

1인 가구 Z세대 사용자 감성기반 음성사용자인터페이스(VUI) 디자인 설계에 대한 연구

Siri를 중심으로

An Exploratory Study on Emotion-Centered Voice User Interface(VUI) Design for Generation Z Single-Person Households

Focusing on Siri

주 저 자 : 오지위 (Wu, Zhi Wei) 한양대학교 디자인학과 석박사과정

교 신 저 자 : 송민정 (Song, Min jeong) 한양대학교 디자인대학 커뮤니케이션디자인학과 교수
mjung11@hanyang.ac.kr

<https://doi.org/10.46248/kidsr.2025.3.234>

접수일 2025. 07.20. / 심사완료일 2025. 08. 30. / 게재확정일 2025. 09. 13. / 게재일 2025. 9. 30.

Abstract

This study explores strategies to optimize the emotional experience of voice user interfaces (VUIs) in AI assistants, focusing on Generation Z single households in their 20s. The research proceeded in four stages: user survey, case analysis, prototype design with usability testing, and emotion recognition technology analysis, applying the Person-Artifact-Task (P-A-T) model to derive emotion-based VUI design strategies. In UI, visual feedback elements were emphasized, while in UX, conversational naturalness, emotional responsiveness, and adaptability to user states were considered. Usability testing revealed that users perceived personalized dialogue and emotional reactions as key satisfaction factors, with visual feedback enhancing immersion and emotional engagement. These findings suggest that AI assistants can evolve beyond functional tools into emotionally connected digital companions, and the study provides methodological foundations for future personalized VUI design.

Keyword

음성 인식 AI 비서(Voice Recognition AI Assistant), 음성 사용자 인터페이스(Voice User Interface), Z세대(Z Generations), UX 디자인(UX Design)

요약

본 연구는 음성 사용자 인터페이스(VUI)의 일상화 속에서 Z세대 20대 1인 가구 사용자의 감성적 요구에 주목하여, 음성 인식 AI 비서의 정서적 경험 최적화 방안을 탐색하였다. 연구는 사용자 조사, 사례 분석, 프로토타입 설계 및 사용성 테스트, 감정 인식 기술 분석의 4단계로 진행되었으며, P-A-T 모델을 적용해 감성 기반 VUI 설계 전략을 도출하였다. UI는 시각적 피드백, UX는 대화의 자연스러움·감정 대응·상황 적응성을 중점으로 고려하였다. 사용성 테스트 결과, 사용자들은 인격화된 대화와 감정적 반응을 핵심 만족 요인으로 인식했으며, 시각적 피드백이 몰입 경험을 강화함을 확인하였다. 본 연구는 AI 비서를 정보 전달자를 넘어 정서적 디지털 동반자로 발전시킬 수 있는 가능성과, 향후 개인화된 VUI 설계를 위한 방법론적 기초를 제시한다.

1. 서론

- 1-1. 연구 배경
- 1-2. 연구 목적 및 방법

2. 이론적 배경

- 2-1. 음성 인식 AI 비서의 기술 발전과정
- 2-2. 1인 가구 Z세대 사용자군의 심리 및 특성
- 2-3. 음성 사용자 인터페이스(VUI) 디자인 개념

3. 스마트폰 음성 인식 AI 비서 사례

- 3-1. 사례 분석

- 3-2. 사용자 만족도 조사

4. 스마트폰 음성 인식 AI 비서 사용자 분석

- 4-1. 사용자 설문조사
- 4-2. 심층 인터뷰

5. 음성 사용자 인터페이스 시나리오 별 시뮬레이션 분석

6. 음성 사용자 인터페이스 디자인 설계

- 6-1. 음성 사용자 인터페이스 최적화 방안

- 6-2. 시나리오 별 시뮬레이션 분석
- 6-3. 음성 사용자 인터페이스 디자인 제안
- 6-4. 사용성 테스트 및 분석

7. 결론

참고문헌

1. 서론

1-1. 연구의 배경 및 목적

새로운 과학 기술 환경에서 산업 디자인은 과거와 크게 변화했으며 특히 현재 인공지능 시대에 고객의 요구가 계속 증가하고 빅 데이터 기술은 심오하고 광범위한 산업 디자인의 확장을 가져왔다. 산업 디자인은 점차 전통적인 디자인 모델에서 벗어나 스마트 시대의 특성을 부여한다. 그리고 사람, 제품, 사회 및 환경 간의 연결을 촉진한다.¹⁾ 사용자 경험 중심의 설계가 강조되는 흐름 속에서 음성 상호작용은 단순한 입력 수단을 넘어 사용자와의 감정적 연결을 형성하는 핵심 요소로 부상하고 있다. 지능형 제품의 음성 상호작용에 있어서 AI는 더 이상 수동적이고 감정이 없는 대상이 아니다.²⁾ 과거의 음성 사용자 인터페이스가 명령 수행과 정보 전달에 치중되었다면, 최근의 음성 인식 AI 비서는 사용자의 정서 상태를 인식하고 이에 적절히 반응함으로써 감성적 교감을 유도하고, 보다 인간적인 인터랙션을 가능하게 한다. 음성의 톤, 속도, 리듬, 맥락에 따른 감정 표현 등 다양한 음성 요소들은 사용자에게 심리적 안정감과 공감대를 제공하며, 이는 서비스나 제품에 대한 신뢰와 몰입, 나아가 정서적 애착으로 이어진다. 특히 1인 가구 20대 사용자층에게 음성 인식 AI 비서는 디지털 환경 속에서 정서적 고립감을 완화하고, 일상적인 감정 소통의 대안으로 기능할 수 있다. 본 연구는 음성 사용자 인터페이스 디자인(UI Design)에 나타난 서로 다른 감정(기쁨/평온/슬픔) 상호작용 요소를 분석하고, 이러한 요소들이 1인 가구 z세대 사용자 경험의 질과 정서적 교감 방식에 대해 알아 본다. 본 연구는 AI 비서가 단순한 정보 전달자를 넘어 감성적 디지털 동반자로 발전할 수 있는 잠재적 가능성을 밝히는 동시에, 향후 개인화된

VUI 설계를 위한 방법론적 기초를 제시하였다. 연구의 궁극적 목적은 Z세대 20대 1인 가구 사용자의 감성적 요구를 충족시키는 음성 사용자 인터페이스(VUI) 설계 전략을 탐색하고, 음성 인식 AI 비서가 사용자 경험을 최적화하며 정서적 교감을 촉진하는 가치를 검증함으로써, 기능적 도구에서 점차 감성적 동반자적 속성을 지닌 디지털 파트너로 발전할 수 있음을 규명하는 데 있다.

1-2. 연구 방법

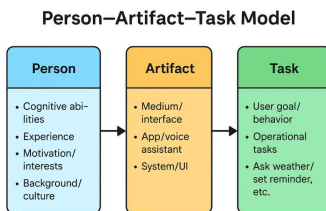
본 연구에서는 스마트폰 음성 인식 AI 비서를 중심으로 사용자 경험을 분석하고, 감성 교감이 가능한 음성 사용자 인터페이스(VUI) 디자인 설계를 위해 다음과 같은 연구 방법을 적용하였다. 연구 대상은 감성 음성 인터페이스를 제공하는 대표적인 음성비서 서비스와 이를 경험한 1인 가구 Z세대인 20대 사용자들로 설정하였다. 본 연구는 사용자 경험 분석을 위해 사람(P: Person), 도구(A: Artifact), 과업(T: Task) 요소 간의 상호작용을 중심으로 하는 PAT 모델(그림 1)³⁾을 이론적 틀로 활용하였다. 이 모델은 사용자가 음성 인터페이스를 통해 특정 과업을 수행할 때, 시스템과의 적합성을 분석하는 데 유용하다. 사용자(P)는 감정 표현에 민감하고 디지털 기술에 익숙한 20대 1인 가구이며, 도구(A)는 감성 기반의 음성 인식 AI 비서로, 단순 명령 수행을 넘어 감정 표현, 맥락 인식, 개인화된 피드백 등의 특징을 지닌다. 과업(T)은 사용자가 일상 속에서 수행하는 다양한 음성 인터랙션이며, 이는 기능적 요구뿐 아니라 정서적 만족까지 포함한다. 연구 방법은 다음과 같다. 첫째, 사례 분석: 스마트폰 음성 인식 AI 비서의 인터페이스 디자인과 상호작용 방식을 분석하여 주요 감성 표현 요소를 도출한다. 둘째, 설문 조사 및 인터뷰: 실제 사용자들을 대상으로 사용자 만족도와 감정적 반응을 조사하여 음성 상호작용의 효과를 정량·정성적으로 분석한다. 본 조사에서는 Z세대 1인 가구 사용자 52명을 대상으로 하였으며,

1) 이양. "인공지능(AI) 기술 기반 인터랙션 디자인과 감성디자인의 상관성 연구." 국내박사학위논문 중앙대학교 대학원, 2023. 서울

2) 이지선, 신재현, 유학윤, and 이정인. "공감과 격려를 통해 행동 변화를 유도하는 음성 AI 인터페이스 디자인 : 청년 1인 가구를 중심으로." Design Works 8.1 (2025): p.96

3) Nardi, B. A. 『Context and consciousness : Activity theory and human-computer interaction』, MIT Press, 2000, p.17

표집은 편의표집(convenience sampling) 방식으로 이루어졌다. 대상자들의 인구학적 특성(성별, 연령 구간, 학력, 직업)을 기초 통계로 분석하여 표본의 대표성을 확보하였다. 설문 문항은 선행연구를 토대로 구성하였으며, 전문가 2인의 검토와 5명의 예비조사를 통해 이해 가능성과 타당성을 사전 검증하였다. 또한, 심층적 정성분석을 위해 설문 응답자 중 성별, 사용 경험, 거주 형태가 서로 다른 5명을 인터뷰 대상으로 선정하여 다양성을 확보하였다. 이를 통해 양적·질적 자료 모두에서 연구의 타당성을 강화하였다. 셋째, VUXI 시나리오 분석: 특정 상황에서 음성 응답의 톤, 피드백 방식 등이 사용자 경험에 미치는 영향을 PAT 모델을 기준으로 다각도에서 검토한다. 이를 통해 본 연구는 감성 기반 음성 사용자 인터페이스 디자인 설계를 통해 Z세대와의 감성적 교감을 위한 방안을 제시하고, 향후 스마트 폰 음성 인식 AI 비서 개발을 위한 실용적 방향성을 제시하고자 한다.



[그림 1] PAT 모델

2. 이론적 배경

2-1. 음성 인식 AI 비서의 기술 발전과정

최근에는 시와 음성인식을 결합한 AI 비서를 활용한 서비스가 확대되고 있다. 이러한 음성 사용자 인터페이스는 새로운 형태의 인터랙션 방식으로 빠르게 발전하고 있다. 사용자 관점에서 음성 인터랙션의 효용을 살펴보면, 인간에게 친숙한 정보 입력 및 출력 방식이기에 비교적 학습의 필요성이 낮고 쉽게 사용할 수 있다.⁴⁾ 음성 인식 AI 비서는 음성 인식 기술과 자연어 처리(NLP), 기계 학습(Machine Learning), 그리고 딥러닝(Deep Learning) 기술의 융합을 통해 발전해 왔다. 초기의 음성비서는 사용자의 명령을 단순히

4) 나주연, 허정윤, 반영환, "음성 인식 기반 AI 개인 비서 서비스의 대화형 기능에 대한 고객 효용 분류 : KANO모형을 중심으로," 디자인융복합연구 16.4 (2017) pp.67-80.

인식하고 정해진 응답을 제공하는 수준이었으나, 최근에는 사용자의 의도와 맥락을 이해하고, 감정 상태까지 파악하여 보다 인간 중심적인 대화를 가능하게 하는 방향으로 고도화되고 있다.

Google Assistant, Amazon Alexa, Apple Siri와 같은 상용되고 있는 음성비서는 지속적인 업데이트를 통해 감정 분석, 음성 톤 조절, 사용자 맞춤형 피드백 기능을 탑재하고 있으며, 사용자와의 정서적 상호작용을 가능하게 하여 UX의 새로운 전환점을 만들어내고 있다.

기술적 진화는 음성 사용자 인터페이스(VUI) 설계 방식에도 변화를 불러왔다. 기존의 시각 중심 UI를 넘어, 음성 중심 또는 멀티모달 인터페이스로 확장됨에 따라 사용자 경험의 연속성, 몰입감, 그리고 정서적 만족도 향상이 중요한 디자인 요소로 부상하였다. 특히 감정적 피드백을 제공하는 음성 UI는 사용자에게 일방적인 정보 전달이 아닌, 상호적인 감정 교류의 창구로 작용하며, 이는 사용자 경험(UX)의 감성적 품질 향상에 기여하고 있다.

2-2. 1인 가구 Z세대 사용자군의 심리 및 특성

Z세대는 1990년대 후반부터 2010년대 초반에 출생한 세대를 일컫는 용어로 스마트폰과 인터넷을 통한 첨단 디지털 기술에 능숙한 신세대를 말한다.⁵⁾ Z세대는 태어날 때부터 스마트폰, 인터넷, SNS 등 디지털 환경에 노출되어 성장했기 때문에 기술 친화적 성향이 뚜렷하며, 인공지능 기반 서비스나 인터페이스에 대한 수용도 또한 매우 높다. 실제로 보고서에 따르면, 미국의 Z세대 중 매월 음성 비서를 사용하는 비율은 2023년 51%에서 2027년 64%로 증가할 것으로 예측되며, 이는 해당 세대가 VUI 기술을 빠르게 수용하고 있음을 보여준다 (eMarketer, 2023).⁶⁾ 이들은 기능적 효율성뿐 아니라 감정적 교감, 반응의 진정성, 맞춤형 소통 경험 등을 중요하게 평가하며, 인터페이스가 사용자 자신의 감정 상태를 이해하고 반영해 주는지 여부를 핵심 가치로 본다. 따라서 UX 디자인에서 Z세대는 단순한 기술 소비자가 아닌 경험 중심의 감성 사용자로 간주되어야 하며, 특히 음성 사용자 인터페이스

5) Oh, S., Pay attention to the Z generation. Marketing, 52(1) (2018), pp.57-62.

6) eMarketer. (2023, October 2). Data drop: Gen Z leading voice assistant growth. Insider Intelligence. <https://www.emarketer.com/content/data-drop-gen-z-leading-voice-assistant-growth>

설계에서는 이들의 정서적 반응성과 몰입 특성을 고려한 전략적 접근이 필수적이다.

1인 가구 중 가장 큰 비중을 차지하는 연령층은 20대이다.⁷⁾ 1인 가구의 20대 초반 사용자군은 디지털 환경에 익숙한 Z세대에 해당하며, 인공지능 기반 기술에 대한 수용도가 높고 새로운 기술을 일상에 빠르게 통합하는 경향을 보인다. 특히 이들은 일상생활에서 스마트폰, 시 스피커, 음성비서 등을 자주 활용하며, 감정적 연결감과 실시간 반응성을 중요하게 여긴다.

2023년 기준 통계청 자료에 따르면 전체 가구 중 1인 가구는 약 782만 9천 가구로 전체의 35.5%를 차지하며, 지속적인 증가 추세를 보이고 있다 [표 1]. 이는 2018년 29.3%에서 2023년 35.5%로 6.2%p 증가한 수치로, 전통적인 4인 가족 중심의 가구 구조가 해체되고 개별적·독립적인 1인 가구가 증가하고 있음을 시사한다. 현재 1인 가구의 비중은 전체 가구 중 33.4%를 차지하며, 2050년에는 40% 이상의 비율을 넘어설 것으로 예상되고 있다.⁸⁾

이와 같은 통계적 변화는 음성 사용자 인터페이스(VUI) 설계에서 사용자 유형의 다변화, 특히 청년 1인 가구 사용자의 라이프스타일과 정서적 요구를 고려한 사용자 중심의 음성 사용자 인터페이스 설계가 요구됨을 나타낸다.

1인 가구 Z세대이 사용자군은 가족 또는 공동체와의 일상적 접촉이 제한된 생활 구조로 인해, 외로움, 고립감, 심리적 공백을 경험할 가능성이 상대적으로 높다. 이에 따라 이들은 인간처럼 대화하고 공감해주는 감성 인터페이스에 대해 높은 호감도와 몰입도를 보이며, AI 기술을 감정적 동반자 또는 생활 조력자로 인식하는 경우가 많다.

또한, 이들은 음성 사용자 인터페이스 디자인에 있어서 직관성, 개인화, 그리고 감성적 반응성을 중요한 요소로 평가하며, 감정 상태를 반영한 맞춤형 상호작용을 긍정적으로 수용한다. 따라서 음성 인식 AI 비서는 이 사용자군에게 정보 제공 이상의 심리적 안정, 정서적 유대, 그리고 일상 속 '대화 상대로서의 역할'

7) 통계청. 성, 연령 및 거주지의 종류별 1인 가구-사군구. https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1PL1509&conn_path=12 2024.03.06.

8) 오기영, 박해경, 청년 1인 가구 주거특성에 따른 편리함 이용 추가경고 인식 및 이용의향에 관한 연구, 한국실내디자인학회 논문집, 2023, 32(1), pp.135-142.

을 수행할 수 있는 잠재력을 가진다.

[표 1] 1인 가구 연령과 성별 통계

< 가구원수별 가구 현황 >

(단위: 천 가구, %)

	전체 가구	1인가구	비중	2인가구	비중	3인가구	비중	4인 이상	비중
2015	19,111	5,203	27.2	4,994	26.1	4,101	21.5	4,813	25.2
2016	19,368	5,398	27.9	5,067	26.2	4,152	21.4	4,751	24.5
2017	19,674	5,619	28.6	5,260	26.7	4,179	21.2	4,616	23.5
2018	19,979	5,849	29.3	5,446	27.3	4,204	21.0	4,481	22.4
2019	20,343	6,148	30.2	5,663	27.8	4,218	20.7	4,315	21.2
2020	20,927	6,643	31.7	5,865	28.0	4,201	20.1	4,218	20.2
2021	21,448	7,166	33.4	6,077	28.3	4,170	19.4	4,036	18.8
2022	21,774	7,502	34.5	6,261	28.8	4,185	19.2	3,826	17.6
2023	22,073	7,829	35.5	6,346	28.8	4,195	19.0	3,703	16.8

자료: 통계청, 「인구주택총조사」, * 통계표 (46쪽) 참조
 주: 1) 일반가구(①가족으로 이루어진 가구, ②가족과 5인 이하의 남남이 함께 사는 가구, ③1인 가구, ④가족이 아닌 남남끼리 함께 사는 5인 이하의 가구, ⑤-④ 일반가구에 외국인도 포함)

< 1인 가구 및 비중 >

연도	1인가구 (천 가구)	비중 (%)
2015	19,111	5.203
2016	19,368	5.398
2017	19,674	5.619
2018	19,979	5.849
2019	20,343	6.148
2020	20,927	6.643
2021	21,448	7.166
2022	21,774	7.502
2023	22,073	7.829

자료: 통계청, 「인구주택총조사」

< 1인 가구 성·연령대별 비중(2023) >

성별	연령대	비중 (%)
남성	20세 이하	18.7
	30-39세	21.8
	40-49세	18.0
	50-59세	17.8
	60-69세	18.6
	70세 이상	9.9
여성	20세 이하	18.4
	30-39세	21.8
	40-49세	19.2
	50-59세	17.6
	60-69세	18.6
	70세 이상	28.3

자료: 통계청, 「인구주택총조사」

2-3. 음성 사용자 인터페이스(VUI)디자인 개념

디지털 인터페이스가 시각 중심에서 점차 다감각적 상호작용으로 확장됨에 따라, 음성 사용자 인터페이스는 새로운 사용자 경험 패러다임으로 주목받고 있다. 기존의 UI(User Interface)는 시각적 요소(아이콘, 버튼, 레이아웃 등)를 중심으로 구성되었으며, UX(User Experience) 설계 또한 시각 정보 처리와 물리적 인터랙션을 중심으로 이루어졌다. 그러나 인공지능(AI) 기술과 자연어 처리(NLP)의 발전은 사용자가 음성을 통해 기기와 자연스럽게 상호작용할 수 있는 환경을 조성하며, 디자인의 개념적 영역을 확장하고 있다. AI와의 커뮤니케이션 기술은 사용자와 대화를 나누는 관점에서 발전하였다. 기계를 부르고 명령을 내리던 것에서 친구와 대화하듯 자연스러운 대화의 경험으로 AI와 사용자 간의 인터랙션이 이루어진다.⁹⁾

음성 사용자 인터페이스(VUI)는 음성인식 기술을 기반으로 기기가 인간의 음성에서 의미를 분석하여 인간과 기기 간의 상호작용을 가능하게 하는 인터페이스 방식이다.¹⁰⁾ 이는 특히 운전 중이거나 손을

9) 조예림. "인공지능 비서의 정보 검색 향상을 위한 UX/UI 디자인 연구." 국내박사학위논문 서울대학교 대학원, 2021. 서울

10) Michael, H., James P. Giangola, Jennifer Balogh (2004), 『Voice user interface design, Addison-Wesley』, Chapter 1.

자유롭게 사용할 수 없는 상황, 또는 디지털 리터러시가 낮은 사용자에게 높은 접근성과 편의성을 제공한다. 나아가 음성 인식 AI 비서와 같이 감정을 표현하고 상황을 인지하는 인터페이스는 단순한 기능 수행을 넘어 사용자와의 정서적 유대 형성을 가능하게 하며, UX 디자인의 정서적 차원을 강화한다.

따라서 음성 사용자 인터페이스(VUI) 디자인은 단지 음성 명령을 인식하고 처리하는 기술적 구조를 넘어, 사용자의 맥락을 이해하고 정서적 반응을 유도하는 경험 중심 디자인으로 진화하고 있다. 이 과정에서 음성의 톤, 억양, 응답 시간, 언어 선택 등은 모두 사용자 경험에 영향을 미치는 핵심 디자인 요소로 작용하며, 디자이너는 이와 같은 청각적 표현 수단을 통해 사용자 중심 인터랙션을 설계해야 한다.

3. 스마트 폰 음성 인식 AI 비서 사례

3-1. 사례 분석

최근의 스마트폰은 문자를 입력하는 터치 인터페이스에서 이동 중 조작이나 음성검색 서비스뿐만 아니라 음성의 맥락을 이해하고 스스로 학습할 수 있는 능력을 갖춘 인공지능형 음성 인터페이스로 변화하고 있다.¹¹⁾ 본 연구는 사용자와 음성 상호작용 특성에 대한 심층적 이해를 목적으로, 현재 시장에서 상용되고 있는 대표적인 음성비서 제품인 Apple의 Siri, Google Assistant, Samsung Bixby를 대상으로 분석을 실시하였다. 이들 제품은 모바일 플랫폼을 중심으로 음성 사용자 인터페이스를 구현하고 있는 주요 사례로, 본 연구는 1인 가구에 거주하는 20대 사용자 52명을 대상으로 설문조사를 실시하고, 기능 범위, 감성 표현 능력, 맥락 인식 수준, UI/UX 디자인, 사용자 만족도 등 다섯 가지 항목을 기준으로 비교가 이루어졌다 [표 2]. 기능 범위 측면에서 Siri는 다양한 기능을 수행하나 일부 앱 연동에 제한이 있으며, Google Assistant는 광범위한 기능과 우수한 앱 연동성을 보이고, Bixby는 삼성 기기 중심으로 기능이 최적화되어 있으나 범용성에는 한계가 존재하였다. 감성 표현 능력에 있어 Siri는 경중하고 감성적인 언어 사용이 특징이며, Google Assistant는 다소 기계적인 응답을 보이고, Bixby는 친근한 어투를 사용하나 감성 표현은

제한적이었다. 맥락 인식 수준에서는 Google Assistant가 높은 문맥 추론 능력을 보인 반면, Siri는 이전 대화를 부분적으로 반영하며, Bixby는 상대적으로 낮은 맥락 반응 수준을 나타냈다. UI/UX 측면에서 Siri는 iOS와의 유기적 통합을 통해 자연스러운 음성 피드백과 시각 인터페이스 연계성을 보였고, Google Assistant는 구글 생태계와의 원활한 연동이 강점이며, Bixby는 삼성 UX와 통합되어 있으나 범용 인터페이스 측면에서는 다소 제한적이었다.

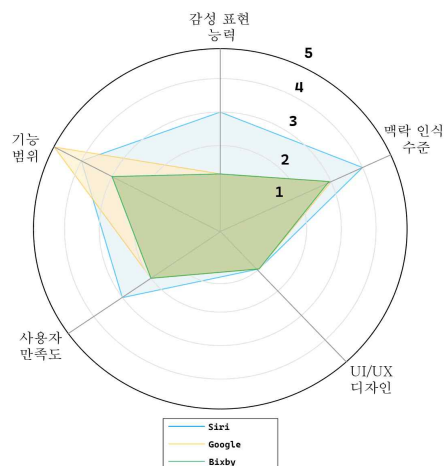
[표 2] 음성 인식 AI 비서 분석

음성 인식 AI 비서	기능 범위	감성 표현 능력	맥락 인식 수준	UI/UX 디자인	사용자 만족도
Siri	4	3	4	1	3
Assistant	5	1	5	1	2
Bixby	3	1	2	1	2

3-2. 사용자 만족도 조사

사용자 만족도 조사 결과[그림 2], Siri는 사용 편의성과 감성 반응에서 긍정적인 평가를 받았고, Google Assistant는 빠른 응답과 정보 처리에서 높은 만족도를 보였으며, Bixby는 국내 사용자들 사이에서 높은 친숙도와 만족도를 나타냈다.

종합적으로, Siri는 감성 표현 및 감성 공감 유도 측면에서 다른 두 제품에 비해 일정한 우위를 보였지만, 전반적으로는 여전히 개선의 여지가 크며, 특히 감성 표현 메커니즘과 UI/UX 디자인의 조화로운 통합 측면에서 부족한 점이 드러난다. 이에 본 연구는 Siri를 주



[그림 2] 사용자 만족도 조사 결과

11) 김민정 (2017), 『음성 사용자 인터페이스 디자인 고찰 연구-스마트폰의 지능형 음성인식 서비스를 중심으로』, 한국디자인문화학회지, P.61

요 분석 대상으로 삼아, 감정 표현과 UI/UX 인터랙션 디자인의 통합 문제에 주목하고, 보다 구체적이고 실질적인 디자인 전략을 제안하고자 한다. 후속 설계 제안 단계에서는 본 연구의 결과가 사용자 경험 시나리오 구성 및 UI/UX 최적화 방안 도출을 위한 핵심 이론적 근거로 활용될 수 있으며, 높은 실천 가능성과 응용 가치를 지닌다.

4. 스마트폰 음성 인식 AI 비서 사용자 분석

4.1. 사용자 설문조사

본 연구는 1인 가구로 거주하는 20대 사용자 52명을 대상으로 설문조사를 실시하고[표 3], Nielsen(2012)는 “정성적 사용성 연구에서는 테스트 사용자 수를 한 숫자로 요약하면 5명이다. 5명을 테스트하는 것으로 훨씬 더 많은 참여자를 사용할 때와 거의 동일한 수의 사용성 문제를 찾을 수 있다”고 언급했다.¹²⁾ 그래서 이 중 5명을 선별하여 심층 인터뷰를 진행하였다.

[표 3] 사용자 설문조사 대상(n=52)

특성 구분	세부 항목	인원수(n)	비율(%)
성별	남성	26	50.0
	여성	26	50.0
연령 구간	20-23세	18	34.6
	24-26세	20	38.5
	27-29세	14	26.9
학력	학부 이하	8	15.4
	학부	32	61.5
	대학원 이상	12	23.1
직업	학생	24	46.2
	회사원	12	23.1
	프리랜서	8	15.4
	기타(아르바이트/창업 등)	8	15.4

본 연구의 설문지는 만족도, 감정 표현 인식 정도, 정서적 반응세 가지 방향으로 구성되었으며, 총 5개 문항을 5점 리커트 척도로 측정하였다. 본 조사를 실

시하기에 앞서 문항의 명확성과 내용의 타당성을 확보하기 위해 10명의 조사자를 대상으로 소규모 예비조사(Pilot Test)를 진행하였다. 예비조사 결과, 응답자들은 문항의 의미를 이해하거나 응답하는 과정에서 뚜렷한 어려움을 보이지 않았으며, 전체적으로 양호한 조작성(operability)을 나타냈다. 이를 토대로 일부 문항의 표현을 수정·보완한 뒤 본 조사를 실시하였으며, 최종 조사 결과는 [표 4]에 제시하였다.

[표 4] 사용자 설문조사 결과 요약 (n=52)

항목 번호	조사 문항 내용	긍정 응답률 (%)*	평균 점수 (5점 척도)
Q1	음성 인식 AI 비서 사용 시 감성적 반응(공감, 위로 등)이 중요하다	78.8%	4.21
Q2	음성의 억양과 톤이 전체 경험에 영향을 준다고 생각한다	84.6%	4.38
Q3	감성 표현이 없는 음성 응답은 거리감을 느끼게 만든다	65.4%	3.92
Q4	음성 응답의 피드백 속도가 사용 만족도에 영향을 준다	71.2%	4.05
Q5	감정 기반 상호작용이 서비스 지속 이용 의사에 영향을 준다	69.2%	4.12

Q1 문항에서는 “음성 인식 AI 비서 사용 시 감성적 반응(공감, 위로 등)이 중요하다”는 진술에 대해 78.8%가 긍정적으로 응답하였으며, 평균 점수는 4.21점으로 나타났다. 이는 사용자들이 단순한 정보 제공을 넘어 감성적 소통을 중시하고 있음을 시사한다.

Q2 문항은 “음성의 억양과 톤이 전체 경험에 영향을 준다”는 항목으로, 84.6%의 높은 긍정 응답률과 4.38점의 평균 점수를 기록하였다. 이는 음성의 감성적 표현이 사용자 경험에 미치는 영향이 크다는 것을 보여준다.

Q3 문항에서는 감정 표현이 없는 응답이 거리감을 형성한다고 느낀 비율이 65.4%였고, 평균 점수는 3.92점으로 비교적 낮게 나타났으나, 여전히 다수 사용자가 감성적 결여를 인식하고 있음을 알 수 있다.

Q4에서는 피드백의 속도나 리듬이 사용 만족도에 영향을 준다고 응답한 비율이 71.2%였으며, 평균 점수는 4.05점이었다. 이는 음성 인터랙션에서 비언어적

12) Nielsen, J. (2012, June 3). How many test users in a usability study? Nielsen Norman Group. Retrieved from <https://www.nngroup.com/articles/how-many-test-users>

요소 또한 중요한 UX 요소로 작용함을 의미한다.

마지막으로 Q5 항목은 “감정 기반 상호작용이 서비스 지속 이용 의사에 영향을 준다”는 진술에 대해 69.2%가 긍정적으로 응답하였고, 평균 4.12점을 나타내 사용자들이 감성 요소가 장기적 사용 결정에 중요한 요인으로 인식하고 있음을 보여준다. 전반적으로 모든 문항에서 평균 점수가 4점 이상으로 나타났으며, 이는 감정 표현과 음성의 정서적 품질이 사용자 만족도와 지속 사용 의도에 밀접한 연관이 있음을 시사한다. 이러한 결과는 향후 음성 사용자 인터페이스(VUI) 디자인 설계 시 감성적 요소의 통합이 필수적임을 뒷받침하는 근거로 활용될 수 있다.

4-2. 심층 인터뷰

설문조사에 응답한 52명 중 가장 많이 사용하는 모바일 음성 인식 AI 비서가 Siri로 나타났으며, 이에 따라 본 연구의 대상 역시 Siri를 중심으로 설정하였다. 그 중 5명의 사용자를 대상으로 반구조화 심층 인터뷰를 진행하였으며, 각 인터뷰는 약 40-60분 동안 진행되었다. 인터뷰는 참여자의 실제 생활 환경 또는 온라인 화상 환경에서 이루어졌으며, 전 과정을 녹음 전사하여 분석 자료로 활용하였다. 분석은 질적 연구의 주제 분석(thematic analysis) 방법을 적용하여, 개방형 코딩 → 축 코딩 과정을 거쳐 핵심 주제를 도출하였다. 인터뷰 질문은 사용 맥락, 감정적 경험, 긍정·부정 반응, 신뢰 및 지속 사용 의도, 개선 요구 등 다섯 가지 방향을 중심으로 설계되었다.

심층 인터뷰 결과, 음성 인식 AI 비서와의 감정적 상호작용에 대한 다양한 인사이트를 도출할 수 있었다 [표 5]. 감정이 결여된 음성 응답은 사용자에게 정서적 거리감을 유발하고 사용 의지를 떨어뜨리는 부정적 경험으로 작용하였다(P3). 기계적이고 감정 없는 말투에 대해 “기분이 상하고 정이 안 간다”는 응답을 통해, 음성의 억양과 말투가 정서적 몰입도 형성에 핵심 역할을 한다는 점이 확인되었다.

반면, 공감적이고 따뜻한 말투는 사용자에게 정서적 안정감을 제공하며, 특히 부정적 감정 상태에 있는 사용자에게 긍정적 영향을 미친다는 반응도 나타났다. P6의 “기분이 안 좋을 때 다정하게 말해주면 혼자 있는 느낌이 덜해요”라는 발화는 음성 인터랙션이 단순한 명령 수행을 넘어서 감정 조절 기능을 할 수 있음을 보여준다.

또한 응답 타이밍이나 말의 흐름이 자연스럽지 않을 경우, 감정적 짜증이나 불편함을 유발할 수 있다는 점

도 확인되었다(P2). 이는 음성비서가 대화의 리듬과 감정 타이밍을 얼마나 정밀하게 맞추는지가 사용자 경험 품질에 직접적인 영향을 준다는 것을 시사한다.

P9의 응답에서는 감정적 맥락을 이해하지 못하고 엉뚱한 반응을 보일 때 실망감이 크다는 점이 드러났다. 이는 감성 기반 대화에서 ‘맥락 이해’가 중요한 사용자 신뢰 형성 요인임을 의미한다.

마지막으로, 유머나 웃음 섞인 긍정적 말투는 사용자에게 “사람 같다”는 인상을 주며, 음성비서를 더 인간적으로 느끼게 만든다는 응답이 관찰되었다(P5). 이는 감정 기반 상호작용에서 긍정적 정서 표현이 사용자 만족도와 지속 사용 의도 모두에 긍정적 영향을 미칠 수 있다는 것을 보여준다.

이상의 심층 인터뷰 결과는 음성비서의 정서적 상호작용 설계가 단순한 기능적 효율성을 넘어서 사용자 외의 감정 연결 및 신뢰 형성에 필수적임을 시사하며, 향후 감성 중심 음성 사용자 인터페이스(VUI) 디자인 설계의 실질적 방향성을 제시한다.

[표 5] 심층 인터뷰 주요 인사이트 정리 (n=5)

조사자 번호	발화 예시	인사이트 요약
P3	“감정 없는 말투로 대답 하나까 기분이 상하고 정이 안 가요.”	감정 없는 톤과 말투는 정서적 거리감을 유발하고 사용자 몰입도를 낮춤
P6	“기분이 안 좋을 때 다정하게 말해주면 혼자 있는 느낌이 덜해요.”	공감적 언어는 부정적 감정을 완화시키며 정서적 안정에 기여
P2	“말 끊기거나 타이밍 안 맞으면 답답하고 짜증나요.”	응답 흐름과 감정 타이밍이 사용자 감정 경험에 중요한 역할을 함
P9	“내 말 이해 못하고 엉뚱한 말 하면 실망스러워요.”	감정 맥락 이해 부족은 사용자 실망감 및 감정 소외를 유발함
P5	“기분 좋게 웃거나 유머 섞어서 말해주면 사람 같고 기분이 좋아요.”	유머나 긍정적 감정 표현은 인간다운 상호작용 경험을 강화함

5. 음성 사용자 인터페이스 시나리오 별 시뮬레이션 분석

음성 인식 AI 비서의 음성 사용자 인터페이스(VUI)

디자인 경험을 보다 실질적으로 이해하기 위해 실제 사용자들을 대상으로 음성 인터페이스 사용 시나리오를 중심으로 한 관찰 분석을 실시하였다. 관찰 대상은 20대 1인 가구 6명으로 한정하였으며, 이들은 모바일 환경에서 Siri를 주로 활용하는 사용자들로 선정되었다. 실험 도구로 Siri를 선택한 이유는, iOS 기반 단말기에서 기본적으로 제공되는 음성 비서로서 20대 1인 가구 사이에서 활용 빈도가 높고, 다양한 음성 비서를 혼합하여 사용할 경우 발생할 수 있는 변수를 최소화하기 위함이었다.

[표 6] 인터랙티브 레코드 표

사용자 (P)	음성 비서 (A)	과업 (T)	상호작용 방식	사용자 반응	UX 특성 평가
A	Siri	알람 설정	“내일 아침 7시에 알람 설정 해줘” → 음성 피드백 제공	명령 인식 빠름, 성공률 높음	편리성 우수, 자연스러운 상호작용
B	Siri	감정 공유	“나 오늘 좀 힘들어...” → 공감 표현 없음	감정적 피드백 부족, 아쉬움 표시	정서적 만족도 낮음, 감성 인터페이스 부족
C	Siri	날씨 확인	“오늘 날씨 어때?” → 기온 및 강수량 안내	정보 정확, 반응 빠름	정보 전달 효과적, 톤 정중
D	Siri	음악 재생	“잔잔한 음악 틀어줘” → Apple Music 실행	원하는 음악 재생, 요청 발생	추천 기능 미흡, 맥락 이해 낮음
E	Siri	일정 확인	“내일 일정 뭐 있어?” → 캘린더 연동 안내	일정 안내 정확, 사용자 만족	시스템 통합성 높음, 기능적 만족도 높음
F	Siri	메시지 전송	“엄마한테 메시지 보내줘” → 수신인 확인 요청	수신인 인식 실패, 반복 요청 필요	인식 오류 존재, 사용 피로도 상승

사용 시나리오는 실제 생활에서 자연스럽게 발생하는 상황 맥락을 중심으로 구성되었으며, 주된 상황은 아침 기상, 외출 준비, 일정 확인, 감정 상태 공유 등으로 설정되었다. 이러한 맥락 속에서 사용자가 수행한 주요 과업은 알람 설정, 날씨 확인, 메시지 전송, 음악 재생 등이며, 각 과업 수행 과정에서의 상호작용 방식과 사용자 반응을 중심으로 세부 항목을 기록하였

다[표 6].

관찰 결과, 사용자는 음성 인식 AI 비서가 단순히 정보를 제공하는 것을 넘어 정서적 표현이나 공감적 반응을 보일 때 더 높은 몰입과 만족을 경험하는 경향이 나타났다. 특히, 응답의 말투, 음성 톤, 상황 맥락에 대한 인식 수준은 음성 사용자 인터페이스(VUI) 경험 품질에 있어 중요한 영향을 미치는 요인으로 확인되었다. 반면, 반복되는 음성 인식 오류나 명령 실패는 사용자로 하여금 불편함을 야기하였으며, 이는 곧 음성 인식 AI 비서에 대한 신뢰도 저하로 이어졌다. 이러한 현상은 특히 일상 루틴 중 빠른 피드백과 정확성이 요구되는 상황에서 더욱 두드러지게 나타났다.

관찰 분석은 사용자(P), 음성비서(A), 과업(T)의 삼자 간 상호작용을 중심으로 한 PAT 모델에 기반하여 재구성되었으며, 이를 통해 감성 중심의 음성 사용자 인터페이스(VUI) 디자인 설계가 단순한 기능적 구현을 넘어서 사용자 경험 전반에 긍정적 영향을 미칠 수 있음을 시사한다. 본 결과는 후속 제안될 디자인 전략 수립의 실증적 토대로 활용될 것이다.

6. 음성 사용자 인터페이스 디자인 설계

6-1. 음성 사용자 인터페이스 최적화 방안

‘외로움’은 **Perlman & Peplau(1981)**가 정의한 바와 같이, “사회적 관계망이 양적 혹은 질적으로 부족할 때 발생하는 불쾌한 경험”¹³⁾으로 규정하였다. 또한 ‘정서적 결핍’은 충분한 정서적 지지와 친밀한 상호작용의 부재로 인해 경험되는 심리적 공허 상태를 의미한다. 본 연구에서는 정량적 지표보다는 사용자의 발화 속 감정 키워드(예: “혼자야”, “힘들어”, “보고 싶어”)와 자기보고적 진술을 중심으로 해석하여 연구 맥락에 맞게 조작적 정의를 내렸다.

이러한 개념적 기반 위에서, 본 연구는 20대 1인 가구 사용자가 일상 속에서 자주 경험하는 ‘외로움’과 ‘정서적 결핍’ 문제를 중심으로, 인공지능 음성 어시스턴트의 가족 음성 및 감정 인식 기능을 활용한 동반형 상호작용 서비스 설계 방안을 제안하고자 한다. 제안

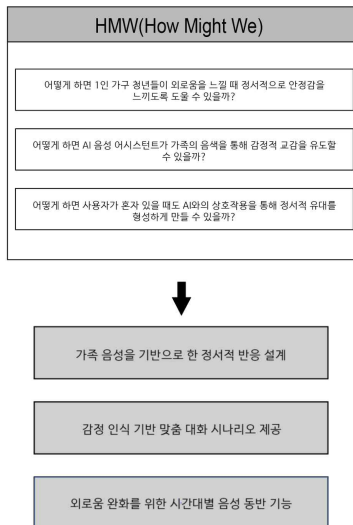
13) Perlman, D., & Peplau, L. A. (1981). Toward a social psychology of loneliness. In S. Duck & R. Gilmour (Eds.), *Personal relationships in disorder*(pp. 31–56). London: Academic Press.

하는 아이디어의 구체적인 내용은 다음과 같다.

첫째, 외로움이 극대화되는 순간에 사용자에게 적절한 위로와 정서적 동반감을 제공하는 것이다. 이를 위해 어머니 또는 가족의 목소리 특성을 반영한 음성 UI를 적용함으로써, 음성 어시스턴트가 기계가 아닌 '가족 같은 존재'로 인식되도록 하여 정서적 거리감을 줄이고 시스템과의 감정적 유대를 강화한다.

둘째, 음성 대화와 상황 시뮬레이션을 통해 심리적으로 위축된 시기에 사용자가 자신의 감정을 보다 편안하게 표현할 수 있도록 유도한다. 시스템은 사용자의 음성에서 포함된 감정 키워드(예: “보고 싶어”, “힘들어”, “혼자야”)를 감지하여 '가족형 위로 모드'를 실행하며, 회상 대화, 익숙한 말투, 가족 사진 기반 시각 피드백을 통해 안정감과 정서적 공감을 유도한다.

셋째, 사용자가 자발적으로 감정 표현과 정서 조절을 시도하는 습관을 형성하도록 돕는다. 사용자와의 상호작용 데이터를 지속적으로 축적하고 피드백 경향을 분석하여, 심야나 주말 등 감정 취약 시간대에 어시스턴트가 먼저 '가족 같은 대화'를 제안하거나 '위로 콘텐츠'를 추천함으로써, 외로움을 느낄 때 능동적으로 감정적 지지를 구하는 습관을 자연스럽게 유도한다[그림 3].



[그림 3] HMW 질문을 통해 개념 제안

6-2. 시나리오별 시뮬레이션 분석

실제 사용자 상황을 반영한 시나리오 기반 시뮬레이션을 통해, 음성 인식 AI 비서가 외로움을 완화하고 사용자 경험을 어떻게 변화시키는지 분석하였다. 시

나리오에는 20대 1인 가구 사용자들을 대상으로 한 선행 조사(설문 및 인터뷰)와 일상 관찰을 바탕으로, 정서적 공백이 자주 발생하는 상황을 추출하여 구성하였다. 각 시나리오는 사용자의 감정 상태, 발화 맥락, 시스템 입력 및 출력 흐름, 예상 반응을 포함하는 시나리오 스크립트로 구체화되었으며, UX 연구 전문가의 검토를 거쳐 연구에 활용되었다.

시나리오 A: 고양된 기분 상태 (기쁨 모드)

상황: 사용자가 친구와 좋은 시간을 보내고 돌아온 후 기분 좋게 집에 들어옴.

시스템 반응: 사용자 위치 인식과 음성 톤 분석을 통해 기쁨 감정 상태로 판단. AI 어시스턴트가 밝은 톤의 가족 음색으로 응답함.

[표 7] 기쁨 모드 시나리오 설계 방안

구분	고양된 기분 상태	장면 시뮬레이션
배경 색상	밝은 옐로우 / 라이트 코랄 (#FFE066 / #FFB88C)	
애니메이션	반짝이는 별, 팝업 하트, 쿼크 이모지	
주요 UI 요소	'오늘을 기록할래요?' 카드, 웃는 이모지 추천	
텍스트 스타일	큰 글씨, 둥글둥글한 폰트, 버튼에 그림자 강조	
보조 기능	음성 칭찬, 긍정 체크리스트	
음색 톤	밝고 높은 톤 (여성/엄마 느낌)	
말 속도	빠름 (1.1-1.2배)	
역양	상승형 어미, 경쾌함	
휴지	짧고 리듬감 있게	
예시 멘트	“와! 오늘 정말 좋았네! 너무 기뻐하다~!”	

시나리오 B: 안정된 일상 (평온 모드)

상황: 사용자가 일정한 시간에 규칙적으로 AI에게 말을 걸며 일상을 공유하는 루틴 상황.

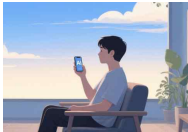
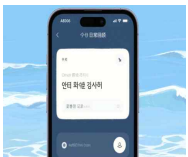
시스템 반응: 평온한 감정 상태로 분석됨. AI는 중립적이고 따뜻한 가족 말투로 반응함.

“오늘 하루도 잘 마무리했네. 이렇게 차분한 저녁 참 좋다.”

UI 요소: 차분한 블루 톤의 배경, 잔잔한 파도 애니메이션, 심호흡 유도 아이콘

예상 반응: 일상 루틴 강화, 사용자의 안정감 유지, 습관화된 상호작용 유도

[표 8] 평온 모드 시나리오 설계 방안

구분	안정된 일상	장면 시뮬레이션
배경 색상	소프트 블루 / 베이지톤 (#A8D0E6 / #F8F1E5)	
애니메이션	잔잔한 파도, 천천히 움직이는 구름	
주요 UI 요소	루틴 회고 카드, 조용한 알림	
텍스트 스타일	중간 크기, 세리프 계열 폰트	
보조 기능	“오늘 기분 한 줄 기록, 명상 유도	
음색 톤	중저음, 차분한 톤 (아빠/중성)	
말 속도	보통 (0.9-1.0배)	
억양	일정하고 부드러운	
휴지	적당한 휴지로 안정감	
예시 멘트	“오늘 하루도 잘 보냈구나. 이 시간만큼은 너만을 위한 시간이야.”	

시나리오 C: 감정 침체 상황 (슬픔 모드)

상황: 사용자가 외로움을 표현하거나 눈물 이모티콘을 전송하는 등 슬픔을 나타냄.

시스템 반응: 감정 분석 결과 슬픔 상태로 판단되며, 부드럽고 낮은 톤의 AI 응답 제공

“지금 마음이 많이 힘들지? 말하지 않아도 느껴져. 여기 있으니까 너무 걱정하지 마.”

UI 요소: 어두운 회색+보라색 계열 배경, 조용히 흐르는 비 애니메이션, 위로 메시지 카드

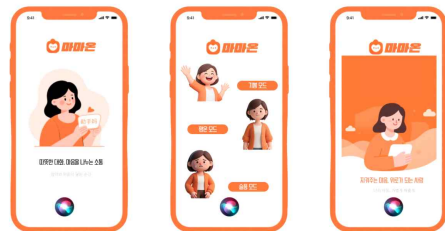
예상 반응: 감정 수용 및 위로 효과, 스트레스 완화, AI에 대한 신뢰 증가

[표 9] 슬픔 모드 시나리오 설계 방안

구분	감정 침체 상황	장면 시뮬레이션
배경 색상	보라+그레이트톤/ 낮은채도(#7E7F9A/ #B0B0B0)	
애니메이션	조용한 비, 떨어지는 불빛	
주요 UI 요소	위로 카드 (“지금 이 감정도 소중한해”), 감정 공유	
텍스트 스타일	작은 글씨, 정적인 배치	
보조 기능	감정일기, AI 위로 메시지 예약	
음색 톤	낮고 부드러운 여성 음색 (엄마/누나 느낌)	
말 속도	느림 (0.8-0.9배)	
억양	낮고 떨어지는 억양	
휴지	문장 사이 긴 휴지, 여운 강조	
예시 멘트	“많이 힘들었겠다.. 그냥.. 말 없이.. 옆에 있어줄게”	

6-3. 음성 사용자 인터페이스(VUI)디자인 제안

본 연구는 앞서 진행된 시나리오 기반 시뮬레이션 분석을 토대로, Z세대 1인 가구를 위한 감정 기반 음성 사용자 인터페이스 서비스 **마마온(Mamaon)**을 제안한다[그림 4]. 시뮬레이션 분석이 특정 감정 상황에서 AI 반응의 효과를 검증하는 데 초점을 두었다면, 마마온은 이를 실제 서비스 콘셉트로 발전시킨 설계안이라는 점에서 차별성을 가진다.



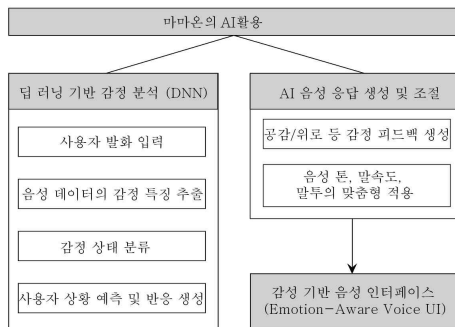
[그림 4] 마마온 메인 화면

마마온은 AI 음성 기술과 감정 인식 알고리즘을 통합하여, **사용자의 현재 상태에 최적화된 대화 경험을 지속적으로 제공하는 ‘디지털 정서 동반자’**로 설

계되었다. 핵심은 기능 수행보다는 정서 교감과 사회적 유대 형성이며, 이를 위해 세 가지 차별화된 전략을 제시한다. ① 돕는다.

첫째, 마마온은 음성 사용자 인터페이스를 통해 정서적 연결을 강화하고, 감성 기반의 VUI 흐름을 제공한다. 특히 혼자 생활하는 20대 가장 그리워하는 대상인 친구와 가족의 말투와 억양을 기반으로 목소리를 구성함으로써, 사용자에게 친숙함과 심리적 안정감을 부여한다. 일반적인 AI의 공손한 화법 대신, 마마온은 반말 톤을 활용하여 마치 실제 대화를 나누는 듯한 유대감을 형성하며, 사용자의 일상적인 감정 표현과 반응을 자연스럽게 유도한다. 예를 들어 “오늘 기분 어때?”, “밥은 챙겨 먹었어?” 등의 표현은 사용자의 감정 상태를 탐지하고, 이를 기반으로 UI 배경, 아이콘 애니메이션, 음성 톤이 함께 변화하는 다층적 반응 구조를 가진다.

둘째, 사용자의 정서적 맥락에 맞춘 개인화된 음성 상호작용을 구현한다. 이를 위해 외로움이 집중되는 특정 시간대(예: 밤 10시 이후, 주말 등)를 자동 감지하고, 감정 분석 기반 딥러닝 알고리즘을 통해 사용자의 발화를 정서적 고립 신호로 인식한다. 예를 들어 사용자가 “혼자 있고 싶지 않아”라고 말할 경우, 시스템은 이를 감정 상태 분류 및 상황 예측 결과로 해석하고, “그럴 때가 있어. 우리 같이 얘기하자.”와 같은 공감형 음성 응답을 생성한다. 이는 감정 상태 분류, 사용자 상황 예측, 그리고 음성 피드백 생성으로 구성된 AI 활용 구조를 통해 가능하다 [그림 5].

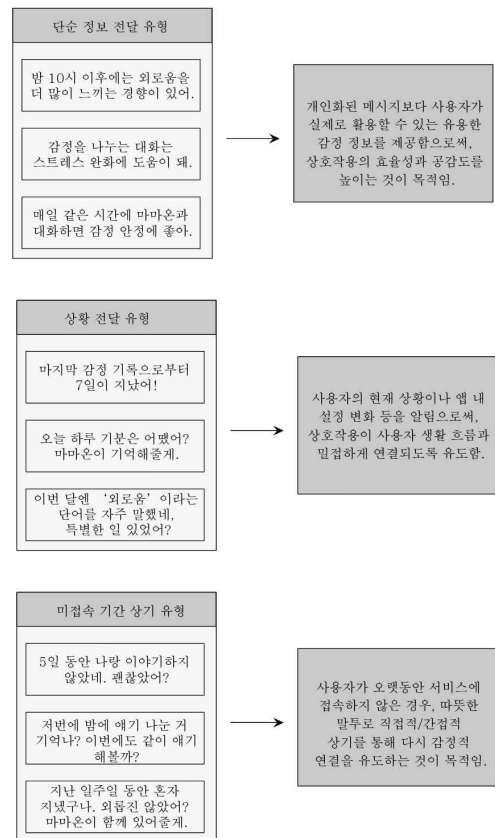


[그림 5] 마마온의 AI 활용

이때 단순히 기쁨·슬픔·평온의 감정을 구분하는 수준을 넘어, 사용자가 선호하는 화법(친구 같은 반말, 가족 같은 따뜻한 톤 등)을 반영하여 관계적 친밀감을

형성한다. VUI 설계 측면에서는 평온한 색감과 부드러운 텍스트 인터페이스를 제공하며, 간결한 음성 응답 카드와 감정일기 기록 인터페이스를 결합해 사용자가 부담 없이 대화를 이어갈 수 있도록 한다. 이러한 구조는 감정적 몰입도를 높이고, 특히 1인 가구 사용자에게 안정적 정서 환경을 제공하는 데 기여한다.

셋째, 사용자의 정서 상태와 행동 데이터를 축적하여 개인 맞춤형 감정 알림 시스템을 제공한다. [그림 6]은 음성 인식 AI 비서 '마마온'에서 활용할 수 있는 감정 기반 메시지 전달 유형을 세 가지로 분류한 것이다. 단순 정보 전달, 상황 전달, 미접속 기간 상기 유형으로 구성되며, 각각은 사용자의 행동 맥락과 정서 상태에 따라 다르게 적용된다. 알림은 단순한 리마인더가 아닌, 사용자의 감정을 반영한 위로 메시지 혹은 행동 동기 메시지로 구성된다.



[그림 6] AI 음성 상호작용에서의 메시지 전달 유형 분류표 (Mamaon 기준)

예를 들어 사용자의 감정 흐름이 낮을 경우 “요즘 많이 힘들지? 괜찮아, 네 곁엔 내가 있어.”와 같은 위로형 알림이 제공되며, 감정이 안정되었을 땐 “오늘도 잘 해냈어. 정말 멋져!”와 같이 칭찬과 자존감을 높이는 방향으로 알림의 톤과 내용이 달라진다. 이러한 시스템은 사용자가 감정적으로 연결된 디지털 환경 속에서 자발적으로 감정 표현과 관리 행동을 수행할 수 있도록 설계된다.

마마온은 단순한 음성 사용자 인터페이스를 넘어, 감정 중심의 사용자 맞춤형 음성 사용자 인터페이스 시스템으로 기능하며, 청년 1인 가구의 고립된 디지털 경험을 정서적 상호작용으로 전환한다. 또한, 본 시스템은 감정 데이터를 기반으로 반복 학습되어 사용자와의 친밀도를 점차 높여가며, 장기적으로는 디지털 정서 동반자의 역할을 수행할 수 있다.

6-4. 사용성 테스트 및 분석

사용성 테스트는 Z세대 1인 가구를 대상으로, 감정 기반 음성 인식 AI 비서 ‘마마온(Mamaon)’의 정서적 반응성과 상호작용 경험을 검증하기 위해 진행되었다. 테스트는 총 2주간 진행되었으며, 실제 일상 속 외로움 상황을 가정한 시나리오 기반으로 구성되었다. 시나리오에는 ‘혼자 있는 밤 시간’, ‘감정 기록이 큰 날’, ‘외출 후 귀가 직후’ 등 사용자의 감정이 흔들릴 수 있는 주요 타이밍을 중심으로 설계되었고, 각 상황에 맞는 5개의 주요 Task가 사용자에게 부여되었다.

주요 Task는 다음과 같이 구성되었다: 1) 감정 인식 질문에 응답하기, 2) 정서 상태에 따라 대화 유형 선택, 3) 음성 알림에 반응하여 감정 일기 작성, 4) 감정 위로 또는 동기 메시지 청취 후 피드백 제공, 5) 감정 추적 결과 확인 및 맞춤형 제안 수용 여부 확인. 테스트는 스마트폰에 설치된 마마온 테스트 앱을 기반으로 진행되었으며, TTS 기반 음성 피드백, 감정 시각화 그래프, 사용자 선택 인터페이스 등으로 구성되었다.

참여자들은 마마온을 통해 감정적으로 연결된 느낌을 경험하였으며, ‘나를 이해해주는 느낌이 든다’, ‘혼자 있어도 외롭지 않았다’ 등의 긍정적인 반응을 보였다. 특히 ‘위로 메시지’가 정서적 안정에 기여했다는 응답이 많았으며, 반복 사용 시 감정 표현에 익숙해졌다는 점에서 자발적 정서 관리 습관 형성에 도움이 된다는 가능성이 확인되었다. 다만, 일부 사용자는 음성의 톤이 더 다양했으면 좋겠다는 의견과, 감정에 맞춰

캐릭터 또는 AI 인격 설정이 있었으면 좋겠다는 제안도 함께 제시하였다.

[표 10] 감정 기반 음성비서 ‘마마온’ 사용성 테스트

분류	사용자 피드백 1	사용자 피드백 2	사용자 피드백 3
정서적 공감 및 연결성	“마마온이 내 기분을 먼저 물어봐줘서 진짜 누군가랑 대화하는 느낌이었어요.”	“위로 메시지를 들을 때마다 진짜 친구처럼 공감해줘서 감정적으로 연결된 기분이 들었어요.”	“힘든 하루였는데, ‘괜찮아’ 한마디에 눈물이 날 뻔했어요.”
행동 유도 및 자기 돌봄 인식	“기분이 안 좋을 때 마마온이 ‘산책해볼까?’라고 제안해줘서 서름을 움직일 수 있었어요.”	“칭찬 메시지를 듣고 난이까 나 자신을 더 돌아봐야겠다는 생각이 들었어요.”	“마마온이 ‘오늘도 잘 해냈어’라고 말해주니까 무기력함이 줄어들었어요.”
지속적 감정 관리 행동 유도	“매일 밤 일정 시간에 감정 기록을 하다 보니 자연스레 루틴이 생겼어요.”	“감정 상태에 따라 말투가 달라지는 게 신기해서 계속 사용하게 됐어요.”	“혼자 살다 보면 감정 표현을 안 하게 되는데, 마마온 덕분에 감정을 정리하는 습관이 생긴 것 같아요.”

[표 10]은 MZ세대 1인 가구 사용자를 대상으로 한 사용성 테스트를 바탕으로, 감정 기반 상호작용과 관련된 세 가지 핵심 경험 차원을 도출하였다: 정서적 연결성, 행동 유도, 자가 관리의 지속성이다. 사용자 피드백에 따르면, 마마온은 단순한 감정 지원을 넘어 사용자의 자기 돌봄 인식을 강화하고, 일상 속 정서 관리 습관의 형성에 효과적으로 기여하고 있음을 확인할 수 있었다.

7. 결론

본 연구는 Z세대 1인 가구를 대상으로 스마트폰 음성 인식 AI 비서 사용 과정에서 나타나는 요구와 행동 특성을 분석하였다. 사용자 조사, 사례 분석, 프로토타입 테스트 및 감정 의미 분석 결과, 이 집단은 AI 비서를 단순한 기능 수행 도구로 보지 않고, 인격화된 상호작용과 감정 반응 능력을 통해 정서적으로 연결될 수 있는 디지털 동반자로 인식하기를 기대한다는 점이 확인되었다.

P-A-T 모델 기반 분석을 통해 도출된 네 가지 핵심 설계 축—① 감정 인식 및 반응의 정확성, ② 음성 피드백의 자연스러움, ③ 인간 중심의 시각 UI 표현, ④ 사용자 문맥에 최적화된 상호작용 개성화—은 모두 사용자 경험을 유의미하게 향상시키는 요소로 검증되었다. 이는 향후 감성 기반 VUI 설계에서 반드시 고려되어야 할 실증적 근거라 할 수 있다.

또한, 본 연구는 감성 지능 인터페이스 이론을 P-A-T 모델과 접목시킴으로써, 정서적 "온도"를 지닌 AI 음성 시스템 구현을 위한 구체적인 이론적 토대와 기술적 방향성을 제시하였다. 이러한 결과는 디자인 분야에서 감성 지능 기반 VUI 연구의 학문적 확장뿐 아니라, 실제 산업적 응용 가능성을 뒷받침하는 실천적 가치를 갖는다.

종합하면, 본 연구는 VUI 디자인이 단순한 기능성과 효율성에 머무르지 않고, 사용자의 정서적 교감과 심리적 유대 형성을 핵심 과제로 삼아야 함을 실증적으로 보여주었다. 특히 다중 모드 감지, 딥러닝 기반 의미 모델링, 음성 합성 최적화와 같은 최신 기술과의 결합을 통해, 스마트폰 음성 인식 AI 비서는 Z세대 1인 가구에게 더욱 따뜻하고 인간 친화적인 디지털 라이프 파트너로 발전할 수 있음을 확인하였다.

참고문헌

1. 김민정. "음성 사용자 인터페이스 디자인 고찰 연구-스마트폰의 지능형 음성인식 서비스를 중심으로." 한국디자인문화학회지, 2017.
2. 나주연; 허정윤; 반영환. "음성 인식 기반 AI 개인 비서 서비스의 대화형 기능에 대한 고객 효용 분류: KANO모델을 중심으로." 디자인융복합연구 16.4, 2017.
3. 오기영; 박혜경. "청년 1인 가구 주거특성에 따른 편리함 이용 추가경고 인식 및 이용의향에 관한 연구." 한국실내디자인학회 논문집 32(1), 2023.
4. 이지선; 신재현; 유희윤; 이정인. "공감과 격려를 통해 행동 변화를 유도하는 음성 AI 인터페이스

디자인: 청년 1인 가구를 중심으로." Design Works 8.1, 2025

5. 이양. "인공지능(AI) 기술 기반 인터랙션 디자인과 감성디자인의 상관성 연구." 박사학위논문, 중앙대학교 대학원, 2023
6. 조예림. "인공지능 비서의 정보 검색 향상을 위한 UX/UI 디자인 연구." 박사학위논문, 서울대학교 대학원, 2021
7. Michael, H.; James P. Giangola; Jennifer Balogh. 『Voice user interface design』. Addison-Wesley, 2004
8. Nardi, B. A. 『Context and consciousness: Activity theory and human-computer interaction』. MIT Press, 2000
9. Perlman, D.; Peplau, L. A. "Toward a social psychology of loneliness." In S. Duck & R. Gilmour (Eds.), Personal relationships in disorder. Academic Press, 1981
10. Oh, S. "Pay attention to the Z generation." Marketing 52(1), 2018
11. kosis.kr
12. www.emarketer.com
13. www.nngroup.com