

TR-I-0310

Labelling Workbench

version 1.0

— User's Guide —

藤原 紳吾
S. Fujiwara

杉山 雅英
M. Sugiyama

1993.3

音素ラベリングワークベンチの操作方法、音素セグメンテーション・サーバーの
起動方法、およびワークベンチ・システムのインストール手順について説明する。

©ATR 自動翻訳電話研究所

©ATR Interpreting Telephony Research Laboratories

目次

1	はじめに	1
2	音素ラベリングワークベンチ・システムの概要	1
3	動作環境	1
3.1	ハードウェア	2
3.2	ソフトウェア	2
3.3	計算機の設定	2
(3.3.1)	音素セグメンテーション・サーバ用マシン	2
4	ラベリングワークベンチの操作方法	4
4.1	ワークベンチを使うための設定	4
4.2	ラベリングワークベンチの起動	4
4.3	ラベリングワークベンチの表示画面	5
4.4	ラベリングワークベンチの各サブウィンドウ	6
(4.4.1)	全波形	6
(4.4.2)	波形 — 音を聴く	6
(4.4.3)	ログパワー	6
(4.4.4)	スペクトル平均変化率	7
(4.4.5)	スペクトログラム — ラベリング: 音素境界入力	7
(4.4.6)	スペクトログラム (濃淡) — ラベリング: 音素境界入力	7
(4.4.7)	ラベル — ラベリング: ラベル記号入力	8
(4.4.8)	自動ラベル	8
(4.4.9)	時間軸 — 時刻表示	8
(4.4.10)	コマンドパネル — 各コマンドの実行	8
4.5	コマンドパネル	9
(4.5.1)	SpeechIn	9
(4.5.2)	WAVE	9
(4.5.3)	LABEL	9
(4.5.4)	Save	9
(4.5.5)	Previous Page	9
(4.5.6)	Next Page	9
(4.5.7)	Voice	9
(4.5.8)	Help	10
(4.5.9)	Quit	10
4.6	ラベリング — 音素境界・ラベル記号の入力、ラベルデータの保存	11
4.7	既存ラベルデータの参照 / 修正	12
(4.7.1)	既存ラベルデータの参照	12
(4.7.2)	既存ラベルデータの修正	12
4.8	操作履歴のログファイルへの保存 / ログファイルからの起動	12
(4.8.1)	操作履歴の保存	13
(4.8.2)	ログファイルからの起動	13
4.9	一括自動ラベリングで作成したラベルデータ・ファイルの利用	13

(4.9.1)	一括自動ラベリングによるラベルデータ作成	13
(4.9.2)	一括自動ラベリングの結果ファイルの利用	14
4.10	ラベリングワークベンチのカスタマイズ	14
(4.10.1)	設定ファイルのキーワード	15
(4.10.2)	サブウィンドウの配置	16
(4.10.3)	ラベル記号入力メニュー	17
5	音素セグメンテーション・サーバ	18
5.1	サーバのための設定	18
5.2	サーバの起動 / 終了	18
(5.2.1)	HMM セグメンテーション・サーバの起動 / 終了	18
(5.2.2)	SRK セグメンテーション・サーバの起動 / 終了	18
5.3	HMM セグメンテーション・サーバのみを使用する	19
5.4	知識に基づく音素セグメンテーション・サーバを使用する	20
6	インストールの手順	21
6.1	ソースファイルの構成	22
参考文献		23
付録 A	設定ファイルの例	24

1 はじめに

ラベリングワークベンチは、音素ラベルデータの効率的作成・保守のための支援ツールであり、エンジニアリング・ワークステーション (EWS) 上で稼働する^[2-5]。本資料では、音素ラベリングワークベンチ・システムの概要、ラベリングワークベンチの操作方法・カスタマイズの仕方、音素セグメンテーション・サーバの起動方法、およびラベリングワークベンチ・システムのインストール手順について説明する。

2 音素ラベリングワークベンチ・システムの概要

図1に音素ラベリングワークベンチ・システムの構成を示す。ラベリングワークベンチ・システム¹は、ラベリングワークベンチ²の本体であるユーザーインタフェース・モジュールと音素セグメンテーション・エンジンから構成される。システム全体はクライアント/サーバー形式^[1]となっており、音素セグメンテーション・エンジンはセグメンテーション・サーバーとして稼働し、ユーザーインタフェース・モジュール(クライアント)は自動ラベリングの結果を受け取る。自動ラベリングの結果が必要なければ、ラベリングワークベンチ(ユーザーインタフェース・モジュール)のみを動作させればよい。

ユーザ・インタフェース部は X Window System を利用しており、ウィンドウ上に波形、FFT スペクトログラムなどの音響パラメータが表示される。作業者は音響パラメータを観察しながら、マウスあるいはキーボード操作により簡単にラベリング作業を行うことができる。

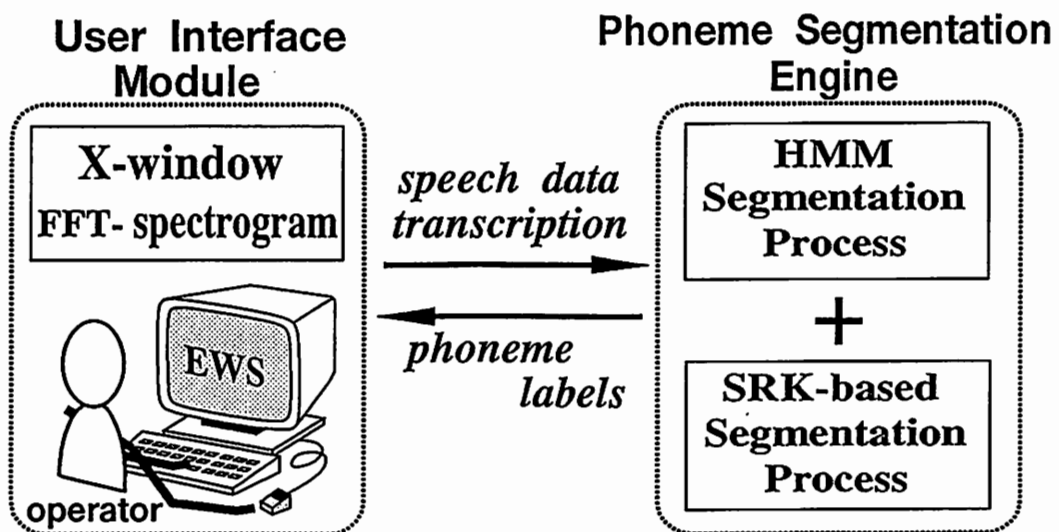


図 1: 音素ラベリングワークベンチ・システムの構成

3 動作環境

ラベリングワークベンチ・システムを動作させるために必要なハードウェア、ソフトウェア、計算機の設定などの計算機環境について説明する。

¹ラベリングワークベンチ・システム = ユーザーインタフェース・モジュール + 音素セグメンテーション・エンジン

²単にラベリングワークベンチといえ、ワークベンチ・システムにおけるユーザーインタフェース・モジュールを表す。

3.1 ハードウェア

- エンジニアリング・ワークステーション (EWS) — 少なくとも 1 台必要です。音素セグメンテーション・サーバを利用する場合は、セグメンテーション・サーバ専用 1 台あるのが望ましい。さらに、十分な swap space を確保できるだけの外部記憶装置 (Hard Disk など) があるのが、理想的である。
- DASBOX — 音声入力・出力用 A/D, D/A 変換装置
- オーディオ系 — Low Pass Filter、アンプ、マイクロフォン、スピーカなど

3.2 ソフトウェア

- X Window System — version 11, release 4 or 5
- SpeechIn プログラム — DASBOX を利用した音声入力プログラム
- daout プログラム — DASBOX を利用した音声出力プログラム

3.3 計算機の設定

(3.3.1) 音素セグメンテーション・サーバ用マシン

• Port 番号

/etc/services に、セグメンテーション・サーバが使用する Port 番号を登録する。これについては、page 18 を参照して下さい。

• System Configuration

音素セグメンテーション・サーバ *SRK-Seg-server*³, *HMM-Seg-Server*⁴ は、かなりのメモリを必要とします。したがって、コンフィギュレーション時に仮想記憶のサイズ (swap space size) をできるだけ大きく設定するのがよい。また、swap space size を設定し直すには、再コンフィギュレーションする必要があります。現在のコンフィギュレーションにおける swap space size は、pstat コマンド (DECstation) でわかります。

— pstat コマンドで swap space size を調べる —

```
% pstat -s
1349022k swap configured
44266k reserved virtual address space
      26104k used (7776k text, 0k smem)
      1322918k free, 920k wasted, 0k missing
avail: 41338*32k 102*1k
```

この例では、1349022k byte (約 1.3G byte) の swap space size です⁵。

参考までに、atrq02 (DECstation 5000, ultrix 4.2a) のコンフィギュレーション例を示す (atrq02:/usr/sys/conf/mips/ATRQ02)。

³*SRK-Seg-server* — スペクトログラムリーディング知識に基づく音素セグメンテーション・サーバ

⁴*HMM-Seg-Server* — HMM に基づく音素セグメンテーション・サーバ

⁵atrq02 では、1.3G byte の Hard Disk Unit を 全て swap space に使用している (1993.March 現在)。

ident	"ATRQ02"
maxdsiz	2550
maxssiz	2550

• limit 値の設定

プロセスで使用可能な swap space は、c-shell の limit 機能で制限されています。したがって、swap space size のコンフィギュレーション値が大きくても、stacksize の limit 値が小さければ、大きな swap space を使えません。この時は、limit⁶ で設定し直して下さい。

limit で limit 値を調べる／設定する

```
% limit
filesize      unlimited
datasize      1048576 kbytes
stacksize      2048 kbytes
coredumpsize  0 kbytes
memoryuse      unlimited
```

このように、stacksize が datasize より小さければ、再設定する。

```
% limit stacksize 1048576
```

これは、\$HOME/.cshrc または \$HOME/.cshrc に書いておくとよい。

注：limit stacksize unlimited では、うまく設定できないので、このようにある値を設定するようにする。

⁶limit は、set などのような c-shell 内のコマンドです。

4 ラベリングワークベンチの操作方法

4.1 ワークベンチを使うための設定

- 音素セグメンテーション・サーバ *SRK-Seg-server*, *HMM-Seg-Server* を利用する場合は、ワークベンチを動作させるマシンの `/etc/services` に次のエントリを登録する。

```
artimsrv          8001/tcp      # SRK Segmentation Server
HmmSeg           8002/tcp      # HMM Segmentation Server
```

- ラベリングワークベンチでは、以下の環境変数を使用される。

変数名	省略値	設定できる値
LABELLING_MODE	Nop	SrkSeg ----- <i>SRK-Seg-server</i> 使用
		HmmSeg ----- <i>HMM-Seg-Server</i> 使用
		ReadFile --- ラベリング結果ファイル使用
		Nop ----- 自動ラベリングしない
ALHmmSegServer	atrq02	ホスト名, ホストアドレス
ALSrkSegServer	atrq02	ホスト名, ホストアドレス
AUTO_LABELS	\$HOME/auto.labels	一括自動ラベリングの結果ファイル

4.2 ラベリングワークベンチの起動

ラベリングワークベンチ・コマンドは、`lwb`⁷ であり、次のように使う。

Usage of lwb

```
% lwb [ options ] wave_file
[ options ]                                [ default ]
-i<customize file> : customize file         : $HOME/.lwbrc
-l<label file>     : label file              : *.LB
-t<transcription>  : transcription in Roman :
-T<trans. table>   : transcription table    :
-f<list>           : wave file list         : .file
                  : (it's used in wave load menu) :
-e<LogFile>        : save history into log file : .lwbdemo
-E<LogFile>        : run by reading log file   : .lwbdemo
-v                : verbose mode
-h                : display this message
```

⁷/usr/local/bin/AL/lwb

—— ラベリングワークベンチの起動例 1 ——

```
% setenv LABELLING_MODE      SrkSeg
% setenv SrkSegServer         atrq02
% lwb -v -t"ikioi" /set1/MAU/WAV/D6/MAU_B_0001.12K
```

—— ラベリングワークベンチの起動例 2 ——

```
% setenv LABELLING_MODE      HmmSeg
% setenv SrkSegServer         atrq02
% lwb -i$HOME/.lwbrc-demo /set1/MAU/WAV/DSC/MAU_SC2_12.12K
```

—— ラベリングワークベンチの起動例 3 ——

```
% setenv LABELLING_MODE      ReadFile
% setenv AUTO_LABELS          $HOME/auto.labels
% lwb -i$HOME/.lwbrc-20K /set1/MAU/WAV/DSC/MAU_SC2_12.20K
```

4.3 ラベリングワークベンチの表示画面

ラベリングワークベンチの表示画面の例を図2に示す。メインウィンドウには複数個のサブウィンドウが設けられ、コマンドパネルのほかに、音響パラメータ、音素ラベル、時間軸がページ単位で表示される。また、マウスカーソルの位置にはサブウィンドウをまたがって縦線

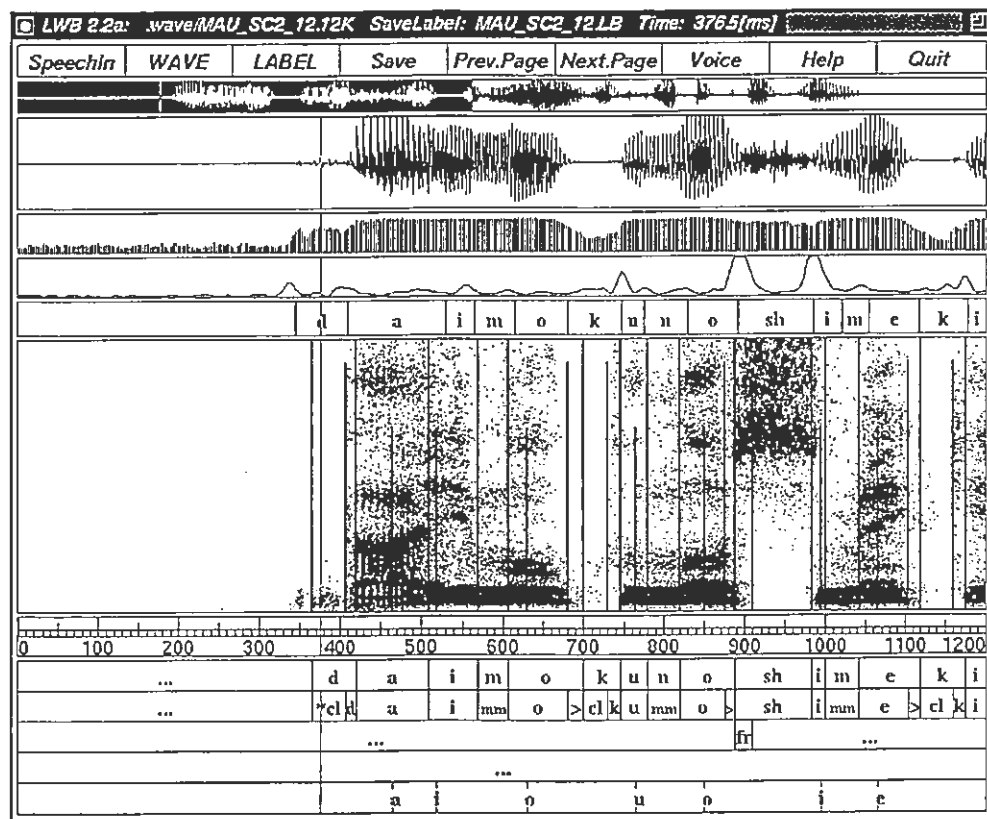


図 2: ラベリングワークベンチの操作ウィンドウ

(ガイドライン)が表示される。カスタマイズすれば、ラベリング操作に必要なサブウィンドウのみを表示することができ、図2の例では、表示ページは0～1200msであり、コマンドパネル、全波形、波形、ログパワー、スペクトル平均変化率、自動ラベル、スペクトログラム、時間軸、5層の音素ラベルが表示されている。

4.4 ラベリングワークベンチの各サブウィンドウ

ラベリングワークベンチ画面に表示可能な項目(サブウィンドウ)について説明する。

(4.4.1) 全波形

波形全体を表示する。図3に示すように、現在の表示ページにあたる波形部分が反転表示される。



図 3: 全波形・サブウィンドウ

(4.4.2) 波形 — 音を聴く

波形を表示する。このサブウィンドウを使用して音声出力ができる。音声出力には、daoutプログラムが使われる。次の3つのいずれかの方法により、マウスの左ボタンで領域を指定した後、

- (1) 開始点、終点を指定
- (2) ダブルクリックで現ページを指定
- (3) 何も指定しない・リセット状態 (→ 波形全体を指定)

マウスの中ボタンを押す(あるいは、**Voice** (音声出力コマンド)を実行する)と指定した領域が音声出力される。また、マウスの右ボタンはリセット用に使用される。



図 4: 波形・サブウィンドウ

(4.4.3) ログパワー

音声のログパワーを表示する。



図 5: ログパワー・サブウィンドウ

(4.4.4) スペクトル平均変化率

スペクトルの平均変化率を表示する。平均変化率として、 Δ ケプストラムのノルムに相当するパラメータを用いている [6]。LPCケプストラムを使用している。



図 6: スペクトル平均変化率・サブウィンドウ

(4.4.5) スペクトログラム — ラベリング: 音素境界入力

短時間FFT分析で求めたスペクトログラムをビットマップ表示する [7]。スペクトログラム強度は、ランダムドットの密度に変換される。このサブウィンドウを使用して音素ラベル境界を入力・削除・移動できる。

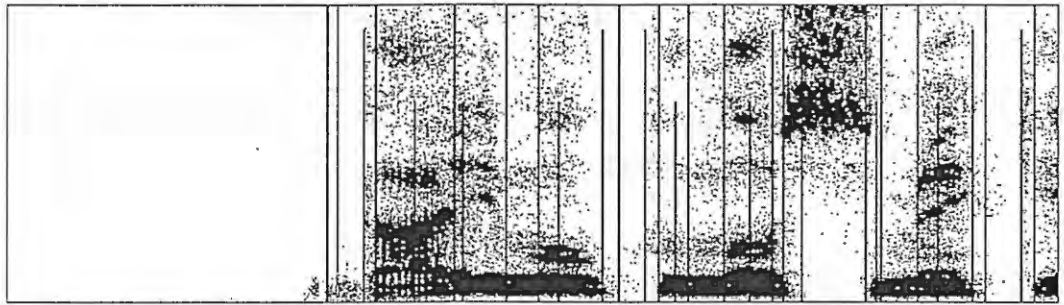


図 7: スペクトログラム・サブウィンドウ

(4.4.6) スペクトログラム (濃淡) — ラベリング: 音素境界入力

短時間FFT分析で求めたスペクトログラムを濃淡表示する。スペクトログラム強度は、濃淡 (グレイスケール) に変換される。このサブウィンドウを使用して音素ラベル境界を入力・削除・移動できる。

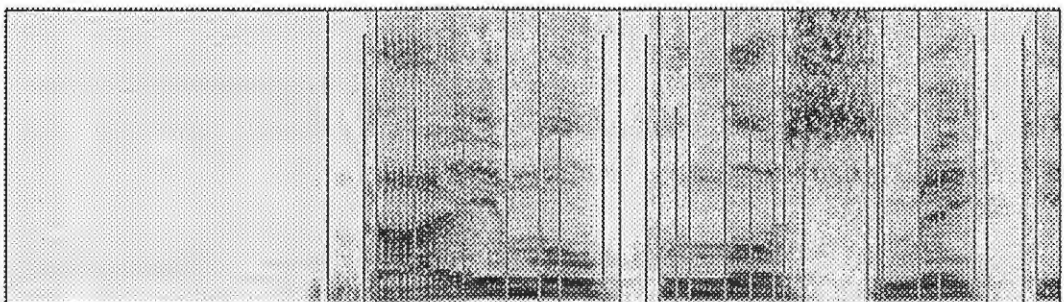


図 8: 濃淡スペクトログラム・サブウィンドウ

(4.4.7) ラベル — ラベリング: ラベル記号入力

作成 (あるいは修正) 中の音素ラベルデータを表示する。この際、ラベル記号は読みやすいように大きさが自動スケーリングされる。このサブウィンドウを使用して、音素ラベル記号を入力・修正できる。

...	d	a	i	m	o	k	u	n	o	sh	i	m	e	k	i						
...	*cl	d	a	i	mm	o	>	cl	k	u	mm	o	>	sh	i	mm	e	>	cl	k	i
	...										fr	...									
	...																				
	a i o u o i e																				

図 9: ラベル・サブウィンドウ

(4.4.8) 自動ラベル

音素セグメンテーション・サーバーによる音素ラベル情報を表示する。

	d	a	i	m	o	k	u	n	o	sh	i	m	e	k	i
--	---	---	---	---	---	---	---	---	---	----	---	---	---	---	---

図 10: 自動ラベル・サブウィンドウ

(4.4.9) 時間軸 — 時刻表示

時間軸を表示する。このサブウィンドウを使用して時刻を読みとれる。マウスの左ボタンをクリックすると、その点の時刻がメインウィンドウのタイトル・バーに表示される。

0	100	200	300	400	500	600	700	800	900	1000	1100	1200
---	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------

図 11: 時間軸・サブウィンドウ

(4.4.10) コマンドパネル — 各コマンドの実行

コマンド実行用メニューを表示する。メニュー項目をクリックすることで、波形データロード・セーブ、ラベルデータロード・セーブ、次ページへのスクロール、前ページへのスクロール、音声入力・出力、終了などのコマンドを実行できる。

SpeechIn	WAVE	LABEL	Save	Prev.Page	Next.Page	Voice	Help	Quit
----------	------	-------	------	-----------	-----------	-------	------	------

図 12: コマンドパネル・サブウィンドウ

4.5 コマンドパネル

コマンドパネルの各コマンドについて説明する。

SpeechIn	WAVE	LABEL	Save	Prev.Page	Next.Page	Voice	Help	Quit
----------	------	-------	------	-----------	-----------	-------	------	------

図 13: コマンドパネル・サブウィンドウ

(4.5.1) SpeechIn

SpeechIn⁸ プログラムを使用して、マイクロフォンから音声データを取り込む。発声テキストの一覧メニューが表示されるので、マウスの左ボタンで選ぶ。発声テキストはカレント・ディレクトリの .speechin というファイルから、発声テキスト一覧メニューが作られる。

(4.5.2) WAVE

波形データファイルを読み込む。波形ファイル名の一覧メニューが表示されるので、読み込むファイル名をマウスの左ボタンで選ぶ。ファイル名はカレント・ディレクトリの .file、あるいは オプション -l で指定したファイルから、一覧メニューが作られる。

(4.5.3) LABEL

ラベルデータファイルを読み込む。まず、カレント・ディレクトリにある *.LB を探す。もしなければ、/set1, /set2, /set3 の下にあるラベルデータファイルを読み込む。読み込まれたファイル名は、ウィンドウのタイトル・バーに表示される。ファイル名の拡張子 .LB は、カスタマイズ・ファイルで指定できる。

(4.5.4) Save

ラベリング結果 (音素境界位置、ラベル記号) をラベルデータとして、カレント・ディレクトリにセーブする。ファイル名は、*.LB である。ファイル名の拡張子 .LB は、カスタマイズ・ファイルで指定できる。

(4.5.5) Previous Page

前ページへのスクロール。

(4.5.6) Next Page

次ページへのスクロール。

(4.5.7) Voice

波形・サブウィンドウ内で指定した領域の音声データを、daout プログラム⁹ を使用して、音声出力する。

⁸SpeechIn は、SpeechIn -b -A 1200 と実行される。

⁹daout は、daout -f 標本化周波数 ファイル名 と実行される。

(4.5.8) Help

Not supported yet.

(4.5.9) Quit

/tmp にある temporary files を削除し、ラベリングワークベンチを終了する。

4.6 ラベリング — 音素境界・ラベル記号の入力、ラベルデータの保存

階層音素ラベルは、第1～5層ラベルまで入力でき、それぞれ次のように用いられる [8]。

- 第1層 (音声記号層) — 発声のローマ字表記を対応する音声区間に記述する。
- 第2層 (イベント層) — 音声現象を出来るだけ詳細に記述するために、各種の記号を付与する。
- 第3層 (異音化層) — 異音化現象を記述する。
- 第4層 (融合化層) — スペクトログラム上で区分化不可能な複数の音声事象を記述する。
- 第5層 (母音中心層) — 母音区間で最も安定した箇所を明示する。

各層のラベルデータは、スペクトログラム (または濃淡スペクトログラム)・サブウィンドウ上で音素境界線、ラベル・サブウィンドウ上で音素ラベル記号を入力することで作成される。音素ラベルの入力手順を以下に示す (図14参照)。

1. ラベル層の選択 — 入力したいラベル層に対応するラベル・サブウィンドウをマウスの右ボタンでクリックする。選択されたサブウィンドウは背景色が変わる。
2. 音素境界線の入力 (挿入) — スペクトログラム・サブウィンドウ上で音素境界を設定したい場所で、マウスの左ボタンをクリックすると境界線が描かれる。選択しているラベル・サブウィンドウにも音素区間の境界線が描かれる。
3. 音素境界線の移動 — スペクトログラム・サブウィンドウ上に描かれている音素境界線をマウスの左ボタンでクリックし (境界線の色が変わる)、移動先でもう一度マウスの左ボタンをクリックする。
4. 音素境界線の削除 — スペクトログラム・サブウィンドウ上に描かれている音素境界線をマウスの左ボタンでダブル・クリックする。
5. 音素ラベル記号の入力 (ポップアップメニュー使用) — 選択しているラベル・サブウィンドウ上の音素区間をダブルクリックすると、ラベル記号入力用のポップアップ・メニューが現れる。ポップアップ・メニュー上の記号をクリックすると、ラベル記号として入力される。
6. 音素ラベル記号の入力 (キーボード使用) — 選択しているラベル・サブウィンドウ上の音素区間をクリックした後、キー入力する。リターンキーで確定すると入力対象の音素区間は右に移動する。入力対象の区間は矢印キー (←→↑↓) でも移動できる。

音素ラベル記号を入力したあとでも、新たな音素境界線の入力、既存の音素境界線の移動・削除が可能である。

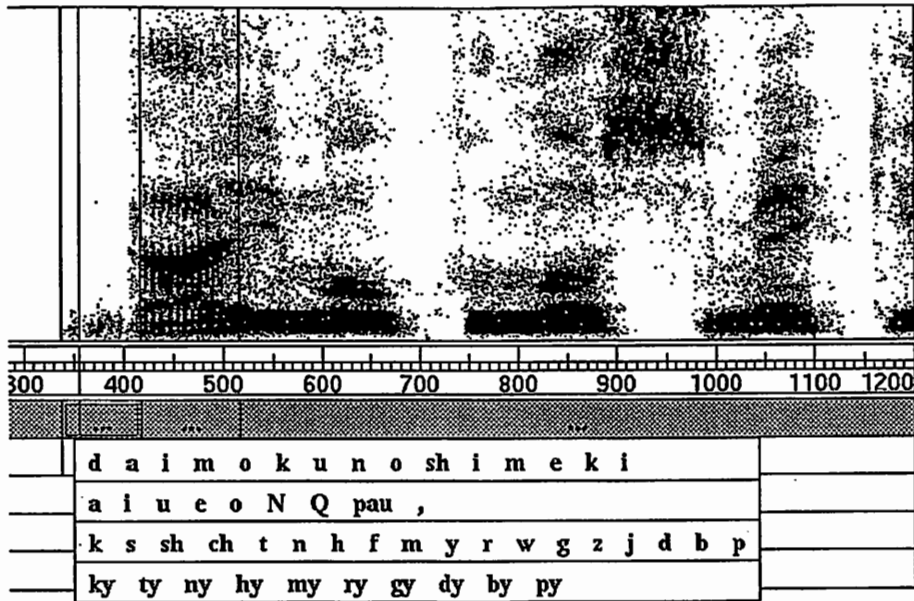


図 14: 音素ラベル記号の入力用ポップアップメニュー

4.7 既存ラベルデータの参照 / 修正

(4.7.1) 既存ラベルデータの参照

オプション -l で、ラベルデータファイル名を指定して、ラベリングワークベンチ lwb を起動する。

—— 既存ラベルデータの参照 ——

```
% setenv LABELLING_MODE      Nop
% lwb -l/set1/MAU/LBL/DSC/MAU_SC2_12.LB MAU_SC2_12.12K
```

(4.7.2) 既存ラベルデータの修正

オプション -l で、ラベルデータファイル名を指定して、ラベリングワークベンチ lwb を起動する。ラベルを修正した後、セーブする。修正したラベルデータは、カレント・ディレクトリにセーブされる。次の例では、./MAU_SC2_12.LB として、セーブされる。

—— 既存ラベルデータの修正 ——

```
% setenv LABELLING_MODE      ReadFile
% setenv AUTO_LABELS          $HOME/auto.labels
% lwb -l/set1/MAU/LBL/DSC/MAU_SC2_12.LB MAU_SC2_12.12K
```

4.8 操作履歴のログファイルへの保存 / ログファイルからの起動

ラベリング操作履歴の保存、異常終了時の入力ラベルデータの復旧、あるいは、ログファイルからの自動デモンストレーションを行うためのヒストリー機能を持つ。

(4.8.1) 操作履歴の保存

ラベリングワークベンチ起動時に、**オプション -e** を指定すると、マウスカーソルの移動・ボタンクリック、キーボード入力など、すべての操作の履歴がログファイルに保存される。

—— 操作履歴をログファイルにセーブする ——

```
% lwb -e/tmp/LWB.fujiwara.log MAU_SC2_12.12K
```

(4.8.2) ログファイルからの起動

ラベリングワークベンチ起動時に、**オプション -E** を指定すると、ログファイルの履歴を読んで、ワークベンチが動作する。ログファイルを全て読んだ後は、通常の使い方ができる。

—— ログファイルからの動作 ——

```
% lwb -E/tmp/LWB.fujiwara.log MAU_SC2_12.12K
```

4.9 一括自動ラベリングで作成したラベルデータ・ファイルの利用

(4.9.1) 一括自動ラベリングによるラベルデータ作成

ALgetAutoLabel プログラムを使用して、自動ラベリングの結果ファイルを作成する。

—— 一括自動ラベリングによるラベルデータの作成の例 1 ——

```
% setenv LABELLING_MODE      SrkSeg
% setenv SrkSegServer         atrq02
% ALgetAutoLabel MAU_B_0001.12K ikioi      > /tmp/labels
% ALgetAutoLabel MAU_B_0002.12K iyoiyo    >> /tmp/labels
% ALgetAutoLabel MAU_B_0003.12K urayamashii >> /tmp/labels
% ALgetAutoLabel MAU_B_0004.12K omoshiroi >> /tmp/labels
...
```

—— 一括自動ラベリングによるラベルデータの作成の例 2 ——

```
% setenv LABELLING_MODE      HmmSeg
% setenv SrkSegServer         atrq02
% ALgetAutoLabel MAU_B_0001.12K PAUikioiPAU      > /tmp/labels
% ALgetAutoLabel MAU_B_0002.12K PAUiyoiyoPAU    >> /tmp/labels
% ALgetAutoLabel MAU_B_0003.12K PAUurayamashiiPAU >> /tmp/labels
% ALgetAutoLabel MAU_B_0004.12K PAUomoshiroiPAU >> /tmp/labels
...
```


(4.9.2) 一括自動ラベリングの結果ファイルの利用

一括自動ラベリングの結果ファイルの利用

```
% setenv LABELLING_MODE      ReadFile
% setenv AUTO_LABELS          /tmp/labels
% lwb -t"ikioi" MAU_B_0001.12K
% lwb -t"iyoiyo" MAU_B_0002.12K
```

4.10 ラベリングワークベンチのカスタマイズ

ワークベンチにおいて、作業に必要な音響パラメータの選択、任意の音響分析条件への対応などを可能にするため、カスタマイズ機能を持つ。動作環境の設定ファイル(.lwbrc)を記述することで、作業者ごとにワークベンチのカスタマイズが可能である。設定ファイルでは、

- メインウィンドウの位置・大きさ
- 表示サブウィンドウとその位置・大きさ
- 音響パラメータの分析条件
- 1 ページの時間幅とページ間隔
- ラベル入力メニューで用いるラベル記号
- 文字フォント、背景色

などを指定できる。

(4.10.1) 設定ファイルのキーワード

表1から表7に、設定ファイルにおけるキーワードと設定内容を示す。

表 1: メインウィンドウ表示関連

MainWindowX	メインウィンドウの左上X座標
MainWindowY	メインウィンドウの左上Y座標
MainWindowWidth	メインウィンドウの幅
MainWindowHeight	メインウィンドウの高さ
MainWindowBackground	メインウィンドウの背景色
PopupMenuBackground	ポップアップメニューの背景色

表 2: サブウィンドウの種類

MenuWindow	メニューサブウィンドウ
WaveWindow	波形サブウィンドウ
WavePositionWindow	全波形サブウィンドウ
PowerWindow	ログパワーサブウィンドウ
DeltaCepWindow	Δ ケプストラムサブウィンドウ
AutoLabelWindow	自動ラベルサブウィンドウ
SpecWindow	スペクトログラムサブウィンドウ
ColorSpecWindow	濃淡スペクトログラムサブウィンドウ
TimeScaleWindow	時間軸サブウィンドウ
LabelWindow1	ラベル (第1レイヤ) サブウィンドウ
LabelWindow2	ラベル (第2レイヤ) サブウィンドウ
LabelWindow3	ラベル (第3レイヤ) サブウィンドウ
LabelWindow4	ラベル (第4レイヤ) サブウィンドウ
LabelWindow5	ラベル (第5レイヤ) サブウィンドウ

表 3: サブウィンドウ表示関連

WindowSeparate	サブウィンドウ間のセパレート幅
SubWindowBackground	サブウィンドウの背景色
LayerHighlight	ラベルサブウィンドウ選択時の背景色
BDLineColorName	音素境界線の色
BDRemarkColorName	音素境界線のリマーク時の色
BDLineWidth	音素境界線の幅
DisplayTimeRange	1 ページに表示する時間 (m 秒)
DisplayScrollTime	スクロール時間幅 (m 秒)
DisplayWaveRange	波形の表示レンジ
DisplayPowerRange	パワーの表示レンジ
DisplayDeltaCepRange	Δ ケプストラムの表示レンジ
DisplaySpecFreqRange	スペクトログラムの表示周波数レンジ
DisplaySpecCoeff	濃淡スペクトログラムの強度変換係数
DisplaySpecRange	濃淡スペクトログラムの表示レンジ

表 4: 音響分析関連

AnaSamplingFreq	サンプリング周波数 (kHz)
AnaFrameWidth	分析フレームサイズ (サンプル数)
AnaFrameShift	分析フレームシフト長 (サンプル数)
AnaFFTLength	FFT 分解点数
AnaLPCOrder	LPC 分析次数
AnaCepOrder	CEP 次数
AnaDeltaCepDelta	Δ フレーム数 (平均変化率)

表 5: ラベルシンボル入力メニュー関連

LabelMenu1	第1層ラベル入力メニューのシンボル
LabelMenu2	第2層ラベル入力メニューのシンボル
LabelMenu3	第3層ラベル入力メニューのシンボル
LabelMenu4	第4層ラベル入力メニューのシンボル
LabelMenu5	第5層ラベル入力メニューのシンボル

表 6: フォント

FontNormal	デフォルトのフォント
FontMenu	メニューで使われるフォント
FontAutoLabel	自動ラベル表示用フォント
FontLabel	ラベル表示用フォント
FontLabelSmall	ラベル表示用フォント (小)
FontLabelTiny	ラベル表示用フォント (極小)
FontTimeScale	時間軸表示用フォント
FontWaveList	波形ファイル名入力メニューのフォント
FontLabelMenu	ラベルシンボル入力メニューのフォント

表 7: 拡張子・コマンド名

ExtensionLabelFile	ラベルファイル名の拡張子
VoiceCommand	音声出力コマンド名

(4.10.2) サブウィンドウの配置

サブウィンドウは、設定ファイルで記述された順番にメインウィンドウの上方から配置される。この際、サブウィンドウ間には *WindowSeparate* で指定したピクセル値分のスペースが取られる。また、各サブウィンドウの大きさはサブウィンドウ名 (表 2) に与える数値で決められる。ただし、数値が 1 未満の時は $\text{MainWindowHeight} \times \text{数値}$ の値が大きさとなる。例えば、

```

MainWindowWidth: 1000
MainWindowHeight: 700
WaveWindow: 100
PowerWindow: 0.05
TimeScaleWindow: 40
SpecWindow: 0.5
Label1Window: 50
Label2Window: 50
WindowSeparate: 5

```

と記述されていれば、

高さ 100 の波形サブウィンドウ、
 高さ 35 のログパワーサブウィンドウ、
 高さ 40 の時間軸サブウィンドウ、
 高さ 350 のスペクトログラムサブウィンドウ、
 高さ 50 の第1層ラベルサブウィンドウ、
 高さ 50 の第2層ラベルサブウィンドウ

がメインウィンドウ (1000×700) 内に上から5ピクセル間隔で配置される。

(4.10.3) ラベル記号入力メニュー

設定ファイルで、

```

LabelMenu1: a i u e o \ N \ k s t n h m y r w
LabelMenu2: < > c l * c l t r d j m m \ a i u e o
LabelMenu3: f r d v

```

と記述すれば (\\ はメニューにおける改行を指示する)、第1層、第2層、第3層のラベル記号入力時のポップアップメニューは、それぞれ次のようになる (第4、5層も同様)。

a i u e o
N
k s t n h m y r w

< > c l * c l t r d j m m
a i u e o

f r d v

5 音素セグメンテーション・サーバ

音素セグメンテーション・サーバとして、

- *HMM-Seg-Server* — HMM に基づく音素セグメンテーション・サーバ
- *SRK-Seg-Server* — 知識に基づく音素セグメンテーション・サーバ

が使用できる。これらのサーバは、ワークベンチ *lwb* の他に *ALgetAutoLabel* プログラムで利用できる。また *HMM-Seg-Server* は、*hmmseg_client* プログラムでも利用できる。

5.1 サーバのための設定

- サーバを起動するマシンの */etc/services* に次のエントリを登録する。

```
artimsrv      8001/tcp      # SRK Segmentation Server
HmmsSeg      8002/tcp      # HMM Segmentation Server
```

5.2 サーバの起動 / 終了

(5.2.1) HMM セグメンテーション・サーバの起動 / 終了

HMM-Seg-Server プログラムである *hmmseg_server* を起動する。

————— *HMM-Seg-Server* の起動 —————

```
% rlogin atrq02
% cd      /usr/local/bin/AL
% setenv ALHmmsConvTbl ./phon25.conv
% hmmseg_server &                                # HMM-Seg-Server 起動
```

————— *HMM-Seg-Server* の終了 —————

```
% rlogin atrq02
% ps -auxw | grep hmmseg_server      でプロセス ID を得る
% kill -9 プロセス ID
```

(5.2.2) SRK セグメンテーション・サーバの起動 / 終了

SRK-Seg-Server は、*HMM-Seg-Server* を使用するので、まず *HMM-Seg-Server* を起動する。次に、環境変数 *ALHmmsSegServer* に *HMM-Seg-Server* を起動したマシンの *hostname* (または IP address) をセットして、*SRK-Seg-Server* プログラムである *ALserver* を起動する。

HMM-Seg-Server, SRK-Seg-Server の起動

```
% rlogin atrq02
% cd /usr/local/bin/AL
% setenv ALHmConvTbl ./phon25.conv
% hmmseg_server & # HMM-Seg-Server 起動
% setenv ALHmSegServer atrq02
% ALserver & # SRK-Seg-Server 起動
```

HMM-Seg-Server, SRK-Seg-Server の終了

```
% rlogin atrq02
% ps -auxw | grep hmmseg_server # プロセス ID を得る
% kill -9 プロセス ID
% ps -auxw | grep ALserver # プロセス ID を得る
% kill -9 プロセス ID
```

5.3 HMM セグメンテーション・サーバのみを使用する

HMM-Seg-Server のみを使用するには、環境変数 LABELLING_MODE に HmmSeg を設定する。また、環境変数 HmmSegServer に、HMM-Seg-Server を起動したマシンの hostname、または IP address を設定する。

HMM-Seg-Server の使用例 1

```
% setenv LABELLING_MODE HmmSeg
% setenv HmmSegServer atrq02
% lwb /set1/MAU/WAV/DSC/MAU_SC2_12.12K
```

HMM-Seg-Server の使用例 2

```
% setenv LABELLING_MODE HmmSeg
% setenv HmmSegServer atrq02
% ALgetAutoLabel MAU_B_0001.12K PAUikioiPAU > /tmp/auto.labels
```

HMM-Seg-Server の使用例 3

```
% setenv LABELLING_MODE HmmSeg
% setenv HmmSegServer atrq02
% hmmseg_client -Sattrq02 -b300.0 -e1100.0 -r"ikioi" MAU_B_0001.12K
```

5.4 知識に基づく音素セグメンテーション・サーバを使用する

SRK-Seg-Server — 知識に基づく音素セグメンテーション・サーバ — を使用するには、環境変数 `LABELLING_MODE` に `SrkSeg` を設定する。また、環境変数 `SrkSegServer` に、*SRK-Seg-Server* を起動したマシンの `hostname`、または IP address を設定する。

SRK-Seg-Server の使用例 1

```
% setenv LABELLING_MODE      SrkSeg
% setenv SrkSegServer        atrq02
% lwb      /set1/MAU/WAV/DSC/MAU_SC2_12.12K
```

SRK-Seg-Server の使用例 2

```
% setenv LABELLING_MODE      SrkSeg
% setenv SrkSegServer        atrq02
% ALgetAutoLabel MAU_B_0001.12K ikioi > /tmp/auto.labels
```

6 インストールの手順

マシンにインストールされている (はずの) X Window System version 11 の release 4 または release 5 を利用して、ラベリングワークベンチ・システムをインストールする。また、以下のマシンでの動作は確認済みである。

machine	O S
DECstation	Ultrix 4.2a, 4.1, 4.0
DECalpha	OSF/1
VAXstation	Ultrix-32 V3.1
HP9000/730, 750, 800	HP-UX
SUN Sparc	SUN OS 4.1

注: DECstation では、Ultrix 4.2 より古い Ultrix (4.1, 3.1 など) に附属の X サーバは古い (X v11, release 3 対応らしい) ので、OS を Ultrix 4.2 にバージョンアップしたほうが良いでしょう。

[0] customize

```
[ ] modify ./LWB.cf      (LATTOPDIR, BINDIR, LWBDIRなどを修正する)
    LATTOPDIR      = top directory of LWB source tree
    BINDIR          = install destination
    LWBDIR          = install destination
```

```
[ ] If necessary, modify ./lib/AL/AL.<machine>.cf
    [ DECstation Alpha ]      ./lib/AL/AL.alpha.cf
    [ DECstation 3000 or 5000 ] ./lib/AL/AL.ultrix.cf
    [ VAXstation ]           ./lib/AL/AL.ultrix.cf
    [ HP-UX ]                ./lib/AL/AL.hpux.cf
    [ SUN Sparc ]            ./lib/AL/AL.sun.cf
    [ Other ]                ./lib/AL/AL.generic.cf
```

[1] make by using installed X11 System

```
% xmkmf (or) /usr/bin/X11/xmkmf (or) /usr/local/X11R5/bin/xmkmf
% make World |& tee make.World.log.
```

[3] install

```
% make install |& tee make.install.log.
```

[4] examples

```
% cd ./examples/specview
% ./specview -x0.6 -y0.6 MAU_B_0002.12K

% cd ./examples/lwb
% ./lwb MAU_B_0002.12K
% ./lwb -i.lwbrc MAU_B_0002.12K
% ./lwb -i.lwbrc-demo MAU_B_0002.12K
```


6.1 ソースファイルの構成

ソースファイルのディレクトリ構造は、次のとおりである。

```
LWB
|
|-- doc
|   |-- lbw
|       |-- fig
|
|-- examples
|   |-- ALtest
|   |-- lbw
|       |-- .LBL
|       |-- .wave
|       |-- .waveDD
|   |-- rootspec
|   |-- specview
|
|-- include
|   |-- AL
|
|-- lib
|   |-- AL
|       |-- X11R5
|           |-- X11
|   |-- SP
|
|-- programs
    |-- bunkail
    |-- hmmseg
    |   |-- 5_MALESxB216+DSABC.phon25.3m.30i
    |   |-- 5_MALESxB216+W1048.phon25.3m.30i
    |   |-- 8_MALESxB216.phon25.3m.30i
    |   |-- Models -> 5_MALESxB216+DSABC.phon25.3m.30i
    |   |-- Models64_for_demo
    |
    |-- lab2wdP
    |-- mklist
    |-- rep
    |-- serverANDclient
    |-- tdcmmhm1-1
    |-- work
        |-- .wave -> ../../examples/lwb/.wave
        |-- .waveDD -> ../../examples/lwb/.waveDD
        |-- ALserver -> /NFS/atrq02/q02/users/artim/bin/ALserver
        |-- Models -> ../hmmseg/Models
```

参考文献

- [1] 村井, 井上, 砂原, “プロフェッショナル UNIX,” アスキー出版, pp.220-221 (1986).
- [2] 藤原, 杉山, “HMMとエキスパートシステムの手法を用いた音素セグメンテーションシステムとラベリングワークベンチ,” ATR Technical Report(非公開), TR-I-0305 (1993-3).
- [3] 藤原, 杉山, “HMMとエキスパートシステムの手法を用いた音素ラベリングワークベンチ,” 信学技報, SP92(132), pp.73-80 (1993-1).
- [4] S. Fujiwara, Y. Komori, and M. Sugiyama, “A Phoneme Labelling Workbench using HMM and Spectrogram Reading Knowledge,” Proc. of ICSLP 92, pp.791-794 (1992-10).
- [5] 藤原, 小森, 杉山, “音素ラベリングワークベンチ,” 音響学会講論集, pp.227-228 (1992-10).
- [6] 嵯峨山, 板倉, “音声の動的尺度に含まれる個人性情報,” 音響学会講論集, pp.589-590 (1979-6).
- [7] 丸山, 村田, 川端, “高精度スペクトログラム表示ルーチン,” ATR Technical Report(非公開), TR-I-0021 (1988-2).
- [8] 武田, 匂坂, 片桐, 桑原, “研究用日本語音声データベースの構築,” 音響学会誌, 44(10), pp.747-754 (1988-10).

付録 A 設定ファイルの例

```
#
#
#       LWB customize file
#
#
MainWindowX:           0
MainWindowY:           0
MainWindowWidth:       0.7
MainWindowHeight:      0.7
MainWindowBackground:  #cb7955
SubWindowBackground:  #d2ad8c
PopupMenuBackground:  #f6e7c2
LabelLayerHighlight:  #f7af79
BDLineColor:          #606060
BDRemarkColor:         red
BDLineWidth:          2
#
MenuWindow:            0.040
WavePositionWindow:    0.040
WaveWindow:            0.115
PowerWindow:           0.05
DeltaCepWindow:        0.05
AutoLabelWindow:       0.042
SpecWindow:            0.345
#ColorSpecWindow:      0.345
TimeScaleWindow:       0.05
Label1Window:          0.04
Label2Window:          0.04
Label3Window:          0.04
Label4Window:          0.04
Label5Window:          0.04
WindowSeparate:        5      [pixel]
#
DisplayTimeRange:      1200    [ms]
DisplayScrollTime:     1000    [ms]
DisplayWaveRange:      ( -10000 10000 )
DisplayPowerRange:     ( 0.0  80.0 )
DisplayDeltaCepRange:  ( 0.0  6.0 )
DisplaySpecFreqRange:  6000    [Hz]
DisplaySpecRange:      ( 17.0 95.0 )
DisplaySpecCoeff:      -10.0
#
# for 12KHz sampling
#
AnaSamplingFreq:       12      [kHz]
AnaFrameWidth:         120     [sample]
AnaFrameShift:         30      [sample]
AnaFFTLenght:          256     [sample]
AnaLPCOrder:           16
```

```

AnaCepOrder:          16
AnaDeltaCepDelta:     8      [frame]

#
# for 16KHz sampling
#
#AnaSamplingFreq:      16      [kHz]
#AnaFrameWidth:        160     [sample]
#AnaFrameShift:        40      [sample]
#AnaFFTLenght:         512     [sample]
#AnaLPCOrder:          16
#AnaCepOrder:          16
#AnaDeltaCepDelta:     8      [frame]

#
# for 20KHz sampling
#
#AnaSamplingFreq:      20      [kHz]
#AnaFrameWidth:        200     [sample]
#AnaFrameShift:        50      [sample]
#AnaFFTLenght:         256     [sample]
#AnaLPCOrder:          17
#AnaCepOrder:          17
#AnaDeltaCepDelta:     8      [frame]

#
# for 48KHz sampling
#
#AnaSamplingFreq:      48      [kHz]
#AnaFrameWidth:        480     [sample]
#AnaFrameShift:        120     [sample]
#AnaFFTLenght:         1024    [sample]
#DisplaySpecFreqRange: 10000   [Hz]
#
#

LabelMenu1: AUTO \\ a i u e o N Q pau , \\ k s sh ch t n h f m y r w g z j d b p \\ ky ty ny hy my ry gy dy by py
LabelMenu2: AUTO \\ a i u e o N Q pau , \\ < > cl *cl tr dj mm \\ k s sh ch t n h f m y r w g z j d b p
LabelMenu4:      a i u e o N Q pau , \\ < > cl *cl tr dj mm \\ k s sh ch t n h f m y r w g z j d b p
LabelMenu5: a i u e o
LabelMenu3: fr dv

#
MenuSpeechIn:          SpeechIn
MenuLoadWave:          WAVE
MenuLoadLabel:         LABEL
MenuPreviousPage:      Prev.Page
MenuNextPage:          Next.Page
MenuSaveLabel:         Save
MenuHelp:              Help
MenuQuit:              Quit
MenuVoice:             Voice
MenuPrint:             Print
#

```

```
#
FontNormal:      -adobe-times-bold-r-normal--18-180-75-75-p-99-iso8859-1
FontMenu:        -adobe-helvetica-bold-o-normal--18-180-75-75-p-104-iso8859-1
FontAutoLabel:   -adobe-times-bold-r-normal--18-180-75-75-p-99-iso8859-1
FontAutoLabel:   -adobe-times-bold-r-normal--24-240-75-75-p-132-iso8859-1
FontLabel:       -adobe-times-bold-r-normal--18-180-75-75-p-99-iso8859-1
FontLabelSmall:  -adobe-times-bold-r-normal--14-140-75-75-p-77-iso8859-1
FontLabelTiny:   -adobe-times-bold-r-normal--12-120-75-75-p-67-iso8859-1
FontTimeScale:   -adobe-helvetica-medium-r-normal--18-180-75-75-p-98-iso8859-1
FontWaveList:    -adobe-courier-bold-r-normal--14-140-75-75-m-90-iso8859-1
FontLabelMenu:   -adobe-times-bold-r-normal--18-180-75-75-p-99-iso8859-1
#
ExtensionLabelFile: .LB
VoiceCommand:      daout
#
```

索引

.file, 4, 9
.lwbdemo, 4
.lwbrc, 4
.speechin, 9
/etc/services, 2, 4, 18
 artimsrv, 4, 18
 HmmSeg, 4, 18

A

ALgetAutoLabel, 13, 18
ALserver, 18

D

daout, 2, 6, 9
DASBOX, 2
datasize, 3

H

hmmseg_client, 18, 19
hmmseg_server, 18

L

limit, 3
lwb, 4

M

make install, 21
make World, 21
maxdsiz, 2
maxssiz, 2

P

phon25.conv, 18
pstat, 2

S

specview, 21
SpeechIn, 2, 9
stacksize, 3
swap space size, 2

X

xmkmf, 21

ゐ

異音化層, 11

イベント層, 11
インストール, 21

お

音声記号層, 11
音声出力, 6, 9
音声入力, 9
音素境界の移動, 11
音素境界の削除, 11
音素境界の入力, 11
音素ラベル記号の入力, 8

か

環境変数

ALHmmConvTbl, 18
ALHmmSegServer, 4
ALSrkSegServer, 4
AUTO_LABELS, 4
LABELLING_MODE, 4

き

キーボード入力, 11

じ

時刻表示, 8

ぼ

母音中心層, 11

ゆ

融合化層, 11