

디스크립터의 의미기술로서 정의를 통한 애매성 제거

The Ambiguity Exclusion Using Definitions as Meaning Description of Descriptors

김 지 훈*
Ji-Hun Kim

차 례

- | | |
|----------------------|----------------------|
| 1. 서 론 | 4. 정의의 작성과 시소러스에의 통합 |
| 2. 시소러스에서 용어와 개념의 관계 | 5. 결 론 |
| 3. 시소러스에서 용어의 의미설명 | • 참고문헌 |

초 록

지금까지 대부분의 시소러스에서 용어의 의미는 주로 관계나 범위주기를 통해 간접적으로 제시되어 왔다. 그러나 시소러스가 양적으로 팽창해짐에 따라, 이들을 통해 용어의 의미를 명확히 파악하는 것이 점차 어렵게 되었다. 이에 일부 시소러스는 사전의 정의를 함께 수록하기도 하였지만, 그 내용이나 형식이 만족스럽지 않은 것으로 인식되어 왔다. 이 연구는 용어의 의미를 제공하는 정의의 필요성을 토대로 표준화된 정의를 시소러스에 통합해 봄으로써, 시소러스가 더욱 발전할 수 있는 가능성을 제시하였다.

키 워 드

시소러스 구축, 용어의미, 표준정의, 정의 모델, 정의규칙

*계명문화대학 사회복지상담과 부교수
(Associate Professor, Dept. of Social Welfare Counseling, Keimyung College, jhkim@km-c.ac.kr)
• 논문접수일자 : 2005년 8월 25일
• 게재확정일자 : 2005년 9월 8일

ABSTRACT

Until now, the meaning of terms in most of thesauri has been indirectly presented through relationships and scope notes. However, as thesaurus has been expanded quantitatively, it has become difficult to understand meaning of terms clearly using those relationships and scope notes. Though some thesauri take the form of 'thesaurus dictionary' in order to present the meaning of descriptors, the contents and forms of provided definitions are not satisfying. As standardized definitions of terms integrate into thesaurus on the basis of necessity of definitions providing meaning of terms, this study has shown a possibility for further development of thesaurus.

KEYWORDS

Thesaurus Constructions, Term Meaning, Standardized Definitions, Defining Model, Defining rules

1. 서 론

정보검색은 문헌을 대표하는 용어와 정보요구를 표현한 용어 간의 일치여부에 기반을 두고 있다. 그러나 하나의 개념은 관점이나 표현에 따라 다양한 용어로 나타나는 경향이 있기 때문에(Bates 1986), 정보검색과정에서 색인자와 탐색자가 동일한 개념에 대해 서로 다른 용어를 사용하여 서로 일치되지 않는 문제가 발생할 가능성이 많다.

이러한 문제를 해결하기 위해, 정보검색에서는 용어의 존재유무에 근거하는 용어기반검색보다는 용어의 개념과 그 관계에 근거한 개념기반검색을 지향하고 있다. 즉, 색인과정에서 문헌을 대표하기 위해 선정한 용어와 탐색과정에서 정보요구를 나타내기 위해 선정한 용어가 서로 일치하지 않더라도, 개념기반검색은

동일한 개념이라면 서로 일치되어 탐색되도록 한다. 정보검색에서 이러한 개념기반검색을 가능하게 하는 주요한 보조도구가 시소러스다.

시소러스는 색인과정에서 문헌의 내용을 완전하고 정확하게 표현하기 위해 개념을 통해 적절한 용어를 선정하고 표준화하도록 하여 색인어를 일관되게 작성할 뿐만 아니라 탐색과정에서 질의를 작성할 때 용어 및 개념 간의 관계를 제시하여 탐색어로서 적절한 용어를 선정하는 데 도움을 주어 전반적으로 정보검색의 품질을 향상시키는 역할을 한다.

시소러스가 이러한 역할을 제대로 수행하기 위해서는 기본적으로 특정주제 분야에서 사용되는 용어의 의미뿐만 아니라 그 주제 분야의 구조를 이해하는 데 도움이 되도록 용어 간의 관계를 명확하게 제시하여야 한다. 이에 따라 오늘날 대부분의 시소러스는 ISO 2788, ISO

5067, ANSI/NISO Z39.19와 같은 표준이나 시소러스 구축에 관한 실제적인 지침에서 제시하고 있는 관계구조에 근거하여 용어의 의미와 용어 간의 관계를 제공하고 있다.

그러나 학문의 발전과 함께 용어의 수가 증가하여 시소러스가 양적으로 팽창해짐에 따라, 기존의 단순한 관계구조로는 용어의 의미를 명확하게 파악하지 못할뿐만 아니라 용어 간의 정밀하고 다양한 관계를 제시하기 어려워지게 되었다. 이러한 문제를 해결하고자 시소러스 연구자들은 시소러스에서 용어의 의미를 제시하는 문제와 관계를 확장하는 문제에 많은 관심을 가지게 되었는데, 지금까지는 주로 관계구조를 확장하는 방안들만 제시되어 왔다.

한편, 용어의 의미를 제시하는 면에서는 지금까지 대부분의 시소러스가 용어가 지니고 있는 개념의 의미를 직접 제시하는 대신 관계구조와 범위주기를 통해 간접적으로만 제시하여 왔다. 그러나 기본적으로 관계구조나 범위주기는 용어의 개념이 지닌 본질적인 속성을 이해하거나 범위를 한정하기 어려운 면이 있어, 관계구조를 더욱 자세하게 설정하기 보다는 용어에 대한 명확한 의미를 제시할 필요성이 지속적으로 제기되어 왔다.

이러한 필요성에 따라 일부 시소러스는 디스크립터로 선정된 용어의 의미를 제공하기 위해 기존의 사전에서 제시된 정의를 함께 수록하고 있지만, 제시되어 있는 정의의 내용이나 형식이 만족스럽지 않은 것으로 인식되어왔다. 이에 시소러스에서 디스크립터의 의미를 나타

내는 정의를 내용이나 형식면에서 표준화된 정의를 체계적으로 포함하려는 생각은 시소러스를 개발한 이래 여러 번 표면화되어 왔다. 그러나 그러한 시도가 이론적인 면에서 확고한 원칙과 절차가 없었을 뿐만 아니라 정의 그 자체가 시소러스에서 어느 정도 유용한 지에 대한 명확한 해답을 도출하지 못해 실제로 거의 적용되지는 않았다.

이 연구는 용어의 정의가 시소러스에서 유용함을 확인하기 위해, 전문용어(terminology) 분야의 이론에서 이용되고 있는 정의 모델과 정의규칙에 의거하여 내용과 형식면에서 일관성 있는 용어의 표준정의를 작성하고, 이를 시소러스에 포함하여 더욱 확장된 시소러스를 구축하는 것을 목적으로 한다. 이를 위해 먼저 시소러스에서 용어와 개념과의 관계를 명확히 검토하고, 시소러스에서 용어의 의미를 설명하는 방법을 토대로 정의의 필요성과 중요성을 고찰하였다. 이러한 이론을 바탕으로 실제로 용어의 표준정의를 작성하여 기존의 시소러스에 통합한 확장된 시소러스를 프로토타입으로 구축하여 정보검색에서 유용한 도구가 될 수 있는 가능성을 테스트하였다.

2. 시소러스에서 용어와 개념의 관계

2.1 용어와 개념의 중요성

정보검색과정에서 시소러스를 이용하는 가장 큰 이유는 용어와 개념과 관련된 문제를 해

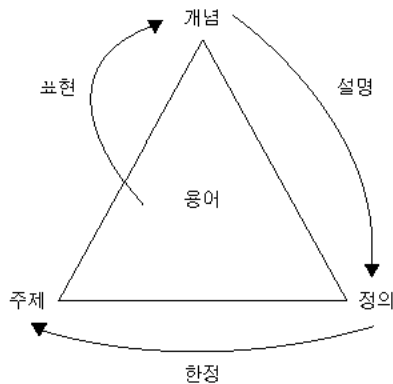
결하기 위함이다. 예컨대, 'Illness'를 검색할 때 'Illness or Disease or Sickness or Ailment'를 고려해야 하지만, 시소러스는 용어와 개념이 일대일관계가 되도록 용어를 제어한다. 그러나 용어는 문맥에 따라 여러 가지 개념으로 나타나 하나의 개념이 여러 용어로 표현되기도 하고, 반대로 하나의 용어가 여러 개념을 나타내기 때문에 본질적이고 분명한 의미를 파악하기 어렵다.

기본적으로 용어는 특정 주제에서 사용되는 하나의 개념을 유일하게 나타내는 주요한 단위로서, 단일어 또는 복합어로 된 어휘를 일컫는다. 개념은 동일한 특성을 가진 대상을 추상화하여 일반화한 관념으로, 그 자체로는 나타낼 수 없기 때문에 기호로 나타낸 것이 용어다. 따라서 개념은 용어의 의미와 동일하며, 용어의 의미가 무엇인지를 나타내기 위해서 다시 기호, 즉 용어를 이용하여 2차적으로 표현한 것이 정의다. 이것을 도식화하면 <그림 1>과 같이

나타낼 수 있다(Bonjour and Falquet 1994).

이러한 관점으로 보면, 개념은 용어와의 관계를 분명히 하여 하나의 용어로만 표현되어야 하며 그 의미가 명확히 정의되어야 한다(Sager 1990). 즉, 하나의 단어가 여러 의미를 가지더라도, 그것이 용어로 사용될 때는 여러 의미 중 단 하나의 의미를 가리켜야 한다는 것이다(Riggs 1981).

이러한 시각으로 볼 때, 구조적으로 유사한 것으로 생각되는 시소러스와 WordNet은 다른 관점에서 만들어 졌음을 볼 수 있다. WordNet은 사전을 개념적으로 찾기 위해 단어의 의미를 중심으로 관계를 표현한 어휘 데이터베이스로서 단어와 의미 간의 관계를 다루는데, 그 관계가 다대다(many to many)의 관계를 이루고 있다(Miller et al, 1990). 하지만 시소러스는 "상위 및 하위개념 사이의 관계를 명백하게 하기 위하여 공식적으로 조직·제어된 색인어의 어휘집 "이고" 색인어는 명사나 명사구의 형



<그림 1> 용어, 개념 및 정의의 관계

태로 나타낸 개념의 표현"이라는 정의에서 볼 수 있듯이(International Organization for Standardization 1986), 시소러스는 개념을 나타내는 용어 데이터베이스로 용어와 개념 간의 관계를 다루고 있어 일대일(one to one)의 관계를 유지해야 한다. 또한 시소러스는 일상적인 언어를 대상으로 하는 것이 아니라 전문가가 공유한 전문지식을 커뮤니케이션하기 위한 용어인 전문용어를 대상으로 하기 때문에 (Aitchison 1994), 시소러스에서 개념을 나타낼 때는 '단어'보다 '용어'로 언급하는 것이 바람직하며, 개념이 하나의 용어로만 표현될 수 있도록 그 관계를 분명히 해야 한다.

이처럼 시소러스에서는 개념을 시소러스의 주요 요소로서 간주하지만, 개념을 나타내는 용어가 자연언어의 본질적인 문제나 일상언어에서 채택되는 문제 등으로 인해 용어와 개념을 분명하게 구별하지 않고, 다만 개념에 대해 디스크립터로서 하나의 용어를 선택하도록 하고 있다. 이와 관련하여 Soergel(1985)은 정보 검색은 개념에 적합한 문헌을 검색하는 것이고, 개념은 적절한 용어를 이용해서만 나타낼 수 있기 때문에, 좋은 시소러스를 개발하기 위해 개념을 한정하는 용어를 만드는데 단초를 제공하는 개념과 용어 간의 관계에 관심을 가져야 함과 동시에 다음의 지적인 작업이 필요하다고 제시하였다.

① 개념 형성(concept formation): 중요한 개념을 서술하고, 용어의 동의어 또는 유사동어들을 확인하고, 두 용어의 의미가 충분

히 밀접한가를 결정한다.

② 개념 정의(definition of concepts): 개념의 내연과 외연 및 관련개념과 구별하는 범위를 결정한다.

③ 개념 배열(arrangement of concepts): 각 개념에 대하여 개념의 계층관계 및 연관관계를 만든다.

이상의 내용으로 보아, 시소러스에서 개념과 용어에 대해 이상적인 것은 하나의 개념이 단지 하나의 용어로만 표현되어야 한다는 것이다. 그러나 시소러스는 개념을 표현하기 위해 분류표나 통제언어를 이용하는 것이 아니라 자연언어를 이용하기 때문에, 시소러스를 이용하면 자연언어 본래의 결점인 용어의 의미가 모호해져 하나의 개념이 하나의 용어로 표현되지 않는 문제를 항상 잠재적으로 가지고 있을 수 있다.

2.2 개념의 특성과 용어의 정의

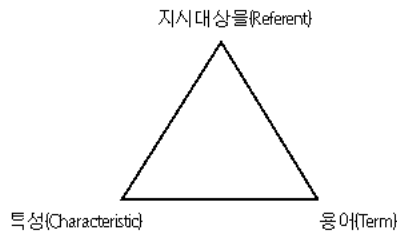
개념에 대해 ISO 1087에서는 "대상물의 공통적인 특성을 기초로 하여 추상을 통해 구성된 사고단위"로 정의하고 있으며(International Organization for Standardization 1990), 철학에서는 하나의 사물을 나타내는 여러 관념 속에서 공통되는 요소를 가려내어 종합한 하나의 관념으로 정의하고 있다. 또한 전문용어 분야에서는 개념을 나타내는 기호가 용어이므로, 용어가 개념을 표현하기 위해서는 용어보다 개념이 먼저 존재해야 하며, 개념은 인간의 사고 속에

많은 중요한 역할을 수행하는 것으로 간주하고 다음과 같은 역할을 한다고 보고 있다(Sager 1990).

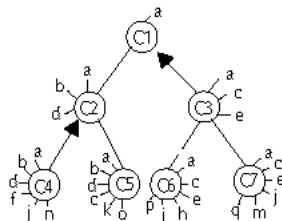
- 사고의 단위, 즉 대상의 정신적 표현 및 인식 수단
- 지식의 단위, 즉 특정 주제 및 학문의 지식구조의 표현 수단
- 커뮤니케이션의 단위, 즉 언어기호를 이용한 지식의 커뮤니케이션 수단

일반적으로, 개념의 구성요소는 <그림 2>와 같이 개념의 지시대상물, 지시대상물에 대한 특성, 그리고 지시대상물의 명칭인 용어 등 세 가지로 구성된 개념삼각형으로 나타내고 있는데, 이 모두가 존재해야만 개념이 될 수 있다(Dahlberg 1981).

한편, 개념을 추상화한다는 것은 각 지시대상물에서 서로 다른 특성을 제거하고 같은 특성만 종합하여 유 종(genus species)의 관계를 만드는 과정이다. 개념은 일반적으로 이러한 추상과정을 거쳐 개념 간의 계층이 피라미드 형식으로 만들어지며, 인접한 상위개념(유)에 따라 정렬되고 관계유형에 따라 하위개념(종)을 가지게 된다. 이러한 유 종관계에 따라, 상위개념의 특성이 하위개념에 상속되는 것을 전제로 하여 하나의 개념을 완전히 설명하려면 상위개념의 특성을 모두 언급해야 한다. 또한 동일한 수준에 있는 개념 간의 차이는 서로 구별되는 특성으로 파악할 수 있는데, 전체적으로 <그림 3>과 같은 감자모양으로 나타낼 수 있다.



<그림 2> 개념의 구성요소



<그림 3> 특성에 의한 개념의 계층구별

예컨대, <그림 3>에서 동일한 수준에 있는 개념 C4와 C5는 상위개념인 C2의 특성인 a, b, d를 모두 공유하며, C4의 특성인 f, j, n과 C5의 특성인 g, k, o에 의해 서로 구별된다는 것이다. 이에 따라 하나의 개념은 인접한 상위 개념의 특성과 동일수준의 다른 개념과 그 개념만이 가지는 구별되는 특성으로 정의될 수 있다. 또한 개념 간의 관계는 특정 주제에서 개념이 지닌 문맥적 의미를 파악하여 나타낸 논리적인 관계로서, 전체적인 개념망을 파악할 수 있게 한다.

전문용어 분야에서는 이러한 개념의 특성을 각 개념을 구별하기 위한 개념의 최소 요소로 간주하여 개념을 기술하고 분류하며 정의하는데 사용하며, 내포(intension)와 외연(extension)으로 구분하고 있다(Wright and Budin 1997). 내포는 개념 간에 구별을 위한 본질적인 속성이며, 외연은 유개념에 포함되는 모든 개념집합으로서 개념이 속하는 동일한 추상수준의 모든 개념 또는 모든 개별대상의 집합이다(International Organization for Standardization 1990).

물론, 개념에 대해 어떤 특성이 적합한지 판단하기 어렵다는 문제가 있기는 하지만, 개념의 특성 중에서 가장 적합한 특성만을 포함한 것이 정의가 되기 때문에, 개념의 특성은 중요하게 간주된다. 일반적으로 개념의 특성에 근거하여 개념을 정의하는 방법으로는 내포적 정의와 외연적 정의가 있다. 전자는 개념의 내포, 즉 구별특성(중차)들을 기술하는 것이고,

후자는 그 개념과 동일한 추상수준의 개념을 열거하는 것이다. 일반적으로 언어사전의 정의는 단어의 의미를 기술하는 반면, 전문용어의 정의는 다른 개념과 구별되는 모든 특성에 근거하여 그 의미를 기술한다. 따라서, 내포적 정의가 전문용어를 정의하는데 가장 이상적인 정의로 인식되고 있다(Sager 1990).

내포적 정의는 분석적 정의로도 알려져 있는데, 그 구조는 아리스토텔레스의 '피정의항 = 정의항(인접류+구별특성)'이라는 정의원칙을 이용한다(Dahlberg 1981). 즉 용어를 분석적으로 정의하는 것은 그 개념이 속한 유와 그 유에 속하는 다른 개념과 구별되는 특성을 나타내는 것이다(Sager and Ndi Kimbi 1995). 이상의 내용으로 볼 때, 용어는 특정 주제에서 정의를 가진 개념을 가리키는 것으로 볼 수 있다.

3. 시소러스에서 용어의 의미설명

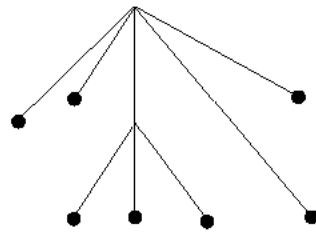
3.1 관계구조를 통한 의미설명

일반적으로 시소러스에서 용어의 의미는 용어 간의 관계 그 자체로 정의된다고 보기 때문에 (International Organization for Standardization 1986), 관계를 통한 용어 간의 의미연결이 특정 주제를 표현하는데 가장 적절한 용어로 충분히 안내할 것이라고 생각하여 왔다(Lancaster 1991). 시소러스에서 관계는 크게 개념 간의 관계 및 용어와 개념 간의 관계 등 두 가지 종류로 구분하고 있다

(Soergel 1985).

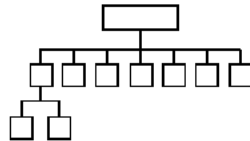
개념 간의 관계는 여러 개념 중에서 하나 이상의 공통적인 특성으로 만들어 지는데, 이는 다시 형식적(formal) 관계와 실체적(material) 관계로 구분된다(Dahlberg 1989). 형식적 관

계는 동일관계, 포함관계, 교차관계, 배제관계 등 네 가지로 구분된다. 실체적 관계는 계층관계, 분할관계, 대립/상보관계, 기능관계 등 네 가지로 구분되는데, 특히 이 관계를 통해 용어의 의미를 파악할 수 있다.



gardening tools
shovel
spade
hoe
 scuffle hoe
 garden hoe
 weeding hoe
rake
 lawn rake
hook

(가) 계층관계

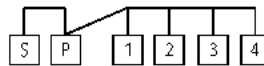


Eye
pupil
 lens
 suspensory ligament
iris
 cornea
 sclera
 vitreous body
retina
optic nerve

(b) 분할관계

- ① 정반대관계 numeric ↔ non unmeccir : presence ↔ absence
- ② 역반대관계 north pole ↔ south pole : large ↔ small
- ③ 긍정 중립 부정관계 favourable ↔ neutral ↔ unfavourable
- ④ 유사/상동관계 arm(of a human being) ↔ wing(of a bird)

(c) 대립/상보관계



Subject; thesaurus
Predicate constructions of
Complement 1: for natural language processing
Complement 2: in the medical field
Complement 3: in Frankfurt
Complement 4: in the year 1976

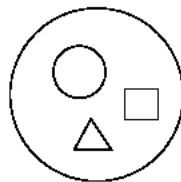
“construction of a thesaurus for natural language processing in the medical field in Frankfurt, 1976”

(d) 기능관계

<그림 4> 개념 간의 실체적 관계

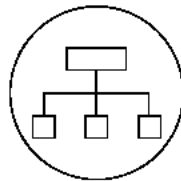
실체적 관계 중에서 계층관계는 형식적 관계 중 포함관계로 될 때 나타나는 것으로, 개념 간에 상하관계가 만들어진다. 즉, 하위개념은 상위개념의 특성 외에 최소한 한 가지 특성을 더 가지며, <그림 4>의 (a)와 같이 대개 나무구조로 묘사된다. 분할관계는 개념 전체가 그 개념 자체 및 부분으로 분할될 때 만들어지며, <그림 4>의 (b)에서 개념 "eye"와 같이 나타난

다. 대립/상보관계는 긍정적 및 부정적 특성이 나 어떤 다른 상반되거나 보완적인 특성을 포함하는 개념에서 나타나며, 그 종류는 <그림 4>의 (c)와 같다. 기능관계는 <그림 4>의 (d)의 주어와 술어처럼 두 개념이 구문적으로 연관되어 여러 수준에서 나타난다. 예컨대, 이 그림에서 첫 번째 수준은 주어와 술어만 다루고, 두 번째 수준은 목적, 조건, 포함된 사람, 공간, 시



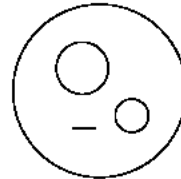
정의 Mane: the long and heavy hair growing around the neck of some animals
 분석 Broad concept(genus): hair
 Differing characteristics 1, 2: long, heavy
 Additional characteristics: growing around the neck...

(a) 계층관계



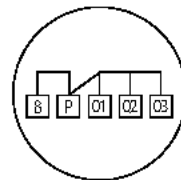
정의 Hammer: a handtool used for pounding, consisting of a solid head set crosswise on a handle
 분석 전체개념: handtool used for pounding
 부분개념: solid head, handle

(b) 분할관계



정의 land: the solid part of the surface of the earth in contrast to water of oceans and seas
 분석 반대는 대조개념으로 표시한다.

(c) 대립/상보관계



정의 Transgression: the covering of land by water on a great scale causing unconformable beds of material
 분석 Subject: land
 Predicate: covering
 Complement 1 (genesis): by water
 Complement 2 (quantity): on a great scale
 Complement 3 (causality): causing unconformable beds of material

(d) 기능관계

<그림 5> 개념 간의 실체적 관계를 통한 정의분석

간 등의 보충요소로서 슬어를 보충한다.

이러한 개념 간의 실제적 관계를 통해 용어의 의미를 나타내는 정의를 다음과 같이 설명할 수 있다. 계층관계를 이용한 정의는 시소러스 구조에서 유 종관계로 중요하게 다루어지는데, 개념을 나타내는 상위어와 하위어가 논리적인 배열로 제시된 것으로, <그림 5>의 (a)와 같이 분석적 정의형태로 만들 수 있다. 분할관계를 이용한 정의는 개념을 대상의 구성요소로서 정의하는 것으로, 정의에서 언급된 첫 번째의 전체개념은 상위개념과 동일하므로 계층관계의 일부분에 해당되며, <그림 5>의 (b)와 같이 만들 수 있다. 대립/상보관계를 이용한 정의는 개념이 부정 또는 반대로 정의되는 것으로, <그림 5>의 (c)와 같이 만들 수 있다. 기능관계를 이용한 정의는 전체개념의 요소 및 특성을 포함할 뿐만 아니라 보충요소의 구문관계를 이용하여 <그림 5>의 (d)와 같이 만들 수 있다.

이처럼 개념 간의 실제적 관계는 개념이 반대로 정의되는 대립/상보관계를 제외하고는 개념에 대해 많은 의미를 포함하고 있음을 볼 수 있다. 이런 이유로 인해 일반적으로 대부분의 일반사전에서는 계층관계와 분할관계를 이용하여 정의하는 반면, 전문사전은 기능관계를 이용하여 정의하고 있다.

한편, 용어와 개념 간의 관계로는 동가관계와 연관관계가 있다. 동가관계는 동의어나 유사동의어가 해당되는데, 시소러스에서 동일한 개념을 나타낸다고 인정되는 경우에는 우

선어와 비우선어간의 관계로 파악하여 USE와 UFR로 관계를 나타내고, 동음이의어는 한 정어를 부가하여 그 의미를 구별하여 일대일 관계를 만든다. 이처럼 시소러스에서 동의어를 이용하여 용어의 의미를 설명하는 것이 일반적이다. 그러나, 개념이 용어의 동의어로 정의될 수 없다는 Sager(1990)의 지적처럼, 동가관계는 기본적으로 개념과 용어 간에 일대일 관계가 아닐 뿐만 아니라 분석적 정의형태로 개념을 정의할 경우에도 이 관계는 적용될 수 없다.

연관관계는 개념적으로 밀접하게 관련되어 있으나 동가관계에 포함되지 않는 용어 간의 관계로 정의하기 대단히 어렵기 때문에, 개념 간에 명시적으로 관련지워야 할 정도로 의미적으로나 심리적으로 연관이 있는 경우에만 만들도록 권고되고 있다(International Organization for Standardization 1986; National Information Standards Organization 1994). 그러나 연관관계는 개념 간의 연관성을 확실하게 제한하는 규칙도 없고, 이해하기 매우 어려우며, 또한 연관관계를 만드는 것이 전적으로 시소러스 편찬자에 달려 있기 때문에, 시소러스 구조에서 주의하여 다루어야 하지만, 그렇지 않은 경우가 많다. 또한 연관관계는 용어사용의 편의성을 위해 만들어 용어의 의미를 모호하게 하는 경우가 많은데, 특히 용어 간의 관계가 다소 정확한 과학기술 분야보다 사회과학 분야에서 심각하게 나타나므로 가치 없는 연결을 하여 시소러스가 확대

되는 위험을 줄이기 위한 제어가 필요하다고 지적되었다(Aitchison and Gilchrist 1987). 그럼에도 불구하고 연관관계는 다른 관계보다 가장 자유롭게 만들어지고 있으며, 많은 시소러스에서 다양한 연관관계를 포함하려는 경향이 있다. 대체로 연관관계는 다음의 여러 가지 기능을 취하는데, 이들 기능 중 두 가지는 의미 설명과 관계가 있다.

- 이용자의 마음속에서 개념적 관계집합을 만들어 디스크립터를 정의한다.
- 이용자에게 유용할 것으로 보이는 다른 디스크립터를 알린다.
- 디스크립터를 문맥에 끼워 넣어 범위를 분명히 하고, 특성을 정의한다.
- 계층을 보충하고 보완한다.
- 밀접한 개념관계를 가진 용어를 함께 연결한다(Willetts 1975).

또한 연관관계가 만약 용어 중 하나가 다른 용어의 정의나 설명에서 필요한 요소로 보여지거나(Aitchison and Gilchrist 1987), 일반적인 요소가 각 용어정의에 있다면 용어가 연관되었다고(Soergel 1974) 고려하는 것으로 볼 때, 연관관계는 용어의 의미설명을 담당한다고 볼 수 있다. 이러한 생각은 시소러스가 용어로 표현된 개념 간에 동등관계와 계층관계 뿐만 아니라 여러 가지 연관관계를 표시하므로, 이러한 관계를 통해 용어의 의미를 한정하는 것은 이론적으로 가능하다고 볼 수 있다.

그러나 시소러스의 관계구조가 개념 간의 관계인지, 아니면 용어 간의 관계인지가 분명

하지 않기 때문에 먼저 개념 간의 관계가 전제되어야 하고(McNaught 1983), 용어의 의미를 관련어들로 파악하는 것이 다소 무리가 있으며(Moores 1985), 특정 주제에서 개념 간의 상호관계에 대한 포괄적인 이론이 아직 없다는(Nkweny Azeh 1994) 등의 지적으로 볼 때, 관계구조는 의미를 완전히 설명하는데 도움이 되지 않고, 다만 용어정의의 보충정보로 볼 수 있다.

한편 전문용어 분야에서는 시소러스의 관계구조를 통한 의미설명은 개념 A는 개념 B를 참조하여 정의하고, 반대로 개념 B는 개념 A를 참조하여 정의하는 것과 같은 순환정의에 해당되어 논리적으로 모순이 있기 때문에(Robert 1985) 이러한 방식을 배제하고 있다. Sager(1990)는 개념을 단지 관계로 의미를 나타내는 것은 극도로 단순하다고 보았다. 그는 이러한 관계구조가 의미를 설명하는 역할을 하여 이용자가 디스크립터의 의미를 결정하는데 약간의 도움을 줄 수 있지만, 정의만큼 정확하게 의미를 제시할 수 없으며, 다만 관계구조가 개념의 의미를 제시하기보다 용어선택을 지원하는 정도로 보아야 한다고 지적하였다. 즉 용어는 그 자체가 정의되어야 하며, 의미를 설명하는데 관계를 이용하는 것은 적절하지 않다는 의미다. 따라서 시소러스에서 용어의 의미를 분명히 전달하기 위해 이러한 기본적인 항목 외에 다른 장치의 필요성이 지속적으로 제기되어 왔다.

3.2 범위주기를 통한 의미설명

일반적으로 범위주기는 디스크립터의 의미를 명확하게 설명하거나 이용법을 제시할 필요가 있을 때 제공하여 왔다. 범위주기에 대해 시소러스 표준에서는 “디스크립터의 활용을 제한 또한 확장하거나, 자연언어에서 의미가 중복되는 디스크립터를 구별하거나, 색인자나 탐색자에게 용어를 사용하는데 조언하기 위해 사용하며, 나아가 디스크립터의 의미를 언급해야 한다”고 기술하고 있다(National Information Standards Organization 1994). 또한 시소러스 개발지침에서는 범위주기를 “색인 시 소러스 용어를 이용하는 방법에 대한 정보를 제공하는 색인지향주기”로 보고, 다음사항을 나타내고 있다고 기술하고 있다(Aitchison, Gilchrist and Bawden 2000).

- 우선어의 의미제한
- 개념의 주제범위
- 색인 시 색인어가 이용되는 방법 지시
- 용어표시
- 용어내역
- 범위주기 내의 다른 우선어에 참조될 때, 상호범위주기

이와 함께 범위주기는 그 역할이 점차 확대되어 용어의 출처에 대한 정보를 제공할 수 있거나(International Organization for Standardization 1986), 보기 드문 용어, 외국어, 최신용어, 일반적인 사용법과 다른 방법으로 사용된 용어가 포함되어야 하고, 심지어 정

확한 정의를 포함할 수도 있다고 제시되었다(Lancaster 1986). 또한, 범위주기는 단지 용어가 사용되는 방법을 설명하는 것으로, 일반적으로 완전한 정의는 아니지만 사용방법에서 부정확하게 해석되거나 한정어로 쉽게 구별할 수 없는 다양한 의미를 가지는 경우에 제시되어야 한다고 주장되었다(Svenonius 1990).

이러한 다양한 역할로 인해 범위주기는 시소러스 개발초기에 예상한 것보다 용어의 명확한 표현을 위해 더욱 중요하게 다루어져 시소러스마다 범위주기의 수가 지속적으로 증가하였으며, 시소러스를 갱신하면 범위주기가 증가하는 경향으로 나타났다(Roberts 1985). 그러나 시소러스 개발지침과 전문가들이 범위주기의 실제적인 내용이나 스타일, 표현방식에 대한 일관성 있는 표준을 제시하기보다는 의미를 명확하게 할 필요가 있을 때마다 범위주기를 부가하도록 권고하였다. 따라서 범위주기가 체계적으로 사용되지 않아 필요하지만 누락되는 경향도 있고, 한편으론 범위주기가 용어에 대한 다양한 정보를 제공하기 때문에 범위주기의 기능이 너무 과다하다는 지적도 제기되었다(Sager, Sommer and McNaught 1982).

한편, Wall(1980)은 시소러스의 범위주기를 조사하여, 그 기능이 한정적 범위주기, 지시적 범위주기, 정의적 범위주기 등 세 가지가 있음을 관찰하였으며, Browne(1992)은 LISA의 범위주기에 대한 연구에서 정의적 범위주기가 모든 범위주기 중에서 상대적으로 매우 적음을 보고, LISA에서 범위주기를 통한 의미설명은

형식적인 것으로 보이므로, 의미설명을 위한 진정한 정의로서 정의적 범주주의의 중요성을 부각시켰다.

이처럼 시소러스 개발에 있어서 용어를 정의하고, 개념의 내포와 외연 및 개념과 관련개념을 구별하는 지적인 문제가 필요하다고 제기되었지만(Soergel 1974), 시소러스에서 의미를 설명하는 정의에 대한 문제는 진지하게 고려되지 않았다. 이러한 문제가 지금까지 시소러스에서 그다지 문제시되지 않은 것은 시소러스에서 의미를 다루려고 하는 의도가 없었기 때문이기보다는, 시소러스가 전문용어를 다루고 있을 뿐 아니라, 그 주제 분야의 전문가들만 이용하기 때문에 의미를 동일하게 인식하고 있다고 생각한 것으로 보인다. 또한 정의형태의 범주주의가 시소러스에 제공되더라도 정의를 작성하는 표준방법이 제시되지 않았을 뿐 아니라 디스크립터로 표현된 개념의 성질에 대한 정보가 많지 않은 것으로 나타났다. 따라서 시소러스에 용어의 의미와 사용법을 명백하게 설명하는 중요성에 의거하여 범주주의와 구분되는 정의를 시소러스에 통합하는 데 대한 관심이 증가하였다(Svenonius 1997; Hudon 1997).

3.3 정의를 통한 의미설명

앞서 기술하였듯이, 시소러스에서 용어의 의미는 주로 관계구조 및 범주주의를 이용하여 파악하고 있지만, 용어의 의미가 모호할 때는 그다지 효과적이지 않은 것으로 나타났다. 따

라서 최근에는 시소러스에서 사전의 정의와 같이 완전하고 상세한 정의의 중요성을 강조하고, 이를 시소러스에 통합하는 데 관심을 가져왔다.

사실 시소러스를 구성하는 것이 궁극적으로 용어를 정의에 근거하기 때문에, 정의가 시소러스에 구체적이고 체계적으로 나타나지 않는 것은 모순이라 할 수 있다. 즉, 복합어를 만드는 것을 결정하기 위해서는 구성되는 개별용어들의 정의를 알아야 하고, 동의어로서 인정하기 위해서는 먼저 개념을 정의해야 하며, 용어를 연관짓기 위해서는 그 개념의 중요한 특징을 파악해야 하며, 아울러 패킷 구조를 만들기 위해서는 먼저 개념의 성질을 확인해야 하기 때문에 정의는 대단히 중요하다 할 수 있다.

시소러스에서 정의를 포함하려는 생각은 시소러스를 처음 이용할 때부터 지속적으로 제기되어왔다. Moores(1985)는 용어가 정의되어 있지 않으면 색인자가 색인어를 선정하기 어렵다고 보고, 문맥에서 결정된 의미를 가진 용어를 제시하기 위해 정의를 부가할 것을 제안하였다. 그는 정의를 부가하는 것은 용어가 다른 의미를 가지고 더 이상 사용될 수 없음을 의미하는 것이 아니라, 그 이용을 제한하기 위한 것으로, 사전의 정의와 달라야 한다고 하였다. 또한 정의가 “알려지지 않은 것과 이미 알려진 것을 연결하는 수단”(Dahlberg 1981)으로서, 시소러스 개발과정에서 개념을 정의하고, 개념의 내포와 외연을 다루는 지적인 문제가 필요하다는 주장도 있었지만(Soergel 1974), 용어의 의

미를 설명하는 정의에 대한 문제는 진지하게 고려되지 않았다. 그러나, 시소러스에서 정의의 중요성은 다음과 같이 여러 번 강조되어 왔다.

Jansen(1970)은 개념에 대한 보충정보로서 정의가 매우 가치있을 것이라고 하였으며, Soergel(1974)은 시소러스의 관계구조를 통해 용어의 의미를 인식할 수 있지만, 한편으로 완전하고 공식적인 정의가 필요함을 인정하였다. Buchan(1989)은 정교한 계층구조로 잘 구성된 시소러스는 비록 정의가 없더라도 그 구조를 통해 그 의미가 이해될 수 있지만, 정의를 통해 더욱 정교한 계층구조를 만들 수 있다고 보고, 디스크립터의 정의를 이용함으로써 정보 검색이 더욱 강화될 것이라고 강조하였다. 아울러 그는 정의를 작성하는 데 시소러스의 관계구조가 도움이 될 수 있으며, 역으로 정의를 통해 시소러스의 관계를 더욱 잘 이해할 수 있다고 하였다. 또한 Svenonius(1997)는 시소러스에서 정의의 중요성을 강조하고 가능한 한 많은 정의를 포함하여 개념을 계층적으로 관련 지워야 한다고 주장하였다.

그러나 시소러스가 문헌에 있는 개념을 반영하는 현실적인 언어도구이고, 용어의 정의를 제시하는 것이 그다지 어렵지 않을 것이라는 생각이 지배적이어서, 이러한 제안들은 많은 주목을 받지 못하였다. 특히, 시소러스 개발에 관한 표준이나 지침에서 정의와 범위주기를 구별하여 이용하는 것을 권고하지 않아, 시소러스에서 정의를 제시하는 것은 그다지 중요하게 취급되지 않았다.

한편, 최근에 들어와 시소러스 개발지침이나 전문가들은 정의의 필요성과 함께 범위주기와 정의를 구별할 것을 제시하였다. 예컨대 ANSI/NISO Z39.19에서 “RT 참조는 ‘ships’와 ‘boats’ 처럼 중복의미를 가진 친족용어에 사용되는데, 이들 용어는 정확하게 정의될 수 있어 등가관계가 아니지만, 거의 상호교환하여 이용되므로 하나의 용어에 관심있는 이용자는 다른 용어도 상기해야”하며 “시소러스 관리 시스템은 레코드에 정의 필드를 제공하여 용어정의 뿐만 아니라 범위주기를 제공해야 하며, 정의는 사전적 의미를 제공하고 범위주기는 디스크립터를 구별해야 한다” (National Information Standards Organization 1994)고 기술하여, 표면적으로 정의의 필요성과 함께 시소러스에 정의를 추가할 것을 제시하고 있다.

또한 시소러스 개발지침에서도 “용어의 의미를 위해 시소러스에 범위주기와 정의를 부가하면 시소러스가 더욱 효율적”이라고 지적하였으며, 나아가 “일반적으로, 시소러스에서 사전 정의와 같은 완전한 정의가 제시되지 않았으나, 제한된 정의와 심지어 확장된 정의가 종종 시소러스 구조에 의해 전달된 의미를 보충하는데 필요하다”고 인정하여 정의에 대해 변화하는 인식을 보여주었다(Aitchison and Gilchrist 1987; Aitchison, Gilchrist and Bawden 2000).

그러나, 이들 표준이나 지침에서는 정의를 사전적 의미로 국한하거나 정의가 용어의 의미

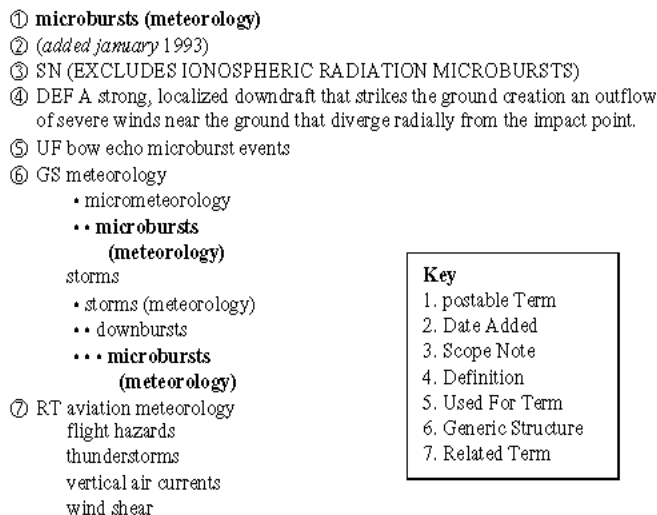
를 보충적으로 설명하는 역할로서만 제시하였을 뿐 구체적인 지시가 없으며, 또한 정의의 이용을 명확하게 권고하고 있지 않아 아직까지 시소러스에 완전한 정의는 거의 없다고 볼 수 있다.

이처럼 시소러스 개발지침에서 정의에 대한 방침이 부족함에도 불구하고, 일부 시소러스는 사전과 통합하거나 범위주기와 구별되는 정의를 부가하여 구축하기도 하였다. 예컨대 “The Computer Database Thesaurus & Dictionary”(1983)와 “The Managements Contents Database Thesaurus & Dictionary”(1984)는 시소러스와 사전을 통합하여 모두 제시하고 있다. 또한 “NASA Thesaurus”도 지금까지 수차례 개정을 거쳐 오면서 정의를 포함한 대표적인 시소러스라 할 수 있는데, 이 시소러스는 1988년과 1994년판

에서는 범위주기와는 별도로 시소러스를 보완하기 위해 용어정의를 포함한 특수한 형태의 사전인 “Definitions”란 표제의 개별권(제3권)을 만들어 정의를 따로 제공하였다.

그러나 이러한 형태의 시소러스 사전(thesaurus dictionary)은 디스크립터의 의미를 더 잘 설명하기 위해 사전과 시소러스를 결합하려는 시도로 보기도 하지만(Buchan 1989), 정의가 동일한 디스크립터에 연결된 다른 의미정보와 구분되어 따로 제시되기 때문에 이런 형태로 정의를 제공하는 것은 시소러스 용어의 의미를 설명하는 중요한 수단이기보다는 기껏해야 시소러스를 보충하는 역할을 하여 품질과 유용성에 문제가 있다고 볼 수 있다.

이런 이유에서인지 NASA Thesaurus는 1998년판부터 <그림 6>과 같이 정의를 ‘DEF’라는 항목으로 부가시켜 제시하고 있다. 그러



<그림 6> NASA Thesaurus의 정의를 포함한 구조

나 용어를 명백하게 설명하기 위해 정의를 제시할지라도 표준화된 정의가 이용되지 않다고 지적되었듯이(Aitchison and Gilchrist 1987), NASA Thesaurus에서도 정의를 기술하는데 표준화된 지침이 없이 소속된 사전편찬자가 여러 자료를 참조하여 만들어 다소 주관적임을 인정하고 있다(National Aeronautics and Space Administration 1998).

여하튼 정의를 제시한 대부분의 시소러스는 정의를 작성하는 표준적인 지침 없이 여러 사전에서 정의를 추출하여 부가하였지만, 시소러스에서 정의를 통해 용어의 의미를 보충할 필요성을 보여주었다. 특히, 1980년대 초반부터 사회과학 분야에서 정보 시스템이 증가하여 시소러스가 더욱 중요하게 부각됨에 따라, 시소러스에서 정의의 기능이 강조되었으나 정의를 작성하는 규칙과 모델을 만드는 데 대한 어려움으로 진전을 보지 못하였다.

한편, 전문용어 분야에서도 시소러스가 부분적으로 관련 있음을 인식하고, 시소러스에서 의미를 설명하는 문제에 관심을 두었다. Sager(1990)는 시소러스에서 용어의 의미를 나타내기 위한 것으로 용어 자체, 다른 개념과의 관계 및 정의를 이용할 수 있는데, 그 중에서 정의가 가장 정확하고 유용한 서술적 묘사를 제공한다고 보고, 용어 간의 의미관계뿐만 아니라 표준화된 정의를 포함하는 전문용어 시소러스(terminological thesaurus)의 개발을 제시하였다.

4. 정의의 작성과 시소러스에의 통합

4.1 디스크립터의 의미정보 수집

이 연구는 시소러스에서 용어의 의미정보로서 정의의 필요성과 중요성에 근거로 실제로 시소러스에 제시된 용어에 대한 정의를 작성하고, 그러한 정의를 기존의 시소러스에 통합하여 확장된 시소러스를 구축하고자 한다.

이러한 확장된 시소러스를 구축하기 위한 기본 시소러스 데이터로 “KEDI 교육 시소러스”의 31개 주제 중 18개 주제에서 선택한 139개의 디스크립터와 관계를 이용하였다. 나머지 13개 주제는 특정 교과교육을 대상으로 한 개별 교과에 대한 주제를 다른 용어가 대다수여서 배제하였다. 선택한 18개 주제 중 17개 주제에서는 교육학 분야에 전체적인 적용가능성에 염두를 두어 74개의 디스크립터를 각 주제마다 약간씩 임의로 선정하였으며, 나머지 한 주제인 “교육과정” 부문에서 기술된 총 65개 디스크립터는 특정 하위주제에서의 적용가능성을 염두에 두고 모두 선정하였다. 여기서 선정한 디스크립터와 관계는 타당한 것으로 가정하였다.

일반적으로 특정 주제의 용어의 정의를 포함한 의미정보의 수집은 그 분야의 정보원에서 추출하거나 주제전문가의 자문을 통해 얻은 정의들을 이용하는데, 이 연구에서는 기본 시소러스의 139개 디스크립터에 대한 의미정보를 얻을 정보원으로 다음의 전문사전, 일반사전

및 텍스트북 등을 이용하였다.

• 전문사전

教育辭典編纂委員會, 教育學大辭典
(1994)

南億祐 등 편, 최신 교육학대사전(1990)

서울대학교 교육연구소 편, 教育學用語
辭典, 전정판(1994)

현종익, 이학춘, 교육학 용어사전(2002)

• 일반사전

엠파스 국어사전

〈<http://kordic.empas.com>〉

엠파스 백과사전

〈<http://100.empas.com>〉

• 텍스트북

김정환, 강선보, 교육학개론(1997)

박준영, 교육학개론(1998)

임창재, 교육학(2001)

이러한 정보원을 이용하여 의미정보를 조사
한 방법은 다음과 같다.

첫째, 139개의 디스크립터에 대한 의미정보
를 얻기 위해 교육학 분야의 전문사전 4개를
조사하여 해당 디스크립터에 대한 정의가 나타
나면, 그 정의를 출처와 함께 기록하였다. 이
과정에서 비록 용어가 정확하게 일치되지 않더
라도 해당 디스크립터의 번역이나 기술방식의
차이를 고려하여 약간 다르게 기술되어 있는 용
어도 동일한 의미로 인정하여 포함하였다. 그
결과 1개 이상의 정의가 나타난 디스크립터는
100개로 조사되었다. 이는 시소러스에 수록된
디스크립터가 전문용어일 것이라는 인식에도

불구하고, 전문사전에 그 용어가 나타나지 않는
것으로 보아 시소러스 구축 시 선택한 용어가
다소 부적절한 것으로 보여지는 부분이다.

둘째, 전문사전에서 의미가 조사되지 않은
디스크립터는 그것을 구성하고 있는 용어를 분
할하여 조사하였다. 이는 대부분의 전문용어가
복합명사로 구성되어 있는 경우가 많기 때문이
기도 하다. 이 과정에서는 분할된 용어에 대한
의미정보를 수집하기 위해 앞서 조사된 100개
의 디스크립터에 대한 정의와 함께 기타 국어
사전이나 백과사전을 이용하여 조사하였으며,
그 외에는 텍스트북을 조사하여 그 용어의 개
념을 설명하는 부분을 조사하여 추가로 30개의
디스크립터에 대한 의미정보를 수집하였다.

이러한 방법을 통해 실험대상인 139개 디스
크립터 중 130개에 대한 의미를 얻었으며, 나
머지 9개 디스크립터는 위의 정보원에서 조사
되지 않아 배제하였다. 배제된 용어들은 대부
분 그 의미를 해석하기에 포괄적이거나 전문용
어로 보기에 다소 문제가 있는 용어로 판단되
었으며, 다른 하위주제에 나타난 디스크립터에
서도 그러한 현상은 나타날 것이므로 시소러스
의 용어선택에서 신중을 기해야 할 부분으로
고려된다.

4.2 표준정의작성을 위한 정의 모델

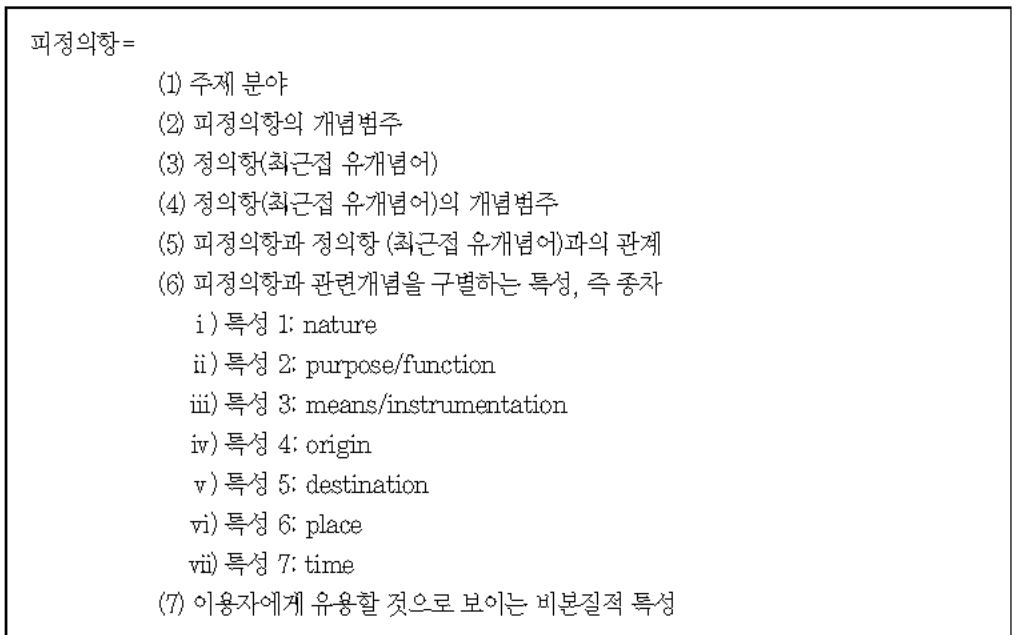
일반적으로 전문사전의 정의는 그에 대한
공식적인 평가나 판단기준은 없지만, 그 분야
의 전문가에 의해 작성되었기 때문에 의미를

수집하기 위한 좋은 정보원으로 볼 수 있다. 그러나 대부분의 전문사전에서 조사한 정의를 보면 일정한 규칙이나 기준이 없이 작성되어 표준적인 정의양식이 없는 것으로 판단되었다.

정의를 작성하는데 표준화 할 필요성은 전문용어 분야에서 많이 논의되었다. 전문용어 분야에서는 정의의 형식과 방법뿐만 아니라 내용면에서 표준화할 필요성을 인식하고 개념의 성질이 동일한 것은 동일한 양식으로 명확하게 정의하기 위해 필요한 정의요소를 조작하는 정의 모델을 연구하여 왔다. 정의를 표준화한다는 것은 정의양식을 일치시키는 것으로, 정의양식을 표준화한 정의 모델은 항상 같은 순서로 용어의 본질적인 요소를 파악하여 정의를 쉽게 구성하고 조직할 수 있게 한다. 따라서 이

연구에서도 분명하고 간결한 표준화된 정의를 작성하기 위해 정의 모델을 설정하였다.

이 연구에서 설정한 정의 모델의 기본구조는 Sager와 L'Homme(1994)의 정의 모델과 Hudon(1996)의 확장정의 모델을 근거로 하였다. Sager와 L'Homme는 정의를 표준화하기 위해 정의를 기술하는 양식을 범주화하고 제한하는 등 분석적 정의를 정규화하여 기존 용어의 정의를 표준화하는 정의 모델을 제시하였다. 그들의 정의 모델은 정의에서 비일관적이고, 불규칙적이며, 불완전한 자연언어적인 요소를 줄이고, 일관성 있고 완전한 정의를 만들기 위해 정의에 본질적으로 필요한 7가지 항목으로 구성되어 있다. 이후 Hudon(1996)은 Sager와 L'Homme가 제시한 정의 모델에서



〈그림 7〉 Hudon의 확장된 정의 모델

<p>용 어:</p> <p>(1) 주제 및 하위주제</p> <p>(2) 최근접 유개념어</p> <p>(3) 특성</p> <p>① 본질</p> <p>② 목적</p> <p>③ 수단</p> <p>④ 상태변화</p> <p>⑤ 유사</p> <p>⑥ 위치</p> <p>⑦ 반대</p>
--

〈그림 8〉 새로 설정한 정의 모델

다른 항목은 그대로 수용하고, 특성항목을 명확히 구분하여 우선순위에 따라 다시 7가지로 세분하여 확장한 정의 모델을 〈그림 7〉과 같이 제시하였다.

이러한 선행 정의 모델을 기초로 하여, 이 연구에서는 표준정의를 작성하기 위해 〈그림 8〉과 같은 새로운 정의 모델을 설정하였다.

이 정의 모델에서 항목(1)은 주제와 그 주제의 하위주제를 다룬다. 이 연구에서 사용한 “KEDI 교육 Thesaurus”는 교육학 분야를 31개의 하위주제로 구분하여 제시하고 있기 때문에, “교육학 하위주제” 형식으로 적용하였다. 항목(2)는 일반적으로 분석적 정의방식에서는 정의항 내에 피정의항의 유개념어를 기본적으로 포함하는 것으로 보기 때문에, 상위어로서 최근접 유개념어를 부여하도록 하였다. 이는 시소러스 구조에서 계층관계를 나타내는

데 사용할 수 있는데, 특별한 경우를 제외하고는 이 모델로 작성한 모든 표준정의에서 최근접 유개념어가 정의문의 마지막에 제시되도록 하였다.

마지막으로 항목(3)은 정의항에서 피정의항과 동일한 수준의 다른 디스크립터와 구별되는 특성을 제시한 것으로 7가지를 설정하였다. 특히 항목(3)의 특성에 포함될 요소는 학문 분야에 따라 그 중요도가 다르게 나타날 수 있는데, 이 연구에서는 실험을 위해 선정한 130개의 디스크립터의 의미정보를 대상으로 선행 연구자들의 정의 모델에서 제시된 모든 특성요소들을 적용하여 하나 이상이 나타나는 특성은 특성요소로 설정하고, 그렇지 않은 것은 특성요소에서 제외하였다. 따라서 여기에 제시된 특성요소는 교육학 분야에 타당성이 있을 것으로 판단된다.

4.3 표준정의작성을 위한 정의규칙

정의규칙은 수집한 의미정보를 앞서 설계한 정의 모델에 의거하여 분석한 내용을 이해하기 쉽고 일관성 있게 표준정의를 작성하기 위해 설정한 것으로, 모든 정의가 동일한 형식을 가지고 이해하기 쉽게 작성하기 위해 규칙이다. 시소러스에 부가할 표준정의를 작성하기 위해, 이 연구에서는 어휘의미적 규칙과 형태구문적 규칙으로 구분하여 설정하였다. 전자는 정의의 내용에 관한 것이고, 후자는 정의의 형식과 표현에 관한 것이다. 대부분 선정된 정의규칙은 전문용어 분야에서 제시한 것(Ndi Kimbi 1994)을 기초로 하였으며, 일부는 표준정의를 만드는 동안 새로 만들거나 수정한 것을 포함하여 전체적으로 다음과 같다.

4.3.1 어휘의미적 규칙

전문용어에서 정의는 기본적으로 피정의항과 같거나 변형된 용어를 재이용하지 않도록 하며, 핵심어는 정의되는 개념의 상위개념이 되도록 하고 있다. 물론, 복합어일 경우에는 예외가 있는데, 그러한 경우 상위개념의 모든 특성은 그대로 상속되기 때문에 모두 열거할 필요는 없다. 이에 의거하여 다음과 같은 규칙을 만들었다.

- 규칙 1. 정의항에서는 피정의항의 상위개념어를 반드시 포함한다.
- 규칙 2. 정의항에서는 피정의항에 이용된 용어를 반복하여 사용하지 않는다.

- 규칙 3. 정의항의 상위개념어가 다른 곳에서 정의될 때, 그것의 특성을 다시 제시하지 않는다.

용어가 정의되더라도 주제와 관점에 따라 다르게 정의되고, 더 이상의 상위개념을 파악할 수 없는 경우는 유개념을 부여하지 않아야 한다. 또한 상위개념으로 선정된 용어가 시소러스에 없을 때는 새로운 용어를 부여하거나 적절한 용어가 없어 범주를 나타내는 경우는 하위개념을 고정하기 위해 그대로 사용한다. 이에 의거하여 다음과 같은 규칙을 만들었다.

- 규칙 4. 상위개념을 파악할 수 없는 개념은 상위개념어를 부여하지 않는다.
- 규칙 5. 상위어가 시소러스에 없을 때는 새로운 용어를 부여하며, 범주를 나타내는 용어는 그 개념을 고정하는데 사용하나 상위어로 인정하지 않는다.

정의항에 피정의항의 동의어가 나타나지 않아야 하며, 간혹 정의에서 피정의항의 의미를 명확하게 나타내기 위해 반대되는 개념을 부여하는 경우도 있으므로 이에 따라 다음과 같은 규칙을 만들었다.

- 규칙 6. 정의항에서 피정의항의 동의어를 사용하지 않는다.
- 규칙 7. 필요에 따라 정의항에 피정의항의 반대개념을 부여할 수 있다.

4.3.2 형태구문적 규칙

표준정의를 일반적으로 서술적인 일반사전이나 백과사전과 달리 하나의 문법적 구조를

가진 문장으로 분명하고 간결하게 그 의미를 전달해야 하며, 또한 이해하기 쉽고 단순한 문법적 규칙을 가져야 한다. 이를 위한 다음과 같은 규칙을 만들었다.

- 규칙 8. 정의는 문법적으로 완전한 하나의 문장으로 구성된 명사구로 한다.
- 규칙 9. 동사는 현재시제로 하고 능동태가 되어야 한다.

4.4 표준정의의 작성

시소러스에서 표준정의의 기능은 개념에 대해 자세한 정의를 쓰거나 언어이용을 규정하는 것이 아니라, 디스크립터로 표현된 개념에 대해 충분한 의미정보를 제공하여 이용자가 디스크립터를 선택하여 이용할 때 다른 디스크립터와 구별을 명확히 하는 것이다. 피정의항이 속하는 개념의 상위개념을 참조하고 피정의항과 유사한 개념을 구별하는 본질적 특성을 언급하는 분석적 정의가 완전한 정의는 아니지만 전문용어를 정의하는 데 가장 이상적인 것으로 인정받듯이, 표준정의는 이러한 분석적 정의형식을 기본으로 하여 내용과 형식면에서 정의를 표준화한 것이다.

따라서 이 연구에서 설정한 정의 모델과 정의규칙은 이러한 요구에 부합한다고 할 수 있으며, 이에 의거하여 정보원을 통해 표준정의를 작성하기 위해 의미정보가 조사된 130개의 디스크립터에 대해 표준정의를 작성하였다. 표준정의를 작성한 방법은 앞서 기술하였듯이,

디스크립터에 대한 정의를 얻기 위한 방법은 크게 관련 전문사전에서 그 용어의 정의가 있는 100개의 디스크립터를 대상으로 먼저 표준정의를 작성하였으며, 나머지 30개의 디스크립터에 대한 표준정의는 이미 작성된 100개의 디스크립터의 표준정의, 일반사전의 정의 및 관련 텍스트북을 조사하여 얻은 정의 및 설명을 토대로 작성하였다

다음은 디스크립터에 대해 조사된 정의와 설명을 어떻게 정의 모델에 적용하고 정의규칙을 이용하여 표준정의를 작성한 사례이다.

4.4.1 <사례 1> '전문교육' (Professional Education)의 표준정의

이 용어의 정의는 조사한 전문사전 중 하나는 두 개의 정의를 가지고 있어, 총 4개의 정의가 조사되었는데, 그 내용은 다음과 같다.

- 정의 1: 전문교육은 특수교육?직업교육 등으로 불리기도 한다. 전문교육은 특정한 직업 등에 요구되는 제 특성을 길러주고, 대상은 특수부분의 문화를 대상으로 하고, 실현하려는 궁극적인 것은 기능가치와 학문가치를 실현해 준다.
- 정의 2: 보통교육에 대립하는 특수한 교육을 뜻하며, 또 일반의 기초적인 수양에 대립해서 어떠한 종류의 업무에 필수적인 지식이나 기능을 양성하는 것을 목적으로 하는 교육을 의미한다.
- 정의 3: 전문가를 양성하는 교육
- 정의 4: 보통교육에 대하여 특수한 교육을

뜻하며, 또 일반의 기초적인 수양에 대해서 어떠한 종류의 업무에 필수적 지식이나 기능을 양성하는 것을 목적으로 하는 교육을 의미한다.

제시된 용어는 모두 교육학 분야에서 주제가 교육과정인 부분들을 다루기 때문에 정의 모델의 항목 중 주제 분야는 ‘교육학 교육과정’이 되며, 이 항목은 이후 제시되는 용어에 모두 동일하게 적용된다. 그리고 조사한 정의를 분석한 결과, 용어의 유개념어는 ‘교육’이며, 다음의 네 가지 특성형태가 나타났다.

- 특성①(본질): 업무에 필수적인 지식이나 기능을 가진 전문가를 양성하는
- 특성②(목적/기능): 기능가치와 학문가치를 실현하려는
- 특성⑤(유사): 특수교육, 직업교육 등으로 불림
- 특성⑦(반대): 보통교육에 대하여 이러한 내용을 토대로 ‘전문교육’의 표준정의를 다음과 같이 작성하였다.
- 표준정의: 보통교육에 대립하여 기능가치와 학문가치를 실현하기 위해 업무에 필수적인 지식이나 기능을 가진 전문가를 양성하는 교육

4.4.2 <사례 2> ‘교육과정개발’(Curriculum Development)의 표준정의

이 용어의 정의는 세 개가 조사되었으며, 그 내용은 다음과 같다.

- 정의 1: 교육목적과 교육내용의 체계, 그리고

이를 효과적으로 전달하기 위한 교육방법·교육평가·교육운영 등에 대한 종합적인 계획이 담긴 문서를 만드는 활동

- 정의 2: 새로운 교육과정안을 만들어 내는 것.
- 정의 3: 아동·학생의 교육을 위한 지침을 기획하는 행위로서 교육과정의 의미와 그것을 행하는 주체, 수준, 상황에 따라 의미의 차이가 있다.

이 용어의 유개념어는 한글복합명사의 성질상 뒷부분이 상위어를 나타내는 경우가 대부분이기 때문에 ‘개발’이 되어야 하지만, 그 용어 자체로는 주제와 비교해 볼 때 상당히 추상적인 용어이고, ‘교육과정’에서의 ‘개발’을 의미하기 때문에 유개념어를 ‘교육과정’으로 하였으며, 다음의 두 가지 특성형태가 나타났다.

- 특성①(본질): 교육과정에 대한 종합적인 계획이 담긴 문서를 만드는
- 특성②(목적/기능): 새로운 교육과정안 만들어 내기 위해

이러한 내용을 토대로 ‘교육과정개발’의 표준정의를 이 용어의 앞부분이 유개념어가 되기 때문에 다음과 같이 작성하였다.

- 표준정의: 새로운 교육과정안을 만들어 내기 위해 교육과정에 대한 종합적인 계획이 담긴 문서를 만드는 것.

이러한 과정으로 정보원에서 조사한 정의 및 설명을 정의 모델에 적용하여 분석한 다음, 정의규칙에 의거하여 130개의 표준정의를 작성하였으며, 표준정의를 작성하면서 나타난 결과를 통해 전체적으로 다음의 사항을 확인하였

〈표 1〉 특성별 사용횟수

특성	사용횟수
특성①(본질)	120개
특성②(목적/기능)	39개
특성③(수단/도구)	5개
특성④(상태변화)	12개
특성⑤(유사)	37개
특성⑥(위치)	20개
특성⑦(반대)	11개

다. 정의 모델에서 유개념어의 확인을 통해 용어 간의 계층관계에서 상위어로서 최근접 유개념어를 확실하게 파악할 수 있었는데, 전체 130개의 용어 중 유개념어가 있는 디스크립터는 125개로 나타났으며, 나머지 다섯 개의 용어는 최상위 상위어로 파악되어 유개념어가 나타나지 않은 것으로 조사되었다. 또한 각 디스크립터의 의미정보를 분석하여 도출한 특성이 〈표 1〉과 같이 이용된 것으로 나타났다.

이들 특성 중에서 특성①(본질)이 가장 많이 이용되었으며, 그 다음으로 특성②(목적/기능)이 많이 이용된 것으로 나타났다. 특히 특성⑤(유사)가 나머지 특성보다 상대적으로 많이 이용된 것으로 나타났는데, 이는 정의방법으로 유사어를 이용하여 용어의 의미를 전달하는 것이 효과적임을 보여준다고 할 수 있다. 이러한 특성형태는 정의에서 다음의 문장형태로 나타나는 것을 근거로 하여 특성을 조사하였다.

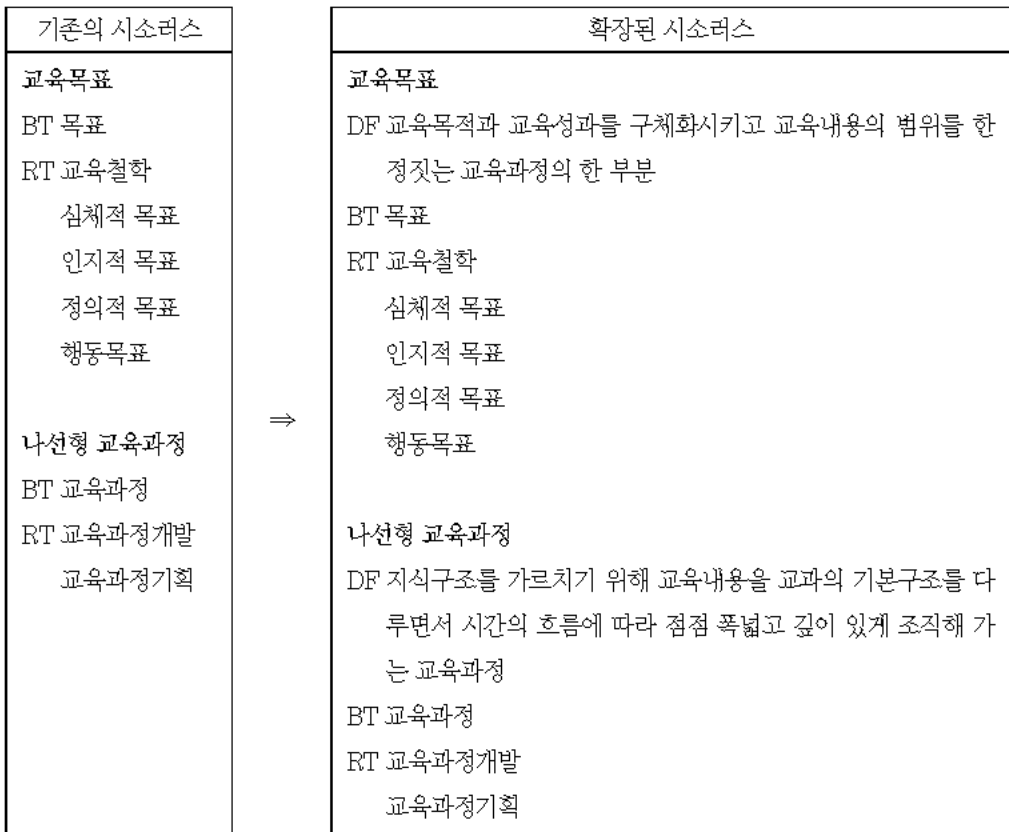
- 특성①(본질): "...을 하는"
- 특성②(목적/기능): "...을 위한", "...하도록"

- 특성③(수단/도구): "...하여", "...를 통해"
- 특성④(상태변화): "...에서 생긴"
- 특성⑤(유사): "...라고 하는"
- 특성⑥(위치): "...에서"
- 특성⑦(반대): "...와 다른", "...와 구분하는", "...에 대립하여", "...이 아니라"

한편, 이러한 과정으로 다시 작성된 표준정의는 해당주제에 영향을 주기 때문에 그 분야 전문가들이 비준하는 것이 필요하다. 그러나 이 연구에서 작성한 표준정의는 적용범위가 제한적이고, 표준정의를 작성하기 위해 새로운 정의를 제시한 것이 아니라 그 분야에서 타당한 것으로 가정한 기존의 전문사전의 정의를 이용하였기 때문에 특별한 검증과정을 거치지 않았다.

4.5 표준정의를 통합한 온라인 시소러스 구현

프로토타입의 확장된 시소러스를 구축하기 위해 기본 시소러스에 새로 작성한 표준정의를



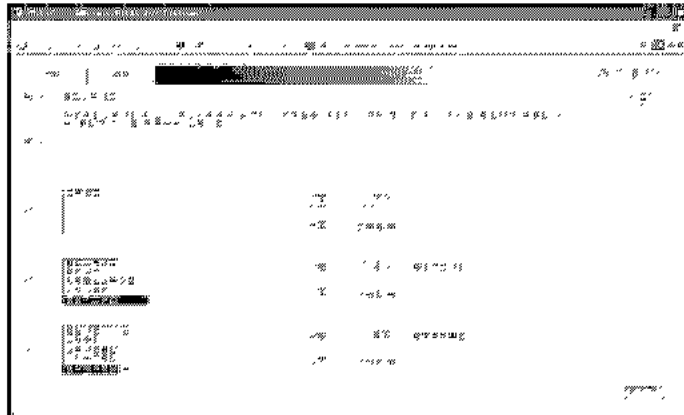
〈그림 9〉 표준정의를 통합한 확장된 시소러스의 사례

통합하였다. 이를 통해 확장된 시소러스의 구조는 기존의 시소러스에 DF라는 제시어를 부여하여 정의가 추가되어 〈그림 9〉와 같이 변경되었다.

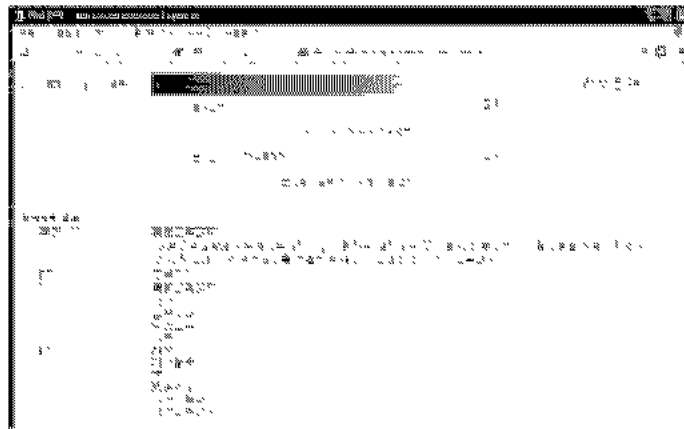
또한 확장된 시소러스를 웹에서 효율적으로 조직하고 이용하기 위해 온라인 환경에서 관리하고 이용하는 시소러스 시스템을 구현하였다. 구현방법은 표준정의를 포함하고 있는 시소러스의 관리와 이용의 효율성에 초점을 두었기 때문에, 웹에서 많이 볼 수 있는 관계형 데이터베이스와 스크립트를 이용하여 HTML 형식으

로 다음과 같이 만들었다. 〈그림 10〉은 관리부문으로 시소러스에 새로운 내용을 입력하거나, 수정하는 데 이용할 수 있다. 새로운 내용을 입력할 때는 시소러스의 디스크립터로 이미 저장되어 있는지를 확인할 수 있으며, 이미 그 내용이 있다면 수정이 가능하고, 없다면 새로운 용어와 관계구조에 필요한 내용을 입력하여 저장할 수 있다.

〈그림 11〉은 온라인 시소러스를 이용하는 '검색' 부문으로, 여기서는 '용어검색'과 '정의검색'의 두 가지 모듈로 구성하였다. 용어검색



〈그림 10〉 온라인 시소러스의 관리화면



〈그림 11〉 온라인 시소러스의 정의검색화면

모듈은 시소러스 구조에서 제시된 우선어 및 비우선어를 포함한 모든 용어를 대상으로 검색하여 해당되는 용어가 나타나는 모든 시소러스 구조를 제시하여 이용자가 전체적으로 파악할 수 있다. 또한 정의검색 모듈은 표준정의를 통해 접근을 가능하게 하는 것으로, 표준정의를 저장할 때 자동으로 색인하여 저장된 명사형태의 색인어를 대상으로 검색하여 표준정의에 이

색인어가 포함된 모든 시소러스 구조를 제시하도록 하였다.

5. 결 론

지금까지 시소러스는 개념을 나타내는 용어의 의미를 제시하기 위해 주로 관계구조나 범위주기를 이용하여 왔으나, 시소러스가 양적으

로 팽창해짐에 따라 기존의 관계구조나 범위주기로는 그 의미를 제시하는 데 한계를 보여 왔다. 즉 관계구조나 범위주기만으로는 용어가 지닌 본질적인 속성을 이해하기 어려울 뿐만 아니라, 그 범위를 한정하기 어려운 면이 있어 정보검색과정에서 용어를 선택할 때 일관성을 유지하고 용어의 의미와 용어 및 개념 간의 관계를 분명히 파악하기 어려운 점이 있었다. 이에 일부 시소러스는 사전의 정의를 포함하는 '시소러스 사전'의 형식을 취하였으나, 수록한 정의의 내용이나 형식이 만족스럽지 않아 명확한 의미를 전달하는 데 문제가 있는 것으로 지적되었으며, 이에 정의를 표준화할 필요성이 제기되어 왔다.

이 연구는 용어의 표준정의를 시소러스에서 유용함을 확인하기 위해, 전문용어 분야에서 거론된 정의 모델과 정의규칙에 의거하여 내용과 형식면에서 일관성 있는 용어의 표준정의를 작성하고, 그러한 표준정의를 기존의 시소러스에 포함하여 더욱 확장된 시소러스를 구축하는 것을 목적으로 하였다. 이를 위해서 각 디스크립터에 대한 의미정보를 해당 분야의 정보원을 이용하여 정의나 설명을 수집하여 기록하고, 전문용어 분야에서 제시된 이론을 근거로 하여 본 연구 목적에 맞게 작성한 새로운 정의 모델과 정의기술방식을 체계화한 정의규칙에 따라 표준정의를 작성하였으며, 이들을 기본 시소러스에 통합하여 더욱 확장된 시소러스를 프로토타입으로 구축하였다.

이러한 과정을 통해 시소러스에서 각 디스

크립터에 대한 의미를 제공하는 도구로서 표준정의를 모두 부가되어 용어의 의미를 명확하게 파악할 수 있게 하였다. 또한 온라인 환경에서 확장된 시소러스의 유용성을 확인하고자 데이터베이스와 웹을 연동한 온라인 시소러스를 구현하여 시소러스에 제시된 모든 용어뿐만 아니라 표준정의를 기술하는 데 사용된 키워드를 탐색키로 이용하게 하는 다양한 접근점을 제시하는 검색부문을 테스트하였다.

이로써, 표준정의를 이용하여 확장된 시소러스는 용어의 의미를 쉽게 이해할 수 있도록 하여 정보검색과정에서 용어의 의미를 정확하게 파악하여 더욱 분별 있게 용어를 선택할 수 있는 가능성을 확인하였다. 따라서 이 연구는 "시소러스는 여전히 존재하지만, 시소러스가 계속 가치 있는 도구가 되기 위해서는 변해야 하고, 어떻게 변화할 지 예상하기 어렵다"는 Milstead(1995)의 지적처럼, 시소러스가 정보환경의 상황에 적응하여 변화하는 시소러스의 사례를 보여주었다.

참고문헌

- 김태수, 2001. 용어 정의를 도입한 시소러스 개발 연구. 『정보관리학회지』, 18(2): 231-254.
- Aitchison, J., 1994. *Words in the Mind: an Introduction to the Mental Lexicon*. Oxford: Blackwell.
- Aitchison, Jean and Alan Gilchrist, 1987.

- Thesaurus Construction: a Practical Manual*, 2nd ed, London: Aslib.
- Aitchison, Jean, Alan Gilchrist, and David Bawden. 2000. *Thesaurus Construction and Use: a Practical Manual*, 4th ed, London: Aslib.
- Bates, M. J. 1986. "Subject access in on line catalogs: a design model." *Journal of the The American Society for Information Science*, 37(6): 357-376.
- Bonjour, M., and G. Falquet. 1994. "Concept Bases: A Support to Information Systems Integration." In *Proceedings of 6th Conference on Advanced Information Systems Engineering(CAiSE*94)*, 242-255.
- Browne, G. 1992. "Scope notes for LISA subject headings." *Online Review*, 6: 3-15.
- Buchan, R. L. 1989. "Intertwining thesauri and dictionaries." *Information Services & Use*, 9: 171-175.
- Dahlberg, I. 1981. "Conceptual definitions for interconcept." *International Classification*, 8: 16-22.
- Dahlberg, I. 1989. "Concept and Definition Theory." In *Classification Theory in the Computer Age: Conversations Across the Disciplines. Proceedings from the Conference, November 18-19, 1988*. Albany, New York: Rockefeller College Press, University of Albany, State University of New York, 12-24.
- Hudon, Michele. 1996. "Preparing terminological definitions for indexing and retrieval thesauri: a model." In Rebecca Green (ed), *Knowledge organization and change. Proceedings of the fourth International ISKO Conference, Washington, DC, 15-18 July 1996*. Frankfurt/Main: Indeks Verlag, 363-369.
- Hudon, Michele. 1997. *An Assessment of the Usefulness of Standardized Definitions in a Thesaurus Through Interindexer Terminological Consistency Measurements*. Ph.D. diss., University of Toronto.
- International Organization for Standardization. 1986. *ISO 2788 1986: Guidelines for the Establishment and Development of Monolingual Thesauri*, 2nd ed. Geneva: ISO.
- International Organization for Standardization. 1990. *ISO 1087 1990:*

- Terminology Vocabulary*. Geneva: ISO.
- Jansen, R. 1970. "Some observations on thesaurus problems." In *Proceedings of the International Conference on General Principles of Thesauri Building, Warsaw, 23-27 March 1970*. Warsaw: Documentation and Scientific Information Center of the Polish Academy of Sciences, 15-30.
- Lancaster, F. W. 1986. *Vocabulary Control for Information Retrieval*. 2nd ed. Arlington, VA: Information Resources Press.
- McNaught, J. 1983. "The generation of term definitions from an on line terminological thesaurus." In *First Conference of the European Chapter of the Association for Computational Linguistics: Proceedings of the Conference, 1-2 September 1983, Pisa, Italy*. Morristown, NJ: Association for Computational Linguistics, 90-95.
- Miller, George A. et al. 1990. "Introduction to WordNet: An on line lexical database." *International Journal of Lexicography*, 3(4): 235-244.
- Milstead, Jessica L. 1995. "Invisible thesauri: the year 2000." *Online and CDROM Review*, 19(2): 93-94.
- Mooers, C. N. 1985. "The indexing language of an information retrieval system." In *Theory of Subject Analysis: a Sourcebook*. Littleton, CO: Libraries Unlimited, 247-261.
- National Aeronautics and Space Administration. 1998. *NASA Thesaurus*. Washington: NASA.
- National Information Standards Organization. 1994. ANSI/NISO Z39.19-1993: *Guidelines for the Construction, Format, and Management of Monolingual Thesauri*. Bethesda, MD: NISO.
- Ndi Kimbi, Augustin. 1994. "Guidelines for terminological definitions: the adherence to and deviation from existing rules in BS/ISO 2382 Data Processing and Information Technology Vocabulary." *Terminology*, 1: 327-350.
- Nkwenty Azeh, B. 1994. "The use of thesaural facets and definitions for the representation of knowledge structures." In *Hanne Albrechtsen and Susanne Oernager (eds.)*,

- Knowledge Organization and Quality Management: Proceedings of the Third International ISKO Conference, 20-24 June, 1994, Copenhagen, Denmark*. Frankfurt: Indeks Verlag, 374-381.
- Riggs, F. W. 1981. *Interconcept Report: a New Paradigm for Solving the Terminology Problems of the Social Sciences*. Paris: UNESCO.
- Roberts, N. 1985. "Concepts, structures and retrieval in the social sciences up to c.1970." *Social Science Information Studies*, 5: 55-67.
- Sager, J. C. 1990. *A Practical Course in Terminology Processing*. Amsterdam: J. Benjamins.
- Sager, J. C., and M. C. L'Homme. 1994. "A model for the definition of concepts: rules for analytical definitions in terminological databases." *Terminology*, 1: 351-373.
- Sager, J. C., H. L. Somers, and J. McNaught. 1982. "Thesaurus integration in the social sciences. Part III: Guidelines for the integration of thesauri." *International Classification*, 9: 64-70.
- Soergel, D. 1974. *Indexing Languages and Thesauri: Construction and Maintenance*. Los Angeles: Melville.
- Soergel, D. 1985. *Organizing Information: Principles of Data Base and Retrieval Systems*. Orlando: Academic Press.
- Svenonius, E. 1990. "Design of controlled vocabularies." In *Encyclopedia of Library and Information Science*. New York: Marcel Dekker, 82-109.
- Svenonius, E. 1997. "Definitional approaches in the design of classification and thesauri and their implications for retrieval and for automatic classification." In: *Knowledge Organization for Information Retrieval, Proceedings of the 6th International Study Conference on Classification Research, University College London, 16-18 June 1997*. The Hague, Netherlands: International Federation for Information and Documentation, 12-16.
- Management Contents. 1983. *The Computer Database Thesaurus & Dictionary*. Northbrook IL: Management Contents.
- Management Contents. 1984. *The*

- Management Contents Database Thesaurus & Dictionary, Northbrook IL: Management Contents.
- Wall, R. A. 1980. "Intelligent indexing and retrieval: a man machine partnership." *Information Processing and Management*, 16: 73-90.
- Willetts, M. 1975. "An investigation of the nature of the relation between terms in thesauri." *Journal of Documentation*, 31: 158-184.
- Wright, Sue Ellen, and Gerhard Budin (ed.). 1997. *Handbook of Terminology Management: Volume 1, Basic Aspects of Terminology Management*. Amsterdam/Philadelphia: John Benjamins Publishing Company.