

언어 장애인의 언어보조 시스템을 위한 아이콘 언어의 구현

추 교 남[†] · 우 요 섭^{††} · 민 흥 기^{†††}

요 약

언어 장애인에게 기존의 키보드에 의한 문자입력 방법보다 원활하고 편리한 의사전달 환경을 제공하기 위한 아이콘 언어 인터페이스를 설계한다. 이를 위하여 활용성이 높은 대화 영역으로부터 구축한 원시 말뭉치를 대상으로 어휘 구사 경향과 특성을 분석하고 형태소, 구문, 의미 분석을 적용하여 아이콘에 부여되는 한국어의 어휘와 의미를 추출한다. 사용자가 직관적으로 인지하고 전달할 수 있는 아이콘 영역을 선별하고 추출한 한국어의 어휘와 의미를 정합시킨다. 이웃하는 아이콘간의 연결로부터 전달하고자 하는 의미적 상황을 만들어내기 위하여, 아이콘 언어의 어휘와 품사, 문법 규칙, 의미체계를 정의하여 아이콘 언어를 설계한다. 아이콘 언어에서 나타날 수 있는 언어적 애매성을 해결하기 위한 방법으로 범용의 한국어 의미사전과 술어 중심의 하위범주화사전으로부터 아이콘 언어에 대한 상황중심의 의미 데이터를 구축한다. 이를 바탕으로 아이콘 언어 인터페이스로부터 한국어를 의미적인 범주에서 생성한다.

키워드 : HCI, 자연언어처리, 의용공학

Implementation of Iconic Language for the Language Support System of the Language Disorders

Kyo-nam Choo[†] · Yo-seob Woo^{††} · Hong-ki Min^{†††}

ABSTRACT

The iconic language interface is designed to provide more convenient environments for communication to the target system than the keyboard-based interface. For this work, tendencies and features of vocabulary are analyzed in conversation corpora constructed from the corresponding domains with high degree of utilization, and the meaning and vocabulary system of iconic language are constructed through application of natural language processing methodologies such as morphological, syntactic and semantic analyses. The part of speech and grammatical rules of iconic language are defined in order to make the situation corresponding the icon to the vocabulary and meaning of the Korean language and to communicate through icon sequence. For linguistic ambiguity resolution which may occur in the iconic language and for effective semantic processing, semantic data focused on situation of the iconic language are constructed from the general purpose Korean semantic dictionary and subcategorization dictionary. Based on them, the Korean language generation from the iconic interface in semantic domain is suggested.

Key Words : HCI, Natural Language Processing, Medical Engineering

1. 서 론

본 논문에서는 언어 장애인이나 일시적 언어 장애가 있는 사람들이 간단한 조작만으로도 한국어 문장을 생성하여 의사소통을 편리하게 할 수 있는 통신보조 시스템을 구현하고자 한다. 언어장애인의 경우, 의사소통은 주로 수화를 이용하는데 수화는 일반인들과의 대화에서 정확한 의미를 전달

하기 어렵기 때문에 다른 의사소통 방법이 필요하다. 그러나 이러한 언어 보조시스템의 경우 주로 키보드에 의한 문자입력 방법에 의존하는 실정이라서 키보드 입력이 어려운 장애인들에게 오히려 정보 소외의 문제를 안겨주고 있다. 또한, 현재 음성인식 등의 기술을 활용하는 시스템 개발도 많은 주목을 받고 있지만, 수화 등을 통해서만 대화가 가능한 청각 장애인들에게는 역시 전혀 도움을 기대할 바가 되지 못한다. 이에 본 논문에서는 자연언어의 특성을 반영하면서 언어 장애인에게 편리한 의사소통 환경을 제공할 아이콘 기반의 언어를 구현하였다.

아이콘 언어는 쉽게 이해할 수 있도록 구성된 아이콘을 선택함으로써 전달하고자 하는 상황을 구성하여 시스템에 입

※ 본 연구는 산업자원부, 한국산업기술평가원 지정 인천대학교 멀티미디어 연구센터의 지원에 의한 것입니다.

† 준 회원: 인천대학교 대학원 정보통신공학과 박사과정

†† 정 회원: 인천대학교 정보통신공학과 교수

††† 종신회원: 인천대학교 정보통신공학과 교수
논문집수: 2006년 1월 19일, 심사완료: 2006년 6월 13일

력하는 방법으로, 선택된 아이콘 집합을 자연언어의 속성으로 재구성하여 그 의미를 파악하고 처리할 수 있는 환경을 제공한다. 기존의 연구에서는 언어 장애인을 위한 통신 보조 시스템에서 심볼을 이용하는 인터페이스를 사용하였으나 [1-3] 각 심볼에 단순히 어휘나 문장이 연결된 형태였고 언어처리를 도입하여 어휘 예측을 시도한 사례가 있었으나 [4-6] 다양한 자연언어의 속성을 처리하기에는 많은 한계를 보였다. 특히 국내에서는 전자공학 및 의공학 분야에서 한국어와 수화간 번역을 시도하는[7, 8] 등의 연구사례가 일부 발표되고 있지만, 그 수가 미비하여 언어 장애인을 위한 시스템 응용과 구현에 많은 어려움이 되고 있다.

이에 본 논문에서 구현하고자 하는 아이콘 언어는 심볼만을 선택하여 정적인 문장만을 생성하였던 기존의 연구와는 다르게 하나의 아이콘에 다양한 품사와 어휘 정보를 저장하여 자연언어와 같이 중의성을 갖도록 구성하였다. 이러한 특징은 하나의 아이콘의 쓰여지는 상황과 이웃하는 아이콘 열에 따라 다른 어휘 특성을 나타내게 되어 동적인 다양한 문장을 생성이 가능하도록 한다. 본 논문에서는 아이콘 언어를 구현하기 위하여 아이콘에 대해 어휘와 문법체계를 정의하고 자연언어 체계의 개념을 적용하였기 때문에 하나의 독립적인 의미기반 언어체계를 갖추며, 이로 인하여 제한적인 아이콘으로부터 다양한 한국어 표현을 생성하는 것이 가능하다. 또한, 언어 장애인의 언어 활동의 특성을 고려하여 적용 영역을 선정하고 어휘를 수집, 분석하였기 때문에 언어 장애인에게 범용적인 응용이 가능하도록 하였다.

2. 아이콘 언어의 설계 및 어휘 데이터의 구축

기존의 언어보조시스템은 심볼에 하나의 어휘만을 할당하여, 구성된 심볼 열로부터 정형적인 문장만을 생성하였다. 이러한 시스템 구조에서는 다양한 언어표현을 위해 많은 심볼을 필요로 하며, 생성되는 문장은 심볼 수에 비하여 그리 많이 없다. 또한 이러한 단일 어휘기반의 언어보조시스템은 문장을 생성하기 위하여 정확히 심볼 열을 구성해야만 했다. 이러한 인터페이스 방식은 언어 장애인에게는 많은 심볼의 의미와 위치를 파악해야 한다는 부담감을 주고, 문장생성은 활용성이 낮다는 점에서 단점으로 지적되어 왔다.

이에 본 논문에서는 자연언어의 중의성을 아이콘에 적용하여 설계함으로써, 아이콘 열이 이루는 정보에 따라 다양한 표현이 가능하도록 하고자 한다. 언어의 중의성이란 문장을 구성하는 구조, 어휘의 의미와 품사 정보가 다양하게 나타날 수 있는 특성을 말한다. 따라서, 본 논문에서 이러한 언어적 중의성을 나타낼 수 있는 아이콘 체계를 아이콘 언어라 정의한다. 아이콘 언어는 사용자가 인지 가능한 의미를 나타내는 그림에 어휘와 품사 그리고 의미 속성이 부여되어 있어, 사용자가 선택한 여러 아이콘 열로부터 전달하고자 하는 상황정보를 재구성하여 다양한 한국어 문장을 생성할 수 있다. 또한, 아이콘 언어는 아이콘으로 명명된 그림에 중의성을 가지고 있는 다양한 언어정보를 가지고 있어, 사

용자는 정확히 아이콘을 조합하지 않아도 전달하고자 하는 적절한 상황을 구성할 수 있다.

그리고 각 아이콘은 외부적으로 사용자에게 일관되고 복잡하지 않은 인터페이스를 제공하고 내부적으로는 다양한 자연언어의 표현을 위하여 여러 품사와 어휘, 의미체계 등 언어 특질의 추가가 용이하다. 아이콘 언어는 장애인의 복지 환경을 위해 연구되고 있는 음성인식과 화상인식 시스템 등과 비교하여 처리 부하가 거의 나타나지 않고 사용자가 직관적으로 인식할 수 있는 아이콘으로 표현되므로 언어 장애인이 전달하고자 하는 정보를 정확하고 빠르게 받아들일 수 있는 장점을 갖고 있다.

2.1 아이콘 언어의 설계 과정

의사전달 수단으로서 아이콘 언어를 활용하기 위해서는 대상 영역의 어휘를 수집하고 분석하여 아이콘에 표현될 어휘 데이터베이스를 구축하는 것과 해당 어휘를 잘 표현해 줄 수 있는 아이콘을 정의하고 아이콘과 어휘간의 연관성을 분석하는 연구가 필수적이다. 한국어 문법의 복잡성은 물론이고, 어휘가 방대하여 품사적인 모호성, 의미적인 중의성을 갖는다는 점과 비교하여 아이콘 언어는 그림으로 구성되는 아이콘을 사용한다. 그러므로 대상 영역이 한정되는 효과가 있어 사용자의 연령이나 장애 정도, 장소나 화제 등 상황에 따라 서로 다른 영역으로 구분이 가능하고 이에 사용되는 어휘의 범위를 한정할 수 있어 언어보조통신에 적합한 모델링 언어라 할 수 있다. 본 논문에서는 다음과 같은 과정으로 아이콘 언어를 설계하였다.

(1) 아이콘 언어의 설계 과정

- 아이콘 언어에 적용될 어휘의 품사 표지와 문말 정보의 정의
- 적용 범주 선정 및 원시 말뭉치 구축
- 아이콘 품사가 부여된 분석 말뭉치 구축: 아이콘 품사가 부여된 말뭉치 구축을 통한 담화 내용의 분석으로부터 아이콘에 부여될 어휘와 품사, 의미적 특성을 추출한다.
- 아이콘과 어휘 및 의미 연결
- 의미사전 구축: 한국어 생성을 위하여 명사 아이콘 어휘를 개념적으로 구분한 시소러스를 구축한다. 그리고 용언 아이콘이 지배하는 명사 아이콘 정보를 나타내는 하위범주화 사전을 구축한다.

2.2 아이콘 언어를 위한 문법 구조의 정의

2.2.1 아이콘 언어의 품사

형태상의 분류로 볼 때 한국어는 의미를 나타내는 실질형태소에 조사와 어미 등이 첨가되는 교착어[9]인 반면, 언어 장애인이 주로 사용하는 수화는 문법 형태소의 발달이 거의 없고 각 어휘들이 독립되어 단어간 위치에 따라 문법적 기능을 갖는 고립어와 유사하다[10, 11]. 아이콘 언어는 한국어와 수화 등에서 공통적으로 나타나는 실질형태소에 그림이 연결되어 표현이 되도록 설계하는 것이 중요하다. 본 논문

에서 설계하고자 하는 아이콘 언어는 하나의 아이콘에 다양한 품사와 어휘를 할당하고 조합되는 이웃 아이콘이 무엇인지에 따라 문장의 생성과 의미를 달리 할 수 있도록 적응적인 형태의 제어 언어를 설계하고자 함이다. 이를 위해서는 각 아이콘이 문장내에서 어떤 속성으로 존재하는지를 구분하여야 할 필요가 있다. 이러한 내용을 고려하여 다음과 같은 사항을 기준으로 품사를 정의하였다.

(1) 아이콘 언어의 품사 특징

- 아이콘 언어는 조사 등의 관계언어나 어미, 접사 등의 의존 형태를 필요로 하는 교착과 첨가어적인 성격의 한국어와는 달리 의미를 나타내는 체언, 용언 등의 중심어와 관형사, 부사 등의 수식어 그리고 문형 정보로 구성된다.
- 각 아이콘은 그림의 속성에 따라 하나 이상의 품사를 가질 수 있다.
- 한국어 생성시 아이콘 언어의 각 품사는 한국어의 중심적 의미를 나타내는 어휘의 품사에 대응될 수 있도록 설계하였다.

<표 1>에 아이콘 언어에 정의한 품사와 한국어 품사를 비교하여 나타내었다. 아이콘 언어의 품사 표지 중, 명사에서는 수사를 분리하여 별도의 품사 표지로 정의하였는데, 숫자는 일상 생활에서 가장 많이 사용하는 빈도가 높은 어휘이기 때문이다. 그리고 부정형 품사 표지를 새롭게 추가하였는데 특정 아이콘의 반의어를 나타내고 싶을 때, 해당 아이콘을 갖는 것보다 NOT 기호를 연결하면 쉽게 관련 어휘를 생성하도록 하였다. 이러한 기능을 위해서 부정형 품사 표지는 아이콘에 특정 아이콘의 부정형이라는 정보를 부여할 때 사용된다.

2.2.2 아이콘 언어의 문말 정보

아이콘 언어의 조합으로부터 한국어를 생성할 때, 보다 정확한 의사 표현을 전달하기 위해서는 종결문형, 시제, 양태 등을 고려해 주어야 한다. 아이콘 언어에는 종결, 시제, 존칭 등을 나타내는 어미 등의 문말 정보가 없이 중심적 어휘로만 구성이 되기 때문에 추가적인 아이콘을 정의하여 문말 정보를 부여하도록 할 수 있도록 하였다.

<표 1> 아이콘 언어와 한국어의 품사 관계

아이콘 품사 표지	한국어 품사 표지	설 명	아이콘 어휘의 예
명사(N)	NG	수사를 제외한 명사 그룹	기차표, 가격
	NA	명사가 부사로 전성	N + 처림
수사(NU)	NU	수사	하나, 둘
용언(V)	VG	일반적인 용언(동사, 형용사)	가다, 가깝다
	DV	관형사로 전성	예쁘다 → 예쁜
	AV	부사로 전성	예쁘다 → 예쁘게
관형사(DT)	DT	관형사	한(하나)
부사(AD)	AD	부사	빨리
부정형(F)	NF	명사의 반의어	not 싸다
	VF	용언의 반의어	*비싸다

<표 2>에 아이콘 언어의 문말 정보에 대한 표지를 나타내었다. 정의한 종결 문형 중 청유형과 부탁은 의미적으로 다른 상황을 나타낼 수 있으나 표현적인 차이가 거의 없어 동일한 태그로 표현한다. 또한, 아이콘 언어의 시제와 필요, 가능, 의지 등의 기타 문형 태그를 정의하였다. 각 표에 나타나있는 정보는 별도의 아이콘으로 제작하였다. 이는 사용자가 문장 생성시 각 문형에 맞는 아이콘을 조합하여 나타냄으로서 사용자의 의사 전달 내용을 분명히 하기 위함이다.

2.3 아이콘 품사부어 말뭉치 구축과 어휘 추출

2.3.1 언어 장애의 특성을 고려한 원시 말뭉치의 구축

아이콘 언어의 활용성을 높이기 위하여 사용자가 이해하기 쉬운 직관적인 그림을 정의하고 여기에 언어적 의미를 부여해야 한다. 그러나 아이콘 언어가 일반적인 자연언어와는 달리 언어 장애인의 의사소통에 목적을 두고 설계된 인공언어이기 때문에 문장생성이 가능한 범주를 정의할 필요가 있다. 그리고 이 범주에서 쓰임새가 높은 문장을 수집하고 특징을 추출해야 한다. 본 논문에서는 적용 대상 범위를 언어 장애인에게 있어 목적성이 분명한 [교통], [쇼핑], [학교], [가정], [병원]으로 결정하였다. 또한, 각 활용 영역에서 발생할 수 있는 대화 샘플을 작업자의 직관력과 어휘의 빈도를 고려하여 공통적으로 많이 나타나는 문장을 수집하여 원시 말뭉치를 구축하였다. 활용 범주를 대표하는 원시 말뭉치를 구축하기 위해서는 여러 작업자가 모든 정의된 영역에 대하여 상황별 발화 문장을 수집하여 활용도의 균형을 고려하여 구축하였고 언어 장애인에게 적용 가능한지를 파악하기 위하여 청각 장애인으로부터 수집한 수화 어휘와 비교, 분석하였다.

본 논문에서는 원시 말뭉치를 구축하는데 있어, 언어 장애를 유발하는 요인에 따라 언어 장애인의 특성을 분류하고 각 특성에 따른 언어 장애인이 정의한 적용 범주에서 사용하는 대화를 수집하였다. 언어 장애를 유발할 수 있는 특성으로 청각 장애, 신체 장애, 정신 장애, 발달 장애로 나누었다. 신체 장애자가 반드시 언어 장애를 갖는 것은 아니지만 언어 장애인이 두가지 이상의 장애를 가지고 있는 경우가

<표 2> 아이콘 언어의 문말 정보

구 분	문말 정보	표 지
시 제	현재형	TN(Time-Now)
	과거형	TP(Time Past)
	미래형	TF(Time-Future)
종결형	평서형	SD(Declarative Sentence)
	의문형	SQ(Question Sentence)
	청유형(부탁)	SR(Request Sentence)
	명령형	SI(Imperative Sentence)
	부정형	SF(False, Not, Negative)
기타 문형 정보	추측, 가정(maybe, perhaps)	MM(Modal-Maybe)
	필요, 의무(have to, need to)	MH(Modal-Have to)
	가능(can)	MC(Modal-Can)
	의지, 소망, 바램(want to)	MW(Modal-Want to)

많고 신체 장애를 수반하는 경우 아이콘을 효율적으로 선택하여 아이콘 열을 구성하기 위한 인터페이스를 설계하는데 참고하고자 함께 고려하였다. 그리고 언어 장애가 없는 사람들을 통해 같은 범주에서 사용하는 대화를 수집하여 비교하였다. 아이콘 언어를 통해 생성되는 한국어는 언어 장애인이 언어 장애가 없는 사람에게 전달될 가능성이 높기 때문이다. 이러한 여러 가지 특성을 고려하여 각 범주마다 500개씩, 모두 2,500개의 문장으로 이루어진 원시 말뭉치를 구축하였다.

2.3.2 아이콘 품사 태깅과 어휘 추출

아이콘에 부여될 어휘를 추출하기 위하여 원시 말뭉치의 각 문장에 대하여 상황을 나타내는 예시 문장으로 간주하고 작업자가 적당한 체언, 용언 등의 중심어에 정의된 품사 표지를 부여한 아이콘 품사부여 말뭉치를 구축하였다. 아이콘 품사부여 말뭉치를 구축하는 이유는 아이콘 후보 어휘 추출을 위한 빈도와 핵심 어휘의 범주별 분석 및 축약 어휘를 구성하는 대표 어휘 구성을 파악하기 위함이다. 아이콘 품사부여시 다음과 같은 사항을 기준으로 정의하여 작업하였다.

(1) 아이콘 품사 태깅 기준

- 원시 발화 문장에서 공통적으로 나타나는 효율적인 축약 어휘로서의 역할이 가능한 대표적인 핵심어휘를 추출하고 아이콘으로 표현되면서 나타날 수 있는 의미와 품사적 애매성을 고려하여 다양한 어휘 형태와 대응하는 품사를 부여하였다.
- 발화 문장의 구성 어휘가 대표 단언어로 분리가 가능한 경우 분리된 각 단어에 품사를 부여하였다.
예) 여행객 = i여행/N + i사람/N, 고속버스 = i빠르다/V + i버스/N(* i: 아이콘)
- 수사와 단위명사로 구성된 단어에 대해서는 아이콘으로 표현 가능한 어휘가 제한적이기 때문에 단일 수사로만 표현이 가능하도록 하였다.
예) 한장, 한개 = i1/NU
- 태깅시 주어진 상황 예문에 대해서는 문형정보까지 나타낼 수 있도록 하였다.
- 명사에 '-하다', '-되다' 등의 접사가 붙어 술어적으로 전성되는 형태는 용언에 대한 품사를 부여하였다.
- 발화 문장의 용언 성분에 보조용언이 연결되어 있거나 필요한 성분인 경우 한국어 생성이 참조 정보로 활용하기 위하여 핵심 어휘와 같이 용언 품사를 부여하였다.
예) 놀러 가다 = i놀다/V + i가다/V
- 문형정보의 태깅시 많이 나타나는 평서형(SD), 현재형(TN) 등에 대해서는 기본 태그로 정하고 "어디"와 같은 의문사에 대해서는 태그 부여시 의문형 태그 생략이 가능하도록 하였다.
- 명사나 부사로 전성이 가능한 용언 성분으로 강조의 의미를 나타내기 위해 아이콘 어휘가 두 번 연속으로 나타나는 형태가 가능하도록 하였다.
예) 최고급 = i좋다/V + i좋다/V

- 동사의 경우에는 시제의 구분 없이 원래의 형태를 기재하였다.
예) 했었다 = i하다/V + /TP
- 축약된 말, 즉 '하면'과 같은 경우에는 이를 두 낱말로 분리하여 '하다'와 '그러면' 으로 기재하였다.
- 존칭의 경우에도 평어의 형태로 바꾸어 기재하였다.
예) 저, 제가 = i나/N

<표 3>에 아이콘 품사부여 말뭉치의 예를 보였다. 그리고 한국어 생성시에 문형 패턴을 참고하기 위하여 아이콘 말뭉치 뿐만 아니라 한국어에 대해서도 형태소 분석기와 수작업을 병행하여 품사부여 말뭉치를 구축하였다. 아이콘 품사부여 말뭉치로부터 아이콘에 나타날 어휘 추출시 장소에 구애받지 않고 널리 쓰이며 공통적으로 나타나는 핵심 어휘와 특정 장소에서 자주 사용되는 주변 어휘로 나누어 분류하였다. 이렇게 분류하는 이유는 사용자가 특정장소에서나 또는 자주 발생하는 상황에서 쉽게 이용할 수 있기 때문이다.

그리고 장소에 따라 분류된 어휘는 어휘의 한계가 설정되기 때문에 어휘 예측의 신뢰성을 좀 더 높일 수 있을 것이다. 특히 언어 장애인이 주된 사용자라는 점에서 기존의 한국어 처리 연구와는 다른 각도에서 어휘를 추출하고 분류할 필요가 있다. 즉, 언어 장애인은 대부분 수화 등에 익숙하기 때문에 한글을 배운다고 하더라도 일반인이 작성하는 문장과는 달리 국어와 수화의 어휘를 혼용하여 문장을 작성하는 경향이 있다는 점이다. 다품사어, 다의어의 경우를 포함하여 한국어 어휘의 중의성과 수화의 중의성이 일대일로 대응하지 않는 문제, 어순의 차이 문제 등이 이러한 원인이 된다.

더구나 수화의 경우 몸짓 어휘수가 3000여개 정도로 적고 이를 여러 방법으로 연결, 혼용하여 새로운 어휘를 표현한다는 점[12]을 고려한다면, 아이콘을 이와 같이 수화 측면에서 중의성을 갖도록 연결할 필요가 있다. 이를 분석하기 위해서는 아이콘 품사가 부여된 문장으로부터 다양한 의미와 어휘를 추출해야만 한다.

<표 3> 아이콘 품사 부여 말뭉치의 예

원시 문장	품사 부여된 아이콘 열
첫 기차는 몇시에 있습니까?	i시작/N i기차/N i시간/N i연제/N /SQ
최고급 호텔로 가주세요.	i좋다/V i호텔/N i가다/V /SR
	i가장/AD i좋다/V i호텔/N i가다/V /SR i좋다/V i좋다/V i호텔/N i가다/V /SR
우등 버스의 요금은 얼마입니까?	i좋다/V i버스/N i얼마/N /SQ
	i좋다/V i버스/N i값/N /SQ
종착역까지는 얼마나 남았습니까?	i끝/N 장소/N i얼마/N i남다/V /SQ
	i끝/N 장소/N i시간/N i얼마/N /SQ
	i끝/N 장소/N i연제/N i도착하다/V /SQ
전전히 가주세요.	i전전히/AD i가다/V /SR
	i늦다/V i가다/V /SR



(그림 1) 아이콘 어휘의 예

2.4 어휘에 맞는 아이콘의 정의

언어 장애인은 인터페이스에 나타난 아이콘의 그림을 선택함으로써 시스템에 자신의 의사를 전달하게 되므로 얼마나 확실하고 명확한 그림을 제시할 수 있는지가 중요한 문제라고 할 수 있다. 아이콘이 내포된 어휘와 의미를 얼마나 잘 전달하느냐가 원활한 한국어의 생성에 중요한 요소이기 때문이다. 자연언어의 어휘가 갖는 중의성과 다의성을 고려하고 직관적인 표현과 효율적인 아이콘의 배열이라는 사용자 인터페이스 측면의 분석을 통해 대상 영역 별로 아이콘을 정의하였다.

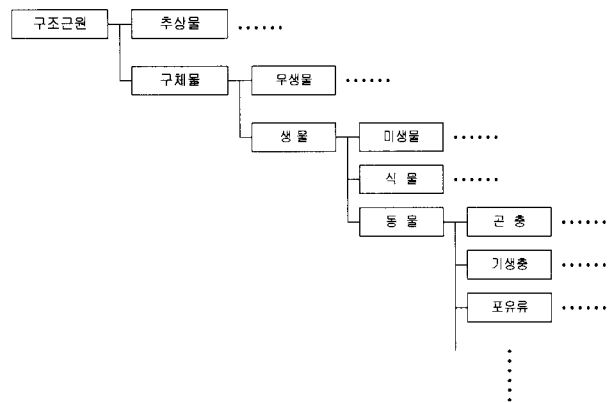
여기에는 앞서 기술한 바와 같이 한국어와 수화, 모두의 고려가 반드시 필요하다고 할 수 있다. 아이콘에 여러 어휘와 각 품사를 다중으로 연결하였으나 동일 품사에 대해서는 중복됨이 없이 하나의 어휘와 의미만을 할당하였다. (그림 1)에 구축된 아이콘 어휘의 예를 보였다. 각 어휘에 대한 그림으로 클립아트를 활용하며 위에서 추출한 아이콘 후보 어휘와 그림을 연결하는 일은 자동화하기 어렵기 때문에 여러 작업자가 많은 논의와 튜닝 작업을 통하여 수행하였다. 예로 “눈”이라는 일반 명사는 적절한 그림에 그 어휘의 의미 그리고 품사를 할당하기 용이하나 “아름답다” 등의 느낌을 전달하는 아이콘은 결정하기 어려운 내용이었다. 이에 여러 공동 작업자의 일치된 의견을 통해 정의하였다.

2.5 아이콘 언어의 의미사전 구축

아이콘 언어 체계는 사용자가 입력한 여러 아이콘간 관계를 분석하고 재구성하여 한국어 생성에 적합한 대응 모델을 도출할 수 있어야 한다. 이를 위해서 명사 아이콘의 의미 관계를 계층화하여 나타난 시소러스와 아이콘간 의존관계 분석 및 용언 아이콘에 필수적으로 의존되는 명사 아이콘의 의미적 관계를 나타내는 하위범주화사전, 그리고 아이콘 어휘의 관용어 사진을 구축하였다.

2.5.1 명사 아이콘 시소러스

명사 아이콘 시소러스를 구축하는 일은 상당한 노력과 경험을 필요로 하고 대부분을 수작업에 의존해야 하는 문제가 있다. 이러한 한계를 극복하면서 빠르고 정확한 구축을 위하여 기구축된 범용적인 한국어 명사 시소러스[13, 14]를 이용하였다. 한국어 명사 시소러스의 상의어 계층을 적용 범주에



(그림 2) 명사 아이콘 시소러스

맞도록 재정의하고 아이콘 품사부어 말뭉치에서 추출한 명사 어휘를 등록하였다. 본 논문에서 구축한 명사 아이콘 시소러스는 각 어휘의 속성을 큰 범주로 나누어 포함시킴으로써 시스템 적용에 용이하도록 하였으며 아이콘이 나타내는 다양한 의미를 동적으로 구성할 수 있도록 설계되었다. 계층을 너무 깊고 세부적으로 나누게 되면 실제 시스템 적용에서 어휘의 각 의미 속성을 결정하는 데 오히려 많은 중의성을 유발시키는 문제점이 나타나게 된다. (그림 2)에 명사 아이콘 시소러스의 상위계층 정보를 나타내었다.

2.5.2 용언 아이콘의 하위범주화사전

아이콘 언어를 한국어로 번역하는 과정에서 가장 먼저 분석하는 정보로는 용언의 역할을 수행하는 아이콘을 분석하는 일이다. 이는 자연언어 분석 이론 중에서, 모든 문장은 용언을 중심으로 상황이 구성되며, 더욱 자세한 정보를 나타내기 위해 각 명사 성분들이 보어으로써 의존한다는 용언 중심론에 근거를 둔 것이다. 하위범주화란 보어의 어휘 개념이 명사된 술어와 보어간 의존관계를 정의하는 중요한 언어 정보로서 구문 및 의미 분석 등에 폭넓게 활용되는 언어 자원이다. 표층문에서 통상 격표지로 표현되는 구문적 의존관계뿐만 아니라 보어가 갖는 의미 정보가 부착되어 있어, 시소러스의 의미 표지와 연동이 가능하도록 설계하였다[15]. 본 논문에서도 용언 아이콘에 대하여 하위범주화의 개념을 적용하여 한국어 생성을 위한 의미적 중심구조 분석에 활용하고자 하였다. 아이콘 말뭉치로부터 술어 아이콘과 이에 의존하는 명사 아이콘의 어휘의 의미적 관계를 기구축된 하위범주화사전[13]의 정보로 정합하고 비교 분석하여 구축하는 시간과 노력을 줄일 수 있었다.

용언 아이콘의 하위범주화사전에서 중요하게 설계된 점은 용언 아이콘에 의존되는 명사 아이콘의 어휘에 대하여 한국어 생성시에 필요한 조사를 연결할 수 있도록 구성된 문형 패턴이다. 아이콘 언어는 실질형태소로만 구성되어 있기 때문에 이 중에서 명사 아이콘 어휘에 대해서는 한국어 생성시 적절한 조사를 복원해주어야 한다. 한국어 문장의 기본 구조를 파악하기 위해 위에서 조사를 바탕으로 형성되는 문형 표준 패턴 64개(동사:47개, 형용사:17개)를 정하였다.

〈표 4〉 표준 문형패턴의 예

구분	문형 패턴
1	N1 이/가
2	N1 이/가 + N2 에/에서
3	N1 이/가 + N2(때문)에/((으)로
4	N1 이/가 + N2 을/를 + N3 에/에게
5	N1 이/가 + N2 을/를 + N3(으)로
6	N1 이/가 + N2 에/에게 + N3 에(대하여)

〈표 5〉 용언의 하위범주화 사전의 예

아이콘 용언	문 형 정 보
가다	[개념] 1 사람 + 2 장소 [문형] 1 이/가 + 2 을/를
사다	[개념] 1 사람 + 2 물건 + 3 장소 [문형] 1 이/가 + 2 을/를 + 3 에서

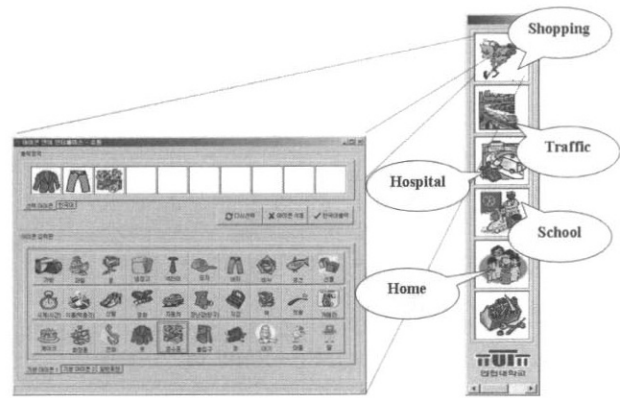
〈표 6〉 아이콘 어휘에 대한 관용어 사전의 종류

아이콘 어휘의 관용어 사전	사전 구성의 예
명사 아이콘간 복합어 사전	i남자/V + 옷/N = 남성복/N
용언과 명사 아이콘의 복합어 사전	i올라가다/V + 장소/N = 승강장/N
명사 아이콘의 어휘 복원 사전	i돼지/N = 돼지고기/N(쇼핑영역)

용언 아이콘의 문형패턴은 아이콘 조합의 한국어 번역을 위해서 매우 중요한 문장 생성의 표준화 수단이다. 용언 아이콘에 의존할 수 있는 수많은 명사 아이콘 중에서 필수적인 성분만을 구문적, 의미적으로 미리 지정해 주는 것이기 때문이다. 본 논문에서의 표준 문형패턴은 기존에 세분화된 패턴을 대표할 수 있도록 정의하여, 계산처리가 용이하고 한국어 문장의 생성 시, 다양한 어휘로 자연스러운 표현이 가능하도록 하였다. 또한 문형패턴의 정의는 일관성있는 한국어 생성 측면에서 바람직하다고 생각된다. 문형패턴은 최대 필수성 성분 4개만으로 구성하였는데, 이는 애매성이 해결된 술어 아이콘이 의존적으로 이끄는 명사 아이콘의 하위 범주가 4개 이상을 벗어나지 않는다는 아이디어를 반영한 것이다. <표 4>에는 용언 아이콘의 문형패턴, <표 5>에는 용언 아이콘의 하위범주화 패턴의 예를 나타내었다.

2.5.3 아이콘 어휘의 관용어 사전

앞서 예를 통해 본 것처럼 [고속버스]라는 명사가 아이콘으로 표현되지 않았다면 [빠르다]와 [버스]를 눌러 표현할 수 있다. 이와 같은 방법이 가능한 이유는 아이콘 인터페이스가 목적성이 분명한 영역에서 연상되는 내용을 아이콘으로 조합이 가능하도록 나타내고 있기 때문이다. 한국어 생성시 필요한 어휘사전으로 활용하기 위하여 각 아이콘의 공기 관계로부터 여러 정보를 사진화하여 관리할 필요가 있다. 본 논문에서는 <표 6>와 같이 구축된 아이콘의 어휘간 관용어 사전을 구축하고 활용한다.



(그림 3) 아이콘 언어 인터페이스

3. 아이콘 언어의 구축 및 평가

3.1 아이콘 어휘의 구축

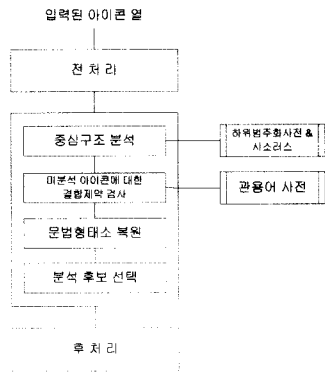
[교통], [쇼핑], [학교], [가정], [병원]에 대하여 활용도가 높은 문장으로 각 500개씩, 2,500여 문장을 원시 말뭉치로 구축하였다. 각 문장으로부터 약 2.3개의 아이콘 후보 어휘열을 얻을 수 있었다. 작업자가 예문으로부터 문장 자체의 분석이 아닌 상황 예시로 간주하여 아이콘으로 구성이 가능한 어휘열을 나열하였기에 한 문장으로부터 여러 아이콘 후보 어휘열을 추출하였다. 각 아이콘 후보 어휘열로부터 약 2.1개의 명사 아이콘 어휘와 1.8개의 술어 아이콘 어휘를 추출하였는데 이처럼 명사가 적게 추출된 이유는 언어 장애인이 사용하는 문장이 단문으로 이루어진 문장 구조가 많았기 때문이다.

이 과정에서 명사 아이콘 어휘는 12,000여개, 용언 아이콘 어휘는 10,300여개를 얻을 수 있었으며, 이 중 명사 아이콘 어휘 600여개, 고빈도 용언 아이콘 어휘 250여개를 약 700여개의 아이콘을 정의하여 정합하고 아이콘 언어를 구축하였다. 아이콘 1개당 2.3개의 명사와 용언에 관한 품사가 부여되었으며 기타 다른 부사, 관형사 등의 품사는 명사 및 용언 속성의 전성을 통해 나타날 수 있도록 하였다.

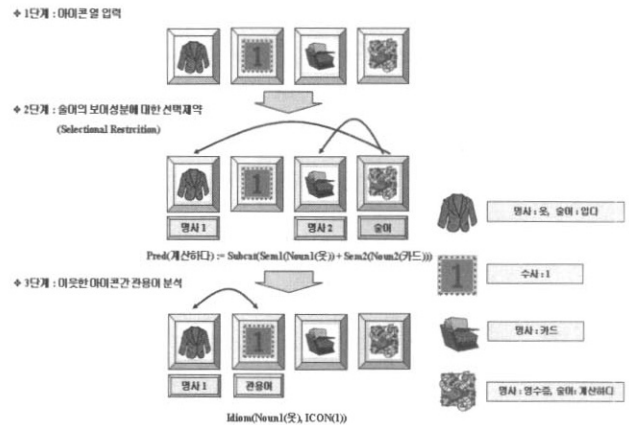
3.2 아이콘 언어 인터페이스

본 논문에서는 언어 장애인이 아이콘 언어를 편리하게 실생활에 활용하기 위한 인터페이스를 (그림 3)과 같이 구현하였다. 아이콘 언어 인터페이스는 (그림 3)의 오른쪽과 같이 사용자가 생성시키고자 하는 범주를 먼저 선택할 수 있도록 하였다. [쇼핑] 선택하면 해당 범주에서 사용되는 아이콘 어휘가 (그림 3)의 왼쪽과 같이 보여지며, 이 중에서 선택된 아이콘은 출력영역에 나열된다. 사용자가 자주 선택하였던 아이콘은 통계적으로 계산되어 선택하기 쉽도록 처음 열에 나타나게 된다. 그러나 사용자의 특성에 따라 고정 위치에서 아이콘을 선택할 수 있도록 옵션을 구성하였다.

선택된 아이콘 열에 대하여 [한국어 출력] 버튼을 누르면 한국어와 음성이 동시에 출력되도록 구성되었다.



(그림 4) 아이콘 언어의 한국어 생성 과정



(그림 5) 아이콘 언어의 한국어 생성 과정의 예

<표 6> 아이콘 간의 구문과 의미적 의존관계

N1+N2	N2 (N1의 지배소)	1 관용어구 파악 2 복합어 파악 3 소유격 파악
	N1 (N1의 지배소)	후위에서 전자로의 의존관계를 고려하지 않는다면 실패
NU1+N1	NU1과 N1의 언어 또는 수식관계 처리	
DT+N1	DT와 N1의 언어 또는 수식관계 처리	
AD+DT	AD와 DT의 활용 또는 수식관계 처리	
AD+V	AD와 V의 활용 또는 수식관계 처리	
V1+V2	V1가 중심어	1 관용어구 파악 2 V2가 보조사의 역할이 가능한지 파악
	V2가 중심어	1 관용어구 파악 2 V1이 절로서 V2에 의존되는지를 파악
V+n	V에 대한 부정어 처리	

3.3 선택 제약을 통한 한국어 문장 생성

사용자가 입력한 아이콘 열에 대하여 먼저 명사 아이콘 시소러스와 하위범주화사전을 기반으로 선택제약(selectional restriction)을 적용하여 술어와 보어 성분간 의미적 중심 구조를 분석하고 선택제약으로 의존되지 못한 기타 아이콘은 이웃한 아이콘간의 관계 분석을 통해 세부적인 어휘 구조를 파악하였다.

한국어 생성을 위해 입력된 아이콘 열은 자연언어처리에서 입력하는 문장과는 다르게 각 아이콘에 다양한 품사와 어휘가 할당되어 있어, 형태소 분석이 이루어진 어휘 집합으로 볼 수 있다. 전처리 단계에서는 아이콘 어휘의 품사 정보를 이용하여 용언(동사, 형용사)으로 판별되는 어휘의 위치 정보를 기록한다. 용언 아이콘의 위치는 용언 아이콘의 앞뒤에서 지배되는 보어 아이콘 성분을 분석하는 기준이 되기 때문이다. 중심구조 분석에서는 용언의 위치 정보가 저장된 배열 공간에서 용언을 선택하고 이를 하위범주화사전에 등록된 용언 정보와 비교하여 해당 용언의 후보패턴을 추출하고 어휘 성분을 중심으로 일치할 수행한다. 즉, 하나의 아이콘 보어 성분에 대하여 시소러스내에 해당 어휘가 등록되어 있는지를 파악한다. 시소러스내에 등록된 단어라면 하위범주화사전의 후보 패턴내에 존재하는 의미와 비교하여 상하위관계가 존재하는 지를 분석한다. 상하위관계가 존재한다면 이 보어성분에 대해서는 어휘의 의미를 분석한 것으로 판별한다.

용언과 보어 성분은 문장을 이루는 중심 구조로써 이를 하위범주화사전, 시소러스와의 비교를 통해 분석하였지만 세부적인 관형어와 복합어는 이웃하는 아이콘간 이루는 어휘적 특성이므로 분석되지 못한 아이콘에 대하여 관용어 사전을 이용하여 결합제약 분석을 수행한다. 각 아이콘이 이루는 결합 특성을 문법적으로 정의하고 이 관계를 통해 관용어 사전을 탐색하여 빠르게 관련 어휘를 나타내도록 하였다. <표 6>에 아이콘간 문법 관계를 정리하였다.

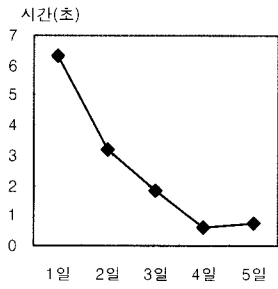
각 아이콘에 대하여 한국어의 실질 어휘 성분이 판별되면 조사와 어미 등의 문법 형태소를 연결하여 기본적인 문장을 생성한다. 이러한 일련의 과정에서 여러 문장 후보들이 생성될 수 있는데, 아이콘 열 중 용언 성분을 할당받은 아이콘어 여러 개 있을 수 있고, 각 용언에 대하여 여러 하위범주패턴이 일치되어 나타날 수 있기 때문이다. 여러 분석 후보 중 하위범주화사전과 가장 많은 일치 성분을 가지고 있는 후보를 선택한다. 후처리에서는 앞서 정의된 문말 표지를 분석하여 시제, 종결형, 기타문형 정보를 선택된 한국어 기본 문장에 적용하여 자연스러운 한국어 생성이 가능하도록 한다. (그림 5)에 아이콘 언어의 한국어 생성 과정의 예를 보였다.

- 외부 인터페이스로부터 (그림 5)과 같이 아이콘 열을 입력받았을 경우를 고려해 본다. 그림에 나타난 아이콘 열은 사용자가 상대방에게 “옷 한 벌을 카드로 계산하고 싶다.” 라는 문장을 생성하기 위해 선택한 것이다.
- 두 번째 아이콘 열을 나타내는 그림에서처럼 술어 속성을 가지고 있는 마지막 아이콘에 대해 하위범주화사전에서 일치하는 한국어 술어 패턴을 찾는다.

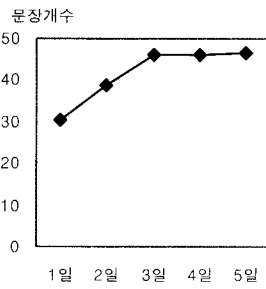
예) 계산하다

→ [사람] AGT {이/가} + [물건] ACC {을/를} + [금전류] INS {로/으로}

- 일치하는 패턴의 정보로부터 앞에 위치한 명사 속성을 아이콘 열과의 의존 관계를 파악하여 대화체 문장의 중심구조를 얻는다. 의존관계를 가지고 있는 명사 속성의 의미가 술어의 하위범주화의 의미에 정합이 되는지를 파악한다. 이때 대화체에서의 주어 생략을 고려할 수 있다.
- 예) 계산하다



(그림 6) 아이콘 조합에 대한 평균 시간



(그림 7) 청각 장애인의 표본 문장에 대한 정확도

→ 옷 [물건] ACC {을/를} + 카드 [금전류] INS {로/으로}

· 중심구조에 대하여 일치 하지 않은 아이콘의 명사 속성에 대해서는 관용어 사전을 참조하여 각 상황에 맞는 아이콘의 속성을 정의한다.

예) 계산하다

→ 한[하나의 수식어] + 옷 [물건] ACC {을/를} + 카드 [금전류] INS {로/으로}

3.4 아이콘 언어의 평가

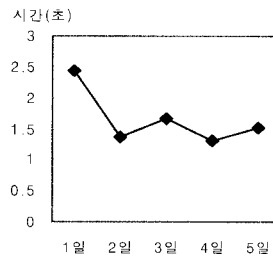
아이콘 언어가 실생활에서 활용이 용이한가를 평가하기 위하여 청각 장애인과 지체 장애인에게 적용하여 검증하였다. 실제 상황에 적용함에 있어 가장 중요한 것은 빠른 시간 내에 정확한 표현의 생성이라 할 수 있다. 아이콘 인터페이스에 대한 훈련을 통해 아이콘 어휘의 위치와 내용을 파악하는 것이 중요하다. 본 논문에서는 활용성 평가를 위하여 다양한 장애 특성을 지닌 장애인에게 아이콘 인터페이스 시스템을 훈련시키고 다양한 환경에서 실전 적용을 통해 검증을 실시하였다. 인천에 소개한 청각장애인협회와 특수 교육 기관인 은광학교의 지체 장애인을 대상으로 상당기간 아이콘 언어 인터페이스를 교육시키고 실전 상황에 적용하여 검증하였다. 장애인의 프로그램 사용 현황을 파악하기 위하여 프로그램에 사용자가 아이콘을 누르는 내용을 기록하는 모니터링 프로그램을 추가하여 결과 파악에 용이하도록 하였다.

3.4.1 평가 조건

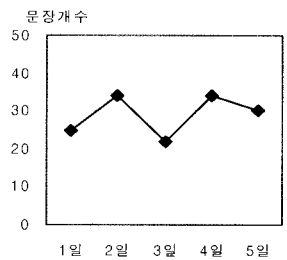
- 목표: 휴대 장치를 사용하는 시스템의 정확도와 문장 생성 시간의 기록과 검증을 통한 활용성 고찰
- 기간: 15일
- 대상: 인천청각장애인협회 소속 청각 장애인 3명(남 1, 여 2) 은광학교(1, 2급 지체 장애인 특수학교) 소속 지체 장애인 3명
- 특이 사항
 - 청각 장애인은 컴퓨터 사용이 대체로 익숙하였으나 지체 장애인은 컴퓨터 사용이 용이하지 않았음.
 - 지체 장애인의 경우 상대방과의 의사 소통이 특히 원활하지 않았음.
- 검증 환경
 - 아이콘 언어 인터페이스가 탑재된 모바일 컴퓨터



(a) 전철역에서 표를 구입하는 모습 (b) 택시에 승차하여 목적지를 알리는 모습
(그림 8) 아이콘 언어의 현장 검증



(그림 9) 아이콘 조합에 대한 평균 시간



(그림 10) 지체 장애인의 표본 문장에 대한 정확도

· 진행 내용

청각 장애인의 경우 프로그램 사용 훈련 기간 동안에는 연구원이 각 장애인에게 개인 지도 훈련을 통해 훈련시켰다. 또한, 문장이 주어진 훈련에서 벗어나 상황을 제시하여 사용자가 연상되는 아이콘은 선택하도록 하였다. 그러나 지체 장애인의 경우에는 의사 소통의 문제가 있어 상황 위주의 검증보다는 문장을 주어 직접 생성하도록 하였다. 실외에서는 아이콘 언어의 영역이 잘 활용되는지를 평가하기 위해 전철이나 버스 등 대중교통, 쇼핑, 식당에서 주문을 하는 등의 정의된 각 영역에 대해 실제 검증을 수행하였다.

3.4.2 아이콘 어휘의 조합 시간

(1) 청각 장애인의 경우

위의 결과는 청각 장애인 3명의 평균 데이터로 표본 상황 50개에 대해 언어 장애인이 조합하여 만들어낸 문장의 의미를 판별하여 결정하였다. (그림 6)은 아이콘 조합을 구성하는데 소요된 시간을 단위일별로 나타내었다. (그림 7)은 청각 장애인이 아이콘을 생성하여 의사를 전달함에 있어 의미를 정확히 조합하였는지를 나타내는 결과이다. 아이콘 조합 시간과는 달리 언어 장애인이 정확한 문장을 만들고자 하는 시간이 길었고 문장의 의미는 거의 정확함을 알 수 있다. (그림 8)에는 아이콘 언어의 현장 검증 모습을 보였다.

(2) 지체 장애인의 경우

(그림 9)는 평가 문장 하나에 대한 평균 아이콘 조합 시간이다. 그리고 지체 장애인 3명의 평균 데이터이다. y축은 분(min)단위 시간이고 x축은 프로그램 훈련 기간이다. 결과를 살펴보면 실행 시간은 크게 차이가 나타나지 않는 것을 알 수 있다. 선행된 검증에서는 초반기에는 조합시간이 많이

걸렸지만 이는 나이가 많은 테스트가 컴퓨터를 익숙하게 다루지 못했던 원인도 있었던 것에 비하여 본 검증에 참가한 지체 장애인들은 컴퓨터 교육을 많이 받은 어린 학생들이었다는데 실행 시간에 차이가 있다. 그러나 훈련 날짜가 지날수록 아이콘의 위치와 도구 사용법에 익숙해져 훈련 마지막 날에는 바로 찾아 조합하는 시간에 도달하였음을 알 수 있다.

(그림 10)을 살펴보면 표본 문장 50개에 대하여 지체 장애인이 조합한 문장의 의미를 판별하여 결정하였다. 그리고 언어 장애인 3명의 평균 데이터이다. y축은 문장의 개수이고 x축은 프로그램 훈련 시간이다. 아이콘 조합 시간과에 대한 한국어 생성의 정확도가 앞선 평가에 비해 현저하게 떨어짐을 알 수 있다. 모바일 도구로 사용한 태블릿 PC는 펜으로 제어할 수 있도록 구성되어 있는데 뇌성 마비 지체 장애자는 손의 떨림 현상이 심하여 아이콘을 선택하기 위한 집중력이 많이 부족하였다. 이로 인하여 같은 아이콘이 여러번 눌러짐으로 인하여 한국어 생성의 정확도를 현저하게 떨어뜨리게 되었다.

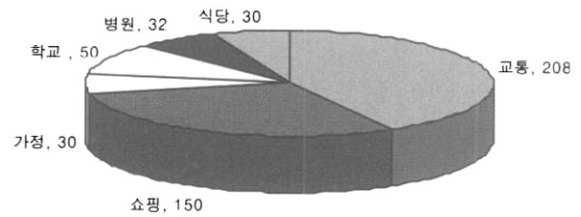
(3) 활용 영역별 분석

(그림 11)은 실험에 참가한 6명의 장애인이 실제로 아이콘 언어 인터페이스를 이용하여 500여개의 상황에 대해 자유롭게 영역별 문장을 생성해 본 결과이다. 위 결과의 단위는 문장 개수이다. 위의 결과로부터 가장 활용도가 높게 나타난 영역은 교통과 쇼핑임을 알 수 있다. 가장 큰 이유로 언어 장애인이 서비스 받고자 하는 목적이 분명하고 많이 이용하는 곳일 수록 높게 나타남을 알 수 있다.

(4) 결과 검토

지하철 등의 교통 상황에서는 대부분 단문의 형식으로 문장을 필요로 하였다. 그러나 목적지의 이름 등 장소가 나타내는 고유명사 등을 모두 어휘 데이터베이스에 등록하는데 한계가 있어 이 부분에 대한 해결이 필요하다. 쇼핑 상황에 있어서는 가장 많은 시간을 할애하였지만 사용자에게 문장의 생성을 맞기다 보니 단조로운 문장들이 생성되었다. 가격을 깎아달라거나 더 큰 사이즈를 보여 달라거나 하는 상황들을 사용자가 생각하지 못하여 “얼마이다”, “필요하다” 등의 동사만을 사용하게 되었다. 프로그램이 한번에 보여줄 수 있는 명사의 개수를 30개로 제한했기 때문에 처음 보여주는 30개 이외의 명사를 잘 찾지 못하는 경우가 발생하였지만 30개 이상의 명사를 보여주더라도 한번에 찾기에는 어려울 것이라 판단된다.

이러한 문제점을 보완하기 위하여 사용자가 많이 사용하는 아이콘 조합은 저장한 후 유사한 조합이 발생할 경우 우선적으로 후보를 보여주는 기법을 추가하였다. 핸드폰 가게에서 핸드폰 가격을 물어볼 때도 이 같은 문제로 “핸드폰” 명사를 찾지 못하여 “물건”이라는 통합적 명사를 사용하였다. 다른 시각에서 본다면 쇼핑에 관련한 수많은 단어들을 모두 데이터베이스에 입력할 수 없으므로 어찌면 “물건”, “물품” 등의 명사를 사용하는 것이 당연할 수 있겠지만 프로그램에서 제시하는 어휘를 빠르게 검색할 수 있는 다양한 방법이 필요할 것으로 생각된다.



(그림 11) 언어 장애인의 영역별 평균 활용도

4. 결론

이상에서 아이콘 언어의 설계 및 구축에 대해 논의하였다. 아이콘 언어는 키보드로부터 문자를 입력받아 제어하는 자연언어 인터페이스보다 직관적이며 사용이 용이한 그림에 어휘 속성이 부여되어 있어 언어 장애인이 아이콘의 조합으로 의사를 쉽고 편리하게 전달할 수 있도록 설계된 언어보조 인터페이스이다. 본 논문에서 제시한 아이콘 언어는 단순히 그림을 정의하고 어휘만 부여했던 기존의 연구와는 달리 아이콘이라는 그림과 어휘가 조합된 단위를 정의하고 여기에 언어적 체계를 갖추었다는 데 의의가 있다. 이를 구현하기 위하여 본 논문에서는 아이콘 언어의 문법적 체계를 정의하였다. 그리고 적용하고자 하는 도메인의 담화 표본을 수집하고 여기에 아이콘 언어의 문법을 적용한 아이콘 말뭉치를 구축하였다. 또한 이로부터 아이콘에 부여될 어휘와 의미기반의 한국어 생성을 위한 사전을 구현하고 구조 모델을 설계하고 효율을 검증하였다.

한국어 어휘와 아이콘과의 조합으로 구성된 아이콘 언어는 그 동안 일상적인 대화가 어려웠던 언어 장애인에게 큰 도움을 줄 수 있으며, 정보통신 분야에서 소외되었던 사람들에게 그들의 잠재력을 개발하고 일상적인 사회생활을 원활히 할 수 있는 분위기를 만들 수 있을 것이다. 언어 장애인외에도 일시적인 장애나 사고로 인한 환자가 사용할 수 있도록 하여 치료에도 도움을 줄 수도 있을 것이라 생각된다. 그 밖에도 현재 널리 활용되고 있는 안내 시스템, 예약 시스템 등 터치스크린 인터페이스를 가진 아이콘 시스템들의 개발에도 기술적 발전을 제공할 수 있다. 아이콘으로 문장을 구성하는 시스템 구현은 어린이들의 어휘향상을 돕는 학습용으로도 쓰일 수 있을 것이다. 학습용으로 적절한 이유는 그 어휘를 가장 잘 표현한 아이콘과 연관시켜 문장을 구성하고, 직접 음성으로도 들을 수 있게 하기 때문이다. 적절한 멀티미디어 환경을 사용하면 학습효과와 흥미 유발에 도움을 줄 수 있을 것이다. 어린이의 학습용뿐만 아니라, 어휘향상을 이루고자 하는 사람들에게도 도움을 줄 수 있으며, 외국어와 적절히 통합한다면, 외국어 학습에도 좋은 방안이 될 것이다.

이를 위하여 향후 해결해 나아가야 할 과제로는 현재 아이콘 언어가 활용 영역에 의존적인 형태로 설계되어 있어 다른 범주의 응용에 많은 노력이 따르게 된다. 그러므로 폭넓은 아이콘 어휘의 연구와 어휘 데이터베이스의 구축을 통해 범용성을 갖출 필요가 있다. 아이콘은 다중적 의미 및 어휘에 대한 품사 등 많은 애매성을 나타내기 때문에 이로부터 효율적인 한국어 생성을 위해서는 아이콘 언어의 문법

체계를 보다 강건하게 보완해야 한다. 그리고 현재 본 논문에서 논의된 아이콘 언어로부터 출력된 중심적인 어휘 구조 모델에 한국어 생성을 위한 규칙을 정의하고 적용하는 연구와 아이콘을 효율적으로 나타내어 빠르게 어휘를 조합할 수 있는 효율적인 인터페이스 연구가 진행 중에 있으므로 언어 장애인이 이용할 수 있는 편리한 언어보조 인터페이스 시스템 개발을 기대해 볼 수 있으리라 생각된다.

참 고 문 헌

[1] 황인정, 민홍기, “언어 장애인을 위한 통신보조기기의 구현”, 대한의용생체공학회, Vol.20, No.6, pp.621-627, 1999.

[2] Eunsil Lee et al, “Sentence Generation by Symbols for People with language disorder”, ICEIC’98, pp37-40, 1998.

[3] Jongshill Lee et al, “Real-Time Tracking and Recognition of Human Hands from a Sing-Language Image Sequence”, ICEIC’98 pp.165-168, 1998.

[4] Kathleen F. McCoy and Patrick Demasco, “Some Applications of Natural Language Processing to the Field of Augmentative and Alternative Communication” In Proceedings of the IJCAI-95 Workshop on Developing AI Applications for People with Disabilities, Montreal, Canada, August, 1995.

[5] Elman, J. L. In R.F. Port & T. van Gelder(Eds.), “Language as a dynamical system”, Mind as Motion: Explorations in the Dynamics of Cognition. Cambridge, MA: MIT Press. pp.195-223, 1995.

[6] Kathleen F. McCoy, “Simple NLP Techiques for Expanding Telegraphic Sentences”, In Proceedings of Natural Language Processing for Communication Aids, an ACL/EACL ’97 Workshop, Madrid, Spain, July 12, 1997.

[7] 박수현, 강석훈, 권혁철, “한국어-수화 번역시스템을 위한 형태소 변환”, 한국정보처리 학회 논문지, 제5권 제3호, pp.688-702, 1998.

[8] 권경혁, 민홍기, “한국어-수화 번역 시스템 설계”, 한국정보처리학회 논문지, 제7권 제3 호, pp.756-765, 2000.

[9] 강승식, “한국어의 형태론적 특성과 형태소 분석기법”, 정보과학회 논문지, 제12권 제8 호, pp.47-59, 1994.

[10] 엄미숙, “한국 수화의 통사론적 특징 분석”, 석사학위 논문, 대구대학교, 1996.

[11] 김승국, ‘한국 수화 연구’, 오성출판사, 1996.

[12] 최선미, “수화의 방법적 유형에 따른 농학생의 문장 이해 능력 분석”, 석사학위 논문, 대구대학교, 1993.

[13] 서영훈외, “토론기반 한국어 분석기 개발-한국어 의미분석 사전 및 하위범주화 사전 구축”, 한국전자통신연구원 보고서, 1998.

[14] 이종인, 한광록, 양승현, 김영섭, “한국어 명사의 시소러스 구축을 위한 시스템 설계 및 구현”, 한국정보처리학회 논문지, 제6권 제2호, pp.347-356, 1999.

[15] 양승현외, “시소러스와 술어 패턴을 이용한 의미역 부착 한국어 하위범주화사전 구축”, 정보과학회 논문지, 제6권 제3호, pp.364-372, 2000.

[16] 추교남, “개념 기반 정보검색을 위한 한국어 어휘의 의미 분석”, 석사학위 논문, 인천대 학교, 1998.

추 교 남



e-mail : kyonom@incheon.ac.kr
 1997년 인천대학교 정보통신공학과(학사)
 1999년 인천대학교 대학원
 정보통신공학과(공학석사)
 2003년~현재 인천대학교 대학원
 정보통신공학과 박사과정

관심분야 : 한국어 정보처리, XML, 인공지능

우 요 섭



e-mail : yswoo@incheon.ac.kr
 1986년 한양대학교 전자통신공학과
 졸업(학사)
 1988년 한양대학교 대학원
 전자통신공학과 졸업(공학석사)
 1992년 한양대학교 대학원
 전자통신공학과 (공학박사)

1992년~현재 인천대학교 정보통신공학과 교수
 관심분야 : 한국어 정보처리, XML, 네트워크프로그래밍

민 홍 기



e-mail : hkmin@incheon.ac.kr
 1979년 인하대학교 전자공학과(공학사)
 1981년 인하대학교 전자공학과(공학석사)
 1990년 인하대학교 전자공학과(공학박사)
 1985년~1991년 한국과학기술연구원
 선임연구원

1991년~현재 인천대학교 정보통신공학과 교수
 관심분야 : 멀티미디어시스템, 생체공학, 센서네트워크