

명령 제어 음성 인터페이스 사용자 중심 설계를 위한 사용성 평가도구의 설계 및 구현

Design and Implementation of a Usability Testing Tool for User-oriented Design of Command-and-Control Voice User Interfaces

이 명 지¹⁾ · 홍 기 형²⁾

Lee, Myeongji · Hong, Ki-Hyung

ABSTRACT

Recently, usability has become very important in voice user interface systems. In this paper, we have designed and implemented a wizard-of-oz (WOZ) usability testing tool for command-and-control voice user interfaces. We have proposed the VUIDML (Voice User Interface Design Markup Language) to design the usability test scenario of command-and-control voice interfaces in the early design stages. For highly satisfactory voice user interfaces, we have to select highly preferred voice commands and prompts. In VUIDML, we can specify possible prompt candidates. The WOZ usability testing tool can also be used to collect user-preferred voice commands and feedback from real users.

Keywords: Voice user interface, Usability testing

1. 서론

스마트 폰과 같은 모바일 기기나 내비게이션과 같은 전자 제품, 가전 제품, IPTV 등 다양한 분야에 음성 인식 기반의 음성 사용자 인터페이스가 적용되어 상용화되는 사례가 증가하고 있다[1][2][3]. 이에 따라서 음성 사용자 인터페이스를 시스템에 적용했을 때 제품의 사용성이 어떻게 증가하는지 또는 어떻게 해야 사용성을 향상시킬 수 있는지에 대한 문제가 논의되고 있다[4][5].

음성 사용자 인터페이스 설계 단계에서 고려되는 요소는 프롬프트, 인식 문법, 대화 로직이 있다. 음성 사용자 인터페이스의 사용성 향상을 위해서는 이 세 가지 구성 요소 설계 시 사

용성 향상에 중점을 두고 작업을 진행해야 한다. 음성 사용자 인터페이스의 구성 요소의 사용성을 향상시키기 위해서는 반복적인 설계가 요구된다[6].

사용자 중심 설계 방법은 반복적인 설계를 원칙으로 한다[7]. 이를 음성 사용자 인터페이스의 구성 요소 설계에 적용하는 경우 사용성 평가를 통해 프롬프트의 선호도 조사, 명령어 수집, 대화 흐름 유형 수집과 그 결과를 활용한 구성 요소의 재설계가 가능하다. 이 과정에서는 프롬프트, 인식 문법, 대화 로직만 설계되고 최종 사용자가 사용하게 될 시스템이 완전히 구현된 상태가 아니므로, 시뮬레이션을 통한 초기 사용성 평가가 실시되는 것이 적합하다. 초기 사용성 평가는 위자드 오브 오즈 방식으로 진행되어 평가에 참가하는 실제 사용자가 평가 대상인 시스템이 제대로 작동된다고 인지하도록 시뮬레이션을 할 수 있게 준비되어야 한다.

본 논문에서는 명령제어형 음성 인터페이스 설계 시 실시할 초기 사용성 평가를 위한 위자드 오브 오즈(Wizard Of Oz)방식의 시뮬레이션 도구를 설계하고 구현하였다. 초기 사용성 평가에서 실시될 시뮬레이션 대상은 음성 사용자 인터페이스의 세 가지 구성 요소 중 프롬프트, 인식 문법으로 제한하였다. 또한

1) 성신여자대학교 revioopp@gmail.com

2) 성신여자대학교 kihyung.hong@gmail.com, 교신저자
이 논문은 2010년도 성신여자대학교 학술연구조성비 지원에 의하여 연구되었음.

VoiceXML[8]을 재구성한 VUIDML(Voice User Interface Design Markup Language)[9]을 설계하고 해석기를 구현하였다. VUIDML은 XML 기반 언어로서 사용자가 사용하기 편한 음성 명령과 선호하는 프롬프트 수집을 위한 다양한 대화 흐름과 후보 프롬프트들을 포함하는 다이얼로그 시나리오를 작성할 수 있다.

본 논문의 구성은 다음과 같다. 2장에서는 음성 사용자 인터페이스의 사용성 평가에 대한 기존 연구 사례를 서술하고 사용성 향상을 위해 보완할 점을 제시한다. 3장에서는 음성 사용자 인터페이스의 사용성 평가 도구를 설계한 내용을 기술한다. 4장에서는 설계를 토대로 구현한 내용을 기술한다. 마지막으로 5장에서 결론을 맺는다.

2. 음성 사용자 인터페이스의 사용성 평가

본 장에서는 음성 사용자 인터페이스의 사용성 평가를 위한 도구를 제안하기 위하여 사용성 및 사용성 평가에 대한 내용과 음성 사용자 인터페이스의 사용성 향상 요소를 기술하고, 기존에 제안된 도구 사례를 통해 보완되어야 할 점과 도구의 설계 배경에 대하여 설명한다.

2.1 사용성과 사용성 평가

사용성은 제품의 기능성처럼 모든 제품의 특성 중 하나이다. 기능성은 제품이 무엇을 할 수 있는지를 가리킨다. 기능성을 시험한다는 것은 제품이 제품 명세서대로 작동하는지를 보는 것이다. 반면 사용성은 사용자가 제품을 어떻게 다루는가를 나타낸다. 넬슨은 사용성이란 사용자 인터페이스가 얼마나 사용하기 쉬운지를 평가하는 질적인 속성이라고 정의하였다[10]. 듀마와 래디쉬(Duma and Radish)는 사용성은 사용자가 제품을 사용하여 직무를 빠르고 쉽게 완수하는 것이라고 정의하였다[11].

사용성 평가란 평가의 대상이 되는 제품의 사용성을 향상시키기 위해 제품의 실제 사용자를 대상으로 제품을 이용하여 특정 작업을 수행하도록 하여 사용성을 측정하는 시험이다. 즉, 사용성 평가는 사용자가 본인의 필요에 의해 제품의 해당 기능을 찾아서 사용하는 것을 보는 것이다[11]. 사용성 평가 방법 중 위자드 오브 오즈(Wizard Of Oz; WOZ) 기법은 아직 존재하지 않는 시스템 사용성 평가 기법이다. 이는 시스템 설계자가 시스템이 실제 구현되기 전에 제품(설계)을 사용성 평가할 수 있도록 해 준다[12]. WOZ 기법은 시스템에 대한 사용자의 기대를 수집하는데 이상적이다. 개발자는 프로그래밍하는 비용 없이 WOZ 기법의 사용성 평가를 통하여 실사용자로부터 설계상의 사용성 문제를 파악하고 이를 보완하는 재설계를 할 수 있다.

2.2 음성 사용자 인터페이스의 사용성 향상 요소

2.2.1 음성 사용자 인터페이스 구성 요소

음성 사용자 인터페이스(Voice User Interface; VUI)란 사람과 음성 언어 어플리케이션이 서로 상호 작용하는 것으로 프롬프트(시스템 메시지), 인식 문법, 대화 로직(call flow)으로 구성된다[13][14]. 프롬프트는 대화 중에 시스템이 사용자에게 들려주는 합성된 음성 또는 미리 녹음된 음성을 말한다. 인식 문법은 각각의 프롬프트에 대한 사용자의 모든 가능한 응답을 정의한다. 시스템은 문법에 포함되어있는 단어, 문장, 구문들만 인식한다. 대화 로직은 시스템에 의해 발생할 수 있는 모든 가능한 행위를 정의한다. 사용자의 발화에 대한 응답이나 데이터베이스에서 조회된 정보를 음성으로 출력해 주는 등의 행위를 예로 들 수 있다.

음성은 인간이 사용하는 가장 자연스러운 의사소통 수단이지만, 이를 인터페이스로 하여 유용하고 신뢰성 있게 설계하기 위해서는 사용자 중심 설계를 통해 사용성을 향상시켜야 한다 [15].

위에서 언급한 세 가지 음성 사용자 인터페이스 구성 요소를 설계하는 과정에서 사용자 중심 설계 방법을 적용하여 초기 사용성 평가를 실시해야 한다[16][17]. 이 때 초기 사용성 평가에서 사용자가 ‘완전히 개발된 시스템을 사용하고 있다.’라고 믿게 하기 위해서 WOZ 기법을 적용하여 사용성 평가를 실시하는 것이 유용하다[7].

2.2.2 프롬프트 사용성 향상 요소

프롬프트의 사용성을 향상시키기 위해서는 사용자로부터 선호하는 프롬프트의 유형을 수집해야 한다. 프롬프트는 실제 사용자가 이해할 수 있는 사용자 어휘들로 구성되어야 한다. 만약 사용자가 접하기 어려운 전문 용어나 시스템 설계자나 개발자들이 사용하는 단어를 포함하여 프롬프트를 구성하였을 경우, 사용자는 프롬프트를 듣고 이해하기 힘든 불편함을 겪게 된다. 이와 같은 문제의 발생을 막기 위해 사용성 평가를 통해 프롬프트에 대한 이해도를 조사해야 한다. 또한 여러 개의 대안이 있는 프롬프트들을 차례대로 사용자에게 들려주어 어느 프롬프트를 선호하는지 선호도를 조사해야 한다.

2.2.3 인식문법 사용성 향상 요소

인식 문법을 설계할 때 인식 명령어를 기억하는 부담을 줄이기 위해서는 사용성 평가 중에 사용자가 실제로 사용하는 명령어들을 수집해야한다. 즉, 사용성 평가 도중 사용자가 주어진 작업을 수행하기위해 발화하는 명령어들을 수집해야 한다. 이 때 수집된 명령어들을 동일하거나 유사한 명령어들로 분류하고, 각 분류별 빈도수를 측정한다. 또는 미리 사용자가 발화할 것이라고 예상한 명령어 후보들을 선정하고, 사용성 평가 중이나 사용성 평가 종료 후 사용자에게 실제로 발화할 것 같은 명

령어 후보가 무엇인지를 질의하여 사용자의 선택을 기록한다. 응답 결과를 분석하여 각 후보들의 빈도수를 측정한다. 마지막으로 빈도수가 높은 명령어들을 음성 인식 명령어로 지정한 후 인식 문법을 설계하면 사용자가 명령어를 기억하는 노력을 감소시키고, 발화하기 자연스러운 명령어 사용을 통해 인식 성공률을 증가시킬 수 있다.

2.3 사용성 평가 도구 사례와 보완점

음성 사용자 인터페이스 사용성 평가를 위한 도구의 요소와 분석에 의한 장단점 도출을 위해 기존 도구의 사례들을 조사하였다. 수집한 도구 사례들의 유형을 진행 방식에 따라 2 가지로 분류하였고, 사용 목적에 따라 4 가지로 분류하였다.

2.3.1 진행 방식에 의한 분류[18][19]

동기식 방식은 1 대 1 방식으로 원격 사용성 평가 도구를 통해 사용성 평가를 진행한다. 사용성 평가 과정의 통제가 가능하고, 문제 발생 시 시험자가 관여할 수 있다. 또한 평가에 참가한 사용자를 직접 관찰하고 사용자의 반응과 의견을 기록할 수 있다. 그러나 사용되는 실제 환경이 아니기 때문에 발견되지 않는 사용성 문제들이 존재할 수 있다.

비동기식 방식은 일반적으로 온라인 형태의 웹 페이지로 기능이 제공된다. 동시에 다수의 사용자가 실제 사용 환경에서 평가를 실시 할 수 있어 비용이 동기식에 비해 적게 들고 수집한 데이터의 양은 많다. 그러나 평가 과정의 사용자를 관찰할 수 없기 때문에 시험자와 사용자간의 상호작용이 발생하지 않는다. 또한 분석된 데이터의 질이 동기식에 비해 낮다.

2.3.2 사용 목적에 의한 분류[20][21]

사용 목적에 따라 설문 제작 및 배포, 전문가 디자인 평가, 인터랙션 평가, 음성 사용자 인터페이스 평가로 분류하였다. 분류에 대한 설명과 사례들은 <표 1>에 나타나 있다.

2.3.2 문제점 및 보완점

일반적인 사용성 평가를 지원하는 도구는 음성 사용자 인터페이스를 구성하는 프롬프트, 인식 문법을 대상으로 사용성을 평가하기 위한 오디오 출력 및 녹음 기능을 제공하지 않는다. 또한 조사된 기존의 VUI 사용성 평가 도구는 설계 단계에서 프롬프트의 선호도를 조사하는 것이 불가능하고 음성 인식의 결과를 선택하여 텍스트를 오디오로 실시간으로 변환하여 시뮬레이션을 실시할 수 없다.

위와 같은 점을 보완하기 위하여, 음성 사용자 인터페이스의 다이얼로그 시뮬레이션을 통한 프롬프트, 인식 문법 구성 요소의 사용성 향상에 특화된 프롬프트 선호도 조사 기능 및 명령어 수집 기능의 추가가 요구된다. 또한 TTS 방식을 적용하여 평가 전에 프롬프트를 모두 녹음할 필요 없이 사용성 평가 중

에 실시간으로 프롬프트의 텍스트를 오디오로 변환하여 재생할 수 있는 기능을 갖춘 WOZ 방식이 지원되어야 한다.

표 1. 사용 목적에 의한 분류
Table 1. Classification based on the target domain

사용목적	설명 및 도구
설문 제작 및 배포	온라인상의 마케팅 자료 수집 또는 연구 목적의 설문 제작 및 배포 (Surveygizmo, SurveyMonkey, Zoomerang, PollDaddy, FormSite, Moodle, LimeSurvey, FormSite 등)
전문가 디자인 평가	전문가에 의한 웹 사이트 디자인 또는 인터페이스에 대한 의견 수집 (Concept Feedback)
인터랙션 평가	사용자의 인터랙션 사용성 평가 (Chalkmark, ClickHeat, ClickTale, Clixpy) 두 웹사이트의 각 트래픽 분석에 의한 평가 (Google Website Optimizer)
음성 사용자 인터페이스 평가	다이얼로그 시나리오 작성 및 실제 평가 과정에서 사용자의 발화 녹음 (SUEDE)

3. VUI 초기 사용성 평가 도구 설계

3.1 구현 범위

본 논문에서 제안한 도구의 구현 범위는 <그림 1>의 점선 부분이다.

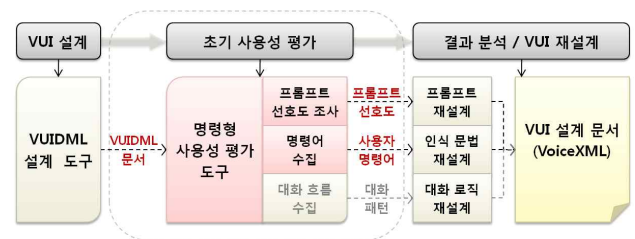


그림 1. VUI 초기 사용성 평가
Figure 1. Early Usability testing for VUI

설계된 음성 사용자 인터페이스를 대상으로 초기 사용성 평가를 실시하기 위해 다이얼로그 시뮬레이션용 시나리오가 요구된다. 이 다이얼로그 시뮬레이션 시나리오용 XML 기반 언어인 VUIDML[9]로 작성한 문서는 명령형 사용성평가 도구에서 해당 다이얼로그가 시뮬레이션 된다. 이때 사용되는 명령제어형 사용성 평가 도구는 WOZ 방식으로 구성되어 있으며, 시뮬레이션 과정에서 사용자를 대상으로 프롬프트 선호도 조사와 명령어 수집을 실시한다. 구현된 도구를 통해 사용자가 선호하는 프롬프트와 사용자 명령어를 초기 사용성 시험 결과로 얻을 수 있다.

3.2 VUIDML

3.2.1 다이얼로그 모델

명령제어형 음성 인터페이스의 사용자 중심 설계를 위해서 VoiceXML[8]의 다이얼로그를 분석하였다. 명령제어형 다이얼로그는 일반적으로 시스템의 프롬프트, 사용자의 발화(응답), 인식의 성공 여부에 따른 다음 작업진행으로 구성된다. 사용자의 응답과 인식 성공여부에 따라, 인식의 성공, 무응답 (응답시간 초과), 인식실패로 나눌 수 있다.

VoiceXML은 설계가 완료된 VUI이므로 초기 설계를 위한 사용자성 분석을 위한 다이얼로그 시뮬레이션에서는 후보 프롬프트 리스트의 제시와 시스템의 오인식(사용자 입장에서)에 대한 시뮬레이션이 필요하다. 이를 위하여 <그림 2>와 같이 초기 사용자성 분석을 위한 다이얼로그모델을 설계하였다.

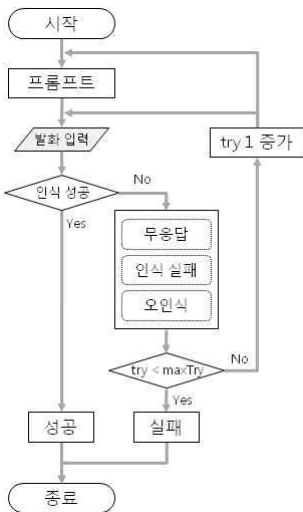


그림 2. 다이얼로그 모델
Figure 2. Dialog simulation Model

초기 사용자성 분석을 위한 다이얼로그 시뮬레이션을 위하여 XML 기반의 VUIDML[9]을 설계하였다. VUIDML은 프롬프트, 무응답/인식실패 프롬프트를 위한 후보 프롬프트 리스트를 명시할 수 있고, 무응답/인식실패/오인식 시뮬레이션을 위한 대화 흐름을 명시할 수 있다. 또한, 사용자 응답 발화(음성인식)를 수집할 수 있는 변수 설정 기능이 있다.

3.2.2 VUIDML 예제

<그림 3>은 VUIDML의 문법을 사용하여 명령어 수집을 위한 사용자성 평가용 다이얼로그 시나리오를 작성한 것이다.

첫 번째 <prompt>는 사용자의 첫 번째 시도 시 출력되며 세 개의 후보 프롬프트를 포함한다. 두 번째 <prompt>는 한 개의 프롬프트를 포함하고, 'count' 속성이 2이므로 두 번째 시도부터 최종 시도까지 'noinput.wav'파일이 재생된다. 사용자가 프롬프트를 듣고 발화를 하면 시험자는 음성 인식 피드백을 사용자에게

제공하기 위해, 사용자의 발화 상태에 따라 <noinput>, <nomatch>, <error>, <match>, <failed> 를 결정하여 프롬프트하게 된다. <error> 의 경우 두 개의 후보 프롬프트 문장을 포함하므로 두 가지 경우로 프롬프트가 가능하다.

이처럼 목적에 맞게 VUIDML을 작성하고, 이 문서에 맞추어 음성 인터페이스가 다양한 대화 흐름으로 시뮬레이션 되도록 시나리오를 설계할 수 있다.

```
<vuidml version="1.0">
<meta writer="홍길동" createDate="2011-01-25"
  modifyDate="2011-03-05"
  title="게임 시작 명령어 수집" goal="게임 시작">
  게임을 시작하기 위한 명령어를 수집한다. </meta>
<form id="form2" keyword="게임 시작"
  object="시작 명령어 수집">
<field id="fd_0" maxTry="3" inputType="recording"
  maxTime="5000">
  <prompt type="alternative">
  <candidate index="1">
    시작할 게임의 이름을 말씀해주세요. </candidate>
  <candidate index="2">
    시작할 게임을 말씀해주세요. </candidate>
  <candidate index="3"> 게임을 시작해주세요. </candidate>
  </prompt>
  <prompt count="2" type="simple">
  <audio src="noinput.wav"/></prompt>
  <noinput> 시간이 초과되었습니다. </noinput>
  <noinput count="3"> 유효한 입력 신호가 없습니다.
  기기 상태를 점검해주세요. </noinput>
  <nomatch> 명령어를 다시 한 번 말씀해주세요. </nomatch>
  <error type="alternative">
  <candidate index="1">
    처음 화면으로 이동합니다. </candidate>
  <candidate index="2"> 도움말이 시작됩니다. </candidate>
  <match> 성함 확인이 완료되었습니다. </match>
  <failed> 사용법을 숙지하신 후 다시 시작해주세요.
  대화를 종료합니다. </failed>
  </field></form></vuidml>
```

그림 3. VUIDML 작성 문서
Figure 3. A sample VUIDML document

3.3. 기능 요구 사항

초기 음성 사용자 인터페이스 사용자성 평가 도구는 다음과 같은 기능을 가져야 한다.

- VUIDML 시나리오 기반 다이얼로그 자동 진행
- 프롬프트 선호도 조사 및 기록
- 사용자 발화 명령어 수집
- 다양한 패턴의 다이얼로그 흐름 제공
- WOZ 방식의 음성 인식 시뮬레이션
- 사용자 발화 실시간 모니터링
- 로그 기록

4. VUI 초기 사용성 평가 도구 구현

4.1 구현 환경

도구를 구현한 환경은 다음 <표 2>와 같다.

표 2. 구현 환경
Table 2. The development environment

구분		내용
하드웨어	CPU	Intel Core2 Duo T7800 2.6GHz
	메모리	4GB
	HDD	200GB
	사운드카드	온 보드 (Realtek High Definition Audio 드라이버)
	네트워크 카드	유선 LAN 100MB
	소프트웨어	운영체제
	구현 언어	Java SDK 1.6.0_22
	구현 도구	Netbeans IDE 6.9.1
	음성합성기	보이스웨어 TTS server

4.2 시스템 구성

도구는 <그림 4>과 같이 크게 사용성 시험자(위자드)가 사용하게 될 위자드 시스템과 참가자(실제 사용자)가 사용하게 될 서브젝트 시스템으로 구성된다. 이 두 개의 시스템은 WOZ 방식에 따라 서로 분리된 환경에서 운영된다.

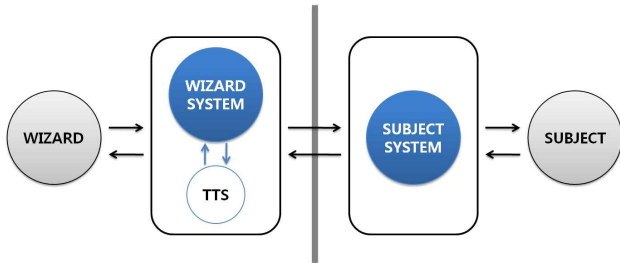


그림 4. 위자드-서브젝트 시스템 개념도

Figure 4. Conceptual diagram of the wizard-subject system.

시험자는 VUIDML 문서와 위자드 시스템을 이용하여 다이얼로그를 시뮬레이션하고 참가자(Subject)를 관찰하는 역할을 한다.

서브젝트(SUBJECT)는 사용자를 가리키며, 시험자가 시뮬레이션하는 시나리오에 따라 발화하여 대화를 진행한다.

위자드 시스템(WIZARD SYSTEM)은 프롬프트를 서브젝트 시스템에 전송하고 참가자의 발화를 모니터하고 데이터(발화내용 및 로그)를 저장한다. 위자드 시스템은 VUIDML의 다양한 프롬프트의 재생을 위하여 음성합성기(TTS)를 포함한다.

서브젝트 시스템(SUBJECT SYSTEM)은 시험자가 시뮬레이션하는 시나리오에 따라 프롬프트를 참가자에게 출력하고, 참

가자가 발화하는 명령어를 위자드 시스템에 전송한다.

4.2.1 위자드 시스템

다음 <그림 5>는 위자드 시스템의 동작을 위한 구성 요소들과 구성 요소들이 동작하는 흐름을 보여준다.

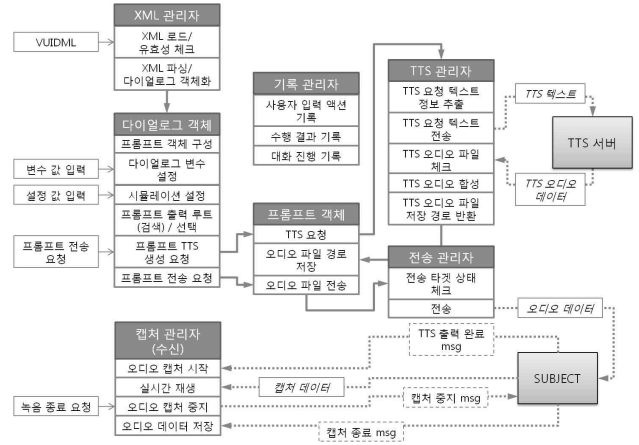


그림 5. 위자드 시스템 구성

Figure 5. Control flow of the wizard system

- XML 관리자

시험자가 VUIDML 문서를 선택하면 XML 관리자는 문서를 로드하고 XML 파싱 과정을 수행한다. XML 파싱을 통해 다이얼로그 객체로 엘리먼트들의 정보를 저장한다.

- 다이얼로그 객체

다이얼로그 객체는 프롬프트 출력을 위한 프롬프트 객체를 포함한다. 다이얼로그에 설정해야할 변수 값이 있다면 시험자로부터 변수 값을 입력받아야 한다. 여기서 변수는 다수의 참가자가 있는 경우, 참가자의 이름이나 호칭 등과 같이 참가자마다 달라하여야 하는 문자 값을 포함한다. 시험자가 다이얼로그 시뮬레이션을 위한 변수 설정 값을 입력하면 프롬프트 문장을 완성한다.

- 프롬프트 객체

프롬프트 객체는 시험자가 프롬프트 전송을 요청했을 때 TTS 관리자에게 텍스트를 전달하여 음성 생성을 요청한다. 생성된 음성은 전송 관리자를 통하여 전송한다.

- TTS 관리자

프롬프트 객체의 요청에 의해 TTS 서버로 TTS 변환을 요청한다. 변환된 오디오 데이터가 전송되면 전송된 데이터를 파일로 저장하고 경로를 프롬프트 객체에 반환한다.

- 전송 관리자

프롬프트 객체의 요청에 의해 오디오 데이터를 전송한다. 전송할 서브젝트 시스템의 상태를 점검하고 전송을 시작한다.

- 캡처 관리자

서브젝트 시스템으로부터 TTS 전송이 완료되었다는 메시지

를 수신하면 캡처 관리자는 서브젝트 시스템에서 전송하는 오디오 데이터 수신을 시작한다. 동시에 수신된 데이터를 실시간으로 시험자가 들을 수 있도록 재생한다. 시험자의 캡처 종료 요청에 의하여 수신을 종료하고 캡처 중지 메시지를 서브젝트 시스템으로 전송한다. 서브젝트 시스템으로부터 캡처 종료 메시지를 수신하면 수신된 오디오를 파일로 저장한다.

• 기록 관리자

시험자를 통해 입력된 액션을 기록하고 다이얼로그 시뮬레이션 결과를 기록한다

4.2.2 서브젝트 시스템

<그림 6>은 서브젝트 시스템의 동작을 위한 구성 요소들과 구성 요소들이 동작하는 흐름을 보여준다.

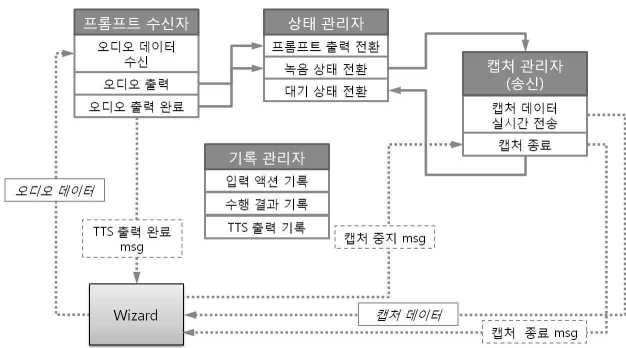


그림 6. 서브젝트 시스템 구성
Figure 6. Control flow of the subject system

• 프롬프트 수신자

위자드 시스템으로부터 오디오 데이터 전송이 시작되면 프롬프트 수신자가 이를 수신하여 오디오로 출력한다. 오디오 출력이 완료되면 위자드 시스템에게 TTS 출력 완료 메시지를 전송한다.

• 기록 관리자

TTS 출력 정보와 입력 액션 및 수행 결과를 기록한다.

• 캡처 관리자

시스템이 녹음 상태로 전환되면 참가자의 발화 내용 캡처를 시작하고 이 때 캡처된 오디오 데이터를 실시간으로 위자드 시스템으로 전송한다.

위자드 시스템으로부터 캡처 중지 메시지가 수신되면 캡처 관리자는 캡처를 종료하고 위자드 시스템으로 캡처 종료 메시지를 보낸다.

• 상태 관리자

시스템 동작에 따라 ‘프롬프트 출력’, ‘녹음 중’, ‘대기 상태’로 시스템의 상태를 전환한다.

4.3 주요 기능 및 동작 과정

구현한 도구에서 사용성 평가를 위해 제공하는 대표적인 기

능에는 ‘프롬프트 전송 및 출력’, ‘발화 녹음 및 저장’, ‘프롬프트 선호도 조사’가 있다. 프롬프트 전송 및 출력 기능은 시뮬레이션 진행을 위한 안내를 제공한다. 프롬프트 선호도 조사 기능은 후보 프롬프트를 제시한 후 사용자가 선택한 선호도 프롬프트를 기록한다. 발화 녹음 및 저장 기능은 사용자가 평가 도중 발화하는 명령어를 오디오 파일로 저장하여 인식 명령어 세트를 수집한다.

이 절에서는 앞에서 소개한 세 가지 주요 기능에 대한 개요와 각 기능이 동작하는 과정을 설명한다.

4.3.1 프롬프트 전송 및 출력

<그림 7>은 시험자와 참가자 사이에서 프롬프트가 처리되는 과정을 설명한다.

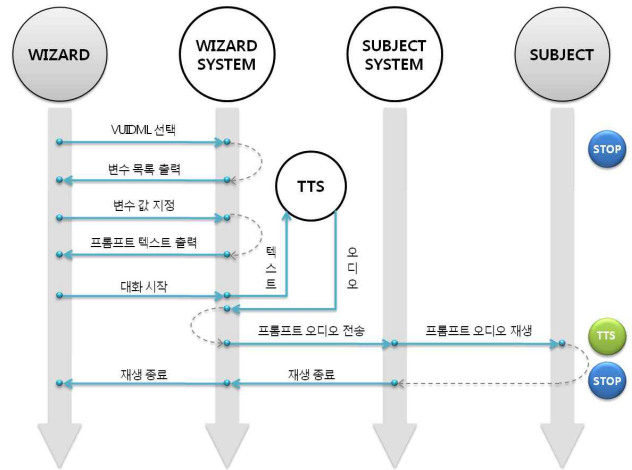


그림 7. 프롬프트 전송 및 출력 동작 과정
Figure 7. Prompt handling process

① VUIDML 선택

시험자는 시뮬레이션을 위한 VUIDML 문서를 선택한다. 시뮬레이션의 시작 시점에서만 필요한 작업이다.

② 변수 목록 출력

위자드 시스템에서는 VUIDML를 파싱하여 변수 목록을 시험자에게 출력한다. 시뮬레이션의 시작 시점에서만 필요한 작업이다.

③ 변수 값 지정

시험자가 변수 값을 지정하면 VUIDML 문서에 입력한 값이 적용된다. VUIDML 문서를 선택하여 로드될 때 또는 시뮬레이션 도중에 시험자의 요청에 의해서 이 작업이 수행된다.

④ 프롬프트 텍스트 출력

시스템은 변경된 VUIDML 문서의 프롬프트를 시험자에게 보여준다.

⑤ 대화 시작

시험자가 대화 시뮬레이션을 시작하면 프롬프트 텍스트가 TTS 모듈을 통해 오디오 데이터로 변환된다.

⑥ 프롬프트 오디오 전송

TTS 시스템이 변환한 오디오 데이터를 위자드 시스템으로 전송하면 위자드 시스템은 서브젝트 시스템으로 오디오 데이터를 전송한다.

⑦ 프롬프트 오디오 재생

서브젝트 시스템은 전송된 오디오 데이터를 참가자에게 재생한다.

⑧ 재생 종료

재생이 완료되면 서브젝트 시스템은 위자드 시스템에게 재생 종료를 알린다.

<그림 8>은 위자드 시스템 수행 화면을 나타내며 프롬프트를 전송하기 전 단계에서 시험자가 선택한 VUIDML 문서가 처리되어 화면에 보이는 모습이다. 시험자가 프롬프트를 선택한 후 ‘전송’ 버튼을 누르면, 선택한 프롬프트의 텍스트가 TTS에 의해 오디오로 변환되고 서브젝트 시스템으로 전송된다.



그림 8. 시험자용 도구 화면
Figure 8. GUI of the Wizard System

<그림 9>는 서브젝트 시스템의 GUI이며, 위자드 시스템에서 전송한 프롬프트가 수신되어 재생되고 있는 모습이다.

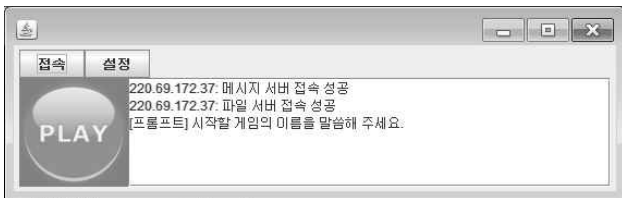


그림 9. 서브젝트용 도구 화면
Figure 9. GUI for Subjects

4.3.2 발화 녹음 및 저장

<그림 10>은 시험자로부터 발화 녹음 명령을 참가자에게 전달하여 사용자의 발화 음성을 오디오 데이터로 저장하고 모니터 하는 동작 과정을 설명한다.

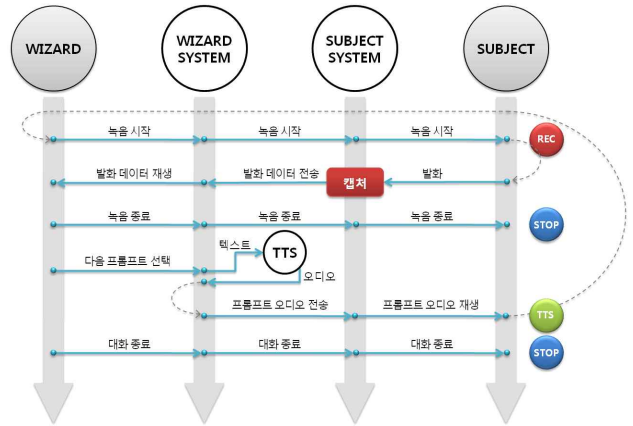


그림 10. 발화 녹음 및 저장 동작 과정
Figure 10. Subject's speech recording process

① 녹음 시작

‘프롬프트 재생’이 종료되었다는 것을 위자드 시스템이 인지하면 녹음을 시작하는 요청을 서브젝트 시스템으로 전송한다. 서브젝트 시스템은 녹음 시작을 시험자에게 알린다. 이 때 시스템은 ‘REC’ 상태로 전환된다.

② 발화

‘REC’ 상태로 전환되면 참가자는 발화를 시작한다. 이 발화는 사용자가 주어진 작업을 수행하기 위한 음성명령이다.

③ 발화 데이터 전송

‘REC’ 상태인 경우 서브젝트 시스템은 발화 내용을 오디오 데이터로 캡처하여 실시간으로 위자드 시스템에게 전송한다.

④ 발화 데이터 재생

위자드 시스템은 전송된 오디오 데이터를 재생하여 시험자가 참가자의 발화 내용을 모니터할 수 있도록 한다.

⑤ 녹음 종료

참가자가 발화를 마치면 모니터 중인 시험자는 녹음 종료 요청을 서브젝트 시스템으로 전송하면 서브젝트 시스템은 캡처를 종료한다. 이 때 시스템은 ‘STOP’ 상태로 전환된다.

⑥ 다음 프롬프트 선택

VUIDML에 설계된 시나리오에서 대화가 종료되지 않았다면 다음 프롬프트를 선택하여 ‘프롬프트 재생’을 한다. 다음 프롬프트는 위자드 시스템에 의해 자동으로 선택될 수 있다.

⑦ 대화 종료

VUIDML에 설계된 시나리오에서 대화가 완전히 종료되었으면 대화가 종료되었음을 서브젝트 시스템에게 알린다. 이 때 시스템은 ‘STOP’ 상태로 전환된다.

4.3.3 프롬프트 선호도 조사

<그림 11>은 후보 프롬프트 리스트에서 참가자의 선호도를 조사하는 과정을 설명한다.

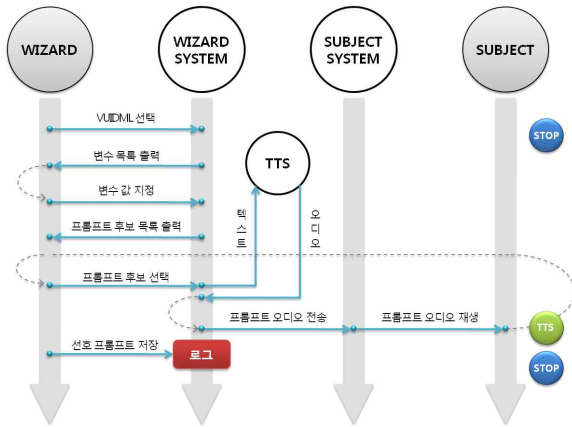


그림 11. 프롬프트 선호도 조사 동작 과정
Figure 11. Preferred prompt survey process

① VUIDML 선택 및 변수 입력

‘프롬프트 출력’에서와 마찬가지로 시험자는 시뮬레이션을 위한 VUIDML 문서를 선택하고, 필요한 경우, 참가자 호칭 등과 같은 변수 값을 입력한다.

② 프롬프트 후보 선택

시험자는 프롬프트 후보 목록 중에서 보이는 순서대로 선택해 참가자가 들을 수 있게 한다.

③ 프롬프트 오디오 전송

선택된 프롬프트는 TTS에 의해 오디오로 변환되고 위자드 시스템은 서버측 시스템으로 오디오 데이터를 전송한다.

④ 프롬프트 오디오 재생

위자드 시스템으로부터 전송이 시작되면 오디오를 재생한다. 재생 완료시 시험자는 다음 프롬프트 후보를 출력한다.

⑤ 선호 프롬프트 저장

시험자는 참가자에게 프롬프트 후보들을 각각 모두 들려주고 선호하는 프롬프트를 조사하여 응답 결과를 저장한다. 이 결과는 <그림 12>와 같은 구조의 XML 문서로 저장된다.

```

</choices>
  <choice promptID="ppt_1">
    <script index="1">
      홍길동씨 반갑습니다.[성함@./ask_name_1.wav]
    </script>
  </choice>
</choices>
  
```

그림 12. 프롬프트 선호도 조사 결과 XML 문서
Figure 12. Prompt preference result XML

5. 결론

본 논문에서는 명령제어형 음성 사용자 인터페이스의 사용성 향상을 위해 실시하는 초기 사용성 평가를 자동화하기 위한 도구를 설계하고 구현하였다. 초기 사용성 평가는 위자드 오브 오즈 방식으로 실시된다는 가정 하에 위자드 시스템과 서버측 시스템으로 각각 구성하였다.

설계한 사용성 평가 도구는 명령제어형 음성 인터페이스의 다이얼로그 시뮬레이션 시나리오가 요구된다. 이를 만족시키기 위해 기존에 개발되어있는 VoiceXML을 분석하여 수정하고 재구성하여 Voice User Interface Design Markup Language (VUIDML) 을 설계하였다.

본문에서는 설계한 VUIDML를 이용하여 작성한 시나리오의 예를 제시하고 시뮬레이션 되는 과정을 설명하였다. 또한 구현한 명령제어형 음성 사용자 인터페이스의 사용성 평가 도구의 구조와 동작 흐름을 기술하였다.

VUIDML과 사용성 평가 도구를 이용하여 음성 사용자 인터페이스의 사용성을 향상시키기 위한 명령어 수집과 사용자의 프롬프트의 선호도를 파악할 수 있다.

앞으로 수행하여야 할 연구는 본 논문에서 제안한 VUIDML 문서를 자동으로 생성하고 문법을 검토해주는 기능을 갖는 VUIDML 설계 도구, 대화 로직을 설계하기 위한 명령형 사용성 평가 도구의 대화 흐름 시뮬레이션 기능, 초기 사용성 평가 단계의 결과 분석과 VoiceXML과 같은 완성된 VUI 설계 문서 생성 기능을 갖춘 도구의 개발이다.

참고문헌

[1] Kim, J. H., Sim, J. H., Song, C. W., and Lee, J. H. (2008). "Development of speech recognition system based on user context information in smart home environment", *The Journal of the Korea Contents Association*, Vol. 8, No. 1, 328-338.
(김종훈, 심재호, 송창우, 이정현 (2008). "스마트 홈 환경에서 사용자 상황 정보 기반의 음성 인식 시스템 개발", 한국콘텐츠학회논문지, 제8권, 제1호, 328-338.)

[2] Kim, B. C., Jung, M. W., and Lee, G. B. (2007). "Current and Future trend of voice user interface technology", *The Magazine of the IEK*, Vol. 34, No. 6, 50-58.
(김병창, 정민우, 이근배 (2007). "음성 인터페이스 기술의 현황과 전망", 대한전자공학회지, 제34권, 제6호, 50-58.)

[3] Lee, Y. G. and Park, J. (2005). "Voice user interface technology", *Trend Analysis on Electronics and Telecommunications*, Vol. 20, No. 5.
(이윤근, 박준 (2005). "음성 인터페이스 기술", 전자통신동향분석, 제20권, 제5호.)

- [4] Kang, Y. H., Jung, S. W., Choi, J. H., and Jung, E. S. (2008). "Research on mobile handset's optimal voice recognition functions while driving", *Proceedings of 2008 Spring Conference of The Ergonomics Society of Korea*, 188-192.
(강윤환, 정성욱, 최재호, 정의승 (2008). "운전 상황에서 최적의 핸드폰 음성인식 기능 조사", 대한인간공학회 2008 춘계 학술대회 논문집, 188-192.)
- [5] Chae, H. S., Hong, J. Y., Jun, M. H., and Han, G. H. (2007). "A study on voice user interface or domestic appliance", *Korean Journal of the Science of Emotion and Sensibility*, Vol. 10, No. 1, 55-68.
(채행석, 홍지영, 전명훈, 한광희 (2007). "가전제품의 음성 인터페이스 디자인 적용에 대한 연구", 감성과학, 제10권, 제1호, 55-68.)
- [6] Nielsen, J. (1993). "Iterative user-interface design", *IEEE Computer*, Vol. 26, No. 11, 32-41.
- [7] Hong, K. H. (2008). "Usability Test Guidelines for Speech-oriented Multimodal User Interface", *Malsori-Journal of the Korean Society of Phonetic Sciences and Speech Technology*, No. 67, 103-120.
(홍기형 (2008). "음성기반 멀티모달 사용자 인터페이스의 사용성 평가 방법론", 말소리, 제67호, 103-120.)
- [8] Edgar, B. (2001). *The VoiceXML Handbook*, Gilroy, CA.: CMP Books.
- [9] Lee, M. J. and Hong, K. H. (2011). "Voice user interface design markup language for user-centered design of command and control voice interfaces", *HCI 2011 Proceedings*.
(이명지, 홍기형 (2011), "명령-제어 형식의 음성 인터페이스 사용자 중심 설계를 위한 XML", HCI 2011 학술대회.)
- [10] Barnum, C. M., Dragga, S. (2001). *Usability Testing And Research*, New York: Longman.
- [11] Dumas, J. S., Redish, J. C. (1999). *A Practical Guide To Usability Testing*, Wilmington, NC.: Intellect Ltd.
- [12] Greenberg, S. (1998). "Prototyping for design and evaluation", <http://grouplab.cpsc.ualgary.ca/saul/681/1998/prototyping/survey.html>.
- [13] Kwon, J. H. and Hong, K. H. (2006). "A survey on voice user interface design methodology", *Communications of KIISE*, Vol. 24, No. 1, 27-36.
(권지혜, 홍기형 (2006). "음성 사용자 인터페이스 설계 방법론", 정보과학회지, 제24권, 제1호, 27-36.)
- [14] Cohen, M. H., Giangola, J. P., and Balogh, J. (2004). *Voice User Interface Design*, Boston, MA.: Addison-Wesley.
- [15] Han, S. H., and Kim, B. S. (1999). "A usability evaluation method for speech recognition interfaces", *Journal of the Ergonomics Society of Korea*, Vol. 18, No. 3, 105-125.
(한성호, 김범수 (1999). "음성인식용 인터페이스의 사용편의성 시험 방법론", 대한인간공학회지, 제18권, 제3호, 105-125.)
- [16] Usability Professionals' Association. "What is user-centered design", http://www.upassoc.org/usability_resources/about_usability/what_is_ucd.html.
- [17] Smith, C., Thorp, J., and Henry, S. L. (2004). "Note on user centered design process(UCD)", <http://www.w3.org/WAI/EO/2003/ucd>.
- [18] Soucy, K. (2010). "Unmoderated, remote usability testing: good or evil?", <http://www.uxmatters.com/mt/archives/2010/01/unmoderated-remote-usability-testing-good-or-evil.php>.
- [19] Oh, K. T. and Lee G. P. (2004). "A study on remote usability test & evaluation for websites - with emphasis on the development of remote interaction observation & analysis software", *Journal of Korean Society of Design Science*, Vol. 17, No. 3, 147-156.
(오기태, 이건표 (2004). "웹 사이트 원격 사용성 테스트에 관한 연구 - 원격 사용자 인터랙션 관찰 및 분석 도구의 개발을 중심으로", 디자인학연구, 제17권, 제3호, 147-156.)
- [20] Tomlin, C. (2009). "24 usability testing tools", <http://www.usefulness.com/24-usability-testing-tools>.
- [21] Klemmer, S. R., Sinha, A. K., Chen, J., Landay, J. A., Aboobaker, N., Wang, A. (2000). "Suede: a wizard of oz prototyping tool for speech user interfaces", *UIST '00: Proceedings of the 13th annual ACM symposium on User interface software and technology*.

- 이명지 (Lee, Myeongji)

성신여자대학교 일반대학원 컴퓨터학과
서울시 성북구 동선동2가 169-1 미디어정보관 403호
Tel: 02-920-7525 Fax: 02-929-7525
Email: revioopp@gmail.com
관심분야: 음성사용자인터페이스, 사용성평가
현재 (주)보이스웨어 음성합성팀

- 홍기형 (Hong, Ki-Hyung) 교신저자

성신여자대학교 자연과학대학 IT학부
서울시 성북구 동선동2가 169-1 미디어정보관 403호
Tel: 02-920-7525 Fax: 02-929-7525
Email: kihyung.hong@gmail.com
관심분야: 음성사용자인터페이스, 사용성평가,
XML 기반 정보시스템
1998~ 현재 성신여자대학교 IT학부 교수