전문용어의 정의문 분석

이해윤(한국과학기술원)

1. 들어가는 말

인터넷 기술의 발달로 우리는 매일 다양한 소스로부터 제공되는 수많은 정보를 접하게 된다. 이러한 정보의 홍수 속에서 인간은 필요한 정보를 재구성 내지는 가공하여 지식을 구축하게 되고, 이를 바탕으로 사회 속에서 다른 구성원들과 의사소통을 하게 된다. 이러한 정보 중에서 특히 언어적 정보는 현대 사회에서 그 중요성이 강조되고 있다.

언어적 정보는 생성-발전-소멸의 발생학적 단계를 밟는 언어기호를 기반 함으로 나름의 역동성을 보여주고 있다. 이러한 역동성을 상대적으로 강하게 느낄 수 있는 분야는 전문용어학 terminology이다. 왜냐하면 이 분야의 정보들은 사회가 발전함에 따라 언어 기호인 용어들이 새로이 생성되거나 유통되기 때문이다. 이러한 역동성으로 인해 기존의 전문용어 사전들이나 용어뱅크 term bank들은 항시 갱신될 필요가 있다.

전문용어 사전 혹은 용어뱅크는 다양한 정보들을 포함하고 있다.[1] 이러한 정보 중에서 핵심이 되는 부분은 전문용어의 정의부분이라고 할 수 있다. 그러나 정의에 대한 견해들은 사람들마다 달라, 각기 다른 내용의 혹은 다른 형태의 정의를 제시하고 있다. 따라서 다양한 코퍼스로부터 정의문 등을 추출하여 용어뱅크를 구축하고자 하는 작업에 있어서는 전문용어의 정의에 대한 명확한 정의가 먼저 이루어져야 일관성 있는 자료를 구축할 수 있다.

본 논문은 어휘 의미론적 관점 하에서 전문용어의 정의문을 분석하고자 한

1) ISO 10241에서 규정하고 있는 전문용어집의 정보 중 필수항목은 다음과 같다: the entry number, the preferred term representing the concept, the definition of the concept. 이외에도 다음의 부가적 정보들을 규정하고 있다: pronunciation; abbreviated form; full form, when the preferred term is an abbreviation; symbol; grammar; subject field; references to sources; non-preferred term(s); other representation(s) of the concept; references to related and other entries; example(s) of term usage; note(s); equivalent terms in other languages.

다. 이러한 연구목적 하에 본 논문은 다음과 같이 구성된다. 다음 2절에서는 정의에 대한 견해들을 사전편찬학적 관점과 전문용어학적 관점에서 살펴보기로 한다. 특히 후자와 관련하여 ISO 표준에서 제시하고 있는 지침을 자세히 논하기로 한다. 3절에서는 기존 전문용어 사전들에 나타나는 정의 부분들을 분석하여 보고, 다음 4절에서는 단어의 의미에 대한 보다 구체적인 분석을 제시하고 있는 생성 어휘론의 특질구조 qualia structure를 이용한 정의문 분석을 제안한다. 그리고 5절에서는 실제 신문기사로부터 특질구조의 부분 정보들이 분포되어 있는 모습을 관찰함으로써, 정의문 분석과 생성에 있어서 특질구조의 이용가능성을 제시하고자 한다.

2. 정의의 정의

'de-fin-e(정의하다)'의 어원에서 암시하듯이, 정의란 어떤 대상의 경계를 설정하는 일이다. 이러한 원칙적 입장은 동일하다 할지라도 실제 사전들에 나타나는 정의는 그 내용과 형태면에서 다양한 모습을 드러낸다. 다음은 표제어 '위궤양'에 대하여 일반사전, 백과사전, 전문용어사전에서 제시하고 있는 정의들이다.

위궤양:
위의 점막이 헐어서 심하면 구멍이 나 쓰리고 아프며 토하는 증세의 병. (연세한국어사전)

위궤양:
위의 점막이 헐어서 점막 밑에 있는 조직이 파괴되는 병. 위궤양은 자극이 심한 음식을 지나치게 먹거나 정신적인 피로가 원인이 된다. 증세로는 식후에 위가 아프거나 대변에 피가 섞이기도 한다. 30~40세의 남자에게 많이 생긴다. 자극이 심한 음식물과 정신적인 피로를 피하고 감정이 상하지 않도록 주의한다. (계몽사 백과사전)

위궤양:
위점막의 결손으로 생긴 양성 함몰(요)성 병변으로 통상 위액의 소화작용이 항진하고 있으므로 소화성 궤양(peptic ulcer)라고도 한다. X선상에서 점막주름의 집중과 위의 변형을 볼 수 있다. 급성기 궤양에서는 부종에 의한 주제를 볼 수 있다. (KORTERM[1])

이와 같이 각기 다른 내용이나 형식으로 정의를 제시하는 것은 해당 사전의 출판목적이나 독자 등 여러 요소들 작용에 의한 것으로 볼 수 있다. 본 절에서는 사전편찬학에서의 정의 개념과 비교하여 전문용어학에서의 정의개념을 살펴보기로 한다.

2.1. 사전 편찬학에서의 정의

언어를 사용한 의사소통 측면에서 정의의 개념은 언어기호와 대상간의 대응관계에서 규정된다. 언어기호를 사용범위에 따라 일반 단어 word와 용어 term로 구분할 경우, 단어는 대상에 대하여 언어기호를 부여하는 어의론적 semasiological 관점을, 그리고 용어는 언어기호에 대하여 대상을 부여하는 명칭론적 onomasiological 관점을 취한다 (Wright & Budin 1997). 그러므로 '정의'에 대한 정의는 단어를 연구대상으로 하는 사전편찬학 lexicography과 용어를 연구대상으로 하는 전문용어학이 서로 다르다고 볼 수 있다.

사전편찬학에서의 방법론이 적용된 결과물로서는 일반사전과 백과사전을 들 수 있다. 일반사전에서 정의부분의 주요 목적은 단어의 다양한 의미들과 용법들을 가능한 한 자세히 설명하여 다른 단어들과 구분시키고자 하는 데 있다. 따라서 동음이의어 homonym나 다의적 polysemous 의미정보를 알아낼 수 있고, 사용빈도수가 낮은 단어들에 대해서도 그 의미와 용법들이 자세히 설명되어 있다. 그리고 정의에 사용되는 메타언어는 그 의미가 일반적으로 파악되는 단어들이다. 이러한 목적 하에 사전편찬학에서는 다양한 형식들이 적용되어 정의를 기술하고 있다. 다음의 예들을 보자(Sager, 1990: 42f):

(1) (a) pneumonia = an inflammation of lung tissue

(b) daisy = bellis perennis
(c) whiteness = the state of being white
(d) metatarsalgia = a painful neuralgic condition of the foot, felt in the ball of the foot.
(e) dial = a clock or watch has a dial divided into segments for hours.
(f) dog = dogs are spanniels, poodles, ..

위의 예들에서 보면, 아리스토텔레스 이래로 가장 광범위하게 사용되어온 방식인 '상위어 + 차이점'으로 기술되거나 (a), 동의어 (b), 다시 쓰기 (c), 종합 (d), 함의 (e), 예시 (f) 등의 형식을 사용할 수 있다. 경우에 따라서는 이러한 형식들이 상호 결합되어 사용되기도 한다. 이러한 정의형식은 가능한 한 의미 명시를 위해서 정의대상에 따라 적절하게 선택된다.

이와는 달리 백과사전은 일반적으로 비전문가에게 해당 표제어에 대한 어느 정도의 이해를 제공하는 것을 목적으로 한다. 따라서 백과사전에 나타나는 정의부분에는 용어와 직간접적으로 관련하는 설명들이 제시되어 있어서 정의라기보다는 일종의 기술에 해당한다고 볼 수 있다.[2] 이러한 기술은 일반사전과는 달리 특정한 형식을 따르지 않으며, 일반적으로 여러 문장으로 구성되어 있어서 장황한 느낌을 받는다.

2.2. 전문용어학에서의 정의

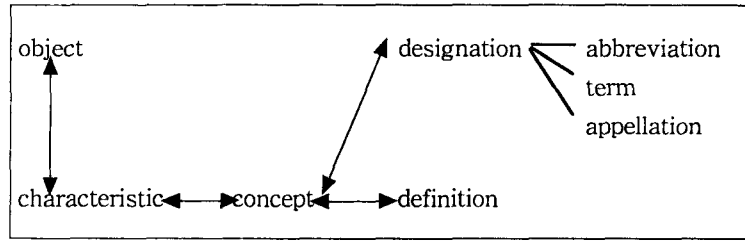
2.2.1. 대상-개념-용어

전문용어학에서 사용하는 '정의'의 개념은 그 출발점에서부터 앞서 살펴본 사전편찬학적 개념과는 다르다. 정의는 대상 object과 개념 concept 간의 관계 속에서 그 의미를 갖는다. ISO 1087에서 제시하는 다음 도식을 보자.

---

2) 정의 definition와 기술 description은 다른 개념들이다. 전자는 언어기호의 의미와 관련된 부분들을 설명하는 반면에, 기술은 이외에도 언어기호와 관련되지 않은 부분들도 포함한다.

(2)



대상은 인지 가능한 혹은 상상 가능한 세계 내에 존재하는 것들로서, '엔진'처럼 물질적인 형태일 수도 있고, '자기장'처럼 비물질일 수도 있다. 개념이란 동일한 속성을 갖는 대상들을 범주화하는 과정에서 그들이 갖는 속성들로 이루어진 사고단위라고 할 수 있다. 특정한 전문분야에서 개념들은 그들 간 일정한 관계를 맺고 있는 개념체계 concept system를 형성한다. 관계로서는 상하관계, 전체-부분 관계, 반의관계 등을 들 수 있다. 이러한 개념을 언어학적으로 표현하는 방식으로는 두 가지가 존재한다. 하나는 용어, 일반명사, 약어 등으로 지시할 수 있고, 다른 하나는 정의를 제시함으로써 표현할 수 있다.

용어나 정의를 이와 같이 개념의 다른 언어적 표현이라고 본다면, 용어나 정의는 개념과 등가관계에 있다고 볼 수 있다. 그러므로 용어나 정의는 개념을 충실하게 반영해야 한다. 즉, 개념체계 내에서 표현하고자 하는 개념이 갖는 위치가를 정확히 지시해 줌으로써, 개념에 대응되는 대상을 올바로 이해할 수 있게 해주어야 한다. 그러나 일반적으로 형태소들의 결합인 단일 명사 혹은 복합명사로 표현되는 용어가 이러한 기능을 충분히 반영하고 있다고 보기에는 무리가 있다. 그러므로 우리는 개념의 또 다른 표현체인 정의에서 이러한 기능을 기대하게 된다.

종합하면, 전문용어학에서 정의란 용어와 동일하게 개념의 언어적 표현체이며, 이상적인 정의란 특정한 개념체계 내에서 기술하고자 하는 개념이 다른 개념들과 차이를 드러내고 있는 점을 표현해줌으로써, 개념에 대응하는 대상을 독자로 하여금 정확히 지시할 수 있어야 한다(Rey, 1995: 42). 이러한 관계를 달리 표현하면 다음과 같다.

(3) (a) 대상 ≈ 개념
(b) 용어 ≈ 정의

즉, 하나의 대상에 하나의 개념이 대응하고, 이러한 개념들이 상호간 관계를 맺음으로써 개념체계를 형성하여 특정한 전문분야의 온톨로지 ontology를 구축한다고 볼 수 있다. 용어나 정의는 이러한 개념체계 내에 존재하는 특정한 개념을 표현하는 것으로서, 용어는 그 형태상 이러한 기능을 수행할 수 없으므로, 우리는 정의를 사용하여 이를 파악하는 것이다. 그러나 개념들이 상호간 관계를 맺고 있는 완전한 개념체계를 이루는 전문분야를 찾기는 쉽지 않다. 즉, 기술의 대상이 되는 개념체계가 완성된 전문분야가 현재로는 존재하지 않으므로, 기술수단인 정의가 현재로서는 완전하게 이루어질 수 없다. 그럼에도 불구하고 우리는 다음 절에서 ISO가 제시하는 표준화된 정의를 살펴봄으로써 이상적인 정의문을 규정해보고자 한다.

2.2.2. ISO 표준[3)]

ISO 표준문서 (ISO 704)에서는 앞 절에서 소개한 전문용어학적 관점에서 정의에 관한 표준을 제시하고 있다.[4)] 정의를 크게 내포적 정의와 외연적 정의로 구분하지만,[5)] 외연적 정의의 경우 내포적 정의가 불가능할 경우에 사용하도록 한정하고 있으므로 본 절에서는 내포적 정의에 제한하여 논의하기로 한다.

ISO 1087, ISO 704에서 제시하는 (내포적) 정의는 전통적인 아리스토텔레스식 방식에 따라 규정되고 있다.

(4) "the superordinate concept, either immediately above or at a higher

3) 전문용어와 관련한 ISO 기술위원회는 ISO/TC 37으로서 '전문용어와 언어자원'의 타이틀 하에 관련 연구 및 표준화 작업을 진행하고 있다.
4) 이와 관련된 ISO 표준으로서는 ISO 860, ISO 1087, ISO 10241 등이 있다.
5) 내포적' 정의와 외연적' 정의의 구분은 의미에 대한 Frege의 구분방식인 의의'와 지시'의 반영이라고 볼 수 있다. 외연적 정의란 개념의 외연에 해당하는 하위 개념들을 나열하여 제시하는 방식이다. 예) Threatened species: Critically endangered species, endangered species or vulnerable species.

level, followed by the characteristic(s) that distinguish the concept from other concepts." (ISO 704)

따라서 정의는 상위개념과, 다른 개념들과 구분되는 특성들로 구성된다고 본다. 그리고 용어가 표현하는 개념과 정의의 상위개념 간에 맺고 있는 관계에 따라 정의 유형들을 세분화하고 있다. 다음 예들을 보자 (ISO 704):

(5) (a) lead pencil: pencil whose graphite core is fixed in a wooden casing that is removed for usage by sharpening.
(b) lead cartridge: that part of a lead advance mechanism which stores and guides the lead refill as it advanced forward.
(c) pencil case: container designed to hold and carry pencils and other writing instruments

위 예에서 나타나는 용어개념 대 상위개념 관계는 다르게 나타난다. (a)에서 'lead pencil'과 'pencil'은 상하위어 관계를, (b)에서 'lead cartridge'와 'a lead advance mechanism'은 전체-부분 관계를, 그리고 (c)에서 'pencil case'와 'container'는 어떤 연상적 관계를 표현하고 있다. 따라서 (4)에서 제시된 정의의 표준에서 상위개념에 관한 한, 개념체계가 완전하게 형성될 경우 우리는 체계적으로 상위개념을 정의 내에 표현할 수 있다. 그러나 정의의 다른 한 구성요소인 특성에 관한 한, 단지 대상 혹은 대상들 집합의 추상화된 속성 (ISO 704: 4)으로 규정되고 있을 뿐 세부적인 설명을 찾아볼 수 없다. 따라서 (5)의 예들에서 우리는 상위개념들을 제외한 나머지 부분들이 용어개념의 특성에 해당된다고 추론할 뿐이고, 이 특성부분에 어떤 내용들이 포함되어야 되고, 어떠한 형식으로 기술되어야 할지는 열려있다고 볼 수 있다.

3. 기존 전문용어 사전에서의 정의문

기존에 출판된 전문용어 사전들에 나타난 정의들을 살펴보면, 편찬자에 따라서 제시된 형태나 내용면에서 차이가 나지만, 전문용어학적 편찬원리가 어느

정도는 지켜지고 있음을 알 수 있다. 형식적인 측면에서 보면, 전문용어 사전에 나타나는 정의는 기술대상에 따라 약 1-3 문장으로 구성되는 차이를 보인다. 일반적으로 첫 번째 문장/구는 정의의 핵심부분을 담고 있으며, 나머지 문장들은 기타정보를 표현하고 있다고 볼 수 있다. 따라서 우리는 첫째 문장을 중심으로 의미론적 관점에서 정의문을 살펴보도록 한다.[6)]

기존 전문용어 사전에 나타나는 정의문은 대부분 상위어 정보를 포함하고 있다. 상위어 정보이외의 다른 정보들은 관계절, 2격 명사구 혹은 전치사구 등으로 표현된다.

(6) (a) Adventivwurzeln - A. sind "zusätzliche" Wurzeln, die aus Gewebe des Sprosses bzw. des Blattes entstehen.
(b) Abundanz - Unter A. versteht man (in der Ökologie) die Zahl der Individuen einer Art pro Flächeneinheit.
(c) Actin - A. ist ein globuläres Protein mit einer Länge von 375 Aminosäuren.

상위어로서 나타나는 명사들은 특정 명사에 한정되지는 않는다. 일반적 현상을 가리키는 보통명사들이 쓰이는 경우도 있다. 예를 들어, *Umstand, Stoff, Gerät* 등이 나타난다.

(7) (a) Allensche Regel - Unter der Allenschen Regel versteht man den Umstand, dass die Größe der Körperanhänge (Schwanz, Ohren, Beine, Arme etc.) bei gleichwarmen Tieren in kälteren Regionen abnimmt.
(b) Antioxidantien - A. sind Stoffe, die eine Oxidation von im Körper vorhandenen Molekülen verhindern.
(c) ADW / A/D Wandler - Gerät, das analoge Signale in digitale umwandelt, mit denen der Computer dann arbeiten kann.

---

6) 본 절에서 제시되는 자료는 다음 인터넷 주소에서 제공하고 있는 컴퓨터 용어사전과 생물학 용어사전에서 발췌한 항목들이다.
(컴퓨터 용어사전 http://www.myglossar.de/index.htm,
생물학 용어사전 http://www.biologie-lexikon.de)

다음은 상위어 정보 이외의 정보 유형을 살펴보기로 하자. 빈번히 접하는 정보유형으로서는 표제어의 기능이나 작용을 표현하는 문장이나 구를 들 수 있다. 이러한 정보는 기능이나 작용을 의미하는 구체적인 단어, 혹은 사역의미의 동사들과 같은 어휘로 직접적으로 표현되거나, 통사구조를 통한 간접적인 방법을 사용하기도 한다.

(8) (a) ABA - Abkürzung für Abscisinsure; ABA gehört zur Gruppe der Phytohormone (Pflanzenhormone); ABA besitzt als Phytohormon eine inhibierende (= hemmende) Wirkung. ABA ist im Pflanzenreich quasi allgegenwärtig vorhanden.

(b) Administrator - Bezeichnung für den Systemverwalter in einem Netzwerk. Er hat uneingeschränkte Zugriffsrechte und ist für die Verwaltung und Betreuung des Netzwerks zuständig.

(9) (a) adäquater Reiz - Unter einem a. R. versteht man den Reiz, der mit der geringsten Energiemenge eine Erregung auslöst.

(b) Agent - Programm, das im Auftrag des Nutzer Anfragen im Netz durchführt und die Antworten bereitstellt.

(10) (a) ADI - ADI ist eine Schnittstellenspezifikation zur Entwicklung von Gerätetreibern für Peripheriegeräte, damit diese mit AutoCAD oder anderen Autodesk-Produkten arbeiten können.

(b) Ames-Test - Mithilfe des Ames-Tests lässt sich nachweisen, ob ein bestimmter chemischer Stoff Mutationen auslöst.

위 예 (8)에서는 어휘들(*Wirkung, zuständig*)을 사용하여 기능의 의미를 직접적으로 표현하고 있다. 이외에도 예 (9)에서 보듯이, 어떤 행위를 유발하는 동사들을 사용함으로써 표제어의 기능을 표현하고 있다(*auslösen, bereitstellen*). 다른 한편, 예문 (10)에서 우리는 통사적인 구조를 통해서 표제어의 기능 내지는 목적을 표현하고 있음을 확인할 수 있다.

이외에도 우리는 표제어의 성격에 따라 드물지 않게 표제어의 구성요소에

관한 정보를 정의문에서 찾아볼 수 있다. 이 경우, 표제어의 부분이나 부품을 암시하는 특정한 어휘들이 사용된다.

(11) (a) Adenosintriphosphat (ATP) - A. ist ein Molekül, das aus Adenin (ein Zuckermolekül) und 3 Phosphatgruppen besteht.
(b) ASP - Abkürzung für "Application Service Provisioning" [...] Das Application Service Provisioning beinhaltet demnach, zusätzlich zum Basis-Dienst, Informationen über Applikations-Zugriffsregeln inklusive Identifizierung und Authentifizierung.

끝으로 다음 예들에서 보는 바와 같이, 표제어의 출현, 생성에 관한 정보들도 정의문에서 찾아볼 수 있다.

(12) (a) alkoholische Gärung - Unter a. G. versteht man die Bildung von Ethanol und Kohlenstoffdioxid aus Glucose unter Abwesenheit von Sauerstoff (also unter anaeroben Bedingungen).
(b) Assimilate - Unter dem Begriff A. versteht man bei Pflanzen die im Zuge der Photosynthese entstandenen körpereigenen Produkte

이상에서 우리는 전문용어 사전에 제시된 정의문들을 그 의미적 유형에 따라 관찰하여 보았다. 다음 절에서는 생성 어휘론에서 제시하는 어휘부를 이용한 정의문 분석의 가능성을 살펴보기로 한다.

4. 생성 어휘론의 특질구조와 전문용어 정의문

자연언어의 의미기술은 오래 전부터 많은 의미론자들이 연구해온 분야이다. 생성 의미론자들의 자질분석 방법, Jackendoff(1990)의 개념구조, Wierbizbica (1996)의 NSM 등이 있지만, Pustejovsky(1995)의 생성어휘론(Generative Lexicon)은 다른 이론들에 비해 의미기술에 있어서 진일보한 이론으로 평가받고 있다.

Pustejovsky (1995)는 자연언어의 다의어에 주목하여, 단어의 의미는 언어적 환경에 따라 결정될 수 있고, 의미적 수용가능성 정도가 상이하게 나타날 수 있다는 관찰 하에 이를 반영한 사전의 구조를 제시한다. 그가 제시하는 사전의 구성은 단순히 평면적인 어휘항목의 나열이 아니라 공유하는 의미를 가진 어휘항목 간에는 보유한 정보의 양에 따라 위계적으로 되어있다. 각 어휘항목은 다음과 같은 세 가지의 각기 다른 유형의 정보를 담고 있는 것으로 본다.

(13) (a) 논항구조 (Argument Structure): 해당 어휘항목이 취하는 논항의 수, 의미유형, 논항들의 통사적 실현정보.
(b) 사건구조 (Event Structure): 해당 어휘항목이 묘사하는 사건유형에 대한 정보.
(c) 특질구조 (Qualia Structure): 해당 어휘항목의 의미기술에 필수적인 정보.

본 고의 주제와 관련하여, 위 세 구조들 중 특질구조에 대하여 자세히 살펴보도록 하자. 특질구조는 다음의 세부 정보들을 포함할 수 있는 것으로 보고 있다.

(14) (a) 형상역 (FORMAL): 관련 영역에서 해당 어휘항목을 다른 항목들과 구분시켜 주는 범주정보.
(b) 구성역 (CONSTITUTIVE): 해당 어휘항목과 그 어휘항목을 구성하고 있는 부분들간의 관계에 대한 정보.
(c) 기능역 (TELIC): 해당 어휘항목의 목적이나 기능에 대한 정보.
(d) 작인역 (AGENTIVE): 해당 어휘항목의 생성이나 유래에 관련한 정 보.

구체적으로 위 정보들이 어떻게 표현되는 지 'book'에 대한 다음의 사전기술을 보자.

(15)

$$\begin{bmatrix} book \\ \text{AGRSTR} & \begin{bmatrix} \text{ARG1} = & y.information \\ \text{ARG2} = & x.phys-obj \end{bmatrix} \\ \text{QUALIA} & \begin{bmatrix} information.phys-obj \\ \text{FORM} = & hold(x,y) \\ \text{TELIC} = & read(e,w,x.y) \\ \text{AGENT} = & write(e',v,x.y) \end{bmatrix} \end{bmatrix}$$

(16)

ISO 표준	특질구조
상위개념	형상역
특징	구성역
	기능역
	작인역

위 기술에서 특질구조를 보면, 먼저 형상정보로서 phys_obj 유형의 x가 information 유형의 y를 담고 있다는 내용이 표현되어 있다. 기능정보로서는 임의의 사건 e에서 w가 객체 x.y를 읽으며, 작인정보로서는 다른 사건 e′에서 v가 개체 x.y를 쓴다는 내용을 담고 있다.

이와 같이 의미에 관한 정보를 네 개의 유형으로 구분하여 기술함으로써 다의적 의미분석 등에 기존의 여타 어휘의미론보다도 진일보한 면을 보여주고 있다. 이러한 특질구조의 세부 정보분류는 전문용어의 정의문 분석에 이용할 수 있다. 앞서 살펴본 바와 같이 ISO 표준에서 제시하고 있는 정의에 관한 규정에서는 상위개념을 제외한 특징 부분의 규정이 불명확함을 보았다. 그러나 앞 절의 전문용어 사전에 제시된 정의문들을 분석해본 결과, 상위개념 이외의 정보들은 특질구조에서 제시하고 있는 구성역, 기능역, 작인역 등을 포괄하고 있음을 보았다. 따라서 우리는 다음과 같은 대응관계를 가정한다면, 특질구조에 의하여 정의문이 포함하고 있는 정보들을 보다 상세히 분석할 수 있다.

그러나 이예식(1999)에서 비판하고 있듯이, Pustejovsky(1995)에서는 특질구조의 정보기술에 대한 자세한 언급이 전혀 없다. 즉, 각각의 정보에 대하여 구체적으로 어떤 경우에 명시해야 하는지 혹은 어떻게 기술해야 하는지에 관한

언급이 없다. 이러한 비판점을 보강하여 실제 사전구축에 이용한 연구로서는 SIMPLE 프로젝트를[7] 들 수 있다(Lenci et al. 2000). 이 프로젝트는 생성 어휘론의 특질구조를 확장하여, 각 요소에 들어갈 유형들을 세분화하고 있다. 예를 들어 기능정보의 경우, 다음과 같이 그 하위 유형들을 가정함으로써 각 요소에 해당하는 정보부분을 좀 더 자세히 규정하고 있다.

(17)

하위 유형	예
Purpose	send: ... receive...
Object of the activity	book: ... read ...
Is the activity of	doctor: ... heal ...
Is the ability of	painter: ... paint ...
Is the habit of	smoker: ... smoke ...
Used for	crane: ... lift ...
Used by	lancet: ... surgeon ...
Used against	chemotherapy: ... cancer ...
Used as	wood: ...material ...

SIMPLE 프로젝트에서 목표로 하는 사전은 전문용어 사전이 아니라 일반 사전이므로 위와 같이 특질구조의 하위 정보마다 세부적인 분류는 광범위한 목록을 포괄할 수 있다. 그러나 이러한 세부적 분류가 과연 전문용어 사전에도 필요한지 검토가 필요하다. 왜냐하면, 전문영역의 경우, 각 하위정보의 구성은 특정한 몇몇 유형에 한정하여 나타난다고 할 수 있기 때문이다. 따라서 우리는 어휘적 의미 혹은 통사적 구조에 의거하여 특질구조의 개개 정보들을 식별하는 방법을 채택하도록 한다.

7) http://www.ub.es/gilcub/SIMPLE/simple.html

5. 코퍼스상에서의 전문용어 정의문

5.1. 정의문 구성

우리는 앞 절에서 정의문을 특질구조의 정보를 이용하여 분석 가능함을 보았다. 전문용어는 일반 단어와 마찬가지로 생성-발전-소멸의 단계를 밟는다. 생성시 전문용어는 그 특성상 정의문을 수반한다고 볼 수 있지만, 전문용어 사전에 등록되어 공식적으로 사용되기까지는 여러 언어자원들에 퍼져 나타난다. 특히 신문기사의 경우 새로운 전문용어와 정의문이 제시되어 나타날 가능성이 높다. 따라서 본 절에서는 신문기사에 나타나는 전문용어 정의문의 유형들을 특질구조의 정보유형에 따라 관찰함으로써 코퍼스에서의 전문용어 정의문 추출 가능성을 살펴보고자 한다.

신문코퍼스는 그 특성상 하나의 사건이나 기사를 중심으로 내용이 구성된다. 보고되는 사건이나 기사가 전문용어에 관계할 경우, 그 용어의 정의를 명시적으로 혹은 암시적으로 표현함으로써 독자의 이해를 도와준다. 특히 용어의 정의문은 일반적으로 기사의 전반부에 나타나 기사의 전체내용 파악에 선지식을 제공한다. 또한 장문의 기사인 경우, 기사의 중간중간에 관련 정의문을 부연설명하거나 반복설명함으로써 기술대상이 되는 전문용어의 이해를 손쉽게 해준다.

정의문은 독립된 문장이나 내포문 형태로 제시될 수 있다. 독립된 문장으로 정의문이 제시되는 경우는 이미 유포된 혹은 발전단계에 있는 전문용어 혹은 앞서 설명된 전문용어에 해당한다고 볼 수 있다. 반면에 새로이 생성된 전문용어의 경우, 아직 그 정의가 확정적이거나 유포되지 않아 독자에게 생소하므로 전문가나 관련자의 보고형식 혹은 인용형식을 취하여 그 정의문을 제시한다고 볼 수 있다.[8)]

다음은 정의문의 내용적 구성을 살펴보자. 텍스트 언어학적 견해에 따르면, 한 문장은 기술대상인 '주제 thema'와 그 기술내용인 '평언 rhema'로 구성되어

8) 정의문은 사태의 선언적 성격을 지니므로 미래, 예정, 진행 등의 시제가 들어있는 문장은 제외된다.

있다. 정의문의 경우, 그 제시형식이 독립적 문장 혹은 내포문이든, 주제는 전문용어에 그리고 평언은 정의부에 해당된다고 말할 수 있다. 따라서 코퍼스상에 나타나는 정의문은 전문용어를 포함하고 있는 문장과 관련된다고 할 수 있다. 그러나 모든 정의문이 전문용어를 포함하고 있는 것은 아니다. 주제에 해당하는 어휘가 전문용어와 공지시 관계를 이룰 경우, 이 어휘를 주제로 한 정의문 구성도 가능한 방법이다.

5.2. 정의문 분석

전문용어 사전에 제시된 정의문과는 달리, 신문기사에서[9] 나타나는 정의문은 여러 정보들이 산재해 나타나기도 한다. 따라서 다음 예에서 보는 바와 같이 하나의 정보만이 표현할 수 있다.[10]

(18) (a) Für die Suche nach neuen Wirkstoffen und die Entwicklung von Medikamenten ist **die Gentechnik**, die aus der von James Watson und Francis Crick angestoßenen molekularbiologischen Wissenschaft hervorgegangen ist, zur Schlüsseltechnologie geworden.

(b) **Die Genomforschung**, die versucht, Zusammenhänge zwischen Mustern des Genoms und Veränderungen einzelner Gene und Krankheiten zu erhellen, ist [...]

위 예에서 정의문 제시로 볼 수 있는 관계절을 보면, 사용된 어휘의 의미를 토대로 하여 각기 작인역, 기능역을 표현하고 있음을 알 수 있다. 다른 한편, 전문용어 사전에서 제시되는 정의문과 마찬가지로 상위어 정보를 기본으로 하여 기타 다른 정보들이 첨가된 형태의 구성을 찾아 볼 수도 있다. 다음은 이에 해당하는 예들이다.

---

9) Frankfurter Rundschau의 Wissen 분야에서 발췌한 기사들이다.

10) 아래 예에서, 굵은 글씨의 단어/구는 전문용어를, 밑줄 친 부분은 관련된 설명을 나타내고 있다.

(19) (a) Mit **molekularen Sonden** - das sind markierte Strangstücke spezifischer Sequenz, die sich an die komplementäre Sequenz anlagern - können Gene lokalisiert und genetische Veränderungen mikroskopisch oder chemisch erkannt werden.

(b) **Ballaststoffe** sind unverdauliche Nahrungsbestandteile, die vor allem in Obst, Gemüse und Vollkornprodukten enthalten sind.

위 예들에서 보면, 형상역과 기능역 (a), 형상역과 작인역 (b) 등이 결합되어 나타남을 볼 수 있다. 이외에도 다음 예에서 보는 바와 같이 여러 정보들이 곳곳에 산재되어 나타나는 경우도 볼 수 있다.

(20) Eine solche Kombination aus MP3-Player und Speicherstick bietet beispielsweise der Soundkartenspezialist Creative Labs aus Singapur mit dem **MuVo** an: Die bei der CeBIT vorgestellte 128-MB-Version des feuerzeuggroßen Gerätes kostet rund 180 Euro. Sie besteht aus zwei Teilen - dem Speicher mit der USB-Schnittstelle und einer Musikeinheit mit Batterie, Schaltern und Kopfhöreranschlüssen. Auf Bedienkomfort oder eine Megabassfunktion müssen Musikfans dabei allerdings verzichten.

위의 기사에서 보면, 첫 번째 문장은 전문용어 'MuVo'에 대한 작인정보를 주어형태로, 그리고 형상정보를 목적어 형태로 제시하고 있으며, 세 번째 문장은 이 전문용어에 대한 구성정보를 표현하고 있다. 따라서 전문용어 'MuVo'에 대하여 위 신문기사로부터 정의문 구성요소들을 다음과 같이 정리/추출해 낼 수 있다.[11]

(21) (a) 형상역: eine solche Kombination aus MP3-Player und Speicherstick

(b) 구성역: ... besteht aus zwei Teilen - dem Speicher mit der

11) 정의문 추출에 관한 연구로서는 신효식 등 (2002), 山田一郎 et al. (2001), Muresan, S. and Klavans, J. (2002) 등을 들 수 있다.

USB-Schnittstelle und einer Musikeinheit mit Batterie, Schaltern und Kopfhöreranschlüssen.

(c) 작인역: ... bietet der Soundkartenspezialist Creative Labs aus Singapur an.

6. 맺는 말

본 논문에서는 어휘 의미론적 관점에서 전문용어의 정의문에 대한 구성을 살펴봄으로써 그간 불명확히 사용되어온 '정의'의 개념을 명확히 정리하고자 시도하였다. 먼저 정의에 관한 사전편찬학적 견해에서 출발하여 ISO 표준의 규정들을 논의하였다. 이러한 논의에서 정의에 대한 기존 정의들의 불명확한 점들이 드러났고, 이에 대한 대안으로서 Pustejovsky(1995)의 특질구조 이용가능성을 살펴보았다. 특히 기존 전문용어 사전들에 나타난 정의문들을 특질구조에 의해 재분석해봄으로써, 우리는 정의문 분석과 생성작업에 있어서 특질구조의 정보들이 유용함을 보았다. 끝으로 신문기사에 나타난 전문용어의 정의문을 특질구조 정보에 의거하여 분석함으로써 용어집 작성이나 용어뱅크의 구축시 활용가능성을 보여주었다.

참고문헌

신효식, 김재호, 이해윤, 최기선 (2002), 텍스트로부터 용어 정의문의 자동 추출 방법. In: *한글 및 한국어 정보처리 자료집 2002.*

이예식 (1999), 어휘 의미론과 다의어의 분석. In: *강범모 등, 형식의미론과 한국어 기술*, 552-592. 한신문화사.

山田一郎, 田正啓, 金淵培 (2001), 뉴스원고를 이용한 용어집 작성 검토. NHK 방송기술연구소, ms.

ISO 704, *Terminology work - Principles and methods.* Geneva: International Organization for Standardization.

ISO 860, *Terminology work - Harmonization of concepts and terms.* Geneva: International Organization for Standardization.

ISO 1087, *Terminology - Vocabulary.* Geneva: International Organization for Standardization.

ISO 10241, *International terminology standards - Preparation and layout.* Geneva: International Organization for Standardization.

Jackendoff, Ray (1990), *Semantic Structures.* Cambridge: MIT Press.

Lenci, Alessandro et al. (2000), Linguistic Specifications: SIMPLE Work Package 2.

Muresan, S. and Klavans, J. (2002), A Method for Automatically Building and Evaluation Dictionary Resources. In: *LREC 2002,* 231-234.

Pearson, Jennifer (1998), *Terms in Context.* Amsterdam: John Benjamins Pub.

Pustejovsky, J. (1995), *The Generative Lexicon.* Cambridge: MIT Press.

Rey, Alain (1995), *Essays on Terminology.* Amsterdam: John Benjamins Pub.

Sager, Juan (1990), *A Practical Course in Terminology Processing.* Amsterdam: John Benjamins Pub.

Wierzbicka, Anna (1996), *Semantics, Primes and Universals.* Oxford: Oxford University Press.

Wright, Sue Ellen & Gerhard Budin (1997), *Handbook of Terminology Management: Vol. 1.* Amsterdam: John Benjamins Pub.

Zusammenfassung

An analysis of terminological definitions

Lee, Hae-Yun(KAIST)

In this paper, we examined various definitions of terminological definition for the extraction of terminological information from corpora. After we reviewed researches at the lexicography and at the terminology, we introduced the qualia structure of Generative Lexicon (Pustejovsky 1995) for the purpose of analyzing terminological definitions. By means of the qualia structure, we analyzed the definitions which are presented at the terminological dictionaries. As a result, we confirmed that the

terminological definitions can be discomposed into 4 subtypes of qualia structure. Based on this examination, we analyzed terminological definitions of articles at a newspaper and showed the usefulness of the qualia structure at the extraction of terminological definitions from the corpora.

[검색어] 용어, 정의문, 전문용어학, 특질구조
Term, Definition, Terminologie, qualia structure

이해윤

305-701
대전광역시 유성구 구성동 373-1 한국과학기술원 전산학과
haeyun@world.kaist.ac.kr