

In The Context Of Conversational Errors With Voice Agents, A Direction For Enhancing Re-response Messages

음성 에이전트와의 대화 오류 상황에서 재응답 메시지 개선 방향

So Yeon Kim¹, Jong Hoon Choe²

김소연¹, 최종훈²

¹ Student, Media Interaction Design Department, Ewha Womans University, Korea,
bgd355@ewhain.net

² Professor, Media Interaction Design Department, Ewha Womans University, Korea,
deadant@ewha.ac.kr

Corresponding author: Jong Hoon Choe

Abstract: With the development of interactive artificial intelligence technology, user experience and satisfaction in voice recognition technology and voice conversation interaction are also becoming important. However, while users expect human-like conversation patterns from voice agents, conversation errors frequently occur, and to offset this experience, it is necessary to focus on conversation design and structure in terms of emotional conversation in addition to technical problems. In this study, we would like to derive feedback and inconveniences that users experience when talking with voice agents and present directions for improving the voice agent's response message accordingly. Through user surveys, five representative types of errors were found for each case of voice agent use, and in-depth interviews were used to derive improvement directions by specifying non-technical inconveniences and user needs in the situation of voice agent conversation errors. Based on this improvement direction, it was possible to present a re-response message strategy framework and re-response message guidelines. This study is meaningful in that it analyzes the structure of conversations with voice agents and identifies users' non-technical inconveniences and clear feedback points, and it is expected that empirical conversation design guides can be presented in the future through the improvement directions derived from this study.

Keywords: Voice Agent, AI Speaker, VUI, Re-response

요약: 대화형 인공지능 기술의 발전에 따라 음성 인식 기술 및 음성 대화 인터랙션에서의 사용자 경험 및 만족도도 중요해지고 있다. 하지만 사용자들이 음성 에이전트에게 인간과 같은 대화 양상을 기대하는 것에 반해 대화 오류는 빈번히 발생하며, 이러한 경험을 상쇄하기 위해서는 단순히 기술적인 문제 외에 감성 대화 측면에서 대화 디자인과 구조에도 집중할 필요가 있다. 본 연구에서는 사용자들이 음성 에이전트와의 대화 시 겪게 되는 피드백과 불편 사항을 도출하고 그에 따른 음성 에이전트의 재응답 메시지 개선 방향을 제시하고자 한다. 사용자 조사를 통해 음성 에이전트 사용 케이스 별 다섯 가지의 대표 오류 유형을

Received: May 23, 2023; 1st Review Result: June 27, 2023; 2nd Review Result: July 30, 2023
Accepted: August 25, 2023

발견하였으며, 심층 인터뷰를 통해 음성 에이전트의 대화 오류 상황에서의 기술 외적인 불편사항과 사용자의 니즈를 구체화하여 개선 방향을 도출하였다. 이 개선 방향성을 토대로 재응답 메시지 전략 프레임워크와 재응답 메시지 가이드라인을 제시할 수 있었다. 본 연구는 음성 에이전트와의 대화 구조를 분석하고 사용자의 기술 외적인 불편함과 명확한 피드백 발생 지점을 파악하였다는 점에서 의의가 있으며, 본 연구에서 도출한 개선 방향을 통해 향후 실증적인 대화 디자인 가이드를 제시할 수 있을 것으로 기대한다.

핵심어: 음성 에이전트, AI 스피커, VUI, 재응답

1. 서론

최근 들어 스마트 스피커나 AI 음성 에이전트, 챗봇 등 다양한 형태의 대화형 에이전트가 등장하고 있다. 특히 인공지능 스피커가 확산됨에 따라 스마트 스피커는 스피커의 기본 기능인 단순 청취뿐만이 아니라 음성인식 기술을 결합하여 쌍방향 대화가 가능하도록 진화하였다. 음성을 통한 인터랙션은 인간에게 친숙한 음성을 통한 소통 방식을 사용하는 정보 입출력 방법으로, 비교적 조작 방식을 학습할 필요성이 낮고 기존의 터치나 타이핑 방식에 비해 동시에 더 많은 일을 수행할 수 있다. 더불어 손이 자유롭지 못한 상황에서도 목소리를 통해 조작이 가능하며, 음성 입력의 속도는 기존 타이핑 방식 속도의 약 3배에 달하는 만큼 다른 인터랙션 조작 방식에 비해 빠르다는 장점이 있다[1]. 이러한 대화형 에이전트는 사용자와 기계 사이에 단순한 정보 제공이나 명령 수행을 넘어서 사람 간 대화처럼 자연스러운 상호작용을 수행하고자 하는 목적을 지닐 수 있게 되었다. 음성 에이전트는 음성 인터랙션을 통해 음악, 일정, 날씨, 상품 등 다양한 정보와 콘텐츠들을 검색, 이용할 수 있는 대화형 서비스를 제공하고 있으며[2], 앞으로 더욱 많은 상황에서 인공지능 음성 에이전트를 활용한 상호작용 시도는 꾸준히 증가할 것으로 기대된다. 하지만 한국 소비자원의 ‘음성인식 스피커 이용 실태 현황 및 소비자 만족도 조사’에 따르면, 일상 대화는 2.78점으로 가장 낮은 만족도를 보였다[3]. 사용자는 음성 에이전트에게 사람과 같은 자연스러운 일상 대화를 기대했지만 이를 적절히 수행하지 못했다는 것을 보여주며, 이러한 대화 오류 상황에서 기술적 한계 외에 사용자의 불편을 줄일 수 있도록 음성 에이전트의 적절한 대처가 필요하다. 실제 사람 간의 대화에서도 의사소통의 오류나 발화 수정을 어렵지 않게 관찰할 수 있고, 특히 인간의 발화에서는 빈번한 휴지나 발화 중단 또는 주저함과 같은 비유 창성 현상이 존재하기 마련이다[4]. 이때 실제로 발화 상의 문제를 발견하면 발화를 중단하고 그 문제를 해결해 나가기도 한다. 이렇듯 음성 에이전트와의 대화에서도 마찬가지로 대화 오류 발생은 불가피하다. 하지만 이러한 상황에서 사용자의 디바이스 사용 중단이나 사용 경험 저하를 최소화할 수 있도록 사용자 몰입 경험이 떨어지는 지점을 파악한 다음 이를 상쇄시키기 위한 적절한 경험을 제공해야 할 것이다[5].

본 연구에서는 음성 에이전트와의 대화 오류 상황을 정의하고, 대화 오류 상황에서의 획일화된 응답과 발화 중단 또는 재명령 상황의 대처를 유형별로 분류해 음성 에이전트의 응답 방식에 대한 개선 방향을 제시하고자 한다. 인공지능 음성 에이전트의 사용 지속성 향상을 위해서는 오류 상황에서 기술적 한계 외의 불편을 줄이는 것이 중요하기에, 이러한 개선 방향에 따라 사용자와의 감성적 인터랙션을 향상할 수 있는 음성 디자인을 제시할 수 있음에 의의를 둔다.

2. 이론적 배경

2.1 CASA 패러다임

CASA(Computers are Social Actors, CASA) 패러다임에 따르면, 컴퓨터는 사회적 행위자로 인식되며 이는 의인화된 신호를 나타내는 컴퓨터와 상호작용할 때 실제 인간과 상호작용하는 것처럼 개인이 무의식적으로 사회적 규칙을 적용한다는 것을 의미한다[6]. 특히, 컴퓨터 또는 디바이스라 할지라도 의인화된 특성을 지니고 있다면 사람들은 무의식적으로 사회적으로 응답하고, 그들이 실제 인간인 것처럼 인식한다[7]. 따라서 인공지능 음성 에이전트와의 대화에서 사용자는 자연스러운 인터랙션 경험을 기대하게 되고, 음성 에이전트의 발화 또한 사람 간의 대화 특성을 따르게 된다.

2.2 AI 음성 에이전트

본 연구에서의 음성 에이전트란, 모달리티를 사용하여 인간과 유사한 방식으로 사용자와 상호작용하도록 디자인된 시스템으로 정의한다[8]. 음성 에이전트는 음성 대화를 통해 스마트 기기를 제어할 수 있는 인터페이스 기술이다. 현재의 음성 에이전트는 사용자가 시동어(wake-up word)를 통해 사용을 시작하면 이를 녹음하여 서버로 보내 분석하고 녹음된 음성을 서버에서 받아와 재생하는 과정을 거친다. 이때, 동음이의어나 구문을 구별하는 문제, 음성 외에 여러 가지 요소를 종합적으로 고려하고 화자가 누구인지, 어떤 상황에서의 대화인지, 앞선 대화의 맥락이 무엇인지 등 분석을 어렵게 하는 요인이 많다. 즉, 음성의 극단적인 유연함은 규칙으로 정리하기 어렵다고 볼 수 있다. 최근 빅데이터와 사용자들의 적극적인 사용을 통한 데이터 누적으로 음성을 잘못 인식하더라도 학습 결과를 토대로 단어를 보정하는 언어 모델을 통해 음성 엔진의 정확도는 올라가고 있다. 하지만 말이나 글의 의미를 알 수 있도록 언어를 구조화하는 자연어 이해 단계에서 사용자의 의도를 파악하고 누락된 정보를 채우는 멀티 턴(multi-turn) 과정에서 여러 번 대화를 반복하다 보면 사용자에게 불편함을 줄 수 있다. 음성 에이전트는 본질적으로 ‘문제 해결용 대화 시스템’이며, 목적이 분명한 대화가 주를 이룬다. 그렇기에 현재까지의 음성 에이전트의 답변은 몇 가지의 템플릿을 두고 번갈아가며 대답하여 대화를 구성해왔다. 이러한 대화 템플릿 작업을 두고 대화를 디자인한다고 할 수 있다. 하지만 여전히 음색 변화이나 감정 표현 등이 어렵다는 단점이 존재한다. 대화에 있어 흐름, 음조, 표현, 감정, 소리 크기, 속도 등 모든 것이 대화의 의미를 상대방에게 전달하는 요소이기 때문이다. 따라서 본 연구에서는 이러한 대화의 감성적 측면 또한 고려한 음성 에이전트의 답변 제공을 통해 사용자에게 만족감과 더 나은 사용 경험을 줄 수 있을지 알아보고자 한다.

2.3 대화형 인터랙션

대화형 인터랙션은 스마트 스피커, 음성 비서, 챗봇과 같은 인공지능과 사람의 대화 상호작용을 의미한다. 이는 앞서 말한 인간과 기기 상호작용에서 인간과 기계의 관계는 사회적이라는 CASA 패러다임에 근간을 두고 있다. 그중에서도 음성 대화형 인터랙션은 음성 인식 기술을 통해 인간의 음성 명령의 의미와 의도를 해석하는 대화형 솔루션인데, 자연어 근간의 소통 방식으로 인간에게 빠르고 간편한 정보 입출력 방식뿐만 아니라

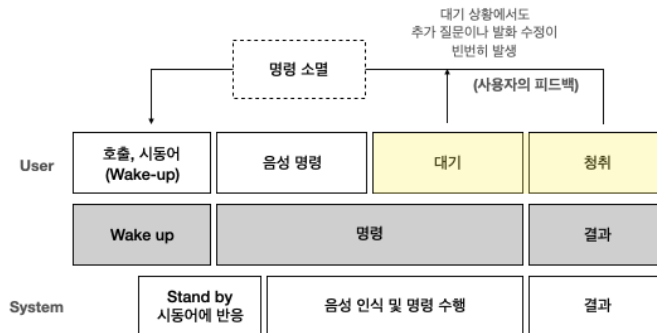
사람과의 자연스러운 감정 교류를 제시한다[9]. 인공지능 음성 에이전트는 커뮤니케이션에 최적화된 사회적 로봇(social robot)의 일종으로서, 인간과 의미 있는 상호작용을 통해[10] 인간과 기계와의 대화가 점차 인간과 인간 사이의 커뮤니케이션 방식을 닮아가고 있다. 인공지능 음성 에이전트가 사람과 유사한 속성을 지닐 수록 더욱 효과적인 상호작용이 가능할 것으로 예상되는 바이다[11].

[표 1] AI 음성비서 대화형 인터랙션 구조에 기반 한 질문[12]

[Table 1] Questions based on the AI voice Assistant Interactive Interaction Structure

AI 음성 어시스턴트 대화형 인터랙션 구조에 기반한 질문 (질문 전 호출에 대한 인지여부)	1. 정말 마음에 들지 않는다 (12명)
	2. 마음에 들지 않는다 (23명)
	3. 상관이 없다 (10명)
	4. 마음에 든다 (5명)
	5. 정말 마음에 든다 (0명)
	평균 2.16점 (50명)

[표 1]은 김민희, 남궁기찬의 선행연구를 통해 알 수 있었던 음성 에이전트와의 대화 인터랙션 구조에 기반한 질문이다. 사용자들은 AI 음성 비서와 대화를 진행하기 전 시동어(wake up word)를 통해 호출하는 것에 대해 평균 2.16점으로 낮은 만족도를 보였다. 이러한 선행 연구를 통해 파악한 음성 에이전트와의 대화 인터랙션 구조는 [그림 1]과 같다.



[그림 1] AI 음성 에이전트와의 음성 대화형 인터랙션 구조

[Fig. 1] Voice Interactive Interaction Structure with AI Voice Agent

사용자는 시동어를 통해 에이전트를 호출한 후, 음성으로 명령을 내리고 대기하는 과정을 거친다. 이때 시스템의 명령 수행 결과를 듣기 전 또는 후에 사용자의 피드백이 자유롭게 발생할 수 있는데, 현재까지의 인공지능 음성 인터랙션은 이러한 피드백은 수용하지 못하는 양상을 보였다. 하지만 이런 예측하지 못한 사용자의 의도에 유연하게 대처하기 위해 대용량의 대화 말뭉치를 반복적으로 학습하는 방안이 제시되고 있다[13].

3. 사용자 조사

3.1 사용자 조사 개요 및 방법

앞서 연구한 내용을 바탕으로 음성 에이전트와의 대화에 있어 오류 상황이 발생했을

때 재응답 메시지 개선 방향을 도출하기 위한 기반이 될 수 있는 사용자 조사를 실시했다. 본 조사 이전, 파일럿 인터뷰에서 음성 에이전트 사용 경험이 있는 참가자 12명(남성 6명, 여성 6명)을 대상으로 음성 에이전트에 대한 사용자의 전반적인 니즈와 사용 상황을 알아보았다. 인터뷰 질문은 음성 에이전트에 대한 인식과 사용 현황, 인상적인 사용 경험, 아쉬운 점과 만족스러운 점을 중심으로 구성되었다. 사례 연구와 파일럿 인터뷰를 통해 정리한 음성 에이전트의 대표적인 사용 상황 및 태스크는 [표 2]와 같다.

[표 2] 관찰에 사용된 전반적인 사용 상황 및 필요한 발화 예시

[Table 2] Typical Use Situation Used in the Observation and Examples of Required Utterances

NO	Case	Task
1	날씨, 뉴스 등 간단한 정보 검색	오늘 날씨랑 미세먼지 알려줘.
2	알람 설정 및 확인	오전 11시 버스 타야하니 미리 알려줘.
3	음악	뉴진스의 'Ditto' 틀어줘.
4	전화	수연이에게 전화 걸어줘.
5	IoT 홈 오토메이션	거실 불 키고 잔잔한 음악 틀어줘.

이후 오류 상황과 사용자의 피드백이 발생하는 순간 및 사용자의 반응을 확인하기 위해 2~30대 참가자 7명(남성 3명, 여성 4명)을 대상으로 사용자 관찰과 심층 인터뷰를 함께 진행하였다. 심층 인터뷰의 참가자는 상용화된 음성 에이전트를 최소 3개월 이상 사용한 경험이 있고 음성 에이전트와의 대화 컨텍스트에 익숙한 사용자로 선정하였다. 아래의 [표 3]는 사용자 조사 세부사항이며, 각 참가자들에 대한 인구통계학적 특성은 [표 4]와 같다.

[표 3] 사용자 조사 세부 사항

[Table 3] User Interview Details

구분	내용
조사대상	음성 에이전트를 최소 3개월 이상 사용한 경험이 있는 사용자
조사방법	사용자가 직접 음성 에이전트와 대화하며 태스크를 수행하는 모습을 관찰한 후 1:1 인터뷰를 통한 대면 심층 조사
피실험자 수	7명 (여성 4명, 남성 3명)
조사기간	2023년 6월 5일~2023년 6월 15일 (10일간)
인터뷰시간	약 50분

[표 4] 피실험자 정보

[Table 4] User Information

참가자	성별	연령	직업	음성 에이전트 종류	사용 기간
U1	여	25세	학생	시리	2년
U2	여	28세	프리랜서	KT 기가지니	6개월
U3	여	30세	사무직	시리, 구글홈	4년
U4	여	33세	자영업	빅스비, 카카오톡 미니	1년 6개월
U5	남	27세	디자이너	시리, 카카오톡 미니	3개월
U6	남	35세	사무직	갤럭시 빅스비, KT 기가지니	2년
U7	남	32세	전문직	KT 기가지니	5개월

실험에 사용한 음성 에이전트는 사용자가 익숙한 아이폰의 시리(SIRI), 갤럭시 빅스비, KT 기가 지니, 카카오 미니를 사용하였으며 사용 상황과 태스크를 제시하여 사용자들이 음성 에이전트와 대화하는 모습을 관찰하였다. 참가자가 제시된 상황과 태스크를 완수하고 난 후에는 각 참가자와 심층 인터뷰를 진행했으며, 전체 절차는 각 참가자 당 약 50분간 진행되었다. 관찰과 인터뷰는 참가자들에게 연구 기록에 대한 동의 후 진행되었음을 밝힌다.

음성 에이전트 사용 상황에서 나타나는 대화 오류 상황을 도출하기 위해 참여자들은 음성 에이전트에게 주요 상황에 대한 동일한 질문을 했으며, 이에 대한 음성 에이전트의 답변과 사용자의 피드백을 바탕으로 오류 상황을 도출하였다. 도출된 오류 상황은 [표 5]과 같이 1) 음성 인식 오류, 2) 의도 오류, 3) 정보 부족, 4) 대화 불일치, 5) 기술적 오류 총 5가지로 구분되었다.

[표 5] 음성 에이전트의 오류 유형 정의

[Table 5] Defining the Error Type for the Voice Agent

오류 유형	오류 내용	예시 (“그 날 날씨랑 미세먼지 알려줘”라는 명령에 대해)
1) 음성 인식 오류	에이전트가 사용자의 음성 인식에 실패함	문장을 인식하지 못함 (죄송하지만 잘 알아듣지 못했습니다.)
2)의도 오류	에이전트가 사용자의 질문이나 명령어를 제대로 파악하지 못하거나, 의도를 잘못 해석함	의도를 잘못 파악하여 미세먼지에 대한 정보만 제공 (오늘 미세먼지는 매우 나쁨입니다.)
3)정보 부족	사용자의 명령과 관련된 정보가 없거나, 요청한 작업을 수행할 수 없음	기상예보 서버에서 특정일의 정보를 제공하지 않는 경우 (죄송합니다. 잘 이해하지 못했습니다.)
4)대화 불일치	에이전트와 사용자 사이의 대화가 일관성 없이 진행됨 이전 대화 내용을 기억하지 못하거나, 사용자의 추가 질문에 적절히 대응하지 못함	“그 날”이 언제인지 맥락을 이해하지 못함 (현재 하늘이 부분적으로 흐린 상태이며, 기온은 19도입니다.)
5)기술적 오류	네트워크 연결 오류, 서버 오류, 음성 인식 엔진의 오작동 등 기술적 결함으로 발생하는 오류	기술적 결함 설명 (네트워크가 연결되지 않았습니다.)

5가지의 오류 상황 중 의도를 잘못 해석하거나 명령어를 제대로 파악하지 못한 2) 의도 오류 유형과 요청한 작업을 수행할 수 없는 3) 정보 부족 유형, 그리고 이전 대화의 맥락을 파악하지 못해 발행하는 4) 대화 불일치 유형에서 사용자의 즉각적인 명령 수정과 같은 피드백이 발생함을 관찰할 수 있었다. 이때 유의한 점은 1) 음성 인식에 실패해 다시 명령을 요청하는 경우와 3) 정보가 부족해 명령을 수행할 수 없는 경우의 응답 메시지가 동일하여 사용자는 두 오류 상황을 구별하지 못한 채 계속해서 명령 수행을 요구하게 된다는 것이다. 더불어 [표 6]과 같이 사용자의 피드백과 요구사항에 따라 해당 연구의 방향성에 맞는 주요 응답과 그에 따른 개선 방향을 정리할 수 있었다.

[표 6] 음성 에이전트와의 대화 오류 상황에서 사용자 응답 및 개선 방향성

[Table 6] User Response and Improvement Direction in Situations of Conversation Errors with Voice Agents

오류 상황	사용자 응답	개선 방향
의도 오류	-‘아니’, ‘그거 말고’ 등 부정어 사용 -시동어를 사용하여 재 호출이 필요하다는 것을 망각함 -성공할 때까지 다양하게 발화 시도	-부정어 인식 및 시동어 없이 대기 상태 돌입
정보 부족	-음성 인식 오류와 구분되지 않아 계속해서 명령 재시도 -반복적인 실패 경험으로 짜증 유발 및 재사용 의도 저하	-음성 인식 오류 메시지와 명확한 구분이 필요 (예: “죄송합니다. 제가 알 수 없는 부분이에요.”) -반복 실패 시 사과 표현 강화 및 추가 방법 제안 (예: “ 또 다시 불편을 드려 정말 죄송해요. 제가 도울 수 있는 다른 방법은 없을까요?”)
대화 불일치	-‘아니’, ‘그거 말고’ 등 부정어 사용 -동문서답을 할 경우 오히려 웃음 유발	-부정어 인식 및 시동어 없이 대기 상태 돌입 -부정어 인식 및 대화 재확인 (예: “제가 이해한 게 ~가 맞을까요?”)

이러한 오류 상황에 따라 적용될 수 있는 메시지 가이드라인 디자인을 위해 Dörnyei and Scott(1997)[14]에서 제시한 의사소통 전략을 참고하여 [표 7]과 같이 재응답 메시지 전략 프레임워크를 도출하였다.

[표 7] 재응답 메시지 전략 프레임워크

[Table 7] Re-response Framework

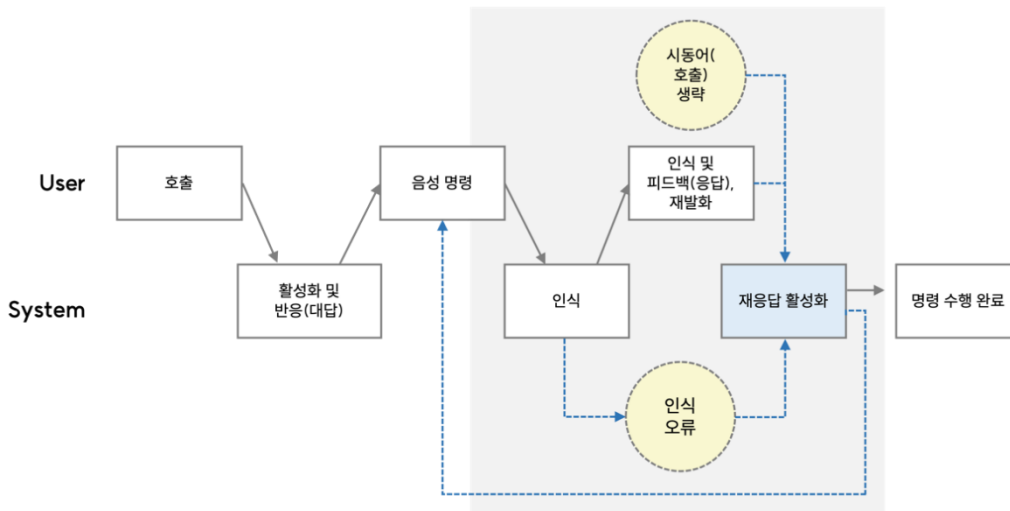
오류 상황	의사소통 전략	재응답 메시지 전략	전략 정의
의도 오류	메시지 포기하기	부정어 인식	즉각 발생하는 사용자의 피드백 중 부정어를 인식하여 올바른 명령 수행
	직접적 도움 요청하기	반복 요청	의도 및 명령을 이해하지 못했음을 표현하고 설명 및 재발화 요구
정보 부족	반응 : 풀어 말하기	대안 제공	특정 명령을 수행할 수 없는 이유를 명확히 설명하고 다른 대안을 제공
대화 불일치	메시지 포기하기	부정어 인식	즉각 발생하는 사용자의 피드백 중 부정어를 인식하여 올바른 명령 수행
	확인 요청하기	해석적 요약	사용자가 발화한 문장을 반복하여 재확인 유도
	간접적 도움 요청하기	추측하기	명령의 범위를 좁혀 재발화 유도

이를 바탕으로 음성 에이전트와의 대화 오류 상황에서 재응답 메시지 제공 시 공통적으로 고려해야 하는 내용을 [그림 2]와 같이 정리하였으며, 이에 따른 디자인 가이드를 구조화하여 [그림 3]과 같이 제시한다.

오류 유형	의도 인지 오류
사용자 응답	· 부정어 사용 · 재호출 생략
메시지 전략	가이드 라인
1. 부정어 인식 2. 반복 요청	1. 아니 / 그거 말고 / 잠깐 / 멈춰 등과 같은 부정어를 인식하고, 호출이 필요하지 않는 대기 상태에 돌입한다. 2. 부정어 사용이 반복적으로 일어날 때, 시스템은 사용자에게 의도를 이해하지 못했음을 명확하게 표현하고 다시 명확하게 발화할 것을 요청한다.
오류 유형	정보 부족 오류
사용자 응답	· 오류 원인 인지 실패 · 재발화 시도
메시지 전략	가이드 라인
1. 고쳐 말하기 2. 대안 제공	1. 사용자에게 명령을 수행할 수 없는 이유를 명확히 설명하며 재발화가 반복적으로 일어날 시 사과 표현을 강화한다. 2. 올바른 결과를 얻는데 도움이 될 수 있는 대안과 같이 대체할 수 있는 기능을 안내한다.
오류 유형	대화 불일치
사용자 응답	· 부정어 사용 · 재호출 생략
메시지 전략	가이드 라인
1. 부정어 인식 2. 반복 / 유도하기 3. 설명 요청하기	1. 아니 / 그거 말고 / 잠깐 / 멈춰 등과 같은 부정어를 인식하고, 호출이 필요하지 않는 대기 상태에 돌입한다. 2. 재발화가 반복적으로 일어날 때, 사용자의 발화 내용을 반복하여 재확인을 유도한다. 3. 사용자가 동일한 단어를 반복적으로 재발화할 경우, 시스템이 이해한 의미와 다른 것으로 간주하고 사용자에게 의미를 명확하게 할 것을 요청한다.

[그림 2] 오류 유형별 VUI 재응답 메시지 가이드라인

[Fig. 2] VUI Re-response Message Guideline by Error Type



[그림 3] 음성 에이전트의 대화 오류 상황 재응답 활성화 흐름도

[Fig. 3] VUI Re-response Flowchart In Conversational Errors

4. 결론 및 제언

본 연구는 음성 에이전트와의 대화 오류 상황에서 사용자의 불만 피드백이 발생하는

부분과 그에 따른 재응답 메시지를 개선하고자 사용자 조사를 실시하여 알아보았다. 사용자 조사 결과, 사용자들의 피드백이 발생하여 명령의 재시도나 수정을 요구하는 부분을 도출할 수 있었다. 앞서 분석한 연구 결과를 토대로 오류 상황에서의 사용자의 불편과 요구사항을 새롭게 도출할 수 있었으며, 선행연구를 기반으로 한 음성 에이전트의 재응답 메시지 개선 방향으로 세 가지 사항을 제시한다.

첫째, 사용자는 음성 에이전트와의 대화에 있어 오류 상황 발생 시 바로 문제를 해결하고자 했기에 음성 에이전트는 오류 메시지 응답 후 대기 상태로 돌아가는 것이 아닌 다음 질문에 대처할 수 있어야 한다.

둘째, 반복적인 실패 경험과 동일한 오류 메시지로 재사용 의도가 저하됨을 알 수 있었기에, 오류 메시지 유형을 세분화하고 사용자의 짜증을 유발하지 않는 선에서 사과 화행 전략을 쓰는 것이 적절할 것으로 보인다.

셋째, 시동어를 통해 재호출하는 과정에서 대화 불일치-맥락 오류가 빈번히 발생하기에 이를 보완하기 위해 음성 에이전트는 대화 기록을 보존하고 분석하여 다음 대화에서 활용할 수 있어야 한다.

이를 통해 음성 에이전트는 사용자의 이전 발화를 기억하고 이해하여 이후 대화에서 그에 맞는 응답을 제공할 수 있어야 하며, 이러한 대화 기록 보존을 통해 사용자의 행동 패턴, 선호도, 개인화 요구사항 등을 파악하고 사용자 경험을 개선하는데 활용할 수 있다.

하지만 현재 음성 에이전트의 대화 흐름과 대화 기록 및 분석은 텍스트 기반 채팅 AI와 비교했을 때 분명 어려움이 존재한다. 음성 입력의 복잡성과 다양성, 발화의 동시성 등의 한계가 있기 때문이다. 그러나 인공지능 음성 인식 기술의 발전과 음성 데이터의 폭발적인 증가에 따라 이러한 기능도 더욱 개선될 가능성이 있으며, 이후 음성 에이전트의 사용자 개인화에 있어 발판이 되는 발화에 대한 가이드라인을 제시할 수 있을 것이다.

한편, 이 연구는 한계점 또한 지니고 있다. 첫 번째 한계점은 음성 에이전트의 종류와 응답이 여러 시도를 거치긴 했지만 어느 정도 제한적이었다는 점이며, 이 연구에서는 특정 음성 에이전트만을 사용했기에 현재까지 개발된 다양한 대화 템플릿이 모두 적용되지 않았다는 한계가 있다. 두 번째 한계점은 음성 에이전트의 비언어적 요소를 다양하게 구성해 각 오류 상황에서의 감성 대화 요소의 효과를 알아보지 못하였다는 점이다. 대화 상호작용에 있어 목소리, 음조, 감정, 소리 크기, 속도 및 시각적 요소 등의 다양한 비언어적 요소는 사용자와의 사회적 관계에 영향을 미칠 수 있기에[15], 이 또한 대화 오류 상황에서의 불편에 있어 영향을 미칠 수 있다. 세 번째로 사용자 조사에 있어 응답 표본이 부족하다는 한계가 있다. 따라서 추후 연구에서 본 연구에서 정리한 프레임워크와 가이드를 참고하여, 많은 수를 표본으로 조사하고 검증할 필요가 있다.

이러한 시사점과 한계점을 바탕으로 후속 연구에서는 실증적인 재응답 대화 가이드 디자인을 통해 사용자 경험이 어떻게 달라지는지 살펴볼 수 있다. 더불어 다양한 비언어적 및 감성적 대화 요소를 활용하여 이에 따른 오류 상황에서의 사용자 경험의 차이를 살펴볼 수 있을 것이므로, 추후 심도 있는 연구가 진행될 것을 제안한다.

References

- [1] K. E. Jo, S. I. Kim, A study on User Experience of Artificial Intelligence speaker, Journal of Korea Convergence Society, (2018), Vol.9, No.8, pp.127-133.

DOI: <https://doi.org/10.15207/JKCS.2018.9.8.127>

- [2] Lee Jihee, Jeon Sowon, Lee Jongtae, A study of Times Pressure Affecting on the User's Acceptance of Convergence Technologies : Focusing on the Diffusion of A.I Speaker, Proceedings of the Korea Technology Innovation Society, Korea Technology Innovation Society, (2018)
Available from: <https://innovation.jams.or.kr/co/main/jmMain.kci>
- [3] M. Y. Kang, Problems and improvement plans for AI home appliances: Focusing on voice recognition speakers, Korea Consumer Agency, (2017)
Available from: <https://www.kca.go.kr/smartconsumer/sub.do?menukey=7301&mode=view&no=1002628710&page=19>
- [4] K. S. Suh, A Study on Self-Repair Speech by Native Korean Speakers, Seoul National University, Doctoral dissertation, (2020)
- [5] Dasom Park, Seung-Min Lee, A Study on the Change of User Expectation by Accumulation of Voice Assistant Usage Experience, Journal of Digital Contents Society, (2022), Vol.23, No.7, pp. 1267-1275.
DOI: <https://doi.org/10.9728/dcs.2022.23.7.1267>
- [6] C. Nass, J. Steuer, E. Tauber, Computers are social actors, Proceedings of ACM CHI '94 Conference on Human Factors in Computing Systems, New York: ACM Press, (1994), Vol.1, pp.72-78.
DOI: <https://doi.org/10.1145/191666.191703>
- [7] B. Reeves, C. Nass, The media equation: How people treat computers, television, and new media like real people and places, New York: Cambridge University Press, (1996)
- [8] J. Cassell, Embodied Conversational Interface Agents, Communications of the ACM, (2000), Vol.43, No.4, pp.70-78.
DOI: <https://doi.org/10.1145/332051.332075>
- [9] Na Ju Yeoun, Heo Jeong Yun, Pan Younghwan, Classification of Customer Utility for Voice Interactive Function of AI Personal Assistant Services : Based on KANO Model, Design Convergence Study, (2017), Vol.16, No.4, pp.67-80.
DOI: <https://doi.org/10.5392/JKCA.2019.19.03.138>
- [10] Jongsoo Lim, Jinho Choi, Hyemin Lee, Measuring the Perceived Anthropomorphism of an AI Conversational Agent : Scale Development and Validation, Korean Journal of Journalism & Communication Studies(Korean Society For Journalism And Communication Studies), (2020), Vol.64, No.4, pp.436-470.
DOI: <https://doi.org/10.20879/kjics.2020.64.4.012>
- [11] Han Youngju, A Study on Communication between AI (Artificial Intelligence) Voice Assistant and AI Speaker Users and Services Usage : Focusing on Anthropomorphism Personality, Social Presence, and Personalization, Journal of Communication Science, (2021), Vol.21, No.3, pp.225-275.
UCI: I410-ECN-0102-2022-000-000745671
- [12] Min Hee Kim, Namkung Kiechan, Deriving a structural model for reactivation of smart speaker services : Cancellation of command and interruption of utterance, Journal of Digital Art Engineering & Multimedia, (2022), Vol.9, No.1, pp.93-104.
DOI: <https://doi.org/10.29056/jdaem.2022.03.09>
- [13] Tae Ra Heo, Study on the Discourse of Artificial Intelligence Speech Recognition Service, Ewha Women's University, Master Thesis, (2018)
- [14] Z. Dornyei, M. L. Scott, Communication Strategies in a Second Language: Definitions and Taxonomies, Language Learning, (1997), Vol.47, No.1, pp.173-210.
DOI: <https://doi.org/10.1111/0023-8333.51997005>
- [15] K. Seaborn, N. P. Miyake, P. Pennefather, M. Otake-Matsuura, Voice in Human-Agent Interaction: A Survey, ACM Computing Surveys, (2021), Vol.54, No.4, pp.1-43.
DOI: <https://doi.org/10.1145/3386867>