

---

# 대화형 음성 인식 항공정보 시스템에서의 ASP 모듈에 관한 연구

윤재석\* · 장준식\*

## A Study On the ASP Module in Conversational Automatic Speech Recognition Flight Information System

Jaeseog Yoon\* · Joonsik Jang\*

### 요 약

본 연구에서는 VoiceXML을 이용하여 전화를 통한 대화형 음성인식 항공정보시스템을 개발하였으며 ASP(Active Server Page)모듈을 작성하여 VXML 코딩을 최적화 하여 효율성을 증대시켰다. 그리고 GSL(Grammar Specific Language)을 사용하여 사람이 말하는 자연어를 어떻게 하면 컴퓨터가 잘 이해할 수 있게 만들 수 있는지에 연구의 중점을 두었으며 이렇게 설계된 시스템을 실험을 통해 그 효과를 검증해 보았다.

### ABSTRACT

In this research, it has been shown that how the computer can recognize and understand spoken natural language and its symbolization using VoiceXML and Grammar Specific Language in developing telephone based conversational automatic speech recognition flight information system. In order for user to hear correct information, ASP Module has been revised and its effectivities has been experimented on the Voice portal airplane information system platform.

### 1. 서 론

음성은 사람들 사이에 가장 기본적인 통신 수단이다. 컴퓨터와의 대화 수단을 보다 인간과 가깝게 하기 위해서는 음성인식(ASR : Automatic Speech Recognition)이 필요하게 되었으며 이러한 음성인식은 매우 다양한 응용분야를 가지며 음성과학과 컴퓨터기술의 발전에 의해 크게 발전되어오고 있다. 컴퓨터형태의 발전 추이가 저용량이면서 휴대 기능으로 변화함에 따라 과거 음성 인식의 주를 이루어왔던 PC 환경 하에서 뿐만 아니라, PDA, 이동 통신 단말기 등의 모바일 환경에서 Web 콘텐츠에 접근 활용할 수 있는 최적화된 음성 인식 기술이 필요하게 되었

다. 음성인식 기술이 실용화 수준으로 발전함에 따라 음성 포털을 비롯한 음성인식 응용 분야가 새로운 이슈로 떠오르고 있다. 이러한 Web 관련 ASR연구가 최근 상당한 연구가 이루어졌으며 어떤 것들은 진행중인 것도 있다[1-5]. 이러한 기존의 마우스나 자판 입력을 대체하는 VUI(Voice User Interface)의 필요성과 이를 가능하게 해줄 새로운 컴퓨터언어인 VoiceXML(Voice eXtensible Markup language)의 필요성이 대두되었다. VoiceXML은 AT&T, IBM, 모토롤라, 루슨트 테크놀로지 등 정보통신 분야에서의 거대 4 기업에 의해 설립된 VXML 포럼에서 제안된 대화형 Markup 언어로서 음성 애플리케이션 개발을 위해 고안된 XML 문서 형식의 일종으로 세계 인터넷

---

\* 대전대학교 컴퓨터공학과

접수일자 : 2002. 6. 18

넷 환경을 주도하고 있는 W3C 컨소시엄에서 VXML 포럼의 제안을 받아들여 표준화된 인터넷과 음성을 바로 연결하여 연동을 할 수 있는 가장 효과적인 언거이다[6-7].

본 연구에서는 VoiceXML의 용이한 확장성을 이용하여 전화를 이용한 대화형 음성인식 항공정보시스템을 개발하였다. 먼저 Call Flow를 작성하고 알고리즘 설계 및 VoiceXML 프로그래밍 했으며 이때 사용자에게 어떤 콘텐츠를 제공하고자 할 때 그 내용이 사용자가 원하는 것에 따라서 상응하는 내용 모두를 제공해야 한다면 VoiceXML로 표현되는 수백, 수천에 이르는 페이지를 작성하여야하는 비효율적인 일이 발생할 것이다. 그래서 본 연구에서는 사용자가 원하는 정보(예를 들면 원하는 날짜의 비행정보)를 들려 줄 수 있게 하기 위한 ASP(Active Server Page) 모듈을 원하는 날짜의 비행 정보를 말해 주는 것을 ASP를 이용하여 사용자가 말한 내용을 데이터베이스에서 검색하여 그 내용을VoiceXML로 변환하여 사용자에게 들려주는 모듈-작성하여 VXML 코딩을 최적화 하여 효율성을 증대시켰다. 그리고 GSL(Grammar Specific Language)을 사용하여 사람이 말하는 자연어를 어떻게 하면 컴퓨터가 잘 이해할 수 있게 만들 수 있는지에 대한 연구에 중점을 두었으며 이렇게 설계된 시스템을 실험을 통해 그 효과를 검증해 보았다.

## II. 전화를 이용한 대화 형 음성인식 시스템의 구축

### 1. 시스템 개요

그림 1은 본 연구에서 사용한 음성 포탈 서비스를 위한 시스템의 개요를 나타낸 것이다. 그림 1에서 VoiceXML Gateway는 VoIP(Voice Over IP), 전화, 모바일 핸드폰 등과 같은 음성 입력을 받을 수 있는 모든 디바이스와 인터페이스를 담당한다.

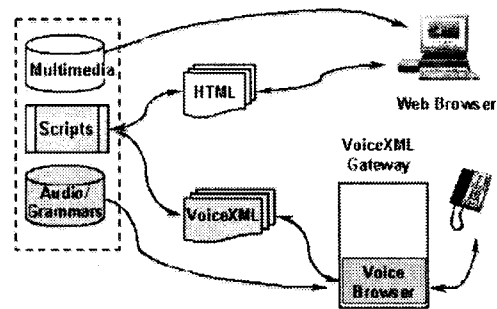


그림 1. 웹 기반 음성 시스템 개요도

웹 서버에는 각종 오디오 파일과 grammar, 스크립트, 멀티미디어 자원 등이 탑재되어 있고 이것을 VoiceXML Gateway를 통해 사용자에게 정보를 제공한다. 이 웹 서버내의 콘텐츠는 VoiceXML형 문서나 HTML형 문서로 변환되어 VoiceXML Gateway나 일반 웹 브라우저로 전송하게 된다. XML과 같은 메타 언어를 사용하게 되면 웹 문서를 원하는 형식으로 변환이 가능하다. 본 연구에서는 이것을 구체적으로 구현하여 가상의 항공사 사이트에서 음성으로 비행 정보를 제공하도록 하였다.

### 2. 플랫폼 환경

본 연구의 실험 환경은 Pentium III 600MHz CPU, 256MB의 윈도우즈 2000 Professional을 웹 서버 및 데이터 베이스 서버로 사용하였다. 데이터베이스는 MS Access 2000을 사용하였으며 AMD Athlon 900MHz CPU와 512MB의 메모리를 탑재한 윈도우즈 2000 advanced 서버시스템을 Voice Web Server로 사용하여 하드웨어 플랫폼을 구축하였다. VWS(Voice Web Server)에는 음성인식엔진으로 Nuance사의 NSR(Nuance Speech Recognition) 7.04 버전을 사용하였고, 웹서버에서 전송되는 VoiceXML 문서의 interpreter로 Nuance VWS 1.2 Beta를 사용하였다. 이것은 그림 1에서 보인 VoiceXML Gateway 역할을 한다. 기타 실험에 사용한 프로그램으로는 자바 실행 환경을 위해 JDK(Java Development Kit)1.3.1버전과 VWS 제어를 위해서 jakarta tomcat 3.2.3 버전을 설치하였다[8].

본 연구의 실험에 사용된 웹 기반 비행 음성정보 시스템의 구축도는 그림 2에 나타났다. 사용자가 음

성 입력이 가능한 입력 디바이스를 이용하여 서비스를 요청하게 되면 Voice Web Server는 음성인식엔진을 통해 요청을 받아들이고 사용자에게 전달할 콘텐츠를 웹 서버에 요청하게 된다. 웹 서버에서는 일반 VXML 문서일 경우에는 바로 Voice Web Server로 전송하게 된다. 본 연구에서는 VXML의 방대한 내용을 ASP 모듈을 이용하여 간결하게 줄였다. ASP 문서를 요청하면 IIS(Internet Information Service)에서 ASP문서를 컴파일 후에 VXML 문서로 변환하여 전송하게 된다. ASP는 Database에 접근하여 요청하는 내용을 검색, 삽입, 삭제, 갱신을 하여 그 내용을 VXML로 변환한다. 사용자가 요청하는 내용에 따라 각각 다른 콘텐츠를 가진 VXML로 전송을 하게 되므로 상황에 따라 다른 콘텐츠를 생성한다[9].

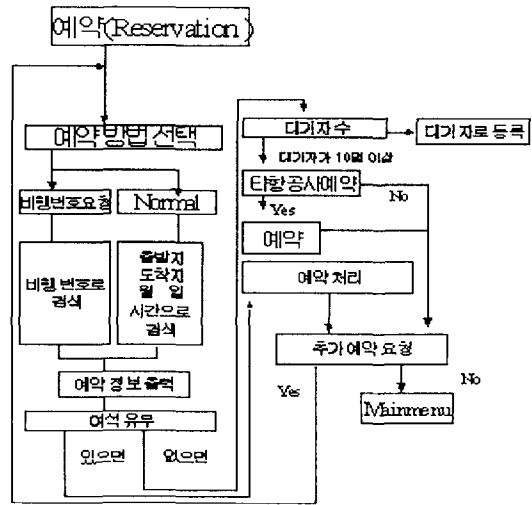


그림 3. 예약부분 Call Flow

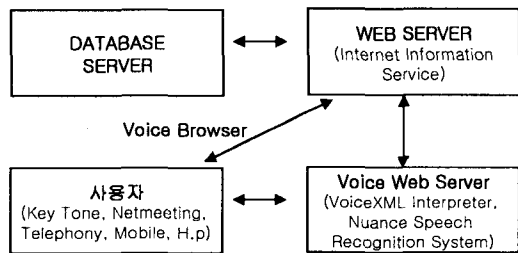


그림 2. 웹 기반 비행 음성 정보 서비스 실험 구축도

### 3. 시스템 설계 및 구현

음성인식 항공정보시스템을 구축하기 위해서 먼저 Call Flow를 작성해야하며 이러한 Call Flow는 시스템 전체의 흐름을 결정하는 중요한 작업이다. 그림 3에는 전체시스템 중 예약부분의 Call Flow를 나타냈다.

이러한 Call Flow의 세부사항인 다이어로그는 grammar와 연동되어 음성인식엔진의 효율과 인식률을 결정하는 중요한 요소 중 하나이다. 본 연구의 음성인식 항공정보시스템은 서비스에 접속한 후 가상의 항공사인 한국항공(가칭)을 대상으로 예약, 운항 스케줄 등과 같은 항공 정보를 제공하는데 그 과정이 대화형식인 다이어로그에 의해 이루어진다. 그림 4는 예약부분의 dialog이다.

한국항공 시스템 : 안녕하세요. 대한항공 안내 시스템입니다. 고객님께서 이용하실 서비스를 말씀해 주시기 바랍니다. 현재 서비스는 예약서비스와 비행정보 도착정보 상담원 연결 및 도움말 서비스를 실시중입니다.  
 예약은 reservation을 비행정보는 Flying information을 도착정보는 arrive를 상담원연결은 sangdamwon을 도움말은 help를 말씀해주시기 바랍니다.  
 사용자 : reservation.  
 한국 항공 시스템 : 회원님의 고유 ID를 말씀해주시기 바랍니다.  
 사용자 : three two one one three.  
 <중략>  
 한국 항공 시스템 : 출발하실 월을 말씀해 주십시오.  
 사용자 : October.  
 한국 항공 시스템 : 출발하실 일을 말씀해 주십시오.  
 사용자 : ten.  
 한국 항공 시스템 : 출발하실 시간을 말씀해 주십시오.  
 사용자 : nine.  
 한국 항공 시스템 : 서울에서 제주까지 운행하는 비행기 로써 10월 10일 9시 15분에 출발하여 10월 10일 10시 30분에 도착하는 비행기입니다. 좌석이 50개 남아 있습니다. 예약시겠습니까?

그림 4. 사용자와 컴퓨터와의 예약을 위한 dialog의 예

### III. ASP(Active Server Page)를 이용한 모듈화

그림 5는 “welcome to service.” 라는 문장을 사용자에게 음성으로 전달하기 위해 VXML으로 작성한 문서이다. 그림 6은 그림 5의 VXML과 같은 기능을 하도록 ASP 스크립트 언어를 사용하여 작성한

것이다. 그림 6과 같이 구현한 프로그래밍이 IIS내에서 컴파일 되면 그림 5와 같은 형태로 변환된다[9].

그림 6에서 ‘<!-- #include file="vxml.asp" -->’ 부분은 본 실험에서 작성한 VXML coding module이 포함된 파일을 포함하는 것이다. vxml.asp란 파일 내에는 위의 ASP 문서를 VXML로 변환하기 위한 함수들이 들어 있다. 원하는 VXML문서를 간편하게 변환하기 위해서는 이 procedure들을 호출하여야 한다. VXML module을 포함시킴으로써 VXML문서를 간단한 procedure 호출로 만들어 낼 수 있다. procedure를 사용하지 않는다면 코딩은 보다 복잡해 질 것이고 이해도 어려워질 것이다. 본 논문에서는 동적인 콘텐츠(비행 정보)를 만들어 낼 수 있도록 ASP모듈을 설계하여 그 효율성을 웹 기반 비행 음성 정보 시스템에서 실험하였다.

```
<?xml version="1.0"?>
<!DOCTYPE vxml PUBLIC "-//Nuance/DTD VoiceXML 1.0b/EN"
'http://community.voxeo.com/vxml/nuancevoicexml.dtd'>
<vxml version="1.0">
  <form id="welcome">
    <block>welcome to service.</block>
  </form>
</vxml>
```

그림 5. VoiceXML의 기본 예

```
<!-- #include file="vxml.asp" -->
<%
Response.ContentType = "text/xml"
call Header
call VxmlStart(NULL, NULL, NULL, NULL)
  call FormStart("welcome", NULL)
  call BlockStart(NULL, NULL, NULL)
  call Prompt("welcom to service", NULL, NULL,
NULL, NULL)
  call BlockEnd
  call FormEnd
call VxmlEnd
%>
```

그림 6. ASP 모듈을 사용한 예

1. ASP 모듈 설계 및 실험

ASP를 이용한 다이얼로그 모듈화는 사용자의 요청에 따라서 전달해야 할 정보가 달라질 경우에 간단한 모듈을 포함하는 프로그래밍과 모듈의 재사용을 가능하게 한다. 이로써 VXML로 voice site

를 구축할 시에 편리성과 개발 시간 단축이라는 효율성을 더하게 된다. 자주 사용되는 다이얼로그를 모듈화하고 이것을 voice site 구축에 이용하게 되면 구축시간과 정보의 표현에서 많은 효율성을 가질 수 있게 된다. 또한 ASP는 일반 웹사이트를 구축하는 방법으로 많이 사용되고 있어 기존의 ASP로 구축한 웹사이트를 그대로 이용할 수 있다.

본 연구에서는 기본적으로 음성인식시스템을 구축하여 항공정보 및 예약 등의 서비스를 제공하게 하였다. 이 시스템의 효율성을 높이기 위해서 ASP로 가상의 항공 정보 사이트를 만들고 음성을 통하여 예약한 내용을 웹 브라우저를 통해서도 예약이나 비행정보를 확인하고 취소할 수 있도록 하였다. 즉 인터넷상에서 예약을 하더라도 전화를 통한 음성인식 시스템으로 예약을 확인할 수 있고 전화로 예약을 하면 웹 site에서 예약을 확인할 수 있게 하였다. 기존의 웹사이트의 음성자료나 데이터베이스 자원을 다시 사용할 수 있도록 하였다.

ASP기본모듈을 확장시켜 비행 정보 콘텐츠를 생성하는 ASP모듈을 설계하고 이를 항공 Voice Portal 음성인식시스템으로 확장·응용하여 전체 시스템에서 구현하는 실험을 하였다. 사용한 비행 정보 데이터베이스에는 비행번호, 출발지, 도착지, 출발 월, 출발 일, 출발 시간, 도착 월, 도착 일, 도착 시간, 여석수 및 대기자 인원의 정보를 저장하도록 설계하였다. 중점을 둔 부분은 기본적인 VoiceXML 태그를 생성하는 VXML 모듈을 응용하여 회원번호를 이용한 회원인증 부문과 숫자를 입력 받아서 숫자의 종류에 따라 각각 다른 음성을 출력하도록 하는 모듈을 설계하고 이를 이용하여 월, 일, 시간, 여석, 대기자수를 각각의 발음을 조사하여 그에 합당한 문장이 생성될 수 있도록 하였다. 여기서 사용된 음성은 일반 사람의 음성을 녹음한 wavefile을 사용하였으며 이를 조합하여 재생하도록 하였으며 영어를 기본 입력어로 사용하였다.

회원인증을 통하여 회원에게만 예약 서비스를 제공하도록 하였는데, 여기서 사용되는 회원번호는 회원이 인터넷을 통하여 항공사에 등록하고 발급 받은 번호라고 가정한다. 5자리의 회원번호를 발급 받으면 이 번호가 전화 서비스에서의 회원 ID로

사용하도록 하였고 비밀번호는 웹에서 신청한 것과 같은 번호를 사용하도록 하였다. 사용자가 회원 번호를 "one two three four five"라고 말하면 음성 인식 엔진이 이를 인식하고 그 결과를 "12345"라는 문자열로 변환하도록 Grammar를 작성하였다. 이렇게 반환된 내용은 ASP를 통하여 데이터베이스에 질의를 하게 되고 위의 번호가 존재하는지를 검색한다. 번호가 없으면 다시 번호를 요구하고 3번의 시도 후에도 검색 내용이 없는 경우에는 서비스를 마치도록 하였다. 회원번호가 있으면 비밀번호를 요구하게 하였고 비밀번호도 3번의 기회만이 주어지도록 하였다. 이와 같은 회원인증 절차를 거쳐야만 예약 및 예약 확인이 가능하도록 설계하였다.

비행정보 중에서 출발 일과 출발 시간 등의 정보는 월, 일, 시간, 분일 경우에 따라 각각 다른 발음을 해야 한다. 예를 들어 1의 경우 시간은 한시라고 하지만 분은 일분이라고 발음하게 된다. 이런 경우를 조사하고 각각의 경우에 맞는 음성을 녹음한 후에 각 경우에 따라서 각기 다른 음성 파일을 링크하도록 설계하였다. 이외에도 사용자에게 음성 정보를 전달하기 위해서 각 단어나 문장을 웨이브 파일로 녹음을 한 후에 이를 조합하여 하나의 완성된 문장으로 출력이 가능하도록 하였다. 본 연구에서는 두 자리까지의 숫자 입력을 대상으로 하였다. 그림 9의 VoiceXML 출력 결과를 보면 월, 일, 시간, 분, 여석수, 대기자 수가 각각의 다른 경로의 음성 파일을 링크하고 있음을 볼 수 있다. 입력이 시간일 경우에는 오전, 오후를 포함하도록 하였다.

그림 7은 ASP모듈 중에서 FlyingInfoPlay라는 procedure로써 비행 정보를 생성하고, 사용자에게 예약 여부를 물어보거나 예약이 힘들 경우 다른 항공사의 비행 예약을 권유하는데 필요한 정보를 세션이 변수에 저장하도록 작성한 ASP 프로그래밍이다. 이 FlyingInfoPlay procedure는 예약 및 예약확인, 비행정보 등에서 공통적으로 사용이 가능하다. 이 procedure는 검색된 비행 정보에 따라서 예약이나 예약 확인, 비행정보 조회를 할 수 있도록 구현하였다. 사용자가 말하는 내용은 서비스 형태, 비행번호, 출발지, 도착지, 출발월, 출발일, 출발시간, 조회기간으로 분류되어 이 procedure의 인

수로 입력된다.

```

Sub FlyingInfoPlay(service_type, flying_num, origin, destination,
d_month, d_day, d_hour, scope)
'중략
'---- 출발지(origin), 도착지(destination), 출발월(d_month), 출발일
(d_day)
'---- 출발시간(d_hour) 각각의 입력값이 있으면 그것에 맞게 쿼리
문에 추가 된다.
'----- database 연결 설정 및 레코드셋 설정 생략 -----
'검색 결과가 없으면
IF rs.EOF or rs.BOF THEN
Call PromptStart(NULL,NULL,NULL)
Call
AudioStart("../prompts/word/flying_infor_no.wav",NULL,NULL,N
ULL,NULL,NULL) '출발지
Call AudioEnd
Call PromptEnd
'결과가 있을 경우
ELSE
Do Until rs.EOF
Session("flying_num") = rs("flying_num") '비행 번호를 세션
변수에 할당
If rs("waiting_list") < 10 Then Session("reservation")="yes"
Else session("reservation")="no"
End If
'중략
'비행 번호가 있을 경우 비행일자, 여석, 대기자 정보 출력
'----- 비행 정보를 말하는 prompts 시작 -----
Call PromptStart(NULL,NULL,NULL,NULL)
Call AudioStart(origin_url,NULL,NULL,NULL,NULL) '출
발지
Call AudioEnd
Call AudioStart(from_url,NULL,NULL,NULL,NULL) '에서
Call AudioEnd
'중략
Call NumeralPlay(rs("arrive_month"),../prompts/"month") '월
Call NumeralPlay(rs("arrive_day"),../prompts/"day") '일
Call NumeralPlay(rs("arrive_hour"),../prompts/"hour") '시간
Call NumeralPlay(rs("arrive_minute"),../prompts/"minute-1")
'분에 도착하는 비행기입니다.
'----- 여기서부터는 여석 및 대기 순서를 말해 준다. -----
IF rs("seat") = 0 Then '여석이 없으면 여석이 존재하지 않는다
고 한다.
seat_url = "../prompts/seat/seat_empty.wav"
Call AudioStart(seat_url,NULL,NULL,NULL,NULL)
Call AudioEnd
'---- 여석이 없더라도 대기자 순번이 10번을 넘지 않았을 때 ---
IF rs("waiting_list") < 10 Then
C a l l
NumeralPlay(rs("waiting_list"),../prompts/"waiting_list") '대기
자 순서
Else '----- 대기자 순번이 10명이 넘었을때. -----
'--- 대기자가 너무 많습니다. 예약이 불가능합니다. ---
Call
AudioStart("../prompts/waiting_list/waiting_x.wav",NULL,NULL,
NULL,NULL,NULL)
Call AudioEnd
End If
Else '----- 여석이 남아 있을 경우 -----
Call NumeralPlay(rs("seat"),../prompts/"seat") '여석의 갯수
End If
Call PromptEnd
rs.MoveNext
Loop
'중략
End Sub
    
```

그림 7. 비행 정보 생성 모듈

FlyingInfoPlay procedure는 서비스의 형태(비행정보, 예약, 예약확인 등)나 비행번호의 유무, 조회할 기간에 따라 각각의 다른 VXML 콘텐츠를 생성한다. 사용자가 비행번호를 알고 있을 시에는 이 고유의 비행 번호를 입력으로 받아서 데이터베이스에서 검색하여 그 내용을 가져와서 비행정보를 형성하도록 하였다. 그러나 비행번호를 모를 경우에는 출발지나 도착지, 출발 월·일·시간 등을 입력으로 받아서 데이터베이스에서 검색하도록 하였다. 검색 기간은 당일이나 주 단위 중에 선택을 할 수 있도록 하고, 그 기간 동안의 비행 정보를 생성하게 한다. 이 모듈을 그림 8과 같이 응용하여 비행 예약을 구현하는데 사용하였다.

그림 8은 그림 7에서 추출된 비행정보를 이용하여 사용자에게 현재 들려주는 비행 정보가 자신이 원하는 예약일 경우에 예약을 할 수 있도록 하였다. 여석이 없을 경우에는 대기자로 등록 가능하도록 하고 대기자로 등록이 불가능할 경우에는 다른 가상의 항공사의 비행정보를 검색하여 예약을 할 수 있도록 하였다. 입력한 내용에 맞는 비행편이 없을 경우에는 재입력이나 다른 서비스를 이용할 수 있도록 하였다. 각 예약 처리여부 상태를 세션 변수에 저장한 후에 이 세션 변수의 값에 따라 각각의 다른 VXML 문서가 생성 되도록 하였다.

그림 9는 그림 8의 소스가 IIS에서 컴파일 된 후에 생성되는 VXML문서중 하나이다. 이 문서 내용은 "서울에서 부산까지 운행하는 비행기로서 9월 11일 오후 6 시 30 분에 출발하여 9월 11일 오후 7시 41분에 도착하는 비행기입니다. 41개의 좌석이 있습니다. 예약하시겠습니까?"라고 사용자에게 문의하는 것이다. 사용자가 원하면 예약 내용을 데이터베이스에 저장하는 ASP문서로 비행 번호와 사용자 아이디를 넘겨준다. 여기서는 1011이란 비행번호가 넘겨지고 사용자 아이디는 서버내의 세션 변수에 저장된 값이 넘어가게 된다.

```
<!--#include file="./vxml.asp"-->
<%
Call Header
Call VxmlStart("./main.vxml",NULL,NULL,NULL)
Call FormStart("PlayInfo", NULL)
Call BlockStart("block1", NULL, NULL)
Call FlyingInfoPlay("reservation_other", Null, request("origin"),
request("destination"), request("month"),
request("day"), request("hour"), request("to_we"))
If Session("reservation") = "no" then '예약할 수 없을 경우 타
항공사로의 예약을 권유
생략
Call IfStart("reservationAsianaYesNo='yes'") '타항공사의 예
약을 원할 경우
Call Assign("flying_num", Session("flying_num"))
'사용자 아이디와 비행번호를 예약을 처리할 페이지에 전달
Call
Submit("http://www.voice-xml.net/asiana/main/vxml/reservation_p
rocess.asp?id="
&
Session("id"),NULL,"flying_num","get",NULL,NULL,NULL,NULL,"a
pplication/x-www-form-urlencoded")
Call VxmlElse '다른 항공사의 예약을 원하지 않을 경우 메뉴
로 돌아감
Call VxmlGoto("./member_check.vxml#service_choice", NULL,
NULL, NULL, NULL, NULL, NULL, NULL, NULL)
Call IfEnd
Call FilledEnd
Call FieldEnd
Elseif Session("reservation") = "yes" then '예약이 가능한 경우
Call BlockEnd
Call FieldStart("reservationYesNo", NULL, NULL, NULL,
NULL, NULL)
Call
AudioStart("./prompts/reservation/reservationYesNo.wav",NULL,
NULL,NULL,NULL,NULL)
Call AudioEnd
Call Grammar("./yes_no.grammar#YES_NO", NULL, NULL,
NULL, NULL)
Call NoMatchStart(NULL, NULL)
Call
AudioStart("./prompts/error/nomatch.wav",NULL,NULL,NULL,N
ULL,NULL)
Call AudioEnd
Call NoMatchEnd
Call FilledStart(NULL, NULL)
Call IfStart("reservationYesNo=='yes'")
Call Assign("flying_num", Session("flying_num"))
'사용자 아이디와 비행번호를 예약을 처리할 페이지에 전달
Call Submit("reservation_process.asp?id=" &
Session("id"),NULL, "flying_num", "get", NULL, NULL, NULL,
NULL, "application/x-www-form-urleenco
ded")
생략
%>
```

그림 8. 비행 정보 생성 모듈을 이용한 비행 예약

그림 9에서 <audio>태그들이 연결된 것에서 볼 수 있듯이 월, 일, 시, 분, 좌석 수 등은 그 종류에 따라 다른 경로의 음성 파일을 연결하고 있음을 알 수 있다. 이런 방법으로 음성을 조합하여 마치 하나의 문장으로 출력되는 것처럼 들리도록 설계하였다.

```

<?xml version="1.0" encoding="TUD-8" >
<!DOCTYPE vxml PUBLIC "-//Nuance/DTD VoiceXML 1.0b/EN"
"http://community.voxeo.com/vxml/nuancevoicexml.dtd" >
<vxml application="./main.vxml" version="1.0" >
<form id="PlayInFo">
<block name="block1">
<prompt>
<audio src="/prompts/city/seoul.wav"></audio>
<audio src="/prompts/word/from.wav"></audio>
<audio src="/prompts/city/busan.wav"></audio>
<audio src="/prompts/word/to.wav"></audio>
<audio src="./prompts/month/9.wav"></audio>
<audio src="./prompts/month/month.wav"></audio>
<audio src="./prompts/day/10.wav"></audio>
<audio src="./prompts/day/1.wav"></audio>
<audio src="./prompts/day/day.wav"></audio>
<audio src="./prompts/hour/pm.wav"></audio>
<audio src="./prompts/hour/6.wav"></audio>
<audio src="./prompts/hour/hour.wav"></audio>
<audio src="./prompts/minute/30.wav"></audio>
<audio src="./prompts/minute/minute.wav"></audio>
<audio src="./prompts/month/9.wav"></audio>
<audio src="./prompts/month/month.wav"></audio>
<audio src="./prompts/day/10.wav"></audio>
<audio src="./prompts/day/1.wav"></audio>
<audio src="./prompts/day/day.wav"></audio>
<audio src="./prompts/hour/pm.wav"></audio>
<audio src="./prompts/hour/7.wav"></audio>
<audio src="./prompts/hour/hour.wav"></audio>
<audio src="./prompts/minute-1/40.wav"></audio>
<audio src="./prompts/minute-1/1.wav"></audio>
<audio src="./prompts/minute-1/minute-1.wav"></audio>
<audio src="./prompts/seat/40.wav"></audio>
<audio src="./prompts/seat/1.wav"></audio>
<audio src="./prompts/seat/seat.wav"></audio>
</prompt>
</block>
<field name="reservationYesNo">
<audio src="./prompts/reservation/reservationYesNo.wav">
</audio>
<grammar src="./yes_no.grammar#YES_NO"/>
<nomatch>
<audio src="./prompts/error/nomatch.wav"></audio>
</nomatch>
</filled>
<if cond="reservationYesNo=='yes'">
<assign name="flying_num" expr="1011"/>
<submit next="/reservation_process.asp?id=" method="get"
namelist="flying_num"
enctype="application/x-www-form-urlencoded"/>
<else/>
<goto next="/reservation_number.vxml"/>
</if>
</filled>
</field>
</form>
</vxml>
    
```

그림 9. 조회된 비행편에 대한 예약 여부를 결정하는 VXML

전화로 예약을 하게 되면 그 내용은 일반 웹 브라우저로도 확인할 수 있도록 하고 웹에서 예약한 것을 전화로도 확인이 가능하도록 설계하였다. 그림 10·11의 과정으로 예약을 한 결과와 전화로 예약을

한 결과는 모두 하나의 데이터베이스 시스템을 사용하게 되므로 웹이나 전화 등으로 예약하고 취소, 비행정보확인 등은 같은 정보가 전달되는 것이다. 사용자는 비행 예약을 위해 여러 가지 방법 중에서 원하는 것을 선택할 수 있게 된다.

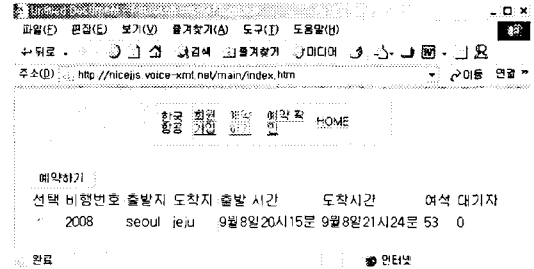


그림 10. 일반 웹 브라우저를 통한 비행 예약

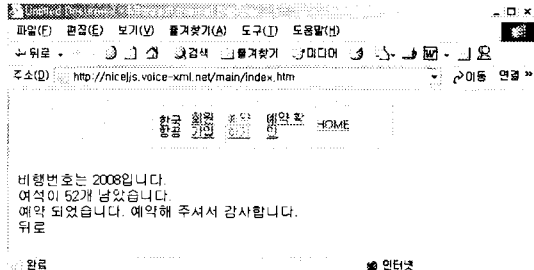


그림 11. 웹을 통한 예약 처리

2. GSL(Grammar Specification Language)

Grammar는 사용자가 요청하고자 하는 것이나 제공할 정보를 문자열로 반환하게 하게 하는 것으로 VXML 표준안에는 그 포맷을 명세하거나 특별한 포맷의 지원을 필요로 하도록 하지 않았다. 따라서 이것은 음성 인식 엔진 자체적으로 포맷을 가지며 반환한 값만을 VXML에서 처리한다. 본 실험에서는 이 grammar 포맷중의 하나인 Nuance사에서 제공하는 GSL(Grammar Specification Language)을 사용하였다.

GSL은 VUI(Voice User Interface)를 구현하는데 아주 중요한 역할을 하게 된다. 사용자는 단지 “예약”이라고 하는 사용자도 있을 것이고 “예약하고 싶어요”, “비행예약” 등과 같이 말할 수도 있다. 이 중에서 예약이란 단어는 꼭 들어가게 된다. 다시 말해 사용자가 수행하고자 하는 것이나 입력하는 내용을

VXML에 전달하게 되고 VWS는 이것을 기반으로 처리한다. 이것은 일반 웹 프로그래머들이 회원 가입·인증 형식의 텍스트 필드에 값을 넣는 것과 비슷한 원리이다. 입력 값으로 음성과 문자열 중에서 어느 것을 선택하느냐의 차이라고 할 수 있다. 음성으로 입력된 값이 VXML의 field 태그의 속성인 name에 지정된 변수에 저장된다. 이것과 비슷하게 html에서는 이것을 `<input type="text" name="memberid">`로 표현할 수 있다. 이 필드의 값이 모여 폼을 형성한다. grammar는 대부분 비슷한 형식을 취하고 있는데, 그림 12는 본 실험에서 GSL을 사용하여 작성한 한 예를 나타낸 것이다. 이것은 예약 체크나 취소 예약 방법을 택하게 하는 grammar로써 사용자가 말한 내용에 따라 각각 다른 출력을 하도록 작성한 것이다. 비행 번호로 예약을 원하면 출력을 number로 시간이나 지역으로 예약을 원하면 nomal을 출력하고, 확인은 check, 취소는 cancel을 반환하도록 하였다.

```

Reservation [
  (? ( i want to ) ?flying number) <choice number>
  [
    (? ( i want to ) area) (? ( i want to ) time) (? ( i
  want to ) nomal)
  ]<choice nomal>
  (? ( i want to ) ?reservation check) <choice
  check>
  (? ( i want to ) ?reservation cancel) <choice
  cancel>
  ]
  ]
    
```

그림 12. 예약 방법 및 확인 체크를 위한 GSL

#### IV. 결 론

본 연구는 VoiceXML이란 표준화된 마크업 언어를 사용하여 웹 기반의 항공정보를 제공하는 음성 인식 시스템을 구축하였다. 전화나 기타 음성 입력이 가능한 디바이스를 사용하여 음성포탈기능인 항공정보 및 예약, 예약 확인 등의 기능 수행을 확인하였으며 또한 기존의 웹과의 browsing이 가능함을 보였다. 특히 기존의 웹 프로그래밍 언어의 하나인 ASP를 사용하여 모듈을 설계함으로써 좀 더 복잡한 구조를 가진 콘텐츠를 생성을 가능하게 하였으며 또한 이를 모듈 화하여 필요시 재사용을 가능하게 하여

시스템의 효율성을 증가시킬 수 있다는 사실을 확인하였다.

이제 웹은 더 이상 눈으로 보여지는 것 뿐만 아니라 음성으로 콘텐츠를 전달하는 것이 가능하다. VXML과 ASP와 같은 기존의 프로그래밍 언어를 사용하면 동적 콘텐츠를 생성 가능하며 이로써 다양한 서비스를 구현할 수 있음을 알 수 있었다. 더불어 VXML이 XML을 기반으로 만들어진 마크업 언어인 만큼 얼마든지 확장 및 변환이 가능하다. 이러한 사실을 바탕으로 XML 기반의 웹 사이트의 콘텐츠를 음성 정보 서비스의 콘텐츠로 변환이 가능하다면 웹 사이트의 내용을 바로 음성으로 browsing하여 그 콘텐츠를 음성으로도 전달가능하게 된다. 본 연구에서는 이러한 연구의 방향을 제시하였으며 향후의 과제로 그 가능성을 열었다하겠다.

#### 참 고 문 헌

- [1] K. Kondo and C.Hemphill. A WWW browser using speech recognition and its evaluation. Systems and Computers in Japan, pages 57-67, Sep. 1998.
- [2] V. Digalakis, L. Neumeyer, and M. Perakakis. Quantization of cepstral parameters for speech recognition over the World Wide Web. IEEE Journal on Sel Areas in Communications, pages 82-90, Jan. 1999.
- [3] Z. Tu and P. Loizou. Speech recognition over the Internet using Java. IEEE International Confrence on Acoustics, Speech, and Signal Processing, Phoeix, AZ: pages 2367-70, Mar. 1999
- [4] C.T.Hemphill and Y.K.Muthusamy. Developing Web-based speech applications. European Conference on Speech Communication and Technology, pages 895-898, Sep. 1997.
- [5] D. Vaufreydaz, J. Rouillard, and M. Akbar. A network architecture for building



applications that use speech recognition and/or synthesis. European Conference on Speech Communiton and Technology, pages 2159-62, Sep. 1999.

- [6] <http://www.w3.org/voice>
- [7] <http://www.voicexml.org>.
- [8] <http://www.nuance.com>
- [9] <http://community.voxeo.com>

### 저 자 소 개



윤재석(Jae Seog, Yoon)

1982년 경북대학교 전자공학과 공학사

1985년 연세대학교 대학원 전자공학과 석사

1991년 Washington Univ. 전기 및 컴퓨터공학과 M.S.

1997년 연세대학교 대학원 전자공학과 공학박사

1998년~현재 대전대학교 컴퓨터공학과 조교수 재직중

※관심분야 : 음성인식 멀티미디어 시스템, e-비즈니스, 디지털반도체회로



장준식(Joon Sik, Jang)

2002년 대전대학교 컴퓨터공학과 졸업 예정

※관심분야 : 음성인식 멀티미디어 시스템, XML 기반 응용 시스템