

# スムーズな多言語交流を実現するためのARによる会話支援システム

— PredicTalk：誰もが外国語を話せる世界へ —

## 1. 背景

異なる言語を話す人々が円滑にコミュニケーションを取るためのツールとして、これまでに多くの翻訳アプリケーションが開発されてきた。これらのアプリケーションは、ユーザーの発話をリアルタイムで翻訳し、画面に表示したり音声として出力することで、外国語を話す必要がなくなるようにして、言語の壁を越えた意思疎通を可能にしている。

しかし、既存の翻訳アプリケーションにはいくつかの課題がある。第一に、ユーザーの発話を代替する形で機能するため、ユーザー自身が話す機会を奪ってしまう。第二に、翻訳した文章を見るために画面を注視する必要があり、会話相手とのアイコンタクトが減ることで、自然なコミュニケーションが阻害される。

実際に我々は、短期留学や海外旅行を通じて、こうした課題を実感した。例えば、フランス留学において、現地の学生と英語で交流をした際に、会話の途中で単語や適切な表現が思い浮かばず詰まってしまうことがあり、そのたびに翻訳アプリケーションに頼らざるを得なかった。翻訳アプリケーションは正確に訳してくれるため助かったが、一度アプリケーションを使い始めると、会話が終了するまで使うのをやめるタイミングがなく、気づけば会話のほとんどが画面越しのやりとりになってしまっていた。結果として、相手と目を合わせる機会が減り、会話の自然な流れが失われるという問題を感じた。

## 2. 目的

本プロジェクトでは、既存の翻訳アプリケーションが抱える課題を解決するため、ユーザーの発話を支援しながらも、自然な会話の流れを維持できるシステムの開発を目指す。本システムでは、ユーザーが会話の途中で適切な単語や表現が思い浮かばず詰まってしまった際に、次の発話の候補を言語モデルがリアルタイムで予測し、ARグラスの視界内に候補を表示する。これにより、ユーザーは会話相手の顔を見ながらスムーズに会話を継続でき、翻訳アプリケーションに依存することなく、自らの言葉で話すことが可能となる。

## 3. 開発の内容

本プロジェクトでは、Brilliant Labs製のARグラスであるFrameを活用して上記の目的を果たす、iOSアプリケーション「PredicTalk」を開発した。PredicTalkは音声認識、次文予測、データ管理などの機能を持ち、FrameとBluetoothで接続することでリアルタイムの音声データのやり取りを行う。次文予測にはOpenAIのAPIを使

用し、ユーザーの発話を解析した上で、次に話すべき文章の候補を予測する仕組みを採用している。

本システムの構成は、iOSアプリケーションとFrameのデバイスドライバーの大きく2つの要素に分かれる。

iOSアプリケーションは、ユーザーインターフェースの設計にも工夫を凝らしており、直感的に操作できるデザインを採用している。

図1に示すホームタブでは、システムが音声認識したユーザーの会話のテキストと、そのテキストの続きを予測した3つの次文予測候補を表示することで、ユーザーに対してシステムの状況をフィードバックしている。

図2に示すFrameタブは、アプリケーションと接続したFrameの情報を表示し、ユーザーへ伝えるための画面である。この画面ではFrameのバッテリー残量、デバイス名、ファームウェアのバージョン、ファームウェアのGit Tag（ファームウェア管理用の番号）、MACアドレスを表示し、また下部にあるボタンを押すことで後述するデバイスドライバーをアプリケーションからFrameへ送信することができる。

図3に示す設定タブは本システムの設定を変更するための画面であり、会話に詰まってから新しい予測を提示するまでの時間と、よりユーザーに最適化された予測を実現するために予測モデルへ入力するユーザー情報を編集することができる。



図1 ホームタブ

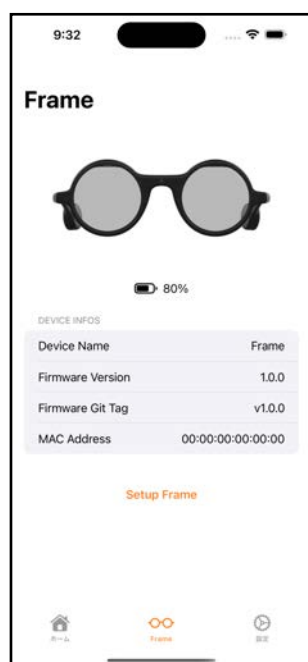


図2 Frameタブ



図3 設定タブ

FrameのデバイスドライバーはLua言語で記述されており、Frameの各状態での画面表示やユーザーからの様々な入力を処理する機能、スマートフォンとのBluetooth通信機能などを担っている。Frameに表示された情報が実際にどのように見えるかを示した画像を図4～6に示している。

ユーザーはFrame起動後（図4）、Frame側面をタップすることでFrameのマイクを起動して録音を開始することができる（図5）。録音された音声はスマートフォン側でリアルタイムに送信・認識され、OpenAIのAPIを用いて次に話す内容が推定される。スマートフォンから受信した予測候補はFrameのディスプレイへ表示され、ユーザーは次に何を言えば良いのかのヒントを得ることができる（図6）。この状態でFrame側面のタップすると、予測候補を切り替えることもできる。こうした構成により、ユーザーが会話中に詰まったときに迅速に次の文章候補を参照でき、自然なコミュニケーションを継続することが可能となっている。

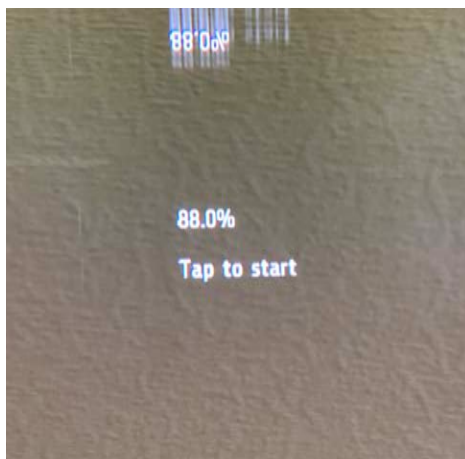


図4 起動直後の状態

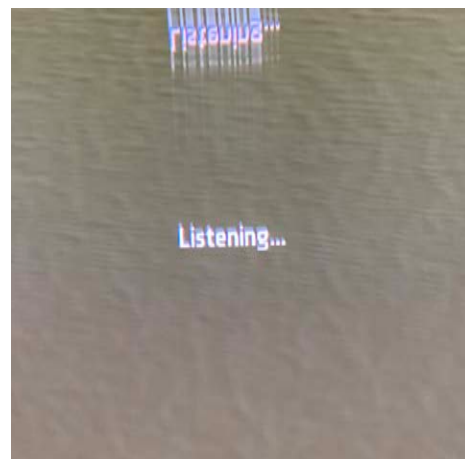


図5 ユーザーの発話を待っている状態



図6 予測したユーザーの発話の続きを表示した状態

#### 4. 従来の技術（または機能）との相違

従来の翻訳アプリケーションは、言語の壁を越えてスムーズな意思疎通を可能にする点で優れているが、一方で「ユーザー自身が英語を話す機会を奪う」という課題がある。翻訳機能が主体となることで、ユーザーは言語を学ぶ意欲を失い、相手と

の会話も「翻訳された情報のやり取り」に留まる可能性がある。また、翻訳の精度や文脈理解の限界により、不自然な会話になってしまうこともある。

本システムは、翻訳アプリケーションとは異なり、ユーザーが「自分で英語を話すこと」を前提にしている。ARグラスに提示されるのは、翻訳された文章ではなく、ユーザー自身が話すためのヒントである。そのため、会話の主体は常にユーザーにあり、翻訳システムのように「機械が会話を代行する」形にはならない。この点で、本システムは単なる翻訳ツールではなく、「会話のスムーズな進行を支援するツール」としての役割を果たす。

## 5. 期待される効果

本システムは、ユーザーが自ら発話することをサポートするものである。これにより、ユーザーは会話の主導権を維持したまま会話を楽しむことができるようになる。また、言葉に詰まる不安が軽減されるため、ユーザーが積極的に他言語での会話ができるようになる。そのため、語学力向上のためのツールとして活用することも可能である。

さらに、ARグラスを通じて情報を視界内に表示する仕組みによって、会話の自然な流れを保ちながらサポートを受けることが可能となり、従来の翻訳アプリケーションが抱える、視線がアプリケーションに集中して会話相手とのアイコンタクトが減少するという課題を解決している。

## 6. 普及（または活用）の見通し

本システムのターゲットは、英語学習者や海外出張者、観光客、ビジネスパーソン、語学教育機関など多岐にわたる。本システムの普及のために、スマートフォンアプリの継続的なアップデートに加え、PCアプリ版の開発を行い、オンライン通話でも利用できるようにする予定である。また、SNSを活用したプロモーションや、学校や企業との連携を通じて、より多くのユーザーに利用されることを目指している。

## 7. クリエータ名（所属）

青原 光（名古屋工業大学 大学院工学研究科 工学専攻 メディア情報プログラム）

伊藤 朝陽（名古屋工業大学 大学院工学研究科 工学専攻 ネットワークプログラム）

## （参考）関連URL

PredicTalkのApp StoreダウンロードURL

<https://apps.apple.com/jp/app/predictalk/id6476206495>