

음성 사용자 인터페이스 내 피드백 유형이 사용자의 주관적 반응에 미치는 영향

이다솜* · 이상원**

성균관대학교

Effects of Feedback Types on Users' Subjective Responses in a Voice User Interface

Dasom Lee* · Sangwon Lee**

Sungkyunkwan University

E-mail : ssom216@naver.com*, upcircle@skku.edu**

요 약

본 연구에서는 음성 사용자 인터페이스가 적용된 시스템 사용 중 오류가 발생했을 때 제공되는 피드백이 시스템에 대한 사용자의 평가에 어떠한 영향을 줄 것인지에 대해 알아보았다. 피드백 유형(확인 피드백, 정교화 피드백)과 오류 유형(인식 실패, 수행 실패)을 독립 변수로 하고 사용자의 주관적 평가와 피드백 수용성, 향후 이용 의도를 종속 변수로 하는 요인설계 실험연구를 진행하였다. 이원 다변량 분산분석을 한 결과, 피드백 유형만 사용자의 주관적 평가와 피드백 수용성, 향후 이용 의도에 영향을 미치는 것으로 나타났다. 또한, 피드백 수용성에 대해 피드백 유형과 오류 유형의 상호작용 효과가 발견되었다. 따라서 VUI 설계 시, 오류 정보를 자세히 알려주는 피드백을 제공하는 것을 고려해 볼 수 있을 것이다.

ABSTRACT

This study aimed to demonstrate the effect of feedback type on users' subjective responses in a voice user interface. Feedback type is classified depend on information characteristic it involves; verification feedback and elaboration feedback. Error type is categorized as recognition error and performance error. Users' subjective assessment about system, feedback acceptance, and intention to use were measured as dependent variables. The results of experiment showed that feedback type has impacts on the subjective assessment(likeability, habitability, system response accuracy) of VUI, feedback acceptance, and intention to use. the results also demonstrated an interaction effect of feedback type and error type on the feedback acceptance. It leads to the conclusion that VUI should be designed with the elaboration feedback about error situation.

키워드

음성 사용자 인터페이스, 피드백 유형, 오류 유형

I. 서론

음성 사용자 인터페이스(VUI; Voice User Interface)는 인간의 기본적 커뮤니케이션 수단인 음성언어를 이용하며 가장 직관적 인터페이스라는 평가를 받고 있다. 특히 인공지능 기술이 발달하면서 지능형 에이전트와 효과적으로 인터랙션 할 수 있는 인터페이스로 손꼽히고 있고, VUI는 영역을 빠르게 확장하고 있다. 그러나 VUI의 다양한 장점에도 불구하고 음성 기능을 매일같이

유용하게 쓰고 있는 사람은 드물며, 일반적 정보 검색이나 네비게이션 같은 단순 기능을 중심으로 사용한다고 한다[1].

사람 간 소통에서도 오해가 발생하듯, VUI를 통한 소통에도 음성인식 실패, 음성언어 이해 실패, 대화처리 실패와 같은 다양한 오류가 발생한다. 그런데 이와 같은 오류가 발생했을 때, 사용자에게 전달되는 오류에 대한 메시지는 알아듣지 못했다거나 무슨 의미인지 모르겠다고 하는 것에 그치는 낮은 수준의 피드백만 포함하고 있다. 무

엇을 못 알아들은 것인지, 왜 명령을 수행할 수 없는지 자세한 원인을 설명하지 않은 채 무엇인가 불가능하다는 결과만 알리는 것은 시스템 사용률을 낮추는 결과를 초래할 수 있다.

음성을 통한 소통에서 오류 발생은 불가피하므로, 최선의 대화는 상황에 따라 알맞은 추가 대화를 유도해 오류 상황을 수정하는 방법일 것이다. Yankelovich는 오류 상황에 대한 적절한 안내멘트를 제공하는 것을 강조하며, 특히 오류가 여러 번 발생할 때는 각 단계마다 다른 안내멘트를 주어 시스템이 오류상황에 대해 적극적으로 대응하고 있다는 인상을 주어야 사용자가 시스템을 신뢰하게 된다고 하였다[2]. 또한 Stephen은 오류 상황을 해결하기 위해서는 단순히 사용자에게 동일한 문장의 재발화를 요구하는 것 보다 오류 발생 원인과 관련된 정보를 포함해 응답하는 것이 효과적이라고 하였다[3].

하지만 현재의 설계는 다양한 오류 상황이 있음에도 불구하고 단조로운 응답 패턴을 보이고 있다. 지금까지의 VUI 관련 연구는 음성 인식을 향상 방안이나 음성 합성에 대한 것들이 주를 이루었고, VUI의 설계와 관련한 연구에서도 오류 상황에 대한 피드백의 중요성을 강조할 뿐, 그 내용에 대해 직접적으로 다루는 연구는 부족하다. 소통 오류 상황에 대한 안정적이고 효과적인 음성 답변 설계는 시스템의 신뢰도 및 사용성에 중요한 영향을 미친다[4]. VUI를 활용한 제품이 쏟아져 나오는 상황에서, 오류 상황에 대한 피드백에 대해 연구해 봄으로써 사용성을 제고할 또 다른 방안을 찾을 수 있을 것이다.

따라서 본 연구에서는 상황에 따라 사용자가 선호하는 피드백에 차이가 있을지 알아보기 위해, 오류 상황을 분류하고 구체성을 다르게 한 두 가지 종류의 피드백을 제공하는 실험을 진행하였다. 그리고 VUI에 대한 사용자의 주관적 평가와 피드백 수용성, 향후 이용 의도를 측정하여 그 효과를 검증하였다.

II. 이론적 배경

1. 피드백 유형

Butler와 Winne는, 피드백은 정보이며 학습자는 그것을 활용해 기존에 가지고 있던 정보를 확인, 추가하여 재구조화 한다고 하였다[5]. 피드백의 유형에 대해서는 다양한 관점이 존재하는데, 특히 내용의 구체성에 대해서는 확인(Verification)과 정교화(Elaboration)로 나눈다[6]. 확인은 학습자의 수행에 대해 맞거나 틀리다는 확인을 해 주는 것을 의미한다. 정교화는 수행에 대한 확인과 더불어 수행해야 할 과제의 주제에 대해 언급하거나 설명 혹은 예시를 제공하는 것이다[7]. 이것을 바탕으로 본 연구에서는 음성으로 전달되는 피드백을 확인 피드백과 정교화 피드백으로 나눈다. 확인 피드백은 현재 상태에 대한 확인과 재발화 요

구로 구성되며, 정교화 피드백은 현재 상태에 대한 확인, 원인 진단, 대안 제시, 재발화 요구를 담고 있다. 사용자가 오류 상황에 대한 피드백을 받은 뒤 대화를 이어나가야 하기 때문에 확인 피드백과 정교화 피드백 모두 공통적으로 재발화 요구를 포함한다.

교육학에서 주로 다루지는 피드백의 효과는 학습자에게 동기를 부여한다는 점이다. 피드백의 구체성은 학습자의 인지적 부하를 감소시켜 생소하거나 난이도가 높은 과제에 대해 인지적 압도감을 감소시키며, 부적절한 과제 전략, 절차상의 오류, 오개념을 수정하기 위한 정보제공의 역할을 한다고 밝혀져 있다[7]. 그런데 피드백 종류의 효과 검증에 대한 연구는 서면으로 이루어진 연구가 다수이고 컴퓨터 환경에서도 디스플레이를 포함한 연구가 대부분이다. 오류가 빈번히 발생할 수 있는 VUI 환경에서 오류에 대한 피드백을 구체성에 따라 검증할 필요가 있다.

2. VUI 환경에서 오류 상황

VUI 환경에서 발생할 수 있는 오류는 사용자의 음성이 아예 인식되지 않은 경우부터 다른 어휘로 잘못 인식하는 경우까지 다양하게 발생할 수 있다. 이에 대해 연구자들은 여러 가지 관점을 가지고 있는데, Schnelle-Walka는 음성인식으로 인해 VUI 환경에서 소통 오류가 발생한다고 하였다[8]. 사용자의 명령이 입력되었을 경우에는 음성 인식 성공, 음성 오인식, 음성 입력 거부가 발생할 수 있고, 명령이 입력되지 않은 경우에는 입력 거부 성공과 입력 거부 실패가 발생할 수 있다고 시스템과 사용자 간 인터랙션을 분류하였다.

Duff는 VUI 환경에서 발생할 수 있는 오류를 단계별로 분류하였다. 레벨 0은 사용자의 음성 입력 실패, 레벨 1은 음성 인식 실패, 레벨 2는 사용자의 음성 입력 해석 불가, 레벨 3은 사용자의 무의미한 음성 입력, 레벨 4는 사용자의 음성 입력에 대한 결과정보 부족, 레벨 5는 사용자의 잘못된 음성 입력, 레벨 6은 서버와 통신 실패, 레벨 7은 기타이다[9].

한편, 김진우는 사람과 디지털 시스템의 인터랙션에 대해 표명과 변환이 사용자로부터 시스템으로 입력하는 부분에서 이루어진다고 하였다. 또한 표현과 평가가 시스템이 사용자에게 결과를 출력하는 부분에서 이루어지며 네 단계가 순환적으로 발생한다고 하였다[10].

이와 같은 기존 연구들을 바탕으로 본 연구에서 정의하고자 하는 오류 유형은 인식 실패와 수행 실패이다. 인식 실패란 사용자가 VUI에 발화하는 단계에서 발생할 수 있는 오류이다. 사용자의 음성이 아예 인식되지 않은 경우, 사용자의 음성이 인식되었으나 유의미하지 않은 소리만으로 인식된 경우, 유의미한 음성이 인식되었지만 인식된 단어가 개념상 서로 관련 있는 것으로 해석되지 않는 경우가 여기에 포함된다. 수행 실패는 사용자의 음성을 성공적으로 인식하고 사용자의 의도를 이

해하였으나, 시스템 상의 문제로 과업을 수행할 수 없는 상태를 의미한다. 두 가지 유형은 각각 VUI의 내부와 외부에서 기인한 것으로 오류 발생 이후 대화가 다른 양상으로 전개된다. 인식 실패의 경우 바로 전 대화에 대한 재발화를 요구하지만, 수행 실패의 경우 요청한 과업을 수행할 수 없으니 다른 명령을 내려달라는 흐름으로 대화가 진행된다.

본 연구에서는 VUI의 오류 유형을 인식 실패와 수행 실패로 나누고 구체성 수준에 따라 피드백 유형을 확인 피드백과 정교화 피드백으로 나누어 제공하는 실험을 통해, 피드백 유형과 오류 유형이 서로 유의미한 연관성을 나타내는지 탐구하고자 한다. 연구 가설은 아래와 같다.

가설 1. VUI 환경에서 피드백 유형(확인 피드백, 정교화 피드백)에 따라 사용자의 주관적 평가와 피드백 수용성, 향후 이용 의도에 차이가 있을 것이다.

가설 2. VUI 환경에서 오류 유형(인식 실패, 수행 실패)에 따라 사용자의 주관적 평가와 피드백 수용성, 향후 이용 의도에 차이가 있을 것이다.

III. 연구방법

1. 실험 설계와 방법

독립변인을 피드백 유형(확인 피드백, 정교화 피드백)과 오류 유형(인식 실패, 수행 실패)로 하는 2x2 요인설계 실험연구를 진행하였다. 오류 유형에 대해서는 피험자 간 설계, 피드백 유형에 대해서는 피험자 내 설계를 하였다. 두 가지 피드백 유형에 대해 순서효과를 최소화 하기위해 역균형화를 하였고 각각 다른 과업을 부여하였다.

인공지능 스피커 혹은 스마트폰 음성인식 비서 기능을 활용해 본 경험이 있는 20~25세의 대학생 8명(남성 7명, 여성 1명)이 실험에 참가하였다. 인공지능 스피커를 사용하는 일상 환경으로 조성한 실험실에서 Wizard of Oz 방식으로 블루투스 스피커를 통해 실험 조건을 조작하였다.

선행 연구의 시나리오와 SK NUGU, 네이버 웨이브의 사용 시나리오를 참고하여 음악 듣기 과업과 쇼핑하기 과업에 대한 시나리오를 작성하였다.

피험자는 인식 실패 오류와 수행 실패 오류 중 한 가지의 오류 상황에 무선으로 할당된다. 이어 피험자는 두 가지 피드백 유형에 노출되는데, 음악을 듣거나 쇼핑을 하는 상황을 설명하는 메시지를 담은 카드로 과업을 부여받는다. 각 조건에서 인공지능 스피커와 대화를 나눈 후 설문지를 작성한다.

2. 종속 변인의 정의와 측정

사용자의 주관적 평가에 미치는 영향을 검증하기 위해 likeability, habitability, system response accuracy의 세 가지 개념을 변인으로 채택하여 측정하였다. VUI에 대한 평가인 SASSI; The Subjective Assessment of Speech System Interfaces에서 추출한 요인으로, 인공지능 스피커에 대한 전반적 평가와, 함께 나눈 대화에 대해 평가할 수 있는 척도들로 구성하였다.

피드백의 수용은 피드백이 자신의 수행을 정확히 묘사하고 있다는 개인의 신념에 따라 결정된다 [10]. 이 연구에서는 피드백 내용이 도움이 된다고 느낀 정도를 측정하였다. 오류 상황에 대해 원인을 진단하고 대안을 제시하는 인공지능 스피커의 피드백이 긍정적 영향을 미칠 것이라고 예측된다.

감정적 변인들과 더불어 피드백의 유형과 오류 유형이 인공지능 스피커를 계속 사용하는 데 영향을 주는지 검증하기 위해 향후 이용 의도를 측정하였다. 모든 변인은 리커트 5점 척도(1=매우 아니다, 5=매우 그렇다)를 사용하였다.

IV. 연구결과

피드백 유형과 오류 유형에 대한 요인설계 실험 연구를 수행하였고 사용자의 주관적 평가와 피드백 수용, 지속 이용 의도를 측정하였다. 이원 다변량 분산분석(two-way MANOVA)을 통해 분석하였고 결과는 표1과 같다.

피드백 유형에 대해서는 다섯 가지 종속 변수 모두 유의한 차이를 보였다. 또한 다섯 가지 종속 변수 모두 확인 피드백보다 정교화 피드백에서 높은 수치를 나타내었다. 효과크기는 피드백 수용

	피드백 유형			오류 유형		
	확인 Mean(S.D)	정교화 Mean(S.D)	F(η^2)	인식실패 Mean(S.D)	수행실패 Mean(S.D)	F(η^2)
Likeability	3.38(0.81)	4.75(0.35)	24.94*** (0.68)	4.25(0.79)	3.88(1.07)	1.86(0.13)
Habitability	2.92(0.90)	4.25(0.53)	15.06* (0.56)	3.83(0.71)	3.33(1.21)	2.12(0.15)
System response accuracy	3.50(0.82)	4.38(0.45)	8.43* (0.41)	4.04(0.52)	3.83(1.01)	0.48(0.04)
Feedback acceptance	2.58(0.71)	4.54(0.31)	76.17*** (0.86)	3.67(0.87)	3.46(1.41)	0.87(0.07)
Intention to use	3.17(0.94)	4.63(0.45)	18.49*** (0.61)	4.08(0.85)	3.71(1.23)	1.22(0.09)

표1. two-way MANOVA summary

성이 가장 높게 관찰되었는데($\eta^2=0.79$), 다섯 가지 종속 변수 중 피드백 유형의 영향을 가장 많이 받는다고 해석할 수 있다.

이에 반해 오류 유형에 대해서는 다섯 가지 종속 변수 모두 유의한 차이가 없는 것으로 나타났다. 실험 뒤 이루어진 피험자 인터뷰를 통해 인식 실패와 수행 실패를 표현하는 실험 조건이 다소 약했음을 알 수 있었다. 또한 피험자 수(8명)가 적었던 점도 영향을 준 원인으로 짐작할 수 있다.

오류 유형과 피드백 유형의 상호작용 효과는 피드백 수용성에서 유의미한 결과를 얻을 수 있었다($F=7.76, p=0.02$). 그림1과 같이, 인식 실패 오류 상황보다 수행 실패 오류 상황일 때 피드백 유형이 피드백 수용성에 미치는 영향이 크다고 할 수 있다.

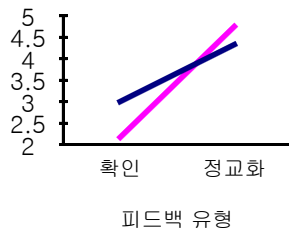


그림1. 피드백 수용성에 대한 상호작용 효과

V. 결론

본 연구에서는 VUI 환경에서 오류 유형과 피드백 유형이 시스템에 대한 사용자의 주관적 평가에 미치는 영향을 알아보려고 하였다. 오류 유형과 무관하게 정교화 피드백을 제공하는 것이 사용자의 주관적 평가에 긍정적인 영향을 주는 것으로 밝혀졌다. 그러므로 VUI 설계 시, 오류가 발생하여 피드백을 줄 때 자세한 오류정보를 주는 것에 대해 고려해 볼 수 있을 것이다.

오류 유형을 나눈 것에 대해 효과를 관찰할 수 없었다는 점은 아쉬운 점이며, 향후 연구에서 오류 상황에 대한 차이를 느낄 수 있도록 더욱 분명한 처치를 하여 진행할 것이다. 또한 음악 듣기와 쇼핑하기로 주어진 과업이 다소 한정된 상황이었다는 점도 본 연구의 한계점으로 볼 수 있다. 실제로 인공지능 스피커나 스마트폰 음성인식 비서를 활용할 때는 특정 목적을 위한 대화뿐만 아니라 일상적 대화도 주고받으므로 이런 상황도 반영할 수 있는 분석을 할 필요가 있다.

본 연구는 교육학에서 주로 다루지던 피드백 유형에 따른 효과를 VUI 환경에 적용하여 검증하였다. 또한 기존 연구에서는 오류 상황에 대한 피드백의 중요성을 강조하는 것에만 그쳤던 것에 비해 피드백의 내용에 대해 직접적으로 다루었다는 점은 본 연구의 의의로 볼 수 있다.

NUI; Natural User Interface의 하나로 VUI가 각광 받고 관련 기술이 빠르게 개발되고 있으므로, 사용성도 발전의 여지가 많다. 본 연구를 바탕으로 향후 인공지능 스피커를 비롯한 VUI 환경에서 대화 패턴 설계와 사용성 향상에 대한 연구에 기여할 수 있을 것이라 기대한다.

사사의 글

이 성과는 2017년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2017R1C 1B 1003650).

참고문헌

- [1] 최재형, 조광수, 사람들은 왜 음성비서를 더 많이 이용하지 않을까, Tech M, 2017. 09. 14
- [2] Yankelovich, Nicole, Gina-Anne Levow, and Matt Marx, Designing SpeechActs: Issues in speech user interfaces, Proceedings of the SIGCHI conference on Human factors in computing systems, ACM Press/Addison-Wesley Publishing Co., 1995
- [3] Choularton, Stephen, Handling speech recognition errors in spoken language dialogue systems, ACL-04 Students Workshop, 2004
- [4] 천재민, et al., 대화형 음성 인터페이스의 인터랙션 오류 상황에 대한 사용자 발화 패턴 분석, 대한인간공학회 학술대회 논문집, 81-84, 2006
- [5] Butler, Deborah L., and Philip H. Winne., Feedback and self-regulated learning: A theoretical synthesis, Review of educational research, 65.3, 245-281, 1995
- [6] Shute, Valerie J., Focus on formative feedback, Review of educational research, 78.1, 153-189, 2008
- [7] 감민영, 피드백 유형과 과제지속의 관계에 대한 피드백 수용성의 조절효과, 교육과정평가연구, 20(1), 143-163, 2017
- [8] Schnelle-Walka, Dirk., A pattern language for error management in voice user interfaces., proceedings of the 15th European Conference on Pattern Languages of Programs, ACM, 2010
- [9] LuperFoy, Susann, et al., An architecture for dialogue management, context tracking, and pragmatic adaptation in spoken dialogue systems, Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics, Volume 2, 1998
- [10] 김진우. Human Computer Interaction 개론, 서울: 안그라픽스, 2005