

# STT 기반 카카오톡 이벤트를 TTS 음성으로 안내하는 IoT 어플 설계

2017104033 최세현

## 요약

위드코로나 시대로 접어들면서 많은 사람들이 SNS를 통해 지인들과 소통한다. 하지만 운전 중이나 보행 중에 텍스트 위주의 SNS를 확인은 사고 발생 확률을 높아지게 만들었다. 이에 대해 본 논문에서는 블루투스와 TTS(Text-to-Speech)를 사용하여 카카오톡 이벤트를 음성으로 안내하는 IoT 어플을 설계한다. 안드로이드 스튜디오에서 사용가능한 블루투스 api와 구글 tts api를 사용하여 헤드셋이나 이어폰을 사용해서도 알림을 음성으로 안내받을 수 있는 어플을 구현한다.

## 1. 서론

### 1.1. 연구배경

위드 코로나 시대로 접어들며 비대면 소통과 디지털 커뮤니케이션이 익숙해 지고 있는 가운데 많은 사람들이 SNS를 통해 지인들과 소통한다. SNS에서 볼 수 있는 텍스트 위주의 정보들을 일반적인 상황에서는 유용하지만 운전 중이나 보행 중에는 확인하기에 불편하다. 다른 무언가를 하면서 동시에 화면에 집중하는 것은 쉽지 않으며 운전 중에는 전방주시태만으로 인해 사고를 초래하기도

한다. 전방주시태만은 운전 중 스마트폰 메시지 확인, SNS와 연관이 깊다고 분석됐으며, 이로 인한 사고발생 확률이 23배까지 증가할 수 있다. 이러한 이유들로 인해 음성기술 AI 시장이 고속 성장하고 있다. 음성기술 AI는 스마트폰이나, 자율주행 차량, 각종 전자기기를 비롯한 다양한 산업군에서 빠른 속도로 진화하며 생활 속에 녹아들고 있다.

본 연구에서는 사람들이 많이 사용하는 SNS인 카카오톡과 음성기술 시장에서 큰 비중을 차지하고 있는 음성합성 기술을 활용하여 카카오톡 메시지를 음성으로 확인할 수 있도록 하여 이동 중에 스마트폰 화면을 확인하여 생길 수 있는 사고를 방지하는 IoT 어플리케이션을 구현하고자 한다.

## 1.2. 연구목표

본 연구에서는 STT기반의 카카오톡의 다양한 이벤트와 음성합성 기술을 활용한 시스템을 구현할 것이다. 사용자가 스마트폰을 보고 있지 않아도 카카오톡으로 수신되는 메시지를 음성합성을 통해 사용자에게 음성으로 알려줌으로써 스마트폰을 볼 수 없는 상황에서도 편하게 사용할 수 있도록 구현하고자 한다. 또한, 블루투스 연결을 가능하게 함으로써 블루투스 스피커나 이어폰, 헤드셋을 통해서도 음성을 들을 수 있도록 할 것이다.

카카오톡으로 전송된 메시지를 어플리케이션으로 전달받기 위해 브로드캐스트 수신자를 등록하고 sms 수신 권한을 추가하여 sms를 수신 가능하도록 한다. 음성합성 기술인 TTS(Text to Speech)를 통해 전달받은 메시지를 음성으로 전달하며 음성의 속도와 크기를 사용자가 조절할 수 있다. 읽어오는 텍스트는 내용 뿐만 아니라 발신자와 발신시각을 사용자의 선택에 따라 읽어올 수 있다. 어플리케이션을 실행시키면 기능이 무조건 동작하는 것이 아니라 사용자의 필요에 따라 사용여부를 선택할 수 있도록 구현할 것이다.

## 2. 관련 연구

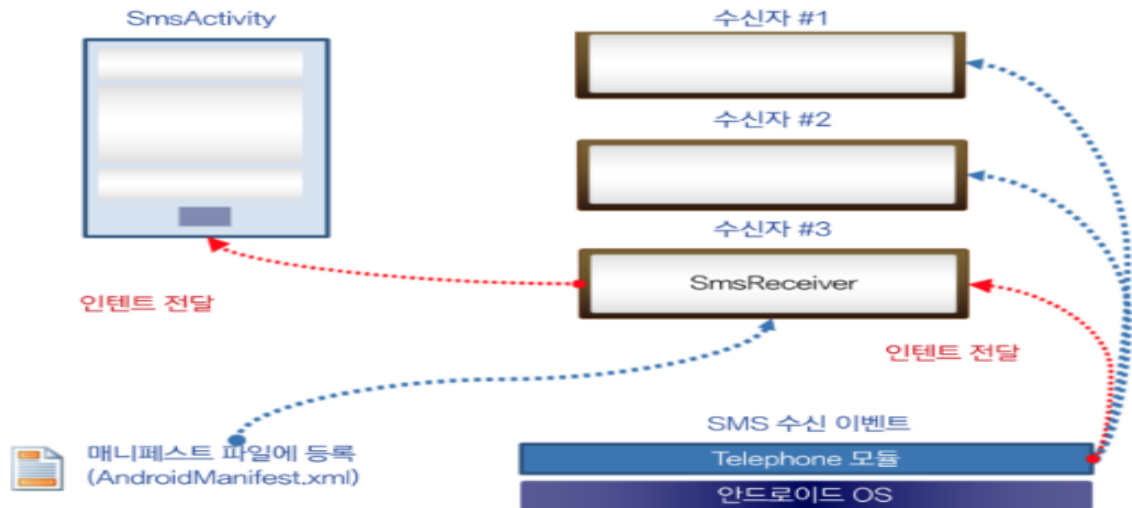
### 2.1. 블루투스

블루투스(Bluetooth)는 스마트폰, 노트북, 이어폰·헤드폰 등의 휴대기기를 서로 연결해 정보를 교환하는 근거리 무선 기술 표준을 의미하는 것으로, 주로 10미터 안팎의 초단거리에서 저전력 무선 연결이 필요할 때 쓰인다. 블루투스를 이용한 무선 네트워크는 언제 어디서나 모든 정보기기 간의 자유로운 데이터 교환이 이뤄질 수 있다.

Android 플랫폼은 블루투스 네트워크 스택에 대한 지원을 포함하므로 기기가 다른 블루투스 기기와 데이터를 무선으로 교환할 수 있다. 애플리케이션 프레임워크는 Android Bluetooth API를 통해 블루투스 기능에 대한 액세스 권한을 제공한다. 애플리케이션은 이러한 API를 통해 다른 블루투스 기기에 무선으로 연결하여 지점 간 무선 기능 및 다중 지점 무선 기능을 활성화할 수 있다.

### 2.2. Broadcast Receiver

안드로이드에서 브로드캐스팅(Broadcasting)이란 메시지를 여러 객체에 전달하는 것을 말한다. 여러 앱 구성 요소에 메시지를 전달할 때 브로드캐스팅을 사용한다. 개발하는 어플에서 브로드캐스팅 메시지를 받으려면 Broadcast Receiver를 만들어 앱에 등록해야한다. 기기 안에서 동작하는 카카오톡으로부터 메시지를 받기 위해 개발하는 어플에 Broadcast Receiver를 등록하면 카카오톡의 메시지가 개발하고 있는 어플로 전달된다.



[그림 1] 브로드캐스트 수신자의 동작방식

단말에서는 다른 사람으로부터 SMS 문자를 받았을 때 텔레포니(Telephony) 모듈이 처리하도록 한다. 이렇게 처리된 정보는 인텐트에 담겨 브로드캐스팅 방식으로 다른 앱에 전달된다. 브로드캐스트 수신자를 manifest 파일에 등록을 해놓은 상태라면 개발하고 있는 앱으로 인텐트를 전달할 수 있다. 브로드캐스트 수신자를 사용하면 어플이 실행되어 있지 않아도 원하는 브로드캐스트 메시지가 도착하면 다른 어플을 실행하고 있는 도중에도 인텐트 안에 들어있는 메시지를 받아볼 수 있다.

### 2.3. NotificationListenerService

카카오톡의 이벤트를 읽어오기 위해서는 카카오톡에서 보내는 알림의 정보를 읽어야 한다. NotificationListenerService는 카카오톡의 Notification에 대한 정보를 받을 수 있는 객체이다. 정보를 받으려면 NotificationListenerService를 상속받은 새로운 서비스를 구현하여야 한다. 구현해야 할 클래스는 Service이기 때문에 AndroidManifest에 등록 후 사용해야 한다. 안드로이드 프레임워크만 접근할 수 있도록 서비스의 permission에 BIND\_NOTIFICATION\_LISTENER\_SERVICE를 설정하고, 그림 2와 같은 인텐트필터를 설정해야 한다.

```

<service
    android:name=".NotificationListener"
    android:permission="android.permission.BIND_NOTIFICATION_LISTENER_SERVICE"
    android:exported="true">
    <intent-filter>
        <action android:name="android.service.notification.NotificationListenerService"/>
    </intent-filter>
</service>

```

[그림 2] intent-filter

intent-filter를 설정하는 이유는 저 action으로 NotificationManager가 단말에 등록된 서비스를 찾는데 사용하기 때문이다. 만약 인텐트필터를 저렇게 등록하지 않으면 notification을 받지 못한다.

구현해야 할 클래스의 주요 메서드는 다음과 같다. onListenerConnected() : 서비스가 연결(실행)되었을 때 콜백되는 메서드이다. 앱이 최초에 실행되거나, 알림 접근 권한을 획득하게 되면 호출된다. onListenerDisconnected() : 서비스가 연결 해제되었을 때 콜백되는 메서드이다. onNotificationPosted(sbn: StatusBarNotification) : 알림이 왔을 때 콜백되는 메서드이다. StatusBarNotification 객체에 해당 알림에 대한 정보가 들어있으며, 객체 내부 getNotification() 메서드로 알림에 대한 정보를 획득할 수 있다. onNotificationRemoved(dbn: StatusBarNotification) : 알림이 삭제되었을 때 콜백되는 메서드이다.

## 2.4. Text-to-Speech

음성합성(TTS)이란 입력문장(Text)을 사람이 청취할 수 있는 음성(Speech)으로 변환시키는 기술이라는 뜻으로 'Text To Speech'의 약자인 'TTS'라는 용어로 오늘날 더 많이 사용되고 있다. 이 때 입력문장은 데이터베이스의 자료일 수도 있고, 교통안내와 같은 문장일 수도 있다. 입력된 문장을 음성으로 변환시키는 방법은 음성의 기본 단위를 먼저 데이터베이스로 구축하여 놓은 뒤 이들을 조합하여 다시 구성하는 제한 어휘 음성합성과 어휘에 제한 없이 어떤 문장이라도 합성하는 무제한 어휘 음성합성 이 두 분류로 나뉜다. 이와 같은 방식을 TTS라고 한다. 이는 억양의 고저나 장단 등 운율에 관한 정보를 결정하는 기능을 갖는다. TTS의 등장은 시각장애인, 외국인, 아이, 노인과 같이

문자를 원활하게 읽지 못하는 상황에 처한 사람들에게 큰 도움을 줄 것이라는 기대를 받았으며, 오늘날 다양한 분야에서 사용되며 주목받는 기술로 부상하고 있다.

본 설계에서는 Google Text-to-Speech를 사용하여 어플리케이션을 구현한다. Google Text-to-Speech는 구글에서 제공하는 텍스트를 음성으로 변환시켜주는 AI기술을 기반으로 API로 40개 이상의 언어 및 방언이 지원된다.

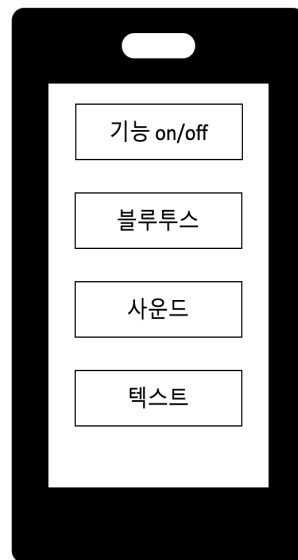
Text-to-Speech는 텍스트 또는 음성 합성 마크업 언어(SSML) 입력을 MP3 또는 LINEAR16(WAV 파일에서 사용되는 인코딩)과 같은 오디오 데이터로 변환한다. Text-to-Speech는 사용자에게 음성 오디오를 재생하는 모든 애플리케이션에 적합하다. 임의의 문자열, 단어, 문장을 동일한 내용을 말하는 사람의 소리로 변환할 수 있다.

텍스트 입력을 오디오 데이터로 변환하는 프로세스를 합성이라고 하며 합성 출력을 합성 음성이라고 한다. Text-to-Speech는 원시 텍스트나 SSML 형식의 데이터와 같은 두 유형의 입력을 사용한다. 음성 합성 프로세스는 원시 오디오 데이터를 base64로 인코딩된 문자열로 생성한다. base64로 인코딩된 문자열을 오디오 파일로 디코딩해야 애플리케이션에서 재생할 수 있다. 또한 대부분의 플랫폼과 운영체제에는 base64 텍스트를 재생 가능한 미디어 파일로 디코딩하는 도구가 있다.

Text-to-Speech는 기존의 다른 합성 음성과 함께 WaveNet에서 생성된 프리미엄 음성도 제공한다. 사용자는 Wavenet에서 생성된 음성이 다른 합성 음성보다 더 따뜻하고 인간과 비슷하다는 것을 알게된다. WaveNet 음성의 주요 차별점은 음성을 생성하는 데 사용되는 WaveNet모델이다. WaveNet 모델은 실제 인간이 말한 원시 오디오 샘플을 사용하여 학습되었다. 따라서 이러한 모델은 음절, 음소, 단어에 인간과 매우 비슷한 강세와 어조를 사용하여 합성 음성을 생성한다.

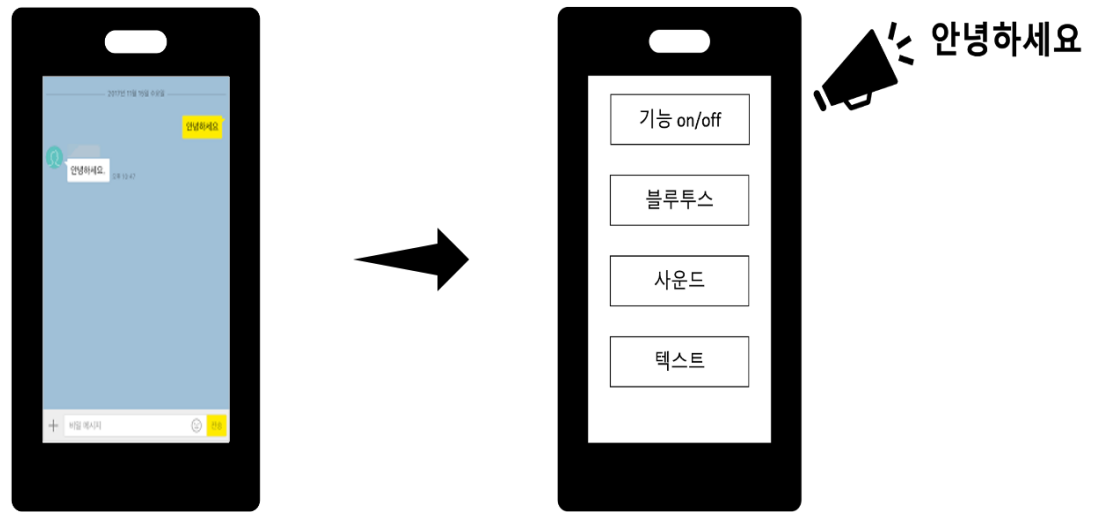
### 3. 프로젝트 내용

#### 3.1. 시나리오



[그림 3] 메인 화면

[그림 3]은 애플리케이션의 메인 화면이다. 메인에서는 이 기능의 사용을 정하는 on/off 스위치를 통해 사용자에게 사용여부를 물어본다. 사용을 원한다면 블루투스와 사운드, 텍스트 설정을 가능하게 한다. 메인 화면에서 블루투스 부분은 블루투스의 사용여부와 연결된 기기의 정보를 알려주고 블루투스 설정 화면으로 들어갈 수 있게 해준다. 사운드 부분에서는 음성합성되는 소리의 볼륨을 조절할 수 있고 속도를 0.5배 부터 2배까지 선택할 수 있다. 그리고 카카오톡에서 오는 메시지 알림을 진동으로 받을지 여부를 선택할 수 있다. 텍스트 부분에서는 메시지의 내용 뿐만 아니라 발신자와 발신 시간을 사용자의 선택에 의해 얻을 수 있다.



[그림 4] 카카오톡 메시지를 받았을 때

카카오톡 메시지를 받으면 메인화면에서 설정했던 사운드와 텍스트 방식을 사용하여 메시지를 음성으로 알려준다. 메시지를 받으면 기본은 속도 1.0배에 메시지 내용만 읽어주지만 사용자가 지정한 설정에 따라 읽어주는 속도를 다르게 할 수 있으며 발신자와 발신 시간까지 읽어줄 수 있다. 또한 블루투스 스피커나 이어폰 등을 연결한 경우에는 음성이 연결한 장치에서 나온다.

### 3.2. 요구사항

블루투스 연결은 Android Bluetooth API를 사용하여 블루투스 기능에 대한 액세스를 가능하게 만든다. 어플 내에서 블루투스의 on/off를 가능하게 하고 연결된 기기의 정보를 출력할 수 있게 한다.



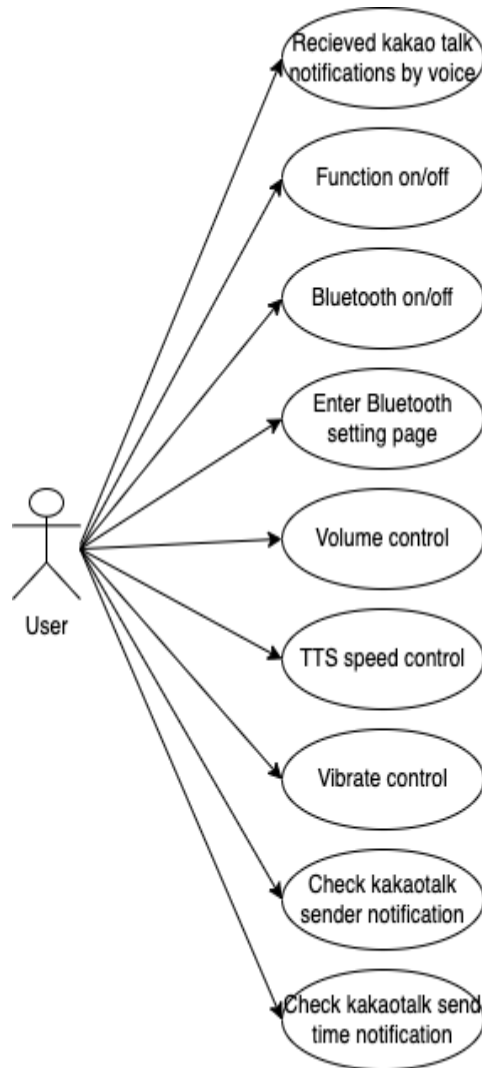
카카오톡 메시지 알림을 받고 그 메시지의 내용과 발신자와 발신시간을 얻기 위해 브로드캐스트 수신자를 등록해야 한다. 그리고 그 정보를 얻어와 화면에 보여주는 것이 아니라 그 정보를 음성으로 전달해야 한다.

음성으로 전달하기 위해서는 TTS, 즉 음성합성 기술을 사용해야 하고 안드로이드에서 구글 TTS API를 내장하고 있기 때문에 그것을 통해 구현한다.

Expandable RecyclerView와 ScrollerView를 사용하여 메인 화면을 구성할 때, 리사이클러뷰를 접고 펼칠 수 있도록 하여 사용자가 직관적으로 사용할 수 있도록 만든다.

### 3.3. 시스템 설계

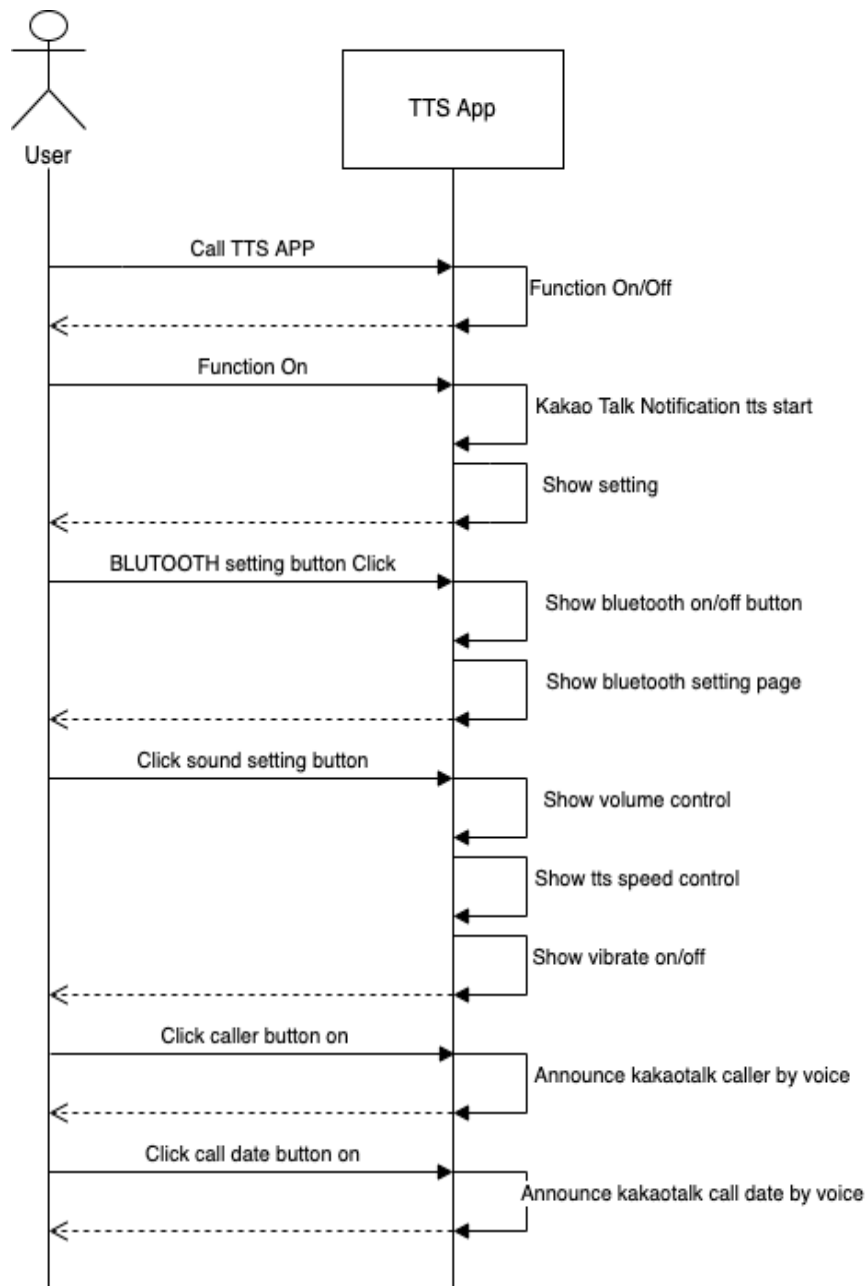
#### 3.3.1. Use Case Diagram



[그림 5] Use Case Diagram

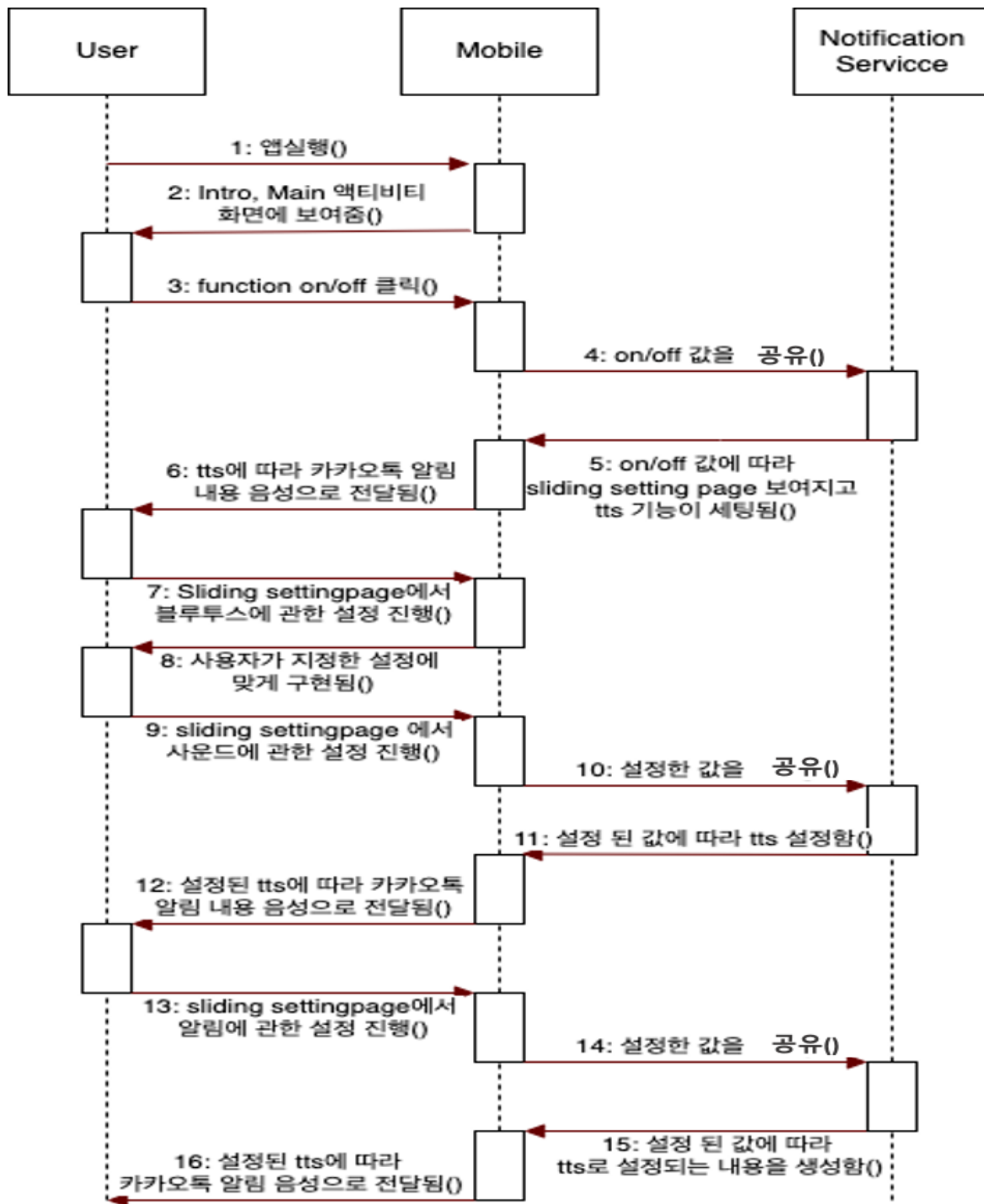
유스케이스는 사용자가 어플을 통해 사용할 수 있는 기능을 나열하였다. 기능의 사용 여부와 블루투스 on/off, 음성, tts에 관한 설정까지 사용자에게 의해 지정된다.

### 3.3.2. Sequence Diagram



[그림 6] 시스템 순서도

추출된 유스케이스 별로 각각의 유스케이스 내의 사건 흐름을 모델링하기 위해 시스템 순서도를 작성하였다. 그리고 시스템 순서도를 기반으로 좀 더 자세하게 흐름을 모델링을 한 sequence diagram을 [그림 7]에 작성하였다.



[그림 7] Sequence Diagram

어플에서 제공하고 있는 기능의 Sequence Diagram을 자세하게 보여주고 있다. Mobile 기기는 사용자의 설정을 기다리고 있으며, 사용자가 지정한 설정 값을 NotificationListenerService에 전달하여 카카오톡 알림을 음성으로 출력해줄 때 사용하는 TTS 설정에 관여한다.

### 3.3.3. Application

#### 3.3.3.1. 블루투스 api

어플리케이션에서 블루투스 기능을 사용하려면 두 개의 권한을 선언해야 한다. 그 중 첫 번째는 BLUETOOTH이다. 이 권한은 연결 요청, 연결 수락 및 데이터 전송과 같은 블루투스 통신을 수행하는데 필요하다. 두 번째로 선언해야 하는 권한은 ACCESS\_FINE\_LOCATION이다. 앱에 이 권한이 필요한 이유는 블루투스 스캔을 사용하여 사용자 위치에 대한 정보를 수집할 수 있기 때문이다. 앱에서 기기 검색을 시작하거나 블루투스 설정을 조작하려면 BLUETOOTH\_ADMIN 권한도 선언해야 한다.

어플리케이션이 블루투스를 사용하여 통신하려면 블루투스가 기기에서 지원되는지 확인하고, 지원되는 경우 활성화해야 한다. 블루투스가 지원되지만 비활성화된 경우에는 BluetoothAdapter를 사용하여 블루투스를 활성화하도록 요청한다. BluetoothAdapter를 사용하면 기기 검색을 통해서 또는 페어링된 기기 목록을 쿼리하여 원격 블루투스 기기를 찾을 수 있다. 본 설계에서는 블루투스를 switch 버튼을 통해 on/off를 조절하고 블루투스를 통해 연결된 기기를 출력한다. 또한 블루투스 설정 화면에 쉽게 들어갈 수 있게 구현한다.

#### 3.3.3.2. TTS api

안드로이드에서는 구글 SpeechToText, TextToSpeech 기술인 stt, tts를 내장 API로 지원하고 있다. 별도의 설치 없이 gradle에 특별한 세팅 없이 사용가능 하다. TextToSpeech 인스턴스는 초기화가 완료된 후에만 텍스트를 합성하는 데 사용할 수 있다. TextToSpeech.OnInitListener을 통해 초기화를 완료한다. TextToSpeech 인스턴스 사용이 끝나면 shutdown() 메서드를 호출하여

TextToSpeech 엔진에서 사용하는 기본 리소스를 해제해야 한다. TTS를 사용하는 어플은 manifest에 TextToSpeech.Engine.INTENT\_ACTION\_TTS\_SERVICE를 선언해야 한다.

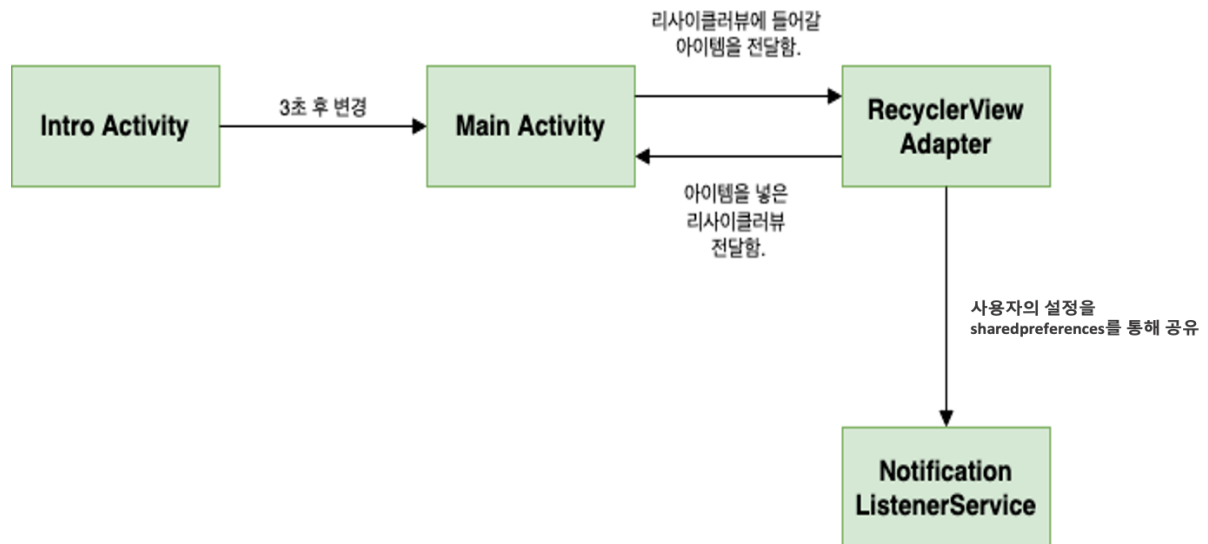
본 설계에서는 카카오톡 이벤트를 읽어주는 언어는 한국어로 설정하기 위해 TextToSpeech 인스턴스를 초기화 할 때, tts.setLanguage(Locale.Korean)로 설정해준다. 또한, spinner 를 통해 읽기 속도를 선택하고 tts.setSpeechRate(); 사용하여 선택된 속도로 알림이 읽어질 수 있도록 설정한다. 텍스트 부분에서는 메시지의 내용 뿐만 아니라 발신자와 발신 시간을 사용자의 선택에 의해 얻을 수 있다. tts.speak()를 통해 카카오톡의 알림이 읽어지도록 구현한다.

#### 3.3.3.3. application 구현



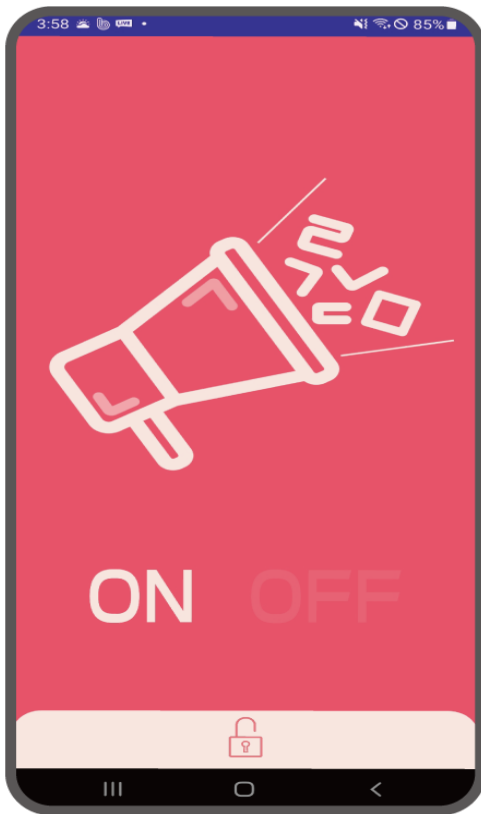
[그림 8] application 시작 화면

[그림8]은 어플리케이션을 실행하면 바로 보이는 화면으로 Introactivity.java를 통해 구현한다. 약 3초정도 보여지며 어플리케이션의 로고와 name을 담고 있다.

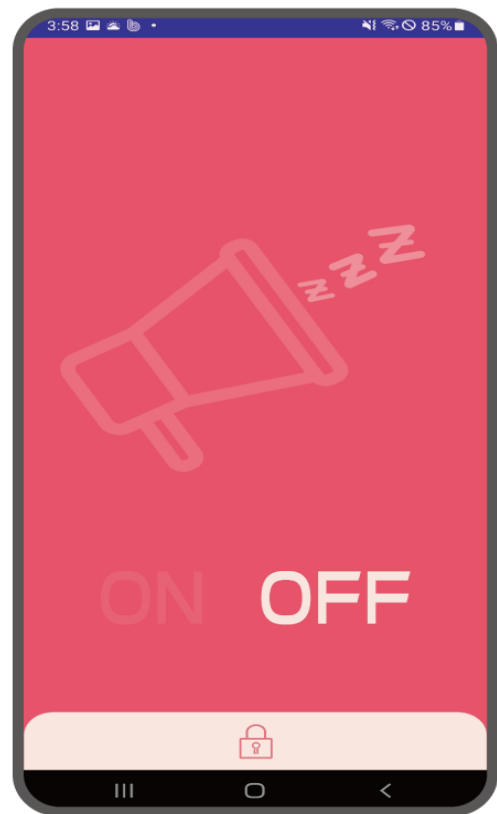


[그림 9] activity 흐름도

어플에서 설정화면을 recyclerview로 나타내기 위해 recyclerview adapter를 사용하여 MainActivity에서 저장할 아이템을 어댑터에 전달하게 하였다. 전달된 아이템을 어댑터에서 리사이클러뷰에 넣고 MainActivity에서 아이템이 든 리사이클러뷰를 볼 수 있도록 하였다. 리사이클러뷰에서 사용자가 선택한 설정들은 sharedPreferences에 공유되어 NotificationListenerService에서 알림에 관한 설정과 TTS에 관한 설정을 할 수 있게 하였다.



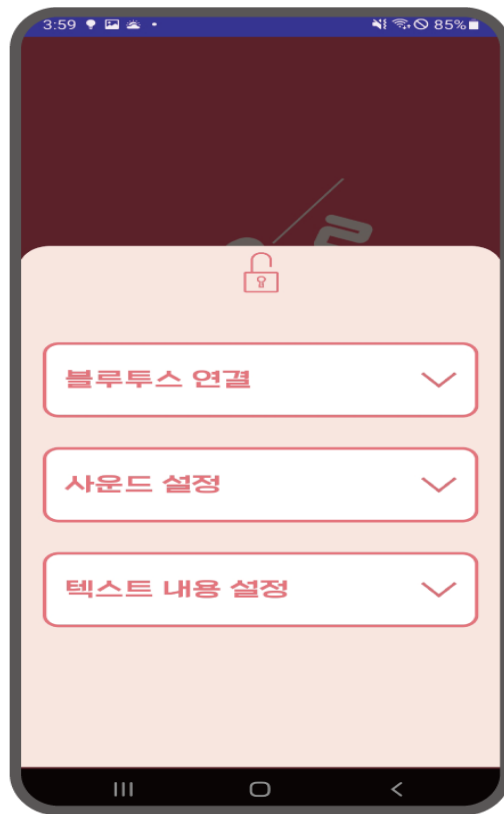
[그림 10] function on



[그림 11] function off

[그림 10]처럼 function on을 누르면 설정을 할 수 있는 슬라이딩 패널이 자물쇠가 열려 [그림 12]에서처럼 슬라이딩 패널을 위로 올려서 사용자가 원하는 설정을 할 수 있다. 또한 카카오톡 알림을 음성으로 제공해주는 기능이 작동하여 카카오톡 알림이 오면 음성으로 안내해준다. [그림 11]은 function off를 눌렀기 때문에 슬라이딩 패널에 자물쇠가 잠겨있고 위로 올리려고 해도 올려지지 않는다. 즉, 기능을 사용하지 않는다고 했기 때문에 사용자가 설정에 들어갈 수 없고, 카카오톡 알림이 와도 음성으로 안내해주지 않는다.





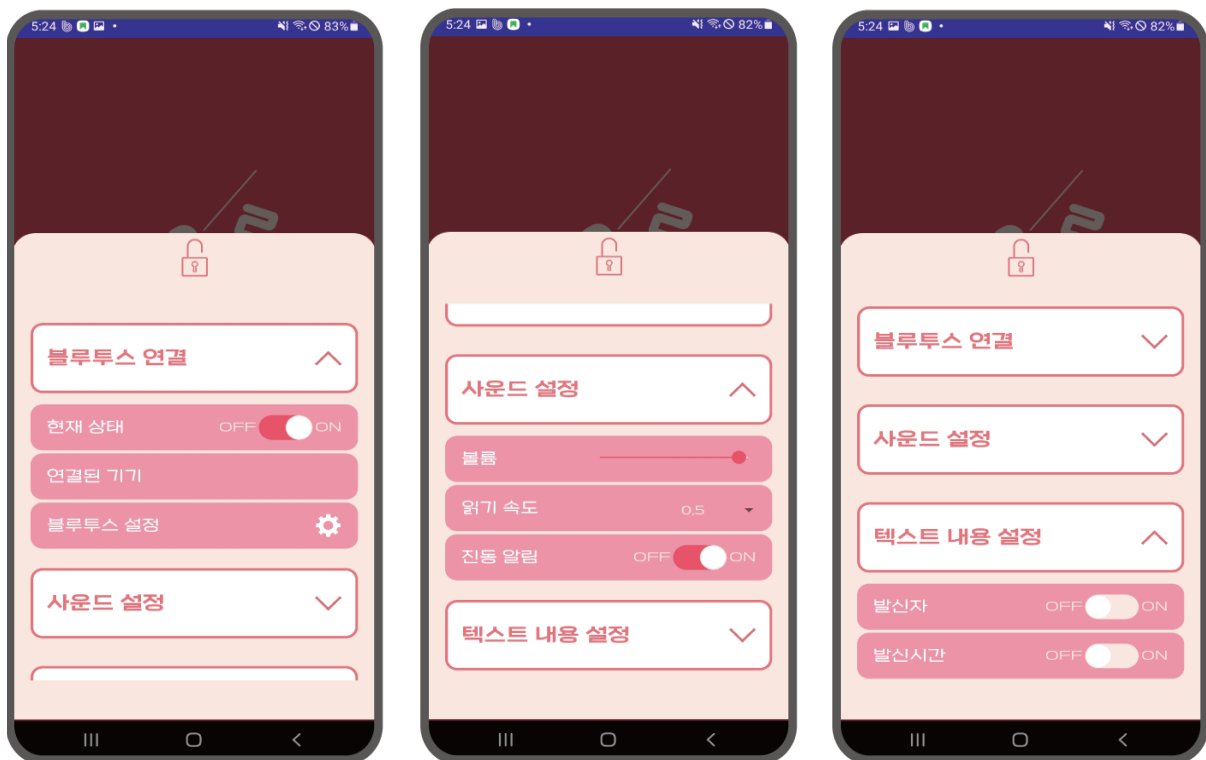
[그림 12] 슬라이딩 패널

function을 on 하면 슬라이딩 패널을 열 수 있고 어플 기능을 설정할 수 있다. 리사이클러뷰를 통해 만들었고 리사이클러뷰를 열고 접을 수 있도록 구현했기 때문에 세부 설정들은 아래 화살표를 누르면 세부 설정에 관한 리사이클러뷰가 열린다.

## 4. 프로젝트 결과

### 4.1. 안드로이드 구현 결과

[그림 13]에서 처럼 블루투스 설정과 사운드 설정 그리고 알림 내용 설정을 사용자가 고를 수 있다. 블루투스 on/off를 통해 어플에서 블루투스를 켜고 끌 수 있으며 연결된 블루투스 기기가 있으면 연결된 기기 이름을 출력해 준다. 볼륨은 seekbar를 통해 조절이 가능하며 tts의 음성 속도는 4단계로 0.5, 1.0, 1.5, 2.0으로 사용자가 원하는 빠르기로 설정 가능하도록 구현했다. 카카오톡 알림이 올 때 사용자의 필요에 따라 진동으로도 올 수 있도록 했다. 어플을 시작해서 기능을 on하고 난 후 카카오톡 알림이 오면 기본으로 카카오톡의 메시지의 내용만 음성으로 알려주는데 사용자가 원하면 발신자나 발신 시간을 음성으로 알려주도록 구현했다. 어플리케이션을 통해 시연한 동영상은 최종보고서와 함께 업로드하였다.



[그림 13] 안드로이드 구현

## 5. 결론

### 5.1. 기대효과

블루투스나 TTS를 이용하여 카카오톡의 이벤트를 음성으로 안내받게 함으로써 운전을 하고 있는 중이나 보행 중에 카카오톡 알림을 편하게 확인 가능한 시스템 설계를 제안하였다. 이를 통해 운전 중이나 보행 중에 중요한 알림을 화면을 보지 않고도 음성으로 알 수 있기 때문에 편리하게 사용 가능하며, 텍스트 알림을 확인하려다 발생하는 사고가 감소할 것으로 기대된다.

블루투스 연결이 가능하기 때문에 이어폰이나 헤드셋을 착용하고 있어도 음성을 들을 수 있고 음성의 속도와 볼륨이 조정 가능해서 사용자의 필요에 맞출 수 있다.

최근에는 음성인식 기술을 탑재한 AI 인공지능 스피커를 많이 사용하는데 그걸 구매하지 않고도 스마트폰의 어플만으로도 메시지를 읽어주기 때문에 장비를 구매해야 하는 비용적인 측면을 감소시킬 수 있다.

### 5.2 추후 연구 방향

향후 연구로는 사람들이 SNS를 읽다가만 사고가 나는 것이 아니라 작성을 하거나 검색을 하다가 사고를 초래하는 경우도 많기 때문에 음성으로 SNS에 메시지를 보낼 수 있도록 할 것이다. STT API와 KakaoTalk API를 활용하여 메시지를 음성으로 읽어오는 것 뿐만 아니라 음성으로 말하면 텍스트로 변환하여 카카오톡에 메시지를 보낼 수 있도록 개발할 계획이다.

또한 이것이 카카오톡에서 뿐만 아니라 다른 SNS에서도 메시지를 음성으로 읽어오고 음성으로 메시지를 보낼 수 있도록 개발하여 운전 중이나 보행 중에 SNS 메시지를 보거나 보내다가 발생하는 사고를 감소시킬 수 있도록 개발할 것이다.

## 6. 참고문헌

[1]정재곤, “Do it 안드로이드 앱 프로그래밍”, 이지스퍼블리싱, 2020.03.09

[2]이세환, 김봉현, 김승연, “TTS를 이용한 음성지원 일정관리 시스템의 구현”, 한국정보처리학회, 2006년

[3]Google Cloud, “Cloud Text-to-Speech 기본사항”,  
[https://cloud.google.com/text-to-speech/docs/basics#wavenet\\_voices](https://cloud.google.com/text-to-speech/docs/basics#wavenet_voices)

[4]Android Developers, “블루투스 개요”, <https://developer.android.com/guide/topics/connectivity/bluetooth?hl=ko>

[5]brunch, “GCP : Google Text to Speech”, <https://brunch.co.kr/@e9c7009de84443b/135>

[6]개발노트, “다른 앱의 알림(Notification) 내용 가져오는 방법(알림 접근)”,  
<https://kobbi-reply.tistory.com/17>

[7]신동훈, “가장 위험한 운전자는? 설문 결과 “1위 음주운전, 2위 운전 중 스마트폰 사용””, CCTV 뉴스, 2018년 2월 7일