

# インカム対話の特徴と対話ログ活用のフィージビリティ検証

黒田 由加 山下 卓規 掛村 篤 佐伯 祐太郎

東芝デジタルソリューションズ株式会社 RECAIUS 事業推進室

{yuka.kuroda, takumi.yamashita, atsushi.kakemura, yuutarou1.saeki}@toshiba.co.jp

## 1. はじめに

音声・画像やセンサーデータから人の行動や状況を把握し知識として蓄え、現在の状況と類似する事例を提示することで人を支援するコミュニケーションAI”RECAIUS”は、様々な現場でワークスタイルの革新に活用されている。そのサービスのひとつである「RECAIUS フィールドボイス インカムエディション」[1]は音声認識技術を中心に、インカムの代わりに音声とテキストを同報送信し、情報共有を可能にするサービスである。本サービスを通じてやり取りされる音声とその音声認識結果のうち、研究開発用途に利用の許諾を受け、整備しているものを「インカム対話コーパス」と呼ぶ。

本稿では、インカム対話コーパスの分析によって明らかになったインカム対話の特徴とその活用のフィージビリティ検証について述べる。

## 2. 関連研究

状況適応・行動型サービスでの音声を使った情報共有の仕組みについては、内平らが、音声入力とマイクロログ型のコミュニケーションを融合し、発話内容と業務状況から自動で配信先を決める「音声つぶやきによるコミュニケーションシステム」を提案している[2]。インカム対話の利用シーンは、内平らの行動型と同じく、施設空間を移動しながら業務を行うものである。対話コーパス分析について、伝は、認知過程と相互行為過程の観

点から、話者交替のタイミングと交代時の発話頭の分析を行っている[3]。また、日本の対話研究では雑談対話（非タスク指向型対話）について対話破綻の種類などについて分析を行い[4]、破たん検出などについて活発に議論されている。

## 3. RECAIUSフィールドボイスインカムエディション

サービス業の現場でのコミュニケーションには、直接の会話、電話、そしてインカムが使われてきた。インカムは行動型サービスの現場で、リアルタイムのやり取りには便利である一方、音声の品質が十分ではない、接客中に聞き逃すことがある、同時に複数の発話を受け付けられないなどの課題のほか、インカム端末が大きくて重たい、高価なためスタッフ全員が持てず、結局は情報を伝えたいスタッフを探しまわっているなど、様々な課題があった。

そこで、インカムに代わって、手持ちのタブレットやスマートフォンにインストールした音声認識アプリケーションを用いて、音声と認識結果のテキストを同報送信するサービスを実現した。これにより、小型で軽量の端末で、同時に複数話者の音声を受け付け、認識し、音声と認識結果のテキストを同時に配信することで、聞き逃しの防止だけでなく、業務の振り返りができるようになることを目指してきた。現在、様々な業種業態の現場で使われ始めている。(RECAIUS フィールドボイス インカムエディション)。

我々はこのようなサービスの現場で使われている音声と音声認識結果を、使用者の許可を得て、音

声認識、知識活用などの研究開発に活用している。

本アプリケーションでは、ユーザはイヤフォンマイクを通じて音声を入力する。発話開始時と終了時には、イヤフォンマイクに既設のスイッチボタンを押すプッシュトーク型音声認識を採用している。ボイストリガー型などの構成も実現できるが、充電なしで勤務時間中継続して使用できる利便性を重視し、プッシュトーク型を採用した。発話開始時に認識の準備ができたことをフィードバックすることで、冒頭の発話切れを防ぐなどの工夫をしている。また、サービス業の現場では、宿泊施設のフロント等、音声認識に適さない環境もあるため、その場合はキーボードからのテキスト入力も受け付ける。受け付けたテキストは音声合成システムが生成した合成音声で、テキストと一緒に配信し、利用者の利便性を高めている。

#### 4. インカム対話コーパスの概要

本節では代表的な旅館の対話を収録したインカム対話コーパスについて紹介する。

老舗旅館において、週5日の営業日に、その日に勤務するスタッフ数十名がタブレットやスマートフォンにインストールしたアプリケーションと、イヤフォンマイクを通じて発話された音声とその認識結果のテキストを収録している。



図1 システム概要

利用開始当初は1日平均400-500発話程度だったが、継続して使用した結果、使用開始から1

年以上を経過した2017年12月には1日あたり平均840発話、繁忙日には1000を超える発話がある。6月や9月に少なく、5月や8月の発話数が多いのは祝日・休日などと連動して宿スタッフ数や業務に変動があるためである。

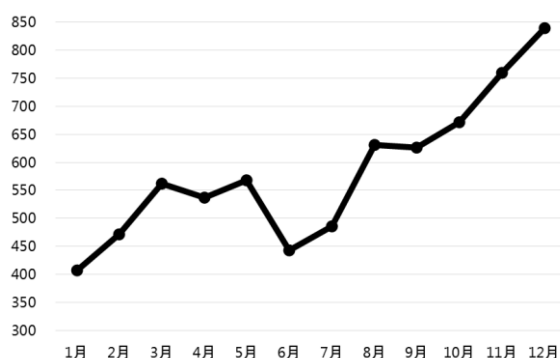


図2 1日あたりの平均発話数

また、時間帯ごとの発話数の変化をみると、図3の通り、11時、15時、18時ごろにピークがあり、それぞれ昼食、チェックイン、夕食での繁忙時となっている。

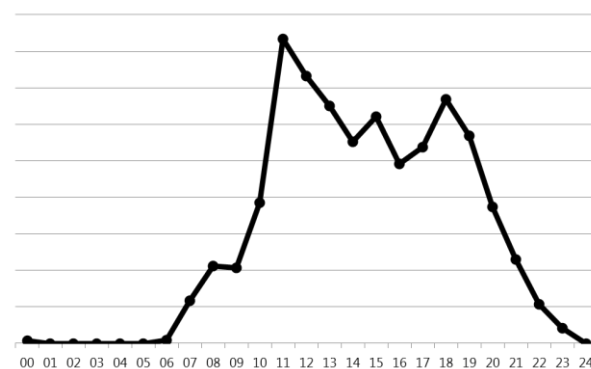


図3 時間帯による発話数のばらつき（単位：時）

このほか、展示会運営スタッフ間の情報共有やキッキング作業における連絡等でも用いられ、18万発話を超える対話データを収録しており、現在も日々蓄積を続けている。

## 5. インカム発話の特徴

以下にインカム発話のサンプルを示す。サンプル中の「しまーす」「えっと」など、音の伸びやフイラーは発話そのままである。

表1 発話サンプル

001	調理場さん、一名様、スタートお願いしまーす。
002	フロントさん、お土産購入、三名様、ロビー案内します。
003	はい、■■さん了解しました。
004	業務連絡です。露天風呂、のれん替え終わりました。お願いします。
005	はい■■さん、ありがとうございます。
006	■■さん、すいません。上に行っていて、もし持てるようでしたら、レストランの炭酸があと1.5しかございませんので降ろしていただけますか。
007	■■さん、片付けどこから入れます、えっと、別館に行った方がいいですかね。
008	デシャップさーん、レストランですー。八〇三のお碗お願いします。
009	はい■■さん、かしこまりました。

インカム発話の特徴を以下に述べる。

- (1) 業務のやりとりのため、基本的には丁寧な話し言葉で発話される。
- (2) 短時間で用件を伝える簡潔な発話である。
- (3) 発話中に相手の発声が割り込まれないため、発話のさえぎりがなく、発話中にあいづちが入ることもない(システムでは1発話最長20秒まで受け付ける)。
- (4) 何らかの目的を持った発話と、それに対するレスポンスが基本の構成であり、非目的指向対話はほぼ含まれない。
- (5) (2)の目的を達成するため、最初に名前を呼び掛けることが多い。
- (6) 一方で、呼び掛けられた人が必ずしも応答す

るわけではなく、業務を遂行できる人が応答する場合がある。

- (7) 発話と応答は必ずしも隣接するわけではなく、サンプル006の応答は009であり、間に他のトピックを挟むことがしばしばある。

これらの特徴は、伝[3]で述べられている、一般的な対話の特徴と異なり、インカムという手段による制約を受けた対話であると言える。

## 6. インカム発話における伝達行為

対話履歴の分析を行うにあたり、伝達行為の分類を行った。人手で分類を付与している。松本ら[5]の分類を参考に、以下の通りの分類を用いた。

挨拶	慣習的な挨拶、発話相手への呼びかけ、作業終了時の声掛けなど。
相槌	相手の発話に対する応答のうち、重要な意味は含まないもの。
応答	相手の発話に対する応答。また、相手の質問に端的に答えているもの。
確認	相手の発話を繰り返したり、聞き返したりして確認をおこなっているもの
報告・説明	相手に、これまでの経緯について何らかの命題内容を伝達しているもの。また、現状を記述する発話。
行為宣言	話者が自分から、これから何事かをおこなうことを言う
指導・忠告	相手に対して未然の物事を勧めるもの
依頼・要求	相手に何事かを頼んでいるもの。命令口調や指示口調も含む
質問	相手、相手の発話に対する質問全般
感性	謝意・相手を安心させる発話・共感

図4 インカムにおける伝達行為の分類

ある1日のデータ500発話を人手で分類したところ、図5のような頻度の分布になった。

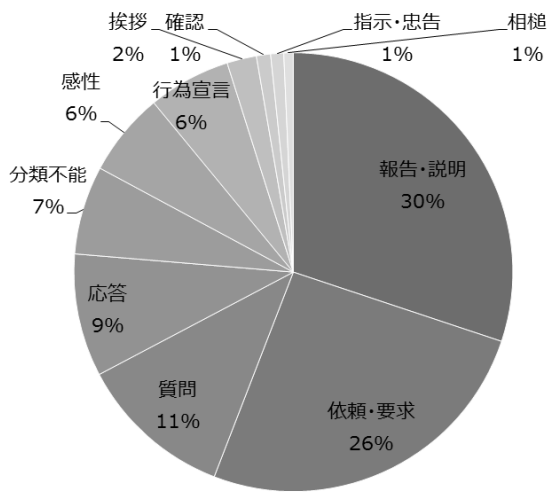


図5 インカムにおける伝達行為の分布

業務の性質上、報告・説明、依頼・要求、質問が多く、その応答が続いて多い。対面対話であれば報告・説明、依頼・要求、質問の多くに応答が必要だと考えられる。一方で、インカムでの対話の場合、発話意図が伝わり、行動を促せば必ずしも応答を必要としないこと、調理場など応答するのが難しい環境があるなど、様々な理由から必ずしも応答要求しないことが、現場観察（エスノグラフィ）から明らかになっており、コーパスからもその傾向が読み取れる。

## 7. ログ活用のフイージビリティ

サービス業において、より質の高いサービスを提供するためには、対話履歴を通じてお客様に関する情報と、提供するサービスに関する情報を共有・記録し、業務に活用することが効果的であると考えられている。

お客様に関する情報と提供するサービスに関する情報は、短時間に活用する情報と、記録することで次のサービスに活かせる情報とに二つに分類可能である。短時間に活用する情報とは、瞬間あるいはそのサービス期間（宿泊ならチェックアウトまでの期間）限り有効な情報を指し、記録することで次のサービスに活かせる情報とは、たとえば食事の嗜好や浴衣のサイズなど、次回のサ

ービス時にも活かすことで、より質の高いサービスにつなげられる情報を指す。このように、現場での発話内容をサービス向上につなげる取り組みを、実証実験を通じて行っている。

## 8. まとめと今後の展望

音声と音声認識結果であるテキストを同報送信するサービスを通じて、サービス業をはじめとする様々な現場で対話履歴を収集し、研究開発用途のコーパスを構築してきた。現在 18 万件を超える発話を収録しており、現在も日々蓄積している。また、情報共有という即時的な効果に加え、対話の中からサービス向上につながる発話を発見し、業務にフィードバックするための取り組みを行っている。インカム使用時や RECAIUS フィールドボイスインカムエディション導入前に抱えていた現場の課題を解決し、現場で働く人の支援につながる取り組みに焦点を当ててゆく。

## 参考文献

- [1] RECAIUS フィールドボイスインカムエディション  
<http://www.toshiba.co.jp/cl/pro/recaius/lineup/fieldvoice.html>
- [2] 内平直志: 音声つぶやきによる気づきの収集と活用で看護・介護サービスの質を向上する, サービスロジー, 1 巻 (2014) 2 号 pp. 14-17
- [3] 伝康晴: コーパス言語学的手法による音声対話の分析: 認知・相互行為背反の観点から (ヒューマンコミュニケーション基礎), 信学技報 116(185), pp.19-24, 2016-08-19
- [4] 東中竜一郎, 船越孝太郎: 対話システムにおけるエラー分析—雑談対話コーパスの収集と対話破綻検出—, 情報処理 57(1), 42-43, 2015-12-15
- [5] 松本斉子, 村井源, 住住彰文: 看護師対話ログにおける行為遂行の分析, 日本認知科学会第 25 回大会発表論文集, pp.25-32, 2008