

高品質な音声収録のための基礎技術

森 大毅
宇都宮大学工学部

発表者の背景



- 現在 宇都宮大学工学部電気電子工学科 助手
- 研究課題
 - 音声対話システム
 - 対話音声に含まれるパラ言語情報
 - 知覚
 - 音響関連量
 - 音声分析
 - 音声合成
 - 音声認識

音声伝えるパラ言語情報の音響的 キュー



- 韻律的特徴

- 基本周波数, パワー
- 時間情報
 - 話速
 - ポーズ
 - タイミング

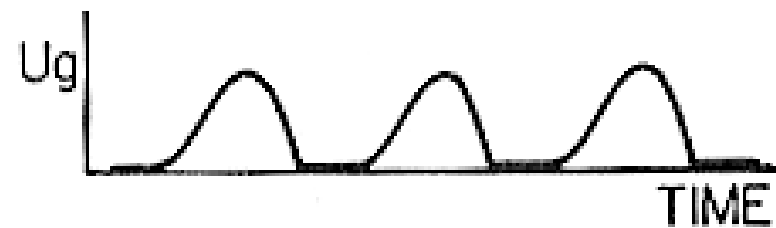
- 分節的特徴（スペクトル）

- 声道特性（フォルマント）
- 喉頭音源特性
 - 喉頭雑音
 - 声門開放率
 - 音源スペクトル傾斜

声質と喉頭音源特性

声門気流

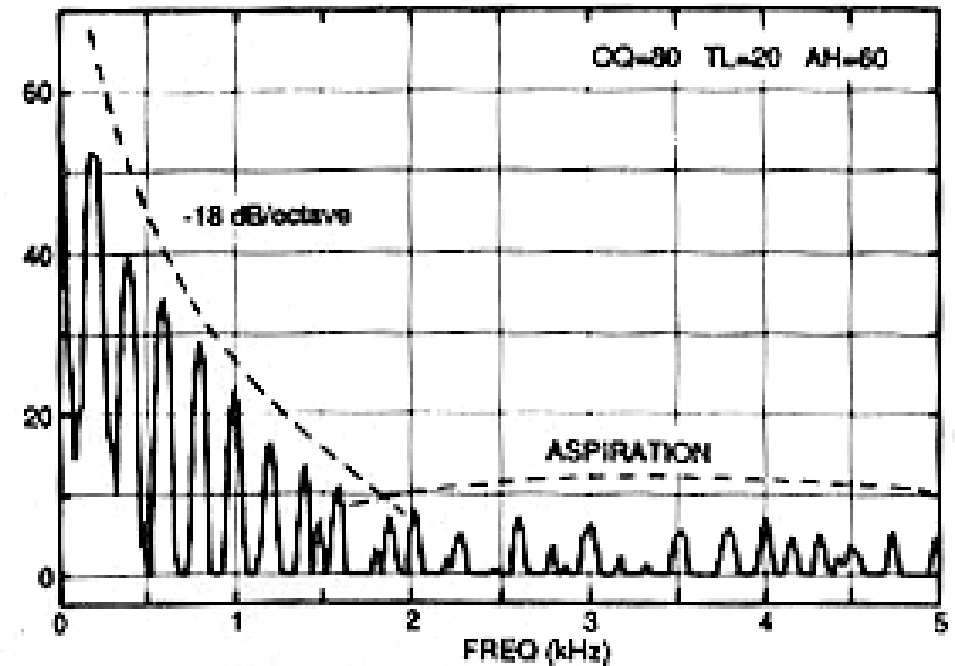
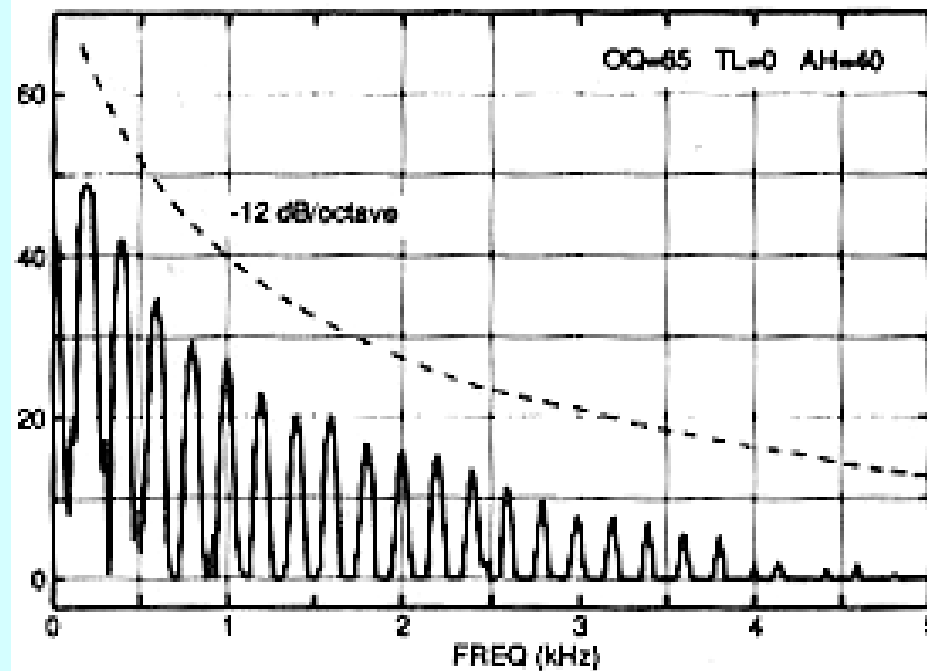
通常 (modal) 発声



気息的 (breathy) 発声



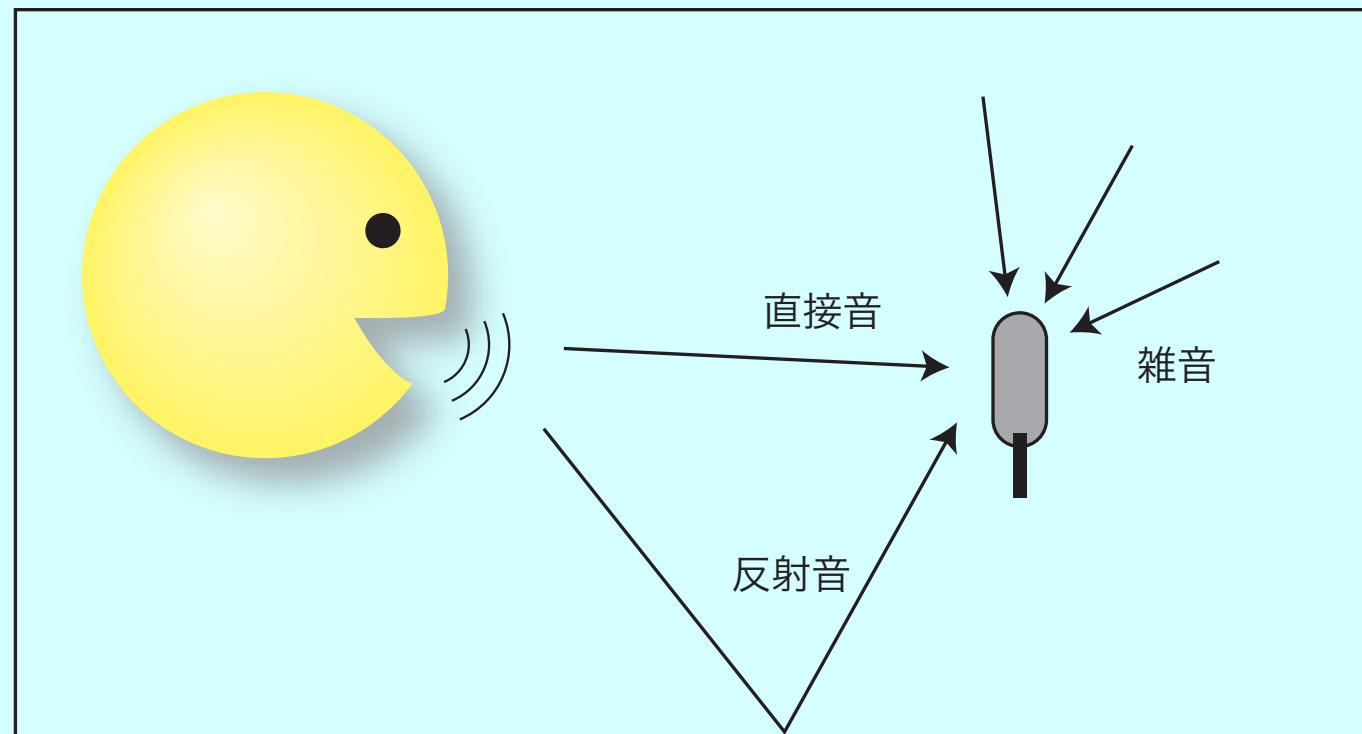
音源スペクトル



(Klatt, D.H., Klatt, L.C., J. Acoust. Soc. Am., 87, p. 822, 1990)

マイクロホンの役割と望まれる性質

- 音波の信号を電気信号に変換する
- 元の情報はできるだけ変化させない
 - 元の情報に含まれていない成分はできるだけ小さく
 - どの周波数成分も同じ感度が理想

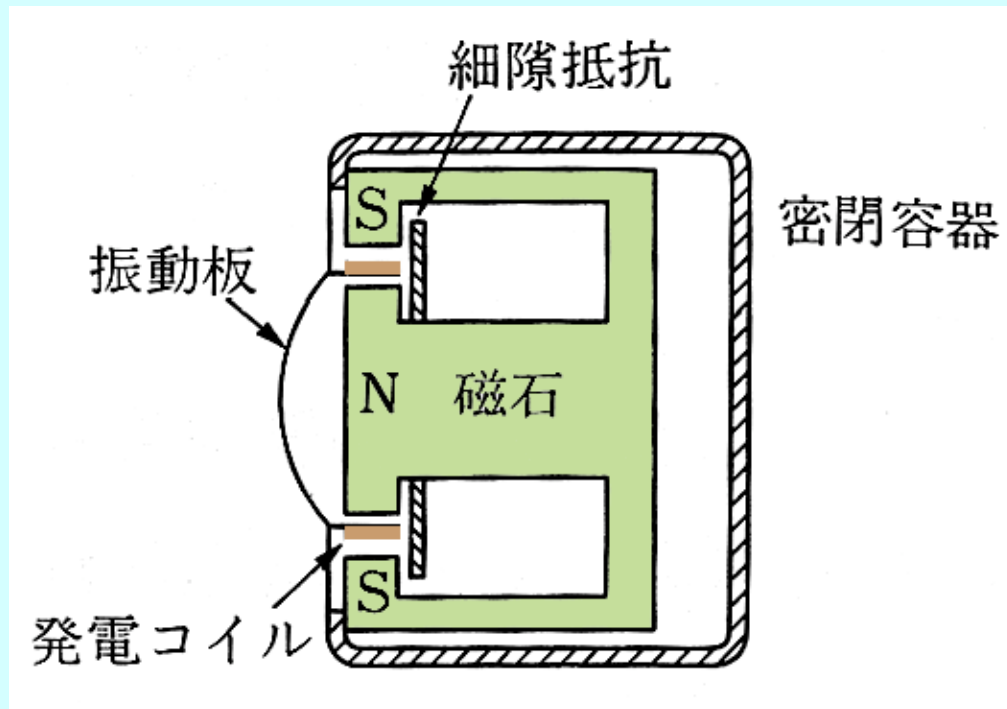


マイクロホンの分類



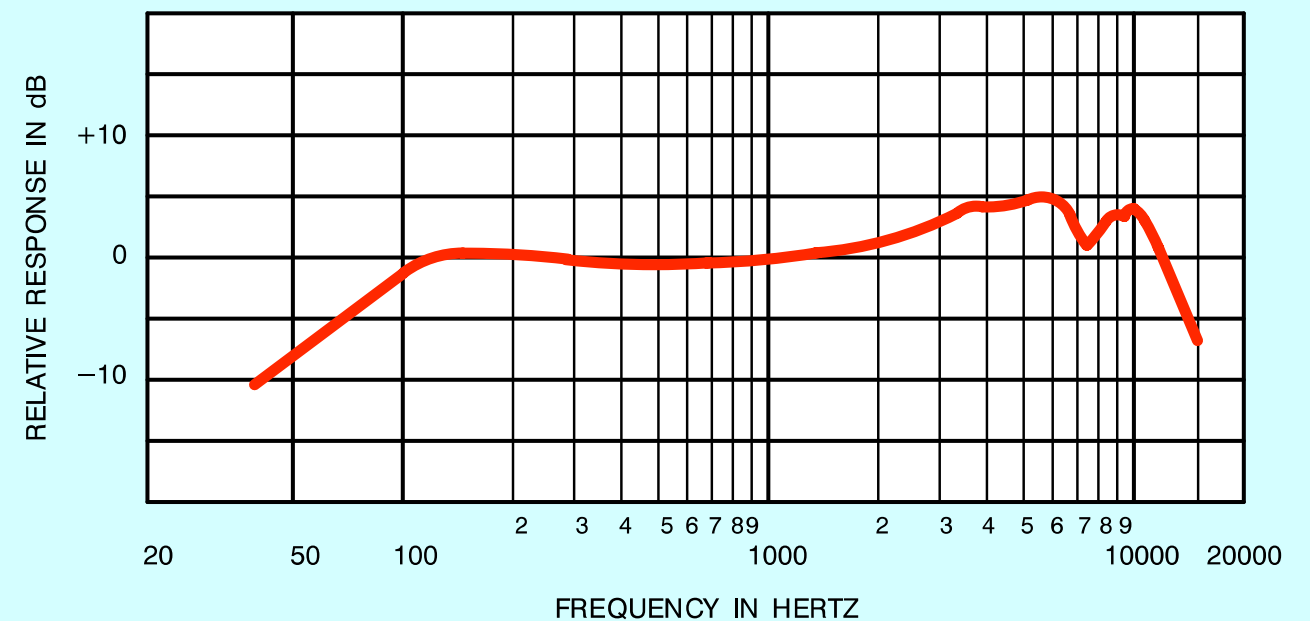
- ダイナミックマイク
 - ムービングコイル型
 - (リボン型)
- コンデンサマイク
 - コンデンサ型
 - エレクトレット型

ダイナミックマイク

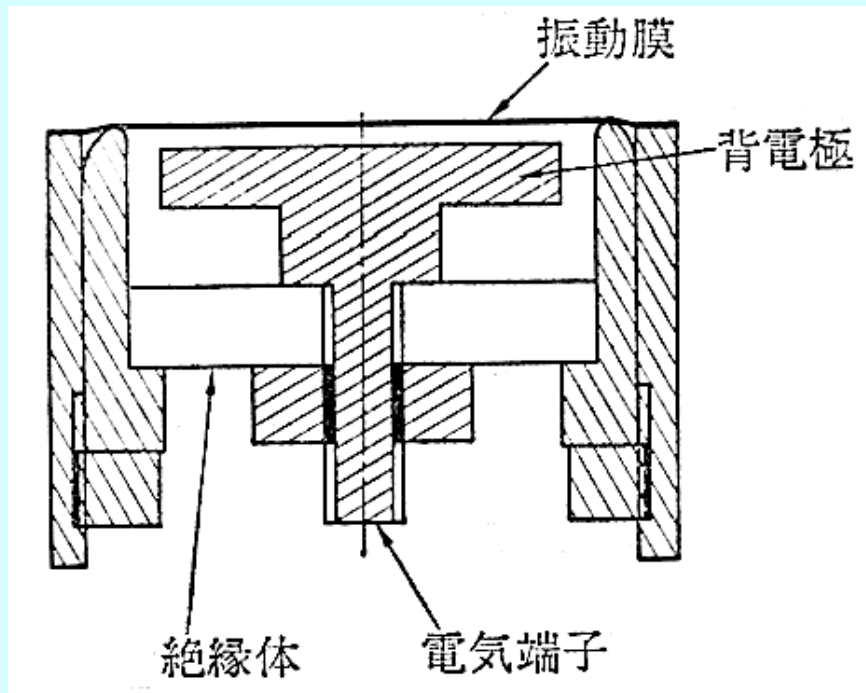


シュアー SM58

- 電源が不要
- 丈夫
- 大きい
- 特性は劣る



コンデンサマイク



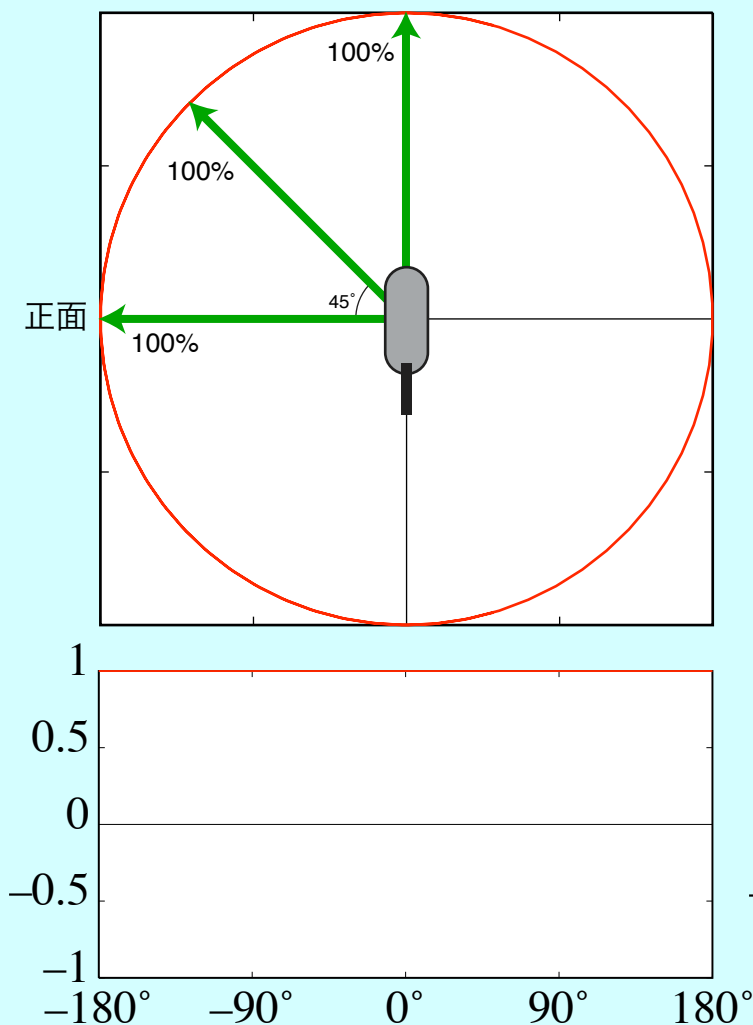
ソニー C-38B

- 電源が必要
 - マイクケーブル経由で 48V を供給するシステムが多い
- 湿気、風などに弱い
- 小型のものがある
- 特性に優れる

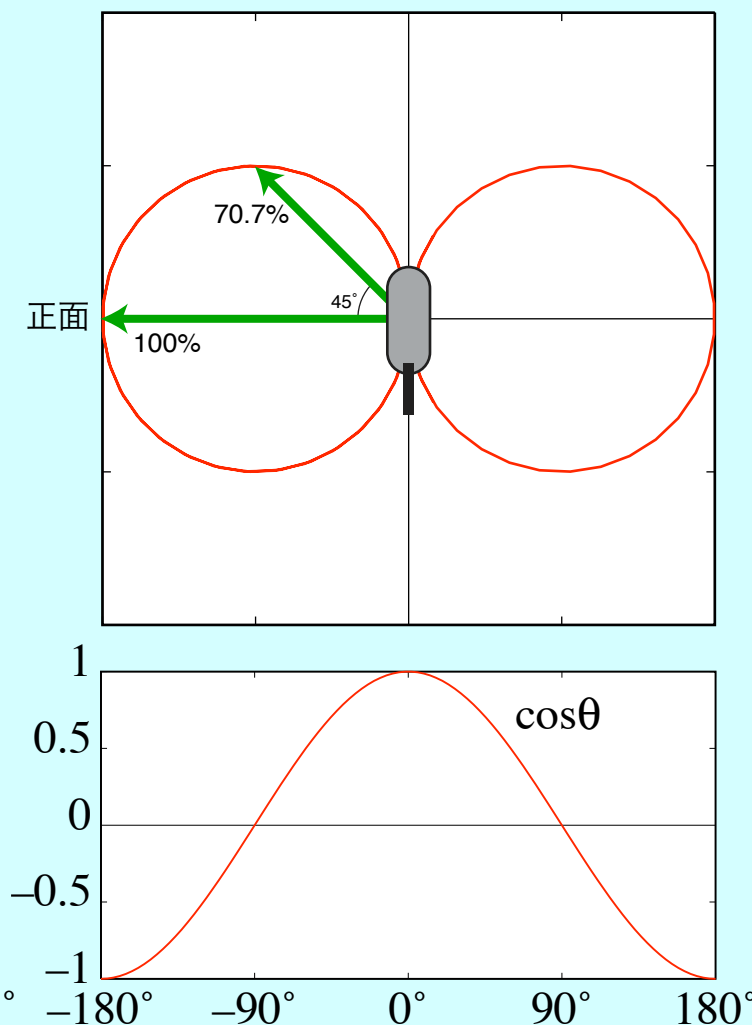
指向性

- 音の到来方向によって感度が変化する性質

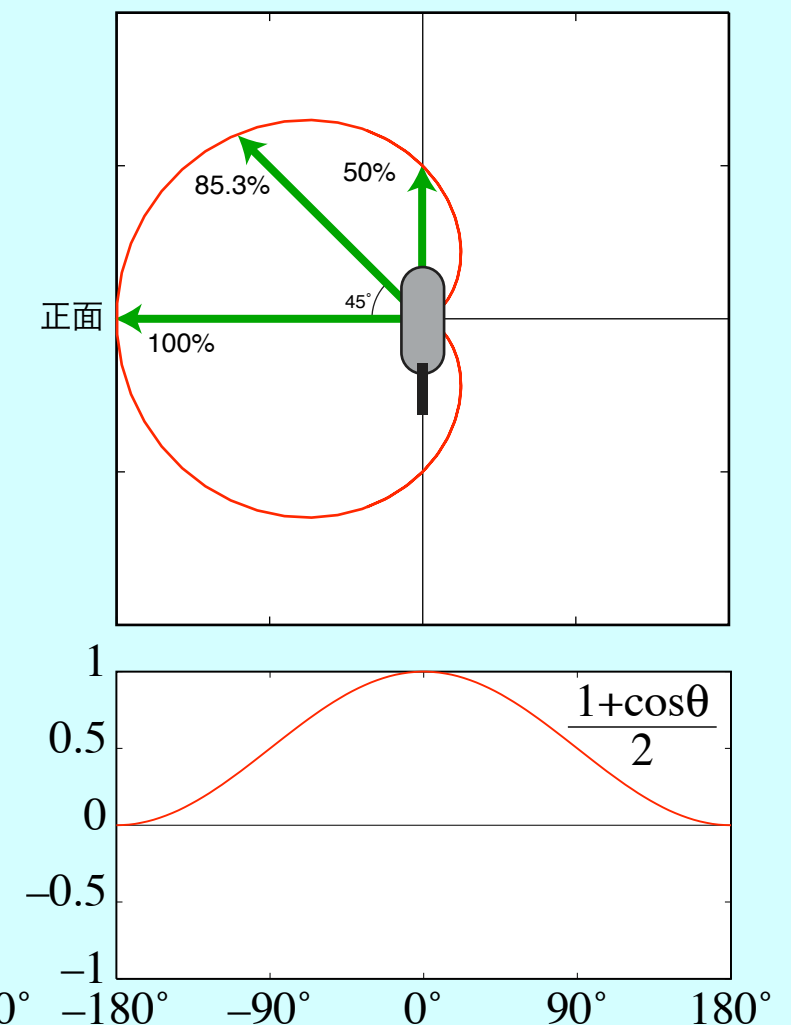
(a) 無指向性



(b) 双指向性

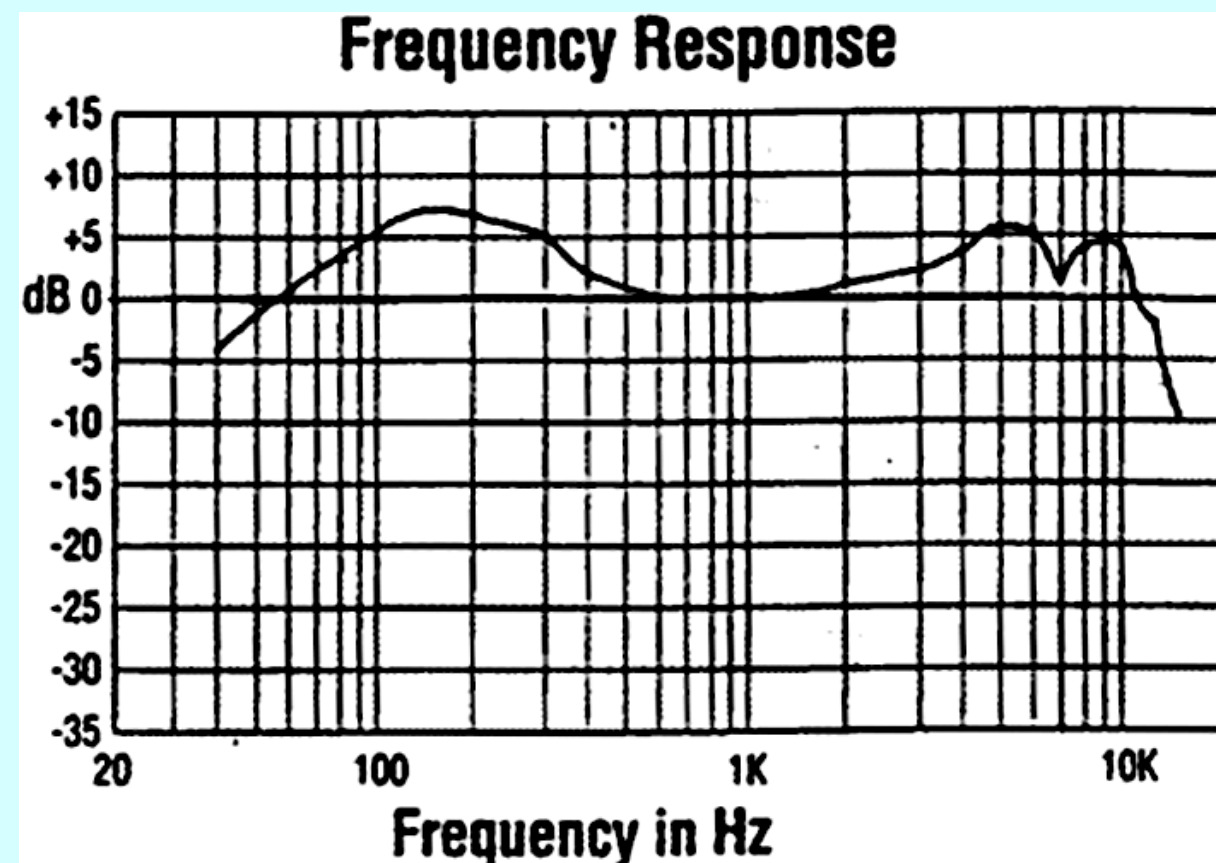


(c) 単一指向性



近接効果

- 音源の近くで指向性マイクを使うと低い周波数成分の感度が増大する現象
- 無指向性マイクではほとんど起こらない



人工口唇から4cmの位置で単一指向性ダイナミックマイクを用いて測定した周波数特性
(Bartlett, B., AES Preprint #4517, 1997)

近接効果のデモンストレーション

- ソニー ECM-360

- エレクトレット型
- 単一指向性



- 口唇からの距離

- 約15cm



- 約2cm



- 「桜は中国やヒマラヤにもありますが、日本の桜は種類が多く、また美しいので有名です。」

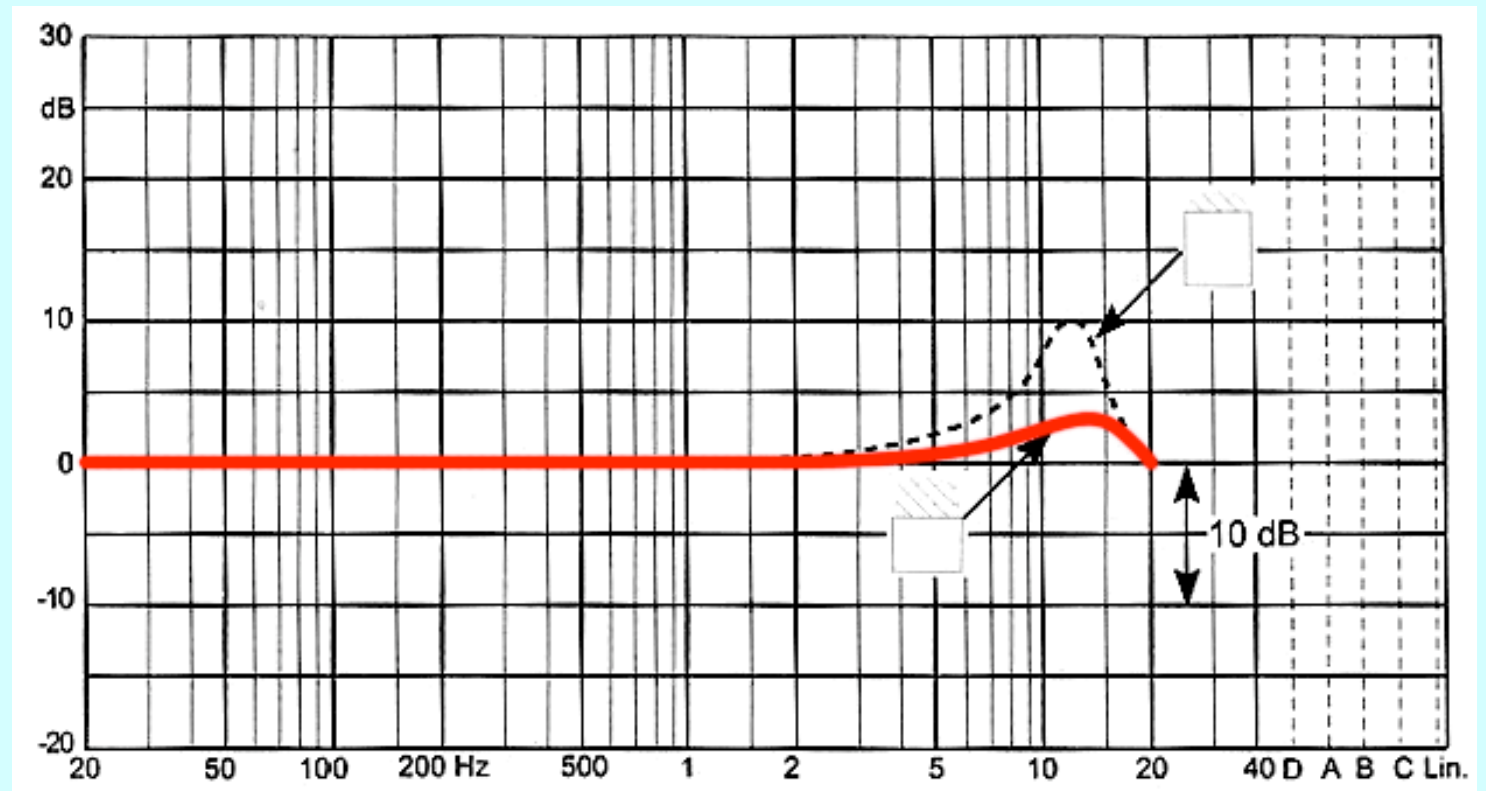
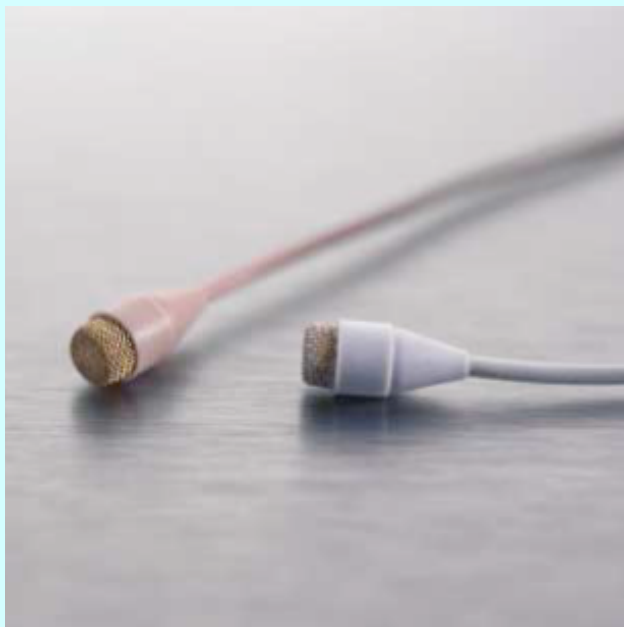
高品質な音声収録のために



- (エレクトレット型)コンデンサマイクを使う
 - 周波数特性が平坦なため
- できるだけ口唇に近い位置にマイクを置く
 - 反射音・外来雑音に対する音声信号の比を大きく取るため
- 無指向性マイクを使う
 - 近接効果を避けるため

DPAミニチュアマイク

- 4060 (高感度型), 4061 (低感度型)
- エレクトレット型
 - 電源はケーブル経由で供給
- 無指向性
- 直径 5.4 mm



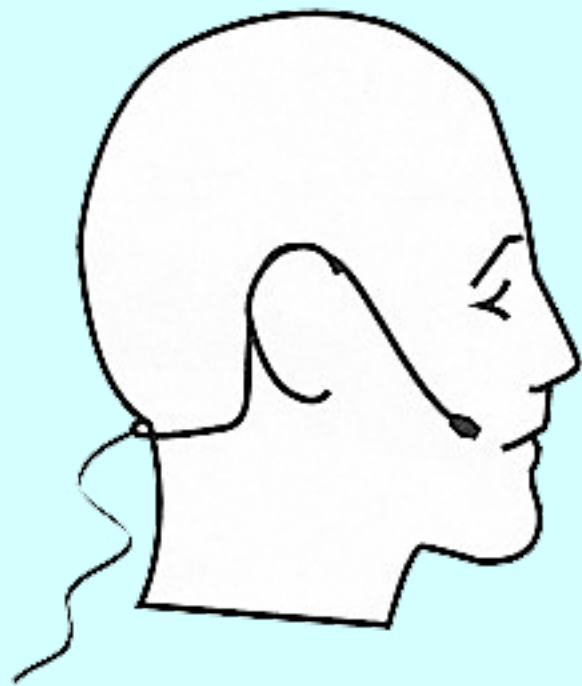
収録時のマイクセッティング

- ブームスタンド
- ヘッドセット
- etc...



DPA 406x のマウント法

- クリップでピンマイク風に
- ヘッドセットにクリップどめ
- ヘッドバンドを利用



音声のデジタル表現



- 符号化法

- **PCM** (lossless)
- ADPCM, ATRAC3, MP3, AAC, GSM... (lossy)

- ファイル形式

- **Microsoft WAVE (*.wav)**
 - PCM, μ -law, A-law, ADPCM, GSM
- AIFF
- Sun Au (*.au)
- MP3
- NIST SPHERE
 - lossy mode, lossless mode

音声収録における推奨方式



- PCM
- 量子化精度: 16ビット
- 標本化周波数: 44.1 kHz or 48 kHz
 - 編集・加工・刺激呈示時には落としても構わない
- ファイル形式: Microsoft WAVE

録音機材選択のポイント



- PCM (無圧縮) で記録可能か
- 録音レベルは十分に大きく取れるか
- 記録時間は十分に長く取れるか
- コンデンサマイクに電源を供給可能か
- 携帯性
- それ自体が音を発生しないか

マランツ PMD670

- ポータブルメモリレコーダ
- コンパクトフラッシュ
 - 1GBで約200分録音可能 (@44.1 kHz PCM, 1ch)
- マイク入力
 - XLR × 2
 - コンデンサマイクに電源供給可能



- TRSバランス × 2
- コンデンサマイクに電源供給可能



音声刺激の評定区間の指示 — QuickTime™を利用した方法 —

自由発話に対する心理実験における 先行条件

A: じゃあシーが最初でしょー。

A: きっと。きっとそうだよね。





[00:00:04.037]

評定区間指示の例



A: うん。

B: 次にスヌーピーが持って来てー、

A: あっそれかもしれない、

B: そっちだよね多分ねー。

A: そうだそうだそうだそうだ。

