

WordCloud を用いた PC 利用時の発話可視化手法の提案

友広 純々野* 塚田 浩二*

概要. 近年、スマートフォンやスマートウォッチの普及により、個人のライフログを残すことが容易になった。またリモートワークや学習活動など、PC で利用する機会が増えている。例えば、オンライン会議や少人数のオンライン対話、eラーニングや授業で動画を視聴するような活動である。本研究ではこうした、PC 本体から流れる音声（PC 音声）と PC を利用しているユーザの発話がライフログとして利用できると考えた。本稿ではこうした PC 音声／ユーザの発話を可視化するシステムの提案、実装について述べる。

1 はじめに

スマートフォンやスマートウォッチの普及により、個人のライフログを残すことが容易になった。これらのライフログには画像や位置情報、身体データなどを利用したものが多い。また、音声認識技術の向上から、会議の円滑な進行を目的とした試み [1] が行われている。しかし利用場面が限定されており、ライフログとして個人が発話情報を利用する試みは少ない。これは、大量の音声／テキストデータの確認に手間がかかることが一つの要因だと考える。

そこで本研究は、話者が気軽に日常を振り返ることを目的に、個人の発話をライフログとして扱い、発話の概要を可視化するシステムを提案する。特に、日常生活の中でも PC の利用場面の発話を対象とした。この理由は、多くの PC はマイク／スピーカーを備えており、かつコロナ禍においてオンライン会議や動画によるオンデマンド学習等の利用機会が増加したためである。本研究では、この PC 本体から流れる音声（PC 音声）と PC を利用するユーザの音声を同時に記録する。さらに、両者を WordCloud 等を用いて可視化することで、PC の利用場面と発話内容を同時に可視化する振り返りシステムを構築する。

2 関連研究

本研究に関連する研究事例として、発話を用いた音声ライフログの事例を示す。

発話音声を利用した研究事例として、Krosnick らは [3] は、ユーザの発話量をリアルタイムに可視化することで、モデリング等の制作プロセスを発話を元に記述する手法 Think-Aloud Computing を提案した。被験者へは、プロトコル法を応用し成果物

作成中の思考を発話するよう教示しつつ、発話量の可視化により積極的な発話を促している。長利ら [5] は、コーディング中の発話とコードを紐づけて記録するシステムを開発した。コード中には記録されなかった話者のアイデアや考察などの確認を目指している。友広ら [4] は、動画視聴中にリアルタイムに発話を可視化する手法を提案し、発話への影響を調査した。

また、音声認識を用いた可視化の事例として、聴覚障害者のコミュニケーションを支援するために、WordCloud を用いた会議のトピック提示 [2] が行われている。

3 提案

本研究では、PC 本体から流れる PC 音声と、その PC を利用しているユーザの発話を同時に記録し、両者を WordCloud 等で可視化するシステムを提案する。

本システムは大きく発話の取得部分と可視化部分に分けられる。まず PC 音声とユーザ発話を分離して記録し、音声認識を用いてテキストデータを取得する。このテキストデータから、インタラクティブに生成される WordCloud を中心として振り返ることができる可視化ツールを実装する。



図 1. システムの概要

Copyright is held by the author(s). This paper is non-refereed and non-archival. Hence it may later appear in any journals, conferences, symposia, etc.

* 公立はこだて未来大学

4 実装

4.1 PC 音声とユーザ発話の個別記録

提案手法では、PC 音声とユーザ発話を分離して取得する必要がある。ユーザ発話は PC 内蔵/外付けマイクを用いることで容易に記録可能であるが、PC 音声は標準設定では記録対象とならない。そこでループバック機能を持つソフトウェアミキサー (VoicemeeterBanana¹) を利用して、PC 音声を個別に記録できるよう工夫した。

次に PC 音声とユーザ発話を、別個に音声認識サービス (GCP Speech-to-Text²) に入力し、テキストデータとタイムスタンプを取得する。これらを一つの CSV 形式にまとめて、PC のローカルストレージに保存した。

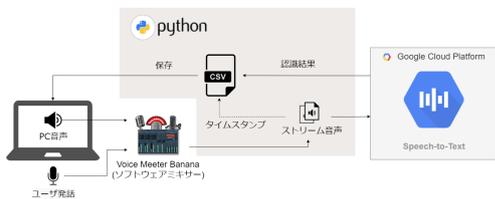


図 2. PC 音声とユーザ発話のテキスト化処理

4.2 WordCloud を利用した可視化ツール

大量の発話情報を一定期間ごとにまとめて振り返るために、インタラクティブな WordCloud を用いた可視化ツールを実装した。WordCloud は前節で取得したテキストデータを Mecab³で形態素解析して、頻出単語と品詞を取得して Python 用ライブラリ⁴を用いて生成した。



図 3. 振り返りツールのスクリーンショット

また図3に、振り返りツールの外観を示す。品詞と時刻範囲を指定することで、対応した WordCloud を PC 側/ユーザ側で同時に生成できる。時刻情報は、現時点では1時間単位でまとめて選択することができる。品詞情報は、発話内容の理解には名詞

を、ユーザの口癖などには形容詞等が利用できると考えた。

著者が3つの異なる状況 (表1) でデータを収集し、名詞、形容詞で視覚化した WordCloud の一例を表2, 3示す。s1 はオンラインで行われた研究室ゼミ, s2 は友人との1対1のオンラインゲーム, s3 は一人で解説動画を視聴しコーディング作業の様子である。

表 1. データ収集の状況

収集時状況の状況	他者と対話の有無	収集される発話	
		利用者側	PC側
s1 ゼミの様子-オンライン	○ (1対多 - 1対1)	著者	教員 生徒 (複数人)
s2 友人とオンラインゲーム中	○ (1対1)		相手の声
s3 解説動画を視聴しながらの コーディング作業	× (一人)		動画内の人物

表 2. 状況ごとに生成された WordCloud (名詞)

状況	s1	s2	s3
ユーザ 発話	Google, 先生, あと, スライド	人, あと, スペシャル	利用可能なデータ なし
頻度	先生, フォーム, データ, スライド	ナイフ, ナイス, あと, ミサイル, 爆弾, アサリ	次, キー, メッセージ, チェット, データ

表 3. 状況ごとに生成された WordCloud (形容詞)

状況	s1	s2	s3
ユーザ 発話	やばい, 早い, なかなか, 懐かしい	難しい, やばい, うまく, やばい, 強い	おかしい
頻度	悪かった, 大きい, なかなか, 早い, やばい	やばい, 難しい, 強い, 早い	新しく, 短い, ほしい, 古い

名詞情報を活用することで、s1~s3でユーザの記憶と合致する内容を含んだ WordCloud が作成された。一方 s3 では、ほぼユーザによる発話はなく、名詞情報で WordCloud の可視化はできなかった。また形容詞情報を活用することで、ユーザが自覚している口癖 (やばい等) や、s3においてコーディングに躓いている様子の単語が含まれた WordCloud が作成された。

謝辞

本研究の一部は、科研費 22H03634 の支援を受けた。

¹ VB-AUDIO, <https://voicemeeter.com/>

² Google Cloud, <https://cloud.google.com/speech-to-text>

³ Mecab, <https://taku910.github.io/mecab/>

⁴ WordCloud, <https://pypi.org/project/wordcloud/>

参考文献

- [1] S. Chandrasegaran, C. Bryan, H. Shidara, T.-Y. Chuang, and K.-L. Ma. TalkTraces: Real-Time Capture and Visualization of Verbal Content in Meetings. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, pp. 1–14. Association for Computing Machinery, 2019.
- [2] R. Iijima, A. Shitara, S. Sarcar, and Y. Ochiai. Word Cloud for Meeting: A Visualization System for DHH People in Online Meetings. In *Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility*, ASSETS '21, pp. 1–4. Association for Computing Machinery, 2021.
- [3] R. Krosnick, F. Anderson, J. Matejka, S. Oney, W. S. Lasecki, T. Grossman, and G. Fitzmaurice. Think-Aloud Computing: Supporting Rich and Low-Effort Knowledge Capture. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–13, 2021.
- [4] 友広 純々野, 塚田 浩二. 動画視聴時の私的発話を可視化するシステムの提案. 第 198 回, pp. 1–6, 2022.
- [5] 長利 槇吾, 寺田 実. 独り言と文字を紐付けて自動記録するテキストエディタ. pp. 710–712. 情報処理学会, インタラクション 2016, 2016.