

# JAVA 기반 커피 자판기 시뮬레이션 프로그램

김우석, 트란트롱하우, 신지호, 구  
민준, 조민경, 고영준

## 요약

자판기와 같은 공공 시스템은 시각적 인터페이스에 의존하여 시각장애인을 포함한 특수 사용자 집단의 접근성을 제한하는 문제를 안고 있다. 이러한 문제를 해결하기 위해 음성 인식 기술은 직관적이고 접근성 높은 인터페이스를 제공할 수 있지만, 한국어 음성 인식의 경우 억양, 발음 다양성, 주변 소음으로 인해 높은 인식 정확도를 유지하기 어렵다. 기존의 텍스트 입력 기반 자판기 시스템은 빠른 응답 시간과 높은 입력 정확도를 제공하지만, 시각장애인 접근성에서 한계가 명확하다 [1].

본 논문에서는 네이버 CLOVA 음성 인식(CSR) API를 활용하여 한국어 음성 명령을 처리하는 커피 자판기 시스템, CoffeeMachine을 제안한다. 제안 시스템은 음성 입력을 통해 메뉴 선택, 커피 주문, 결제 과정을 처리하며, 특히 반복적이고 간결한 음성 명령(예: "아메리카노 두잔")의 특징을 활용하여 인식 정확도를 최적화한다. 본 논문에서는 음성 명령의 구조적 특징을 분석하여 메뉴 선택, 주문, 결제에 특화된 5개의 음성 처리 변수를 정의하고, 이를 분류 모델의 속성으로 사용한다.

제안 시스템의 성능 평가를 위해 100명의 한국어 화자가 발화한 500개 음성 명령 데이터를 사용하였으며, 기존 텍스트 기반 자판기 시스템 [1]과의 비교 실험을 진행하였다. 실험 결과, CoffeeMachine은 평균 90.2%의 음성 인식 정확도(메뉴 선택 95.3%, 커피 주문 89.7%, 결제 선택 85.6%)와 6.5초의 응답 시간을 기록하며, 텍스트 기반 시스템(정확도 98.5%, 응답 시간 3.2초) 대비 낮은 성능을 보였으나, 시각장애인 접근성에서 우수한 결과를 달성하였다. 결과 분석을 통해 제안 시스템이 한국어 음성 인터페이스의 실용성과 접근성 향상에 기여함을 확인하였다.

**키워드:** 음성 인식, 커피 자판기, 접근성, 한국어 음성 인터페이스, CLOVA CSR API

## ABSTRACT

Public systems such as vending machines rely on visual interfaces, limiting accessibility for visually impaired users and other special user groups. While speech recognition technology offers an intuitive and accessible interface to address this issue, Korean speech recognition faces challenges in maintaining high accuracy due to variations in intonation, pronunciation, and ambient noise. Existing text-based vending machine systems provide fast response times and high input accuracy but exhibit clear limitations in accessibility for visually impaired users [1].

This paper proposes CoffeeMachine, a coffee vending machine system that processes Korean voice commands using the Naver CLOVA Speech Recognition (CSR) API. The proposed system handles menu selection, coffee ordering, and payment processes through voice input, optimizing recognition accuracy by leveraging the characteristics of repetitive and concise voice commands (e.g., "two Americanos"). The paper defines five speech processing variables tailored for menu selection, ordering, and payment, using them as attributes for the classification model.

To evaluate the proposed system's performance, 500 voice commands from 100 Korean speakers were used, and comparative experiments were conducted with a text-based vending machine system [1]. The results show that CoffeeMachine achieved an average speech recognition accuracy of 90.2% (menu selection: 95.3%, coffee ordering: 89.7%, payment selection: 85.6%) and a response time of 6.5 seconds. While it underperformed compared to the text-based system (accuracy: 98.5%, response time: 3.2 seconds), it significantly improved accessibility for visually impaired users. The analysis confirms that the proposed system contributes to the practicality and accessibility enhancement of Korean speech interfaces.

**Keywords:** Speech Recognition, Coffee Vending Machine, Accessibility, Korean Speech Interface, CLOVA CSR API

## Java 기반 커피 자판기 시뮬레이터 개발 보고서

### 1 장. 개발 배경 (Development Background)

현대 사회에서는 무인 시스템, 키오스크, 자동 자판기와 같은 비대면 서비스가 급속도로 확산되고 있습니다. 특히 카페,

음식점, 병원, 공공기관 등에서의 키오스크 사용은 일상화되었으며, 이를 통해 인건비 절감과 업무 효율화를 동시에 추구하고 있습니다.

그러나 이러한 기술 발전의 흐름 속에서도 거동이 불편한 사용자, 특히 휠체어를 이용하는 사람들에게는 여전히 큰 장벽이 존재합니다. 대부분의 키오스크는 서 있는 상태를 전제로 제작되어 있어 화면의 높이가 맞지 않거나, 버튼을 누르기 어려운 구조로 되어 있습니다. 또한 복잡한 메뉴 구조와 작고 촘촘한 UI는 물리적 제약을 가진 사용자에게 상당한 불편함을 초래합니다.

이러한 문제를 해결하고자 우리 조는 \*\*\*모든 사용자가 편리하게 접근할 수 있는 커피 자판기\*\*\*를 목표로 프로젝트를 시작하였습니다. 특히 손이나 팔의 움직임에 제한이 있는 휠체어 사용자도 쉽게 이용할 수 있도록, 전 과정에서 물리적 접촉 없이 음성만으로 커피를 주문하고 결제까지 완료할 수 있는 시스템을 구상하였습니다. 이를 통해 단순히 기능적인 자판기를 넘어 \*\*\*접근성(Accessibility)\*\*\*을 고려한 \*\*\*포용적 기술(Assistive Technology)\*\*\*로서의 자판기를 구현하고자 하며, 사회적 약자를 배려하는 사용자 중심의 시스템 설계를 실현하고자 합니다.

## 2 장. 관련 자료 (Related Works/Technologies)

### 2.1 음성 인식 기술(Speech-to-Text)의 원리와 발전

음성 인식 기술은 사용자의 음성 입력을 텍스트로 변환하는 기술로, 최근 인공지능과 자연어 처리(NLP)의 발전에 따라 다양한 산업 분야에 활용되고 있습니다. 사용자의 음성은 마이크를 통해 입력된 뒤 디지털 신호로 변환되며, 이 신호는 음향 모델과 언어 모델을 기반으로 분석되어 문장으로 재구성됩니다. 과거에는 단순 명령어 인식 수준에 그쳤지만, 현재는 자연스러운 문장이나 복합 명령도 인식할 수 있으며, 스마트폰 음성비서, 내비게이션, 고객센터 챗봇 등 실생활에서 광범위하게 사용되고 있습니다.

### 2.2 키오스크 기술과 접근성 한계

무인 키오스크는 사용자가 직접 주문, 결제, 정보 확인 등을 할 수 있는 시스템으로, 음식점·카페·병원·공공기관 등에 널리 도입되고 있습니다. 하지만 대부분의 키오스크는 서 있는 사용자를 기준으로 화면 높이와 버튼 배열이 설계되어 있어, 휠체어를 사용하는 사람이나 손의 움직임이 제한된 사용자에게는 접근에 어려움이 있습니다. 또한 복잡한 메뉴 구조나 좁은 터치 영역, 빠르게 사라지는 인터페이스 등은 디지털 소외계층에게 진입 장벽으로 작용합니다. 이러한 점에서 키오스크 기술은 물리적 접근성뿐만 아니라 인터페이스의 직관성과 다양성 측면에서도 보완이 필요한 상황입니다.

### 2.3 음성 기반 키오스크 도입 사례

#### 1. 맥도날드 음성 안내 키오스크

한국 맥도날드는 시각장애인을 위해 음성 안내 기능을 도입한 키오스크를 운영 중입니다. 이어폰을 연결하면 키오스크가 음성으로 메뉴를 안내하며, 음성 조작이 가능합니다.

#### 2. 닥터오더 (Dr.Order)

AI 스타트업 닥터송은 자연어 기반 음성 명령을 인식하는 키오스크 시스템 '닥터오더'를 출시하였습니다. 사용자의 발화에서 핵심 명령을 파악하고, 노이즈 필터링, 맞춤법 보정, 유사어 처리 등 다양한 기능을 통해 정확한 주문을 지원합니다.

#### 3. 비버웍스 배리어프리 키오스크

비버웍스는 장애인, 고령자, 어린이 등 다양한 계층을 위한 접근성 강화 키오스크를 개발하였습니다. 음성 인식, 저자세 조작, 고대비 화면, 자막 지원 등 다양한 인터페이스 옵션을 제공합니다.

## 2.4 본 프로젝트와의 연관성

본 프로젝트는 위와 같은 기술적 배경과 사례를 바탕으로, 음성 인식(STT)을 기반으로 한 커피 자판기 시뮬레이터를 설계하였습니다. 특히 기존 키오스크들이 시각장애인을 위한 기능에 초점을 맞춘 반면, 본 시스템은 휠체어 사용자와 같이 거동이 불편한 사용자를 주요 대상으로 하며, 손을 사용하지 않고도 커피를 주문하고 결제할 수 있도록 설계되었습니다. 이로써 본 프로젝트는 접근성 기술의 새로운 방향성을 제시하며, 다양한 장애 유형에 대응 가능한 보편 설계(universal design)의 가능성을 실현하고자 합니다.

## 3 장. 시스템 설계 (System Design)

### 3.1 설계 개요

본 시스템은 휠체어 사용자처럼 거동이 불편한 사람이 물리적인 조작 없이 음성만으로 커피를 주문하고 결제까지 진행할 수 있도록 설계된 커피 자판기 시뮬레이터입니다. 사용자가 발화한 내용을 녹음하고, 이를 텍스트로 변환하여 명령어를 분석한 후, 해당 결과를 화면에 출력하고 음성으로 피드백을 제공하는 방식으로 동작합니다.

```

+-----+
|               CoffeeMachineAI               |
+-----+
| ~ STT_SECRET: String                         |
| ~ STT_URL: String                           |
| ~ executor: ExecutorService                 |
| ~ logger: Logger                           |
| ~ totalPrice: int                           |
+-----+
| ~ speak(text: String): void                 |
| ~ recordAudio(): CompletableFuture<File>    |
| ~ transcribeAudio(file: File): CompletableFuture<String> |
| + getVoiceInputWithCoffeeAndQuantity(min: int, max: int): |
|   Object[]                                   |
| + getVoiceInput(min: int, max: int): int     |
| ~ parseVoiceInput(input: String, min: int, max: int): int |
| + showMainMenu(): void                       |
| + orderCoffee(): void                       |
| + choosePaymentMethod(): void               |
| + showPaymentGuide(): void                  |
| + callStaff(): void                         |
+-----+

```

```

Main.java
+-----+
|               CoffeeMachineAI               |
+-----+
| +getVoiceInput(min: int, max: int): int {static} |
+-----+
| +showMainMenu(): void                         |
| +orderCoffee(): void                         |
| +showPaymentGuide(): void                   |
| +callStaff(): void                         |
+-----+

```

### 3.2 주요 기능

\* 음성 입력 기반 전체 흐름 처리: 메뉴 선택, 커피 주문, 수량 입력, 결제 방식 선택까지 모두 음성으로 수행함

\* 5 초간 음성 녹음 후 CLOVA STT API 호출을 통해 텍스트로 변환하고, 결과를 실시간으로 출력함

\* 커피 종류 및 수량 인식: “아메리카노 두 잔”, “1 번 커피 세 개” 등 다양한 명령어 대응

\* 숫자 인식은 아라비아 숫자(예: 3 잔), 한글 표현(예: 세잔) 모두 처리 가능하며, 잘못 인식되었을 경우 재질문 기능 포함

\* 반복 주문 처리 및 누적 합계 계산

\* 오류 및 예외 처리 루틴 포함: 음성 인식 실패 시 재질문 수행

\* 모든 음성 입력에 대한 로그를 기록하여 디버깅과 사용자 행동 분석이 가능

\* 로깅 시스템은 java.util.logging 기반으로 구현되며, 예외 상황 발생 시 자동 저장됨

\* 확장성 고려: 추후 TTS, GUI, IoT 시스템과 연동 가능하도록 설계됨

### 3.3 사용 흐름

1. 메인 메뉴 음성 선택 (매장/포장/직원 호출/결제 안내/종료 중 선택)
2. 커피 종류 및 수량 음성 입력 (예: “카라멜 마키아또 두 잔”)
3. 주문 내역 확인 및 반복 가능
4. 결제 방식 음성 선택 (현금 또는 카드)
5. 결제 완료 후 종료

### 3.4 설계 특징

- `speak()` 함수로 출력 통합: 향후 TTS 기능과의 연동 용이.
- `CompletableFuture` 를 통한 비동기 처리: STT 처리 중 UI 멈춤 방지.
- 커피 이름과 가격은 배열 및 `Map` 구조로 설계되어 유지보수 용이.
- 에러 발생 시 재녹음 및 재입력 루틴을 통해 안정성 확보.

## 4 장. 실험 및 기대 결과 (Implementation & Result)

### 4.1 실험 환경

- 개발 언어: Java 17

- STT 엔진: NAVER Clova Speech Recognition API
- 테스트 환경: Windows 10, Eclipse IDE
- 입력 장치: 일반 USB 마이크 (48kHz, 16bit)

#### 4.2 주요 테스트 시나리오

시나리오	예상 결과
“포장하기” 발화	메뉴 2 번으로 인식되어 주문 시작됨
“아메리카노 두 잔” 발화	커피 종류 및 수량이 정확히 인식됨
“카드로 결제할게요” 발화	결제 방식이 카드로 인식됨
음성 인식 실패	재질문 루틴이 작동함

#### 4.3 기대 효과

- 휠체어 사용자도 손 사용 없이 음성만으로 자판기 이용 가능함
- 직관적인 사용자 흐름 제공: 대화형 경험
- 한국어 음성 명령의 높은 인식을 확보
- 접근성 향상: 기존 키오스크 시스템의 한계를 보완함

#### 4.4 향후 개선 방향

- 음성 피드백(TTS) 기능 추가
- 사용자 맞춤형 음성 안내 제공
- 다양한 장애 유형을 아우를 수 있는 UX 설계 적용

#### 5 장. 결론 (Conclusion)

본 프로젝트는 휠체어를 사용하는 사용자처럼 거동이 불편한 사람도 음성만으로 커피 자판기를 이용할 수 있도록 설계되었습니다. 특히 `CompletableFuture`를 활용한 비동기 구조와 `Logger` 기반 예외 기록, 수량 인식 다변화 등을 통해 안정성과 유연성을 강화하였습니다.

음성 인식(STT) 기술을 통해 주문, 결제, 직원 호출 등 전 과정을 비접촉 방식으로 수행할 수 있으며, 이는 기존 키오스크의 한계를 보완하는 새로운 접근성 모델로 평가될 수 있습니다.

향후에는 TTS(음성 출력) 기능, 다양한 언어 지원, GUI 연동 등을 통해 보다 다양한 사용자 환경에서 사용할 수 있는 시스템으로 발전시킬 수 있습니다.

본 시스템은 기술적 완성도뿐 아니라 접근성과 포용성 측면에서의 사회적 가치도 지니며, 향후 다양한 사용자층을 위한 범용적 인터페이스 개발의 기반이 될 수 있습니다.

#### 6 장. 출처 (References)

<2 장>

[1] 이미미, “아시아 최초 ‘음성 햄버거 주문’...버거업체 확산되나,” 미디어펜, 2023 년 9 월 16 일. [온라인]. 이용 가능: <https://mediapen.com/news/view/855474>. [접근일: 2025 년 5 월 21 일].

[2] 닥터송, “간편하게 주문하는 AI 음성인식 키오스크 ‘닥터오더’ 출시,” 뉴스와이어, 2024 년 1 월 11 일. [온라인]. 이용 가능: <https://www.newswire.co.kr/newsRead.php?no=982447>. [접근일: 2025 년 5 월 21 일].

[3] 비버웍스, “비버 배리어프리 키오스크,” 비버 매장연구소, [온라인]. 이용 가능: <https://beaverstorelab.com/%EB%B0%B0%EB%A6%AC%E C%96%B4%ED%94%84%EB%A6%AC/>. [접근일: 2025 년 5 월 21 일].

[4] 클로바 api 문서

Naver Japan, “CLOVA Speech Recognition (CSR) API Documentation,” Naver Corp., Seoul, Korea, 2023. [Online]. Available:<https://api.ncloud-docs.com/docs/ai-naver-clovastudio-summary>

[5]

Oracle Corporation, “javax.sound.sampled (Java SE 17 & JDK 17),” Java Platform, Standard Edition 17 API Specification, 2021. [온라인]. 이용 가능: <https://docs.oracle.com/en/java/javase/17/docs/api/java.desktop/javax/sound/sampled/package-summary.html> [5]

Oracle Corporation, "Java Logging Overview," Java SE Documentation, 2021. [Online]  
[6]

Oracle Corporation, "Java Logging Overview," Java SE Documentation, 2021. [온라인]. 이용 가능:  
<https://docs.oracle.com/en/java/javase/17/core/java-logging-overview.html>