



感情情報を含む 音声行動コーパスの構想

桐山 伸也, 大谷尚史, 石川翔吾,
仲川淳, 竹林洋一, 北澤茂良

静岡大学



コモンセンス研究のアプローチ

- ◆ 言語コーパス (岡田先生)
- ◆ 仮想世界モデル (Minsky)
- ◆ 実世界データに根ざした現場主義研究



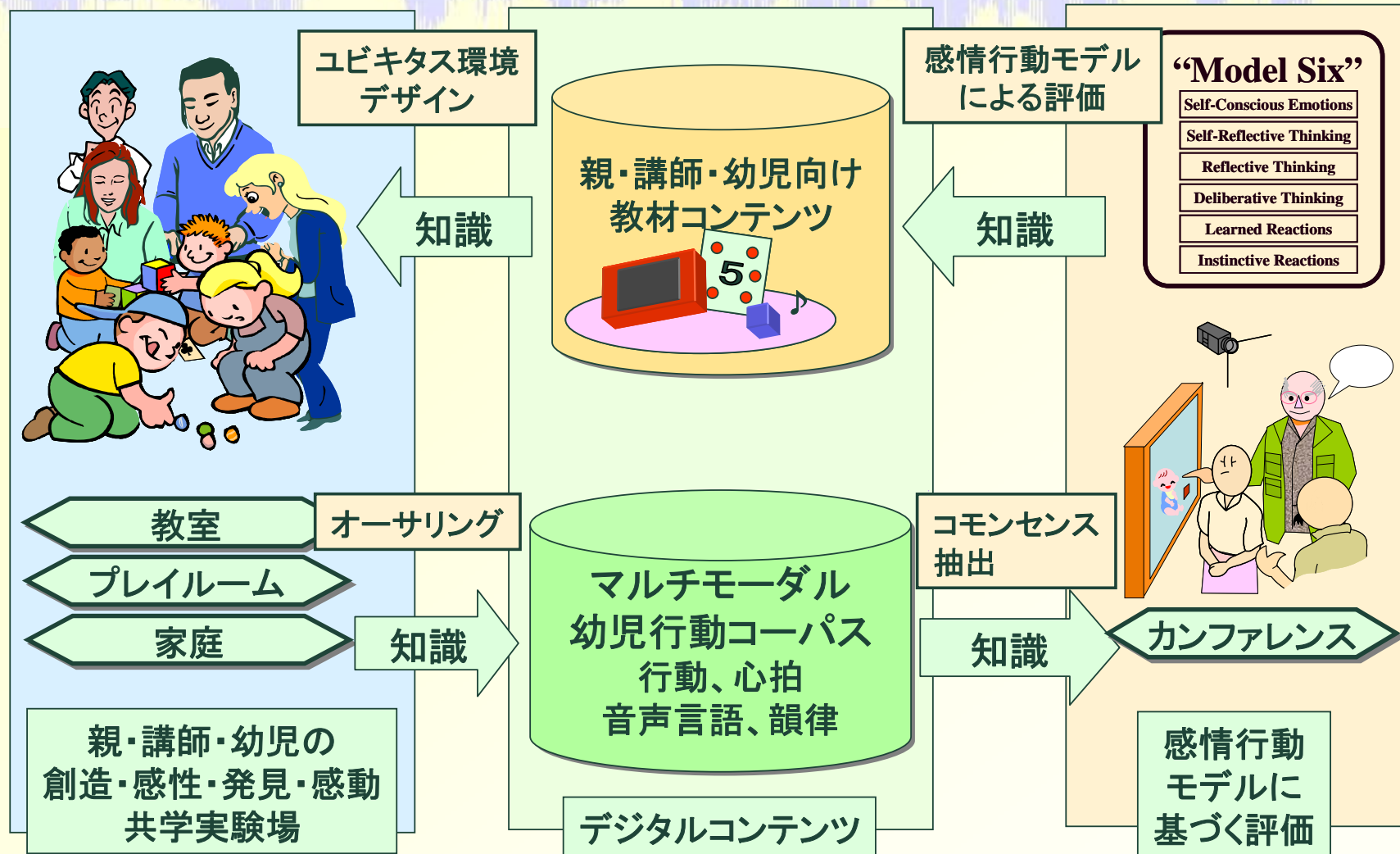
マルチモーダル幼児行動コーパス

- ◆ ナイーブな幼児の行動に着目
- ◆ 感情・思考・意図の内面が観測しやすい
- ◆ 幼児教室を定期開催し映像・音声収録
- ◆ 行動に注釈付けをしてコーパス化

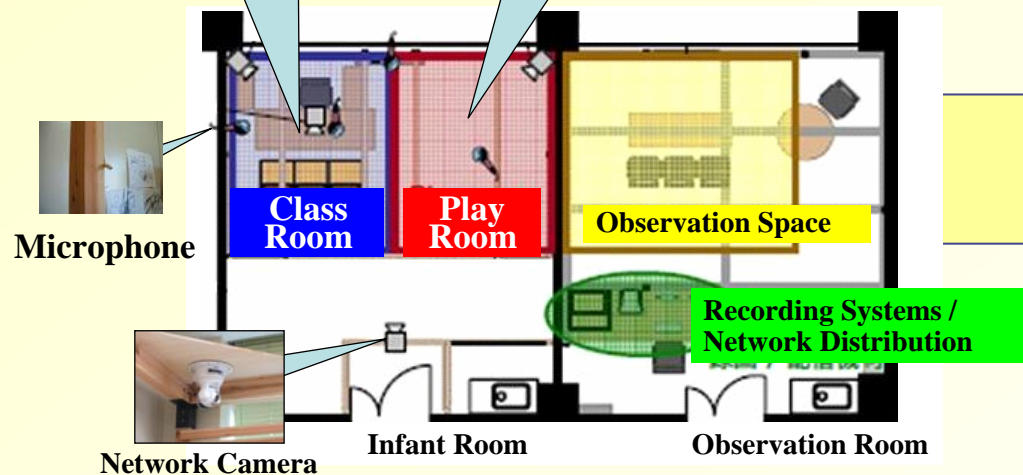
学内に幼児教室を開設



幼児学習プロジェクト



マルチモーダル幼児行動コーパス



マルチモーダル
幼児行動コーパス



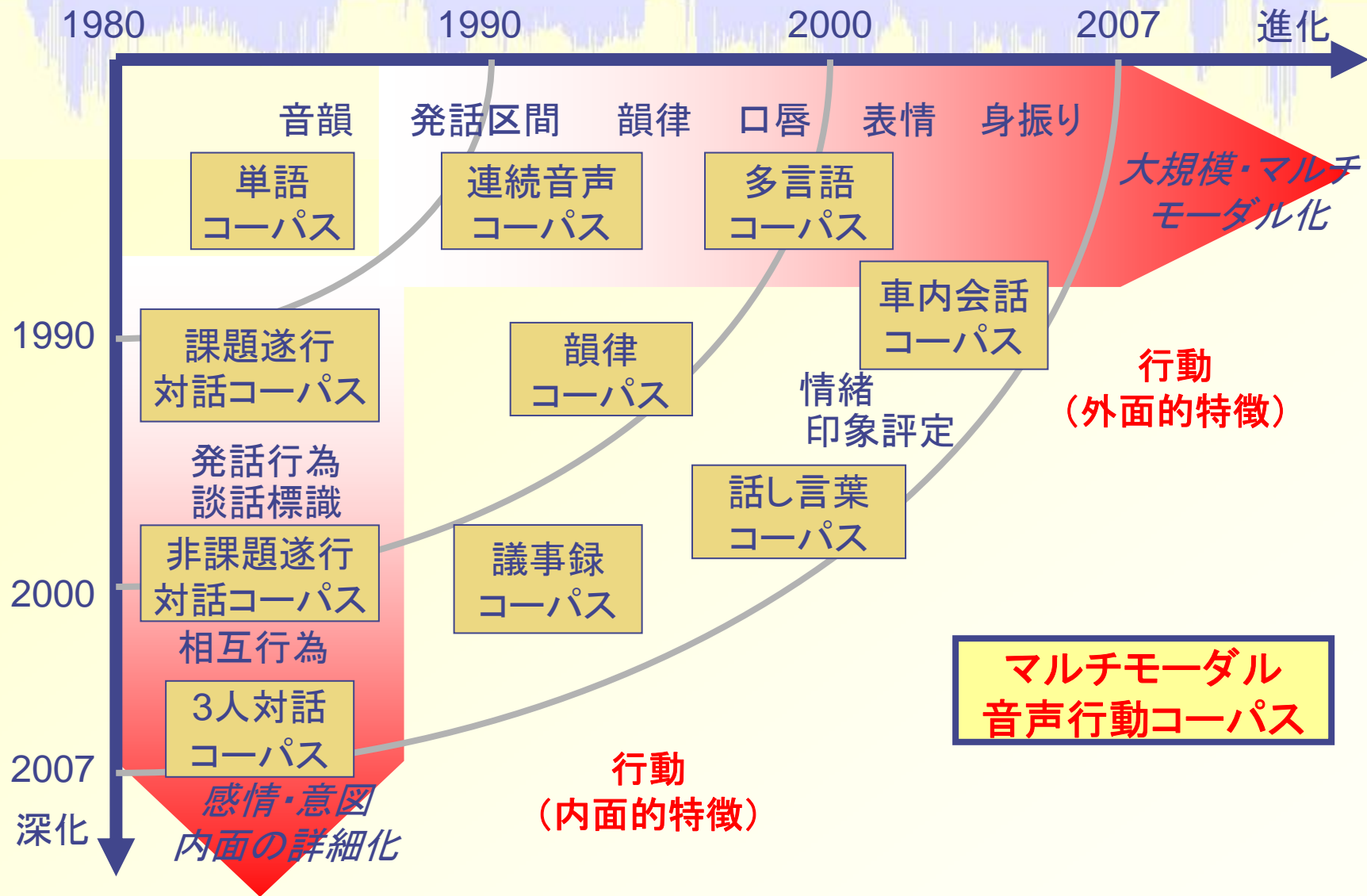
発話、しぐさ
目線、表情などの
マルチモーダル注釈

- ◆ マルチアングルカメラとウェアラブルマイクで高品位収録！
- ◆ 同一幼児の経年データをこれからも継続して蓄積可能！

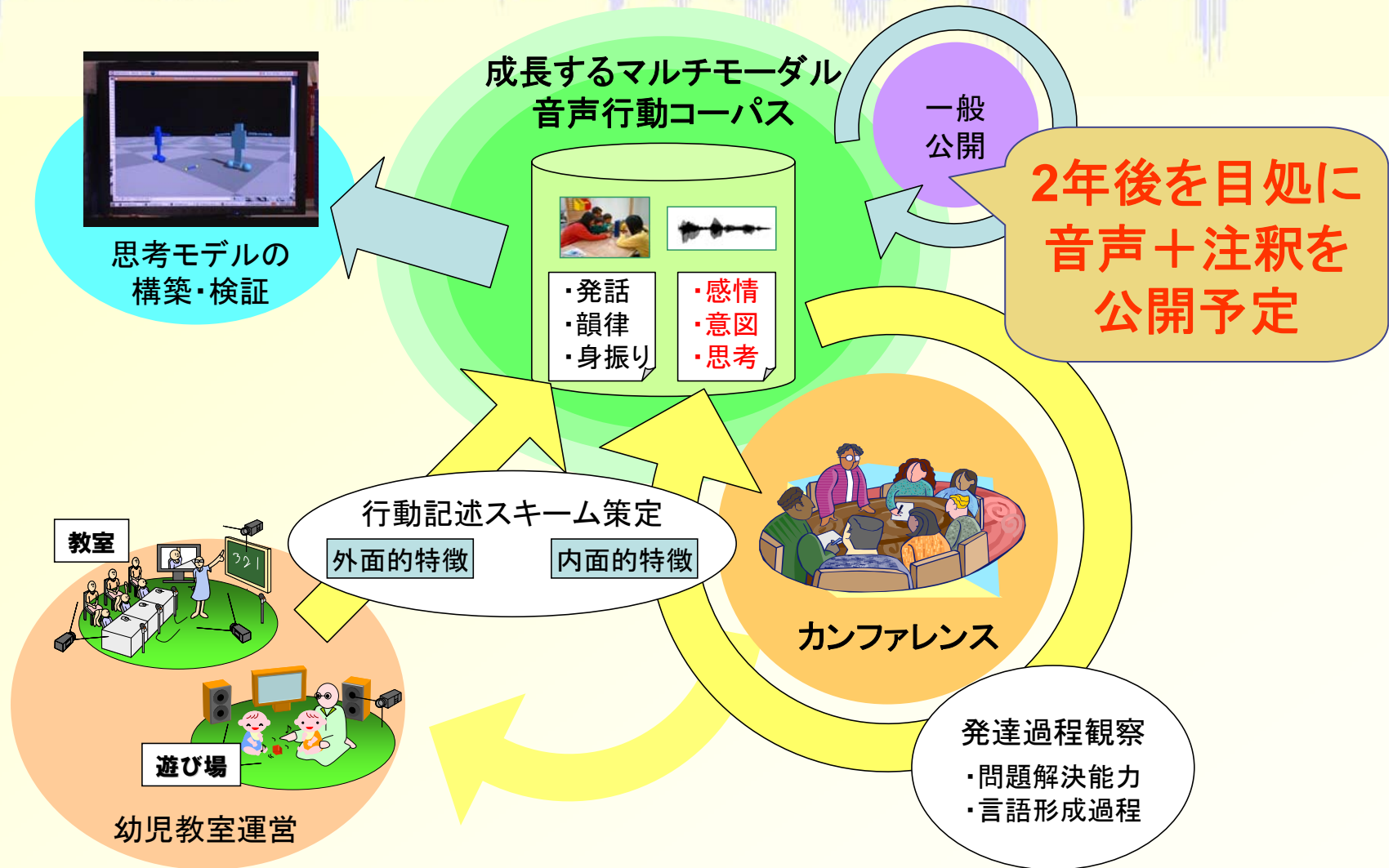
行動コーパス構築で見えたこと

- ◆ 音声が生の内面を捉える重要な手掛かり
 - 身振り、視線、動作は解釈があいまい
 - 音声は一次元で文字シンボルに転記可＋韻律
- ◆ 主観の客観化プロセスでコーパスの深化が可能
 - 生の内面の解釈に正解はない
 - 研究者・保護者・教室講師など異なる視点の主観を持ち寄って議論すると最善の解釈が可能
 - 患者の治療方針を決める病院のカンファレンスと同じ

音声コーパスの発展



カンファレンス主体のコーパス構築



現状コーパスの概要

- ◆ 3年半に亘り300時間の映像・音声収録
- ◆ 「おいしい」場面を**ゴール(目標)**主導で選択
 - 心の状態が変化する「アクシデント」の場面
- ◆ 客観化が容易な外面的特徴を記述
 - 授業内容・場面説明・発話区間・発話内容・話し手・聞き手・自然言語のコメント
 - 約1000場面・3000発話を書き起こし
- ◆ **上記を手掛かりに「感情」の記述に着手**



感情に着目するメリットは？

◆感情は問題解決のスイッチ (Minsky)

- 怒りの状態： 反射的に高速な解を優先
- 恐れの状態： 熟考し慎重に解を選択

⇒ 意図・思考を推察する重要な手掛かり

◆感情音声の分析・認識・合成研究が盛ん

- 感情は、音声研究コミュニティとの共通言語



感情ラベリングの目的

- ◆ アクシデントに直面し、問題解決する場面における思考過程の解析
 - 感情ラベルをトリガに類似場面を検索し、
 - 感情状態と思考過程の関係を分析したい！
- ◆ 成長に伴う感情表現の発達変化の解析

感情ラベリングの実践

- ◆ アクシデントが多いプレイルームに着目
 - 1歳児・3歳児クラスの各300場面・1000発話
 - トータルで2000発話にラベリング
- ◆ 「基本6感情」に「中立」を加えた7種を付与
 - 一人のラベラが全ての発話に付与
 - 音声に加え映像で場面を把握した上で判断
 - 可能性のある感情カテゴリを全て選択
 - 「中立」以外には、弱・中・強の強度も付与

幼児行動分析支援ツール

ビデオ再生

The screenshot displays the InfantViewer application window. The main area is a video player showing a classroom scene. A red box highlights the video player and a table of time-stamped comments. Below the video player is a control panel with play, stop, and seek buttons. At the bottom, there is a comment analysis panel with tabs for 'social', 'event', and 'behaviors'. A red box highlights this panel, which shows a detailed comment in English and its corresponding behavioral analysis.

2007/01/22 (MON) 14:55:22

Target: 053_070122_age3
Title: 053_070122_age3

開始時間	終了...		
<input checked="" type="checkbox"/>	00:00:00	00:07:42	
<input checked="" type="checkbox"/>	00:07:42	00:47:58	
<input checked="" type="checkbox"/>	00:12:45	00:13:01	
<input checked="" type="checkbox"/>	00:28:58	00:29:13	0. ○ふくと:表現おも...
<input checked="" type="checkbox"/>	00:29:13	00:29:24	0. ふくと母「何色って...
<input checked="" type="checkbox"/>	00:29:27	00:29:44	0. ふくと「二個に、こ...
<input checked="" type="checkbox"/>	00:32:18	00:32:35	0. ○ふくと:説明ふく...
<input type="checkbox"/>	00:47:58	00:47:58	0. 授業後
<input checked="" type="checkbox"/>	00:48:18	00:48:56	0. あやね観察あまり...
<input checked="" type="checkbox"/>	00:48:56	00:49:11	0. あやね拳動不審...
<input checked="" type="checkbox"/>	00:49:11	00:49:26	0. あやね許す?ふく...
<input checked="" type="checkbox"/>	00:49:51	00:50:02	0. あやねチョコボール...
<input checked="" type="checkbox"/>	00:50:12	00:50:30	0. あやね先生に反...
<input checked="" type="checkbox"/>	00:57:08	00:57:38	0. あやね回転だんだ...

一時停止 50:25

映像ソース 053_070122_3.wmv 50 開始位置 00:50:12

音声ソース 50 終了位置 00:50:30

コメント 書き起こし 教材評価 詳細記述 指差し

social

event

<story>
Girl reacted to teacher. And she said "I'll come!". Then she moved close to chair.
After that, she told teacher "Hi!"

<behaviors>
(sequential
(says girl teacher "I'll come!")
(moves-to Girl Teacher)
(is-visible-to Girl Teacher)
(observes girl (

tells
asks

the ACTOR tells something
how to use: (tells told-by told-to thing-told)

行動分析パネル

残り
続

write

映像の時間情報に
コメントを関連付け

自然言語を用いたコメントの記述

←コメント記述時の指針やヒント、過去の記述事例を表示

感情ラベルによるコーパス分析

発話ID	話し手	聞き手	発話内容	喜び	怒り	悲しみ	驚き	恐れ	嫌悪	中立
034_20060	K	K母親+先生	とられちゃったからんだ							○
034_20060	F	世界	とられちゃちゃったーい	△						
034_20060	K	世界	オッカッタツツッケー	△						○
034_20060	K	F	こないで		△					
034_20060	F	K	ないよーだ							○
034_20060	F	F	こないで		△					
034_20060	F	世界	おかたづけーじゃないよ							○
034_20060	K	F	こないでこないで、こないで		○			○		
034_20060	K	F	こないでこないで		○			○		
034_20060	K	F	ぎゃー、はー(以後泣き)		◎			◎		
034_20060	F	世界+先生	がーん							○
034_20060	K	F母親	うーあーん		△	△				
034_20060	K	K母親	うわーん		△			△		
034_20060	K		いたーい		△			△		

◆感情ラベルの時間変化に着目することで、
問題解決の種類を整理できる見通し



感情を主観でつけてよいのか？

- ◆ 音声パワー、基本周波数などの物理量が客観データになるのでは？
- ◆ 定型的なパターンの抽出には効果あり
 - ⇒ 従来の感情音声認識・合成研究のアプローチ
- ◆ 激しい怒りを押し殺す、悲しみをこらえて空元気で振舞うなど、複雑な心的状態が格好の分析場面
- ◆ 両者を併用すると、上記が抽出可能に！

◆心の状態の計算機表現

- 所有に関する心的状態表現モデル
- 問題解決に関する階層的思考モデル
- CGによる思考モデルのシミュレーション



思考モデルの構築・検証

成長するマルチモーダル音声行動コーパス

一般公開

◆主観の客観化プロセスの具現化

- カンファレンスの実践
- カンファレンス支援ツールの開発

◆音声言語獲得過程分析

行動記述スキーム策定

外面的特徴

内面的特徴



カンファレンス

◆新たな教材取り組みの提案



教室



遊び場

幼児教室運営

発達過程観察

- ・問題解決能力
- ・言語形成過程



学術研究への貢献

◆言語発達研究

- 韻律の発達変化
- モダリティの獲得過程

◆インタラクション研究

- 個人の認知・思考・行動からインタラクションを捉える新パラダイム提供



音声行動コーパスの実世界応用

- ◆ 医療現場のカンファレンス・インフォームドコンセント
 - ◆ 放送局・工場・発電所等でのチーム連携業務
 - ◆ 学校・オフィス等での会議・プレゼンテーション
 - ◆ ...etc
-
- ◆ コミュニケーションがある現場全てが対象
 - ◆ 役割分担・ワークフローにより行動のゴールが設定可
 - ◆ ウェアラブル機器の普及により応用場面が拡大



操船現場の音声行動コーパス

- ◆ イージス艦を筆頭に船舶事故多発
- ◆ 原因はヒューマンコミュニケーションエラー
 - 他船通信中に目視報告を遠慮
 - 指示に対する確認応答の欠落
- ◆ ブリッジ業務音声行動コーパス構築
 - ヒューマンエラー候補の自動検出教示
 - シミュレータによるトレーニング支援

船舶ブリッジの音声行動分析

The screenshot displays an audio analysis software interface with the following components:

- Timeline:** A horizontal axis at the top showing time from 1:48.0 to 2:08.0 in 0.1-second increments.
- Label Track:** A track containing labels for communication events: "B-1: 操船の指示", "C-1: 他船への注意喚起", "C-2: 次のコースの報告", and "B-2: 操船の完了報告".
- Audio Tracks:** Multiple tracks labeled "1_r", "2_r", "3_r", and "4_r" showing waveform data. Each track has a corresponding "ラベルトラック" (Label Track) with text labels such as "20. Starboard_20.", "O.K. Thank you.", "(F a) * Starboard ... 20?", "Captain! confirm are small vessel starboard <?>.", "two three.", "small vessel.", "O.K. Now next to * course (F wa-) <?>", "Next course around 252, s", "board", "Starboard 20.", "Starboard 20, sir.", "C-1 => ok (※)", "C-2 => ok", and "ok (※ 2/O、3/Oの二人から同時に報告が入ったが、Captainは一度しか返事をしていない。二人に対しての返事なのかどうか?)".
- Bottom Panel:** Contains project settings including "プロジェクトのサンプリング周波数 (Hz): 44100", "選択開始:" (00 h 01 m 52 s), "終了" (00 h 01 m 52 s), "長さ" (00 h 00 m 00 s), and "音の位置". A checkbox for "スナップモードを有効" is present.
- Footer:** A note in Japanese: "オーディオを選択するためにはクリック後にドラッグして下さい。"



まとめ

- ◆ **現場主義**のコモンセンス研究
- ◆ **音声行動コーパス**という新概念提唱
- ◆ 幼児教室ドメインでコーパス構築
 - 感情・意図・思考のラベリング
 - 主観の客観化でコーパスを進化・深化
- ◆ 学術研究 * 実世界応用に貢献