

## 기계번역의 오류 양상과 사후수정(post-editing)○] L2 쓰기 적절성에 미치는 영향\*

신동광\*\* · 전유아\*\*\*  
교수, 광주교육대학교 · 교수, 한양대학교

---

Shin, Dongkwang & Chon, Yuah Vicky (2022). Machine translation errors and the effect of post-editing on the appropriateness of revised L2 texts. *Journal of the Korea English Education Society*, 21(1), 121-140.

Although machine translation (MT) such as *Google Translate* (GT) may be used by second language (L2) learners to solve lexical and syntactic problems for L2 writing, the use of MT requires L2 learners to notice errors in MT output and post-edit MT errors. The purposes of the study were to 1) analyze the types of MT errors; 2) examine how post-editing of MT errors would have an effect on sentence adequacy scores of post-edited texts; and 3) examine any differences between the MT output and post-edited texts on measures of lexical and syntactic complexity. To this end, 57 L2 university learners were asked to compose a text in L1, translate it with GT, and post-edit the text for any mistranslations, ungrammatical sentences, or incomplete expressions. The MT output produced a total of 931 sentences of which 600 sentences (64.5%) were without any errors. Three hundred and thirty-one sentences had errors in the areas of mistranslation (65.56%), ungrammaticality (16.11%), missing word (16.94%), and extra words (1.39%). When the effectiveness of post-editing was examined, the total mean of sentence adequacy scores was found to have improved significantly in comparison to those of the machine translator. Examination of lexical and syntactic complexity measures indicated an increased number of sentences but a drop in the number of coordinate phrases in the post-edited texts. The results are discussed as to how MT can be used to help learners improve their L2 writing skills.

[machine translation/error types of MT/post-editing/Google Translate/기계번역/기계번역 오류 유형/사후수정/구글 번역]

---

\* ○] 논문은 한양대학교 교내연구지원사업으로 연구되었음(HY-2019년도).  
\*\* 제1저자  
\*\*\* 교신저자

## I. 서론

최근 OpenAI가 만든 인공지능(AI) 언어 예측 모델인 GPT-3(Generative Pre-trained Transformer 3)가 언어 교육 분야에서 주목을 받고 있다. 인간의 언어 수준에 가까운 언어 능력을 가진 것으로 평가되는 GPT-3의 출현은 대화 파트너로서 소셜 챗봇(social chatbot)의 기능을 수행할 수도 있고, 기계번역이나 소설을 창작하는 등의 과업을 수행할 수 있다. 이는 언어 교육의 입장에서는 상당한 위협될 것이라는 우려의 목소리도 있다(Brown et al., 2020). 그러나 또 다른 한편에선 GPT-3의 기계학습 과정이 대규모 데이터에 지나치게 의존하고 있어 천문학적인 비용이 소요되고 맥락적 상황 인식의 한계로 인해 인간과 같은 사고 과정에 이르기는 쉽지 않을 것이라는 지적도 있다(Nair & Bashir, 2020).

이러한 상반된 시각을 언어 교육의 측면에서 정리하면, 후자는 AI가 결코 인간을 능가할 수는 없을 것이고 AI의 도움을 받는 언어활동 또한 고유한 언어능력으로 간주하기 힘들다는 입장이라 볼 수 있다. 이에 반해, 전자는 AI를 활용하여 구사하는 언어활동을 새로운 언어능력으로 간주할 필요가 있으며 현재의 AI 기술 수준에서도 언어교육에 유용하게 활용할 수 있다는 입장이다.

기계번역기를 활용하는 맥락에서도 이와 동일한 논란이 일고 있다. 학습자가 기계번역기에 과도하게 의존하게 될 경우, 특히 초보 L2 학습자는 자신의 언어 능력 향상을 위한 노력을 소홀히 할 수 있다고 우려를 제기한 연구(Ahn & Chung, 2020; Kol et al., 2018)가 있는 반면, L2 쓰기 활동에서 기계번역기 활용의 장점을 제시하고 있는 연구(임희주, 2017; Alhaisoni & Alhaysony, 2017; Kim, 2018; Lee, 2020; Tsai, 2019) 또한 다수 존재한다.

이러한 논란에도 불구하고 최근 인공지능 활용 교육(Artificial Intelligence in Education, AIED)이라고 불리는 학습 모델은 언어 교육의 주요 트렌드로 자리를 잡아가고 있다(Hwang et al., 2020; Pokrivcakov, 2019). 실제로 기계번역기가 아니더라도 L2 쓰기에서 코퍼스, 사전, 문법 자동 교정기와 같은 참고 자료를 활용하는 쓰기 모델은 꾸준히 제안되고 있다(예, Daniels & Leslie, 2013; Gaskell & Cobb, 2004; O'Sullivan & Chambers, 2006; Shin et al., 2021). 이는 21세기에 요구되는 쓰기 능력에 대한 재정의(redefining)의 필요성을 의미하는 동시에 교사 역할 제고에 대한 필요성 또한 시사하는 것이다.

이에 본 연구는 AIED 도구로서 기계번역기의 활용에 초점을 두고 기계번역의 오류 양상과 그 원인을 분석하고, 학습자의 사후수정(post-editing)이 적절한 목표 표현의 구사에 실제 기여하였는지 살펴보고자 한다.

## II. 이론적 배경

### 1. 기계번역의 오류 분석

기계번역의 정확도가 점점 향상되고 그에 따라 기계번역기의 활용도 증가하고 있지만 기계번역은 아직도 다양한 오류를 포함하고 있다(Chon & Shin, 2020). 기계번역 오류에 대한 분석은 Vilar 외(2006)의 연구로 거슬러 올라간다. 초창기 기계번역은 현재의 신경망 기반(neural network machine translation)이 아닌 통계 분석을 통한 규칙 기반 방식(rule-based machine translation)을 채택하였고 번역 시 신경망 방식에 비해 상대적으로 더 많은 오류가 발생했다. Vilar 외의 연구는 통계 기반 방식의 기계번역에서 나타난 오류 분석에 초점을 두었다. 이 연구의 오류 분석에서는 기계번역의 오류를 ‘어휘 누락 오류,’ ‘어휘 사용 오류,’ ‘어순 어류,’ 그리고 ‘구두점 오류’로 분류하였다.

Ferris(2006)는 Vilar 외(2006)의 오류 분류 기준을 보다 세분화하였다. 현재 기계번역의 오류 분석에 대한 다수의 연구는 Ferris의 기준을 참고로 하고 있다(윤채원, 2021). Ferris가 제시한 15개의 오류 유형은 ‘어휘 선택 오류,’ ‘동사 시제 오류,’ ‘동사 형태 오류,’ ‘어휘 형태 오류,’ ‘주어-동사 일치 오류,’ ‘관사 오류,’ ‘명사의 접미사 오류,’ ‘대명사 오류,’ ‘문장 연속 오류,’ ‘문장 단절 오류,’ ‘구두점 오류,’ ‘철자 오류,’ ‘문장 구조 오류,’ ‘비공식적 문체 오류,’ ‘숙어 오류’이다.

Daems 외(2015)의 연구에서는 기계번역 오류의 유형뿐만 아니라 오류 빈도수 까지 분석하였고 그 결과, 대부분의 오류는 ‘문법 및 구문 오류,’ ‘표현의 적절성 오류,’ ‘번역의 일관성 오류,’ ‘어휘 사용 오류,’ ‘철자 및 스타일 오류’ 순으로 나타났다. 이 밖에도 구글 번역(Google Translate)을 활용한 연구 중 하나인 Groves와 Mundt(2015)의 연구는 구글 번역가 말레이어나 중국어를 영어로 번역 할 때 상당한 정확성을 보였고 대부분의 기계번역 오류는 ‘문장 구조’와 ‘어휘 선택’에서 발생했다고 주장하였다.

가장 최근의 연구 중 하나인 윤채원(2021)의 연구에서는 구글 번역을 활용한 기계번역의 오류를 분석하였다. 그 결과를 살펴보면, 한글 원문에 대한 인식의 문제에서 발생한 ‘오역 오류’가 가장 많았고, ‘문법 오류,’ ‘구 수준 오류,’ ‘문장 수준 오류,’ ‘어순 오류’ 순으로 오류의 빈도가 높게 나타났다. 또한 기계번역기는 주어 누락, 고유명사, 동음이의어 등이 한글 원문에 포함되었을 때 이를 제대로 번역하지 못하는 것으로 나타났다. 이는 기계번역기를 효과적으로 활용하기 위해서는 먼저 언어 간의 차이를 인지하는 것이 중요하며 이를 위한 체계적인 교육이 필요하다는 것을 시사한다.

### 2. 기계번역의 사후수정

앞서 살펴본 바와 같이 기계번역이 여전히 오류를 포함하고 있기는 하지만

기계번역기의 활용은 사용자의 문법 오류 등을 줄이는 데 유용한다는 의견도 있다(Lee & Briggs, 2021). 하지만 기계번역기를 보다 효과적으로 사용하기 위해서는 결국 기계번역의 오류를 학습자들이 인지하여 수정하는 사후수정의 과정이 반드시 필요하다는 결론에 이른다.

Bowker와 Ciro(2019)에 따르면 사후수정(post-editing)은 기계번역의 초벌을 언어 전문가가 수용 가능한 수준으로 오류를 수정하는 활동을 의미한다. 이들의 연구에서 사후수정의 의미는 기계번역기의 활용 주체를 언어 전문가로 한정하고 있지만 최근의 추세는 기계번역기의 활용 주체가 일반 사용자로 일반화되고 있다. Kol 외(2018)와 Lee와 Briggs(2021)의 연구에서는 실제 기계번역의 오류를 수정하는 활동을 L2 교수학습의 방법으로 활용한 사례를 제시하였다.

이준호(2018)의 연구에서는 기계번역기에 기반한 사후수정의 과정에서 나타난 학습자들의 특정한 경향성을 제시하였다. 즉, 학습자들은 기계번역의 초벌 번역이 적절하다고 판단하게 되면 어떤 수정도 없이 그대로 수용하는 경향을 보였다. 반면 오류나 만족스럽지 못한 표현을 인지했을 경우에는 ‘교체,’ ‘추가,’ ‘삭제’의 사후수정 유형 순의 선호도를 보였다. 또한 간단한 수정이 요구되는 경우에는 ‘추가’ 또는 ‘삭제’를 선호하였으나 문장을 전반적으로 수정해야 하는 경우는 수정보다는 문장을 새로 재작성하는 편을 택했다.

이윤재와 이동주(2020)는 고등학생들을 대상으로 기계번역기 활용의 효과성을 검증하기 위해 일반 영어 쓰기와 기계번역기 활용 영어 쓰기의 오류 발생 빈도를 비교·분석하고 기계번역기에 대한 학습자들의 활용 인식을 조사하였다. 이 연구가 기존 기계번역의 사후수정 연구와 다른 점은 기계번역기를 전체 한 글 원문 번역이 아닌 쓰기 활동의 참고 도구로 활용하였다는 점이다. 이들의 연구 결과를 살펴보면, 일반 영어 쓰기에서는 ‘관사 오류’가 전체 오류의 22.6%를 차지했고 ‘구두점 오류(15%),’ ‘어휘 선택 오류(4.4%),’ ‘동사 형태 오류(10.9%)’ 등이 그 뒤를 이었다. 반면 번역기 활용 영어 쓰기에서는 ‘문장 구조’와 ‘어휘 선택 오류’가 각각 18.6%를 차지해 가장 높은 빈도를 보였고 ‘구두점 오류(14.6%),’ ‘관사 오류(14.2%)’ 등의 오류가 그 다음 순위에 올랐다. 전체적으로는 기계번역기의 활용이 학습자의 쓰기 오류의 수를 줄이는 효과가 있는 것으로 나타났다. 학습자들은 대체로 단순한 문법 수정의 참고 자료로 기계번역기를 활용하였고 이는 상위권 학생들이 기계번역기를 보다 효과적으로 사용한다는 기존 연구들(Chung, 2020; Kol et al., 2018)과는 달리 영어 능숙도가 낮은 학습자 집단에서 보다 두드러진 효과를 보였다. 다만 이윤재와 이동주의 연구에서는 두 가지 다른 주제의 글쓰기를 비교하여 동일한 주제에 대하여 기계번역기 사용이 쓰기결과에 미치는 영향을 정확히 판단하는 데는 한계가 있었다.

위에서 살펴본 윤채원(2021)의 연구에서는 기계번역기의 오류 유형뿐 아니라 학습자들의 사후수정에 대한 양상 또한 분석하였다. 이 분석의 주요 결과로, 먼저 연구에 참여한 중학생들은 이윤재와 이동주(2020)의 연구와 마찬가지로 기계번역기를 주로 문법 표현의 정확성 여부를 확인하기 위해 사용한다고 응답하

였다. 또한 학습자들은 자신들의 문법 지식에 대해 대체로 확신이 없는 것으로 나타났다. 실제 시제, 태(voice), 관사, 수일치 등 한글 작성에서 민감하지 않은 요인과 관련된 기계번역의 오류일수록 학습자들의 오류 인식률은 낮았고 사후 수정을 통해서도 이를 교정하지 못했다. 그럼에도 불구하고 연구에 참여한 중 학생들은 자신들이 예측한 수정의 정확도에 비해 높은 사후수정의 정확도를 보여 학습자들의 자신감을 키워주는 등의 정의적 요인도 기계번역기의 활용에 중요한 조건이 될 수 있다는 것을 시사했다.

지금까지 살펴본 사후수정 능력은 PACTE Group(2005)이 제안한 ‘도구적 하위 역량(Instrumental Sub-competence)’라고 규정할 수 있으며 이는 번역기와 같은 디지털 도구 활용을 언어능력의 한 부분으로 인정할 필요가 있다는 것을 의미한다. 하지만 이러한 사후수정 전략은 주로 번역학을 중심으로 연구가 이루어졌을 뿐 언어교육의 분야에서는 아직 연구가 미진한 상태이다. 이와 같은 맥락에서 Jia 외(2019)는 기계번역기를 보다 효율적으로 활용하여 더 향상된 쓰기 결과물을 산출하기 위해서는 체계적인 사후수정 교육이 필요하다고 주장하였다.

본 연구는 기계번역의 언어교육적 활용 측면과 관련하여 다음과 같은 연구 질문을 설정하였다.

첫째, 기계번역의 주요 오류 유형 및 분포는 어떠한가?

둘째, 기계번역의 사후수정은 쓰기의 적절성에 어느 정도 영향을 미치는가?

셋째, 기계번역의 사후수정은 쓰기의 어휘 및 통사적 특성 변화에 어느 정도 영향을 미치는가?

### III. 연구 방법

#### 1. 연구 대상

본 연구에서는 2020년 2학기 영어작문 과목을 수강한 호남권 1개 대학( $n = 27$ )과 수도권 1개 대학( $n = 30$ )에 재학 중인 2-3학년 57명의 L2 학습자가 참여하였다. 호남권 학습자들은 교육학을 전공(이하 ‘영어 비전공 학습자’)하였고 영어 수준은 중·상급에 해당하였고 수도권의 학습자들은 영어교육 전공자(이하 ‘영어 전공 학습자’)로 영어 수준이 상급에 해당되었다(아래 표 1 참조). 실험을 위한 모든 쓰기 과업은 강의 중 수행되었다. 먼저 영어 쓰기 능숙도를 평가하기 위해 “아침 식사를 하는 습관과 거르는 습관에 대한 장점 중 자신의 입장을 선택하여 서론을 쓴 후, 세 가지 근거와 함께 결론을 쓰시오. (300단어 내외)”라는 문항으로 1시간 동안 사전 평가를 실시하였다. 채점에는 신동광 외(2012, p. 188)가 개발한 ‘자기의견쓰기’ 유형에 대한 채점 기준표를 사용하였다. 이 채점 기준표는 ‘과제완성,’ ‘내용,’ ‘구성,’ ‘언어사용’ 4개의 영역을 포함한 분석적 채

점 기준표이며 최종 점수는 5점 척도를 적용한 4개 영역의 평균인 5점을 만점으로 하였다. 채점은 연구자들이 참여하여 복수채점으로 수행하였고 채점 간 신뢰도는 피어슨 상관계수(Pearson's Correlation)를 기준으로  $p < .01$  수준에서 .895로 높게 나타났다. 다음의 표 1은 두 집단의 쓰기 평균 점수를  $t$ 검정을 통해 분석한 결과이다.

**표 1 사전 평가를 통한 두 실험 집단의 영어 쓰기 능숙도 비교**

	M	SD	t	df	Sig.
영어 전공 학습자 (n = 30)	4.387	0.272			
영어 비전공 학습자 (n = 27)	4.077	0.286	4.190	55	.000***

Note. \*\*\* $p < .001$

영어 전공자 집단의 경우 평균 4.387로 평균 4.077의 영어 비전공자 집단보다 유의한 수준( $t = 4.190, p < .001$ )에서 더 높은 영어 쓰기 능숙도를 보였다.

## 2. 연구 도구

### 1) 쓰기 과업

학습자들은 ‘한글 원문 작성하기’ → ‘기계번역기로 초벌 번역하기’ → ‘초벌 번역물 사후수정하기’의 3단계로 구성된 쓰기 과업을 강의 시간 중 1시간 동안 수행하였다. 1단계에서는 한국인 학습자에게 친숙한 주제 중 하나로 인터넷 사용의 장점 또는 단점 중 한 가지 입장을 선택한 후 자신을 주장을 뒷받침하는 세 가지 이유를 포함하여 한글로 300단어内外로 작성하게 하였다. 학습자들의 인지적 부담을 줄이고 언어능력의 측정에 초점을 두기 위해 세 가지 근거 중 두 가지 이유는 키워드로 제시하였다. 2단계에서는 작성된 한글 원본을 구글 번역(<https://translate.google.com/?hl=ko&tab=TT>)을 통해 영문으로 전환하였다. 마지막 3단계에서는 기계번역물에 나타난 미흡한 표현, 오역, 문법적 오류 등을 찾아 학습자들이 각자 수정하여 제출하도록 하였다.

### 2) 기계번역의 오류 분석과 기계번역 및 사후수정의 적절성 평가 기준

기계번역에 나타난 오류의 종류와 빈도를 살펴보기 위해 먼저 기계번역의 초벌 샘플을 바탕으로 기존 연구들(예, 윤채원, 2021; Groves & Mundt, 2015; Vilar et al., 2006)을 참고하여 다음의 표 2와 같이 오류 유형을 확정하였다.

표 2  
기계번역 오류 유형(error scheme)

오류 유형	코드	예문
어휘 누락 (word missing)	OS	인터넷을 연결해보면 할 수 있는 것들이 많은데 (정보를 검색하기, 이메일 보내기, 지난 tv프로그램 시청하기 등) 다양한 활동을 집에 앉아서 몇 번의 클릭만으로 즐길 수 있어 우리의 삶의 질이 향상됐다고 본다면 <u>그건 오산이다.</u>
	MW MS	If you connect to the Internet, there are many things you can do (search for information, send an email, watch past tv shows, etc.) and enjoy a variety of activities at home with just a few clicks. Ø
오역 (Mistranslation)	CS	If you connect to the Internet, there are many things you can do (search for information, send an email, watch past tv shows, etc.) and enjoy a variety of activities at home with just a few clicks. And we may think of it as improving the quality of our life, but <u>that is a mistake.</u>
	MT MS	이렇듯 인터넷은 사람들로 하여금 사이버 폭력에 <u>가해자 혹은 피해자가 될 수도 있는</u> 위험한 상황에 노출시킨다.
비문법성 (Ungrammaticality)	OS	As such, the Internet exposes people to dangerous situations <u>that could be victims</u> or victims of cyberbullying.
	CS	As such, the Internet exposes people to dangerous situations <u>in which they can become perpetrators</u> or victims of cyberbullying.
불필요한 어휘 (Extra Word)	UG OS	인터넷은 우리에게 새로운 여가생활을 창조할 수 있도록 도와줬다
	MS	The Internet <u>helped</u> us to create new leisure life.
	CS	The Internet <u>has helped</u> us create a new leisure life.
	OS	하지만, 인터넷의 도입으로 인해 단 몇 초 혹은 몇 분 만에 서로의 안부를 물거나 하는 일상적인 대화가 가능하고, <u>심지어는</u> 급박한 상황에서도 빠르게 연락을 취할 수 있기 때문에 도움이 되는 경우가 많다.
	EW MS	However, the introduction of the Internet is often helpful because it allows you to talk to each other in just a few seconds or minutes, and <u>even</u> get in touch quickly, <u>even</u> in urgent situations.
	CS	However, the introduction of the Internet is often helpful because it allows you to talk to each other in just a few seconds or minutes, and get in touch quickly, <u>even</u> in urgent situations.

Note. OS: 원본 문장, MS: 기계번역 문장, CS: 올바르게 수정된 문장

‘어휘 누락’은 한글 원본의 표현 가운데 기계번역 중 누락된 표현이 있을 경우를 의미하고 ‘오역’은 한글 원본에서 의도한 의미와 차이가 있을 경우, ‘비문법성’은 ‘동사의 시제 불일치 및 굴절 오류’, ‘관사 생략 및 오류’, ‘문장의 단절’, ‘어순 오류’, ‘대명사 사용 시의 응집성 오류’ 등 문법적으로 맞지 표현을

의미한다. 끝으로 ‘불필요한 어휘’는 한글 원본에 없는 표현이 추가되거나 동일한 어휘를 연이어 반복하는 등과 같은 오류를 의미한다. 즉 오류가 없을 경우, ‘Translation is Acceptable,’ 누락된 어휘가 있는 경우는 ‘MW’로, 오역된 표현은 ‘MT,’ 비문법적인 표현은 ‘UG,’ 그리고 한글 원문에 없는 표현이 추가된 경우는 ‘EW’라는 5개 코드를 활용하여 태깅(tagging)하였다. 본 연구에서는 이 기준을 활용하여 A대학에서 다년간 교양영어를 담당한 영어 원어민으로 하여금 기계번역 초벌의 각 문장에 대하여 오류 여부와 오류의 종류를 태깅하게 하였다.

위에서 설명한 바와 같이 실험에 참여한 57명이 작성한 총 931개의 문장에 대하여 오류 코딩이 완료된 후 본 연구의 두 연구자는 오류의 개수와 번역의 정확도를 고려하여 57개 한글 에세이의 기계번역 초벌과 사후수정본에 대하여 쓰기의 적절성 평가를 실시하였다. 적절성 평가에는 5점 척도(1점: 매우 부적절, 2점: 다소 부적절, 3점: 보통, 4점: 대체로 적절, 5점: 매우 적절)를 적용하였으며 기계번역 초벌과 사후수정본에 대한 두 연구자의 채점 신뢰도는 Pearson’s Correlation을 기준으로 각  $p < .01$  수준에서 .945와 .884로 매우 높게 나타났다. 이를 통해 두 연구자가 채점한 결과의 평균 점수를 기준으로 기계번역 초벌과 사후수정본 사이의 적절성 변화 정도를 측정하였다.

### 3) 어휘 및 통사적 특성 분석 도구

기계번역 초벌과 사후수정본 사이의 어휘 및 통사적 특성 변화의 분석에서는 학습자별 쓰기 답안( $N = 57$ ), 즉 쓰기의 적절성 평가와 마찬가지로 에세이 단위로 분석하였다. 먼저 어휘의 특성 분석에는 Lu(2012)가 개발한 온라인 어휘 복잡성 분석 도구인 Lexical Complexity Analyzer(LCA, <https://aihaiyang.com/software/lca/single/>)를 적용하였다. LCA는 쓰기 답안을 구성하는 총 어휘 수, 평균 문장 길이, 어휘의 다양성 등 어휘 사용에 관련된 총 25개 요인의 특성을 측정할 수 있다.

통사적 특성 분석에는 역시 Lu(2010)가 개발한 L2 Syntactic Complexity Analyzer(L2SCA, <https://aihaiyang.com/software/l2sca/single/>)를 사용하였다. L2SCA는 글의 통사적 복잡성을 분석하는 도구로서 총 문장 수, 동사구의 수, 등위구의 수, T-unit의 수 등 14개의 통사 요인을 측정할 수 있다.

본 연구에서 이중 3개의 어휘 요인과 8개의 통사 요인을 분석하였다(표 7 참조).

## IV. 연구 결과 및 논의

### 1. 기계번역의 오류 분석 결과

기계번역기를 활용한 초벌 번역본의 오류 분석 결과를 보면, 표 3과 같이 57

개의 쓰기 답안을 구성하고 있는 총 931개의 문장 중 600개의 문장(64.5%)이 오류가 없는 수용 가능한 번역이었다. 그 외 303개 문장(32.55%)은 1개의 오류를 포함하고 있었고 27개 문장은 2개의 오류, 1개의 문장은 3개의 오류를 포함하고 있었다.

**표 3  
오류 개수별 문장 수 및 비율**

문장당 오류 수	3개	2개	1개	0개	계
해당 문장 수	1(0.11%)	27(2.90%)	303(32.55%)	600(64.45%)	931(100%)

전체 931개 문장을 기준으로 보면 문장당 평균 0.39개의 오류가 발생한 것이고, 오류를 포함하고 있는 331개의 문장만을 기준으로 산정하면 문장당 1.09개의 오류가 발생하였다고 볼 수 있다. 오류의 개수만 보면 실험에 사용된 구글 번역으로 한정되기는 하지만 기계번역의 수준이 상당히 정확하다는 것을 확인할 수 있다. 다음의 표 4는 오류 유형별 빈도수와 그 비율을 정리한 것이다.

**표 4  
오류 유형별 빈도수 및 비율**

유형	어휘 누락	오역	비문법성	불필요한 어휘	계
빈도(%)	61(16.94%)	236(65.56%)	58(16.11%)	5(1.39%)	360(100%)

기계번역에서 가장 빈도수가 높은 오류 유형은 총 360개의 오류 중 236개로 65.56%를 차지한 오역이었다. 오역은 주로 한국어에서 흔한 주어 생략, 한자 표현, 태(voice), 어휘 선택 등에서 발생했다. 이는 Kim(2019)의 연구에서 지적한 구글 번역의 오역 원인과도 상당히 일치한다. 그 다음으로 빈도수가 높은 오류 유형은 16.94%를 차지한 어휘 누락이었다. 비문법성 오류는 16.11%로 세부 유형별로 보면 관사와 시제 오류가 많았다. 특히 한국어의 과거형은 영어에서는 현재완료 시제가 되어야 하는 경우가 많은데 문자그대로의 기계번역이 이루어지는 경향이 있어 오류가 빈번하게 발생한 것으로 보인다. 끝으로 불필요한 어휘가 추가되는 경우가 5회 나타났다. 각 오류 유형에 대한 실제 사례를 살펴보면 다음과 같다.

### #발췌 1

OS: 더욱 문제인 것은 이러한 허위사실이나 비전문적인 지식이 다수에 의해 빠르게 소비된다는 것이다.

MS: Even more problematic is that this false or unprofessional knowledge is quickly consumed by the majority.

CS: Even more problematic is that this false or unverified knowledge is quickly consumed by the majority.

#발췌 1의 예시는 기계번역에서 가장 빈번하게 나타나는 오역의 사례이다. ‘false knowledge’는 자연스런 어휘 조합이나 ‘unprofessional knowledge’는 어색하다. 정확한 표현은 ‘knowledge which is not verified by experts’로 위의 맥락에서는 ‘unverified knowledge’ 정도가 적절해 보인다. 이는 문자 그대로 번역하긴 했지만 문맥에 적합한 어휘 선택에 한계를 보인 경우라고 할 수 있다.

다음의 사례는 한자어 인식의 문제와 인칭대명사의 생략으로 지칭대상이 바뀐 오역의 예시이다.

### #발췌 2

OS: 예를 들어, 교환학생으로 미국에 간 친구와 근황을 주고받고 영상통화를 통해 얼굴을 볼 수 있다.

MS: For example, you can exchange faces with a friend who has been to the United States as an exchange student and see your face through a video call.

CS: For example, you can communicate with a friend who went to the United States as an exchange student and see each other through a video call.

#발췌 2의 예시에서 ‘근황’은 ‘current status’를 의미하는 한자 표현으로 상대적으로 사용 빈도수가 낮은 어휘라고 판단되며 그로 인해 기계번역기가 제대로 인식하지 못한 것으로 보인다. 또한 한글 원문에서 주체 및 지칭 대상이 생략된 관계로 기계번역기가 ‘you’라는 주어를 임의로 부여했고 ‘얼굴’의 지칭 대상 또한 ‘your’로 통일하면서 오역이 발생했다. 주어나 지칭 대상의 생략으로 인한 오류는 엄격히 말해 기계번역기의 오류라기보다는 한글 원문 작성의 모호성에 기인한 것으로 기계번역의 정확성을 높이기 위해서는 한글 원문 또한 영어 문장 형식의 특성을 고려하여 명확히 작성해야 한다는 점을 시사한다.

사용 빈도수가 낮은 한자 표현과 마찬가지로 사용 빈도수가 낮은 순우리말 또한 기계번역기가 제대로 인지하지 못했다. 다음은 #발췌 3은 ‘조리돌림’이라는 표현을 오역한 사례이다.

### #발췌 3

OS: 파격적인 행보로 인터넷 조리돌림을 당하던 가수 00는 루머 기사들에 계속해서 시달려 왔다.

MS: Singer 00, who was faced with Internet cooking because of his unprecedented move, has been plagued by rumor articles.

CS: Singer 00, who had been ridiculed on the Internet for her unconventional actions, has continued to suffer from rumors.

‘조리돌림’은 현재 잘 사용되지 않는 표현으로 ‘집단으로 괴롭히다’는 의미로 사용된다. 당연히 사용 빈도수가 낮은 이 표현을 기계번역기는 ‘조림’이라는 일

부 표현을 들어 ‘cooking’으로 오역하였다.

그밖에도 한글 원문의 표현이 번역에서 누락되는 사례도 다음의 예시와 같이 발생하였다.

#### #발췌 4

OS: 인터넷을 통해 쇼핑, 뱅킹, 교통편, 숙박시설, 공연, 전시회 등의 예약과 관리, 신문구독이나 우편 등 다양한 서비스를 가정에서 해결할 수 있다.

MS: Ø Various services such as shopping, banking, transportation, accommodations, performances, exhibitions, Ø etc can be solved at home.

CS: Through the Internet, various services such as shopping, banking, transportation, accommodation, reservations and management of performances and exhibitions, newspaper subscriptions and mail can be solved at home.

#발췌 4의 예시에서는 한글 원문의 ‘인터넷을 통해’와 ‘예약과 관리’가 번역에 반영되지 않고 누락되었다.

끝으로 다음은 앞서 언급한 바와 같이 비문법성의 대표적인 사례인 시제 오류와 불필요한 표현이 추가된 예시이다.

#### #발췌 5

OS: 인터넷을 통해 활발한 의사소통이 가능해졌다.

MS: Active communication was possible through the Internet.

CS: Active communication has become possible through the Internet.

#### #발췌 6

OS: 인터넷이란 정보를 교환할 수 있도록 전 세계의 컴퓨터가 연결된 통신망을 말한다.

MS: The Internet is a network of computers connected to computers around the world to exchange information.

CS: The Internet is a communication network where computers all over the world are connected so that information can be exchanged.

#발췌 5에서 과거시제 ‘was’는 현재완료 ‘has been’으로 수정되는 것이 더 자연스럽고 #발췌 6의 ‘of computers’ 뒤쪽에 ‘computers’가 있는 관계로 불필요한 표현이라고 할 수 있다.

## 2. 기계번역 초벌-사후수정본 사이의 번역 적절성 차이 분석 결과

학습자의 사후수정이 실제 쓰기 답안의 적절성에 어떠한 영향을 주었는지 살펴보기 위해 본 연구에서는 5점 척도에 근거하여 채점한 기계번역 초벌의 적절성과 사후수정본의 적절성 결과를 *t*검정을 통해 분석하였다.

**표 5  
기계번역 초벌-사후수정본의 대응표본 *t*검정 결과**

(N = 57)	M	SD	t	df	Sig.
기계번역 초벌 적절성	4.468	0.248			
사후수정본 적절성	4.553	0.276	-2.590	56	.012*

Note. \* $p < .05$

표 5에서 보듯 쓰기 적절성의 점수에서는 사후수정본의 점수가 기계번역 초벌의 점수에 비해 높았으며 표준편차 또한 커진 것을 확인할 수 있다. 즉 사후수정의 효과가 유의한 수준( $t = 12.590, p < .05$ )에서 나타났으며 사후수정 단계에서 개인의 능력차이가 반영된 것으로 판단된다. 하지만 영어 전공 집단( $n = 30$ )과 영어 비전공 집단( $n = 27$ )으로 나누어 분석해 보면 그 결과는 상이했다. 다음의 표 6은 두 집단이 작성한 문장들을 구분하여 분석한 결과이다.

**표 6  
학습자의 전공별 기계번역 초벌-사후수정본의 대응표본 *t*검정 결과**

	M	SD	t	df	Sig.
<u>영어 전공 학습자(<math>n = 30</math>)</u>					
기계번역 초벌 적절성	4.430	0.246			
사후수정본 적절성	4.579	0.274	-2.904	29	.007**
<u>영어 비전공 학습자(<math>n = 27</math>)</u>					
기계번역 초벌 적절성	4.509	0.248			
사후수정본 적절성	4.524	0.281	-.408	26	.686

Note. \*\* $p < .01$ , M: 적절성 평가 점수의 평균

두 집단으로 나누어 분석을 실시한 결과, 두 집단 모두 적절성 평균 점수는 사후수정 후 상승하였다. 하지만 영어전공 집단은 통계적으로 유의한 수준( $p < .01$ )에서 사후수정본의 적절성 점수가 상승한 반면 영어 비전공 집단에서는 유의한 변화를 확인할 수 없었다( $p = .686$ ). 이러한 이유로 기계번역기의 산출물보다 더 뛰어난 사후수정본을 작성하기 위해서는 미묘한 표현상의 차이나 오류 까지 인식할 수 있는 상위수준의 영어 능숙도가 요구된다고 판단된다. Chung(2020)의 연구에서도 기계번역기를 활용한 사후수정은 상위 영어 능숙도를 가진 학습자일수록 더 효과적이었다고 주장하였다. 이는 Tian과 Zhou(2020)의 주장, 즉 상위수준의 학습자일수록 자신의 L2 능력에 대한 믿음을 바탕으로

사후수정을 통해 자신만의 문체로 글을 작성하려는 성향이 강하다는 주장과도 맥락을 같이한다.

### 3. 기계번역 초벌-사후수정본 사이의 어휘 및 통사적 복잡성 차이 분석 결과

본 연구에서는 사후수정의 효과성 측정을 위해 기계번역 초벌과 사후수정본 사이의 어휘 및 통사적 변화를 분석하였다. 하지만 어휘 수, 문장 길이 및 어휘 다양성 등과 같은 어휘 요인뿐만 아니라 통사 요인인 동사구, 절(clause) 수, 종속절 수 등에서도 유의한 차이를 확인할 수 없었다.

**표 7  
어휘/통사 요인의 변화 측정을 위한 대응표본 t검정 분석 결과**

(N = 57)	기계번역 초벌		사후수정본		t	Sig.
	M	SD	M	SD		
어휘 요인 분석						
어휘 수	276.316	131.642	275.298	129.555	.529	.599
문장 길이	16.796	3.295	16.384	3.679	1.198	.236
어휘의 다양성(MTLD)	63.548	19.184	63.209	20.453	.336	.738
통사 요인 분석						
문장 수	16.474	7.127	16.684	7.224	-2.120	<b>.038*</b>
동사구 수	36.702	17.221	36.684	16.771	.060	.952
절 수	26.895	12.327	27.000	12.265	-.504	.616
T-unit	18.368	7.932	18.439	8.049	-.468	.641
종속절 수	8.491	5.275	8.439	5.230	.268	.789
등위구 수	7.474	4.676	7.123	4.326	2.231	<b>.030*</b>
T-unit 평균 길이	14.820	2.227	14.741	2.033	.490	.626
절 평균 길이	10.244	1.367	10.173	1.268	.662	.511

Note. \* $p < .05$

사후수정 후 차이를 보인 요인은 ‘문장 수’와 ‘등위구의 수’가 유일하다. 그 이유로는 이미 언급한 바와 같이 수용 가능한 번역이 전체 문장의 64.45%, 1개의 오류만 포함된 경우가 32.55%로 이미 번역기의 초벌 번역만으로도 상당한 정확성을 보였기 때문에 학습자들은 대대적인 수정보다는 의미의 변화나 통사적 복잡성에 큰 변화를 가져오지 않는 단어 교체(approximation)나 구 교체(phrase paraphrase) 수준에서 수정이 이루어졌고 긴 문장을 간결하게 쪼개어 작성하는 경향을 보여 이것이 표 7의 분석 결과에 영향을 준 것으로 파악된다. 이러한 현상은 쓰기 능숙도가 상대적으로 낮은 집단에서 보다 두드러지게 나타난다.

**표 8  
학습자의 전공별 어휘/통사 요인의 변화 측정을 위한 대응표본 *t*검정 분석 결과**

	a. 어휘 요인 b. 통사 요인	기계번역 초별		사후수정본		<i>t</i>	Sig.
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
영어 전공 학습자 ( <i>n</i> = 30)	a. 총 문장 수	15.033	7.522	15.100	7.503	.571	.573
	b. T-unit	17.133	8.328	16.900	8.206	1.045	.305
영어 비전공 학습자 ( <i>n</i> = 27)	a. 총 문장 수	18.074	6.421	18.444	6.594	-2.294	.030*
	b. T-unit	19.741	7.378	20.148	7.660	-2.275	.031*

Note. \**p* < .05

영어 전공과 비전공으로 구분하여 분석한 위의 표 8에서 보듯 영어 쓰기 능숙도가 상대적으로 낮은 영어 비전공 집단에서 기계번역 초별의 문장 수(*M* = 18.074, *SD* = 6.421)와 비교하여 사후수정본의 문장 수(*M* = 18.444, *SD* = 6.594)가 통계적으로 유의한 수준(*p* < .05)에서 증가하는 것을 확인할 수 있었다. 영어 비전공 집단에서는 비슷한 맥락에서 단문이나 등위접속사로 연결된 하나의 독립절을 의미하는 T-unit(Hunt, 1965) 또한 기계번역 초별(*M* = 19.741, *SD* = 7.378) 보다 사후수정본(*M* = 20.148, *SD* = 7.660, *p* < .05)에서 높게 나타나는 것을 확인할 수 있다. 즉 영어 쓰기 능숙도가 상대적으로 낮은 학습자들일수록 사후수정의 과정에서 문장을 여러 개로 짧게 나눠 수정하는 경향을 보인 것으로 이는 기존의 여러 연구 결과(예, Grant & Ginther, 2000; Jung et al., 2019)와도 일치한다.

## V. 결론 및 제언

본 연구는 기계번역의 오류 분석과 더불어 사후수정의 효과성을 사후수정본의 적절성 개선 정도와 어휘 및 통사적 측면의 변화를 통해 살펴보았다. 그 결과를 요약하면 다음과 같다.

첫째, 총 931개의 한글 원문 문장에 대한 기계번역에서 중 64.5%인 600개의 문장은 오류가 없었고 32.55%인 303개 문장은 1개의 오류를 포함 문장이었고 27개 문장은 2개, 그리고 3개의 오류를 포함한 문장은 1개였다. 전체적으로 보면 문장당 평균 0.39개의 오류가 발생한 것이고 오류를 포함한 문장만 보면 문장당 1.09개의 오류가 발생한 것이다. 가장 빈도수가 높은 오류 유형은 65.56%를 차지한 오역이었고 그 다음은 16.94%를 차지한 어휘 누락 오류였다. 비문법성 오류는 16.11%이고 불필요한 어휘 추가 오류는 1.39%였다. 오류의 주요 원인은 기계번역기의 기술적 한계에서도 기인했지만 한국어-영어의 언어적 특성 차를 고려하지 않은 한글 원문에서 기인한 경우도 많았다.

둘째, 문장 단위로 살펴본 사후수정의 효과성 검증에서는 기계번역 초별에

비해 사후수정본의 적절성 평균 점수가 통계적으로 유의한 수준( $p < .05$ )에서 상승하였다. 하지만 영어 쓰기 능숙도에서 차이를 보인 영어 전공 집단과 영어 비전공 집단의 비교·분석에서는 영어 전공자 집단의 사후수정본에서만 적절성 점수가 통계적으로 유의한 수준( $p < .01$ )에서 개선되었다. 이는 사후수정의 전제 조건이 오류를 인식하는 데 있는 만큼 이미 상당한 정확도를 보이고 있는 기계번역 초벌에서 미묘한 오류를 인지하는 것이 상위권 학습자들을 제외한 중하위권 학습자들에게는 쉽지 않았기 때문이라고 추정된다.

셋째, 기계번역 초벌과 사후수정본 사이의 어휘 및 통사적 복잡성의 변화는 거의 보이지 않았다. 반면, 영어 비전공자 집단의 사후수정본에서는 문장 수와 T-unit의 수가 통계적으로 유의한 수준( $p < .05$ )에서 증가하였는데 이는 영어 쓰기 능숙도가 상대적으로 낮은 학습자들일수록 사후수정의 과정에서 문장을 여러 개로 짧게 나눠 수정하는 경향이 있다는 것을 보여준다.

이러한 연구 결과는 기계번역기를 활용한 L2 쓰기 활동이 학습자가 L2 쓰기 과정 중에 직면할 수 있는 어휘 구사 및 통사구문과 관련한 문제를 극복하는데 도움을 줄 수 있다는 점을 시사한다. 이와 같은, 기계번역 기술의 발전은 영어 쓰기 지도에도 다양하게 활용될 수 있을 것으로 기대된다. 기계번역기의 몇 가지 교육적 활용 방안을 제안하면 다음과 같다.

기존의 사전(dictionary)은 검색한 표현의 의미와 관련 예문을 제공해 주지만 그 예문이 검색한 표현의 다양한 쓰임과 의미를 모두 제공하는 데는 한계가 있다. 반면 기계번역기는 사전보다 확장된 맥락적 정보를 제공해 줌으로써 사전 보다 효과적인 참고 자료로 활용할 수 있다. 이윤재와 이동주(2020)의 연구에서도 기계번역의 활용 시 어휘 선택 오류의 수가 56%나 줄어들었다고 보고한 것을 보면 기계번역기가 연어(collocation)와 같은 어휘 조합의 측면에서 보다 정확한 맥락적 정보를 제공할 수 있다는 것을 알 수 있다(예, 높은 빌딩, 높은 산 > tall building, high mountain). 또 다른 기계번역기의 활용 방안으로는 기계번역에 포함된 오류 표현을 오류 인지 활동 자료로 활용하는 것이다. 학습자들이 오류를 인식하는 것은 매우 중요한 능력 중 하나이나 이를 지도하기 위해 오류가 포함된 자료를 확보하거나 개발하는 것은 많은 시간과 노력이 필요하다. 하지만 영어교육 전문가들이 위에서 살펴본 기계번역기의 오류 특성을 파악하다면 어느 정도 의도된 오류 문장을 만드는 것은 어렵지 않기 때문에 기계번역기를 사용하면 두 언어 간 차이가 반영된 오류 문장 확보에 유용하게 활용할 수 있다. 이에 더하여, 기계번역기 활용 전략을 체계적으로 지도한다면 학습자들은 언어 간의 차이를 보다 실제적으로 인지할 수 있을 것이다. 실제 기계번역기를 보다 효과적으로 사용하기 위해서는 한글 원문 작성 시부터 영어식 문장 형식을 고려해야 한다. 여기서 영어식으로 표기하는 것이 한글 표현을 파괴하면서 까지 지도해야 한다는 의미하는 것은 아니며 두 언어 간 차이의 지도에 초점을 두어야 한다는 것을 말한다. 예를 들어, 한글 작성 시 주어를 생략하는 경우가 많지만 영어에서는 주어가 없으면 비문이 되는 경우가 많기 때문에 앞서 언급

한 바와 같이 번역기 활용을 위한 한글 원문 작성에서는 오류의 예방을 위해 문장에서 주체가 되는 주어를 반드시 작성할 필요가 있다. 하지만 무엇보다도 기계번역기의 효과적 사용을 위해서는 학습자는 기계번역기가 산출한 초별 번역에 포함된 오류를 인지할 수 있는 어휘 및 문법능력을 갖추고 있어야 하며 언어 간 표현 방식의 차이를 충분히 이해하여 이를 사후수정에 반영할 수 있어야 한다는 점이 전제된다.

본 연구는 여러 흥미로운 결과를 제시했음에도 불구하고 추후 연구 설계에서는 개선해야 할 몇 가지 제한점들이 있다.

첫째, 연구의 수행 시점과 비교하여 현재의 구글 번역은 다시 업그레이드되었다. #발췌 3에서는 ‘조리돌림’이란 표현이 ‘cooking’으로 기계번역 되었지만 현재는 ‘being ridiculed’로 번역되어 기계번역의 오류가 지속적으로 개선되고 있다는 것을 알 수 있다. 따라서 연구 결과는 연구의 수행 시점을 고려하여 해석할 필요가 있다.

둘째, 연구 대상이 된 두 집단을 대상으로 사전 평가를 통해 영어 쓰기 능숙도를 진단하였지만 진단 평가에 적용한 쓰기 주제와 실제 쓰기 실험 과업에 적용한 쓰기 주제가 상이한 관계로 기계번역기 미사용 시와 사용 시의 쓰기 결과 비교에 대한 정확한 통계분석은 어려웠다. 따라서 본 연구와 유사한 연구를 수행할 때는 실험 대상을 두 집단으로 구분하여 진단 평가의 주제와 실험 과업의 주제 배정 시 역균형화(counterbalancing) 기법의 적용을 제안하고자 한다.

끝으로, 본 연구에서는 기계번역의 오류를 분석하였지만 사후수정의 유형은 분석하지 못했다. 따라서 사후수정의 유형 또한 분석하여 어떤 유형이 최종 쓰기 결과물의 질적 측면에 보다 영향을 미치는지도 살펴볼 필요가 있다.

#### 참고문헌

- 신동광, 민호기, 주현우. (2012). 국가영어능력평가시험 대비 영어 말하기·쓰기 평가매뉴얼: 고등학교용(연구자료 ORM 2012-67). 서울: 한국교육과정평가원.
- 윤채원. (2021). 기계 번역의 오류 유형 분석과 학습자들의 오류 수정 전략. 석사학위논문, 한양대학교.
- 이윤재, 이동주. (2020). 영어자동번역기 활용이 고등학생 영어 글쓰기에 미치는 영향. *영어교과교육*, 19(2), 159-180.  
<http://doi.org/10.18649/jkees.2020.19.2.159>
- 이준호. (2018). 포스트에디팅 교육을 위한 포스트에디팅과 인간번역의 차이 연구. *통역과 번역*, 20(1), 73-96. <http://doi.org/10.20305/it201801073096>
- 임희주. (2017). 교양영어 수업에서 영어자동번역기 사용에 대한 대학생의 인식 및 태도연구: 영작문 수업을 중심으로. *교양교육연구*, 11(6), 727-751.

<https://j-kagedu.or.kr/upload/pdf/kagedu-11-6-727.pdf>

- 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.

[http://journal.kamall.or.kr/wp-content/uploads/2020/07/23\\_2\\_01.pdf](http://journal.kamall.or.kr/wp-content/uploads/2020/07/23_2_01.pdf)

- Alhaisoni, E., & Alhaysony, M. (2017). An investigation of Saudi EFL university students' attitudes towards the use of Google Translate. *International Journal of English Language Education*, 5(1), 72-82.

<https://doi.org/10.5296/ijele.v5i1.10696>

- Bowker, L., & Ciro, J. B. (2019). *Machine translation and global research: Towards improved machine translation literacy in the scholarly community*. Emerald Publishing.

<https://www.emerald.com/insight/publication/doi/10.1108/9781787567214>

- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C. … Amodei, D. (2020). Language models are few-shot learners. *ArXiv*. <https://arxiv.org/pdf/2005.14165v4.pdf>

- Chon, Y. V., & Shin, D. (2020). Direct writing, translated writing, and machine translated writing: A text level analysis with Coh-Metrix. *English Teaching*, 75(1), 25-48. <https://doi.org/10.15858/engtea.75.1.202003.25>

- Chung, E. S. (2020). The effect of L2 proficiency on post-editing machine translated texts. *Journal of Asia TEFL*, 17(1), 182-193.

<http://dx.doi.org/10.18823/asiatefl.2020.17.1.11.182>

- Daems, J., Vandepitte, S., Hartsuiker, R., & Macken, L. (2015). The impact of machine translation error types on post-editing effort indicators. In S. O'Brien & M. Simard (Eds.), *4th Workshop on Post-Editing Technology and Practice* (pp. 31-45). Miami, USA: Association for Machine Translation in the Americas.

[https://www.lt3.ugent.be/media/uploads/publications/2015/Daems\\_WPTP4Proceedings.pdf](https://www.lt3.ugent.be/media/uploads/publications/2015/Daems_WPTP4Proceedings.pdf)

- Daniels, P., & Leslie, D. (2013). Grammar software ready for EFL writers? *OnCUE Journal*, 9(4), 391-401.

[https://jaltcue.org/files/OnCUE/OCJ9.4/OCJ9.4\\_pp391\\_401\\_AC\\_Daniels\\_Leslie.pdf](https://jaltcue.org/files/OnCUE/OCJ9.4/OCJ9.4_pp391_401_AC_Daniels_Leslie.pdf)

- Ferris, D. (2011). *Treatment of error in second language student writing*. University of Michigan Press.

- Gaskell, D., & Cobb, T. (2004). Can learners use concordance feedback for

- writing errors? *System*, 32(3), 301-319.  
<https://doi.org/10.1016/j.system.2004.04.001>
- Grant, L., & Ginther, A. (2000). Using computer-tagged linguistic features to describe L2 writing differences. *Journal of Second Language Writing*, 9(2), 123-145. <https://doi.org/10.1016/S1060-3743%2800%2900019-9>
- Groves, M., & Mundt, K. (2015). Friend or foe? Google Translate in language for academic purposes. *English for Specific Purposes*, 37, 112-121.  
<https://doi.org/10.1016/j.esp.2014.09.001>
- Hunt, K. (1965). *Grammatical structures written at three grade levels* (NCTE Research Rep. No. 3). Champaign, IL, USA: NCTE.  
<https://files.eric.ed.gov/fulltext/ED113735.pdf>
- Hwang, G-J., Xie, H., Wah, B. A., & Gašević, D. (2020). Vision, challenges, roles and research issues of Artificial Intelligence in Education. *Computers & Education: Artificial Intelligence*, 1, 100001.  
<https://doi.org/10.1016/j.caeari.2020.100001>
- Jia, Y., Carl, M., & Wang, X. (2019). How does the post-editing of neural machine translation compare with from-scratch translation?: A product and process study. *The Journal of Specialised Translation*, 31, 60-86.  
[https://jostrans.org/issue31/art\\_jia.pdf](https://jostrans.org/issue31/art_jia.pdf)
- Jung, Y.-J., Crossley, S., & McNamara, D. (2019). Predicting second language writing proficiency in learner texts using computational tools. *The Journal of Asia TEFL*, 16(1), 37-52. <https://doi.org/10.18823/asiatefl.2019.16.1.3.37>
- Kim, N.-Y. (2018). Effects of the use of bilingual dictionary and machine translation on EFL students' English proficiency. *Journal of the Korea English Education Society*, 17(1), 21-42.  
<http://dx.doi.org/10.18649/jkees.2018.17.1.21>
- Kim, S. (2019). Playing with machine translation in language classroom: Affordances and constraints. *Multimedia-Assisted Language Learning*, 22(2), 9-28.  
[http://journal.kamall.or.kr/wp-content/uploads/2019/07/sungwoo-kim\\_22\\_2\\_01.pdf](http://journal.kamall.or.kr/wp-content/uploads/2019/07/sungwoo-kim_22_2_01.pdf)
- Kol, S., Schcolnik, M., & Spector-Cohen, E. (2018). Google Translate in academic writing courses? *The EuroCALL Review*, 26(2), 50-57.  
<https://doi.org/10.4995/eurocall.2018.10140>
- Lee, S. M. (2020). The impact of using machine translation on EFL students' writing. *Computer Assisted Language Learning*, 33(3), 157-175.  
<https://doi.org/10.1080/09588221.2018.1553186>
- Lee, S. M., & Briggs, N. (2021). Effects of using machine translation to mediate the revision process of Korean university students' academic writing.

- ReCALL*, 33(1), 18-33. <https://doi.org/10.1017/S0958344020000191>
- Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, 15(4), 474-496.  
<https://www.jbe-platform.com/docserver/fulltext/ijcl.15.4.02lu.pdf?>
- Lu, X. (2012). The relationship of lexical richness to the quality of ESL learners' oral narratives. *The Modern Language Journal*, 96(2), 190-208.  
[https://doi.org/10.1111/j.1540-4781.2011.01232\\_1.x](https://doi.org/10.1111/j.1540-4781.2011.01232_1.x)
- Nair, S., & Bashir, D. (2020, August 5). GPT-3: An AI breakthrough, but not coming for your job. *Skynet Today*. <https://www.skynettoday.com/briefs/gpt3>
- O'Sullivan, I., & Chambers, A. (2006). Learners' writing skills in French: Corpus consultation and learner evaluation. *Journal of Second Language Writing*, 15, 46-68. <https://doi.org/10.1016/j.jslw.2006.01.002>
- PACTE Group. (2005). Investigating translation competence. Conceptual and methodological issues. *Meta: Translators' Journal*, 50(2), 609-619.  
<https://doi.org/10.7202/011004ar>
- Pokrivcakov, S. (2019). Preparing teachers for the application of AI-powered technologies in foreign language education. *Journal of Language and Cultural Education*, 7(3), 136-153. <https://doi.org/10.2478/jolace-2019-0025>
- Shin, D., Kwon, S. K., & Lee, Y. (2021). The effect of using online language-support resources on L2 writing performance. *Language Testing in Asia*, 11(4). <https://doi.org/10.1186/s40468-021-00119-4>
- Tian, L., & Zhou, Y. (2020). Learner engagement with automated feedback, peer feedback and teacher feedback in an online EFL writing context. *System*, 91, 102247. <https://doi.org/10.1016/j.system.2020.102247>
- Tsai, S.-C. (2019). Using Google Translate in EFL drafts: A preliminary investigation. *Computer Assisted Language Learning*, 32(5-6), 510-526.  
<https://doi.org/10.1080/09588221.2018.1527361>
- Vilar, D., Xu, J., D'Haro, L. F., & Ney, H. (2006). Error analysis of statistical machine translation output. In N. Calzolari, K. Choukri, A. Gangemi, B. Maegaard, J. Mariani, J. Odijk & D. Tapias (Eds.), *Proceedings of the 5th International Conference on Language Resources and Evaluation* (pp. 697-702). Genoa, Italy: European Language Resources Association (ELRA).  
[http://www.lrec-conf.org/proceedings/lrec2006/pdf/413\\_pdf.pdf](http://www.lrec-conf.org/proceedings/lrec2006/pdf/413_pdf.pdf)

**적용가능 수준 (Applicable Levels): Primary/Secondary/Tertiary**

신동광

광주교육대학교 영어교육과  
61204 광주광역시 북구 팔문대로 55  
Tel: (062) 520-4211  
Email: sdhera@gmail.com

전유아

한양대학교 영어교육과  
04763 서울특별시 성동구 왕십리로 222  
Tel: (02) 2220-1144  
Email: vylee52@hanyang.ac.kr

Received December 31, 2021

Revised February 5, 2022

Accepted February 20, 2022