

## 대화형 에이전트의 오류 상황에서 사회적 전략 적용: 사전 양해와 사과를 이용한 사례 연구

Applying Social Strategies for Breakdown Situations of Conversational Agents:  
A Case Study using Forewarning and Apology

이유미<sup>\*,\*\*</sup> · 박선정<sup>\*\*</sup> · 석현정<sup>\*\*†</sup>  
Yoomi Lee<sup>\*,\*\*</sup> · Sunjeong Park<sup>\*\*</sup> · Hyeon-Jeong Suk<sup>\*\*†</sup>

\*경일대학교 디자인학부  
\*School of Design, Kyungil University

\*\*한국과학기술원 산업디자인학과  
\*\*Department of Industrial Design, KAIST

### Abstract

With the breakthrough of speech recognition technology, conversational agents have become pervasive through smartphones and smart speakers. The recognition accuracy of speech recognition technology has developed to the level of human beings, but it still shows limitations on understanding the underlying meaning or intention of words, or understanding long conversation. Accordingly, the users experience various errors when interacting with the conversational agents, which may negatively affect the user experience. In addition, in the case of smart speakers with a voice as the main interface, the lack of feedback on system and transparency was reported as the main issue when the users using. Therefore, there is a strong need for research on how users can better understand the capability of the conversational agents and mitigate negative emotions in error situations. In this study, we applied social strategies, “forewarning” and “apology”, to conversational agent and investigated how these strategies affect users' perceptions of the agent in breakdown situations. For the study, we created a series of demo videos of a user interacting with a conversational agent. After watching the demo videos, the participants were asked to evaluate how they liked and trusted the agent through an online survey. A total of 104 respondents were analyzed and found to be contrary to our expectation based on the literature study. The result showed that forewarning gave a negative impression to the user, especially the reliability of the agent. Also, apology in a breakdown situation did not affect the users' perceptions. In the following in-depth interviews, participants explained that they perceived the smart speaker as a machine rather than a human-like object, and for this reason, the social strategies did not work. These results show that the social strategies should be applied according to the perceptions that user has toward agents.

**Key words:** Conversational Agents, Smart Home, Agent-based Interface, Social Interface

---

※ 이 논문은 2017년도 정부(과학기술정보통신부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임 (No.2016-0-00564, 사용자의 의도와 맥락을 이해하는 지능형 인터랙션 기술 연구개발).

† 교신저자 : 석현정 (KAIST 산업디자인학과)

E-mail : color@kaist.ac.kr

TEL : 042-350-4523

FAX : 042-350-4502

## 요 약

음성인식 기술의 비약적 발전으로 최근 몇 년 사이 대화형 에이전트는 스마트폰, 인공지능 스피커 등을 통해 널리 보급되었다. 음성인식 기술의 인식의 정확도는 인간의 수준까지 발전하였으나, 여전히 말의 의미나 의도를 파악하는 것과 긴 대화를 이해하는 것 등에는 한계를 보이고 있다. 이에 따라 사용자는 대화형 에이전트를 사용함에 있어 다양한 오류 상황들을 경험하고 있으며 이는 사용자 경험에 부정적인 영향을 미칠 수 있다. 또한, 목소리를 주 인터페이스로 하는 인공지능 스피커의 경우, 대화형 에이전트의 기능 및 한계에 대한 피드백의 부족은 지속적 사용을 저해하는 요소로 꼽히고 있다. 따라서 사용자가 대화형 에이전트의 기능 및 한계를 보다 잘 이해하고 오류 상황에서 부정적인 감정을 완화할 수 있는 방안에 대한 연구에 대한 필요성이 높으나, 아직 관련 연구는 미비한 실정이다. 이에 본 연구에서는 사회적 전략 중 ‘사전 양해’와 ‘사과’를 대화형 에이전트에 적용하고 이러한 전략이 사용자가 에이전트에 대해 느끼는 인식에 어떠한 영향을 미치는지 조사하였다. 이를 위해 사전 양해와 사과 여부를 나누어 사용자가 대화형 에이전트와 대화하는 데모 영상을 제작하였고, 실험참가자들에게 영상을 보여준 뒤 느끼는 호감도와 신뢰도를 설문을 통해 평가하도록 하였다. 총 104명의 응답을 분석한 결과, 문헌조사를 토대로 한 우리의 예상과는 상반되는 결과를 얻었다. 사전 양해는 오히려 사용자에게 부정적인 인상을 주었으며, 특히 에이전트에 대한 신뢰도에 부정적인 영향을 주었다. 또한 오류 상황에서의 사과는 사용자가 느끼는 호감도나 신뢰도에는 유의미한 영향을 미치지 않았다. 심층 인터뷰를 통해 원인을 파악한 결과, 실험참가자들은 인공지능 스피커를 사람과 같은 인격체보다는 단순한 기계에 가깝다고 인식했기 때문에 인간관계에 작용하는 사회적 전략이 영향력을 발휘하지 못한 것으로 해석된다. 이러한 결과는 사용자가 에이전트를 얼마나 기계, 혹은 사람에 가깝게 인식하는지에 따라 오류 상황에 대한 에이전트의 대처 방식 또한 달라져야 함을 보여준다.

**주제어:** 대화형 에이전트, 스마트홈, 에이전트 기반 인터페이스, 사회적 인터페이스

## 1. 서론

음성인식 기술과 인공지능 기술의 발전으로 최근 몇 년 사이 대화형 에이전트(Conversational Agent)는 스마트폰, 인공지능 스피커 등을 통해 일상생활에 널리 보급되고 있다. Apple의 Siri, Google의 Google Now, Microsoft의 Cortana, Amazon의 Alexa, 삼성의 빅스비 등에 이르기까지 다양한 대화형 에이전트들이 등장하였으며, 이러한 에이전트들은 제품의 부가 기능이 아니라 운영 시스템을 작동하는 주 인터페이스로써 자리잡고 있다. Apple의 HomePod, Amazon의 Echo, Google의 Google Home, SKT의 누구, 네이버의 웨이브, KT의 기가지니 등에 이르기까지 국내외 주요 IT 업체들은 대화형 에이전트를 탑재한 인공지능 스피커를 경쟁적으로 출시하고 있다. 대화형 에이전트는 정보 검색, 음악 재생 등의 기본 기능에서부터 쇼핑에 이르기까지 다양한 기능을 수행하고 있으며, 에이전트와의 상호작용은 사용자의 경험에 크게 영향을 미치고 있다.

이와 같은 대화형 에이전트의 보급은 급속도로 발전하고 있는 음성인식 기술 덕분이다. Microsoft는 5.1%, Google은 4.9%의 인식 오류를 보이고 있으며, 이는 인간의 오류 비율과 유사한 수치이다(Price, 2017; Li, 2017). 이러한 약 95%의 정확도 수치는 2013년과 비교하였을 때 약 20% 정도 증가한 수치이다. 하지만 이처럼 빠른 발전 속도에도 불구하고 음성인식 기술은 아직 완벽하지 않다. 상업적으로 이용 가능한 시스템의 경우 음성인식 오류 비율은 12% 정도로 추정되며, 말의 의미와 의도를 이해하는 것에는 아직 한계를 보이고 있다. 또한 긴 대화를 이해함에 있어서도 한계를 보이고 있다(Jarnow, 2016). 이러한 한계는 대화형 에이전트를 사용함에 있어 다양한 오류 상황들을 야기하고 있으며, 이는 사용자의 경험에 부정적으로 작용할 수 있다.

목소리 위주로 의사소통하는 인공지능 스피커의 경우, 시스템의 한계 자체의 문제뿐만 아니라 이에 대한 피드백 부족은 사용자의 대화형 에이전트 이용을 막는 주요 이유로 꼽히고 있다. Luger & Sellen(2016)의

연구에서 사용자들은 에이전트가 무엇을 할 수 있는지 파악하기 어렵다고 응답했고, 이는 에이전트의 무궁무진한 가능성에 압도되거나 에이전트의 기능이 매우 제한적이라고 느끼는 등의 상반된 반응들을 불러 일으키며 지속적 사용을 저해하는 요소로 작용하였다. 이러한 피드백 부족이 야기하는 사용자의 부정적 감정은 사용자의 음성인식 기술에 대한 사전 지식이 적을수록 크게 나타났다.

따라서 더 나은 대화형 에이전트의 사용자 경험을 위하여 대화형 에이전트의 더 나은 피드백과 오류 상황에 대응하는 방안에 대한 연구가 매우 필요하다. 하지만 대화형 에이전트에 대한 높은 관심에도 불구하고 관련 연구는 아직 매우 미흡한 실정이다. 본 연구에서는 사용자가 대화형 에이전트의 기능과 한계를 잘 이해할 수 있도록 돕고, 오류 상황에서 사용자의 부정적 감정을 완화할 수 있는 방안에 대해 탐색하고자 한다. 그 방안으로써 이전의 서비스 로봇의 연구를 통해 사용자에게 미치는 효과가 검증된 방식인 ‘사전 양해’와 ‘사과’를 대화형 에이전트에 적용하여 그 효과를 검증하고자 한다.

본 연구는 기술적 측면에서 오류 상황을 줄이기 위한 연구가 아니라, 현재의 기술적 한계로 인해 어쩔 수 없이 발생할 수밖에 없는 오류 상황에 대해 대처하는 방안에 대한 연구이다. 본 연구를 통해 사용자의 관점에서 대화형 에이전트에 대해 가지고 있는 기대와 실제 경험에 대하여 더 깊게 이해할 수 있을 것이다. 본 연구의 결과를 통해 완벽하지 않은 음성인식 기술을 사용함에 있어 사용자가 에이전트에 대한 적절한 기대를 설정할 수 있도록 돕고, 오류 상황이 생겼을 때 부정적 감정을 완화할 수 있는 방안을 제안할 수 있으리라 기대한다. 이를 통해 음성인식 기술의 사용에 있어 발생할 수 있는 부정적 감정을 완화하고 보다 긍정적인 사용자 경험을 디자인하는데 기여할 수 있을 것이다.

## 2. 연구 목적

본 연구는 대화형 에이전트의 오류 상황에 대한 사전 양해가 사용자의 에이전트에 대한 인식에 미치는

영향을 이해하고, 오류 상황에서의 정중한 사과가 사용자의 부정적 감정 완화에 미치는 영향을 검증하는 것을 목적으로 한다. 이 연구의 가설은 다음과 같으며, 이 가설은 문헌 연구 결과를 통해 도출하였다.

가설 1: 대화형 에이전트의 발생 가능한 오류 상황에 대한 사전 양해는 사용자가 에이전트에 대해 느끼는 호감도와 신뢰도에 긍정적인 영향을 줄 것이다.

가설 2: 오류 상황에서 대화형 에이전트의 정중한 사과는 사용자가 에이전트에 대해 느끼는 부정적 감정을 완화하는데 긍정적인 영향을 끼칠 것이다.

## 3. 관련 연구

본 연구는 사람과 사람 간의 상호작용에서 적용되는 사회적 전략을 사람과 대화형 에이전트 간의 상호작용에 적용시키고 그 효과를 검증하고자 하는 연구이다. HCI 분야에서는 컴퓨터를 사회적 행위자(social actor)로 인식하고 컴퓨터와 사람과의 관계에 어떠한 사회적 전략을 적용할 수 있는지에 대한 수많은 연구가 이루어져 왔다. 또한, 최근에는 스마트 홈 환경에서 목소리 기반의 대화형 에이전트가 어떻게 인식되고 어떠한 역할을 하는지에 대한 연구들도 활발하게 이루어지고 있다.

### 3.1. 사회적 행위자로서의 컴퓨터

Nass & Moon(2000)의 연구는 사용자들이 대화 상대가 컴퓨터인 사실을 아는 경우에도 자연스럽게 상대에 대해 사람과 사람 사이에 적용되는 사회적 규칙들에 대한 기대를 가진다는 사실을 밝혔다. 사람들은 컴퓨터에게 사람과 사람 사이에 기대되는 기본적인 예의를 갖추었으며, 사람들의 반응은 에이전트의 성격에 따라 다르게 나타났다. 또한 에이전트의 행동이 얼마나 실재처럼 느껴지는지에 따라 사람들의 반응이 다르게 나타나기도 하였다(Appel et al., 2012; von der Pütten et al., 2010). 이러한 사실들을 바탕으로 가상

에이전트의 디자인에 인간관계에서 작용하는 사회적 전략(social strategies)을 적용하는 많은 연구가 이루어져 왔다. Bickmore & Cassell(2001)의 연구에서는 에이전트가 사용자와 가벼운 잡담을 하도록 하여 에이전트에 대한 신뢰를 쌓도록 하였고, Cassell(2000), Park(2007), Song et al.(2008)의 연구에서는 시선 처리나 몸동작 등과 같이 사람처럼 느껴질 수 있는 행동들을 가상의 에이전트에 적용하였다. 이러한 사회적 전략들은 로봇에도 적용되었다. Lee et al.(2010)의 연구에서는 인간의 모습과 유사한 서비스 로봇의 실수 상황에 대해 사전 양해, 사과, 보상, 제안 등의 사회적 전략을 적용하였고 모두 사용자의 로봇의 실수에 대한 부정적 감정을 완화하는데 효과를 보였다.

위와 같이 가상의 에이전트나 로봇에 대해 사람들이 어떻게 인식하고 반응하며, 에이전트의 행동과 대화에 어떠한 사회적 전략을 적용할 수 있는 지에 대한 연구는 많이 이루어져 왔다. 하지만 대부분의 연구가 에이전트의 얼굴 표정, 몸의 움직임, 말투 등을 복합적으로 다루었으며, 목소리 기반의 대화만을 단일 인터페이스로 다룬 연구는 충분히 이루어지지 않았다. 대화형 에이전트가 내장된 인공지능 스피커의 경우, 사용자가 에이전트에 대한 사회적 단서를 수집할 여지가 로봇이나 가상 에이전트 등에 비해 매우 제한적이므로 사회적 전략에 대한 효과는 다를 수 있으며 이에 대한 더 많은 연구가 필요할 것으로 보인다.

### 3.2. 대화형 에이전트의 성격과 사용자 인식

스마트홈 환경에서의 대화형 에이전트에 대한 연구도 최근 활발하게 이루어지고 있다. Mennicken et al.(2016)의 연구는 스마트홈 환경에서 사용자가 선호하는 대화형 에이전트의 성격을 파악하는 연구를 진행하였다. 성실하고 친절하고 차분한 성격과 외향적이고 쾌활한 성격을 비교한 결과, 실험참가자의 61%는 전자를 선호하였으며, 전자의 경우는 후자에 비해 사용자가 주도권을 가진 느낌을 주었다. Luger & Sellen(2016)의 연구에서 인공지능 스피커 사용자들의 대부분은 에이전트와의 장난 같은 대화를 통해 사용을 시작하였고, 에이전트의 유머 있는 답변들은 에이전트에 대한 친근감을 높이고 이후의 지속적인 사용을 유

도하는 역할을 하였다. Purington et al.(2017)의 연구에서는 사용자가 대화형 에이전트를 인격체로 인식하고 대화 상대나 친구와 같이 사회적 관계 정도가 높은 용도로 이용할수록 제품에 대한 만족도가 높아지는 상관관계를 보였다.

위와 같이 대화형 에이전트 관련 연구들을 살펴본 결과, 사회적 행위자로서의 대화형 에이전트의 특성과 이러한 특성이 사용자 경험에 미치는 영향의 중요성이 언급되고 있다. 하지만 단일 인터페이스로서의 대화형 에이전트에 사회적 전략을 적용하는 방안에 대한 연구는 그 중요성에도 불구하고 아직까지 충분히 이루어지지 않았음을 확인할 수 있었다. 이에 본 연구에서는 대화형 에이전트의 오류 상황에 주목하여 사람과 사람사이에 작용하는 사회적 전략 중, 서비스 로봇 연구(Lee et al., 2010)를 통해 검증된 방식인 ‘사전 양해’와 ‘정중한 사과’를 대화형 에이전트에 적용하여 그 효과를 사용자 평가를 통해 검증하고자 한다.

## 4. 오류 상황에서의 대화형 에이전트의 사전 양해와 사과의 효과 탐색

### 4.1. 연구 목표

본 실험에서는 대화형 에이전트의 한계에 대한 사전 양해와 오류 상황에서의 정중한 사과가 사용자의 에이전트에 대한 호감도 및 신뢰도에 미치는 영향을 탐색하고자 하며, 이를 위해 설문조사와 심층인터뷰를 실시하였다.

### 4.2. 실험참가자

본 실험의 참가자는 온라인 커뮤니티 및 개인 네트워크를 통하여 모집하였다. 총 119명이 설문에 참여하였으나, 이 중 모든 설문 항목에 대해 일률적으로 응답한 15명의 참가자의 경우 불성실한 응답으로 간주하여 분석에서 제외하였다. 15명의 불성실한 응답자를 제외한 104명의 참가자 중 남성은 34명(32.7%), 여성은 70명(67.3%)이었으며, 평균 나이는 24.63세였다. 또한 우리는 실험에 앞서 실험참가자들의 IT기술



과 음성인식 서비스 관련 사전 지식을 조사하였다. IT 기술 관련 사전 지식의 경우 자주 접하며 잘 알고 있다(32.7%), 가끔 접해본 적이 있으나 잘 알지 못한다(56.7%), 거의 접해본 적 없다(10.6%)로 나타났으며, 음성인식 서비스 관련 사전 지식의 경우 잘 알고 있다(74.0%), 들어본 적이 있으나 잘 모른다(24.0%), 들어본 적이 없다(1.9%)로 나타났다.

#### 4.3. 자극물 구성: 사전 양해와 사과 여부에 따른 영상

본 실험은 네이버의 인공지능 스피커 ‘웨이브’를 이용하여 대화형 에이전트 ‘샬리’와 사용자가 대화하는 가상의 상황이 담긴 데모 영상을 시청하고(Fig. 1 참조) 이에 대한 설문에 응답하는 방식으로 설계하였다. 이러한 데모 영상을 통한 시나리오 기법은 HCI 분야에서 흔히 적용되는 연구 방법이다. Woods et al.(2006)의 연구는 실험참가자가 직접 제품을 사용해보는 사용자 조사 방법이 지닌 한계를 지적하며(상대적으로 적은 실험참가자 수, 완벽하게 통제하기 어려운 실험 환경 등), 그 대안으로써 데모 영상을 통한 시나리오 기법을 제안한다. 특히 예상과 상반된 연구 결과를 얻는 경우가 흔한 탐색적 성격을 띠는 많은 HCI 연구의 특성 상, 시나리오 기법은 초기 가설을 검증하기 위해 불필요한 노력을 최소화할 수 있는 대안이 될 수 있다. 다만, 실험참가자의 직접 사용 실험과 데모 영상을 통한 실험 간의 얼마나 동일한 결과를 얻을 수 있는가가 관건이나, Woods et al.(2006)는 직접 사용 실험과 데모 영상 실험을 비교하여 실험참가자의 반응이 정도의 차이는 있으나 높은 수준의 일치성을 보이고 있음을 증명하였다. 본 실험에서도 대화형 에이전트에 사회적 전략 적용 가능성 여부를 파악하기 위한 초기 탐색적 연구임을 고려하였을 때, 이러한 데모 영상을 통한 시나리오 기법이 더 적합하다고 판단하여 적용하였다. 서비스 로봇의 사전 양해와 정중한 사과의 효과를 검증한 사전 연구(Lee et al., 2010)에서도 동일하게 데모 영상을 통한 시나리오 기법이 적용되었다.

본 실험에서 사용한 데모 영상은 두 가지 과업을 기반으로 구성하였는데, 각각은 데모 영상 1과 데모 영상 2로 구분하였다. 데모 영상 1은 사용자가 지시한 과업을 성공적으로 수행하는 내용으로, 에이전트의

사전 양해 여부에 따라 FO(사전 양해 포함)와 FX(사전 양해 미포함)의 두 가지 버전이 준비되었다. 반면 데모 영상 2에서는 사용자가 지시한 과업에 대해 에이전트가 실수를 하게 되고, 실수에 대한 사과의 여부에 따라 AO(사과 포함)와 AX(사과 미포함)의 두 가지 버전으로 구분하였다. 데모 영상은 소리와 더불어 자막을 함께 보여주어 실험참가자가 충분히 내용을 숙지할 수 있도록 하였다.

#### 데모 영상 1



Fig. 1. Video demonstrations presented in size of 1600 by 1200 pixels

먼저 데모 영상 1의 경우 에이전트의 간단한 자기 소개로 시작되는데, 모든 실험참가자들이 같은 내용을 듣게 된다:

샬리: “안녕하세요, 저는 에이전트 샬리예요 저는 사용자님과 대화하며 일상적인 이야기를 나누거나, 사용자님이 부탁하는 간단한 일을 수행할 수 있어요.”

이어서 사전 양해를 부연적으로 하는 경우와 그렇지 않은 경우로 나뉜다. 사전 양해를 덧붙이는 경우는 다음의 내용이 추가된다. 이 사전 양해가 추가되는 경우가 FO, 그렇지 않은 경우가 FX에 해당된다:

샬리: “하지만 저는 가끔 사용자님의 말을 이해하지 못하거나, 사용자님이 부탁한 일을 정확히 수행하기 어려울 수도 있어요.”

이렇게 사전 양해의 여부에 따른 FO, FX 두 가지 종류의 데모 영상 1은 다음의 내용을 동일하게 이어 가며 종료된다.

샬리: “제가 필요하실 때에는 저의 이름을 불러주세요.”

영상 속 사용자: “샬리아, 오늘 날씨 어때?”

샬리: “오늘 최고기온은 13도, 최저기온은 2도입니다.”

영상 속 사용자: “샬리아, 요즘 마음이 쓸쓸해. 유행하는 노래 틀어줘.”

샬리: “힘든 일이 있으신가 봐요. 요즘 유행하는 따뜻한 노래를 틀어 드릴게요.”

(노래 재생)

### 데모 영상 2

데모 영상 2에서는 에이전트가 사용자가 지시한 작업을 제대로 수행하지 못하는 상황이 재현된다. 우선 아래의 대화가 이어진다.

영상 속 사용자: “샬리아, 근처 맛집 알려줘.”

샬리: “근처 맛집에는 원조 할매국밥, 플라워가든 비스트로, 중화반점이 있어요.”

영상 속 사용자: “플라워가든 비스트로 예약할 수 있어?”

여기까지 이어진 대화 후, 사과를 하는 영상 AO에서는 다음과 같은 대화가 이어진다:

샬리: “죄송해요 무슨 말인지 잘 이해하지 못했어요. 다음에는 좀 더 귀 기울여 들어볼게요.”

반면, 사과를 하지 않는 영상 AX에서는 다음의 대화가 이어진다:

샬리: “무슨 말인지 잘 이해하지 못했어요.”

실험은 데모 영상 1과 데모 영상 2에서 각각 시청하는 종류에 따라 다음과 같이 구분되었다.

그룹 1: FO(사전 양해 O)-AO(사과 O)를 시청

그룹 2: FO(사전 양해 O)-AX를 시청

그룹 3: FX-AO(사과 O)를 시청

그룹 4: FX-AX를 시청

실험참가자는 네 개의 그룹 중 한 개에 무작위로 배정되었고, 본인이 어떠한 그룹에 소속되었는지 알지 못하는 상태에서 실험에 참여하였다. 실험에 성실하게 응답한 총 104명의 응답자의 그룹별 구성은 그룹 1 25명, 그룹 2 32명, 그룹 3 22명, 그룹 4 25명이었다. 각 그룹별 성별 구성은 카이검정 goodness-of-fit 테스트 기준 1:1과 0.05 유의수준에서 차이가 없었으며 네 그룹 연령 평균값에도 그룹 간 일원분산분석 결과 유의수준 0.05 기준에서 통계적 차이가 없는 것으로 나타났다.

### 4.4. 연구 과정

실험은 기본 조사, 데모 영상 1 시청 후 중간 설문, 데모 영상 2 시청 후 최종 설문의 순서로 진행되었다. 기본 조사에서는 실험참가자의 성별과 나이와 같은 기본 정보 및 IT기술과 음성인식 서비스에 대한 사전 지식의 정도를 조사하였으며, 중간 설문과 최종 설문은 각 데모 영상을 시청한 후 에이전트에 느끼는 친근감 및 신뢰도에 관한 질문으로 구성되었다.

중간 설문과 최종 설문은 동일한 문항으로 구성되었으며, 각각 호감도를 측정하는 질문 3개, 신뢰도를 측정하는 질문 3개로, 총 6개의 문항으로 구성되었다. 각 질문에 대한 응답은 5점 단위의 리커트 척도로 측정하였다. 호감도에 대한 질문은 ‘에이전트와 대화해보고 싶다’, ‘에이전트가 친근하게 느껴진다’, ‘에이전트는 내 말을 잘 들어줄 것 같다’로 에이전트에 대해 느끼는 호감 정도를 평가하도록 하였고, 신뢰도의 경우 ‘에이전트는 내 말을 잘 이해할 수 있을 것 같다’, ‘에이전트는 내가 부탁한 일을 잘 수행할 것 같다’, ‘에이전트는 일상생활에서 유용하게 사용될 수 있을 것 같다’와 같이 에이전트의 수행능력에 대한 신뢰의 정도를 평가하도록 하였다.

이와 같이 실험 전체 구조는 사전 조사 이후 데모 영상 1에 대한 중간 평가 1회, 그리고 데모영상 2에 대한 최종 평가 1회 등으로 Fig. 2와 같이 구성되었다. 데모 영상 1과 2에서 사전 양해 및 사과의 여부에 따라 네 개의 실험참여자 그룹이 구분되었다.

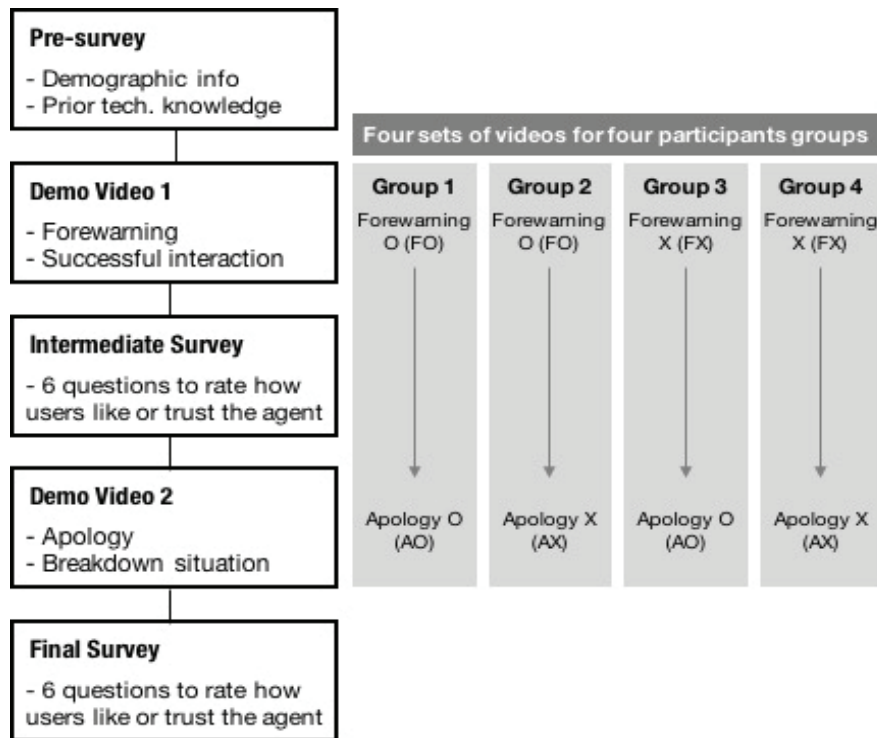


Fig. 2. Flow of the study

## 5. 실험 결과

실험참가자들의 응답을 토대로 대화형 에이전트의 한계에 대한 사전 양해와 오류 상황에서의 에이전트의 정중한 사과가 에이전트에 대한 호감도와 신뢰도에 미치는 영향을 분석하고자 하였다. 이를 위해 먼저 실험참가자들의 설문 응답을 정량적으로 분석하여 사전 양해와 사과가 미치는 영향 정도를 파악하고자 하였으며, 이어진 심층 인터뷰를 통해 정량적 데이터가 의미하는 바를 구체적으로 이해하고자 하였다.

### 5.1. 설문 응답에 대한 통계 분석 사전 양해가 호감도와 신뢰도에 끼치는 영향

우선 각 그룹 별로 중간 설문 및 최종 설문에서 호감도와 신뢰도에 대한 평균 점수를 산출하였다. Table 1에 제시된 바와 같이, 중간 설문에서는 사전 양해여부에 따라 그룹 1과 그룹 2가 동일한 내용을 시청하였고 그룹 3과 그룹 4가 동일한 내용을 시청하였으므로 그룹 1, 2의 평가값과 그룹 3, 4의 평가값을 비교해 보았다. 그런데 문헌을 통해 예측한 바와는 달리 사전

양해의 여부는 에이전트 제품에 대한 선호도에 유의한 영향을 끼치지 않는 것으로 나타났다 [ $F(1, 102) = 1.97, p = 0.16$ ]. 또한 에이전트의 신뢰도에 대한 평가에서도 오히려 사전 양해가 없는 경우에 더 신뢰하는 것으로 나타났으며 그 차이는 통계적으로 유의하게 나타나기도 하였다 [ $F(1, 102) = 6.10, p = 0.02$ ]. 즉 사전 양해의 여부는 에이전트에 대한 호감도나 신뢰도에 긍정적인 영향을 끼치지 않으며 오히려 에이전트가 사전 양해를 덧붙이는 경우, 에이전트의 첫인상에 대한 신뢰도를 낮출 수 있음을 시사한다.

### 5.2. 설문 응답에 대한 통계 분석 사과가 호감도와 신뢰도에 끼치는 영향

데모 영상 2에서는 에이전트가 사용자로부터 요청 받은 과업을 제대로 수행하지 못하며, 이에 대해 사과를 하는 경우와 하지 않는 AO 및 AX 종류로 나뉜다. 따라서 최종 설문 분석에서는 Fig. 2에 도식화된 네 그룹 각각에 대하여 호감도와 신뢰도 평가 수치를 분석해 보았다.

그러나 Table 1에 제시된 평가 점수의 평균을 살펴

Table 1. Average assessments about how the users like and trust the agent when the agent expresses forewarning and/or an apology. Standard deviations are in parentheses. FO: when a forewarning is mentioned; FX: when no forewarning is mentioned; AO: when an apology is made; AX: when no apology is made.

Survey	Subjective judgement	Average values of mid- and final survey assessed by four groups			
		Group1 (FO-A0) (N = 25)	Group2 (FO-AX) (N = 32)	Group3 (FX-AO) (N = 22)	Group4 (FX-AX) (N = 25)
intermediate survey	Like it (1~5)	3.78 (0.84)		3.59 (0.64)	
		F(1, 102) = 1.97, p = 0.16			
	Trust it (1~5)	3.19 (0.84)		3.55 (0.59)	
F(1, 102) = 6.10, p = 0.02					
final survey	Like it (1~5)	3.15 (0.93)	3.13 (1.04)	3.20 (0.72)	3.19 (0.72)
		F(3, 100) = 0.04, p = 0.99			
	Trust it (1~5)	2.96 (0.78)	2.82 (1.13)	2.79 (0.75)	2.91 (0.66)
F(3, 100) = 0.20, p = 0.90					

보면 호감도와 신뢰도 모두 네 그룹 간 평균 크기에 큰 변화가 관찰되지 않았다. 일원분산분석을 통한 통계적 분석에서도 호감도와 신뢰도 모두 그룹 간 차이가 유의하지 않았다 [호감도:  $F(3, 100) = 0.04, p = 0.99$ ; 신뢰도:  $F(3, 100) = 0.20, p = 0.90$ ]. Sheffe 방법을 이용한 사후분석에서도 통계적 차이를 보이는 그룹들은 파악되지 않아, 모든 네 그룹에 대한 평가 점수는 차이가 없는 것으로 확인되었다. 따라서 에이전트의 사회적인 역할에 대한 기존 연구나 로봇을 대상으로 살펴본 최근 연구와는 달리 에이전트의 경우 사전 양해와 사과의 여부가 에이전트 제품에 대한 호감이나 신뢰를 형성하는 데 유의미한 기여를 하지 못하는 것으로 파악된 것이다.

한편 실험 연구의 과정에서 중간 설문과 최종 설문 평가가 동일한 항목으로 진행된 만큼, 평가 시점에 따른 비교해볼 수 있다. Table 1에 나타난 바와 같이 최종 설문에서 호감도와 신뢰도 두 가지 항목의 평균 점수가 모두 감소하는 경향이 관찰된다. 모든 실험참가자가 중간 및 최종 설문에 참여하였으므로 반복측정 t 테스트를 실시한 결과, 호감도 및 신뢰도 모두 유의한 수준에서( $p < 0.01$ ) 감소한 경향이 통계적으로 유의한 것으로 밝혀졌다 [호감도:  $t(103) = 4.29, p < 0.01$ ; 신뢰도:  $t(103) = 5.58, p < 0.01$ ]. 이는 데모 영상 2에서 에이전트가 과업을 제대로 실행하지 못한 내용을 접하게 된 데 원인이 있다고 사료된다.

### 5.3. 실험 결과에 대한 논의

사전 연구 (Lee et al. 2010)에서는 서비스 로봇이 사용자가 요구한 음료캔이 아닌 다른 캔을 가져오는 실수 상황에 대하여 다양한 사회적 전략을 적용한 데모 영상을 보여주고 실험참가자가 로봇과 로봇의 서비스를 평가하도록 하였다. 2 (사전 양해 여부) x 4 (실수 상황 이후 사과, 보상, 대안 제공, 언급없음) x 2 (로봇 생김새 차이) 등 총 16 그룹으로 나누어 총 317명의 실험참가자 (그룹별 14명 이상)에게 데모 영상을 보여주고 서비스와 로봇에 대해 평가하도록 한 결과, 사전 양해는 실험참가자들의 로봇에 대한 평가에 있어 긍정적인 영향을 보였다. 또한, 실수 상황에 대한 사과, 보상, 대안 제공 등의 사회적 전략은 아무런 언급이 없는 경우 보다 모두 긍정적인 영향을 주었다. 즉, 서비스 로봇의 실수 상황에 대한 사전 양해, 사과, 보상, 대안 제공 등의 사회적 전략은 사용자의 부정적인 감정을 완화하는데 긍정적인 효과를 보였다.

이러한 사전 연구 결과를 근거로 본 연구에서는 대화형 에이전트의 경우에도 사전 양해가 사용자의 에이전트에 대한 평가에 긍정적인 영향을 주고 오류 상황에서의 정중한 사과가 사용자의 부정적 감정 완화에 도움이 될 것이라 예상하였다. 하지만 실험 결과, 예상과 상반된 결과를 얻었다. 에이전트가 사전 양해를 하는 경우 오히려 신뢰도가 크게 감소하는 것을 발견하



였고 오류 상황에서 에이전트의 사과는 사용자의 부정적 감정 완화에 큰 영향을 보이지 않았다. 이러한 차이는 사용자들의 대화형 에이전트에 대한 인식과 로봇에 대한 인식의 차이에서 비롯된 것으로 보인다. 인공지능 스피커와 로봇은 형태에 있어 실재감과 사람과 가깝다고 느껴지는 정도가 달라, 사람 간의 사회적 전략의 효과가 다르다고 추측할 수 있다. Mennicken et al. (2016)의 연구에서도 대화형 에이전트의 “좋은 하루 보내세요.”와 같이 아주 가벼운 사회적 소통의 표현에 대해서 대부분의 사용자가 불편함을 표현하였다. 해당 연구에서 한 사용자는 단순히 기계일 뿐인 에이전트가 이러한 말을 하는 것은 가짜처럼 느꼈다고 설명하였다. 이 밖에도 본 연구의 가설이 기각된 데 대해서는 다양한 원인이 존재할 것이라 예상되며, 인공지능을 이용한 제품의 호감도와 신뢰도를 높이기 위한 방법은 로봇 등 기존의 다른 제품에서 확인된 바를 그대로 접목하는 데는 새로운 탐색이 필요함을 시사한다. 이에 설문지 평가의 정량적 분석에 그 원인을 파악하고자 실험에 참여한 평가자들을 무작위로 선택, 심층 인터뷰를 진행하였다.

## 6. 심층 인터뷰

실험에서 발견된 현상의 원인을 찾아보고자 설문에 참여했던 참가자 중 5명 (P1-P5)을 대상으로 심층 인터뷰를 추가로 진행하였다. 인터뷰는 특별한 형식 없이 설문 중에 시청했던 동영상을 다시 보며 자유롭게 의견을 이야기하는 형식으로 진행하였다.

### 6.1. 과업 수행 전 사전 양해를 미리 구하는 에이전트

대화형 에이전트의 사전 양해에 대하여, 참가자 P1은 “말이 많아서 지루했고, 자신감이 없어 보인다고 생각했다. 자신감 없는 사람이 ‘저 실수 할 수도 있으니 좋게 봐 달라’고 하는 것 같아서 좋게 보이지 않았다. [...] 내가 무엇을 물어볼 지도 모르는데 미리 양해를 구하는 것이 좀 이상하다고 생각했다.”고 설명했다. 하지만 사전 양해에 대해 긍정적으로 생각하는 참가자도 있었다. P2와 “처음 사용할 때만 나오는 문

구이고 매번 나오는 문구가 아니므로 (말이 길더라도) 상관없다고 생각한다. 어떻게 사용하는지 알 수 있으니 좋다고 생각했다.”고 했으며 P5도 “같은 말이더라도 하는 말투에 따라 스피커가 알아들을 수도 있고 없을 수도 있는데, 사전 양해를 통해 (정확한 기계의 사용법을) 알려주는 것은 좋은 것 같다. 사전 양해를 통해 사람들이 말을 할 때 스피커가 잘 알아들을 만하게 문장을 잘 생각해서 에이전트에게 말할 수 있을 것 같다.”고 설명했다.

### 6.2. 과업 실패에 대해 사과를 하는 에이전트

오류 상황에서 에이전트의 사과에 대한 인터뷰 참가자의 의견은 조금씩 다르게 나타났다. P3은 “사과를 하면 자신이 무엇을 잘못된 건지 안다는 의미이니 이후의 사용에서 데이터가 쌓여서 더 좋아지지 않을까 하는 기대감이 있다.”고 하였다. P3은 “사과를 하는 것은 괜찮지만, 실수할 때마다 사과를 하게 되면 짜증이 날 것 같다”고 했고, 특히 P5의 경우, “(오류 상황을 야기한) 동일한 질문을 반복하더라도 못 알아듣는 것은 마찬가지로 텐데, ‘다음에는 좀 더 귀 기울여 들어 볼게요.’라는 사과는 의미가 없을 것 같다”고 하였다. P4의 경우, “굳이 기계인데 사과를 해야 하나 생각이 든다. [...] 오류 상황이 생겼다면, 질문을 제대로 하지 못한 사람의 실수라고 생각한다. 기계의 작동법을 잘 알지 못한 사람의 실수인데, 기계가 그 상황에 대해 미안해하는 것이 의미가 있을까 싶다. 사과를 한다면, 그냥 그런 기능이 있구나 정도로 생각이 들 것 같다”고 하였다. 또한 이러한 반응은 실험에 사용한 스피커의 형태의 영향을 받는다는 의견도 있었다. P5는 “이 기계를 단지 기계로만 보는지 아님 더 사람 같은 어떤 것으로 보는지에 따라서 다를 것 같다. 이 스피커는 로봇처럼 생긴 게 아니라 그냥 기둥처럼 생겼고 얼굴이 있는 것도 아니고... 만약 귀여운 토끼 같은 인형 형태라면 또 다를 수 있을 것 같다”고 설명했다.

### 6.3. 심층 인터뷰에 대한 논의

심층 인터뷰 결과는 실험참가자들이 대화형 에이전트를 어떻게 인식하였는지 잘 보여준다. 인터뷰 참가

자들은 개인 간 정도의 차이는 있었으나 대체로 대화형 에이전트를 사람과 가까운 인격체보다는 단순한 기계로 인식하는 경향이 매우 강하게 나타났다. 이에 따라 사람 간에 작용하는 사회적 전략은 하나의 부가적인 기능 정도로만 인식되었고 크게 영향력을 발휘하지 못한 것으로 보인다. 사전 양해의 경우, 에이전트의 사전 양해는 참가자들에게 부정적인 첫인상을 주는 것으로 드러났다. 하지만 기계 사용법을 설명해 준다는 측면에서는 긍정적으로 받아들여지기도 하였다. 다만, 데모 영상에서의 “하지만 저는 가끔 사용자님의 말을 이해하지 못하거나, 사용자님이 부탁한 일을 정확히 수행하기 어려울 수도 있어요.”와 같이 애매한 표현보다는 정확하게 사용 지침을 주는 것이 더 유용할 것이라는 의견이 있었다. 오류 상황에서의 에이전트의 사과에 대한 인터뷰 응답은 대체로 부정적으로 나타났다. 인터뷰 참가자들은 진심에서 우러나온 것이 아닌 단순 설계된 사과 반응에 대해서 불필요하다는 의견을 보였다. 다만, 제품 형태에 대해 언급한 P5의 응답에서 볼 수 있듯이 제품 형태에 따라 사용자의 에이전트에 대한 인식은 달라질 수 있고 이는 에이전트의 사회적 전략에 대한 반응에도 영향을 미칠 수 있을 것으로 보인다.

## 7. 종합 논의 및 결론

본 연구는 대화형 에이전트의 오류 상황에 대한 사전 양해가 사용자의 에이전트에 대한 인식에 미치는 영향을 이해하고, 오류 상황에서의 정중한 사과가 사용자의 부정적 감정 완화에 미치는 영향을 검증하는 것을 목적으로 진행하였다. 총 104명을 대상으로 한 설문조사 결과, 에이전트의 사전 양해 여부는 에이전트의 호감도에는 영향을 미치지 않지만 신뢰도에는 크게 영향을 미쳤고, 에이전트의 오류 상황 이후 정중한 사과 여부는 에이전트의 호감도와 신뢰도에 영향을 미치지 않음을 발견할 수 있었다. 이는 기존의 로봇을 이용한 유사한 실험을 진행하였을 때와 상반되는 결과이다.

설문 및 심층 인터뷰 결과를 바탕으로 볼 때, 사용자가 대화형 에이전트와 로봇에 갖는 인식의 차이가

있음을 알 수 있었다. 사람들은 로봇 혹은 얼굴이 있는 가상 에이전트는 보다 사람 같은 존재로 인식을 하는 반면 대화형 에이전트의 경우 단순한 기계에 가깝다고 느끼는 경향이 크게 나타났다. 이는 본 실험의 데모 동영상에 사용된 대화형 에이전트 및 사람들의 기존 인식 속의 대화형 에이전트가 단순한 외형을 갖고 있기 때문인 것으로 판단된다. 사람들이 대화형 에이전트를 기계에 가깝게 인식하였기 때문에 사전 양해나 사과와 같은 인간관계에 흔히 사용되는 사회적 전략들이 큰 영향력을 발휘하지 못하였다. 사전 양해의 경우 오히려 양해를 구함으로써 기계로서의 수행 능력에 대한 한계를 드러내어 부정적인 영향을 미치게 되었고, 사과의 경우 기계가 하는 사과이기 때문에 진심이 전해지기보다는 형식적이라는 느낌을 주어 부정적 감정을 완화시키지 못한 것으로 보인다.

이러한 결과는 사용자가 에이전트를 얼마나 기계, 혹은 사람에 가깝게 인식하는지에 따라 오류 상황에 대한 에이전트의 대처 방식 또한 달라져야 함을 의미한다. 이는 에이전트와의 상호작용을 디자인할 때에 사용자로 하여금 에이전트를 어떠한 존재로 인식하게 할지에 대한 고려가 충분히 되어야 하며, 사용자 개인의 인식의 차이 또한 고려해야 함을 의미한다.

본 실험은 실험참가자가 대화형 에이전트와 직접 상호작용하는 것이 아니라, 데모 영상을 시청한 후의 감정을 바탕으로 호감도와 신뢰도를 평가하였기 때문에 경험의 한계가 있을 수 있다. 또한 실험에서 사용된 데모 영상의 상황이 최근 방영되고 있는 인공지능 스피커의 광고들을 떠올리게 하여 실험참가자들의 평가에 영향을 끼쳤을 가능성을 배제할 수 없다. 또한 본 연구의 실험은 모든 그룹이 대화형 에이전트의 사전 양해 여부에 따른 데모 영상 1을 먼저 접하도록 동일하게 설계되어, 뒤에 접한 데모 영상 2의 에이전트의 실수 상황에 대한 평가에 영향을 미쳤을 가능성도 감안해야 할 것이다. 이와 같은 본 연구가 가지는 한계점에도 불구하고 본 연구는 최근 활발히 연구되고 있는 대화형 에이전트와의 상호작용 분야에서 좋은 시발점이 될 수 있으리라 기대한다. 특히, 에이전트의 오류 상황에 대한 사용자의 인식 및 경험 변화에 대한 이해 및 에이전트의 대응방향에 대한 연구의 시발점이 될 수 있을 것이라 생각한다.

차후 연구에서는 실험참가자가 대화형 에이전트를 직접적으로 사용하도록 하고, 보다 장기적으로 에이전트와의 상호작용 경험을 추적하여 호감도와 신뢰도에 관한 전반적인 변화를 관찰해볼 수 있을 것이다. 또한 실험참가자가 대화형 에이전트에 대해 느끼는 호감도와 신뢰도를 보다 정확히 측정할 수 있는 설문 방식과 구체적인 설문 문항에 대해서도 추가적인 고려가 필요할 것으로 보인다. 이를 통해 대화형 에이전트와의 단편적인 상호작용을 넘어 실제 사용 경험 전반을 관찰할 수 있다면 대화형 에이전트의 오류 상황에 대한 사용자 경험을 보다 깊게 이해할 수 있을 것이다.

## REFERENCES

- Appel, J., von der Pütten, A., Krämer, N. C., & Gratch, J. (2012). Does humanity matter? Analyzing the importance of social cues and perceived agency of a computer system for the emergence of social reactions during human-computer interaction. *Advances in Human-Computer Interaction*, 13(2012), 1-10. DOI: 10.1155/2012/324694
- Bickmore, T., & Cassell, J. (2001). Relational agents: a model and implementation of building user trust. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 396-403. DOI: 10.1145/365024.365304
- Cassell, J. (2000). Embodied conversational interface agents. *Communications of the ACM*, 43(4), 70-78. DOI: 10.1145/332051.332075
- Jarnow, J. (2016, April 8). *Why our crazy-smart AI still sucks at transcribing speech*. Retrieved from <https://www.wired.com/2016/04/long-form-voice-transcription/>
- Lee, M. K., Kielser, S., Forlizzi, J., Srinivasa, S., & Rybski, P. (2010). Gracefully mitigating breakdowns in robotic services. *Proceedings of the 5th ACM/IEEE International Conference on Human-robot Interaction*, 203-210. DOI: 10.1145/1734454.1734544
- Li, A. (2017). *Google's speech recognition is now almost as accurate as humans*. Retrieved from <https://9to5google.com/2017/06/01/google-speech-recognition-humans/>
- Luger, E., & Sellen, A. (2016). "Like Having a Really bad PA": The Gulf between User Expectation and Experience of Conversational Agents. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 5286-5297. DOI: 10.1145/2858036.2858288
- Mennicken, S., Zihler, O., Juldaschewa, F., Molnar, V., Aggeler, D., & Huang, E. M. (2016). It's like living with a friendly stranger. *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 120-131. DOI: 10.1145/2971648.2971757
- Nass, C., & Moon, Y. (2000). Machines and mindlessness: social responses to computers. *Journal of Social Issues*, 56(1), 81-103. DOI: 10.1111/0022-4537.00153
- Park, J.Y. (2007) Effects of the interaction with computer agents on users' psychological experiences. *Science of Emotion & Sensibility*, 10(2), 155-168.
- Price, R. (2017). *Microsoft's AI is getting crazily good at speech recognition*. Retrieved from <http://uk.businessinsider.com/microsofts-speech-recognition-5-1-error-rate-human-level-accuracy-2017-8>
- Purinton, A., Taft, J. G., Sannon, S., Bazarova, N. N., & Taylor, S. H. (2017). "Alexa is my new BFF": Social roles, user satisfaction, and personification of the Amazon echo. *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, 2853-2859. DOI: 10.1145/3027063.3053246
- Song, H.S., Kim, M.J., Jeong, S.H., Suk, H.J., Kwon, D.S., & Kim, M.S. (2008). The behavioral patterns of neutral affective state for service robot using video ethnography. *Science of Emotion & Sensibility*, 11(4), 629-636.
- von der Pütten, A. M., Krämer, N. C., Gratch, J., & Kang, S.-H. (2010). "It doesn't matter what you are!" Explaining social effects of agents and avatars. *Computers in Human Behavior*, 26(6), 1641-1650. DOI: 10.1016/j.chb.2010.06.012
- Woods, S., Walters, M., Koay, K. L., & Dautenhahn,

K. (2006). Comparing human robot interaction scenarios using live and video based methods: towards a novel methodological approach. In *Proceedings of the 9th International Workshop on Advanced Motion Control*. 750-755.  
DOI: 10.1109/AMC.2006.1631754

원고접수: 2018.01.02

수정접수: 2018.02.21

게재확정: 2018.02.23