

# 참조용어(Reference Terminology) 모델 확장을 위한 한의학용어 정형화(Normalization) 연구

진병욱, 홍성천  
원광대학교 한의학전문대학원

---

## A study of Traditional Korean Medicine(TKM) term's Normalization for Enlarged Reference terminology model

Byoung-Uk Jeon, Seong-Cheon Hong  
Professional Graduate School of Oriental Medicine, Wonkwang University

The discipline of terminology is based on its own theoretical principles and consists primarily of the following aspects: analysing the concepts and concept structures used in a field or domain of activity, identifying the terms assigned to the concepts, in the case of bilingual or multilingual terminology, establishing correspondences between terms in the various languages, creating new terms, as required.

The word properties has syntax, morphology and orthography. The syntax is that how words are put together. The morphology is consist of inflection, derivation, and compounding. The orthography is spelling.

Otherwise, the terms of TKM(Traditional Korean Medicine) is two important element of visual character and phonetic notation. A visual character consist of spell, sort words, stop words, etc. For example, that is a case of sort words in which this '다한', '한다', '多汗', '汗多' as same. A phonetic notation consist of palatalization, initial law, etc. For example, that is a case of palatalization in which this '수족랭', '수족냉', '手足冷', '手足冷' as same.

Therefore, to enlarged reference terminology is a method by term's normalization. For such a reason, TKM's terms of normalization is necessary.

---

**key word** : *reference terminology, Traditional Korea Medicine terminology*

---

### I. 서론

용어는 인간의 의사소통의 도구로서 중요한

위치를 점하고 있다. 이미 형성된 용어라 할 지라도 그 개념은 변화할 수 있으며, 그 반대로 어떤 개념을 표현하기 위한 수단으로서, 용어는 새롭게 생성 될 수도 있다.

특히, 광범위한 의료정보학의 영역에서, 참조용어(Reference terminology) 모델은 데이터 교환 및 정보공유 측면에서 중요한 의의가 있으며, 다양한 의료정보시스템의 개발은 효율적인 데이터 교환 및 정보 공유에 필요한 표준화된 참조용어 모델을 요구한다.

SNOMED CT는 질병의 여러 가지 특성을 기반으로 용어를 결합하여 복잡한 개체나 필요한 개념을 만들 수 있는 특성을 가지고 있다. 때문에 여러 가지의 용어 조합방식에 의해 개념의 확장성을 제공할 수 있는 장점을 가지고 있는 참조용어 모델이다.

UMLS는 SNOMED CT를 포함한 148개의 의학 관련 분야의 소스를 기반으로 동일 개념을 표현하는 용어를 그룹핑(grouping) 하여 메타시소러스(Metathesaurus)를 구축하고 있으며 SNOMED CT 보다 광범위한 참조용어 모델이라고 할 수 있다.

즉, 참조용어 모델은 의료정보의 전문화 및 세분화에 따른 정보시스템의 발전과 더불어 전자의무기록(EMR), EHR, e-PHR 도입 등 효율적인 데이터 교환 및 정보공유에 활용하기 위한 목적으로 다양한 사용자 계층의 요구를 수용할 수 있어야 한다.

한편, 우리나라는 이원화된 의료체계의 현실과 외래어가 주를 이루는 의학용어의 한글화에 대한 문제점 등 현재의 의료 현실에 적합한 참조용어 모델이 요구되고 있다. 이에 의료정보시스템 개발에 있어, 임상에서 수용도가 높으며 비교적 표현력이 풍부하고 쉽게 적용할 수 있는 SNOMED CT가 참조용어 모델로서 채택되고 있다. 하지만, 한의학에 관련된 참조용어 모델이라고 할 수 있는 체계는 미흡한 실정이다.

참조용어 모델은 다양한 어휘 유형을 분석하여 동일한 개념을 중심으로 용어를 그룹화 하는 것은 동일 개념을 표현하는 확장된 용어를 제공하여 사용자 계층의 수용성을 증대시키고, 효율적인 데이터 교환 및 정보공유에 도움이 된다.

이러한 동일 개념의 문자열 확장의 방안으로는 용어 정형화(Normalization, 定型化) 과정이 포함된다. 정형화란 ‘일정한 형식이나 틀을 마련하는 것’으로서, ‘근거나 기준을 마련하기 위한 일련의 과정’이라는 점에서 표준화의 개념과 유사하다.

UMLS에서는 SPECIALIST Lexicon & Tools을 사용하여 용어 정형화에 대한 연구를 수행하고 있다. 이렇게 정형화된 용어는 다른 용어들과 연결시키고, 각각의 용어들을 어떤 개념으로 매핑(mapping) 하는데 사용된다.

즉, 용어 정형화는 문자열의 구조를 분석하고 일련의 규칙을 정하여 표기가 가능한 문자열을 생성하여 참조용어 모델을 확장하기 위한 하나의 방안이다.

본 연구는 참조용어 모델의 확장을 위한 방안으로서 한의학용어의 정형화 규칙을 마련하는데 그 목적이 있다.

## II. 연구범위 및 방법

한의학과 관련된 용어 표준화 연구로는 국제 표준용어집(WHO-IST)<sup>1)</sup>, 한의학표준용어집(KST)<sup>2)</sup>, 보건의료정보 표준화연구 등이 있지

1) IST : International Standard Terminology WHO 서태평양지구 주도하여 한국, 중국, 일본의 3개국 주도하여 만들어진 한의학용어집.

2) KST : Koran Standard Terminology. 대한한 의사협회 주관하여 한의학 분과학회별로 용어를 수집

만, SNOMED CT, UMLS에 비해 동일 개념을 표현하기 위한 용어의 빈도는 지극히 낮다. 즉, 기존의 연구결과는 참조용어 모델로서는 적합하지 않다고 생각된다. 그러나 이러한 연구결과는 사회적 타협이 이루어진 용어들로서 어떤 개념을 대표하는 용어이기 때문에 정형화 과정 등을 통해 보다 많은 유형의 용어 생성이 가능하며 참조용어의 확장이 가능하다.

특히, UMLS SPECIALIST Lexicon & Tools의 용어 정형화 방안을 분석하였으며, 용어의 속성을 3가지로 정의하고, 한의학용어의 정형화 방안을 제시하고자 한다.

- 문법구조(Syntax) : 용어의 구조적 배열에 의한 정형화 방안
- 언어형태(Morphology) : 언어의 변형, 언어의 어원, 복합어 등 용어의 형태적 측면에 의한 정형화 방안
- 표기법(Orthography or spelling) : 용어의 표기 또는 철자에 의한 정형화 방안

이에 본 연구에서는 의학용어의 정형화 과정을 사전분석한 후 동일 개념을 표현하는 용어들의 표기방법을 분석하고 일련의 규칙을 정하고, 한의학용어에 적용이 가능한 정형화 방안을 제시하고자 한다.

### III. 선행연구

의학용어는 일반영어, 그리스어, 라틴어에 기원을 두고 있으며, 그리스어 또는 라틴어로 이루어진 의학용어는 복합어의 경우에도 기원

---

하여 만들어진 한의학용어집.

이 같은 용어가 서로 결합하여 하나의 개념을 형성하고 있다.

따라서 의학용어는 동일 개념의 문자열 형태가 용어의 기원에 따라 다양한 표현 방안을 가지고 있다. 특히 그리스어나 라틴어를 일반 영어로 표기하는 사용자 계층에 있어서는 그 표현 방법이 다양하며, 특히 UMLS에서는 일반 영어로 표현되는 용어를 대상으로 사용자 계층의 수용성을 확보하기 위하여 용어 정형화 과정을 통해 다양한 문자열을 확보함으로써 참조용어의 확장이 가능하다.

첫째, 문법구조 측면에서의 용어의 정형화는 용어의 순서 배열에 따라서 참조용어의 확장에 대한 가능성을 제시한다.

예를 들면 '*hodgkin disease - disease hodgkin*'은 동일한 개념을 가지며 용어의 정형화 과정에 일정한 규칙으로 포함되어 있다.

둘째, 언어형태 측면에서의 용어의 정형화는 대소문자, 단·복수 등 용어의 고유한 어원을 그대로 유지하면서 용어의 변형에 따른 참조용어의 확장에 대한 가능성을 제시한다.

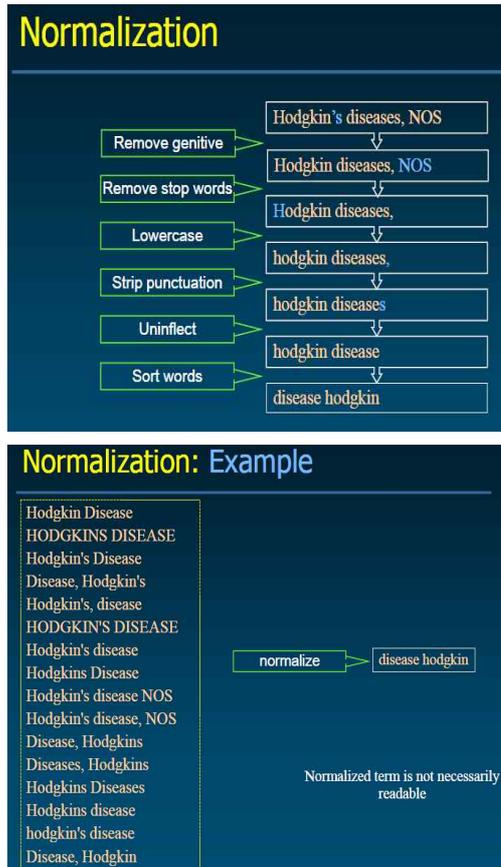
예를 들면 '*HODGKINS DISEASE - Hodgkin Disease - Hodgkins Diseases - Hodgkins Disease*' 등은 용어의 정형화 과정을 통해 확장될 수 있는 용어이며 참조용어 모델에 있어서 동일한 개념을 표현하는 용어이다.

셋째, 표기법 측면으로서 동일한 의미를 가지는 다양한 철자법에 따라서 생성이 가능한 용어를 대상으로 참조용어의 확장에 대한 가능성을 제시한다.

예를 들면 '*oe/e', 'ae/e', 'ise/ize', 'genitive*

mark(소유격 표현기호=), 'British-American'에 대한 다양성을 포함한 참조용어의 확장에 대한 가능성을 제시한다.

[그림 1]은 'disease hodgkin'에 대한 용어 정형화 과정 및 정형화 과정을 통한 용어의 결과에 대한 것이다.



[그림 1] UMLS의 Term Normalization

따라서, 의학용어에서 사용되는 용어 정형화 방안은 일반영어, 그리스어, 라틴어에 초점을 두고 있기 때문에 언어가 다른 한자문화권의 한의학용어에 그대로 적용할 수는 없다.

즉, 한의학용어의 정형화는 의학용어의 관점과 다른 측면에서 접근되어야 정형화 방안을 통해 참조용어의 확장가능성에 대한 접근방안

을 살펴보았다.

#### IV. 한의학용어 정형화 방안

의학용어와 달리 한의학용어의 정형화는 문법구조, 언어형태, 표기법에 대한 보다 전문적인 연구가 수행되어야 한다.

어떤 개념을 중심으로 용어를 조사하고 매핑시키는 것은 한의학 지식 전문가에 의해서 가능할 것이다.

그러나, 이와 같은 방법은 인력과 시간적 측면에서 비효율적이며, 가령 동일한 개념을 표현하는 용어라 할지라도 동양철학을 기반으로 하는 한의학용어는 전문지식인간의 이해관계에 있어서 사회적 타협이 이루어져야 한다.

즉, 어떤 개념에 새로운 용어를 추가하여 참조용어를 확장하는 방안이 아니라, 사회적 타협이 이루어진 용어를 대상으로 참조용어의 확장이 위하여 일정한 규칙을 세우고 일관된 형식을 갖추기 위한 정형화 방안을 제시하고자 한다.

##### 4.1 한의학용의 특수성

한의학용어는 한글과 한자로 구성된 것으로서, 그 표기법은 한자의 음을 한글로 표기하거나, 한자를 그대로 표기하여 사용하는 것이 일반적이다.

한자의 음을 한글로 표기하는 과정에서는 발음에 따라 한글을 음차하여 표기할 때 두음법칙, 구개음화 등 국어문법에 따른 표기법의 차이가 존재하며, 한자를 표기하는 과정에서는 한자의 형태를 변형하여 표기하는데, 변형되는 한자의 형태는 번자, 간자, 속자로 구분

되고 이는 사용자 계층에서 각자의 편의성에 따라 사용되고 있다.

즉, 한의학용어는 사용자 계층에서 각 상황에 따라 혼용하여 사용되고 있기 때문에 적합한 참조용어를 개발에는 많은 노력이 있어야 하며, 다양한 표기법에 의한 용어를 추가하여 참조용어 개발 및 확장에 고려해야 한다.

즉, 어떤 동일한 개념을 표현하는 용어의 표기법이 존재하며, 사용자 계층에서 동일 개념의 용어를 표기할 수 있는 부분을 어느 정도 예측이 가능하기 때문에 일련의 형태를 갖추어 다양한 표기법에 따른 용어의 추가가 가능해진다.

따라서 한의학 용어의 정형화 유형을 분석하여 정리하고, 예측이 가능한 용어의 다양한 표기법을 추가하여 참조용어를 확장시킬 수 있다.

## 4.2 한의학용어 정형화 방안

### ◆ 한글용어 표기 정형화

手足冷, 수족랭, 수족냉처럼 한자의 발음을 한글로 음차 할 경우 발생하는데 이처럼 한자어 하나가 병음을 가지고 있는 경우를 고려하여야 한다. 즉, 한자음이 2개 이상인 경우 한의학용어를 표기할 때 발생 가능한 용어의 추가가 필요하다.

대표적인 한자어로서 ‘수’-‘數’와 ‘삭’-‘數’은 시각적인 측면에서 한자어는 같지만 한글용어의 표현에 있어서는 시각적 차이가 존재한다. 또한 개념적 차이도 발생하는데 ‘수’로 발음할 경우에는 개수, 숫자와 관련이 있지만 ‘삭’으로 발음할 경우에는 빈도, 횟수의 개념이다. 예외적인 상황도 고려하여 용어 추가가 이루어져

야 한다.

### ◆ 한자용어 표기 정형화 방안

氣, 陰, 陽, 風, 手足癱瘓, 麻黃, 등

气, 陰, 陽, 風, 手足癱瘓, 麻黃, 등

위의 예시처럼 한자는 표기법에 따라 번자와 간자 그리고 속자로 구분되며, 사용자 계층의 표현에 대한 각자의 편의성에 의하여 그 변형이 가능하다.

따라서, 어떤 개념을 대표하는 대표어 외에 한자어의 형태변화에 따른 용어 추가가 필요하다.

### ◆ 한글과 한자 병행표기 정형화

한글과 한자의 병행 표기 및 한자어의 배열 순서에 상관없이 개념이 변하지 않는 경우에 대한 용어 추가가 필요하다.

한글과 한자의 병행 표기는 ‘손발이 냉함’등 자연어로서 나타나는데 명사화 보족절인 ‘口’, ‘기’를 사용하여 표현이 가능한 용어로서 어떤 개념을 표현하는 용어들은 더욱 풍부해질 수 있다.

또한, 한자어의 배열 순서에 관계없이 그 개념이 변하지 않는 경우를 고려하여야 하는데 ‘多汗’, ‘汗多’등은 대표적인 사례이다.

그러나, 한자어의 재배열시 한글음만을 가지고 표기를 할 경우에는 주의해야 한다. ‘열한’의 경우 ‘烈寒’, ‘熱汗’으로 표기가 가능하며, 각 용어는 개념에 차이가 있다. 따라서 이런 경우의 용어 정형화는 한글과 한자의 병행표기에 대한 용어 추가를 하여야 한다.

## V. 결론 및 제언

정형화란, 어떤 일련의 규칙과 틀을 만드는 것으로서 용어의 정형화에 대한 방안이 있다면, 참조용어의 확장에 도움이 될 것이다. 특히, 용어 사용의 오용 및 오자에 대해서도 정형화 과정을 통해 잘못 표기된 용어라 할지라도 그 용어의 기원과 생성에 대한 추론이 가능하기 때문에 정형화 과정을 통한 참조용어의 확장은 잘못 사용된 용어에 대한 교정이 용이할 수 있으며, 사용자가 원하는 용어의 확장성을 제공할 수 있기 때문에 정형화에 대한 방안의 규칙을 마련하는 것은 필요할 것이다.

이처럼 참조용어 모델 확장을 위한 용어의 정형화 방안은 다음의 사항이 고려되어야 한다.

첫째, 동일 범주 내에서 사용되는 개념간의 구조를 분석하는 것.

둘째, 어떤 개념을 나타낼 수 있는 고유한 용어를 식별하는 것.

셋째, 용어를 표기하는 다양한 어휘 유형을 분석하여 동일한 개념중심으로 용어를 그룹화하는 것.

마지막으로 새로운 용어 생성과 확장이 가능해야 하는 것으로 볼 수 있다.

특히, 한의학용어는 한글과 한자로 표현되는 용어의 형태뿐만 아니라, 한글과 한자의 혼용으로 표현하는 용어의 풍부한 표현력에 대한 문제를 정형화 범주에 포함시켜야 할 것으로 생각되며, 더 나아가서는 한의학의 영역화에 대한 표기 방안도 정형화 범주에 포함하여 일련의 규칙을 마련하고 다양한 사용자계층이 수용할 수 있는 확장된 참조용어의 모델이 필

요하다고 생각되며, 이를 위해서 정형화 방안을 제시함으로써 확장된 참조용어의 모델 개발이 이루어져야 한다고 생각한다.

또한, 용어 표준화는 어떤 개념을 대표하는 대표어를 규정하는 것 이외에 그 용어의 변형이 가능한 형태를 추가함으로써, 수용자 계층에 풍부한 표현력을 가지는 참조용어의 모델에 지속적인 연구가 필요하다고 생각된다.

## 참고문헌

1. Rachel Kleinsorge and Jan Willis. Unified Medical Language system Basics, National Library of Medicine. 2008
2. K. Donnelly. SNOMED CT : The advanced terminology and coding system for eHealth. Stud Health Technol Inform. 2006;121
3. K. Giannangelo and L. Berkowitz, SNOMED CT helps drive EHR success. J Ahima. 2005;76
4. 최진욱. UMLS(Unified Medical Language System)를 이용한 한글진단 증상용어 Terminology Server의 개발. 보건복지부. 2004
5. 박영진. 의료 영역에서 대용량 연속 음성 인식의 성능 향상. 서강대학교 대학원. 석사학위논문. 2002.
6. 국립국어원. 표준국어대사전. 2009
7. 서정수. 현대 국어문법론. 한양대학교 출판원. 서울. 1996
8. 한국전통의학연구소. 한의학사전. 정보사. 2002
9. 대한의사협회. 의학용어집. 아카데미아. 2001
10. <http://www.wikipedia.org>