

# 수화 애니메이션 자동 생성을 위한 한국어 복문의 수화 스크립트 변환 방법

김상하    장은영    박종철  
한국과학기술원 전자전산학과 전산학전공  
{shkim, euny, park}@nlp.kaist.ac.kr

## Translating a Complex Sentence in Korean into a Sign Language Script for an Automatic Sign Language Generation

Sangha Kim Eunyoung Chang Jong C. Park  
Computer Science division, EECS department, KAIST

### 요 약

한국 수화는 시각, 공간언어로 한국어와는 상이한 문법체계를 가진 언어로 수화를 일차 언어로 사용하는 농인들에게 있어 복잡한 구조의 한국어 문장은 부담이 된다. 본 논문은 이런 한국어 문장의 복잡한 구조를 농인들이 이해하기 쉬운 구조의 전개방식으로 변환하는 수화 스크립트 생성 시스템을 제안한다. 시스템은 세 단계로 구성되는데, 첫 번째 단계는 한국어 문장의 결합범주문법을 이용한 구문 분석이며, 두 번째 단계는 농인들이 이해하기 수월한 전개방식으로서의 절단위 재배열이고, 세 번째 단계는 공간이동을 고려한 스크립트 형태로의 변환이다. 본 논문은 한국 수화의 복문 실현 방법에 대해 살펴본 후, 이를 처리하는 시스템의 단계별 처리 방안에 대해 구체적으로 논의한다.

### 1. 서론

사회가 점차 복지 사회로 나아감에 따라 장애인에 대한 많은 관심과 개선의 노력을 보이고 있지만, 누구나 정보를 쉽게 이용하고 공유할 수 있는 정보화 사회에서 이런 노력은 아직 많이 부족한 실정이다. 특히 수화를 일차언어로 사용하는 많은 수의 농인들은 청각 손실에 따른 문제 이외에도 언어간 문법 체계의 차이로 정보의 원활한 획득에 많은 어려움을 겪고 있는 실정이다. 이에 따라 그 격차를 줄이고자 농학교 교육에서 한국어 문법을 가르치고 있으나, 청각장애 학생의 독해력 연구결과를 살펴보면 한국어 문법과 다른 시각, 공간언어인 수화를 사용하는 농인들에게 한국어 문장은 부담이 되고 있음을 알 수 있다[1]. 특히 대다수의 정보가 텍스트의 형태로 보관, 전달되고 있으며, 문헌상의 많은 문장들이 높은 비율로 복문 구조를 보이고 있다는 점을 감안한다면, 복잡한 구조의 문장을 이해하는데 어려움을 보이는 농인들을 위한 한국어 문장의 수화 변환 시스템에 대한 필요성은 더욱 높아진다. 최근에는 일반인들도 이해하기 어려운 법률 용어와 문장 구조로 이루어진 법령문을 시각, 청각장애인들이 이해하기 쉬운 문장으로 바꿔야 한다는 주장이 제기되고 있어 이런 노력은 더욱 의미를 갖는다.

수화를 사용하는 농인들이 한국어를 이해하도록 무조건 요구하기 보다는 농인들이 이해하기 쉬운 한국어 문장이나 한국 수화로 표현해준다면, 농인들의 정보 접근에 많은 도움을 줄 수 있을 것이다. 본 논문에서는 인터넷과 문헌상에 존재하는 한국어 문법을 따르는 복잡한 구조의 문장을 농인들이 이해하기 쉬운 단순한 구조의 수화문으로 바꾸어 농인들이 텍스트의 형태보다 쉽고 빠르게 정보를 습득할 수 있도록 문장 구조를 변환하는 시스템을 제안하고자 한다. 본 연구는 자연수화 생성을 목표로 이해하기 쉬운 구조의 수화 스크립트를 생성하는 방안에 대해 논의한다.

본 논문은 다음과 같은 순서로 진행된다. 2절에서는 기존의 한국어 문장을 수화로 변환하는 연구에 대해 살펴보고, 3절에서는 수화의 특성에 대해 살펴보고 수화의 복문표현의 어려움과 해결방안에 대해 살펴본다. 4절에서 앞 절에서 논의된 내용을 바탕으로 이를 처리하는 시스템에 대해 논의하고, 5절에서는 결론 및 향후 계획에 대해 논의한다.

### 2. 관련 연구

본 절에서는 한국어 문장을 한국 수화로 변환해주는

기존 시스템들에 대해 알아본다.

[2]에서는 수화 단어의 종류를 세분화하여 8가지로 정의하고 이를 품사별로 5종류의 사전으로 구성하여 이 사전을 통해 한국어 문장을 문법식, 혼합식 수화로 변환하는 시스템을 제안한다. 이 시스템은 각 어절에 대하여 품사별로 분리한 후, 수화 사전을 참조하여 수화어휘로 변환해 주는 방식을 제안한다. 이 시스템은 다수의 수화 어휘를 가지는 수화 사전을 구축하고 동음어휘를 처리하는 방안도 제시한다. 그러나 문법식, 혼합식 수화로 의 변환은 국문법의 체계를 그대로 따르거나 그 틀에서 크게 벗어나지 못한다는 한계가 있다.

[3, 4]에서는 한국어 문장의 구문 분석을 통해 자연수화를 생성하는 시스템을 제안해 기존 연구에서 주로 이루어지던 단어 별 수화 어휘 매핑 방식에서 벗어나 수화의 특유의 동시성과 공간성을 살려 변환하는 방안을 제시한다. 또한 한국어와 수화의 표현상에서 차이를 보이는 여러 언어현상을 재현, 생략, 변형, 이동 네 가지로 나누어 정리하여 이론적인 기틀을 마련하였다.

[5]에서는 수화 표현을 위한 수화 아바타의 손 모양을 편집할 수 있는 인터페이스를 제안한다. 손 모양 데이터를 코드번호로 표시하고 저장할 수 있어 수화 단어에 대한 사전 구성이 좀더 쉬워지며 데이터 용량을 줄이는 효과를 얻을 수 있다. 또한 수화 단어 생성에 있어서도 기존 수화 사전들이 텍스트 형식의 모호한 설명 방식을 택하여 정확한 수화 동작을 재현하기 힘들었으나 코드번호에 의한 아바타를 통해 정확한 표현을 제공할 수 있는 가능성을 보인다.

본 연구에서는 [3]의 시스템을 바탕으로 수화로의 변환 과정에서 나타나는 생략현상과 이동현상에 중점을 두어 이를 절 개념과 연계하여 구체적으로 논의하고자 한다.

### 3. 수화 표현의 특성

본 절에서는 수화에서 나타나는 특성들 중 하나인 공간성과 동시성에 대해 살펴보고, 이런 특성이 복문을 표현하는데 어떤 영향을 미치는지, 이에 따라 발생할 수 있는 의미 혼동을 막기 위해 어떤 표현 방식을 사용하는지 살펴보도록 한다.

#### 3.1 공간성과 동시성

한국 수화는 한국어와 달리 조사와 어미 같은 기능어에 해당하는 형식 형태소가 거의 없다. 따라서, 한국 수화에서는 어순을 문장 성분을 결정짓는 중요한 요인으로 해석하여 한국 수화를 일종의 고립어로 보는 견해도 있다[6]. 한국 수화에는 SV(주어-동사) 또는 SOV(주어-목적어-동사) 라는 기본 어순이 존재하는 것으로 알려져 있다. 그러나, 사실 수화는 공간개념과 가역/비가역성을 이용해 이보다 더 다양한 어순을 가진다[7,8]. 가역/비가역성은 개체간 의미적인 관계를 통해 어순에 관계 없이 주체와 대상이 분명하게 드러나는 경우로 자

유로운 표현이 가능하다. 공간개념은 공간상에 개체들을 배치하고 이들 개체들을 참조하여 수화 서술어의 공간이동을 취하는 것으로, 이 때 공간이동 하는 수화 동작의 시작점과 끝점을 통해 주체와 대상을 구별한다. 따라서 주어의 먼저 이야기 하거나 목적어를 먼저 이야기 하는 것에 관계없이 주체와 대상을 명확히 구분할 수 있다. 또한 서술어를 표현하는 동안 주어와 목적어도 공간 상에 유지되어 한 시점에 서술어와 함께 동시에 표현할 수 있는 것도 수화의 독특한 특징이다[9]. 수화의 이런 공간성과 동시성은 수화의 문장 구조에도 영향을 미치게 되는데, 수화 문장에서의 서술어는 주어 또는 목적어와 멀리 떨어져 표현하기보다는 한꺼번에 묶어 표현하는 경향을 보인다. 하나의 주술 관계 속에 또 다른 주술 관계가 들어있는 구조의 복문과 일부 성분의 탈락 현상을 보이는 병렬문의 경우 이런 수화의 공간성과 동시성을 잘 표현할 수 있도록 순서를 바꿔 의미 단위로 묶어 표현하는 경우를 볼 수 있다.

#### 표기법

본 논문에서는 지금까지 논의한 수화의 특성을 반영하여 다음과 같은 표기법을 사용하도록 한다.

- (1) 나' + 그' + 사랑하다'[나'->그']
- (2) 나' + 그' + 사랑하다'[그'->나']

(1)은 나'가 주체가 되어 그'를 사랑한다는 의미로, (2)는 그 반대를 의미한다. (위 표기법은 [3]이 제안한 시스템의 action script를 변형한 형태이다.)

#### 3.2 한국 수화의 복문

한국 수화의 복문과 중문에 대한 연구 결과는 단문에 비해 상당히 부족한 상황으로[6] 한국 수화에는 복문이 없을 것이라는 견해가 있었으나, 농학생 수화의 통사구조에 관한 연구[8]에 따르면, 한국수화에도 복문 구조의 문장이 있다고 한다. 그는 한국수화 복문의 성분절을 한국어 복문의 다섯 가지 유형으로 분류해 분석한 결과, 절의 성분을 표시하는 기호는 전혀 발견되지 않았으며 대신 의미, 화용론에 의지하여 표현하고 있는 것으로 분석하였다. 그러나 한국어 복문을 수화로 표현해 이에 대한 농학생의 이해력을 테스트한 연구[10] 결과를 살펴보면, 의미, 화용론에 의지하는 한국 수화의 복문 구조는 수신자가 그 의미를 명확하게 이해하기 쉽지 않은 구조임을 드러낸다. 따라서 복문을 표현할 때 그 구조가 내포구조를 갖는 경우 의미 관계를 명확히 전달하기 위해 이를 절 단위로 분리하여 표현하는 특징을 보인다. 3.3 절에서는 이에 대해 좀더 자세히 살펴본다.

#### 3.3 복문 내포 구조의 나열화

수화는 시각적인 언어이기 때문에 긴 문장을 계속적으로 나열하면 산만해지고 누가 누구에게 말하는 것인지 모호해진다. 이는 3.1에서 설명한 바와 같이 서술어

에 해당하는 수화 동작을 취할 때 앞에서 언급했던 객체들을 수시로 참조한다는 사실에서도 분명히 드러난다. 따라서 장황하게 표현하기 보다는 의미 단위로 묶어서 표현하는 것이 수화자나 수신자의 입장에서 의미 해석을 쉽고 명확하게 이해할 수 있다[11]. 이를 위해 한 문장 안에 긴 수식 내용이 들어 있는 (3)의 (a)와 같은 한국어 문장의 경우 주절의 주어와 서술어를 가까이 두기도 하며, (4)의 (a)와 같은 겹문장을 홑문장으로 나누어 표현하기도 한다. 홑문장으로 분리하는 작업은 의미 단위로 묶을 수 있으며, 수신자가 이해하기 쉬운 순서로 내용을 전개할 수 있다는 장점이 있다. 사건의 발생 순서, 인과 관계, 상태의 변화에 따라 내용을 전개하는 것은 수신자의 입장에서는 부담을 덜 느끼며 들을 수 있고, 수화자 입장에서는 보다 분명하게 의미 전달을 할 수 있다는 효과가 있다[11].

- (3) (a) 딸은 결혼식이 다음 주로 다가오자 분주해졌습니다.  
 (b) 딸' + 분주하다' + 다음주 + 결혼식' + 가깝다'
- (4) (a) 상호는 언제나 바다 이야기를 하는 혁이에게 산골에서 태어났다고 하였다.  
 (b) 혁이' + 언제나' + 바다' + 이야기하다' + 상호' + 산골' + 태어나다' + (상호') + (혁이') + 말하다'[상호->혁이']

그러나 (4)와 같이 관형절이 수식하는 명사가 관형절 안에서 탈락된 성분인 '관계 관형절'과 달리 (5)과 같은 '동격 관형절'은 무리하게 분리하여 표현하고자 할 경우 수식관계에 대한 정보를 잃게 되는 결과를 초래할 수 있어 수식관계를 유지하면서 일부 성분의 어순 변형으로 표현해주는 것이 필요하다. 즉, (5)의 경우 S[S'O'V]OV 구조를 [S'O'V]OSV 구조로 어순 이동을 하여 수식관계를 유지하면서 주술 호응관계도 명확히 보여줄 수 있다. 이는 [5]의 수화 동격관형절에 대한 연구<sup>1</sup> 결과와도 일치한다.

- (5) (a) 당신은 새끼 거북이가 바다를 헤엄치는 모습을 본 적이 있습니까?  
 (b) 새끼' + 거북이' + 바다' + 헤엄치다' + 모습' + 너' + 보다'[너->모습]' + 경험' + 까'

미국 수화(ASL)의 경우 수화의 문법적인 기능을 전달한다고 알려져 있는 비수지신호<sup>2</sup>를 통해 관계절을 표시

<sup>1</sup> 수화 복문에 관한 연구에서 동격관형절에 국한되는데, 이유, 방법, 장소 등의 구체적이지 않은 명사와 옆, 앞, 위 등의 위치를 표현하는 명사 그리고 냄새, 소리 등의 감각 명사를 수식하는 동격관형절(보문관형절)은 주명사 앞에 선행한다고 하였다.

<sup>2</sup> 손을 제외한 얼굴 표정, 입술의 움직임, 눈썹의 움직임, 어깨의 움직임을 의미한다.

한다는 연구[12] 결과가 있으나 한국 수화의 경우엔 아직 알려진 바가 없어 논의에서 제외한다<sup>3</sup>.

또한 주술 호응 관계 사이에 다른 주어, 술어가 들어 있는 명사절, 인용절, 부사절 (SS'V'V, SS'O'V'V 등)의 경우 그대로 수화로 표현하면 그 주술관계를 정확하게 파악하기 쉽지 않으며 오히려 어색하다. 따라서 예문 (6), (7)의 (b)에서와 같이 안쪽 내포절을 바깥쪽 절의 서술어와 위치를 바꿔 뒤로 빼는 것이 수화에서 일반적으로 표현하는 방식이다.

- (6) (a) 유전자를 연구한 과학자들은 인간이 계획적으로 설계되어 있음에 놀랍니다.  
 (b) 유전자' + (현미경') + 연구하다' + 박사' + 놀라다' + 무엇' + 사람' + 원래' + 누구' + 로부터' + 기술' + 계획'
- (7) (a) 공장장이 작업을 중지하라고 하셨습니다.  
 (b) 공장' + 장' + 지시하다' + 일' + 중지'

그러나 부사절의 하나로 분류되기도 하는 문장 안에 삽입된 종속절의 경우 중문의 형태로 바꿀 수 있는데, 이 경우 수화에서 사건의 발생 순서, 인과 관계, 상태의 변화에 따라 내용을 전개한다는 점을 고려한다면 삽입된 종속절은 주어 앞으로 빼는 것이 더 자연스럽고 의미 관계를 분명히 드러낼 수 있다. 예를 들어, (8)의 (a)는 '~아서'라는 인과관계 연결어미로 종속절이 내포되어 있는 문장이다. 이 경우 종속절이 원인이 되어 결과가 나타난 형태로 (b)처럼 표현한다.

- (8) (a) 나는 그가 돌아와서 기뻐다.  
 (b) 그' + 돌아오다' + 나' + 기쁘다'

다음 절에서는 앞에서 살펴본 수화 표현상의 특성들을 고려하여 내포구조의 복문을 보다 수화로 표현하기 적절한 구조로 재배열하는 시스템을 제안한다.

#### 4. 처리 및 구현

본 절에서는 결합범주문법을 이용하여 앞 절에서 정리한 한국어 복문을 수화 생성을 위한 스크립트로 변환해주는 처리 방법을 제안하고 실행 결과를 보인다.

##### 4.1 결합범주문법

결합범주문법(Combinatory Categorical Grammar, CCG)은 범주문법(Categorical Grammar, CG)에 결합자(combinator)가 접목된 것으로, 소수의 축약 규칙에 의해 구문 분석이 이루어지는 문법 체계로서, 병렬 구조 등의 복잡한 문형을 별도의 제약 조건 없이 처리할 수

<sup>3</sup> [13, 14] 연구를 살펴보면 비수지신호를 통해 조건문, 의문문, 부정문 등의 문종은 표시되고 있으나 관계관형절을 비롯한 내포문에 대한 연구결과는 아직 밝혀진 바가 없다.

있으며, 어휘 범주에 통사 정보 외에도 의미 정보, 담화 정보 등을 추가하면 축약 규칙을 통해 여러 단계의 분석을 한 단계의 유도과정으로 해결할 수 있는 장점을 갖는다[15].

## 4.2 시스템

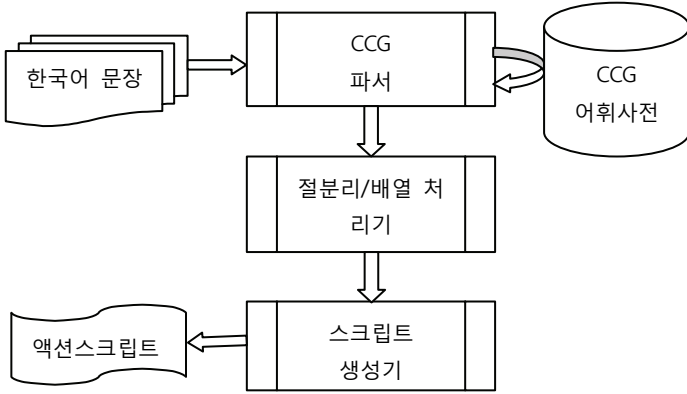


그림 1. 시스템 구조

전체적인 시스템의 구조는 그림1과 같이 세 부분으로 나누어 처리된다. 한국어 문장을 입력 받은 뒤 CCG 파서가 CCG 어휘 사전을 참조하여 결합범주문법을 통해 구문 분석을 한 후 그 결과를 절분리/배열 처리기에서 절의 종류, 개수, 구조에 따라 단순한 구조로 변환한다. 이를 액션 스크립트 생성기에서 공간을 할당하고 서술어 수화 어휘의 방향(수향)을 설정해 액션 스크립트를 생성한다. 본 시스템은 자바로 구현되었으며, CKY 알고리즘을 사용하였다.

- STEP1. CCG 파서를 통한 구문 분석
- STEP2. 절분리/배열 처리기를 통한 문장 구조 변환
- STEP3. 액션 스크립트 생성

### 4.2.1 어휘에 할당되는 범주

다음은 어휘사전 구성에 필요한 두 가지 기본 요소와 그 예이다.

```

np(case):word##h_num*class
(9) np(nom):나#one*per           [→ 나는]
s:[pred^word##h_num*drt, case^word##h_num*drt, ...]
(10) s:[pred^먹다#one*drtn, acc^자장면#one*etc,
nom^나#one*per]           [→ 나는 자장면을 먹는다.]
    
```

np는 명사구, s는 문장을 지칭하며, np의 case는 조사가 결합된 명사구의 격<sup>4</sup>을 표시한다. word는 수화 사전

<sup>4</sup> nom(주격), acc(대격), dat(여격), loc(처소격)을 의미한다.

의 기본어휘를 의미하며, h\_num, class, drt는 해당 어휘의 자질정보를 담고 있다. h\_num은 수화 어휘 표현시 필요한 손의 수 또는 지문자를 표시하고, class는 분류사로서 word의 의미에 따라 사람(per), 동물(ani), 장소(loc), 이동성(mov), 기타(etc)로 분류하여 표시되며, drt는 방향변화동사나 상태이동동사의 방향성유무(drty/drtn)를 표시한다. s에 표현된 어휘들은 인용절, 부사절, 부사를 제외하고 문장 안의 어순을 그대로 따른다(역방향).

### 4.2.2 구문 분석

한국어의 경우 내포절의 주어가 빈번하게 생략되거나, 관계관형절의 경우 수식하는 명사가 관계절 안에서 생략되는 현상을 보여 문장을 분석하는데 어려움이 있다. 절의 경계를 인식하는 연구(clause identification)는 구문 분석을 위해 활발히 연구된 분야 중 하나로, 대부분 규칙 기반이거나 기계 학습을 통해 처리한다[16].

본 연구에서는 결합범주문법을 이용해 구문 분석을 하여 빈번히 발생하는 생략 현상을 처리할 수 있도록 CCG 어휘사전에 하나의 서술어에 대해 여러 개의 가능한 어휘 항목(lexical entry)들을 구성해 구문 분석이 가능하도록 하였다(표1 참조). 또한 표2에 할당된 조사와 전성어미, 연결어미에 할당된 범주를 통해 절의 종류와 끝을 인식할 수 있도록 하였으며, 그와 동시에 결합범주문법 규칙을 통해 결합하는 동안 그 수를 제한시킬 수 있다. 그리고 잘못된 구문 분석 결과를 줄일 수 있도록 의미 정보 자질과 내포문 성분의 탈락규칙<sup>5</sup>[17]을 역이용하여 우선순위를 부여하였다.

표 1. 서술어 자릿수별 범주<sup>6</sup>

자릿수 서술어	범주
한 자릿수 서술어	s:[pred^word, -^]\np(-):- s:[pred^word1, case2^word2]\np(case2):word2
두 자릿수 서술어	s:[pred^word, -^, -^]\np(-):-\np(-):- s:[pred^word1, case2^word2, -^]\np(-):-\np(case2):word2 s:[pred^word1, case2^word2, case3^word3]\np(case3):word3\np(ca

<sup>5</sup> (1) 관계 대명사 탈락 규칙 - 관계절의 수식을 받는 명사구와 관계절 내의 어떤 명사구와의 사이에 상호지시성이 성립될 때 적용되는 규칙이다.

(2) 동일 명사구 탈락 규칙 - 일정한 조건 아래서 상위문의 주어들을 삭제시키는 통사규칙이다.

조건 1) 상위문장의 주어와 보문의 주어의 주어가 같을 때  
조건 2) 상위문장의 간접목적어(여격어)와 보문의 주어의 주어가 같을 때

<sup>6</sup> 라인(-)은 빈자리를 의미하며, 생략된 성분의 위치를 의미하기도 한다.

	se2):word2
세 자릿수 서술어	s:[pred^word, -^-, -^-, -^-\]np(-):-\np(-):- np(-):-
	s:[pred^word1, case2^word2, -^-, -^-\] np(-):-\np(-):-\np(case2):word2
	s:[pred^word1, case2^word2, case3^word3, -^-\]np(-):- np(case3):word3np(case2):word2
	s:[pred^word1, case2^word2, case3^word3, case4^word4]\np(case2):word2\np(ca se3):word3\np(case4):word4

표 2. 조사와 전성어미의 범주<sup>7</sup>

조사/어미	범주
조사	np(case):word\np(-):word np(case):word\np_c(-):word
관형사형 전성어미	np_c(case):[mn^word1, adnc^X]\np(case):word1\s:X np_c(case):[mn^word1, adnc^X]\np(case):word1\sz:X
명사형 전 성어미	np_c(-):[nc^X]\s:X np_c(-):[nc^X]\sz:X
인용조사	s:[Y, qc^X]\s:Y\s:X s:[Y, qc^X]\s:Y\sz:X sq:[Y, qc^X]\s:Y\s:X sq:[Y, qc^X]\s:Y\sz:X
부사형 전 성어미	s:[Y, advc^X]\s:Y\s:X s:[Y, advc^X]\s:Y\sz:X sv:[Y, advc^X]\s:Y\s:X sv:[Y, advc^X]\s:Y\sz:X
종속절 연결어미	s:[Y, subc^X]\s:Y\s:X s:[Y, subc^X]\s:Y\sz:X sc:[Y, subc^X]\s:Y\s:X sc:[Y, subc^X]\s:Y\sz:X
sq:[Y, qc^X]\s:Y	(s:[Y, qc^X]\np(C1):A)/(s:Y\np(C1):A) (s:[Y, qc^X]\np(C1):A\np(C2):B)/ (s:Y\np(C1):A\np(C2):B)
sv:[Y, advc^X]\s: Y	(s:[Y, advc^X]\np(C1):A)/(s:Y\np(C1):A) (s:[Y, advc^X]\np(C1):A\np(C2):B)/ (s:Y\np(C1):A\np(C2):B) (s:[Y, advc^X]\np(C1):A\np(C2):B\np(C3): C)/(s:Y\np(C1):A\np(C2):B\np(C3):C)
sc:[Y, subc^X]\s: Y	(s:[Y, subc^X]\np(C1):A)/(s:Y\np(C1):A) (s:[Y, subc^X]\np(C1):A\np(C2):B)/ (s:Y\np(C1):A\np(C2):B) (s:[Y, subc^X]\np(C1):A\np(C2):B\np(C3): C)/(s:Y\np(C1):A\np(C2):B\np(C3):C)
sz:X	s:X\np(-):- s:X\np(-):-\np(-):- s:X\np(-):-\np(-):-\np(-):-

<sup>7</sup> 의미 정보에 나타나는 qc(인용절), adnc(관형사절), nc(명사절), mn(수식받는명사), advc(부사절)을 의미하며, 통사 정보에 나타나는 s(문장 또는 절), np(명사구), np\_c(절을 내포한 명사구)을 의미한다. 그리고 인용절, 부사절, 종속절에 나타나는 sq, sv, sc, sz 는 표기의 편의를 위한 약자로 sq, sv, sc는 결합되는 범주에 따라 달라지며, sz는 생략된 성분을 지닌 절을 의미한다.

#### 4.2.3 절 분리/배열

구문 분석 결과를 통해 절의 종류와 절의 경계를 인식하고 이를 앞에서 논의한 대로 수신자가 이해하기 쉬운 구조로 전환한다. 이를 위해 표3과 같은 기본적인 규칙을 정의한다. 같은 수위의 문장구조에서 내포된 관형사절이 하나인 경우 어순 이동을 통해 앞으로 보내고, 둘 이상인 경우 하나를 제외한 나머지 절들은 수식 구조를 분리하여 별도의 문장으로 처리한다. 그리고 이로 인해 발생할 수 있는 문장과 문장 사이의 어색한 전개 흐름을 막아 매끄럽게 이어질 수 있도록 분리한 나머지 문장들은 완전한 문장으로 바꾸어 기존에 수식하고 있던 명사에 대한 추가 설명을 하는 구조로 배치한다. 다만 관계관형사절을 별도의 문장으로 분리시킬 경우, 수식하는 명사가 해당 관계관형사절 안에서 생략되어 있기 때문에 이를 관형절의 서술어 앞으로 복원시킨 후 분리한다. 내포된 명사절, 인용절, 부사절의 경우 3절에서 논의된 대로 서술어 뒤로 이동시키며, 문장 안으로 삽입된 종속절의 경우 주어 앞으로 이동시킨다. 절 안에 절이 두 번 이상 내포된 재귀적인 구조의 경우 안쪽에서부터 위에서 논의된 규칙을 적용하여 재귀적으로 처리한다.

표 1. 문장 구조 변경<sup>8</sup>

문장 구조	변형 전	변형 후
관형사절 내포 문장	S[S'V']OV S[S'O'V']OV	[S'V']OSV [S'O'V']OSV
	[S'V']S[S'O'V'] OV	[S'V']SOV + w(의문 사) + S'O'V"
명사절, 인용절 내포 문장	S[S'O'V']V S[S'V']V	SV[S'O'V'] SV[S'V']
부사절 내포 문장	S[S'O'V']V S[S'V']V SO[S'V']V	SV[S'O'V'] SV[S'V'] SOV[S'V']
종속절 내포 문장	S[S'O'V']V S[S'V']V SO[S'V']V	[S'O'V']SV [S'V']SV [S'V']SOV
나머지 구조의 문장	[S'O'V']SV [S'V']SV [S'V']SOV	그대로

#### 4.2.4 스크립트 생성

이해하기 쉬운 단문의 나열식으로 전환된 결과를 바탕으로 향후 수화 애니메이션을 생성할 수 있도록 액션 스크립트를 생성한다. 액션 스크립트는 그림 2와 같이 네 개의 필드로 구성된다. sign\_word 필드는 수화 동작 데이터베이스를 참조하기 위한 수화 기본 어휘를 표시하며, position|direction 필드는 서술어를 제외한 논항의 경우 어휘의 class 자질 정보와 case 자질 정보를

<sup>8</sup> 일반적으로 한국어 복문의 경우 내포절 안에 다른 내포절이 나타나는 구조를 보이지만, 같은 수위에서 내포절이 3개 이상 등장하는 경우는 드물다.

이용하여 공간 할당 여부를 결정하여 공간상의 특정 위치를 할당 받는다. 공간 할당은 일인칭 ‘나’, 이인칭 ‘너’는 각각 ‘1’, ‘2’를, 나머지는 언급되는 객체 순서로 ‘3’부터 순차적으로 할당 받는다. 방향 변화 서술어(drty) 자질 정보를 지닌 서술어)의 경우 논항들의 position 정보와 case 자질 정보를 이용하여 주체가 대상을 향하도록 서술어의 공간 이동 방향(direction)을 지정해준다. 공간 또는 방향을 지정할 필요가 없는 나머지 경우 ‘0’으로 표시한다. 세 번째 손 할당 필드는 position 정보와 h\_num 자질 정보를 고려하여, 한 손 수화의 경우 ‘left\_h’ 또는 ‘right\_h’로 지정하며, 양 손 수화의 경우 ‘both\_h’로 지정한다. 또한 h\_num 자질 정보가 ‘fing’로 표시된 경우 sign\_word 필드에 ‘f\_spell(sign\_word)’로 표시하여 지문자로 표현될 수 있도록 하며, 손 정보는 ‘right\_h’로 지정한다. 마지막으로 next\_start\_time 필드는 이어진 단문의 나열, 끊어진 단문의 나열, 끊어진 문장을 구별해 주기 위해 다음 수화 동작이 시작되는 상대적인 시간 간격을 순서대로 각각 ‘0’, ‘1’, ‘2’의 값으로 지정한다.

Act(sign\_word, position|direction,  
left\_h|right\_h|both\_h, next\_start\_time)

그림 2. 액션 스크립트 폼

#### 4.2.5 결과

그림 3은 관계 관형절과 인용절이 내포된 예문 (4)를 본 시스템에서 실험한 결과이다. <STEP1>은 CCG 파서를 통해 구문 분석된 결과로 생략현상처리를 위해 서술어에 여러 어휘 항목을 할당한 결과 <0>~<2> 세 가지 구문 분석 결과가 얻어진다(그림 4, 5 참고). 이 중 전체 문장의 주어가 생략된 <2>의 결과가 내포문의 주어가 생략된 <0>, <1> 결과보다 우선 순위가 낮으므로 <0>, <1>를 이용해 절분리/배열 시킨다. <STEP2>는 절분리/배열 처리기에서 구문 분석 결과를 통해 절의 종류와 위치, 수를 파악하여 관형사절(adnc)과 수식 받는 명사는 주어 앞으로 이동시키고, 인용절(qc)은 서술어 뒤로 재배열 시킨다. 그러나 위의 두 결과는 부사의 위치를 제외하고는 같은 구조를 지니므로 같은 결과가 유도된다. <STEP3>은 <STEP2>의 결과를 이용해 액션스크립트로 변환한 결과이다. ‘혁이’와 ‘상호’의 경우 고유명사로 ‘fing’이라는 정보를 통해 f\_spell로 표현이 바뀌었으며, ‘per’과 ‘loc’라는 class 정보를 이용해 ‘사람’과 ‘장소’라는 동작이 추가되었다. class와 case 자질 정보에 따라 공간 위치 정보가 숫자 ‘3’, ‘4’, ‘5’로 표시되었다. 또한 drty 자질을 가지고 있는 동사 ‘말하다’의 경우 case 와 position 정보를 활용해 이동 방향이 ‘4->3’로 표시되었다. 어휘와 어휘 사이, 분리되지 않은 관형사절과 수식 받는 어휘 사이는 ‘0’, 나머지 절과 절 사이의 간격은 ‘1’, 문장과 문장 사이는 ‘2’로 다음 수화 동작과의 시간간격이 설정되었다.

(4) 상호는 언제나 바다 이야기를 하는 혁이에게 산골에서 태어났다고 말했다.

#### <STEP1>

<0>

```
s:[pred^say#one*drty, dat^[mn^hyeki#fing*per,
adnc^[pred^do#two*drtn, acc^[n2^story#two*etc,
n1^sea#two*loc], nom^-]], nom^sangho#fing*per],
qc^[pred^born#two*drtn,
loc^mountain_region#two*loc, nom^-]],
adv^always#two*etc]
→ [상호는 언제나 [바다 이야기를 하는] 혁이에게
[산골에서 태어났다고] 말했다]
```

<1>

```
s:[pred^say#one*drty, dat^[mn^hyeki#fing*per,
adnc^[pred^do#two*drtn, acc^[n2^story#two*etc,
n1^sea#two*loc], nom^-, adv^always#two*etc]],
nom^sangho#fing*per], qc^[pred^born#two*drtn,
loc^mountain_region#two*loc, nom^-]]
→ [상호는 [언제나 바다 이야기를 하는] 혁이에게
[산골에서 태어났다고] 말했다]
```

<2>

```
s:[pred^say#one*drty, dat^[mn^hyeki#fing*per,
adnc^[pred^do#two*drtn, acc^[n2^story#two*etc,
n1^sea#two*loc], nom^sangho#fing*per],
adv^always#two*etc]], nom^-,
qc^[pred^born#two*drtn,
loc^mountain_region#two*loc, nom^-]]
→ [[상호는 언제나 바다 이야기를 하는] 혁이에게
[산골에서 태어났다고] 말했다]
```

#### <STEP2>

```
adv^always#two*etc n1^sea#two*loc
acc^story#two*etc pred^do#two*drtn
dat^hyeki#fing*per nom^sangho#fing*per
pred^say#one*drty
/ loc^mountain_region#two*loc pred^born#two*drtn
```

#### <STEP3>

```
Act[always, 0, both_h, 0]
Act[sea, 0, both_h, 0]
Act[story, 0, both_h, 0]
Act[do, 0, both_h, 0]
Act[f_spell(hyeki), 0, right_h, 0]
Act[person, 3, right_h, 0]
Act[f_spell(sangho), 0, right_h, 0]
Act[person, 4, left_h, 0]
Act[say, 4->3, left_h, 1]
Act[mountain_region, 0, both_h, 0]
Act[location, 5, right_h, 0]
Act[born, 0, both_h, 2]
```

그림 3. 예문 (4) 처리 과정과 결과

본 절에서 논의된 사항들이 잘 처리되기 위해서는 CCG 어휘 사전의 체계적인 구성이 무엇보다 중요하며, 구문 분석 결과의 수를 줄이기 위해서는 앞에서 제안한 방법 외에 보다 정교한 어휘 자질 정보들이 요구된다. 서술어 어휘 항목의 논항 자질 정보를 보다 구체화하여 제한한다면 잘못된 구문 분석 결과를 줄일 수 있을 것으로 생각된다. 또한 관형사절 처리의 경우 본 시스템에서는 관계 관형절과 동격 관형절을 구별하지 않고 처리하고 있으나 이 역시 동격 관형절의 수식을 받는 어휘가 제한적이라는 특징을 고려할 때 어휘 자질 정보를 추가하면 해결될 수 있으리라 생각된다. 그리고 절분리/배열의 경우 재귀적인 구조는 안쪽부터 해결하며 절이 절을 내포한 경우 절의 종류에 관계없이 서술어 뒤로 분리하여 처리하고 있는데 뒤로 보내진 절들의 순서는 먼저 분석된 순서를 따르고 있다. 이에 따라 복잡한 구조의 문장의 경우 뒤로 보내 표현된 절 간에 다소 흐름이 어색해질 수도 있다는 문제가 있다. 따라서 이들간의 순서를 규정하는 규칙이 필요하나 이에 대한 수화 문장을 수집하는데 어려움이 있어 이는 향후 과제로 남긴다. 본 시스템에서 테스트한 복문은 [10]에서 논의된 문장들로서 이를 시스템에서 입력으로 받아 생성된 표현이 실제 수화 표현 방식과 일치하는지 여부를 살펴보았다. 그러나 이와 같은 방식의 타당성을 보다 객관적으로 검증하기에 복문의 유형과 수효가 많이 부족하여 다른 방안으로 시스템의 성능을 확인할 필요가 있는 것으로 보인다.

## 5. 결론 및 향후 계획

한국 수화는 시각, 공간언어로 한국어와는 상이한 문법체계를 가진 언어이다. 따라서 수화를 일차언어로 사용하는 농인들에게 있어 복잡한 구조의 한국어 문장은 부담이 따르게 된다. 이를 해결해주기 위해서는 한국어 문장을 수화로 변환해주는 시스템이 요구된다.

본 논문에서는 이를 위해 수화에서의 표현 특징을 살펴보고, 이를 바탕으로 한국어 복문 구조를 분석하고 이를 농인들이 이해하기 쉬운 구조로 변환한 후 이를 단순 수화 생성을 위해 공간 할당과 공간이동이 가능하도록 하였다. 이를 위해 분류사 정보를 활용하여 공간 위치 정보와 손을 할당하고, 서술어의 공간 이동 방향을 지정하는 액션스크립트 생성 시스템을 제안하였다. 그러나 제안된 방법이 잘 수행되기 위해서는 보다 정교한 어휘 사전 구축 작업이 추가적으로 이루어져야 하며 테스트 자료 수집을 통한 보다 실질적인 검증 작업도 요구된다.

수화 애니메이션 생성을 위해서는 스크립트 생성 외에 수화 동작 데이터베이스 구축도 함께 이루어져야 한다. 향후 연구에서는 수화 동작 데이터베이스를 구축하는 방안과 데이터베이스에서 불러온 기본 동작을 액션 스크립트를 통해 조정해주는 방안에 대한 연구도 진행

할 예정이다.

## Acknowledgements

본 연구는 첨단정보기술 연구센터를 통하여 한국과학재단의 지원을 받았음

## 참고문헌

- [1] 최영주. 청각장애 학생의 독해력 분석, 대구대학교 석사학위논문, 1990.
- [2] 이용원, 최창석. 수화 자동 애니메이션을 위한 국어문장을 수화문장으로의 변환 방법, HCI, 2002.
- [3] 최지원, 박종철. 수화 자동 생성을 위한 한국어 문장 분석과 처리, 제 15회 한글 및 한국어 정보처리 학술대회, 2003.
- [4] 최지원, 박종철. 결합범주문법을 이용한 수화 자동 생성, HCI, 2003.
- [5] 오영준, 박광현, 변증남. 수화 표현을 위한 손 모양 편집 프로그램의 개발, 전자공학회논문지, 제44권 SC호, pp. 48~54, 2007.
- [6] 엄미숙. 한국수화의 통사론적 특성 분석, 대구대학교 석사학위논문, 1996.
- [7] 석동일. 한국수화의 언어학적 분석, 대구대학교 박사학위논문, 1989.
- [8] 황도순. 농학생 수화의 통사 구조와 통사론적 특징, 특수교육논집, 1994.
- [9] Peng, F.C. Sign language and language acquisition in man and ape, AAAs selected symposium, 1975.
- [10] 최선미. 수화의 방법적 유형에 따른 농학생의 문장 이해 능력 분석, 대구대학교 석사학위논문, 1993.
- [11] 박인선 외. 수화 업그레이드, 이한출판사, 2000.
- [12] Liddell, S.K. Nonmanual Signals and Relative Clauses in American Sign Language, Academic Press, 1978.
- [13] 손천식. 자연수화실현에 대한 이해. 수화언어의 미래를 생각하는 세미나. 한국수화연구회, 2000.
- [14] 윤병천. 한국수화의 비수지신호에 대한 언어학적 특성 연구, 특수교육저널:이론과 실천, 제5권 1호, pp. 253~277, 2004.
- [15] Steedman, M. The Syntactic Process, The MIT Press, 2000.
- [16] Kim, M.Y., Lee, J.H.. Two-Phase S-Clause Segmentation, IEICE TRANS. INF. & SYST., Vol.E88-D, No.7, 2005.
- [17] 남기심. 언어학 개론, 탑출판사, 1985.

상르는	언제나	바다이야기를	하	는	력이	에게	산골에서	태어났	다고	말했다.
np(nom):sang ho#fing*per	{s:[X, adv^always#two o*etc]np(Case 1):Y}(s:X)np(C ase1):Y	np(acc):[n2^story# two*etc, n1^sea#two*loc]	s:[pred^do#two o^drtn, acc^X, nom^-]np(-): Y(s:X)np(-): Y	np_c(-):[mn^Y, adnc^X]np(-): Y(s:X)np(-): Y	np(-): hyeki #fing* per	np(dat): Xnp_c(-) )X	np(loc):moun tain_region# wo*loc	s:[pred^born# wo^drtn, loc^X, nom^-]np(-): Ynp(loc):X	{s:[Y, qc^X]np(Case 2):Znp(Case1 ):W}(s:Y)np(C ase2):Znp(Ca se1):W}(s:X)np(-): Y	s:[pred^say#o ne^dtry, dat^X, nom^Y]np(nom): Ynp(dat):X
		s:[pred^do#two^drtn, acc ] [n2^story#two*etc, n1^sea#two*loc] , nom^-]np(-): np_c(-):[mn^Y, adnc^X]np(-): acc^[n2^story#two*etc, n1^sea#two*loc] , nom^-]np(-):Y					s:[pred^born#two^drtn, loc^ mountain_region#two*loc, nom^-]np(-): (s:[Y, qc^[pred^born#two^drtn, loc^ mountain_region#two*loc, nom^-]]np(Case2):Znp(Case1)W) /(s:Y)np(Case2):Znp(Case1)W) s:[pred^say#one^dtry, dat^X, nom^Y], qc^[pred^born#two^drtm, loc^ mountain_region#two*loc, nom^-]]np(nom):Ynp(dat):X			
		np_c(-):[mn^ hyeki #fing*per, adnc^X]np(-): acc^[n2^story#two*etc, n1^sea#two*loc] , nom^-]								
		np(dat):[mn^ hyeki #fing*per, adnc^X]np(-): acc^[n2^story#two*etc, n1^sea#two*loc] , nom^-]								
		s:[pred^say#one^dtry, dat^[mn^ hyeki #fing*per, adnc^X]np(-): acc^[n2^story#two*etc, n1^sea#two*loc] , nom^-], qc^[pred^born#two^drtm, loc^ mountain_region#two*loc, nom^-]]np(nom):Y								
		s:[[[pred^say#one^dtry, dat^[mn^ hyeki #fing*per, adnc^X]np(-): acc^[n2^story#two*etc, n1^sea#two*loc] , nom^-], qc^[pred^born#two^drtm, loc^ mountain_region#two*loc, nom^-]], adv^always#two*etc]np(nom):Y								
		s:[[[pred^say#one^dtry, dat^[mn^ hyeki #fing*per, adnc^X]np(-): acc^[n2^story#two*etc, n1^sea#two*loc] , nom^-], nom^sangho#fing*per, qc^[pred^born#two^drtm, loc^ mountain_region#two*loc, nom^-]], adv^always#two*etc]								

그림 4. 생략현상에 대한 서술어 복수 범주를 활용한 예문 (4) 구문 분석 결과 <0>

상르는	언제나	바다이야기를	하	는	력이	에게	산골에서	태어났	다고	말했다.
np(nom):sang ho#fing*per	{s:[X, adv^always#two o*etc]np(-): Y}(s:X)np(-): Y	np(acc):[n2^story# two*etc, n1^sea#two*loc]	s:[pred^do#two o^drtn, acc^X, nom^-]np(-): Y(s:X)np(-): Y	np_c(-):[mn^Y, adnc^X]np(-): Y(s:X)np(-): Y	np(-): hyeki #fing* per	np(dat): Xnp_c(-) )X	np(loc):moun tain_region# wo*loc	s:[pred^born# wo^drtn, loc^X, nom^-]np(-): Ynp(loc):X	{s:[Y, qc^X]np(Case 2):Znp(Case1 ):W}(s:Y)np(C ase2):Znp(Ca se1):W}(s:X)np(-): Y	s:[pred^say#o ne^dtry, dat^X, nom^Y]np(nom): Ynp(dat):X
		s:[pred^do#two^drtn, acc ] [n2^story#two*etc, n1^sea#two*loc] , nom^-]np(-): np_c(-):[mn^Y, adnc^X]np(-): acc^[n2^story#two*etc, n1^sea#two*loc] , nom^-]np(-):Y					s:[pred^born#two^drtm, loc^ mountain_region#two*loc, nom^-]np(-): (s:[Y, qc^[pred^born#two^drtm, loc^ mountain_region#two*loc, nom^-]]np(Case2):Znp(Case1)W) /(s:Y)np(Case2):Znp(Case1)W) s:[pred^say#one^dtry, dat^X, nom^Y], qc^[pred^born#two^drtm, loc^ mountain_region#two*loc, nom^-]]np(nom):Ynp(dat):X			
		np_c(-):[mn^ hyeki #fing*per, adnc^X]np(-): acc^[n2^story#two*etc, n1^sea#two*loc] , nom^-], adv^always#two*etc]								
		np(dat):[mn^ hyeki #fing*per, adnc^X]np(-): acc^[n2^story#two*etc, n1^sea#two*loc] , nom^-], adv^always#two*etc]								
		s:[pred^say#one^dtry, dat^[mn^ hyeki #fing*per, adnc^X]np(-): acc^[n2^story#two*etc, n1^sea#two*loc] , nom^-], qc^[pred^born#two^drtm, loc^ mountain_region#two*loc, nom^-]]np(nom):Y								
		s:[[[pred^say#one^dtry, dat^[mn^ hyeki #fing*per, adnc^X]np(-): acc^[n2^story#two*etc, n1^sea#two*loc] , nom^-], adv^always#two*etc], nom^ sangho#fing*per, qc^[pred^born#two^drtm, loc^ mountain_region#two*loc, nom^-]]								

그림 5. 생략현상에 대한 서술어 복수 범주를 활용한 예문 (4) 구문 분석 결과 <1>