

# 한국어 대화의 표상구조

이동영  
세종대학교 영어영문학과  
dylee@sejong.ac.kr

## Representation Structure of Korean Dialogue

Dong-Young Lee  
Dept. of English Language and Literature, Sejong University

### 요약

한국어 대화에서는 상황에 따라 주어나 목적어가 대화 속의 발화문에서 완전히 생략되기도 하고, 존대현상이 일어나기도 하고, 존대대명사가 사용되기도 한다. 본 논문은 이러한 특징적인 언어현상이 일어나는 한국어 대화를 처리하고 그것의 표상구조를 만들기 위해서는 대화참석자에 관한 정보, 발화문의 화행에 관한 정보, 대화에 관련된 사람들의 사회적 지위에 있어서의 상대적 순위에 관한 정보, 대화에 나타나는 발화문 사이의 정보흐름에 관한 정보 등을 명시적으로 표시하고 이용하여야 한다고 주장하며, 또한 이러한 상황정보를 한국어대화표상구조에 구현하는 방법을 제안한다. 본 논문에서 한국어대화표상구조의 구축은 담화표상이론(Discourse Representation Theory)과 분할담화표상이론(Segmented Discourse Representation Theory)을 수정, 확대하여 이루어진다.

### 1. 서론

담화(discourse)에서는 문장에 상관없이 서술자와 청중이 고정되어 있지만, 대화(dialogue)에서는 발화문(utterance)마다 화자(speaker)와 청자(addressee)가 바뀌는 것이 일반적이다. 즉, 대화가 담화보다 더 역동적(dynamic)이다. 또한, 한국어 대화에는 주어나 목적어가 빈번히 생략되고 존대현상이 항상 관련되어 있고 존대대명사가 쓰이기도 한다. 이러한 특징을 지닌 한국어를 처리하고 그것의 표상구조를 적절하게 구축하기 위해서는 대화에 관련된 상황정보를 이용하여야 한다.

담화의 표상구조와 관련된 이론으로 담화표상이론(Discourse Representation Theory)[1,2]과 분할담화표상이론(Segmented Discourse Representation Theory)[3]이 있지만, 이들 이론은 원래 담화를 다루기 위해 주장되었기 때문에 대화, 더군다나 한국어 대화에 그대로 적용시킬 수 없는 한계성이 있다. 본 논문은 한국어 대화를 적절히 표상하기 위해서 어떠한 정보가 어떠한 구조 속에 구체적으로 포함되어야 하는지를 보여준다.

### 2. 담화표상에 관련된 이론

대화를 살펴보기 전에 담화를 다루고 그것의 표상구조를 구축하는 이론들을 알아보고자 한다.

#### 2.1 담화표상이론

담화표상이론[1,2]은 문장의 연속인 담화를 해석하여 그 결과를 담화표상구조(Discourse Representation Structure)로 나타낸다. 이 담화표상구조는 두 가지 구성요소(즉, 담화지시체(discourse referent)와 담화표상구조조건(DRS-condition))로 이루어져 있다. 예를 들면, 표 1에 보여진 영어담화는 그림 1에 보여진 구조로 표상되어진다.

표 1 영어담화의 예

- |   |
|---|
| a. A violinist played the sonata.<br>b. The audience applauded him.<br>c. He was happy. |
|---|

v s d
violinist(v)
sonata(s)
played(v,s)
audience(d)
applauded(d,v)
happy(v)

그림 1 담화표상구조의 예

그림 1에 나타난 구조는 3개의 담화지시체(즉,  $v, s, d$ )와 6개의 담화표상구조조건(한 예는  $played(v, s)$ 임)으로 이루어져 있다. 담화표상이론에서 대명사의 해결(resolution)은 대명사에 관련된 담화표상구조조건에서 대명사의 지시체가 접근가능(accessible)면 성공적으로 이루어진다. 예를 들면, 표 1의 문장 c에 나타난 대명사 *he*와 관련된 담화표상구조조건은  $happy(v)$ 인데, 지시체  $v$ 가 이 조건으로부터 접근가능하기 때문에 (즉, 여기서는 동일한 담화표상구조 내에 있기 때문에) 이 대명사는 해결될 수 있다.

#### 2.2 분할담화표상이론

분할담화표상이론[3]은 담화를 분할담화표상구조(Segmented Discourse Representation Structure)로 나타냄으로써, 담화표상이론에서의 달리 담화에 일어나는 각 문장을 독립된 표상구조로 나타낸다. 분할담화표상구조는 크게 두 가지 구성요소로 이루어져 있다. 하나는 담화표상구조(DRS)나 분할담화표상구조(SDRS)이고 다른 하나는 이들 구조에 대한 분할담화표상구조조건(SDRS-condition)이다. 이 조건의 형태는  $\delta(RS1, RS2)$ (여기서  $\delta$ 는 문장사이의 담화관계(discourse relation)[4,5]를 가리키고,  $RS1$ 이나  $RS2$ 는 DRS나 SDRS를 가리킴)이다. 예를 들면, 표 1에 이미 나타난 영어담화는 다음의 그림 2에 나타난 분할담화표상구조로 표상된다. 그림 2에서 DRS인  $S11, S12, S13$ 은 표 1에 나타난 문장 *a, b, c*를 각각 표상한다(represent). 따라서, 분할담화표상이론에서는 담화 속의 문장들이 별개의 표상구조로 분할되어 나타내어진다.

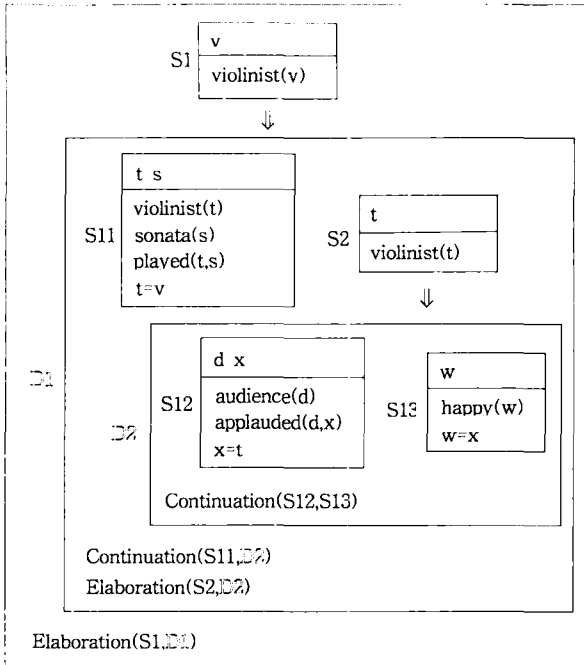


그림 2 분할담화표상구조의 예

이 절에서 담화표상에 관한 두 가지 이론을 살펴보았다. 이 이론들이 한국어대화에 그대로 적용되어 한국어대화표상이 적절하게 이루어질 수 있는지 논의하기 전에 한국어대화의 특징을 다음 절에서 알아보기로 한다.

3. 한국어대화의 특징

한국어대화의 특징을 표 2에 나타난 대화를 바탕으로 하여 고찰하기로 한다.

표 2 한국어대화의 예

- a. 최 과장이 어제 조감도를 완성하였어.  
(화자: 김 부장, 청자: 남 부장)
- b. 홍 과장이 그 조감도를 검토하였어.  
(화자: 남 부장, 청자: 김 부장)
- c. 박 과장, ≡ 남 부장에게 그 조감도를 전했니?  
(화자: 김 부장, 청자: 박 과장)
- d. 제가 오늘 오전에 그분께 ≡ 보여 드렸어요.  
(화자: 박 과장, 청자: 김 부장)

첫째로, 한국어대화는 존대현상이 항상 관련되어 있다. 존대현상의 발생은 대화에 관련된 사람들의 사회적 지위에 있어서의 상대적 순위에 의해 제약을 받고, 대화에서 존대는 존대형태소(즉, 존대접미사 남, 시와 존대격조사 께서, 께, 그리고 존대동사어미 (어)요, 습니(다/까) 등)를 통하여 실현된다. 대화에서 화자가 존대형태소를 써서 존디하는 사람의 사회적 지위는 화자보다 높다. 예를 들면, 표 2에 나타난 발화문 d에서 존대동사어미 어요가 쓰였기 때문에 청자의 사회적 지위는 화자보다 높다는 정보의 추론이 가능하다. 이와 반대로, 발화문 c에서는 존대동사어미가 쓰이지 않았기 때문에 그 발화문의 화자의 사회적 지위는 청자와 같거나 더 높다는 정보를 추론할 수 있다.

둘째로, 한국어대화에서는 대화문맥상으로 파악이 가능하면

주어나 목적어가 완전히 생략된다. 예를 들면, 표 2에 나타난 발화문 d에서 목적어(기호 ≡ 로 표시되어 있음)가 생략되어 있다. 이 생략된 목적어가 무엇인지 알기 위해서는 발화문 d보다 앞선 발화문들에서 얻은 정보를 이용해야만 한다. 심지어 어떤 경우에는 대화의 첫 번째 발화문에서도 주어나 목적어가 생략되기도 한다. 이런 경우에는 그 발화문의 화행(speech act)[6-9]에 의존하여 생략된 부분을 되찾을 수 있다. 예를 들면, 대화 첫 번째 발화문의 화행이 허락요청(ask-permission)이고 주어가 빠져 있으면, 이 빠진 주어는 화자를 가리킨다.

마지막으로, 존대현상과 관련하여 대화에서 존대대명사(발화문 d에 나타난 그분이 그 예임)가 쓰이기도 한다. 이 존대대명사의 지시체는 존대대명사가 쓰인 발화문의 화자보다 사회적 지위가 높아야 한다. 따라서, 이러한 조건을 만족시키는 어떤 지시체를 앞선 발화문들에서 얻은 사회적 지위에 관한 정보를 이용하여 찾아내는 것이 곧 존대대명사를 해결하는 것이 된다.

4. 담화표상에 관련된 이론의 한국어대화 적용에 있어서의 한계성

이 절에서는 2절에서 논의한 이론들이 한국어대화에 일어나는 특징적인 현상들을 다루고 대화표상구조를 구축하는 데 적용될 때 어떠한 한계성이 있는지 알아보기로 한다.

4.1 담화표상이론의 한계성

첫째로, 화자와 청자 같은 대화참석자에 대한 정보는 한국어대화에 일어나는 존대현상을 설명하고 생략된 부분을 찾는 데 필요하나, 담화표상이론에서는 이러한 정보를 전혀 고려하지 않고 담화표상구조에도 포함시키지 않고 있다.

둘째로, 한국어대화에 나타나는 특징적인 현상을 설명하기 위해서는 어느 구조가 어느 발화문을 표상하는지 명확히 구분되어 있어야 하는 데, 담화표상구조에서는 그림 1에 보여진 바와 같이 그러한 구분이 되어 있지 않다.

마지막으로, 대화의 첫 번째 발화문에서 생략된 주어나 목적어는 그 발화문의 화행에 의존하여야만 되찾을 수 있는데, 화행에 관한 정보가 담화표상이론에서는 고려되지 않고 있다.

4.2 분할담화표상이론의 한계성

분할담화표상이론에서는 그림 2에 나타난 바와 같이 어느 구조가 어느 문장을 표상하는지 구분이 되어 있는 점에서는 담화표상이론보다는 한국어대화표상에 유리하나 여전히 아래와 같은 한계성을 지니고 있다.

첫째로, 분할담화표상이론이 담화표상이론에서와 같이 대화참석자에 대한 정보를 고려하고 있지 않아서 한국어대화표상에 그대로 적용시킬 수 없다.

둘째로, 발화문의 화행에 관한 정보를 고려하지 않아 한국어대화의 첫 번째 발화문에서 생략된 부분을 찾아낼 수 없다.

마지막으로, 분할담화표상이론에서 쓰이는 담화관계(continuation, elaboration 등)의 개념은 한국어대화표상에 아무런 역할을 하지 못한다.

5. 한국어대화의 표상구조

앞 절에서 살펴본 바와 같이, 담화표상이론이나 분할담화표상이론은 그대로 한국어대화표상에 적용될 수 없는 한계성을 지니고 있다. 이 절에서는 담화표상이론의 담화표상구조와 분할담화표상이론의 분할 개념을 이용하여 한국어대화의 특징적인 현상을 설명할 수 있는 한국어대화표상구조를 제안하고자 한다.

한국어대화에 일어나는 존대현상을 설명하기 위해서, 그리고 존대대명사를 해결하기 위해서는 대화참석자에 대한 정보와 대화에 관련된 사람들의 사회적 지위에 있어서의 상대적 순위에 관한 정보가 대화표상구조에 명시되어야 한다.

한국어대화에서 생략된 주어나 목적어를 되찾기 위해서는 발화문의 화행, 앞선 발화문들에 관한 정보의 후속 발화문으로의 흐름 등이 대화표상구조에 명시되어야 한다.

이러한 상황적 정보가 명시되고 정보의 흐름이 발생하는 한국어대화표상구조의 골격은 아래에 제시되어 있다.

발화내용1 발화내용2 ... 발화내용(n-1) 발화내용(n) 화행1(화자1,청자1,발화내용1:DRS1) 화행2(화자2,청자2,발화내용2:DRS2) ... 화행(n-1)(화자(n-1),청자(n-1),발화내용(n-1):DRS(n-1)) 화행(n)(화자(n),청자(n),발화내용(n):DRS(n))
--

그림 3 n개의 발화문으로 구성된 한국어대화의 표상구조의 골격 한국어대화에 나타나는 각 발화문은 컴퓨터 프로그래밍언어의 하나인 프롤로그(Prolog)[10-14]의 프레디카트(predicate)의 형태로 표상된다. 발화문의 화행은 프레디카트의 펑터(funcor)로 나타내고, 발화문의 화자와 청자는 프레디카트의 첫 번째와 두 번째의 논항(argument)으로 각각 나타낸다. 발화문의 내용은 담화표상구조로 나타내고 발화내용이라는 표지를 붙인다. 이 표지가 대화표상구조에서 대화지시체로 사용됨으로써 앞선 발화문의 내용이 뒤에 따르는 발화문의 처리에 활용될 수 있게 된다. 즉, 정보의 흐름이 일어나게 된다. 대화표상구조의 예를 들면, 표 2에 나타난 한국어대화의 전산적 표상구조는 아래 그림과 같다.

[ [발화내용1,발화내용2,발화내용3,발화내용4], [문의(김부장,남부장,발화내용1( [K,N,C,B,T], [ 이름(K,김부장),이름(N,남부장), 이름(C,최과장),조감도(B), 어제(T),완성하였다(C,B,T), 동등_상위(K,C),동등_상위(K,N) ])], 진술(남부장,김부장,발화내용2( [H], [ 이름(H,홍과장),검토하였다(H,B), 동등_상위(N,H),동등_상위(N,K) ])], 문의(김부장,박과장,발화내용3( [P], [ 이름(P,박과장),전했다(P,N,B), 동등_상위(K,P),동등_상위(K,N) ])], 진술(박과장,김부장,발화내용4( [M], [ 오늘오전(M),보여드렸다(P,N,B,M), 상위(K,P),상위(N,P) ])] ] ]
---

그림 4 한국어대화의 전산적 표상구조의 예

그림 4에 명시된 바와 같이, 한국어대화의 각 발화문의 내용을 나타내는 담화표상구조 속에 그 발화문에 관련된 사람들의 사회적 지위에 있어서의 상대적 순위에 관한 정보(예를 들면, 상위(K,P)는 지시체 K의 사회적 지위는 지시체 P보다 높다는 것을 의미함)가 들어있다. 또한, 대화전체에 대한 표상구조와 대화에 나타나는 각 발화문에 대한 표상구조, 그리고 대화지시체의 집합 등은 모두 프롤로그의 리스트(list) 형태로 표시되어 있다.

6. 결론

본 논문은 한국어대화에 일어나는 특이한 현상들(주어나 목적어의 생략, 존대현상의 발생, 존대대명사의 사용 등)을 설명하고 한국어대화를 알맞게 표상하기 위해서는 대화에 관련된

상황정보들(대화참석자에 관한 정보, 화행에 관한 정보, 사회적 지위에 관한 정보)을 이용하고 이러한 정보들 표상구조에 명확하게 표시하여야 하며 정보흐름이 대화의 발화문 표상 사이에 이루어지도록 해야 한다는 것을 주장하였다. 나아가서, 본 논문은 이러한 주장이 수용된 한국어대화표상구조의 기본 뼈대를 제시하였고, 프롤로그를 사용하여 전산적으로 구축된 한국어대화표상구조의 구체적인 예를 보여주었다.

이러한 한국어대화표상구조는 한국어대화에서와 비슷한 언어 현상이 일어나는 일본어대화(발화문의 구성요소가 생략되고 존대현상이 일어남)의 표상에도 큰 수정을 가하지 않고 적용될 수 있으리라 생각된다. 나아가서 다른 자연언어의 대화를 표상할 때에도 그 대화에서 일어나는 언어현상에 따라 표상구조에 포함시킬 정보를 달리하면 되므로, 본 논문에서 제시한 한국어대화표상구조는 다른 자연언어의 대화표상구조에 대한 하나의 모형(template)이 된다고 할 수 있다.

참고문헌

- [1] Hans Kamp. "A Theory of Truth and Semantic Representation." In Jeroen Groenendijk, Theo Janssen, and Martin Stokhof, editors, *Formal Methods in the Study of Language*, pp. 277-322, Mathematical Centre, Amsterdam, 1981.
- [2] Hans Kamp and Uwe Reyle. *From Discourse to Logic: Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. Kluwer Academic Publishers, Dordrecht, 1993.
- [3] Nicholas Asher. *Reference to Abstract Objects in Discourse*. Kluwer Academic Publishers, Dordrecht, 1993.
- [4] Alex Lascarides and Nicholas Asher. "Discourse Relations and Defeasible Knowledge." In *Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics (ACL-91)*, pp. 55-63, Berkeley, 1991.
- [5] Alex Lascarides, Nicholas Asher, and Jon Oberlander. "Inferring Discourse Relations in Context." In *Proceedings of the 30th Annual Meeting of the Association for Computational Linguistics (ACL-92)*, pp. 1-8, Newark, 1992.
- [6] Edward Finegan. *Language: Its Structure and Use*. Third Edition. Harcourt Brace College Publishers, New York, 1999.
- [7] Stephen C. Levinson. *Pragmatics*. Cambridge University Press, Cambridge, 1983.
- [8] John R. Searle. *Speech Acts: an Essay in the Philosophy of Language*. Cambridge University Press, Cambridge, 1969.
- [9] John R. Searle. *Expression and Meaning: Studies in the Theory of Speech Acts*. Cambridge University Press, Cambridge, 1979.
- [10] Ivan Bratko. *Prolog Programming for Artificial Intelligence*. Third Edition, Addison-Wesley Publishing Company, Reading, 2001.
- [11] William F. Clocksin and Christopher S. Mellish. *Programming in Prolog*. Third, Revised and Extended Edition. Springer-Verlag, New York, 1987.
- [12] Clive Matthews. *An Introduction to Natural Language Processing through Prolog*. Addison-Wesley Longman Ltd., Boston, 1998.
- [13] Richard A. O'Keefe. *The Craft of Prolog*. The MIT Press, Cambridge, 1990.
- [14] Leon Sterling and Ehud Shapiro. *The Art of Prolog*. Second Edition. The MIT Press, Cambridge, 1994.