

한국어 동사구 대용어의 제약기반문법적 분석

이용훈
(충남대학교)
김종복
(경희대학교)

Lee, Yong-hun & Kim, Jong-bok. 2012. A Constraint-based Analysis of VP-anaphora in Korean. *The Journal of Studies in Language* 28.1, 111-132. While VP-ellipsis is frequent in English, VP-anaphora prevails in Korean. The goal of this paper is to provide a constraint-based analysis for the VP-anaphora in Korean. This paper adopts Egg et al. (2001)'s Constraint Language for Lambda Structure (CLLS) and develops the algorithms for finding the antecedent of the conjunction VP-anaphora. This paper shows that the English VP-ellipsis within the CLLS system can be adopted for the Korean VP-ellipsis in a similar fashion even though the two languages are typologically different. (Chungnam National University, Kyung Hee University)

Key word: VP-anaphora, constraint-based, CLLS, NP-anaphora

1. 서론

생략현상(ellipsis)은 자연언어에서 빈번하게 나타나는 현상으로 자연언어들이 경제적으로 사용될 수 있도록 하는데 도움을 준다. 그러나 이러한 생략 현상은 임의로 일어나는 것이 아니라, 생략된 부분을 복원(recoverability) 할 수 있을 경우에만 일어난다(Sag, 1976). Johnson(2008:3)은 영어의 생략 현상을 다음과 같이 분류하고 있다.

(1) 생략현상의 유형

a. Sluicing

She read something, but she won't say what [_{NP}].

b. Verb Phrase Ellipsis

She read something and he did [_{VP}] too.

c. Pseudogapping

She'll read something to Sam, but she won't [_{VP}] to Bill.

d. Gapping

She read something to Sam and others [VP] to Bill.

e. Right Node Raising

She deliberately [VP], and he accidentally, read something.

f. Comparative Deletion

Mary read more books than Bill has [VP].

(1b)에서 보는 바와 같이, 영어에서는 동사구 생략현상(VP ellipsis)이 생략 현상 중의 하나로 나타나지만, 한국어에서는 동사구 생략현상이 영어와는 조금 다르게 나타난다. 다음의 예문을 보자.

- (2) a. 그녀는 무엇인가를 읽었고, 그 역시 [VP 무엇인가를 읽었다].
 b. *그녀는 무엇인가를 읽었고, 그 역시 [VP [V]었다].
 c. *그녀는 무엇인가를 읽었고, 그 역시 했다.
 d. 그녀는 무엇인가를 읽었고, 그 역시 그랬다.

(2a)에서 보면 동사구 ‘무엇인가를 읽었다’가 두 번 나와 있다. 여기서 두 번째 문장의 동사구를 무조건 생략해 버리면 (2b)와 같이 비문이 된다. 그렇다고 영어에서 일어나는 do-support처럼 한국어에서도 (2c)와 같이 생략된 부분에 ha-support를 한다고 해서 정문이 되지 않는다. 이 문장을 정문으로 만들기 위해서는 (2d)와 같이 ‘그렇다’라는 용언을 생략된 동사의 자리에 넣어주어야 비로소 정문이 된다. (‘그랬다’는 ‘그렇다’ 또는 ‘그러게 하였다’의 구어적 표현으로 볼 수 있다.) 한국에서는 이처럼 용언이 생략된 뒤에는 대형태인 ‘그렇다’가 생략된 동사구의 자리에 들어가 주어야 정문이 된다. 즉, 한국어에서는 동사구 대용현상(VP-anaphora)이 나타나는 것이다.

본 논문에서는 한국어에 나타나는 이러한 동사구 대용현상을 제약기반문법에서는 어떻게 분석할 수 있는지를 살펴보고자 한다. 동사구 대용현상의 분석과 그 전산학적 구현을 위하여 본 논문에서는 Egg et al. (2001)에 제시되어 있는 CLLS를 한국어의 분석에 적용한다. Egg et al. (2001)에서는 λ-표현을 위한 제약기반 언어를 제시하고, Egg and Erk (2001)에서 이를 영어에 적용하여 동사구 생략현상을 분석하고 있다. 물론 한국어에서는 동사구 생략현상이 아닌 동사구 대용현상이 나타나지만, Egg and Erk (2001)에 제시된 분석기법은 한국어의 동사구 대용현상에도 충분히 적용될 수 있다. 따라서 본 논문에서는 Egg and Erk (2001)의 분석방법을 적용하여 동사구 대용현상, 그 중에서도 접속부사에 의해서 두 문장이 연결되어 있는 문장들을 설명해 보겠다.

본 논문은 다음과 같이 구성되어 있다. 먼저 2장에서는 한국어에서 나타나는 동사구 대용현상에 대하여 그 특징들을 고찰해 보고 이를 세종전자사전에서 실제 예문들을 추출하여 살펴보겠다. 3장에서는 Egg et al. (2001)에 기술되어 있는 λ -표현을 위한 제약기반 언어(CLLS)를 제시하고, 이것을 활용하여 영어의 동사구 생략현상이 어떻게 분석되는지 살펴보도록 하겠다. 4장에서는 CLLS를 한국어를 설명하는 데 도움이 되도록 수정한 후, 이를 적용하여 한국어의 동사구 대용현상, 그 중에서도 접속부사에 의해서 두 문장이 연결되어 있는 문장들에서 일어나는 동사구 대용현상들을 분석해 보겠다. 5장에서는 이러한 설명방법의 장점을 제시하고 본 논문을 마무리하겠다.

2. 한국어의 동사구 대용어

2.1. 대용어 ‘그렇다’

본 절에서는 대용어 ‘그렇다’가 가지는 여러 가지 특징들을 살펴보기로 하겠다. 먼저, 대용어 ‘그렇다’는 (2d)와 같이 동사를 대신하여 쓰이기도 하지만 형용사도 대신할 수 있다.

- (3) a. 승리도 순간이고, 패배도 그러하다.
b. 그녀는 무엇인가를 읽었고, 그 역시 그랬다. (=2d)

(3b)에서는 대용어 ‘그렇다’가 동사인 ‘읽다’를 대신하고 있지만, (3a)에서는 형용사인 ‘순간이다’를 대신하고 있다. ‘그렇다’ 자체는 형용사이지만, 이 단어는 형용사만을 대신하는 것이 아니라 동사도 대신할 수 있다. 이는 한국에서는 동사와 형용사가 ‘용언’이라는 하나의 범주로 함께 묶일 수 있기 때문인데, 이러한 점을 고려한다면 대용어 ‘그렇다’는 ‘용언’을 대신하는 대형태로 보는 것이 타당할 것이다.¹⁾

또한 대용어 ‘그렇다’는 여러 가지 형태론적 변이형을 취하고 있다. 다음의 문장들을 보자.

- (4) a. 그녀는 무엇인가를 읽었고, 그 역시 그랬다.
b. 그녀는 무엇인가를 읽었고, 그 역시 그러했다.

1) Sohn(1999:404)에서도 같은 언급을 하고 있다. 비록 ‘그렇다’의 품사가 형용사이기는 하지만, 형용사도 대신할 수 있을 뿐더러 동사도 대신할 수 있기 때문에 ‘pro-predicate’라고 언급하고 있다.

- c. 그녀는 무엇인가를 읽었고, 그 역시 그러하였다.
 d. 그녀는 무엇인가를 읽었고, 그 역시 그렇게 하였다.
- (5) a. 그녀는 무엇인가를 읽었다 그리고 그 역시 그랬다.
 b. 그녀는 무엇인가를 읽었다 그리고 그 역시 그러했다.
 c. 그녀는 무엇인가를 읽었다 그리고 그 역시 그러하였다.
 d. 그녀는 무엇인가를 읽었다 그리고 그 역시 그렇게 하였다.

먼저 (4)와 (5)의 차이점은 (4)에서는 접속어미가 동사에 부착되어 있으나 (5)에서는 접속 부사가 하나의 단어로 독립되어 있다는 것이다. 즉, (4)는 이분지 등위접속(binary co-ordination)에 해당되고 (5)는 삼분지 등위접속(ternary co-ordination)에 해당된다. (4)와 (5)의 예문들이 보여 주듯이, 대용어 ‘그렇다’는 ‘그랬다’, ‘그러했다’, ‘그러하였다’, ‘그렇게 하였다’ 등 여러 가지 형태적 변이형을 취하고 있다. ‘그렇다’가 지니는 이러한 특성들은 한국어의 동사구 대용현상을 체계적이고 일관되게 설명하는 데 어려움을 주고 있다.²⁾

2.2. ‘그렇다’의 선행동사구: 문장 vs. 담화

본 논문에서는 ‘그렇다’의 쓰임을 살펴보기 위하여 세종전자사전 ‘그렇다’가 사용된 문장들을 추출하고, 이 문장들을 분석하여 ‘그렇다’가 지칭하는 용언이 주어진 문장 내에 있는지 혹은 담화상에 있는지, 그리고 ‘그렇다’가 가리키는 선행용언이 주어진 문장 내에 존재한다면 주로 어떠한 문장의 패턴과 많이 나타나는 지를 조사해 보았다. 물론, 세종전자사전의 모든 예문들을 대상으로 조사하는 것이 바람직하겠으나, ‘그렇다’는 용언을 대신하여 쓰이는 단어이므로 본 논문에서는 형용사 부분과 동사 부분만을 선택하여 예문들을 추출해 보았다.

2) 한국어의 대용어 ‘그렇다’가 형용사와 동사를 모두 대신하여 쓰일 수 있다고 하더라도 형용사를 대신할 때와 동사를 대신할 때에 다른 통사적인 특성을 보인다. ‘그렇다’가 형용사를 대용할 때는 다른 활용 변화없이 그대로 ‘그렇다’와 시제어미의 결합만으로도 충분하지만, 동사의 경우에는 그렇지 않다. 동사의 경우에는 ‘그렇게 했다’나 ‘그렇게 할 것이다’와 같이 ‘그렇 + 게’라는 부정사형이 필요하게 된다. 본 논문에서는 이러한 사실을 부정하는 것이 아니라, 동사의 경우에는 ‘그랬다’, ‘그러했다’, ‘그러하였다’, ‘그렇게 하였다’ 등 여러 가지 형태적 변이형들이 나타나서 이들을 모두 다루기는 어렵지만, 동사의 경우 ‘그렇 + 게’라는 부정사형으로 쓰이더라도 여기의 ‘그렇다’는 앞에 나오는 동사를 가리킨다는 점은 변하지 않기 때문에, 이러한 점에서 ‘그렇다’가 형용사와 동사를 모두 대신하여 쓰일 수 있다고 한 것이다.

먼저, ‘그렇다’가 가리키는 선행용언이 주어진 문장 내에 존재하는 경우를 살펴보면, 가장 빈번하게 나타나는 경우가 다음과 같이 접속어미를 사이에 두고 두 문장이 등위접속된 상태에서 두 번째 문장에 쓰인 ‘그렇다’가 첫 번째 문장의 용언을 가리키는 경우이다.³⁾

- (5) a. 승리도 순간이고, 패배도 그러하다. (형)
 b. 배움이 다양성을 받아들이는 것처럼 우리의 삶도 그러하다. (형)
 c. 자다가 이불을 내차면 배가 아프니 그러지 마라. (동)
 d. 많은 직원들이 김 과장에 대해 사장에게 불평하고 싶지만 그러지 못했다. (동)

이 문장들에서 대응어 ‘그렇다’는 모두 접속어미로 연결된 첫 번째 문장의 형용사나 동사를 대신하고 있다.

그러나 ‘그렇다’가 접속어미로 연결된 문장들에서만 쓰이는 것은 아니었다. 다음 문장들에서와 같이 ‘그렇다’가 명사를 수식하는 관형절 안에 있는 용언을 지칭하는 경우들도 있었다.

- (7) a. 기영이는 수진이가 준호에게 잘하는 것을 준호에게 부탁할 것이 있어서 그러는 것으로 빗보았다. (동)
 b. 친구들한테는 기영이가 수진이에게 잘하는 것이 수진이한테 관심이 있어서 그러는 것으로 빗보이고 있어요. (동)
 c. 내가 보기에 학생들은 박수를 치고 환호를 하고 하는 것이 마음에서 우리나라서 그러는 것 같지가 않았다. (동)
- (8) 교수님은 주례를 봐 달라는 제자의 부탁에 그러자고 승낙해 주셨다. (동)

(7)에서는 의존명사 ‘것’을 수식하는 관형절 안에 있는 용언을 가리키는 경우이고, (8)은 ‘것’이외의 명사를 수식하는 관형절 안에 있는 용언을 가리키는 경우이다.

한편, ‘그렇다’가 지칭하는 용언이 주어진 문장 내에 있는 것이 아니라 담화상에 존재하는 경우에는 ‘상태가 그렇다’ 또는 ‘그렇게 행동하다’ 정도의 의미를 갖는 경우가 많았으며, 그 예들은 다음과 같았다.

3) 여기서 각각의 문장 뒤에 ‘(형)’이라고 붙어있는 문장들은 주어진 문장이 세종전 자사전의 형용사 부분에서 추출되었음을 의미하며 ‘(동)’이라고 붙어있는 문장들은 주어진 문장이 동사 부분에서 추출되었음을 의미한다.

- (9) a. 비가 오려고 그러는지 날씨가 영 침울하다. (형)
 b. 너에게 그러는 게 아니었는데, 내가 경솔했다. (형)
 c. 너 아무리 어머니에게 앵똥아줬다고 해서 그러는 것 아니다. (동)
 d. 나는 언니가 책상에 엎드려 우는 것을 단지 힘들어서 그러는 줄로 알았다. (동)

이 문장 속에 나오는 ‘그렇다’는 주어진 문장 내의 어떠한 특정 용언을 가리키는 것이 아니라 담화상에 존재하는 어떤 용언을 가리키며, ‘상태가 이리이러하다’ 또는 ‘그렇게 행동하다’ 정도의 의미를 갖게 된다. 그리고 이러한 용례는 평서문에서만 나타나는 것이 아니라 다음과 같이 의문문에서도 나타난다.

- (10) a. 누구 볼기맞는 꼴을 보려고 그러시오? (동)
 b. 너 내가 망하는 꼴을 보고 싶어서 그러나? (동)
 c. 이만하면 남부러울 것 없는데 또 무엇이 탐나서 그러는 거요? (동)
 d. 내가 시간이 없어서 그러는데 전기요금을 좀 은행에 대납해 줄 수 있을까? (동)

마지막으로 다음과 같이 특별한 의미가 없이 관용적으로 쓰이는 ‘그렇다’도 있다.

- (11) a. 그러지 말고 사정 한 번 봐 주세요. (동)
 b. 어제 산 옷이 작아서 그러는데 환불됩니다까? (동)

여기서의 ‘그렇다’의 의미를 굳이 해석하자면 ‘상태가 그렇다’ 또는 ‘그렇게 행동하다’ 정도로 해석할 수도 있겠지만, 그렇게 해석하기 보다는 이제는 너무 많이 쓰이어져서 상투화된 표현으로 생각할 수도 있다.

동사구 대용어 ‘그렇다’의 이러한 특징들은 다음과 같은 명사구 대용어와도 비교가 된다.

- (12) a. 영희는 자신을 거울에 비추어 보았다.
 b. 자신의 할 일을 남에게 미루지 말라.

(12a)의 재귀대명사 ‘자신’의 선행사는 문장 안에 존재하는 ‘영희’이지만 (12b)의 ‘자신’의 선행사는 담화상에서만 존재하는 일반적인 사람이다. 즉, 명사구 대용어인 재귀대명사 ‘자신’의 선행사는 (12a)와 같이 주어진 문장

내에 존재할 수도 있지만 (12b)와 같이 담화상에서만 해석이 가능할 수 있다. 이러한 점을 고려하여 본다면, 동사구 대용현상은 명사구 대용현상과 유사한 특징을 보인다고 할 수 있을 것이며, 따라서 동사구 대용현상과 명사구 대용현상이 유사한 분석방법으로 분석할 수 있는 근거가 된다.⁴⁾

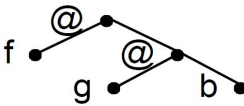
3. Constraint Language for Lambda Structure와 생략현상

3.1. Constraint Language for Lambda Structure

CLLS은 Egg et al. (1998), Erk and Niehren (2000), Egg and Erk (2001), Egg et al. (2001), Erk et al. (2002) 등에 의해 발전된 이론으로서 λ -표현들을 비명세(underspecification)의 방법을 사용해서도 올바르게 분석할 수 있도록 만들어진 이론이다. CLLS의 기본 아이디어는 모든 λ -term들을 트리형태의 그래프(a tree-like graph)로 표현되며, 이 트리구조 안에서 각각의 노드들은 λ -term들의 subterm을 나타낸다는 것이다.

CLLS가 자연언어의 문장들을 어떻게 분석하는지를 보기 전에 CLLS의 가장 기본적인 개념을 살펴보면 다음과 같다. 첫째, 문장의 모든 term들은 트리의 형태로 표현된다. 예를 들어, $f(g(b))$ 와 같은 term은 다음과 같이 분석된다. (Egg and Erk, 2001:164)

(13) $f(g(b))$ 의 CLLS식 표현

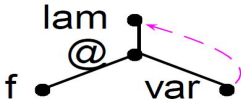


여기서 함수적용(functional application)은 '@'이라는 라벨이 붙어있는 노드로 나타내며, 'apply'의 의미를 갖는다. 이 노드는 두 개의 자노드를 갖는데, 첫 번째는 λ -term이 적용되는 함수이고 두 번째는 그 함수의 논항이다.

둘째, 모든 λ -term들은 λ -link라는 것과 함께 나타난다. 예를 들어, $\lambda x.f(x)$ 는 다음과 같이 나타난다. (Egg and Erk, 2001:164)

4) Egg et al. (2001:6-7)에서는 동사구 생략현상과 명사구 대용현상의 유사성을 밝히고, 이 두 현상이 비슷한 알고리즘으로 분석될 수 있음을 밝히고 있다. 본 논문에서는 영어의 동사구 생략현상을 CLLS로 분석한 것을 한국어의 동사구 대용현상에 적용하고 있는데, 이것이 가능하다면, 한국어에서도 동사구 대용현상과 명사구 대용현상이 비슷한 알고리즘으로 분석될 수 있으리라 생각된다.

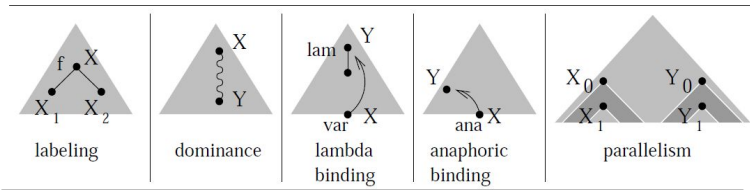
(14) $\lambda x.f(x)$ 의 CLLS식 표현



여기서 λ -abstraction과 결속되는 변수들은 각각 lam과 var로 나타내며 λ -link는 빨간색 점선화살표로 나타낸다. 이 λ -link의 목적은 어느 변수가 어디에 결속되는지를 도식적으로 보여주는 데 있다.

셋째, CLLS는 (14)와 같은 가장 기본적인 리터럴(literal)을 비롯하여 다음과 같은 형태의 리터럴들로 구성되어 있다. (Egg and Erk, 2001:164)

(15) CLLS의 리터럴들



여기서 ana는 대용어(anaphoric expressions)를 가리키고 있으며, λ -term이 λ -link에 의해 λ -binding되는 것과 같이 ana는 **ante**-link에 의해 **ante**-binding이 되도록 되어 있다.

넷째, CLLS의 모든 문장들(formula)은 이러한 리터럴들의 복합체로 구성되어 있으며 다음과 같이 형식화할 수 있다. (Egg and Erk, 2001:164)

(16) CLLS의 문장

$$\begin{aligned}
 \varphi ::= & X:f(X_1, \dots, X_n) \mid X \triangleleft^* Y \\
 & \mid \lambda(X) = Y \mid \text{ante}(X) = Y \\
 & \mid X_0/X_1 \sim Y_0/Y_1 \mid \varphi \wedge \varphi'
 \end{aligned}$$

먼저 labeling 리터럴은 (14)에 제시된 것으로서 $X:f(X_1, \dots, X_n)$ 은 X 가 f 라고 레이블링되어 있는 노드가 X_1, \dots, X_n 로 구성되어 있어야 함을, 즉 X_1, \dots, X_n 가 f 의 자노드들이어야 함을 나타낸다. Dominance 리터럴은 X 라는 노드가 Y 라는 노드를 지배하고 있어야 함을 나타낸다. 그 다음, binding 리터럴 λ

$(X)=Y$ 와 $\text{ante}(X)=Y$ 는 λ -term이 λ -link에 의해 λ -binding이 되어야 하며 ana 는 ante -link에 의해 ante -binding이 되어야 함을 나타낸다. 마지막으로 나오는 parallelism 리터럴은 CLLS에서는 두 개 이상의 문장들이 연결되어 하나의 문장이 될 수 있음을 나타내며, 이때 X_0 에서 X_I 을 제외한 부분(segment)은 Y_0 에서 Y_I 을 제외한 부분과 동일해야 함을 나타낸다.

자, 이제 CLLS가 어떻게 자연언어의 문장들을 분석하는지 살펴보기 위하여 다음의 예문을 보자.

(17) Every linguist attends a workshop.

이 예문은 전형적인 양화사 작용역(scope)의 중의성을 나타내는 문장으로 서 다음과 같이 두 개의 의미를 갖는다.

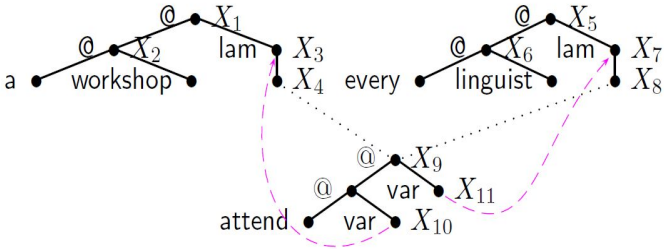
- (18) a. (a workshop) (λx
 (every linguist) (λy
 (attend x) y))
 b. (every linguist) (λy
 (a workshop) (λx
 (attend x) y))

Copestake et al. (2005)의 Minimal Recursion Semantics(MRS)에서는 이러한 양화사의 작용역에 나타나는 중의성을 처리하기 위하여 비명세를 사용한다.

이를 CLLS를 사용하여 분석한다면 다음과 같다. 먼저 (17)의 문장을 세 부분으로 나누어 보면, (every linguist)(λx (.)), (a workshop)(λy (.)), 그리고 ((attend x) y)로 나뉘어진다.⁵⁾ 이 세 개의 λ -term들을 트리형태의 그래프로 표현해 보면 다음과 같다(Egg et al. 2001:4).

5) 여기서 y 는 목적어인 a workshop에 그리고 x 는 주어인 every linguist에 해당한다. 타동사 attend는 주어인 every linguist보다는 목적어인 a workshop과 먼저 결합하기 때문에, 논리식에서도 y 가 x 보다 attend와 먼저 결합한다.

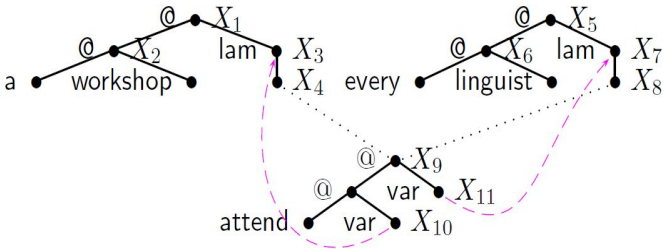
(19) (17)의 CLLS분석



여기서 X_4 와 X_8 은 X_9 를 지배할 수 있고 ($X_4 \triangleleft^* X_8$, $X_4 \triangleleft^* X_9$), X_{10} 은 X_3 에 연결되며 X_{11} 은 X_7 에 연결되어 있다 ($\lambda(X_{10})=X_3$, $\lambda(X_{11})=X_7$).

이를 사용하여 (17a)의 의미를 나타내보면 다음과 같다. (Egg et al. 2001:3)

(20) (17a)의 의미



이 그림에서는 a workshop의 작용역이 (every linguist)($\lambda x(((\text{attend } x) y))$)이고 every linguist의 작용역이 ((attend x) y)이기 때문에 (17a)의 의미를 갖게 된다. 이 둘의 작용역이 바뀌게 되면 (17b)의 의미를 도출할 수 있다. 따라서 CLLS에서는 이 둘의 의미를 모두 도출할 수 있는 (19)의 분석을 가장 기본적인 분석으로 삼고 있다.

Egg et al. (2001)와 Egg and Erk (2001)에 따르면 CLLS는 이와 같은 양화사의 분석뿐만 아니라 생략현상(ellipsis)이나 대용어(anaphora)의 분석에도 효과적으로 사용될 수 있다고 언급하고 있다. 따라서 영어의 동사구 생략현상 역시 CLLS로 부석이 가능하며 이를 응용하면 한국어의 동사구 대용현상도 효과적으로 설명하는 것이 가능하리라 생각된다. 또한 이렇게 분석된 결과들은 LKB시스템(Linguistic Knowledge Builder; Copestake 2002) 등을 활용하여 전산적으로 구현이 가능하리라 생각된다.

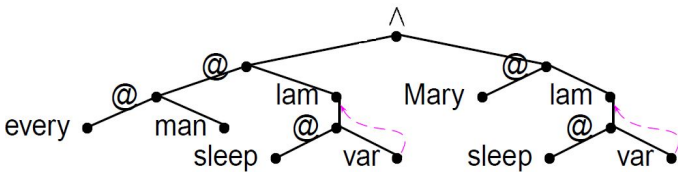
3.2. CLLS를 이용한 영어의 동사구 생략현상의 구현

이제 CLLS를 사용하여 영어의 동사구 생략현상이 어떻게 분석되는지 살펴보자. 먼저 다음의 문장을 보자. (Egg and Erk, 2001:166)

- (21) a. Every man slept and Mary did. too.
- b. Every man -ed [VP sleep] and Mary did [VP sleep]. too.

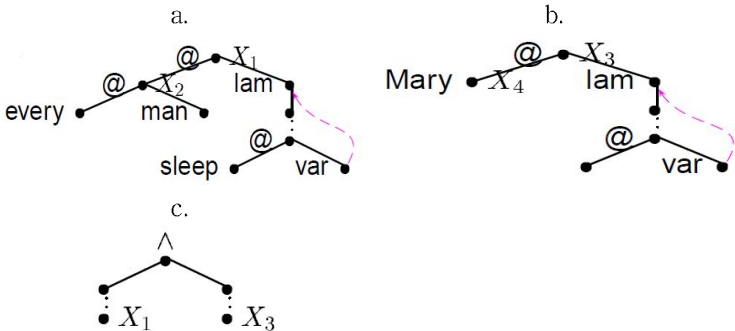
(21a)의 문장은 (21b)에서 ‘sleep’이라는 동사가 생략되어서 나온 문장이다. 이 문장을 CLLS로 분석해 보면 다음과 같이 된다. (Egg and Erk, 2001:167)

(22) (21)의 CLLS 분석



이 문장은 세 개의 부분으로 구성되어 있는데, 각 부분을 나누어서 나타내 보면 다음과 같다. (Egg and Erk, 2001:166)

(23) (22)의 세 부분



두 개의 문장 (23a)와 (23b)를 ‘and’로 연결하여 하나의 문장을 만드는 데에는 (23c)가 사용된다. 즉, (23c)의 X_1 에 (23a)가 연결되며, X_3 에 (23b)가 연결된다. 그리고, 여기에 parallelism 리터럴이 적용된다. 이를 형식화하여 나타내면 다음과 같다. (Egg and Erk, 2001:166)

$$(24) X_1/X_2 \sim X_3/X_4$$

즉, X_1 의 트리에서 X_2 가 의미적으로 기여하는 부분을 제외한 나머지 부분은 X_3 의 트리에서 X_4 가 의미적으로 기여하는 부분을 제외한 나머지 부분과 동일하다는 것이다. 이는 (21)에서 ‘did’가 의미적으로 ‘slept’와 동일하다는 것으로, ‘did’가 ‘slept’를 가리킨다라는 것을 알 수 있다. CLLS에서는 이러한 방법을 사용하여 영어의 동사구 생략현상을 분석하고 있다.

그러면 이러한 분석방법을 LKB 등이 사용하는 자질구조로 구현하면 어떻게 될까? Egg and Erk (2001:167)에서는 이러한 현상을 분석하기 위하여 다음과 같은 Attribute-Value Matrix(AVM)를 제시하고 있다.

(25) SEM의 AVM

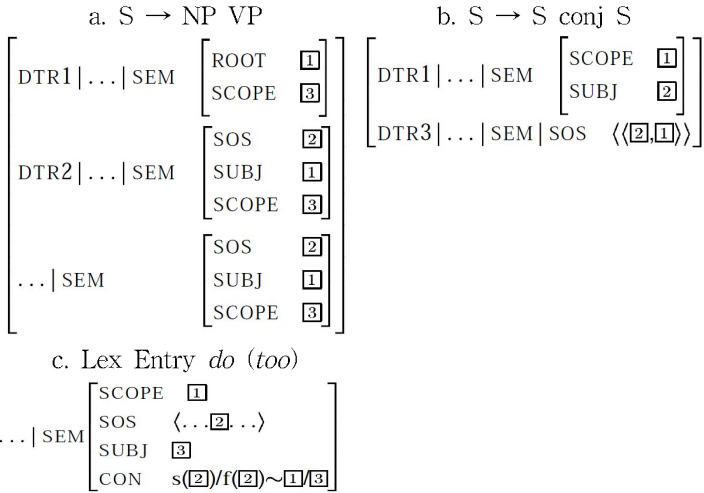
SEM	<table style="border-collapse: collapse; width: 100%;"> <tr> <td style="padding: 2px 5px;">SCOPE</td> <td style="padding: 2px 5px;">the scope variable</td> </tr> <tr> <td style="padding: 2px 5px;">ROOT</td> <td style="padding: 2px 5px;">the root variable</td> </tr> <tr> <td style="padding: 2px 5px;">CON</td> <td style="padding: 2px 5px;">the constraint for this constituent</td> </tr> </table>	SCOPE	the scope variable	ROOT	the root variable	CON	the constraint for this constituent
SCOPE	the scope variable						
ROOT	the root variable						
CON	the constraint for this constituent						

여기서 SCOPE는 작용역(local scope domain)을 나타내는 변수로 보통은 통사구조에서 가장 작은 문장을 지칭한다. ROOT는 각 구성소의 루트를 나타내며, CON은 각각의 구성소에 적용되는 여러 가지 제약조건들을 나타낸다.

이 AVM은 (26)과 같이 ‘do (too)’라는 조동사는 물론 동사구 생략현상과 관련이 있는 여러 문법규칙들에 적용된다.⁶⁾

6) 영어의 ‘do so’나 ‘do it’도 본 절에 나오는 ‘do too’의 분석방식과 비슷하게 분석된다.

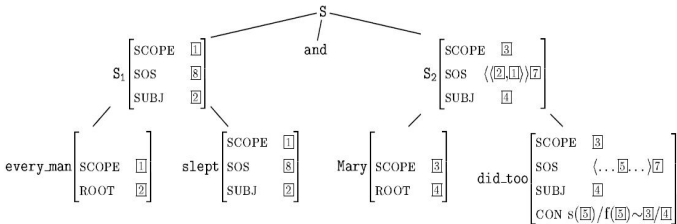
(26) 문법규칙과 *do (too)*에 적용되는 제약조건들



여기에 나오는 SOS는 원시문장(SOURCE Sentence; 선행문장)을 나타내며, SUBJ는 해당하는 문장의 주어, 그리고 $f(X)$ 와 $s(X)$ 는 각각 X 의 첫 번째 멤버와 두 번째 구성소를 나타낸다.

이 제약조건들을 사용하여 (21a)를 분석해보면 다음과 같다.

(27) (21a)의 분석



이 분석은 다음과 같이 설명될 수 있다. 먼저 *every man*과 *slept*가 결합될 때에는 (26a)가 적용된다. 명사구 *every man*은 동사 *slept*의 주어가 되므로 *slept*의 SUBJ의 값은 *every man*의 ROOT의 값과 같게 된다. 그리고 *every man*과 *slept*는 한 문장 S1의 구성소이므로 local domain이 같아서 *every man*의 SCOPE의 값은 *slept*의 SCOPE의 값과 같게 된다. 그리고 이 문장의 핵어는 동사 *slept*이므로 S1의 SOS값은 *slept*의 SOS의 값과 같

게 된다.

이 문장은 접속사 *and*에 의해 다음 문장 *Mary did too*로 연결된다. 여기에 (26b)가 적용되어 첫 번째 문장의 SUBJ와 SCOPE의 값들이 <SUBJ, SCOPE>의 순서쌍을 이루어 두 번째 문장의 SOS값이 된다.

두 번째 문장은 *Mary*와 *did too*로 이루어져 있는데, 명사구 *Mary*가 동사구 *did too*의 주어가 되므로 *did too*의 SUBJ의 값은 *Mary*의 ROOT의 값과 같게 된다. 그리고 *Mary*와 *did too*는 한 문장 S_2 의 구성소이므로 local domain이 같아서 *Mary*의 SCOPE의 값은 *did too*의 SCOPE의 값과 같게 된다. 마지막으로 *did too*의 CON에 의하면 첫 번째 문장 *every man slept*에서 *every man*을 제외한 나머지 부분의 의미는 *Mary did too*에서 *Mary*를 제외한 나머지 부분의 의미와 동일하게 된다. 따라서 두 번째 문장의 *did too*는 첫 번째 문장의 *slept*를 가리키게 된다.

4. CLLS를 이용한 동사구 대용어의 분석

4.1. 영어의 동사구 생략과 한국어의 동사구 대용어

3절에서 보듯이 영어의 동사구 생략현상은 Egg et al. (2001)의 CLLS로 분석할 수가 있었다. 그러면 한국어의 동사구 대용현상도 CLLS로 분석할 수 있을까? 그 가능성을 타진하기 위해서는 먼저 이 두 현상의 공통점과 차이점을 살펴볼 필요가 있다.

먼저 다음의 두 문장을 보자.

- (28) a. John slept and Mary did. too.
b. 승리도 순간이고 패배도 그러하다.

이 두 문장은 완전히 동일하지는 않지만 그 구조만큼은 비슷하다. 두 문장 모두 등위접속으로 연결되어 있으며 각각의 문장들은 ‘주어+자동사/형용사’의 형태를 취하고 있다. 그러나 (28a)의 문장은 다음과 같이 동사구 생략현상에 의해 만들어진 반면 (28b)의 문장은 동사구 대용현상에 의해서 만들어졌다.

- (29) a. John slept and Mary did [_{VP} sleep]. too.
b. 승리도 순간이고 패배도 [_{VP} 순간이다/그러하다].

이러한 차이가 (28a)와 (28b)의 가장 큰 차이점이다.

그러나 이 문장들을 자세하게 살펴보면 영어의 동사구 생략현상과 한국어의 동사구 대용현상을 비슷한 알고리즘으로 분석할 수 있는 단서를 얻게 된다. 다음의 문장들을 보자.

- (30) a. *John slept and Mary did. too.
b. *승리도 순간이고 패배도.

이 문장들은 (28)의 문장들에서 각각 'did'와 '그리하다'를 생략해본 문장들이다. 위에서 보듯이 두 문장 모두 비문이 된다. 그렇다면 한국어의 대형태 '그리하다'는 영어에서 VP가 생략되고 난 이후의 조동사 'did (too)'와 동일한 역할을 하고 있다고 할 수 있다. 따라서 영어의 'did (too)'에 부여했었던 제약조건들을 한국어의 '그리하다'에 부여한다면 영어의 동사구 생략현상과 한국어의 동사구 대용현상을 비슷한 알고리즘으로 분석할 수 있을 것이다.

4.2. 한국어 동사구 대용어의 분석

이제, CLLS를 사용하여 한국어의 동사구 대용현상을 분석해보자. 그러나, 위의 2절에서도 보듯이 대용어 '그렇다'가 나타내는 형태론적 변이형들도 다양할뿐더러 이 대용어가 가리키는 선행용언이 주어진 문장 속에 존재하지 않을 수도 있기 때문에, 본 논문에서 이 모든 경우를 다루기는 어렵다. 따라서 본 논문에서는 여러 가지 동사구 대용현상 중에서 가장 기본적인 등위접속 구문에서의 동사구 대용현상만 다루는 것으로 하겠다. 그리고 본 논문에서는 '그렇다'의 선행용언이 문장 안에 존재하는 경우로만 국한하여 설명하도록 하겠다.

본 논문에서는 한국어의 동사구 대용현상을 올바르게 분석하기 위하여 Egg and Erk(2001)의 분석과는 다르게 몇 가지의 변화를 주었다.

첫 번째는 SOS에 대한 것이다. Egg and Erk(2001)의 분석에 따르면, SOS는 선행하는 문장을 가리킨다. 그런데, (27)의 분석을 보면 첫 번째 문장도 SOS의 값을 가지고 있을 뿐 아니라, 이 값은 등위접속을 통하여 두 번째 문장으로 전달되지도 않는다. 두 번째 문장으로 전달되는 값들은 첫 번째 문장의 SOS값이 아닌 ROOT와 SUBJ값이다. 따라서 본 논문에서는 첫 번째 문장의 SOS의 값을 없는 것으로, 즉 첫 번째 문장의 선행문장은 없는 것으로 하겠다.

두 번째, Egg and Erk(2001)에서는 parallelism 리터럴의 제약조건인

‘ $X_1/X_2 \sim X_3/X_4$ ’를 구현하기 위해서 문장 전체의 인덱스(SCOPE)와 주어부분의 인덱스(SUBJ)를 모두 사용하고 있다. 왜냐하면 제약조건인 ‘ $X_1/X_2 \sim X_3/X_4$ ’이 말해 주듯이, 문장 전체의 인덱스와 주어부분의 인덱스를 제외한 부분이 동사구가 되기 때문이다. 그러나 핵어주도구구조문법(Head-driven Phrase Structure Grammar; Pollard and Sag, 1994; Sag and Wasow, 1999; Sag et al., 2003; Kim and Sells, 2008)에서는 동사가 문장의 핵어(head)가 되기 때문에, 목표문장(target sentence; 대응어를 포함하고 있는 문장)으로 인덱스를 넘겨줄 때에 문장의 핵어가 되는 동사(구)의 인덱스를 넘겨준다면 한국어의 동사구 대응현상을 좀 더 용이하게 구현할 수 있을 것이다. 따라서 본 논문에서는 원시문장 전체의 인덱스와 주어부분의 인덱스를 모두 목표문장으로 넘겨주는 것이 아니라, 원시문장의 동사(구)의 인덱스를 넘겨주도록 하였다. 즉, 목표문장의 SOS의 값은 원시문장의 동사(구) 인덱스가 되는 것이다.

세 번째, 본 논문에서는 Copestake et al. (2005)의 MRS에 쉽게 적용하도록 하기 위하여 (25)에 제시된 자질들을 CONST라는 자질 안에 넣어서 나타내겠다. 즉, 기존의 MRS의 SYNSEM|LOCAL|CONT에는 HOOK, KEY, RELS, HCONS 등 네 개의 자질값이 들어가 있었으나, 본 논문의 분석에서는 여기에 CONST라는 자질을 추가하고 여기에 (25)에 기술된 자질들을 묶어서 넣도록 하겠다.

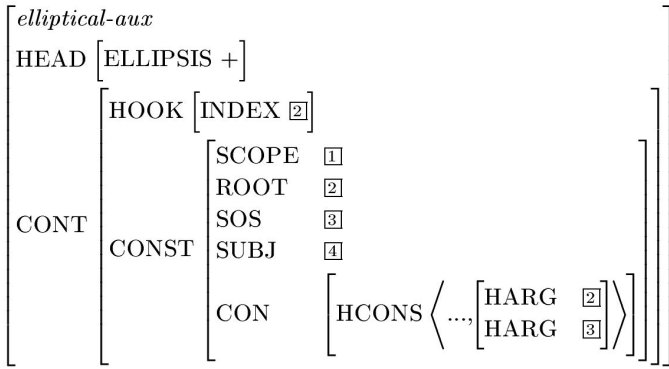
마지막으로, (28b)의 문장에서 두 번째 문장의 ‘그러하다’는 물론 첫 번째 문장의 ‘순간이다’를 가리키기는 하지만, 이 두 용언은 엄격하게 말하면 두 개의 다른 용언이다. 따라서 본 논문에서는 이러한 대응관계를 HCONS의 *qeql*를 사용하여 나타내도록 하겠다.

영어의 동사구 생략현상과 비슷하게 한국어의 동사구 대응현상을 분석하기 위해서는 (26)과 같이 대응어 ‘그러하다’와 구구조규칙들에 대한 제약조건들이 필요하다. 본 논문에서는 (26)에 제시되어 있는 세 가지 제약조건들 중에서 (26a)만 그대로 사용하고 (26b)와 (26c)는 다음과 같이 변형한다.

(31) ‘S → S-and S’에 대한 제약조건

$$\left[\text{ARGS} \left\langle \left[\text{ROOT} \square \right], \text{conjunction}, \left[\begin{array}{l} \textit{elliptical-aux} \\ \text{SOS} \square \end{array} \right] \right\rangle \right]$$

(32) 대용어 ‘그렇다’의 AVM

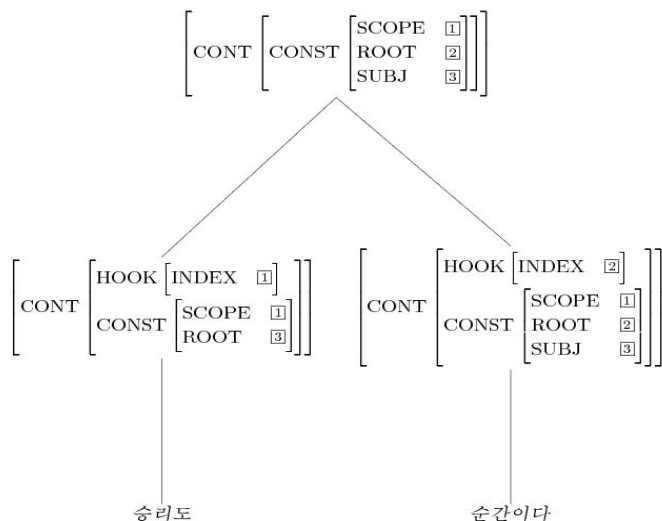


(31)은 첫 번째 문장의 ROOT값은 두 번째 문장의 SOS의 값과 같아야 함을 나타낸다. 즉, 이 제약조건은 첫 번째 문장의 ROOT값을 두 번째 문장의 SOS의 값으로 넘겨주는 역할을 한다. 이때 두 번째 문장의 핵어는 반드시 *elliptical-aux*이어야 한다. 즉, 이 제약조건은 동사구 대용현상에만 사용되며 일반적인 등위접속현상의 분석이나 기술에는 사용되지 않는다. 한편, (32)는 대용어 ‘그렇다’의 AVM을 나타내는데, 여기서 CON|HCONS가 나타내듯이 두 번째 문장의 ROOT의 값은 SOS의 값과 동일하다는 사실을 나타낸다. 즉, (32)의 대용어 ‘그렇다’는 등위접속을 통해서 넘어온 첫 번째 문장의 동사(구)의 인덱스와 같아야 함을 나타낸다. 이를 다시 해석하면 두 번째 문장의 대용어 ‘그렇다’는 첫 번째 문장의 동사(구)를 지칭한다는 것이다.

자, 이제 (28b)의 문장이 실제로 어떻게 분석되는지 살펴보자. 이 문장에서는 이분지등위접속이 사용되고 있으나, 아래의 설명에서는 삼분지등위접속을 기본으로 하여 설명하겠다. Kim and Yang (2006), Kim and Yang (2010), Kim et al. (2011) 등을 보면 한국어에서 이분지등위접속과 삼분지등위접속이 같은 역할을 하고 있음을 알 수 있다. 따라서 본 논문의 분석이 삼분지등위접속을 기반으로 하고 있지만, 이러한 분석방법은 이분지등위접속에도 그대로 적용될 수 있다.

먼저, 첫 번째 문장 ‘승리도 순간이다’에는 (26a)의 제약조건들이 적용되며 다음과 같이 분석된다.

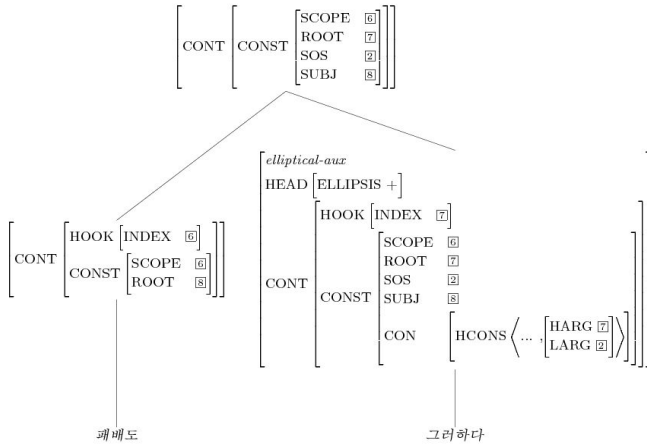
(33) ‘승리도 순간이다’의 분석



먼저 명사구 ‘승리도’는 첫 번째 문장의 주어이므로 ‘순간이다’의 SUBJ의 값에 ‘승리도’의 ROOT의 값이 주어진다. 그리고 ‘승리도’와 ‘순간이다’는 첫 번째 문장의 주부와 술부에 해당하므로 이 둘은 동일한 SCOPE값을 갖게 된다. 그 다음, 첫 번째 문장의 핵어는 동사(구) ‘순간이다’이므로 ‘순간이다’의 ROOT 값이 문장의 ROOT 값으로 올라가게 된다.

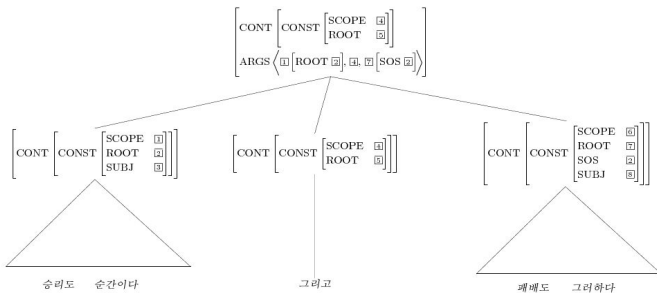
다음, 두 번째 문장 ‘패배도 그러하다’에는 (32)의 제약조건들이 적용되며 다음과 같이 분석된다.

(34) ‘패배도 그러하다’의 분석



앞의 문장과 같이 명사구 ‘패배도’는 두 번째 문장의 주어이므로 ‘그러하다’의 SUBJ의 값에 ‘패배도’의 ROOT의 값이 주어진다. 그리고 ‘패배도’와 ‘그러하다’는 첫 번째 문장의 주부와 술부에 해당하므로 이 둘은 동일한 SCOPE값을 갖게 된다. 그 다음, 두 번째 문장의 핵어는 동사(구) ‘그러하다’이므로 ‘순간이다’의 ROOT 값이 문장의 ROOT 값으로 올라가게 된다. 그 다음 CON을 보면 HCONS에 HARG의 값 [7]이 LARG의 값 [2]를 가리키도록 되어 있다. 즉, 두 번째 문장의 동사구 ‘그러하다’가 첫 번째 문장의 동사구의 인덱스 [2]를 가리키도록 되어있는 것이다. 그 다음, (31)의 제약조건을 사용하여 두 개의 문장들이 다음과 같이 연결된다.

(35) 두 문장의 등위접속



이 분석을 보면 알 수 있듯이 원시문장의 ROOT의 값(동사구의 인덱스)가 등위접속을 통하여 목표문장의 SOS의 값으로 전달된다. 이때, 두 번째 문장의 핵어는 반드시 elliptical-aux이어야 한다.

전체적인 분석을 보면 첫 번째 문장과 두 번째 문장의 SUBJ와 ROOT의 값이 각각 다르게 되어있다. 즉, 이 두 문장의 각각 별개의 주부와 술부로 구성되어 있는 것이다. 그러나 두 번째 문장의 대용어 '그러하다'가 첫 번째 문장의 '순간이다'를 가리킨다는 사실을 '그러하다'의 CON의 제약조건에 나타나있다. HCONS의 HARG 값이 LARG의 값과 동일하도록 되어 있으므로, 두 번째 문장의 동사구 '그러하다'가 첫 번째 문장의 동사구의 인덱스를 가리킬 수 밖에 없는 것이다.

이렇게 CLLS를 활용한 동사구 생략현상의 분석방법을 한국어에 맞도록 변형하여 사용하게 되면, 한국어의 동사구 대용현상을 효과적으로 설명할 수 있게 된다. 또한 이러한 설명방법은 다음과 같은 두 가지의 사실을 정확하게 포착하고 있다. 첫째는 이 두개의 문장들이 각각 별개의 주부와 술부로 구성되어 있는 것이며, 둘째는 두 번째 문장에 나오는 대용어 '그러하다'가 첫 번째 문장의 동사구를 가리킨다는 사실이다.

5. 결론

본 논문에서는 한국어에서 나타나는 동사구 대용현상을 살펴보고 이를 효과적으로 분석할 수 있는 방법을 찾아보았다.

우리는 먼저 한국어의 동사구 대용현상에 나타나는 여러 가지 특징들을 살펴보았다. 앞 문장의 동사구가 반복되는 경우, 영어에서는 조동사 다음에 나오는 동사구가 생략되는 현상이 나타나지만, 한국어에서는 동사가 생략된 자리에 '그렇다'라는 대용어가 사용되게 된다. 대용어 '그렇다'는 동사뿐 아니라 형용사도 받을 수 있으며, 여러 형태론적 변이형을 가질 수 있다. 또한 '그렇다'의 선행어가 앞 문장이나 주어진 문장내에 존재할 수도 있으나 담화상의 어떤 용언을 지칭할 수도 있음을 알 수 있었다.

그런 다음 우리는 Egg et al. (2001)의 CLLS를 적용하여 동사구 생략현상이 어떻게 분석될 수 있는 지 살펴보았다. Egg and Erg (2001)에 따르면 첫 번째 문장의 정보가 등위접속을 통하여 두 번째 문장으로 전달되며, 두 번째 문장의 조동사에 사용된 제약조건을 통하여 생략된 조동사의 의미가 복원된다.

4절에서는 Egg and Erg (2001)의 분석을 변형하여 한국어 동사구 대용구문의 분석에 적용하여 보았다. 이 분석에 의하면 첫 번째 문장의 동사구에

대한 정보가 등위접속을 통하여 두 번째 문장으로 전달되며, 두 번째 문장의 대응어 ‘그렇다’에 사용된 제약조건을 통하여 치환된 동사구의 의미가 복원하게 된다.

본 논문에서 기술된 방법은 핵어구조도문법을 기반으로 개발되었기 때문에 Kim and Yang (2003)이나 Kim et al. (2011)에 기술된 한국어 자원 문법(Korean Resource Grammar; KRG)에 적용하여 전산적으로 구현이 가능할 것으로 기대된다.

참고문헌

- Copestake, A. 2002. *Implementing Typed Feature Structure Grammars*. Stanford: CSLI.
- Copestake, A., D. Flickinger, C. Pollard, and I. Sag. 2005. Minimal Recursion Semantics: An Introduction. *Research on Language and Computation* 3, 281-332.
- Egg, M. and K. Erk. 2001. A Compositional Account of VP Ellipsis. *Proceedings of the 8th International HPSG Conference*, 162-179.
- Egg, M., A. Koller, and J. Niehren. 2001. The constraint Language over Lambda Structures. *Journal of Logic, Language, and Information* 10, 1-29.
- Egg, M., J. Niehren, P. Ruhrberg, and F. Xu. 1998. Constraints over Lambda-Structures in Semantic Underspecification. *Proceedings of the 17th International Conference on Computational Linguistics and 36th Annual Meeting of the Association for Computational Linguistics (COLING/ACL'98)*, 353-359.
- Erk, K. and J. Niehren: 2000, Parallelism Constraints. In Bachmair, L. (eds.), *Rewriting Techniques and Application*. 110-126.
- Erk, K., A. Koller, and J. Niehren. 2002, Processing Underspecified Semantic Representations in the Constraint Language for Lambda Structures. *Journal on Research in Language and Computation* 1.1, 127-169.
- Johnson, K. 2008. *Topics in Ellipsis*. Cambridge: Cambridge University Press.
- Kim, J. and Sells, P. 2008. *English Syntax: An Introduction*. Stanford, CA: CSLI.
- Kim, J. and Yang, J. 2003. Korean Phrase Structure Grammar and Its Implementations into the LKB system. *Proceedings of the 17th Pacific Asia Conference on Language, Information, and Computation (PACIC-17)*, 88-97.
- Kim, J. and Yang, J. 2006. Coordination Structures in a Typed Feature Structure Grammar: Formalization and Implementation. *Lecture Note in Computer Science* 4319, 194-205.

- Kim, J. and Yang, J. 2011. Symmetric and Asymmetric Properties in Korean Verbal Coordination: A Computational Implementation. *Language and Information* 15.2, 1-21.
- Kim, J., J. Yang, S. Song, and F. Bond. 2011. Deep Processing of Korean and the Development of the Korean Resource Grammar. *Linguistic Research* 28.3, 635-672.
- Pollard, C. and Sag, I. 1994. *Head-driven Phrase Structure Grammar*. Chicago: University of Chicago Press.
- Sag, I. 1976. *Deletion and Logical Form*. Ph.D. Dissertation. MIT.
- Sag, I. and Wasow, T. 1999. *Syntactic Theory: A Formal Introduction*. First Edition. Stanford, CA: CSLI.
- Sag, I., Wasow, T. and Bender, E. 2003. *Syntactic Theory: A Formal Introduction*. Second Edition. Stanford, CA: CSLI.
- Sohn, H. 1999. *The Korean Language*. Cambridge: Cambridge University Press.

이용훈

대전광역시 유성구 대학로 99

충남대학교 영어영문학과

E-mail: ylee@cnu.ac.kr

김중복

서울특별시 동대문구 경희대로 26

경희대학교 영어영문학과

E-mail: jongbokkim@gmail.com

접수일자: 2012. 03. 13

게재결정: 2012. 05. 14

※ 본인이 투고한 논문은 다른 학술지에 게재된 적이 없으며 타인의 논문을 표절하지 않았음을 서약합니다. 추후 중복게재 혹은 표절된 것으로 밝혀질 시에는 논문게재 취소와 일정 기간 논문 제출의 제한 조치를 받게 됨을 인지하고 있습니다.