

2T-8 音声入力機能を備えた英会話用知的CAIシステム

田川 忠道† 山本 秀樹† 平山 輝† 明神 知†

† 沖電気工業(株) †† 大阪ガス(株)

‡ (株) オージー情報システム総研

1 はじめに

近年の技術革新に伴い、動画像技術や音声入出力技術などが発達し、これらマルチメディアを扱うことのできるパーソナルコンピュータやワークステーションが安価に入手することができるようになった。

そこで本論文では、音声入力機能を組み込んだ英会話用知的CAIシステムについて述べる。以下ではまず、英会話用知的CAIシステムに対し、音声入力機能を導入することの特質について考察し、さらにそのシステムの概要について述べる。

2 音声入力機能導入の特質

まず、音声入力機能が導入されていない従来システムの問題点について考察する。従来システムとしては、会話シミュレーションを基にした英会話用知的CAIシステム(文献[1])などがある。これらのシステムは、学習者側の発話入力をキーボードを用いて行なうために、

- ・ 単語綴りも正確に入力する必要があり面倒である
- ・ 入力速度の限界などから、会話的な即応性を養うことができない

という問題点がある。これらの問題点は、音声入力機能を導入することにより容易に解決することができる。

つぎに、音声入力機能を導入するために、最も重要な部分である音声認識装置の方式について検討する。最近の音声認識装置の方式は、文献[2]によると、

- 単語または文節の孤立発声かつ特定話者の発声訓練を必要とするが、多くの語彙を扱えるもの

- 文型や話題の制約があるが、連続音声で不特定話者に対応するもの

に大別される。ここで、会話の練習をするという本システムの目的から考察すると、一文の連続入力が必要であるため、bの方式を採用した音声認識装置を採用する必要がある。

最後に、上記bの方式の音声認識装置による音声入力機能を、従来システムへ導入した場合について考察する。その導入により新たに発生する問題点としては、

1. 入力文の制約により、完全に自由な会話に対応することができない
2. 音声認識装置において認識誤りが発生する

が考えられる。まず1については、場面設定などを適宜行なうことにより、会話教育として十分な会話が実現できる。また2については、音声認識装置の認識誤りの可能性が高いとシステムが判断した場合には、再入力を促すという方針をとる。発音がかなり悪い場合には認識誤りが発生しやすいことを考えれば、会話教育という意味で再入力を促すことは比較的自然である。

3 システム概要

音声入力機能の概略図を図1に示す。本章では、各モジュールの概要について説明する。

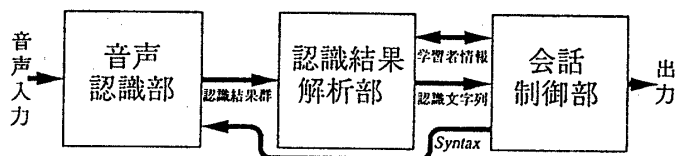


図1: 音声入力機能概要図

An Intelligent CAI System for Conversational English with a Speech Recognition System.

Tadamichi TAGAWA† Hideki YAMAMOTO† Hikaru HIRAYAMA†† Satoru MYOJIN†

† Oki Electric Industry Co., Ltd. †† Osaka Gas Co., Ltd. ‡ Osaka Gas Information System Research Institute

3.1 音声認識部

音声認識部では、音声認識装置 [3] を用いて音声入力を実現している。この音声認識装置は、不特定話者の連続音声に対応しており、前章で考察した仕様に合致している。装置の概略仕様は、*Syntax* とよばれる有限状態文法で、認識すべき文を記述し予め登録しておけば、その *Syntax* から生成される認識候補文とその認識評価点の対からなる認識結果群が得られるというものである。その *Syntax* の例を図 2 に、認識結果群の例を図 3 に示す。

```

S      -> _GREETING _DOE
      |  _GREETING
_GREETING -> good afternoon
      |  good evening
_DOE     -> mr john doe
      |  mr doe

```

図 2: *Syntax* の例

認識候補文	認識評価点
good afternoon mr john doe	652
good evening mr john doe	640
good afternoon	540

図 3: 認識結果群の例

3.2 認識結果解析部

認識結果解析部は、音声認識部から得られる認識結果群の中から、有意な認識候補文を選択する。有意性は、学習者レベルなどから設定された閾値と、認識評価点とを比較することにより判断している。選択された認識候補文の中に、認識誤りが含まれている可能性が高いと本モジュールが判断した場合には、有意な候補がなかったとして、再入力要求を次の会話制御部へ渡す。また正しく認識されたが、学習者レベルなどからもっと正確な発話が期待できると本モジュールが判断した場合にも、上記と同様の処理を行なう。

3.3 会話制御部

会話制御部は、学習者とシステムの会話を制御するモジュールである。ここでは、音声入力機能を導入す

ることにより、従来のシステムと大きく異なる点について簡単に説明する。

1 つは、学習者の次発話を予測してその *Syntax* を生成する機能である。

2 つめは、学習者の入力部分の解析方法の変更である。従来のシステムではキーボード入力であったために、会話制御部は綴誤りの訂正や構文解析を行ない、学習者からの入力を得ていた。しかし、音声入力機能の導入により、学習者の入力は認識結果解析部より認識結果文字列の形で得られるため、本モジュールではその文字列が予測した次発話のどれに該当するかの一致検出をするだけでよくなっている。

4 評価

本システムを、実際にワークステーション上に構築して試用した。会話シミュレーションの応答時間は、*Syntax* の大きさや音声認識解析レベルにより変化するが、通常の使用状態で 500(msec) 程度であった。これは、会話をしているという感覚を与えるのに十分であり、高い会話学習の効果が期待できる。

5 おわりに

音声入力は、その認識率の向上と認識装置の低価格化が進むにつれ、ますます多用されると考えられる。

あるアプリケーションに音声入力機能を導入する際に考慮すべき点は、認識誤りと再入力がかつて許容できるかである。本論文でも述べたように、英会話用知的 C A I システムは、上記許容範囲が大きい。このため音声認識装置のアプリケーションとしては最適なものであり、実用システムとなり得るものであると考える。

参考文献

- [1] 山本ほか: 会話シミュレーションを基にした語学訓練用知的 C A I システムの構成, 情報処理学会論文誌, Vol.30, No.7, pp.908-917 (1989).
- [2] 藤崎: 音声認識・理解の目標と将来課題, 電子情報通信学会誌, Vol.73, No.12, pp.1264-1268 (1990).
- [3] 平山, 平島: 不特定話者, 連続音声認識システムの開発とその応用, International Symposium "Computer World '91," pp.180-196 (1991).