

音声合成モジュール

山下 洋一

yama@media.ritsumei.ac.jp

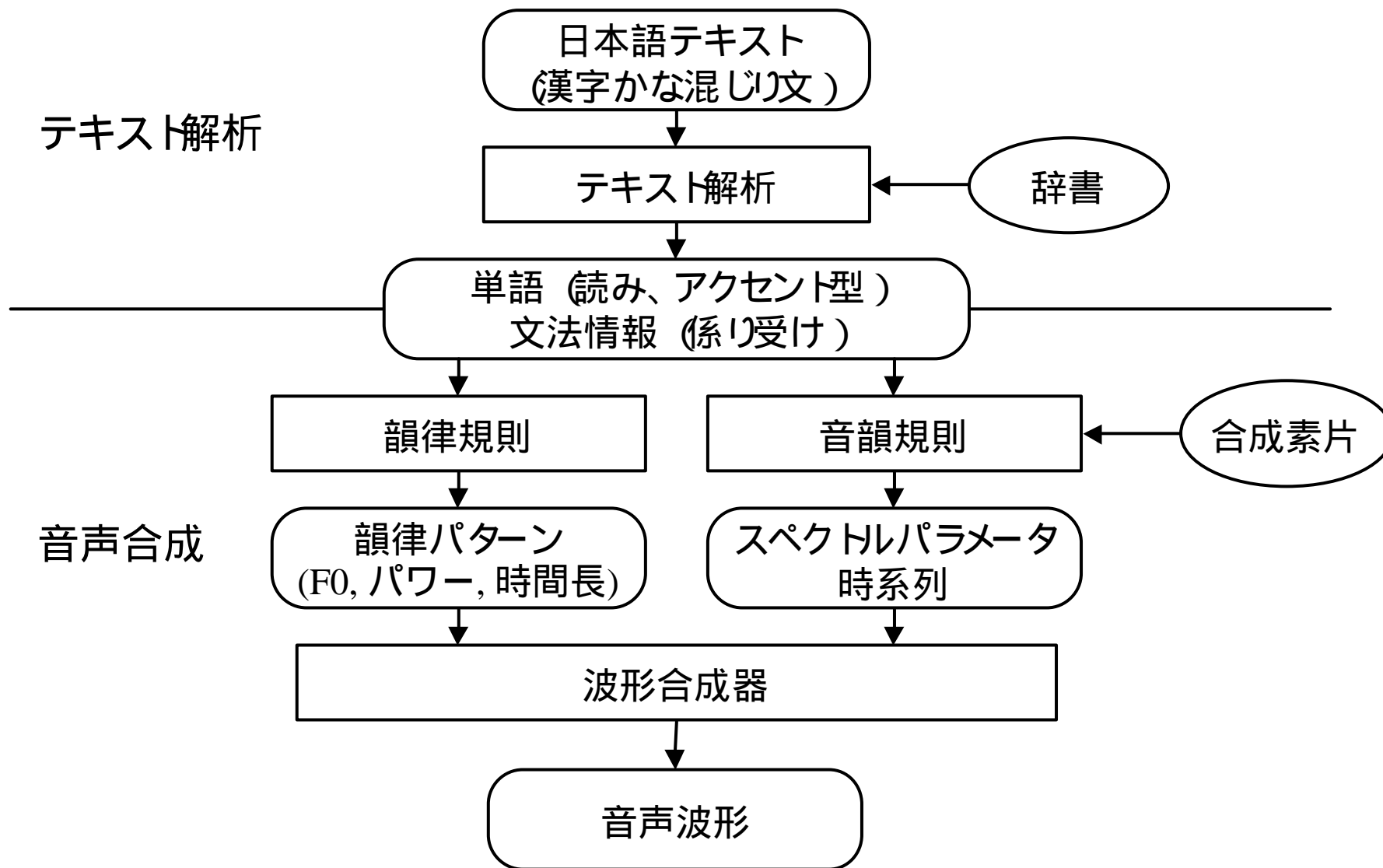
立命館大学 情報理工学部

2007年3月9日

音声合成モジュール: GalateaTalk

- 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



GalateaTalk での処理

■ 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

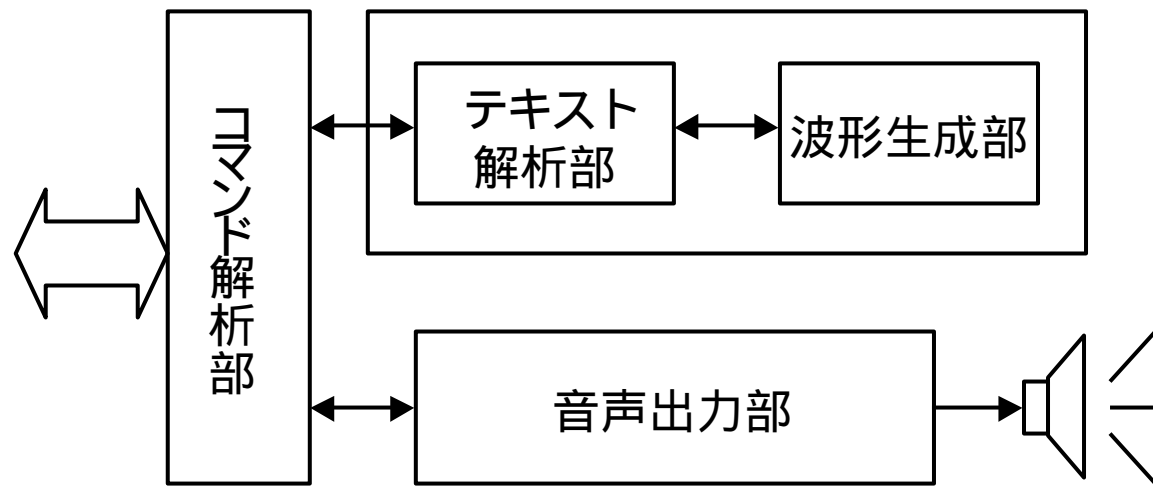
■ 波形生成

- HMMに基づく音声合成

GalateaTalk の構成

■ 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



GalateaTalk の機能

- 基本的な日本語テキストト音声合成
 - JEIDA-62-2000 日本語テキスト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

初年度(1年目, 2003.11 ~)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

前年度(2年目,2004.11~)の活動報告

1. 辞書の拡充

- 語彙数: 約35,000

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

6. Linux版とWindows版のソースの共有

今年度(3年目,2006.4~)の活動予定 (2006年3月総会より)

1. 辞書の拡充

- 固有名詞などの追加
- 55,000 ~ 60,000語

2. 多様な声質での音声合成

- 話者モデルの学習

3. 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

活動報告 (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-1.2.0
 - 名詞 , 固有名詞 , 動詞などを中心とした追加
 - 124,671語
- unidic-1.3.0pre2

活動報告 (2)

2. 多様な声質での音声合成

- 話者モデル作成スクリプトの提供
 - 音声収録 , 分析 , 話者モデルの作成
 - 夏の講習会でも利用

3. 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
- GalateaTalk の生成した韻律情報ファイルに対する修正ツール (ProsEdit) の提供

活動報告 (3)

■ 韻律修正ツール ProsEdit

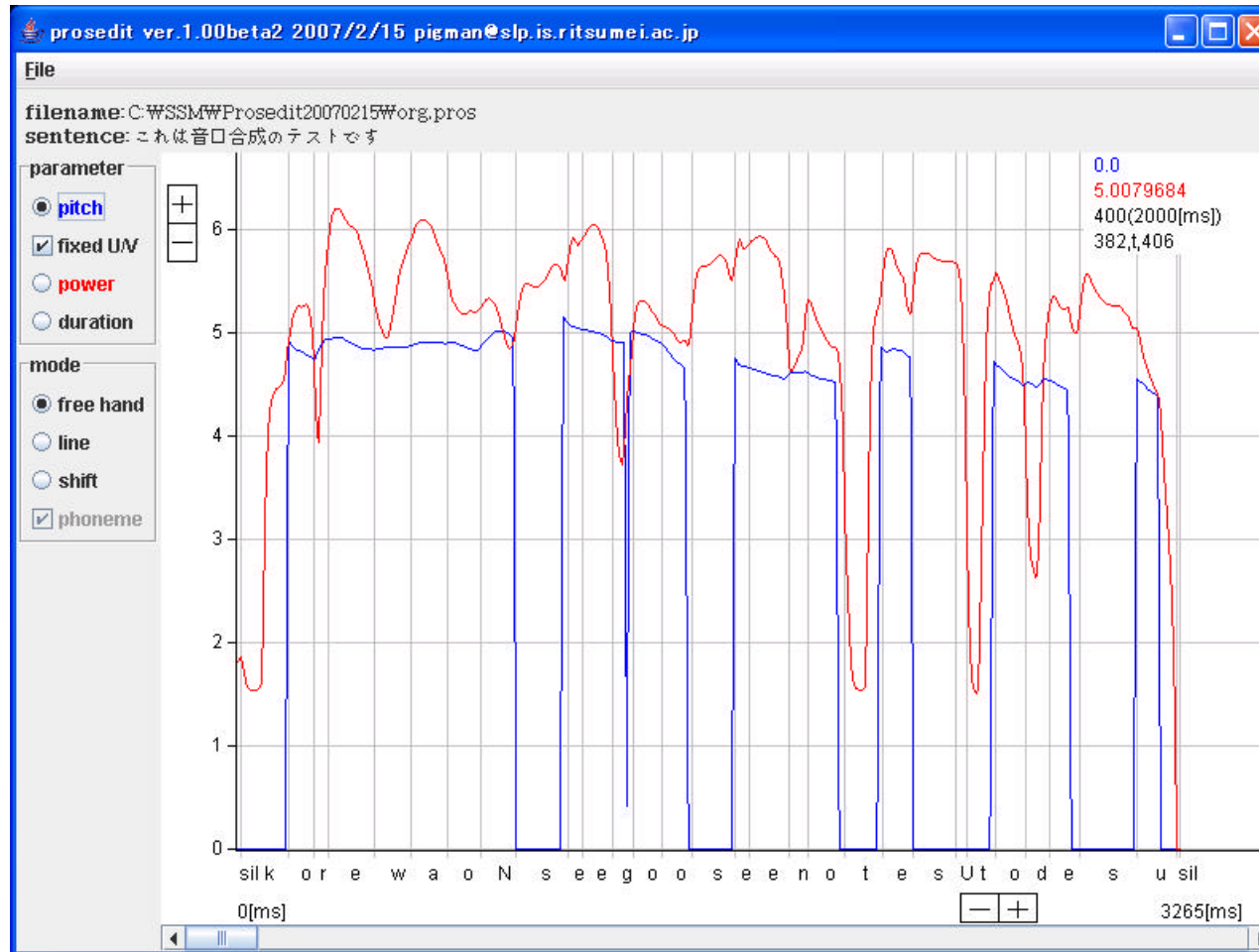


図 ProsEdit の動作画面

次年度(4年目, 2007.4~)の活動予定

1. 多様な声質での音声合成

- 話者モデルの学習

2. 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
- gtalk での時間長変更への対応
- gtalk のフロントエンドに

3. 辞書の整備

- unidicプロジェクトとの連携
 - 3月末に unidic-1.3.0 正式版をリリース予定
 - chaone (音韻交替処理, アクセント処理) の開発