

=.53zw plus 3pt minus 3pt =.53zw plus 3pt minus 3pt

目次

| | |
|--------------------|----|
| 第 1 章 序論 | 3 |
| 1.1 研究の背景 | 3 |
| 1.2 研究の目的 | 4 |
| 1.3 論文の構成 | 5 |
| 第 2 章 飲食店での接客について | 6 |
| 2.1 声とは | 6 |
| 2.2 各社音声認識用ソフトの調査 | 8 |
| 2.3 音声認識とは | 11 |
| 2.4 各社音声合成用ソフトの調査 | 12 |
| 2.5 音声合成とは | 14 |
| 第 3 章 検討用のシステムについて | 15 |
| 第 4 章 音声認識実験 | 18 |
| 4.1 実験目的 | 18 |
| 第 5 章 結言 | 30 |
| 5.1 本研究のまとめ | 30 |
| 5.2 今後の課題 | 30 |
| .1 今回使用したもの | 30 |
| 参考文献 | 33 |
| 謝辞 | 34 |

第1章

序論

1.1 研究の背景

近年ソフトバンク社が Pepper を開発した。また、長崎県に存在するテーマパーク「ハウステンボス」で宿泊する際に利用すると考えられる「変なホテル」の受付には女性タイプと恐竜タイプの接客ロボットや話しかけことで照明のオンオフを行ってくれるロボットなどが配備されている。本研究は飲食店にて接客業務での人件費削減を行うため、ロボットを制作してほしいという要望があった。そこで音声でのやりとりをするためのシステムを充実する必要があると考え、本研究を開始した。飲食店で注文を取るロボットに使用した音声を認識させるシステムについて今から述べていく。図 1.1 に pepper、図 1.2 にホテルに存在する女性タイプの接客ロボット、図 1.3 に恐竜タイプの接客ロボットの外観を示す。

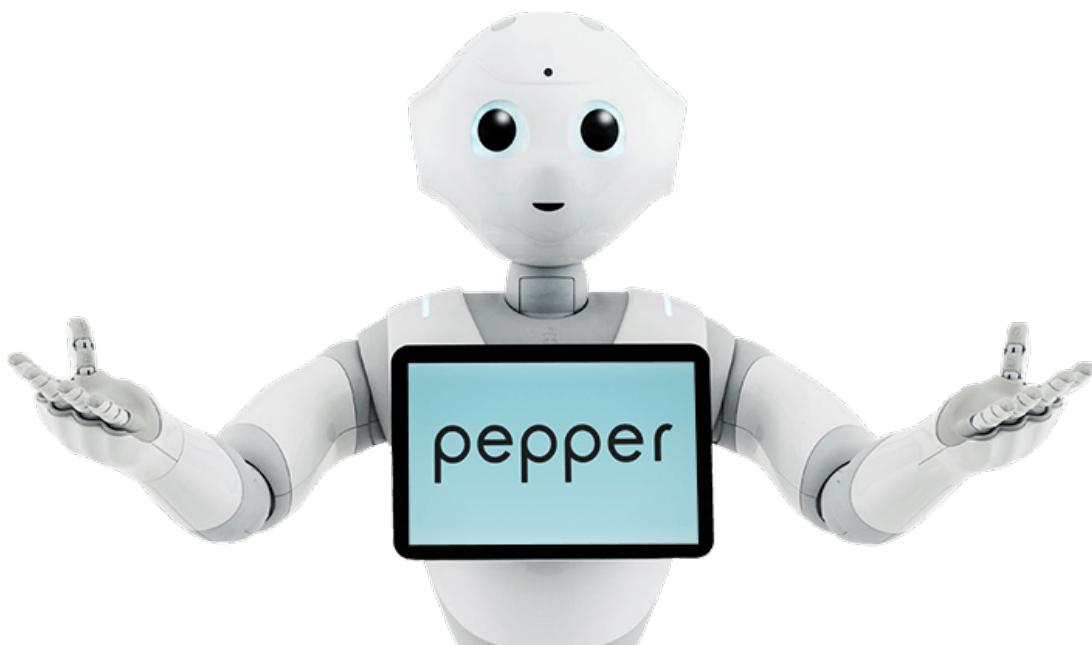


図 1.1 pepper



図 1.2 女性タイプの接客ロボット



図 1.3 恐竜タイプの接客ロボット

1.2 研究の目的

飲食店にて PC を実装したロボットに音声認識をするため Julius,Open-JTalk, 認識すべき単語, マイクを準備し飲食店で禁煙か喫煙かの判断や「あちらの席へどうぞ」「こちらの席へどうぞ」のようなレ

ベルの案内をすることを目的とする。図 1.4 に音声認識をしようとしたときのイメージ図を示す。



図 1.4 様子

1.3 論文の構成

この論文では以下のように構成されている。

1章では、音や音声について、研究背景について考察したものである。

2章では、本研究に使用する、音声を認識するための道具及び発生する音声について研究したものである。

3章では、本研究のために構築した検討用のシステムについて研究したものである。

4章では、前章で述べた検討用のシステムを使用して行った実験について研究したものである。

5章では、本研究のまとめや今後の課題について研究したものである。

6章では、本研究において大きな力となった関係者各員に謝辞を記す。

第2章

飲食店での接客について

2.0.1 接客の業務内容

お客さまから注文を取り、出来上がった料理や飲み物を運び、空いた皿の片付け、レジ業務を行うのが主な仕事である。その接客の業務内容の流れをフローチャートを図2.1に示す。

2.0.2 研究として検討する接客の業務内容

前述では接客の仕事内容をフローチャートとして挙げたが今回、接客をするロボットの開発を行うにあたり、店の入口付近で客がタバコを吸うかどうかの判断や「あちらの席へどうぞ」「こちらの席へどうぞ」の案内をすることが研究として検討する接客の業務内容となる。

2.1 声とは

声とは空気の振動であり波である。つまり \sin カーブで表すことができ、人間が声として発する「あ」という音一つから「卒業研究」のような単語、「私の職業は学生であるため勉強をすることが仕事である。」と言った文章を声で発している限り \sin カーブで表すことができる。人間の声をマイクに入れてそのデータをグラフのように近いグラフを書くためには特別な手段を用いる必要がある。それは周波数成分をチェックすることであり、音声にフーリエ変換を掛けることである。

2.1.1 フーリエ変換とは

まず、フーリエ変換は式2.1.1で示すことが出来る。式2.1.1より関数 $f(t)$ と $e^{-j\omega t}$ を掛けあわせて積分することにより2つの関数が似ていれば絶対値が大きな値、似ていなければ小さな値を取りますので関数 $f(t)$ は時間の関数でもあるためオシロスコープでも確認することが可能である。また、回転の成

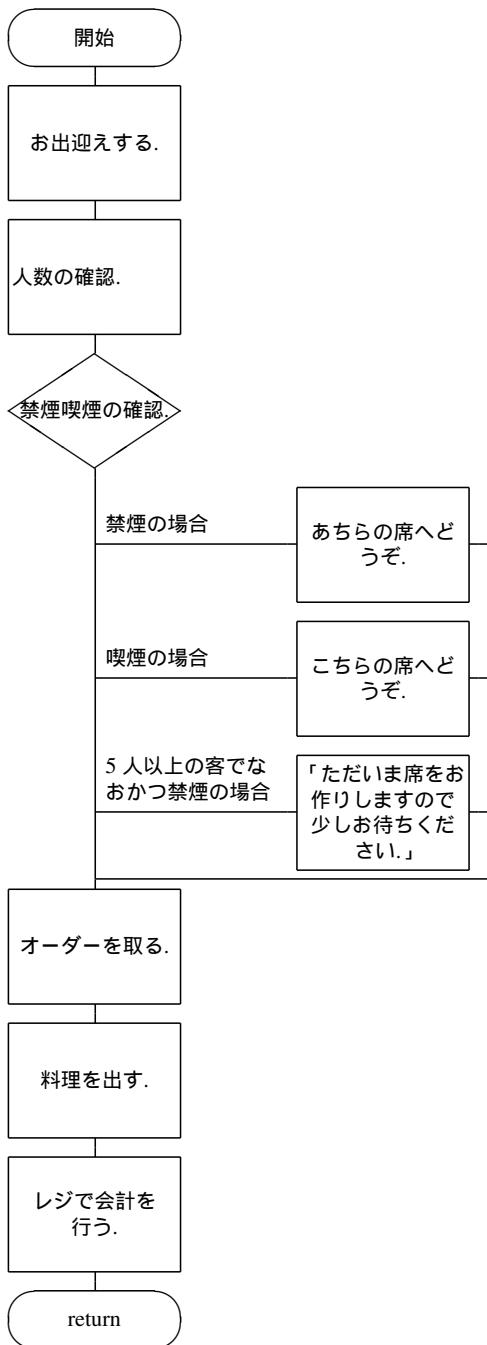


図 2.1 接客方法の流れ

分としてスペクトラムアナライザーでも確認することが出来る.[1]

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-j\omega t} dt = 0 \quad (2.1)$$

2.1.2 高速フーリエ変換とは

前述の通り声は空気の振動のため実際に視覚的に見ようとした場合、特別な手段を用いなければならない。ここではマイクを使い視覚的に表示することで音声の波形を見る能够があると考え、実際に audacity という音声の編集ソフトを使用して音の波形を視覚的に確認することに成功した。ここで、音の波形を高速フーリエ変換にかける。高速フーリエ変換は英語で Fast Fourier Transform と呼ばれていることから頭文字を取って FFT とも呼ばれている（以下 FFT とする）。前述に示したフーリエ変換に各音専用の係数をかけることで FFT を行うことができる。図 2.2 に「あ」の音声波形、図 2.3 に「い」の音声波形、図 2.4 に「う」の音声波形、図 2.5 に「え」の音声波形、図 2.6 に「お」の音声波形、図 2.7 に「あ」に FFT をかけた波形、図 2.8 に「い」に FFT をかけた波形、図 2.9 に「う」に FFT をかけた波形、図 2.10 に「え」に FFT をかけた波形、図 2.11 に「お」に FFT をかけた波形を示す。

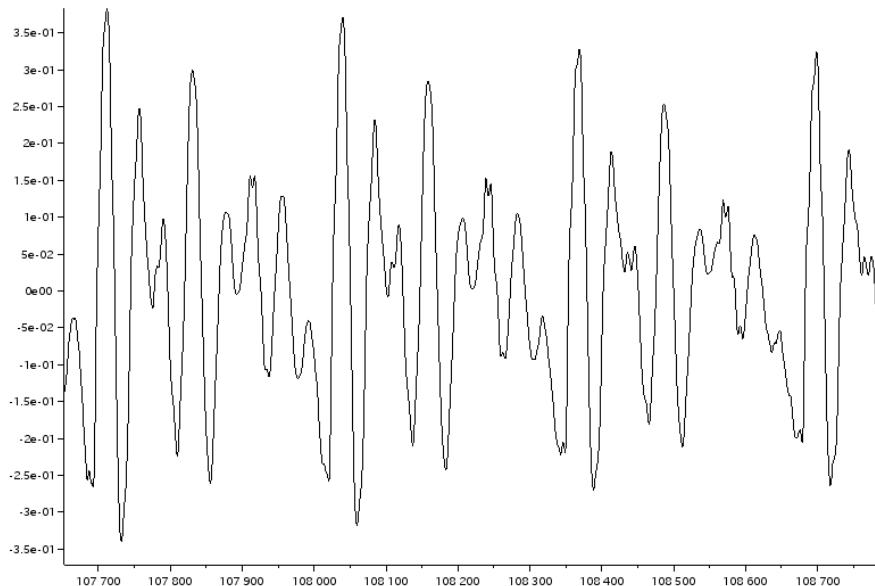


図 2.2 「あ」の音声波形

2.2 各社音声認識用ソフトの調査

音声によるロボットを制作するにあたり、まずは音声を認識することが必要条件である。そのため、音声を認識するプログラムを作成し使用することが考えられるのだが時間的コストがかかることが

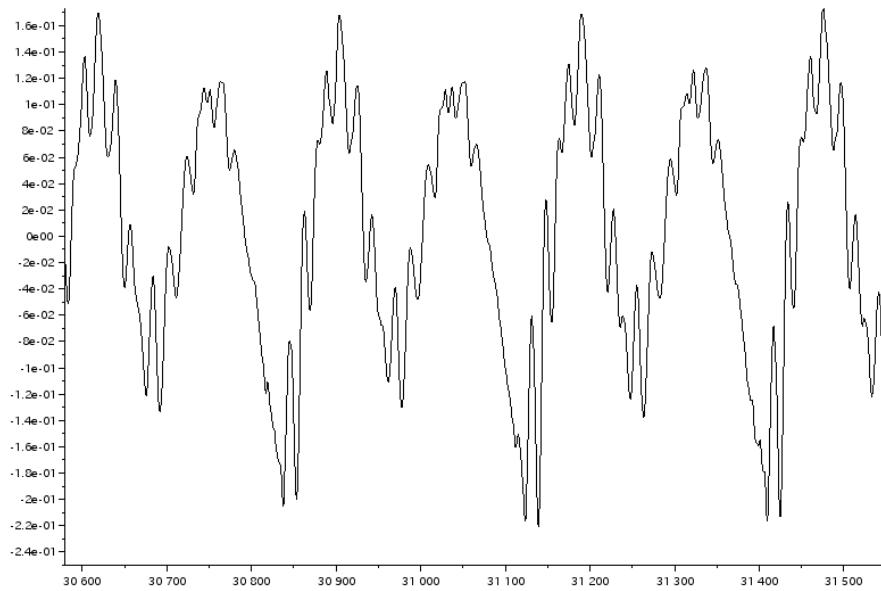


図 2.3 「い」の音声波形

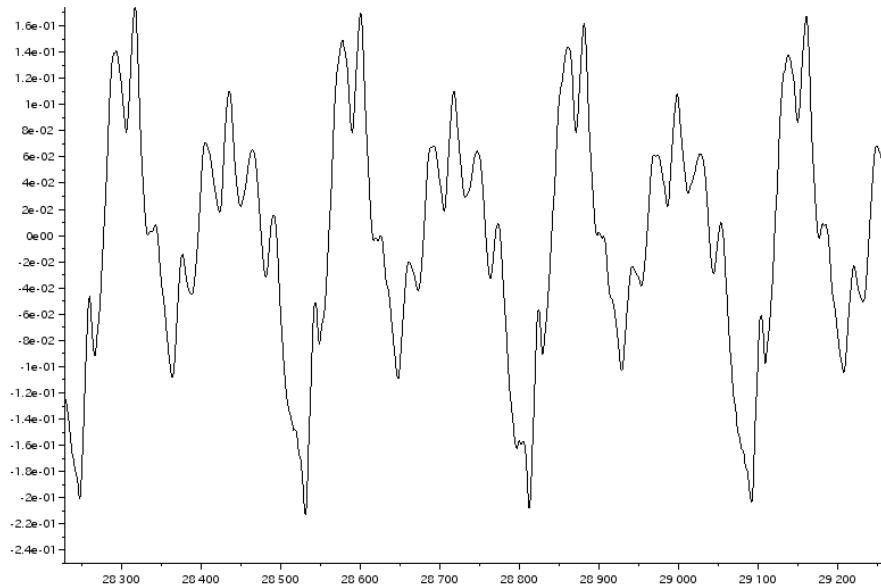


図 2.4 「う」の音声波形

予想された。そのため音声を認識するためのソフトを調査した。調査した結果,Julius, ドラゴンスピーチ,docomo 音声認識 API,Ami Voice,Google Speech API の 5 つが候補として挙がった。本研究では

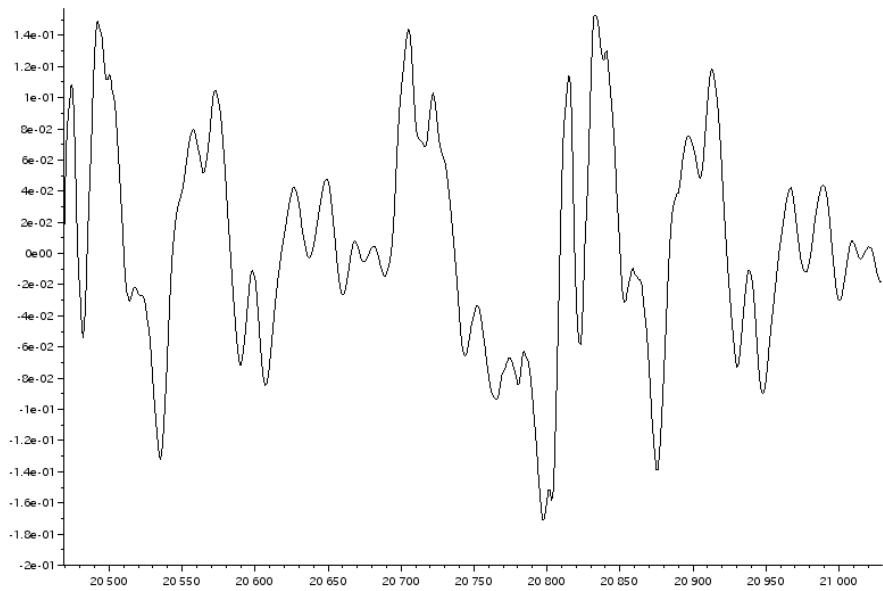


図 2.5 「え」の音声波形

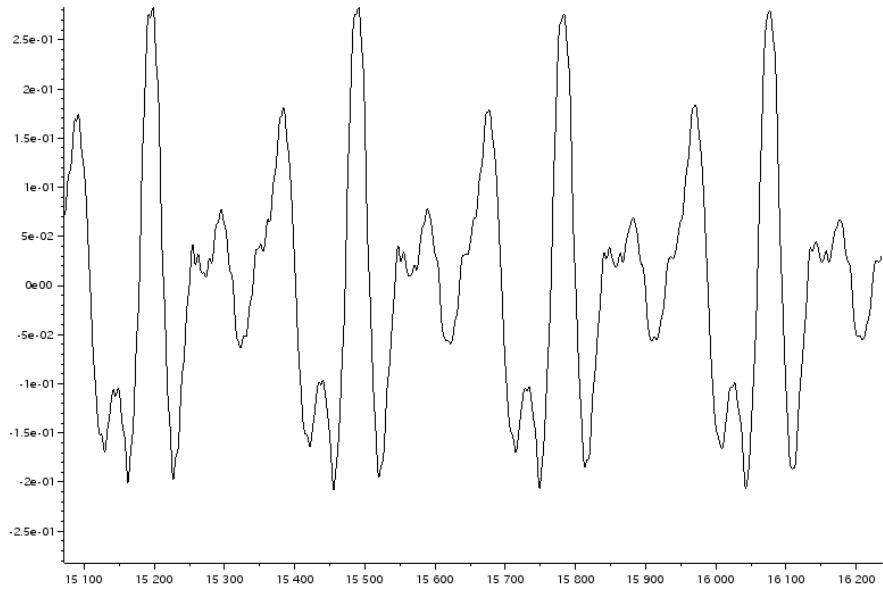


図 2.6 「お」の音声波形

linux でも使用でき無料かつ使用期間の制限がない Julius を選択した。表 2.1 に音声認識一覧を示す。また、本研究で使用することとした音声認識ソフト Julius のアイコンを図 2.12 に示す。

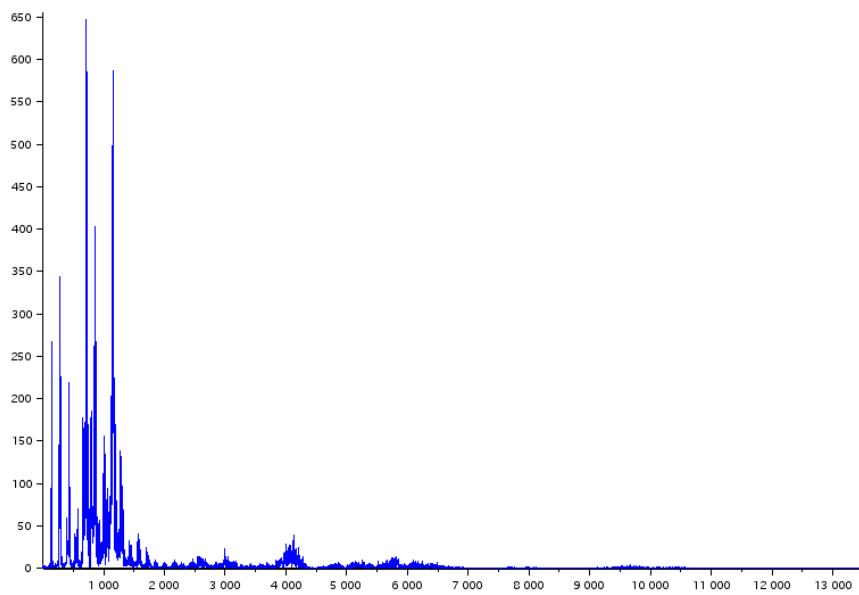


図 2.7 fft を掛けた「あ」の波形

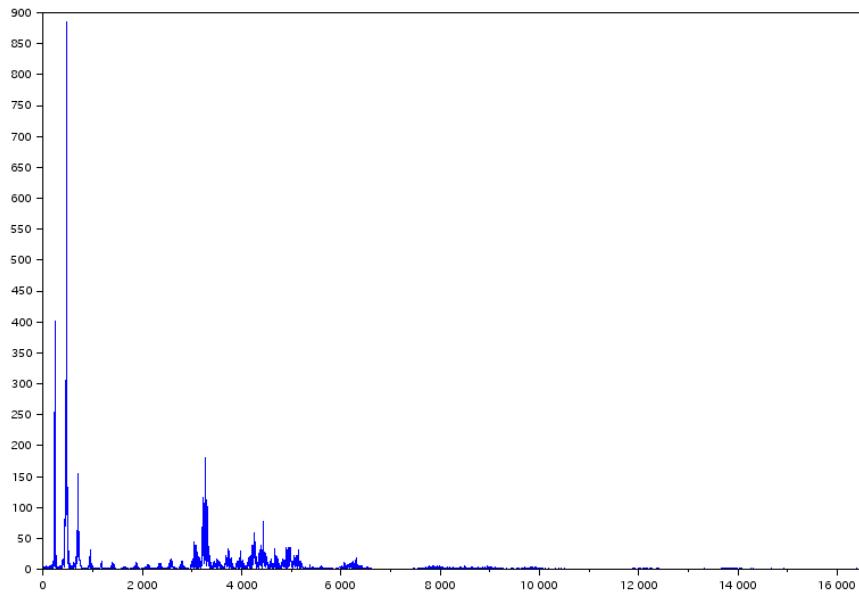


図 2.8 fft を掛けた「い」の波形

2.3 音声認識とは

音声認識とは音声をマイクより入力し、入力された音声を文字として表すことを指す。

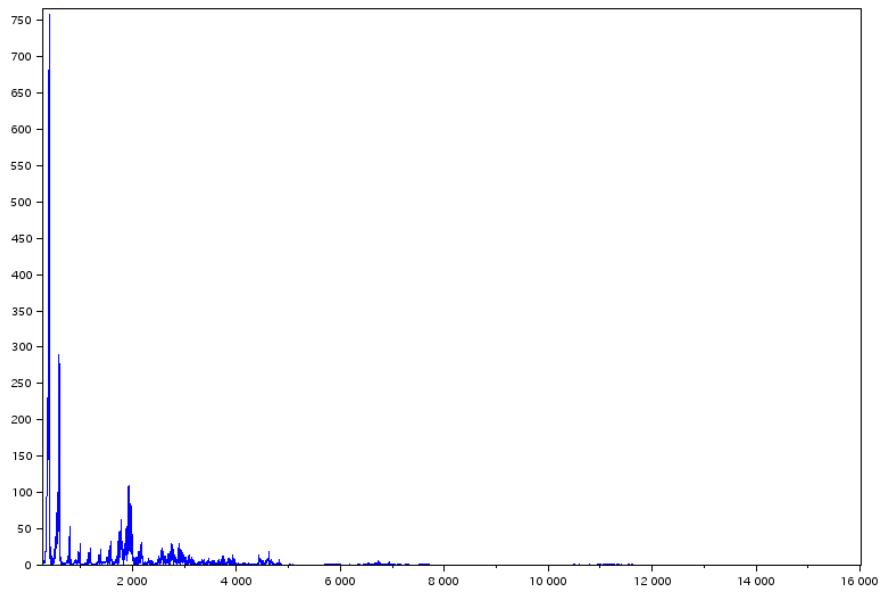


図 2.9 fft を掛けた「う」の波形

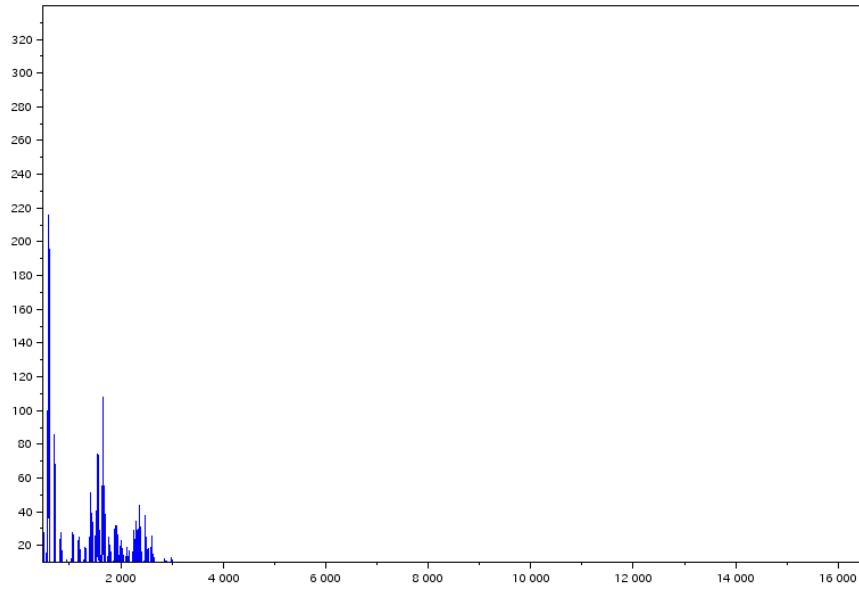


図 2.10 fft を掛けた「え」の波形

2.4 各社音声合成用ソフトの調査

接客ロボットは音声を認識した後に音声を発して店内での案内も行わなければならない、よって音声合成ソフトを使用することで音声を作成し、言語を用いて発声させることを考えた。そこで音声を作成

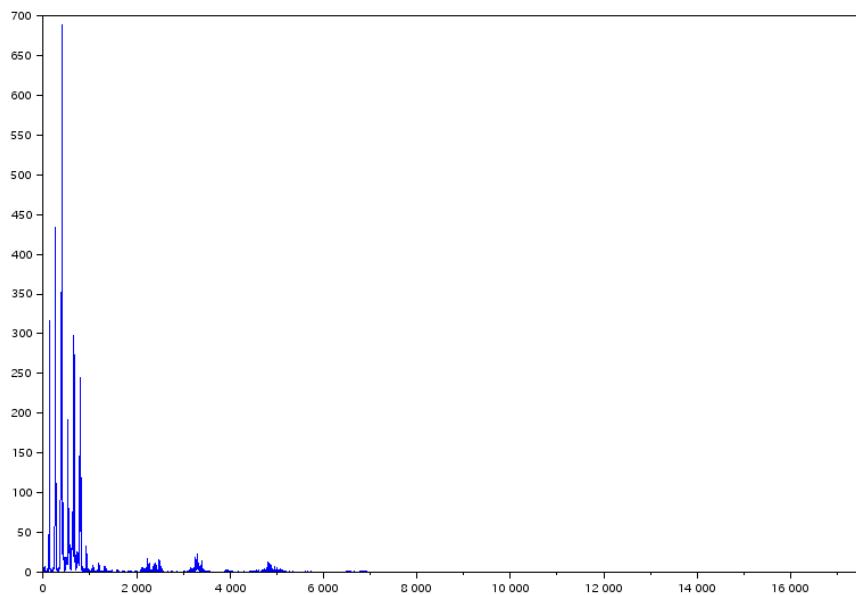


図 2.11 fft を掛けた「お」の波形

表 2.1 音声認識一覧

| ソフト名 | 無料である | 使用期限がない | 認識率 |
|-------------------|-------|---------|------|
| Julius | | | 90 % |
| ドラゴンスピーチ | × | × | 不明 |
| docomo 音声認識 API | × | × | 不明 |
| Ami Voice | × | × | 不明 |
| Google Speech API | × | × | 不明 |

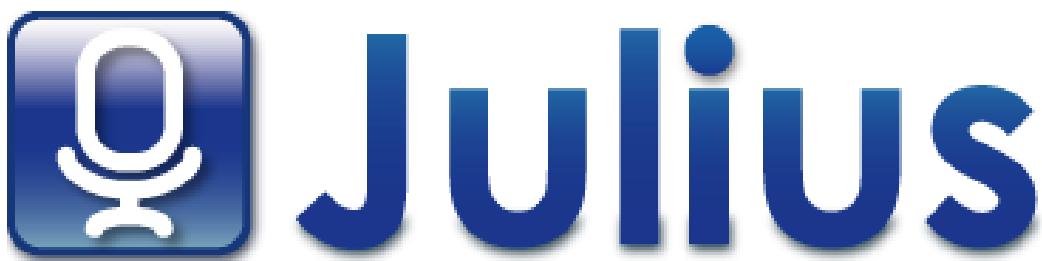


図 2.12 Julius

するためのソフトの調査を行った結果,Open-JTalk,Vocaloid,AITalk,VoiseText, テキストオーディオ,Clipboard Translator, 棒読みちゃん, 歌声合成ツール UTAU の 8 つのソフトが候補に挙がった. 本研究ではまず音

声を発声させることができ、なおかつ音声の速さが変更できる Open-JTalk を選択した。表 2.2 に音声合成一覧を示す。また、本研究で使用することとした音声合成ソフト Open-JTalk のアイコンを図 2.13 に示す。

表 2.2 音声合成一覧

| ソフト名 | 音の早さが変更可能 | 無料である | 使用期限がある | 男女性の声で話せる | 感 |
|----------------------|-----------|-------|---------|-----------|---|
| Open-JTalk | | | × | | |
| Vocaloid | | × | | 不明 | |
| AITalk | 不明 | × | | | |
| VoiceText | 不明 | × | | 不明 | |
| テキストトーク | | × | | × | |
| ClipBoard Translator | 不明 | × | | 不明 | |
| 棒読みちゃん | | × | | 不明 | |
| 歌声合成ツール UTAU | 不明 | × | | 不明 | |



図 2.13 Open-JTalk

2.5 音声合成とは

音声合成とは人間の音声を人工的に作り上げることを指す。

第3章

検討用のシステムについて

今年度のシステムは図 3.1 に示すシステムを使用して音声認識を行う。



図 3.1 使用したシステム

3.0.1 実験で使用するマイク

本研究では buffalo 製のマイク (BSHSM04BK) を使用した。その図を図 3.2 に示す。



図 3.2 使用したマイク

3.0.2 Julius

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。^[2]Julius はコマンド 1 つで起動させることができ、オプションでサーバモードで使用できるため、今回はサーバモードで起動し、クライアント側で認識した単語に対応した音声を発生させることを目指す。また本研究で使用した Julius は最新版としてバージョン 4.4.2 が平成 28 年 9 月 12 日に Github にて更新されている。しかし、研究を始めた時点でバージョン 4.3.1 が最新版だったためこれを継続して使用することとした。

Julius の仕組み

Julius の音声認識アルゴリズムは、ツリートレリス探索方式を基礎とするアルゴリズムである。全体は 2 パス構成となっており、2 段階に分けて認識処理を行う。まず、第 1 パスでは入力全体に対して荒い認識処理を行い、有望な候補の集合をある程度絞り込む。このとき、簡便なモデルや近似計算を用いることで高速に処理を行う。第 2 パスでは詳細な認識処理を行うが、その際に第 1 パスの結果

を参照しながら探索を行うことで，必要な部分にだけ精密な再計算を行って，最終的な最尤解を求め
る . [3]

3.0.3 Open-JTalk

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. ここでは案内のために使用する音
声の内容を WAV ファイルとして保存するために使用する。「今日はあちらの席へどうぞ」，「こちらの
席へどうぞ」の 2 つを WAV ファイルとして保存した.

第4章

音声認識実験

4.1 実験目的

今回 Julius を使用して音声を認識させ, 100 回分の認識結果を集計し平均を取ることを目的とする.

4.1.1 実験方法

マイクを実装したパソコンを使用し, Julius を端末より起動させる. Julius を使用し, 「禁煙です」, 「喫煙です」, 「吸います」, 「吸いません」, 「吸わないです」の認識を行う. このとき 100 回分の認識したときと認識しなかったときの結果を集計する. 100 回分の認識結果の平均を式 4.1.1 より算出する.

$$x = a/100 \quad (4.1)$$

4.1.2 実験結果

4.1.3 Julius を用いて行った実験結果

今回 Julius を用いて実験を行い, 前節に示した式より算出した認識率は禁煙が 99 %, 喫煙が 100 %, 吸いますが 100 %, 吸いませんが 30 %, 吸わないですが 100 % という結果が得られた. また, マイクと被験者との距離を遠く取り, 実験を行った. そのときの認識率は禁煙が 90 %, 喫煙が 86 %, 吸いますが 43 %, 吸いませんが 69 %, 吸わないですが 71 % だった. 別条件での認識結果を表 4.1 に禁煙, 表 4.2 に喫煙, 表 4.3 に吸いますが, 表 4.4 に吸いません, 表 4.5 に吸わないですがの認識結果, 4.6 に認識結果一覧を示す. また, そのときのグラフを図 4.1 に示す. また、マイクとの距離を遠くして認識実験を行った結果, という結果が得られた. そのときの表を表 4.7, 表 4.8, 表 4.9, 表 4.10 表 4.11, に示す. グラフにした

ものを 4.2 に示す.

表 4.1 禁煙の認識結果

| 回数 | 禁煙です | 回数 | 禁煙です | 回数 | 禁煙です | 回数 | 禁煙です |
|----|------|----|------|----|------|-----|------|
| 1 | | 26 | | 51 | | 76 | |
| 2 | | 27 | | 52 | | 77 | |
| 3 | | 28 | | 53 | | 78 | |
| 4 | | 29 | | 54 | | 79 | |
| 5 | | 30 | | 55 | | 80 | |
| 6 | | 31 | | 56 | | 81 | |
| 7 | | 32 | | 57 | | 82 | |
| 8 | | 33 | | 58 | | 83 | |
| 9 | | 34 | | 59 | | 84 | |
| 10 | | 35 | | 60 | | 85 | |
| 11 | | 36 | | 61 | | 86 | |
| 12 | | 37 | | 62 | | 87 | |
| 13 | | 38 | | 63 | | 88 | |
| 14 | | 39 | | 64 | | 89 | |
| 15 | | 40 | | 65 | | 90 | |
| 16 | | 41 | | 66 | | 91 | |
| 17 | | 42 | | 67 | | 92 | |
| 18 | | 43 | | 68 | | 93 | |
| 19 | | 44 | | 69 | | 94 | |
| 20 | | 45 | | 70 | | 95 | |
| 21 | | 46 | | 71 | | 96 | |
| 22 | | 47 | | 72 | | 97 | |
| 23 | | 48 | | 73 | | 98 | |
| 24 | | 49 | | 74 | | 99 | |
| 25 | 0 | 50 | | 75 | | 100 | |

4.1.4 考察

前述の音声認識実験での結果より認識率が 99 %, 30 % の単語が存在するがこれは「禁煙です」が「吸います」、「吸います」が「吸いません」と間違えたためである。また、「禁煙です」が「吸います」と誤認識をした原因は不明だが「吸います」と「吸いません」が極めて似ている単語であるため誤認

表 4.2 嘸煙の認識結果

| 回数 | 喰煙です | 回数 | 喰煙です | 回数 | 喰煙です | 回数 | 喰煙です |
|----|------|----|------|----|------|-----|------|
| 1 | | 26 | | 51 | | 76 | |
| 2 | | 27 | | 52 | | 77 | |
| 3 | | 28 | | 53 | | 78 | |
| 4 | | 29 | | 54 | | 79 | |
| 5 | | 30 | | 55 | | 80 | |
| 6 | | 31 | | 56 | | 81 | |
| 7 | | 32 | | 57 | | 82 | |
| 8 | | 33 | | 58 | | 83 | |
| 9 | | 34 | | 59 | | 84 | |
| 10 | | 35 | | 60 | | 85 | |
| 11 | | 36 | | 61 | | 86 | |
| 12 | | 37 | | 62 | | 87 | |
| 13 | | 38 | | 63 | | 88 | |
| 14 | | 39 | | 64 | | 89 | |
| 15 | | 40 | | 65 | | 90 | |
| 16 | | 41 | | 66 | | 91 | |
| 17 | | 42 | | 67 | | 92 | |
| 18 | | 43 | | 68 | | 93 | |
| 19 | | 44 | | 69 | | 94 | |
| 20 | | 45 | | 70 | | 95 | |
| 21 | | 46 | | 71 | | 96 | |
| 22 | | 47 | | 72 | | 97 | |
| 23 | | 48 | | 73 | | 98 | |
| 24 | | 49 | | 74 | | 99 | |
| 25 | | 50 | | 75 | | 100 | |

識をしたのではないかと考えられる。また同じ被験者の方に協力していただき、マイクとの距離を変更して行った音声認識実験では「禁煙です」、「喰煙です」、「吸います」、「吸いません」、「吸わないです」の認識が成功した数が減少したのは input rejected by short input という認識不可を示す結果が出たためである。「吸いません」の音声認識実験を行った際に認識不可以外にも「喰煙です」と認識されることがあった。

表 4.3 吸いますの認識結果

| 回数 | 吸います | 回数 | 吸います | 回数 | 吸います | 回数 | 吸います |
|----|------|----|------|----|------|-----|------|
| 1 | | 26 | | 51 | | 76 | |
| 2 | | 27 | | 52 | | 77 | |
| 3 | | 28 | | 53 | | 78 | |
| 4 | | 29 | | 54 | | 79 | |
| 5 | | 30 | | 55 | | 80 | |
| 6 | | 31 | | 56 | | 81 | |
| 7 | | 32 | | 57 | | 82 | |
| 8 | | 33 | | 58 | | 83 | |
| 9 | | 34 | | 59 | | 84 | |
| 10 | | 35 | | 60 | | 85 | |
| 11 | | 36 | | 61 | | 86 | |
| 12 | | 37 | | 62 | | 87 | |
| 13 | | 38 | | 63 | | 88 | |
| 14 | | 39 | | 64 | | 89 | |
| 15 | | 40 | | 65 | | 90 | |
| 16 | | 41 | | 66 | | 91 | |
| 17 | | 42 | | 67 | | 92 | |
| 18 | | 43 | | 68 | | 93 | |
| 19 | | 44 | | 69 | | 94 | |
| 20 | | 45 | | 70 | | 95 | |
| 21 | | 46 | | 71 | | 96 | |
| 22 | | 47 | | 72 | | 97 | |
| 23 | | 48 | | 73 | | 98 | |
| 24 | | 49 | | 74 | | 99 | |
| 25 | | 50 | | 75 | | 100 | |

表 4.4 吸いませんの認識結果

| 回数 | 吸いません | 回数 | 吸いません | 回数 | 吸いません | 回数 | 吸いません |
|----|-------|----|-------|----|-------|-----|-------|
| 1 | | 26 | | 51 | | 76 | |
| 2 | | 27 | | 52 | | 77 | |
| 3 | | 28 | | 53 | | 78 | |
| 4 | | 29 | | 54 | | 79 | |
| 5 | | 30 | | 55 | | 80 | |
| 6 | | 31 | | 56 | | 81 | |
| 7 | | 32 | | 57 | | 82 | |
| 8 | | 33 | | 58 | | 83 | |
| 9 | | 34 | | 59 | | 84 | |
| 10 | | 35 | | 60 | | 85 | |
| 11 | | 36 | | 61 | | 86 | |
| 12 | | 37 | | 62 | | 87 | |
| 13 | | 38 | | 63 | | 88 | |
| 14 | | 39 | | 64 | | 89 | |
| 15 | | 40 | | 65 | | 90 | |
| 16 | | 41 | | 66 | | 91 | |
| 17 | | 42 | | 67 | | 92 | |
| 18 | | 43 | | 68 | | 93 | |
| 19 | | 44 | | 69 | | 94 | |
| 20 | | 45 | | 70 | | 95 | |
| 21 | | 46 | | 71 | | 96 | |
| 22 | | 47 | | 72 | | 97 | |
| 23 | | 48 | | 73 | | 98 | |
| 24 | | 49 | | 74 | | 99 | |
| 25 | | 50 | | 75 | | 100 | |

表 4.5 吸わないですの認識結果

| 回数 | 吸わないです | 回数 | 吸かないです | 回数 | 吸わないです | 回数 | 吸わないです |
|----|--------|----|--------|----|--------|-----|--------|
| 1 | | 26 | | 51 | | 76 | |
| 2 | | 27 | | 52 | | 77 | |
| 3 | | 28 | | 53 | | 78 | |
| 4 | | 29 | | 54 | | 79 | |
| 5 | | 30 | | 55 | | 80 | |
| 6 | | 31 | | 56 | | 81 | |
| 7 | | 32 | | 57 | | 82 | |
| 8 | | 33 | | 58 | | 83 | |
| 9 | | 34 | | 59 | | 84 | |
| 10 | | 35 | | 60 | | 85 | |
| 11 | | 36 | | 61 | | 86 | |
| 12 | | 37 | | 62 | | 87 | |
| 13 | | 38 | | 63 | | 88 | |
| 14 | | 39 | | 64 | | 89 | |
| 15 | | 40 | | 65 | | 90 | |
| 16 | | 41 | | 66 | | 91 | |
| 17 | | 42 | | 67 | | 92 | |
| 18 | | 43 | | 68 | | 93 | |
| 19 | | 44 | | 69 | | 94 | |
| 20 | | 45 | | 70 | | 95 | |
| 21 | | 46 | | 71 | | 96 | |
| 22 | | 47 | | 72 | | 97 | |
| 23 | | 48 | | 73 | | 98 | |
| 24 | | 49 | | 74 | | 99 | |
| 25 | | 50 | | 75 | | 100 | |

表 4.6 認識結果一覧

| 単語 | 禁煙です | 喫煙です | 吸います | 吸いません | 吸わないです |
|----|------|------|------|-------|--------|
| 計 | 99 | 100 | 100 | 30 | 100 |
| 平均 | 0.99 | 1 | 1 | 0.3 | 1 |

音声認識実験

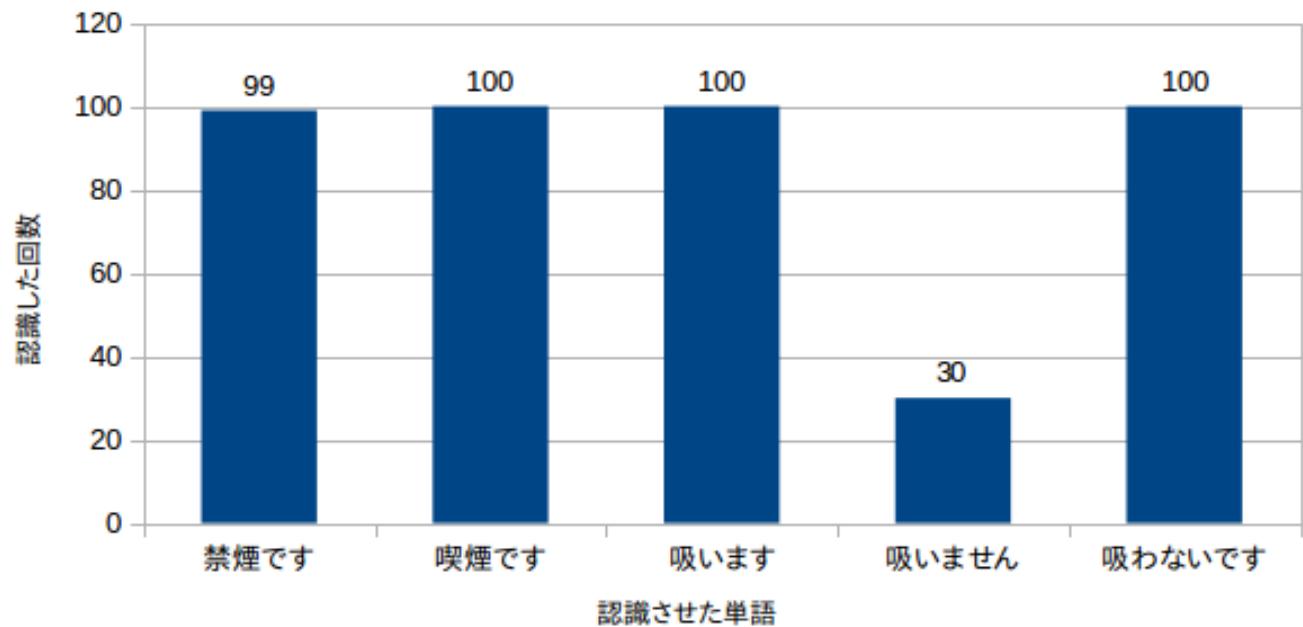


図 4.1 認識結果グラフ

音声認識実験

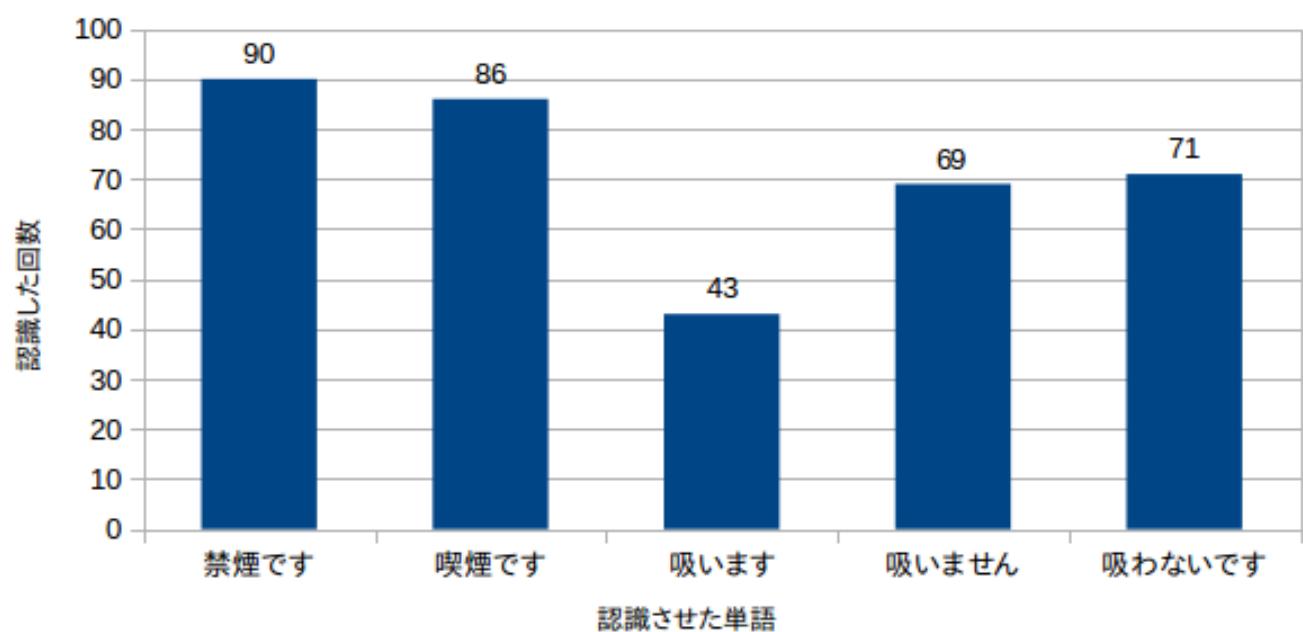


図 4.2 別条件での認識結果グラフ

表 4.7 別条件での禁煙の認識結果

| 回数 | 禁煙です | 回数 | 禁煙です | 回数 | 禁煙です | 回数 | 禁煙です |
|----|------|----|------|----|------|-----|------|
| 1 | | 26 | | 51 | | 76 | × |
| 2 | | 27 | | 52 | | 77 | |
| 3 | × | 28 | | 53 | | 78 | |
| 4 | × | 29 | | 54 | | 79 | |
| 5 | | 30 | | 55 | | 80 | |
| 6 | | 31 | | 56 | | 81 | |
| 7 | | 32 | | 57 | | 82 | |
| 8 | | 33 | | 58 | | 83 | |
| 9 | | 34 | | 59 | × | 84 | |
| 10 | | 35 | | 60 | | 85 | × |
| 1 | | 36 | | 61 | | 86 | × |
| 12 | | 37 | | 62 | | 87 | × |
| 13 | | 38 | | 63 | | 88 | |
| 14 | | 39 | | 64 | | 89 | |
| 15 | | 40 | | 65 | | 90 | × |
| 16 | | 41 | | 66 | | 91 | |
| 17 | | 42 | | 67 | | 92 | |
| 18 | | 43 | | 68 | | 93 | |
| 19 | | 44 | | 69 | | 94 | |
| 20 | | 45 | | 70 | | 95 | |
| 21 | | 46 | | 71 | | 96 | |
| 22 | | 47 | | 72 | | 97 | |
| 23 | | 48 | | 73 | | 98 | |
| 24 | | 49 | | 74 | | 99 | |
| 25 | | 50 | | 75 | × | 100 | × |

表 4.8 別条件での喫煙の認識結果

| 回数 | 喫煙です | 回数 | 喫煙です | 回数 | 喫煙です | 回数 | 喫煙です |
|----|------|----|------|----|------|-----|------|
| 1 | | 26 | | 51 | × | 76 | |
| 2 | | 27 | | 52 | × | 77 | |
| 3 | | 28 | | 53 | × | 78 | |
| 4 | | 29 | | 54 | × | 79 | |
| 5 | | 30 | | 55 | | 80 | |
| 6 | | 31 | | 56 | | 81 | |
| 7 | | 32 | | 57 | | 82 | |
| 8 | | 33 | × | 58 | | 83 | |
| 9 | | 34 | × | 59 | × | 84 | |
| 10 | | 35 | | 60 | | 85 | |
| 11 | | 36 | | 61 | | 86 | |
| 12 | | 37 | | 62 | | 87 | |
| 13 | | 38 | | 63 | | 88 | |
| 14 | | 39 | × | 64 | | 89 | |
| 15 | | 40 | | 65 | | 90 | |
| 16 | | 41 | | 66 | | 91 | |
| 17 | | 42 | | 67 | | 92 | |
| 18 | | 43 | | 68 | | 93 | |
| 19 | | 44 | | 69 | | 94 | |
| 20 | | 45 | | 70 | | 95 | |
| 21 | × | 46 | | 71 | | 96 | |
| 22 | | 47 | | 72 | | 97 | × |
| 23 | | 48 | | 73 | | 98 | |
| 24 | | 49 | × | 74 | | 99 | × |
| 25 | | 50 | × | 75 | | 100 | × |

表 4.9 別条件での吸いますの認識結果

| 回数 | 吸います | 回数 | 吸います | 回数 | 吸います | 回数 | 吸います |
|----|------|----|------|----|------|----|------|
| 1 | × | 26 | × | 51 | × | 76 | × |
| 2 | | 27 | | 52 | | 77 | × |
| 3 | × | 28 | | 53 | × | 78 | |
| 4 | × | 29 | | 54 | × | 79 | |
| 5 | × | 30 | × | 55 | × | 80 | |
| 6 | × | 31 | | 56 | × | 81 | |
| 7 | × | 32 | × | 57 | × | 82 | |
| 8 | × | 33 | | 58 | | 83 | |
| 9 | | 34 | × | 59 | × | 84 | |
| 10 | × | 35 | | 60 | × | 85 | |
| 11 | | 36 | × | 61 | × | 86 | |
| 12 | | 37 | × | 62 | | 87 | |
| 13 | × | 38 | | 63 | × | 88 | |
| 14 | × | 39 | | 64 | × | 89 | |
| 15 | | 40 | | 65 | × | 90 | × |
| 16 | × | 41 | | 66 | × | 91 | × |
| 17 | × | 42 | | 67 | × | 92 | × |
| 18 | | 43 | × | 68 | × | 93 | × |
| 19 | | 44 | × | 69 | | 94 | × |
| 20 | | 45 | × | 70 | × | 95 | × |
| 21 | × | 46 | | 71 | | 96 | × |
| | | | | | | | |

表 4.10 別条件での吸いませんの認識結果

| 回数 | 吸いません | 回数 | 吸いません | 回数 | 吸いません | 回数 | 吸いません |
|----|-------|----|-------|----|-------|-----|-------|
| 1 | | 26 | | 51 | | 76 | × |
| 2 | | 27 | | 52 | × | 77 | |
| 3 | | 28 | | 53 | | 78 | × |
| 4 | | 29 | | 54 | | 79 | |
| 5 | | 30 | | 55 | | 80 | × |
| 6 | | 31 | | 56 | × | 81 | × |
| 7 | × | 32 | | 57 | | 82 | × |
| 8 | | 33 | | 58 | × | 83 | |
| 9 | | 34 | | 59 | | 84 | × |
| 10 | | 35 | | 60 | | 85 | × |
| 11 | | 36 | | 61 | × | 86 | |
| 12 | | 37 | | 62 | | 87 | |
| 13 | | 38 | × | 63 | × | 88 | |
| 14 | | 39 | | 64 | | 89 | |
| 15 | | 40 | × | 65 | | 90 | |
| 16 | × | 41 | × | 66 | | 91 | |
| 17 | × | 42 | | 67 | | 92 | × |
| 18 | | 43 | × | 68 | × | 93 | |
| 19 | | 44 | × | 69 | | 94 | |
| 20 | | 45 | | 7× | | 95 | |
| 21 | | 46 | × | 71 | × | 96 | |
| 22 | × | 47 | | 72 | × | 97 | |
| 23 | | 48 | | 73 | × | 98 | |
| 24 | × | 49 | | 74 | × | 99 | |
| 25 | × | 50 | | 75 | × | 100 | |

表 4.11 別条件での吸わないですの認識結果

| 回数 | 吸わないです | 回数 | 吸かないです | 回数 | 吸わないです | 回数 | 吸わないです |
|----|--------|----|--------|----|--------|-----|--------|
| 1 | | 26 | | 51 | | 76 | × |
| 2 | | 27 | | 52 | × | 77 | |
| 3 | | 28 | | 53 | × | 78 | × |
| 4 | | 29 | | 54 | × | 79 | |
| 5 | | 30 | | 55 | × | 80 | |
| 6 | | 31 | | 56 | | 81 | |
| 7 | | 32 | | 57 | | 82 | |
| 8 | | 33 | | 58 | × | 83 | |
| 9 | | 34 | | 59 | | 84 | |
| 10 | | 35 | | 60 | | 85 | × |
| 11 | | 36 | | 61 | | 86 | × |
| 12 | | 37 | | 62 | | 87 | × |
| 13 | | 38 | | 63 | × | 88 | × |
| 14 | | 39 | | 64 | × | 89 | × |
| 15 | | 40 | | 65 | | 90 | × |
| 16 | | 41 | | 66 | | 91 | × |
| 17 | | 42 | | 67 | | 92 | |
| 18 | | 43 | | 68 | × | 93 | |
| 19 | | 44 | | 69 | | 94 | × |
| 20 | | 45 | × | 70 | | 95 | |
| 21 | | 46 | | 71 | × | 96 | × |
| 22 | | 47 | × | 72 | | 97 | × |
| 23 | | 48 | | 73 | | 98 | × |
| 24 | | 49 | × | 74 | × | 99 | |
| 25 | | 50 | × | 75 | × | 100 | × |

表 4.12 別条件での認識結果一覧

| 単語 | 禁煙です | 喫煙です | 吸います | 吸いません | 吸わないです |
|----|------|------|------|-------|--------|
| 計 | 90 | 86 | 43 | 69 | 71 |
| 平均 | 0.9 | 0.86 | 0.43 | 0.69 | 0.71 |

第 5 章

結言

5.1 本研究のまとめ

音声を認識する接客ロボットのシステムを構築するにあたり Julius を用いることで音声認識を行うことが可能となった。また、音声の認識結果より平均を算出することで Julius の有用性を確かめることができた。今回は音声を作成するのみとなったが音声を認識した後に発声させる予定だった音声を WAV ファイルとして作成し、保存することができた。これによって音声を認識した後に対応した音声を発声させることが出来る。

5.2 今後の課題

今後は研究者だけでなく研究者以外の人間を実験対象とし、認識率を算出する。また音声を認識したあとに認識した単語に対応した音声を発生させるプログラムを TCP/IP を介して組み合わせることにより、音声を認識し、認識した結果より対応した音声を発声させるためのシステムとして扱うことできるシステムの確立がなされるのではないかと考える。また、このシステムを用いることで音声認識を用いた接客ロボットの開発を進めることができるのではないかと考える。

.1 今回使用したもの

yomi ファイルとは認識させるために準備した単語を書き込んだものである。今回使用した yomi ファイルを図 1 に示す。voca ファイルとはカテゴリごとに単語の表記と読み（音素列）を登録する [1]。今回使用した voca ファイルを図 2 に示す。grammar ファイルとは構文制約を単語のカテゴリを終端規則として記述したファイルである。[1] 今回使用した grammar ファイルを図 3 に示す。dfa ファイルと

はオートマトン定義ファイルであり，grammar で記述される文法制約を有限個の状態と状態遷移のある振る舞いの抽象的なモデルに変換したものである.[1] 今回使用した dfa ファイルを図 4 に示す.term ファイルとは単語カテゴリ番号と実際のカテゴリ名の対応をさせるために用いるファイルである [2]. 今回使用した term ファイルを図 5 に示す.dic ファイルとは単語を認識させるための辞書ファイルである. 後述の dict ファイルとの相違点は jconf ファイルに指定する単語を認識させるためである. 今回使用した dic ファイルを図 6 に示す.dict ファイルとは単語を認識させるための辞書ファイルである.[3] dic ファイルとの相違点は dfa ファイルと連動させる点にある. 今回使用した dict ファイルを図 7 に示す.

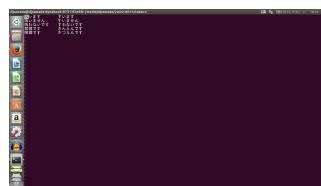


図 1 yomi ファイル

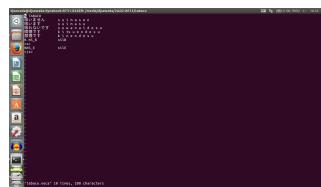


図 2 voca ファイル

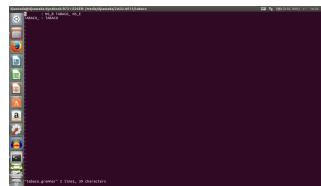


図 3 grammar ファイル

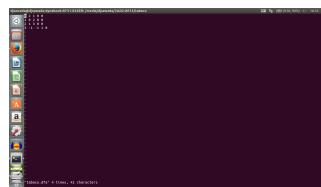


図 4 dfa ファイル

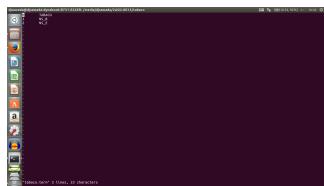


図 5 term ファイル

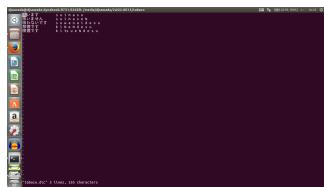


図 6 dic ファイル

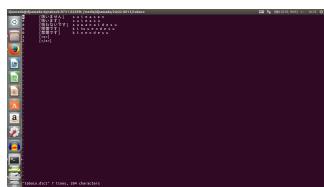


図 7 dict ファイル

参考文献

- [1] ディジタルアナログ信号処理のためのやさしいフーリエ変換 松尾 博 著 森北出版株式会社 1986 年
3月 31 日出版 P4-P5
- [2] Julius book 第 7 章 言語モデル・https://julius.osdn.jp/juliusbook/Ja/desc_lm.html・2016/09/02 閲覧
- [3] Julius book 第 8 章 認識アルゴリズムとパラメータ https://julius.osdn.jp/juliusbook/ja/desc_search.html?id=253869
閲覧

謝辞

本論文作成にあたりテーマの決定, 研究の考え方, 方法のまとめ方など全てにおいて長期にわたって厳しくも熱意のあるご指導, ご鞭撻していただいた, 伊藤恒平教授に厚く御礼申し上げます.

特に分析においても論文の書き方においても論文を何度も読んでいただき, 指導していただいた伊藤恒平教授に大変ご苦労をかけてしましましたことにも心よりお詫び申し上げたいです.

その他, 助けていただいた多くの皆様に心から感謝しております. ありがとうございました.