

## 중한 전문분야 통역 준비를 위한 코퍼스 기반 용어집 구축에 관한 예비 연구: 연설문에 등장한 경제 및 금융 용어를 중심으로

이지혜 · 이용훈  
(이화여대 · 충남대)

### 1. 서론

통역사는 고객으로부터 통역 의뢰를 받는 순간부터 통역 준비 업무가 시작된다. 통역 전 준비 업무는 회의 주제 관련 자료를 읽으며 주제 지식을 습득하는 ‘주제 준비’, 통역해야 할 분야의 전문용어를 검색하고 정리, 암기하는 ‘용어 준비’, 회의 직전 곧 발표될 연설문 내용 확인하는 ‘마지막 순간 준비’(최정화 2004:4)로 세분화된다. 회의 주제 관련 자료를 보면서 용어를 정리하는 동안 주제 지식도 함께 축적되므로 주제와 용어 준비는 동시에 진행된다.

통역사가 모든 분야의 전문가가 아니기 때문에 통역 준비 기간 동안 관련 용어와 주제 지식을 축적하기 위해 많은 시간과 노력을 기울인다. 문제는 통역사에게 전문용어 및 배경지식을 충분히 쌓을 수 있는 시간이 많이 주어지지 않

는다는 점이다. 회의 자료가 사전에 공개되지 않거나 수정되는 등 여러 가지 이유로 인하여 회의에서 무슨 내용이 나오는지도 모른 채 회의 주제만 알고 준비를 해야 하는 경우도 많다. 통역을 해 본 적이 없는 생소한 주제이거나 전문분야 통역일수록 짧은 준비 기간이라는 시간적 압박 하에 통역사는 출발어와 도착어의 언어 지식과 전문 지식 등을 익혀 전문가의 담화를 이해하고 유창하게 통역해야 하므로 준비과정에서 심적 부담을 느낄 수밖에 없다. 아무리 언어 지식이 풍부한 통역사라도 특정 분야 및 특정 단체에서 사용하는 주제별 구문 및 문체 표현, 전문적인 정보가 담긴 전문용어(Kalina 2006; Fantinuoli 2017: 27)와 주제 준비를 소홀히 할 경우 전문가 화자의 말을 제대로 이해하기 어려우며 인지적 부하가 걸리기 쉽다.

그러나 통역 준비 과정에서 축적한 용어 및 주제 지식은 출발어에서 도착어로의 재구성 단계에서 통역사가 보다 정보를 빠르고 정확하게 예측하고 올바르게 의미를 구성하도록 돕고(Stoll 2009; De Groot 1999; Fantinuoli 2017: 27) 올바른 용어 사용은 정확하고 완전한 의사소통에 영향을 준다(Amtz, Picht and Mayer 2009; Fantinuoli 2017: 28). 게다가 전문통역사의 통역에 대한 평가 항목 중 통역사의 직업의식에 초점을 맞추고 있는 용어 사용의 정확도, 전문지식의 이해도, 자료 준비 성실성 (Bühler 1986; Chiaro and Nocella 2004; Kurz 1993; Moser 1996; Pöchhacker 2001; 원종화 2014)은 통역 평가에서 빠지지 않는다. 청중과 연사를 대상으로 한 통역 평가에서도 올바른 용어 사용을 중요하게 여기고 있으며(Kurz 2001; 박현일 2019) 과거보다 훨씬 더 정확한 용어로 통역하기를 바라는 청중의 요구도 커지고 있다(Kalina 2007; Fantinuoli 2017: 28). 이처럼 통역사의 인지적 측면, 시장의 니즈 및 통역 품질 측면 등에서 전문용어와 주제 지식이 중요한 요소임을 감안할 때 통역 수행 시 전문가 연사와 준전문가 통역사의 언어 및 주제 지식 등의 격차를 최소화하기 위해 통역 수행 전 준비 단계가 중요하다.

통역 준비 과정에서 전문용어 구축과 주제 지식 축적을 보다 효율적으로 할 수 있도록 기술적 도구를 활용할 것을 주장한 환티누올리(Fantinuoli 2017)는 ‘코퍼스 기반 통역 준비(CDIP, Corpus-driven Interpreter Preparation)’라는 개념을 제시하였다. CDIP는 용어추출도구를 통해 대량의 텍스트에서 전문용어를 추출하고, 통역사가 학습자가 되어 대량의 데이터를 코퍼스 분석 도구인 콘

코댄서(concordancer)를 통해 용례를 확인하며 언어와 패턴을 인식하고 언어, 주제에 대해 통찰력을 얻는 것을 추구한다(Fantinuoli 2017). 물론 통역 평가 기준을 단순히 어휘 의미 차원이 아닌 담화 차원으로의 평가를 지향해야겠지만 어휘가 지식 습득에 있어서 가장 기본적인 단위이자 전문분야 담화의 언어적 특징을 결정하므로 통역사에게 전문 분야 담화를 구성하는 어휘 능력은 중요하다(Garzone 2000: 82 ; 최문선 2019: 248). 본고에서도 용어집 구축은 전문 분야 통역 준비를 위한 최소한의 준비이자 최선의 준비 방법이라는 판단 하에 어떠한 용어를 어떻게 준비를 할지, 그리고 준비한 용어가 통역 수행에 어느 정도 기여할 수 있는지를 간접적으로 판단하고자 본 연구를 설계하였다.

이러한 배경 하에 본고에서는 경제 및 금융 전문분야 회의에서 사용된 중국어 연설문으로 구축한 자가코퍼스를 대상으로 AntConc 코퍼스 분석 도구를 활용하여 코퍼스에 출현한 경제 및 금융 분야의 단일명사와 합성명사를 중심으로 용어집을 구축한 뒤 용어의 연도별 출현 양상을 비교 분석한다. 나아가 용어집의 효용성을 간접적으로 확인하기 위해 특정 연도의 경제 및 금융 회의를 준비한다는 가정 하에 회의에 등장한 경제 및 금융 관련 용어가 앞서 구축한 용어집에서 얼마나 동일하게 출현하였는지를 살펴보고자 한다.

이를 위해 2장에서는 전문용어 구축에 관한 선행연구를 살펴보고, 3장에서는 AntConc를 통해 경제 및 금융 회의 연설문 텍스트에 출현한 용어 추출 및 용어집 구축 과정을 기술하겠다. 4장에서는 경제 및 금융 관련 단일명사 및 합성명사의 연도별 출현 양상을 살펴본 뒤 같은 회의의 특정 연도 회의에 등장한 용어와 비교하여 용어집의 효용성을 간접적으로 파악하고자 한다. 5장에서는 본고의 결론과 한계점을 논의하는 것으로 마무리한다.

## 2. 선행연구 분석

### 2.1 통역학 내 코퍼스 기반 전문용어 관련 연구

전문분야 회의 통역 준비를 위한 용어집 구축에서 전문용어는 빠질 수 없는 중요한 요소이다. 전문용어의 사전적 의미는 특정한 전문 분야에서 주로 사

용하는 용어로서(국립국어원 표준국어대사전 2020) 회의장에서의 연사는 주제 분야의 전문가이므로 전문용어 또는 직업어로 특징지어질 수 있는 이들의 발화를(이미경 2018: 3) 준전문가인 통역사가 이해하고 전문가 및 일반인 등 청중을 대상으로 도착어로 전달해야 한다.

통역학에서의 코퍼스 기반 전문용어에 관한 연구는 전통적인 방식의 용어집 구축 과정에 기술을 접목하여 용어집 구축 과정에 효율화를 추구한 연구가 진행되었다. 환티누올리(Fantinuoli 2017)는 통역사의 직관에 따라 부정확하게 통역하는 것을 방지하기 위해 실제 사용 맥락이 담겨 있는 대량의 데이터에서 추출한 전문용어 목록에서 통역사가 언어 지식을 학습할 것을 주장하였다. CDIP의 목표는 대량의 데이터에서 코퍼스를 통해 추출한 전문용어 목록을 구축하여 문맥에서 의미와 용례를 발견하고 지식을 축적함으로써 통역 준비 과정을 최적화하는 것이다. 단순히 용어 정리만으로는 실제 통역을 수행하는 과정에 큰 도움이 되지 않기 때문에 통역사가 필요로 하는 ‘지식과 언어 학습’을 하는 준비 과정에서 실제 맥락을 통해 용어와 내용을 익힐 것을 주장하였다(Fantinuoli 2006: 174). 고르양크(Gorjanc 2009; 최문선 2018)는 용어 지식이 통역 수행의 성과를 가르는 의료 통역 분야에서 코퍼스 분석 도구인 ‘스케치엔진(SketchEngine)’을 통해 웹문서 기반 코퍼스를 구축하고 도구가 제공하는 다양한 기능을 활용하는 코퍼스 기반 통역 준비 과정을 연구하였다. 코퍼스를 활용하여 전문분야 통역을 위한 용어집 구축에 관한 대표적인 국내 연구로서 코퍼스 기반 통번역 학습자용 금융 용어 언어 사전 구축 방안(이지은·최효은 2013)과 전문분야 통역을 위해 기술 도구를 활용하여 용어집을 구축한 연구(최문선 2018, 2019)등이 있다.

이지은과 최효은(2013)은 코퍼스의 언어 정보와 실사 용례를 활용할 수 있는 가능성을 살펴보고자 <이진영의 동시통역 기초 사전>에 수록된 주식 관련 전문용어를 중심으로 세종 코퍼스와 ‘The Corpus of Contemporary American English(COCA) 코퍼스를 활용하여 통번역 학습자용 한국어 언어 사전과 한영 언어 사전 구축을 시도하였다. COCA 및 세종 코퍼스 내 통계 메뉴의 ‘언어 구성어’를 통해 언어 및 용례를 추출하고 검색어와 함께 자주 쓰이는 언어 표현을 제공함으로써 기존의 통번역 사전에 비해 전문용어의 의미 파악이 용이하고 훨씬 다양한 언어 조합과 풍부한 실사 용례를 제공하였다. 그러나 코퍼스의 한

계로 인해 기존의 통번역 사전에서 제시하고 있는 전문용어 중 상당수에 대해 검색 결과를 얻을 수 없었으며, 특히 영어 코퍼스의 경우 ‘한국식 전문용어’의 용례를 찾을 수 없었다고 지적하며 적절한 코퍼스를 구축하는 것이 중요한 선행과제임을 언급하였다(이지은·최효은 2013:141). 본고에서는 중간 전문분야 통역 준비를 위한 용어집 구축을 위해 실제 연사 발언에서 쓰인 전문용어를 추출하고자 2008년~2018년 경제 및 금융 관련 회의에서 사용한 연설문 텍스트를 대상으로 자가코퍼스를 구축하였다.

최문선(2018)은 환티누올리의 CDIP 개념을 토대로 전문분야 통역을 위한 용어집 구축에 기술 도구의 적용 가능성을 고찰하였다. 국내의 전문 통역사는 보통 인터넷상에서 회의 주제를 중심으로 키워드를 만들어 출발어와 도착어로 된 자료를 검색한 후 유용한 자료를 선별하여 읽으면서 추출한 용어를 엑셀과 같은 문서 프로그램이나 종이에 정리하며 수작업으로 용어집을 구축한다(최문선 2018). 이러한 용어집 구축 과정에 BootCaT 기술과 용어추출도구를 적용하여 생성한 용어집과 수작업으로 구축한 기존 용어집을 비교한 결과 용어집 구축의 속도 향상과 추출 범위의 확대는 물론 실제 IT 통역에서 사용된 용어와의 적중률도 75% 나왔다. 또한 최문선(2019)은 IT 분야를 중심으로 코퍼스에 대한 키워드 분석 후 정제 과정을 거쳐 영어에서 한국어로 통역하는 데 필수적인 용어를 도출하고 그 특징을 살펴보았다. IT 전문 담화 용어 중 품사별 비중을 분석한 결과 일반 명사가 절반 이상을 차지하였으며 축약어, 고유명사, 동사, 형용사와 부사 순으로 많이 쓰였다. 텍스트 커버율에서 명사가 많은 비중을 차지한 분석 결과는 본고에서도 확인되었으며 단기간의 통역 준비에서 기술적 도구를 활용하여 용어집을 구축하고 용어 출현 양상과 그 특징을 파악하여 통역 준비에 활용할 필요가 있다.

통역용 용어집 구축 및 활용은 전문용어의 적절한 도착어 대응 표현을 정리하고 정비하는 데 충분한 시간을 할애할 수 있는 번역(이미경 2018)에서의 용어집 구축 및 활용과는 조금 다르다. 통역 당일까지 통역사는 용어집을 준비할 수 있는 시간적 제약과 통역 수행 과정에서 용어를 검색하고 활용할 수 있는 시간 및 환경적 제약을 받는다. 심지어 연사가 어떤 발화를 할지도 모르는 상황이 비일비재하기 때문에 통역 준비 기간 동안 해당 분야의 용어 지식과 배경지식을 최대한 빠르고 폭넓게 쌓아야 한다. 이를 위해 통역사는 문어 텍스트

에서 등장하는 용어보다는 ‘연사 발언, 구어 텍스트에서 많이 등장할 가능성이 높은 용어’, 즉 ‘통역용 용어집’을 구축하고 활용할 필요가 있다. 수시로 검색할 수 있는 번역용 용어집의 활용 목적과 달리 통역용 용어집의 용어는 실제 통역에서 높은 적중률을 보여야 할 것이며 도착어로의 정확하고 신속한 전환에도 도움이 되도록 구축해야 한다. 따라서 통역용 용어집 구축은 전공서적, 사전, 신문 등 문어 텍스트에 등장하는 용어를 정비하고 사용 양상을 분석하는 등 기존의 언어학적, 번역학 관점과 다르게 접근할 필요가 있다. 이는 전문용어의 사용이 통역 과정에 미치는 영향에 관해 연구한 이미경(2018)의 연구 결과와도 맥락을 같이한다.

이미경(2018)은 전문분야 회의임에도 불구하고 실제로 전문용어나 표현이 차지하는 비율이 극히 소수였으며 연사가 사용한 용어에는 전문용어의 사용 빈도가 그리 높지 않아 통역사가 숙달해야 하는 전문용어의 양이 생각보다 많지 않고 한정되어 있다고 언급하였다. 즉, 전문분야 회의더라도 일반적으로 이해하기 쉬운 용어와 표현을 많이 사용하고 있기 때문에 기존의 통번역 사전에서 제시하고 있는 전문용어가 실제 연사 발화와 전문서적 및 사전 등 문어 텍스트에서 얼마나 많이 출현하는지에 대해서도 고민해 볼 필요가 있음을 시사한다. 예를 들어 경제 분야 관련 전문분야 회의를 준비한다고 가정했을 때 통번역 사전의 경제 및 금융 부분에 나온 전문용어를 암기하는 것은 비효율적이며, 주로 인터넷과 회의 자료를 대상으로 용어집을 구축하더라도 시간적 제약이 있기 때문에 인터넷에서 나온 용어를 모두 정리할 수도 학습할 수도 없다. 전문분야 회의 통역 준비를 위한 용어집 구축은 단기간의 통역 준비 기간 동안 문어 텍스트가 아니라 연사가 실제 발화한 연설문인 구어 텍스트를 연구 대상으로 하여 통역에 많이 나오거나 또는 많이 나올 확률이 높은 용어를 중심으로 구축하는 것이 중요하다. 또한 구어 텍스트, 연사 발언에는 꼭 전문용어만 등장하는 것이 아니라 전문 분야라 하더라도 대중적으로 알려진 공공용어적인 속성을 지닌 용어도 많이 등장한다. 특히 일반 청중이 많은 회의일수록 연사 발언에서 전문용어의 비중이 적을 확률이 높기 때문에 청중에 따라 전문용어의 출현 비중도 달라질 것이다. 따라서 ‘통역 준비 차원’에서의 ‘특정 분야 연설문 텍스트에 등장한 전문용어를 포괄한 용어’로 범위를 확장하여 통역용 용어집을 구축해야 한다. 통역용 용어집 구축을 위한 용어 선정 및 분류 기준을 마련하기 위해 다음 몇 가

지 분류 기준을 고려해 보았다.

## 2.2 전문 분야 용어집 구축 기준

통역 전달까지 전문분야에서 어떤 용어를 대상으로 용어집을 구성할지, 몇 개의 용어를 우선적으로 선택할지, 어떤 기준으로 용어를 선택할지 등은 중요한 고려 사항이다(Todd 2017).

첫째, 품사적 차원에서 전문용어는 주로 명사/명사구로 이루어지므로 전문 담화를 이해하는 데 핵심요소인 명사/명사구는 용어집의 기본 구성 대상이다. 물론 명사/명사구와 자주 호응되며 원문 이해에 핵심적인 특정 동사, 형용사와 같은 서술어(최문선 2018; 이미경 2018; 남슬기 2018 외)도 용어집에 포함될 수 있다. 전문용어는 체언에만 국한되지 않으며 언어학적 관점에서 일정한 개념을 나타내는 용어가 늘 일정한 용언과 공기하여 쓰이는 경우가 많으므로 언어 정보도 기술되어야 한다(강현화 1998). 통역사는 명사만 듣고 통역을 하는 것이 아니라 명사 주변에 있는 어휘 덩어리를 대상으로 하기 때문에 후속 연구에서는 명사/명사구와 자주 공기되는 서술어, 언어 정보가 담긴 언어 사전 구축을 통해 통역 준비와 연습에 활용하고자 한다. 해사 분야 관련 공식 문서, 학문, 미디어 등 영어 문어 텍스트로 구축한 해사 영어 코퍼스를 대상으로 장세은과 변현정(2011)은 코퍼스 분석 프로그램 WordSmith 5.0을 통해 합성명사의 품사별 누적 빈도를 산출하였다. 모든 합성어(합성명사, 합성형용사, 합성동사)를 재분류하는 데 어려움이 있다고 언급한 것처럼(장세은·변현정 2011) 본고 역시 품사 태깅 후의 정제 과정에서 명사이면서 동사인 어휘 등 태깅 오류 및 복잡성으로 인해 단일명사와 합성명사를 중심으로 분석하고자 한다.

둘째, 어떤 용어에 학습 시간을 우선적으로 배분할지 결정해야 한다. 학술 텍스트 분야를 대상으로 출현 빈도수에 따라 고빈도 어휘, 학술적 어휘, 기술적 어휘, 저빈도 어휘로 구분한 것처럼(Nation 2001)처럼 ‘출현 빈도수’에 따라 학습할 용어의 우선순위를 분류할 수 있다. 본고는 AntConc의 워드리스트(Word

List)와 엔그램(N-Gram)을 활용하여 목록화한 단일명사와 합성명사를 연도별 출현 빈도수에 따라 우선적으로 살펴볼 용어를 정하고자 한다. 물론 출현 빈도수는 적더라도 특정 분야에서 중요하거나 또는 의미가 다르게 쓰이는 전문용어도 있기 때문에 전문분야의 관련성에 대한 평가 척도(Chung & Nation 2003)를 적용해 보거나 자주 사용되는 전문용어를 해당 분야의 전문가에게 의뢰하여 필수 용어를 목록화해 달라고 요청할 수도 있다(최문선 2019). 하지만 통역 준비 기간 동안 전문가에게 자문을 요청하는 것이 쉽지 않은 경우도 많기 때문에 용어 분류의 기준 및 범위 설정에서 가장 많이 사용하는 출현 빈도수를 용어의 기본 분류 기준으로 채택하였다.

셋째, 통역용 용어집에 수록되는 전문용어의 범위를 전문용어에만 국한하지 않는다. 물론 전문용어 사전 등에 등재되었거나 특정 분야 전문가만 쓰는 전문용어는 용어집의 주요 구성 대상이다. 그러나 전문용어를 전부 이해했다고 하여 전문분야 회의에 등장한 모든 출발어를 도착어로 완벽하게 전환할 수 없다. 앞서 언급하였듯이 통역용 용어집에는 전문가만 쓰는 전문용어, 직업어만을 대상으로 하지 않는다. 질(Glie 2002; 최문선 2018)은 영-불 의학 회의 통역을 준비하는 과정에서 의학 교과서에 실린 영어 용어를 추출하였지만 실제 통역 시 거의 활용되지 못했고 용어집에 포함되지 않았던 전문용어도 매우 많았다고 보고하였다. 따라서 통역용 용어집은 최대한 해당 통역 상황과 유사하고 실제성 있는 언어 자료를 활용할 필요가 있으며(최문선 2018: 252) 난해한 전문용어만을 대상으로 할 필요가 없다.

본고에서는 학술적 깊이를 가지는 난해한 전문용어와 상반되는 뜻의 ‘일상 생활에서 일반인이 쉽게 접할 수 있는 공공언어’ 역시 용어집의 중요 대상으로 보았다(민재훈·강현화 2012). 전문용어의 사용자를 일반인에게까지 확대하는 경우와 전문가 집단에만 한정하는 경우에 따라 전문용어의 모습을 달라질 수 있는 것처럼(민재훈·강현화 2012) 통역에서도 연사와 청중이 전문가인지, 일반인인지에 따라 용어집의 용어 대상이 달라질 수 있다. 전문분야 회의더라도 일반 청중도 대상으로 하는 회의에서는 전문용어의 출현 빈도수보다는 경제 및 금융 분야에서 일반인이 쉽게 접한 공공언어의 출현 빈도수가 더 많을 것이다. 왜냐하면 통역은 전문용어만을 대상으로 통역하는 것이 아니며 전문용어를 포함한 텍스트에 등장한 모든 용어와 구문을 대상으로 하므로 공공언어의 비중이 전문

1) ‘합성명사’란 둘 이상의 말이 결합된 명사(국립국어원 표준 국어대사전)로서 명사와 명사 또는 형용사, 동사가 명사와 결합되어 하나의 의미 단위를 이루는 명사임. 본고에서는 기본적으로 ‘명사’와 ‘명사’가 결합된 용어를 합성명사로 정의함.

용어의 비중보다 높을 수밖에 없다. 연사가 특정 전문가 집단의 청중을 대상으로 하여 난해한 전문용어만 등장하는 회의 상황을 제외한다면 일반인을 청중으로 하는 전문분야 회의에서는 난해한 전문용어가 등장하더라도 전문용어를 공공언어로 쉽게 표현하거나 풀어서 설명하는 경우가 더 많다. 본고의 경제 및 금융 관련 전문분야 회의 역시 특정 전문가 집단만을 대상으로 한 회의가 아니기 때문에 난해한 전문용어의 출현 빈도수보다 경제, 금융 관련 일반 용어적인 공공언어가 더 많이 등장할 것으로 가정하였으며 실제 용어 목록에서도 경제 및 금융 관련 일반적인 용어가 더 많이 출현하였다.

### 3. 연구 방법

#### 3.1 경제 및 금융 전문회의 연설문 텍스트의 자가코퍼스 구축

자가코퍼스 구축 대상은 중국 상하이(上海) 루지아주이(陸家嘴) 금융 중심지에서 2008년부터 매년 개최한 경제 및 금융 전문분야 회의에서 실제 사용한 연설문 텍스트이다. 용어목록을 추출하기 위해 코퍼스 언어학 입문자뿐 아니라 전문 연구자들도 널리 사용하며 무료로 사용할 수 있는 최고의 콘코덴서로 꼽히는(박명수 2013) AntConc를 사용하였다. 연설문 텍스트는 회의의 공식 홈페이지의 공개 자료를 다운로드한 뒤 AntConc에서 읽을 수 있도록 전처리 작업을 거쳤다.

본고에서 구축한 자가코퍼스의 자격 여부를 판단하기 위해 코퍼스 자격 기준인 코퍼스 규모의 유한성(finite size), 표준 참고로서의 기준(standard reference), 자료의 실제성(authentic texts), 기계 가독성(machine-readable)에 따라 적합성을 확인하였다(박명수 2017). 본고의 자가코퍼스는 경제 및 금융 전문분야 회의에 관한 목적과 표준 참고로서의 기준에 부합하며 회의 참석자들이 실제로 사용한 언어를 가공 없이 전사한 구어 텍스트이며 오프라인 콘코덴서에서 검색이 모두 가능한 형태로서 코퍼스로서의 기본 자격을 갖추었다고 판단하였다. 본고의 자가코퍼스 기본 특징은 특정 시기(2008년~2018년), 특정 장르(경제 및 금융), 특정 유형(구어 코퍼스)의 언어를 관찰하는 것을 목표로 하는 벨

런스 코퍼스이자 특수 코퍼스에 해당한다. 또한 중국어로만 이루어진 단일어 코퍼스이며 중국인 연사의 구어를 모아서 만든 원어민 코퍼스에 해당한다.

#### 3.2 경제 및 금융 전문회의 연설문 텍스트의 용어집 구축

통역 준비를 위한 용어집 구축은 크게 단일어 용어 목록과 이중 용어 목록 과정으로 나뉜다. 단일어 용어 목록은 회의 주제와 관련된 자료를 선별한 뒤 수집한 자료를 검토하면서 출발어 용어를 추출하는 단계이며, 이중 용어 목록은 사전, 전문가 자문, 통역사 개인의 가번역 등을 통해 단일어에 대응하는 도착어를 작성하는 단계에 해당한다(최문선 2018). 환티누올리(Fantinuoli 2017)는 단일어, 이중어 목록 추출 과정을 통역사가 직접 하는 것이 아니라 개발한 용어추출도구를 활용할 것을 제안하였다. 문제는 이중 용어 목록을 구축하는데 있어 텍스트 집합으로부터 특정 분야의 전문용어에 해당하는 항목을 자동으로 신속하고 편리하게 추출해 주는 용어추출도구가 유용하지만 아직 유럽언어 중심의 프로그램이고, 병렬 코퍼스 구축 자체에 시간과 노력이 많이 요구되며(최문선 2018) 본고에서 활용 가능한 한중 병렬 코퍼스, 중국어 경제 및 금융 관련 비교 코퍼스 및 참조 코퍼스가 없는 상황이다. 따라서 본고는 단일어 용어 목록을 추출하기 위해 경제 및 금융 회의의 관련 연설문 텍스트로 구축한 자가코퍼스를 대상으로 오프라인 콘코덴서 AntConc를 통해 용어 목록을 추출하였다. 추출한 단일어 용어 목록을 대상으로 파파고 번역기를 통한 기계번역과 연구자의 가번역을 통해 이중 용어 목록을 구축하였다. 용어집 구축에 관한 기본적인 작업 과정 및 내용은 다음 <표 1>과 같다.

<표 1> 용어집 구축 과정 및 작업 내용

단계	작업명	작업 내용
1단계	텍스트 수집	수집과 전처리를 한 연도별 텍스트를 텍스트 파일(txt) 형태로 저장
2단계	품사 태깅(POS-tagging) <sup>2)</sup>	한자를 품사 단위로 인식하지 못하는 오류 발생을 방지하기 위해 품사 태깅

## 4. 결과 및 논의

### 4.1 경제 및 금융 연설문 텍스트의 자가코퍼스 특징

위의 연구 방법에 따라 2008년~2018년(2014년~2016년 데이터 확보가 어려워 제외함) 총 8년 동안 자가코퍼스에 출현한 평균 어종(types)수와 평균 출현형(tokens)수의 현황은 다음 <표 2>와 같다.

<표 2> 2008년~2018년 전체 및 명사의 평균 어종(types)과 출현형(tokens) 현황

전체 평균 types수(개)		전체 평균 tokens수(개)	
2,630		15,169	
명사types수&비중		명사tokens수&비중	
832	31.6%	6,558	43.2%
경제·금융 단일명사types수&비중		경제·금융 단일명사tokens수&비중	
156	5.9%	2,565	16.9%
경제·금융 합성명사types수&비중		경제·금융 합성명사tokens수&비중	
716	27.2%	2,826	18.6%

자가코퍼스의 평균 어종(types)은 2,630개, 평균 출현형(tokens)은 15,169개였다. 이중 명사의 평균types비중은 약 31.6%, 평균tokens의 비중은 약 43.2%로서 전체 자가코퍼스에서 명사가 30% 이상의 높은 비중을 차지하였다. 명사 중에서도 경제 및 금융 관련 단일명사의 전체 코퍼스 대비 평균types와 tokens의 비중은 각각 5.9%, 16.9%였으며, 경제 및 금융 관련 합성명사의 전체 코퍼스 대비 평균types와 tokens의 비중은 각각 27.2%, 18.6%였다. 특히 경제 및 금융 관련 합성명사types수가 단일명사types수 보다 5배 이상 많았다. 예를 들어 ‘금융’이라는 단일명사로 ‘금융센터, 금융시장, 금융위기’등 다양한 합성명사를 만들 수 있기 때문에 연설문에 경제 및 금융 관련 많은 합성명사types가 등장한 것으로 판단하였다. 아래 <표 3>은 08년~18년 연도별 어종 및 출현형의 상세 현황을 빈도수(개)를 기준으로, <표 4>는 비중(%)을 기준으로 정리하였다.

3단계	AntConc 탑재	일부 수작업 <sup>3)</sup> 을 거친 품사 태깅 자료 AntConc에 탑재
4단계	워드리스트를 활용하여 용어후보군 추출-명사 추출	AntConc 워드리스트에서 명사만 추출
5단계	Stopwords를 활용하여 용어후보군 추출-경제 및 금융 관련 명사추출 <sup>4)</sup>	경제 및 금융 관련 용어 목록을 작성한 Stopwords를 AntConc 내 Tool preference의 Word List Range Options에 Add Word form file로 탑재
6단계	단일명사 용어집 구축	엑셀 파일로 다운로드한 파일을 보면서 명사 이외의 다른 품사 어휘 제거 등 가공 및 가변역 작업
7단계	N-Gram 중 바이그램을 통해 합성명사 용어집 구축	바이그램으로 용어 추출 후 가공 및 가변역 작업

단일어 용어 목록을 구축하기 위해 단일명사는 AntConc의 메뉴 중 코퍼스에 출현한 어휘 목록을 보여주는 워드리스트를 이용하고, 합성명사는 연속된 단어나 철자를 보여주는 엔그램 중 연속된 두 단어나 두 개의 철자를 보여주는 바이그램(bi-grams 또는 2-grams)을 이용하였다. 워드리스트는 명사, 동사, 부사 등 하나의 품사로 이루어진 어휘를 추출할 수 있는 반면 합성명사 목록을 추출하기 위해서는 엔그램의 바이그램으로 추출한 뒤 복잡한 가공 과정이 필요하다. 왜냐하면 품사 태깅을 하면 명사, 동사, 형용사 등 각 품사별로 태깅이 되어 나오기 때문이다. 합성명사 목록을 추출하기 위해서는 ‘명사’+‘명사’로 추가적인 품사 태깅을 거친 뒤 바이그램으로 목록화한 것을 눈으로 보면서 태깅 오류 및 주제와 관련 없는 것 등을 최대한 걸러낸다. 가공 및 정제 작업 후 구축한 중국어 단일명사 및 합성명사 목록을 대상으로 연구 분석을 위해 파파고 번역기를 이용하여 가변역을 한 뒤 어색한 번역 부분을 수정 편집하였다.

- 2) 품사 태깅은 각각의 단어 옆에 품사를 지정해서 보여주는 것임(박명수 2017).
- 3) 중국어 전문 품사 태깅 사이트(语料库在线)를 이용하였지만 품사 태깅 결과 완벽한 분류가 되지 못하는 경우가 많아 육안으로 확인하면서 오류 부분을 최대한 수작업으로 조정함.
- 4) Stopword의 본래 의미는 텍스트에서 의미보다는 기능적인 역할을 하는 관사, 접속사, 전치사 등의 어휘를 지칭하지만(박명수, 2017:232) 본고에서는 경제 및 금융과 관련되지 않은 명사를 Stopword로 구축 후 stoplist를 통해 주제 관련 명사만 추출함.

<표 3> 연도별 전체 및 명사의 어종(types)과 출현형(tokens) 현황(개수 기준: 개)

구분	2008	2009	2010	2011	2012	2013	2017	2018
전체types	2,753	2,096	2,656	1,667	2,924	2,332	3,296	3,312
전체tokens	16,841	10,702	15,107	6,796	19,508	12,874	20,244	19,281
명사types	869	715	653	529	926	753	1,054	1,156
명사tokens	7,665	4,731	4,908	3,459	8,514	5,983	8,233	8,967
경제·금융 단일명사types	167	151	138	124	159	143	220	144
경제·금융 단일명사tokens	2,897	2,305	1,671	1,167	3,179	2,639	3,594	3,064
경제·금융 합성명사types	818	614	634	451	911	905	987	406
경제·금융 합성명사tokens	3,524	1,310	2,198	1,564	3,856	3,206	4,154	2,792

<표 4> 연도별 전체 코퍼스 대비 명사의 어종(types)과 출현형(tokens) 현황(비중 기준: %)

구분	2008	2009	2010	2011	2012	2013	2017	2018
명사types	31.6	34.1	24.6	31.7	31.7	32.3	32.0	34.9
명사tokens	45.5	44.2	32.5	50.9	43.6	46.5	40.7	46.5
경제·금융 단일명사types	6.1	7.2	5.2	7.4	5.4	6.1	6.7	4.3
경제·금융 합성명사types	29.7	29.3	23.9	27.1	31.2	38.8	29.9	12.3
경제·금융 단일명사tokens	17.2	21.5	11.1	17.2	16.3	20.5	17.8	15.9
경제·금융 합성명사tokens	20.9	12.2	14.5	23.0	19.8	24.9	20.5	14.5

<표 2>와 마찬가지로 연도별 합성명사types의 비중이 단일명사types의 비중보다 보통 3~6배로 많다는 것을 확인할 수 있다. 용어 목록을 추출하기 위해서는 단일명사만이 아니라 합성명사를 추출하는 것도 중요하며 합성명사 어종이 매우 다양하다는 것을 알 수 있다. 비록 단일명사를 추출하는 작업에 비해 합성명사를 추출하는 작업이 복잡하지만 합성명사의 높은 비중을 고려해보면 코퍼스 분석 도구를 활용한 합성명사 용어집 구축의 중요성과 유용성을 보여준다.

#### 4.2 경제 및 금융 관련 '단일명사' 용어 목록 특징

워드리스트와 Stopwords로 경제 및 금융 관련 용어만을 추출한 단일명사의 연도별 출현 빈도수 현황은 아래 <표 5>와 같다.

<표 5> 단일명사 연도별 출현 빈도수 현황(상위 20위 목록)

No	한국어	중국어	08년	09년	10년	11년	12년	13년	17년	08년~17년
1	금융	金融	586	448	224	214	569	172	747	2,960
2	시장	市场	243	187	149	77	136	194	223	1,209
3	경제	经济	158	151	121	101	191	115	209	1,046
4	리스크	风险	159	126	108	56	148	49	179	825
5	은행	银行	92	62	112	41	233	135	104	779
6	위안화	人民币	13	5	1	11	254	295	34	613
7	보험	保险	77	79	54	49	57	51	130	497
8	기업	企业	74	54	18	22	51	34	121	374
9	제품	产品	81	48	10	21	37	59	83	339
10	자본	资本	37	27	45	10	77	56	34	286
11	업무	业务	18	62	17	9	66	61	31	264
12	자산	资产	62	33	19	14	16	79	27	250
13	거래	交易	30	11	36	10	39	37	85	248
14	회사	公司	35	63	35	14	22	18	50	237
15	재산(부)	财富	2	1	1	-	13	202	2	221
16	화폐(통화)	货币	41	9	16	12	55	52	34	219
17	업종	行业	14	33	21	11	29	17	61	186
18	채권	债券	69	10	10	5	20	25	36	175
19	인재	人才	69	19	10	3	34	19	6	160
20	자금, 자본금	资金	18	22	3	9	19	34	55	160

2008년~2017년 총합의 빈도수에 따라 상위 20권에 해당하는 용어를 보면 금융, 시장, 경제, 리스크, 은행, 위안화, 보험 등의 순으로 높게 나왔다. 하지만 음영 처리한 부분의 위안화와 재산(부)의 빈도수는 상위권이더라도 한두 해 특정 연도에만 100회 이상 고빈도로 출현한 것을 알 수 있다. 이는 회의 주제, 국내의 경제 및 금융 정책과 환경 등에 따라 특정 용어의 출현 빈도수가 달라졌을 것으로 판단하였다. 이처럼 용어마다 연도별 출현 양상이 다양한 가운데 본 고에서는 몇 가지 기준에 따라 용어 분류 작업을 진행하였다. 연도별 용어 출현 양상을 확인하면서 우선적으로 학습할 용어 등 용어의 중요도를 부여하기 위해 다음 <표 6>과 <표 7>과 같이 경제 및 금융 관련 단일명사를 '연도별 출현 여부'와 '연도별 출현 빈도수'에 따라 2차 분류를 하였다.

〈표 6〉 연도별(2008~2017년) '출현 여부'에 따른 용어 분류 기준

분류	기준 설명	중요도
상	6~7년 지속적으로 출현한 용어	필수
중	3~5년 간헐적으로 출현한 용어	필수
하	1~2년 특정 한두 해만 출현한 용어	선택

과거 7년간의 연설문에 몇 년 동안 지속적으로 출현했는지, 즉 연도별 지속적인 출현 여부에 따라 '상중하'로 분류하였다. 예를 들어 6~7년 동안 해마다 빠지지 않고 등장한 '상' 용어는 경제 및 금융 관련 회의에서 어떠한 연설문 텍스트가 나오지 모르더라도 비슷한 유형의 경제 및 금융 통역을 준비하기 위해 반드시 준비해야 할 필수 용어라고 간접적으로 판단하였다. 반면, 매년 등장하지 않았지만 기본적으로 3~5년 정도 지속적으로 출현한 '중'에 해당하는 필수 용어 역시 '상'에 해당하는 필수 용어와 함께 살펴봐야 한다고 판단하였다. 마지막으로 '하'에 해당하는 용어는 한두 해 특정 연도에만 출현하였기 때문에 '상'과 '중'에 비해 중요도가 다소 떨어져 선택 용어로 분류하였다. 통역 준비 시 '상'과 '중' 필수 용어를 우선적으로 학습하되 시간적 여유가 있다면 '하'에 해당하는 용어도 살펴볼 수 있다. 하지만 용어의 연도별 출현 여부만을 기준으로 용어의 중요도를 파악하기에는 부족하므로 다음 <표 7>과 같이 해마다 실제로 텍스트에 얼마나 출현했는지 '출현 빈도수'에 따른 분류 기준을 마련하였다. 본 고에서는 출현한 용어의 전체 빈도수를 보면서 '연도별 10회 미만, 10회 이상~100회 미만, 100회 이상 출현한 용어' 세 가지 임의적인 기준을 정하여 분류 작업을 하였다.

〈표 7〉 연도별 '출현 빈도수'에 따른 용어 분류 기준

분류	기준 설명	중요도
고빈도	100회 이상으로 출현한 용어	필수
중빈도	10회 이상~100회 미만으로 출현한 용어	필수
저빈도	10회 미만으로 출현한 용어	선택

위의 <표 6>과 <표 7>에서 제시한 분류 기준을 결합하여 아래 <표 8>과 같은 용어 분류 기준을 마련하였다. 즉, 자가코퍼스에 몇 번 출현하였는지 빈도

수에 따른 '고빈도, 중빈도, 저빈도' 기준과 연도별 지속적인 출현 여부를 보여주는 '상, 중, 하'를 결합하여 다음과 같이 분류 기준을 마련하였다.

〈표 8〉 연도별 '출현 여부'와 연도별 '출현 빈도수'에 따른 용어 분류 기준

분류	기준 설명	중요도
(상)고빈도	6~7년 100회 이상 출현한 용어	필수
(상)중빈도	6~7년 10회 이상~100회 미만 출현한 용어	필수
(상)저빈도	6~7년 10회 미만 출현한 용어	필수
(중)고빈도	3~5년 100회 이상 출현한 용어	필수
(중)중빈도	3~5년 10회 이상~100회 미만 출현한 용어	필수
(중)저빈도	3~5년 10회 미만 출현한 용어	필수
(하)고빈도	1~2년 100회 이상 출현한 용어	필수
(하)중빈도	1~2년 10회 이상~100회 미만 출현한 용어	선택
(하)저빈도	1~2년 10회 미만 출현한 용어	선택

<표8>의 중요도 기준의 경우 3년 이상 지속적으로 출현함과 동시에 10회 이상 출현한 용어를 필수 용어, 1~2년에만 100회 미만으로 출현한 용어는 선택 용어로 분류하였다. 다만 1~2년만 등장하였더라도 향후 많이 쓰일 수도 있으며 반대로 과거 7년 동안 해마다 출현하였지만 앞으로 아예 출현하지 않거나 많이 쓰이지 않을 수 있는 등 과거 추이만으로 미래 역시 동일한 양상을 보일 거라고 확신할 수 없다는 점을 밝힌다.

위의 기준에 따라 분류한 단일명사의 출현 현황은 다음 <표 9>와 같으며 경제 및 금융 관련 단일명사 중 상위 목록은 <표 10>과 같다. 비록 연도별 출현 여부와 출현 빈도수에 따라 분류한 필수 용어와 선택 용어 중에서 필수 용어는 반드시 살펴봐야 하지만 아래 <표 10>의 목록 중 선택 용어 역시 주목해야 할 용어임을 알 수 있다. 선택 용어는 필수 용어처럼 다년간 출현했거나 고빈도로 출현한 용어가 아니지만 연설문에 등장한 전문용어이다. 만약 필수 용어를 이미 잘 숙지한 전문통역사라면 용어 준비 과정에서 오히려 (하)중·저빈도에 해당하는 선택 용어를 중심으로 준비하는 것도 전략적인 통역 준비의 한 방법이라고 생각한다.



<표 9> 단일명사 출현 현황(개)

빈도수	연도별 출현 현황(과거 7년)			
	상(6~7년)	중(3~5년)	하(1~2년)	합계
고빈도(100회 이상)	(상)고빈도 6	(중)고빈도 0	(하)고빈도 0	6
중빈도(10회 이상~100회 미만)	(상)중빈도 41	(중)중빈도 0	(하)중빈도 2	43
저빈도(10회 미만)	(상)저빈도 60	(중)저빈도 84	(하)저빈도 24	168
합계	107	84	26	217

<표 10> 단일명사 출현 목록 예시 (상위 목록)

분류	용어 목록(상위권)	중요도
(상)고빈도	금융, 시장, 경제, 리스크, 은행, 위안화	필수
(상)중빈도	보험, 기업, 제품, 자본, 업무(사업), 보험업, 자본시장, 자산, 거래, 금융시장 등	필수
(상)저빈도	재산(부), 원가, 선물, 재정, 외환, 채무, 예금, 레버리지, 신용, 외자(외국자본) 등	필수
(중)고빈도	3~5년 100회 이상 출현한 용어 없음	필수
(중)중빈도	3~5년 간 10회~100회 미만 출현한 용어 없음	필수
(중)저빈도	효익, 권익, 법인, 현물, 어음, 수출입	필수
(하)고빈도	1~2년 100회 이상 출현한 용어 없음	필수
(하)중빈도	지폐	선택
(하)저빈도	보험증권, 샤오강 사회 <sup>5)</sup> , 확장성 등	선택

사실 용어 분류 작업 과정에서 용어마다 연도별 빈도수 증감에 따른 등락 변화가 심하기 때문에 이를 일괄적으로 분류하는 게 쉽지 않다. 예를 들어 어떤 해에 100회 이상 출현했던 용어가 다른 해에는 10회 미만으로 출현하는 등 연도별 용어의 출현 현황이 다르기 때문에 다양한 변수가 존재하기 때문이다. 따라서 본고에서는 7년 중 4년 이상 지속적으로 출현하고 출현 빈도수가 10회 이상인 용어를 ‘고빈도’로 분류하고, 3년 이하 출현하되 출현 빈도수가 10회 미만인 용어는 ‘중빈도’ 등으로 분류하였다. 분류한 용어의 연도별 출현 양상을

5) 샤오강(小康)사회는 장쩌민 국가주석이 2002년 16차 당 대회에서 “2020년까지 전면적인 샤오강사회를 달성하겠다”고 말한 이후 중국 발전의 상징어로 자리잡은 용어로 모든 국민이 풍족하고 편안한 생활을 누리는 사회를 일컬음(네이버 지식백과)

눈으로 보면서 중요도(선택, 필수)를 판단하였다.

### 4.3 경제 및 금융 관련 ‘합성명사’ 용어 목록 특징

합성명사를 목록화하기 위해 품사 태깅 이후 가공 정제 작업을 해야 하는데 문제는 어휘 한 개 당 품사가 태깅되므로 합성명사를 바로 추출할 수 없다는 점이다. 예를 들어 ‘사과나무’의 경우 ‘사과\_n(명사) 나무\_n(명사)’로 품사 태깅 되며, ‘사과나무’ 같은 합성명사 상태로 추출할 수 없는 경우가 많아 별도의 가공 및 정제 작업이 필요하다. 본고에서는 합성명사를 추출하기 위해 바이그램으로 연속된 두 명사(‘명사+명사’)를 추출한 뒤 stopwords로 stoplist를 적용하여 경제 및 금융 관련 용어후보군을 추출하여 가공 작업을 하였다.

<표 11> 합성명사 연도별 출현 빈도수 현황 (상위 20위 목록)

No	한국어	중국어	08년	09년	10년	11년	12년	13년	17년	08년~17년
1	금융센터	金融中心	258	69	24	26	216	62	190	845
2	글로벌금융	国际金融	196	68	66	36	202	46	144	758
3	금융기관	金融机构	128	63	80	76	90	28	132	597
4	금융위기	金融危机	16	78	114	24	90	20	36	378
5	금융시스템	金融体系	78	21	22	38	70	16	42	287
6	실물경제	实体经济	12	9	16	6	110	32	102	287
7	자산관리	财富管理	-	-	-	-	10	258	-	268
8	보험업	保险业	26	74	61	33	19	12	37	262
9	자본시장	资本市场	60	59	35	12	31	16	25	239
10	금융시장	金融市场	78	20	22	14	57	23	38	171
11	채권시장	债券市场	90	3	12	4	10	14	166	166
12	일대일로	一带一路	-	-	-	-	-	-	20	165
13	시중은행	商业银行	10	3	34	20	38	40	8	141
14	경제체	经济体	14	5	14	2	50	48	22	137
15	통화정책	货币政策	50	5	18	18	6	18	62	136
16	금융리스크	金融风险	34	6	12	14	8	-	10	119
17	경제사회	经济社会	32	9	30	26	12	-	18	114
18	보험시장	保险市场	14	12	32	18	14	6	24	108
19	글로벌 경제	全球经济	4	20	8	8	36	8	36	99
20	인프라시설	基础设施	24	5	2	2	16	14	2	97

위의 <표 11>은 08년~17년 빈도수 총합에 따라 연도별 상위 20권의 합성명사 목록을 보여준다. ‘금융센터, 글로벌 금융, 금융기관’ 등 용어는 중, 고빈도로 해마다 출현한 반면에 ‘자산 관리, 일대일로<sup>6)</sup>’ 등 일부 용어는 특정 연도에만 등장하였다. 특히 1~2년 특정 연도에만 출현하였지만 100회 이상 고빈도

로 출현한 용어 중 ‘일대일로(一帶一路)’는 2017년에 처음 등장한 중국 경제 정책의 신조어이다. 이처럼 연도별 용어 출현 양상을 통해 신조어나 특정 연도에 자주 쓰인 용어 등도 파악할 수 있다.

아래 <표 12>는 합성명사의 출현 현황에 관한 분석 결과이다. 1~2년만 저빈도로 등장한 용어가 많았지만 ‘저, 중빈도’에 해당하더라도 7년 동안 꾸준히 등장한 용어는 약 200개로 다른 경제 및 금융 관련 회의를 준비할 때 우선적으로 봐야 할 용어라고 보았다.

<표 12> 합성명사 출현 현황

빈도수	연도별 출현 횟수(7년:08년~17년)			
	상(6~7년)	중(3~5년)	하(1~2년)	합계
고빈도 (100회 이상)	(상)고빈도 0	(중)고빈도 0	(하)고빈도 2	2
중빈도 (10회 이상~100회 미만)	(상)중빈도 20	(중)중빈도 6	(하)중빈도 51	77
저빈도 (10회 미만)	(상)저빈도 14	(중)저빈도 186	(하)저빈도 3,252	3,452
합계	34	192	3,305	3,531

<표 12>에 해당하는 경제 및 금융 관련 합성명사의 분류별 용어 목록은 다음 <표 13>과 같다.

<표 13> 합성명사 출현 목록(상위권) 현황

태그	용어 목록(상위권)	중요도
(상)고빈도	6~7년 100회 이상으로 출현한 용어 없음	필수
(상)중빈도	금융센터, 글로벌 금융, 금융기관, 금융위기 등	필수
(상)저빈도	글로벌통화, 금융시장시스템, 시장메커니즘 등	필수
(중)고빈도	3~5년 100회 이상으로 출현한 용어 없음	필수
(중)중빈도	금융인재, 금융자산, 금융도시 등	필수
(중)저빈도	자본항목, 기관투자자, 금융자원, 재정정책 등	필수
(하)고빈도	자산관리, 일대일로	필수
(하)중빈도	자유무역지대, 역외 위안화, 금융요소 등	선택
(하)저빈도	책임보험, 보험시장시스템, 역외시장 등	선택

6) 일대일로(一帶一路, One belt, One road)는 내륙과 해상의 실크로드 경제벨트를 구축하려는 중국 주도의 신 실크로드 전략이 담긴 중국의 중요한 경제 정책 임.

<표 13>의 상위권에 속하는 필수 용어를 보면 경제 뉴스에서 흔히 등장하는 공공언어적인 용어도 많았다. 한편 (하)저빈도, (하)중빈도 등 연도별 출현 여부 횟수와 출현 빈도수는 적지만 ‘역외 위안화, 역외 시장’ 등 경제 및 금융 관련 전문용어가 많이 등장하였다. 앞서 단일명사 분석 결과와 마찬가지로 빈도수가 많고 연도별 자주 출현하는 필수 용어를 우선적으로 봐야하지만 필수 용어의 경우 전문통역사로서 이미 알고 있는 공공용어 속성의 용어도 많이 포함되었다. 따라서 전문분야 통역에서 저빈도에 해당하는 선택 용어가 오히려 통역에 어려움을 야기할 수 있기 때문에 통역 준비 시 선택 용어를 우선적으로 학습 및 준비하는 것도 고려해 볼 수 있다.

#### 4.4 단일명사 용어집과 특정 연도 단일명사 용어 간 출현 일치도

해마다 용어의 출현 양상이 다양한 가운데 용어집에서 나온 용어가 2018년에도 동일하게 출현했는지 즉, 용어 간 출현 일치도(적중률)를 확인하기 위해 다음과 같은 통역 상황을 가정해 보았다. 이러한 의문을 갖게 된 배경은 전문분야 통역을 처음 맡는 통역사보다 비슷한 분야이거나 해마다 동일한 회의를 맡아온 통역사는 해당 분야에 자주 등장하는 용어에 어느 정도 익숙해져 있는 상황이므로 처음 해당 전문분야의 통역을 준비할 때 문맥 파악이 좀 더 용이할 것이라고 생각했다. 본 연구에서는 특정 전문 분야에 오랜 기간 동안 통역을 맡거나 통역 경험이 많은 전문 통역사가 그렇지 않은 통역사에 비해 전문용어와 문맥 파악 측면에서 유리한 이유 중의 하나로써 용어 간 출현 일치도로 확인하였다. 예를 들어 2009년~2017년까지 회의를 맡아온 통역사가 2018년에도 동일하게 회의를 의뢰 받은 경우 2018년 회의 자료를 사전에 제공받지 않은 상황에서 할 수 있는 통역 준비 사항으로는 2018년도 회의 주제 관련 내용을 인터넷 등을 통해 용어 지식과 주제 지식을 쌓고, 과거 회의 자료 및 용어집을 참고하면서 준비할 것이다. 과거 연설문 텍스트를 기반으로 준비한 용어집이 2018년 회의 준비에 얼마나 유용할지의 여부를 간접적으로 판단하기 위해 다음 <표 14>와 같이 분류 기준을 마련하여 용어집과 2018년 용어 목록을 비교 분석하였다.

<표 14> 용어집과 특정 연도 용어 간 일치도에 따른 분류

분류(1차-2차)	기준 설명	중요도
일치-고빈도	7년 총합 100회 이상 출현함과 동시에 2018년에도 100회 이상 출현	필수
일치-중빈도	7년 총합 10회~100회 미만 출현함과 동시에 2018년에도 10회 이상~100회 미만 출현	필수
일치-저빈도	7년 총합 10회 미만 출현함과 동시에 2018년에도 10회 미만	필수
불일치-고빈도	7년 총합 100회 이상 출현하였으나 2018년에는 출현하지 않음	필수
불일치-중빈도	7년 총합 10회~100회 미만 출현하였으나 2018년에는 출현하지 않음	필수
불일치-저빈도	7년 총합 10회 미만 출현하였으나 2018년에는 출현하지 않음	선택

1차 분류 기준으로 사전에 구축한 용어집에 출현했던 용어가 2018년도에도 출현했는지에 따라 ‘일치’ 또는 ‘불일치’로 구분하였다. 기존 용어집의 용어가 2018년도도 동일하게 출현했다는 것은 빈도수에 상관없이 해당 분야 통역 준비를 위해 봐야 할 필수 용어로 판단하여 중요도에 ‘필수’로 분류하였다. 분류 기준에 따라 분석한 결과 7년 동안 단일명사 목록 217개 중 2018년도도 동일하게 출현한 용어는 141개였으며, 2018년에는 나타나지 않은 용어는 76개였다. 즉, 기존 단일명사 용어집의 약 65%는 2018년도도 준비한 용어집에서 동일하게 출현하였으며 약 35% 용어는 2018년에는 출현하지 않았다.

1차 분류 기준 다음으로 2차 분류 기준으로써 용어집의 용어 출현 빈도수 총합(100회 이상, 10회~100회 미만, 10회 미만)과 2018년 용어 출현 빈도수(100회 이상, 10회~100회 미만, 10회 미만)에 따라 ‘고빈도, 중빈도, 저빈도’로 구분하였다. 예를 들어 ‘일치-고빈도’는 7년간 출현함과 동시에 18년도도 동일하게 출현하였으며 기존 용어집과 18년도 용어집에서도 100회 이상 출현한 용어를 의미한다.

<표 15> 단일명사 용어집과 특정 연도 용어집 간 비교한 용어 목록(상위권)

분류(1차-2차)	기준 설명	중요도
일치-고빈도	금융, 시장, 경제, 리스크 등	필수
일치-중빈도	산업, 달러, 거시적, 원가(비용), 신용, 선물 등	필수
일치-저빈도	중국 자본, 신용카드, 주기성, 수익률 등	필수
불일치-고빈도	18년도에는 출현하지 않았으며 용어집 출현 빈도수 100회 이상 분류에 해당하는 용어 없음	필수
불일치-중빈도	지폐	필수
불일치-저빈도	업계, 기업가, 투자자, 자본금, 공급과 수요 등	선택

<표 15> 상세 용어 목록 중 ‘금융, 시장, 경제’ 등이 ‘일치-고빈도’에 해당하였다. 또 다른 예로 ‘불일치-저빈도’ 용어의 경우 18년에는 출현하지 않았으며 기존 용어집에서 저빈도에 해당하는 용어로서 ‘기업가, 투자자, 자본금’ 등의 용어는 18년에는 출현하지 않았다.

#### 4.5 합성명사 용어집과 특정 연도 합성명사 용어와의 일치도

마찬가지로 합성명사 역시 1차 작업으로 7년 동안 출현했던 용어가 2018년도에도 출현했는지에 따라 ‘일치’ 또는 ‘불일치’로 구분하였다. 전체 합성명사 용어 약 3,500개 중 2018년도도 동일하게 출현한 합성명사 용어는 394개, 불일치한 용어 목록은 3,136개였다. 즉, 준비한 용어집에서 11.3%는 2018년도도 동일하게 출현하였지만 89.6%는 용어집에서 나오지 않았다. 단일명사에 비해 합성명사의 일치한 비율이 낮게 나온 것은 2018년에는 기존 용어집과 다른 다양한 합성명사가 등장했다는 것을 보여준다. 2차 기준에 따라 출현한 빈도수(100회 이상, 10회~100회 미만, 10회 미만)에 따라 분류한 결과는 다음 <표 16>과 같다.

<표 16> 합성명사 용어집과 특정 연도 용어집 간 용어 비교 목록(상위권)

분류(1차-2차)	기준 설명	중요도
일치-고빈도	금융센터, 글로벌금융, 금융기관, 일대일로	필수
일치-중빈도	인프라 시설, 글로벌 경제, 자본 항목 등	필수
일치-저빈도	신용 시스템, 리스크 요인, 주식시장 등	필수
불일치-고빈도	부의 관리, 보험시장 등	필수
불일치-중빈도	유로존, 거시경제 등	필수
불일치-저빈도	법인기구, 가치사슬, 채권잔액 등	선택

<표 16>에서 ‘일치-고빈도’에 해당하는 ‘금융센터, 금융기관’ 등은 합성명사 용어집에서는 물론 18년에도 고빈도로 등장한 합성명사였다. ‘불일치-중빈도’에 해당하는 ‘유로존, 거시경제’는 18년에 등장하지 않았으나 기존 용어집에는 중빈도로 등장했던 용어였다. 한편 기존 용어집에는 출현하지 않았으나 2018년도에 처음 등장한 ‘블록체인’과 같은 신조어도 있었다. 블록체인과 같은 용어는 2019년, 2020년에 주요 용어인 만큼 과거 추이만 보고 미래의 추이를 속단할 수 없지만 경제 및 금융 환경 변화에 따라 과거에 자주 사용하던 용어가 어느 해부터는 덜 사용되거나 사라질 수 있으며 과거에 자주 사용하지 않던 용어가 신조어로 등장하거나 자주 사용될 수 있음을 보여준다. 과거의 추이를 통해 필수 또는 선택 용어를 판단하는 데 기초적인 근거를 제시하고자 한 본 연구를 통해 관련 분야의 환경 변화에 따라 용어 역시 생성과 소멸, 사용 등이 달라지므로 통역사는 해당 분야에서 이슈가 되는 용어에 항상 많은 관심을 기울여야 한다.

### 5. 결론

본고는 중한 경제 및 금융 관련 전문 분야 회의 준비에서 활용할 수 있는 경제 및 금융 분야 용어집의 구축을 시도한 예비연구이다. 용어집 구축을 위해 중국의 한 경제 및 금융 전문회의(2008~2018년)에서 사용한 연설문 텍스트를 자가코퍼스로 구축한 뒤 AntConc 코퍼스 분석 도구라는 기술적 도구를 활용하였다. 본고의 용어집은 자가코퍼스에서 30% 이상을 차지한 명사를 중심으로 단일명사와 합성명사로 구분하여 연도별 출현 빈도수를 조사하였다. 경제 및 금융 회의인만큼 전체 명사에서 경제 및 금융 관련 단일명사와 합성명사가 높은 비중을 차지하였다. 또한 연도별 경제 및 금융 관련 명사의 출현 양상을 비교 분석한 결과 해마다 중빈도, 고빈도로 출현하는 용어 또는 특정 해에만 출현한 용어 등 연도별로 용어의 출현 양상이 매우 다양하였다. 단일명사와 합성명사 목록의 각 용어에 중요도를 부여하기 위해 연도별 출현 빈도수와 출현 여부라는 기준을 마련하여 살펴본 결과 연도별 공통적으로 중, 고빈도로 출현한 용어는 다른 경제 및 금융 회의를 준비할 때에도 참고할 수 있는 기본 필수 용어일 가능성이 높은 용어라고 판단하였다. 기본 필수 용어 목록에는 경제 및 금융

관련 전문용어들도 있지만 전문가뿐 아니라 일반 대중들도 상식적으로 아는 공공언어에 속하는 경제 및 금융 관련 용어도 많았다. 따라서 전문통역사는 1~2년 정도 저빈도로 출현한 선택 용어 중에서 어려운 용어를 중심으로 준비하는 것도 용어집을 활용한 통역 준비의 한 방법이라고 본다.

나아가 구축한 용어집의 효용성을 간접적으로 확인하기 위해 동일한 회의의 2018년도 회의를 준비한다는 가정 하에 용어집에서 과연 2018년도의 연설문에 얼마나 많은 용어가 동일하게 출현하였는지 확인해보았다. 단일명사는 준비한 용어집에서 약 65% 동일하게 출현한 반면 합성명사는 약 11.3%로 단일명사에 비해 낮게 나왔다. 비록 합성명사에서 일치한 비율이 단일명사 용어에 비해 낮았지만 전체 텍스트에서 합성명사가 차지하는 비중이 크고, 합성명사types가 일반명사types보다 4~5배 많다는 점을 감안하여 텍스트에서 다양하게 출현하고 있는 합성명사의 용어 추출 역시 중요하다. 이처럼 단일명사와 합성명사 용어별로 연간 출현 양상이 다양하기 때문에 용어 준비 과정에서 다양한 용어에 대한 학습과 준비 전략이 필요하다는 것을 알 수 있었다.

하지만 본고의 데이터 분석 과정 및 품사 태깅에서 수작업을 통해 최대한 분류한다고 하였지만 정제 과정에서 완벽히 분류되지 못한 부분이 존재하며 용어의 분류 기준과 중요도에 대한 판단 역시 본고에서 임의적으로 정한 점 등이 본 연구의 한계점이기도 하다. 후속 연구에서는 이러한 한계점을 보완하고, 본고에서 구축한 용어 목록이 다른 경제 및 금융 회의에서 얼마나 일치하게 나타나는지, 명사명사구와 자주 공기되는 서술어 등 언어 정보에 대한 추가적인 연구 등을 진행할 수 있을 것이다. 아직까지 국내 중한 전문분야 통역을 연구 대상으로 코퍼스 기반 통역 준비(CDIP) 개념을 활용하여 통역 용어집을 구축한 사례 연구가 없음을 감안하였을 때 본 연구가 코퍼스 기반 중한 통역 준비 용어집 구축 및 활용의 시작이 되길 바란다.

### 참고문헌

강현화 (1998) 「전문용어 사전에 대한 국어학적 제안: 무역 경제 용어를 중심으로」, 『연세대학교 언어정보연구원』 98(0): 309-327.

남슬기, 성미경, 손은은, 신나정, 최수연, 한성숙, 정혜연 (2018) 통역사가 선택한 핵심정보의 형태적 특징, 『통역과 번역』 20(2): 73-95.

민재훈, 강현화 (2012) 「기업 및 시장경제 용어의 사용 양상 분석: 조어 분석과 코퍼스의 공기 분석을 바탕으로」, 『언어사실과 관점』, 30(0): 5-34.

박명수 (2017) 『코퍼스 언어학의 이해와 실제』, 서울: 한빛문화.

박현일 (2019) 청중의 통역품질 평가에 대한 고찰, 『통번역학연구』 23(1): 45-64.

원종화 (2014) 영한 통역 정확도 평가를 위한 아이디어 단위 연구, 『번역학연구』 15(2): 123-148.

이미경 (2018) 전문용어의 사용이 통역의 과정에 미치는 영향에 대하여, 『통번역교육연구』 16(3): 127-146.

이지은, 최효은 (2013) 코퍼스기반 통번역 학습자용 언어 사전 구축 방안: 주 식용어를 중심으로 한 한국어 및 한국어-영어 언어 사전 구축에 관한 예비조사, 『통번역학연구』, 17(2): 117-147.

장세은, 변현정(2011) 코퍼스를 활용한 해사영어 어휘 분석, 『새한영어영문학』, 53(4): 247-268.

최문선(2018) 전문분야 통역을 위한 용어집 구축 -기술 도구의 적용가능성 고찰, 『통역과 번역』, 20(2): 163-191.

최문선(2019) 「전문분야 통역을 위한 어휘지식: IT분야를 중심으로」, 『통역과 번역』, 20(3): 247-283.

최정화 (2004) 「국제회의통역 사전 준비방법에 대한 고찰」, 『국제회의 통역과 번역』 6(2): 187-206.

Chung, Teresa M. and Paul Nation (2003) 'Technical Vocabulary in Specialised Texts', *Reading in a Foreign Language* 15(2): 103-116.

Fantinuoli, Claudio(2017) 'Computer-assisted preparation in conference interpreting', *Translation & Interpreting* 9(2): 24-37.

Gile, Daniel (2002) 'The Interpreter's Preparation for Technical Conferences: Methodological Questions in Investigating the Topic', *Conference Interpretation and Translation* 4(2): 7-27.

Gorjanc, Vojko (2009) 'Terminology resources and terminological data management for medical interpreters' in D. Andres & S. Pöllabauer

(eds), *Spürst Du, wie der Bauch rauf-runter? Fachdolmetschen im Gesundheitsbereich. Is Everything all Topsy Turvy in Your Tummy? Healthcare Interpreting*, München: Martin Meidenbauer. 85-95.

Nation, I.S.P. (2001) *Learning vocabulary in another language*. Cambridge, England: Cambridge University Press 6-21.

Todd, Richard Watson (2017) 'An opaque engineering word list: Which words should a teacher focus on?', *English for Specific Purposes* 45: 31-39.

<부록 1> 경제 및 금융 연설문(08년~18년)에 출현한 고빈도 용어 목록

\* 단일명사(좌)와 합성명사(우) (각 50개)

No	한국어	중국어	No	한국어	중국어
1	금융	金融	1	금융센터	金融中心
2	시장	市场	2	금융기관	金融机构
3	경제	经济	3	금융위기	金融危机
4	리스크	风险	4	금융시스템	金融体系
5	은행	银行	5	자산관리	财富管理
6	위안화	人民币	6	실물경제	实体经济
7	보험	保险	7	시중은행	商业银行
8	기업	企业	8	보험업	保险业
9	제품	产品	9	자본시장	资本市场
10	자본	资本	10	경제체	经济体
11	업무	业务	11	채권시장	债券市场
12	자산	资产	12	경제 사회	经济社会
13	거래	交易	13	통화정책	货币政策
14	회사	公司	14	보험시장	保险市场
15	재산(부)	财富	15	금융인재	金融人才
16	화폐(통화)	货币	16	금융시장	金融市场
17	업종	行业	17	금융리스크	金融风险
18	채권	债券	18	금융상품	金融产品
19	인재	人才	19	경제금융	经济金融
20	자금	资金	20	금융자산	金融资产
21	실물	实体	21	자본항목	资本项目
22	수요	需求	22	인프라 시설	基础设施
23	자원	资源	23	경제구조	经济结构
24	증권	证券	24	중앙은행	中央银行
25	거래처	客户	25	글로벌 통화	国际货币
26	투자자	投资者	26	재정정책	财政政策
27	유동성	流动性	27	기관투자자	机构投资者
28	거래소	交易所	28	거시경제	宏观经济
29	비즈니스	商业	29	채무위기	债务危机
30	금리	利率	30	은행업	银行业
31	가격	价格	31	금융업	金融业
32	저축	储蓄	32	유로존	欧元区
33	산업	产业	33	자산증권	资产证券
34	사업, 프로젝트	项目	34	금융업	金融行业
35	달러	美元	35	금융시장시스템	金融市场体系
36	거시적	宏观	36	글로벌 금융시장	全球金融市场
37	원가	成本	37	국가 채무	主权债务
38	가격 책정, 정가	定价	38	농촌금융	农村金融
39	신용	信用	39	금융도시	金融城
40	선물	期货	40	체제 메커니즘	体制机制
41	펀드	基金	41	자유무역지대	自贸区
42	경쟁력, 우위	优势	42	글로벌은행	世界银行
43	재정	财政	43	은행 시스템	银行体系
44	외환	外汇	44	글로벌경제	国际经济
45	채무	债务	45	위안화시장	人民币市场
46	예금	存款	46	역외 위안화	境外人民币
47	레버리지	杠杆	47	시장주체	市场主体
48	신용	信贷	48	시장 메커니즘	市场机制
49	외자	外资	49	보험회사	保险企业
50	이익	利益	50	자유무역지대	自贸区

[Abstract]

**A Study on the Establishment of Corpus-Based terminology list for the Preparation of Specialized Interpretations: Focused on Economic and Financial domains in the Speech**

Lee, Ji-hye and Lee, Yong-hun

(Ewha Womans University, Chungnam National University)

For interpretation in the economic and financial domains, this study attempts to extract single nouns and compound nouns based on the study corpus was created with transcripts of the economic and financial conference. This paper attempts to examine the patterns of terminology over the seven years and select basic technical words according to the frequency and the annual usage patterns.

The results show that the single noun and the compound nouns are very diverse every year. The words that appear every year are grouped into essential vocabulary that must be studied. Even if it doesn't appear every year, the words that appear at high frequencies are defined as 'essential vocabulary'. In addition, the terminology contains not only technical words, but also many common economic and financial words. The result of confirming the terminology hit rate for a given year based on the construction of the terminology list shows that the hit rate for a single noun is 65%, compared to 11% for a compound noun. However, considering that compound nouns account for a very high proportion of the entire text, compound nouns are also a factor that must be included in terminology preparation. I hope the limitations of this study will be further supplemented so that we can build and utilize a corpus-based terminology for interpreters' preparation in the economic and financial domains.

- ▶ Keywords: terminology list, word List analysis, specialized interpretation, transcripts, corpus analysis tool
- ▶ 주제어: 어휘 목록, 워드리스트 분석, 전문분야 통역, 연설문, 코퍼스 분석 도구

이지혜

이화여자대학교 통역번역대학원 통역번역학과 박사과정

julia429@hanmail.net

관심분야: 통번역교육, 코퍼스언어학, 문학번역, 기계번역

이용훈

충남대학교 영어영문학과 강사/한국코퍼스영어학통계연구소 소장

yleeuiuc@hanmail.net

관심분야: 코퍼스언어학, 전산언어학, 머신러닝/딥러닝

논문투고일: 2020년 11월 8일

심사완료일: 2020년 11월 26일

게재확정일: 2020년 11월 30일