

2.7 多言語コミュニケーション技術

2.7.1 言語処理・複数言語翻訳技術の研究開発

ICT を使って言語・文化・能力などの壁を越えて自由にコミュニケーションが行える環境を実現するための多言語コミュニケーション技術の研究開発を目標として、主に人と人との言葉の壁を越えるための翻訳技術の研究を行った。特筆すべき成果としては、平成20年度から23年度にわたる総合科学技術会議の社会還元加速プロジェクトの1つに選定されたネットワーク音声翻訳プロジェクトや「高度言語情報融合フォーラム (ALAGIN)」への貢献が挙げられる。本研究開発では、日本と世界の間にある言語の壁の克服に向けて、多言語・多分野の高精度翻訳システムを社会へ提供することを目標として、以下の研究開発を行った。

①2,800万文の日本語文とその翻訳文の対から成る対訳コーパスの構築

②上記対訳コーパスを用いた高精度翻訳の実現

本研究開発は、統計翻訳手法を採用するという研究戦略に立脚して行った。①の成果を活用しつつ、高度なアルゴリズムを創出することを同時に実行するという戦略によって初めて②も実現できる。統計翻訳の概要を図2.7.1に示す。統計翻訳では、対訳コーパスから翻訳に必要な知識・モデル・データを自動的に抽出する。例えば、対訳コーパスから、2言語の単語間の翻訳という観点での対応付けを自動的に行い、統計処理することによって、(翻訳の確率が付与された)対訳辞書を生成する。このような自動生成された知識・モデル・データによって、新規入力を翻訳するシステムが構築できる。

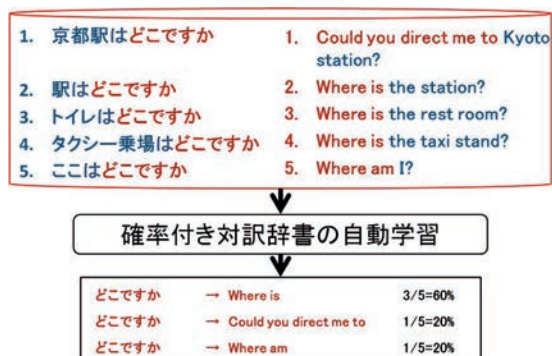


図2.7.1 統計翻訳の概要

(1) 対訳コーパスの構築

対訳コーパスを効率的に収集するために、以下のような2つの補完的なアプローチがある。

- ① Web から対訳コーパスをクロールすること (図2.7.2) や文章レベルの対訳から自動的に文レベルで対応付けする技術などのコンピュータ中心のアプローチ

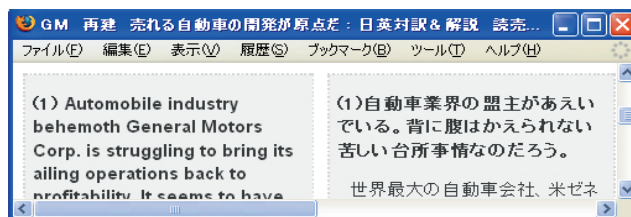


図2.7.2 Webの単一ページ内対訳データ

- ② ボランティア翻訳のホスティング・サービスや外部機関との提携等、人や社会中心のアプローチ

NICTでは、両方のアプローチを併用して精力的に対訳コーパスを集め、第2期中期計画において2,800万文の大規模対訳コーパスを構築した後に、可能なものから順次、ALAGINを通じて、公開を開始した。

さらに、②の手法を実現するために翻訳者支援サイト「みんなの翻訳」を開発した。図2.7.3に利用画面を示した。左側の窓が原文を表示し、右側の窓が訳文を表示している。原文にある下線は自動的に辞書引きされた語彙やイディオムを表す。利用者が下線にアクセスすると、辞書引き結果がポップアップし、COPY・PASTEで訳文画面に入力することができる。翻訳時間の1/3程度が辞書引き時間であることがわかっているので、この自動辞書引きの利用によって、翻訳作業が大幅に効率化できる。「みんなの翻訳」は、クリエイティブ・コモンズ・ライセンスの考えに基づき、翻訳情報を共有することで、近年爆発的に活発になっているオンライン個人翻訳者の翻訳、NPO/NGOによる翻訳の効率改善と発展を促進する。平成21年4月の公開以降、利用者数は増加し続けて、平成27年1月現在で2,990人を超え、また、日英対訳を大量に(英語単語数で126万語)集めることができた。

また、「みんなの翻訳」は平成22年6月にアジア太平洋

機械翻訳協会 (AAMT) 第5回長尾賞を受賞した。いわゆる学会の学術賞ではなく、たとえば、高性能の機械翻訳システムを商品化した、機械翻訳システムを使った新しいサービスを開始した、といった貢献を対象とした賞であり、「みんなの翻訳」が、社会に資するものと認められたといえる。

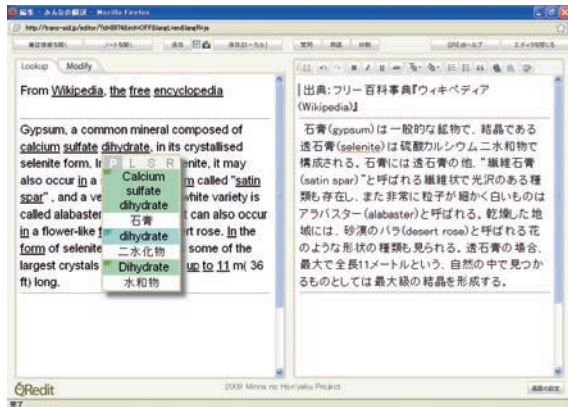


図2.7.3 みんなの翻訳

(2) 高精度翻訳

a) 翻訳技術の開発

辞書や対訳コーパスに現れない未知語は翻字(発音をなるべく変えずに2言語間で文字を翻訳すること)で処理することができる。これに関して、ディリクレ過程を用いた新しいモデルを開発した(図2.7.4)。本手法の利点はモデルがコンパクトになることと過学習しない点である。本技術は、平成22年に、翻字に関する国際コンペ ACL/NEWS2010で8つの言語対のうち5つの言語対で1位の世界最高性能を達成した。

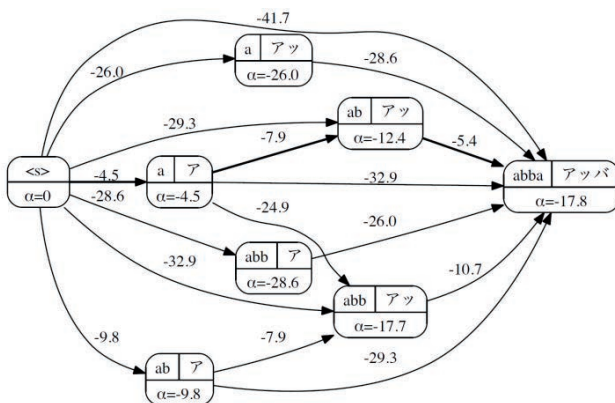


図2.7.4 日英間の翻字モデルの例

b) 単語切り出し技術の開発

翻訳技術に必須である各言語の単語切り出しプログラムは、研究が遅れていたり、よいプログラムが存在しても、種々の制約から入手困難な場合もあったりする。また、既存のプログラムが翻訳に最適とは限らない。アラビア語、タイ語、ベトナム語をはじめとして、このような問題を抱える言語は少なくない。そこで分割の初期値として文字を設定し、翻訳スコアが上昇するように分割の単位を大きくする方向で学習する手法を提案し、表2.7.1にあるような様々な言語で、高精度翻訳ができる単語切り出し技術を確立した(数字は翻訳スコアであり、値が大きいほど品質が高いことを表す)。

表2.7.1 翻訳向け単語切り出しの多言語対応

| 言語 | サンプル | 文字 | 学習 |
|------|------------------|-------|-------|
| アラビア | نعم ، انه كذلك . | 58.60 | 63.70 |
| タイ | ใจมันเป็นเช่นนี้ | 44.41 | 55.00 |
| ベトナム | Vâng, đúng rồi. | 49.91 | 60.56 |

c) 旅行会話翻訳技術の開発

対訳コーパスから翻訳システムが自動的に構築できる手法のメリットの1つに多言語化の容易性がある。N個の言語から成る多言語対訳コーパスを用意すれば、その全ての組合せ、N(N-1)個の翻訳システムが自動的に構築できる。我々は、旅行会話の分野で多言語対訳コーパス(N=21)を構築し、420通りの翻訳システムを実現、実用レベルの翻訳品質を確認した(図2.7.5)。さらに、音声認識、音声合成と組み合わせて、スマートフォン用の多言語音声翻訳アプリケーションVoiceTraとして全世界に向けて公開した。

(3) 学術的な貢献

音声翻訳に関する国際会議 IWSLT の第1回を平成6年10月に開催した。この会議では、多言語対訳コーパス BTEC を提供して音声翻訳技術の比較を目的として運営した。参加・参照が年々増加しており、標準的な会議として認知されている。また、国際会議 NTCIR の一部として特許翻訳に関する PatentMT を主催した。NTCIR7、NTCIR8は日英対訳コーパスを提供して特許翻訳技術を比較するものである。NTCIR9は The Hong Kong Institute of Education (香港教育學院) と共同で日英・日中对訳コー

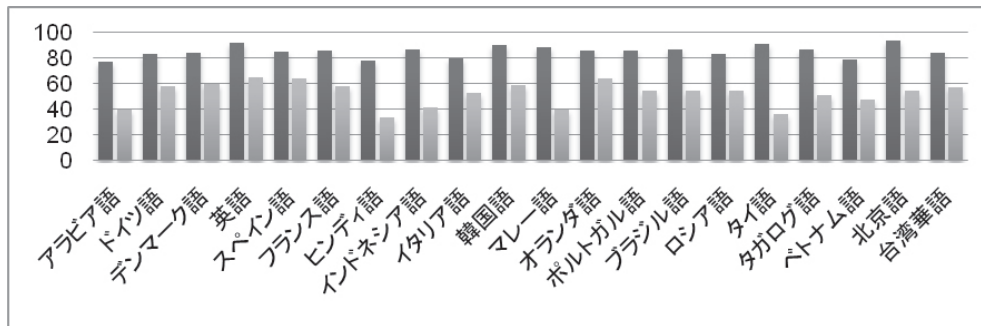


図2.7.5 多言語翻訳での翻訳品質比較 (縦軸が日本語への翻訳率、横軸が翻訳元の言語である。黒はNICTの翻訳精度で80%前後の高精度を実現している。灰色は従来技術である。)

パスを提供して特許翻訳技術を比較、多数の機関の参加を得て、相互比較により、新たな知見を明らかにした。

2.7.2 自然言語処理技術・言語グリッド技術の研究開発

平成13年より、言語障壁を起因とするデジタルデバイドの解消を目指し、将来の言語情報処理の基盤となる大規模な言語資源やツールの構築・公開を中心に、その構築・活用に資する言語処理技術や応用システム、それらを統合しサービスとして実現する言語グリッドの開発を行った。平成14年には、タイに拠点を設置し、研究成果の実証の場として活用した。

(1) コーパス・ツールの構築と公開

言語情報処理の基盤として、過去にはなかった、次のような言語資源・ツールを作成・公開した。

a) 日本語話し言葉コーパス (CSJ)^{*1}

平成11年度から5年間にわたって、文部科学省科学技術振興調整費による「話し言葉の言語的・パラ言語的構造の解明に基づく『話し言葉工学』の構築」プロジェクトを推進し、国立国語研究所と共同で650時間以上(700万語以上)の規模の日本語話し言葉コーパスを作成した。NICTでは、形態素情報(単語区切りや品詞等の情報)の自動付与、節境界、係り受け、談話構造、要約(重要文)の情報を付与した(図2.7.6)。このコーパスは現在も国内外で言語研究等に広く使われている。本プロジェクトでコーパスに形態素情報を付与するために開発した解析ツールは、その後も国立国語研究所のプロジェクトで拡張・活用され、その一部はオープンソースとして公開されている(<http://comainu.org/>)。

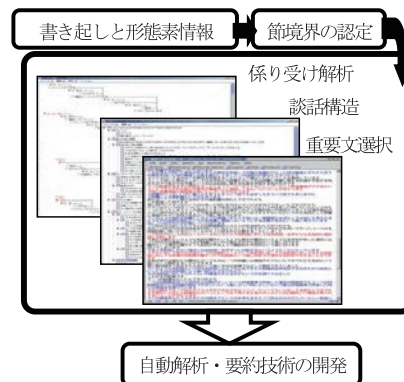


図2.7.6 日本語話し言葉コーパスへの情報付与

b) 日英新聞記事対応付けデータ^{*2}

対訳記事から自動的に文の対応を抽出するツールを開発し、平成14年には、読売新聞とThe Daily Yomiuriから自動作成した約20万文の日英対応付けコーパスを公開した。本ツールとコーパスは広く商用ライセンスされている。

c) 日本人1200人による英語コーパス (The NICT JLE Corpus)^{*3}

平成16年には、日本語を母語とする英語学習者1,281名分の英語インタビューテスト(株式会社アルクによるStandard Speaking Test)における発話を書き起こしたデータに基づく話し言葉の学習者コーパスを構築・公開した(図2.7.7)。世界最大規模の学習者コーパスであり、発話者の熟達度レベルや文法的・語彙的誤りを対象とした47種類のエラータグ等の情報も付与されている。平

*1 日本語話し言葉コーパス
http://www.ninjal.ac.jp/corpus_center/csj/

*2 日英新聞記事対応付けデータ
http://www2.nict.go.jp/univ-com/multi_trans/member/mutiyama/jea/index-ja.html

*3 日本人1200人による英語コーパス
https://alaginrc.nict.go.jp/nict_jle/

成24年にオープンソースとして公開後、約600件がダウンロードされ、研究等に活用されている。



図2.7.7 一般書籍として公開

d) 日本語 WordNet^{*4}

平成18年より、広く利用されている英語の意味辞書 Princeton WordNet を日本語に半自動的に翻訳することにより、日本語 WordNet と呼ばれる10万語規模の意味辞書を構築し、改良を重ねた。平成21年にオープンソースとして公開後、10万件超のダウンロードがあり、世界中で研究やビジネス等に活用されている。

e) 科学技術文書を対象とした日英・日中対訳コーパス^{*5}

平成18年から22年にかけて、文部科学省科学技術振興調整費による「日中・中日言語処理技術の開発研究」プロジェクトを推進し、科学技術文書を対象とした翻訳システム・辞書、及び翻訳システム用に400万文規模の日英・日中対訳コーパスを構築した。対訳コーパスの一部はアジア言語の機械翻訳に関する評価型ワークショップ (<http://orchid.kuee.kyoto-u.ac.jp/WAT/>) で活用されている。

f) EDR 電子化辞書の拡張(日中対訳辞書)^{*6}

基盤技術研究促進センターと企業の共同出資のもとに9年間のプロジェクト(昭和61年度～平成6年度)で開発された、知的情報処理のソフトインフラ開発用の電子化辞書を拡張・公開した。平成18年より、日中対訳辞書の構築を進め、平成22年に23万語規模の日中対訳辞書として公開した。商用利用も含め広く活用されている。

(2) 研究成果の実証の場：タイ自然言語ラボラトリー

平成14年には、言語処理技術や構築したコーパス・ツールをアジア言語へと拡張・適用する活動の拠点としてタイにラボが設置された。アジア言語の言語資源の構築には、各国の参加者による共同作業が不可欠である。平成19年には、インターネット上での共同作業を支援するツール KUI(図2.7.8)を構築した。このツールはタイ科学技術省の公式ホームページでも活用された。



図2.7.8 共同作業支援ツール(KUI)

また、同年、アジア圏での自然言語処理の研究コミュニティの構築を目標に、アジアの10か国を超える国々から参加者を募り、自然言語処理に関するスクールを開催した(図2.7.9)。



図2.7.9 スクール風景

(3) 統合サービスの実現：言語グリッド

平成18年には、言語の壁の克服に向けて、インターネット上の言語インフラストラクチャーである言語グリッド

*4 日本語 WordNet

<http://nlpwww.nict.go.jp/wn-ja/>

*5 科学技術文書を対象とした日英・日中対訳コーパス

<http://lotus.kuee.kyoto-u.ac.jp/ASPEC/>

*6 EDR 電子化辞書の拡張(日中対訳辞書)

http://www2.nict.go.jp/out-promotion/techtransfer/EDR/J_index.html

の開発及びそれを利用した異文化コラボレーションツールの研究開発を行った。従来、Webは、文書や画像をはじめとする、いわゆる「コンテンツ」が提供される場であった。近年になり、ソフトウェア機能やヒューマンリソースなどを含めた、無形の「サービス」の提供が行われるようになってきている。言語グリッドは、このような変化を背景として、サービスを単位として集合知の形成を図り、インターネット上に多言語サービス基盤を実現するものである。

言語グリッドの主な貢献として、サービス指向の言語基盤の構築、参加型デザインの実践、連邦制運営、の3点が挙げられる。

① サービス指向の言語基盤の構築

言語グリッドでは、世界各地の大学・研究機関・企業が持つ言語資源が、言語サービスとして提供されている。プロジェクトが終了した平成23年3月時点までに、世界18か国138組織が参加し、94の言語サービスが言語グリッド上で共有された。図2.7.10に、参加組織の数の推移と内訳を示す。サービスの種類は、機械翻訳、対訳辞書、用例集、形態素解析など多岐に渡る。

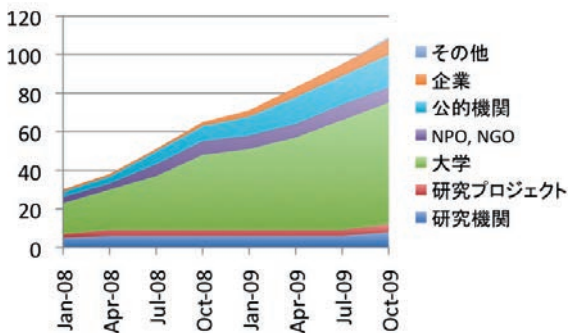


図2.7.10 参加組織数の推移

共有されたサービスは、非営利目的であれば、参加組織間で自由に相互利用できる。これにより、例えば汎用の機械翻訳を、コミュニティ特有の専門用語の対訳辞書と組み合わせ、特定のコミュニティに特化した翻訳の精度を向上させるといったことが可能となる。

② 参加型デザインの実践

言語グリッドは当初から、NPO、NGOなどと共同で開発と利用を推進してきた。そのため、プロジェクト開始時にユーザーグループ「言語グリッドアソシエーション」を発足させ、医療受付支援・教育支援など様々な目的を

持つサブグループが活動を行った。言語グリッドのサービスを利用し、平成19年頃より各グループが現場に特化した多言語コミュニケーションツールを開発していった(図2.7.11)。



多言語医療受付システム(2007)
(開発は和歌山大学と多文化共生センターきょうと)
21年度バリアフリー・ユニバーサルデザイン推進功労者表彰を首相官邸で受賞



国際NPOのための多言語掲示板(2008)
(NPO パンゲア)

図2.7.11 言語グリッドを用いた多言語コミュニケーションツール

さらに、ユーザによる多言語コミュニケーションツール利用を促進するため、カスタマイズ可能な多言語コミュニケーションツールである言語グリッドToolboxを、平成21年9月に公開した。図2.7.12は、言語グリッドToolboxに含まれる機能の1つである、多言語掲示板である。日本人学生と留学生に対して、それぞれの母語でメッセージを表示する。



図2.7.12 言語グリッドToolbox(多言語掲示板)

その他、辞書作成機能、翻訳サービスの組み合わせ機能など、15種の機能を順次開発していった。各機能は

カスタマイズが容易なように設計され、利用者コミュニティごとに、そのニーズに特化したデザインや機能セットが実現できる。

この特徴を生かし、平成22年3月、大学やNPO等に向けて言語グリッド Toolbox をクラウドサービスとして提供を開始した。このクラウドサービスは、国立大学法人東京外国語大学や国立大学法人京都大学など約30組織において、留学生支援や国際交流活動などに活用された。

③連邦制運営

言語グリッドの基盤ソフトウェアを複数の拠点に導入することで、それぞれの参加組織が相互に言語サービスを連携して利用することが可能である。言語グリッドの運営は当初、国立大学法人京都大学大学院情報学研究所社会情報学専攻によってのみ行われていた。しかし、海外の地域でも言語グリッドの運営組織を立ち上げることで、より多様な言語サービスを共有できる。そこで、タイの国立電子・コンピュータ技術研究所(NECTEC)において、平成23年2月に、言語グリッド運営組織を発足させると共に、2つの言語グリッドを接続し、連邦制運営を実現した(図2.7.13)。これにより、新たに東南アジアの言語資源を中心として、13か国語21の言語サービスが、従来の言語グリッドのユーザからも利用可能になった。

NICTにおけるプロジェクトは平成23年3月で終了したが、言語グリッドの運営は国立大学法人京都大学を中心として継続されており、ベトナムにおける農業の国際的支援活動など、応用を広げながら現在も発展を続けている。(http://langrid.org/jp/)

2.7.3 対話システムの研究開発

人間と機械のコミュニケーションにおいて、人間の自然なコミュニケーション手段である音声を用いることができれば、特別な訓練をすることなく機械との対話が可能である。本研究プロジェクトでは、いつでも、どこでも、何語であっても、音声や身振り・手振りなどの人間にとって自然な言語・非言語表現を用いて、コミュニケーションを実現するための技術の研究開発を行った。本稿では、音声対話システムを構築するために行った、研究基盤となる対話コーパスの整備、非言語情報を利用するための画像処理技術の開発、音声認識・合成技術の改良及び多言語化など、対話システムを構成する要素技術の開発、さらに、これらの技術を統合した対話処理のプロトタイプシステムの開発、実証実験について述べる。

(1) 音声対話のための音声認識・合成の研究開発

音声対話では、多様な話者や発話スタイルの音声を実時間で認識・理解し、システムから応答を返す必要があることから、高精度な実時間音声認識・合成の機能が必須である。

a) 話者非依存の多言語音響モデルの構築

高精度な音声認識モデルを構築するために実利用データの大規模な音声コーパスが必要であるが、不特定話者の大量の実利用データを収集することは困難であることから、利用場面を想定して実験室で被験者から収録して構築した模擬会話音声コーパスにより、平成23年10月に音声認識評価実験を行った(表2.7.2)。日本語を単語正解精度で評価したところ、実利用ログ94%の性能を得た。

b) 音声対話に適した自然な合成音声

音声合成技術に関しては自然な対話音声の実現、多様

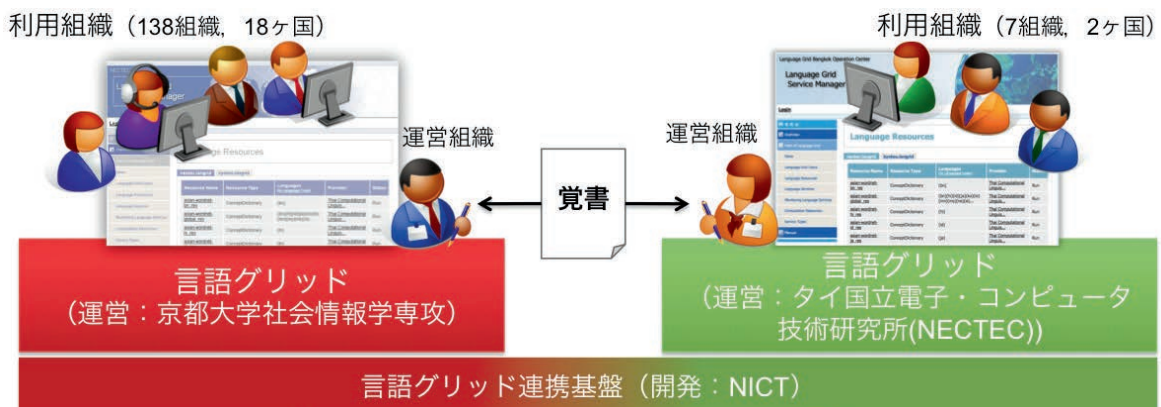


図2.7.13 言語グリッドの連邦制運営

な話者への適応を目的として、対話音声コーパスから音響モデルを学習することで対話に適した韻律で音声を合成する方法を研究開発した。混合励振源モデルの枠組みを確立し、合成音声のブザー感を与える雑音を軽減した。

表2.7.2 多言語音声認識の性能評価

| 言語 | 話者数 | 文章数 | 発話時間 | 学習データ | テストデータに対する性能 (単語正解精度 [%]) | |
|-----|------|---------|-------|--------|------------------------------|------------|
| | | | | | 旅行会話文 読み上げ音声 | 模擬 旅行会話 |
| 日本語 | 4500 | 226,673 | 387.6 | 模擬旅行会話 | 94% | 89% |
| 英語 | 930 | 236,737 | 257.6 | 模擬旅行会話 | 91% | 80% |
| 中国語 | 540 | 213,352 | 251.0 | 模擬旅行会話 | 91% | 80% |

(2) 音声対話コーパスの整備

a) 発話行為タグ付き音声対話コーパスの構築

より人間らしく応答するシステムを実現するため、京都観光案内をタスクとしてガイドとユーザの人間対人間の音声対話コーパスを構築し、統計的な対話制御モデルを構築した。学習データとして、人と人の対面対話の分析を行い、発話行為タグ・コンテンツタグ付与によるデータを整備した。実利用場面ではコンピュータと人間が対話することから、コンピュータとの対話における人間の自然な振る舞いを調べるため、ガイドがコンピュータのふりをしてユーザと対話を行う WOZ (Wizard of Oz) 形式のデータも収集した。研究成果として、対面対話100対話(約50時間)を京都観光案内対話データベースとして、平成22年3月に ALAGIN を介して公開した。

b) 京都観光案内データベースの構築

京都観光案内の大規模実証実験を目的として、京都市内の主要な100か所の観光スポットに関する詳細なデータベースを整備した。既存の簡易情報とあわせて合計840件のデータベースを平成21年11月に構築した。

c) 京都観光案内システムの多言語化

多言語音声対話システムの実現に向けて、英語話者による対話音声コーパスの収録、ソフトウェアプラットフォームの多言語対応化、対話シナリオの英語化を平成24年3月に行った。

(3) 音声対話制御の研究開発

a) 統計的対話制御システム

システムによる自然な応答を自動生成するため、構築した音声対話コーパスより統計的対話制御モデル(重み付き有限状態トランスデューサ: WFST)を学習し、これを用いて対話制御を行うことで、人間に近い自然な対話を実現する手法を研究開発した。WFSTに基づく音声対話制御は、発話意図 WFST、システム応答を決定するシナリオ WFST を合成することにより対話制御 WFST を生成することからシステムの拡張性が高く、さらに対象分野、タスク間の移植性も高いという特徴がある。図2.7.14に WFST に基づく統計的対話制御モデルの状態遷移図を示す。この例では、自然な観光案内対話を実現するためには、ユーザのリクエストとして177種類の入力を理解し、471種類以上の異なる応答をする必要がある。

b) 嗜好モデルに基づく対話制御

2,000名の被験者を対象に評価グリッド法及びWeb調査により抽出した嗜好間の優先順位、因果関係などを表現する選好構造に基づいた調査を行い、京都の観光スポットを推薦するシステムを平成22年11月に開発した。

(4) 非言語情報を用いた自然な音声対話の研究

人間の胸部の高さに設置したカラーステレオカメラの画像を入力として、複数の人物が融合した3次元点群のクラスタを分割して個人領域に切り分けることにより、頭部位置を高精度(混雑環境で93%、疎な環境で

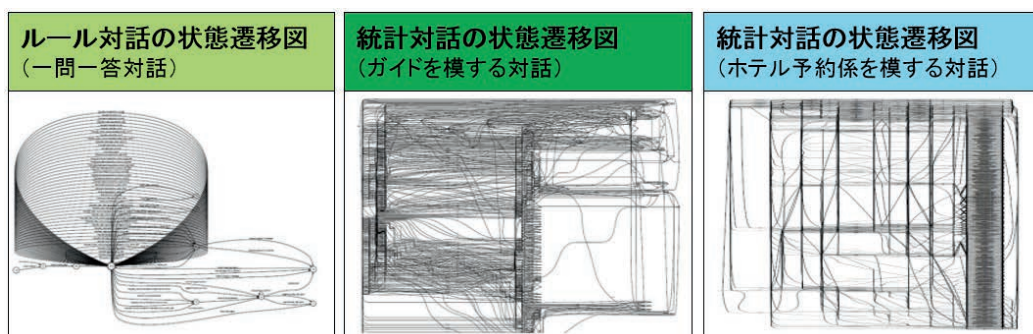


図2.7.14 WFST に基づく統計的対話制御モデルの状態遷移図

99.7%)に推定するシステムを平成22年11月に開発した。また、複数台カメラの誤差を最急降下法で最小化することにより、特別な照明なしにユーザの顔の向きを高精度(誤差は約5度)に推定することに成功した。さらに、ユーザの頭部検出に基づく話者の属性推定、顔の向き推定とそれに基づく興味推定、システムの誤応答への反応の検出を行う技術を平成22年11月に開発した。

(5) 音声対話システムを用いたスマートフォンによる実証実験

日本語音声での受け答えを通じて観光コンシェルジュのように京都観光をサポートする、iPhone向けアプリ“AssisTra”(アシストラ)を平成23年6月に公開し、実証実験を行った(図2.7.15)。本システムの開発に関しては平成24年にドコモ・モバイル・サイエンス賞を受賞した。



図2.7.15 iPhone向けアプリ“AssisTra”

音声対話の研究結果を統合して大画面音声対話システム(図2.7.16)を開発し、学会・展示会で動態展示すると共に、被験者実験により非言語情報処理を考慮した音声対話システムの有効性を平成22年12月に確認した。

(6) 音声対話技術のロボットへの応用

平成22年7月、音声対話、機械学習技術の機能実証としてロボカップ世界大会家庭用ロボット部門に参加した結果、家事動作の模倣学習技術が評価され、24チーム中で優勝した。

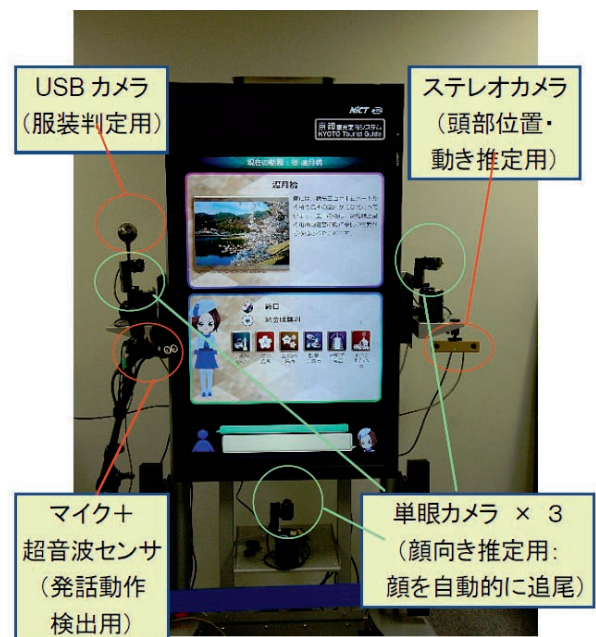


図2.7.16 大画面音声対話システム

2.7.4 音声コミュニケーション技術の研究開発

① Web上の音声データに対する字幕付与、検索のための自動インデキシング技術、②人間と人間、人間と機械の簡便なコミュニケーションを実現するための音声インタフェースの研究を行った。本稿では、音声認識の研究基盤となる学習データに用いる音声コーパスの収集とそれに基づく音声認識性能の改善、及び多言語音声合成システムの研究開発について述べる。

(1) 高精度かつ高速な大語彙音声認識技術

a) WFSTを用いた音声認識システム

音声認識技術を実世界へ応用するには、人間が話す豊富な語彙を対象として実時間での認識が必須となる。実時間で高精度な音声認識を実現するために行った、WFSTに基づく音声認識エンジンの研究開発において、従来多くの処理時間を要していた新語彙追加の手続きをほぼゼロにする高速なアルゴリズムを発明し、それをNICT独自の音声認識システムSprinTraに平成23年3月に導入した。図2.7.17にWFSTに基づく音声認識アルゴリズムを示す。その結果、10倍の語彙サイズの認識で速度を保ちつつ従来手法に対して単語正解精度で30%の性能改善を達成した。SprinTraはNTTドコモの「しゃべってコンシェル(平成24年3月公開)」、KDDIの「おはなしアシスタント(平成24年11月公開)」に採用された。

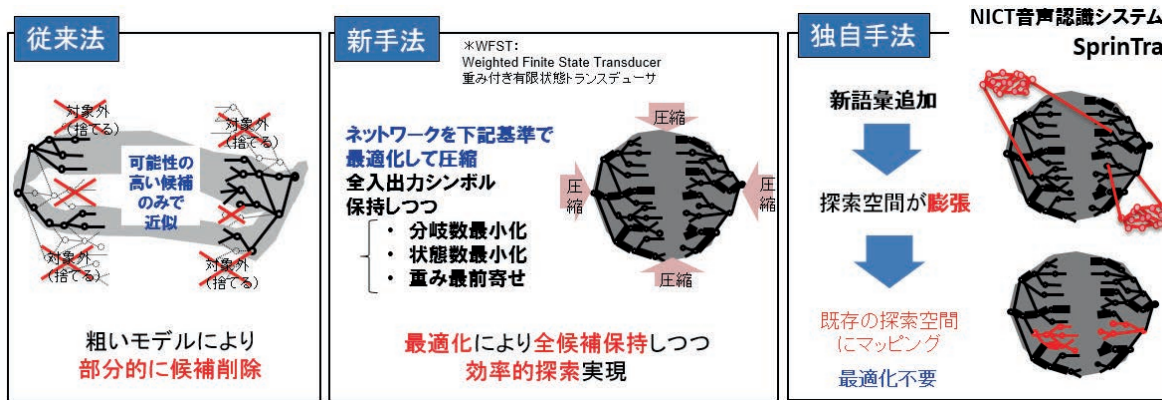


図2.7.17 WFSTに基づく音声認識アルゴリズム

b) 英語講演の音声認識

TEDの英語講演音声を対象として音声認識性能技術を競う評価型国際ワークショップ IWSLT に参加し、英語講演音声認識において世界第1位を2年連続で獲得した(平成24年12月、平成25年12月)。表2.7.3に示すように IWSLT には、世界トップの研究機関が参加している。

(55.4% → 61.0%)、中国語(67.5% → 77.6%)の単語正解精度を平成25年3月に達成した。

表2.7.3 IWSLT における音声認識性能競争

英語音声認識の評価結果

| 参加研究機関 | 評価セット 数字は単語誤り率 (%) | | |
|-------------|----------------------------|-----------------------------|-----------------------------|
| | 評価セットA tst2011 (8講演) | 評価セットB tst2012 (11講演) | 評価セットC tst2013 (28講演) |
| NICT | 7.9 | 8.6 | 13.5 |
| KIT | 9.3 | 9.6 | 14.4 |
| MIT-LL/AFRL | 10.6 | 11.3 | 15.9 |
| RWTH | 10.2 | 11.3 | 16.0 |
| NAIST | 9.1 | 10.0 | 16.2 |
| UEDIN | 10.2 | 11.6 | 22.1 |
| FBK | 13.6 | 16.2 | 23.2 |
| PRKE/IOIT | 14.6 | 16.2 | 27.2 |

KIT: カールスルーエ工科大学 (ドイツ)
 MIT-LL/AFRL: マサチューセッツ工科大学リンカーン研究所/空軍研究所 (アメリカ)
 RWTH: アーヘン工科大学 (ドイツ)
 NAIST: 奈良先端科学技術大学院大学 (日本)
 UEDIN: エディンバラ大学 (イギリス)
 FBK: ブルーノ・ケスラー財団-研究所 (イタリア)
 PRKE/IOIT: ベトナム科学技術アカデミー情報技術研究所 パターン認識及び知識工学部 (ベトナム)

(2) VoiceTra による実証実験

音声翻訳アプリ VoiceTra (図2.7.18) で実証実験を実施し、11,136,285 発話、857,257 ダウンロードを達成し(平成22年8月～平成25年3月)、従来の20倍規模の約8,000時間の音声データを取得した。約600万文(約5,000時間)の実利用ログの一部を用いて学習することにより、実証実験のデータに対して日本語(71.1% → 83.4%)、英語



図2.7.18 音声翻訳アプリ VoiceTra

(3) 音声翻訳プロトコルの世界標準化

NICT が主導する国際研究共同体「U-STAR」は、25か国・30研究機関(平成27年1月現在)が共同で自動音声翻訳技術の研究を推進している。U-STAR は、平成22年に NICT が国際標準化を主導した ITU-T 勧告書 F.745 及び H.625 に準拠した技術を用いて、加盟機関の音声翻訳サーバをネットワーク型音声翻訳通信プロトコルで相互接続し、クライアントアプリケーションを介した多言語音声翻訳を平成24年7月に VoiceTra4U として実現した(図2.7.19)。17言語の音声認識、14言語の音声合成、23言語のテキスト翻訳を実現し、音声データを収集した(図2.7.20)。タイ語の収集音声を用いて学習したところ、単語正解精度が30%向上し、60%になった。

(4) Web 上の音声データを用いた認識性能改善

Web 上の大量の音声・テキストデータを用いて音響モデル、言語モデルを学習し、自動字幕付与及び自動インデキシングを実現する技術の研究開発を行っている。

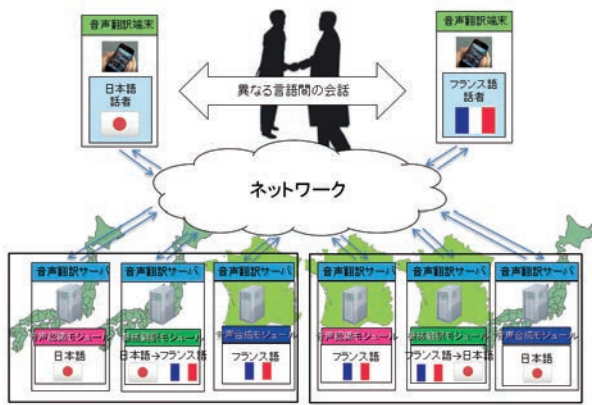


図2.7.19 ネットワーク型音声翻訳システム

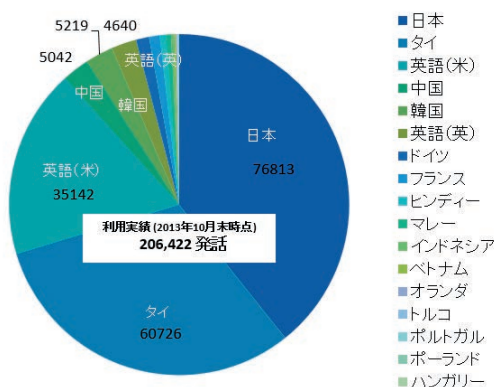


図2.7.20 VoiceTra4Uによる音声データ収集

インターネットを通じて、合計5,000時間の多言語音声データを収集する目標に対して、中国語音声データを約800時間、英語音声データを約6,000時間収集し、音響特徴量を抽出した。その特徴量を用いて中国語のニュース音声認識精度の単語正解率を平成25年度に77.2% (平成24年度59.4%)、英語ニュース音声認識82.9% (平成24年度63.8%)に向上させた。

(5) 多言語音声合成システム

隠れマルコフモデル (HMM) に基づき、7か国語 (日本語、英語、中国語、韓国語、インドネシア語、ベトナム語、マレー語) の音声合成システムを構築した。また、収録音声进行处理するフィルタを高精度化し合成音声の自然性を改善し、主観評価により従来法に比べ40%の改善を平成23年6月に確認した。

さらに、多言語音声翻訳システムでは、ユーザ本人の声質を用いて多言語で音声合成できれば、自然性が高まることが期待できる。そこで、合成音声の研究では、異なる言語間で類似した性質を選択するヴォイス・セレクターの研究開発を行った (図2.7.21)。

2.7.5 多言語コンテンツ処理技術の研究開発

人と人との言葉の壁を克服するため、日本語と英語のように文法や語彙が大きく異なり、翻訳が最も困難な言語対の翻訳の研究を実施している。この成果としての汎用アルゴリズムには、より翻訳が容易な言語対にも適用できるという大きなメリットがある。NICTは対訳データ (原文と訳文の対を集積したもの) に基づいて翻訳する統計翻訳手法を採用している。平成18年から、自動収集やコミュニティとの協業など多数の手法によって、基盤である対訳データ構築を効率化し、大規模化してきている。成果は、多数の、技術移転、受賞、国際会議・国際論文誌採録、という形で、産学官から広く評価されている。

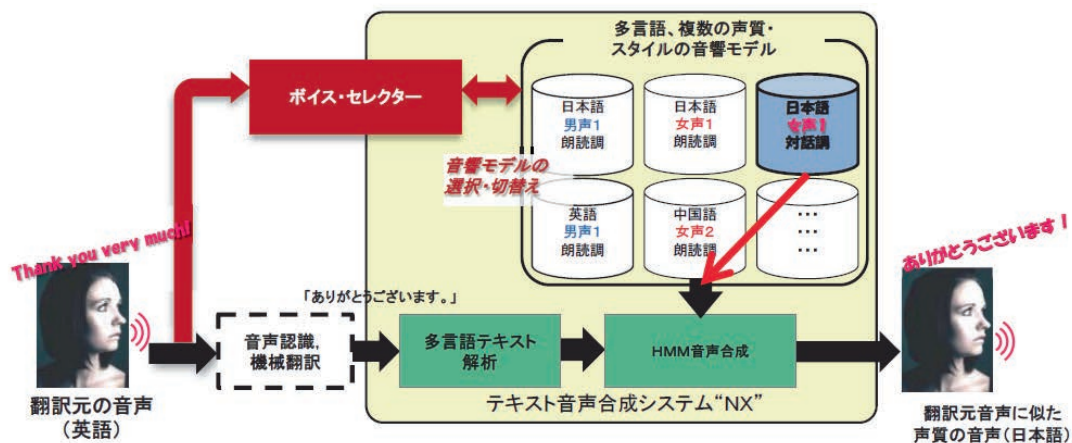


図2.7.21 異なる言語間で類似した声質を選択するヴォイス・セレクターシステム

(1) 翻訳アルゴリズム

専門分野向け高精度自動翻訳システムを多分野・多言語で実現する技術は、波及性も高く、社会経済的に重要である。この背景の下、特許などの長文テキストの自動翻訳の基礎技術の研究に注力した。

a) 長文の自動翻訳の高精度化

文法が異なる言語間の翻訳の際、NICTの自動翻訳方式では、自動学習した語順変換規則に従って、語順を変換した後、自動学習した確率付きの対訳辞書を用いて訳語を選択する(図2.7.22)。また、対訳辞書の規模は約10億に上り、従来ソフトの専門用語収録数の約100～1,000倍に相当する大規模なものになっている。中日、英日の平均約25語の特許データで翻訳実験をし、同手法は翻訳率で従来法に大きく優る80%、85%の高精度を平成25年に実現した。

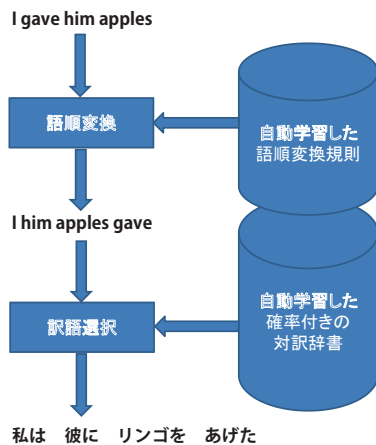


図2.7.22 長文の自動翻訳のNICT提案の新方式

b) 長文の翻訳に不可欠な解析技術を改良

精度の向上と記憶容量の削減を同時に実現し、特に、平成26年には、40語以上の長文翻訳の翻訳率を大幅に改善した(表2.7.4)。その方法は、(i) 構文データを翻訳に適した構造に変換、(ii) 変換後のデータの単語ラベルを、別途TEXTから学習したクラスで置換、(iii) 潜在的な文法記号を使った確率的文脈自由文法を学習、である。

表2.7.4 長文の自動翻訳の改良

| | 翻訳率 (通じる%) | | |
|-----|------------|-------|-------|
| | 全体 | 40語未満 | 40語以上 |
| 改良前 | 67 | 71 | 54 |
| 改良後 | 84 | 87 | 75 |

c) 自動翻訳のための係り受け木における教師なし品詞推定を提案

原言語の単語に加えて目的言語の単語から、隠れ状態として品詞タグZを推定する(図2.7.23)ことにより、単言語の既存知識に基づく品詞タグにある翻訳に向いていない点を自動的に修正する。

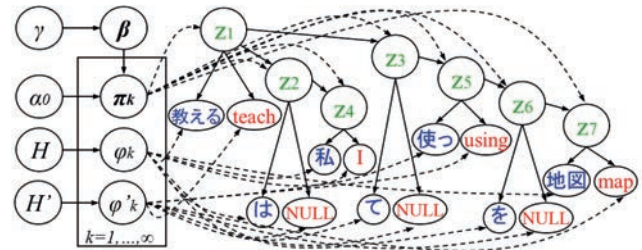


図2.7.23 機械翻訳のための品詞推定

(2) 技術移転

a) テキスト翻訳の事業化

日本特許情報機構(Japio)と共同研究し、自動翻訳システムの訳文に対して、特許の専門家が目を見た「通じる翻訳」に向けた評価によって効率的な改良を行うことができた。新技術に基づいて開発した、「中日自動翻訳ソフトウェア」の翻訳者が判定した精度は、従来技術の3倍以上の値を達成した。この「中日自動翻訳ソフトウェア」によって、Japioは中国の特許文献を日本語に翻訳及びデータベース化し、特許情報検索サービス(図2.7.24)の拡張版として平成25年4月に事業化した。企業の知財部や弁理士の知財調査や特許庁審査官の先行技術調査や、日本の企業の特許侵害のリスクの軽減などに役立つことができる。

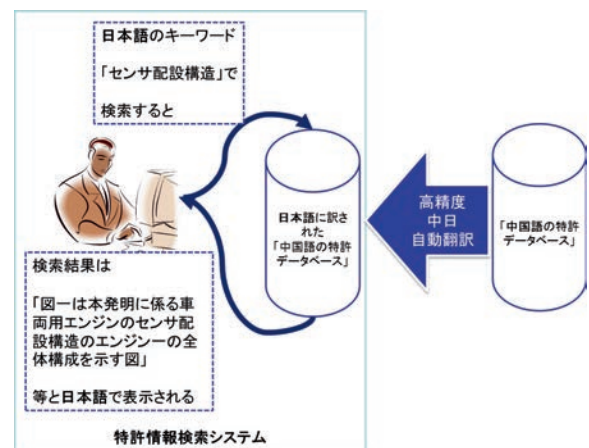


図2.7.24 翻訳の検索システムでの利用

日本発明資料株式会社(ニッパツ)と共同で開発した「英日自動翻訳ソフトウェア」では、(特許要約1件あたりの)訳語誤り数を従来技術と比べて、約12分の1に削減するという高い品質を実現した。開発した「英日自動翻訳ソフトウェア」によって、ニッパツは、①英語特許文献を日本語に翻訳するサービス、②日本語に自動翻訳した英語特許文献データベース、などとして平成25年5月に事業化した。

b) 音声翻訳の技術移転

平成22年8月に公開した音声翻訳アプリ VoiceTra は、技術の見える化、ひいては事業化に大きく貢献した。最初に、平成23年12月に成田国際空港株式会社での NariTra がサービスを開始した(図2.7.25)。その後、KDDI(おはなしアシスタント、平成25年7月、図2.7.26)を含む多数の民間事業者に音声翻訳のプログラム・特許を技術移転した。



図2.7.25 成田空港の音声翻訳



図2.7.26 KDDIの音声翻訳

c) 開発したプログラムのオープンソース化

開発した機械翻訳のプログラム*7をオープンソースとして公開し、講習会を開催(平成26年3月4~7日)、自動翻訳技術の教科書を出版(コロナ社、平成26年1月22日)するなど、アカデミアへの貢献も活発に行った。

(3) 受賞

音声翻訳プロトコルをITU-Tで国際標準化して多言語化の基礎を整え、23か国26機関とU-STARを組織し、世界人口の約95%をカバーするシステム「VoiceTra4 U-M」で世界実験を開始し、活発に民間へライセンス供与したという一連の音声翻訳に関わる業績が評価され、平成24年度第58回前島密賞を受賞した。同賞史上初めて受賞者に外国籍の研究者(Dixon Paul Richard 研究員、FINCH

Andrew Michael 主任研究員、PAUL Michael 主任研究員)が含まれていたことも注目された。

「音声翻訳技術の実用化」で、第11回産学官連携功労者表彰総務大臣賞(平成25年8月、NICT多言語翻訳研究室室長・隅田英一郎、株式会社フィート代表取締役社長・小林照二、成田国際空港株式会社代表取締役社長・夏目誠)受賞(図2.7.27)。



図2.7.27 総務大臣賞受賞

(4) 特記事項

社会還元加速プロジェクト「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」を5年計画(平成20年度に開始、平成24年度末終了予定)で実施していたところ、研究計画を上回る成果を出したため、1年前倒しで平成23年度末に成功裏に終了した。

*7 機械翻訳プログラム

●cicada、http://www2.nict.go.jp/univ-com/multi_trans/cicada/

●expgram、http://www2.nict.go.jp/univ-com/multi_trans/expgram/