

AquesTalk Android Manual

1. 概要

本文書は、規則音声合成ライブラリ AquesTalk Android をアプリケーションに組み込んで使用するためのプログラミングに関しての方法および注意点を示したものです。

AquesTalk はかな表記のよみ記号にアクセント等の情報を付与した音声記号列から、音声波形データを生成する日本語の音声合成ライブラリです。

なお、漢字かな混じり文からの音声合成には、別途、言語処理ライブラリ AqKanji2Koe を利用します。

特長

- ・耳なじみの良い声
聴きやすさや明瞭性を重視し、人の声にこだわらずに開発
声種の切り替えは簡単で、システムに複数の声種を入れることもできます
- ・超小型軽量
他に類を見ない少ないリソース量で音声合成が実現できます。
- ・簡単に実装
簡単なインターフェース関数。
ファイルアクセスも不要なため、ライブラリをリンクするだけで簡単に使用できます。

2. 仕様

AquesTalk

ライブラリ形式	so 形式ダイナミックライブラリ (JNI 実装) 各種 ABI に対応 armeabi-v7a, arm64-v8a, x86, x86_64
対応 OS	Android 4.1 (API レベル 16) 以降
入力データ形式	音声記号列 (UTF-8)
出力データ形式	WAV フォーマットデータ (8KHz サンプリング, 16bitPCM)
開発環境	Android NDK r19
プログラムサイズ	約 110KByte (音素片データ含む)

3. 関数 API

3.1. Android 版 Java ラップクラス

JNI をラップした Java クラスを用いてアクセスします

syntheWav

package AquesTalk

説明	音声記号列から音声データを生成します
構文	byte[] syntheWav(String kanaText, int speed)
引数	
kanaText	音声記号列 (UTF-8)

speed	発話速度(%) 50 - 300
戻り値	wav フォーマットのデータ エラー時には、長さ1で、先頭にエラーコードが返される

4. ファイル

ビルド及び実行時に下記のファイルを使用します。

libAquesTalk.so : AquesTalk ライブラリの実体(ネイティブコード)
AquesTalk.java : libAquesTalk.so を Java からアクセスするためのラッパークラス

任意のアプリケーションを開発するときは、libAquesTalk.so と AquesTalk.java ファイルを、アプリケーションのプロジェクトに追加します。Android Studio の環境であれば、プロジェクトディレクトリに、上のファイルを下記のようなディレクトリ構造上に(ディレクトリがなければ作成)コピーします。

「AquesTalk Android」の実体はネイティブコードで書いてあり、JNI 経由で呼び出しています。そのため、使用するシステム環境(ABI)に応じたライブラリを使用する必要があります。たとえば、使用端末が ARM CPU の場合は arm64-v8a か armeabi-v7a を、Intel CPU の場合は x86, x86_64 などのライブラリを使用します。

Andorid では FAT-Binary(シングル APK)によって、複数の異なる ABI のライブラリを1つの apk に含めることができ、動作マシン側で自動的に選んで実行します。そのため、apk のサイズさえ気にしなければ、SDK パッケージの jniLibs 以下すべてをアプリのプロジェクト側にコピーしても良いでしょう。

```

/<app/src/main>
  /java
    /<アプリのソース>
    /aquestalk
      /AquesTalk.java
  /jniLibs
    /arm64-v8a
      /libAquesTalk.so
    /armeabi-v7a
      /libAquesTalk.so
    /x86
      /libAquesTalk.so
    /x86_64
      /libAquesTalk.so
  /res
    .
    .

```

なお、呼び出しの関係は、次のようになります。

* アプリ -> AquesTalk.java -> libAquesTalk.so

5. アプリへの実装方法

任意のアプリに AquesTalk を組み込む方法を SDK に含まれているサンプルアプリのコードを参照しながら以下に示します。(samples¥AqTkApp¥app¥src¥main¥java¥com¥a_quest¥aquestalk¥AqTkApp.java)

AquesTalk のクラスをインポートします。

```
import aquestalk.AquesTalk;
```

AquesTalk.java の中を見ればわかりますが、静的メソッドと、通常のメソッドが用意されています。いまのところ、関数は1つだけなので、どちらを使っても違いはありません。

AqTkApp.java では、アプリ起動時の onCreate() のなかで、インスタンスを生成しています。

```
aquestalk = new AquesTalk();
```

そして、発声ボタンが押されたときに、音声記号列 と発話速度を取得して、次のように音声データをメモリ上に生成します。なお、音声記号列の文字コードは UTF-8 で指定します (Android のデフォルトのままで ok)。

```
onPlayBtn():  
    byte[] wav = aquestalk.syntheWav(koe, speed);
```

ここで得られる音声データは wav フォーマット (8KHz, Mono, Straight PCM) です。音声記号列が不正な場合など生成エラーの時は、byte[] のサイズが 1 になり、配列の先頭にエラーコードが返されます。エラーコードは、下章のエラーコード表を参照ください。

現時点では、「AquesTalk Android」に直接音声を出力するメソッドは用意されていないので、音声再生はアプリ側で記述する必要があります。AudioTrack クラスを用いて音声出力するコードの例がサンプルアプリ中の AqTkApp.java にありますので、参考にしてください。

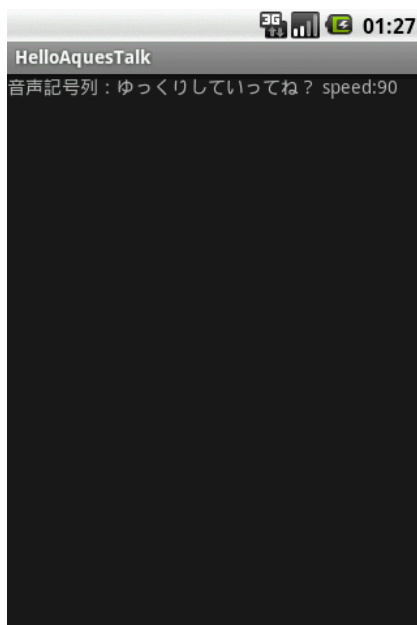
注意すべき点は、AquesTalk の出力は wav フォーマットであるのに対し、AudioTrack で再生するデータは、ヘッダを含まない生のデータ列を指定する必要があります。そこで、生成した音声データの先頭にある WAV ヘッダ (44 バイト) を除いて、それ以降を AudioTrack に渡します。また、データの長さ (サンプル数) は、1 つのサンプルが 2 バイトですので、次のようになります。

```
サンプル数 = (wav.length - 44) / 2
```

6. サンプルアプリ

SDK パッケージには、ライブラリと一緒に、サンプルアプリが 2 つ (HelloAquesTalk, AqTkApp) 入っています。サンプルアプリのソースコード (ライブラリやライブラリのヘッダファイルを除く) は、ご自由に改変してご使用いただけます。

- * HelloAquesTalk:
最も単純なサンプルで、起動時に規定のメッセージを合成出力するだけのものです。
- * AqTkApp:
GUI を持たせ、任意の音声記号列と話速を指定して音声合成するものです。



サンプルアプリの動かし方

Android Studio の開発環境が構築されていることが前提です。以下に HelloAquesTalk アプリを動かす手順を示します。もうひとつのサンプル AqTkApp や、独自のアプリの場合も同様の手順で使用できます。

1. SDK パッケージ内の samples/HelloAquesTalk フォルダを任意の場所にコピー
2. SDK パッケージ内の java/aquestalk フォルダを HelloAquesTalk/app/src/main/java フォルダにコピー
3. SDK パッケージ内の jniLibs フォルダ以下を HelloAquesTalk/app/src/main フォルダにコピー
4. Android Studio を起動
5. menu>File>Open でコピーした HelloAquesTalk フォルダを Open
6. 各種設定ファイルや build が自動的に行われます
7. あとは、menu>Run>Run で実行

7. エラーコード表

関数が返すエラーコードの内容は、次の通りです。

値	内容
100	その他のエラー
101	メモリ不足
102	音声記号列に未定義の読み記号が指定された
103	韻律データの時間長がマイナスになっている
104	内部エラー(未定義の区切りコード検出)
105	音声記号列に未定義の読み記号が指定された
106	音声記号列のタグの指定が正しくない
107	タグの長さが制限を越えている(または[>]が見つからない)
108	タグ内の値の指定が正しくない
109	WAVE 再生ができない(サウンドドライバ関連の問題)
110	WAVE 再生ができない(サウンドドライバ関連の問題 非同期再生)
111	発声すべきデータがない
200	音声記号列が長すぎる
201	1つのフレーズ中の読み記号が多すぎる
202	音声記号列が長い(内部バッファオーバー)
203	ヒープメモリ不足
204	音声記号列が長い(内部バッファオーバー)

8. ライセンス

本ライブラリを含んだアプリケーションを使用する場合は「使用ライセンス」の購入が必要となります。また、本ライブラリを含んだアプリの配布には事前に当社と『配布ライセンス契約』が必要となります。ライセンスの詳細は、SDK パッケージ内のライセンスドキュメントを参照ください。

9. 履歴

日付	版	変更箇所	更新内容	更新者
2011/01/31	1.0	新規作成		N.Y
2016/03/11	1.4		SDK Ver. 1.4、IDE を AndroidStudio に変更	N.Y
2019/05/29	1.4.1		Android Studio 3.4 用に修正	N.Y