

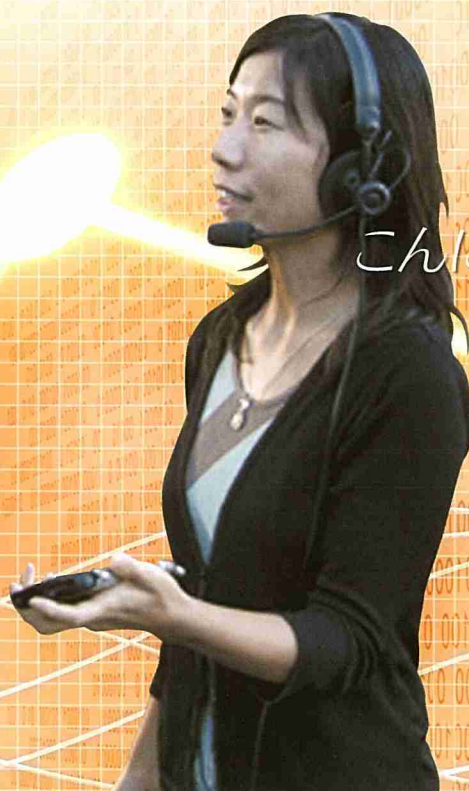
ATR Journal

AUTUMN 2005

No.54



Hello



こんにちは

ATRの音声翻訳技術

ATR

●特集

ATRの音声翻訳技術

大規模コーパスベース対話音声翻訳技術の実現にむけて

1 中村 哲

Speech Translation is Growing up:

4 Alex Waibel

From Concept Demonstration to Lasting Impact

多言語音声認識技術

6 中村 哲
菊井玄一郎

音声翻訳処理の全体を統合する技術

8 菊井玄一郎
清水 徹

コーパスに基づくマルチエンジン機械翻訳技術

10 佐々木 裕
隅田英一郎
今村 賢治

コーパス・ベース多言語音声合成システムXIMERA

12 平井 俊男
河井 恒

●脳活動イメージングセンタから

ATR脳活動イメージングセンタ利用の音声研究と
その成果展開

14 正木 信夫

●特許紹介

注目特許

15 荒木 晃司

●トピックス

ATRグループ企業情報

16

●所員往来

平成17年4月～平成17年9月

17

●受賞 等

19

●編集後記

20

大規模コーパスベース対話音声翻訳技術 の実現にむけて

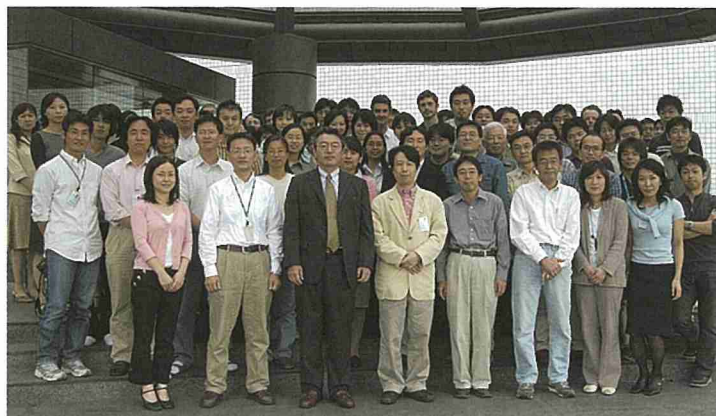


音声言語コミュニケーション研究所所長
中村 哲

平成13年から開始したNiCT民間基盤技術促進制度受託課題「大規模コーパスベース音声対話翻訳技術の研究開発」の終了を半年後に控え、これまで研究開発を行ってきた音声翻訳技術の評価と総まとめに入っている。このプロジェクトは、平成12年から開始し途中で終了したKTCプロジェクト「知識利用型音声言語通信の基礎研究」の中から、音声翻訳に関わる部分について当該研究受託制度に応募して採択されたものである。

著者は、平成12年度に奈良先端科学技術大学院大学助教授から転身して当初研究室長の立場でプロジェクトに参画した。平成13年に突然KTC制度がなくなりTAOの民間基盤受託制度に変わると言うことでドキッとしたが、幸い提案が採択され、無事当初の研究を進めることが出来た。現在は、社長、所長として長年音声翻訳プロジェクトを率いてこられた山本誠一氏が同志社大学に教授として転出されたため、所長職を引き継いでいる。

平成13年当初は、ATRでは用例翻訳に基づく音声翻訳技術が研究開発されていたが、音声認識で培われた統計的手法が言語処理分野にも徐々に適用され始めた時期であった。この用例翻訳には、話題のポータビリティの問題、新しい言語対への展開の問題があった。そこで、広い話題、日英に加え日中を始めとする新しい言語対への対処を目的として、用例ベースと統計ベースを含めたより広い概念であるコーパスベース音声翻訳技術の研究と、それらを実際の実証、稼働させるための大規模コーパス構築を目的に「大規模コーパスベース音声対話翻訳技術の研究開発」を提案したのである。



この研究開発の結果、日英 100 万文、日中 50 万文の対訳コーパスの完成、マルチエンジン型コーパスベース音声翻訳により、日英では TOEIC スコア 650 点と同等の翻訳性能を達成するに至り、音声翻訳技術は飛躍的に進歩した。関西空港における実証実験での利用者アンケートでは、「ほとんど通じた」が 50 % を超えた。

個別要素技術毎に述べると、音声認識については、実際の環境での利用を目指し、音声認識においては、雑音、発話スタイルに頑健で接話マイクのいない多言語（日本語、英語、中国語）音声認識技術、不適切な発話を棄却する技術開発を目指した。音声認識の研究で著名な米国の国防総省（DARPA）のプロジェクトの騒音下英語音声認識評価プロジェクト（SPINE: Speech Recognition in Noisy Environment）にも 2 年続けて挑戦し、英語音声認識の経験のない中、日本から初参加ながらまずまずの成績を収めた。また、ハンズフリー音声通信に関する国際ワークショップ（HSC）を創設し、今年第 2 回 HSC が米国で開催された。今年度は、小型携帯端末用 8ch のマイクフォロンアレーシステムの試作を行い、雑音抑圧フィルタの併用により実雑音環境で非常に高い性能の音声認識を達成するに至った。

音声合成については、ATR は *talk*, CHATR などコーパスベース音声合成の提案者として有名であったが、実際には実利用可能な音声合成システムの開発が求められていた。そこで、大規模コーパスを構築し、日本語、中国語コーパスベース音声合成システム（XIMERA）を構築した。英語については、現在、鋭意システム開発中である。

言語翻訳に関しては、日英 100 万文対、日中 50 万文対の対訳コーパスを構築するとともに、これらのコーパスに基づき、マルチエンジン機械翻訳システムを構築した。翻訳エンジンとしては、ATR が従来から取り組んできた用例翻訳エンジンに加え、近年、欧米で盛んになってきている統計翻訳技術をいち早く導入することにより日中・日英の統計翻訳エンジンを構築し、現在もこの分野における日本の先導的役割を担っている。

また、要素技術を統合して音声翻訳としての総合性能を向上させる技術として、音声認識結果の信頼性に基づいて音声翻訳処理を制御する技術や複数の音声認識候補を使って翻訳する技術、さらに、システム全体の実験・評価などの研究を精力的に行った。

特に、昨年度、音声翻訳に関する国際コンソーシアム C-STAR の機関と協力して、音声翻訳技術に関する世界初の評価型の国際ワークショップ IWSLT2004 を企画し ATR にて開催した。このワークショップでは、ATR で開発したコーパスの一部を参加者に無償で提供し、同一のデータで翻訳システムの学習、チューニング、テストを行い、結果を統一的に評価して技術の有効性を分析するものである。今後の音声翻訳技術に非常に重要な意味を持つ活動であり、平成 17 年の 10 月にも第 2 回の IWSLT2005 が米国カーネギーメロン大学で開催された。

海外においても、音声翻訳の研究が加速してきている。欧州では、TC-Star と呼ばれる講演、講義の音声翻訳プロジェクトが昨年から 3 年計画で始まり、来年度からは米国で国防総省（DARPA）の大型プロジェクトとしてアラビア語、中国語と英語のテキスト翻訳および音声翻訳プロジェクトが開始される。軒並み各地域で大型のプロジェクトが開始され、我々を追撃する様相を呈している。

一方、技術の成果展開についても、新たに、マルチクライアントプロジェクトという形式の技術移転プロジェクトを行った。2002 年から 2 年間は音声認識合成技術、2004 年から 2 年間は音声翻訳技術を対象とし、それぞれ、10 社、4 社の参加を得て、ATR のコーパス、ソフトウェアをパッケージ化することに成功した。これまで研究者がバラバラに作成し再利用性の低かったソフトウェアを、ソフトウェア技術者の導入により、統一した設計、一貫したプログラミング、日本語、英語のマニュアル整備などにより完成度の高い物とした。現在、この成果はプロジェクト参加企業に移転され、各企業で実用化の検討がなされている。さらに、2005 年 8 月には、音声翻訳関連技術を核とした事業展開を行う孫会社 ATR-LANG を ATR 自ら立ち上げた。この孫会社を通じて、技術移転、ライセンス、商品販売、サービス事業を行っていく予定である。

ATR が音声翻訳の研究を開始して約 19 年になる。音声翻訳技術が夢の技術であった時代から、音声翻訳技術の実用化を目指し、ニーズが明確な旅行対話にフォーカスしながら音声翻訳システムの研究開発を進めてきた。現在では、日常旅行会話に対し、1 秒程度の処理時間で音声翻訳が行え、音声翻訳の実用性を議論する技術レベルまで到達した。今後、さらに翻訳サービスを必要

とする観光地などに常時設置して、実証実験、改良を進めていく必要がある。また、実際のサービスに必要な固有名詞の問題もインフラを含めた形で解決法を探る必要がある。

現在のATRの音声翻訳は、文が比較的短く単純な日常の旅行会話を対象にし、文脈を考慮せず逐次翻訳をするシステムである。しかし、いざ、人間の同時通訳者と比較すると、足元にも及ばない。今後の基礎研究の目標として、連続して発話される音声を自動的に適切な単位に区切り理解しながら聞き取りやすい表現で、五月雨式に音声翻訳していく同時通訳のような音声翻訳技術の実現を目指して研究を進めていきたいと考えている。また、音声翻訳だけでなく、情報検索、対話、要約、アーカイビングなどより広い音声言語情報処理にも研究の幅を広げていく予定である。

この場をお借りして、これまでATRの音声翻訳技術の研究に携わってこられた一人一人の研究者、技術者の皆様に感謝したい。また、本プロジェクトに関して、ご指導、ご協力頂きました皆様には、心から感謝し、さらに、ATRの音声言語研究に引き続きご支援をお願いしたい。時代が変わり、基礎研究には厳しい風が吹く昨今ではあるが、こんな時こそ、全員一丸となってこの風に立ち向かう所存である。



●
●
●
●
●

Speech Translation is Growing up: From Concept Demonstration to Lasting Impact



Professor of Computer Science
Director, International Center for Advanced
Communication Technologies (InterACT)
Carnegie Mellon University &
University of Karlsruhe
Alex Waibel

It is indeed a great honor to write an article for ATR Up-to-Date, and it is indeed a great moment in time to do so. For it is just about 15 years, that our joint adventure to tackle a difficult problem between ATR in Japan and our laboratories in Japan, the US and Europe began. Due to ATR's early (1987) leadership, its support, vision and friendship, a new research field could begin. ATR had just begun as a company and speech translation was chosen as one of four key technology areas. To address the difficulties and to consider the many aspects of many languages, the international Consortium for Speech Translation Advanced Research (C-STAR) was formed in 1991. It included just four partners (including ATR and our laboratories at CMU and UKA) at first, but it has since grown to include 20 of the leading laboratories around the world.

The problem of speech translation is enormously difficult due to the combined difficulty of accurate speech recognition, acceptable machine translation and naturally sounding speech synthesis, of which neither can be considered a solved problem by any measure. Researchers in the component fields often derided the early efforts as unmanageable, intractable, impractical, and even not useful (!), given the poor solutions and poor performance that existed in the component technologies at the time. Undeterred, however, those early researchers persisted, first demonstrating feasibility (1991), then capabilities for spontaneous speech (1993-1999) and now practical fieldable solutions. In an age of globalization, where business, humanitarian, healthcare and security needs have rapidly grown beyond national boundaries, the enormous importance, yes the absolute necessity of cross-lingual technologies in every form (text, speech, image), was to be recognized before too long.

Now, 15 years later, virtually all governments and research sponsoring agencies in the developed world, support significant efforts in the area of speech translation. Indeed, perhaps the largest ongoing research programs that still fund speech and language research at all in Europe (TC-STAR, CHIL) and in the US (DARPA-GALE, DARPA-TRANSTAC, NSF-STR-DUST), are now committed to crack the speech translation problem and related cross-lingual language requirements. The rapidly growing need for fast, effective response to multilingual information, the need for effective cross-lingual communication necessitate technical solutions to deliver the required greater speed, broader language coverage and lower cost than what can possibly be made available by human language services alone.

Given 15 year history of speech translation research, and the tremendous effort and investment currently underway, one might ask, if the problem is almost solved, and if not, what challenges remain. In my view there are 4 remaining challenges:

- **Robustness** - Speech Translation must be reliable in all circumstances for which it is to be employed and it must deliver trusted output. How can the output of a technical device be trusted? Unlike humans, machines are woefully inadequate in judging the plausibility of their own output and articulating their own self-doubt. Robustness also remains a challenge, when we consider not only clean speech input, but highly disfluent conversational speech, noisy environments, distant microphones and stressed or emotional speech.
- **Domain Unlimited Capability** - While a number of practical applications can be fielded that require only translation capability in limited domains, the domain restriction of most of today's systems must be removed. This is necessary, if we wish to provide translation of open-domain spoken language tasks such as Broadcast News, Lectures/Seminars, Parliamentary Speeches, Meetings and Telephone Conversations. Domain unlimited speech translation in turn must cope with disfluent, conversational speech *as well as* large open domain language and vocabulary coverage.
- **Language Portability** - Sadly, most current efforts are concentrated around only a few languages of general interest: English, Chinese, Arabic, Spanish, Japanese, German, ... Perhaps the greatest social impact of translation technology, however, could come from capabilities in less commonly spoken languages and language pairs, where language tools and human translation services are less readily available. Short of covering 6,000 languages of the world, however, managing even 20 languages of an expanding Europe, already presents great difficulty and cost. Can more advanced machine learning techniques help to lower the cost of development and language portability?
- **Human Delivery** - For the language barrier to become invisible, we also have to be concerned with appropriate human interfaces that deliver language services in an unobtrusive way. Clearly, spoken input is preferable in mobile situations or meeting situations, but images may require photo or video input, or a mixture of image and voice. How should output be presented? By voice? By text? Should it be delivered via headphones, heads-up displays, speakers? Should it run on a PDA, mobile phone, laptop, or be implanted in a ubiquitous intelligent environment? Numerous intriguing possibilities exist.

Speech translation as a research field has grown up. It has been a privilege to collaborate with ATR for 15 years on the problem of speech translation and help build foundations in a field of growing importance. Looking toward the future, many open challenges remain, but the excitement does not let up: Where else can scientists be offered the dual benefit of scientific fascination with a grand challenge problem and a guaranteed opportunity to change the world for the better? As our efforts have been so fruitful, growing international recognition of the problem also brings growing intense world-wide competition over new systems, solutions and standards. ATR's pioneering benchmarking exercises as launched under IWSLT, provide a mechanism and a forum for these forces and for the best laboratories around the world to advance the state of the art rapidly and jointly. International exchange will continue to refresh and deepen our understanding of the problems and accelerate the turn-around in implementing viable solutions. Our laboratories look forward to continuing and further deepening the strong collaboration and friendship that we have begun with ATR 15 years ago.



音声言語コミュニケーション研究所
音声音響処理研究室、音声言語処理研究室
中村 哲、菊井 玄一郎

1. はじめに

機械による音声の認識は、人類の長年の夢であり、古くから研究開発が行われてきました。音声の認識は非常に複雑な問題で、個人ごとの口や歯などの発話器官の構造の違い、発話器官の動かし方の違い、前後の発話に依存した発話器官の動かし方の違いにより引き起こされる音声信号の違いが、音声認識を困難にしてきました。さらに、複数の単語からなる文発声の認識は非常に難易度の高い課題でした。これらの課題に対し、音声の音の高さの成分分析を精密に行うスペクトル分析法、線形予測分析法が開発され、また、発話毎に長さの異なる発話の照合を行うために、時間構造正規化を含んだパターンマッチング法である動的計画法が開発されました。特に、1970年後半に開発された隠れマルコフモデル（HMM：Hidden Markov Model）は、時間構造の揺らぎと周波数的な音声の特徴の揺らぎを統計的に表現でき、データに基づく学習と認識を一貫して行える枠組みであるため、今日の音声認識の基本手法となっています。本稿では、これらの手法をさらに発展させた実環境における多言語連続音声認識技術の研究成果について述べます。

2. 音声認識のモデル

図1に、音声認識部のブロック図を示します。音声認識は、ある単語列を仮定したときの音声の確率（声 | 単語列）（音響モデルと呼ぶ）とその単語列の確率（単語列）（言語モデルと呼ぶ）の積を最大化する単語列を見つける問題と考えることが出来ます。

語列の確率（単語列）（言語モデルと呼ぶ）の積を最大化する単語列を見つける問題と考えることが出来ます。音声の特徴は、10ミリ秒周期、20ミリ秒窓の短時間の周波数スペクトル分析により抽出されます。この特徴を元に、隠れマルコフモデルを用いて、単語や音素の音のモデルを構成します。また、単語列のモデルには、単語の連鎖出現確率を表すN-gramと呼ばれるモデルが用いられ、このモデルは対象分野の大量のテキストから学習が行われます。

3. 音響モデル

当研究所では、統計的モデルによる音声認識の研究にいち早く取り組み、主として隠れマルコフモデルの研究を行ってきました。このモデルは、通常、3つの状態を有し、状態遷移しながら観測された音声の特徴ベクトルの確率を計算する方法です。各状態では、複数の正規分布の重み付き和により複雑な音声の特徴量分布を表現します。当研究所では、特に音声の特徴を前後の音素の関係を考慮して最適に表現する音響モデル^[1]、さらに、HMMの学習データ量に応じて音響モデルの構造を自動的に変化させ、最適な状態数を割り当てる最小記述長による音響モデル構築法を開発しました^[2]。これにより、学習データ量が変わっても、自動的に最適な音響モデルを構築することが出来るようになりました。この方法を用いて、現在、日本語については地域を考慮した400人発話者、英語に

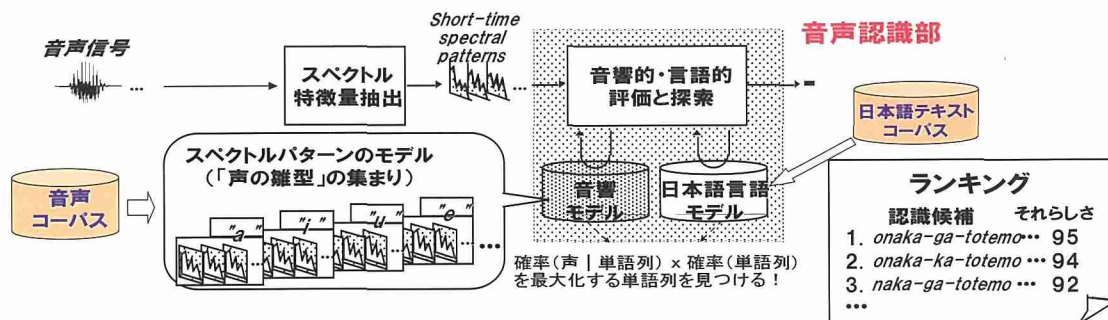


図1 音声認識処理図

については約570人発話者、中国語については約540人音声コーパスを用い、音響モデルを構築しています。

4. 言語モデル

言語モデルについては、現在の単語が生起する確率を過去N個の単語の連鎖確率により表すN-gramと呼ばれる手法に注目して研究を行ってきました。当研究所では、この方法を発展させ、単語でなく品詞などの単語クラスを単位として確率を計算するクラス言語モデル、前後の文脈を分離し前方文脈と後方文脈とを別々に考慮する多重クラスN-gram、言語モデルの単位を可変長にする複合N-gramを開発してきました^[3]。また、地名、人名などの辞書に登録されていない単語については、音節を用いた階層型言語モデルを用いることで認識を行う手法を開発しました^[4]。現在、言語モデルの学習には、旅行文コーパス（BTEC：Basic Travel Expression Corpus）の学習セット約60万文と疑似旅行対話文を使用しています。クラス言語モデルのクラス数は8000ないし20000としています。

5. 音声認識性能

音声認識性能をBTEC旅行文コーパスの日本語、英語、中国語510文をテストセットに選び評価致しました。音声認識実験では、各言語40人より発話されたものを評価データとしました。認識語彙は約35000語で、発音の変形も入れると約50000語です。音響モデルとしては、最小記述長による音響モデル設計法を用いました。言語モデルには、前後の文脈を分離し前方文脈と後方文脈とを別々に考慮する多重クラス複合2-gramを用いました（MCC 2gram）。従来の単語2-gram（Word 2gram）と比較した場合の認識結果を表1に示します。テスト文の複雑度、つまり、現在の単語に続く平均の単語数を表すテストセットパープレキシティは、単語2-gram、単語3-gram、前後の文脈を分離し前方文脈と後方文脈とを別々に考慮する複合2-gramに対し、30.6,17.5,24.8でした。

6. 雑音抑圧処理

音声認識を実際の音響環境で使用する際には、さらに音響的な外乱への対策が必要となります。音響的な外乱には、対象話者以外の話者の音声、雑音などが存在します。当研究所では、図2に示す携帯情報端末用の8個のマイクロフォン素子を取付けた専用マイクロフォンアレー信号処理ユニットを新たに開発し、指向性制御、雑音抑圧により、雑音下における音声認識を実現しています^[5]。このシステムでは、まず、8チャンネルの信号をA/D変

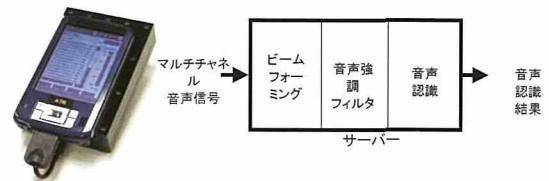


図2 マイクロフォンアレーを備えた携帯情報端末と雑音抑圧処理

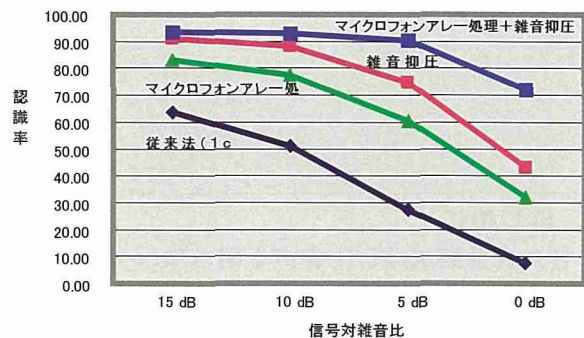


図3 マイクロフォンアレー処理と雑音抑圧による認識率改善

換した後、ワイヤレス LAN を介してサーバーに送ります。処理は、雑音源からの信号を抑圧し、目的方向の音を強調する一般化サイドローブキャンセラと呼ばれるマイクロフォンアレー信号処理と、混合ガウスモデルに基づく雑音除去フィルタの従属接続の2段構成になっています。図3に1ch使用時、マイクロフォンアレー処理時、雑音抑圧フィルタ処理時、両手法使用時の認識結果を示します。

7. まとめ

音声翻訳の要素技術の一つである多言語音声認識部に関して述べました。音声翻訳の対象にしている旅行会話の認識性能として現在、日本語94.6%、英語89.1%、中国語87.3%の単語認識率が得られました。また、マイクロフォンアレーと雑音除去フィルタにより雑音のある環境でも動作するシステムを構築しました。しかし、実環境での性能向上、安定化のためには、さらに、実環境での実証実験を行い、データを収録し、そのデータを使用して各モジュールを改良してゆく必要があると考えています。

参考文献

- [1] Ostendorf, et al. Computer Speech and Language, vol.11, pp.17-41, 1997.
- [2] Jitsuhiro, et al IEICE Trans. Inf. & Syst., vol. E88-D, no. 3, pp.391-400, 2005.
- [3] Yamamoto, et al, Speech Communication, Vol.41-2003, pp.369-379, 2003.
- [4] 谷垣, 山本, 匂坂, 信学論, Vol.J84-DII, No.11, pp.2371-2378, 2001.
- [5] Herbordt, et al, 秋季音講論, 3-7-20, pp. 141-142, 2005.

表1 日本語、英語、中国語単語音声認識率

日本語		英語		中国語
Word 2-gram	MCC 2-gram	Word 2-gram	MCC 2-gram	MCC 2-gram
93.8	94.6	89.1	91.6	87.3

音声翻訳処理の全体を統合する技術



音声言語コミュニケーション研究所
音声言語処理研究室、統合システム研究室
菊井 玄一郎、清水 徹

1. はじめに

音声自動翻訳は、①音声を取り取って、②別の言語に翻訳し、③音声にして出力する、という3つの要素技術から出来ています。これら3つは音声翻訳にとっては「部品」ですが、従来からそれぞれ独立した技術として研究開発されており、すでに音声を取り取って文字にする音声認識ソフトやWEB ページを翻訳するソフトなどが市販されています。

では、これらのソフトを買ってきてつなぐと音声翻訳ができるのでしょうか？

答えはNOです。例えば、通常の自動翻訳ソフトは正しい文が入力として与えられることを前提として設計されているため、音声認識結果が誤っていても無理やり正しいものと解釈して翻訳しようとします。その結果しばしばとんでもない誤訳が起きます。これを避けるには、例えば、認識結果の信頼性が低い場合には翻訳せずに処理を中断して再度入力を求める、といった処理が有効です。

また、通常の音声認識では認識結果を一つに絞り込んで出力しますが、この絞り込みの誤りを避けるため、複数の可能性を残したまま翻訳まで試み、最終結果をみて1つに絞り込むと性能が上がるかも知れません。

本稿では、特に、音声認識と自動翻訳の統合に焦点をあて、認識結果の信頼性を評価する技術、および、音声認識と翻訳を組み合わせる最も正しいような結果を出力する技術を紹介します。さらに、これらを含むシステム全体を設計・評価するための前提となるデータ収集について紹介します。

2. 音声認識結果の信頼性を評価する技術

音声認識結果の正誤を自動的に検知しユーザに誤った結果を提示しない技術のリジェクションと言います。音声認識では、入力された音声と音響モデル（音声の特徴を表すモデル）・言語モデル

（言葉のつながりを表すモデル）・辞書との照合が行われ、照合の結果、他と比べて最も尤度（ゆうど）が高い単語列を認識結果として出力します。尤度とは統計的観点から見た尤もらしさの度合いで、尤度が高い単語は正しく、尤度が低い単語は誤っている傾向があります。この尤度の値（絶対値）をリジェクションに用いることができるでしょうか？実は、尤度に基づいて正誤を判定するのは一般的には難しいと考えられています。たとえば、音響モデルの尤度は、母音・子音などの音の違い、話者、話し方、周囲の雑音によって大きく影響を受け、結果が正しいのに尤度が低い、あるいは、誤っているのに尤度が高いということが起こるからです。

そこで私たちは、音声認識の過程で生成される単語グラフから単語の信頼度を出す方法を考案しました。単語グラフとは、入力音声と音響モデル・言語モデル・辞書との照合結果をネットワーク状に表現したものです。図1が簡略化された単語グラフの例で、左端のSと右端のEを結ぶ1つの経路が1つの認識結果候補に対応します。

ある単語の信頼度はこの単語を通る経路の尤度が単語グラフ内の全経路（上記の例では6つ）の尤度の総和に対してどの程度の割合であるか（事後確率）という値で定義します¹⁾。正しい単語であれば対立候補が現れにくいので事後確率は高く、誤った単語であれば事後確率が小さくなるので、単語のリジェクションのための有効な指標になり

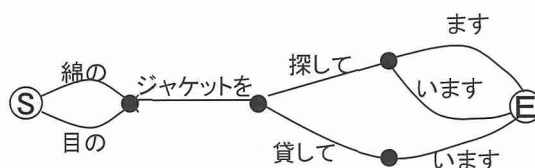


図1 単語グラフ（簡略化版）の例

ます。この事後確率は Generalized Word Posterior Probability (GWPP：一般化単語事後確率) と呼ばれています。

3. 複数の認識候補からの最適翻訳

上で述べたように音声認識処理の内部では、複数の認識結果候補に対して音響的・言語的観点から尤度を計算し、この尤度の尤も高いものを1つ選択しています。しかし、この尤度には様々な要因で誤差が含まれており、2位や3位のものが実は正解ということがしばしば起こります。旅行会話表現集の読み上げ音声を認識した結果によると、上位10位までの認識候補に正解の含まれる率は上位1位だけに比べて11%も向上します。

そこで、我々は音声認識候補の2位以下を捨ててしまうのではなく、複数個（たとえば上位10個）の認識結果に対して全て自動翻訳を試みることを考えました。翻訳までやってみてから音声認識時に得られる尤度と翻訳時に得られる尤度を組み合わせた「統合スコア」を使って最終的に一つの結果を選ぶというわけです（図2）。

ここで重要な「統合スコア」は、音声認識結果に対する音響的・言語的尤度、自動翻訳結果に対する訳語の対応の良さをあらわす尤度、翻訳結果の言語的な正しさを表す尤度など、様々な尤度に重みを付けて加えたものです（正確にはlog-linear modelというものを使います）。この重みは、入力音声データとこれに対する正解翻訳結果の組を大量に（数百以上）用意して、正解候補の統合スコアが極力高くなるように自動調整します^[2]。

この方法で認識結果の上位20位を使うことで、1位のみ使った場合に比べて最終的な翻訳結果の誤りを4%以上削減させることができました。なお、スコアの近い認識結果同士は一部しか変わらないのが普通です。そこで、複数の認識結果を別々に翻訳するのではなく、共通部分をまとめた図1の単語グラフのような形式のままで翻訳することによって高速化を図っています。

4. 音声翻訳実験とデータ収集

音声翻訳全体の研究を実証的に進めるために、我々は日本語と外国語（英語や中国語）のネイテ

ィブを被験者として、様々な条件下で音声翻訳システムを使った対話実験を行い、データを収集してきました。このデータは先に述べたリジェクションの研究などに使えるだけでなく、システムに対する利用者の発話の性質を解明する上でも有用です。

たとえば、翻訳システムを通した対話の言語的な特徴（言語表現の頻度などの統計量）は、「人間の通訳を介した対話」、および、「旅行会話集のような表現」の両方の特徴を含む（合成したもの）であること、音声翻訳システム利用法の説明によって入力音声の言語的、音響的特徴が体系的に変化すること、などが明らかになってきました^[3]。

前者の結果は音声認識や自動翻訳のパラメータを決定する際に参照する言語データを選ぶ手になります。後者の結果は、音声翻訳システムの利用者に対して、使い易くかつ音声翻訳性能が向上するようなガイダンスを設計するのに直接利用できます。

さらに、2004年の年末には、関西国際空港等で我々の音声翻訳システムを外国人旅行者に実際に使ってもらってデータ収集を行いました。アンケートによれば英語話者の50%、中国語話者の53%が自分の言いたいことがほぼ伝わった、という評価をしています。

5. まとめ

本稿では音声自動翻訳のシステム全体の統合にかかわる研究の一端をご紹介しました。これらの研究は多岐にわたりますが、いずれも要素技術の研究だけでは出てこないものです。その内容も、要素技術を単にうまく組み合わせることだけに留まらず、個々の要素技術の設計や最適化までも含みます。これらは要素技術からシステムの利用実験まで一体的に研究しているATRだからこそできるユニークな研究分野といえます。

参考文献

- [1] Frank K. Soong et al., "Optimal Acoustic and Language Model Weights for Minimizing Word Verification Errors", Proc. of ICSLP 2004.
- [2] Ruiqiang Zhang et al., "A unified approach in speech-to-speech translation: Integrating features of speech recognition and machine translation," Proc. of Coling 2004.
- [3] Genichiro Kikui et al., "Multi-lingual Corpora for Speech-to-speech Translation", Proc. of ICSLP 2004.

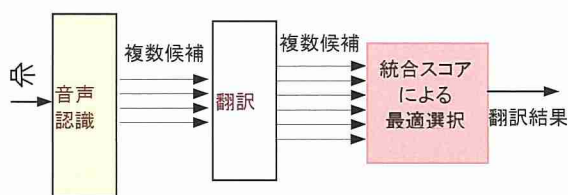
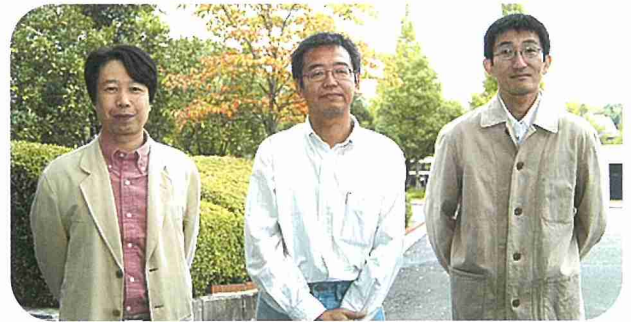


図2 複数の認識候補を用いた音声翻訳

コーパスに基づくマルチエンジン 機械翻訳技術



音声言語コミュニケーション研究所
自然言語処理研究室

佐々木 裕、隅田 英一郎、今村 賢治

1. はじめに

音声翻訳における機械翻訳の役割は、音声認識により出力される単語列を目的の言語に翻訳し、単語列として音声合成に渡すことです。ここで、あえて「文」を目的の言語に翻訳すると書かないで、単語列を翻訳すると表現している理由は、対象が新聞記事のように明確な文の単位になっていないため、1発話という単位で翻訳の対象を考える必要があるためです。

この点も含めて、音声翻訳における機械翻訳技術が通常のテキスト翻訳と異なる点は以下の通りです。

- 入力の話言葉であるため、口語的な表現が頻繁に使われます。
- 翻訳の単位が文頭～句点（。）といった単純な文ではなく、発話が翻訳の単位になります。
- 音声認識結果であるため、大文字・小文字の区別がなく、句読点やかぎ括弧等の記号が含まれていません。
- 英字や数字は、読み方に合わせてスペルアウトされています（たとえば、ATRはエーティーアールと表現されます）。
- 音声認識誤りや発話者の言い間違いを含めて、文法的に間違っている文も翻訳する必要があります。

特に、最後の項目は、機械翻訳システムの設計に大きな影響を与えます。入力として、文法的に正しい文を想定して、翻訳規則を手で作成することは可能ですが、どのような文法的な誤りが含まれるかわからない状態で、様々な場合を想定しながら、網羅的な翻訳規則を手で作成することは非常に困難です。

そこで、旅行先での会話を対象にした対話表現とその翻訳例を大量に収集し、これに基づいて、誤りに強い機械翻訳システムを自動作成するアプローチを採用しました。このような大量の対訳文

集は、対訳コーパス（音声言語データベース）と呼ばれます。

ATRでは、旅行会話に関して日英約100万文対、日中約50万文の対訳コーパスの構築しました。

2. マルチエンジン機械翻訳技術

翻訳の対象である旅行会話文とは、海外旅行中に使われる比較的短い文のことです。たとえば、「窓を開けてもいいですか」という内容を述べている英語としては、以下のようなものが考えられます。

1. may i open the window
2. ok if i open the window
3. can i open the window
4. could we crack the window
5. is it okay if i open the window
6. would you mind if i opened the window
7. is it okay to open the window
8. do you mind if i open the window
9. would it be all right to open the window
10. i'd like to open the window

これらの文は、音声認識結果を表現していますので、記号や大文字・小文字の区別がなくなっています。さらに、実際には、音声認識誤りによる、単語の追加、削除、置換が発生します。

コーパスベース・アプローチでは、人間が翻訳規則を入力するのではなく、コーパスに含まれる大量の対訳文を分析し、翻訳方法を計算機が自動的に学習するというアプローチを採用しています。ATRの機械翻訳システムの特徴は、セレクト^[1]と呼ばれる選択器の下に複数の翻訳エンジンを配置したマルチエンジン構成をとっている点にあります（図1）。翻訳エンジンには、統計翻訳エンジンSAT^[2]と用例翻訳エンジンHPATR^[3]の2種類の方式の異なるエンジンが採用されています。

以下、それぞれの手法を紹介します。

2.1 用例翻訳エンジン

用例翻訳エンジン HPATR は、原文および訳文の構文を考慮しながら、精度の高い置き換えを繰り返し行なうことにより翻訳を実現します。置き換えに使えるような構文の部品は、対訳文例集からあらかじめ自動的に収集しておきます。HPATRでは、生成される文の品質は高いのですが、文法による制約が強いため、翻訳結果を生成できないケースが存在します。

2.2 統計翻訳エンジン

統計翻訳エンジン SAT (Statistical ATR Translator) は、対訳文の間の統計的な単語列の対応・変換関係（モデル）をあらかじめ分析しておき、その統計情報に基づいて入力に対する翻訳結果を生成します。

このため、入力に対してロバスト（頑健）で、ほとんどの場合、なんらかの翻訳結果を出力しますが、原文にない不要な語を付加してしまったり、局所的に原文と全く異なる訳語を出してしまうことがあります。

2.3 セレクタ

選択には統計翻訳のモデルを利用します。ここで、統計翻訳のモデルを使うのであれば、統計翻訳の結果が常に選ばれるのではないかという疑問が沸きますが、翻訳文の生成（デコード）は、メモリと処理時間の制約によって、統計的翻訳モデルによって規定された膨大な探索空間の一部しか探索できていないため、準最適解を求めているに過ぎません。このため、全く異なるアプローチで翻訳文を作成した用例翻訳エンジンの出力が選択される場合が存在するのです。

このようにして特性の異なる翻訳エンジンによる翻訳結果の中から一番尤もらしい結果を選択することにより、全体の翻訳精度を単体の翻訳エンジンよりも平均5%～15%向上させることができます。

3. おわりに

マルチエンジン翻訳方式は、高品質の音声翻訳の実現に適したアプローチです。

今後の課題としては、各翻訳エンジンの個々の性能を向上させるとともに、多様な翻訳エンジンの出力から、品質の良い出力をよりうまく選択する方法を見つけることがあげられます。各翻訳エンジンの性能向上のためには、さらに深い統語解析や意味解析の情報を活用する必要があり、そのためには大規模コーパスやオントロジといった言語資源を充実させることが急務となっています。

参考文献

- [1] 秋葉泰弘, 渡辺太郎, 隅田英一郎, SMT 指標を用いた複数翻訳エンジンからの最適訳選択, 情報処理学会研究報告「自然言語処理」研究会, No. NL-150, pp.63-68 (2002)
- [2] Watanabe, Taro; Sumita, Eiichiro. Example-based Decoding for Statistical Machine Translation. In Machine Translation Summit IX, pp. 410-417, New Orleans, Louisiana (2003)
- [3] 今村賢治, 大熊英男, 渡辺太郎, 隅田英一郎, 統計翻訳指標を導入した構文トランスファに基づく用例翻訳, 情報処理学会研究会報告, 2004-NL-162, pp.71-77 (2004)

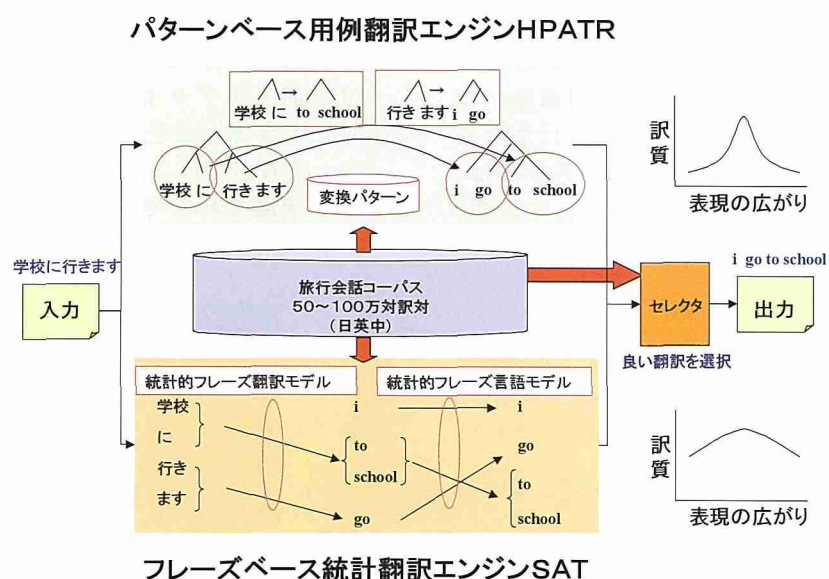
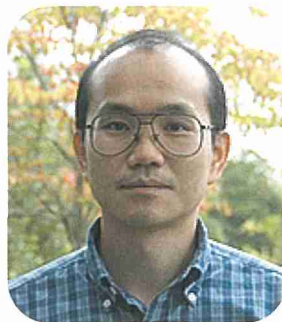


図1 マルチエンジン翻訳システムの概要

コーパス・ベース多言語音声合成システム XIMERA



音声言語コミュニケーション研究所
音声音響処理研究室
平井 俊男、河井 恒

あらまし

ここでは、まず音声合成のしくみと高品質な合成音を得ることが可能なコーパス・ベース手法について説明し、次に本研究所で開発されたコーパス・ベース音声合成システム XIMERA（キメラ）、システムで用いられる音声コーパス、ならびにシステムの多言語対応などについて述べます。

1. はじめに

人間にとって最も基本的な情報伝達手段の一つである音声言語には、一度習得してしまえば特別な訓練なしに利用でき、手や目を使う別の作業をしていても使える、という利点があります。これらの利点から、人間と機械との情報伝達手段としても古くから注目されてきました。音声翻訳システムに代表される音声対話システムでの音声出力や、機械から人間への情報伝達に音声言語を用いる際には、テキスト情報を音声情報に変換する「音声合成技術」が用いられます（図1）。

近年、新しい音声合成の方式として、コーパス・ベース音声合成が急速に広まりつつあります。コーパス・ベース音声合成とは 1980 年代から ATR で研究開発されてきた方式で、最近の計算機パワーの飛躍的な向上により実現できるようになった方式です。コーパス・ベース方式での処理手順は、人間の生の声をあらかじめ大量に録音しておき、

入力されるテキストに応じて、録音された音声の中から適切な音声の断片（音声素片。可変長で最短は数十ミリ秒程度）を選び出して（素片選択）滑らかに繋ぎ合わせ、合成音声として出力する、というものです。ここでは最初に音声合成のしくみを説明し、続いて音声言語コミュニケーション研究所で研究開発されたコーパス・ベース多言語音声合成システム XIMERA（キメラ）について述べます。

2. テキストを音声に変えるしくみ

音声合成のしくみは図2のようになっています。入力テキストで、対応する音声合成音声として出力されます。それぞれの処理を以下に示します：

- (1) 入力テキストを形態素（意味を持つ最小の言語単位）に分解します。
- (2) 形態素ごとに読みとアクセントを付けます。複数の形態素が連なって複合語となる時には読みもしくはアクセントが変わることがありますので、変形規則に従ってこれらを変換します。
- (3) 読みとアクセントの情報から、声の高さ、大きさ、音色をどのように変化させれば自然な音声になるかを予測します。
- (4) 予測された特徴の変化パターンに近い音声



「約10キロ・メートル先に故障車があり渋滞中です。回避しますか？」



「5メートルの津波の恐れがあります。至急非難してください。」

図1 音声合成の利用例

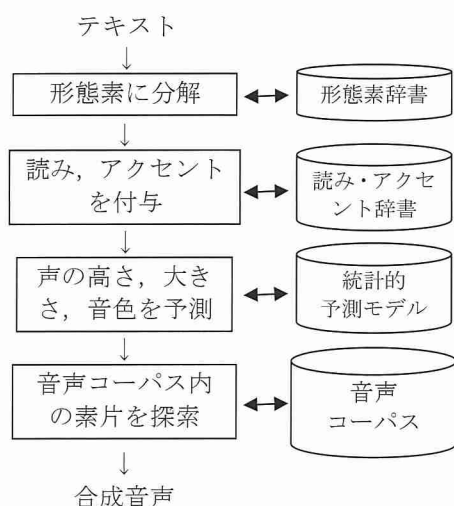


図2 テキストを音声に変えるしくみ

素片を音声コーパス内で探索します。その結果得られる素片列を接続して合成音声を得ます。探索の時には、接続される素片の接続点でのスムーズさなども考慮されます。

3. コーパス・ベース音声合成^[1]

最初にも述べた通り、ATRにおけるコーパス・ベース音声合成の研究は20年ほど前からスタートしました。しかしながら、当時の計算機パワーの限界により、(1) コーパスのサイズを大きくすることが困難、(2) 声の高さなどを推定するためのモデルの学習データを増やすことが困難、などの問題があり、理論的には高い品質の合成音声を得られるはずでしたが、十分な自然性を持つ合成音の生成は困難でした。

これに対して、近年のCPUパワーの飛躍的向上や記憶装置の容量増大、ならびに大量のテキスト・データ、音声データの広範な収集を前提とした統計的なモデルの学習が可能となったことにより、コーパス・ベース音声合成を実現するための環境が整い、結果として高品質の合成音声を得られる大規模コーパス・ベース日本語音声合成システムXIMERAの構築が可能となりました。XIMERAでは男性音声約100時間、女性音声約50時間という、世界的に見ても最大規模の音声コーパスが用いられています。

4. 中国語と英語の音声合成^{[2], [3]}

コーパス・ベース手法の特徴として、異なる言語への展開が容易、ということも挙げられます。このことを確認するため、XIMERAを他の言語(中国語ならびに英語)に展開することが試みられています。それぞれ女性音声約20時間(中国語)、男性音声約15時間(アメリカ英語)の収録が完了しており、中国語については既にシステム構築も

完了しています。英語合成システムは現在構築中で今年度中にプロトタイプが完成する予定です。

5. 処理スピードの向上^[4]

音声対話システムでは、対話の流れをスムーズにするため、システムのレスポンスをできる限り速くすることが求められます。大きなコーパスを擁するコーパス・ベース音声合成では、その処理のほとんどが素片選択に費やされているため、処理スピードの向上には、素片選択処理の速度向上が重要となります。そのためには、探索するコーパス内のデータの範囲を絞る方法が通常採られますが、これによりスムーズに接続できる素片列を見つけにくくなり、結果的に合成音声の品質劣化に繋がります。

XIMERAではこの問題を解決するため、ある程度探索が進んだ時点で選択素片を確定していくと同時に、確定された素片を音声データとして逐次出力する、というアプローチを採用しています。

6. まとめ

ここでは、まず音声合成のしくみについて簡単に説明した後で、音声言語コミュニケーション研究所で研究開発されたコーパス・ベース多言語音声合成システムXIMERAについて、特に多言語音声合成の試みと処理能力向上について述べました。コーパス・ベース方式を採用することにより非常に高品質の合成音を得られるようになりましたが、自然音声の示す自然性には到達できていないのが実情です。また、言語的な情報だけではなく意図や感情も合成音で伝達したいという要求も強まっています。ATRではこのような課題に取り組んでいきます。

参考文献

- [1] 平井ら, 「コーパス・ベース音声合成システムXIMERA」, 電子情報通信学会技術研究報告 SP2005-18 (2005, 5)
- [2] 倪ら, 「ATRにおける中国語テキスト音声変換システム」, 日本音響学会 2005年 春季研究発表会講演論文集, pp.287-288 (2005, 3)
- [3] 倪晋富, 平井俊男, 河井恒, 「英語音声合成のためのコーパス設計」, 日本音響学会 2005年 秋季研究発表会講演論文集, pp.385-386 (2005, 9)
- [4] 西澤信行, 河井恒, 「短遅延音声合成のための素片選択法」, 電子情報通信学会技術研究報告 SP2004-48 (2004, 8)



ATR 脳活動イメージングセンタ利用の音声研究とその成果展開

脳活動イメージングセンタ 正木 信夫

ATR 脳活動イメージングセンタのMRI装置は、主に脳の機能を調べる研究に利用されています。しかし、MRIは本来、外から見るできない体内の臓器の形や組織を3次的に視覚化する装置です。今回はこの「本来の機能」を音声研究に用いた取り組み、およびその成果展開について紹介します。

本号の特集で解説されているように、ATRでは音声認識・音声合成をはじめとする音声の応用技術を研究開発しています。それらの基礎になるのは音響学的な分析処理技術です。しかし現在、その技術の中に音声が生み出される仕組みが十分活かされているとは言えません。それは、この仕組みが喉から口や鼻にかけての外から見えない部分にあるために、その形や動きの観測が難しいことが一つの原因です。

そこで、脳活動イメージングセンタは、人間情報科学研究所の生体イメージング研究室を支援して、MRIの機能を活かした音声生成の研究を推進しています。口腔顔面領域の微細な構造を捉えることのできる高感度コイルの開発（図1）、発話器官の運動をダイナミックに再現できる3次元動画撮像技術の確立はその例です。3次元MRIデータに基づく声道の実体模型（図2）の音響測定実験からは、話者毎の声の特徴を決定づける機構が明らかとなるなど、その成果は音声の応用研究の分野からも注目されています。

これらの成果は学术论文や学会発表として公表するばかりではなく、母音生成時の発話器官のMRI画像データ（図3）の無料提供（教育・研究目的に限る）などの形で展開しています。今後もMRI技術を用いた音声研究への貢献と、成果の展開を継続的に実施していきます。

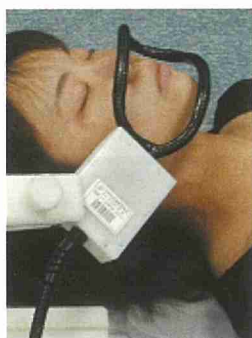


図1 口腔顔面用高感度コイル
唇や喉頭を高精度で撮像可能



図2 声道実体模型
母音生成の教育・研究用に販売予定



図3 母音発話MRIデータ
教育・研究用に無料提供中

脳活動イメージングセンタ Q & A コーナー

Q：脳活動イメージングセンタの利用を検討していますが、被験者はどのように手配するのでしょうか？

A：被験者の手配は実験者側でお願いしております。しかし、実験者の所属機関が遠方で、被験者を同行することが難しい場合があります。そのような場合、人材派遣会社を通して被験者を募集する方法等を提案させていただいております。実績ある会社の紹介もいたします。

このページに記載された内容に関するお問い合わせは、ATR 脳活動イメージングセンタ（担当：中川）まで、電話やE-mailでお問い合わせください。お待ちしております。

TEL: 0774-95-2620, FAX: 0774-95-1281

E-mail: info@baic.jp, URL: <http://www.baic.jp/>

\$ 注目特許 \$

経営統括部 企画担当
荒木 晃司

日本特許第 3672770 号「アレーアンテナ装置」

出願日：1999 年 7 月 8 日

発明者：大平 孝、行田 弘一

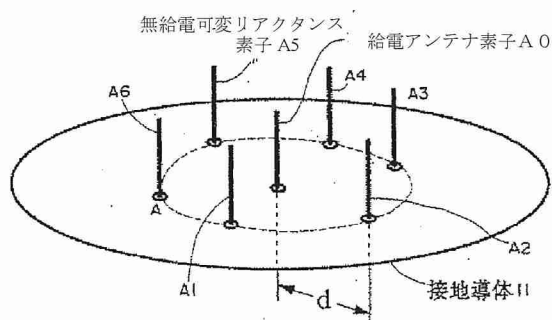
[発明の概要]

この発明は、電氣的に指向性が制御できるアレーアンテナ装置に関するもので、接地導体 11 の中央部にこの接地導体 11 に電氣的に絶縁された状態で設けられた 1 本の給電アンテナ素子 A0 と、その周囲に $\lambda/4$ の間隔を置いて接地導体 11 に電氣的に絶縁された状態で配置された 6 本の無給電アンテナ素子 A1 ～ A6 と、を主構成要素とし、無給電アンテナ素子 A1 ～ A6 の夫々は、接地導体 11 の裏面側で可変リアクタンス素子を介して接地導体 11 に高周波的に接地されており、この可変リアクタンス素子に可変直流電源から逆バイアス電圧を印加することによって、アレーアンテナ装置の指向性を制御している。

[期待される効果]

フェーズドアレーアンテナ装置は所望の放射パターンが得られる高機能なアンテナとして指向性が要求される分野で多用されているが、各アンテナ素子のそれぞれに低雑音増幅器や可変位相器を対応付けて設けなければならず、構成が複雑で製造コストが嵩む問題を抱えていた。これに対して本発明によれば、構成が極めて簡単で、製造コストが軽減され、しかも指向性の制御を容易に行うことができる。

第 1 の実施形態



カナダ特許第 2341172 号「舌圧測定システム」

出願日：1998 年 8 月 19 日

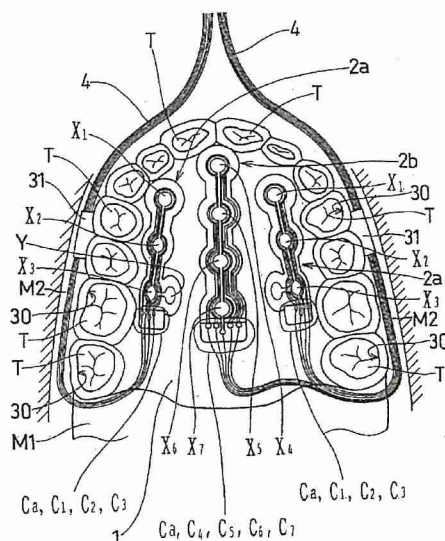
発明者：和久本 雅彦、正木 信夫、大上 壽一

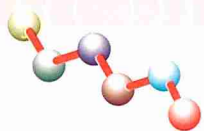
[発明の概要]

この発明は、発話時における舌と口蓋との接触圧を測定するシステムであって、口蓋の形状に合わせて形成された薄い口蓋床板 1 は各歯 T に孔 30 を嵌め込むことによって口蓋の上部に固定されており、この口蓋床板 1 の中央部には圧力センサシート 2a、2b が貼り付けられていて、発話時にこれらのセンサシート 2a、2b に設けた圧力検出セル X1 ～ X7 が検出する舌の接触圧を舌圧信号としてケーブル 4 を介して口蓋外に導出している。

[期待される効果]

本発明によれば、自然な発話を妨げることなくリアルタイムで発話時の口蓋への舌の接触圧を測定することができるので、音声科学の研究に極めて有用である上に、病気や事故で舌に欠損を持つ構音障害者のリハビリ訓練などに応用することができ、また発話時に限ることなく、液体や固形物を飲み込むときの舌の動きや食物の動きの検出にも利用できる。





ATR グループ企業情報

株式会社 ATR-Waves

平成 17 年 8 月 1 日 ATR ビル内に ATR-Promotions100 % 出資の株式会社 ATR-Waves（代表取締役社長 太郎丸眞）が、ATR-Robotics に続く一連のベンチャー企業の第二号として設立されました。＜URL:<http://www.atr-waves.com>＞

ATR-Waves は ATR 波動工学研究所の指向性制御アンテナ技術を基に商品化開発を行い TV アンテナを販売しています。手軽に移動でき、薄型テレビに最適な「地上波デジタル用室内アンテナ」です（ブラスター内蔵）。また、リモコンスイッチひとつで映りが良くなる方向にアンテナの向きを電子的に切り替えられる、指向性制御機能付き平面アンテナです。このほか、アンテナ等波動技術の商品化と ATR 技術の一般市場への普及を図るとともに、研修、セミナー、コンサルティング、技術普及に関する企画、研究支援等幅広く活動しています。



薄型テレビにマッチする平面タイプの室内アンテナ

株式会社 ATR-Lang

平成 17 年 8 月 1 日 ATR ビル内に ATR-Promotions100 % 出資の株式会社 ATR-Lang（代表取締役社長 袋谷丈夫）が、ATR-Robotics に続く一連のベンチャー企業の第三号として設立されました。＜URL:<http://www.atr-lang.com>＞

ATR-Lang は ATR 音声言語コミュニケーション研究所の音声翻訳技術と関連する技術である、音声認識、音声合成、機械翻訳、ならびに音響信号処理等をベースにした製品サービス、及び関連するコーパス等データベースの商品化と販売を行っています。また、ATR 技術の一般市場への普及を図るとともに、ATR-waves と同様に研修、セミナー、コンサルティング、技術普及に関する企画、研究支援等幅広く活動しています。

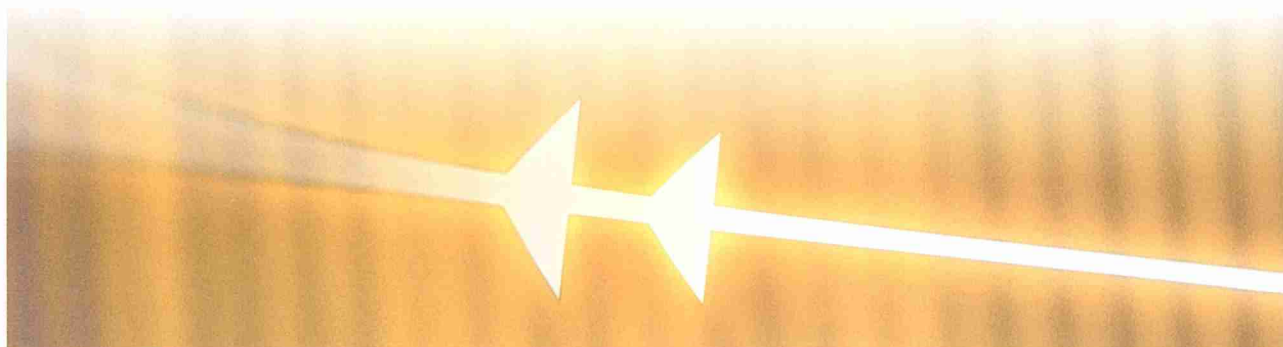


音声翻訳技術の利用例

所員往来

平成 17 年 4 月 1 日から平成 17 年 9 月 30 日までの所員往来

採用年月日	ATR 所属	氏 名
2005. 4. 1	経営統括部 総務 担当課長	北坂 一三
2005. 4. 1	音声言語コミュニケーション研究所 音声音響処理研究室 専任研究員	榎本 成悟
2005. 4. 1	音声言語コミュニケーション研究所 音声音響処理研究室 専任研究員	川本 真一
2005. 4. 1	適応コミュニケーション研究所 自律システム研究室 室長	山口 明
2005. 4. 1	波動工学研究所 電波研究室 研究技術員	天野 信之
2005. 4. 1	波動工学研究所 非線形科学研究室 専任研究員	中榮 穰
2005. 4. 1	波動工学研究所 非線形科学研究室 客員研究員	崔 武漢
2005. 4. 1	波動工学研究所 電波研究室 研究技術員	山本 真由美
2005. 4. 1	メディア情報科学研究所 インタラクシオンメディア研究室 主幹研究員	鳥山 朋二
2005. 4. 1	メディア情報科学研究所 インタラクシオンメディア研究室 専任研究員	岡田 昌也
2005. 4. 1	メディア情報科学研究所 感性データベース研究室 室長	井ノ上 直己
2005. 4. 1	知能ロボティクス研究所 コミュニケーションロボット研究室 専任研究員	吉川 雄一郎
2005. 4. 1	知能ロボティクス研究所 コミュニケーションロボット研究室 専任研究員	小泉 智史
2005. 4. 1	知能ロボティクス研究所 ネットワークロボット研究室 主任研究員	西尾 修一
2005. 4. 1	知能ロボティクス研究所 ロボットサービス研究室 専任研究員	MOLDOVAN, Daniel
2005. 4. 1	知能ロボティクス研究所 ロボットサービス研究室 専任研究員	佐藤 哲
2005. 4. 1	脳情報研究所 認知神経科学研究室 研究技術員	木村 真弘
2005. 4. 1	脳情報研究所 認知神経科学研究室 研修研究員	樋口 さとみ
2005. 4. 1	脳情報研究所 計算神経生物学研究室 契約社員	野中 晶子
2005. 4. 15	音声言語コミュニケーション研究所 音声言語処理研究室 専任研究技術員	王 海波
2005. 7. 1	経営統括部 経営企画 担当課長	松浦 徹
2005. 9. 1	知能ロボティクス研究所 ロボットサービス研究室 専任研究員	VERGES, LLAHI Jaume





退職年月日	氏 名	ATR 所属
2005. 4. 14	LIU, Min	音声言語コミュニケーション研究所
2005. 4. 30	亀山 善文	経営統括部
2005. 5. 31	LO, Wai Kit	音声言語コミュニケーション研究所
2005. 6. 21	二谷 明	経営統括部
2005. 6. 30	西川 智	経営統括部
2005. 6. 30	渡辺 太郎	音声言語コミュニケーション研究所
2005. 7. 12	GOETZ, Melanie Beth	音声言語コミュニケーション研究所
2005. 7. 19	HOULET, Lionel Fabrice	波動工学研究所
2005. 7. 31	NETHAJI, Dharmarasu	波動工学研究所
2005. 8. 5	LEVEILLE, Jasmin	脳情報研究所
2005. 8. 31	斎藤 信雄	波動工学研究所
2005. 8. 31	片桐 恭弘	メディア情報科学研究所
2005. 8. 31	鮫島 和行	脳情報研究所
2005. 9. 30	山幡 章司	経営統括部
2005. 9. 30	澤谷 琢磨	波動工学研究所
2005. 9. 30	鈴木 雅実	メディア情報科学研究所
2005. 9. 30	荒川 佳樹	知能ロボティクス研究所
2005. 9. 30	NOORANI, Ansa Benazeer Saira	知能ロボティクス研究所



受賞等

(2005年4月以降)

年 月	学会・賞名等	受賞者	受賞対象
2005/5/6	情報処理推進機構 天才プログラマー/スーパークリエイター	坂本 竜基	メタデータの豊饒化に向けた公共的アノテーションシステム
2005/5/28	電子情報通信学会 論文賞	中島 淑貴 柏岡 秀紀 ニック キャンベル	「非可聴つぶやき認識」
2005/5/28	電子情報通信学会 猪瀬賞	中島 淑貴 柏岡 秀紀 ニック キャンベル	「非可聴つぶやき認識」
2005/5/29	日本認知心理学会 優秀発表賞	西野由利恵 安藤 広志	3次元物体の学習とトップダウン処理に関わる脳部位の特定
2005/5/31	中日新聞社 第58回中日文化賞	川人 光男	脳科学研究とロボット技術への応用
2005/6/1	情報通信月間推進協議会 志田林三郎賞	川人 光男	脳科学分野の発展、情報通信技術への応用への先駆的成果
2005/6/4	人工知能学会 第六回MYCOM (AI 若手の集い) 優秀プレゼンテーション賞	中山 功一	「概念獲得システムの構成法」
2005/8/5	情報処理学会 音楽情報科学研究会 平成17年度 夏のシンポジウム ベストプレゼンテーション賞	中山 功一	「人の身体性に着目した音階の生成過程に関する考察」
2005/9/9	電子情報通信学会 フェロー	萩田 紀博	文字認識の先駆的研究とメディア情報に関する教育指導の貢献
2005/9/9	電子情報通信学会 情報・システムソサイエティ活動功労賞	桑原 和宏	人口知能と知識処理研究専門委員会幹事、副委員長、委員長としての貢献
2005/9/21	日本神経回路学会 論文賞	佐藤 雅昭 吉岡 琢 銅谷 賢治 川人 光男	Hierarchical Bayesian Estimation for MEG Inverse Problem
2005/9/21	日本神経回路学会 奨励賞	田中 沙織	セロトニンが異なる時間スケールの報酬予測に関わる線条体の活動を制御する
2005/9/21	日本神経回路学会 研究賞	田中 沙織 銅谷 賢治	セロトニンが異なる時間スケールの報酬予測に関わる線条体の活動を制御する
2005/9/21	日本神経回路学会 論文賞	田中 沙織 銅谷 賢治	Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops
2005/9/21	日本神経回路学会 論文賞	中西 淳	Feedback error learning and nonlinear adaptive control



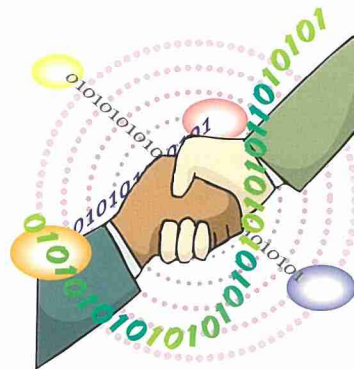
編集後記

今回はATR 音声言語コミュニケーション研究所で取り組んでいる「音声翻訳技術」の特集を組みました。最近「国際化」や「ボーダーレス」といったことが日常的に言われるようになりましたが、ビジネスや海外旅行、国内に滞在する海外の人達とのコミュニケーションにおいても、言葉の問題は依然大きな障害として残っていると思います。ATRでは音声翻訳に関する研究開発を精力的に行うとともに、新たにATR-Langというグループ企業を設立し、製品化や技術の普及に努めています。この技術の普及により使用言語が異なっても情報交換の壁がなくなることが期待しています。

ATR ジャーナルは、1987年春の創刊号の発刊以来、年2回の定期購読雑誌として、読者の皆様にATRの最新の情報をお届けして参りました。しかしながら、昨今のインターネットの普及等世の中の情勢の変化に伴い、読者の皆様への利便性も考慮致しまして、WEB化への移行を検討することとなりました。つきましては雑誌としてのATR ジャーナルは今回をもちまして最終号とさせていただきます。今後、再び皆様に有益な情報をタイムリーにお届けできるよう努力していく所存でございます。長い間、ATR ジャーナルをご愛顧頂き、誠にありがとうございました。

(ATR ジャーナル編集長 松岡茂登)

編集長：松岡 茂登
編集委員：安井 久恵、帆前加奈子、松浦 徹、若井 浩
編集スタッフ：安井 久恵、帆前加奈子、松浦 徹



ATR Journal 第54号 2005年11月7日発行

- 発行・編集 株式会社 国際電気通信基礎技術研究所
〒619-0288
「けいはんな学研都市」光台二丁目2番地2
(京都府相楽郡精華町)
(0774) 95 1111 (大代表)
- 製作 有限会社トータルマップ

本誌記事の無断転載を禁じます。

©2005 (株)国際電気通信基礎技術研究所

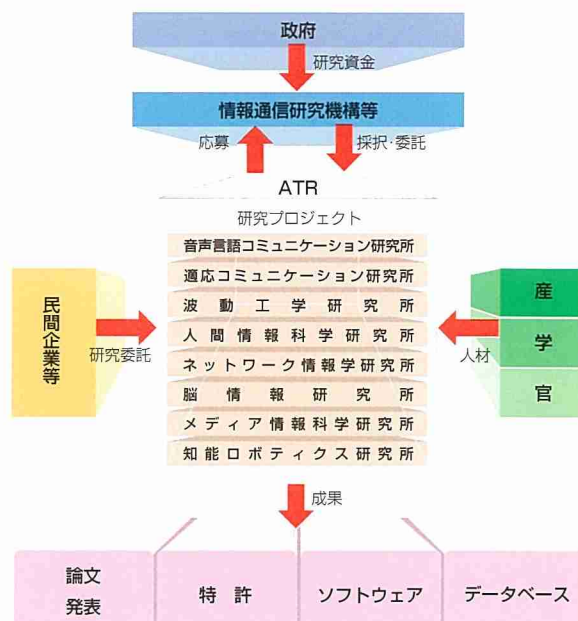
国際電気通信基礎技術研究所のご紹介

国際電気通信基礎技術研究所は電気通信分野における基礎的・独創的研究の一大拠点として内外に開かれた研究所を設立する構想のもとに産・学・官の幅広いご支援をいただき昭和61年（1986年）3月に設立いたしました。

大阪市内での3年間の暫定研究所を経て、平成元年（1989年）4月関西文化学術研究都市の中核的施設として本研究所を開所いたしました。

発足以来、基盤技術研究促進センター（KTC）および多くの民間企業から各研究開発会社への出資に支えられて研究活動を行ってきました。

平成13年度（2001年度）下期からは通信・放送機構（現・情報通信研究機構）に新設された通信・放送基盤技術に関する試験研究の促進制度に基づく委託研究等により研究資金を得て、21世紀を拓く電気通信分野の基礎的独創的研究開発を積極的に推進しています。



ATR ホームページ

<http://www.atr.jp>

お役に立つ様々な情報を公開しています。

今後も随時拡充していきますので、皆様のアクセスをお待ちしております。

※2004年1月1日よりATRのドメインネームを「atr.co.jp」から「atr.jp」へ変更いたしました。

ATR ジャーナル担当宛

TEL : (0774) 95 1183

FAX : (0774) 95 1178

E-mail : editor@atr.jp

ご連絡内容（いずれかに印をお願いします。）

☐ 送付先変更連絡

☐ ご意見、ご要望等

ATR ジャーナルの送付先変更等をお寄せ下さる場合には、下記にご記入の上、FAX 等でご送付下さい。

		変 更 後	変 更 前	変更事由
送 付 先	フリガナ お 名 前			☐ 人事異動 ☐ 住所変更 ☐ その他
	送 り 先			
	会 社 名			
	部 署 名			
	役 職 名			
	Tel / Fax			
	E-mail			
ご意見ご要望				

ATR

株式会社 国際電気通信基礎技術研究所

ATR Journal No.54

