

# 知識科学研究分野 駒谷研

教授 駒谷和範 准教授 武田龍

<http://www.ei.sanken.osaka-u.ac.jp/>



## 音声対話システム: 人の音声を聞いて話すコンピュータ

### 研究目的

機械が人間にとって身近で使いやすい存在となるには、人間が本来備えている音声対話機能が必要です。本研究分野では音声や言語、マルチモーダル情報を用いて**人間と対話するシステム**を研究します。機械学習技術に基づき、音響信号処理や音声認識、自然言語処理、知識グラフ、マルチモーダルインタラクションやユーザモデリングなど、幅広い観点から**賢く話せるシステムの実現**を目指します。

### 対象とするシステムと技術

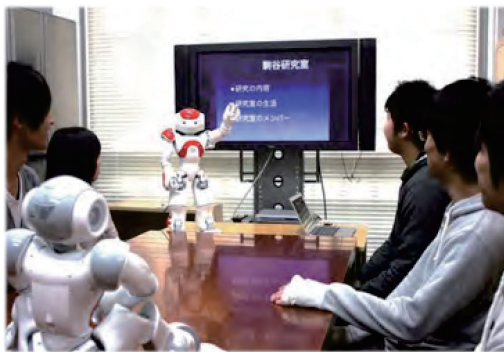
#### テキストチャット

2021年に東京でオリンピックが開催されます

何の競技が人気なの？

最近は色々人気ですよ

確かにメディアも色々取り上げるようになったね

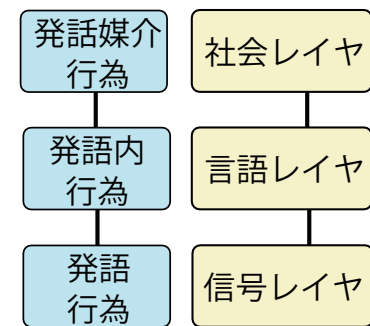


複数の人間と話す対話ロボット

#### マルチモーダル対話



### 対話理論

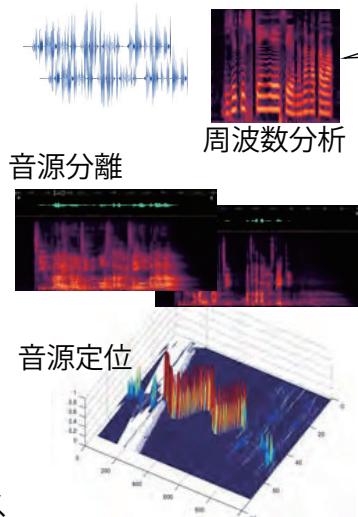


発話行為理論 [Austin, 1962]



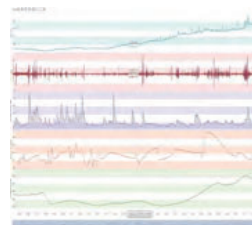
### モデル・知識 対話戦略・実装

#### 信号処理

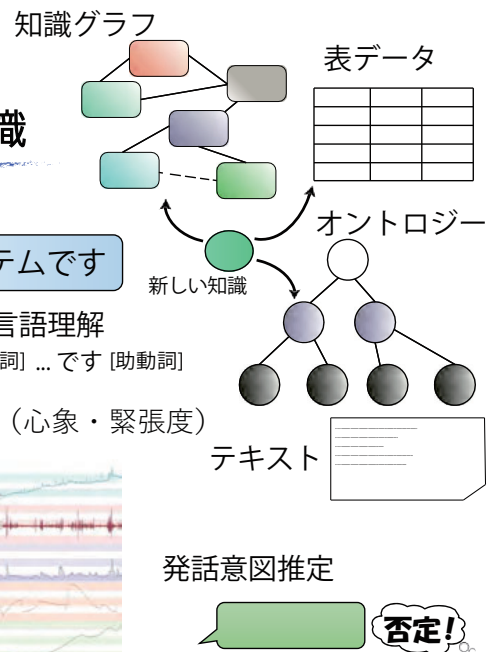


#### パターン認識

音声認識  
音声対話システムです  
テキスト分析・言語理解  
音声[名詞] 対話[名詞] ... です [助動詞]  
ユーザ状態推定 (心象・緊張度)



#### 知識表現・拡充



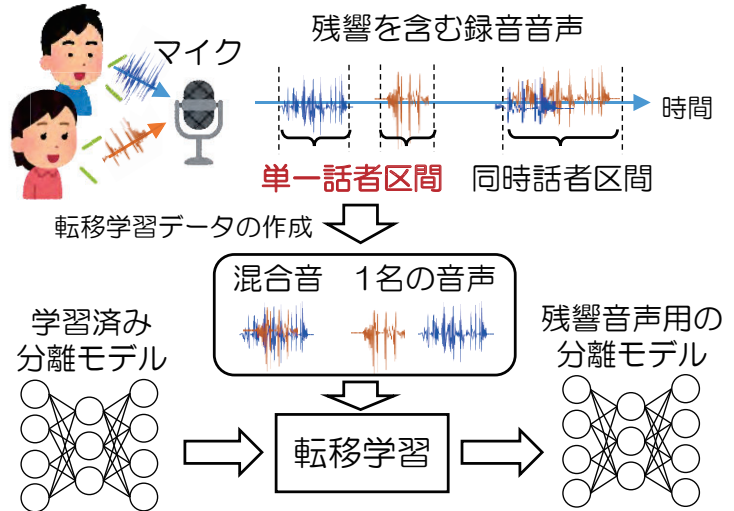
# 研究内容の例

## 教師ありモノラル音源分離のための残響音声データ内の単一発話区間を活用した転移学習 (R2年度学部生卒論)

**背景** モノラル音声分離は、複数話者の声が混ざったモノラルの音響信号を各話者の音声に分ける技術です。音声認識の前処理や音声の明瞭度向上に使われます。**深層学習に基づく音声分離**は高性能ですが、録音環境の残響(反射音)を含むデータを用いた教師あり学習が必要です。

**本研究** 学習済みの音声分離モデルを残響音声に適應する**転移学習**に着目しました。分離したい音声データに含まれる**単一話者区間**を用いて、**学習に必要なデータのペア**を作成します。その有効性と転移学習に必要なデータ量を検証しました。

### IPSJ 全国大会 学生奨励賞受賞



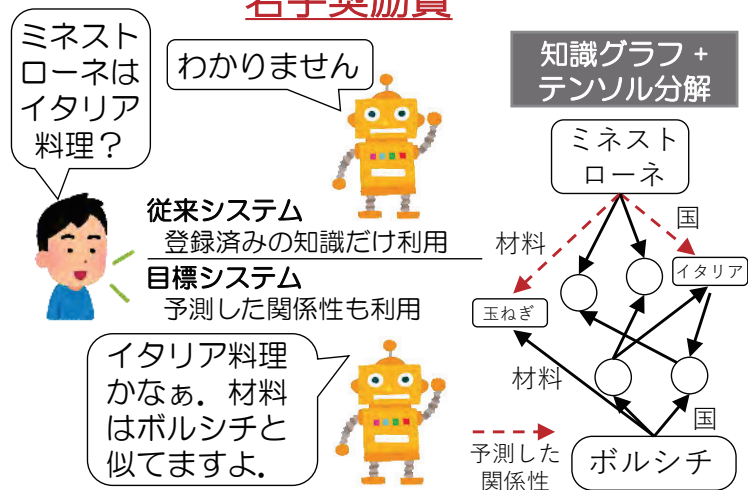
## 未知の属性の問い合わせに応答可能な対話システム (R元年度修士学生修論)

**背景・課題** 現状の対話システムは、**システムにとって未知の項目や関係性(属性)**に関する質問に上手く応答できません。ユーザからのこのような質問への対応は不可欠ですが、事前にシステムの持つ知識を完全にすることは困難です。

**本研究** システムの持つ既知の**知識をグラフで表現**し、項目間の**未知となっている関係性(属性)**を**モデルに基づき予測**します。その結果に基づき、合理的な応答生成を目指します。

例: 断定はしないが可能性を仄めかす応答  
類似した語を推測し応答に活用

### JSIAI 対話システムシンポジウム 若手奨励賞



## 学習時の計算量を考慮した対話行為設計に基づく強化学習による発話選択 (R2年度学部生卒論)

**背景** ユーザの話の聞き役となる傾聴対話システムでは、**話の流れに合った質問や反応**を返す必要があります。ある話の流れのもとで、どのシステム発話が適切かを対話を繰り返して学習します(**強化学習**)。システム発話と話の流れの定義によっては、対話の破綻や学習時の計算量の増加を招いてしまいます。

**本研究** システムの発話を対話行為に基づいて適切な粒度で分類し、計算量の増加を抑えました。また、過去のシステム発話との整合性を考慮することで、対話破綻を抑えました。

### IPSJ 全国大会 学生奨励賞受賞

