

6. 【言語の壁を乗り越える音声コミュニケーション技術の実現】

(1) 概要

国際化の進展の中で諸外国の相互理解の増進の必要性が益々重要になってきている今日、特にアジア圏の近隣諸国の人々と会話による直接的なコミュニケーションを図り、国民レベルでの相互理解を深めることが必要とされている。

この課題を解決していくため、言語の壁を越えて、アジア圏等の海外の人々と直接会話による交流を可能とすることのできる自動音声翻訳システムに関し、当面の利用ニーズと今後5年程度で期待できる技術向上レベル等に考慮して、海外旅行、外国人向け観光・ショッピング、国際交流イベント等の分野における音声翻訳システムの実証を企画・推進し、プロジェクト終了後短期間で民間ベースでのサービスにつながるよう、その成果の社会還元を加速を目指す。

【5年後】普通の旅行者が、日本、英語、中国語圏でほとんど支障なく海外旅行を楽しめる。

【10年後】多様な言語で、ビジネス等を含む、多様なコミュニケーションを実現。

(2) 海外の動向

コーパスベース言語処理や音声翻訳などの日本の技術は世界最先端であり、その性能は国際的に高い水準にある。一方、海外において大型プロジェクトがスタートすることから、世界に先んじてブレークスルーにつながる技術の開発を加速・推進する必要がある。以下、海外のプロジェクトについて簡潔に記す。

1. GALE (米)

国防総省の新 大規模プロジェクトであり、音声、ドキュメントを包括した分野に依らない音声言語翻訳、要約、コミュニケーションより情報収集を目的とする。

予算： 46 億円/年 (推定)。期間： 2006-

2. TC-Star (EU)

EU の FP6 プロジェクトであり、会議、講演の音声翻訳を行う。

予算：20 億円/年 × 3 年。期間：2004-2006

3. 言語情報処理プロジェクト（韓国）

MIC(韓国情報通信省)により、ETRI 主体に進められているプロジェクトであり、音声翻訳、・情報獲得、音声インターフェースに関する研究開発を行う。

予算： USD 7 million/2005。期間： 2003 - 2005 (3 年間)

(3) プロジェクトの目標

①音声翻訳コミュニケーションプロジェクトにより目指すべき姿

世界に開かれた社会の実現を目指し、IT、人文科学の融合によって、言語の壁を乗り越える音声コミュニケーション技術の研究開発および実証を行う。

(i) 10年後の社会

1) 『国民誰でも必要に応じいつでも自動音声翻訳サービスを利用できる社会』の実現

- ・音声翻訳コムにより、国際観光交流の一層の拡大に加え、これまで言語の壁により国際展開が進まなかった草の根レベルでの国際経済活動、文化交流活動の拡大・活性化
- ・医療・教育等を含む公共サービス機関等での在留外国人支援活動の充実、強化
- ・また、専門（特殊）用語の充実等に伴い、例えば、近年特に商品開発などで企業間の関係強化が著しい中国など英語圏以外の国々の企業との間でのオフショア開発など、国際的ビジネス拡大のためのツールとして活用
- ・期待されるアウトカム
 - a 国際観光産業の振興（外国人旅行者の受入れ拡大）
 - b 海外旅行者、旅行先の拡大
 - c 海外取引、オフショア開発、外国人向け商品販売等国际ビジネスの拡大
 - d 海外移住、就労、留学範囲の拡大
 - e 外国人労働者の受入れ拡大
 - f 文化紹介イベント等国际相互文化交流活動の拡大
 - g 海外、または、外国人と共同での様々ボランティア活動の展開
 - h 病気、怪我や非常災害等で被災した外国人救助対応の充実
 - i 初等中等教育機関での外国移住者の学習活動支援の充実、等

2) 音声翻訳システムを使った新ビジネスの創出

(A) オプションサービスが基本

- ・自動音声翻訳サービスは、国民生活向上、地域産業の振興等に重要な役割。
- ・一方、常時継続的ニーズは、国際観光産業、外国人就労者、留学生等特定者を除けば限定的。また、必要性も語学能力等個人資質によ

り大きく異なる。

→基本的には、利用者個々人のニーズに応じて多様なサービス形態で提供されるオプションサービス

(B) 民間事業者により、創意工夫を活かした競争的サービス展開

- ・競争的個々人のニーズを発掘・具現化する形で、的確迅速により良質なサービスを提供し、利用者拡大を図ることが普及発展に不可欠
- 事業インセンティブの働く民間ビジネスベースの事業展開が基本

(C) 現時点で想定されるサービス形態

IT 機器ベンダー、アプリケーション提供事業者、通信事業者等によるサービス展開が期待。

- ・家電ショップ等で、会議通訳等対応可能な高機能機器から語彙翻訳レベルの電子辞書型機器まで、多様なレベルのシステム（専用端末またはソフトウェア）の販売、リース
- ・携帯電話等通信事業者や、通信インフラを活用したアプリケーション提供事業者（ASP、SaaS 事業者）等が提供する音声通訳アプリケーション提供サービス
- ・海外旅行者等が空港等で、簡単に翻訳システム（端末）のレンタルサービス、等

3) 地域発展に役立つ全国サービス基盤の構築

以下のような自動音声翻訳サービス基盤を活用することにより、地域ぐるみの国際化による地場産業等の振興につなげる。

(A) 新たな情報格差を生み出すことのないサービス展開

- ・地域格差のないサービス環境
- ・年齢、ICT 利用経験による格差の生じないサービス提供

→全国的に整備される通信インフラ整備との連携によるサービス展開

(B) 地域の情報、文化の多様性にも配慮したサービス内容

→地域社会・文化に密着した翻訳情報データベースの構築・提供

(ii) 自動音声翻訳システムの姿

1) 誰でも簡単に使える（容易性、携帯性・移動性、安定性）

(A) ICT の幅広い普及は、使い勝手が決定的要因

- ・新しい ICT 普及の成否は、多くの場合、その機能の不足よりも、使い勝手の悪さが決定的な阻害要因になる場合が多い。

- ・特に「言語の壁」の問題に関しては、これまで外国人とのコミュニケーションに馴染みの薄い層への浸透が重要な鍵となるが、このような利用層は、ICTの利用に通じていない子供や高齢者といった人達も多い。
- これらの層に対しても、簡単なマニュアルと短時間の準備作業等により利用方法を習得できる簡便性や使い勝手の良さが不可欠
- (B) 屋外や移動中での利用が不可欠
 - ・初期の段階では、例えば特定の観光施設や商業施等における案内や応対といった半固定的な利用が中心となると想定。
 - ・一方、一般個人レベルへの普及段階では、海外旅行への対応や災害現場での利用等、常時携帯できること、また、車等での移動中にも活用できることが必須。
- (C) 常時携帯しての利用シーンにも耐えられる堅牢性
 - ・ICTに馴染みのない人々による利用、常時携帯利用といった点から、一般の家電商品以上に安定性・堅牢性も重要。

2) 多様な言語に (英語以外の言語によるコミュニケーション)

- (A) アジア諸国の人々との連携強化の必要性
 - ・国際化の進展の中で、国民の英語会話能力もかなり向上しつつあり、英語が共通語として浸透している欧米先進諸国の人々との間では、一般的かつ定型的な旅行会話やショッピング会話などは機械翻訳等の支援をあまり要しない状況になってきている。
 - ・一方で、邦人の海外旅行や移住・留学先の多様化、様々な国々からの旅行者、留学者、就労者の拡大などの変化は、英語圏以外の国々の人々との交流支援手段に対するニーズを高めている。
 - ・とりわけ、我が国は、ロシアを含むアジア諸国と社会経済面で幅広く地域的關係強化を進めてきているところ、草の根レベルでの相互理解の増進や経済關係の強化が重要な課題と
 - 英語だけでなく、中国語、韓国語、インドネシア語、タイ語、ベトナム語、ロシア語といった日本で馴染みの薄い近隣諸国の言語にも対応できるものであることが必要。
- (B) 日系外国人との關係拡大への対応
 - ・また、日系外国人を中心とした南米諸国の人々との社会経済面での交流も急速に拡大
 - スペイン語やポルトガル語への対応も重要。

3) 個人レベルの多様な用途・ニーズに応える

- (A) 個人の会話能力、利用場面ごとに対応できる多様性・拡張性が必要
- ・外国人との会話に対するニーズは、利用者ごとに異なる。また、それをサポートする自動音声翻訳システムに対するニーズは、個人の会話能力等に大きく依存する。
 - ・初めからユニバーサルシステムとして機能の画一化を図る必要性は少ない（機能の画一化は、汎用であるがどの用途にも中途半端な有用度になりやすい。）
- 利用者のニーズや利用場面に合わせて柔軟に機能を変化（拡張）できることが重要。
- (B) 利用者個人の特徴に合わせた機能拡張が必要
- ・当初は、旅行、ショッピング、観光案内等の利用から、教育、就労での利用、災害対策、医療活動での利用といった利用目的や場面ごとの専門化（チューニング）機能が重要
 - ・会話の内容が高度になるにつれ、場面重視の定型翻訳から、利用者の個性による属人性の高い翻訳が求められる。
- 利用者個人の理解（翻訳）能力、活用履歴なども考慮した語彙や翻訳情報の追加修正、表現方法など、利用者の多様性に応えられる発展機能を有することも重要になる。

②音声翻訳コミュニケーションプロジェクト終了時（5年後）までの目標

(i) 自動音声翻訳システム機能の検証

1) 翻訳性能

最終実証研究終了時点で、実利用に供し得るものとして、以下の性能を有することを検証。

- (A) 外国語による応対・サービスがほとんど行えていなかったホテル、旅館、物産店等において外国人客を相手に、また、外国語がほとんど話せない邦人が、海外旅行先の交通機関やホテル、店舗等の現地スタッフを相手に、10語（ワード）程度の会話による応対がほぼ問題なく行える（利用者より、ほぼ実用に供することができるとの評価が得られる）。
- (B) 世界遺産等の海外でも関心の高い観光スポットにおいて、当該地域に関係する固有名詞や社会事象を含む多様な語彙を備えた案内や質問応答をほぼ問題なく行える（利用者より、ほぼ実用に供することができる）。

できるとの評価が得られる)。

- (C) 発話者の文章理解が不完全な場合でも、場所、利用状況や他の前後の文脈からの情報を加味した翻訳を可能とする。
- (D) 対応言語については、英語に加え、中国語で日英翻訳同等の機能を確保できる。韓国語については当該技術の適用可能性について検討する。

2) 使い勝手

- (A) 小学生高学年レベルの ICT 熟練度でも、簡単なマニュアル（ビデオ等）により、短時間で使用方法を習得できる。
- (B) 観光スポットや駅、空港等の程度の雑踏においても概ね利用できる。
- (C) 簡単なチューニングプロセスにより、通常の会話で問題とならない程度の発声上の差異を吸収することができる。
- (D) 携帯電話端末による全国的なサービス利用も可能とする。
- (E) バス等車両移動中でも利用できる。
- (F) 通信インフラがない場所でも、端末だけで一般の旅行会話程度の翻訳が可能とする
- (G) 個人利用向けサービス用端末として、携帯電話、PDC 端末のような携帯性、堅牢性のある端末での利用を可能とする。

3) 高機能化・多様化に向けた拡張可能性

- (A) 自動学習機能により、翻訳知識や語彙、文例の自動拡大を可能とする。
- (B) 将来的（10年後頃まで）には、言語種類10言語程度、また、語彙についても100万程度の対応が可能となるシステム機能拡張を可能とする。

(ii) 社会普及可能性（社会的有用性）の検証

1) 社会生活面、産業活動面への寄与の観点で期待しうる効果の予測

- ・実証実験を通じて観光・旅行産業の振興等の面での影響について分析し、将来（10年後）の本システムの普及浸透による効果についてとりまとめる。

2) 商用サービス化に向けた制度的課題の整理と解決方策の検討

- ・実用サービスに近い形での実証を通じ、研究開発フェーズから商用サービスフェーズに移行するに当たって、新たに障害となる制度的課題がな

いか検証し、障害等があれば、その解決方策についてとりまとめる。

3) 自動音声翻訳システムの有効性に関する社会認知度の向上

- ・ 実証実験及びその成果発表（マスコミへの公開）等を通じて、特に、国際観光業や海外旅行用等の用途において、自動音声翻訳技術が概ね実用レベルにあること、また、今後のビジネス展開や国際交流拡大の可能性が大きいこと等をアピールする。

(4) プロジェクトにおける現状と課題

①技術開発の現状と課題

幅広い話題へ対応するため、ネットワーク上に分散する翻訳知識を活用し、翻訳端末と組み合わせるネットワークベース音声翻訳技術の確立が必要。

(i) 音声技術

- ・ 端末内に内蔵された汎用の音声認識、合成知識に加え、話者、環境に応じて自動的に適応学習し、頑健な音声認識、状況に応じた音声合成を可能とする。音声翻訳ネットワーク上の多数の話者の大量の音声を用い、信頼度を考慮して音声認識、合成知識を学習することにより、「使うほど賢くなる」音声認識、合成を可能とする。
- ・ 未知語を検出し、Webを利用して発音、アクセント、表記を付与し音声認識、音声合成を可能にする。
- ・ 場所、状況に合わせて、音声認識、音声合成知識を適応的に切り替えることにより発音の曖昧性解消、語彙の絞り込みによる性能改善を実現する。

(ii) 翻訳技術

- ・ 端末内に内蔵された汎用の音声翻訳知識に加え、音声翻訳ネットワークからダウンロード可能な専門の音声翻訳知識、翻訳者や集合知との連携機能を開発・整備する。これにより、地名などの固有名詞や専門用語を有する文章も翻訳を可能とする。
- ・ 端末内に内蔵された汎用の音声翻訳知識に加え、音声翻訳ネットワークからダウンロード可能な専門の音声翻訳知識、翻訳者や集合知との連携機能を開発・整備する。これにより、地名などの固有名詞や専門用語を有する文章も翻訳を可能とする。
- ・ 場所、状況に合わせて、翻訳知識を適応的に切り替えることにより適切に翻訳することを可能とし翻訳精度の向上を図る。
- ・ Webから自動的に翻訳知識を学習する手法、音声翻訳の信頼度を開発し、信頼度に基づいた「使うほど賢くなる」機構を確立する。
- ・ 概念階層辞書を構築・活用し固有名詞等の未知語がある場合でも適切に翻訳することを可能とし翻訳精度の向上を図る。

(iii) 標準化

- ・ 翻訳ネットワークへのインターフェースを標準化することで、様々な専門知識を取得可能にする。

②システム改革課題（実用化に向けた課題）

- ・自動翻訳システムの実用化に向け、既に民間レベルでの動きもあるところであり、本システムの普及に重大な障害となるシステム改革要因（社会制約要件）は見あたらない。
- ・しかしながら、本システムの円滑な実現・普及を加速する観点から、以下のような課題に取り組んでいくことが重要。

（i）音声翻訳コミュニケーション技術によるイノベーションの「見える化」

音声翻訳については、会話によるコミュニケーションに活用するものであるが、一般的な会話コミュニケーションは、記述コミュニケーションと異なり、その行為が無意識に近い状況で自然に行えなければ返ってコミュニケーションに負担を感じるもの。

その一方で、会話によるコミュニケーションは話者の個性に強く根ざすもので定型化しにくいものあるのに加え、日本独特の文化に根ざした言語の複雑さや多様性などにも対応することが重要な要素となることから、「機械化はかなり困難」との認識が定着。

この結果、当該システム利用に関して、「是非このような用途に。」といった具体的ニーズが明確にならず、このため、一般の人々からは、「とりあえずいろいろな用途に簡単に利用できるようなものであれば」といった漠然として、返って技術的には厳しい要求がつけつけられる。

このような状況から、開発サイドにあっては、技術的にはかなりの高度なシステムを開発しても、通常会話のような「自然さ」を評価基準とする利用者側のニーズに応えられず、そのミスマッチが、実用システム開発、ビジネス展開といった動きに繋がらないのが現状。

このデッドロックを打破するためには、技術レベルに合わせて敢えて利用者のニーズ（用途）を十分に絞込み、その中で当該システムの有効性を実証してみせる（体感してもらう）、いわゆる「見える化」を進めることが必要。

また、「見える化」にあたっては、利用者側の意図に沿って研究開発した技術性能を誇示するのではなく、利用者にとっての有効性が判断できるよう、事前に想定される利用者側と協議し、目的に合わせて具体的な利用場面、利用方法、利用範囲、語彙、精度等を利用環境条件などを敢えて厳選し、それに適した機能とユーザーインターフェースを持ったシステムを開発して実証することが重要である。

（ii）高速無線通信インフラ

自動音声翻訳技術を国民の生活のツールとして長期間にわたって受入れられるようにするためには、必要な時必要な場所でそれぞれの必要度に応じ適切なサービスが受けられること、また、その一方で地域にも根ざした会話ができるようにすることが不可欠。

このため、これに必要な全国的な高速無線通信インフラが早期に整備されることが重要なポイントである。また、高速無線通信インフラに関しては、拡張型第3世代携帯電話（HSDPA、Super3G等）、第4世代携帯電話等、次々と新しいサービスが開発・導入されてくることから、今後、整備されるこれらの高速無線インフラを、状況に応じて最大限活用できるようにしていくことも重要なポイントとなる。

(iii) 翻訳コーパス構築等に係る著作権法上の課題の整理

常に性能の向上を追求する利用者に応じて、社会に普及定着させるためには、機械的な機能充実以上に、複雑・多様に変化する会話にも適切に応じられるも向上性が重要。

これには、対訳例コーパスの機能的・効率的な充実化が鍵。

対訳例コーパスの構築に伴うサーバーへの複製・加工等や検索結果の表示については著作権法上未整理の部分があり、現在この点については、文化審議会著作権分科会で議論されており、近々結論がまとめられることとなっているところ、この検討結果を待って、将来の商用サービスに向けて制度面での課題を整理し、今後新たな制度上の対応が必要となれば、別途対応策について検討を行うこととする。

(5) プロジェクトの課題解決に向けた取組み

①技術開発面での取組み

(i) 音声技術^{※1} 【総務省、経済産業省】

- ・ 2006～2008 年度 ; ユーザビリティ向上のための音声認識技術の研究開発^{※2}
- ・ 2008 年度 ; 話者、環境に適応する音声認識技術の研究開発
- ・ 2008～2011 年度 ; 場所・状況に応じて適応的に辞書などを切り替える音声認識、音声合成の研究開発
- ・ 2009～2011 年度 ; 音声認識、音声合成知識の信頼度を用いた自動学習技術の研究開発
- ・ 2012 年度 ; 自動学習技術、場所・状況に適応する音声認識、音声合成技術の改良

(ii) 翻訳技術^{※1} 【総務省】

- ・ 2008～2009 年度 ; 旅行・観光知識の集合知的学習技術の研究開発
- ・ 2010～2011 年度 ; 声翻訳知識の信頼度準拠学習技術の研究開発
- ・ 2008～2010 年度 ; 場所・状況に応じて適応的に知識を切り替える翻訳技術の研究開発
- ・ 2012 年度 ; 集合知的学習技術、場所・状況に適応する翻訳技術の改良

(iii) 技術性能確認 (Ⅲ-2 (1) の実証と併せて実施) ^{※3} 【総務省】

- ・ 2009 年度 ; 実証用プロトタイプ制作 (小規模、上海万博など向け)
- ・ 2011 年度 ; 実証用プロトタイプ制作 (大規模、主要観光地向け)

(iv) 標準化 【総務省】

- ・ 2008～2013 年度 ; ネットワークベース音声翻訳技術の進展に併せて標準化活動を実施

②システム改革課題 (実用化に向けた課題) 面での取組み

(i) 音声翻訳コミュニケーション技術によるイノベーションの「見える化」^{※1}

- ・ 2010 年度 ; 小規模実証実験を実施。上海万国博覧会開催期間中、日本の公式参加事業との連携の下、自動音声翻訳端末を使った日本紹介イベントの実施を検討。

- ⇒ 日本館担当部局等の協力を得て、日本館敷地等での実証イベント案を検討（2008～）
- ⇒ 実証イベント設計、モニターの公募（2009～2010）
- ⇒ 上記検討のため、TFの下に支援班を設置（2008～2010）

【総務省、経済産業省、国土交通省】

- ・ 2012 年度 ; 外国人観光客誘致拡大を目的として総合的な観光振興施策を実施または計画している観光名所地区等の自治体、商工会等業界団体、企業等より提案を公募し、その中からモデル地区を設定し、当該地区において、官民協力しつつ外国人旅行者を対象とした実証実験を実施。

- ⇒ 実証イベント公募（具体的活用方法等提案）、モデル地区の指定（2010）
- ⇒ 上記検討のため、TFの下に利用サイドの専門家を含めた支援班を設置（2010～2011）

【内閣府、総務省、国土交通省、他観光関連産業関係省庁】

(ii) 高速無線通信インフラ【内閣府、総務省】

- ・ 2010～2012 年度 ; 音声翻訳コミュニケーションシステムの活用に必要な第3世代携帯電話システム等高速無線インフラに関し、5年乃至10年程度における整備見通しを調査し、特に、音声翻訳コミュニケーションシステムのパイロット的展開に重要でありながら、必要なインフラ整備の遅れにより計画に影響が生じるような地域やエリアがあれば、当該地域における早期インフラ整備等に関する要望をまとめ、関係機関に対して早期整備を促す。

(iii) 翻訳コーパス構築等に係る著作権法上の課題の整理

- ・ 対訳例コーパスの構築に必要なクローリングによるデータ収集、複製、編集、検索、表示活用等に関する内外の制度について整理。
- ・ 関係府省の協力を得て、著作権問題関係専門家等を含めた課題検討会を設置し、システム実用化に向けて必要となる制度の在り方について提言をとりまとめる。

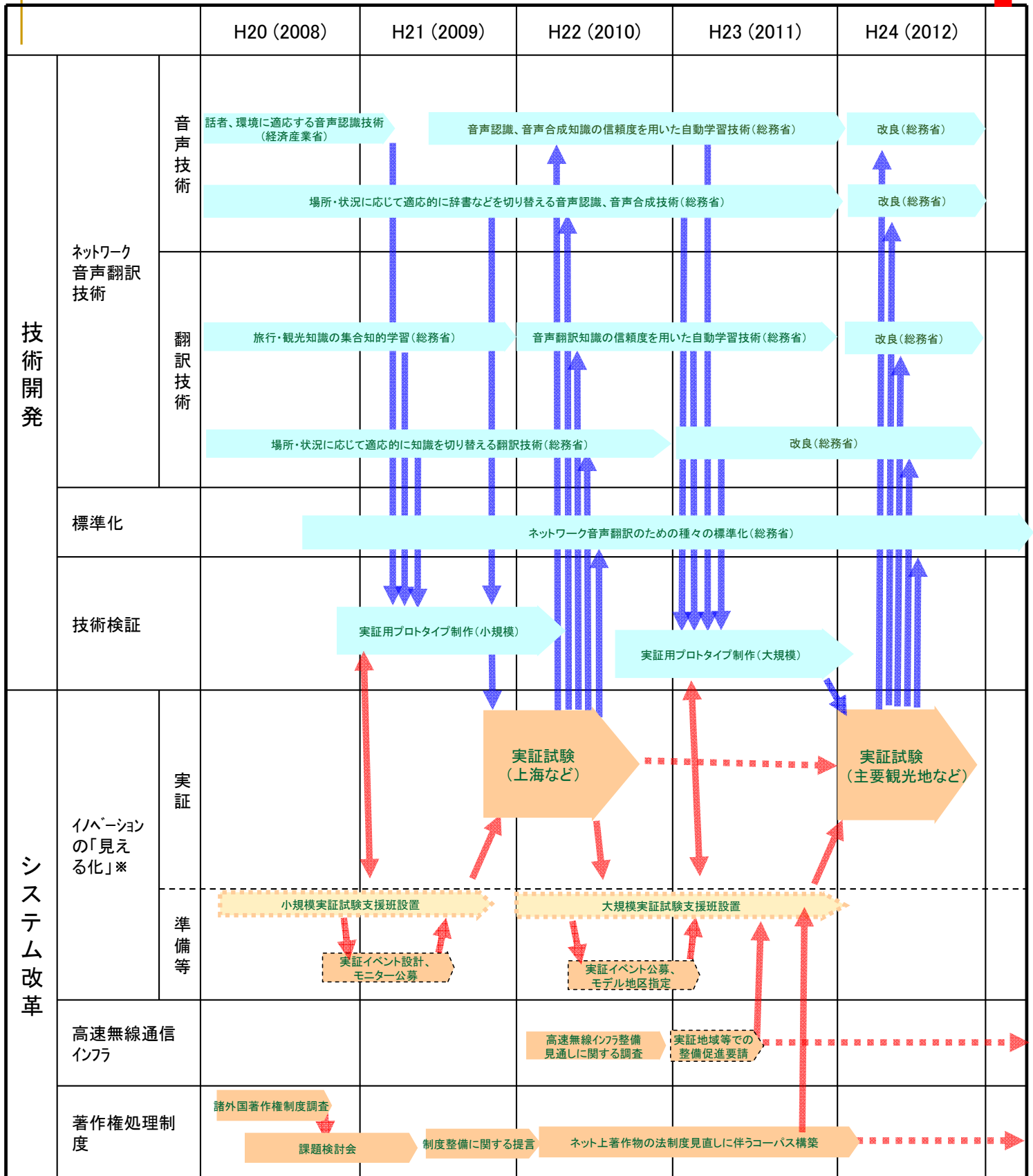
【内閣府、総務省】

※1 本計画は、研究開発の進捗状況、地域社会ニーズの変化、財政事情等を考慮しつつ、適宜見直しを行う。

※2 経済産業省では、2006～2008 の計画で「情報家電センサー・ヒューマンインターフェイスデバイス活用技術の開発」の中で音声認識技術の研究開発を行っており、2008 は本社会還元PJに位置付けられる。

※3 総務省(NICT)で対応できるのは技術性能確認までで、社会実験用総合システムの開発を行うのは予算の都合上困難。

TF活動線表



※本計画は、研究開発の進捗状況、地域社会ニーズの変化、財政事情等を考慮しつつ、適宜見直しを行う。

民間利用

(参考)

2008 年度施策一覧

(単位：百万円)

省庁	施策名	2008 年度 予算額	施策の概要
総務省	自動音声翻訳技術の研究開発	697	ネットワーク上に分散する翻訳知識を活用し、翻訳端末と組み合わせることにより、幅広い話題への対応を可能にする技術（ネットワークベース翻訳技術）等の研究開発を行い、翻訳性能の高度化（幅広い分野・話題への対応、正確でより自然な翻訳）を目指す。
経済産業省	情報家電センサー・ヒューマンインターフェイスデバイス活用技術の開発	321	情報家電の使い勝手を飛躍的に向上するヒューマンインターフェイスデバイス等消費者の利便性に直結する音声認識技術を確立し、音声認識によるヒューマンインターフェイスを核に、メーカーの違いを超えて各機器が相互連携できる環境を整え、その開発成果の普及を促進する。その結果、利用者の実生活をより充実させる環境を提供する。
	小計	1,018	

音声翻訳ロードマップ概要

