

採 択 番 号 : 18001

課 題 名 : 自治体向け音声翻訳システムに関する研究開発

副 題 : 自治体窓口業務に対応した音声翻訳システムとサービス化に向けた研究

(1) 研究開発の目的

本研究開発は、NICTが開発した音声翻訳システムをベースに自治体向けの音声翻訳システムの基盤技術の確立を行うため、実施するものである。NICTが開発した音声翻訳システムは、対訳コーパスから翻訳システムに必要な統計的な知識を自動的に構築する技術であり、音声認識／翻訳できる語彙の不足等は音声翻訳システムの性能に直接関わる。現在、NICTでは、訪日・在留外国人を対象にした生活会話等のコーパス整備を進めており、実用場面・局面毎のコーパスが拡張されているところである。

本研究開発では、音声翻訳システムの導入により窓口業務が大きく改善が図られるものと、音声翻訳システムの導入した場合においても改善が図られないものを把握し、自治体窓口業務の構造の把握と体系化を行う。コーパスデータの整備では、各種窓口等における生活会話文等を参考に対訳コーパスの基礎整理を図るとともに、音声コーパスでは、シナリオ策定に向けたシーンの設定等を行う。これらの成果を踏まえ、自治体窓口向け音声翻訳システムの開発と実証実験を行い、翻訳精度とユーザビリティの向上を図る。あわせて、研究開発成果の社会展開に向けて、サービス開発にかかる基盤の確立を行う。

これらから、本研究開発では、NICTで整備されたコーパス情報等も積極的に活用し、自治体窓口における音声翻訳システムの精度向上と研究開発の効率化を図る。

(2) 研究開発期間

平成 27 年度から平成 31 度 (5 年間)

(3) 実施機関

凸版印刷株式会社<代表研究者>
株式会社フィート

(4) 研究開発予算 (契約額)

総額 749 百万円 (平成 30 年度 150 百万円)
※百万円未満切り上げ

(5) 研究開発項目と担当

研究開発項目 1 : 自治体窓口業務に関する研究 (凸版株式会社)

研究開発項目 2 : コーパスデータの収集に関する研究

1. 音声・対訳コーパス収集研究 (凸版株式会社)
2. クラウドソーシングによるコーパス収集研究 (凸版株式会社)

研究開発項目 3 : 音声翻訳システムに関する研究

1. 音声翻訳システム研究 (株式会社フィート)
2. ユーザーインターフェース開発 (凸版株式会社)

研究開発項目 4 : 音声翻訳システムの実証実験及び検証

1. 音声翻訳システムの実証実験 (株式会社フィート)
2. 音声翻訳システムの評価 (凸版株式会社)

研究開発項目 5：ビジネスモデル等統括研究（凸版株式会社）

(6) 特許出願、論文発表等

		累計（件）	当該年度（件）
特許出願	国内出願	0	0
	外国出願	0	0
外部発表	研究論文	0	0
	その他研究発表	2	2
	プレスリリース・報道	7	4
	展示会	17	6
	標準化提案	0	0

(7) 具体的な実施内容と成果

研究開発項目 1：自治体窓口業務に関する研究

本研究開発項目は、平成 27、28 年度に自治体窓口の業務分析を実施し、外国人が自治体窓口に来訪した際の頻度の高い代表的な業務（住民登録・税金・年金、子育て）を抽出し、ロールプレイ等から窓口会話のサンプルの収集を行った。平成 29 年度以降に、本項目に関する研究の予定はない。

研究開発項目 2：コーパスデータの収集に関する研究

①目標

音声・対訳コーパス収集では、音声・対訳コーパス収集では、平成 28 年度に研究開発項目 1 で実施した自治体窓口業務に関する研究成果を踏まえ、平成 29 年度に 23 万文の日英、日越対訳コーパスを完成させた。平成 30 年度は簡体字（18 万文）、伯語（8 万文）を作成し、コーパスの多言語を進める。また、平成 30 年度の音声コーパスの収集は、NICT において、本研究開発で対象とする言語のデータの蓄積が進んでいることから実施せず、翻訳システムの多言語展開のトライアルとして、日韓 8 万文、日泰 8 万文、日尼 8 万文、日緬 8 万文の対訳コーパスの多言語化を行う。

また、クラウドソーシングでは、アマゾンメカニカルターク（AMT）で、スマートフォンや PC 等を利用した音声収集を引き続き実施し、音声データの品質や収集方法に関する評価を行う（音声収録目標は 20 名）。また、言語の拡張性を検討し、対象となる言語の検討および課題抽出を行う。

②実施内容

対訳コーパスでは、優先業務 B に関するコーパス（子育て・年金コーパス）18 万文を日簡にて作成した。また、優先業務 A に関するコーパス（住民登録・国保コーパス）計 8 万文の日伯コーパスを作成するとともに、翻訳エンジンの多言語化トライアルを目的として、韓・泰・尼・緬の言語に関してそれぞれ 8 万文（住民登録・国保コーパス）のコーパスを作成した。

また、クラウドソーシングについては、AMT（amazon メカニカルターク）での音声収集において、1）音声収録の歩留まり率、2）作業品質、3）音声収録タスクに係る費用等の把握を行った。

③成果

今年度の成果として、①対訳コーパスでは、研究項目 1 で実施した優先業務 B（国民年金、税金、子育て）の日簡対訳コーパス（180,000 文）、優先業務 A（住民登録・国民健康保険）の日伯対訳コーパス（80,000 文）を作成した。②翻訳エンジンの多言語化トライアルの目的にて優先業務 A（住民登録・国民健康保険）の日韓・日泰・日尼・日緬対訳コーパ

ス（80,000 文）を作成した。

クラウドソーシングについては、ベトナム語音声収録歩留まり率は 20%前後であり、ベトナム語のネイティブでない作業者も極めて多く見受けられた（マスタース資格を有する話者が少ない等も含む）。これらから、今回のタスク内容（クオリティコントロールの仕組みの考慮なし）では、1 時間の正しく発話されたベトナム語の音声時間を得るためには、概ね、170 万円程度が収録コストとして発生することが推計された。今回の収録結果から、海外クラウドソーシング、特にアメリカ合衆国を中心に展開されている Amazon Mechanical Turk（AMT）を利用した、ベトナム語音声の収集では、作業者への収録費用のみならず、品質保証のためにかなりの追加コストを支払う必要があることがわかった。従って、クラウドソーシングで音声収集を行う場合、大規模の録音でなければコストに見合わない可能性が高いと考えられる。特に、作業者の登録が少ない場合は、その影響が顕著になると思われる。収録したい音声データの規模に合わせて収録手法、例えば国内でのアウトソーシングやベトナムでのアウトソーシング等、品質を保証しやすい手法での実施の検討が必要と考える。

表 1 ベトナム語音声収録に係るタスク

タスク	内容
音声収録、聴音アプリMturk 上での作成作業費用	アプリ開発工数0.5人月程度
収録コスト	20人×3回のタスク ● Mturk作業費用(20人×3回タスク, US\$435(約49,000円)) ● 聴音判定作業費用(ベトナム人<ネイティブ>による判定, 150 円/1発話)
収録音声時間	7名、350ファイル、1,065秒

研究開発項目 3：音声翻訳システムに関する研究

①目標

自治体窓口コーパスを利用した翻訳モデルの構築と評価、及び対訳辞書の登録を行い、自治体窓口向け音声翻訳システムの改善を図る。このため、本項目では、実証実験用システムの開発を行う。

音声翻訳システム研究では、研究開発項目 2 で作成した自治体窓口業務を踏まえたコーパス、自治体用語辞書等から、平成 29 年度に構築した自治体窓口向け音声翻訳システムの改善を図る。実証実験用システム（H30 版音声翻訳システム及びアプリケーション）では、日英、日越に加え、研究開発項目 2 の対訳コーパスで作成した言語を対象として多言語化を図る。

また、自治体窓口向け業務支援アプリケーションのユーザーインターフェースの研究では、研究開発項目 4 において平成 29 年度に行ったユーザビリティ評価で抽出された課題である、専門用語・定型フレーズ機能のユーザーインターフェースの見直しを行う。また、より利便性の高いアプリケーションを目指すために、外付けマイクに代表される音声入力の外部インターフェースについて比較調査を行い、より最適な音声入力デバイスについて検討する。

②実施内容

前年度の研究成果を踏まえ、音声翻訳アプリケーションのユーザーインターフェースを設計し、アプリケーションとして実装した。

音声翻訳システム研究では、研究開発項目 2 で作成した対訳コーパスを利用して自治体窓口向け翻訳モデルを構築した。具体的には、NICT の開発による既存の旅行会話分野の翻訳モデルをベースとし、本研究開発で作成した自治体窓口分野の対訳コーパスを利用して H30 版音声翻訳システムで使用する翻訳モデルを構築した。H30 版音声翻訳システムの翻訳言語対は、日英、英日、日中、中日、日越、越日、日伯、伯日、日韓、韓日、日尼、尼日、日緬、緬日、日泰、及び、泰日の合計 16 言語対である。構築した音声翻訳システムを対象と

して自動評価、及び、主観評価を実施した。更に、研究項目 2 で作成した対訳辞書データを音声翻訳システムに登録した。

ユーザーインターフェースの研究では、専門用語・定型フレーズ機能の検索性について改善方法を検討し、目的の言葉を検索しやすいユーザーインターフェースを設計した。その後、設計したユーザーインターフェースを現状のアプリに対して実装した。さらに、昨年度までは 16 語であった登録単語を増量し、翻訳言語も英越以外に対応させた。また、3 種のマイクデバイスを対象とした比較調査を行い、現状の最適なマイクデバイスおよび今後のマイクデバイスに必要な要件の抽出を行った。

③成果

新たに構築した H30 版音声翻訳システムを対象として、翻訳性能の自動評価、及び、主観評価を実施した。翻訳言語対それぞれの評価結果は下記の通りである。

表 2 翻訳言語対の評価結果

翻訳モデル	自治体コーパス サイズ	自動評価	主観評価	
		BLEU	SAB	SABC
日英	200k ¹	28.8	68.0	81.0
英日	200k ²	36.3	78.0	86.0
日中	230k	40.0	91.0	98.0
中日	230k	50.9	95.0	100.0
日越	230k	38.2	88.0	97.0
越日	230k	39.9	75.0	81.0
日伯	80k	27.6	58.0	74.0
伯日	80k	35.6	66.0	77.0
日韓	80k	58.6	96.0	100.0
韓日	80k	63.8	95.0	99.0
日尼	80k	30.5	94.0	94.0
尼日	80k	39.3	66.0	72.0
日緬	80k	33.2	87.0	93.0
緬日	80k	45.4	79.0	84.0
日泰	80k	30.4	90.0	94.0
泰日	80k	35.5	88.0	94.0

自動評価は、正解訳と機械翻訳システムが出力した訳との類似度を 0 から 100 の間の数値で表す BLEU 値により実施した。BLEU 値は、経験的に概ね 30 以上が目標とされるが、いずれの言語対についても、ほぼ目標値に到達していることが確認できた。

一方、主観評価は、以下の表の基準を用いた SABCD の 5 段階評価における SABC の文数の比率で表される。同様に SABC の評価結果が 80%以上となることが目標値とされる

¹ 評価時のもの。現在のコーパスサイズは 230k となっている。

² 評価時のもの。現在のコーパスサイズは 230k となっている。

が、日伯、伯日、及び、尼日以外の言語対については目標値に到達していることが確認できた。目標値に到達していない三つの言語対については、いずれも自治体コーパスサイズが80k と少ないことが、主観評価結果が目標値に到達していない要因のひとつと推測される。これらの言語対については、次年度においてコーパスサイズを増量し、改めて翻訳モデルの構築、及び、評価を実施することとする。

表 3 主観評価 SABCD

S) ネイティブ並み	原文の情報が漏れ無く翻訳されており、訳出に文法的な間違いがない。使われている語彙もネイティブから見て自然である。
A) 申し分ない	使われている語彙はネイティブから見て不自然であるが、原文の情報が漏れ無く翻訳されており、訳出に文法的な間違いがない。
B) まずまず	原文のあまり重要でない情報が一部漏れていたり、または間違っていたり、あるいは訳出に文法的な間違いが若干あるが、容易に理解できる。
C) 許容範囲	原文の重要な情報が漏れていたり、訳出に文法的な間違いがだいぶあって、かなり崩れた訳出であるが、よく考えれば理解できる。
D) 意味不明	重要な情報が誤訳されており、理解不能である。

ユーザーインターフェースの開発では、検索性の課題を改善したユーザーインターフェースを具備する専門用語・定型フレーズ機能を設計、実装した。具体的には、様々な自治体のウェブサイトを開覧し、自治体が役所の手続きをどのように分類しているかを参考に設計した。これによって、ユーザーである職員が用語を検索する際に、その用語がどのカテゴリにあるかを想起しやすいユーザーインターフェースを開発することができた。また、分類後の階層構造においても、カテゴリ検索（大分類）＞業務や手続き（中分類）＞単語といった3階層に分けることで、目的の言葉まで迷わず到達できるように工夫を行った。

さらに、昨年度までは16語であった登録済単語を79語まで増やした。それらは英・越・中・韓・ブラジルポルトガル語に対応できるように翻訳を行った。これによって、次年度の実証利用において、本機能をより多くの窓口で利用可能となった。

マイクデバイスの比較調査では iPad に実装されたマイクデバイスが最も欠点が少ないとの結論となった。他の対象デバイスであるカード型ハンズフリーマイクと外付けスタンドマイクについては長所があるものの欠点となる点も散見され、総合的には iPad のマイクが最もよいと結論付けた。



図 1 改善後の専門用語・定型フレーズ機能の画面イメージ

研究開発項目 4：音声翻訳システムの実証実験及び検証

①目標

研究開発項目3で開発した音声翻訳システムの実証・検証として、平成30年度は実際の窓口業務での利用する実証利用（複数自治体）による評価を実施する。

これまでの実証実験においては、会議室等に模擬的な窓口環境を用意し自治体職員と外国人のモニターを被験者としていたが、より実際の利用シーンに近い環境での評価を行うべく、平成30年度は実際の窓口業務で本システムを利用する実証利用という形式を採用することとする。具体的な手法としては、項目5にて接点のある自治体へ実証利用参加の募集を行い、協力可能な自治体に対して一定期間本システムを利用してもらう。協力自治体については1～10自治体程度で行い、音声翻訳システムの評価については、実証利用協力自治体に対して利用現場を観察する調査やインタビュー等を行い、実利用の中での本システムの有用性や課題について明らかにする。

この他に、これまでの自治体へのヒアリングにおいては、通訳者がいないと外国人とのコミュニケーションがうまくできないとの声が多いにも関わらず、一方で日本語がわからない外国人が一人で来庁するケースは多くないとの声が聞かれたため、日本が得意ではない外国人が単独で来庁するケースの理由や背景等の実態について調査を行い、今後の本システムの利用シーンへの影響を探る。

②実施内容

音声翻訳システムの実証利用については、甲府市、綾瀬市、前橋市、新潟市西区、広島市、岩国市、岐阜市、枚方市、袋井市の9自治体の協力を得て実施することができた。

評価については、前橋市、甲府市、新潟市西区の3自治体の協力を得て、実際の利用現場を観察することができた。その他の自治体についてはヒアリング形式でフィードバックを得た。

また、単独で来庁する外国人の実態調査（ウェブアンケート）を実施し、6つの言語圏に対し、合計618名へのウェブアンケートを通して回答を得た。言語圏は英語、中国語、ブラジルポルトガル語、ベトナム語、タガログ語、ネパール語の6言語圏である。

③成果

実証利用においては、実業務にて利用してもらうことで新たな気づきや課題を得ることができた。具体的には、本システムのユーザーインターフェース（操作性）においては実利用においても大きな課題は見られないことがわかった。その一方で、窓口業務におけるアプリ利用の実態観察を通して、これまでの模擬環境で行った実証実験ではみられなかった「利用前」や「継続利用」段階における様々な課題が明らかとなった。例えば、利用前の本システムへの期待値がユーザーによって異なるため、実際の利用時のインタラクションで上手く翻訳できなかった場合に利用を止めてしまうケースや、システムを介した会話によって時間がかかることに不満を感じて利用を止めてしまうケースが見られた。他方、職員からの一方的な説明による従前の手続きスタイルに対する問題意識が強い職員は、誤翻訳や手続き時間というマイナス面よりもコミュニケーションの深さを利用価値として捉え、継続的な利用を行っていた。

これらの結果を踏まえて、本システムを実業務においてより活用してもらうためには、翻訳精度や操作性だけではなく、利用前のレクチャーやアプリを介した会話体験、利用による成功体験の共有といったことを行っていく必要があることがわかった。

また、日本語がわからない外国人が単独で来庁するケースのWEB調査においては、これまでの想定を超えるような意外性のある回答は見られなかった。来庁理由は住民登録や国民健康保険といった手続きが多く、単独来庁の理由は日本語がわかる知人等の同席の都合がつかなかったからといった回答が多かった。

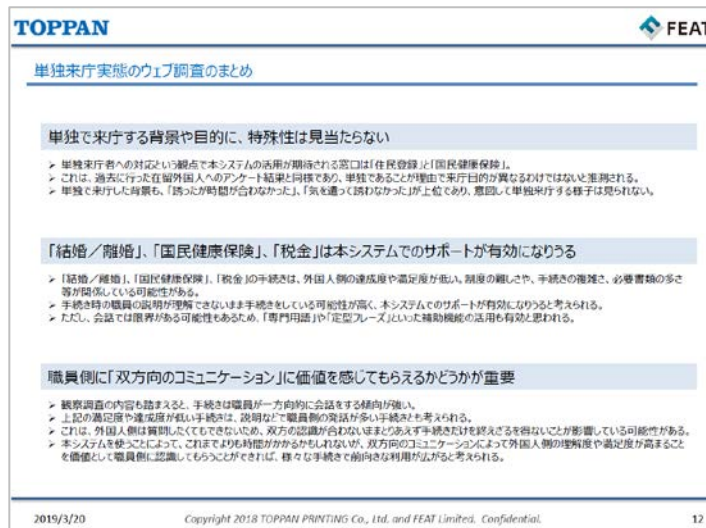


図2 単独来庁実態の調査結果

研究開発項目5：ビジネスモデル等統括研究

①目標

平成30年度は、自治体音声翻訳システムのあり方を、①現行モデルと、②総合行政ネットワーク連携モデル、の2つのビジネスモデルがあると考え、これらの社会実装に向けた調査・活動を実施する。具体的には、音声翻訳システムの導入・社会実装化調査（ヒアリング調査）を実施する。さらに、昨年度、ベトナムで海外市場への応用展開の可能性についての調査・検証を行ったが、平成30年度も引き続き、海外調査を実施し、自治体音声翻訳システムの成果をベースに、訪日予定のベトナム人向けに自治体音声翻訳システムのアプリを提供し、訪日外国人の送り出し機関を対象に、当該システムの利活用状況に関する調査（アンケート調査等）を実施する。

②実施内容

本研究開発の「アプリ型」モデルの社会実装化に向けては、実証実験の自治体への勧奨活動を通じて、2019年度以降の社会実装を図るべく、各自治体へのヒアリング（普及啓発、予算動向等の把握を含む）を実施した。また、音声翻訳システムの窓口業務の応用展開として、他の公的機関（教育委員会、消防署等）への探索活動も引き続き実施した。本研究開発の推進に向けては、展示会の出展、セミナーの実施、ホームページによる情報発信を実施した。

また、総合行政ネットワークと連携した運用モデルの可能性を把握するため、自治体業務の効率化・コスト削減の可能性、個人情報に対するセキュリティの確保等について検討を行った（行政サービスと関連した新たな付加価値の可能性等を含む）。これらの把握に当たっては、実証実験等で深く関与している自治体や総合窓口化・行政改革等で実績のある自治体、防災・介護・教育等で特筆する活動を行っている自治体等を対象に深堀調査を実施した。

日本語ーベトナム語音声翻訳システムに関連した取組みでは、ベトナム行政当局を中心にヒアリング調査を実施した。また、日本語ーベトナム語音声翻訳アプリでは、訪日前の日本語学習や訪日後の生活全般に活用を想定、翻訳精度向上を図り、アプリケーションの利用状況等についての検証を行った（利用者アンケートを含む）。当該アプリの普及啓発として、ベトナムの日本語研修センター等での広報活動、ベトナム人コミュニティサイトへの告知掲載、現地展示会（ハノイ市）や日本留学生フェア等を活用し認知度向上を図った。

③成果

「アプリ型」モデルの社会実装化に向けて、各自治体向け実証実験の勧奨活動を実施した。勧奨活動は、実証実験実施対応と当社エリア事業部と連携し、外国人集住自治体へ「地域特性」を活かした実証・普及体制構築を図った。他の公的機関での利活用では、横浜市、浜松市の実証実験の支援を行った。「総合行政ネットワーク連携モデル」では、袋井市、草津市、高浜市、越前市等の情報部門へヒアリングを実施した。併せて、政策動向等を把握するため、公開情報調査を実施した。

日本語ーベトナム語音声翻訳システム (NhaTra) では、日本政府観光局 (JNTO) ハノイ事務所の Facebook、ベトナム人向け日本情報サイト「wappuri」に掲載され、利用促進 (約 8000 ダウンロード) に寄与したほか、訪日前ベトナム人を対象にアンケート調査を実施し、コミュニケーション課題や音声翻訳への期待等を把握した。