

# 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

2008年3月7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

2

ISTC総会 2008/3/7@ ASTEM

## GalateaTalk での処理

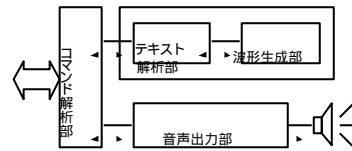
- テキスト解析
  - chasen の利用
  - 辞書の整備 :アクセント型、アクセント結合様式
- 韻律生成
  - アクセント結合 :先行研究を実装
  - 韻律パラメータ:コーパスベース
- 波形生成
  - HMMに基づく音声合成

3

ISTC総会 2008/3/7@ ASTEM

## 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: コメント解析、韻律・波形生成、音声出力
    - HMMに基づいた音声合成



4

ISTC総会 2008/3/7@ ASTEM

## GalateaTalk の機能

- 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

5

ISTC総会 2008/3/7@ ASTEM

## (当初の) 3年間の予定

- 辞書の拡充
  - 50,000語程度に
    - 当初、約23,000語
- 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
- 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
- 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

6

ISTC総会 2008/3/7@ ASTEM

## 初年度 (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

7

ISTC総会 2008/3/7@ ASTEM

## 第2年度 (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

8

ISTC総会 2008/3/7@ ASTEM

## 第3年度 (2006.4~) の活動報告

1. 辞書の拡充
  - 名詞, 固有名詞, 動詞などを中心とした追加
  - UniDic1.3.0pre2
2. 多様な声質での音声合成
  - 話者モデル作成スクリプト
3. 韻律パラメータの柔軟な制御
  - 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

9

ISTC総会 2008/3/7@ ASTEM

## 2007年度 (第4年度) の活動予定から (2007年3月総会)

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

10

ISTC総会 2008/3/7@ ASTEM

## 活動報告 (1)

1. 多様な声質での音声合成
  - 話者モデル作成スクリプト
    - 音声収録, 分析, 話者モデルの作成
  - 2007年度夏の講習会での利用
    - 作成された話者モデルの品質向上

11

ISTC総会 2008/3/7@ ASTEM

## 活動報告 (2-1)

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
  - 韻律情報ファイルに対する修正ツール (ProsEdit) の改良
    - ver 1.00 → ver.1.10
    - GalateaTalk の呼び出し

12

ISTC総会 2008/3/7@ ASTEM

## 活動報告 (2-2)

### ■ 韻律情報ファイル

```
input_text: こんにちは
spoken_text: こんにちは
number_of_phonemes: 11
total_duration: 680
-----
sil [10]
k [100]
o [85]
N [90]
n [25]
i [60]
ch [105]
i [50]

w [50]
a [95]
sil [10]
-----
total_frame: 136
-----
0: 0.000000 1.723706
1: 0.000000 1.779953
2: 0.000000 1.819527
3: 0.000000 1.737764
4: 0.000000 1.607652
...
```

13

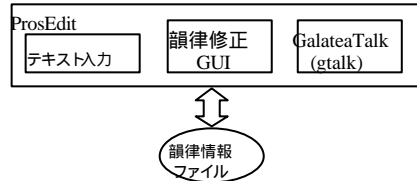
ISTC総会 2008/3/7@ ASTEM

## 活動報告 (2-3)

### ■ ProsEdit 1.00



### ■ ProsEdit 1.10



14

ISTC総会 2008/3/7@ ASTEM

## 活動報告 (2-4)

### ■ 韻律修正ツール ProsEdit

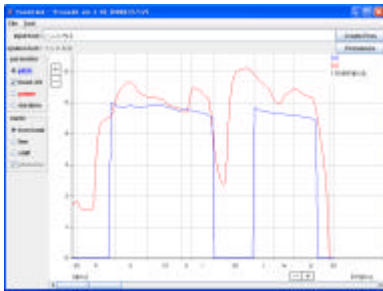


図 ProsEdit の動作画面

15

ISTC総会 2008/3/7@ ASTEM

## 活動報告 (3-1)

### 3. 辞書の整備

- UniDic プロジェクトからの辞書の提供
- UniDic 1.2.0pre (2006年4月)
  - 約46,000語
- UniDic 1.3.0 (2007年4月)
  - 約106,000語
- UniDic-chasen 1.3.5 (2007年10月)
  - 約110,600語
  - 連体形と終止形を区別
  - 名称を UniDic-chasen に変更 (UniDic-mecab の開発を開始)

16

ISTC総会 2008/3/7@ ASTEM

## 活動報告 (3-2)

### 3. 辞書の整備

- ChaOne 1.3.2
  - 一部の環境におけるrun time errorへの対処
    - 属性ノードの追加と子ノードの追加の順序を明示的に指示
  - ChaOne 1.3.1の音変化処理結果のチェックと不具合の修正
    - 助数詞「日」に対するad hocな対応 (本来は語彙選択の問題)
  - アクセント句生成規則の不具合への対応

17

ISTC総会 2008/3/7@ ASTEM

## 活動報告 (4)

### 4. GalateaTalk における機能追加

- アクセント句を強制分割するタグの導入
  - アクセント句生成規則の不具合
  - アクセント句境界を<APB/>タグで指定可能に
  - 例  
山下洋一 → 山下<ARB/>洋一

18

ISTC総会 2008/3/7@ ASTEM

次年度(5年目,2008.4~)の活動予定

---

1. 韻律修正ツールの改良・機能追加
  - ユーザによる韻律パラメータのカスタマイズ機能
  - GalateaTalk のフロントエンドに
2. 多様な声質での音声合成
  - 話者モデルの学習
3. 辞書の整備
  - UniDicプロジェクトとの連携