納得度を表す韻律情報の調査

早稲田大学　人間科学部　人間情報科学科

学籍番号：2J19F508

川崎　咲希

研究指導教員：菊池　英明

2021年1月14日

目次

[第1章　はじめに 3](#_Toc116152763)

[1.1本研究の背景 3](#_Toc116152764)

[1.2先行研究 3](#_Toc116152765)

[1.2.1　納得の定義 3](#_Toc116152766)

[1.2.2　対話における韻律情報を調査した研究 3](#_Toc116152767)

[1.3予備実験 4](#_Toc116152768)

[1.4本研究の目的 5](#_Toc116152769)

[1.5本論⽂の構成 5](#_Toc116152770)

[第2章　研究 5](#_Toc116152771)

[2.1方法 5](#_Toc116152772)

[2.2データ 5](#_Toc116152773)

[2.3評価項目 6](#_Toc116152774)

[2.4実験方法 6](#_Toc116152775)

[第3章　結果 7](#_Toc116152776)

[第4章　考察 7](#_Toc116152777)

[第5章　まとめと今後の課題 7](#_Toc116152778)

# 第1章　はじめに

## 1.1本研究の背景

患者を対象とした臨床研究を行う際、患者に対して医療者が直接説明を行い参加を打診し、同意を得る場合がある。患者にとって研究の説明は難解である。同時に、患者は医療者に対し気軽に質問や意見を述べられない場合がある。吉田ら(2017)は臨床試験に参加したがん患者へのインタビューから「説明過程に対する記憶も鮮明ではなく、わからないとしても試験実施者に質問をしていなかった」と指摘する。そして、「臨床試験を患者がどのように認識しているのかについては、試験実施者は慎重に確認する必要がある」と述べている。

患者の発話から納得度を測り可視化することができれば、医療者は説明の効果を確認でき、納得度が不十分な場合はそのポイントに関し再度説明や話し合いができるであろう。

## 1.2先行研究

### 1.2.1　納得の定義

　今井ら(2016）は「納得」を以下のように定義している。

「ある事象に対して，自分のもつ価値や自分への利益を明確にすることで理解を深め，認知的にも感情的にも受容した状態であり，主体的かつ他者との信頼関係のなかで生み出される流動的な状態」

### 1.2.2　対話における韻律情報を調査した研究

対話における韻律情報を調査した先行研究を調査した。

横谷ら（2020）は、構造化臨床インタビューのデータを用いて、休止間単位(IPU)と発話単位(LUU)での音響同調を比較し、精神科面接中の信頼をより良く予測できる単位を検討している。その中で、調査協力者とセラピスト間のRMS、モーラ、MFCCの同調とセラピストへの信頼の高さに関連があることを見出している。

　宮﨑ら（2021）は、韻律と動作および頭部特徴量のセットで主体性の推定精度が高いことを示唆している。そして、韻律情報において主体性はインテンシティの幅と関連しているという。

岡登ら(1999)は、相槌における韻律情報の特徴を利用し対話中の相槌を検出するシステムを構築している。この研究では聞き手側の相槌のうち無意味発話のみを対象としており、発話の基本周波数や発話長のほか、発話のタイミングからも特徴を見出していた。

戸上ら(2007)は研究協力者が簡易防音室内で問題を解いているときの音声を使用し句末のイントネーションに関する音響的特徴と学習者の理解度との関係を分析している。方法は、収録した音声を第3者が「驚き」「疑問」「困惑」「理解」を表しているか主観評価を行う。そして、句末のイントネーションの様子がどのように変化しているのか音響的な分析を行っている。分析に用いたパラメータは、句末のF0の傾きである。戸上ら(2007)は分析の結果、「理解」を表す音声は、句末がほぼ一貫して平坦であると示唆している。

　渡邉(2018)はインタビューにおける言語・非言語情報を用いて共感的反応の分析を行っている。また、登張(2014)によると共感の要素として「他者の感情と一致する感情を感じること」「他者の感情や心的状態が分かること」「他者に対する気遣い」が存在する。石澤ら(2014) は共感的反応尺度および応答的反応尺度として5段階評価を行っている。

豊田ら(2012)の研究では、２者間対話における発話状態時間長に着目し「単独発話時間」「無音時間（話者交替時に生じる）」「同時発話時間」を特徴量に用いている。そして、対話内の合計発話時間が長い話者をA、もう一方をBとして分析している。Aは対話をリードする側とし、分けて分析している。

西村ら(2009)は人間同士の対話のリズムに着目し対話者間の相互作用を調査・分析している。調査方法は被験者が対話音声を聴収し、「親しさ」「盛り上がり」「同意反発」「年齢差」「かみ合い」の５段階評価を行なっている。

須藤(2008)は「うん系」感動詞の韻律的特徴の分析、考察を行っている。調査方法は調査協力者に同意および非同意の「うん」を発話してもらい分析対象としている。この方法では調査協力者間の個人差が大きかったため、分析対象音声のうち最も「同意らしい」「不同意らしい」音声を抽出し特徴を検討する必要があるとしている。

渋谷ら（2007）は親近感を表す特徴量として基本周波数を挙げている。そして、不特定話者の基本周波数の分布形状から特徴量を抽出し、親近感を判定している。

藤江ら(2005)はユーザの発話態度を韻律情報から認識するために以下の方法を用いている。まず、音声合成装置を用いて合成した発話に対して調査協力者が肯定的、否定的な態度で応答した音声をそれぞれ収録する。発話者が収録時に想定した態度を正解として、肯定的な発話と否定的な発話における特徴量の違いを学習し識別器とする。一方で発話した者とは別の協力者が発話を聞き否定的か肯定的か判定し、識別器で得られる結果との一致率も検証している。学習には2000発話、一致率の検証には収録したデータのうち、肯定的、否定的の各20発話を用いている。認識に用いる特徴量は「モーラの母音部分のF0の傾き」「発話全体のF0レンジ」「最終モーラの継続長」であった。

天利ら(2021)はカウンセリングにおけるクライアントの納得の推定を行っている。カウンセラーの言語特徴量のうち、納得群と非納得群では名詞や動詞の単語数および繰り返し数、そしてポジティブな単語数に有意差があった。

「納得しているかどうか（納得度）」に近い「理解度」「肯定的態度」「親近感」などの事象と韻律情報の関係を明らかにした研究および納得と言語特徴量の関係を明らかにした研究は行われているが、「納得しているかどうか（納得度）」と韻律情報の関係を明らかにした研究は存在しない。

## 1.3本研究の目的

本研究の目的は発話から「納得しているかどうか（納得度）」を知ることである。先⾏研究から発話の韻律情報から納得度を推定できるかをリサーチクエスチョンとした。予備実験の結果から、特にF0の最大値と最小値の差に着目する。

## 1.4予備実験

予備実験の目的は発話から「納得しているかどうか（納得度）」を知るにあたって必要と思われる手続きの試行である。予備実験は自動車販売員対話コーパス「c1\_f\_06」の音声を聴収し、はい/あー/そうですね等の自分が相槌と判断した28の発話を対象に評価した。「納得しているかどうか」を第３者である自分が評価する方法をとった。評価尺度は「納得を示す反応を表出していない」を１とし「納得を示す反応を表出している」を５とし、度数の間隔は均一であるとした。以下に評価結果を該当発話の横に示す。

Ａ６：［はー］　2

Ａ１４：（な）なるほどね。［そん］違いますね。 5

Ａ１６：［あ］そうですね。やっぱ気持ちいいと思いますので。 5

Ａ１９：［はー］ 3

Ａ２０：か，子供ですか。［はははー］ 4

Ａ２１：［んー］ 3

Ａ２４：気にならない。［ははは］ 3

Ａ２８：なるほど，その点やっぱいいですね。 3

Ａ２９：［へえ］3

Ａ３６：［あー］ 4

Ａ４１：［はー］ 2

Ａ４２：［あ］確かに安いですね。 4

Ａ４３：［あ］すごい安いですね。 5

Ａ４４：普通の，ごく普通のセダン，ですね。 3

Ａ４５：［ふーん］ 2

Ａ４６：［はー］ 2

Ａ５０：分かりました。ベンツはまあ，高いですからね。［はははー］ 4

Ａ５２：［ふーん］ 3

Ａ５３：なるほど。 3

Ａ５６：［あー］じゅうもう，そうなんですか。 3

Ａ５７：やっぱ標準なってるんですね。 4

Ａ５８：［んー］ 2

Ａ６２：分かりました。 4

Ａ６５：ええ。 4

Ａ６８：［あ］そうですか。［はー］ 4

Ａ７１：［あは］ 4

Ａ７２：［はー］ 2

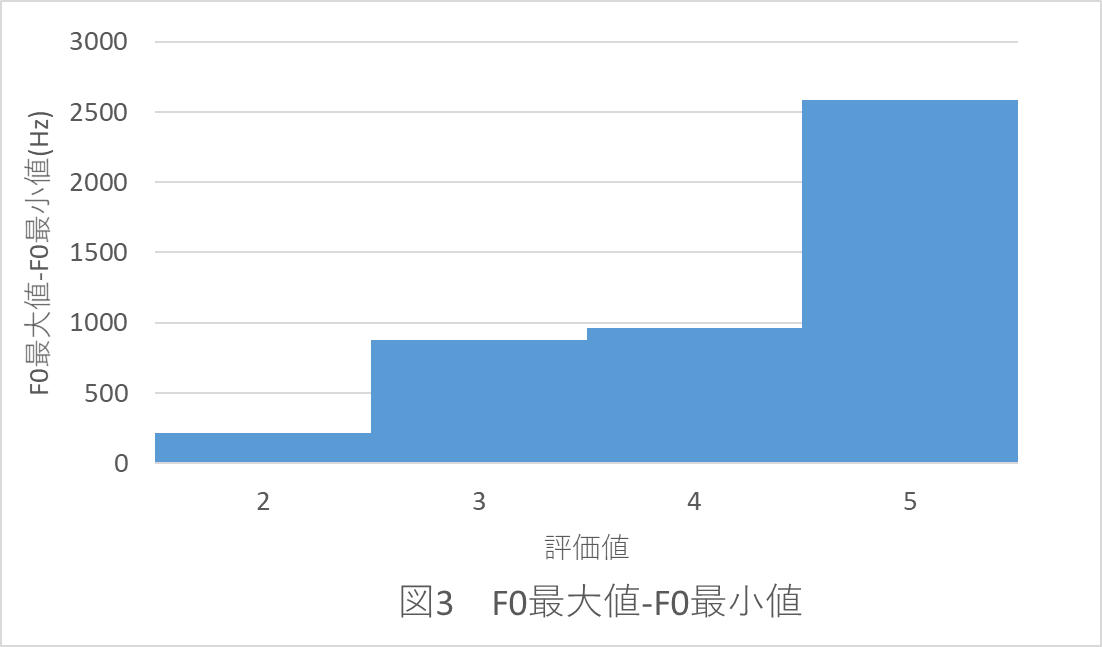
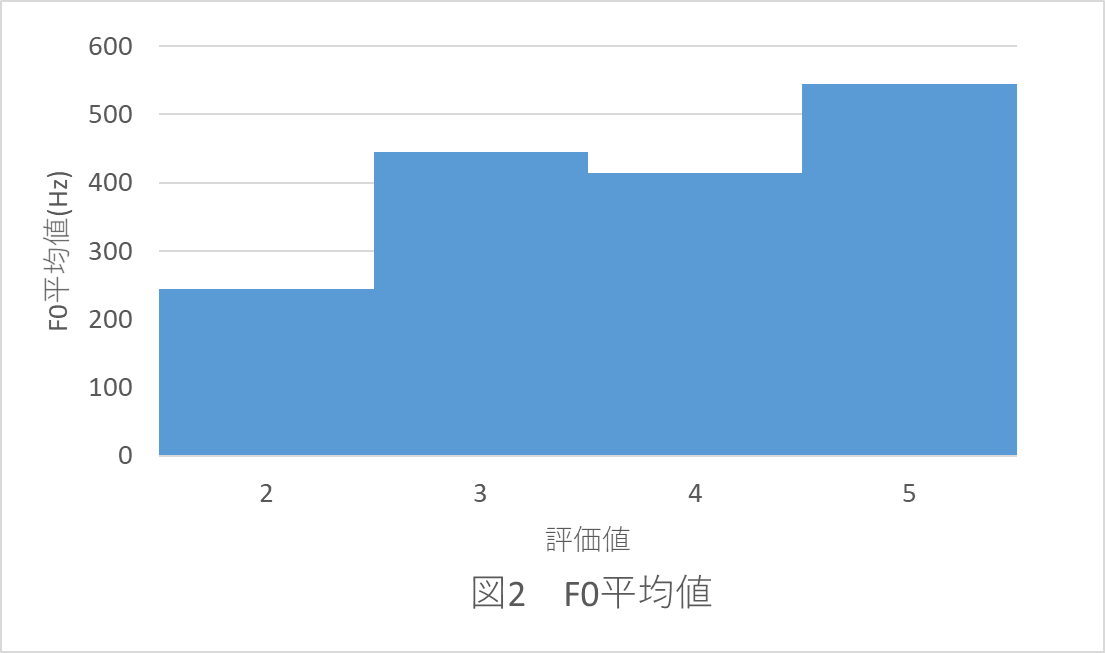
Ａ７３：［あ］分かりました。 4

上記の発話ごとに特徴量を数値化し、納得評価値と特徴量との関係を分析した。自分が評価を行なった際、評価値の高かった発話の特徴として抑揚の有無が挙げられた。そこで抑揚の度合いを判断するため、予備実験において検討する特徴量は基本周波数とした。基本周波数の平均値、最大値と最小値の差を調べた。

納得評価値と基本周波数をグラフ化したものを以下に示す。図1（評価値ごとのF0平均）より、評価値の低い発話よりも評価値の高い発話の方がF0が大きいと示された。また、図2（評価値ごとのF0最大値とF0最小値の差）より評価値の低い発話より評価値の高い発話の方がF0の差が大きいと示された。

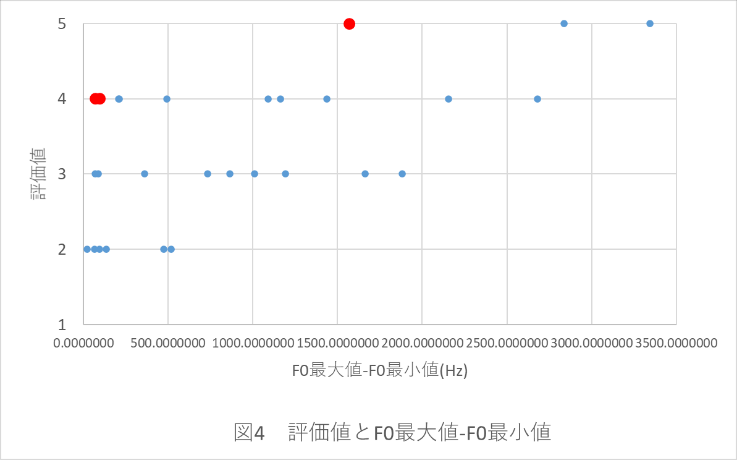
図2

図1



評価値とF0最大値－F0最小値に関し、図3の２次元散布図を作成した。図3において赤くプロットした値を視覚的に外れ値と考え、該当するA14、A36、A62の発話を改めて聴収し考察を行った。

図3



改めて聴収すると、３つの発話は息を吐きながら発音していると思われた。そのため他の発話よりもF0の値が振れないと予想される。F0の差以外の納得度評価に影響を与える要因として、３つの発話は理解や受容の意味を示していることが考えられる。A14、A62は「なるほど」「わかりました」といった理解を示す文言を発している。A36は文字起こし上では「あー」とされているが聴収した際には「ああー」という受容ととれる文言と判断していた。

以上予備実験の結果では、納得度評価に影響を与える要因は「F0の最大値と最小値の差」および「発話内容」であった。予備実験において評価対象とした発話は自分が相槌と判断した発話であった。そのため、研究では相槌にも様々な形があることを踏まえて評価対象を選定する必要がある。

## 1.5本論⽂の構成

本論⽂は次のように構成されている。

第2章では研究⽅法、使用するデータおよび評価項目を述べる。

第3章では結果を述べる。

第4章では第３章の結果に基づき考察を述べる。

第5章ではまとめと今後の課題を述べる。

# 第2章　研究

## 2.1方法

音声を聴いて納得しているかどうかを第3者評価し、その音声の韻律情報と納得度の関係を調査する。

## 2.2データ

研究に使用するデータについて、データベース内の説明文書より概要を以下に述べる。データは旅行代理店のカウンターにおける店員と客の対話ロールプレイングを収録したものである。客役の被験者が旅行の計画を店員役の被験者に相談する設定である。客役の被験者は必ずしも実生活で旅行を計画しておらず、収録環境に用意した旅行パンフレットを見て店員役に相談する内容を決めている。

・RWCP 音声対話データベース（RWCP-DB-SPEECH-96-1）計24対話

被験者は男性、女性各12名

・RWCP 音声対話データベース（RWCP-DB-SPEECH-97-1）計13対話

被験者は男性7名、女性6名

データベースのうち、以下の5つの対話データを研究に使用する。

T1\_M\_05.L16、T1\_M\_12.L16、T2\_F\_01.L16、T2\_F\_04.L16、T2\_F\_05.L16

以下はT2\_F\_05の対話の一部である。

　客：で，あとパスポートなんですけども，一回，行ってちょっと切れちゃってるんですね。そうすると取り直し，てゆうと，何日ぐらいかかりますか。

店員：まあ十日間みていただければ。

客：十日ですか。

店員：はい，お取りできますので。

客：今，十年でした，け，一回。

店員：五年物と十年物と，うん，二つあります。

客：もうそれはどちらでも好きな方で。

店員：五年，の方は一万円で，十年になりますと一万五千円になります。

客：そうですか。うん，十年の方が得ですねきっとね。

店員：よく行かれる方は十年の方がう，ん，得だと思いますけれども。

客：そうですか。カナダの食べ物，てゆうのは，どうです日本人に合いますか，ね。

店員：海の物でしたら，サーモン類が。

客：サーモン類ですか。

　選定した5つの対話はいずれも、どちらかが一方的に話すのではなく、上記の対話のように客からの質問・希望に対して店員が回答・提案しながら対話が展開されているため、納得度の評価に適している。

## 2.3評価項目

　本研究で評価する韻律情報は以下である。

　F0の最大値と最小値の差、句末のF0の傾き、発話時間

　また、納得度は検討に際し数値化する必要がある。西村ら(2009)の研究を参考に以下の尺度で評価する。

「納得を示す反応を表出していない」を１とし「納得を示す反応を表出している」を５とする。度数の間隔は均一であるとする。

## 2.4実験方法

旅行代理店の店員と客の対話データを第3者である自分が聴き、納得度を５段階で評価する。評価対象発話の前後の対話音声を含めて聴収する。また、聴収する回数の制限は設けない。

納得度評価結果と韻律情報の関連を調べる。

音声聴収の手順を以下に示す。

①発話評価に使用するデータの音声ファイル（拡張子".l16"のファイル）をNASのcorporaにあるRWCPフォルダよりダウンロードする。

②フリーソフトウェア（Wavesurfer）をダウンロードする。

③Wavesurferを起動し、"File "> "Open "でダウンロードした音声ファイルを開く。開くファイルを選択する際には、ファイルの種類を"All files"に変更し拡張子".l16"のファイルを選ぶ。

④表示されたダイアログボックスで以下の設定にし"OK"ボタンを押す。

Sample rate: 16000

Sample encoding: Lin16

Channels: Stereo

Byte Order: Big Endian

⑤別のダイアログボックスが表示される。"n-waveforms"を選んで"OK"ボタンを押す。

⑥音声波形が表示される。右上の黒三角のボタンをクリックし音声を再生する。

⑦対話音声を聴収し納得度を評価する。

特徴量を数値化する手順を以下に示す。

①基本周波数の値をWavesurferからダウンロードするため、"Transform"＞"Convert"を開き"Channels"を"Mono"に変更しOKをクリックする。

②データウインドウ内で右クリックし"Create Pane">"Pitch Contour"を選択する。

③基本周波数が表示されたウィンドウ内で右クリックし"Save Data File..."を選択してテキスト形式のファイルをダウンロードする。このファイルに基本周波数の値が入っている。

④③でダウンロードしたテキストファイルをExcelで開き検証に利用する。

⑤Excelで開いた際に左から3つ目のセルに基本周波数F0が表示される。セル１行には0.01秒間のデータが表示されるため、①で記録した時間に該当するF0値を抽出する。

⑥抽出したF0の平均および最大値、最小値、句末のF0の傾き、発話時間を求め、納得度評価値と紐付ける。

# 第3章　結果

コーパスの対話音声における韻律情報と納得度の関係を検証する。

# 第4章　考察

# 第5章　まとめと今後の課題

参考文献

天利安志,岡田将吾,松本妹子,貞光九月,中元淳「美容カウンセリングにおけるクライアントの納得場面の自動分析」『人工知能学会研究会資料 言語・音声理解と対話処理研究会91回』(人工知能学会,2021)pp.78-84.

有本泰子,大野澄雄,飯田仁「「怒り」の発話を対象とした話者の感情の程度推定法」『自然言語処理』14.3 (言語処理学会,2007)pp.147-163.

石澤亜耶乃,島田英昭「ワーキングメモリの負荷が共感的反応に及ぼす影響－二重過程理論に基づく検討－」『認知科学』21.2, pp.245-253.

稲川恵一,新宅弘幸,小倉茂,他「韻律情報の制御を行った合成音声による感情の考察」『工学院大学研究報告 = Research reports of Kogakuin University』86(工学院大学,1999)pp.199-205.

今井芳枝,雄西智恵美,板東孝枝「納得の概念分析－国内文献レビュー－」『日本看護研究学会雑誌』39.2 (日本看護研究学会,2016)pp.73-85.

岩井淳,矢野良和,大熊繁「韻律的特徴による感情音声の分類と認識に関する基礎検討」『ファジィシステムシンポジウム講演論文集』20,pp.192-195.

大島寿美子「婦人科がんの患者会を対象とした医師とのコミュニケーションに関する基礎的研究: 医師による情報提供を中心として」『北星学園大学文学部北星論集』44.1(北星学園大学文学部, 2006)pp.93-101.

岡登洋平,加藤佳司,山本幹雄,板橋 秀一「韻律情報を用いた相槌の挿入」『情報処理学会論文誌』40.2(情報処理学会,1999) pp.469-478.

越智雅人,黒田久泰「アクセント型とモーラを考慮することによる感情を含んだ音声変化の解析」『人工知能学会全国大会論文集』JSAI2014(人工知能学会,2014)pp.2C33-233.

小野貴大,斎藤博人,金子博,他「ポスター講演 音声パラメータと感情印象変化の対応モデルの構築 : 話速とピッチを制御された音声の感情印象の改善」『電子情報通信学会技術研究報告 = IEICE technical report : 信学技報』114.440 (電子情報通信学会,2015)pp.193-198.

苔縄陽子,津崎実,加藤宏明,匂坂芳典「基本周波数パタンに見られる発話態度の分析」『情報処理学会研究報告ヒューマンコンピュータインタラクション(HCI)』2004.74 (情報処理学会,2004)pp.87-92.

笹目友香,小渕千絵,山本弥生,城間将江,野口佳裕「音韻情報と韻律情報の一致･不一致が感情判断に与える影響」『国際医療福祉大学学会誌』26(抄録号) (国際医療福祉大学学会,2021) p.140.

重永實「感情の判別分析からみた感情音声の特性」『電子情報通信学会論文誌.A, 基礎・境界=The IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences (Japanese edition).A』83.6(電子情報通信学会,2000)pp.726-735.

渋谷貴紀,川端豪「音声対話システムのための親近感特徴量の探索『電子情報通信学会技術研究報告 = IEICE technical report : 信学技報』106.122(電子情報通信学会,2006)pp.25-30.

渋谷貴紀,益永祐吾,川端豪「F0分布に基づく親近感特徴量の話者適応」『情報処理学会研究報告音声言語情報処理 (SLP)』2007.(129(2007-SLP-069))(情報処理学会,2007)pp.67-72.

須藤潤「「うん系」感動詞の韻律的特徴に関する―考察―「受け入れ」にかかわる意味・機能をめぐって―」『ポリグロシア』15(立命館アジア太平洋大学,2008)pp.99-108.

多田和彦,矢野良和,道木慎二,大熊繁(2010) 「感情遷移における急激な韻律特徴変化の検出による感情遷移判別法」『知能と情報』22.1,pp.90-101.

鶴真紀子,武田昌一「アナウンサーが発声した「怒り」の韻律的特徴と聴覚的印象の主成分分析を用いた比較」『久留米信愛女学院短期大学研究紀要』30(久留米信愛女学院短期大学,2007)pp.65-70.

鶴真紀子,武田昌一「種々の度合の感情音声における発話者の意図と聞き手の受容の一致率と韻律的特徴との関係」『日本感性工学会論文誌』13.2(日本感性工学会,2014)pp.381-389.

戸上雅夫,飯田仁「学習者の理解度把握へ向けたeラーニング学習中の音声の音響的特徴分析」『第69回全国大会講演論文集』2007.1(情報処理学会,2007)pp.575-576.

登張真稲「共感の神経イメージング研究から分かること」『発達心理学研究』25.4, pp.412-421.

豊田薫,宮越喜浩,山西良典,加藤昇平「発話状態時間長に着目した対話雰囲気推定」『人工知能学会論文誌』27.2(人工知能学会,2012)pp.16-21.

西村良太,北岡教英,中川聖一「音声対話における韻律変化をもたらす要因分析」『音声研究』13.3(日本音声学会,2009)pp.66-84.

藤江真也,江尻康,菊池英明,小林哲則「肯定的/否定的発話態度の認識とその音声対話システムへの応用」『電子情報通信学会論文誌D』J88-D2.3(電子情報通信学会, 2005)pp.489-498.

藤木大介,若杉佳彦,楞野祥子,岩本理沙,島田英昭「作動記憶負荷が物語への共感的反応に及ぼす影響」『心理学研究』88.4, pp.390-395.

前川喜久雄「韻律とコミュニケーション」『日本音響学会誌』55.2(日本音響学会, 1999)pp.119-125.

宮﨑健斗,片上大輔「マルチモーダル情報に基づく就職面接場面における被面接者の評価モデルの提案」『人工知能学会論文誌』36.5(人工知能学会,2021)pp.1-9.

宮澤幸希,佐藤可直「音声対話システムとのコミュニケーション円滑化を目的とした韻律情報からの態度推定」『言語・音声理解と対話処理研究会』95(言語・音声理解と対話処理研究会,2022)pp.75-80.

横谷謙次,高木源,若島孔文「心理面接中の韻律の類似性が信頼関係を予測する」『第82回全国大会講演論文集』2020.1(情報処理学会,2020)pp.47-48.

吉田幸恵,中田はる佳,武藤香織「臨床試験に関与した, がん患者の語り-｢治療｣と｢研究｣を区別することの困難さに関する考察」『生命倫理』27.1(日本生命倫理学会,2017)pp.122-131.

李遠超,井上昂治,中村静,高梨克也,石井カルロス寿憲,河原達也「ヒューマンロボットインタラクションにおける韻律とテキスト情報を組み合わせた感情認識と評価応答選択」『人工知能学会研究会資料』79(言語・音声理解と対話処理研究会,2017)p.9.

渡邉綾「日本在住外国人の医療体験に関するインタビュー:言語・非言語資源を用いた共感的反応の協働構築」『言語文化共同研究プロジェクト』2017, pp.23-31.

RWCP音声対話データベース（RWCP-DB-SPEECH-96-1、RWCP-DB-SPEECH-97-1）説明文書