

스마트 스피커와 요리하기: 음성기반 레시피 제공 서비스의 사용자 경험

정구민^o,

정혜선*

한림대학교 디지털콘텐츠융합스쿨

심리학과 & 한림응용심리연구소

wjdrnals16@naver.com^o, heis@hallym.ac.kr

Cooking with a smart speaker: User experience of cooking with a voice-only recipe service

Gumin Jung^o,

Heisawn Jeong*

Hallym University

요약

음성 기반 컴퓨팅 기술과 서비스의 수준이 높아지면서, 이러한 음성 기반 기술을 활용할 분야도 넓어지고 있다. 그 중 스마트 스피커의 보급이 증가하면서 다양한 음성기반 서비스가 출시되고 있고, 1인 가구 및 요리 방송이 증가하면서 스마트 스피커를 사용한 레시피 서비스 제공이 시작되었다. 전통적으로 레시피 서비스는 시각적인 정보제시를 주로 사용하였으나, 사용자의 경험 증진을 위해서는 음성만으로도 동일한 정보를 적절하게 제공해야 한다. 스마트 스피커가 가진 장점은 '핸즈프리'가 가능하다는 것인데, 이는 요리 상황에서 효율적으로 쓰일 수 있으며, 시각을 사용하지 못하는 사용자에게 도움이 될 수 있다. 본 연구에서는 음성 기반 레시피 서비스가 사용자의 요리 활동을 얼마나 효과적으로 지원하는지 살펴보았다. 현장연구로 요리 숙련도별로 새로운 레시피를 조리하도록 하였고, 소리내어 생각하기(think-aloud), 관찰 및 인터뷰를 실시했다. 그 결과 음성 기반 레시피 서비스는 사용자의 손을 자유롭게 하여 조리 활동을 중단하지 않으면서 레시피를 확인할 수 있도록 하였으나, 조리지시가 한꺼번에 주어져서 사용자가 기억 부담을 하는 것이 관찰되었다. 시각적인 정보의 부재는 심각한 단점으로 작용하지는 않았으나, 조리의 효율성과 완성도에 영향을 주었다. 본 연구의 결과는 HCI의 발전과 음성을 사용한 레시피 서비스 사용자 경험 연구에 참고 될 수 있을 것으로 기대된다.

Abstract

This study examined how users use smart speakers in cooking situations. Skilled and unskilled participants cooked a new recipe while following voice instructions delivered by a smart speaker. The results from video recordings of their cooking, think-aloud protocols, and interviews showed that the smart speakers freed users' hands, allowing them to cook while checking recipes. The lack of visual information did not pose a serious challenge to the cooking task, but impacted cooking quality. The implications for VUI-based recipe service designs are discussed.

키워드: 스마트 스피커, 레시피 서비스, 소리내어 생각하기, 음성 사용자 인터페이스, 핸드프리

Keywords: Smart Speaker, Recipe Service, Think-aloud, Voice User Interface, Hands-free

자연어처리 기술의 발전으로 최근 음성으로 조작 가능한 VUI(Voice User Interface) 사용이 증가하고 있다. 대표적으

1. 서론

* corresponding author: Heisawn Jeong/Hallym University(heis@hallym.ac.kr)

*corresponding author: Heisawn Jeong/Hallym University(heis@hallym.ac.kr)

Received : 2021.08.03. / Review completed : 1st 2021.09.06. / Accepted : 2021.11.16.

DOI : 10.15701/kcgs.2021.27.5.13

ISSN : 1975-7883(Print)/2383-529X(Online)

로 스마트폰의 AI비서, 자동차의 음성 AI, 스마트 키친, 스마트 스피커 등이 있다. VUI 기기 중 하나인 스마트 스피커는 사용자들이 일상에서 접할 수 있는 기기로 정보 검색, 음악 듣기에서부터 스마트 가전의 제어까지 다양한 기능을 제공하고 있다. 스마트 스피커의 보급이 증가하면서 음성으로 지원되는 다양한 서비스가 생겨나고 있는데, 그중 하나로 레시피 서비스를 들 수 있다. 최근 1인 가구 및 음식 배송의 증가와 함께 다양한 요리를 직접 하고자 하는 사용자들이 증가하고 있다. 스마트 스피커의 레시피 서비스는 요리책이나 스마트폰 내 앱과 달리 음성을 기반으로 한다는 점에서는 다른 레시피 제공 서비스와 구분된다. 본 연구에서는 스마트 스피커의 레시피 서비스가 사용자들의 조리 활동을 얼마나 효과적으로 안내하는지를 소리내어 생각하기 기법, 관찰, 인터뷰를 통해 살펴보고, 음성기반 레시피 서비스 설계에서 고려해야 할 점을 탐색하였다.

2. 선행 연구

2.1 VUI와 레시피 서비스

음성 인식 기술과 인공지능의 발전에 따라 VUI의 사용이 증가하고 있다[1, 2]. VUI는 사용자가 음성으로 기기를 제어할 수 있도록 한다. 또한, 디스플레이를 사용하기 어려운 상황(예, 어둠 속) 또는 다른 활동을 수행 중이어서 기기 조작이 어려운 상황(예, 요리, 운전)에서 유용하다. VUI는 인간의 자연 언어로 상호작용한다는 점에서 다른 인터페이스보다 사용이 직관적이다. 디지털 기기에 적응이 어려운 사람들은 물론 시각 장애인들을 위해서도 효과적인 인터페이스가 될 수 있다[3, 4].

‘Designing Voice User Interface’의 저자 Cathy Pearl은 VUI가 가진 이점을 다음과 같이 요약하고 있다[5]. 첫째, 명령어를 말하는 속도가 타자하는 속도보다 빠르기 때문에 사용자는 VUI를 사용할 때 더욱 빠르게 기기와 상호작용할 수 있다. 둘째, 손이 자유로워지는 상태를 헨즈프리라고 하는데, 이는 기기와 상호작용하면서 동시에 다른 작업을 할 수 있는 상태이다. 셋째, 사람의 언어로 상호작용하기 때문에 디지털 기기에 익숙하지 않은 사람들도 직관적으로 상호작용할 수 있다. 넷째, 음성의 톤(tone), 강약 등과 같은 준언어(paralinguistic) 단서를 사용하여 더 효과적으로 의사소통할 수 있다.

Pearl은 VUI가 유용한 상황의 하나로 운전이나 요리처럼 사용자가 손을 사용해야 하는 상황을 들었다[5]. 운전을 하면서 기기 조작을 위해 손을 사용할 경우 사고의 위험을 증가시킬 수 있다[6]. 요리를 할 때에도 손을 사용하는 작업이 많기 때문에 기기를 조작해야 하는 경우 조리 활동의 지연과 중단을 발생시킨다. 김보민 등은 혼자 사는 싱글들이 스마트폰이나 종이로 레시피를 보며 요리할 때 발생하는 문제점을 살펴보았다. 조리 순서나 계량의 어려움과 함께 스마

트폰과 종이를 손에 닿기 꺼려 요리하는 순서를 놓쳐서 이 때문에 흐름이 끊겨 시간이 지체되는 경우가 관찰되었다[7]. 레시피는 음식을 만드는 과정을 안내하는 역할을 한다. 처음 요리하는 음식은 물론 이전에 해본 음식이라도 필요한 재료나 조리 순서를 모두 기억하기는 쉽지 않다. 레시피는 기억의 보조 장치 역할을 하는데, 사람들은 조리하면서 필요한 재료, 조리 순서 등을 확인하기 위해 레시피를 참조한다. 효과적인 레시피 서비스는 이 확인 작업이 매끄럽게 일어나도록, 즉, 현재 진행되는 조리 활동을 방해하지 않으면서 일어나도록 해야 한다. 이런 점에서 VUI기기는 레시피 서비스를 위한 효과적인 방안이 될 수 있다. 종이나 디스플레이 대신 음성으로 레시피가 제시되면 진행 중인 활동을 중단하거나 시선을 이동시키지 않아도 되기 때문에 조리 활동이 더욱 매끄럽게 진행될 수 있다.

2.2 기억 부담과 시각 정보의 부재

최근 1인 가구의 증가와 요리에 관한 관심의 증가로 요리 방송이 대세로 자리 잡으면서 방송에 나온 요리를 시도하는 인구가 늘어나고 있다. 과거에는 요리책을 보며 요리했지만, 최근에는 스마트 스피커를 포함한 다양한 디지털 기기와 앱을 이용하여 레시피를 검색하고 조리에 참고하는 것이 일상화되었다[8]. 스마트 스피커 레시피 서비스는 음성으로 레시피 정보를 제공한다는 점에서 다른 레시피 제공 서비스와 구분된다. 이는 사용자의 손을 자유롭게 하여 조리 활동을 보다 효과적으로 지원할 수 있음을 의미한다.

첫째, 레시피는 상당히 많은 정보를 담고 있다. 레시피는 통상 재료와 조리 단계로 구분되는데, 재료는 사용되는 재료의 종류 및 크기, 질량 등을 명세하고, 조리 단계는 이들 재료를 변형하는 과정을 기술하며, 정해진 방식으로 재료를 손질하고, 찌거나 굽는 등의 활동이 필요하기도 하다. 일상적으로 하는 간단한 요리도 여러 개의 재료와 조리 단계로 구성되는데, 이는 상당히 많은 양의 정보로 레시피에서 이 정보를 어떻게 전달하는지에 따라 사용자들에게 기억과 처리 부담을 야기할 수 있다.

레시피가 부과하는 기억 부담은 스마트 스피커처럼 정보가 청각적으로 제시되는 경우 더 문제가 될 수 있다. 음성 정보의 휘발성으로 말미암아 시연(rehearsal)과 같은 기억전략을 사용하지 않으면 작업기억에서 사라지기 때문이다. 기억 부담은 스마트 스피커 사용자들이 종종 지적하는 문제인데[9], 과연 어느 정도의 정보가 기억 부담을 야기하는지는 분명하지 않다. Bigot 등은 음성 상호작용 시스템에서 사용자가 5개의 단어가 한꺼번에 제시되면 이전 단어를 기억하는데 어려움을 겪는 것을 보고하였다[10]. Hamada와 동료들은 단어 대신 행동 단위(Action Units; AUs) 개념을 제안했다[12]. 행동 단위는 레시피에서 조리 행동의 단위를 지칭하는데, 예로 “양파를 잘라라”나, “양파를 볶아라”가 하나의

행동 단위가 된다. “양파를 5분간 볶아라”와 같이 해당 행동이 얼마 동안 일어나는 것이 좋은지에 대한 시간 정보를 포함하기도 한다. Hamada등은 조리 지시 당 하나의 행동 단위를 제안하였으나, 조리 지시 당 어느 정도의 정보 제시가 적절한지는 상황에 따라 달라질 것으로 보인다. 예로, 나이가 증가하면 레시피에 대한 기억 능력이 감소할 수 있다 [13, 14]. 또한, 조리에 대한 지식과 경험이 많은 경우 정보를 더 묶어서(chunking) 처리할 수 있기 때문에[15], 한 번에 기억하는 양이 많아질 수 있다. 사용자마다 기억할 수 있는 정보의 양은 달라지겠지만, 분명한 것은 사용자들이 한 번에 기억하고 실행에 옮길 수 있는 정보는 제한적이라는 것이다. 특히 청각적으로 정보가 제시되는 경우 사용자의 제한된 기억 용량에 대한 고려가 필요한데, 수적이다.

한 번에 제시되는 정보의 양뿐만 아니라 정보의 제시 시점에 대한 고려 또한 필요하다. 예로, 레시피에서 ‘단호박을 깨끗이 씻어 씨를 파내고, 전자레인지에 2분간 익히주세요.’와 같은 지시를 하는 경우, 이 지시에는 단호박을 씻고, 씨를 파내고, 전자레인지에 익히는 3개의 행동 단위로 이루어진다. 각 행동 단위를 실시하는 데는 시간이 걸리기 때문에 마지막 지시를 완결하기까지는 몇 분 또는 그 이상이 걸릴 수도 있다. 여러 조리 활동에 대한 지시가 한꺼번에 이루어지면 사용자는 마지막 행동이 완결되기 전까지 해당 지시를 작업기억에 유지하고 있어야 하는데, 조리에 집중하다 보면 종종 정보가 작업기억에서 손실되는 일이 발생한다. 음성 기반의 레시피 정보는 손을 자유롭게 하지만, 손의 자유로움이 기억 부담으로부터의 자유로움을 의미하지는 않는다. 효과적인 레시피 서비스 디자인을 위해서는 레시피 참조 상황에서 발생하는 기억 부담의 정도와 성격을 명확하게 이해하는 것이 필요할 것으로 보인다.

둘째, 스마트 스피커는 기본적으로 음성을 사용하여 정보를 전달하지만, 조리 활동은 오감을 요구하는 활동이다. 재료의 상태를 확인하고, 적절하게 준비하는데, 조리 활동이 적절한 정보로 수행되었는지, 요리가 완성되었는지를 가늠하는 시각, 후각, 청각 등의 감각이 모두 관여한다. 이러한 정보를 글이나 말로 모두 전달하는 데는 한계가 있고, 이를 보완하기 위해 요리책이나 웹사이트에서는 종종 시각 정보를 추가한다. 이미지나 영상정보는 재료의 모양과 상태, 준비 방법, 조리 도구 사용, 조리의 중간 및 완성된 상태 등에 대한 다양한 정보를 제공하는데, 익숙하지 않거나 새로운 음식을 만들 때 이러한 시각 정보의 역할은 결정적이다. VUI 기반으로 레시피를 제공할 때는 이러한 이미지 정보를 전달할 수 없게 된다. 영상 기기와 연동하여 이를 보완할 수 있으나, 음성 정보만으로 레시피를 전달하고, 조리 활동을 충분히 안내할 수 있는지, 이때 발생하는 문제점에는 무엇이 있는지에 대한 이해가 필요하다.

3. 연구 문제

본 연구에서는 스마트 스피커 레시피 서비스를 이용하여 요리할 때 발생하는 사용자의 경험을 살펴보았다. 본 연구는 현장 연구로 참가자들의 집에서 실시되었다. 요리 숙련자와 비숙련자 각 3명의 참가자가 연구에 참여하였고, 스마트 스피커 레시피 서비스를 사용하여 참가자들이 이제까지 조리해보지 않은 음식인 ‘단호박 버터구이’를 조리하였다. 사용자 경험 자료를 얻기 위해, 소리내어 생각하기 방식, 관찰, 인터뷰 방식을 사용하여 자료를 수집하고, 내용 분석을 실시했다.

스마트 스피커 레시피 서비스의 경우 아직 서비스 제공 초기 단계로, 기존의 서비스 방식(예, 책 또는 요리 영상)과의 비교는 아직 이른 것으로 보인다. 또한 음성 기반 서비스는 기존의 서비스와 대척점에 존재하기 보다는 보완적인 관계에 있는데, 음성기반 서비스 제공에 따르는 사용자의 경험을 보다 정교하게 살펴보는 것을 통해서 보다 나은 음성기반 서비스를 디자인하고 효과적으로 영상기반 서비스와 통합하는 것이 가능하다. 본 연구에서 사용자 경험의 탐색은 다음 세 질문을 중심으로 이루어졌다.

첫째, 음성기반 레시피 서비스가 가진 장점 중 하나는 진행 중인 활동을 방해하지 않는 것이다. 실제 조리 장면에서 핸드프리(‘손이 자유로워지는 상태’) 상태가 얼마나 나타나는지 살펴보았다. 음성 기반 레시피 서비스를 사용하면서, 레시피를 확인하기 위해 진행 중인 활동을 중단하거나, 시선을 이동하는 노력을 하지 않아도 되는 가능성을 확인하고자 했다. 또한, 요리를 하는 사람의 요리 능력이나 기기에 대한 경험에 따라서 이 장점이 제대로 활용되지 않을 수 있다.

둘째, 레시피를 사용하여 조리를 진행하는 동안 사용자가 기억 부담을 경험하는지 살펴보았다. 인간의 제한된 작업기억 용량은 잘 알려져 있으나, 이러한 사용자의 정보처리 특성이 조리 상황에서 어떤 영향을 주는지는 아직 제대로 연구된 바가 없다. 작업기억용량제한으로 인한 기억 부담은 VUI 사용 시 더 부각될 것으로 보이는데, 기억 부담을 느끼는 경우 되뇌기와 같은 정보유지 전략을 사용하거나 이전에만 들은 정보를 재요청하는 일이 증가할 것으로 보인다.

셋째, 시각 정보 없이도 음성만을 이용하여 레시피를 제공하는 것이 사용자의 요리 과정을 충분히 안내할 수 있는지 살펴보았다. 시각 정보는 음성 정보보다 많은 정보를 가지고 있고 완성된 음식 또는 중간 단계에 대한 이미지 정보의 부재는 조리의 성공에 결정적인 영향을 미칠 수 있다. 하지만, 음성 정보만으로 레시피를 제공하는 것은 불가능한 일만은 아닌데, 사진이나 영상이 보편화되기 이전에는 레시피가 글로만 구성되었고 여전히 이러한 형태의 레시피가 존재한다. 본 연구에서는 조리 상황에서의 음성 정보만으로 조리 과정을 충분히 안내할 수 있는지, 문제가 발생한다면 어떠한 유형의 문제가 발생하는지를 알아보았다.

넷째, 요리 숙련자와 비숙련자에 따라 레시피 서비스를 사용하는 데 있어 차이가 나타나는지 탐색하였다. 요리 정보 레시피 서비스를 사용하는 사용자들은 연령, 성별, 요리 경험, 기기 조작 능력 등에 있어서 다양한 차이가 존재할 것으로 기대된다. 이들 변인들의 영향을 체계적으로 알아보는 것이 본 연구의 목적은 아니지만, 사용자의 요리 숙련도는 조리 과제 수행에 큰 영향을 미칠 수 있기에 이를 통제하기 위해서 숙련자와 비숙련자를 구분하여 모집하였고 이들의 사용 경험을 구분하여 살펴보았다. 단순히 어느 집단의 수행이 더 뛰어난지가 아니라 각 집단이 경험하는 사용성의 문제가 무엇인지를 알아보는 것이 목적이었다.

4. 연구 방법

4.1 참가자

요리 숙련자와 비숙련자 각 3명, 총 6명의 참가자가 연구에 참여하였다. 요리 숙련도는 사전 설문 조사를 통해 최근 지속해서 요리한 기간과 회수를 조사하였다. 지속해서 4년 이상 동안 주당 1회 이상 요리를 한 참가자를 숙련자로, 그렇지 않은 참가자들을 비숙련자로 분류했다. Table 1에 참가자들의 성별과 나이가 제시되었는데, S1-3은 숙련자, NS1-3은 비숙련자를 지칭한다. Table 1에서 볼 수 있듯이 두 집단 간에 성별 및 나이 구성에 있어 차이가 존재하는데, 숙련자의 평균 나이는 47.3세, 비숙련자는 19.6세였다. 숙련자 집단은 모두 여성으로 구성되었으나, 비숙련자 집단은 2명의 남성과 1명의 여성 참가자였다. 이는 요리 숙련도를 쌓기 위해서는 일정 기간이 필요한 것과 세대 간 요리에 관한 관심의 차이를 반영하는 것으로 보인다. 참가자들은 모두 자신의 집에서 이루어지는 자료수집과 녹화에 동의했고, 본 연구는 한림대학교 IRB 심의를 거쳤다(심의승인번호: HI RB-2020-030).

Table 1. Gender and age of participants

Participant	Gender	Age
S1	Female	50's
S2	Female	40's
S3	Female	40's
NS1	Female	20's
NS2	male	20's
NS3	male	20's

사용성 평가의 목적은 모집단의 차이를 규명하거나, 변인의 인과성을 증명하는 것이 아닌 특정 제품이나 디자인의 문제를 발견하여 개선하는 데 있다. 따라서, 통계적 검증력이 아닌 사용성 문제를 충분히 탐지할 수 있는지가 표본의 크기를 결정하는데 중요한 기준으로 작용한다[16]. 대상이 되는 제품이 무엇인가에 따라 적절한 표본의 크기가 달라진다[17], 일반적으로 컴퓨터를 기반한 인터페이스의 경우 5명

내외가 충분하며, 동일한 기기에 대해서 여러 사용자 집단이 존재하는 경우 집단당 3~4명 정도도 충분한 것으로 간주된다[18, 19]. 선행 연구에서도 유사한 수의 참가자들을 대상으로 자료가 수집되었는데, 김보민 등(2014)의 경우 총 3명(요리 경험 상·중·하 각 1명)에 대한 관찰을 하였고, Hamada 등(2005)의 연구에서는 총 8명(요리 경험 상 2인, 중 및 하 3인)에 대한 자료를 수집했다.

4.2 도구

국내의 레시피를 제공하는 스마트 스피커는 대표적으로 SKT의 '누구', KT의 '기가지니'가 존재하는데, 두 기기 모두 비슷한 방식으로 레시피 정보를 제공한다. 사용자가 특정 레시피 정보를 요청하면, 스마트 스피커가 해당 레시피를 찾아 제공하는데, 이때 정보제시는 조리 단계별로 이루어진다. 사용자가 '다음', '다시' 등의 명령을 사용하여 조리 지시를 다시 듣거나 다음 조리 단계로 이동할 수 있다. 본 연구에서는 스마트 스피커는 SKT의 '누구 mini'가 사용되었다(앞으로 '누구'로 지칭).

레시피 출처는 '(주)이밥차'이다. '누구'의 호출명은 '아리아'로 참가자의 "아리아, 000레시피 알려줘", "아리아, 000 어떻게 만들어?"와 같은 요청으로 서비스가 시작된다. '누구'를 호출하면, 우선 레시피에 필요한 재료를 알려주고 준비가 되면 "요리 시작"이라고 말하라고 알려준다. 사용자가 재료를 준비하고, "아리아, 요리 시작"이라고 하면, 조리 지시가 단계별로 제시된다. 다음 단계로 진행하기 위해서는 "다음", 이전 단계의 지시를 다시 듣고 싶으면 "다시" 명령어를 사용할 수 있다. 조리 중간에는 클래식 곡이 재생된다.

4.3 조리 과제

'누구' 사용자 안내서에 적힌 레시피 371개를 검토하여 사전 설문을 거쳐 참가자 6명이 공통으로 만들어보지 않은 요리인 '단호박 버터구이'를 선정하였다. Table 2에서 볼 수 있듯이 5개의 재료와 4단계의 조리 단계로 구성된다. 또한, 각 조리 단계는 각기 다른 행동 단위의 수로 구성된다.

Table 2. Butter roasted sweet pumpkin recipe

Ingredients	Half sweet pumpkin, 1 tablespoon butter, 1/4 cup sliced almonds, sugar, cinnamon powder	
Steps	1	wash the sweet pumpkin, scoop out the seeds, and cook it in the microwave for 2 minutes. (3)*
	2	cut the cooked sweet pumpkin into thick C-shapes. (1)
	3	melt the butter in the microwave for 30 seconds, then it with the sweet pumpkin. (2)
	4	put the pumpkin on a frying pan and roast it over medium heat and sprinkle

with sugar, cinnamon powders and almonds to your liking. (3)

* 괄호 안 숫자는 각 단계를 구성하는 행동 단위의 수를 지칭

사용자가 “아리아, 요리 시작”이라고 하면, 1단계부터 제시한다. 1단계를 예를 들면, “1단계, 단호박을 깨끗이 씻어, 씨를 파내고, 전자레인지에 2분간 익혀주세요. 다음 단계를 들으시려면, 아리아, 다음. 이라고 말씀해주세요.”라고 한다. 사용자가 해당 단계에 필요로 하는 행위를 한 뒤, “아리아, 다음 단계”라고 하면, “2단계, 익힌 단호박은...(생략)”라고 한다. 조리 과제를 수행하는 동안 참가자들은 소리내어 생각하기를 했다. 소리내어 생각하기는 과제 수행 중 사용자들의 내적인 사고를 파악하기 위한 것으로 사용성 테스트에 종종 사용되는 자료수집 절차이다[20]. Ericson과 Simon에 의해 1984년에 확립된 방법으로 단기 기억의 내용을 밖으로 외현화하는 방식으로 간주한다[21].

4.4 사용자 경험 평가 절차

실험 참가자의 동의를 얻어 참가자의 집에서 자료수집이 이루어졌다(Figure 1). 연구자가 사전에 조리 과제에 필요한 재료와 도구를 준비하였으나, 참가자가 원하는 경우 자신의 집에 보유하고 있는 것을 사용할 수 있었다.

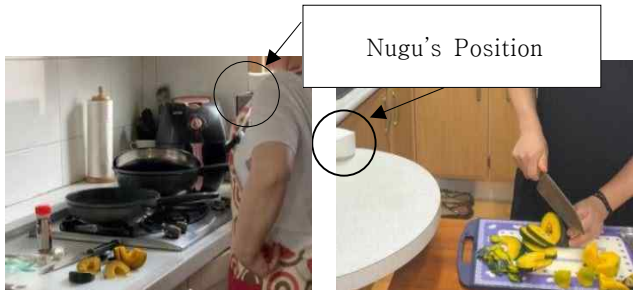


Figure 1. Screenshot of participants' cooking environment

조리를 시작하기 전에 참가자들은 약 10분 동안 레시피 서비스 호출에 사용되는 명령어를 습득했다. 이후 소리내어 생각하기가 익숙하지 않은 참가자들을 위해 5분 동안 연습을 실시했다. 연습과제로는 조리 과제와 무관한 ‘종이접기’가 사용되었다.

참가자들은 스마트 스피커에서 제공되는 음성 레시피만을 사용하여 ‘단호박 버터구이’를 조리했으나, 요리 도중에 완성된 요리 사진을 보고 싶을 때는 연구자에게 요청할 수 있었다. 이는 시각 정보의 부재로 조리 활동을 지속하지 못하는 일이 발생하는 것을 방지하고 어느 시점에서 시각 정보가 필요한지 알아보기 위함이었다. 참가자들의 조리 활동과 발화는 모두 녹화되었다.

조리를 마친 후, 약 5분 동안 반 구조화된 인터뷰를 실시했다(Table 3). 첫 번째 질문은 스마트 스피커를 사용한 조리

활동에 대한 난이도에 대한 것으로 7점 척도로 답변하게 했다. 어렵다고 답변한 경우(즉, 중간값이 4점 이하) 이유를 추가적으로 질문하였다. 두 번째 질문에서는 평소 조리 행동에 대한 것으로 평소 레시피를 사용하여 조리하는지, 레시피를 사용하여 조리하는 경우 사용 매체와 스마트 스피커와의 차이를 질문했다. 마지막 질문은 스마트 스피커 레시피 서비스에 대한 추가 사용 의도와 이유에 대한 것이었다. 추가적으로 조리 과제 수행 중 특이한 행동을 한 경우 참가자별로 질문을 하였다. 예로, 조리 완성 사진을 요청한 경우 이유를 질문하였고, 기타 의미와 의도가 불분명한 발화와 행동(예, 조리 중 이유없는 웃음)에 대해서 추가 질문을 통해 참가자들의 발화 의도를 파악했다.

Table 3. Interview questions

	Questions
1	How do you feel about the difficulty of cooking using the recipe service using the smart speaker? (Answer on a 7-point scale of 7-points very easy and 1-point very difficult)
1-1	(If the participant answered that question 1 was difficult) You answered the difficulty level with 4-points or less, why do you think so?
2	Do you usually check recipes before cooking?
2-1	If you check recipes, what medium do you usually check through?
2-2	Compared with the traditional way, what are the pros and cons of smart speaker recipe service?
3	Would you like to try the service you used today again next time?
3-1	If you are willing (or not) to try it, why?

5. 결과

연구 질문 1-4에 대한 분석 결과를 섹션 5-1부터 5-4에 걸쳐 살펴보았다. 5-5에서는 음성기반 레시피 제공 서비스에서 추가적으로 고려해야 할 사항, 5-6에서는 사용자의 인터뷰 답변을 살펴보았다.

5.1 스마트 스피커의 사용 및 핸즈프리

참가자들은 조리 과제를 수행하면서 활발하게 스마트 스피커를 사용했다. 코딩을 단순화하기 위해 조리 녹화영상을 1분 단위 구간으로 나누어 스마트 스피커 사용 빈도를 분석했다. 참가자들의 평균 조리 시간은 26분으로 이에 상응하는 평균 26개의 구간이 존재했다(영상 길이가 1분 미만일 때, 10초가 넘어가는 경우 독립된 구간으로 간주하였다. 예로, 영상의 길이가 24분 2초인 경우 24구간으로, 24분 20초인 경우 25구간으로 구분하였다.)

각 구간에서 참가자가 스마트 스피커를 사용했는지 살펴보았

는데, 각 구간에서 한번이라도 스마트 스피커 레시피 서비스를 호출한 경우 ‘사용’으로 코딩되었다. 대부분 참가자가 전체 조리 구간 동안 스마트 스피커를 활발하게 사용했다. 참가자 평균 26개 구간 중 평균 11.8개 구간(45%)에서 스마트 스피커 사용이 관찰되었다. 즉, 조리 활동을 하는 구간의 약 반 가까운 구간에서 레시피를 참조하는 활동이 일어났다. 또한, 참가자들이 요리 활동을 하면서, 핸드프리 상태가 나타나는지 살펴보았다. 각 구간에서 스마트 스피커를 사용하면서 동시에 조리 활동을 수행하는 것(예, 단호박을 C모양으로 썰면서, 동시에 “아리아, 다음”과 같이 요청하는 것)이 한 번이라도 관찰되면 해당 구간에서 핸드프리가 일어난 것으로 코딩했다. 스마트 스피커가 사용된 11.8개 구간 중 평균 10개 구간에서 핸드프리 사용이 관찰되었는데, 이는 스마트 스피커를 사용한 구간의 85%에 해당한다. 이러한 결과는 스마트 스피커가 조리 과정에서 활발하게 사용되고, 대부분의 사용이 조리 활동과 병행하여 이루어짐을 보여준다.

5.2 재요청 및 되뇌기

조리 활동 동안에 참가자가 경험하는 기억 부담을 알아보기 위해서 스마트 스피커와 참가자의 상호작용 중 레시피 재요청과 되뇌기 전략 사용을 살펴보았다. ‘누구’와의 상호작용은 ‘사용자 요청-누구 응답’을 한 턴(turn)으로 진행되는데, 사용자가 “레시피 시작”, “다음”, “다시” 등을 요청하면, ‘누구’는 해당 요청에 대한 응답을 한다. 재요청은 참가자들이 ‘누구’의 대답을 듣고, “아리아, 다시”와 같이 앞에 들었던 정보를 다시 요청한 경우를 지칭한다. 참가자들은 스마트 스피커와 상호작용하면서 평균 26.5회(턴) 발화하였다. 이 중 ‘누구’가 응답 실패한 경우를 제외하여 평균 6.3회(23.8%)턴에서 재요청이 관찰되었다.

사용자들은 또한 ‘누구’의 조리 지시를 듣고, 종종 되뇌는 행동을 보였다. 되뇌기(rehearsal)는 정보 부하 상황에서 사용하는 기억전략의 하나로[22], 참가자들은 조리 지시를 듣고 기억 부담을 느끼고, 적극적으로 기억전략을 사용한 것으로 보인다. 참가자들이 각 턴에서 ‘누구’의 대답을 듣고 제시된 정보를 되뇌기 한 경우 코딩했다. 한 턴에서 1회 이상의 되뇌기를 하였을 때, 되뇌기를 1회만 한 것으로 간주했다. 코딩 결과 평균 4.8개(18.1%) 턴에서 되뇌기가 관찰되었다. 어떠한 정보를 되뇌기하였는지 살펴본 결과 1단계와 3단계에서 ‘2분간’, ‘30초’와 같은 시간 정보가, 2단계에서는 ‘C 모양’에 대한 되뇌기가 나타났다.

재요청과 되뇌기가 조리 단계별로 다르게 발생하는지를 살펴본 결과, 재료 제시 단계에서 가장 적게 발생하였다(Figure 2). 조리 단계에서는 재요청과 되뇌기가 고르게 나타났다. 이는 조리 과제인 ‘단호박 버터구이’의 경우 조리 단계조리 단계별로 2-3개 정도의 행동 단위로 구성되어 단계별 행동 단

위의 편차가 크지 않았기 때문으로 보인다.

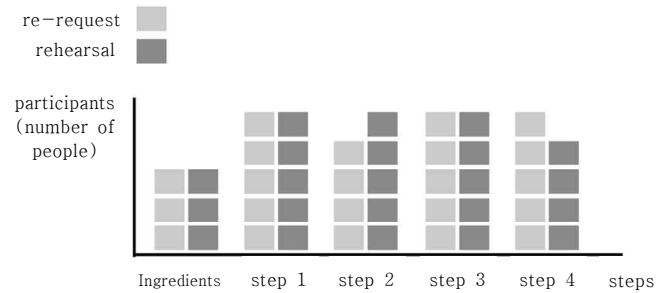


Figure 2. Participants distribution of re-request and rehearsal by cooking steps

5.3 시각 정보의 부재와 요리 완성도

대부분의 참가자들은 스마트 스피커의 음성 정보만으로 조리 과제를 완성했으나, 두 명의 참가자(S3, NS2)는 완성된 조리 사진을 요청했다. 사진을 요청한 지점은 S3의 경우 조리 2단계에서 단호박을 ‘도톰하게 C모양으로 썰어주세요’라는 지시를 받고 난 다음 요청했다. ‘C 모양’이라는 지시를 듣자, S3는 “이거 어떻게 해야 해. 씨 모양이 아니, 껍질 벗기라는 소리도 없고”라는 발화를 했다. 이후에도 “씨 모양이 어느 모양인지 하기 전에 사진을 봤으면 좋겠는데”라고 했다. 이런 발화에서 볼 수 있듯이 S3는 ‘C(씨)모양’이라는 지시에 대한 의문을 가진 것인데, 음성으로만 들었을 때, ‘씨’는 씨앗을 의미할 수도 있고, 알파벳을 지칭할 수도 있다. 이런 의문에도 불구하고 S3는 씨앗처럼 보이는 모양으로 일단 단호박 썰기를 시작했다. 하지만, 결국 단호박 썰기를 마치기 전에 사진을 요청했다. S3의 최종 완성된 요리는 이러한 혼란을 반영하는데(Figure 3.a), 일부의 단호박(접시 왼쪽)은 C모양으로 썰려있지만, 일부(접시 오른쪽)는 씨앗과 같은 모양으로 썰려있다.

NS2는 조리 1단계에서 사진을 요청했다. 1단계는 단호박을 썰기 전, 단호박을 씻고 씨를 파내고, 익히는 단계이다. NS2는 단호박을 씻은 뒤, 씨를 파내기 전에 ‘누구’에게 “아리아, 이거 어떻게 파?”라고 질문했는데, ‘누구’로부터 적절한 답변을 받지 못하자 사진 요청을 했다. 사진을 확인하고 단호박을 반으로 자른 뒤, 씨를 파냈다. 조리가 끝나고 진행된 인터뷰에서 NS2는 “단호박 씨를 빼라고 했는데, 할로윈 때 호박 만드는 것처럼 원형으로 파서 빼는 것인지 자른 다음에 씨를 빼는 것인지 헷갈려서” 사진을 요청했다고 응답했다. NS2의 경우 ‘씨’와 ‘C’의 혼돈이 아니라 ‘씨를 파라’는 지시에 대한 의문을 가지고 있었는데, 이처럼 지시 자체의 의미가 분명한 경우에도 여러 방식으로 실시가 가능한 경우 사용자들에게 혼돈을 야기할 수 있는 것으로 보인다.

사진을 요청하지 않은 나머지 참가자들도 종종 모양에 대한 의문(예, “씨 모양으로? 설명이 좀 부족한데”)을 표시했다.

예외적으로 S2의 경우 사진을 요청하지도 ‘씨 모양’에 대한 의문을 표현하지 않았지만, 최종 조리 사진은 단호박 씨는 모양을 알파벳 C모양이 아닌 씨앗 모양으로 이해한 것으로 보인다(Figure 3.b).



(a) S3

(b) S2

Figure 3. Butter roasted sweet pumpkin made by participants

5.4 요리 숙련도의 차이

집단별 참가자 수가 충분하지않지만, 숙련도에 따라 스마트 스피커 레시피 서비스 사용에 차이가 존재하는지 살펴보았다. 일련의 관찰치에서 다음과 같은 차이가 관찰되었다.

첫째, 두 집단의 조리 시간 차이가 존재했다. 숙련자 집단은 평균 21분, 비숙련자 집단은 평균 31분으로 숙련자 집단이 더욱 빨리 완성했다. 조리 시간 차이를 보다 구체적으로 알아보기 위해 집단별로 단호박 자르기를 하는 데 걸리는 시간을 동영상에서 측정하였다. 그 결과, 숙련자 집단의 경우 모든 참가자가 30초 이내에 해당 단계의 조리 행동을 완결하였으나, 비숙련자의 경우 모두 30초 이상이 걸렸다. 특히 NS1의 경우 단호박이 단단하여 자르기에 어려움을 경험하여 실험자의 도움을 받아야 했고, NS2는 씨는 모양에 대한 의문으로 사진을 요청하여 시간이 더 소요되었다.

둘째, 숙련자 집단은 비숙련자 집단보다 스마트 스피커 레시피 서비스를 더 자주 사용하고 스마트 스피커 사용과 핸드프리를 더 활발하게 사용하였다(Figure 4). 또한, 숙련자 집단은 비숙련자 집단과 비교하면 핸드프리가 12% 더 나타났다.

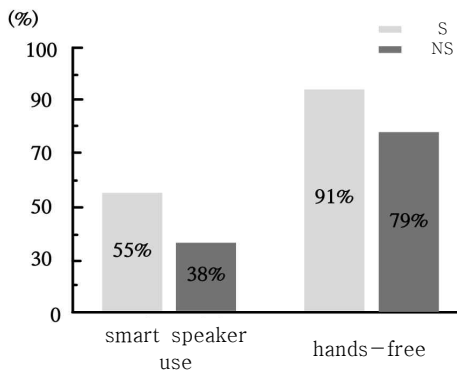


Figure 4. Percentage of smart speaker use and hands-free section by group

셋째, 비숙련자 집단의 경우, 숙련자 집단보다 레시피를 재요청하는 빈도가 28%로 두 배 이상 더 높았다. 숙련자 집단에서는 스마트 스피커와의 상호작용 중 13%에서 앞에 제시된 레시피를 재요청한 반면, 비숙련자 집단에서는 41%에서 재요청이 일어나 두 배 이상 차이가 존재했다(Figure 5). 두 집단의 연령 차이를 고려할 때, 기억 실패가 숙련자 집단에서 더 많이 일어날 것으로 기대되지만, 실제로는 비숙련자 집단에서 재요청이 더 많았다. 숙련자들의 경우 과거의 조리 경험에 바탕하여 레시피의 정보를 더 효과적으로 군집화(chunking)할 수 있었던 것으로 보인다[15].

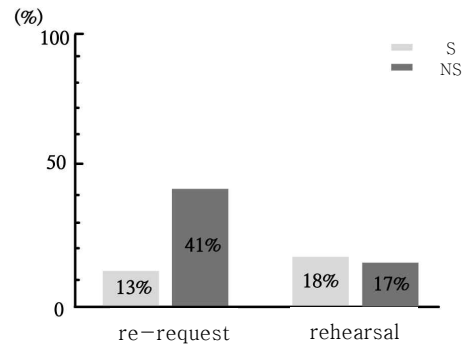


Figure 5. Percentage of re-request and rehearsal by group

숙련자 집단과 비숙련자 집단 간에 스마트 스피커 레시피 서비스 사용의 뚜렷한 차이가 관찰되었지만, 되뇌기 행동의 경우 숙련자와 비숙련자 간에 뚜렷한 차이가 존재하지 않았다(Figure 5). 되뇌기는 기억전략의 사용을 의미하는데, 이는 두 집단 모두 조리 지시를 듣고 기억을 하기 위한 노력을 비슷한 정도로 기울였음을 의미한다.

또한, 스마트 스피커 레시피 서비스에서 시각 정보의 부재는 두 집단에 차별적으로 작동하지 않은 것으로 보인다. 완성된 요리 사진 요청, 모양에 대한 의문 표시 등에 있어서 집단 차이가 존재하지 않았다.

5.5 사용자의 정보 요구와 언어적 상호작용

참가자들은 조리 과정에서 스마트 스피커에게 다양한 질문과 요청을 했는데 이들 질문은 사용자들이 가지고 있는 레시피에 대한 추가적인 정보 요구와 기대를 보여준다(Table 4). 예를 들어, 레시피에 언급되지 않은 세부 정보에 대한 질문이 상당 존재했는데(예, “호박에 씨 빼고, 껍질도 베껴내니?”, “후라이팬 필요해?”), 이들 질문은 레시피에서 언급되지 않거나, 불충분하게 언급된 내용들이다.

참가자들의 질문과 관련하여 흥미로운 점은 참가자들이 ‘누구’가 이런 질문에 답할 수 없다는 것을 알고 있음에도 이런 질문을 하였다는 것이다. 현재 ‘누구’가 알아들을 수 있는 명령어들은 “요리 시작”, “다음”과 같은 요청에 제한되어 있다.

참가자가 Table 4에 제시된 것과 같은 질문을 하면, ‘누구’는 ‘말씀하신 것을 찾지 못했어요’, ‘이해하지 못했어요’와 같은 답변을 한다. 참가자들이 제대로 답변을 듣지 못함에도 이러한 질문을 하는 것이 관찰되었다.

Table 4. Questions participants asked smart speakers

	Type	Example	Questioner
1	Recipe Details	“How thick is it?” (S1) “Do you remove the seeds from the pumpkin and also remove the pumpkin skin?” “Need a frying pan?”	S1
2	Cooking Method	“Tell me how to cut a pumpkin” “How do you dig this?”	NS1, NS2
3	etc.	“Can you play a trot other than the classics?”	S1

5.6 인터뷰

조리 과제 난이도에 대한 첫 번째 질문에 대해서 참가자들은 대체로 쉽다는 응답을 하였다(M=5.33). 조리 행동에서 보여진 일련의 차이들과 일관되게 숙련자(M=6.33)들이 비숙련자들(M=4.33)이 더 쉽다고 답변하였다. 4점 이하로 난이도가 보통 이상이라고 답변한 참가자들(NS2, NS3)에게 이유를 물어본 결과, 두 참가자 모두 레시피 서비스의 더욱 자세한 설명이 필요함을 언급했다.

두 번째 질문에 대해서 참가자 6명 중 한 명(S1)을 제외하고는 평소에 레시피를 사용하여 조리한다고 응답하였다. 레시피를 확인하는 경우, 포털 사이트 검색창(S2, S3, 및 NS3), 유튜브나 블로그(NS1과 NS2), 레시피 앱(NS3)이 언급되었다. 기존에 레시피를 얻는 방식과 비교할 때, 스마트 스피커 레시피 서비스의 장점으로 레시피를 확인하기 위해 번거롭게 핸드폰을 확인하지 않아도 된다는 점을 언급하였다(예, “유튜브는 핸드폰 영상을 보면서 하니까 손에 기름이 묻은 경우 불편한데 이걸 그렇게 안해도 되니까 좋다”, “블로그는 요리를 할 때 보면서 하기 힘든데 이걸 원하는 부분을 말하면 다시 들려주니 좋은 것 같다”). 5명의 답변자 중 4명이 관련된 언급을 하였을 뿐만 아니라 스마트 스피커 레시피 사용을 어려워한 비숙련자 집단 또한 모두 이 기능을 언급하여 조리에 어려움을 느낀 사용자조차도 핸즈프리 기능을 편리하게 느꼈음을 보여준다. 기타 중간에 음악이 나와 마음이 편해진다는 것도 장점으로 언급되었다. 반면, 단점으로는 음성 인식 오류, ‘C’ 모양으로 자르라는 모호한 안내와 같이 레시피 내의 조리 방식에 대한 설명이 부족한 것, 재료나 모양의 상태를 확인할 수 없는 점, 그리고 전체 조리 과정에 대한 정보를 한눈에 확인할 수 없는 점 등이 지적되었다.

세 번째 질문에 대해서 6명 중 4명(S1, S2, S3, 및 NS2)

이 재사용 의사를 밝혔다. 재사용 이유로는 모르는 요리를 할 때 도움이 될 것 같으며, 요리하는데 지루하지 않을 것 같다는 의견이 있었다. 반면, 재사용하지 않겠다는 답변의 경우 시각 정보의 부재, 핸드폰의 익숙함, 레시피를 외워야 한다는 부담감 등이 이유로 제시되었다. 숙련자 집단의 참가자들 모두는 재사용 의향이 있었으나 비숙련자 집단에서는 한 명의 참가자만이 재사용 의향을 밝혔다.

6. 논의 및 결론

본 연구는 요리 상황에서 음성 기반 레시피 서비스가 사용자의 요리 활동을 얼마나 효과적으로 지원을 하는지 알아보기 위해, 소리내어 생각하기 기법과 관찰 및 인터뷰를 시행했다. 본 연구의 결과는 HCI의 발전과 음성을 사용한 레시피 서비스 사용자 경험 연구에 참고 할 수 있을 것으로 기대된다. 연구 결과를 바탕으로 몇 가지 논의점을 도출하였다. 이는 다음과 같다.

1. 스마트 스피커의 사용 및 핸즈프리

사용자들은 요리 도중 손이 자유로운 것을 경험하였다. 이는 음성 기반 레시피 서비스가 사용자의 진행 중인 요리 활동을 방해하지 않고, 시선을 이동하게 만들지 않는다는 것을 의미할 수 있다. 이는 레시피 서비스에는 음성 정보 제공이 효과적일 수 있음을 시사한다.

2. 재요청 및 되뇌기

사용자들은 앞선 조리 지시를 기억하는데 종종 어려움을 경험하였고, 스마트 스피커와의 상호작용 중 상당 부분은 앞선 조리 지시를 다시 듣는 데 사용되었다. 재요청은 특히, 비숙련 사용자들이 조리하는 과정에서 음성 정보를 처리하는 부담을 온전히 제거하지는 못하는 것으로 보인다. 한꺼번에 너무 많은 정보를 제시하게 되면 기억 부담이 초래되고 정보가 손실되기도 한다. 이러한 기억 부담은 동일한 조리 단계에 대한 재요청 및 정보를 잊지 않으려는 되뇌기 기억전략의 사용을 촉발시켰다. 효과적인 스마트 스피커 레시피 서비스의 사용성을 개선하기 위해서는 사용자가 경험하는 기억 부담을 완화시켜주는 것이 필요한 것으로 보인다. 기억 부담을 완화하는 방법 중 하나는 한 번에 제시하는 정보의 양을 줄이는 것으로 예를 들어 Hamada등이 제안하였듯이 한 조리 활동이 하나의 행동 단위로 이루어지도록 레시피를 제시하는 것을 고려할 수 있다[12]. 그러나 한 번에 한 조리 동작만 지시하는 것이 항상 효과적인지는 아직 불분명한데, 해당 음식을 여러 번 조리해본 경우나 요리 경험이 많은 경우에 오히려 레시피 참조를 더 번거롭게 만들 수 있다. 조리 과정에서 동시에 여러 개의 조리 활동이 병행해서 일어나거나 다음 활동을 준비하고자 하는 경우(예, 물을 끓이면서 재료를 손질) 한 번에 하나의 조리 행동만 지시하게 되면 조리 활동의 지연이 발생할 수 있다. 이는 사용자

변인, 레시피의 종류, 상황적 요인 등에 의해서 효과적인 정보제시 단위가 어떻게 영향받는지에 대한 연구가 선행되어야 할 것으로 보인다.

3. 시각 정보의 부재와 요리 완성도

글이나 말만 사용하여 레시피가 제공되는 경우가 여전히 존재하나 현대의 사용자들은 이미 다양한 이미지와 영상정보에 매우 익숙해져 있고, 음성 정보와 함께 시각 정보를 연동하여 제시하는 것이 더욱 정확한 레시피 전달을 위해서 도움이 될 것으로 보이며, 이미 핸드폰이나 TV 모니터 등으로 레시피의 영상 이미지를 함께 전송하는 방식이 시도되고 있다[23][24]. 하지만 동시에 음성만으로도 최대한 완전적인 레시피 서비스를 설계하는 것이 필요하다. 영상기기의 사용이 모든 상황에서 가능하지 않을 수도 있고, 시각 정보를 사용할 수 없는 사용자들도 존재한다. 이 경우 레시피의 전달력을 높이기 위해 조리 지시의 의미를 명확하게 하는 것이 중요한데, 본 연구에서 관찰되었듯이 동음이의어가 포함된 레시피(‘씨’, ‘C’)의 경우 음성 정보만 의존할 때 의미 혼돈이 발생할 수 있다. 의미의 모호성은 본 연구의 조리 과제와 완성도를 떨어뜨리는 요인으로 작용했다. 따라서, 이러한 동음이의어가 존재하는 경우에는 의미가 명확하게 전달되도록 레시피를 수정하는 것이 필요할 것으로 보인다. 청각 정보를 사용하여 레시피의 전달력을 높이는 방안도 생각해 볼 수 있다. 예를 들어 음식이 조리될 때 발생하는 소리 등을 배경에 제공하는 것은 지시의 모호성을 해소하는데 도움이 될 수 있을 것으로 보인다.

4. 요리 숙련도의 차이

본 연구에서 보았듯이 사용자의 조리 경험에 따라서 레시피 서비스에 대한 사용 경험이 달라졌다. 비숙련자의 경우 조리 경험의 부족으로 기본적인 조리 동작에 대해서도 추가적인 설명이 필요했거나, 조리에 더욱 시간이 걸렸고 추후 서비스 재사용 의사도 낮았다. 두 집단의 연령 차이를 고려할 때, 오히려 비숙련자들이 기기의 사용에 더 능숙할 것으로 추정되나, 디지털 기기에 대한 친숙도 보다는 조리 경험이 사용에 더 큰 영향을 미치는 것으로 보인다. 또한, 비숙련자 집단이 숙련자 집단보다 더 큰 기억 부담을 경험한 것으로 보인다.

조리 비숙련자가 레시피 정보를 어떻게 이해하고 조리에 옮기는지에 대한 이해를 통해 이들의 요구가 좀 더 서비스 제공에 반영되도록 하는 것이 필요할 것으로 보인다. 또한, 조리 숙련도뿐만 아니라 다양한 사용자 특성이 조리 활동과 레시피 서비스 사용에 영향을 줄 것으로 보인다. 연령이 높은 사용자들의 경우 그렇지 않은 사용자들보다 기억에 대한 도움이 더 필요하다는 연구가 존재하고[14], 그 외 디지털 기기에 대한 친숙도, 사용자의 언어 등이 레시피 이해와 조리 행동에 영향을 미칠 수 있다. 궁극적으로 사용자 특성에

대한 정확한 이해를 바탕으로 맞춤형 레시피 제공을 생각해 볼 수도 있다.

조리 지시를 어떻게 수행하는지에 대한 질문을 하는 경우, 조리 지시의 모호성보다는 지시를 행동에 옮기는 것과 관련된 의문에 속한다. 주로 비숙련자들이 이러한 질문을 했는데, 이들은 ‘자르기’, ‘파기’와 같은 동작을 어떻게 실행해야 하는지 경험을 많이 해보지 못했기 때문으로 보인다. 이는 숙련도에 따라서도 정보 요구가 달라지는 것으로 보인다. 연구 참가자 중 S1은 ‘누구’가 자신의 질문에 대답하지 못하는 것을 반복적으로 경험했음에도 반복적으로 질문했다. 이는 평소 사람과 상호작용할 때의 특성을 반영하는 것일 수도 있다.

음성 인식 기술과 자연어처리기술의 발달과 함께 음성을 사용하는 VUI 사용이 점차 증가할 것으로 기대된다. 효과적인 VUI 디자인을 위해서는 음성인식이나 자연어처리의 발달 못지 않게 VUI가 사용자에게 촉발하는 인지적, 정서적, 행동적 요구가 무엇인지를 이해하는 것이 필요한 것으로 보인다. VUI 기반 서비스의 대부분은 아직 GUI 기반 콘텐츠를 음성으로 제시하는 것에 머물러있는데, 동일한 정보가 제시되더라도 GUI 기반인지 또는 VUI 기반인지에 따라서 사용성에 대해 고려가 달라진다. GUI 기반으로 제시되는 경우 정보가 화면에 남아 있을 수 있으나, VUI 기반일 때는 정보는 제시된 직후 사라진다. 따라서 VUI 디자인에서는 사용자가 정보를 필요로 하는 시점을 파악하여 적합한 시점에 정보를 제공하는 것이 중요해질 것으로 보인다. 사용자의 정보처리 특성과 요구에 대한 이해를 바탕으로 서비스를 구축할 때 VUI가 가진 장점을 십분 활용하는 디자인이 가능해질 것으로 보인다. 본 연구는 최소한의 참가자를 통해 사용자 경험을 확인한 것으로 추후 연구로는 더 세부적 분류를 통해 타겟 집단을 선정하고, 참가자를 늘려 연구해 볼 수 있다. 또한, 나이에 대한 변수나 기기 친숙도와 관련하여 인터뷰 및 설문조사를 통해 자세하게 파악할 필요성도 존재한다.

감사의 글

본 연구는 한림대학교의 지원을 받아 수행되었음(HRF-202103-002).

References

- [1] Won Moon Song, Myung Won Kim, “Performance Improvement of Speech Recognition Using Context and Usage Pattern Information”, *Korea Information Processing Society*, Vol.5, No.108, pp.553-560, 2006.
- [2] Harksoo Kim, “Natural Language Processing

- Techniques of Spoken Assistant Systems based on Artificial Intelligence”, *Communications of the Korean Institute of Information Scientists and Engineers*, Vol.35, No.8, pp.9–18, 2017.
- [3] B. Hatscher, M. Luz, L. E. Nacke, N. Elkmann, V. Müller, and C. Hansen, “GazeTap: towards hands-free interaction in the operating room”, *Multimodal Interfaces and Machine Learning for Multimodal Interaction*, pp.243–251, 2017.
- [4] G. Shalene, “For the disabled, smart homes are home sweet home”, *Fortune*, 2015.
- [5] C. Pearl, “Designing Voice User Interfaces”, O’Reilly Media, 2016.
- [6] J. E. Törnros, & A. K. Bolling, “Mobile phone use—effects of handheld and handsfree phones on driving performance”, *Accident Analysis & Prevention*, Vol.37, No.5, pp.902–909, 2005.
- [7] Bobin Kim, Eunjin Lee, Dongwook Kim, Minkyoo Lim, Misook Yu, So yeon Kim, “Proposal for an SNS-based sauce-mixing device for improvement of single consumers cooking environment – Centering on a behavior survey of single people”, *The HCI Society of Korea*, pp.225–229, 2014.
- [8] Jiyu Choi, Gyusang Han, “Structural Analysis of Cooking Recipe Texts – Based on Kimchi Jjigae Recipe”, *The Korean Journal of Community Living Science*, Vol.28, No.2, pp.191–201, 2017.
- [9] Dongwon Kang, Ki-Young Nam, “Design Criteria for Smart Speakers Multi-modal Interface based on VUI and GUI Characteristics”, *Korean Society of Design Science*, pp.6–7, 2019.
- [10] L. Le Bigot, L. Caroux, C. Ros, A. Lacroix, & V. Botherel, “Investigating memory constraints on recall of options in interactive voice response system messages”, *Behaviour & Information Technology*, Vol.32, No.2, pp.106–116, 2013.
- [11] W. Kintsch., & J. Keenan, “Reading rate and retention as a function of the number of propositions in the base structure of sentences”, *Cognitive Psychology*, Vol.5, No.3, pp.257–274, 1973.
- [12] R. Hamada, S. Satoh, J. Okabe, S. Sakai, I. Ide, & H. Tanaka, “Cooking Navi: Assistant for daily cooking in kitchen”, *International Multimedia Conference*, pp.371–374, 2005.
- [13] F. I. Craik, & E. Bialystok, “Planning and task management in older adults: Cooking breakfast”, *Memory & Cognition*, Vol.34, No.6, pp.1236–1249, 2006.
- [14] Q. Tran, G. Calcaterra, & E. Mynatt, “How an older and a younger adult adopted a cooking memory aid”, *Proceedings of HCII: Human Computer Interaction International*, 2005.
- [15] W. G. Chase, & H. A. Simon, “Perception in chess”, *Cognitive Psychology*, Vol.4, No.1, pp.55–81, 1973.
- [16] D. Wickens, “Statistics.” *Ergonomics in Design*, Vol. 6, No. 4, pp.18–22, 1998.
- [17] P. Henstam, “How many participants are needed when usability testing physical products?”, Dissertation, 2018.
- [18] J. Nielsen, & T. K. Landauer, “A mathematical model of the finding of usability problems”, *Conference on Human Factors in Computing Systems*, ACM, pp. 206–213, 1993.
- [19] J. Nielsen, “How many test users in a usability study?”, Nielsen Norman Group, 2012.
- [20] L. Cooke, “Assessing concurrent think-aloud protocol as a usability test method: A technical communication approach”, *IEEE Transactions on Professional Communication*, Vol.53, No.3, pp.202–215, 2010.
- [21] K. A. Ericsson, & H. A. Simon, “Protocol Analysis: Verbal Reports as Data”, MIT Press, 1984.
- [22] K. J. Turley-Ames, & M. M. Whitfield, “Strategy training and working memory task performance”, *Journal of Memory and Language*, Vol.49, No.4, pp.446–468, 2003.
- [23] K. Doman, C. Y. Kuai, T. Takahashi, I. Ide, & H. Murase, “Video CooKing: Towards the synthesis of multimedia cooking recipes”, *In International Conference on Multimedia Modeling*, pp.135–145, 2011.
- [24] E. Nouri, A. Fourney, R. Sim, & R. W. White, “Supporting complex tasks using multiple devices”, *In Proceedings of WSDM’19 Task Intelligence Workshop*, 2019.

〈 저 자 소 개 〉



정 구 민

- 2020년 8월 : 한림대학교 심리학과 (학사) & 융합소프트웨어학과 (공학사)
- 2020년 9월 ~ 현재 : 한림대학교 디지털콘텐츠융합스쿨 석사 과정 중
- 관심분야 : UX, VUI, GUI, 챗봇
- <https://orcid.org/0000-0002-7281-9117>



정 혜 선

- 2001년 ~ 현재 : 한림대학교 심리학과 교수
- 2013년 ~ 현재 : 한림대학교 대학원 디지털콘텐츠융합스쿨 교수
- 관심분야 : 인지과학, 학습과학, CSCL
- <https://orcid.org/0000-0003-3998-3086>