

文音声変換ソフトのための音韻列生成方式の開発

5G-8

- 音韻列生成方式と開発環境 -

小林賢一郎* 新田恒雄** 原義幸** 山中紀子**

*東芝AVE(株) ***(株)東芝 マルチメディア技術研究所

1. はじめに

ソフトウェアによるPC上での文音声変換ソフトが各社で開発されている^{[1][3]}。優れた文音声変換ソフトを開発する上で、単語辞書、日本語解析、音声記号の生成、音声合成器のバランスのとれた開発が必要である。

開発の段階で、これらの要素のどれかに対して修正を加えると、修正した部分に関しては改善されるが、修正の仕方によっては他の部分に対して悪影響を与える場合も生じる。

本報告では文音声変換ソフトにおける音韻列生成方式と、文音声変換ソフトに必要なこれらの要素を、バランスよく開発するために相互の影響を評価することができる開発環境について報告する。

2. システムの構成

評価システムは図1に示すような構成をとっており、単語辞書、単漢字辞書、形態素解析部、音声記号生成部、生成規則、音声合成器、音声素片、影響評価部から構成されており影響評価部を除いた部分が基本的な文音声変換ソフト部分である。各部分は自由に取り替えることができるようになっている。

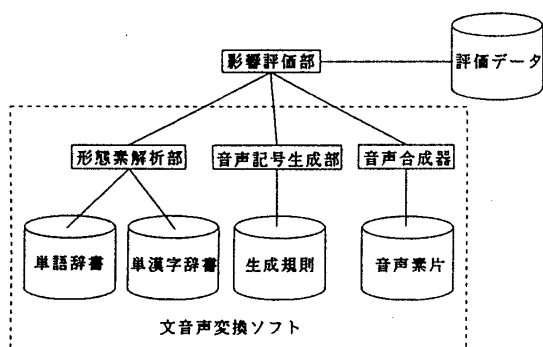


図1 システム構成図

Development of Phonetic and Prosodic Symbol Generation Method for Text-to-Speech Software.

Ken-ichiro Kobayashi*, Tsuneo Nitta**, Yoshiyuki Hara**, and Noriko Yamanaka**

*Toshiba AVE Co., Ltd. **Multimedia Engineering Lab., Toshiba Corp.

影響評価部においては、これらの各部を変更した際に生じる音韻記号列への影響を評価する。

次に各部分について説明する。

2-1 辞書

単語辞書には形態素解析を行うための単語と、単語に付随した音声合成に必要な情報が記述されている。記載されている情報としては、品詞、読み、アクセント情報などがある。漢字単位の読み情報は単漢字辞書にある。

2-2 形態素解析

入力された文章を単語辞書と単漢字辞書を参照することにより形態素に分割する。解析結果には各形態素の読み、アクセント情報などが情報が付与される。

2-3 音声記号生成

形態素解析結果をもとに音韻列生成規則に基づき読みの変更、アクセント情報の付与、ポーズ位置及び長さの設定を行い、音声合成器に渡すための音声記号列の生成を行う。

2-4 音声合成器

音声記号列をもとにケプストラム方式^[2]によって音声波形の生成を行い、合成音の出力を行う。

2-5 影響評価部

音声合成ソフトの評価を行うための評価データの作成、評価、分類作業を行う。

3. 処理の流れ

文音声変換ソフトは入力された漢字かな混じりの文字列を、形態素解析により単語分割を行い、得られた解析結果より音韻列生成規則に基づいて読み、アクセント、ポーズなどを制御して音声記号列を生成し、その音声記号列をもとに合成音を生成する。

4. 修正に伴う影響の評価

音声合成ソフトを開発の過程で、辞書、形態素解析、及び音声記号生成部のいずれかに修正を加えた場合、他の文章の読み上げ結果に対してに影響を与える場合がある。この影響の中で、従来正しく読み上げることができなかったものが正しく読み上げられることは歓迎されるが、正しく読めていたものが正しく読めなくなることはできる限り避ける必要がある。

このような修正に伴う他の文章に対する悪影響を少なくするためには、修正による他の文章への影響を評価する必要がある。

今回影響の評価を行う方法として正しく読めた実績のある文とその文に対応する音声記号データをテンプレートとして収集し、修正を加えたときに従来正しく読み上げることのできた部分がどのように変化するかを評価することを基本方針として評価を行った。

評価は図2に示す手順により行われ、大きく分けて評価データの作成、分類、自動評価の3つの行程により処理される。

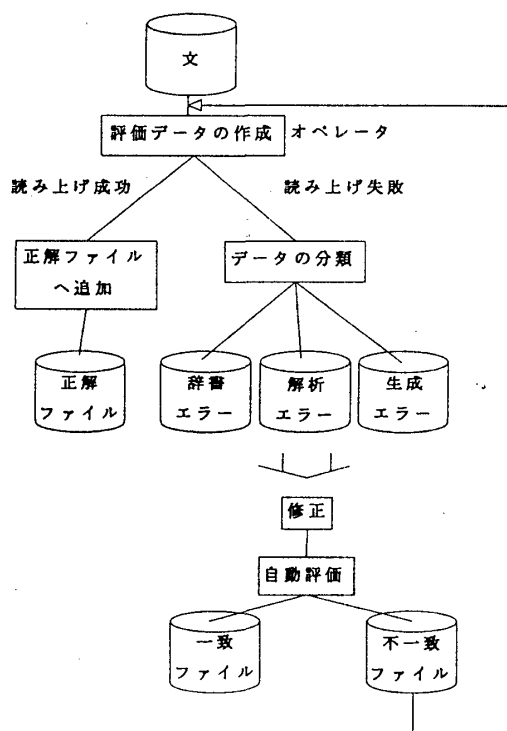


図2 評価処理の流れ図

4-1 評価データの作成

ここでは評価の対象となる文章を文単位に切り分け、各々の文に対して音声記号列の生成を行う。ここで正

しく生成されたものとそうでないものに分類作業を行い、正しく生成されたもののみを正解データとして蓄積していく。

4-2 分類

正しく読み上げることのできなかった文を読み上げることのできなかった原因によって分類する。分類される原因としては、辞書情報の間違い、形態素解析の解析ミス、音声記号生成規則の誤りの3つがある。

4-3 自動評価

辞書、形態素解析、音声記号生成のどれかに対して修正が加えられたとき、音声記号列に生じる変化を調べる。テンプレートとして蓄えてある正解文を音声記号に変換し、従来の記号列との比較を行う。このとき、正解の音声記号列は一つとは限らず複数存在する場合もある。

5. 評価

評価ツールを用いて新聞記事2.2万文、学会論文1.2万文、に対しての評価データの作成を行った。この評価データを用いることにより、従来の評価ツールを用いずに行っていた開発に比べて安定した修正が行えるようになった。

6. 今後の課題

今回の評価方式では評価データの作成に対して多くの人手と時間を必要とした。音声合成の場合各個人の生まれ育った環境や主観が大きく影響するため定量的に評価することはなかなか難しい。

しかし、偏りの少ないデータを大量に収集し、この評価ツールを使用することにより安定した合成が期待できる。

今後は会話文などの評価データを増やしていくことにより、より自由度の高い音声合成ソフトの開発に取り組みたい。

7. 参考文献

- [1]小林他 日本語解析処理を用いた音声読み上げ方式の開発 情報全大, 7N-6 (1992. 3)
- [2]原他 文音声合成ボードの試作 信学全大, A-231 (1992. 3)
- [3]岩田他 パソコン向けソフトウェア日本語テキスト音声合成 音学講論集 2-8-13 (1993, 10)