

문장생성에 의한 통신보조시스템의 설계 및 구현

황인정[†], 우요섭^{**}, 민흥기^{***}

요 약

본 논문은 통신보조시스템을 위한 문장생성의 구현과 설계에 관한 것이다. 통신보조시스템은 언어장애인을 위한 보조 시스템으로서 시간과 키의 수를 줄여 문장을 생성하는데 그 목적이 있다. 본 논문에서는 기존의 문장생성의 장단점을 보완하여 문장생성을 하였다. 문장생성을 위하여 동사와 조사에 따라 명사가 한정되는 한글 구조를 이용하였다. 본 논문의 특징은 도메인 개념을 이용하여 명사와 동사를 연결하였다. 문장생성을 위해 한글의 특성으로 구축한 어휘정보를 이용하였다. 또한 현재 문장생성에 관한 여러 방법을 비교하였다. 문장생성은 문장특징 추출에 의한 어휘정보에 바탕을 둔다.

Design and Implementation of a Augmentative and Alternative Communication System Using Sentence Generation

Ein-Jeong Hwang[†], Yo-Seop Woo^{**}, Hong-Ki Min^{***}

ABSTRACT

This paper designs and implements a sentence generation for an augmentive and alternative communication system(AAC). The AAC system is assistive communication device to help the mute language disorder communicate more freely and the system have an objected to reduce time and keystrokes for sentence generating. The paper of sentence generation make up for merits and demerits in the existing sentence generation method and in order to sentence generation. One aspect of Korean language that confines nouns depending on the verbs or postpositional words is used for sentence generation. The distinctive feature of this paper is to connect verbs to nouns using domain knowledge. We utilize the lexical information that exploits characteristics of Korean language for sentence generation. A comparison with other approaches is also presented. This sentence generation is based on lexical information by extracting characteristics of sentences.

Key words: Augmentative and Alternative Communication(보완 대체 통신), Sentence Generation(문장 생성)

1. 서 론

본 논문은 통신보조시스템에 적용할 수 있는 문장생성에 관한 것이다. 통신보조시스템은 언어장애인이 일상적인 대화를 할 수 있도록 도와주는 기기이

다. 그러므로, 통신보조시스템은 언어장애인이 일상 생활에서 필요한 문장을 빠르게 생성하는데 그 목적이 있다. 통신보조시스템은 보완 대체 통신(Augmentative and Alternative Communication : AAC)의 고급 공학체계 중 하나이며, 보완 대체 통신은 언

※ 교신저자(Corresponding Author) : 황인정, 주소 : 인천시 남구 도화동 177번지(402-749), 전화 : 032)772-2289, FAX : 032)770-8284, E-mail : ejhwang@incheon.ac.kr
접수일 : 2004년 11월 28일, 완료일 : 2005년 5월 3일

[†] 인천대학교 대학원 정보통신공학과(공학박사)

^{**} 인천대학교 정보통신공학과 교수

(E-mail : yswoo@incheon.ac.kr)

^{***} 인천대학교 정보통신공학과 교수

(E-mail : hkmin@incheon.ac.kr)

※ 본 연구는 인천대학교 멀티미디어 연구센터의 일부지원으로 수행되었음.

어장애인들의 언어생활을 도와주는 총체적인 학문 분야이다[1]. 보완 대체 통신은 고급 공학체계와 하위 공학체계로 나누어지며, 고급공학체계는 통신보조시스템, 인터넷에서 글을 읽어주는 가라사대, 음성합성장치, 언어학습도구 등 정보기술을 이용한 제품과 기술을 말한다. 하위 공학체계는 보완 대체 보드와 책 등 사용자가 휴대하면서 손가락으로 의미심볼이나 문장을 지적하여 의사소통을 할 수 있는 단순한 제품을 의미한다[2].

통신보조시스템을 실용화하기 위해 현재 우리에게 필요한 연구부분은 문장생성이다. 통신보조시스템은 이미 선진외국에서 개발되어 있고, 지속적인 발전을 거듭하고 있지만, 언어문제로 인하여 우리에게 직접 적용할 수 없다. 문장생성은 한글을 이용하여 언어장애인이 필요로 하는 문장을 빠르게 생성할 수 있는 방향으로 연구되어야 한다. 그러므로 본 논문에서는 적은 키의 수를 사용하여 문장을 생성하고, 개인 시스템이 될 수 있도록 외부입력 또는 사용자가 필요한 단어를 시스템에 직접 입력할 수 있는 방안을 마련하였다.

통신보조시스템에 맞는 문장생성이 될 수 있도록 어휘정보를 구축하여 이용하였다. 어휘정보는 명사와 동사, 조사의 연결정보를 이용한 것이다. 문장생성은 구축된 어휘정보를 이용하였으며, 사용자 인터페이스를 고려하여 어휘의 입력, 저장, 검색 기능을 추가하였다. 문장생성에 관한 모든 과정에 구축된 어휘정보가 관여한다.

2. 통신보조시스템의 문장생성

2장에서는 통신보조시스템의 문장생성의 여러 방법을 설명하였다. 보완 대체 통신 중 고급 공학체인 통신보조시스템의 문장생성 방법은 의미심볼에 단어나 문장을 1대1로 배치시켜 단순하게 문장을 생성하는 방법, 한정된 영역의 단어를 이용한 단어예측, 의미심볼을 나열하여 문장을 생성하는 방법, 한글 문장의 동사예측 등이 있다.

2.1 어휘와 의미심볼의 연결

단어 또는 문장 한 개를 하나의 의미심볼과 연결하는 방법은 가장 단순한 문장생성이다[1]. 의미심볼은 그림이나 기호로 이루어진 것으로써 어휘의 의미

를 쉽게 표현하고자 하는 수단이다. 이 방법은 한 개의 의미심볼로 하나의 뜻을 나타낸다. 이 방법은 다양한 어휘를 표현할 수는 없으나, 필수어휘를 명확히 표현할 수 있어서 어휘능력이 떨어지는 사용자에게 적용할 수 있는 방법이다. 아직 우리에게는 일상생활에서 사용할 수 있는 적절한 의미심볼이 개발되어 있지 않지만, 구현이 가장 쉽다.

2.2 단어예측

단어예측은 사용자가 문장을 생성하기 위해 단어를 만들 때 그 단어의 글자와 조합 가능한 글자들이 연속적으로 제시되는 방식이다. 그러므로 모든 키를 다 사용하지 않더라도 단어를 만들기가 쉽다. 통신보조시스템에 사용되는 단어예측은 전체 어휘를 기반으로 하는 단어예측이 아니라, 한정된 영역에서 발췌된 어휘를 이용한 단어예측이다[3]. 한정된 영역은 쇼핑, 병원, 교통 등과 같은 영역을 말하는 것으로써, 그 영역에서 발췌된 단어를 이용하여 단어예측을 한다면 사용자에게 필요한 단어를 적은 키의 수로 제시할 수 있다. 통신보조시스템에 사용되는 단어예측은 사용자가 이용하는 장소나 상황에 따라 많이 사용되는 단어를 기반으로 한다. 단어예측은 문장을 생성하기 위해 단어를 쉽게 만드는 방식이므로, 명사와 동사 등의 단어를 만들기 위해 사용된다. 그러므로 문장을 생성할 때 명사와 동사를 만들 때 사용하므로, 전체 문장패턴에서 부분적으로 사용되는 방법이다. 이 방법은 사용자가 원하는 거의 모든 문장을 생성할 수 있는 장점을 가지고 있다. 그러나 어휘능력이 떨어지거나 손가락 사용이 자유롭지 못한 사용자에게는 적합하지 않으며, 문장을 생성할 때 키의 수를 많이 줄일 수 없는 단점을 가지고 있다. 키의 수를 많이 줄이지 못하면 일상대화에서 문장을 생성할 때 시간이 많이 소요된다. 통신보조시스템의 문장생성 방법 중 가장 많이 사용되는 방식이다.

2.3 의미심볼의 배열에 의한 문장생성

의미심볼을 배열하여 문장을 생성하는 방법은 의미심볼이 가지고 있는 특성을 이용한 것이다. 의미심볼은 그림이나 기호로 표시된 것으로써, 문장 내에서 사용된다면 품사에 따라 다양한 의미로 표현될 수 있다[4]. 의미심볼을 여러 개 배열하면 다양한 문장 표현이 가능하기 때문이다. 예를 들어 '사과'를 표시

하는 의미심불의 경우, '사과'라는 표현 외에 '떡다', '빨간색', '맛있다' 등의 표현이 가능하다. 그러므로, 사용자는 의미심불을 배열하여 다양한 문장을 생성할 수 있을 뿐 만 아니라, 키의 수를 줄일 수 있는 장점을 가지고 있다. 통신보조시스템은 휴대용기기이므로, 한 화면에 표현할 수 있는 단어가 많지 않기 때문에 이 방법은 매우 유용하다. 사용자 입장에서도 다양한 문장생성뿐만 아니라 어휘능력을 향상할 수 있는 학습효과도 거둘 수 있다. 이러한 점점에도 불구하고, 현재 이 방법은 좀 더 연구가 필요하다. 한글 문장의 경우, 품사의 위치가 어느 정도 자유로운 구조를 가지고 있기 때문에 문장 내 의미심불의 위치에 따라 그 의미심불이 표현하고자 하는 정확한 품사를 구별하기가 쉽지 않다. 사용자가 배열한 의미심불을 이용하여 문장을 생성했을 때, 사용자가 표현하고자 하는 문장과 뜻이 일치하지 않는 문장이 출력되기가 쉽다. 이 방법을 적용하기 위해서는 구어체 문장패턴에 대한 연구가 진행되어야 하고 문장 내 의미심불이 내포하는 의미와 품사에 관한 연구가 필요하다. 문장패턴에 대한 연구도 일반인과 수화를 사용하는 사용자를 구분하여야 한다. 수화의 경우 일반문장표현과 다른 순서로 문장을 생성하는 경우가 있기 때문이다. 수화는 경우에 따라 의문문, 과거 등의 표현을 문장 앞 또는 뒤에 표현할 수 있기 때문이다.

2.4 동사예측

동사예측에 의한 문장생성은 한글문장의 특성을 이용한 경우로서 명사에 따라 연결될 수 있는 동사를 단계적으로 제시해주는 방법이다[5]. 한글 문장의 경우 조사와 동사에 따라 명사의 종류를 한정할 수 있는 특성이 있기 때문이다. 사용자는 교통, 식당 등의 영역을 정한 후 명사를 선택하게 되면 그 명사와 연결될 수 있는 동사가 제시되면 그 중 하나의 동사를 선택하게 된다. 단어예측보다는 적은 키의 수로 문장을 생성할 수 있으며 올바르게 생성된 문장의 생성을 방지할 수 있는 장점이 있다. 그러나 이 방법은 영역을 나누어 단어를 발체한다 하더라도 한 개의 명사에 연결될 수 있는 동사의 수가 많기 때문에 사용자가 원하는 동사가 도출될 때 까지 여러 단계가 소요될 수 있어 키의 수가 증가될 수 있다. 또한 문장패턴이 명사 하나와 동사하나로 고정되어 있기 때문에 좀 더 자유로운 문장패턴을 만들기 어렵고 의미가 많이

포함되어 있는 동사를 단독으로 도출할 수 없다. 그리고 사용자가 필요한 단어를 직접 입력할 수 있는 방안이 마련되어 있지 않고, 통신 장치에 의해 새로운 단어가 시스템에 입력되면 기존 단어와의 연결을 재작성해야 하는 불편이 따른다. 통신보조시스템은 사용자 중심의 시스템이 되어야 하기 때문에 사용자가 필요로 하는 어휘의 입력과 문장생성이 원활히 이루어져야 한다.

3. 제안하는 문장생성

본 논문의 문장생성은 기존의 문장생성의 단점을 보완하여 연구되었다. 문장생성을 할 때 키의 수가 적도록 하고, 다양한 문장과 올바른 문장을 생성하며, 사용자가 필요로 하는 단어를 직접 시스템에 입력할 수 있도록 하는데 주안점을 두었다. 적은 키를 소모하도록 하기 위해서 단어예측 방법이 아니라 명사와 동사를 사용자가 선택하여 문장이 완성되도록 하였다. 사용자가 명사와 동사를 선택하면 적절한 조사가 붙어 완전한 문장을 출력하는 것이다. 이것은 한글 문장에서 조사와 동사에 따라 명사종류를 한정할 수 있다는 특징을 이용한 것이다[6]. 이러한 특징을 이용하여 본 연구는 올바른 문장생성이 되도록 명사에 따라 연결 가능한 동사종류가 제시될 수 있도록 하였다. 동사종류는 동사에서 비슷한 의미와 특성을 갖는 동사들의 모음이다. 문장은 연결구조(명사종류)+조사+(동사종류)에 의해 생성된다. 사용자는 명사에 따라 적절한 동사종류가 제시되어 문장을 생성하며, 어휘입력을 할 때 사용자가 명사종류와 동사종류 내에 단어를 입력하게 되면 연결구조에 따라 문장을 생성할 수 있는 것이다. 이 연결구조는 문장생성과 어휘입력, 수정, 검색을 쉽게 할 수 있도록 도와준다.

이러한 목표를 달성하기 위하여 어휘정보를 구축하였다. 어휘정보는 어휘의 미시적 특징이 아니라, 한글 문장의 특성을 이용하여 동사에 연결될 수 있는 명사종류와 조사를 분석하여 구축한 것이다.

3.1 어휘정보의 구축

어휘정보를 구축하기 위하여 먼저 각 영역에 따라 문장발체를 한다. 영역은 특정 환경이나 상황을 나타낸 것으로서 본 논문에서는 교통, 식당, 쇼핑, 병원,

집을 영역으로 정하였다. 특정한 영역을 정하여 문장을 발췌하는 것은 사용자가 특정 환경이나 상황에서 필요한 문장을 만들기 위해서이다[1]. 또한 보완 대체 통신 중 통신보조시스템은 주로 휴대하면서 사용하기 때문에 통신보조시스템의 화면에 특정한 영역에서 자주 사용하는 단어를 배열하는 것이 사용자에게 유리한 환경을 제공해준다. 그리고 전 영역에 대해 문장을 생성하는 것 보다, 제한적인 영역의 단어를 이용한 문장생성이 좀 더 명확하기 때문이다.

본 논문에서 문장생성을 하기 위해 각 영역에 따라 문장을 발췌하였다. 그리고 영역 외에 도움말, 인사말을 포함한 중심어휘도 발췌하였다. 중심어휘의 예를 표 1에 나타내었다.

중심어휘와 특정영역에서의 어휘 외에 대화에서의 의미를 분명하게 해주는 색, 시제, 숫자, 단위를 의미하는 단어를 발췌하였다. 색, 숫자, 단어, 시제를 나타내는 단어의 예를 표 2에 나타내었다.

영역에 따라 발췌된 문장은 명사+조사+동사의 연결정보를 얻기 위해 문장 분류를 한다. 문장 분류는 문장을 명사+조사+동사로 분리한 후, 명사는 명사종류, 동사는 동사종류의 대표적 이름을 붙인 것이다. 명사종류, 동사종류의 이름은 분류자의 의도에 따라 달라질 수 있다. 분류명은 명사와 동사를 연결하는 중요한 요소가 된다. 사용자가 명사를 선택했을 때 연결 가능한 동사종류를 제시해줄 수 있기 때문이다.

표 1. 중심어휘의 예

안녕하세요. 도와주세요. 고맙습니다. 미안합니다, 괜찮습니다. 반갑습니다. 예. 아니오.

표 2. 색, 시제, 숫자의 예

분류명	어휘
색	검정, 하양, 밤색, 카키, 노랑, 빨강, 보라, 분홍, 하늘색, 파랑, 갈색, 회색, 연두색, 주황색, 녹색,
숫자	0, 1, 2, 3, 4, 5, 6, 7, 8, 9
단위	개, 장, 묶음,
요일, 날짜	월-일요일, 공휴일, 휴일, 생일, 년-월-일
시제	올해, 작년, 내년 - 월중, 월, 월초, 오늘, 내일, 어제, 현재, 지금, 과거, 현재, 미래, 일년, 한달, 새벽, 아침, 점심, 저녁, 밤, 낮, 오전, 오후, 어둡다, 밝다, 하루, -전, -중, -후, 며칠, -동안, -사이

각 영역에 따라 발췌된 문장을 분류한 예를 표 3에 나타내었다. 표 3은 식당영역의 문장 분류의 예이다.

각 영역에 따라 문장을 분류한 후, 명사종류와 동사종류의 계층구조를 생성한다. 명사종류와 동사종류는 영역 내에서 그 위치가 지정되며, 경우에 따라 세분화될 수 있다. 세분화는 분류된 어휘에 따라 달라질 수 있는 유동적인 구조를 가지고 있다. 식당영역에서 세분화된 계층구조를 표 4에 나타내었다.

3.2 어휘정보

어휘정보는 영역, 명사종류, 동사종류, 조사를 이용하여 구축하는데, 동사의 경우, 동사 기본형과 동사활용을 수집한다. 동사활용은 문장의 형태를 평서문, 의문문, 청유문 등으로 변화시킬 수 있다. 그러나 동사활용은 문장유형이 평서문이라 하더라도 의미 기능에 따라 서술, 의문, 명령, 청유의 의미로 나타낼 수 있고, 의문문도 질문의문문, 요청의문문, 수사의문문으로 분류할 수 있고 더 상세한 분류도 가능하다. 문장 유형은 연구자에 따라 여러 형식으로 분류 [7,8]한다. 그러므로 본 논문에서 평서문은 답변형식으로 의문문은 질문형식으로 문장유형을 수집하였다. 그 밖에 요청, 허용, 허가 등의 문장유형은 동사기

표 3. 문장 분류의 예

이것(지시, 목록)을 주세요. (주문)
어느것(지시, 음식)이 좋을까요. (주문)
음료수(음식)의 종류는 무엇이 있습니까? 있습니다. (주문)
안내해주세요. (안내)
어디에 앉으면 됩니까? (안내)
()명 예약해주세요. (안내)
예약 자리는 어디입니까? (안내)
맛있게 먹었습니다.(인사)

표 4. 식당도메인의 계층구조

영역	상위도메인	하위도메인	
식당	명사	지시	이것, 저것, 그것
		지시-식당	어떤 것, 어느 것
		목록 :음식&기타	물, 음료수, 후식, 반찬...
	동사	목록	메뉴판, 계산서, 영수증
		주문	
		안내	
	인사-식당		

표 5. 동사 어휘정보

문형번호	구 분	단어 또는 문장				
1	기본형	걸리다.				
	동사 하위 도메인명:동사종류	시간				
	명사 하위 도메인명:명사종류	기본	사람	수단	위치	행위
	조사	은, 는	이, 가	으로, 로	까지	하는데
	평서:답변	걸립니다.				
	요청	걸리게 해주세요.				
	의문	얼마나 걸립니까?				
	요청:허가	걸려도 됩니다.				
	평서 부정	걸리지 않습니다.				
	요청 부정	걸리지 않게 해 주세요.				
	의문 부정	걸리지 않습니까?				
	요청 부정	걸리면 안 됩니다.				

본형을 바탕으로 변형하였다. 수집된 동사는 모든 문장유형으로 동사활용하지는 않았다. 어휘정보는 동사 어휘정보와 명사 어휘정보로 나눌 수 있으며, 시스템을 구현한 후 생성되는 저장정보와 보완정보가 포함된다.

동사 어휘정보, 명사 어휘정보, 저장정보, 보완정보는 서로 상호 연결되어 있는 형태를 가지고 있다. 동사 어휘정보에는 동사기본형, 동사종류, 상 하위 도메인, 연결 가능한 명사종류, 조사, 동사활용을 포함하고 있다. 동사 어휘정보의 예를 표 5에 나타내었다.

동사 어휘정보와 명사 어휘정보를 구축할 때 관용표현과 언어 문장[9,10]의 경우, 명사와 동사로 나누지 않고 동사 어휘정보에 포함시켰다. 관용표현과 언어는 특정한 명사와 동사가 결합했을 때 특별한 의미

를 갖는 문장이므로, 이런 경우 명사와 동사로 분류하여 문장을 여럿 만드는 것이 불필요하다. 관용표현과 언어는 사용자가 한 번의 선택으로 사용할 수 있도록 하기 위하여 동사 어휘정보에 포함시켰다. 관용표현이나 언어의 예를 표 6에 나타내었다.

어휘정보는 영역을 정하여 문장발췌와 분류를 하였으나, 발췌된 문장 중 여러 영역에 중복되어 발견되는 문장이 있다. 이런 경우에도 해당 영역에 따라 사용이 가능하도록 하였다. 그리고 조사의 경우, 동사 어휘정보에 있는 명사종류와 동사종류를 비교하여 조사가 연결되나, 명사의 중성에 따라 달라지는 조사 은/는, 이/가, 을/를 의 경우, 유니코드를 이용하여 명사의 중성의 유무를 검색한 후 명사에 맞는 조사를 붙인다.

표 6. 관용표현, 언어의 예

문형번호	구 분	단어 또는 문장			
122	기본형	감기에 걸리다.			
	동사 하위 도메인명:동사종류	건강			
	명사 하위 도메인명:명사종류	기본	사람	행위	시간
	조사	이	이	하는데	동안
	평서:답변	감기에 걸렸습니다.			
	요청				
	의문	감기에 걸렸습니까?			
	요청:허가				
	평서 부정				
	요청 부정				
	의문 부정				
	요청 부정				

어휘정보에 구축된 명사는 현재 284개로 하였다. 명사는 많은 수로 늘릴 수 있으나, 개인사용자에게 필요한 명사를 무한정 늘리는 것이 좋은 것만이 아니기 때문에 필요에 따라 입력할 수 있는 방안을 마련하였다. 그리고 명사에 대응하는 동사는 기본형 103개로 하였다. 기본형에서 파생되는 동사활용은 포함하면 약 300개 이다. 조사의 수는 32개이며, 중심어휘 문장은 8개이다. 시스템에 입력된 단어는 보고서[11]을 활용하였다. 보고서에서 수집된 낱말 수는 총 3090개였으며 그 영역은 영화관, 병원, 미용실, 은행, 비디오, 쇼펍, 음식점, 가정 이었다. 수집된 낱말은 명사가 대부분을 차지하고 있었고, 동사의 경우 그 수가 많지 않았으며, 발견된 동사라 하더라도 동사기본형, 동사활용에 포함되는 경우가 많았다.

4. 시스템의 구현

어휘정보를 이용하여 문장생성을 하기 위해 시스템을 구현하였다. 시스템을 구현하기 위해 먼저 문장생성 조건과 문장생성 방법을 다음과 같다.

4.1 문장생성 원리

본 논문에서 통신보조시스템에 적용하기 위한 문장생성은 영역내의 명사종류와 동사종류의 연결이다. 영역 내에는 여러 개의 명사종류와 동사종류가 존재하며, 그 명사종류와 동사종류는 연결구조가 형성되어 있다. 명사종류와 동사종류의 연결에 따라 알맞은 조사가 있다. 그림 1은 문장생성의 원리를 설명한 것이다. 문장을 생성하기 위한 연결은 어휘정보에 의해 명사(명사종류, 상위도메인)와 동사(동사종류), 조사에 의해 문장이 생성된다. 명사 어휘정보 $N(n+domain+k)$ 로 나타내었다. 해당 명사(n)와 동사연결에 대한 정보와 명사의 상위 도메인(domain), 명사종류(하위도메인 k)를 포함한다. 동사 어휘정보 $V(v+domain+k)$ 는 동사(v)와 동사활용의 정보와 상위도메인(domain) 그리고 동사종류(하위도메인 k)를 포함한다. 조사는 동사 어휘정보와 명사 어휘정보의 비교하여 얻어진 결과를 나타낸다.

문장은 다음과 같은 패턴으로 생성될 수 있다. 첫째, “명사+조사+명사+조사+동사” 둘째, “명사+조사+동사” 셋째, “동사”이다.

문장생성의 원리를 그림 1에 나타내었다.

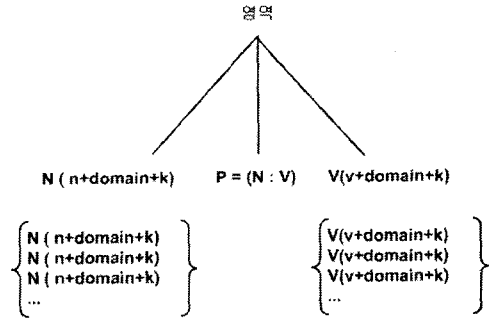


그림 1. 문장 생성의 원리

문장생성 절차는 그림 2에 나타내었다. 사용자가 문장을 생성할 때, 저장, 검색 등에 어휘정보가 이용된다. 생성된 문장을 저장할 때 저장유무의 판단은 어휘정보를 이용하였다. 어휘정보에 포함된 연결구조가 맞지 않는 문장은 저장하지 않고 삭제한다. 저장된 문장은 검색기능을 이용하여 쉽게 재사용될 수 있도록 하였다. 문장의 생성과 저장, 재 출력에 관한 것을 그림 2에 나타내었다.

4.2 구현

문장생성에 의한 통신보조시스템을 다음과 같이 구현하였다. 구현된 시스템은 윈도우 환경에서 C#언어를 사용하였다. 문장생성 시스템은 통신보조시스템의 일반적인 구성요소인 어휘 입력부분, 출력부분, 제어부분으로 구성되어 있다. 각 영역은 왼쪽에

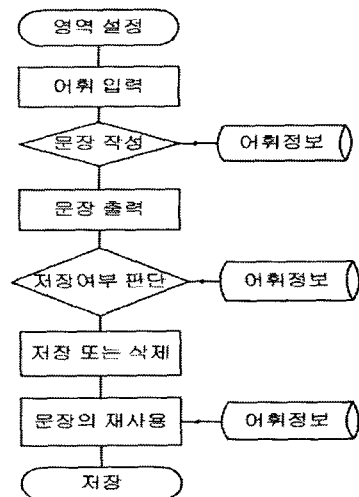


그림 2. 문장생성 절차

배열하였고, 명사종류 그리고 동사종류를 배열하였다. 동사종류는 명사종류에 따라 연결 가능한 경우 색이 바뀌는 형태로 구성하였다. 그림 3은 시스템 전체를 보여주는 것으로서, 교통영역을 선택한 후 명사종류 중 명사를 선택하면 그에 연결 가능한 동사종류가 색이 바뀌어 사용자가 선택하기 편리하도록 구성한 것이다. 사용자가 선택한 단어는 아래에 있는 표시부분에 나타나며, 단어를 수정하고자 할 때에는 명사종류와 동사종류를 열어 쉽게 수정할 수 있도록 하였다. 단어 또는 문장을 사용자가 직접 입력할 수

있도록 입력부분을 시스템 아래에 추가하였다.

그림 4는 단어 수정에 관한 것이다. 사용자가 '주안역'에서 '부평역'으로 단어를 수정한 것으로서, '지명'으로 명명된 동일한 명사종류를 열어 단어를 쉽게 변경하였다.

그림 5는 보완정보의 활용에 관한 것이다. (명사종류)+조사+(동사종류)를 연결하기 때문에 생성된 문장 중 문장이 어색하거나, 조사에 따라 뜻이 달라지는 경우가 발생할 수 있다. 그런 경우 보완정보에 의해 적절한 문장을 보여주어 사용자가 선택할 수 있고

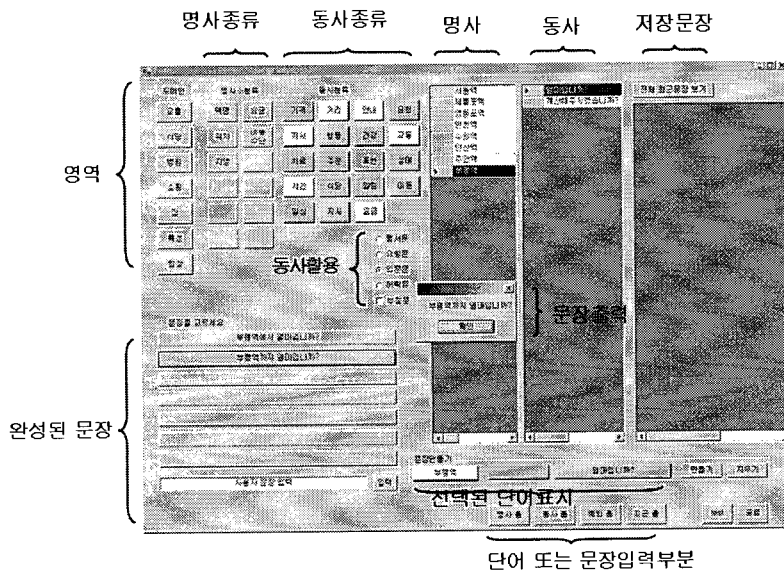


그림 3. 문장생성 시스템

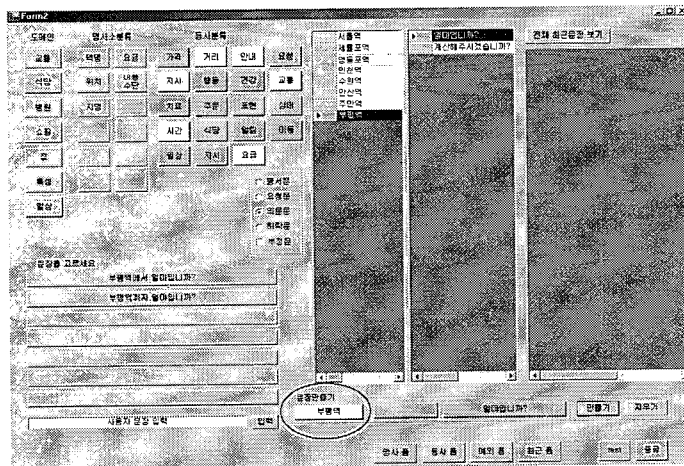


그림 4. 단어의 변경

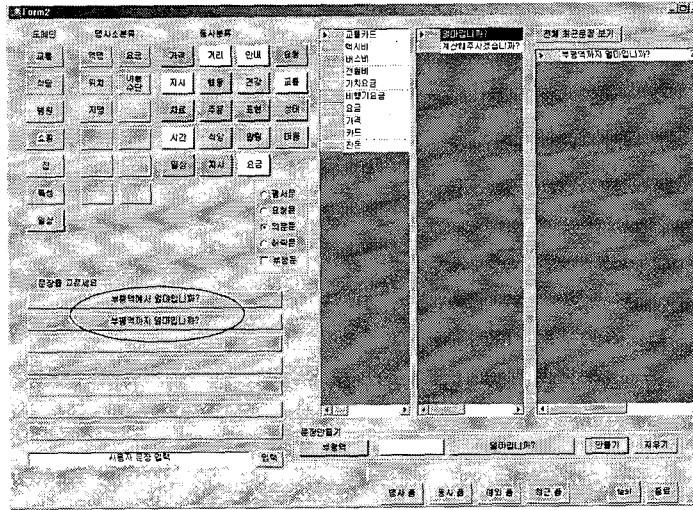


그림 5. 보완정보의 활용

록 한 것이다. 사용자가 명사와 동사를 선택하여 문장을 생성할 때 같은 명사와 동사를 가지고 있으나, 다른 조사가 포함되어 뜻이 달라지는 경우의 예를 그림 5에 나타내었다. 이런 경우 한 문장으로 생성할 수 없기 때문에 보완정보를 이용하여 사용자의 선택을 기다리는 과정을 보여준 것이다.

그림 6은 시스템에서의 검색기능을 보여준 것이다. 사용자가 문장을 생성하기 위해 명사 또는 동사를 선택하면 저장 문장 중 그 단어를 포함하고 있는 문장을 빈도수에 맞춰 보여준 것이다. 검색기능은 이미 사용된 문장을 다시 생성하지 않기 때문에 편리하

다. 문장의 생성과 동시에 저장문장에서 검색이 이루어지기 때문에, 사용자가 따로 검색하지 않아도 된다.

4.3 검토

본 논문의 문장생성은 일정한 키의 수를 유지한다. 문장생성을 위한 최대 소요 키의 수는 9개 이다. 그것은 영역선택, 명사종류, 명사선택(1), 명사선택(2), 동사종류, 동사선택, 동사활용, 보완정보, 출력까지 총 9단계이다. 시스템에서 주로 사용되는 키의 수는 명사선택(2), 동사활용, 보완정보를 제외하고 6~7의 키가 소요되며, 저장문장을 활용할 경우 2~3개의

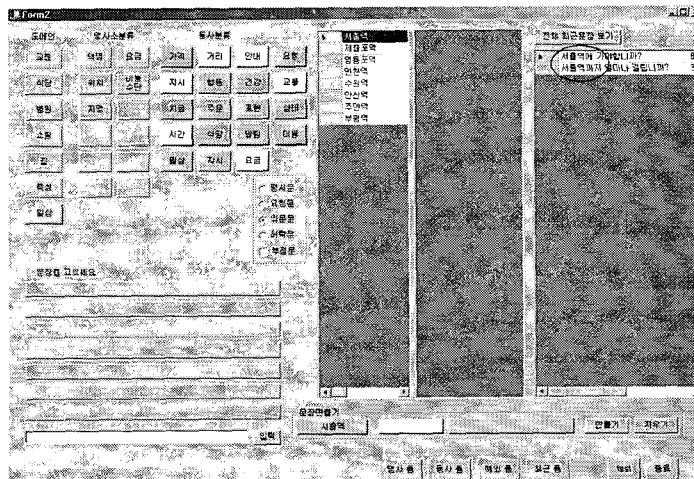


그림 6. 문장 검색 기능

표 7. 문장생성 방법의 비교

	키의 수	생성문장 수	구현가능성	어휘입력기능	문장패턴
의미심볼(일대일 연결)	가장 적음	가장 적음	가장 쉬움	없음	고정됨
단어예측	많음	많음	쉬움	있음	자유 문장패턴
의미심볼의 배열	적음	많음	어려움	있음	자유 문장패턴
동사예측	적음	많음	쉬움	없음	고정됨
본 논문의 문장생성	적음	많음	쉬움	있음	약간 허용

* 본 논문의 문장생성의 패턴은 동사만을 출력하거나, 명사 2+동사, 명사 1+ 동사의 출력이 허용된다.

키가 필요하다.

문장생성 시스템의 효율을 알아보기 위해 본 시스템에서 생성된 문장을 키보드에 의한 자모수와 시스템의 키의 수를 비교하였다. 시스템에서 생성되는 문장 중 명사와 동사로 이루어진 문장의 경우 총 자모수는 15에서 34를 보였고, 평균 24.2의 자모수를 가지고 있다. 관용어와 연어의 경우는 평균 18.2의 자모수를 가지고 있다. 총 자모수는 띄어쓰기, 마침표, 물음표를 포함한다.

이러한 결과를 이용하여 시스템의 문장생성의 효율을 계산해 보았다.

효율 계산은 식 [12] 이다.

$$\left\{ 1 - \frac{(\text{시스템의 키의 수})}{(\text{총 자모수})} \right\} \times 100$$

명사와 동사로 이루어진 문장의 경우, 본 시스템의 최대 키의 수는 9개이며, 평균 자모수가 24.2 이므로 효율은 약 62.8%를 보인다. 약 62.8%의 효율은 최소 효율이라고 할 수 있다. 본 시스템에서 주로 사용되는 키의 수가 6-7 정도이고, 관용어와 연어의 경우는 키가 3-4 정도 소요되므로 더 효과적일 것이다. 저장문장과 검색기능을 이용하면 가장 높은 효율을 보일 것이다. 표 7은 여러 문장생성 방법을 비교한 것이다.

5. 결 론

본 논문은 언어장애인을 위한 문장생성에 의한 통신보조시스템의 설계 및 구현이다. 통신보조시스템에 관한 연구 부분 중 문장생성에 주안점을 두었다. 문장생성은 통신보조시스템에 적용할 수 있는 것으로서 키의 수를 줄이면서 다양한 문장을 생성할 수 있고, 어휘입력이 가능하도록 하였다. 문장생성을 하기 위하여 명사와 조사, 동사의 연관성에 관한 정보

를 이용하여 어휘정보를 구축하였으며, 어휘정보는 명사 어휘정보와 동사 어휘정보 그리고 시스템 구현 후 보안정보와 저장정보가 있다. 시스템은 사용자가 문장의 생성과 저장, 입력을 할 수 있도록 구현되어 있다.

향후과제는 문장 유형을 다양화 하고, 영역의 확대 그리고 의미심볼과 연관된 문장생성에 관한 연구가 진행되어야 하며, 연구가 원활히 진행되기 위한 연구로는 사용자의 연령과 각 영역에서 필요한 단어가 체계적으로 수집되어야 할 것이다.

본 논문의 문장생성은 휴대용 통신보조시스템에 적용하기 위한 것이기 때문에, 실용화되기 위해서는 사용자에게 적합한 하드웨어장치와, 소프트웨어 그리고 서비스를 위한 정보통신 기술의 통합이 필요하다.

참 고 문 헌

- [1] Sharon L, Glennen; Denise C, and DeCoste, *The Handbook of Augmentative and Alternative Communication*, Singular Publishing Group, 1996.
- [2] 강운주, 원성옥, "발달장애 아동의 보완대체 의사소통 도구로서 수화언어의 활용," *한국정서학습장애아교육학회, 정서·학습장애 연구*, Vol. 18, No. 2, pp. 137-157, 2002.
- [3] Matthew E.J. and Wood, "Syntactic Pre-processing in Single-Word Prediction for Disabled People," *Department of Computer Science University of Bristol*, pp. 206-216. 1996.6.
- [4] ZYGO Industries, Inc. <http://www.zygo-usa.com>
- [5] 이은실, 제한된 공간내에서 동사예측기법을 이

용한 언어장애인용 문장발생장치의 구현, 인하대학교 박사학위 논문, 2001.2.

- [6] 홍재성 등 9인, *현대 한국어 동사 구분 사전*, 두산동아 1997.1.10.
- [7] 박영순, *한국어 문장의미론*, 도서출판 박이정, 2001.11.25.
- [8] 남기심, 고영근, *표준국어문법론, 개정판*, 탑출판사, pp. 96-108, 1997.
- [9] 김진해, “언어 (collocation)의 의미 관계에 대하여,” *한국어어미학회, 한국어 의미학*, Vol. 4, pp. 239-280, 1999.
- [10] 고평주, “관용어의 논항구조와 형성제약,” *민족어문학회 (구 안암어문학회), 어문논집*, Vol. 42, pp. 261-284, 2000.
- [11] 민홍기 등 10인, *언어장애인을 위한 언어생활지원 시스템 개발*, 보건복지부, 최종 연구결과 보고서, 2003. 11.
- [12] Ann Copestake. “Augmented and alternative NLP techniques for augmentative and alternative communication,” *In Proceedings of the ACL workshop on Natural Language Processing for Communication Aids*, pp. 37-42, Madrid. 1997.



황 인 정

1990년 인천대학교 물리학과 졸업(이학사)
 1999년 인천대학교 대학원 정보통신공학과(공학석사)
 2004년 인천대학교 대학원 정보통신공학과(공학박사)

관심분야: 보완 대체 통신(AAC), HCI, WEB



우 요 섭

1986년 한양대학교 공학사
 1988년 한양대학교 공학석사
 1992년 한양대학교 공학박사
 1992년~현재 인천대학교 정보통신공학과 교수

관심 분야: 한국어 정보처리, 데이터베이스, 인터넷 프로그래밍



민 홍 기

1979년 인하대학교 전자공학과 졸업(공학사)
 1981년 인하대학교 대학원 전자공학과 정보공학 전공(공학석사)
 1985년 인하대학교 대학원 전자공학과 정보공학 전공(공학박사)

학박사)

1985년~1991년 한국 과학 기술 연구원 선임연구원
 1991년~현재 인천대학교 정보통신공학과 교수
 관심분야: 디지털 신호처리, AAC, 정보검색