

# 傾聴システムにおける韻律特徴を活用した非言語的な共感応答の開発

徳田 裕紀<sup>†</sup> 佐伯 幸郎<sup>††</sup> 中村 匡秀<sup>†,†††</sup>

<sup>†</sup> 神戸大学 〒657-8501 神戸市灘区六甲台町 1-1

<sup>††</sup> 高知工科大学 〒782-8502 高知県香美市土佐山田町宮ノ口 185

<sup>†††</sup> 理化学研究所・革新知能統合研究センター 〒103-0027 東京都中央区日本橋 1-4-1

E-mail: <sup>†</sup>{tokup,masa-n}@cs.kobe-u.ac.jp, <sup>††</sup>sachio@carp.kobe-u.ac.jp

**あらまし** バーチャルエージェントを活用した傾聴カウンセリングシステムでは、発話者へ共感的応答をすることが重要である。これまでに言語的な共感表出に関しては検討されているが、非言語的な共感表出は考慮できていない。本稿では、傾聴カウンセリングシステムに対して、非言語的な共感表出実現に向けた機能拡張を行う。提案手法では、発話者の音声を元に韻律的に同調した相槌を生成する。また、提案システムが共感応答を出来ているかシステムとの対話を通じて確認している。

**キーワード** 在宅介護, バーチャルエージェント, 積極的傾聴, 共感応答

## Development of nonverbal empathic response utilizing prosodic features for the active listening agent

Yuki TOKUDA<sup>†</sup>, Sachio SAIKI<sup>††</sup>, and Masahide NAKAMURA<sup>†,†††</sup>

<sup>†</sup> Kobe University Rokkodai-cho 1-1, Nada-ku, Kobe, Hyogo, 657-8501 Japan

<sup>††</sup> Kochi Institute of Technology Miyanokuchi 185, Tosayamada-cho, Kami-shi, Kochi, 782-8502 Japan

<sup>†††</sup> Riken AIP, 1-4-1 Nihon-bashi, Chuo-ku, Tokyo 103-0027

E-mail: <sup>†</sup>{tokup,masa-n}@cs.kobe-u.ac.jp, <sup>††</sup>sachio@carp.kobe-u.ac.jp

**Abstract** In the active listening counseling system that utilizes a virtual agent, it is important to respond empathically to the speaker. So far, verbal empathic response has been examined, but nonverbal empathic response has not been considered. In this paper, we will extend the functions of the active listening counseling system to realize nonverbal empathic response. In the proposed method, prosodically synchronized responses is generated based on the speaker's voice prosody. In addition, it is confirmed through dialogue with the system whether the proposed system is able to respond appropriately to empathy.

**Key words** Home care, Virtual Agent, Active listening, empathic response

### 1. はじめに

現在、日本は超高齢化社会に直面している。厚生労働省の調査によると、総人口は減少し続けるにもかかわらず、65歳以上の高齢者数は上昇し続け、2036年には高齢者の割合が33.3%になると予測されている[1]。高齢者の増加に伴い認知症を持つ人の数も増加し、高齢者や認知症当事者のための有効かつ持続的な在宅での支援が必要とされている。

認知症の症状に対する非薬物療法としては、患者一人一人の立場に立ってその人を理解し、その人に寄り添った支援を行うパーソンセンタードケア[2]がある。パーソンセンタードケアの一つとして、積極的傾聴が知られている。積極的傾聴は、「話を

聞いていること」(一致)「共感していること」(共感的理解)「相手を受け入れていること」(受容)を相手に示しながら話を聴くことで、受容の対応を示し、自分を受け入れてくれているという安心感・自己肯定感によって話者(クライアント)の心理的安定を図る手法である。積極的傾聴はパーソンセンタードケアの中でも、自分を見失い不安に陥る認知症当事者には効果的であるとされている。

積極的傾聴用いたケアではケア提供者とクライアントとの継続的な会話が重要となる。しかし、専門医が日常的にカウンセリングを行うのは経済的・時間的に困難である。また、今日では在宅介護への移行が進んでおり、家族の介護負担も増えている。よって会話やカウンセリングに十分な時間を費やせない状

況にある。

そこで我々の研究グループでは、音声対話が可能なロボットプログラムであるバーチャルエージェント (VA) を活用して、在宅でクライアントがいつでも傾聴対話のコミュニケーションを行えるケアシステムを開発している [3]。この傾聴ケアシステムでは、人感センサーがクライアントを感知し定期的に、ないしはクライアント側から呼びかけることによって、クライアントが話を行い VA が相槌を返すという傾聴対話を行い、これによって継続的な傾聴ケアを実現している。

しかしながら、現段階のシステムでは相手の応答に関わらず汎用的な相槌・動作をランダムに返しており、積極的傾聴に重要となる共感性が不足している。「なるほど」「そうなんですな」などの汎用的な相槌は相手の発話内容に関わらず一定の受容を示すことが可能であり、現状のシステムにおいても傾聴に重要な「話を聞いていること」「受け入れていること」をある程度満たすことは可能である。しかし「共感すること」に関しては満たしておらず、共感性の不足はクライアントが自分の感じていることに自信を持てなくなり、自分への否定の恐怖に繋がる [4]。そのため傾聴システムを実現するために共感応答は必要な機能である。

そこで本研究では、より高齢者の感情に寄り添った傾聴の実現を目的とし、共感応答機能を有する傾聴ケアシステムを提案する。共感を示す方法としては、言語的に返答の内容によってクライアントの立場・感情の状況への理解を示す方法と、クライアントが発露した感情と同一のそれを表情や、声の韻律特徴などを通して発露させることで示す方法の二種類が存在する [5]。ケア提供者とクライアント間の感情的な繋がりや、クライアント側の情報開示への精神的敷居の引き下げには後者のほうが重要であり [6]、また前者においては言語の幅広さ故、返答の一般化・システム化が困難であることから今回は非言語的な共感応答を生成する手法を提案する。

より具体的な方法として、まず同調して有効とされる一部の韻律特徴を取得する。またこれに加え、韻律特徴とテキスト情報を組み合わせた感情解析によって発話において表出した感情の大きさとその方向性（ポジティブな感情であるかネガティブな感情であるか）を取得する。取得したこれらの特徴を用い、発話者と同じ韻律・感情発露した相槌を生成する。

本稿では、既存解析手法の実現性の確認とパラメータの検討のため感情解析手法の検証を予備実験として行う。また、提案手法に基づいた応答による傾聴システムが、従来のシステムよりも共感性の高い応答が実現できているか確認するため、被験者実験による客観評価を行う。感情解析の予備実験においては、あらかじめ感情ラベルを付与した音声テストデータを用意し、これらを感情解析、結果から今後のシステム設計において望ましい閾値を検討する。共感応答の有効性の確認においては、あらかじめ対話内容を決められたシステムとの傾聴対話を録画し、それを第三者によって視聴してもらい、アンケートを通じてシステムの印象評価を行う。対話内容として既存システムを想定した応答と提案手法を想定した応答の対話をそれぞれ用意し、これによりベースラインである既存システムと提案手法の傾聴



図 1 傾聴対話の様子

対話の印象を比較評価した。分析の結果、提案手法による対話の方がより話者の感情に寄り添っているという印象を与えるということが示され、傾聴システムにおいて共感応答の有効性が示された。

## 2. 準備

### 2.1 在宅認知症のための傾聴ケアシステム

我々の研究グループでは、在宅の高齢者や認知症当事者に対して、音声対話が可能なロボットプログラム VA (バーチャルエージェント) を用いた傾聴ケアシステムを開発している。図 1 に傾聴ケアシステムで実際に対話している様子を示す。傾聴ケアシステムは、クライアントが日常的にその日の出来事や気分を VA に話し、VA がそれに対して相槌を返すことで傾聴を行う。

具体的には、まずクライアントが傾聴をしてもらおうとシステムに話しかける、ないしは人感センサーが反応しかつ正午や午後 7 時などの特定の時間になった際に、VA 傾聴システムがその日食事、あった出来事、気分など日常の話題を問いかける。問いかける話題は、特定の時間での傾聴の場合はその時間に合わせた「お昼ご飯は何を食べましたか?」や「朝の出来事を教えてください」などの話題が、クライアントから話しかけた場合は「私に何でも話して下さい」という汎用的な話題が選択される。これに対しクライアントがその話題に関して自由に発言し、発話が一定時間以上止まるごとにそこまでの音声の一つの発話として区切る。そしてその発話に対し VA が汎用的な相槌・動作をランダムに返す。その後、引き続きクライアントが自由に発言する。これをクライアントが傾聴を終了する旨の発言をするまで繰り返す。

これにより時間によらない認知症当事者のケアを実現でき、人間の介護者の負担を軽減することが期待できる。またこれらの傾聴対話中のクライアントの音声および傾聴中のクライアントの動画を継続的にデータベースに保存することで、クライアントの日々の精神状態をモニタリングし、それらのデータを人間の手によるケアへの補助として用いることも期待できる。

### 2.2 共感応答

共感とは 2 人の人物間で一方が他方の感情と同一の感情を体験することであり、共感応答とはそのように共感したことを言葉、表情、動作などによって相手に示すことである。共感応答は



表 1 音声テストデータの一例

感情ラベル	テキスト内容
中立	「そうですね、今日は昼に散歩をしました」
ポジティブ弱	「そうですね今日は気分がいいですよ」
ポジティブ強	「いや景色がきれいすぎてもう興奮しちゃって！」
ネガティブ弱	「お墓参りにも行けなくなってしまいましたね」
ネガティブ強 (悲)	「なんで私こんなのだろうって思っちゃって…」
ネガティブ強 (怒)	「話ちゃんと聞いている?? 怒らせたいの?」
自虐的発言	「その時私本当に頭悪くて (笑)」

パラメータを調整することによって生成する。また“感情の方向性”においてポジティブないしはネガティブの特徴が判定がされている場合、VA の動作においてポジティブな発話に対しては笑顔を浮かべる、ネガティブな発話には悲しい表情をするなどの同調するような動作を合わせて応答する。

#### 4. 評価実験

本節では提案システムの実現性を確認する 2 種類の予備実験を行う

##### 4.1 既存の感情解析手法の実現性検証

提案システムにおける、既存の感情解析手法の実現性の確認とパラメータ検討のために感情解析手法の検証を予備実験として行う。あらかじめ“ネガティブ強 (悲しみ)”, “ネガティブ強 (怒り)”, “ネガティブ弱”, “中立”, “ポジティブ弱”, “ポジティブ強”, “自虐的発言” の 7 種類の感情ラベルを付与した音声テストデータを用意し、これらを感情解析し、結果から今後のシステム設計において望ましい閾値を検討する。テストデータの一例を表 1 に示す。今回、韻律の感情解析には python の myprosody [12] を、テキスト感情解析には AzureTextAnalytics [13] を利用した。韻律の解析では感情の方向性が「Passionately」「Reading」「Normal」の 3 値で返され、それぞれを「ポジティブ」「中立」「ネガティブ」な感情の結果として扱う。テキストの解析では感情の大きさが 0 (ネガティブ) 1 (ポジティブ) の値で返される。

##### 4.2 感情解析手法の検証結果

テストデータに対しテキスト感情解析を行ったスコアの、各感情ラベルごとの集計結果を図 3 に箱ひげ図で示す。またテストデータに対し韻律感情解析を行った結果に対し、各感情ラベルごとに集約し、解析で得られた結果のラベルを割合で示したものを図 4 に示す。

まずテキスト感情解析においては、ポジティブの音声の判定値の第一四分位数がネガティブの音声の判定値の第三四分位より小さく、2つの解析結果はともに 7 割 5 分以上が重ならないという結果が得られた。また自虐の音声はテキスト感情解析においてはネガティブの音声のものと平均値・最大値・最小値において差が 0.1 以下という結果が得られた。また中立の音声データはいずれも 0.5~0.65 の範囲に収まっている。

次に韻律感情解析に関しては、ポジティブの音声をポジティブであると判定したものは 4 割、ネガティブの音声をネガティブであると判定したものは 2 割であった。また自虐の音声に対

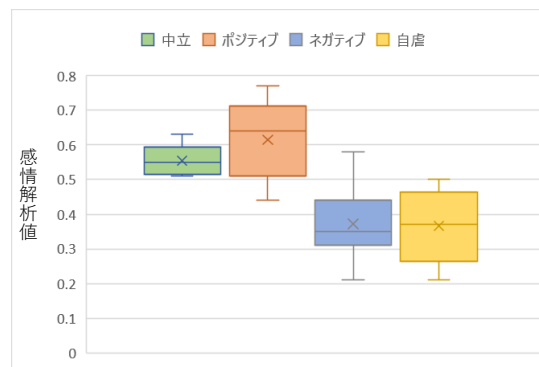


図 3 各感情ラベルごとのテキスト感情解析の結果

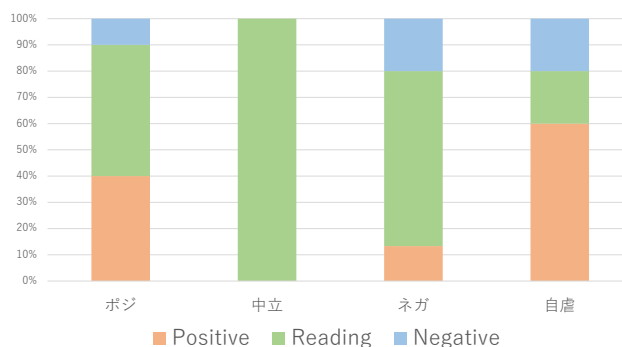


図 4 各感情ラベルごとの韻律感情解析の結果

しては 6 割の音声はポジティブであると判定されている。ポジティブである音声をネガティブと判定したものは 1 つ、ネガティブである音声をポジティブと判定したものは 2 つあった。

##### 4.3 提案手法の被験者実験による客観評価

提案手法に基づいた応答による傾聴システムが、従来のシステムよりも共感性の高い応答生成を実現できているか確認するため、被験者実験による客観評価を行う。あらかじめ対話内容を決められたシステムとの傾聴対話を録画し、それを第三者によって視聴してもらい、アンケートを通じてシステムの印象評価を行う。対話内容に関しては、ポジティブな話題とネガティブの話題の 2 種類、そしてその話題それぞれに対して聞き手 (バーチャルエージェント) の応答の種類として、既存システムを想定したものと提案システムを想定したものの 2 種類ずつの、計 4 種類の対話を用意する。例としてポジティブな話題で提案システムを想定した応答の対話内容の一部を下に示す。聞き手 (バーチャルエージェント) からの発話を A、話し手 (クライアント) からの発話を C とする。

C: 今日の朝はすごく天気が良くて気持ちよかったです。  
 A: まあそうなんですね。(笑顔の動作)  
 C: いやあ、綺麗に晴れていましたよ。  
 A: ええ、ええ  
 C: それで気分が良かったからね、いつもよりも結構遠くまで歩いて、楽しかったですね。  
 A: それはそれは、良かったですね(笑顔の動作)



図 5 アンケートに用いた傾聴動画の画面

また被験者はアンケートは 5 段階尺度の SD 法を用いた。対話への印象を質問し、尺度のラベルとしては全ての質問に共通して「そう感じなかった」「感じた」を使用した。また付随して自由記述によって対話への感想を質問した。用いた質問の一覧を下に示す。

- 質問 1: 聞き手の応答は自然でしたか。
- 質問 2: 聞き手は真面目に話を聴いていると感じましたか。
- 質問 3: 聞き手は積極的に話を聴いていると感じましたか。
- 質問 4: 聞き手は話し手の感情や状況を否定せず受け入れていると感じましたか。
- 質問 5: 聞き手に感情は感じましたか。
- 質問 6: 聞き手は話し手の感情に寄り添った応答をしていると感じましたか。
- 質問 7: 聞き手は親しみやすいと感じましたか。
- 質問 8: 話し手は話を通じて安心感を得ていると感じましたか。

質問の 1 によって対話が成り立っているかを確認し、質問 2・3 で積極的傾聴の必要要素である「一致」を、質問 4 で「受容」を、質問 5, 6 で「共感」を傾聴ケアシステムが実現できているかを確認する。また質問 7, 8 によって傾聴ケアの効果が得られているかを確認する。

#### 4.4 アンケート結果

アンケートで得られた、提案システムのスコアと既存システムのスコアの差分を値ごとに集計したもの（ポジティブな話題とネガティブな話題のものを合計する）を表 2 に、またそれらの平均・分散を表 3 に示す。これにより提案システムと既存システムそれぞれが与える印象の差を分析する。また 4.3 で述べた各質問の意図を確認するため、一部質問における各アンケートの差異のヒストグラムを図 6 に、またポジティブな話題とネガティブな話題での結果を比較するため、各話題ごとの質問 6 のヒストグラムを図 7 に示す。

アンケート結果の差分の平均ではいずれも正の値を示しており、提案システムの印象評価が総じて高くなっている。中でも

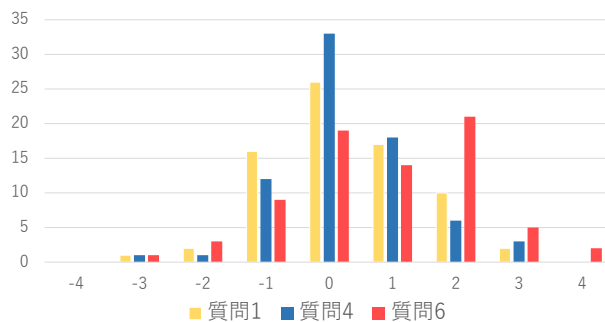


図 6 アンケートスコアの差分ごとのヒストグラム (抜粋)

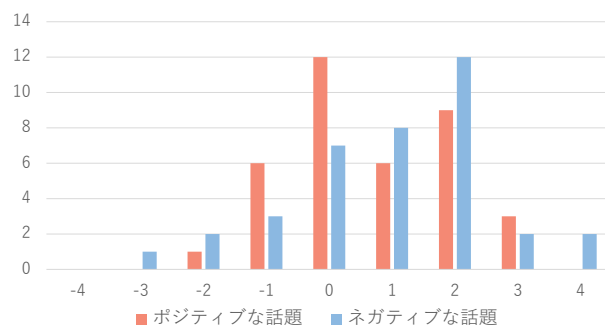


図 7 質問 6 におけるアンケートのスコアの差分ごとのヒストグラム

表 2 客観評価の差分による集計結果

		差分								
		-4	-3	-2	-1	0	1	2	3	4
0	0	0	1	9	16	6	5	0	0	
0	1	1	7	16	10	1	1	0		
0	0	1	6	13	12	4	1	0		
0	1	0	8	16	10	2	0	0		
0	0	1	6	12	6	9	3	0		
0	1	0	5	11	4	10	6	0		
0	0	1	6	12	6	9	3	0		
0	0	3	7	10	4	12	1	0		
0	0	1	8	14	9	4	1	0		

質問 5 「聞き手に感情を感じましたか」と質問 6 「聞き手は話し手の感情に寄り添った応答をしていると感じましたか」は他の質問よりも大きな差を示した。また図 6 からは、質問 4 は提案システムと既存システムに差を感じなかった人が多くを占め、質問 1 は他の質問より差異が正のものと負のもの両方が比較的多く見受けられる。また図 7 よりポジティブとネガティブの話題では、ネガティブの話題の方が感情に寄り添った応答をしていると感じる人が多いということが示された。

#### 4.5 考察

##### 4.5.1 感情解析の手法検証

まずテキストによる感情解析ではポジティブとネガティブを十分に判別できると考える。これはポジティブの音声の判定値の第一四分位数がネガティブの音声の判定値の第三四分位数より大きいこと、この間の数値を感情判定の閾値とすることで大

表 3 各質問ごとの差分の平均・分散

話題	平均			分散		
	ポジ	ネガ	合計	ポジ	ネガ	合計
質問 1	0.135	0.405	0.270	1.036	1.701	1.386
質問 2	0.676	0.973	0.824	1.679	2.459	2.091
質問 3	0.081	0.811	0.446	1.156	1.829	1.625
質問 4	0.405	0.324	0.365	1.106	1.841	1.475
質問 5	0.486	0.649	0.568	1.925	1.633	1.786
質問 6	0.919	0.892	0.905	2.183	2.637	2.410
質問 7	0.081	0.514	0.297	0.939	1.331	1.182
質問 8	0.270	0.432	0.351	1.170	1.489	1.336
平均	0.382	0.625	0.503	1.399	1.865	1.661

きな間違いは避けられると考えられる。基本的にテキスト解析の結果によって判定してよいと考えられ、特にポジティブの音声ネガティブの音声ともに平均値以上/以下は異なる感情を判定するという結果が見られないことから、この値を強い指標として用いられると考えられる。しかし、それでもなお感情的な音声の約半数が中立と判定されることや、自虐的発言のように感情的に反応することが躊躇われる音声がテキストのみでは判断がつかないことから、韻律的解析の結果を補助的に用いることによってより正確に感情を取得できると考えられる。

#### 4.5.2 傾聴ケアシステムの実現性

4.4 の実験結果より、提案システムは既存システムより共感性の高い応答が可能であると考えられる。各質問ごとに結果を見ていくと、まず質問 1 では提案システムと既存システムに大きな差は見られず、人によっては既存システムのほうがより自然であるという印象を抱いている。これは応答の感情に起伏が存在することで、人によっては対話に不自然さを感じるのではないかと考える。次に質問 2, 3, 4 について、これらは積極的傾聴に必要な共感以外の要素を満たしているかの確認であるが、いずれもある程度提案システムの方が良い印象を与えたという結果であり、特にポジティブな話題の質問 2・4 においてはほぼ差がなく、既存システムでもポジティブな話題ならばこれらの要素を満たせることが確認できた。そして質問 5・6 においてはともに提案システムが既存システムより高い共感性が見られ、傾聴ケアシステムの実現性が示唆された。また共感応答によって質問 8 においては今回は第三者による客観評価であるため、話し手の感じたことを聞くこの質問では既存システムとの差異が得られなかったのだと考えられる。今後実際に提案システムとの対話を通した主観的な評価が必要であると考えられる。

## 5. おわりに

本稿では、より高齢者の感情に寄り添った傾聴ケアシステムを提案した。提案システムでは、同調して共感応答に有効とされる韻律特徴を取得し、さらに韻律特徴とテキスト情報を組み合わせた感情解析によって、発話に表出した感情の大きさとその方向性を取得する。これらの韻律特徴および感情特徴を用い、発話者と同じ韻律・感情を発露した応答を生成する。

既存解析手法の実現性の確認のため感情解析手法の検証を予

備実験として行った。また、提案手法に基づいた応答が実現できているか確認するため、被験者実験による客観評価を行った。その結果、感情解析手法においてはポジティブとネガティブの感情を判別できる指標が確認された。そして被験者実験の結果として、提案システムが既存システムより高い共感性を持った応答を生成でき、よりクライアントの感情に寄り添った傾聴システムの実現性が示唆された。

今後の課題として、提案システムによる傾聴対話が発話者にとってどのような印象を与えるのか確認する必要があると考えられる。

謝辞 本研究の一部は JSPS 科研費 JP19H01138, JP18H03242, JP18H03342, JP19H04154, JP19K02973, JP20K11059, JP20H04014, JP20H05706 および、立石科学技術振興財団の研究助成を受けて行われている。

## 文 献

- [1] Government of Japan Cabinet office. annual report on the aging society. <http://www.cao.go.jp/>, June 2019. (Accessed on 11/13/2019).
- [2] パーソン・センタード・ケア — 健康長寿ネット. <https://www.tyojyu.or.jp/net/byouki/ninchishou/person-care.html>. (Accessed on 10/14/2021).
- [3] Seiji Sakakibara, Sachio Saiki, Masahide Nakamura, and Kiyoshi Yasuda. Generating personalized dialogue towards daily counseling system for home dementia care. In *International Conference on Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management*, pages 161–172. Springer, 2017.
- [4] 植田千晶. 「共感」の概念をどう教えるか. *和歌山大学教育学部教育実践研究指導センター紀要*, (8):73–80, 1998.
- [5] 福田正治. 看護における共感と感情コミュニケーション. *富山大学看護学会誌*, 9(1):1–13, 2009.
- [6] 長岡千賀. 対人コミュニケーションにおける非言語行動の 2 者相互影響に関する研究. *対人社会心理学研究*, 6:101–112, 2006.
- [7] 梅田聡. 共感の理論と脳内メカニズム. *高次脳機能研究*, 38(2):133–138, 2018.
- [8] 小森政嗣, 長岡千賀, and 中村敏枝. 音声対話における反応潜時の同調傾向が目標達成の評価に及ぼす影響. *感性工学研究論文集*, 6(4):91–98, 2006.
- [9] Marianne Sonnby-Borgström, Peter Jönsson, and Owe Svensson. Emotional empathy as related to mimicry reactions at different levels of information processing. *Journal of Nonverbal behavior*, 27(1):3–23, 2003.
- [10] Richard E Maurer and Jeffrey H Tindall. Effect of postural congruence on client's perception of counselor empathy. *Journal of counseling psychology*, 30(2):158, 1983.
- [11] Open jtalk. <http://open-jtalk.sourceforge.net/>. (Accessed on 10/14/2021).
- [12] Github - shahabks/myprosody: A python library for measuring the acoustic features of speech (simultaneous speech, high entropy) compared to ones of native speech. <https://github.com/Shahabks/myprosody>. (Accessed on 10/14/2021).
- [13] Text analytics — microsoft azure. <https://azure.microsoft.com/ja-jp/services/cognitive-services/text-analytics/>. (Accessed on 10/14/2021).