

2005年日本心理学会ワークショップ
心理学者のための音声収録・分析・呈示テクニック

合成音声に対する印象は 呈示条件で変化するか？ —携帯電話の音声—



芝崎朱美
産業技術総合研究所

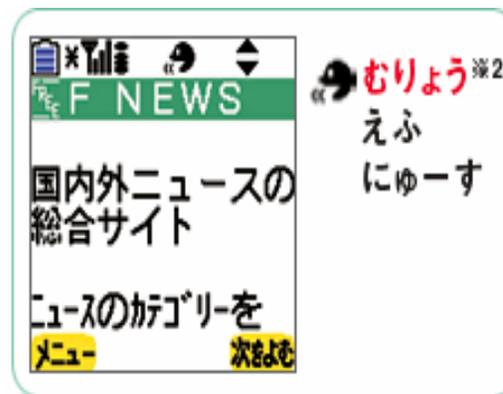
本日の内容

- 研究の背景
- 合成音声の紹介 2, 3点
- 例 携帯電話を使った実験での刺激音声作成

研究の背景

音声を使った製品やサービス

合成音声
による読み
上げ



屋内での利用

屋外での利用

より便利に！

現実

理想

GAP



合成音声技術
の向上



携帯電話への応用
音声読み上げ機能

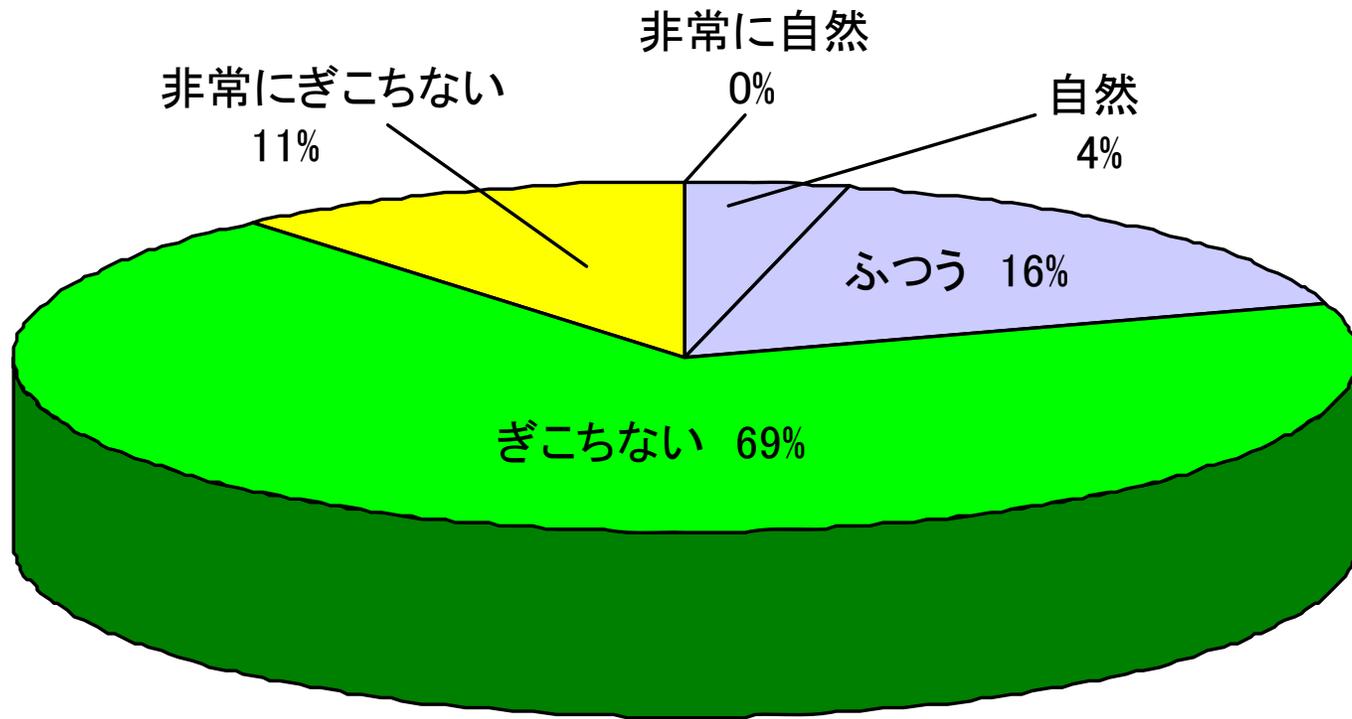


デジタルデバ
イド解消！
高齢者や障害
者にやさし
い！

興味あるけど、
不自然そう...
使いにくそ
う...



コンピュータの音声合成による「読み上げ音声」の自然さ



2003/9/18 20代～60代のインターネット機能付き携帯電話ユーザー300人を対象とした調査
「携帯を使ったコンテンツの「音声読み上げサービス」、4人に1人は興味あり」より

<http://japan.internet.com/research/20030919/1.html>

合成音声を用いた製品の利用・販売が先行

問題意識

聞き取りやすさ
使いやすさ
など、検証されているのか？

検証実験

- ① 合成音声の印象評価(ガイドラインに沿って)
- ② 聴取条件での印象の変化
- ③ 聴取条件での了解度の差



PC上で再生される音声の聴取印象

- Protalker 97 (IBM)
- Finespeech (富士通)
- 携帯電話 ラクラクホン

→ 純粹に音情報だけで印象を比較・・・①



携帯電話の読み上げ音声の聴取印象

- 静条件 (椅子に座って聴き取り)
- 動条件 (階段昇降しながら聴き取り)

→ 携帯電話を利用する実環境における印象

聴取者が動かない場合・・・②

聴取者が動く場合 (より実環境に近い状況)・・・③

合成音声の紹介

IBM Protalker97

- ホームページ：
<http://www6.ibm.com/jp/voiceland/pt20/intro.html>
- 特長：
 - 人間に近い自然な声で日本語の文章を読み上げる
 - クリップボードに文章をコピーしたり、ドラッグ & ドロップ操作だけで読み上げる
 - 男性・女性、読み上げる速さ、声の高さなどを自由に設定
 - 固有名詞や特殊な専門用語も、ユーザー辞書に登録できるので、とても便利です。また、簡単な英単語なら、ユーザー辞書に登録することなく、読むことが可能。
 - 読み上げた音声をWAVE形式のファイルにする
- 価格：
ProTALKER 97 V2.0 本体価格 9,800円

クックポートリーダーの設定

[声の設定]

キャラクター  読み秀君

遅 速度 速 小 音量 大

低 ピッチ 高 弱 抑揚 強

これはテストですよ。 テスト

- クックポートリーダーを手前に表示(吹き出しウィンドウを表示)
- 合成音声をWAVファイルに出力

[読み上げのタイミング]

- 複写後すぐに読み上げ開始
- フレイ・マークを押して読み上げ開始

[息継ぎ位置の設定]

- 自動
- 記号による指定

[数字の読み方]

- 自動
- 桁読み
- 棒読み

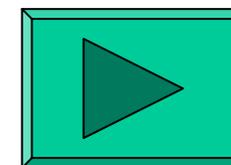
OK キャンセル デフォルト

男声か女声か

話速

ピッチ

音声ファイル
作成



Animo Finespeech



- ホームページ: <http://www.animo.co.jp/product/fs/>

- 特長

- 感情音声合成のサポート (V2.1 Emotional Option)
- 幅広いプログラミングインターフェース

- 価格

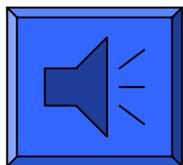
| | |
|--------------------------------------|----------|
| FineSpeech SDK V2.1 | ¥200,000 |
| FineSpeech SDK V2.1 Emotional Option | ¥100,000 |
| FineSpeech Basic SDK | ¥50,000 |

- 応用例

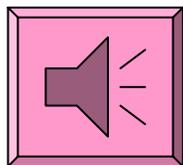
- 自動応答システム
- カーナビゲーション
- メールの内容読み上げ
- 各種音声ガイダンス
- ボイスポータルサービス など

Finespeechのデモンストレーション

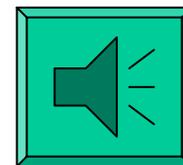
<http://www.animo.co.jp/product/fs/demo.html>



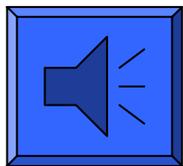
男声8kHz



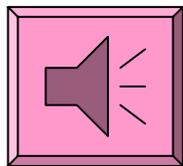
女声8kHz



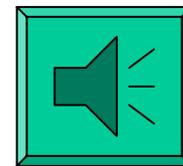
平常(女声)



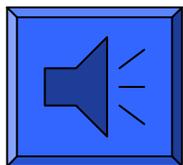
男声16kHz



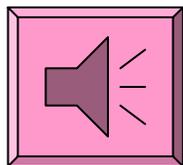
女声16kHz



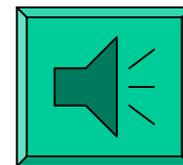
喜(女声)



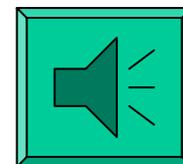
男声22kHz



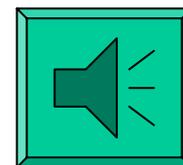
女声22kHz



怒(男声)



悲(女声)



悲(男声)

携帯電話 ラクラクホン



● 音声読み上げボタン

表示中の操作説明や、iモードサイトの表示内容などを音声で確認できます。

国内外ニュースの
総合サイト
ニュースのカテゴリを
メニュー

むりょう^{※2}
えふ
にゅーす



● ホームページ:

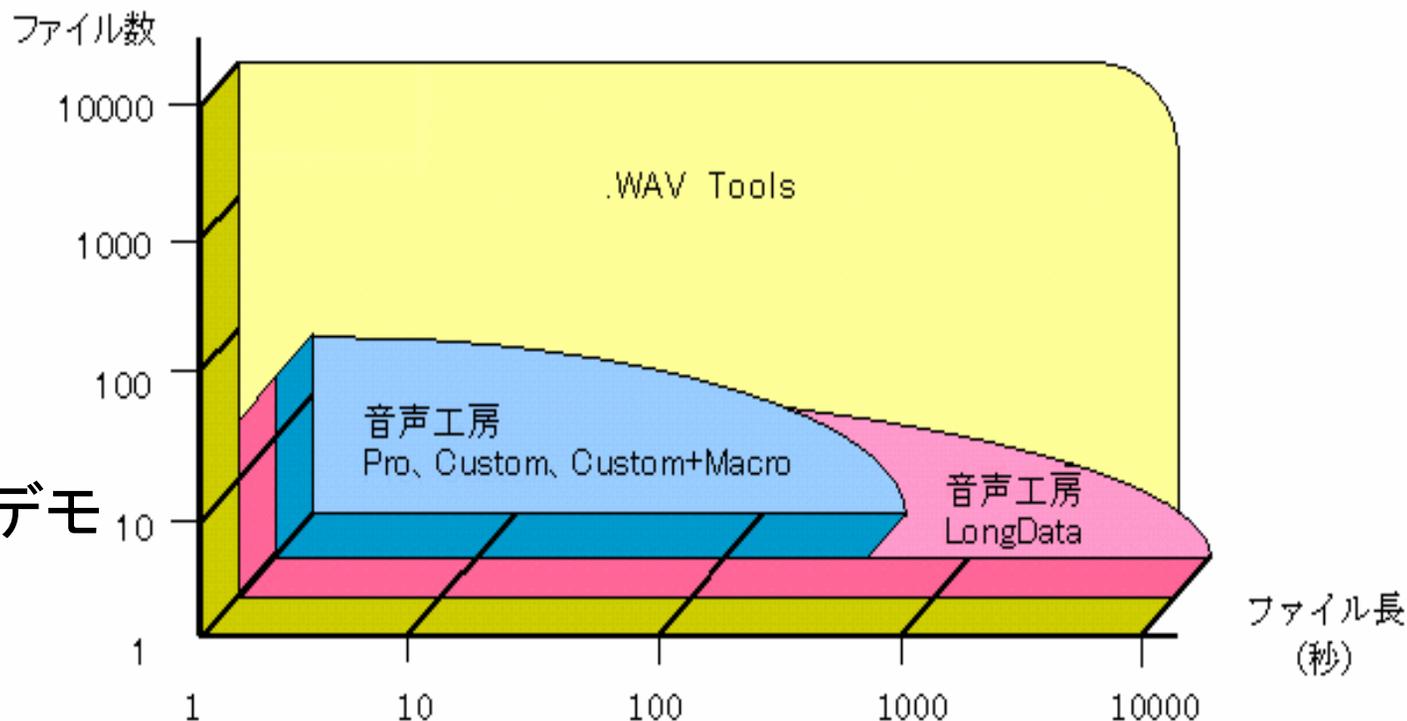
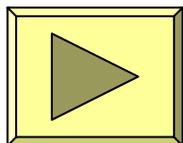
- <http://www.fmworld.net/product/phone/f672i/index.html> (富士通)
- http://www.nttdocomo.co.jp/product/easy_phone/ (NTTドコモ)



音声工房 (NTTアドバンステクノロジー株式会社)

- ホームページ:
<http://www.sp4win.com/>
- 特長: 扱いたい音声の特徴別に適用領域の異なるラインナップが充実している.

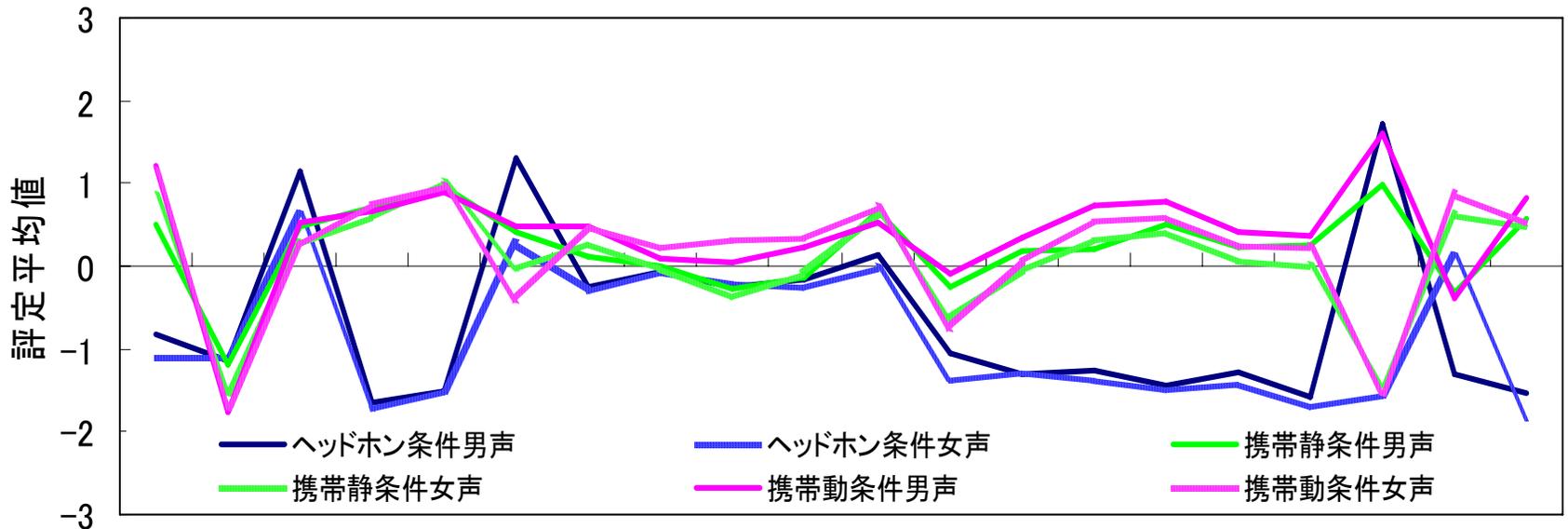
音声工房録音デモ



実験条件による印象評価の違い (らくらくホンのみ)

携帯電話の音声として
適している

内容がよくわかる
読み取りがある
力強い
きれいな
はっきりした
太い
リズムが自然な
しゃべり方が速い
抑揚が自然な
アクセントが自然な
言葉が流暢な
人間らしい
好き
総合的に良い
聴きやすい
快い
気持ちの良い
男声らしい
高い声

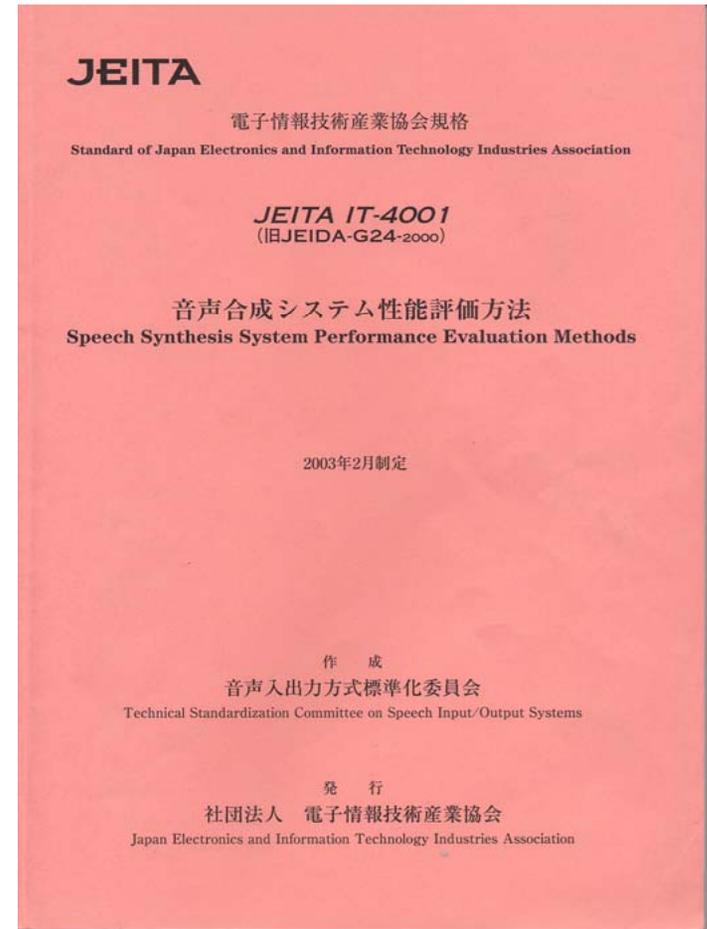


おまけ

印象評価の統制 ①

聴取印象をいかに統一して評価するか？

- 合成音声システム性能評価方法
(電子情報技術産業協会規格 JEITA IT-4001)
- 日本語音声合成システムの性能評価方法を規定するもの
 - 正読率
 - 了解性
 - 総合評価など, 規定項目について



8000円程度