

音声言語シンポジウム10周年企画
「音声言語研究関連分野の10年の歩み」
音声対話システム・音声インタフェース

(株)ホンダ・リサーチ・インスティテュート・ジャパン

中野 幹生

nakano@jp.honda-ri.com

1999年には何が研究開発されていたのか

研究

開発

音声対話システムの基本構成が共有化(次頁)

マルチモーダル入力
(OGI等)
マルチモーダル出力
(KTH等)

VoiceXML Forum
立ち上げ(3月)
0.9策定(9月)

さまざまな音声対話システムの試作

音声認識カーナビ
が普及

ARISE
(EU 列車時刻案内)

観光案内
(豊橋技科大 中川研)

Communicator
(フライト予約 DARPA)

Jijo-2 (産総研)

DUG-1 (NTT)

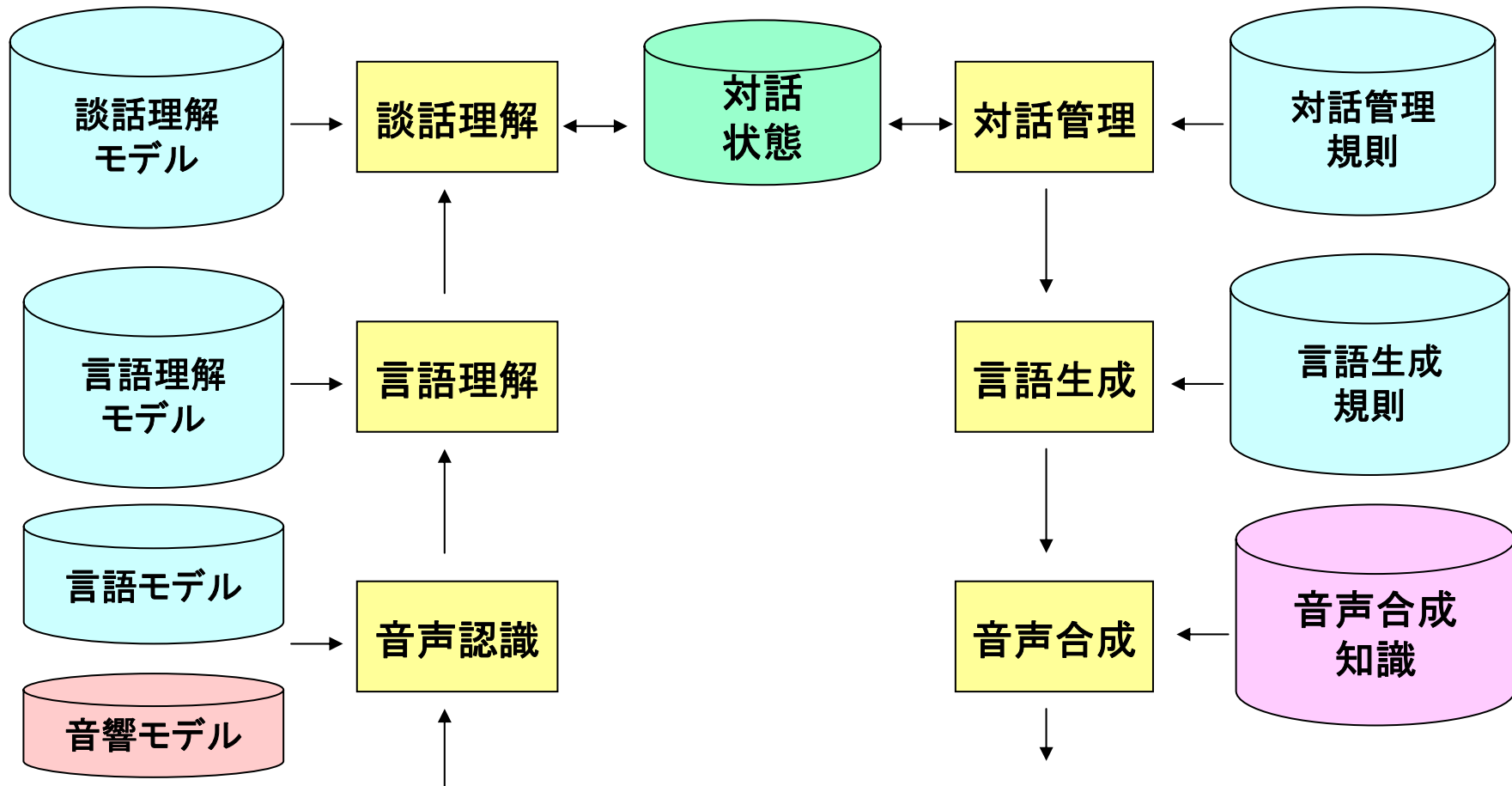
ROBITA
(早大 小林研)

文献検索
(東大 広瀬研)

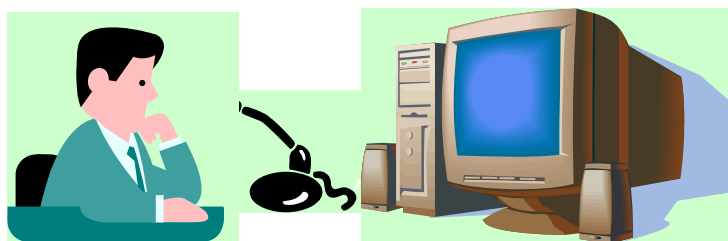
技術課題のリストアップ

音声対話システムの基本構成

(中野・堂坂, 人工知能学会誌2002; cf. 河原・荒木, 音声対話システム 2006)



1999年当時、何が課題だったのか



1999年当時、何が課題だったのか

タスク達成率・
効率

- 理解率向上
- エラー検出 & リカバリ
(平沢他, SLP99)
- 確認戦略
(新美他, SLP99)

1999年当時、何が課題だったのか

タスク達成率・
効率

自然性

- 話し言葉理解
- 柔軟な発話交代
(Nakano et al. Eurospeech 99)
- 韻律の利用
- 応答生成
(桐山&広瀬, SLP99)

1999年当時、何が課題だったのか

タスク達成率・
効率

自然性

- 複数ドメインでの対話
(Lin et al. Eurospeech 99)
- 複数人との対話
(Matusaka et al. Eurospeech 99)

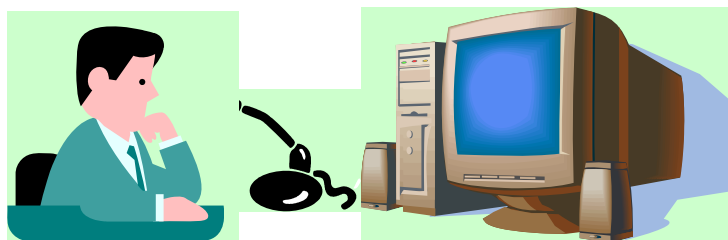
対話の種類

1999年当時、何が課題だったのか

タスク達成率・
効率

自然性

音声認識イン
タフェースの利
用者増加



対話の種類

システムに対
する人間の行
動分析
(小林 SLP99)

システム開発
効率
(小暮他 SLP99)

10年間の進展

タスク達成率・
効率

自然性

音声認識イン
タフェースの
利用者増加

研究が進展

•認識理解信頼度

(e.g. Hazen他, Comp. Speech & Lang.'02, 駒谷他 情
処論02)

•文脈を利用した音声理解

(e.g. Higashinaka他 ACL03, 藤原他 信学論06, Bohus
PhD Thesis 07)

•動的, 確率的な対話制御

(Singh AAAI00, Dohsaka et al Eurospeech 03,
Williams他 Comp. Speech & Lang'07)

対話の種類

システムに対
する人間の行
動分析

システム開発
効率

10年間の進展

タスク達成率・
効率

自然性

音声認識イン
タフェースの
利用者増加

研究が進展

- 種々の話し言葉音声理解法

(サーベイ: De Mori, ASRU07)

- 話者交代

(e.g. Sato et al. ICSLP02, Kitaoka et al. J.JSAI05, Raux, SIGDial08, Fujie et al. Interspeech05)

- 韻律を利用した態度認識

(e.g. Fujie et al. Interspeech04)

対話の種類

システムに対
する人間の行
動分析

システム開発
効率

10年間の進展

タスク達成率・
効率

自然性

様々なタイプの対話システムの 研究

- **ドキュメント検索** (e.g. Misu et al. ICASSP07)
- **非タスク指向** (e.g. Bernsen et al. ADS04)
- **クイズ** (Higashinaka et al. Interspeech07)
- **マルチドメイン対話** (e.g. O' Neill et al. ICSLP04, Komatani et al. Sigdial06)
- **対話的プレゼンテーション** (Nishimura et al. IUI07)

音声認識イン
タフェースの
利用者増加

対話の種類

システムに対
する人間の行
動分析

システム開発
効率

10年間の進展

タスク達成率・
効率

自然性

音声認識イン
タフェースの
利用者増加

ツール構築 & 研究進展

- 音声認識, 合成, 対話ツール充実
- 対話知識・モデルのUnsupervised学習, 能動学習
(Tur et al. Speech Communication '05, Sudoh et al. Speech Communication '05)
- WEBからの知識獲得
(e.g. Misu et al. Interspeech06)
- 未登録語などのオンライン知識獲得
(e.g. Choueiter et al. Interspeech07)

対話の種類

システムに対
する人間の行
動分析

システム開発
効率

10年間の進展

タスク達成率・
効率

自然性

音声認識イン
タフェースの
利用者増加

データ収集用システム

- Communicator (DARPA)
- 京都バス案内(京大)
- CIAIR(名大)
- たけまるくん(奈良先端大)

対話の種類

対話システムを用いたuser study
(e.g. Hu et al. CHI 2007)

システムに対
する人間の行
動分析

システム開発
効率

10年間の進展

タスク達成率・
効率

自然性

音声認識イン
タフェースの
利用者増加

商用システム

- 音声カーナビ
- ボイスポータル
- 電話応答
(チケット予約, フライト情報etc.)
- ハンドヘルドデバイス

対話の種類

システムに対
する人間の行
動分析

システム開発
効率

ブレイクスルーはあったか？

■ 技術課題

- 各技術課題の研究は着実に進んだが、ブレイクスルーといえるほどのインパクトのある新技術はなかったのではないか

■ 研究パラダイムのシフトがあった

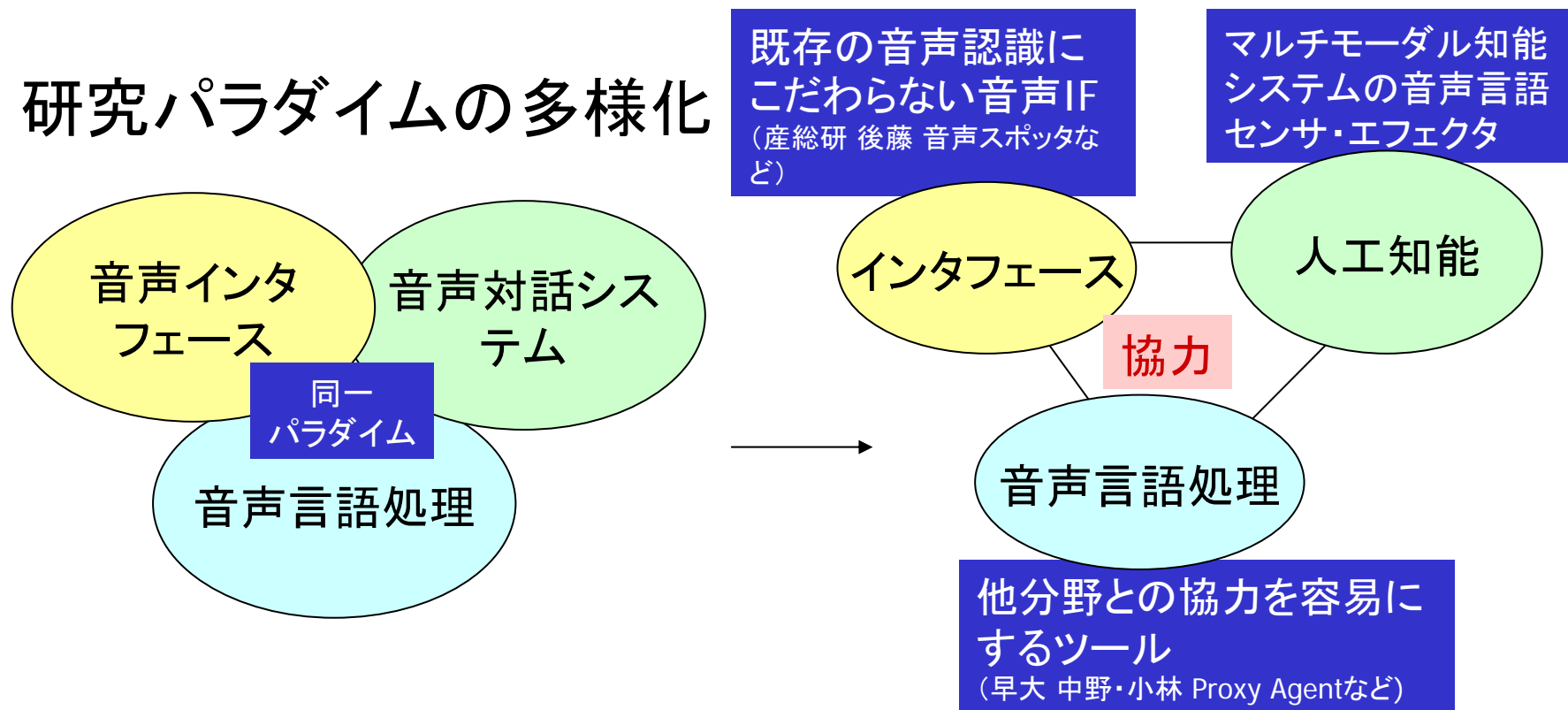
- 音声対話ツール
 - 例: Galatea(含Julius), Galaxy, VoiceXML, etc.
 - 音声インタフェース利用の裾野を広げた
- 人-システム対話の収集・分析事例の増加
- 対話の種類が増加

できなかったこと

- 各技術の研究段階から実用段階への移行
 - 特に統計的手法
 - データの収集 & アノテーションのコストが膨大
(cf. Knight et al. Eurospeech 2001)
- 音声対話・音声認識インタフェースを日常的に使う人を増やすこと(特に日本で)
 - 誰とでも対話できる人工知能対話システムはまだ難しい
 - 限定されたインタフェースとして使う場合でもインタフェースとしての課題が解決されていない (cf. Shneiderman&Plaisant 2005)
 - 誤りの取り消しなど

今後の展開予想

研究パラダイムの多様化



対話知識の自動構築・オンライン学習

語彙, 言語モデル, 音声理解, 対話コンテンツ, etc.