

관계절 두어(head noun)의 정보와 조사의 정보를 이용한 한국어 문장의 재분석과정¹

권유안* · 김영진** · 남기춘*
(*고려대학교 심리학과, **아주대학교 심리학과)

권유안 · 김영진 · 남기춘. 관계절 두어(head noun)의 정보와 조사의 정보를 이용한 한국어 문장의 재분석과정. 『언어청각장애연구』, 2003, 제8권, 제1호, 41-65. 이 연구는 한국어 문장의 관계절 재분석에 관한 것이다. 한국어는 핵-말언어(head final language)로 그 특징이 영어와 다르다. 따라서 한국어는 특징적인 문장처리 과정이 있을 수 있다. 본 연구는 중의적인 관계절 명사구의 구조적 해결과정을 적용하여 한국어 문장의 중의성 해결과정을 추론하였다. 이를 위해 영어를 바탕으로 형성된 수정모형과 진단모형을 각각 한국어 문장에 적용하여 한국어 문장의 중의성 해소에 적절한 문장모형을 찾고 더 나아가 한국어 문장의 중의성 해소에 보다 적절한 모형을 수정 및 제시하고자 하였다. 실험 1은 설문지 연구를 통해 구조중의적 문장과 관계절 두어의 논제적 역할(thematic role) 및 중의적 문장간에 난이도를 평정하였다. 실험 1의 결과, 구조적 중의성이 높은 문장이 보다 어려운 문장으로 평정되었지만 두어의 논제적 역할중의성은 조건간 차이가 발생하지 않았다. 실험 2는 실시간 과제인 자기조절읽기과제(self-paced reading task)를 통해 구조적, 논제적 역할중의성이 탈중의지역(disambiguation region)에서 반영되는지를 측정하였다. 실험 2의 결과, 논제적 역할의 중의성이 탈중의지역인 관계절 두어에서 나타났다. 실험 3은 주격조사 '-가/-이'와 주제표지 '-는/-은'이 재분석에 미치는 영향을 측정하였다. 주격조사의 경우 주절의 주어 및 관계절의 주어 모두 가능한 기능을 가지고 있지만 주제표지의 경우 관계절의 주어로 사용되지 않는다. 따라서 구조적 중의성이 다소 적은 주제표지가 있는 문장의 탈중의지역에서 더 짧은 읽기시간이 나타날 것으로 예측되었다. 그러나 실험 3의 결과, 주제표지가 있는 문장의 탈중의지역이 주격조사가 있는 문장의 그것 보다 읽기시간이 더 짧지 않았다. 그러나 초기분석과정에서 주제표지의 영향은 나타났다. 본 연구를 통해 한국어 문장의 재분석과정은 어휘적 정보를 재분석 과정에 적용한다고 가정하는 진단모형이 보다 적절하며 또한 문장에 사용된 어휘적 정보(예를 들어 명사의 논제적 역할)와 문장의 구조적 정보가 상호작용하는 모형의 가능성을 제시하였다.

핵심어: 문장 재분석과정, 수정모형, 진단모형, 구조적 정보, 논제적 역할 정보, 주격조사, 주제표지

I. 서론

한 문장의 의미를 정확히 파악하기 위해선 제시된 문장안의 단어들 간의 관계성을 정

¹ 본 연구는 한국과학기술부 특정연구개발사업(M1-0107-00-0013) 지원으로 수행되었음.

확히 파악해야 한다. 문장안의 단어들이 서로 관련이 없다면 비문이 된다. 그러나 단어들이 서로 관련되어 있지만 이해하는 자가 그 관련성을 정확히 파악하지 못해 정확한 의미를 파악하지 못하는 경우도 있을 수 있다. 후자의 경우를 통사분석실패라고 하며 이것은 문장의 재분석과정에 관련이 있다고 할 수 있다. 과거부터 문장의 초기분석(initial parsing)에 대한 이론적 검증은 많이 수행되어 왔지만 정작 핵심이라 할 수 있는 문장의 재분석(reanalysis, revision, re-parsing 등)은 그 중요성에 비해 연구가 많지 않다. 설령 연구가 있더라도 대부분은 영어를 위주로 수행되어 왔다. 영어위주의 연구는 자칫 영어위주의 모형개발로 언어처리의 일반성이라는 대주제에 역행하는 결과를 낼 수 있다. 따라서 영어와는 그 성격이 다른 한국어, 일본어 등을 소재로 그 처리의 특징을 밝히는 것은 언어처리의 일반성을 밝히는 데 중요하다(Inoue & Fodor, 1995; Yamasita, Stowe & Nakayama, 1993).

본 논문은 한국어 관계절 문장의 재분석과정을 통해서 영어에 적용되는 재분석모형들이 한국어에도 적용될 수 있는지, 만약 적용된다면 어떤 방식으로 적용될 수 있는지, 반면 적용되지 않는다면 한국어 특유의 처리방식이 있는지를 찾아 볼 것이다. 본론에 들어가기 앞서 한국어 문장의 특징을 잠시 언급하고 재분석모형들을 논할 것이다.

한국어 문장의 경우 영어와 다른 다음과 같은 특징이 있다. 핵-말(末)언어(head final language), 좌연접언어(left-branching), 동사 이전의 논항들의 어순의 상대적 자유로움, 조사의 존재, 마지막으로 공범주(empty category)의 존재 등이 있다(Kim, 1999; Mazuka, 1991; Yamasita, 1995).

핵-말언어란 핵이 보어를 후행하는 언어를 의미하며(박경자 외, 2001), 좌연접언어란 왼쪽에서부터 오른쪽으로 단어를 순차적으로 결합해가는 방식으로 수형도 상에서 가지가 좌측으로 더해져 간다(이기용, 1999). 또한 영어의 어순에 비해 상대적으로 동사 이전 논항의 어순이 자유롭다는 것은 동사 이전에 논항들의 순서가 바뀌어도 문장 전체의 의미는 변하지 않는다는 것을 의미한다. 예를 들면 (1)-(4) 등과 같은 경우 모두 비문이 아니며 의미 역시 동일하다.

- (1) 순이가 철수를 영희에게 소개한다.
- (2) 철수를 순이가 영희에게 소개한다.
- (3) 영희에게 순이가 철수를 소개한다.
- (4) 철수를 영희에게 순이가 소개한다.

하지만 동사가 논항 앞에 오는 경우는 문법적으로 비문이다(화용론적 차원에서는 가능

하다). 그리고 조사의 존재가 있는데 문장에서 조사의 역할은 문장의 동사 이전에 격을 할당하는 것으로 문장처리에서 그 중요성은 이미 Kim (1999)에서 언급된 바 있다. Kim (1999)에 따르면 동일한 구조의 문장(관계절 문장)임에도 불구하고 주격조사 ‘-가/-이’와 한정사 ‘-은/-는’에 따라 관계절에서 다른 읽기시간이 보였는데 이는 동사 이전에 나타나는 조사가 문장의 구조를 어느 정도 예측하게 하는 역할이 있음을 시사했다. 마지막으로 공범주의 존재로 공범주는 Chomsky (1981)의 지배-결속이론에서 음성적으로 아무것도 없는 일련의 공(null) 요소를 의미한다. 즉 영어의 wh-movement의 흔적(trace)은 자신이 이동한 위치를 남기지만 한국어의 경우 그 흔적을 남기지 않고 공범주라는 표면적으로 제시되지 않는 요소를 남긴다. 다음 예문에서 ‘e_i’라는 기호가 공범주의 위치를 의미하며 ‘i’는 참조하는 논항을 의미한다.

(5) 순이가 (e_i 철수를 때린) 영수_i에게 전화했다.

위 예문 (5)에서 괄호 안은 주어의 위치가 공범주로 있는 관계절이다. 영어의 경우 이와 같은 경우 반드시 e_i에 그 흔적을 남겨야 한다. 한국어 문장의 특징은 영어의 그것과 매우 다르다. 영어의 경우 핵-선행(先行)언어, 우연접언어(right-branching language), 동사 전의 논항들의 어순의 비자유로움, 조사의 부재, 공범주가 아닌 어떤 흔적(trace)을 남겨야 하는 특징 등 한국어의 특징과 정반대되는 특징을 가지고 있다. 이에 영어와는 다른 처리방식이 있을 수도 있다는 주장은 우리 한국어와 유사한 특징을 갖고 있는 일본어에서 제기되었다. 일본어를 대상으로 한 실험에서 일본어 역시 영어의 처리방식과 다르지 않다는 주장(Inoue & Fodor, 1994; Mazuka, 1991; Yamasita, 1995)과 영어와는 다를 수 있다는 주장(Hirose & Inoue, 1998)이 양립하고 있다.

여러 문헌들은 이러한 주장을 위해 다양한 중의적 문장을 사용하지만 그 중에서 관계절 문장을 통해 각각의 주장을 하고 있다. 특히 재분석에 관련된 문헌의 경우 관계절이 주를 이루고 있다. 그 이유로는 중의적 문장을 만들기 용이하고 각기 다른 언어간에 관계절이라는 동일한 소재로 서로 다른 통사적 규칙이 적용되므로 서로 그 차이점이나 유사점을 비교할 수 있기 때문이다. 본 논문에서도 역시 관계절 문장을 통해 문장의 재분석과정을 논할 것이다. 관계절의 처리방식을 논하기 전에 문장의 중의성을 잠시 설명하자. 문장의 중의성은 문장의 전체적인 의미에는 관련 없지만 문장처리 중간에 지역적으로 발생하는 지역적인 중의성(local ambiguity)이 있으며 문장의 의미를 전체적으로 분석하기 힘들게 하는 전체적인 중의성(global ambiguity)이 있다(Pritchett, 1992). 이 중의성은 흔히 오인현상(garden path)으로 설명하는데 특히 통사적 중의성이 논의의 쟁점이다. 통사적 중의성이란 정확한 통사분석을

포함한 여러 개의 통사분석의 대안들이 공존할 경우를 의미한다. 이 경우 오인현상을 해결하기 위해 재분석을 실시하게 되는데 재분석을 통해 오인현상을 해결하는 부분을 재분석지역 또는 탈중의지역(disambiguity region)이라고 한다. 이 탈중의지역에서 일반적으로 인지적 처리가 필요하게 되며 이 인지적 비용은 읽기시간이나 안구고정시간 등으로 표출된다.

문장실험 연구에서 중의적 지역과 중의적 부분이 해결되는 탈중의지역이 가장 중요한 부분으로 그 읽기시간 차이로 처리의 과정을 추론하게 된다. 본 논문 역시 중의적 문장과 통제 문장을 바탕으로 문장처리의 과정을 추론할 것이다.

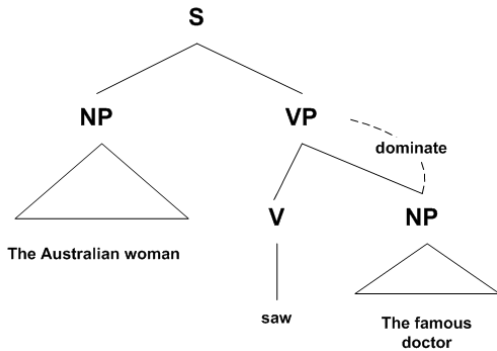
문장재분석모형은 초기분석모형과 매우 밀접한 관계를 갖는다. 왜냐하면 초기분석모형에서 형성한 구조(통사적 구조)가 잘못 되었음을 알고 다시 분석을 하는 과정이 재분석과정이기 때문이다(Sturt & Crocker, 1999). 만약 초기분석모형이 순수하게 문장의 구조적 정보를 바탕으로 가장 단순한 구조만을 선택하여 분석을 진행했다면 재분석의 첫 번째 근거 역시 문장의 구조적 정보가 될 것이다. 예를 들면 최소수정모형(minimal revision model)이 이에 해당한다(Frazier & Clifton, 1998). 반대로 초기분석에서 단어의 구조적, 의미적 정보를 모두 사용한다면 재분석에 있어서 모두의 정보를 사용할 것이다. 이에 대한 모형은 제약-바탕이론(constraint-based models)에 해당하는 모형들이다(Trueswell & Tanenhaus, 1994).

또한 재분석모형은 수정모형(revision model)과 진단모형(diagnostic model)으로 나눌 수 있다(김영진, 2001; Inoue & Fodor, 1995). 수정모형의 경우 재분석의 근거는 문장의 구조적 정보이고 인지적 처리의 어려움은 문장의 구조를 수정하는데, 어떤 방식으로 수정하는가에 따라 달라진다. Gorrell (1995)의 경우 재분석시 반드시 초기분석에서 형성한 지배, 선행관계가 지켜져야만 처리의 부담이 없다고 하였다. 그는 입력되는 구가 이미 형성된 구에 추가되거나 구가 인하(lowering)되는 경우 초기분석 구조의 지배, 선행의 관계가 계속 유지되므로 처리의 부담이 없지만 반대로 인상(raising)되는 경우 지배, 선행의 관계가 깨짐으로 처리의 부담이 발생한다고 하였다.

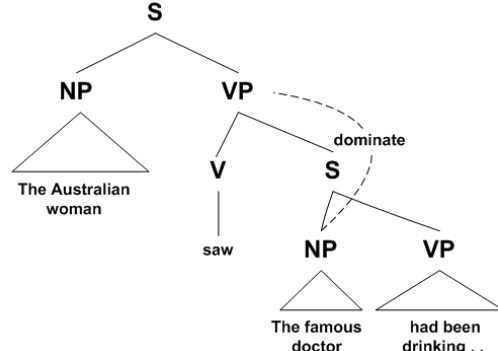
(6) The Australian woman saw the famous doctor had been drinking quite a lot.

위 예문 (6)의 구조는 <그림 - 1>과 <그림 - 2>와 같은 수형도로 나타낼 수 있다. 예문 (6)의 'the famous doctor'는 중의적인 구로 두 가지의 재분석 대안이 가능하다. 첫째는 앞선 'saw'의 목적어, 둘째는 새로 나올 수 있는 종속절의 주어이다. 그러나 초기분석에서 보다 간단한, 또는 보다 선호되는 구조가 선택되는데 그것이 <그림 - 1>의 구조이다. 이 <그림 - 1>의 구조에서 동사구(VP)는 명사구(NP)를 지배(dominate)하고 있다. 예문 (6)에서 'had'는

관계절 두어(head noun)의 정보와 조사의 정보를 이용한 한국어 문장의 재분석과정



<그림 - 1> 예문 (6)의 초기분석 수형도

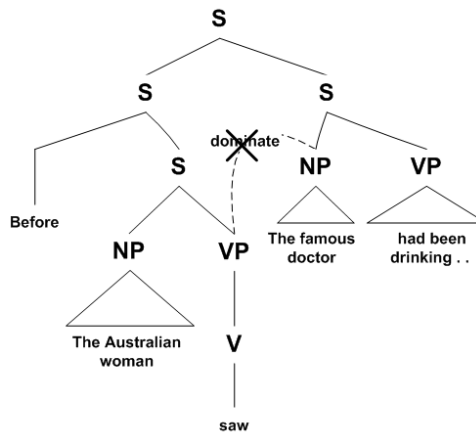


<그림 - 2> 예문 (6)의 재분석 수형도

‘the famous doctor’는 ‘saw’의 목적어가 아니라 종속절의 주어라는 증거가 되는 중의성해소 지역으로 재분석이 일어나는 지역이다. 재분석 결과는 <그림 - 2>와 같다. <그림 - 2>의 경우 Gorrell (1995)이 상정한 지배관계가 유지됨을 알 수 있다. 이러한 경우 처리의 부담이 발생하지 않는 것이다. 그러나 예문 (7)의 경우에는 처리의 부담이 발생한다.

(7) Before the woman saw the famous doctor had been drinking quite a lot.

예문 (7)의 초기분석과정은 <그림 - 1>과 거의 비슷하다. 그러나 재분석과정은 <그림 - 3>과 같을 것이다.



<그림 - 3> 예문 (7)의 재분석 수형도

<그림 - 3>에서는 초기분석에서 형성된 동사구와 명사구 간의 지배관계가 재분석시 유지되지 못하는 것을 알 수 있다. 이와 같은 경우에 재분석의 어려움(오인현상)이 발생한다. Suh (1994)는 지배와 선행관계의 유지라는 두 가지의 원칙 중 지배관계가 더 중요한 요인임을 주장한바 있다.

진단모형(diagnostic model)에서도 수정모형에서처럼 재분석이 동일하게 일어나지만 어떤 재분석이 더 명확하게 분석을 도출하는 정도에 따라 처리의 부담이 변한다. 한국어 예문(김영진, 2001에서 인용) (8)과 (9)를 살펴보고자 한다.

- (8) 그 운전사가 청소부를 설득한 건축가를 비판했다.
- (9) 그 운전사는 청소부를 설득한 건축가를 비판했다.

예문 (8)와 (9)는 구조적으로 동일한 문장이다. 그러나 탈중의지역인 ‘건축가를’을 이해하는 데 소요되는 시간인 안구고정시간에서 예문 (9)의 경우가 예문 (8)에서보다 더 짧았다 (김영진, 2001). 두 문장이 구조적으로 동일하나 예문 (8)은 ‘운전사가’로 주절의 주어도 될 수 있고 종속절의 주어도 될 수 있는 주격조사 ‘가’가 쓰이지만 예문 (9)의 경우 주절의 주어만 가능한 ‘는’이라는 주제표지(topic marker)가 사용되었다. 이 정보는 ‘운전사는’을 주절의 주어로만 가정하게 하게 한다. 그러므로 재분석의 시점(탈중의지역인 ‘건축가를’)에서 주절의 주어만이 가능하다는 정보를 활용하게 되어 재분석이 좀더 쉬워진다.

반대로 예문 (8)의 경우 ‘운전사가’는 주절의 주어, 종속절의 주어 모두 가능한 상태로 유지되다가 ‘건축가를’에서 ‘운전사가’를 주절의 주어로 재분석해야 한다. 그러므로 예문 (8)의 ‘건축가를’에서 보다 긴 안구고정시간이 발생했다고 볼 수 있다.

수정모형은 위와 같이 조사(또는 격표지)의 차이로 인한 처리의 어려움을 예측하지 못한다. 왜냐하면 초기분석에서 최소부착의 원리를 적용하여 두 예문 모두 복합명사절(‘그 운전사가(는) 청소부를 설득한’)을 형성하기 때문이다. 하지만 진단모형의 경우 주제표지의 정보가 재분석의 방향을 제시해주므로 ‘그 운전사는’이 종속절의 주어가 아니라 주절의 주어라는 최종적인 재분석 결과를 도출하게 돕는다. 따라서 주절의 동사 전에 지역적인 중의성이 해소되면서 문장분석이 이루어지는 한국어에 보다 적합한 모형이 진단모형이라고 추론할 수 있다. 본 연구는 이 주장이 관계절 재분석에서도 마찬가지로 적용가능한 지를 확인해 보았다. 본 논문의 조건문장은 Hirose & Inoue (1998)에서 사용한 일본어 문장을 수정하여 사용하였다. 이들의 실험을 반복한 이유는 한국어 문장의 경우 관계절 동사에 ‘-는’이라는 관계절화어미가 표면적으로 제시되기 때문에 그들의 결과와 다를 수 있기 때문이고 또한 보다 정교한

실험을 통해 다시 한번 진단모형의 타당성을 살펴보기 위함이다.

II. 실험 1

1. 설문지 연구

우선 Hirose & Inoue (1998)의 일본어 문장을 수정한 한국어 문장을 설문지로 작성하여 실제적으로 그들의 주장과 일치하는 지를 검증하였다. 굳이 설문지를 통해 일차적인 검증을 한 이유는 Hirose & Chun (1998)에서 일부 한국인 화자가 보다 어려울 것이라고 예측한 문장에 대해 더 쉽다고 평정했다는 언급이 있었기 때문이다. 또한 일본어와 한국어의 차이로 인해 그들이 예측한 주장이 없을 수도 있다는 걱정 때문이었다. 실험에서 사용된 문장은 다음과 같다.

Null: 영희가 아이를 교차로에서 본 사실이 알려졌다.

Single: 영희가 아이를 교차로에서 본 소녀에게 전화했다.

Double-D: 영희가 아이를 교차로에서 본 소녀에게 말했다.

Double-E: 영희가 아이를 광장에서 본 그네에 태웠다.

첫 번째 문장조건은 초기분석의 최소부착의 원리에 의해 복합명사절로 ‘본’ 이전의 논항들이 모두 하나의 절로 분석된다. 그리고 ‘사실이’에서 재분석이 일어나지 않고 마지막 어절인 ‘알려졌다’로 이어진다. 왜냐하면 복합명사절 구조가 ‘사실이’가 읽힘으로 해서 아무런 구조적 변형 없이 구문분석이 되기 때문이다. 두 번째 문장조건의 경우 초기분석은 첫 번째 조건과 동일하다. 그러나 관계절의 두어(head noun)인 ‘소녀에게’에서 재분석이 일어나야 올바른 구문분석이 된다. 즉 하나의 구조로 되어 있던 초기분석 구조(‘영희가 아이를 교차로에서 본’)에서 ‘영희가’가 빠져나와 주절로 이동해야만 적절한 구문분석이 된다. 이때 복합명사절에서 하나의 논항(‘영희가’)이 빠져나오기 때문에 조건명칭을 ‘Single’이라고 하였다. 반대로 첫 번째 조건문장은 아무런 논항도 빠져 나오지 않으므로 ‘Null’이라고 명명하였다. ‘Single’조건의 마지막 어절인 동사 ‘전화했다’에서 이미 재분석을 통해 ‘영희가’가 주절로 나왔기 때문에 ‘영희가 소녀에게 전화했다’로 쉽게 분석 가능하다. 따라서 어떤 재분석도 필요하지 않다. 세 번째 조건문장의 경우 두어 ‘소녀에게’에서 ‘Single’ 조건과 마찬가지로의 구문 분

석과정이 가정된다. 그러나 마지막 어절인 ‘맡겼다’의 경우 다른 구문분석과정이 필요하다. ‘맡겼다’의 경우 ‘영희가 소녀에게 맡겼다’도 가능하지만 맡긴 대상이 있는 문장이 보다 자연스럽다. 따라서 ‘영희가 아이를 소녀에게 맡겼다’로 재분석해야 한다. 이 재분석을 하기 위해 선 관계절 안에 있던 ‘아이를’을 다시 주절로 옮겨야 한다. 그러므로 마지막 어절에서 또 한번의 재분석이 필요할 것으로 예측된다. 세 번째 조건문장의 경우 총 두 개의 논항이 주절로 이동됨으로 ‘Double’이라고 명명하였다. 마지막으로 네 번째 조건문장의 경우 세 번째 조건문장과 동일한 구문분석과정을 겪는다. 그러나 두어의 논제적 역할(thematic role)이 세 번째 조건과 다르다. 이 두어의 논제적 역할의 차이는 다음과 같이 다른 재분석과정을 낳는다. 세 번째 조건문장의 두어인 ‘소녀’는 유정명사(animate noun)로 관계절 동사인 ‘본’의 주어와 목적어가 모두 가능하다. 그러나 네 번째 조건문장의 두어인 ‘그네’는 ‘본’의 목적어만 가능하다. 다시 말하면 세 번째 조건문장은 ‘소녀가 본’과 ‘소녀를 본’이 모두 가능하며 네 번째 조건문장의 경우 ‘그네를 본’만 가능하다. 따라서 관계절 주어로 구문분석되는 재분석의 대안이 제외되는 단서가 된다. 네 번째 조건문장 역시 마지막 어절에서 두 개의 논항이 필요한 ‘태웠다’가 제시됨으로 두 개의 논항이 주절로 옮겨가는 구조가 된다. 그러나 세 번째 조건문장처럼 관계절 주어 분석과 목적어 분석 모두가 유지되다가 마지막 어절에서 이 중의성이 해결되는 것이 아니라 관계절의 목적어 분석만이 유지됨으로 네 번째 조건문장의 마지막 어절의 재분석처리의 어려움은 세 번째 보다 덜할 것으로 예측할 수 있다. 네 번째 조건문장의 경우 세 번째와 마찬가지로 주절로의 논항이동이 총 두개가 됨으로 ‘Double’이라고 명명하지만 두어의 논제적 역할로 인해 보다 쉬운 재분석과정이 예측됨으로 ‘Double-easy (E)’라고 명명하고 세 번째는 반대로 ‘Double-Difficult (D)’라고 다시 명명하였다.

수정모형의 경우 Null과 Single, Single과 Double의 재분석처리의 어려움을 예측할 수 있지만 Double-E와 Double-D의 차이는 예측할 수 없다. 왜냐하면 두 조건 모두 동일한 구조를 갖고 있기 때문이다. 반면 진단모형의 경우 재분석의 명확한 정보를 제공하는 Double-E 조건이 더 쉽다고 예측할 수 있다. 결과적으로 재분석이 발생하지 않는 Null 조건이 가장 쉬울 것이고 한 번의 재분석(두어 ‘소녀에게’)이 일어날 것을 예측하는 Single 조건이 다음으로 쉬울 것이며, 재분석이 두 번 발생(두어인 ‘소녀에게’와 마지막 어절인 ‘태웠다’)하지만 두어의 정보가 재분석을 더 쉽고 명확하게 하는 Double-E 조건이 다음으로 쉬울 것이다. Double-D 조건의 경우 두 번의 재분석과 두어의 중의성으로 인해 가장 어려운 문장으로 평가될 것이다.

2. 실험방법

가. 실험참가자

아주대학교에 재학중인 학부생으로 인지심리학, 통계학을 수강하고 있는 76명의 학생을 대상으로 설문지 조사를 하였다. 4개의 설문지 유형에 각각 19명씩 무선 할당하였으며 총 12명의 무성의 응답자는 분석에서 제외하고 64명의 자료를 분석하였다.

나. 실험재료

앞에서 설명한 문장조건 네 개를 우선 설문지로 작성하여 과연 일반인들이 어떤 문장을 더 쉽게 또는 어렵게 이해하는지를 알아보았다.

4개의 조건문장을 동일한 구조이지만 어휘만 다르게 하여 각 조건당 12개씩 총 48개를 만들고 각각 4개의 설문지 목록에 12개씩 역균형화(counter balancing)시켜 할당하였다. 그리고 29개의 매끔질문장(filler)을 만들어 총 41개의 문장을 무선적으로 섞은 다음 각 4종류의 설문지에 넣었다. 따라서 한 설문지 대상자는 조건문장 12개와 매끔질문장 29개를 무선적인 순서로 읽게 된다. 그리고 설문지 대상자는 문장을 읽고 문장의 난이도를 5점 척도로 표시하였다(1-아주 쉽다. 2-쉽다. 3-보통이다. 4-어렵다. 5-아주 어렵다).

다. 실험절차

실험자는 우선 간단한 실험소개를 하고 실험자의 지시 없이 표지를 넘기지 말도록 당부하였다. 그리고 정확한 문장의 이해가 아니라 문장을 한번 보고 그 문장이 얼마나 어려운지를 즉각적으로 평정해야 한다고 강조하였다. 설문지 작성 중간에는 아무런 질문을 하지 못하게 하였으며 설문지 먼저 끝나더라도 끝났다는 의사표시를 못하게 지시하였다. 설문지 응답시간은 약 20분에서 30분 정도 소요되었고 설문지 끝나고 난 다음에는 설문지의 취지를 간단히 설명하였다

라. 실험설계

문장조건 변인이 피험자내 변인이고 목록변인이 피험자간 변인인 4 × 4 혼합설계를 사용하였다.

3. 실험결과

<표 - 1>은 설문조사 결과 나온 난이도 평정에 대한 평균값이다.

<표 - 1> 조건별 난이도 평정 평균표

문장조건	사례수	평균	표준편차
Null	64	2.95	0.95
Single	64	3.52	0.79
Double-D	64	3.71	0.74
Double-E	64	3.69	0.89

<표 - 1>의 결과에서 Null 조건에 비해 구조적으로 더 어려운 Single, Double의 평균값이 더 어려운 방향으로 나왔다($F_{(1, 189)} = 13.88$, $MSe = 10.12$, $p < .001$; $F_{(1, 189)} = 25.32$, $MSe = 18.47$, $p < .001$). 그리고 Single에 비해 Double이 더 어려운 것으로 나왔다. 그러나 통계적으로 유의미하지는 않았다($F_{(1, 189)} = 2.05$, $MSe = 1.49$, $p = 0.15$). 그리고 Double-D와 Double-E의 평균차이에서 Double-D가 좀 더 어렵다는 경향이 나왔으나 역시 통계적으로 유의미하지는 않았다($F_{(1, 189)} = 0.02$, $MSe = 0.01$, $p = 0.89$). 이러한 결과는 Inoue & Chun (1998)에서 추측했던 Double-D 조건이 더 쉬울 것이라는 예측을 반증하는 것이다. 오히려 Hirose & Inoue (1998)의 결과를 지지한다.

4. 논의

실험 1의 결과, 우선 구조적인 재분석의 복잡성이 문장의 난이도에 주요한 영향을 끼친다는 것을 알았다. 그러나 예측했던 관계절 두어의 논제적 역할의 차이는 문장의 난이도에 영향을 거의 끼치지 않는 것으로 나왔다. 결과적으로는 수정모형의 예측만이 지지되었고 진단모형의 예측은 그 경향성만이 나왔다. 그러나 실험 1의 경우 실시간과제가 아니므로 단지 어떤 문장이 더 어렵다 또는 쉽다만을 알 수 있을 뿐 정말 구조적 수정의 복잡성이 재분석처리의 어려움을 반영했는지, 아니면 두어의 논제적 역할 정보가 중요 원인인지는 알 수 없다.

본 설문지 조사를 통해 한국어 화자 역시 일본어 화자와 비슷한 문장난이도 평정을 보이는 것으로 보였다. 이 설문지조사를 바탕으로 문장재분석의 구체적인 과정을 추론하기 위해 실시간 과제를 실험 2에서 실시하였다.

II. 실험 2

1. 관계절 두어의 역할: 실시간과제

실험 2를 통해 구체적으로 어떤 어절에서 어떤 분석이 고려되는 지를 추론하였다. 실험 2에서는 실험 1에서 사용한 문장을 다시 한번 수정하여 보다 자연스러운 문장들을 만들고 문장의 수를 늘려 일반화를 고려하였다.

실시간과제를 통해 밝히고자 한 것은 구체적으로 어떤 어절에서 재분석이 발생하는지, 그리고 그 재분석과정이 어떤 정보를 바탕으로 진행되는지를 밝히고자 하였다. 우선 재분석이 첫 번째로 일어나는 위치는 네 문장조건의 다섯 번째 어절일 것이다. Null 조건의 경우 재분석이 일어나지 않을 것으로 예측되고, Single, Double-D (이하 D-D), Double-E (이하 D-E) 조건들에서는 재분석이 일어날 것이다. 이미 실험 1에서 언급한 두어의 논제적 역할 차이가 재분석에 명확한 단서로 제공된다면 D-E와 D-D간의 읽기시간의 차이가 두어(‘소녀에게’)에서 발생할 것이다. 그리고 마지막 어절에서 Null 및 Single은 재분석 없이 종결될 것이지만 D-D 및 D-E 조건에서는 재분석이 한 번 더 발생할 것으로 보인다. 하지만 두어의 정보가 재분석에 적극적으로 쓰여 관계절의 목적어로만 구문분석되는 D-E의 경우가 D-D보다 더 짧은 읽기시간을 마지막 어절에서 보일 것으로 예측된다. 만약 조건 D-E와 조건 D-D의 두어에서의 읽기시간이 D-E가 더 짧게 나타난다면 진단모형을 지지하는 것이며, 반대로 읽기시간 차이가 나지 않는다면 수정모형이 지지되는 것이다. 다음은 실험 2에서 사용된 수정된 조건문장의 예이다.

Null: 영화가 아이를 놀이터에서 본 사실이 알려졌다.

Single: 영화가 아이를 놀이터에서 본 소녀에게 손짓했다.

D-E: 영화가 아이를 놀이터에서 본 그네에 태웠다.

D-D: 영화가 아이를 놀이터에서 본 소녀에게 맡겼다.

2. 실험방법

가. 실험참가자

고려대학교에 재학중인 학부생 92명이 피험자로 참가하였으며 이 중에서 실수가 높은

12명을 제외한 80명의 자료를 분석하였다.

나. 실험재료

실험 2에 사용한 문장은 실험 1을 수정한 문장이다. 각 조건당 12개(총 48개)의 조건문장을 만들고 네 개의 목록에 역균형화하여 12개씩 할당하였다. 그리고 32개의 메꿈질문장을 각각의 목록에 추가하여 한 실험참가자는 총 44개의 문장을 보았다.

다. 실험도구

실험에 사용된 과제는 어절별 자기조절읽기과제(word by word self-paced reading task)였다. 실험에 사용한 컴퓨터 사양은 CPU가 Intel Pentium II 166 MHz였고 메모리는 64 M였다. 모니터는 1024×768의 17" color 모니터였다. 그래픽카드의 메모리는 2 M인 것을 사용하였다. 이와 같은 사양의 PC를 두 대 사용하여 실험하였다. 실험에 사용한 프로그램은 Visual Basic 6.0으로 연구자가 직접 만들었으며 Windows 환경에서 실행하였다. 읽기시간을 1/1000초 단위까지 측정할 수 있는 타이머를 사용하였다.

라. 실험절차

실험참가자는 설치된 PC 두 대 앞에 각각 앉아 실험자의 지시사항을 듣고 어절별 자기조절읽기과제에 익숙해지기 위해 연습시행을 8문장에 대해 실시했다. 연습시행이 끝난 후 실험참가자는 다시 한번 지시사항을 실험자를 통해 들은 다음 본 실험으로 들어갔다.

자극의 제시는 한 어절씩 총 여섯 어절을 제시하였고 마지막 어절을 읽은 후에 읽은 문장에 대한 이해정도를 파악하기 위해 질문문장을 제시하였다. 질문문장은 ‘누가 누구에게 무엇을 했습니까?’로 주어와 목적어를 동수로 만들어 제시하였다. 읽기시간은 한 어절별로 1/1000 sec 단위로 키보드를 통해 측정되었는데, 즉 어절이 제시되면서 타이머가 작동하고 반응키(방향키의 오른쪽 버튼)를 누르면 해당 어절을 읽으면서 소요된 시간이 텍스트로 기록되었다. 제시된 문장을 정상적으로 이해했는지를 조사하는 질문문장에 대한 반응도 키보드를 통해 입력되었다.

마. 실험설계

실험 설계는 4(조건문장) × 4(목록조건)로 조건문장 요인은 피험자내 변인(within-

subject variable)이고 목록조건은 피험자간 변인(between-subject variable)인 혼합설계를 사용하였다.

3. 결과

전체 실험참가자 92명 중 질문문장에 대한 정확률과 실험 과정 동안에의 태도를 판단해 12명을 제외한 80명의 결과를 분석했다. 종속치는 읽기시간으로 한 어절의 읽기시간을 음절수로 나누어 사용하였다(예: 철/수/는 = 3음절, 만약 읽기시간이 300 msec이었다면 $300/3 = 100$). <표 - 2>는 정답수의 분석 결과이다.

<표 - 2> 조건별 정답수 (최대값 = 3)

조건	사례수	정답수
Null	80	2.71
Single	80	2.60
D-D	80	1.88
D-E	80	1.92

정답수의 조건간 차이가 통계적으로 유의미하였다($F_{(3, 228)} = 39.38$, $MSe = 15.18$, $p < .001$). 조건 Single의 경우 최대값 3에 가까운 2.60으로 Double과의 차이가 통계적으로 유의미하였다($F_{(1, 288)} = 66.56$, $MSe = 25.66$, $p < .001$). 그러나 Null과 Single, D-D와 D-E간의 차이는 통계적으로 유의미하지 않았다. 이 결과는 문장이해의 난이도가 Null과 Single이 D-D와 D-E에 비해 더 쉬웠다는 것을 의미한다.

<표 - 3>은 조건문장의 어절별 평균 읽기시간이다. 재분석이 발생하는 어절 5에서 조건 D-D와 D-E간의 차이는 $F1$ 에서만 통계적으로 유의미하였고($F_{(1, 228)} = 9.60$, $MSe = 1093196.68$, $p < .002$; $F_{2(1, 24)} = 2.1$, $MSe = 37561.36$, $p < .16$) Single과 Double (D-D, D-E)간의 차이는 유의미하지 않았다. 그리고 D-D와 Single간의 차이는 $F1$ 에서만 유의미하였다($F_{(1, 228)} = 8.58$, $MSe = 223690.16$, $p < .003$; $F_{2(1, 24)} = 1.87$, $MSe = 33555.28$, $p < .18$). $F2$ 의 결과가 지지되지는 않았지만 어절 5에서 Null 조건이 가장 빨랐고 다음으로는 D-E가 오히려 Single보다 빨랐다. 그리고 가장 느린 조건은 D-D 조건이었다. 예측한 대로 두어의 논제적 역할의 차이가 재분석에 영향을 끼친 것으로 나타났다.

<표 - 3> 어절별 평균 읽기시간

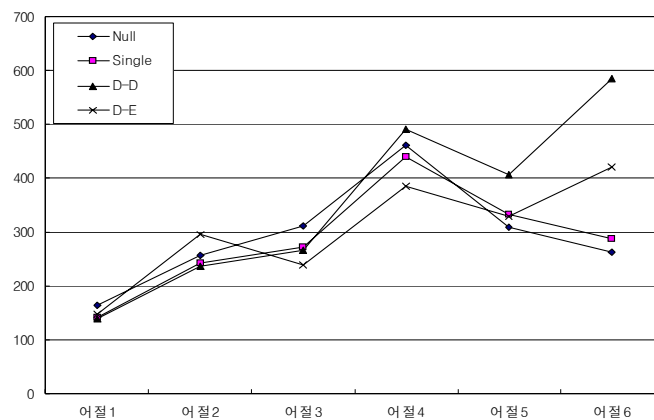
(단위 = msec)

조건	어절1	어절2	어절3	어절4	어절5	어절 6
Null	영희가	아이를	놀이터에서	본	사실이	알려졌다.
	163.49	255.41	310.30	460.92	307.98	262.36
Single	영희가	아이를	놀이터에서	본	소녀에게	손짓했다.
	140.79	243.10	271.46	439.49	331.84	287.80
D-D	영희가	아이를	놀이터에서	본	소녀에게	말했다.
	139.29	235.95	266.69	491.29	406.62	585.37
D-E	영희가	아이를	놀이터에서	본	그네에	태웠다.
	146.96	296.03	239.11	386.17	327.53	420.06

D-D: 재분석과정이 두 번 발생하며 관계절 두어의 논제적 역할이 중의적인 경우

D-E: 재분석과정이 두 번 발생하지만 관계절 두어의 논제적 역할이 명확한 경우

마지막 어절 6에서 읽기시간의 차이는 D-D와 D-E간의 차이가 $F1$ 및 $F2$ 에서 모두 유의미했으며($F1_{(1, 228)} = 17.29$, $MSe = 1093196.68$, $p < .001$; $F2_{(1, 24)} = 8.8$, $MSe = 163977.60$, $p < .001$), 또한 Single과 D-E, Single과 D-D간의 차이가 모두 유의미했다(Single, D-D: $F1_{(1, 228)} = 56.01$, $MSe = 3541957.85$, $p < .001$; $F2_{(1, 24)} = 28.58$, $MSe = 531287.42$, $p < .001$. Single, D-E: $F1_{(1, 228)} = 11.06$, $MSe = 699646.14$, $p < .001$; $F2_{(1, 24)} = 5.65$, $MSe = 104945.66$, $p < .01$). 어절 6의 차이는 우선 재분석에 있어서 구조적 복잡성이 확실히 주요한 요인 중의 하나임을 보이고 있다. 또한 두어의 논제적 역할이 마지막 재분석에서도 영향을 지속적으로 끼치고 있음을 알 수 있다. <그림 - 4>는 평균 읽기시간을 꺾은선 그래프로 제시한 것이다.



<그림 - 4> 실험 2의 어절별 평균 읽기시간의 꺾은선 그래프

4. 논의

실험결과를 요약하면 재분석이 일어날 것이라고 예측한 Single, D-D, D-E 조건의 어절 5의 평균 읽기시간이 Null보다 길었다. 이것은 재분석이 일어남을 나타낸다. 또한 두어의 정보가 재분석을 보다 명확하게 해주는 D-E의 읽기시간이 그렇지 않은 D-D 조건 보다 더 길었다. 이 결과는 두어의 어휘적 정보를 사용하지 않는다는 수정모형보다는 진단모형을 지지하는 결과라고 할 수 있다. 그리고 또 한번의 재분석이 발생할 것이라고 예측한 D-D 및 D-E의 읽기시간이 Null 및 Single 조건보다 더 길었다. 이것은 재분석과정에서 구조적 수정의 복잡성이 더 클수록 처리의 부담이 가중된다는 것을 의미한다. 마지막으로 두어의 정보가 마지막 어절까지 영향을 끼치게 되어 D-E 조건이 D-D 조건보다 더 짧은 읽기시간을 보였다.

구체적인 재분석과정을 추론하면 다음과 같다. 모든 조건문장에서 네 번째 어절까지 동일한 어순과 어휘로 구성되어 있기 때문에 동일한 구조를 갖을 것이라고 가정할 수 있다. 여기서 말한 동일한 구조란 실험 1에서 언급하였듯이 초기분석의 원리인 최소부착의 원리에 따라 ‘영희가 아이를 놀이터에서 본’이 하나의 절로 형성됨을 의미한다. Single 및 D-D 조건의 경우 어절 5(‘소녀에게’)에서 오인현상(garden path)에 빠지게 되는데, 그 이유는 ‘소녀’가 관계절의 두어가 되지만 관계절의 주어가 되는지, 관계절의 목적어가 되는지 모호하기 때문이다. 수정모형에 따르면 여하튼 어떤 한 구조를 선택하여 분석을 이어나가야 한다. 그러면 어떤 구조를 선택하는지에 관한 의문이 남는데, 최소수정의 원리에 의하면 문장의 구조를 최소한 수정하는 방식으로 재분석이 진행 될 것이다(Frazier & Clifton, 1998; Inoue & Fodor, 1995; Mazuka, 1991). 즉, ‘소녀’가 관계절의 주어로 분석되는 경우만을 선택할 것이다(두어인 ‘소녀’가 관계절의 목적어로 참조되는 경우는 수정된 구조가 더 복잡하다). 그리고 마지막 어절에서 선택한 구조가 정확함을 확인하면서 Single구조는 또 한 번의 재분석 없이 문장 분석이 종결된다. 그러나 D-D의 경우 마지막 어절에서 주절의 동사가 논항을 3개 필요로 하는 동사임을 알고 다시 재분석을 해야 한다. 주절의 동사에서의 재분석의 경우, 관계절 ‘아이를 놀이터에서 본’에서 ‘아이를’을 다시 주절로 옮기는 과정이 더 추가된다.

여기까지의 분석과정은 수정모형으로도 설명이 가능하다. 그러나 D-E 조건을 고려하면 다른 분석과정이 있음을 알 수 있다. 즉, 두어의 논제적 역할이다. D-E 조건의 두어인 ‘그네’는 무정명사로 관계절 동사 ‘본’의 목적어로 밖에 쓰일 수 없다. 이는 D-D 조건의 두어에서처럼 오인현상에 빠질 가능성을 대폭 감소시키는 단서가 된다. 따라서 문장분석기(parser)는 D-D에서처럼 틀린 분석을 선택하지 않고 옳은 분석인 관계절의 목적어분석을 선택하게 된다. 관계절의 목적어 분석을 선택함으로써 D-E 조건의 마지막 어절인 ‘태웠다’에서 D-D 조

건보다 더 짧은 읽기시간이 나온 것이라 할 수 있다.

요약하면 오인현상에 접했을 때 어떤 분석을 선택할 것인가에 대한 결정을 도와주는 정보가 있다면 그 정보를 적극적으로 사용하는 진단모형이 한국어 관계절 재분석에서 더 적절한 모형이라고 할 수 있다.

Ⅲ. 실험 3

1. 주격조사와 주제표지의 효과(실시간과제)

실험 3에서는 한국어의 특징 중 하나인 조사(또는 격표지)가 재분석에 도움을 줄 수 있는 정보인지를 살펴보았다. 실험 3에서 사용한 조사는 주격조사인 ‘-이/-가’와 주제표지(topic marker)인 ‘-은/-는’이다.

주격조사 ‘-이/-가’의 경우 주절의 주어 및 종속절(예: 관계절)의 주어로 기능할 수 있다. 그러나 주제표지인 ‘-은/-는’은 주절의 주어로만 기능한다(Kim, 1999). 이러한 기능의 차이가 재분석에 영향을 끼친다면 실험 2에서 논의한 진단모형의 또 하나의 증거가 될 수 있을 것이다.

실험 3에서 사용한 조건문장은 실험 2에서 사용한 Single 조건(복합 명사절에서 재분석시 한 개의 논항이 주절로 이동)과 Double 조건(복합 명사절에서 두 개의 논항이 주절로 이동하는 구조)을 변형하여 사용하였다. 다음은 조건문장의 예이다.

Single-가: 갯단의 두목이 부하를 어렵게 찾아낸 사람에게 어제 사례했다.

Single-은: 갯단의 두목은 부하를 어렵게 찾아낸 사람에게 어제 사례했다.

Double-가: 갯단의 두목이 부하를 어렵게 찾아낸 사람에게 어제 보냈었다.

Double-은: 갯단의 두목은 부하를 어렵게 찾아낸 사람에게 어제 보냈었다.

수정모형에 의하면 주격조사가 쓰이든, 주제표지가 쓰이든 우선 ‘두목이(은)’를 절의 주어로 분석할 것이다. 왜냐하면 문장의 구조적 정보를 바탕으로 한 모형이기 때문에 초기분석에서 최소 부착의 원리를 적용했다고 보기 때문이다. 그러나 진단모형의 경우 주격조사 ‘-이/-가’는 주절의 주어 또는 종속절의 주어가 모두 가능하다. 따라서 중의적인 부분이 된다. 반대로 주제표지인 ‘-은/-는’의 경우는 주절의 주어로만 사용된다는 제한점 때문에 한 가지 가능성만을 가정하게 된다.

이 정보가 재분석에서 사용된다면 우선 재분석이 발생하는 관계절의 두어에서 읽기시간의 차이가 발생할 것으로 예측된다. 초기분석에서 single과 double의 ‘-가/-이’조건은 절의 주어로 우선 선택하고 분석을 진행하지만 관계절의 주어로의 가능성을 배제시키지는 못했다. 그러므로 관계절임을 확인하고 재분석을 해야 하는 ‘사람에게’에서 ‘두목이’를 주절의 주어로 옮겨가는 재분석과정을 거쳐야 한다. 반면 ‘-은/-는’의 경우 확실하게 관계절 주어의 역할은 배제되기 때문에 ‘사람에게’에서 ‘두목은’을 주절로 배제시키는 과정이 불필요할 것으로 예측된다. 결과적으로 ‘사람에게’에서 ‘-이/-가’ 조건이 ‘-은/-는’ 조건에 비해 읽기시간이 더 짧을 것으로 예측된다.

두 번째 재분석이 일어나는 마지막 어절에서 ‘single-가’ 조건의 경우 두어에서 재분석한 결과가 그대로 옳은 분석이 되므로 필요없게 되어 그대로 문장분석을 종결하게 된다. 역시 ‘single-은’ 조건도 재분석 없이 종결되지만 주제표지의 역할로 인해 ‘single-가’ 보다는 짧은 읽기시간을 보일 것이다. ‘double-가’와 ‘double-은’ 조건 모두 ‘부하를’이 관계절에서 배제되어 주절로 옮겨가는 재분석과정을 거쳐야 한다. 따라서 두 조건간의 차이는 마지막 어절에서 없을 것으로 예측된다.

2. 실험방법

가. 실험참가자

아주대학교에서 심리학개론을 수강한 90명이 실험에 참가하였다. 이 중에서 80명의 자료만을 분석에 포함시켰다.

나. 실험재료

실험 3에서 사용한 조건문장은 실험 2의 Single 조건과 Double 조건의 문장구조와 동일한 구조를 사용했고 단 주절의 주어와 주격조사를 갖는지, 주제표지를 갖는지만을 조작하였다. 그리고 보다 자연스러운 문장을 만들기 위해 주절의 주어 앞에 수식어를 붙이고 주절의 동사 앞에 부사어를 붙였다.

실험문장은 조건당 8개씩 32개를 네 개의 목록에 역균형화시켜 배당하고 매꿈질문장 60개를 추가하였다. 한 실험참가자는 총 92개의 문장을 보았다.

다. 실험도구

실험에 사용한 PC는 CPU가 AMD 700 MHz이고 모니터는 해상도가 1024 × 786인 20" 모니터 였으며 E-prime으로 만들어진 단어별 자기조절읽기과제(word by word self-pace reading task)를 사용하였다.

라. 실험절차

실험 2와 동일한 절차를 사용하였다.

마. 실험설계

실험 2와 동일한 설계를 사용하였다.

3. 실험 3의 결과

<표 - 4>는 실험 3의 문장에 대한 질문의 정답수이다.

<표 - 4> 조건별 정답수 (최대값 = 8)

조건	사례수	정답수
Single-가	80	7.36
Single-은	80	7.22
Double-가	80	5.02
Double-은	80	5.07

정답수 분석 결과 Single 조건이 Double 조건에 비해 약 2점 정도 점수가 높았고 통계적으로 유의미하였다 ($F_{(1, 228)} = 66.59$, $MSe = 25.66$, $p < .001$; $F_{2(1, 228)} = 13.01$, $MSe = 171.12$, $p < .01$). 그러나 주격조사와 주제표지간의 차이는 무의미했다.

<표 - 5>는 어절별 평균 읽기시간이다. 첫 번째 재분석이 발생하는 어절 6에서 조건간 차이는 통계적으로 유의하진 않았지만 예측대로 주격조사 조건(Single-가, Double-가)의 읽기시간이 주제표지 조건(Single-은, Double-은)보다 길었다.

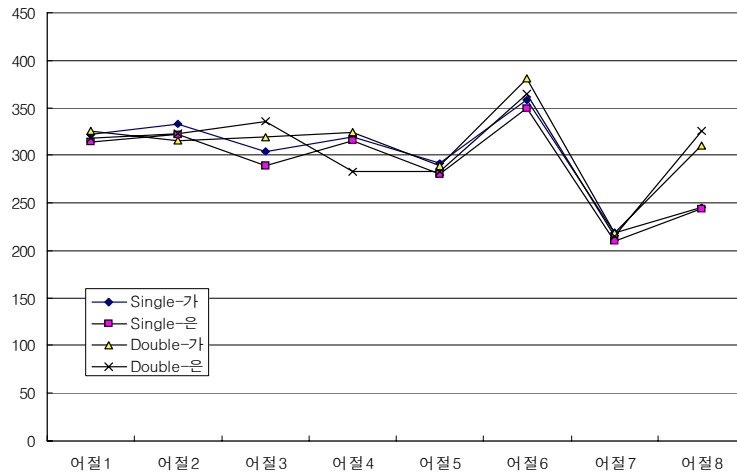
<표 - 5> 어절별 평균 읽기시간

(단위 = msec)

조건	어절1	어절2	어절3	어절4	어절5	어절 6	어절 7	어절 8
Single-가	갱단의 321.45	두목이 332.86	부하를 304.14	어렵게 319.63	찾아낸 292.02	사람에게 358.39	어제 218.22	사레했다. 244.84
Single-은	갱단의 314.16	두목은 322.41	부하를 289.51	어렵게 315.18	찾아낸 279.84	사람에게 349.63	어제 210.02	사레했다. 243.75
Double-가	갱단의 325.17	두목이 315.57	부하를 318.76	어렵게 324.32	찾아낸 289.63	사람에게 381.14	어제 218.64	보냈었다. 311.10
Double-은	갱단의 318.35	두목은 323.61	부하를 335.41	어렵게 283.2	찾아낸 283.2	사람에게 364.61	어제 214.55	보냈었다. 325.15

Double-가: 어절 6과 어절 8에서 재분석이 일어나고 관계절 두어의 논제적 역할이 중의적인 조건
 Double-은: 어절 6과 어절 8에서 재분석이 일어나고 관계절 두어의 논제적 역할이 명확한 조건

재분석이 두 번째로 발생하는 마지막 어절에서 조건간 차이가 통계적으로 유의미하였
 다($F_{(3, 288)} = 14.95$, $MSe = 148005.24$, $p < .001$; $F_{2(3, 84)} = 8.52$, $MSe = 59202.09$, $p < .001$).
 이 차이는 문장재분석에서 구조수정 복잡성의 차이로 보이고 주격조사나 주제표지의 차이가
 연장된 것으로 보이지는 않는다. <그림 - 5>는 어절별 평균 읽기시간의 꺾은선 그래프이다.



<그림 - 5> 실험 3의 어절별 평균 읽기시간의 꺾은 선 그래프

4. 논의

실험 3의 결과를 요약하면 주격조사와 주제표지간의 차이는 첫 번째 재분석 위치인 ‘사람에게’에서 예측대로 ‘Double-가’ 조건이 가장 긴 읽기시간을 냈고 ‘Single-은’ 조건이 가장 짧은 읽기시간을 보였다. 또한 마지막 어절에서는 재분석이 발생하지 않는 Single 조건이 재분석이 일어나는 Double 조건에 비해 더 짧은 읽기시간을 냈다. 그러나 주격조사와 주제표지간의 차이는 없었다. 실험 3의 결과는 재분석에서 조사의 정보가 사용되어 어떤 재분석으로 향할지를 안내하는 역할을 함을 나타낸다. 즉 수정모형에서는 예측할 수 없는 결과를 진단모형은 예측할 수 있음을 나타낸다. 실험 2와 같이 실험 3에서도 역시 진단모형을 지지하는 결과가 나타났다.

그러나 ‘-가/-이’나 ‘-는/-은’이 즉각적으로 통사분석에 사용된다면 직접적인 읽기시간의 차이가 있을 것으로 보였으나 그 차이는 해당 어절에서 없었다. 그렇지만 주어 다음의 어절인 세 번째 어절(‘부하를’)에서 주격조사와 주제표지간의 차이가 나타났다(single-가: 304.14, single-은: 289.51, Double-가: 318.76, Double-은: 296.67). 이것은 효과가 spill-over 된 것으로 보인다. 이런 패턴은 Kim (1999)에서도 나타났다.

IV. 논 의

본 논문을 통해 검증하려고 했던 것은 한국어 문장의 재분석에 어떤 모형이 적용될 수 있는지를 검증하는 것이었다. 결과적으로 수정모형보다는 진단모형이 더 지지를 받는 결과를 얻었다. 그러나 몇 가지 의문점이 여전히 남아 있다. 첫째는 실험 2에서 두어의 논제적 역할에 대한 것이다. 실험 2에서 실제적으로 초기분석이 이루어지는 부분은 어절 5(관계절의 두어)까지일 수 있다. 왜냐하면 하나의 절로 묶이는 ‘영희가 아이를 놀이터에서 본’까지 처리하고 재분석을 즉각적으로 ‘소녀에게’ 또는 ‘그네에’에서 시작한다고는 볼 수 없다. 즉 두어인 ‘소녀’나 ‘그네’를 우선 읽고 앞선 문장구조가 틀렸음을 알고 재분석에 들어가야 한다. 따라서 두어는 초기분석과정과 재분석과정이 겹쳐져 있다고 할 수 있다. 이렇게 생각한다면 초기분석에서 활성화되는 정보가 과연 구조적(통사적) 정보인지 의미적 또는 논제적 정보인지가 다시 중요한 쟁점이 된다. 이 쟁점은 여전히 풀리지 않고 있으나 여기서 어떻다라고 할 수 없다. 그러나 실험 2의 결과에 따르면 재분석의 시점에서 어휘에 포함된 여러 정보(의미적, 맥락적, 참조적 등)들 중에서 논제적 역할이라는 정보가 도출되었다고 할 수 있다. 단지 논제적

역할 정보가 가장 현출할 것이다라는 추측은 근거가 부족하다. 아직 통사적, 의미적 정보 중에서 어떤 정보가 재분석에서 우선시된다는 증거는 없기 때문이다. 설령 논제적 정보가 재분석의 시점에서 중요하다고 추측할 수 있다고 하더라도 또 다른 의문점이 있다. 그 의문점은 논제적 역할 안에서도 어떤 것이 더 선행하는지 그 위계에 관한 것이다. Hirose & Inoue (1998)는 논제적 역할의 선호의 순위는 Camire (1981)의 의미역(semantic role)의 순위를 따를 수 있다고 하였다.

즉, 행위자(agent)가 가장 상위이고 경험자(experiencer)가 다음이고 원인(cause)이 세 번째이며 주제(theme)가 마지막 순위라는 것이다. 따라서 두어의 의미역에서 행위자(관계절의 주어)로 선택되는 D-D 조건이 오인현상에 빠지게 된다고 하였다. 이 가정을 실험 2의 결과에 적용하여도 타당할 것 같다. 그러나 아직 실험적으로 증명된 사실이 아님으로 확인하기 힘들다. 또한 논제적 역할에 대한 Trueswell & Tanenhaus (1994)의 실험에서 일반적인 문장의 주어자리에 유정명사가 오느냐, 무정명사가 오느냐에 따라서 관계사 'that'이 생략된 관계절의 중의성해소에 영향을 받는다는 결과가 있었다. 이들의 실험은 재분석이 아닌 초기분석에서의 중의성해소에 초점이 있지만 순수하게 통사적 정보만이 문장의 처리에 관여한다는 모형들(Frazier, 1978, 1987)은 설명할 수 없는 부분이기도 하다. 이러한 가정들을 바탕으로 한다면 한국어 문장의 처리, 특히 재분석과정을 설명하기 위해선 초기분석에서 어휘의 의미적 요소도 고려하는 상호적인 모형의 개발이 필요할 것으로 보인다(김영진, 1996).

두 번째 의문은 조사의 효과가 재분석에 영향을 끼친다면 초기분석에서 조사의 역할이 결정되었다는 것을 의미한다. 이것은 초기분석에서 주격조사와 주제표지의 기능(관계절의 주어로 쓰일 수 있다 또는 없는데 대한 정보)이 이미 활성화되었다는 것을 또한 의미한다. 이렇게 본다면 실험 3의 '두목이'와 '두목은'에서 읽기시간의 차이가 있었어야 한다. 그러나 실험 3의 결과에서 보듯 두 조건의 해당 어절에서 읽기시간의 차이는 없었다. 그렇다면 가능한 분석을 모두 유지하는 것일까? 아니면 주격조사나 주제표지가 모두 동일하게 처리하는 것일까?

Inoue & Fodor (1995)에 의하면 한국어나 일본어와 같이 지역적 중의성이 많은 언어는 우선 재분석에 대한 처리의 비용이 없다고 가정한다. 그리고 중의적인 시점에서 분석가능한 구조들에 대한 일종의 탐침(flag)을 설정하면서 일반적인 통사적 분석의 원리(예: 최소 부착의 원리)를 적용해 간다고 한다. 이때 탐침의 설정은 인지적 에너지를 필요로 하지 않으며 그 설정의 순위는 해당 분석의 확실한 정도에 비례한다고 하였다. 재분석을 이 모형으로 설명하자면 재분석의 시점에서 탐침을 설정한 부분으로 돌아가 처음에 선택했던 즉, 가장 확신한 분석방법 다음의 차선의 분석을 선택하여 다시 최소부착의 원리로 진행해 가는 것이다. 이

모형으로 두 번째 의문을 설명하면 다음과 같을 것이다. ‘두목이’를 접했을 때 (10a) - (10c)와 같은 구조들이 활성화될 것이다.

(10a) (두목이 부하를 찾아낸)

(10b) 두목이 (부하를 찾아낸)

(10c) 두목이 부하를 (찾아낸)

이들 구조들 중에서 가장 확신하는 구조는 가장 단순한 구조인 10a가 될 것이다. 그러나 ‘사람에게’가 ‘찾아낸’ 다음에 나타나면 재분석이 발생하게 되어 ‘두목이’를 괄호 밖으로 옮겨야 한다(주절의 주어로의 이동을 의미한다). 그러나 ‘두목이’가 괄호 밖으로 설정된 구조는 확신이라는 가중치가 두 번째(10b)인 것으로 차선의 구조이다. 따라서 재분석의 비용이 들 것이다.

반대로 주제표지인 ‘-은’이 사용된 경우를 보자.

(11a) (두목은 부하를 찾아낸)

(11b) 두목은 (부하를 찾아낸)

(11c) 두목은 부하를 (찾아낸)

주제표지의 기능이 반영되어 (11a)는 가능한 분석의 구조가 아님으로 탈락된다(Inoue & Fodor의 주장에서는 이 구조가 탈락됨을 예측할 수는 없지만 일단 탈락된다고 보자). 그러면 (11a)와 (11b)만이 가능한 구조가 된다. 두 구조 중에서 (11b)가 더 단순 구조임으로 이 구조에 대한 가중치가 더 크다. 따라서 (11b)구조를 선택하고 문장 분석이 진행된다. ‘찾아낸’ 다음에 ‘사람에게’라는 두어가 제시되고 재분석이 시작되면 이미 주절의 주어로 할당된 (11b)를 재분석할 필요는 없다. 따라서 ‘두목이’ 예문에 비해 ‘두목은’의 예문의 두어 ‘사람에게’가 읽기시간이 짧아야 한다. 실험 3에서 이와 동일한 결과가 나왔다.

그러나 두 번째 의문은 초기분석에서 조사의 기능이 과연 즉각적으로 반영되는가이다. Inoue & Fodor의 모형은 이러한 가정은 하지 않았지만 재분석에 대해 높은 설명력을 갖는다. 이 장점을 유지하면서 이 모형에 조사의 정보를 참고하는 기능이 추가된다면 ‘두목이’와 ‘두목은’의 즉각적인 처리의 차이는 설명할 수 있을 것이다. 실험 3에서 어절 3(‘두목이/은’ 다음 어절인 ‘부하를’)에서 두 조건간 읽기시간의 차이를 보인 것은 (11a) 분석이 배제되어 탐침을 설정해야 하는 경우의 수가 하나 줄어들어 인지적 처리의 비용이 줄어든 것으로 해석된다.

세 번째 의문점은 한국어와 일본어의 차이점에 대한 것으로, 관계절화 어미 '-ㄴ'의 존재가 있는 한국어와 없는 일본어의 처리의 차이가 있는지이다. 차이가 있는지 없는지는 본 실험의 결과로 구분할 수 없다. 우선은 그 결과의 경향이 동일하고 두 언어가 비교를 위한 실험조건이 없으므로 직접적인 비교는 할 수 없다. 그러나 다음과 같은 처리의 차이는 추론할 수 있다.

우선 일본어의 경우 관계절화 어미가 없이 일반적인 문장종결어미 '-ta'가 관계절이 될 수 있다. 이 특징은 일본어가 더 재분석이 어려울 수 있다는 추측도 가능하게 하지만 실제적으로 일본어 화자는 별 어려움을 갖지 않는다고 한다(Hirose & Inoue, 1998). 그러나 관계절화 어미가 있는 경우 '두목이 부하를 찾은'에서 관계절의 존재를 보다 명확하게 암시할 수 있다. 즉, '-ㄴ'을 접했을 때 적어도 종속절이 있음을 신호할 수 있다. 따라서 일본어보다 관계절의 존재를 적어도 한 어절 앞서 표면적으로 알려준다.

일본어의 경우 관계절이나 일반 단문이 모두 '-ta'로 끝나므로 표면적으로 더 중의적이라고 할 수 있다. 그러나 이런 추측은 일본어 환경을 현재 모르고 있으므로 확신할 수 없다. Inoue & Chun (1998)는 이런 차이가 있을 것으로 논의하고 있으나 그 구체적인 실험적 차이는 제시하지 못하였다.

결론적으로 한국어 문장을 처리하는 데 있어서 특히 재분석과정은 수정모형처럼 문장의 통사적 정보만으로 설명할 수 없는 언어적 특징을 가지고 있다. 따라서 제안된 진단모형의 경우, 한국어의 특징을 반영할 수 있어서 적용가능한 모형이라고 할 수 있다.

그러나 문장의 재분석은 초기분석의 결과를 바탕으로 하는 것이므로 재분석과 초기분석의 유기적인 정보의 소통이 필요할 것이다. 통사적 정보만으로 형성된 초기분석 결과로는 한국어의 재분석과정을 설명하는 데 무리가 따르므로 초기분석 단계에서 어휘적 정보와 구조적 정보의 병렬적인 활성화를 가정하는 것이 한국어 재분석모형에 적절할 것으로 보인다. 이에 한국어 재분석모형은 '통사-어휘정보 상호작용적 모형'이라고 할 수 있을 것이다.

참 고 문 헌

- 김영진(1996). 조사 '는/은'이 포함된 관계절 문장의 이해 과정. 『한국심리학회지: 실험 및 인지』, 10, 119-133.
- 김영진(1996). 한국어 통사처리와 논제/참조처리: 조사의 역할. 『한국심리학회지: 실험 및 인지』, 8, 207.
- 김영진(2001). 한국어 문장처리과정의 보편성과 특수성. 『한국심리학회지: 실험 및 인지』, 13, 339-359.

- 박경자(2001). 『응용언어학 사진』. 서울: 경진문화사.
- 이기용(1999). 『전산형태론』. 서울: 고려대학교 출판부.
- Camire, B. (1981). *Language universals and linguistic typology*. Chicago: The University of Chicago Press.
- Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht, Netherland: Foris.
- Frazier, L. (1978). The sausage machine: A new two stage parsing model. *Cognition*, 6, 1-34.
- Frazier, L. (1987). Syntactic processing: Evidence from Dutch. *Natural Language and Linguistic Theory*, 5, 519-559.
- Frazier, L. & Clifton, Jr., C. (1998). Sentence reanalysis and visibility. In J. D. Fodor & F. Ferreira (Ed.), *Reanalysis in sentence processing*. Dordrecht Netherland: Kluwer.
- Gorrell, P. (1995). *Syntax and Parsing*. Cambridge: Cambridge University Press.
- Hirose, Y. & Chun, S. A. (1998). Attachment ambiguity in head final language. *Japanese and Korean Linguistics* (Vol. 7). Stanford, CA: Stanford University.
- Hirose, Y. & Inoue, A. (1998). Ambiguity of reanalysis in parsing complex sentences in Japanese. In D Hillert (Ed.), *Syntax and Semantics (vol. 31)*. New York: Academic Press.
- Inoue, A. & Fodor, J. (1995). Information-paced parsing of Japanese. In R. Mazuka & N. Nagai (Eds.), *Japanese sentence processing*. Hillsdale, NJ: Erlbaum.
- Kim, Y. J. (1999). The effects of case marking information on Korean sentence processing. *Language and Cognitive Processing*, 14, 687-714.
- Mazuka, R. (1991). Processing of empty categories in Japanese. *Journal of Psycholinguistic Research*, 20(3), 215-232.
- Pritchett, B. (1992). *Grammatical competence and parsing performance*. Chicago IL: University of Chicago Press.
- Sturt, P. & Crocker, M. (1999). Structural change and reanalysis difficulty in language comprehension. *Journal of Memory and Language*, 40, 136-150.
- Suh, S. K. (1994). *The syntax of Korean and its implication for parsing theory*. Unpublished doctoral dissertation, University of Maryland.
- Trueswell, J. C. & Tanenhaus, M. K. (1994). Toward a lexicalist framework for constraint-based syntactic ambiguity resolution. In C. Clifton, Jr., L. Frazier. & K. Rayner. (Eds.), *Perspective on sentence processing*. Hillsdale, NJ: Erlbaum.
- Yamasita, H. (1994). Processing of Japanese and Korean. Doctoral dissertation, Ohio State University.
- Yamasita, H. (1995). Verb argument information used in prodrop language: An experimental study in Japanese. *Journal of Psycholinguistic Research*, 24(5), 333-347.
- Yamasita, H., Stowe, L. & Nakayama, M. (1993). Processing of Japanese relative clause construction. *Japanese and Korean Linguistics*, 2, 248-263. Stanford, CA: Stanford University.

ABSTRACT

The Effects of Thematic Role in Head Noun and Case Marking Information on Korean Relative Clause Re-parsing

Youan Kwon*, Youngjin Kim**, Kichun Nam*

(*Korea University, **Ajou University)

Head-final languages are different from head-initial languages in many respects. Korean is a head-final language. A number of sentence re-parsing models have been proposed for English, which is a head-initial language. The thematic roles of both head-nouns and case markers are characteristic of Korean sentences. The purpose of the present study was to investigate the thematic roles in Korean that affect the sentence re-parsing. Three experiments were carried out. The first off-line experiment was conducted in order to define the type of sentences that entail higher processing cost. It was found that the double displacement condition (D - D) brought about the highest degree of difficulty. In the second experiment, we manipulated the thematic roles of head-nouns in such a way that we could detect the different effects of the thematic roles on the re-parsing of head-nouns. The condition thus created was D - E, that is, the double displacements in the initial parsing with a single thematic role of "theme" in the head-nouns, which is contrasted with the D - D condition, which contains two thematic roles of "agent" and "theme" in the head-nouns. The result of the second experiment was that the reading time of the re-parsing region of the D - D condition was longer than that of the D - E condition. A third experiment was conducted to determine the effect of the topic marker "-nun," which is used only in the matrix sentence in Korean, in reducing the cost of sentence re-parsing. It was found that the topic marker didn't affect re-parsing; it affects only the initial parsing, which is in conflict with other studies advocating the role of the topic marker in sentence re-parsing in Korean. We could tentatively conclude that there seems to be interactions between lexical and syntactic modules.

Key Words: re-parsing, relative clause, thematic role, sentence structure, revision model, diagnostic model

-
- ▶ 게재 신청일: 2003년 2월 12일
 - ▶ 게재 확정일: 2003년 2월 21일

- ▶ 권유안 (제 1 저자): 고려대 심리학과 박사과정, e-mail: thot@korea.ac.kr
- ▶ 남기춘 (교신 저자): 고려대 심리학과 교수, e-mail: kichun.ac.kr
- ▶ 김영진 (공동 저자): 아주대학교 심리학과 교수, e-mail: yjkim@hanmail.net