

クラウド型音声認識 API を用いて適切な話速を定量的に評価・改善するセルフチェックサービス

杉山 渉[†] 中村 亮太[‡] 上林 憲行[†]
 東京工科大学[†] 湘北短期大学[‡]

1. はじめに

厚生労働省と文部科学省の調査により、2013 年 10 月における大卒者の就職内定率は 64.3%と前年の同時期よりも 1.2 ポイント上回ったと発表されたが、現在も厳しい状態が続いている[1]。書籍やインターネット等で就職活動の面接試験への対策について述べられているが、面接時の回答内容に関する記述がほとんどである。身体動作や話し方などの非言語行動に関しては具体的に述べられていることは極めて少ない。したがって面接受験者は回答内容を向上させることばかりに力を注ぐ傾向にあるが、先行研究[2]によると初対面の相手の印象を形成する要因は、視線行動、身だしなみ、話す内容、表情、話し方の 6 つの要因があり、その中で本研究では話し方に着目している。話し方の構成要素には発話する速度（以下、話速）、声の大きさ、「えー」、「あー」などの音声フィラー、繰り返し、沈黙があり、実験結果から話速が聞き手の印象に大きな影響を与えることが明らかになっている[3]。

本論文では、就職活動を控えた者を対象にクラウド型音声認識 API を用いて話速を定量的に評価・改善するセルフチェックサービス「Wasokun」を開発及び評価実験について述べる。

2. 話速の定量的評価

話速が速すぎると聞き手に「落ち着きがない」、「威圧的である」といった印象を与え、一方で話速が遅すぎると退屈感を与え、印象を悪くしてしまうことが従来研究により報告されている。対面対話においてどの程度の話速が聞き手に良い印象を与えるのかを定量的に調査するために以下の予備実験を実施した。

就職採用面接における受験者の話速を評価するために学生同士で行う模擬面接を実施した。受験者は 14 名であり、学部 3 年生 14 名(男女比は男子学生 12 名に対し、女子学生 2 名)という構成である。なお面接官役を担当した学生は模擬面接官養成講座にてトレーニングを積んだ学部 4 年生 2 名である。

発話した文字数を数えるためにビデオ撮影した模擬面接を視聴しながら発話内容を平仮名で書き出した。話速は発話した文字数を発話時間で割って算出している。今回の実験では自己 PR の発話内容を対象とした話速を算出している。採用面接では面接官から最初の質問の多くは自己 PR なので用いた。話速の評価方法は聞き取りやすさを 4 段階の主観評価で回答させた。回答は面接官が担っている。

予備実験の結果、話速が 330 字/分に以下および 450 字/分以上の場合、評価点が低くなることが分かった。聞き手に良い印象を与える基準を 3 以上としたとき、話速が 340 から 440(字/分)であることから、本研究では上記の範囲を適切な話速としてサービスを構築することにした(図 1)。

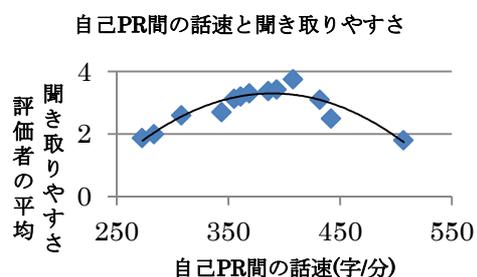


図 1 話速を定量的評価・計測

3. サービスの実装

Wasokun の実装について述べる。以下の 4 つの機能を組み合わせたサービスを開発した(図 2)。

- ① 発話記録(WebSpeechAPI)
- ② 平仮名変換 (YahooAPI)
- ③ 時間計測(ストップウォッチ)
- ④ 話速計測

1 の「発話記録」は、Google が無料で提供しているクラウド型音声認識 API であり、発話内容を Web 上で記録させる。

2 の「平仮名変換」は 1 で出力された文字を、YahooAPI を用いて平仮名変換を行う。以前まで模擬面接動画を視聴し発話内容を書き出していたので自動で実装できるようにするため用いた。

3 の「時間計測」は話速を求める上で、発話時間を記録する必要があるため用いた。

4 の「話速計測」は 2 と 4 から得られるデータをもとに話速を計測し、定量的数値をもとに評価した。

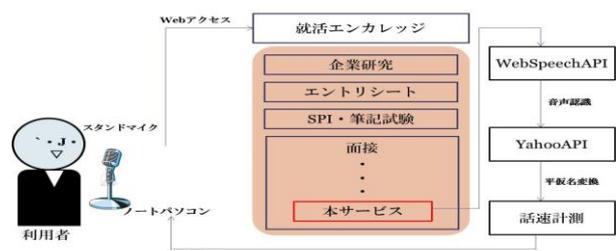


図 2 サービスコンセプト

“Self-check service to evaluate and improve quantitatively ad measurement of appropriate rate by introducing speech recognition technology”

[†]WATAYARU SUGIYAMA, [‡]RYOTA NAKAMURA, [†]NORIYUKI KAMIBAYASHI

[†]Tokyo University of Technology, [‡]Shohoku College

4. 模擬面接実験実施

4.1. 模擬面接概要

模擬面接とは、11月末日に本学主催で行われた模擬面接会のことである。模擬面接会の受験者は14名となっており、学部3年生14名(男子14名、女子3名)という構成である。面接官2名、受験者1名、観察者1名からなる個人面接形式となっている。面接評価者は事前に本学主催の模擬面接間養成講座を受講した学部4年生が担っている。模擬面接は合計2回行い、第2回目は1回目の模擬面接から1週間後に実施した。1回目と2回目の間に、Wasokunを利用してもらい、話速を比較し改善されたか分析した。

4.2. 実験結果

第1回目と第2回目の自己PR間の話速を計測し、2回目で改善された割合を出したのが図3で示している。なお、第1回目で話速が適切な受験者は14名中7名いたためデータから除いた。実験結果から、第1回目から話速の改善が診られなかった割合が67%、話速が改善されたのが33%である。

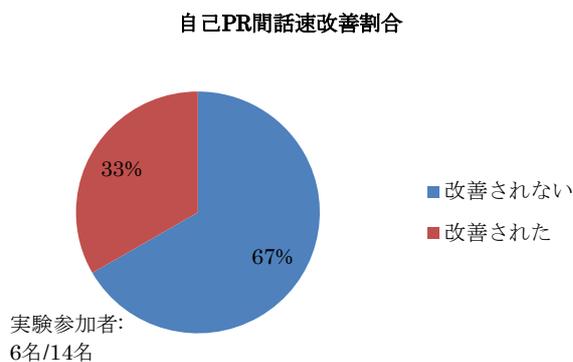


図3 話速改善割合

4.3. Wasokun 利用継続時の話速変化

模擬面接実施結果から、話速が適切にならない受験者が67%いるということで、本サービス利用時は話速が適切になっているのか、示したのが図4である。サービス利用継続時は、話速が改善傾向にあることが判明した。

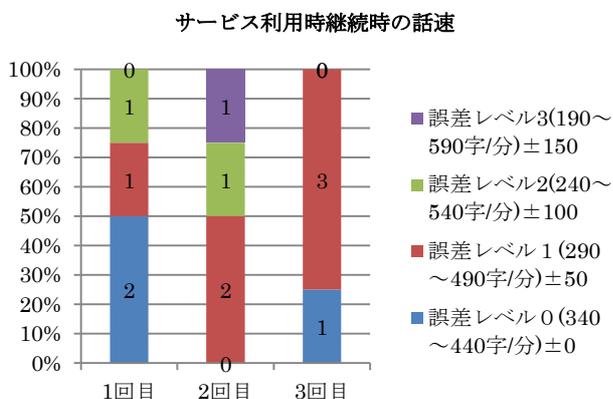


図4 サービス利用継続時の話速

5. 考察

Wasokun 利用時は話速が改善傾向にあっても、実際の面接では、緊張してしまい、話速が速くなり、頭が真っ白になり沈黙時間が増えて話速が遅くなる。本サービスは画面に向かって、発話するだけなので相手との対面式ではないため、本番同様の緊張感を持ったトレーニングが出来ないということが考察できる。

話速が改善出来れば、自然と話し方の構成要素である音声フィラーや、沈黙時間、繰り返しが減るのではないかと考えられる。話速が適切であると、音声フィラーや、沈黙時間、繰り返しが少ないということは実験データから示唆される。話し方の印象値を上げるためには、まず話速から直すことが今後期待される。

Wasokun 利用時には話速が改善傾向にあったので、サービス利用回数を重ねることで、実際の面接でも改善が見込めると考えられる。

5.1. Wasokun の今後

現状、話速しか計測出来ないが、話し方の構成要素である音声フィラー、沈黙時間、繰り返しなど計測出来れば、話速だけを改善するサービスでなく、話し方を改善できるサービスになると考えられる。そのためにはまず、それぞれの要素の適切な範囲を求める必要がある。今回の話速も定量的にして、適切な範囲を出しているのと同様に実験して、定量的に表し、適切な範囲を決めることができれば、実現は可能かもしれない。

6. おわりに

本論文では面接官が重視する非言語行動の中で話し方について着目し、話し方の構成要素の1つである、話速を定量的に評価し、改善できるセルフチェックサービスの開発「Wasokun」である。

模擬面接実施による結果、話速が適切になった割合は33%である。話速が適切にならなかった人が、67%である。サービス利用継続時は、話速が改善傾向にあるが、サービス利用時でも、話速が適切になっていない。

利用回数を重ねることで話速が改善傾向にあるということが示唆される。

今後の課題として、Wasokun の機能が話速を計測出来るだけでなく、話し方の構成要素である、音声フィラー、沈黙時間、繰り返しなども計測出来るより、話し方そのものを改善できるサービスが実用化されることを望む。

参考文献

- [1]時事ドットコム, 大卒の就職率と内定率
http://www.jiji.com/jc/graphics?p=ve_soc_tyosa-koyou-college-naitei (2013/12/19 参照)
- [2] 籠宮隆之, 音声を手掛かりにした対人印象形成にかかわる要因の研究
<http://www.lib.kobe-u.ac.jp/repository/thesis/d1/D1004289.pdf> (2013/12/19 参照)
- [3] 渡辺智美: “採用面接における非言語行動の印象改善方法の提案—話速改善による面接評価への影響調査と分析—” (2013/11/1 参照)