

TR-IT-0341

## KaNsai ToBI Labelling

Kiyoka Tsuchiyama & Nick Campbell

14th February 2000

### ABSTRACT

Kansai accent differs from the version of Japanese spoken in general use across the country. This technical report proposes a system for labelling the prosodic characteristics of Kansai Japanese in accordance with and extending the J-ToBI Tones and Break Index prosodic labelling system, and presents a comparison of a well-known twenty-minute short story read in the two different accent styles.

関西方言と共通語の、アクセントの比較を目的としている。JAPANESE ToBI に基づいたラベリング方法では、関西方言には、不十分な箇所がある。そのため、どの様な違いがあるのか、同じ文章のテキストで比較をしている。サンプルテキストは共通語を読んだもので、関西方言のアクセントは自然会話的な強いものにはなっていない。

©ATR Interpreting Telecommunications  
Research Laboratory.

©ATR 音声翻訳通信研究所

# KaNsai ToBI Labelling

Kansai accent differs from the version of Japanese spoken in general use across the country. This technical report proposes a system for labelling the prosodic characteristics of Kansai Japanese in accordance with and extending the J-ToBI Tones and Break Index prosodic labeling system, and presents a comparison of a well-known twenty-minute short story read in the two different accent styles.

The read speech is not typical of conversational spoken kansai Japanese, and the text is not that of the Kansai region, but the fact that the reader is from the Kyoto area is clear from listening to the speech. We present the following pages as a first step in the direction of labelling the prosody of Kansai speech, and are aware that the system we currently use may be in need of improvement and revision. We welcome and comments and suggestions.

## Note:

(1) The data that follows is read speech. It is difficult for a Kansai speaker to read non-Kansai text in 'pure' Kansai accent. Conversational speech data is required before a full KaNsai-ToBI can be determined and the difference from J-ToBI can be known.

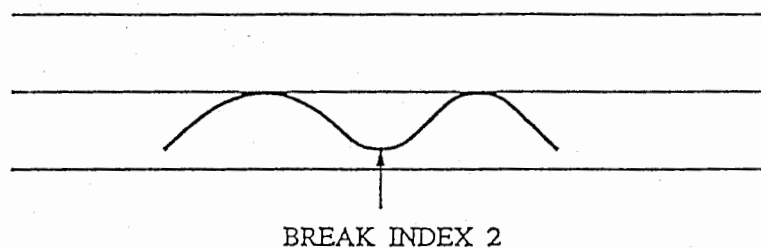
(2) The reader of the 'standard' texts was Kuroyanagi Tetsuko; that of thr Kansai version was Yoko C ita (now Mrs Ken Fujusawa).

## JAPANESE ToBI LABELLING GUIDELINES (English version)

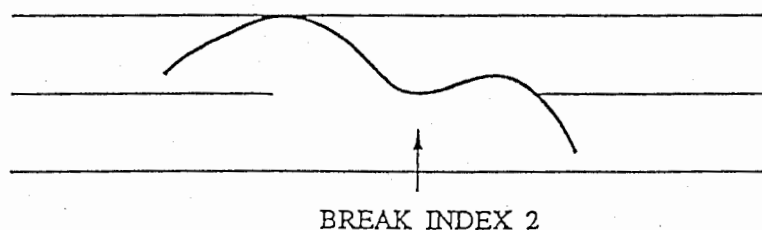
### 1. BREAK INDICES(BI)

- BI1 : word (default specification) boundary ; includes particles
- BI2 : Accentual phrase (AP) boundary

(1) In the case of two nuclear accents occurring one after the other and are of equal height:

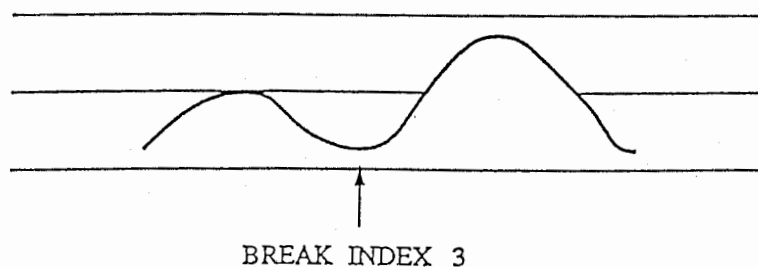


(2) In the case of two nuclear accents occurring one after the other, where the second accent is lower than the first :



- BI3 : Intermediate phrase (IP) boundary

In the case of two nuclear accents occurring one after the other, where the second accent is higher than the first :



- BI4 : Intonation phrase (IP) boundary

The end of the accentual phrase with a pause after it,  
or the end of the sentence

## 2. TONES

- H\*+L

Marks lexical accent

Place at the beginning of the pitch fall.

(In the sample labelling, it doesn't necessarily correspond to the exact beginning of the pitch fall because of the segmentation by mora.)

- H-

Marks an unaccented accentual phrase or the "phrasal high" in an accented phrase, if this can be distinguished from the lexical accent in the F0 pattern.

Place at the end of the initial pitch rise.

---

- %w L

Post-pausal initial boundary tone

Place just before the beginning of the accentual phrase

Use in the case of either of the following :

- (1) there is a syllabic nasal sound on the second mora
- (2) the first accentual phrase begins with a long syllable
- (3) the first mora is accented

- %L

Post-pausal initial boundary tone

Place just before the beginning of the accentual phrase

Use in the case of either of the following :

- (1) unaccented accentual phrases
  - (2) the second mora or the remaining mora is/are accented
- 

- WL%

Weak final boundary tone of accentual phrase

Place at the same interval as break index 2 or 3

Usage : refer to (1), (2), (3) of W%L above

- L%

Final boundary tone of accentual phrase

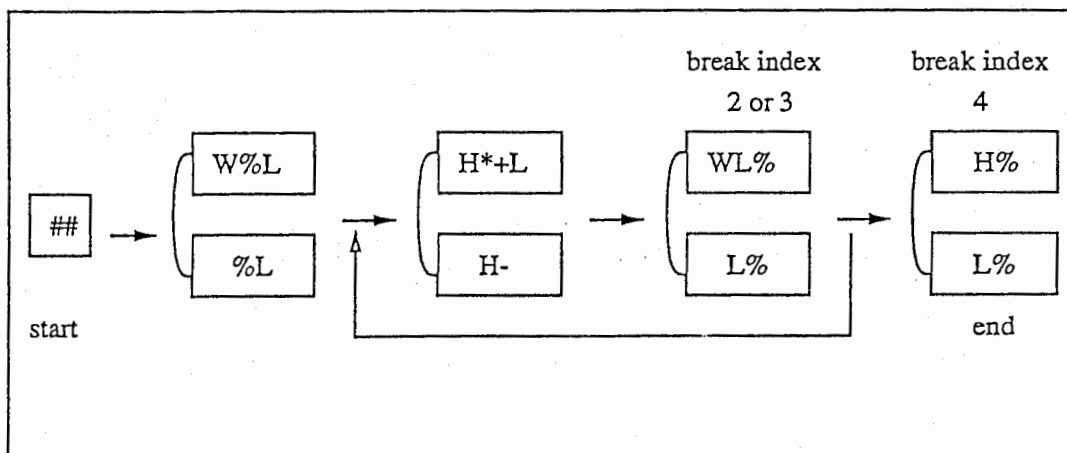
Place just at the break index 2, 3 or 4

Break index 2 or 3 : refer to (1) & (2) mentioned in %L above

Break index 4 : use in the case of no final rise

- H%  
Final boundary tone of intonational phrase  
Place at the same interval as break index 4  
Use in the case of a final rise

- Flow chart of tone labels



## KaNsai ToBI Labels

The symbols which have been newly introduced for Kansai Japanese are as follows:

- H\* + L same as J-ToBI (see following pages)
- H\* + LH Rise-fall-rise on a sequence of three mora
- H- same as J-ToBI (see following pages) but can also be marked on the first mora of a phrase which starts with a high tone.
- L- New. Can be marked on the first mora of a phrase which starts with a low tone.
- %wL same as J-ToBI (see following pages) but is not as syntactically-determined (see examples)
- %L same as J-ToBI (see following pages) but is used in contrast to %wL to indicate an initial low-tone in the following samples. This usage should be replaced by L- in future.
- %wH New. See examples in the following pages. Used in conjunction with a 2- break index when (as in ...de-arū) there is a continuous high tone over a syntactic discontinuity.
- wL% same as J-ToBI (see following pages)
- L% same as J-ToBI (see following pages)
- H% same as J-ToBI (see following pages) but can be used independently, while in J-ToBI it is frequently used in conjunction with L% (see examples).
- wH% Frequently found on phrase-final particles (and elsewhere) and is used to indicate a brief phrase-level rise which is not as strong as a question rise.

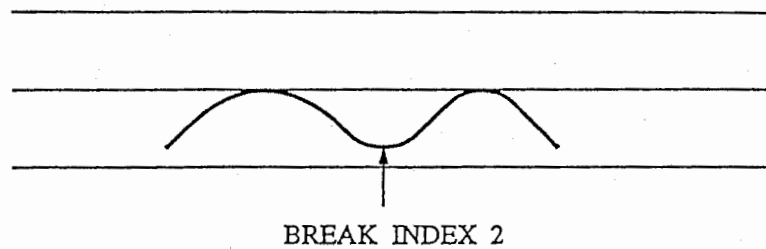
Note:

(1) Three labels ( L- · %wH · H\*+LH ) are currently under consideration and do not appear in the following samples.

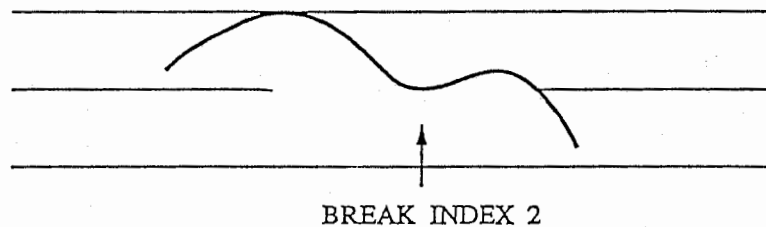
1. BREAK INDEX

- BI 1 : 単語、助詞の境界に置く。
- BI 2 : アクセント句の境界に置く。

(1) 境界前後のアクセント核の F0 が、ほぼ同じ場合 :

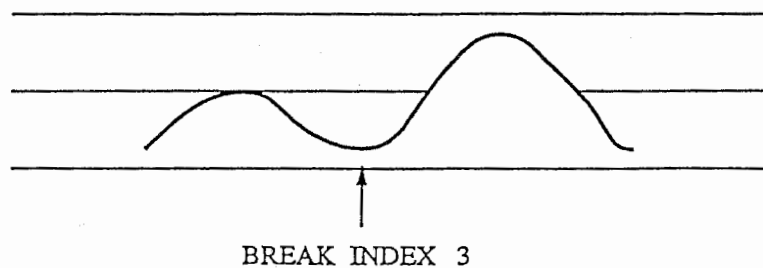


(2) 後ろのアクセント核の F0 方が低い場合 :



- BI 3 : アクセント句の境界に置く。

境界後ろのアクセント核の F0 が、前のアクセント核の F0 より、高い場合 :



- BI 4 : 長いポーズのある文節境界及び、文末に置く。

## 2. TONES

- H\*+L

ピッチの下降始めに置く。

(サンプルラベリングでは、モーラ単位で区切っているため、必ずしも下降始めと一致していない)

「起伏型アクセント句」に用いる。

- H-

ピッチの上昇終りに置く。

「1型、2型以外のアクセント句」に用いる。

---

- %wL

文頭または、BREAK INDEX 4の後の文節頭(「##」上)に置く。

次にくるアクセント句が(1)(2)(3)のいずれかの場合に用いる。

(1) 2 mora 目が「撥音」である

(2) 1 または 2 mora 目が「長母音」を構成する

(3) アクセント型が1型である

- %L

文頭または、BREAK INDEX 4の後の文節頭(「##」上)に置く。

次にくるアクセント句が(1)(2)のいずれかの場合に用いる。

(1) 平板型である

(2) 1型以外の「起伏型アクセント」である

---

- WL%

アクセント句の終り (BREAK INDEX 2 or 3の上) に置く。

使用条件は上記 W%L の(1)(2)(3)と同じ。

- L%

アクセント句の終り (BREAK INDEX 2, 3 or 4の上) に置く。

(BREAK INDEX 2 or 3の場合)

使用条件は上記 %L の(1)(2)と同じ。

(BREAK INDEX 4の場合)

BREAK INDEX 直前が上がり調子でない場合に用いる。

- H%

イントネーション句の終り (BREAK INDEX 4の上) に置く。

BREAK INDEX 直前が上がり調子の場合に用いる。



## KaNsai ToBI

### TONES

---

- ・ H\*+L 語句のアクセントを示す。  
f0 曲線のピーク、下降始めに使用する。
  - ・ H\*+LH 1 mora ずつ上・下・上となっている場合に使用する。
  - ・ H- 高起語の 1 mora に使用する。  
(サンプルラベリングでは、高起語というより、  
高くから始まっている語句に使用している。)
  - ・ L- 低起語の 1 mora に使用する。
- 
- ・ %wL 文頭または、**BREAK INDEX** のアクセント句に使用する。  
次にくるアクセント句の 1 mora に、H\*+L、高起語の  
H- がくる場合に使用する。
  - ・ %L 文頭または、**BREAK INDEX** のアクセント句に使用する。  
次にくるアクセント句の 1 mora に、アクセントがなく、  
低起語の L- がくる場合に使用する。
  - ・ %wH アクセント句で、助詞上がりが後ろの語句につく場合に、  
使用できる可能性がある。  
**BREAK INDEX 2-** の上に置く事が多いと思われる。
-

- ・ wL%      アクセント句の終り (**BREAK INDEX 2 or 3**) に使用する。  
使用条件は、上記 %wL と同じ。  
文末には使用しない。
  
- ・ L%      アクセント句の終り (**BREAK INDEX 2 or 3 or 4**) に使用する。  
使用条件は、上記 %L と同じ。
  
- ・ H%      疑問文などの文末に使用する。  
また、共通語では、**L%H%** として使用していたが、  
2拍の疑問文などで最終 **mora** がその直前の音より  
高い場合には **H%** のみでも使用できる。  
  
例)    まだ?
  
- ・ wH%      アクセント句の最終 **mora** が、その直前の音より高い  
場合に使用する。  
疑問文には、使用しない。

---

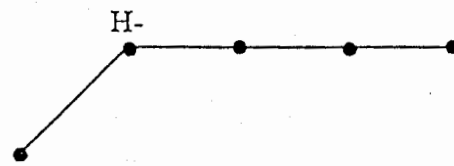
※注      **L · %wH · H\*+LH** は、現在確定していない為、  
サンプルラベリングでは使用していない。

共通語のアクセントも混在しているので、共通語スタイルの  
アクセントの **ToBI** も使用する。

# アクセント型

## 共通語のアクセント型

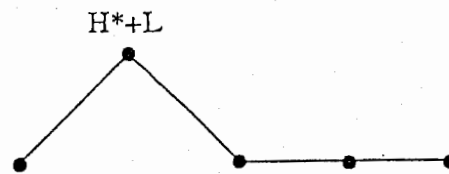
・ 0 型



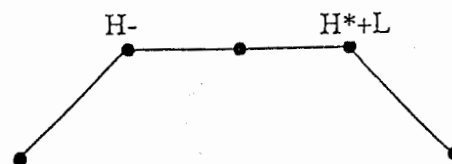
・ 1 型



・ 2 型



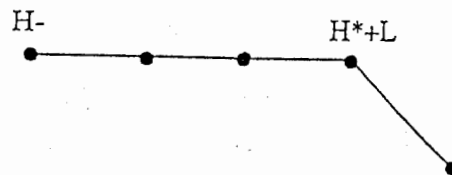
・ 3 型



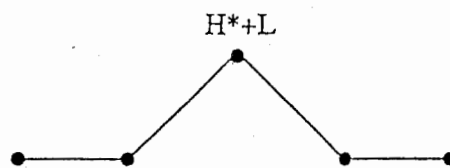
# アクセント型

## 共通語に見られないアクセント型

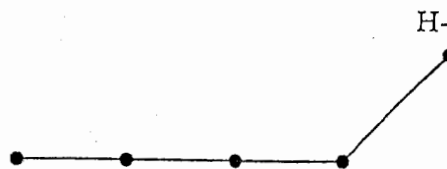
・ i 型



・ ii 型



・ iii 型



・ iv 型



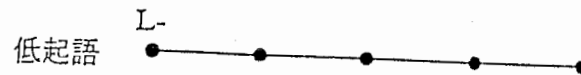
## アクセント型

### 共通語に見られないアクセント型

#### ・v型



#### ・vi型



#### ・vii型



The numbering of types from i to vii is arbitrary and no meaning or naming is intended or implied.

Types vi and vii are not used in the following sample labelling but are perhaps to be preferred (please watch this space). To represent L- in the following pages, we used %L followed by an H- on the first mora.

※注 vi・vii型はサンプルラベリングでは使用していない。  
L-は使用しないで、H-になっている。

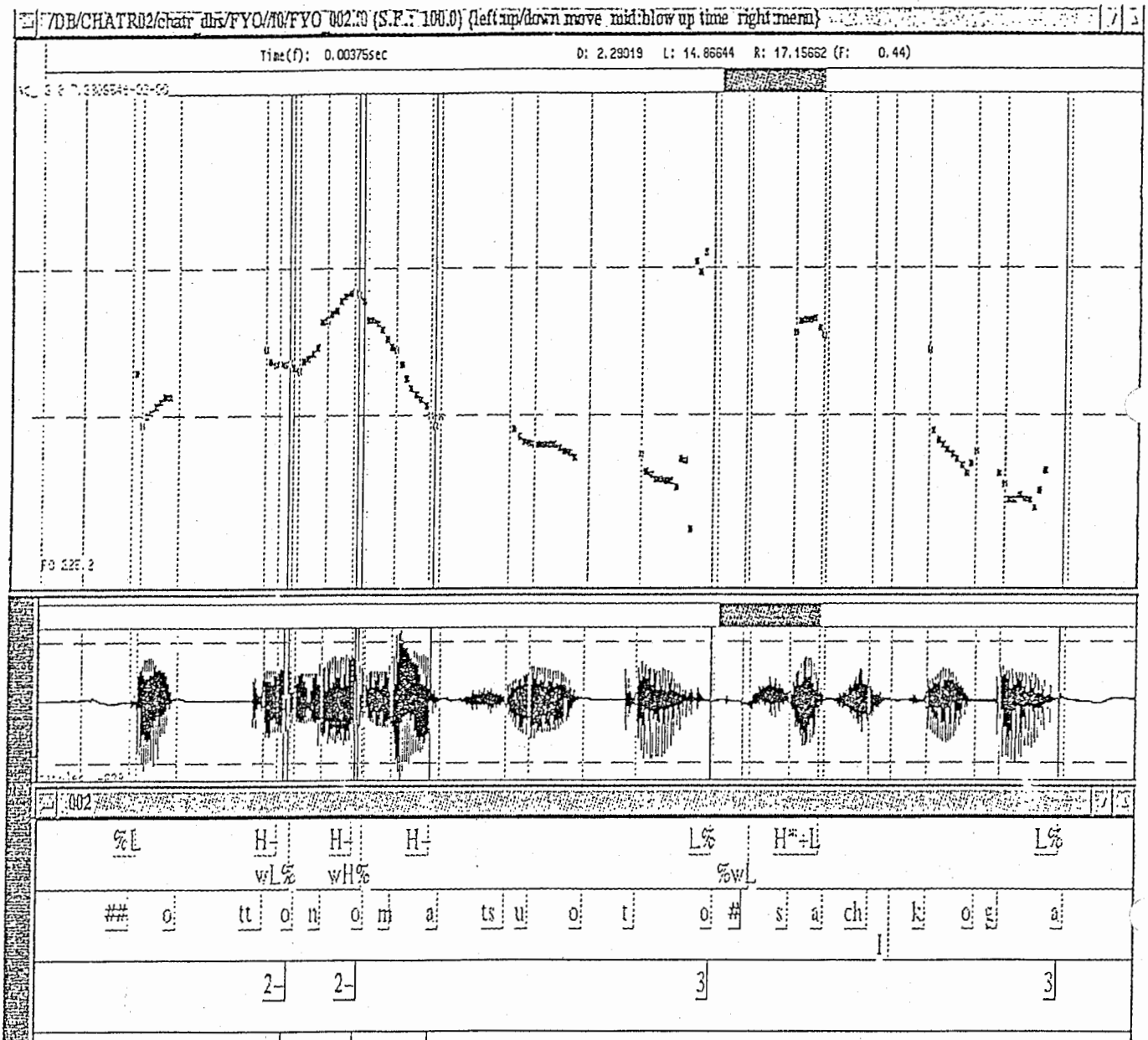
i～vii型はここで示す便宜上の名称であって、  
現在確定していない。

## SAMPLE LABELLING

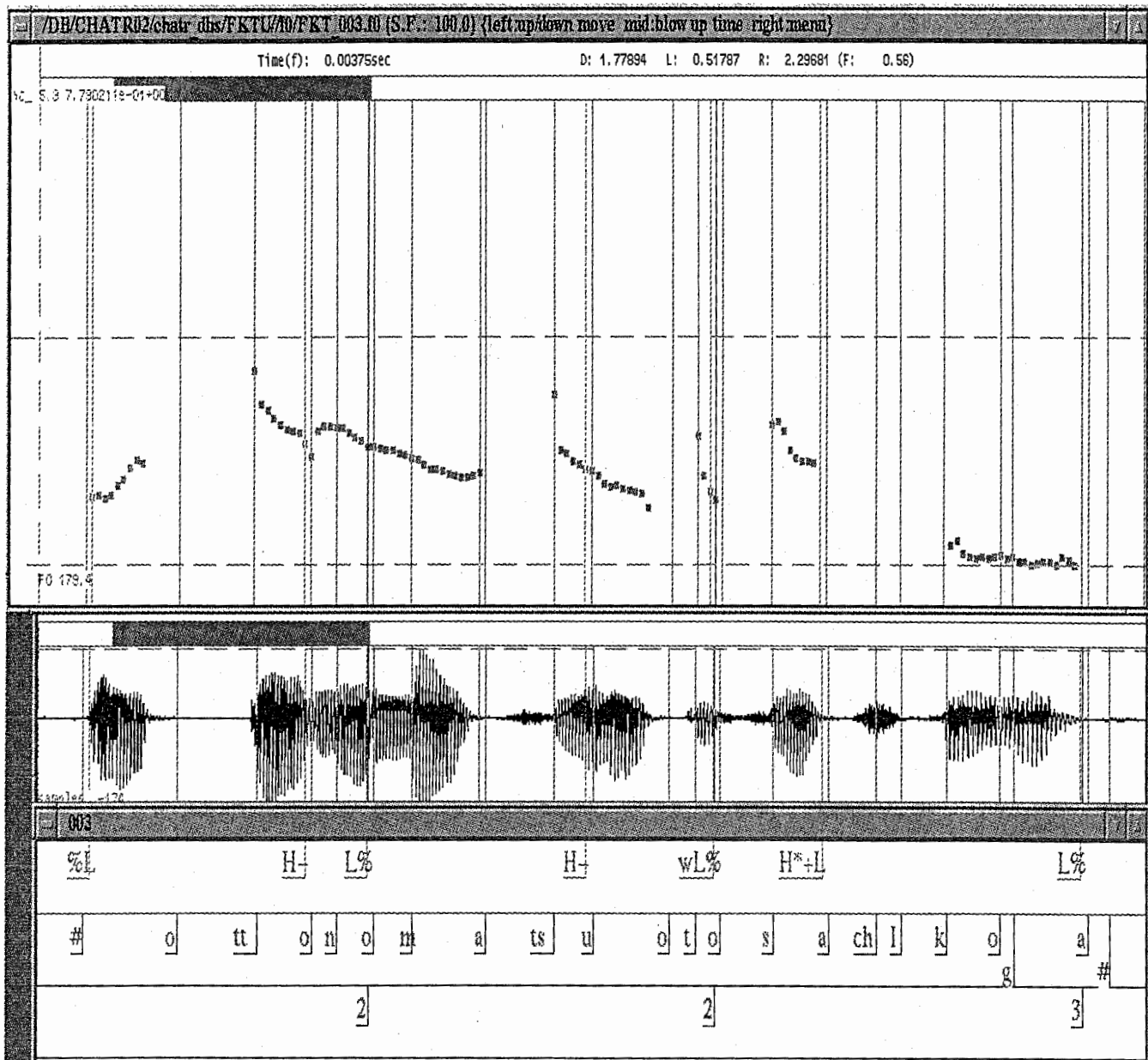
夫の松夫と左知子が (関西)	1
夫の松夫と左知子が (共通語)	2
ご飯にかける分だけ皿にのせ (関西)	3
ご飯にかける分だけ皿にのせ (共通語)	4
固くて重たいような気がする (関西)	5
固くて重たいような気がする (共通語)	6
食卓で朝刊をひろげている松夫が (関西)	7
食卓で朝刊をひろげている松夫が (共通語)	8
もつとも、待たれたのは三年目までで (関西)	9
もつとも、待たれたのは三年目までで (共通語)	10
こうなると、いっそあけすけに催促されているほうが (関西)	11
こうなると、いっそあけすけに催促されているほうが (共通語)	12
食卓で卵を割るのは (関西)	13
食卓で卵を割るのは (共通語)	14
温かくなっているが (関西)	15
温かくなっているが (共通語)	16
いつもより冷蔵庫の温度を下げたせい (関西)	17
いつもより冷蔵庫の温度を下げたせい (共通語)	18
いつもより余計にかき廻し (関西)	19
いつもより余計にかき廻し (共通語)	20
おれのと取り替えてやろうかと言ったが (関西)	21
おれのと取り替えてやろうかと言ったが (共通語)	22
おれのと取り替えてやろうかと言ったが (共通語)	22
赤いものの浮いた卵を (関西)	23
赤いものの浮いた卵を (共通語)	24
左知子は昆布の佃煮を (関西)	25
左知子は昆布の佃煮を (共通語)	26
子供の時分、見ちゃったのよ (関西)	27
子供の時分、見ちゃったのよ (共通語)	28
なかから、ヒヨコになりかけのが出て来たの (関西)	29
なかから、ヒヨコになりかけのが出て来たの (共通語)	30
今でも卵を割るときは (関西)	31
今でも卵を割るときは (共通語)	32
今どきそんな卵はないよ (関西)	33
今どきそんな卵はないよ (共通語)	34
すぐに意味が判って (関西)	35
すぐに意味が判って (共通語)	36
田舎じゃないんだから (関西)	37
田舎じゃないんだから (共通語)	38
雄鶏と雌鶏が一緒になって (関西)	39
雄鶏と雌鶏が一緒になって (共通語)	40
雄は早いとこ焼鳥にでもなっさ (関西)	41
雄は早いとこ焼鳥にでもなっさ (共通語)	42
雄鶏と一緒にいても、無精卵を産む雌鶏がいるんじゃないの (関西)	43
雄鶏と一緒にいても、無精卵を産む雌鶏がいるんじゃないの (共通語)	44
出来なきゃ出来ないでもいいじゃないかと言って (関西)	45
出来なきゃ出来ないでもいいじゃないかと言って (共通語)	46
夫婦揃って検査を受けようという (関西)	47
夫婦揃って検査を受けようという (共通語)	48
たしか小学校のそばの瀬戸物屋の店先だった (関西)	49
たしか小学校のそばの瀬戸物屋の店先だった (共通語)	50
迎えに来た母親が一緒だったような気がする (関西)	51
迎えに来た母親が一緒だったような気がする (共通語)	52

茶碗や湯呑み、などの置かれた台の下に (関西)	53
茶碗や湯呑み、などの置かれた台の下に (共通語)	54
その中に卵が沢山入っていた (関西)	55
その中に卵が沢山入っていた (共通語)	56
高校時代の友達 (関西)	57
高校時代の友達 (共通語)	58
もうすぐ三十に手が届くというのに (関西)	59
もうすぐ三十に手が届くというのに (共通語)	60
うまくゆかないもんねえ (関西)	61
うまくゆかないもんねえ (共通語)	62
欲のない人に当たるのよ (関西)	63
欲のない人に当たるのよ (共通語)	64
結果が出るまでに二週間かかった (関西)	65
結果が出るまでに二週間かかった (共通語)	66
かぎつけられはしないかと不安だった (関西)	67
かぎつけられはしないかと不安だった (共通語)	68
私には責任がなかったのだ (関西)	69
私には責任がなかったのだ (共通語)	70
こういう結果を予測していたからではなかったか (関西)	71
こういう結果を予測していたからではなかったか (共通語)	72
知らないうちにお多福風邪にかかっていたって こともあるんですってよ (関西)	73
知らないうちにお多福風邪にかかっていたって こともあるんですってよ (共通語)	74
とにかく、あたし一人が罪人みたいに言われるの (関西)	75
とにかく、あたし一人が罪人みたいに言われるの (共通語)	76
松夫は煙草に手を伸ばした (関西)	77
松夫は煙草に手を伸ばした (共通語)	78
ことばの代わりに煙を吐いておしまいにする (関西)	79
ことばの代わりに煙を吐いておしまいにする (共通語)	80
次の日の夕方 (関西)	81
次の日の夕方 (共通語)	82
何度もそう言い、あたしも頂こう (関西)	83
何度もそう言い、あたしも頂こう (共通語)	84
グラスの置き方も、お祝いというより (関西)	85
グラスの置き方も、お祝いというより (共通語)	86
あのはなしが本当だとしたら (関西)	87
あのはなしが本当だとしたら (共通語)	88
道にせり出した看板の感じが変わっていた (関西)	89
道にせり出した看板の感じが変わっていた (共通語)	90
子供が出来たというのは本当だったんですか (関西)	91
子供が出来たというのは本当だったんですか (共通語)	92
年は三十二、三 (関西)	93
年は三十二、三 (共通語)	94
安直なのではなかった (関西)	95
安直なのではなかった (共通語)	96
大型の手帳を出した (関西)	97
大型の手帳を出した (共通語)	98
「スケジュール、スケジュール」とひとりごとを言いながら (関西)	99
「スケジュール、スケジュール」とひとりごとを言いながら (共通語)	100
一週間後の同じ時間に (関西)	101
一週間後の同じ時間に (共通語)	102

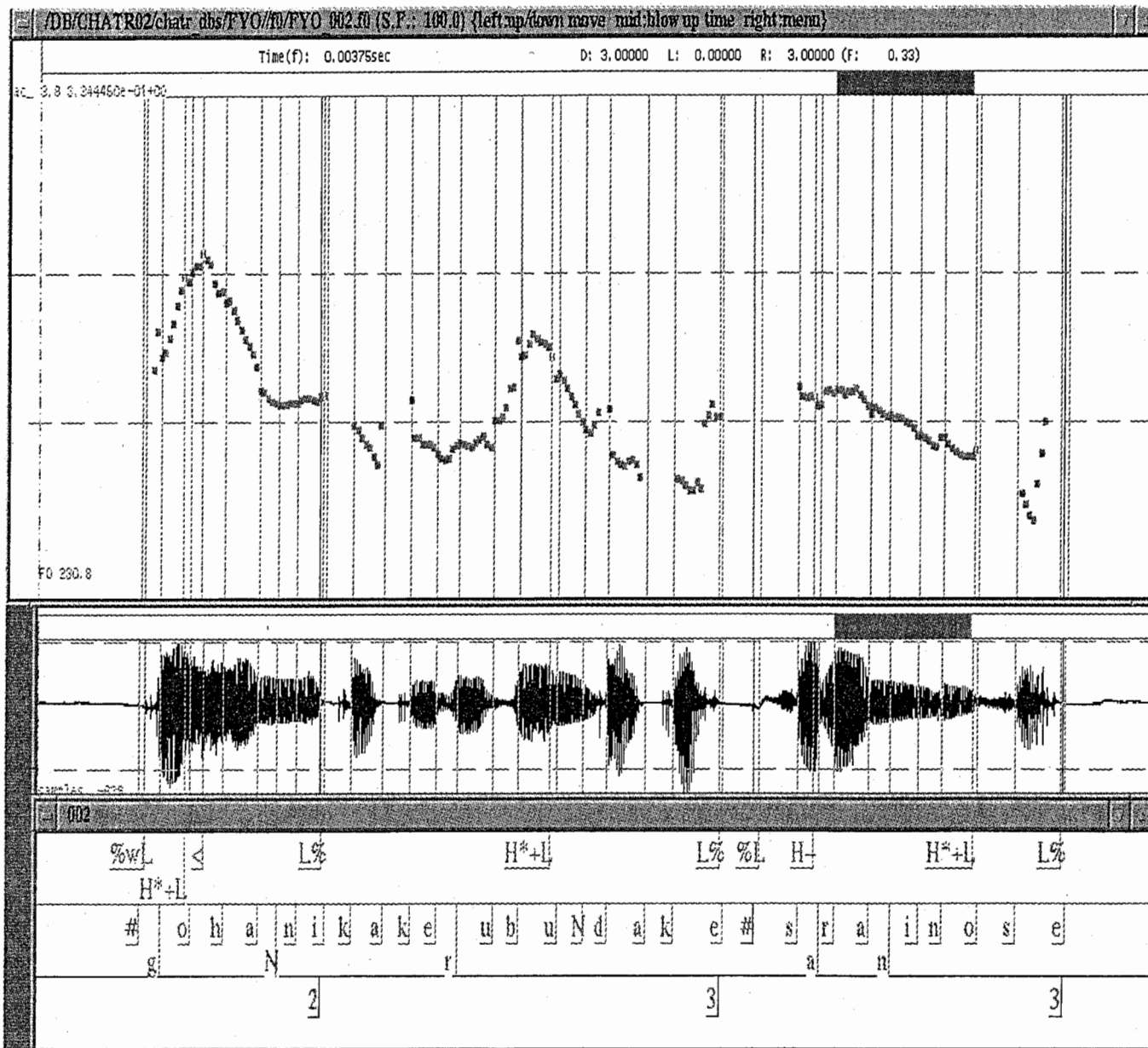




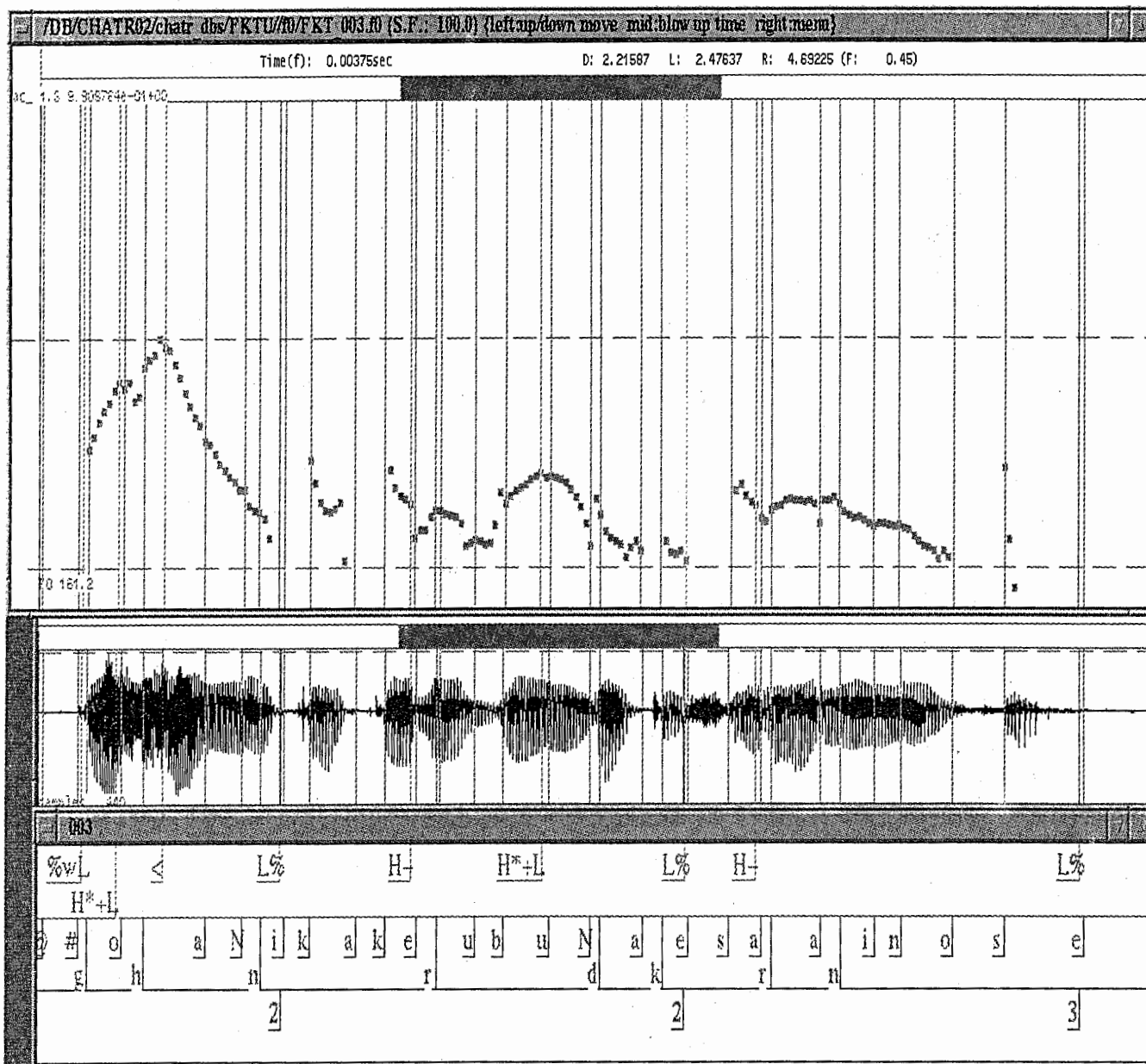
“夫の松夫と左知子が” (関西)



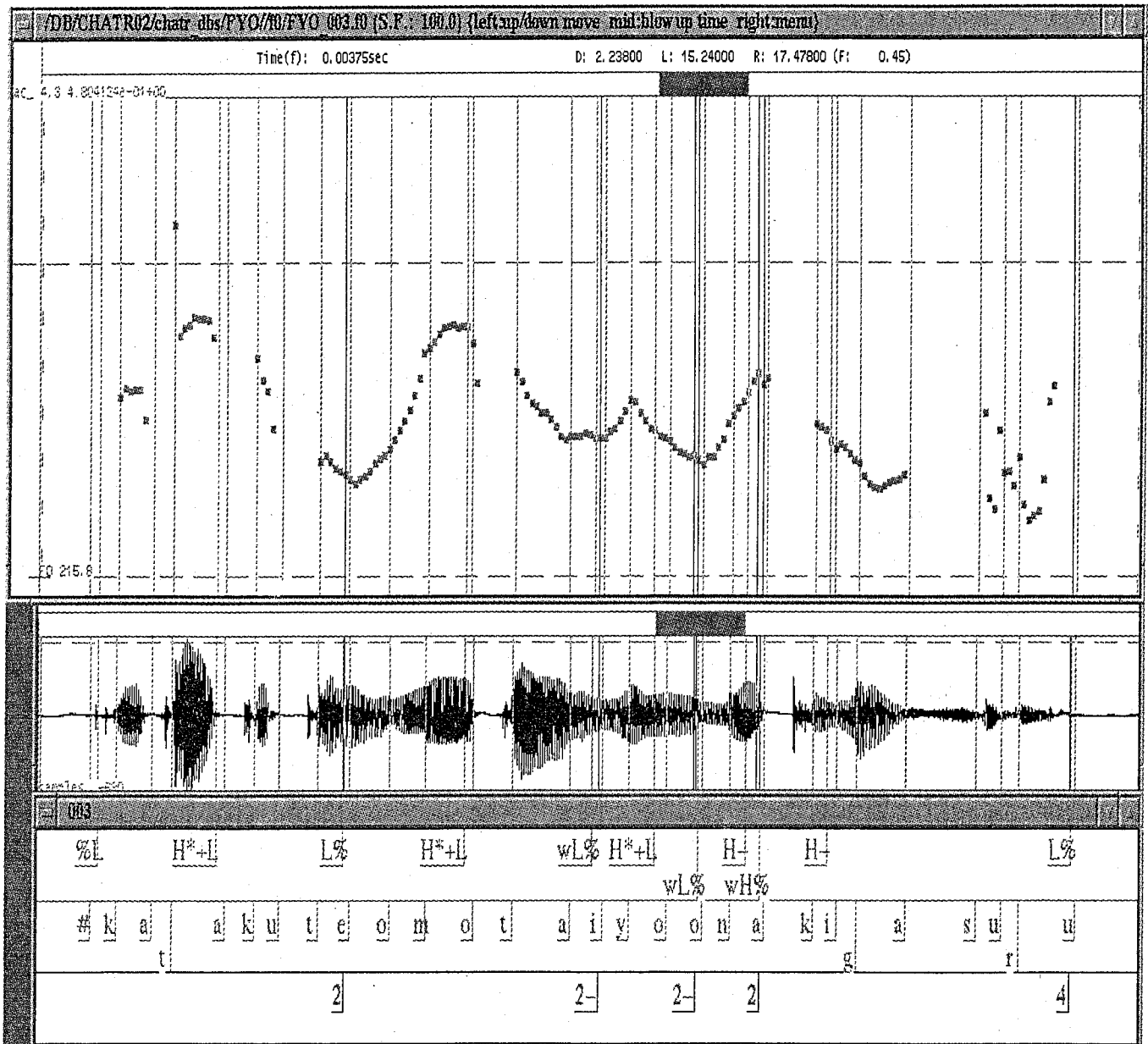
“夫の松夫と左知子が” (共通語)



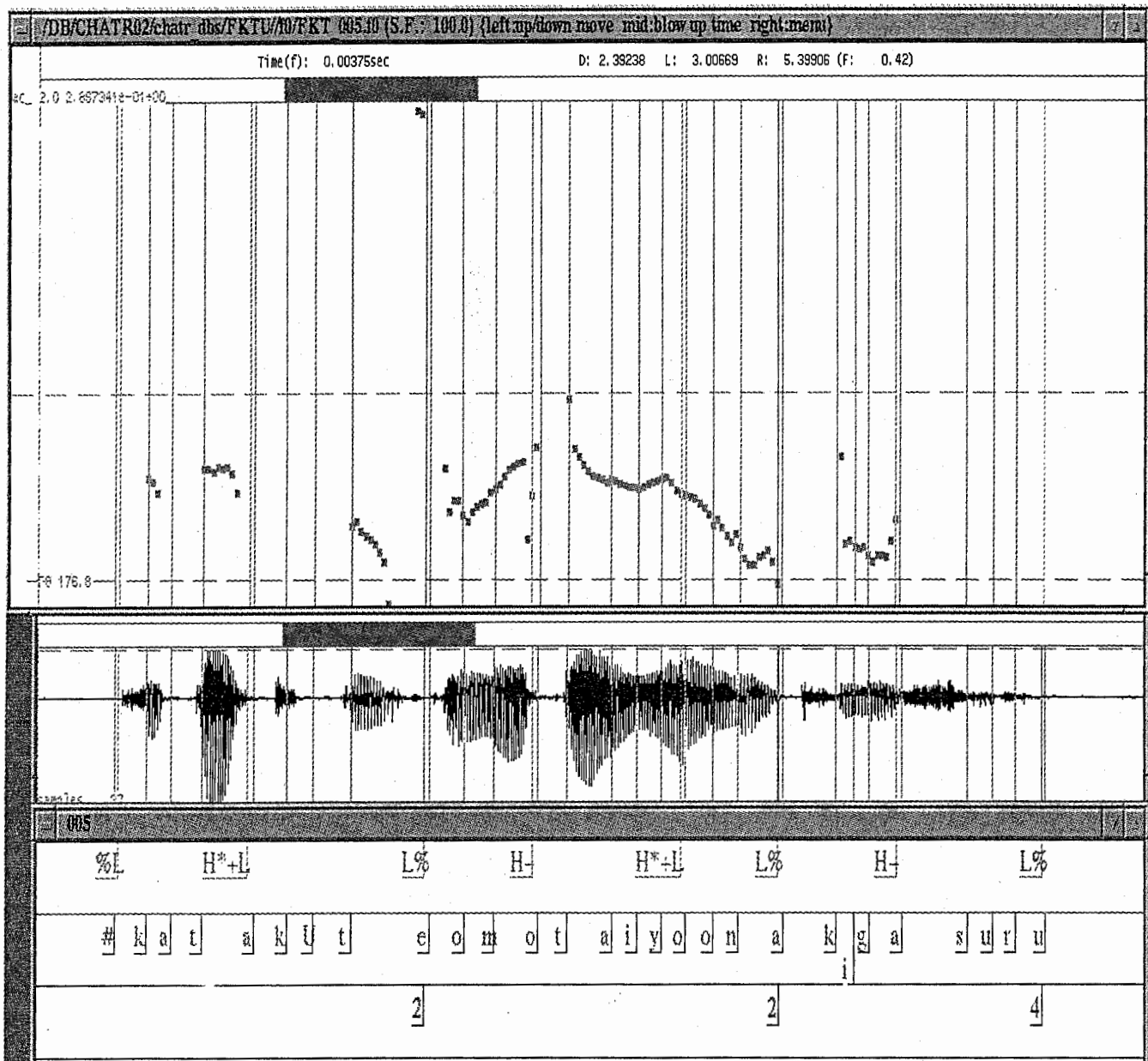
“ご飯にかける分だけ皿にのせ” (関西)



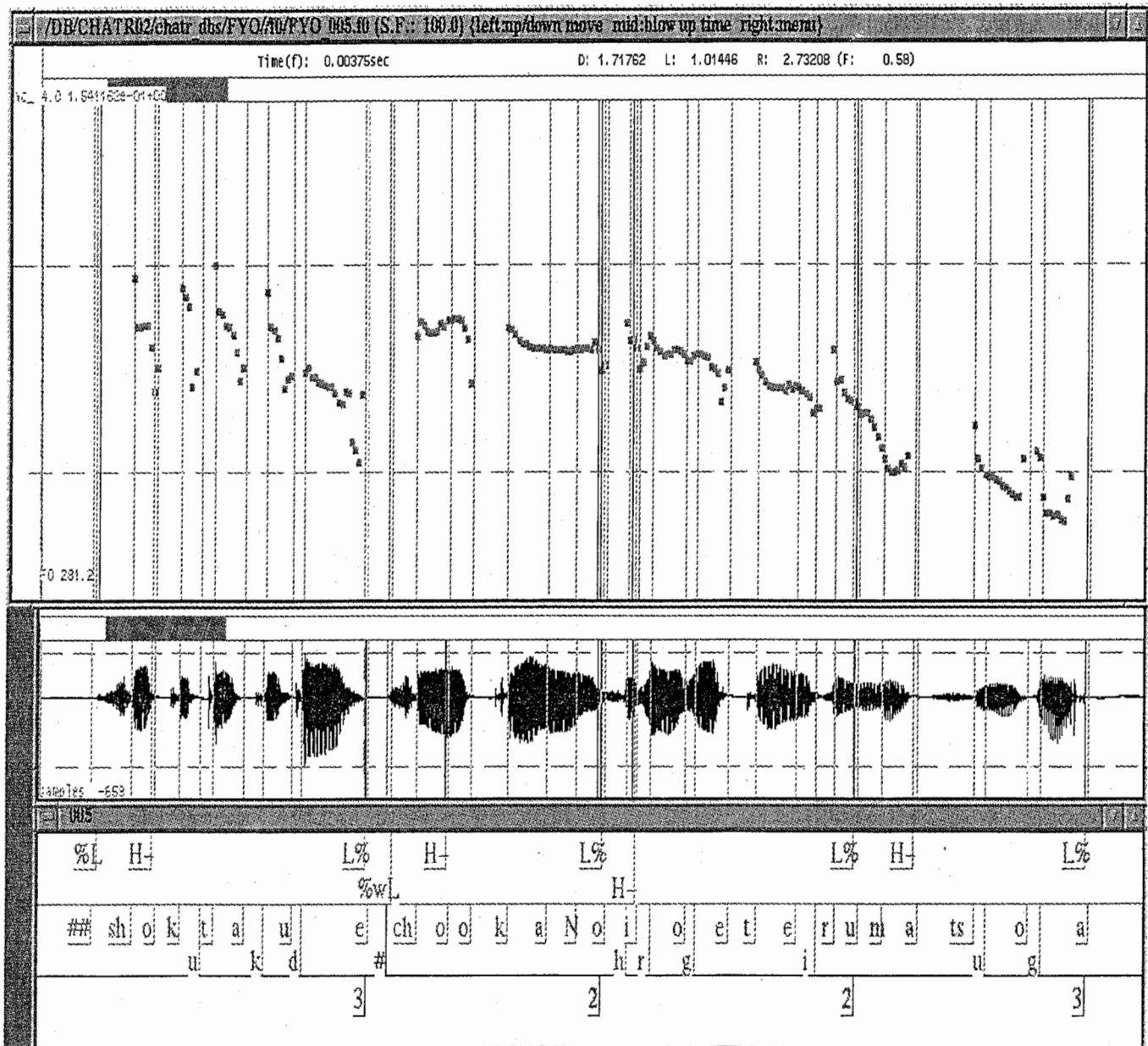
“ご飯にかける分だけ皿にのせ” (共通語)



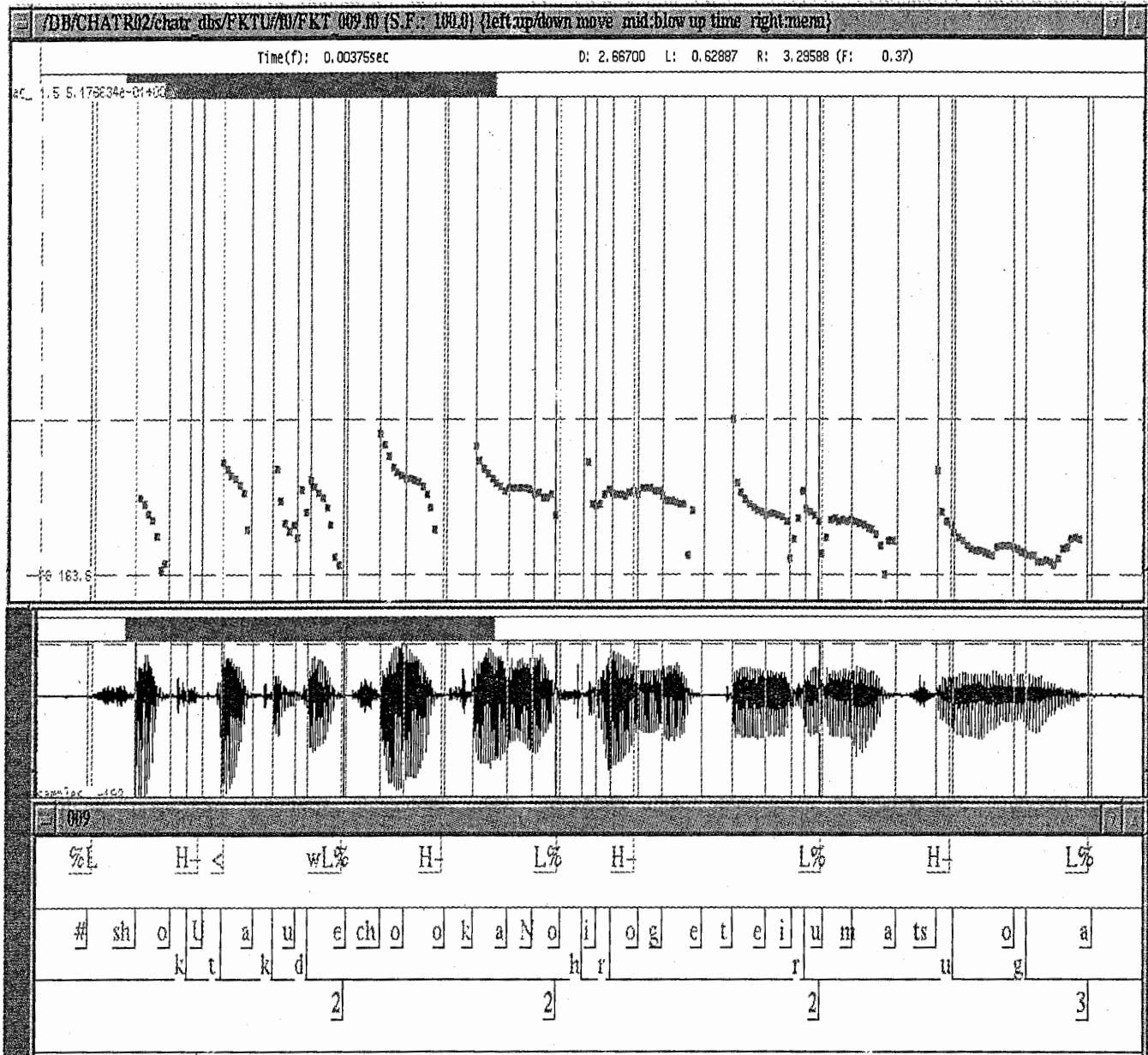
“固くて重たいような気がする” (関西)



“固くて重たいような気がする” (共通語)

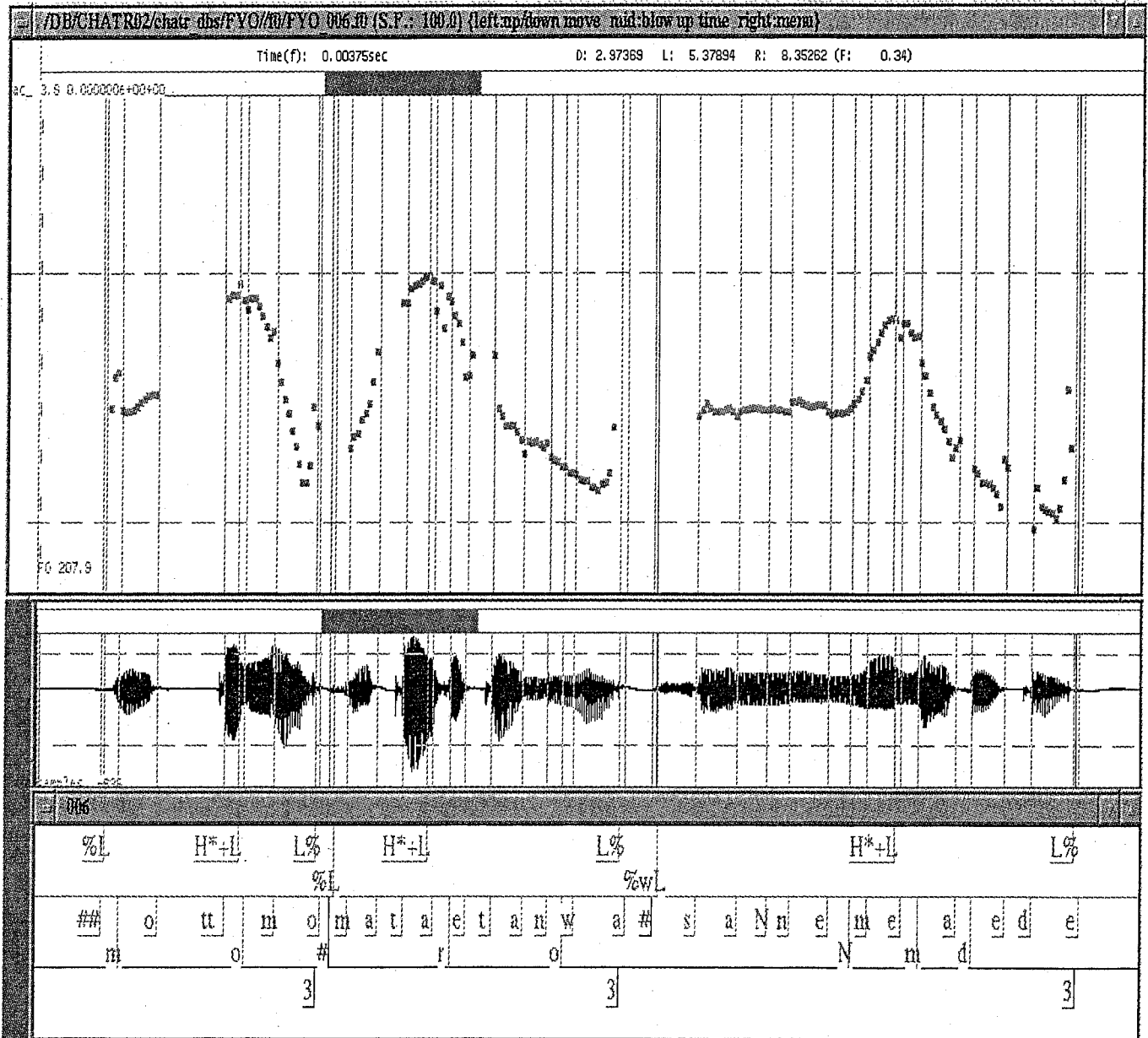


“食卓で朝刊をひろげている松夫が” (関西)

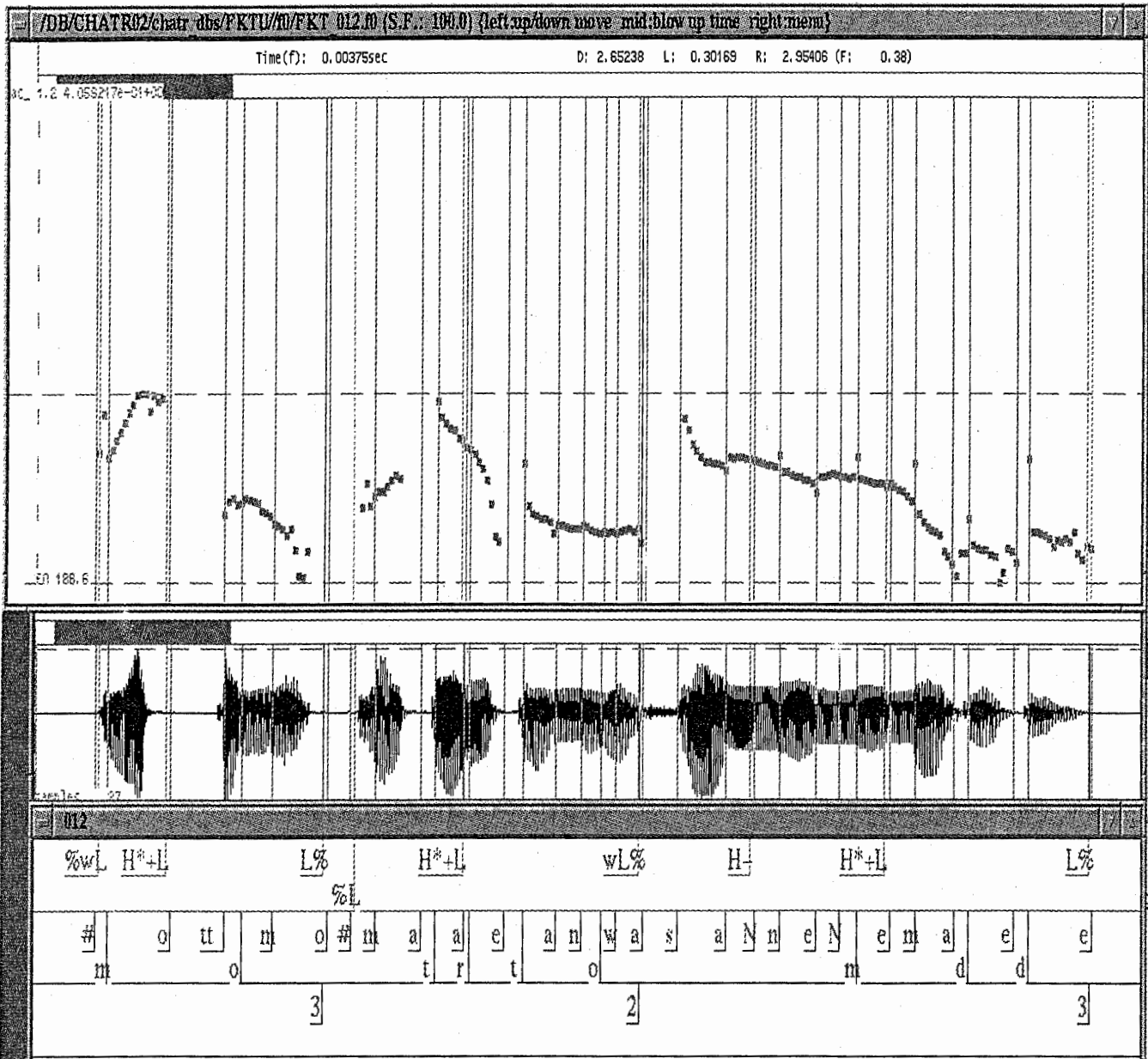


“食卓で朝刊をひろげている松夫が” (共通語)

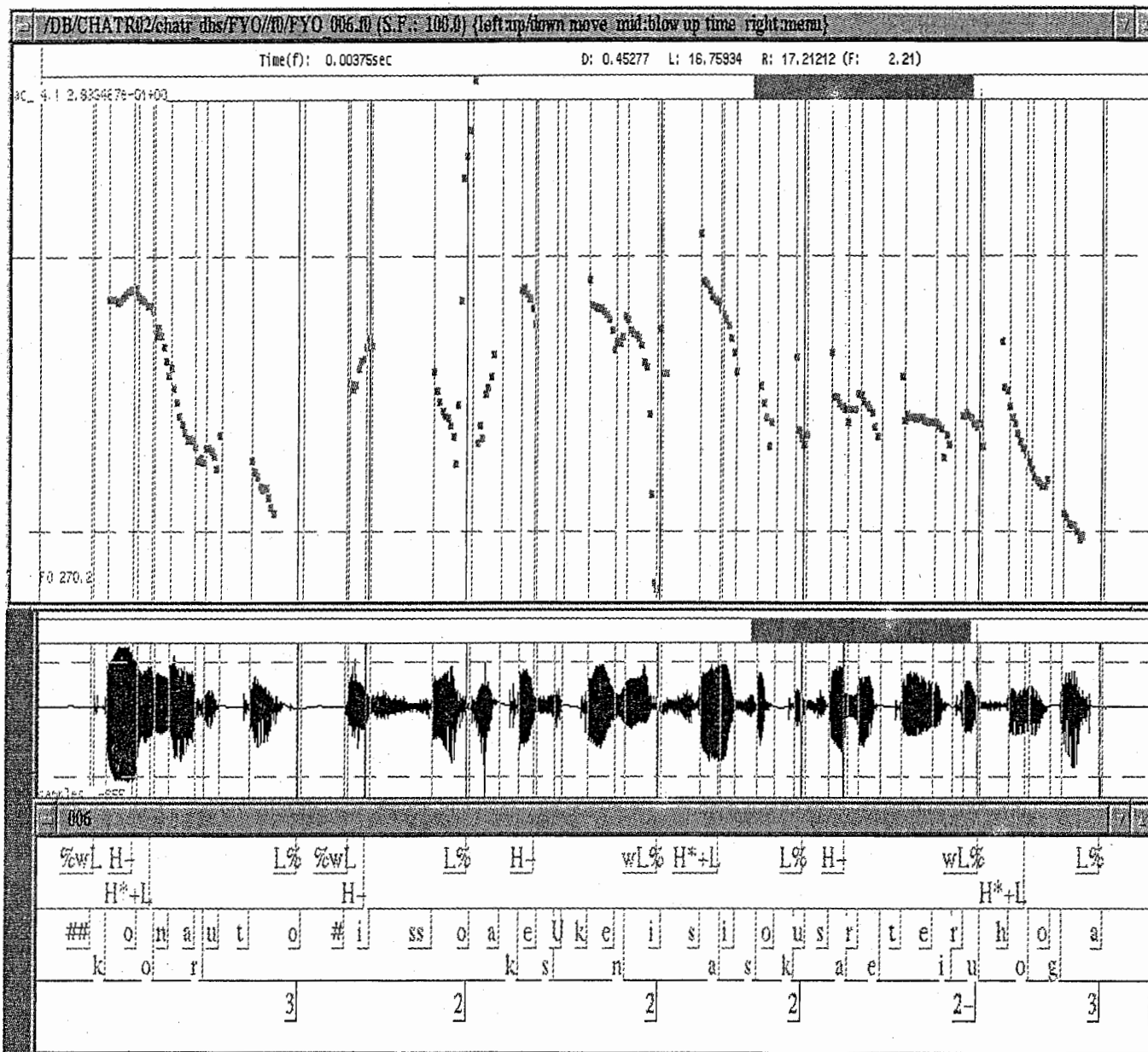




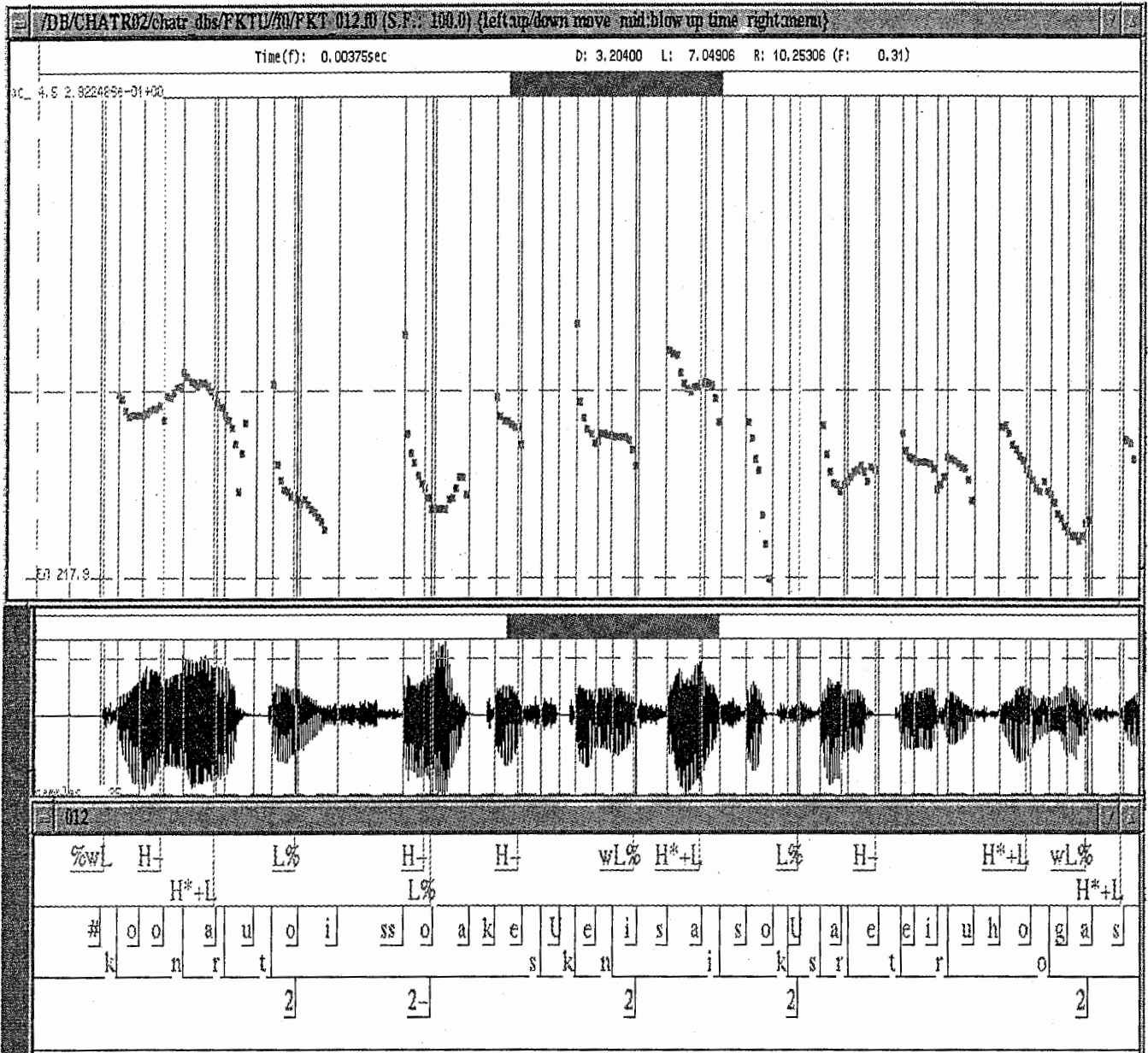
“もともと、待たれたのは三年目まで” (関西)



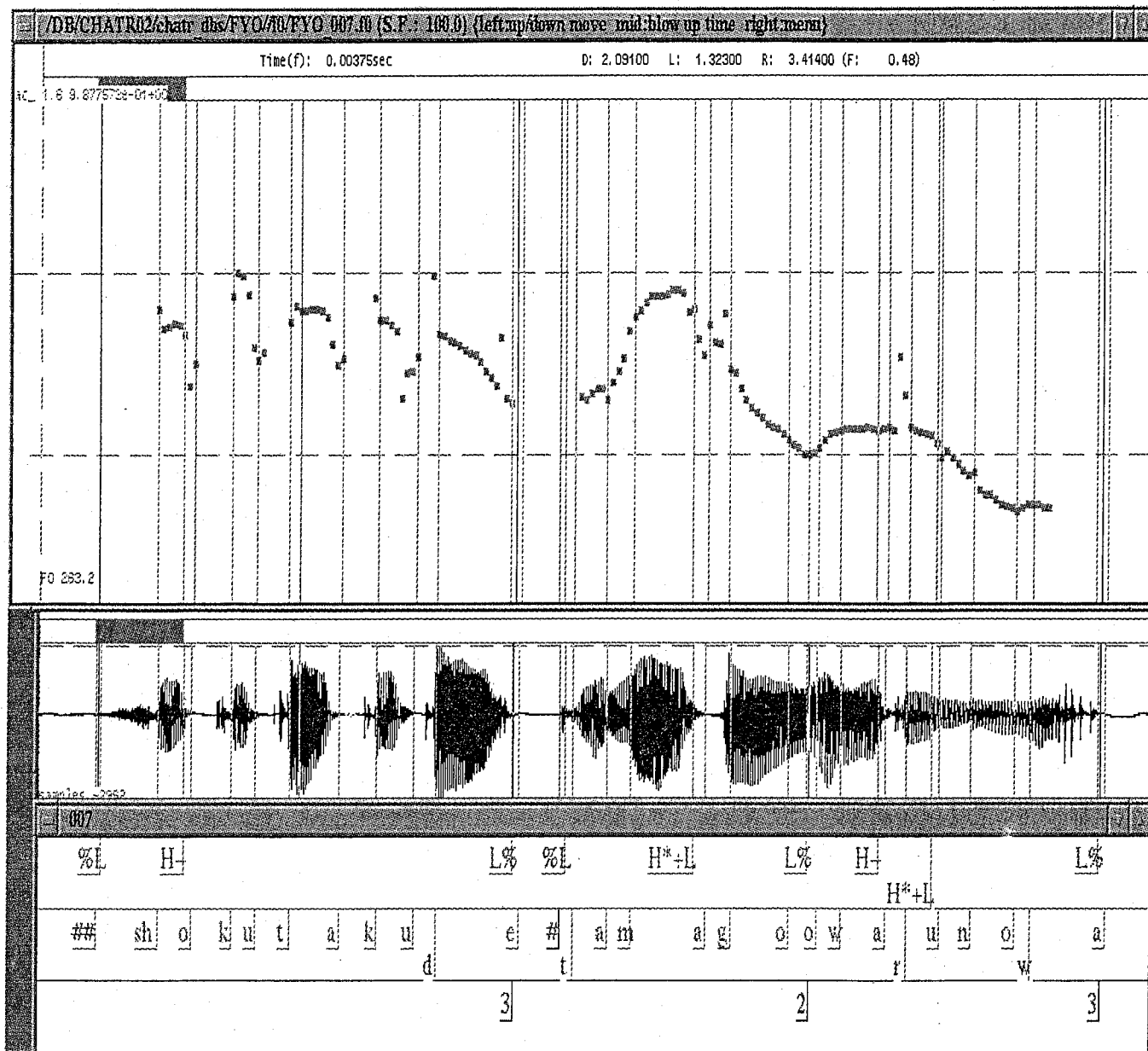
“もっとも、待たれたのは三年目まで” (共通語)



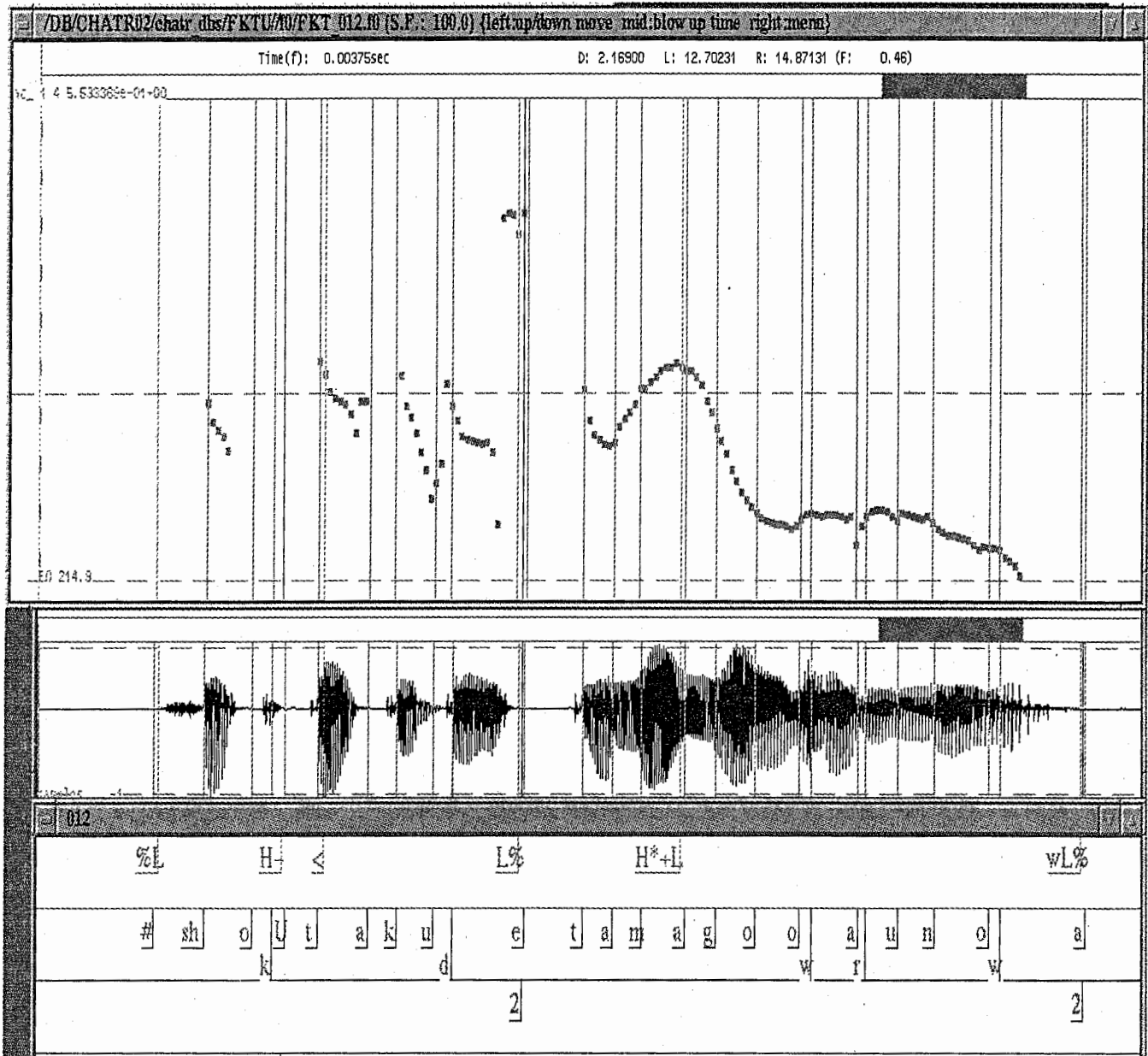
“こうなると、いっそあけすけに催促されているほうが” (関西)



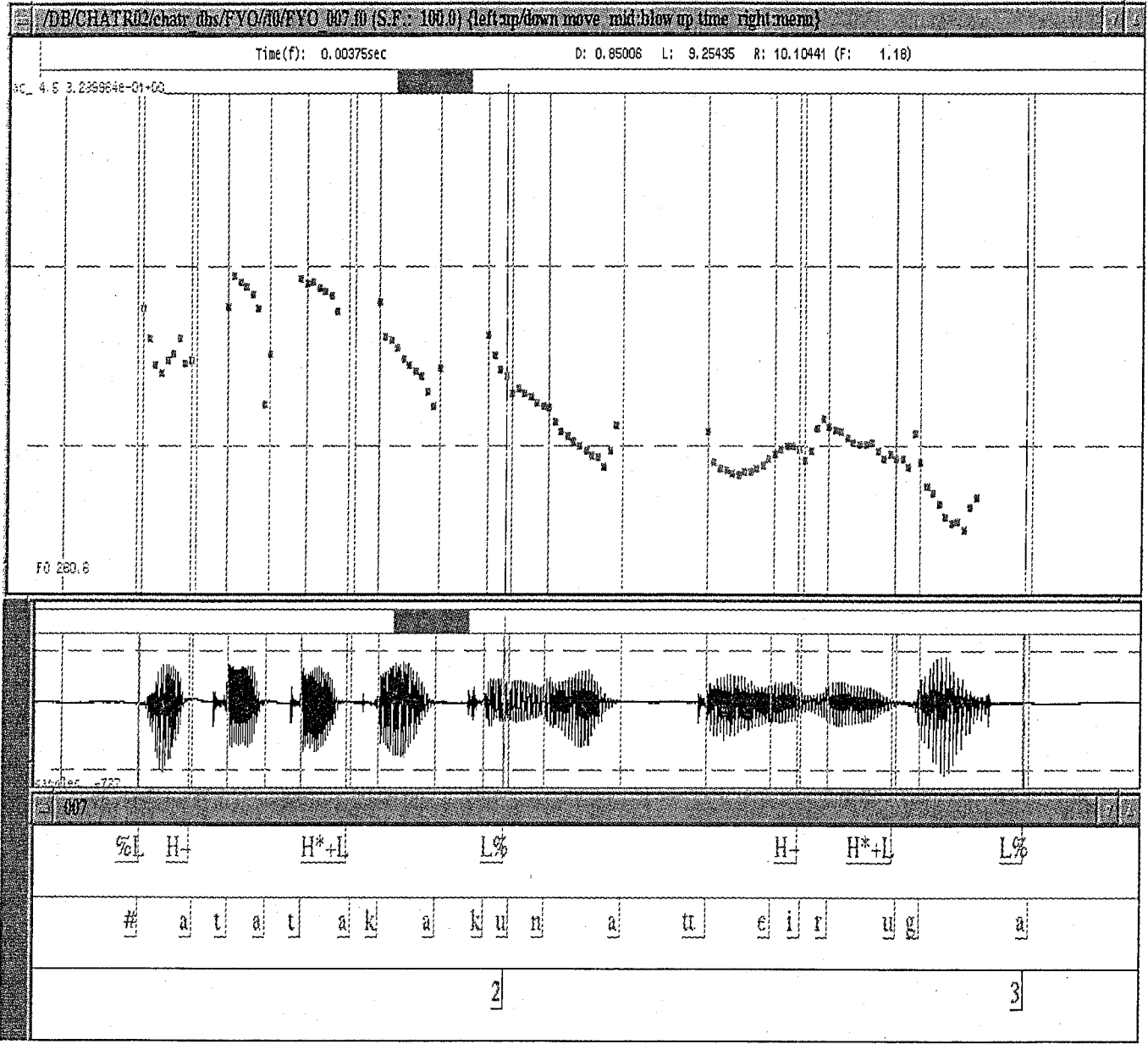
“こうなると、いっそあけすけに催促されているほうが” (共通語)



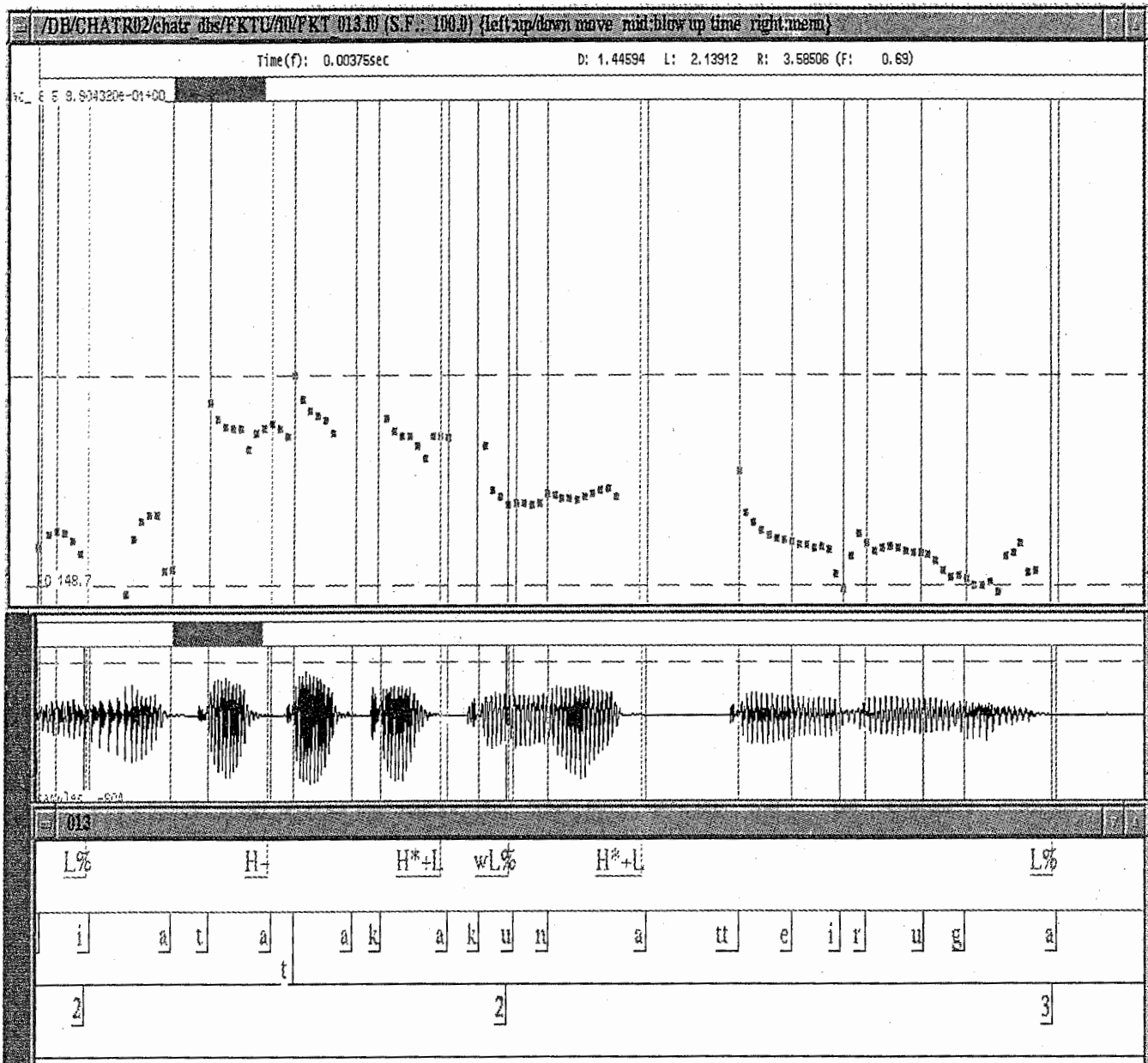
“食卓で卵を割るのは” (関西)



“食卓で卵を割るのは” (共通語)

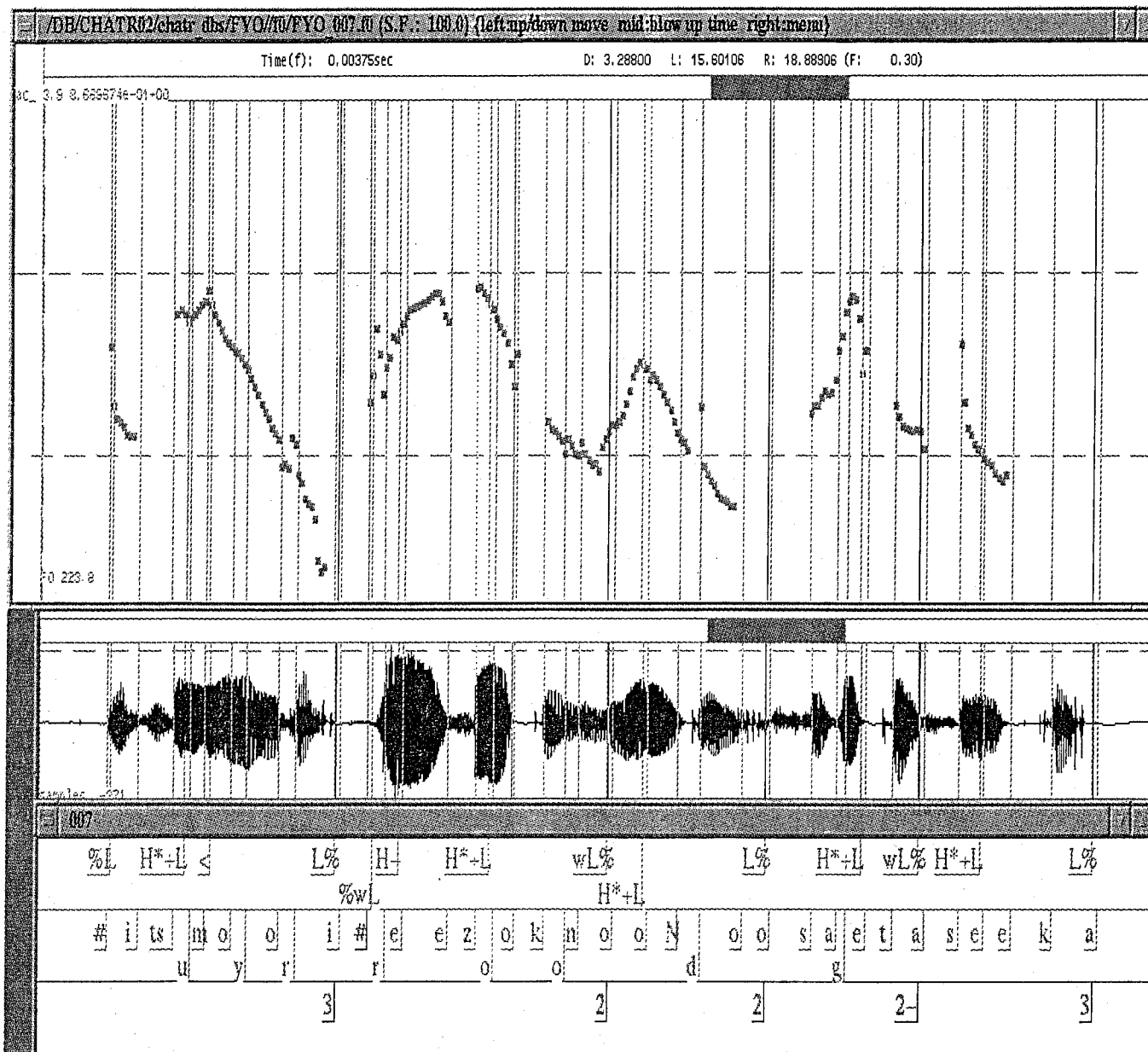


“温かくなっているが” (関西)

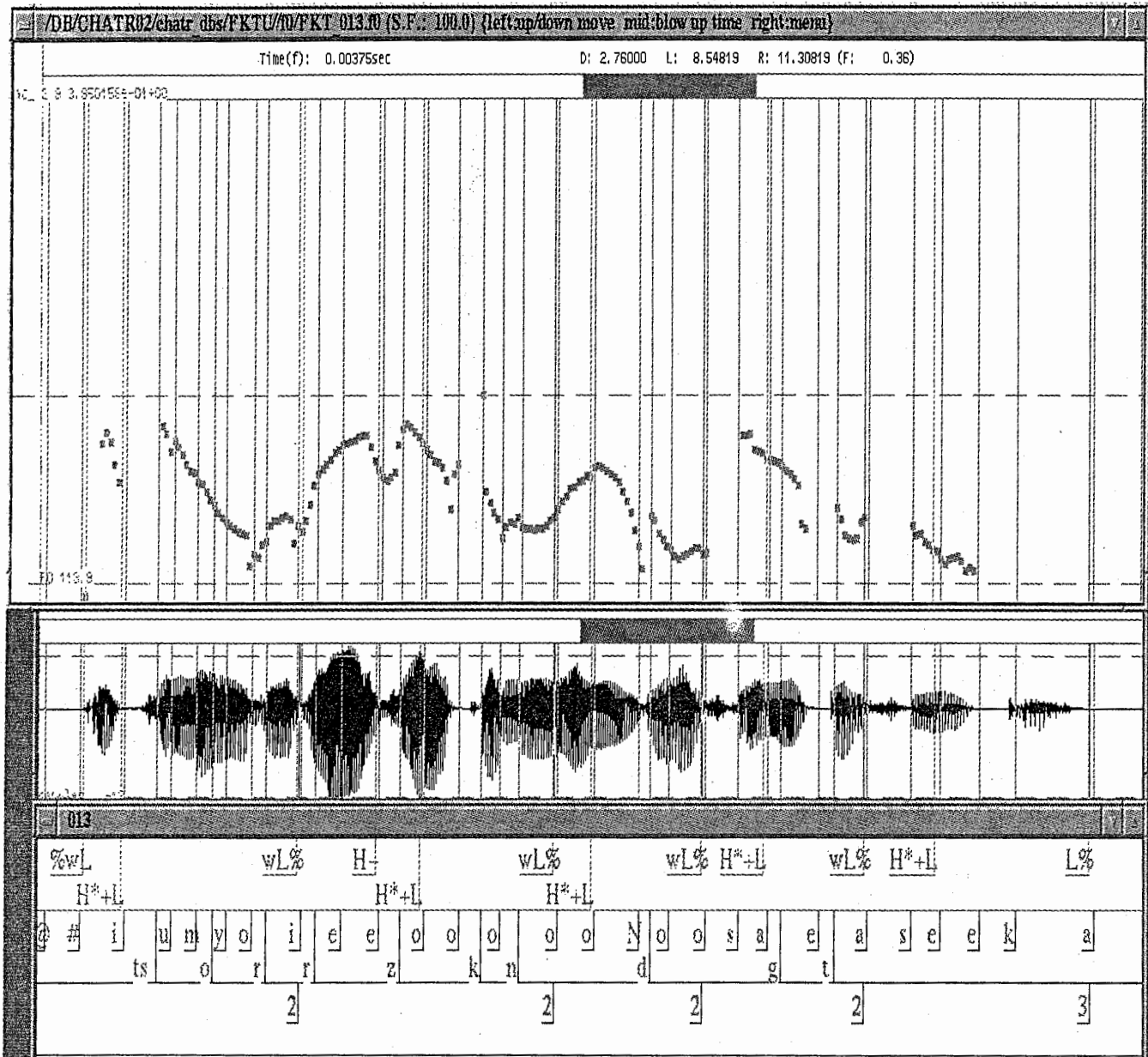


“温かくなっているが” (共通語)

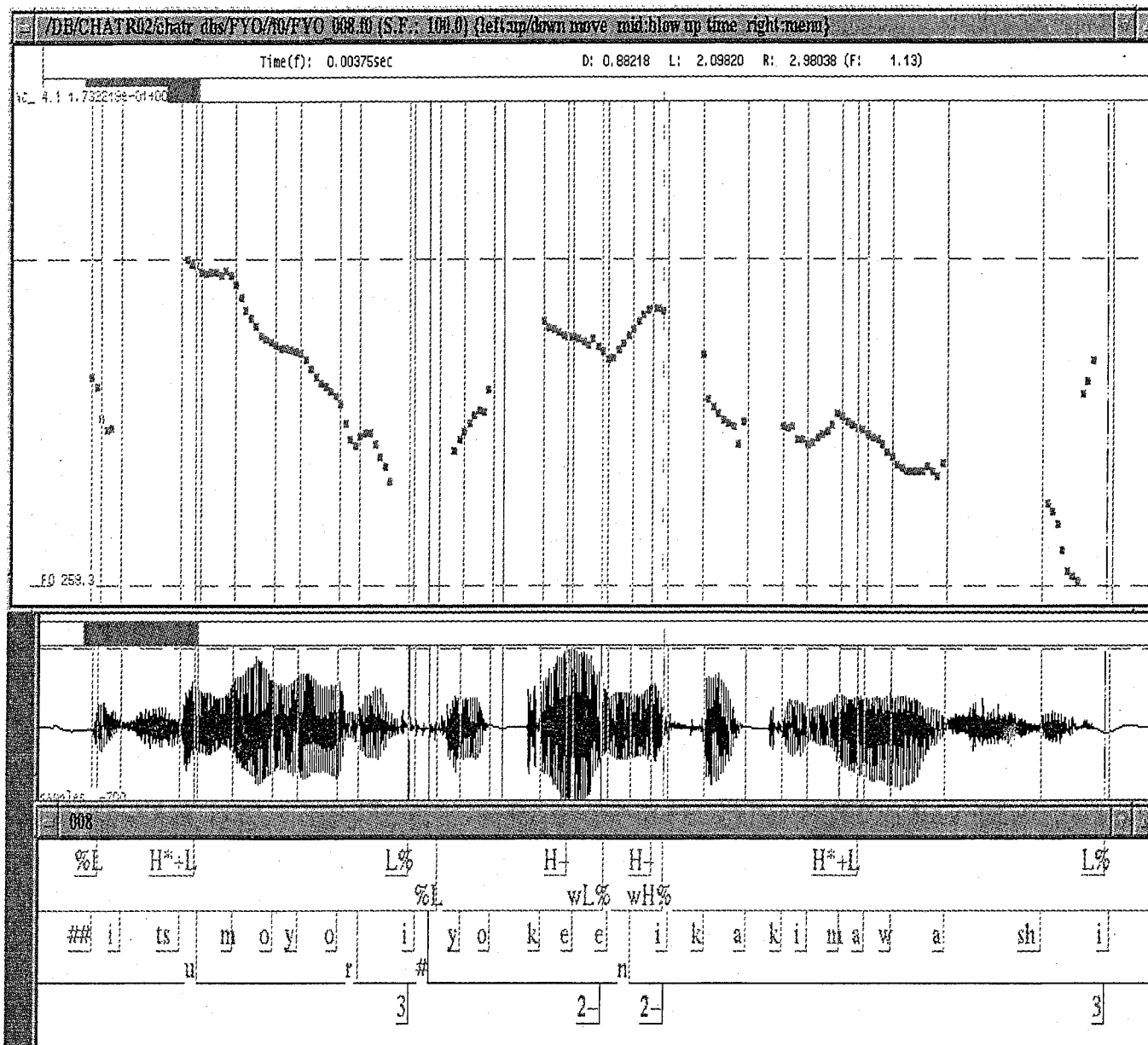




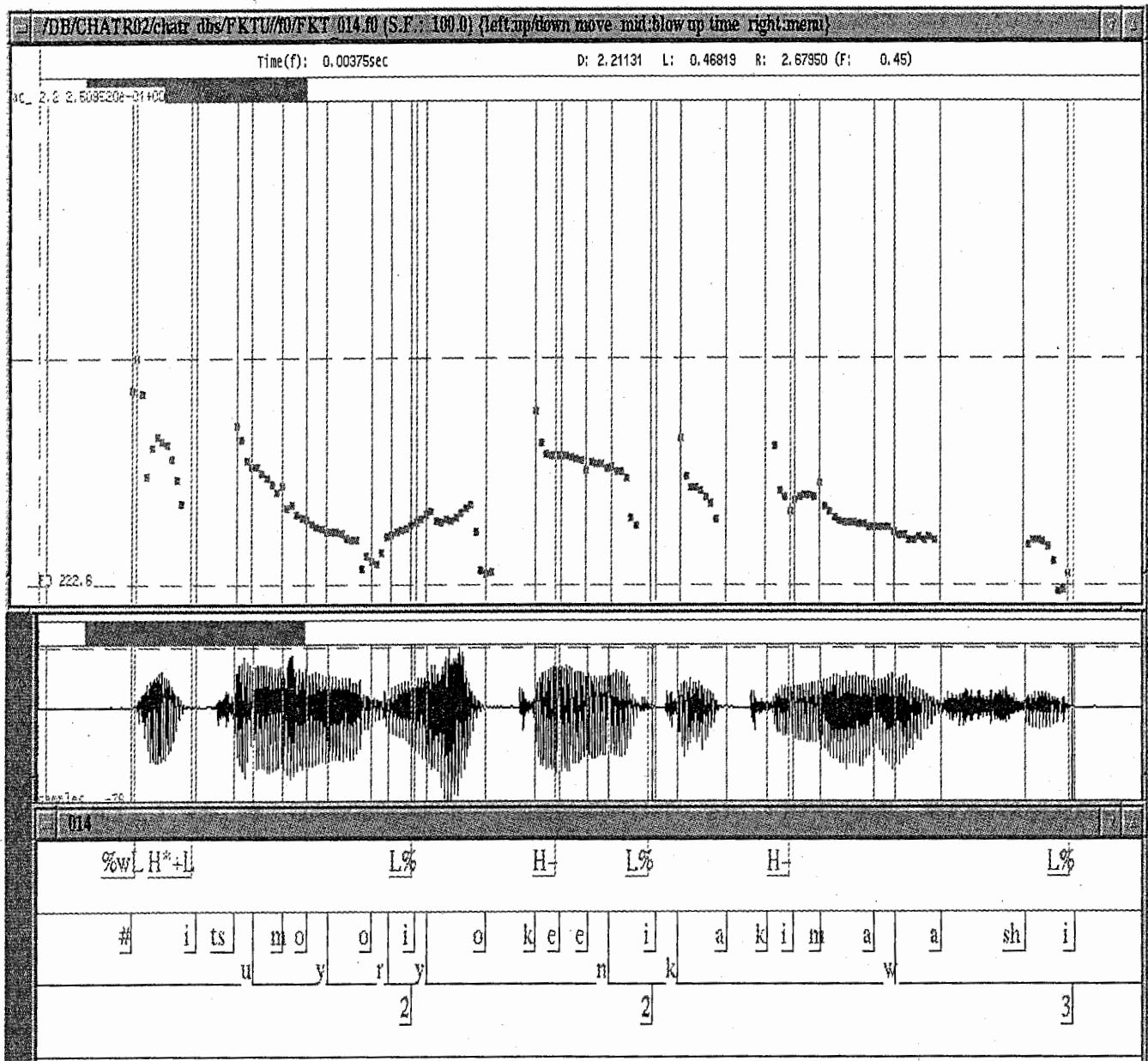
“いつもより冷蔵庫の温度を下げたせいかな” (関西)



“いつもより冷蔵庫の温度を下げたせいか” (共通語)

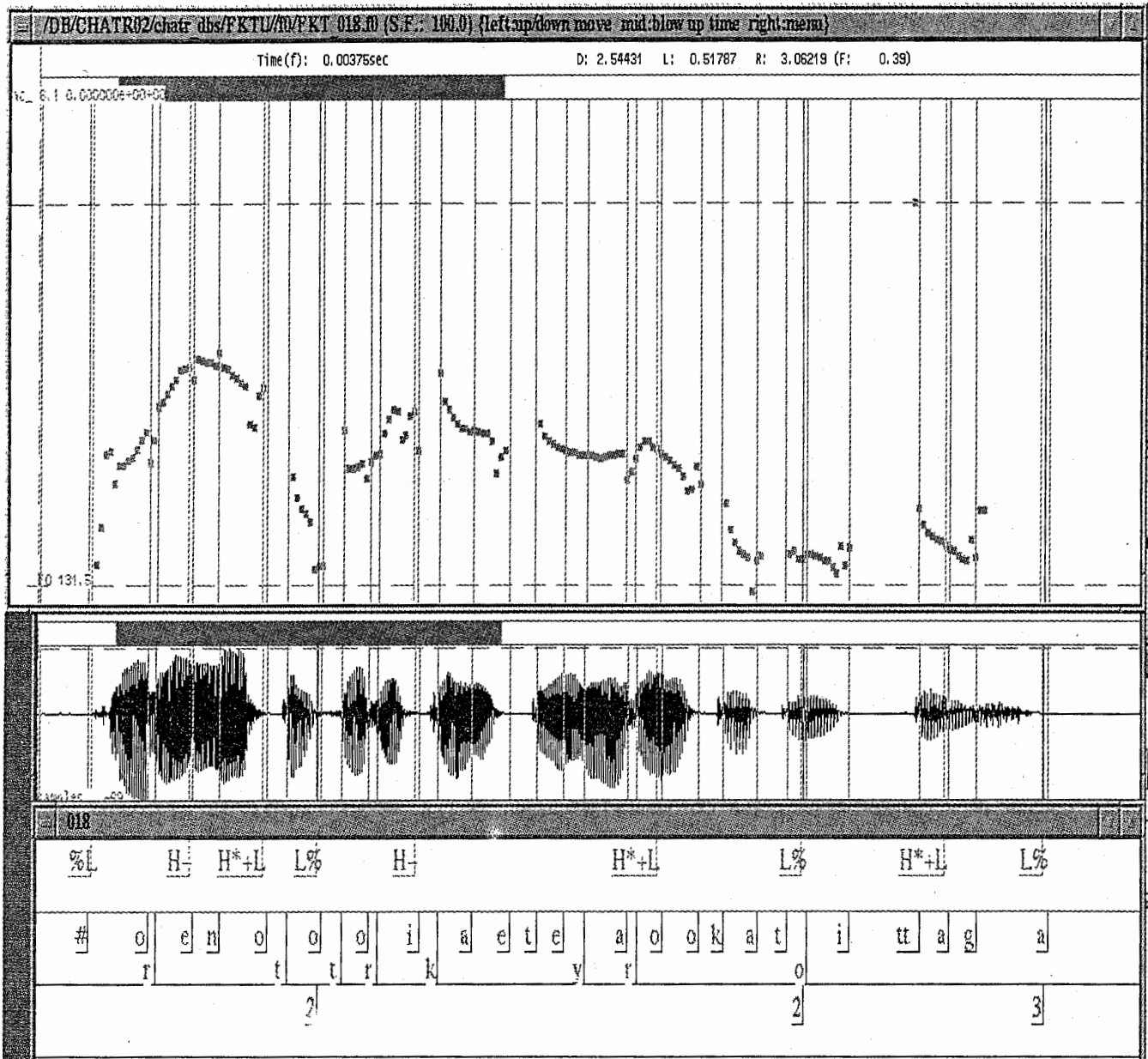


“いつもより余計にかき廻し” (関西)

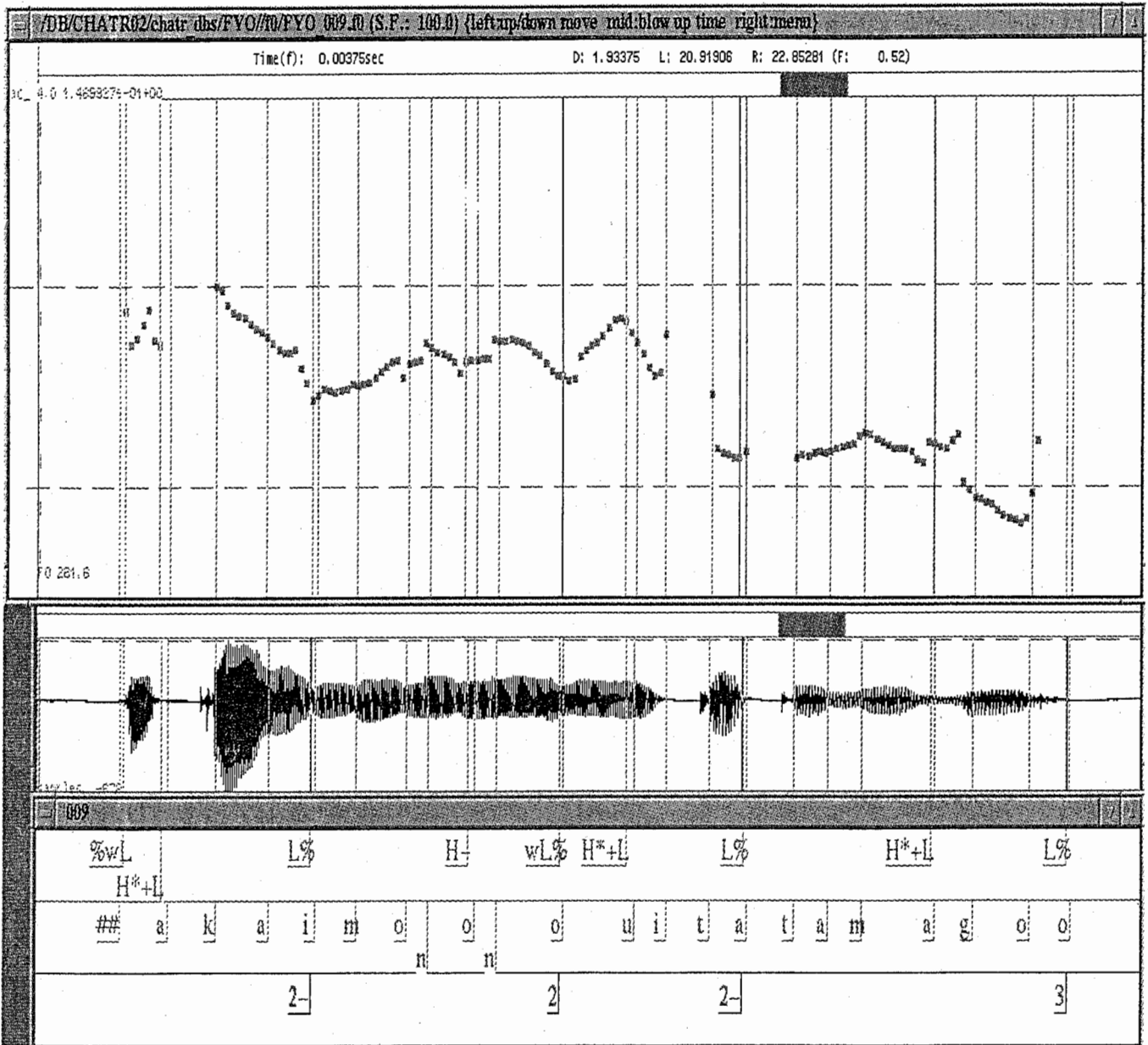


“いつもより余計にかき廻し” (共通語)





“おれのと取り替えてやろうかと言ったが” (共通語)

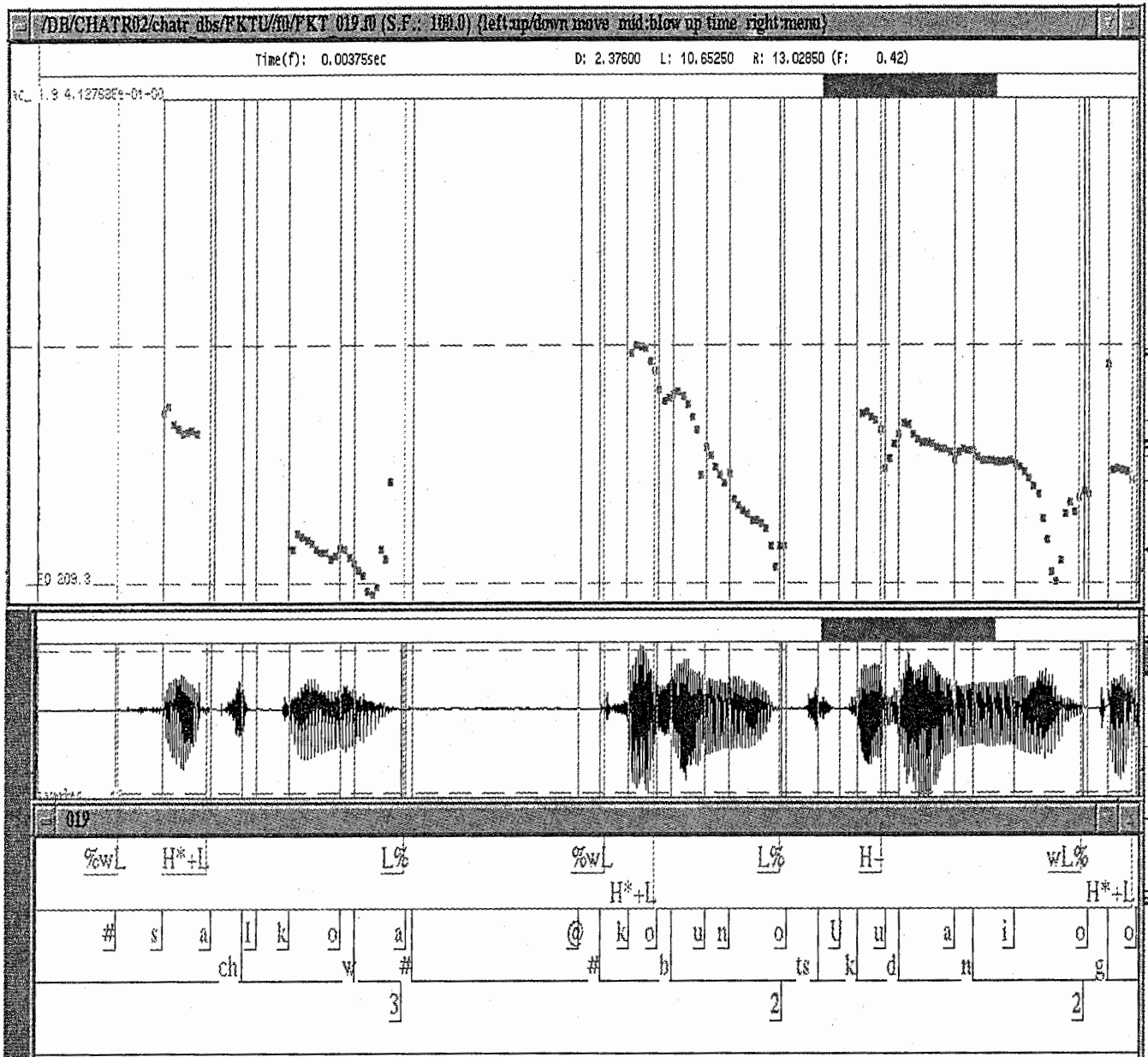


“赤いものの浮いた卵を” (関西)

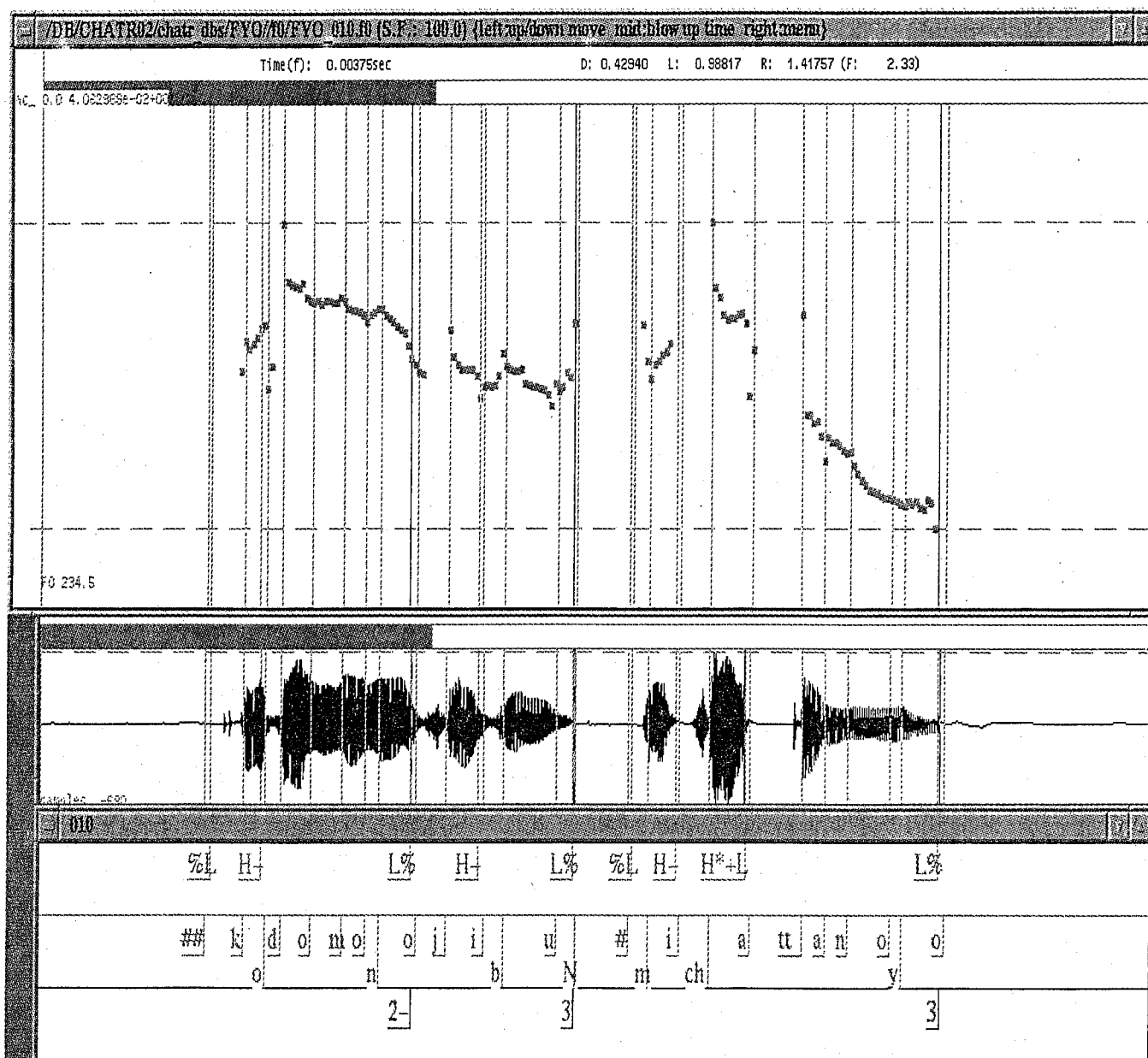




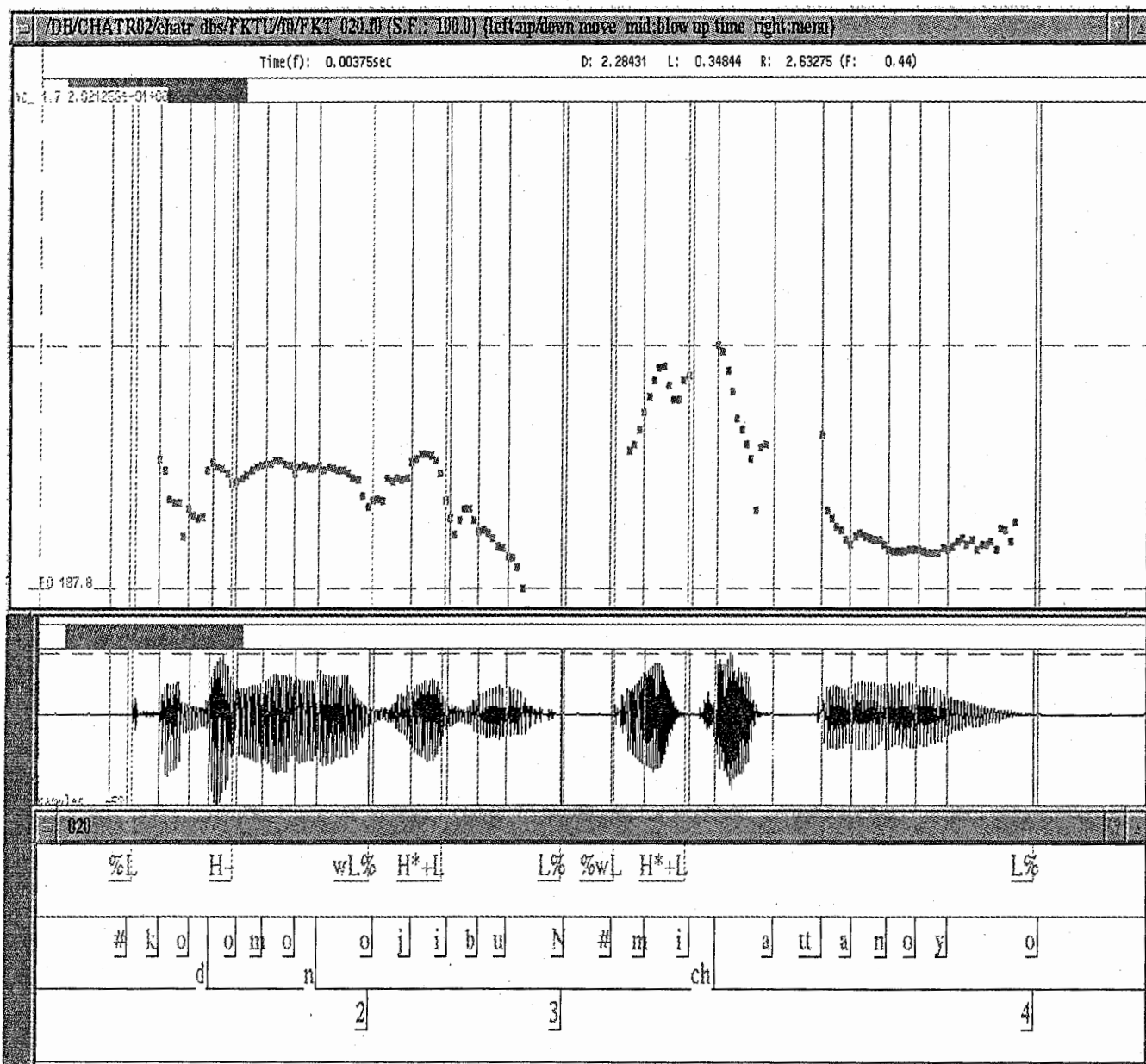




“左知子は昆布の佃煮を” (共通語)

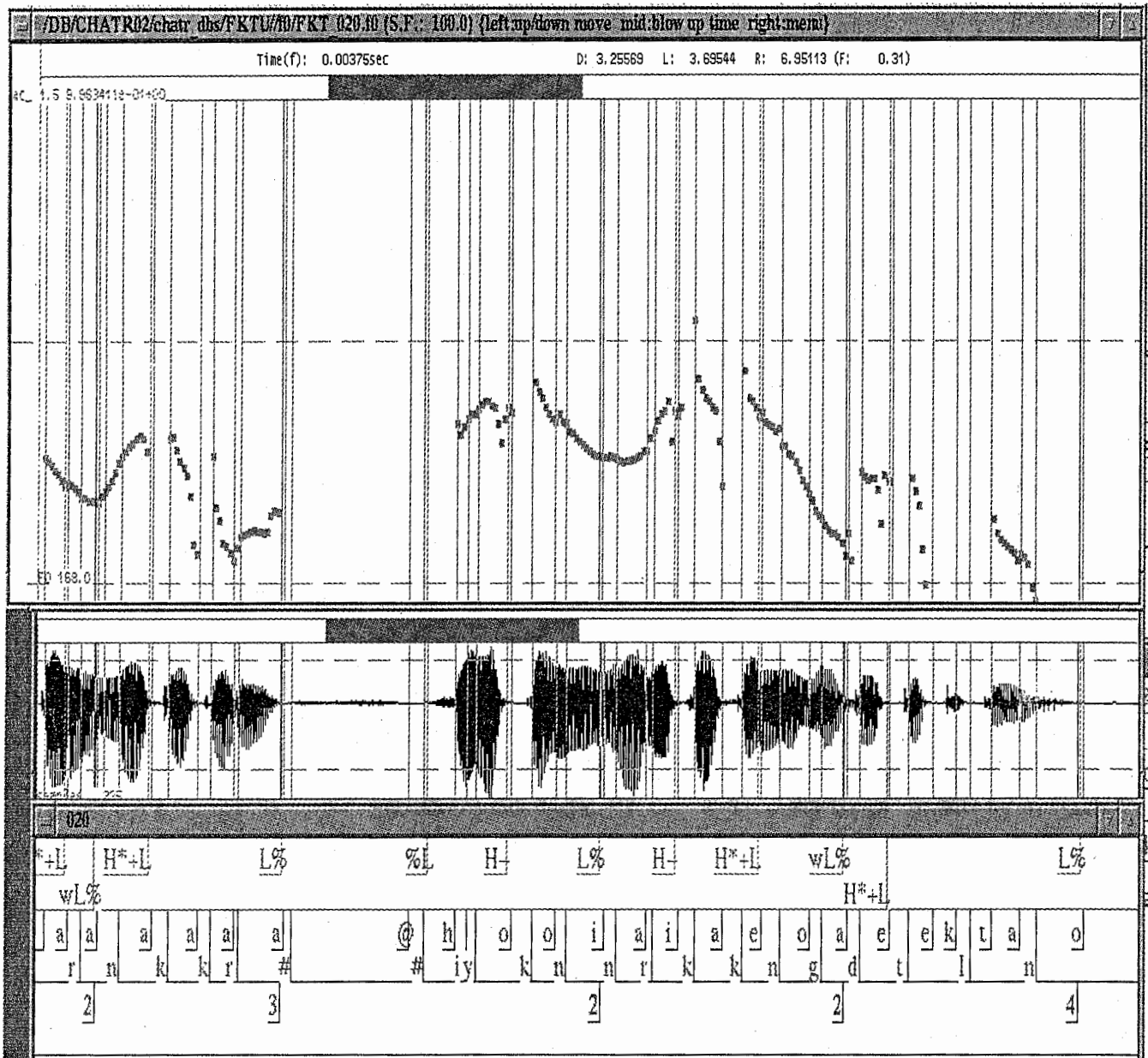


“子供の時分、見ちゃったのよ” (関西)



“子供の時分、見ちゃったのよ” (共通語)



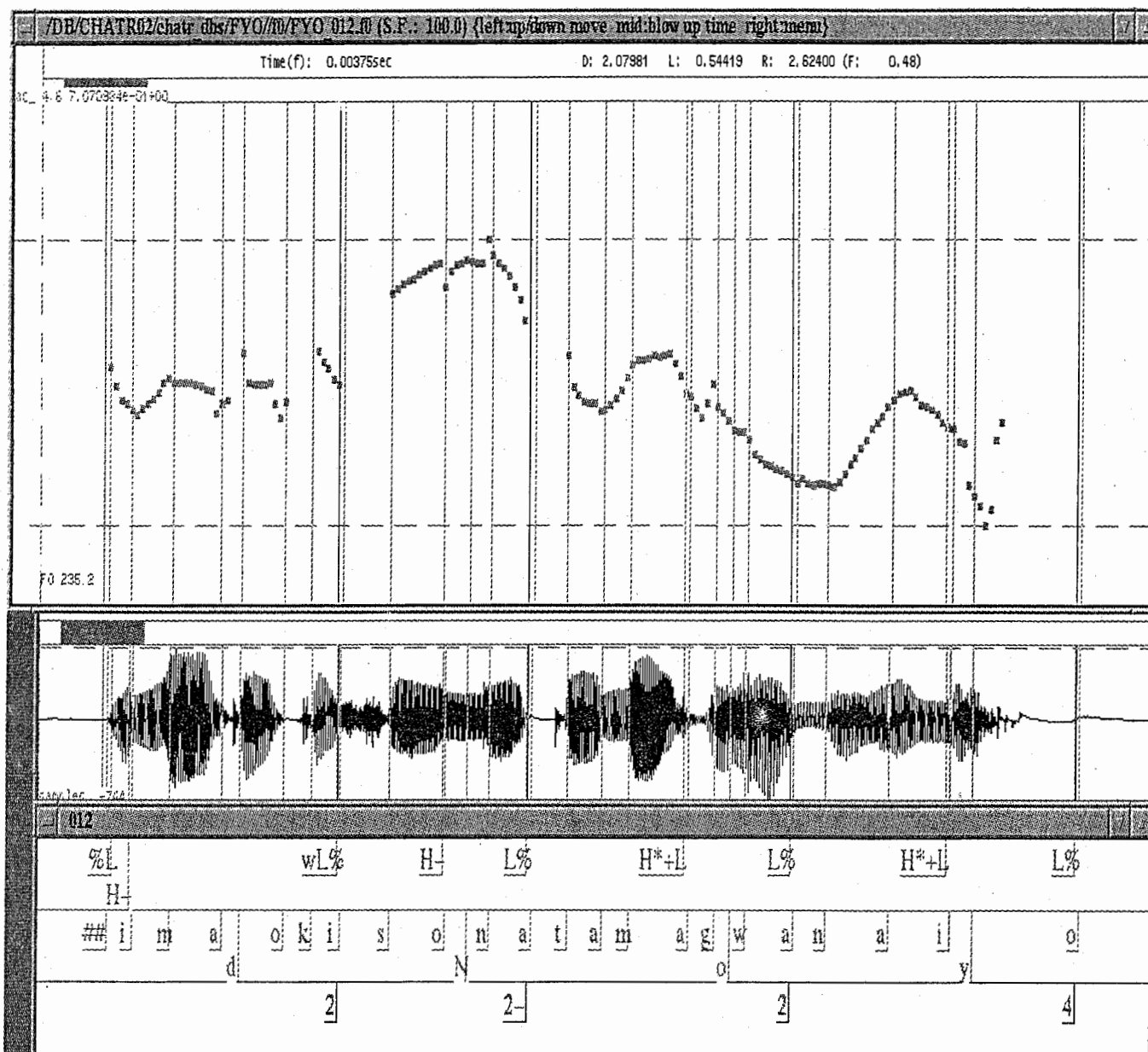


“なかから、ヒヨコになりかけのが出て来たの” (共通語)

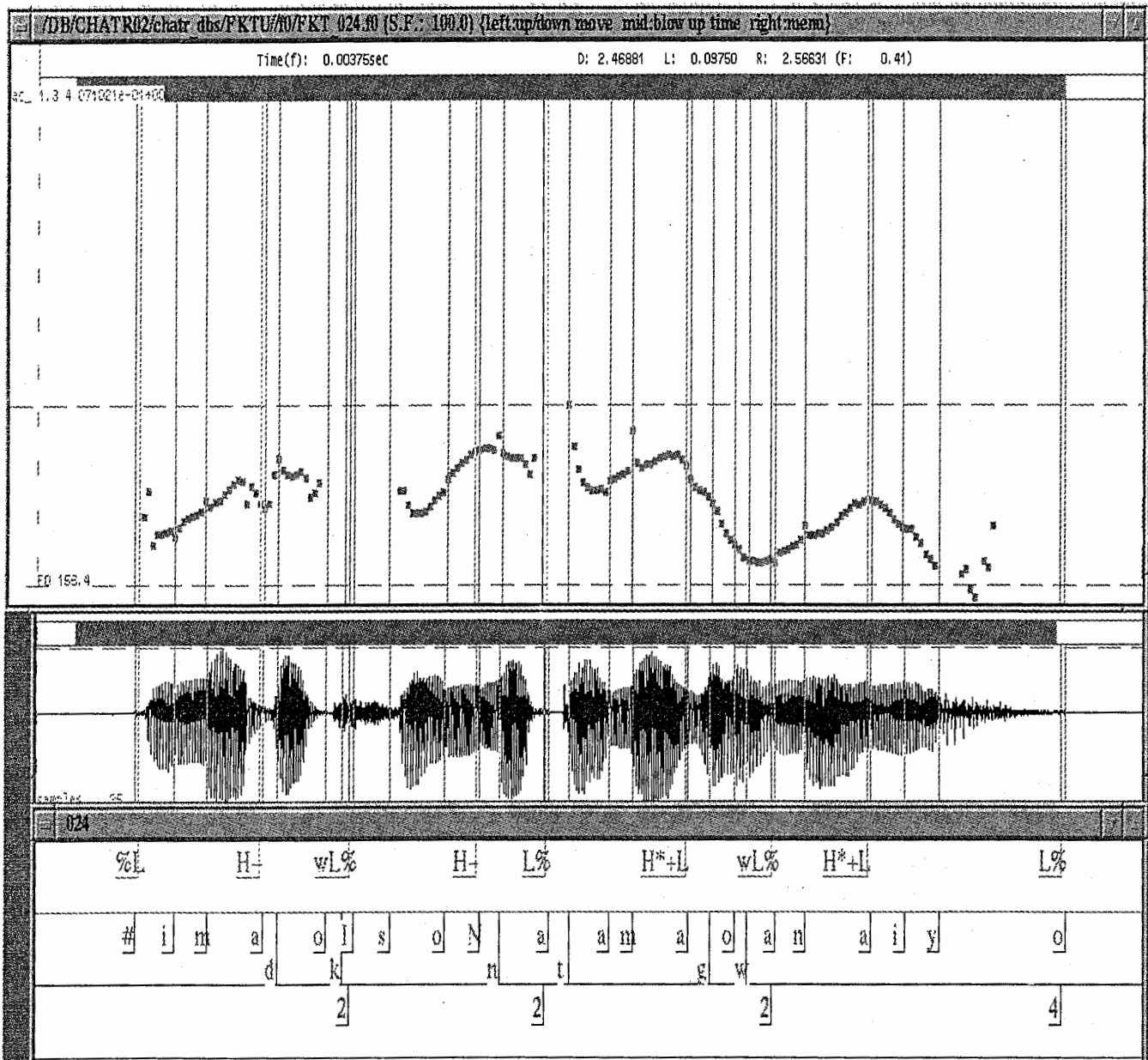




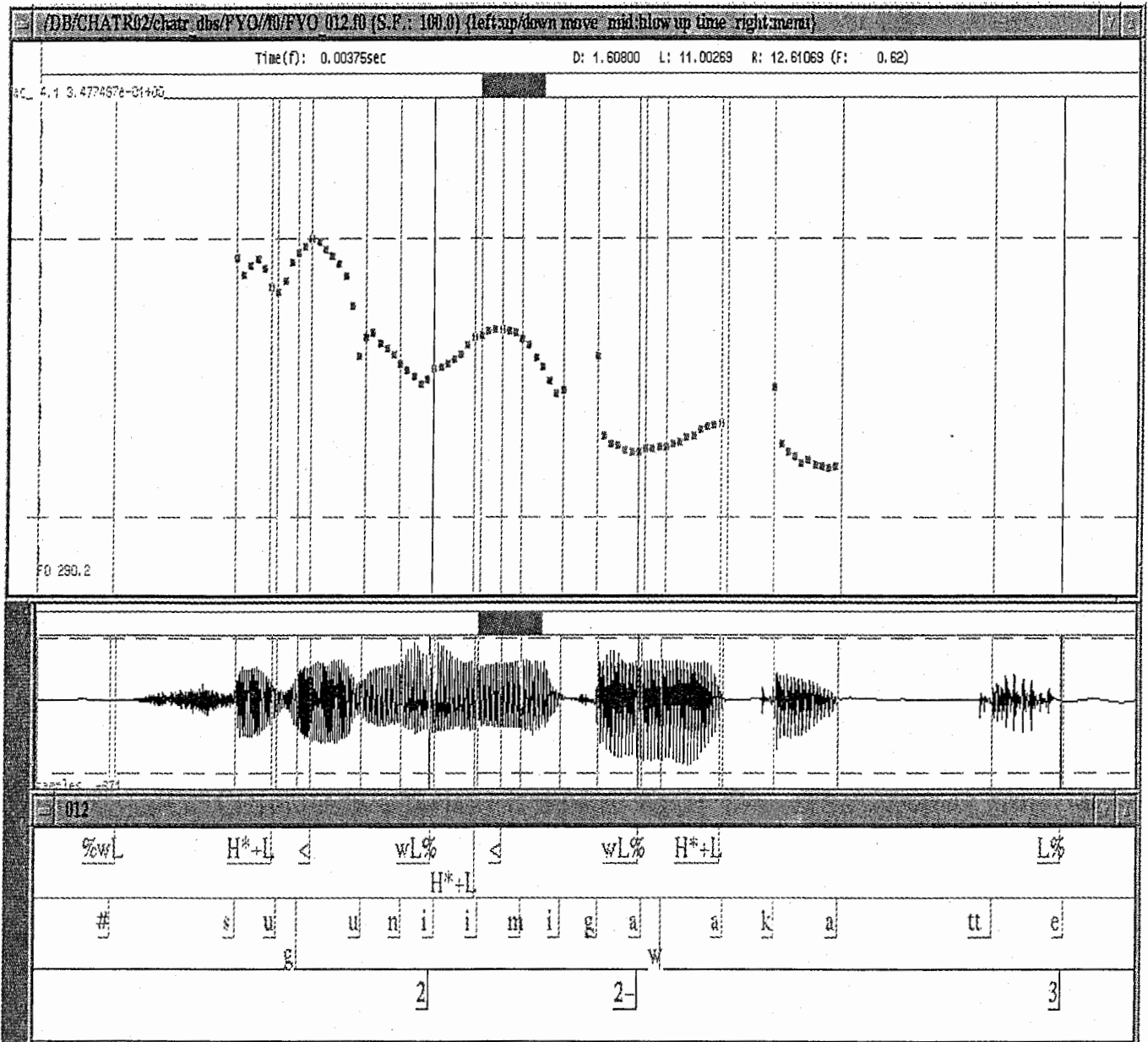




“今どきそんな卵はないよ” (関西)



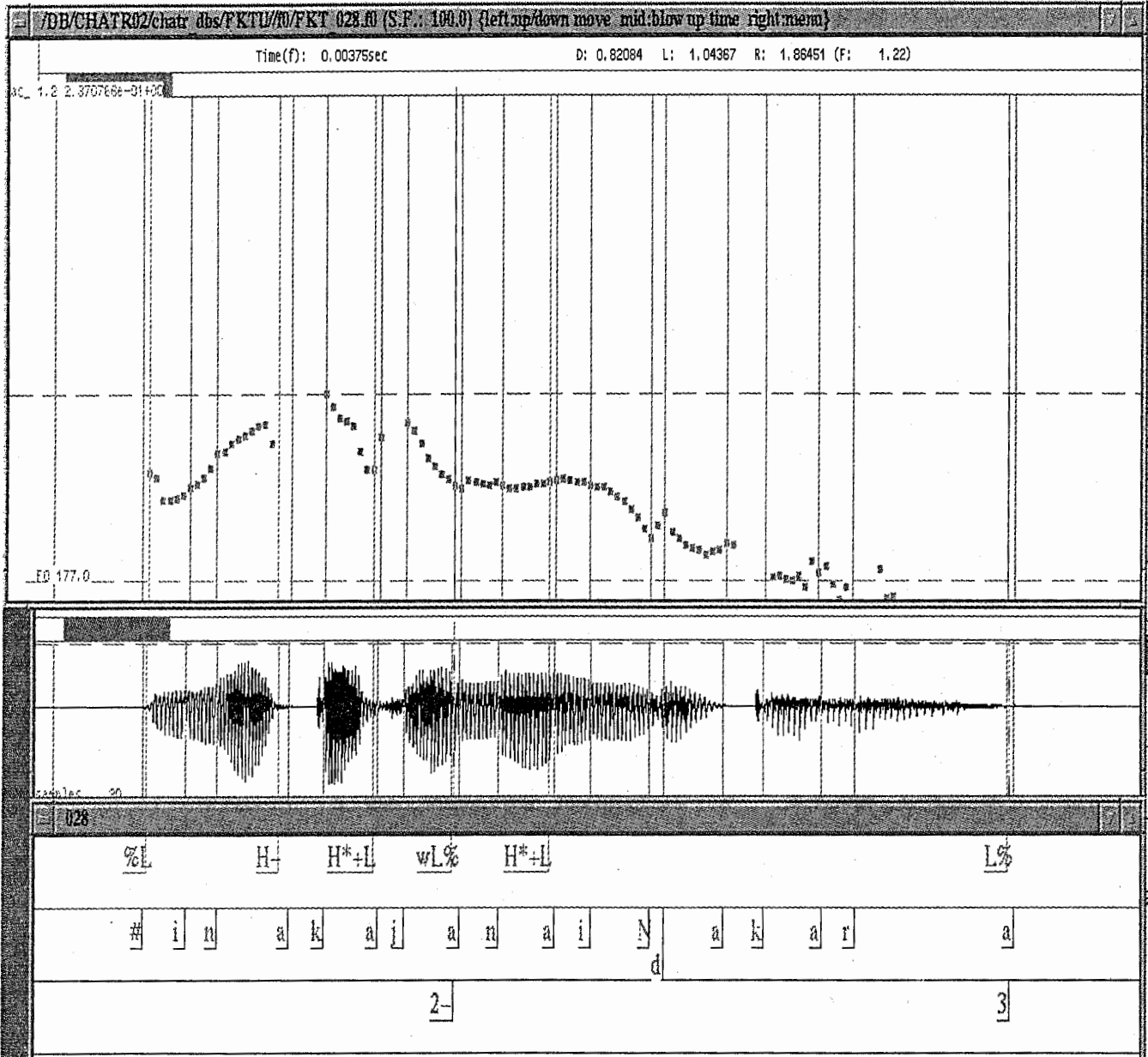
“今どきそんな卵はないよ” (共通語)



“すぐに意味が判って” (関西)

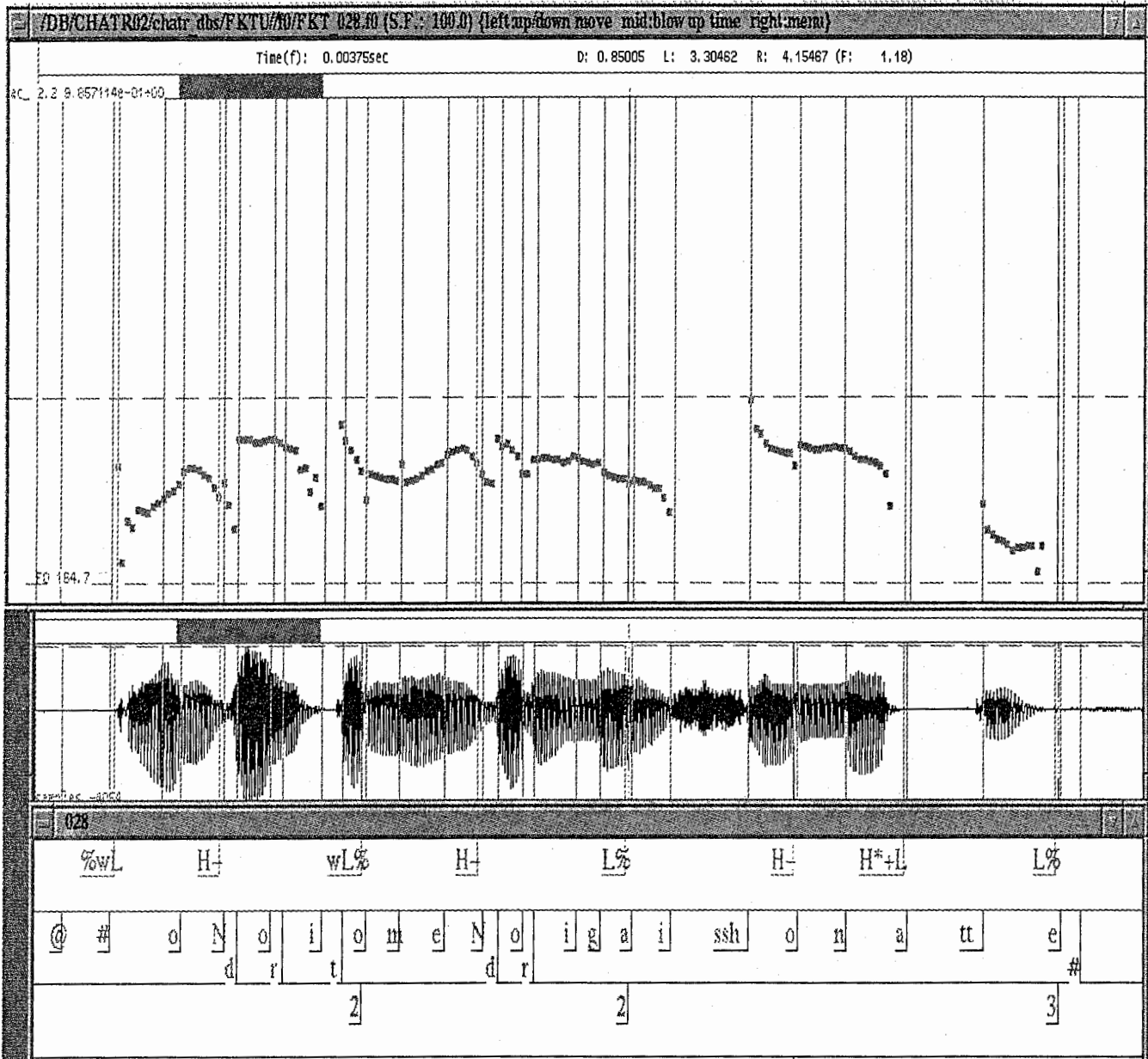






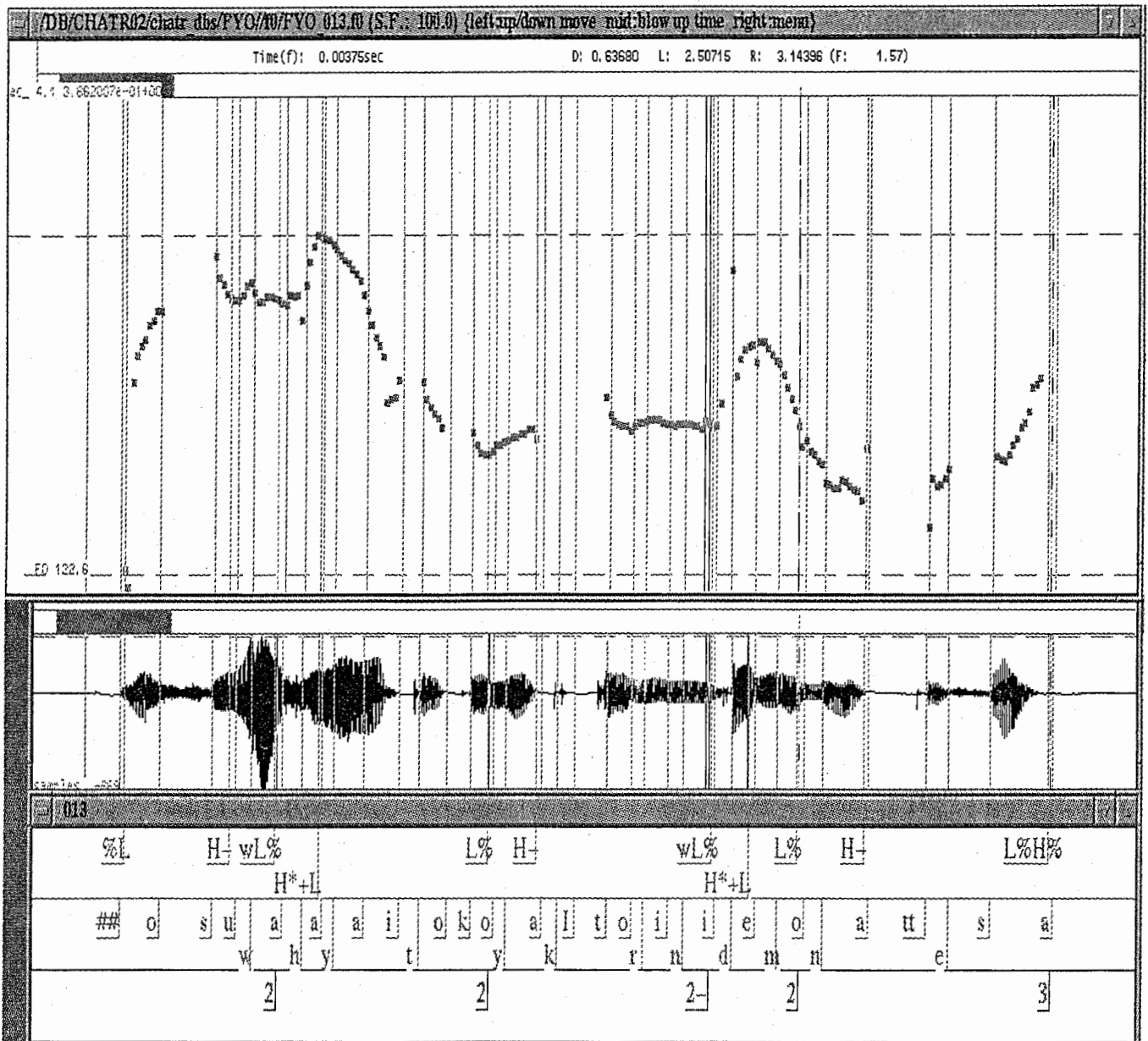
“田舎じゃないんだから” (共通語)



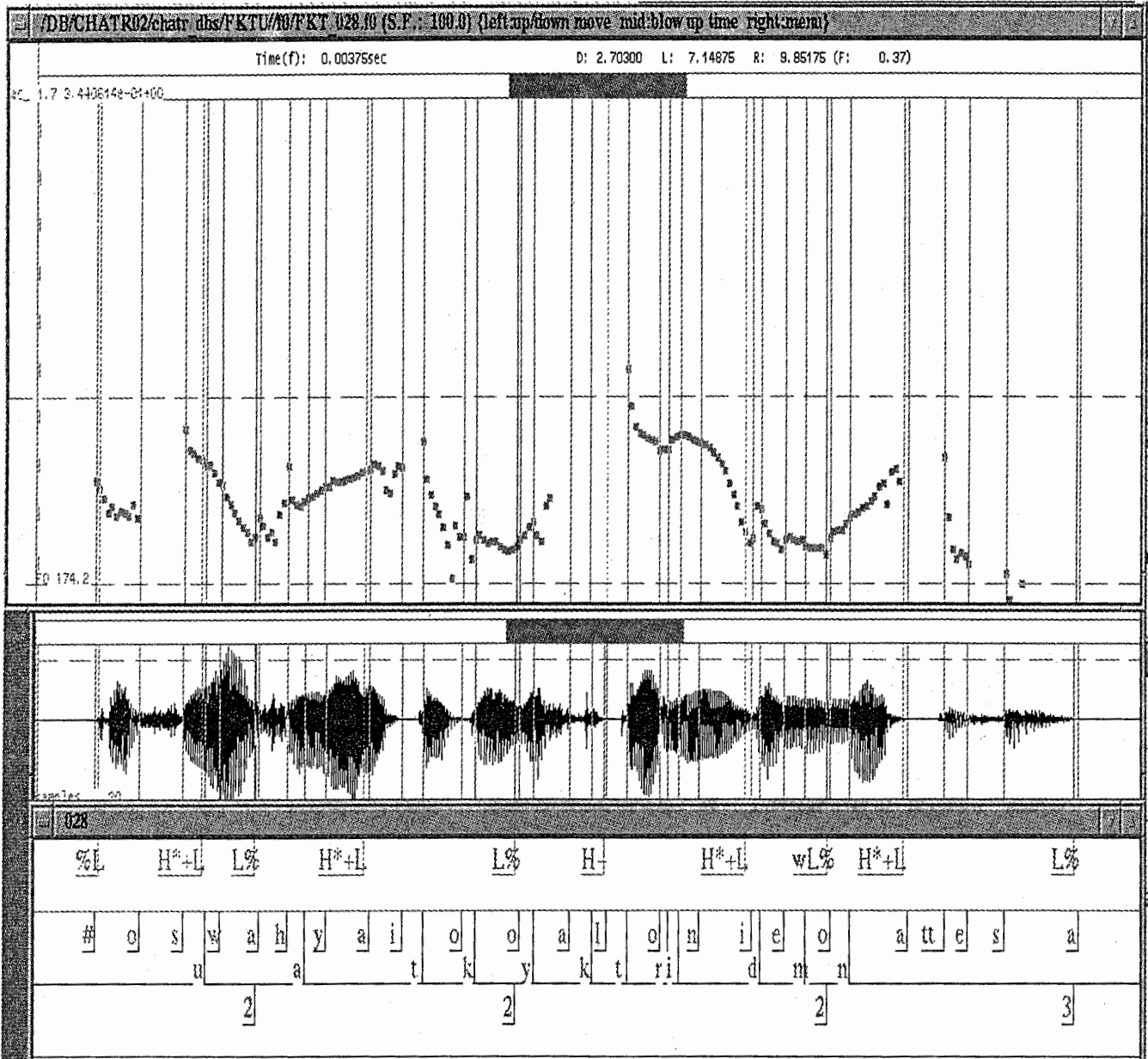


“雄鶏と雌鶏が一緒になって” (共通語)

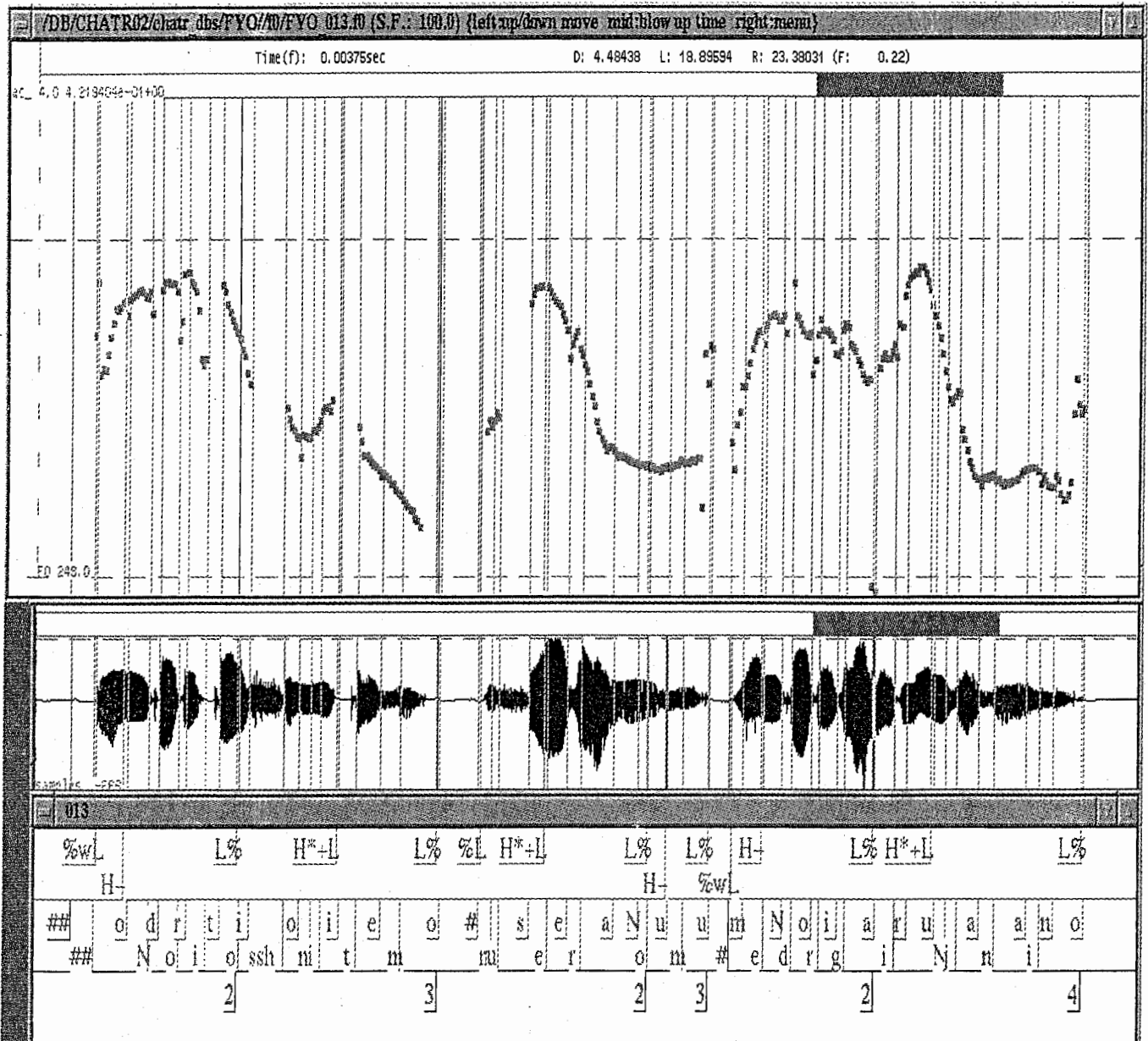




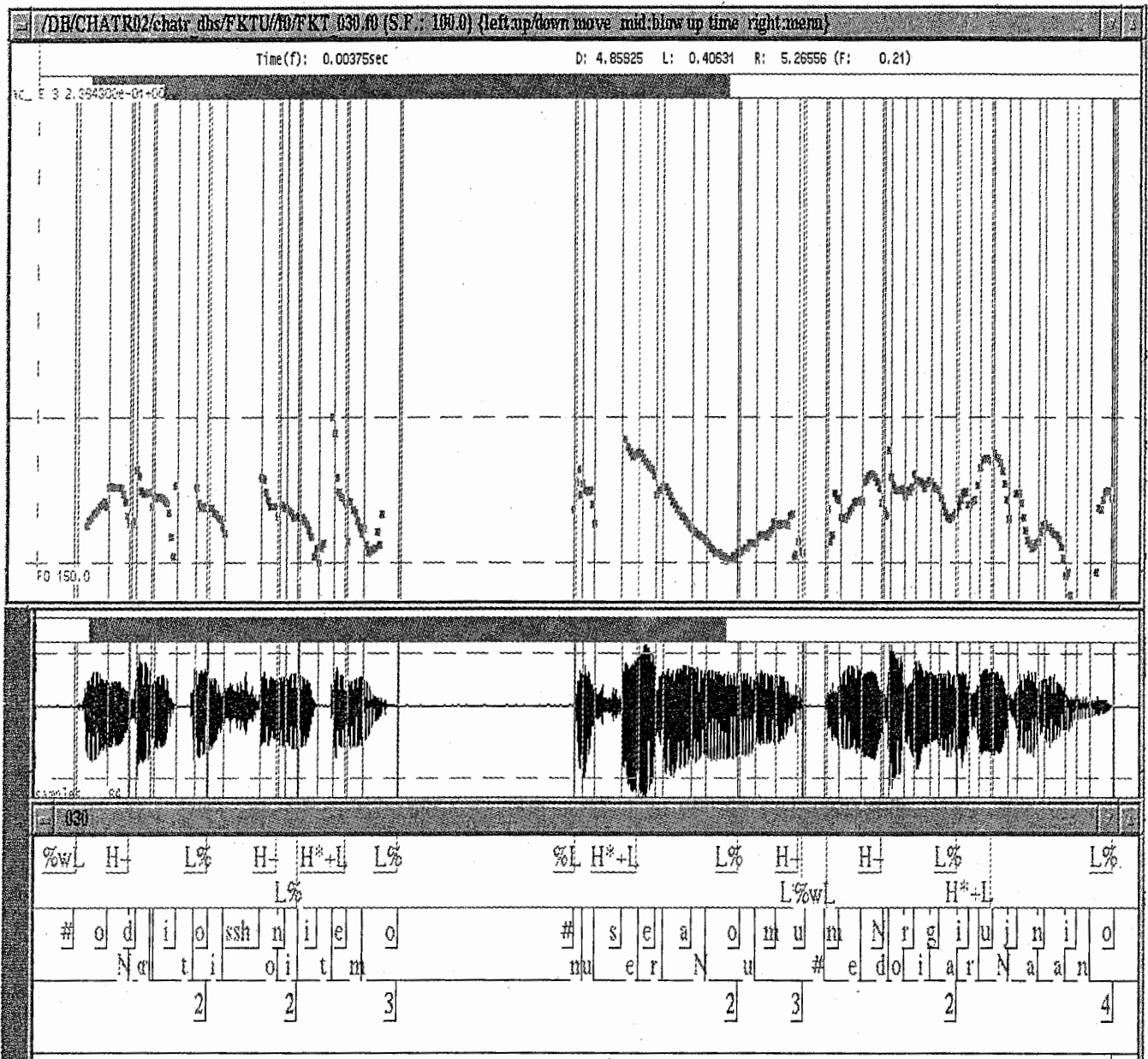
“雄は早いとこ焼鳥にでもなってさ” (関西)



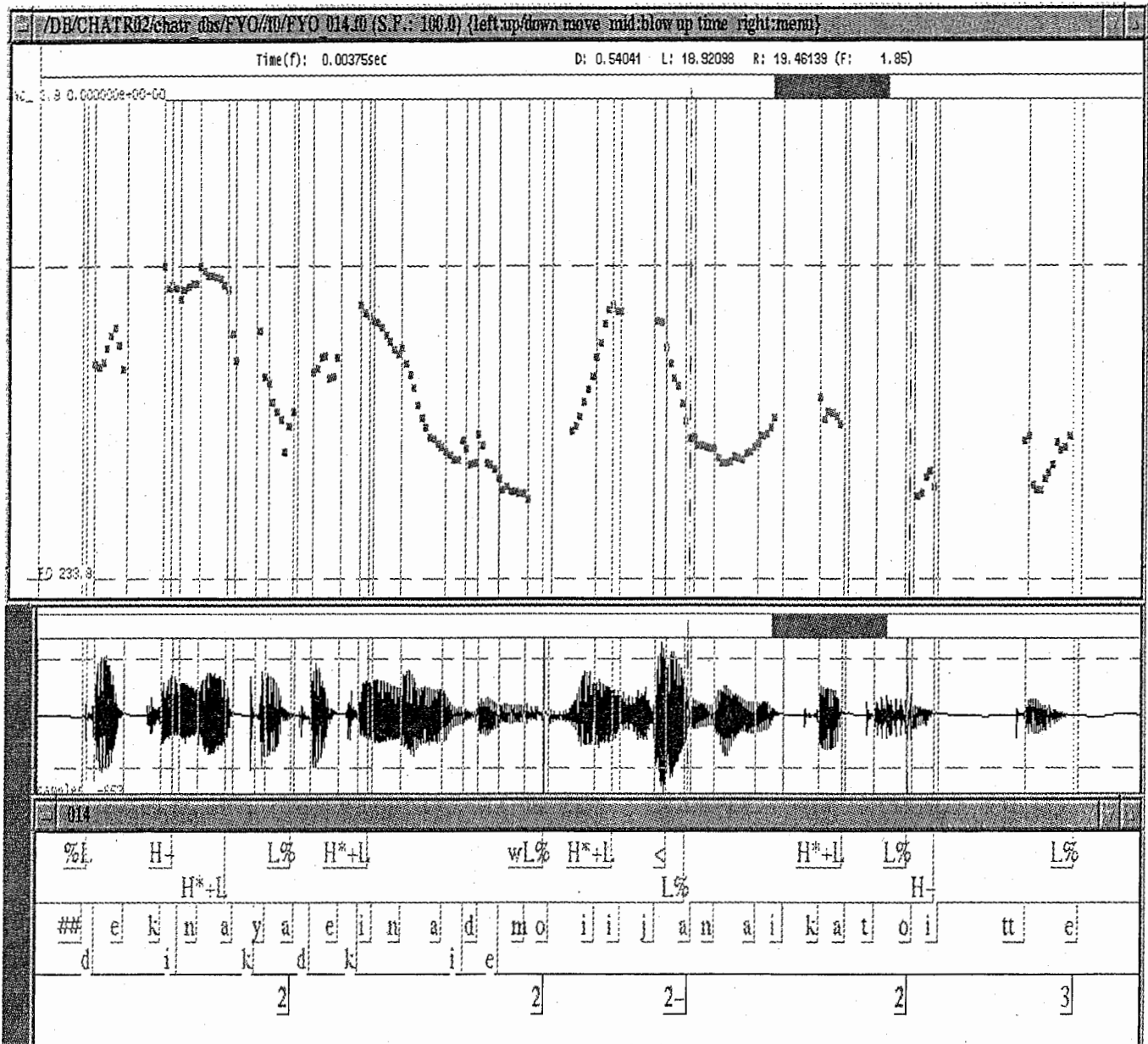
“雄は早いところ焼鳥にでもなつてさ” (共通語)



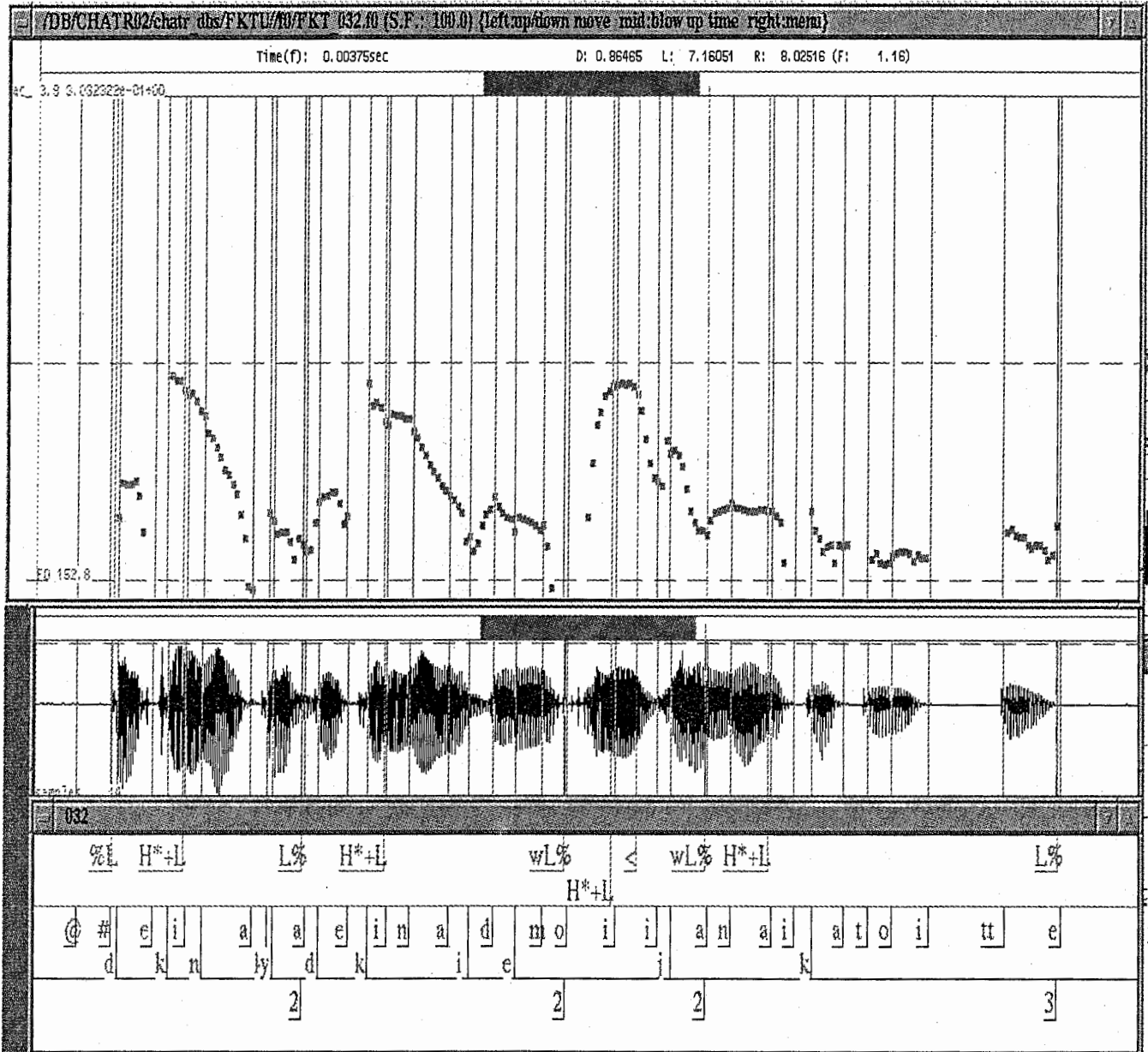
“雄鶏と一緒にいても、無精卵を産む雌鶏がいるんじゃないの” (関西)



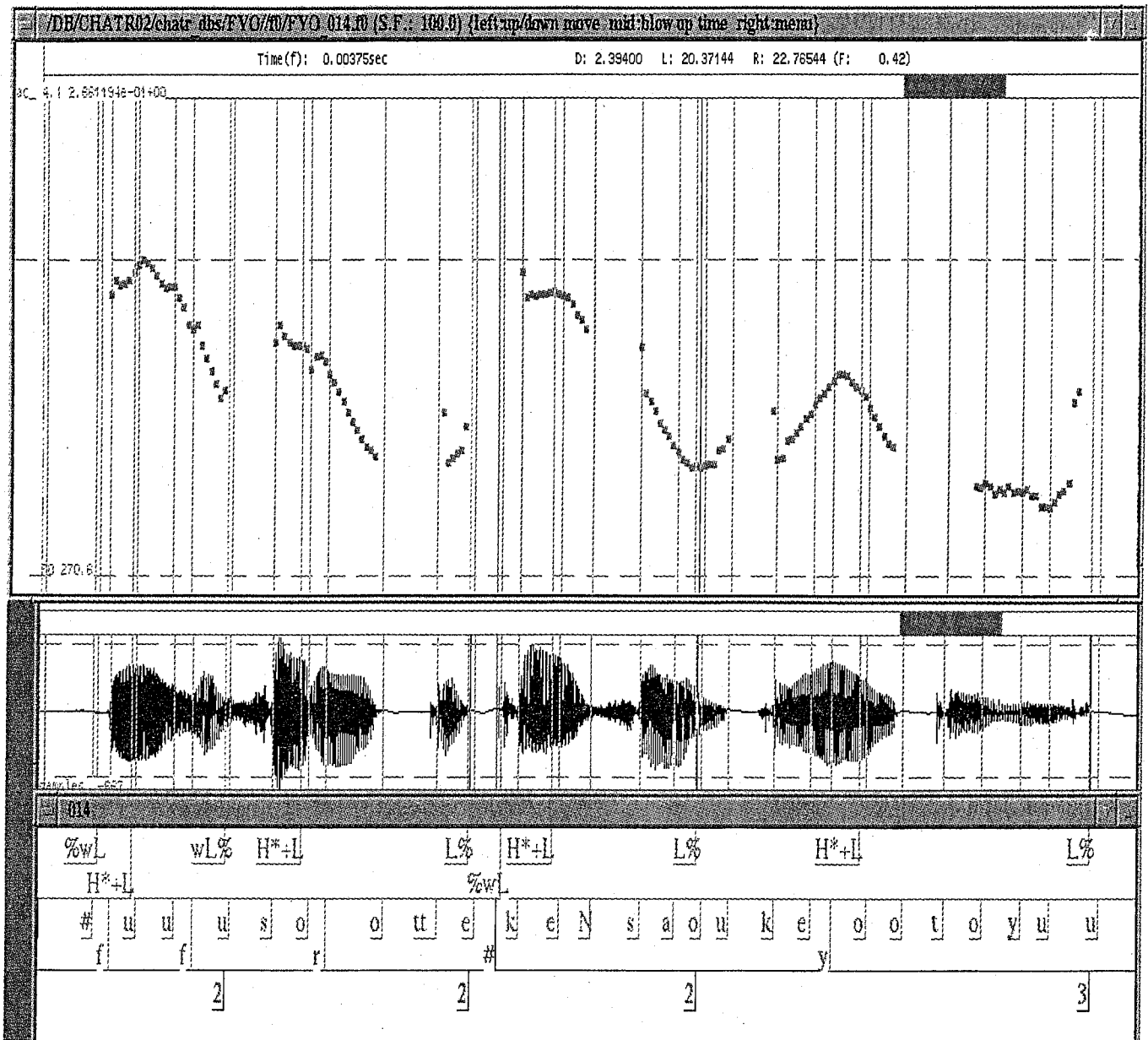
“雄鶏と一緒にいても、無精卵を産む雌鶏がいるんじゃないの” (共通語)



“出来なきゃ出来ないでもいいじゃないかと言って” (関西)



“出来なきゃ出来ないでもいいじゃないかと言って” (共通語)

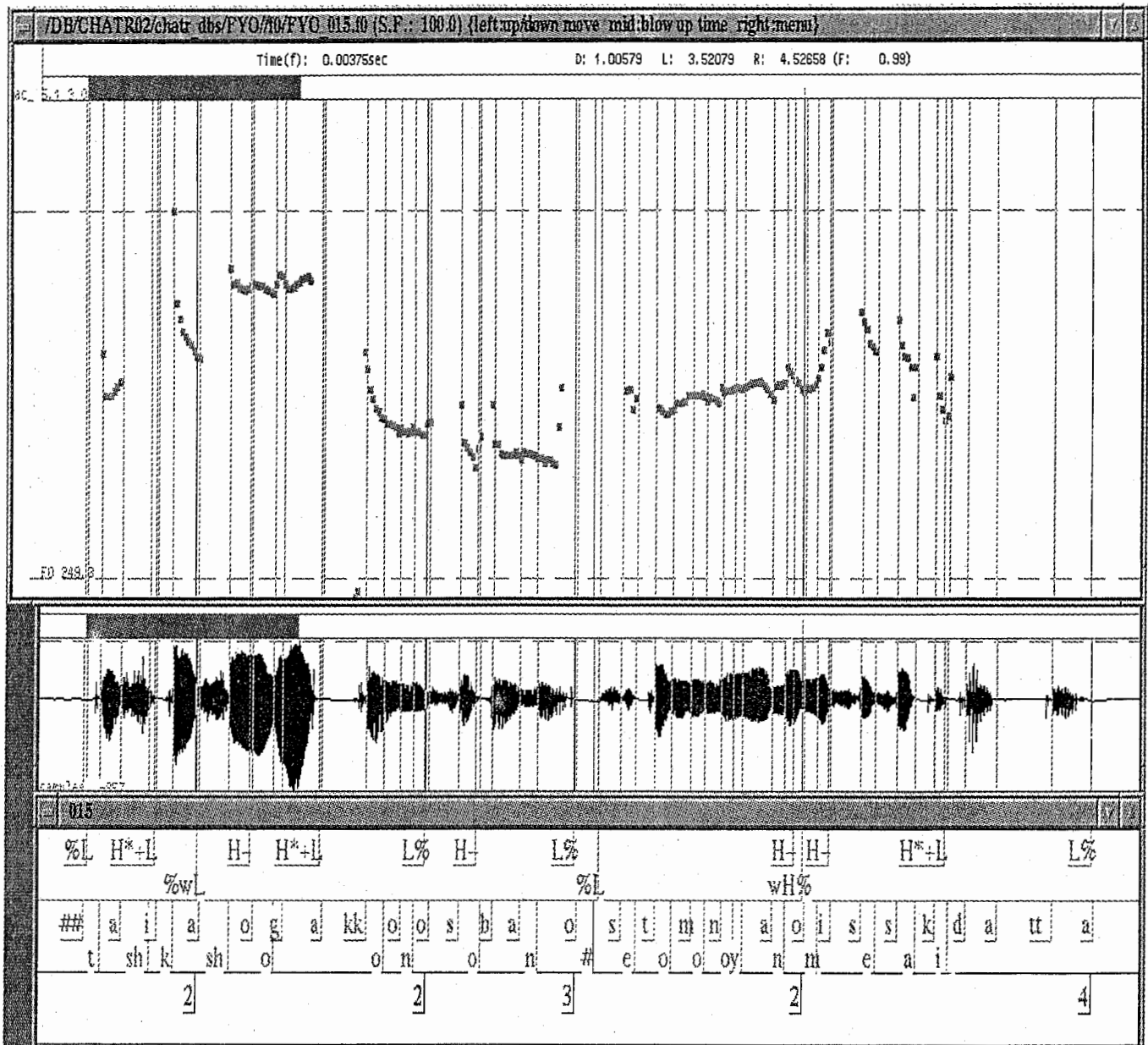


“夫婦揃って検査を受けようという” (関西)

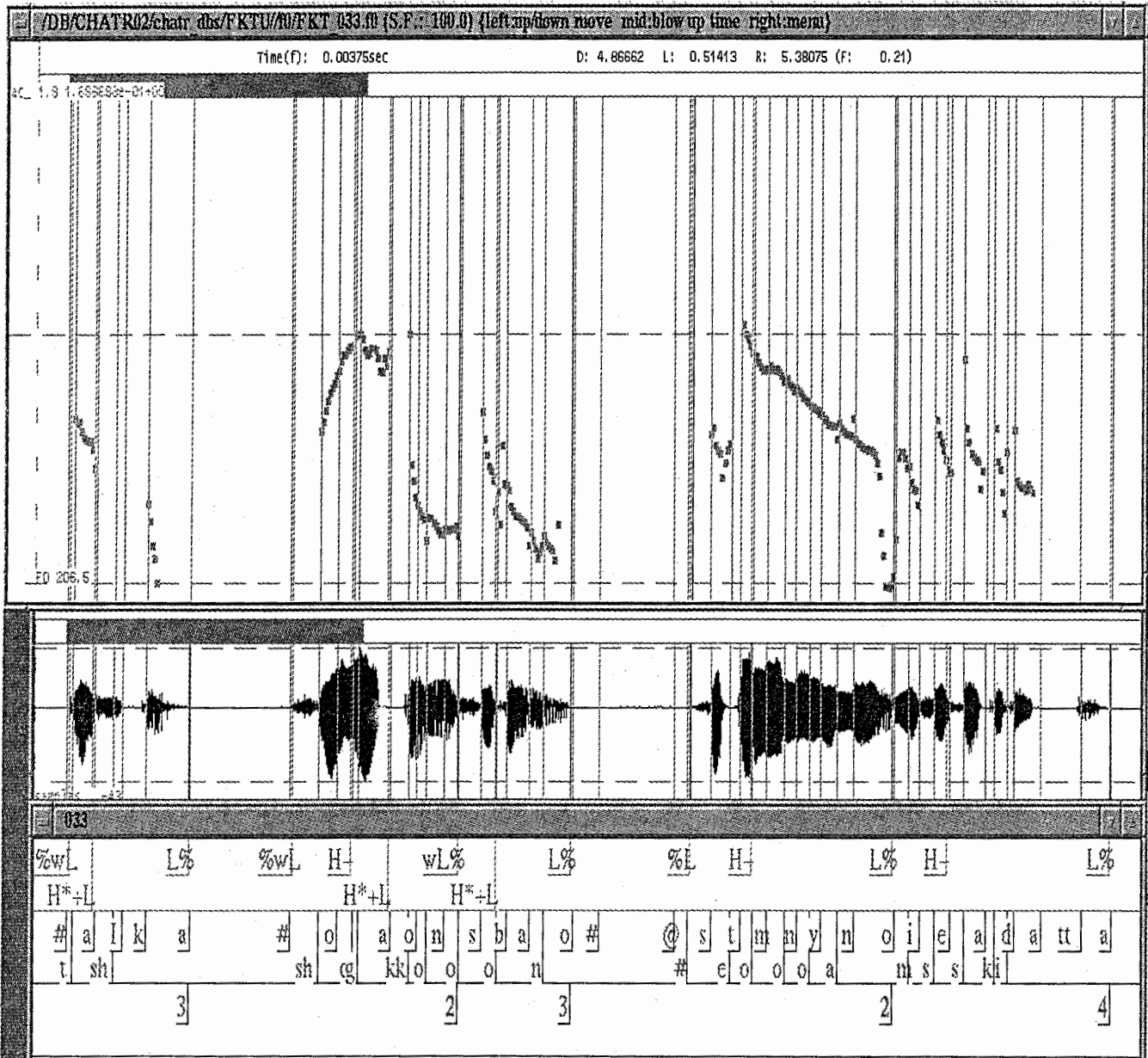




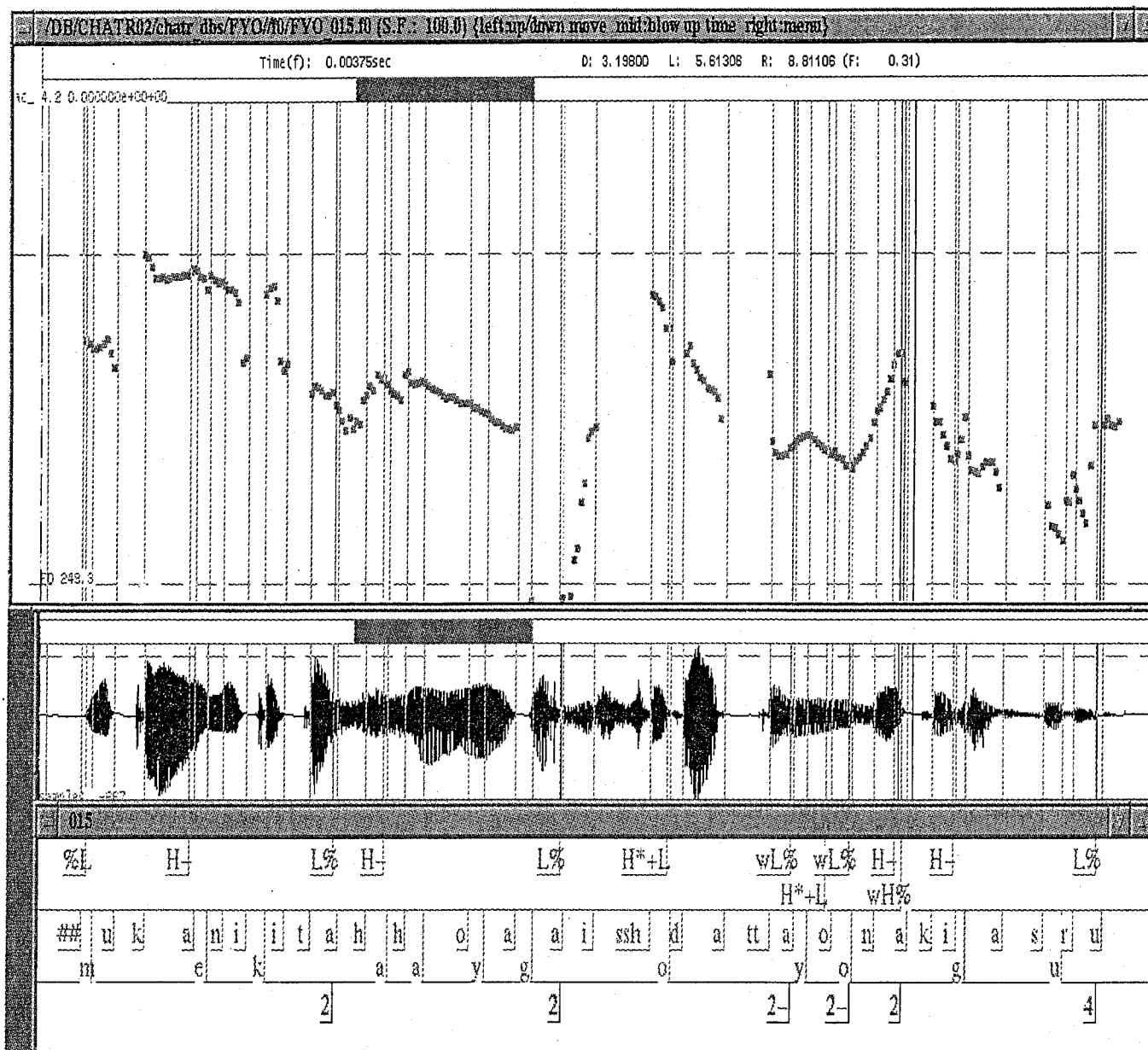




“たしか小学校のそばの瀬戸物屋の店先だった” (関西)

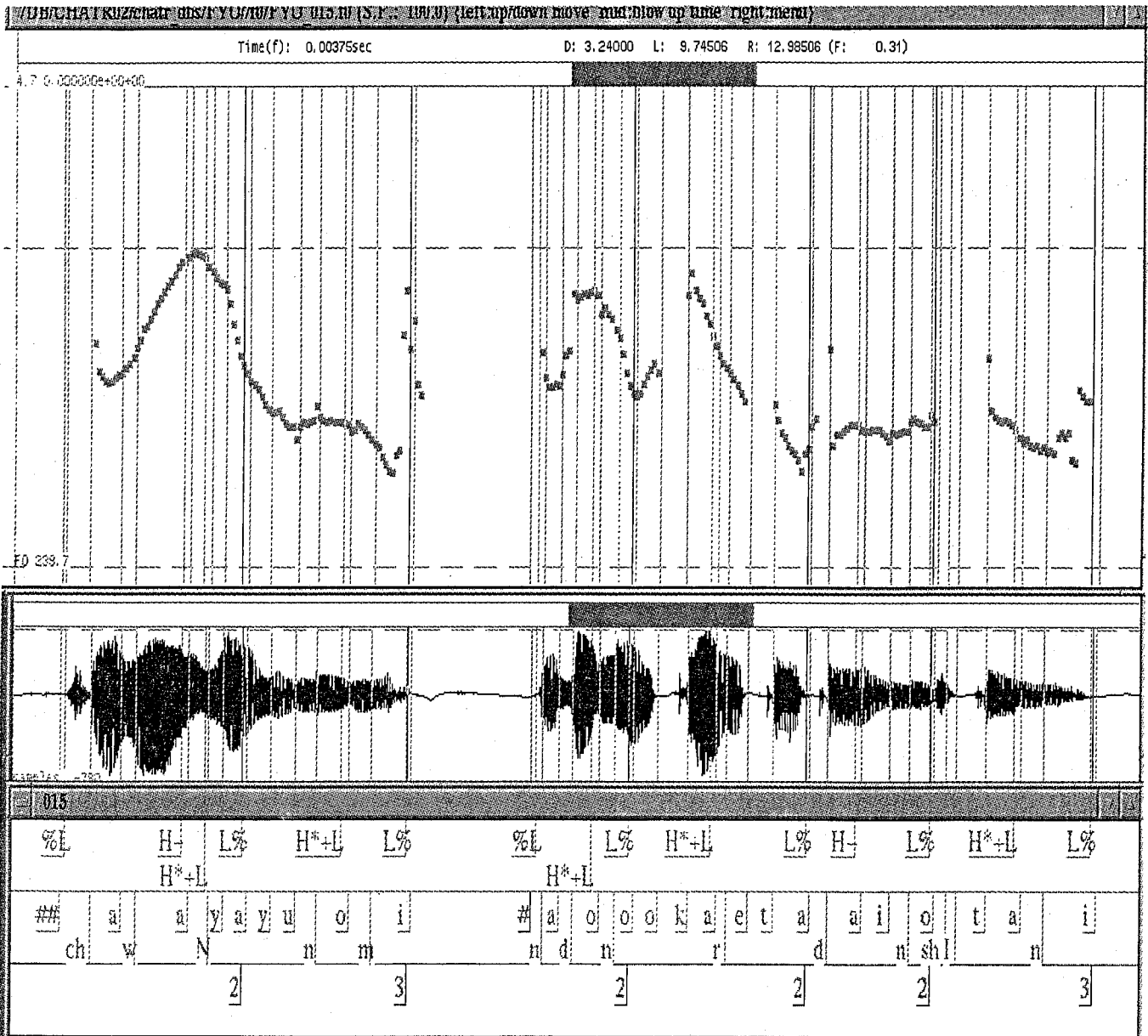


“たしか小学校のそばの瀬戸物屋の店先だった” (共通語)

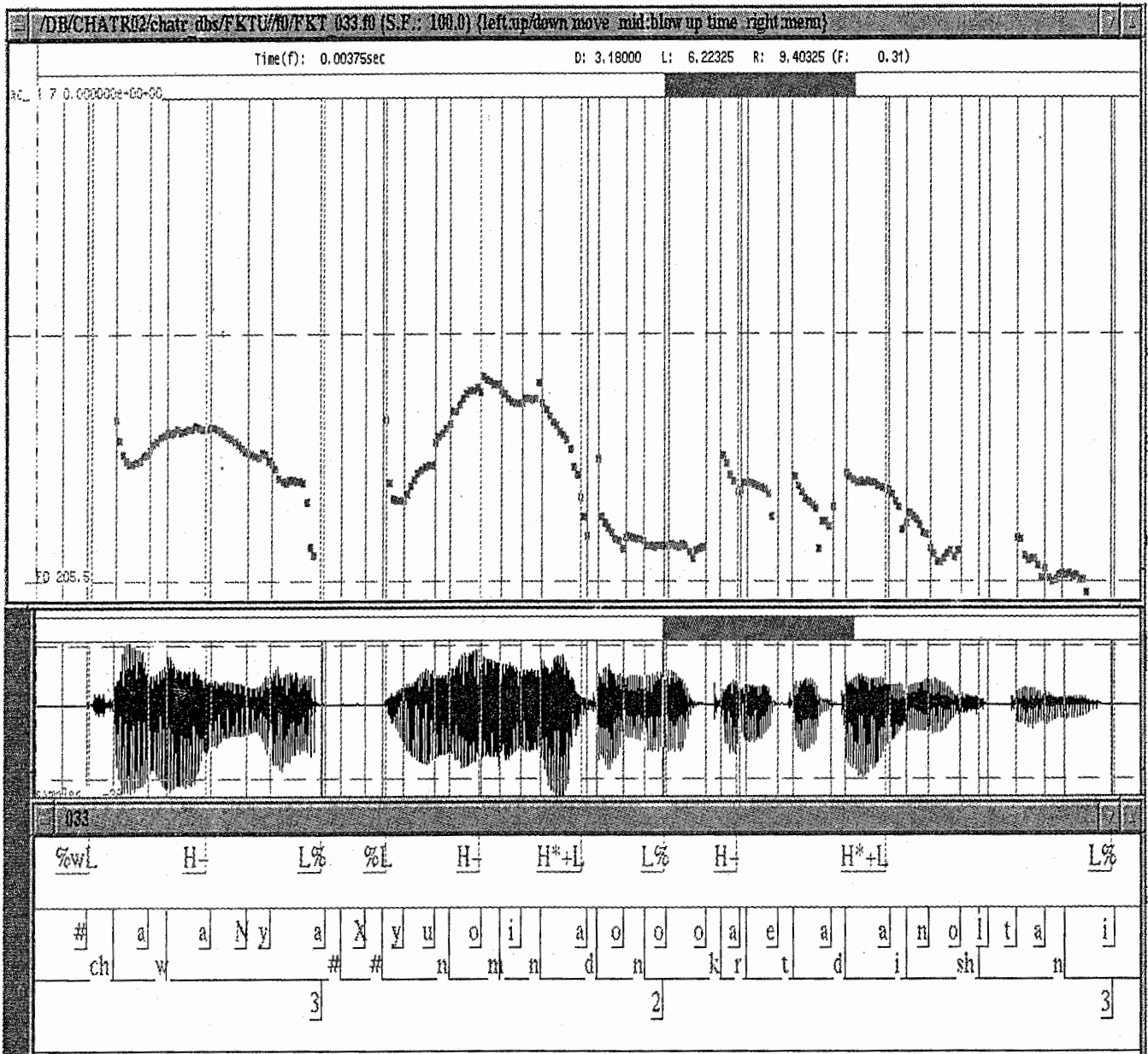


“迎えに来た母親が一緒だったような気がする” (関西)

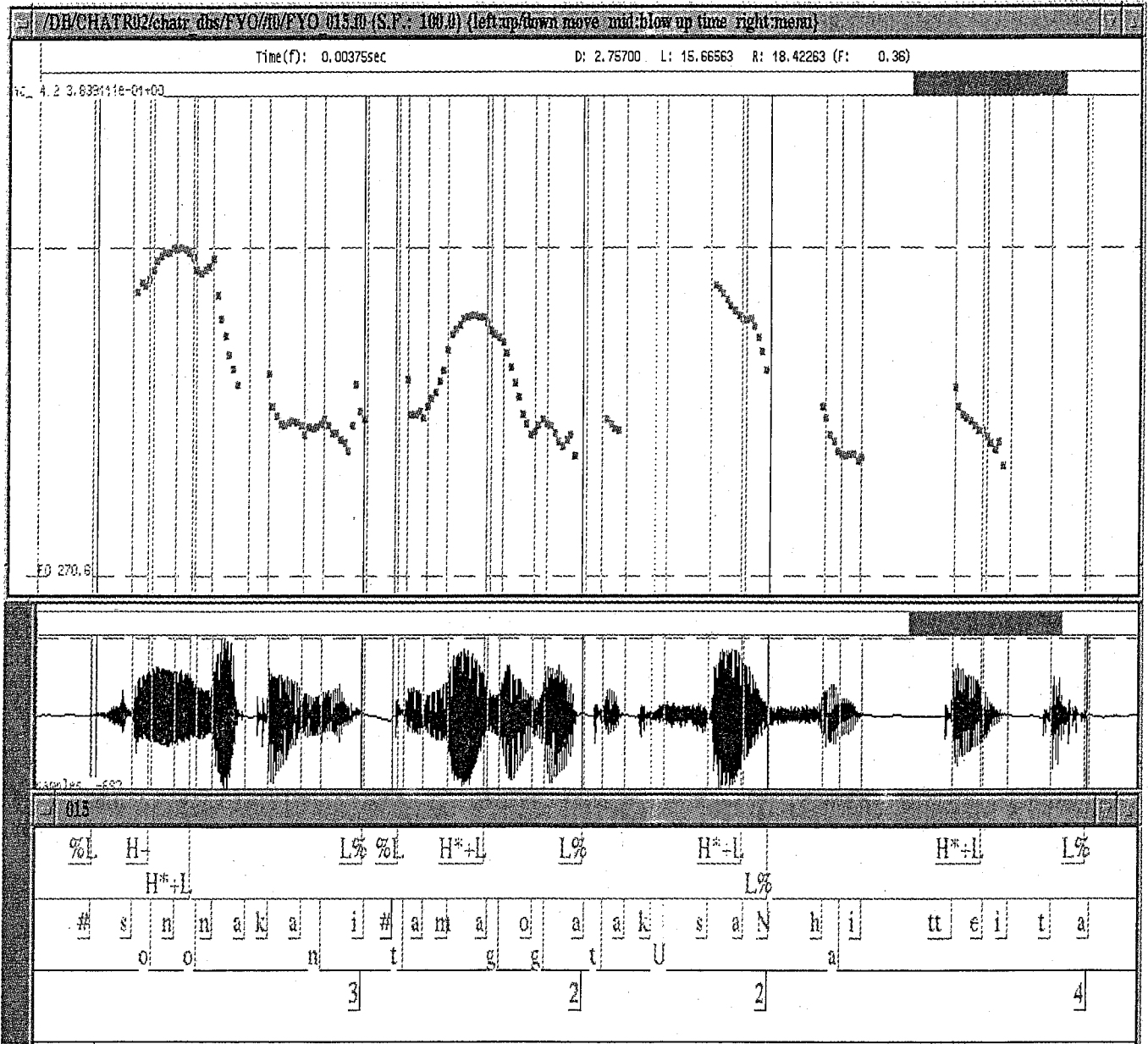




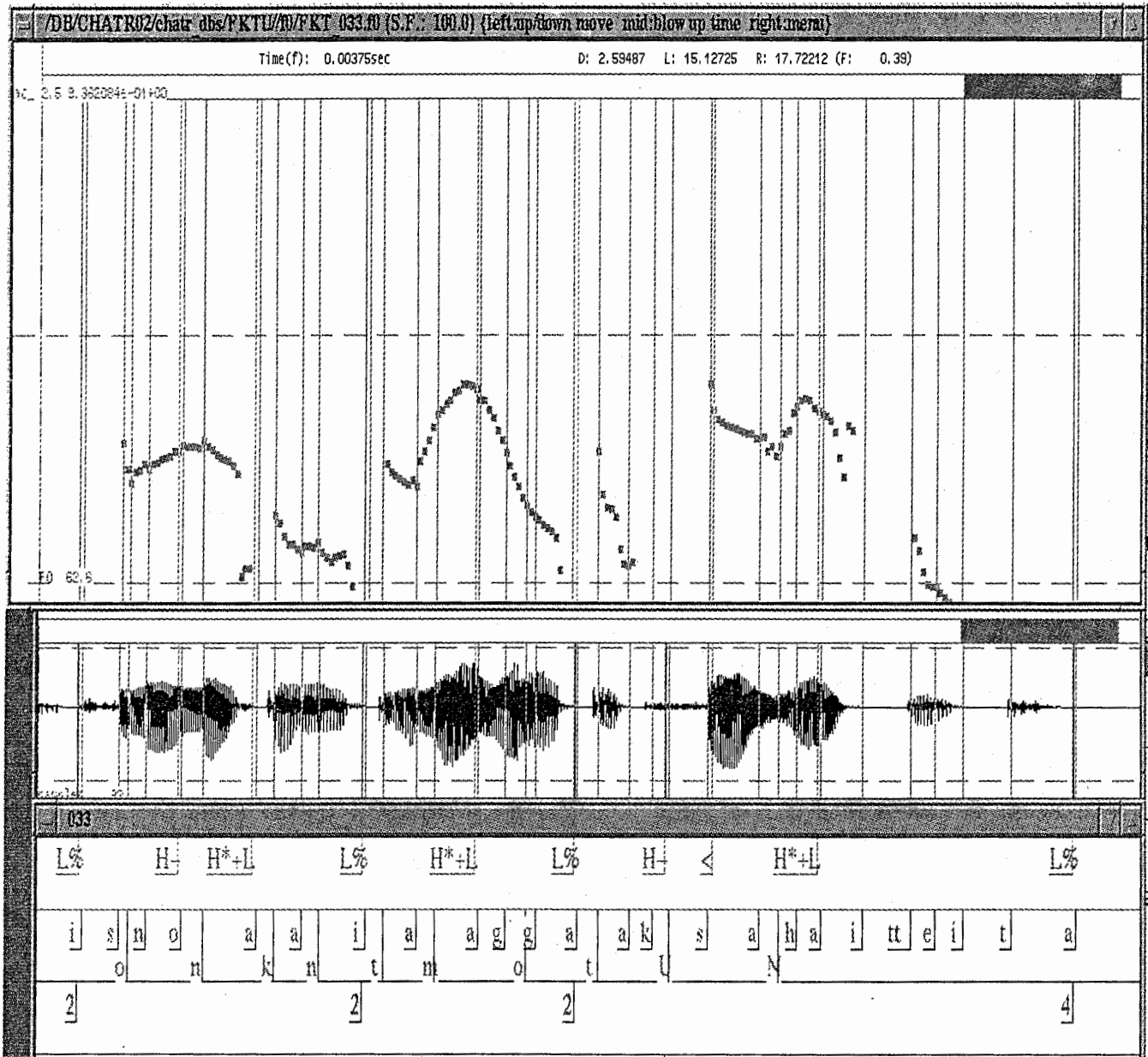
“茶碗や湯呑み、などの置かれた台の下に” (関西)



“茶碗や湯呑み、などの置かれた台の下に” (共通語)



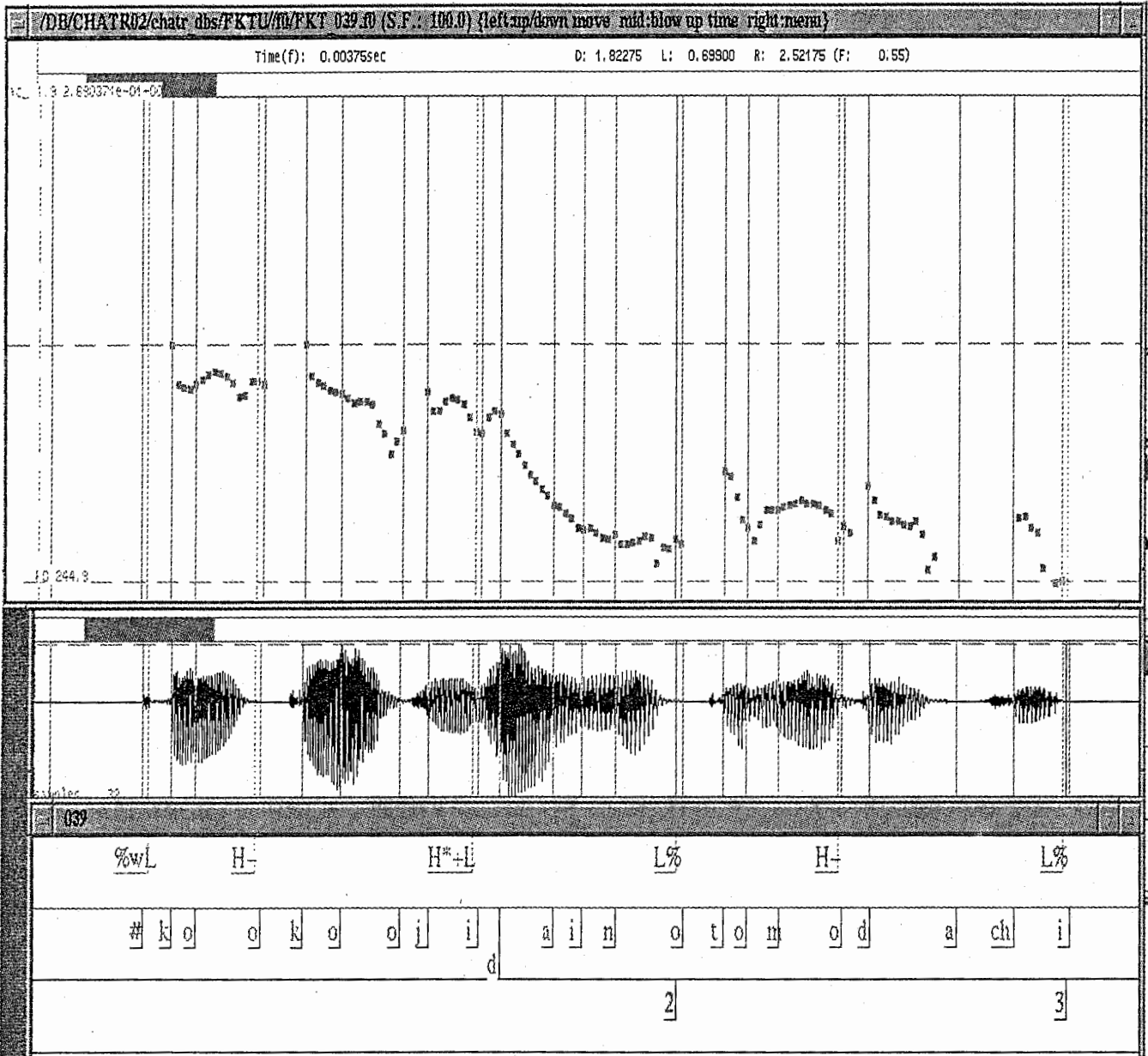
“その中に卵が沢山入っていた” (関西)



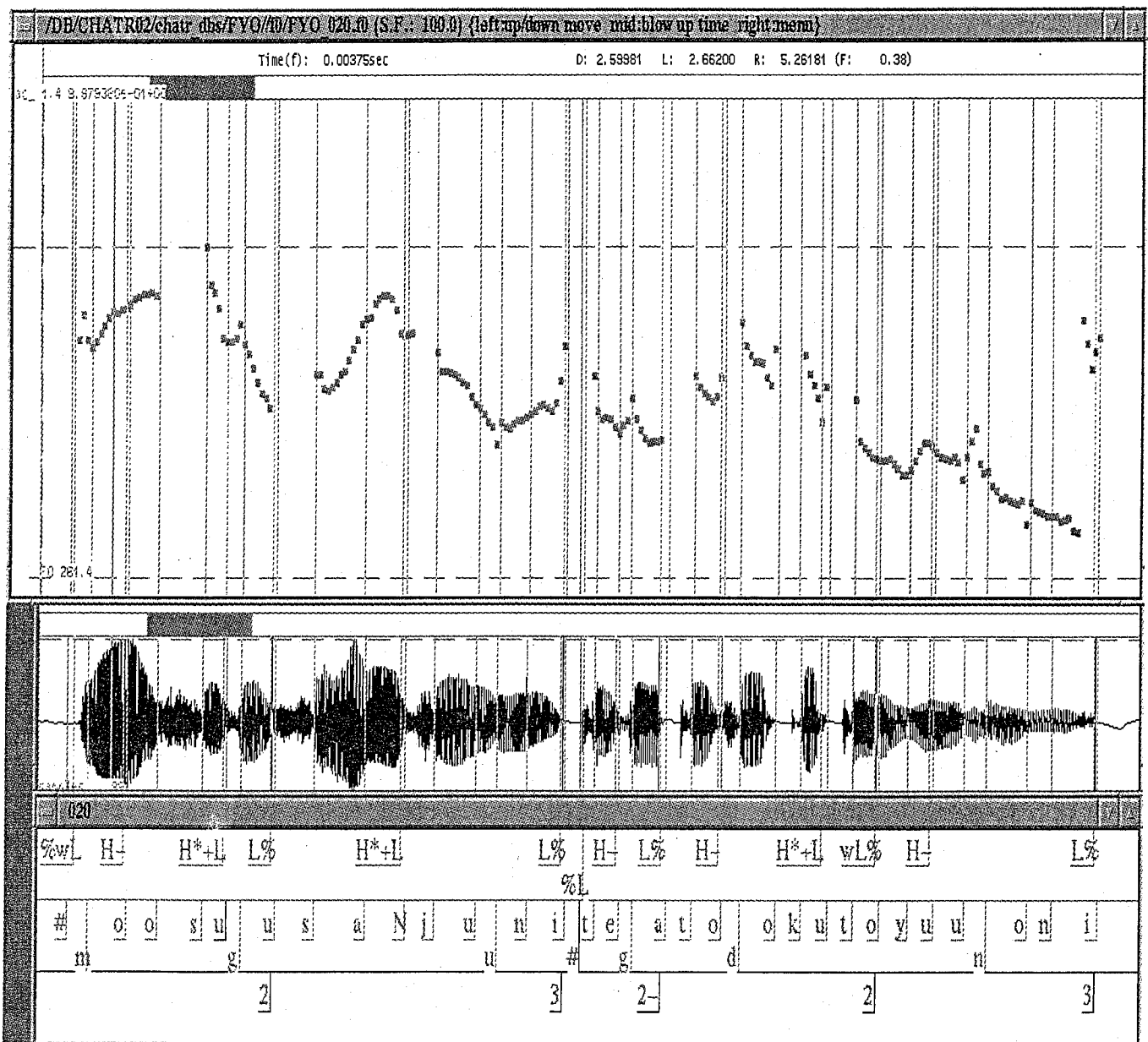
“その中に卵が沢山入っていた” (共通語)



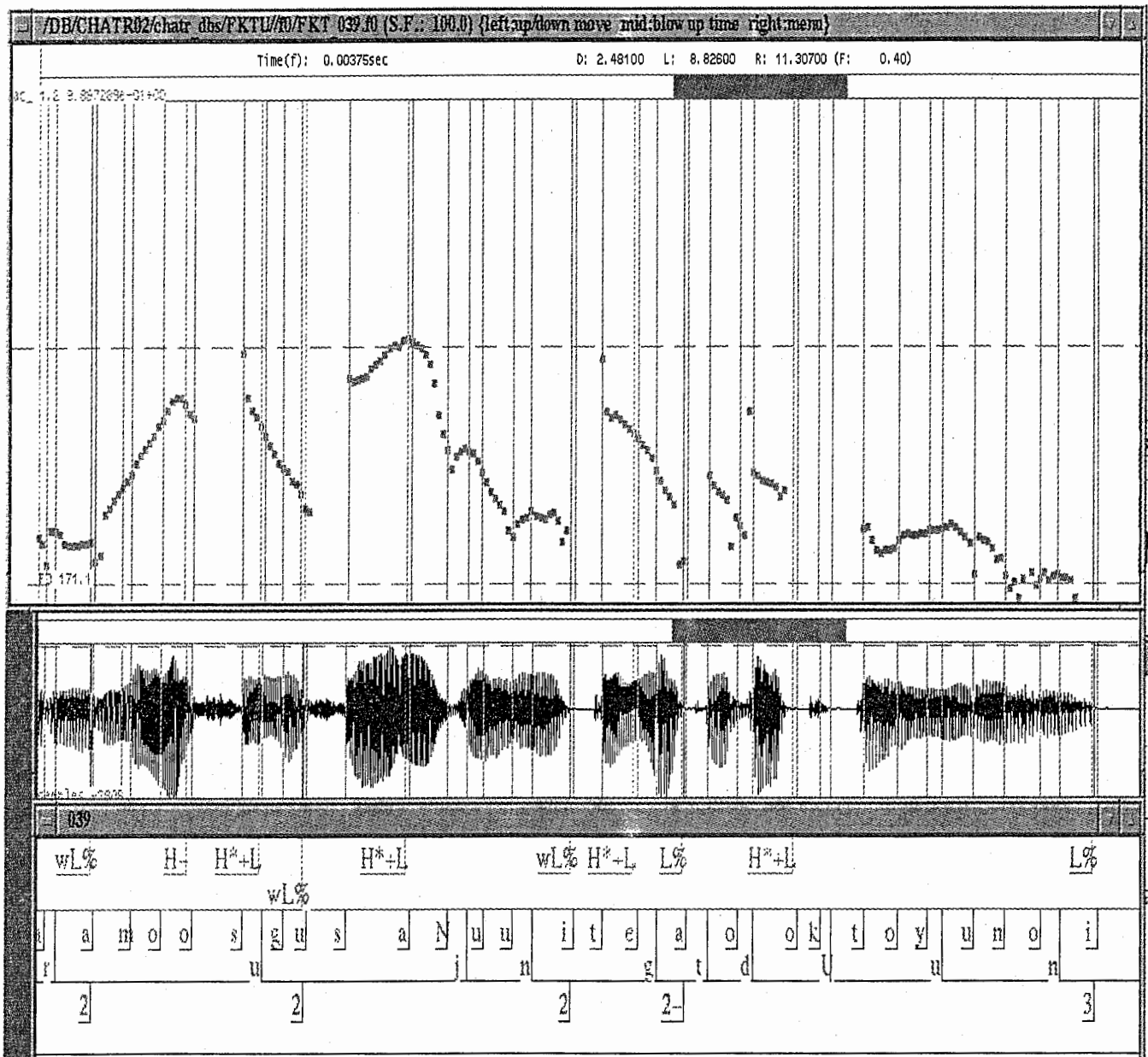




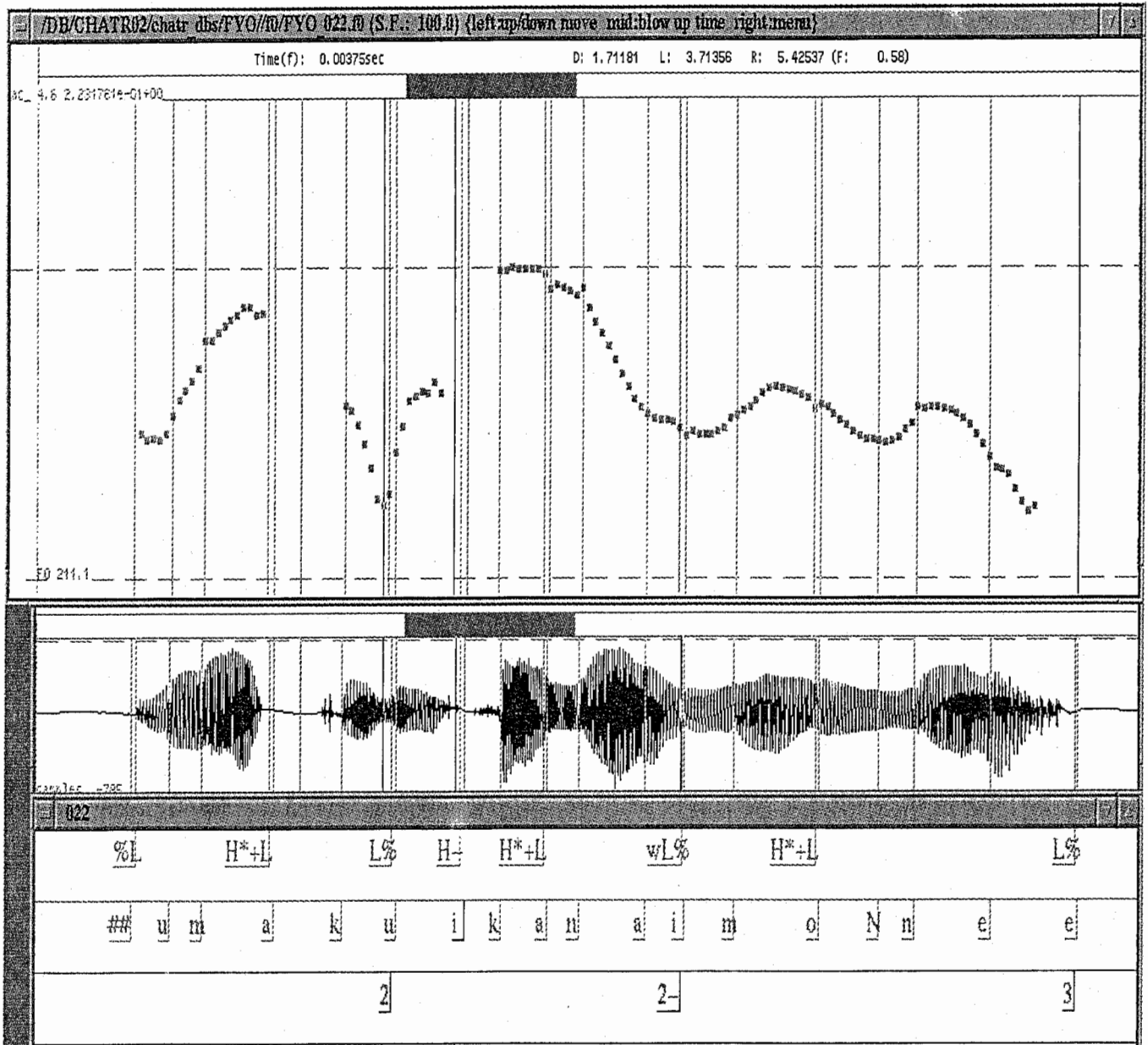
“高校時代の友達” (共通語)



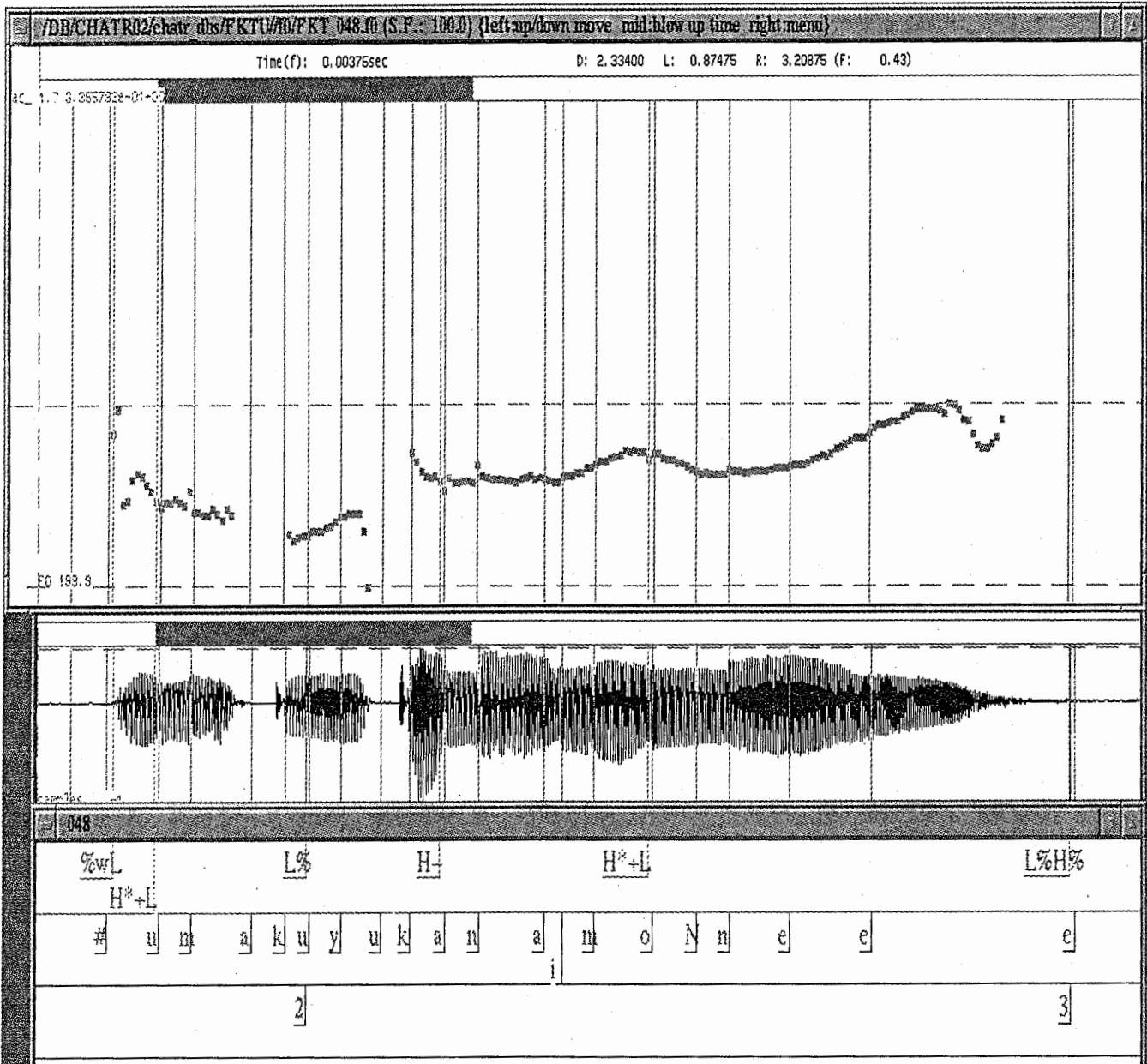
“もうすぐ三十に手が届くというのに” (関西)



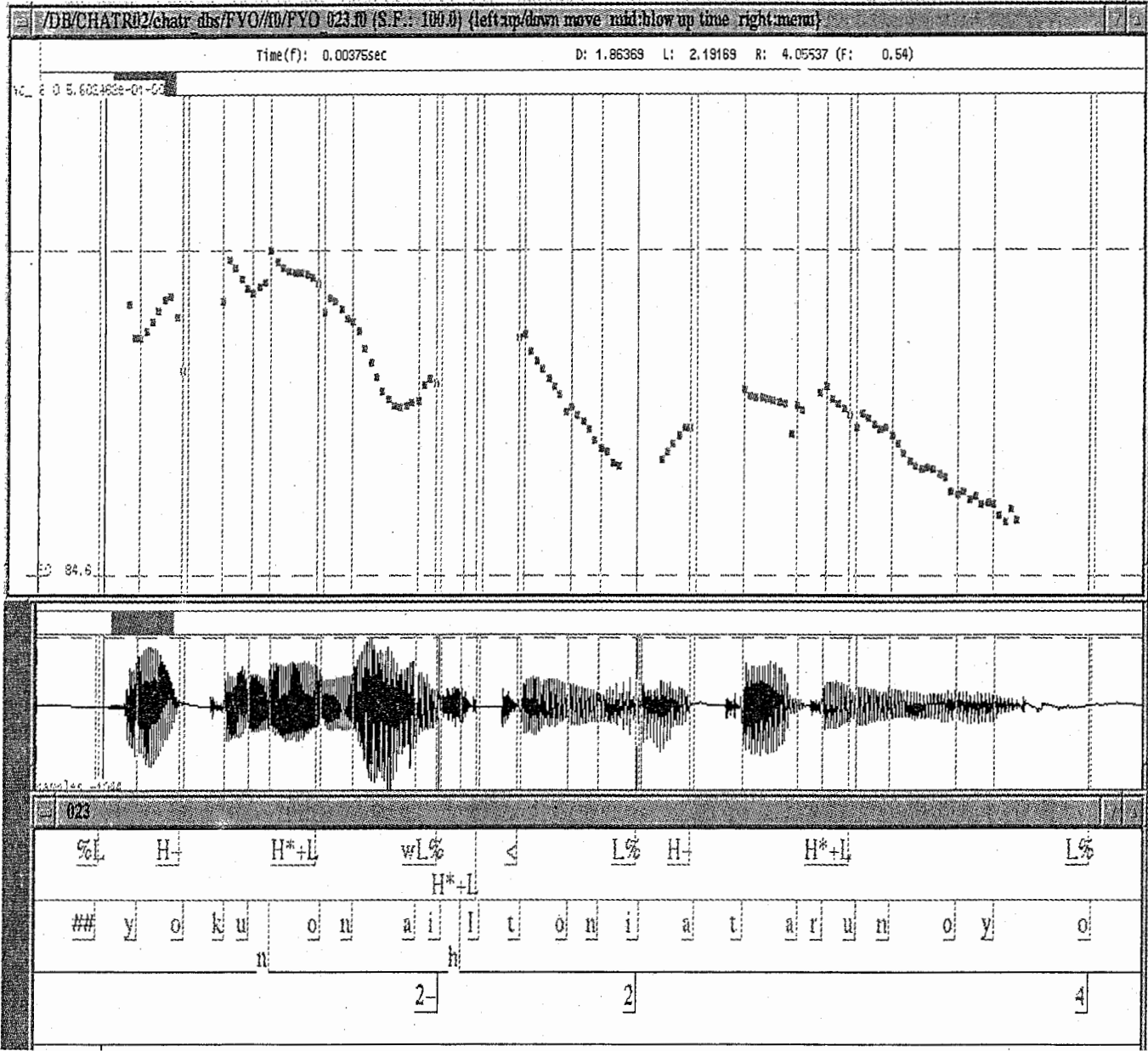
“もうすぐ三十に手が届くというのに” (共通語)



“うまくいかないもんねえ” (関西)



“うまくいかないもんねえ” (共通語)

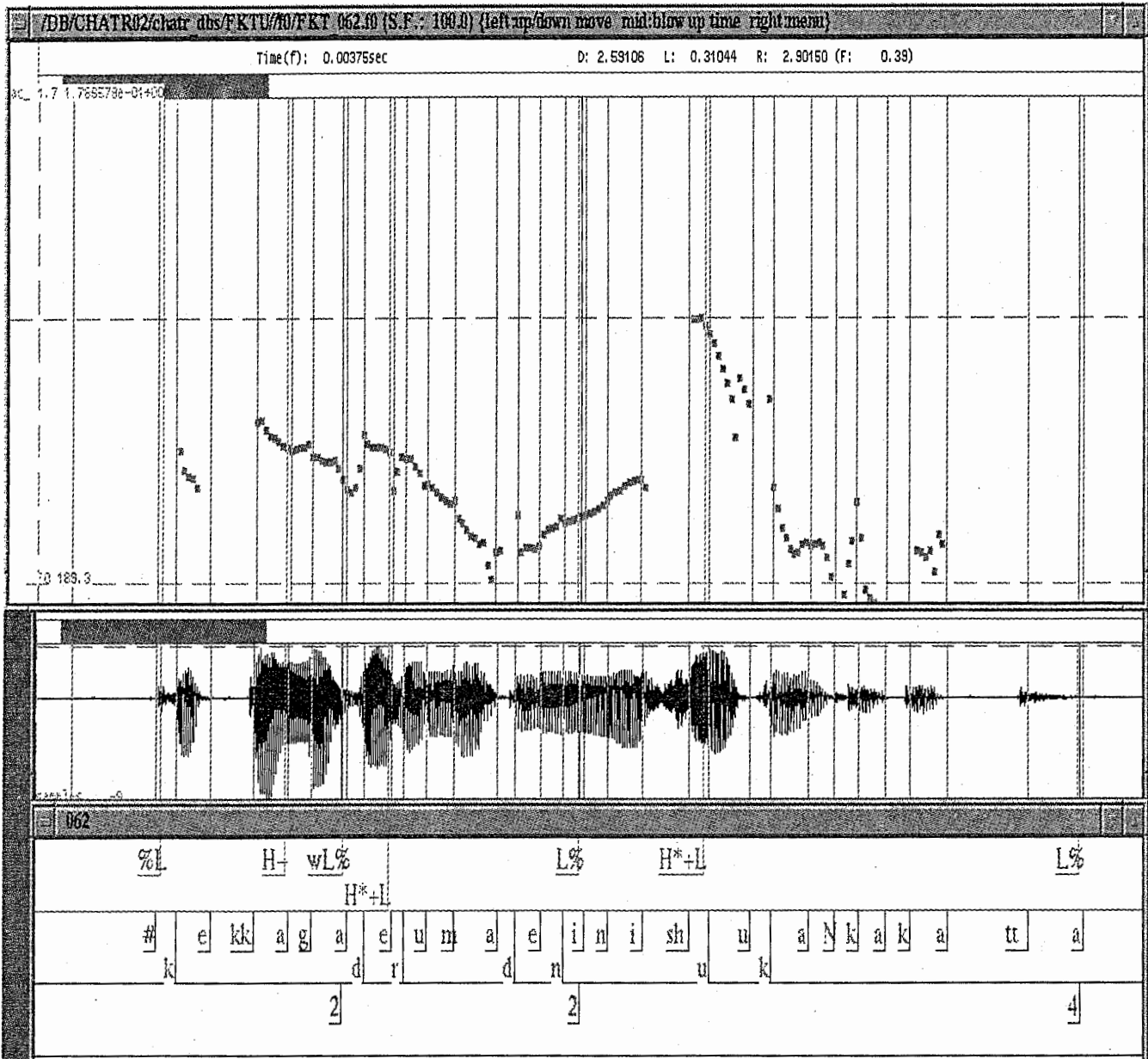


“欲のない人に当たるのよ” (関西)

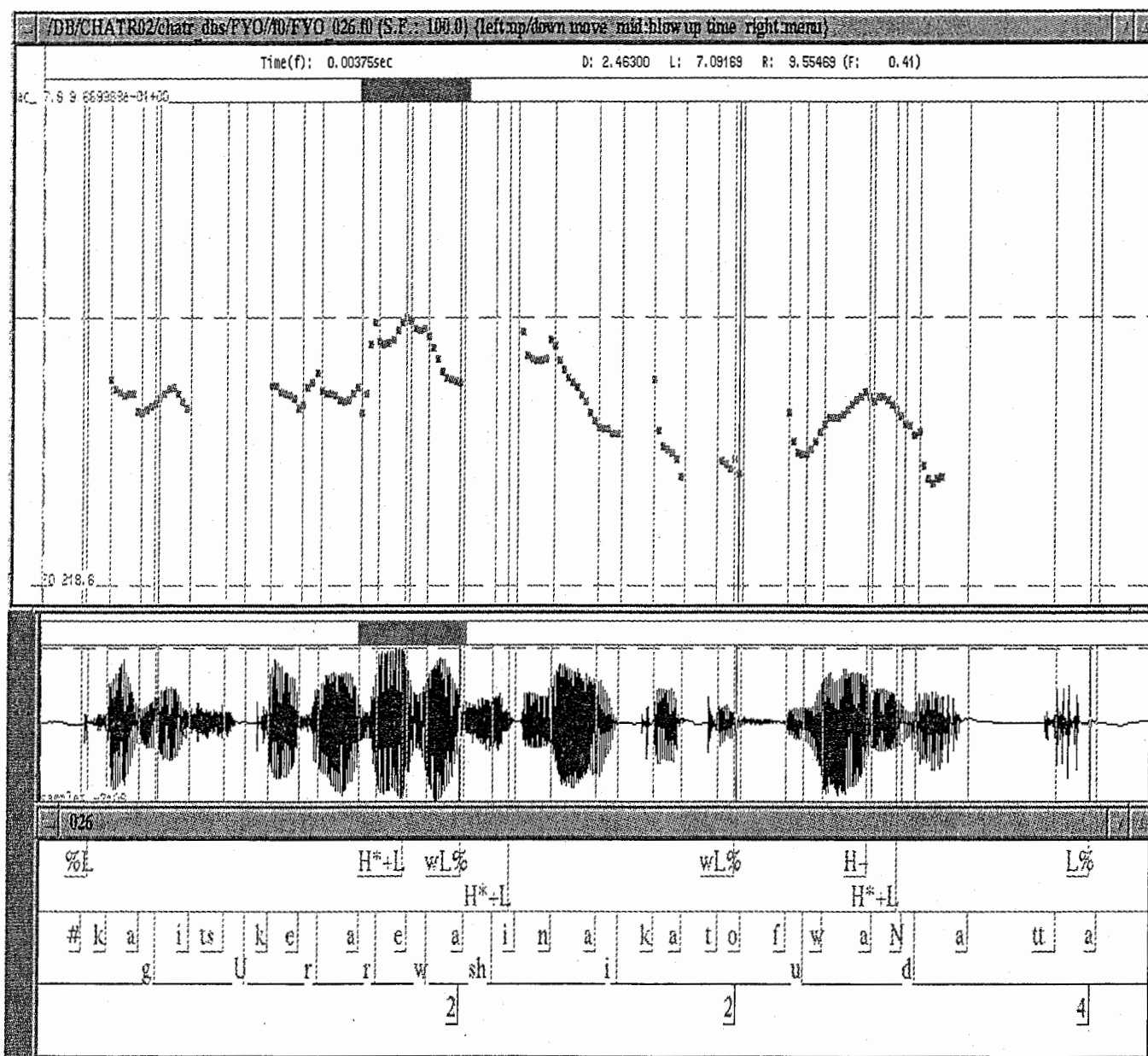




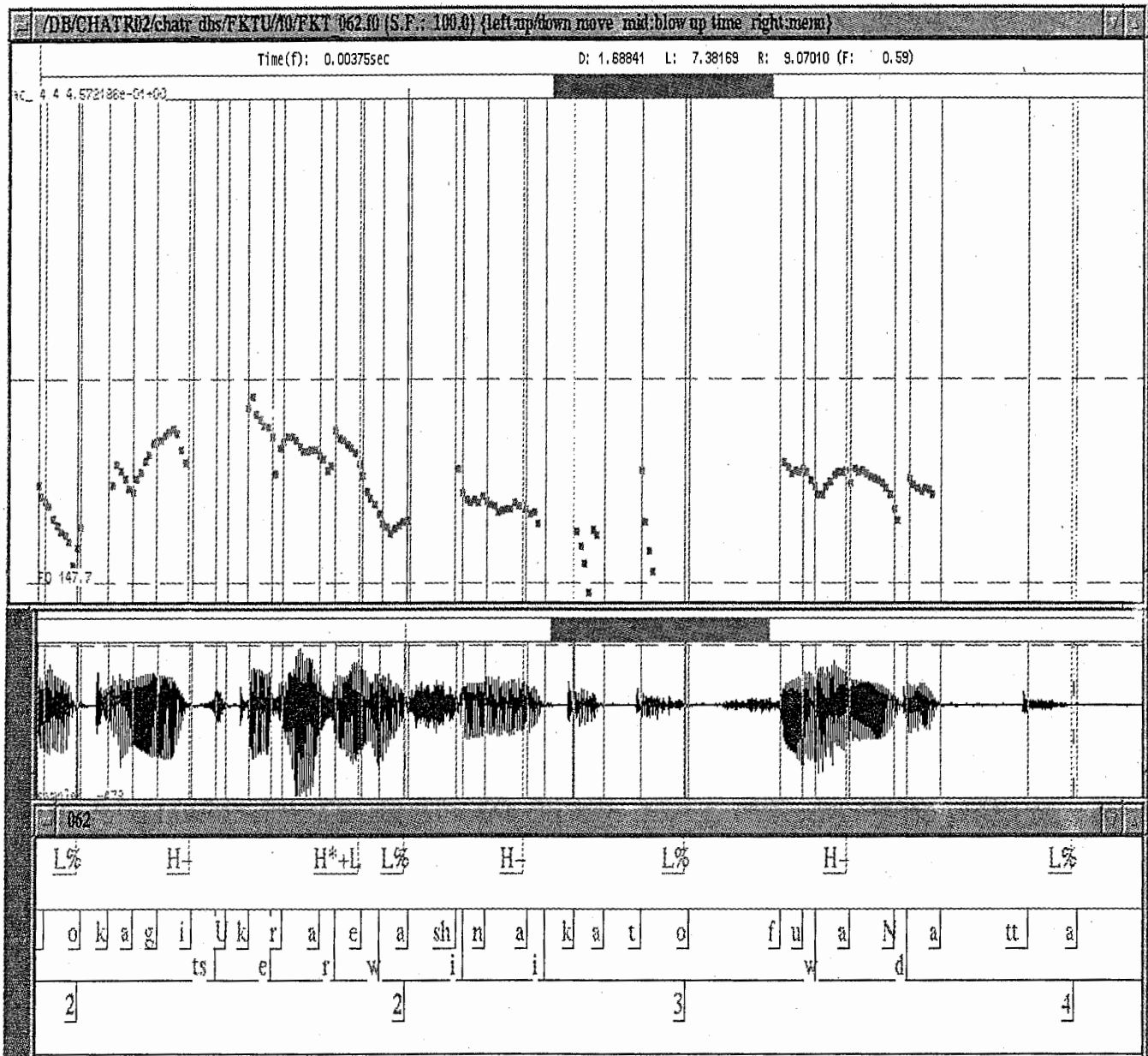




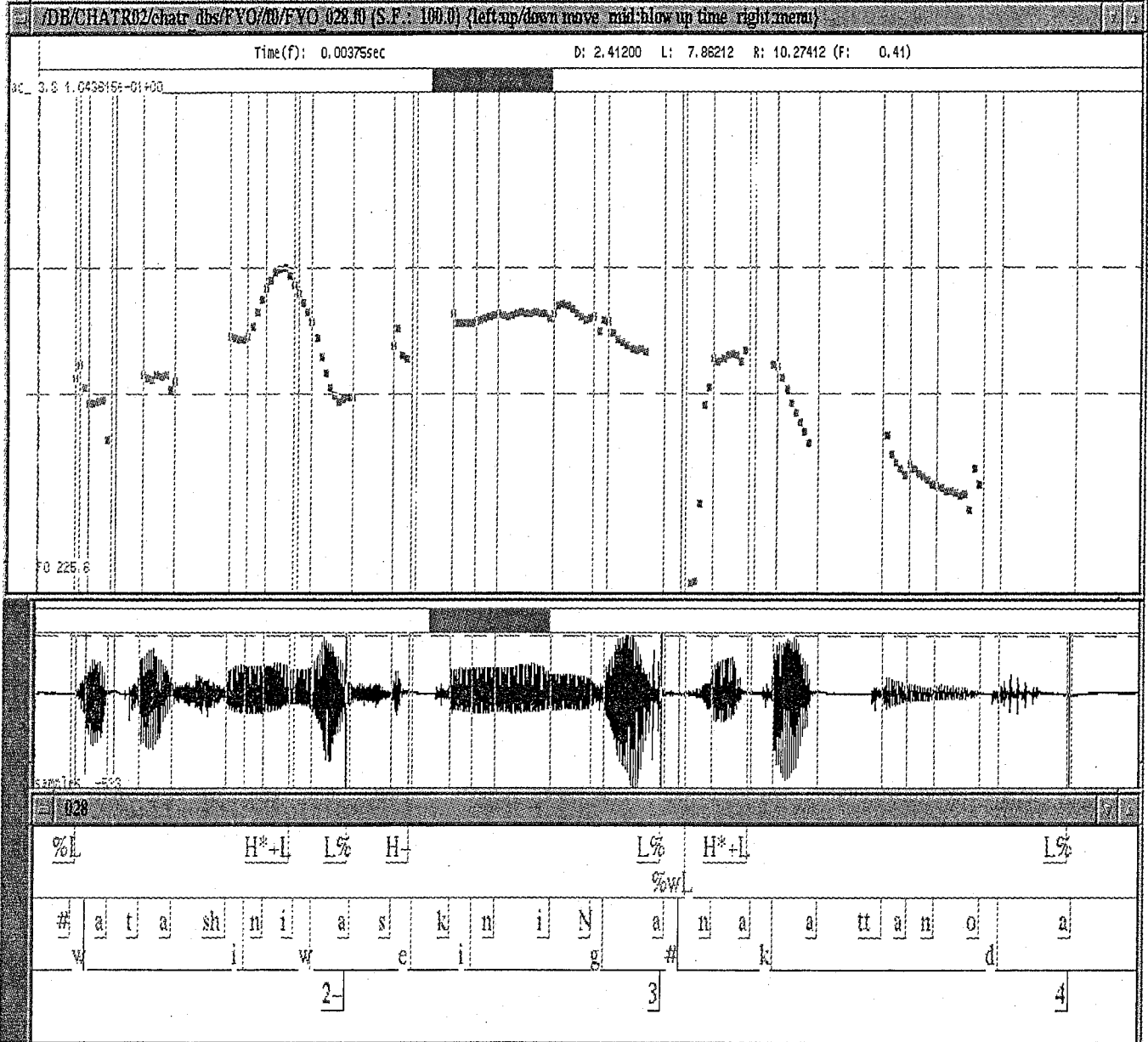
“結果が出るまでに二週間かった” (共通語)



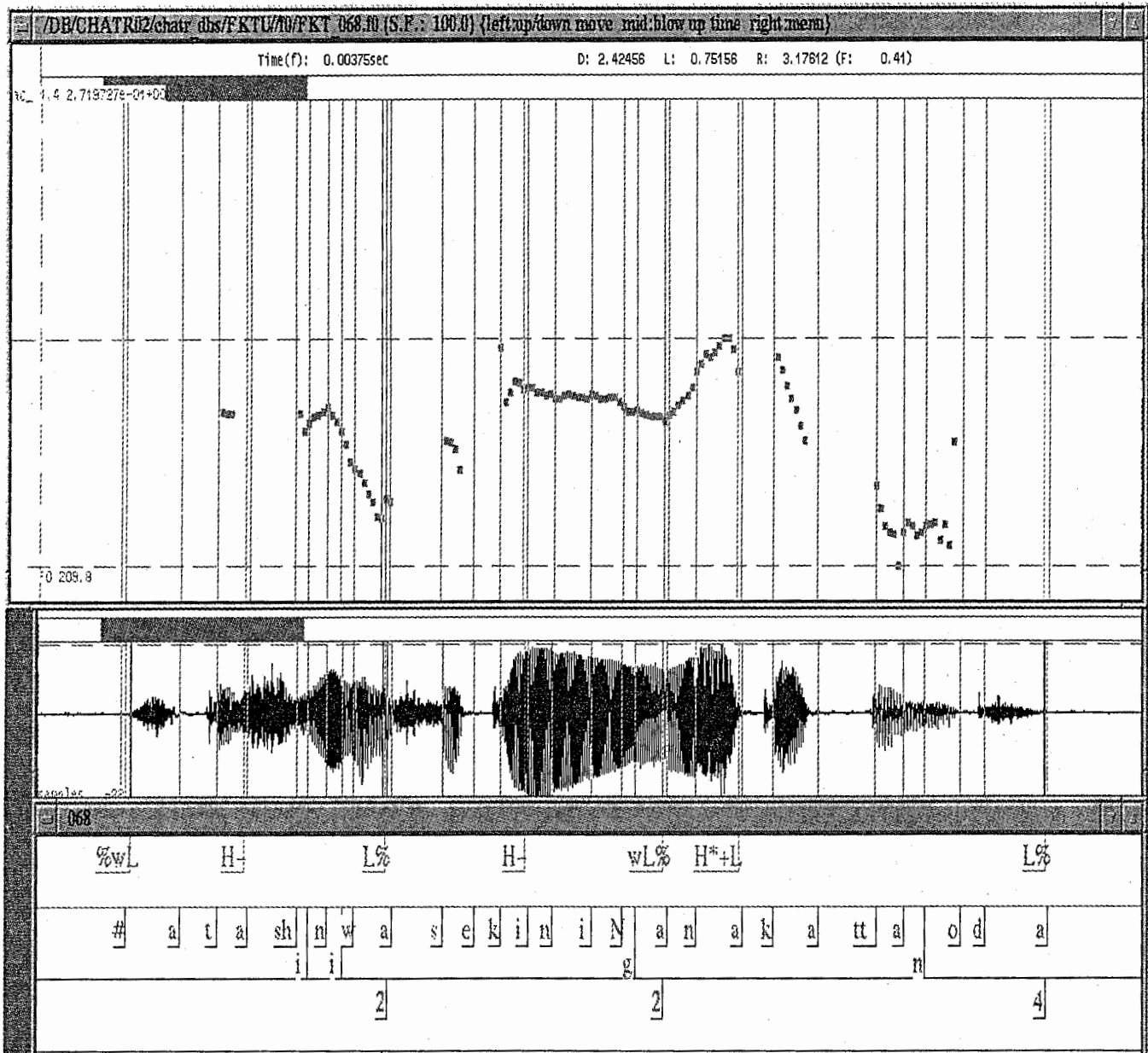
“かぎつけられはしないかと不安だった” (関西)



“かぎつけられはしないかと不安だった” (共通語)

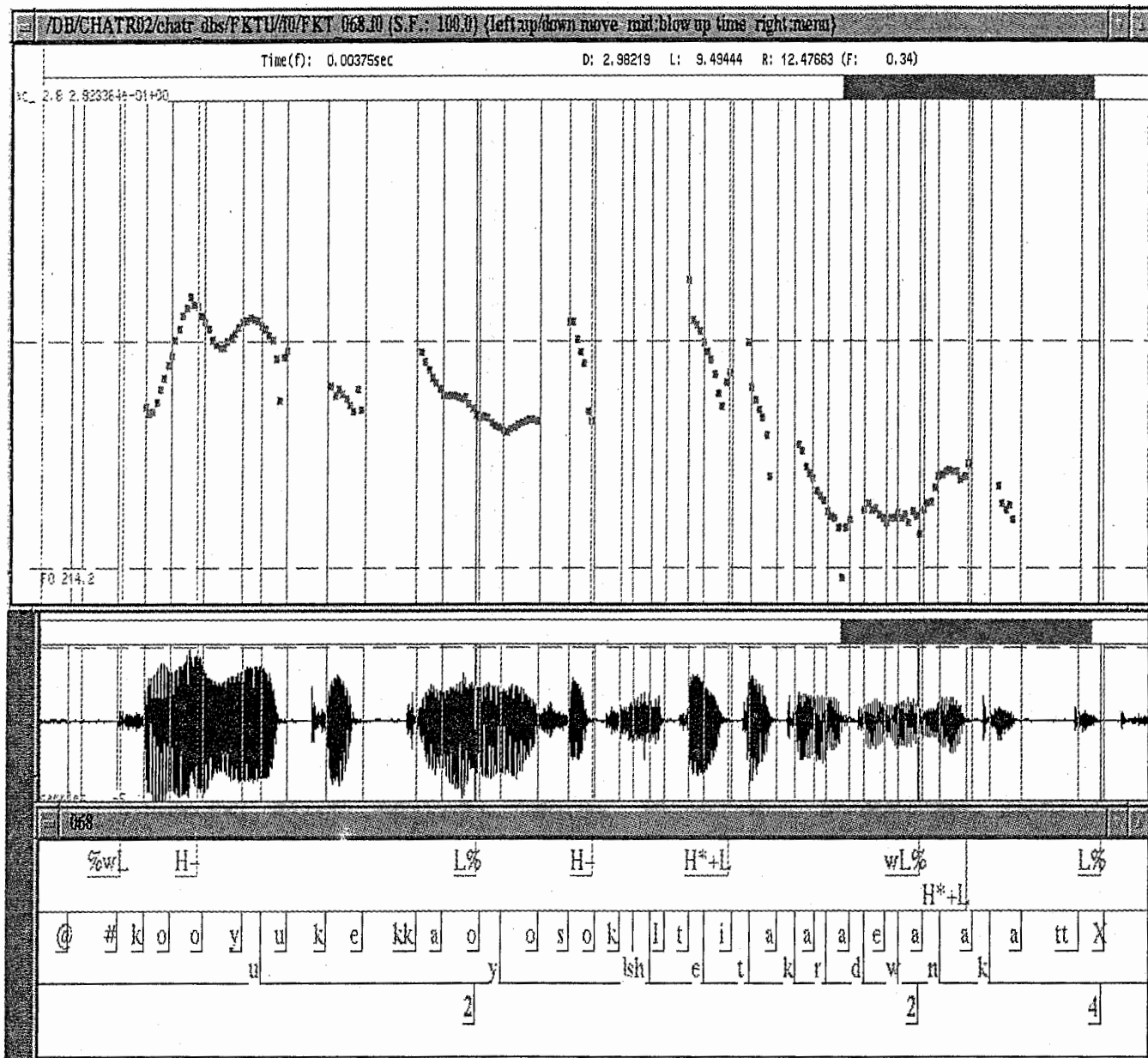


“私には責任がなかったのだ” (関西)



“私には責任がなかったのだ” (共通語)

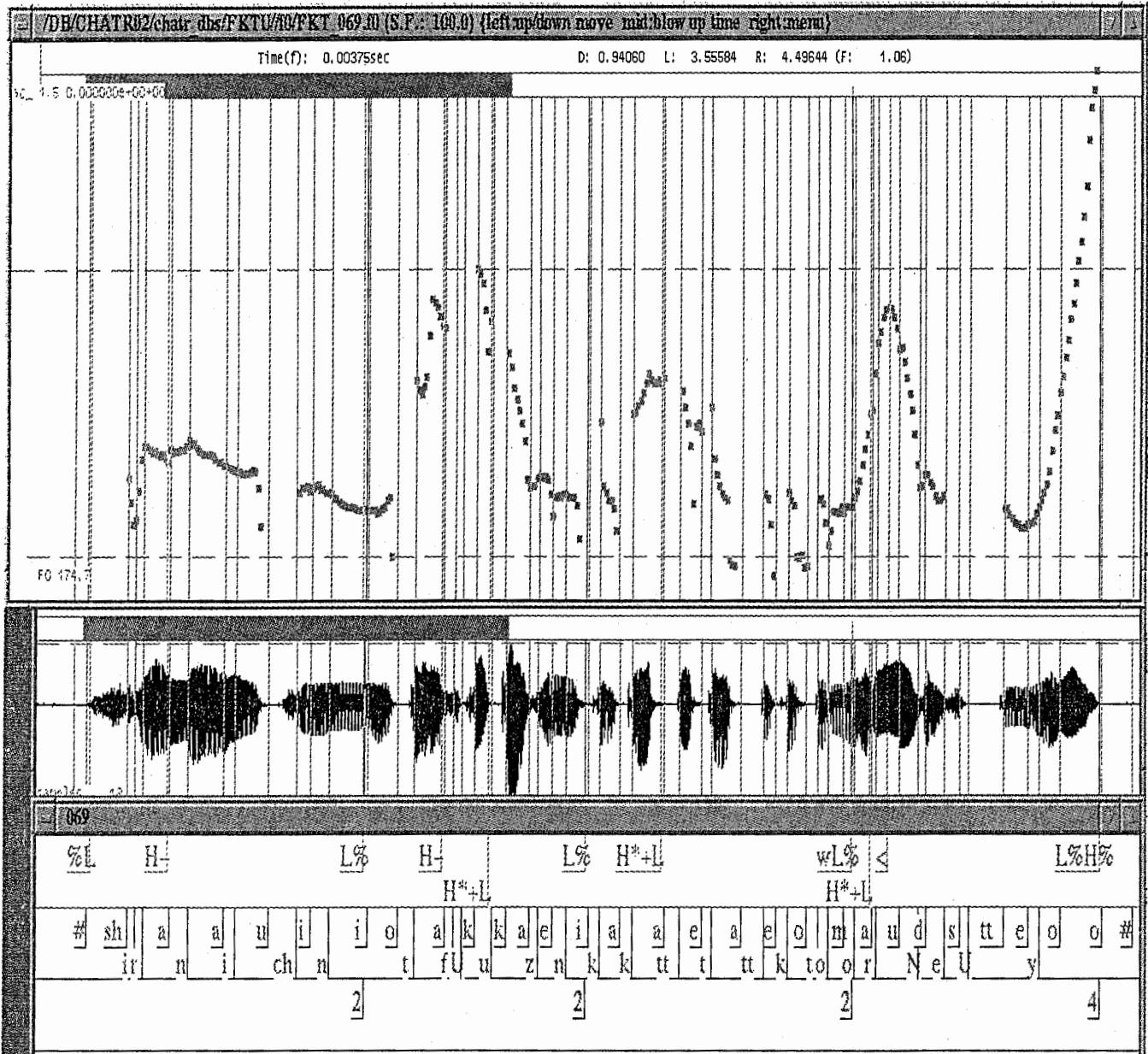




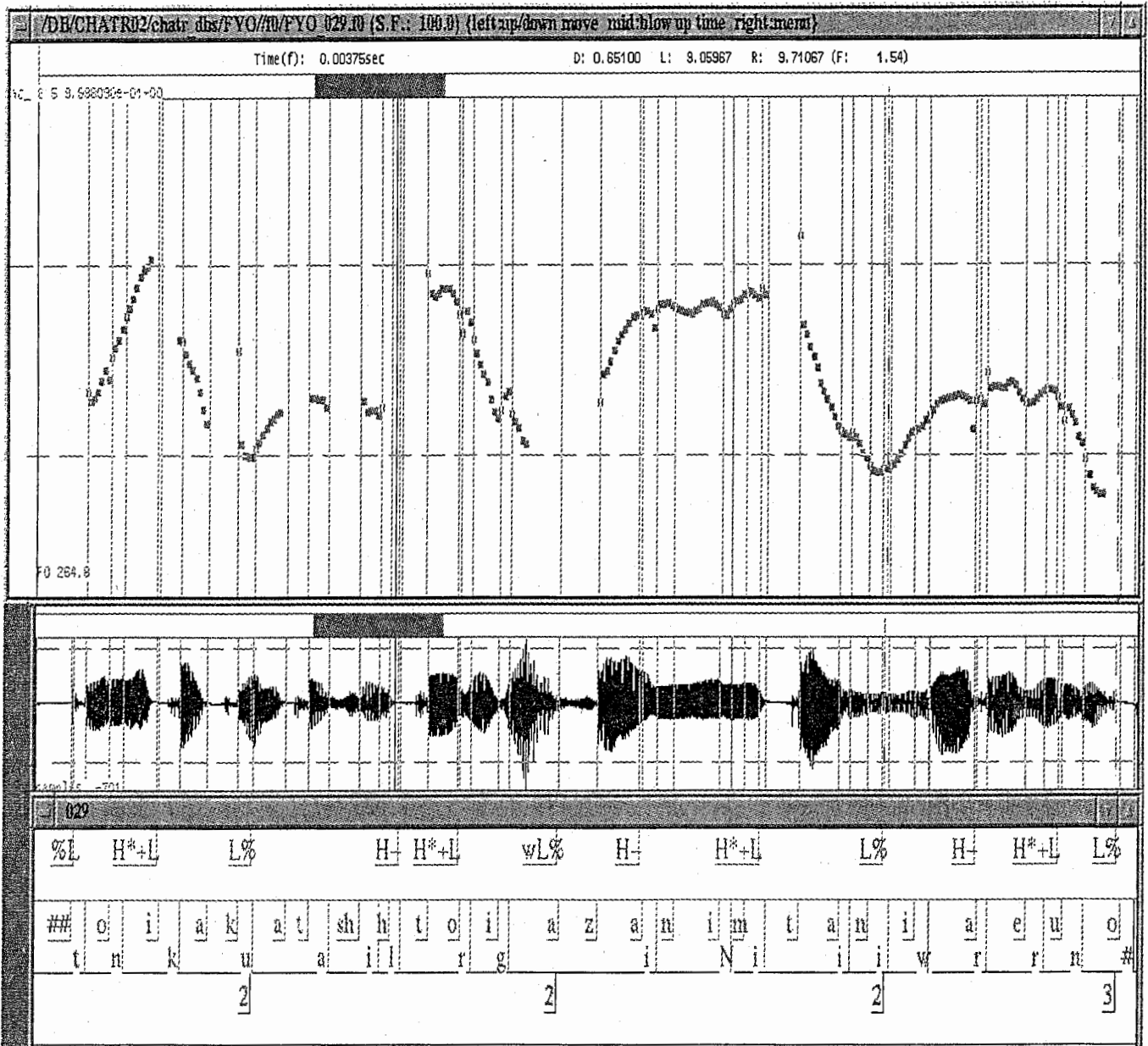
“こういう結果を予測していたからではなかったか” (共通語)



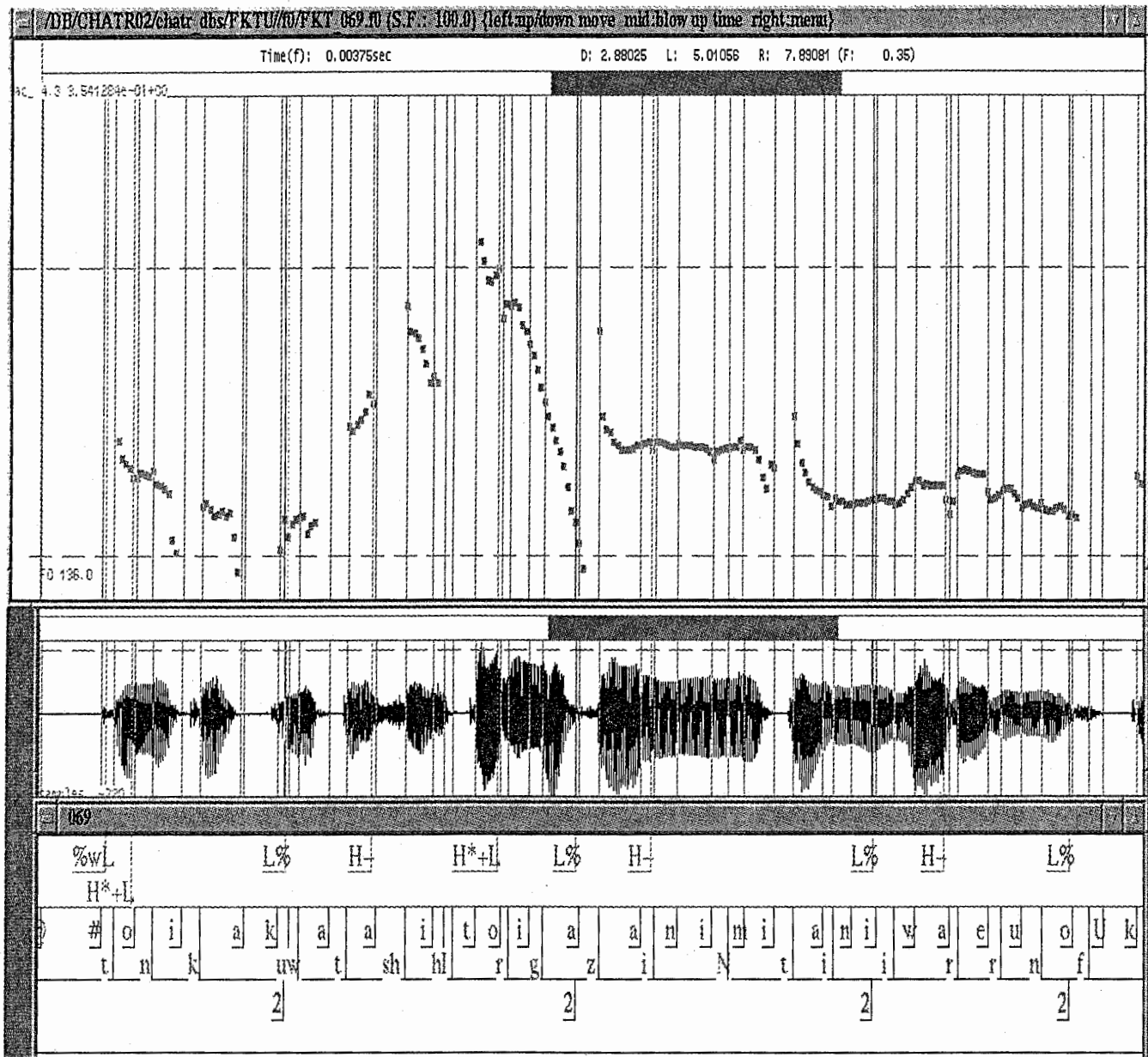




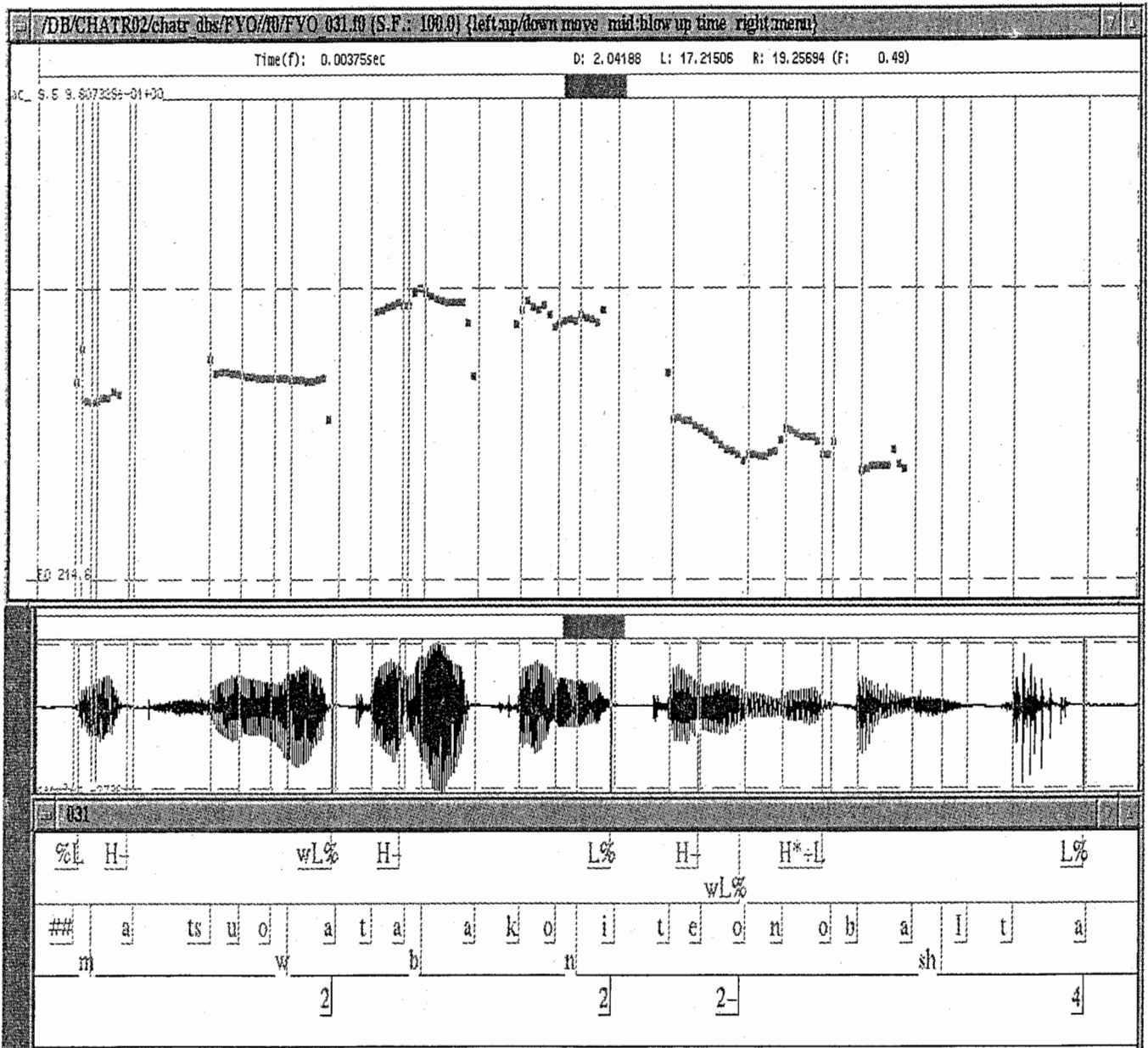
“知らないうちにお多福風邪にかかっていたってこともあるんですけどよ”  
 (共通語)



“とにかく、あたし一人が罪人みたいに言われるの” (関西)

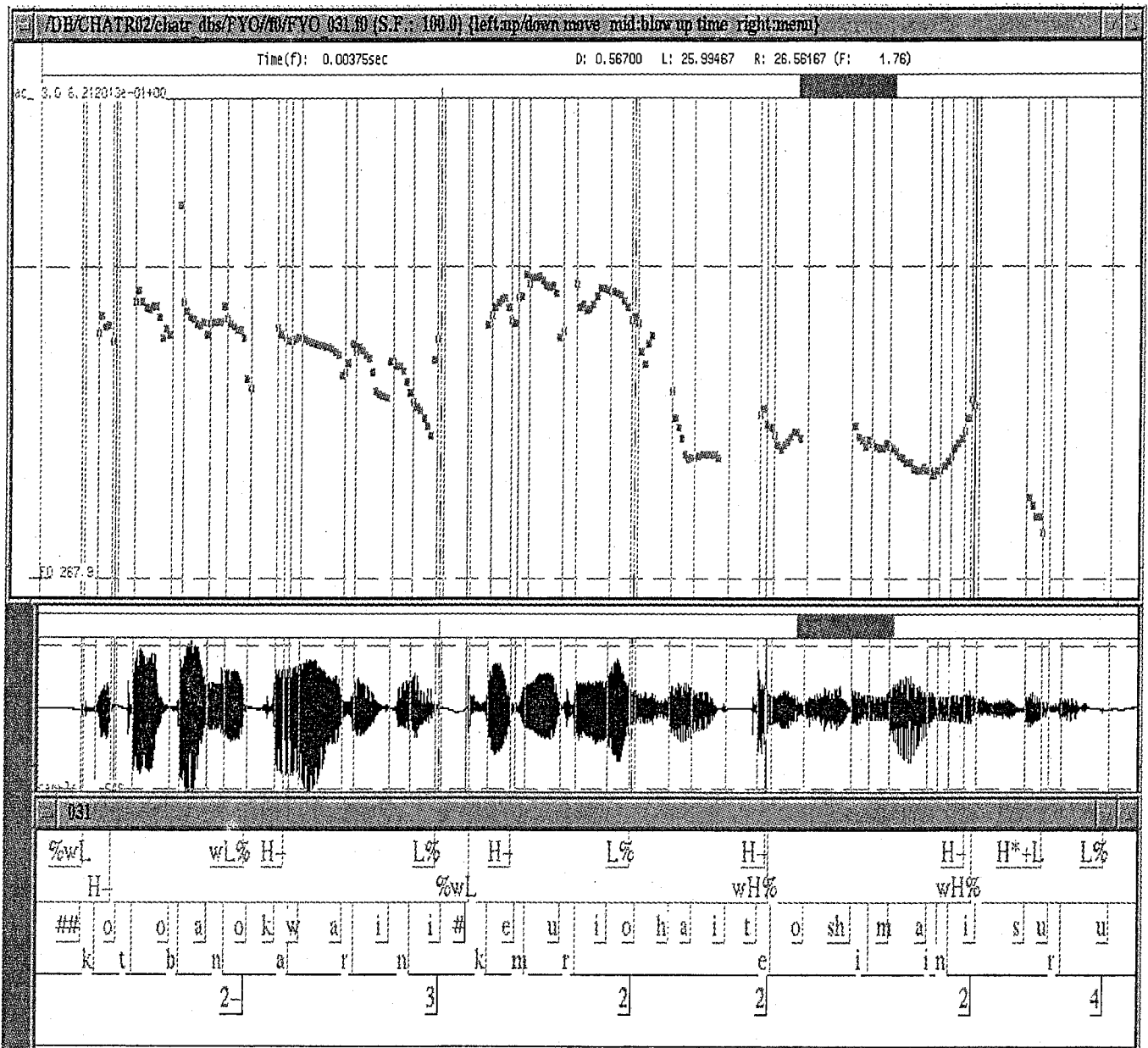


“とにかく、あたし一人が罪人みたいに言われるの” (共通語)



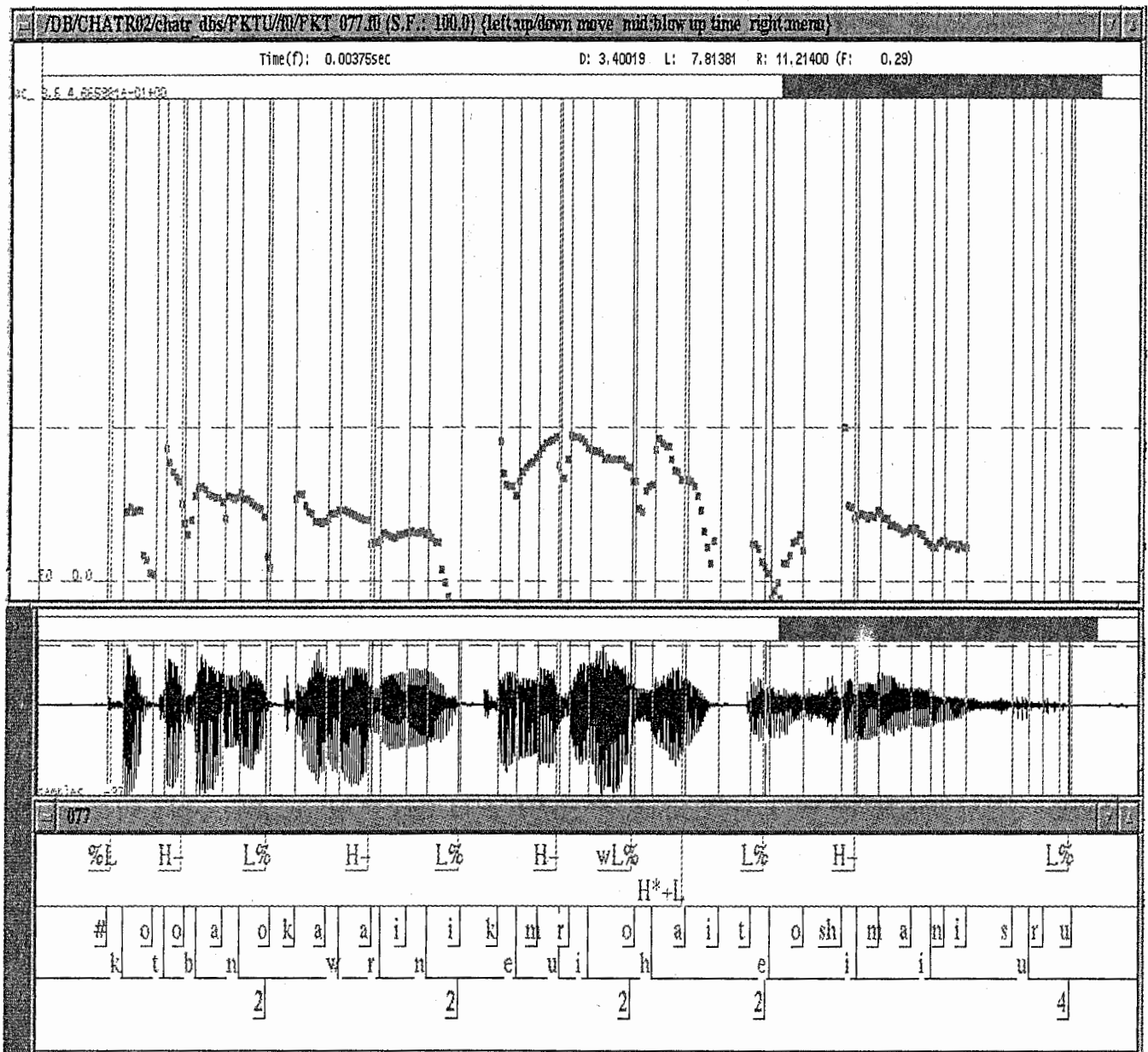
“松夫は煙草に手を伸ばした” (関西)





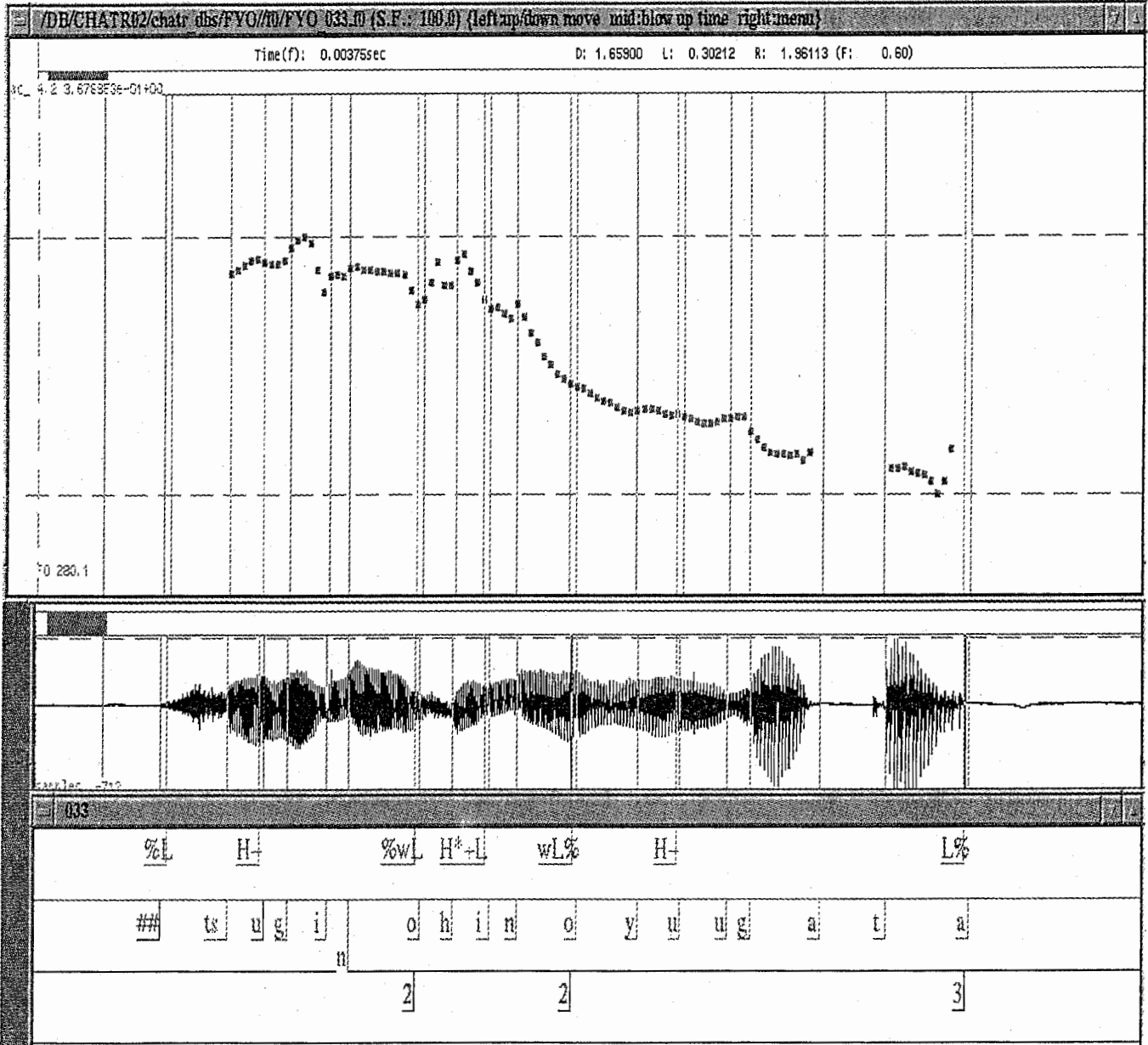
“ことばの代わりに煙を吐いておしまいにする” (関西)



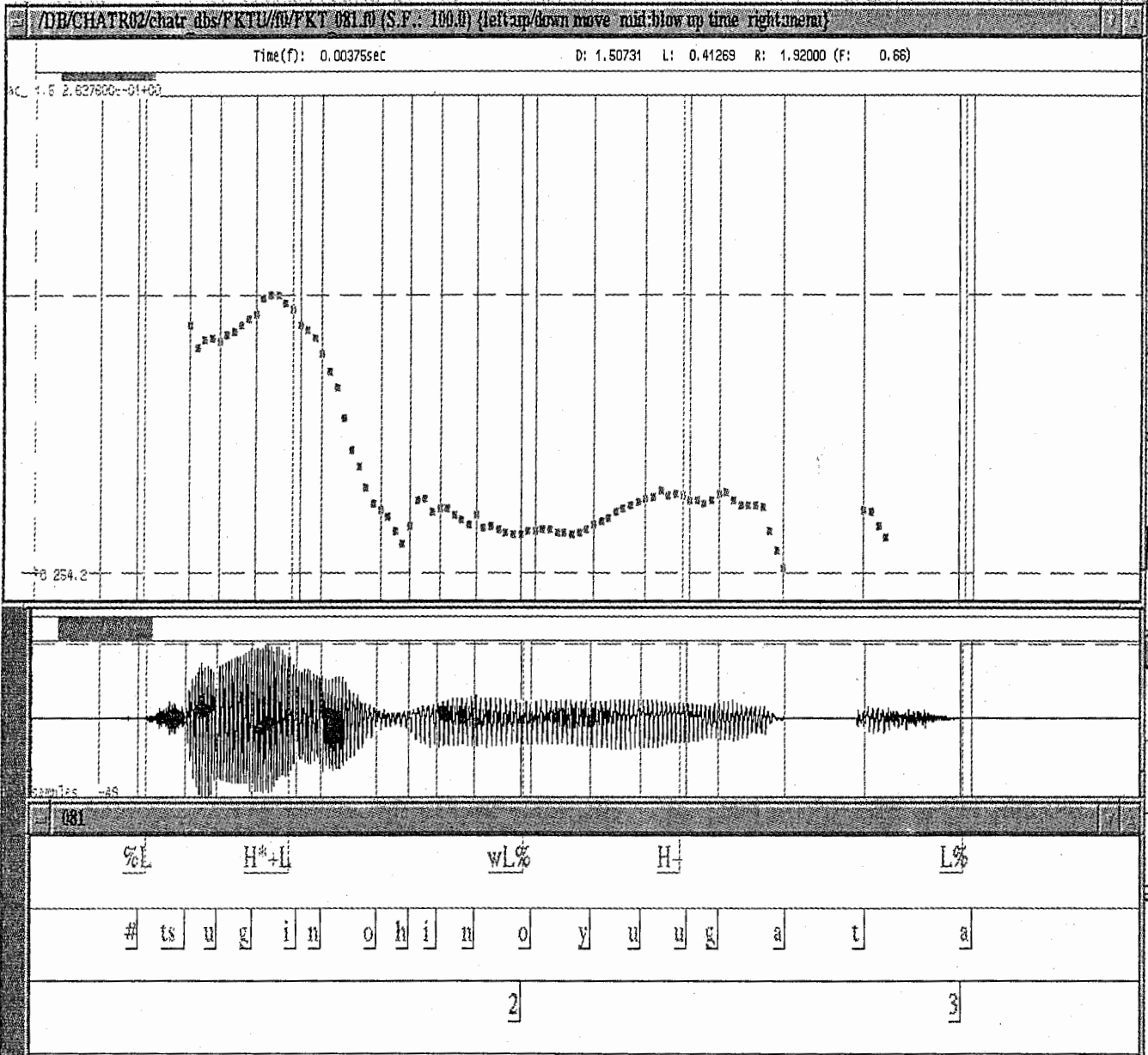


“ことばの代わりに煙を吐いておしまいにする” (共通語)

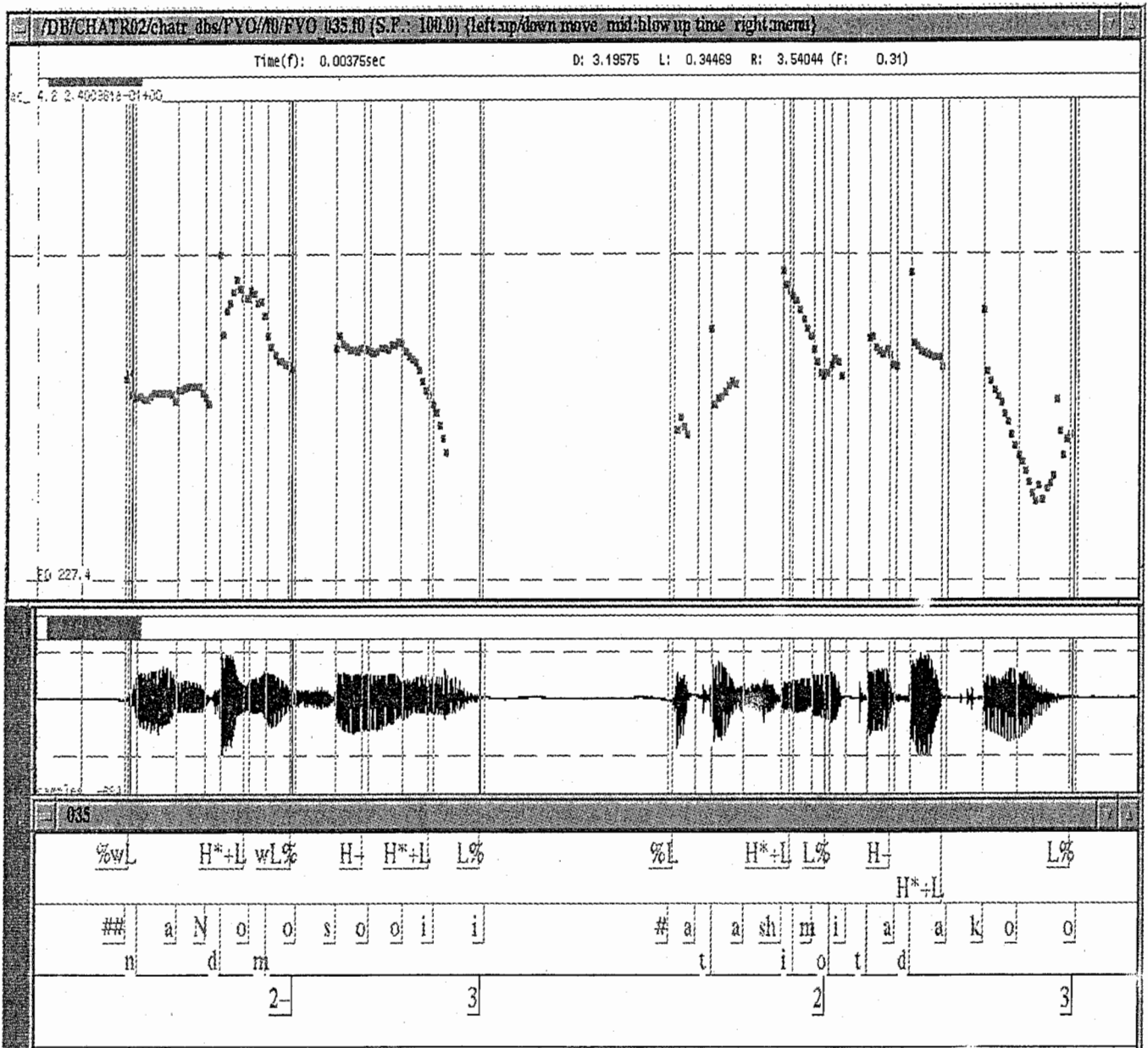




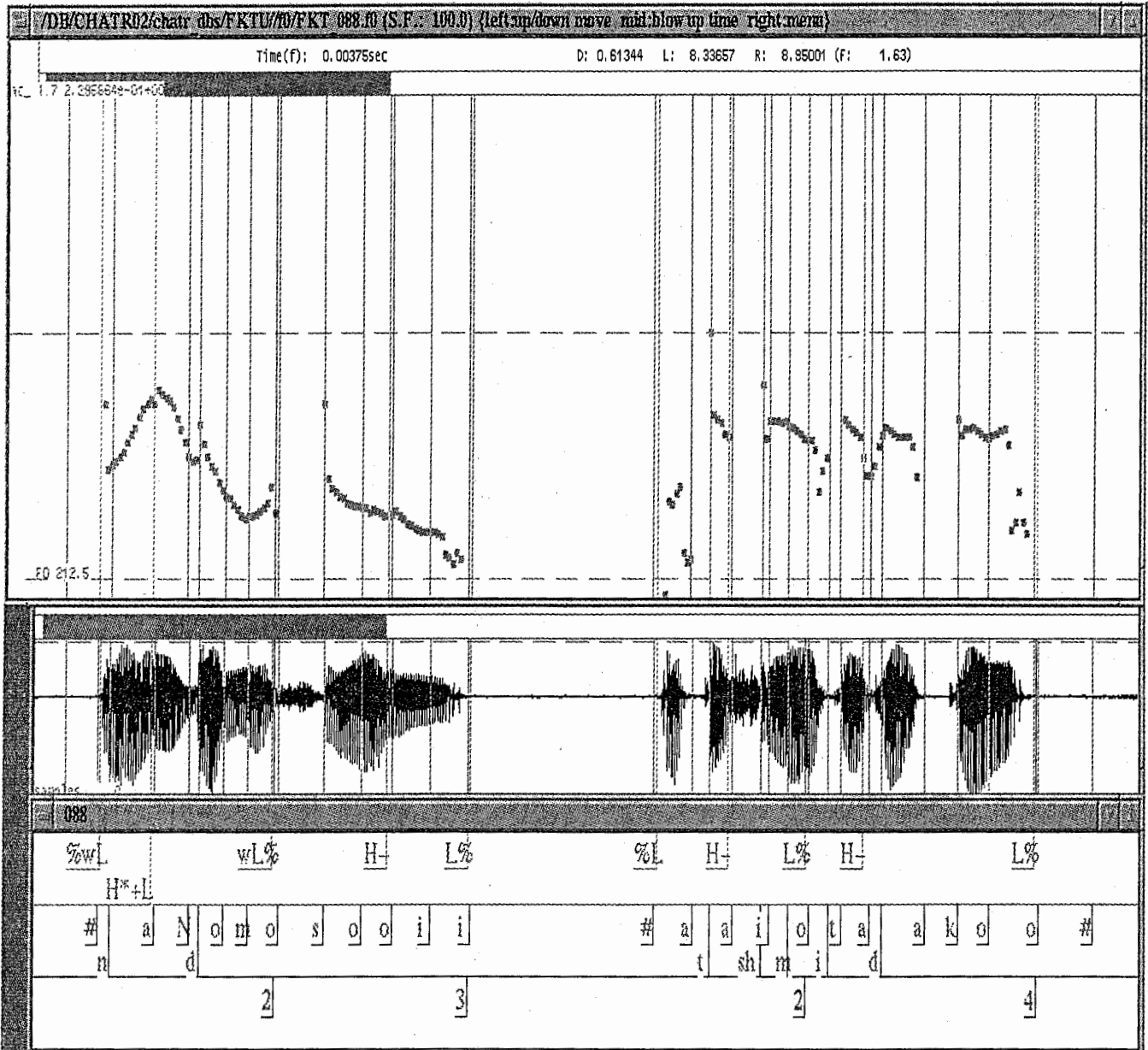
“次の日の夕方” (関西)



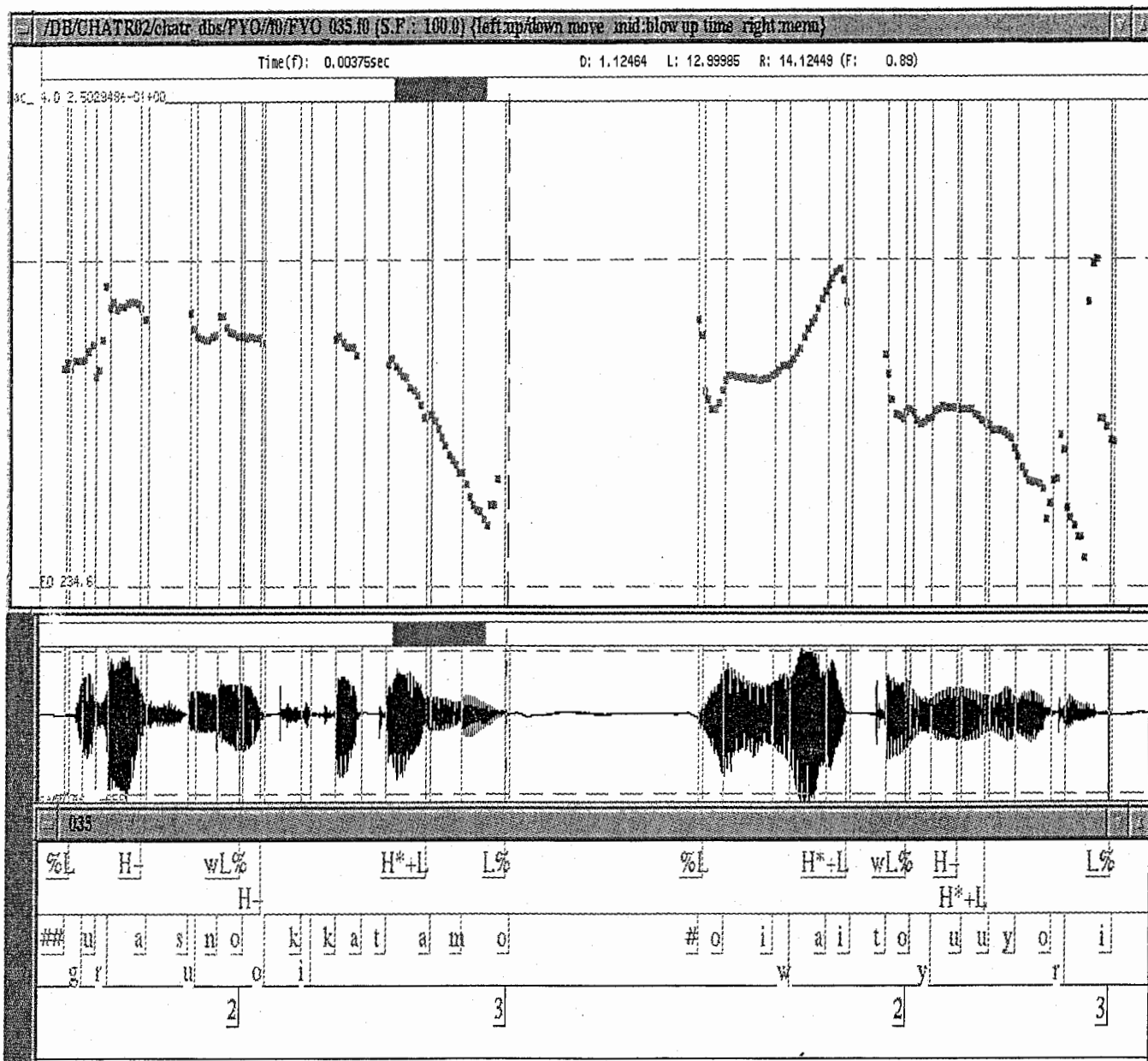
“次の日の夕方” (共通語)



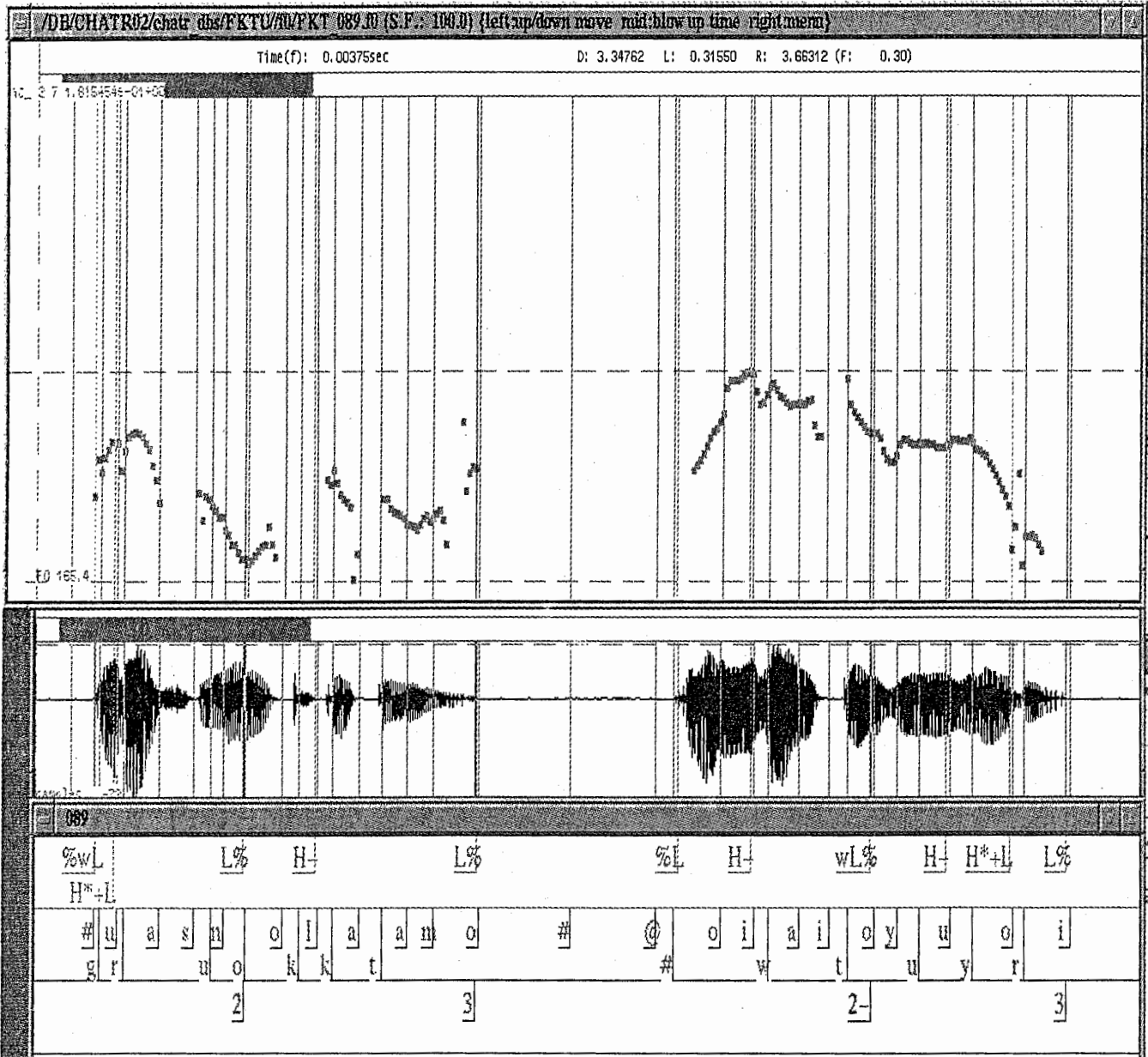
“何度もそう言い、あたしも頂こう” (関西)



“何度もそう言い、あたしも頂こう” (共通語)

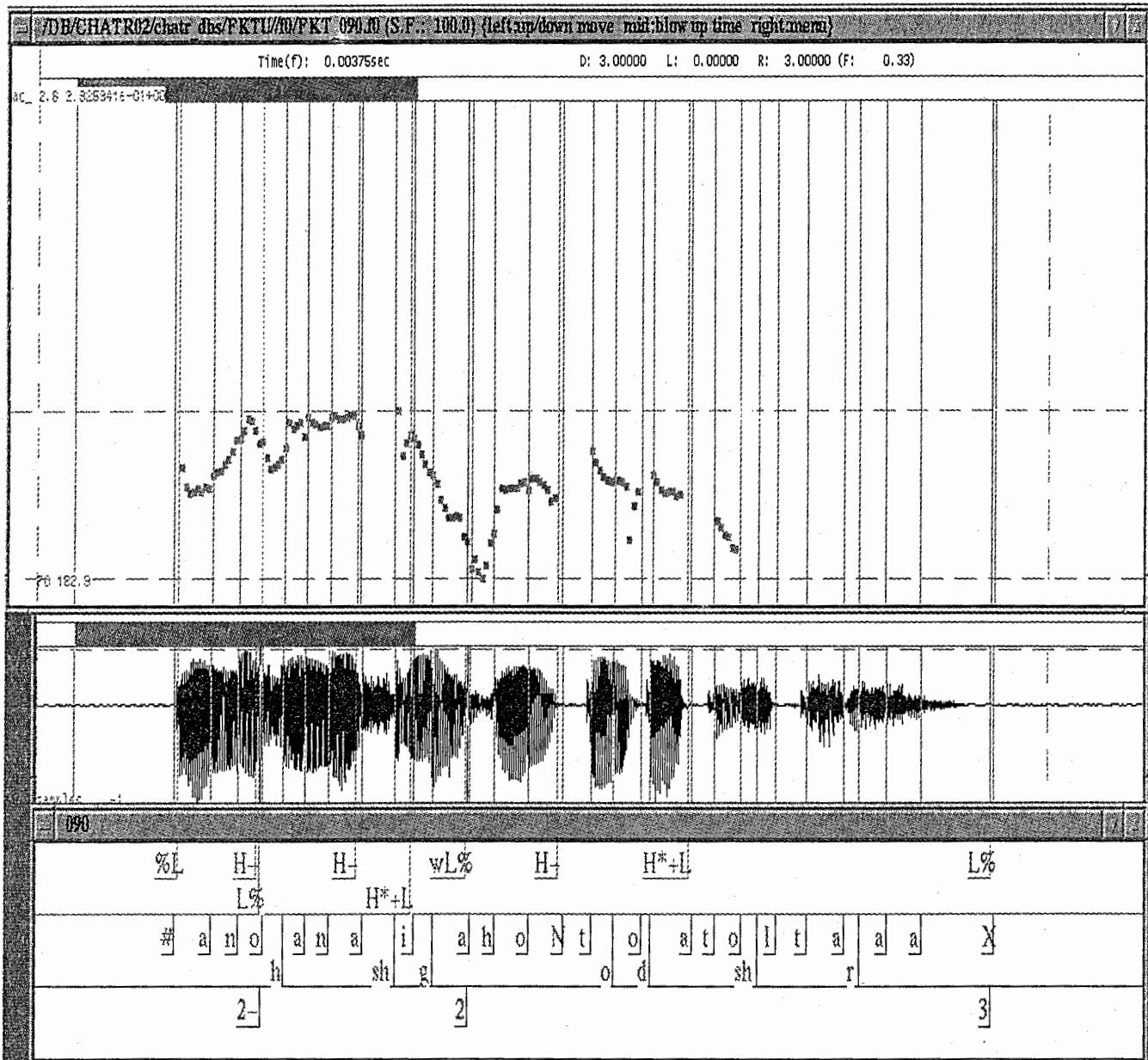


“グラスの置き方も、お祝いというより” (関西)



“グラスの置き方も、お祝いというより” (共通語)

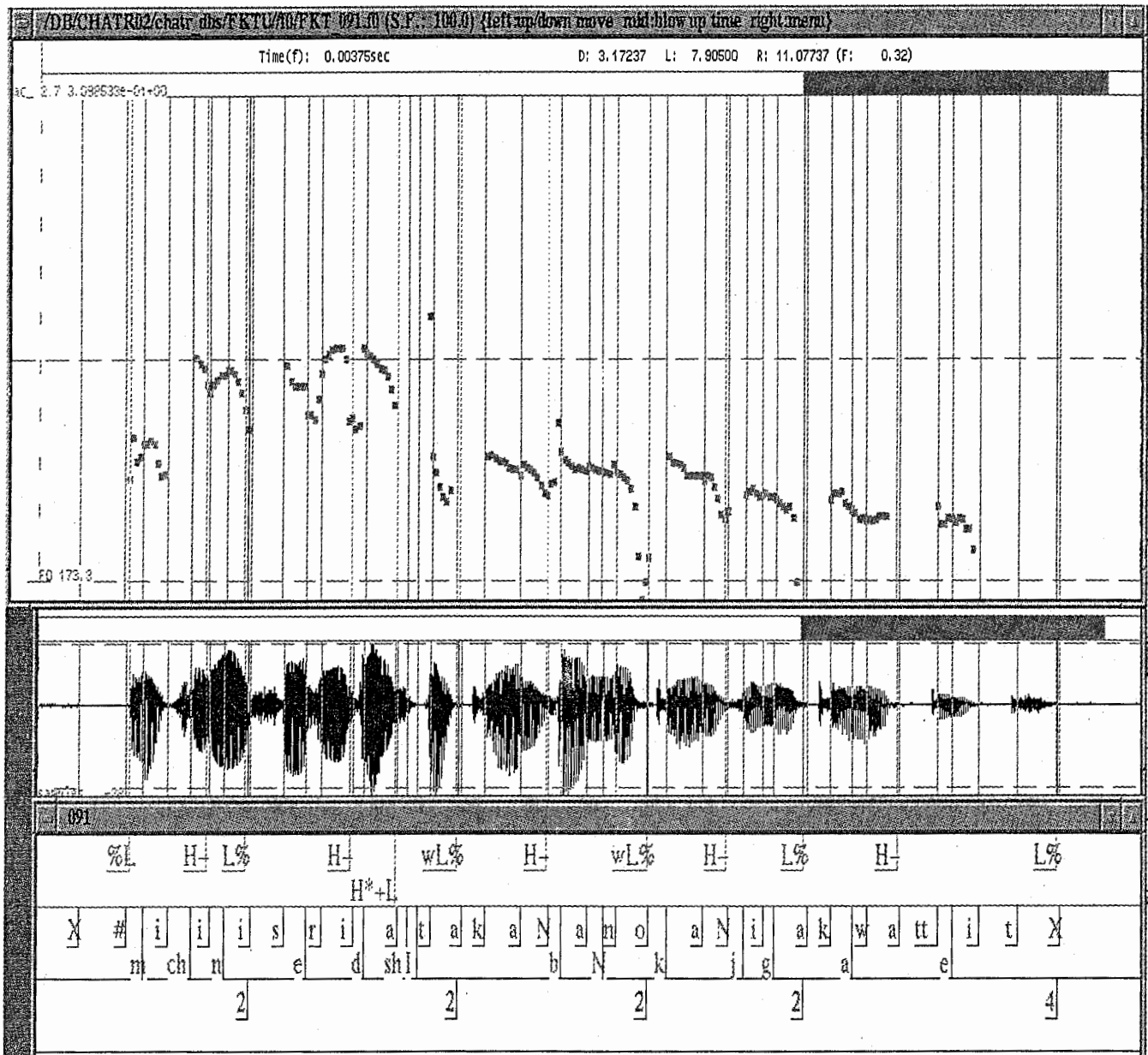




“あのはなしが本当だとしたら” (共通語)

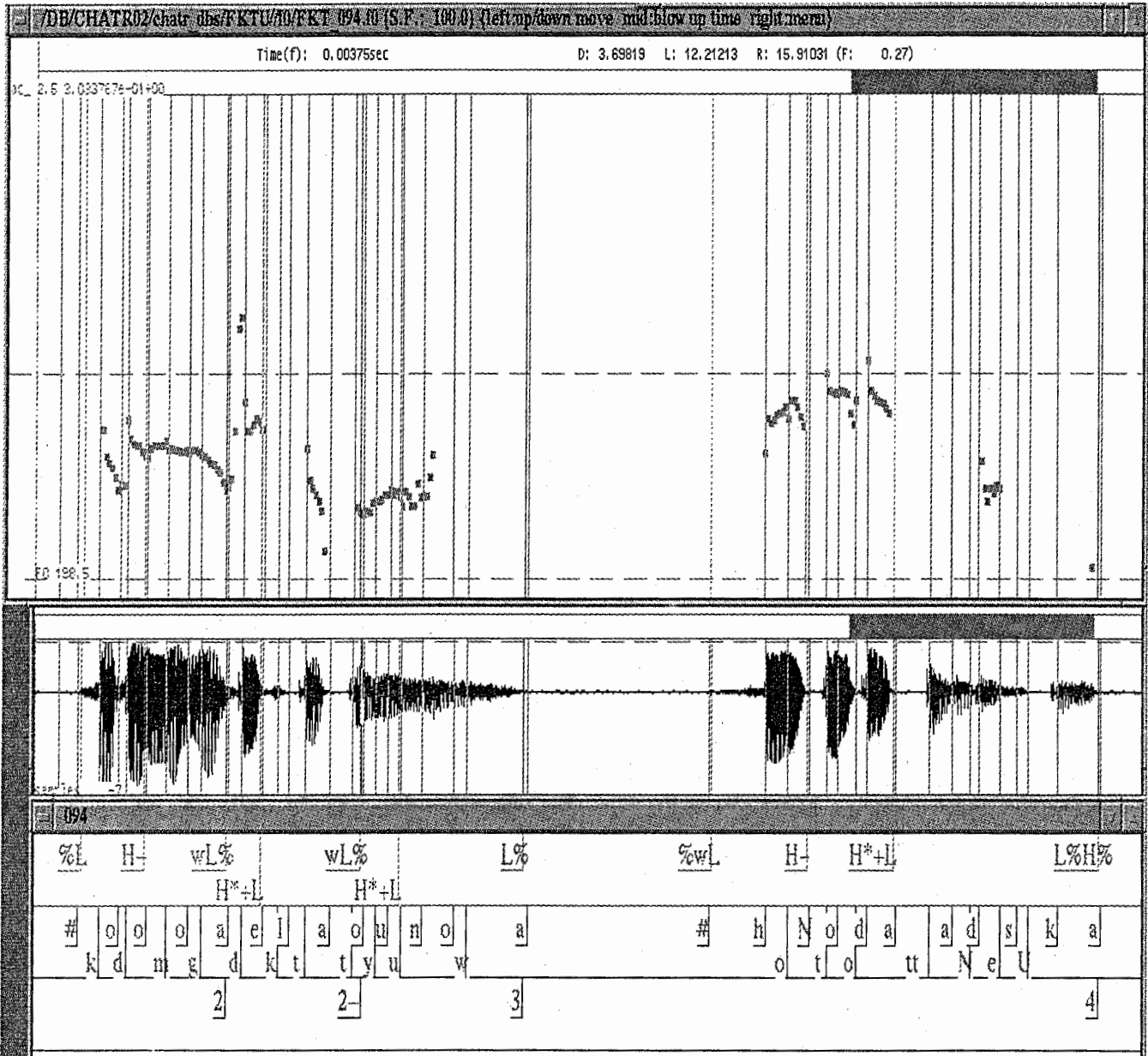




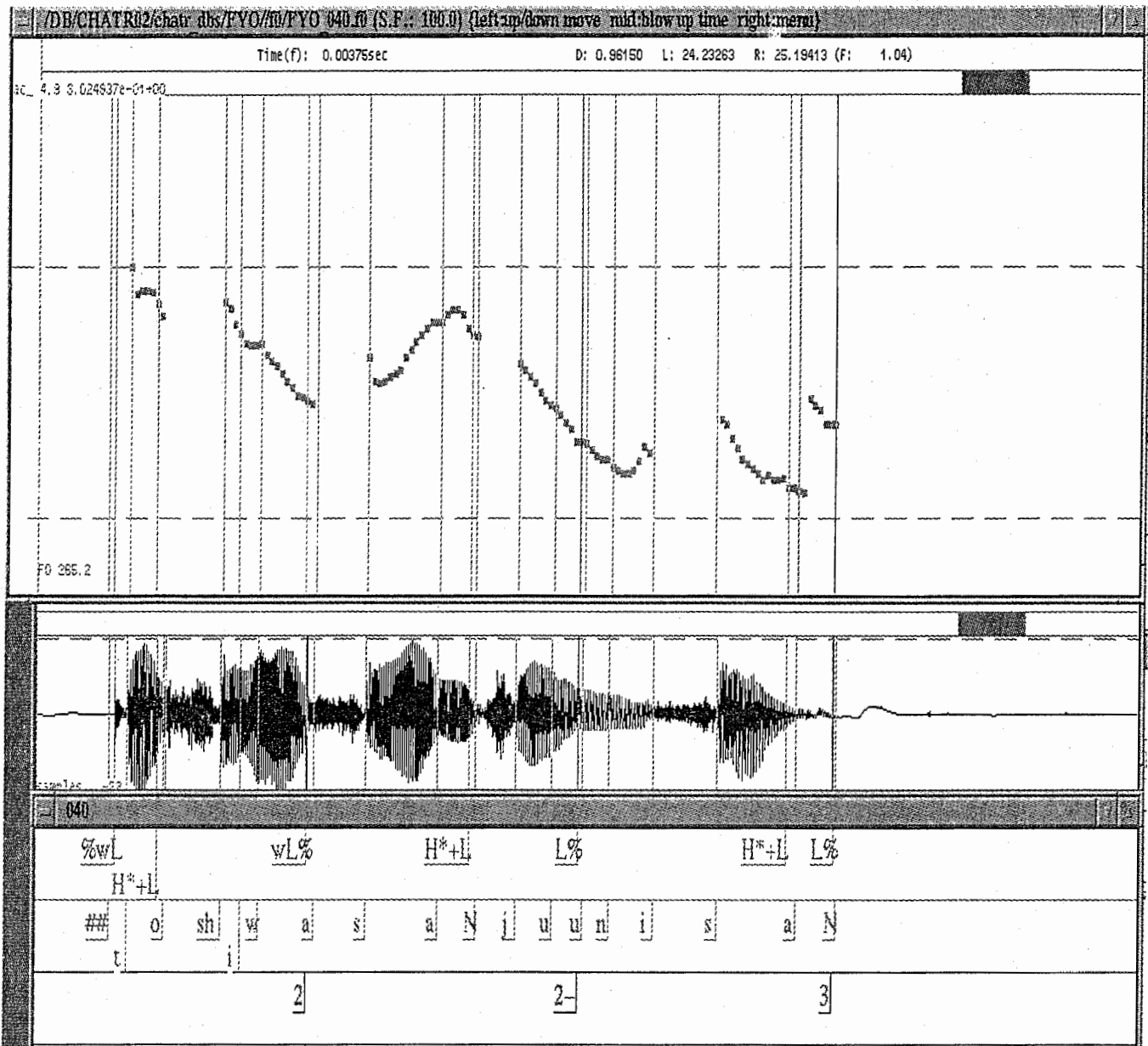


“道にせり出した看板の感じが変わっていた” (共通語)

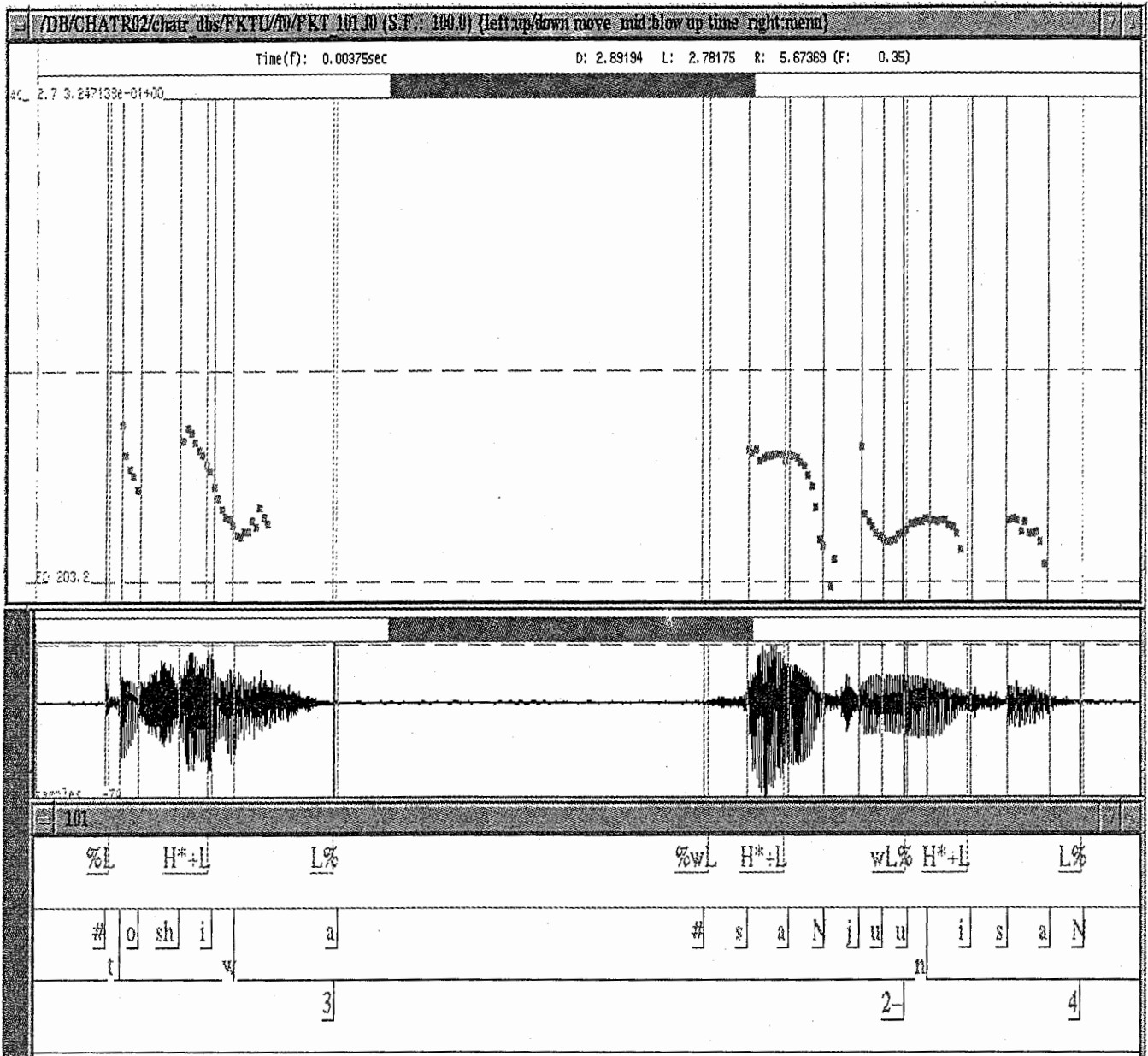




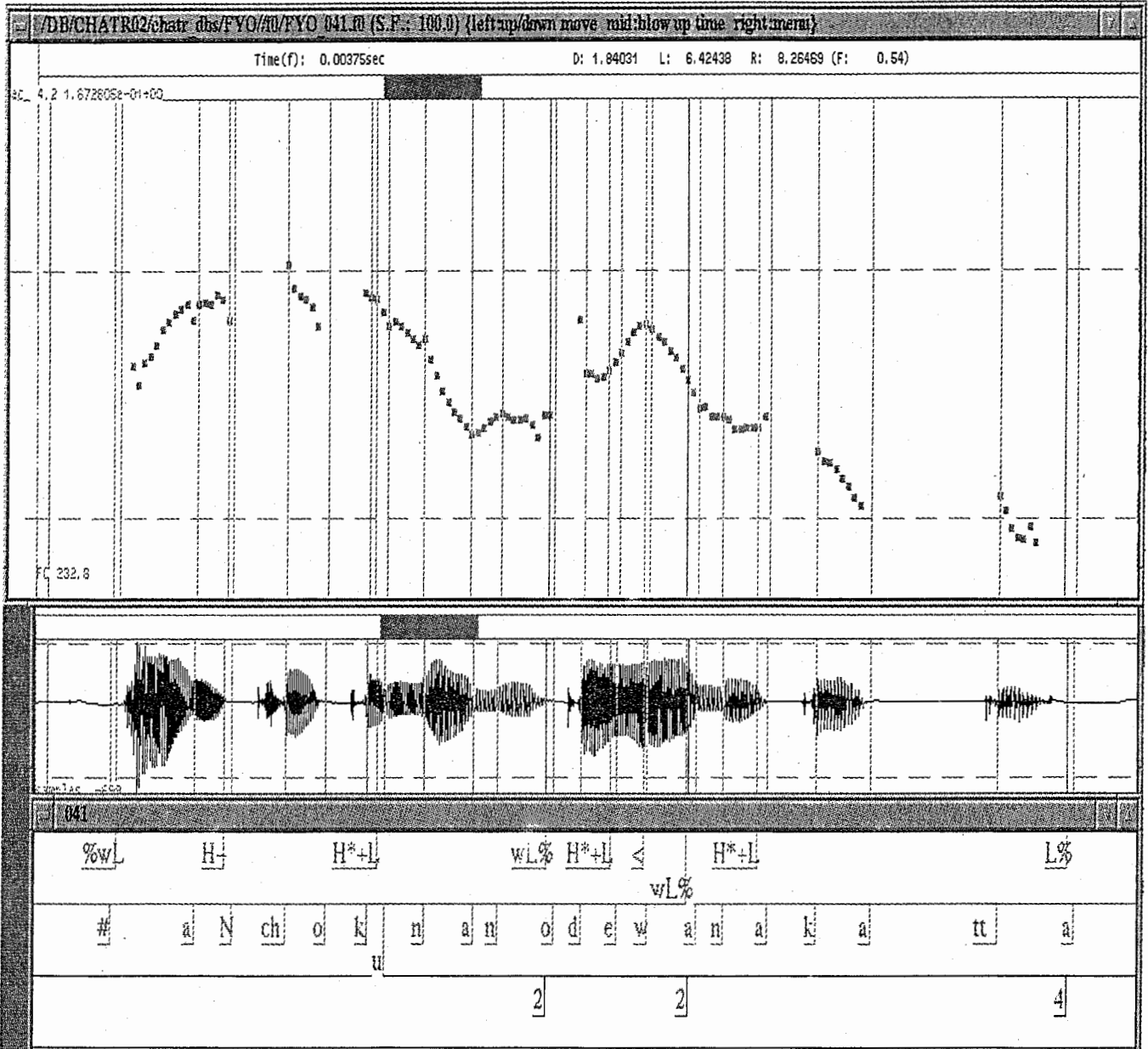
“子供が出来たというのは本当だったんですか” (共通語)



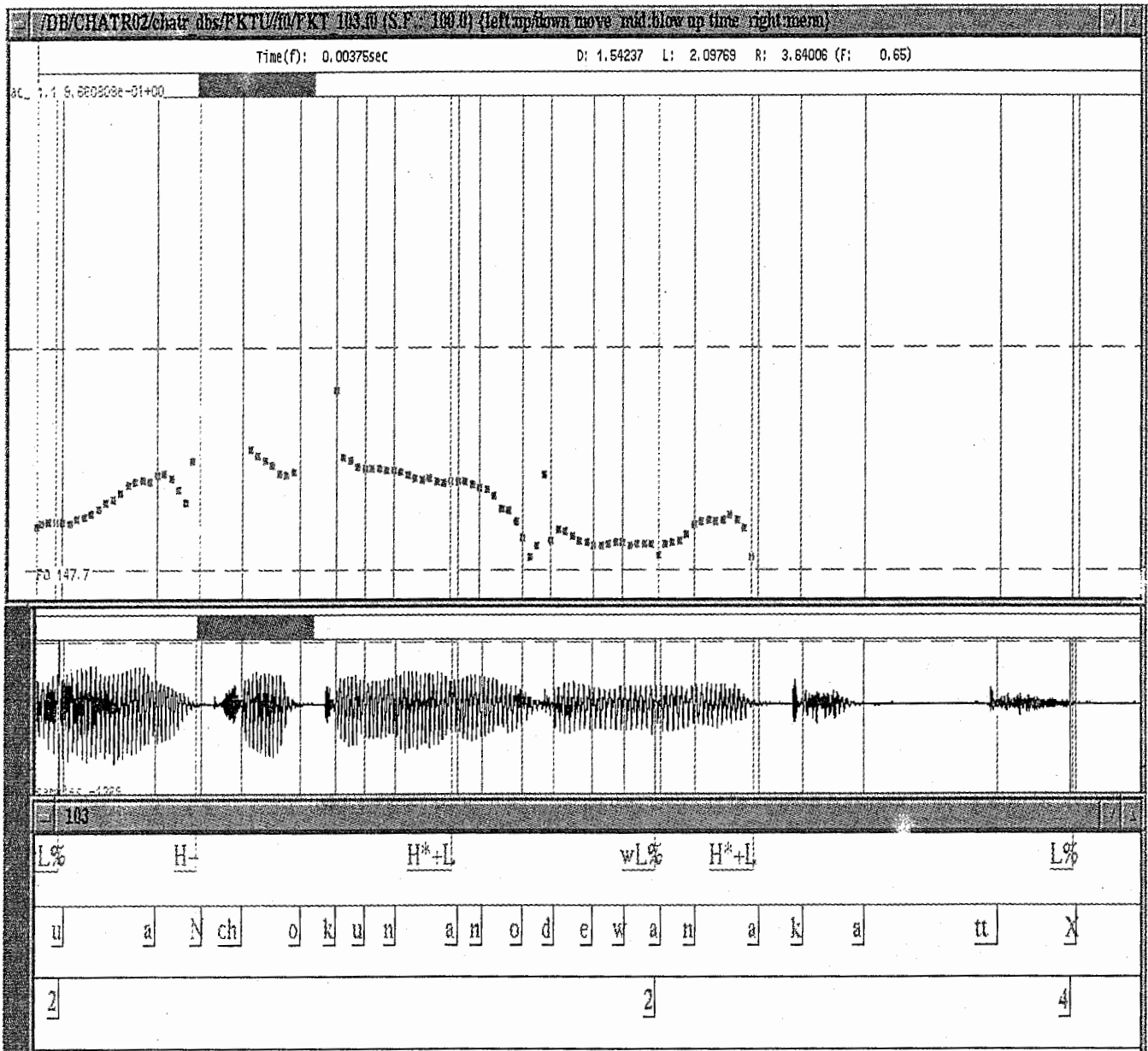
“年は三十二、三” (関西)



“年は三十二、三” (共通語)



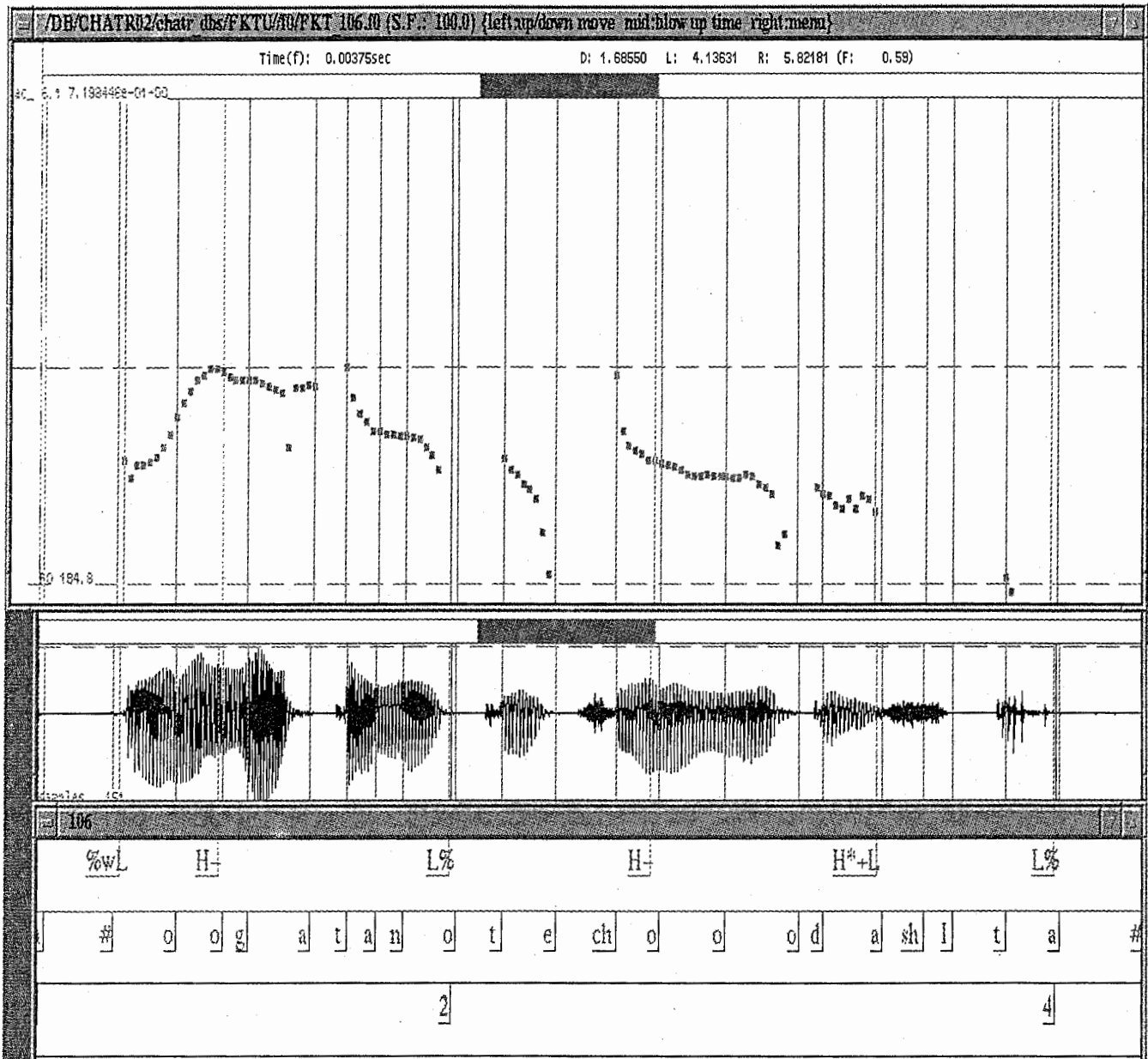
“安直なのではなかった” (関西)



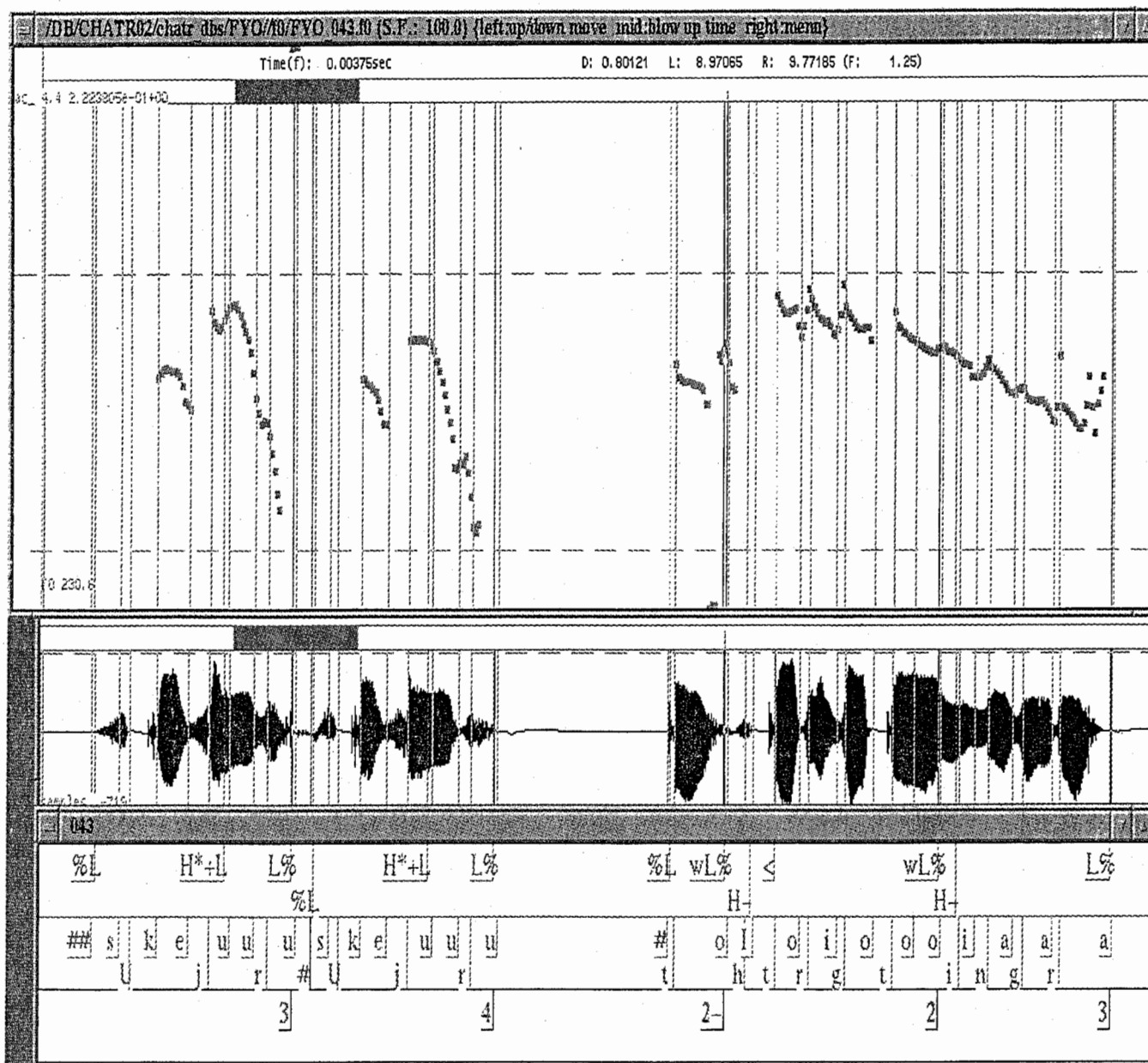
“安直なのではなかった” (共通語)



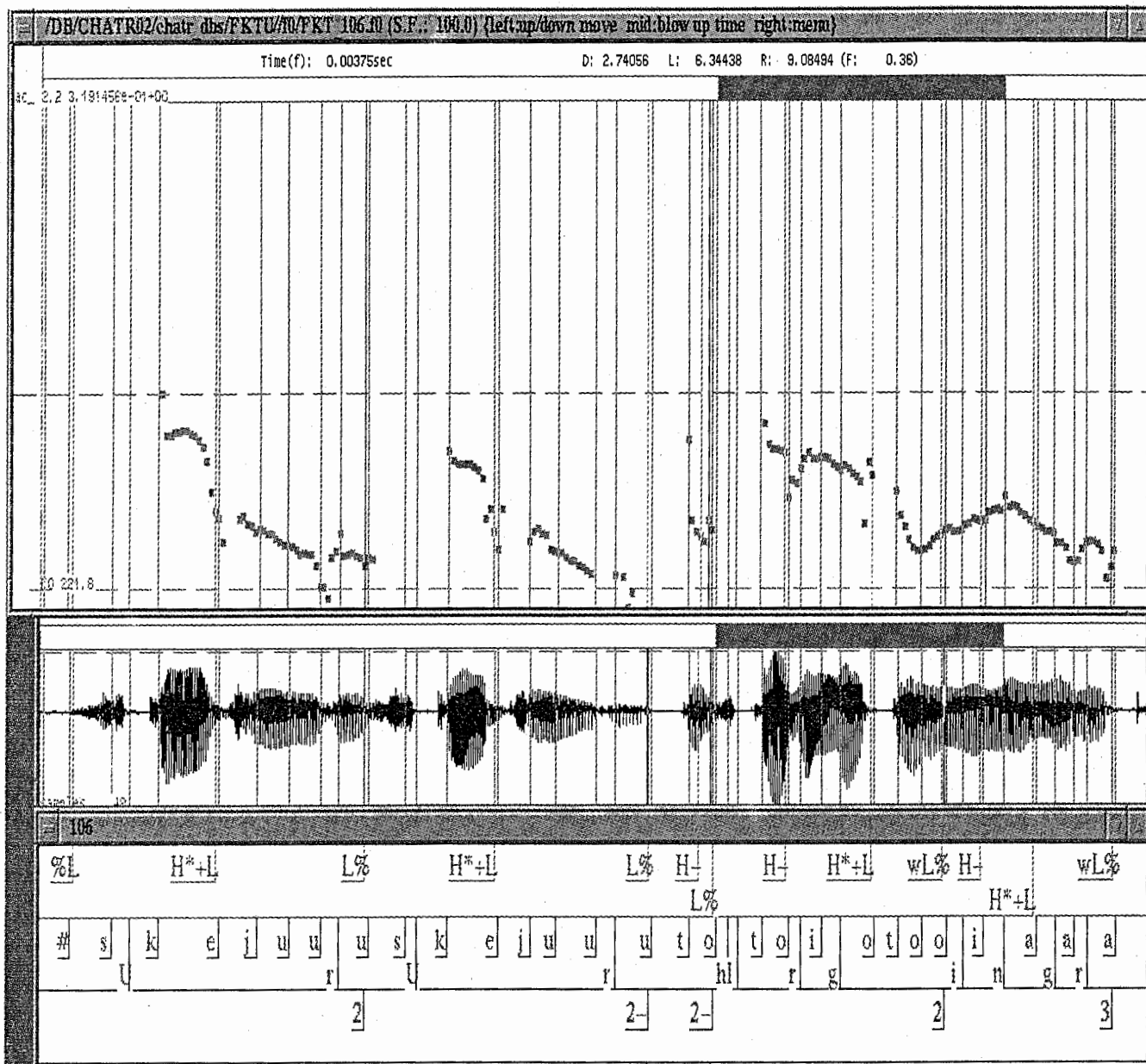




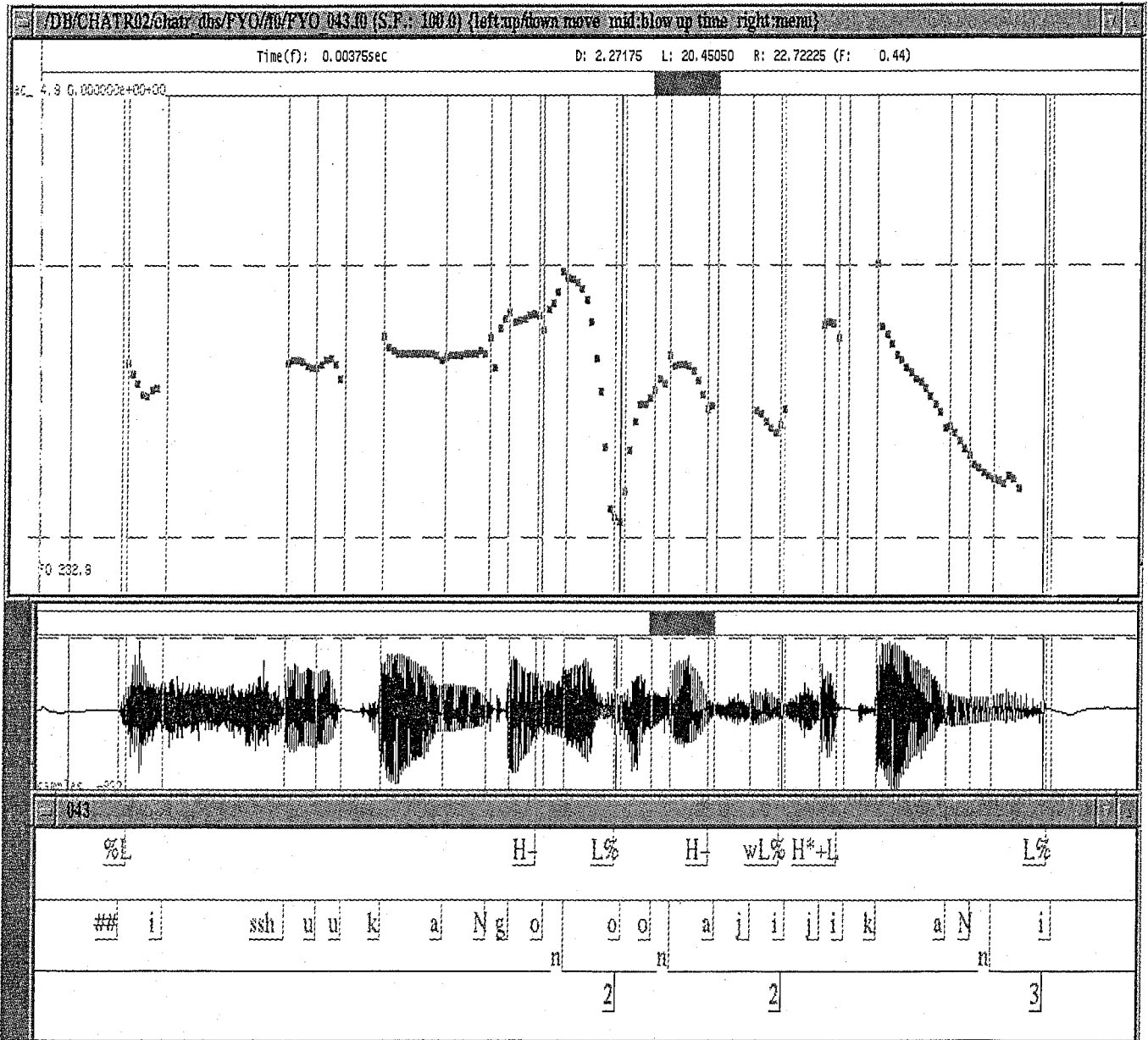
“大型の手帳を出した” (共通語)



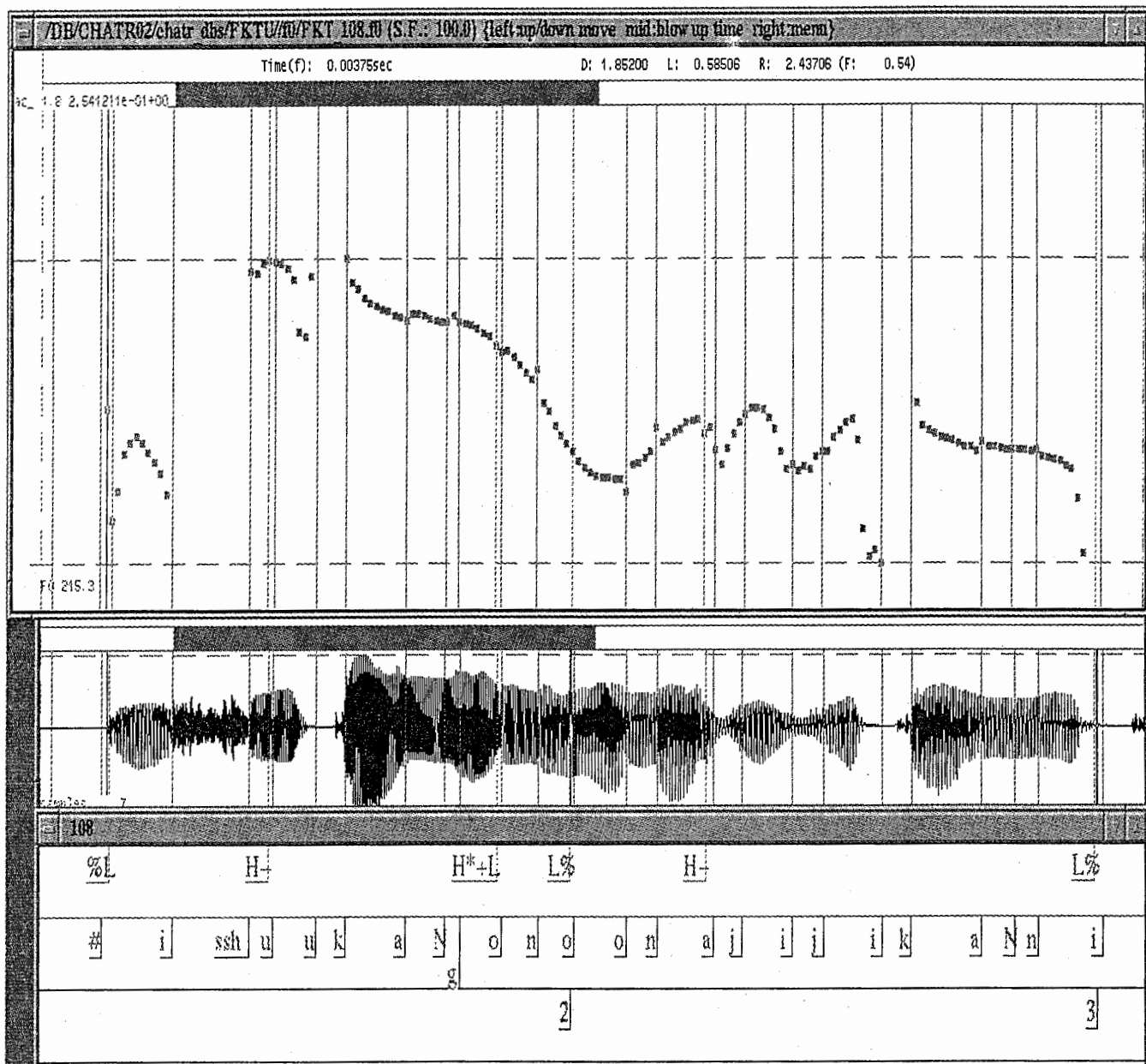
“「スケジュール、スケジュール」とひとりごとを言いながら” (関西)



“「スケジュール、スケジュール」とひとりごとを言いながら” (共通語)



“一週間後の同じ時間に” (関西)



“一週間後の同じ時間に” (共通語)