

용어관계의 분류 모형 개발에 관한 연구*

A Study on the Development of a Classification Model for Terminological Relationships

백지원(Ji-Won Baek)** 정연경(Yeon-Kyoung Chung)***

초 록

본 연구는 모든 지식조직체계의 근간인 용어관계가 동일성, 계층성, 연관성이라는 세 가지 포괄적인 기준에 의해 정의되어 사용됨으로써 정보의 정확성이 증시되는 오늘날의 정보 환경에서 제 기능을 다하지 못하고 있으므로, 그 해결 방안의 하나로 용어관계의 분류 모형을 제시하고자 하는데 목적이 있다. 이를 위해 기존의 여러 지식조직체계에 나타나는 각종 용어관계의 사례와 용어관계에 대한 이론적 연구들을 광범위하게 수집하여 다양한 용어관계의 유형을 파악하였다. 그리고 이를 바탕으로 용어관계를 명확하게 정의하고 범주화할 수 있는 용어관계 분류의 근거를 세우고 용어관계의 분류 모형을 개발하였다. 더 나아가 이 분류 모형을 정보검색을 비롯한 다양한 방면에 활용할 수 있는 방안을 모색하고 향후 용어관계 분류 연구에 대한 제언을 했다.

ABSTRACT

The purpose of this study is to present the limitation of terminological relationships in the current information environment and to propose a solution to result the richer and refined terminological resources. For this, various kinds of terminological relationships in knowledge organization systems and theoretical researches were collected and analyzed. Based upon the analysis, a methodology for classification of terminological relationships was suggested and classification models were presented. Additionally, four suggestions were made for the practical uses of the classification models.

키워드: 용어관계, 분류, 시소러스, 의미관계, 용어 관계, 분류모형, 지식조직

* 본 논문은 이화여자대학교 대학원 박사학위논문을 요약한 것임.

** 이화여자대학교 사회과학대학 문헌정보학 전공 강사(jw_baek@naver.com)

*** 이화여자대학교 사회과학대학 문헌정보학 전공 교수(ykchung@ewha.ac.kr)

1. 서론

1.1 연구의 목적

현행 지식조직체계에서 나타나는 용어관계는 특정한 원칙 없이 일관성 없이 설정됨으로써 색인·검색 등 정보처리에서 지식조직체계의 효율성 및 활용성을 떨어뜨리고 있으며, 지식조직체계간의 호환과 통합 등 상호교류를 어렵게 하는 원인이 되고 있다. 이를 개선하기 위해 기존에 관행적으로 지속되어 온 용어관계 설정의 근거와 본질을 재조명하고, 기존의 여러 지식조직체계에 실재하는 용어관계 유형의 양상을 파악하여, 이를 보다 명확하고 상세하게 범주화할 수 있는 용어관계의 분류 모형을 개발하고자 했다. 더 나아가 이렇게 개발된 모형을 정보검색을 비롯한 다양한 방면에 있어서 활용할 수 있는 방안을 모색하였다.

1.2 연구의 방법 및 범위

본 연구는 용어관계 분류 모형을 설계하고 그 활용 방안을 밝혔다. 분류 모형 설계를 위한 분류의 방법과 원칙을 세우고, 대표적인 용어관계의 범주인 동등, 계층, 연관관계에 있어서 각 범

주별로 관계성의 본질을 분석하고 각각의 특성에 맞게 용어관계의 분류 모형을 설계하였다. 용어관계의 분류모형 구축을 위하여 7가지 범주의 총 29종의 연구 사례 및 각종 지식조직체계로부터 실재하는 용어관계의 유형을 수집하여 ISO 2788-1986(E) *International Organization for Standardization Documentation: Guidelines for the Establishment and Development of Monolingual Thesauri*(이하 ISO 2788)의 세 가지 용어관계 범주로 범주화시킨 결과, 동등관계 138개, 계층관계 132개, 연관관계 454개, 총 724개의 유형을 분석할 수 있었다. 이러한 분석의 결과는 다음과 같이 요약할 수 있다(분석대상의 목록과 각각의 용어관계의 수는 부록으로 수록함).

- 실재하는 용어관계의 유형의 수는 매우 많고 다양하다.
- ISO 2788을 기준으로 봤을 때, 서로 유사한 경우도 있고 전혀 다른 경우도 있다.
- 동등, 계층, 연관관계 중 어디에 속하는지 불분명한 용어관계도 많다.
- 각 용어관계 내에서의 하위 구분도 불분명한 경우가 많다.

따라서, 용어관계의 양상에 대한 보다 상세한 분석을 바탕으로 용어관계

구분의 논리적 기반을 마련하고 이를 통해 체계화된 분류 구조를 제시할 필요가 있음을 파악할 수 있었다.

2. 용어관계의 분석

2.1 동등관계

동등관계는 언어학, 철학, 논리학 등 여러 영역에서 다방면의 연구가 진행되어 온 연구 영역이므로 용어관계의 유형도 다양하게 나타난다. 동등관계를 구성하는 다양한 용어관계의 유형들을 정리하면 동등관계는 크게 동의어(혹은 유의어)와 반의어의 두 가지로 나누어질 수 있으며 다음과 같이 분석할 수 있다.

첫째, 동의어 및 유의어와 관련된 문제이다. 동의어와 관련된 용어관계로는 ‘인접동의, 완전동의, 인지적 동의, 인지 동등, 문맥적 동의, 동일 지시물 동의, 동일 의미 동의, 참된 동의, 협의적 동의, 준 동의, 광의적 동의’ 등이 있다. 또 유사동의어와 관련해서는 ‘협의적 유의, 근사유의, 광의적 유의’ 등이 존재한다.

언어학적으로 유의어는 의미가 서로 비슷하나 가리키는 대상의 범위 차이나 쓰이는 상황의 차이로 미묘한 느낌의 차이를 보이기 때문에 같이 쓰일 수 없는 경우도 있다고 본다. 또한 동의어와 유의어는 자유롭게 바꾸어 쓸 수 있는

나 없느냐로 구분하기도 한다. 예를 들어, ‘책방/서점’은 동의어이고 ‘색별/금성’은 유의어라는 것이다. 그런데 동의관계와 유의관계는 서로 지칭하는 바가 다를 뿐, 동의성의 정도에 차이가 있는 것이라고 볼 수 있다. ‘완전한 동의관계’가 아닌 것은 사실상 유의하다고 할 수 있으므로 ‘동의어’와 ‘유의어’의 구분이 모호해진다. 일례로 TEST에서는 강한 유사동의관계는 동등관계로, 약한 유사동의관계는 연관관계로 보았는데, ‘강함’과 ‘약함’ 사이의 기준은 사실상 없다.

ANSI/NISO Z39.19-X *Guidelines for the Construction, Format, and Management of Monolingual Thesauri*의 개정안(2005)은 ‘Quasi-synonyms’를 ‘Plesionyms’와 동의어로 취급했다. Edmonds와 Hirst(2002)는 ‘Plesionyms’의 ‘Plesio-’가 ‘near’를 뜻하는 접두사라는 점을 들어 ‘Plesionyms’와 ‘Near synonyms’를 같은 것으로 간주하였다. ANSI Z39.19 개정안(2005)과 Edmonds와 Hirst(2002)의 견해를 종합하면 ‘Quasi-synonyms’, ‘Plesionyms’, ‘Near synonyms’는 모두 동일한 것으로 볼 수 있다. 그러나 ALA에서 작성한 주제관계 분류인 *Taxonomy of Subject Relationships. Appendix B. Part 1 and 2*(이하 ALA 분류)에서는 ‘Quasi-synonyms’를 동의어와 구분하여 반의(Antonyms), BT/NT 관련 관

계(BT/NT issue relationships), 인접 동의(Near synonyms) 등의 상위어(BT)로 간주하였다.

이처럼 연구 사례별로 동일한 용어관계에 대해 상하위 관계로 간주하기도 하고 동등관계로 간주하기도 하는 등 각기 다른 해석을 하고 있다. 결국 동의관계는 동의관계를 지칭하는 명칭 보다는 동의성에 차이를 가져오는 요인 자체에 초점을 맞춰야 할 것임을 알 수 있다.

둘째, 반의어의 문제이다. 언어학에서 반의어는 두 단어 사이에 공통적인 의미 요소가 있으면서 하나의 의미요소가 다른 것이다. 예를 들어, ‘남자/여자’는 사람이라는 점에서 공통적이지만 성별이라는 요소에서 차이를 보이므로 반의어이다. 이처럼 직관적으로 반의관계는 어휘적 대립쌍을 연결시킨다. 그러나 반의 관계가 대칭적 관계라는 사실을 제외하면, 광범위한 현상들을 포괄할 수 있는 ‘대립’이 의미하는 것이 정확히 무엇인지는 여전히 불분명하다(Cruse 1986). 따라서 WordNet 1.5에서는 반의 관계를 ‘단어 X의 반의어는 때로 not-X 이지만, 항상 그런 것은 아니다’와 같이 포괄적으로 정의했다(Miller, et al. 1990).

이러한 반의관계의 모호한 성격을 반영하듯이, 시소러스의 반의관계는 매우 다양한 유형으로 나타난다. ‘보완적 반

의어, 환위적 반의어, 인접 반의어, 매우 느슨한 반의어, 스칼라 반의어, 품사간 위사 반의어, 비동등 반대, 상반, 보충반의, 비등가 반의, 대립적 반의, 역전(reversals), 극성 대립, 관계 대립, 정도 대립, 상보성, 역, 상호친족관계’ 등 다양한 관계유형이 존재한다.

개별 용어관계의 정의가 내려진 사례가 드물고, ‘매우 느슨한 반의’와 같이 명칭에서부터 막연한 설정을 하고 있기 때문에 반의관계의 분류 기준을 기준 사례에서 찾는 것은 쉽지 않다. 따라서 정의와 예가 주어져 있는 사례들을 중심으로 분석하면 반의관계 유형은 <표 1>과 같이 동일하거나 매우 유사한 종류의 예들이 서로 다른 관계유형으로 지정되어 있는 것을 알 수 있다. 표 중의 점선은 유사하다고 판단되는 종류의 예들을 구분하기 위한 것이다.

반의관계의 분석 결과, 반의관계간의 구분이 불분명한 것은 물론, 유사동의어와의 구분도 명확하지 않은 것을 알 수 있다. 또한 반의관계간의 계층관계가 언급되고 있으나 이 또한 사례 간에 공통된 것이 없다.

2.2 계층관계

ISO 2788에서 계층관계를 구성하는 세 가지 범주 중 종/속관계와 사례관계는 상대적으로 범위가 명쾌하게 정해져 있어 용어관계의 판정이 쉽다.

그러나 전체/부분관계는 하위분류가 다양하게 존재할 뿐 아니라 연관관계와의 구분 또한 쉽지 않다.

ALA 분류에서는 크게 다음과 같이 네 가지 항목으로 계층관계를 나누고 있는데, 다음과 같이 이들 중 셋은 ISO 2788의 계층관계 분류와 정확하게 일치한다.

- 종/속 쌍(Genre/species pairs)
= BTG/NTG
- 부분 관계(Partitive relation)
= BTP/NTP
- 류/사례 쌍(Class/instance pairs)
= BTI/NTI
- 비포함 계층관계(Non-inclusion hierarchical relation)
- 해당되는 것 없음

ALA 분류의 계층관계의 경우 각 범주의 하위 수준이 ISO 2788의 계층적 전체/부분관계와 유사성을 보인다.

ISO 표준에서는 계층적 전체/부분

관계를 ‘신체의 체계나 기관’, ‘지리적 위치’, ‘학문분야’, ‘계층적 사회구조’ 등 네 가지로 나누고 있는데, 이들은 각각 ALA 분류의 ‘해부학적 전체/부분 쌍’, ‘지리적 전체/부분 쌍’, ‘주제포함-학문/하위 학문 쌍’과 유사한 개념 범주로서 대응된다. 마지막의 ‘계층적 사회구조’는 비포함 계층관계로 분류된 ‘조직적 보고(Organizational reporting)’와 내용상 유사한 것으로 볼 수 있다.

Lu(1990)는 자신의 연구에서의 ‘전체/부분관계’가 계층관계에 해당하는 것이라고 밝히고 있으나, ‘개체와 그 특징’, ‘재료’의 경우에는 ISO 2788이 명시하고 있는 계층관계의 요건 보다는 일반적인 연관관계의 성격과 더 유사한 것으로 판단되었다.

Wang, Vandendorpe, Evens(1985)의 연구에서 제시한 ‘물질’과 ‘기원’ 역시 계층관계 보다는 연관관계에 가까운 것으로 판단되었다.

<표 1> 예시를 통한 동등관계 비교

예시	용어관계(출처)
hardness/softness	상반관계(Mandersloot 외 1970)
smoothness/roughness	연속체의 한 지점을 나타내는 유사동의어(ANSI 2005) 동일한 특성의 연속체로 관점의 상이를 표현한 용어의 경우 (TEST 1967)
dryness/wetness	유사동의어(ISO 1986) 반의어(ANSI 2005) 보충관계(Mandersloot 외 1970)
husband/wife	상대적 보완(ALA 분류 1999) 상호 친족관계(Wang, Vandendorpe, Evens 1985)
male/female	대립적 반대(Mandersloot 외 1970)

Storey(2001)는 부분 관계를 구분하는 기준을 기능(function), 분리가능성(separable), 동질성(homogeneous), 동시발생(contemporaneous)이라고 하여 부분관계를 크게 일곱 가지로 나눈 후 다시 하위에서 세분하였다.

따라서 계층관계의 유형은 전체적인 체계로서 ISO 2788을 기반으로 삼고, 그 하위단계를 사례유형으로 분류할 수 있다는 가능성과 함께 연관관계와의 구분이 필요하다는 사실을 파악하였다.

2.3 연관관계

연관관계의 표준 및 실제 유형에서는 명확하게 정의된 연관관계의 유형이나 목록을 제시하는 대신 연관관계에 속할 수 있는 대표적인 유형을 제시하는 방식으로 연관관계에 대한 지침을 대신하고 있다. 그런데 앞선 유형 분석의 결과, ISO 2788과 ANSI Z39.19 등 시소러스 구축 표준을 구성하는 연관관계의 대표적인 속성 15가지가 전체 연관관계 유형 중 약 25%이고, 이들 15가지 속성과 유사한 개념으로 볼 수 있는 속성들까지 한데 모으면 30%를 약간 넘는다는 것을 파악했다. 따라서 70%에 가까운 연관관계는 이른바 ‘기타 연관관계’에 해당하는 것이므로, 기존에 지침으로 사용되어 온 연관관계의 대표 속성만으로는 연관관계를 효과적으로 분류하지 못한다는 사실을 알 수 있었

다.

연관관계에 대해 관련 표준에서 구체적으로 명시된 내용은 다음과 같이 정리할 수 있다.

- 연관관계로 연결되는 용어 간에는 강한 인지적 연상이 있어야 한다.
- 연관관계로 연결되는 용어의 선택에 대한 엄격한 통제를 수행하는 것과 주관적 판단을 피하는 것이 중요하다.
- 연관관계의 설정은 색인 이용자가 공유하는 참조의 틀에 따라야 한다.
- 두 용어 중 한 용어가 다른 용어의 설명이나 정의의 필수적인 구성성분일 때 연관관계로 연결할 수 있다.

앞의 두 가지 지침은 매우 막연한 일반론일 뿐이며, 세 번째 지침은 용어관계 설정의 목적이나 영역에 따라 크게 좌우되는 것이다. 네 번째 지침은 이중 가장 명확한 내용을 담고 있다. 그러나 이 방법만으로는 용어의 의미요소에 대한 분석으로는 부족하고, 참고자료의 신뢰성에 전적으로 의지할 수밖에 없으므로 연관관계 설정을 위한 단일 지침으로는 부족하다. 이처럼 일단 막연한 ‘인지적 연상’에 의해 두 용어사이에 연관성이 있음을 판단하고 나면, 이들 관계는 모두 ‘RT’라는 단일한 표지 밑으로 소속되게 되는 것이므로, 연관관계 설정 및 용어관계 상세화에 관한 보다 명확한 지침이 마련되어야 할

것임을 알 수 있다.

이상과 같이 연관관계의 유형 정의 및 범주화에 대한 문제제기가 계속적으로 있어왔고 그 결과 많은 연구가 있었으나 완전한 해답을 제시하고 있는 사례는 없다. 따라서 이러한 연구들을 기반으로 한 새로운 접근 방법이 필요하다고 사료된다.

3. 분류 모형의 설계

용어관계에 관한 표준 및 실제 유형을 분석한 결과, 동등관계와 계층관계는 서로의 같음과 다름에 대한 구별 없이 명칭을 달리한 다양한 유형의 용어관계가 혼란스럽게 사용되고 있다는 점에 문제가 있다는 것을 알 수 있었다. 따라서 여러 관계유형들을 하나의 단일한 체계로 정의하고, 필요에 따라 사용할 수 있는 상세한 유형 분류 방안을 마련하는 연구가 필요한 것으로 파악되었다.

용어관계의 정의 및 구분에 대한 논리적 근거와 신뢰할만한 유형 분류가 부재하다시피 한 연관관계에서는 여타의 용어관계와는 다른 접근법이 필요하다고 판단하였다. 연관관계의 분류를 위한 가장 현실적인 방법으로 패킷의 개념을 적용했다.

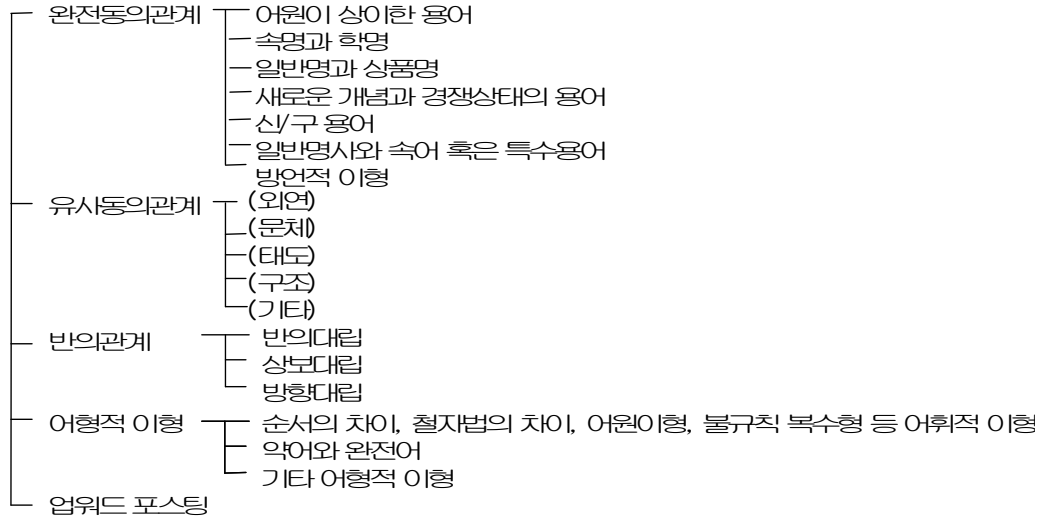
3.1 동등관계

동등관계는 동의성의 정도에 따라 완전 동등관계와 유사동의관계로 나누고, 유사 동의관계의 하위 분류가 필요한 경우에는 언어의 의미자질 분석에 기반한 패킷을 적용하는 방안을 제안하였다. 하위 분류를 위한 패킷으로는 언어학 분야에서 제안하고 있는 여러 유형의 분류 방식 중 가장 포괄적이고 명백한 기준에 의해 분석된 방식을 선택했다. <그림 1>은 이러한 동등관계의 의미적 분류와 함께 본 연구에서 직접적으로 다루지 않은 비의미적인 동등관계인 어형적 이형과 업워드포스팅(upward posting)까지 포함시켜 동등관계 범주의 전체 구성을 나타낸 것이다.

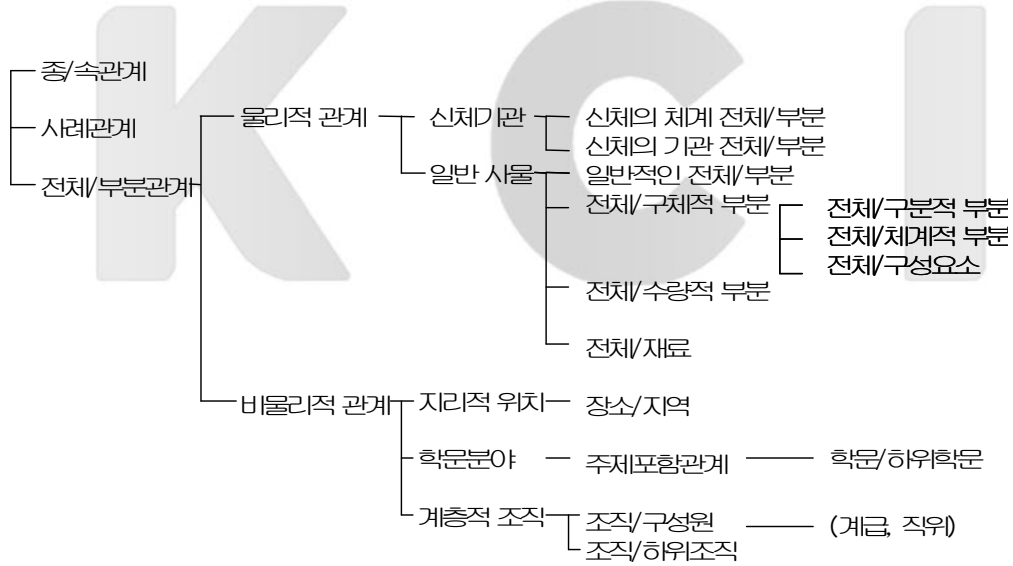
이러한 모형화를 통해 동등관계를 구성하는 다양한 유형을 명확하게 정리하였고, 그 내부적인 하위 분류를 위해서는 기존의 용어관계 유형이나 용어관계 범주 구분이 아닌 의미적인 분석에 의한 접근이 필요하다는 점을 밝힐 수 있었다.

3.2 계층관계

계층관계에 있어서는 계층관계에 속하는 용어관계들을 체계화된 조직으로 구조화한 계층관계 유형의 분류 모형을 제안하였다. <그림 2>는 계층관계에 있어서 본 연구가 주된 논의의 대상으로 삼은 전체/부분관계와 함께 중/속관계와 사례관계까지 포함하여 계층관계의 전



<그림 1> 동등관계의 용어관계 분류 모형



<그림 2> 계층관계의 용어관계 분류 모형

체적인 분류 모형을 나타낸 것이다.

이러한 모형화의 결과 ISO 2788에서 제시한 유형은 비교적 직관적인 구분이 가능하다고 여겨지나, ISO 2788

에서 간과하고 있는 부분으로서 본 연구에서 신설한 ‘물리적 관계’의 ‘일반 사물’이 특히 연관적 계층관계와의 구분이 어려운 부분이라는 것을 파악할

수 있었다. 이는 이 부분에 속하는 용어쌍의 범주에 대한 해석이 다양하기 때문일 것이다.

이 분류 모형을 통해 전체/부분관계의 소속에 대한 혼란을 일으키는 부분을 확인하여 분리해 내고, 그의 하위분류를 통해 계층관계의 경계를 분명히 할 수 있다는 의의가 있다. 또한 ISO 2788의 구분으로 시작한 하향식의 방법이 실제 다양한 용어관계의 사례로 접근한 상향식 방법과 대범주에 있어 거의 일치한다는 점을 확인함으로써 다양한 계층관계의 유형들을 하나의 틀로 체계화하고 상세화 할 수 있다는 근거가 되었다.

계층관계를 맺는 용어쌍 내부에서의 하위 분류를 위해서는 패킷지시어의 삽입을 통한 의미적 구분이 필요하다. 일례로 신체기관과 신체조직, 조직-구성원의 관계에서 패킷 분류를 도입한다면 각각 영역 특정한 패킷 구분의 방법과 의미상 특정한 패킷 구분을 이용한 하위 분류의 가능성을 제시할 수 있다. 이는 사례별로 특정한 수밖에 없으므로 완벽한 패킷지시어의 목록을 작성하기란 어렵다. 따라서 계층관계를 갖는 용어쌍의 하위 분류를 위해 범용적으로 사용될 수 있는 패킷지시어의 대표적인 유형 및 그 범주화의 사례를 통합적으로 제시하였다.

본 연구에서는 패킷 분석의 시초라 할 수 있는 Ranganathan(1976)의 'PMEST'

를 비롯하여 Perreault(1994)가 제안한 용어관계의 분류 및 한상길(1999)의 연구에서 제안된 패킷지시어를 비교하여 계층관계를 이루는 용어쌍의 하위 분류를 위한 패킷지시어 집합을 마련하였다.

<표 2>는 선행 패킷 구분의 방법에 제시된 개별 사례들을 가능한 한 인정하면서 서로 중복이 없도록 정리하고, 앞선 여러 계층관계의 사례 분석 결과로 보충하여 작성한 것이다. 이 통합된 패킷지시어의 목록은 개별 용어쌍 내부에서의 하위 분류가 필요할 경우에 참조할 수 있는 도구로 사용할 수 있다. 이는 대표적인 사례로서 제시한 것이므로 완벽한 목록은 아니지만, 하위 분류로 이용할 수 있는 대표적인 속성들을 범주화하여 제공함으로써 패킷지시어의 선택을 용이하게 하려는 목적으로 마련된 것이다.

이상에서와 같이 계층관계를 구성하는 다양한 용어관계들의 체계적 구조화가 이루어지고, 그 구조 내에서 필요에 따라 각 용어쌍의 관계에 패킷지시어를 부여하여 내부적인 세분이 이루어진다면, 계층관계의 설정 및 이용이 용이하게 이루어질 수 있을 것이다.

3.3 연관관계

연관관계의 유형 분류에 대해서는 지침 뿐 아니라 참고할 만한 선행연구 또한 거의 없다. 단지 일련의 연관관계

<표 2> 계층관계 분류를 위한 패킷지시어의 비교 및 통합

Ranganathan (1976)	Perreault (1994)	한상길 (1999)	통합
P (개체·핵심주제)	상태 형태 정도 크기 지속기간 지속일사 유사 다름	속성 분야 부분 형태, 크기·수준 조건 조직·위치 상태·현상	부분 상태·현상 속성 정도, 크기, 수준 조직·위치 지속기간 태도 형태
M (재료·물리적 속성) 물질-속성 물질-방법 물질-재료	전체/부분 · 유기체/기관 · 복합물/성분 · 매트릭스/조각 주체/소유 · 물질/성질 · 소유주/소유 · 동반 · 존재 · 부재	재료·설비 도구 구성요소	기관 도구·설비 동반 방법 성분·구성요소 성질·속성 재료 조각
E (에너지)	주동/사동 · 생산적/생산된 · 제한/제한된 · 파괴적/파괴된 상호작용 · 조화 · 다름 · 반대	행위자 목적 대상·수령인 과정 방법·수단 적용 작용 상호 작용 이용 결과·생산물 측정	결과·생산물 과정·수령인 목적 방법 영향 원인 이용(이익,害, 상호) 적용 측정·평가 행위·행위자
S (공간)	수평 · 좌·중·우 축 · 앞·중앙·뒤 수직 · 위·수평·아래	장소	공간적 위치(수평, 축, 수직) 장소
T (시간)	전 동시 후	시간	시간 시간적 위치(동시, 전, 후)

유형 목록을 나열한 연구와 연관관계 분류를 위한 패킷을 제안한 연구가 있을 뿐이다. 따라서 그러한 이론적 연구들이 과연 실제 연관관계의 양상을 어느 정도 포괄할 수 있는지, 어떤 조건에서 어느 정도의 효용성을 가지고 적

용될 수 있는지 등에 대한 이론적, 실증적 기반은 없는 실정이다.

ISO 2788은 “체계적으로 배열된 시소러스(systematic thesaurus)에서 서로 다른 유형의 연관된 개념을 이용자의 편의를 위해 한곳에 두기 위해 노드

레이블(또는 패킷지시어)을 사용할 수 있다”고 명시하고 있다. ANSI Z39.19 역시 ISO 2788의 지침과 동일하다. ANSI Z39.19의 개정안(2005)에서는 이러한 이전 판의 내용에 더하여, 연관관계에 관한 지침의 서두에서 “의미적으로 세분된 연관관계의 개발이 가능하다”는 점을 별도로 밝히고 있다는 점이 특징적이다. 이는 연관관계의 세분에 대한 필요성이 더욱 강조되고 있는 추세를 반영하는 것으로 판단된다.

이처럼 연관관계 세분의 방법으로 패킷지시어의 적용 가능성이 이론적으로 제시되어 왔으나 표준에서도 그 이상의 지침은 제공하지 못하고 있다는 점에 한계가 있다. ANSI Z39.19는 연관관계의 세분에 있어서 “연관관계 세분이 필요할 경우, 이를 위한 관계 지시 부호를 독자적으로 개발할 수 있다”고 밝히고 있으나, 사실상 부호 보다 먼저 필요한 것은 연관관계 세분을 위한 원칙과 방법론이다.

이처럼 연관관계에 대한 논의를 뒷받침해줄 만한 선행 연구사례가 드물기 때문에 본 연구는 이론적으로 제안되어 온 패킷을 이용한 분류의 가능성을 확인하기 위하여 연관관계에 한하여 일련의 실험 평가의 방법을 도입했다.

실험 대상 용어로서 일반적인 수준의 <고등학교 사회·문화> (대한교과서 2002)와 <고등학교 과학> (대한교과서 2004)의 색인 용어를 선정하였고, 보다

용어의 주제 특정성이 강한 용어로서 <국립중앙도서관 주제명표목표>(2002)와 *Thesaurus of ERIC Descriptors*(2001), *Unesco thesaurus : a structured list of descriptors for indexing and retrieving literature in the fields of education, science, social science, culture and communication*(1977)을 대상으로 사회분야 중 ‘Social Problems’에 속하는 용어를 추출하였다. 또한 주제영역별 용어관계의 차이를 파악하기 위하여 사회분야와 대비되는 용어로서 자연과학 분야의 ‘Plant Pathology’에 관한 용어를 선정하여 *AGROVOC Multilingual Agricultural Thesaurus*(2005)를 대상으로 연관관계를 추출하여 그 패킷 분류의 가능성을 평가하였다. 이러한 실험의 결과 패킷 분석을 이용한 연관관계 분류의 가능성과 한계를 밝히고 연관관계의 패킷 분류 방안을 제안하였다.

연관관계의 분류에 관한 일련의 실험 분석의 결과, 연관관계는 영역 특정성과 개별 용어쌍의 특정성을 최대한 고려한 패킷 분석 방법이 필요하다는 결론을 내렸다. 따라서 연관관계의 분류를 위해 활용할 수 있는 패킷지시어로 구성된 분류 모형을 구축했다.

<표 3>은 본 연구에서 다룬 연관관계를 위한 패킷 범주의 사례와 제안을 비교·분석하여 통합적인 패킷 구분의 양상을 밝히고자 작성한 것이다. 이는 Ranganathan의 ‘PMEST’ 범주와

BC2(Bliss Bibliographic Classification version 2)의 13가지 범주, CRG(Classification Research Group)의 7가지 범주를 통합한 항목과, 연관관계의 분석을 위한 전말숙(1998)의 제안, 그리고 UMLS 의미망의 용어관계를 패킷으로 간주한 사례, 한상길(1999)의 패킷지시어 제안을 기반으로 하여 본 연구의 사례 분석 결과를 종합한 항목 등 크게 네 가지 사례를 비교·분석하여 통합적으로 제시한 것이다. 제안된 모형의 범주명으로는 이해가 쉽고 명확한 UMLS 의미망의 연관관계 대범주명을 따랐다. 한편, UMLS 의미망의 용어관계 패킷은 영역 특정성이 강해서 다른 사례들에 비해 많은 차이가 난다. 따라서 영역 특정성이 뚜렷한 사례, 예를 들어 ‘치료, 분열, 심화, 징후, 진단’ 등은 제외하였다.

<표 3>에서 평가, 척도, 측정 등의 개념을 UMLS에서는 개념적인 연관성으로 보고 있으나, 다른 사례에서는 대부분 행동에 관한 기능적 연관성으로 보고 있다. 본 연구의 제안에서는 이를 기능적인 데에 중점을 두어 기능적인 관계로 간주하였다.

본 연구는 어떤 패킷이 많이 쓰이고 적게 쓰이는지의 분포를 밝혀 패킷을 단순화하는데 목적을 두지 않았다. 실재하는 가능한 한 모든 용어관계를 논리적인 모순이 없는 수준에서 포괄할 수 있는 패킷을 마련하는데 목적이 있

다. 따라서 <표 3>에 제시한 패킷 범주는 단순히 같은 항목끼리 통합·정리한 것이므로 분야와 목적별로 일정한 정도의 가공을 거쳐야 적절한 효율을 얻을 수 있다.

이와 같이 비교·통합된 패킷 구분의 결과를 바탕으로 하여 연관관계의 분류를 위한 패킷지시어의 모형을 설계하였다. <표 3>에서 제시한 패킷 범주의 비교 분석 결과를 기준으로 하향식으로 설계하고, 이를 실제 용어관계 분석 결과에 따라 상향식으로 수정 및 보충하였다. 모형의 설계는 UMLS 의미망의 용어관계 정의 형식과 방법을 참조하여 패킷지시어의 정관계와 역관계를 설정한 모형을 작성하였다. <표 4>는 연관관계의 분류를 위한 패킷지시어의 모형이다. <표 5>는 패킷지시어 모형의 정의에 해당하는 것이다. <표 4>와 <표 5>에서 UMLS 의미망의 용어관계를 수용하되, 용어관계가 속하는 관계 범주를 변화시킨 사례(*로 표시)와 용어관계의 의미적 내용을 변화시킨 사례(**로 표시)가 있다. 굵은 글씨로 표시된 부분은 UMLS 의미망에 새롭게 추가하거나 변형시킨 부분이다.

<표 4> 연관관계 분류를 위한 패킷지시어 모형

관계 범주	패킷지시어와 대표적인 용어관계/역관계			
개념적 연관 conceptually _related_to / conceptually _related_to	속성	property_of / has_property		
	구성요소	element_of / has_element		
	기원 · 파생	derivation_of / has_derivation, generation_of / generated_by		
	종류	{kind, type, species, brand}_of / has_{kind, type, species, brand}		
	사례	instance_of / has_instance		
	(개념적)부분	conceptual_part_of / conceptually_related_to		
	정도 · 수준	{degree, level, grade}_of / has_{degree, level, grade}		
	형태 · 외형	{form, appearance}_of / has_{form, appearance}		
	구조	structure_of / has_structure		
	대상 · 현상	phenomenon_of / has_phenomenon		
	상태	{condition, state, situation}_of / has_{condition, state, situation}		
	이슈 · 주제**	issue_in / has_issue		
물리적 연관 physically_ related_to / physically_ related_to	추상적 방법	연결 · 부착**	connected_to / connected_to	
		구성	consists_of / constitutes	
		분기	branch_of / has_branch	
	구체적 구성	합류	tributary_of / has_tributary	
		부분	part_of / has_part	
		재료	consists_of / constitutes	
연관 associat ed_wit h / associat ed_wit h	과정	생산	produces / produced_by	
		방법	method_of / has_method	
		영향 · 작용 · 기능**	affects / affected_by function_of / has_function operate / operated_by	
	기능적 연관 functionally_ related_to / functionally_ related_to	이로운	해로운	harmful
			상호간의	reciprocal
			원인	causes / caused_by
		결과 · 산물	이용	use_of / used_by, utilize_of / utilized_by
			적용	apply / applied_by
			측정*	measurement_of / has_measurement_of
		평가*	수행**	evaluation_of / has_evaluation
			대항 · 억제	performs / performed_by, carries_out / carried_out_by, practices / practiced_by
				opposition_of / has_opposition, constraint_of / has_constraint
분석	analyzes / analyzed_by			
동작주	관리	manages / managed_by		
	보완	supplement_of / has_supplement		
	주체	agent		
공간적 연관 spatially_ related_to / spatially_ related_to	수동체	patient		
	위치	location_of / has_location		
	근접	adjacent_to / adjacent_to		
	경계	surrounds / surrounded_by		
시간적 연관 temporally_ related_to / temporally_ related_to	교차	traverses / traversed_by		
	동시발생	co-occurs_with / co-occurs_with		
	선행	precedes		
	후행	follows		

<표 5> 연관관계의 분류를 위한 패킷지시어와 정의

관계 범주		패킷지시어와 정의			
개념적 연관	속성	특성, 질			
	구성요소	어떤 사물이나 개념을 이루는 성분이나 조건			
	기원·파생	사물의 근원이 되거나 어떤 본체에서 다른 사물이 갈라져서 나와 생김			
	종류	어떤 기준에 따라 나눈 갈래			
	사례	어떤 일이나 개념의 실례			
	(개념적)부분	개념적으로 부분을 의미하는 것			
	정도·수준	사물의 가치, 품질, 등급 등의 특정한 수준(정도) 과정의 상대적 집중도 혹은 질이나 속성의 상대적 분량			
	형태·외형	형태, 모양, 색깔			
	구조	어떤 물건이나 조직체 따위의, 전체를 이루고 있는 부분들의 서로 짜인 관계나 그 체계			
	대상·현상	(본질과의 상관 개념으로서) 시간과 공간 속에 나타나는 대상			
	상태	사물이나 현상이 처해 있는 현재의 모양 또는 형편			
	이슈·주제**	논점, 논쟁점, 화제			
물리적 연관	추상적 방법	연결·부착**	다른 물리적 단위로 이어지거나 결합됨		
		구성	구조적으로 다른 물질이나 사물의 부분이나 전체를 구성		
		분기	어떤 부분으로부터 갈라져 나옴		
	구체적 구성	합류	하나로 모임. 융합. 병합		
		부분	더 큰 전체인 다른 물리적 단위를 구성		
		재료	물건을 만드는 감		
성분	구성성분				
기능적 연관	과정	행동, 기능, 상대			
	방법	어떤 목적을 달성하기 위하여 취하는 수단			
	생산	생산, 창조, 발생, 산출			
	영향·작용·기능**	이론	이로운	유익한 영향	
			해로운	부정적인 영향	
			상호간의	상호적인 영향	
	원인	어떤 상태나 효과를 초래함			
	결과·산물	어떤 활동이나 과정의 결과, 영향, 결론으로서 발생하는 상태, 산물, 상황			
	이용	어떤 물건을 쓰거나 방편으로 함			
	적용	응용하거나 맞추어 씀			
	측정*	차원, 양, 정도, 능력 등을 조사하거나 확인함			
	평가*	가치나 속성, 과정을 평가			
수행**	어떤 활동, 기능, 과정을 실행, 수행, 달성함				
대항·억제	어떤 작용이나 기능이 발생하지 못하도록 맞서거나 억누름				
분석	설정된 정량적, 정성적 방법으로 연구하거나 검사. 기능이나 행동의 영향이나 결과를 분석				
관리	관리하거나 돌보는 것				
보완	모자라는 것을 더하여 채우거나 완전하게 함				
동작주	주체	어떤 기능의 행위자, 직업 종사자			
	수동체	어떤 기능의 대상이 되거나 당하는 사람			
공간적 연관	위치	개체나 과정의 위치, 장소, 지역			
	근접	다른 물리적 단위에 근접하거나 이웃하고 있음			
	경계	경계를 설정하거나 다른 물리적 구조의 한계를 정의함			
교차	다른 물리적 구조나 지역을 교차함				
시간적 연관	동시발생	동시에 혹은 함께 발생함			
	선행	먼저 발생함			
	후행	나중에 발생함			

한관계
유연관계
강연관계
친연관계

4. 활용방안

본 연구에서 제안하는 바와 같이 일정한 체계로 세분화된 용어관계 분류의 활용 방안은 다음과 같다.

첫째, 정보검색의 측면에서 전통적인 시소러스에 세분된 용어관계를 설정함으로써 그 가치와 활용도를 높일 수 있다. 제안된 용어관계의 분류 모형을 검색용 시소러스에 적용하면 검색의 확장 수단으로 세분화된 용어관계를 이용할 수 있다.

직접적인 질의어 확장 및 축소 외에도 큰 규모의 결과 집합을 다루는 참조 도구로도 활용할 수 있다. 한번에 다루기 힘들 정도로 큰 결과 집합을 여러 측면에 따라 분류하여 보여준다면 이용자는 최종 선택하는 결과물을 얻기 이전에 한결 개선된 브라우징을 제공받을 수 있게 된다. 이처럼 검색결과에 대한 분류의 역할을 하는 것이 해당 질의어에 대한 용어관계이며, 이를 이용하여 결과집합의 재가공이 용이하게 된다.

둘째, 용어관계의 세분된 분류 모형은 용어관계의 레지스트리로서의 역할을 할 수 있다. 본 연구가 의도하는 바와 같이 신뢰할만한 상세한 용어관계 정의에 기반한 용어자원이 있으면 이를 활용하여 온톨로지를 생성하는 데 수반되는 의미적 문제들을 상당 부분 해소할 수 있을 것이다. 용어관계가 체계적이고 망라적으로 마련되면 시소러스 뿐

아니라 온톨로지, 분류체계, 주제분류번호, 주제명표목 등 지식조직체계의 여러 유형 및 영역별로 다양하게 활용될 수 있다. 따라서 용어관계를 상세하게 분류한 용어관계 분류 모형을 마련하여 두고, 상황에 맞게 선택하여 사용할 수 있도록 해야 할 것이다. 이는 ISO ISO/DIS 12620:1999 *Computer Applications in Terminology--Data Categories* 및 ISO CD 12620:2003 (Part 2) *Terminology and other language resources--Data Categories --Part 2: Data Categories for Terminological Resources*에서 DCR (Data Category Registry)에 대한 DCS(Data Category Selection)의 개념과 유사한 것이라 할 수 있다.

셋째, 용어관계의 분류 모형은 언어 자원 개발의 기반으로 사용될 수 있다. 자연언어처리 등 다양한 응용분야에서의 활용을 위해서는 명사 이외의 품사간의 용어관계를 설정하는 것도 필요하다. 본 연구에서와 같이 명사를 대상으로 한 용어관계가 먼저 확실하게 마련되면 명사 용어체계를 활용하여 동사와 형용사 등 다른 품사에까지 확대하여 자연언어 처리를 비롯한 의미를 기반으로 한 지식처리에 광범위하게 활용할 수 있을 것이다.

넷째, 용어관계의 분류 결과는 한글 자원 개발의 표준 기반이 될 수 있다. 한글 용어자원의 특색을 반영한 용어관

계가 세밀하고 체계적으로 정립되면 외국의 유수의 언어자원에 필적하는 한글자원을 마련할 수 있는 기반이 된다. 향후 보다 잘 정비된 국제표준으로 제정될 예정인 ISO TC 37의 DCR은 개방형 범주이다. 따라서 사용자가 필요로 하는 항목이 없으면 추가를 제안하도록 규정되어 있다. 그러므로 한국어자원을 기준으로 만든 용어분류체계는 국제표준으로는 미처 포괄하지 못하는 한국어 현실을 반영할 수 있을 것이다.

5. 결론 및 제언

본 연구는 용어관계 분류 모형의 설계에서 각 용어관계의 본질을 보다 상세히 밝히고 각각의 특성에 맞게 개별적인 분류 방안을 제시했다. 이는 용어관계의 세분이 필요함을 이론적으로 주장하여 단순히 용어관계의 세분화된 목록만을 제시하거나 일률적으로 패킷 분석이 적용되어야 함을 주장하는 기존의 연구 사례의 한계를 보완하고자 하는 것이다.

패킷을 이용한 분류에 있어서는 논리적인 완전성 못지않게 실질적으로 활용을 용이하게 하는 방안이 마련되어야 한다. 본 연구에서는 다음과 같은 세 가지 방안을 제안하였다.

첫째, 패킷을 이용한 용어관계의 분류를 위하여 패킷과 실제 용어쌍을 잇

는 중간 역할로 실제 용어관계 유형을 두는 방안이다. 둘째, 다양한 이용 요구를 가정하여 미리 여러 가지 패킷 모형을 마련하여 두는 것이다. 셋째, 패킷지시어 집합에서의 선택에 있어서 영역과 목적, 수준 등에 따라 패킷지시어를 선택할 수 있는 틀을 마련하여 그 활용을 도울 수 있다.

한편, 앞으로 용어관계의 연구를 위해서는 기존의 용어관계 연구에서 보다 학제적인 협력이 더욱 중요하게 된다. 용어관계의 연구에 있어서 특히 학제적 협력이 강조되는 연구들은 다음과 같다. 첫째, 용어관계 분석에 관련된 모든 요소에 관한 정의가 필요하다. 둘째, 보다 명확한 용어관계의 분류를 위해서는 용어의 분류를 통한 개념적 체계도가 마련되어야 한다. 셋째, 용어관계의 세분이 이루어지면 시스템적인 처리와 인터페이스 설계의 문제가 더욱 중요하게 된다.

이러한 연구들이 특정한 학문 영역에서 고유하게 담당해야 할 몫이라고 경계를 짓는 것은 사실상 어렵다. 정의 작성과 개념 분석의 작업은 언어학적인 측면의 성격이 강하고, 시스템 설계나 인터페이스의 설계는 컴퓨터공학적인 측면이 강하지만, 이들이 개별적인 학문 영역에서 배타적으로 연구될 수 있는 것은 아니다. 오히려 용어관계 연구는 여러 관련 분야에서 협력적으로 추진해 나가야 할 과제이다.

참 고 문 헌

- 국립중앙도서관. <국립중앙도서관 주제명 표목표>. <<http://www.nl.go.kr>>.
- 전말숙. 1998. 시소러스의 연관관계 유형에 관한 연구. 『정보관리연구』, 29(1): 20-39.
- 한상길. 1999. 『시소러스 용어관계의 확장에 관한 연구』. 박사학위논문, 중앙대학교 대학원, 문헌정보학과.
- American Library Association, Subject Analysis Committee, Subcommittee on Subject Relationships/Reference Structures. 1999. *Final Report to the ALCTS/CCS Subject Analysis Committee*. [cited 2004. 4. 6]. <<http://archive.ala.org/alcts/organization/ccs/sac/msrscu2.pdf>>.
- Cruse, D. A. 1986. *Lexical Semantics*. Cambridge and New York: Cambridge University Press.
- Edmonds, Philip and Graeme Hirst. 2002, "Near-synonymy and lexical choice." *Computational Linguistics*, 28(2): 105-144.
- ERIC. 2001. *Thesaurus of ERIC Descriptors*. 14th ed. [Phoenix, Ariz.] : Oryx Press.
- FAO. *AGROVOC Multilingual Agricultural Thesaurus*. <<http://www.fao.org/agrovoc/>>.
- ISO 2788-1986(E). *International Organization for Standardization Documentation: Guidelines for the Establishment and Development of Monolingual Thesauri*. 2nd ed.
- ISO/DIS 12620:1999 *Computer Applications in Terminology --Data Categories*.
- ISO CD 12620:2003 (Part 2) *Terminology and other language resources -- Data Categories -- Part 2: Data Categories for Terminological Resources*.
- Lu, X. 1990. "Document retrieval: a structural approach." *Information Processing & Management*, 26(2): 209-218.
- Mandersloot, Wim. G. B., Eleanor M. B. Douglas, and Neville Spicer. 1970. "Thesaurus control: the selection, grouping, and cross referencing of terms for inclusion in coordinate index word list." *Journal of the American Society for Information Science*, 21(1):

- 49-57. *Information Science*, 36(1): 15-27.
- Miller, G., R. Beckwith, C. Felbaum, D Gross and K. Miller. 1990. "Introduction to WordNet: an on-line lexical database." *International Journal of Lexicography*, 3(4): 302-312.
- National Information Standards Organization. 1993. ANSI/NISO Z39.19-1993. *Guidelines for the Construction, Format, and Management of Monolingual Thesauri*.
- . 2005. ANSI/NISO Z39.19-200X. *Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies*.
- Perreault, Jean. 1994. "Categories and relators: a new schema." *Knowledge Organization*, 21(4) : 189-198.
- Ranganathan, Shiyall Ramamrita. 1976. *Colon classification*. 6th ed. Bombay : Asia Publishing House.
- Storey, Veda C. 1993. "Understanding semantic relationships." *VLDB Journal*, 2: 455-488.
- Wang, Yih-Chen, James Vandendorpe, Martha Evens. 1985. "Relational thesauri in information retrieval." *Journal of the American Society for*

부록. 분석 대상 및 용어관계의 유형별 개수

사례 유형	연구사례	동등관계	계층관계	연관관계
언어학	임지룡(1992)	11	5	*
시소러스 표준(지침)	ISO 2788(1986)	(11)	(6)	(12)
	ANSI Z39.19(2005)	11	6	16
	Aitchison 외(2000)	(11)	(6)	(14)
용어관리표준	ISO 12620:1999 (Part 1)	10	2	10
	ISO 12620:2003 (Part 2)	15	9	8
	GB/T 16786-1997(1997)	(10)	(2)	(10)
이론적연구	Perreault(1994)	0	5	78
	Willetts(1975)	1	1	7
	Wang 외(1985)	5	8	0
	Lancaster(1986)	0	0	10
	Maniez(1988)	0	0	3
	Lu(1990)	3	16	8
	Schmitz-Esser(1999)	0	4	9
	Foskett(1996)	3	16	8
	ALA 분류(1999)	56	26	121
	시소러스	한국문화예술진흥원(1994)	*	*
IRTET(1968)		0	0	15
Thesaurofacet(1970)		0	0	9
TCT(1985)		1	2	4
기타(1999)		4	1	15
의미망과 온톨로지	WordNet(1998)	2	2	0
	EuroWordNet(1998)	5	18	23
	HowNet(2005)	3	2	11
	EDR(1990)	2	2	*
	UMLS 의미망(2005)	*	1	53
	기타 온톨로지	*	*	*
데이터베이스	Storey(1993)	2	5	20
	Neelameghan(2001)	4	0	28
합계	29종	138	132	454

(‘) 표시는 같은 범주의 동일 사례와 중복되지 않도록 계수에서 제외한 것이다.
 ‘*’ 표시는 수가 너무 많거나 무한하여 셀 수 없는 항목임을 표시한 것이다.)