

# 영어문장제공시스템에서 한국어문장의 색인방법 Indexing Methods of Korean Sentences in the English Sentences Offering System for English Composition

이 태 영, 전북대학교 문헌정보학과

Tae Young Lee, Dept. of Library and Information Science,  
Chonbuk National University

한국어문장과 의미가 같은 영어문장을 검색해 내기 위하여 한국어문장을 분석하고 색인언어를 고안하였다. 명사와 더불어 용언, 보조용언, 조사, 접속사 등이 색인어 및 기호로 선정되었다. 색인어 수를 줄이는 데 용언과 명사의 유사의미 단어들의 통제가 필요하였다.

## 1. 서론

인터넷을 통해서 수많은 정보를 쉽게 접하는 세상이 도래하였다. 한국의 대학 및 대학도서관에서의 인터넷 통신은 이제 당연한 서비스의 한 품목으로 자리잡았다. 아울러 전세계의 여러 이용자와 영어로 의사를 주고받는 일이 많아지고 있으며 앞으로 더욱더 이러한 현상은 증가하고 계속될 전망이다. 따라서 표현하고자 하는 영어문장이 잘 떠오르지 않거나 그러한 류의 표현양식에 익숙해 있지 아니 할 때 그 고민을 해결해 줄 도구가 있으면 매우 편리할 것이다. 다시 말해서 여러 가지 경우(추천문, 상업문, 개인적 서신, 알림말 등)의 영작문에 대한 전문 또는 일부 문장의 범례적 모델을 제공하는 시스템이 필요하다는 말이다. 1)

이러한 영문제공시스템은 현재 한영번역시스템이 나와 있지 않은 상태에서<sup>2)</sup> 매우 유용하게 쓰일 수 있으며 설사 한영번역시스템이 갖추어져 있다 해도 하나의 한국어 문장에 대해 여러 개의 영어 유사문장의 수수 가능성이 있으므로 보다 효율적이며 여러 방향의 검색 접근점을 제공하여 편리함을 도모할 수 있다.

이용자의 검색 접근점은 크게 세가지로 설정할 수 있다. 첫째, 하나의 명제로 접근하는 경우; 즉 어떤이를 초청하려고 초청문안을 작성

하기 위해 기존의 문안을 참고하려할 때, 둘째, 하나의 문장을 구체적으로 작성하려는 경우; 문장을 세부적으로 구체화해 놓고 이를 영역하려 할 때, 셋째, 하나의 문단을 구성하려는 경우; 한 명제에 얽힌 개념들을 부분적으로 정리하여 그 부분을 작성하려 할 때이다.

여기서 (1)명제적 접근은 초청에 대한 내용을 여러개의 문장으로 나타낸 완전한 것이며 내용의 주제에 대한 대변도 '초청장', '초청장--생일'과 같이 단순히 표현될 수 있다. (2)문장적 접근은 어떠한 문장을 완성하려고 하나 잘 이루어지지 않을 때 참조할 수 있는 방법으로 써 질문된 한국어문장을 분석 색인하여 영문 DB에서 이에 상응하는 영어문장을 검색해낸다. (3)문단적 접근은 작문을 하다보면 어느 부분에 가서(초청문의 시작-인사말 부분 등) 그곳을 어떻게 표현할 것인가를 고민할 때 이용된다.

본고에서는 (1)과 (3)은 차후의 연구기회로 돌리고 (2)의 문장적 접근을 다루기로 한다. 한국어문장과 영어문장이 서로 만나 의미가 같은 것이 한데 묶여 나옴으로써 상호 문장 간에 부여된 색인언어가 같으면 된다. "나는 간다", "I go"는 색인어가 "나, 가다"로 공통으로 되기 마련이며 "나, 가다"의 색인어로 검색시에 서로

1)영어의 제반 서식 및 문장의 범례를 보여주는 책자는 여러 가지가 있으나 보다 편리하게 다양한 경우를 검색해 볼 수 있는 컴퓨터화된 시스템이 필요하다.

2)참고로 영한번역시스템은 현재 "애플(IBM solution provider), Word Change(정소프트), Trannie(서울대 언어공학연구소), 세계로(주 성운시스템), 등이 있다.

만나 출력되어진다. 본 논문의 목적은 이렇게 용언이 포함되는 한국어문장의 색인형태를 색인언어 차원에서 구명 정리하려는 것이다.

## 2. 한국어와 영어의 문장 구성요소

품사(9종 기준)에서 한국어와 영어가 다른 것은 종류에서 한국어의 “조사, 수사”, 영어의 “전치사, 접속사”이고 쓰임새에서는 ‘형용사’이다.

이것들을 살펴 보면 한국어의 ‘수사’는 영어의 명사, 형용사 등에 그대로 포함되어 있고, 한국어의 ‘조사’와 영어의 ‘전치사’는 그 쓰임새가 비슷하다. 영어의 ‘접속사’는 한국어에서 용언이 어미변화를 일으켜 접속어구 구실을 하거나 명사 또는 조사에 속한 단어가 그 역할을 대신한다.(영어의 관계대명사, 관계부사도 접속 역할을 한다) 한국어 ‘형용사’와 영어 ‘형용사’는 뜻하는 바는 같으나 용법상에는 현격한 차이가 있다. 한국어의 ‘형용사’는 동사와 같이 어미변화를 일으키므로 한국어에서는 ‘동사’와 ‘형용사’를 묶어 용언으로 통칭한다.

한편 한국어나 영어 문장은 공히 기본 5형식이 있으므로 영어와 한국어 문장을 색인할 때에는 이 기본 5형식을 먼저 살펴 색인어 발체에 중심으로 삼아야 한다. 색인시 구체적으로 고려해야 할 한국어와 영어의 문장 구성요소들을 분석하여 보면 다음과 같다.(고딕체 단어는 변수이고, 샘플체는 상수이며, ‘-’는 어간이나 단어를 나타낸다.)

### (1) 조사와 전치사

①의; of, ②으로; to, ③으로부터; from, 등

### (2) 접속어구(관계대명사와 관계부사 포함)

(가)대등접속 : ①-고; and, ②-(으)나; but,

③-나; or, ④보다; more-than, 등

(나)종속접속 : ①-면; if, ②전예; before,

③때문예; because, ④-는 한; as long as, 등

(다)관계대명사 : ①-누, -는(사람, 인명); who,

②-누, -는(사물); which, ③-는 것은(을);

what, ③-(이)든(지) 간예; whatever, 등

(라)관계부사 : ①-누, -는(장소); where,

②-누, -는(방법); how, 등

### (3) 선어말어미

(가)존대 : -부니-, -습니-, -시-, -하여

주시-, -하여 주십니-, 하고 있습니-;

Would - please,

(나)과거 : -ㅁ-, -았-, -였-, -었-, -ㅅ-(왔),  
-ㅆ-(땀); -ed.

(다)현재 : -ㄴ-, -는-; -.

(라)미래 : -겠-; will -.

(마)과완(과거완료) : -ㅁ 었다, -하엿었다, -  
했었다, -되엿었다; had -ed

### (4) 보조용언

(가)미래 : -ㄹ 것이다, -려고 한다; will -.

(나)현진(현재진행) : -고 있다, -는 중이다,  
-가다; be -ing.

(다)현완(현재완료) : -(해) 버렸다, -적이  
있다, -고 있다; have -ed.

(라)부정1 : -지 않-; do not -, 등

(마)가능(불가능) : -ㄹ 줄 알다(모르다),  
-ㄹ 수 있다(없다), -ㄹ 수가 있다(없  
다); can -, 등

(바)추측 : -는 듯 하다, -는 것 같다, -ㄹ  
것이다; seem to -, would seem -, 등

(사)허가 : -(하여, 해)도 좋다, -어도 된  
다; may -, 등

(아)당위 : -야 한다, -지 않으면 안된다,  
-는 것이 당연하다; must -, 등

(자)습관 : -곤 하였다, -기도 하였다; used  
to, 등

(차)희망 : -고 싶다, -기를 원한다; like  
to -, 등

(카)권유 : -는 것이 좋다, -ㄹ 필요가 있  
다; had better -, need to -, 등

(타)가정 : -면 -되겠다, -ㅁ 을지도 모른다;  
if-were/would -, would have -, 등

(파)양태 : -게 되다, -게 하다, -게 시키  
다, -게 지다(곱게 지다), -해 지다  
(깨끗해 지다); let -, 등

### (5) 형용사와 형용사 상당어구

한국어에서 형용사는 ‘파랳다’에서 보는 바와 같이 동사처럼 어미가 있으며 어미변화를 한다. 한편 “이다, 있다, 없다”는 어떤 명사의 뒤에 붙어 용언구실을 한다. (예; 산있다, 산이다, 맛있다, 맛이다, 맛없다, 등) :

①파랳다; is blue, ②산이다; is mountin, 등

### (6) 부정사와 동명사

①부정2 : -기 위하여, -ㄴ 것, 등; - to  
-(eat to live, is to make)

②동명 : -는 것을, 등; -ing (like  
smoking)

### (7) 관용구

①아지 : 아무리 강조하여도 지나침이 없

다; cannot - too

②**않안** : 하지 않으면 안된다; must do

③**않없** : 하지 않을 수 없다; cannot help -ing, 등

④**최것** : 최대한 발휘하는 것이다.

(8)기본 5형식

①S(주어) + V(술어)i : 별이 빛나다;  
Stars Twinkle

②S + Vi + C(보어) : 나는 소년이다;  
I am a boy

③S + Vt + O(목적어) : 어린이는 우유를 먹는다; Children eat milk

④S + Vd + I.O + D.O : 그는 우리에게 영어를 가르친다; .He teaches us English

⑤S + Vt + O + C : 나는 그녀를 행복하게 만들었다; I made her happy

3. 절의 색인방법

3.1 절분석 색인언어

어떤 명제에 대한 문장이 생성되는 과정은 다음과 같이 유추할 수 있다. 처음에는 먼저 대주제 개념이 떠오르고 다음에 이 대주제 개념에 속하거나 연관이 있는 몇몇 개의 소주제 개념들이 그 아래 모여 각각의 소주제 개념이 언어적인 표현을 시도하여 절들을 이루고 그 각각의 절들이 연결되어 하나의 문장을 형성한다. 여러 개의 문장이 이러한 과정을 거쳐 형성되면 개념의 흐름 순서대로 모여 문단을 이루고 문단들이 또 모아져서 문헌을 만든다.

절은 개념서술의 기본단위라고 할 수 있으며 개념서술이 한개 절로 완성되면 그 절은 바로 '단일문'이 된다. 절이나 단일문의 성립요건은 2장의 (8)에서 본 바와 같이 'S(주어)'와 'V(술어)'를 필수적으로 갖고 있어야 하며 따라서 기본 5형식이 절에 그대로 적용되고 9종류의 품사가 절안에서 언어적 규칙에 따라 결합될 수 있어 부분적 소주제 개념을 서술하는데 부족함이 없다. 그래서 보통 문헌들이 색인될 때 명사(기본 5형식에서의 S,C,O)만이 색인어로 채택되지만 개념서술이 중요한 절색인에서는 서술을 최종적으로 결정짓는 V도 색인어로 채택하지 않으면 안된다.

그러므로 절문 문장의 한국어 구문분석의 기본은 절(단일문)단위로 이루어지며 절단위로 이루어진 분석된 구문 중에 주요한 것들을 원래

의 순서대로 나열하면 절 구문의 대변값 즉 색인언어가 된다. 그런데 개념이 발화되어 분화될 때 절마다 명사와 동사가 구성이 되면 명사에 주어지는 조사는 저절로 결정이 된다. 예를 들면 아래의 표1과 같다.

표1. 문장생성 예

(사람, 글, 쓰다)-->사람이 글을 쓰다
(사람, 글, 쓴다)-->사람이 글을 쓴다, 사람은 글을 쓴다.
(돼지, 글, 쓰다)-->돼지를 글로 쓰다.

여기서 “돼지가 글을 쓰다”가 성립한다고 보는 것은 상식적으로 올바른 쓰임이 아니며 가정적 바람에 근거한 예술적 표현이다. 즉 언어적으로 보았을 때 진실이 아니라 허위를 얘기하는 것이 된다. 따라서 색인할 때 글의 의미를 손상치 않는 조사는 생략하도록 한다. 그 결과 색인언어의 구성요소는 (1)명사(주어, 목적어, 보어, 등), (2)용언(어간만 사용), (3)조사(특별한 것만), (4)접속어구, (5)선어말어미, (6)보조용언, (7)용언 상당어구, (8)관용구로 정한다.

3.2 색인작성

앞절의 구성요소를 가지고 먼저 2장의 (8)의 예들을 표2와 같이 작성하였다. 각 예문의 번호 우측의 집합괄호안에 색인어를 기술하였다.

표2. 색인예1

①{별, 빛나}, ②{나[인대], 소년[이]}, ③{어린이, 우유, 먹+현재}, ④{그[인대], 우리[인대], 영어, 가르치+현재}, ⑤{나[인대], 그녀[인대], 행복하게, 만들+과거}
--

샘플체로 쓰여진 단어가 상수 색인어이고 각 괄호안에 필기체로 쓰여진 단어는 변수 색인어로서 상수 색인어의 범주어인데 탐색확장을 할 시에 탐색어로 사용된다. 그리고 '+'기호와 함께 고딕체로 쓰여진 것은 시제와 양상을 나타내는 선어말어미의 변수 색인기호이다. 이어서 조사, 접속어구, 보조용언, 관용구, 복합문의 색인작성 예를 나열한다.

(1)조사의 경우

예문1) 우물은 동에서 서로 22미터이고, 남에서 북으로 12미터이다; The well is 22m from east to west, 12m from north to south

작성예1) { 우물, 등; 에서, 서; 로, 미터이},  
 { 우물, 남; 에서, 북; 으로, 미터이}

(2)접속어구의 경우

예문2) 당신이 말하고자 하는 것을 안다: I see what you mean

작성예2) { 나, 알+현재 },  
 { 당신[인대], 말하+what }

(3)보조용언의 경우

예문3) 회의에서는 메모를 하는 것이 좋다:  
 It has better for you to take notes at the meeting.

작성예3) { 당신, 메모, 회의, 하+권유 }

(4)관용구의 경우

예문4) 친구 선택은 아무리 주의해도 지나침이 없다: We cannot be too careful in the choice of our friends

작성예4) { 우리, 친구, 선택, 이+아지 }

(5)복합문의 경우

예문5) 젊은이가 세상에서 할 수 있는 가장 큰 일은 아마 그에게 부여된 능력을 최대한으로 발휘하는 것일 것이다: The great thing a youth can possibly do in this world is to make the most possible out of the ability that has been given him

작성예5) ①{ 젊은이, 세상, 하+가능+which},  
 ②{ 일, 능력, 최것, 이+부정2 },  
 ③{ 그, 부여되+현완+that }

#### 4. 문장의 색인작성

문장의 색인은 단일문의 경우 절 색인이 그대로 문장 색인으로 적용이 되지만 복합문의 경우는 절 색인들을 통합해야 한다. 통합은 영어에서 관계대명사, 관계부사로 이어질 경우는 주절과 부절에 나타난 명사와 접속어구를 맞추어 가며 연결을 시키고 그렇지 않은 경우는 절의 순서대로 연결한다.

3장의 (5)복합문의 색인작성 예를 하나의 문장 색인으로 통합하면 다음의 표3과 같은 문장 색인이 생성된다. 부절인 ①과 ③의 'which', 'that' 접속어구에는 '일', '능력'이라는 색인어를 추가시켜 주절 ②와의 정확한 연결을 꾀하였다.

표3. 색인예2

{ ①( 젊은이, 세상, 하+가능+which, 일) + ②( 일, 능력, 이+최것 ) + ③( 그, 부여되+현완+that, 능력 ) }
---

#### 5. 결론

영어문장과 한국어문장을 대비 분석하여 절과 문장 색인을 작성할 때 필요한 색인언어를 (1)명사(주어, 목적어, 보어, 등), (2)용언(어간만 사용), (3)조사(특별한 것만), (4)접속어구, (5)선어말어미, (6)보조용언, (7)용언 상당어구, (8)관용구로 정하였다. 본 색인언어는 전칭(명사의 경우)과 어간(용언의 경우) 및 약칭을 색인어나 색인기호로 사용하고 영자와 한글이 혼용된다. 문장을 색인하는 것은 단문이나 절단위로 작성되는데 그 기준은 문장의 기본 5형식이며 이를 중심으로 주변의 요소를 추가하였다. 복합문일 때는 그 문에 속한 절 색인들의 접속형식을 강화하여 순서에 따라 결합시켜 집합적 색인을 만들었다.

본 색인언어에서 앞으로 강조되는 것은 용언과 체언들의 동의어를 묶어 색인어 종수를 줄여야 한다. 그리고 등위접속사와 종속접속사들을 관계대명사의 'which'나 'that'처럼 변수 색인어로 묶을 수 있는 연구가 이루어져야 한다.

#### 참고 문헌

권혜진, 「범주문법과 논리구조에 기반한 자연어 질의의 의미분석」, 한국과학기술원 석사학위논문, 1997.  
 김나리, 「패턴정보를 이용한 한국어 구문 분석」, 서울대학교 대학원 박사학위논문, 1997.  
 김민수, 「국어문법론」, 서울; 일조각, 1989.  
 김민수, 「국어의미론」, 서울; 일조각, 1981.  
 남기심 & 고영근, 「표준 국어문법론」, 서울; 탑출판사, 1987.  
 은종진, 「효율적인 구문분석을 위한 전처리 기구현과 복합명사의 구조분석」, 한국과학기술원 석사학위논문, 1995.  
 Bonura, L.S., 「the Art of Indexing」, New York; John Wiley & Sons, Inc., 1994.