

音声合成モジュール

山下 洋一

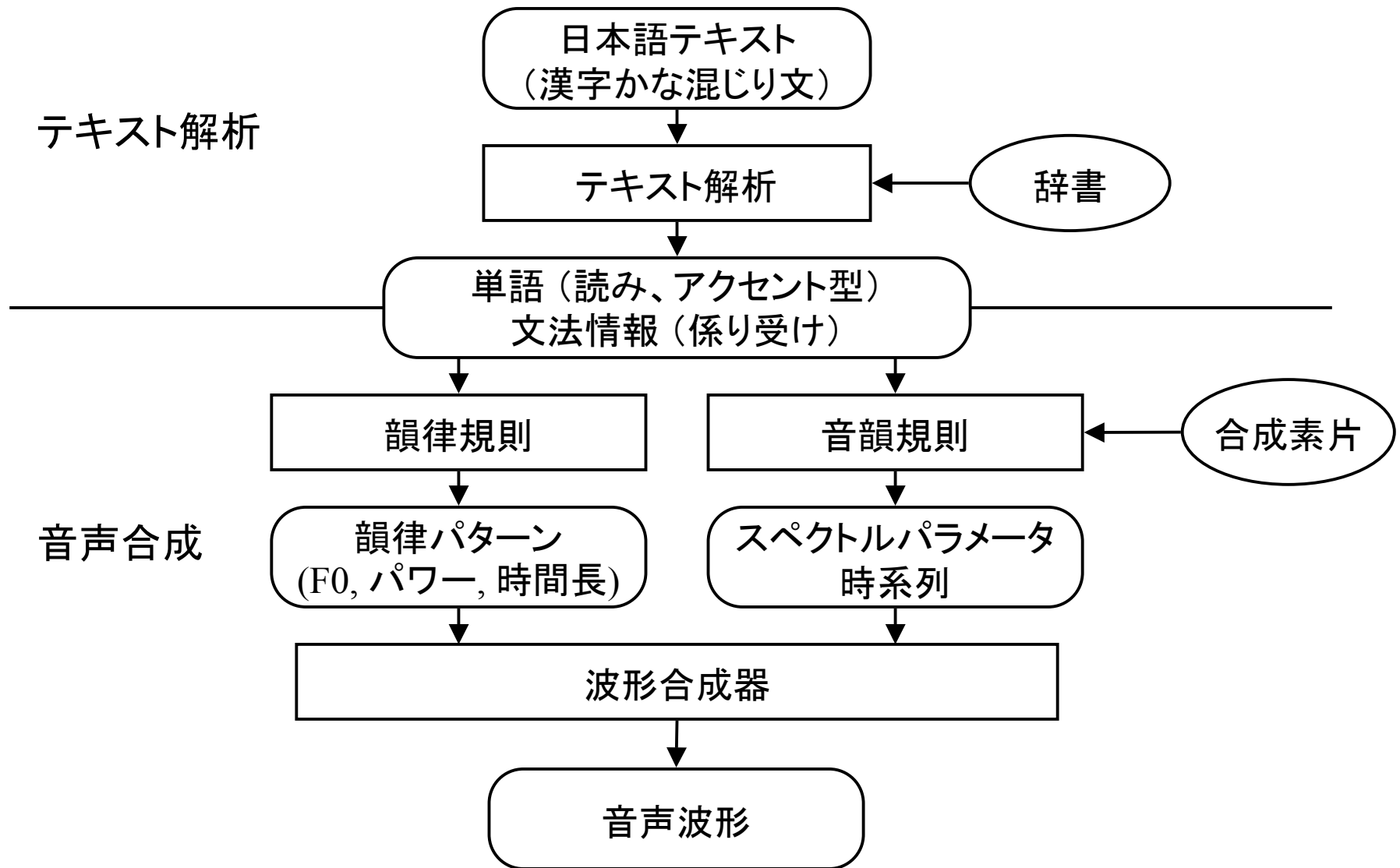
yama@media.ritsumeai.ac.jp

立命館大学 情報理工学部

2005年7月21日

- Galatea における音声合成モジュール
 - 単体の日本語テキスト音声合成システムとしても動作
 - 無償で利用可能、ソースコードの公開
 - 顔画像との同期
 - テキスト解析部、韻律生成部、波形生成器のモジュール化
 - 対話システムにおける音声合成
 - IPAのプロジェクトから継続

音声合成における処理の流れ



■ テキスト解析

- chasen の利用
- 辞書の整備: アクセント型、アクセント結合様式

■ 韻律生成

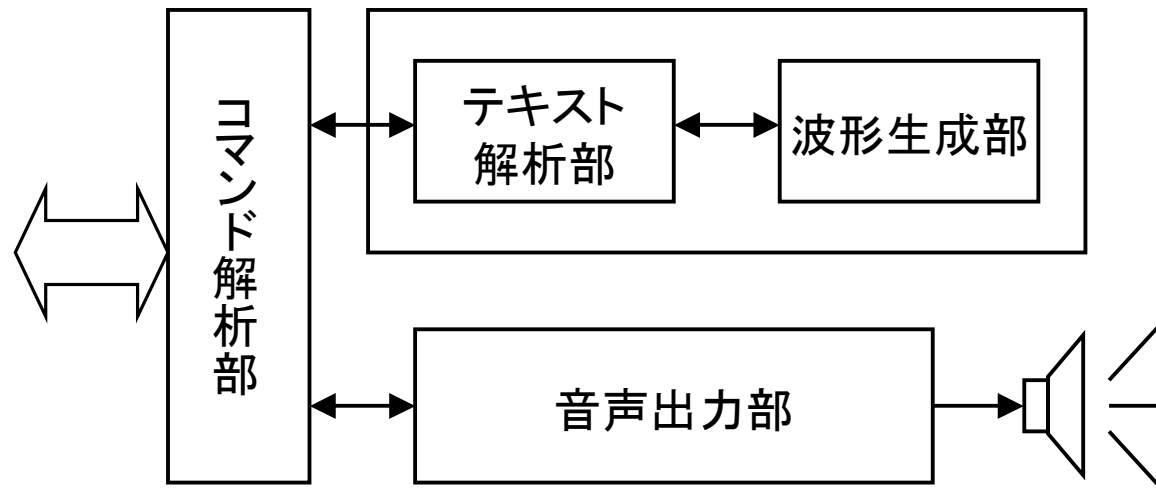
- アクセント結合: 先行研究を実装
- 韻律パラメータ: コーパスベース

■ 波形生成

- HMMに基づく音声合成

■ 構成

- chasen: 形態素解析
- chaone: 複合語処理、音韻交替処理
- gtalk: コマンド解析、韻律・波形生成、音声出力
 - HMMに基づいた音声合成



- 基本的な日本語テキスト音声合成
 - JEIDA-62-2000 日本語テキスト音声合成用記号の規格
- 音声出力の途中中断
 - 中断までに音声出力した音素列の出力
- 合成音声における音素列の時間長の出力
 - 顔画像出力との同期
- 話者の変更
- 韻律の制御
 - 基本周波数、パワー、発話速度の部分的変更

(当初の) 3年間の予定

1. 辞書の拡充

- 50,000語程度に
 - 当初、約23,000語

2. 音声出力処理の改良

- 合成音声出力開始までの時間短縮
 - 波形生成と音声出力の並行処理
 - ファイルにセーブされた音声データの利用

3. テキスト解析精度の向上

- 読み決定の精度向上
 - 読みわけ、単漢字の読み付与
- アクセント句決定の精度向上

4. 合成音声のカスタマイズ機能の実現

- 声質のカスタマイズ
- 韻律のカスタマイズ

前年度(1年目)の活動報告

1. 辞書の拡充

- サ変名詞の追加
- 語彙数: 約35,000
- 単漢字読み(第一水準のみ)

2. アクセント句生成、アクセント型決定処理の改良

- chaone の改良
- gtalk 内部処理の変更

3. ファイルにセーブされた音声の再生

- gtalk で合成しセーブしたファイルの利用

4. 韻律パラメータのファイルからの読み込み

- 音素時間長と基本周波数(F0)

5. XML 形式によるアクセント句情報の出力

今年度(2年目,2004.11~)の活動予定 (2005年2月総会より)

1. 辞書の拡充

- 約50,000語程度に

2. 音声出力ルーチンの改良

- 波形生成との並行処理

3. 読み決定処理、アクセント句処理の改良

- 宣言的記述によるアクセント処理
- ユーザによる辞書登録

4. 多様な声質での音声合成

- 話者モデルの学習

5. 韻律パラメータの柔軟な制御

- ユーザによる韻律パターンのカスタマイズ機能

進捗状況 (1)

1. 辞書の拡充

- unidic-1.1.0
 - 約23,000語
- unidic-1.1.1
 - サ変名詞の追加 (5,761語)
- unidic-1.1.2
 - 漢字2字、3字からなる一般名詞の追加 (4,788語)
 - 34,463語
- unidic-1.1.3(?) 近いうちに公開の予定
 - 固有名詞などの追加
 - 55,000～60,000語の語彙数になる見込み

2. 音声出力ルーチンの改良

- 波形生成との並行処理
 - 検討中

3. 読み決定処理、アクセント句処理の改良

- 宣言的記述によるアクセント処理
 - 完了
 - XMLベースでのアクセント句決定処理
- ユーザによる辞書登録
 - 検討中

アクセント句の決定

<S>

<AP orth="私は" pron="ワタシワ" aType="0" silence="NON">

<W2 orth="私" pron="ワタシ" pos="代名詞" aType="0" aConType="">

<W1 orth="私" pron="ワタシ" pos="代名詞" ... />

</W2>

<W2 orth="は" pron="ワ" pos="助詞-係助詞" ...>

<W1 orth="は" pron="ワ" pos="助詞-係助詞" .../>

</W2>

</AP>

<AP orth="音声合成器です" pron="オンセーゴーセーキデス" aType="7" silence="NON">

<W2 orth="音声合成器" pos="名詞-普通名詞-一般" ...>

<W1 orth="音声" pron="オンセー" pos="名詞-普通名詞-一般" .../>

<W1 orth="合成" pron="ゴーセー" pos="名詞-普通名詞-サ変可能" .../>

<W1 orth="器" pron="キ" pos="接尾辞-名詞的-一般" .../>

</W2>

<W2 orth="です" pron="デス" pos="助動詞" aType="0" ...>

<W1 orth="です" pron="デス" pos="助動詞" cType="助動詞-デス" .../>

</W2>

</AP>

</S>

アクセント句決定規則の記述

<!-- フィラーは単独のアクセント句に -->

```
<rule id="0" priority="1">
```

```
  <cond>
```

```
    <current pos="感動詞-フィラー"/>
```

```
  </cond>
```

```
  <then ap="alone"/>
```

```
</rule>
```

<!-- '?' の前は別のアクセント句に -->

```
<rule id="1" priority="50">
```

```
  <cond>
```

```
    <current orth=" ? "/>
```

```
  </cond>
```

```
  <then ap="new"/>
```

```
</rule>
```

...

アクセント型決定規則の記述

```
<!-- 接頭辞(pre) + 自立語(current) -->
<rule id="P1">
  <!-- 一体化型 -->
  <if>
    <pre aConType="P1"/>
  </if>
  <then>
    <if>
      <current no_nucleus="yes"/>
    </if>
    <then aType="0"/>
    <else aType="pre/@mora + current/@aType"/>
  </then>
</rule>
```

...

4. 多様な声質での音声合成

- 話者モデルの学習
 - 検討中

5. 韻律パラメータの柔軟な制御

- ユーザによる韻律パターンのカスタマイズ機能
 - 検討中

進捗状況 (4)

- Linux版とWindows版のソースの共有
 - 完了