

用例読み上げ機能を持つ多言語医療受付対話支援システムM³宮部真衣[†] 吉野 孝^{††}和歌山大学大学院システム工学研究科[†] 和歌山大学システム工学部[‡] 情報通信研究機構^{††}

1. はじめに

現在、在日外国人数の増加に伴い、日本における多言語コミュニケーションの機会が増加している。多言語コミュニケーションを正確かつ円滑に行うためには、言語の違いを克服する必要がある。

本研究では、医療分野を対象とした多言語コミュニケーション支援の検討を行う。医療分野は人命に大きく関わるため、コミュニケーションに極めて高い正確性が要求される。機械翻訳技術を利用した多言語コミュニケーション支援に関する研究は進んでいるものの、現在の機械翻訳の精度は医療分野において利用できるほど高くはなく、翻訳精度の保障された支援が必要である。

我々はこれまでに、多言語医療受付対話支援システムM³の開発を行ってきた[1]。M³はタッチパネルを用いて操作を行うシステムであり、テキストを介した対話や受診支援が可能である。M³では正確に翻訳された用例対訳を利用し、医療従事者および外国人患者の母国語による高精度なコミュニケーションの支援を実現している。しかし、在日外国人の中には、日本語を話すことのできない人や、日本語や母国語を話すことはできるが、読み書きのできない人が存在する[2]。M³はテキストベースの支援を行っており、このような非識字者の支援は困難である。

そこで、M³に用例読み上げ機能の組み込みを行った。本システムは、多言語発話収集Webシステム[3]との連携により、音声を用いた非識字患者の理解の補助を実現する。

本稿では、これまでのシステムの概要と問題点、用例読み上げ機能の組み込みおよび今後の課題について述べる。

2. 関連研究

テキスト読み上げ機能は、視覚障害者支援への応用が多くなされており、これまでにコンピュータの画面やメニューの読み上げ、施設案内などの様々な分野において視覚障害者のためのインタフェースの検討が行われている[4, 5]。一方、非識字者を支援するためのインタフェースに関しては十分な検討が行われていない。視覚障害者と非識字者では、音声による支援が必要であるという点では同じであるが、その他の部分については必要とされる支援方法が異なると考えられる。したがって、非識字者に適したインタフェースを検討する必要がある。

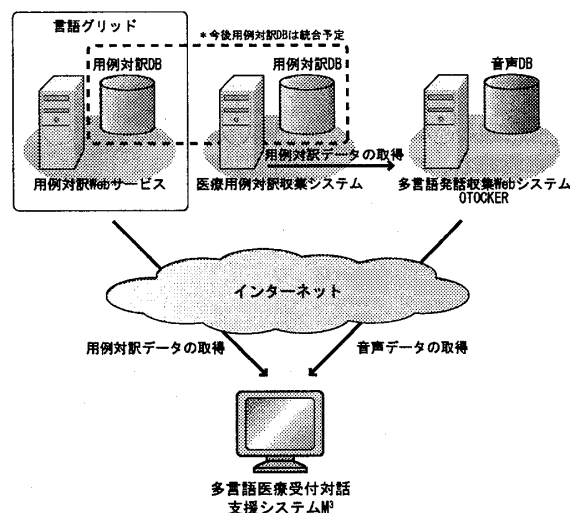


図1 システム構成

3. 従来の多言語医療受付対話支援システムM³の概要と問題点

これまでに開発したM³は、以下の機能により外国人患者の受診を支援する。

- (1) 対話機能
- (2) 問診機能
- (3) 受診科選択機能
- (4) 道案内機能
- (5) Q & A機能
- (6) 受診手続き支援機能

(1)～(3)は医療従事者が主導し、外国人患者から情報を引き出すための機能である。(4)～(6)の機能は、患者のみの操作による、受診や院内における問題の解決を支援するための機能である。

これらの機能は、全てテキストベースであり、非識字者のコミュニケーション支援は難しい。しかし、世界的に非識字者は珍しくなく[2]、非識字者に対する支援が必要である。

4. 用例読み上げ機能の組み込み

非識字者のコミュニケーションを支援するために、対話で用いる用例対訳の読み上げ機能を実現する。本研究の支援対象は医療分野であり、用例対訳の精度の高さだけではなく、読み上げる音声にも適切なニュアンスの表現が望まれる。本章では読み上げ機能についての検討および機能の実装について述べる。

4.1 読み上げ機能の検討

テキストの読み上げを行う技術として、古くから音声合成の研究がなされている。近年では、合成音声の品質

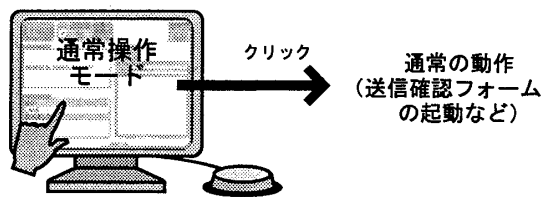
M³: Multilingual Interactive Medical Support System at Medical Reception with the Function of Reading Parallel Texts
Mai Miyabe[†], Takashi Yoshino^{††}

[†]Graduate School of Systems Engineering, Wakayama University

[‡]Faculty of Systems Engineering, Wakayama University

^{††}National Institute of Information and Communications Technology

(a) モード切替スイッチを押していない場合



(b) モード切替スイッチを押している場合

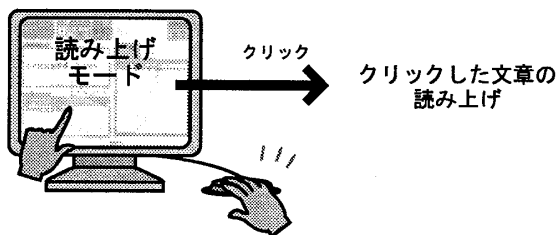


図2 読み上げ機能のイメージ

を自然音声に近い努力がなされ、高品質な合成音を実現されてきている[6]。さらに、音声合成の多言語への応用もなされており、東芝は9ヶ国語への音声合成に対応し、言語特有の韻律制御と発音を実現している[7, 8]。これらの技術を用いれば、対応している言語については高品質なテキストの読み上げが可能であると考えられる。しかし、新たな言語の高品質な音声合成を実現するには、言語固有の処理等が必要になるため、対応言語の増加に関する敷居は高い。本システムの現在の対応言語は4言語であるが、今後対応言語の増加を予定しており、新規言語の用例作成と併せて、新規用例の読み上げにも対応していく必要がある。

そこで、本研究では多言語発話収集 Web システム OTOCKER[3]との連携による読み上げ機能の実装を行う。OTOCKERはWebブラウザ上でユーザがテキストを発声することにより、正確な多言語音声データを収集・提供するシステムであり、現在、本システムの対応言語（英語、中国語、ポルトガル語、韓国語、日本語）の音声登録・利用することができる。また、対応言語を増加する場合、新規言語のネイティブがいれば音声データを登録可能であるため、OTOCKERは対応言語の増加が容易である。さらに、本システムは予め用意された用例対訳を用いるため、動的に音声生成する必要はない。また、新たな用例対訳を登録する場合、登録作業を行うことができる医療用例対訳収集システム[9]とOTOCKERは連携しているため、対訳作成者に音声登録を促すことも可能である。また、人の発声した音声データを用いるため、医療に根差した適切なイントネーションの音声データを実現できる可能性が高い。

4.2 用例読み上げ機能の実装

(1) システム構成

本システムの構成を図1に示す。本システムは、ネットワークを介して、情報通信研究機構（NICT）の言語グリッドプロジェクトが提供している言語グリッド[10]上の用例対訳 Web サービスおよび多言語発話収集 Web システム OTOCKER と連携し、多言語の用例対訳および音声データを取得する。

(2) 用例読み上げのインタフェース

非識字者を支援するためには、システムで表示される全てのテキストの読み上げを可能にする必要がある。医療従事者からの質問のみを読み上げるのであれば、質問文を自動で読み上げることが可能であるが、本システムでは複数の回答候補から自分の回答を選択する必要がある。また患者主体の機能では、多数の選択肢やボタン等があり、自分のわからないテキストを選択し、読み上げを可能にする必要がある。また、読み書きができる場合、テキストを読み上げる必要はない。

そこで、本システムでは通常操作モードと読み上げモードを自由に切り替えてそれぞれの操作を行う。モードの切り替えにはモード切替スイッチを使用する。読み上げ機能のイメージを図2に示す。通常操作モードを基本とし、モード切替スイッチを押している間のみタッチパネル上でクリックした文章を読み上げる。モード切替スイッチを押していない場合は、通常の動作（回答の送信確認や各ボタンの動作）が行われる。

5. おわりに

本稿では、用例読み上げ機能を持つ多言語医療受付対話支援システムM³の開発を行った。

今後は、開発した用例読み上げ機能の評価実験を行い、適切な用例読み上げインタフェースの検討を行う。また、非識字者の対話を円滑化するために、応答文の多言語音声認識機能の検討を進める。

謝辞

本研究は、日本学術振興会科学研究費 基盤研究 (B) (19300036) の補助を受けた。

参考文献

- [1] 宮部 真衣 他：病院受付における多言語間コミュニケーション支援システムM³の開発，情報処理学会，マルチメディア，分散，協調とモバイル（DICOM2007）シンポジウム，pp. 355-363（2007）。
- [2] 小林 米幸：外国人患者診療・看護ガイド，エルゼビア，ジャパン（2002）。
- [3] 福島 拓 他：医療分野を対象とした多言語発話収集 Web システム OTOCKER の開発，電子情報通信学会技術研究報告，AI2007-14，pp. 17-22（2007）。
- [4] 渡辺 隆行：視覚障害者のコンピュータ利用と日英2カ国語音声化システム BEP，情報処理学会誌，Vol. 43，No. 8，pp. 873-879（2002）。
- [5] 清田 公保 他：視覚障害者のためのカメラ映像内の文字情報読み上げシステム，情報処理学会研究報告，2006-UBI-010，pp. 139-143（2006）。
- [6] 広瀬 啓吉：音声合成研究への招待-自由な合成の実現に向けて-，情報処理学会誌，Vol. 43，No. 3，pp. 321-324（2002）。
- [7] 平林 剛 他：TOS-Drive TTS の米・英・独・仏・中国語への応用，日本音響学会講演論文集，pp. 239-240（2002）。
- [8] mmlab：多言語音声認識・合成，<http://www.toshiba.co.jp/rdc/mmlab/tech/w25.htm>
- [9] 岩部 正明 他：Web サービス連携を用いた医療用例対訳の収集と利用，電子情報通信学会技術研究報告，AI2006-28，pp. 17-22（2006）。
- [10] Toru Ishida: Language Grid: An Infrastructure for Intercultural Collaboration, IEEE/IPSJ Symposium on Applications and the Internet (SAINT-06), pp. 96-100 (2006)。