

インタラクティブなテキスト 合成音声メッセージの生成法

5G-9

長 倉 恵 一 伊 藤 真 人
NTT グループ事業推進本部

【1】 はじめに

テキスト音声変換技術を用いた音声メッセージの生成は、音声蓄積&再生技術ベースの方法に比べて、音声録音作業を省けるため、更新頻度の高い音声メッセージの生成法として注目されており、筆者らも応用方法について報告した[1]。しかし、日本語テキストから合成音声メッセージを生成する場合、①漢字の読み誤りが生じる、②音声メッセージが肉声に比べて単調で、厚みに欠けるメッセージになり易い等の問題がある。筆者らは、良好な合成音声メッセージの生成を狙いとして、入力した日本語テキストに対して、読みがな表示による確認と修正、GUIによるテキストへの音声制御符号の挿入、合成音声の試聴等をインタラクティブに行えるシステムを検討したので報告する。

【2】 技術課題

日本語の規則音声合成の研究は、幅広く行われているが、前項に述べた視点から問題解決に取り組んだ例は少なく、音声強調インタフェースの設計として報告された[2]程度である。

インタラクティブにテキスト修正/音声試聴を繰り返しながら良好な合成音声メッセージを生成するシステム実現のための第1の課題は、漢字の読み誤りへの対処法である。次の課題が厚みのある合成音声とするための、テキスト中の音声制御符号の設定法である。そして、

複数の音声制御符号を設定した場合には、それらの設定状況の確認が必要になる。

【3】 実現法

テキストデータ入力から合成音声出力までの処理、読み誤りチェック、音声制御符号の設定及びそのチェックの流れを図1に示し、処理の内容を以下に述べる。

(1) 漢字かな変換と音声合成

本システムにおけるテキスト音声合成機能は、「しゃべりん坊HG」[3]により実現した。本装置は、漢字かな混じりテキストを合成音声に変換するほか、入力テキストに挿入する音声制御符号により音量や発声速度を変えることができる。

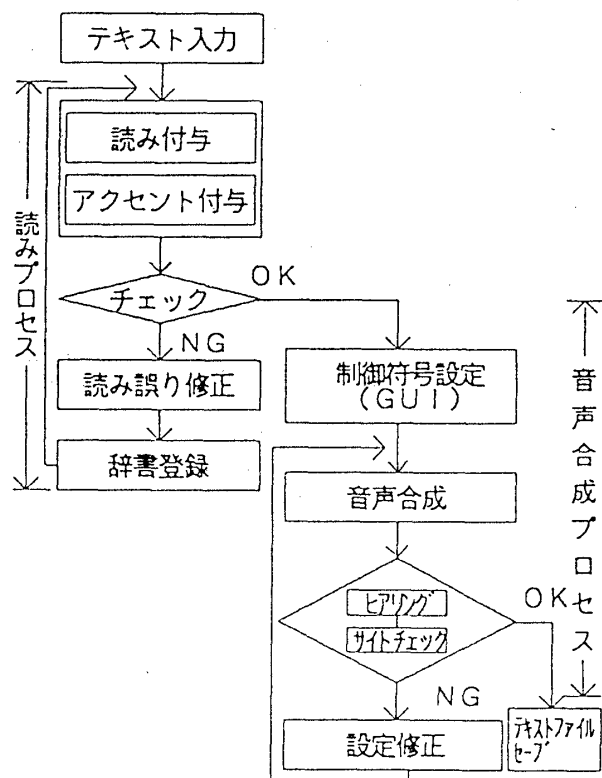


図1 合成音声メッセージ生成の流れ

Text-to-Speech Message Generation System With
Interactive Text Editing and Speech Control
Keiichi NAGAKURA · Masato ITO

NTT Affiliated Business Development Headquarters

(2) 漢字の読み誤り修正

漢字かな混じりテキストからカナテキストへの変換結果の表示、及びエディタによる誤り修正を図2の様に Windows 環境 [4] で実現した。また同じ誤りを繰り返さないように誤り修正語を単語登録（私用語登録）する機能を実現（図3）して、読み誤り対策とした。

(3) 合成音声コントロール

テキストへの音声制御符号の設定により、合成音声をコントロールする機能の実現には、音声強調インタフェースの設計法 [2] をベースに、改善を加えて実現した。改善の主なポイントは、①二次元座標による抑揚、ピッチ、音量設定を項目毎に画面切り替え表示し（図4）、どう設定したか分かる様に、これを可視化した（図5）②GUIの設計を Windows 環境で実現し、馴染みのあるGUI環境とした。

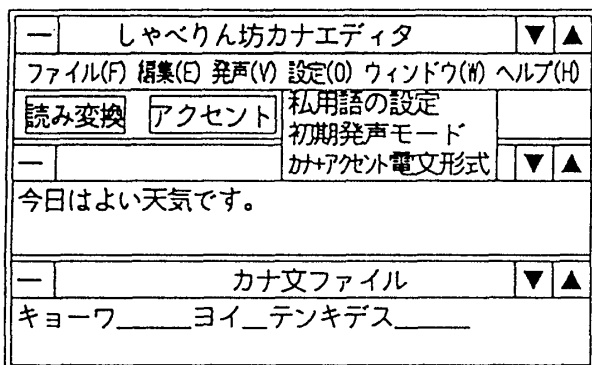


図2 漢字カナ変換結果表示と修正機能

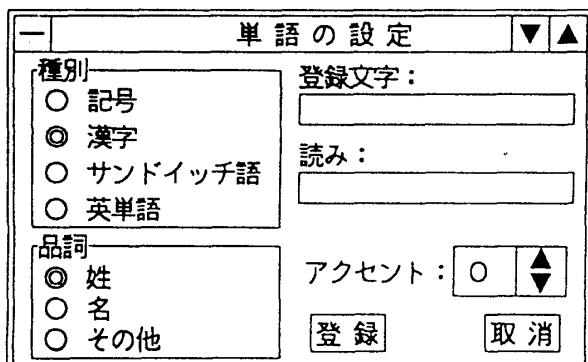


図3 私用語登録機能

【4】本システムの応用

新曲CDを電話で聴く「ミュージック オンデマンド」サービスにおいて、リクエスト曲の確認用音声メッセージ生成に、本システムが活用され、有効性が確認された。

【5】おわりに

良好な合成音声メッセージを得るために、テキスト修正/音声試聴をインタラクティブに行うシステムの検討を行い、漢字の読み誤りへの対処、GUIの音声制御符号設定による音声コントロールを盛り込んでシステムを実現した。

（参考文献）

[1] 長倉他：“表形式データを情報源とする音声情報提供の実現法”第47回情処全大5V-8

[2] 浜田他：“音声合成における音声強調インタフェースの設計法”情処学会HI研究会資料 vol 92 No 15 (1992)

[3] NTTインテリジェントテクノロジー “しゃべりん坊HG操作マニュアル” (1991)

[4] Microsoft “Windows Ver3.1”

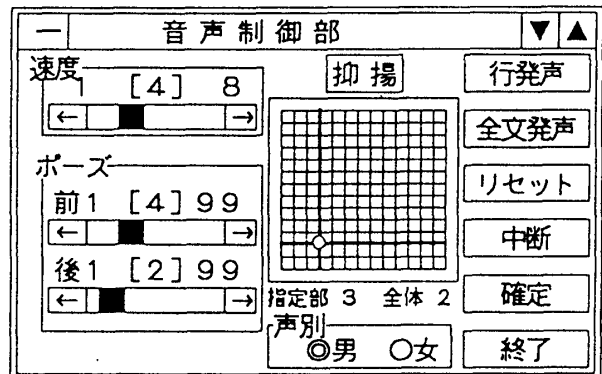


図4 GUIによる合成音声コントロール設定

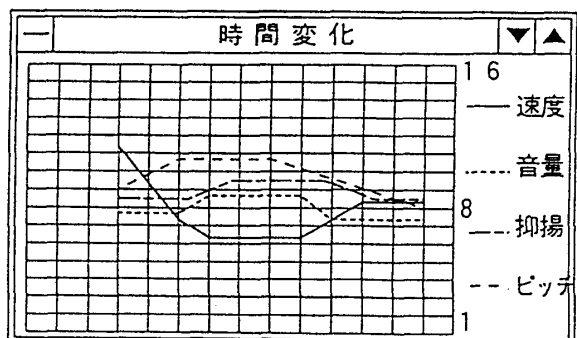


図5 合成音声コントロール設定状況の表示