

어휘함수를 이용한 한국어 어휘 관계의 기술*

이 병 근 · 박 진 호 · 김 진 형

1. 서 론

최근 실용적인 측면에서나 이론적인 측면에서 어휘 관계(lexical relation)를 총체적으로 기술하고 분류하는 일이 중요한 과제로 대두되고 있다. 본고는 한국어 어휘 체계를 총체적으로 기술하는 데에 I. A. Mel'čuk이 제안한 어휘함수(lexical function)라는 개념이 어떻게 이용될 수 있는지에 대해 탐색해 보려 한다. 우선 2절에서 한 언어의 어휘 관계를 총체적으로 표상하는 일의 필요성을 살펴보고, 3절에서 어휘 관계에 대한 기존의 연구가 지닌 한계를 지적하고, 4~6절에서 어휘함수의 개념과 유형을 소개하고, 7절에서 어휘함수를 한국어에 구체적으로 적용하는 문제를 살펴보려 한다. 그리고 부록에서는 어휘함수의 실제 활용 사례로서 Mel'čuk의 결합 설명 사전의 포맷에 따라 '눈'이라는 단어를 기술한 것을 제시하였다.

2. 어휘 관계 記述의 필요성

2.1. 실용적 측면

2.1.1. 정보 검색

최근 정보화 사회로 넘어가면서 인터넷이 우리 생활의 구석구석까지 침투하고 있고, 이에 따라 인터넷의 여기저기에 널려 있는 정보의 바다로부터 필요한 정보를 최소의 시간과 노력을 들여서 찾아내는 일이 많은 사람들의 관심사로 대두되고 있다. 정보가 곧 돈이고 힘이고 국력인 시대에서 이것은 개인의 차원을 넘어 국가적인 관심사라고도 할 수 있다. 그래서 누구나 검색 사이트를 자주 이용하게 된다. 그러나 검색 엔진을 만들기 위한 여러 가지 기반 기술들이 아직은 미흡한 형편이기 때문에 검색 엔진의 성능도 여러 가지 면에서 만족스럽지 못한 것이 사실이다. 검색 시스템의 성능에 대한 평가는, 사용자가 원하는 정보의 총량 중 얼마나 많이 찾아냈는가를

* 이 연구는 1998년도 한국학술진흥재단 자유공모과제 공동연구지원을 받아 이루어졌다.

측정하는 재현율(recall)과, 검색 시스템이 뽑아낸 정보 중 사용자에게 실제로 필요한 정보가 얼마나 되는가를 측정하는 정확률(precision)에 따라 이루어지므로 이 두 가지에 대해 살펴보겠다.

우선 현재의 검색 엔진들은 사용자가 원하는 정보를 모두 다 뽑아 주지 못한다. 즉 재현율이 낮은 것이다. 여기에는 여러 가지 원인이 있겠지만, 가장 중요한 원인은 현재의 검색 엔진들이 단순한 문자열 비교(string matching)의 수준에 머물러 있다는 것이다. 현재의 검색 엔진들은 수많은 웹 문서들에서 색인어(index word)를 추출하여 이에 따라 색인을 만들어 놓고, 사용자가 입력한 질의어(query word)와 일치하는 색인어를 가진 웹 문서들을 사용자에게 제시하는 식으로 되어 있다. 그런데 사용자가 입력한 질의어가 사용자가 원하는 문서에 반드시 포함되어 있으리라는 보장이 없다. 좀 단순화시켜서 예를 들어 보면, 사용자가 ‘주식 시세’라는 질의어를 입력했을 때 검색 엔진은 ‘주식 시세’라는 문자열이 들어 있는 문서만을 찾아 주지만, 사실 ‘주식 시세’라는 문자열은 포함하고 있지 않더라도 ‘증권 시장’, ‘주가지수’, ‘기관 투자자’, ‘선물 시장’, ‘코스닥 시장’ 등의 문자열을 포함하고 있는 문서에 사용자는 관심을 가지고 있을 가능성이 높은 것이다. 그렇다면 단순한 문자열 비교에 머무를 것이 아니라 사용자가 입력한 질의어와 밀접한 관계가 있는 다른 단어들을 포함한 문서도 찾아 줄 필요가 있는 것이다. 실제로 현재 일부 검색 엔진에서는 이러한 사실을 인식하고 부분적으로 그런 시도를 하고 있다. 예를 들어, 사용자가 입력한 질의어에 대해 일단 문자열 비교에 입각해서 검색 결과를 보여 주면서, 거기에 덧붙여 “당신이 A라는 단어에 관심이 있다면 B, C, D 등등의 단어에도 관심이 있을 가능성이 있는데 이들 중 관심이 있는 것이 있으면 그 단어를 다시 검색하십시오” 하는 식이다. 이것을 질의어 확장이라고 한다. 이런 일들이 원활히 이루어지기 위해서는 사용자가 질의어로 입력할 가능성이 있는 단어에 대해(이것은 한 언어의 거의 모든 實辭가 될 것이다) 이 단어와 밀접한 관련이 있는 단어들을 미리 다 지정해 놓아야 할 것이다. 다시 말해 그 언어의 어휘 체계 전반에 대해 단어와 단어 사이의 어휘 관계를 총체적으로 표상해 놓은 데이터베이스가 필요한 것이다.

요즘 검색 엔진 사용자들은 재현율보다는 정확률에 더 큰 불만을 가지고 있는 듯하다. 웹 문서의 수가 엄청나게 많고 또 나날이 늘어나다 보니 웬만한 질의어에 대해서는 수만, 수십만 건의 문서가 뽑혀져 나온다. 대부분의 사용자는 이 많은 문서를 다 뒤질 수 없기 때문에 맨 앞의 몇 페이지에 나온 극히 일부의 사이트에만 들어가 보고 마는 것이 보통이다. 정확률이 낮은 원인도 여러 가지가 있는데 이 역시 근본적으로는 모두 문자열 비교에 머무르고 있다는 것으로 귀착된다. 사용자가 입력한 질의어와 일치하는 문자열을 지니고 있는 문서라 하더라도 동형어(homonym)의 존재 때문에 사실은 사용자가 원하는 문서가 아닐 수 있다(예: bank). 또한 형태소 분석의 잘못에서 기인하는 경우도 있다(예: 대학생선교회). 한 번에 사용자가 원하는 문서만을 선별하기 어렵기 때문에, 몇 단계를 거치게 하는 방법도 생각해 볼 수 있고, 실제

로 일부 검색 엔진에서는 이런 방법을 사용하고 있다. 즉, 사용자가 처음에 A라는 질의어를 입력하면 A를 포함한 문서를 일단 다 찾아 놓고서, “이 많은 문서 중에서 진짜 당신이 원하는 문서만을 골라내기 위한 단서를 더 제공해 주십시오. A와 관련된 단어로 다음의 B, C, D 등등의 단어들이 있는데 이들 중 어느 단어가 들어 있는 문서들만을 원하십니까?” 하는 식이다. 이것을 질의어 정교화라고 한다. 예를 들어, 사용자가 ‘파충류의 생태’라는 질의어를 입력하면 ‘파충류’를 포함한 문서들을 일단 다 찾아 놓고서(이들 문서의 수는 엄청날 것이다) ‘파충류에는 이러이러한 것들이 있는데 이들 중 어느 동물에 관한 문서만을 여기서 다시 추릴까요?’ 하는 식으로 묻는 것이다. 이런 일이 원활히 이루어지기 위해서도 어휘 관계를 총체적으로 표상한 데이터베이스는 필수적이다.

요컨대 정보 검색 시스템의 재현율과 정확률을 향상시키는 데에, 어휘 관계를 총체적으로 표상한 데이터베이스가 꼭 필요하다고 할 수 있다.

2.1.2. 기계 번역

기계 번역은 자연언어처리 관련 기술들이 집대성되어야 제대로 이루어질 수 있는 일종의 종합 예술이기 때문에 요소 기술 중 어느 하나라도 제대로 되어 있지 않으면 성공하기가 어렵다. 이런 점을 감안할 때 현재 나와 있는 기계 번역 시스템들이 만족스러운 수준이 아닌 것은 당연하다고 할 수 있다. 기계 번역 시스템의 성능이 낮은 데에도 여러 가지 원인이 얽혀 있지만, 성능 향상의 중요한 결림들은 역시 어휘 관계를 총체적으로 표상한 데이터베이스의不在라고 할 수 있다.

다음 한국어 문장을 한영 기계 번역 시스템이 영어로 어떻게 번역할지에 대해 생각해 보자.

- (1) 정부는 오늘 회의에서 대우자동차의 회생이 불가능하다고 결론을 내렸다.

현재 대부분의 기계 번역 시스템은 출발 언어(source language)의 입력 문장에 대해 형태소 분석과 구문 분석을 수행하고 대역 사전(이 경우에는 한영 사전)을 참조하여 각각의 한국어 단어들을 영어 단어들로 전환한 뒤 이 단어들을 가지고서 영어의 문법 및 입력 한국어 문장의 구조를 참조하여 영어 문장을 생성하는 방식으로 이루어진다. 그럼 위의 한국어 문장에 대한 형태소 분석이나 구문 분석이 정확하게 이루어졌다고 가정하고(이것도 현재로서는 쉬운 일이 아니다), 그 다음 일들이 어떻게 진행될지를 생각해 보자. 대역 사전을 참조하여 각각의 단어들을 영어 단어로 전환할 텐데, 아마도 ‘결론’은 ‘conclusion’으로 ‘내리’는 ‘lower’, ‘descend’, ‘get off’ 등등의 것들 중 어느 것이든 대역 사전에서 ‘내리’의 대역어로 제시되어 있는 것으로 전환될 것이다. 그러나 이렇게 해서는 제대로 된 영어 문장이 생성될 수 없다.

문장의 통사 구조 분석에서나 생성에서나, 그리고 이론적으로나 실제적으로나 문

장의 述語가 문장의 구조에 대한 정보의 열쇠를 쥐고 있기 때문에 문장의 술어가 무엇인지를 결정하는 일이 대단히 중요하다. 위의 예문 (1)의 번역에서 핵심은 명사 ‘결론’과 동사 ‘내리’를 따로따로 번역할 것이 아니라, 문장의 주된 술어 역할은 ‘결론’이 하고 있고 ‘내리’는 의미상의 술어인 ‘결론’의 역할을 통사적으로 보조하기 위해 존재하는 기능동사(support verb)라는 사실을 파악하는 데 있다. (1)에서 명체에 해당하는 보문을 취하는 것도 동사 ‘내리’ 때문이 아니라 명사 ‘결론’ 때문이라는 사실을 보아도 명사 ‘결론’의 서술 기능이 의미 차원에 머무르는 것이 아니라 통사 차원에까지 미치고 있음을 알 수 있다. 그래서 ‘결론’을 명사로 번역할 것이 아니라 ‘결론을 내리’라는 술어 명사 + 기능동사 연쇄가 영어의 하나의 동사 ‘conclude’로 번역되도록 해야 할 것이다. 이런 일이 원활히 이루어지기 위해서는 한국어의 술어 명사에는 어떤 것들이 있고 그들이 취할 수 있는 기능동사는 어떠한 것들인지에 대한 정보를 미리 구축해 놓아야 할 것이다. 이것은 술어 명사 + 기능동사 연쇄뿐 아니라, 각각의 단어를 단어 대 단어(word-to-word)로 번역해서는 안 되고 출발 언어의 둘 이상의 단어로 이루어진 연쇄에 대해 대역어를 제시해야 하는 경우 모두에 적용되는 일반적인 사실이다. 이런 현상을 기계 번역 시스템에서 적절히 처리하기 위해서는 해당 언어의 連語(collocation) 전반에 대한 총체적인 기술이 이루어져야 할 것이다. 앞의 정보검색에서 필요로 하는 어휘 관계가 주로 계열적인(paradigmatic) 어휘 관계였다면, 기계 번역에서 필요로 하는 어휘 관계는 주로 통합적인(syntagmatic) 어휘 관계라고 할 수 있다.

또한 다음의 세 예가 영어로 어떻게 번역될지를 생각해 보자.

- (2) a. 상한 달걀
 b. 상한 생선
 c. 상한 치즈

한영 대역 사전에 ‘상하다’의 대역어로 ‘addle’, ‘stale’, ‘rancid’ 등의 단어가 그냥 나열되어 있어서는 위의 세 경우 각각에 어떤 대역어를 선택해야 할지 알 수 없다. 셋 중 아무거나 선택하면 대강의 뜻은 통할 수 있을지라도 품질이 좋은 번역은 될 수 없다. 피수식어가 무엇이냐에 따라 수식어가 선별적으로 사용된다면 이에 대한 정보를 미리 표상해 놓아야 자연스러운 번역문을 생성해 낼 수 있다.

2.13. 언어 교육

최근 외국어로서의 한국어 교육에 대한 관심이 고조되고 많은 한국어 교재 및 교육 자료가 만들어지고 있는 것은 매우 다행스러운 일일 것이다. 외국인들에게 한국어를 제대로 가르치기 위해서는 정확한 문법 교육과 한국어 어휘에 대한 교육이 기본적으로 필요하겠지만, 고급 과정에서 한국어 학습자의 유창성(fluidity)을 향상시

키기 위해서는 어휘 관계에 대한 교육, 그리고 어휘 관계 정보를 이용한 교육이 필수적이다.

예를 들어 “옷이나 장신구 등을 몸에 착용하다”의 의미를 갖는 동사, 즉 영어의 ‘put on’, ‘wear’에 대응되는 한국어 동사가 무엇인지를 가르친다고 치자. 이 때 목적어가 ‘옷’ 부류에 속하면 ‘입다’, ‘모자’, ‘안경’이면 ‘쓰다’, ‘양말’, ‘신발’이면 ‘신다’, ‘허리띠’이면 ‘띠다, 두르다’, ‘망토’이면 ‘걸치다’, ‘귀걸이’이면 ‘달다’, ‘목걸이’, ‘팔찌’이면 ‘하다’ 하는 식으로 목적어와 그에 어울리는 동사의 쌍을 가르쳐 주어야 할 것이다. “연주하다”의 의미를 갖는 동사도 영어는 ‘play’ 정도이지만 한국어에서는 ‘피아노’는 ‘치다’, ‘북’은 ‘치다, 울리다’, ‘바이올린’은 ‘켜다’, ‘가야금’은 ‘타다, 뜯다’, ‘나팔’은 ‘불다’ 등으로 다양하다. 이 또한 통합적 어휘 관계의 예로서, 어휘 관계에 대한 정보가 체계적으로 표상되어 있으면 교육에 효과적으로 이용될 수 있음을 쉽게 짐작할 수 있다.

2.2. 이론적 측면

Chomsky 학파의 대두 이래로 현대 언어학자들은 화자가 자신의 모국어에 대해 지니고 있는 지식(언어 능력)을 명시적인 형태로 표상하는 것을 언어학의 주된 목표로 생각하고 있다. 그런데 화자의 모국어에 대한 지식 중에서 통사론적 지식, 음운론적 지식에 대해서는 많은 연구가 이루어져 왔고 이를 표상하기 위한 매우 구체적인 정밀한 이론적 제안들이 많이 제시되어 왔지만, 사람들이 자신의 모국어 어휘에 대해 가지고 있는 어휘적 지식, 즉 어휘부에 대해서는 상대적으로 관심이 적었고 최근야 이에 대한 관심이 차츰 고조되고 있는 형편이다. 즉, 어휘부가 전체적으로 어떤 구조를 지니고 있으며 여기에 저장되어 있는 단어들에 관한 갖가지 정보가 어떤 방식으로 표상되어야 하는가 하는 것은 현대 언어학에서 매우 중요한 이론적 문제라 할 수 있다.

또한 화자들의 어휘에 대한 지식은 단어 각각에 대한 파편화된 지식이 아니라 단어와 단어 사이의 관계를 매우 다양한 방식으로 인식하고 있는 네트워크 비슷한 것이다. 따라서 이론 언어학적 관점에서 화자의 어휘적 지식을 제대로 표상하려면 어휘 내향 각각의 미시구조에 대한 이론뿐 아니라 단어들 사이의 관계에 대한 논의, 즉 어휘부의 거시구조에 대한 연구가 필요한 것이다.

단어와 단어가 맺고 있는 관계의 양상이 매우 다양하고 복잡하기 때문에 어휘부의 거시구조에 대한 탐구는 언어학의 어느 한 분야에만 해당되는 것이 아니라 여러 분야의 接面(interface) 비슷한 성격을 지니고 있다. 지금까지는 어휘의미론에서 이에 대한 연구를 많이 해 왔지만 형태론에서 담당해야 할 부분이 많이 있다고 생각된다. 최근 형태론을 규칙보다는 유추라는 기제에 입각해서 바라보고 형태론 및 심리언어학의 관점에서 어휘부의 내부 구조를 탐구하려는 시도가 일어나고 있는 것은 매우 고무적인 일이라 할 수 있다.

3. 어휘 관계에 대한 기존 연구

어휘 관계에 대한 연구는 지금까지 주로 어휘의미론에서 담당해 왔다고 할 수 있다. 의미론 교과서를 보면 대개 의미 관계(sense relation)라는 제목 아래 동의관계(synonymy), 반의관계(antonymy), 상하위관계(hyponymy), 부분관계(meronymy) 등으로 분류하고 각각에 대해 대표적인 예 몇 개를 제시하는 정도이다. 간혹 이들 의미 관계와 관련된 개념적 문제점 등에 대한 논의가 있을 뿐이다. 이런 논의의 문제점은, 한 언어의 복잡다단한 어휘 관계의 그물조직을 표상하기에는 그 가짓수나 개념적 장치가 턱없이 부족하다는 것이다. 그리고 이러한 논의에 입각해서 한 언어의 어휘 전체에 대해 철저한 기술이 행해진 것도 아니기 때문에 2절에서 말한 실용적 목적을 위해서는 아무런 소용이 없다. 이 두 가지 문제는 서로 긴밀히 맞물려 있다고 생각한다. 한 언어의 어휘 전체에 대해 철저한 탐구가 행해진다면 어휘 관계의 복잡다단한 양상이 좀더 분명하게 드러날 것이고 어휘 관계에 대한 이론적 논의도 전통적인 어휘의미론에서 논하던 것보다는 훨씬 넓고 깊은 수준에서 이루어질 수 있을 것이다.

몇 해 전에 영어를 대상으로 비교적 대규모의 어휘에 대해 어휘 관계를 표상한 WordNet이라는 데이터베이스가 완성되어 공개되었다. WordNet은 Princeton 대학의 George Miller를 중심으로 한 여러 사람들이 다년간의 노력 끝에 내놓은 결실로서 이 정도의 대규모 어휘 데이터베이스가 만들어졌다는 것 자체가 대단한 일이고, 그 이름에서 드러나듯이 각 단어에 대한 정보를 원자적으로 제시하는 데 그치지 않고 단어들 사이의 어휘 관계를 표상하는 데에도 많은 힘을 기울였다는 데에 의의가 있다. 그러나 WordNet에서 표상하고 있는 어휘 관계도 전통적인 어휘의미론에서 논의된 수준을 거의 넘어서지 못하고 있다. 명사의 경우 동의어, 반의어, 同位語(coordinate), 上位語, 下位語를 제시하고 있고, 동사의 경우 이 밖에 troponym이라는 것을 제시하고 있는데 이것도 전통적인 의미의 下位語에서 크게 벗어난 것은 아니다.¹

4. 어휘함수의 소개

4.1. 결합 설명 사전에 대한 간략한 소개

결합 설명 사전(Explanatory Combinatorial Dictionary, ECD)은 Canada의 Université de Montréal에 있는 I. A. Mel'čuk과 그의 동료들이 개발해 온 사전으

1. WordNet에 관한 자세한 사항은 Fellbaum(1998) 참고.

로서 지금까지 나와 있는 어떤 사전보다도 어휘적 지식을 가장 철저하게 표상하고자 하는 사전이다. Mel'čuk은 오래 전에 러시아어의 결합설명 사전을 낸 바 있으며 1970년대에 Canada에 정착한 뒤로는 프랑스어의 결합설명 사전 개발에 주력해 오고 있다. 지금까지 네 권의 현대 프랑스어 결합설명 사전이 출간되었다. 결합설명 사전은 의미나 통사적 특성 등을 기술하는 데에 있어서도 매우 정밀한 방법론을 내놓고 있지만 가장 두드러진 특징은 단어들 사이의 어휘 관계를 표상하는 데에 매우 많은 노력을 기울인다는 점이다. 이를 위해 제안된 것이 어휘함수(lexical function)라는 개념이다. Mel'čuk의 어휘함수에 대한 연구는 그 어떤 어휘 관계의 모형보다도 자세한 유형 분류와 깊이 있는 개념적 천착을 보여 주고 있으며, 많은 수의 언어에 대한 정밀한 어휘 기술을 통해 그 타당성을 검증해 가고 있다.²

4.2. 어휘함수의 개념

어휘함수는 어휘 관계 중에서도 제도화된 어휘 관계(institutionalized lexical relation)를 체계적으로 표상하기 위해 고안된 장치이다. 제도화된 어휘 관계란, 두 어휘 단위 L_1 과 L_2 가 있을 때 L_2 에 대해 M이라는 특정 의미 관계를 나타내기 위해 L_1 이 자동적으로 선택되는 경우에 L_1 과 L_2 사이에 성립되는 관계를 말한다. 여기서 의미 M에 해당하는 어휘함수를 F라고 하면 이 어휘관계는 'F(L_2)= L_1 '으로 표시된다. 예를 들어 “매우 높은 정도”라는 의미(M)를 나타내는 표현으로 피수식어(L_2)가 ‘거짓말’일 때에는 수식어(L_1)로 ‘새빨간’이 선택되고 피수식어가 ‘사의(謝意)’일 때에는 수식어로 ‘심심(甚深)한’이 선택된다. 이 때 ‘새빨강’과 ‘거짓말’, ‘심심하’와 ‘사의’는 제도화된 어휘 관계를 맺는다고 한다. 즉 어떤 의미를 나타내는 단어가 여럿 존재하며 그들 중 어느 것이 선택되는가가 다른 단어에 의해 결정될 때 이들 단어 사이의 어휘 관계는 제도화된 것이라고 할 수 있는 것이다. 여기서 “매우 높은 정도”라는 의미에 해당하는 어휘함수를 Magn이라고 하면 위의 어휘관계는 'Magn(거짓말)=새빨간', 'Magn(사의)=심심한'과 같이 표시할 수 있다.

제도화되지 않은 어휘관계는 그런 어휘적 민감성을 보이지 않는 경우를 말한다. “더러운”의 의미를 나타내는 표현으로 ‘더러운’, ‘不潔한’, ‘不淨한’, ‘지저분한’ 등이 있지만 이들은 피수식어가 무엇이냐(‘손’, ‘집’, ‘옷’ 등)에 따라 어느 하나가 자동적으로 선택되는 것은 아니다(이들의 쓰임에 화용론적 제약이나 차이가 있기는 하지만 그것은 어휘적 민감성과는 다르다). 따라서 이들 사이의 관계는 제도화된 것이 아니고 어휘함수에 의한 기술 대상이 되지 않는다.

형식적 측면에서 볼 때 어휘함수는 수학적 의미에서 함수이며($f(x)=y$), x를 함수

2. Mel'čuk 등의 연구에서 기반이 되고 있는 의미-텍스트 대응 이론(Meaning-Text Theory)에 대해서는 홍재성·박동호(2000)를 참고하고, 결합 설명 사전에 대한 자세한 사항은 이병근·박진호(2000)를 참고하기 바란다.

의 논항(argument) 또는 키워드(key word), y를 함수값(value)이라 한다. ‘어휘’라는 수식어가 붙은 것은, 함수의 키워드도 어휘요소이고 함수값도 어휘요소들의 집합이기 때문이다.

내용적 측면에서 볼 때 어휘함수는 매우 특징적인 의미로서, 다음의 3가지 특성에 의해 특징지어진다.

(3) 어휘함수의 내용적 특성

- ㉠ 어휘함수와 관련된 의미는 매우 추상적이고 일반적이어서, 매우 많은 어휘요소를 키워드로 가질 수 있다.
- ㉡ 함수값이 되는 어휘요소도 매우 다양해야 한다.
- ㉢ 함수값이 되는 어휘요소는 키워드에 의해 결정되어야 한다.

5. 어휘함수의 유형

어휘함수는 여러 가지 기준에 따라 분류될 수 있다. 우선 함수의 정의구역과 치역의 크기가 크나 작으나에 따라 표준 어휘함수와 비표준 어휘함수로 나뉜다. F가 나타내는 의미가 일반적이어서 매우 많은 수의 단어에 적용되어 매우 다양한 함수값을 갖는 것은 표준 어휘함수이고, 제한된 적은 수의 단어에 적용되거나 매우 적은 함수값밖에 갖지 않는 것은 비표준 어휘함수이다. 위에서 보았듯이 “매우 높은 정도”라는 의미에 해당하는 어휘함수 Magn은 많은 단어에 적용되어 많은 함수값을 낳기 때문에 대표적인 표준 어휘함수로 분류된다. 반면에 영어에서 “썩은, 부패한”의 의미를 나타내는 어휘함수는 비교적 적은 수의 단어에 적용되고 함수값도 많지 않으므로 비표준 어휘함수의 예라 할 수 있다. 표준 어휘함수는 많은 언어에 나타나고換言(paraphrasing) 규칙에서 중요한 역할을 하지만,³ 비표준 어휘함수는 특정 언어에 한정되어 나타나고 환언 규칙에도 참여하지 않는다. Mel'čuk은 대부분의 언어에 보편적으로 나타나는 표준 어휘함수가 60여 개에 이를 것으로 보고 있다.

둘 이상의 어휘함수가 결합되어 사용될 수도 있다. 反義관계를 나타내는 어휘함수 Anti와 Magn이 합성된 AntiMagn은 “낮은 정도”를 의미한다. AntiMagn과 같이 둘 이상의 어휘함수가 결합된 것은 복합 어휘함수(complex lexical function), Anti나 Magn 같이 하나만 홀로 쓰인 것은 단순 어휘함수(simple lexical function)라고 한다. 단순 어휘함수 중에는 혼자서는 실현되지 않고 항상 복합 어휘함수의 일부로만 실현되는 것도 있다. 강도의 강화, 약화를 뜻하는 Plus, Minus, 기동, 지속, 종결 등 사태의 相의 국면을 나타내는 Incep, Cont, Fin 등이 그 예이다.

3. 환언 규칙에 대해서는 Mel'čuk(1988b, 1988c) 참고.

복합 어휘함수는 수학적 의미의 합성함수와는 성격이 다르다. 수학에서는 $f(g(L))=f(g(L))$ 이 성립되나 복합 어휘함수의 경우에는 반드시 이 등식이 성립되는 것은 아니다. 예를 들어 $\text{IncepOper}_1(\text{'협상'})=\text{'들어가다'}$ 이고 $\text{Oper}_1(\text{'협상'})=\text{'별이다, 하다'}$ 이지만 $\text{Incep}(\text{'별이다'})=\text{'들어가다'}$ 의 등식은 성립되지 않는다.

복합 어휘함수와 좀 다른 것으로 어휘함수의 배열(configuration of lexical functions)이란 것이 있다. 서로 통사적으로 관련되어 있지는 않지만 동일한 키워드를 취하는 두 어휘함수의 결합으로서, 해체할 수 없는 하나의 의미를 들이 함께 나타낸다. 다음과 같은 예를 들 수 있다.

Magn+A(병) = 위독한

Magn+Manif(기쁨) = 넘치다

한편 어휘함수에 의해 맺어지는 두 단어 사이의 관계가 계열관계나 통합관계나에 따라 계열적 어휘함수와 통합적 어휘함수로 나누기도 한다. 반의관계를 나타내는 Anti, 동의관계를 나타내는 Syn 등은 계열적 어휘함수이고 Magn은 통합적 어휘함수의 예이다. 지금까지 어휘론은 계열관계에 대한 연구에 치중해 왔다. 그래서 유의 관계, 반의관계, 상의·하의관계 등의 계열적 의미 관계에 대해서는 많은 연구가 이루어졌지만 통합관계에 대한 어휘론적 연구는 많지 않다. Mel'čuk의 이론, 특히 어휘함수는 통합관계에 대한 체계적 기술을 가능하게 했다는 점에서 매우 큰 의의를 지닌다.

Mel'čuk의 어휘 기술에서 가장 중요한 역할을 하는 것은 단순 표준 어휘함수라 할 수 있다. 이들은 여러 언어에서 반복해서 나타나므로 언어간 어휘 대조 연구나 기계번역 프로그램의 개발과 같은 목적에도 유용하게 이용될 수 있다.

(4) 단순 표준 어휘함수: 어휘요소 L에 대해 어휘요소들의 집합 $f(L)$ 을 연결지어 주는 함수 f로서; 다음 네 조건을 만족시킨다.

-
4. 여기서 어휘함수의 함수값에 대해 그냥 '하나의 어휘요소'라고 하지 않고 '어휘요소들의 집합'이라고 한 것에 대해 설명이 필요할 듯하다. 어휘함수의 가장 전형적인 경우는, 키워드가 주어지면 그에 대한 함수값이 하나의 어휘요소로 되는 경우이다. 즉 키워드와 함수값 사이의 관계가 1대1 또는 다대1이 되는 것이 가장 이상적인 경우이다. 그리고 그래야만 함수의 정의에 부합된다. 그러나 실제로 자연 언어의 어휘 관계를 살펴보다 보면 하나의 키워드에 대해 둘 이상의 어휘요소를 어휘함수의 함수값으로 주는 것이 좋은 경우가 많이 발생한다. 예를 들어 '도망을 가다'나 '도망을 치다'나 '도망을 하다'가 인지적 의미(cognitive meaning)에 있어서 별다른 차이가 없다고 할 수 있으므로 '가다, 치다, 하다'를 모두 'Oper(도망)'의 함수값으로 주고 싶어지는 것이다. 그러나 그렇게 하면 함수의 일반적 정의에서 벗어나게 된다. 이런 문제를 해결하기 위해 Mel'čuk이 선택한 방안은 어휘함수의 함수값을 그냥 '어휘요소'라고 하지 않고 '어휘요소들의 집합'이라고 한 것이다. 이렇게 하면 하나의 키워드에 대해 함수값이 둘 이상의 어휘요소인 경우를 포괄할 수 있게 된다. 그리고 하나의 원소만으로 이루어진 집합도 가능하기 때문에 함수값이 하나의 어휘요소인 경우에도 아무 문제가 발생하지 않는다. 본고에서 함수값을 제시할 때 집합 표시를 하지는 않지만(그리고 Mel'čuk의 논고에서도 집합 표시는 사용하지 않고 있다) 엄밀하게는 집합 기호가 있는 것으로 생각해야 할 것이다.

- ㉠ 두 어휘요소 L_1 과 L_2 의 모든 쌍에 대해, L_1 과 $f(L_1)$ 의 관계는 L_2 와 $f(L_2)$ 의 관계와 동일하다.
- ㉡ 일반적으로 $f(L_1) \neq f(L_2)$
- ㉢ 함수 f 의 키워드 수가 많아야 한다.
- ㉣ 함수 f 의 함수값 수가 많아야 한다.

6. 단순 표준 어휘함수의 예시

6.1. 계열적 어휘함수

1. Syn [synonymum]: 동의어.

Syn(결혼)=혼인 Syn(미워하다)=증오하다 Syn(어리둥절하다)=얼떨떨하다
 Syn(사랑)=애정 Syn(쓰다)=사용하다 Syn(맑다)=환하다

의미 영역이 완전히 일치하지 않는 擬似 동의관계의 경우 집합 관계 표시를 이용하여 Syn_{\supset} , Syn_{\subset} , Syn_{\cap} 와 같이 나타낸다. $Syn_{\supset}(X)=Y$ 는 Y의 정의가 X의 정의를 포함한다는 것(즉 Y의 의미성분이 더 풍부하다는 것, 다시 말해 Y가 더 특정한 의미를 나타낸다는 것)을 의미하며, $Syn_{\subset}(X)=Y$ 는 Y의 정의가 X의 정의에 포함된다는 것을 의미하며, $Syn_{\cap}(X)=Y$ 는 X의 정의와 Y의 정의가 부분적으로 일치함 [intersecting]을 의미한다.

Syn_{\supset} (좋아하다)=사랑하다 Syn_{\subset} (대화하다)=이야기하다 Syn_{\cap} (부엌)=주방

2. Conv [conversivum]: 상대관계에 있는 단어. 의미상의 논항들의 지위가 바뀌는 양상을 숫자로 표시.

Conv₂₁(남편)=아내 Conv₂₁(이기다)=지다 Conv₃₂₁(주다)=받다 Conv₂₃₁(평가)=평판

3. Figur [figuraliter]: 같은 의미를 나타내는 상투적인 비유적 표현.

Figur(침체)=늪 [침체=침체의 늪] Figur(홍분)=도가니 [홍분=홍분의 도가니]
 Figur(비난)=화살 [비난=비난의 화살] Figur(절망)=구렁텅이 [절망=절망의 구렁텅이]

4. S_0 , V_0 , A_0 , Adv_0 [영 substantival, verb, adjective, adverb]: 어휘적 의미가 같은, 대응하는 명사, 동사, 형용사, 부사.

S_0 (죽이다)=살해 A_0 (지혜)=지혜롭다 V_0 (승리)=이기다 Adv_0 (빠르다)=빨리

5. [] 안에 제시한 것은 어휘함수의 이름의 출처이다. 대개 라틴어나 프랑스어와 영어도 몇 개 있다.

6.1.1. 명사 어휘합수

5. S_i [영 substantival]: 표제항이 가리키는 상황의 i 번째 의미상의 논항을 나타내는 표준적인 이름.

S_1 (가르치다)=교사 S_2 (가르치다)=과목 S_3 (가르치다)=학생
 S_1 (이야기)=화자 S_2 (이야기)=내용 S_3 (이야기)=청자
 S_1 (팔다)=상인 S_2 (팔다)=상품 S_3 (팔다)=고객 S_4 (팔다)=가격

6. Sing [singulus]: 표제항의 하나의 예, 단위.

Sing(합대,선단)=배 Sing(의회)=의원 Sing(내각)=각료
Sing(당)=당원 Sing(총격)=발 Sing(나무)=그루 Sing(책)=권

7. Mult [multum]: 표제항들의 집합체. Mult와 Sing은 서로 역합수 관계에 있음.

Mult(위원)=위원회 Mult(사절)=사절단 Mult(군인)=군대

6.1.2. 형용사 어휘합수

8. A_i [영 adjective]: 표제항이 가리키는 상황의 i 번째 논항의 (상황 내에서의) 역할을 정의해 주는 속성을 나타내는 형용사.

A_1 (즐기다)=즐겁다 A_2 (속썩이다)=속상하다 A_2 (괴롭히다)=괴롭다

9. $Able_i$ [habilis]: 가능한 i 번째 의미상의 논항을 정의해 주는 일반적 속성.

$Able_1$ (타다)=可燃性이다 $Able_1$ (발생하다)=가능하다 $Able_1$ (웃기다)=우습다
 $Able_2$ (놀라다)=놀랍다 $Able_2$ (이해하다)=이해할 만하다
 $Able_2$ (무서워하다)=무섭다 $Able_2$ (존경하다)=존경스럽다
 S_1Able_1 (올다)=올보 $S_1Able_1, Caus$ (웃다)=코미디언

10. $Qual_i$ [qualitas]: 蓋然的인 i 번째 논항을 정의해 주는 질적 속성. $Qual_i$ 는 $Able_i$ 의 개연성을 함의하나 그 逆은 성립하지 않음.

$Qual_1$ (올다)=슬프다 $Qual_1$ (웃다)=즐겁다 $Qual_2$ (놀라다)=색다르다,유별나다
 $Qual_2$ (이기다)=약하다 $Qual_2$ (이해하다)=당연하다,명백하다

6.2. 통합적 어휘합수

6.2.1. 형용사 어휘합수

11. Magn [magnus]: “높은 정도로”, “높은 정도의”를 의미하는 수식어.

Magn(발전)=눈부신 Magn(경쟁)=치열한,불꽃튀는 Magn(지지)=열렬한,뜨거운
Magn(戰果)=혁혁한,빛나는 AntiMagn(목소리)=기어들어가는

12-13. Plus, Minus : “더한 정도로”, “덜한 정도로”. 이들은 주로 복합 어휘함수의 일부로만 사용됨.

IncepPredPlus(가격)=치솟다, 급등하다, 폭등하다

IncepPredMinus(기온)=급강하하다, 뚝 떨어지다

CausPredMinus(규제)=완화하다

14. Bon [bonus]: 표제항에 대해 기대되는, (화자의 관점에서 볼 때) 좋은 속성.

Bon(제안)=매력적이다, 그럴듯하다

Bon(조건)=유리하다

Bon(음식)=맛있다

Bon(정리하다)=말끔히, 깨끗이

6.2.2. 부사 어휘함수

15. Adv₁ [영 adverb]: 표제항이 가리키는 상황의 i번째 논항의 (상황 내에서의) 동작을 정의해 주는 속성을 나타내는 부사.

Adv₁(질주하다)=빠른 속도로

Adv₁(폭락하다, 폭등하다, 급락하다, 급등하다, 급강하하다)=큰 폭으로

Adv₁(변모하다, 달라지다, 성장하다, 발전하다)=몰라보게

Adv₂(급파하다)=갑자기, 급작스럽게

6.2.3. 동사 어휘함수

16. Pred [praedicatum]: 표제항임을 나타내는 동사. Pred(L)=L+Copul(L).

Pred(이웃)=이웃하다 Pred(독재자)=독재하다 Pred(폭군)=폭정을 휘두르다

Pred(교사)=가르치다 Pred(효자, 효녀)=효도하다

17-19. 기능동사

17. Oper₁ [operari]: 술어명사인 표제항을 목적어(또는 제1 보어)로 갖고 표제항의 i번째 논항을 주어로 갖는 기능동사.

Oper₁(결론)=내리다, 짓다, 내다

Oper₁(전화)=하다, 걸다

Oper₁(노력)=하다, 기울이다, 경주하다

Oper₁(도망)=하다, 치다, 가다, 오다, 다니다

Oper₁(성공)=하다, 거두다

Oper₁(실패)=하다, 돌아가다

IncepOper₁(토의, 심사, 심의, 협상)=들어가다

IncepOper₁(절망)=빠지다

IncepOper₁(파업)=돌입하다, 들어가다

Oper₂(공격, 따돌림)=받다, 당하다

Oper₂(비웃음, 부러움, 미움)=사다

Oper₂(의심, 의혹)=받다, 사다

Oper₂(잔소리, 걱정, 꾸중, 핀잔)=듣다

Oper₂(저항)=받다, 부딪히다

Oper₂(조사, 검사, 진단, 비난, 놀림, 버림)=받다

Oper₃(약속)=받다, 받아내다

18. Func_i [functionare]: 술어명사인 표제항을 주어로 갖고 표제항의 i번째 논항을 목적어(또는 제1 보어)로 갖는 기능동사.

Func₀(결론)=나다,나오다

Func_{2i}(명령,지시)=내려오다,떨어지다

Func₀(침묵)=흐르다

Func_{2i}(연락,전화,무전,전보)=오다

Func_i(부패,부정)=만연하다

MagnFunc_{2i}(비난,질문)=쏟아지다 MagnFunc_{2i}(문의,전화)=뱃발치다,쇄도하다

19. Labor_{ij(k)} [laborare]: 술어명사인 표제항을 제2 보어로 갖고 표제항의 i번째 논항을 주어로 갖고 표제항의 j번째 논항을 목적어(또는 제1 보어)로 갖는 기능동사.

Labor₁₂(고려,계산)=넣다

Labor₁₂(염두,기억)=두다

Labor₁₂(심사,심의,토의,토론,표결)=부치다

20-22. 실현 동사

20. Real_i [realis]: 표제항 명사를 목적어(또는 제1 보어)로 가지면서 표제항의 요구·목표를 실현함을 나타내는 동사. 표제항의 i번째 논항을 주어로 가짐.

Real₁(목적,목표)=달성하다,이루다,성취하다 Real₁(소원)=풀다,성취하다

Real₁(약속)=지키다,이행하다

Real₁(주장)=증명하다,입증하다

Real₂(명령)=이행하다,수행하다,따르다

Real₂(요구)=만족시키다,충족시키다,이행하다

Real₂(법,법규,규제)=지키다,따르다

Real₂(시험,테스트)=붙다,합격하다

21. Fact_i [factum]: 표제항 명사를 주어로 가지면서 표제항의 요구·목표가 실현됨을 나타내는 동사. 표제항의 i번째 논항을 제1 보어로 가짐.

Fact₀(꿈,희망,소망)=실현되다,이루어지다

Fact₀(영화)=상영되다

Fact₀(구름,먹구름)=끼다

Fact₁(신)=내리다

Fact₁(귀신)=들리다

22. Labreal_{ij} [Labor+Real]: 표제항 명사를 제2 보어로 가지면서 표제항의 요구·목표를 실현함을 나타내는 동사. 표제항의 i번째 논항을 주어로 갖고 j번째 논항을 제1 보어로 가짐.

Labreal₁₂(絞首臺)=매달다

Labreal₁₂(칼,톱)=자르다,썰다

23-25. 相 동사. Incep [incipere], Fin [finire], Cont [continuare]: 사태의 시작, 끝, 계속의 相의 국면을 나타냄.

Incep(있다)=생기다, 발생하다	IncepOper ₁ (이야기, 말)=걸다, 붙이다
Fin(P) = Incep(nonP)	FinOper ₁ (오해)=풀다
FinOper ₁ (희망)=없다, 버리다	FinOper ₂ (신임, 신용)=없다
Cont(P) = nonFin(P) = nonIncep(nonP)	
ContOper ₁ (침묵)=지키다	ContOper ₁ (침착성)=유지하다

26-28. 사동사

26. Caus [causare]: 사태의 발생을 야기함.

Caus(자다)=재우다	CausIncep(있다)=만들다, 창조하다, 세우다, 건립하다
CausOper ₁ (절망)=빠뜨리다	CausOper ₁ (희망)=주다, 붙여넣다
CausFinOper ₁ (희망)=빼앗다	CausOper ₂ (공격, 침략)=부르다, 초래하다

27. Liqu [liquidare]: 사태가 중지되도록 함.

Liqu(P) = Caus(nonP)	Liqu(있다)=없애다
LiquFunc ₁ (부패, 부정)=일소하다, 척결하다	

28. Perm [permittere]: 사태가 발생하는 것을 방해하지 않음.

Perm(p) = nonLiqu(p) = nonCaus(nonP)

PermOper₁(공격, 침략)=허용하다

PermOper₁(입장, 판매)=허락하다, 허가하다, 용인하다, 묵인하다, 못본체하다

notPermOper₁(입장, 판매, 공격, 침범)=불허하다, 금지하다, 막다, 방해하다

29. Degrad [degradare]: 표제항 명사를 주어로 취하면서 표제항의 상태가 나빠짐을 의미하는 동사.

Degrad(음식)=상하다	Degrad(옷)=헤지다, 헐다	Degrad(꽃)=시들다
Degrad(기강)=헤이해지다	Degrad(기억력)=쇠퇴하다, 떨어지다	
Degrad(달걀)=푹다	Degrad(색)=바래다	Degrad(쇠)=녹슬다

30. Son [sonare]: 표제항 명사의 특징적 소리가 남을 나타내는 동사.

Son(개)=짖다	Son(바람)=윙윙거리다	Son(시계)=똑딱거리다
Son(돼지)=꿀꿀거리다	Son(자동차, 엔진)=붕붕거리다	
CausSon(자동차, 엔진)=붕붕거리다		

7. 어휘함수의 활용

7.1. 어휘함수의 개념과 목록에 대한 수정

Mel'čuk은 단순 표준 어휘함수가 모든 언어의 어휘 체계에서 보편적으로 나타나 는 매우 일반적인 어휘 관계를 묘사하는 데 유용하게 쓰일 것이라는 의견을 피력한 바 있다. 따라서 단순 표준 어휘함수가 한국어 어휘 기술에의 적용 가능성이 가장 높을 것이라고 예상할 수 있다. 60여 개에 이르는 단순 표준 어휘함수의 목록을 검토 해 보면 대부분은 한국어의 어휘 관계 기술에도 그대로 사용될 수 있음직하나, 한국 어에서는 별로 쓸모가 없을 듯한 것도 있다. 예를 들어 전치사를 함수값으로 갖는 어휘함수인 Loc_{in} , Loc_{ab} , Loc_{ad} , Loc_{tempin} , $Instr$, $Propt$ 등은 한국어에서 함수값이 다양하지 못하기 때문에 표준 어휘함수의 자격이 있다고 보기 어려우며, 전치사나 조사가 문법적 성격이 강한 요소임을 상기할 때 어휘함수로 다룰 만한 성질의 것이 아니라는 생각도 든다.

한편 Mel'čuk은 표준 어휘함수의 목록에 포함시키지 않았지만 한국어 어휘 기 술을 위해서는 유용하다고 생각되는 어휘함수를 추가할 수도 있을 것이다. 그런 후보 를 몇 개 들어 보면 다음과 같다.⁶

1. Son_{adv} : 키워드가 지시하는 동물이나 물건이 발하는 특징적인 소리를 나타내는 의성어를 키워드와의 대응 관계 속에서 형식화하는 어휘함수. 한국어에 풍부한 의성어의 일부를 체계적으로 묘사하기 위해 Son 과 구별하여 설정할 필요가 있을 것이다.

Son_{adv} (개)=멍멍	Son_{adv} (귀뚜라미)=귀뚜라미	Son_{adv} (개구리)=개골개골
Son_{adv} (까치)=까각	Son_{adv} (눈)=뿌드득	Son_{adv} (돼지)=꿀꿀
Son_{adv} (소)=음매	Son_{adv} (시계)=똑딱똑딱	Son_{adv} (전화)=따르릉
Son_{adv} (종)=맵맵	Son_{adv} (쥐)=찍찍	Son_{adv} (차)=부르릉
Son_{adv} (참새)=쩍쩍	Son_{adv} (호랑이)=어흥	

2. $Modul$ [modulari]: 악기를 연주하는 행위의 동사와 악기명사 사이의 관계를 표시.

$Modul$ (가야금)=뜯다	$Modul$ (거문고)=타다	$Modul$ (기타)=치다
$Modul$ (나팔)=불다	$Modul$ (바이올린)=켜다	$Modul$ (북)=치다
$Modul$ (피리)=불다	$Modul$ (피아노)=치다	

6. 다음의 세 어휘함수(1~3)에 대해서는 홍계성(1995b) 참고.

3. Indu [induere]: 몸에 착용하는 행위의 동사와 착용하는 의복, 장신구 등의 키워드 명사 사이의 관계를 나타냄.

Indu(귀걸이)=달다,하다	Indu(백타이)=매다,하다	Indu(꽃)=달다, 꽃다
Indu(모자)=쓰다	Indu(목걸이)=하다	Indu(반지)=끼다
Indu(시계)=차다	Indu(신)=신다	Indu(안경)=쓰다,끼다
Indu(양말)=신다	Indu(옷)=입다	Indu(장갑)=끼다
Indu(칼)=차다	Indu(허리띠)=매다, 하다	

한편, 한국어에는 의성의태어가 매우 풍부하게 존재하는데, 이들 중 일부는 특정 용언(또는 특정 언어 표현)과 공기하는 특성이 있다. 예를 들어 ‘칭칭 감’, ‘빙빙 돌’, ‘사뿐사뿐 건’, ‘강충강충 뛰’, ‘줄줄 흐르’, ‘꿀꺽 삼키’, ‘펼펼 꿇/꿇이’, ‘버럭 소리를 지르’, ‘버럭 화를 내’, ‘둘둘 말’, ‘눈물이 핑 돌’ 등이 그런 예이다. 따라서 이들 의성의태 부사와 용언 사이의 관계는 한국어에서 매우 대규모로, 그리고 체계적으로 존재하는 어휘 관계라 할 수 있으므로 이것을 어휘부에서 어떻게 표상해 줄 것인가가 문제가 된다. 그런데 이들 관계는 의미가 경우에 따라 다르므로 엄밀히 말하면 본래적인 의미의 어휘함수로 기술할 수 있는 관계는 아니다. 그러나 한국어에 풍부하게 존재하는 이들 관계를 어떻게든 표상해 주어야 하므로, 어휘함수의 개념을 조금 느슨하게 적용하여 이들을 기술할 수도 있을 것이다. 예를 들어 의태 부사와 이의태 부사와 전형적으로 공기하는 용언 사이의 관계를 나타내기 위해 Shape이라는 어휘함수를 설정하고, 의성 부사와 용언 사이의 관계를 나타내기 위해 Sound라는 어휘함수를 설정할 수 있다.⁷

4. Shape: 의태 부사와 이 의태 부사가 수식하는 용언 사이의 관계를 표시.
 Shape(칭칭)=감다 Shape(빙빙)=돌다 Shape(사뿐사뿐)=건다
 Shape(강충강충)=뛰다 Shape(펼펼)=꿇다, 꿇이다 Shape(둘둘)=말다
 Shape(버럭)=소리를 지르다, 화를 내다 Shape(핑)=돌다, 눈물이 돌다
 Shape(철철)=넘치다
5. Sound: 의성 부사와 이 의성 부사가 수식하는 용언 사이의 관계를 표시.
 Sound(줄줄)=흐르다 Sound(꿀꺽)=삼키다

이들 관계는 표준 어휘함수 중 Adv로 기술할 수 있을 듯도 하다. 그러나 의성의태 부사의 특수성을 감안하여 별도의 함수를 마련하는 것이 좋다고 판단된다. Adv의

7. 의성의태어와 피수식 용언 사이의 관계를 어휘함수로 표상하는 문제는 21세기 세종계획 전자사전 본과의 언어 사전 소과제 회의(소과제 책임자: 임홍빈)에서 논의된 사항을 많이 참고했음을 밝혀 둔다.

경우 그 부사가 용언의 몇 번째 논항과 관련되는지를 숫자로 표시해 주는데, 이 기제를 Shape이나 Sound에 도입할 수는 있을 것이다. 예를 들어 'Shape(필필)=끓다', 'Shape(필필)=끓이다' 같은 식이다. 또한 Adv의 경우 동사나 명사가 키워드이고 부사가 함수값이 되는데, Shape과 Sound에서는 반대로 되어 있다. 어느 방법이 더 나은가는 좀더 연구가 필요한데, 의성의태 부사가 주어지면 그것과 공기하는 용언은 하나 또는 몇 개로 제한되는 데 비해, 어떤 용언이 주어졌을 때 그와 공기할 수 있는 의성의태 부사의 집합은 열려 있고(사뿐사뿐 건-, 성큼성큼 건-, 터벅터벅 건-, ……) 이들 각각의 의성의태어가 피수식 용언과 맺는 의미 관계도 동일하지가 않다. 실제로 사전에서 어휘함수를 이용하여 어휘관계를 표상할 때에, 함수값이 너무 많으면 일일이 나열해 주기가 어렵지만 몇 개의 어휘요소로 한정되어 있으면 쉽게 나열해 줄 수 있다. 직관적으로 봐도, '사뿐사뿐'이라는 부사가 어떤 용언과 결합되는가 하는 것은 '사뿐사뿐'이라는 어휘요소에 관한 중요한 정보이고 따라서 사전의 '사뿐사뿐' 항목에서 이것을 제시해 주어야 할 것으로 생각되지만, '건다'가 어떠한 의미태 부사와 결합할 수 있는가 하는 것은 '건다'에 관한 그만큼 중요한 정보라고 생각되지는 않는다. 따라서 의성의태부사를 키워드로 하고 이들의 수식을 받는 용언을 함수값으로 주는 것이 실용적인 관점에서 더 낫다고 생각된다.

Mel'čuk이 제시한 어휘함수의 본래적 정의에는 딱 부합되지 않지만 특정 언어의 특정 어휘 관계를 표상하기 위해 좀 느슨하게 어휘함수의 개념을 적용한 사례는 본고 말고도 다른 데서도 찾아볼 수 있다. 예를 들어 Fontenelle(1997)에서는 Gener의 역함수로 Spec(상위어를 키워드로, 하위어들을 함수값으로 갖는 어휘함수)을, 그리고 전체-부분 관계를 나타내기 위해 Part를 상정하고 있다. 예를 들면,

Spec(새)=참새, 제비, 딱따구리, 까치, 비둘기, ……

Part(자동차)=브레이크, 엔진, 트렁크, 범퍼, ……

와 같은 식이다. Mel'čuk은 어느 학술 회의에서 왜 부분관계(meronymy)를 나타내는 어휘함수를 설정하지 않았느냐는 질문에 대해 부분 관계의 경우 하나의 어휘함수로 포괄하기에는 너무나 이질적인 의미 관계들이 뒤섞여 있기 때문이라고 답한 바 있다. 실제로 흔히 거칠게 부분관계로 통칭되는 것들 속에는 '물리적 부분(physical part)', '기능적 요소(functional element)', '구성원(member)', '차원적 끝(dimensional edge)' 등 많은 것들이 들어 있는 것이 사실이다. 따라서 앞으로의 연구에 따라 Part라는 하나의 어휘함수 대신 여러 개의 어휘함수를 설정하는 것이 더 좋을지도 모른다. 그래도 아무런 조치를 취하지 않는 것보다는 잠정적으로라도 Part 같은 어휘함수를 설정하여 실제의 어휘 관계 기술에 활용하는 것이 좋다고 생각된다.

7.2. 기계 번역에서의 활용

2.1.2.에서 어휘 관계를 잘 표상해 놓은 어휘 데이터베이스가 기계 번역에서 꼭 필요하다라는 것을 지적했었는데, 어휘함수를 이용하여 이 문제를 어떻게 해결할 수 있

는지 살펴보겠다.

예문 (2) 같은 경우 “상하다”라는 의미는 표준 어휘함수는 아니지만 ‘egg’, ‘fish’, ‘cheese’ 등 상할 수 있는 음식물을 가리키는 명사에 대해서는 사전에 이것이 상했음을 나타내는 형용사가 무엇인지를 밝혀 주어야 할 것이다. 이것은 각 명사의 어휘내향에 이에 해당하는 어휘함수(예를 들어 Degrad)를 통해 나타낼 수 있다. 이것은 영어 문장 생성시에 유용하게 이용될 것이다.

예문 (1) 같은 경우는 한국어 문장 분석시에 발생하는 문제이다. 이런 경우를 처리하기 위해서는 한국어의 모든 술어 명사에 대해 이것이 어떤 기능동사를 취할 수 있는지에 대한 정보를 각 술어 명사의 어휘내향에 제공해 주어야 할 것이다. 이것은 6.2.3.에서 살펴본 기능동사 어휘함수 및 실현 동사 어휘함수를 통해 나타낼 수 있다. ‘결론’의 경우 ‘Oper(결론)=내리다, 짓다’, ‘Funcd(결론)=나오다’와 같은 정보를 ‘결론’의 어휘내향에 주고, 입력 문장에 ‘결론’ 같은 술어 명사가 들어 있으면 우선 ‘결론’이 기능동사 어휘함수의 함수값에 해당하는 단어와 共起했는지 여부를 점검해야 할 것이다. 기능동사와 공기하지 않았으면 술어 명사를 그냥 영어의 명사로 번역하면 될 것이고, 기능동사와 공기했으면 이 둘을 영어의 하나의 동사로 번역하면 될 것이다.

영한 기계번역에서 한국어 문장 생성시에 발생하는 많은 문제들도 어휘함수로 해결할 수 있다. 영어의 ‘put on’이나 ‘wear’를 한국어로 번역할 때 무조건 ‘입다’로 번역해서는 ‘양말을 입다’, ‘안경을 입다’ 같은 이상한 문장이 생성될 것이다. 신체 착용물을 나타내는 한국어 명사의 어휘내향에 앞에서 살펴본 Indu 어휘함수를 이용해, 표제항 명사에 대해 착용을 나타내는 동사가 어떤 것이 쓰일 수 있는지를 명시해 놓으면 이 정보를 활용해서 자연스러운 한국어 문장을 생성할 수 있을 것이다. “연주하다”의 의미를 나타내는 ‘play’의 번역도 똑같은 방식으로 이루어질 수 있다.

7.3. 정보 검색에서의 활용

정보 검색 시스템의 재현율을 향상시키기 위해서는 단순한 문자열 비교에 머무를 것이 아니라 사용자가 입력한 질의어와 밀접한 관련이 있는 다른 단어를 포함한 문서도 뽑아 줄 수 있어야 한다고 했다. 질의어와 밀접한 관련이 있는 단어들 중에서도 가장 우선적으로 고려해야 할 것은 동의어일 것이다. 이것은 어휘함수 Syn을 이용하여 해결할 수 있다. 질의어의 어휘내향에서 Syn의 함수값에 해당하는 단어들을 뽑아서 이 단어들을 추가 검색 단어 후보로 사용자에게 제시해 주면 될 것이다. 동의어 이외에 질의어와 어떤 관계에 있는 단어들이 추가 검색 후보로서 유용한지에 대해서는 경험적인 연구가 더 필요할 것이다. 이런 연구를 통해 정보 검색 시스템을 위해 유용한 새로운 어휘함수를 고안해 낼 수도 있을 것이다.

정보 검색 시스템의 정확률 향상을 위해서, 일단은 질의어를 포함한 문서들을 찾아 놓은 뒤 이들 중 추가적으로 어떤 단어를 포함한 문서를 추릴 것인지를 사용자에게

게 다시 물어보는 방법을 앞에서 언급한 바 있다. 이 때 애초의 질의어에 덧붙여 추가할 단어들도 다양한 유형이 있겠지만 가장 중요한 것은 질의어의 하위어(hyponym)일 것이다. 하위어에 대한 정보는 현재 단순 표준 어휘함수에서는 제공하고 있지 않다. 그러나 반대로 상위어에 대한 정보를 제공하는 Gener라는 어휘함수가 있다. 정보 검색 시스템에서는 우선 off-line에서 Gener를 이용해서 상위어-하위어 쌍들에 관한 정보를 뽑아 놓고 이로부터 상위어에 따라 색인을 만들어 놓으면 될 것이다. 또는 7.1에서 제시한 대로 Gener의 역함수로 Spec을 상정하여 이를 이용할 수도 있을 것이다. 상위어-하위어 관계 외에 어떤 어휘 관계가 정확률 향상에 도움이 될지도 앞으로 경험적인 연구가 더 필요할 것이다.

참 고 문 헌

- 이병근·박진호 (2000) '결합설명 사전의 어휘 기술 방법론,' 인문논총 43, 서울대학교 인문학연구소.
- 홍재성 (1995a) '어휘함수 개념에 의한 한국어 어휘기술과 사전편찬', 해방 50주년 - 세계 속의 한국학, 인하대학교 한국학연구소.
- 홍재성 (1995b) '의미·텍스트 대응모형에 따른 불어어휘의 연구: 어휘함수 개념의 활용을 중심으로', 인문논총 33, 서울대학교 인문학연구소.
- 홍재성·박동호 (2000) '멜축의 의미·텍스트 대응 모형 연구', 인문논총 43, 서울대학교 인문학연구소.
- Fellbaum(ed.) (1998) *WordNet: An Electronic Lexical Database*, Cambridge, The MIT Press.
- Fontenelle, T. (1997) *Turning a Bilingual Dictionary into a Lexical-Semantic Database*, Tübingen, Max Niemeyer.
- Mel'čuk, I. (1988a) *Dependency Syntax: Theory and Practice*, New York University Press.
- Mel'čuk, I. (1988b) 'Paraphrase et lexicque dans la théorie linguistique Sens-Texte', Vingt ans après, Ire partie, *Cahier de lexicologie* 52.1, 5-50.
- Mel'čuk, I. (1988c) 'Paraphrase et lexicque dans la théorie linguistique Sens-Texte', Vingt ans après, 2e partie, *Cahier de lexicologie* 53.2, 5-53.
- Mel'čuk, I. (1996) 'Lexical Functions: A Tool for the Description of Lexical Relations in a Lexicon', in L. Wanner, ed. (1996), 37-102.
- Mel'čuk, I. et al. (1984) *Dictionnaire explicatif et combinatoire du français contemporain: Recherches Lexico-Sémantiques I*, Montréal, Les Presses

- de l'Université de Montréal.
- Mel'čuk, I. et al. (1988) *Dictionnaire explicatif et combinatoire du français contemporain: Recherches Lexico-Sémantiques II*, Montréal, Les Presses de l'Université de Montréal.
- Mel'čuk, I. et al. (1992) *Dictionnaire explicatif et combinatoire du français contemporain: Recherches Lexico-Sémantiques III*, Montréal, Les Presses de l'Université de Montréal.
- Mel'čuk, I. et al. (1999) *Dictionnaire explicatif et combinatoire du français contemporain: Recherches Lexico-Sémantiques IV*, Montréal, Les Presses de l'Université de Montréal.
- Mel'čuk, I., A. Clas and A. Polguère (1995) *Introduction à la Lexicologie Explicative et Combinatoire*, Editions Duculot.
- Steele, J.(ed.) (1990) *Meaning-Text Theory: Linguistics, Lexicography and Implications*, University of Ottawa Press.
- Steele, J. and I. Meyer (1990) 'Lexical Functions in an Explanatory Combinational Dictionary: Kinds, Descriptions and English Examples', in J. Steele, ed. (1990).
- Wanner, L. ed. (1996) *Lexical Functions in Lexicography and Natural Language Processing*, John Benjamins Publishing Company.
- Wanner, L. ed. (1998) *Recent Trends in Meaning-Text Theory*, John Benjamins Publishing Company.

<부록> 명사 ‘눈’에 대한 결합 설명 사전 식의 記述 예시

명사 ‘눈’(眼)에 대해 DEC식으로 시험적으로 기술한 것을 아래에 제시한다. 어휘합수가 어휘 관계를 기술하는 데 어떻게 도움을 줄 수 있는지의 구체적인 예를 제시하기 위해서이다. 각 정보 구획에 대한 자세한 설명은 이병근·박진호(2000)으로 돌린다. 아래의 기술은 Mel'čuk et al.(1988)의 EIL 항목을 많이 참고하였다. Mel'čuk et al.(1988)의 EIL 항목에서와 마찬가지로 아래의 기술에서도 비표준 어휘합수가 많이 사용되었다. 비표준 어휘합수에 대해서는 적절한 이름을 붙이기가 어려우므로 자연 언어로 그 의미를 풀이한 것으로 이름을 대신하였다. 그리고 몇몇 경우에는 적절한 의미를 기술하기가 어려워 어휘합수 이름을 아예 생략한 곳도 있다.

눈

I.1a [Y를 보는 것을 가능하게 하는] X의 (Z한) 눈 = 사람 X의 얼굴의 두 부분. 이 두 부분은 코를 중심으로 하여 서로 대칭적이다. 각각은 희끄무레한 움직일 수 있는 구와 두 개의 수평적인 피부 주름으로 이루어져 있다. 희끄무레한 구의 앞 가운데에는 유색의 원이 있고 이 원 안의 가운데에는 검은 원이 있다. 두 개의 수평적인 피부 주름 중 위의 것은 움직일 수 있어서 희끄무레한 구를 덮을 수 있다. 이 구는 보는 것을 가능하게 하는 가상의 빛을 發하는 것으로 생각된다. (이 구는 X의 상태 또는 속성 Z를 나타낸다.) X의 보는 기관으로서, X가 Y를 보는 것을 가능하게 한다.

지배 패턴

I = X	II = Y	(III = Z)
N _{poss}	—	A _{adm}

- C_I : 철수의 눈, 내 눈
- C_{II} : 슬픈 눈, 우수에 젖은<찬> 눈
- C_I+C_{II} : 철수의 슬픈 눈

어휘합수

- Syn_▷ : 눈알, 안구
- 눈의 투명한 액체 분비물 : //눈물
- 많은 눈물을 포함하다 : 눈물이 글썽글썽하다
- 눈의 진득진득한 분비물 또는 그것이 말라붙은 것 : //눈곱
- 눈곱을 포함하다 : 눈곱이 끼다
- Labreal₁₂ : 보다
- 가리다
- 비비다

눈의 형태와 자원과 관련된 속성

- 크다
- 작다
- (특) 튀어나오다
- 푹 들어가다, 푹 꺼지다, 쑥 들어가다
- 짜지다, 찢어지다
- 실눈1
- 왕눈이

눈의 색깔과 관련된 속성

- 갈색, 담갈색
- 푸르다
- 검다

눈의 광채

초롱초롱하다
빛나다

눈의 감정 표현

홀기다, 찌리다, 찌려보다
노려보다
부라리다

Excess^{dim}(눈)-Sympt₁(놀라움) : 휘둥그
래지다

눈빛
눈웃음
눈초리

눈의 이동과 위치

Prepar : 뜨다
LiquiFacto : 감다
깜박이다, 꿈벅꿈벅하다
실눈2 뜨다
치뜨다
내리깔다, 깔다

눈을 함축하는 몸짓

눈짓
윙크
눈인사
눈싸움
눈대중, 눈짐작, 눈어림
눈속임

눈과 관련된 장애

Obstr : 침침하다, 무시다
붓다
충혈되다
계습츠레하다
하나의 눈만 가진 사람 : //애꾸, 애꾸눈,

애꾸눈이, 외눈박이

안질, 눈병
다래끼
백내장
녹내장

눈 돌보기

안대
안약
안경
콘택트렌즈
눈가리개
마스카라

눈의 부분

안구(眼球), 눈알, 눈깔
눈동자
눈망울
눈꺼풀
눈구멍, 안공, 안과, 안와, 안확
눈썹
속눈썹
눈두덩
눈탱이
눈언저리
눈꼬리
눈시울

눈높이
눈썰미
눈엣가시

I.1b [Y를 보는 것을 가능하게 하는] X
의 (Z한) 눈 = 동물 X의 머리 앞부분의
두 부분. 이것의 구조는 눈 I.1a의 구조
와 같다. (이것은 X의 상태 또는 속성 Z

를 나타낸다.) X의 보는 기관으로서, X가 Y를 보는 것을 가능하게 한다.

지배 패턴

I = X	II = Y	(III = Z)
N _{poss}	—	A _{adn}

C_I : 고양이의 눈, 파리의 눈
 C_I+C_{III} : 개의 충직한 눈

어휘함수

I.2 (Z를 표현하는) Y에 대한 X의 눈 = Y에 대한 X의 시선. (X의 상태 Z 또는 Y에 대한 X의 태도 Z를 나타낸다.)

지배 패턴

I = X	II = Y	(III = Z)
N _{poss}	N에 대한	A _{adn}

C_I : 사람들의 눈을 의식하다
 C_{II} : 자신에 대한 눈을 의식하다
 C_{III} : 원망 어린 눈
 C_I+C_{II} : 선생님에 대한 아이들의 눈이 곱지 않다
 C_I+C_{III} : 아이들의 원망 어린 눈
 C_{II}+C_{III} : 선생님에 대한 원망 어린 눈
 C_I+C_{II}+C_{III} : 선생님에 대한 아이들의 원망 어린 눈

어휘함수

Syn : 시선, 눈길
 눈초리
 Oper₁ : 주다, 두다 [눈 둘 곳을 모르겠다 / 눈을 어디다 두어야 할지], -에/에게 돌리다

AntiOper₁ : -에서/에게서 돌리다
 Oper₂ : 띄다
 AntiOper₂ : 피하다
 의식하다
 맞추다
 마주치다
 어리다
 어른어른하다
 선하다

II.1 X의 눈 = 사람 X가 멀리까지, 작은 것까지 정확하게 보는 힘.

지배 패턴

I = X
N _{poss}

C_I : 철수의 눈이 나빠졌다

어휘함수

Syn : 시력

눈의 기능

Bon : 좋다, 밝다
 S₁MagnBon : 천리안

눈의 기능 장애

AntiBon : 나쁘다, 어둡다
 FinFact₀ : 멀다
 S₁FinFact₀ : 장님, 소경, 맹인

근시
 원시
 난시
 노안
 색맹
 색약

야맹증

눈의 기능 장애의 치료

II.2 (Y에 대한) X의 눈 = 사람 X가 Y를 보고 판단하는 힘.

지배 패턴

I = X	II = Y
N _{poss}	N에 대한

- C_I : 나는 그의 눈을 믿는다
- C_{II} : 미술작품에 대한 눈이 있는 사람
- C_I+C_{II} : 미술작품에 대한 그의 눈을 믿는다

어휘함수

- Syn : 안목, 식견
- Pred : 있다
- AntiPred : 없다

III (Y에 대한) X의 눈 = 사람 X가 Y를 보고 판단하는 기준

지배 패턴

I = X	II = Y
N _{poss}	—

- C_I : 그의 눈에 들다

어휘함수

- Real₂ : 들다, 차다
- AntiReal₂ : 거슬리다
- Magn : 높다
- AntiMagn : 낮다

◇ 관용표현

- 눈이 맞다
- 눈이 멀다
- 눈을 붙이다
- 눈이 빠다
- 눈을 속이다
- 눈에 설다
- 눈에 익다
- 눈에 흠이 들어가다
- 눈에 넣어도 아프지 않다
- 눈 밖에 나다
- 눈 깜짝할 사이
- 눈도 깜짝 안 하다
- 눈 딱 감고
- 눈을 씻고 보다
- 눈에 불을 켜고
- 눈이 빠지게
- 첫 눈에 반하다
- 한 눈에 알아보다
- 한 눈에 들어오다
- 눈에 밝히다
- 눈 앞이 캄캄하다
- 눈이 시퍼렇게 살아 있다
- 눈을 속이다
- 눈에 쌍심지를 켜다
- 눈에 불이 나다
- 눈 가리고 아웅하다
- 눈을 의심하다
- 눈이 무섭다
- 눈이 별쨍다
- 눈이 튀어나오다
- 눈이 뒤집히다

복합어

- 눈감아주다
- 눈여겨보다

ABSTRACT

The Description of Lexical Relations in Korean Using Lexical Function

Lee, Pyong-Geun · Park, Jin-ho · Kim Jin-hyung

Recently, the task of describing the totality of lexical relations in a language has come to the fore. This paper discusses how the mechanism of lexical function, proposed by I. Mel'čuk, can be used to describe the lexical relations in Korean. In section 2, we present the necessity of describing the totality of lexical relations from practical viewpoints, i.e. information retrieval, machine translation, and language teaching. In section 3, we point out the limits of previous studies on lexical relations. In sections 4~6, we introduce the concept and typology of lexical function. In section 7, we discuss how lexical functions can be applied to concrete examples of Korean. In appendix, the lexical entry '눈(eye)' in the format of Explanatory Combinatorial Dictionary is presented.

이 병근

151-745

서울특별시 관악구 신림동 산 56-1
서울대학교 인문대학 국어국문학과
전자우편 : leepg@snu.ac.kr

김진형

151-745

서울특별시 관악구 신림동 산 56-1
서울대학교 인문대학 국어국문학과
전자우편 : quirk@chollian.net

박진호

151-745

서울특별시 관악구 신림동 산 56-1
서울대학교 인문대학 국어국문학과
전자우편 : synpjh@eoneo.co.kr