

논문 2005-42CI-2-3

# VoiceXML을 이용한 Web 연동 학습 시스템 설계 및 구현

## (Design and Implementation of Web Interworking Learning System Using VoiceXML)

김 동 현\*, 조 창 수\*, 신 정 훈\*, 홍 광 석\*

(Dong-Hyun Kim, Chang-Su Cho, Jeong-Hoon Shin, and Kwang-Seok Hong)

### 요 약

멀티미디어 및 통신망 기술의 발전은 학습 시스템 구현 측면에서의 많은 변화를 이룩하였다. 보다 창의적이며, 보다 효율적인 학습 시스템 구축을 위한 연구가 Web 및 전화망을 이용하여 시도 되었다. 그러나 현재까지 구현된 학습 시스템의 경우 웹 또는 전화망을 이용한 단독 시스템으로 각각의 장 단점을 가지고 있다. 즉, Web을 이용한 학습 시스템의 활용을 통한 학습 시 단점으로는 컴퓨터를 활용한 정적인 상태에서의 학습만 가능하며, 컴퓨터를 사용하지 못하는 사람들에게겐 이용하기 위한 새로운 시스템 활용법을 익혀야 하는 문제점이 존재 하였다. 또한, 전화망을 이용한 학습 시스템의 경우 언제 어디서나 전화를 활용학습이 가능하다는 장점을 가지고 있으나, 컴퓨터를 활용한 시스템과 같이 심도있게 효율적으로 정보의 전달이 이루어지지 않는다는 단점을 가지고 있다. 이에 본 논문에서는 VoiceXML을 활용하며 전화망을 경유한 학습 시스템과 Web 기반 학습 시스템의 연동을 통한, 효율적이며 편리한, 언제 어디서나 활용 가능한 학습 시스템을 제안 한다. 또한 전화망을 사용한 학습 시스템에서 사용자 ID, password, 이름 등록 기능의 새로운 알고리즘을 제안하고 전화망과 Web을 사용한 학습시스템에서는 개인학습 진도 저장 기능을 제안 한다.

### Abstract

Development of both multimedia technology and communication network technology has accomplished many changes through the field of learning system. For the construction of a more efficient and clever learning system, there is a research being done by the use of the Web and the telephone network. But until now, the case of current implemented learning system is single system, and so it has each merits and demerits. That is to say, when we use the learning system through the Web, the demerit is only possible by the static states using computer. For those who do not use the computer, the demerit is that the user must learn the use of the new system. Also, the case of using telephone network has merits that one can use the system anyplace, anytime by the telephone. But, it has the problem of not being able to transmit information very efficiently. From these, this paper proposes the learning system that can be used efficiently and conveniently anyplace, anytime by connecting both telephone network and web. Also, we propose a new algorithm of user ID, password and name registration function using learning system using VoiceXML and individual learning progress save function using VoiceXML and web.

**Keywords :** VoiceXML, 학습시스템, 원격학습

### I. 서 론

폭발적인 인터넷의 성장은 많은 사람들에게 정보를 쉽게 다양하게 접할 수 있는 기회를 주었으나 그 이면에는 이용자에게 새로운 시스템의 활용법을 배워야 한다는 부담감을 제공하고 있다. 즉, 원하는 정보를 획득하기 위한 과정이 쉽지 않으며 정보기기와 정보전달 창

구가 늘어가면서 정보처리를 위한 더 많은 사전교육을 필요로 하고 있다. 이러한 문제의 해결을 위해 인간과 컴퓨터의 인터페이스 중 정보전달 측면에서의 신속성과 사용상의 용이성으로 인하여 음성언어를 이용한 사용자 친화 인터페이스 기술 개발의 필요성에 대한 인식이 널리 확산, 연구되고 있다. 이러한 연구 결과로 미국, 유럽 등의 기술선진국에서는 음성인식 기술을 활용한 서비스의 제공이 보편화 되고 있는 실정이다. 정보통신 기기 중 전화는 일상생활에서 누구나 특별한 훈련 없이 쉽게 사용할 수 있는 편리한 정보 전달 수단이며 이러한 편리성은 통신기술의 발전과 함께 전화망 기반의 음

\* 정희원, 정균관대학교 정보통신공학부  
(School of Information and Communication Engineering Sungkyunkwan University)  
접수일자: 2004년11월29일, 수정완료일: 2005년3월8일

성인식 기술을 활용한 서비스 개발을 가속화 하고 있다. 그러나 전화망 기반의 음성서비스 구현을 위해서는 자원 재활용 측면의 문제점에 대한 해결이 필요하였다. 즉, 기존의 방대하게 구축 되어진 웹 기반 자원의 활용에 관한 해결책이 필요하였으며 이의 해결을 위하여 VoiceXML이 개발되었다<sup>[1]</sup>.

인터넷이 HTML 문서의 폭발적인 증가로 활성화되었던 것처럼 음성을 이용한 웹 서비스의 경우도 VoiceXML 문서의 표준화로 활성화 될 것이다. 이에 따라 일부 회사들은 기존의 HTML 문서를 VoiceXML 문서로 변환하는 기술을 개발하고 있다. 마치 웹 서버가 HTML 문서를 처리하듯 VoiceXML 문서를 처리하고 전화망, 음성인식기, 음성합성기와의 인터페이스를 제공하는 VoiceXML 서버 관련 기술은 차세대 음성 포털 사이트의 핵심기술이다<sup>[2][3]</sup>. 이러한 장점으로 인하여 실질적인 응용서비스 구현 시 VoiceXML을 활용하는 사례가 증가하고 있다. 현재까지 구현된 학습 시스템의 경우 웹 또는 전화망을 이용한 단독 시스템으로 각각의 장, 단점을 가지고 있다. Web을 이용한 학습 시스템은 웹 브라우저라는 편리한 사용자 인터페이스 환경에서 기존의 인터넷 서비스를 포함하여 멀티미디어 정보를 취급 할 수 있으며 하이퍼텍스트 형식의 편리한 정보탐색구조로 구성되어 있기 때문에 초보자라도 원하는 정보에 쉽게 접근할 수 장점을 가지고 있으나 단점으로는 컴퓨터를 활용한 정적인 상태에서의 학습만 가능하며, 컴퓨터를 사용하지 못하는 사람들에게는 이용하기 위한 새로운 시스템 활용법을 익혀야 하는 문제점이 존재 하였다<sup>[4]</sup>. 또한, 전화망을 이용한 학습 시스템의 경우 언제 어디서나, 전화를 활용하여 학습이 가능하다는 장점을 가지고 있으나, 컴퓨터를 활용한 시스템과 같이 심도있게 효율적으로 정보의 전달이 이루어지지 않는다는 단점을 가지고 있다.

일반적인VoiceXML기반 학습 시스템의 회원 등록 과정에서 사용자의 이름, 아이디, 패스워드는 엄청나게 많은 수의 인식후보를 필요로 하고, 인식률을 현저히 떨어뜨린다. 그러나 본 논문에서는 이러한 인식률의 저하를 방지하기 위하여 새로운 알고리즘을 제안하였다. 한글은 초성 19개, 중성 21개, 종성 28개에 해당하는 인식 후보 수를 가지고 모든 한글을 조합 할 수 있고, 조합된 한 글자의 한글을 조합하여 이름을 입력 받았다. 아이디와 패스워드는 알파벳 26개와 숫자 10개의 인식 후보 수를 가지고 조합하여 사용자가 원하는 아이디와 패스워드를 사용할 수 있는 새로운 알고리즘을 제

안하였다. 전화를 활용한 로그인 과정에서 사용자의 아이디를 인식하기 위해서는 VoiceXML의 grammar 파일에 사용자의 아이디가 인식후보로 저장 되어 있어야 인식할 수 있다. 기존의 VoiceXML을 이용한 회원 인증 과정에서는 관리자가 직접 grammar 파일에 등록된 사용자의 아이디를 써줘야 하는 단점을 가지고 있다. 본 논문에서는 이러한 단점을 해결하고자 회원 등록 시 자동적으로 grammar 파일에 사용자의 아이디가 저장 되도록 구현했다. 또한 개인별 개인진도 저장 기능을 두어 서비스에 접속하였을 때 이전에 학습한 내용을 알 수 있도록 구성하여 사용자에게 맞춤형 서비스를 제공한다. 마지막으로 온라인 학습의 단점을 보완하기 위하여 웹에서는 Q & A 게시판을 활용하고 전화망에서는 상담원과의 연결이 가능하게 하여 학습자가 학습 내용의 궁금한 사항을 질문할 수 있도록 구현했다.

본 논문의 구성은 다음과 같다. II장에서 VoiceXML의 전반적인 사항에 대해서 알아보고 III장에서는 기존 학습시스템의 구현 유형에 대해서 설명한다. IV장에서는 본 논문에서 제안한 학습 시스템의 구성 환경과 각 모듈의 기능에 대해서 살펴보고, V장에서는 학습 시스템의 평가를 위해 주관적인 평가 방법 중 널리 쓰이는 MOS 방법을 사용하여 학습자의 만족도를 조사하고, 결론 및 향후 연구 방향을 제시한다.

## II. VoiceXML

VoiceXML은 AT&T, IBM, 루슨트 테크놀로지, 모토로라 등 정보통신 분야의4개 기업에 의해 설립된 VXML Forum에서 제안한 대화형 마크업 언어로써, 오디오, 음성인식, DTMF, 음성입력, 전화 등 음성기반의 사용자 인터페이스 어플리케이션 개발을 위해 고안된 XML(eXtensible Markup Language) 문서 형식의 일종이다. 1999년 8월에 "VoiceXML Forum"에서 버전 0.9를 발표한 후, 2000년 3월 보완하여 버전 1.0을 제안하였다. 2000년 5월에 VoiceXML spec 1.0이 발표되었고, 2001년 8월에 VoiceXML 2.0 Working Draft가 발표되었다.

VoiceXML의 장점은 음성서비스 시나리오 저작자를 다양한 문제에서 해방시켜 서비스 내용에만 집중할 수 있게 해 준다. 기존의 음성서비스 구현 과정과는 달리 VoiceXML 문서로 음성서비스 시나리오를 작성한 사람은 음성 입출력의 기술적인 문제에 대해 거의 알 필요가 없다. 또한 시나리오를 통해 전달할 정보의 창출과

정이나 수집한 정보의 처리과정은 웹 서버에 연결된 DB를 통해 가능하므로 대화 시나리오 저작자는 서비스 로직으로부터 독립될 수 있다. 쉬운 저작이 가능하고 복잡한 상황을 표현하기에 별다른 어려움 없는 VoiceXML 문법체계 또한 저작자에게 큰 도움이 된다. 자원할당과 탄압문제도 인터프리터에서 자동 처리되므로 저작자가 전혀 신경 쓸 필요가 없다. 또한, VoiceXML은 콘텐츠 구축에 소요되는 인력을 크게 줄여 음성 포털의 실현을 가능하게 해 준다.

### 1. VoiceXML 게이트웨이

이종망간의 연동을 위해서는 망간 연동을 가능하게 해주는 게이트웨이 장비가 필수적이다. VoiceXML 또한, 전화망과 인터넷망을 연동하고 웹 서버로의 URL 전송하며 웹 서버로부터 수신된 VoiceXML 문서를 분석하고 렌더링 기능을 수행하는 VoiceXML Gateway를 필요로 한다<sup>[5][6]</sup>.

그림 1은 VoiceXML의 구조 모델을 나타내는 그림으로 문서 서버는 웹 서버이고, HTTP 클라이언트 응용 프로그램에 해당하는 VoiceXML 해석기가 VoiceXML 해석기 컨텍스트를 통해서 요청하는 것을 처리한다. 문서 서버는 VoiceXML 해석기의 요청에 따라 적절한 처리 과정을 거쳐서 VoiceXML 문서를 만들고, 그 문서를 HTTP 요청에 의한 응답으로 VoiceXML 해석기에 전달한다. 다음절에서 각각의 기능에 대하여 자세히 설명하도록 하겠다.

#### 가. Document Server

웹 서버를 이용하여 VoiceXML Interpreter가 요청한 문서나 자원을 HTTP 클라이언트 응용 프로그램에 해당하는 URI(Uniform Resource Identifier) 형태를 유지한 채 인터프리터로 전송하는 역할을 한다.

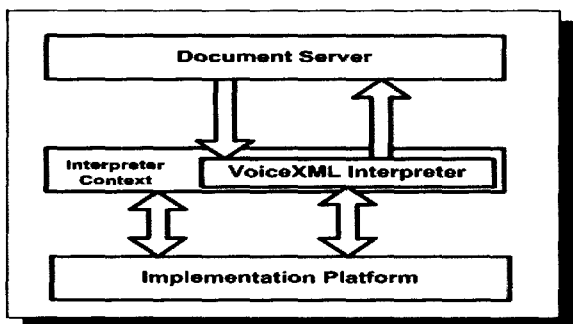


그림 1. VoiceXML의 Architecture 모델  
Fig. 1. Architecture model of VoiceXML.

#### 나. VoiceXML Interpreter

VoiceXML 문서를 적재하고 그 내용을 해석해 실행하는 역할을 담당한다. VoiceXML 문서로 표현된 음성 애플리케이션 시나리오를 해석하여 다이얼로그, 문법, 이벤트, 오디오출력, 콜 제어, 오디오 입력, 흐름 제어와 관련된 47종의 각 태그에 설정된 기능에 따라 문서 실행의 순차적 흐름을 제어하고, 음성 입출력 내용을 결정해 음성플랫폼에 필요한 명령을 내린다. 또한 문서 서버를 통하여 필요한 자원을 다운로드 하거나 다른 문서로 전이하는 등 VoiceXML 문서 실행을 총괄적으로 제어하는 역할을 한다.

#### 다. VoiceXML Interpreter Context

웹 서버와 HTTP를 통해 데이터를 주고받는 HTTP 통신을 담당하며 VoiceXML 해석기가 VoiceXML 문서를 해석하게 한다. 또한, VoiceXML 해석기와는 독립적으로 실행 플랫폼과 상호작용하는 역할을 한다.

#### 라. Implementation Platform

VoiceXML Interpreter Context에 의해 제어되며 하드웨어와 소프트웨어를 모두 포함한다. 주요 기능으로는 전화 수신 기능, 전화호 전환 기능, 음성인식 기능, 음성합성 기능, 음성과 오디오 재생 기능, 음성과 오디오 녹음 기능 등을 들 수 있다.

## III. 기존 학습 시스템

학습 시스템은 원격교육이라는 용어로 가르침의 질을 높이는 것, 교육에의 접근을 용이하게 하는 것, 교육의 비용을 억제하는 것을 목표로 하고 있으며, 1988년 M.Moore는 "원격교육은 의사소통을 매개하기 위해 기술 공학 매체를 필요로 하는 독특한 대화 형식이며, 그 조직은 자율적인 체제이다. 이는 보다 작은 직접적인 대화 그리고 보다 큰 공간적 전이성을 특징으로 하는 모든 교육 프로그램과 학습을 통칭하는 것"이라고 하였다<sup>[7][8]</sup>. 기존의 학습 시스템은 텍스트와 그래픽 중심의 시각적 학습 시스템과 음성 기술을 이용한 음성 학습 시스템으로 크게 나눌 수 있다. 텍스트와 그래픽 중심의 학습 시스템은 텍스트에 그래픽 또는 오디오파일을 추가하여 학습을 진행하는 방식으로 시각적인 면을 충분히 활용하여 사용자에게 학습 능력을 효율적으로 높일 수 있는 장점을 가지고 있으나 시각 장애인 사용자가 불가능하고 컴퓨터와 인터넷이 연결된 환경에서 사

용할 수 있다는 단점을 가지고 있다<sup>[9][10]</sup>. 현재 웹 기반의 학습 시스템은 많은 기업과 연구소에서 개발 상용화하고 있다. 대표적인 예로는 <http://math.kongju.ac.kr/calculus/data/> 이 사이트는 웹에서 텍스트와 간단한 음성 파일을 들음으로써 학습할 수 있는 서비스를 제공한다. 이외에 <http://cyber.edunet4u.net/hanja/> 이 사이트는 텍스트와 그래픽을 사용하여 학습자가 쉽게 학습과 평가 받을 수 있는 서비스를 제공하고 있다. 이 외에도 많은 기업과 대학에서 컴퓨터를 활용한 학습 시스템을 제공한다. 그러나 이러한 웹 기반의 학습 시스템은 학습할 수 있는 공간의 제약을 가지고 있다.

그림 2는 기존 Web기반 학습 시스템의 구성도로 인터넷을 통하여 컴퓨터와 웹서버가 연결된 형태이고 사용자는 컴퓨터를 이용하여 학습시스템에 접속해 학습할 수 있다. VoiceXML을 이용한 학습 시스템은 가장 대중적인 통신 수단인 전화기만 가지고 있으면 컴퓨터와 인터넷이 설치되지 않은 장소에서도 학습할 수 있고 컴퓨터를 사용하지 못하는 사용자와 시각 장애인들도 학습을 할 수 있다는 장점을 가지고 있다. 반면 시각적인 면을 활용하지 못하고 청각적인 면만을 활용하여 학습 효율을 높이지 못하는 단점을 가지고 있다.

그림 3은 기존 VoiceXML기반 학습 시스템의 구성도로 VoiceXML 게이트웨이와 웹 서버가 하나의 서버에

설치된 형태의 구성도를 나타낸다. 또한 로그인 과정을 위하여 관리자가 중간에 개입하여 사용자의 아이디를 VoiceXML grammar 파일에 등록하는 등 관리자가 시스템의 유지에 꼭 필요로 하고 있다.

#### IV. 제안한 학습 시스템

본 논문에서 구현한 학습 시스템은 웹과 전화망을 연동한 학습 시스템으로 사용자는 컴퓨터나 전화를 선택적으로 이용하여 학습할 수 있도록 구현하였다. 시스템의 이용은 회원 등록 후 사용이 가능하고 사용자의 편의를 위하여 개인진도 저장 기능을 두었고 웹과 전화망에서 사용자 DB 는 하나를 사용하여 웹과 전화망이 연동되도록 구성했다. 기존의 VoiceXML 시스템은 사용자 등록과정을 가지고 있지 않았지만 본 논문에서 제안한 학습 시스템은 사용자 정보 등록이 가능하도록 구현하였다.

그림 4는 본 논문에서 구현한 학습 시스템의 구성도로 VoiceXML 게이트웨이와 Web서버는 인터넷을 통하여 연결된 구성도를 보여주고 있다. 그림에서 나타난 바와 같이, 관리자가 추가적으로 필요하지 않고, 사용자 등록을 할 수 있도록 구성하였다.

##### 1 구성 환경

본 논문에서 구현한 음성인식 학습 시스템의 OS는 윈도우 2000 professional을 사용하였으며, 웹서버는 윈도우 NT전용 Web 서버인 IIS(Internet Information Server) 5.0을 사용하였다. 사용자 등록과 로그인 처리 그리고, 개인진도를 저장 기능을 구현하기 위해 ASP(Active Server Page)를 이용하여 시스템을 구현하고 사용자 정보와 개인진도에 관련된 정보를 DB에 저장하기 위해 Microsoft Access를 사용하였다. 전화망 인터페이스를 위해서 Intel Dialogic 41JCT/LS<sup>[11]</sup>를 사용

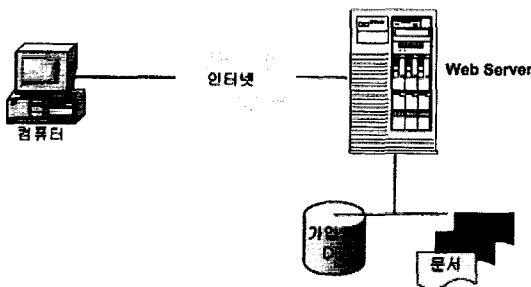


그림 2. Web기반 학습 시스템 구성도  
Fig. 2. Configuration of learning system based on web.

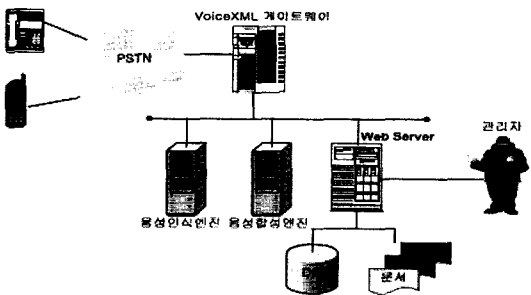


그림 3. VoiceXML기반 학습 시스템 구성도  
Fig. 3. Configuration of learning system based on VoiceXML.

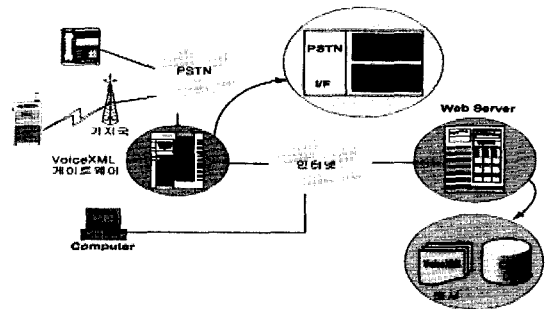


그림 4. 학습 시스템 구성도  
Fig. 4. Configuration of learning system.

하였고 Interpreter는 KT의 HUVOIS 1.0<sup>[12][13]</sup>을 사용하였다.

2. 학습 시스템 메뉴 구성

그림 5는 본 논문에서 구현한 학습 시스템의 메뉴를 보여주고 있다. 메뉴는 웹과 전화망이 동일한 구성 체계를 가지도록 설계 하였다. 각 단원의 학습 내용을 설명해주는 개요설명, 학습을 한 후에도 생소한 용어나 더 자세한 설명이 필요한 용어를 설명해주는 용어 설명이 있다. 또한 학습 내용을 평가하기 위하여 O/X문제와 주관식 문제 두 가지 유형의 문제를 두어 사용자가 자신이 그 단원의 내용을 평가 받을 수 있도록 구현했다.

가. 전체 학습 시스템

그림 6은 본 논문에서 제안하는 학습시스템 서비스의 전체적인 흐름도를 나타낸다. 전체 시스템은 회원등록과 회원 인증 과정, 학습 단원과 학습 범위를 선택하는 과정, 개인진도 저장과 VoiceXML에서 상담원 연결을 하는 기능으로 구성되어 있다. 다음 절에서 각각의 모듈에 세부 기능을 설명한다.

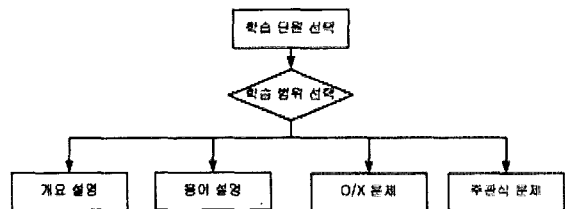


그림 5. 학습 시스템 메뉴  
Fig. 5. Menu of learning system.

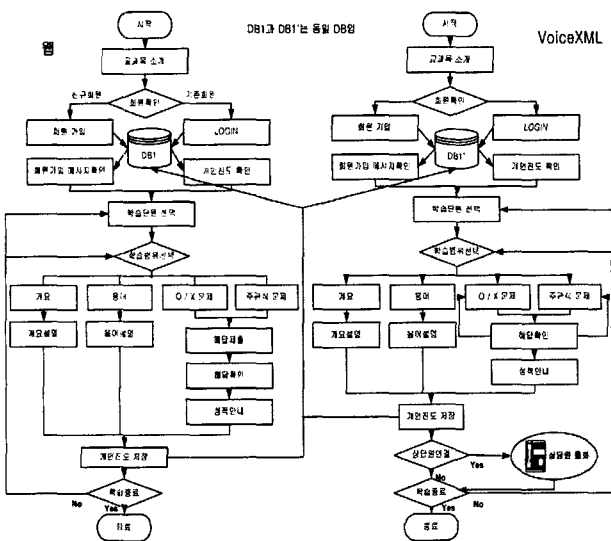


그림 6. 전체 시스템 흐름도  
Fig. 6. Flow of entire system.

나. 회원 등록, 회원 인증 모듈

본 논문에서 VoiceXML상에서 회원등록을 효율적으로 하기 위해 새로운 알고리즘을 제안한다.

그림 7과 8은 웹에서 사용자 등록과 등록결과를 보여주는 화면이다. 사용자 등록 결과에서 개인 정보의 누출 위험성을 대비하여 패스워드는 등록 결과 화면에서 보여주지 않음. 또한 정상적으로 등록을 했는지 확인하기 위해 사용자 등록 과정 후 다시 로그인하는 과정을 가지도록 구성했다.

그림 9는 전화망에서 사용자의 아이디와 패스워드를 등록하는 과정과 사용자 인증 과정의 흐름도를 보여주

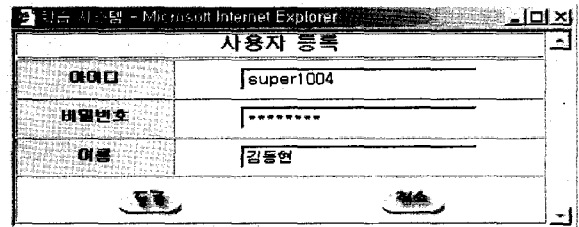


그림 7. 사용자 등록 화면  
Fig. 7. Screen of user register.

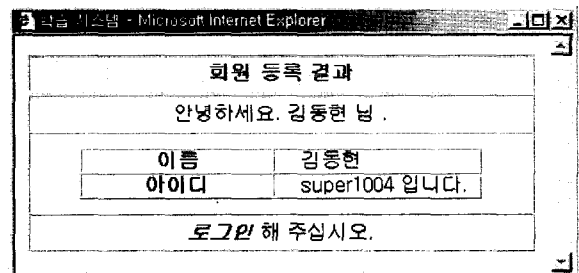


그림 8. 사용자 등록 결과 화면  
Fig. 8. Result screen of user register.

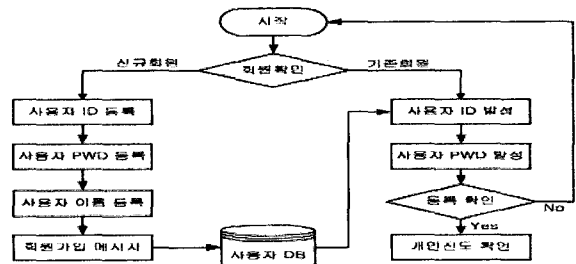


그림 9. 사용자 등록, 사용자 인증 흐름도  
Fig. 9. Flow of user register, User certification.

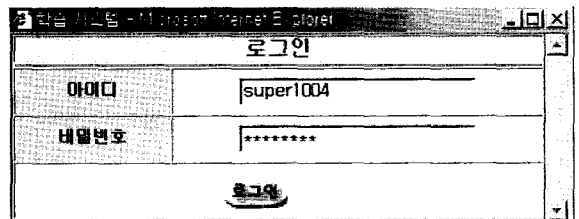


그림 10. 사용자 인증  
Fig. 10. User certification.

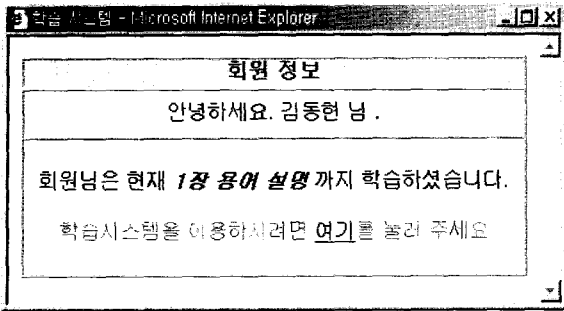


그림 11. 사용자 인증 결과  
Fig. 11. Result of user certification.

고 있다. 그림 9에서처럼 사용자가 전화를 이용하여 학습 시스템 활용을 시도하면 회원에 대한 안내 메시지를 받게 된다. 이때 신규회원일 경우 사용자 등록과정으로 이동되며 기존회원인 경우에는 회원 인증과정으로 이동된다. 사용자의 아이디와 패스워드는 알파벳과 숫자와의 조합으로 등록할 수 있도록 구성하였고 전화기의 특성을 고려하여 알파벳과 숫자는 음성으로 한 글자씩 입력 받는 과정을 반복하여 원하는 아이디와 패스워드를 등록할 수 있도록 했다. 사용자 등록 과정을 완료 후 각각의 사용자의 아이디, 패스워드와 이름은 사용자 DB에 등록된다. 회원에 등록된 학습자가 전화를 이용하여 학습을 할 경우 자신의 아이디를 발성하여 등록된 사용자인지 확인하는 사용자 인증 과정을 거치게 되며 이때 사용자가 발성한 아이디를 인식할 수 있는 grammar파일을 필요로 한다. 따라서 웹과 전화를 사용하여 회원 등록 시 grammar파일에 사용자의 아이디가 자동 저장되어 관리자의 개입 없이 사용자 인증과정을 할 수 있도록 했다. 사용자 등록과정을 거친 사용자와 기존회원은 사용자 인증 과정으로 이동한다. 인증과정은 사용자가 아이디와 패스워드를 발성하면 사용자 DB를 검색하여 아이디와 패스워드가 일치하는지 확인 과정을 가진다. 아이디와 패스워드가 일치하지 않으면 전 단계로 되돌아가 아이디와 패스워드를 다시 발성할 수 있도록 하였고 아이디와 패스워드가 일치하면 "OOO님 학습시스템을 찾아주셔서 대단히 감사합니다."라는 안내 메시지를 받은 후 신규 회원인 경우 다음단계로 이동하여 학습단원을 선택하는 과정을 거치게 되며 기존 회원인 경우 전에 학습한 개인진도 사항을 안내 받는다. 사용자 인증 과정의 마지막 단계로 기존 회원은 저장된 진도 이후부터 혹은 새로운 학습과정을 선택하여 학습할 수 있도록 구현했다.

그림 10과 11은 회원 인증과 인증 결과를 보여주고 있다. 회원 인증 과정도 학습을 위해서 필수적으로 거

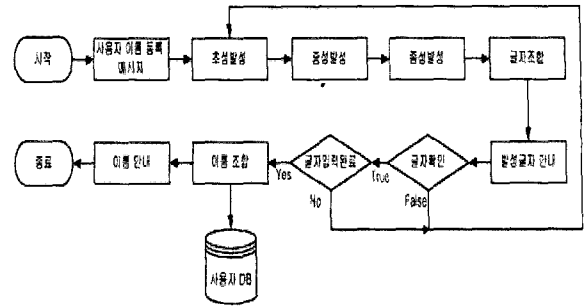


그림 12. 사용자 이름 등록 과정  
Fig. 12. Register process of user name.

쳐야 하는 단계이다. 그림 11 회원 인증 결과에서 회원 확인 메시지를 받고 기존 회원의 경우 이전에 어디까지 학습을 하였는지 회원에게 보여준다. 사용자의 이름 등록 과정은 다음절에서 자세히 설명하도록 하겠다.

#### 다. 이름 등록 모듈

사용자 이름 등록 과정에서 음성 인식 엔진을 활용하여 인식해야 할 음절의 수는 사용자가 발성한 음성의 표시 가능한 수인 완성형 한글 음절의 개수와 같은 2,350개의 인식 후보를 둔 인식을 수행한다. 또한 단어 나 이름단위 인식 시에는 인식 후보의 개수가 무한대의 인식 후보를 가져야 한다. 기존 시스템의 이러한 음성 인식 기능의 수행을 위해서는 음성인식 엔진의 성능 구현이 어려운 상황이며, 구현 후 인식 기능을 수행 시 인식률의 저하에 따른 원인으로 실용적으로 활용하기 어려운 실정이다.

이러한 문제점 해결을 위하여 그림 12와 같은 사용자 이름 등록 알고리즘을 제안한다. 즉, 2,350 음절에 대한 인식기능의 수행에 따른 인식률 저하를 방지하기 위하여 사용자 이름을 예로 들면 다음과 같은 기능을 수행한다. 즉, "홍길동" 이란 사용자가 자신의 이름을 등록하기 위해서는 먼저 인식 음절의 수를 선정 후 각 음절의 초성, 중성, 종성 보유 여부를 확인한다. 본 예에서는 음절의 수는 3음절을 보유하며, 각 음절은 초성, 중성, 종성을 모두 보유하고 있으며 따라서 인식기능의 수행은 초성으로 올 수 있는 자음군 "ㄱ, ㅋ, ㆁ, ㄷ, ㅌ, ㄹ, ㅁ, ㅂ, ㅅ, ㅆ, ㅇ, ㅈ, ㅊ, ㅊ, ㅋ, ㅌ, ㅍ, ㅎ" 등 19개 후보군에 대한 발성으로 인식 기능을 수행한다. 즉, 사용자가 초성으로 "ㄱ"을 입력하기 위해서는 "기억" 이라 발성을, "ㄴ"을 입력하기 위해서는 "니은"을 발성하면 되며 이러한 경우 19개 단어에 대한 음성인식 기능을 수행하면 된다. 중성의 경우 동일한 방법을 적용 시 단모음과 이중모음을 모두 고려 시에도 20개의 단어에 대한 인식기능의 수행으로 가

능하며, 중성의 경우 한글의 7중성 법칙에 따른 7개 단어에 대한 인식기능을 수행하는 것만으로 사용 가능한 모든 한글 자모의 입력이 가능하게 된다. 즉, 초성, 중성, 중성을 모두 합하여도 인식후보의 개수는 46개뿐이 되지 않으며 이는 2,350개의 인식후보와 비교 시 인식률에 있어서 높은 결과를 나타낼 수 있다. 또한, 초성, 중성, 중성의 인식과정을 VoiceXML 기술을 이용하여 각 페이지별 설정을 따로 하게 되면, 인식후보의 수는 더욱더 적어지게 되어 초성의 경우 19개, 중성은 20개, 중성은 7개 단어에 대한 인식기능만 수행하면 된다.

라. 학습 모듈

학습은 각 단원의 학습 내용을 설명해주는 개요설명과 학습을 한 후에도 생소한 용어나 더 자세한 설명이 필요한 용어를 설명해주는 용어설명으로 구성되어 있다.

그림 13은 개요설명의 일부분을 보여주는 그림으로 회원은 자신이 학습하고 싶은 단원의 내용을 선택하여 학습할 수 있고 전화망에서는 개요설명에 대한 내용을 합성기가 수화기를 통하여 회원에게 합성음을 들려주어 회원은 합성음을 듣고 학습이 가능하다.

그림 14는 웹에서 각 단원별 학습 내용에 대한 생소한 용어나 설명이 더 필요한 용어의 목록을 보여주는 그림이다. 회원은 목록을 보고 용어설명을 원하는 용어를 선택하면 그림 15의 결과를 얻는다.

그림 15는 "인간-컴퓨터 상호작용(HCI)"을 회원이 선택하였을 때 용어에 대한 설명을 보여주는 그림이다. 용어에 대한 학습 후 회원은 학습 내용에 대해 저장하기를 원하면 "여기"를 선택하면 회원 DB에 저장할 수 있다. 전화망에서는 용어를 들려준 후 회원은 자신이 학습하고자 하는 용어를 발성하면 해당용어에 대해 합성기의 합성음으로 수화기를 통하여 용어 설명을 학습할 수 있다. 또한 "저장하시겠습니까" 라는 안내 메시지가 나오면 회원의 선택에 의해 저장하거나 저장하지 않을 수 있도록 구성했다.

마. 학습 평가 모듈

학습 과정을 마친 후 학습한 내용에 대해 스스로 평가할 수 있도록 학습 평가 과정을 두었고 학습 평가는 각 단원별 O/X 문제와 주관식 문제의 두 가지 유형의 문제로 구성했다.

그림 16은 전화망에서 해답 확인 및 성적을 확인하는 과정을 나타내는 그림을 나타낸다. O/X 문제와 주관식 문제가 동일하게 그림16과 같은 과정을 가지고 있다.

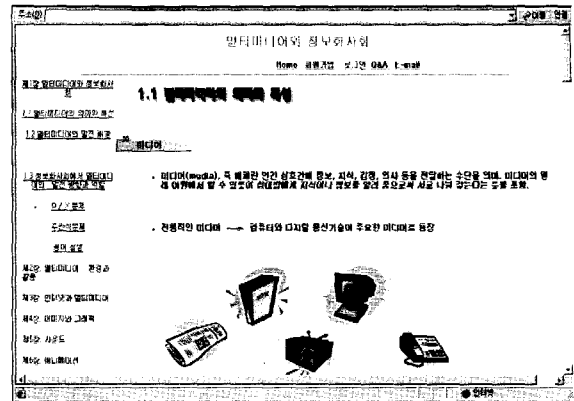


그림 13. 개요 설명의 일부분  
Fig. 13. One part of summary explanation.

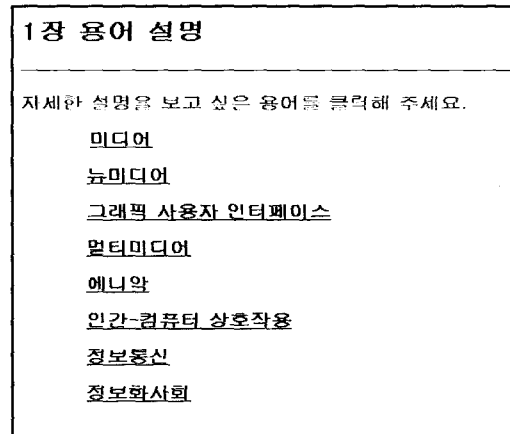


그림 14. 용어 설명  
Fig. 14. Terms explanation.

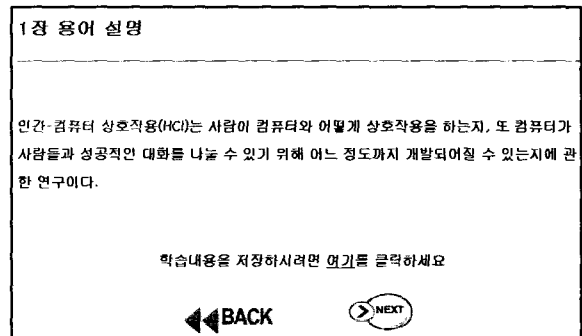


그림 15. 용어 설명 결과  
Fig. 15. Result of terms explanation.

문제는 한 문제씩 출제 되고 사용자는 O/X 문제의 경우 각 문제에 대해 "예", "아니오"라고 대답을 하고 주관식 문제는 문제에 해당하는 용어를 직접 발성하면 정답 DB에서 정답을 확인하고 다음 문제로 넘어가는 과정을 가지고 있다. 마지막으로 모든 문제가 끝나면 해당 문제 유형의 최종 성적을 계산하여 사용자에게 안내한 후 성적 DB에 저장한다.

그림 17은 1장의 O/X문제를 예로써 나타낸다. O/X

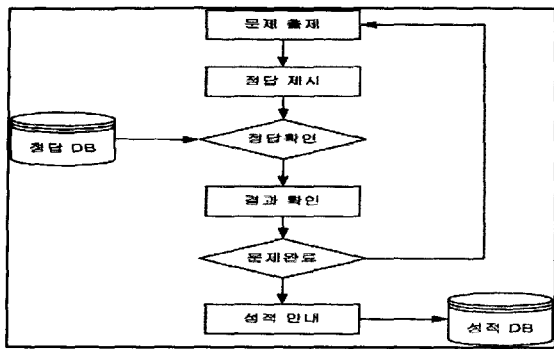


그림 16. 해답 확인 및 성적 확인  
Fig. 16. solution confirmation and score confirmation.

**1장 O/X 문제**

1. 사형자의 목소리를 인식하여 멀티미디어 정보를 검색할 수 있는 기술은 음성형성 기술이다.  
 A 예  
 B 아니요
2. 초고속 정보통신망은 전송선의 전송속도에 비해 만배 정도의 전송속도를 기대할 수 있으므로 분산 멀티미디어 환경에서 핵심 인프라 역할을 담당한다.  
 A 아니요  
 B 예
3. 대체한 인간 상호간에 정보, 지식, 감정, 의사 등을 전달하는 수단이다.  
 A 아니요  
 B 예
4. 멀티미디어의 기존의 대체에 대하여 전자 및 통신 신기술에 의하여 개발된 새로운 대체이다.  
 A 아니요  
 B 예
5. 최초의 컴퓨터라고 할아버지는 컴퓨터는 아니었다.  
 A 아니요  
 B 예

그림 17. O/X 문제  
Fig. 17. O/X quiz.

문제의 특성으로 하나의 해답만 체크하도록 라디오 버튼으로 구성했다. 학습자는 문제를 읽고 정답을 각 문제에 대해 체크를 한 후 해답제출 버튼을 누르면 사용자가 선택한 정답이 결과 페이지로 전송된다.

그림 18은 O/X 문제의 결과와 성적을 출력하는 화면을 나타내었다. 정답은 문제에 대한 정답을 나타내고 제출한 답은 그림 17에서 학습자가 문제에 대한 답을 제출한 것이다. 정답과 비교하여 맞은 개수, 획득 점수, 그리고 패스 여부의 성적 결과를 학습자에게 보여주고 학습 내용을 저장한다.

그림 19는 1장의 주관식 문제의 예를 나타내고 각 문제에 대한 해답은 텍스트박스 안에 직접 학습자가 입력하여 모든 문제를 입력하였을 때 해답제출 버튼을 눌러 주관식 문제 결과 페이지로 해답을 전송한다.

그림 20은 주관식 문제의 결과와 성적을 보여주는 그림이다.

**바. Q & A 모듈**

오프라인 학습에서 학생이 선생님에게 모르는 문제나 학생의 이해가 필요한 부분에 대해 질문하듯이 본 학습 시스템의 웹에서는 Q & A 게시판을 통하여 질문을 하도록 하였고 전화망에서는 VoiceXML의 call

이름		김동현	
경기단원		1장 O/X문제	
평가일시		2004.11.11	
문제 번호	정답	제출한답	정답여부
1	아니요	아니요	O
2	예	예	O
3	예	아니요	X
4	예	예	O
5	아니요	예	X
6	아니요	아니요	O
7	예	예	O
8	아니요	아니요	O
9	예	예	O
10	아니요	예	X

맞은 개수 7/10    획득점수 70    컷라인점수 60    Pass 여부 Pass

학습진도를 저장하시려면 여기를 클릭하세요

그림 18. O/X 문제 결과  
Fig. 18. Result of o/x quiz.

**1장 주관식 문제**

1. 인간 상호간에 정보, 지식, 감정, 의사 등을 전달하는 수단인 의미는 특은 무엇인가?
2. 세계 최초의 전자의 컴퓨터로 인정되고 있는 기계는?
3. 사형기 순자와 문자 위주뿐만 아니라 컴퓨터 연이달을 벗어나 소리, 그림, 그래픽, 문자화상(이미지) 및 동영상(비디오)과 같은 여러 형태의 정보를 컴퓨터를 이용하여 획득, 처리, 통합, 제어 및 표현하는 분야는?
4. 사용자가 그래픽을 통해 컴퓨터와 정보를 교환하는 다른 환경은 어느 분야인가?
5. 멀티미디어 정보용 입력, 처리, 제어하기 위하여 컴퓨터에서 활용하는 기술은?

그림 19. 주관식 문제  
Fig. 19. Subjective quiz.

이름		김동현	
경기단원		1장 주관식 문제	
평가일시		2004.11.11	
문제 번호	정답	제출한답	정답여부
1	미디어	미디어	O
2	에니악	에니악	O
3	멀티미디어	멀티미디어	O
4	GUI	GUI	O
5	HCI	HCI	O
6	뉴미디어	미디어	X
7	정보통신	정보통신	O
8	단원미디어	단원미디어	O
9	그래픽	그래픽	O
10	정보화사회	정보화사회	O

맞은 개수 9/10    획득점수 90    컷라인점수 60    Pass 여부 Pass

학습진도를 저장하시려면 여기를 클릭하세요

그림 20. 주관식 문제 결과  
Fig. 20. Result of subjective quiz.

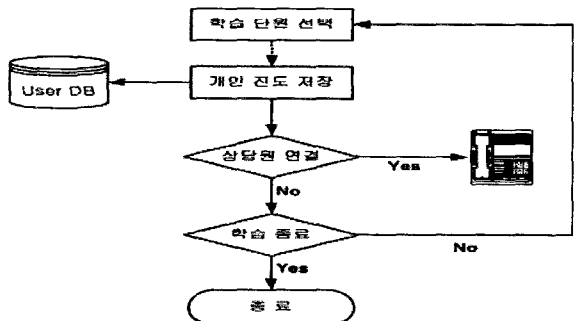


그림 21. 상담원 연결  
Fig. 21. Counselor connection.

transfer 기능을 활용하여 상담원 연결 기능을 구현했다. 이는 선생님과 직접 통화하여 학습의 효율을 높이고 온라인 학습의 단점을 해결하기 위한 목적이다.



그림 21은 전화망에서 개인진도 저장과 상담원과 연결하는 과정의 흐름도를 보여주는 그림이다. 학습자의 편의를 위해 학습 종료 전에 개인진도 사항을 저장하여 다음 번 접속하였을 경우 이전에 학습한 내용을 알 수 있도록 개인진도 저장 기능을 구성했다.

### V. 학습 시스템 평가

본 학습 시스템의 평가하기 위해 성균관대학교 재학생 10명을 대상으로 학습의 만족도를 조사하였다. 본 논문에서는 주관적 평가 방법 중 널리 쓰이는 MOS(Mean Opinion Score)의 방법을 사용했다<sup>[14]</sup>.

표 1은 본 논문에서 제안한 학습 시스템의 평가 기준 방법으로 사용된 MOS의 평가표이다.

표 2는 10명의 학습자가 Web기반, VoiceXML기반, 본 논문에서 제안한 Web과 VoiceXML기반의 학습 시스템을 학습한 후 사용자의 만족도를 나타낸 표이다. 평가 결과 전반적으로 본 논문에서 제안한 웹 연동 VoiceXML 기반 학습 시스템이 다른 시스템 보다 학습자의 만족도가 높은 것으로 나타났다. 이러한 조사 결과는 본 논문에서 제안한 시스템은 Web과 음성서비스의 장점을 모두 활용할 수 있도록 구현 되어져 있기 때문으로 판단되어 진다.

### VI. 결 론

음성은 인간의 가장 기본적인 의사소통 수단으로서, 편리함과 경제성의 측면에서 다른 방법에 비해서 우수한 특성을 가진다. 최근 음성 인식, 음성 합성 기술의 발달과 정보통신 기술의 발전으로 유무선 전화망을 이용하여 컴퓨터와 인터넷을 연결한 음성서비스를 제공하는 업체가 늘어나고 있는 추세이다. 기존의 학습 시스템은 웹과 전화망이 연동되지 않고 각각 독립적인 시스템으로 개발 되어져 왔다.

본 논문에서 제안한 학습시스템은 다수의 클라이언트가 접속하여 학습할 수 있고 텍스트와 이미지 데이터를 지원하는 웹 기반 학습 시스템의 장점과 컴퓨터를 전혀 알지 못하는 학습자뿐만 아니라 시각 장애인도 학습할 수 있고 장소에 구애 받지 않고 언제 어디서나 전화를 활용하여 학습할 수 있는 VoiceXML기반 음성서비스의 장점을 살려 학습자의 학습 효율성을 높일 수 있도록 구현했다. 전화망을 활용한 학습자가 회원 등록 과정에서 한글 한 글자를 인식하기 위해서는 2,350 음

표 1. MOS 평가 기준  
Table 1. valuation standards using MOS.

Rating	1	2	3	4	5
Quality	Unsatisfactory	Poor	Fair	Good	Excellent

표 2. 학습 만족도  
Table 2. Studying satisfaction.

학습자	Web	VoiceXML	Web + VoiceXML
1	4.0	3.6	4.2
2	4.2	3.7	4.3
3	3.9	3.6	4.1
4	4.3	3.8	4.2
5	4.2	3.9	4.1
6	3.7	3.7	4.2
7	3.7	3.9	4.3
8	4.0	3.8	4.3
9	3.9	4	4.2
10	3.8	3.9	4.2
평균	3.97	3.79	4.21

절에 대한 인식기능을 수행 하여야 하고, 아이디와 패스워드는 알파벳과 숫자의 조합 길이에 따라 엄청나게 많은 수의 인식 후보를 가져야 한다. 본 논문에서는 이러한 인식 저하를 방지하기 위하여 새로운 알고리즘을 제안하였다. 한글은 초성 19개, 중성 21개, 종성 28개에 해당하는 인식 후보 수를 가지고 모든 한글을 조합 할 수 있고, 아이디와 패스워드는 알파벳 26개와 숫자 10개의 인식 후보 수를 가지고 사용자가 원하는 아이디와 패스워드를 사용할 수 있는 새로운 알고리즘을 제안했다. 전화를 활용한 로그인 과정에서 사용자의 아이디를 인식하기 위해서는 VoiceXML의 grammar 파일에 사용자의 아이디가 저장 되어 있어야 한다. 회원 등록 후 전화를 활용하여 로그인하기 전에 관리자가 grammar 파일에 등록된 사용자의 아이디를 써줘야 하는 불편한 점을 가지고 있다. 본 논문에서는 이러한 단점을 해결하고자 회원 등록 할 시 자동적으로 grammar 파일에 사용자의 아이디가 저장 되도록 구현하였다. 또한 개인별 개인진도 저장 기능을 두어 서비스에 접속하였을 때 이전에 학습한 내용을 알 수 있도록 구성하여 사용자에게 맞춤형 서비스를 제공한다. 마지막으로 온라인 학습의 단점을 보완하기 위하여 웹에서는 Q & A 게시판을 활용하고 전화망에서는 상담원과의 연결이 가능하게 하여 학습자가 학습 내용의 궁금한 사항을 질문할 수 있도록 구현했다.

본 학습 시스템의 효율성 검증 및 서비스 활용성 검증을 위하여 성균관 대학교 10명의 학생을 대상으로 실험 하였다. 실험 결과 회원 등록 과정에서 한글의 "ㄱ"

과 "ㄱ"의 유사발음으로 인하여 "김"이 "킴"으로 저장되는 경우가 발생하였으나 사용자가 정확한 발음을 한다면 이러한 오류는 거의 발생하지 않았다. 또한 실험자가 학습 후 기존의 학습 시스템과 본 논문에서 제안한 학습 시스템의 만족도를 비교한 결과 본 논문에서 제안한 학습 시스템이 우수한 결과를 나타내었다.

본 논문에서 구현된 학습 시스템을 기반으로 음성 인식기와 음성 합성기의 성능이 높아진다면 사용자의 욕구를 더욱 충족시킬만한 학습 시스템이 될 것이며 상용화도 가능해질 것이라고 판단된다.

### 참 고 문 헌

[1] W3C, "Voice Extensible Markup Language VoiceXML"(Version1.0),<http://www.w3.or/TR/voicexml>, 2000.

[2] P.J, Danielsen "The promise of a voice-enabled Web," Computer vol.3, no.8, pp.104 - 106, 2000.

[3] 홍광석, "음성인터페이스 위한 VXML 기술", 한국인터넷정보학회지, 제2권, 1호, 2001.

[4] 이인숙, 홍기형, "유무선 전화 음성 기반 VoiceXML 학습 평가 시스템," 한국멀티미디어학회지, vol.005, no.004, 2001.

[5] J.A. Quiane Ruiz and J.R. Manjarrez Sanchez, "Design of a VoiceXML gateway," Computer Science, 2003. ENC 2003. Proceedings of the Fourth Mexican International Conference on, vol.49, no.53, pp. 8-12 Sept. 2003.

[6] 박섭형, "VoiceXML: 음성 웹 애플리케이션 구축을 위한", 한빛미디어, pp. 59, 2001.

[7] L.B. Holcomb, S.W. Brown, and F.B. King, "Examining self-efficacy and self-regulation levels across gender in business distance education courses," Computers in Education, 2002. Proceedings. International Conference on, pp.1463 - 1464, 2002.

[8] M. Simonson, "Distance education: trends and redefinition," Frontiers in Education Conference, 1996. FIE '96. 26th Annual Conference., Proceedings of, no.2, pp. 549 - 552, 1996.

[9] D.A. Harris, "Online distance education in the United States," Communications Magazine, IEEE, vol.37, no.3, 1999.

[10] J. Garvin, E. Alsheimer, Z. Bradshaw and P. Anselmo, "Simplified Authoring for Instructors using Distance Education (SAIDE)," Information Technology: Coding and Computing Proceedings. ITCC 2003. International Conference on, pp.47, 2003.

[11] [http://www.computerhq.com/Intel\\_Dialogic\\_D\\_41JCT\\_D41JCTLS\\_99999999999999/network/partinfo-id-18655.html](http://www.computerhq.com/Intel_Dialogic_D_41JCT_D41JCTLS_99999999999999/network/partinfo-id-18655.html)

[12] 구명완, "VoiceXML 기반 음성 응용 시스템," TELECOMMUNICATIONS REVIEW, 제13권, 5호, 2003년 10월

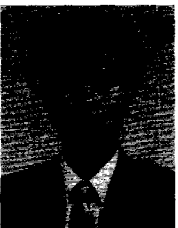
[13] 구명완, 김재인, "음성정보처리기술 응용서비스," 정보처리학회지, 제11권, 2호, 2004년 3월

[14] <http://www.autoeversystems.com/File/skill/VoIP%20%20C5%EB%8AD%C7%B0C1%FA%20%C7%A5C1%D8%B1%E2%BC%FA%20%B5%BF%C7%E2.pdf>

### 저 자 소 개



**김 동 현(정회원)**  
2003년 강원대학교 컴퓨터 정보통신공학부 졸업.  
2005년 성균관대학교 전자전기공학과 석사과정 졸업.  
2005년 ~ 현재 삼성전기.  
<주관심분야 : 음성인식 및 음성합성, VXML>



**신 정 훈(정회원)**  
1992년 성균관대학교 전자공학과 졸업.  
1994년 성균관대학교 대학원 전자공학과 석사과정 졸업.  
2002년 성균관대학교 대학원 전기전자 및 컴퓨터공학과 박사과정 졸업.

1994년 (주)SKC 중앙연구소 연구원  
1995년~2002년 (주)DACOM 종합연구소 주임연구원  
2003년~현재 인덕대학교 정보통신 전공 겸임 교수  
<주관심분야 : 유무선 통신, 음성 인식 및 합성, 오감융합, CTI>



**조 창 수(정회원)**  
2002년 호원대학교 컴퓨터공학과 졸업.  
2004년 성균관대학교 전기전자 및 컴퓨터공학과 석사과정 졸업.  
2004년~현재 성균관대학교 정보통신 기술 연구소.  
<주관심분야 : 음성인식 및 합성, VXML>



**홍 광 석(정회원)**  
1985년 성균관대학교 전자공학과 졸업.  
1988년 성균관대학교 전자공학과 석사과정 졸업.  
1992년 성균관대학교 전자공학과 박사과정 졸업.  
1990년~1993년 서울보건전문대학 전산정보처리과 전임강사

1993년~1995년 제주대학교 정보공학과 전임강사  
1995년~현재 성균관대학교 정보통신공학부 교수  
<주관심분야 : 오감인식, 융합 및 재현, HCI>