

## 指示に従った発話や必要な情報の自動取得を行い、人間の様に自然な対話 ができる音声対話技術を確立

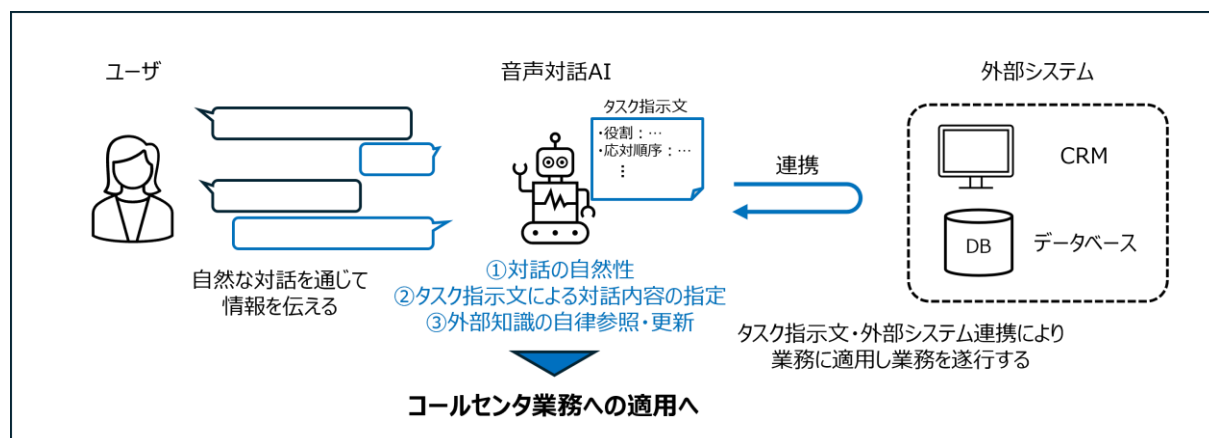
～コールセンタの対応品質向上とオペレータ負荷軽減をめざす～

発表のポイント:

- ◆ 相手の発話を聞きながら応答し、相槌、発話待ち、割込みに対応できる双方向型音声対話 AI 技術を開発し、他社サービスと比較して平均応答時間を 50%以上短縮、人間のオペレータに迫る応答速度を実現したことを検証しました
- ◆ 指示文の変更により業務ごとのトークフローへ対応し、対話中にデータベースや CRM などの外部システムを参照・更新します
- ◆ 人手不足が深刻化するコールセンタ業務への適用をめざし、自動対応の品質向上による顧客体験(CX)の向上と、オペレータの負荷軽減による運営効率化への貢献が期待されます

NTT 株式会社(本社:東京都千代田区、代表取締役社長:島田 明、以下「NTT」)は、相手の発話を聞きながら応答し、相槌、発話待ち、割込みに対応できる双方向型音声対話 AI 技術を開発しました。本技術は、業務ごとの指示文に沿って対話を進めるとともに、対話中にデータベースや CRM などの外部知識を参照しつつ指示に従った発話が可能です。

本研究成果は、2026 年 11 月 18 日～25 日に開催される NTT R&D FORUM 2026 ※1 に展示予定です。



## 1. 背景

近年、人手不足や業務負担の増加を背景に、コールセンタや窓口などの顧客対応業務では、AIを活用した業務の自動化・効率化への期待が高まっています。特に音声による対応は、電話や対面で幅広く利用されており、人に代わって自然に対応できる音声対話 AI の活用が期待されています。

一方、実際の顧客対応に AI を活用するためには、単に質問に回答だけでなく、人間同士の会話のような自然なやりとりが求められます。例えば、ユーザの発話を適切に待つ、自然なタイミングで相槌を打つ、ユーザからの割込みに応じるなど、会話の流れに合わせた適切なタイミングでの発話が重要です。

また、実際の顧客対応業務への活用には、業務や役割に応じたトークフローで会話を進めるとともに、ユーザの意図や対話の状況を把握し、必要な情報を外部から参照・更新しながら正確に対応することも求められます。

NTT では、人間同士の会話のようなやり取りである対話の自然性に加え、タスク指示文による対話内容の指定や外部知識の自律参照・更新による対話内容の制御が可能な双方向型音声対話技術の確立に取り組みました。

## 2. 技術のポイント

### ①対話の自然性

ユーザの音声を聞きながら同時に発話できる双方向型音声対話により、自然なタイミングでの応答や相槌、適切な発話待ち、ユーザからの割込みへの対応を実現します。人間同士の会話に近い、相手の発話や会話の流れに応じて「話す・聞く・待つ」を柔軟に切り替えることで、テンポのよい自然な音声対話を可能にします。

### ②タスク指示文による対話内容の指定

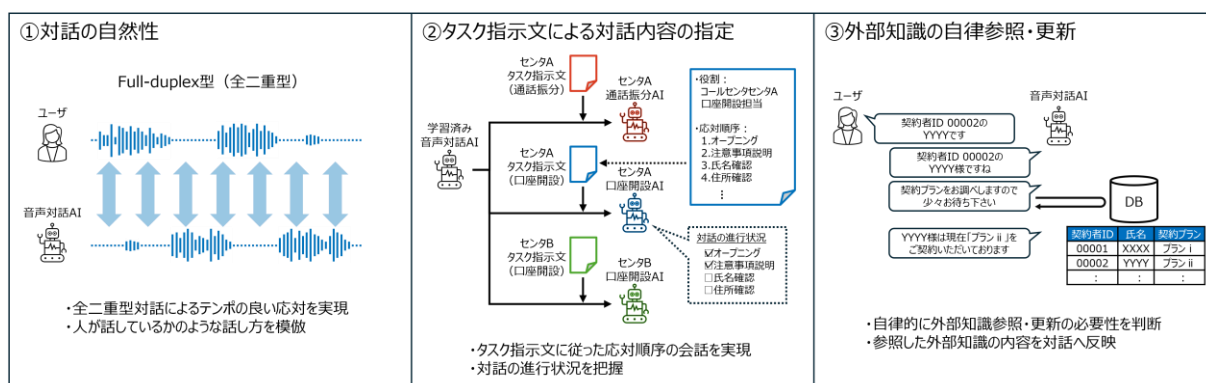
AI に与えるタスク指示文に基づいて、役割や業務内容、対応方法を理解し、それぞれの業務に適した対話を行います。タスク指示文を変更することで、モデルを再学習することなく、予約受付や問い合わせ対応など多様な対応業務に応じたトークフローでの会話を可能にします。また、対話の進行状況やユーザから得られた情報を把握・出力することで、業務の流れに沿った一貫性のある対応を実現します。

### ③外部知識の自律参照・更新

対話の内容やユーザの意図に応じて、AI が必要な外部機能を自律的に呼び出し、データベースや外部システムの情報を参照・更新します。AI が持つ知識だけに依存せず、最新の商品情報や予約状況、顧客情報などを対話中に活用することで、実際の業務に必要な正確な回答や手続きを可能にします。

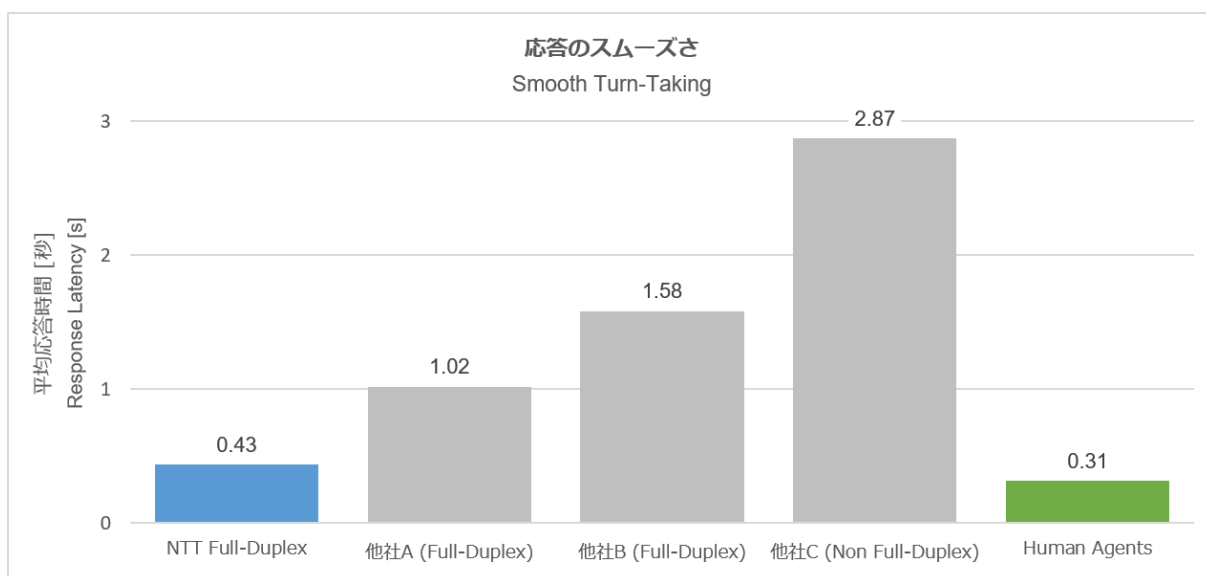
これら 3 点の技術確立においては、特に AI に学習させる音声対話データの準備が重要です。NTT が、長年の研究により培ってきた音声強調や音声認識、音声合成などの音声処理技術を活用し、対話音声データの高品質化に取り組むと共に、AI に期待する振る舞いを含んだ疑似音声対話データの生成にも取り組み学習へ活用します。それにより、対話タイミングの自然性においてより高品質

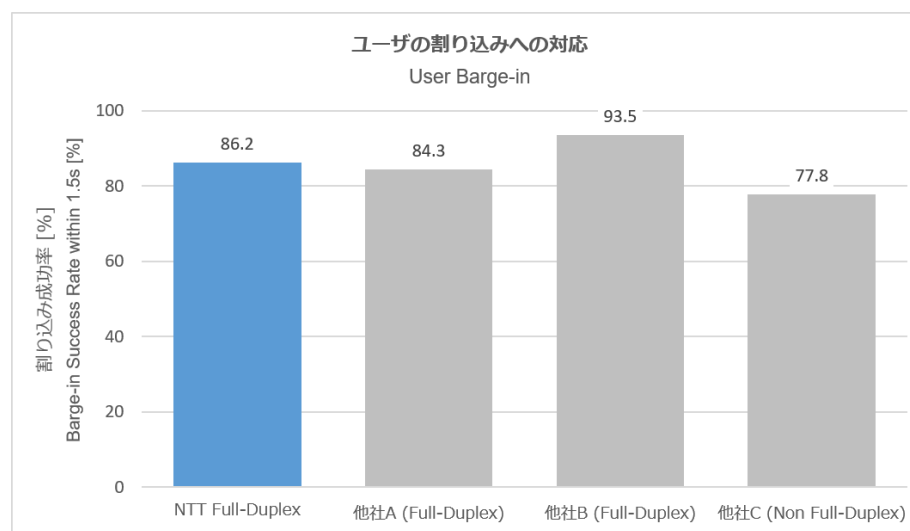
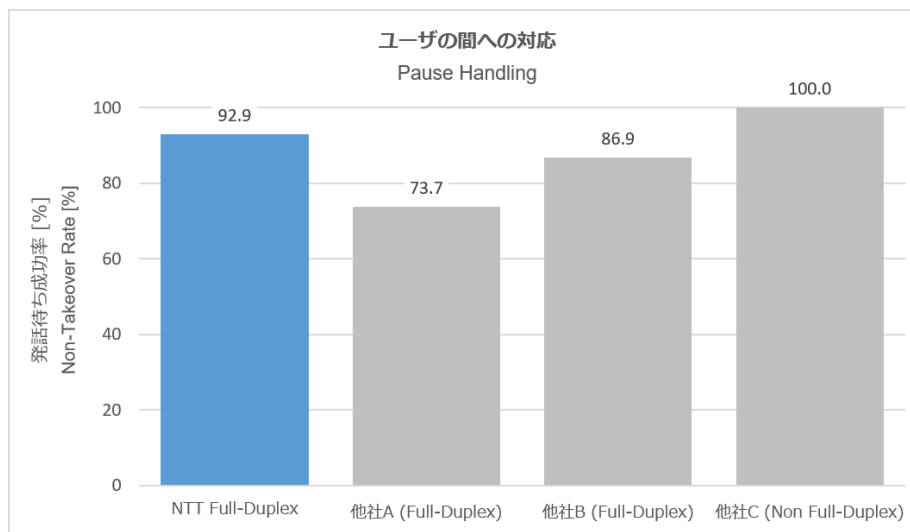
な音声対話技術を確立しました。



### 3. 実験の概要

対話の自然性を構成する要素のうち、人間と対話 AI との対話のテンポ（応答の時間的な特徴）を評価しました。具体的には、対話 AI の評価において用いられる、平均応答時間、発話待ち成功率（ユーザが発話中に間を空けた際に、発話の続きを待つことができる割合）、割り込み成功率（対話 AI の発話中にユーザが割り込んだ際に、対話 AI が発話を停止できる割合）の3つを評価指標とし、日本語のコールセンタ対話を模した 400 対話を用いて評価しました。





市中の商用対話 AI との比較評価を行った結果、開発した対話 AI は平均応答時間 0.43 秒を達成し、他社サービスと比較して 50%以上の短縮を実現しました。これは、日本オペレータの平均応答速度 0.31 秒に迫る水準です。さらに、発話待ち成功率や割り込み成功率においても、他社サービスに匹敵する性能を達成しました。これらの結果から、本技術に基づく対話 AI が、日本語での会話において人間に近い自然なテンポで対話するために必要な性能を備えていることを確認しました。

#### 4. 今後の展開

##### 1.)技術の応用範囲

本技術は、人手不足が深刻化するコールセンタ業務において自動対応の品質向上による顧客体験(CX)の向上と、オペレータの負荷軽減による運営効率化への貢献が期待されます。本技術を活用することで、自治体や医療機関における窓口案内、ホテルや商業施設での受付・接客、企業内の問い合わせ対応など、人との自然なコミュニケーションが求められる幅広い業務への展開につながります。

## 2.)次のステップおよびロードマップ

今後、応用検討を進めるとともに、様々な外部システムとの連携を進めることで、より充実した内容の対話実現に貢献します。

本研究成果は、2026 年 11 月 18 日～25 日に開催される NTT R&D FORUM 2026 に展示予定です。  
また、コールセンタでの実運用を見据え、コールセンタ業務への適用可能性を検証するための共同実験を現在計画中です。

### 【用語解説】

※1. NTT R&D FORUM 2026 <https://www.rd.ntt/forum/2026/>



■ 本件に関する報道機関からのお問い合わせ先

NTT 株式会社

サービスイノベーション総合研究所

広報担当

[問い合わせフォームへ](#)