02784.x>
- Schuck, A. R. T., & de Vreese, C. H. (2006). Between risk and opportunity: News framing and its effects on public support for EU enlargement. *European Journal of Communication*, 21(1), 5-32. <https://doi.org/10.1177/0267323106060987>
- Vieira, L. N. (2020). Machine translation in the news: A framing analysis of the written press. *Translation Spaces*, 9(1), 98-122. <https://doi.org/10.1075/ts.00023.nun>

부록: 코드북

1. 코딩 원칙

- 각 항목은 독립적으로 판정한다.
- 한 기사가 여러 개의 긍정 및 부정 항목에 동시에 1을 받을 수 있다.
- 부정 및 긍정의 판단은 단순한 한 단어가 아니라 문장 또는 단락 수준을 기준으로 한다.
- 인용되는 화자의 평가도 기사의 프레임으로 간주한다.
- 제목만으로 코딩하지 않고 반드시 본문 전체를 읽고 판정한다.

2. 코딩 기준

프레임	기준	설명
긍정	P1	정의
		AI 통번역의 정확도, 성능 향상, 기술적 진보·혁신을 부각
		판정 단서 “정확도가 높아졌다”, “비약적 발전”, “업계 최고 성능”, 신제품·신기술 출시 성과 강조
		예시 “생성형 AI의 정확도가 높아지면서...” / “실시간 번역 성능을 대폭 개선”
	P2	정의
		비용 절감, 효율화, 접근성 향상, 언어장벽 해소 등 실용적 편익을 부각
		판정 단서 “비용 효율화”, “언어장벽을 없앤다”, “누구나 손쉽게”, 외국인·이주민 소통 지원
		예시 “QR만 찍으면 바로 번역” / “공사장 언어 소통문제를 해결”
	P3	정의
		산업 성장, 시장 기회 언급. 이용자와 사회에 대해 긍정적인 전망을 제시.
		판정 단서 “시장 선점”, “새로운 기회”, “확대·성장 전망”, 투자·수출 성과
		예시 “베트남서 현장 계약 성과” / “글로벌 시장 공략 본격화”
	P4	정의
		AI를 인간 통번역(사)의 대체가 아닌 보조·협업 도구로 긍정적 서술
		판정 단서 “통번역사의 생산성을 돕는”, “인간과 협업”, “보조 도구로 활용”
		예시 “AI를 활용할 수 있는 개발자가 살아남는 구조” / “후편집

			으로 품질을 높인다”
부정	N1	정의	오역, 정확도 부족, 맥락·취향스 전달 실패 등품질 한계 부각
		판정 단서	“오역”, “맥락을 못 살린다”, “아직 부정확”, 실패, 오작동 사례
		예시	“그 시대 맥락을 AI가 알고 번역하는 것은 불가능” / “엉뚱한 번역으로 혼선”
	N2	정의	통번역사의 일자리 대체, 수요 감소, 직업적 위기 부각
		판정 단서	“번역가가 사라진다”, “일자리 위협”, “수요 감소”, 인력 대체, 구조 조정
		예시	“번역가 대신 AI를 이용하려는 사람들을 경계” / “인력 의존도가 감소하는 추세”
	N3	정의	문화적 맥락 상실, 과의존, 문해력 저하, 윤리, 오·남용 등 사회적 위험과 우려 부각
		판정 단서	“문해력 외주”, “과의존”, “문화가 사라진다”, 프라이버시, 윤리우려
		예시	“AI에 문해력을 외주 주변...” / “인간 고유의 판단이 사라질 우려”
	N4	정의	AI 통번역에 대한 과장 광고를 비판하거나 회의적·냉소적 태도를 드러냄.
		판정 단서	“과대 포장”, “기대에 못미친다”, “허상”, 마케팅 과장 지적
		예시	“‘모든 언어’ 번역이라는 주장은 과장” / “실상은 기대와 달랐다”

3. 코딩 예시

기사(요지)	P1 P2 P3 P4	N1 N2 N3 N4	
대우건설 AI, CCTV 실시간 번역으로 건설 현장 외국인 근로자 안전 소통 지원	1 1 1 0	0 0 0 0	긍정
서울국제도서전: 번역가들이 AI의 맥락 및 정서 번역 한계와 문해력 외주 우려	1 0 0 0	1 1 1 0	부정
게임업계: AI가 QA, 번역반복 업무를 대체해 생산성 상승, 인력 의존도와 채용 감소	1 1 1 0	0 1 0 0	긍정

Media framing of translation and interpreting discourse in the age of AI: A Korean news big data analysis

Jimin Lee

Department of English, Gyeongsang National University

Abstract

This study examines how Korean news media frame translation and interpreting (T&I) in the age of AI, using a corpus of 481 news articles collected via BigKinds over one year (July 1, 2025–June 30, 2026). Combining keyword frequency analysis, LDA topic modeling (K=5), valence frame coding, and semantic network analysis, the study identifies both the thematic content and the positive, negative, or neutral valence of media discourse. Results show that Korean coverage is overwhelmingly organized around technology- and industry-solution frames, with positive framing predominating overall. By contrast, discourse addressing interpreters' and translators' professional roles and identities was comparatively marginal, yet concentrates most of the negative coverage found in the corpus—revealing an asymmetric structure in which technology and services are represented positively while negative framing is displaced onto human professionals. Coverage also emphasizes AI translation as a tool for resolving language barriers faced by migrant workers and foreign residents, foregrounded "real-time" processing, and referenced translation nearly three times more often than interpreting. These findings suggest that Korean media construct AI-era T&I primarily through economic and industrial-efficiency frames, with implications for public perceptions of T&I professionals and the social value of T&I education.

Keywords: Media framing; AI translation and interpreting; topic modeling; valence frame; semantic network analysis; news big data

키워드: AI 통번역, 토픽모델링, 가치 프레임, 언어 네트워크 분석, 뉴스 빅데이터

이지민

경상국립대 영어영문학부 부교수

ke9836@hanmail.net

논문 투고일: 2026년 8월 15일

1차 심사 완료일: 2026년 9월 1일

2차 심사 완료일: 2026년 9월 8일

게재 확정일: 2026년 9월 15일