

# 스피치 인터페이스가 사용자-로봇 인터랙션에 미치는 영향

## -실버 로봇 사용자의 수행도 및 선호도 분석을 중심으로-

### Effects of Robot Speech Interface on Human-Robot Interaction

우명희

한국과학기술원 산업디자인학과

김명석

한국과학기술원 산업디자인학과

Woo, Myoung-Hee

Dept. of Industrial Design, KAIST

Kim, Myung-Suk

Dept. of Industrial Design, KAIST

• Key words: Human-Robot Interaction, Speech Interface, Silver Robot

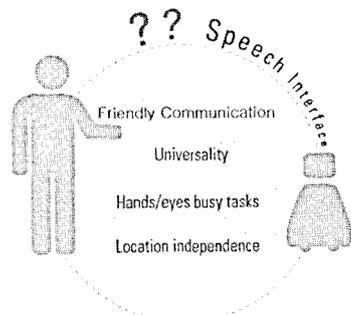
## 1. 서론

최근 단일 기능의 서비스 로봇이 이미 시장에 출시되고 있다. 이 서비스 로봇들은 물리적 혹은 지능적 업무들을 수행하기 위해 인간과 다양한 매체를 통해서 인터랙션 해야 한다. 단순히 물리적으로 만지거나, 스크린을 조작하는 등의 행동을 통해서도 가능하며, 언어를 통하여 사용자가 내린 명령에 따라 로봇이 응답하는 대화방식도 가능하다. 그 중 스피치 인터페이스(대화 음성)는 인간의 가장 자연스러운 의사전달 수단으로 로봇이 앞으로 인간과 공존하고 원활한 상호작용을 하기 위한 핵심적인 모달리티가 될 것이다. 지금까지 음성 합성 및 음성 인식 등의 기술적 접근이 대부분임에 비해 본 연구는 사용자 중심의 디자인 관점에서부터 스피치 인터페이스가 사용자-로봇 인터랙션에 미치는 영향을 살펴보고자 한다.

## 2. 이론적 고찰

### 2-1. 스피치 인터페이스의 개념 및 특징

스피치 인터페이스는 보이스 및 대화 방식을 포함하고 있는 음성 언어를 이용한 메인 청각인터페이스다. 특히, 자율적인 움직임을 갖고 사용자와 직접적인 인터랙션을 해야 하는 로봇에 있어서 시각 중심에서 벗어나 우리의 일상생활처럼 대화음성이라는 청각 요소도 함께 인터페이스로 구성되어져야 한다.



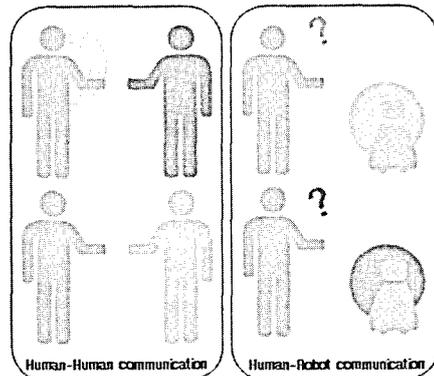
[그림 1] 로봇 스피치 인터페이스

리브스와 나스(Reeves and Nass, 1996)의 연구에 의하면 인간은 컴퓨터에 나오는 사람의 목소리의 억양이나 톤에 따라 컴퓨터 사용자가 사용하는 프로그램이 제시하는 내용에 대한 공신력의 차이가 보임을 보여주었다. 예를 들어 요리법을 가르쳐 주는 소프트웨어에는 프랑스어 억양이 심한 영어를 구사하는 목소리가 효과적이라는 것이다. 이처럼 스피치 인터페이스에 관련된 HCI

(Human-Computer Interaction) 선행연구들을 고찰해본 결과, 대화의 콘텐츠 및 보이스의 특징에 따라 사용자의 인지에 영향을 끼침을 알 수 있었다. 그 특징을 결정짓는 요인들은 기본 생리적인 화자의 성, 연령 및 문화적 배경 그리고 화자의 성격이나 감정상태 등으로 추출되었다.

### 2-2. 사용자-로봇 인터랙션의 커뮤니케이션학적 요소

로봇은 특성상 기존 제품 디자인상에서 중요하지 않았던 사회적 디자인 요소를 가지면서 사용자-로봇 커뮤니케이션에 관한 연구를 필요로 한다. 대표적인 대인 커뮤니케이션의 특징으로 인간은 외국인이나 어린아이를 상대할 때 커뮤니케이션 파트너에 맞추어 적응하는 경향이 있는데 이런 사용자의 태도는 로봇과의 인터랙션 설계에 매우 중요한 고려사항이라 할 수 있다. (Thora Tenbrink)



[그림 2] 휴먼-로봇 커뮤니케이션의 이해

## 3. 실증적 연구

본 연구의 실험은 앞서 언급한 스피치 인터페이스의 특징 요인 중 커뮤니케이션 파트너 구분에 초점을 두어 성과 연령별로 설정하였으며, 로봇의 태스크 수행에 따른 사용자 만족도와 선호도로 그 영향을 관찰하였다.



[그림 3] 실험 개요

### 3-1. 연구 대상

실험 대상은 서비스 로봇 시장의 많은 수요층대를 가진 실버

로봇 사용자로 하였다. 이들은 또한 신체적으로 시각이 우선 퇴화함으로써 청각적 인터랙션에 민감하기 때문이다. 본 실험은 유성구 노인 복지회관을 직접 방문하여 청각에 이상이 없는 학습능력이 가능한 노인들을 대상으로 하였다.

### 3.2. 연구 설계

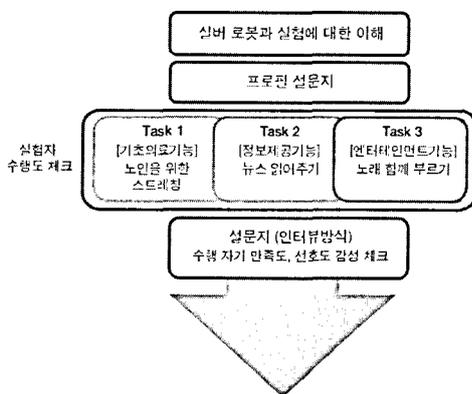
실험을 위한 음성녹음은 남녀 연극배우를 이용하여 KAIST Music and Sound Design Lab 에서 동일 음질 수준에서 제작 (44KHz)되었으며 음성의 평균크기도 일정하게 만들었다. 그리고 스크립트는 연령대의 가장 전형적인 특징을 반영하도록 하였다. 이는 실험과정에서 피험자에게 6개의 목소리를 구분하게 함으로써 음성녹음에 대한 확인 검증을 실시했다.

독립변인	변인 수준	수준결정근거	측정치	
			선호도	수행도
로봇 대화 음성	연령	어린 아이	[선호만족도] 5점 척도법	[수행여부] 실험자 관측 [수행만족도] 5점 척도법
		20대	실버 로봇 사용자의 사회적 관계 (주요 커뮤니케이션 파트너) 중 구분가능한 스피치인터페이스 연령대	
		50대	편한 유쾌한 매력있는 친절한 명랑한 따뜻한 책임감있는 똑똑한 강한 믿음만한	
	성	남	스피치 인터페이스 구분 우선 생리요인	
		여	사용자 특성에 따른 분석을 위한 데모그라픽 기본 요소	
		성	남	
피험자	남			
	여			

[표 1] 실험 요인 overview

실험 과정은 아래와 같다.

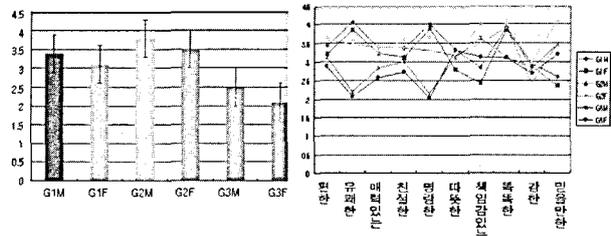
- 1차 실험 : Task1을 수행하지 않고 6개의 목소리를 들은 후 가장 선호하는 목소리와 감성을 체크하도록 하였다.
- 2차 실험 : Task 1,2,3 각각 30명씩 참가하여 Task를 수행하였다. 단, 학습효과 및 집중효과의 바이어스를 방지하기 위해 6개의 목소리를 randomizing 하였다.



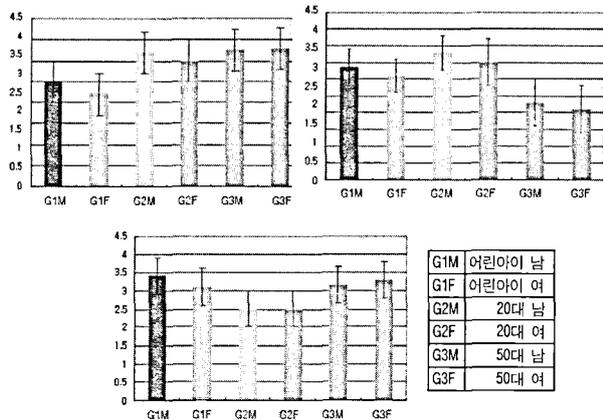
[그림 4] 2차 실험 Process

0) 본 연구는 정보통신부 지원으로 수행하는 대학 IT연구센터 육성. 지원 사업 중 Information Technology Research Center (ITRC)의 지원으로 수행되었음.

### 3.3. 연구 결과 및 분석



[그림 5] 1차 실험 선호도 및 감성평가 결과



[그림 6] 2차 실험 선호도 결과 (시계방향으로 Task1, Task2, Task3)

태스크 수행을 실시하지 않은 1차 실험 경우 20대 남녀의 대화 음성이 가장 선호된 반면 2차 태스크를 수행하면서 다른 양상을 보였다. 태스크1과 같은 교육기능에 있어 노인들은 50대의 편안함과 비교적 느린 대화속도에 큰 만족감을 보이며 태스크를 수행하였고 어린 아이 음성엔 신뢰감을 느끼지 못하였다. 한편 태스크2 뉴스를 들 때는 20대의 명랑하고 책임감이 느껴지는 음성을 선호, 50대에겐 지루함마저 느꼈다. 태스크3 엔터테인먼트 기능에서는 개인적 성향에 확연한 차이를 보이며 어린 아이를 좋아하는 부류와 공감대를 형성하는 50대 양쪽 모두 유사한 선호도를 보였다.

### 4. 결론 및 향후 연구

본 연구를 통해 로봇 스피치 인터페이스의 성별 연령 차이가 로봇의 기대역할을 결정지며, 그 기대역할에 따라 분명히 사용자의 태도에 유의한 영향을 끼침을 알 수 있었다. 이는 로봇에 적합한 스피치 인터페이스를 찾기 위한 사용자 참여 평가의 일환으로 사람과 로봇의 보다 효과적인 인터랙션을 유도하기 위한 기초 자료로 제시한 바이다. 향후 음성의 속성에 기반하여 보다 객관적으로 검증할만한 연구를 지속적으로 연계 추진하고자 한다.

#### 참고문헌

- 김현진, 로봇 디자인에서 사용자의 사회적 니즈 추출에 관한 연구, 한국디자인학회 봄 학술대회, 2003
- Thora Tenbrink, Communicative Aspects of Human-Robot Interaction, Proceedings of the International Futuristic Conference on Language Development, 2001
- Clifford Nass, Wired for Speech, 2005