

## 2010년대 국내 통번역학 연구 동향에 대한 키워드 분석\*

이 창 수

(한국외국어대)

### 1. 들어가는 말

국내에서 통번역에 대한 학문적 연구가 뿌리를 내리기 시작한 것은 1990년대 말에서 2000년대 초 사이에 통번역연구를 발표할 수 있는 전문학술지들이 출범한 이후라고 할 수 있다. 구체적으로 한국외국어대학교 통번역연구소가 1997년에 『통번역연구』를 창간하였고 뒤이어 1998년에 한국통역번역학회의 『통역과 번역』(초기 이름은『국제회의통역과 번역』), 2000년에 한국번역학회의 『번역학연구』, 2003년에는 한국통번역교육학회의 『통번역교육연구』가 창간되었다. 그 후 20년 가까운 세월이 흘렀다. 이 시기를 반으로 나뉘볼 때 전반 10년은 통번역 연구가 독자적인 연구 영역으로 자리매김을 한 시기라고 본다면, 2000년 이후는 통번역 연구가 본격적으로 꽃을 핀 시기라고 할 수 있다. 본 연구는 이 중 후반기에 해당하는 시기인 2010년에서 현재 시점까지 국내 통번역 연구

---

\* 본 연구는 2019년도 한국외국어대학교 교내연구비 지원을 받아 작성되었음.

가 연구 주제나 방법론 면에서 어떤 흐름을 형성해왔는가를 분석해보는 것을 목적으로 한다.

분석 방법으로는 위에 언급된 4종의 학술지에서 수집한 총 1299개의 영문 초록을 코퍼스로 구성해서 시대별 특징을 나타내는 용어, 즉 키워드를 분석하였다. 이와 같은 코퍼스 분석방법은 사전에 분석항목 목록을 만들고 그에 따라 수작업으로 코딩을 해 나가는 콘텐츠 분석기법과 대비된다. 후자는 분석항목을 세분화할 경우 매우 상세한 내용분석이 가능하다는 장점이 있지만 수작업 코딩을 하다보니 많은 인력과 시간이 소요되고 샘플 수가 제한되어 전체 시스템을 분석하지 못하는 단점이 있다(포르투나 외(Fortuna et al.) 2009: 27). 이에 비하여 코퍼스를 활용한 컴퓨터 분석은 분량에 구애받지 않고 전체 시스템을 분석대상으로 삼을 수 있다. 또한 사전에 정한 분석 항목에 제한 받지 않고 데이터 분석에서 드러나는 특징을 해석한다. 이런 특징 때문에 코퍼스를 활용한 텍스트 분석은 기본적으로 ‘데이터 중심적(data-driven)’이며 ‘탐구적(exploratory)’이고 ‘귀납적인(inductive)’ 특징을 갖고 있다. 그 결과 기존 콘텐츠 분석 방식에서는 사전에 예측하지 못한 특징을 찾아낼 수 있는 장점이 있다(왈데르 외(Waldherr et al.) 2016: 231).

본 연구에서는 2000년부터 2019년까지 10년을 초기-중기-후기와 전반부-후반부로 나누고 각 시기별로 키워드를 추출하여 해당 시기의 특징적 연구 동향을 분석하였다. 본 연구는 그 동안 통번역 분야의 연구성과물을 코퍼스로 분석한 연구가 별로 없기 때문에 새로운 연구방법론을 소개한다는 점과 더불어 2010년대를 마감하는 시점에서 국내 통번역연구의 큰 흐름을 정리해본다는 점에서 의의를 찾을 수 있다.

## 2. 국내 통번역학 동향 연구

그동안 국내 통번역학 연구 동향을 분석하려는 시도는 간헐적으로 이어져 왔다. 그중 가장 중요한 기여를 한 연구로는 정호정(2013)과 김혜림(2013)을 들 수 있다. 정호정(2013)의 경우는 국내학술지 뿐만 아니라 국제학술지를 분석 대상에 포함하였는데, 국내연구에서는 『번역학연구』와 『통번역교육연구』 두 종의

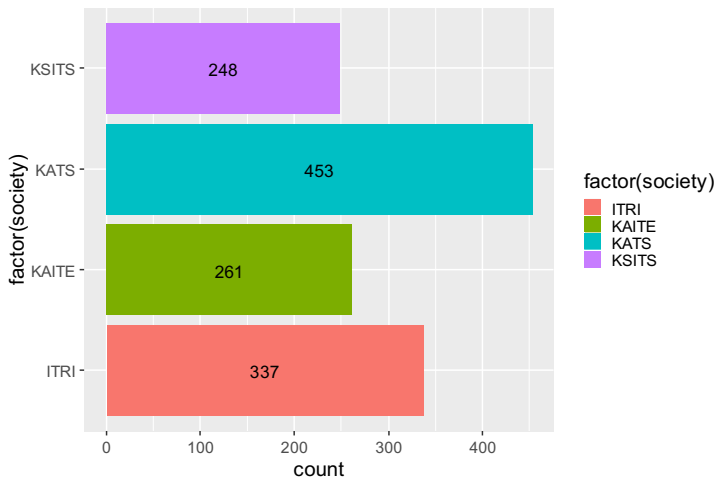
학술지의 번역관련 논문을 분석하였다. 분석 결과 7개의 대주제 중에 번역이론에 관한 연구가 전체 연구의 1/4에 달할 정도로 가장 큰 비중을 차지하였다. 연구대상 기간이 1999년에서 2012년도까지로 번역학의 초기 도입기에 해당하기 때문에 해외 번역이론을 소개하고 활용한 연구가 많았던 것으로 해석된다. 동 연구에서 관심을 끄는 또 다른 연구결과는 2000년 후반기에 들어서면서 『통번역교육연구』에서는 비문학번역(또는 전문번역)에 비하여 문학번역의 비중이 높게 나타난 반면 『번역학연구』에서는 반대 현상이 발견되었다는 점이다. 이는 학술지별로 발표되는 연구의 관심 주제나 분야가 다를 수 있다는 점을 시사한 것으로 해석할 수 있다. 김혜립(2013) 연구에서는 번역 뿐만 아니라 통역까지 분석 대상에 포함하였으며 연구성과물을 11개 주제별로 분류하였다. 그 결과 텍스트 분석에 기초한 연구가 가장 큰 비중을 차지하였고 다음으로 언어학과 교육학적 관점에서 통번역을 분석한 연구가 비슷한 비중을 차지하였다. 그 외에도 한현희(2016)는 통역연구에 초점을 맞춰 국내와 해외의 통번역학 연구 동향을 비교 분석하였다. 그 결과 국내 통역연구에서는 직업론 측면에서의 연구가 대다수를 차지한 반면, 해외 연구는 통역 윤리, 통역사 정체성, 통역 과정과 평가 등 연구 주제와 범위가 훨씬 더 다양한 것으로 나타났다. 이들 연구 외에는 특정 연구분야의 연구 동향을 분석한 연구들이 있는데, 영상번역 연구에서 조성은(2014), 코퍼스 활용 연구에서 최희경(2015), 기계번역 연구에서 이건일과 이종혁(2015), 교육 분야 연구에서 한현희(2017) 등을 들 수 있다.

이상의 연구동향 연구는 기본적으로 콘텐츠 분석에 기초한 연구이다. 이와는 대조적으로 이창수(2018)는 본 연구와 유사한 코퍼스 분석 접근을 시도하였다. 동 연구는 2000~2003년과 2015~2017년 양 기간에 『번역학연구』에 실린 논문을 분석데이터로 사용하여 어휘 빈도와 네트워크 분석을 시행하였다. 동 연구의 분석 결과에 따르면 양 기간 사이에 연구 흐름은 (1) 언어구조적 연구에서 담화-텍스트 분석으로 (2) 등가적 연구에서 기능적 연구로 (3) 실례 위주 분석에서 코퍼스를 활용한 실증적 연구로 (4) 일반번역에서 문학번역과 전문번역으로 바뀐 것으로 나타났다. 이 연구와 비교하여 본 연구에서는 분석 기간을 국내 통번역학 역사의 후반기라고 할 수 있는 2010년 이후로 선정했다는 점, 분석방식으로 비교 분석에 좀 더 효율적인 키워드 분석을 사용하였다는 점, 연구재단에 등재된 통번역관련 전문 학술지 4종 전체를 연구 대상으로 삼았다는 점에서 차이가 있다.

### 3. 분석 데이터 및 방법

분석 데이터는 2010년부터 2019년 최근호까지 10년 동안 『통번역연구』, 『통역과 번역』, 『번역학연구』, 『통번역교육연구』 등 4개 학술지에 실린 논문 전체의 초록을 대상으로 하였다. 다만, 영문 초록을 분석대상으로 삼았기 때문에 한글 초록이 실린 몇 편의 논문은 제외되었다. 논문의 편의를 위하여 상기 4종의 학술지를 발행기관과 학술지 명을 반영하여 각각 임의로 ITRI, KSITS, KATS, KAITE로 명하기로 한다. 분석에 포함된 총 초록 수는 1,299개로 각 학술지별 분포는 그림 1과 같다.

그림 1 학술지별 초록 수



초록은 학회, 학술지, 발간연도 등의 메타 데이터 및 논문 키워드와 함께 엑셀에 정리하였으며, 해당 엑셀 데이터를 R 통계 프로그램에 탑재하여 컴퓨터로 분석하였다. 텍스트 분석 과정에서는 R의 `quanteda` 패키지가 제공하는 텍스트 분석 기능 중 키워드 분석 기능을 중점적으로 활용하였다. 키워드 분석은 참조코퍼스(reference corpus)에 비하여 분석코퍼스(target corpus) 내에서 특징적으로 나타나는 어휘를 추출하는 것으로 이런 어휘적 특징을 키값(keyness)라고 부른다.<sup>1)</sup>

분석코퍼스 어휘의 키값을 계산하는 것은 간단하다. 가령, 분석코퍼스를 T, 비교코퍼스를 R이라고 하고 ‘translation 이란 어휘의 키값을 구한다고 가정하자. 이 경우 각 코퍼스에서 발생하는 translation의 빈도수, XT와 XR 그리고 각 코퍼스의 총 어휘수에서 XT와 XR을 뺀 나머지 어휘 수 -XT와 -XR을 사용하여 카이제곱검정을 실시하면 통계유의치인 p값과 더불어 카이제곱값이 구해지는데 바로 이 카이제곱값이 키값이 된다. 즉, 해당 어휘가 R에 비하여 T에서 상대적으로 많이 발생할수록 키값이 높아지게 된다. 반대로 상대적으로 적게 발생하면 키값은 마이너스 숫자가 된다.

데이터분석은 2단계로 진행하였다. 먼저 2010~2019년의 분석대상 연도를 2010~2012, 2013~2015, 2016~2019 등 ‘초기’, ‘중기’, ‘후기 3개의 시기로 나눠 분석하였다. ‘초기’와 ‘중기’는 3개 연도가 포함된 반면, ‘후기’만 4개 연도로 구성되었는데 이는 2019년도가 아직 끝나지 않은 상황을 고려하였다. 두 번째 단계에서는 분석대상 연도를 상반기(2010~2014)와 하반기(2015~2019)로 나눠 비교 분석하였다. 분석에서는 분석 대상 시기에 해당하는 초록을 분석코퍼스, 나머지 시기에 해당하는 것은 참조코퍼스로 나눠 각 시대별 키워드를 산출하였다.

## 4. 분석 결과

### 4.1. 어휘빈도 분석

키워드 분석에 앞서서 전체 초록의 어휘빈도를 분석해본다. <그림 2>를 보면 빈도 상위 20개 어휘가 나와있다. 이 어휘목록은 조동사, 전치사, 관사 같은 기능어들과 연구 주제와 큰 상관이 없다고 여겨지는 형용사나 동사 등을 제거한 후 만들어졌다. 전체 초록코퍼스에서 빈도가 가장 높은 단어는 translation(번역)으로 4,913번 발생한다. 그 외에 번역과 관련된 어휘들인 translators(번역사

1) 이 기능은 원래 텍스트 분석프로그램인 워드스미스(WordSimth)에 포함된 기능을 R에 구현한 것이다. 따라서 워드스미스 프로그램을 사용해도 같은 결과를 얻을 수 있다. 다만, R에서는 필요에 따라 데이터를 가공하는 것이 쉽기 때문에 연도, 시대, 학술지 별로 나눠 비교하는 것이 훨씬 용이하다. 그와 더불어 R이 제공하는 그래프나 기타 코퍼스 분석패키지를 통합해 활용할 수 있는 장점이 있다.

들), translations(번역들), translated(번역된) 등이 포함되어 있다. interpreting, interpretation(통역)은 각각 8, 17위에 올라 있다. 그 외에 study, studies(연구), paper(논문), analysis(분석) 등 초록에 일반적으로 등장하는 어휘들이 눈에 띈다. 또한 korean(한국어), english(영어)는 논문에서 한-영 통번역을 다룬 논문이 상대적으로 많음을 시사한다. 그 외에 통번역을 논할 때 일반적을 등장하는 text, texts(텍스트), target(도착(어)), source(출발(어)) 등이 있다. students(학생들)는 학생을 대상으로 연구가 많았음을 보여준다. 그렇지만 그림 2의 최빈도 어휘는 대부분 통번역 논문 초록에 등장하는 어휘로 특정한 연구 주제와 관련이 없다.

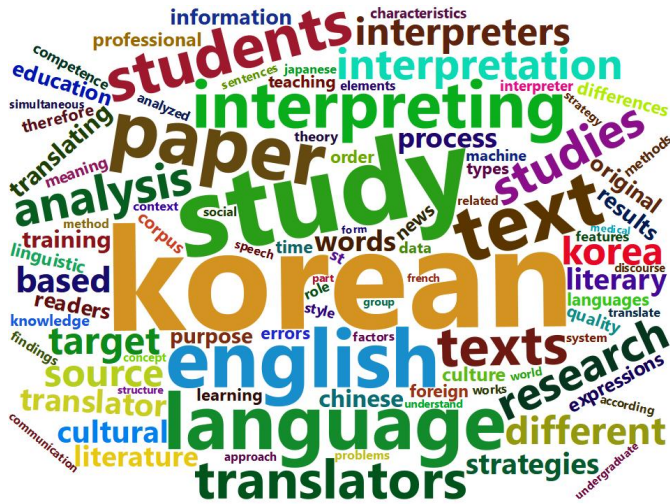
그림 2 초록코퍼스의 최빈도 어휘 20개

	feature	frequency	rank
1	translation	4913	1
2	korean	1890	2
3	study	1493	3
4	english	1066	4
5	text	1045	5
6	paper	1034	6
7	language	1027	7
8	interpreting	863	8
9	students	814	9
10	texts	737	10
11	translators	712	11
12	translations	698	12
13	translated	688	13
14	analysis	649	14
15	studies	648	15
16	research	610	16
17	interpretation	565	17
18	target	513	18
19	korea	510	19
20	source	509	20

이번에는 그림 3의 워드클라우드에 나와 있는 대로 최상위 어휘수를 100개로 늘려 검토해보자. 여기서 글자 크기는 빈도수에 비례하는데, 최상위 어휘인 translation은 글자 크기가 너무 커지기 때문에 제외되었다. 그림 3을 보면 연구 주제와 관련된 것으로 보이는 어휘들이 다수 등장한다. 가령, literature, literary는 ‘문학번역’, cultural, culture는 ‘문화’, training, linguistic은 ‘언어적 연구’, education, teaching, learning은 ‘통번역교육’, quality, error 등은 ‘품질 평가’, machine은 ‘기계번역 과 관련된 어휘들이다. 그 외에 corpus, based는 ‘코퍼스 분석을 방법론으로 택한 연구와 관련 있고, strategy, strategies는 ‘통번역 전략’, style은 ‘문체 등과 연관 있다. 이와 같이 최빈도 어휘를 살펴보면 대충

2010년대의 통번역연구에서 어떤 주제들이 큰 관심을 끌었는지를 알 수 있다.

그림 3 최빈도 어휘 100개를 활용한 워드클라우드



#### 4.2. 초·중·후기 비교 분석

이번에는 앞서 살펴본 2010년대 전체의 연구 주제들이 시기별로 어떤 흐름을 형성하는지를 알아보기 위하여 2010년대를 초·중·후기 등 3개 시기로 나눠 시기별 키워드를 분석해본다. 3절에서 간단하게 소개한 quanteda 패키지를 사용하여 키값을 구하면 그림 4와 같은 결과가 나온다. 이 표에서 카이검정값인  $\chi^2$ 가 킷값이며, 이 수치를 중심으로 순위가 매겨져있다. 그 옆에는 각 어휘별 카이검정의 p값, 분석코퍼스( $n_{target}$ )와 대조코퍼스( $n_{reference}$ )의 발생빈도수가 표시되어 있다. 여기서 분석코퍼스는 2010에서 2013년까지 3년 동안 발표된 초록, 대조코퍼스는 이 3년을 제외한 나머지 7년 동안 발표된 초록을 의미한다. 그런데 그림 4의 목록을 보면 wordplays(언어유희), hangul(한글), numeral(숫자), footnotes(각주)와 같이 전반적 연구의 주제나 방법보다는 특정 논문의 연구문제와 연결된 어휘들이 보인다. 이런 어휘는 키값은 높지만 발생빈도가 작고, 특정 시점에 나온 논문이기 때문에 대조코퍼스에서 거의 나타나지 않는 것

이 특징이다. 이들 어휘는 해당 시기를 대표하는 키워드로 보기 어렵기 때문에 분석코퍼스와 대조코퍼스에서 최소발생빈도수를 정해 그 수치 이하의 어휘는 분석에서 제외할 필요가 있다. 본 연구에서는 목록을 검토한 결과 양 코퍼스를 합쳐 발생빈도수 30개 이하를 제거하기로 하였다. 이렇게 해서 얻어진 키워드 목록의 상위 어휘 30개를 그래프로 나타내면 그림 5와 같다. 여기서 막대의 길이는 해당 어휘의 키값을 나타낸다.

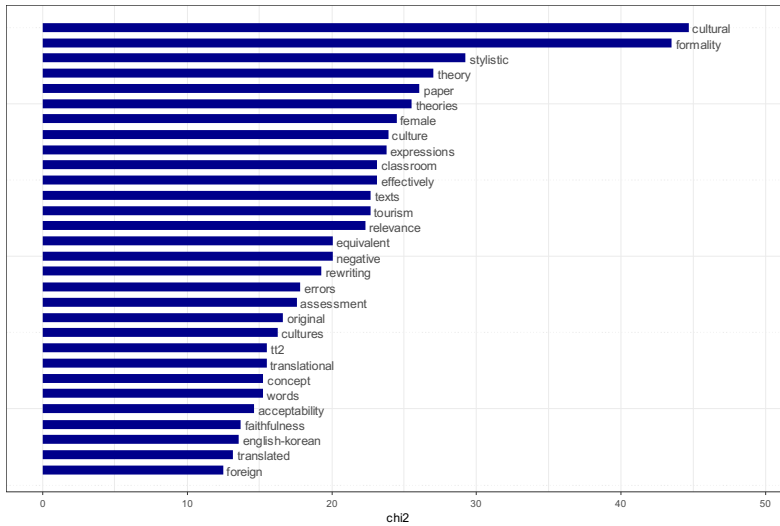
그림 4 도출된 어휘별 키값의 예

	feature	chi2	p	n_target	n_reference
1	cultural	44.66049	2.343381e-11	172	227
2	formality	43.48740	4.266987e-11	35	15
3	wordplays	40.03388	2.495965e-10	17	0
4	hangul	37.47810	9.244538e-10	16	0
5	numeral	32.94102	9.499751e-09	16	2
6	stylistic	29.26027	6.328002e-08	40	30
7	footnotes	28.66251	8.615554e-08	18	5
8	tqa	27.26291	1.775857e-07	12	0
9	theory	27.03336	1.999740e-07	83	100
10	paper	26.07206	3.289085e-07	364	670
11	theories	25.52246	4.372628e-07	50	48
12	female	24.47376	7.532867e-07	28	18
13	brand	24.32949	8.118666e-07	15	4
14	culture	23.91268	1.008053e-06	106	147
15	expressions	23.78225	1.078725e-06	108	151

그림 5를 보면 가장 두드러진 키워드는 문화와 관련된 어휘로 cultural, culture, cultures 등이 있다. 이를 보면 2010년대 초기에는 다른 시기와 대비하여 문화 관련 연구가 매우 높은 비중을 차지했음을 알 수 있다. 이 키워드들은 키값 뿐만 아니라 어휘빈도수에서도 매우 높게 나타난다. 가령, cultural의 경우는 분석코퍼스에서 172번, 대조코퍼스에서는 227번 발생한다. 대조코퍼스의 전체 어휘수는 104,628이고 분석코퍼스는 40,870으로 대조코퍼스가 두 배 이상 많다. 이같은 코퍼스 크기 차이를 고려하여 대조코퍼스에서의 cultural 어휘수를 대조코퍼스에 맞추면 89가 된다. 이는 분석코퍼스에서의 cultural의 빈도가 대조코퍼스에 비하여 두 배 가량 높다는 의미로 2010대년 초기의 통번역 연구에선 문화가 큰 화두였음을 알 수 있다.



그림 5 2010~2012의 키워드 (최소발생건수=30)



문화 관련 키워드에 이어 그림 5에서 상위를 차지한 키워드 중에는 formality(격식체), stylistic(문체적인)이 있다. 이는 문체와 관련된 어휘로 2010년대 초기에는 번역문의 문체를 분석한 연구 비중이 높았다는 것을 보여준다.

다음에 theories(이론들)란 키워드는 이 시기에는 각종 이론을 논한 연구가 많았음을 시사한다. 구체적으로 어떤 theories들이 언급되었는가를 알아 보기 위하여 quanteda 패키지의 콘코던스기능을 사용하여 그림 6과 같이 theories의 주변 텍스트를 검색해보았다. 그림 6은 전체 콘코던스의 일부만 캡처한 그림이다. 이를 보면 대부분 번역과 관련된 이론이고 일부 통역이나 언어학 이론과 관련된 경우도 있다. 어쨌든 2010년대 초는 통번역 이론을 논하는데 관심이 많은 시기였음을 확인할 수 있다.

그림 6 2010~2012, theories의 콘코던스 (일부)

on translation and incorporate translation	theories	into the practice of reviewing
are known to develop target-oriented	theories	in Translation Studies. How
, then, can their	theories	be connected with" acceptability
tries to reconsider these three	theories	around this concept. On
from the discussion of the	theories	of translation norms and the
one of the most prominent	theories	of interpretation( Interpretation is
, and the diverse translation	theories	that formed when many schools
A Metareflection on the Translation	Theories	of the Humanities How should
? scientificity? of translation	theories	of the humanities in section
, we describe briefly several	theories	of explicit and implicit translation
of explicit and implicit translation	theories	in the humanities. In
of metareflections on the translation	theories	of the humanities not only
theory of foregrounding among cognitive	theories	which are universally stylistic features
national literature and the translation	theories	. They greatly influenced the
. We should examine more	theories	of interpretation as broadly as
Cheonan. Drawing on the	theories	and concepts of frames,
Korean. According to earlier	theories	on Korean plurality, nouns
theory, mainly with the	theories	which Gutt has adopted in
acceptability is used in other	theories	in translation studies and in

female(여성의)이란 키워드는 번역에서 페미니즘을 다룬 연구와 연관이 있을 것으로 추정되었는데, 이를 확인하기 위하여 그림 7과 같이 콘코던스를 살펴본 결과 대부분이 female translator(여성번역가)나 female language(여성 언어)와 연결되어 있어 그와 같은 추정이 맞는 것임을 확인하였다.

그림 7 2010~2012, female의 콘코던스 (일부)

translation styles of male and	female	translators, in order to
used by both male and	female	translators' works, this
Chinese characters by male and	female	translators. Currently, in
Chines characters. However,	female	translators prefer to utilize Korean
professional identity. Male and	female	translators practice different styles of
Generally, the style of	female	translator, is that of
is that of the TL's	female	language, with a tendency
a tendency to utilize a	female	language when translating from the
life of O-Lan, the	female	protagonist of, is indicative
a Christian missionary and a	female	American writer who was not
innately or explicitly expressed through	female	infinitives in China's Confucian cultureare
marked gender differences within the	female	infinitives of the source text
do women translators work in	female	language? Several studies have
case; men translators use	female	language in their translations,
as written language?;	female	language as a chosen translation

그 외에 classroom(교실)은 통번역교육과 관련된 어휘로 문맥을 확인한 결과 수업활동이나 수업을 활용한 실험연구와 연관되어 있었다. tourism(관광)은 관광분야 텍스트 분석이 많았다는 것을 보여준다. relevance는 구트(Gutt 2000)의 relevance theory와 연관된 어휘로 해당 이론을 활용한 연구의 비중이 상대적으로 높았음을 보여준다. equivalent는 나이다(Nida 1965; Nida & Taber 1969)의 등가이론(equivalence theory)과 연관된 연구가 두드러졌음을 보여준다. 그 외에 errors(오류), assessment(평가), acceptability(수용도), faithfulness(충실도) 등은 번역 품질 평가와 관련된 어휘들로 분류된다. 특히 original이란 키워

드와 관련하여 원문과 비교한 번역 오류 분석이나 품질 평가가 많이 다뤄졌음을 알 수 있다. 이상을 종합하면 2010년대 초기는 문화, 이론, 페미니즘, 품질평가 등과 관련된 연구가 두드러진 시기였다고 할 수 있다.

다음에는 그림 8에 나와 있는 2010년 중기에 해당하는 2013~2015년의 키워드를 살펴보기로 한다. 이 시기의 초록, 즉 분석코퍼스의 총 어휘수는 50,974이고 대조코퍼스는 94,644로 대조코퍼스가 분석코퍼스에 비하여 두 배 정도 더 많다. 먼저 목록 상단에 있는 news와 websites는 분석텍스트 장르 중 뉴스 기사와 웹사이트가 두드러진 시기였음을 보여준다.

키워드 목록 상단에 있는 corpus와 목록 맨 아래 statistical(통계적인), corpus-based(코퍼스 기반의)는 코퍼스를 기반으로 한 정량분석 연구가 활성화된 시기라는 것을 보여준다. 분석코퍼스와 대조코퍼스에서 corpus의 어휘수는 각각 112, 120으로 대조코퍼스 어휘수를 분석코퍼스 크기에 맞추면 58이 된다. 즉, 이 키워드는 대조코퍼스에 비하여 분석코퍼스에서 2배 정도 더 많이 발생한다. 따라서 2010년대 중기에 들어서 코퍼스 관련 연구가 급격히 늘었음을 알 수 있다.

그림 8 2013~2015의 키워드 (최소발생건수=30)

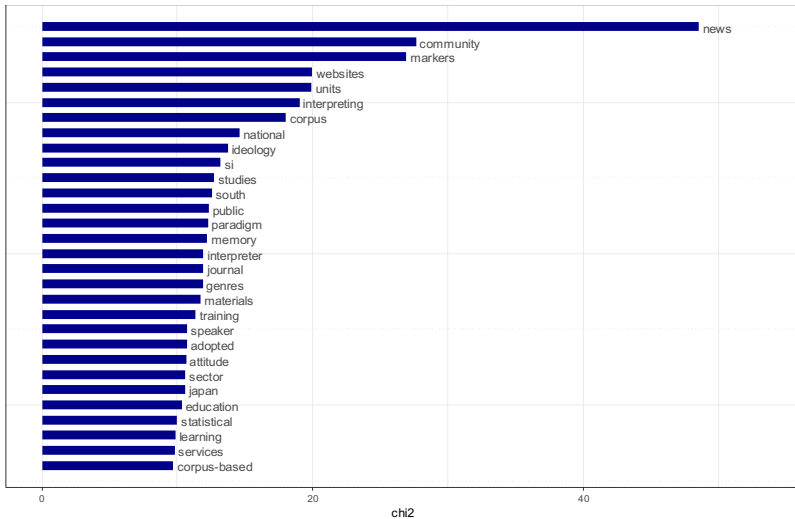


그림 8에서 interpreting(통역), si(동시통역), memory(기억), interpreter(통역사) 등의 키워드는 2010년대 중기에 통역에 관한 연구가 특별히 많았음을 보여준다. 어휘빈도수로 보면 interpreting의 경우 분석코퍼스는 363, 대조코퍼스는 500으로 대조코퍼스 어휘수를 분석코퍼스 크기에 맞추면 265가 된다. 즉, 분석코퍼스에서의 interpreting 어휘수가 대조코퍼스에 비하여 1.3배 더 많다. 이는 앞서 corpus 키워드에서의 2배 차이가 난 것에는 미치지 못한다. 따라서 다른 시기에도 통역에 관한 연구가 꾸준히 있었지만 2010년 중반에 상대적으로 많이 늘어났다는 정도로 해석할 수 있다.

중간에 public(공공)과 하단의 sector(분야)는 2010년대 중기는 공공번역에 관한 연구가 관심을 끈 시기였음을 보여준다. public 어휘수는 분석코퍼스 65, 대조코퍼스 66으로 대조코퍼스 어휘수를 분석코퍼스에 맞추면 34로 분석코퍼스에서 두 배 가량 많이 발생한다. 즉, 공공번역이란 주제는 2010년도 중반을 규정하는 핵심 이슈였다고 볼 수 있다.

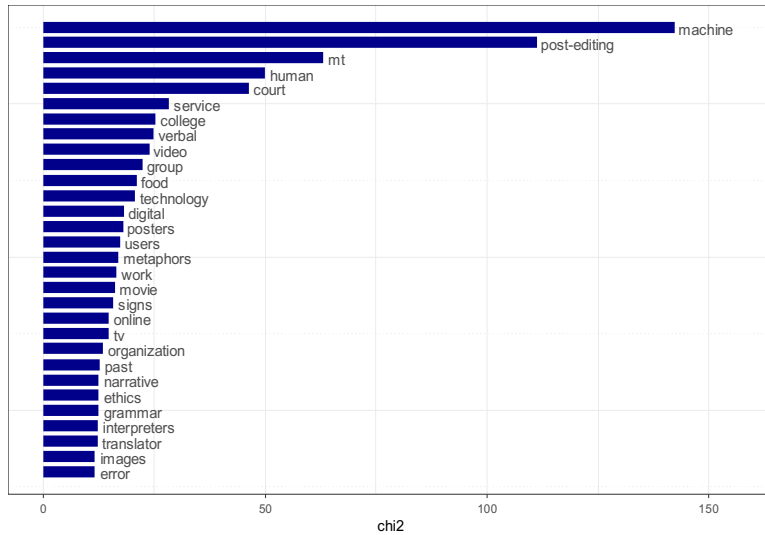
다음으로는 training(훈련), education(교육), learning(학습)으로 이어지는 통번역교육 관련 키워드는 2010년대 중기에 들어 통번역교육 연구에 대한 관심이 고조되었음을 보여준다. training의 경우 분석코퍼스 어휘수는 136, 대조코퍼스는 172이다. 대조코퍼스 어휘수를 분석코퍼스에 맞추면 91로 대조코퍼스가 1.5배 정도 더 많다. 따라서 2010년대 중기는 통번역교육 연구의 분수령을 이룬 시기였다고 볼 수 있다. 2010년대 초기 키워드 목록에 나타난 classroom과 비교해보면 2010년 초기 교실 위주의 교육과 연구 관심이 중기로 넘어오면서 전반적인 교육 훈련과 학습에 관한 연구로 외연이 확장되었다고 볼 수 있다.

그 외에 ideology는 통번역 행위를 이데올로기와 연결시킨 연구가 상대적으로 많았음을 시사하고, services는 그림 9에서 보듯이 주로 medical(의료)이나 community(지역사회), marriage migrant women(결혼이주여성) 같은 어휘와 어울려 의료 통역이나 결혼이주 여성들을 위한 지역사회 통역 등에 관한 연구가 관심을 끌기 시작하였음을 보여준다. 이상을 정리해보면 2010년 중기는 코퍼스 기반 연구가 붐을 이룬 시기이며 동시에 통역 과정, 특수 통역 및 통번역교육 관련 실증적 연구가 크게 약진한 시기라고 할 수 있다.

그림 9 2013~2015, services의 콘코던스 (일부)

started offering interpreting and translation	services	for marriage migrant women through
the community interpreting and translation	services	for marriage migrant women and
assessment, perception of the	services	performance and areas for further
and future of the relevant	services	in South Korea. A
translation quality of automated translation	services	as the user base for
lead in improving legal interpreting	services	in South Korea, Ewha
level of Korea's military language	services	.
as to offer favorable medical	services	for foreign patients. In
determining the quality of medical	services	. Some of suggestions for
agencies are providing community interpreting	services	. This paper provides an
provides an overview of interpreting	services	, including language services,
interpreting services, including language	services	, which are offered in
with managers of community interpreting	services	provided by seven government and
of the seven agencies'	services	and also the management of
to the current community interpreting	services	in South Korea and proposes
matic and professionalized community interpreting	services	.

그림 10 2016~2019의 키워드 (최소발생건수=30)



다음에는 그림 10에 나와 있는 2010년대 후기에 해당하는 2016~2019년의 키워드를 살펴보자. 이 시기에 분석코퍼스의 어휘수는 53,752, 대조코퍼스는 91,868이다. 이 시기에서 가장 눈에 띄는 키워드는 단연 기계번역과 관련된 것들이다. 키워드 목록에서 상위를 점하고 있는 machine(기계), post-editing(후편집), mt(기계번역), human(인간) 외에 technology(기술) 등 기계번역과 관련된 키워드들이 다수를 점하고 있다. 특히 이들 어휘는 다른 시기에서는 거의 발견되지 않는다는 점에서 2010년 후기에 새롭게 등장한 연구 문제인 것을 알 수 있다. 가령, machine의 경우 분석코퍼스에서 어휘수는 154, 대조코퍼스에서는 44

이다. 대조코퍼스의 machine 어휘수를 분석코퍼스에 맞춰보면 23으로 줄어든다. 154에 대비하여 23은 0.14퍼센트로 기계번역이 2010년 후기 이전에는 전혀 관심을 끌지 않았다는 점을 보여준다. 이는 기계번역이 통번역 연구의 화두로 등장한 것이 2016년 구글번역기에 신경망기계번역 기술이 도입된 이후란 점을 반영한다. 또한 human이 키워드로 등장한 것은 기계번역에 관한 연구의 상당수가 인간번역과 기계번역을 대비하였다는 점을 시사한다.

이와 관련하여 그림 10에서 관심을 끄는 키워드는 목록 맨 하단에 error이다. 이는 그림 11의 콘코던스에서 보듯이 주로 error rate(오류발생율), error types(에러 형태) 같은 구와 연결되어 발견되는데 상당부분이 기계번역의 품질 평가와 연결되어 있다. 통번역 연구에서 ‘오류’에 관한 연구는 연구 초기에 유행했던 연구인데 2010년대 후기에 들어 기계번역이 부상하면서 번역에서의 오류 문제가 다시 관심을 끈 것은 흥미로운 현상이다.

그림 11 2016~2019, error의 콘코던스 (일부)

presented the criteria for translation	error	analysis by inductive reasoning.
enable qualitative evaluation. The	error	scores gained in this manner
measuring the severity of an	error	. To strengthen the objectivity
different evaluators for an identical	error	. An additional element of
. However, the punctuation	error	rate by machine translation was
was 24.1% while the	error	rate by human translation was
the creation of machine translation	error	typology based upon theoretical reviews
. The paper also analyzes	error	types with high frequency in
human translation. The major	error	types observed in machine translation
Korean and in Russian,	error	occurred as the Korean grammar
Russian syntax incurred the most	error	( for example, commas
interpretation, based on an	error	analysis of Korean-Chinese consecutive interpretati
nterpreters. After including additional	error	categories that did not fall
than Google Translate. The	error	analyses of incorrect translations from

그 외에 그림 10에서 court(법정), video(비디오), food(음식), posters(포스터), metaphor(은유), movie(영화), tv, narrative(나레티브), ethics(윤리), signs(기호), images(이미지) 등의 키워드는 법정 통역, 비디오 게임/사막, 한국음식 번역, 나레티브 구조 분석, 통번역의 윤리, 이미지에 대한 기호학 분석 등의 연구 주제와 연관 있다. 다른 시기와 비교하여 2010년대 후기에 이같이 세부 분야의 연구 주제와 관련된 키워드들이 부각된 점은 최근 들어 통번역의 연구 주제가 다양화되고 있음을 보여준다. 이상의 분석을 종합해보면 2010년 후기는 기계번역 연구가 봄을 이룬 시기이며 동시에 연구주제가 다양화-전문화된 시기로 정의할 수 있다.

그런데 한 가지 흥미로운 사실은 2018년도에서 2019년도로 넘어오면서 기

계번역 연구 분야가 점차 가라앉고 있는 경향을 보인다는 점이다. 이것은 그림 12, 그림 13의 연도별 키워드를 비교해 보면 알 수 있다. 2018년도 키워드를 보여주는 그림 12에서는 post-editing과 machine이 키값에서 압도적인 상위를 차지하고 있는데 반하여 그림 13의 2019년도 키워드 목록에서는 기계번역과 관련된 키워드는 mt만 남고 political, food, proper, images와 같이 다른 세부 연구 주제와 관련된 키워드들이 상위에 올라있다. 이는 2018년도를 정점으로 연구 주제로서 기계번역에 대한 관심이 줄어들기 시작했음을 보여준다.

그림 12 2018년도 키워드 (최소발생건수=30)

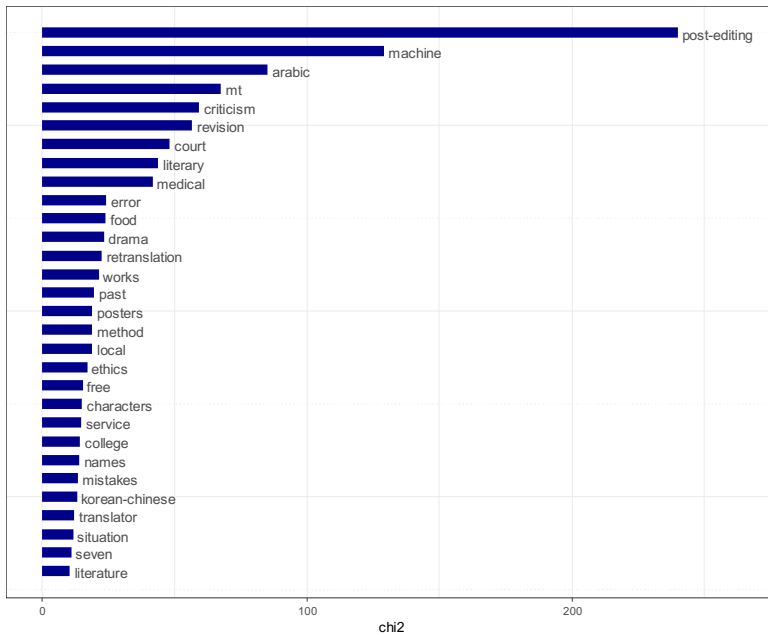
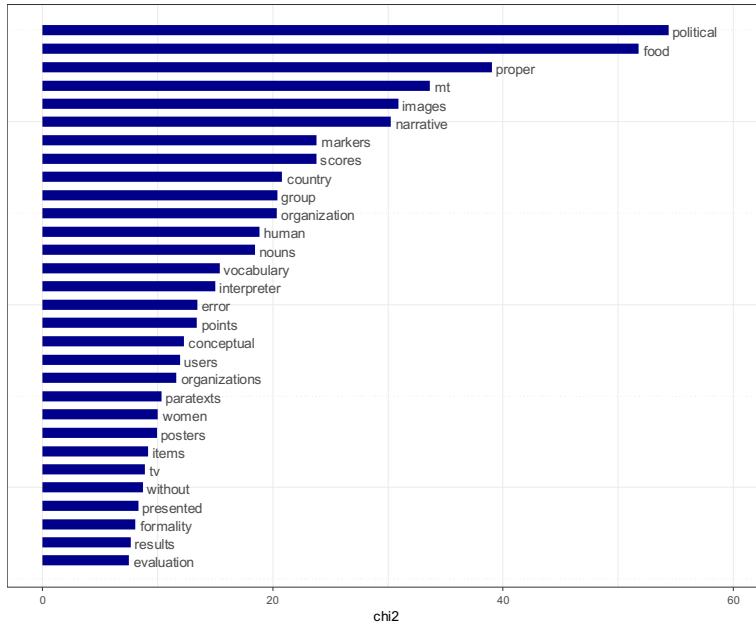


그림 13 2019년도 키워드 (최소발생건수=30)

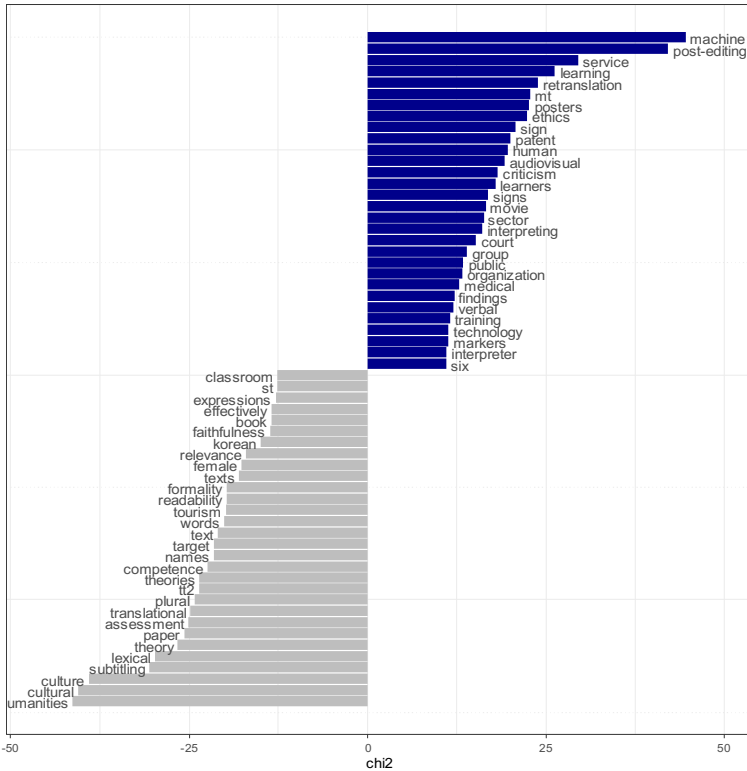


#### 4.2. 상-하반기 비교 분석

이번에는 2010년대를 5년 단위로 나눠서 상-하반기의 키워드를 비교해보자. 분석 결과는 그림 14에 나와 있는데 그래프 왼쪽 하단이 상반기, 오른쪽 상단이 하반기 키워드이다. 이 그래프에는 4.2절에서 분석했던 초-중-후기 시대별 분석에서 초기와 후기에 두드러졌던 키워드들이 포함되어 있는 반면 **public**, **corpus** 같은 중기의 키워드들이 빠져있다. 그것은 중기에 해당하는 데이터가 전-후반기로 나뉘지면서 키워드로서 위치를 잃어버렸기 때문이다. 이는 국내 통번역 연구의 흐름을 분석하는데 5년 단위는 너무 길다는 점을 시사한다. 그만큼 연구 주제의 변화가 빠르다고 볼 수 있다. 그렇지만 그림 14와 같이 상-하반기로 나뉘었을 때 몇 가지 새로운 키워드가 눈에 보인다.



그림 14 2010년대 상-하반기 키워드 비교 (최소발생건수=30)

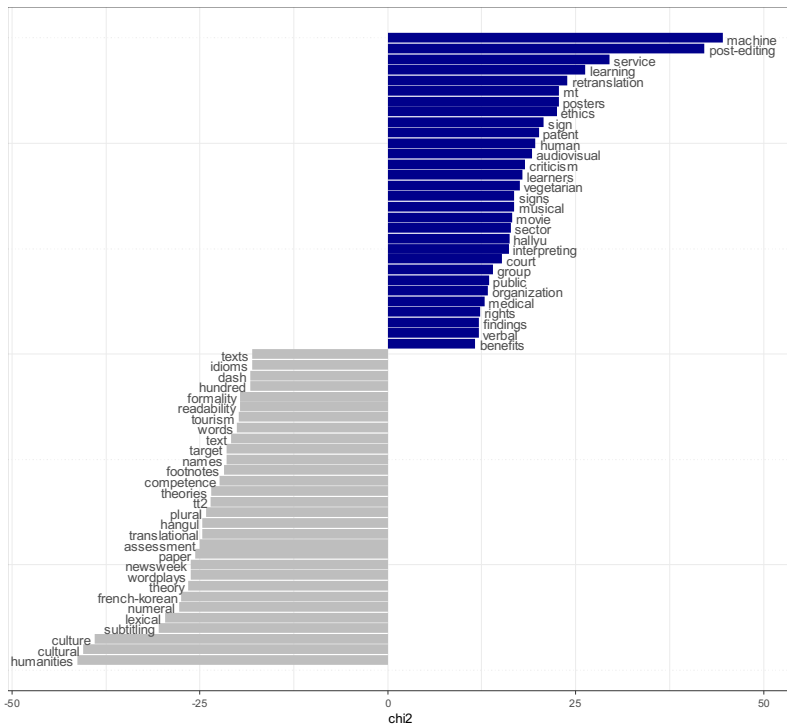


먼저 상반기에서는 subtitling(자막), competence(능력), readability(가독성) 등의 키워드가 새로 포함되었다. 이는 자막번역, 통-번역사 능력 및 번역시 가독성에 관한 연구가 상반기에 꾸준한 맥을 유지했다 것을 의미한다. 하반기에서는 audiovisual(시청각)이란 키워드가 새로 등장했는데 이는 전반기의 book 키워드와 대비되어 2010년 하반기에서는 번역연구의 분석대상이 다매체로 확장된 것을 반영하는 것으로 해석된다. 또한 medical(의료), court(법정)가 하반기를 특징짓는 키워드로 포함되었는데, 이 역시 의료-법정통역에 대한 관심이 2010년도 하반기에 일어난 것을 반영한다. 또한 retranslation(재번역), criticism((번역)비평)등도 하반기에 주로 진행된 연구를 반영한다.

이번에는 그림 15와 같이 최소발생건수를 제한하지 않은 상태에서 키워드

를 살펴보면 상-하반기의 특이한 연구 주제들이 발견된다. 상반기 경우는 hangul(한글), plural(복수형), footnotes(각주), names(이름), words(단어), dash(대시), idioms(관용구)와 같은 키워드가 새로 등장한다. 이는 2010년대 상반기에 특정한 표기나 표현 양식 등 텍스트의 특정한 부분에 초점을 맞춘 연구가 유행했음을 보여준다. 하반기에는 vegetarian(채식주의자), musical(뮤지컬), hallyu(한류) 같은 키워드가 눈에 띈다. 이중 vegetarian은 데보라 스미스(Deborah Smith)의 영역으로 맨부커 상을 수상한 ‘채식주의자 번역과 관련된 연구를 지칭한다. hallyu는 번역에서 ‘한류를 관심 주제로 다룬 연구를 칭한다. 이를 보면 2010년 하반기에는 사회적으로 이슈가 된 현상을 쫓는 연구도 하나의 흐름을 형성하였음을 알 수 있다.

그림 15 2010년대 상-하반기 키워드 비교 (최소발생건수=0)



## 5. 결론

4절의 분석에서 2010년을 초-중-후기로 나뉘었을 때 각 시대별로 뚜렷한 연구 경향을 발견할 수 있었다. 이런 경향 중 가장 괄목할만 한 것을 추려내어 큰 흐름을 짚어 본다면 초기는 ‘통번역 이론’, 중기는 ‘실증적 연구’, 후기는 ‘기계번역의 시기’라고 할 수 있다.

세계적 차원에서의 통번역연구는 일련의 전환기를 거쳐 왔다. 소위 어떤 ‘턴(turn)이라고 명명되는 변화로 1960년대와 70년대는 언어학적 관점에서 번역을 논하던 ‘언어학적 턴’의 시대라고 할 수 있다(스넬-혼비(Snell-Hornby) 2006). 이 시대의 대표적인 학자로는 촘스키의 변형생성문법에 기초하여 번역이론을 정립한 나이다(1964)를 들 수 있다. 1980년대에 기술론적번역학(descriptive translation studies: DTS)이나 스코포스 이론을 중심으로 한 기능주의 이론 등이 등장하면서 언어학적 접근법에서 벗어나는 ‘문화적 턴’을 맞이하였다. 그리고 90년대에는 실증적 연구를 중시하는 ‘실증적 턴’이 도래하였다(스넬-혼비 2006). DTS는 1990년 말에서 2000년 초에 베이커(Baker 1996; 2004)에 의하여 도입된 코퍼스분석 기법과 맞물려 발전하였다.

앞서 4절에서 분석한 2010~2012년도의 키워드 목록을 보면 *stylistic, equivalent, theories, relevance, faithfulness, acceptability* 같은 키워드들이 등장하는데 이들은 언어학 또는 화용론적 관점에서 번역에 접근한 이론과 연결된다. 이를 보면 2010년대 초기의 국내 통번역연구는 국외에서는 60년대와 70년대에 유행하였던 언어학적 연구와 번역이론을 뒤늦게 따라가는 경향이 아직 강하게 남아있던 시기라고 할 수 있다. 그 와중에 문화 번역에 대한 관심이 강하게 표출되기도 하였다. 2010년대 중기에 해당하는 2013~2015년도의 키워드 분석에서는 *interpreting, memory, si나 corpus, statistical*과 같은 키워드가 등장했는데, 이는 2010년 중반 대에 이르러서야 국내 연구가 해외의 ‘실증적 턴’을 따라가기 시작했다는 것을 의미한다. 중기에 두드러진 *training, education*과 같은 교육 훈련 분야도 대부분 실험이나 실증 데이터를 기반으로 한다는 점에서 궤를 같이 한다. 2010년도 후기인 2016~2019년에는 기계번역이 태풍처럼 몰아치면서 최소한 국내에서는 ‘기술적 턴(technological turn)’을 한 시기로 볼 수 있다. 그 와 동시에 연구 주제가 다양화되면서 연구 범위가 확장되는 양상이 나타났다.

그러나 2019년 들어 연구 주제로서 기계번역에 대한 관심이 줄어들고 있어 기술적 턱이 일시적 관심사로 끝날지 아니면 통번역연구의 중요한 흐름으로 자리매김 할지는 좀 더 지켜볼 일이다.

본 연구에서는 분석기간을 몇 개의 시기로 나눠 상대적으로 많이 등장한 키워드를 분석하였기 때문에 분석기간 전체에 걸쳐서 꾸준한 경향을 보인 연구 주제는 키워드에서 빠진 점은 유의할 필요가 있다. 가령, 문학번역을 지칭하는 literary(문학적)란 어휘는 개별 분석 시기의 빈도수에서 항상 상위권에 들었지만, 바로 그 이유 때문에 시기별 키워드 분석에는 이름을 올리지 못했다. 본 연구의 제약점으로는 통역과 번역을 나누지 않았기 때문에 시대별 연구 동향이 둘 중 어느 분야를 더 반영하는지가 명확하지 않다는 점을 들 수 있다. 또한 모든 학술지를 하나의 데이터로 통합 분석하였기 때문에 학술지 간에 연구 경향에 차이가 있는지도 앞으로 규명이 필요한 부분이라 하겠다.

### 참고문헌

- 김혜림 (2013). 「한국의 당대(1999-2013) 번역학 연구 동향」, 『번역학연구』 14(5): 95-115.
- 이건일, 이종혁 (2015) 「한국 심층학습을 이용한 기계번역 연구 동향」, 『정보과학회지』 33(10): 40-47.
- 이창수 (2018). 「국내 번역연구의 어제와 오늘: 텍스트 마이닝 분석 - 『번역학연구』를 중심으로」, 『번역학연구』 19(2): 233-252.
- 정호정 (2013). 「번역학 전문학술지의 발달과 최근 연구 동향 - 1999년 이후 국내외 대표 학술지 분석」, 『통역과 번역』 15(1): 235-256.
- 조성은 (2014). 「영상번역 연구의 동향과 전망」, 『번역학연구』 15(2): 205-224.
- 최희경 (2015). 「코퍼스를 활용한 번역 연구」, 『번역학연구』 16(2): 317-338.
- 한현희 (2016). 한국 통역학 연구의 동향과 과제: 해외 연구 동향과의 비교를 기반으로., 『번역학연구』 17(2): 251-286.
- 한현희 (2017) 「한국의 통번역 교육 연구 동향에 관한 메타분석 연구」, 『번역학연구』 18(1): 171-202.

- Baker, Monda (1996). 'Corpus-based Translation Studies: the Challenges that Lie Ahead', in Harold Somers (ed.) *Terminology, LSP and Translation, Studies in Language Engineering in Honour of Iuan C. Sager*, Amsterdam: John Benjamins, 175-186.
- Baker, Mona (2004). 'A Corpus-Based View of Similarity and Difference in Translation', *International Journal of Corpus Linguistics*, 9(2): 167-193.
- Fortuna, Blaz, Carolina Galleguillos and Nello Cristianini (2009). 'Detection of Bias in Media Outlets with Statistical Learning Methods', in Ashok N. Srivastava and Mehran Sahami (eds.) *Text Mining: Classification, Clustering, and Applications*, Boca Raton: CRC Press, 27-50.
- Gutt, E.A. (2000). *Translation and Relevance: Cognition and Context*, St. Jerome Publishing, Manchester.
- Nida, Eugene A (1964). *Toward a Science of Translating, Bible Translating*, Leiden: Brill.
- Nida, Eugene A, and Charles R. Taber. (1969). *The Theory and Practice of Translation, With Special Reference to Bible Translating*, Leiden: Brill.
- Snell-Hornby, Mary (2006). *The Turns of Translation Studies: New Paradigms or Shifting Viewpoints*, Amsterdam: John Benjamins.
- Waldherr, Annie., Gerhard Heyer, Patrick Jähnichen, Andreas Niekler and Gregor Wiedemann (2016). 'Mining Big Data With Computational Methods', in Gerhard Vowe and Philip Henn (eds.) *Political Communication in the Online World: Theoretical Approaches and Research Designs*, London: Routledge, 201-217.

[Abstract]

### A 'Keyword' Analysis of Major Trends in Domestic Interpretation & Translation Research of 2010s

Lee, Chang-Soo

(HanKuk University of Foreign Studies)

The current research adopts a corpus-based approach to uncovering major trends in the domestic interpretation and translation (T&I) research of the 2010s. The corpus analyzed consists of 1,299 abstracts collected from four major domestic academic journals. The corpus was divided, first, into three periods and, second, into two halves, and keywords were extracted from the subcorpora to identify most prominent research trends for each period. The analysis showed that the early 2010s (2010~2012) was marked by strong interest in theories, culture, feminism and quality assessment. The mid-2010s (2013~2015) saw a surge of empirical studies in interpreting and corpus-based research in translation, along with marked attention given to training and education. The late 2010s (2016~2019) was characterized by a proliferation of research related to machine translation, which centered around comparing human and machine translation at various levels and exploring post-editing as a way of compensating for observed weaknesses of machine translation. This period also witnessed the scope of T&I research widening to include specialized interpretation and translation services and a focus of attention on multimodal texts. The two-period analysis largely duplicated the findings from the three-period investigation.

▶ Key Words: research trends, interpretation, translation, corpus-based, keywords

▶ 주제어: 연구 동향, 통역, 번역, 코퍼스 기반, 키워드

이창수

한국외국어대학교 통번역대학원 교수

soolee@hanmail.net

관심분야: 코퍼스언어학, 사회언어학, 전산문체학, 담화분석

논문투고일: 2019년 11월 3일

심사완료일: 2019년 12월 2일

게재확정일: 2019년 12월 15일