

바벨탑 3.0의 시대, 인공지능과 번역을 논하다*

이상빈(한국외국어대학교)

1. 바벨탑 3.0의 시대

아랫글은 김상균 교수가 『초인류』에서 제시한 미래 사회의 모습이다.

레슬리는 한국어로 발표를 진행했다. 물론 블랙스버그에 모인 프로젝트 멤버들에게는 각자 자신이 선호하는 언어로 발표가 들렸다. 레슬리가 한국식 농담을 던지면, 멤버의 문화권, 연령 등을 바탕으로 자동으로 이야기가 일부 각색이 되어 전달되었다. 예전과 달리 멤버들의 얼굴에 미소가 가득한 모습을 보면서 레슬리는 더 신나게 발표를 마칠 수 있었다. (김상균, 2023, p. 128)

공상과학소설에서나 나올법한 이 장면이 관심을 끄는 이유는 인공지능이 맥락을 파악하여 “각색[번안]”까지 한다는 데 있다. 한국식 농담이 번안되어 의도된 반응을 이끌어 낸다?!¹⁾ 정말 이런 미래가 올까? 4~5년 전만 하더라도 가볍게 넘겼을 장면이지만, 챗GPT 광풍이 불고 있는 요즘에는 이

* 이 글의 일부는 한국외대 통번역연구소 학술대회(2024. 7. 3.)에서 소개되었다. 이 글에서 제시하지 못한 인공지능 관련 논의는 필자의 다른 에세이에서 확인할 수 있다. 지면을 빌려 편집위원장과 심사위원들께 감사의 뜻을 표한다.

1) 화행이론(speech act theory)에서 말하는 발화 행위(locution), 발화 의도(illocution), 발화 효과(perlocution)가 모두 달성된 상황이다(Austin & Urmson, 1962).

런 시나리오를 믿을 만한 근거가 충분해 보인다. 적어도 일반 대중에게는 그런 것 같다.

인간은 언어 간 장벽을 넘기 위해 몇천 년 동안 각고의 노력을 해왔다. 이런 노력의 중심에는 늘 번역가²⁾가 있었고, 그들의 노동은 보상의 종류와 관계없이 역사적으로 그 가치를 인정받았다. 20세기 들어서는 번역을 자동화하려는 시도가 시작됐다. 이러한 노력의 선봉에는 영국, 미국 등을 비롯한 일부 선진국 정부와 혁신을 내세운 대학 연구팀이 있었다. 하지만 최근까지도 이들의 활동은 우리 일상을 획기적으로 바꿀만한 결실로 이어지지 못했고, 대개는 그저 실험실 결과로만 남게 되었다. 그래서 번역가는 여전히 절대적인 위상을 누렸다. 바로 바벨탑 1.0 시대의 모습이다.

그러다 바벨탑 2.0의 시대가 열렸다. 일부 IT 기업이 대량의 말뭉치를 활용하여 자동기계번역기(FOMT)를 선보였고, 이제 일반인도 무료로 기계번역의 가능성을 직접 확인할 수 있는 시대가 온 것이다. 하지만 기계번역의 수준은 여러 논문과 기사 등에서 드러났듯이 우스꽝스럽거나 형편없을 때가 많았다. 일부 연구자는 기계번역의 품질을 인간번역(human translation)과 비교하면서 “기계번역의 시대는 멀었다”라고 선언했다. 그럼에도 불구하고 일반인들 사이에서는 기계번역 사용이 증가했고, 특정 분야에서는 그 효용이 분명하게 드러났다. 일단 속도 면에서는 인간이 기계를 따라갈 수 없었고, 특정 언어 쌍이나 사용 환경에서는 조잡한 수준이라도 기계번역이 큰 도움이 되었다. 이처럼 바벨탑 2.0의 시대에는 기계번역에 대한 불신과 신뢰가 공존했다. 사람들은 이 탑이 무너질 수 있다고 생각하면서도 최소한 어느 높이까지는 충분히 올라갈 수 있다고 믿었다.

이제 바벨탑 3.0의 시대가 도래한 것일까? 2022년 챗GPT가 세상에 모습을 드러내면서 일반 대중도 인공지능의 엄청난 잠재력을 두 눈으로 직접 확인할 수 있었다. 언어학자들도 챗GPT의 ‘언어 구사력’에 놀라움을 금치 못했고, 일부 IT 전문가들은 인공지능이 인류의 미래에 큰 위협이 될 수 있다고 경고했다. 기술적 숭고함(technological sublime)이 인간의 마음속에 경이와 공포를 동시에 불러일으킨 것이다. 다만, 이 글의 독자가 관심 있어 할 번역만을 놓고 보면, 생성형 인공지능이 기존의 신경망 기계번역을 크게 넘어서

2) 이 글에서 “번역”은 통역의 개념을 포함할 때가 많다.

지는 못하고 있다. 하지만 생성형 인공지능은 번역 과정에 도움을 줄 수 있는 각종 교정 교열과 원문 분석 등을 지원한다. 게다가 바로 이 순간에도 학습과 진화를 거듭하면서 그 잠재력과 능력을 끊임없이 키워가고 있다. 그래서 생성형 인공지능의 시대(이하 “인공지능 시대”)에는 기계번역을 조롱하는 분위기가 생겨날 틈이 없다. 기계번역에 대한 신뢰나 의존이 기계번역에 대한 불신을 조금씩 대체하기 시작한다.

인공지능 시대가 본격적으로 막을 올리면서 기존의 통번역 전문가와 교육자들이 큰 고민에 빠졌다. 빠르게 확장 중인 인공지능 바벨탑을 보면서 교육 철학과 노선을 수정해야 하고, 인공지능 발전에 따른 번역(가)에 대한 인식 변화에도 대응해야 하기 때문이다. 이제 기존의 통번역 전문가들은 <그림 1>의 ‘인공지능 바벨탑’(오른쪽 그림) 앞에 모여 있는, 시대에 뒤떨어진 사람처럼 보일 처지다.

그림 1

Pieter Bruegel the Elder의 바벨탑(1563)과 DALL·E로 만든 인공지능 바벨탑(2024)



이 글에서 나는 인공지능 시대에 번역장(translation field)이 어떤 변화를 겪고 있는지, 나아가 번역가와 번역 교육자들이 어떤 도전에 직면해 있는지를 다각도로 살펴보고자 한다.⁴⁾ 이를 위해 먼저 신기술의 발달 정도를 보여

3) 피터르 브뤼헬 1세의 바벨탑(왼쪽 그림)을 이용해 오른쪽 그림을 제작하였다. 엄청난 규모의 데이터 센터와 비교되도록 그림 속 사람들을 비슷하게 만들었다.
 4) 이 글은 비문학 번역에 초점을 맞춘다. 이에 대해 한 심사위원은 다음과 같은 논평을 제시하였다. “AI의 등장으로 어찌 보면 통역과 번역의 경계도, 문학번역과 비문학 번역의 경계도 무너지고 있다 [...] 개인적으로 문학번역이 비문학 번역보

주는 몇 가지 사례를 소개한 후, 기술 진보로 인해 번역 전문가들이 어떤 상황에 놓여 있는지를 생각해 본다. 그다음에는 인간과 기술의 상호관계에 주목하여, 기계번역의 발전으로 대두된 번역가의 실존적 가치와 번역기계의 주체성 행위성 문제를 다룬다. 이어서는 인공지능 시대에 번역 역량의 의미를 따져보고, 번역 교육이 직면한 인지적 상황을 간략히 논한다. 끝으로는, 미래의 번역가와 번역의 운명을 상상해 보며 이 글을 마무리한다.

이 학술 에세이의 목적은 인공지능 시대에 번역 전문가가 처한 현실을 진단하고 미래를 예측해 봄으로써 적절한 대응 방안을 모색하기 위한 담론의 장을 마련하는 것이다. 미리 선언하는데, 이 글에서 나는 어떤 문제에 대한 해결책을 제시하지 않는다. 나의 궁극적인 목적은 나와 비슷한 처지에 있는 다른 전문가들이 급변하는 현 상황을 어떻게 인식하고 대응 중인지를 논의할 수 있게끔, 하나의 계기를 마련하는 데 있다.

2. 기술 발전의 속도와 번역

우리는 데이터 시대에 살고 있다. 스마트폰 사용을 비롯한 일상의 거의 모든 활동이 데이터로 만들어지고, 이렇게 생산된 데이터는 또 다른 데이터를 대량으로 산출한다. 데이터의 생산 규모는 일반인이 상상하기 힘든 수준이다. 포츠머스 대학의 Melvin Vopson 교수는 디지털 데이터의 양이 현재와 같은 속도로 증가하면 비트(bit) 수가 지구의 원자 수를 초과할 수 있다고 전망했다. 그는 최신 데이터 메모리 기술에서 1비트가 대략 25제곱 나노미터를 차지한다고 가정한 후, 약 150~200년 후에는 디지털 데이터가 지구의 전체 부피보다 더 많은 공간을 필요하게 된다고 추산했다(김성태, 2023, p. 86). 이처럼 데이터가 지수 함수의 곡선처럼 확장하면, 챗GPT를 능가하는 신기술도 머지않아 등장할 수 있다.

데이터를 처리하는 기술 역시 무섭게 발전하고 있다. 인공지능의 성능을

다 안전지대에 있다고 생각하지 않는다. ‘문학번역’이라는 개념은 사실 매우 모호하고 그 경계가 불분명하다(Cronin, 2013).” 문학번역과 기계번역에 관한 또 다른 관점은 전미세리(2023)에서 확인할 수 있다.

높이기 위해서는 몇 가지 기술적 전제가 필요한데, 그중에는 고도의 컴퓨팅 능력을 갖추는 것도 포함된다. 사실 나 같은 일반인도 이미 엄청난 컴퓨팅 파워를 경험하고 있다. 우리가 사용하는 스마트폰의 컴퓨팅 능력, 메모리 용량 등은 인류 최초로 달 착륙에 성공한 아폴로 11호의 컴퓨터(AGC)보다 우수하다. 인간 대다수가 이미 첨단기술의 선봉에 있다고 해도 과언은 아닌 것이다. 이런 상황에서 게임 체인저가 될 수 있는 양자컴퓨팅 연구도 활발히 진행 중이다. 구글 인공지능 퀀텀팀을 이끈 John Martinis UC샌타바바라대 교수 등은 『Nature』에 게재한 논문에서 “현존 최고 슈퍼컴퓨터로 1만 년 소요되는 연산을 양자컴퓨터로 200초 만에 해결했다”라고 주장한다(김용주, 2019). 이것도 자그마치 5년 전의 일이다.

통번역 업계도 기술 발전의 직접적인 영향을 받고 있다. 특히 거대언어 모델(LLM)의 도입은 우리 삶에 많은 변화를 줄 수 있다. 우선 통번역 자동화를 일으켜 외국어 역량의 중요성을 더욱 감소시킬 것이다(카이스트 문술미래전략대학원 미래전략연구센터, 2023, p. 109). 일부 전문가는 인공지능이 아무리 발전해도 번역에 미치는 영향에는 한계가 있다고 지적한다. 가령 기계는 문학 번역가를 대체할 수 없고, 다른 유형의 번역에서도 반드시 인간의 개입이 필요하다고 주장한다. 그러나 이보다 더 실질적인 문제가 있다. 바로 번역에 대한 사람들의 인식이다. 외국어를 모르는 사람도 이문화 의사소통에서 불편함을 못 느끼면 번역 자체를 중요하게 생각하지 않을 것이다. 즉, 노력 휴리스틱(effort heuristic)⁵⁾에 따라 번역을 인간이 힘써야 할 분야로 보지 않는다는 뜻이다. 주변을 둘러보자. 인공지능이 거의 모든 일을 다 해 줄 수 있을 것처럼 묘사되지 않던가? 하물며 번역이라고 예외로 둘까? 기술적 특이점(technological singularity)이 실제로 온다면 인간의 인지 활동 중 가치 있는 일이 뭐가 있을까 싶다.

기술적 특이점은 인공지능이 모든 인간의 지능보다 뛰어난 초인공지능으로 성장하는 시점이다. 『특이점이 온다』(The Singularity Is Near)의 저자 Ray Kurzweil은 특이점의 도래를 2045년 전후로 예측한다.⁶⁾ 이 용어가 제법

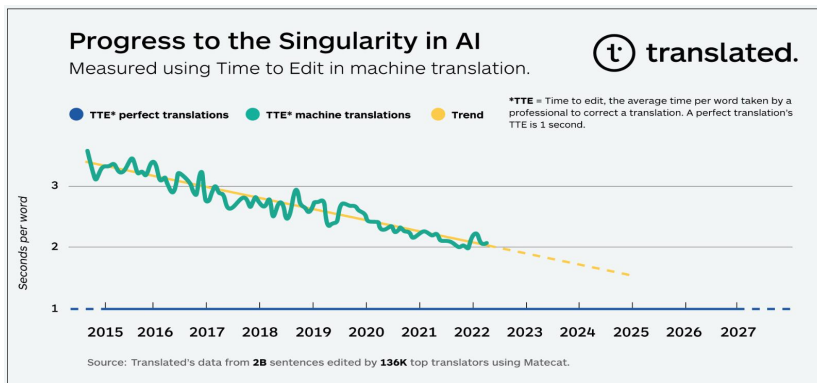
5) 노력 휴리스틱은 개체의 품질이나 가치를 그 개체의 생산에 들어간 노력에 따라 판단하는 정신적 경험 법칙이다. 노력 휴리스틱에 따르면 제작에 더 많은 시간이 소요된 물건을 더 가치 있는 것으로 판단한다.

유명해졌을 때, 문득 이런 생각이 들었다. 번역의 기술적 특이점도 존재할까? 번역의 특이점은 모든 인간보다 번역 능력이 뛰어난 기계번역이나 인공지능이 도래한다는 뜻이다. 이런 정의는 이 글 도입 부분에서 언급한 레슬리의 상황, 혹은 언제 봤을지 모를 SF 영화의 한 장면을 떠오르게 한다. 그럼에도 불구하고 한편으로는 가능하지 않을까 싶은 생각도 든다. 통번역 실무 공부를 한창 하던 2000년대 초로 돌아가 나 자신에게 챗GPT와 신경망 기계번역을 소개해 준다면, 과거의 나는 어떤 반응을 보일까?

라이트 형제가 1903년에 만든 비행기는 59초를 비행했다. 지금 기준으로 생각하면 비행이라기보다는 허공에 잠시 머물렀다고 보는 게 맞을 듯싶다. 하지만 1909년에는 Louis Blériot가 영국 해협을 가로지르는 비행에 성공했고, 1920년대에는 상업용 항공기도 등장했다. 특히, 라이트 형제가 비행에 성공한 지 약 50년 뒤에는 상업용 제트 여객기의 대중화를 이끈 보잉 707이 탄생했다. 라이트 형제가 노스캐롤라이나 키티호크에서 비행기를 날릴 때 이런 변화를 예상이나 했을까? 위와 같은 진보의 사례는 수없이 많다. 딥마인드가 2010년에 설립됐으니, 그로부터 50년 뒤에는 어떤 시대가 열릴지 상상해 보자. 수익 가속화의 법칙(Law of Accelerating Returns)⁷⁾을 고려하면 새로운 시대는 훨씬 더 앞당겨질 수 있다.

그림 2

기술적 특이점을 향한 진전(Translated, 2022)



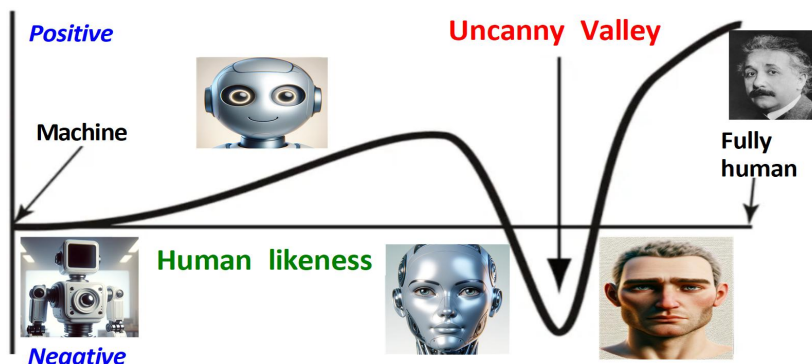
- 6) 특이점의 도래 가능성과 시점에 대한 이견도 많다.
- 7) 무어의 법칙(컴퓨터 칩의 트랜지스터 수가 약 2년마다 두 배로 증가)을 기초로 한, 커즈와일의 개념이다. 기술의 발전이 기하급수적으로 일어난다는 뜻이다.

번역의 특이점은 나만의 생각이 아니다. <그림 2>는 한 업체가 발표한 관련 자료인데, 여기서 우리는 특이점에 관한 실마리를 얻을 수 있다. TTE(Time To Edit)는 전문번역가가 번역을 수정할 때 소비하는 총시간을 ‘단어당 평균 시간(초)’으로 환산한 것이다(따라서 수정이 필요 없는 완벽한 번역은 TTE가 0이다). 이 그림에서 “기계번역 TTE”는 2015년 기준, 3초를 조금 웃돈다. TTE는 해를 거듭하면서 감소세를 보였고 2022년에는 2초 지점에 도달했다. 이런 추세가 지속되면 기계번역 TTE는 1초대로 진입하고 중국에는 0에 가까워질 수 있다. 물론 현재로서는 TTE가 0에 수렴하는 시나리오를 쉽게 상상하긴 어렵지만, TTE가 0.2~0.3초 이하만 되더라도 특이점을 거론할 수 있을 것 같다.

번역 기술의 혁신은 일반 사람들에게 희소식이겠으나, 번역가나 번역 교육자들에게는 우울한 소식이 아닐 수 없다. 어쩌면 우리 전문가들은 이미 ‘불쾌한/섬뜩한 골짜기’(uncanny valley)에 빠져 있는 걸지도 모른다. 불쾌한 골짜기란 일본의 로봇공학자 마사히로 모리(森政弘)가 제시한 용어로, 로봇의 진화와 그에 대한 인간의 대응을 압축적으로 묘사한 개념이다. <그림 3>이 보여주듯이 개발 초기 단계의 로봇은 인간과 비교해 여러 면에서 부족하다. 그러다 로봇에게 적절한 외모와 능력이 갖추지면, 인간은 그런 로봇을 호감 있게 바라본다. 하지만 로봇이 인간과 너무 흡사해지면 오히려 불쾌감이나 거부감을 일으킨다. 몇몇 사람들은 이런 로봇 때문에 정체성의 혼란을 경험하고, 심지어는 실존적 위협까지 느낄 수 있다.

그림 3

불쾌한 골짜기(DALL·E의 도움을 일부 받아 제작)



기계번역에 관한 초기 번역학 연구를 떠올려 보자. 연구자들이 기계번역의 ‘순수함’을 얼마나 높였는가? 기계번역의 효율성은 논외로 한 채, 기계번역의 오류(오역)에 집중하고 인간번역의 우수성을 피력한 연구 말이다. 하지만 실제 기계번역을 사용하다 보면 감탄은 아니더라도 수긍하는 경우나 편리함을 느낄 때가 적지 않다. 어떤 용도로든 최근에 기계번역을 사용해 본 사람이라면 기계번역의 진화를 가볍게 넘기진 못할 것이다. 그러니 한 차원 높은 인공지능 번역이 등장하면서 번역 실무자나 교육자들이 불편해하는 것은 어찌 보면 당연하다. ‘불쾌하다’라는 말을 일차적 의미로만 해석하지 말자. 번역가나 번역 교수자가 인공지능의 발전에 어떻게 대응해야 할지 모르고 있다면, 이런 상황 또한 ‘불쾌한 골짜기’를 표상하는 것이다.

인공지능의 발전이 심화할수록 통번역 전문가의 고민도 그만큼 깊어진다. 챗GPT와 같은 ‘도구’를 적극적으로 받아들이기만 하면 될까? 번역 노동의 가치가 떨어지는 상황을 감내하고 노동 환경의 변화에 적응해야만 할까? 초인적 지식을 대가로 기존 정체성을 포기한 ‘新 파우스트’가 되도 괜찮을까? 솔직히 대안이 있는지 모르겠다. 그렇다고 기계번역을 완강히 거부하는 ‘新 러다이트(neo-Luddite)’로 남을 수도 없다. 통번역 전문가라면 두 극단의 어느 한 지점에서 자신만의 생존법을 찾고 있을 것이다. 어떤 방법을 선택하든, 한 가지 사실은 분명하다. Yuval Harari가 말했듯, 인공지능과 함께 살아가야 할 우리에게 변화만이 유일한 상수이다.

3. 번역의 주체(들) — 누가 주인인가?

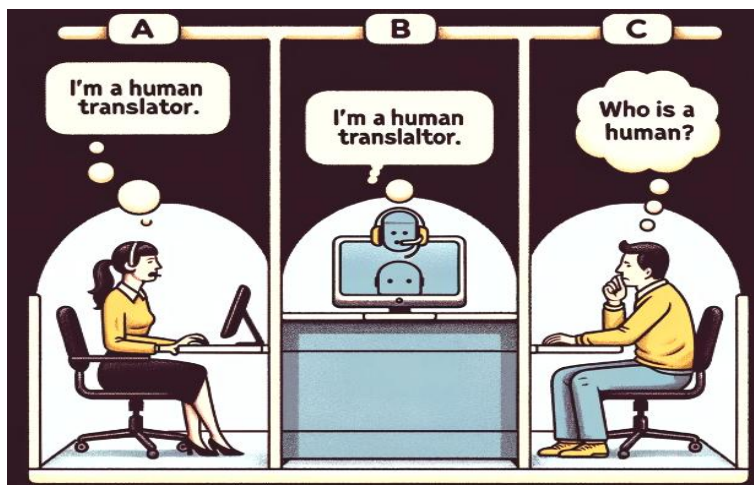
기계번역이 지금보다 훨씬 더 좋아져 인간번역과 구분할 수 없다면 어떻게 될까? 조금은 엉뚱하지만, 튜링 테스트(Turing Test)⁸⁾의 핵심 논리를 번

8) 이 테스트에는 한 명의 질문자와 두 명의 응답자(컴퓨터와 인간)가 참여한다. 질문자는 어느 쪽이 컴퓨터인지 모르며, 응답은 오직 키보드를 통해서만 이루어진다. 질문자가 컴퓨터와 인간 응답자를 구별할 수 없다면, 컴퓨터는 테스트를 통과한 것으로 간주한다. 이 테스트는 기계가 인간과 얼마나 유사하게 대화할 수 있는지를 기준으로 기계의 지능을 판별하려는 것이며, ‘이미테이션 게임’으로도 알

역 상황에 적용해 보자. <그림 4>는 ‘번역의 튜링 테스트’를 상징적으로 보여주기 위해 DALL E로 제작한 그림이다. 이 그림에서처럼 기계가 산출한 번역물과 인간이 만든 번역문⁹⁾을 비교하게 한 후 둘 간의 구분이 어렵다면, 기계의 번역지능은 인간의 번역지능과 비견할 만한 수준이 된 것이다.¹⁰⁾

그림 4

기계번역 ‘튜링’ 테스트(DALL·E에서 프롬프트로 제작)



물론 이 같은 사고 실험을 진행하기는 당분간 어렵다. 기계번역이 아직까진 전문번역가의 수준에 미치지 못하기 때문이다. 그렇다면 상황을 조금만 바꿔보자. 인간번역과 기계번역을 있는 그대로 비교하지 말고, 인간이 교정한 기계번역 포스트에디팅(MTPE) 산물과 인간번역을 비교하는 것이다. 만일 인간의 손을 약간 거친 포스트에디팅 결과물과 순수 인간번역을 구분할 수 없다면 어떨까? 이런 실험은 제한적으로나마 현재도 가능하리라 본다. 만일 라이트 포스트에디팅(light post-editing)의 수준이 인간번역과 구분

려져 있다. 1950년 앨런 튜링이 「Computing Machinery and Intelligence」라는 논문에서 처음 제시하였다.

9) 국어사전에서 ‘번역문’(翻譯文)과 ‘번역물’(翻譯物)의 뜻은 거의 비슷하다. 이 문장에서는 한자 ‘물(物)’의 기본 뜻에 초점을 맞춰 기계가 만든 번역을 번역물로 표현한다.

10) <그림 4>에는 철자 오류가 있다. 쉽게 찾을 수 없는, 기계의 오류를 상징적으로 보여주기 위해 의도하였다.

할 수 없을 만큼 좋다면, 번역기계는 ‘튜링 테스트’를 통과했다고 보는 것이다. 물론 여기에는 한 가지 전제가 있다. 포스트에디팅 결과물의 주인을 기계로 본다는 점이다.

번역 주체의 문제는 그리 간단하지 않다. 포스트에디팅을 거친 번역은 궁극적으로 인간의 산물인가, 기계의 산물인가? 기계든 인간이든, 어느 정도 참여해야 진정한 주체가 될 수 있나? 기계번역을 활용하는 번역가를 이제는 ‘교열자’로 불러야 할까? 인간이 포스트에디팅을 자신의 산물로 주장한다면, 그 반대 논리도 타당하지 않을까? 다시 말해, 인간번역을 챗GPT와 같은 인공지능 ‘도구’로 교정하면, 인공지능도 그 텍스트의 주인이 될 수 있다는 뜻이다. 인공지능이 저작권을 가질 수 없다곤 하지만, 오직 인간만이 ‘번역자’¹¹⁾가 될 수 있다는 주장은 논리적으로 타당하지 않다.

기술 수행(techno-performance)은 기본적으로 다음과 같이 몇 단계로 구분할 수 있다.

- 0단계: 인간은 기술의 도움 없이, 독립적으로 수행한다.
- 1단계: 인간이 주도적으로 수행하고, 기술은 보조적 역할만 한다.
- 2단계: 인간과 기계가 동등한 수준으로 수행한다.
- 3단계: 기술이 주도적으로 수행하고, 인간은 보조적 역할만 한다.
- 4단계: 기술이 인간의 도움 없이, 독립적으로 수행한다.

여기서 “기술”은 우리가 쉽게 떠올리는 첨단기술만을 지칭하지 않는다. 신석기 시대의 뾰족한 돌도 인류에게는 첨단(尖端)을 자랑하는 기술이었다. 따라서 0단계는 오늘날 우리 관점으로는 불가능한 상황이다. 인간은 아주 오랫동안 1단계 상황에서 삶을 영위해 왔다. 지금도 상당수의 번역가가 인터넷과 사전만으로 번역하며, 이때도 기술은 보조적 역할만을 수행한다. 2단계에서는 인간과 기계가 비슷한 수준으로 수행한다. 다만 2단계의 경우, 구체적인 사례를 떠올리는 것보다 상황을 개념적으로만 이해하는 게 편할 수 있다. ‘동등하다’라는 게 양적으로 같다는 것인가, 아니면 질적으로 비슷하다는 것인가? 어떻게 동등함을 평가할 것인가? 번역과 관련하여 3단계의 전

11) 번역가와 번역자를 구분하고 싶다. 기계가 번역가는 될 수 없어도 번역자는 될 수 있다는 게 나의 판단이다. 이에 대해서는 다른 에세이에서 논하겠다.

형적인 사례는 일부 포스트에디팅이나 컴퓨터 보조 번역(CAT)이 활용되는 상황이다. 여기에도 문제는 있다. 인간의 개입이 양적으로 적다 하더라도 인간의 수행을 절대적으로 보는 시각이 있다. 즉, 포스트에디터는 기존의 번역가와 마찬가지로 절대적인 위치에 있다고 보는 것이다. 마지막으로, 4단계는 기술이 더욱 발전하여 인간의 개입이 불필요한 상황을 가정한다. 사실, 4단계의 초기 형태는 지금도 존재한다. 인터넷에서 자신이 모르는 외국어를 개략적으로 이해할 수 있는 것은 인터넷 브라우저에 자동번역 기능이 있기 때문이다. 일단 브라우저에서 간단한 설정만 해주면 그다음부터는 기계가 모든 번역을 알아서 해준다. 미래에는 이처럼 인간과 기계 사이에 탈매개화 현상이 심화하여 기계의 단독 수행이 좀 더 보편화될 수 있다.

이제 기계는 단순한 타자, 즉 아더(other)가 아니라, Shoshana Zuboff 등의 표현을 빌자면, 빅 아더(big other)이다. 미래에는 기계가 초행위 주체성(hyper-agency)을 얻게 될 것이다. 즉, 매우 발달한 알고리즘 덕분에 인간의 생각이나 의도 등을 건너뛸 수 있는 비인류 발생론적(non-anthropogenic) 방식¹²⁾이 번역 실무의 밀바탕이 될 수도 있다. 이때 인공지능은 번역가에게 실존적 위협이 된다. 기계가 번역의 주체가 된다는 것은 단순히 일자리를 빼앗는다는 뜻이 아니다. 번역장에서 인간이 주변부로 밀려나는 상황을 감당할 수 있는냐의 문제다.

8년 전 알파고-이세돌 대국 — 챗GPT를 생각하면 당시 알파고는 장난 수준이었다 — 직후, 나는 한 논문에서 다음과 같이 언급한 바 있다.

번역기계는 팬 번역가, 전문번역가 등과 교류하는 거대 행위자(macro-actor)로 성장할 것이다. 지금까지의 번역기계는 번역가의 완전한 지배를 받는, 단순한 툴로써만 인식됐다. 하지만 가까운 미래의 번역기계는 번역가가 참고할 수도 있는 보조도구가 아니라 번역가가 반드시 동원해야 할 행위자로 기능할 것이다. 특히 게임에서와 같이 제한적 형태의(문학작품 등과 비교해 봤을 때 ‘제한적인’) 번역이 이루어지는 영역에서는 궁극적으로 인간의 번역과 기계의 번역에 차이를 두기가 점점 더 어려워질 수도 있다. 이 경우 번역가는 번역 초기 단계에서부터 완성 단계에 이르기까지 번역기계와 밀접하게 교류할 수밖에 없

12) 알고리즘에 의해 인간의 생각과 의도를 건너뛰는 절차, 방식 등

고, 그러한 과정을 겪은 번역기계는 (번역 데이터의 누적으로) 점점 더 완성도 높은 번역을 산출할 수 있다. (이상빈, 2016, p. 130)

윗글을 썼을 당시, ‘인공지능’ 하면 대개 바둑이나 체스만을 떠올렸고, 번역기계는 종종 놀림감 정도로 취급받았다. 당시만 해도 챗GPT와 같은 생성형 인공지능은 상상조차 할 수 없었다. 위 인용문에서 나는 기계의 수행가능성을 높이 평가하며 인간과 기계가 동등한 행위자로서 수행할 수 있다고 말했다. 이제 이 글을 다시 읽어보니, 어쩌면 기계가 “[인간과] 교류하는 거대 행위자”임을 넘어, 인간과 교류하고 때로는 인간을 넘어서는 존재가 될 수 있겠다는 생각이 든다. 만일 그런 상황을 기정사실로 간주한다면, 번역가는 어디서 자신의 실존적 가치를 찾아야 할까? 실존적 가치가 죽으면 그 직업도 죽은 것이나 다름없다.

데카르트는 이성 중심의 자아를 중시했다. 자아의 본질을 ‘사고하는 나’에서 찾았고, 따라서 ‘나는 생각한다. 고로 나는 존재한다’(Cogito, ergo sum)라는 명제를 지식의 출발점으로 삼았다. 이런 사고방식이 그간의 인간 번역가를 지탱해 왔다. 즉, ‘나는 생각한다. 고로 나는 번역한다’(Cogito, ergo transfero)였다. 번역의 본질과 주체는 인간의 블랙박스를 가정하지 않고서는 논할 수 없었다. 그런데 최근 다른 ‘블랙박스’가 나타났다. 바로 챗GPT와 같은 생성형 인공지능이다. 이 신종 블랙박스¹³⁾에게 번역은 어떤 의미가 있을까? 번역의 실존적 개념과 이론적 틀이 기계에게 의미가 있을 리 만무하다. 예컨대 발터 벤야민의 계생(繼生, *Fortleben*),¹⁴⁾ 조지 스타이너의 해석학적 운동(Hermeneutic Motion) 등이 ‘철학 하지’ 않는 기계에게 무슨 소용이 있을까? 기계번역은 추천과 클릭 그리고 알고리즘으로 대체할 수 있는 확률적 등가일 뿐이다.

그렇다면, 사유하지 못하는 인공지능은 번역의 실질적 주체가 될 수 없

13) 인공지능 전문가들도 챗GPT가 어떤 과정을 통해 결론을 도출하는지 설명하지 못한다. 이런 점에서 생성형 인공지능도 (인간의 블랙박스보다 더 캄캄한) 블랙박스로 묘사할 수 있다.

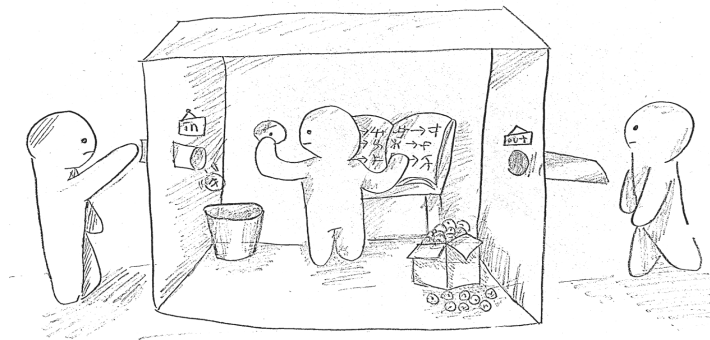
14) 이 독일어는 당초 *afterlife*로 번역되어, 국내 연구자들이 오랫동안 ‘사후세계’, ‘내생’, ‘후생’ 등으로 썼다. *afterlife*가 오역이므로 새로운 용어를 제안한다. 계(繼)가 독일어 *Fort*-(*forth*)의 의미를 거의 모두 담고 있다.

을까? 말하자면, 인공지능은 번역 도구 그 이상도, 그 이하도 아닌 걸까? 인공지능 시대에 번역 실무자나 교육자가 고민해할 첫 번째 문제가 바로 기계의 주체성 문제이다. 주체성 문제에 어떤 답을 내놓느냐에 따라 그다음 방향이 결정되기 때문이다.

4. 외국어/번역 역량과 교육의 문제

외국어를 못해도 번역을 할 수 있을까? 인공지능 시대라면 완전히 허황된 질문은 아니다. 먼저 <그림 5>에 나타난 중국어 방 논증(“Chinese Room” argument)을 생각해 보자.

그림 5
중국어 방(“Chinese room”)



중국어 방 논증은 철학자 John Searle이 고안한 사고 실험이다. 방 안에 중국어를 모르는 참가자가 들어가서 중국어 질문과 그에 대한 응답이 적힌 매뉴얼을 받는다. 또한 다른 사람과 소통하기 위한 필기도구도 함께 받는다. 심사관이 방 안으로 중국어 질문을 보내면, 참가자는 매뉴얼을 참고하여 올바른 중국어 답변을 작성한다. 방 밖에 있는 관찰자는 이 참가자가 중국어를 구사할 수 있다고 생각하겠지만, 사실 참가자는 중국어 질문과 답변을 이해하지 못한 채, 매뉴얼에 따라서만 답변을 작성한다. 중국어 방 논증은 이 실험을 데이터 처리 과정에 비유한 것이다. 참가자가 중국어 질문에 적절한 답변을 할 수 있어도 그가 중국어를 이해했는지는 알 수 없다. 마찬가지로

지로 기계가 튜링 테스트를 통과해도 그것이 진정한 지능인지 단순한 모방 인지는 확실치 않다는 게 이 논증의 주장이다.

비슷한 맥락에서 William J. Rappaport는 ‘한국어 방’을 소개했다. 한국에 사는 한 영문과 교수는 영어를 읽지도 쓰지도 못한다. 그는 한국어로 번역된 셰익스피어 작품을 읽고 셰익스피어에 관한 논문을 썼는데 이 논문이 영어로 번역되어 세계적인 인정을 받았다. 만일 이 과정에서 기계번역이나 인공지능의 도움을 받았다면, 우리는 이 교수를 어떻게 받아들여야 할까?

번역을 제대로 하려면 해당 언어 쌍, 나아가 언어 전반에 대한 지식과 이해가 필요하다. 우리는 외국어 구사 능력이 없는 사람을 번역가라 칭하지 않는다. 영어 실력이 형편없는 영문학자를 상상하기 어렵듯 말이다. 그러나 인공지능 시대에는 이런 전통적인 시각이 조금씩 흔들리고 있다. 외국어에 대한 이해가 부족해도 번역은 할 수 있다는 소위 ‘결과 중심론적 시각’이 확산하고 있다. 나는 일본어를 전혀 할 줄 모른다. 하지만 최근 구글과 딥엘(DeepL)의 도움을 받아 일본어 번역문을 분석하고 논문을 작성한 적이 있다. 마치 ‘일본어 방’에 있는 실험 참가자처럼 말이다. 학생들도 이처럼 자신만의 ‘방’에 있는 걸까? 그러면 외국어 교육은 근본부터 흔들린다.

생성형 인공지능의 발전으로 외국어 교육이 심각한 도전을 받고 있다. “외국어, 배워서 뭐 해”와 같은 자극적인 기사 제목도 심심찮게 보인다. 어학 전공의 필요성에 대한 자조 섞인 농담도 들린다. 이런 상황에서 교육자인 우리는 언어와 번역의 가치를 어떻게 피력해야 할까? 번역은 단지 외국어를 뛰어나게 잘한다는 증거가 아니라, 다른 언어·문화권의 세계관을 이해하려는 인문학적 성찰을 의미합니다. 하이데거는 “언어가 존재의 집이다”라고 말했고, 비트겐슈타인도 “내가 아는 언어의 한계가 곧 내가 사는 세계의 한계다”라고 말했죠. 외국어의 중요성을 강조하기 위해 이런 잔소리를 한다 치자. 이런 말들이 요즘 학생들에게 제대로 전달될 수 있을까? 현시점에서 나 같은 사람이 가장 고민하는 문제는 외국어와 번역 교육의 기본 방향을 설정하는 일이다. 나를 포함해 주변의 많은 사람이 새로운 교육 방법을 시도하고 있고, 학교 당국자 또한 외국어 교육에 변화를 가하고자 여러 방안을 강구 중이다. 하지만, 내가 아는 한, 다수가 호응하는 교육 전략이나 전술은 없는 상태다. 기껏해야 외국어 학습의 규모를 줄이면서 특정 영역을

통폐합하고, 챗GPT 같은 도구를 접목하여 전통적인 주제를 다루는 게 보통이다. 그 어느 때보다도 다양한 교실 현장의 목소리를 듣고 새로운 교육 철학을 논해야 하는데, 교육자 대부분은 — 히브리어 ‘바벨’이 뜻하듯이 — 혼란에 빠져 있는 듯싶다.

걱정되는 점이 또 있다. 나는 대학생의 ‘번역’ 역량(competency)¹⁵⁾이 차츰 떨어질 거로 생각한다. 기계번역과 인공지능 도구에 익숙해지면서 번역을 전공하는 학부생조차도, 예전 같으면 할 수 있는 번역을 제대로 하지 못하고, 하더라도 상대적으로 많은 시간을 써야 할지도 모른다. 휴대 전화 덕택에(?) 전화번호 암기가 어려워지고 암기의 필요성조차 느끼지 못하듯이, 기계번역의 확산은 번역 역량의 감퇴로 이어지고 인간번역의 필요성을 사전에 차단해 버린다. 근본적으로는, 학생들의 외국어 역량이 줄어들 위기에 처했다. 요즘 국내 학생들은 중학교 때부터 영어를 공부한 나 같은 세대와 달리, 훨씬 이른 나이에 영어 공부를 시작한다. 그래서 그런지 문법적인 부분은 몰라도 학생들의 영어 구사력은 예전보다 훨씬 좋아졌다. 하지만 인공지능 도구가 활용되는 요즘 상황을 보면 학생들의 어학 역량이 지속될지 의문이 든다. 외국어 역량을 키우기 위해서는 외국어로 사고하는 습관이 필요한데, 번역기계는 그런 인지적 훈련의 기회를 줄이기 때문이다. 전화번호도 외우지 못하는 마당에 외국어의 세부 내용을 기억할 수 있을까?

문자 능력이 사라진 시대를 그린 구병모의 「오토 포이에시스」가 생각난다. 소설을 쓰는 인공지능 로봇 ‘백지’가 문명이 파괴된 먼 미래에서 깨어난다. 하지만 사람들은 더 이상 글을 쓰거나 읽지 못하고 구전의 형태로만 이야기를 주고받는다. 여기서 조금 엉뚱한 상상을 해보자. 수백 년 동안 기계번역에 의존해 온 인류가 인공지능의 거부로 인해 통번역을 직접 해야 한다면? 외국어에 관한 지식이 백지상태가 된 인류에게 어떤 상황이 펼쳐질까?

15) 역량(competency)은 능력(competence)과 달리, 오랜 기간에 걸쳐야만 강화할 수 있는 인지적 기반이다. 대학생들에게 필요한 ‘번역 역량’은 전문번역가에게 필요한 번역 능력이라기보다 그들의 경력을 발전시키고 삶을 풍요롭게 하는 데 필요한 핵심 외국어 역량에 가깝다.

5. 미래의 번역가

국내 번역학자라면 2017년 2월을 기억할 것이다. 세종대에서 경력 5년 이상인 전문번역가 4명이 모여, 구글 번역기, 네이버 파파고, 시스트란 번역기와 대결을 펼쳤다. 30점 만점인 이 경연에서 번역가들은 평균 24.5점을 획득한 데 반해, ‘(구세대) 인공지능’은 10점에 그쳤다. 언론은 이 경연과 관련해 “인공지능(AI) 번역기계와 인간이 벌인 ‘번역 대결’에서 인간이 압승하며 싱겁게 끝났다”라고 보도했다(김은정, 2017, para 1).

2017년 2월은 알파고 사건 때문에 일반인들도 인공지능의 힘을 실감하고 있던 때다. 번역 전문가들도 약간의 위기감을 느끼고 이런 대회를 준비했을 것이다. 하지만 이런 시도는 현재와 같은 챗GPT 시대에서는, 불가능하거나 매우 어색한 이벤트로 남을 공산이 크다. 번역 업체는 물론 공기업, 지자체도 인공지능 번역 쪽으로 사업을 적극 추진하고 있으며, 기계번역의 품질이 예전보다 크게 향상되었기 때문이다. 번역기가 좋아졌다고 해서 문학 텍스트로 번역 대결을 펼칠 수는 없다. 문학작품은 언어 사용 면에서 가장 인간적인 텍스트이므로 애당초 대결은 기계에 불리하다. 어떤 연구자는 인간 번역가의 우위를 주장하기 위해 독특한 원문을 활용하는데, 이는 불공정하다는 인상만을 줄 뿐이다.¹⁶⁾ 맥락, 문체 등이 복잡한 문학작품을 선택해 (이미 출판된) 전문번역가의 번역과 기계번역의 품질을 비교하면 승자를 예상하기란 어렵지 않다.

기술이 발전하면서 기계 번역자가 인간 번역가를 조금씩 대체하기 시작했다. 우리는 여전히 번역 노동을 수행하고 있지만, 그 가치가 조금씩 하락하고 있다. 급속도로 진행 중인 번역의 탈숙련화(de-skilling) 때문이다. 탈숙련화란 산업혁명 이후 인간의 노동이 이전만큼 전문성을 요하지 않고 인간의 노동이 점차 파편화되어, 기계의 보조적 역할에 그치는 현상을 말한다. 컴퓨터 보조 번역 수업이 기술에 관한 부분을 다룬다고는 하지만, 이런 학습 영역이 번역의 본질을 깊게 다루는 건 아니다. 즉, 번역 능력이나 역량을

16) 세종대 경연에서도 공정성 논란이 있었다. 가령, “[원문이] AI가 많이 학습하는 주요 시사 기사와 거리가 있었다 [...] 평가 배점도 AI의 특성을 무시한 채 너무 사람에게 유리하게 되어 있다” 등의 주장이 있었다(김태균, 2017. 2. 22.).

확장하기보다는 번역의 외적인 부분, 즉 기계 조작 능력에 집중하는 경향이 있다. 번역이 단순 기술 조작으로 치환되면 그때는 정말 번역가라는 직업에서 남는 것이 없을지도 모른다. 내가 우려하는 점은 미래의 ‘번역가들’이 반복적인 디지털 노동에 빠져 있는 모습이다. 번역 과정에서 느끼는 문제 해결의 기쁨을 누리지 못하고, 기계에서만 답을 뽑아내려는 무기력한 ‘번역가’ 말이다. Colin McGinn의 개념을 빌자면, 이런 번역가에게는 진정한 ‘내적 경험’¹⁷⁾이 없다. 충분한 내적 경험이 수반되지 않으면 임금이나 작업 효율도 더 떨어질 수밖에 없다. 그래서 나는 우려한다. 기술에 관한 교육이 자칫 번역의 내적 경험을 앗아가고 직업 환경과 조건도 훼손할 수 있음을.

미래의 번역가는 어떤 사람들일까? 앞서 잠깐 언급했듯이, 오늘날 우리가 생각하는 진정한 의미의 번역가는 아닐 수 있다. 원초적 형태의 기술을 활용하면서 텍스트와 장기간 씨름하는 그런 번역가는 문학번역 등 일부 영역에서만 제한적으로 존재할 것이다. 미래의 ‘번역가’는 교정 교열, 프로젝트 및 인적자원 관리 등을 주로 수행하는, 다른 의미의 전문가가 될 가능성이 크다. 마치 제조업 공정의 일부로, 대개 분업 하에서 일하며, 텍스트 생산을 보조하거나 대행하는 사람일 수 있다. 그래서 향후 ‘번역가’는 기계번역의 오류를 찾아 표시하는 주석자(error annotator), 작업의 전 과정을 조직하거나 감독하는 큐레이터(curator)/감독자(supervisor), 번역물의 상태나 품질을 인증하는 검증자 validator) 등의 역할을 모두 수행하거나 이 중 일부만을 수행하는 사람일 수 있다. 물론 다른 가능성도 있다. 가령 어떤 사람은 문학 작품을 포스트에디팅 하여 출판권을 따내고, 실제 번역의 필요성이 생기면 그때 전문번역가를 연결해 주는 업무를 할지도 모른다. 어떤 상황에서든 기계번역이 발전할수록 포스트에디팅을 수반하는 공정 프로세스는 더욱더 세분화할 것이다.

지금까지 언급된 내용을 토대로 “translation”의 어원(tranlatio)을 생각해 보자. 미래의 번역가는 ‘(힘들게) 강을 건너 물건을 가져오는 사람’으로 남을 수 있을까? 인공지능 시대가 본격적으로 열리면, 강을 건너는 실질적인 주체는 기계이고, 우리는 그저 물건만 배에서 싣고 내리는 하역부가 되는 건 아닐까? 방향타를 이리저리 바꾸며 노를 힘겹게 젓는 인간의 모습은 어

17) 심사위원의 말대로 Antoine Berman의 “체험으로서의 번역”과 상통한다.

디서 볼 수 있을까? 번역에 종사하는 사람들에게 비관적이고 마음을 짓누르는 비유일 수도 있겠다. 하지만 바로 지금이 번역가의 정체성과 미래를 솔직히 고민해야 할 시점이다.

참고문헌

- 김상균. (2023). 초인류. 웅진지식하우스.
- 김성태. (2023). 데이톨로지. 이른비.
- 김용주. (2019. 10. 24.). 구글 “슈퍼컴으로 1만 년 걸리는 문제, 양자컴으로 200초면 끝.” 전자신문. bit.ly/42fLNnr
- 김은정. (2017. 2. 21.). 바둑은 졌지만 번역은 이겼다... 인간 vs AI 번역대결서 인간 ‘압승.’ 조선일보. bit.ly/3vXXNOj
- 김태균. (2017. 2. 22.) 논란에 흔들리는 AI 인간 번역 대결... “평가 과정 불공정해.” 연합뉴스. <https://bit.ly/47biBAq>
- 이상빈. (2016). 번역기계, 팬번역가, 로컬라이저의 네트워크: 게임번역기 MORT를 통해 살펴본 게임번역 네트워크의 미래. 번역학연구, 17(3), 117-137.
- 전미세리. (2023). 기계 번역이 인간 번역을 대신하게 될까? 조의연, 이상빈. (공편), K 문학의 탄생: 한국문학을 K 문학으로 만든 번역 이야기 (pp. 174-190). 김영사.
- 카이스트 문술미래전략대학원 미래전략연구센터. (2023). 카이스트 미래전략 2023. 김영사.
- Austin, J. L., & Urmson, J. O. (1962). *How to do things with words*. Oxford University Press.
- Cronin, M. (2013). *Translation in the digital age*. Routledge.
- Translated. (2022, December 21). Translated measured the speed to singularity in AI using trends in machine translation. bit.ly/3ueQO30

AI and the future of translation: An analytical perspective

Sang-Bin Lee (sblee0110@naver.com)

Department of English for International Conferences and Communication, Hankuk
University of Foreign Studies

Abstract

The rapid advancement of artificial intelligence (AI) is poised to revolutionize various fields of expertise, including translation. This essay explores the evolution of the translation ecosystem in the age of AI, focusing on five key areas. First, I introduce the term *Tower of Babel 3.0* to describe the potential impact of GenAI on the public's perception of translation. Second, I discuss the implications of technological advancements for educators and practitioners in the field of translation. I suggest the notion of *translational singularity*, which signifies a transformative threshold where AI overtakes human intelligence in every facet of translation. Third, I investigate the *hyper-agency* of machine translators, a term that denotes their growing dominance and autonomy in the translation industry. This shift in the status of machines raises several concerns, including ethical dilemmas and the potential displacement of human translators. Fourth, I extend John Searle's "Chinese Room" metaphor to the present context, highlighting the rise of translators with limited proficiency in the target language. This metaphor is used to critically assess the essence of studying and performing translation. Finally, I examine the changing roles of human translators in an AI-driven world, focusing on how their responsibilities and skill sets may diverge from traditional expectations.

Keywords: Tower of Babel 3.0; artificial intelligence; ChatGPT; machine translation post-editing (MTPE); translation education

키워드: 바벨탑 3.0, 인공지능, 챗GPT, 기계번역 포스트에디팅(후편집),
번역 교육

이상빈(<https://orcid.org/0000-0003-2140-3266>)

한국외국어대학교 EICC학과 교수

sblee0110@naver.com

논문 투고일: 2024년 6월 21일

1차 심사 완료일: 2024년 8월 30일

2차 심사 완료일: 2024년 9월 9일

게재 확정일: 2024년 9월 14일