

한국어 피동형의 전산적 구현*

홍정하

(고려대학교 언어과학과)

kleist@korea.ac.kr

요약

이 논문에서는 피동접사 ‘이, 히, 리, 기’와 결합하는 피동형과 관련된 형태·통사적 문제를 전산적 관점에서 다룬다. 전산처리에서 이러한 피동형의 형태적 문제는 다음과 같다. 첫째, 피동접사 ‘이, 히, 리, 기’와 결합할 수 있는 타동사 어간의 분포가 제한되어 있다 둘째, 타동사 어간이 결합할 수 있는 피동접사는 고정되어 있다 셋째, 피동형 중에 타동사 어간과 피동접사가 결합할 때 형태적으로 변화하는 것들이 있다 ‘나누다/나뉘다, 모으다/모이다, 잡그다/잡기다, 자르다/잘리다’ 등이 여기에 해당된다.

이러한 형태적 문제 외에도 전산처리에서 피동형과 관련된 통사적 문제는 다음과 같다. 첫째, 능동형의 타동사가 피동형이 되면서 논항구조도 함께 변화한다. 둘째, 피동문의 행동주가 문장에서 생략되는 경우가 종종 있다. ‘문제가 쉽게 풀리었다’, ‘소리가 잘 들린다’ 등이 이에 해당된다.

이 논문은 한국어 피동접사 ‘이, 히, 리, 기’와 결합하는 피동형의 형태·통사적 특징을 전산적으로 처리하는 것이 목적이다. 이를 위해 표상모형으로는 차질구조를, 구현도구로는 Malaga를 사용한다.

1. 서론

영어와 같은 인구어의 피동(passive)은 주로 통사적 차원의 문제로 다루어지지만, 한국어의 피동문은 일차적으로 문장의 서술어에 따라 결정되기 때문에 형태적 차원의 문제라 할 수 있다. 지금까지의 연구에서 피동문을 구성하는 한국어 서술어의 피동형식을 형태적 차원에서 네 가지 유형으로 정리할 수 있는데, 그 유형과 예는 다음과 같다¹⁾.

(1) 피동형식의 유형

- 가. 타동사 어간 + ‘이, 히, 리, 기’
 도둑이 포교한테 잡히었다.
- 나. 자동사
 소가 농부에게 맞았다.
- 다. 용언어간 + ‘아/어’ + ‘지다’
 수양체가 읊지문덕에게 깨뜨려지다.
- 라. 명사성어근 + ‘되다’
 그 사람의 글이 신문에 게재되었다.

* 이 논문은 1999년도 한국학술진흥재단 선도연구자 지원사업(연구책임자: 이기용)의 연구비를 지원 받아 이루어졌다. 이 논문을 읽고 조언을 해 주신 이기용 선생님과, 피동형 목록을 정리해준 고려대 언어학과 대학원 최승철군에게 감사 드린다.

1) 박양규 (1990), 배희임 (1988) 참조.

이 네 가지 피동형식의 유형 중에서 타동사 어간에 피동접사 ‘이, 히, 리, 기’가 결합한 형태가 한국어에서 피동문을 구성하는 가장 대표적인 서술어이다. 실제로 대부분의 국어사 전에서도 이 피동형을 피동사로서 표제어로 수록하고 있다²⁾.

그러나 피동접사 ‘이, 히, 리, 기’가 개입된 동일한 유형의 피동형태일지라도 피동성의 정도 차이에 따라 행위자를 상정하기 어려워 심지어 자동사로 분류되어야 하는 것들도 상당수 있다. 다시 말해서 동일형태의 피동형이더라도 통사적으로 행위자와 행위를 받는 대상을 모두 나타낼 수 있는 것이 있는가하면, 행위자를 나타낼 수 없는 것들이 있다. 혼히 전자의 유형을 진피동이라 한다.

이 논문에서는 타동사 어간에 피동접사 ‘이, 히, 리, 기’가 결합한 진피동형의 형태·통사적 특징을 전산적으로 처리하는 것이 목적이다. 이를 위해 표상모형으로는 자질구조를³⁾, 구현도구로는 Malaga⁴⁾를 사용한다.

2. 피동형 처리의 문제

전산적으로 피동형을 처리할 때의 문제는 형태적 측면과 통사적 측면으로 구분하여 생각해 볼 수 있다. 먼저 형태적 측면의 문제는 적형의 피동형 처리와 직결되는 것으로 다음과 같다. 첫째, 피동접사 ‘이, 히, 리, 기’와 결합할 수 있는 타동사 어간의 어휘 수는 극히 제한되어 있고⁵⁾, 실제 피동형을 형성하는 것과 그렇지 못한 것간의 규칙성을 포착하기가 어렵다. 다음의 (2-가, 나, 다, 라)와 (2-가’, 나’, 다’, 라’)는 각각 피동형이 되는 것과 불가능한 것인데 그 기준 설정이 쉽지 않다.

(2) 피동형의 불규칙성

가. 뽑다-뽑히다	가’. 줍다-*줍(이/히/리/기)다
나. 팔다-팔리다	나’. 사다-*사(이/히/리/기)다
다. 몰다-몰리다	다’. 넣다-*넣(이/히/리/기)다
라. 떠먹다-떼먹히다	라’. 흘겨보다-*흘겨보이다

둘째, 각각의 타동사 어간이 결합할 수 있는 피동접사는 고정되어 있다. 즉, 각각의 타동사 어간이 결합할 수 있는 피동접사는 ‘이, 히, 리, 기’ 중 하나뿐이고 항상 고정되어 있다.

2) 동사어간에 ‘이, 히, 리, 기’가 결합한 피동형을 피동사로 기술하는 국어사전은 다음과 같다(우인혜 (1995) 참조).

이희승 (1982), 「국어대사전」, 서울: 민중서림.

신기철, 신용철 (1988), 「새우리말큰사전」, 서울: 삼성출판사.

김민수 외 (1991), 「국어대사전」, 서울: 금성출판사.

사회과학원언어연구소 (1992), 「조선말대사전」, 평양: 동광출판사.

국립국어연구원 (1999), 「표준국어대사전」, 서울: 두산동아.

3) 이 논문의 자질구조는 Sag & Wasow (1999)를 기반으로 하여 수정하였다.

4) Malaga는 독일 Erlangen-Nurnberg 대학교 컴퓨터언어학과에서 좌연접문법(Left-Associative Grammar)을 표상할 수 있도록 C-언어를 보강하여 개발한 프로그래밍 언어이다(Beutel 1997 참조). 좌연접문법에 대해서는 Hausser (1989), 이기용 (1999), 홍정하 (1998) 참조.

5) 우인혜 (1993)에 따르면 타동사에 ‘이, 히, 리, 기’가 결합하는 피동사의 개수는 약 150여 개이지만, 실제 국어사전에 공통으로 피동사로 기술되는 어휘수가 95개에 불과하다(우인혜, 1995).

셋째, 피동형 중에 타동사 어간과 피동접사가 결합할 때 형태적으로 변화하는 것들이 있다. ‘나누다/나뉘다, 모으다/모이다, 잡그다/잡기다, 자르다/잘리다’ 등이 여기에 해당된다.

한편, 통사적 측면에서 고려되어야 할 문제는 다음과 같다. 첫째, 능동형의 타동사가 피동형이 되면서 논항구조도 함께 변화하는데, 어떻게 체계적으로 표상하고 처리하는가의 문제이다.

둘째, 피동문에서 종종 행위주가 생략된다. 다음의 (3-가)와 같이 행위주와 행위의 대상이 모두 나타나기도 하지만, (3-나)와 같이 행위주의 생략이 가능하다.

(3) 피동문 구성의 차이

- 가. 도둑이 경찰에게 쫓긴다.
- 가'. 경찰이 도둑을 쫓는다.
- 나. 음악소리가 (누구에게) 들린다.
- 나'. (누가) 음악소리를 듣는다.

3. 피동형의 형태적 처리

그간의 피동 연구에서 피동형이 될 수 있는 기준은 아니지만, 피동형이 될 수 없는 기준, 즉 피동형의 제약에 대한 연구가 있어왔다⁶⁾. 그러나 그 제약에는 예외가 있어서 절대적 기준이 될 수 없으며, 추상적인 것들이어서 전산적으로 처리하기에는 부적합하다고 하겠다.

이 논문에서는 ‘이, 히, 리, 기’ 피동형이 될 수 있는 타동사 어간의 어휘부를 속성 DERV를 이용하여 정의한다. 속성 DERV는 i_pass, hi_pass, li_pass, ki_pass 중 하나의 값을 취하는데, 각각 ‘이, 히, 리, 기’ 파생접사와 결합하여 피동형으로 파생될 수 있음을 나타낸다. (4)는 파생형이 될 수 있는 타동사 어간의 어휘정보이다.

(4) 피동형 어간의 어휘정보

- 가. 파생접사 ‘-이-’와 결합하는 어간의 예

[PHON: “보”,
MOR: [STEM: [BASEF: “보”, DERV: i_pass]],
SYN: [HEAD: [POS: verb]],
ARG_ST: <[TAG: 1, CASE: nom], [TAG: 2, CASE: acc]>,
SEM: [ROLE: [RELN: “see”, AGENT: 1, THEME: 2]]];

- 나. 파생접사 ‘-히-’와 결합하는 어간의 예

[PHON: “먹”,
MOR: [STEM: [BASEF: “먹”, DERV: hi_pass]],
SYN: [HEAD: [POS: verb]],
ARG_ST: <[TAG: 1, CASE: nom], [TAG: 2, CASE: acc]>,
SEM: [ROLE: [RELN: “eat”, AGENT: 1, THEME: 2]]];

6) 배희임 (1988)에서는 하다형 동사, 수여동사, 수혜동사, 지각동사, 대칭동사, 공동격, 자동과 타동을 겸하는 동사, 파생에 의한 타동사, 복합동사는 피동형이 될 수 없거나 되기 힘든 경향을 갖고 있다고 한다. 또한 강명윤 (1997)에서는 형태론적 피동의 불규칙성은 통사·의미적 원리, 즉 피영향성 조건(affectedness condition)에 의해 설명될 수 있다고 제시하였다.

다. 파생접사 ‘-리-’와 결합하는 어간의 예

[PHON: “풀”,
MOR: [STEM: [BASEF: “풀”, DERV: li_pass]],
SYN: [HEAD: [POS: verb]],
ARG_ST: <[TAG: 1, CASE: nom], [TAG: 2, CASE: acc]>,
SEM: [ROLE: [RELN: “solve”, AGENT: 1, THEME: 2]]];

라. 파생접사 ‘-기-’와 결합하는 어간의 예

[PHON: “찢”,
MOR: [STEM: [BASEF: “찢”, DERV: ki_pass]],
SYN: [HEAD: [POS: verb]],
ARG_ST: <[TAG: 1, CASE: nom], [TAG: 2, CASE: acc]>,
SEM: [ROLE: [RELN: “tear”, AGENT: 1, THEME: 2]]];

따라서 어휘정보에 피동형으로 파생될 수 있는 속성 DERV과 이에 해당하는 값을 갖는 타동사어간만이 피동접사와 결합할 수 있고, 그렇지 않은 동사는 피동접사와 결합할 수 없게 된다⁷⁾. 또한 이 속성과 같은 특정한 피동접사와의 결합만을 허용한다. (5)는 타동사 어간 ‘먹, 보, 찢, 물’과 피동접사 ‘이, 하, 리, 기’, 그리고 종결어미 ‘다’간의 가능한 결합에서 이러한 제약이 적용되어 정형의 어형만이 생성되고 있음을 보이고 있다.

(5) 정형(well-formed) 피동형의 생성

malaga> mg 4 먹 보 잘 찢 물 이 히 리 기 다

1: “먹히다”

2: “먹다”

3: “보이다”

4: “보다”

5: “찢기다”

6: “찢다”

7: “물리다”

8: “물다”

여기서 mg(morphological generation)는 구축된 결합규칙에 따라 mg 뒤의 형태소로 결합 가능한 어형을 모두 생성하는 명령어이다. 또한 숫자 4는 최대의 형태소 결합 개수를 제한하는 것이다.

한편, ‘나누다/나뉘다, 모으다/모이다, 잡그다/잡기다, 자르다/잘리다’와 같이 어간과 피동접사가 결합할 때 어형변화를 수반하는 예들이 있다. 이러한 문제의 처리는 속성 VARIAT를 이용한다. 이 속성의 값으로는 i_contract('이' 모음의 축약), u_del('으' 모음의 탈락), ku_del('그' 탈락)을 취할 수 있다. (6)은 이형태 생성 및 이형태의 결합을 요구하는 속성 VARIAT와 값을 포함하는 어간의 어휘정보이다.

7) 이런 방법의 사용은 피동형과 같이 불규칙성과 생산성이 낮은 어휘처리시에 적절하다. 또한 이기용 (1999)에서도 ‘하다’형 용언을 이러한 방식으로 처리한 바 있다.

(6) 어형변화를 수반하는 어간의 어휘정보

가. '이' 모음의 축약

[PHON: “나누”,
MOR: [STEM: [BASEF: “나누”, DERV: i_pass, VARIAT: i_contract]],
SYN: [HEAD: [POS: verb]],
ARG_ST: <[TAG: 1, CASE: nom], [TAG: 2, CASE: acc]>,
SEM: [ROLE: [RELN: “divide”, AGENT: 1, THEME: 2]]];

나. '으' 모음의 탈락

[PHON: “모으”,
MOR: [STEM: [BASEF: “모으”, DERV: i_pass, VARIAT: u_del]],
SYN: [HEAD: [POS: verb]],
ARG_ST: <[TAG: 1, CASE: nom], [TAG: 2, CASE: acc]>,
SEM: [ROLE: [RELN: “gather”, AGENT: 1, THEME: 2]]];

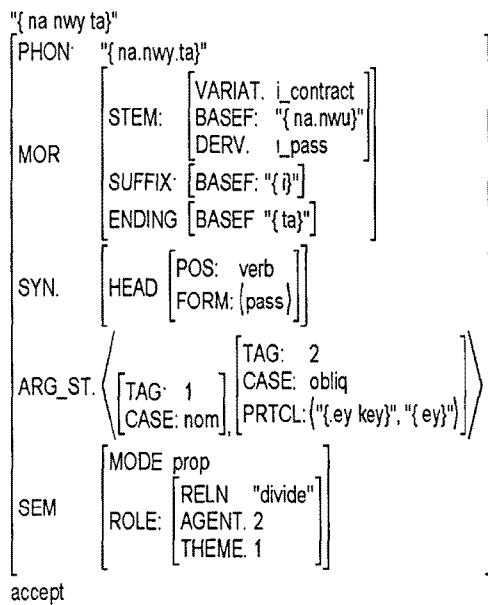
[PHON: “자르”,
MOR: [STEM: [BASEF: “자르”, DERV: li_pass, VARIAT: u_del]],
SYN: [HEAD: [POS: verb]],
ARG_ST: <[TAG: 1, CASE: nom],
[TAG: 2, CASE: acc]>,
SEM: [ROLE: [RELN: “cut”, AGENT: 1, THEME: 2]]];

다. '그' 탈락

[PHON: “잠그”,
MOR: [STEM: [BASEF: “잠그”, DERV: ki_pass, VARIAT: ku_del]],
SYN: [HEAD: [POS: verb]],
ARG_ST: <[TAG: 1, CASE: nom],
[TAG: 2, CASE: acc]>,
SEM: [ROLE: [RELN: “lock”, AGENT: 1, THEME: 2]]];

속성 VARIAT를 포함하는 어휘의 처리를 위해 이형태생성규칙에 따라 그 값에 적절한 이형태들을 생성한다. 속성값이 i_contract인 어간의 처리를 위해서 피동접사 '이'의 이형태로 'ㅣ'를 생성하여 처리한다. 즉 '나뉘다'는 「나누 + 이」의 결합이 아니라 '이'의 이형태인 '나누 + ㅣ'의 결합이 되는 것이다. 이에 비해 u_del과 ku_del과 같이 모음과 한 음절이 탈락하는 경우는 피동접사가 아닌 타동사 어간의 이형태를 생성해낸다. 따라서 이형태생성규칙에 따라 '모으, 자르, 잠그'의 이형태 '모, 잘, 잠'이 각각 생성된다. (7)은 이러한 예들 중 하나인 '나뉘다'의 형태분석결과이다. (7)의 형태정보 MOR에서 접사 SUFFIX는 기저형 BASEF이 '이'인 것을 알 수 있다.

(7) ‘나뉘다’의 형태분석결과



그러나 원래 어간과 그 이형태들은 각각 결합할 수 있는 형태소의 차이가 있다. 다시 말해서 속성 VARIAT를 포함하는 어간은 피동접사와 결합할 수 없지만 종결어미 ‘-다’와 직접 결합할 수 있는 반면, 그 이형태들은 종결어미와 직접 결합할 수는 없지만 피동접사와 결합할 수 있다. 따라서 이러한 제약을 결합규칙에 부여함으로써 다음과 같은 비정형의 어형 분석 및 생성을 방지해야한다.

- (8) 가. *나누이, *모으이, *자르리, *잠그기
 나. *모다, *잘다, *잡다

다음의 (9)는 이러한 제약이 적용되어 어간, 이형태, 피동접사, 종결어미간의 가능한 결합으로 정형의 어형만이 생성됨을 보이고 있다.

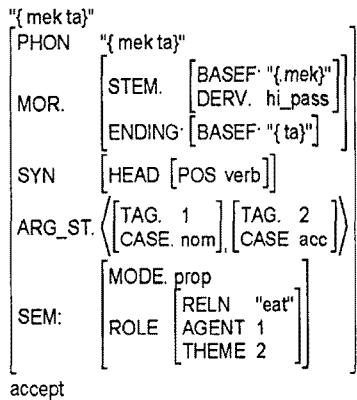
(9) 이형태를 이용한 정형의 피동형 생성

- malaga> mg 4 나누 모으 모 잠그 자르 잘 이 । 히 리 기 다
 1: “나뉘다”
 2: “나누다”
 3: “모으다”
 4: “모이다”
 5: “잠그다”
 6: “잠기다”
 7: “자르다”
 8: “잘리다”

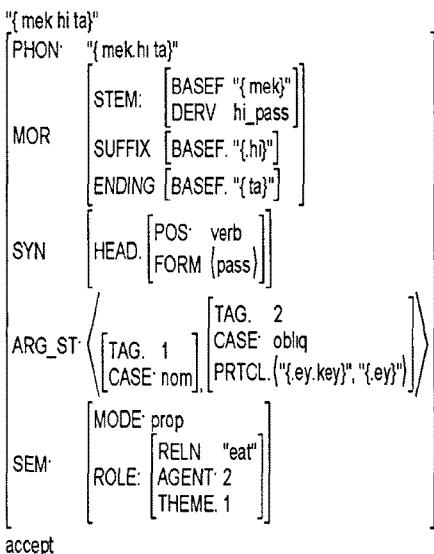
4. 피동형의 통사적 처리

피동문을 능동문과 비교해 보면, 행위자와 행위의 대상은 동일하지만 해당하는 각 논항의 위치와 위치에 따른 격의 차이가 있다는 것이다. 따라서 타동사의 피동형은 논항구조의 차이를 반영하면서도 동일한 의미관계를 유지해야 한다.

(10) 타동사 '먹다'의 분석결과



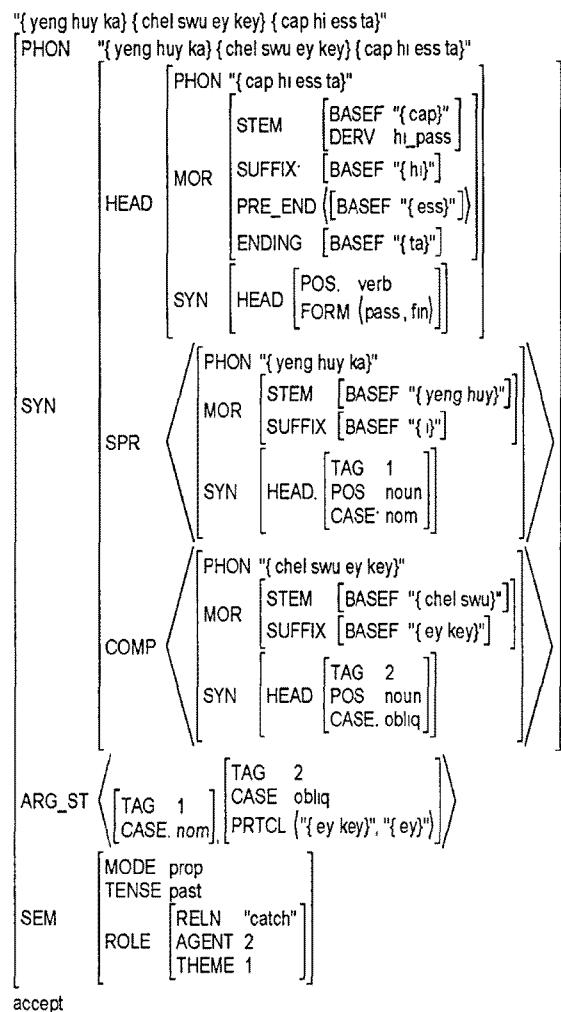
(11) 피동형 '먹히다'의 분석결과



타동사 '먹다'의 자질구조 (10)과 이것의 피동형 '먹히다'의 자질구조 (11)을 비교해 보면, 형태정보를 표상하는 MOR에 접사 '히'가 결합되어 있다는 것, 통사정보를 표상하는 SYN에서 동사가 피동형이라는 것, 그리고 논항구조 AGR_ST와 의미정보 SEM에서 AGENT, THEME의 값들에 차이가 있다는 것이다. 여기서 (10)과 (11)의 AGR_ST와 SEM을 자세히 살펴보면, AGR_ST의 TAG 값에 의해 행위역 AGENT과 대상역 THEME으로

지시되는 논항의 위치와 격이 달라졌다는 것뿐이다. 즉, (10)에서는 AGENT의 값 1과 동일한 TAG 값을 취하는 AGR_ST의 첫 번째 논항의 격은 주격 nom이고, THEME의 값 2와 동일한 TAG 값을 취하는 AGR_ST의 두 번째 논항의 격은 대격 acc이다. 그러나 (10)의 피동형인 (11)에서는 AGENT의 값 2와 동일한 TAG 값을 취하는 AGR_ST의 두 번째 논항의 격은 사격 oblique이고 조사 PRTCL이 '에게, 에' 중 하나이어야 함을 나타낸다⁸⁾. 반면에 (11)에서 THEME의 값 1과 동일한 TAG 값을 취하는 AGR_ST의 첫 번째 논항의 격은 주격이다. 따라서 각 의미역을 표상하는 논항의 위치의 차이와 그 논항의 격이 틀리지만 두 의미역은 그대로 유지된다.

(12) '영희가 철수에게 잡히었다'의 분석결과



8) 피동문에서 행위역 논항의 조사는 이 밖에도 여러가지가 가능하나, 이 논문에서는 '에게, 에'로 한정한다.

이러한 피동형의 자질구조는 (12)와 같은 피동문의 분석을 허용한다. (12)는 동사 ‘잡히었다’가 HEAD이고, 명사구 ‘영희가’는 SPR, 명사구 ‘철수에게’는 COMP로 분석되었음을 나타내고 있다. 또한 SPR로 분석된 명사구 ‘영희가’는 SEM의 THEME 값 1과 동일한 TAG 값을 취하고, COMP로 분석된 명사구 ‘철수에게’는 AGENT 값 2와 동일한 TAG 값을 취하므로 각각 대상역과 행위역의 논항임을 나타내고 있다.

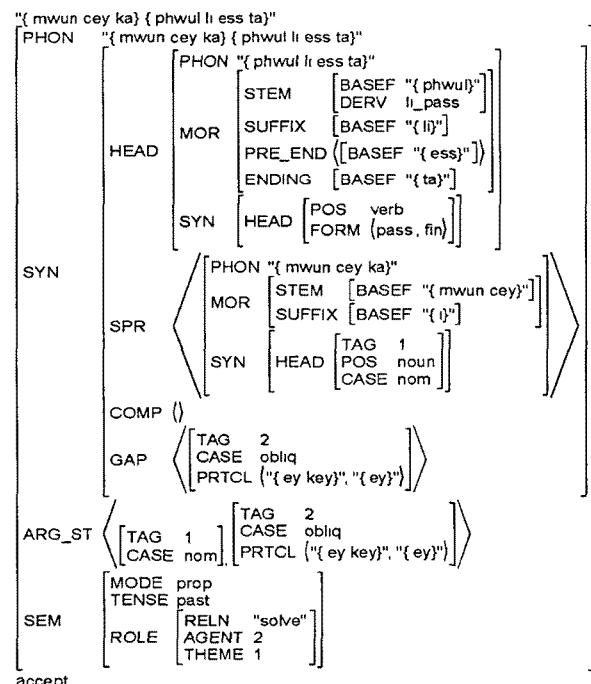
한편, 모든 피동문이 (12)와 같이 행위역과 대상역의 논항들을 모두 표면적으로 나타내는 것은 아니다. (13)과 같이 피동문에서 행동주가 생략되는 경우가 있다.

(13) 행동주의 생략

- 가. 문제가 (누구에 의해) 풀리었다.
- 가’. (누군가가) 문제를 풀었다.
- 나. 음악소리가 (누구에게) 들린다.
- 나’. (누군가가) 음악소리를 들었다.

이러한 생략은 한국어에서 흔히 나타나는 현상으로 학용적 추론에 의해 화자, 청자 또는 담화상의 제3자로 해석될 수 있는 논항이다. 따라서 이러한 생략현상을 처리하기 위해서 속성 GAP을 이용한다⁹⁾.

(14) ‘문제가 풀리었다’의 분석결과



9) LFG, GPSG, HPSG와 같은 비변형이론에서는 음성적으로 실현되지 않지만, 논항위치에 나타나는 빈요소들을 공백(gap)이라하여 기술하고 있다.

(14)는 (13-가)를 분석한 결과이다. (14)에서 COMP의 값은 빈리스트로, 문장에서 COMP로 분석된 논항이 없음을 나타낸다. 그러나 문장에서 표면적으로 나타나있지는 않지만 속성 GAP에 의해 화용적 추론에 의해 복원될 수 있는 논항이 있음을 나타내고 있다. GAP은 TAG 값이 2이면서 사격 oblique인 논항을 값으로 취하는데, 이 논항은 AGENT와 동일한 값을 가지므로 피동문에서 행위자 역할을 하게된다. 또한 이 논항의 자질구조는 AGR_ST 값 중 두 번째 논항과 일치하므로 논항구조를 위반하지 않는다.

5. 결론 및 향후 연구과제

지금까지 속성 DERIV를 이용하여 '이, 히, 리, 기' 진피동형의 분포적 제약, 고정된 피동접사와의 결합 제약 문제를 해결하였고, 속성 VARIAT을 취하는 피동형의 이형태를 생성해서 형태적 변화형들을 처리하였다. 또한 의미역은 그대로 유지시키면서 논항구조 변화를 성공적으로 표상할 수 있도록 TAG 값을 이용하여 능동문과 피동문의 관련성을 포착하였다. 그리고 속성 GAP을 이용하여 행동주가 생략된 피동문을 처리하였다. 결론적으로 말하자면, 피동형과 피동문에 관한 핵심적인 몇 가지 문제해결을 통해 이 논문에서 제시한 자질구조와 구현한 분석기의 타당성을 제시하였다.

그러나 여기서 다룬 문제는 피동의 핵심적 특징이기는 하지만 피동의 문제를 총괄하지는 못하였다. 따라서 앞으로 다음과 같은 연구가 수행되어야 한다.

첫째, 피동접사 '이, 히, 리, 기'와 결합하지만 피동문에서 행위주를 설정할 수 없는 피동형 처리에 관한 연구가 필요하다.

둘째, 앞서 언급한 것처럼 '이, 히, 리, 기'의 진피동형이 차지하는 분포는 미약하기 때문에 분포적으로 한국어 피동의 대다수를 차지하고 있다고 할 수 있는 자동사, '지다'형, '되다'형의 피동형들의 연구가 필요하다.

셋째, 각 피동형 사이에 존재하는 피동성 정도의 차이는 서술어뿐만 아니라, 각 문장의 논항의 의미적 특성에서도 기인한다. 따라서 이러한 논항의 연구가 의미적 차원에서 연구가 수행되어야 한다. 이 연구는 또한 피동형 사이에 존재하는 피동성 정도차이의 문제를 해결 할 수 있을 것이다.

참고문헌

- 강명윤 (1997), “한국어의 형태론적 피/사동 현상의 소형구절구조 이론적 접근”, 「여학연구」 33-1: 79-100.
- 박양규 (1990), “피동법”, 서울대학교 대학원 국어연구회 편, 「국어연구 어디까지 왔나」, 동아출판사, 493-499.
- 배희임 (1988), 「국어피동연구」, 고려대학교 민족문화연구소.
- 우인혜 (1993), 「국어의 피동법과 피동 표현의 연구」, 한양대학교 박사학위논문.
- 우인혜 (1995), “국어 피동의 범위”, 「국어학」 26: 99-124.
- 이기용 (1999), 「전산형태론」, 고려대학교 출판부.
- 홍정하 (1998), 「좌연접문법을 이용한 영어명사구의 전산처리」, 고려대학교 석사학위 논문.

- Beutel, Björn & Roland Hausser (1997), "Malaga 4.0"(unpublished), Abteilung für Computerlinguistik, Universität Nürnberg-Erlangen, Germany.
- Hausser, Roland (1989), *Computation of Language: an Essay on Syntax, Semantics and Pragmatics in Natural Man-Machine Communication*, Berlin: Springer-Verlag.
- Sag, Ivan & Thomas Wasow (1999), *Syntactic Theory: A Formal Introduction*, CSLI.