

音声分析と対話型 AI による ACP 会話の振り返り支援

星 温樹^{1,a)} 角 康之^{1,b)} 佐藤 生馬¹ 塚本 美穂² 上原 浩文³

概要: 本研究は、ACP (Advance Care Planning) における会話の振り返りを支援することを目的とする。日本では ACP の認知度や実施率が依然として低く、医療従事者の教育機会の不足や記録作業の負担、記述内容のばらつきが課題となっている。本研究では、臨床データの利用に伴う個人情報保護や当事者への心理的負担といった課題を伴うことを踏まえ、仮想的なロールプレイング会話を対象とした。その上で、重要な発言を項目別に自動抽出・要約し、対話形式により内容を確認可能なシステムを構築した。これにより、会話全体の把握や重要事項の整理が容易となり、記録の正確性と効率性の向上が期待できる。さらに、会話内容の充実度や沈黙時間などの非言語的特徴を可視化して客観的に分析することで、会話の質を多面的に評価し、医療従事者が効果的に振り返りを行えるよう支援する。加えて、要約結果や評価を教育資源として活用することも可能であり、医療従事者のトレーニングや ACP 教育への応用も期待される。

1. はじめに

ACP (Advance Care Planning) とは、将来の医療とケアについての目標や希望を定義し、それを家族や医療従事者と話し合い、必要に応じて記録・見直しをするプロセスである。患者の意思を尊重したケアが実現するほか、患者本人や家族の心理的負担を軽減する取り組みとして重要視されている。Karen ら [1] は高齢患者に対して ACP を実施することで患者の終末期の希望がどれだけ尊重されるか、心理的負担が減るかを検証することを目的に比較試験を行った。309 名の患者を、通常ケアのみを行う対照群と、ACP を加えた介入群に無作為に割り付け、患者の満足度を定量的に評価した。その結果、ACP は患者に対して多面的な利益をもたらす有効な手段であると結論づけられた。

ACP の有用性が示される一方で、日本においては実施率が低い現状にある。その背景には、ACP を担当する医療従事者の技術的・時間的制約や教育機会の不足といった課題が存在する。ACP は情報収集、会話、記録など複数の段階から成るが、特に記録の段階では、会話から患者の価値観や希望を過不足なくまとめる必要がある。しかし、会話の内容を一枚の資料に集約する現行の手法では作業に長時間を要するうえ、医療従事者の記録スキルに依存するため、内容の質にばらつきが生じやすい。さらに、重要な発言や微妙なニュアンスが抜け落ちやすいという課題がある。

そこで本研究では、ACP における会話音声から重要な発言を項目別に抽出・要約し、それをもとに対話形式で会話内容を確認できるシステムを構築するとともに、発話内容および会話音声を分析して会話の質や注目点を評価し、その結果をフィードバックすることで、ACP 会話の記録および振り返りを支援することを目的とする。ただし、本研究では、臨床データの利用に伴う個人情報保護や当事者の心理的負担といった課題を踏まえ、実際の ACP 会話ではなく、看護師や医師のトレーニングを兼ねた仮想的な ACP ロールプレイング会話を対象とする。

2. 関連研究

会話の記録・振り返りを目的とした研究は、LLM や対話分析手法を用いた取り組みが報告されている。

例えば、Fraile Navarro ら [2] は医療会話を対象とした研究で、ChatGPT を用いた自動要約が人手要約に近い品質を示し、臨床的にも有用であることを示した。また、Daryanto ら [3] は面接練習を対象とした研究で、LLM との会話を通じて発話内容を振り返るプロセスを設計することで、利用者の自己省察や理解が促進されることを示した。さらに、辻本・角 [4] によりチュータリングにおいて発話のタイミングや量といった音声的特徴を用いて、対話の質や参加者の関与度を推定する試みが行われている。また、澤本ら [5] は、発話や視線といった観測可能な行動単位の組み合わせや時間的關係に着目し、それらを階層的に分類することで、対話インタラクションをボトムアップに構造化する手法を提案している。これらの知見は、本研究においても、会話中の沈黙や相槌といった行動の組み合わせや時間的關係を捉

¹ 公立はこだて未来大学

² 介護老人保健施設ゆとりろ

³ 函館五稜郭病院

a) h-hoshi@sumilab.org

b) sumi@acm.org

えることが、会話構造の理解に有用であることを示唆している。

これらの研究は、LLM が会話要約に有効であること、AI との対話が振り返りを支援し得ること、会話の音声の特徴から対話の質を評価できる可能性を示している。一方で、ACP 会話のように価値観や意思決定を扱う対話を対象とし、要約結果や分析結果をもとに当事者が内容を確認しながら対話的に振り返るプロセスまでを統合的に支援する研究は十分に検討されていない。

3. 提案手法

本研究では、ACP における会話音声から重要な発言や項目を自動的に抽出・要約した後、医療従事者が対話形式で会話内容を確認でき、さらに発話内容と会話音声进行分析して会話の質や注目点を評価し、その結果をフィードバックするシステムを構築する。システムの流れを図 1 に示す。

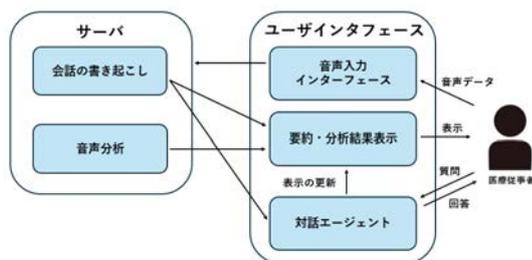


図 1 システムの構成図

本提案手法 (図 1) を実装するにあたり、計算負荷の大きい音声分析および会話の書き起こし処理は Google Colaboratory (以下、Colab) 上で実行し、ユーザが操作する対話型インターフェースは Streamlit^{*1} を用いた Web アプリケーションとして構築している。ただし、この形態は開発・評価の初期段階において手軽に利用できる利点がある一方で、実際の医療現場での運用やユーザビリティの面からは、使用環境や要求に応じて検討し直す必要がある。

3.1 データの収集

本研究で用いるデータは、協力いただいている病院から提供を受ける。提供される音声データは、実際の患者との会話記録ではなく、医療スタッフの教育や研修を目的として実施されている仮想 ACP の会話を録音したものである。録音には Zoom を使用し、話者ごとに個別の音声トラックを取得することで、話者別の音声分析を可能にする。この仮想 ACP データをシステムの入力として用い、システムの各機能の性能評価および検証を行う。

*1 Streamlit (version 1.44.0), <https://streamlit.io/>

3.2 書き起こしと音声分析

ユーザが Streamlit のインターフェースから音声データをアップロードすると、データは Colab 環境に送信され、OpenAI の Whisper (medium モデル)^{*2} を用いて音声データの書き起こしが実行される。

次に、分離された各音声トラックに対して webrtcvad^{*3} を用いた音声区間検出 (VAD) を実行し、「発話量」、「沈黙区間」、「会話の衝突 (オーバーラップ)」などの会話構造に基づく定量的指標を算出する。具体的には、沈黙については、全参加者の音声トラックがいずれも無音と判定された区間を抽出し、その総時間と回数を算出する。特に 3 秒以上継続する沈黙を「長い沈黙」として抽出する。また、会話の衝突については、2 名以上の話者が同時に有声と判定された区間を抽出する。

さらに、音響特徴量の解析として、parselmouth^{*4} ライブラリを用いて各話者の声のピッチ (基本周波数) や音量 (Intensity) を抽出する。ここでは、性別や個人による声の高さの違いを正規化するため、基本周波数 (Hz) を対数尺度であるセミトーン (半音) へ変換した上で標準偏差を算出している。これにより、話者の「抑揚」や「感情の表れに関連する特徴」を、客観的に評価する指標として用いている。

これらの処理が完了すると、生成された「書き起こしテキスト」と音声分析結果が JSON 形式で Streamlit アプリケーションに返送され、要約・分析結果の可視化やチャット機能の参照データとして利用される。

3.3 会話の要約

Colab から「書き起こしテキスト」を受け取ると、OpenAI API の GPT モデル (GPT-5)^{*5} を用いて要約を生成・提示する。当初、本システムでは ACP 会話の要約を文章形式で生成していた。しかし、文章形式の要約では、プロンプト上で要約に含めるべき情報を明確に定義することが難しく、モデルが着目する情報の観点が要約ごとにばらつくという課題があった。その結果、ACP 記録として重要な内容が欠落したり、逆に重要度の低い情報が強調されたりする場合があります。要約結果の一貫性を確保することが困難であった。

また、文章形式の要約では、要約文全体を対象として評価を行う必要があり、どの情報が適切に要約されているかを項目ごとに確認することが難しいという課題があった。

そこで本研究では、協力機関の看護師が実際に作成した ACP 会話要約文を分析し、ACP の記録として重要となる

*2 Whisper (v20240930), <https://github.com/openai/whisper>

*3 webrtcvad (version 2.0.10), <https://github.com/wiseman/py-webrtcvad>

*4 parselmouth (version 0.4.7), <https://parselmouth.readthedocs.io/>

*5 OpenAI Python SDK (version 2.11.0), <https://github.com/openai/openai-python>

要素を整理した上で、ACP 要約のための 9 項目を以下に定義した。

- (1) 本人が楽しみにしている事、趣味
- (2) 病気や怪我に対する理解
- (3) 本人の治療に対する希望
- (4) 心配している事
- (5) 本人の動向
- (6) 支えになる人、代わりに意思決定できる人
- (7) 状況変化に応じて見直す事
- (8) 退院後の希望
- (9) 今後病院側がやること

システムはこの 9 項目に基づき、書き起こしテキスト全体から関連する発言を項目ごとに抽出・要約するよう、プロンプトを設計している。また、要約の一貫性と精度を高めるため、仮想的な ACP 会話例とその要約を Few-shot 形式でプロンプトに含める工夫を行った。さらに、ハルシネーション防止のため、要約は書き起こしテキストのみを参照するようにし、該当する会話が存在しない項目については「該当なし」と明示するように指示している。

3.4 チャット機能

本システムでは、要約・分析結果を確認しながら会話内容を深く振り返ることを目的として、チャット形式の対話型インターフェースを構築する。チャットボットは要約ではなく、書き起こしテキストそのものを直接参照する。これにより、要約過程で生じる情報の欠落や抽象化の影響を受けずに、会話の詳細について自由に質問することが可能となる。応答の生成には、要約機能と同様に OpenAI API の GPT モデルを用いる。チャットボットは会話全文の文脈を踏まえて応答を生成するため、要約だけでは把握しきれない発話の意図や背景を対話的に確認できる。これにより、ACP 記録の正確性向上と会話理解の深化を支援する。

3.5 会話の評価とフィードバック

本システムでは、会話の評価を LLM による定性評価と音声分析に基づく定量評価の二つの側面から行う。定性評価では、LLM が会話の書き起こしテキスト全体を入力として、ACP 会話としての質を会話内容に基づいて評価する。具体的には、あらかじめ定義した評価観点に基づき、LLM による採点およびコメント生成を行う。定性評価に用いた評価観点および評価基準、ならびに採点・コメント生成の指示を含むプロンプトの詳細を以下に示す。

- (1) はじめの挨拶と終わりの挨拶が適切か

評価基準：

- 会話ができる体調なのか、時間はあるかなどを確認し、患者の理解を得てから会話を始めているか。(5

点満点で採点して下さい)

- 会話を終わるときは時間を使ったことへの感謝や今後気軽に相談できる雰囲気を作って会話を終わっているか (5 点満点で採点してください)
- はじめと終わりの点数を足して評価項目 1 の点数にしてください。また、コメントでは何が足りなかったかを挙げてください。

- (2) 直接的な表現を避けて聞くべきことを引き出しているか

評価基準：

- 趣味の話や家族の話など、世間話をしながら、退院後に楽しみにしていることや支えになる人などの患者の価値観を聞き出しているかを判断して、10 点満点で採点してください。
- 患者が楽しく会話をしながら必要な項目を話してくれたら高得点です。コメントでは、世間話から引き出したことを記述してください。

- (3) 必要なことを聞き出しているか

評価基準：

- 要約に必要な 9 項目を聞き出しているかを判断してください。「治療に対する希望」の項目では、2 点 (延命治療の有無や蘇生処置などの希望が含まれていなかったら減点)
- それ以外の項目はそれぞれ 1 点満点で採点し、合計した点数を評価項目 3 の点数としてください。コメントは、聞き出しが足りなかった項目を挙げてください。

- (4) 患者の意思を尊重しているか

評価基準：

- 「ですよね?」「はい」など確認作業のような会話になっていたら減点。判断を誘導するような表現をしていたら減点。
- 満点を 10 点とし、減点方式で採点して下さい。コメントには減点に該当した会話を抜き出して下さい。

評価項目は、協力を得た看護師から共有された ACP 会話の一般的な進行や、実践において重要とされるコミュニケーション上の工夫や留意点を基に、書き起こしテキストから判断可能な内容として設定した。各評価観点を点数化することで、会話における「良かった点」および「改善が必要な点」とその程度を直感的に把握でき、振り返りに有用であると考えた。評価方式については、ACP 会話において実施されていることが望ましい要素については加点方式を採用した。一方で、患者の意思決定を損なう可能性があるなど、できるだけ避けるべき要素については減点方式とした。特に、「必要なことを聞き出しているか」の評価項目においては、ACP において中核的な位置づけにある「治療に対する希望」をより重要と捉え、配点を高く設定

している。また、単なる点数付けにとどまらず、LLM によって各評価項目の採点理由や、不足している点を言語的にフィードバックするよう設計した。これにより、具体的な改善方針を把握しやすくし、次回以降の ACP 会話においてどの点を意識すべきかを具体的に理解できるよう支援することを目的としている。

一方、定量評価では、音声分析によって得られた客観的な指標（発話比率、話速、沈黙の割合、会話の重複、声の抑揚など）を用いて、会話の構造と質を評価する。本システムは、これらの数値を可視化してユーザーに提示するとともに、あらかじめ設定された閾値や推奨範囲との比較を行う。これにより、例えば「沈黙が少なく一方的に話している可能性がある」や「抑揚が乏しく事務的な対応になっている」といった、データに基づいた具体的なフィードバックを提供する。また、閾値を逸脱するなど指標に顕著な特徴がみられた会話区間（例：3秒以上の沈黙、会話の衝突）を自動的に抽出し、その前後の会話内容とともに提示することで、ユーザーが会話の改善点や特徴的なやり取りを効率的かつ客観的に振り返ることができるようにする。

4. 進捗と考察

現在、会話の定量評価機能を除き、システムの構築を完了している。また、音声データを本システムで解析し、要約・評価・振り返りを実施した結果を協力いただいている看護師に共有し、フィードバックを得た。

要約機能では、ACP 要約のための9項目に基づき関連発話を自動的に整理して出力する。図2は、実際に仮想 ACP 会話に対して本機能を適用した出力例を示している。

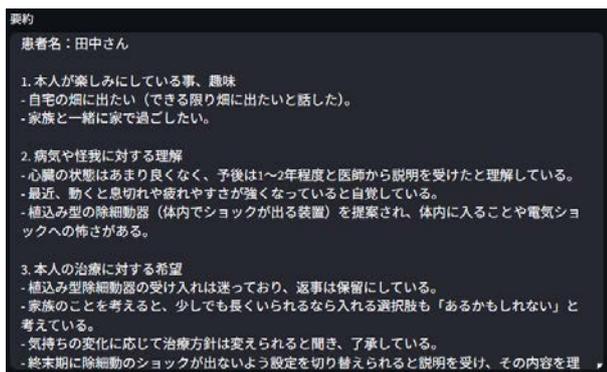


図2 要約機能

図2に示すように、各項目は見出し付きの箇条書き形式で提示され、ACP 記録において重要な観点ごとに簡潔に情報が整理されている。これにより、会話全体を確認することなく、ACP 会話の要点を効率的かつ体系的に把握することが可能である。看護師からは「要約は項目ごとに良くまとまっている」、「まとめて振り返ることで、同じ方に別の場面で聴く際にとっても参考になる」との評価が得られた。

チャット機能では、チャットボットが書き起こしテキストを直接参照しながら応答を生成し、ユーザーは出力された要約を基に、気になった点について対話的に追加の確認を行うことができる。図3は、図2における「植え込み型除細動器の受け入れは迷っている」という要約結果を起点として、その理由について詳細な確認を行っている例を示している。

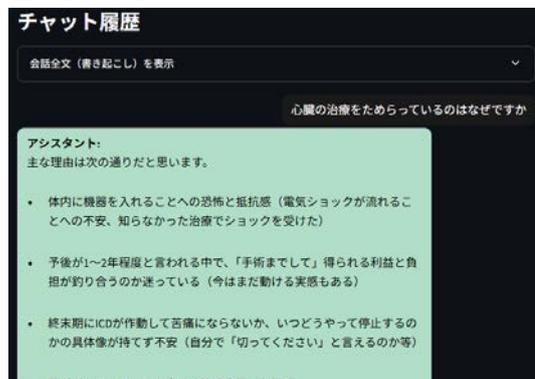


図3 チャット機能

図3に示すように、利用者は要約結果を踏まえつつ、「なぜその判断に至ったのか」といった点を質問することが可能であり、要約のみでは十分に把握できない背景情報や補足的な意図理解を支援する。

定性評価機能では、ACP 会話の書き起こしテキスト全体を入力として、あらかじめ定義した評価観点および評価基準に基づき、LLM による採点およびコメント生成を行う。仮想 ACP 会話に対して本機能を適用した出力例を以下に示す。

(1) 会話の挨拶 10点

- コメント: 冒頭で時間の都合を確認し同意を得て開始。終了時は感謝を伝え、追加の話題有無を確認し、今後も気軽に相談できる旨を伝えて締めている。プライバシー確保（場所・同席者）の確認は明示なし。

(2) 直接表現の回避 9点

- コメント: 趣味（登山・庭いじり）、家族（妻・息子・孫）、食の話題、地域サービス（デイサービス）などの世間話から、価値観（家族に迷惑をかけたくない、家に帰りたい）と支えになる人、退院後の楽しみを自然に引き出せている。延命観などの核心的価値観の掘り下げは今後の課題。

(3) 必要項目 9点

- コメント: 1 趣味○、2 理解○、3 治療希望は延命治療/蘇生の希望が未確認のため 1/2 点、4 心配○、5 動向○、6 支え○、7 見直し○、8 退院後○、9 病院側○。不足: 延命治療や蘇生の希望、具体的な代理意思決定者の指名。

(4) 患者の意思の尊重 8点

- コメント: 全体として尊重的だが、やや誘導的と受け取れる表現が散見。「そうなるとちょっとおじいちゃん頑張って手術受けてもらいたいなという気持ちはありますけどね。」「…前向きに検討していただければ…」 「ぜひぜひそうしてみてください」など。確認型の「～ですよね?」の多用も一部あり。

このように、各評価項目について点数とともに、評価理由や不足している点が文章として提示される。特に、相対的に低い評価となった項目である「患者の意思の尊重」では誘導的と受け取られた具体的な発話が抜き出されており、どの表現に改善が必要なのかを容易に把握できる。また、「必要なことを聞き出せているか」の評価項目では、ACPの要約項目ごとに十分に聞き取れているかが示され、不足している項目が明示される。これにより、次回のACP会話において新たに確認すべき内容が明確となり、継続的なACP実践を支援するフィードバックとして機能する。

看護師からは「評価も“確認内容別”に整理されており、確認されていない部分が明確になっている」との意見が得られる一方で、「倫理的配慮を点数化するのは難しい」との指摘もあった。

全体として、「音声録音を期待してメモを取らないため、話に集中でき、結果として質の高い会話につながっている」との意見が寄せられた。これらの結果から、本システムはACP会話の記録および振り返りを支援し、医療従事者の自己省察や会話の質の向上に寄与する有効なツールとなる可能性が示唆された。

音声分析による定量的評価機能としては各話者の発話時間と比率、話速およびターン数、沈黙、会話の衝突、そして声のトーンと抑揚の表示機能が実装済みである。開発したシステムでは、これらの指標をユーザーが直感的に把握できるよう、Streamlit上でダッシュボード形式で提示する画面を構築した(図4)。

现阶段のシステムには、算出された数値に対して自動的な評価までは行わず、客観的な値の提示にとどめている。しかし、複数の会話データを比較分析することで、これらの指標が会話の質を反映する有意な特徴量となり得ることが確認された。特に興味深い結果として、ACPに熟練した看護師が医療従事者役として対話を行ったケース(通常ケース)と、同じペアで役割を交換し、不慣れな立場で対話を行ったケース(役割交換ケース)の約30分の会話の比較において、会話の衝突(オーバーラップ)の発生回数に顕著な差が見られた。通常ケースでは会話の衝突回数は47回にとどまった。これは、熟練者による「傾聴」の姿勢が表れたものと考えられる。対して役割交換ケースでは衝突回数は154回に達した。役割の不慣れさからくる焦りや、会話の間合いの取りづらさが、相手の発話への割り込みや同時発話の増加として明確に現れた結果と言える。



図4 評価機能(定量評価)

さらに、本システムでは長い沈黙の前後の会話を提示する機能を備えている。この機能により、単に沈黙の発生回数や継続時間を把握するだけでなく、沈黙が生じた文脈を具体的な発話内容とともに振り返ることが可能となる。ACP会話において長い沈黙は、重要な意思決定や価値観の表出に先立つ内省の時間として機能する 경우가多く、その直前には核心的な問いかけや感情的負荷の高い発話が含まれることが少なくない。したがって、沈黙前後の会話区間を明示的に提示することは、会話の転換点や参加者の心理的反応を特定する手がかりを与える。特に、感情的負荷の高い発話に対して長い沈黙が生じた例として、4.05秒間の沈黙が観測された会話区間を以下に示す。

8:4.05 秒間の沈黙(25分53秒付近)

前(SPEAKER_1): そんな時は殺してくれって思ってます。

後(SPEAKER_2): そうですね。まあ、在宅三層をその酸素ボンベ使いながらですね。まあ、その旅行とかあのさされている方も、まああのいらっしやるには。

この発話区間では、病状が強く出現した状況を想定した強い苦痛や切迫感がSPEAKER_1によって言語化されている。このような感情的負荷の高い発話に続いて沈黙が生じている点は、沈黙が単なる会話の停滞ではなく、聞き手が発話内容を受け止め、応答の方向性を模索するための時間として機能していた可能性を示唆する。実際、この沈黙の後に続くSPEAKER_2の発言は、「そうですね。まあ、在宅三層をその酸素ボンベ使いながらですね。まあ、その旅行とかあのさされている方も、まああのいらっしやるには。」というように、感情的表現に直接応答するのではなく、状況を緩和する具体例を提示する方向へと話題が展開されてい

る。このように、長い沈黙の前後の会話を具体的に提示することで、沈黙が生じた背景や会話の転換過程を明確に把握することが可能となり、ACP 会話における支援的応答のあり方を振り返るための有効な手がかりを提供する。

5. 今後の方針

今後は、本システムの分析機能および実証的検証の両面において、さらなる拡張を行う予定である。

まず、音響特徴量に基づく分析機能の高度化として、ピッチ変動が顕著に生じた会話区間を自動的に抽出・提示する機能を追加する。これにより、感情表出や強調が音的に顕在化した場面を特定し、長い沈黙や会話の重複といった既存指標と併せて、ACP 会話における重要な局面を多角的に把握できるようにする。

次に、定量的指標および抽出された会話区間のデータを継続的に蓄積・分析することで、患者の価値観や治療に対する希望を尊重し、将来の医療・ケアについて必要な事項を対話的に共有・確認できている会話を、本研究における「理想的な ACP 会話」として相対的に特徴づける。具体的には、熟練した医療従事者による ACP 会話において観測される発話内容および会話構造（発話量、沈黙、会話の重複等）の指標値の分布や傾向を明らかにし、評価の目安となる基準値の設定を目指す。将来的には、これらの基準に基づき、会話に対する具体的な改善点をフィードバックとして提示する機能へと発展させることを想定している。

さらに、システムに対する評価方法についても段階的な拡充を行う。これまでのように解析結果を共有し意見を得る段階にとどまらず、実際に医療従事者に本システムを使用してもらい、その使用体験に基づく評価を実施する。具体的には、使いやすさ、有用性といった観点からの定性的評価に加え、会話構造指標や音響指標の変化といった定量的評価を行う。これらを総合的に分析することで、本システムが ACP 会話の振り返り支援ツールとして、実臨床においてどの程度有効に機能するかを検証する予定である。

6. おわりに

現在、日本における ACP (Advance Care Planning) の認知度は依然として低く、実施率も伸び悩んでいる。その要因として、ACP を担う医療従事者の経験や教育の不足に加え、患者との会話を記録する際の技術的課題が挙げられる。特に、患者の価値観や希望を過不足なく簡潔にまとめる「記録」の段階では、記述の質にばらつきが生じやすく、重要な情報が見落とされるおそれがある。本研究で構築したシステムは、音声から重要な発言を項目別に自動抽出・要約し、対話形式で内容を確認できる機能を備えており、これらの課題を解決する一助となる。特に、チャットボットによる柔軟かつ双方向的な情報参照機能により、従来の紙媒体や静的な要約では困難であった動的な振り返り

を実現した。これにより、医療従事者の記録作業の効率化と内容の質の安定化に寄与する可能性が示唆された。

さらに、本システムに音声分析を組み合わせることで、発話量、沈黙、発話の重複、およびピッチ変動といった、会話の進行や感情表出に関わる指標を定量的に把握することが可能となる。これにより、記録内容だけでは捉えにくい、会話の間合いや応答の適切さ、感情的負荷の高い局面といった側面を可視化でき、ACP 会話の質を多角的に振り返るための新たな視点を提供する。

本研究では、臨床データの利用に伴う個人情報保護や当業者への心理的配慮といった課題を考慮し、看護師や医師のトレーニングを兼ねた仮想的な ACP ロールプレイング会話を、本システムの検証および調整を行うための入力データとして用いた。その結果、本システムが ACP 会話の振り返り支援として有効に機能し得ることを確認するとともに、仮想 ACP 会話、その要約・評価結果、および音声分析に基づくフィードバックを教育資源として活用することで、医療従事者のトレーニングや ACP 教育への応用可能性を示した。これらの結果を踏まえると、本研究で得られた知見は、将来的に実際の ACP 会話を対象とした振り返り支援および ACP 教育へと展開するための基盤として位置づけられる。その際には、個人情報を含む会話データを安全に処理するための技術的要件を満たす必要があり、外部サービスに依存しないローカル LLM をはじめとした、情報セキュリティに配慮した技術の選定およびシステム構成が求められる。

参考文献

- [1] Detering, K. M., Hancock, A. D., Reade, M. C., & Silvester, W. (2010). *The impact of advance care planning on end of life care in elderly patients: Randomised controlled trial*. *BMJ*, 340, c1345.
- [2] Fraile Navarro, D., Coiera, E., Hambly, T. W., Triplett, Z., Asif, N., Susanto, A., Chowdhury, A., Azcoaga Lorenzo, A., Dras, M., & Berkovsky, S. (2025). *Expert evaluation of large language models for clinical dialogue summarization*. *Scientific Reports*, 15, Article 1195.
- [3] Daryanto, T., Ding, X., Wilhelm, L. T., Stil, S., Knutson, K. M., & Rho, E. H. (2025). *Conversate: Supporting reflective learning in interview practice through interactive simulation and dialogic feedback*. *Proceedings of the ACM on Human-Computer Interaction*, 9(1), Article GROUP9, 1-32.
- [4] 辻本 海成, 角 康之. (2020). チュータリング対話の質評価に寄与する非言語情報の検討. インタラクシオン 2020 (情報処理学会), オンライン開催, 2020 年 3 月 9 日~11 日, pp. 39-47.
- [5] 澤本 祐一, 神山 祐一, 平野 靖, 梶田 将司, 間瀬 健二, 鈴木 富雄, 勝山 貴美子, 山内 一信. (2010). モチーフ抽出によるマルチモーダルインタラクシオン解釈手法の提案と医師-患者対話要約への応用. *情報処理学会論文誌*, 51(2), pp. 334-345.