

## 전문분야 통역을 위한 어휘 지식: IT 분야를 중심으로

최 문 선  
(이화여대)

### 1. 서론

일반적으로 특수한 분야의 전공자가 아닌 제너럴리스트(*generalist*)로서의 통역사가 전문지식을 지닌 전문가(*specialist*)들 간의 담화를 통역한다는 것은 결코 쉬운 일이 아니다. 모국어 화자라고 하더라도 전문적인 배경지식이 없을 경우 전문분야의 담화를 이해조차하기 어렵다. 그렇기 때문에 통상적으로 해당 분야의 전문가가 외국어를 습득하거나 혹은 이중 언어자인 전문가가 통역을 하는 것이 이상적이라는 말을 하기도 한다. 그러나 통역에 필요한 기술을 훈련받지 않은 상태로 이중 언어자인 전문가가 전문 통역사를 대체하려 했을 때 오히려 내용의 완결성 저하, 연사 의도 왜곡, 통역의 신속성에 대한 압박을 견디지 못해 중도 포기하는 등의 문제가 발생하여 소통 실패로 이어질 가능성이 높다.

실제로는 현재 많은 실무 통역사들이 해당 분야의 전문가가 아님에도 불구하고 고도로 전문적인 영역에서 활동하고 있다. 이것이 어떻게 가능할까? 통역사들의 지식 습득 방식과 그 방식에 있어서 어휘의 중요성에서 부분적으로나마

답을 찾을 수 있다. 전문분야의 지식을 습득할 때 상위 개념에서 시작해서 세부적인 기술적 지식으로 내려가는 하향식 방식과는 달리 통역사들은 사용가능한 시간적 제약, 통역을 특정 분야로 한정짓기 어려운 현실적인 제약 등으로 인하여 주어지는 입력물(*input*) 중심으로 정보를 습득하는 상향식 지식 습득 (*Fantinuoli 2006: 172*)의 방식을 취한다. 통역사들은 해당 분야의 비전문가로서 개념적 지식(*conceptual knowledge*)에 접근하기가 쉽지 않고, 특히 매번 다른 주제를 다루어야 하는 시간적 압박 하에서 효율적인 준비를 위해 언어적 지식 (*linguistic knowledge*)을 중심으로 통역을 준비하게 되기 때문이다. 이러한 상향식 지식 습득에 있어서 가장 기본적인 단위가 되는 것은 어휘이다. 전문분야 담화의 가장 중요한 언어적 특징은 그 어휘에 있고, 전문 담화를 구성하는 어휘와 관련 지식에 대한 준비 없이는 전문분야의 통역이 제대로 이루어질 수 없으며, 따라서 전문 담화를 구성하는 어휘 능력은 통역사의 핵심 능력 중 하나이다 (*Garzone 2000: 82*).

그렇다면 소위 ‘전문용어 목록’만 숙지하면 통역이 가능할까? 전문분야를 다루는 통역사에게 필요한 어휘가 무엇인지는 단순한 문제는 아니다. 어휘를 무엇으로 볼 것인가부터 통역에 필요한 이해의 수준에 도달하기 위한 어휘 지식(*lexical knowledge*)의 상대성도 고려해야 한다. 뿐만 아니라 통역의 정보 수신 단계와 재표현 단계에서 요구되는 어휘 지식의 폭과 깊이도 다르다. 흔히 전문분야에서 해당 분야의 주요 개념을 표현하는 어휘를 전문용어라 하는데 통역 시 이해와 재표현에 필요한 어휘는 통상의 전문용어보다 범위가 더 넓다. 대표적인 예가 고유명사이다. 기업이나 기관명, 제품명, 약물명, 공정의 명칭 등을 지칭하는 고유명사의 경우 통역 이해 단계에서 해당 고유명사가 구체적으로 어떤 사물이나 대상을 지칭하는지에 대한 지식이면 충분하며 고차원적인 개념 이해나 용례에 대한 지식 등을 아우르는 깊은 수준의 어휘 지식이 필수적이지는 않다. 그러나 해당 고유명사가 출발어에서 어떻게 발음되며 도착어에서는 어떻게 발음되는지 등 해당 어휘의 소리 정보가 통역사의 지식 속에 있지 않다면 이해 단계에서 제대로 포착되지 못하거나 재표현 단계에서 도착어 청중에게 제대로 전달되지 못할 가능성이 매우 높다.

일반적인 전문용어는 아니나 성공적인 통역에 필수적인 어휘의 유형 중 또 한 가지는 동사이다. 동사는 문장 내의 서술어로서 기능하며 하나의 문장을 완

결하기 위해서 반드시 요구되는 성분이다. 그런데 전문분야에서 쓰이는 동사의 경우 일반적으로 용례를 찾기 어려운 저빈도 어휘여서 이해나 재표현에 어려움을 야기하거나, 고빈도 어휘임에도 불구하고 해당 분야에서만 특정한 의미를 취하는 경우 문제가 될 수 있다. 따라서 문장의 완결적인 서술어에 해당하는 동사 역시 전문분야 통역을 위해 습득이 필요한 핵심 어휘군에 해당한다.

본 연구에서는 전문분야 담화를 성공적으로 통역하는데 필수적인 어휘를 도출하고 특징에 따라 분류하여 어휘의 유형별로 어떠한 어휘 지식이 필요한지 고찰함으로써 향후 전문분야 통역의 교육과 실무에 유용한 자료를 제공하고자 한다. 이를 위하여 현재 전 세계적으로 진행 중인 4차 산업혁명의 핵심 분야이자 전문 통역사에 대한 수요가 꾸준히 이어지고 있는 정보통신기술(이하 IT) 분야의 통역 대상 텍스트로 간이 코퍼스를 구축한 후 이를 바탕으로 전문분야 담화의 어휘적 특징을 고찰하고, 해당 코퍼스의 키워드 분석을 바탕으로 IT 통역시 통역사들이 습득해야 하는 어휘를 종합적으로 제시하고자 한다. 또한 IT 통역의 준비와 수행에 어려움을 야기할 수 있는 유형의 어휘 유형을 사례 중심으로 고찰함으로써 IT 전문 담화의 통역에 필요한 어휘 지식이 어떠한 성격을 지니는지 분석하고자 한다.

전술한 본 연구의 목적을 달성하기 위하여 이어지는 2장에서는 전문분야 담화의 이해에 필요한 어휘 및 이를 연구하기 위한 방법과 관련된 이론적 배경을 살펴보고, 3장에서는 IT 전문 담화 어휘 분석을 위한 연구 방법을 기술한다. 4장에서는 IT 전문 담화의 어휘적 특징과 도출된 어휘의 성격을 고찰함으로써 IT 전문 담화의 통역에 필요한 어휘 및 요구되는 지식의 성격에 대해 종합적으로 분석 및 논의한 후 5장에서 결론으로 마무리한다.

## 2. 이론적 배경

### 2.1. 통역과 전문분야 담화의 이해

철(Gile)은 통번역사 교육을 위한 기본 개념과 모델을 설명한 저서(Gile 2009)에서 통번역사가 전문분야 담화를 이해하는 과정을 상세히 풀이하였다.

그의 논의는 통역사가 해당 분야 전문가가 아닌 ‘일반인(layperson)’일지라도 언어적 지식, 언어외적 지식, 분석을 통해 고도로 전문적인 분야의 텍스트나 구두 담화를 충분히 이해할 수 있다는 전제에서 출발한다(Gile 2009: 89). 철(2009: 92-93)에 따르면, 정보적 텍스트를 구성하는 문장은 다음의 세 가지 요소로 구성된 일종의 네트워크 구조로 표현할 수 있다.

(1) 명사적 객체(Nominal Entity 혹은 NE): 인물, 사물, 아이디어, 행위 등을 지칭하는 명사 및 명사구.

(2) 속성(Attribute 혹은 A): 형용사, 형용사 유사어, NE에 해당하는 인물, 사물 등을 설명하는 절(clause), NE의 존재, 소멸, 성장 등 NE에 대한 진술(예를 들어 ‘X가 존재한다’, ‘Y가 성장했다’ 등 X, Y라는 NE에 대한 설명이 여기에 해당).

(3) 연결고리(Link 혹은 L): 문장 구조와 문법 규칙으로서 NE에 해당하는 인물, 사물, 개념 간의 논리적 혹은 기능적 관계를 나타냄(‘A는 B와 비교된다’, ‘A가 B에 작용한다’ 등 복수의 NE 간의 관계를 설명).

철의 주장에 따르면 아무리 전문적이고 복잡한 문장이라도 이를 하나의 의미 네트워크(semantic network)로 보고 위의 세 가지 구성요소로 분해할 수 있다. 단적인 예로 철이 제시한 ‘A good methodology results in powerful algorithm’이라는 문장(Gile 2009: 92)의 경우 ‘a good’, ‘powerful’은 속성(A), ‘methodology’, ‘algorithms’는 명사적 객체(NE), ‘results in’은 연결고리(L)로 분석할 수 있다. 이러한 분석을 통해 각 어휘의 의미는 접어두더라도 ‘methodology’와 ‘algorithms’가 원인과 결과의 관계를 형성한다는 논리를 파악할 수 있는 것이다. 이러한 원리를 따를 때 고도로 기술적이고 복잡한 문장이라고 할지라도 문장이 전달하는 아이디어의 논리적 구조는 일반인이라도 충분히 이해 가능하다. 통역사는 이처럼 문장의 논리적인 구조를 분석함으로써 상당한 수준의 이해에 도달하게 된다. 철이 예시한 것처럼 비단 단일 문장 내에서 뿐만 아니라 여러 문장에 걸쳐 NE, A, Link를 분석하고 이를 다시 연결시킴으로써 하나의 의미 네트워크를 형성할 수 있다. 나아가 극도로 전문적인 텍스트라 하더라도 속성과 연결고리의 대부분은 비전문적 정보성 텍스트에 흔히 볼 수 있

는 것과 동일하며, 전문적이지 않은 일반적인 단어와 구조를 통해 표현된다는 점(Gile 2009: 93) 역시 전문분야 텍스트 혹은 담화의 이해가 분석을 통해 가능하다는 점을 시사한다.

따라서 복잡하고 기술적인 전문분야의 담화라고 할지라도 기본적으로 문장의 의미 네트워크, 즉 기능적, 논리적 구조에 대한 분석을 바탕으로 명사/명사구, 혹은 동사 등의 목표어 대응어를 안다면 전문분야의 통역이 가능하다(Gile 2009: 96). 다시 말해서 비전문가인 통역사라고 할지라도 의미 네트워크를 구성하는 요소인 NE, A, Link를 구분해 내고 이들 각 요소에 해당하는 어휘에 대한 지식을 갖추으로써 전문분야 담화의 통역이 가능하다는 것을 알 수 있다.

## 2.2. 통역에 필요한 어휘

일반적으로 기술적이고 생소한 주제의 통역을 준비할 때 거의 모든 통역사가 용어집을 구축한다(Jiang 2013). 전문분야의 통역을 위해 준비하는 용어 혹은 어휘로서 가장 먼저 떠오르는 것은 전문용어이다. 특정 영역(domain)을 기술하고 정의하는 데 사용되는 일련의 어휘(Thomas 1993: 43)인 전문용어는 전문분야 담화를 이해하고 도착어로 재표현하는 데 필수적인 빌딩블록임에 틀림없다. 그러나 통역에 필요한 어휘는 소위 전문용어사전에 실린 용어 항목에 국한되지 않는다. 예를 들어 최문선(2018)의 연구에서 실제로 전문통역사들이 만든 용어집의 내용을 살펴보면 명사나 명사구 위주의 전문용어만 포함된 것이 아니라 동사, 형용사와 같이 ‘술어적 특성’을 지니는 항목도 포함되어 있었다. 이는 남슬기 외(2018)가 통역사가 선택한 핵심정보의 형태적 특징을 고찰한 결과 통역사의 원문 이해에서 가장 핵심적인 역할을 하는 핵심 정보는 ‘서술어’인 것으로 나타난 것과 무관하지 않다.

이미경(2018)은 통역 과정에서 소위 ‘전문용어’가 미치는 영향을 고찰하면서 실제 연사의 발화에 나타난 용어, 즉 통역사의 주의 집중을 요하는 대상을 고유명사, 일반 명사, 전문용어, 동사로 유형화하였다. 특히 일반 명사가 전문적인 맥락에서 사용될 때 고빈도로 호응하는 동사(이미경 2018: 137)를 통역 과정에 영향을 미치는 중요한 항목으로 지적하였다. 이 중에서 ‘일반 명사가 전문적인 맥락에서 사용된 경우’는 일반 언어 학습자들이 가장 어려움을 겪는 항목

으로도 주목된 바 있다. 토드(Todd 2017)는 다의어 중 해당 문맥에서의 의미가 일반적인 의미가 아닌 용례를 보이는 어휘를 ‘불투명 어휘(opaque words)’라고 명명하고 이것이 학습자에게 가장 큰 어려움을 야기한다고 보았다. 피터스와 페르난데즈(Peters & Fernández 2013) 역시 특수목적영어(ESP) 학습자가 가장 어려움을 겪는 어휘는 전문용어가 아니라 일반적인 영어 어휘이면서 전문분야에 특화된 의미를 갖는 어휘라고 밝혔다. 농업 분야의 기술적 어휘 목록을 제시한 무뇨스(Muñoz 2015)가 ‘특정 분야에 의미적으로 연관되어 분야 특수적인 개념을 가리키는 용어’(Muñoz 2015: 34)라고 정의한 기술적 어휘(technical word)도 이와 유사한 개념이다.

이처럼 관련 선행연구에 비추어 볼 때 전문분야 통역에 필요한 어휘는 전문용어뿐만 아니라 명사/명사구 및 술어적 특성을 지니는 동사 및 형용사(최문선 2018), 원문 이해에 핵심적인 서술어(남슬기 외 2018), 전문 담화 내에서 일반 명사와 자주 호응하는 특정 동사(이미경 2018), 전문 담화의 맥락에서 사용될 경우 일반적인 의미에서 벗어나 학습 난이도가 높은 어휘(Peters & Fernández 2013; Todd 2017), 기술적 어휘(Muñoz 2015)를 중심으로 구축될 수 있다.

전문분야 통역에 필요한 어휘를 도출하는 데 있어서 또 한 가지 중요한 고려 사항은 어휘를 추출하는 대상을 어떻게 설정할 것인가라는 문제이다. 칠(2002)은 영-불 의학 회의 통역을 준비하는 과정에서 의학 교과서에 실린 영어 용어를 추출하였는데 실제 통역 시 거의 활용되지 못했고 용어집에 포함되지 않았던 전문용어도 매우 많았다고 보고하였다. 이러한 결과는 구어적인 특수목적언어(LSP) 커뮤니케이션 상황인 의학 통역, 그리고 의학 교과서라는 매우 정제된 문어적 언어 자료와의 차이에서 기인하는 것일 수 있다. 예를 들어 길고 복잡한 의학 용어의 경우 발음과 용례를 익혀야 유용하게 활용될 수 있는데 문어 자료로는 이러한 정보의 획득이 제한적이기 때문이다. 이에 비추어 볼 때 통역에 필요한 어휘를 효과적으로 추출하기 위해서는 최대한 해당 통역 상황과 유사하고 실제성 있는 언어 자료를 활용할 필요가 있다.

## 2.3. 전문분야 어휘 목록

특정한 전문분야에서 자주 사용되는 어휘를 확인하는 방법 중 하나는 전문

가 자문으로서 주어진 분야에서 성공적인 커뮤니케이션에 필수적인 어휘 항목을 목록화해 달라고 요청하는 것이다. 그러나 이 방법은 전문가의 주관성에 크게 의존한다는 문제를 안고 있다(Szudarski 2018: 188). 주관성을 보완하기 위한 방법으로 평가 척도를 이용하는 방법도 가능하다. 정과 네이션(Chung & Nation 2003)은 전문분야와의 관련성을 4점 척도로 구분하여 전문가들로 하여금 평가하도록 하였으며 다른 기존의 방식에 비해 정확도와 평가자간 신뢰도도 높다는 것을 발견하였다. 그러나 평가 척도를 적용할 어휘 목록이 미리 작성, 제시되어야 한다는 점에서 이 방법도 독립적으로 적용하기에는 어려움이 따른다.

이와 관련하여 언어 학습의 기초 어휘를 확립하기 위한 어휘 목록(word list) 연구들을 참고할 필요가 있다. 언어 학습자가 보다 효율적으로 어휘를 학습할 수 있도록 돕기 위하여 가장 자주 사용되는 어휘를 목록화하는 다양한 시도가 이어져 왔다. 일반적으로 웨스트(West 1953)가 영어 학습자를 위해 가장 흔히 사용되는 2,000개 어휘군(word family)을 목록화한 GSL(General Service List)를 출발점으로 보고 있으며, 최근에는 브레지나와 가블라스바(Brezina & Gablasova 2015)가 GSL의 단점을 보완하여 NGSL(New General Service List)를 발표하였다. 국내에서도 한국어의 기초 어휘를 선정, 등급화하거나(이삼형 2018) 외국어로서의 한국어 학습자들을 위한 한국어 어휘 목록이 개발되었다(김창구 2016).

일반적인 언어 사용에서 한 걸음 더 나아가 보다 구체적으로 분야를 설정한 후 해당 분야의 어휘 목록을 제안하는 연구도 다수 이루어졌다. 대표적인 예가 콕스헤드(Coxhead 2000)의 AWL(Academic Word List)로서 GSL에 포함되지 않은 단어 중 주로 학술 영어에 자주 사용되는 570개의 어휘군과 3107개의 어종(type)을 목록화하였다. 보다 최근에는 공학(신창원 2012, 하명호 2017, Todd 2017), 농업(Muñoz 2015), 환경(Liu & Han 2015), 금융(Tongpoon-Patanasorn 2018), 간호학(Yang 2015), 해사(장세은과 변현정 2011) 등 특정 전문분야별 어휘 목록을 제시하는 연구가 폭넓게 이루어졌다.

어휘 목록을 개발한 일련의 연구(Liu & Han 2015; Muñoz 2015; Tongpoon-Patanasorn 2018; Yang 2015)를 살펴보면 특정 전문분야별 어휘 목록을 구축하기 위해서는 구체적인 방법론적 고려가 필요함을 알 수 있다. 우선

어휘를 추출하기 위한 언어 자료를 수집해야 한다. 이를 위해서 어휘 목록의 구체적인 목적 하에 특수목적 코퍼스를 구축하는 기준에 맞추어 자료의 크기, 대표성, 출판/발표의 시점 등을 고려해야 한다. 다음으로 수집된 자료는 연구의 목적에 맞는 전산 처리가 가능하도록 정제하고 처리한다. 마지막으로 연구목적에 맞는 분석용 도구를 활용하여 일련의 기준에 따라 어휘 목록을 도출한다.

정제, 처리된 코퍼스를 바탕으로 어휘목록을 도출하는 방법은 일반적으로 빈도(frequency), 범위(range), 산포도(dispersion)를 바탕으로 어휘의 중요도 혹은 기본도(basicness)를 산출하는 것이다. 이는 네이션 교수 및 그를 중심으로 하는 학파에서 제안한 이론으로 세계의 영어 교육계에서는 이 이론이 매우 널리 사용되고 있다(이삼형 2018: 56). 한국어의 기초어휘목록을 포괄적으로 작성한 이삼형(2018)의 연구, 일본 대학의 교양학부 한국어 과정 교수학습을 위한 참조 어휘목록을 작성한 김창구(2016)의 연구가 그러한 사례이다. 네이션 학파에서 빈도, 범위, 산포도를 어휘 점수 산출의 지표로 제안하기는 했으나 이 지표들을 어떻게 조합하여 최종적인 어휘 점수를 산출할 것인지에 대해서는 다양한 견해가 존재한다(이삼형 2018: 58). 일례로 리우와 한(Liu & Han 2015: 5)의 경우에는 빈도와 텍스트상 분포를 결합하는 ‘이용률(usage)’ 점수를 개선하여 ‘최적 이용률(optimized usage)’ 공식을 자체적으로 개발, 적용한 바 있다.

빈도 중심의 어휘 선정 방법과는 달리 코퍼스 언어학의 도구를 이용한 키워드 분석(Keyword Analysis)을 중심으로 특정 분야의 어휘 목록을 도출할 수도 있다. 빈도와 분포를 중심으로 어휘목록을 도출할 경우 대상 코퍼스에 골고루 분포하며 일정 빈도 이상으로 출현하는 어휘를 선정한다. 그러나 키워드 분석은 절대적인 빈도가 아니라 비교 대상이 되는 참조 코퍼스(reference corpus)와 비교하여 상대적으로 높은 빈도를 보이는 어휘를 뽑아내는 방식으로서 연구 대상 코퍼스(study corpus)의 특징을 보여주는 어휘 목록을 산출하는 데 용이하다. 공학 분야의 ‘불투명 어휘’ 목록을 개발한 토드(2017), 키워드 분석에 평가 척도를 결합하여 금융 용어 목록을 개발한 톱푼-파타나손(Tongpoon-Patanasorn 2018)의 연구에서 이러한 사례를 볼 수 있다.

지금까지 본 장에서 고찰한 이론적 배경을 정리하면 다음과 같다. 전문 담화의 통역에 필요한 통역사의 이해 과정은 문장을 명사적 객체, 속성, 연결고리라는 개별 구성요소로 분해한 후 이를 연결하여 하나의 의미 네트워크를 구성

함으로써 가능하며, 이는 일반인의 이해 과정과 크게 다르지 않다(Gile 2009). 이것은 전문 담화의 통역에 있어서 의미 네트워크를 구성하는 개별 요소에 대한 통역사의 지식이 매우 중요함을 의미한다. 이러한 개별 요소의 단위를 어휘로 볼 때 통역에 필요한 어휘가 무엇인지에 대해 이해할 필요가 있으며, 기존의 관련 연구(남슬기 외 2018; 이미경 2018; 최문선 2018; Jiang 2013 등)에서 밝혀진 바와 같이 전문 담화 통역에 필요한 어휘는 소위 ‘전문용어’에 제한되지 않으며 특히 서술어 역시 중요한 어휘로 고려되어야 함을 알 수 있다. 보다 구체적으로 어휘 목록을 구축하는 방법론과 관련하여 언어 학습자를 위한 기초어휘목록, LSP 학습자를 위한 용어목록 연구를 살펴보았다. 이어지는 장에서는 이러한 관련 연구에서 적용된 어휘 목록 산출 방법을 기반으로 전문분야 통역에 필요한 어휘를 도출하고 주요 유형별 특성을 고찰하고자 한다.

### 3. 연구 방법

#### 3.1. IT 전문 담화 코스퍼스 구축

본 연구는 서론에서 언급한 바와 같이 연구 대상인 전문분야로 IT 분야를 선정하였다. IT는 전문통역에 대한 수요가 왕성히 이어지고 있는 분야로서 대학원 수준의 전문적인 통역 교육 과정에서도 비중 있게 다루어질 필요가 큰 분야이기 때문이다.

IT 분야의 전문적인 내용을 다루는 통역은 다양한 목적 하에 여러 가지 형식을 띠 수 있다. IT 기업이나 일반 기업의 IT 부서 혹은 IT 프로젝트팀에 소속되어 일하는 사내 통역사의 경우 현업의 실무진과 IT 담당자 간의 의사소통을 통역하거나, 프로그래밍을 담당하는 개발자들 간의 개발 업무와 관련된 고도로 기술적이고 구체적인 내용에 대한 통역을 수행할 수도 있다.

프리랜서 통역사의 경우에는 일반에게 공개되는 대규모 회의에서 이루어지는 각종 발표에 대한 통역을 의뢰받는 경우가 많다. 최근 마이크로소프트, 아마존웹서비스, 구글 등의 글로벌 IT 기업들은 전 세계 주요 도시에서 자사 기술 및 신제품의 최신 업데이트와 향후 사업 방향성을 공유하는 대규모 기술 컨퍼

런스를 개최하는 경향을 보이고 있다. 일반적으로 본사가 위치한 미국 내에서 먼저 컨퍼런스를 개최하고 거의 유사한 내용을 가지고 주요 시장 별로 투어를 하듯 순차적으로 현지 컨퍼런스를 개최하는 방식을 취한다. 영어 발표가 대부분을 차지하여 최신 기술 발전 내용을 영어가 아닌 현지어로 듣고자 하는 고객이 많을 경우 해당 언어로 동시통역이 이루어진다. 한번에 100여개가 넘는 세션으로 구성되는 경우도 있어 통역을 요하는 세션의 수도 상당하며 이에 따라 다수의 통역사가 동시에 투입된다. 이와 같은 IT 통역의 구체적인 의사소통 상황(communicative situation)에 따라 사용되는 언어의 어휘적, 문법적, 담화적 특징도 달라진다.

본 연구의 목적이 전문통역의 교육과 실무에 유용한 어휘를 도출하는 것임을 감안할 때 코스퍼스를 구축하는 단계에서 중요한 고려사항 중의 하나는 수집되는 자료의 실제성(authenticity)이다. 가장 이상적인 방법은 실제로 통역이 이루어진 발표자료를 수집하는 것이나 대부분 저작권 문제로 인하여 통역이 발표 원문과 함께 공개되는 경우는 극히 드물다. 따라서 차선책으로 일반에 공개된 행사의 프로그램을 살펴보고 동시통역이 제공되었다는 표시가 있는 세션을 기준으로 관련 자료를 수집하는 방법을 취하였다. 이에 본 연구에서는 표1과 같은 특징을 지니는 의사소통 상황을 연구의 대상으로 한정하고 IT 분야를 다루는 일방향 구어 발표 중 영어로 된 담화로 코스퍼스를 구축하였다.

〈표 1〉 IT 전문 담화 코스퍼스 특징

어역(register <sup>1)</sup> )	다수의 청중을 상대로 공식적인 상황에서 구어로 이루어지는 일방향성 발표
분야(discipline)	IT
출처	글로벌 IT 기업이 개최하는 대형 컨퍼런스의 발표 세션
시기	2019년 4월
연사	IT 전문가 (tech professionals)
청중	IT 전문가 (tech professionals)
목적	정보 전달, 설명, 교육
언어	영어

1) 어역은 ‘특정한 상황에서 사용되는 언어의 종류’를 뜻한다(Biber & Conrad 2009, Szudarski 2018:138에서 재인용).

보다 구체적으로는 행사의 발표 세션을 녹화하여 인터넷 상에서 주문형 콘텐츠(VOD)로 제공하고 있는 마이크로소프트사(社)의 연례기술 컨퍼런스인 ‘Microsoft Ignite | The Tour’ 시리즈<sup>2)</sup>(이하 ‘이그나이트’) 중에서 2019년 4월 서울에서 개최된 ‘Microsoft Ignite | The Tour 서울<sup>3)</sup>’의 행사 아젠다(세션 카탈로그<sup>4)</sup>)에서 ‘한국어 통역 제공’이라는 표시가 있는 세션을 확인하였다. 해당 행사는 이틀간의 일정으로 약 120여개의 발표 세션이 진행되었으며 이 중에서 영한 동시통역으로 진행된 세션 중 표1과 같이 일방향성 발표의 형식을 취한 세션은 총 47개로 파악되었다. 해당 발표들의 경우 연사와 청중 모두 IT 전문가로서 특정 주제에 대한 정보 전달이나 설명 및 교육을 목적으로 하는 전문가 간의 의사소통이었다.

전술한 절차대로 선택한 47개 세션 중 이그나이트 웹사이트 상 주문형 콘텐츠가 제공되는 세션은 총 30개였다<sup>5)</sup>. 이 중에서 기술적 문제가 있었던 2개 영상을 제외하고 총 28개 영상을 연구 코퍼스에 포함시켰다. 연구 코퍼스에 포함된 동영상의 총 재생시간은 약 26시간(25시간47분34초)으로 집계되었다. 각 동영상의 자막을 다운 받아 특수 기호 등 불필요한 부분을 제거하고 오달자를 수정하는 정제 과정을 거친 후 텍스트 파일로 저장하여 IT 전문 담화 코퍼스(이하 ‘IT 코퍼스’)를 구축하였다.

### 3.2. IT 전문 담화 어휘 목록 구축

2장에서 고찰한 바와 같이 주어진 코퍼스를 바탕으로 전문분야의 어휘 목록을 도출하는 방법에는 코퍼스 상에서 해당 어휘의 빈도, 범위, 산포를 고려하여 가장 폭넓게 자주 출현하는 어휘를 선정하는 방법과 참조 코퍼스 대비 키워드 분석을 수행함으로써 연구 코퍼스의 특징을 반영하는 어휘를 선정하는 방법이 있다. 기본적으로 빈도를 중심으로 하는 전자의 방법은 고빈도 어휘일수록

2) 본 행사의 개요는 <https://www.microsoft.com/ko-kr/ignite-the-tour/> 참조.  
 3) 서울 행사는 [https://news.microsoft.com/ko-kr/2019/03/27/ms\\_ignite\\_the\\_tour\\_seoul/](https://news.microsoft.com/ko-kr/2019/03/27/ms_ignite_the_tour_seoul/) 참조.  
 4) <https://sessioncatalog.myignitetour.techcommunity.microsoft.com/seoul>에서 검색 가능.  
 5) 동영상 콘텐츠 검토 결과 서울 행사를 녹화한 자료는 아니었으나 세션 제목이 완전히 일치하여 서울 행사의 발표 내용과 대동소이할 것으로 판단하여 그대로 사용하는데 무리가 없을 것으로 판단하였다.

언어 학습자에게 유용성이 높다는 점을 전제로 한다. 그러나 공학 등 특정 전문분야의 어휘를 학습하는 데 있어서는 일반적인 언어 사용 환경 대비 해당 전문분야에서 더 자주 출현하는 어휘들이 더욱 유용하다(Gries 2015: 55). 즉 절대 빈도가 아닌 상대빈도를 고려하는 키워드 분석 방식을 적용함으로써 전문분야를 다루는데 보다 유용한 어휘 목록이 도출될 수 있다. 이에 따라 IT라는 특정 전문분야 통역의 교육과 실무에 유용한 어휘를 고찰하고자 하는 본 연구에서는 3.1.에서 구축한 IT 코퍼스를 연구 코퍼스, BNC를 참조 코퍼스로 두고 워드스미스(Version 7)<sup>6)</sup>를 이용하여 키워드 분석을 실시하였다. 로그 라이클리후드(Log Likelihood) 값 기준으로 상위 500개의 키워드를 추출한 후 다음과 같은 절차를 거쳐 IT 전문분야 담화의 특징을 반영하는데 불필요하거나 중복되는 항목을 제거하였다.

- 콘코던스 라인 확인 후 전문분야(IT)와 내용 관련성이 낮은 일반 어휘라고 판단되는 경우 이중언어 구사력이 높은 수준에 도달해 있는 통역학습자 및 통역사에게 유용성이 매우 낮으므로 목록에서 제거(예: nice, faster, favorite, see, then)
- 대명사 및 대명사 포함 어구 제거(예: you, we, it's)
- 동사의 굴절형, 명사의 단수/복수형은 하나의 항목으로 통합

## 4. 결과 및 논의

### 4.1. IT 전문 담화의 어휘적 특징

앞의 3.1.의 방법을 통해 최종적으로 구축된 IT 코퍼스의 특징을 살펴보기 위하여 워드스미스를 이용해 산출한 통계를 기존의 선행 연구에서 분석한 코퍼스와 비교하였다. 비교를 위하여 일반 영어의 특징을 보여주는 것으로 널리 활용되는 BNC, 그리고 본 연구의 연구 코퍼스와 유사한 분야인 전자공학 및 컴퓨터 공학 참고문헌과 저널논문을 수집한 공학 코퍼스의 어휘적 특징을 비교한 크하미스와 호-압둘라의 분석 데이터(Khamis & Ho-Abdulla 2017: 112)를 원용하였다.

6) Scott, M., (2019), WordSmith Tools. Stroud: Lexical Analysis Software.

〈표 2〉 IT 코퍼스의 어휘적 특징(BNC, 공학 포커스와 비교)

구분	IT 코퍼스	BNC	전자공학 및 컴퓨터 공학 코퍼스
출현형(token)수	256,859	97,860,872	601,481
어종(type)수	7,093	512,588	12,458
STTR	32.27	43	30
평균단어길이 (글자수)	4.27	5	4.85
1~4개 글자로 이루어진 단어 비중	65.19%	58%	54.16%

〈표 2〉에서 출현형(혹은 토큰)수는 각 코퍼스에 등장한 단어수로서 각 코퍼스의 크기를 보여준다. 어종(혹은 타입)수는 중복 출현을 배제한 단어수를 의미하며 사용된 어휘의 다양성을 가늠하는 지표인 타입 대 토큰 비율(type to token ratio, 이후 TTR)을 계산하는 데 사용된다. 각 코퍼스의 크기에 큰 차이가 있으므로 토큰 1,000개당 포함된 타입수로 표준화한 STTR(standardized TTR)을 산출하였다. 그 결과 IT 코퍼스의 STTR은 BNC에 비하여 크게 낮고, 공학 코퍼스와는 유사한 수준인 것으로 나타났다. 다시 말해 본 연구에서 구축한 IT 코퍼스는 또 다른 전문분야 코퍼스인 공학 코퍼스와는 어휘 다양도가 비슷한 수준이나 일반 영어에 비해서는 상대적으로 낮은 수준으로 나타났다. 이는 특정 전문분야의 담화(기일 경우 더욱 한정된 수의 특수한 어휘를 사용하기 때문인 것으로 해석할 수 있으며, 전문분야 담화의 통역에 있어 상대적으로 한정된 수의 분야 특수적인 어휘를 파악하고 이를 학습할 필요성을 확인해 준다.

세 코퍼스의 평균단어길이를 비교해 보면 IT 코퍼스의 평균단어길이가 가장 짧은 것을 알 수 있다. 일반적으로 단어길이는 이독성(readability)에 영향을 미치며 평균단어길이가 길수록 일반적으로 텍스트 난이도가 높아지는 것으로 알려져 있다. 따라서 IT 코퍼스에 포함된 IT 전문가 발표는 BNC나 공학 코퍼스에 비해 이독성이 낮은, 상대적으로 이해하기 쉬운 텍스트일 것으로 추정할

수 있다. 이러한 결과는 BNC와 공학 코퍼스가 문어 텍스트인데 반해 IT 코퍼스가 정보 전달을 목적으로 하는 대중연설 형태의 구어 코퍼스라는 점에 기인하는 것으로 보인다. 1~4개 글자로 이루어진 단어 비중에서도 IT 코퍼스가 가장 높은 수치를 기록하고 있어 문어 코퍼스인 나머지 2개 코퍼스에 비해 짧은 어휘의 비중이 높다는 것을 알 수 있다.

일반 영어인 BNC, 전문분야 영어인 공학 코퍼스와 비교했을 때 본 연구에서 구축한 IT 코퍼스의 어휘적 특징은 첫째, STTR을 기준으로 본 어휘 다양도는 상대적으로 낮게 나타나 전문분야 영어의 특징을 그대로 반영하고 있음을 확인하였으며, 둘째, 사용된 어휘의 평균길이가 짧고, 짧은 어휘의 비중이 높은 것으로 나타났다. 한 마디로 본 연구를 위해 구축한 IT 코퍼스는 어휘 다양도가 높지 않고 상대적으로 짧은 어휘가 다수 사용된 담화라는 특징을 지닌다. 이러한 특징을 볼 때 IT 전문가 발표의 어휘 난이도는 대체로 평이한 수준일 것이라는 추론이 가능하다. 그렇다면 IT 코퍼스의 전문성, 기술성을 높이는 것은 결국 분야 특수적인, 상대적으로 한정된 수의 어휘일 가능성이 높으므로 전문분야의 성공적인 통역을 위해서는 이러한 분야 특수적인 어휘에 더욱 집중할 필요성이 제기된다.

IT 코퍼스의 어휘 난이도별 구성을 살펴보기 위하여 AntWordProfiler<sup>8)</sup>를 사용하여 GSL 어휘 중 빈도 기준 상위 1,000위에 속하는 GSL1과 그 아래부터 2,000위에 속하는 GSL2 어휘, 학술 영어인 AWL 어휘, GSL과 AWL에 수록되지 않은 저빈도 어휘(Off-list)의 비중을 살펴보았다.

〈표 3〉 IT 코퍼스의 어휘 구성

구분	토큰(%)	타입(%)
GSL1	82.89	30.22
GSL2	3.88	13.25
AWL	5.58	14.83
Off-list	7.65	41.69

7) 본고에서 분석한 코퍼스는 마이크로소프트사의 이그나이트 행사에서 이루어진 발표로 구성되었으므로 ‘특정 전문분야’라는 표현은 이러한 범위에 한정하여 해석될 필요가 있다.

8) Anthony, L. (2014). AntWordProfiler (Version 1.4.0) [Computer Software]. Tokyo, Japan: Waseda University. <https://www.laurenceanthony.net/software>에서 다운로드.

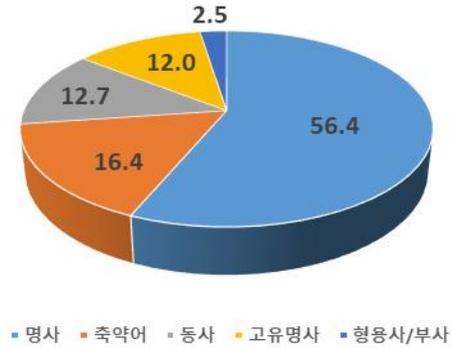
<표 3>에서 볼 수 있듯이 IT 코퍼스는 출현형(token) 기준으로 일반 어휘가 약 86.7%(GSL1과 GSL2 합계)를 차지하였고 학술 어휘(AWL)가 5.6%, 저빈도 어휘(GSL과 AWL에 포함되지 않은 어휘)가 7.6%를 차지하고 있는 것으로 나타났다. 그러나 출현형의 중복을 제거한 어종(type) 기준으로 코퍼스의 어휘 구성을 살펴보면 일반 어휘는 약 43.5%, 학술 어휘는 14.8%, 기초 어휘목록에 포함되지 않는 어종의 비중은 약 41.7%로 크게 달라진다. 다시 말해 토큰 기준으로 볼 때는 대부분이 빈출 어휘에 속하는 일반적인 단어로 구성된 듯 보이나, 타입 기준으로 보면 GSL, AWL 등 빈출 어휘에 포함되지 않는 저빈도 기타 어휘의 비중이 절대적으로 높은 것을 알 수 있다. 이러한 결과는 금융 분야의 전문어휘 목록을 개발하고자 했던 통푼-파타나손(2018)의 연구에서 구축한 금융 코퍼스에서도 기타 어휘의 비중이 토큰 기준으로는 약 14퍼센트였으나 타입 기준으로는 무려 93퍼센트를 차지하는 것으로 나타난 결과(Tongpoon-Patanasorn 2018: 50)와도 일맥상통한다. 즉, 금융 코퍼스와 같이 전문분야에 특화된 코퍼스로서 본 연구에서 구축한 IT 코퍼스는 저빈도 기타 어휘가 다수 포함되어 있다는 것을 알 수 있으며, 담화 내에서 빈출 어휘로 구성되는 부분에 비해 저빈도 기타 어휘 출현 시 통역에 어려움을 증가시킬 것을 예상할 수 있다. 통역, 특히 전문분야를 다루는 통역의 경우 이처럼 한, 두 번 등장하는 저빈도 어휘를 처리하지 못할 때 통역 실패로 이어질 위험이 높다. 그러므로 전문분야 통역을 위한 어휘를 확인하는 데 있어 저빈도 어휘 역시 중요하게 고려되어야 한다. 이러한 결과는 빈도를 기준으로 한 어휘 목록보다 키워드 분석을 바탕으로 한 어휘 목록이 전문분야 통역에 필요한 어휘 니즈를 보다 더 잘 충족시킬 수 있음을 방증한다.

#### 4.2. IT 전문 담화 통역에 필요한 어휘

앞의 3.2.에서 기술한 절차를 거쳐 최종적으로 총 275개 항목으로 구성된 IT 전문 담화 어휘 목록이 구축되었다(부록 1). 품사 중심으로 분류한 결과 어휘 목록의 구성은 명사 155개, 축약어<sup>9)</sup> 45개, 동사 35개, 고유명사<sup>10)</sup> 33개, 형

9) 축약어는 약어(abbreviation)와 두문자어(acronym)를 모두 합친 유형이다. 일반 명사나 고유명사와는 달리 축약어를 이해하기 위해서는 축약 전의 원래 표현이나 어휘에 대

용사 6개, 부사 1개로 집계되었다. 그림1은 전체 어휘 목록 중 각 품사의 구성 비중을 보여준다.



<그림 1> IT 전문 담화 어휘 목록의 품사 구성(단위:%)

그림 1에서 볼 수 있듯이 IT 전문 담화의 키워드 분석 결과에서 IT 분야와 특별한 관계가 없는 일반 어휘와 대명사형을 제거한 후 남은 핵심 어휘의 구성에 있어서 고유명사를 제외한 일반 명사가 절반 이상(56.4%)을 차지하는 것을 볼 수 있었다. 축약어 역시 약 16%를 차지하고 있어 IT 담화 이해에 필요한 분야 특수적 어휘 중 상당 부분을 차지하였다. 앞서 IT 전문 담화의 어휘적 특성을 분석한 결과 평균단어길이가 일반 코퍼스나 공학 코퍼스에 비해 짧게 나타났는데, 이것이 IT 코퍼스의 축약어 비중이 상대적으로 높았기 때문이라는 주장을 제기할 수 있다. 그러나 이를 확인하기 위해서는 추후 다른 분야의 전문 담화 코퍼스의 축약어 비중과 비교가 이루어져야 할 필요가 있다. 이어서 동사가 12.7%, 고유명사 12%, 형용사와 부사가 2.5%를 차지하였다. 키워드 분석이 일반 언어사용(참조 코퍼스) 대비 전문 담화에서 빈도가 두드러지는 어휘를 도출하는 방법임을 감안할 때, 여타 품사 대비 현격히 미미한 비중을 차지한 형용사와 부사는 IT 전문 담화에서도 사실상 일반적 담화와 큰 차이가 없는 부분이라 볼 수 있다.

한 지식이 요구된다는 점에서 별도의 유형으로 취급하였다.  
10) IT 관련 제품과 기술의 명칭을 포함한다.

IT 코퍼스의 키워드 분석을 바탕으로 어휘 목록(부록1)을 도출함으로써 IT 전문 담화를 구성하는 주요 어휘를 확인할 수 있었다. 이러한 어휘 목록은 IT 전문 담화의 전반적인 어휘적 특징을 보여준다는 점에서 그 자체로서 유용성을 지닌다. 그러나 일련의 어휘가 제시된다고 해도 해당 어휘가 사전적으로 어떤 의미인지를 아는 정도의 어휘 지식으로는 전문 담화의 통역에 크게 부족하며, IT 전문 담화의 통역이라는 관점에서 이러한 어휘들의 어떠한 측면이 중요한지에 대해 보다 구체적인 논의가 필요하다. 특히 통역 준비에 들일 수 있는 시간적 여유가 만성적으로 부족한 상황을 고려할 때 전문 담화의 통역을 준비하는 통역사, 혹은 통역 학습자가 효율적으로 통역을 준비하기 위해서는 각각의 어휘와 관련된 지식을 보다 전략적으로 습득할 필요가 있다. 실질적으로 시간적, 인지적 제약이 존재하는 상황에서 어휘의 전략적인 습득이란 최소의 노력을 들여 최대의 효과를 내는 방법에 집중한다는 것을 의미한다.

전문 담화의 성공적인 통역에 필요한 어휘의 전략적 습득은 어휘 지식의 폭과 깊이라는 관점에서 접근해 볼 수 있다. 어휘 지식의 폭이란 ‘알고 있는 어휘의 수’를 의미하며, 어휘 지식의 깊이는 ‘어휘를 얼마나 잘 알고 있는지’를 의미한다(Szudarski 2018: 39). 먼저 어휘 지식의 폭이라는 관점에서 얘기한다면 주어진 어휘 목록에서 가장 많은 부분을 차지하는 유형의 어휘를 학습하는 것이 전문 담화의 통역을 준비하는 데 가장 효과적일 것이다. 본 연구에서 도출한 IT 전문 담화 어휘 목록 중 절대 다수를 차지하는 유형은 명사와 축약어로써 전체 어휘의 약 73%를 차지한다. 이러한 어휘들은 철(2009)이 전문 담화 통역에 필요한 요소로 꼽은 유형 중 NE(명사적 객체)에 해당한다. 다시 말하면 전체 담화를 포괄하는 영역(coverage)을 기준으로 볼 때 IT 전문 담화의 통역을 위해서는 해당 담화를 구성하는 명사와 축약어 등과 같은 NE를 우선적으로 습득할 필요가 있다.

그런데 어휘를 ‘안다’라고 하는 것은 어휘의 형태(소리, 발음, 철자, 형태소 구성 등), 의미(형태와 의미의 결합, 개념, 지시대상 등), 사용(문법적 기능, 언어, 활용 제약 등) 등 다차원적으로 이루어지기 때문에(Szudarski 2018: 37) 특정 어휘를 통역하는 데 요구되는 어휘 지식의 깊이라는 측면에서 어휘의 유형 간 차이가 있을 것으로 판단된다. 이어지는 항에서는 통역에 요구되는 어휘 지식의 깊이를 기준으로 어휘 유형을 구분하여 IT 전문 담화의 통역 시 어휘의

어떠한 측면이 중요한지에 대해서 사례 중심으로 논의하고자 한다. 축약어와 고유명사는 통역에 요구되는 어휘 지식의 깊이가 얇은 유형, 그리고 명사와 동사는 요구되는 어휘 지식의 깊이가 깊은 유형으로 나누어 고찰하고자 한다.

#### 4.2.1. 축약어와 고유명사

축약어와 고유명사는 지칭하는 대상이나 객체에 일대일로 대응되며 그 지칭하는 것 자체가 해당 어휘의 의미를 이루고 어휘의 활용 제약이 적다는 면에서 요구되는 어휘 지식의 깊이가 상대적으로 얇은 유형의 어휘라고 할 수 있다. 대부분의 축약어나 고유명사는 인터넷 검색 한 번 만에 사실상 통역에 필요한 대부분의 어휘 지식을 쉽게 확인할 수 있다. 간단한 인터넷 검색이나 사전 검색을 통해 해당어가 지칭하는 대상이나 객체가 무엇인지에 대한 지식만 확인하면 통역이 가능하다. 예를 들어 IT 전문 담화 어휘 목록(부록1) 중 일부 고유명사, 그리고 해당 고유명사로 인터넷 검색 시<sup>11)</sup> 첫 화면에 제시되는 설명의 첫 문장을 제시하면 다음과 같다.

AZURE: 마이크로소프트 애저는 2010년 시작된 마이크로소프트의 클라우드 컴퓨팅 플랫폼...

KUBERNETES: 쿠버네티스는 디플로이 자동화, 스케일링, 컨테이너화된 애플리케이션의 관리를 위한 오픈 소스 시스템...

LINUX: 리눅스는 리누스 토르발스가 커뮤니티 주체로 개발한 컴퓨터 운영 체제...

MICROSOFT: 마이크로소프트 코퍼레이션는 미국의 세계 최대의 다국적 소프트웨어 및 하드웨어 기업...

SHAREPOINT: 셰어포인트는 마이크로소프트 오피스 서버 제품군의 웹 애플리케이션 플랫폼...

위의 문장에 포함된 개념이나 내용을 모두 이해하지 못하더라도 ‘AZURE는 플랫폼이다’, ‘KUBERNETES는 오픈 소스 시스템이다’, ‘LINUX는 컴퓨터 운영 체제이다’, ‘MICROSOFT는 기업이다’, ‘SHAREPOINT는 플랫폼이다’라는 지식 정도면 통역에 큰 어려움을 야기하지 않는다.

11) 2019년 7월 현재 구글(www.google.com) 검색을 활용하였다.

그런데 여기서 유념해야 할 부분은 통역이 구어로 발화된 언어를 듣고 이를 구어로 재표현하는 방식이라는 점이다. 축약어나 고유명사의 경우 조어 방식이나 특이한 철자법 등으로 인해 발음이나 소리 정보가 일반적인 규칙을 따르지 않는 경우가 있다. 따라서 ‘소리’에 대한 지식이 없으면 통역에 문제가 발생할 수 있다. 일례로 ‘IAAS’라는 약어의 경우 각 철자를 하나하나 따로 발음하여 ‘아이 에이 에이 에스’라고도 하고 두문자어로 취급하여 ‘이아스’라고 발음하기도 한다. 이러한 소리에 대한 지식이 없다면 ‘IAAS’에 대해 아무리 많은 개념적인 지식을 갖추고 있다고 하더라도 형태(어휘의 소리)와 의미를 연결시킬 수 없기 때문에 이해 단계에서 어려움이 초래될 수 있다. <표 4>에서 이러한 예를 더 제시하였다.

<표 4> IT 전문 담화 어휘 목록에 포함된 축약어(예시)

축약어	발음	축약 전 원래 표현	의미
MNIST	[엠니스트]	Modified National Institute of Standards and Technology database	손으로 쓴 숫자들로 이루어진 대형 데이터베이스 <sup>12)</sup>
SECOPS	[세캡쓰]	Security Operations	IT 보안팀과 운영팀 간의 협업. 양 팀의 도구, 프로세스, 기술을 통합하여 조직의 보안을 유지하면서 리스트 저감과 비즈니스 민첩성 향상이라는 공동의 목표를 달성하고자 하는 노력. <sup>13)</sup>
SQL	[씨퀄] 혹은 [에스큐엘]	Structured Query Language	관계형 데이터베이스 관리 시스템(RDBMS)의 데이터를 관리하기 위해 설계된 특수 목적의 프로그래밍 언어 <sup>14)</sup>
YAML	[얘믈]	YAML Ain't Markup Language	‘사람이 쉽게 읽을 수 있는’ 데이터 직렬화 양식 <sup>15)</sup>

12) 위키백과 참조. <https://ko.wikipedia.org/wiki/MNIST>

13) <https://www.carbonblack.com/resources/definitions/what-is-secops/>

14) 위키백과 참조. <https://ko.wikipedia.org/wiki/SQL>

‘MNIST’의 경우 자음 2개가 연속되어 철자별로 따로 읽는 약어(abbreviation)인지 아니면 한 단어처럼 읽는 두문자어(acronym)인지 불분명하나 ‘NIST’가 마치 하나의 단어처럼 읽힐 수 있는 부분이므로 소리를 예측하기가 크게 어렵지는 않다. 그러나 두 번째 축약어인 ‘SECOPS’를 문서로 처음 접하는 경우 발음을 추정하기 쉽지 않다. 일례로 IT 전문지 기사<sup>16)</sup>에서 ‘세크옵스’라는 표현을 보고 이 용어는 영어를 그대로 음차한다는 점을 알게 되었다 하더라도 실제 화자가 어떻게 발음하는지 모르는 상태에서 연사 발화 내에서 처음 접한다면 단 번에 알아듣고 통역하기 극히 어렵다. 소리 정보가 없을 때 이것이 통역사가 모르는 일반 명사인지 고유명사인지조차 구분하기 어렵기 때문이다. ‘SQL’은 약어로 읽는 방법도 있으나 모음이 없는 상태에서도 하나의 단어처럼 ([씨퀄]) 읽히기도 한다. 따라서 이 두 가지 발음에 대한 지식을 통역사가 미리 보유하고 있어야 청각적 인풋에 포함된 ‘SQL’을 정확하게 포착할 수 있다. 마지막으로 ‘YAML’은 하나의 단어처럼 발음되는 두문자어로서 분야 전문가가 아닌 통역사의 입장에서 사전에 형태, 즉 발음에 대한 지식이 없을 경우 알아듣지 못해 놓치기 쉽다. 다시 말해 해당 어휘 항목을 눈으로 읽어 알고 있다 해도 그것이 청각적으로 전달될 때 인식이 어려운 경우가 발생하는 것이다.

이처럼 축약어와 고유명사는 의미라는 측면에서만 보면 통역에 요구되는 어휘 지식이 알지만 어휘의 구어적 형태, 즉 소리 정보 면에서는 직관이나 규칙을 넘어서는 어휘 지식이 요구되는 경우가 많다. 따라서 단순 사전 검색이나 일반 문서로 된 참고자료만으로는 성공적인 통역에 요구되는 어휘 지식을 습득하기에 불충분하며, 동영상이나 음성 자료 등을 참조하여 해당 어휘가 실제 화자들에 의해 어떻게 발음되는지 확인할 필요가 있다.

#### 4.2.2. IT 분야 특수적 의미로 사용되는 명사와 동사

IT 전문 담화 어휘 목록의 약 2/3를 차지하는 명사와 동사 중에는 IT라는 전문분야 이외의 분야에서도 일반적인 의미로 사용되는 어휘가 포함되어 있다. 예

15) 위키백과 참조. <https://ko.wikipedia.org/wiki/YAML>

16) <클라우드 서비스 업체가 알려주지 않는 보안 기법 2가지>, IT World, 2019년 3월 18일자.

를 들어 ‘connect(연결하다)’, ‘delete(삭제하다)’, ‘manage(관리하다)’, ‘location(위치)’, ‘scenario(시나리오)’, ‘user(사용자)’ 등의 어휘는 IT 전문 담화에 상대적 으로 자주 등장하는 일반 어휘라고 볼 수 있다.

이와는 달리 IT 전문 담화에 쓰일 때 문맥에 따라 특수한 의미를 취하거나 일반적인 의미와 다른 의미로 사용되는 어휘들이 있다. 일반 어휘로 보이면서도 IT라는 전문분야에서 사용될 때 특수한, 혹은 일반적 의미와는 다른 의미를 띠게 되는 어휘들의 경우 어휘의 형태와 의미 간 일대일 대응이 용이하지 않으므로 해당어를 통역하는데 요구되는 지식의 깊이가 상대적으로 깊게 요구된다. 이러한 어휘는 여러 가지 의미를 지니는 다의어, 혹은 다의어 중 출현 빈도가 낮은 의미로 사용된 불투명 어휘(Todd 2017)라고 할 수 있다. 예를 들어 어휘 목록에 포함된 명사 중 ‘container’의 사례를 살펴보면 다음과 같다.

(a) The first thing we're going to do is create a resource group. A resource group is going to act as maybe like a logical container of where we put all of our items. (MIG10<sup>17</sup>, CC<sup>18</sup>47)

(b) So even though you're running a container, a Linux container, it's running in its own virtual machines, so you do have that hard virtualization security boundary. (MIG40, CC84)

문장(a)와 (b)는 모두 ‘마이그레이션(migration)’이라는 주제로 이루어진 발표에서 가져온 콘코던스 라인이다. 두 문장에 사용된 ‘container’라는 단어 중 문장(a)에 쓰인 용례는 일반적인 의미인 ‘(무언가를 담는) 용기’로 해석되며, 문장(b)는 ‘컴퓨터 운영 체제 수준의 가상화 기술 혹은 그러한 기술로 만들어진 운영체제 상의 사용자 공간<sup>19)</sup>’을 의미한다. 따라서 문장(b)를 일반적 의미로 해석해서는 통역 시 재표현을 위한 이해가 적절히 이루어질 수 없다. 통역 과정에서 이는 단순히 대응어의 문제에 그치지 않는다. 일반적인 의미의 컨테이너가

17) 해당 콘코던스 라인이 포함된 파일명.  
 18) ‘CC’는 콘코던스 라인을 의미하며, 뒤에 오는 숫자는 워드스미스 검색 결과 표시된 콘코던스 라인 번호이다. 내용 이해를 돕기 위해 콘코던스 라인의 앞뒤 문장을 함께 제시하였다.  
 19) 장독대 역(2014: 13) 참조.

아님을 직감한 통역사가 ‘컨테이너’라고 음차를 해서 순간의 위기를 모면한다고 하더라도, 통역사의 머릿속에 ‘용기’에 해당하는 심상을 띄우게 된다면 앞뒤 맥락 파악에 큰 어려움이 발생할 수 있다. 우선 ‘run’이라는 동사와의 연어 관계를 파악하기 어려우며(‘용기를 run한다는 것이 어떤 의미인가?’), 뒤이어지는 ‘a Linux container’의 경우 리눅스가 대표적인 오픈 소스 운영체제라는 배경지식이 있다고 가정하더라도 리눅스라는 추상적인 기술과 구체적 사물인 ‘용기’ 혹은 ‘컨테이너’와의 관계를 파악하고자 순간적으로 추가적인 노력이 소요되며 이는 뒤이어 들어오는 입력물의 이해와 처리에 부담을 준다. 따라서 성공적인 통역을 위해서는 ‘container’의 IT 특수적 의미에 대한 지식이 갖추어져 있어야 하며, 주위 문맥에 따라 일반적 의미와 IT 특수적 의미를 구분할 수 있는 능력이 필요하다. 어휘 목록(부록1)에 포함된 ‘blob’이라는 명사도 ‘방울, 덩어리, 얼룩’ 등을 의미하는 단어와 발음은 동일하나 IT 전문 담화에서는 ‘바이너리 라지 오브젝트(binary large object)’라는 특정한 지시물을 가리키는 어휘일 가능성이 매우 높아 통역에 어려움을 야기할 수 있다. 이처럼 IT 특수적 의미로 사용된 명사의 경우 연어나 문맥 등 해당 어휘의 실제적 활용 환경이 주어진 명사의 의미를 결정짓는데 중요하게 작용한다. 위의 사례와 유사하게 IT 특수적인 의미를 취하는 일반 명사의 사례를 몇 가지 더 살펴보면 다음의 <표 5>와 같다.

<표 5> IT 특수적 의미를 취하는 명사(예시)

명사	일반적 의미 <sup>20)</sup>	IT 특수적 의미	한국어 표현
instance	1. 경우; 2. 실례, 예	객체지향 프로그램의 주요 개념 중 하나로써 개념적 정의가 하나의 실제적인 객체로 표현된 것 <sup>21)</sup>	(a) ‘예’를 든다는 일반적 표현 (b) 사례’, ‘예’가 아닌 ‘인스턴스’로 음차
			(a) If you're familiar with SQL for <u>instance</u> , think of all the things that you can do with SQL. (DAT10, CC11) (b) So if I go back into SQL Database Managed <u>Instance</u> , I will have one actually another <u>instance</u> . (MIG20, CC119)
function	1. 기능, 역할; 2. 의식, 행사	소프트웨어에서 특정 동작을 수행하는 일정 코드 부분 <sup>22)</sup>	‘기능’이 아닌 ‘함수’로 번역하거나 음차

	So I have here a list of kernels. Essentially, we take our image. I load it up, and we define this convolution <u>function</u> . (DAT50, CC21)		
production	1. 생산, 제작, 2. 제품	소프트웨어 개발 완료 후 실제 사용 환경을 지칭.	‘생산 환경’, ‘제조 환경’이 아닌 ‘실제 운영 환경’과 같이 풀어서 번역하거나 ‘프로덕션 환경’이라고 음차
	You'll use an automated test suite and other tools to ensure proper functionality before actually delivering that software to a <u>production environment</u> or your customers. (SRE10, CC1)		

지금까지 살펴본 명사 외에도 IT 전문 담화 목록상의 상당수 동사가 일반적 의미와 달리 IT 분야 특수적인 의미를 취하는 것을 확인할 수 있었다. 질 (2009)이 말하는 전문 담화의 이해에 있어 필요한 요소 중 명사적 객체(nominal entity)를 제외한 속성(attribute)과 연결고리(link)의 상당 부분이 동사로 표현될 수 있으며, 동사와 같은 서술어는 성공적인 통역을 위한 원문 이해에 있어 반드시 필요한 핵심정보(남슬기 외 2018)임을 고려할 때 전문 담화 통역 시 동사에 대한 정확한 어휘 지식은 매우 중요한 역할을 한다. 그런데 동사는 통역 상황에서 특히 더 문제가 될 수 있다. 원문의 동사에 해당하는 도착어 대응어를 찾지 못할 경우 한국어와 같이 SOV의 어순을 갖는 도착어로 통역할 때 문장의 맨 끝에 놓이는 서술어를 확정하지 못하여 문장을 완결할 수 없는 문제가 발생하기 때문이다.

다음에서는 IT 전문 담화 어휘 목록 중 일반 동사이면서 IT 특수적 의미를 취하는 ‘deploy’, ‘build’, ‘migrate’, ‘call’, ‘scale’의 사례를 살펴보고자 한다. IT 특수적인 의미를 취한다는 것은 해당 동사가 IT 맥락 속에 쓰일 때 원래 일반 어휘로서 지니는 의미를 완전히 잃지 않으면서도 보다 구체적인 어떤 과정이나 행위를 지칭하는 것을 뜻한다. 일반적 의미와 IT 특수적 의미의 차이를 살펴보

기 위해 일반 영어에서의 용례와 IT 전문 담화에서의 용례를 비교하였다. 이를 위하여 일반 영어 코퍼스인 BNC와 본 연구의 IT 코퍼스 상에서 나타나는 각 동사의 연어 관계를 조사하였다(부록2 참조).

‘deploy’의 경우 일반 영어에서는 ‘troop’, ‘missile’, ‘tactic’, ‘force’, ‘weapon’ 등과 연어 강도가 높게 나타나 주로 군사 분야에서 사용되는 용어임을 확인할 수 있었다. 이때 ‘deploy’는 한국어로 ‘병력, 미사일, 무기’ 등을 ‘배치하다’라고 번역할 수 있다. IT 전문 담화에서는 ‘resource’, ‘application’, ‘VM’, ‘workload’ 등과 자주 공기하는 것으로 나타났다. IT 분야에는 ‘신규 소프트웨어나 하드웨어를 주어진 환경에서 적절히 구동되도록 만드는 모든 과정’<sup>23)</sup>을 일컫는 독특한 의미를 취하게 되었다. 다시 말해 군사 분야에서처럼 물리적인 실체를 특정 물리적 공간에 배치한다는 기본적인 의미와 무관하지 않으나 배치의 대상이 소프트웨어나 하드웨어와 같은 IT 관련 자산으로 특정되는 경향을 보인다. 이러한 의미의 ‘deploy’는 일반적으로 한국어로는 ‘배포하다’라고 번역된다.

‘build’는 일반 영어에서는 ‘house’, ‘station’, ‘church’, ‘bridge’, ‘wall’, ‘plant’ 등 구체적인 물리적 실체로서 눈에 보이고 만질 수 있는 건축 혹은 건설의 대상을 연어로 취하는 것을 볼 수 있다(부록2 참조). 따라서 일반적인 의미로 쓰인 ‘build’는 ‘건설하다’, ‘짓다’ 등 사전 상의 대응어를 그대로 가져다 쓸 수 있다. 반면 IT 전문 담화 내에서는 ‘build’라는 동사의 대상이 되는 목적어가 ‘application’, ‘system’, ‘dashboard’, ‘pipeline’ 등으로 나타났는데, 이는 상대적으로 훨씬 추상적인 기술이나 개념을 지칭한다. 따라서 ‘짓다’, ‘건설하다’와 같은 대응어로 번역하기에 부적절하며 ‘개발하다’, ‘만들다’, ‘구축하다’ 등 연어 관계에 있는 목적어에 맞추어 번역하는 것이 바람직하다. 이러한 경우는 큰 문제가 되는 않는데, 여전히 일반 영어에서 사용되는 의미를 거의 유지하고 있어서 통역 시 원문 맥락에 따라 보다 자연스러운 대응어를 찾는 정도의 노력이 요구되기 때문이다.

그런데 IT 코퍼스에서 ‘build’와 가장 높은 연어 강도(10.86)를 보인 ‘definition’의 경우 콘코던스 라인을 확인한 결과 동사로 사용된 경우가 아닌 것

20) 다음 영어사전(<https://dic.daum.net/index.do?dic=eng>)을 참조하였으며, 상위에 기재된 주요 의미 항목을 제시하였다.

21) 김치수(2019: 296-299) 참조.

22) 위키백과([https://ko.wikipedia.org/wiki/함수\\_\(프로그래밍\)](https://ko.wikipedia.org/wiki/함수_(프로그래밍)))참조.

23) <https://whatis.techtarget.com/definition/deploy> 참조.

으로 나타났다. 일례로 아래 예문(c)에서 ‘build editor’, ‘build definition’처럼 구체적 대상이나 개념을 지칭하는 명사적인 용법으로 사용된 것을 볼 수 있었다.

(c) Yes, sorry the question is to paraphrase was, if you already have that visual build editor, that visual build definition, can you pull that into your source code? (DEV20, CC57)

IT 맥락에서 ‘build’가 명사로서 어떠한 의미를 지니는지 조사하기 위해 위키피디아 상의 정의를 참조한 결과 ‘컴퓨터 소프트웨어 분야에서 소프트웨어 빌드(software build)는 소스 코드 파일을 컴퓨터에서 실행할 수 있는 독립 소프트웨어 가공물로 변환하는 과정을 말하거나 그에 대한 결과물’<sup>24)</sup>을 지칭하는 것으로 확인되었다. 이 정의를 볼 때 ‘build’는 IT 맥락에서 위의 정의에 표현된 과정을 지칭하는 동사나 그 결과물을 지칭하는 명사로 사용될 수 있음을 알 수 있다. 위의 정의에서 드러난 것처럼 IT 맥락에서 ‘build’는 다음과 같이 ‘빌드하다(동사)’로 음차되어 사용되고 있음을 볼 수 있으며,<sup>25)</sup> 일반적 의미로 번역할 경우 IT 맥락에서 의미하는 프로그래밍 상의 특정 과정이나 결과물을 지칭하지 못하게 된다.

- 고객에게 판매되는 제품을 개발자 각자가 빌드하다 보니 바이러스에 감염된 파일이 제품에 포함되어 배포되는 사고가 발생하기도 합니다.  
- 사실, 위의 선행 단계를 하지 않고 그냥 PCBuild.sln만 빌드하다가 오류가 발생해도 상관없습니다. python.exe까지는 빌드가 된 상태이기 때문에 웬만한 것은 테스트해 볼 수 있는데요. 그래도 웬지 모든 것이 빌드 안되면 시원하지 않기 때문에 마무리를 지어보았습니다.

‘migrate’는 일반 어휘와의 의미 차이뿐만 아니라 문법적 쓰임새 면에서 IT 특수적인 양상을 보이는 것을 알 수 있었다. 일반적 용례의 언어 순위(부록2)에서 최상위를 차지한 ‘complex’는 ‘이동성 위장관 복합운동(migrating motor

24) 위키피디아([https://ko.wikipedia.org/wiki/소프트웨어\\_빌드](https://ko.wikipedia.org/wiki/소프트웨어_빌드)) 참조.  
25) SketchEngine([www.sketchengine.eu](http://www.sketchengine.eu))에서 제공하는 한국어 코퍼스 koTenTen18에서 ‘빌드하다’로 찾은 콘코던스 라인 중 일부.

complex)’라는 의학 전문용어였으며, 두 번째 연어인 ‘sub-routine’이 IT 관련 텍스트였으므로 일반 영어의 쓰임새에서 배제한다면, 언어 강도가 높은 항목은 대부분 부사(westward, outward, eastward, northward)로서 migrate가 자동사로 활용되고 있음을 알 수 있었다. 세 번째 연어인 ‘bird’ 역시 대부분의 용례가 ‘migrating bird’로서 자동사로 활용된 예였다. 이외는 대조적으로 IT 코퍼스에서 ‘Mongo’, ‘centers’, ‘database’, ‘application’ 등을 목적으로 취하는 타동사로 활용되는 차이를 보였다.

‘call’이 일반 영어에서 공기하는 연어(‘police’, ‘meeting’, ‘election 등)와는 달리 IT 코퍼스에서 가장 강도 높은 연어는 ‘function’이었다(부록2). 앞서 IT 특수적 의미로 쓰인 명사의 사례에서 본 ‘function’이 ‘함수’를 우선적으로 의미함을 고려할 때 ‘call’-‘function’은 ‘함수’-‘호출’이라는 IT 작동의 한 단계를 의미하는 용례로 자주 등장한 것으로 추측할 수 있다. ‘scale’의 경우 일반 영어에서 ‘height’, ‘fence’, ‘ladder’, ‘wall’, ‘everest’, ‘peak’, ‘mountain’ 등을 목적으로 취하는 것으로 볼 수 있었고 (‘높은 곳에) 오르다’라는 의미로 해석되었다. 그러나 IT 전문 담화에서는 ‘horizontally’, ‘instantly’, ‘down’, ‘out’, ‘automatically’, ‘up’ 등의 부사 수식어와 함께 쓰이면서 무엇인가를 ‘단계적으로 확장하다’, (‘규모를) 늘리다’라는 의미를 주된 의미로 취하였다.

지금까지 논의한 바와 같이 IT 전문 담화 어휘 목록의 일부 동사들의 언어 관계를 바탕으로 해당 동사가 IT 특수적 어휘로 사용될 때 지니는 특징을 정리하면 <표 6>과 같다.

<표 6> IT 특수적 의미를 취하는 동사(예시)

단어	IT 특수적 의미	IT 코퍼스에서 나타난 활용 특징	한국어 번역
deploy	신규 소프트웨어나 하드웨어를 주어진 환경에서 적절히 구동되도록 만드는 모든 과정	‘resource’, ‘application’, ‘VM’ 등 소프트웨어나 하드웨어 관련 대상을 목적으로 취함	배포하다
build	소프트웨어의 소스 코드를 배포 가능한 상태로 개발하는 것(혹은 그 결과물)	동사 및 명사로 활용	빌드하다(동사) 빌드를 수행하다(동사) 빌드(명사)

migrate	한 운영환경으로부터 다른 운영환경으로 옮기는 작업 <sup>26)</sup>	일반 영어에서 자동사 용법이 주되나 IT 맥락에서는 타동사 용법이 우세	마이그레이션하다 마이그레이션을 수행하다
call	컴퓨터 프로그램의 작동을 위해 함수 등을 불러오는 작업	‘함수(function)’과 높은 언어강도를 보임	호출하다
scale	사용자 수 증대에 따라 컴퓨터 시스템이 더 커지고 강력하게 대응하는 것 <sup>27)</sup>	‘horizontally’, ‘up’, ‘down’, ‘out’ 등의 수식어와 공기	확장하다

이러한 유형의 어휘에 더하여 일상적인 맥락에서는 거의 쓰이지 않고 IT 분야에서만 배타적으로 사용되는 동사도 발견되었다. 어휘 목록(부록1)에 포함된 동사 중 ‘configure’와 ‘lint’는 일반 영어에서는 용례가 미미하였던 반면 IT 전문 담화에서는 상당한 빈도로 출현하였다.

<표 7> ‘configure’와 ‘lint’의 출현 빈도 (백만 단어당, 단위: 횟수)

구분	configure	lint(동사)
BNC	2.05	0
IT 코퍼스	520.06	23.79

<표 7>에서 볼 수 있듯이 두 동사 모두 BNC로 대표되는 일반 영어에 비해 IT 전문 담화에서 현격이 많이 사용되었다. BNC의 콘코턴스 라인 확인 결과 포함된 일부 사례에서도 ‘configure’는 컴퓨터 관련 텍스트에서 등장하고 있는 것을 확인할 수 있었고, ‘lint’의 경우 동사로 활용된 경우는 전무하였다. 이러한 빈도 상의 현격한 차이를 고려할 때 이 두 동사는 충분히 IT 전문용어로 간주될 수 있을 것이다. ‘configure’는 일반 영어사전<sup>28)</sup>에서 ‘(특별한 목적에 맞추어) 설계, 개조, 수정하다’, 혹은 ‘[컴퓨터] 구성하다(필요한 하드웨어와 소프트웨어

26) <https://life-with-coding.tistory.com/68>의 설명 참조.

27) 동사로서의 scale의 정의가 제시된 사례를 찾을 수 없어 명사형 ‘scalability’의 정의를 원용하였다. 다음 사진 참조.

(<https://dic.daum.net/word/view.do?wordid=ekw000145665&q=scalability>)

28) 다음 영어사전(<https://dic.daum.net/index.do?dic=eng>) 참조.

를 상호 접속하거나 로드하여 시스템을 형성’라고 정의되어 있다. 사전적 정의에 따르면 ‘설계하다’, ‘개조하다’, ‘수정하다’, ‘구성하다’라는 복수의 한국어 번역이 가능하며 연어 관계를 이루는 목적어에 따라 적절한 대응어가 선택되어야 한다. 예를 들어 IT 코퍼스 상에서 임의로 추출한 다음의 용례 중 (d)에서는 ‘(매개변수를) 설정하다’, (e)에서는 ‘(애저 파이프라인을) 구성하다’로 번역할 수 있다. 이처럼 IT 전문용어이면서 공기하는 언어별로 번역어가 달라져야 하는 경우 순간적인 술어 선택이 요구되는 통역 과정에 추가적인 어려움을 야기할 수 있다.

(d) So one of the things that the Azure Machine Learning Service provides is, it provides the ability to do auto hyper optimization and configure those parameters for me and run multiple copies of my experiments with different range sets of hyper parameters to find the ideal value. (DAT40, CC2)

(e) So I just need to configure Azure Pipelines, this application, and then just give it access to the repo that I've created or the repo that has the code in it. ((DEV20, CC9)

다음으로 ‘lint’는 일반 어휘로는 명사로 ‘(천에 생기는) 보푸라기’라는 의미이나, IT 전문용어로는 ‘소스 코드를 분석하여 오류를 표시하는 도구’<sup>29)</sup>(명사의 용법)나 이 도구를 동작하여 분석, 오류를 확인하는 행위 자체(동사의 용법)를 의미한다. ‘lint’는 음차하여 ‘린트를 수행하다’, ‘린트검사를 수행하다’ 등으로 번역할 수 있다<sup>30)</sup>. 다음과 같은 lint의 용례 중 (f)는 ‘(나의 파이썬 프로젝트를) 린팅하다, (나의 파이썬 프로젝트에) 린트를 수행하다’, (g)는 ‘단순히 린팅을 하는 것이 아니다’, ‘단순한 린팅이 아니다’, ‘단순히 린트검사를 수행하는 정도에 그치지 않는다’와 같이 표현될 수 있다.

29) 위키백과([https://ko.wikipedia.org/wiki/린트\\_\(소프트웨어\)](https://ko.wikipedia.org/wiki/린트_(소프트웨어))) 참조.

30) 한국어 용례 참고: <Lint의 개념과 Eclipse에서의 사용법> <https://belll.tistory.com/entry/Lint-%EC%9D%98-%EA%B0%9C%EB%85%90%EA%B3%B0-Eclipse-%EC%97%90%EC%84%9C%EC%9D%98-%EC%82%AC%EC%9A%A9%EB%B2%95>.

(f) ...we're gonna see this in the demonstration so things like IntelliSense the ability to, like, I said we've already talked about the linting, you know, the ability to, like, lint my Python project and see where I might have issues as well as some other features that we'll see in the demonstration. (FUN30, CC1)

(g) ...so this is the Accra Lynx one right here I don't have it hooked up but what you can see here is that we've actually had like a little application inside a Visual Studio code now so it's not just linting things that we perform on the text files themselves...(FUN30, CC6)

## 5. 결론

일반적인 기준으로 통역사는 이중 언어 구사력이 매우 뛰어나지만 전문분야 통역 시에는 일반적인 언어 구사력으로 해결하기 어려운 부분이 많다. 해당 분야의 비전문가인 통역사가 특수한 기술적 지식이 전제되는 전문분야의 담화를 통역하기 때문이다. 해당 분야의 전문 지식을 개념부터 시작하여 총체적으로 학습할 시간적 여유가 없고 전문 지식을 완전히 흡수하는데 필요한 기술적 소양도 부족한 통역사가 전문분야의 담화에 접근하여 이를 이해하고 통역할 수 있는 과정의 시작점은 바로 전문 담화를 구성하는 어휘이다.

본 연구에서는 IT 전문분야 담화를 영어에서 한국어로 통역하는데 필수적인 어휘를 도출하고 특징에 따라 분류하여 어휘의 유형별로 어떠한 어휘 지식이 요구되는지 고찰하였다. 실제 통역이 이루어졌던 IT 관련 발표와 유사한 자료로 IT 코퍼스를 구축한 뒤 어휘적 특징을 살펴본 결과, 일반 언어 대비 어휘 다양성이 낮고 평균어휘길이가 상대적으로 짧았으며 어종 기준으로 저빈도 어휘의 비중이 높아 관련 선행 연구에서 보고한 전문분야 코퍼스와 유사한 특징을 나타내었다. 이는 특정 전문분야의 담화일 경우 더욱 한정된 수의 특수한 어휘를 사용하기 때문인 것으로 해석할 수 있으며, 전문분야 담화의 통역에 있어 상대적으로 한정된 수의 분야 특수적인 어휘를 파악하고 이를 학습할 필요성을 확인하였다.

IT 코퍼스에 대한 키워드 분석 후 정제 과정을 거쳐 도출한 총 275개의 IT 전문 담화 어휘(부록1)는 일반 명사가 절반 이상(56.4%)을 차지하였고, 축약어

(16%), 동사(12.7%), 고유명사(12%), 형용사와 부사(2.5%)가 뒤를 이었다. 어휘의 유형별로 통역에 요구되는 어휘 지식을 고찰한 결과, 축약어와 고유명사는 의미라는 측면에서만 보면 통역에 요구되는 어휘 지식이 않지만 어휘의 구어적 형태, 즉 소리 정보 면에서는 직관이나 규칙을 넘어서는 어휘 지식이 요구되었다. 명사와 동사의 경우에는 본고에서 고찰한 다양한 어휘 사례에 비추어 볼 때 일반 언어의 맥락에서 사용될 때와는 달리 IT 특수적 의미를 취하는 경우가 발견되었다. 따라서 문맥에 따라 일반적 의미와 IT 특수적 의미를 구분할 수 있는 능력이 필요하므로 통역에 요구되는 어휘 지식의 깊이가 깊다고 판단되었다. 이와 더불어 일상적인 맥락에서는 거의 쓰이지 않고 IT 분야에서만 배타적으로 사용되는 동사도 발견되었다. 이러한 명사와 동사들을 일반 언어 사용 시의 용례와 비교하여 IT 특수적인 의미를 고찰하였고, 코퍼스 상의 코컨텐츠 라인을 바탕으로 전문 담화의 맥락에서 쓰이는 명사나 동사의 경우 문법적 활용이나 주로 쓰이는 품사도 달라짐을 확인하였다.

본 연구에서 도출한 어휘 목록과 핵심 어휘 유형별 어휘 지식은 향후 IT라는 전문분야 통역의 교육과 실무에 유용한 자료를 제공할 수 있을 것이다. 현실적으로 다양한 전문분야를 모두 다루기 어려운 실무 통역사는 어휘 목록에 담긴 개별 어휘를 출발점으로 상향식 지식 습득의 과정을 보다 효율적으로 시작할 수 있다. 또한 전문분야 통역 수업의 교수학습 자료로 활용함으로써 교육 과정에 실질적이고 직접적인 도움을 기대할 수 있다. 나아가 본고의 어휘 목록을 활용하여 특정 전문분야 통역 준비도를 평가하는 검사지를 개발하는 등 평가의 기준으로 활용하는 것도 가능할 것이다.

그러나 본 연구는 다음과 같은 한계를 지닌다. 우선 본 연구의 결과는 표1의 특징을 갖는 언어 자료에만 적용될 수 있다는 점에서 일반성에 한계를 지닌다. 또한 본 연구에서는 IT 전문 담화를 특징짓는 어휘를 선정하는 데 있어 키워드 분석이라는 방식에 의존하였다. 이를 통해 도출된 어휘 목록은 통계를 바탕으로 한 객관성이 확보된 반면, 전문가의 정성적인 평가가 있었다면 반드시 추가되었을 주요 어휘들이 누락되기도 하였다. 예를 들어 본 연구에서 구축한 IT 코퍼스의 내용과 밀접한 관련이 있는 'virtualization(가상화)'이나, IT 맥락에서 동사로 활용되며 IT 특수적 의미를 지니는 'provision(프로비저닝하다)'이라는 어휘는 키워드 분석으로 검출되지 못하였다. 이는 본 연구의 한계로서

향후 전문 통역 실무와 교육의 참고자료로서 제공되는 어휘 목록을 구축하는데 있어 전문가의 주관적 평가가 결합될 필요성을 보여준다. 향후 이러한 점이 보완된다면 전문지식 면에서나 언어적인 면에서 통역의 실무 및 교육에 매우 유용한 자료가 개발될 수 있을 것으로 기대된다.

## 참고문헌

- 김창구 (2016) 「KFL 교육과정을 위한 한국어교육 학습용 참조 어휘목록과 검색 시스템 개발」, 『언어과학연구』 78:37-65.
- 김치수 (2019) 『쉽게 배우는 소프트웨어 공학(5쇄)』, 서울: 한빛아카데미(주).
- 남슬기, 성미경, 손영은, 신나정, 최수연, 한성숙, 정혜연 (2018) 「통역사가 선택한 핵심정보의 형태적 특징」, 『통역과 번역』 20(2): 73-95.
- 신창원 (2012) 「ESP 교육을 위한 소규모 코퍼스 구축: 기계공학과 관련 논문초록을 중심으로」, 『언어와 정보사회』 17: 175-205.
- 이미경 (2018) 「전문 용어의 사용이 통역의 과정에 미치는 영향에 대하여」, 『통번역교육연구』 16(3): 127-146.
- 이삼형 (2018) 『2018년 국어 기초 어휘 선정 및 어휘 등급화 연구』, 국립국어원.
- 장독대 (역) (2014). 『더 도커 북. 제임스 턴블의 The Docker Book』, 부천: 루비페이퍼.
- 장세은, 변현정 (2011) 「코퍼스를 활용한 해사영어 어휘 분석」, 『새한영어영문학회 학술발표회 논문집』 33-55.
- 최문선 (2018) 「전문통역사의 용어집 구축에 관한 예비연구」, 『통번역연구』 22(2): 179-206.
- 하명호 (2017) 「공학영어 어휘 목록 개발」, 『언어과학』 24(3): 167-185.
- Brezina, Vaclav and Dana Gablasova (2015) 'Is There a Core General Vocabulary? Introducing the New General Service List', *Applied Linguistics* 36(1): 1-22.
- Chung, Teresa M. & Paul Nation (2003) 'Technical Vocabulary in Specialised Texts', *Reading in a Foreign Language* 15(2): 103-116.
- Coxhead, Averil (2000) 'A New Academic Word List', *TESOL Quarterly* 34(2): 213-238.
- Fantinuoli, Claudio (2006) 'Specialized Corpora from the Web for Simultaneous Interpreters', in M. Baroni & S. Bernardini (eds.), *Wacky! Working Papers on the Web as Corpus*, Bologna: GEDIT, 173-190.
- Garzone, Giuliana (2000) 'Textual Analysis and Interpreting Research', *Interpreters' Newsletter* 10: 69-88.
- Gile, Daniel (2002) 'The Interpreter's Preparation for Technical Conferences: Methodological Questions in Investigating the Topic', *Conference Interpretation and Translation* 4(2): 7-27.
- Gile, Daniel (2009) *Basic Concepts and Models for Interpreter and Translator Training(Revised edition)*. Amsterdam/Philadelphia: John Benjamins.
- Gries, Stefan (2015) 'Quantitative Designs and Statistical Techniques', in D. Biber & R. Reppen (eds.), *The Cambridge Handbook of English Corpus Linguistics*, Cambridge: Cambridge University Press, 50-72.
- Jiang, Hong (2013) 'The Interpreter's Glossary in Simultaneous Interpreting: A Survey', *Interpreting* 15(1): 74-93.
- Khamis, Noorli and Imran Ho-Abdullah (2017) 'Lexical Features of Engineering English v. General English', *Journal of Language Studies* 17(3): 106-119.
- Liu, Jia and Lina Han (2015) 'A Corpus-based Environmental Academic Word List Building and Its Validity Test', *English for Specific Purposes* 39: 1-11.
- Muñoz, Verónica L. (2015) 'The Vocabulary of Agricultural Semi-popularization Articles in English: A Corpus-based Study', *English for Specific Purposes* 39: 26-44.
- Peters, Pam and Trinidad Fernández (2013) 'The Lexical Needs of ESP Students in a Professional Field', *English for Specific Purposes* 32: 236-247.
- Szudarski, Paweł (2018) *Corpus Linguistics for Vocabulary: A Guide for Research*. New York: Routledge.

Thomas, Patricia (1993) 'Choosing Headwords from Language-for-special-purposes(LSP) Collocations for Entry into a Terminology Data Bank(Term Bank), in H. B. Sonneveld & K. L. Loening (eds.), *Terminology: Applications in Interdisciplinary Communication*, Amsterdam/Philadelphia: John Benjamins, 43-68.

Todd, Richard W. (2017) 'An Opaque Engineering Word List: Which Words Should a Teacher Focus On?', *English for Specific Purposes* 45: 31-39.

Tongpoon-Patanasorn, Angkana (2018) 'Developing a Frequent Technical Words List for Finance: A Hybrid Approach', *English for Specific Purposes* 51: 45-54.

Yang, Ming-Nuam (2015) 'A Nursing Academic Word List', *English for Specific Purposes* 37: 27-38.

West, Michael (1953) *A General Service List of English Words*. London: Longman.

<부록 1> IT 전문 담화 어휘 목록

1. 축약어	DOCS	CLOUD	LATENCY	SCROLL	HOST(v.)
ACI	EXCHANGE	CLUSTER	LINK	SEARCH(n.)	INSTALL
ACR	EXPLORER	COMMAND	LOAD(n.)	SECRETS	INTEGRATE
ADMIN	GITHUB	COMPLIANT	LOCATION	SECURITY	LINT
AI	GREMLIN	CONNECTION	LOG	SEQUEL	LOAD(v.)
AKS	HYPERVISOR	CONNECTIVITY	LOGIC	SERVER	MANAGE
API	IGNITE	CONTAINER	LOGIN	SERVERLESS	MIGRATE
APP	JAVASCRIPT	CONTENT	MACHINES	SERVICE	OPTIMIZE
ASR	JS	CREDENTIALS	MAIL	SETTINGS	PASTE
ATP	JSON	CUSTOMERS	MAILBOX	SOLUTION	REFRESH
AUTH	KEYVAULT	DASHBOARD	MALWARE	STAGING	REPLICATE
AZ	KUBERNETES	DATA	MARKDOWN	STORAGE	RESTORE
CLI	KUSTO	DATABASE	METADATA	STRING	RUN
CPU	LINUX	DATABRICKS	METRICS	SUBNET	SCALE
DB	MICROSOFT	DATACENTER	MICROSERVICES	SUBSCRIPTION	SEARCH(v.)
DEMO	MONGO	DEFAULT	MILLISECONDS	SYNC	SELECT
DEV	MONGODB	DEPENDENCIES	MODEL	SYSTEMS	STORE
DEVOPS	MYSQL	DETECTION	MONITORING	TAB	TALK
DNS	NET (.NET)	DEVELOPERS	NETWORK	TAG	TRIGGER
GEO	NOSQL	DEVICES	NODE	TELEMETRY	UPLOAD(v.)
HCW	ONEDRIVE	DIRECTORY	ONLINE	TEMPLATE	VALIDATE
HTML	POSTGRESQL	DOCKER	ORGANIZATION	TERABYTES	ZOOM
HTTP	POWERPOINT	DOCUMENTATION	OUTAGE	TIER	5. 형용사/부사
IAAS	POWERSHELL	DOWNLOAD(n.)	PASSWORD	TOOLS	ACTIONABLE
IIS	PYTORCH	DOWNTIME	PAYLOAD	TRACKING	CUSTOM
IOT	SHAREPOINT	ENCRYPTION	PHISHING	TROUBLESHOOTING	HYPER
IP	SHELL	ENDPOINT	PIPELINES	UPLOAD(n.)	MANUALLY
MNIST	TENSORFLOW	ENVIRONMENT	PLATFORM	USER	SECURE
NGINX	UBUNTU	EVENT	POD	USERNAME	VIRTUAL
NPM	WINDOWS	EXCHANGE	POLICIES	VAULT	VISUAL
OAuth	3. 명사	FAILOVER	PORTAL	VERSION	
RBAC	ACCESS	FEATURE	PREM (on-prem)	WEB	
RDP	ACCOUNT	FILE	PREVIEW	WEBHOOK	
REPO	AGENT	FIREWALL	PRODUCTION	WEBSITE	
SDK	ALERTS	FOLDER	QUERY	WORKFLOW	
SECOPS	ANALYTICS	FRONTEND	RECOVERY	WORKSPACE	
SKU	APPLICATION	FUNCTION	REGISTRY	4. 동사	
SLI	ARTIFACTS	FUNCTIONALITY	RELIABILITY	ASSIGN	
SLO	AUTHENTICATION	GATEWAY	REMIEDIATION	BUILD	
SQL	AUTOMATION	GRID	REPLICAS	CALL	
SRE	AVAILABILITY	HOST(n.)	REPOSITORY	CLICK	
URL	BACKUP	HUB	REQUESTS	CODE	
VM	BALANCER	HYBRID	RESILIENCY	COMPUTE	
VPN	BLOB	INCIDENT	RESOURCE	CONFIGURE	
WDATP	BLOG	INFRASTRUCTURE	RESPONDER	CONNECT	
YAML	BLUEPRINTS	INSIGHTS	ROUTING	CREATE	
2. 고유명사	BOX	INSTANCE	RUNBOOK	CUSTOMIZE	
AZURE	BROWSER	INTERNET	RUNTIME	DELETE	
COSMOS	CANARY	INVENTORY	SCENARIOS	DEPLOY	
CRON	CAPABILITIES	LAPTOP	SCRIPT	DOWNLOAD(v.)	

<부록 2> IT 담화 내 특수한 의미를 취하는 동사의 언어(IT 코퍼스와 BNC 비교)

중심어	IT 코퍼스			BNC			중심어	IT 코퍼스			BNC		
	언어	빈도	연어강도	언어	빈도	연어강도		언어	빈도	연어강도	언어	빈도	연어강도
deploy (+목적어)	resource	11	10.49	troop	30	8.84	build (+목적어)	definition	15	10.86	house	686	9.9
	application	14	10.47	missile	12	8.55		application	14	10.23	society	203	8.74
	VM	7	10.12	tactic	9	7.94		system	13	10.23	station	162	8.52
	workload	6	10.06	force	49	7.93		dashboard	5	9.45	church	169	8.51
	gateway	5	9.91	weapon	17	7.86		pipeline	4	9.17	bridge	134	8.35
	machine	7	9.84	argument	24	7.5		query	5	9.1	wall	148	8.33
	service	9	9.77	satellite	4	7.38		model	5	9.07	plant	110	7.87
	image	7	9.76	resource	20	7.27		alert	4	8.88	road	108	7.8
	staging	4	9.71	soldier	9	7.26		context	3	8.78	factory	87	7.78
	code	5	9.35	battalion	3	7.19		artifact	3	8.76	nest	78	7.67

중심어	IT 코퍼스			BNC			중심어	IT 코퍼스			BNC		
	언어	빈도	연어강도	언어	빈도	연어강도		언어	빈도	연어강도	언어	빈도	연어강도
migrate (+목적어)	(vt)			(vi)			call (+목적어)	function	12	10.42	police	334	8.54
	Mongo	1	10.6	complex	32	10.16		Center	4	9.12	the	302	8.41
	phase	1	10.35	sub-routine	2	8.45		trader	4	9.1	meeting	193	7.64
	permission	1	10.14	bird	24	8.37		Learn	3	8.74	election	159	7.49
	center	1	10.14	westward	2	7.83		Agent	3	8.72	man	224	7.15
	VMs	1	9.95	outward	3	7.79		group	4	8.7	strike	115	7.12
	login	1	9.91	eastward	2	7.75		API	3	8.68	halt	101	6.98
	database	2	9.28	northward	2	7.72		something	6	8.56	name	144	6.96
	application	3	9.01	contraction	3	7.67		Swan	2	8.18	doctor	87	6.61
	user	1	8.96	larva	2	7.53		Review	2	8.17	girl	93	6.56
	datum	4	8.93	motor	2	6.85							
	everything	1	8.11										

중심어	IT 코퍼스			BNC			중심어	IT 코퍼스			BNC		
	언어	빈도	연어강도	언어	빈도	연어강도		언어	빈도	연어강도	언어	빈도	연어강도
scale (+수식어)	horizontally	1	10.3	appropriately	2	6.35	scale (+목적어)	side	1	11.3	height	20	8.98
	instantly	4	10.24	suitably	2	6.21		App	1	10.24	fence	10	8.76
	well	4	10.05	substantially	2	5.13		app	1	9.75	ladder	7	8.53
	down	3	9.57	down	4	3.37		number	1	8.93	fee	15	7.84
	easily	1	9.16	first	3	2.8		model	1	8.35	wall	21	7.72
	out	1	8.65	back	8	2.44		thing	1	5.83	everest	2	7.61
	automatically	1	8.3	simply	2	2.18					peak	4	7.01
	also	5	7.89	then	2	0.17					drawing	3	6.35
	up	1	7.73	also	4	0.11					palace	2	6.21
	actually	2	6.6								mountain	2	6.09
	so	2	6.14										
	really	1	5.51										

빈도: 중심어와 언어의 공기 빈도  
연어강도: LogDice 통계값

[Abstract]

**Lexical Knowledge for Specialized Interpretation:  
Key Words for Interpretation of IT Discourse**

Choi, MoonSun  
(Ewha Womans University)

This study aims to examine the lexical knowledge required for an interpreter to perform specialized interpretation in the field of information technology (IT). A study corpus (“IT Corpus”) was created with transcripts of presentations made at a major IT conference and keyword analysis was performed to develop a word list containing core vocabulary reflective of the lexical features of the corpus. It was found that IT Corpus was characterized by lower lexical variety, shorter word length and higher proportion of low frequency words relative to a general corpus. A closer examination of the word list revealed that abbreviations/acronyms and proper names represented a group of core IT vocabulary that required relatively shallow lexical knowledge for an interpreter to process, but their sound information was deemed critical during comprehension. It was also found that certain nouns and verbs on the word list showed IT-specific meanings and grammatical usages, which were different from the ones usually found in general language discourse. The specific examples of such words were discussed with their concordance lines and collocations extracted from IT Corpus compared with those from a general corpus (BNC).

- ▶ Key Words: specialized interpretation, lexical knowledge, word list, keyword analysis, words with domain specific meanings
- ▶ 주제어: 전문분야 통역, 어휘 지식, 어휘 목록, 키워드 분석, 분야 특수적 어휘

최문선

이화여자대학교 통역번역대학원 통역학과 조교수

choimoonsun@ewha.ac.kr

관심분야: 통역교육, 통역전략, 통번역연구방법론, 통번역과 기술

논문투고일: 2019년 8월 3일

심사완료일: 2019년 8월 28일

게재확정일: 2019년 9월 16일