

학부번역전공자의 인간번역과 기계번역 포스트에디팅 품질 비교*

이선우** · 이상빈***

Sun-Woo Lee and Sang-Bin Lee (2021). Comparing human translation and post-edited machine translation: A case study of Korean undergraduate students. *The quality of machine translation outputs, including those of Google Translate and Naver Papago, has increased rapidly over the last few years. This development has raised the question of whether the current undergraduate translator training program is still relevant and feasible. Against this backdrop, a small-scale experiment was conducted in this study to assess the quality of human translation and post-edited machine translation undertaken by university students. The participants were two groups of 30 Korean undergraduates who attended a basic translation course in the spring semester of 2021. They were asked to translate two Korean texts, Text A (171 characters) and Text B (160 characters), both of which were extracted from relatively easy newspaper articles, into their B language (English) under two different conditions. One of the participant groups translated Text A for 10 minutes using machine translation outputs and translated Text B for 20 minutes without the help of any translation machine. The other group performed the same tasks, save that the source texts were assigned in reverse order. A total of 117 translations were collected and rated independently on a five-point scale by the two teacher-researchers. The results showed that the post-edited translations were much better than the human translations and that the human translations contained more errors in vocabulary and grammar. (Hankuk University of Foreign Studies, Republic of Korea)*

* 본 연구는 2021학년도 한국외국어대학교 대학혁신지원사업비 지원으로 수행되었음

** 한국외국어대학교 일반대학원 영어번역학과/EICC학과(제1저자)

*** 한국외국어대학교 EICC학과(제2저자)

Keywords: free online machine translation (FOMT), post-editing (MTPE), Google Translate, Naver Papago, translation quality assessment

주제어: 기계번역, 포스트에디팅, 인간번역, 구글/파파고 번역, 번역평가

1. 서론

2016년 딥러닝(deep learning)에 기반한 인공지능 기반 기계번역(Neural Machine Translation, NMT) 모델이 도입되면서 전 세계는 이른바 기계번역 시대를 맞이하였다. NMT는 기존의 통계기반 기계번역(Statistical Machine Translation, SMT) 모델과는 차원이 다른 비약적인 품질을 선보이며, 기계번역이 가까운 미래에 사람의 번역을 대체할 수 있을 것으로 전망되었다.¹⁾ 상황이 이렇다 보니 “앞으로 더 많은 데이터가 축적되고, 또 새로운 데이터가 생성되며, 딥러닝과 인공지능 기술이 발전하는 속도를 고려하면, 언어의 장벽이 사라지는 시대가 올지도 모를 일이다”라는 사회적 인식도 퍼지고 있다(김현정, 2021). 게다가 최근에는 COVID-19라는 세계적 팬데믹을 겪으며 물리적 제약이 없는 가상공간에서의 모임이 활성화되었고, 그로 인해 AI 번역은 언어 장벽을 해소하고 혁신적 소통에 기여하며 더욱 주목받고 있다.

번역을 둘러싼 사회적 분위기는 번역(학)에 대한 학술적 관심을 끌어올리기에 충분했다. 그 중에서도 기계번역 관련 연구는 번역학뿐만 아니라 교육학, 언어학, 문학, 공학 등 여러 분야에서 어느 때보다 활발하게 진행되고 있다. 일례로 국내외 외국어 교육(학) 분야에서 기계번역 관련 연구를 살펴보면 다음과 같은 뚜렷한 흐름이 파악된다. 외국어 교육 현장에서 구글 번역(Google Translate)과 같은 온라인기계번역(free online machine translation, FOMT)은 그 품질이 현재와 같은 수준에 도달하지 못한 시기에도 이미 교육적 차원에서 꾸준히 활용되고 있었다(Garcia & Pena, 2011; Groves & Mundt, 2015; Bahri & Mahadi 2016; 임희주, 2017). 그리고 AI의 성능이 획기적으로 개선된 최근에 와서는 기계번역을 활용한 외국어 교육을 시대적 변화에 따른 필수 과정으로 인식하는 경향이 있으며, 외국어 학습자가 기계번역을 적극적으로 수용할 수 있도록 기계번역의 긍정적인 측면을 부각시킨 연구가 많다(Briggs, 2018; Tsai, 2019; Ahn

1) 본고에서 기계번역은 구글 번역, 네이버 파파고와 같은 인공지능 기반의 온라인기계번역을 가리킨다.

& Chung, 2020; 조인희, 2018/2021; 김은영, 2021; 변길자 2021). 즉, 외국어 교육에서 기계번역은 교육적으로 효용 가치가 있고, 특히 문어(written text)와 연계한 외국어 학습 시에 유용한 도구로 자리잡고 있다.

한편 번역학계는 NMT가 도입된 이후에도 기계번역의 안정적이지 못한 품질과 기술의 한계로 나타나는 다양한 오류 유형에 큰 관심을 보여왔다(Tezcan et al., 2019; Haque et al., 2020; 박옥수, 2017/2018; 서보현과 김순영, 2018; 한현희, 2019a). 이 때문에 국내 학부 번역 교육 현장에서 기계번역을 긍정적으로 수용하고 공개적으로 다루는 경우는 아직까지 흔하지 않다. 하지만 기계번역 품질이 빠르게 개선되고 있고, 생산성과 효율성 측면에서 높은 경쟁력을 갖췄다는 점은 부정할 수 없는 사실이다. 이에 학부 차원에서 기계번역 활용 방안 및 실효성 등을 논의하고 검증하는 경험적 연구가 이목을 끌었고, 대부분 번역교육에서 기계번역을 어떻게 활용할 수 있을 지에 주목한다(김순미, 2017; 이상빈, 2017; 2018a; 2018b; 마승혜, 2018). 이와 더불어 기계번역 품질이 향상됨에 따라 기계번역 결과값을 사후에 인간이 수정하는 포스트에디팅(post-editing)과 기계번역 결과값을 최적화하기 위해 사전에 원문을 수정하는 프리에디팅(pre-editing)의 중요성이 부각되면서 관련 교육과 가이드라인 등이 다양한 차원에서 논의 및 제안되고 있다(Flanagan & Christensen, 2014; Hu & Cadwell, 2016; 윤미선 외, 2018; 한현희, 2019b; 남보리, 2021; 박건영, 2021).

이러한 사회적, 학문적 배경 하에 필자는 기계번역이 등장하면서 국내 학부 번역 전공생들의 번역 품질이 기계번역과 비교해서 경쟁력을 갖추고 있는지 확인해 볼 필요성을 느꼈다. 본 연구와 유사한 맥락의 기존 연구들은 학생들의 기계번역 포스트에디팅(machine translation post-editing, MTPE) 결과물을 분석하여 기계번역을 번역교육에 활용할 때 어떤 부분에 주의해야 하는지를 주로 논의했다면(김순미, 2017; 마승혜, 2018), 본 연구의 목적은 만약 인간번역(human translation)의 품질이 기계번역 포스트에디팅 품질보다 못하다면 국내 학부 번역 교육은 어떤 방향으로 나아가야 하는지 논의하는 것이다. 이를 위해 본고는 우선 학부 번역전공자를 대상으로 한영번역에서 인간번역과 기계번역 포스트에디팅의 품질을 비교할 것이다. 구체적으로 다루고자 하는 연구 문제는 다음과 같다.

- 1) 국내 학부 번역전공자의 인간번역과 기계번역 포스트에디팅은 한영번역 상황에서 어떤 것이 더 우수한가?

2) 위 연구 질문을 통해 확인된 품질 차이가 국내 학부 번역 교육에 주는 함의는 무엇인가?

본고의 구성은 다음과 같다. 우선 논문 2장에서는 인간번역과 기계번역 포스트에디팅의 품질을 비교한 선행연구를 간략히 조사하고, 학부 번역전공자의 포스트에디팅에서 나타나는 문제점을 살펴보도록 한다. 3장에서는 실험참여자들과 실험에 활용된 텍스트를 소개하고, 구체적인 실험 방법과 번역결과물의 평가 방식에 관해 기술한다. 이어 4장에서는 수집된 내용을 정량·정성적으로 분석하고 연구문제에 대한 답을 구하고자 한다. 마지막 5장에서는 연구 내용과 결과를 요약하고, 본 연구의 한계와 의의를 간략히 논의할 것이다.

2. 선행연구

기계번역의 품질이 획기적으로 개선되고 온라인 기계번역 사용이 보편화되면서 기계번역은 여러 학계와 연구분야에서 주목을 받고 있다. 이는 속속 등장하는 여러 편의 기계번역에 관한 메타연구만으로도 무리 없이 이해된다(원은하, 2019; 최문선, 2019; 이상빈, 2020; 신지선, 2020). 이들의 연구에 의하면, 상대적으로 짧은 번역학의 역사를 고려하면 최근 국내외 기계번역에 대한 관심과 연구 성과는 양적 질적으로 괄목할 만한 성장을 이루었다. 본 연구는 기계번역과 관련된 주제 가운데 먼저 인간번역과 기계번역 포스트에디팅의 품질을 비교한 연구들의 특징을 정리하고, 국내 학부 번역전공자들이 수행한 포스트에디팅에서 나타나는 문제점에는 주로 어떤 것들이 있는지 간략히 살펴보도록 한다.

2.1. 인간번역과 기계번역 포스트에디팅 품질 비교 연구

인간번역과 기계번역 포스트에디팅의 품질 차이를 비교한 국내외 연구는 주로 다음과 같은 특징을 보인다. 첫째, 기계번역 포스트에디팅의 주체는 주로 포스트에디팅을 전문적으로 훈련받은 전문번역가(포스트에디터 포함)이거나 포스트에디팅에 상대적으로 경험이 많은 대학원생이다(Fiederer & O'Brien, 2009; Lee, 2018; 이준호, 2018/2021a/2021b). 모삽(Mossop, 2007)은 번역텍스트의 오류를

수정하거나 검열하는 작업을 교정교열(revision)의 개념으로 설명한다. 그에 따르면 학부생들은 번역 자가교정(self-revision)이나 다른 사람의 번역을 교정할 수 있는 준비가 거의 되어 있지 않다. 오브라이언(O'Brien, 2002)도 번역텍스트를 수정해본 경험과 그로 인해 축적되는 자신감을 포스트에디팅의 필수 요소로 강조하며, 포스트에디팅 수업을 수강하기 앞서 언어 능력, 번역 능력 등에서 높은 기량을 쌓는 것이 필요하다고 설명한다. 그에 따르면 포스트에디팅 교육은 학부 마지막 과정에서 제공되거나 대학원 과정에서 이루어지는 것이 보다 이상적이다(O'Brien, 2002: 105). 이러한 이유로 그간 인간번역과 기계번역 포스트에디팅 간 품질 비교 연구에서 포스트에디팅의 주체는 대부분 포스트에디팅에 익숙하고 숙련된 전문번역가와 대학원생인 경우가 많았다. 하지만 지난 연구에서 논의된, 기계번역과 관련된 사회문화적 조건은 현재의 사회문화적 조건과 크게 다르다. 특히 지금의 국내 학부생들은 대학에 입학하기 전부터 인공지능경망 기반의 기계번역을 접했기 때문에 기계번역에 대한 이해와 경험이 이전 번역전공자들과는 매우 다를 수 있다(cf. Carl et al., 2015; Guerberof & Moorkens, 2019).

둘째, 인간번역과 기계번역 포스트에디팅의 품질 비교 연구는 주로 생산성 차원에서 다루졌다. 여기서는 포스트에디팅이 번역 생산성에 얼마나 긍정적으로 기여하는지, 나아가 생산성 향상은 대개 어떤 요소들에 의해 결정되는지에 초점을 맞춘다(Guerberof, 2014; 이준호, 2021a/2021b). 플리트와 마셀로(Plitt & Masselot, 2010)는 영어를 프랑스어, 이탈리아어, 독일어 등으로 번역하는 12명의 전문번역가를 참여시켜, 그들의 인간번역과 기계번역 포스트에디팅을 비교한 바 있다. 그 결과, 포스트에디팅은 평균적으로 작업량의 74% 이상을 증가시키는 데 도움이 되었다. 하지만 모든 언어쌍과 번역주체에 이 같은 결과를 적용하기는 어렵다. 예를 들면 구에베로프(Guerberof, 2014)는 같은 번역가라 하더라도 포스트에디팅에 숙달한 사람과 그렇지 못한 사람 간의 차이는 미미하다고 주장했다. 또한 수습 번역가(trainee translator)를 대상으로 영중번역을 살펴본 가르시아(García, 2010)의 연구에서도 생산성은 별 차이가 없었다. 반면 최근에는 이준호(2021b)가 국내에서 포스트에디팅을 훈련받은 석사생 9명을 대상으로 인간번역 대비 영한 포스트에디팅의 생산성 증가율을 살펴보았다. 그는 “인간번역 대비 평균 34%의 생산성 증가율 개선이 있음을 관찰했으며, 모든 참여자가 포스트에디팅에서 인간번역 대비 높은 생산성을 보였다”고 언급하였다. 또한 “텍스트에 따라 포스트에디팅의 생산성과 품질이 영향을 받을 수 있

음을 명확하게 관찰할 수 있었기에, 포스트에디팅 작업을 위한 텍스트 선택에도 주의가 필요”하다고 제안한다. 이상의 내용을 간단히 정리해 보면, 포스트에디팅의 생산성은 언어쌍, 텍스트 유형, 포스트에디팅 작업 환경에의 친숙함 정도 등에 따라 상이하게 나타날 수 있다(Koponen, 2016).

셋째, 인간번역과 기계번역 포스트에디팅의 품질 차이는 평가 영역에 따라 다르게 나타난다. 바꾸어 말하면, 인간번역이 반드시 기계번역보다 우수하다고 단정지을 수 없다. 피데러와 오브라이언(Fiederer & O'Brien, 2009)은 기계번역을 향한 태도에 기반하여 ‘기계번역 품질은 인간번역 품질보다 반드시 낮은가’라는 질문을 던진다. 그들은 전문분야 텍스트(specialised text)인 사용설명서(user guide)를 활용하여 영독 인간번역과 기계번역 포스트에디팅의 품질 차이를 명료성(clarity), 정확성(accuracy), 표현성(style) 측면에서 평가하였다. 그 결과, 인간번역은 표현성 영역에서만 비교적 나은 품질을 보였고, 기계번역 포스트에디팅은 명료성과 정확성 부분에서 최소 인간번역과 동등하거나 높은 품질을 보였다. 이와 같은 결과는 언어 전문가의 일반적인 예측과 다르다는 점에서 매우 흥미롭다(Fiederer & O'Brien, 2009: 68-9). 선행연구의 결과를 종합해 보면, 기계번역의 성능이 지속적으로 향상되고 있는 시점에서 인간번역과 포스트에디팅의 품질 격차가 언어쌍, 번역 방향, 번역 주체 등에 의해 어떻게 달라지는지도 확인할 필요가 있다.

2.2. 학부 번역전공자의 포스트에디팅에서 나타나는 문제점

국내 학부 번역전공자가 수행한 포스트에디팅에서 주로 나타나는 문제점의 특성은 다음과 같이 정리할 수 있다. 첫째, 국내 학부 번역전공자의 포스트에디팅은 언어조합, 번역방향, 기계번역 엔진 등의 차이와 관계없이 유사한 문제점을 보인다. 가령 이상빈(2017)은 구글 번역을 활용하여 학부 번역전공자가 수행한 영한번역 포스트에디팅 결과를 모삽의 교정교열 파라미터(revision parameters)로 분석하였다. 이 연구에서 사용된 원문은 비교적 통사 구조가 간단한 의학 텍스트였고, 최종 확인된 포스트에디팅 상의 문제점은 11가지 항목(메시지 전달의 정확성과 완결성, 논리와 사실 관계, 언어, 관용구, 문법, 문서의 레이아웃과 타이포그래피 등)으로 정리할 수 있다. 이상빈에 따르면 학생들은 원문과 기계번역 결과물을 충분히 비교하지 않은 채 불완전한 단일 언어 포스트에디팅을 수

행했고 텍스트 맥락을 제대로 파악하지 않은 채 단어나 구 차원의 수정에만 몰두했다. 또한 메시지 전달에는 집중했으나 문체적 측면 및 전문용어 처리에는 소홀했다. 이러한 연구 결과는 구글 번역과 네이버 파파고를 활용하여 한영번역 포스트에디팅을 살펴본 김순미(2017)의 연구, 네이버 파파고에서 중한번역 포스트에디팅을 다룬 김혜림(2020)의 연구 등에서도 다수 확인된다.

둘째, 연구에 따라 동일한 텍스트 유형일지라도 상반된 문제점이 나타나기도 한다. 일례로 마승혜(2018)는 텍스트 유형에 기반하여 한영 기계번역 포스트에디팅 상황에서의 문제점을 논의하였다. 그는 정보적(informative) 텍스트의 경우 텍스트적 연결성이 중요하고, 표현적(expressive) 텍스트에서는 화용적 명시화가 중요하며, 설득적(persuasive) 텍스트에서는 적절한 단어 선택 및 표현 선택에 신경 써야 한다고 주장한다. 특히 텍스트 유형에 따라 포스트에디팅에서 나타나는 문제점을 세분화하고 이에 따라 가이드라인도 정교하게 제시되어야 함을 강조한다. 김혜림(2020) 역시 정보적 텍스트와 설득적 텍스트를 활용하여 중한 기계번역 포스트에디팅을 연구하였다. 그는 의미(message) 항목에서의 정확성(accuracy) 문제를 지목하며 “정보적 텍스트[보고서]에서는 주로 단어의 문제가 나타났으나 설득적 텍스트[연설문]에서는 단어보다 연결 절의 논리 관계, 강조어 등 화자 태도 표현 문제가 주로 나타났다”고 설명한다. 마승혜(2018)와 김혜림(2020)의 연구 결과를 비교해 보면, 텍스트 유형에 따른 문제점이 두 연구간 상충되는 경향을 보인다. 예컨대 마승혜(2018)는 정보적 텍스트에서 텍스트적 연결성을, 설득적 텍스트에서 적절한 단어 및 표현 선택을 지적하였는데 반해, 김혜림(2020)은 정보적 텍스트에서 단어의 문제를, 설득적 텍스트에서 연결 절의 논리 관계를 지적하였다. 즉, 기계번역 포스트에디팅에서 주로 나타나는 문제점은 텍스트 유형보다는 언어조합과 번역 방향 등의 변수로 더 분명해질 가능성이 있다.

3. 연구방법

3.1. 실험대상 및 텍스트 선정

본 연구에서는 국내 (통)번역 전공 학부생을 대상으로 인간번역과 기계번역 포스트에디팅의 품질 차이를 비교해 보는 실험을 진행하였다. 여기서 “인간번

역”은 실험참여자가 한국어 원문을 영어로 직접 번역하는 것이고, “기계번역 포스트에디팅”은 한국어 원문의 온라인기계번역(영어) 결과값을 실험참여자가 수정하는 것을 의미한다. 실험참가자는 2021학년도 1학기 서울의 한 대학에서 번역을 (이중) 전공하는 2학년 학생들로, 연구자 1인이 담당할 동일 과목 두 개 분반에서 수업을 수강한 60명이다.

실험참여자는 실험 전 수업에서 기계번역과 포스트에디팅 개념을 학습했고 영한 기계번역 포스트에디팅 사례도 살펴보았다. 또한 구글번역을 활용해 국문으로 작성된 짧은 설명문과 광고문의 영어 결과값을 텍스트 유형 차원에서 비교하고 이를 포스트에디팅해보는 과업도 마친 상태였다. 따라서 실험참여자들은 최소한 기계번역 포스트에디팅에 대한 기초적인 이해와 경험을 가지고 있었다.

실험에는 두 개의 한국어 원문이 사용되었는데, 모두 정보적 텍스트인 신문 기사이다. 정보적 텍스트는 주로 사실과 정보만 전달하기 때문에 기계번역 포스트에디팅을 활용하여 최종 번역물을 만들어내는 데 효율적이다(마승혜, 2018). 텍스트 선별 시, 한국어교육연구소에서 제공하는 한국어 읽기 텍스트 난이도 측정 프로그램(출처: <http://korean-lab.org>)을 활용하여 어휘, 문법, 이독성 등을 고려하였다. 그 결과, 주제는 상이하지만 난이도 측면에서 유사한 텍스트 두 개를 각각 조선일보 <사회> 면과 조선비즈 <금융·증권> 면에서 발췌할 수 있었다. 선정된 텍스트는 표 1과 같다.

표 1. 실험에 사용한 한국어 원문

Text A (171자)	문제는 꿀벌이 자연 생태계 유지에 결정적 역할을 하는 곤충이란 점이다. UN 식량농업기구(FAO)에 따르면 꿀벌이 세계 100대 농작물의 71%를 매개(가루받이)하고 있다. 꿀벌 의존도가 100%인 아몬드든 벌이 사라지면 재배하기가 불가능해지고, 사과나 양파 등도 수분의 90% 이상을 꿀벌에 의존하고 있다.
Text B (160자)	주식은 올해 코로나 쇼크로 가장 많은 조정을 받은 자산이다. 유럽과 미국을 중심으로 바이러스 확산세가 높게 유지되는 등 코로나 불안감은 여전하다. 전문가들은 선진국 주식은 더 지켜보되 코로나 사태가 소강 내지 진정 국면에 접어들고 있는 중국과 한국 주식은 “담아볼 만하다”고 입을 모았다.

3.2. 실험방법

인간번역과 기계번역 포스트에디팅 간의 한영 번역 품질 차이를 비교하기 위해 필자는 학기 중 2주에 걸쳐 두 번의 번역 과업을 실시하였다. 해당 과업이 있었던 학기에 COVID-19의 영향으로 전면 비대면 수업이 진행되었기 때문에 번역 과업은 e-Class <설문> 기능을 활용하여 익명으로 실시하였다.

1차 과업에서 학생들은 앞서 살펴본 표 1의 Text A를 번역하였다. 분반 (1)은 기계번역의 도움 없이 직접 번역하였고, 분반 (2)는 기계번역 포스트에디팅을 실시하였다. 반대로 2차 과업에서는 Text B를 기반으로, 분반 (1)은 포스트에디팅을 수행했고 분반 (2)는 직접 번역을 수행하였다. 다시 말하면, 실험참여자는 다른 텍스트를 가지고 한 번은 인간번역을 했고 다른 한 번은 포스트에디팅을 수행했다. 이때 과업에 부여된 시간은 인간번역에는 20분, 포스트에디팅에는 10분을 할당하였다. 포스트에디팅의 최대 장점과 존재 이유가 효율성과 생산성임을 고려할 때 이러한 시간 배분은 타당하다고 판단했다. 실험참여자는 포스트에디팅 과업 직후 15분 동안 포스트에디팅의 근거를 간단히 기술하도록 했다. 모든 과업은 계획된 시간에 시작되고 마쳤으며 지정된 시간 외에는 참여할 수 없도록 설계되었다.

기계번역 포스트에디팅에 사용한 온라인기계번역은 실험참여자가 직접 선택하도록 하였고, 어떤 것을 수정하고 어느 정도까지 수정해야 하는지 등 세부 가이드라인은 제시되지 않았다. 왜냐하면 실험참여자가 기계번역 포스트에디팅에 숙련된 학생들도 아니고 경험도 많지 않기 때문에 특정 기계번역 엔진에 대한 이해가 부족할 수 있고, 인간번역과 기계번역 포스트에디팅의 품질을 비교하는 실험의 특성상 어느 한쪽에서만 구체적인 가이드라인이 제공된다면 평가에도 영향을 미칠 수 있기 때문이다. 과업 지침을 정리하면 아래 표 2, 표 3과 같다.

표 2. 인간번역 과업 지침

과업 지침	제한시간
제시된 한국어 텍스트(ST)를 영어로 번역하십시오. (반드시 직접 번역할 것, 기계번역 사용 금지)	20 분

표 3. 기계번역 포스트에디팅 과업 지침

단계	과업 지침	제한시간
1	제시된 한국어 텍스트(ST)를 온라인기계번역을 활용한 후 포스트에디팅하시오. (본인의 포스트에디팅 결과물을 반드시 저장해 두었다가 두 번째 단계에서 활용하시기 바랍니다.)	10 분
2	기계번역을 수정한 부분을 표시하고 수정한 이유에 대한 간단한 의견을 첨부해 주세요. 이때, 본인의 최종결과물을 함께 게시해 주세요.	15 분

3.3. 결과물 평가

평가는 연구자 2인이 독립적으로 진행하였다. 본격적인 평가에 앞서 전체 과업 결과물은 제3자에 의해 무작위 순으로 재정리되었다. 그리고 인간번역과 기계번역 포스트에디팅을 구분하지 않은 채 연구자 2인에게 각각 전달되었다. 개별 평가가 끝난 후에는 다시 제3자가 인간번역과 기계번역 포스트에디팅 별로 번역 결과물을 정리하였고, 이렇게 재정렬된 데이터를 연구자 2인이 최종 분석에 활용하였다.

평가는 텍스트의 길이와 과업 시간을 고려하여 일괄평가 방식으로 진행되었다. 연구자들은 과업에 활용된 원문의 특성을 고려할 때 텍스트의 표현적 측면보다는 어휘·문법의 정확성과 텍스트의 완결성 차원에서 총체적인 평가가 이루어져야 함에 동의했다. 또한 철자나 띄어쓰기와 같은 오류를 잡기보다는 내용 전달 측면에서 얼마나 정확하고 적합한 표현을 찾아 주었는지를 평가하는 것이 과업의 특성, 학생들의 역량, 텍스트 유형 등에도 적합하다고 판단하였다. 이에 선행연구를 검토하여 5점 척도의 평가 기준표(표 4)를 고안하여 평가에 활용하였다(Lai, 2008; Tzu-yi Lee, 2013).

평가를 앞두고는 평가자 간 신뢰도(inter-rater reliability)를 높이기 위해 평가 연습이 선행되었다. 이 과정에서 평가자 1과 평가자 2는 번역 샘플을 활용하여 평가기준표에 따라 점수를 매기고 그 결과를 비교하였으며 최대한 일관된 기준을 적용, 공유하고자 논의하였다(Doherty, 2017). 또한 평가점수의 차이가 2점 이상 난 경우, 재평가를 하거나 제3의 평가위원을 영입해 세 점수를 평균내기로 결정하였다.

표 4. 평가기준표(Tzu-yi Lee, 2013 재구성)

점수	내용
5 점	메시지 전달 및 어휘-문법적 측면에서 오류가 없고 잘 읽힌다. 비교 대상 가운데 추천할 만한 번역이다.
4 점	어휘-문법적 측면에서 작은 실수나 오류가 소수 확인된다. 하지만 원문과 번역문을 비교할 때, 중대한 오류라고 보기는 어렵다.
3 점	어휘-문법적 측면에서 중대한 실수나 오류가 소수 확인된다. 또한 작은 실수나 오류도 소수 있어 메시지를 이해하기가 조금 힘들다.
2 점	어휘-문법적 측면에서 크고 작은 실수나 오류가 여러 개 있다. 메시지 전달에 문제가 있고 이독성(readability)도 떨어진다.
1 점	어휘-문법적 측면에서 문제가 많고 메시지 전달이 제대로 이루어지지 않아 텍스트로서 기능할 수 없다.

4. 연구결과

4.1. 평가 결과

인간번역과 포스트에디팅 결과물은 총 113개가 수집되었다. 실험 참가자 수를 고려하면 120개의 번역 결과물이 수집되었어야 했지만, 과업을 제출하지 않았거나 번역을 수행하지 않고 원문만 제출한 경우가 있었기에 7개는 최종 데이터에서 제외되었다. 전체 번역 결과물 113개 중 Text A의 인간번역과 포스트에디팅 결과물은 총 59개이고, Text B의 인간번역과 포스트에디팅 결과물은 총 54개이다.

Text A와 Text B에 대해 평가자 1과 평가자 2의 평가 평균은 표 5와 같다. 이 표에서 확인할 수 있듯이 평가자 2가 두 개 텍스트 모두를 평가자 1보다 조금 높이 평가하였다. 하지만 평가자 간의 점수 차는 Text A에서 0.05, Text B에서 0.04이므로 그 차이는 아주 미미하다. 또한 113개의 개별 평가에서 5점 척도를 사용한 두 평가자가 2점 이상 차이를 보인 경우는 단 한 건도 없었다. 이는 평가자 간 신뢰도가 어느 정도 지켜졌다는 뜻이고 평가자 모두 평가기준표를 잘 준수한 결과이다.

표 5. Text A와 Text B에 대한 두 평가자의 평점(5점 척도 기준)²⁾

원문	평가자 1	평가자 2	점수 차
Text A (번역 59 개)	2.81	2.86	0.05
Text B (번역 54 개)	2.35	2.39	0.04

4.2. 인간번역과 기계번역 포스트에디팅 품질 비교

인간번역과 포스트에디팅 품질 평가 결과는 표 6과 같다. 한눈에 알 수 있듯이 한영번역 맥락에서 볼 때 기계번역 포스트에디팅은 Text A와 Text B 모두에서 학생이 직접 번역한 인간번역보다 나은 품질을 보여준다. 특히 Text A의 경우, 기계번역 포스트에디팅의 품질이 월등히 높은 것을 확인할 수 있다(1.48 점 차이). 인간번역 대비 기계번역의 품질은 평균 1점(5점 척도에서 1구간) 이상 높았다.

표 6. 인간번역과 포스트에디팅의 품질 평가 결과

원문	인간번역	포스트에디팅	점수 차
Text A	2.09	3.57	1.48
Text B	1.98	2.79	0.81
평균	2.04	3.15	1.11

표 6과 같은 결과는 인간번역과 포스트에디팅에 차등 할애된 시간을 고려하면 그 차이가 훨씬 부각된다. 즉, 기계번역 포스트에디팅은 과업 수행 시간도 인간번역에 비해 절반밖에 되지 않는데 그 품질은 인간번역보다 훨씬 높은 것이다. 만약 전문분야의 텍스트와 같은 수준으로 원문의 난도가 높아진다면 기계번역 포스트에디팅과 인간번역의 품질 차이는 더욱 벌어질 가능성도 있다. 게다가 학생들이 포스트에디팅을 전문적으로 교육받아 필요한 지식과 요령을 터득한다면 점수 차는 훨씬 더 커질 수도 있다.

2) 본고의 모든 수치는 소수점 두 번째 자리까지 반올림한 것이다.

4.3. 텍스트 차원에서의 품질 비교(질적 분석)

실험참여자가 수행한 인간번역과 기계번역 포스트에디팅에서 확인할 수 있는 텍스트 차원에서의 차이는 대략 다음과 같다.

첫째, 포스트에디팅은 문법 차원에서 상대적으로 안정적인 품질을 보여줬지만, 인간번역은 단순한 문법에서도 크고 작은 실수나 오류가 많았다. 아래 사례 1과 사례 2는 각각 Text A와 Text B의 인간번역에 해당한다. 강조된 부분은 수(number), 시제/상(tense/aspect) 등에서의 오류를 보여준다. 우선 사례 1에서 첫 번째 that절의 주어는 복수 명사 “honeybees”이다. 여기서 주어가 복수이기 때문에 두 번째 that절의 동사 “plays”는 play로 수정되어야 문법적으로 타당하다. 사례 2의 경우 해당 학생은 원문의 “(전문가들은) ...입을 모았다”를 현재진행형인 “(Analysts) are saying”으로 번역하였다. 원문은 특정 시기의 주식 시장 상황을 염두에 두고 전문가들의 공통된 의견을 전달하는 맥락이므로 번역에서는 과거 또는 현재완료 시제를 사용하거나 맥락을 고려해 현재시제를 사용하는 것이 문법 규칙에 어긋나지 않고 원문 내용을 왜곡 없이 전달할 수 있다.

사례 1 (Text A 인간번역)

The problem is that honeybees are insects that **plays** [수 일치 오류] a major role in maintaining natural eco-system. According to U.N. FAO, honeybees are pollinating 71% of the world’s top 100 crops.

사례 2 (Text B 인간번역)

Analysts **are saying** [시제/상 오류] that stocks of South Korea and China are “worth a shot” due to declining COVID cases, whereas in other developed countries, it is safer to wait.

둘째, 포스트에디팅 수행 시 실험참여자들은 어휘 차원에서 잘못된 기계번역 결과값을 그대로 수용하였다. 아래 사례 3은 Text A와 두 개의 포스트에디팅

결과물이다. 학생 상당수는 <포스트에디팅 1>처럼 원문의 “매개(가루받이)하다”를 “mediate(pollinate)”로 포스트에디팅 하였다. 영어 mediate는 문제에 대한 해결책을 찾기 위해 ‘중재하다’, ‘조정하다’의 뜻이므로, 본 맥락에서는 어색하다. 이 경우 독자는 학생들이 남겨둔 괄호 속 어휘 “pollinate”(수분受粉 하다)로 의미를 보완해 맥락을 이해해야 한다. 심지어 사례 3의 <포스트에디팅 2>에서 볼 수 있듯이, 학생들이 제출한 결과물에는 오염을 뜻하는 “pollution”도 확인된다.

사례 3

<Text A 원문>

UN 식량농업기구(FAO)에 따르면 꿀벌이 세계 100 대 농작물의 71%를 매개(가루받이)하고 있다.

<포스트에디팅 1>

According to the Food and Agriculture Organization of the United Nations (FAO), honeybees mediate(pollinate) 71% of the world’s top 100 crops.

<포스트에디팅 2>

According to the UN Food and Agriculture Organization (FAO), honeybees mediate (pollution) 71% of the world’s top 100 crops.

셋째, 원문을 가급적 유지하는 포스트에디팅과 달리, 인간번역에서는 장문을 활용하거나 구조적으로 복잡한 문장을 단순화시켜 원문의 의미를 누락시키거나 뉘앙스를 바꾸기도 한다. 이로 인해 의미의 정확성에도 영향을 주었다. 사례 4에서 실험참여자는 하나의 문장을 두 개로 분리해 “꿀벌에 100% 의존하는 아몬드는 꿀벌 없이 재배가 불가능하다. 게다가 사과 또는 양파는 물공급을 위해 꿀벌에 90% 이상 의존한다”(필자의 직역)로 번역한다. 이때 ‘(환경 등의 문제로) 꿀벌이 사라지면’이 전제가 되어야 하는 내용인데 그 부분을 번역하지 않아서 원문의 뉘앙스가 달라졌다.

사례 4(Text A 인간번역)

꿀벌 의존도가 100%인 아몬드든 벌이 사라지면 재배하기가 불가능해지고, 사과나 양파 등도 수분의 90% 이상을 꿀벌에 의존하고 있다.

Almonds, which is relying 100% on honeybees, cannot be grown without them. Besides, apples or onions are dependent on them for supplying water more than 90%.

한편 사례 5의 실험참여자는 문장 전체를 번역하는 대신, “전문가들은 선진국 주식은 볼 필요가 있더라도 중국과 한국 주식은 살 가치가 있다고 전했다”(필자의 직역)고 문장을 단순화시킨다. 여기서도 원문의 ‘코로나 사태가 소강 내지 진정 국면에 접어들고 있는’이 번역되지 않아서 번역 독자가 맥락을 제대로 파악하는 데 필요한 정보를 충분히 담지 못했다.

사례 5(Text B 인간번역)

전문가들은 선진국 주식은 더 지켜보되 코로나 사태가 소강 내지 진정 국면에 접어들고 있는 중국과 한국 주식은 “담아볼 만하다”고 입을 모았다.

The experts said it is worth to get Chinese and Korean stocks even if we need to watch stocks in developed countries.

넷째, 학생들은 포스트에디팅 과정 중에 기계번역 결과값의 품질을 떨어뜨리기도 한다. 가령 사례 6에 나타난 한 학생의 포스트에디팅과 이에 대한 학생의 설명을 확인해 보자.

사례 6

The problem is that honeybees are insects that play an important role in maintaining the ecosystem. According to the UN Food and Agriculture Organization (FAO), honeybees pollinate 71% of the world’s top 100 crops. Almonds, which are 100% dependent on honeybees, are **unable to**

raise when honeybees disappear. **Apples and onions also depend on honeybees for 90% of their stored water.** (강조는 필자의 것)

<위 포스트에디팅에 대한 학생 자신의 설명>

Decisive → important decisive 는 물론 “결정적”이라는 뜻을 갖고 있긴 하다. 하지만 결단을 내릴 때 있어서 결정적이라는 뜻이지 꿀벌이 일을 할 때 비판적인 사고를 가지고 결정을 내리지는 않으니 어색하다고 느껴 important 로 바꾸어 줬다. natural ecosystem → ecosystem: 굳이 natural 을 붙일 필요없이 ecosystem 만으로 생태계의 의미는 전달된다. Food and Agriculture Organization → UN Food and Agriculture Organization: UN 을 추가해서 원문에 가깝게 수정했다. become impossible to grow → unable to raise 번역기가 번역한 것도 틀리진 않다고 생각했으나 문장이 깔끔하지 못하다고 생각하여 다듬는 과정에서 바꿨다. and apples and onions depend on honeybees for more than 90% of the water → **Apples and onions also depend on honeybees for 90% of their stored water.** 고치는 과정에서 more than 90%라고 했어야 했는데 빼먹었다. 콤마와 and 로 두 문장이 이어져 있길래 원문이 이어져 있어서 이렇게 출력되었다는 걸 알았다. 하지만 영어에서 어색하므로 두 문장을 분리했다. 수분을 stored water 로 번역하면 더 정확하다고 생각했다. moisture 도 생각해봤는데 이건 식물 겉의 수분기를 나타내는 것 같아서 사용하지 않았다. (강조는 필자의 것)

이 학생은 맥락을 고려해 어휘를 수정하거나 이독성을 높이기 위해 원문의 복잡한 문장을 나누는 등, 주로 어휘와 문장 구조를 포스트에디팅 했다고 설명한다. 하지만 원문의 마지막 문장에서 기계번역 결과값을 “문장이 깔끔하지 못하다”는 이유로 수정하다가 오류를 범했다. 이를 구체적으로 확인하면, 기계번역 값인 “Almonds ... become impossible to grow”를 “Almonds ... are unable to raise”로 수정했는데, 이러한 교정은 원문 내용이 “불가능해지고”임을 고려할 때 정확성이 떨어진 것이다. 또한 주어가 식물에 해당하는 Almonds라는 점에서 동사 raise보다는 기계번역 출력값인 grow가 바람직하다. 여기서 특이한 점은 이미 원문의 “매개(가루받이)”를 “pollinate”로 잘 번역해 주었음에도 불구하고 마지막 문장에서 “수분”을 pollination 대신 “[stored] water”로 남겨두었다는

것이다.

4.4. 실험 결과 논의

앞서 살펴보았듯이 국내 학부생의 한영번역 포스트에디팅은 기계번역의 도움 없이 혼자서 하는 번역보다 품질이 높았다. 번역 내용 면에서는 인간번역의 경우 수, 시제/상 등 문법에서의 문제가 대부분의 번역문에서 확인되었고, 복잡한 문장을 분리하는 과정에서 원문의 메시지가 누락되는 경우가 제법 있었다. 반면 기계번역 포스트에디팅은 문법 측면에서 상대적으로 안정적인 품질을 보여주고 원문의 구조를 대체로 유지해서 메시지 누락은 거의 없었다. 또한 학부생들은 포스트에디팅 과정에서 기계번역 결과값의 어휘를 그대로 수용하거나 오히려 기계번역 결과값을 잘못 수정하여 최종 번역 품질을 저하하기도 했다.

이러한 결과는 다음과 같은 논의로 이어진다. 첫째, 일반적으로 학부생들이 외국어로의 번역을 수행하는 것은 기계번역 시대에 실무 관점에서 적절하지 않을 수도 있다. 특히 학부생이 보수를 받고 기계번역의 도움 없이 한영번역을 맡는 것은 생산성 측면에서 효율적이지 않고 품질 면에서도 손해가 될 수 있다. 실험 결과만 놓고 보면, 학부생이 실무에서 한영번역을 할 때 기계번역의 사용은 선택이 아닌 필수일지도 모른다.

둘째, 학부 과정에서 한영번역이나 한영번역 포스트에디팅을 가르칠 때 어휘·문법 등의 기초 분야에 신경 써야 한다. 인간번역과 포스트에디팅 모두에서 학생 개인의 어휘·문법에서의 영어 실력은 번역 품질을 결정짓는 주된 요소이다. 학부생에게 번역이든 포스트에디팅이든 번역과 관련된 실무 지식을 가르칠 때는 보다 근본적인 관점에서 접근할 필요가 있다. 즉, 새로운 것, 기술적인 것에 집중할 것이 아니라 언어, 특히 외국어 역량의 기반을 탄탄하게 갖추 수 있도록 도와야 한다.

셋째, 번역학과에서 외국어로의 번역을 가르칠 때 교육 목표나 방향을 다시 고심해 볼 필요가 있다. 예컨대 지금도 기계번역의 품질이 인간번역(학생번역)보다 우수헌데, 기계번역의 지능이 날로 높아지는 상황에서 국내 학부생들에게 한영번역을 교육하는 것이 어떤 의미를 갖는지 고민해야 한다. 또한 포스트에디팅을 수반하지 않은 한영번역이 경쟁력을 갖추려면 어떤 부분에 더 신경을 써야 하는지, 나아가 현재의 교육 방식이 번역업계의 업무 방식과 어떤 점에서

유사하고 다른지 등을 충분히 논의해야 한다. 이와 같은 질문들은 새로운 목표와 방향을 설정하는 데 유용한 힌트를 줄 수 있기 때문이다.

넷째, 본 연구의 결과가 더욱 의미 있으려면 학부생과 교수자를 대상으로 기계번역 사용 실태 및 인식 조사를 병행해야 한다. 가령 실제 번역 상황에서 기계번역이 어떤 점에서 유용하다고 생각하는지, 학부 과정에서 포스트에디팅 교육은 어떠해야 하는지, 향후 기계번역 관련 진로를 선택할 때 어떤 입장을 취해야 하는지 등을 면밀하게 살펴볼 필요가 있다. 이와 같은 실증적 조사가 이루어진다면 우리가 현재 간과하고 있는 부분을 보완할 수 있고, 미래의 기계번역 교육도 전략적으로 설계할 수 있을 것이다.

5. 결론

본고에서는 국내 학부생 60명을 대상으로 두 번의 번역 과업을 통해 인간번역과 기계번역 포스트에디팅의 품질 차이를 정량·정성적으로 분석하였다. 분석 결과에 따르면, 학부생들은 정보적 텍스트에 기반한 한영번역 상황에서 직접 번역할 때보다 기계번역을 활용한 포스트에디팅에서 더 나은 번역을 수행하였다. 시간 대비 기계번역 포스트에디팅의 품질은 매우 안정적인데 반해, 인간번역은 어휘, 문법 측면에서 오류가 많았고 그로 인해 이독성도 낮아졌다. 한 가지 흥미로운 사실은 실험참여자는 포스트에디팅에 상대적으로 익숙한 대학원생이나 전문번역가가 아니라는 점이다. 더군다나 실험참여자가 소속된 학과에는 다양한 분야의 통번역 교과목이 운영되고 있지만 기계번역 포스트에디팅을 훈련할 수 있는 정규 수업은 컴퓨터 보조 번역인 “Computer Aided Translation” 단 한 과목에 불과하다. 결과적으로 실험참여자들은 포스트에디팅 경험이나 훈련이 충분하지 않을 가능성이 있는 상태에서도 직접 번역할 때보다 기계번역을 활용할 때 번역 품질을 높일 수 있었다.

본 연구에는 몇 가지 한계가 있다. 첫째, 본 연구는 국내 학부생 60명만을 대상으로 진행하였고 실험에 사용한 텍스트도 분량이나 내용 면에서 대표성이 높지 않기 때문에 연구의 결과를 일반화할 수 없다. 향후 연구에서는 표본 수를 확대하고 다양한 유형의 텍스트를 활용해 본 연구 결과를 재검증하고 보완할 필요가 있다. 둘째, 필자는 한영번역 포스트에디팅만을 고려하였다. 따라서

후속 연구자들은 모국어로의 번역이나 다단계 번역을 고려하는 등, 번역방향과 언어쌍을 보다 다각화할 필요가 있다. 셋째, 본고의 실험은 COVID-19의 영향으로 비대면 수업에서 진행되었기 때문에 비대면 환경에서 과생될 수 있는 여러 문제점을 완전히 통제할 수 없었다. 향후 연구에서는 면대면 교육 환경에서도 비슷한 결과가 도출되는지 신뢰성 검사를 할 필요가 있다.

하지만 본 연구의 결과는 국내 학부 번역 수업을 통해 입증된 경험적 데이터라는 점에서 실제 한영 번역 및 외국어 교육뿐만 아니라 기계번역에 기반한 다양한 교실 환경에서 학문적 기초자료로 활용할 수 있다. 또한 본 연구는 가속화되는 기계번역 시대에서 번역 교육의 목표와 방향을 다시 생각하게 하므로 시의성이 있고, 비록 제한적이긴 하지만 학부 번역전공자들에게 기계번역이 어느 정도 효용 가치가 있는지 정량·정성적으로 검증해 보았다는 것에 의미가 있다. 추후 연구에서는 번역 교실 현장에서의 기계번역 활용 방안과 관련하여 실증적이고 구체적인 논의가 이루어지길 기대한다.

참고문헌

- 김순미. (2017). 「신경망번역기(NMT) 활용 학부 번역교육의 가능성 연구」 『통번역교육연구』 15(3): 5-37.
- 김은영. (2021). 「AI기계번역을 활용한 외국어교육의 방향성-포스트에디팅 수업 사례를 통해」 『언어학연구』 26(1): 23-42.
- 김혜림. (2020). 「학부생 번역자의 중한 기계번역 포스트에디팅 사례연구」 『중국언어연구』 91: 383-414.
- 남보리. (2021). 「LSP 관점에서 바라본 기계번역 포스트에디팅의 전략적 지향점: 스타일 가이드 지침 연구 및 LSP 기계번역 서비스 사례」 『번역·언어·기술』 2: 5-43.
- 마승혜. (2018). 「한영 기계번역 포스트 에디팅에 대한 경험적 고찰: 학부 교육 과정 및 결과를 중심으로」 『통번역학연구』 22(1): 53-87.
- 박건영. (2021). 「정보성 텍스트의 한영 기계번역 포스트에디팅 가이드라인 제시 - 신경망 기계번역(NMT)을 사용한 뉴스 기사문 번역의 사례」 『번역학연구』 22(1): 109-137.
- 박옥수. (2017). 「한영 기계번역에서 ST의 유형적 특징에 따른 번역 오류 분석」 『동아인문학』 41: 155-183.
- 박옥수. (2018). 「원천언어의 관점에서 살펴본 기계번역의 오류분석과 수정 원칙: 비문학 텍스트의 통사적 특징에 근거해서」 『동아인문학』 44: 151-171.
- 변길자. (2021). 「기계번역을 활용한 한영번역에서의 번역품질 비교연구」 『영미영문학』 26(3): 233-259.
- 서보현·김순영. (2018). 「기계번역 결과물의 오류유형 고찰」 『번역학연구』 19(1): 99-117.
- 신지선. (2020). 「기계번역 포스트에디팅에 관한 해외 연구 동향」 『번역학연구』 21(4): 87-114.
- 원은하. (2019). 「국내 기계번역 연구 현황 및 전망」 『인문-IT융합논총』 1(1): 11-28.
- 윤미선, 김택민, 임진주, 홍승연. (2018). 「영어-한국어 언어쌍에 적합한 포스트에디팅 가이드라인: 포스트에디터를 위한 가이드라인과 그 실례」 『번역학연구』 19(5): 43-76.
- 이상빈. (2017). 「학부번역전공자의 기계번역 포스트에디팅, 무엇이 문제이고, 무엇을 가르쳐야 하는가?」 『통역과 번역』 19(3): 37-64.
- 이상빈. (2018a). 「학부 번역전공자의 구글 기계번역 포스트에디팅에 관한 현상학 연구」 『통번역학연구』 22(1): 117-143.
- 이상빈. (2018b). 「국내학부 번역전공자의 구글 기계번역 포스트에디팅 과정(process) 및 행위 연구」 『번역학연구』 19(3): 259-286.
- 이상빈. (2020). 「기계번역에 관한 KCI 연구논문 리뷰: 인문학 저널 논문(2011~2020년 초)의 논의내용과 연구방법을 중심으로」 『통역과 번역』 22(2): 75-104.
- 이준호. (2018). 「포스트에디팅 교육을 위한 포스트에디팅과 인간번역의 차이 연구」 『통역과 번역』 20(1): 73-96.
- 이준호. (2021a). 「한영 포스트에디팅, 누구나 수행할 수 있는가?: 포스트에디팅 수업 설계를 위한 예비 연구」 『번역학연구』 22(1): 171-198.
- 이준호. (2021b). 「영한 포스트에디팅 생산성에 대한 고찰-시간적 노력을 중심으로」 『통번역학연구』 25(2): 55-83.

- 임희주. (2017). 「교양영어 수업에서 영어자동번역기 사용에 대한 대학생의 인식 및 태도연구: 영작문 수업을 중심으로」 『교양교육연구』 11(6): 727-751.
- 조인희. (2018). 「기계번역을 활용한 영어작문 수업의 효과: 자기 효능감과 글쓰기의 양적 증가를 중심으로」 『영어영문학연구』 60(2): 253-279.
- 조인희. (2021). 「EFL 대학 작문 수업에서 기계번역 활용의 학습 효과」 『언어학 연구』 58: 385-416.
- 최문선. (2019). 「국내 번역학 기계번역 연구 동향: 내용 분석과 키워드 분석을 중심으로」 『언어학연구』 24(1): 275-297.
- 한현희. (2019a). 「한노 기계번역의 오류 유형화 및 품질 개선을 위한 프리에디팅(pre-editing) 규칙 제안」 『통번역학연구』 23(3): 291-327.
- 한현희. (2019b). 「노한 기계번역의 프리에디팅 (pre-editing) 및 포스트에디팅 (post-editing) 가이드라인 제안」 『노어노문학』 31(4): 65-96.
- Ahn, S., & Chung, E. S. (2020). Students' perceptions of the use of online machine translation in L2 writing. *Multimedia Assisted Language Learning* 23(2): 10-35.
- Bahri, H., & Mahadi, T. S. T. (2016). Google Translate as a supplementary tool for learning Malay: A case study at Universiti Sains Malaysia. *Advances in Language and Literary Studies* 7(3): 161-167.
- Briggs, N. (2018). Neural machine translation tools in the language learning classroom: Students' use, perceptions, and analyses. *The JALT CALL Journal* 14(1): 2-24.
- Carl, M., Silke, G., & Silvia, H. S. (2015). Post-editing machine translation-efficiency, strategies and revision processes in professional translation settings. In Aline, F., & Schweiter, J. W. (Eds.), *Psycholinguistic and cognitive inquiries into translation and interpreting* (pp. 145-174). Amsterdam: John Benjamins.
- Doherty, S. (2017). Issues in human and automatic translation quality assessment. In Kenny, D. (Ed.), *Human issues in translation technology* (pp. 131-148). London: Routledge.
- Fiederer, R., & O'Brien, S. (2009). Quality and machine translation: A realistic objective? *The Journal of Specialised Translation* 11: 52-74.
- Flanagan, M., & Christensen, T. P. (2014). Testing post-editing guidelines: How translation trainees interpret them and how to tailor them for translator training purposes. *The Interpreter and Translator Trainer* 8(2): 257-275.
- Garcia, I. (2010). Is machine translation ready yet? *Target* 22(1): 7-21.
- Garcia, I., & Pena, M. I. (2011). Machine translation-assisted language learning: Writing for beginners. *Computer Assisted Language Learning* 24(5): 471-487.
- Groves, M., & Mundt, K. (2015). Friend or foe? Google Translate in language for academic purposes. *English for Specific Purposes* 37: 112-121.
- Guerberof A. (2014). Correlations between productivity and quality when post-editing in a professional context. *Machine Translation* 28(3-4): 165-186.
- Guerberof A., & Moorkens, J. (2019). Machine translation and post-editing training as part of a master's programme. *The Journal of Specialised Translation* (31): 217-238.
- Haque, R., Hasanuzzaman, M., & Way, A. (2020). Analysing terminology translation errors in statistical and neural machine translation. *Machine Translation* 34(2): 149-195.
- Hu, K., & Cadwell, P. (2016). A comparative study of post-editing guidelines. *Baltic Journal of*

Modern Computing 4(2): 346-353.

- Koponen, M. (2016). Is machine translation post-editing worth the effort? A survey of research into post-editing and effort. *The Journal of Specialised Translation* 25: 131-148.
- Lai, T. (2008). Comparisons of four translation assessment scales. *Compilation and Translation Review* 1(1): 71-92.
- Lee, I. (2018). A quality comparison of English translations of Korean literature between human translation and post-editing. *International Journal of Advanced Culture Technology* 6(4): 165-171.
- Lee, T. (2013). Incorporating translation into the language classroom and its potential impacts upon L2 learners. In Tsagari, D., & Floros, G. (Eds.), *Translation in language teaching and assessment* (pp. 3-22). Newcastle upon Tyne, UK: Cambridge Scholars Publishing.
- Mossop, B. (2007). *Revising and editing for translators*. Manchester: St. Jerome.
- O'Brien, S. (2002). Teaching post-editing: A proposal for course content. *Proceedings of the 6th EAMT Workshop on Teaching Machine Translation* (pp. 99-106).
- Plitt, M., & Masselot, F. (2010). A productivity test of statistical machine translation post-editing in a typical localisation context. *The Prague Bulletin of Mathematical Linguistics* 93: 7-16.
- Tezcan, A., Hoste, V., & Macken, L. (2019). Estimating post-editing time using a gold-standard set of machine translation errors. *Computer Speech & Language* 55: 120-144.
- Tsai, S. (2019). Using Google Translate in EFL drafts: A preliminary investigation. *Computer Assisted Language Learning* 32(5-6): 510-526.

<인터넷 자료>

- 김현정. (2021/06/24). 번역, 인간의 일? 기계의 일? The Science Times. Retrieved October 22, 2021 from <https://www.sciencetimes.co.kr/news/번역-인간의-일-기계의-일/>.

This paper was received on 27 October 2021; revised on 21 November 2021; and accepted on 3 December 2021.

Author's email address

layout81@naver.com

tandistudies@daum.net

About the authors

Sun-Woo Lee (first author) is a doctoral student at Hankuk University of Foreign Studies. She teaches translation in the Department of English for International Conferences and Communication.

Sang-Bin Lee is a full professor at Hankuk University of Foreign Studies. He teaches translation and interpreting in the Department of English for International Conferences and Communication.