

가변어휘 단어 인식기를 사용한 음성 명령 웹 브라우저

Voice Command Web Browser Using Variable Vocabulary Word Recognizer

이 항 섭*
(Hang Seop Lee*)

*이 연구는 정보통신부 출연 "HCI를 위한 음성 입출력 처리기술 개발" 과제의 연구 결과물입니다.

요 약

본 논문에서는 웹 브라우저 상에서 한국어 음성인식을 이용하여 정보검색을 할 수 있는 가변어휘 단어 인식기를 사용한 음성 명령 웹 브라우저에 대하여 기술한다. 이 시스템의 특징은 웹 브라우저 상에서 보여지는 링크를 가지는 HyperText Word들과 웹 브라우저 메뉴를 음성으로 인식할 수 있는 것으로, 마우스 click 뿐만이 아니라 음성인식을 이용하여서도 웹 브라우저를 사용할 수 있다는 것이다. 웹 브라우저를 통해서 보여지는 문서에서 추출되는 인식 후보들은 각 문서에 따라 고정되지 않고 계속하여 변화하므로, 이러한 가변적인 인식 후보들을 인식하기 위해 가변어휘 단어 인식기를 사용하였다. 가변어휘 단어 인식기는 훈련용 음성 데이터와 무관한 임의의 새로운 어휘를 훈련 없이 인식해 낼 수 있는 인식기로 POW (Phonetically Optimized Words) 3,848 단어를 사용하여 훈련한 결과 32단어에 대해 93.8%의 단어 인식률을 보인다. 음성 명령 웹 브라우저는 Windows 95/NT 환경에서 Netscape Navigator를 사용하여 개발되었으며, 사용자가 음성을 사용하는 새로운 인터페이스를 배울 필요 없이 바로 사용할 수 있도록 사용자 편의성 부분도 고려하여 개발되었다. 개발된 음성 명령 웹 브라우저는 환경 독립, 화자 독립에 대해 On-line으로 실험한 결과 평균 90%의 인식성능을 보인다.

ABSTRACT

In this paper, we describe a Voice Command Web Browser using a variable vocabulary word recognizer that can do Internet surfing with Korean speech recognition on the Web. The feature of this browser is that it can handle the links and menus of the web browser by speech. Therefore, we can use speech interface together with mouse for web browsing. To recognize the recognition candidates dynamically changing according to Web pages, we use the variable vocabulary word recognizer. The recognizer was trained using POW (Phonetically Optimized Words) 3,848 words. So that it can recognize new words which did not exist in training data. The preliminary test results showed that the performance of speaker-independent and vocabulary-independent recognition is 93.8% for 32 Korean words. The Voice Command Web Browser was developed on windows 95/NT using Netscape Navigator and reflected usability test results in order to offer easy interface to users unfamiliar with speech interface. In on-line experiment of speaker-independent and environment-independent situation, Voice Command Web Browser showed recognition accuracy of 90%.

1. 서 론

인터넷의 소개 이후 인터넷의 사용은 지금도 전세계적으로 급격히 늘어나고 있으며 1998년 현재 전세계의 인터넷 호스트는 2,967만대에 이르며, 그 사용자는 3억에 달하고 있다. 인터넷은 정보 검색 및 정보 공유의 표준적인 방법으로 사용되고 있다. 국내에서도 인터넷 이용자는 급격히 증가하고 있으며, 그 영역은 전문적인 내용에서부터 기업 소개, 홈 쇼핑, 레저, 문화 및 전자 상거래로 계

속 확장되고 있는 추세이다. 인터넷 뿐만 아니라 인트라넷의 사용도 사내 전자결재 시스템의 도입으로 크게 확산되고 있는 실정이다. 이러한 인터넷과 인트라넷의 사용은 웹 브라우저를 통해 이루어진다. 이러한 웹 브라우저에 음성을 이용한 인터페이스가 추가 된다면 보다 쉬운 인터넷의 사용으로 사용자 층을 더욱 넓힐 수 있을 것이다.

이러한 이유에서 ETRI에서는 음성인식을 이용하여 웹 브라우저 상에서 정보를 검색할 수 있는 시스템인 음성 명령 웹 브라우저를 개발 하였으며, 이를 EVANS (ETRI Voice Activated Navigation System)라 명하였다. 음성인식을 사용하여 웹 브라우저 상에서 정보검색을 하는 데는 아직 해결해야 할 여러 가지 문제점들이 남아 있지만

* 한국전자통신연구원 음성신호처리팀
접수일자: 1998년 9월 29일

우리의 이러한 시도가 음성인식의 실용화를 앞당길 수 있기를 기대한다.

본 논문에서는 EVANS의 특징 및 시스템 구성 등 개발에 대한 전반적인 내용을 소개하고, 구현된 시스템에서의 On-line 인식실험 결과를 기술한다.

II. EVANS 개요

EVANS는 웹 브라우저 상에서 음성인식을 이용하여 정보검색을 할 수 있게 해주는 음성 인터페이스 시스템으로 EVANS의 특징은 다음과 같다.

1) 어휘독립, 화자독립 인식

사용자는 별도의 훈련 없이 접속 가능한 모든 인터넷 사이트에서 음성 인식을 이용하여 정보검색을 할 수 있다.

2) Browser 메뉴에 대한 음성 명령:

사용자는 브라우저 윈도우 control을 포함한 웹 브라우저의 모든 메뉴를 음성으로 접근할 수 있다. 예) "인쇄", "열기", "최대크기", "스크롤업"

3) Link에 대한 음성 명령:

사용자는 현재의 웹 페이지가 제공하는 모든 Link들을 음성으로 접근할 수 있다.

4) 음성북마크:

사용자는 음성으로 접근할 수 있는 자신만의 음성북마크를 만들 수 있다.

5) 음성명령을 현재의 Web page에 표시:

사용자는 음성인식 후보들이 웹 페이지 상에 표시되므로 어떠한 것들이 음성인식 후보인지, 그리고 이러한 음성으로 접속할 수 있는 Link들에 대한 발성 방법을 쉽게 알 수 있다.

6) 메뉴와 음성북마크에 대한 음성명령 재 정의:

사용자는 웹 브라우저 메뉴와 음성북마크에 대한 음성명령을 사용자가 원하는 대로 재 정의 할 수 있다.

EVANS는 세 가지 기능을 가지고 있다. 첫째는 웹 브라우저가 현재 보여주고 있는 웹 페이지에서 인식후보를 추출하여 새로운 단어 사전을 생성하는 것이고, 둘째는 생성된 단어 사전을 사용하여 인식을 수행하는 것이고, 셋째는 인식결과에 따라 웹 브라우저를 동작시키는 것이다. EVANS가 동작되면 음성 인식을 사용하여 웹 브라우저를 동작시키기 위한 음성인식후보를 얻기 위해 현재의 웹 페이지에 대한 parsing을 수행한다. 웹 페이지에 대한 parsing 결과로 얻어진 음성인식 후보는 HTML 파일에 기록되고 웹 브라우저를 통해 보여진다. 이제 사용자는 자신이 가고자 하는 site로 가기 위해서 그 site를 나타내는 하나 또는 몇 개의 단어를 발성하면 된다.

현재의 웹 페이지에서 다른 곳으로의 link를 제공하는 HyperText word들은 웹 페이지에 따라 계속하여 변화하기 때문에 이들을 인식하기 위해서 EVANS에서는 미리 훈련되지 않은 단어들에 대한 인식이 가능한 가변어휘

단어 인식기를 사용하였다.

III. 가변어휘 단어 인식기

가변어휘 단어 인식기[1][2]는 그림 1과 같은 구조를 갖는다. 이 인식기는 기존의 인식기와 달리 미리 훈련되지 않은 단어를 인식할 수 있는 기능을 가지고 있다. 인식 대상이 되는 단어 목록이 바뀌게 되면 인식할 어휘에 대한 음성 훈련을 수행하지 않고 단지 발음 사전만을 새로 교체하여 단어 모델들을 재구성 하므로 이론적으로 무제한의 임의의 단어를 주어진 단어 목록 내에서 인식할 수 있게 된다. 이러한 단어 인식기를 구현하려면, 우선 한국어에 존재하는 모든 음소를 충분한 음소환경에서 정확히 모델링 해야 한다. 이렇게 하기 위해서는 먼저 각 음소를 정확히 모델링하기 위한 훈련 데이터를 다양한 음소 환경하에서 수집해야 하며, 또 이를 음소 모델에 적절히 반영시키기 위하여 이러한 다양성을 포용할 수 있는 음소 모델 구조를 가져야 한다. 이러한 조건을 충족시키기 위하여 본 연구에서는 훈련용 음성 데이터로써 당 연구실이 제안한 POW 3,848 단어 목록[3]을 사용하여 다수의 화자로부터 음성을 수집하여 사용하고, 음소의 다양성을 모델 구조에 반영하기 위하여 음소만이 아닌 변이음까지의 상세 모델링[4]을 함으로써 각 모델의 정확도를 향상시켰다.

사용된 POW 3848 DB는 어휘수가 총 3,848개로 구성되었으며, 이를 8명이 481개씩 나누어 발성한 것을 1개의 set으로 하였다. 음소 모델 훈련을 위해 사용한 것은 수작업으로 labelling 되어있는 남성음 3 set과 여성음 2 set (약 19,240 단어)이다. 이 음성 DB는 비교적 조용한 녹음실에서 수집되었으며, 16 kHz, 16Bit로 양자화 되었다.

특징벡터 추출 과정은 다음과 같다. 먼저, 10 msec 마다 265 point FFT를 수행하고, 이로부터 PLP (Perceptually linear prediction) 특징벡터를 구한다. 구해진 특징벡터로부터 dynamic feature를 구하기 위해 FIR Filter를 사용하여 first-order dynamic feature를 얻고, 이 두 가지 벡터를 연결한 26차 벡터에 mean-subtraction을 이용한 정규화를 거쳐 최종적인 26차 특징벡터를 구한다.

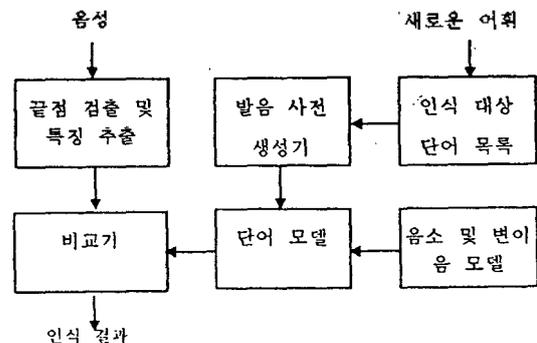


그림 1. 가변어휘 단어 인식기의 구조
Fig. 1. Block diagram of variable vocabulary word recognizer.

IV. EVANS시스템 구성 및 구현

4.1 시스템 구성

EVANS는 그림 2와 같이 HTML parser, 인식후보 추출 모듈, 가변어휘 단어 인식기, 인식결과 분석기로 이루어졌다. 웹 브라우저가 새로운 HTML 파일을 open하면 이 파일은 HTML parser로 전달된다. HTML 파일은 다른 site로의 접속을 제공하거나 같은 페이지내의 특별한 곳으로 이동하기 위한 link(URL)를 가지고 있다. 인식후보를 추출하기 위해서 이러한 link들을 가지는 hypertext word를 사용하였다. Hypertext word들은 서로 다른 길이를 가지며, 단어, 이미지, 숫자 또는 이런 것들의 혼합으로 이루어진다. 이러한 다양한 형태의 hypertext word로부터 실제로 사용될 인식후보를 추출하기 위해 "인식후보 추출 모듈"이 사용되었다. 인식후보 추출 모듈은 hypertext word들과 음성북마크 명령, 웹 브라우저 메뉴, 웹 브라우저 윈도우 제어 명령들을 사용하여 현재의 웹 브라우저 상에서 사용될 인식후보를 만들어 낸다. 표 1은 사용된 웹 브라우저 윈도우 제어 명령들이다.

표 1. 웹 브라우저 윈도우 제어 명령
Table 1. Web Browser window control commands.

Browser Window Control Command	
Normalize	Scroll up
Maximize	Scroll down
Minimize	Scroll right
Page up	Scroll left
Page down	

음성북마크는 웹 브라우저가 제공하는 북마크와 같은 기능을 제공하는 것으로 사용자가 미리 등록된 site에 음

성으로 접속할 수 있게 해준다. 또한, 사용자는 새로운 음성북마크를 "북마크설정"이라는 음성명령을 사용하여 추가할 수 있으며, 텍스트 기반의 음성북마크 파일을 별도로 관리하므로 음성북마크의 내용을 text 편집기를 사용하여 쉽게 수정 할 수 있다. 사용자가 "북마크설정" 명령을 내리면 시스템은 현재 접속중인 site를 등록하기 위한 창을 제공한다. 여기에 사용자는 현재의 site를 음성으로 접속하기 위한 text 이름을 적어주면 된다.

HTML parser로부터 얻어지는 hypertext word로부터 음성인식 후보를 생성하는 과정은 다음과 같다. 첫째, 만일 hypertext word가 3 단어 이상으로 이루어져 있으면 3 단어까지만 추출하며, 추출된 단어들 전체를 하나의 단어로 취급을 한다. 이렇게 하는 이유는 hypertext word들이 다양한 길이를 가지고 있으므로, 만일 각각의 단어를 모두 독립적인 인식 후보로 취급할 경우 똑같은 인식 단어를 가지는 link(URL)들이 존재하게 된다. 이러한 상황에서는 현재 인식된 단어가 어떠한 link를 가리키는지 알 수 없게 된다. 예를 들어, "증권", "증권 시세", "증권 정보 속보"와 같은 hypertext word들이 같은 웹 페이지에 존재한다면, "증권"이라는 단어를 인식했을 때 이것이 위의 3개의 link중 어디를 가리키는지 알 수 없게 된다. 그래서, 이들을 "증권", "증권-시세", "증권-정보-속보"와 같이 하나의 단어처럼 취급하여 서로를 구분하였으며, 또한 긴 단어를 발성하는 불편함을 없애기 위해 단어의 길이를 3 단어로 제한하였다. 둘째, 위와 같이 하여도 이미 똑같은 이름의 인식후보가 존재 한다면 이를 "하나", "둘", "셋"과 같은 연속적인 숫자 인덱스로 바꾸어 준다. 또한 이러한 숫자 인덱스는 hyperlink를 제공하는 것이 한글이 아닌 그림이나 숫자 또는 영어 등 한국어를 제외한 다른 언어의 경우에 대해서도 적용된다. 이 과정을 거쳐 생성된 인식 후보들은 각 인식 후보들이 가리키는 link(URL)과 함께 인식후보 리스트에 기록된다. 가변어휘 단어 인

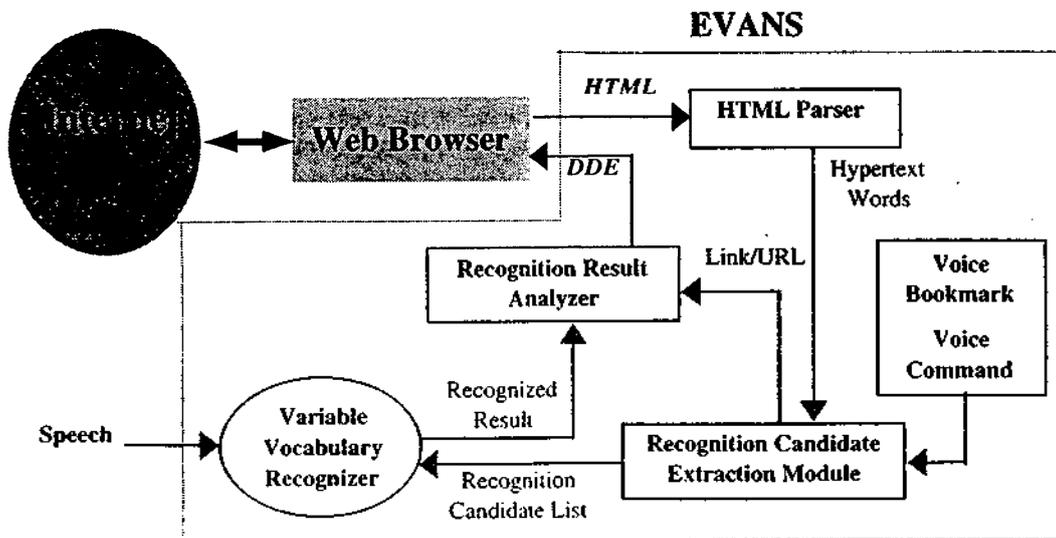


그림 2. EVANS 구성
Fig. 2. Configuration of EVANS.

식은 현재의 웹 페이지에서 추출된 인식후보 리스트를 사용하여 인식을 수행하며 인식된 결과는 인식결과 분석기로 전달된다. 인식결과 분석기는 인식기의 인식결과에 해당하는 적절한 명령을 웹 브라우저에 보낸다. 웹 브라우저로는 Netscape Navigator를 사용하였다. Netscape Navigator와 통신을 위해서는 Dynamic Data Exchange (DDE) method [5]를 사용하였다.

4.2 시스템 구현

EVANS는 Windows 95/NT 환경에서 Pentium pro 200 MHz CPU, 일반적인 사운드 카드, Netscape Navigator를 갖춘 PC에서 구현되었다. EVANS는 자동음성검색[6] 기능을 가지고 있어 음성 입력의 시작과 끝을 시스템에 알리기 위한 keyboard나 mouse 입력을 필요로 하지 않는다. 이 기능은 사용자에게 편의성을 제공하는 반면, 주변 소음이 너무 크거나 일정하지 않은 상황에서는 시스템의 성능에 나쁜 영향을 주기도 한다. EVANS는 또한 음성을 인터페이스로 사용하기 위한 여러 가지 요구사항[7] 들을 갖추고 있다. EVANS에서 사용된 인식기는 ETRI에서 개발한 화자독립, 어휘독립 SCHMM-base의 가변어휘 단어 인식기이다. 이 인식기는 문법 없이 간단한 문장을 인식할 수 있으나, 이 시스템에서는 단어인식 기능만을 사용하였다.

V. 사용 편의성 평가

EVANS에 대한 사용 편의성 평가는 사용자의 불편함을 최소화 하고 쉽게 배울 수 있으며 음성인식 기능의 효율을 극대화 할 수 있도록 문제점을 파악하고 이를 수정하는데 목적을 두었다. 평가는 전문가 의견에 의한 평가, Benchmark Test, Hands-on Experiment 등 3개 부분의 평가를 실행하여 평가 결과를 체계적이고 종합적으로 도출하였다[8]. 다음은 각 평가에 대한 간략한 내용을 기술한다.

- 전문가 의견에 의한 평가: 음성인식시스템의 각 평가 대상들이 Usability Principle이나 Guideline에 따라 설계되었는지 여부를 사용 편의성 전문가의 판단으로 평가한다.
- Benchmark Test: 시스템에서 사용자가 수행하게 될 대표적인 작업을 Benchmark Task로 선정하여 사용자에게 수행해 보도록 함으로써 시스템 사용상의 문제점을 파악한다.
- Hand-on Experiment: 사용자에게 시스템을 자유롭게 사용하게 함으로써 사용 중 발생하는 문제점을 파악한다. 이러한 사용 편의성 평가 결과 visual feedback과 인식어휘 선정 그리고, 인식후보 제시 방법이 크게 개선되었으며 이 중 인식어휘 선정 및 제시 방법에 대해 도출된 의견을 표 2와 같은 방법으로 반영하였다.

VI. EVANS 성능 실험

EVANS 인식 성능 테스트는 on-line 가변어휘 단어 인식 상황에서 인식률과 인식 속도의 변화에 중점을 두었다.

표 2. 사용 편의성 평가 결과 반영 내용

Table 2. The application of Usability Test Results.

평가 결과 제안된 내용	반영된 내용
긴 어휘는 Learning하기 쉽고 일관되게 발음이 가능하며, 더 많은 음성 정보를 제공하므로 짧은 어휘보다 인식되기 쉽다.	링크를 3단어까지 발성할 수 있도록 함
자연스럽고(Natural) 친밀한 어휘를 사용한다.	화면에 보이는 단어 그대로 발성하게 함.
사용자가 자신의 어휘를 선정하도록 하는 것이 좋다.	웹 브라우저 메뉴와 음성복마크의 발성을 사용자가 쉽게 재정의 할 수 있는 기능 제공.
단어 선택의 유연성을 높이기 위해 'Multiple Mapping'을 이용한다. (한 기능을 여러 어휘로 가능하게 한다)	웹 브라우저 메뉴의 경우, 하나의 메뉴를 사용자가 여러 개의 어휘로 발성할 수 있는 기능 제공.
어떠한 것이 음성인식 후보이고, 인식된 단어가 어떠한 것인지를 쉽게 인지할 수 있도록 하는 것이 좋다.	· 음성인식 후보를 HTML 문서상에 표시. ([]나 [] 같은 것을 사용하여 표시) · 전체 인식 후보를 쉽게 인지할 수 있도록 별도의 "사용가능 인식후보창"을 제공.
인식기의 상태를 사용자가 쉽게 인지할 수 있게 하는 것이 좋다.	"인식상태 창"을 두어 음성입력 상태에서 웹 페이지 loading까지의 전 과정을 표시함.

실험은 EVANS의 실제적인 사용 상황에서의 인식을 계산을 위해 실제로 존재하는 인터넷 사이트를 접속한 후 인식 대상인 링크와 Netscape 메뉴 그리고, 음성복마크 명령을 내리고 그 결과를 기록하는 방법으로 실험을 수행하였다.

6.1 시험 절차

인식 실험에 참여한 화자와 시스템 규격 그리고 환경은 다음과 같다.

- ◆ 화자: 표준말을 사용하는 30대 초반의 남성 2인과 30대 후반의 남성 1인
- ◆ 시스템:
 - 화자 1: Pentium Pro 200MHz Dual CPU, 128Mbyte RAM
 - 화자 2: Pentium II 266MHz Dual CPU, 256Mbyte RAM
 - 화자 3: Pentium 166MHz, 32Mbyte RAM
- ◆ 환경: 다수의 PC와 Workstation 그리고 printer가 있는 사무실 환경
- ◆ 마이크: SENNHEISER HMD 410 (Close-Talk Microphone)

실험은 사이트를 바꾸어 가며 총 4번을 실시 하였다. 각 사이트는 전체 어휘의 종류, 어휘의 개수 및 평균 단어의 길이에 변화가 있는 곳으로 선정하여 단어 수와 단어 길이의 변화에 따른 인식률과 인식시간의 변화를 기록하였다. 사용된 사이트는 "http://www.seoul.net"와 그 밖의 사이트들이다.

EVANS의 인식대상은 크게 3가지로 Hyperlink를 제공

하는 단어, Netscape 자체 메뉴, 음성복마크이다. 실험에서는 이 3가지 인식 대상을 그 수에 비례하게 선택하여 발생하도록 하였다.

6.2 실험 결과

실험 결과는 표 3에 나타내었다.

표 3. On-line 인식 실험 결과
Table 3. Performance evaluation results of on-line recognition.

	전체 어휘수	가변 어휘수	평균 단어 길이	발성 단어수	인식률 (%)	인식 시간(sec)
실험 1	129	90	3.76	50	86.7	1.73
실험 2	116	77	4.03	50	88.7	1.70
실험 3	119	80	4.8	50	93.3	1.76
실험 4	146	106	5.11	50	92	1.948
평균	127.5	88.3	4.25	50	90.2	1.78

실험 결과에 의하면 가변어휘 상황에서는 가변어휘의 수에 비례하여 인식률이 달라지기 보다는 가변어휘의 종류가 훈련된 데이터에 얼마나 유사하는가가 인식률에 더 큰 영향을 주는 것을 알 수 있다. 또한 발성 단어의 길이가 길수록 인식률이 좋은 결과를 보이고 있어 짧은 단어 보다는 긴 단어가 인식시간은 오래 걸려도 인식률은 더 좋음을 알 수 있다. 여기에서의 인식시간은 사용된 모든 시스템들의 평균 인식시간이다.

VII. 결 론

본 논문에서는 웹 브라우저 상에서의 음성인식을 이용한 정보검색 시스템인 EVANS를 소개하였다. EVANS에서 인식 가능한 대상은 웹 브라우저 자체 메뉴와 브라우저 윈도우 제어 메뉴, 그리고 현재 웹 페이지에서 보여지고 있는 HTML 파일에서 다른 곳으로의 링크를 제공하는 hypertext word들이다. 인터넷 사이트 마다 달라지는 웹 페이지내의 hypertext word들을 인식하기 위해서 화자 독립, 어휘독립 인식기능을 가지는 가변어휘 단어 인식기를 사용하였다. EVANS를 사용함으로써 인터넷을 통한 정보 검색시 mouse와 함께 음성을 이용한 인터페이스가 가능하게 되었다.

가변어휘 단어 인식기의 시뮬레이션 인식률(93.8%)에 비해 on-line 인식률(90.2%)이 나쁘게 나온 이유는 인식기 사용 환경의 차이와 이에 따른 음성검출기의 결과가 시스템의 전체적인 인식률에 영향을 미치는 것으로 생각된다. 그러므로, 실제로 음성인식 기술이 실생활에 사용되기 위해서는 우선 인식기 자체의 성능이 좋아야 하고, 인식기의 성능을 저하시키지 않는 전처리 과정의 개발 그리고 인식기의 성능을 최대한 발휘할 수 있는 응용 분야의 선택 및 시스템화에 보다 많은 노력을 기울여야 할 것이다.

참 고 문 헌

1. Hoi-Rin Kim and Hang-Seop Lee, "Variable vocabulary word recognizer using phonetic knowledge-based allophone model," *Jour. of Acoustical Society of Korea*, vol. 16, no. 2, pp. 31-35, Feb. 1997.
2. 김희린, 이항섭, "음성학적 지식 기반 변이음 모델을 이용한 가변 어휘 단어 인식기," *한국음향학회지* 제 16권 제 2호, pp. 31-35, 1997.
3. Yeonja Lim and Youngjik Lee, "Implementation of the POW (Phonetically Optimized Words) algorithm for speech database," *Proc. of ICASSP*, pp. 89-91, 1995.
4. 서영주, 상철재, 이장철, 한민수, 이영직, "음성학적 지식에 기반한 한국어 변이음 집단화 수행도의 구현," 제13회 음성통신 및 신호처리 워크샵(KSCSP'96)논문집, 13권, 1호, pp. 344-347, 1996
5. "Netscape's DDE implementation," Version 0x00020002, available from the Internet site <http://developer.netscape.com/docs/manuals/communicator/DDE/abtdde.htm>.
6. H. S. Lee and M. S. Hahn, "Development of a real-time endpoint detection algorithm," *Proc. of ICSPAT'93*, vol. 2, pp. 1547-1553, Sept. 1993.
7. Alexander I. Rudnicky, "Speech interface guidelines," available from the Internet site <http://www.speech.cs.cmu.edu/rspeech-1/air/papers/SpInGuidelines>.
8. 어홍준, 김범수, 한성호, 이영직, 이항섭, "음성인식용 이용한 사용자 인터페이스의 평가 지침," HCI 98 학술대회 논문집, pp. 336-341, 1998.

▲이 항 섭(Hang-Seop Lee)



1990년 2월 : 광운대학교 컴퓨터공학과 졸업(학사)
 1992년 2월 : 광운대학교 컴퓨터공학과 졸업(석사)
 1992년 1월 ~ 현재 : 한국전자통신연구원 음성신호처리팀 선임연구원
 ※주관심분야 : 음성응용시스템, 음성인식, HCI