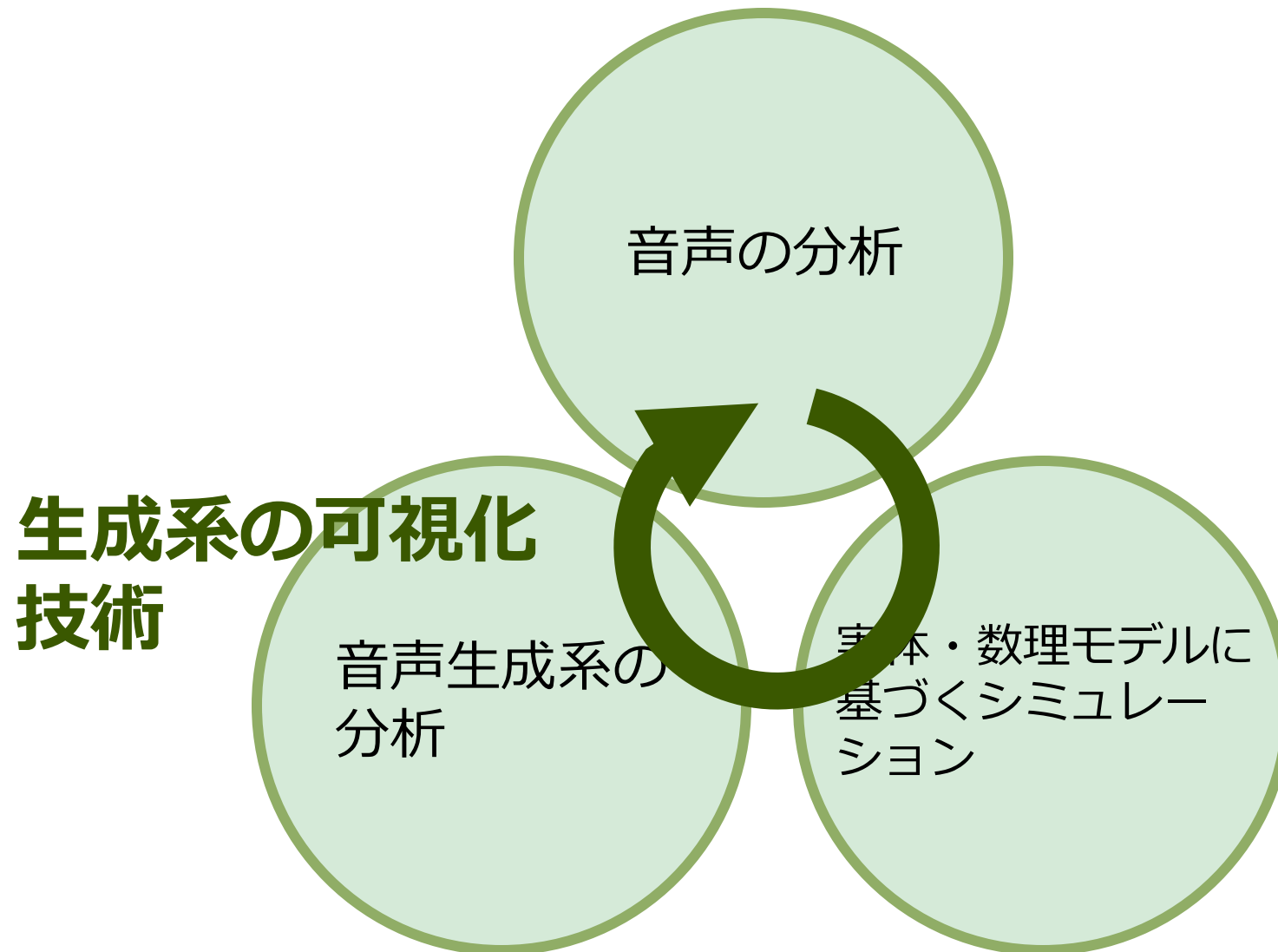


音声コミュニケーション
スペシャルセッション「調音の可視化とその応用」

音声生成系の様々な可視化手法

北村達也 (甲南大)

調音音声学の研究手法



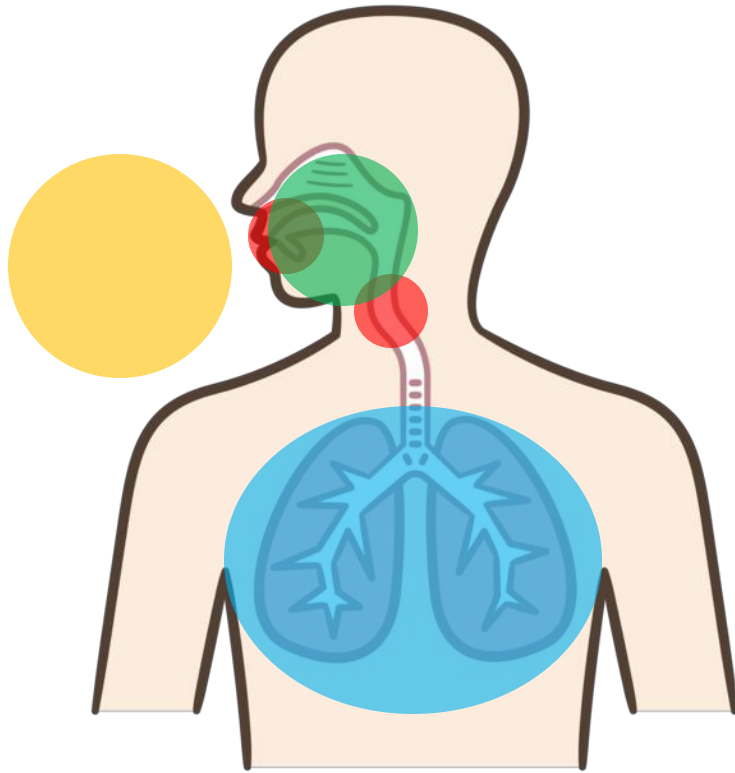
生成系の可視化の意義（魅力）

研究対象の音声が生じる原因や原理に直接迫ることができる

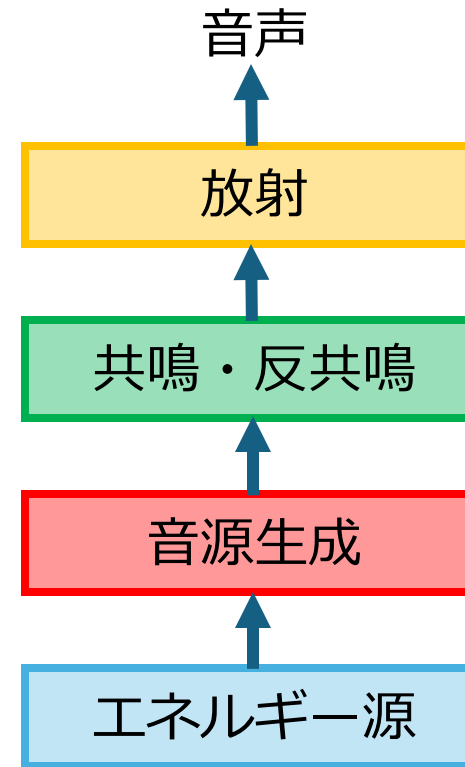
- 従来見えなかったもの，現象を
- 従来を超える精度，規模，速さで

可視化

音声生成と可視化

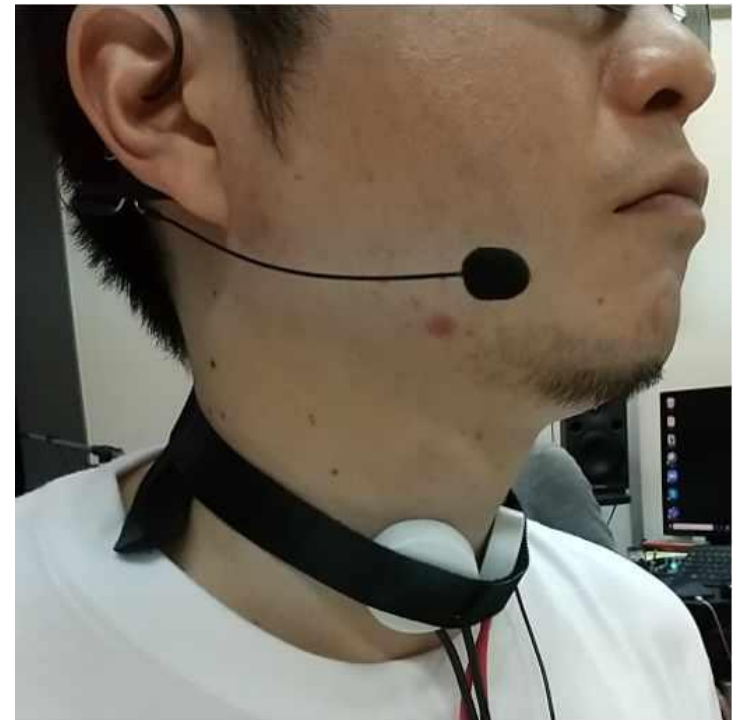


© PIXTA



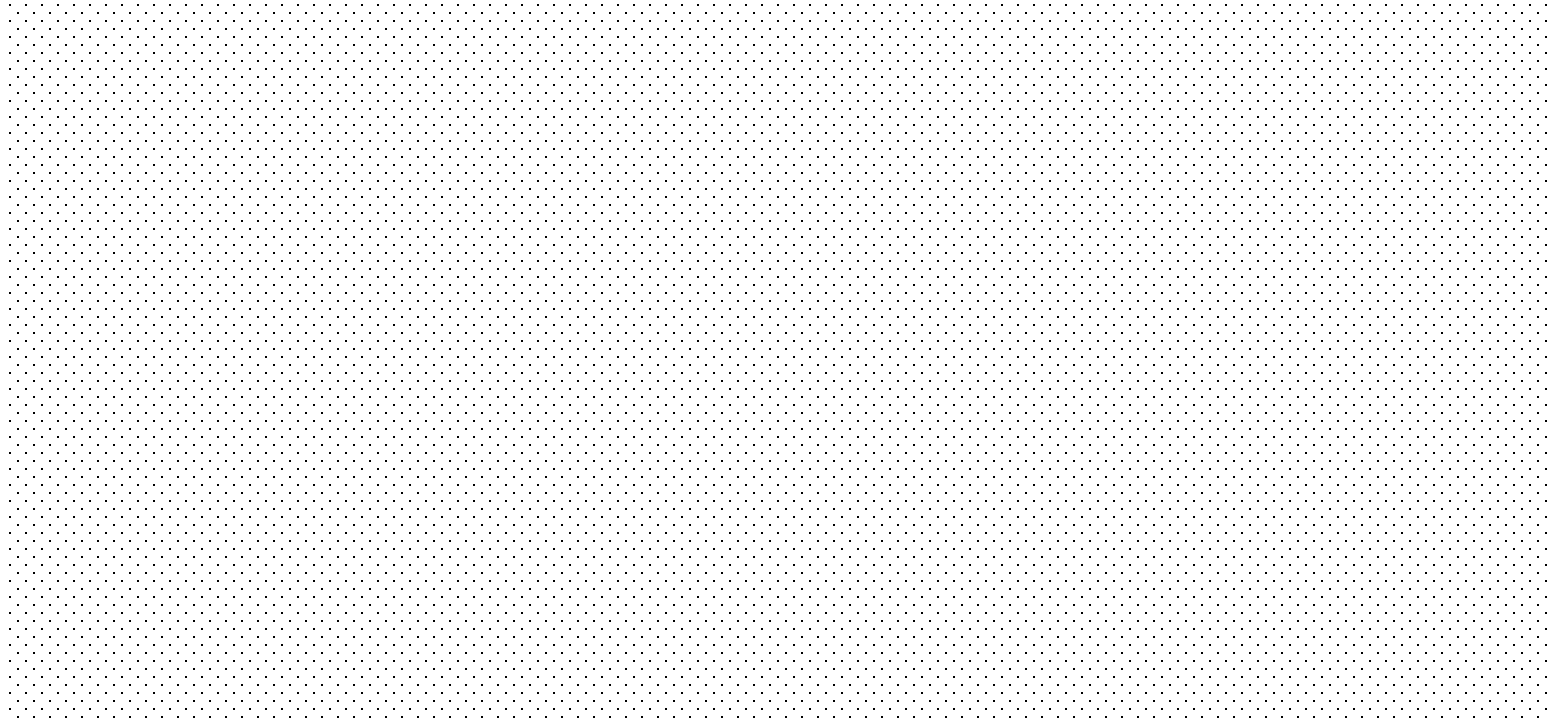
発声器官の可視化

- 間接的観測法（非侵襲的）
 - Electroglottography (EGG)
 - Photoglottography (PEG)
- 直接的観測法
 - ファイバ스코ープ
 - 硬性側視鏡



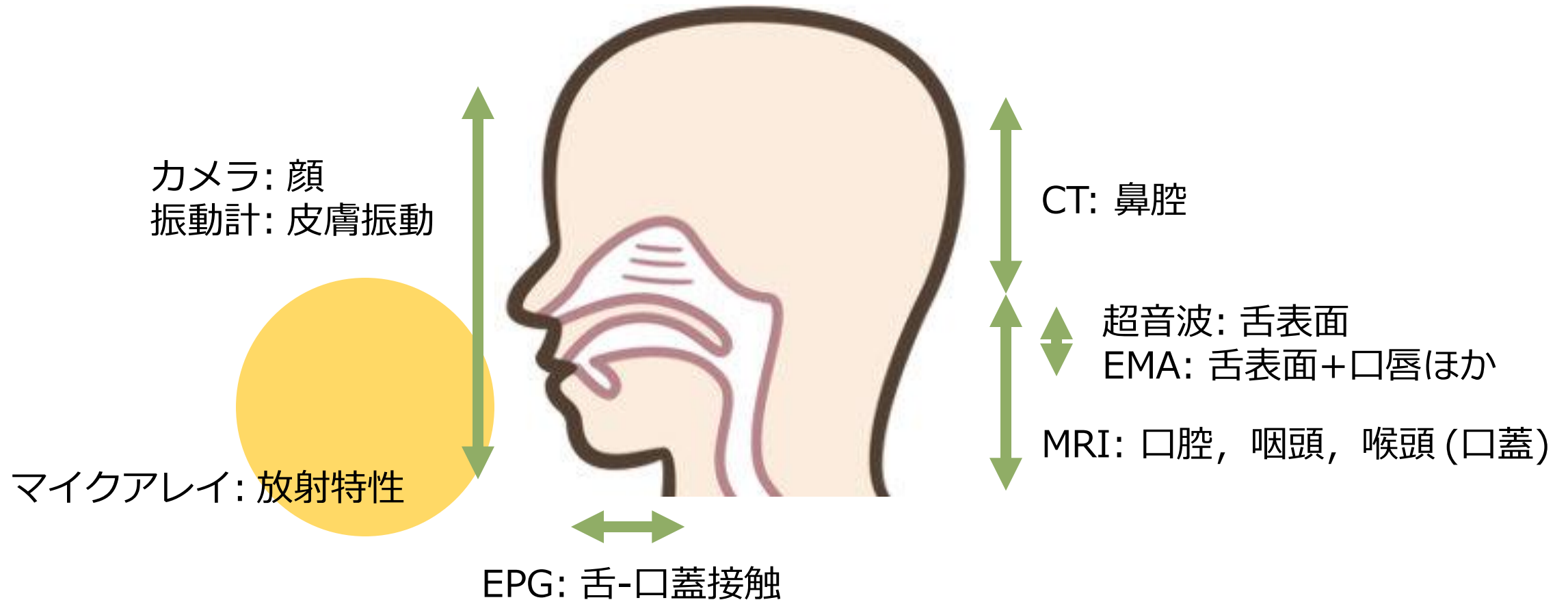
EGGの使用例

参考：超音波画像診断法による可視化例



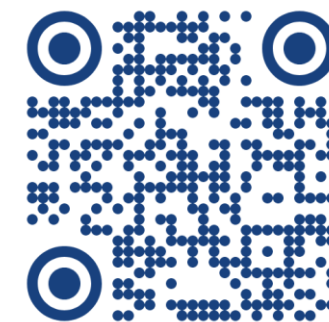
Cho *et al.* (2013)

調音の観測法



リアルタイムMRI調音運動データベース (rtMRIDB)

- 国立国語研究所が公開
- 14/27 fpsにて撮像した日本語分節音生成時のrtMRI
- 話者22名
- 26000発話以上
- 種々の条件により検索可能
- 動画のダウンロードも可能

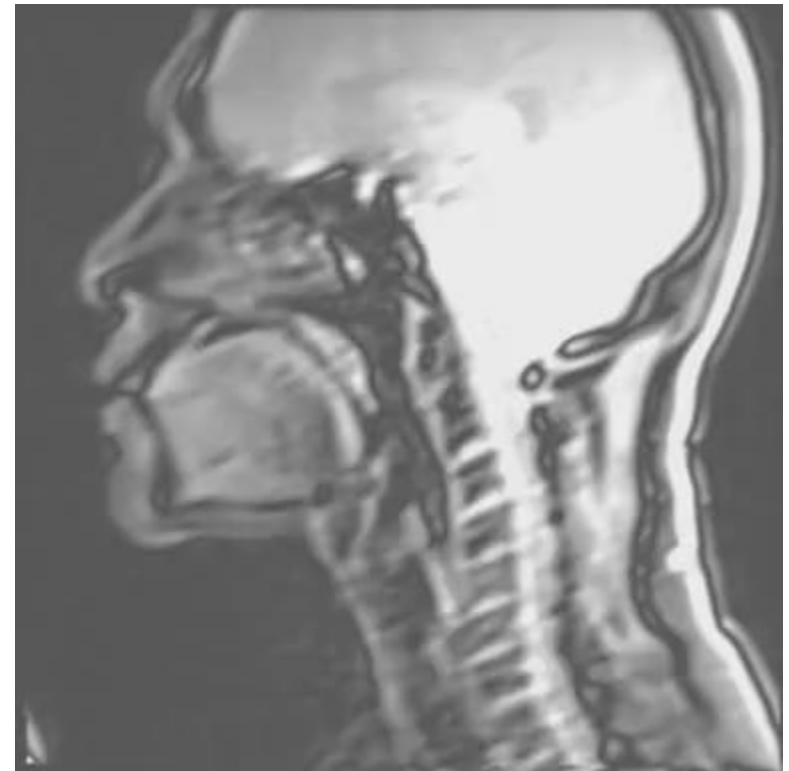


rtMRIDB

愛 /ai/

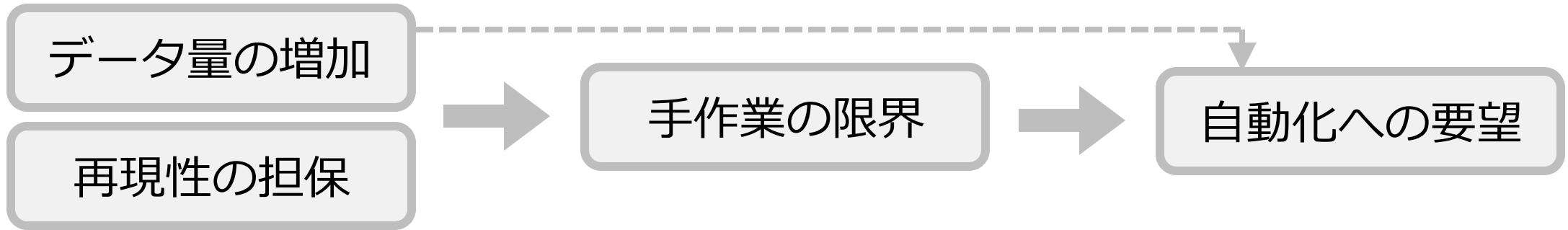
rtMRIのフレームレート

14/27 fps → 54 fps

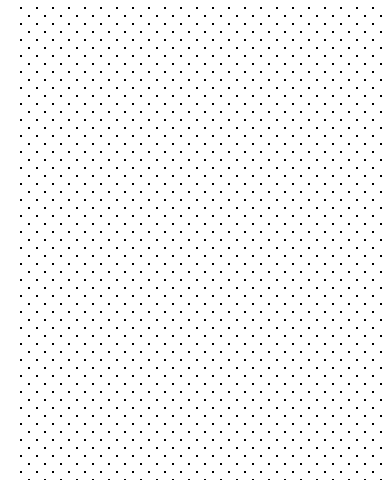


グイ語の調音運動 (音なし)

輪郭線の自動抽出



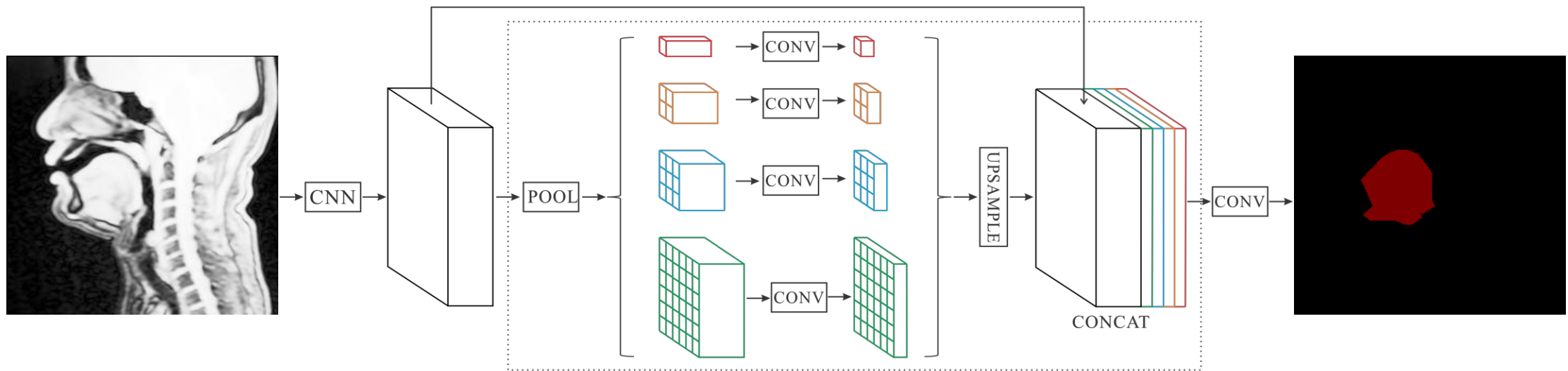
- 調音器官の輪郭の自動抽出
- 後藤らの手法
 - 機械学習ライブラリDlibを使用
 - 少ない学習データ, 少ない演算量
 - 高精度



後藤ら (2019)

セマンティックセグメンテーション (1/2)

- PSPnet (Pyramid Scene Parsing Network) (Zhao et al., 2017) による舌領域抽出

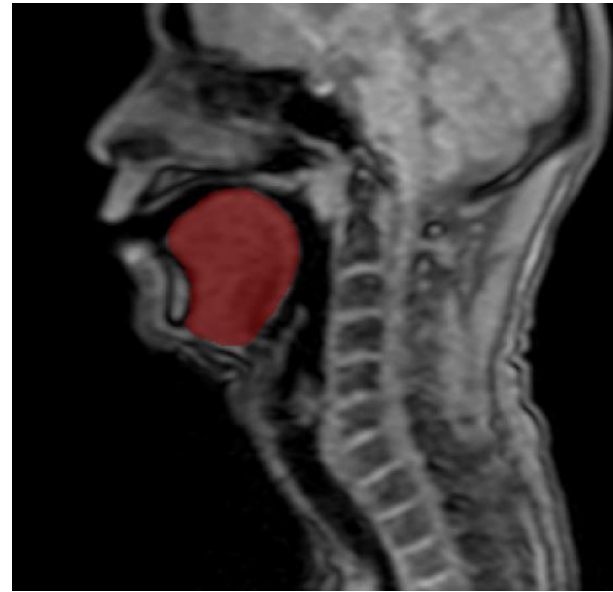


セマンティックセグメンテーション (2/2)

- 話者1名(A)の2000枚の学習データで学習
- IoU : 話者A 93.4%, 話者B 91.9%



話者A



話者B

AIへの期待

- 空間分解能・時間分解能の向上, 2D→3Dに利用できないか？
- アップスケーラー(Topaz Labs)による時間解像度の向上(4倍)



オリジナル



アップスケール後

超音波画像診断法

- 比較的安価，使用法容易
- レンタル可
- ソフトウェアの整備が進んでいる
- 矢状面以外の断面も観測可
- 制約
 - 撮像領域
 - 口蓋の造影が難しい
 - 下顎の動きによりプローブが動く

本日午後に2件の招待講演



富士フィルムメディカル iViz Air



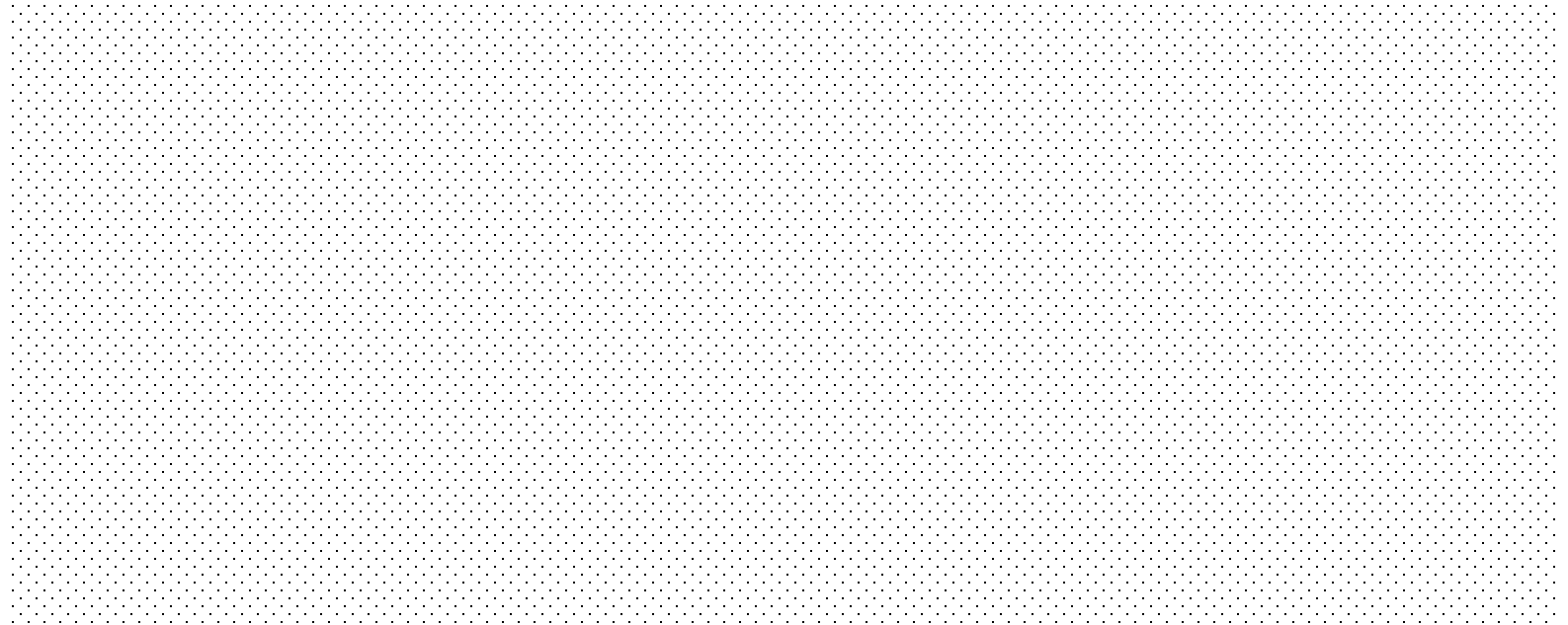
GE Healthcare
LOGIQ e Premium

プローブの固定 (1/2)

- プローブの固定法は悩みの種



Ultrafit headset
(Pucher *et al.*, 2020)



Li *et al.* (2023)

ALPHUS (Chen *et al.*,
2023)

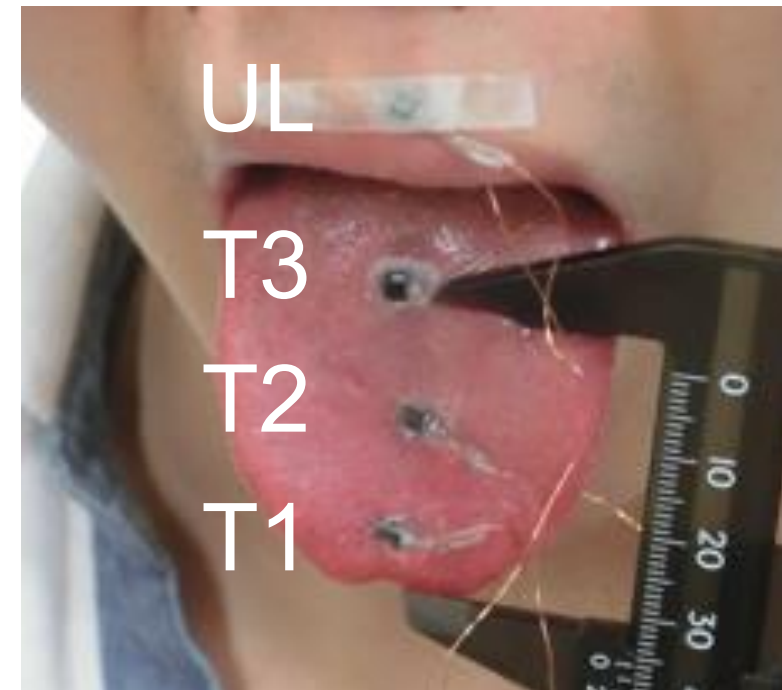
プローブの固定 (2/2)

- プローブ固定部にダンパ構造を入れてはどうか？
- 国語研にて試作中
- 市販のフェイスガード部品の流用により低コスト化
- 左右のアームで固定
- プローブ位置のトラッキングも実現したい



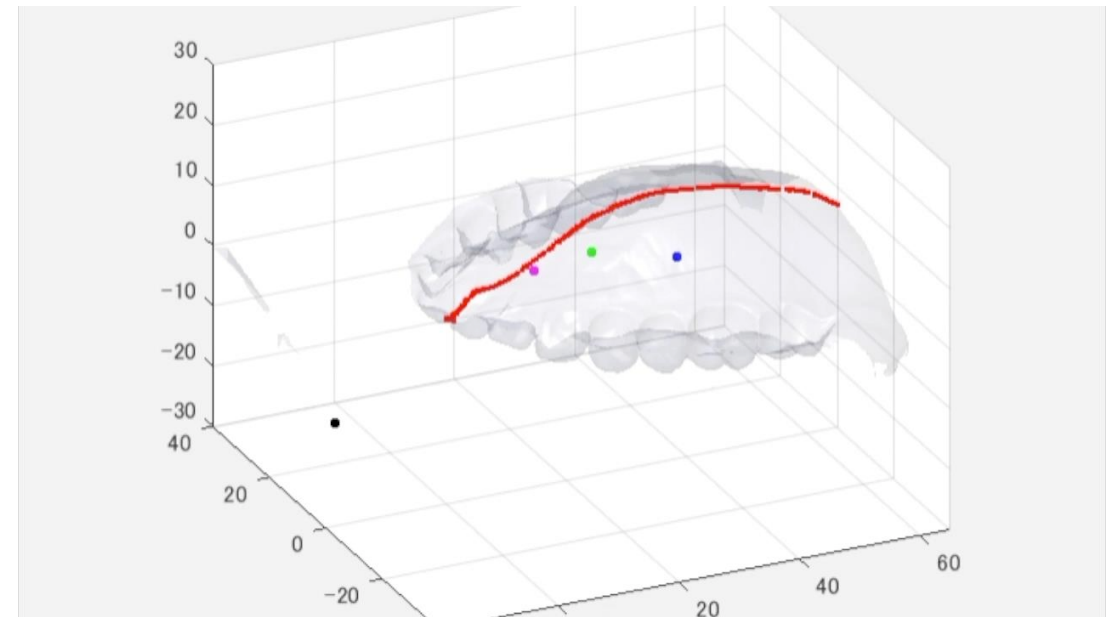
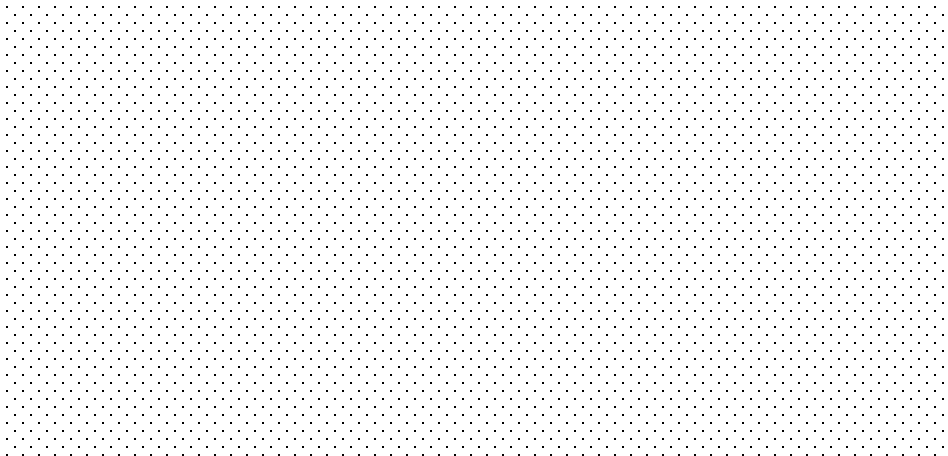
磁気センサシステム (EMA)

- 磁気を利用して調音器官に貼り付けたペレット/センサの位置をトラッキング
- 高時間分解能, 高精度
- 矢状面に限らない
- 点の情報のみ
- 計測点の再現性に難あり



口蓋座標系へのマッピング

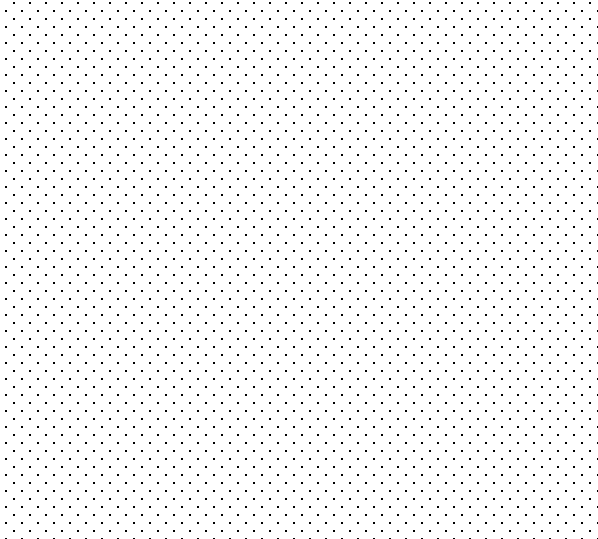
- センサ付きの口蓋印象採取トレイを開発
- 口蓋印象の形状を3Dスキャナで計測し，調音運動データにマッピング



「たちつてと」

顔面の動き

- 顔面の動きは，調音器官の動き，音声と関連が高い
- 画像処理AIによりマーカレスでランドマークの抽出が可能



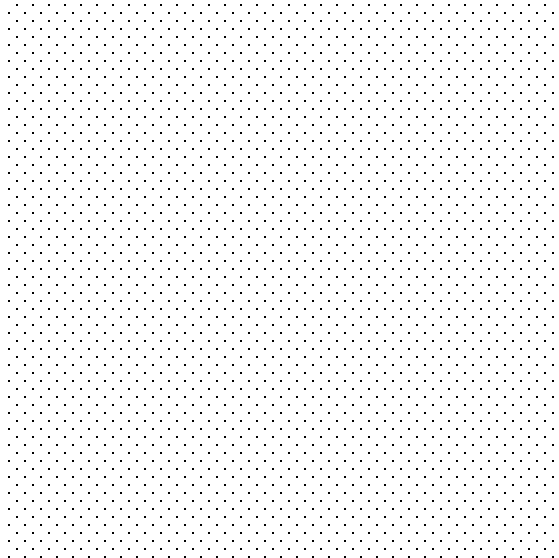
Yehia *et al.* (1998)



Mediapipe Face Landmark Detection
による抽出例

顔面皮膚振動

- レーザドップラ振動計による非接触な振動計測
- 発声様式により振動パターンが変化



© Polytec



ハミング



チューブ発声

国立国語研究所共同利用型共同研究(B)



大学共同利用機関法人 人間文化研究機構
国立国語研究所
National Institute for Japanese Language and Linguistics

English | 交通案内 | サイトマップ   Google 提供 

国語研案内 | **研究活動** | 大学院教育 | データベース・コーパス・資料 | 催し物

共同利用型共同研究 (B) の公募について

趣旨・目的

国立国語研究所が保有する研究資料・言語資源・分析装置を利用対象とし、日本語・言語・日本語教育に関する研究を行うものです。研究経費を配分するものではありませんが、国立国語研究所の専任研究者(研究教育職員)から研究のための助言を受けることができます。また、年度途中でも随時申請が可能です。

〔研究期間：研究開始日から1年以内。年度をまたがっても可〕

公募要項・申請書類

- 公募要項
 - 「共同利用型共同研究 (B)」 公募要項 [PDF 形式](#) [271KB]
- 申請書類
 - 国立国語研究所「共同利用型共同研究 (B)」 研究申請書 [PDF 形式](#) [183KB]、[Word 形式](#) [28KB]
 - 所属機関の長の承諾書 (ただし、非常勤等所属機関のない者については提出の必要なし) [PDF 形式](#) [96KB]、[Word 形式](#) [21KB]



トップへ

超音波
EMA
EGG

ハンズオンセミナー

- 実際に可視化技術を体験する機会を提供 (計測, 解析)
 - EGG, MRI, EMA, 超音波, EPG
 - コミュニティ構築の場にも
 - 国語研にて今年も開催するそうです
- 講義動画 (YouTube)



おわりに

- 音声生成系の可視化は調音音声学を支える技術
 - 日本音声言語医学会編, 「新編声の検査法」(2009) が詳しい
- 他領域と同様, AIとの連携が今後のポイント
- 可視化技術の利用者, 開発者のどちらも増えてほしい
- 関連の研究者の連携が必要

アイディアの先取権を保証しつつ, ネットを活用してゆるやかに連携し, 議論, 助言, 協力しあう体制の構築に期待