

한국어 병렬문의 중의성 분석

이호동, 박종철*†

한국과학기술원

Hodong Lee and Jong C. Park. 2003. Analysis of Coordination Ambiguity in Korean. *Language and Information* 7.2, 59-79. Coordination in natural language is often syntactically explained by the rule schema $X \rightarrow X' \text{ and } X''$, where X' and X'' refer to expressions of a like syntactic type. While this schema does explain a lot of instances of coordination, it is well known that this schema also gives rise to serious structural ambiguity, leading not only to numerous unintended interpretations but also to a high degree of unnecessary computational complexity. In order to identify intended interpretations, we need to take into account a number of clues at all the levels of morphology, syntax, semantics, and pragmatics/discourse. In this paper, we propose, with an accompanying data analysis, principles of symmetry and coherency for an effective and reasonable identification of intended interpretations from coordination-bearing sentences. (Korea Advanced Institute of Science and Technology)

Key words: 병렬(coordination), 구조적 중의성(structural ambiguity), 접속항 선정(conjunct identification), 대칭성(symmetry), 응집성(coherency)

1. 서론

병렬문은 일반적으로 둘 이상의 구나 문장이 대등하게 연결된 문장을 말한다. 그러나 이러한 병렬문의 접속항은 다양한 형태로 나타날 수 있기 때문에 그 처리가 어렵다. 병렬문은 일반적으로 Schachter (1977)에서 언급된 바와 같이 접속항간의 구문적인 유사성을 지니고 있다는 사실을 전제로 처리되고 있지만 병렬문에서 나타나는 접속항에서 구문적으로 유사한 후보가 여러 개 존재할 경우 이들 중 올바른 후보를 선택해야 하는 어려움이 따르게 된다. 이러한 병렬문의 처리에 대해 Postal (1971)은 병렬은 적어도 양쪽 접속항 사이의 의미적 차이를 전제하거나 유발하며, 많은 경우에 억양법이 이러한 차이를 판별하는 수단이 될 수 있다고 했다. 이와 유사하게, Cremers (1993)은 문법적으로 올바른 병렬문에서 양쪽 접속항은 접속항 밖의 구문에 대해 유

* 305-701 대전광역시 유성구 구성동 373-1 한국과학기술원 전산학전공,
E-mail: {hdlee, park}@nlp.kaist.ac.kr

† 첨단정보기술연구소센터를 통해 본 연구를 지원해 준 한국과학재단에 감사드립니다.

형적인 유사성이 있다고 지적했다. 그러나 현재 대부분의 자연언어처리를 위한 문법 이론에서 도입되고 있는 병렬에 대한 통사 규칙인 $X \rightarrow X' \text{ and } X''$ 은 직관적이기는 하지만 그 함축성이 높아 병렬문의 처리에서 많은 중의성을 유발하는 원인이 되고 있다. 이러한 중의성의 예는 다음 문장에서 살펴볼 수 있다.

1. [설비 [투자]]와 [[대외] 투자]를 증가시키는 효과가 있다고 분석되었다.

이 문장에서는 명사구 병렬을 보이고 있는데 각 접속항의 명사구에서 통사 범주가 일치하는 여러 후보를 가질 수 있기 때문에 이로 인한 중의성이 나타난다. 이러한 중의성은 영어 전치사구의 구조적 중의성과 같이 구조적인 이유에 의해 발생하는데 병렬문의 전산처리 과정에서 오분석을 유도할 뿐만 아니라 처리대상이 되는 후보의 수를 늘림으로서 처리의 복잡도를 증가시킨다 (Park and Cho, 2000). 또한 영어의 전치사구에서보다 더욱 다양한 양상으로 나타나기 때문에 그 중의성을 분석하고 처리 방법을 제안하는데 어려움을 주는 요인으로 작용할 수 있다.

이렇게 병렬문의 통사적 중의성으로 인해 나타나는 여러 후보 중 올바른 후보를 선택하기 위해서는 형태소 및 통사, 의미 단계의 접두사, 접미사, 통사 범주 및 자질, 의미 범주 등의 정보들뿐 아니라 화용 및 문맥, 억양 단계에서의 주제 구조, 강세 및 휴지(pause) 등의 정보까지도 필요하다고 알려져 있다 (Steedman, 2000; Hoeks, Vonk, and Schriefers, 2002). 그러나 실제 문장의 처리에 있어서 이렇게 다양한 정보들을 개별적으로 적용하여 병렬문에서의 올바른 후보를 선택하기 힘들기 때문에 이러한 정보들을 종합하여 일관성있게 고려할 수 있는 방법이 필요하다. 이러한 목적을 위해 기존의 방법에서는 주로 병렬항 간의 대칭성을 이용하고 있다 (Agarwal and Boggess, 1992; Okumura and Muraki, 1994; Kurohashi and Nagao, 1994; Park and Cho, 2000).

2. (a) 설비 투자와 대외 투자

(b) 예쁜 소녀와 노인

대칭성은 병렬항이 형태소 및 통사, 의미에서 대칭적인 요소를 지닌다는 점에 기반한 것으로, 2a에서의 투자가 양쪽 접속항에서 대칭적으로 나타나 병렬 구조를 명확히 하고 있다. 이와 달리 2b의 명사구에서는 소녀와 노인이 의미상으로 대칭이 되지만 이러한 대칭성만으로는 정확한 병렬항의 범위를 찾을 수 없고 수식어인 예쁜의 수식 범위에 대한 정보가 필요하다는 것을 알 수 있다. 이러한 정보는 복합명사나 영어 전치사구에서의 구조적 중의성을 해결하는데에도 사용되는 것으로 (Hindle and Rooth, 1993; Kobayasi, Tokunaga, and Tanaka, 1994), 2b에서와 같이 병렬문에서도 수식 관계는 병렬항의 범위를 알아내는데 도움을 줄 수 있다 (Resnik, 1999). 이와 같이 명사병렬에서의 수식관계는 병렬항에 의해 공유되는 요소를 찾는데 사용될 수 있는데 공유되는 요소는 병렬항의 범위에 들지 않으므로 병렬문에서 병렬항의 내의 요소들은 서로 공유되지 않고 의미적으로 병렬항 내의 요소들이 뭉치게 된다고 볼 수 있다. 앞

서 보인 2b에서와 같이 이러한 점은 병렬항의 범위를 알아내는데 도움을 줄 수 있는데 이 성질을 본 논문에서는 병렬항의 의미적 응집성이라고 한다. 2a에서도 투자의 대칭성 외에 왼쪽 병렬항의 설비가 공유되는지를 파악해야 병렬항의 범위가 명확해 질 수 있는데 이 예에서 보인 바와 같이 병렬문에서는 이러한 대칭성과 응집성의 성질이 항상 공존한다는 것을 알 수 있다.

본 논문에서는 한국어 병렬문에서 나타나는 중의성으로 인한 병렬 접속항 선정의 문제점을 해결하기 위해 병렬문의 분석을 통해 중의성의 원인을 파악한다. 이러한 중의성을 지닌 병렬문의 올바른 접속항에 대해 대칭성과 응집성을 통해 설명함으로써 이 두 가지 성질을 통해 올바른 접속항이 선정될 수 있음을 보이고 두 성질간의 관계에 대해 살펴본다.

2. 병렬문의 분석

병렬문에서 생성되는 병렬 후보들 중에 올바른 후보를 선택하기 위해서는 병렬항 구조의 선정에 영향을 줄 수 있는 요소들과 다수의 병렬 후보를 생성하는 중의성의 원인을 파악하는 것이 필요하다. 본 절에서는 병렬구조의 선정에 필요한 성질을 도출하기 위해 병렬문의 분석을 통해 이들을 살펴본다. 표 1은 나타나는 병렬문의 유형으로 한국어에서 병렬문이 광범위하게 사용되며 이에 대한 분석도 다양하게 이루어져야 함을 보여 준다. 이 표는 KAIST 구문 분석된 말뭉치 총 321,550어절 31,109문장으로부터 추출한 병렬문 14,641문장과 전자상거래와 관광상품정보 분야의 자연언어질의 시스템 구축을 위해 질의문 원시 말뭉치 2,130문장 중에서 병렬문 424문장을 대상으로 하였다 (Lee and Park, 2002). 특히 KAIST 말뭉치는 해설서, 일반 소설, 역사, 산문, 지침서 등의 일반 분야를 대상으로 하므로 전체 문장 중 47%가 병렬문이라는 사실은 자연언어에서의 병렬문 처리의 중요성을 내포하는 수치라고 할 수 있다. 또한 질의 말뭉치는 전문 분야를 그 대상으로 하는데 이러한 전문 분야의 질의문에서는 사람들이 문장 구조를 간단하게 표현하려는 경향이 있어 전체 문장 중 병렬문의 비율이 20%로 상대적으로 작게 나타났다. 한국어에서 병렬문은 크게 체언연결과 용언연결로 나눌 수 있는데 체언연결에 대해 문의 연결로 보는 견해 (최재희, 1985a), 구의 연결로 보는 견해 (김진수, 1987), 문의 연결과 구의 연결로 보는 견해 (김영희, 1988)가 있는 반면 용언연결은 공통적으로 문의 연결로 인정하고 있다. 이와 함께 한국어에서는 보통 주어나 목적 등과 같은 주변성분은 때에 따라 탈락되기도 하지만 중심성분인 서술어는 탈락되지 않는데 병렬문에서는 서술어가 탈락되기도 한다 (이관규, 1992). 용언연결에서 다양하게 발생하는 이러한 탈락을 고려하기 위해 이 표에서는 병렬항간에 주어를 공유하는 용언 병렬을 동사구 병렬로 분류하였고 각각 주어를 따로 가지는 용언 병렬과 독립된 문장간의 병렬을 문장 병렬로 분류하였다. 그리고 서술어가 탈락된 병렬을 비정규성분 병렬에 포함시켜 분류하였다. 여기서 비정규성분 병렬은 정규 성분 단위가 아닌 성분간의 병렬을 의미하고 표에서의 수식어구는 관형사구나 부사구간의 병렬을 의미한다. 또한 견해에 따라 대등접속소에 나, 지만과 같은 역접접속어미를 포함

통사범주	KAIST	Query
명사구	54.6% (7,997/14,641)	82.1% (348/424)
동사구	47.0% (6,882/14,641)	6.8% (29/424)
문장	16.9% (2,329/14,641)	12.5% (53/424)
수식어구	0.1% (12/14,641)	0.2% (1/424)
형태소	0.0% (2/14,641)	0.2% (1/424)
비정규성분	0.1% (15/14,641)	0.7% (3/424)
병렬문 (/총문장수)	47.1% (14,641/31,109)	20.0% (424/2,130)

[표 1] 통사범주별 병렬문 발생빈도

할 수 있는지는 여러 가지 견해가 있지만 이 표에서는 포함시켜 분류하였다 (서정수, 1985b; 서정수, 1996; 이관규, 1992; Cho, 2000).

2.1 병렬구조 선정의 요소

병렬 구조의 선정이 어려운 이유 중 하나는 형태소, 통사, 의미, 억양, 문맥 등의 단계에서의 다양한 요소들이 이에 영향을 미칠 수 있기 때문이다. 이러한 요소들을 분석하는 것은 올바른 병렬 구조의 선정에 필수적인 작업으로 본 절에서는 이러한 요소들이 병렬항 선정에 어떤 영향을 주는지를 살펴보기로 한다.

2.1.1 형태소적 요소. 한국어는 한자어나 조사, 어미 등에 의한 형태소적 요소의 활용이 다양하게 나타나는데 이는 병렬문에도 그대로 반영이 된다. 다음은 병렬문에서 나타나는 형태소적 요소의 예를 보이고 있다.

3. (a) 냉, 난방 시설

(b) 차익이 생산자와 유통업자의 몫으로 흡수 되었다.

(c) 놀이와 먹기, 잠자기가 하루 일과의 전부이다.

(d) 철수가 책을 좌측으로, 영희가 서류를 우측으로 옮겨라.

예 3a에서는 형태소단위로 병렬이 일어 나는 것을 보여주고 있다. 이 경우 병렬이 되는 형태소의 의미가 반대가 되어 대칭적으로 나타나 병렬 구조를 명확히 하고 있다. 3b에서는 접미어 자가 공통적으로 사용되어 병렬항간 의미적인 유사성을 높임으로써 병렬 구조의 중의성을 줄이고 있다. 3c에서도 파생접미사 기에 의한 어미활용의 유사성을 보여주고 있다. 이러한 점을 이용하기 위한 병렬항 어휘간의 일부 문자열을 비교하여 이를 병렬 구조의 판단에 적용할 수 있다 (Okumura and Muraki, 1994; Kurohashi and Nagao, 1994). 3d는 비정규성분 병렬을 보이고 있는데 양쪽 병렬항에서 조사의 대칭성으로 병렬이 되는 대상을 지정할 수 있음을 보이고 있다. 이와 같이 한국어에서는 통사적인 기능이나 특수한 의미의 표지의 역할을 하는 조사나

어미는 병렬문의 구조를 알아내는 중요한 요소로 사용된다 (Park and Park, 1999). Resnik (1999)는 영어 복합명사에서 병렬항간에 명사의 수에 의한 대칭성을 이용하였고 Kurohashi and Nagao (1994)에서는 일본어 병렬문에서 나타나는 기능어간의 대칭적 요소를 이용하였다.

2.1.2 통사적 요소. 병렬문에서 통사적 정보에 해당하는 통사범주나 자질 정보는 병렬 규칙의 제약조건으로 사용될 수 있다. 즉 영어에서의 시제나 수의 일치를 위한 자질이나 한국어에서도 나타나는 격표지나 허가 조건을 위한 자질 등은 병렬규칙에도 적용되어 비문을 제약하는 수단으로 사용될 수 있지만 기본적으로 병렬규칙에 의해 발생하는 중의성을 해소하지는 못한다. 이는 다음 예 4에서도 쉽게 알 수 있다.

4. (a) *맛있는 [NP_{nom} 사과가] 그리고 [NP_{acc} 배를] 먹었다.¹

(b) [NP_{acc} 맛있는 [NP_{acc} 사과를]] 그리고 [NP_{acc} 배를] 먹었다.

예 4a에서 격표지에 의해서 병렬문이 성립되지 못하는 예이고 4b에서는 격이 일치하지만 수식어에 의해서 중의성이 발생하는 경우를 보이고 있다. 이 예에서와 같이 병렬문의 구조적인 이유에서 유발되는 중의성은 통사적인 조건만으로는 구분하기 어렵다. 단지 영어 명사구 병렬항간에도 수의 일치와 같은 특징이 하나의 선호조건으로 사용될 수는 있다 (Resnik, 1999). 이와 비슷하게 통사범주의 대칭성도 선호되는 요소로 작용할 수 있다. 5에서는 이러한 예를 보이고 있는데 5a는 대칭적인 수식구조를 보이고 있다.

5. (a) 오스트레일리아의 면양이나 네덜란드의 홀스타인도 그렇다.

(b) 철수가 책을 좌측으로, 영희가 서류를 우측으로 옮겨라.

(c) Consumer prices for the first 10 months of 1989 surged by 5% and wholesale prices by 1.3%.

5b는 앞서 4b를 다시 보인 예인데 이와 같은 비정규성분 병렬에서는 문장 성분간의 구조적인 대칭이 강하게 나타남을 알 수 있다. 5c와 같은 영어의 간극(gapping) 현상이 발생하는 비정규성분 병렬에서도 이와 같은 사실을 발견할 수 있다. 이렇게 병렬 구조를 판별하는데 문장에서 나타나는 통사 구조의 대칭적 요소를 이용할 수 있다 (Agarwal and Boggess, 1992; Okumura and Muraki, 1994). Park (1998)에서는 품사의 패턴만을 이용하여 병렬 구조를 판단하기 위한 시도를 하였다.

2.1.3 의미적 요소. 병렬문에서 나타나는 중의성은 통사 단계에서 일어나므로 이를 판별하기 위해서 의미 단계의 정보를 도입해야 할 필요성이 제기된다. 이러한 점은 앞서 살펴보았던 형태소적 요소나 통사적 요소들이 의미 단계의 요소들에 영향을 미치어 병

¹ *는 통사적으로 허용되지 않음을 나타낸다.

렬항의 구조를 판단하는데 도움을 줄 수 있기 때문이다. Munn (1993)은 비정규성분 병렬이나 상이범주 병렬에서 일어나는 통사적 비대칭성을 근거로 병렬규칙이 의미 범주에 대한 것이라는 점을 주장하였다. 이 주장은 그 전산처리 방법의 구현에 있어서는 어려움이 예상되지만 병렬문의 처리에 있어서 의미 정보의 이용이 중요하다는 점을 지적하고 있다. 다음의 예 6에서 병렬에 나타나는 의미적 요소를 살펴보도록 한다.

6. (a) 주가와 지가가 연쇄적으로 올라갔다.
 (b) 엘지 냉장고와 삼성 텔레비전
 (c) 엘지 21인치 평면 텔레비전과 디오스 냉장고
 (d) 기존의 관행인 현물거래가 사라지고 선물거래가 도입되어 투기성이 커졌다.

이 예에서 6a-6c는 명사구 병렬을 보이고 있다. 6a에서는 앞서 설명한 바와 같이 형태소적인 요소가 의미적 요소에 영향을 미칠 수 있다는 것을 보이고 있다. 6b와 6c의 경우는 복합명사의 병렬을 보이고 있는데 이와 같은 복합명사는 명사의 연속으로 의해 구성되기 때문에 그 구성 어휘들 간의 관계를 파악하기가 쉽지 않다 (Kobayasi, Tokunaga, and Tanaka, 1994; Yoon, Jong, and Song, 1998). 이러한 특징은 병렬문에서도 나타나서 복합명사를 구성하는 어휘들이 명사의 연속으로 수식관계를 이루기 때문에 구조적 중의성이 증가하는 요인이 된다. 복합명사로 접속항이 구성될 경우에는 올바른 접속항의 후보를 결정하기가 쉽지 않다. 6b와 같이 양쪽의 접속항이 의미적으로 대칭된 형태로 나타날 수도 있지만 6c의 예에서와 같이 비대칭적으로 나타날 수도 있는데 6b의 경우도 대칭적인 구조를 이룬다는 것을 파악하기 위해 단어의 의미를 비교하는 것이 필요하다. 특히 6c와 같은 경우는 병렬항간의 의미적 대칭성 뿐만 아니라 병렬항에 의해 공유되는 엘지의 의미적 결합성을 파악하여 그 범위를 결정해야 한다. 6d는 서술절간의 병렬을 보이고 있는데 이 예에서는 문장 성분 현물거래와 선물거래, 사라지고 도입되는 대칭성과 기존의 관행인의 현물거래에 대한 수식 관계를 통해서 병렬 구조를 파악할 수 있다. 이러한 의미적 대칭성을 파악하기 위해 어휘의 의미적 유사성을 기반으로 미리 정의된 의미 범주나 의미 계층과 같은 정보를 이용한다 (Agarwal and Boggess, 1992; Kurohashi and Nagao, 1994). 또 다른 의미적 요소로 단어의 결합성을 이용하는 연구도 있다 (Resnik, 1999).

2.1.4 문맥 및 억양적 요소. Cremers (1993)은 병렬에서 양쪽 접속항이 의미적 차이를 전제하며 이러한 차이는 많은 경우에 억양에 의해 구분될 수 있다는 점을 지적했는데 이러한 의미적 차이는 Steedman (2000)에서의 병렬규칙에서도 지적되고 있다. 또한 병렬에 의한 중의성이 존재하는 문장에서 병렬 구조는 이러한 억양 정보나 문맥 정보에 의해 명확해질 수 있다. 이러한 점을 예문 7을 통해 알 수 있다.

7. (a) [_SThe model embraced [_{NP}the designer]] and [_S[_{NP}the photographer] laughed].

(b) [_S 영희가_i [_{VP} 철수는 사랑하]]고_j [_{VP/S} 상호는 좋아한다].

예 7a는 embraced와 laughed가 같은 시제의 형태를 지님으로써 명사구 병렬 또는 문장 병렬이 일어나는 구조적 중의성을 지닌다. 한국어에서도 7b와 같이 동사구 병렬과 문장 병렬에 의한 구조적 중의성이 발생할 수 있다. 7a의 예에서는 의미적인 분별만으로는 명사구 병렬인지 문장 병렬인지를 판별하기 힘들지만 일반적으로 독자는 명사구 병렬에 대한 선호 경향이 있는 것으로 보여지며 이를 설명하기 위한 노력이 있었다 (Frazier, 1987). Hoeks, Vonk, and Schriefers (2002)는 이러한 중의성이 있는 문장에서 주제 구조(topic structure)의 분석을 기반으로 하여 더욱 간단한 주제 구조로 사람이 해석하고자 하는 경향이 있다는 점을 지적하였다². 이러한 점은 병렬 구조의 결정에 있어 화용 정보도 하나의 요소로 작용될 수 있음을 보여주는 연구로 병렬 구조의 결정에 다양한 단계의 정보들이 관여되고 있음을 알 수 있다. 이와 비슷하게 7b에서 억양에서 나타나는 휴지 정보가 *i*나 *j*의 위치에 따라 병렬 구조가 다르게 파악될 수도 있다는 것을 알 수 있다. 이러한 휴지 정보는 담화 분석을 통한 화제(theme)/평언(rheme) 구조에 따라 파악할 수 있다 (Steedman, 2000; Lee and Park, 2001).

위에서 논의한 바와 같이 병렬 구조를 결정에 있어서는 형태소, 구문, 의미, 문맥 등 다양한 단계의 요소들이 작용될 수 있음을 보였다. 본 연구에서는 이러한 요소들을 기반으로 병렬 구조의 범위를 파악하기 위하여 이러한 특징들이 모델링될 수 있는 일관성 있는 병렬문형의 성질을 파악하고 병렬문을 전산처리할 때 생성되는 가능한 후보들에 대해 이를 적용함으로써 병렬에 의한 구조적 중의성 문제의 해결을 시도하고자 한다.

2.2 병렬문의 구조적 중의성

영어의 전치사구에서 발생하는 구조적 중의성은 오분석을 유도하고 처리에서의 복잡도를 증가시키는 것으로 알려져 있다(Hindle and Rooth, 1993; Resnik and Hearst, 1993). 이러한 문제는 복합명사에서도 나타나는데 복합명사에서는 같은 통사 범주가 계속 반복해서 나오기 때문에 구조적 중의성이 일어나게 된다(Lauer, 1995; Barker and Szpakowicz, 1998). 하지만 병렬문에서는 이러한 전치사구나 복합명사를 접속함으로써 가질 수 있기 때문에 전치사구나 복합명사에서 발생하는 구조적 중의성을 포함할 수 있으므로 병렬문에서 더욱 구조적 중의성 문제가 심각할 수 있다. 병렬문에서 발생하는 구조적 중의성은 다음과 같은 네 가지 유형으로 분류될 수 있다. 표 2는 각 유형들이 병렬문에서 발생하는 빈도수를 보인 것이다. 이 표는 앞서 보였던 KAIST 말뭉치의 병렬문 중 874문장과 질의문 말뭉치의 병렬문 중 142문장에 대해 자동으로 분류한 후 수동으로 확인하여 분석한 것이다. 이 표의 각 유형별 빈도수는 중첩된 유형에 의한 중의성이 65%로 한 가지 유형에 의한 32%보다 높게 나타나 병렬에 대한 중의성의 원인이 복잡하게 발생함을 보여준다. 또한 중의성이 없는 경우는 2.5%로 병렬문에

² 기존의 문맥 정보가 없다고 가정했을 때, 명사구 병렬 구조는 주제의 역할을 하는 주어와 하나이고 문장 병렬 구조는 주제가 두개가 나타난다.

중의성 유형	KAIST	Query	합계
(1) 수식관계	19.1% (167/874)	12.7% (18/142)	18.2% (185/1,016)
(2) 논항구조	3.8% (33/874)	2.1% (3/142)	3.5% (36/1,016)
(3) 연속된 병렬	0.3% (3/874)	0.0% (0/142)	0.3% (3/1,016)
(4) 다가능 병렬접속사	10.0% (88/874)	12.0% (17/142)	10.3% (105/1,016)
합계	33.3% (291/874)	26.8% (38/142)	32.4% (329/1,016)
(1) + (2)	12.7% (111/874)	3.5% (5/142)	11.4% (116/1,016)
(1) + (3)	1.1% (10/874)	0.0% (0/142)	1.0% (10/1,016)
(1) + (4)	29.5% (258/874)	52.1% (74/142)	32.7% (332/1,016)
(2) + (3)	0.1% (1/874)	0.0% (0/142)	0.1% (1/1,016)
(2) + (4)	0.8% (7/874)	0.0% (0/142)	0.7% (7/1,016)
(3) + (4)	0.2% (2/874)	1.4% (2/142)	0.4% (4/1,016)
(1) + (2) + (3)	3.1% (27/874)	0.0% (0/142)	2.7% (27/1,016)
(1) + (2) + (4)	3.5% (31/874)	0.0% (0/142)	3.1% (31/1,016)
(1) + (3) + (4)	9.3% (81/874)	7.0% (10/142)	9.0% (91/1,016)
(2) + (3) + (4)	0.0% (0/874)	0.0% (0/142)	0.0% (0/1,016)
(1) + (2) + (3) + (4)	4.9% (43/874)	0.0% (0/142)	4.2% (43/1,016)
합계	65.3% (571/874)	64.1% (91/142)	65.2% (662/1,016)
중의성 없음	1.4% (12/874)	9.2% (13/142)	2.5% (25/1,016)
합계	100.0% (874/874)	100.0% (142/142)	100.0% (142/1,016)

[표 2] 중의성의 유형에 대한 병렬문 빈도수

서 중의성이 없는 경우는 거의 발생하지 않는 것으로 분석되어 중의성의 해소가 병렬문 분석에서 중요한 문제임을 보여준다.

수식 관계에 의한 중의성

이 유형은 표 2에서 보인 것과 같이 중의성의 유형들 중 가장 자주 발생하는데 명사나 관형사, 서술어, 부사 등과 같은 수식어에서부터 구나 절 등에 이르기까지 수식의 역할을 할 수 있는 표현들에 의해 같은 범주가 반복됨으로써 일어나는데 이런 수식어의 병렬항에 대한 공유 여부에 따라 후보들의 병렬 구조가 결정된다. 앞서 보인 6b나 6c의 명사구 병렬에서는 명사 수식에 의한 중의성을 나타낸다. 수식어와 명사로 이루어진 명사구에서 내부적인 의미 구조상의 중의성을 고려하지 않고 병렬 관계에 의한 중의성만을 고려할 때 다음과 같은 명사구 패턴에서 n 과 m 이 양쪽 병렬항의 명사 개수이면 $n \cdot m$ 의 비율로 중의성이 증가한다³.

³ 복합명사의 명사간의 관계는 복잡하기 때문에 이에 대한 분석은 중요한 연구 주제이다 (Lauer, 1995; Yoon, Jong, and Song, 1998). 수학적 관점에서 순수한 복합명사의 경우 그 의미 구조가 영

$$NP_1 NP_2 \dots NP_n \text{ and } NP_1 \dots NP_m$$

위의 패턴은 단지 설명을 위한 것으로 이 패턴에서 수식 관계에 있는 명사는 가능한 다른 수식어로 대체될 수 있다. 따라서 일반적으로 양수인 n 과 m 은 양쪽 병렬항에 나올 수 있는 수식어를 포함한 어휘의 개수를 나타낸다.

논항 구조에 의한 중의성

문장 병렬이나 동사구 병렬에서 논항이 양쪽 병렬항의 서술어와 관계가 있는지 혹은 한쪽에만 관계가 있는지에 따라 중의성이 발생할 수 있다. 즉 서술어와 논항의 목적어 관계나 주어 관계가 병렬에 의해 중의성이 나타날 경우 병렬 구조에서도 중의성이 나타나게 된다. 이러한 예는 앞서 보인 7에서 나타나는데 영어에서의 간극이 나타난 문장이나 7a의 명사구/문장 병렬이 있는 문장, 한국어에서의 7b와 같은 동사구/문장 병렬이 있는 문장이 이에 해당하게 된다. 그러므로 이 유형은 다른 요인에 의한 중의성이 없을 때 동사구/문장 병렬의 경우 2개의 병렬 구조 후보를 생성한다. 그러나 이 유형은 보통 다른 요인에 의한 중의성을 동반하므로 동사구 병렬이나 문장 병렬의 양쪽에서 발생하는 중의성에 의한 후보만큼 총 후보의 개수가 나타나게 된다.

연속된 병렬에 의한 중의성

콤마나 와/과/나 등과 같은 대등 접속사에 의한 병렬이 연속되어 발생할 때 이러한 병렬 관계에 의해 그 구조가 중첩된 범위를 가질 수 있다. 예를 들어, NP and NP and NP와 같은 통사 패턴의 경우 and를 이진 관계로 보면 $[[NP \text{ and } NP] \text{ and } NP]$, $[NP \text{ and } [NP \text{ and } NP]]$ 와 같은 병렬 구조의 후보를 가지게 된다. 예제 8에서 8a는 대등 접속사 와와 과를 갖고 이 문장은 8b와 같은 후보들을 가진다. 이러한 유형의 중의성에 대해 다음과 같은 명사구 패턴을 고려해 본다.

$$NP_1 \text{ and } NP_2 \text{ and } \dots \text{ and } NP_n$$

양수 n 이 증가함에 따라 병렬항을 이루는 명사구의 내부적인 의미 구조를 고려하지 않을 때 위의 연속된 병렬 패턴은 병렬 구조의 후보 개수를 Catalan number의 비율로 증가시킨다고 알려져 있다⁴ (Church and Patil, 1982).

8. (a) 부동산 시장의 변화와 주식 시장의 활황과 불황

(b) [부동산 시장의 변화]와 [주식 시장의 [활황과 불황]]

[[부동산 시장의 변화]와 [주식 시장의 활황]]과 [불황]

어의 전치사 자체에 대한 의미를 고려치 않은 전치사구 패턴과 동일하다고 보면 그 중의성의 개수는 Catalan number인 1, 1, 2, 5, 14, 42, 132, 469의 비율로 증가한다 (Church and Patil, 1982). 이 논문에서는 병렬항의 내부적인 의미 구조에 대해서는 고려하지 않기로 한다.

⁴ Church and Patil (1982)는 영어의 전치사구를 가지는 문장의 분석(parse) 개수가 지수로 증가하는 괄호삽입된 수식 표현의 개수와 같은 비율로 증가한다는 점을 보였다. 여기서 개수는 조합 함수인

Catalan 함수 $\binom{2n}{n} - \binom{2n}{n-1}$ 에 의해 계산될 수 있다.

[부동산 시장의 변화]와 [주식 시장의 활황]과 [불황]

대등 접속사의 다기능에 의한 중의성

영어나 한국어에서 콤마는 종속절, 대등절, 삽입 등의 다양한 통사적 기능으로 사용될 수 있다 (Bayraktar, Say, and Akman, 1998). 이러한 다양한 기능은 문장에서 통사 정보만으로 분석이 불가능한 경우가 있다. 한국어에서 대등 접속사로 자주 사용되는 와/과, 고도 각각 공동격 조사나 종속 접속사와 같은 기능을 가지고 있는데 그 예로 9에서 와/과에 대해 9a의 대등 접속사의 역할과 9b의 공동격 조사의 역할을 보이고 있다. 9에서 보인 바와 같이 두 가지 기능에 대해 와의 통사적 환경이 같기 때문에 통사 정보만으로 대등 접속사 와의 기능을 구분할 수 없으므로 이에 의해서 중의성이 나타나게 된다. 이 유형의 중의성은 주어진 문장에 대해 대등 접속사의 여러 통사 범주 중에 문장으로 분석이 되는 수만큼 문장 구조의 후보를 가지게 된다.

9. (a) 사과와 빵을 먹었다.

(b) 친구와 빵을 먹었다.

이러한 중의성의 유형 중에 마지막 다기능 대등 접속사에 의한 유형은 다른 유형들에 비해 그 중의성의 원인이 다른데 나머지 유형들이 병렬 관계에 의한 통사 구조상의 중의성을 갖는데 반해 마지막 유형은 어휘의 통사 범주 상의 중의성에 의해 유발되기 때문에 병렬 관계와 관련 없는 문장 구조의 후보들을 생성할 수 있다. 그러나 이러한 중의성은 영어에서 발생하는 전치사구의 구조적 중의성도 전치사의 통사 범주의 문제까지도 포함하여 다루고 있는 것을 고려하여 구조적 중의성의 한 유형으로 분류하였다 (Hindle and Rooth, 1993).

표 2는 이러한 구조적 중의성의 유형과 이와 함께 발생하는 경우에 대한 문장에서의 발생 빈도도 보이고 있다. 이렇게 한 문장에서 여러 유형이 중복해서 발생하게 되면 병렬 구조의 후보의 수가 급격히 증가할 뿐만 아니라 병렬 구조의 범위를 알아내는 데에도 어려움을 가중시킬 수 있다. 이렇게 유형이 중복된 경우에는 그 통사 패턴이 다양하기 때문에 이러한 경우에 대한 후보의 개수를 계산하기가 어렵지만 표 3에서 보인 각각의 유형에 대한 후보의 개수로부터 예측해 볼 수 있다. 특히 표 3에서의 유형 (1)과 (3)은 후보의 개수를 증가시키는 비율이 높으므로 유형 (1)과 (3)이 다른 유형과 함께 문장에서 발생할 경우 병렬 구조의 후보는 급격히 증가할 것으로 예상된다. 참고로 표 3에서 각 병렬 후보의 증가 비율은 예제 패턴에서 보인 명사구의 내부 의미 구조를 고려하지 않은 것이다.

3. 병렬항 선정을 위한 특성

앞서 살펴본 병렬항 선정을 위한 요소들을 이용하여 병렬문의 다양한 중의성에 의해 생성되는 병렬 구조의 후보들 중 올바른 범위를 가진 후보를 선택하기 위해서

중의성 유형	예제 패턴	생성 후보 수
(1) 수식 관계	$NP_1 NP_2 \dots NP_n \& NP_1 \dots NP_m$	$n \cdot m$
(2) 논항 구조	$NP NP VP \& NP VP$	2
(3) 연속된 병렬	$NP_1 \& NP_2 \& \dots \& NP_n$	Catalan numbers
(4) 다가능 병렬접속사	-	$\#Cat_{\&}$

- 접속사는 & 으로 표시.
- NP와 VP는 각각 명사구와 서술어구를 가리킨다.
- $\#Cat_{\&}$ 은 병렬 접속사 & 에 의해 문장으로 인식되는 범주의 개수.

[표 3] 구조적 중의성의 유형에 따른 병렬구조의 후보 수

는 여러 요소들을 적용할 수 있는 방법이 필요하다. 그러나 이러한 방법의 모델링을 위해서는 병렬문의 올바른 병렬 구조가 가지는 일관된 성질을 파악해 내는 작업이 요구된다. 대칭성과 응집성은 대부분의 문법에서 채용하는 병렬성에 기반한 규칙인 $X \rightarrow X'$ and X'' 로부터 유도되는 성질로 본 절에서는 앞서 분류한 병렬 구조의 중의성을 대칭성과 응집성으로 설명함으로써 이 두 가지가 올바른 병렬 구조의 선정에 여러 요소들과 함께 일관되게 적용될 수 있음을 보이도록 한다.

3.1 병렬항의 대칭성

병렬 규칙에 기반한 병렬항의 대칭성은 앞서 보인 예제들에서 병렬항 사이의 다양한 요소들간에 대칭으로 나타나고 있다. 이러한 대칭은 병렬항에서의 형태소간, 어휘간, 병렬항간의 통사 범주나 의미에 있어서 다양한 대칭 관계를 구조적으로 이루고 있기 때문에 이러한 대칭성을 구조적 유사성이라고 표현하기로 한다. 이러한 구조적 유사성은 병렬 구조에 대해 다음과 같은 관계로 정의한다.

구조적 유사성 병렬항이 형태소적, 통사적, 의미적으로 서로 대칭적일수록 병렬 구조가 명확해진다.

이 관계는 선호되는 병렬 구조일수록 대칭적인 경향을 갖는다는 것으로 영어에서 나타나는 상이범주 병렬은 비록 통사 범주는 일치하지 않지만 의미적으로 대칭적인 경향이 강하다는 것을 Munn (1993)에서 지적한 바 있다. 실제 예를 통해 구조적 유사도의 이러한 관계를 살펴 보도록 한다.

10. (a) 철수가 책들을 좌측으로, 영희가 서류를 우측으로 옮겨라.
- (b) 철수가 책들을 좌측으로, 영희가 우측으로 옮겨라.
- (c) 철수가 책들을 좌측으로, 서류를 우측으로 옮겨라.
- (d) 철수가 책들을 좌측으로 옮겨고 영희가 서류를 우측으로 옮겨라.

10a는 앞서 보인 비정규성분 병렬의 예로 양쪽 접속항간에 논항의 조사간 대칭성으로 전체적인 병렬항의 대칭 관계가 강하게 드러난다. 10b는 문장 성분들간의 대칭

성으로 인해 특별한 문맥 정보가 없을 경우 기본적으로 책들이 우측 병렬항에서 공유 혹은 생략된 성분으로 독자에게 이해될 가능성이 많다. 10c도 마찬가지로 주어를 공유하는 경우를 보이고 있다. 10d 예에서와 같이 문장 성분의 생략이 일어나지 않은 경우를 생각해 볼 때 양쪽 접속항간의 대칭성은 더욱 두드러지게 나타난다. 이러한 대칭성은 병렬이 일어나는 양쪽 항의 구조가 동일함에 따라 발생하므로 병렬 구조를 명확히 파악할 수 있도록 한다. 다음의 예를 통해 이러한 비정규성분 병렬에 대해 좀 더 자세히 살펴보도록 한다.

11. (a) 영희가 빵을, 철수가 우유를 먹었다.
 (b) 영희가 빵을, 우유를 철수가 먹었다.
 (c) #영희가 빵을, 철수에게 우유를 먹었다.⁵
 (d) #영희가 빵을, 철수에게 우유를 먹었다.
 (e) 영희가 빵을 그리고 우유를 철수에게 먹었다.
12. (a) Dexter eats beans, and Warren, potatoes.
 (b) #Dexter eats beans, and potatoes, Warren.

11, 12 에서 보인 예는 비정규성분 병렬에서 나타나는 병렬문의 예인데 일반적으로 이러한 병렬문은 한쪽 접속항에서 동사가 생략되어 발생한다. 이 예문들에서는 11a와 같이 논항이 뭉치로 나타나게 되어 구문적 대칭성이 접속항간에 강하게 일어남을 알 수 있다. 한국어에서는 이러한 구문적 대칭성이 조사를 통해서 나타나는데, 11b와 같은 경우에도 위치와 관계없이 일어날 수 있다. 반면에 11c, 11d와 같이 다른 격의 조사가 사용되어 격정보에 의한 구문적 대칭성이 깨진 경우에는 문장의 의미가 통하지 않는 것을 알 수 있다. 이러한 것은 11e와 같이 병렬되는 접속항의 구조를 동일하게 만들어 주면 병렬문이 가능해지는 것에서도 나타난다. 영어의 경우 12a와 같이 비정규성분 병렬이 일어날 수 있으며, 이 예문에서도 논항의 위치에 의한 구조적 유사성이 강하게 일어남을 볼 수 있다. 영어에서는 한국어와 다르게 접속항에서 어휘간의 대칭이 일어나는 논항의 위치가 중요함을 볼 수 있다. 따라서 12b와 같은 경우에는 의미적인 오류가 발생한다. 이러한 사실은 일반적으로 병렬 구조의 범위를 찾는 문제에 있어서 양 접속항의 대칭성을 비교할 때, 양 접속항에서 어휘들의 위치 정보도 고려되어야 하며 이러한 비교를 위해 접속항의 대응될 어휘들간의 정렬이 이루어져야 함을 의미한다. 이러한 점은 한국어에서는 논항의 위치가 비교적 자유롭게 나타날 수 있지만 그 대신 격표지 정보를 이용하여 대응 어휘간의 비교를 할 수 있다. 그러나 보어를 필요로 하는 문장간의 병렬에서는 주어와 보어간의 순서가 바뀔 수 없고 보어에 대한 격조사와 주어에 대한 격조사가 일치할 수 있다. 또한 한국어에서는 조사 정보가

⁵ #은 의미적으로 허용되지 않는 경우를 나타낸다. 이 경우에는 격표지에 의하여 통사적으로도 허용되지 않을 것으로 보인다.

빈번히 생략되어 나타난다는 점을 고려할 때 이러한 경우 논항의 순서 정보가 중요하게 고려되어야 함을 알 수 있다. 병렬항간에 나타나는 구조적 동일성 (또는 유사성)은 13의 병렬문 예에서도 살펴볼 수 있다.

13. (a) 냉, 난방 시설
 (b) 이태원과 왕십리 지역
 (c) 엘지 냉장고와 삼성 텔레비전
 (d) 기름기가 있는 코 양쪽이나 이마

13은 수식어에 의한 중의성이 있는 병렬문의 예를 보이고 있다. 13a는 앞서 보인 형태소 병렬의 예인데 형태소 냉과 난의 의미적 대칭성에 의해 병렬 구조가 드러난다. 13b에서는 지역명을 나타내는 고유 명사의 대칭성으로 부터 지역이 공유되는 병렬 구조를 알 수 있다. 13c는 복합 명사간의 구조도 대칭적이지만 복합 명사를 이루는 어휘들의 의미상 대칭 관계로 그 병렬 구조가 명확해 진다. 13b와 13c를 통해 복합 명사에서는 그 통사 패턴의 대칭성보다 의미적인 대칭 관계가 병렬 구조를 판별하는데 더욱 유용하다는 것을 알 수 있다. 이러한 점은 13d에서도 알 수 있는데 코와 이마의 의미간 대칭성을 통해 병렬 구조를 파악하는데 도움이 될 수 있다.

14. (a) [부동산 시장의 변화]와 [주식 시장의 [활황과 불황]]
 (b) [[부동산 시장의 변화]와 [주식 시장]의 [[활황]과 [불황]]
 (c) [[부동산 시장의 변화]와 [주식 시장의 활황]]과 [불황]

14는 앞서 보였던 연속된 병렬 접속사에 의한 중의성이 나타나는 예이다. 이 예에서는 실제 복합 명사나 관형격 조사의 사용에 따른 수식어에 의한 중의성이 함께 나타나기 때문에 발생하는 병렬 구조의 후보들은 더욱 많지만 그 중 일부를 보인 것이다. 이 예에서는 좌우 병렬항에서 어휘간에 부동산과 주식, 시장과 시장, 활황과 불황의 대칭성이 발생하게 되는데 이를 통해 부동산 시장과 주식 시장, 활황과 불황이 병렬되는 구조를 파악할 수 있다. 또한 관형격 조사 의의 대칭에 의한 수식 구조로 부터 변화와 활황과 불황이 의미적으로 대칭 관계일 가능성이 높다는 사실을 파악할 수 있다. 이러한 대칭성을 이용하여 14a의 구조가 14b나 14c보다 더욱 선호되는 병렬 구조로 파악할 수 있다.

15. (a) 사과와 빵을 먹었다.
 (b) 친구와 빵을 먹었다.

예 15은 외에 의한 범주의 중의성이 존재하는데 와가 공동격 조사로 사용될 경우 일반적으로 함께나 같이 등과 같은 부사와 사용되거나 교호성(reciprocity)이 있는

동사들과 함께 사용되어 공동의 의미를 명확히 하는 경우가 많다. 그러나 15b와 같은 경우에는 와가 병렬의 용법으로 사용되어도 통사적으로는 허용될 수 있다. 이러한 차이는 구분해 주기 위한 방법으로 병렬항의 의미적인 대칭성을 고려해 볼 수 있는데 15a와 같이 병렬 접속사로 사용된 경우에는 의미적으로 대칭되지만 15b와 같은 경우에는 의미적인 유사성이 낮아서 대칭성이 떨어지게 된다. 이와 비슷하게 고도 다양한 용법을 지니는데 특히 접속사로 사용되는 경우와 보조 연결어미로 사용되는 경우에 중의성이 존재하게 되는데 이 경우에는 연결어미로 사용되는 고의 용법이 한정되어 접속사의 용법과 구분하기가 비교적 용이하다.

16. (a) 사랑하고 존경을 구별하여라.

(b) 사랑하고 존경을 구별하여라.

하고에서 하가 동사파생접사로 사용되는 경우와 하고가 병렬 접속사로 사용되는 경우에 의해 중의성이 일어난다. 이에 의한 예를 16에서 보이고 있는데 이 경우에도 의미적 대칭성이 의해 병렬 구조의 선정에 도움을 줄 수 있을 것으로 고려된다. 이 예는 16a의 명사구 병렬과 16b의 동사구 병렬로 16b의 경우 사랑하와 구별하의 의미적 대칭성이 16a의 경우보다 떨어지기 때문에 16a의 경우가 선호되는 것으로 고려된다.

3.2 병렬항의 응집성

병렬항은 서로 병렬 환경 외부에 나타난 어휘를 공유하지 않는다고 고려할 때 병렬항 안의 요소들은 서로 의미적으로 응집된다고 생각할 수 있다. 이러한 점은 병렬항의 범위를 판단하는데 있어 병렬항간에 병렬 환경 외부에 공통적으로 적용되는 조건이라고 할 수 있다. 이러한 병렬의 성질을 병렬항 응집성이라고 하고 선호되는 병렬항에 대한 관계를 다음과 같이 정의한다.

병렬항 응집성 병렬항이 형태소적, 통사적, 의미적으로 응집될수록 병렬 구조는 명확해진다.

이 관계는 병렬 규칙에 기반한 것으로 각 병렬항에서 어휘들의 관계가 통사적으로 동일한 범주로 응집되고 하나의 의미단위로 응집되어서 다른 접속항에 대해 공유되는 요소가 없다는 것을 말하고 있다. 여기서 통사 범주에 대한 응집은 결합범주문법에서 자연스럽게 설명될 수 있는 것으로 비정규성분 병렬에 있어서도 문법 규칙의 결합에 의해 한 통사 범주로 묶이게 된다 (Steedman, 2000). 또한 이러한 통사 범주의 응집은 병렬 규칙의 적용시에 요구되는 통사 조건으로 사용되기 때문에 여기서는 주로 의미 단계에서의 응집성에 초점을 두도록 한다. 이러한 응집성은 Milward (1994)에 의해 비정규성분 병렬의 처리를 위해 제안된 방법에서 전제로 하는, 비병렬 환경에 대해 병렬항이 공통되는 성질인 음운적 요소, 통사 구조, 어휘의 의미를 공유한다는 주장과도 연관이 되는데 본 논문에서의 응집성이 의미적인 부분과 병렬 구조와의 연관성을 강조한 것이라고 할 수 있다. 또한 Cremers (1993)에서도 병렬 규칙이 의미 단계에 더욱 잘 적용된다는 것을 보이기 위해 병렬항을 한 의미 단위로 간주하여 처리하였다. 병렬항 응집성의 올바른 병렬 구조에 대한 관계를 예문을 통해 살펴보기로 한다.

17. (a) 엘지 21인치 평면 텔레비전과 디오스 냉장고
 (b) 엘지 21인치 평면 텔레비전과 디오스 냉장고
 (c) 엘지 21인치 평면 텔레비전과 디오스 냉장고
 (d) 엘지 21인치 평면 텔레비전과 디오스 냉장고

17은 앞서 보였던 예문으로 복합명사로 이루어진 병렬문이다. 이 예는 명사들로 구성되어 있기 때문에 수식어에 의한 구조적 중의성이 심하여 그 병렬 구조의 후보도 많아지게 된다. 그러나 17a와 같은 병렬 구조를 선호하게 되는 이유는 접속항의 의미적인 응집성에 기인한다. 즉, 이 예에서 엘지는 디오스 냉장고의 제조 회사로 양쪽 병렬항에서 수식이 가능하므로 17a의 병렬 구조가 17b-17d 보다 더욱 강한 의미적 응집성을 좌,우측 병렬항에서 보이게 된다. 17b와 같은 병렬 구조에서 엘지는 우측 병렬항에서 수식이 일어날 수 있으므로 좌측 병렬항에서 의미적 응집을 방해하는 요소로 작용한다. 또한 17c와 같은 경우는 21인치가 우측 접속항에는 관계되지 않고 좌측 접속항에만 수식어로 나타날 수 있기 때문에 좌측 병렬항의 범위를 제대로 표현하고 있지 못하므로 17a보다 우측 접속항의 의미적 응집성이 떨어진다고 할 수 있다. 17d의 경우도 17c와 마찬가지로 이유로 응집성이 떨어지므로 17a가 후보들 중 가장 응집성이 높은 병렬 구조로 선택될 수 있다. 다음은 17의 병렬문에 변형을 가하여 병렬항에 대한 응집성에 대해 좀 더 자세히 살펴보도록 한다.

18. (a) 엘지 21인치 흰색 텔레비전과 디오스 냉장고
 (b) 엘지 21인치 흰색 텔레비전과 디오스 냉장고

18은 좌측 병렬항에 양쪽 병렬항을 모두 수식할 수 있는 명사 수식어 흰색을 추가한 예문으로 18b 보다 18a가 올바른 병렬 구조로 인식될 수 있는 이유는 흰색이 비록 우측 병렬항에 대한 수식이 가능 하지만 좌측 병렬항의 21인치에 의해서 흰색이 수식할 수 있는 범위가 제한되기 때문이다. 따라서 18a의 경우가 18b의 경우보다 의미적인 결속력이 더욱 강하다고 볼 수 있다.

19. (a) 수염이 긴 노인과 아이
 (b) 수염이 긴 노인과 아이
 (c) 수염이 긴 노인과 청년
 (d) 수염이 긴 노인과 신선

19는 관형사절에 의한 수식에 의해 중의성이 생기는 예문으로 19a 같은 병렬 구조가 19b 보다 선호되는데 그 이유는 수염이 아이보다 노인에 의미적인 연관성이 있기 때문이다. 비록 19b의 병렬 구조가 의미적으로 대칭 관계를 이루고 있지만 19a의

병렬 구조가 수염에 의해 응집성이 높아지게 된다. 이러한 점은 19c와 19d의 예에서 좀 더 살펴볼 수 있는데 특별한 문맥 정보가 가정되지 않을 경우 19b에서 19c, 19d로 갈수록 더욱 선호될 가능성이 높아진다는 것을 알 수 있다.

20. (a) 밥을 먹고 출발하여라.

(b) 영희가 철수는 사랑하고 상호는 좋아한다.

20은 논항 구조에 의한 동사구 병렬과 문장 병렬의 중의성이 있는 예문이다. 20a에서는 밥이 먹다의 논항으로 결합될 가능성이 높기 때문에 20a와 같은 병렬 구조가 선호됨을 알 수 있다. 20b에서는 영희가 양쪽 병렬항의 서술어에서 결합이 일어날 수 있으므로 서술어와의 의미적인 결속력으로 판별할 수 없는데 양쪽 병렬항의 조사가 대칭됨으로써 20b와 같은 구조가 선호된다. 그러나 이와 같은 예문에서는 앞서 논의한 것과 같이 문맥 정보의 역할이 중요할 것으로 보인다. 이러한 문맥 정보 중 앞에서 언급했던 휴지 정보는 병렬항의 의미적인 성분성과 연관되어 의미 단위로 인식될 수 있도록 하기 때문에 의미적인 응집성에 도움을 주는 정보로 작용할 수 있다.

21. (a) [1년은 365일], [하루는 24시간], [1시간은 60분], [1분은 60초]라는 개념

(b) [삼성과 엘지]의 [텔레비전과 냉장고]

(c) [부동산 시장의 변화]와 [주식 시장의 [활황과 불황]]

21은 연속된 병렬 접속사에 의한 중의성이 일어나는 예문이다. 21a에서는 콤마가 병렬 접속사로 사용되는데 각 병렬항이 대등한 관계로 묶이고 있다. 이 예에서 각 병렬항은 의미상으로 대칭적인 구조를 가지고 있지만 의미상으로 각 병렬항이 응집되어 있다. 특히 한국어에서 콤마는 이렇게 다수로 발생하는 대등한 관계의 병렬에 자주 사용되어 한 의미적인 단위로 묶일 수 있는 구문의 병렬을 표시하는 역할을 하게 된다. 21b와 21c는 연속된 병렬 접속사에 의한 중의성과 함께 관형형 어미 의가 사용되어 중의성이 증가하게 된다. 21b에서는 텔레비전과 냉장고가 삼성과 엘지에 의해 모두 수식될 수 있기 때문에 21b의 구조를 가지게 되고 21c에서는 활황과 불황이 변화에 의해 수식이 될 수 없고 시장의와의 결속력이 강하기 때문에 21b와 같은 구조가 아닌 21c의 병렬 구조를 가지게 된다. 위의 예들을 통해 의미적 응집성이 강하게 일어날수록 병렬 구조를 명확히 선택할 수 있음을 보였는데, 의미적 응집성은 서술어와 논항간의 관계를 파악할 수 있는 성질도 될 수 있지만 특히 병렬문에서 자주 발생하는 수식에 의한 구조적 중의성을 해결하기 위한 성질로 유용하게 사용될 수 있다.

4. 대칭성과 응집성

앞서 살펴보았던 병렬항의 대칭성과 응집성은 두 가지 성질이 모두 병렬성에 기반하는 성질이기에 때문에 이 두 가지가 병렬문에 함께 존재하여 병렬 구조의 선정에 영향을 줄

수 있다는 점을 의미하고 있다. 이러한 점은 병렬 구조의 선정에 있어서 대칭성과 응집성이 함께 고려되어야 한다는 것을 내포하며 이러한 성질은 올바른 병렬 구조에 대해 다음과 같은 관계로 정의될 수 있다.

대칭과 응집의 조화 병렬문에서 대칭성과 응집성이 함께 나타나고 두 가지 성질이 강하게 나타날수록 병렬 구조는 명확해 진다.

이 성질은 구조적 대칭성과 병렬항의 응집성에 작용하는 형태소적, 구문적, 의미적인 요소들이 병렬문에서 영향을 미치고 있음을 의미하므로 병렬문의 구조를 파악하기 위해서는 이러한 복합적 요소들을 일관된 방법으로 함께 고려하여 처리할 수 있는 방법이 필요하다. 이러한 점은 병렬문의 구조를 파악하고자 했던 기존의 연구들에서도 전제되고 있는 것인데 다만 각 연구에서 모델링하는 요소들의 종류나 적용 방법에 있어 차이를 보이고 있다. 이러한 연구들은 대칭성에 기반하여 병렬 구조의 범위를 파악하고자 하는 연구와 (Agarwal and Boggess, 1992; Kurohashi and Nagao, 1994; Okumura and Muraki, 1994; Yang, 1996; Park and Cho, 2000), 응집성에 기반하여 병렬 구조의 범위를 파악하고자 하는 연구 (Resnik, 1999)가 있다. 이외에 통사 구조의 패턴으로 병렬 구조를 파악하고자 하는 패턴 기반 방법이 시도되었다 (Park, 1998). 다음의 예문을 통해 대칭성과 응집성이 병렬문에 함께 나타남을 보이고 병렬 구조와의 관계를 살펴보도록 한다.

22. (a) 엘지 21인치 평면 텔레비전과 디오스 냉장고
 (b) 엘지 21인치 평면 텔레비전과 디오스 냉장고
 (c) 엘지 21인치 평면 텔레비전과 디오스 냉장고
 (d) 엘지 21인치 평면 텔레비전과 디오스 냉장고
 (e) 엘지 21인치 평면 텔레비전과 디오스

위의 예 22는 앞서도 설명한 바와 같이 수식어에 의한 중의성에 존재하며 22a의 병렬 구조가 선호된다. 그러나 22b와 비교하여 22a가 선호되는 이유는 텔레비전과 냉장고의 통사적, 의미적 대칭성이 강하기 때문으로 보여진다. 즉 이 명사구 병렬에서 디오스는 고유 명사이지만 22b와 같은 구조를 가지기 위해서는 의미적인 응집성이 우측 병렬항에서 디오스만으로 강하게 나타날 수 있어야 하는데 냉장고가 양쪽 병렬항에서 공유되지 못하고 우측 병렬항에서만 결속될 수 있기 때문에 우측 병렬항의 의미적인 응집성이 디오스 냉장고에서 더 강하게 나타난다. 그러므로 22a가 22b에 비해 선호되는 구조가 될 수 있다. 22c와 22d에서의 병렬항 범위는 의미적인 응집성에서 22b보다도 강하게 나타나지 못하므로 선호되는 구조가 될 가능성이 떨어진다. 22e와 같이 냉장고가 나타나지 않을 경우 비교할 대상이 디오스로 한정되므로 22e의 병렬 구조가 선호될 것으로 예측할 수 있다. 이와 같이 22와 같은 명사구의 병렬 구조를 파악하기 위해서는 앞서 주장된 구조적 대칭성이나 의미적 응집성이 함께 고려되어 분석

되어야 한다. 이러한 분석을 통해 22의 예에서 22a와 같이 선호되는 병렬 구조는 구조적 유사성이나 병렬항 응집성이 다른 후보 구조에 비해 강하게 나타나게 된다.

23. (a) [삼성과 엘지]의 [텔레비전과 냉장고]

(b) [부동산 시장의 변화]와 [주식 시장의 [활황과 불황]]

이 예는 앞서 보인 연속적인 병렬 접속사에 의해 중의성이 발생하는 예이다. 이와 같은 유형의 중의성은 보통 수식어에 의한 중의성을 동반하는 경우가 많은데 이러한 경우 대칭적 측면과 응집적인 측면이 병렬 구조에 함께 존재하므로 어느 한가지 성질만으로 올바른 병렬 구조를 판단하기가 힘들다. 23a에서는 대칭성에 의해 삼성과 엘지, 텔레비전과 냉장고의 대칭 구조를 파악할 수 있고 응집성에 의해 삼성과 엘지와 텔레비전과 냉장고의 수식 구조를 파악할 수 있다. 이러한 분석을 통해 전체 병렬 구조를 파악할 수 있고 23b에서도 이러한 방식은 마찬가지로 적용될 수 있다.

24. (a) 밥을 먹고 출발하여라.

(b) 영희가 철수는 사랑하고 상호는 좋아한다.

이 예는 논항 관계에 의한 중의성이 존재하는 예로 24a과 24b는 동사구 병렬과 문장 병렬 구조로 후보들이 생성된다. 이 예에서도 마찬가지로 대칭성과 응집성이 동시에 존재하는데 대칭성은 양쪽 병렬항의 격이 같은 논항 사이와 서술어 사이의 구조를 파악하는데 도움이 될 수 있고 응집성은 논항이 양쪽 병렬항의 서술어에 의해 공유되는지를 판단하는데 도움이 될 수 있다. 24a의 병렬항은 밥을이 먹다에 대해서만 논항으로 결속될 수 있고 24b에서는 영희가 사랑하다와 좋아하다에 모두 결속될 수 있기 때문에 24에서 보인 것과 같은 병렬 구조가 선호된다.

25. (a) 사랑하고 좋아하는 것은 다르다.

(b) 사랑하고 좋아하는 것은 다르다

(c) 사랑하고 좋아하는 것은 다르다

25는 병렬 접속사의 범주에 의해 중의성이 발생하는 예이다. 이 예에서는 25a의 구조가 선호되는데 그 이유는 25b의 구조에 비해 25a가 더욱 대칭적이고 것이 사랑에도 수식될 수 있기 때문에 응집성도 떨어지게 되기 때문이다. 25c도 25a에 비해 구조적으로도 대칭성이 떨어지지만 의미적으로도 좌측항에 다르다가 결속되어 대칭성을 떨어뜨린다. 이와 같이 대칭성과 응집성을 이용한 병렬 구조의 분석을 통하여 구조적 중의성이 나타나는 병렬문의 후보들을 분석함으로써 선호되는 구조를 파악하였다. 앞서 보인바와 같이 이러한 선호되는 병렬 구조는 두 가지 성질에 일관성 있게 적용되므로 병렬문의 전산처리를 위한 요소들을 이 두가지 성질에 적용시켜 병렬문의 후보들 중 선호되는 구조를 선택할 수 있을 것으로 기대한다.

5. 결론

한국어 병렬문에서는 구조적 중의성으로 인해 다수의 통사 구조 후보들이 생성되는데 이 후보들 중에서 올바른 병렬 구조를 찾는 것은 병렬문의 정확한 의미를 이해하고 전산처리에 있어서 효율성을 증가시키는 데에 도움을 줄 수 있다. 이러한 방법을 위해 본 논문에서는 병렬 구조 선정의 단서가 될 수 있는 요소들을 일관성 있게 적용하기 위한 병렬의 중요한 두 가지 성질로 대칭성과 응집성을 제안하였다. 이 두 가지 성질은 병렬에 기반한 것으로 올바른 병렬 구조와의 관계를 통해 병렬문에 있는 다양한 정보들로부터 그 병렬 구조를 파악할 수 있도록 하는데 대칭성은 병렬항간의 구조적 유사성으로, 응집성은 병렬항의 비병렬 환경에 대한 결속력으로 병렬 구조를 파악할 수 있음을 보였다.

이를 위해 본 연구에서는 병렬문의 구조를 파악하는데 영향을 미치는 요소들에 대해 형태소, 통사, 의미, 담화 단계에 걸쳐 살펴보았고 병렬문에서 나타나는 구조적 중의성의 유형을 파악하여 여러 단계의 요소들이 대칭성과 응집성의 성질에 일관성 있게 적용됨을 구조적 중의성이 있는 병렬문의 분석을 통해 알아 보았다. 또한 이러한 대칭성과 응집성이 함께 병렬 구조의 분석에 적용될 수 있음을 보였다. 병렬문 중의성의 전산처리를 위해 제안된 이러한 성질들의 실제적인 적용을 위해서는 본 논문에서 파악한 핵심적인 요소들을 형태소 비교나 통사 자질, 공기 정보, 갈래말 사전 등의 정보들을 통해 측정하고 이들을 구조적 유사성과 병렬항 응집성에 적용시키는 작업이 필요하다. 본 논문은 이러한 실제 적용을 위한 체계를 제안하고 그 가능성을 보였다는 점에서 그 의의를 찾을 수 있다.

<참고문헌>

- 김영희. 1988. 등위접속문의 통사특성. 한글 201: 83-117.
- 김진수. 1987. 국어 접속조사와 어미 연구(I). 탑출판사.
- 서정수. 1985. 국어의 접속어미 연구(I). 한글 189: 51-86.
- 서정수. 1996. 국어문법. 한양대학교 출판원.
- 이관규. 1992. 대등구성의 요건과 유형. 한글 217: 65-81.
- 최재희. 1985. 국어 명사구 접속의 연구. 한글 188: 91-115.
- Agarwal, R. and L. Boggess. 1992. A Simple But Useful Approach to Conjunction Identification. In *30th ACL*, pages 15-21.
- Barker, K. and S. Szpakowicz. 1998. Semi-Automatic Recognition of Noun Modifier Relationships. In *COLING-ACL*, pages 96-102.
- Bayraktar, M., B. Say, and V. Akman. 1998. An Analysis of English Punctuation: The Special Case of Comma. *International Journal of Corpus Linguistics*, 3.1: 33-57.
- Cho, H. 2000. Coordinate Constructions in Korean and Parsing Issues in Combinatory Categorical Grammar. Master's thesis, Korea Advanced Institute of Science and Technology.

- Church, K. and R. Patil. 1982. Coping with Syntactic Ambiguity or How to Put the Block in the Box on the Table. *American Journal of Computational Linguistics*, 8.3-4: 139-149.
- Cremers, C. 1993. *On Parsing Coordination Categorially*. Ph.D. thesis, Leiden University.
- Frazier, L. 1987. Syntactic Processing: Evidence from Dutch. *Natural Language and Linguistic Theory*, 5: 519-559.
- Hindle, D. and M. Rooth. 1993. Structural Ambiguity and Lexical Relations. *Computational Linguistics*, 19.1: 229-236.
- Hoeks, J. C., W. Vonk, and H. Schriefers. 2002. Processing Coordinated Structures in Context: The Effect of Topic-Structure on Ambiguity Resolution. *Journal of Memory and Language*, 46: 99-119.
- Kobayasi, Y., T. Tokunaga, and H. Tanaka. 1994. Analysis of Japanese compound nouns using collocational information. In *COLING*, pages 865-869.
- Kurohashi, S. and M. Nagao. 1994. A Syntactic Analysis Method of Long Japanese Sentences Based on Detection of Conjunctive Structures. *Computational Linguistics*, 20.4: 507-534.
- Lauer, M. 1995. Corpus Statistics Meet the Noun Compound: Some Empirical Results. In *the 33rd ACL*, pages 47-54.
- Lee, H. and J. C. Park. 2002. Interpretation of Natural Language Queries for Relational Database Access with Combinatory Categorical Grammar. *International Journal of Computer Processing of Oriental Languages*, 15.3: 281-303.
- Lee, W. J. and J. C. Park. 2001. Computational Generation of Context-based Intonation for Korean with Combinatory Categorical Grammar. In *International Conference on Computer Processing of Oriental Languages*, pages 415-420.
- Milward, D. 1994. Non-Constituent Coordination: Theory and Practice. In *COLING*, pages 935-941.
- Munn, A. B. 1993. *Topics in the Syntax and Semantics of Coordinate Structures*. Ph.D. thesis, University of Maryland.
- Okumura, A. and K. Muraki. 1994. Symmetric Pattern matching analysis for English Coordinate Structure. In *4th Conference on Applied NLP*, pages 41-46.
- Park, J. 1998. Analysis of Coordinate Noun Phrase Using Part-of-Speech Pattern. Master's thesis, Korea Advanced Institute of Science and Technology.
- Park, J. C. and H. J. Cho. 2000. Informed Parsing for Coordination with Combinatory Categorical Grammar. In *COLING*, pages 593-599.
- Park, J. K. and J. C. Park. 1999. An Analysis of the Semantic and Discourse Functions of the Korean special marker '-to'. In *Hangul and Korean Information Processing*, pages 324-331.
- Postal, P. 1971. *Cross Over Phenomena*. MIT Press.
- Resnik, P. 1999. Semantic Similarity in a Taxonomy: An Information-Based Measure and its Application to Problems of Ambiguity in Natural Language. *Journal of Artificial Intelligence Research*, 11: 95-130.

- Resnik, P. and M. Hearst. 1993. Structural Ambiguity and Conceptual Relations. *the Workshop on Very Large Corpora: Academic and Industrial Perspectives*, pages 58-64.
- Schachter, P. 1977. Constraints on Coordination. *Language*, 53: 86-103.
- Steedman, M. 2000. *The Syntactic Process*. MIT Press.
- Yang, J. 1996. Structural Ambiguity Resolution in Korean Noun Phrase Conjunctions Using Co-occurrence Similarity. *Journal of KISS*, 23.3: 311-321.
- Yoon, J., E. Jong, and M. Song. 1998. Analysis of Korean Compound Noun using Lexical Information between Nouns. *Journal of KISS*, 25.11: 1716-1725.

접수일자: 2003년 11월 7일

게재결정: 2003년 11월 30일