

기계번역 포스트에디팅에 관한 해외 연구 동향

신 지 선
(이화여대)

1. 서론

기계번역의 전반적인 품질이 향상되면서 기계번역에 대한 업계와 학계의 관심이 뜨겁다. 번역을 의뢰하는 사용자들은 시간과 비용을 줄이고자 기계번역에 주목하고, 실제로 번역을 수행하는 번역사들은 시장에서 살아남기 위해 관심을 기울이지 않을 수 없다. 번역학자들은 연구와 실무가 불가분의 관계인 번역학의 특성상, 기계번역과 관련된 실무상의 여러 가지 변화를 간과할 수 없다.

기계번역 사용이 일반화되고 사용자 수가 더욱 확대되면 인간번역사의 지위 및 위상에 지대한 영향을 미칠 것이다. 그 기저에는 번역료 인하라는 경제적 측면이 자리 잡고 있다. 기계번역에 대한 정확한 이해가 없는 상태에서 그릇된 정보가 무분별하게 확산되면, 번역에 종사하는 많은 사람들에게 불이익이 돌아갈 수 있다. 뛰어난 실력을 갖춘 전문 번역사가 적절한 평가와 대우를 받지 못할 수 있고, 그 결과 번역 시장의 왜곡과 혼란을 초래하여 전반적인 번역 품질이 하향평준화로 치달을 수도 있다.

본 연구는 기계번역과 관련하여 다각적인 연구가 시급한 상황에 주목하여 현재 국내의 기계번역 연구 현황을 간략히 소개한 후, 연구가 전무(全無)하거나 미진한 분야에 집중하여 연구의 필요성을 강조하고자 한다. 기계번역의 장단점, 활용 가능성, 한계 및 문제점, 개선 사항 등을 정확히 파악하고 이해하기 위해서, 연구자가 유념해야 할 기계번역의 특징을 구체적으로 제시하는 것이 본 연구의 목표이다. 우선 기계번역 연구의 역사, 국내 번역학에서의 선행연구를 간략히 살펴보고, 기계번역 관련 연구가 오랜 기간에 걸쳐 광범위하게 진행된 해외의 사례를 참고하여 해외의 연구 동향을 고찰할 것이다. 그 과정에서 국내에서 아직 연구가 이루어지지 않은 주제 및 분야, 해당 연구를 진행할 때 유의할 사항, 관련 내용이 향후 학계와 업계에 시사하는 바를 모색해 보겠다.

2. 기계번역 관련 국내외 선행연구 분석

2.1. 기계번역 연구의 역사

기계번역 포스트에디팅에 관한 연구가 서구에서는 지금으로부터 이미 70여 년 전에 이루어졌다. 포스트에디팅 연구사(史)를 기술한 가르시아(Garcia 2012)의 논문에 따르면, “포스트에디터는 도착어 실력은 있으나 출발어는 전혀 모르는 사람이 할 수도 있다. 언어 전문가보다 주제 전문가가 더 잘할 수 있다. 포스트에디터가 얼마나 생산적으로 될 수 있는가는 기계가 얼마나 완벽한 번역물을 만들어내는지와, 사용자가 얼마나 완벽한 번역을 원하는지에 달려있다 (Yngve 1954: 21; Garcia 2012: 299에서 재인용)”라는 연구 결과를 찾아볼 수 있다. 국내에서는 최근에 고민하기 시작한 사항들에 대해 서구에서는 이미 1954년에 관련 분석이 이루어진 것이다.

1950년대 중반에 어떤 사람이 포스트에디터가 되어야 할지 논의하였고, 프리에디터는 출발어 전문가가, 포스트에디터는 주제 전문가가 잘할 수 있다고 가정하였으며, 품질평가에서는 언어적 등가보다 번역의 목적을 고려한 적합성을 기준으로 삼았다(Garcia 2012: 300). 또한 기계번역 평가와 관련하여 상대적으로 단순하면서도 정확도는 더 높은 평가방식을 개발하려고 시도하였고, 실험

을 통해 단일언어 구사자가 이중언어 구사자보다 번역 품질을 더 빠르게 잘 구분한다고 밝혔다(Carroll 1966: 55; Garcia 2012: 300에서 재인용). 이 외에 포스트에디팅의 이용 가능성을 고찰하기 위하여 인간이 번역한 결과물, 기계로 번역한 결과물, 기계번역 후 포스트에디팅한 결과물을 각각 비교하는 실험도 진행하였다. 그 결과, 정확성과 속도 면에서 인간번역이 제일 뛰어나고, 다음이 포스트에디팅 결과물, 마지막이 기계번역 결과물임을 보고하였다(Orr and Small 1967: 9; Garcia 2012: 301에서 재인용).

이상의 연구가 지금으로부터 60년 전인 1960년대에 진행되었다는 사실을 생각하면, 연구자들이 당시에 다음과 같은 결론에 도달했다는 점이 전혀 놀랍지 않다. “포스트에디팅 과정은 매우 힘들고 비용이 많이 드는 과정이다. 포스트에디터를 쓰기보다는 사용자가 기계번역 결과물을 스스로 읽도록 훈련하는 것이 더 효율적일 수 있다”(Orr and Small 1967: 9; Garcia 2012: 301에서 재인용).

기술의 한계로 기계번역의 생산성과 품질이 기대에 못 미치자 1970년대에는 기계번역 연구가 침체기에 접어들었다. 그러다가 1980년대에 기술의 발전으로 기계번역의 가능성을 알아본 대기업과 국제기구들이 연구에 관심을 두게 되면서, 연구의 주체가 대학에서 기업으로, 순수 연구자에서 이해관계 당사자로 바뀌게 되었다(Garcia 2012: 299). 연구 초창기에는 기계번역에 활용할 수 있는 신기술이 대두하기를 기다리며 연구자들이 기계번역의 가정 설정 및 실험 설계를 고민했다면, 이후로는 관심이 실용성에 쏠리며 사례 연구가 주를 이루게 되었다(Garcia 2012: 301).

1980년대에 진행된 몇 가지 연구를 소개하자면, 제록스에서 시행한 연구에서는 통제언어 사용의 중요성과 포스트에디팅 보조 툴의 필요성을 강조했고, 제너럴 모터스는 번역사 한 명의 1일 작업량과 포스트에디터의 생산성을 비교한 후, 스타일 측면의 수정은 하지 말 것을 지적하였다. 지멘스는 포스트에디터가 작업에 익숙해지는 시점까지는 생산성이 줄어들다가, 초기 단계(몇 개월에서부터 1년 이상까지)가 지나면 생산성이 현저히 높아진 조사 결과를 제시하였다. 유럽연합에서는 포스트에디팅 관련 에러를 관사나 전치사의 ‘마이너’ 에러, 문법이나 구문의 ‘메이저’ 에러, 스타일 문제인 ‘그레이’ 에러로 구분하였다가, 처리 속도가 관심사로 부각하자 에러를 다음과 같이 새롭게 구분하였다. 즉, 충실성 차원의 ‘수정 필요’, 가독성 차원의 ‘수정 가능’, 선호도 차이 측면의 ‘수정

불필요’이다. 한편 전미(全美)보건기구에서 진행한 연구에서는 이상적인 포스트에디터는 주제 전문가와 전문 번역사 가운데 후자일 것이나, 기계번역에 대한 번역사의 태도가 매우 중요하다고 강조하고 포스트에디팅에 익숙해지는데 필요한 훈련 기간은 30일 내외로, 연습할 분량은 10만 단어 내외로 추정하였다(Garcia 2012: 303-305).

가르시아(Garcia 2012 : 305)는 기계번역 연구가 처음 시작된 이후 50년간 논의된 사항을 다음과 같이 요약하였다. “언제 포스트에디팅이 필요한가? 이 텍스트가 애초에 번역 가능한 텍스트인지와 이 텍스트가 기계번역하기에 적합한 텍스트인지는 어떻게 측정할 수 있는가? 포스트에디팅은 어떻게 하는 것이며, 고품질용, 요약번역용 등으로 목적을 달리 할 것인가? 포스트에디팅은 누가 해야 하나? 전문 번역사인가? 아니면 주제 전문가인가? 이중 언어 구사자인가? 아니면 단일 언어 구사자인가?”

그간 해외에서 포스트에디팅에 관해 언제부터 어떠한 연구가 이루어졌는지 대략만 살펴봐도, 서구에서는 우리의 관심사에 대해 훨씬 오래전부터 연구하고 있었음을 알 수 있다. 이는 본 논문의 의의를 재확인할 수 있는 대목이기도 하다. 관련 선행 연구가 이미 많이 이루어졌다면, 관련 내용을 찾아 읽고 분석하여 국내 연구에 십분 활용해야 할 것이기 때문이다. 기계번역 품질이 날로 높아지고 사용자가 급증하는 이때, 무엇보다 시간 대비 효율적인 연구가 절실한 상황이다. 선행연구를 꼼꼼히 분석하면 향후 연구에서 시행착오를 줄이는데 크게 도움이 될 것이다. 이러한 필요성에서 본 논문에서는 해외 연구 동향을 고찰하여 국내에서 연구가 필요한 주제를 제안하고, 연구를 진행할 때 특히 유의할 사항을 구체적으로 제시해 보려고 한다.

2.2 국내외 선행연구의 주제 분석

그동안 국내에서는 기계번역에 관하여 어떤 연구가 이루어졌는지 연구 동향을 간략히 살펴보겠다¹⁾. 김순미 외(2019)는 2016년 말 신경망 기계번역 등장 전후로 국내 통번역 관련 주요 등재 학술지, <통역과 번역>, <통번역학연구>,

1) 국내의 기계번역 연구 동향을 고찰한 논문으로 본고에서 소개하는 두 논문 외에 원은하(2019)와 이상빈(2020)의 논문도 있음.

<번역학연구>, <통번역교육연구>에 인공지능과 기계번역 관련하여 실린 연구를 조사하여 주제별로 분류하였다. 총 32편의 연구를 대상으로 분류한 결과, 기계번역의 오류 유형과 인간번역과의 차이(9편)²⁾, 통번역의 변화하는 현황과 미래(8편)³⁾, 포스트에디팅 교육 관련 실험(7편)⁴⁾, 기계번역의 한계와 가능성(4편)⁵⁾, 구글, 파파고 등 엔진 성능 비교(3편)⁶⁾, 포스트에디팅의 기준(2편)⁷⁾, 윤리 담론(2편)⁸⁾으로 구분하였다.

한편 최문선(2019)은 번역학 학술지 창간 년인 1997년에서 2018년까지 국내에서 발표된 기계번역 관련 논문 총 49편을 수집하여 연구 주제를 중심으로, 크게 기계번역의 실체를 개괄적으로 조망한 연구, 기계번역의 생산 관점의 연구, 기계번역의 소비 관점의 연구, 세 가지 범주로 구분하였다. 기계번역의 실체를 개괄한 연구는 총 10편으로 주로 문헌 조사를 바탕으로 기계번역의 현황을 파악하고 관련 개념을 정립하며 미래의 전망을 개진한 연구들이고, 기계번역의 생산 관점의 연구는 총 8편으로 실제 연구자가 개발에 참여한 기계번역 시스템의 특징을 기술하거나, 기계번역 성능 개선을 위한 조사를 하고 그 효과를 검증하는 연구들이다. 마지막으로 기계번역의 소비 관점의 연구는 사용자 입장에서 기계번역 결과물이나 입력물에 대한 기술과 분석을 바탕으로 하는 연구로서 논문 수에서 가장 큰 비중을 차지하고 있다(최문선 2019: 285).

국내의 기계번역 연구 동향을 조사한 후, 최문선(2019: 284)은 기계번역이라는 주제와 이를 둘러싼 제반 환경에 대한 개괄적 연구가 다수 수행되었다는 점은 아직 연구가 충분히 축적되지 않은 상태에서 상호 학습의 과정을 거치는 연구의 초기 단계라는 점을 보여준다고 지적하였다. 한편 김순미 외(2019: 48)

2) 박영란 2017, 최효은, 이지은 2017, 한승희 2017, 광중철 한승희 2018, 서보현 김순영 2018, 이주리에 2018, 이준호 2018, 진용주 서보현 김순영 2018, 한현희 2018
 3) 이노신 이신재 이재영 이주희 2016, 이상빈 2016, 박지영 2017, 신지선 2017, 신지선 김은미 2017, 장애리 2017, 전현주 2017, 김순미 2018
 4) 김순미 2017, 이상빈 2017, 마승혜 2018a, 박혜경 2018, 이상빈 2018, 이성화 김세현 2018, 이준호 2018
 5) 마승혜 2018b, 마승혜 2018c, 박영란 2018, 이준호 2019
 6) 장애리 2017, 박영란 2018, 이주리에 2018
 7) 윤미선 김택민 임진주 홍승연 2018, 정혜연 2018
 8) 송연석 2018, 이승일 2018

는 특이한 점으로 국내에서 기계번역 포스트에디팅의 시간, 노력 기준 생산성 연구가 한 편도 없다는 점과, 연구가 기계번역 품질과 교육, 통번역의 미래에 치중되었음을 지적하였다. 최문선(2019: 286)에서도 국내의 연구 주제가 주로 언어적 완성도 측면의 평가라는 점을 언급하며, 기계번역의 장점은 대량의 결과물을 단시간에 산출해 낸다는 생산성에 있으므로 향후 생산성 연구가 필요함을 강조하였다.

이상에서 살펴본 국내의 연구 동향에 유념하며 그간 해외에서는 기계번역 포스트에디팅과 관련하여 어떠한 연구가 진행되었는지 고찰해보겠다. 우선 주제별로 크게 다음과 같이 열다섯 항목으로 분류하여 간략히 소개한 후, 이어지는 장에서 특히 주목해야 할 네 분야를 중심으로 자세히 살펴보겠다⁹⁾.

① 기계번역의 생산성 (인간번역과의 비교)

기계번역이 전통적인 인간번역 방식과 비교하여 생산성에 별 차이가 없다면 애초에 이렇게 뜨거운 관심과 우려의 대상이 되지 않았을 것이다. 번역과 연관된 모든 사람에게 최대 관심사는 기계번역의 생산성이 아닐 수 없다. 그간 해외에서 생산성 연구는 경제적, 시간적 차원에서 효율성이라는 측면에 집중하여 연구가 이루어졌다. 기계번역 후 포스트에디팅하여 ‘수용할만한’ 번역문으로 만들어내는데 총 소요되는 시간과 포스트에디터가 기울이는 노력을 측정하여 생산성의 구체적인 증감 정도를 분석하는 것이다(Garcia 2011).

② 포스트에디팅과 기존 번역과의 차이 (인지 과정 및 작업 방식)

기계번역에 대한 이해가 부족하면 포스트에디팅은 번역과 크게 다를 바 없다고 생각할 수 있다. 그리고 번역을 잘하는 사람이 포스트에디팅도 잘 할 것이라고 예상하기 쉽다. 그러나 포스트에디팅은 번역과 유사점보다는 차이점이 훨씬 많아서 번역과는 성격이 전혀 다른 작업으로 보아야 한다. 포스트에디팅을 할 때와 일반적인 번역을 할 때는 각각 인지적으로 상이한 노력이 필요하다. 따라서 포스트에디팅과 전통적인 인간번역의 차이를 다양한 측면에서 여러 가지 지표를 기준으로 비교 대조하는 연구가 많이 이루어지고 있다(O’Curran 2014).

③ 포스트에디팅 노력 구분 (소요 시간, 인지적 노력, 물리적 동작 측면)

9) 각 주제에 해당하는 선행연구가 매우 많아서 필자가 개인적으로 추천하고 싶은 논문 한 편씩만 제시하기로 함.

앞서 포스트에디팅은 작업 방식과 인지적 노력면에서 일반적인 번역과는 다르다고 언급했다. 여기에 주목하여 포스트에디팅을 할 때 그 과정에 들어가는 포스트에디터의 노력을 크게 세 가지로 나누어 각각을 고찰하는 연구가 많이 있다. 즉, 소요 시간, 인지적 노력, 키보드나 마우스를 움직이는 등의 물리적 동작 측면으로 구분하여 각 요소를 관찰하고 분석하는 것이다. 포스트에디팅을 하는데 시간은 얼마나 소요되는지, 인지적으로 어떠한 노력을 얼마나 기울여야 하는지, 에러를 수정하는 데는 얼마나 많이 클릭하는지 등을 실험 조사한다 (Vieira 2014).

④ 과정 중심 연구 (키스트로크, 아이트래킹, 스크린리코딩 등 활용)

이 유형은 바로 위의 포스트에디팅 노력 관련 연구를 진행할 때 실험에 이용하는 최신 툴에 초점을 맞추어 분류한 경우이다. 관련 테크놀로지가 비약적으로 발전하며 포스트에디터가 기울이는 노력을 과학적으로 측정할 수 있는 첨단 소프트웨어들이 속속 선보이고 있다. 기능이 업그레이드된 첨단 툴의 개발에 힘입어, 몇 년 전이라면 불가능했던 요소들을 관찰할 수도 있고, 관찰의 범위, 시간, 분량 등이 늘어나게 되었다. 과정 연구에 관심이 있다면 한층 정교해진 첨단 툴의 현황을 파악하고 꾸준히 업데이트하는 작업이 필수적일 것이다 (Vieira 2016).

⑤ 결과물의 품질 (포스트에디팅 노력과 결과물 품질 간의 상관관계)

포스트에디팅 노력을 고찰할 때 반드시 고려해야 할 사항이 바로 결과물의 품질이다. 포스트에디팅 과정에 들어간 노력과, 그 노력의 결과로 생성해낸 결과물의 품질 간에 어떤 상관관계가 있는지를 연구하는 것이다. 앞에서 포스트에디팅 노력을 세 가지 요소로 구분했는데, 시간, 인지, 물리적 동작, 이들 각각을 대상으로 소요된 시간과 품질 간의 관계, 인지적 노력과 품질 간의 관계, 물리적 동작 빈도와 품질 간의 관계 등을 분석한다. 또한 이러한 실험에 다양한 변수가 존재하므로 각각의 변수를 달리하여 여러 가지 연구가 가능하다(Arenas 2014).

⑥ 번역 메모리¹⁰⁾와 기계번역의 비교 (생산성, 품질)

기술 번역에서 컴퓨터 보조 번역(CAT, Computer Aided Translation) 툴은

이미 광범위하게 이용되고 있다. 통역번역대학원에서 CAT 수업이 개설될 정도로 번역사가 갖추어야 할 필수 능력의 일부가 되었다. 기계번역의 품질이 향상되며 기존의 번역 메모리와 최근의 기계번역 엔진을 비교하는 연구가 늘고 있다. 두 가지 가운데 무엇을 이용했을 때 생산성이 더 높은지, 그 때 각각의 품질은 어떠한지를 분석하는 연구이다. 최근에는 트라도스에서 기계번역 결과물을 바로 참고하여 번역할 수도 있어서, 트라도스, 기계번역, 기계번역을 통합한 트라도스, 세 가지를 비교하여 유용성, 생산성, 편의성 측면에서 분석한 연구 등도 있다(Escartin 외 2015).

⑦ 번역 단가 책정 (생산성과의 상관관계)

위의 연구와 불가분의 관계에 있는 것이 바로 번역 단가 책정의 문제이다. 생산성 연구를 업계에서 진행한다면 연구의 최종 목적은 적절한 번역료 책정의 문제로 귀결될 수 있다. 현재 가능한 과학적인 방법으로 생산성을 측정하여 가격 책정의 가이드라인을 제안하려는 연구가 많이 진행되고 있다. 연구자들은 타당성과 신뢰성을 갖추지 못한 연구 결과로 불합리한 단가가 책정되고 그 단가가 업계의 관행으로 자리 잡는 일이 발생하지 않도록 유의해야 할 것이다. 특히 번역 단가의 책정 기준이 생산성의 증감 정도이므로, 생산성을 측정할 때 유의할 사항을 명확히 인지하고 연구를 설계해야 한다.

⑧ 사용자 인터페이스 (각종 옵션, 선호도)

포스트에디팅 연구에서는 생산성, 효율성의 극대화를 목표로 하는 연구가 많다 보니, 번역사가 더욱 신속 정확하게 번역할 수 있도록 물리적인 환경을 개선하려는 연구도 많이 이루어진다. 그 가운데 사용자 인터페이스에 초점을 맞춘 연구들을 많이 찾아 볼 수 있다. 직관적으로 빨리 에러를 수정할 수 있는 인터페이스를 제공하기 위해, 기존의 인터페이스를 사용하며 불만인 사항을 조사하기도 하고, 다양한 옵션을 제공하여 결과를 관찰하기도 한다. 이러한 연구가 필요한 이유는 기계번역에서 발생하는 에러는 인간번역에서와는 전혀 다른 양상을 보이기 때문이다(Moorkens 외 2017).

⑨ 기계번역의 품질에 영향을 미치는 변수 (텍스트 종류 및 주제, 번역 방향, 언어 조합)

기계번역의 최종 품질은 포스트에디팅의 결과에 달려있고, 포스트에디팅의 결과는 기계번역이 만들어낸 최초 결과물의 품질에 달려있다. 즉, 기계번역의

10) 대표적인 예로 트라도스가 있음.

최초 결과물이 일정 수준 이상이면 포스트에디팅으로 적절한 번역물을 만들어 낼 수 있다. 그런데 기계번역이 인간번역과 비교하여 대체로 품질이 좋다 나쁘다 한마디로 평가하기 어려운 이유는 텍스트 주제, 번역 방향, 언어 조합 등에 따라 품질의 편차가 크기 때문이다. 텍스트가 어떤 종류의 텍스트로 주제는 무엇인가, 모국어로의 번역(BA 방향)인가, 외국어로의 번역(AB방향)인가, 사용자가 다수인 언어에서 소수인 언어로의 번역인가, 아니면 그 반대로의 번역인가에 따라 품질과 완성도가 천차만별일 수 있다. 그러한 이유로, 번역 품질에 영향을 미치는 다양한 변수를 달리하여 조사 분석한 연구들이 많다(Sánchez-Gijón 2014).

⑩ 원문의 언어적 요소가 미치는 영향 (통제언어, 프리에디팅)

기계번역이 고도화될수록 통제언어나 프리에디팅에 대한 관심이 커질 것으로 예상된다. 원문의 어떠한 요소가 기계번역의 품질에 직접적으로 영향을 미치는지 다양한 실험을 통해서 밝혀낸 후, 품질 저해 요소를 제거하거나 수정하여 품질 변화를 관찰하는 것이다. 기계번역 엔진이 출발어 텍스트를 이해하는데 문제가 없도록, 구체적인 어휘에서부터 통사적 특징, 관형적 표현, 화용론적 특성까지 고려하여 애초에 통제언어로 작성하거나, 아니면 프리에디팅을 거쳐 수정하는 방안에 대해 연구가 계속될 것으로 보인다(O'Brien 2006).

⑪ 번역평가 (자동평가방식)

기계번역 결과물을 평가하는데 인간 평가의 주관적인 특성을 개선하고, 기계번역 결과물 평가에 소요되는 시간을 단축하고자 자동평가방식에 관심이 쏠리고 있다. 자동평가 틀인 BLEU, GTM, TER, METEOR¹¹⁾ 등¹²⁾을 이용하여 평가한 결과를 분석하는 연구가 많이 진행되고 있다. 자동평가방식은 평가가 확실히 신속하게 이루어진다는 장점이 있으나, 평가의 정확성이나 적절성 측면에서는 한계가 있다. 인간의 평가가 주관적이라고 하지만, 자동평가 틀도 인간의 번역문을 참조 기준으로 삼아 어휘 수 비율이나 에디팅 횟수 등을 비교하는 방식이다. 따라서 자동평가를 실제로 이용하거나 관련 연구를 진행하고자 한다

11) BLEU(Bilingual Evaluation Understudy), GTM(General Text Matcher), TER (Translation Edit Rate), METEOR(Metric for Evaluation of Translation with Explicit Ordering)

12) 이 외에도 NIST, ROUGE, BLANC, ROSE, AMBER, LEPOR, PORT 등이 있음.

면, 자동평가방식의 과정 및 문제점, 한계 등을 명확히 파악하는 것이 무엇보다 중요하다(Doherty 2017).

⑫ 온라인 커뮤니티 사용자를 위한 포스트에디팅 연구

기계번역과 포스트에디팅에 관심이 많은 집단으로 전문번역사, 연구자, 언어서비스제공업체, 기업 등이 있겠으나, 기계번역의 중요한 사용자로 온라인 커뮤니티도 빼놓을 수 없다. 이러한 커뮤니티에는 특정 과학기술에 관심이 있어 각국의 관련 정보를 각자의 모국어로 올리는 전 세계 과학도들의 커뮤니티도 있고, 방탄소년단의 전 세계 팬들이 각 나라의 언어로 정보를 교류하는 커뮤니티도 있다. 이러한 다양한 온라인 커뮤니티에서 각종 글들을 일일이 번역 할 수 없으므로 기계번역 후 포스트에디팅을 거쳐 올리기도 하고, 회원들이 필요에서 자체적으로 기계번역을 이용하기도 한다. 번역학의 관심 주제인 클라우드 소싱 번역, 비전문 번역 등을 온라인 커뮤니티의 기계번역 포스트에디팅과 연계하여 고찰할 수도 있을 것이다(Mitchell 2014).

⑬ 포스트에디터 탐구 (프로필, 필요한 기술, 자질, 성향, 만족도 등)

포스트에디팅을 하는 주제인 ‘인간’에 대한 연구이다. 지금까지는 기계번역과 관련하여 주된 관심사가 기계번역의 생산성, 품질 변화, 급변하는 번역 환경 등이었다면, 이제는 포스트에디팅을 행하는 인간에 관심을 기울이기 시작하였다. 어떤 자질과 능력을 갖추고 어떤 교육(또는 훈련)을 받은 사람이 포스트에디팅을 효율적으로 잘 할 것인가에 대한 연구도 필요하겠지만, 그 이전에 보다 근본적인 인간의 문제를 심층 탐구해 보아야 할 것이다. 기계번역과 인간번역이 공존하는 시대, 관련 테크놀로지의 강압적 사용이 아닌 자발적 활용으로 가는 시대, 인간의 장점이 평가절하 되기보다 고평가되는 시대로 나아가기 위해 필요한 연구 분야라고 사료된다(LeBlanc 2017).

⑭ 교육 및 훈련 (적정 훈련 기간, 커리큘럼 제안)

시장에서 포스트에디팅 수요가 증가하면 교육현장에서도 이러한 상황을 간과할 수만은 없을 것이다. 국내 통번역대학원 교육 과정에 CAT 틀을 활용한 번역 과목이 포함되었다. 이제 한 걸음 더 나아가 기계번역 포스트에디팅을 정규 과목으로 편성하여 조만간 교육할 수도 있다. 앞서서도 언급했듯이 번역과 포스트에디팅은 전혀 다른 작업이므로 교육 내용, 교수법, 물리적 환경, 기자재, 교재 등 수업을 구성하는 모든 요소에서 번역 수업과는 다른 준비가 필요할 것

이다(Doherty 외 2014).

⑮ 산학 협력 방안 (기계번역 시스템 개발, 고품질 데이터 제공)

기계번역이 번역 서비스 제공자와 사용자 모두에게 만족스러운 방안으로 발전하기 위해서는 산학 협력이 중요한 사안으로 대두할 것이다. 해외에는 기계번역과 번역 툴에 관심이 있는 사용자, 연구자, 개발자들의 국제기계번역협회인 IAMT(The International Association for Machine Translation)가 1991년에 설립되었고 지역별로 산하에 EAMT(유럽기계번역협회), AMTA(미국기계번역협회), AAMT(아시아태평양기계번역협회)¹³⁾를 두고 각 지부에서 정기적으로 학회, 세미나, 워크숍 등을 개최하고 있다(Sanchis-Trilles 외 2014).

이외에도 EU의 지원 프로그램으로 진행된 CasMaCat 프로젝트¹⁴⁾를 일례로 보면, 대학과 산업체 간의 협력으로 인터랙티브 번역 예측이 가능한, 웹 기반, 오픈소스, 기계번역 지원 툴을 개발하는 성과를 거두었다. 기계번역에서 번역 엔진 개발을 담당할 공학 전문가와, 엔진 성능 향상에 핵심적인 고품질 병렬 데이터를 제공할 수 있는 번역전문가 간의 긴밀한 협력이 필요하다. 실무자와 연구자가 서로의 축적된 경험과 노하우를 공유하며 공동의 문제에 대처할 수 있도록 공론의 장을 활성화하고, 공동 프로젝트를 추진하는 방향으로 나아가야 할 것이다.

이상으로 기계번역 포스트에디팅과 관련하여 향후 연구가 필요한 분야를 열다섯 가지로 분류하여 간략히 소개해 보았다. 이 가운데, 관련 연구를 진행할 때 연구자가 특별히 유의할 사항을 강조하고 싶은 네 가지 주제에 대해서 자세히 살펴보겠다. 네 가지 주제란 번역의 생산성 비교, 포스트에디터 탐구, 결과물 평가, 포스트에디팅 교육 및 훈련으로 우선 생산성 비교에 관한 연구부터 검토해 보겠다.

3. 주요 연구 주제의 집중 탐구

3.1 생산성 비교

기계번역의 생산성은 기계번역 관련 연구에서 번역 생산자와 사용자 양측 모두가 촉각을 곤두세우고 있는 초미의 관심사이다. 기계번역이 인간번역보다 생산적인지, 생산적이라면 얼마나 생산적인지를 파악해야 번역작업에 기계번역을 사용할지 안 할지 결정할 수 있다. 또한 같은 원문을 두고 인간번역과 기계번역을 비교하여 기계번역의 생산성이 얼마나 향상되는지 알아야 번역 단가를 산정할 수 있다. 이처럼 기계번역의 생산성 문제는 번역에 드는 비용과 직결된 사안이어서 의뢰인, 번역업체, 전문 번역사, 연구자, 교수자, 학습자 모두에게 매우 중요한 문제이다.

김순미 외(김순미 외 2019: 42)에 따르면 점차 많은 언어서비스 제공업체들이 기계번역 포스트에디팅을 활용한 번역을 고객들에게 제공하고 있으나 아직 도입 초기 단계여서 여러 가지 문제가 불거지는 상황으로 가장 큰 문제는 번역료 책정의 어려움이라고 한다. 기계번역 포스트에디팅의 서비스 및 과금 방식이 구체적이고 세부적으로 정해지지 않아 업체마다 상황에 따라 다르게 단가를 책정하기도 하고, 고객들은 기계번역을 이용했다는 이유만으로 텍스트 난이도나 기계번역 엔진 성능과 특징 등 다른 변수를 고려하지 않은 채 단가를 낮추려고 하여 불이익을 당하는 경우가 있다고 한다.

해외에서는 기계번역의 생산성을 파악하는 연구가 많이 이루어졌으나, 연구 결과에 차이가 큰 경우도 많아서 참고할 때 주의가 필요하다. 생산성을 보여주는 지표의 하나인 번역시간의 경우, 번역시간의 증감 여부 및 비율이 다를 수 있고 번역결과물의 에러를 고찰하는 연구에서도 결과가 상이한 경우가 많다. 일례로 플리트 외(Plitt 2010)는 인간번역문에서보다 포스트에디팅 결과물에서 에러가 더 적다고 하고 그린 외(Green 2013)는 포스트에디팅으로 번역 품질이 전반적으로 향상되었으며 번역사의 생산성도 제고되었다고 밝힌 반면, 데프라테레(Depaetere 2014)에서는 포스트에디팅 결과물의 품질이 저하되었다는 결과를 제시하였다(Vieira 2017에서 재인용).

이렇게 연구 결과에 차이가 생기는 이유는 무엇일까? 그 이유는 바로 기계

13) EAMT(The European Association for Machine Translation, AMTA(the Association for Machine Translation in the Americas), AAMT(the Asia-Pacific Association for Machine Translation)

14) <https://cordis.europa.eu/project/id/287576/reporting>

번역 관련 연구에서 결과에 지대한 영향을 미칠 수 있는 중요한 변수들을 충분히 고려하지 않았기 때문이다. 이러한 변수에 유의하지 않고 생산성이 높아졌다는 연구만을 보거나 기계번역 결과물의 품질이 저하되었다는 논문만을 참고하게 되면, 기계번역에 대해 왜곡된 견해를 갖기 쉽다. 연구자나 실무자가 이러한 사항을 인지하고 있어야, 기계번역의 생산성을 실제 이상으로 과대평가하거나 과소평가하는 일이 발생하지 않을 것이다.

그렇다면 생산성 변화에 관한 연구를 진행할 때 반드시 유의해야 할 중요한 변수에는 어떤 것들이 있을지 생각해보자. 흔히 기계번역의 생산성을 파악하고자 할 때, 인간번역과 비교하여 총소요시간과 결과물의 품질을 기준으로 생산성이 높아졌다고 결론을 내리기 쉽다. 그런데 기계번역 결과물의 품질을 분석할 때 고려해야 할 사항은 다음과 같다.

첫째, 출발어 텍스트의 종류, 주제 분야, 난이도 등이다. 정보적 텍스트가 표현적 텍스트보다 번역 품질이 높을 수 있고, (어법상) 난이도가 높은 텍스트보다 난이도가 낮은 텍스트의 번역이 더 정확할 수 있다. 둘째, 분석대상이 되는 두 언어가 어떤 언어와 어떤 언어인가가 중요하다. 원문과 번역문의 병렬 데이터가 충분히 확보된 메이저 언어 간의 번역에서는, 사용자가 소수인 마이너 언어와의 번역에서보다 기계번역의 품질이 좋을 수 있다. 셋째, 어느 방향으로의 기계번역인가에 유의한다. A언어에서 B언어로의 번역인가, 그 반대방향으로의 번역인가를 따져본다. 넷째, 연구에 활용한 기계번역 엔진이 양질의 데이터로 훈련을 한 맞춤형 엔진인가, 누구나 자유롭게 사용할 수 있는 개방형 엔진인가에 따라 품질 차이가 크다. 다섯째, 번역 품질을 평가하기 위해 이용한 평가 방식이 인간 평가자의 전통적인 방식인가, 기계에 의한 자동평가방식인가를 살펴본다. 여섯째, 실험에 참여한 포스트에디터들의 프로필이 매우 중요하다. 출발어와 도착어 언어 실력, 포스트에디팅 경험 여부, 숙련된 정도, 기계번역에 대한 인식 등에 따라 작업의 효율성과 생산성이 크게 달라질 수 있기 때문이다.

다음으로 기계번역과 인간번역의 생산성을 비교할 때, 흔히 포스트에디팅을 하는데 걸린 총 시간을 기준으로 비교 분석한다. 즉, 기계번역이 번역 품질면에서는 인간번역과 별 차이가 없거나 더 좋으면서, 번역시간은 대폭 감소하여 생산성이 매우 높아졌다고 결론을 내리는 식이다. 그런데 여기에서 시간적 요소만을 고려해서는 기계번역의 생산성을 제대로 파악하기 어렵다. 산치스 트리에

스 외(Sanchis-Trilles 2014: 233)는 번역에 총 소요된 시간을 단순 비교하기 어려운 이유로 다음의 변수들을 지적한다. 즉, 실험 참가자 간의 번역 경험 차이(학생과 전문 번역사), 실험에 이용하는 기계번역 시스템의 차이, 특정 포스트에디팅 인터페이스에 익숙한 정도의 차이, 언어 조합 및 텍스트 주제, 데이터 수집과 필터링 방식의 차이가 작업시간에 크게 영향을 끼친다는 것이다. 포스트에디팅에 들어가는 노력은 단순히 측정할 수 있는 과정이 아니어서 시간적 요소만 가지고는 생산성을 분석할 수 없다. 따라서 포스트에디터의 인지적 노력을 연구하는 것이 필요하다.

포스트에디팅 과정의 인지적 노력과 관련하여 어떤 연구가 가능할 수 있을지 생각해보자. 예를 들어보면, ‘인지적 노력에 영향을 미치는 요인에는 어떤 것들이 있는가? 인지적 노력의 정도를 측정할 수 있는 지표에는 무엇이 있을까? 인지적 노력과 결과물 품질 간에는 상관관계가 있는가? 포스트에디팅 시간과 인지적 노력 간의 관계는? 인지적 노력을 조사할 때 어떤 연구방법들을 활용할 수 있을까? 아이트래킹과 키로깅을 지표로 삼았을 때 연구의 한계를 보완할 방안은 무엇인가? 언어조합 또는 번역방향에 따라 인지적 노력에 어떤 차이가 발생하는가? 포스트에디터의 포스트에디팅 경력에 따라 인지적 노력은 어떻게 달라지나? 인지적 노력을 줄이는데 적어도 훈련시간이 얼마나 필요한가? 초보 번역사와 전문 번역사, 전문 번역사와 전문 포스트에디터, 주제 전문가와 번역 전문가 가운데 인지적 노력이 더 필요한 측은 누구인가? 포스트에디팅 가이드라인은 인지적 노력에 어떤 영향을 미치는가? 기계번역에 대한 포스트에디터의 긍정적/부정적 인식과 인지적 노력 간에는 어떤 상관관계가 있나?’ 등 포스트에디팅 과정의 여러 가지 요소와 인지적 노력을 연계하여 다양한 연구가 가능할 것이다.

3.2 포스트에디터 탐구

기계번역을 논할 때 주로 작업의 생산성에만 관심이 집중된다. 이 모든 변화의 한 가운데 있으며, 변화의 영향을 가장 크게 받는 ‘인간’에 대한 진지한 논의는 찾아보기 어렵다. 인간에 관심을 둘 때는 기계번역과의 품질을 비교하거나 기계번역에 대한 인식 등을 조사하기 위해서인 경우가 대부분이다. 그러

나 점차 인간 측면의 문제가 부상할 것이다. 인간에게 초점을 맞춘 연구가 빠르면 빠를수록, 급변하는 환경 속에서 번역사나 사용자가 겪게 될 혼란과 시행착오를 줄이는 데 도움이 될 것이다.

이런 맥락에서 르블랑(LeBlanc 2017)이 번역 관행의 변화에 번역사들이 어떻게 반응하는지를 심층 조사한 연구는 우리 모두에게 시사하는 바가 크다. 번역 툴 사용과 관련하여 이전에는 번역 툴의 기술적인 문제를 주로 연구했으나, 르블랑(LeBlanc 2017)은 번역 툴 자체의 문제보다 비즈니스 관행의 변화가 더 문제라는 점을 강조하였다. 그는 캐나다의 대형 번역서비스업체 세 곳을 대상으로 테크놀로지가 번역과정에 통합되며 인간에게 어떤 영향을 미치는지를 조사하였다. 자율성과 만족도 조사를 중심으로 300시간 동안의 관찰 및 심층 인터뷰, 대규모의 설문 조사를 시행하였다. 본 연구는 향후 국내에서도 기계번역과 인간번역사가 협업할 때, 관계를 어떻게 정립할지에 대해 시사하는 바가 크므로 자세히 소개하고자 한다.

본 연구에서 조사 대상으로 삼은 세 번역업체는 텍스트 유형에 상관없이 거의 모든 번역에 툴을 사용한다. 각 업체는 생산성 가이드라인을 명확하게 규정하고 생산성 요건을 확실하게 부과하고 있다. 생산성 가이드라인이란 다음과 같다. 번역할 새로운 텍스트가 들어오면 관리자가 우선 번역 메모리와 비교 분석하여 이 번역에 들어갈 노력의 정도를 파악한 후에, 담당 번역사가 그 텍스트를 번역하는데 필요할 시간을 몇 시간 몇 분까지 구체적으로 명시하여 번역사에게 그에 따라 번역하게 하는 것이다. 번역과정에서 번역사들은 번역 메모리의 기존의 번역어를 재활용하도록 강요당하고, 기존의 번역어를 도저히 사용할 수 없는 경우에도 번역사가 직접 수정해서 안 된다. 매니저나 검수자에게 보고하여 수정 승인을 받아야 하며, 마음대로 고치는 경우가 반복되면 후에 나쁜 평가를 받을 수 있다. 가이드라인대로 재활용하는지 세밀하게 모니터 당하고, 명백한 오역이 아닌데도 스타일을 고치는 경우는 엄격히 금지된다(LeBlanc 2017: 49-51).

르블랑(LeBlanc 2017)은 번역사들과의 심층 인터뷰를 통해 작업방식의 변화에 대한 반응, 작업환경의 변화로 인한 효과, 전문가로서 번역사의 직업 만족도에 관해 고찰하였다. 번역작업에서 번역 툴 활용이 지나치게 강요될 때, 번역사들에게 어떤 일이 발생하는지, 그 결과 번역물에는 어떤 영향을 미치는지, 관련 인터뷰 내용을 구체적으로 소개하겠다.

① 작업방식 변화에 대한 반응

“어떻게 번역할까 더 시간을 들여 생각해보고 싶지만 그럴 수 없다.”

“처음에는 힘들었지만 이제는 완전 매치, 퍼지 매치를 보면 머리를 쓰지 않는다. TM 제안을 그냥 따른다. 왜냐하면 이런 단어에 신경 쓰는 것은 전혀 노력으로 카운트되지 않기 때문이다. 관리자는 전체적인 생산성을 높이는 데만 집중하라고 강조한다. 이것이 문제다. 당연히 품질에 영향을 미친다.”

“기계는 내가 번역하고 있는 텍스트의 난이도를 전혀 모른다. 그래서 번역이 얼마나 복잡한 과정인지 이해 못 한다. 완전 매치, 퍼지 매치라고 해서 좋은 품질과 동의어는 아니다.”

“최근에 보도 자료를 번역했는데, 완전 매치에서 엄청난 오역을 발견했다. 가이드라인에 따르면 완전 매치는 손대지 말아야 한다. 그러니 그 오역이 6개월 이상 그대로 있으면서 계속 사용된 것이다.”

② 작업환경의 변화에 대한 반응

“자율성 상실, 비전문적, 대량생산, 노동자, 만족도 저하, 품질 저하, 조립 라인, 단순 기술자, 조각 작업, 관리자와의 단절, 품질에 대한 무관심.”

“TM은 만능통치약으로 간주된다. 우리는 어휘 생산자에 불과하다.”

“품질은 아무도 신경 쓰지 않는 듯하다. 그저 생산성만이 중요. 번역기술자가 되어가고 있는가? 번역기술을 갖고 닦을 수는 있는가?”

③ 전문 번역사로서의 직업 만족도에 대한 반응

“우리는 단순히 세그먼트를 작업, 다른 사람의 번역을 위해 관리자가 요구하면 그대로 재활용할 뿐이다.”

“전문가로서 우리 번역사는 왜 새로운 작업가이드라인을 만들고 이행하는 과정에서 적극적인 역할을 하지 않는가? 우리는 번역전문가로서 중요한 피드백을 줄 수 있다. 툴 자체가 문제가 아니라 툴을 사용하도록 강요하는 방식에 문제가 있다.” (LeBlanc 2017: 51-57)

이상의 조사를 참고하면, 번역작업에 테크놀로지를 어떻게 적용하느냐에 따라 어떠한 결과를 초래할 수 있는지 가늠할 수 있다. 물론 위의 연구는 번역 툴을 사용하는 번역사들을 대상으로 하였으나, 포스트에디팅이 점차 일반화된다면 동일한 상황에 직면하게 될 것이다. “번역사가 일과 작업환경을 스스로 통제한다고 인식하면 테크놀로지를 자산으로 여기나, 테크놀로지 때문에 본인이 일에서 통제권을 잃었다고 생각하면 테크놀로지를 긍정적으로 보지 않는다 (Marshman 2012: 10; LeBlanc 2017: 59에서 재인용)”라는 연구 결과도 있듯이,

진정한 의미에서의 생산성 제고를 목표로 한다면 먼저 인간을 중심으로 고민해 보아야 할 것이다. 번역작업에 테크놀로지 이용이 불가피하다면, 처음부터 번역사가 능동적으로 통제할 수 있는 작업환경을 체계적으로 마련해 나가는 것이 중요하다.

그동안 포스트에디팅 연구에서 인간에 주목할 때는 생산성을 측정할 목적으로 포스트에디터의 인지적 노력, 마우스와 키보드를 물리적으로 움직이는 동작, 에러 수정을 위해 모니터 화면에 시선을 고정한 횟수 등을 분석하는 경우가 대부분이었다. 그러나 이제부터는 기계번역을 중심에 두고 주변부 인간의 역할을 분석할 것이 아니라, 인간을 연구의 중심에 두고 기계번역을 고찰해야 할 것이다. 번역작업에서 번역사의 자율성을 보장하고, 의사결정 과정에 번역사의 전문성을 반영하는 것이야말로 번역업체와 번역사가 상생의 길로 나아가는 첫 걸음일 것이다.

3.3 결과물 평가

기계번역 포스트에디팅 연구에서는 결과물 평가가 필요한 경우가 특히 많다. 기계번역의 생산성, 번역 메모리와의 품질 비교, 프리에디팅의 효과, 새로 개발된 번역 툴의 성능, 맞춤형 기계번역 엔진의 정확성, 포스트에디팅 능력 등 주제가 무엇이든 간에 어느 시점에서 결과물의 평가가 이루어진다. 평가가 정확해야 연구 결과가 유의미해지므로 평가를 제대로 하는 것이 매우 중요하다. 기계번역 평가는 전통적인 번역평가와 비슷한 점도 있지만 다른 점이 훨씬 많다. 그렇다면 연구자들이 기계번역을 고찰하며 결과물을 평가할 때 특히 유의할 점은 무엇인지 살펴보자. 특히 ‘평가자’와 관련된 사항을 중심으로 살펴보겠다.

포스트에디팅 결과물을 평가할 때 일반적으로 고려해야 할 사항은 다음과 같다. 기계번역 엔진의 특성, 포스트에디팅 가이드라인, 사용한 툴, 작업조건, 평가 단위, 분량, 평가 목적, 평가 기준, 평가자의 특성, 참여자 수, 참여자 프로필, 원문의 특성, 언어 조합, 번역 방향 등을 꼽을 수 있다.

평가 방식과 관련하여 유의할 사항으로 도허티(Doherty 2017: 140)는 “기계번역 품질 연구에서는 평가물이 맥락과 상관없이 제시되고, 평가를 문장 단위

로 하게 된다. 업계에서의 평가 방식은 무작위로 추출된 샘플을 평가하는 게 일반적이다. 따라서 평가자가 사전에 그와 같은 특수성을 숙지해야 한다. 평가자가 전체 텍스트를 보고 싶거나 이전에 평가한 문장으로 돌아가고 싶어도 허용되지 않는 경우가 많다. 또한 세그먼트를 어떤 순서로 보여주는가에 따라 평가에 영향을 미칠 수 있다”고 강조한다.

흔히 기계번역 연구에서 평가자가 명확히 드러나지 않는 경우가 많다. 평가자에 따라 연구 결과가 완전히 달라질 수 있으므로, 연구의 신뢰성을 높이기 위해서 전문성을 갖춘 평가자가 평가했음을 제시해야 한다. 물론 그 이전에, 기계번역을 평가하기에 적합한 평가자의 요건을 파악하여 거기에 맞추어 평가자를 선정해야 할 것이다. 도허티(Doherty 2017)는 평가 관련 연구를 진행할 때 유의할 사항으로 다음의 세 가지를 지적하였다.

첫째, 논문에서 평가자 관련 정보, 구체적 평가 가이드라인, 조작적 정의 등을 꼭 밝혀야 한다.

둘째, 평가자는 전문지식을 갖추고 있어야 하고 품질평가 연습을 해서 평가방식의 특수성을 이해하고 있어야 한다. 평가자로 학생을 이용하는 경우 특히 유의해야 하고, 기계번역 연구 커뮤니티에서는 연구에 평가 연습을 한 전문 번역사를 이용할 필요가 있다.

셋째, 기계번역에 대한 평가자의 긍정적 혹은 부정적 인식이 평가 결과에 영향을 미칠 수 있다(Doherty 2017: 140).

평가자 개개인의 전문지식, 사전교육, 인식도 중요하지만, 평가자가 여러 명일 때 평가자들 간의 일치도 중요하다. 번역평가가 본질적으로 매우 주관적인 작업이긴 하나, 생산성을 측정하는 중요한 실험에서 신뢰할 수 있는 평가 결과를 도출할 수 있도록 연구를 정교하게 설계해야 할 것이다.

3.4 포스트에디팅 교육 및 훈련

신지선(2019)은 번역 테크놀로지의 혁신적인 변화에 주목하여 번역 교육에서의 새로운 패러다임으로 ‘통합’ 패러다임을 제안하였다. 학생들이 졸업 후 변화된 환경에 잘 적응할 수 있도록 기존의 수업 구성요소에, 새로 습득이 필요한

요소들을 추가하거나 보완하는 통합적인 패러다임을 제시하였다. 기계번역과 관련해서는 수업에 테크놀로지 관련 콘텐츠 보강과 실제 연습 활동을 추가할 수 있을 것이다.

오브라이언(O'Brien 2002)은 번역 전공 학생들에게 포스트에디팅 기술을 가르쳐야 하는 이유로 다음과 같은 사항을 지적하였다. 첫째, 기계번역 시스템의 비용대비 효과는 포스트에디팅 과정의 효율성에 달려있다. 둘째, 포스트에디팅 기술은 연습을 충분히 했을 때 점진적으로 향상된다. 셋째, 번역을 잘한다고 해서 포스트에디팅을 잘 하는 것은 아니다. 넷째, 포스트에디팅 기술을 익혀서 졸업하면 기계번역 환경에서 곧바로 활동할 수 있다. 다섯째, 번역사가 기계번역의 방식과 기계번역의 장단점을 이해하는 데 도움이 된다.

이러한 필요에서 포스트에디팅 수업 경험을 토대로 수업 사례를 소개하는 연구가 많다. 학부에서 언어교육 목적으로 포스트에디팅을 지도하는 경우가거나 포스트에디팅 자체가 목적인 경우 등, 학습 목표가 무엇이나에 따라 다양한 선행연구들이 있다. 기계번역을 수업에서 다루지 않던 교수가 포스트에디팅을 가르쳐야 할 때 한 학기 동안 무엇을 어떻게 가르쳐야 할지 막막할 수 있다. 오브라이언(O'Brien 2002)은 포스트에디팅은 학생들이 경험과 자신감을 갖추었을 때 수업을 받는 것이 효과적이므로, 학부의 마지막 학기나 대학원 과정에서 지도하는 것이 적절하다고 밝히고 강의계획안을 제안한 바 있다. 그 내용을 표로 정리하면 다음과 같다.

〈표 1〉 기계번역 포스트에디팅 강의계획안(O'Brien 2002: 103-104)

이론 탐구	실무 연습
① 포스트에디팅 소개	① 포스트에디팅 실습
② 기계번역 테크놀로지 소개	② 다양한 기계번역 시스템으로 여러 유형의 텍스트를 포스트에디팅해보는 연습
③ 통제언어로 텍스트 작성하는 법 소개	③ 포스트에디팅의 목표를 달리하여 해보는 연습(참고용, 출판용 등)
④ 용어 매니지먼트 고급 과정	④ 기계번역 시스템의 설정을 달리하여 연습(특정 용어의 코드화, 언어규칙의 변경 등)
⑤ 텍스트 언어학 고급 과정	
⑥ 프로그래밍 기술 기초 과정	

위의 강의 내용을 보면 포스트에디팅 수업에서 다룰 이론적 측면이 제시되

어 있다. 이론적 측면에서 학생들에게 소개해야 할 가장 기본적인 내용에는 어떤 것이 있을지 생각해보자. 번역을 전공하는 학생들에게 포스트에디팅을 가르칠 경우, 첫째, 무엇보다 번역과 포스트에디팅의 차이를 작업 과정을 중심으로 이해시키고 둘째, 포스트에디팅에 필요한 기술과 능력을 번역 능력과 비교 대조하여 소개하며, 셋째, 포스트에디팅 결과물에 영향을 미치는 다양한 변수들을 강조하는 강의가 필요할 것이다.

한편 국내의 연구로 진실희(2019: 199)가 소개한 로컬라이제이션과 포스트에디팅 강의 일정에는 통번역산업의 기술적 전환, 기계번역의 특성 이해, 기계번역후처리 가이드라인, 기계번역후처리 실무, 라이트 포스트에디팅/풀 포스트에디팅 등이 포함되어 있다. 다음으로 학생들에게 부여할 과제와 관련하여 코포넨(Koponen 2015: 6)은 다양한 기계번역 엔진의 결과물 비교, 프리에디팅, 원문 없이 하는 포스트에디팅, 품질 수준을 달리하는 포스트에디팅, 기계번역 결과물의 평가 연습 등을 제안하였다.

이 외에 포스트에디팅을 지도할 때 교수가 유의해야 할 사항으로, 포스트에디팅 가이드라인을 효과적으로 제시하는 방법의 문제가 있다. 포스트에디팅은 주지하다시피 라이트 에디팅부터 풀 에디팅에 이르기까지 의뢰인의 요구에 따라 작업의 내용과 정도를 달리 할 수 있다. 모든 포스트에디팅의 목표가 출판이 가능할 정도의 최상의 품질인 것은 아니기 때문에, 품질 요건에 따라 구체적인 작업과정이 완전히 달라질 수 있다. 라이트 에디팅을 해야 할 때 ‘얼마나 라이트할지’에 대해 사용자와 번역사 간에 서로 이해한 바가 다르다면 나중에 문제가 될 수 있다. 따라서 포스트에디팅 의뢰를 받을 때, 의뢰인과 결과물의 품질 및 비용에 관해 구체적으로 논의하는 과정이 매우 중요하다. 실제로 작업이 시작되면 양측이 합의한 결과물의 품질 조건을 맞추기 위하여 명확한 포스트에디팅 가이드라인이 반드시 제공되어야 한다.

또한 학생들은 포스트에디팅을 연습할 때 어떤 항목을 어느 수준까지 수정해야 할지 몰라서 혼란과 좌절을 느낄 수 있다. 교수는 포스트에디팅 수업에서 학생들에게 구체적인 가이드라인을 반드시 제공해야 할 것이다. 가이드라인에서 포스트에디팅이 필요한 에러를 유형화한 후, 유형별로 실제 예시와 함께 제시해야 포스트에디팅 교육이 효율적으로 이루어질 것이다.

4. 결론

기계번역의 품질 향상으로 번역학계와 업계에서 기계번역 포스트에디팅에 대한 관심이 날로 높아지고 있다. 국내에서는 관련 연구가 진행된 지 그리 오래 되지 않았지만 해외에서는 이미 70여 년 전에 이미 기계번역에 대한 연구가 시작되었다. 기계번역 포스트에디팅과 관련된 연구 결과는 업계에서 번역 단가를 책정할 때 토대가 될 수 있어 연구에 한층 더 신중하게 접근해야 한다. 기계번역 사용에 대한 요구는 증대하는데 관련 연구가 아직 부족한 상황에 주목하여, 본 연구에서는 해외의 기계번역 연구 동향을 고찰하였다.

우선 해외의 기계번역 역사와 국내의 기계번역 연구 동향을 간략히 살펴본 후, 그간 해외에서 이루어진 관련 연구들을 크게 열다섯 가지 주제로 분류하여 개괄적으로 소개하였다. 열다섯 가지 주제는 기계번역의 생산성, 포스트에디팅과 기존 번역과의 차이, 포스트에디팅 노력, 과정 중심 연구, 포스트에디팅 노력과 결과물 품질 간의 상관관계, 트라도스와 기계번역의 비교, 번역료 책정, 사용자 인터페이스, 기계번역의 품질에 영향을 미치는 변수, 원문의 언어적 요소가 미치는 영향, 자동평가방식, 온라인 커뮤니티 사용자를 위한 포스트에디팅 연구, 포스트에디터 탐구, 교육 및 훈련, 산학 협력 방안 등이다.

이 가운데, 관련 연구가 필요한 주제이나 국내에서는 아직 관심을 기울이지 않았거나, 연구 결과가 신뢰성과 타당성을 갖추기 위하여 특히 연구 방법에 유의해야 할 분야를 네 가지 선정하여 집중 고찰하였다. 네 가지 분야는 생산성 비교, 포스트에디터 탐구, 결과물 평가, 포스트에디팅 교육 및 훈련으로, 이들 연구를 진행할 때 기계번역의 특징을 고려하여 유의해야 할 사항들을 중심으로 탐구해보았다. 앞으로, 기계번역에 관심 있는 학계와 업계, 연구자와 실무자, 공학자와 번역사가 서로의 전문지식과 경험을 공유하며 상호 협력할 수 있는 시스템이 구축되기를 기대해본다.

참고문헌

- 곽중철, 한승희 (2018) 「포스트에디팅 측정지표를 통한 기계번역 오류 유형화 연구」, 『통번역학연구』 22(1): 1-25.
- 김순미 (2017) 「신경망번역기(NMT) 활용 학부 번역교육의 가능성 연구」, 『통번역교육연구』 15: 5-37.
- 김순미 (2018) 「AI 시대 인간번역과 기계(NMT)번역의 공존-경영학 ‘확장(Augmentation)전략’ 중심」, 『통역과 번역』 20(2): 1-32.
- 김순미, 신호섭, 이준호 (2019) 「번역학계와 언어서비스업체(LSP)간 산학협력연구」, 『번역학연구』 20(1): 41-76.
- 김순영, 서보현 (2018) 「기계번역 결과물의 오류유형 고찰」, 『번역학연구』 19(1): 99-117.
- 마승혜 (2018a) 「한영 기계번역 포스트에디팅에 대한 경험적 고찰: 학부 교육과정 및 결과를 중심으로」, 『통번역학연구』 22(1): 53-87.
- 마승혜 (2018b) 「기계번역의 외연 확대 - 신문기사 특정 은유표현 포스트 에디팅 전략 모색」, 『번역학연구』 19(2): 117-145.
- 마승혜 (2018c) 「문학작품 기계번역의 한계에 대한 상세 고찰」, 『통번역학연구』 22(3): 65-88.
- 박영란 (2017) 「영한번역에서의 시제의 등가성과 기계번역과의 비교」, 『통번역교육연구』 15: 165-194.
- 박영란 (2018) 「기계번역에서의 영어속담 분석」, 『통번역교육연구』 16(1): 131-150.
- 박지영 (2017) 「언어기술 혁신과 통번역 산업의 미래: 20 대 글로벌 LSP의 통번역 서비스 분석을 바탕으로」, 『통번역학연구』 21(1): 137-68.
- 박혜경 (2018) 「석사 과정의 기계번역 수업에 대한 소고: 한일번역 전공생의 포스트에디팅 사례를 통하여」, 『번역학연구』 19(3): 163-193.
- 송연석 (2018) 기계번역 담론에 대한 비판적 고찰, 『번역학연구』 19(1): 119-145.
- 신지선 (2017a) 「테크놀로지 패러다임에서의 번역능력 재조명」, 『통번역학연구』 21(4): 51-71.
- 신지선 (2017b) 「포스트에디팅: 최적의 협업방식을 찾아서」, 『KIGO 소식지』 7:

- 38-53.
- 신지선, 김은미 (2017) 「인공지능 번역 시스템의 출현에 대한 소고」, 『번역학연구』 18(5): 91-110.
- 신지선 (2019) 「번역 교육의 새로운 ‘통합’ 패러다임 제안」, 『통번역교육연구』 17(2): 43-62.
- 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」, 『번역학연구』 19(5): 43-76.
- 이노신, 이신재, 이재영, 이주희 (2016) 「통번역의 미래지평 : 인공지능과 소통형 융합통번역 연구」, 『번역학연구』 17(2): 65-89.
- 이상빈 (2016) 「트랜스크리에이션, 기계번역, 번역교육의 미래」, 『통역과 번역』 18(2): 129-152.
- 이상빈 (2017) 「학부번역전공자의 기계번역 포스트에디팅, 무엇이 문제이고, 무엇을 가르쳐야 하는가?」, 『통역과 번역』 19(3): 37-64.
- 이상빈 (2018) 「학부 번역전공자의 구글 기계번역 포스트에디팅에 관한 현상학 연구」, 『통번역학연구』 22(1): 117-143.
- 이상빈 (2020) 「기계번역에 관한 KCI 연구논문 리뷰: 인문학 저널 논문(2011~2020년 초)의 논의내용과 연구방법을 중심으로」, 『통역과번역』 22(2): 75-104.
- 이성화, 김세현 (2018) 「영-한 및 한-영 기계번역 품질향상을 위한 프리에디팅 기법 제안」, 『번역학연구』 19(5): 121-154.
- 이승일 (2018) 「AI 번역의 속성 및 번역주체에 대한 논의」, 『통번역학연구』 22(4): 184-209.
- 이주리아 (2018) 「인공신경망 기계 번역의 한일/일한 번역 품질에 대한 예비연구 -품질향상 검토와 교열코드 적용-」, 『통역과 번역』 20(1): 32-71.
- 이준호 (2018) 「포스트에디팅 교육을 위한 포스트에디팅과 인간번역의 차이 연구」, 『통역과 번역』 20(1): 73-96.
- 장애리 (2017) 「국내 기계 통번역의 발전 현황 분석」, 『번역학연구』 18(2): 171-206.
- 전현주 (2017) 「4 차 산업혁명과 한국의 번역산업 현황 및 통번역 교육의 미래」, 『통번역교육연구』 15(3): 235-261.
- 정혜연 (2018) 「번역의 자동평가: 기계번역 평가를 인간번역 평가에 적용해보기」, 『통번역학연구』 22(4): 265-87.
- 진실희 (2019) 「4 차 혁명시대의 소프트웨어 융합 통번역교육 사례연구」, 『번역학연구』 20(4): 189-218.
- 진용주, 서보현, 김순영 (2018) 「프리에디팅 (pre-editing) 이 기계번역 품질에 미치는 영향 고찰: 유창성, 충실성, 가독성 비교를 중심으로」, 『통번역학연구』 22(3): 27-64.
- 최문선 (2019) 「국내 번역학 기계번역 연구 동향: 내용 분석과 키워드 분석을 중심으로」, 『언어학연구』 24(1): 275-297.
- 최효은, 이지은 (2017) 「특허 기계번역 결과물의 평가—KIPRIS 의 무료 한영기계번역을 중심으로」, 『통역과 번역』 19(1): 139-78.
- 한승희 (2017) 「기계번역·인간번역·트랜스크리에이션의 문체 비교: 광고 번역을 중심으로」, 『통번역학연구』 21(2): 163-188.
- 한현희 (2018) 「한노 기계번역의 문장 부호 처리 양상 및 오류 개선을 위한 실용적 함의: 프리에디팅 규칙 수립과 효용성의 관점에서」, 『통번역학연구』 22(3): 227-260.
- Arenas, Guerber (2014) ‘Correlations Between Productivity and Quality When Post-editing in a Professional Context’, *Machine Translation* 28(3-4): 165-186.
- Doherty, Stephen and Dorothy Kenny (2014) ‘The Design and Evaluation of a Statistical Machine Translation Syllabus for Translation Students’, *The Interpreter and Translator Trainer* 8(2): 295-315.
- Doherty, Stephen (2016) ‘The Impact of Translation Technologies on the Process and Product of Translation’, *International Journal of Communication* 10: 947-69.
- Doherty, Stephen (2017) ‘Issues in Human and Automatic Translate Quality Assessment’, in Dorothy Kenny (ed.) *Human Issues in Translation Technology*, London & New York: Routledge, 131-148.
- Escartín, Carla and Manuel Arcedillo (2015) ‘Living on the Edge: Productivity

- Gain Thresholds in Machine Translation Evaluation Metrics’, *Proceedings of 4th Workshop on Post-Editing Technology and Practice (WPTP4)*, 46-56.
- Garcia, Ignacio (2011) ‘Translating by Post-Editing: Is It the Way Forward?’, *Machine Translation* 25: 217-237.
- Garcia, Ignacio (2012) ‘A Brief History of Post-editing and of Research on Post-editing’, *Anglo Saxonica* 3(3): 291-310.
- Guerberof, Ana (2014) ‘Correlations Between Productivity and Quality When Post-Editing in a Professional Context’, *Machine translation* 28(3-4): 165-186.
- Koponen, Maarit (2015) ‘How to Teach Machine Translation Post-editing? Experiences from a Post-editing Course’, *Proceedings of 4th Workshop on Post-Editing Technology and Practice (WPTP4)*, 2-15.
- Krings, Hans (2001) *Repairing Texts: Empirical Investigations of Machine Translation Post-Editing Processes*, Kent: The Kent State UP.
- LeBlanc, Matthieu (2017) ‘I Can’t Get No Satisfaction!’ Should We Blame Translation Technologies or Shifting Business Practices?’, in Dorothy Kenny (ed.) *Human Issues in Translation Technology*, London & New York: Routledge, 45-62.
- Mitchell, Linda, Sharon O’Brien and Johann Roturier (2014) ‘Quality Evaluation in Community Post-editing’, *Machine Translation* 28(3-4), 237-262.
- Moorkens, Joss and Sharon O’Brien (2017) ‘Assessing User Interface Needs of Post-editors of Machine Translation’, in Dorothy Kenny (ed.) *Human Issues in Translation Technology*, London & New York: Routledge, 45-62.
- O’Brien, Sharon (2002) ‘Teaching post-editing: a proposal for course content’, *Proceedings of the 6th EAMT Workshop on Teaching Machine Translation*, 99-106.
- O’Curran, Elaine (2014) ‘Translation Quality in Post-Edited versus Human-Translated Segments: A Case Study’, *Proceedings of the 3rd Workshop on Post-editing Technology and Practice*, 113-118.
- Sanchis-Trilles, German, Vicent Alabau, Christian Buck, Michael Carl, Francisco Casacuberta, Mercedes García-Martínez, Ulrich Germann, Jesús González-Rubio, Robin L. Hill, Philipp Koehn, Luis A. Leiva, Bartolomé Mesa-Lao, Daniel Ortiz-Martínez, Herve Saint-Amand, Chara Tsoukala and Enrique Vidal (2014) ‘Interactive Translation Prediction versus Conventional Post-editing in Practice: A Study with the CasMaCat Workbench’, *Machine Translation* 28(3-4): 217-235.
- Vieira, Lucas (2014) ‘Indices of Cognitive Effort in Machine Translation Post-editing’, *Machine translation* 28(3-4): 187-216.
- Vieira, Lucas (2017) ‘From Process to Product: Links Between Post-editing Effort and Post-edited Quality’, in Lykke Jakobsen, A. & Mesa-Lao (eds.) *Translation in Transition: Between Cognition, Computing and Technology*, Amsterdam: John Benjamins Publishing Company, 162-186.
- <인터넷 자료>
- O’Brien, Sharon (2006). ‘Controlled Language and Post-Editing’, Available at <https://216.18.156.115/multilingual/downloads/screenSupp83.pdf>

[Abstract]

Trends of Overseas Research on Machine Translation Post-editing

Shin, Jisun

(Ewha Womans University)

This paper aims to identify the topics and issues for further exploration in machine translation research. In order to provide insights into numerous challenges machine translation poses, this study analyzed a significant number of research papers published abroad and presented fifteen research topics. They include: post-editing productivity, post-editing vs. traditional human translation, post-editing efforts, process-oriented research, correlations between productivity and quality, translation memory vs. machine translation, post-editing rates, design of post-editing interface, variables affecting the quality of machine translation, the impact of linguistic elements of source text, automated machine translation evaluation, role of post-editors, post-editing in translator training, educational academia-industry cooperation. Among them, four research topics were selected for detailed discussion about the difficulties faced by translators and researchers at home and abroad.

▶ Key Words: machine translation, post-editing, research topics, productivity, post-editing efforts

▶ 주제어: 기계번역, 포스트에디팅, 연구주제, 생산성, 포스트에디팅 노력

신지선

이화여자대학교 통역번역대학원 번역학과 부교수

jisunshin@ewha.ac.kr

관심분야: 기계번역, 포스트에디팅, 번역교육, 번역평가

논문투고일: 2020년 9월 6일

심사완료일: 2020년 9월 22일

게재확정일: 2020년 10월 8일