

수화 자동 생성을 위한 한국어 문장 분석과 처리

최지원 박종철

한국과학기술원 전산학전공 및 첨단정보기술 연구센터

{identity, park}@nlp.kaist.ac.kr

Analysis and Computational Processing of Sentences in Korean for Automatic Sign Language Generation

Jiwon Choi Jong C. Park

Computer Science Division & AITrc, KAIST

요 약

한국 수화는 한국어에 대한 기본적인 유사성을 가지고 있지만, 교착어이자 청각-음성 체계 언어인 한국어와는 달리 고립어이자 시각-운동 체계 언어로서의 특성을 동시에 나타내고 있다. 그러므로 텍스트 형태의 한국어 문장으로부터 수화를 자동 생성하기 위해서는 한국어를 위해 미리 정의된 문법에 수화 표현을 무리하게 연계 시키려고 하기 보다, 수화 고유의 의미 전달 체계를 분석하고 활용하여야 할 필요가 있다. 본 논문에서는 수화 표현상의 언어학적 특징을 재현·생략·변형·이동의 네 가지로 구분하여 분석하고 결합범주문법을 이용한 이 같은 현상의 처리 방법 및 구현 방안에 대하여 논의한다.

1. 서 론

사회가 발전함에 따라 사람들의 생활 수준과 함께 의식 수준이 향상되면서 장애인의 복지에 대한 사회적 관심이 점차 높아져 가고 있으며, 이에 따라 청각 장애인의 모국어라 할 수 있는 수화도 이전 어느 때보다 많은 주목을 받고 있다.

최근에는 텔레비전에서 건청인에 비해 정보 획득에 많은 어려움을 겪고 있는 청각 장애인의 이해를 돕기 위해 일부 프로그램에서 자막 방송을 내보내는 등 여러 의미 있는 시도가 이루어지고 있지만, 대부분의 청각 장애인들은 수화를 문자에 비해 훨씬 친근하게 느끼고 잘 이해하는 경향이 있다. 이처럼 사회적 관심과 함께 자연언어 문장의 수화 변환에 대한 필요성이 대두됨에 따라 한국어 단어와 수화 단어를 일대일 방식

으로 대응시키고 형태소 변환 규칙을 적용한 후 종류별 수화 사전 검색을 통하여 수화 문장을 생성하는 방식의 연구가 진행되어 왔으나, 한국어를 위해 정의된 문법에 수화 표현을 연계 시키고자 하는 구조적인 한계점으로 인해 자연언어 문장을 수화 고유의 의미 전달 체계로 표현하는 데 문제점을 드러내고 있는 것이 사실이다. 자연언어 문장이 가지고 있는 최초의 정보를 최대한 유지하면서 이를 동일한 의미의 수화 체계로 변환하기 위해서는 고립어이자 시각적 언어로서 나타나는 수화 의미·표현상의 특징이 기존의 문장 변환 관련 연구에 비해 보다 다양하게 고려되어야 할 필요성이 있다.

본 연구에서는 텍스트 형태의 한국어 문장을 수화로 변환할 때 나타나는 수화 표현의 언어학적 특징을 재현·생략·변형·이동의 네 가지로 구분하여 정리하고 이에 대한 처리 방법을 제안한다. 본 논문의 2절에서는 수화

* 본 연구는 첨단정보기술 연구센터를 통하여 한국과학재단의 지원을 받았음

에 대한 기존의 언어학적 연구 및 자연언어 문장을 수화 문장 형태로 변환하는 연구에 대해 살펴보고, 3절에서는 한국어 문장을 수화로 변환할 때 일어나는 수화 표현의 재현·생략·변형·이동 현상 및 그 특징을 예문을 바탕으로 분석한다. 4절에서는 3절에서 제시한 문제에 관해 결합범주문법을 이용한 처리 방법을 제안하고 구현 시스템을 살펴본다. 5절에서는 결론 및 향후 계획에 대해 논의한다.

2. 관련 연구

본 절에서는 수화의 언어학적 연구에서 나타난 수화의 방법적 유형 및 특징과 함께 자연언어 문장을 수화 문장으로 변환하는 기존의 연구에 대해 살펴본다.

2.1 수화의 방법적 유형 및 특징

수화의 유형은 수화 언어의 구성 요소 및 표현 방법에 따라 자연 수화와 문법 수화로 나눌 수 있다. 문법 수화는 문자 언어의 어순에 따라 필요한 문법적 기호나 단어의 변화를 수반하는 수화이며, 자연 수화는 청각장애인의 마음에서 생긴 관념에 따라 형성된 자연 발생적 수화이다[표 1].

구분	자연 수화 (전통적 수화)	문법 수화
국어와의 일치도	말의 의미, 문법을 국어와 일치시킴을 전제하지 않고, 다소 독자적인 문법을 가지고 있다.	말의 의미, 문법이 국어와 일치하며 조사나 용언의 활용, 어미 사용이 특히 현저하다.
이해의 용이성	국어를 이해하지 못해도 직관적으로 이해하기 쉽다.	국어 미습득자는 이해에 다소 어려움을 가진다.
지문자 사용	사용하나 적다.	적극적으로 사용한다.
구어병용	병용을 전제하지 않는다.	병용을 전제한다.

[표 1] 한국의 자연 수화와 문법 수화의 언어적 특징[4]

이 두 유형에 대한 장단점은 지금까지도 활발히 연구되고 있으나 아직도 어느 것이 더 바람직하다는 뚜렷한 결론을 얻지 못하고 있다. 미국이나 일본의 경우에는 1970년대부터 자연 수화와 문법 수화의 차이를 좁히기 위하여 지속적인 연구를 진행하여 왔으며 그 결과로 양

자의 간격이 점차 좁혀져 가고 있는 추세이다[4]. 이에 비해 한국 수화에 대해서는 상대적으로 이러한 연구가 활발히 이루어지지 못하였다. 몇몇 연구에서는 문법 수화의 틀을 그대로 유지하면서 품사별로 형태소 변환 규칙을 적용한 혼합식 수화를 대안으로 제시하고 있으나, 이 또한 기본적으로 문법 수화의 틀에서 많이 벗어나지 못하고 있다.

한국 수화는 한국어와 기본적인 유사성을 지니지만, 다음의 두 가지에서 커다란 차이점을 나타내고 있다. 한국어가 의미 형태소에 조사·어미와 같은 형식 형태소가 붙음으로써 문법적 기능을 하게 되는 교착어(agglutinative language)인 반면, 한국 수화는 문법적인 관계를 나타내는 형태소의 발달을 거의 찾아보기 힘들고 각각의 말이 독립되어 일반적으로 어순에 따라 문법적인 기능을 하는 고립어(isolated language)와 유사하다[5]. 또한 수화는 청각으로 이해되고 음성으로 표현되는 청각-음성 체계의 음성 언어와 달리 시각으로 이해되고 손 운동으로 표현되는 시각-운동 체계로[11], 그 형태와 의미는 가지고 있으나 태생적 특성상 문자 언어 체계를 갖고 있지 않은 독특한 언어이다. 때문에 고립어의 형태를 띠면서 시각적으로 표현하는 언어인 수화의 근본적인 속성으로 인한 의미·표현상의 특징을 고려하지 않고서는 제대로 된 번역을 기대하기 어렵다.

2.2 자연언어 문장의 수화 변환 연구

[6]에서는 국문법 체계의 수화 문장을 수화 문법 체계의 문장으로 변환하는 방법을 제안하고 있으며 이를 위해 수화 단어의 종류를 8가지로 정의하고 품사별로 5가지의 수화 사전을 구성하였다. 여기에서는 입력 문장을 우선 어절별로 분리한 후, 체언과 조사, 용언의 기본형과 어미, 접두사와 접미사, 문장 기호, 숫자의 단위별 구문분석을 거친 다음 분석된 단어 및 음소별로 종류별 수화 사전을 참조하여 문법식 또는 혼합식 수화 체계로 변환한다.

[9]에서는 한국어의 각 품사별로 형태소 해석 규칙에 대응하는 수화 형태소 규칙을 제안하고 있는데, 자연

수화의 특징을 이용하고자 한 노력이 돋보이지만 그 적용 범위가 형태소 단위에 머무르고 있고, 하나의 형태소에 대해 제안한 규칙이 복수 개일 경우 실제 시스템에서 어떤 방식으로 규칙을 선택할 것인지에 대해서는 자세히 다루고 있지 않다.

[14]는 일본어를 일본 수화로 변환해 주는 연구로, 주된 내용은 일본 수화 어휘가 일본어에 비해 많이 부족하기 때문에 일어나는 수화 번역상의 어려움을 해결하기 위해 입력 단어가 수화 사전에 없을 경우 일본어 단어 사전을 검색, concept identifier와 정의 문장, concept hierarchy를 활용하여 적당한 대역어를 선택하게 하는 것이며 일본어에 맞는 수화 기술 방법도 제안하고 있다.

3. 문장 변환 상에서 나타나는 수화 표현의 재현·생략·변형 및 이동 현상

본 절에서는 한국어 문장을 수화로 변환할 때 일어나는 수화 표현의 재현·생략·변형·이동 현상과 그 특징을 예문을 바탕으로 분석한다.

3.1 함축적 표현의 수화 재현

한국어 문장에서 독립적 어휘로 나타나 있지 않은 표현이 수화에서 하나의 독립 어휘로 재현되는 경우는 크게 다음의 세 가지로 나누어 볼 수 있다.

3.1.1 형식 형태소의 독립 어휘 재현

한국어는 고립어인 수화와는 다르게 어미나 조사, 보조용언 등의 형식 형태소가 중요한 문법적 기능을 많이 담당한다. 일반적으로 어순에 따라 결정되는 문장 성분과 문장 내 내용어의 의미 관계 유추를 통해 전체 문장의 의미 파악에 문제가 없을 경우 종종 생략하기도 하나 그렇지 않은 경우에는 문장에서 나타나는 형식 형태소를 별도의 수화 어휘로 재현해야 할 필요성이 생기게 된다.

시제 표현의 경우 한국어 문장에서는 시제어미나 관형화된 어미의 굴곡적 방법에 의해 실현되지만 어미의 사용이 거의 없는 한국 수화에서는 다른 방법으로 시제

를 실현하게 되는데[5] 여러 가지 실현 방법이 있으나 가장 간단하게는 용언의 선어말 어미 또는 시제와 관련된 관형사형 어미를 통해 문장의 시제를 파악하여 내용어 뒤에 해당 시제를 나타내는 적절한 수화 어휘를 별도로 표현해 줌으로써 표현이 가능하다. (1)에서 (a)와 같은 과거형의 문장의 경우, (b)와 같이 ‘끝’이라는 어휘를 표현해주는 방식이다(cf.[9]).

- (1) (a) 영화가 왔다.
(b) [영화 + 오다 + 끝]

이 밖에도 문장의 의미를 정확히 전달하기 위해서는 조사와 어미를 상응하는 의미의 독립된 수화 어휘로 표현해야 하는 경우가 많이 있다. (2)의 (a)와 같은 조사는 ‘입장’, (b)와 같은 조사는 ‘이용하다’, (c)와 같은 조사는 ‘따옴표’와 같이 별도의 어휘를 사용하여 조사의 의미를 나타낸다[2].

- (2) (a) ‘~로서’ : 지위나 신분, 자격을 나타냄
(b) ‘~로써’ : 수단이나 도구를 나타냄
(c) ‘~라고’ : 인용을 나타냄
- (3) (a) 비에 젖었다
(b) 학교에 가자
(c) 오후에 만나자

하지만 문자상으로는 동일하게 표현된다 하더라도 그 의미가 다른 조사의 경우에는 앞서 언급한 조사들과 같이 단순한 대응으로는 정확한 의미를 전달하기 어렵다. 예를 들어 ‘~에’와 같은 조사의 경우, (a)에서는 ‘원인’의 의미를, (b)에서는 ‘장소’의 의미를, (c)에서는 ‘시간’의 의미를 각각 나타내기 때문에, 조사를 대치할 어휘를 정하기 이전에 문장의 의미 분석이 우선적으로 이루어져야 한다.

3.1.2 수화 어휘의 능동적 삽입

수화 문장의 의미 관계를 보다 명확하게 표현하기 위해서는 앞서 언급한 것과 같은 기능어를 독립적 어휘로 표현할 뿐 아니라 한국어 문장상에는 나타나 있지 않은 어휘를 상황에 따라 능동적으로 수화 표현에 삽입해 주어야 한다.

- (4) (a) 어렸을 때 어른들께서 장래 희망이 무엇이나고 물으시면 대통령이 되고 싶다고 대답했습니다
 (b) [나+ 아이+ 때+ 어른들+ 묻다+ 앞으로+ 희망+ 무엇+ 나+ 대통령+ 되다+ 원하다+ 대답하다]

예를 들어 (4)의 (a)처럼 비교적 길고 복잡한 문장의 경우, (b)와 같이 원래 문장에서 호응하는 서술어 보다 앞에 각각 생략된 주어 ‘나’를 삽입하여 표현함으로써 (cf.[2]) 원 문장에서 생략된 주어 이후에 나타난 ‘어른’이라는 인물과 주어와 서술어의 의미적 호응이 잘못 이루어지는 것을 피할 수 있게 된다.¹ 수화에서는 일반적으로 처음 표현되는 동사를 생략된 주어의 것으로 판단하나, 새로운 주어가 언급될 때까지의 모든 후속 동사들을 직전에 표현된 주어와 관련된 것으로 간주하기 때문이다[12].

3.1.3 설명적 표현

어휘가 발달한 언어일수록 함축적 단어가 많이 존재한다. 한국 수화의 어휘는 한국어에 비해 매우 부족하기 때문에 한국어 문장에 나타나는 함축적 어휘를 수화로 나타내기 위해서는 가능한 한 해당 어휘를 풀어서 설명하여야 한다. 특히 일부 한자성어와 신조어처럼 별도의 수화 어휘가 없는 개념을 풀어서 설명하는 형식의 표현이 절실히 필요하다.

- (5) (a) [ㅇ+ ㅍ+ ㅅ+ ㅅ+] + ㅁ+ ㅍ+ ㅎ+ ㅅ+]+ ㅅ
 (b) 평소에 준비가 철저하면 후에 근심이 없다

예를 들어 ‘유비무환’과 같은 사자성어의 경우 마땅히 대응되는 수화가 없다고 해서 (5)의 (a)와 같이 지화를 나열할 경우 원래의 의미를 전달하기 어렵기 때문에, 각 한자의 뜻을 풀이한 후 한자어와 한국어의 어순 관계를 고려하여 재배치해야 하거나, 아예 (b)와 같이 국어 사전의 해당 정의 문장(definition sentence)으로 대체(cf.[14])한 다음, 이를 다시 수화 표현으로 바꾸어

¹ 어휘의 삽입 대신 조사 ‘~면’으로 의미를 나타내고자 할 수도 있겠으나, 수화에서의 ‘~면’ 어휘는 ‘만약’이라는 가정의 개념이 강하기 때문에 본 예문에서의 활용은 적절치 않다.

주는 방식이 필요하다.

3.2 독립된 어휘의 수화 표현 생략

한국어 문장에서 독립적으로 나타나 있는 어휘가 수화 표현시 생략되거나 다른 어휘의 일부로 표현되는 경우는 다음과 같다.

3.2.1 형식 형태소의 생략

고립어라는 수화의 특성상 한국어 문장에서 나타난 형식 형태소의 상당 부분이 수화 표현에서 생략된다. 격조사의 경우 주격, 목적격, 보격, 서술격, 호격 조사는 일반적으로 생략되어 어순에 따라 격을 파악하며, 생략시 의미가 불분명해지는 관형격조사와 부사격조사의 일부만이 실제로 사용되고 있다(cf.[8]). 관형격이나 부사격조사의 경우에도 그 의미가 시각적 표현상 다른 어휘 표현의 일부로 포함되어 있을 때에는 의미적 중복을 제거하기 위해 해당 조사를 생략하게 된다.

- (6) (a) 어머니로부터 온 편지
 (b) [어머니 + 오다 + 편지]

일례로 (6)의 (a)는 (b)로 표현될 때 ‘어머니’를 표현한 위치에서 화자 쪽으로 ‘오다’ 수화를 하게 되면서 ‘어머니께서 편지를 보내셨다’는 의미를 가지게 되므로, ‘부터’는 생략된다[2].

3.2.2 통합 형태의 통사 구조

수화의 어휘는 단순히 음성 언어의 단어에 대응하는 것이 아니라 관념을 표시하는 것이기 때문에, 수화에서 명사는 명사의 의미를 지닐 뿐만 아니라 명사와 해당 명사의 가장 일반적인 서술어가 통합된 개념을 가지고 있는 경우가 많다.

- (7) (a) 밥을 먹었다
 (b) [밥 + 끝]

(7)의 (a)는 ‘밥’이라는 수화 표현에 ‘밥을 먹다’의 개념이 포함되어 있다고 보기 때문에 (b)와 같이 ‘밥’ + ‘완료’ 두 개의 수화 어휘 만으로 표현된다.

그리고 시각-운동적으로 처리되는 수화는 계열적(serial) 통사 규칙만을 가지고 있는 음성 언어와 달리

동시적(simultaneous)-계열적(serial) 통사 규칙을 모두 갖는 것으로 추정되고 있는데, 동시적 통사 규칙은 통사 표지(syntactic marker)를 계열적으로 나타내지 않고, 수화 공간을 이용하여 동시적으로 나타내는 원리이다. 이러한 통사 규칙은 수화소가 결합된 수화 어근 기호의 공간적 요소(운동 방향, 기하학적 배열, 평면 장소 등)나 수동의 특징(종료 방법, 긴장도, 속도 등)을 변화시켜 수화문의 통사 표지를 공간에 동시적으로 나타나게 되며, 수화에서 거의 발견되지 않는 기능어나 용언의 어미, 조동사, 품사의 전성 기호 등이 이러한 원리에 따라서 표현되는 것으로 추정되고 있다(cf.[8]). 이러한 표현을 통해 자연언어 문장에서 나타나는 특정 단어를 독립적 어휘로 표현하지 않고 수화 동작을 구성하는 일부만으로 표현하여 동일한 효과를 얻음으로써 정보를 효율적으로 전달하는 것이다[1].

- (8) (a) 버스가 뒤로 가다
(b) [버스(뒤로)]
(9) (a) 맥박을 재다
(b) [맥박]
(10) (a) 관계 없다
(b) [관계(고리를 품)]

(8)의 (a)는 (b)와 같이 ‘버스’ 어휘에서 버스 모양을 하고 있는 손이 뒤쪽 방향으로 운동하는 것만으로 실현된다. 또한 (9)의 (a)에서는 목적어의 어휘가 서술어의 특정한 수화소와 의미적 연결성을 가짐으로 인해 서술어의 독립된 형태가 나타나지 않는 S+O+∅(영형태)² 구조로, (b)만으로 실현되는데, 이는 ‘맥박’ 어휘에서 나타나는 손목에서의 수위(手位)가 ‘맥박을 재다’라는 통사적 구성을 가지고 있기 때문이다(cf.[1,5]). 또한 (10)의 (a)에서는 ‘없다’ 어휘를 따로 표현하는 대신 (b)와 같이 양손 엄지와 검지의 고리로 ‘관계’ 어휘를 표현하고, 그 상태에서 손 고리를 끊는 동작으로써 문장의 의미를 효과적으로 실현한다.

² 영형태(∅)는 목적어와 서술어가 연결되어 있어 수화소가 생략된 형태를 지칭한다.

3.3 수화 표현의 변형

수화에서는 동일한 어휘에서 수화소 등 수화 표현을 위한 구성 요소의 일부만을 변형시킴으로써 문맥에 따른 의미의 변화를 반영할 수 있다.

3.3.1 동사의 방향 변형

수화에서는 화자가 직접 청자에게 무엇을 하도록 하거나 그 반대의 경우에는 직접 그 동작이 상대방에게 전하여지며, 전하고자 하는 의미의 수화 방향이 주동자에게서 사동자로 향한다(cf.[5]). 우선 주동자와 사동자, 또는 능동자와 피동자가 각각 먼저 표현되고³, 전달하고자 하는 의미의 방향을 향해 동사에 해당하는 수화를 표현하는 것이다.

- (11) (a) 민수는 선생님께서 미움을 받았어요
(12) (a) 어머니께서 나에게 청소를 하라고 하셨다

(11)의 (a)에서는 먼저 ‘민수’와 ‘선생님’을 표현하고, ‘선생님’을 표현한 위치에서 ‘민수’를 표현한 위치를 향해 ‘미워하다’ 라는 표현을 하게 된다. (12)의 (a)에서는 원래는 화자를 기준으로 바깥 방향을 향하게 되는 ‘~을 하라고 하다(명령하다)’라는 수화 표현을 지시 받는 대상인 ‘나’, 즉 화자를 향해 안쪽으로 표현하게 된다. 이는 직접적인 어휘에 의해서도 실현 가능하나, 주체와 객체 간에 이루어지는 수향과 수동의 변형에 의해 실현되는 것이 더 정확한 실현 방법이다(cf.[1,5]).

3.3.2 부사의 영향으로 인한 용언의 변형

부사는 일반적으로 문장의 중심 역할을 하는 서술어를 수식 한정하기 때문에, 문장의 이해와 표현에 있어서 그 통사·의미론적 특성으로 인한 적절성과 다양성이라는 효율적 가치를 지니게 된다. 따라서 부사의 특성을 고려하지 않고서는 용언의 문법적 관계나 의미 관계를 효율적으로 파악하기가 매우 어렵다[3]. 부사 가운데 특히 통사론적 제약 부사의 일종인 양태 부사, 정도 부사는 어휘로도 실현되지만 시각적 언어라는 수화의

³ 사동자 혹은 피동자가 화자일 경우에는 수향에 의해 의미가 전달되므로 따로 먼저 표현하지 않아도 된다.

특징을 심분 활용하여 부사의 의미에 따라 수식하고 있는 용언의 표현에 직접적인 변화를 일으킴으로써 보다 실감 나는 표현이 가능하다. 양태 부사의 경우 ‘천천히 가다’에서는 수식하는 부사가 없을 때보다 ‘가다’ 표현의 운동 속도를 늦추고, ‘빨리 가다’에서는 반대로 속도를 높이는 식으로 표현할 수 있다. ‘바람이 세계 분다’ 같은 문장 또한 ‘바람이 불다’의 수화 표현⁴에서 움직임의 강도를 높이는 방식으로 실현된다(cf.[4]). ‘철수는 느리게 걷고, 순이는 더 느리게 걷는다’라는 문장에서는 처음의 ‘걷다’ 동사는 ‘느리게’라는 양태 부사가 없는 일반적 ‘걷다’ 동사보다 움직임의 속도가 느리게 표현되고, 두 번째 ‘걷다’ 동사는 비교성을 가지는 정도 부사 ‘더’를 반영하여 직전에 표현된 ‘걷다’ 동사에 비해 더욱 느린 속도로 움직이게 된다.

3.4 수화 표현의 이동⁵

수화는 시각적 언어이므로 문장이 길어질수록 청자는 앞쪽에서 표현한 수화 어휘를 망각하거나 의미 관계를 혼동할 가능성이 높다. 때문에 수화에서는 어휘 표현의 순서를 적절하게 이동시킴으로써 청자의 이해를 돕는다.

3.4.1 주어와 서술어의 위치

수화는 시각적 언어이기 때문에 주어가 표현된 다음 긴 수식의 내용이 이어질 경우 주어를 기억하지 못할 가능성이 커진다.

- (13) (a) 딸은 결혼식이 다음주로 다가오자 분주해졌습니다.

(13)의 (a)에서는 주어인 ‘딸’의 바로 뒤에 ‘분주하다’를 표현하고, 그 후에 분주함의 이유를 표현하는 순으로 주어와 서술어를 가까이에 두고 나타내는 방식을 취함으로써 의미의 혼동을 방지할 수 있다(cf.[2,4]).

3.4.2 정보에 중요도에 따른 이동

수화 문장에는 한국어 문장보다 도치된 표현이 많이

나타난다. 일련의 시각적 수화 표현상에서 내용상으로 강조하고자 하는 부분을 문장의 뒤쪽에 표현함으로써 정작 중요한 정보가 기억에서 지워지는 것을 방지하기 위한 수단으로 볼 수 있다(cf.[2]).

- (14) (a) 내가 가장 늦게 출발했다
(b) 가장 늦게 출발한 건 나이다
(15) (a) 왜 학생들이 싸움을 합니까?
(b) [학생 + 싸움 + 왜?]

이는 한국어의 (14)의 (a)에서 ‘나’를 강조하고 싶을 경우 문장을 도치시켜 (b)와 같이 표현하는 것과 유사하다. 이런 방식의 어순 변동은 상대방에게 질문을 하는 경우에도 쉽게 적용된다. 일반적인 의문문에서 가장 중요한 정보는 ‘어떤 것을 묻고 있는가’ 이므로 수화에서는 의문사를 문장의 맨 뒤에 위치시키는 경우가 많다. (15)의 (a)를 (b)와 같이 표현하는 것이 일례다.

3.4.3 사건의 발생 순서에 따른 이동

수화 문장에서 시점이 다른 두 가지 이상의 사건이 언급되는 경우에는 되도록 사건의 발생 순서에 따라 재배열하여 표현하는 것이 오해의 소지를 없애는 방법이다.

- (16) (a) 3호선으로 바꿔 타는 곳은 어디입니까?
(b) [지하철+3+호+바꾸다+어디+입니까?]
(c) [지하철+바꾸다+3+호+어디+입니까?]

지하철에서 (16)의 (a)를 표현할 때 시간 순서에 따른 어순의 재배열 없이 (b)처럼 표현한다면 ‘3호선에서 바꿔 타면 어디입니까?’라는 의미로 잘못 전달될 가능성도 있다. 이런 경우 ‘지하철을 바꿔 타고 나면’ 3호선이 되는 것이므로, (c)와 같이 사건이 발생하는 순서대로 ‘3호’라는 표현 이후에 ‘바꾸다’라는 어휘가 오도록 수화 표현 순서를 이동시켜 준다.

4. 처리 및 구현

본 절에서는 결합범주문법을 이용하여 앞 절에서 정리한 수화 표현의 재현·생략·변형·이동 현상을 적용한 한국어 문장의 처리 방법을 제안하고 실행 결과를 보인다.

4.1 결합범주문법

⁴ 본 논문의 3-2-2절에서 언급된 바와 같이 ‘바람’ 명사가 해당하는 가장 일반적인 서술어 ‘불다’와 통합된 구조이다.

⁵ 여기서 이동이란 수화의 표현 순서, 즉 문장 어순상의 이동을 의미한다.

결합범주문법(Combinatory Categorical Grammar, CCG)은 범주문법(Categorial Grammar, CG)에 결합자(combinatory)가 접목된 것으로, 소수의 축약 규칙에 의해 구문 분석이 이루어지는 문법 체계로서, 병렬 구조 등의 복잡한 문형을 별도의 제약 조건 없이 처리할 수 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[10]. 하지만 결합범주문법은 구문 분석에서 나타나는 가능한 모두 후보를 결과로 생성해 주기 때문에 과생성의 문제가 나타날 수 있으며 CCG 어휘 사전의 자동 확장이 까다롭다.

4.2 어휘에 할당되는 범주

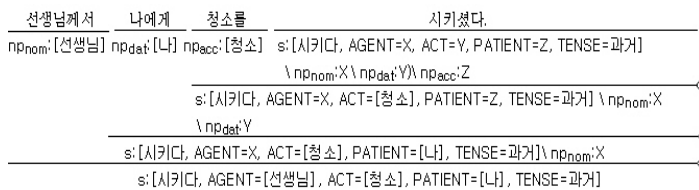
본 연구에서 각 어휘에 할당되는 범주는 다음과 같이 구성되어 있다.

np : (CASE, ACTION/STATE) : [NAME]
 s : (ACTION/STATE) : [NAME,AGENT,ACT,PATIENT,TENSE]

다음은 예문 ‘선생님께서 나에게 청소를 시키셨다’에 해당하는 CCG 어휘 항목 사전이다.⁶

lex(‘선생님께서’, np(nom) : [선생님]).
 lex(‘나에게’, np(dat) : [나]).
 lex(‘청소를’, np(acc) : [청소]).
 lex(‘시키셨다’, s : [시키다, AGENT=X, ACT=Y, PATIENT=Z, TENSE=과거] \ np(nom):X \ np(dat):Y \ np(acc) : Z).

[그림 1]은 예문을 구문 분석한 결과이다.



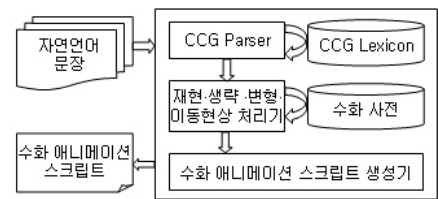
[그림 1] 예문 분석 결과

4.3 시스템 구조 및 애니메이션 스크립트 생성

전체적인 시스템 구조는 [그림 2]와 같다. 텍스트 형태의 한국어 문장이 시스템에 입력되면 먼저 CCG

⁶ 여기에서는 어휘 항목을 어절단위로 하였으나 형태소 단위로 어휘 항목을 다루는 방안에 대한 연구가 진행 중이다.

Parser에 의해 통사·의미적 구문 분석이 이루어지고, 그 결과는 재현·생략·변형·이동 현상 처리기에 입력되어 분석된 문장의 구조와 의미를 적용하여 단어별로 재현·생략·변형·이동의 과정을 거친 후, 수화 사전에서 각 수화 단어의 기본 수화소 정보를 가져온다. 이는 다시 수화 애니메이션 스크립트 생성기에 입력되어 수화 애니메이션 스크립트를 생성하게 된다.



[그림 2] 시스템 구조

[그림 3]과 [그림 5]는 각각 재현·생략·변형·이동 처리기의 출력으로 나타나는 중간 처리 스크립트와 수화 애니메이션 스크립트 예시이다.

중간 처리 스크립트의 첫 번째 요소는 기본 단어와 표현하는 손(왼손 또는 오른손, 양손 수화일 경우 우세손⁷이 있으면 표시하고 없으면 neutral을 표시함)을 나타내고, 두 번째 요소는 수화 표현의 방향성(방향성 있는 수화의 경우 동작의 시작점과 끝점을 표시하고, 없으면 unidirectional을 표시함)을 나타내며, 세 번째 요소는 명사형 수화 단어의 위치 지정 표시이다. [그림 3]이 의미하는 것은 양손으로 ‘선생님’ 수화를 표현(STEP 1)하면서 위치를 지정하고, 오른손으로 ‘명령하다’를 표현하되 방향 요소에 따라 직전에 ‘선생님’이 표현되었던 위치 3에서 디폴트로 화자 자신을 나타내는 위치 1을 향해 ‘시키다’ 수화를 표현(STEP 2)한 다음, 양손으로 ‘청소’의 기본 수화를 표현(STEP 3)하라는 것이다.⁸

- STEP 1 hand(teacher-neutral, unidirectional, setloc-3)
- STEP 2 hand(command-right, 3-1, null)
- STEP 3 hand(clean-neutral, unidirectional, null)

[그림 3] 중간 처리 스크립트

⁷ 한 손만 움직이는 양손 수화의 경우, 정지한 손을 ‘비우세손’, 움직이는 손을 ‘우세손’이라 한다[4].

⁸ 서술어인 ‘시키다’가 주어 쪽으로 이동하고, 수화의 운동 방향 변화로 인해 간접 목적어 ‘나’의 의미가 실현되었다.

hand(NUM, RIGHT, LEFT, RLOC, LLOC, SETLOC)
 RIGHT or LEFT = NAME(EXP|EXP*)
 EXP = (SHAPE, POSITION, DIRECTION, MOTION)
 NUM = 해당 수화 표현을 위해 필요한 손의 수
 NAME = 수화 표현의 명칭
 SHAPE = 수형 POSITION = 수위
 DIRECTION = 수향 MOTION = 수동
 RLOC or LLOC or SETLOC = [HORIZONTAL, VERTICAL, FRONTREAR]
 HORIZONTAL = [left | mid | right]
 VERTICAL = [low | mid | high]
 FRONTREAR = [front | mid | rear]

[그림 4] 애니메이션 스크립트 구성 요소

수화 애니메이션 스크립트에는 위 처리기의 결과물을 바탕으로 기본 수화 단어에 변화를 적용한다.⁹

```
STEP 1 hand(two, [[sg2, p19a, dbbl, m0333]], [[sa, p1, drfb,
mn]], null, null, [mid, mid, mid])
STEP 2 hand(one, null, [[sg, p5, [mid, mid, mid], m20]], null,
null, null)
STEP 3 hand(two, [[sbl, p30, dblr, m45],[sa, p11, dblf, m155]],
[[sbl, p20, dulf, mn],[sa, p11, dblf, m155]], null, null,
null)
```

[그림 5] 수화 애니메이션 스크립트

5. 결론 및 향후 계획

수화는 청각장애인의 모국어로서 시각적 언어와 고립어라는 특징을 가지는 하나의 독립적 언어 체계이며 때문에 한국어 문장을 수화 문장으로 변환하기 위해서는 번역의 관점에서 두 언어간의 차이점을 분석하여야 한다.

본 연구에서는 사용자로부터 텍스트 형태의 한국어 문장을 입력 받아 수화 애니메이션을 자동으로 생성하는 시스템에 사용하기 위한 한국어 문장의 분석 과정을 통하여 한국어를 수화로 변환하고자 할 때 나타나는 언어 현상을 재현·생략·변형·이동의 네 가지 관점에서 정리하였고, 결합범주문법을 사용하여 이러한 현상을 반영한 수화 문장의 처리 방법에 대해 논의하였다. 구현된 시스템에서는 아직 비교적 간단한 문장의 처리만을 다루고 있으나 복잡한 구조를 가지고 있는 문장의 현상도 제안하는 방법으로 처리가 가능할 것으로 보인다. 그러나 이렇게 처리된 수화 표현은 문장의 구조와 의미에 따라 수화소에 변화를 주기 위해 실시간 애니메이션

을 생성해야 하기 때문에 시간 복잡도의 문제가 커질 수 있다. 이의 해결을 위해서는 아바타에서 실제 동작이 일어나지 않는 부분은 2D 정보만을 갖도록 하고, 움직임이 있는 3D 부분은 각 부분별로 따로 저장되 기본적인 단어에 대해서는 미리 정의된 애니메이션을 내보내고, 수화소에 변화를 주어야 하는 경우에만 해당 요소의 조합만을 고려하여 실시간 애니메이션 되도록 함으로써 상당 부분 완화될 수 있을 것으로 생각한다.

앞으로는 본 논문에서 제시한 이론을 기반으로 애니메이션 생성 모듈과의 연동 작업을 통해 현재 부분적으로 구현되어 있는 시스템을 전체적으로 완성할 예정이다. 또한 본 연구에서는 논의하지 않았으나 수화 표현의 주요한 구성 요소인 비수지 신호에 관한 처리 부분이 추가되면 보다 효과적으로 자연스러운 수화를 표현할 수 있을 것으로 기대된다.

6. 참고문헌

- [1] 최지원, 박종철. 결합범주문법을 이용한 수화 자동 생성, HCI, 2003.
- [2] 박인선 外. 수화 업그레이드, 이한출판사, 2000.
- [3] 손남익, 국어 부사 연구, 도서출판 박이정, 1995.
- [4] 석동일. 한국수화의 언어학적 분석, 대구대학교 박사학위논문, 1989.
- [5] 엄미숙. 한국수화의 통사론적 특징 분석, 대구대학교 석사학위논문, 1996.
- [6] 이용원, 최창석. 수화 자동 애니메이션을 위한 국어문장을 수화문장으로서의 변환 방법, HCI, 2002.
- [7] 변지원. 자연 수화의 문형 특성에 관한 연구, 단국대학교 석사학위논문, 1992.
- [8] 황도순. 농학생 수화의 통사 구조와 통사론적 특징, 특수교육논집, 1994.
- [9] 박수현, 강석훈, 권혁철. 한국어-수화 번역 시스템을 위한 형태소 변환, 한국정보처리학회 논문지, 1998.
- [10] Mark Steedman. The Syntactic Process, The MIT Press, 2000.
- [11] H. Davis & Silverson, Hearing & Deafness, N.Y. Holt Rinehart and Winston, 1978.
- [12] L.A. Friedman. The manifestation of subject, object, and topic in American sign language. In C. N. Li, (Eds.) Word Order and Word Order Change, Austin : Univ. of Texas Press, 1976.
- [13] Liwei Zhao *et al.* A Machine Translation System from English to American Sign Language, Proceedings of the Association for Machine Translation in the Americas, 2000.
- [14] Masaaki Tokuda & Manabu Okumura, Towards Automatic Translation from Japanese into Japanese Sign Language, 1998

⁹ 본 연구에서는 [4]에서 제시한 수화소를 기호만 재정의하여 사용하였다.