

# 音声の収録から効率的な転記・音声分析用 データの作成まで

榎本 美香

千葉大学大学院自然科学研究科

# 発表者の背景

- 所属
  - 千葉大学大学院自然科学研究科 情報科学専攻
- 専門
  - 言語心理学
- 研究内容
  - 会話における相互行為分析
    - 話者交替のための言語的・非言語的規則の構築
    - 円滑なコミュニケーションのための会話連鎖分析
    - 聞き手の会話運用原理の解明

# 音声収録の方法

- パソコンに直接録音
  - 後日パソコンに取り込む必要がない
  - 音質が落ちない
- MDに録音→音声をパソコンに取り込む
  - 録音時間に制限がない
  - 収録時に持ち運ぶ機材が少ない
  - 被録音者が動き回ってもかまわない

# 直接パソコンに録音するセット1

マイク延長用コード  
Victor CN-220A  
(1,365円)

AudioTechnica AT9820X  
(8,500円)



ミニプラグ付マイク×2

↑  
モノミニ

SONY PC-202M

(マイクに付属)



↑  
変換プラグ



Roland UA-20  
(20,000円)

USB2.0ケーブル  
(ABタイプ)  
UA-20に付属



パソコン  
wavesurfer(2チャンネルまで録音可)

# 直接パソコンに録音するセット2(高品質・多人数用)

USB2.0ケーブル  
ル(ABタイプ)  
UAに付属



Sennheiser MKE104



(68,000円)

端子  
キヤノン



Roland UA-1000

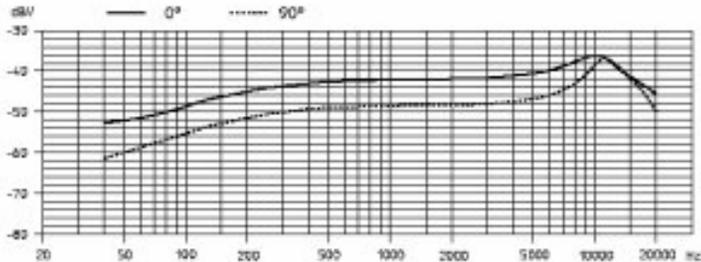
(オープンプライス  
100,000円程度)

パソコン  
DigionSound5(10  
チャンネルまで録  
音可)

(スタンダード45,000円)



Frequency response ME 104



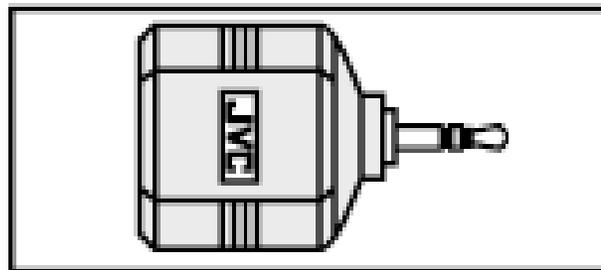
# MDに録音するセット

AudioTechnica AT9820X  
(8,500円)



↑  
モノ  
ミニ

ミニジャック×2 —— ステレオミニプラグ



Victor AP-212A(1,000円)

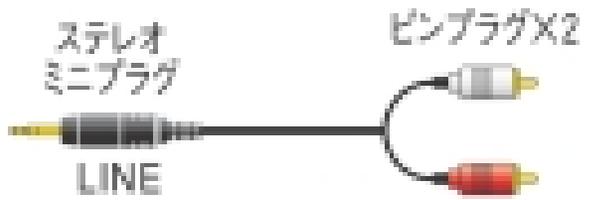
ポータブル  
MDレコーダー



[マイク端子へ]

# MD・DATからパソコンへ録音するセット

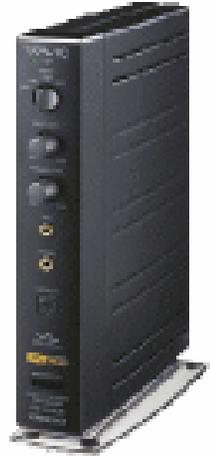
Roland UA-1X  
(9,000円)



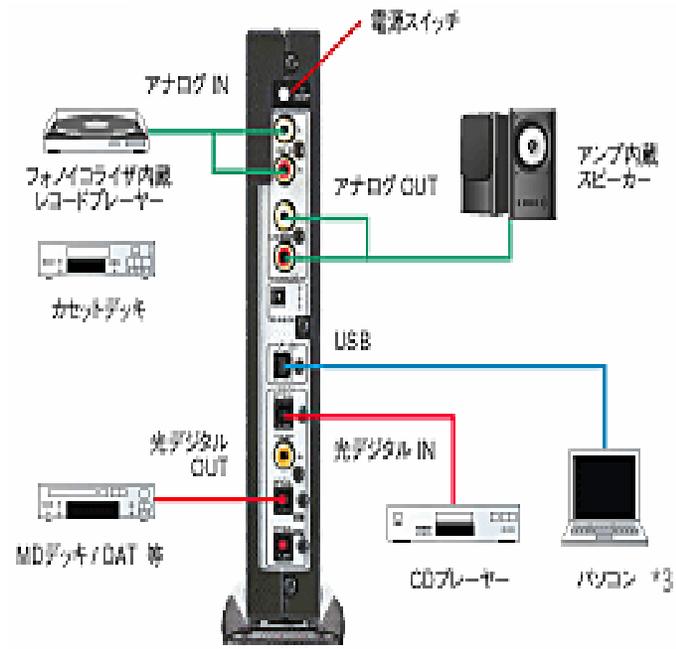
Panasonic RP-CAPM3G15 (1,000円)



光デジタルケーブル

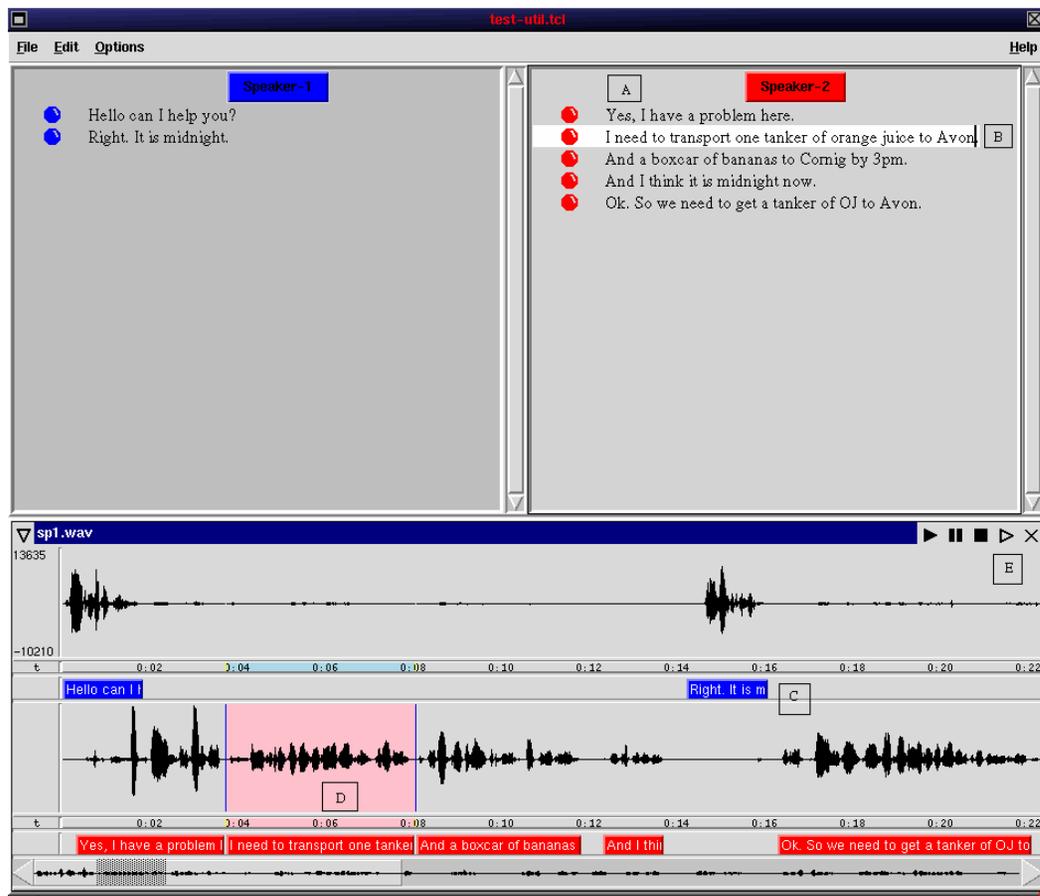


ONKYO SE-U55GX(B)  
(22,000円)



# MultiTransを使った転記の作成

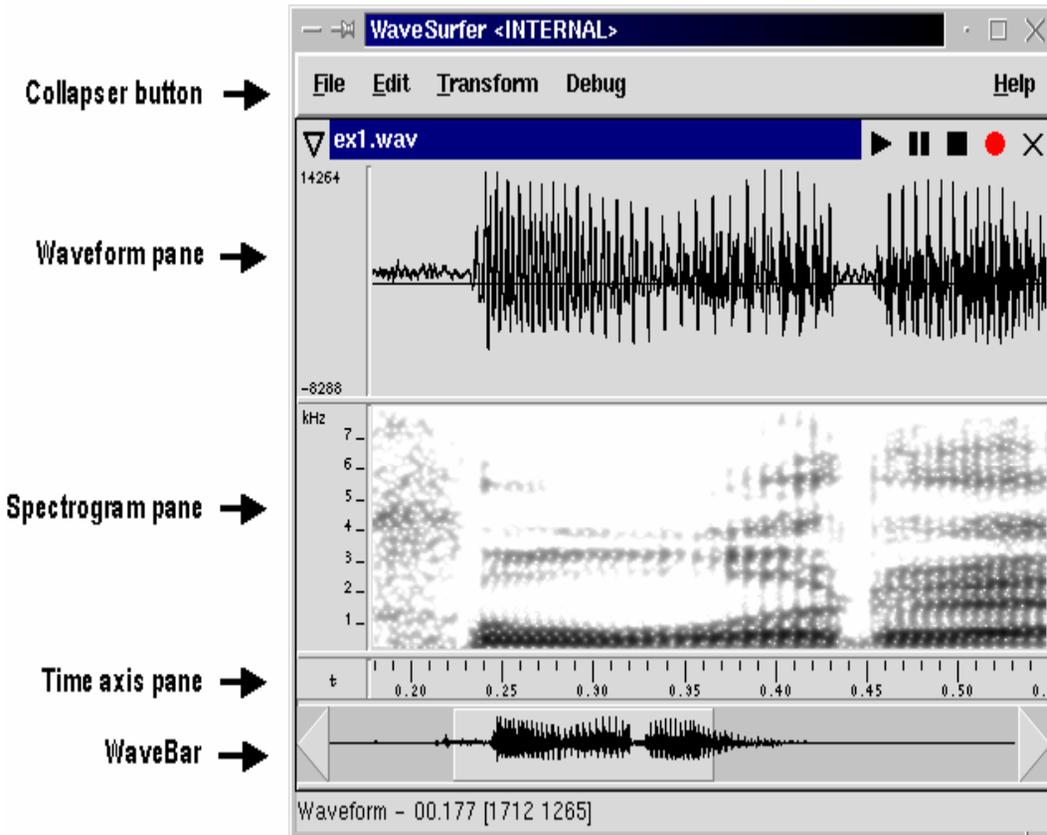
[http://sourceforge.net/project/showfiles.php?group\\_id=25216&package\\_id=20111](http://sourceforge.net/project/showfiles.php?group_id=25216&package_id=20111)



- 音声を聞きながら、発話内容を転記するためのソフト
- 日本語を扱うためには、<http://cogsci.l.chiba-u.ac.jp/~den/Tools/MultiTrans/MultiTrans.html> より更新ファイル入手する必要がある。
- 日本語文字コードにutf-8を使用しているため、他のソフト(Microsoft Word 等)で加筆・修正をした場合、Charset Converter 等のソフトを用いてコード変換する必要がある。

# WaveSurferを使った音声分析

<http://www.speech.kth.se/wavesurfer/download.html>



- スペクトログラム・ピッチ曲線を描画・保存したり、それらを参照して分節音や音調のラベリングをするためのソフト
- MultiTransで作成したLCF形式のラベルファイルを扱うためには  
<http://cogsci.l.chiba-u.ac.jp/~den/Tools/WaveSurfer/WaveSurfer.html>より設定ファイルを入手する必要がある。
- 日本語を読み込む/書き込む場合は、Property のLabel file encoding でutf-8を指定する(utf-8と手で書き込む)必要がある。

# ラベル間の関連付け・映像のラベリング・評価には Anvil

<http://www.dfki.de/~kipp/anvil/description.html>

The screenshot displays the Anvil 3.6 software interface, which is used for video analysis and gesture recognition. The main window shows a video frame of a man in a suit, identified as 'ZDF'. The interface is divided into several panels:

- File Edit View Tools Bookmarks ?**: The top menu bar.
- Video: lq1-7-reich.avi**: The main video player showing the man's gesture.
- Track: gesture.phrase**: A panel on the right showing track details for 'gesture.phrase', including 'Referenced track: gesture.phase' and 'Time: 03:07:63 - 03:10:27 (66 frames)'. It also lists attributes: 'category: iconic', 'iconic type: smash', 'handedness: 2H', 'cocc: Rivas', 'function: emblematic', and 'timing: direct'. A comment box contains 'compare with lq1-8 at 0:28'.
- Annotation: lq1-7-reich.avi**: A bottom panel showing a timeline with various tracks: 'wave', 'praat', 'trl', 'trl2', 'ling', 'posture' (pose, shift), 'phase', 'phrase', and 'gesture' (compound). The 'phrase' track shows 'metaphoric, heart, 2H' and 'iconic smash, 2H'. The 'gesture' track shows 'emblem, so-what, 2H'.

The video player shows a man in a suit and glasses, identified as 'ZDF', with his arms raised in a gesture. The timeline below the video shows various tracks, including 'wave', 'praat', 'trl', 'trl2', 'ling', 'posture', 'phase', 'phrase', and 'gesture'. The 'phrase' track shows 'metaphoric, heart, 2H' and 'iconic smash, 2H'. The 'gesture' track shows 'emblem, so-what, 2H'.

# その他の音声分析ツール

- 無料ツール
  - SpeechAnalyzer  
<http://www.sil.org/computing/catalog/index.asp>
  - Praat  
<http://www.fon.hum.uva.nl/praat/>
- 有料ツール
  - Acoustic Core (198,000円)  
株式会社アルカディア <http://www.arcadia.co.jp/index.html>
  - 音声録聞見 for Windows(18,000円)  
デイテル株式会社 <http://www.cd4power.jp/onsei/home.htm>
  - SUGI SpeechAnalyzer(94,800円)  
株式会社アニモ <http://www.animo.co.jp/>