

合成音声を用いてインタラクティブに教示するペン習字 CAI システムの構築

篠崎 翼 阿部 匡伸

NTT ヒューマンインタフェース研究所
〒239 神奈川県横須賀市光の丘 1-1

Tel: 0468-59-4264, 0468-59-2547. Fax: 0468-56-1054

E-mail: tsubasa@ntspch.hil.ntt.co.jp, ave@ntspch.hil.ntt.co.jp

1. はじめに

我々はシステムを使う気になるとかシステムに親しみを持てるといったユーザフレンドリなシステムを構築したいと考えている[1][2]。このようなマンマシンインタフェースのユーザフレンドリ化の検討を進める上では、具体的なタスクを考える必要がある。本報告では、具体的なタスクとして合成音声によってインタラクティブな教示[3]を行う CAI ペン習字システム「CALTURE」(CALLigraphy Training system Using voice Response)の構築について述べる。

2. CAIペン習字システム「CALTURE」

音声によってインタラクティブに指示を行う CAI ペン習字システム「CALTURE」を図 1 に示す。このシステムは液晶画面付きペンタブレットを用いて、画面に表示されたお手本をペンでなぞることによりペン習字を行うものである。表示されたお手本をユーザがなぞると、書きはじめの位置やストローク(一筆)のずれについてインタラクティブに音声で指示を行う。また、正しくなぞれている場合には褒めるといったことを行う。

3. お手本と指示発話の収録

ペン習字の題材として「ひらがな」を選んだ。まず、ペン習字の指示発話に関する基本的な知見を得るために、「ひ」「お」「ほ」の 3 つの文字についてお手本の収集と指示発話の収録を行った。指示発話の収録では、ペン習字の方法を「お手本をなぞる練習」に限定し、なぞり書きにおいてお手本とのずれを指摘、書き直させることを想定して行った。

3.1. お手本の収集

図 1 のペンタブレットを用い補助線によって 16 マスに区切られた正方形の文字枠の中に、一文字ずつお手本の字を先生役の被験者に書かせ時系列データとして収集した。お手本の収集は先生役の被験者としてペン習字経験のある成人女性一名で行った。

3.2. リアルタイムな書き直し指示の収録

文字の「書きはじめ」や「跳ね」、「止め」といった各部分について、間違ったその場で指示を行う場合の発話の収録を行った。収録のための実験システムを図 2 に示す。実験は、先生役側にディスプレイを置き、お手本を表示しておく。操作側では先生の指示の様子を VTR カメラによってモニタできる。操作者がペンタブレットの液晶画面に表示されたお手本をペンでなぞると、先生側のディスプレイになぞり書きの様子が表示される。先生には予め、操作側ではお手本が見えない状態で文字を書いていると伝えておく。そして、操作者は、予め

“Computer Assisted Instruction calligraphy training system to teach interactively using synthesized speech”

Tsubasa SHINOZAKI, Masanobu ABE

NTT Human Interface Laboratories

1-1 Hikarinooka Yokosuka-shi Kanagawa 239 Japan

Tel: +81-468-59-4264, +81-468-59-2547 Fax: +81-468-56-1054

いくつかの書き誤りのパターンを想定しておき、お手本をなぞるときわざと書き誤る。先生にはディスプレイに映るなぞり書きの様子を見て音声で指示させた。このとき、画面への指差しによる指示についても許した。この指示発話と指差しの様子を DAT 及び VTR で収録した。

3.3. 指示発話に関して得られた知見

指示発話の収録を通して得られた知見のうち、指示発話に関する知見を以下に示す。

- 指示は「書きはじめ」の位置や「跳ね」、「止め」といったストローク(一筆)のはじめと終わり及び、「曲げ」や「結び」といった特徴的な点について主に行われる。
- 指示の内容は「跳ね」や「止め」、特徴的な点について各ストロークごとに細かく指示するものと、バランスやつながりといった全体に関する指示とに大きく分けられる。
- 「書きはじめ」の誤りは直ぐに指示する。
- 「書きはじめ」以外は、1 ストローク書き終わってから指示を出すことが多い。
- 指示を出す際、動作にインタラプトをかけるために「あつ」や「えつ」、「そこっ」といった間投詞や指示語が付くことが多い。
- 誤りを指摘するとき、「そこ」か「ここは」といった指示語に指差しを添えた指示がよくみられる。
- 誤りがない場合には「そうそう」や「そう」といった指示発

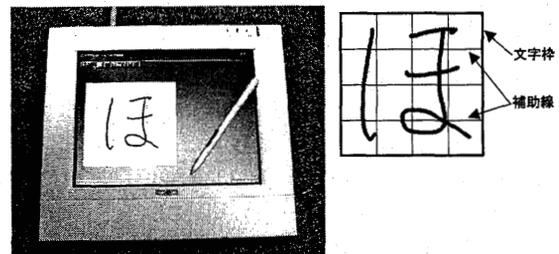


図 1. CAI ペン習字システム「CALTURE」
液晶画面付きペンタブレットを用い、音声指示によってインタラクティブにペン習字を行うシステム。ペン習字のための文字枠(図右上)。

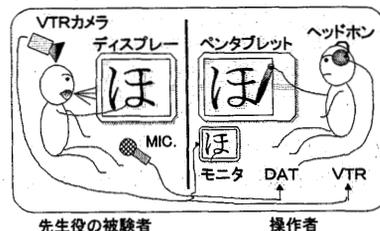


図 2. 指示発話収録のための実験システム
被験者はディスプレイに映った操作者のお手本をなぞる様子を見て指示を行う。その様子を DAT 及び VTR で記録する。操作者は被験者の様子をモニターとヘッドホンを通して知ることができる。

話がなされる。

4. ペン習字の練習ステップ

3 節で得られた知見から、ペン習字には文字全体の大きな形やバランスを練習するモードと各ストロークの「跳ね」や「曲げ」といった部分を細かく練習するモードがある。そこで、「CALTURE」では次の 3 つのステップで練習を行う。

STEP1 書き方のポイント教示とお手本をなぞる練習

STEP2 リアルタイムな指示による 1 ストロークごとの練習

STEP3 1 文字単位のバランス練習

5. リアルタイムな指示発話の生成

「CALTURE」ではユーザのペン入力に対してリアクション音声を返すことによりインタラクティブに教示を行う。また、同時にリアクション音声を通して「やさしい先生」であるとか「厳しい先生」といったように、システムに対して個性を与えることを考えている。このようにインタラクティブな教示と個性を与えることによって、ユーザがシステムに対して人間っぽさや親しみを感じ、楽しくペン習字を行えるといったユーザフレンドリなシステムを目指している。

5. 1. リアクション音声の生成

リアクション音声の生成フローチャートを図 3 に示す。

2 つの異なるメッセージ生成パス

リアクション音声の生成では、「書きはじめ」の位置のずれなどのユーザのペン入力によって発生したイベントに対して、文脈スクリプト(b)にもとづいてインタラプトメッセージ(処理 A)と指示メッセージ(処理 BCD)とが生成される。文脈スクリプト(b)には 3 節の知見をもとに、予め文字ごとに指示を行うポイント(表 1)及びインタラプトメッセージを出す出さない等の制御情報が記述してある。インタラプトメッセージは「あっ」とか「そこっ」といったような間投詞や指示代名詞を用いてユーザの注意を引くことによりユーザの動作をインタラプトする。一方、イベント入力(a)と文脈スクリプト(b)を比較することによってどういった指示を出すかの決定(処理 B)、「右から」とか「大きく跳ねて」といった指示メッセージが生成される(処理 C)。

指示メッセージの修飾

生成された指示メッセージはリアクションモードの決定(処理 F)にもとづいて修飾される(処理 D)。リアクションモードの決定(処理 F)では、予め設定しておいた「やさしい先生」や「厳しい先生」といった教師モード(c)及び出力したリアクション音声の履歴(処理 E)をもとに指示メッセージに感情表現やバリエーションをつけることにより、システムに対して個性を与える。例えば、何度も同じ誤りを繰り返していると次第に口調が厳しくなり「怒った」印象を与えるだとか、うまく書けるようになると次第に口調が柔らかくなり「明るくのりのよい」印象を与えるといったようにユーザへ与える印象に変化を付ける。また、教師モードが「きびしい先生」と「やさしい先生」とでは、「怒る」までの閾値を変えると行ったことを行う。このような人同士の会話でも見られるモードの切り替えによってユーザへ与える印象を変化させ、知的な振る舞いや人間っぽい振る舞いをシステムで表現する。

リアクションセット

修飾された指示メッセージに対応する音声データをリアクションセット(d)から取り出し、リアクション音声(e)として出力する。リアクションセットは大きく 2 つに分類される。1 つは「汎用リア

クションセット」であって、「怒った感じ」とか「のりのよい感じ」といったように合成できる相槌や合いの手、間投詞といった汎用的な音声パラメータを集めたセットである。他は、「右から」や「長く」といった文脈スクリプトに対応した音声を集めた「特定リアクションセット」である。こうしておくことで「文脈スクリプト」と「特定リアクションセット」の交換により、リアクション音声の生成の機構を他のシステムへ適用することができる。また、リアクションセットの音声データは、Speed97 [4]を用いて韻律を適切に調整した合成音声の韻律パラメータの形で用意することで、自然で多様な韻律を持ちかつコンパクトなものも利用できる。

以上のようにこのリアクション音声生成の特徴は、指示を行う前にインタラプトメッセージの出力を行えること、そして、教師モードと出力履歴をもとに感情表現やバリエーションの付けられる機構を持つことにある。

6. まとめ

本報告では、ペン習字における誤り指示についての収録から得られた知見をもとに合成音声によってインタラクティブに教示を行う CAI ペン習字システム「CALTURE」の構築について紹介した。今後は、「CALTURE」を用い、更にユーザフレンドリ化のための検討を行っていく予定である。

参考文献

- [1] 阿部：“誤り指示音声の特徴分析と音声出力への適用,” 電子情報通信学会論文誌 D-2 Vol. J79-D-2 No. 12 pp. 2191-2198.
- [2] 篠崎, 阿部：“規則合成音声で躍動感を実現する方略について,” 情報処理学会研究報告, 97-SLP-17, pp. 81-88 (1997).
- [3] 篠崎, 阿部：“合成音声によるインタラクティブな教示音声を CAI システムへ導入するための検討,” 情報処理学会研究報告, 98-SLP-20, (1998).
- [4] 阿部, 水野, 中島：“様々な音声表現を実現できる音声作成ツール - Speed97 -,” 情報処理学会研究報告, 97-SLP-17, pp. 67-72 (1997).

表 1. 例、ひらがな「ま」について指示を行うポイント

ストローク	指示を行うポイント
1	「書きはじめ」、「縦棒の曲げ」、「跳ね」
2	「書きはじめ」、「跳ね」
3	「書きはじめ」、「止め」
4	「書きはじめ」、「結び」、「止め」

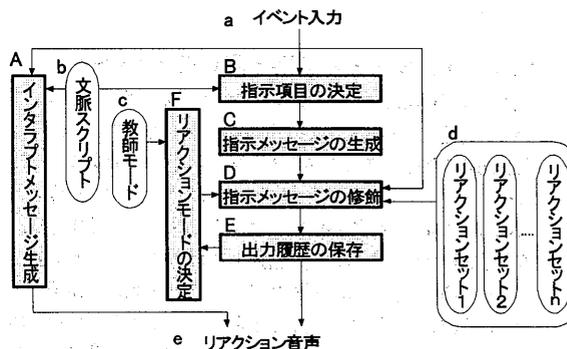


図 3. リアクション音声生成のプロチャート

イベントに対しインタラプトメッセージを生成し、文脈スクリプトに沿って指示項目を決定する。教師モードと指示の出力履歴をもとに、「怒る」や「驚める」といったリアクションモードを決定、これにしたがってメッセージを生成、修飾し、リアクションセットから選ばれた音声によってリアクション音声生成される。