

第9回コーパス利用講習会

Praat・ELAN講習会

2019年8月31日







000	Praat Objects		• •	•	1	· ·		Praat	Picture			
New Open Save		Help	File	Edit	Margins	World	Select	Pen	Font			Help
New Open Save Objects:	Praat Objects	Help	- 1 - 1 - 2 - 3 -	Edit	Margins 1	World 2	Select	Praat Pen 3	Picture Font	4	ŝ	Help 6
Rename Copy Inspect Info Remove			-4 -5 - - - - - - 7 - 7									

主要な操作対象 ➡ オブジェクトウィンドウ

ピクチャーウィンドウ



$[Praat] \rightarrow [Preferences]$

National Institute for Japanese Language and Linguistics

Praat Edit Wi	ndow	Help		
About Praat		Praat Objects		
New Praat script Open Praat scrip	t ot			
Goodies	•			
Preferences	>	Buttons		
Technical	•	Text reading preferences		
Quit	ЖQ	Text writing preferences CJK font style preferences		UTF-8 を選択
		Sound recording preferences	😑 🔘 🔘 Te	ext reading preferences
		Sound playing preferences	Encoding of 8-bit	text files: 💿 UTF-8
		LongSound preferences		O try UTF-8, then ISO Latin-1
				O ISO Latin-1
				O try UTF-8, then Windows Latin-1
				O Windows Latin-1
				O try UTF-8, then MacRoman
				MacRoman
≈ NIN	IJAL	,	Help Standards	Cancel Apply OK

音声ファイル・TextGridファイルを読み込む

オブジェクトウィンドウから、[open] → [Read from file...] で、 音声・TextGridファイルを選択 (音声ファイルは Open long sound file でも可)



※今回の講習会では以下のファイルを読み込む



C001_002_IC01.wav IC01(玲子)の音声ファイル C001_002-luu.TextGrid 発話単位で区切ったTextGrid







音声ファイル・TextGridファイルの表示



TextGridにおける2種類の層

Interval 層(tier):開始・終了時刻をもつ範囲
 (日常会話コーパスの転記はInterval tierのみ)

ा ज 1	おはようご p ざいます。 z	(L)	pz	あー。	喉がらがら。	pz	う ん ^(168/843)
2	(D イッカ)(0.295)おは ようござい。	(L)	p おはよ z う。		pz tə	.:., p	超 r 痛 (615)

2 Point 層(tier):1つの時刻のみ(範囲ではなく点)





範囲あるいはポイントの指定方法①



波形上でクリックしてポイント指定



Interval 層の該当ラベルのクリックで指定



Interval 層の該当ラベルの境界 or Point 層のラベルのクリックで指定



音声の再生



以下のいずれかの方法で範囲指定し ①a 聞きたいラベルをクリック ①b 波形上のドラッグ&ドロップ

National Institute for Japanese Language and Linguistics

② 画面下の一番上のバーをクリック(or タブキー)で再生 (1段目:選択範囲、2段目:表示画面の範囲、3段目:ファイル全体の再生)

※ ラベルや層を移動する場合は、マウス操作のほか、 Alt/Option + 矢印($\leftarrow \rightarrow \uparrow \downarrow$)でも可

範囲の拡大・縮小など



all in out set bak National Institute for Japanese Language and Linguistics

アノテーション内容の修正



National Institute for Japanese Language and Linguistics

アノテーション境界(始端・終端)の移動・削除

②a 移動の場合:ドラッグ&ドロップで任意の位置に移動 ②b 削除の場合:Alt / Option + backspace(削除した境界前後の範囲が結合)



① 修正したいラベルの境界をクリックで指定 ※選択した境界が赤くなっていることを確認









0 Hz	楽しそ う。	pz	なんか れとか いの。	あな(F あの す)) 今度出る やつの	(W (D チ うかな	-ラ) チラシ)チラシってゆ んか要綱みたいのって。	(G a 英語	まあ ま) 吾だろう †ど。	pz	UMUSA AN	75 Hz IC01 _玲 子 (186/843)
2	pz		pz	何。	pz		えー。	あ 。	ウエブない。	しか う 。 ん	日本の会が	IC02 _美 沙 (615)
			3.78542	26		-	5.184574					[
	Ĩ	Nation	INJ/		e and Linguistics							





Point層へのラベルの挿入



② 挿入したい層の上部の ○ をクリックするとラベル挿入用のマーク 1 が表示



Interval層へのラベルの挿入



ラベルの境界位置の同期

① 新規作成:「Point層へのラベルの挿入」①bに記した方法で作成 (Interval層の場合も範囲の始端と終端をこの方法で指定)

② 他のラベルに同期させるように境界・ポイントを移動



1層目と3層目が同期していない状態 (1層目の始端を選択すると3層目始端 の上部に⁰があることが非同期の印)

③他のラベルと同期させたまま同時に移動: 移動させたい境界を選択し、 Shift キーを押したまま ドラッグ&ドロップで移動させると、 同期している境界が一緒に移動



上部の⁰に向けてラベルの境界を ドラッグ&ドロップで移動させると同期 (上部に⁰がないことを確認)



ピッチの設定画面





00	Pitch s	ettings					
	Pitch range (Hz):	100.0	450.0		🛑 個	人の声の高	さに
	Unit:	Hertz		0	合	わせて調整	
The autocorrela	tion method optimizes f	or intonation re	search;				
and the cross-c	orrelation method optim	izes for voice r	esearch:				
	Analysis method:	autocorrelatio	n				
		Cross-correla	tion				
	Drawing method:	speckles		0			
(all of your "adv	anced settings" have th	eir standard val	ues)				
(your "time step	strategy" has <mark>i</mark> ts standa	ard value: auton	natic)				
Help	Standards	Cancel	Apply	ок			





 File
 Edit
 Search
 Convert
 Font

 376.2919019648211 Hz
 (interpolated pitch at CURSOR)

FO値を取得したい箇所をマウスでクリックすると① 画面右端にFO値が表示② あるいは、[Pitch] → [Get Pitch] ③ とすると、上の図のようにFO値が細かく表示④





当該範囲の最大値や最小値も [Get minimum/maximum pitch] で表示



National Institute for Japanese Language and Linguistics

.....

Query -

Save as semicolon-separated file..

音声ファイルの追加表示

追加で音声ファイルを読み込み [View & Edit]で波形表示



C001_002_IC02.wav

IC01(玲子)の会話相手の IC02(美沙)の音声ファイル を追加で開く



 ② 右下の Group がチェックされていると 最初に開いた Window と同期して表示







複数のTextGridを読み込み選択した上で [Merge]をクリックすると



Mergeされたファイル

Objects:

- 1. Sound C001_002_IC01
- 2. TextGrid C001_002-luu
- 3. TextGrid C001_002-transUnit

5. TextGrid merged



[Save] → [Save as text file] で textgridとして保存 (自動で拡張子は TextGrid となる)

New	Open	Save
Ohior	ater	Save as text file
Jojec	SIS:	Save as short text file
1. Sou	und COO	Save as binary file
2. Tex	tGrid CO	
3. Tex	tGrid CO	Save as chronological text file
5 Tov	tGrid m	arged







オブジェクトウィンドウズ上で音声ファイルを選択し 新規TextGrid作成 右のメニューの [Annotate -] → [To TextGrid] を選択 Praat Objects New Open Save Help Objects: Sound help View & Edit Play Draw -Query -Modify -Annotate -Annotation tutorial Anal To TextGrid... Ana To TextGrid (silences)... ①全ての層の名前を記入 Tomononym Manipulate -Convert -Sound: To TextGrid word seg break tone prm AP All tier names: Which of these are point tiers? break tone prm Standards OK Help Cancel Apply Rename... Copy... Inspect Info ② 上で指定した層のうち Remove Point層とするものの名前を記入して [OK] National Institute for Japanese Language and Linguistics





ELANで転記ファイルを開く

※今回の講習会では以下のファイルを使用

C001_002-luu.eaf 発話単位で区切ったELAN用のファイル(拡張子:eaf)

- 方法1) eaf という拡張子のファイル(アイコンが 🎾)をダブルクリックして開く
- 方法2) 上の方法でうまく開かない場合は、ELAN を起動した上で、 下図のように[ファイル] → [開く] で 対象の eaf ファイルを選択

	ELA	AN 5.4				
ファイル 編集 注釈 注釈層	言語タイプ	検索	表示	オプション	ウィンドウ	ヘルプ
新規作成	жN					
開く	жо					
最近使ったファイルを開く	•					
閉じる	ЖW					
保存	ЖS					
名前をつけて保存	企業S					
テンプレートとして保存	℃企業S					
選択部分を.eafファイルとして	保存					
Save a Copy as EAF 2.7						
ELANファイルの統合						
自動バックアップ	•					
ページ設定	企業P					
印刷プレビュー	∖сжр					
印刷	ЖP					
Multiple File Processing	•					
別ファイル形式で保存	•					
複数ファイルを別ファイル形式	で保存 ▶					
読み込み	•					
Import Multiple Files As	•					
終了	жQ					





再生する音声の選択や 注釈(転記)の一覧表示など

ファイル 編集 注釈 注釈層 言語タイプ 検索 表示 オプション ウィンドウ ヘルプ 音声認識 メタデータ 同生的な Comments . 再生音量: 25 C001_002_IC01.wav Mute Solo C001_002_IC02.wav Mute Solo C001_002_MIX.mp4 Mute Sola 26 50 25 映像 再生スピード: 100 200 再生コントロール 選択モード 繰り返し再生 音声 00:01:52.000 00.01:54.000 00:01:55.000 00.01 59.000 00.02 00.000 10:01:48.000 00:01:49.000 00.01 50.000 00.01.51.000 00.01 53.000 00:01:56.000 00:01 57.000 00.01 \$8.000 (個人IC) 20:01:48.000 注釈層 00:01:55:000 0 中口にも直せる。 00:01:49.000 00-01-50 000 00.01:51.000 00:01:52.000 00-01-53 000 00-01-54 000 00:01:56.000 00:01:57.000 00:01:58.000 00:01:59.000 00-02 00.000 00.02.01 0 うんうんうん。 31. 万歩計の ふーん ふーん。 すごいな:。 IC01_ 种子 うキロに直せる で: えーと マイル(0.107) これマイルになってるけど ほんとはキロに直したいんだけど キロ。 そうなの。 (各話者の 転記テキスト)

National Institute for Japanese Language and Linguistics



ファイル 編集 注釈 注釈層 言語タイプ 検索 表示 オプション ウィンドウ ヘルプ

	THE		-	< 音声認識	メタデータ	再生調整	
		0:0:33.555 ■ -M ▶ P+ ▶F ▶1 ▶I ▶I	※許: 00:00:33.39	再生音量: 100 0 C001_002_1 Mute C001_002_1 ✓ Mute C001_002_1 ✓ Mute C001_002_1 ✓ Mute C001_002_1 ✓ Mute C001_002_1 ✓ Mute C001_002_1 ✓ Mute C001_002_1	50 C01.wav Solo 0 25 C02.wav Solo 0 25 C02.wav Solo 0 25 MIX.mp4 Solo 0 25 C04.wav C04.wav	・・・・・・・ 5 50 75 5 50 75 5 50 75 5 50 75 5 50 75 6 75	このバーの付近(どの位置でもOK) でアイコンが変わったら クリックしながら上下に動かすことで 上部(映像など)と下部(波形・注層) の大きさが変わる
 C001_002 ᅌ	00:00:36.000	00:00:36.500	00:00:37.000	00:00:37.500	00:00	0:38.000	
▲▼ IC01_玲子	00:00:36.000	00:00:36.500 ガそうなんだ。	.• 00:00:37.000	00:00:37.500 うんうん。	00:00	0:38.000	
IC02_美注 [53	少 (D ケ)。 4]	4	うん。	なんかね で これを自	x車のほうにワイヤL		同様に、このバーの付近で アイコンが変わったら クリックしながら上下に動かすことで

波形と注釈層の大きさが変わる



再生(1)

ファイル 編集 注釈 注釈層 言語タイプ 検索 表示 オプション ウィンドウ ヘルプ



再生したい箇所をクリックで選択し 再生コントロールで再生ボタン ▶ 停止ボタン ■

※ Control + Space で再生・停止 NINJAL National Institute for Japanese Language and Linguistics

ファイル 編集 注釈 注釈層 言語タイプ 検索 表示 オプション ウィンドウ ヘルプ





再生したい発話区間をドラッグで選択するか、 注釈層の部分をクリックして範囲を選択した上で、 選択部分再生ボタン♪ 5 で再生♪ 5 で停止

※ Shift + Space で選択再生 停止



その他の再生コントロール

-▶+ 8 M 1 F > F 1 M ÞS \rightarrow \rightarrow $\mathbf{1}$ T. ←

	現在位置から1ピクセル前後に移動	Cntr + Shift+ ←/→	Shift + \Re + \leftarrow / \rightarrow
F◀	現在位置から1フレーム前後に移動	Cntr + ←/→	Ж+←/→
1	現在位置から1秒前後に移動	Shift+ ←/→	Shift + ←/→
	現在位置から1画面前後に移動	Cntr + Fn + \uparrow/\downarrow	光 + Fn + 个/↓
	会話の先頭・末尾に移動	Cntr+b/e	
► →	選択した範囲の先頭・末尾に移動	Cntr + / (backslash)	
$\leftarrow \rightarrow \downarrow \uparrow$	注釈層における上下左右の移動	Alt + $\leftarrow / \rightarrow / \uparrow / \downarrow$	Opt + $\leftarrow / \rightarrow / \uparrow / \downarrow$



ショートカットの確認



	ショートカット	
General 注釈モード	メディアの同期化画面	Transcription Mode
ショートカット	内容	
メディアのナビゲーション		
^Space	再生 / メディアを一時停」	<u>F</u>
ት∺	1ピクセル前へ戻る	
ዮ∺→	1ピクセル次へ進む	
∺ ←	1フレーム前へ戻る	
ж→	次のフレームへ	
ራ←	1秒前へ戻る	
合→	1 秒次へ進む	
жв	メディアの開始位置へ	
жE	メディアの終了位置へ	
жG	時間を指定して移動	
ſ	印刷 Print All	



再生する音声の調整





波形表示する音声メディアの選択・拡大縮小









波形部分で右クリックするとメニュー (上の図:波形の縦方向の拡大・縮小選択)



ファイル 編集 注釈 注釈層 言語タイプ 検索 表示 オプション ウィンドウ ヘルプ



映像音声ファイル追加



表示する映像の選択・映像の表示切り離し

[表示] → [メディアプレイーヤー] で表示する映像を選択



複数映像が表示されている場合、いずれかの映像上で右クリックでメニュー表示し [ビデオ画面を切り離す]を選択することえ、その映像を独立したウィンドウで表示





			層(話	诸)ごと
	①[格子]	を選択	② プル ┃ 層(√ダウンで 話者)を選択
	格子 テロップ Lexicon Co	mments 音声認識	メ・データ ▶	
	1002_美沙			
> 梢	号 注釈	開始時間 終了時間	間隔	
	61 あ 。	00:02:57.183 00:02:5	7.407 00:00:00.224	
	62 取れる 取れる。	00:02:57.407 00:02:5	8.223 00:00:00.816	
	63 うん。	00:02:59.509 00:02:5	9.697 00:00:00.188	
	64 なんかね。	00:02:59.697 00:03:0	0.295 00:00:00.598	
	65 何で 計ってんのかね:。	00:03:03.312 00:03:0	5.533 00:00:02.221	
	66 取れるんだねって思うよね。	00:03:08.227 00:03:0	9.805 00:00:01.578	
-	67 なんかでやっぱり体に密着させな	00:03:09.805 00:03:1	2.902 00:00:03.097	
	68 なんか(0.389)で(0.112)たぶん(0.36	00:03:14.672 00:03:1	9.929 00:00:05.257	
	69 それが受信機で(0.883)(D ハシ) (U	00:03:20.440 00:03:2	3.813 00:00:03.373	
	70 データ取れる感じ。	00:03:25.260 00:03:2	6.638 00:00:01.378	
	71 なんか心拍トレーニングとかなん	00:03:26.811 00:03:3	1.420 00:00:04.609	<u> </u>
	72 なん(0.16)(K ヒャ%ク 百)三十まで	00:03:32.564 00:03:3	9.576 00:00:07.012	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
	73 で(0.373)もうまったく新しい世界	00:03:39.576 00:03:4	4.932 00:00:05.356	
	74 謎。	00:03:46.913 00:03:4	7.462 00:00:00.549	
	75 <i>5</i> h.	00:03:48.650 00:03:4	8.834 00:00:00.184	
	76 何。	00:03:50.096 00:03:5	0.521 00:00:00.425	
	77 z	00:03:52.243 00:03:5	4.521 00:00:02.278	
	78 a .	00:03:54.521 00:03:5	4.811 00:00:00.290	ファイル 編集 注釈 注釈層
	79 ウエブしかない。	00:03:54.811 00:03:5	5.899 00:00:01.088	and the second second
	80 つつん。	00:03:55.899 00:03:5	6.286 00:00:00.387	CA VEAL
	81 日本の会社が主催なんだよ。	00:03:56.286 00:03:5	8.032 00:00:01.746	The Party of the second

層(話者)ごとの注釈(転記)の一覧表示

参考) クリックすると注釈層の該当位置に移動 (逆に注釈層 → 格子の移動も)

1				テキント 格子	テロップ Lexi	con Comments	音声認識	メタデータ	再生調整
(A)	CANO.			- IC01 0-7					
F		26.25		• ICO1 pJ		REAL DURING	10 7 IA		
5-1 - 2			The Contract of Sparly	51 1 10		00.01	58 337 0	0-01-59.459	00-00-01
and the second	100	Comment of the		52 万歩計の自転車/	「ージョンみたいな感し	だよ	00.865 0	0:02:04.049	00:00:03
A STATE OF STREET		- 11 C	and the second	53 5-4.		00.02	04 722 0	0-02-06 428	00:00:01.
	~ 11TH			54 うん。		万歩計の自転車パージ	ヨンみたいな感	じだよね。 10	00:00:00.
and the second		5 10.	Part Charles	55 うん。		00:02	07.677 0	0:02:08.088	00:00:00
No. 1 Marine	-		1 1 1	56 (Dン)はー。	•	00:02	09.340 0	0:02:09.919	00:00:00.
Section The	APRIL OF	1 114	ALANA	57 すごいね。	•	00:02	10.449 0	0:02:11.190	00:00:00
and the second		The lot of	4	58 うん。	A 100 million	00:02	11.667 0	0:02:12.388	00:00:00
	No la		and the second	59 わかんない。		00:02	14.544 0	0:02:15.269	00:00:00
-	00:02:0	0.865	選択: 00:0	2:00.865 - 00:02:04.049 318	4 選択モード	繰り返し再生	b		
		10110-10-11						A 1910 Million & 1910 1910	
M M	-	LANDACH, LINE, BUILDAND, A							
C001_002	I 00:01:58.000	00:01:59.000	00:02:00.000 00:0	2:01.000 00:02:02.000	00:02:03.000	0:02:04.000 00:0	2:05.000 0	0:02:06.000	00:02:07.0
	00:01:58.000	00:01:59.000 0	00:02:00.000 00:0	2:01.000 00:02:02.000	00:02:03.000	0:02:04.000 00:0	2:05.000 0	0:02:06.000	00:02:07.0

参考) 格子上での右クリックで メニューを開き、表示項目 や文字サイズなど調整可能





層(話者)を選択



(D イッカ) · おはようござい。 · (L) · おはよう。 · ね:。 · 超痛い。 · え。 · これね。 · あ。 · 最近なんか · (F あの) 自転車用に · ケイデンス 計りたくて回転数とか。 ・で ・で: ・なんかやっぱ九十ぐらいで 九十回転。 ・速く回す:って(W ユ|ゆう)か · 軽く(D #) 軽いギアで(W (D ハヤ)|速い) · (Fあの) (Dイッ)・速い回転数で回すのがいんですって。・(Fあの) 重く 踏まないで。 ・で ・ (D ケ)。 ・うん。 ・なんかね で これを自転車のほう にワイヤレスで ・ えーと: ・ なんだ 回転数を見るのを取り付けて。 ・ で こ れで えーと: ・ブルートゥースだと思うんだけどな。 ・ワイファイか。 ・ ワイファイ ・で ・繋いで。 ・で 二千円ぐらいなの。 ・これ 中国産だか ら。 · で なんか · 使ったことない(U ん) · 使ったことないって(W ユ)ゆ う)か · いまいちよくわかってないから ちょっと使ってみるんだけど。 · こう取り付けはこの間したんだけど。 そうなんです: そうなんです。 ・ なんかね トレーニングにやっぱり。 ・あと 距離どれくらい? ・乗ったか。 (D #)。 · いろいろ。 · (F あの)ね · なんかね これを押すと · COツー とかも出るよ。@CO2 · よくわかんねえ。 · (D タ) 気温も出るの。 · すご くない?。 ・なんか。 ・で ・で いろいろディスタンスとか設定し(U たい しな)。 ・ここら辺がよくわかんないんだよね。 ・まだあたしも。 ・で ・ スキャンし(U た/て)メモリー · (D ター) (G まあ)ま) 時計出る(W (U # #))(U けど))。 · で ここに: (F あの) 自転車を回すと: · 数字が上がってくの。 · で: えーと マイル・これマイルになってるけど。・ほんとはキロに直した いんだけどね。・キロ。・うん。・キロに直せる。・ん。・そうなの。 ああ。 · そう そう。 · まさに まさに そう。 · (U なんかね) · 最近は 心 拍?。 ・ね。 ・なんかさ (W (D スン)|スント) ・ (R 玲)ちゃんスントとかわ かる?。 ·ガーミンとか なんかあんのよ。 · (F あの:) トレーニングする人 たちが。 · うん うん。 · よく · 計測 · 計測って(W ユ|ゆう)かなんか · んー アウトドア 山とか登ったりとか:普通に · 海も山も川(U も) · あ 走っ たり いろいろ · (U その)人たちが 着けてまあ · 心拍をなんかね 心拍計る やつとかここら辺に着けて:。 ・で ・(U ト)レーニングするのに(U 使う)。 うん。・ある ある。・なんだっけ あれ トレッドミルじゃなく%て トレッ ドミル?。 · (L) · トレッドミルだ: · かな?。 · あ。 · 取れる 取れる。 · うん。 · なんかね。 · 何で 計ってんのかね:。 · 取れるんだねって思うよ



参考) クリックすると注釈層の該当位置に移動 (逆に注釈層→テキストの移動も)

0 00-00 +1.000 00 00 46.000

GAUMERERADERVIIIT. TONTA-L



|うん| |なんかね て これを自転車のほうにワイヤレスで

36.000 00.00 24.2 24

14-6

(0 イッカ) おは うござじ。 (L) おはよう。 む:, 眉間り。 え, これな。 の: 最近なんか (F 20) 前転専用に・ウイデンス計りたくて調整数とか。 で で : なんかやつ ばカナくらい 細転。 法く罰すって(W ユ)ゆう)か 毎く(D 3) 軽い中アで(W (D ハヤ)送い) り、うん。なんかねでこれを自転車のほうにワイヤレスで、えーと、存んだ 画転数を見 うの変別付けて、こで これでえーと、ブルートゥースだと思うんだけどな、フイフィイボ。 フイファイ・で 着いで、で 二千円くらいなの、これ 中華産だから、 でなんか 使ったことない(Y_A) ・使ったことないって(W ユ)ゆう)か ・ いまいちよくわかってないから あんだけど。 こう取り付けはこの間したんだけど。 そうなんです: そうなん ちょっと使っ トレーニングにやっぱり。 あと 距離どれくらい? ・乗ったか。 (D #)。 の)ね なんかね これを押すと COツーとかも出るよ。@CO2 よくわかん ろいろ。 ための)ね なんかね これを押すと COツーとかも出るよ。GCU2 よくれの は、 (D 5)気温も出るの。 すごくない?。 なんか。 で でいろいろディスタンスと はた。 (D-7 X MB O D D U) が設定し(D-1) しな)。 ここら辺がよくわかんないんだよね。 まだあたしも。 で スキャ ンし(D た(アメモリー・(D ター) (G まあ(ま)時計出る(W (U # #))((U けど))。 で ここに:(F あの) ンしじた(17天年) ー・(0 ター)(3 年高)(時間)118代(10)(1 年高)(10)(1 だ)、でここに(5 あの) 国際業務によう、数字がよがってくる。 でえーとマイム これでイム)にでなってるださ。 単 んとはそれご思したいんだけでね。 キロ、 ラム。 キロに間様も、 ム、 きうなの、 ある。 ひううく。 まざこをさっ、(10) ムれの)) 素加はなわれ、 ね、 なんかさ (10)(1 クン)(スント) (13)(1 ちゃんスントとかわか)) 素加はなわれ、 なんかさ (10)(1 クン)(2 く)(1 ちょうん)、 カム ふ、 ム、 なく 1首。(1前)(1 ちょうん)(1 ちん)(1 ちん) 人 一 フラトドア山とかきっとりとか 雪風に、 雪心は高川(1 ちょうに)(1 ふうら)(1 しょう) レーニングするのに(U使う)、うん、あるある、なんだっけあれトレッドミルじゃなく% レーングするのに(U使う)、うん、あるある、なんだっけあれトレッドミルじゃなく% レンドミル。(L)・トレッドミルだ、かな?。あ。取れる取れる。うん、な んかね。(何で計ってんのかね)、取れるんだねって思うよね、なんかでやっぱり体に密度 ログル。 「「ビコントのの方法」、取れる人となってあっため。 なんが でやっぽり体に装置 せないといけなく%て。 なんか で たぶん こうゆう腕時計のとかも最近多いけど。 イガダ気増率 (D/3-2)(ビックなの)。 データ取れる感じく なんかふれトレーニングとか かんか これませどまたんだりゃ かん パレーとくが(不)ニュキャトピキー リストキ かんや

Lexicon Comments 日本語識 メタデータ 再生調整

0000 000056.000 000051.000 000052.000 0000 53.000

JA-19-2868 94774 9477



[注釈層]→[新規追加]

			注釈層0	D追加											ž
			注釈層の	の追加				ファイル	編集	注釈	注釈層	言語タイプ	検索	表示	オプション
Eの注釈層									1	11	新規連	自加			жт-
釈層名)1_玲子)2_美沙	上位注釈層 - -	言語タイプ default-lt default-lt	話者	注釈者	既定言語 - -	Content Lang - -					新規注 上122 注釈層 注釈層	³ の性地で更 <mark>釈層の追加(第T)</mark> Eボ層の変史 Mana B M Mana B M M M M M M M M M M M M M M M M M M M	nt		∵жт
		追加	変更 肖	削除 読み込み							注釈層	層内の文字列分	割		
釈層名 者		comment	-	注釈層 を入力	層の名 ^つ	前(ほかの) 追加](こ	項目は任意 こでは c	त्त) omment	とい	う層	を追力	0)			
釈者		koiso		_ ,						•••					
位注釈層		none													
語タイプ		default-It													
定言語		None				0									
ontent Lang 他のオプシ	guage ション	None				0									
へこ回 ontent Lang 他のオプシ	guage ション	None	10 to 1			0			-						



注釈の入力



NINJAL National Institute for Japanese Language and Linguistics

注釈の修正・移動・削除

A) 注釈の内容の修正: 修正したい注釈を選択しダブルクリックすると編集可能となる

00:00:49.000	00:00:50.000	00:00:51.000	00:00:52.000	00:00:53.00	00:00:54.000	00:00:55.000	00:00:56.000
<u>IC01_玲子</u> [468]	うんえーすごい		うん。		ふーん。		
IC02_美沙 ^[363]	ワイファイ(0.64)で(0	.114)繋いで。	で 二千円くら	いなの。	これ 中国産だから。	で なんか(0.264)	使ったことない(U ,

B) 注釈の境界位置の修正

alt/optionを押しながら注釈の境界にポインターをかざすとアイコンが変わるので、 そのまま alt/optionを押したままマウスでドラッグ&ドロップで境界を移動

C) 注釈をそのまま移動

alt/optionを押したまま注釈の中央あたりでドラッグ&ドロップして注釈全体を移動

D) 注釈の削除

削除対象の注釈を選択し [注釈] → [注釈の削除]

※ その他の操作は[注釈] のメニュー参照





A) <u>eaf ファイルとして保存する場合</u> .eaf メインファイル [ファイル] → [保存] .pfsx フォントなどの環境設定





C) <u>会話分析風のトランスクリプション</u>

[ファイル] → [別ファイル形式で保存] → [旧来の記録文書形式で]





[注釈層] → [Create Annotations from Gaps]



Select All で玲子と美沙の二人の層を選択
 算出した沈黙長を gap という新しい層に表示
 沈黙長をミリ秒形式で表示

gapという新しい層ができ、 発話間の沈黙長が表示





[注釈層] → [Create Annotations from Gaps]

elect Tier(s) ICO1 给子	
ICO2_美沙 gap	
Select All Select None	
オプション	
Create annotations on	
● the same tier IC01_玲子 ♀ ● a new tier	
Value for the new annotations	
O Empty	
O Specific value:	
 Duration of the gaps 	
● ミリ秒	
○秒.ミリ秒	
○ 時:分:秒.ミリ秒	

指定した層(玲子)に、 玲子の発話間の沈黙長がミリ秒単位で表示



一つの層(ここでは玲子)を選択
 算出した沈黙長を同じ層(玲子)に表示
 沈黙長をミリ秒単位で表示
 ※同様のことを美沙の層でも行う

National Institute for Japanese Language and Linguistic











① 選択した注釈層における注釈(発話) の頻度や全体に占める割合などの統計

•	•	統計					
<u>}</u>	家 注釈 注釈	积層 Tier Type	Participa	ant Annotator	Language		
主釈層							
注釈層の選択:	IC02_美沙	0					
🗹 Show onl	y root tiers						
Count co	ntiguous annotatio	ons with the same	value as 1				
Use med	ia duration as obse	rvation period					
統計変数							
注釈	Occurrences 🔻	頻度	注釈時間	時間の比率	待ち時間		
うん。	18	0.017170657	0.34411	0.00590861394	36.901		
(L)	16	0.015262806	1.21575	0.01855575693	1 850		
<u><u> </u></u>					1.035		
あ。	13	0.012401030	0.16684	0.00206906419	11.278		
、 あ。 そう。	13 12	0.012401030	0.16684	0.00206906419	11.278 404.916		
あ。 そう。 あっ。	13 12 5	0.012401030 0.011447104 0.004769627	0.16684 0.36858 0.1528	0.00206906419 0.00421921205 7.28799007917	11.278 404.916 269.116		
() あ。 そう。 あっ。 はい。	13 12 5 5	0.012401030 0.011447104 0.004769627 0.004769627	0.16684 0.36858 0.1528 0.3548	0.00206906419 0.00421921205 7.28799007917 0.00169226366	11.278 404.916 269.116 722.129		
() あ。 そう。 あっ。 はい。 うん うん。	13 12 5 5 4	0.012401030 0.011447104 0.004769627 0.004769627 0.003815701	0.16684 0.36858 0.1528 0.3548 0.6455	0.00206906419 0.00421921205 7.28799007917 0.00169226366 0.00246303539	11.278 404.916 269.116 722.129 140.604		

③各注釈層の注釈数などの統計

		•	•	統計					
ž	È釈 注釈	注釈層	Tier T	ype Par	ticipant	Annotator	Langua	ge	
統計変数一									
注釈層	注釈数	最短の注	最長の注	注釈時間	注釈時間	注釈時間	注釈時間	待ち時間	
				APPENDING TO DO THE OWNER	Version and teaching	And the state of t	APPENDING TO DO THE PARTY OF TH	10 - 00 100	
IC01_玲子	468	0.098	8.613	1.069372	0.838	500.466	47.726	0.03	

② 選択した注釈層における注釈の継続 時間の最小・最大・平均値などの統計

				紛	た言十			
(注釈	注釈 II	注釈層	Tier Type	Participant	Annotator	Language	
È釈層								
主釈層の選	択: [IC02_美》	▶ ᅌ	🗹 Show	only root tiers			
計変数								
注釈	Occurr	ences 🔻	最短の注釈	時間間隔	最長の注釈時間間	隔 注釈時間	間隔の平均	
うん。	18		0.184		0.853	0.344		
(L)	16		0.274		3.531	1.216		
あ。	13		0.057		0.29	0.167		
そう。	12		0.224		0.641	0.369		
あっ。	5		0.131		0.184	0.153		
はい。	5		0.301		0.432	0.355		
うんうん。	4		0.522		0.809	0.646		
なんか。	4		0.495		0.555	0.527		
ね。	4		0.208		0.471	0.353		
あー。	3		0.276		0.763	0.598		
(D #)。	2		0.128		0.271	0.199		
ああ	2		0 224		0 201	0.257		





範囲を選択し波形上で右クリックし [Praatで選択部分を開く]を選ぶ



該当範囲の波形がPraat上で表示





PraatとELANの連携② Praatで作成した TextGrid を読み込む

[ファイル] → [読み込み] → [Praat TextGridファイル]





[閲覧] から対象のTextGridを指定 (Encoding は UTF-8)

:oiso/Desktop/C001_002/C001_002-transUnit.TextGrid

閲覧...

✓ポイント層(point tier)注釈期間の初期値(ミリ秒)

40 ⁽⁾ Praatのポント層(point tier)を含む。

──値が入力されていない区間/注釈は省く。

ELANには interval tier 相当の注釈しかないため、 読み込み対象の TextGridに point Tier がある場合 [ポイント層...] にチェックしたうえで 継続長の初期値を設定

(本来、継続長0のポイントに一定の長さを与える)



次へ進んで「言語タイプの選択」

(基本的には何もせず)[終了]

ELANに TextGrid の層が追加で読み込まれる

PraatとELANの連携③ ELANで作成した注釈をTextGrid で書き出す

[ファイル] → [別ファイル形式で保存] → [Praat TextGridファイル形式で] ファイル 編集 注釈 注釈層 言語タイプ 検索 表示 オブション ウィンド 新規作成... ₩N 80 開く.... 最近使ったファイルを開く 閉じる #w 保存 **#**S 名前をつけて保存... **û**#S ℃**企業S** テンプレートとして保存... 選択部分を.eafファイルとして保存... Save a Copy as EAF 2.7... ELANファイルの統合 自動バックアップ ページ設定 **☆#**P 印刷プレビュー **₹**%P EDRI ЖP Multiple File Processing 別ファイル形式で保存 Toolboxファイル(UTF-8)形式で... 複数ファイルを別ファイル形式で保存 FLEx File... 読み込み CHATファイル... Import Multiple Files As タブ区切り文書ファイル形式で.... Tiger-xmlファイル形式で... 終了 #O インターリニアー文書... HTML C001 002... 00:04.000 旧来の記録文書形式で Praat TextGridファイル形式で. 言葉のリスト... SMILファイル形式で Export a OuickTime文書ファイル形式で... Subtitles文書ファイル形式で... 00:04.000 Tiers for Recognizer... IC01_玲子 メディアクリップ... Media Clip using Script... IC02_美沙 ELANのウィンドウ画像... pz Filmstrip Image... IC01 玲子-1 Shoeboxファイル形式で... 1000 羊沙 -----

保存対象の層を選択して [OK]

✓ IC01_用于 ✓ IC02_美沙				
基礎注釈層のみを	·表示 7 Une	do Sort	Soloct All	Salact None
				Select None
オプション				
限定された時間間	隔に制限せう	3		
	時間のオフセ	ットを注釈時間に追	加	
マスターメディア				
マスターメディア	0	OK 閉	6	

保存ファイルを指定し UTF-8 を選択して保存

Praat は ELAN のファイルを読めないため ELAN 側で一旦 TextGrid 形式で書き出す 必要がある

検索①作業中のファイルを単純に検索



検索② 高度な検索: 正規	見表現を使ってみる
[検索] → [複数のeafファイルを詳細条件で検索]	正規表現
(c)	 任意の1文字 [abc] カッコ内の文字のいずれか [a-c] カッコ内で範囲指定された文字のいずれか [^ab] カッコ内の文字以外のいずれか (abc) 文字群の指定("abc"という文字列) (ab[cd) 複数文字群の指定("ab"あるいは"cd") (注釈の)先頭 (注釈の)先耳 (注釈の)末尾 直前の文字(群)が1つ有るか無いか 車前の文字(群)の1回以上の繰り返し * 直前の文字(群)の0回以上の繰り返し [n] 直前の文字(群)のn回以上の繰り返し [n,] 直前の文字(群)のn回以上の繰り返し [n,] 直前の文字(群)のn回以上の繰り返し [あそ]の 注釈に「あの」か「その」を含む (うん」はい) 注釈に「うん」か「はい」を含む ^あ[っ一]?。\$ 「あ。」「あっ。」「あー。」である注釈

検索③高度な検索:複数tier間の条件

[検索] → [複数のeafファイルを詳細条件で検索]

検索例1



検索例2

	Substring Search	Single Layer Search	Multiple Layer Search	
Domain:	1 eaf files		Define	Domain
Query Hi	istory: <	> New (Query Save query	Load query
Mode:	case insensitive	ᅌ 🛛 regular exp	ression ᅌ	Clear
Minir	mal Duration	Maximal Duration	Begin After	End Before
1)a	ta. \$	1 b	Tier Name: IC01_玲子	-
(3 Overlap	-	Must be in same file	-
2 a	^うん	2 b	Tier Name: IC02_美沙	-
Find	Hide query	Fewer Columns	More Columns	Fewer Layers
Found 4	hits in 4 annotations (of	831) R	eady	Cancel
<	> Hit	1 - 4 of 4 Sa	ve hits	
	#1 なんかガー #1	ミンて聞いたことある	#2 うん うん。 うん _	
	#1 塩水た	ビステレストロット #21 ビ(WカンIから)ね。1 #2	21うん うん。1	
	#1 もうね:	十:五分撮ってるね。」	#2 うん うん。1	

「ね。」で終わる①a 玲子の発話①b に対し 「うん」で始まる②a 美沙の発話②b が Overlap ③している場合



「うん」で終わる<u>1</u>a 玲子or美沙の発話<u>1</u>b に対し 「うん」で始まる<u>2</u>a 相手の発話<u>2</u>b が 200ミリ秒以内の間隔で後続<u>3</u>している場合

検索④高度な検索:複数ファイルを検索

検索等の処理をする「ファイルの集合」であるドメインを指定してから検索 ※個々に指定した複数のファイル、あるいは指定フォルダに含まれる複数のファイルの集合

「検索」→「複数のeafファイルを詳細	甲冬件で検索1	検索②③に記	した方	法で検索
		Substring Search	Strigle Lay	er Search Multiple Layer Search
→ [Define Domain] を迭択		Domain: 252 eaf files	ドメイン中	ロのファイル数 ノルが検索対象)
Substring Search Single Layer Search Multiple L	ayer Search	Query History: <	<pre>\25277</pre>	New Query Save query Load query
Domain: 1 eaf files	Define Domain	Mada		
Query History: < > Loaded 18 Queries Sa	ve query Load query	MODE: Annotation	◇検索語	case insensitive of substring match of shares of substring match of
Mode: Annotation Case insensitive	substring match	Find 夏休み		All Tiers 0
Find All Tiers	•		4	
Found 0 hits in 0 annotations (of 0)	Cancel			
< > No hits Save hits		ドメインに名前	うをつけ	ナて [OK] で保存
		Specify a name for the domain	ok to save	保存したドメインは [Define Domain]で ドメインリストから選択可
[New Domain]				
Specify new domain	[Add Folder] でこ	フォルダを指定		
New Domain			マイリ	を指定)
New Domain from IMDI Search			7 1 70	
	メロイバ /Volumes/share/プロジェクト/会話コー	ーパス/CEJC-monitor/data	Ad	ld File
National Institute for Japanese Language and Linguistics	ここでは CEJC-D 全てのデータを	VD版に含まれる 指定	Add	Folder emove

管理語の設定①

管理語(注釈のリスト)を設定し、そのリストから注釈を選択できるようにする



+

+ -

-

管理語の設定②





上位/下位の層の設定①









言語タイプを作り直す

[言語タイプ] → [言語タイプの追加]

目ケッテ語タノブ	Ī	言語タイプの通	自力口			
言語タイプ名	ステレオタイプ	管理語を使う	データカテゴリのI	時間調整		
default-lt	-			*	1 言語ダ (ここて 2 ステレ	タイプ名の指定 では act-type) ・ オタイプの選択 ロータイプ1を選択
	追加 default	変更 削除 -lt	読み込み	0	3 管理語(ここで)	高の選択 は act)
言語タイプ名(act-type				言語タイプ名	act-type
ステレオタイプ	None	de alto de terre	タイプつ		ステレオタイプ	Symbolic Association
管理語を使う 🤅	Include	d In	タイプ3		管理語を使う	act
Lexicon connect	ion Symbol Symbol	ic Subdivision ic Association	タイプ4 タイプ1			
ISOデータカテゴ	J			+ -		
時間調整が可能	\checkmark					

上位/下位の層の設定③

スライド「管理語の設定②」も参考に 新しい注釈層(IC01_行為)を作成





上位層と完全に範囲が一致した 注釈が作成される!

<u>IC01_玲子</u>	おはようございま	₹ す 。
1C02_美沙	(D イッカ)(0.295)	おはようござい。
IC01_行為	挨拶	



テンプレートの保存・適用

設定した「管理語」「言語タイプ」などを保存(拡張子: .etf)

[ファイル] → [テンプレートとして保存]

ファイル	編集	注釈	注釈層	言語タイ
新規作成	t			ЖN
開く				жo
最近使っ	たファ	イルを	開く	•
閉じる				жw
保存				₩S
名前をつ	けて保	存		<mark></mark> ዮ፝፝
テンプレ	ノートと	して保	存	ጚ <mark>፞</mark> ଜିଞ୍ଚ
课报部分	₹ eaf	ファイノ	しとして傷	存



新しい EAF ファイル (例:CO01_002-transUnit.eaf)を開く

nstitute for Japanese Language and Linguistics



「管理語の設定②」の手順に従い 新しい注釈層(IC01_行為)を作成



[閲覧]①で保存したテンプレートを読み込むと 登録してあった言語タイプ act-type が現れる2

言語ダイノ名	ステレオタイプ	管理語を使う	データカテゴリのI	時間調整
default-lt	5	-	-	~
	2			
	追加	変更 削除 🐻	おひみ	

新規ファイル作成

[Add Media File] をクリックしメディアファイルを指定



• • •	ELA	AN 5.7				
ファイル 編集 注釈 注	L 釈層 言語タイプ	検索	表示	オプション	ウィンドウ	ヘルフ
新規作成 開く 最近使ったファイルを開 閉じる	೫N ೫O < ► ೫W					
保存 名前をつけて保存 テンプレートとして保存 選択部分を.eafファイル/ Save a Copy as EAF 2.7	業S 企業S て企業S として保存					

新規作成 選択されたファイル: Add Media File... Add Template File... ストリーミングファイルを追加... Remove

[注釈層] → [新規追加] で任意の注釈層を作成 (「注釈層の追加」参照)



