

# C-STAR 音声翻訳 国際共同実験

横尾 昭男 竹沢 寿幸 山本 誠一  
(株) エイ・ティ・アール 音声翻訳通信研究所

(株) ATR 音声翻訳通信研究所は1999年7月22日(木)にけいはんなプラザ(京都府精華町)において、音声翻訳研究の国際コンソーシアムC-STAR (Consortium for Speech Translation Advanced Research) のメンバを接続し、日本語、英語、ドイツ語、韓国語の4カ国語間の音声翻訳実験を実施した。多言語間を同時に接続した会話実験と、ホテルの予約を想定した会話実験を行った。本稿ではその概要と結果および今後の課題を述べる。

## 会話の自動翻訳をめざして

21世紀を間近に迎え、異なる言語を話す人同士のコミュニケーションが広い範囲にわたって増大すると思われる。そのコミュニケーションを支援する技術として音声による会話を自動的に翻訳する音声翻訳技術が期待されている。

今回の実験は、1993年1月にATR自動翻訳電話研究所が実施した実験に続く2回目の国際共同実験であり、多言語間を同時につなぐ音声翻訳実験と、ふだん我々が話すような自然な会話の音声翻訳実験を実施することで、6年間の技術進歩を検証するのが目的である。

## C-STAR 音声翻訳研究国際コンソーシアム

ATR自動翻訳電話研究所が1986年に設立され、音声翻訳技術の基礎研究が進展するに伴い、他のいくつかの研究

機関でも音声翻訳の研究が開始された。これらの研究機関間の研究交流を促進するために、ATRを中心に音声翻訳の研究に関する国際的なコンソーシアムC-STARが1992年に発足した。ATR以外のメンバは、米国・ピッツバーグのカーネギーメロン大学、ドイツ・ミュンヘンのシーメンス社、ドイツ・カールスルーエのカールスルーエ大学であった。C-STARメンバは音声翻訳に関する研究活動を世界に広くアピールするために、1993年1月に世界で初めての音声翻訳に関する国際共同実験を実施した。

1993年にATR音声翻訳通信研究所が設立された後、このコンソーシアムはATRを始めとする従来の4研究機関を中核として1994年に改編され、自然な話し言葉の実時間で音声翻訳技術の研究の推進を目的とし、研究協力を効率的に進めるため1999年に研究成果をまとめることを目標としてC-STARの第2フェーズが開始された。

1999年に研究成果をまとめる一環として国際共同実験を実施することに合意した研究機関はパートナーメンバと呼ばれる。その後、パートナーメンバとして韓国・テジョンの韓国電子通信研究所、イタリア・トレントのIRST、フランス・グルノーブルのCLIPSが加わった。

C-STARでは、各国の研究機関が本国語の音声認識と音声合成、本国語から相手言語への翻訳を分担することにより効率よく研究を進めており、各研究機関の研究成果をまとめることにより多言語間の接続を可能にしている。

## 実験システムの構成と特徴

ATRのシステムは、自然な話し言葉を対象とした日本語から英語・ドイツ語・韓国語への音声翻訳システムATR-MATRIX<sup>1)~3)</sup>である。「あのー」「えーと」といった言葉がはさまれたり、「しょうがないですね」のような話し言葉独特の表現や、「予約お願いします」のような助詞が省略された表現があっても認識し、翻訳することができる。音

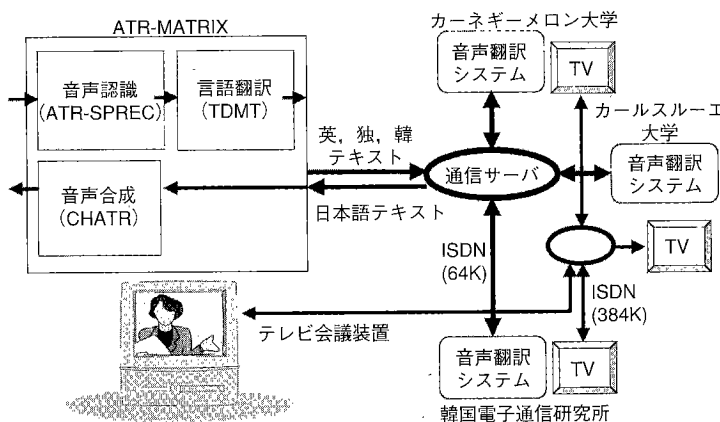


図-1 多言語間を同時に接続する音声翻訳実験のシステム構成

声認識・翻訳ともに約1万3千語の語彙を扱うことができる。これは人名等の固有名詞を除けば、ホテルの予約等の旅行会話を扱うのに十分なサイズである。

ATR-MATRIXはパソコンないしワークステーション1台で動作し、ほぼ実時間で処理を行うことができる。システム構成を変更することにより、ウェアラブルコンピュータやノートパソコンを使った構成でも同じように動作させることができる。

図-1に多言語間を同時に接続して行った音声翻訳実験のシステム構成を示す。ATR、カーネギーメロン大学、カールスルーエ大学、韓国電子通信研究所の各研究機関の音声翻訳システムの間をISDNにより64kb/sで接続した。多数の研究機関のシステム間のデータ転送を効率的に実現するために専用の通信サーバを介した。各国の音声翻訳システムは自国語の音声認識を行い、相手言語へ翻訳した後、結果をテキスト形式で通信サーバに送る。通信サーバはそれを各国のシステムに送り、各システムはその中から自国にとって必要な情報を取り出し、自国語への翻訳結果を合成音声で出力する。実験に際しては、話者や会場を映すテレビ会議装置を併用した。テレビ会議装置はISDNにより384kb/sで接続した。

## 音声翻訳実験

今回の実験は、多言語間を同時につなぐ音声翻訳実験とふだん我々が話すような自然な会話の音声翻訳実験からなる。

多言語を同時につなぐ音声翻訳実験は、ATR、カーネギーメロン大学、カールスルーエ大学、韓国電子通信研究所の間で実施した。複数の言語が用いられる国際会議のように、たとえば日本語話者が日本語で話すと、それを英語、ドイツ語、韓国語に翻訳し、英語話者が英語で発話すると日本語、ドイツ語、韓国語に同時に音声翻訳できることを示した。内容は主に日常的な挨拶会話である。具体的には「アメリカではいま何時ですか」「何時ですか、ドイツは」「日本はいま晴れていますが、韓国はいまどんな天気ですか」といったものである。図-2に実験風景(写真)を示す。

自然な会話の音声翻訳実験はATRとカーネギーメロン大学の間で実施した。日本人旅行者がワールドシリーズ観戦のためにニューヨークのホテルを予約する会話をアメリカの旅行代理店と行うという想定で会話音声の翻訳実験を実施した。途中で何度か認識誤りが生じたが、再発話や相手からの確認質問等により、会話が成立することが確かめられた。

その後、ATRの日英双方向音声翻訳システム<sup>4)</sup>の構成を変えて、音声認識や音声合成の処理の一部をクライアントであるウェアラブルコンピュータ上で動作させ、音声翻訳サーバのデスクトップ型パソコンとの間で無線LANを使ってデータを伝送して音声翻訳をする実験を行った。道に迷った英語話者が日本語話者に道を尋ねるといった設定で会話を行い、携帯型端末での実現性を示した。

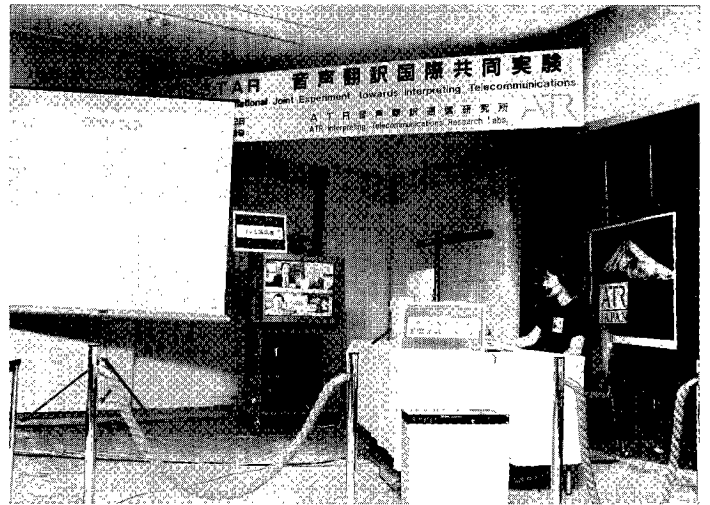


図-2 音声翻訳の実験風景

最後にけいはんな都ホテルのフロント係による音声翻訳実験を行い、一般の人でも使用できることを示した。

## 関連情報と今後の課題

来訪者数は、官公庁、大学、企業関係者等約100名とテレビや新聞等の報道関係者であった。実験の後の質疑応答では、「システム実用化まではどれくらいか」、「最終的には外国人と議論できるようになるのか」といった質問が出された。これに対し、前者には「どういう使い方をするかによって実用までの期間が変わるので予測しがたい。実験で示したような利用なら遠くはない」と答えた。最近のサーベイ<sup>5)</sup>、<sup>6)</sup>によれば2010年から2020年に実用化がなされると予測されている。後者には「日常的な議論についてはできるようになるだろうが、ATRのようなコーパスベースのシステムで特異な内容や特異な論理展開を扱うのは困難であろう」と回答した。

今後は、音声翻訳システムという機械の助けを借りた異言語コミュニケーションについて、対話をする上での人間の振舞いの変化の究明や、具体的な応用を念頭に置いた場合の性能評価と課題の解明を進めていく。

ATR-MATRIX 音声翻訳システムの概要説明のWWWホームページを開設しているので興味のある方はご覧いただきたい (<http://www.itl.atr.co.jp/matrix/>)。

### 参考文献

- 1) Takezawa, T., Morimoto, T., Sagisaka, Y., Campbell, N., Iida, H., Sugaya, F., Yokoo, A. and Yamamoto, S.: A Japanese-to-English Speech Translation System: ATR-MATRIX, In Proceedings of ICSLP '98, pp.2779-2782 (1998).
- 2) Takezawa, T., Sugaya, F., Yokoo, A. and Yamamoto, S.: A New Evaluation Method for Speech Translation Systems and a Case Study on ATR-MATRIX from Japanese to English, In Proceedings of Machine Translation Summit VII (1999), (to appear).
- 3) Sumita, E., Yamada, S., Yamamoto, K., Paul, M., Kashioka, H., Ishikawa, K. and Shirai, S.: Solutions to Problems Inherent in Spoken-language Translation: The Approach of ATR-MATRIX, In Proceedings of Machine Translation Summit VII, (1999), (to appear).
- 4) Sugaya, F., Takezawa, T., Yokoo, A. and Yamamoto, S.: End-to-end Evaluation in ATR-MATRIX: Speech Translation System between English and Japanese, In Proceedings of EUROSPEECH '99 (1999), (to appear).
- 5) <http://www.mpt.go.jp/policyreports/japanese/telecouncil/tsusin/koudojouhou/FS5.html>
- 6) <http://www.iftech.or.jp/hiroba/chronotab/chronol2.html>

(平成11年8月11日受付)