

## 이어콘을 적용한 음성 메뉴의 사용성 평가에 관한 연구

임치환\* · 이재인\*\* · 이성수\*\*\*

\* 서원대학교 경영학부

\*\* 한국과학기술원 산업공학과

\*\*\* GS 홈쇼핑

## A Study on the Usability Evaluation of Earcon Applied to Voice Menu

Chee-Hwan Lim\* · Jae-In Lee\*\* · Sung-Soo Lee\*\*\*

\* Management Information Systems, School of Business Administration, Seowon University

\*\* Dept. of Industrial Engineering, KAIST

\*\*\* GS home shopping

This paper describes the experiment that investigated the possibility of design and evaluation of the usability of earcon applied to voice menu. The earcon has functions in providing navigational cue in the hierarchical menu, and also can be applied to voice menu for improving the recall rate and the response time. In this experiment, participants identified their location with the help of earcon applied to the voice menu with the various earcon parameters. In detail, some participants listened to the good quality sound of voice menu using the structured earcon, and they recalled the location they heard. Other participants listened to the good quality sound of voice menu without earcon, and they recalled the location they heard in the same manner. And the response times were checked through their answers. On analyzing the results, we found the earcon applied to voice menu showed the increase of recalling rate from what they heard during experiment. That is, the performance of task was better with earcon applied to voice menu than with voice menu without earcon. In the earcon applied to voice menu test, it showed the accuracy of 92.50%, but in voice menu without earcon test, people could only recall 66.25% among given questions. The response time was reduced from 4.98 sec to 3.85 sec. In addition, this experiment showed the 87.5% of participants preferred the earcon applied to voice menu.

**Keywords :** Earcon, Voice Menu, Usability, Evaluation

### 1. 서 론

언어는 인간에게 말과 글의 두 가지 형태로 병행하여 이용되어 왔다. 이중 말에 해당하는 음성은 인간 상호간의 의사소통 수단으로서 커다란 역할을 해왔으며, 최근에는 음성인식 및 음성합성 기술의 발전으로 인간과 컴퓨터와의 대화식 상호작용이 가능해졌다. 또한 Dictation, Reader, Command System, Voice Portal, Voice Commerce

등 다양한 분야에서 응용 발전되어 왔다[4]. 이러한 대화형 음성 처리 기술을 바탕으로 ARS (Automatic Response System) 서비스 분야에서는 정보를 음성 메뉴 형태로 제공한다. 하지만 시각적 디스플레이(Visual Display)의 메뉴와 달리 음성 메뉴에서는 짧은 시간을 청취한 후 메뉴 항목을 선택하게 되므로 탐색(Navigation)할 때 자신이 선택한 메뉴의 깊이(Depth)와 폭(Breadth)을 파악하기 어려워 다음 선택할 메뉴가 어디쯤 있는지 혼돈이 생기

기 쉽다. 이러한 문제점을 해결하기 위한 방법으로 이어콘(Earcon)이 제시되고 있다.

이어콘(Earcon)은 본래 컴퓨터의 Graphical User Interface (GUI)상의 아이콘(Icon)이나 메뉴의 보조 사운드로 사용되었고, 소리 형태로 정보를 전달하는 효과적인 수단으로 연구되어 왔다[3]. 그리고 Leplatre and Brewster (1998)는 이어콘을 시각적 디스플레이에서 메뉴 계층(Menu Hierarchy)에 대해 탐색 큐(Navigational Cue)를 제공하려는 시도를 하였다[5]. 이러한 연구들을 기반으로, 음성 메뉴가 가지는 정보 표현 영역을 넓히고 탐색에 따라 선택된 메뉴의 구조와 위치를 쉽게 파악하도록 하기 위해서 청각적 큐(Auditory Cue)를 제공하는 개선된 이어콘을 음성 메뉴에 적용하는 연구가 필요하다. 또한 이어콘이 적용된 합성 음성(Synthesized Speech) 메뉴에 대한 사용성 평가를 통하여 현실적으로 효과적인지도 확인이 필요하다.

본 연구에서는 시각적 디스플레이의 보조적 음향으로 사용되던 음악적 모티브(Motive)인 이어콘에 관한 연구와 음성 합성 기술로 만들어진 합성 음성에 관한 연구를 근간으로 음성 메뉴에 이어콘의 적용과 그것에 대한 사용성을 평가하였다. 즉, 이어콘을 적용한 음성 메뉴와 이어콘이 없는 음성 메뉴와의 비교 실험을 한 뒤, 음성 메뉴의 구조를 쉽게 파악하고 사용상 효과적으로 기억에 도움을 주는 이어콘을 음성 메뉴에 적용시킴으로써, 음성 메뉴의 사용성 개선 가능성을 확인하였다.

## 2. 연구 내용 및 방법

### 2.1 이론적 배경

음성 메뉴(Voice Menu)는 합성 음성을 노드(Node)에 사용함으로써 다른 노드들과 상호 연관성 있는 음성들로 이루어져 있는 계층도를 이루고 있다. 일반 메뉴와 마찬가지로 음성 메뉴에서도 사용자는 자신이 목표로 하는 메뉴 항목을 찾아가기 위해 음성 메뉴를 중첩으로 다녀야 한다. 그런데 보이지 않는 연속된 소리의 음성 메뉴에서 그 깊이와 폭을 파악하기는 쉽지 않다. 이러한 기존 음성 메뉴의 단점을 보완하기 위해 이어콘을 적용한 새로운 음성 메뉴를 고려하였다.

음성 메뉴는 인간의 음소 데이터로 이루어진 합성 음성으로 구성되어 있다. 그 대표적인 예로 TTS (Text-To-Speech)를 들 수 있는데, 이는 작은 합성 단위 음성과 언어 처리를 이용하여 임의의 문장에 대한 음성을 생성한다. TTS 절차에는 단위음 합성, 언어 처리, 운율 조절 등으로 나뉘어 지는데, 언어 처리를 이용하여 입력된 문

장을 적당한 합성 단위의 조합으로 대응시키고, 문장으로부터 적당한 억양과 지속시간을 추출하여 합성음의 운율을 결정한다[1].

계층적 메뉴 구조는 오랫동안 시각적 디스플레이에서 사용되어 왔으며 컴퓨터를 이용하는 많은 사람들에게 이미 익숙해진 구조이다. 많은 시각적 디스플레이상의 메뉴는 정보를 체계화하기 위해 계층화 구조와 인간의 사고 패턴을 따름으로써 메뉴의 의미 파악과 기억을 하게 된다. 하지만 이러한 시각적 메뉴와 달리 음성 메뉴는 탐색할 때 메뉴의 의미만 가지고 깊이와 폭을 파악하기 어렵다. 따라서 이어콘을 적용한 음성 메뉴에서 이어콘은 청각적 큐를 제시하여 위치 파악을 도와주고, 합성 음성 메뉴의 의미와 결합되어 탐색에 필요한 강한 큐를 발생시켜 그 문제점을 해결해 줄 수 있다. 즉, 이어콘은 음성 메뉴의 위치적 관계에 대한 정보를 제공하도록 만들어진 메시지로 사용되며, 이를 통해 사용자는 메뉴의 깊이와 폭을 파악할 수 있게 되어 음성 메뉴를 좀 더 효과적으로 사용할 수 있을 것이다.

Blattner, Sumikawa, and Greenberg(1989)는 음악적 파라미터(Musical Parameters)를 이용해 변화되는 다양한 종류의 연관된 소리를 얻어 내기 위해 다양한 음악적 파라미터를 변화시킨 모티브(Motive)라는 음악 소절을 소개한 바 있다[2]. 여기서, 이어콘은 사용자에게 정보를 제공하기 위한 청각적 큐 또는 메시지(Message)라고 정의된다[6]. 사운드 메시지는 어떤 대상에 대한 다양한 측면을 나타내는 청각적 표현이므로 개별적으로는 서로 다른 모양을 하고 있으나 전체적으로 연관된 의미를 나타낼 수 있도록 만들어져야 한다. 이러한 통일되고 연관된 의미를 갖기 위해 본 연구에서는 음역(Register), 음색(Timbre), 리듬(Rhythm), 템포(Tempo) 등을 음악적 파라미터로 사용하였다.

### 2.2 이어콘 파라미터

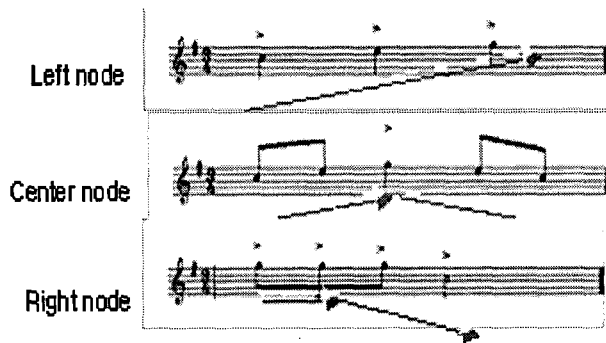
본 절에서는 이상의 음악적 파라미터에 해당되는 네 가지 파라미터를 이용하여 이어콘을 디자인하는 방법을 나타낸다.

음역(Register) 파라미터는 첫 번째 수준(Level 1)과 두 번째 수준(Level 2)에 걸쳐 음색(Timbre)이 결합되어 그 특징을 나타내도록 하였다. 이는 기존의 이어콘이 단음이나 한 가지 음색만으로 만들어진 것을 보완하여 최소한 하나 이상의 파라미터가 결합된 형태를 지향하기 위함이다. 또한 단순음이 아닌 복합적 이어콘을 만들어 청각적 큐(Auditory Cue)를 강화하기 위한 것이다.

음역(Register)과 함께 Level 1과 Level 2를 구분 짓는 파라미터로 사용된 음색(Timbre)은 여러 가지 기구를 이

용할 수 있겠으나 가장 널리 사용되고 인지도가 높은 Roland사의 기구를 가지고 제작하였다. 다양한 악기의 음색 중에서 서로 음색이 뚜렷이 분리되는 대표적인 'French Horn', 'Drum', 'Oboe', 'Violin', 그리고 'Electric Organ'을 선택하였다.

세 번째 수준(Level 3)의 파라미터로써 리듬(Rhythm)을 사용하였고 장단에 의해 구분되도록 하였다. 즉 부모 노드(Parent Node)인 Level 2에서 만들어진 3개의 자식 노드(Child Node)를 좌측 노드(Left Node), 중앙 노드(Center Node), 그리고 우측 노드(Right Node)로 나누어 음성 메뉴가 나타날 수 있는 각각의 경우를 만들고 이를 기준으로 계층적 구조를 갖는 이어콘 디자인과 구현을 하였다.

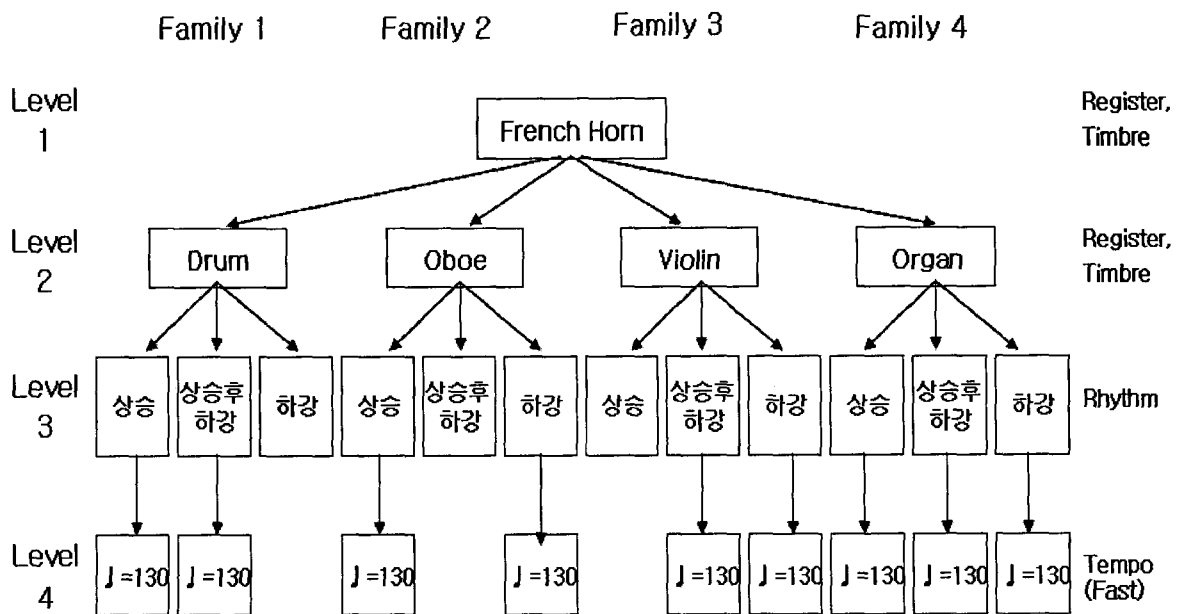


<그림 1> 리듬(Rhythm)

<그림 1>은 Level 3과 4에 해당되는 자식 노드(Child node)에서의 이어콘으로, 좌측 노드는 상승하는(Ascending) 음의 진행 모습을 갖도록 하며, 중앙 노드는 상승하던 음이 중앙에서 다시 하강하는(Descending) 모습을 갖도록 한다. 그리고 우측 노드는 짧은 3분의 1박자의 시작음으로 하여 하강하는 모양을 갖도록 디자인하였다. 이는 정보의 양이 늘어나 메뉴 아이템이 많아질 경우 이러한 구조의 이어콘을 적용한 음성 메뉴를 기초로 노드의 추가에 따른 문제에 대한 해결하기 위한 것이다.

Level 4를 구분 짓는 템포(Tempo) 파라미터에서는, Level 3에서와 같은 방식으로 이어콘을 만든 후에 템포를 '♩=130'으로 만들어 준다. 이것은 유일하게 Level 3과 Level 4를 구분 짓게 하는 파라미터이다. <그림 2>는 앞에서 설명한 이어콘 파라미터에 대한 전체적인 개요를 나타낸 것이다. <그림 2>에서 Level은 메인메뉴-커뮤니케이션-전자우편-드럼위즈 등으로 이어지는 계층의 수준이다. 그리고, Family는 각 메뉴의 계층을 내용별로 분류하여 비슷한 내용의 메뉴는 같은 Family안에 포함되도록 묶은 그룹이다. 또한, Level 1에서 Level 4로 내려갈수록 복잡한 수준으로, Family 1에서 Family 4로 갈수록 익숙하지 않은 내용으로 설정하였다.

<표 1>은 이어콘이 어떻게 구성되어 있는지를 도표로 나타낸 것이다. 이는 이어콘을 이용한 음성 메뉴가 갖는 계층적 구조의 연관성과 통일성을 나타내며, 각각의 이어콘 파라미터의 쓰임새에 대한 구성표이다.



<그림 2> 이어콘 파라미터의 적용

<표 1> Level에 따른 이어콘 파라미터 구성표

	Timbre	Register	Tempo	Rhythm Left node	Rhythm Center node	Rhythm Right node
Level 1	French Horn	A	J = 100			
Level 2	Drum, Oboe, Violin, Organ	F	J = 100			
Level 3	Drum, Oboe, Violin, Organ	F-A-C	J = 100	상승음	상승후 하강음	하강음
Level 4	Drum, Oboe, Violin, Organ	F-A-C	J = 130	상승음	상승후 하강음	하강음

- Pentium 4 PC
- Sound Level Meter (SM-9600)

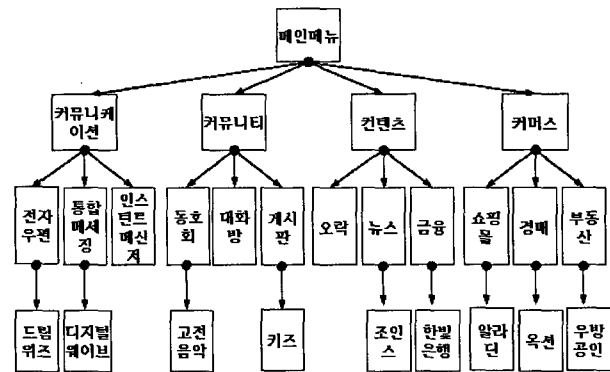
실험용 소프트웨어는 Microsoft사의 Visual C++를 이용하여 제작하였다. <그림 3>은 본 실험에서 사용된 음성 메뉴의 구조도이며, <그림 4>는 실험용 소프트웨어의 주요화면이다. 여기서 이어콘이 있는 음성 메뉴는 상단에 배치하고, 이어콘이 없는 음성 메뉴를 하단에 배치함으로써 두 그룹간 실험을 원활하게 수행하도록 하였다. 그리고 <그림 3>의 음성 메뉴 구조도와 같이 Level과 Family에 맞는 배치를 위해 <그림 4>에서 보는 것처럼 26개 버튼을 이용하도록 하였다.

### 3. 실험 계획

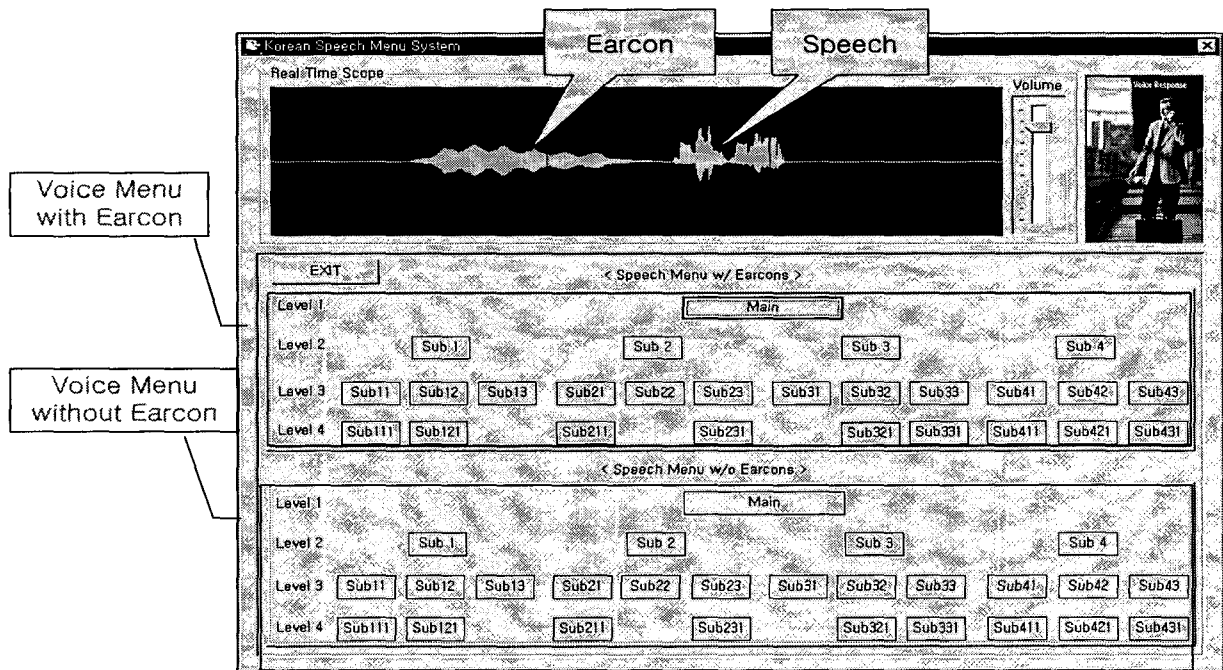
#### 3.1. 실험 장치

본 실험에서 사용한 하드웨어는 다음과 같다.

- Sound Blaster Live DE 5.1 Card
- SHURE VR250 BT Headset



<그림 3> 실험에 사용된 음성 메뉴의 구조도



<그림 4> 실험용 소프트웨어

### 3.2 실험 방법

피실험자는 16명의 대학원생이 실험에 참여하였고, 남 14명과 여자 2명으로 구성되었다. 평균 나이는 25.8세이고 표준편차 2.47이다. 모든 피실험자는 한국에서 태어나 정상적인 교육과정을 거쳐 음성 메뉴를 듣고 기억하고 답하는데 문제가 없는 사람들이었다.

먼저 예비 실험으로 합성 음성에 대한 피실험자의 이해도(Intelligibility)를 평가하고 실험 주변의 소음 정도에 대한 측정을 실시하여 피실험자 및 주변 환경이 실험에 적합한지 파악을 한 후에 본 실험을 실시하였다.

본 실험은 우선 피실험자들의 학습효과를 제거하기 위하여, <그림 4>에서 8명의 피실험자(그룹 1)들은 상단부의 이어콘이 적용된 음성 메뉴를 대상으로 실험을 시작하였고 또 다른 8명의 피실험자(그룹 2)들은 하단부의 이어콘을 이용하지 않은 음성 메뉴를 대상으로 실험을 시작하였다. 즉, 피실험자 그룹 간에 시작하는 메뉴만 다를 뿐 각 피실험자들은 2개의 메뉴를 모두 실험하였다. 그리고 각 피실험자에게 음성 메뉴를 청취한 후 위치와 메뉴를 기억하고 답변하도록 하여 음성 메뉴의 정확성과 반응 시간을 측정하였다.

이런 방식으로 피실험자의 상기율(Recall Rate), 반응 시간(Response Time), 그리고 선호도(Preference)에 대한 실험 데이터를 수집하였다. 또한 피실험자의 답변을 위한 실험용 데이터 시트(Data Sheet)를 가지고 사용성 평가의 분석 자료로 활용하였다. 상기율(Recall Rate)은 각 피실험자에게 제시되는 10개의 문제 중 위치를 맞춘 개수로 하였으며, 반응 시간은 피실험자가 음성 메뉴를 듣고 데이터 시트에 기록하기 시작하는 때까지의 시간을 측정하였다.

이렇게 측정된 데이터는 Family, Level별로 각각 상기율을 분석하였으며, 이어콘이 있는 음성 메뉴와 없는 음성 메뉴를 전체 비교하는 분석도 하였다. 또한 반응 시간의 단위는 msec(1/100초)까지 측정하여 두 그룹 사이에 반응 시간의 차이가 있는지를 분석하였다. 그리고 선호도 조사를 위해 이어콘이 있는 음성 메뉴를 가지고 실험한 사람은 이어콘이 없는 음성 메뉴를 같은 방법으로 테스트 해 본 뒤 둘을 비교하여 피실험자의 선호도를 표시하도록 하였다. 또한 이어콘이 없는 음성 메뉴를 먼저 시도한 피실험자도 이어콘이 있는 메뉴에 대해서 테스트 해 본 다음 선호도를 표시하도록 하였다.

### 3.3 실험 절차

모든 실험절차는 실험방법에서 설명된 바와 같이 그룹 간에 순서만 다를 뿐 같은 방식으로 수행되었다. 다

만, 이어콘이 있는 메뉴에 대한 피실험자들의 연습을 위한 단계(단계 1)가 추가되는 차이는 발생하였다. 따라서 구체적인 절차를 살펴보면 다음과 같다. 단계 1에서는 피실험자 그룹 1에 대하여만 실시하였는데, 이어콘만을 메뉴별로 청취함으로써 이어콘의 계층성을 피실험자가 숙지하였다. 단계 2에서 피실험자는 실험 데이터 수집용 시트를 받고, 메뉴 전체를 한 번 청취하였다. 단계 3에서는 임의의 하나의 메뉴를 청취 후, 위치를 기억하여 기록하였으며, 총 10개의 문제에 대하여 수행하였다. 단계 4에서는 반대의 경우인 이어콘이 있는(또는 없는) 경우에 대하여 단계 1(필요시), 2, 3을 수행하였다.

## 4. 실험 결과

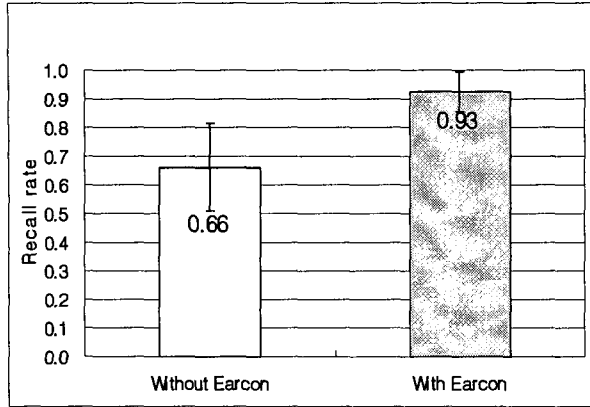
피실험자에게 음성 메뉴에 대한 훈련을 시킨 후 각각 10개의 태스크(Task)를 부여하여 상기율을 측정하는 실험에서 이어콘이 있는 음성 메뉴가 이어콘을 이용하지 않은 음성 메뉴보다 상기율이 좋게 나와 유의한 차이가 있는 것으로 나타났다.

피실험자의 음성 메뉴에 대한 반응 시간을 측정하는 실험에서는 이어콘 있는 음성 메뉴에서 평균 3.85초, 표준편차 ±1.10초가 나왔고, 이어콘 없는 음성 메뉴에서 평균 4.98초, 표준편차 ±1.52초가 나와서 두 그룹 간에 이어콘이 있을 때 피실험자의 반응 시간이 줄어드는 것으로 나타났다.

선호도 조사에서는 87.5%(16명 중 14명)가 이어콘이 적용된 한글 음성 메뉴를 선호했으며, 적용된 이어콘 파라미터 중 리듬(Rhythm), 템포(Tempo), 음역(Register), 음색(Timbre) 순으로 이 실험의 태스크를 수행하는데 효과가 있었다고 응답하였다. 자세한 실험 분석 결과는 다음과 같다.

### 4.1 상기율(Recall rate)

두 그룹(이어콘이 있는 음성 메뉴로 테스트한 그룹과 이어콘이 없는 음성 메뉴를 테스트한 그룹)간의 상기율을 비교한 결과 <그림 5>에서 보는 바와 같이 각각의 평균은 0.93(표준편차 ±0.07)과 0.66(표준편차 ±0.15)로 나타났다. 그리고 두 그룹간의 차이에 대한 유의성 검정을 실시한 결과, 유의 수준 0.05에서  $t = 4.46$ ,  $p < 0.0012$ 로 나타나 두 그룹 간 상기율에 있어서 유의한 차이가 있는 것으로 나타났다. 따라서 상기율 관점에서 이어콘이 있는 음성 메뉴가 없는 것보다 사용성이 좋다고 할 수 있겠다.



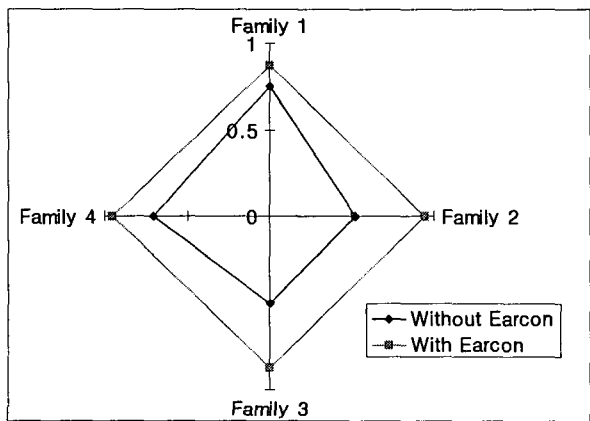
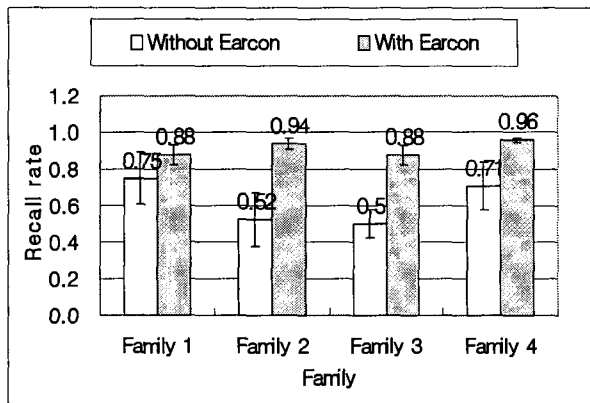
<그림 5> 두 그룹 간 상기율의 비교

또한 두 그룹 간 Level 별 분석 결과에 대한 추세선 그래프를 작성했을 때 두 그룹의 Level 3과 Level 4의 상기율 간격의 차이가 있음을 알 수 있었다. 이 차이에 대한 유의성 검정을 실시해 본 결과, 유의 수준 0.05에서 Level 3의  $t=2.75, p=0.0157$ 로 나타났고, Level 4에서는  $t=2.73, p=0.0177$ 로 나타나 Level 1과 2와 달리 이어콘

의 유무에 따라 Level 3과 Level 4에서 유의한 차이가 있음을 알 수 있었다.

이러한 결과는 상위 메뉴에서 하위 메뉴로 복잡해질수록 이어콘의 효과가 나타나기 때문으로 설명된다. 상위 메뉴에 해당하는 Level 1과 Level 2에서는 이어콘의 유무에 따른 상기율의 평균에 차이가 없으나, 하위 메뉴에 해당하는 Level 3과 Level 4에서는 유의한 차이가 나타나고 있는 점으로 미루어 보아 이어콘이 있는 음성 메뉴의 효과를 확인할 수 있었다.

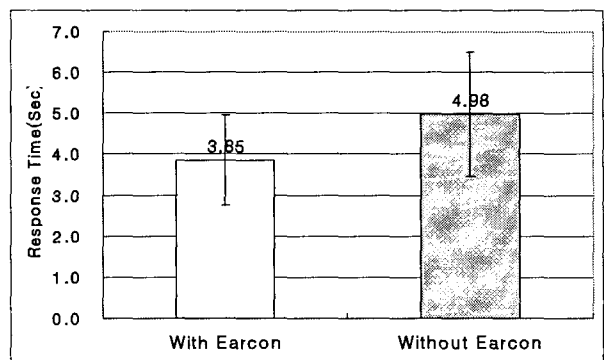
<그림 6>은 두 그룹 간 Family별 상기율을 비교해 본 결과이다. Family 2와 Family 3의 상기율 차이에 대한 유의성 검정을 위해 T-Test를 실시한 결과, Family 2와 Family 3의 경우 각각  $t=2.80, p=0.0191$ 과  $t=2.91, p=0.0118$ 로 나타나 이어콘이 적용되었을 경우 상기율이 개선되었음을 알 수 있었다. 이는 일반 대학원생들에게 익숙한 “커뮤니케이션”과 “커뮤니티”에 대한 내용의 메뉴보다는 상대적으로 덜 익숙한 “컨텐츠”와 “커머스”에 대한 메뉴에서 이어콘의 효과가 더 크게 나타나는 것으로 설명될 수 있으며, 이로부터 익숙하지 않은 내용에 대한 이어콘이 포함된 메뉴가 효과가 있음을 알 수 있었다.



<그림 6> 두 그룹 간 Family별 상기율의 비교

#### 4.2 반응 시간(Response Time)

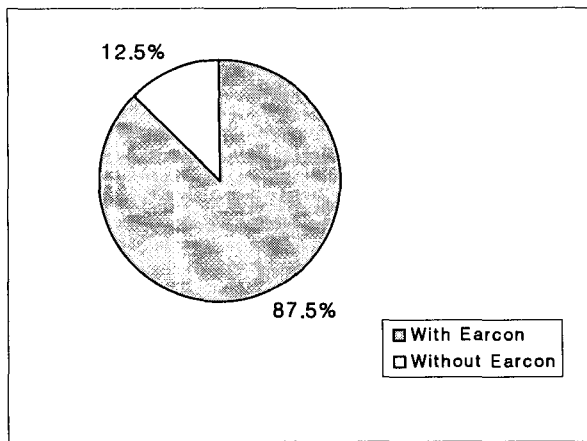
<그림 7>은 이어콘이 있는 음성 메뉴와 이어콘이 없는 음성 메뉴간 피실험자의 반응시간을 비교한 것이다. 이어콘이 있는 메뉴와 이어콘이 없는 메뉴 사이에서 반응 시간이 각각 3.85초(표준편차  $\pm 1.10$ )와 4.98초(표준편차  $\pm 1.52$ )로 나타났으며, 이 차이에 대한 유의성 검정 결과  $t=7.0, p<0.0001$ 로 나타났다. 음성 메뉴에 이어콘을 이용함으로써 반응 시간이 줄어든 것을 알 수 있으며, 피실험자가 음성 메뉴를 듣고 메뉴와 위치를 알아맞히는 데 따르는 어려움이 감소되었음을 알 수 있다. 따라서 반응시간 관점에서도 이어콘이 있는 음성 메뉴가 없는 것보다 사용성이 좋다고 할 수 있겠다.



<그림 7> 두 그룹 간 반응 시간의 비교

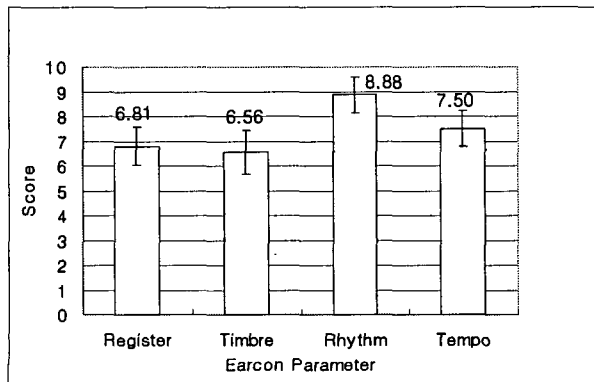
### 4.3 선호도(Preference)

<그림 8>은 이어콘이 있는 음성 메뉴와 이어콘이 없는 음성 메뉴 사이의 선호도를 조사해 본 결과이다. 이어콘이 있는 음성 메뉴에 대한 피실험자의 선호도가 87.5%(16명중 14명이 선호)로 나타났다. 즉, 단조로운 합성 음성만으로 만들어진 메뉴보다 계층적 정보를 제시하는 이어콘이 있는 합성 음성 메뉴를 더 선호하는 것으로 나타났다.



<그림 8> 두 그룹 간 선호도의 비교

<그림 9>는 피실험자에게 실험에서 이용되었던 이어콘을 설명하고, 실험 후 파라미터에 대한 선호도를 조사한 결과이다. 선호도는 10점 척도로 조사하였는데, 각 파라미터에 대하여 0점(좋지 않다) ~ 5점(보통) ~ 10점(좋다)으로 표시하도록 하였다. 여기서 파라미터에 대한 선호도는 리듬(Rhythm; 8.88±0.72)이 가장 많았고, 다음은 템포(Tempo; 7.50±0.73), 그 다음은 음역(Register; 6.81±0.75), 그리고 음색(Timbre; 6.56±0.89)의 순서로 나타났다. 그리



<그림 9> 이어콘 파라미터에 대한 선호도의 비교

고 유의수준 0.05에서 가장 높은 점수의 리듬(Rhythm)과 그 다음 점수인 템포(Tempo)간에는 유의한 차( $t=3.07$ ,  $p=0.0065$ )가 나타났으며, 템포(Tempo), 음역(Register), 음색(Timbre) 간에는 차이가 없었다. 따라서 이어콘을 제작할 때 리듬(Rhythm) 파라미터를 많이 사용하는 것이 이어콘이 있는 메뉴의 선호도 면에서 효과가 있음을 알 수 있었다.

### 5. 결론

본 연구에서 이어콘을 적용한 음성 메뉴의 상기와 이어콘이 없는 음성 메뉴의 상기는 각각 92.50%와 66.25%로 나타나 이어콘이 적용된 음성 메뉴의 경우에 좋은 결과가 나타났다. 또한 각 Level별 분석과 Family별 분석을 통해 이어콘이 적용된 음성 메뉴가 유의하게 더 좋은 결과가 나타났다. 이를 바탕으로 이어콘을 적용한 음성 메뉴는 사용자로 하여금 메뉴의 계층적 구조를 쉽게 파악하게 하고 탐색하려는 메뉴의 위치와 내용을 쉽게 기억하게 하는 조력성(Helpfulness)이 있다고 판단된다.

이어콘을 적용한 음성 메뉴의 반응 시간과 이어콘이 없는 음성 메뉴의 평균 반응 시간이 각각 3.85 초와 4.98초로 나타나 이어콘이 있는 음성 메뉴에 대한 반응 시간이 짧게 나타났다. 이로부터 이어콘을 음성 메뉴에 적용함으로써 음성 메뉴에 대한 반응 시간을 단축시켜 전체적으로 탐색에 소요되는 시간을 줄여주는 것을 알 수 있었다.

한편 피실험자의 선호도 조사를 통해 이어콘이 적용된 합성 음성 메뉴가 사람들에게 더욱 선호된다는 것을 알 수 있었다. 그리고 파라미터는 리듬(Rhythm)과 템포(Tempo)를 상대적으로 선호함을 알 수 있었다. 따라서 앞으로 더욱 세련되고 정교한 파라미터를 갖는 이어콘을 만들고 그 의미를 정확히 이해할 수 있는 명료도 높은 합성 음성을 통해 이용한다면 선호도는 높아지고 사용성은 개선될 것이다.

본 연구에서는 이어콘 파라미터를 적용하여 기존 음성 메뉴가 가지고 있었던 문제점을 보완 개선하기 위한 방법을 모색하려고 하였으며, 이어콘을 적용한 음성 메뉴에 대한 많은 가능성을 발견하였다. 현대는 멀티미디어 시대를 맞고 있으며, 사용자는 예전 같은 단순한 자극을 주는 소리보다 좀 더 다양한 음색과 리듬이 있는 소리를 원하고 있다. 따라서 음악적 요소를 풍부히 담고 있고 소리로서 정보 전달의 큰 역할을 할 수 있는 이어콘의 활용이 확대되어야 할 것이다.

한편 현재의 합성 음성은 아직도 미흡한 점이 많아 사용자에게 친숙하지 못한 실정이므로, 합성 음성에 관

해서 사용자의 이해도(Intelligibility)를 높이기 위한 노력이 필요하다고 본다. 이러한 점을 고려할 때 앞으로는 이어콘을 적용한 음성 메뉴의 설계는 사용자 측면을 고려하여 진행되어야 할 것이다. 그리고 이러한 이어콘이 적용된 합성 음성 메뉴는 Voice Portal, Voice Commerce, Interactive Voice Response(IVR) 등 많은 응용 분야에서 활용될 것으로 기대된다.

### 참고문헌

- [1] 오영환, 음성 언어 정보 처리, 홍릉 과학 출판사, 1998.
- [2] Blattner, M., Sumikawa, D. & Greenberg,; “R. Earcon and icons: Their structure and common design principles”, *Human Computer Interaction*, 4: 11-44, 1989.
- [3] Brewster, S.A., Wright, P.C. and Edwards, A.D.N.; “A detailed investigation into the effectiveness of Earcon. In G. Kramer (Ed.), *Auditory display, sonification, audification and auditory interfaces.*” *The Proceedings of the First International Conference on Auditory Display*, Santa Fe Institute, Santa Fe, NM: Addison-Wesley, pp. 471-498, 1999.
- [4] Juang, B.H. and Tsuhan Chen,; “The past, present, and future of speech processing,” *IEEE Signal Processing Magazine*, 15: 24 -48, 1998.
- [5] Leplatre, G., and Brewster, S.A.; “An Investigation of Using Music to Provide Navigational Cues.” In *Proceedings of ICAD'98 (Glasgow, UK)*, British Computer Society, 1998.
- [6] Meera M. Blattner, & Roger B. Dannenberg,; “Multi-media Interface Design”, *ACM Press frontier series*, 1992.