

分析のための音声・歌声の収録

榊原 健一 (北海道医療大学)

20H00291
21H01596
21H00497

簡単に自己紹介

miniCV

- — 1994-03 京都大学大学院理学研究科 数学専攻 (前期課程修了)
- 1994-04 — 2006-03 日本電信電話株式会社
 - おもにNTT コミュニケーション科学基礎研究所
 - 客員研究員: ATR人間情報科学研究所(2003—2005), 東京大学大学院医学研究科耳鼻咽喉科(2003—2004)
 - 非常勤講師: お茶の水女子大学文教育学部 芸術・表現行動学科 音楽コース(2004—2006)
- 2006-04 — 北海道医療大学
 - 客員研究員: 東京大学大学院医学研究科耳鼻咽喉科
 - 非常勤講師: 札幌大谷大学 音楽学部(2008—2018), 昭和音楽大学音楽学部 (2014),

研究テーマ (始めた順)

- 代数幾何学 (射影多様体上の楕円曲線の格子点の研究)
- コンピュータ音楽 (楽音合成, 作曲)
- 歌声研究 (生理, 合成, 分析, 知覚)
- 音声研究 (生成, 喉頭音源, 病的音声)
- 雪崩リスクマネジメント

音楽歴

- 作曲, 和声, 対位法を塩見直子に師事
- バンド活動でライブハウス, クラブに出演多数

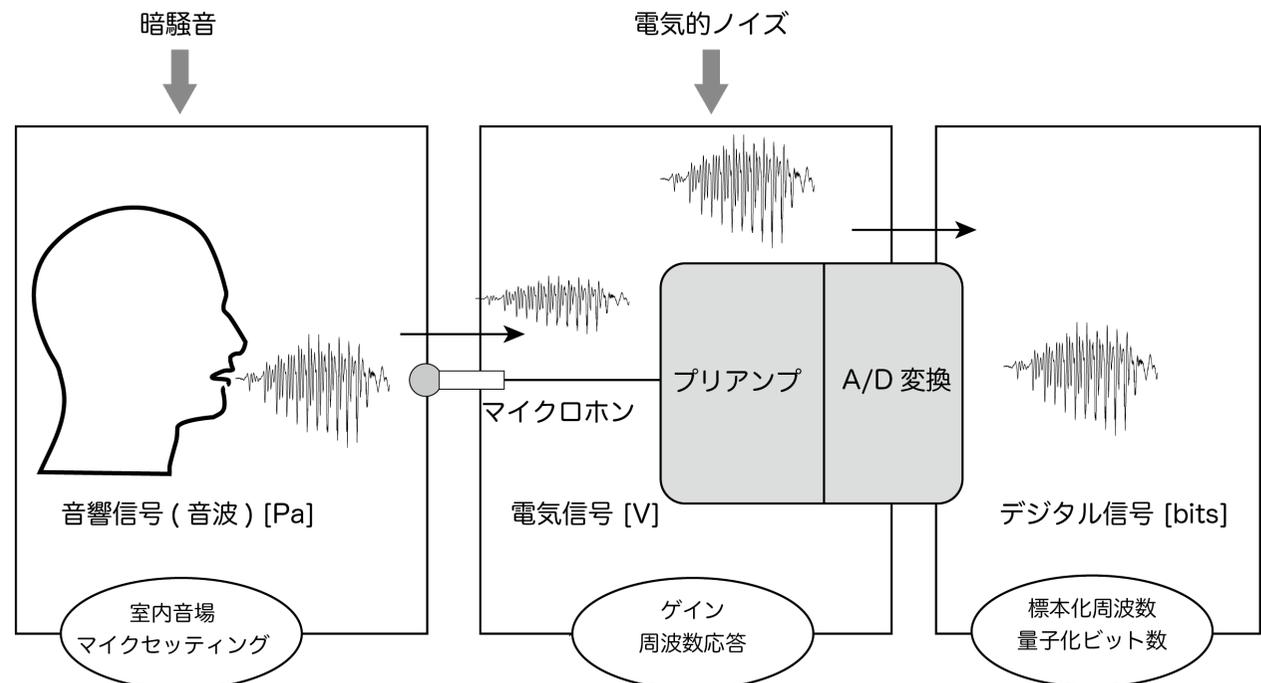


音声・歌声録音に関する Q & A

- 方法, 手順
- 使用機材
- 担保すべき品質
- 記録すべきこと

音声録音におけるキーポイント

- 機器
 - マイク, プリアンプ, A/D変換
 - 機器設定
- 音圧較正
 - 音圧レベル Sound Pressure Levelを記録し, デジタルデータの波形振幅を音圧 [Pa]に換算する.
- 環境
 - 暗騒音
 - 音場



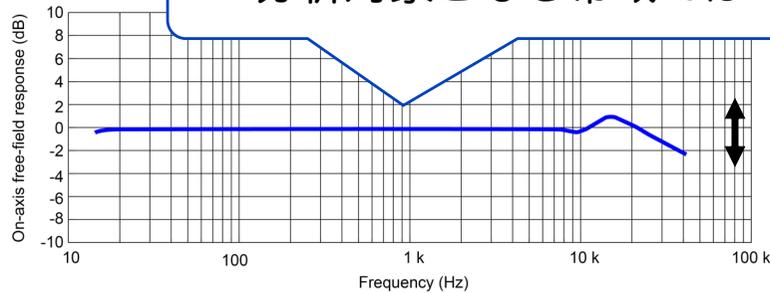
マイクホン microphone

電源供給が必要. 通常
48 V, 高電圧のものや
低電圧(エレクトレット
型など) もある

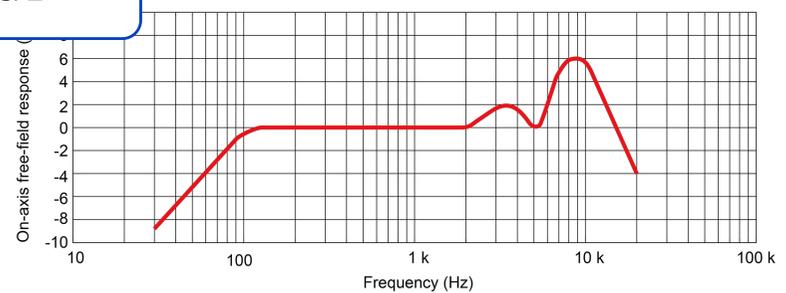
• コンデンサマイク (condenser microphone) を用いる

- ダイナミックマイク (dynamic microphone) は性能で劣る

分析対象となる帯域では < 2 dB



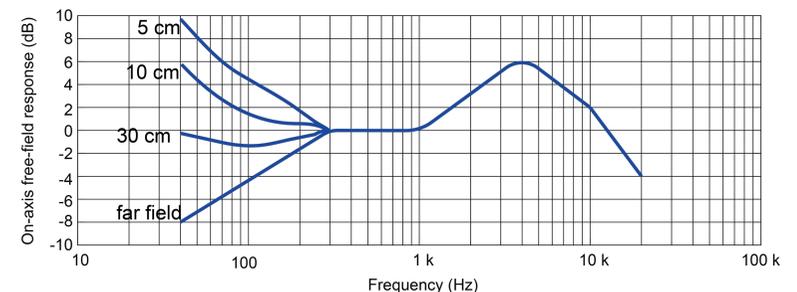
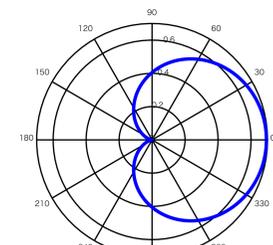
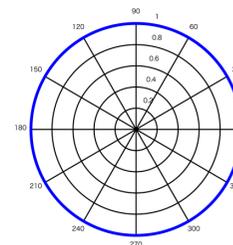
コンデンサマイクの周波数応答例



ダイナミックマイクの周波数応答例

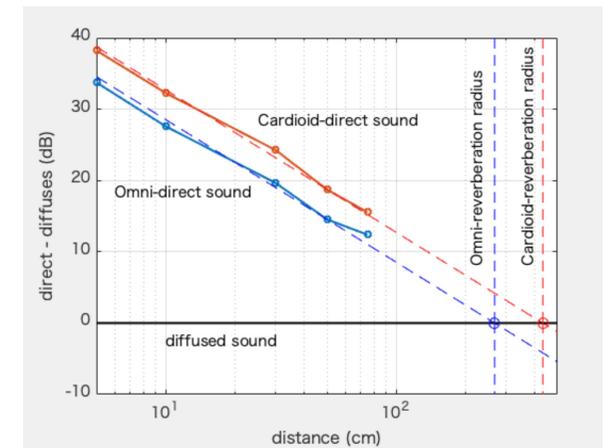
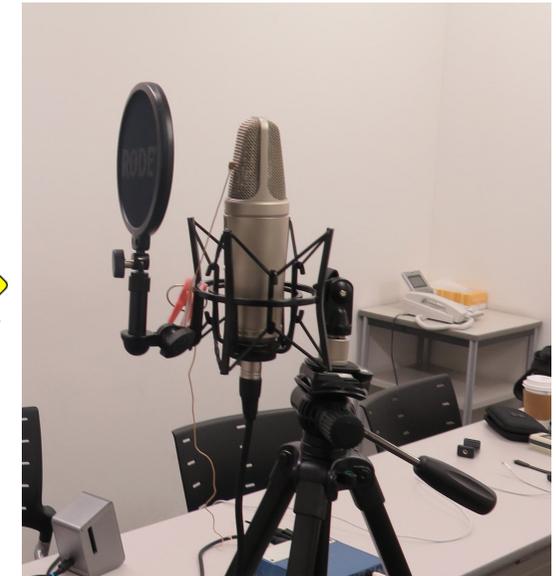
• 指向性

- 無指向性 (全指向性, omnidirectional)
 - すべての方向からの音波が入力する
- 有指向性
 - 特定の方向からの音波が選択的に入力す
 - Cardioid, Hypercardioid, Supercardioid, ...
 - 近接効果有り: 音源との距離が近いと
低音の利得が上がる



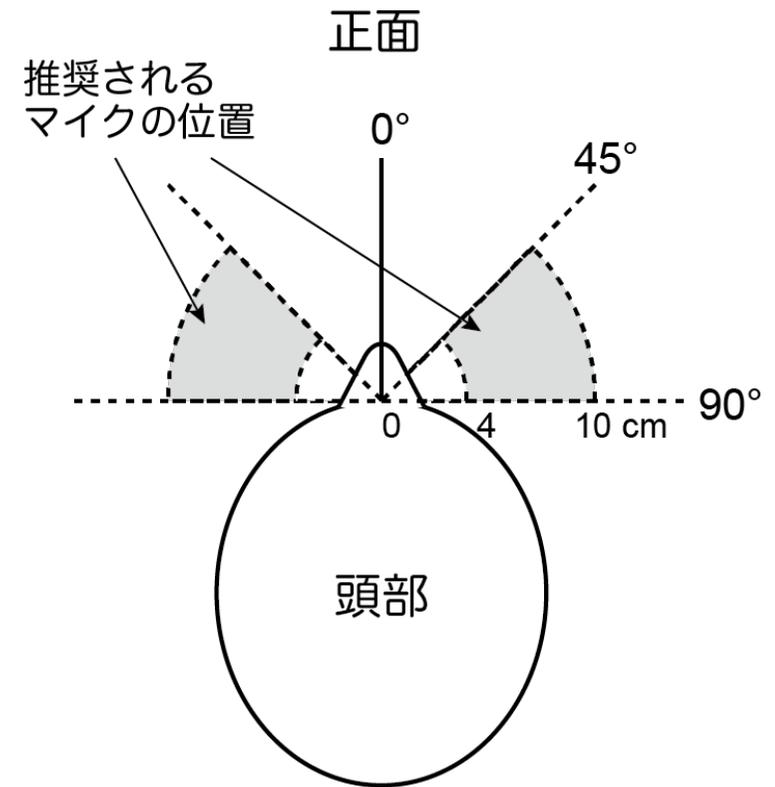
マイクのセッティング

- 良いSNRでの録音
 - 音声信号の音圧を上げる
 - 話者の口とマイクとの距離を近づける
 - 口とマイクの距離を近づける場合は吹かれ(popping noise)に注意
 - ポッピングスクリーンなど用いるのは可だがスクリーンの音響特性に注意
 - 指向性マイクは近接効果に注意（＜残響半径以内）
- 暗騒音の入力音圧を下げる
 - 指向性マイクも有効
 - 振動などの影響を避けるため、サスペンション付きのマイクホルダなどを用いるとよい



マイクセッティング推奨例

- ヘッドセット型の無指向性コンデンサマイクを用いる
 - 利点: 口とマイクとの距離を一定に保てる
- 口とマイクとの距離をなるべく近づける
 - 4—10 cm
 - 問題: 吹かれ 
 - マイクと口との距離は記録する
 - 較正方法によっては…
- マイクを正面から45°—90°の位置にする
 - 吹かれの防止



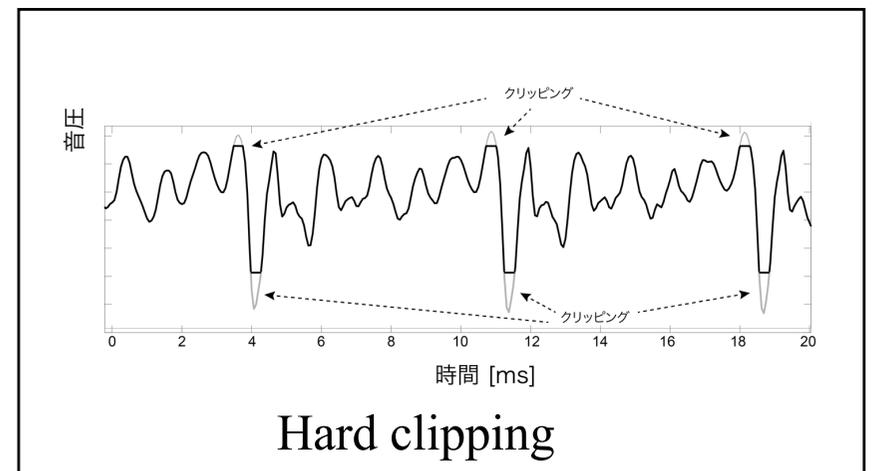
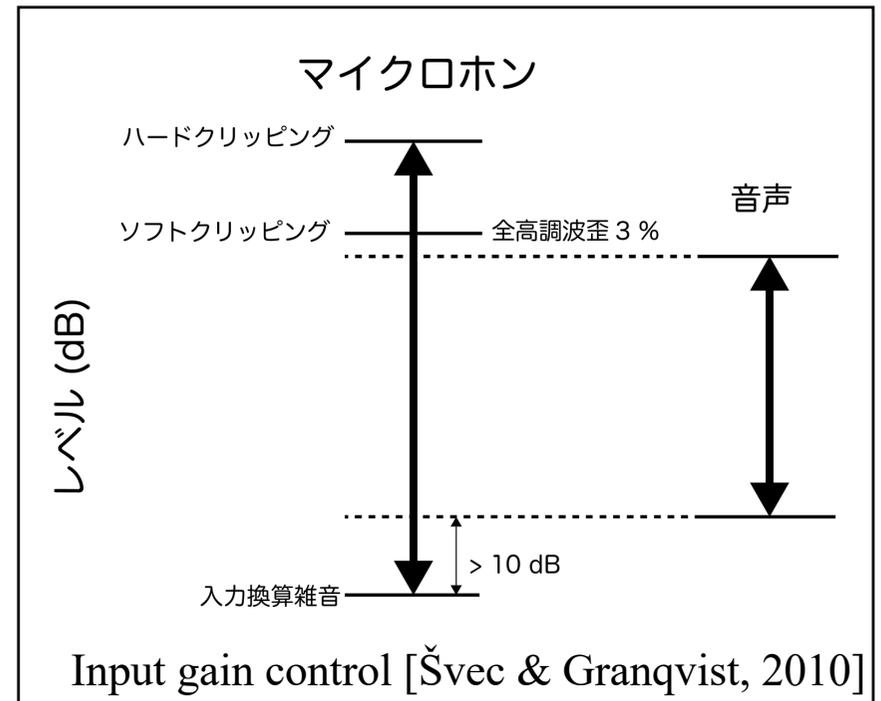
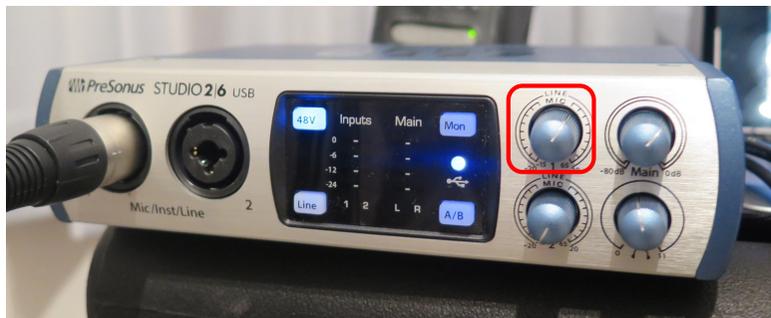
コンデンサマイクロホン

- 無指向性 vs 指向性
 - 無指向性の長所:
 - 近接効果無し, ポッピングノイズの感度低
 - 指向性の長所:
 - 背景雑音に強い
- スモールダイアフラム vs ラージダイアフラム
 - スモールの長所:
 - 許容音圧レベル高, 帯域広*, 音場への影響小, ダイナミックレンジ広
 - ラージの長所:
 - 等価雑音レベル低, 感度高,

*音声の収録で < 20 kHzならラージダイアフラムでも帯域は十分

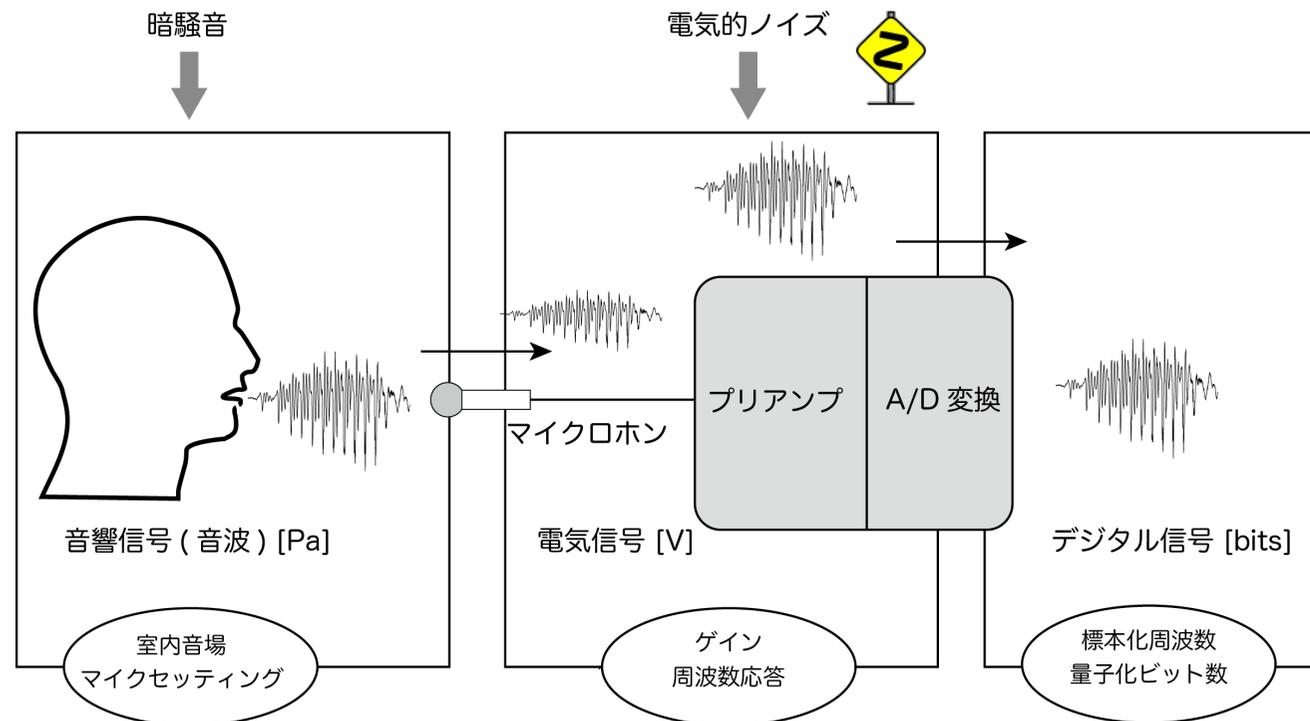
プリアンプ + A/D変換器

- 標本化周波数
 - 44.1 kHz 以上
- 量子化ビット数
 - 16 ビット以上
 - 入力ゲインを適切に設定
 - クリッピングに注意
 - 最も大きい声で決める
 - クリッピングしない程度になるべく大きく
- ファイル形式
 - 非可逆的 デジタル圧縮フォーマット(MP3, AAC, WMAなど)は厳禁
- その他
 - AGC (オートゲインコントロール) OFF



その他機器に関連して

- アナログ電気信号を搬送するケーブルは、バランスケーブルを用いる
- プリアンプなど用いる機器の電源供給はなるべく他と独立になるように切り離す
 - 安定化電源, ノイズカットトランスなどを用いる (++)
 - アイソレータなどを入れてグラウンドループを遮断

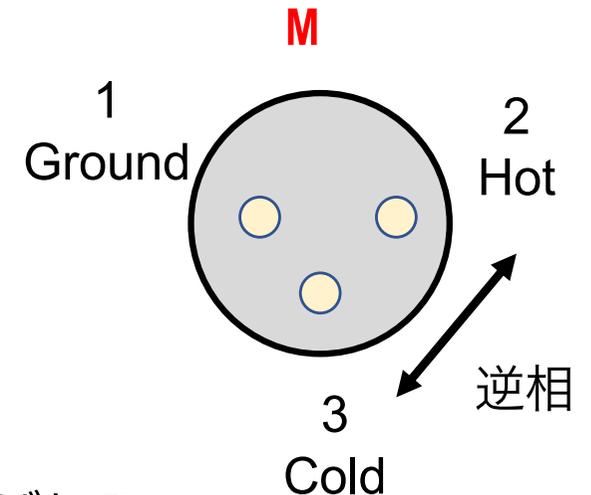
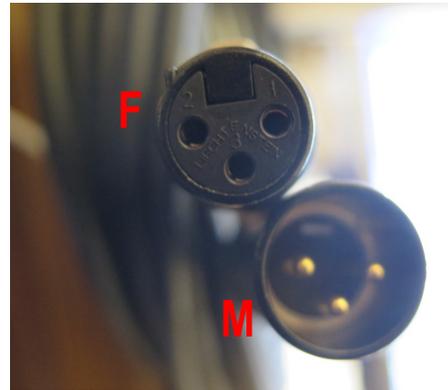


バランスコネクタ

電氣的ノイズに強い

バランスのコネクタ

XLR



Canon社が生産していたためキャノンとも呼ばれる
ロックできる

TRS



T: Hot

R: Cold

S: Ground

標準 (6.3 mm = 1/4 inch)

ミニ (3.5 mm)

ミニミニ (2.5 mm)

モノラルではバランスのコネクタ
(ステレオではアンバランス T: left, R: right)

フォンプラグ (phone plug)

アンバランスコネクタ

TS



T: Signal
S:Ground

標準 (6.3 mm = 1/4 inch)
ミニ (3.5 mm)

RCA pin



Pin:Signal
Ring :Ground

BNC



- アンバランス同軸ケーブルのコネクタ
- ロックできる

microdot



- 小型のアンバランス同軸ケーブルのコネクタ
- ロックできる
- ヘッドセットマイクやピンマイクなどの端子として使われる

mini XLR

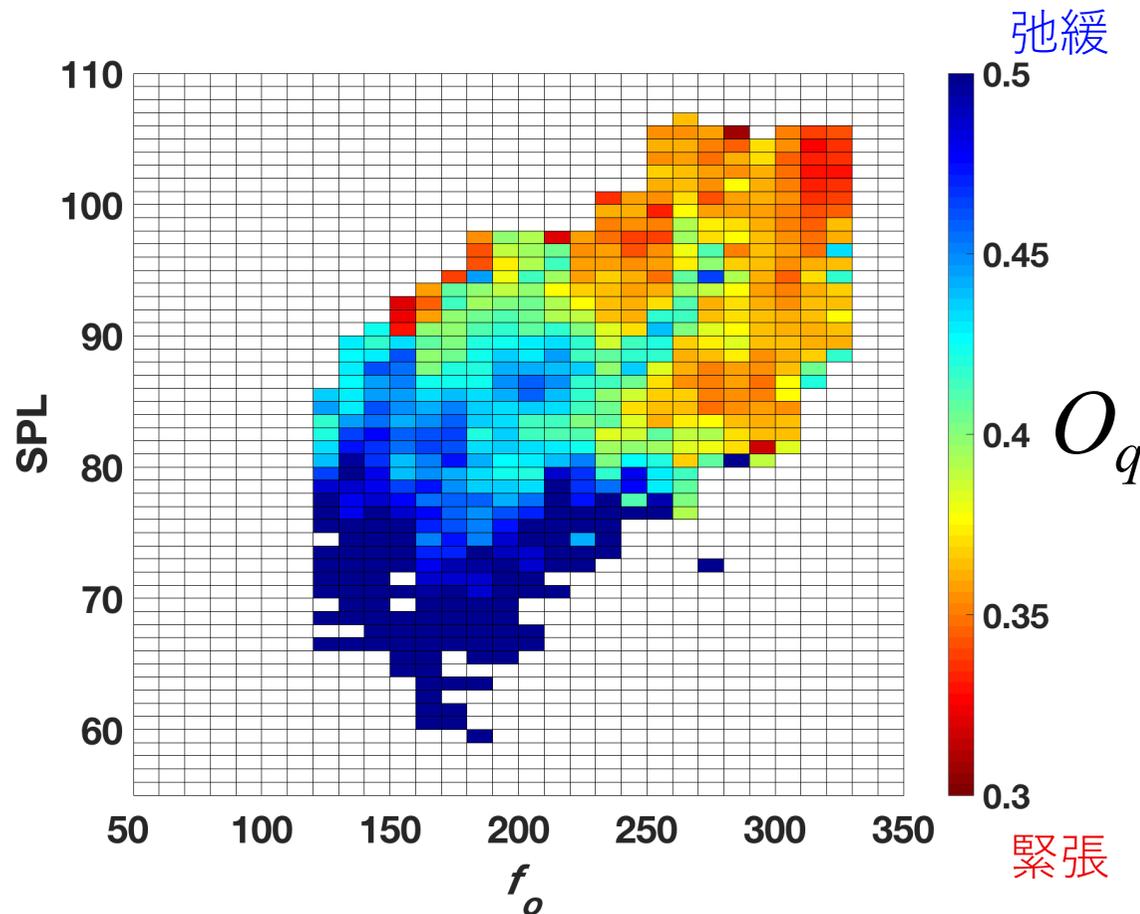


TRRS



- RRの機能割当は機器によって違う

声の音圧レベルは重要なパラメータ



- 声の音圧レベルは声の様々な特徴量と従属関係にある

- f_0
- 声質
- 声門下圧



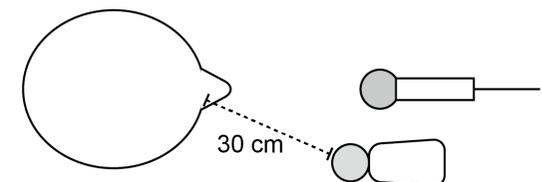
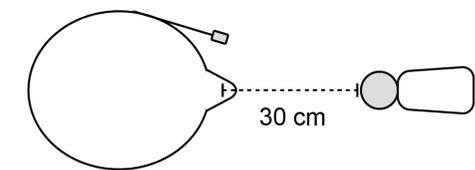
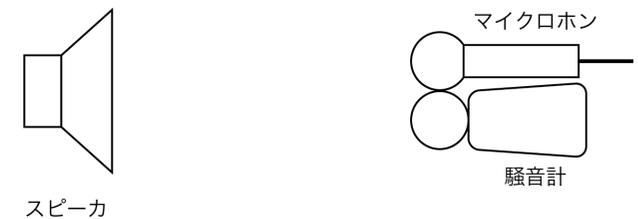
発声努力 (vocal effort) と関連する最重要パラメータ

若狭, 寺澤, 河原, 榊原, 音講論, 2019春

特に歌声, 感情音声など声質研究では音圧較正は不可欠

マイクの音圧校正

- 音圧校正専用機器（ピストンホンなど）を用いた校正
 - 高価
 - 対応可能なマイクのみ
- スピーカ + 騒音計
 - スピーカの特性に依存はしない
 - 校正信号が再生できる周波数特性のもの
 - 校正精度は騒音計の精度に依存する
 - 校正信号
- 声 + 騒音計
 - 持続母音を用いる



音声の音圧レベルは @ 30 cmでの音圧レベルで表示する (Patel et al., 2018)

音声収録用のマイクロフォン



騒音計

試験信号用のスピーカー



マイク音圧校正: 校正信号と騒音計

校正信号

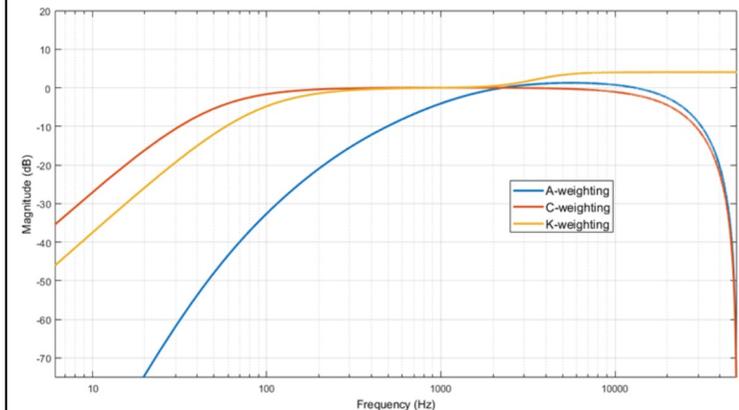
- 1000 Hz純音(おすすめではない)
 - A特性やC特性の重み付けを考慮しないくていい
 - 反射がある環境ではマイクの位置によって定在波の影響を受ける
- 中心周波数1000 Hzの狭帯域雑音
 - A特性やC特性の重み付けをほとんど考慮しないくていい
 - 定在波の影響によるマイク位置の制約はほとんど無い



https://drive.google.com/drive/folders/1iQrNyWB_25UtGl-XS65i64jQ_NUP4dGH?usp=sharing

騒音計

- IECクラス1または2が望ましい
- スマホのAppsの騒音計を用いる場合精度は落ちる (Faber, 2017)
- 重み付けを逆算すればA特性でもC特性でもOK
- 重み付けを考慮しない場合はC特性の方が誤差が小さい



較正信号 60秒



Thanks to Hideki Kawahara

音圧校正の手順

録音機材を設置しマイクロホンと話者との位置を決める

最大音圧となるタスクでクリップしないように入力ゲインを設定

マイクを話者から離し騒音計をマイクの直ぐ側に設置

試験信号をスピーカーから出力し音量を調節

QRコード

A特性でSlowまたは時間平均Leqを記録
&
試験信号を録音

マイクと話者の口との間の距離（位置関係）を記録

環境

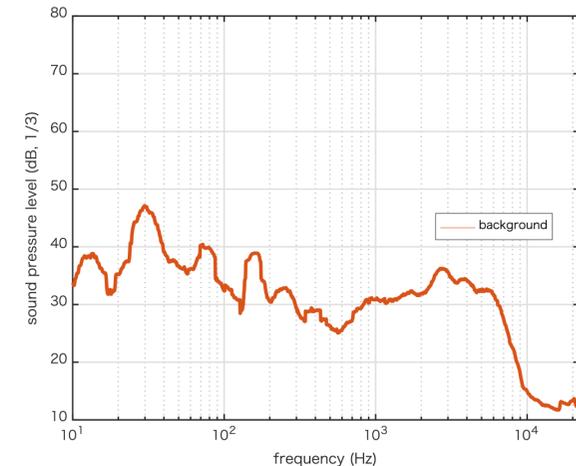
例えば…

「反射の少ない静かな部室で録音」 「X社の防音室内で録音」
音響的な情報は殆ど無いに等しい

- 暗騒音をデータとして収録
 - 本来は周波数特性などは必要な情報
- 声質分析の場合は、暗騒音レベルが録音する最も小さい声よりも10 dB以上小さいことが必要 (Deliyski et al, 2005)
 - 目安: $L_A < 25$ dB , $L_C < 38$ dB, at 30 cm

残響:

- 可能であれば録音環境のインパルス応答を残しておくといい (スピーカなど再生系の特性も含んでしまうのには注意)
- 残響半径でもいい



部屋が狭い場合

- ノイズ源から遠かる
- 指向性のマイクの場合、ノイズ源から軸をずらす
- 壁など反射の影響を下げるように遠ざかる
- 吸音の工夫
- 共振する物体を部屋から出す（なるべく録音する部屋にものを置かない）

手順（例）

1. 適切な機器, 環境のセッティング



2. 機器の入力レベルの調整



3. 音圧較正



4. 暗騒音の録音



5. 音声収録

録音ツール---ソフト編

- 録音ツール
 - Audacity: よく使われている
- 切り出しツール: 作っておくとよい
 - 立ち上がり, 立ち下がりの閾値, 前後のマージン, 無音時間の長さなどをパラメータに自動的にファイルを切り出すようなツール
- フォーマット変換
 - Resampling
 - ファイル形式
 - チャンネル変換
- モニタリング
 - 波形, スペクトログラム
 - オシロスコープ....

録音に関する情報

- 機材の型番
 - ファームウェアバージョン
 - スマホだとハードウェアとOS, 騒音計Appsのバージョン
 - 機材のスペック表は保存
- 録音設定
 - サンプリング周波数, 量子化ビット数, マイクと人との位置関係
- 発話者 (年齢, 性別, 既往歴, 歌唱経験など)
- 録音者
- 日時, 場所
- 音圧校正の情報
- 録音環境 = 暗騒音, 音場情報

まとめ

- マイクはコンデンサマイクを用いる
- 指向性マイクは近接効果に注意
- 無指向性マイクを近づけて使う場合は吹かれに注意
- 暗騒音を録音しておく
- マイクの音圧較正は必須
 - 難しく無い
 - 正しい手順で手際よく
 - L_p @ 30 cmの値で表す

Remark

- GithubのHideki Kawaharaにあるコードがいろいろ使えます
- 録音の品質と音響特徴量の分析精度についてはSSの河原, 榊原他を参考に
- 雑音除去技術も進歩している (特徴量によっては有効)

さいごに

- 録音が終わってから悔やんでも遅いので、周到に準備をすること