

音声イメージの 製品開発への応用

2000/11/8

オムロン(株) IT研究所

大本浩司

発表内容

- 1) オムロンにおける音声関連事業
- 2) 心理学的知見の応用
- 3) Media Equation
- 4) Voice User Interface
- 5) 音声対話システムの開発プロセス
- 6) 音声対話システムにおけるユーザ評価
- 7) オムロンの音声合成技術
- 8) まとめ

電話音声自動応答装置

(<http://www.omron.co.jp/cma/index.html>)

- 1) チケット予約
- 2) 資料請求受付
- 3) 株価照会(野村証券さま)
- 4) 電話の自動振り分け(SONYさま)
- 5) 生産状況と納期の回答(DELLさま)



オムロンの音声関連事業と技術開発



【事業】

オムロン株式会社 事業開発本部 CMAプロジェクト
The Customer Management Automation Project
CMA プロジェクト

サポート

【技術】

オムロン株式会社 技術本部 IT研究所



音声対話研究室

- ・Voice User Interfaceの設計/評価技術
- ・音声合成技術
- ・対話制御技術
- ・音声認識技術

1. デザイン指針 (心理学的理論)

設計者は、どのように音声対話システムをデザインすればよいのか

2. 印象評価 (心理学的実験手法)

ユーザは、音声対話システムを使った時、どのような印象を受けているのか

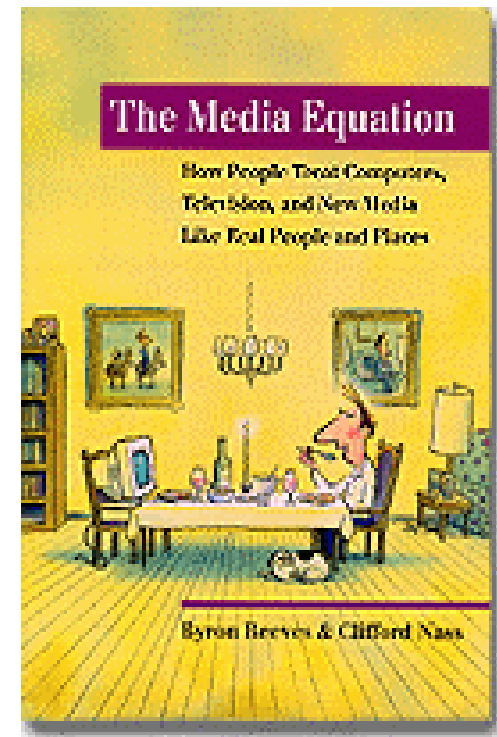
Media Equation

「人間とテクノロジー(メディア)のインタラクションは人間同士のインタラクションと基本的に同じである」

メディアの設計に社会的ルール(パーソナリティ、礼儀、お世辞など)を適用することが重要であると主張

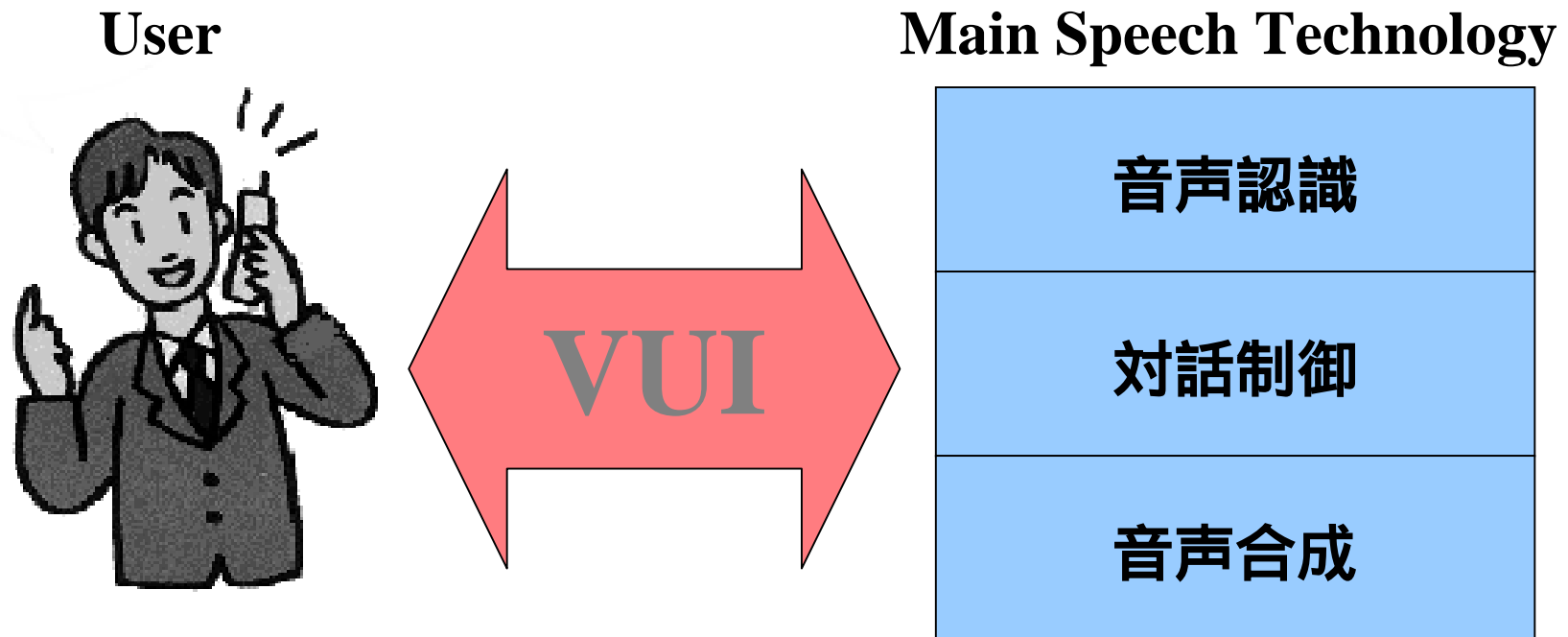
Stanford Univ.

Byron Reeves & Clifford Nass

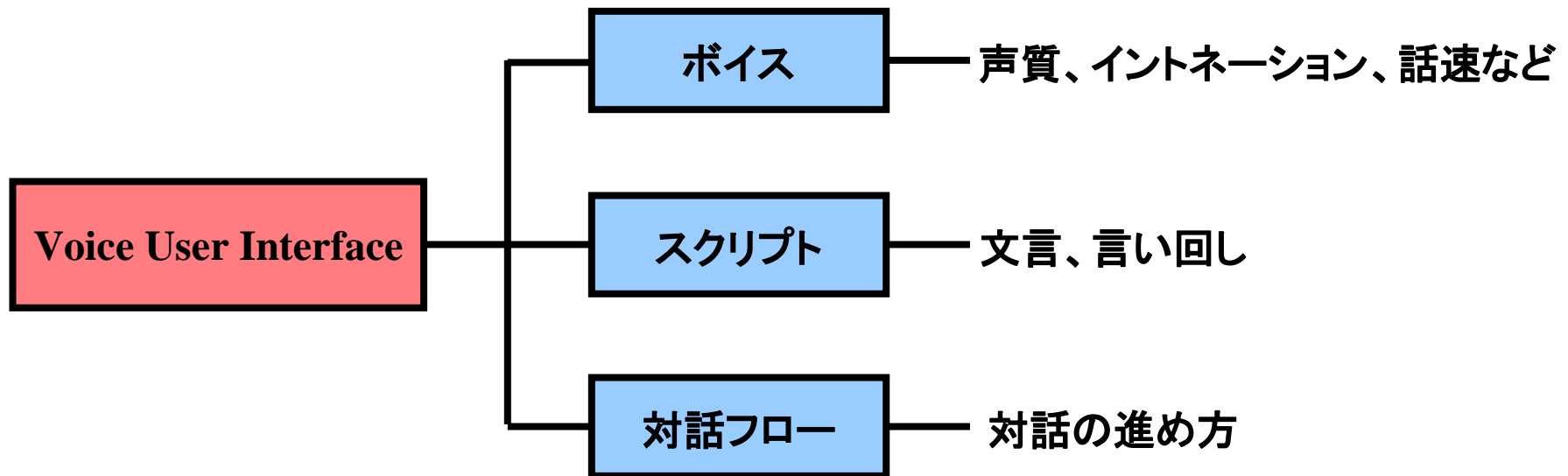


Voice User Interface (VUI)

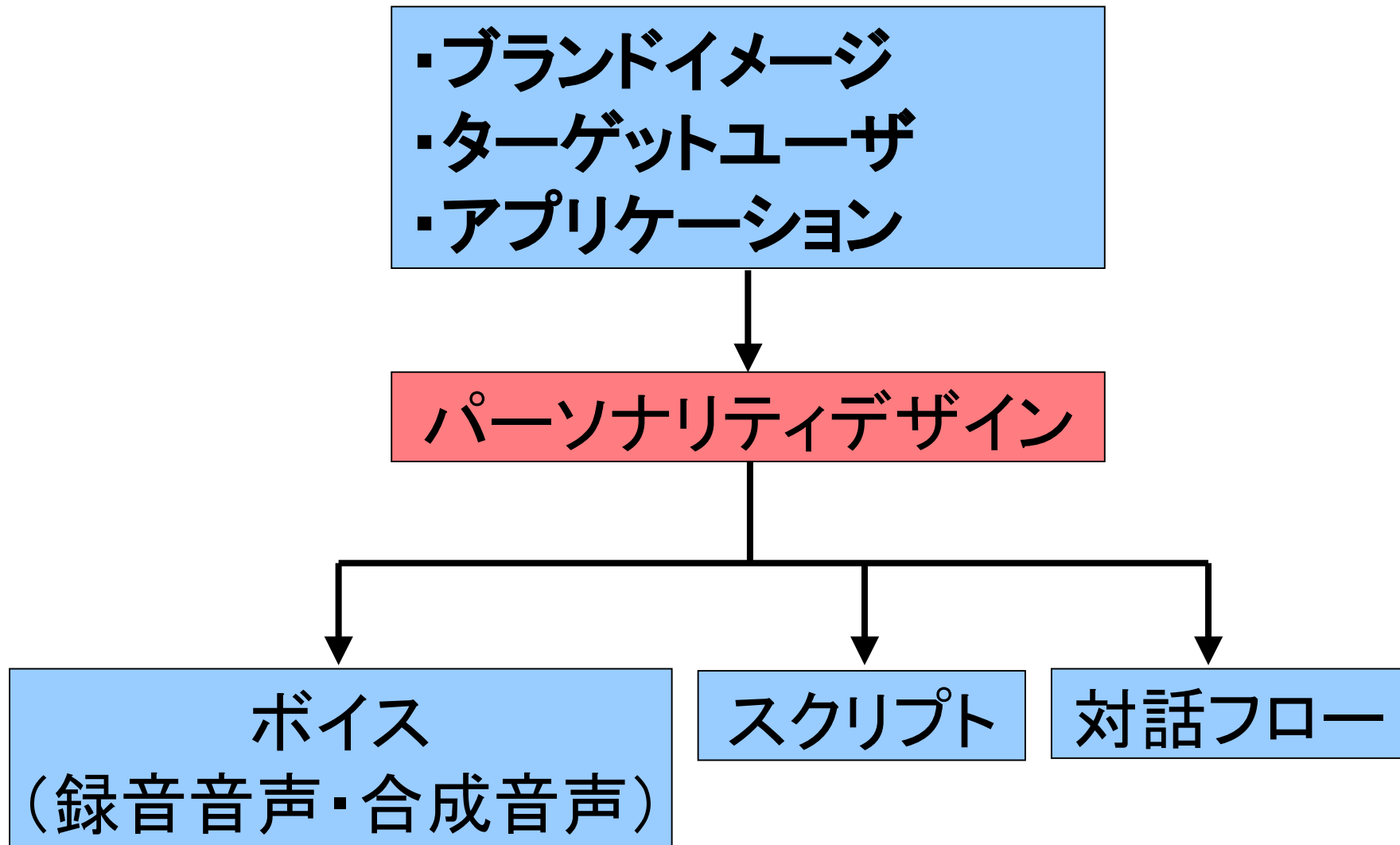
- Voice User Interface は、Graphical User Interface と同様に重要である



Voice User Interfaceの構成要素



Voice User Interfaceの設計手法

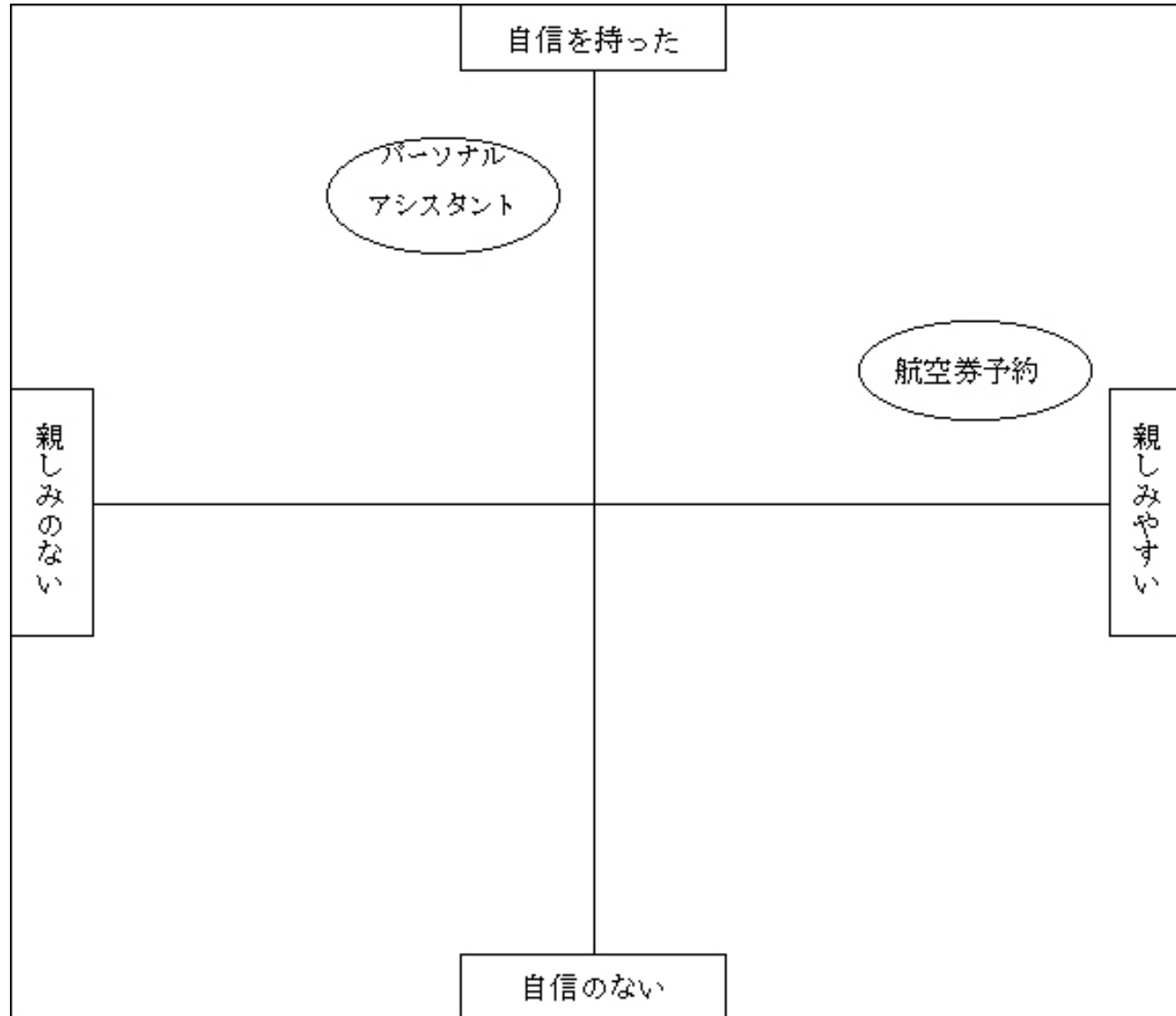


音声対話システムの開発プロセス



開発ステップ	開発成果物
(1)システムとパーソナリティの明確化	・システムの要求仕様
(2)パーソナリティデザイン、および声の選定	・パーソナリティマップ ・パーソナリティシナリオ ・声の選定結果
(3)対話フローの作成	・フローチャートあるいは状態遷移図 ・対話制御に用いる発話内容定義リスト
(4)発話内容の作成	・発話内容リスト
(5)声の録音と編集	・録音音声
(6)アプリケーション開発	・システム
(7)パイロットデータの収集	・評価データ
(8)データ分析およびチューニング	・課題リスト ・システムの改訂版
(9)システムの導入およびメンテナンス	・システムのログ結果の報告書 ・改善報告書

パーソナリティマップ



パーソナリティシナリオの例



[設定されたパーソナリティの背景情報]

- ・彼女の名前は中谷よしこ、勤めて1年目のオペレータで、21歳である。

[システムの目標]

- ・彼女の業務は、電話対応であり、資料の発送を行うことである。

[システムの振る舞い方]

- ・対応は、丁寧で印象が良く、いい加減な対応はしない。
- ・自分の仕事には責任を持ち、他人に頼ることは少なく、自分で最後までやり終える。
- ・電話はよくかかってくるので、できるだけ短時間で対応しようと努力している。
- ・お客様の要求が理解できない場合は、よりスキルのある人に対応を任せる。

[対象ユーザ]

- ・お客様の年齢層は広く、はじめて電話する人が多い。

[具体的なユーザとシステムのインタラクション]

システム:お電話ありがとうございます。資料発送です。資料番号をおっしゃって下さい。

ユーザ:04269

システム:お名前をお願いします。

ユーザ:山田たろう

システム:次に、住所をお願いします。

ユーザ:京都府長岡京市伊賀寺...

ボイスサンプル

- ナレーション調の録音音声
- パーソナリティデザインを行った録音音声

- 設計者に
➡ インタフェース設計の判断基準
- 音声対話システムに
➡ 一貫性と良い印象

- ユーザの視点からの評価技法
 - 1) 主観評価
 - 2) パフォーマンス評価
 - 3) インタラクション評価

ボイス
(録音音声・合成音声)

スクリプト

対話フロー

- 波形接続型音声合成

[特徴]

- 1) 録音音声の話者らしい1個性的な音声を再現できる
- 2) パーソナリティデザインに合った特定の音声を設定できる (規則合成方式では困難)

合成音声の利用例

- ・住所認識の自動応答対話

システム：「お客様のご住所をお願いします。」

ユーザ：「東京都練馬区中村1 - 2 - 3。」

システム：「ご住所は『東京都練馬区中村1 - 2 - 3』

よろしいですか？

ユーザ：「はい。」



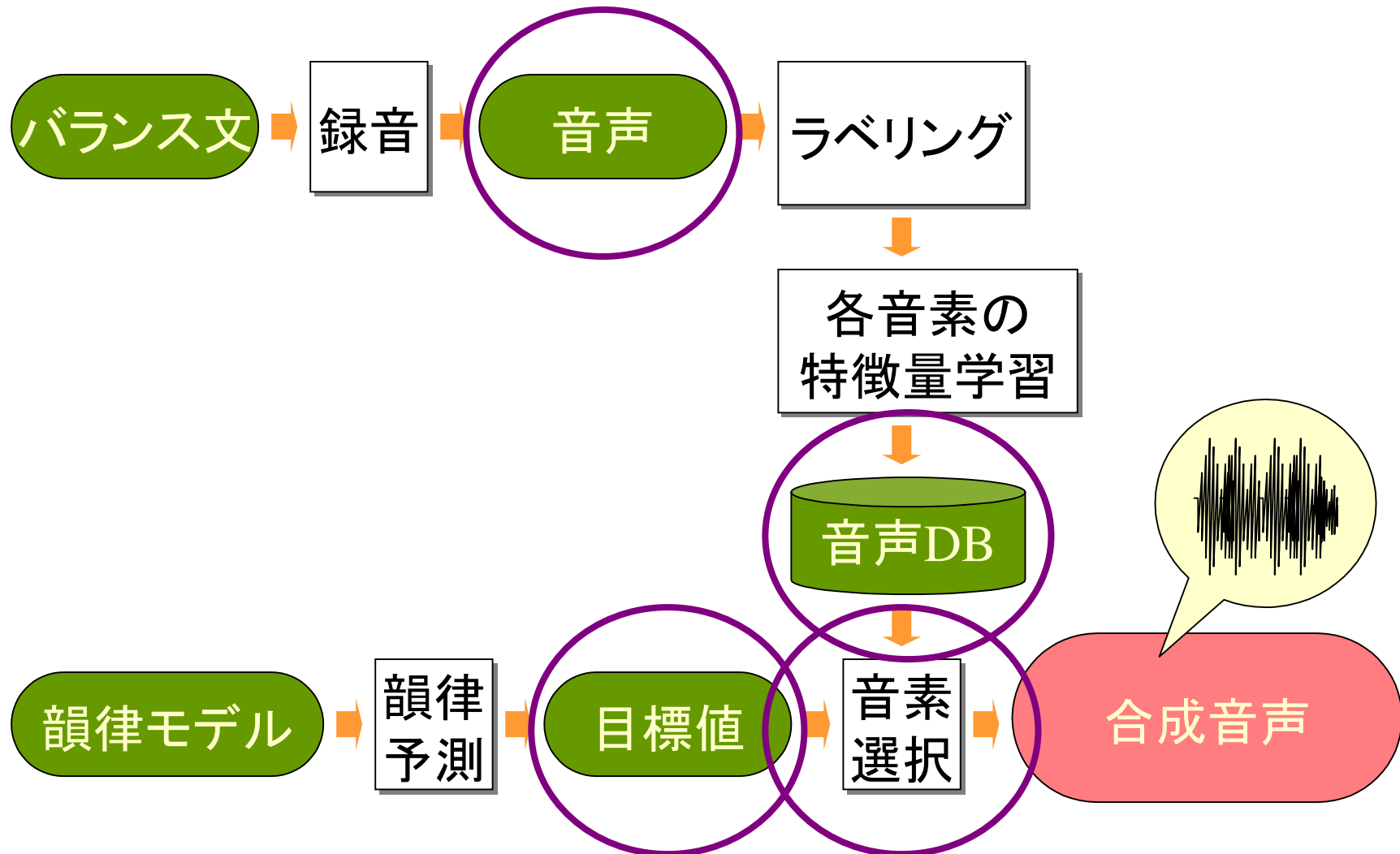
合成音声

目的:

- 1) 現状の技術レベルを把握
- 2) 不具合個所の特定
- 3) 改善効果の検証
- 4) 改善方法の明確化
- 5) パーソナリティデザインとの適合性

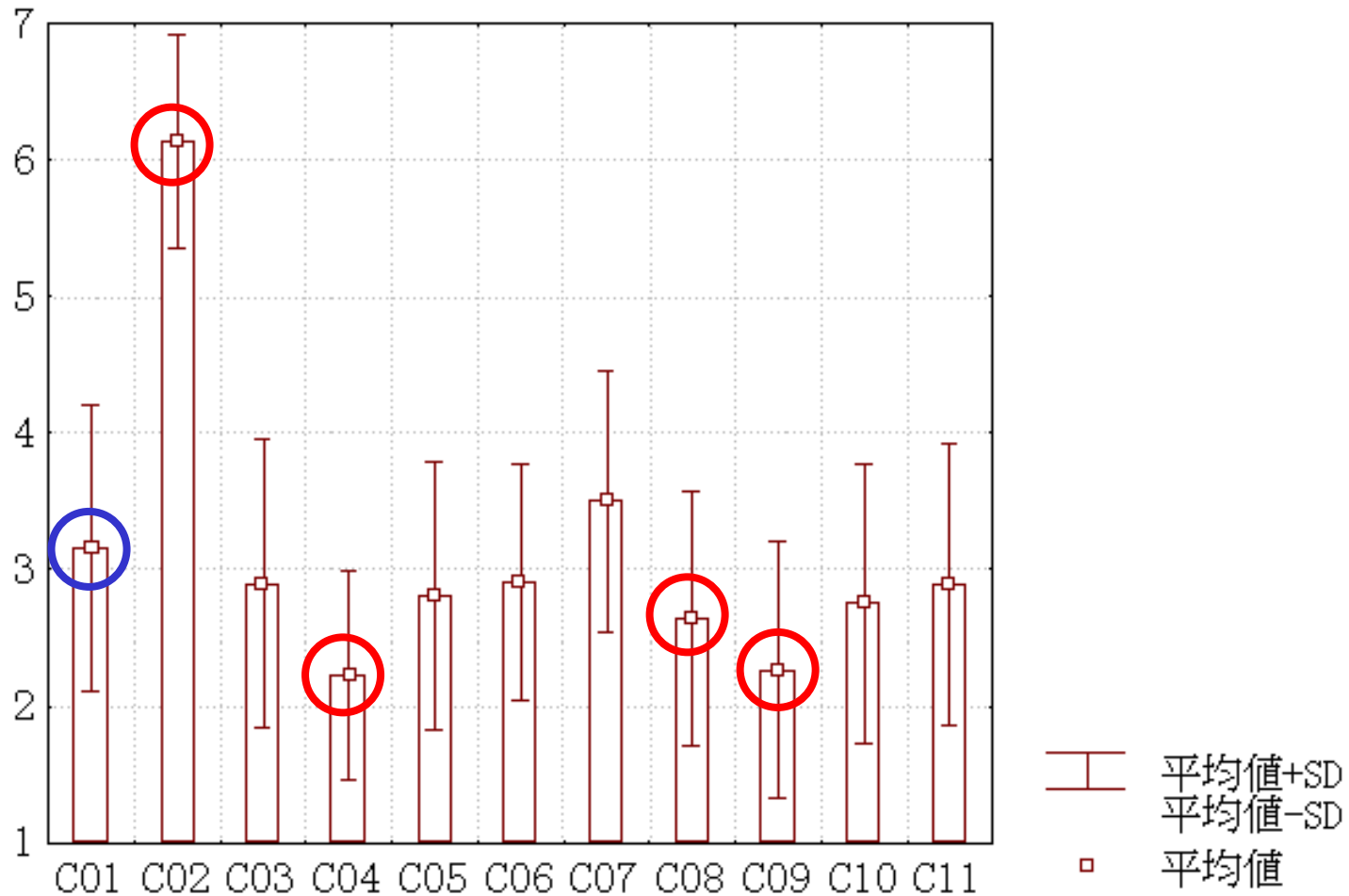
波形接続型音声合成の方法

○ = チューニングポイント



各刺激音声における「親しみやすさ」の程度

- ・C01(基準合成音声)は C02(録音音声)、 C04(音素接合無視)、 C08(音声DB1/5)、 C09(音声DB1/10)との間に有意差が認めれた。



まとめ

- 1) 心理学的知見は、音声対話システムの開発において、設計指針および評価で役立っている
- 2) 音声から伝わるパーソナリティの印象を考慮して設計すれば、より印象の良い音声対話システムが実現できる
- 3) 評価に実験心理学的な手法を適用することで、音声合成のチューニングパラメータが「親しみやすさ」に与える影響を明確にできる

心理学的な実験研究により、「音声から伝わるパーソナリティの印象」と「物理的な音響的な特徴」の関係を明確にすれば、下記のようなところで音声対話システムの開発に役立てられる

- a) ボイスタレントの選定
- b) 音声録音時のディレクション
- c) 音声合成の「間」、「高さ」、「話速」等の制御方法