

音声分析実験テキスト

(最終改訂 2016/03/18)

実験の目的

本課題は、音声信号をコンピュータに取り込み、観察、計測、分析する実験である。音声は、日常生活において欠くことのできないコミュニケーション手段であるが、音声を直接目で見ることはできなかった。しかしながら、近年のコンピュータ、デジタル信号処理技術の進歩によって、音声信号をコンピュータに取り込み、処理する（例えば音声認識）ことが容易になった。音声は、人の調音器官（参考文献1）によって生成され、空気（媒質）を通して、人の耳に到達する。この音声信号をコンピュータで取り込んだ場合、デジタル信号として、便宜上波形で表現される。発声している音声が、いかに変化にとんだものかを、まずはじっくり観察して欲しい（参考文献2）。また母音などでは、同じような波形が繰り返されていることに注目して欲しい。すでに実験（情報学実験Ⅱ②オシロスコープと信号処理）を行ったように、時間領域の波形情報は、周波数領域の情報（周波数スペクトル）に変換できる。音声の情報を周波数領域で観察し、音の種類（母音・子音）によって、周波数スペクトルの形状が異なることに着目して欲しい。

参考文献1 大島義和、日本語を考える 音声・音韻(1)

http://www.gsid.nagoya-u.ac.jp/oshima/teaching-old/IU_2009-2010/IU-docs/IUJL_03.pdf

参考文献2 峯松信明、音声の音響分析の「いろは」 ～初めて音声波形を見る方へ～

<http://www.gavo.t.u-tokyo.ac.jp/~mine/japanese/acoustics/lecture-02.pdf>

注意

本テキストでは実験を行う上で必要な項目について記述しており、音声分析の詳細については音声関係の教科書に譲る。本課題ではPCのサウンド機能を使用するため、情報学実験Ⅰ課題②（以下「**課題Ⅰ②**」と言い、他の課題についても同様とする。）、課題Ⅰ⑦と同様に、**Local PC (ローカルPC) でログイン**する。ただし、再生での**既定の形式を変更する必要は無い**（備考 p.20「スピーカーのプロパティについて」参照）。共用ヘッドセットの使用に抵抗のある学生は、各自で用意するかまたは備考 p.25「共用ヘッドセット以外での收音について」を参照の上、分析用音声ファイルを事前に準備しておくこと。

実験ノート以外に用意するもの

方眼紙：母音五角形を描く

USBメモリ：音声ファイル、フォルマント・基本周波数データの保存用

事前準備

必ず、本テキストに最後まで目を通して、全体の作業の流れを把握しておく。
wavesurfer-1.8.8p4-win-i386.zip をダウンロードして（既にダウンロードして実験課題で使用した WaveGene と WaveSpectra を本課題でも使用するのでこれも併せて）**USBメモリ**に保存しておくこと。

使用機器・部品類

ヘッドセット：サンワサプライ MM-HS515SV

仕様 ステレオヘッドホン（ダイナミック型、ボリュームコントローラー付）＋モノラル

マイクロホン（エレクトレットコンデンサ型プラグイン給電方式）詳細はメーカー製品ページ参照

<http://www.sanwa.co.jp/product/syohin.asp?code=MM-HS515SV>

情報科学研究教育センターPC：[Local PC](#)でログインすることに注意。

註：音響の学術用語では phon[e] はホンと表記し、本テキストもこれに従う。このヘッドセットメーカーのページでも忠実に表記している。

使用ソフトウェア

WaveSurfer

開発元ページ <http://www.speech.kth.se/wavesurfer/>

ダウンロード先 <http://sourceforge.net/projects/wavesurfer/files/latest/download>

マニュアル <http://www.speech.kth.se/wavesurfer/man.html>

参考ガイド <http://www.linguistics.ucla.edu/people/hayes/103/WavesurferHints/>

<http://www.f.waseda.jp/kikuchi/tips/wavesurfer.html>

WaveGene

ダウンロード先 <http://efu.jp.net/soft/wg/wg.html>

WaveSpectra

ダウンロード先 <http://efu.jp.net/soft/ws/ws.html>

表紙交付基準

以下の事項の全ての完成を確認して表紙を交付する。レポート作成については、実験手順中の記述の他、本テキスト末尾の備考 p.19「レポート作成について」を参照のこと。

- 6節④ (p. 19) の分析対象音声ファイルが適正レベルで収録されたことを、PC上でWaveSurferで音声ファイルを開きペーンの配置、レベルメータの振れを確認する
- 6節⑤⑥ (p. 19) の分析対象音声ファイルが適正レベルで収録されたことを、保存した音声ファイルのファイル名 (拡張子も明記のこと)、ファイルサイズ (プロパティでバイト数を確認して正確に記録すること)、Waveform ペーン 縦軸上下限の値を実験ノートで確認する。
- 6節④ (p. 19) の分析対象音声ファイルに基づき、5母音のそれぞれについて、基本周波数と第1～第4フォルマント周波数を計測できたことを実験ノートで確認する。
- 前項の結果に基づき作成した母音五角形の図を方眼紙で (等間隔、片対数、両対数の何れの目盛でもよいが、単に実験ノートに概形を描くのではなく、必ず方眼紙を使用して正確にプロットすること) 確認する。

スペクトルの時間変化について

課題 I ⑦前半の正弦波合成編では周期波形が正弦波の重ね合せで得られることを、WaveGeneとWaveSpectraを用いて実際に三角波と矩形波を合成して波形とスペクトルにより確認し、課題 II ②では、信号の表現で表裏の関係にある時間領域の波形と周波数領域のスペクトルとの関係について理解を深めた。スペクトルが無限の過去から未来まで時間的に変化しない完全な周期波形では、予め定めた以外の情報を追加して伝達することはできず、人間の重要な情報伝達手段である音声も当然時間的に変化する。このような信号の特徴を把握するには、有限な分析区間でスペクトルを求めてその時間変化を見ることが有効である。言うまでもなく時間分解能の追求と周波数分解能の追及は相反し両立しない。

これまで WaveGene、WaveSpectra と PC のサウンドカードの機能を用いた実験課題では、20kHz まで忠実に再現するために標本化周波数を CD 音質の 44100Hz としたが、本課題では音声の特徴量を効率よく高い精度で処理するため**標本化周波数 16000Hz**（上限周波数は 8kHz と低くなるが、同じ FFT 分析サンプル数での周波数分解能はほぼ 3 倍になる）の**モノラル**で収録、分析をする。16000 と 48000 とは 1:3 の単純な整数比にあり Windows7 オーディオエンジンの標本化周波数変換アルゴリズムでも十分な精度で変換できるため、本課題では、サウンドカードの再生プロパティで**既定の形式を変更する必要は無い**（備考 p.20 「スピーカーのプロパティについて」参照）。ただし、Virtual Desktop のサウンド機能は上限 8kHz であっても不十分なため **Local PC でログインする**ことは変らない（備考 p.21 「OS の違いによる PC サウンド機能の比較」参照）。

図 1 は WaveSpectra のスペクトル画面を 3D 表示（横軸が周波数でここではリニアスケールに設定、縦軸がパワーでここでは 80dB レンジに設定、奥行方向が時間で手前が現在）に設定して音声スペクトルの時間変化をレリーフマップで表示した例である。左は、民放連リファレンス音源のナレーション、右は、課題Ⅱ⑤の VOCALOID による大学歌で、それぞれ標本化周波数 16000Hz のモノラルに変換した音声ファイルを再生している。

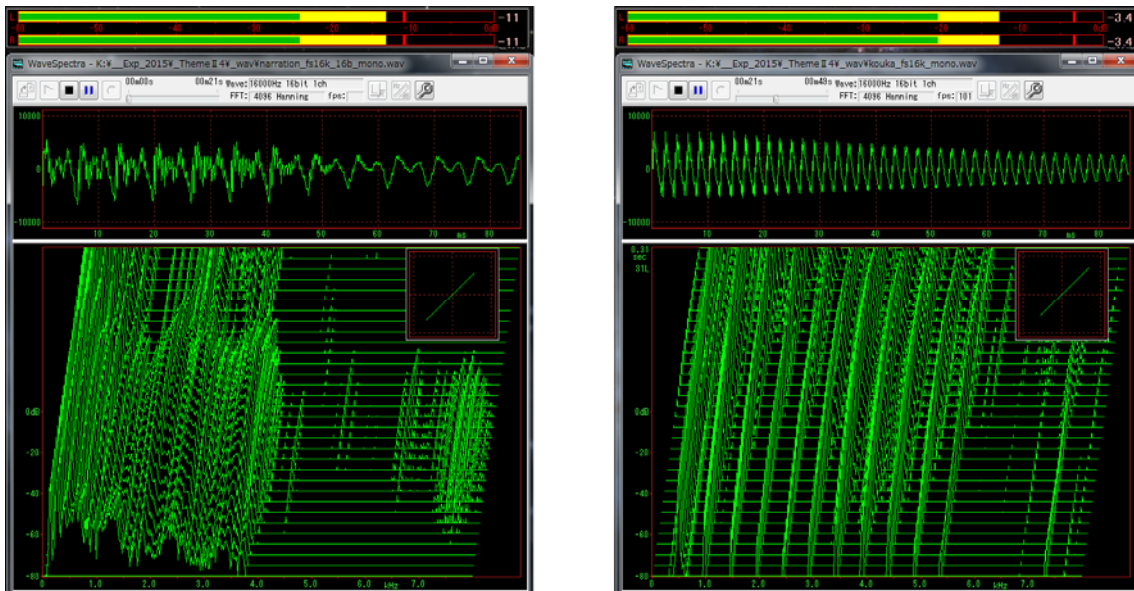


図 1 ナレーション（左）と合成音声の歌唱（右）でのスペクトルの時間変化

フォルマントと基本周波数について

図 1 左で、母音区間では**各母音に対してそれぞれ特定の周波数**にスペクトルのパワーの強い部分があり、これをフォルマント（formant：ホルマントとも表記する）と呼び、母音の特徴となっている。周波数の低い方から第 1～第 4 フォルマント（記号では F1～F4）と称するが、第 1 と第 2 は特に重要で本課題ではこれを両軸に目盛って母音五角形を描く。

p.4 図 2 左は WaveSpectra の画面を左に 90° 回転して縦軸を周波数、横軸を時間としたもので、図 1 左で再生しているファイルの先頭から 641ms までのスペクトルの時間変化を示している。図 2 右は同じ区間の WaveSurfer の Formant Plot ペーン（p.16 後述）を WaveSpectra の画面に周波数スケールを揃えて対比したものである。3D のレリーフマップではなく、パワーの強弱を濃淡（濃い方が強い）で表し、F1～F4 の時間変化をそれぞれ赤、緑、青、黄のカーブ（レリーフマップの稜線に相当する。無音区間の値は便宜上 500Hz、1500Hz、2500Hz、3500Hz としている：p.19 参照）で示している。

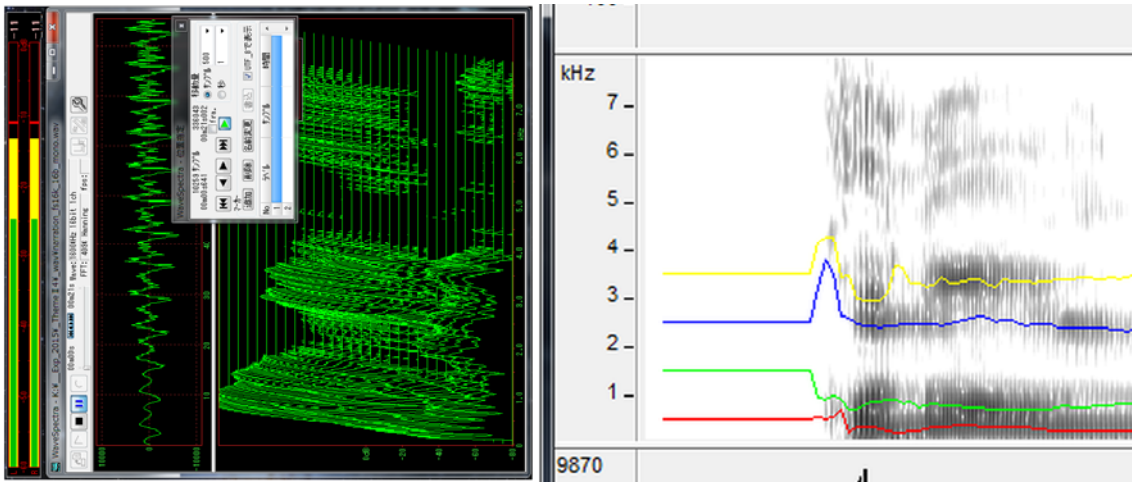


図2 民放連リファレンス音源ナレーションの先頭から 641ms のスペクトル時間変化

周波数リニア 8kHz レンジの図 1、2 では、はっきり分らないが、F1 よりも低い周波数にもスペクトルのパワーの強い部分がある。これは母音の特徴ではなく声の抑揚に対応するもので、基本周波数（記号は F0）と呼ばれる。図 3 左は、WaveSpectra の設定で周波数レンジを 400Hz に変更し、p.2 図 1 左のファイルの先頭から 724ms までを再生した低域部分のスペクトルの時間変化を示している。図 3 右は、同じ区間の WaveSurfer の Pitch Contour ペーン（p.16 後述。註）を図 2 と同様に周波数スケールを揃え（そのままでは幅が大きくなるので）横方向を 1/3 に圧縮して対比したものである。

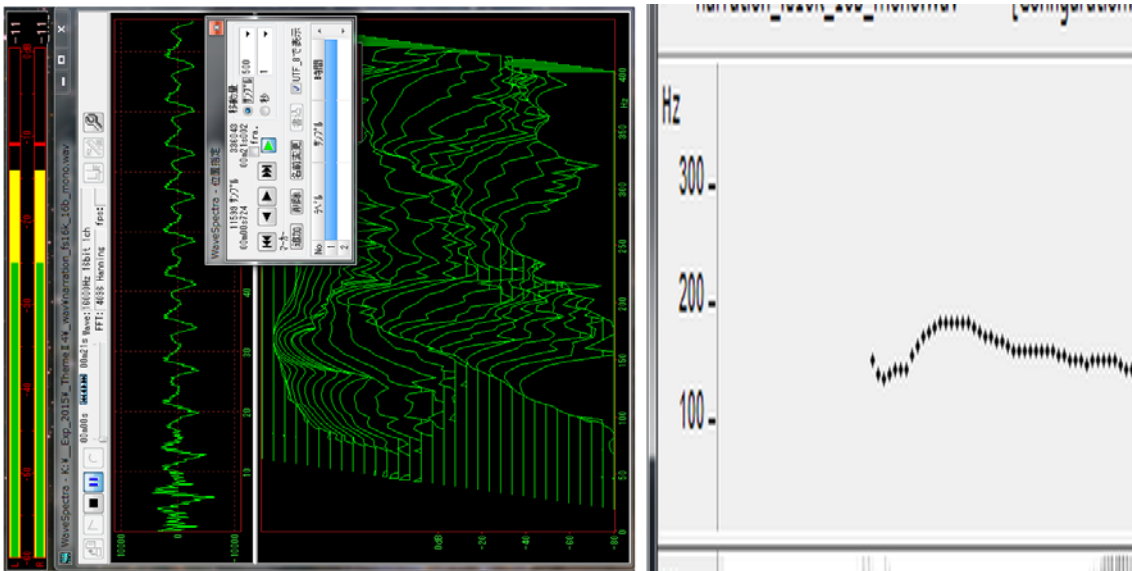


図3 図2と同音源の先頭から 724ms の低域部分のスペクトル時間変化

註：WaveSurfer の **Pitch Contour** は通常の会話音声として発声される母音の基本周波数を分析しており、歌唱のデータを分析しても音楽編集制作ソフト（DTM ソフト）の機能にある「ピッチ検出」による旋律情報は得られないことに注意する。興味のある人は課題 II ⑤ で作成した歌唱のファイルを WaveSurfer で処理してみることを推奨する（標準化周波数の

変換、モノラル化については、備考 p.25「共用ヘッドセット以外での收音について」参照。

図4、5は p.3 図1 で使用した音声ファイルを再生した WaveSurfer の分析画面の例（操作については、p.15 実験手順5 節を参照）で p.4 図2 右の Formant Plot、p.4 図3 右の Pitch Contour の両ペーンに加えて波形表示の Waveform のペーンが表示されている。

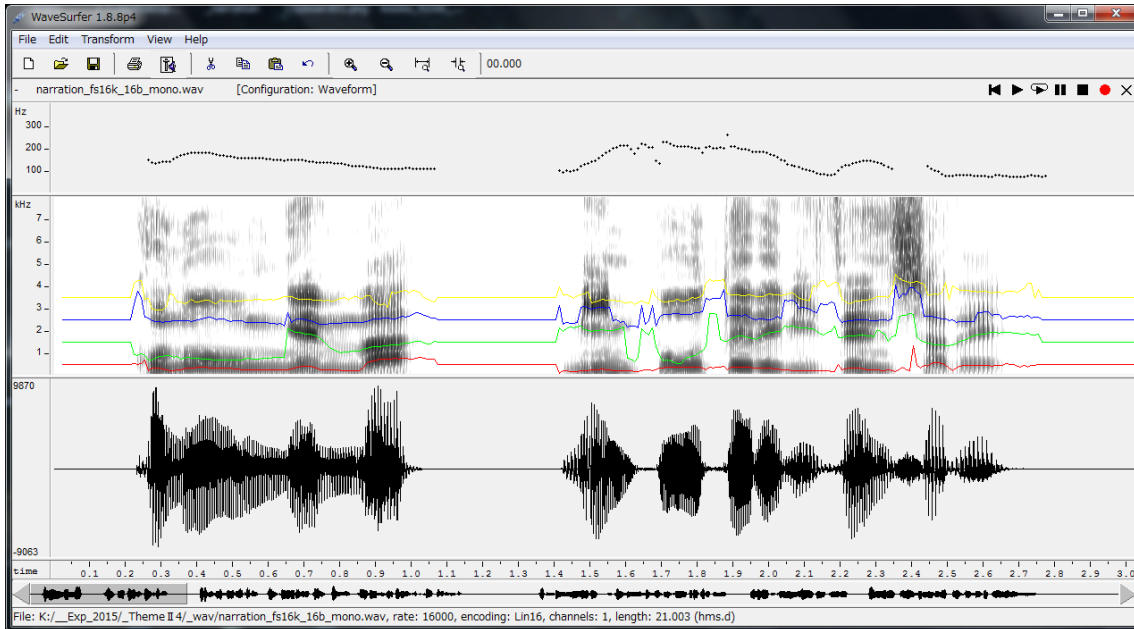


図4 男声のナレーション 「この音源は、民放テレビ全社で…」

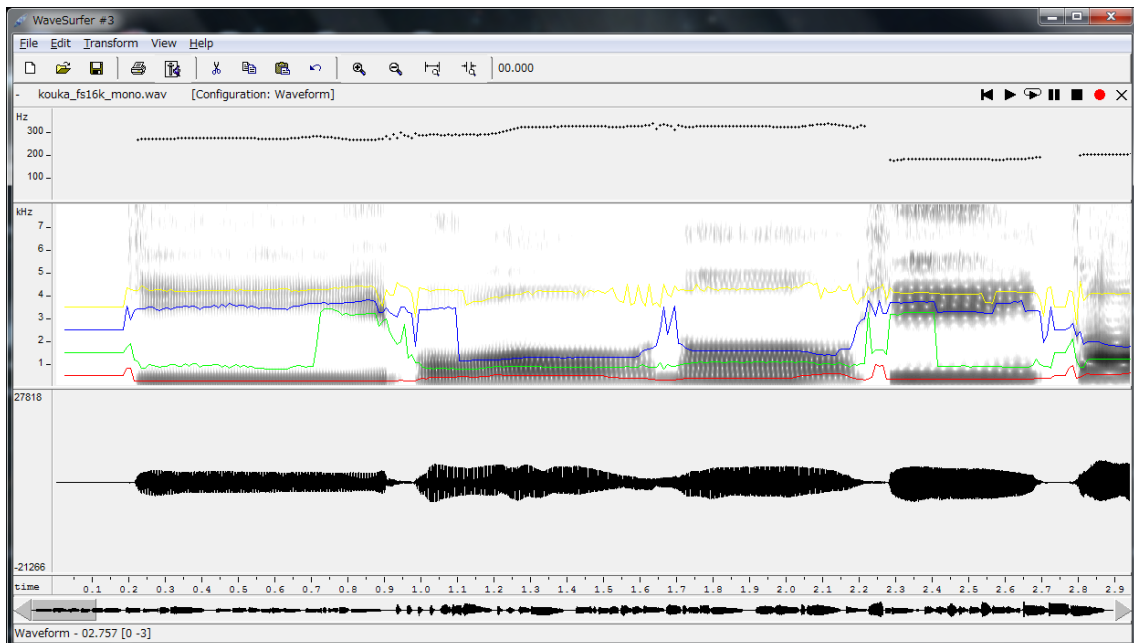


図5 音声合成による歌唱 「きぼうのひか…」

フォルマント周波数の値は声道の形状に関係し、舌の位置、唇の形で決る共振（共鳴）周波数に対応する。舌の位置と発声される母音の関係を示す母音図（vowel chart）では左

その他の参考 URL

日本音響学会 Q and A (050) : Q: 男性の声, 女性の声, 子供の声は, 聞けば区別がつかうのですが, 具体的にはどこが異なっているのでしょうか?

<http://www.asj.gr.jp/qanda/answer/50.html>

日本音響学会 Q and A (101) : Q: ピッチと基本周波数はどう違うのですか。

<http://www.asj.gr.jp/qanda/answer/101.html>

図 8 は 7 歳から成年までの男女による日本語 5 母音の第 1、第 2 フォルマント周波数の分布を示す。

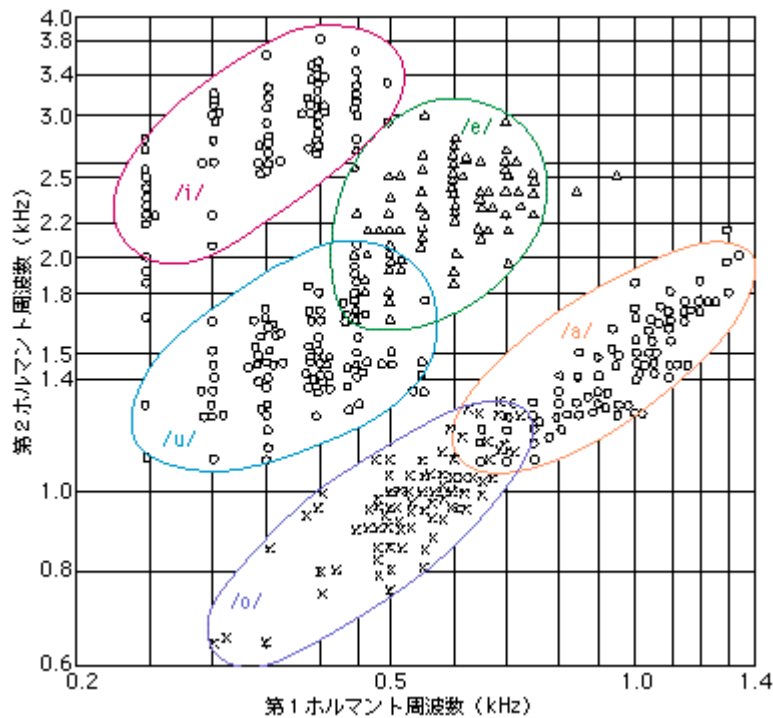


図 8 日本語母音のフォルマント 【国立天文台編 理科年表 2006】

録音レベルと雑音について

録音レベルが適正でない場合、意味のある分析結果が得られないので特に注意を要し、実験手順の 1 節 (p. 9) では WaveSurfer による收音に先立ち WaveGene と WaveSpectra を用いて再生レベルと録音レベルの調節を行うこととしている。p. 8 図 9 は p.3 図 1 左で再生しているナレーションのファイル (左) とこれを 60.206dB 減衰させたファイル (16 ビット深度の内、10 ビット分の情報が失われたことになる: 右) の「この音源は」の発声部分を WaveSurfer で分析した例で Waveform の上に Power Plot ペーンを表示している。

右の図では、フォルマントの表示で濃淡がはっきり分れず全体がノイズに埋もれている様子が、またプロットされたカーブからはフォルマント周波数を正しく検出できていないことが分る。波形表示の振幅レンジは自動的に調整されるため、録音レベルが低過ぎる場合も普通に録音できているものと誤解しやすいが図右ではパワーがレンジ外であることを示している。Waveform ペーン左端の波形の縦軸の値が 16 ビット量子化に相応しい値であることを確認すること。【この値は表紙交付基準 (p.2) の確認事項に含まれている】

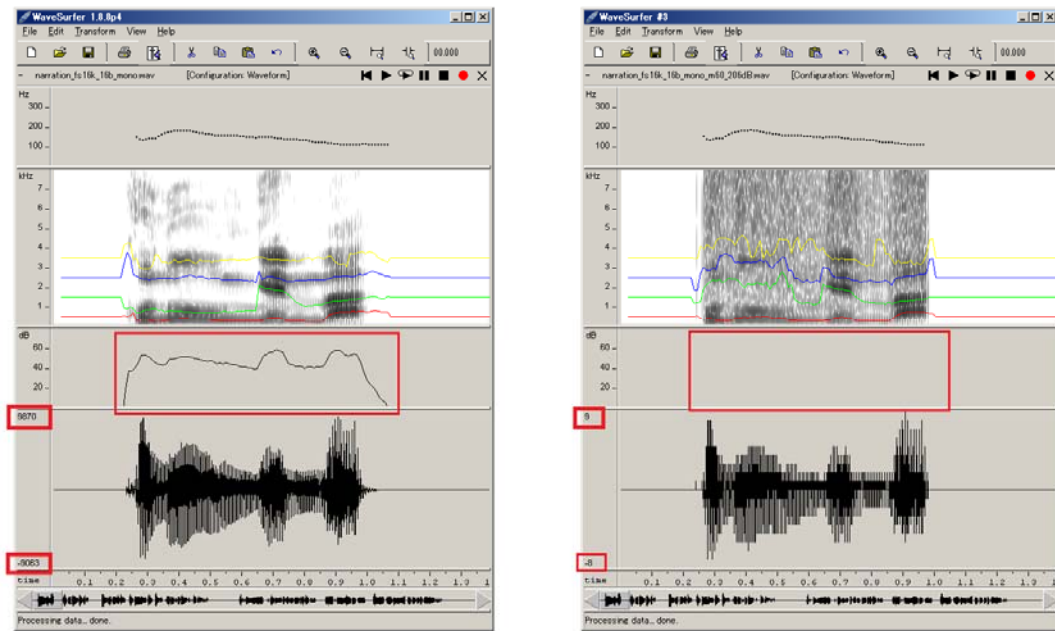


図9 適正な録音レベル（左）と低すぎる録音レベル（右）での分析結果の違い

本実験の分析音声の録音作業では、他人の発声もヘッドセットのマイクロホンを通して録音され、これが自分の音声データに対する雑音となる。図10は雑音が発声部分に及ぼす影響を示す例で、民放連リファレンス音源のBGM入りナレーション（左）、女声ナレーションを発声の先頭を揃えてS/N 10dBで加えたもの（中）、白色雑音をS/N 10dBで加えたもの（右）の図9と同じ「この音源は」の発声部分である。作業中は声高な私語を慎むなど他の実験者への雑音とならない様に心がけていただきたい。

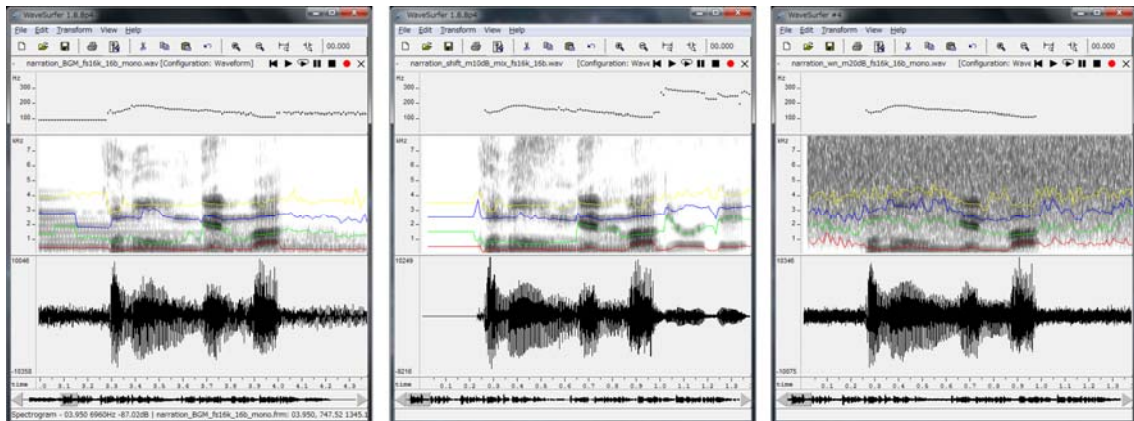


図10 雑音の影響 左からBGM、女声ナレーション、白色雑音

実験手順

録音・分析の対象音声については、本手順最後の p. 19 第6節に掲げる。必ず、一度、最後まで目を通して、本課題で必要なデータの意味を理解し、全体の作業の流れを把握した上で、本手順に従うこと。

0. USBメモリとヘッドセットをPCに接続する。

第1演習室PCに **Local PC (ローカルPC) でログイン**し (Virtual Desktop ではない)、**wavesurfer-1.8.8p4-win-i386.zip** を保存した USBメモリとヘッドセットを接続する。ヘッドセットのプラグにはそれぞれのシンボルが記入されているが、暖色のピンクが電力供給デバイス (hot) のマイクロホン、グリーンが電力消費デバイス (cold) のヘッドホンである。ヘッドホンとマイクロホンを接触させた状態 (図 11 左) で、ヘッドホンから WaveGene で再生した白色雑音をマイクロホンで收音して WaveSpectra で表示した両者の総合特性 (図 11 右) にはピークとディップが見られるが、この形態で測定したときの特性としてはごく普通のものである (ヘッドホンとマイクロホンの特性を見るため、次節の作業と異なり WaveGene/WaveSpectra の標本化周波数は 44100Hz に設定している)。

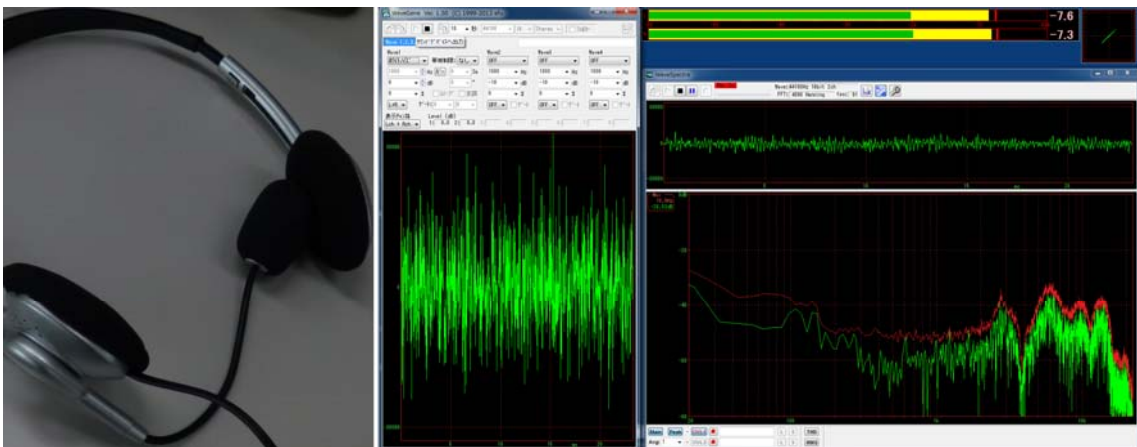


図 11 WaveGene で再生した白色雑音を WaveSpectra の録音モードで表示

1. WaveGene と WaveSpectra を用いて、レベル調節を行う。

「録音レベルと雑音について」(p. 7) で、正しい分析のためには録音レベルを適正にする必要があることを述べた。WaveSurfer は音声ファイルの録音再生を行う一般のアプリケーションとは異なり録音時にレベルメータが表示されない仕様である。このため收音時のレベル調節作業は試行錯誤による効率の悪いものとなり、ここでは、**WaveSurfer を起動する前に標準的な入力に対する適正レベルを設定しておくこと**とする。使用する WaveGene および WaveSpectra については課題 I ②、課題 I ⑦のテキストおよびソフトウェアのヘルプを参照のこと。本節の作業では、WaveGene/WaveSpectra の標本化周波数は WaveSurfer で分析を行う標本化周波数である 16000Hz に設定する。

① 再生レベルの調整

ヘッドセットの**ボリュームコントローラーのつまみを最小の位置**に設定して、**WaveGene** (実行ファイル: **WG.EXE**) を開く。標本化周波数を既定の 44100 から 16000 (プルダウンメニューのリストに含まれないので直接指定する) に変更するため、「サンプリング周波数」表示窓をクリックする。BackSpace でも Delete でも表示窓は空白ではなく 1000 となるので(この標本化周波数では再生できない既定の周波数欄の 1000 が赤く表示される:p.10 図 12 左) 右向きカーソルキーでカーソルを 1 と 0 の間に置いて 6 を挿入する。

再生ボタン (▶) をクリックする。既定では**-10dB** の 1kHz 正弦波が再生される (図 12 中) が「おトノイズ」を再生してもよい。ボリュームコントローラーの音量つまみを少しずつ上げて、最大の位置で適切な音量で聞こえる様に、コントロールパネル→個人設定→サウンド→再生の頁でスピーカーのプロパティを表示し、レベルを調節する (図 12 右)。

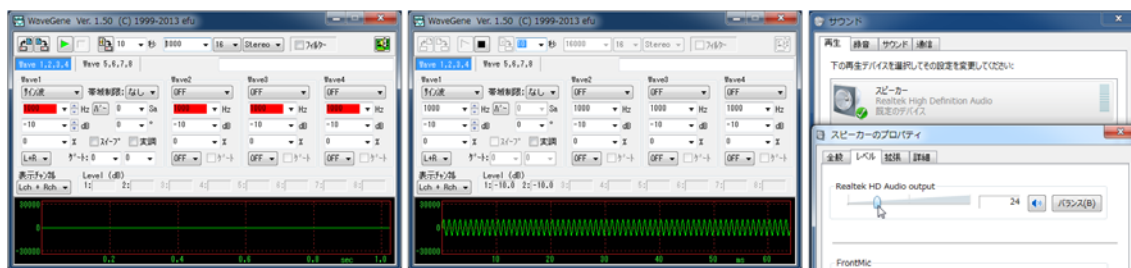


図 12 WaveGene の標本化周波数を 16kHz に変更して**-10dB** 1kHz 正弦波を再生する。

② 録音レベルの調整


WaveSpectra (実行ファイル: WS.EXE) を開き、 をクリックして設定画面の再生/録音頁で「フォーマット」を 16000 にする (図 13 左)。Wave 頁でレベルメーター (既定は「表示しない」) の指定を「表示(横)」に変更し、設定をクリックする。メーターの表示場所、サイズを調整後、録音ボタン (●) をクリックする (既定では入力音声のスペクトル表示のみでファイルには保存されない)。マイクロホンを口元に近付けた状態で発声し、メーターのピーク値が**-10dB**程度 (図 13 中) になる様にコントロールパネル→個人設定→サウンド→録音の頁でマイクのプロパティを表示し、レベルを調節する (図 13 右)。なお、**PC の入力端子に何も接続していない状態では「マイク」は認識されない**ため、プロパティを表示できない他、WaveSpectra を録音状態にしたときにエラーとなる。この様なときは、WaveSpectra を一度閉じて、マイクロホンを接続後に開き直す。



図 13 WaveGene の標本化周波数を 16kHz に変更して、録音状態でレベル調節する。

2. WaveSurfer を起動し、録音モードの設定をする。

はじめに、WaveSurfer の**録音パラメータを正しく設定**する。音声分析では音声の特徴を的確に分析する**標準的な形式として標本化周波数 16000Hz、モノラル**が使われる。

wavesurfer-1.8.8p4-win-i386.zip を開くと、C:\¥work¥wavesurfer-1.8.8p4-win-i386 に展開されるので各自のドキュメントライブラリにコピーしておくこと。作業中は、実行フ

ファイル wavesurfer.exe のショートカットを作り、デスクトップに置いておくとよい。

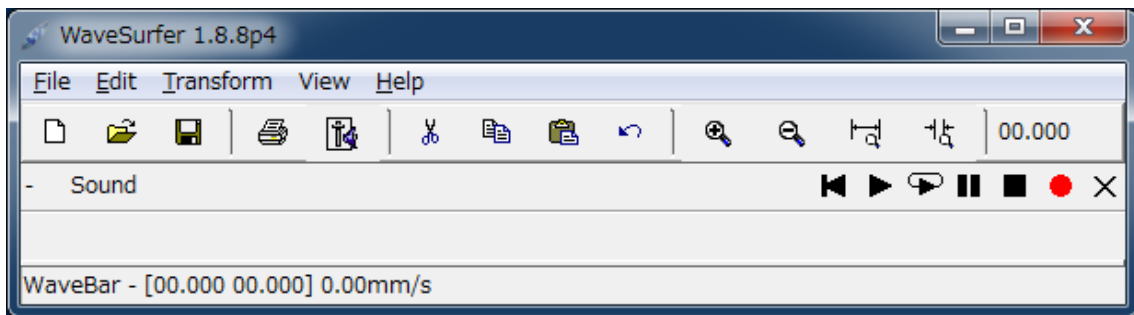


図 14 WaveSurfer を最初に起動した状態

wavesurfer.exe を開き (図 14)、**File**→**Preferences...**を選択し、**Sound I/O** 頁で分析用録音パラメータを、New sound default rate:を既定の 48000 から 16000 (標準化周波数 16000Hz)、New sound default channels:を既定の 2 から 1 (モノラル) にそれぞれ変更設定する (図 15)。

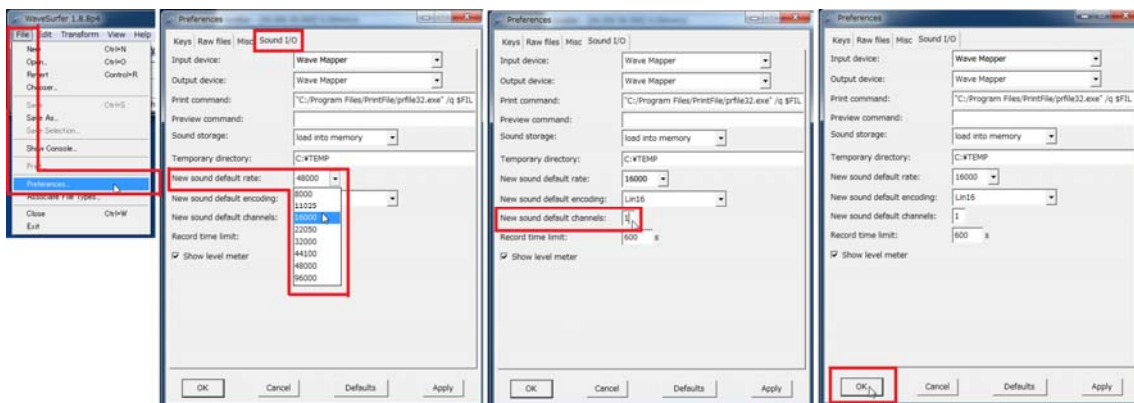


図 15 File メニューで Preferences を選び、Sound I/O で録音パラメータを設定する。

設定を終えたら、WaveSurfer を一旦閉じて、開き直し、マウスポインタを”Sound”の行 (初期状態では”Sound”と表示されるこの行には、録音内容を保存したとき、または音声ファイルを開いたときに音声ファイル名が表示される) に置きパラメータが正しく設定されていること (rate: 16000 channels: 1) を確認する (図 16)。設定が正しく行われていない場合には、以降の作業がすべて無駄になるので注意する。

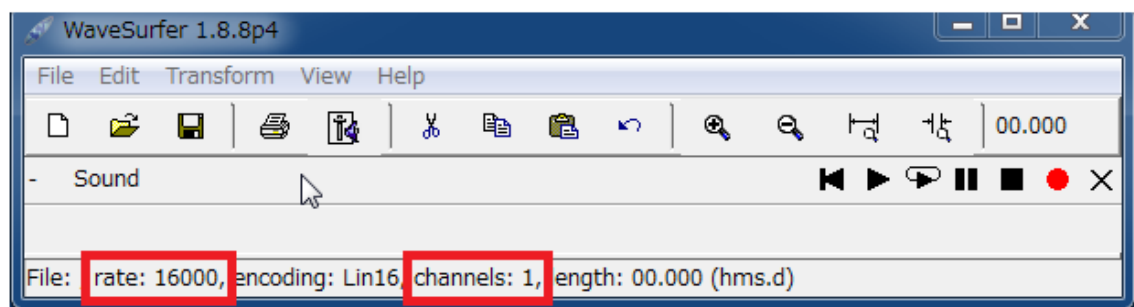


図 16 標準化周波数 16kHz、モノラルに設定されていることを確認する。

ヘッドセットを接続すると、これが既定のサウンドデバイスとなるので、Sound I/O 頁の Input device、Output device の項目は既定の **Wave Mapper** のままにしておく。ヘッドセット（コントロールパネル→個人設定→サウンドでは”Reaktek High Definition Audio”と認識される）を明示的に指定すると、「スピーカー」、「マイク」が文字化けするため、次回以降起動しなくなる。設定ファイルはドキュメントライブラリの¥.wavesurfer¥1.8 に作られログインし直しても設定は保存されるため、ログインし直しても起動できない。もし、起動しなくなった場合には、設定ファイルをフォルダごと削除してやり直す。

3. 分析用音声を録音する。

注意すべきことは、録音レベルを適正にすること（p. 7「録音レベルと雑音について」参照）の他、録音中はヘッドセットのボリュームコントローラーのつまみを最小の位置にすることである（備考 p.23「音声出力の入力への漏洩について」参照）。

WaveSurfer の操作ボタンは、●で録音開始（IC レコーダや DTM ソフトの●ボタンとは異なり録音スタンバイではなく、いきなり録音され、クリックする度にそれまでの内容が破棄される）、■で一時停止・再開、■で停止である。スペースバーで一時停止と解除（停止時からは再生、録音一時停止時には録音続行）のオルタネイト操作ができるのは通常の音楽・映像の再生ソフトと同様である。スペースバーを押しながら●をクリックすると録音スタンバイとなるが、この状態でレベルメーター（再生時のみ表示される）を表示してレベル調節することはできない。

分析課題の音声は p. 19 6 節に掲げる通りである。

以下の節の説明では、マイクロホンからの收音ではなく電子辞書の音声出力を PC の入力ジャックに接続し NHK 日本語アクセント辞典に収録された「箸」、「橋」「端（はし、はじ）」の音声を録音した例を対象に記述している。録音が終わったら、**File→Save As...**で拡張子は.wav を指定してファイルに保存する。

4. 開始、停止のノイズ、余分な無音部分を編集で切り取る。

録音後そのまま継続してもよいが、ここでは一般的な編集作業として新たに WaveSurfer を起動した状態から記す。**File→Open...**で編集対象のファイルを開くと、ペーン（p. 15 5 節後述）配置のプロンプトで **Choose Config...**のウィンドウが開くので、編集作業用の Waveformを選び編集作業を見易くするため、**View→250 mm/s**で時間軸を拡大する（p. 13 図 17）。ファイルパスに日本語文字を含まない場合には音声ファイルを WaveSurfer 実行ファイル **wavesurfer.exe** のショートカットにドラッグして WaveSurfer を起動することもできる（パスに日本語文字が含まれる場合もメニューの **File→Open...**では開ける）。

Waveform ペーン左端の縦軸の値が 16 ビット量子化に相応しい値であることを確認し（p. 8 図 9 参照）、実験ノートに記録する。録音レベルが適正でない場合は 3 節の作業をやり直す。

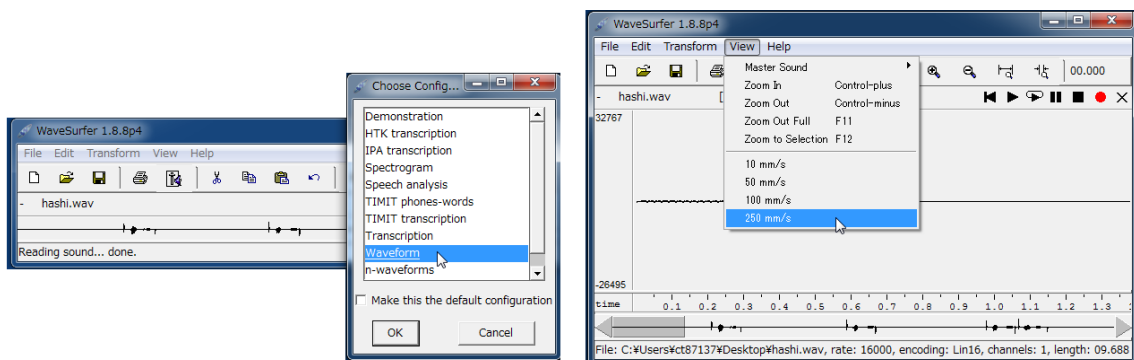


図 17 Choose Config...で Waveform を選び、View→250 mm/s で時間軸を拡大する。

ウィンドウ最下段の WaveBar に音声ファイル全体の波形が表示され、WaveBar で灰色 (図 18 の赤枠内) の部分がペーン (ここでは Waveform ペーンのみ開いているが、開かれたすべてのペーンについて) のウィンドウ内に表示される部分である。図 18 は、ウィンドウサイズを横に拡大し、WaveBar の灰色の部分で左クリック・ドラッグして編集対象箇所をスクロールした例である。

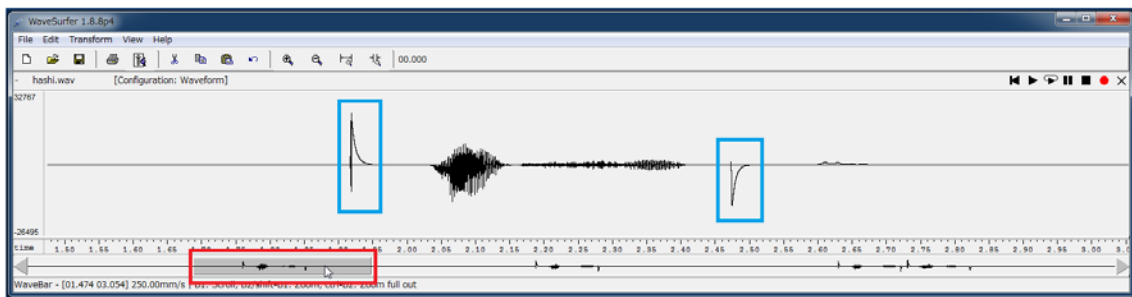


図 18 時間軸を拡大し編集対象箇所をスクロールで表示した例。

この例のノイズ波形 (図 18 の青枠内) は、PC の入力端子がマイクロホンの動作電源でもあるプラグインマイク用であるため電子辞書の音声再生の切替時に生じた電氣的なノイズであり、マイクロホンが拾うキー操作の音響ノイズとは異なるが編集操作は全く同じである。(a) スクロールして編集対象箇所を表示し (図 18)、(b) 左クリック・ドラッグで範囲を指定 (範囲の解除は任意の場所で左クリック) して (図 19)、(c) **Edit→Cut** で切り取る (p. 14 図 20、21)。この操作を繰り返し、編集が終了 (p. 14 図 22) したら **File→Save As...** でファイルに保存する (p. 14 図 23)。

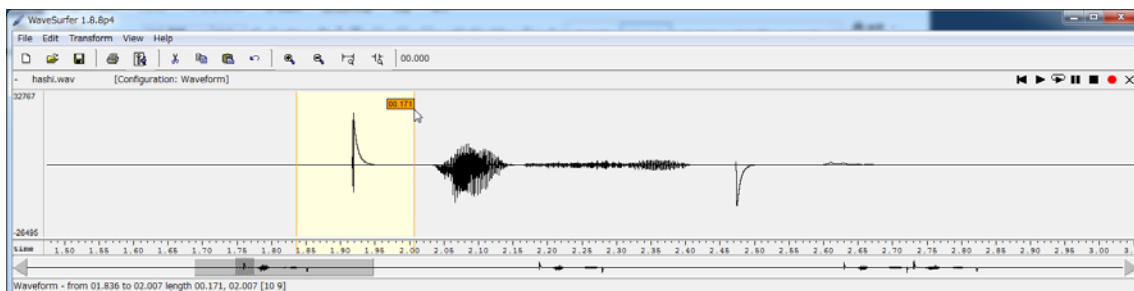


図 19 左クリック・ドラッグで範囲を指定する。

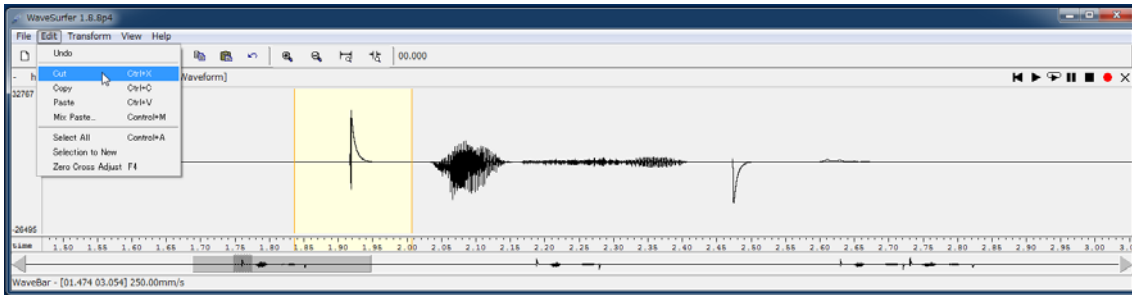


図 20 指定箇所を切り取る。

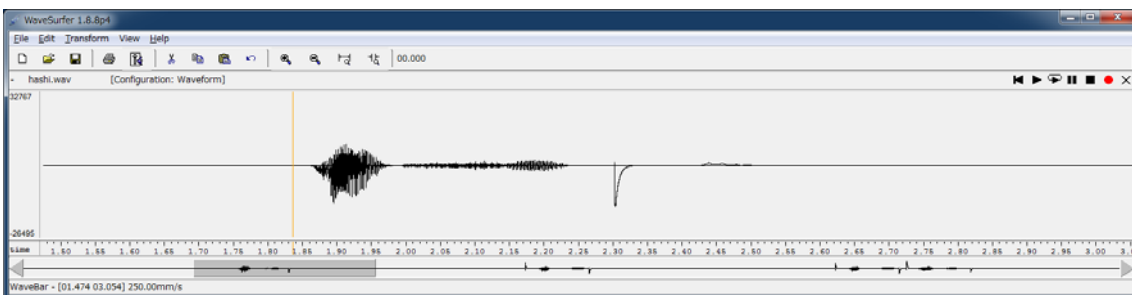


図 21 最初のノイズを切り取った状態

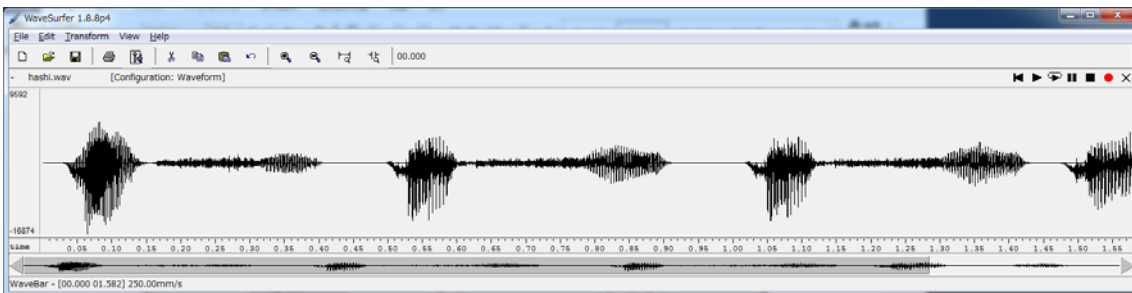


図 22 ノイズ、余分な無音部分をすべて切り取った状態

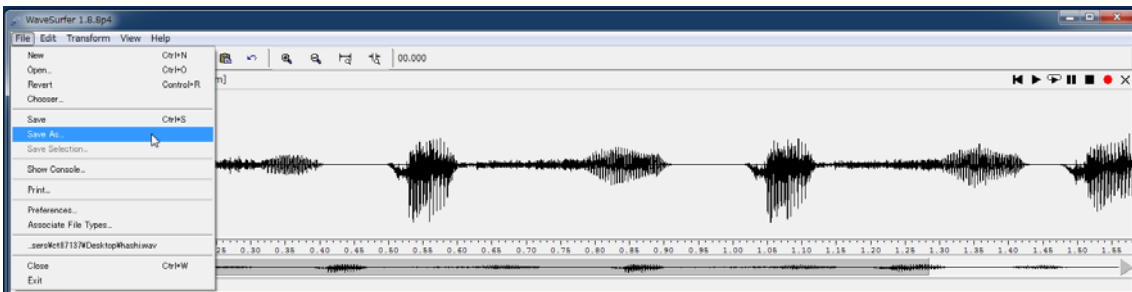


図 23 編集結果をファイルに保存する

保存ファイルの既定の拡張子は.smp であるが、ここでは、wav を指定する。

5. 分析ペーンを開き、必要なデータを保存する。

編集完了の状態からそのまま継続しても構わないが、一般の編集済の分析対象を保存した音声ファイルを開くものとして記述する。前節と同様の手順で開き (p. 13 図 17 の説明に対応)、これに分析用のペーンを追加する。ペーン (pane) とは窓枠などを意味し、WaveSurfer では編集・分析用の独立なウィンドウ区画を示す。ペーンが作成される場所は、右クリックした場所が属するペーンの上側である (上を先頭、下を末尾と見てエディタコマンドの append ではなく insert に相当)。最初に右クリック **Create Pane**→**Spectrogram** でスペクトログラム (縦軸に**標本化周波数の 1/2** である 8kHz まで等間隔に周波数を取り、音声の各周波数成分の強度の時間変化を描いたもので、既定では白黒で黒が強い部分を示す) を作成する (図 24)。

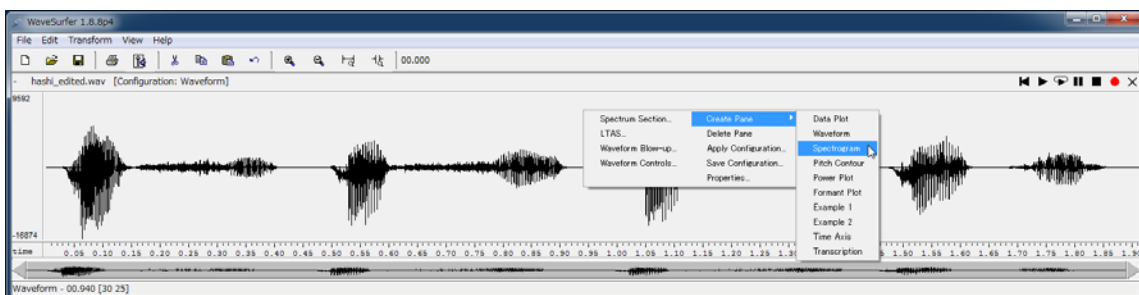


図 24 スペクトログラムペーンの作成

このままでもよいが、スペクトログラムペーン上で右クリックして **Properties...** を選び、**Spectrogram color.** で color を選ぶと (図 25) カラー表示となる (赤が強く、青が弱いことを示す)。更に、分析用に 2 つのペーンを追加する。

右クリック **Create Pane**→**Formant Plot** でフォルマントの時間変化をプロット (p. 16 図 26)、右クリック **Create Pane**→**Pitch Contour** で基本周波数変化のペーンを追加する (p.16 図 27)。前述の通り、新たなペーンは右クリックしたペーンの上に挿入される。

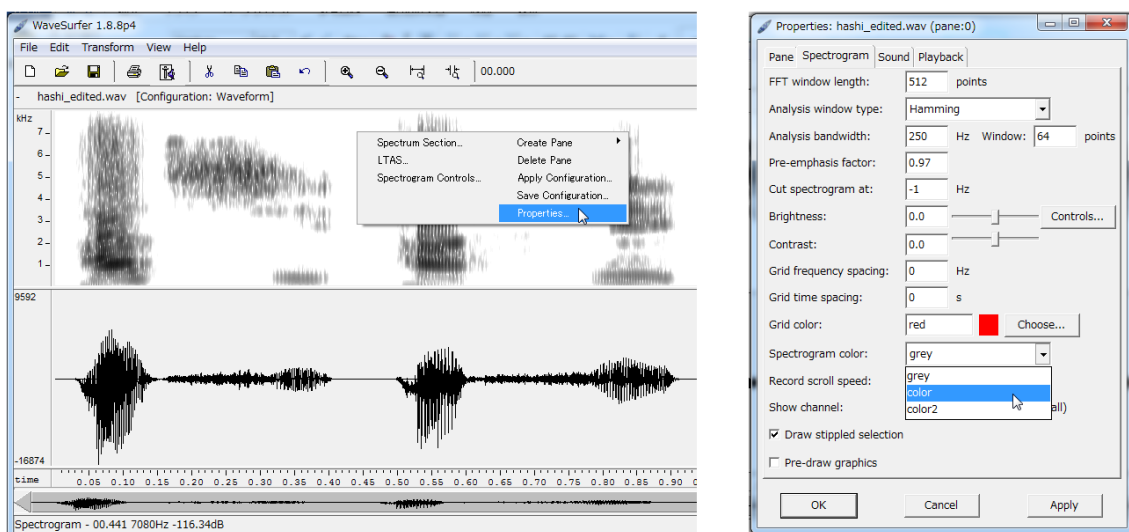


図 25 スペクトログラムの表示をカラーに変更する。

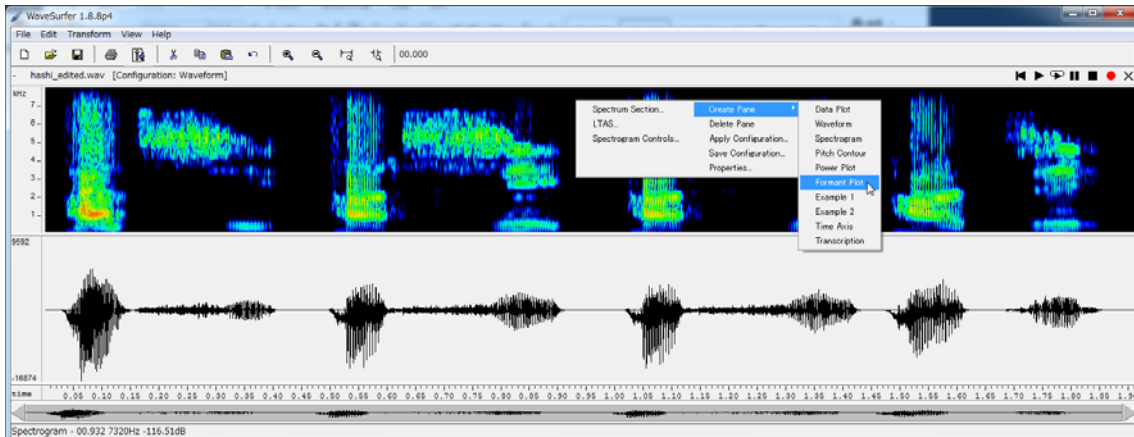


図 26 フォルマントプロットのペーンを開く。

フォルマントプロットのペーンで濃淡がはっきり分れない場合は録音レベルが適正でない。右クリック **Create Pane**→**Power Plot** で Power Plot ペーンを作成してパワー（高域通過フィルタを通して計算されており最大値は約 91dB）を確認し、母音区間のピークで 40dB 未満であれば録音をやり直すこと。

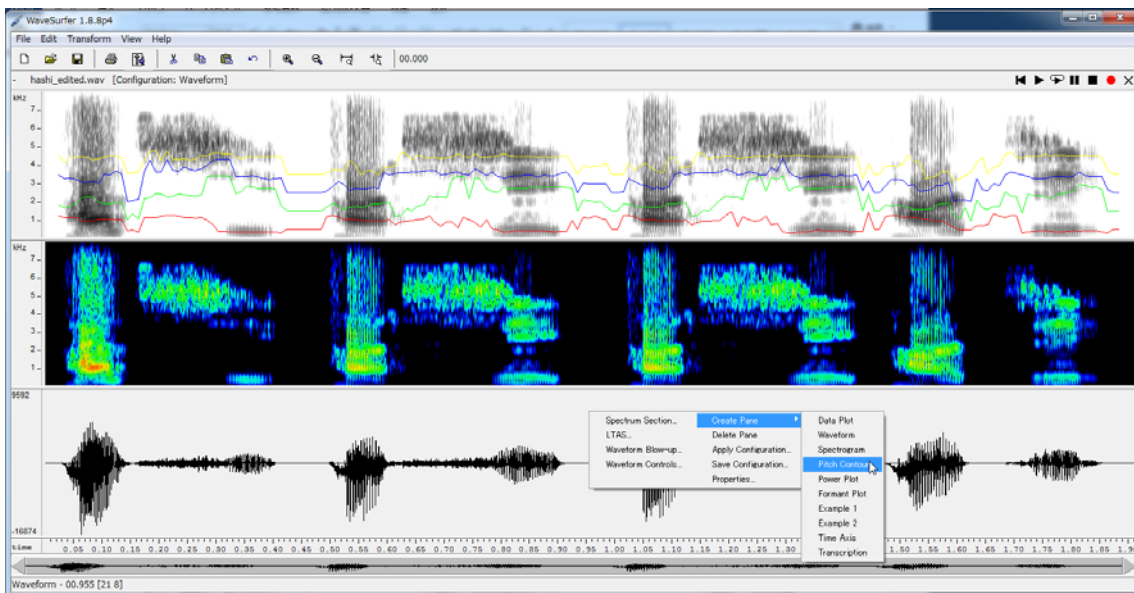


図 27 基本周波数の時間変化のペーンを開く。

p. 17 図 28 は再生ボタン (▶) をクリックして音声ファイルを再生している例で、縦の赤いカーソルが再生箇所、下の赤いバーがレベルメーターである。**Pitch Contour** ペーンの表示で「はし」のアクセントの違いがよくわかる。

全体を通して再生するのではなく再生範囲を指定するには、編集と同様に左クリック・ドラッグ（範囲の解除は任意の場所で左クリック）、再生開始箇所のみを指定するには、左クリック（解除は ◀ をクリック）をする。再生ボタン（または停止状態であればスペースバーでもよい）をクリックすると再生開始箇所に戻って再生するので、次に述べる **Statistics...** データ（分析範囲の 位置決めを慎重に行うこと）の取得・保存に際し、活用す

ること。

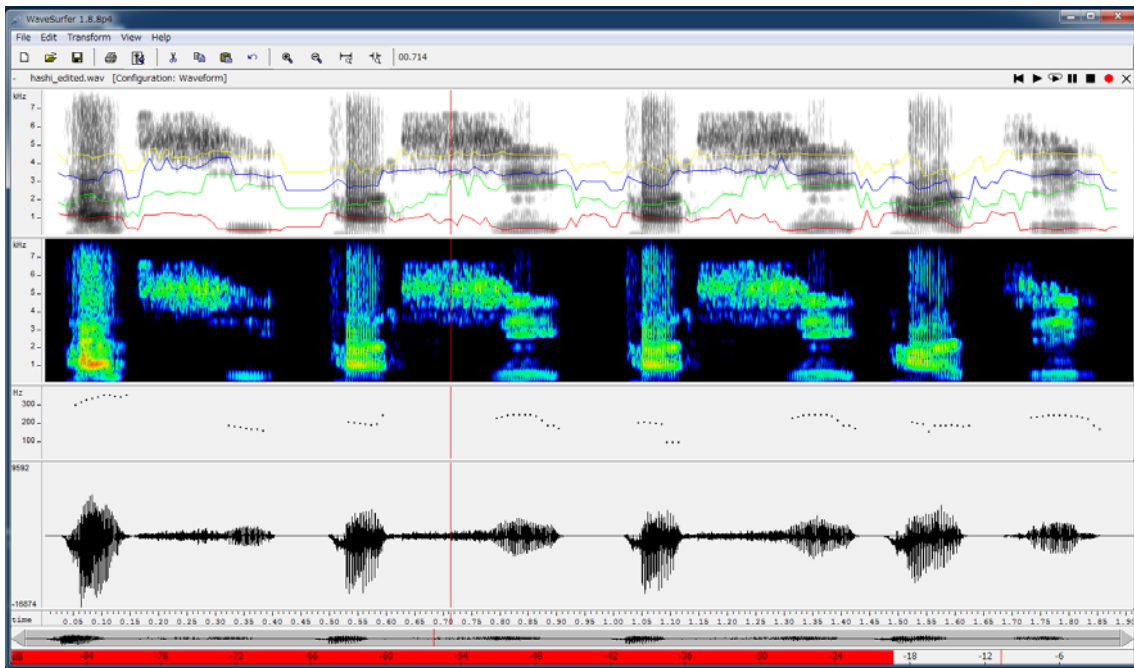


図 28 各種のペーンを開き、再生している画面

毎回音声ファイルを開く度に必要な分析ペーン（定番の分析ならば同一）を追加するのは煩雑である。ペーン配置の情報を設定ファイルに保存して使用するには、次の様にする。

ペーン（どのペーンでもよい）で右クリック **Save Configuration...** を選び（図 29 左）、設定を保存するファイル名（例では **Preset1.conf**）を指定する（図 29 中）。設定ファイルは **S:\Documents\wavesurfer\1.8\configurations** フォルダに保存され（これをコピーして個人 PC での実行環境にも反映させられる）、次回以降はファイルを開くとき **Choose Config...** で選べばよい（図 29 右）。

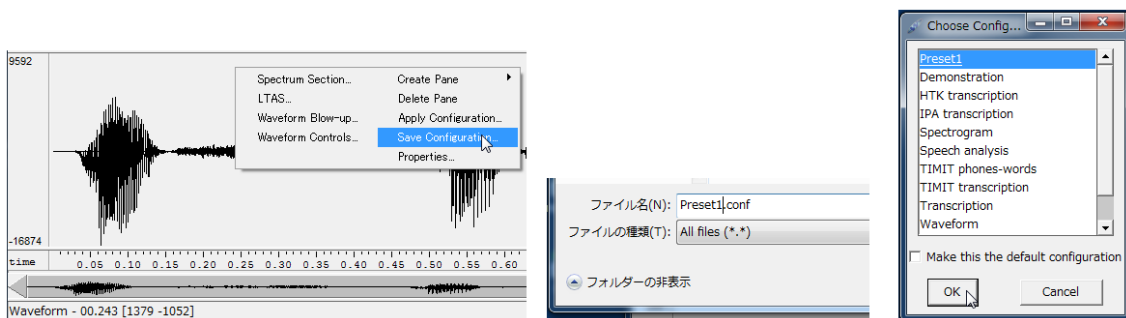


図 29 ペーン配置情報の保存と利用

フォルマント周波数と基本周波数のデータの取得、保存の方法は次の通りである。

- ① ペーンで右クリック **Statistics...** を選ぶと、左クリック・ドラッグで **指定した範囲**（p. 18 図 30 左の例では「箸」の「し」の母音「い」の区間。指定していない場合にはファ

イル全体) の当該ペーンの値 (**Formant Plot** ではフォルマント周波数、**Pitch Contour** では基本周波数) の平均値 (mean) と標準偏差 (sd) が計算される。Copy でクリップボードに送り、メモ帳に貼り付けてテキストファイルとして保存できる (図 30 右上)。

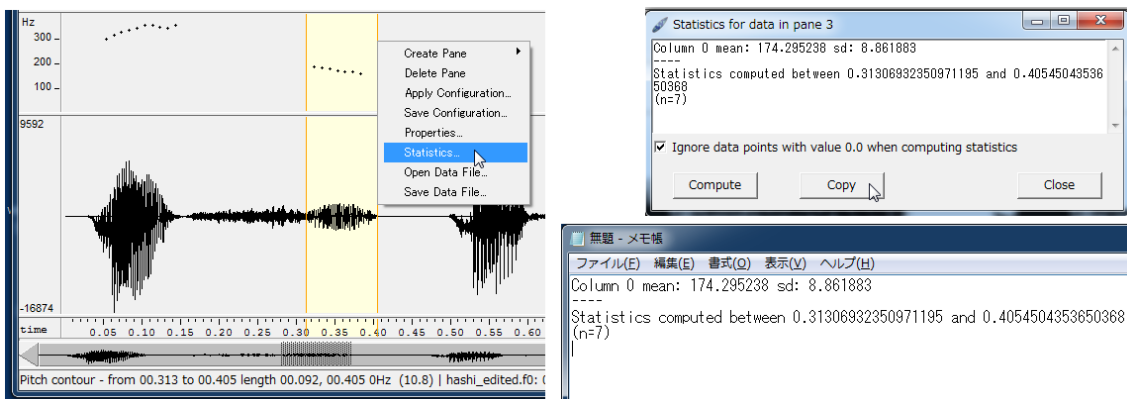


図 30 Pitch Contour ペーンで範囲を指定し Statistics...を表示した例

- ② ペーンで右クリック **Save Data File...** を選ぶと、当該ペーンのファイル全体の値 (**Formant Plot** ではフォルマントの周波数と帯域幅、**Pitch Contour** では基本周波数) がファイルに保存される。図 31 は、ペーンで右クリック **Properties...** で **Formants** 頁で **Add header in export file** にチェックを入れ (図 31 右赤枠内: Excel で処理する場合に有用である)、**Save Data File...** で保存する例である。

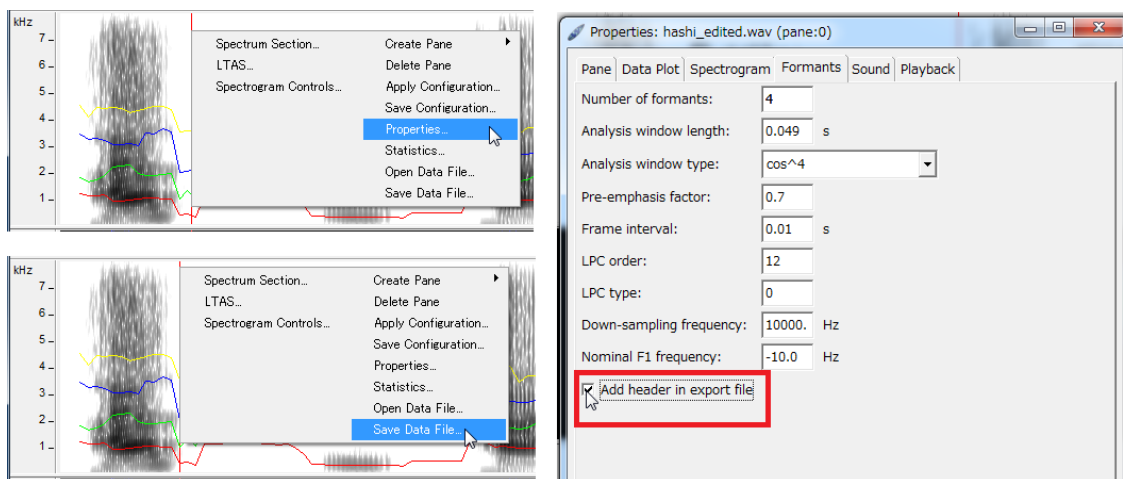


図 31 Properties...でヘッダ付加を指定してフォルマントデータを保存する。

保存したフォルマントデータ (既定の拡張子は .frm) を Excel の「テキストファイルのインポート」で開くときは、「すべてのファイル (*.*)」を指定して開き、「テキストファイルウィザード」では、「カンマやタブなどの区切り文字によってフィールドごとに区切られたデータ(D)」にチェックを入れ、「区切り文字」には「タブ(T)」と「スペース(S)」のみにチェックを入れ、「列のデータ形式」は「G/標準(G)」とする。p.19 図 32 はインポートした例で、ヘッダ部分の”Frame interval: 0.01 s”から本体のデータが 0.01 秒毎の値であること、”

Number of formants: 4”から各行の値は、第1～第4フォルマン트의周波数と帯域幅を表すことを示している。なお、無音部分のデータは”500.0 1500.0 2500.0 3500.0 1000.0 1000.0 1000.0 1000.0”となる。

	A	B	C	D	E	F	G	H
1	File:	C:\Users\ct87137\Desktop\hashi_edited.wav						
2	Window	type:	Hamming					
3	Window	length:	0.049 s					
4	Frame	interval:	0.01 s					
5	Pre-emphasis:	0.7						
6	Number	of	formants:	4				
7	LPC	order:	12					
8	LPC	type:	0					
9	Down-sampling	frequency:	10000 Hz					
10	Nominal	F1	frequency:	-10 Hz				
11	Range:	00.000-01.916	hms.d					
12	---	End	of	header	---			
13	1228.296387	1852.622803	3460.237793	4575.588867	123.9206543	150.3979187	285.7467651	521.2536452
14	1149.856812	1696.695801	3292.432861	4105.834961	801.9908905	192.9830017	369.7494812	435.2852478
15	1166.625854	1744.207275	3324.761475	4499.82666	83.60075378	160.7248993	250.5199585	725.6669312
16	1198.077637	1910.086792	3259.686768	4437.683105	199.4601746	561.7342529	627.1330566	570.8774414
17	963.2443237	2224.233643	3096.715332	4485.530762	698.7885742	303.6361389	298.5454102	703.5552368

図 32 Excel にインポートしたフォルメントデータの例

6. 分析課題音声について作業をする。

録音レベルが適正で雑音の混入が無ければ、Pitch Contour のデータが表示されている区間を母音区間と見てよい。これを元に課題音声の各母音区間の範囲を指定して（例 p.18 図 30）**再生して当該母音であることを確認し**、フォルメント周波数と基本周波数のデータを取得する。課題音声は次の通りである。

- ① 日本語の 5 母音を単独で区切って発声・録音し、フォルメント周波数と基本周波数を求める。第 1 フォルメント周波数を横軸、第 2 フォルメント周波数を縦軸に取って、「い」、「え」、「あ」、「お」、「う」の順に（この順で時計回りとなる）結んだ「母音五角形」を作成する。
- ② 同じ母音だけから成り、各母音から始まる「頭」「意識」「写る」「エゲレス（スは殆ど無声音）」「男」を発声・録音し（下線部の条件を満たせば、これ以外の言葉でも構わない）、①と比較する。
- ③ 「はし」（箸と橋）、「あめ」（雨と飴）、「はる」（春と貼る）、「こと」（琴と事）について、アクセントの異なるそれぞれ 2 つの言葉を発声・録音し、フォルメント周波数と基本周波数を求め、①と比較する。

注：フォルメント周波数の軸については、p. 7 図 8 の様に共に対数で目盛るもの、共に等間隔のもの、F1 が等間隔、F2 が対数目盛のものとは色々であるが、軸の目盛り方を変えたときの五角形の形の変化を確かめていただきたい。

備考

レポート作成について

6 節課題①の母音五角形については 2 種類以上（例えば等間隔と両対数）の目盛で作成

既定の形式を「16 ビット、44100 Hz (CD の音質)」に変更すると、THD 0.0033%、THD+N 0.0558%と僅かではあるが悪化する (図 34)。

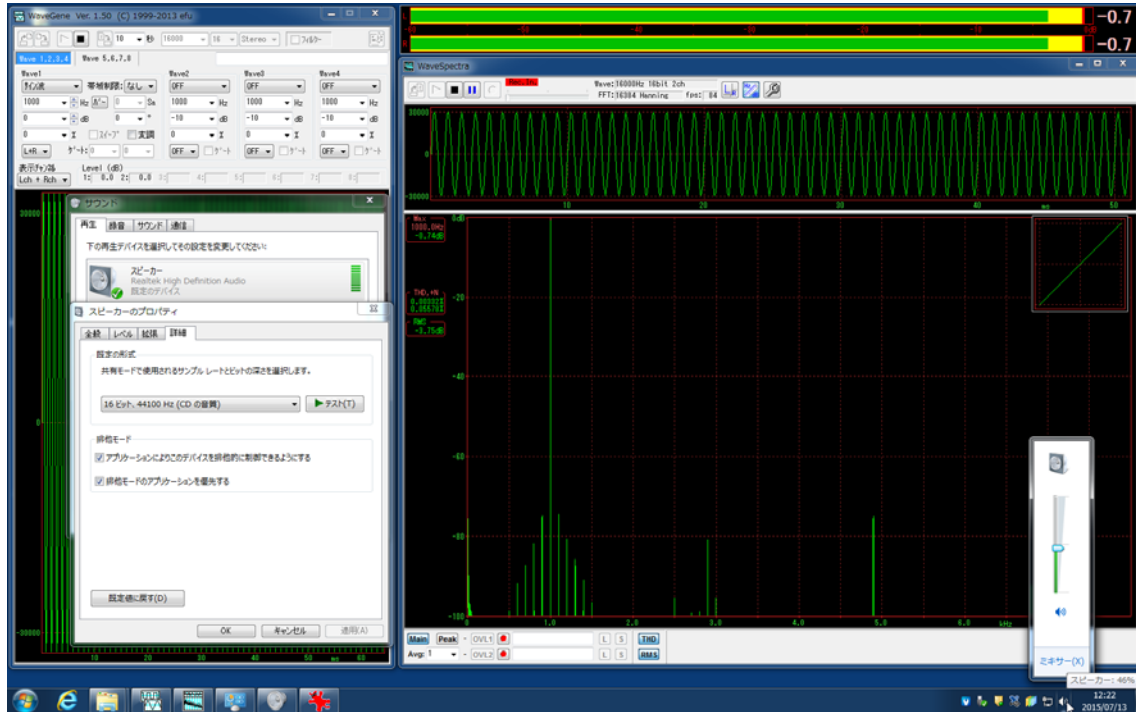


図 34 16 ビット、44100 Hz (CD の音質)

OSの違いによるPCサウンド機能の比較

サウンドカードの機能を使用する情報学実験の実験課題では、必ず Local PC でログインすることと注意を促している。本実験課題では分析に必要な周波数帯域が 8kHz であり CD 音質を必要としないが、現時点では Virtual Desktop の仮想オーディオデバイス (註) はこれも満たしてはいない。

註：仮想デバイスの性能は、学期毎、または学期中にも授業の実施形態・内容により必要に応じて変更される 情報科学研究教育センターの設定に依存するが、以下の記述は 2016 年 2 月 29 日の状況に基づいている。

サウンド機能を見るため、WaveGene の再生と WaveSpectra の録音フォーマットで標準化周波数を 44100Hz とし、WaveGene で白色雑音を出力し、PC 前面の入出力ジャックを直結して WaveSpectra でスペクトルを見る。ここでは高域の応答を比較しやすい様に、周波数スケールを対数ではなく等間隔目盛にしてある。Local PC (p. 22 図 35) がほぼ CD 音質を満足しているのに対し、Virtual Desktop の既定のデバイスでは、通常の電話帯域 (話が明瞭に聞き取れる 300Hz~3.4kHz) をカバーする程度である (p. 22 図 36)。人間が話す音声帯域は 7kHz 辺りまで伸びており (高音質電話の帯域は 100Hz~7kHz)、音声分析の目的には不十分と分る。コントロールパネル→個人設定→サウンド→録音で既定のデバイスを「マイク配列」から「マイク」に変更することで特性は 7kHz まで伸ばせるが、定期的

にパルスノイズが入るので本課題の目的には好ましくない。

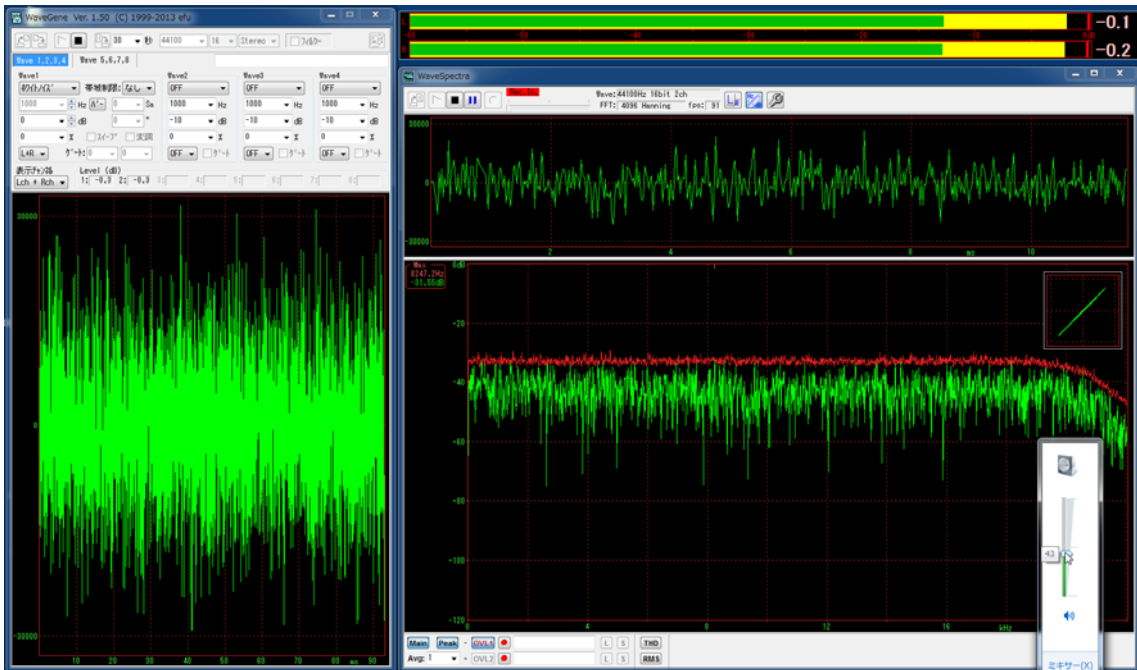


図 35 Local PC (Realtek High Definition Audio) の再生録音総合特性

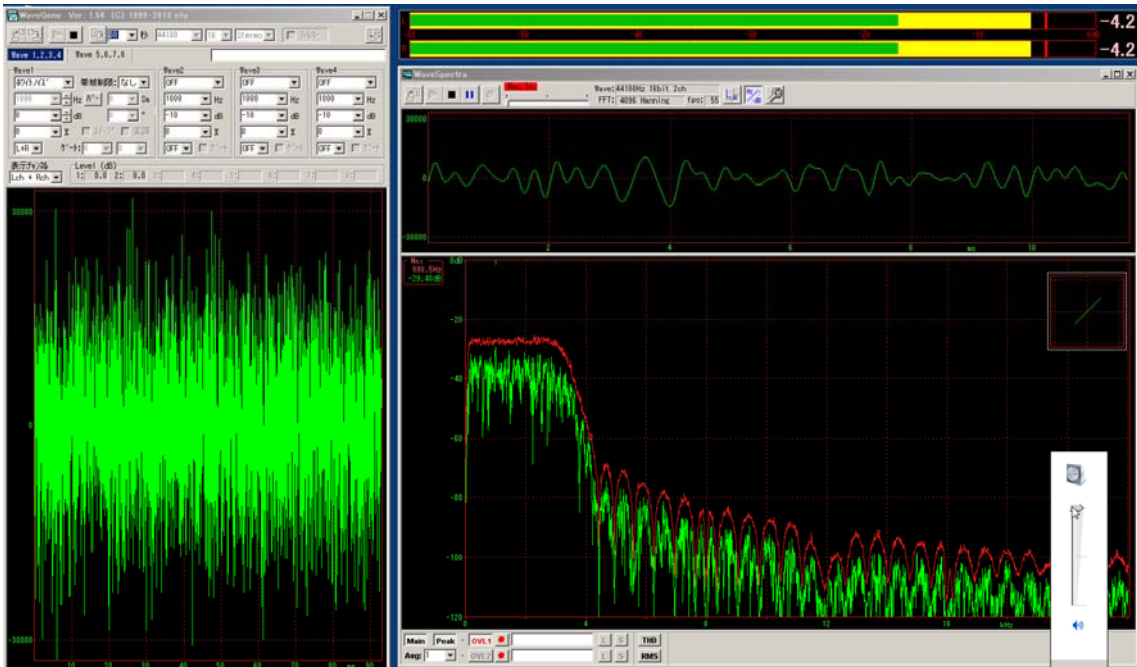


図 36 既定の録音デバイスでの Virtual Desktop の再生録音総合特性

音声出力の入力への漏洩について

録音時にボリュームコントローラーを最小の位置にする (p. 12) のは、ヘッドホンから出る音がマイクロホンに入る音響的な混入を防ぐ以外に、電気的な混入を減らす目的がある。電気的な混入には、(a) **PC 回路内部**の漏洩と(b) **ヘッドセットケーブル**内の誘導がある。

本実験と同様にサウンドカードの機能を使用する課題 I ②、課題 I ⑦では、(b)は無関係で、(a)についても PC の音声入力端子に加える信号がラインレベルの電圧であったため問題とはならなかった。これに対して、微弱なマイクロホンの出力電圧を扱う本実験ではマイクブーストをかけてゲイン (利得) を上げており、(高級オーディオ機器の様な対策をしていない PC では) 音声出力端子への信号が無視できないレベルで入力側に混入する。

漏洩による混入は、(a)(b)何れも出力回路に流れる電流で決まり、出力電圧と PC 側から見た負荷抵抗の値により変化する。(a)について確認するため、PC のヘッドホン出力の左チャンネルに 10Ω の抵抗 (電流は流れるが音は出ない)、右チャンネルを開放 (電流は流れず音も出ない) として、マイクロホン入力にはステレオプラグインマイクロホン (本実験で使用するマイクロホンはモノラルであるが、演習室 PC の音声入力回路はステレオに対応している) を接続する。

図 37 は「スピーカー」のレベルを 100 (最大)、「マイク」のレベルは当該マイクロホンの感度に合せてレベル 100 ブースト 24dB として WaveGene で 1kHz 0dB 正弦波を両チャンネルに出力し、これを WaveSpectra で測定した例である。スペクトルでは演習室内の環境音 (殆どが空調の騒音である) 以外に、1kHz の電気的な混入が見られるが、**出力側では電流の流れない右チャンネル**のマイクロホンにも左チャンネルと同量の混入が見られ、WaveGene で左チャンネルのみに出力した場合もこの図と同じになる。ここでは、サウンドカードの特性をそのまま示すため、スピーカーのプロパティ詳細の既定の形式を「16 ビット、44100Hz (CD の音質)」に変更し、WaveGene と WaveSpectra の標準化周波数も本実験の 16000Hz ではなく 44100Hz に設定している

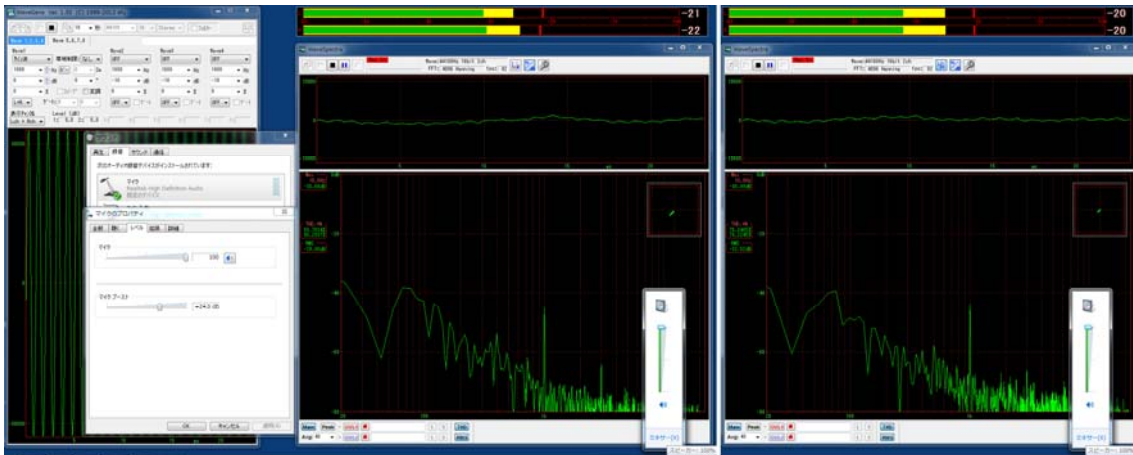


図 37 PC 回路内で出力信号の電流が流れない右チャンネルにも混入する

(b)の影響について確認するには、(a)を無視できる PC を必要とする。p. 24 図 38 はあるノート PC のヘッドホン出力に 10Ω 、マイクロホン入力に $1k\Omega$ の抵抗をそれぞれ両チャンネルに接続して再生、録音共にレベルを最大にして測定した例で 1kHz の入力側への混入は見られない。p. 24 図 39 はこの PC に実験で使用するヘッドセットを接続し、ボリュームコントローラーを最小、「スピーカー」と「マイク」のプロパティでレベルを本実験での再生録音レベルに設定した例で、これがサンワサプライ MM-HS515SV の(b)単独の漏洩を示している。音声入出力を細いケーブルにまとめた機器ではこの程度の漏洩は避けられない。



図 38 出力信号の入力信号への混入対策が十分な PC の例

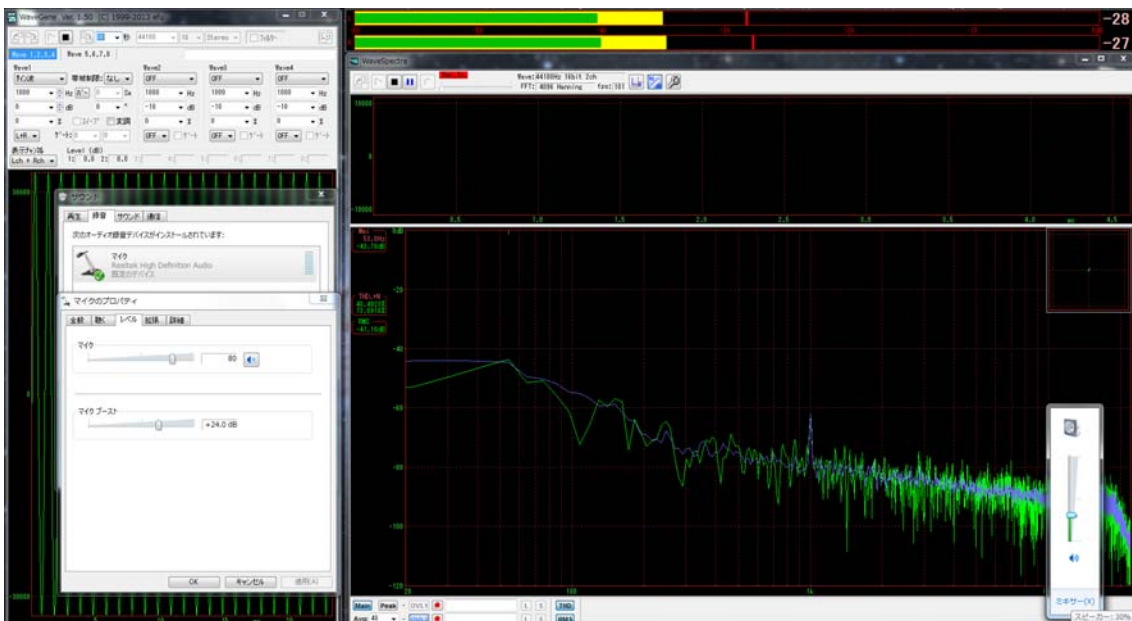


図 39 サンワサプライ MM-HS515SV のケーブルによる誘導

電気的な混入を完全に防ぐには(a)(b)の何れについてもヘッドホンのプラグを PC の音声出力ジャックから抜くかスピーカーのプロパティでレベルを 0 に絞ることであるが、本実験で使用するヘッドセットには**ボリュームコントローラー**が付いており（直流抵抗はテキスト執筆用の個体の場合、最大位置で左右共 38Ω、最小位置で左 1043Ω、右 1052Ω）、これを**最小の位置にすれば十分**である。p. 25 図 40 は、「スピーカー」と「マイク」のプロパティでレベルを本実験での再生録音レベルに設定し、ボリュームコントローラー最大の位置（左）と最小の位置（右）で 1kHz 0dB 正弦波を再生した様子である。演習室の環境音に埋もれているが右の図で 1kHz の混入は図 39 の値とほぼ同じである。

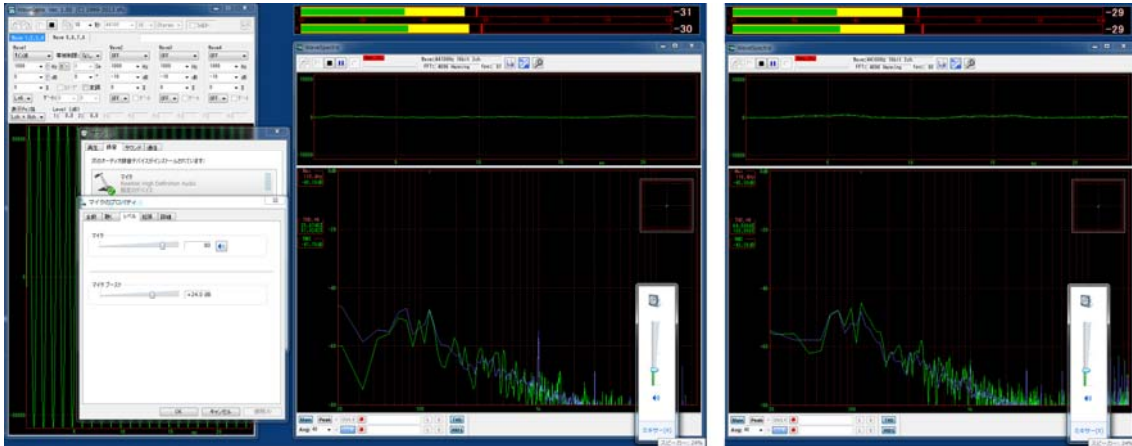


図 40 情報科学研究教育センターPC に接続した MM-HS515SV

共用ヘッドセット以外での収録について

録音機で課題の分析用音声を録音して、標本化周波数 16000Hz、モノラルに変換したファイルを USB メモリに保存しておくこと（ハンディレコーダの貸出は台数に限りがあるので申請は実験日の 2 週間前、貸出期間は 1 週間とする。実験日にイヤホンを用意すること。

事前に申し出てハンディレコーダ（課題Ⅱ⑨で使用する Zoom H4n）により、課題の音声（p. 19）を収録し、以下の手順で標本化周波数 16000Hz、モノラルの音声ファイルを作成する。

① Zoom H4n の説明書を熟読し、課題の音声を 44100Hz ステレオのフォーマットで録音する。録音機を単独で使用する場合、SD カードは各自で用意すること。PC に接続してオーディオインターフェースとして使用する場合、カードは不要である。

② 保存した音声ファイルをフリーの波形編集ソフト Wavosaur を用いて標本化周波数 16kHz モノラルに変換し保存する。

● 保存した音声ファイルを開く（File メニューの Open...またはドラッグ&ドロップ）。p.26 図 41 は p.3 図 1 左で再生している 16kHz モノラル音声ファイルの変換前のファイルを開いた状態である。

● Process メニューの Resample で標本化周波数を 16000Hz に変換する（p. 26 図 42）。

● Process メニューの Convert to mono でモノラルに変換し保存する（p. 26 図 43）。

図 41 のファイルでは 24 ビットから 16 ビットへの変換作業があるがここでは割愛する。

Zoom H4n 説明書

https://www.zoom.co.jp/sites/default/files/products/downloads/pdfs/J_H4n.pdf

Wavosaur

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

[http://www.wavosaur.com/download/files/Wavosaur.1.1.0.0-x86\(en\).zip](http://www.wavosaur.com/download/files/Wavosaur.1.1.0.0-x86(en).zip)

<http://www.wavosaur.com/download/manual/Wavosaur-quick-start-guide-v1.0.2.0.zip>

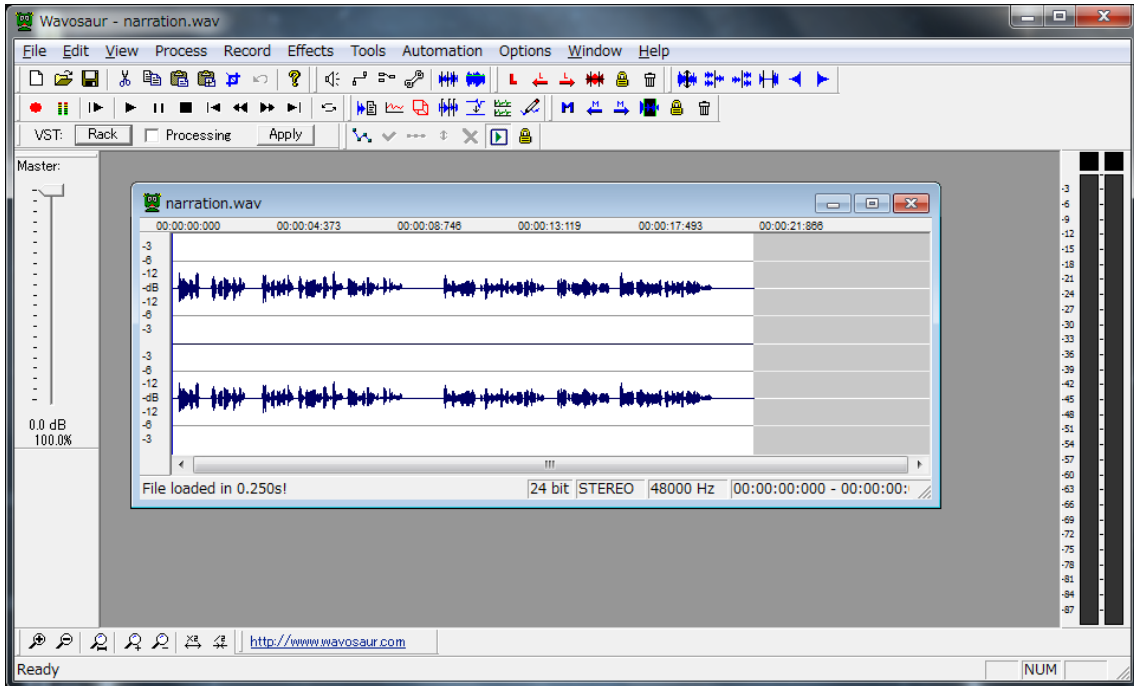


図 41 保存した音声ファイルを Wavosaur で開く

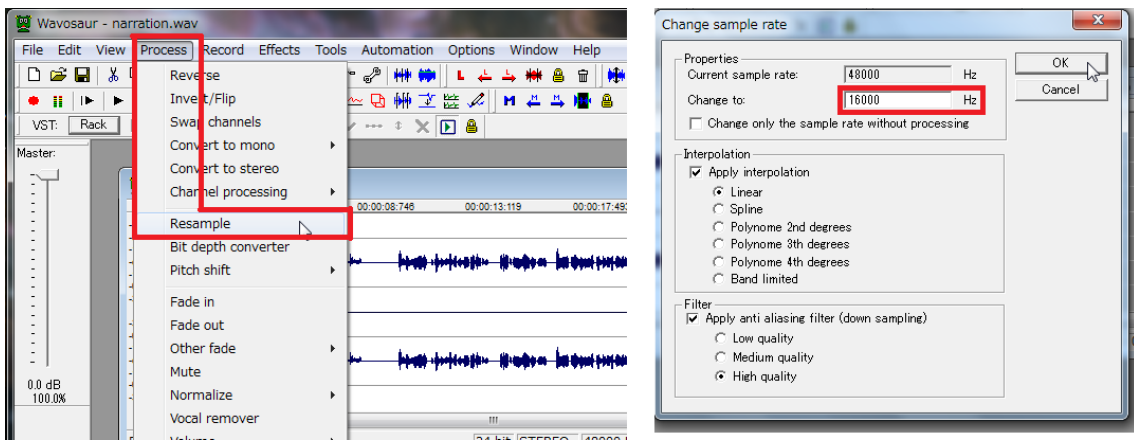


図 42 標本化周波数を 16000Hz に変換する

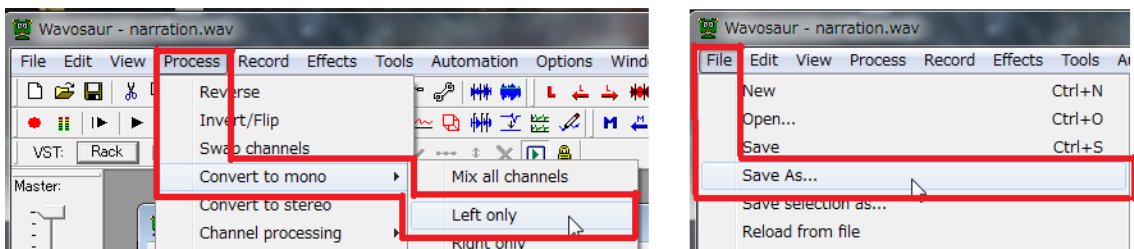


図 43 モノラルに変換し保存する