
発話の開始・終了の同定のために： 分析法と機器の選定

日本学術振興会・京都大学
長岡千賀

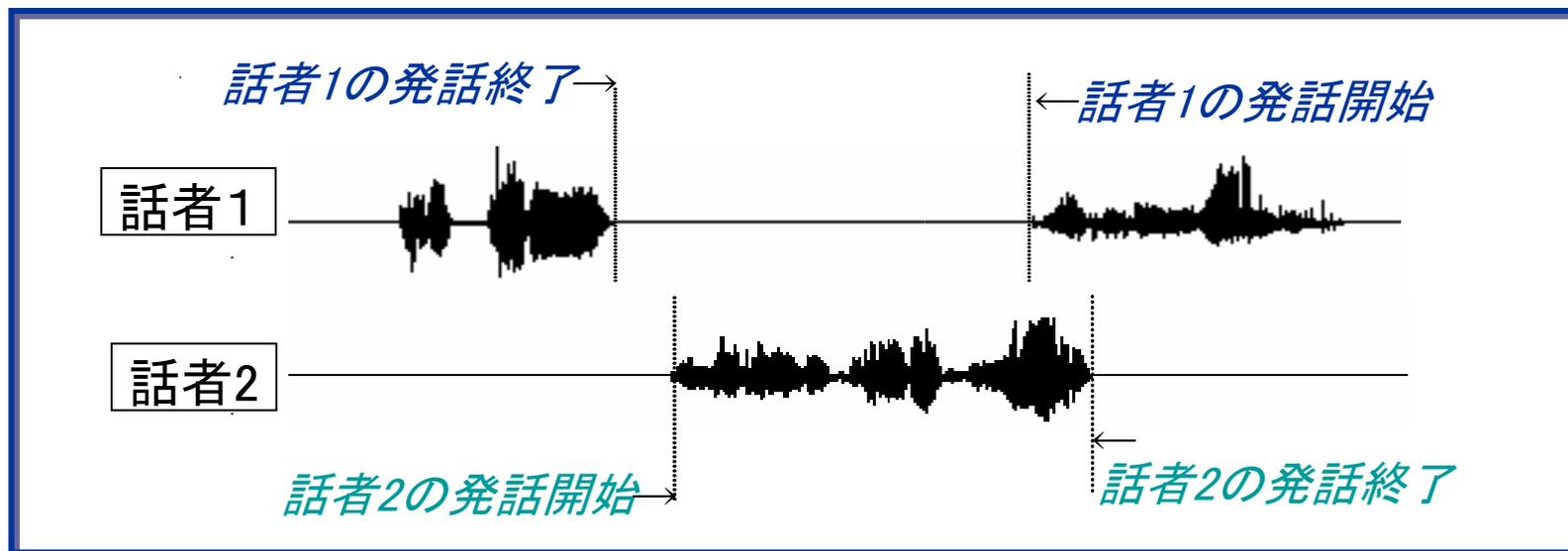
発表者の背景

- 大阪大学人間科学研究科
感性情報心理学講座(中村敏枝教授) 出身
(04年03月まで)
- キーワード: 音声対話, 対話の「間」(反応潜時)

発表の内容

- 発話開始・終了の同定
 - 音の強さ・大きさの分析方法の紹介
- 2話者の音声を別チャンネルに収録するために有効なマイク
- 防音室の設置

発話開始・終了の同定



手順

1. 声の強さ・大きさを分析する
2. 分析データを手がかりに発話の開始, 終了を同定する

声の強さ・大きさの分析方法

- 音圧レベル (Sound Pressure Level: dB)
 - 比較的一般的なソフトで分析できる.
 - これまで多くの研究で基準として使用されている.

デメリット: 物理量である. 人間の聴覚特性に忠実とはいえない.
- 騒音レベル (A-weighted Sound Pressure Level: dB(A))
 - 人間の聴覚特性に合わせて補正されている.

デメリット: これまで騒音レベルを分析できる機器が高価だった.

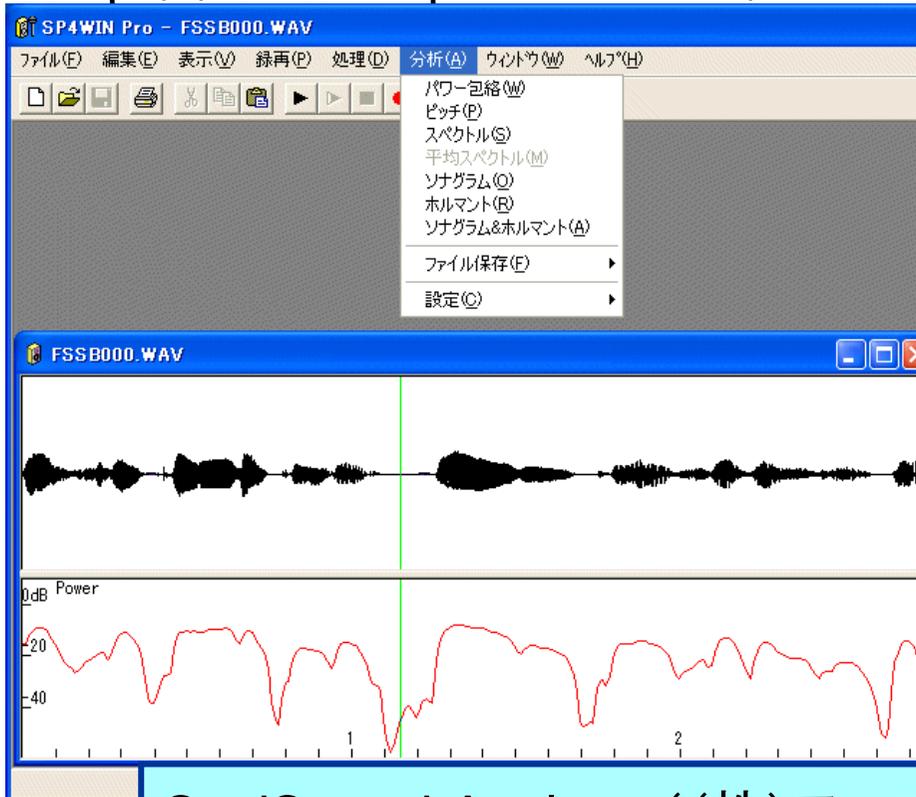


SA-29 (1chタイプ)
大きさ: 約21 × 30 × 8 cm
価格: 100万前後

音圧レベルの分析:

例えば, 音声処理ソフトウェア「音声工房」を用いて

<http://www.sp4win.com/>



表示された音声データに対しパワーの時間変化を表示する. また, テキストデータを出力する.

2chの音声ファイルを扱える. 長時間の音声ファイルを扱える.

ピッチ, ほかを分析できる.

標準価格 ¥60,900(税込み)
NTTアドバンステクノロジー(株)

SugiSpeechAnalyzer((株)アニモ)は対話データには不向き...テキストデータに出力できない. 長時間の音声ファイルを扱えない. 扱えるのは1チャンネルのみ.

騒音レベルの分析

- 音響測定分析システム DSSF3(吉正電子株式会社: YMEC)を用いて



**DSSF3 Full-system
Version ¥78,000**

<http://www.ymec.com/store/jp/>

この会社のサイトでは音響計測の事例を多数紹介.

また, メールでの問い合わせにも丁寧に答えてくれる.

DSSF3を用いた分析法

- 対話の「間」の検討では、
時定数10msec*で、
10msec *ごとに分析する
(*時間分解能が高い例. 研究目的により変える)

1. リアルタイムアナライザ(RA)
のランニングACF測定

↓ 連携させる

2. サウンドアナライザ(SA)

ランニングACF

積分区間 0.01 sec

計算周期 0.01 sec

聴感補正 A特性

最大遅れ時間 0.05 sec

τe

ピーク検出時間間隔 5 ms

回帰終了レベル -5 dB

回帰終了時間 50 ms

原点を含める

測定条件設定

サンプリング 44.100 kHz

測定時間 800 sec

チャンネル
 モノラル ステレオ

表示設定

積分時間 0.005 sec

計算周期 1 sec

横軸(τ) 5 ms

聴感補正 A特性

時間軸データ数 800

時間軸方向
 前方
 後方

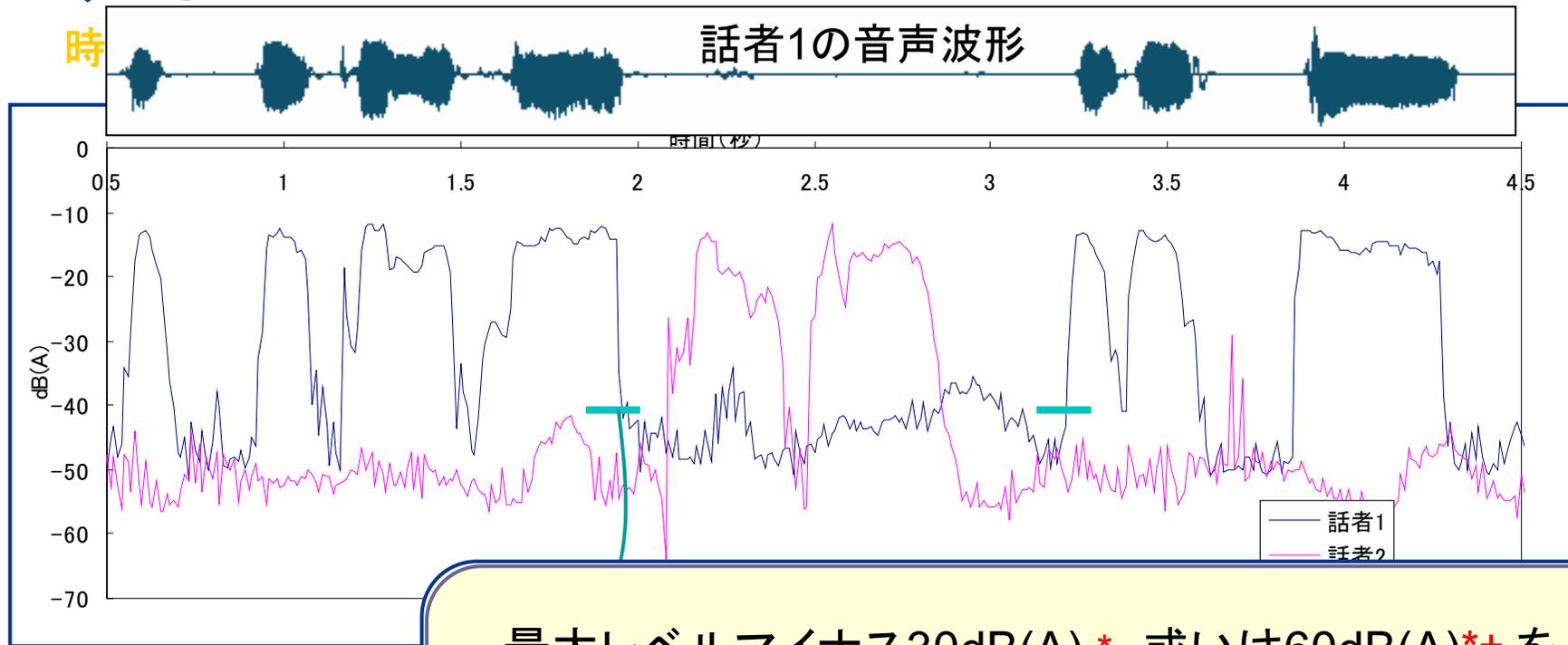
縦軸
 直線 絶対値
 対数 -40 dB

チャンネル
 Lch
 Rch

再表示(R)

○の数値を設定

分析データ中に発話開始と終了をマークする



51	0.49	-36.07	
52	0.5	-44.08	-1.#J
53	0.51	-31.81	
54	0.52	-44.82	16
55	0.53	-39.72	1
56	0.54	-50.78	18
57	0.55	-49.84	-1.#J
58	0.56	-49.24	70
59	0.57	-46.82	9
60	0.58	-50.65	-1.#J
61	0.59	-50.12	10.5
62	0.6	-50.81	-1.#J

最大レベルマイナス30dB(A) *, 或いは60dB(A)*† を
発話ON・OFFの基準としてマークする。

*ヘッドホンで聴取して分析した場合とよくフィットする。

†絶対的な値を分析するためには収録時に基準音の収録と測定が必要。

2話者の音声を別チャンネルに収録するのに有効なマイク

■ ヘッドセットマイク

- AKG: c420



メーカー希望小売価格：¥37,000

二人の声を独立系統でとる
動く話者の声を安定してとる ために

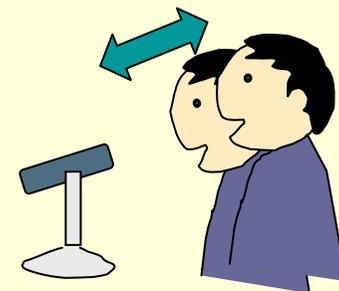
■ ピンマイク

- ソニー: ECM-66B



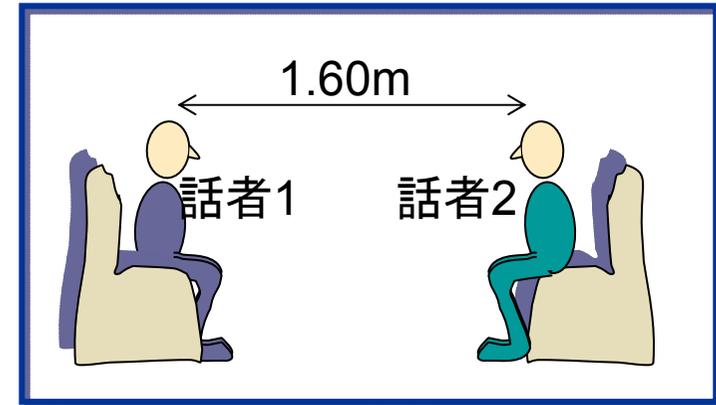
標準価格(税込): ¥63,105

据え置きマイクの場合、口からマイクまでの距離により声のレベルが変化する。

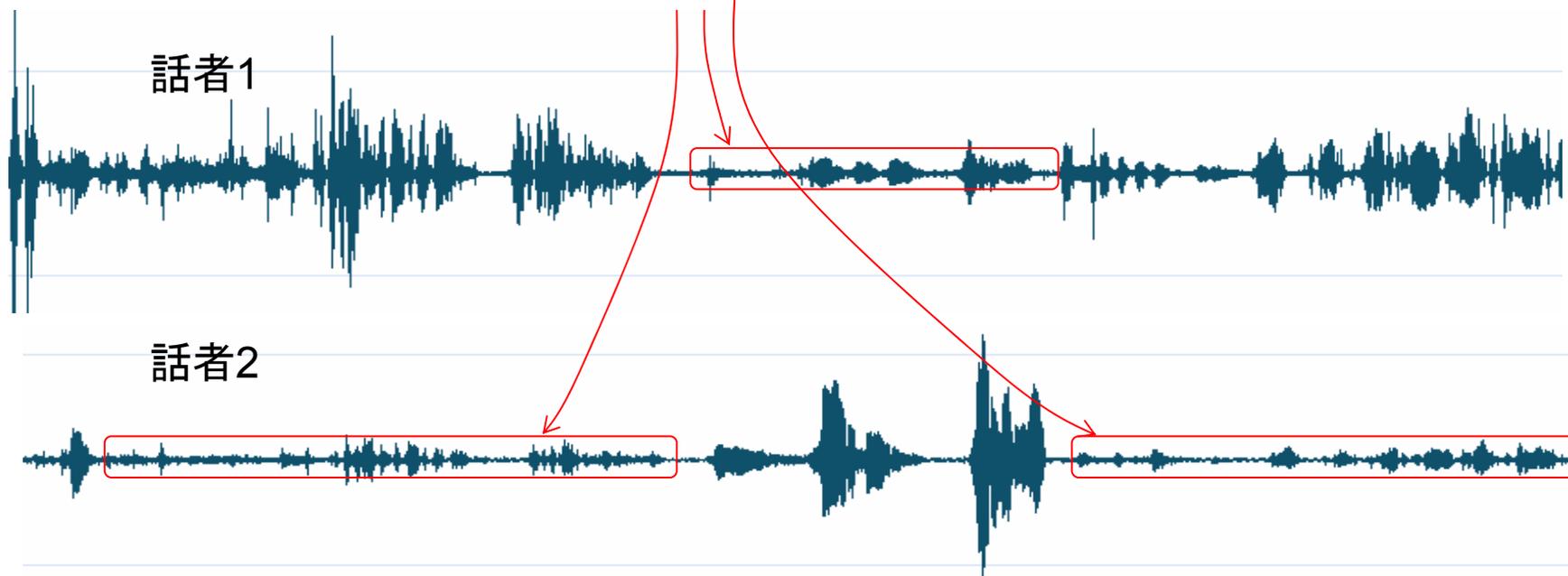


ピンマイクを使った収録音声の例： 同室にいる2者の対面対話

- ピンマイク(ソニー: ECM-66B)を使用



別Chへのうつり込み



防音室の設置

- 防音室の2つのタイプ
 - 研究室に設置するタイプ
 - 建物に作り付けるタイプ

研究室に設置するタイプの防音室

- ミュージックキャビン株式会社
 - サイエンスキャビン
 - <http://www.music-cabin.co.jp/sc/index.html>

Tokuhachi
Corporation Co., Ltd.
**SOUND
PROOF**

サイエンスキャビンは、
セミオーダーメイドの高性能遮音空間です。
あなたの実験研究室内に手軽にコンパクト
に設置することができます。

無響室・残響箱・多目的音響測定室 等

サイエンスキャビンは、おおよその
規格をもとに、設置条件に合わせた
形状・寸法に製作・加工して設置現
場にお届けするセミオーダーメイド
の防音室です。

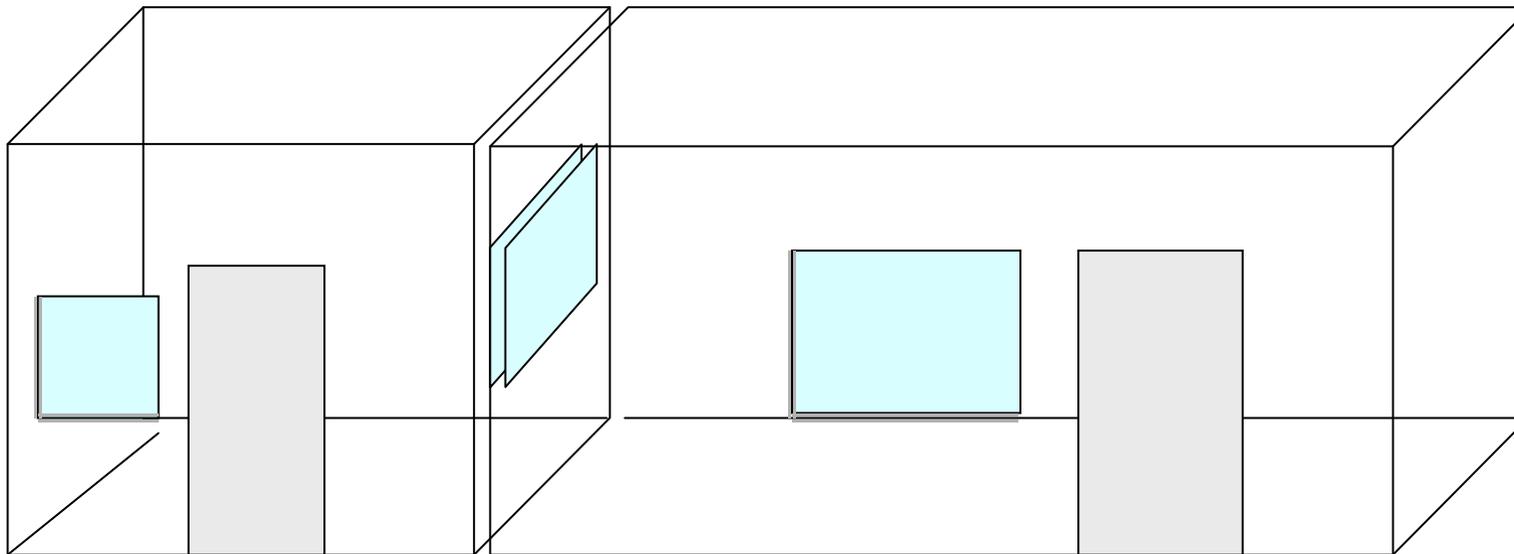


価格:1畳弱で140万～

作り付けの防音室

例：阪大人間科学研究部感性情報心理学講座防音室・視聴覚実験室

- 窓をはさんで2防音室が隣接



阪大中村研防音室の使用例

- 音声収録
- 対話収録
 - 同室で対面
 - 別室で窓をはさんで対面
 - 別室で非対面
 - 部屋を分けると,2話者の音声を完全に分離して収録できる



作り付けの防音室

例：阪大人間科学研究部感性情報心理学講座防音室・視聴覚実験室

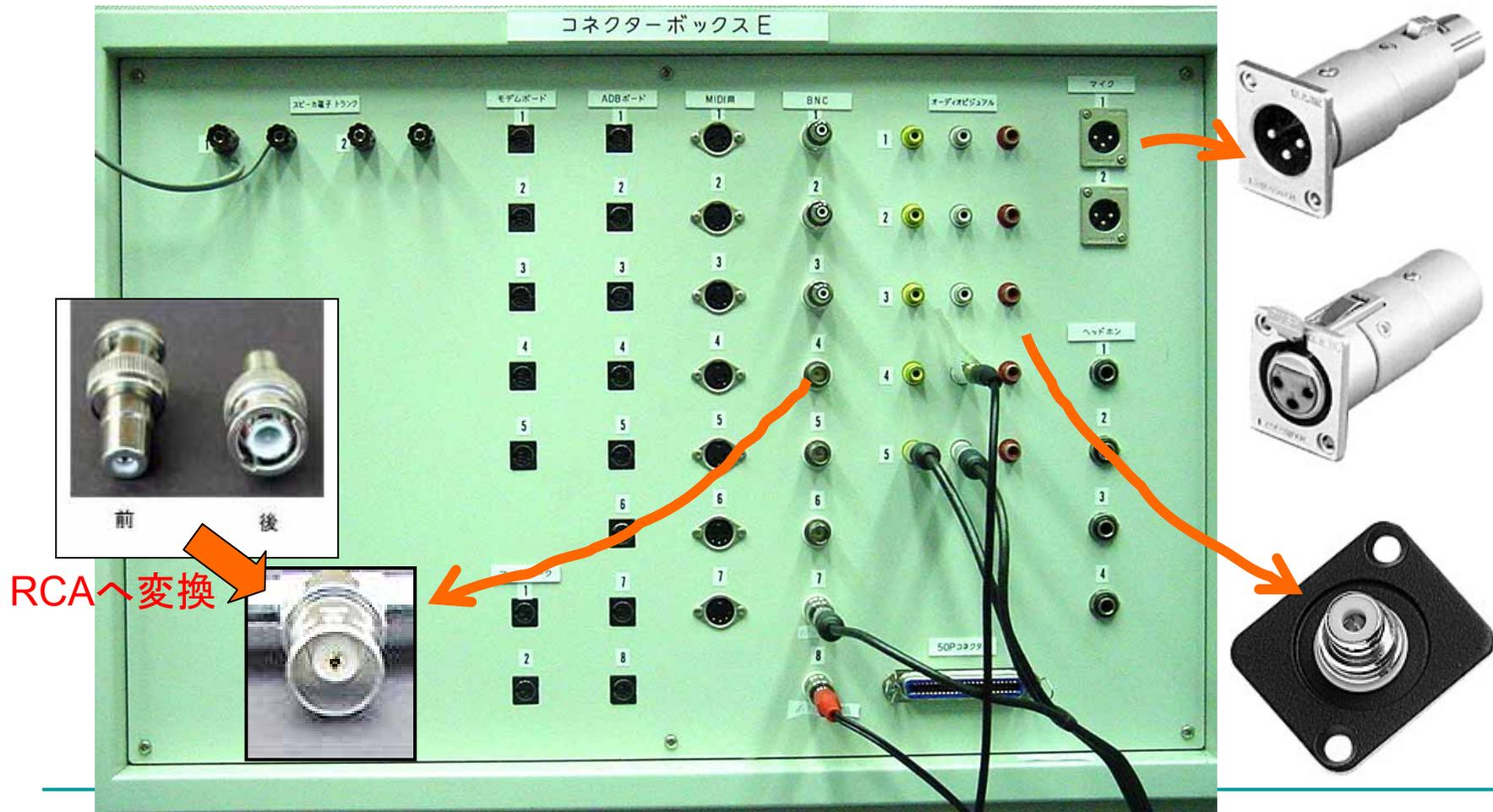
- 窓をはさんで2防音室が隣接
- 居室のような内装
- 防音のための2重扉, 3~4重ガラス窓
- 録音再生機器は防音室外に設置.



オーディオワークステーション

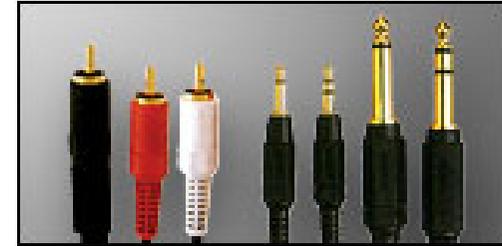
コネクタパネル

- 十分な種類・個数のコネクタが必要.



コネクタ

- XLR端子・キャノン端子＝マイク端子
- RCA(ビジュアル・オーディオL/R)
＝ビデオ端子
- ミニジャック(モノ・ステレオ)
- 標準プラグ(モノ・ステレオ)
- スピーカーケーブル

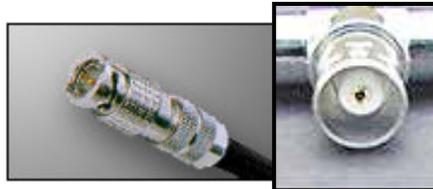


RCA

標準プラグ

ミニジャック

- BNC



- IEEE1394 (6ピン・4ピン)＝i-Link, FireWire



- USB

- S画像



S端子
(Mini Din 4Pin)オス

マイク(&付属ケーブル)のコネクタもマイク選定の際の1つの手がかり

- XLR端子・キャノン端子



SONY



よりノイズがのりにくい
より業務用

- 標準プラグ
- ミニジャック



ファントム電源(接続先の機器からマイクケーブル経由で電源を得る). 多くのオーディオキャプチャ, 業務用のカメラにはファントム電源がある.

まとめ

- 発話開始・終了の同定法
 - 声の強さ・大きさの分析方法の紹介
- 2話者の音声を別チャンネルに、
かつ、動く話者の音声を安定して収録するために
 - 単一指向性ヘッドセットマイクやピンマイク
- 防音室の設置
 - 十分な種類・個数の端子を備えたコネクタパネルを！

補遺：発話速度の算出法

1. モーラ数をカウントする.
 - モーラは拍。「チョコレート」なら5モーラ, 「切手」なら3モーラ.
2. 発話速度を計算する.
 - 平均モーラ長(ミリ秒) = 発話の長さ(ミリ秒) ÷ モーラ数
 - モーラ/秒 = モーラ数 ÷ 発話の長さ(秒)