

번역 교육을 위한 대안사전 개발의 방법론적 고찰*

최병진(목포대)

1. 이끄는 말

최근 들어 일반적으로 사전에 대한 관심이 높아지면서 사전편찬학에 대한 많은 연구가 이루어지고 있는 것은 매우 고무적인 현상이다. 사전편찬학은 유럽에서는 어휘학과 더불어 이미 오래 전부터 학문적인 이론과 체계를 갖추고 있지만 우리나라에서는 아직 20년도 채 안되는 짧은 기간 동안 연구가 이루어지다 보니 자연스럽게 외국의 이론을 수용하게 되었다. 그 수용과정 또한 국내의 언어학적인 필요성에서보다는 인접학문에 의해서 주도되었으며, 특히 자연 언어처리 분야와 관련하여 사전의 활용과 이에 대한 외국의 연구 동향을 소개하고 그와 유사한 전자사전의 모델을 개발하고 있는 실정이다(김민수 외 1995, 문유진 1996, 이재윤 & 김태수 1999, 최기선 외 2002, 2003).

이러한 사실은 아마도 인터넷의 급진적인 발달과 보급에 발맞추어 국제화와 정보화가 빠른 속도로 이루어지면서 컴퓨터를 이용하여 여러 나라의 다양한 문서를 처리할 수 있는 시스템의 개발이나, 또는 자동번역시스템의 개발이라는 필연적인 시대적 조류에 부응한 것이라고 할 수 있다.

그러나 자동번역시스템이 실질적으로 활용되기 위해서는 아직도 많은 부분들이 수정되고 보완되어야 할 필요가 있다. 특히 컴퓨터를 이용하여 자동번역을 하기 위해서는 실제로 언어학적인 지식과 전산학적인 방법론 외에도 또 다른 지식이 필요하다. 즉 문화권이 서로 다른 언어를 실제로 번역하는 과정에서 출발언어 Ausgangssprache와 목표언어 Zielsprache 사이에서 본래의 뜻이 보존된 상태로 번역을 하기 위해서는 인지적인 개념구조 체계에 기반을 둔 지식이 필요하다. 이러한 사실은 실제로 인간이 번역을 하는 과정 속에서도

* 이 연구는 한국과학재단 특정기초연구(과제번호: R01-2003-000-11618-0)지원으로 수행되었음.

이 논문을 심사해 주시고 적절한 지적과 유익한 조언을 해 주신 심사위원님들께 감사드립니다.

종중 직면하는 문제이다.

개념구조 체계에 기반을 둔 사전을 구축할 수 있다면 이것을 번역 작업에도 유용하게 사용할 수 있을 것이며, 또한 이를 자연언어처리에 응용할 수 있는 형태로 바꾸어 놓으면 기계번역에도 활용할 수 있을 것이다.

이와 같이 개념체계에 기반을 둔 사전은 매우 유용하게 사용될 수 있지만, 이런 사전을 구축하기 위해서는 인간이 직접 수동으로 작업을 해야 하므로 많은 시간적, 경제적 비용을 요구한다. 때문에 아직까지도 이에 대한 충분한 연구가 이루어지지 못하고 있다.

설령 연구가 이루어졌다 해도 주로 국어나 영어에 편중되어 있으며,¹⁾ 독일어와 관련하여서는 충분한 연구가 이루어지지 못한 실정이다.²⁾ 이러한 점을 고려해 볼 때 독일어 분야에 있어서도 사전편찬학과 관련된 연구가 독한사전이나 한독사전을 중심으로 이루어질 필요가 있다. 이러한 이유로 이 논문에서는 번역이나 통역, 외국어 교육, 나아가 자연언어처리에 사용될 수 있는 개념 기반의 독한사전이나 한독사전을 개발하기 위하여 워드넷을 활용한 새로운 형태의 대안사전을 구축하는 방법론을 제안하고자 한다.

2. 본론

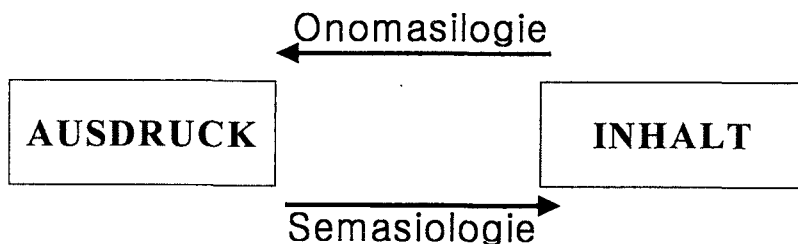
2.1. 개념사전으로서의 워드넷

사전과 관련된 의미와 개념의 문제는 전통적인 언어학적 관점에서 어의론 Semasiologie과 명칭론 Onomasiologie에 바탕을 두고 있다.

1) 국내에서는 프린스턴 Princeton 대학의 워드넷 WordNet을 기반으로 한영사전과 국어사전을 이용하여 한국어 명사의 개념체계를 자동으로 구축함으로써, 이미 구축된 다른 언어의 개념체계를 이용하여 새로운 언어의 개념체계를 자동으로 구축할 수 있는 연구들이 이루어졌다.

2) 독일어 분야에서 워드넷과 관련된 연구로는 이민행(1999, 2004)과 오장근(2002)을 들 수 있다.

표현에서 출발하여 의미를 연구하는 것이 어의론이라고 한다면, 의미 또는 개념에서 출발하여 그에 상응하는 표현을 연구하는 것이 명칭론이다.



지금까지의 의미론 연구는 주로 어의론적인 관점에서 이루어졌으며, 사전 편찬의 방법 역시 어의론적인 시각이 주류를 이루었다. 어의론적인 방법론은 단일어의 의미를 기술하는 데 가장 적합하다고 할 수 있으며, 한 언어의 의미 관계를 기술하는 데 유리하다고 할 수 있다. 그러나 하나의 표현이 다의성을 지닐 때는 사실상 분석과 기술이 용이하지 않은 점도 있다. 이에 반해 명칭론적인 연구방법은 전통 언어학에서는 큰 주목을 끌지 못했으나, 최근 들어서 정보검색이나 자동번역과 같은 자연언어처리시스템의 개발과 관련하여 워드넷이나 시소러스(Thesaurus³⁾)와 같은 다양한 연구결과가 나오고 있다.

워드넷(Fellbaum 1998)이란 어휘들의 의미 관계를 망으로 체계화한 것이다. 본래 워드넷이란 인간의 어휘지식에 대한 심리언어학 연구의 성과를 토대로 1985년부터 미국 프린스턴 대학 인지과학 연구소의 조지 밀러 G. Miller 교수를 중심으로 구축해 온 영어어휘 데이터베이스를 말한다.⁴⁾ 워드넷은 단순하게는 관련 있는 어휘들을 분류해서 동의어 정보를 제공하는 수준의 것도 있지만 대개는 개념간의 여러 가지 관계 및 체계적인 의미풀이를 포함함으로써 다양한 자연언어처리시스템의 지식 베이스로 이용된다.⁵⁾

3) 시소러스라는 용어는 국내에서 정착되어 사용되고 있는 전문용어이다.

4) <http://www.cogsci.princeton.edu/~wn/> 참조.

워드넷에 대한 이론적인 개념 소개는 본 학회지의 남유선(2004)을 참조하기 바람.

5) 워드넷은 인간의 어휘 기억을 모델화한 온라인 어휘 참조 시스템이라고 할 수 있는데, 영어의 명사, 동사, 형용사, 부사를 동의어 집단으로 구성함으로써 어휘 개념

인간의 어휘부 Lexikon에 대한 심리적 모델인 워드넷의 기본 단위는 어휘이다. 자연어의 어휘는 다의성을 띄고 있으며 다른 어휘들의 의미와 일정한 관계 Relation를 지니고 있다.

워드넷에서의 어휘는 개념을 표상한다. 일반적으로 개념은 대상, 사건, 상태, 상황, 관념 등의 복합적인 구조를 지닌 지식을 표상한다. 그래서 언어처리를 연구하는 언어공학이나 인공지능 분야에서는 이 개념들로 구성되는 지식의 의미구조를 연구하는데, 그 중의 하나가 의미망 *semantisches Netzwerk*이다. 의미망은 개념들의 의미관계를 표상하는 모델이다. 즉 개념 단위를 표상하는 노드 *Knoten*와 이 개념 단위들의 관계를 표상하는 경로 *Pfad*로 구성되는 의미망에 이 개념 구조를 표상한다. 그리하여 이러한 개념들을 분류하고 계층화하기 위해 어떤 개념과 범주들이 있는 지, 그 속성은 무엇인지, 이들의 관계가 무엇인지를 주로 연구하여 이를 처리 가능한 데이터로 구축한다. 이것이 바로 존재론 *Ontologie*이며, 이러한 존재론적 분류체계를 바탕으로 어휘를 체계화한 것이 시소러스이다. 시소러스에서는 어휘들이 의미범주에 따라 분류된다. 존재론적 의미분류체계를 기본으로 한 대표적인 것으로 *Roget*의 시소러스가 있다.

자연언어처리 분야에서 사용하는 의미는 바로 이러한 의미분류체계에 기반한 개념을 가리키며, 개념의 시소러스나 워드넷을 중심으로 연구가 이루어진다.

이 논문에서는 자연언어처리에서 주로 연구되는 워드넷을 수용하여 컴퓨터에 의해서 뿐만 아니라 인간에 의한 언어처리과정, 즉 통역이나 번역, 그리고 외국어 학습에 이용할 수 있는 대안사전으로 워드넷을 발전시킬 수 있는가 하는 가능성을 살펴보고자 한다.

자연언어처리용 워드넷이 실질적인 통번역 작업에 활용될 수 있다면 무엇보다도 다양하고 실용적인 한국어 자료를 독일어로 번역하는 데 큰 기여를 할 수 있을 것이며, 번역이나 인지적 외국어 학습방법에도 크게 공헌하게 될

을 표현하였다. 동의어 집단은 신셋 *synset*이라 불리며 워드넷 구성의 기본이 되고, 각각의 신셋은 전체 워드넷 상에서 특정한 계층에 놓이게 된다. 이러한 워드넷은 특히 정보검색, 기계번역 등과 같이 어휘의 의미를 다루는 데 있어서 지식 베이스로서 매우 중요한 역할을 한다. 워드넷의 주된 특징은 어휘형이 아닌 어휘의 의미를 구성요소로 하였다는 점이다. 워드넷은 현재 계속해서 연구되고 있으며 다국어론을 대상으로 하는 유로워드넷 *EuroWordNet*(<http://www.let.uva.nl/~ewn>)도 개발되고 있다.

것이다. 나아가 자동번역에 필요한 어휘지식과 구문 패턴들을 중심으로 자연 언어 처리의 실질적인 면에서도 도움이 될 것이다.

2.2. 번역을 위한 대안사전의 필요성

외국어를 번역할 때 출발 텍스트에 나타난 의미 단위를 목표 텍스트에서 적절히 표현하기 위해서 대부분 외국어 사전을 참고하지만, 대역어에 대한 전형적인 의미를 중심으로 사전 정보가 주어지기 때문에 적절한 도움을 받지 못하는 경우가 많다. 앞서서도 언급했듯이 사회·문화적인 관점에서 뿐만 아니라 언어학적인 관점에서 구조적으로 또는 유형적으로 서로 다른 언어를 번역할 때는 출발 언어의 고유한 특성을 유지하면서도 동시에 그 어휘나 문구에 적합한 문장으로 번역을 하기 위해서는 서로 상이한 언어구조와 인지적인 개념구조의 체계에 기반을 둔 어휘 의미적 특성들이 어떻게 대응되는 지에 대한 지식이 필요하다.

기존의 사전들은 단지 어휘의 의미에서 출발하고 있으므로, 문맥에 따라 매우 큰 융통성을 발휘해야 하는 통역과 번역의 문제를 해결하는 데는 한계를 보인다. 외국어를 공부하는 학생들이 자연스러운 번역과 통역을 하지 못하는 경우를 자주 접하게 되는데, 이는 학생들이 기존의 사전에서 얻은 지식과 학습태도에 적응되어 있기 때문이다. 학생들이 자연스러운 목표 텍스트를 생산하도록 하기 위해서는 지금까지 외국어 사전에서 터득한 등가 관계⁶⁾, 나아가서는 그러한 등가 관계를 토대로 익힌 외국어 텍스트 해석방식을 탈피하는 것이 통번역에서 가장 중요한 첫 번째 과제이다.

특히 이중언어사전의 경우 개별 어휘 항목의 의미기술을 여러 사전에서 대조해 보면, 해당 역어의 의미기술에도 많은 차이점이 있는 것을 발견할 수 있

6) 특정 문화 속에서 사용되는 언어를 다른 문화의 언어로 번역할 때 등가 개념을 찾는 일이 쉬운 것은 아니며, 더욱이 어족이나 구조가 다른 독일어와 한국어 사이에서 완전히 일치하는 등가 개념은 찾아보기가 힘들다. 다음은 출발어인 A언어와 목표어인 B언어 사이에서의 개념의 일치 가능성에 대한 구분이다(안인경 1995 참조).

- 1) A개념과 B개념이 완전히 일치(A와 B가 등가 개념)
- 2) A개념이나 B개념 중 하나가 없다.
- 3) A개념과 B개념이 부분적으로 일치.

다. 또한 이중언어사전에 나와 있는 표현 역시 그 어휘의 기본적인 의미만을 나열하고 있을 뿐 실제 생활에 쓰이는 표현이나 상황에 따라서 적절하고 구체적인, 그리고 보다 세련된 표현이 취급되지 않은 경우가 많다.

전문 통번역사는 외국어에 능통하기 때문에 사전의 도움 없이 통역이나 번역을 쉽게 할 수 있으리라는 사고는 편견일 것이다. 통역사와 번역사들은 통역이나 번역 업무를 수행할 때 출발 텍스트 표현들을 목표 텍스트 표현들로 나타내기 위하여 많은 노력을 해야 한다. 따라서 이러한 노력의 결과물들을 제대로 정리하고 분류하여 체계화시킨다면 적어도 다음 작업 때 같은 수고를 반복하지 않아도 될 것이다. 이와 같은 작업의 반복은 작가는 통번역사들 개인들의 시간과 노력의 낭비일 뿐 아니라, 크게는 국가적인 낭비라고 하지 않을 수 없다. 특히 기업체나 단체, 정부 기관 등에서 필요로 하는 통역과 번역이 대부분 짧은 시간 안에 이루어져야 한다는 상황을 감안하면, 이러한 노력을 감소시킬 수 있는 대안사전의 출현은 통번역사들의 불필요한 수고를 덜어줄 뿐 아니라, 좀더 유연하고 정확하며 적절한 문체를 만드는 데도 크게 기여할 것이다.

또한 외국어 학습에서도 학습자가 일정 수준에 도달한 이후부터는 모국어와 외국어의 낱말들이 1:1로만 대응되는 것이 아니다. 자신의 생각을 외국어로 표현하는 데는 모국어에서와 마찬가지로 의미적, 통사적, 형태적 측면에서 여러 가지 가능성이 있다는 것을 알려주고, 또 그러한 가능성들을 보여줄 수 있어야 효과적이고 수준 높은 학습을 기대할 수 있을 것이다.

이러한 일련의 목적을 충족시킬 수 있는 “대안사전”을 만들기 위해서는 여러 가지 언어적, 화용적 문맥들을 보여주는 텍스트들을 실제로 관찰하여, 목표 텍스트와 출발 텍스트를 비교하여 동일한 표현들을 찾아내고 정리하는 작업이 선행되어야 한다.

이를 위해서는 실제 자료를 바탕으로 광범위한 자료를 추출해야 하며, 다시 자료들을 정선하여 새로운 표현들을 수록하여야 한다.

전문 번역가들은 오랫동안의 경험과 한국어에 대한 풍부한 어휘력을 바탕으로 나름대로 전문기술과 어휘를 축적해 놓고 있다. 이런 어휘들은 기존 사전의 기본적으로 규정된 역어와는 달리, 상황과 문맥에 따라 적절하게 나타낸 표현들이다. 이러한 전문 번역가들의 도움을 받아 기존의 독한사전을 보충하

면서 대안사전을 구축할 수 있다.

이 논문에서 선택한 방법은 전문 번역가에 의해 번역된 독일어와 한국어의 병렬 코퍼스 Korpus를 구축하고, 거기에서 일단 한 언어의 특정 어휘를 선별하여 해당 어휘가 출현하는 문장들을 모은 뒤, 한 어휘가 다른 언어에서는 어떻게 번역이 되는지를 살펴 이를 사전 구축에 이용하는 것이다.

병렬 코퍼스는 2개국어로 구축할 수도 있고, 다국어로 구축할 수도 있다. 2개국어의 병렬 코퍼스에서는 개별 텍스트가 상호 번역된 것으로 이해할 수 있다.

자연언어처리분야에서는 병렬 코퍼스가 자동번역을 위해 다양하게 구축되고 있다. 그러나 우리나라에서는 아직 독일어와 한국어에 대한 병렬 코퍼스는 구축되어 있지 않다. 이 논문에서는 독일어-한국어 병렬 코퍼스를 구축하고, 이를 교육적으로 활용할 수 있고 또한 자연언어처리에도 활용할 수 있는 대안사전으로서의 워드넷을 구축하려고 한다.

Neues Deutschland; Was weiß Donald Rumsfeld von Korea ?

/ 립즈펠드는 한반도에 대해 무엇을 알고 있나?

Debatte über Gerüchte, Tatsachen und Tendenzen

! 한반도를 둘러싼 소문과 진상, 추세에 대한 토론

Ein halbes Jahrhundert nach dem Ende des opferreichen Korea-Krieges, der am 25. Juni 1950 ausbrach, ist die koreanische Halbinsel immer noch ein Spannungsherd.

! 1950년 6월 25일 발발해 많은 인명 피해를 야기시켰던 한국전쟁이 종전된 지 벌써 반세기가 지났지만 한반도는 여전히 긴장 상태에 있다.

Am Mittwoch widmete sich ein internationales Symposium in Berlin der Frage, wohin Nordkorea - die KDVR - steuert.

! 6.25(수) 독일 베를린에서는 “북한, 어디로 향하고 있나”라는 주제로 한 국제 심포지움이 열렸다.

Rund 130 Wissenschaftler, Politiker und Korea-Interessierte unterschiedlicher Profession waren der Einladung des Korea-Verbandes gefolgt, um über soziale Verhältnisse, Entwicklungstendenzen und Perspektiven der Koreanischen Demokratischen Volksrepublik zu debattieren.

! 독일 코리아협회 주최로 개최된 이번 국제 심포지움에는 130여명의 학자와 정치인, 한반도 전문가들과 여러 교수들이 참석하여 북한의 사회적 상황, 현재의 발전 추세 및 향후 전망에 대해 토론을 가졌다.

Leider fehlten Vertreter aus der KDVR. Wegen der SARS-Gefahr sei gegenwärtig keine Ausreise erlaubt, hieß es.

! 유감스럽게도 북한측 대표는 참석하지 않았는데, 호흡기 질병 사스(SARS)로 인해 출국이 어려웠다고 한다.

(출처: 27. 6. 2003, Koea Heute)

<병렬 코퍼스의 예>

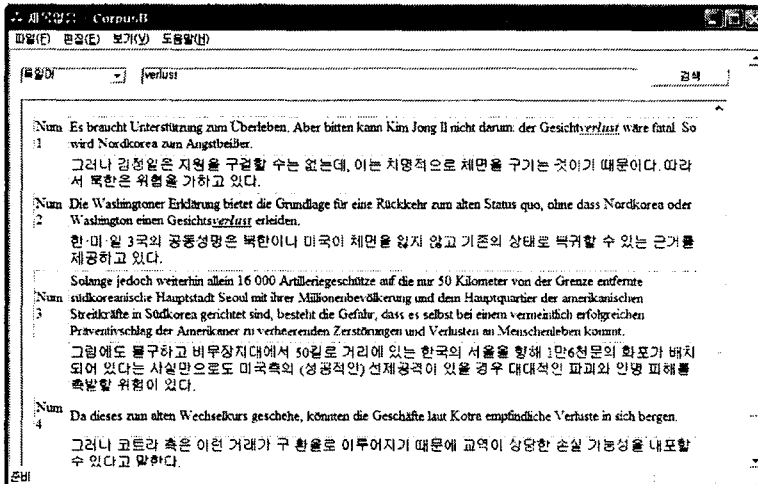
병렬 코퍼스를 구축하기 위해서는 해당 언어의 전문가가 직접 텍스트를 번역하여 코퍼스를 구축할 수도 있지만 여기서는 기초 연구 자료의 수집을 위하여 번역되어 있는 텍스트를 병렬 코퍼스로 수집하는 것을 우선적으로 시도하였다.

1차적으로 구축된 병렬 코퍼스의 대상은 Korea Heute(<http://www.koreaheute.de>)의 2000년 4월부터 2003년 10월까지의 기사를 중심으로 하였다. 수집된 데이터의 양은 독일어 코퍼스는 총 654,655 어절이며 한국어 코퍼스는 491,792 어절의 양으로 이루어져 있다.

콩코던스 Konkordanz 프로그램을 이용하여 병렬 텍스트를 분석하면 문장의 패턴과 대응어의 선정을 위한 충분한 자료를 얻을 수 있을 것이다.

물론 여기서 문제가 되는 것은 구축된 병렬 코퍼스의 질적인 문제이다. 경우에 따라서는 지나치게 직역이 되거나 또는 완전히 의역이 된 경우도 있기 때문이다. 본 논문에서는 이러한 문제를 제시하기 위하여 독일어 원문에 대한 한국어 번역 원문을 수정하지 않고, 원 번역자가 번역한 원문을 그대로 보여 주었다. 이러한 문제를 수정하기 위해서 이 연구에서는 코퍼스를 구축한 후 후처리 작업 Post-Editing을 통해 번역 예문의 내용을 수정하려고 한다.

다음 그림은 구축된 코퍼스를 바탕으로 코퍼스를 분석할 수 있는 화면 인터페이스이다.



구축된 병렬 코퍼스에 대한 검색을 가능하게 하기 위하여, 파일메뉴에서는 코퍼스를 불러올 수 있으며, 한 개 또는 여러 개의 코퍼스를 동시에 불러와서 검색을 할 수 있다.

검색 화면에서는 좌측의 언어를 독일어나 한국어로 설정하여 해당 언어의 키워드를 중심으로 검색을 하고 그 결과를 다른 파일로 저장하여 결과에 대한 분석을 가능하게 한다.

2.3. 연어 Kollokation 분석의 중요성

문장을 번역할 때 문장에서 가장 핵심적인 요소는 동사이므로 번역의 이해도를 높이기 위해서는 이에 대한 정확한 대응어를 선택하는 일이 중요하다. 특히 한국어는 명사와 동사 사이의 조사가 생략되는 현상이 빈번하고(겉먹다, 밥먹다 등), 동사와 동사의 연결구성이 많으며(들어보다, 사먹다 등), 한자어와 고유어의 결합이 많아(결합하다, 결합시키다, 결합되다 등) 이를 통사적으로 보아야 할지 복합어로 한정해야 할지를 판정하기가 쉽지 않다. 이러한 단위의 결정과 한정에는 언어이론 뿐만 아니라 사전편찬의 실제적이고 응용적인 부분 또한 고려되어야 한다.⁷⁾

한편 독일어의 경우에도 동사가 어떤 어휘와 결합되어 문장에서 나타나는 지에 따라 각각 다르게 한국어로 번역이 될 수 있다. 예를 들면 동사 *nehmen* 은 일반적으로 ‘취하다’, ‘택하다’, ‘받다’의 의미를 지니고 있지만 함께 나타나는 어휘에 따라서 각각 달리 해석될 수 있다. 다음은 이와 같은 현상을 보여 주는 예이다.⁸⁾

1a. Drei Kilo von dem deutschen Rindfleisch hat die Familie zugeteilt bekommen.

Dankbar *nimmt* die alte Frau *die Hand* eines deutschen Helfers.

1b. 송 할머니 가족은 독일이 지원한 쇠고기 3킬로를 배급받았는데, 할머니는 감사해 하며 독일 원조요원의 손을 *잡았다*

7) 김현권(2003): 한국어의 동사(III). 국어학의 이해. 1쪽 참조

8) 이 논문에 제시된 용례는 Korea Heute의 2000년 4월부터 2003년 10월까지 구축된 코퍼스에서 검색어를 통해 검색된 결과를 보여 준 것이다.

- 2a. Dann *nimmt* das jüngste Politbüro-Mitglied *das Wort* - Michail Gorbatschow, 52 der spätere Kreml-Chef.
- 2b. 그러자 가장 나이 어린 정치국원이자 나중에 크렘린궁에 입성할 고르바초프(52살)가 발언하였다.
- 3a. Eher lässt sich sagen: Bush *nimmt Kim Jongil ernst*.
- 3b. 오히려 부시 대통령이 북한을 *중요시하고* 있다고 말할 수 있다.
- 4a. Die jubelnden Massen und roten Banner sind ein deutliches Zeichen dafür, wie *wichtig* Nordkorea *dieses Gipfeltreffen nimmt*.
- 4b. 환호하는 군중들과 붉은 깃발들은 북한이 이번 정상회담을 얼마나 *중요하게 생각하는* 가를 분명 보여주는 것이다.
- 5a. Außerdem dürfte die Entscheidung des Nordens auch dazu dienen, der Kritik an der "Sonnenscheinpolitik" in Südkorea *die Spitze zu nehmen*.
- 5b. 아울러 북한측의 대화재개 결정은 한국내에서의 "햇볕정책"에 대한 비판도 *수그러들게* 할 것이다.
- 6a. Unabhängig von der Darstellung des nordkoreanischen Regimes, die USA hätten ihre vertraglichen Verpflichtungen nicht erfüllt, *nimmt* das Regime also eine verheerende Hungersnot *in Kauf*.
- 6b. 북한 정권은 미국이 핵합의상의 의무를 준수하지 않았다는 주장을 펴고 있으나, 이와는 무관하게 심각한 기아 상황을 스스로 감수하고 있는 것이다.
- 7a. Auch erlaubt es die Präsenz am Ort, zumindest in geringem Umfang *Einfluss auf Reformen zu nehmen*.
- 7b. 그리고 문제 지역에 직접 뛰어들 경우 적어도 소규모의 범위이기는 하지만 개혁작업에 *영향력을* 줄 수 있다는 것이다.

위의 예문에서 *nehmen*의 목적으로 *die Hand*와 *das Wort*가 나오는 경우 하나는 ‘(손을)잡다’로, 다른 하나는 ‘(발언)하다’로 번역이 된다. 또한 *die Spitze nehmen*의 경우에는 ‘(비판을) 없애다’라는 의미로 번역이 된다.

그러나 *etw. ernst nehmen*이나 *etw. wichtig nehmen*에서는 *nehmen*이 ‘받아들이다’로 번역된다. 한편 전치사와 명사와 함께 결합하여 *in Kauf nehmen*, *Einfluss auf etw. nehmen*에서는 각각 ‘감수하다’와 ‘영향을 미치다’라는 뜻으로 동사구에 의미가 용해되어 번역이 된다.

여기서 알 수 있는 것은 동일한 어휘라 하더라도 언어에서 서로 공기하여 나타나는 어휘에 따라 각기 다른 의미를 가질 수 있다는 것이다. 따라서 동사의 대역어를 선택하기 위해서는 동사가 결합하여 나타나는 언어정보를 이용한다면 대역어 선택이 훨씬 수월해질 수 있다. 이를 위해서 언어정보를 사전에 함께 구축해 놓는 것이 유용하게 쓰일 수 있다.

이러한 언어정보를 추출하기 위해서는 대량의 코퍼스에서 동사와 공기하는 목적어구, 명사구, 전치사구 등의 쌍을 추출한다. 그리고 동사와 목적어, 또는 기타 필수 성분어와의 결합된 형태의 의미가 문장의 문맥을 통해 결정되면, 이렇게 결정된 의미를 개념사전에 추가하면서 대안사전을 구축해 나간다. 또한 동사 번역 역시 문맥을 통하여 결정하며, 동사의 번역 부류는 각 동사마다 미리 정의하여 이를 워드넷과 연결시킨다. 이 단계에서 명사 의미와 동사 번역의 결정은 수동으로 이루어진다.

2.4. 코퍼스 분석을 통한 번역 대응어 선택의 문제

독일어와 한국어 또는 한국어와 독일어 사이의 번역에서는 유사한 개념의 어휘가 1:1로 대등하게 번역되기보다는 목표언어에서 여러 어휘가 결합한 구의 형태로 번역되는 경우를 종종 볼 수 있다. 이러한 현상은 명사의 번역에서 보다 동사의 번역이나 성구소 Phraseolexem에서 더 많이 발견된다.

가령 우리가 흔히 속어 Idiom라고 일컫는 언어현상은 주로 동사와 부사구, 전치사구 등의 여러 어휘가 한 묶음으로 결합되어 우리말에서 한 가지 뜻을 나타내는 경우가 대부분이다. 이런 속어들에서는 동사와 함께 부사구, 전치사구가 모두 의미를 결정하는 제한 요소가 되기 때문에, 이러한 점도 함께 고려해야 한다.

8a. Aber wird diese Abstimmung, wenn sie denn stattfindet, die militärischen Pläne Washingtons *beeinflussen*? fragte die "Iswestija" am Donnerstag und beantwortete die Frage gleich selbst:

8b. 러시아 신문 "이스베스티아"지는 "그러나 표결이 이루어진다면 과연 미국 정부의 군사적 계획에 *영향력을 발휘*할 수 있을까?"라고 자문하면서 동시

에 스스로 답변도 제시했다.

- 9a. Aus Sorge darüber, daß Nordkorea zu einem zweiten Irak werden könnte, versuchen Tokio und Seoul, weiter auf Pjöngjang *einzuwirken*.
- 9b. 북한이 제2의 이라크가 될 수도 있다는 우려에서 일본과 한국 정부는 계속 북한에 대한 *영향력 행사*를 시도해 왔다.
- 10a. Ungeachtet aller Schwächen sind Südkoreas Gewerkschaften eine *einflussreiche* Kraft geworden.
- 10b. 한국의 노조는 여러 취약한 점들이 있음에도 불구하고 *영향력이* 큰 세력으로 자리를 잡았다.

앞의 예문에서 “*einflussreich* = 영향력이 큰”, “*einwirken* = 영향력을 행사하다”, “*beeinflussen* = 영향력을 발휘하다”에서처럼 독일어 어휘가 한국어에서는 명사와 용언이 결합한 형태로 대응되는 것을 볼 수 있다.

반대로 한국어의 경우에는 하나의 어휘에 대하여 여러 개의 다의어가 대응되는 등의 상당히 복잡한 대응관계를 찾아 볼 수 있다. 물론 이러한 대응어가 기존 사전에 나와 있는 것도 있지만 그렇지 못한 경우도 있다. 그리고 실제로는 기존 사전에서는 찾을 수 없는 번역 가능성들이 다양한 텍스트에서 발견된다.

예를 들어, 기존 한독사전에서 ‘감소’라는 표제어를 찾아보면 *Verminderung*, *Abnahme*, *Abnehmen*, *Verkleinerung*, *Verringerung* 등의 명사들을 발견하게 된다.

그러나 다음 예문들을 보면 그 밖의 다른 어휘들도 자연스럽게 ‘감소’로 번역됨을 알 수 있다.

- 11a. Für Hyundai Motor ein schwerer Schaden: Die Koreaner bezifferten den Umsatzverlust auf bislang gut 1 Mrd. \$.
- 11b. 이는 현대차에 거대한 손실을 가져다 주었는데, 그 동안 매출이 10억 달러 상당 *감소*했다.
- 12a. Als Gründe für den *Rückgang* des Viehbestands gelten die kommunistische Mißwirtschaft, Dürren und Überflutungen.
- 12b. 가축수의 *감소*원인으로는 공산주의 계획경제의 실패, 한발과 홍수 등 자연재해가 거론되고 있다.

- 13a. Sorgenkind bleibt die Bauindustrie die im Jahr 2000 weiter um 3,7 % schrumpfte nach einem *Einbruch* im Vorjahr von bereits 9,1 %.
- 13b. 분야별로 보면 특히 건설분야가 1999년 9.1% 감소에 이어 지난 해에도 3.7% 감소를 기록하는 불황을 보이고 있다
- 14a. Das Ziel des Infineon-Chefs Ulrich Schumacher: Transfer von Technologie, Marktanteile in Japan und *Abbau* von Kapazitäten, damit sich die Preise erholen.
- 14b. 울리히 슈마허 인피니온 CEO의 목표는 기술 이전, 일본 시장 점유율 제고, 가격 정상화를 위한 생산량 감소 등이다.
- 15a. Einen Rückzug von US-Einheiten, heißt es, könnte der Norden als *Minderung* der Bedrohung werten, denn nicht der Süden, sondern die US-Streitmacht gilt in Pjöngjang als Feind.
- 15b. 한편 북한은 한국이 아닌 미군을 적으로 간주하고 있는 만큼 주한 미군의 철수는 북한으로서는 위협의 감소를 의미할 수도 있다.

복합어의 경우에는 어떤 어휘와 결합하느냐에 따라 ‘감소’에 대한 다양한 독일어 표현을 발견할 수 있다.

Produktionsausfälle	생산 감소
Gewichtsverlust	체중 감소
Schmerzlinderung	고통 감소
Produktionsreduktion	생산 감소
Einkommenseinbuße	수입 감소(* 타격)
Einkommenseinbruch	수입 감소
Nachlassen des Gedächtnisses	기억력 감소

이때 ‘감소’에 내포된 의미는 각 경우에 따라 차이가 있으며, 독일어의 경우에는 각 어휘를 사용할 수 있는 통사적, 의미적, 화용적 환경이 다르다. 따라서 “대안사전”은 반드시 각 예들의 사용 용법을 문장이나 문장 사슬의 형태로 보여주어야 할 것이다.

그리고 이러한 감소에 해당하는 어휘를 독일어 워드넷의 한곳에 분류하여 체계화해 놓으면, 우리말로 번역하거나 반대로 독일어로 번역할 때 유용하게 사용할 수 있다.

한편 번역에서는 통사적 구조를 바꾸거나 어휘를 합성함으로써 번역이 가능해지는 경우가 많으며, 때로는 이를 통해서만 좋은 문체를 만들어 낼 수 있다. 아래 예에서는 ‘감소’에 해당하는 독일어 표현이 각각 형용사, 합성어의 일부, 또는 분사로 나타나고 있다.

매출 감소	rückläufiger Umsatz
(실업 감소로 인한) 지출 감소	Minderausgaben (wegen des Rückgangs der Arbeitslosigkeit)
영향력 감소	verminderter Einfluss

이와 유사한 예로 다음 예문에서는 ‘역사교과서 왜곡’과 관련된 표현들이 독일어로 어떻게 실현되는지를 살펴보기로 한다.

- 16a. Daß der entschiedene Protest aus Seoul, verbunden mit der Aufforderung, 35 "fehlerhafte" Passagen in den Geschichtsbüchern zu ändern, ignoriert wird, empfinden die Nachbarn als eine Erniedrigung, als eine Enttarnung freundschaftlicher Lippenbekenntnisse.
- 16b. 한국인들은 *역사교과서에 왜곡 기술된 35개항*을 일본측에 수정하도록 강력한 항의를 폈으나 무시당하자, 이를 입으로만 우호를 외치는 일본의 진짜 모습을 보여주는 무례한 행동으로 느끼고 있다.
- 17a. Tokio hat seinen Teil der Vereinbarung nicht eingehalten: die *Geschichtsklitterung* geht weiter.
- 17b. 그러나 일본측은 양국간의 합의 내용을 이행하지 않았고 일본인들의 *역사왜곡*은 계속되고 있다.

일반적으로 ‘역사교과서왜곡’을 독일어로 표현하면 ‘Geschichtsschulbuchkritzelung’이라고 할 수 있는데, 위 예문에서는 이와 관련하여 ‘fehlerhafte Passagen in den Geschichtsbüchern’이나 ‘Geschichtsklitterung’이란 표현으로 실현되고 있다.

한독사전에는 표제어 ‘왜곡’에 대해 *Verbiegung, Krümmung, Verdrehung* 등의 역어를 나열하고 있을 뿐, ‘Klitterung’이나 ‘fehlerhaft’를 사용한 표현은 등재되어 있지 않다.9)

한편 기존 사전의 정형화된 표현에 의지하여 ‘왜곡하다’라는 한국어를 독일어로 번역한다면 ‘verbiegen’, ‘krümmen’, ‘verdrehen’ 등의 정형적 어휘만을 사용하기가 쉬운데, 앞에서 언급했듯이 통사적 유연성이 허락되는 ‘falsch darstellen’이나 ‘ein falsches Bild abgeben’ 등과 같은 유용한 표현도 사용할 수 있다는 것을 알려줄 필요가 있다.

물론 사전에 가능한 모든 표현을 다 등재시켜야 한다는 것은 아니다. 또 그렇게 할 수도 없다. 그렇게 되면 표제어의 서술에 대한 한계를 설정하는 것에도 많은 논쟁이 야기될 수 있다.

단지 이 논문에서 주장하고자 하는 바는 기존 사전에 제시된 표제어의 의미뿐만 아니라 코퍼스 분석을 통한 보다 폭넓은 표제어의 사용을 제시하여야 하며, 이를 통하여 외국어 교육이나 번역에서 한 언어의 의미를 다른 의미로 올바르게 전달하고 앞의 예문에서와 같이 통사적 구조를 변형하면서도 문법에 맞고 의미가 자연스러운 문장으로 바꿀 수 있는 유연성을 키워주어야 한다는 점이다.

또한 이 연구에서 추구하는 ‘대안사전’에서는 우리말에서 요즘 흔히 사용되는 조어 방식의 허와 실을 보여줌으로써 외국어 학습자나 번역가의 어려움을 덜어주고자 한다.

예를 들어 ‘-화’, ‘-형’ 등은 독일어에서 특정한 형태소로 실현될 수도 있지만, 경우에 따라서는 전혀 실현되지 않는다.

다음의 예에서는 ‘-화’가 ‘-ieren, -ierung’으로 실현되었다.

내실화 정책	Konsolidierungspolitik
경영 안정화	Stabilisierung des Managements

그러나 다음 예에서는 ‘-화’의 의미 요소가 전혀 실현되지 않거나, ‘Belebung’이라는 어휘에 포함되어 있다.

경영활동의 현지화	Betrieb auf örtlicher Basis
정보화산업	Informationsindustrie
지역경제 활성화사업	Projekte zur Belebung der regionalen Wirtschaft

9) 독한사전에는 ‘klittern’이 ‘일룩점을 묻히다, 글씨를 서투르게 쓰다, 피발개발 쓰다’로 설명되어 있다.

아래 예들을 보면 ‘-형’은 독일어에서는 전혀 실현되지 않은 것을 볼 수 있는데, 이러한 표현들은 번역가나 외국어 학습자들을 혼란시키는 주범 중의 하나이다.

수출전략형 현지 모델	exportstrategisches örtliches Modell
고부가가치형 사회	Mehrwertgesellschaft
현지 밀착형 판매전략	ortnahe Verkaufsstrategie
ITS 지능형 교통시스템	Intelligentes Verkehrssystem?

마지막으로 성구소와 관련하여 대안사전의 효율성을 살펴보면 다음과 같다. 독한사전에서 표제어 ‘Ohr(귀)’를 찾으면 다음과 같은 표현들이 나온다.

jn. übers Ohr hauen	속이다
bis über die Ohren in Schulden sitzen	빚더미에 앉다

그러나 한독사전에서 ‘속이다’나 ‘빚’을 찾으면 이러한 독일어 표현들이 나오지 않는다. 이러한 것들을 반대 방향으로 찾을 수 있다면, 학습과 통번역 모두에 매우 도움이 될 것이다.

특히 외국어 학습이나 번역에 있어 가장 어려운 언어 영역 중의 하나가 소위 말하는 속어인 것을 고려할 때 이러한 찾기 방법은 요원하기 이를 데 없다. 속어는 1) 통사적인 측면에서 볼 때 2개 이상의 어휘로 고정 결합되어 있어(다어휘성 Polylexicalizität/ 고정성 Stabilität) 이들을 한 어휘처럼 외어야 한다는 부담감이 있고, 2) 속어의 전체적인 뜻이 어휘 구성요소들의 개별 뜻에서 전혀 유추되지 않거나 또는 일부만 유추될 수 있으며(속어성, Idiomatizität), 3) 속어에는 외국인들이 알 수 없는 화용적 정보가 담겨져 있다. 화용적 뜻을 모르고 사전에 제시된 뜻만 갖고 속어를 사용하게 되면 언어 규범에 적절치 않은 민망한 상황을 초래하거나 적절치 않은 번역, 나아가 오역을 하게 된다. 그러므로 능동적(외국어→모국어 번역) 상황은 물론이요, 수동적(모국어→외국어) 상황에서 속어를 적절하고 올바르게 사용/이해하기 위해서는 해당 속어의 다양한 정보를 제시해 줄 수 있는 외국인 학습자 친화적인 사전이 무엇보다

다 필요한데(천미애 2001 참조), 대안사전은 이러한 요구도 많은 부분 충족시킬 수 있을 것이다.

3. 결론

이중언어 사전은 출발언어의 한 어휘에 대한 목표언어의 의미적, 화용론적 사용을 중심으로 하는 대역어의 내용을 담고 있다. 그리고 이러한 사전이 번역을 하는 데 가장 기본적인 도움을 제공하고 있다. 그러나 번역을 해 본 경험이 있는 사람이라면 번역에서 대역어를 선별할 때 단순히 사전에 나오는 대역어를 택하여 직접 번역이나 통역을 할 수 있는 것이 아니라 그와는 다른 개념을 기반으로 대역어를 선택해야 하며, 때로는 대역어를 사전에서 찾을 수 없는 상황도 있다는 점을 알고 있다.

지금까지 사전편찬을 위해 어의론적인 관점에서 연구가 이루어짐으로써 오늘날의 사전편찬학의 발전을 이루었다는 것은 팔목할 만한 성과이다. 그러나 그에 반해 명칭론적인 관점에서는 최근에 들어서야 비로소 많은 연구가 시도되고 다양한 연구 성과를 보여 주고 있다.

이 논문에서는 명칭론적인 관점에서 번역이나 외국어 교육에 도움이 될 수 있는 개념사전을 개발하기 위하여 독일어와 한국어의 병렬 코퍼스를 분석하고 그 분석 결과를 워드넷에 접목하여 새로운 형태의 대안사전을 개발할 것을 제안하였다.

이를 위하여 독일어와 한국어의 병렬 코퍼스를 구축하고 이 병렬 코퍼스에서 특정 어휘에 대한 대역어의 실현 가능성을 분석해 봄으로써 기존 이중언어사전을 보완할 수 있는 방법론을 부분적으로나마 살펴보았다.

여기서는 특히 번역에서 미리 설정된 의미범주에 의존하지 않고 일반적이고 각 문장에서의 쓰임에 적합한 역어를 사전에 수록하기 위한 새로운 사전 정보 형식으로 필수적인 언어정보에 주안점을 두었다.

병렬 코퍼스는 사전편찬을 위한 여러 이론에 중요한 정보를 줄 수 있고 다의어나 언어정보 등 사전편찬을 위한 구체적이고 실증적인 증거를 제공해 준다. 이러한 병렬 코퍼스의 구축과 분석을 통한 대안사전의 개발은 사전 편찬

뿐만 아니라 외국어 교육을 위한 다양한 연구를 위해 활용될 수 있다. 이 외에도 한독, 독한 병렬 코퍼스의 구축은 최초의 시도라는 점에서 그 자체의 작업만으로도 충분한 가치를 지닌다고 할 수 있다.

이 논문에서는 병렬 코퍼스의 구축과 이 코퍼스의 활용을 위한 경험적 연구방법의 일부만을 제시하였는데, 앞으로 이를 확장시켜 그 연구 결과를 구체적으로 워드넷의 개념체계와 연결시킴으로써 보다 다양한 후속 연구를 파생시킬 수 있다고 기대한다.

참고문헌

- 김민수 외 (1995): 국어사전을 이용한 한국어 명사에 대한 상위어 자동 추출 및 WordNet의 프로토타입 개발. 정보처리학회논문지, Vol.2, No. 6.
- 김현권 (2000): EuroWordNet 의 구성 원리와 설계. 언어학, Vol.27, No.1.
- 남유선 (2004): 한독 워드넷 구축을 위한 기본 방법론 고찰. 독어학 9집.
- 문유진 (1996): 의미론적 어휘개념에 기반한 한국어 명사 Wordnet의 설계와 구축. 서울대 박사학위 논문.
- 안인경 (1995): 독문학 용어번역과 용어사전. 번역연구.
- 오장근 (2002): 유로워드넷 기반의 어휘 데이터베이스 활용을 위한 한국어-독일어 ILI 대응 방법론 연구. 독어학 6집.
- 이민행 (1999): 독일어 어휘부에 대한 연구. 독일문학 69집.
- 이민행 (2004): 독일어와 영어의 감정명사들의 의미관계에 대한 연구. 독일문학 89집.
- 이재윤 & 김태수 (1999): WordNet과 시소러스. 언어정보의 탐구 1. 연세대학교 언어정보개발연구원.
- 이창기 & 이근배 (1999): WordNet을 이용한 한국어 시소러스 자동 구축. 제11회 한글 및 한국어 정보처리 학술대회 논문집.
- 천미애 (2001): 독어 성구소의 번역 문제. 텍스트 언어학 11.
- 최기선 외 (2002): 명사 워드넷과 단일어 사전을 이용한 한국어 동사 워드넷 구축. 제14회 한글 및 한국어 정보처리 학술대회 논문집.
- 최기선 외 (2003): 한국어 워드넷. “개념체계”. 한국과학기술원 전문용어언어공학연구센터.

- Ch. Fellbaum (1998): WordNet. An Electronic Lexical Database. MIT Press.
A. Wichmann et al. (1997): Teaching and Language Corpora. Addison Wesley Longman.
H. E. Wiegand (1998): Wörterbuchforschung. Walter de Gruyter.
EuroWordNet home page. <http://www.hum.uva.nl/~ewn>

- 사전 -

- 지명렬(편): 동아 프라임 독한사전. 서울 1996.
한국 독어독문학회(편): 모델 독한사전. 서울 1995.
허형근(편): 옛센스 독한사전. 서울 1997.
Duden Deutsches Universalwörterbuch A-Z 1996. 3. neu bearbeitete Auflage.

Zusammenfassung

Eine methodologische Untersuchung zur Erstellung eines Ersatzlexikons zum Zweck der Übersetzungserziehung

Choi, Byung-Jin(Mokpo National Univ.)

Ziel dieser Arbeit ist es, Lexkion zu erstellen, das nicht nur der menschlichen Übersetzung, sondern auch der natürlichsprachlichen Verarbeitung dienen soll.

Um eine automatische maschinelle Übersetzung durchführen zu lassen, braucht man noch ein umfangreiches Wissen in Linguistik und Informatik.

Bei der Verarbeitung, aufgrund derer ein Wort oder eine Phrase aus der Ausgangssprache durch ein passendes Äquivalent der Zielsprache überstzt wird, sollte man auch auf die kulturellen Hintergründe Rücksicht nehmen.

Während des Übersetzens versuchen Übersetzer oder Dolmetscher, zweisprachige Wörterbücher zu konsultieren, um die angemessensten Äquivalente der Zielsprache zu finden. In vielen Fällen aber macht man die Erfahrung, dass solche Bemühungen vergeblich sind.

In dieser Arbeit wird zunächst das parallele Korpus (Koreanisch-Deutsch) untersucht, um das Desiderat der herkömmlichen Wörterbücher festzustellen. In den parallelen Koreanisch-Deutsch Korpora wird die Wörter oder Phrasen der Ausgangssprache mit denen

der Zielsprache verglichen und auf dieser Grundlage eine Liste von Mängeln erstellt.

Die Ergebnisse dieser Untersuchung sollten nicht nur auf die menschliche Übersetzung, sondern auch auf die natürlichsprachliche Verarbeitung angewendet werden, damit die Effizienz der Übersetzung noch erhöht wird.

Außerdem soll die weitere Untersuchung mit den Korpora auch für den Fremdsprachenunterricht Anwendungsmöglichkeiten bieten.

[검색어] 번역, 워드넷, 대안사전
Übersetzung, Wordnet, Lexikon als Alternative

최병진

534-729

전남 무안군 청계면 도림리 61

목포대학교 어문학부 독일언어문화학 전공

bjchoi@mokpo.ac.kr