# 音声認識による接客ロボットの開発

金沢工業高等専門学校　澤田茂人

2016-8-23

**概要**

音声認識について今から解説する。音声認識とは、人間の声などをコンピューターに認識させることであり、話し言葉を文字列に変換したり、あるいは音声の特徴をとらえて声を出している人を識別する機能を指す。

## 目次



## 1　はじめに

今回の音声認識は食堂にて注文を取るロボットがある場合、人件費等の経費削減ができるため制作してほしいという要望があった。食堂で注文を取るロボに使用した音声を認識させるシステムについて述べていく。

### 1.1　研究の背景

食堂にて注文を取るロボットがある場合、人件費等の経費削減ができるため制作してほしいという要望があり音声でのやりとりをするためのシステムを究明する必要があると考える。

## 1.2 研究の目的

linuxPC に音声認識をするため julius、Open-JTalk を用意して人間の声をマイクから拾い、実際に店で使用することを目的とする。図 1 に音声認識をしようとした時のイメージを示す。

### 1.2.1 julius,Open-JTalk

Julius は，音声認識システムの開発・研究のためのオープンソースの高性能な汎用大語彙連続音声認識エンジンである。数万語彙の連続音声認識を一般の PC 上でほぼ実時間で実行できる。また，高い汎用性を持ち，発音辞書や言語モデル・音響モデルなどの音声認識の各モジュールを組み替えることで，様々な幅広い用途に応用できる。[1] ここではオリジナルの言語モデルを使用したため、図 2、図 3、図 4、図 5、図 6、図 7、図 8 に示す。Open-JTalk は、日本語テキストを音声に変換するシステムである。[2]

# 2 研究手順

【julius】

https://github.com/julius-speech/julius より julius 4.3.1 を clone する。

https://osdn.jp/projects/julius/downloads/60416/dictation-kit-v4.3.1-linux.tgz/ よりディクテーションキットをダウンロードし、解凍する。

https://osdn.jp/projects/julius/downloads/51159/grammar-kit-v4.1.tar.gz/ よりディクテーションキットをダウンロードし、解凍する。

cd julius-4.3.1/

./configure
make
sudo make instal
テキストファイルにて grammar ファイルと yomi ファイル図 2、図 3 に習って書く。
yomi2voca.pl tabaco.yomi ＞ tabaco.voca と打ち込むと図 4 が出力される。
mkdfa.pl tabaco と打ち込むと図 5、図 6、図 7、図 8 が出力される。
generate tabaco と打つことで文法から文をランダム生成することができる。[3] その様子を図 9 に示す。
tabaco.jconf と tabaco.dic を grammar-kit-v4.1 に移動させる。
dictation-kit 中で julius -C tabaco.jconf -input mic -charconv EUC-JP UTF-8
と打ち込むことで音声を入力できる。
【Open-JTalk】
cpp でプログラムを作成し、コンパイルする。今回は voice という名前で保存した。プログラムの中を図 10 に示す。

## 3　研究手順説明

julius は研究手順に示したコマンドを上から順に vi にて打ち込む。Open-JTalk は図 8 に示したプログラムを組み、コンパイルさせることで音声を出力させることができる。

## 4　研究結果

julius で音声を入力し、認識することを確認した。そのときの様子を図 11 に示す。Open-Jtalk では自作のプログラムを組み、必要な音声を出力することに成功した。そのときの様子を図 12 に示す。

## 5　考察

今回は julius を使用して音声の入力や Open-JTalk を使用して音声の出力を行うことができたがそれぞれ独立して操作を行ったため、julius を使用する際にテキストファイルを出力させ、そのテキストファイルを Open-JTalk が音声として出力させるためのツールを見つけて使用することができればこの研究は完成すると考えられる。

## 6　おわりに

今回ロボットに音声認識をさせることで人件費の削減につなげられるのではないかと考え、julius や Open-JTalk で音声の入出力を行うことと決め julius での音声の認識や Open-JTalk での音声の出力に成功した。今後は julius と Open-JTalk を連携する手段を見つけ、人間とロボットの対話を行えるようにする。

## 7　参考文献

[1] julius book http://julius.osdn.jp/juliusbook/ja/pr01.html

[2] Open JTalk - HMM-based Text-to-Speech System http://open-jtalk.sourceforge.net/
[3]generate https://julius.osdn.jp/juliusbook/ja/generate.html


図 1　様子

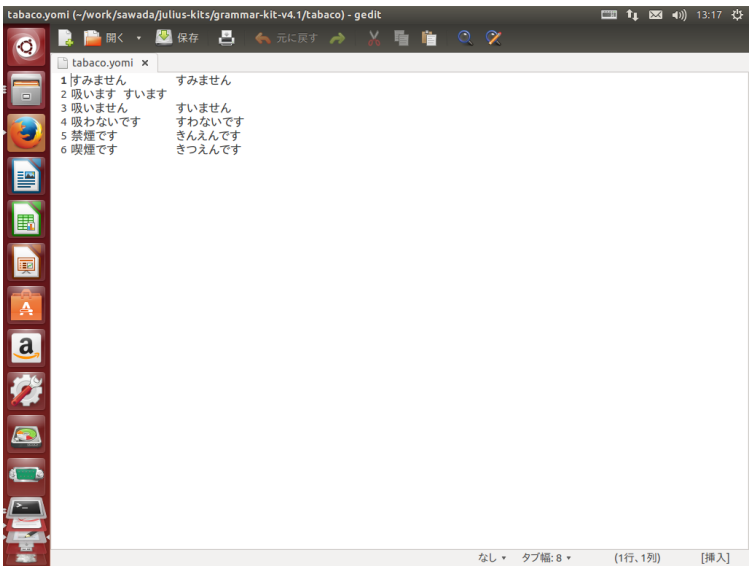


図 2　yomi ファイル

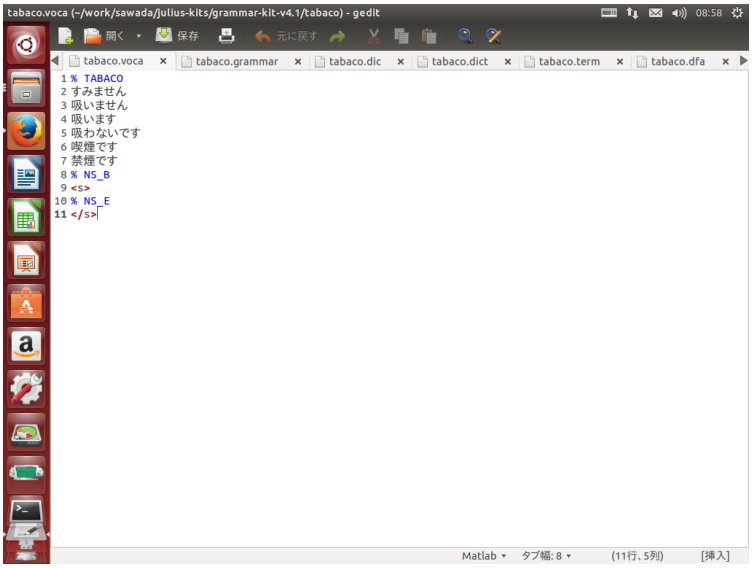


図 3　voca ファイル

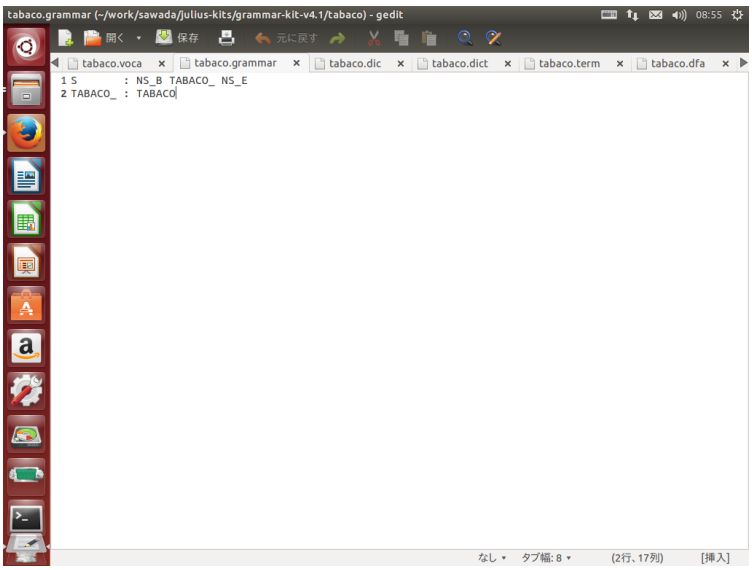


図 4　grammar ファイル

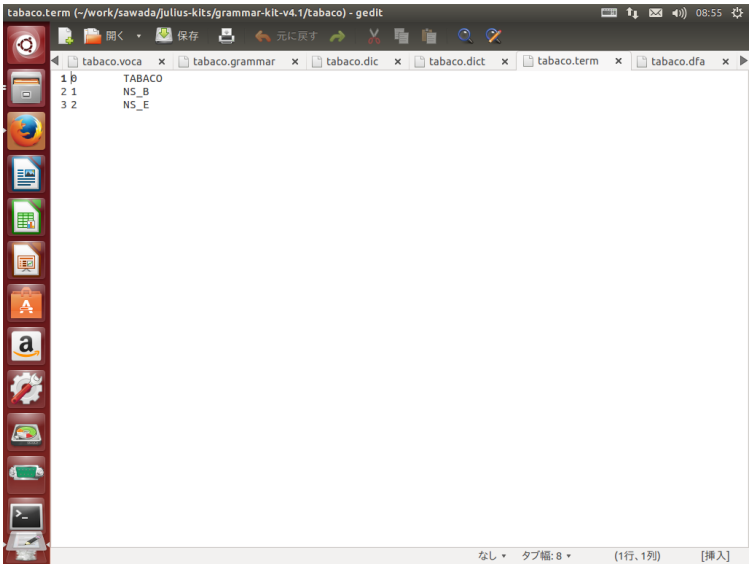


図 5　term ファイル

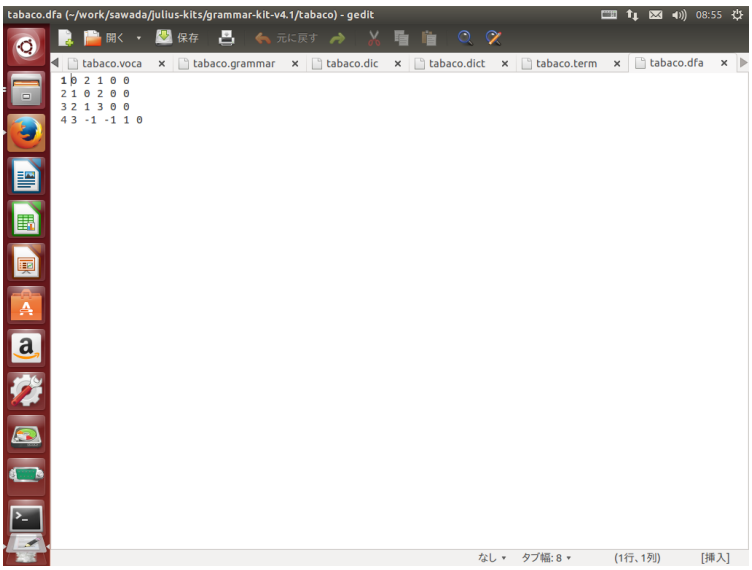


図 6　dfa ファイル

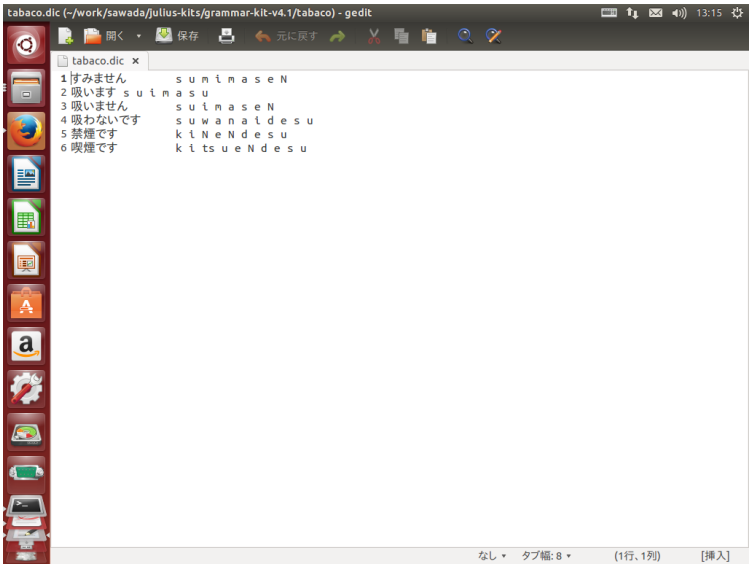


図 7　dic ファイル

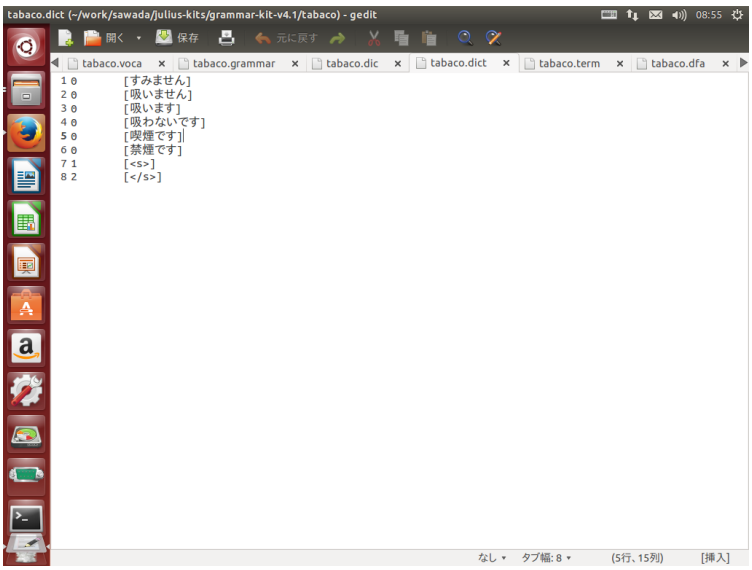


図 8　dict ファイル

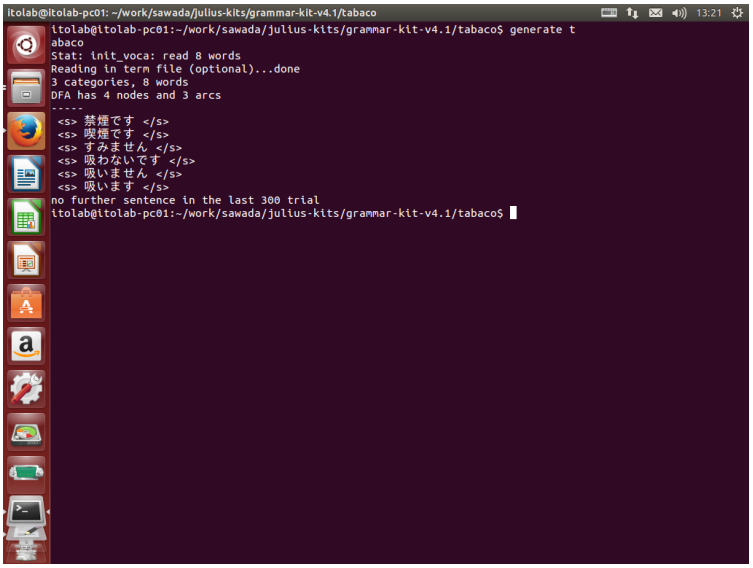


図 9　generate

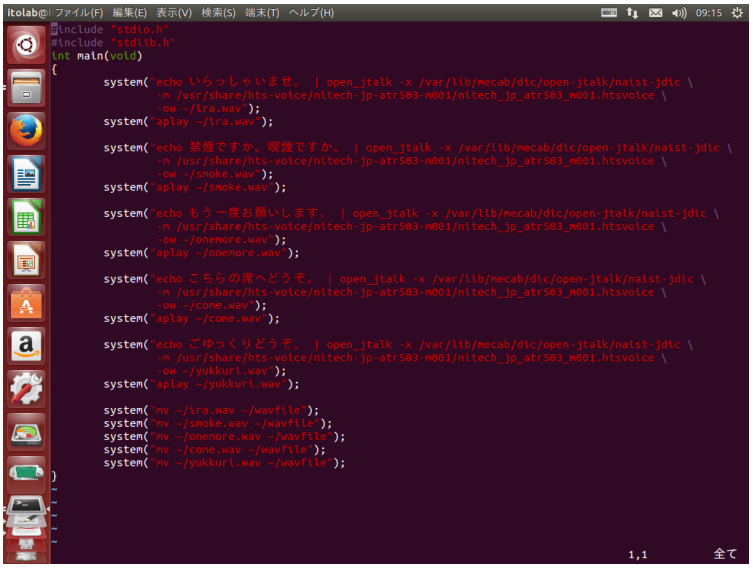


図 10　プログラム

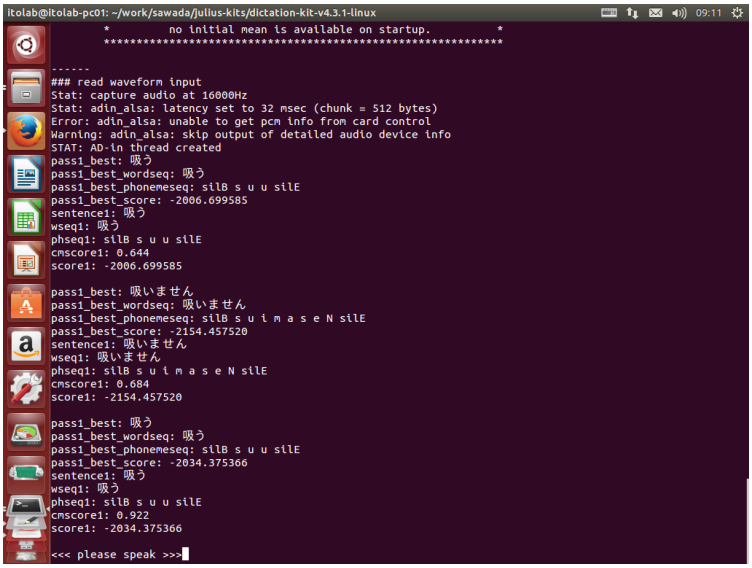


図 11　認識結果

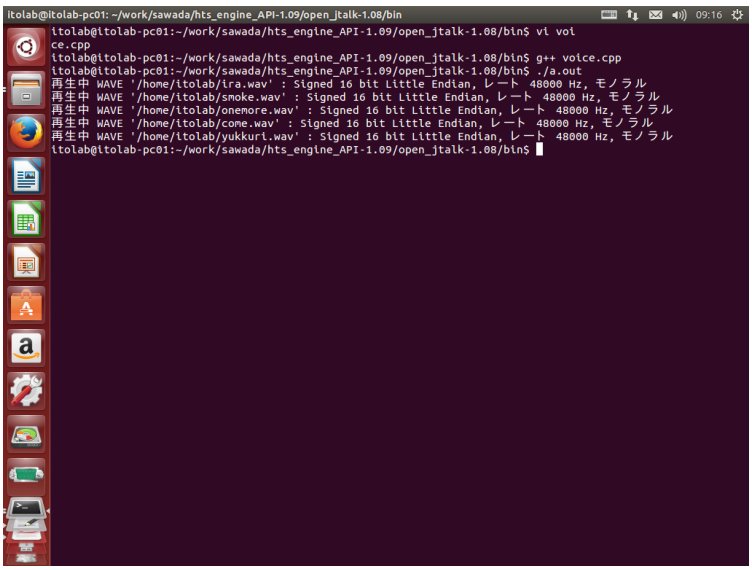


図 12　音声の出力