**Speech Translation SDK**

**User's Guide for**

**Sample Application**

**(Android)**

Version1. 1

Table of Contents

[1. About This Application 3](#_Toc374973807)

[1.1. Overview 3](#_Toc374973808)

[1.2. Functions 3](#_Toc374973809)

[1.3. Released materials 3](#_Toc374973810)

[2. Environments 4](#_Toc374973811)

[2.1. Operating environment 4](#_Toc374973812)

[2.2. Development environment 4](#_Toc374973813)

[3. Screen View 5](#_Toc374973814)

[3.1. Language-selecting list 6](#_Toc374973815)

[3.2. Result display area 6](#_Toc374973816)

[3.3. Level meter 6](#_Toc374973817)

[4. Setting and Logs of Application 7](#_Toc374973818)

[4.1. Setting of the server to communicate 7](#_Toc374973819)

[4.2. Logs 7](#_Toc374973820)

[5. Screen Transition 8](#_Toc374973821)

[5.1. Normal operation 8](#_Toc374973822)

[5.2. Abnormal operation 10](#_Toc374973823)

Update History

|  |  |  |
| --- | --- | --- |
| Version | Release Date | Features |
| 1.0 | Dec. 20, 2013 | Released |
| 1.1 | Oct. 31, 2014 | Added ADPCM encoding for input speech and ADPCM decoding process for output speech |

# About This Application

This document is a tutorial to use the sample application (Android version) of the open-source Speech Translation SDK (Software Development Kit) for based on the MCML system.

## Overview

This application is a sample application for speech recognition, translation, and speech synthesis, based on the MCML system.

## Functions

The sample application has the following functions:

* The application transmits and receives MCML requests and responses of the MCML system (server) necessary for speech recognition, translation, and speech synthesis.
* The application performs a series of three processes, i.e., speech recognition, translation, and speech synthesis.
* The speech recognition function includes:
  + Speech recording via the microphone.
    - * Recording triggered by a proximity sensor is not supported.
  + Transmission of sound data in a linear data format and in an asynchronous transfer mode.
* The translation function includes:
  + Back translation to verify the accuracy of translation results.

※ Translation of input text is not supported.

* The speech synthesis function includes:
  + Playing the sound data received from the MCML system in a linear data format.

※ Speech synthesis of input text is not supported.

## Released materials

Released materials include the sample application and related documents as follows:

* Source codes
* User's Guide for Sample Application (Android) (this document)
* API reference (JavaDoc)

# Environments

## Operating environment

This application requires Android 2.2 or later.

## Development environment

The source code of this application is a part of the Eclipse project structure.

External libraries are needed for development and operation of the application.

* + 1. MCML Libraries

The following jar files are used as a library of the MCML system

* MCMLLib.jar
* MCMLLibAU.jar
  + 1. External Libraries

The following library is required for the application to work.

* JavaMail API

Download the three jar files from:

<http://code.google.com/p/javamail-android/downloads/list>

・mail.jar

・activation.jar

・additional.jar

* commons-codec

Download the compressed file including the jar file from:

<http://archive.apache.org/dist/commons/codec/binaries/>

(Download and unfreeze"commons-codec-1.5-bin.zip".

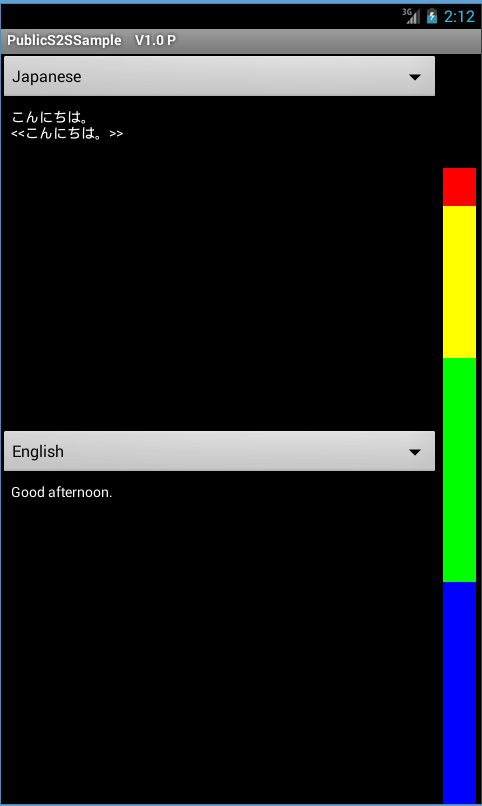
・commons-codec-1.5.jar

* + 1. Inclusion of the library into the sample application project

Add the six files described in Sections 2.2.1. and 2.2.2. to the sample application project as external jar files.

# Screen View

A screen view of the sample application is shown below.



Language-selecting list

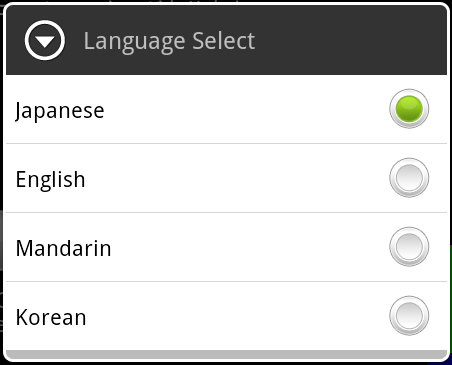
Language-selecting list

Result display area

Level meter

Result display area

Language-selecting list



## Language-selecting list

This list is used to select a language for recognition and translation.

Languages options include Japanese, English, Mandarin, and Korean. For translation, select two different languages separately in the upper and lower lists. If you choose the same language, only speech recognition, but not translation, will be performed.

## Result display area

The result display area is used to display results of speech recognition and translation, and also to start/stop speech recording.

* + 1. What appears in the display area

The result display area provides a view of:

* Speech recognition results

Results of speech recognition by a language selected in the list are displayed.

* Translation results

Results of translation by a language selected in the list are displayed.

Back translations, enclosed in angle brackets << >>, are shown below the recognition results.

* Processing status

The following messages appear to indicate the processing status.

* + Recording…
  + Recognizing…
  + Translating…
    1. Starting and stopping speech recording

Recording of speech is initiated by pressing a result display area, and is continued while the area is being pressed. If you release the area, recording is terminated and recognition is started. This function works in both of the upper and lower result display areas.

* + By pressing the upper result display area, speech is recognized by a language selected in the upper language-selecting list, and results of translation into another language selected in the lower language list will appear in the lower display area.
  + By pressing the lower result display area, speech is recognized by a language selected in the lower language-selecting list, and results of translation into another language selected in the upper language list will appear in the upper display area.

## Level meter

The level meter indicates the sound input level from the microphone.

# Setting and Logs of Application

## Setting of the server to communicate

Specify the address of the MCML system (server) in the source code.

* Source file

…/src/jp/go/nict/S2SSample/PublicS2SSendThread.java

* Code to edit

private final String serverUrl = "**Insert the server's URL here**";

Speech codec for input speech

* Source file

…/src/jp/go/nict/S2SSample/PublicSpeechTranslator.java

* Code to edit

The ‘Encode’ method in SignalAdpcm class encodes input speech to ADPCM. Delete this process if you do not wish to encode input speech to ADPCM.

Note: The output speech is processed automatically in accordance with the attributes of ‘AudioFormat’ in the ‘Signal’ tag of the MCML response (i.e. ADPCM would be decoded to raw PCM) and therefore, the codec settings is not necessary.

## Logs

Operating status of the application is output as a log.

You will find it on the debugging tool, DDMS (Dalvik Debug Monitor Service), of the Android SDK.

* Example of output logs

12-01 19:42:05.470: I/S2SSample(956): 分割送信先頭 https://myserver/ControlServer/ControlServer XML <?xml version="1.0" encoding="UTF-8"?><MCML xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" Version="1.0"><User><Transmitter><Device><Location><URI>dummyUtteranceID</URI><GlobalPosition Longitude="135.7548" Latitude="35.00913"/></Location></Device><UserProfile ID="test07"/></Transmitter><Receiver><Device><Location><URI>dummyUtteranceID</URI></Location></Device><UserProfile ID="nict-dev-111"/></Receiver></User><Server><Request Service="ASR" ProcessOrder="101"><InputUserProfile ID="dummyUserID" Gender="Male" Age="29"><InputModality><Speaking><Language ID="en" Fluency="5"/></Speaking></InputModality></InputUserProfile><InputUserProfile ID="dummyUserID" Gender="Male" Age="29"></InputUserProfile><TargetOutput><HypothesisFormat NofN-best="5"/></TargetOutput><Input><Data><Audio ChannelID="1"><ModelType><Domain>Travel</Domain><Task>Dictation</Task></ModelType><Signal SamplingRate="16000" ValueType="integer" AudioFormat="raw PCM" BitRate="16" Endian="Little" ChannelQty="0"/></Audio></Data><AttachedBinary ChannelID="1" DataID="ja"/></Input></Request></Server></MCML>

# Screen Transition

Screen transition through the course of speech recognition, translation, and speech synthesis is presented below.

## Normal operation

Recognizing…

Japanese　　　　▼

English 　　　 　▼

5.1.3

5.1.2

Recording…

Japanese　　　　▼

English 　　　 ▼

Japanese　　　　▼

English 　　　 　▼

5.1.1

Releasing the display area

Pressing the display area

Recognition completed

Processing translation

本日は晴天なり

<<今日は良い天気です>>

Today is fine day.

Japanese　　　　▼

English 　　　 　▼

5.1.6

本日は晴天なり

Today is fine day.

Japanese　　　　▼

English 　　　 　▼

5.1.5

5.1.4

本日は晴天なり

Translating..

Japanese　　　　▼

English 　　　 　▼

Translation completed

Processing back translation

Processing speech synthesis

Back translation completed

本日は晴天なり

<<今日は良い天気です>>

Today is fine day.

Japanese　　　　▼

English 　　　 　▼

5.1.7

Speech synthesis completed

* + 1. Before starting speech

When the application is started, the startup screen will appear.

Recording of speech is initiated by pressing the result display area.

* + 1. Recording speech

Recording is continued while the result display area is being pressed.

Text of messages being recorded will appear on the result display area of each language.

* + 1. Processing speech recognition

If you release the result display area, the recorded speech data is sent to the server and processed for speech recognition.

* + 1. Speech recognition completed/Processing translation

Results of speech recognition returned from the server are shown in text on the result display area. The recognition results are automatically transmitted to the server for machine translation.

* + 1. Translation completed/Processing back translation/Processing speech synthesis

Results of machine translation returned from the server are shown in text on the result display area. The translation results are automatically transmitted to the server for back translation and speech synthesis.

* + 1. Back translation completed/Processing speech synthesis

Results of back translation returned from the server are shown on the result display area.

* + 1. Speech synthesis completed

Results of speech synthesis are returned from the server, and the synthesized speech is played on your device.

## Abnormal operation

Recognizing…

Japanese　　　　▼

English 　　　 　▼

本日は晴天なり

Translating..

Japanese　　　　▼

English 　　　 　▼

本日は晴天なり

Today is fine day.

Japanese　　　　▼

English 　　　 　▼

Translation completed

Processing back translation

Processing speech synthesis

Recognition completed

Processing translation

Back translation successful

Synthesis failed

Translation failed

Recognition failed

Back translation failed

Synthesis successful

Back translation failed

Synthesis failed

本日は晴天なり

<<There is no back

translation result.>>

Today is fine day.

Japanese　　　　▼

English 　　　 　▼

本日は晴天なり

<<There is no translation result.>>

Japanese　　　　▼

English 　　　 　▼

<<There is no recognition result.>>

Japanese　　　　▼

English 　　　 　▼

本日は晴天なり

<<今日は良い天気です>>

Today is fine day.

Japanese　　　　▼

English 　　　 　▼

本日は晴天なり

<<There is no back

translation result.>>

Today is fine day.

Japanese　　　　▼

English 　　　 　▼

* Messages displayed in case of process failure
  + Failure in recognition

Message in the result display area: << There is no recognition result.>>

* + Failure in translation

Message in the result display area: <<There is no translation result.>>

* + Failure in back translation

Message in the result display area: <<There is no back translation result.>>