The *Corpus of Interactional Data*: a Large Multimodal Annotated Resource

Philippe Blache

Laboratoire Parole et Langage
Brain and Language Research Institute
CNRS & Aix-Marseille Université

Outline

- Multimodal annotation: general overview
- The formal background
- Annotation of the different domains in the CID

Part I

Multimodal Annotation: General Overview

Multimodality

- Goals
  - Description of modalities and their interaction
  - Analysis of natural communication
- Different sources of information
  - Different modalities: verbal, non verbal, context, etc.
  - Different domains: phonetics, prosody, syntax, pragmatics, etc.
- Issues
  - Representation, encoding
  - Diversity of annotation tools and formats
  - Alignment vs. synchronization
  - Data manipulation, querying
- Method
  - Rich annotation for each domain
  - Homogeneous framework

Multimodal Corpora: a survey

- Switchboard in NXT (NITE XML Toolkit)
  - 642 conversations; 830,000 words.
  - *Syntax, turns, disfluency, information status, coreference, phonemes, syllables, prosodic phrases, breaks, accents*
- LUNA (Spoken Language Understanding in Multilingual Communication Systems)
  - 8100 human-machine dialogues and 1000 human-human dialogues in Polish, Italian and French.
  - *Turns, POS, chunks, dialogue acts, reference*
- SAMMIE (Saarbrücken Multimodal MP3 Player Interaction Experiment)
  - Multimodal dialogue system, human-machine multimodal interaction (Wizard of Oz)
  - *Transcription, turns, clauses, discourse entities, dialogue acts*

Multimodal Corpora: a survey

- AMI (Augmented Multi-party Interaction)
  - 100h meeting, full manual transcription
  - *Dialogue acts, focus of attention, movement (hand, head, leg), named entities, topic segmentation*
- The ITC Corpus
  - 11 groups of 4 people (25 minutes each). Task: decision making scenario
  - *No transcription, functional role, socio emotional, speech activity, body activity*
- The ATR Corpus
  - 10 meetings, 1 hour each
  - *No transcription, speech activity, body movements, activity type*

Multimodal Corpora

Part II

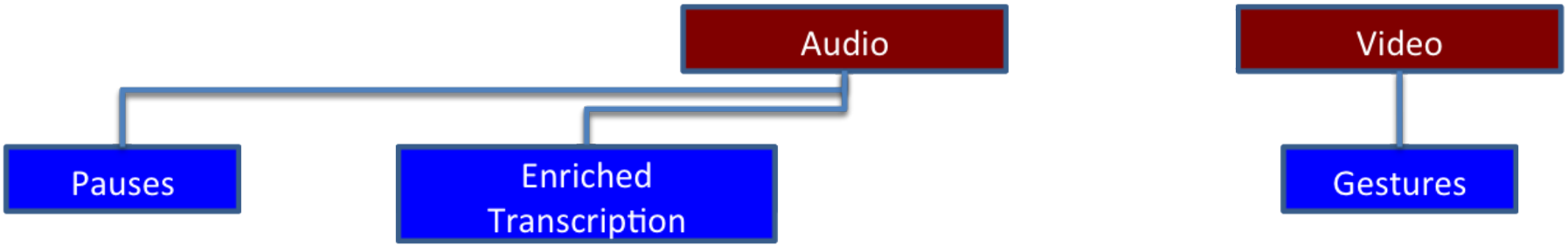
The *Corpus of Interactional Data*: a Large Scale Experiment

CID: main features

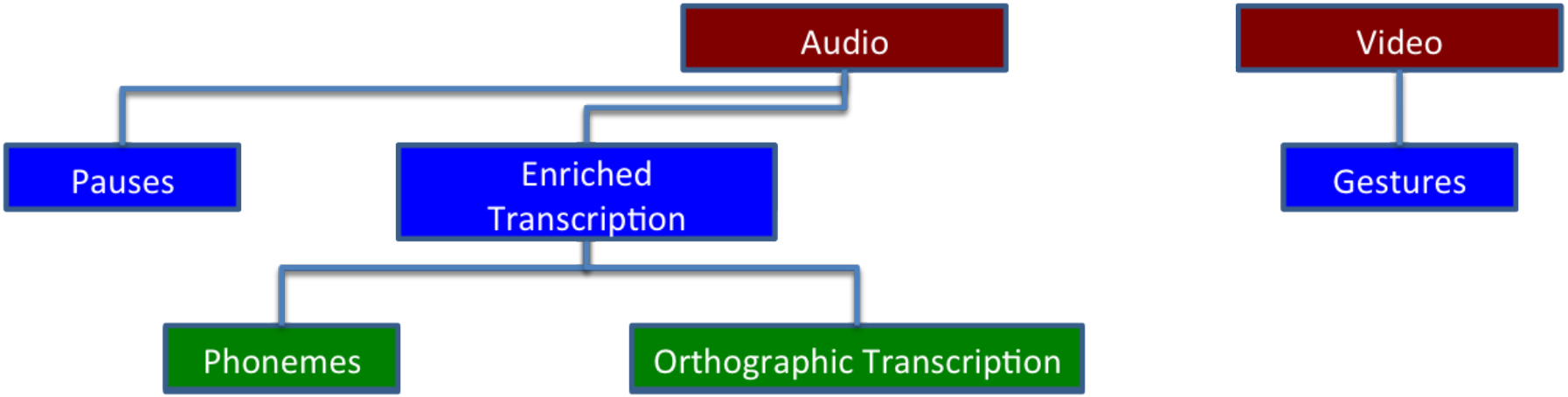
- 8 dialogs, 1 hour each (4 male/male; 4 female/female)
- Task: - *"Tell something unusual which happened to you"*
  - *"Tell about professional conflicts you may have met"*
- Setting
  - Anechoic room
  - 1 camcorder / 2 microphones
- Annotations (aligned on the signal)
  - Phonetic and orthographic transcription
  - Prosody (units, intonation, contours)
  - Morphosyntax, syntax
  - Discourse (markers, turns, etc.)
  - Gestures

Example

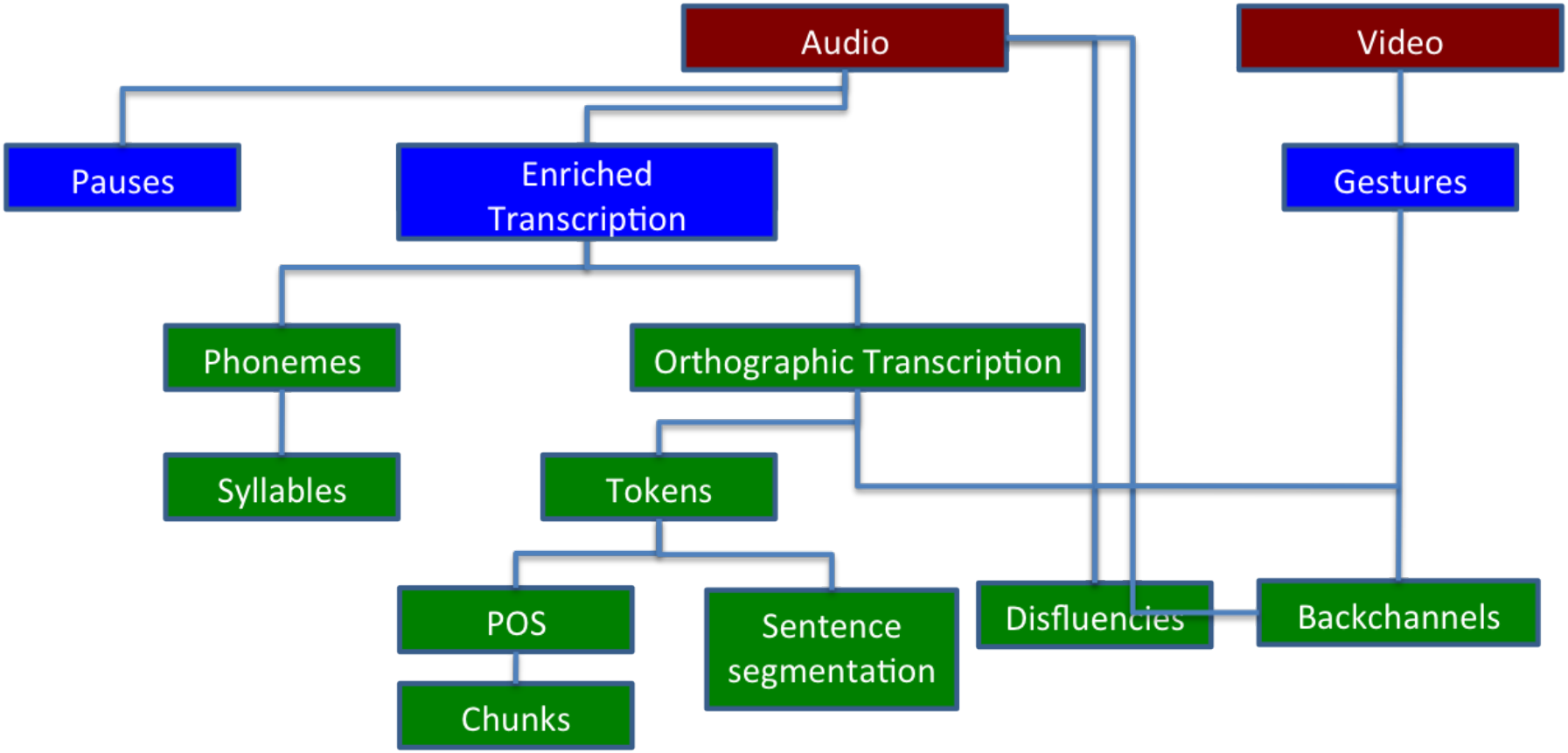
The Annotation Wokflow



The Annotation Architecture



The Annotation Architecture



Main steps and contributions

1. Primary Data Preparation
   - Transcription Convention <<< CID Convention
   - Generation of orthographic and phonetic transcriptions
   - Aligning transcriptions with the signal <<< CID
2. Automatic Annotation
   - Syllabification <<< CID
   - Intonation
   - Sentence segmentation <<< CID
   - POS-tagger
   - Chunker
   - Shallow parser

Main steps and contributions

3. Manual Annotation <<< CID
   - Gestures: hands, head, arms
   - Prosody: phrasing, contours
   - Disfluences
   - Discourse: turns, backchannels, reported speech, information structure

4. Formal representation
   - Abstract schema: Typed Feature Structures <<< CID
   - Generation of the XML schema <<< CID
   - Formatting data
   - Querying

Some descriptions

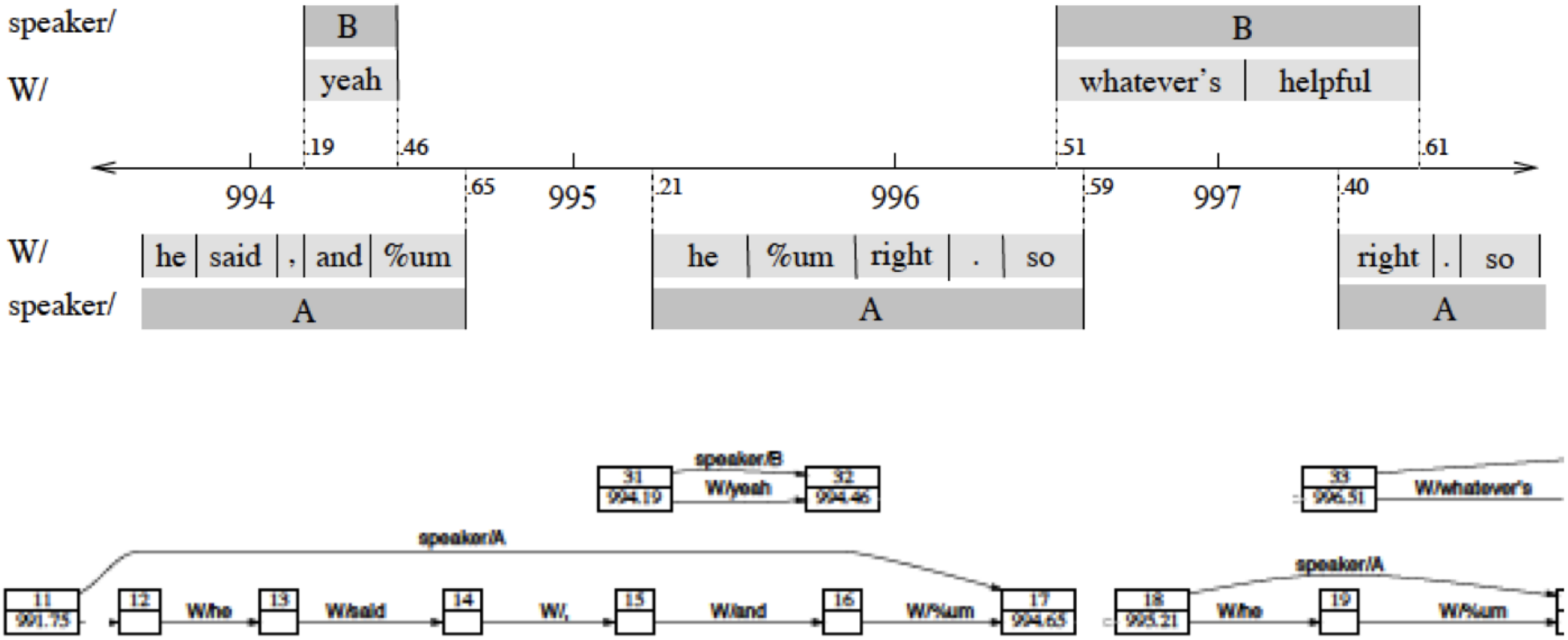
1. Backchannels <<< CID
   - Vocal and gestural
   - Description in terms of prosody, discourse, morpho-syntax

2. Detachments <<< CID
   - Dislocation, cleft, topicalization
   - Annotation of the detachment type, the category, the function, the anaphor

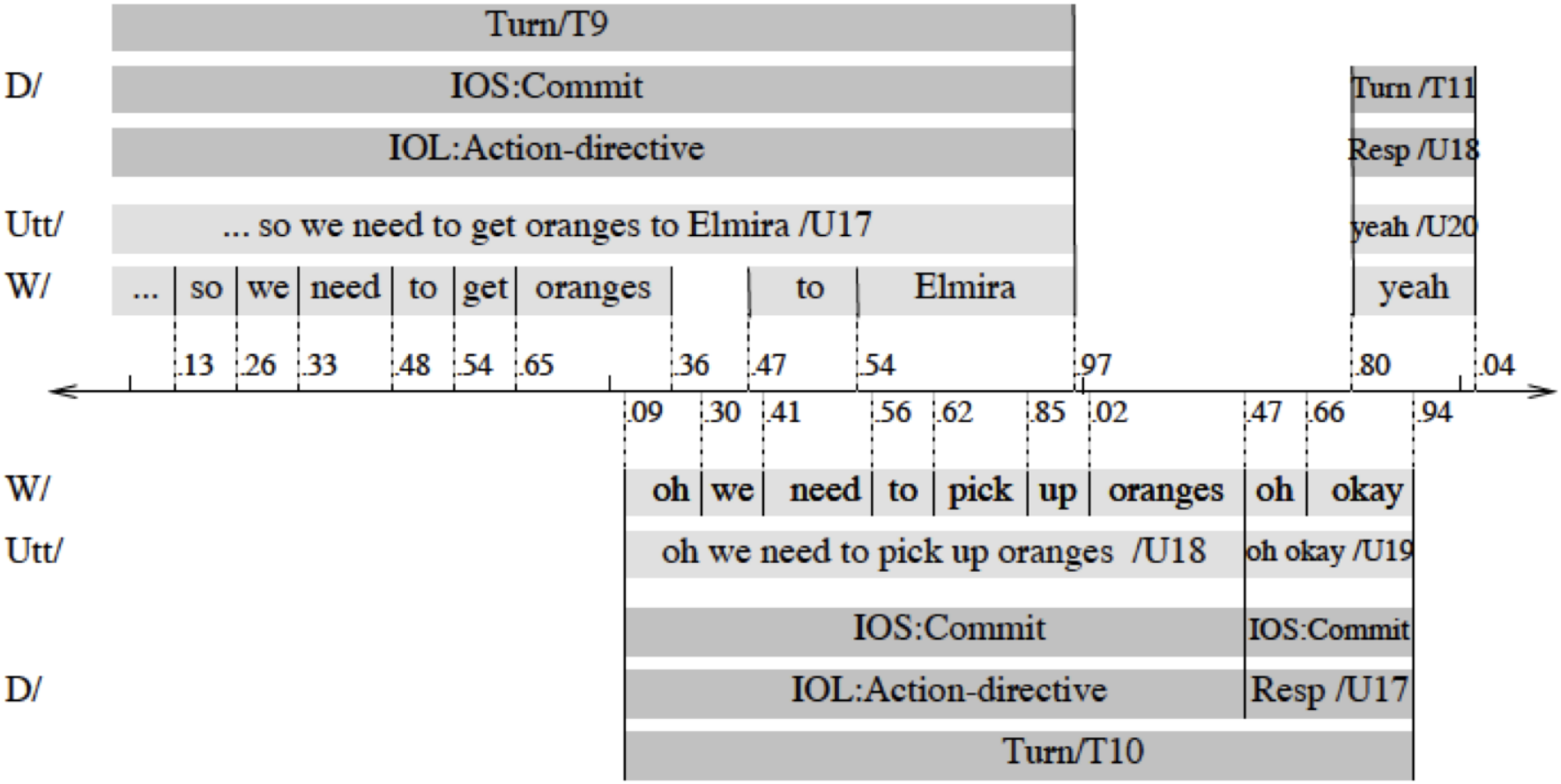
Part III

The Formal Background

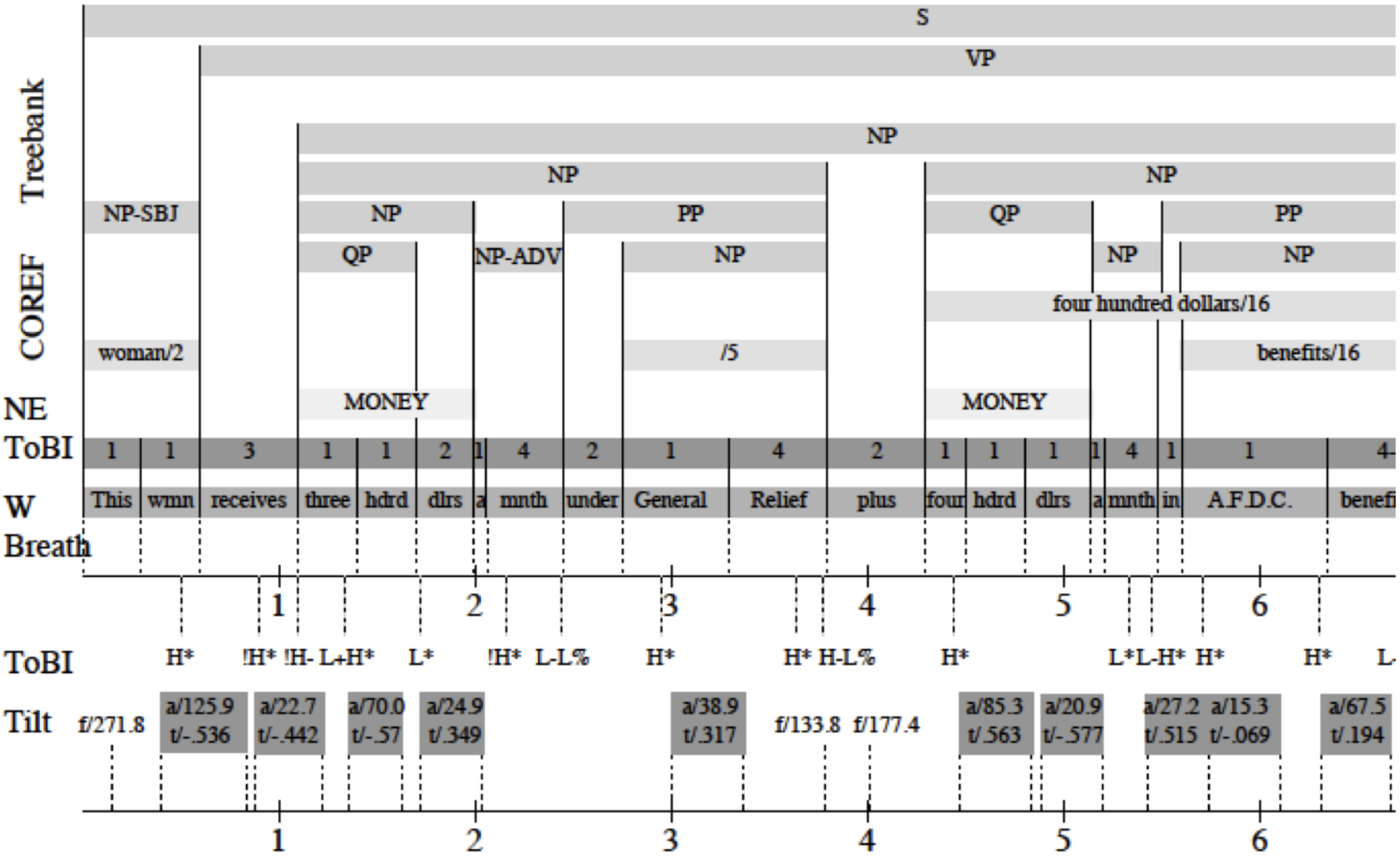
Annotation Graphs



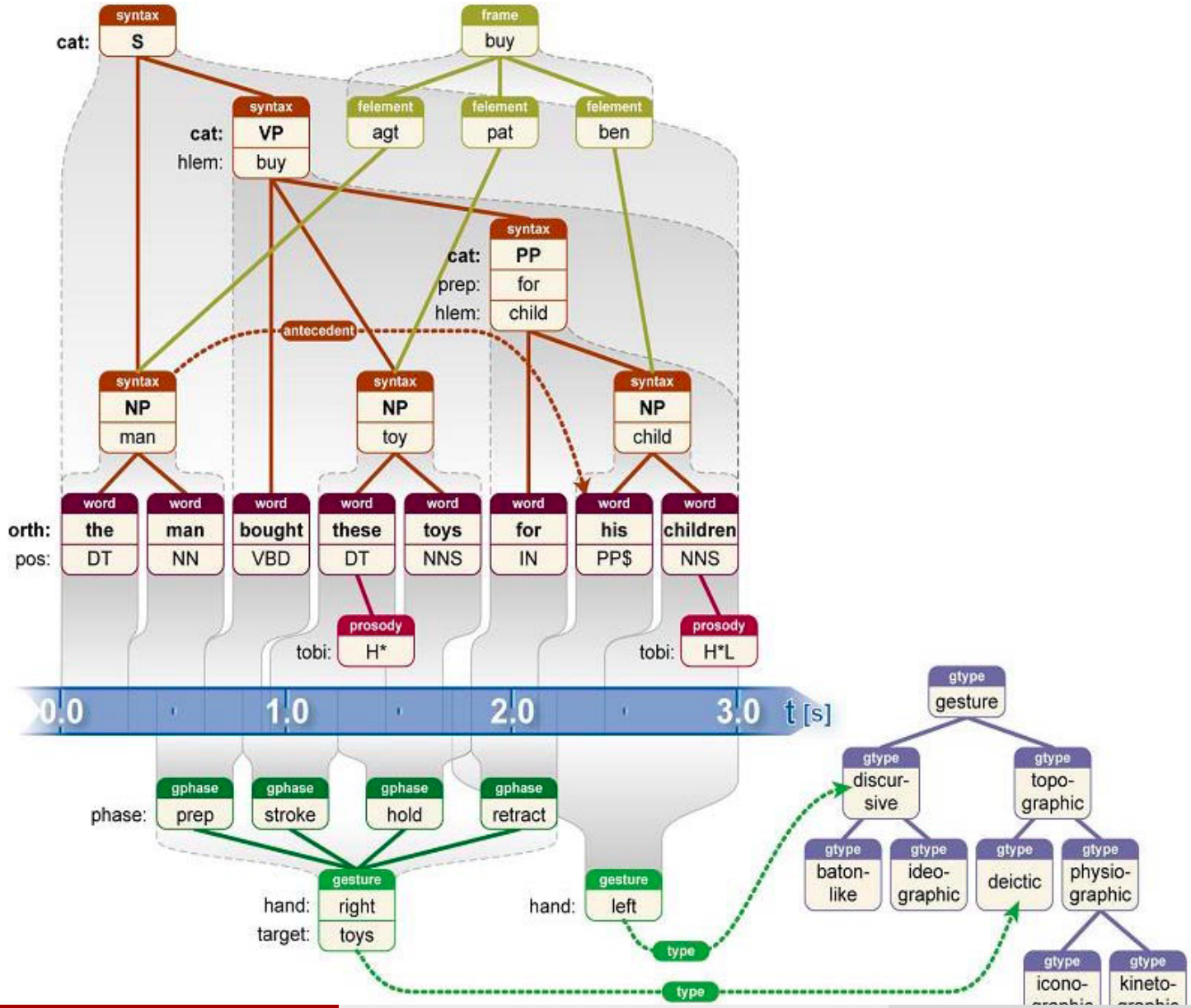
Annotation Graphs



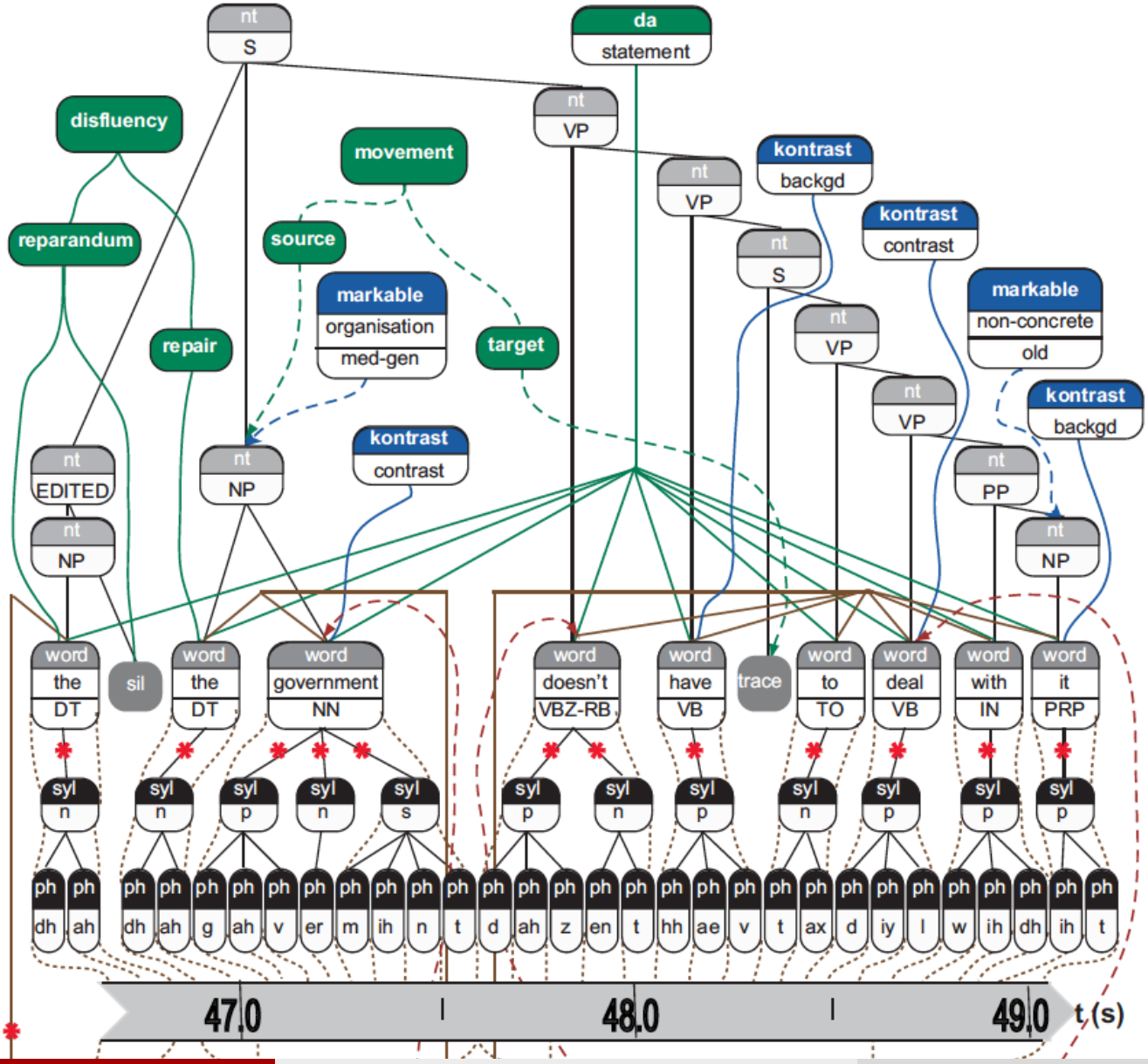
Annotation Graphs



NXT gestures

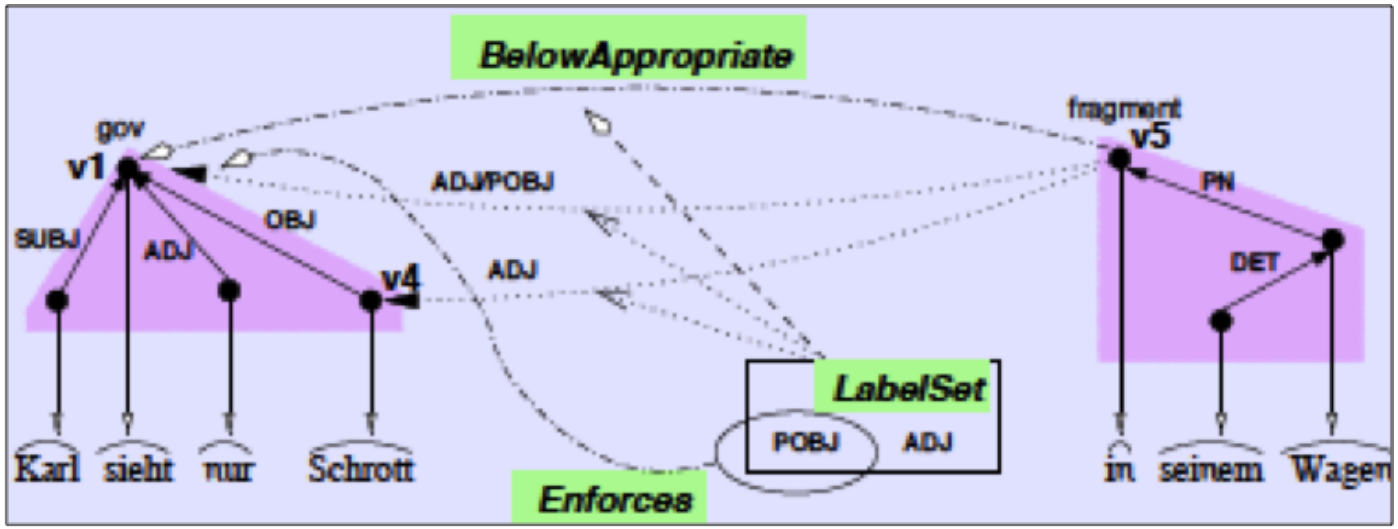


NXT-format Switchboard



Graph Annotation Format (GrAF)

- GrAF: nodes and edges, decorated with feature structures
- Annotations associated to the nodes (rather than the edges as in AG)
- Nodes may be linked to:
  - Primary data
  - Other nodes in the graph



Graph Annotation Format (GrAF)

- Base segmentation:

```
<seg:sink seg:id="42" seg:start="24" seg:end="35"/>
```

- Annotation over the base segmentation:

```
<msd:node msd:id="16">
    <msd:f name="cat" value="NN"/>
</msd:node>
<msd:edge from="msd:16" to="seg:42"/>
```

- Annotation over another annotation:

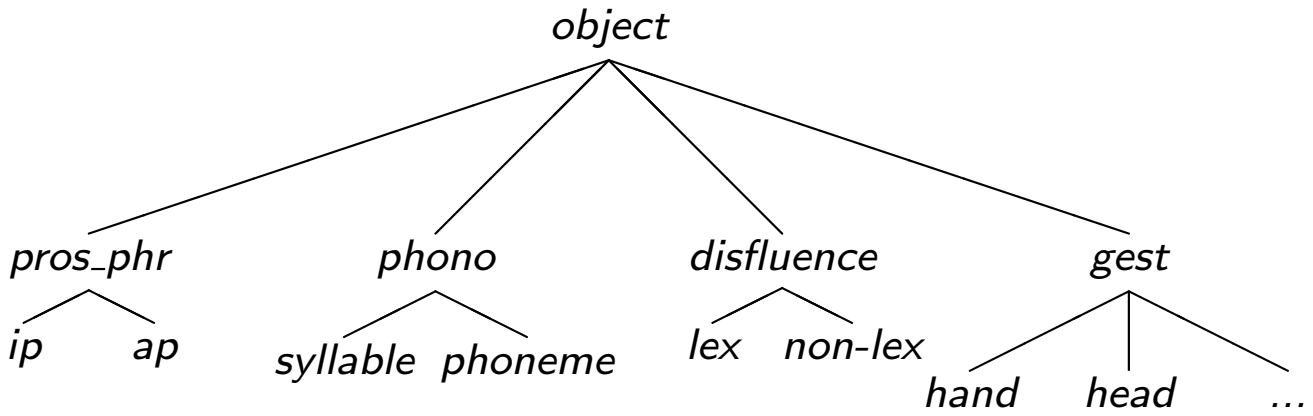
```
<ptb:node ptb:id="23">
    <ptb:f name="type" value="NP"/>
    <ptb:f name="role" value="SBJ"/>
</ptb:node>
<ptb:edge from="ptb:23"to="msd:16"/>
```

A generic scheme

- Needs
  - A mean to describe the information to be encoded, its organization
  - A precise description of:
    - the categories or objects in each domain
    - the organization of each domain
    - the relations between the domains
  - An homogeneous framework for representing all sources of information
  - Independent from a specific tool or formalism
- Solution: *Typed Feature Structure*
  - Description of the objects and their properties
  - Description of the hierarchical structure

An Annotation Scheme in terms of *TFS*

- Type hierarchy:



- Constituency hierarchy:

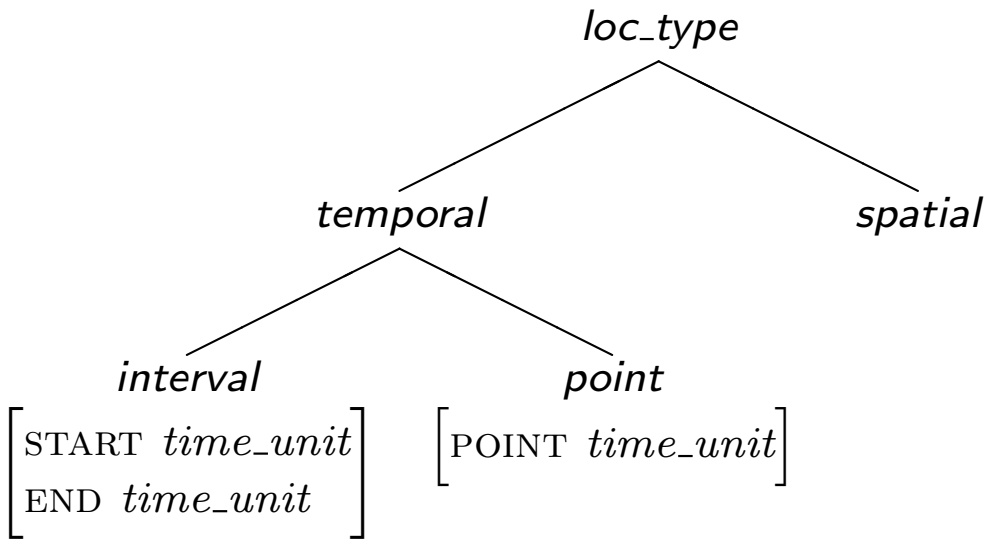
IP ::= AP*
AP ::= SYL$^+$
SYL ::= CONST_SYL$^+$
CONST_SYL ::= PHON$^+$
DISF ::= REPRANDUM BREAK REPARANS

The TFS Schema

- Object type :

$$\textit{object} \begin{bmatrix} \text{INDEX} & \textit{integer} \\ \text{LOCATION} & \textit{loc\_type} \end{bmatrix}$$

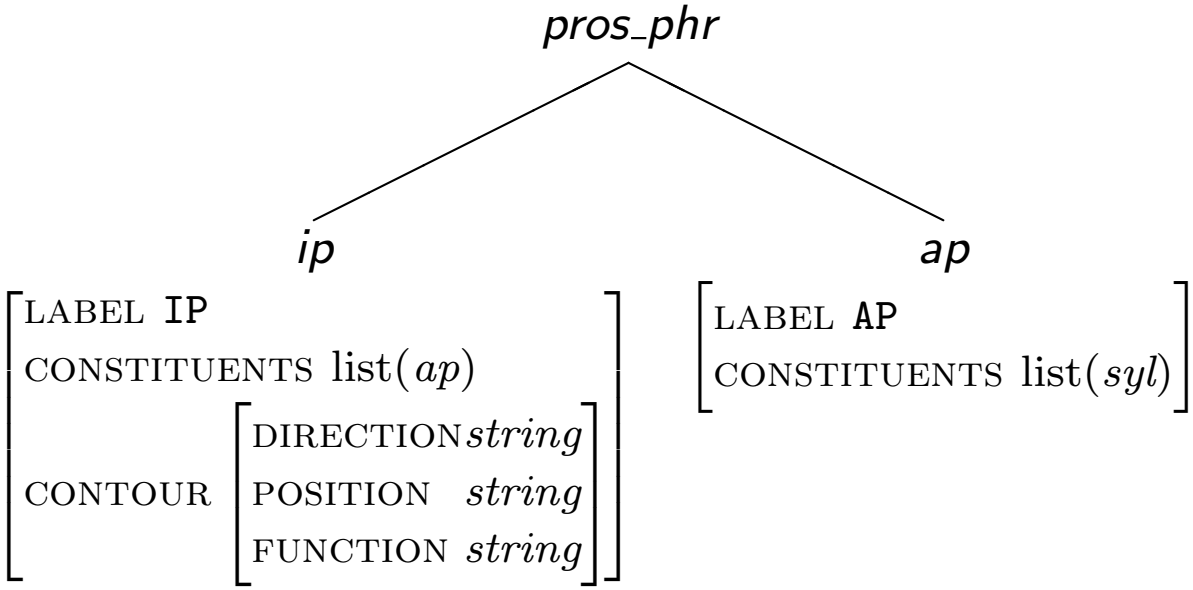
- Location type:



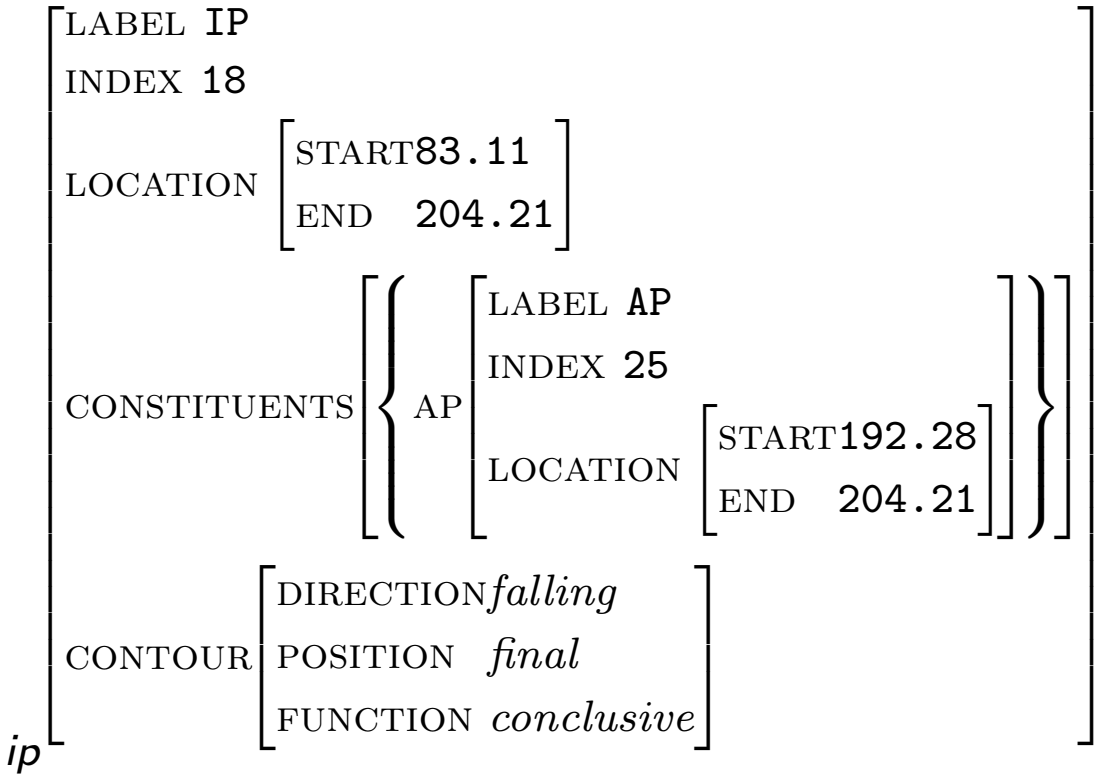
Phonetics

$$
\textit{phon}\left[\begin{array}{l}
\textsc{sampa\_label}\ \textit{sampa\_unit} \\
\textsc{cat}\ \{\textit{vowel},\ \textit{consonant}\} \\
\textsc{type}\ \{\textit{occlusive},\ \textit{fricative},\ \textit{nasal},\ \ldots\} \\
\textsc{articulation}\left[\begin{array}{l}
\textsc{lip}\left[\begin{array}{l}\textsc{protusion}\ \textit{string} \\ \textsc{aperture}\ \textit{aperture}\end{array}\right] \\
\textsc{tongue}\left[\begin{array}{l}\textsc{tip}\left[\begin{array}{l}\textsc{location}\ \textit{string} \\ \textsc{degree}\ \ \textit{string}\end{array}\right] \\ \textsc{body}\left[\begin{array}{l}\textsc{location}\ \textit{string} \\ \textsc{degree}\ \ \textit{string}\end{array}\right]\end{array}\right] \\
\textsc{velum}\ \textit{aperture} \\
\textsc{glottis}\ \textit{aperture}
\end{array}\right] \\
\textsc{role}\left[\begin{array}{l}\textsc{epenthetic}\ \textit{boolean} \\ \textsc{liaison}\ \ \ \ \ \ \textit{boolean}\end{array}\right]
\end{array}\right]
$$

Prosody


pros_phr
ip
ap
LABEL IP
CONSTITUENTS list(ap)
CONTOUR
DIRECTION string
POSITION string
FUNCTION string
LABEL AP
CONSTITUENTS list(syl)

Prosody: Example



Syllabic structure

$$\begin{bmatrix} syl \\ \text{STRUCT } syl\_struct \\ \text{POSITION } \begin{bmatrix} \text{RANK } \{integer\} \\ \text{SYL\_NUMBER } \{integer\} \end{bmatrix} \\ \text{ACCENTUABLE } boolean \\ \text{PROMINENCE } boolean \\ \text{CONSTITUENTS } \text{list}(const\_syl) \end{bmatrix} \begin{bmatrix} const\_syl \\ \text{PHON } \text{list}(phon) \\ \text{CONST\_TYPE } \{onset,\ nucleus,\ coda\} \end{bmatrix}$$

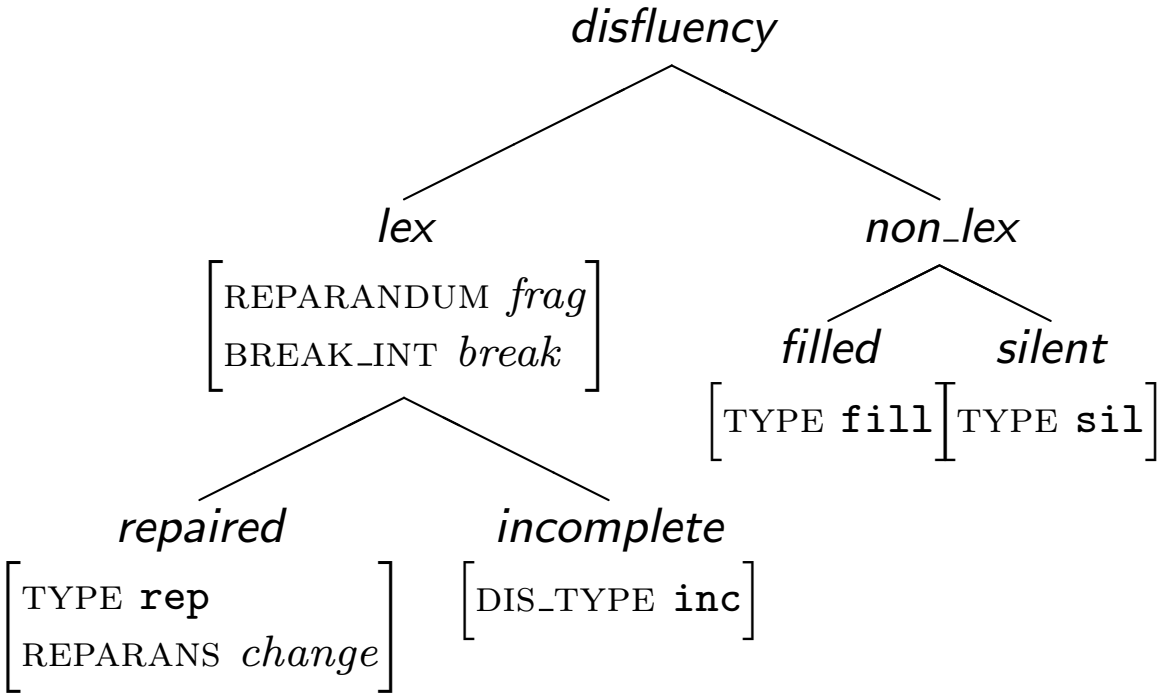
$$\begin{bmatrix} \text{LABEL} & \text{syl} \\ \text{INDEX} & 42 \\ \text{LOCATION} & \begin{bmatrix} \text{START } 195.12 \\ \text{END } 204.21 \end{bmatrix} \\ \text{CONSTITUENTS} & \left\{ \begin{bmatrix} \text{CONST\_TYPE onset} \\ \text{PHON /f/} \end{bmatrix}, \begin{bmatrix} \text{CONST\_TYPE nucleus} \\ \text{PHON /u/} \end{bmatrix}, \begin{bmatrix} \text{CONST\_TYPE coda} \\ \text{PHON /l/} \end{bmatrix} \right\} \\ \text{STRUCT} & \text{CVC} \\ \text{POSITION} & 3/3 \\ \text{ACCENTUABLE} & \text{false} \\ \text{PROMINENCE} & \text{false} \end{bmatrix}$$

Syllabic structure

$$\begin{bmatrix} \textit{syl} \\ \text{STRUCT } \textit{syl\_struct} \\ \text{POSITION } \begin{bmatrix} \text{RANK } \{\textit{integer}\} \\ \text{SYL\_NUMBER } \{\textit{integer}\} \end{bmatrix} \\ \text{ACCENTUABLE } \textit{boolean} \\ \text{PROMINENCE } \textit{boolean} \\ \text{CONSTITUENTS } \text{list}(\textit{const\_syl}) \end{bmatrix} \begin{bmatrix} \textit{const\_syl} \\ \text{PHON } \text{list}(\textit{phon}) \\ \text{CONST\_TYPE } \{\textit{onset},\ \textit{nucleus},\ \textit{coda}\} \end{bmatrix}$$

$$\begin{bmatrix} \text{LABEL} & \texttt{syl} \\ \text{INDEX} & 42 \\ \text{LOCATION} & \begin{bmatrix} \text{START } 195.12 \\ \text{END } 204.21 \end{bmatrix} \\ \text{CONSTITUENTS} & \left\{ \begin{bmatrix} \text{CONST\_TYPE onset} \\ \text{PHON } /\texttt{f}/ \end{bmatrix}, \begin{bmatrix} \text{CONST\_TYPE nucleus} \\ \text{PHON } /\texttt{u}/ \end{bmatrix}, \begin{bmatrix} \text{CONST\_TYPE coda} \\ \text{PHON } /\texttt{l}/ \end{bmatrix} \right\} \\ \text{STRUCT} & \text{CVC} \\ \text{POSITION} & 3/3 \\ \text{ACCENTUABLE} & \text{false} \\ \text{PROMINENCE} & \text{false} \end{bmatrix}$$

Disfluency


disfluency
lex
REPARANDUM frag
BREAK_INT break
non_lex
filled
TYPE fill
silent
TYPE sil
repaired
TYPE rep
REPARANS change
incomplete
DIS_TYPE inc

Part IV

The annotations

Transcription

EOT on y va avec des copains on a(v)ait pris l(e) ferry en *Normandie*, *T*/ p(ui)sque j'avais un frère qui était en $Normandie, T/$ on traverse on a(v)ait passé [une, uneu] nuit épouvantab(le) sur le ferry et euh on arrive à $Londres,T /$ on voit ma soeur e(lle) nous amène dans le [B&B, biainbi] où ...

Tokens on y va avec des copains on avait pris le ferry en Normandie puisque j' avais un frère qui était en Normandie on traverse on avait passé une nuit épouvantable sur le ferry et on arrive à Londres on voit ma soeur elle nous amène dans le B&B où ...

Segmentation

on y va avec des copains /Wm/ on avait pris le ferry en Normandie puisque j' avais un frère qui était en Normandie /Wd/ on traverse /Wm/ on avait passé une nuit épouvantable sur le ferry /Wm/ et on arrive à Londres /Wm/ on voit ma soeur /Wm/ elle nous amène dans le B&B /Wm/ où on devait loger /Wd/ on se promène /Wm/ moi /Wm/ j' étais déjà crevée au bout de trois jours /Wm/ parce qu' on voyageait vachement à pied /Wm/ donc j'en pouvais plus

Phonetic transcription

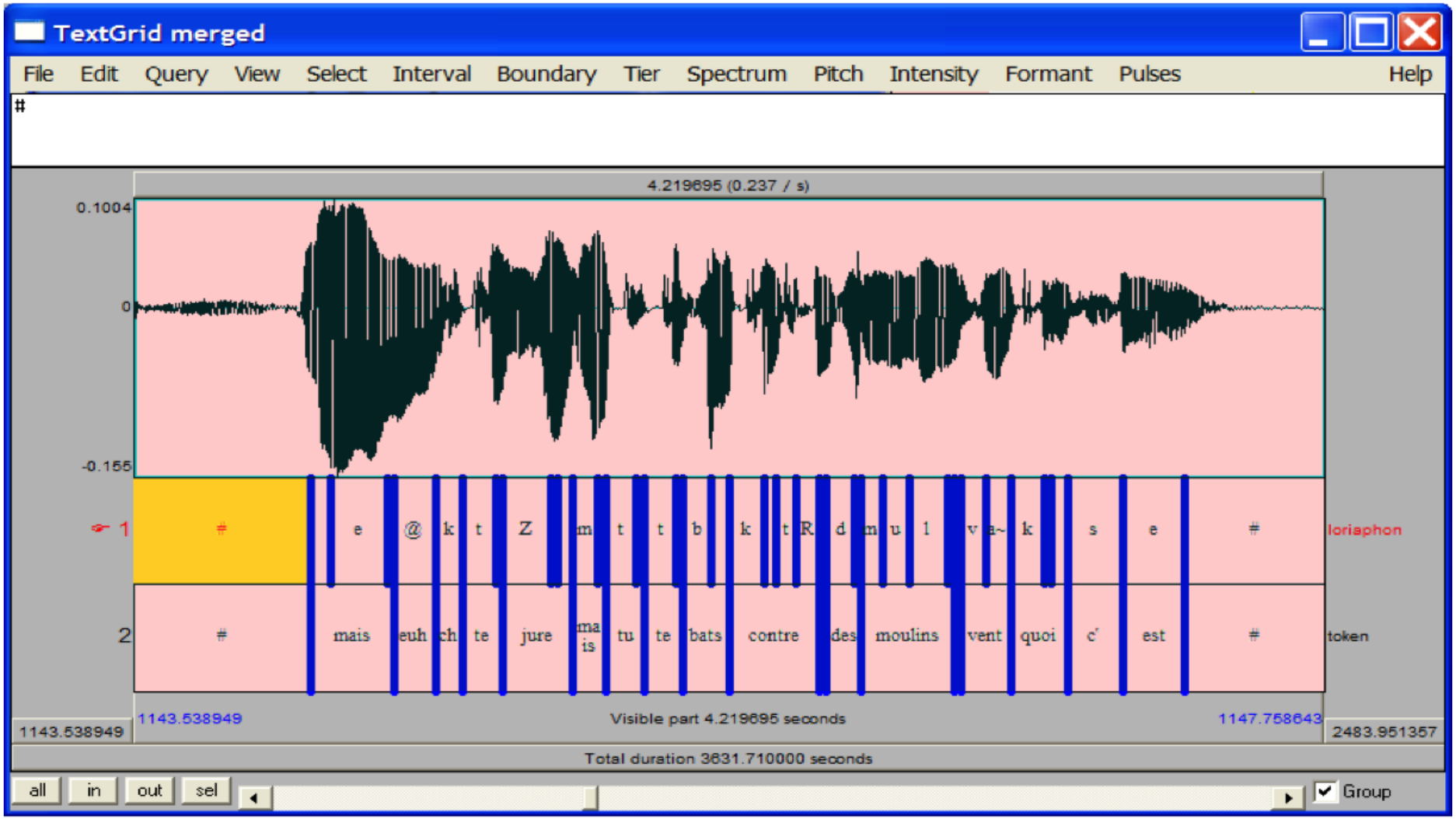
- Grapheme-phoneme conversion
  - Input: enriched transcription
  - Output: list of phonemes, with liaisons

- Example

  et c'est comme en anglais te rappelles pas en anglais quand euh tu épelais ton nom euh tu sais quand tu apprends les lettres

  e s e k o m a~ n a~ g l e t @ R A p e l p A a~ n a~ g l e k a~ @ t y e p @ l e t o~ n o~ t @ t s e k a~ t A p R a~ l e l e t R #

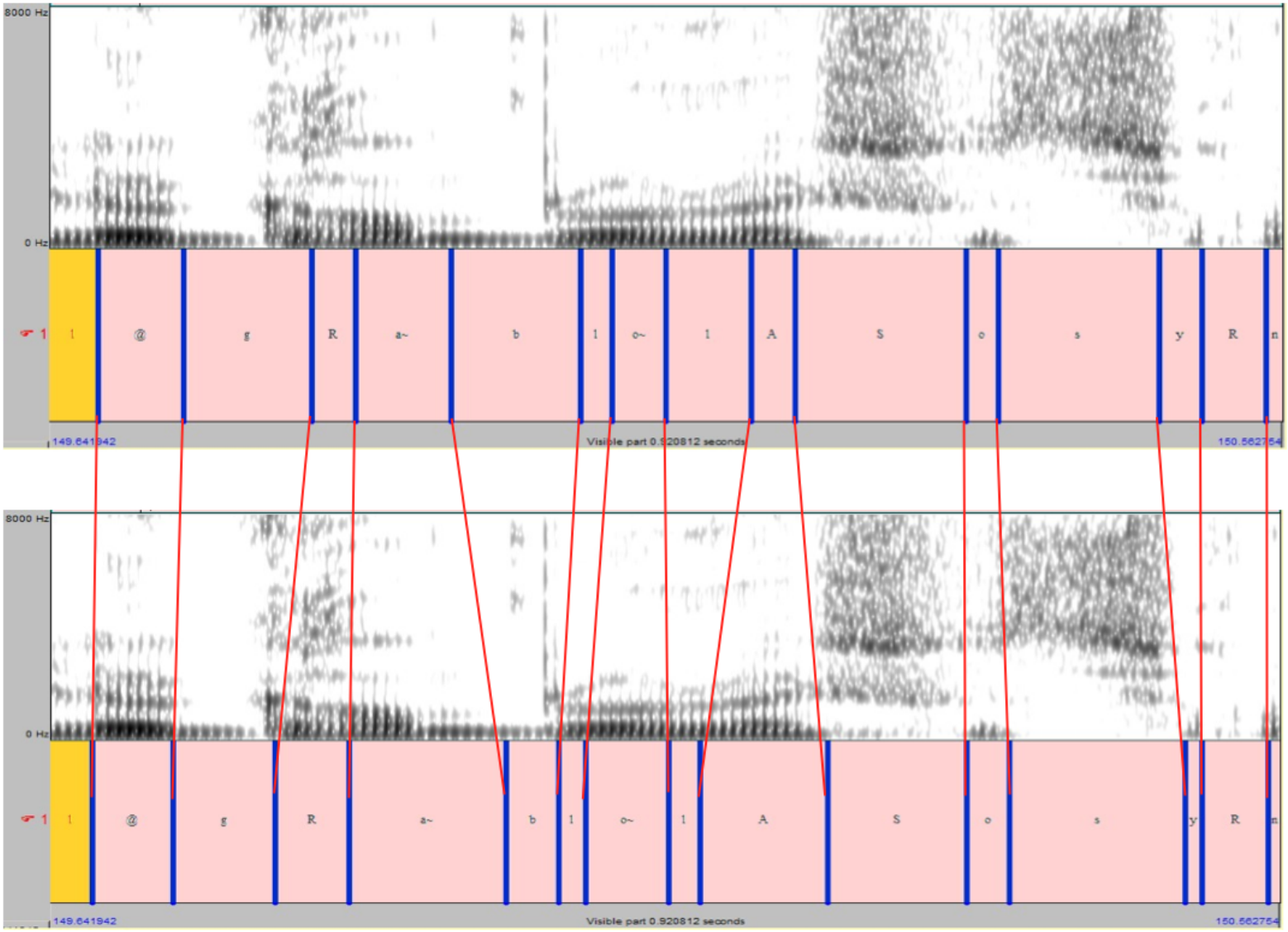
Phonetics: example in Praat



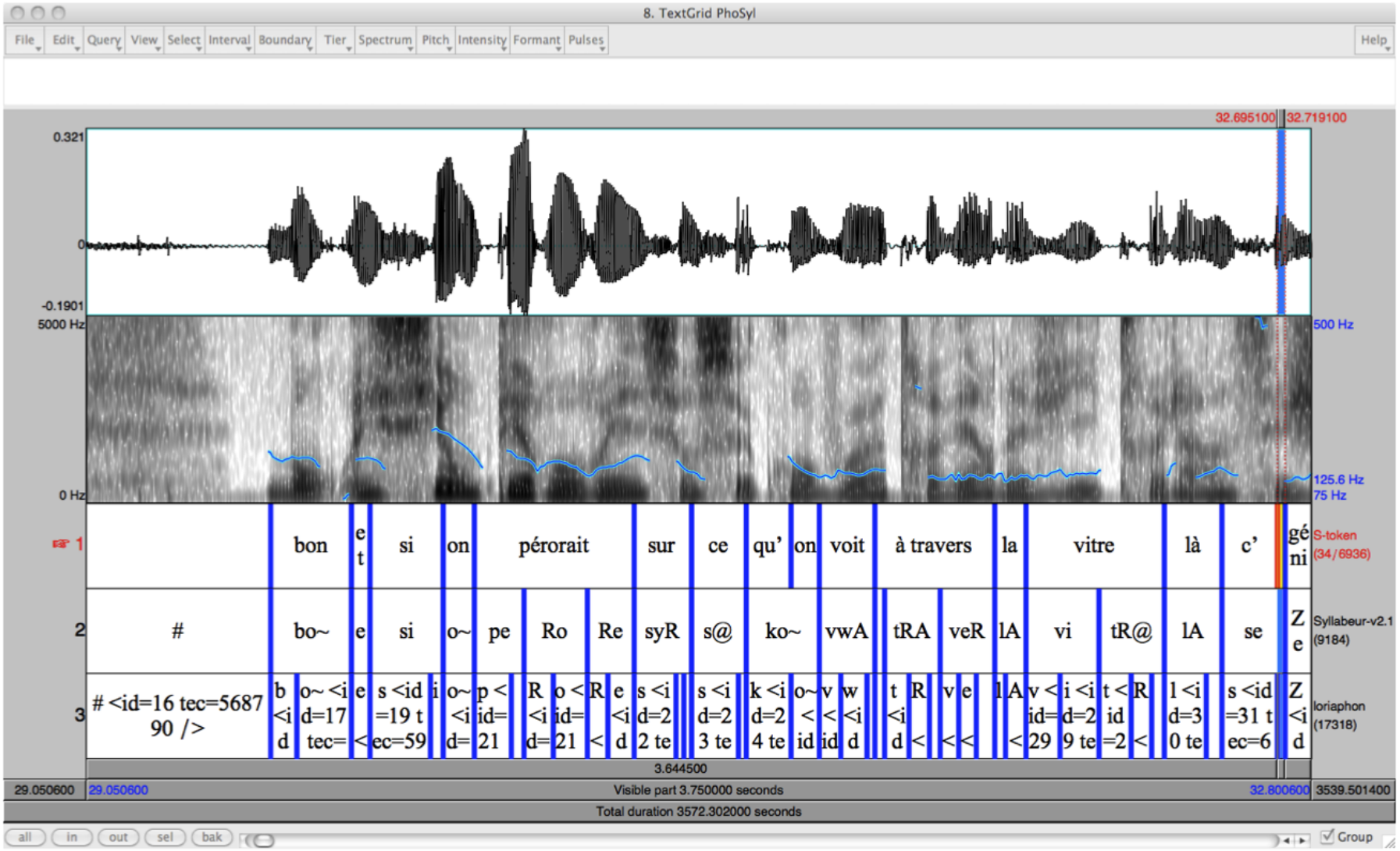
Phonetics: some figures

Phenomenon	Number
Elision	11,058
Word truncation	1,732
Standard liaison missing	160
Unusual liaison	49
Non-standard phonetic realization	2,812
Laugh seq.	2,111
Laughing speech seq.	367
Single laugh IPU	844
Overlaps $>$ 150 ms	4,150

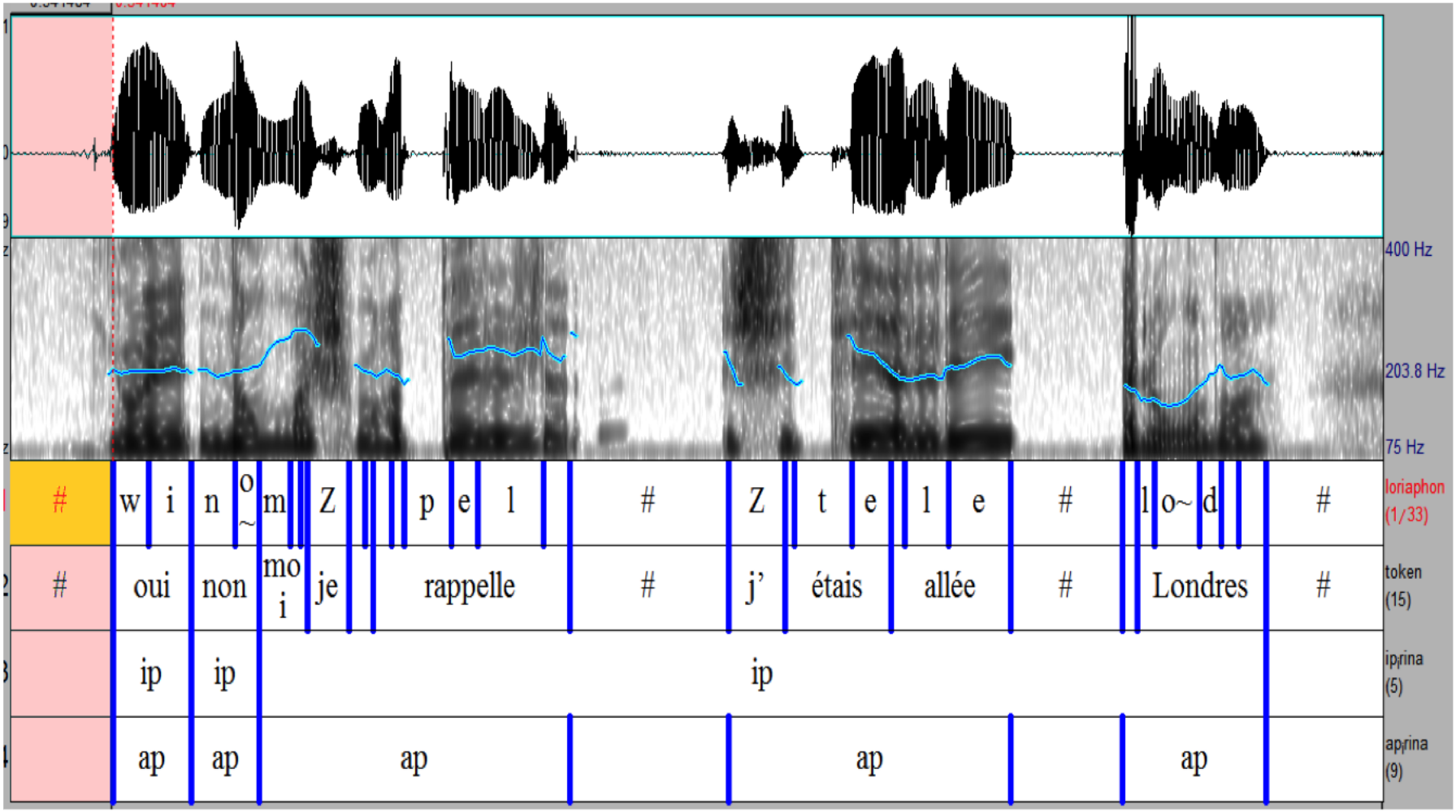
Alignment


8000 Hz
0 Hz
1
@
g
R
a~
b
1
o~
1
A
S
o
s
y
R
n
149.641942
Visible part 0.920812 seconds
150.562754
8000 Hz
0 Hz
1
@
g
R
a~
b
1
o~
1
A
S
o
s
y
R
n
149.641942
Visible part 0.920812 seconds
150.562754

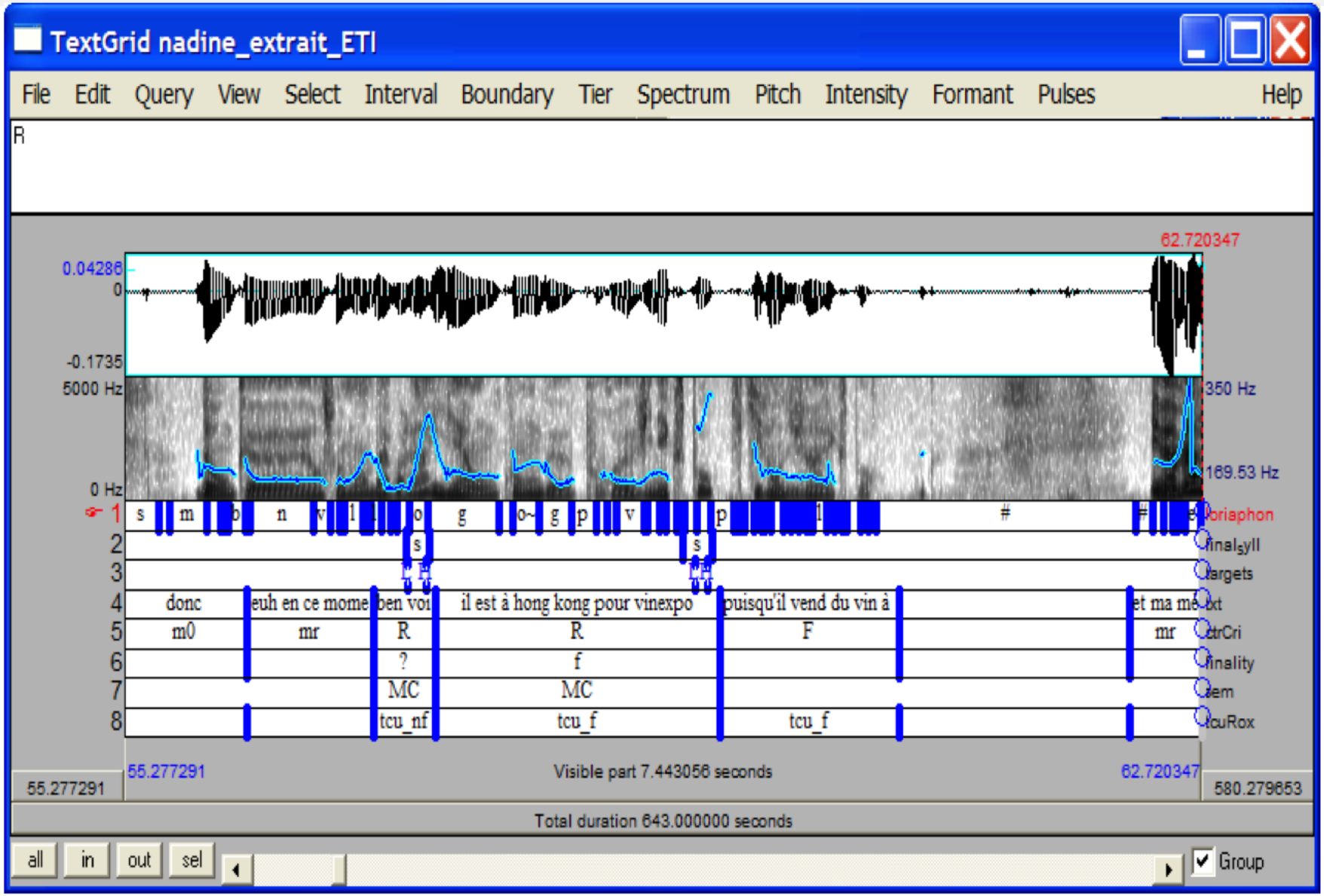
Syllables


8. TextGrid PhoSyl
File Edit Query View Select Interval Boundary Tier Spectrum Pitch Intensity Formant Pulses Help
32.695100 32.719100
0.321
0
-0.1901
5000 Hz
0 Hz
500 Hz
125.6 Hz
75 Hz
1 bon et si on pérorait sur ce qu' on voit à travers la vitre là c' gé ni
S-token (34/6936)
2 # bo~ e si o~ pe Ro Re syR s@ ko~ vwA tRA veR lA vi tR@ lA se Z e
Syllabeur-v2.1 (9184)
3 # <id=16 tec=568790 /> b <id o~ <id d=17 tec= e < s <id =19 tec=59 i o~ <id d= p < id= 21 R <id d= o < id= 21 R < e <id d s <id d=2 2 te s <id d=2 3 te k <id d=2 4 te o~ < id v < id w <id d t <id d R < v < e < l A < v <id= 29 i <id d=2 9 te t < id =2 R < l <id d=3 0 te s <id =31 tec=6 Z <id d
loriaphon (17318)
3.644500
29.050600 29.050600 Visible part 3.750000 seconds 32.800600 3539.501400
Total duration 3572.302000 seconds
all in out sel bak Group

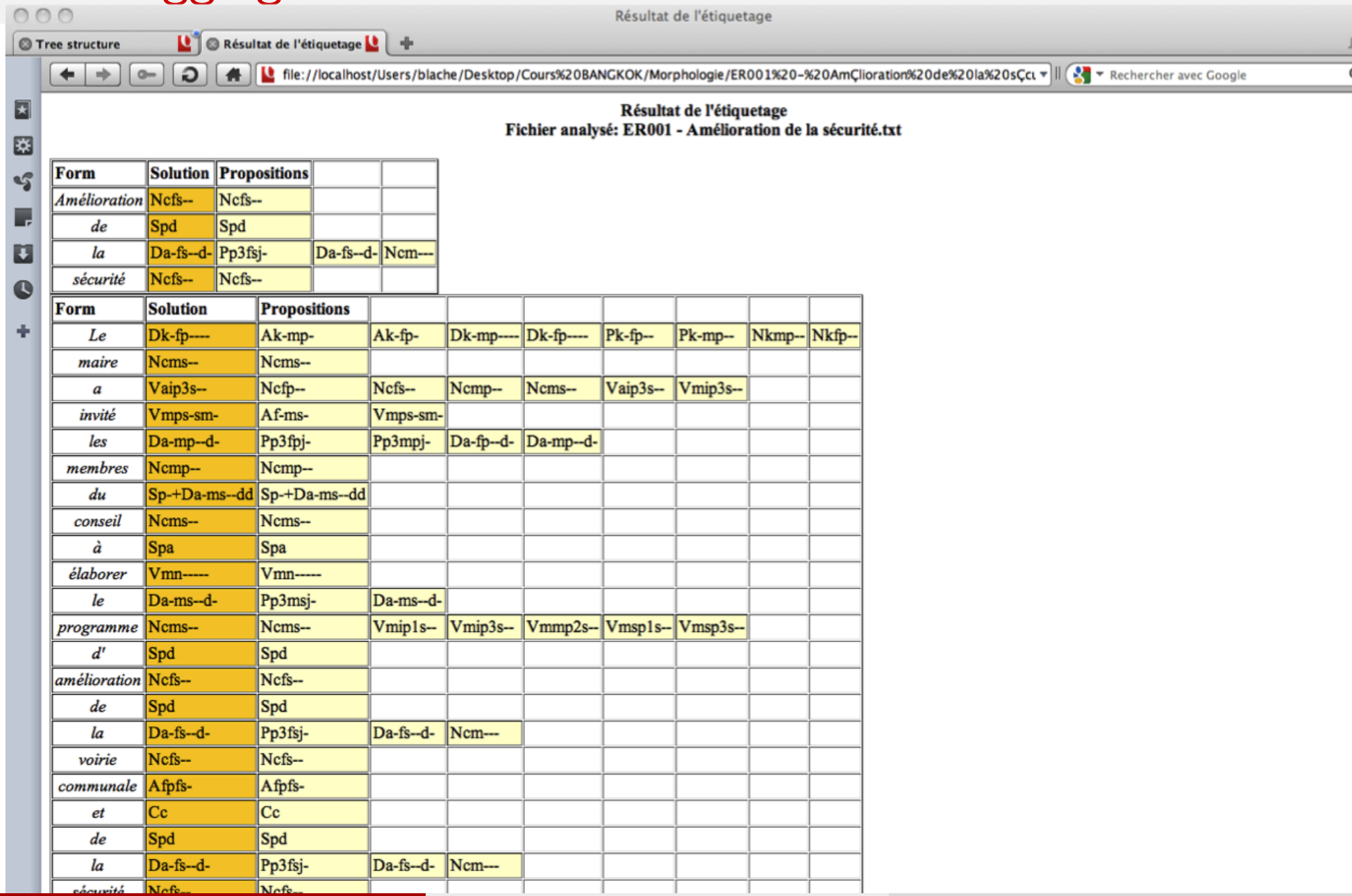
Prosody


400 Hz
203.8 Hz
75 Hz
loriaphon (1/33)
token (15)
ipɾina (5)
apɾina (9)
w i n o~ m Z p e l # Z t e l e # l o~ d
oui non moi je rappelle # j' étais allée # Londres
ip ip ip
ap ap ap ap ap

Prosodic contours



POS-tagging



Résultat de l'étiquetage
Fichier analysé: ER001 - Amélioration de la sécurité.txt

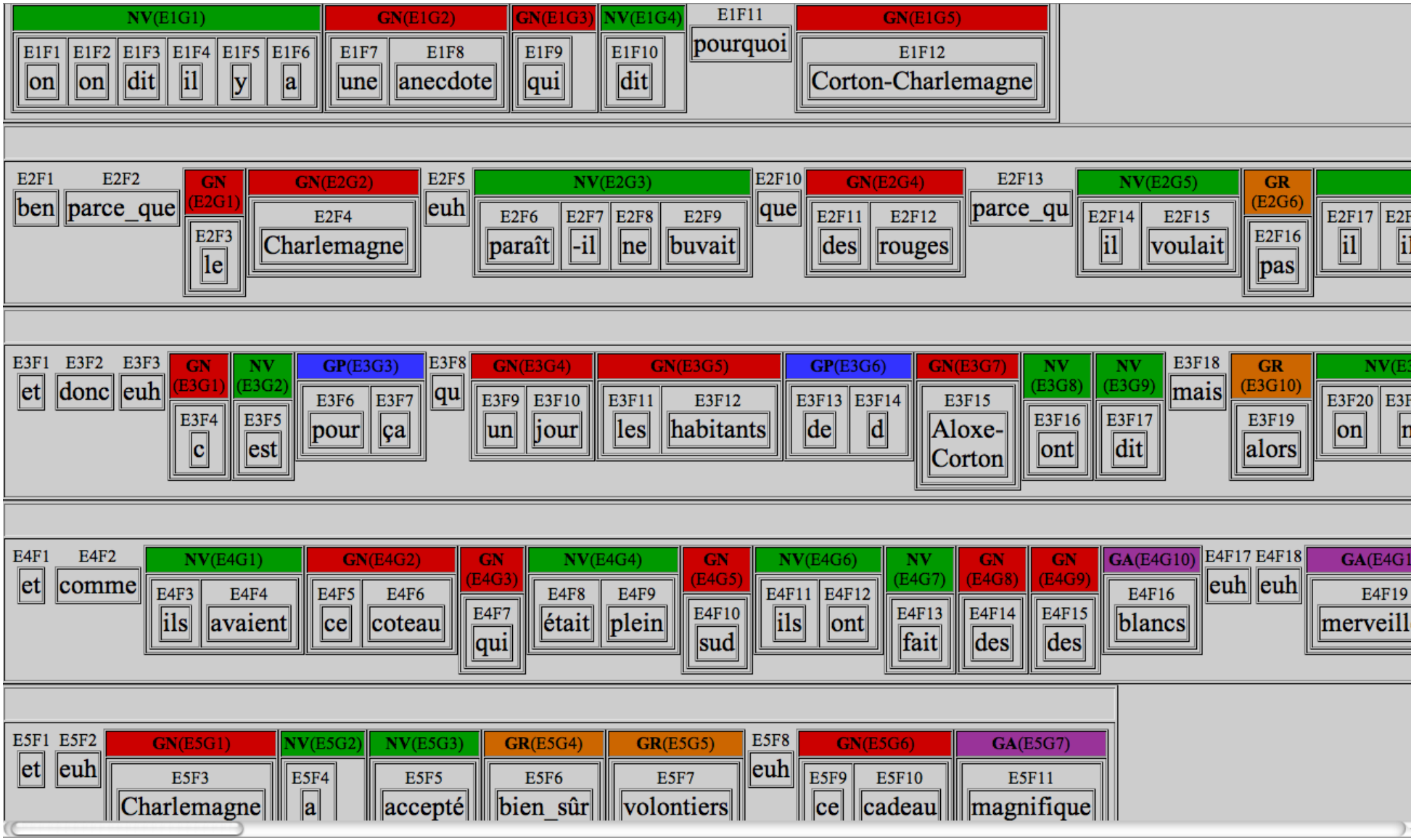
Form	Solution	Propositions		
Amélioration	Ncfs--	Ncfs--		
de	Spd	Spd		
la	Da-fs--d-	Pp3fsj-	Da-fs--d-	Ncm---
sécurité	Ncfs--	Ncfs--		

Form	Solution	Propositions							
Le	Dk-fp----	Ak-mp-	Ak-fp-	Dk-mp----	Dk-fp----	Pk-fp--	Pk-mp--	Nkmp--	Nkfp--
maire	Ncms--	Ncms--							
a	Vaip3s--	Ncfp--	Ncfs--	Ncmp--	Ncms--	Vaip3s--	Vmip3s--		
invité	Vmps-sm-	Af-ms-	Vmps-sm-						
les	Da-mp--d-	Pp3fpj-	Pp3mpj-	Da-fp--d-	Da-mp--d-				
membres	Ncmp--	Ncmp--							
du	Sp-+Da-ms--dd	Sp-+Da-ms--dd							
conseil	Ncms--	Ncms--							
à	Spa	Spa							
élaborer	Vmn-----	Vmn-----							
le	Da-ms--d-	Pp3msj-	Da-ms--d-						
programme	Ncms--	Ncms--	Vmip1s--	Vmip3s--	Vmmp2s--	Vmsp1s--	Vmsp3s--		
d'	Spd	Spd							
amélioration	Ncfs--	Ncfs--							
de	Spd	Spd							
la	Da-fs--d-	Pp3fsj-	Da-fs--d-	Ncm---					
voirie	Ncfs--	Ncfs--							
communale	Afpfs-	Afpfs-							
et	Cc	Cc							
de	Spd	Spd							
la	Da-fs--d-	Pp3fsj-	Da-fs--d-	Ncm---					
sécurité	Ncfs--	Ncfs--							

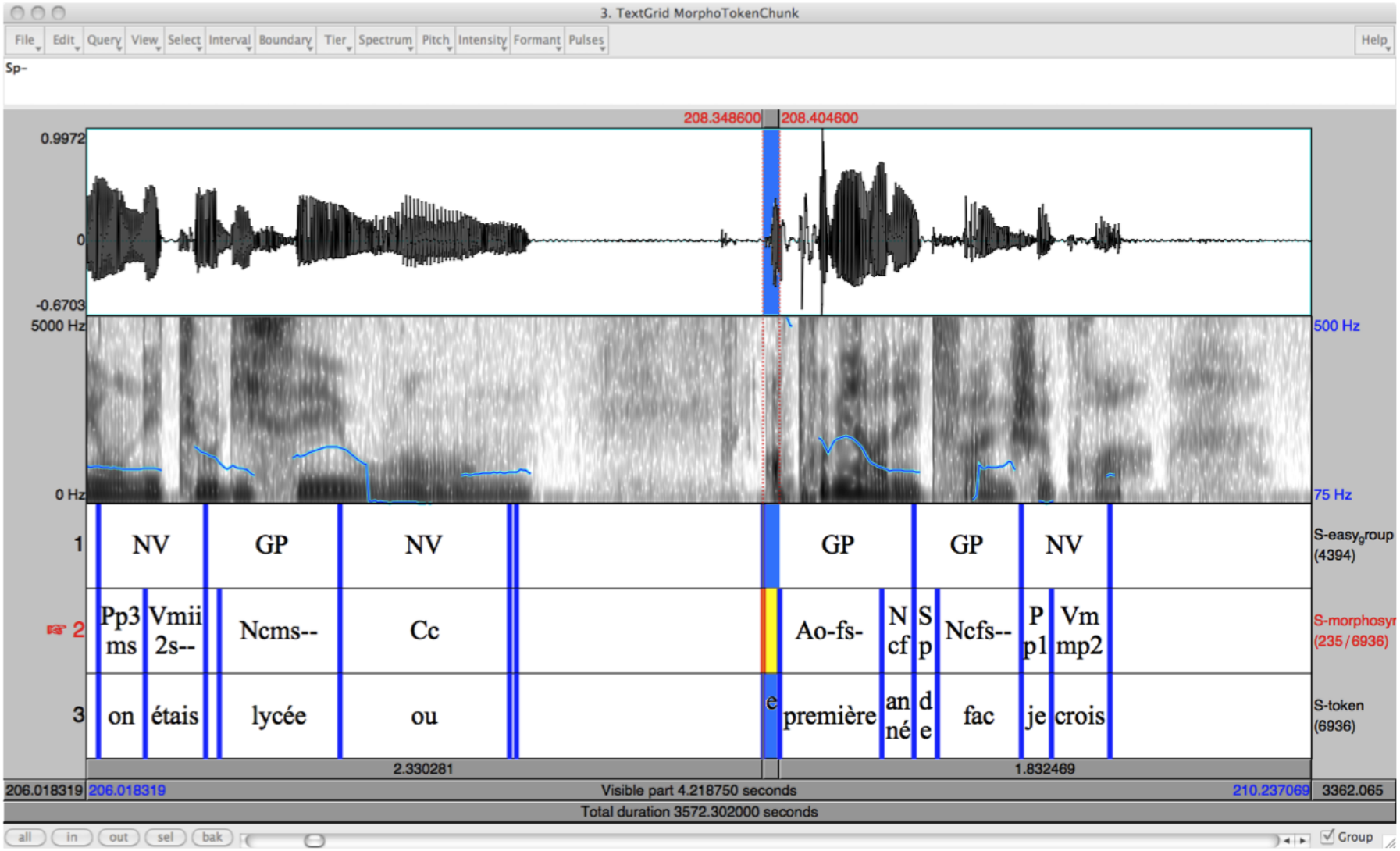
Lexicon

acceptable	0	aksEptabl@	1061	Afpfs-	acceptable
acceptable	0	aksEptabl@	703	Afpms-	acceptable
acceptables	0	aksEptabl@	374	Afpfp-	acceptable
acceptables	0	aksEptabl@	119	Afpmp-	acceptable
acceptai	0	aksEptE	1	Vmis1s--	accepter
acceptaient	0	aksEptE	348	Vmii3p--	accepter
acceptais	0	aksEptE	23	Vmii1s--	accepter
acceptais	0	aksEptE	0	Vmii2s--	accepter
acceptait	0	aksEptE	625	Vmii3s--	accepter
acceptâmes	0	aksEptam@	1	Vmis1p--	accepter
acceptant	0	aksEpta~	1403	Vmpp----	accepter
acceptante	0	aksEpta~t@	0	A--fs-	?
acceptante	0	aksEpta~t@	0	Ncfs--	?
acceptantes	0	aksEpta~t@	0	A--fp-	?
acceptantes	0	aksEpta~t@	0	Ncfp--	?
acceptas	0	aksEpta	0	Vmis2s--	accepter
acceptasse	0	aksEptas@	0	Vmsi1s--	accepter
acceptassent	0	aksEptas@	1	Vmsi3p--	accepter
acceptasses	0	aksEptas@	0	Vmsi2s--	accepter
acceptassiez	0	aksEptasje	0	Vmsi2p--	accepter
acceptassions	0	aksEptasjo~	0	Vmsi1p--	accepter
acceptât	0	aksEpta	6	Vmsi3s--	accepter
acceptâtes	0	aksEptat@	0	Vmis2p--	accepter
acceptation	0	aksEptasjo~	1551	Ncfs--	acceptation
acceptations	0	aksEptasjo~	13	Ncfp--	acceptation
accepté	0	aksEpte	1684	Af-ms-	accepter

Chunking


NV(E1G1)
E1F1 on
E1F2 on
E1F3 dit
E1F4 il
E1F5 y
E1F6 a
GN(E1G2)
E1F7 une
E1F8 anecdote
GN(E1G3)
E1F9 qui
NV(E1G4)
E1F10 dit
E1F11 pourquoi
GN(E1G5)
E1F12 Corton-Charlemagne
E2F1 ben
E2F2 parce_que
GN (E2G1)
E2F3 le
GN(E2G2)
E2F4 Charlemagne
E2F5 euh
NV(E2G3)
E2F6 paraît
E2F7 -il
E2F8 ne
E2F9 buvait
E2F10 que
GN(E2G4)
E2F11 des
E2F12 rouges
E2F13 parce_qu
NV(E2G5)
E2F14 il
E2F15 voulait
GR (E2G6)
E2F16 pas
E2F17 il
E3F1 et
E3F2 donc
E3F3 euh
GN (E3G1)
E3F4 c
NV (E3G2)
E3F5 est
GP(E3G3)
E3F6 pour
E3F7 ça
E3F8 qu
GN(E3G4)
E3F9 un
E3F10 jour
GN(E3G5)
E3F11 les
E3F12 habitants
GP(E3G6)
E3F13 de
E3F14 d
GN(E3G7)
E3F15 Aloxe-Corton
NV (E3G8)
E3F16 ont
NV (E3G9)
E3F17 dit
E3F18 mais
GR (E3G10)
E3F19 alors
E3F20 on
E4F1 et
E4F2 comme
NV(E4G1)
E4F3 ils
E4F4 avaient
GN(E4G2)
E4F5 ce
E4F6 coteau
GN (E4G3)
E4F7 qui
NV(E4G4)
E4F8 était
E4F9 plein
GN (E4G5)
E4F10 sud
NV(E4G6)
E4F11 ils
E4F12 ont
NV (E4G7)
E4F13 fait
GN (E4G8)
E4F14 des
GN (E4G9)
E4F15 des
GA(E4G10)
E4F16 blancs
E4F17 euh
E4F18 euh
E4F19 merveill
E5F1 et
E5F2 euh
GN(E5G1)
E5F3 Charlemagne
NV(E5G2)
E5F4 a
NV(E5G3)
E5F5 accepté
GR(E5G4)
E5F6 bien_sûr
GR(E5G5)
E5F7 volontiers
E5F8 euh
GN(E5G6)
E5F9 ce
E5F10 cadeau
GA(E5G7)
E5F11 magnifique

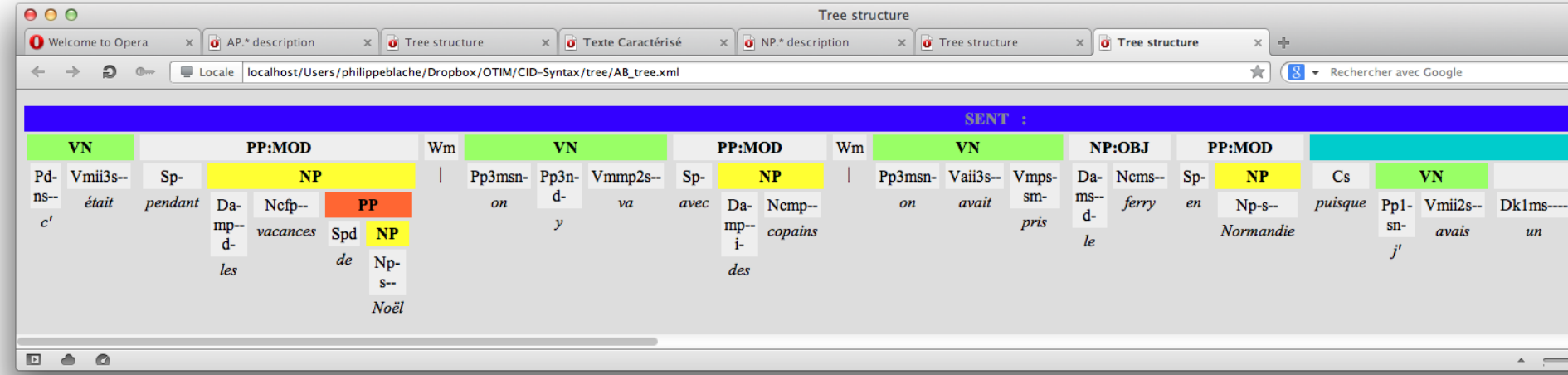
Chunking (2)


3. TextGrid MorphoTokenChunk
File Edit Query View Select Interval Boundary Tier Spectrum Pitch Intensity Formant Pulses Help
Sp-
208.348600
208.404600
0.9972
0
-0.6703
5000 Hz
0 Hz
500 Hz
75 Hz
1
NV GP NV GP GP NV
S-easy_group (4394)
2
Pp3 ms Vmii 2s-- Ncms-- Cc Ao-fs- N cf S p Ncfs-- P p1 Vm mp2
S-morphosyn (235/6936)
3
on étais lycée ou e première an né d e fac je crois
S-token (6936)
2.330281
1.832469
206.018319
206.018319
Visible part 4.218750 seconds
210.237069
3362.065
Total duration 3572.302000 seconds
all in out sel bak
Group

Some results

Category	Count	Group	Count
Adverb	15 123	AP	3 634
Adjective	4 585	NP	13 107
Auxiliary	3 057	PP	7 041
Determiner	9 427	AdvP	15 040
Conjunction	9 390	VPn	22 925
Interjection	5 068	VP	1 323
Preposition	8 693	Total	63 070
Pronoun	25 199		
Noun	13 419	Soft Pct	9 689
Verb	20 436	Strong Pct	14 459
Total	11 4397	Total	24 148

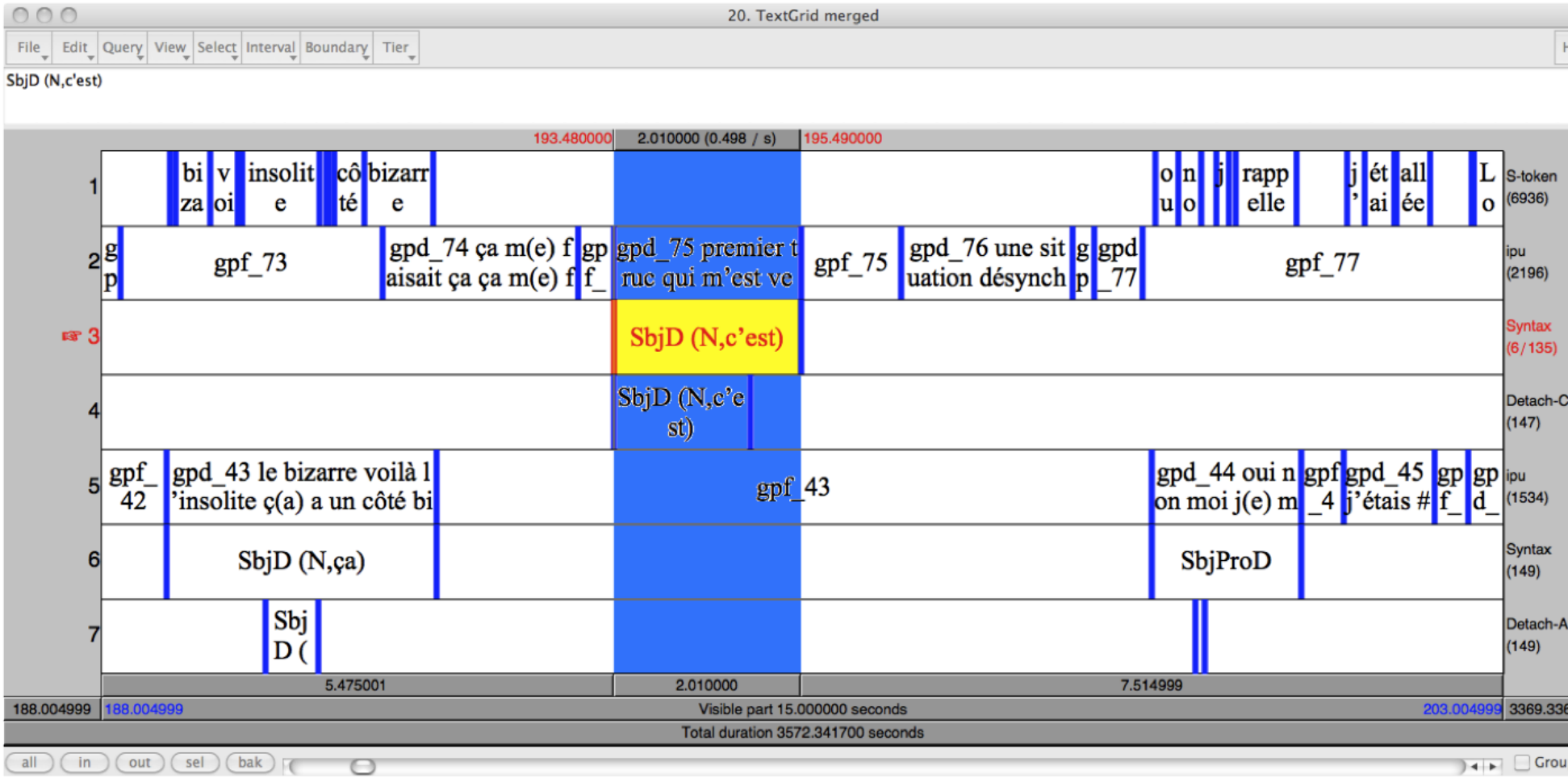
Trees


Tree structure
Welcome to Opera
AP.* description
Tree structure
Texte Caractérisé
NP.* description
Tree structure
Tree structure
Locale
localhost/Users/philippeblache/Dropbox/OTIM/CID-Syntax/tree/AB_tree.xml
Rechercher avec Google
SENT :
VN
PP:MOD
Wm
VN
PP:MOD
Wm
VN
NP:OBJ
PP:MOD
Pd-ns-- c'
Vmii3s-- était
Sp- pendant
NP
Da-mp--d- les
Ncfp-- vacances
PP
Spd de
NP
Np-s-- Noël
Pp3msn- on
Pp3n-d- y
Vmmp2s-- va
Sp- avec
NP
Da-mp--i- des
Ncmp-- copains
Pp3msn- on
Vaii3s-- avait
Vmps-sm- pris
Da-ms--d- le
Ncms-- ferry
Sp- en
NP
Np-s-- Normandie
Cs puisque
VN
Pp1-sn- j'
Vmii2s-- avais
Dk1ms---- un

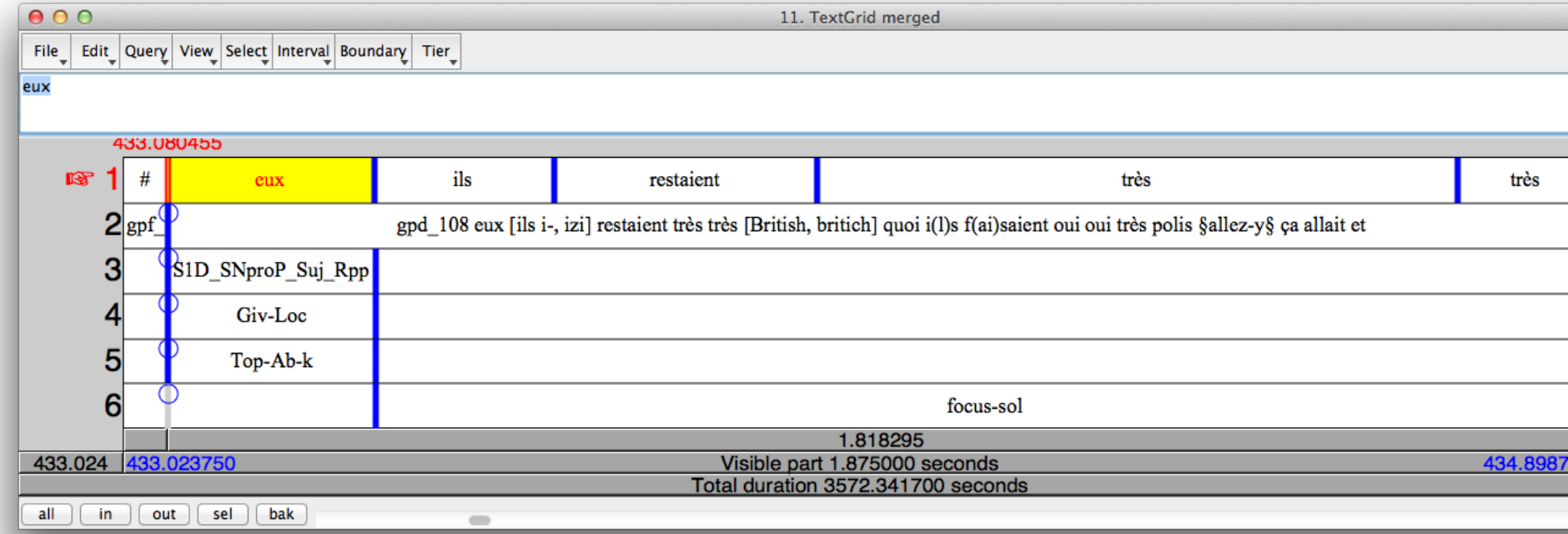
Detachment: annotations

- Dislocation: "*Chocolate, I hate*"
- Cleft: "*It is John who married Ann*"
- Pseudo-cleft: "*What he wanted to do was to travel*"
- Binary constructions: "*Being happy, it is not always*"
- Features
  - Detachment type: `D`, `CV`, `PSCV`, `B`
  - Detached category: `NP`, `NPrel`, `NPproP`, `NPproD`, `NPproQ`, `PP`, `AP`, `AdvP`, `VP`, `S`
  - Function: `Subj`, `Odir`, `Oind`, `Loc`, `Adj`
  - Resumptive element: `Rxx` (xx : type of the res. element)

Detachment


20. TextGrid merged
File Edit Query View Select Interval Boundary Tier
SbjD (N,c'est)
193.480000
2.010000 (0.498 / s)
195.490000
1
bi za
v oi
insolit e
cô té
bizarr e
o u
n o
j
rapp elle
j '
ét ai
all ée
L o
S-token (6936)
2
g p
gpf_73
gpd_74 ça m(e) f aisait ça ça m(e) f
gp f_
gpd_75 premier t ruc qui m'est ve
gpf_75
gpd_76 une sit uation désynch
g p
gpd _77
gpf_77
ipu (2196)
3
SbjD (N,c'est)
Syntax (6/135)
4
SbjD (N,c'e st)
Detach-C (147)
5
gpf_ 42
gpd_43 le bizarre voilà l 'insolite ç(a) a un côté bi
gpf_43
gpd_44 oui n on moi j(e) m
gpf _4
gpd_45 j'étais #
gp f_
gp d_
ipu (1534)
6
SbjD (N,ça)
SbjProD
Syntax (149)
7
Sbj D (
Detach-A (149)
5.475001
2.010000
7.514999
188.004999
188.004999
Visible part 15.000000 seconds
203.004999
3369.336
Total duration 3572.341700 seconds
all in out sel bak
Grou

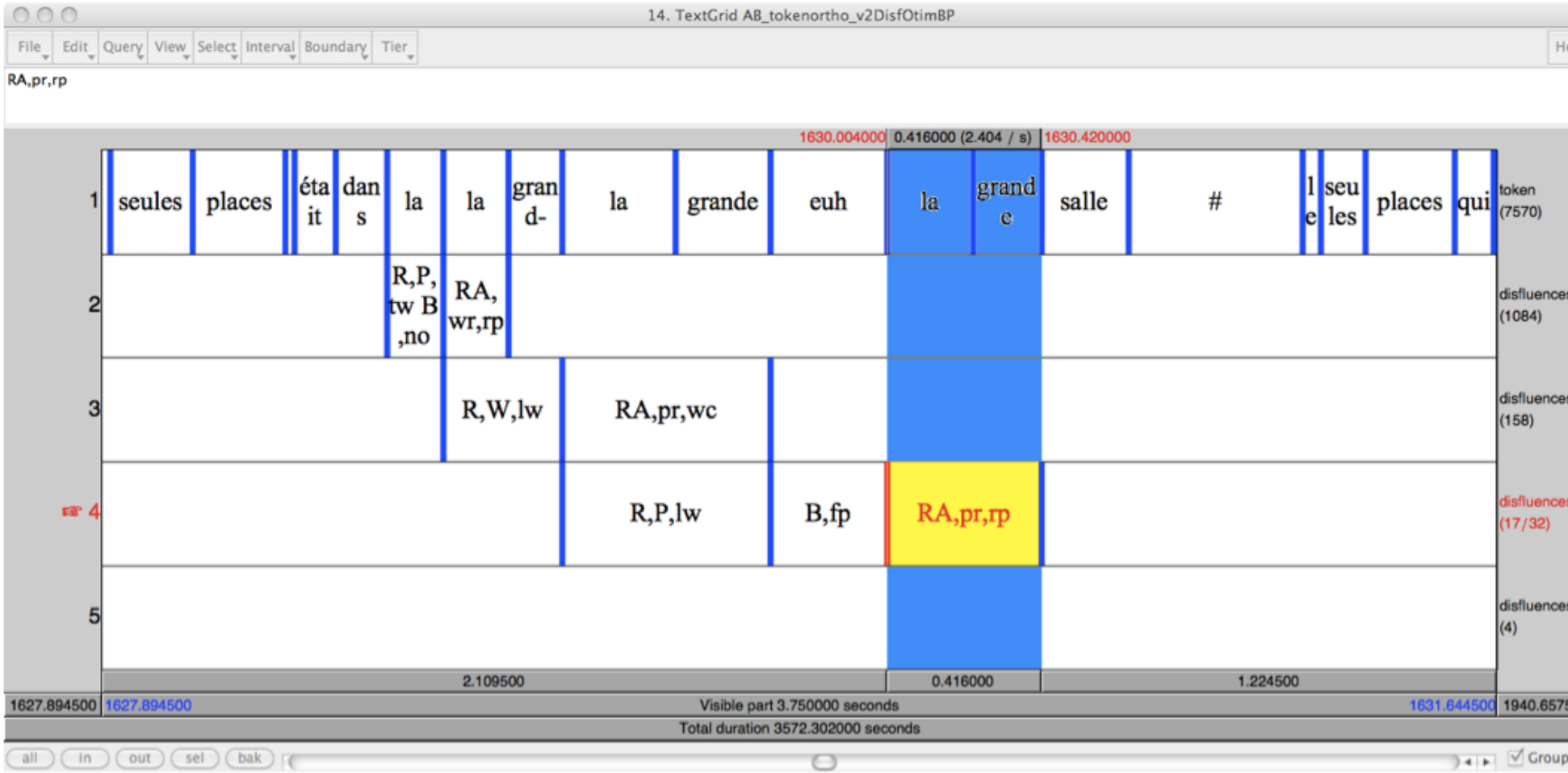
Detachment (2)



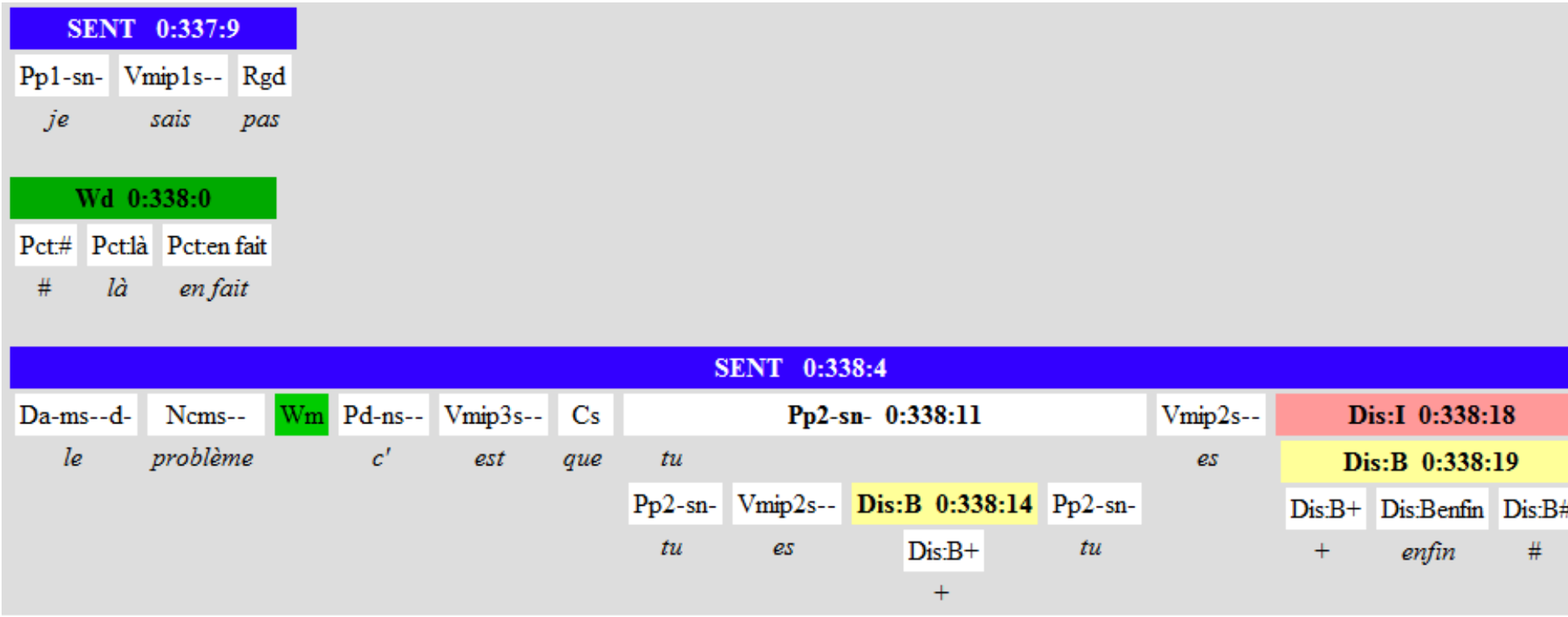
Disfluencies

Reparandum	
Reparandum Type	`R` *Temporary interruption* `I` *Definitive Interruption*
Reparandum category	`W` *Word reparandum* `P` *Phrase reparandum*
Lexical type	`tw` *Tool word* `lw` *Lexical word*
Break type B	
	`no` *no interval* `sp` *silent pause (> 200ms)* `fp` *filled pause* `dc` *discursive connector* `ps` *parenthetical statement* `rt` *truncation repetition*
Reparans RA	
Reparans position type	`nr` *no restart* `wr` *word restart* `dr` *determinant restart* `pr` *phrase restart* `or` *other restart*
Reparans type	`co` *continuing the item* `wc` *repairing without change* `rp` *Repairing through repeating* `rc` *repair with change in the truncated word* `rm` *repair with multiple change*

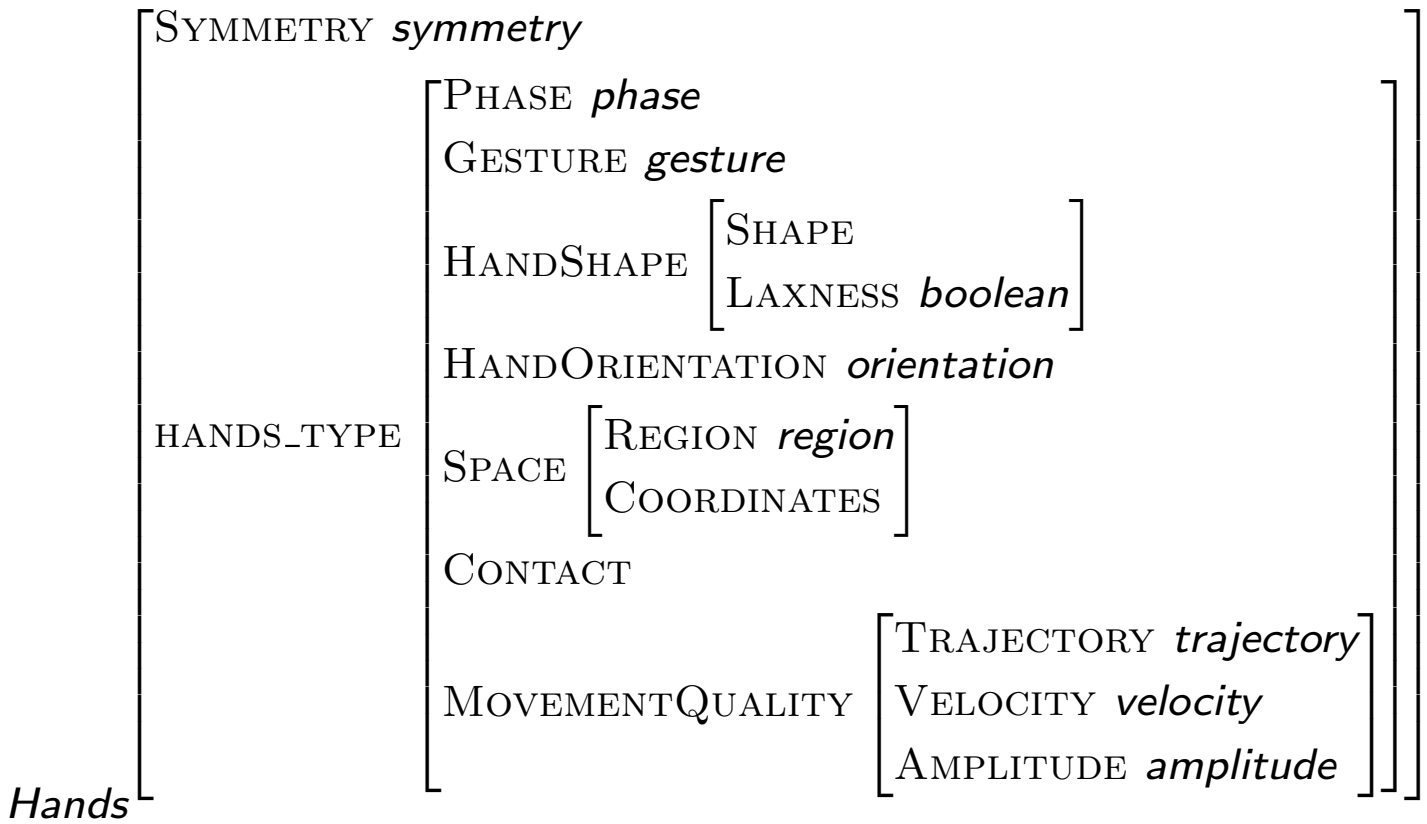
Disfluencies



Disfluencies (2)



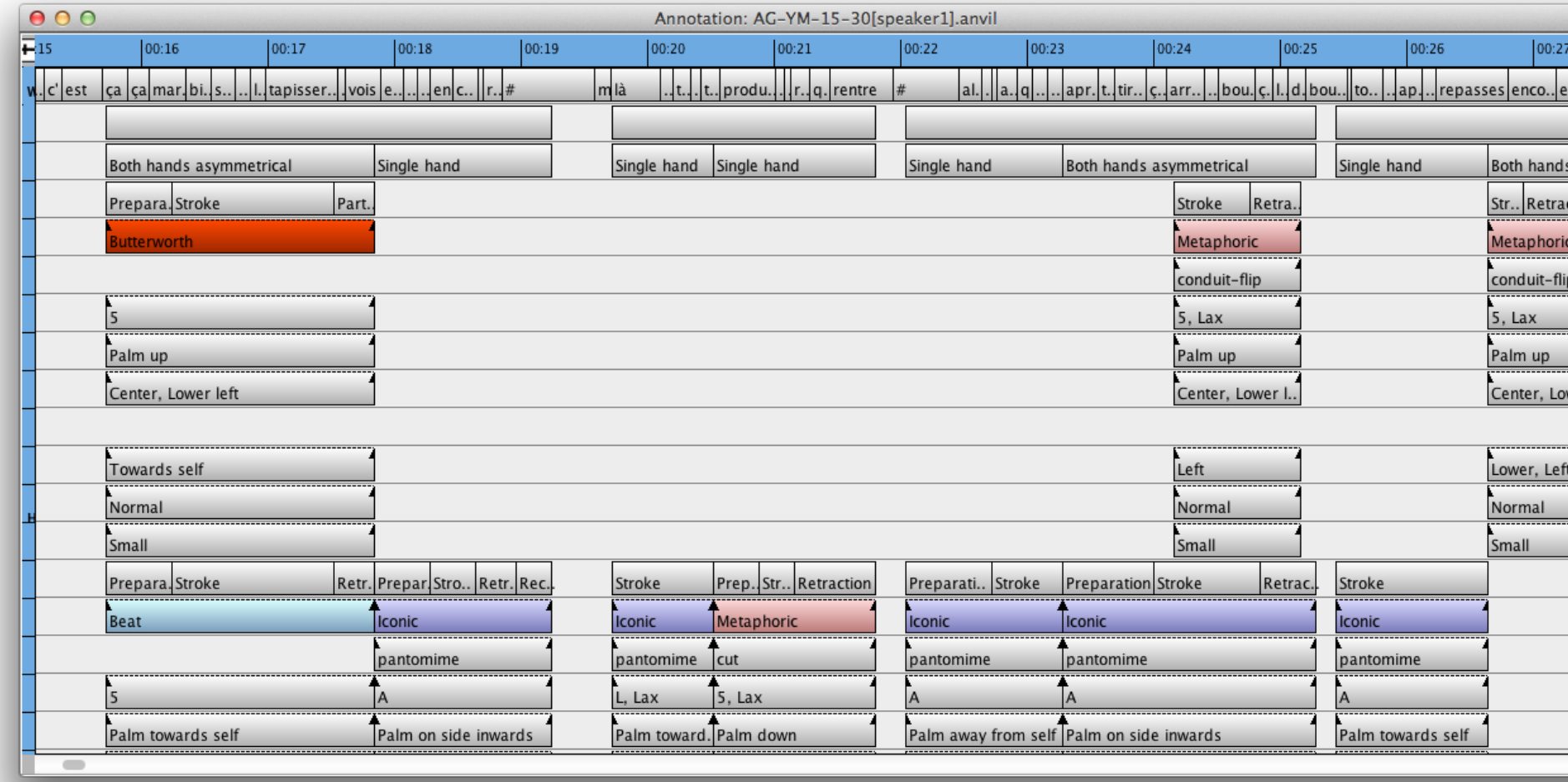
Gestures: hands



Gestures: hands

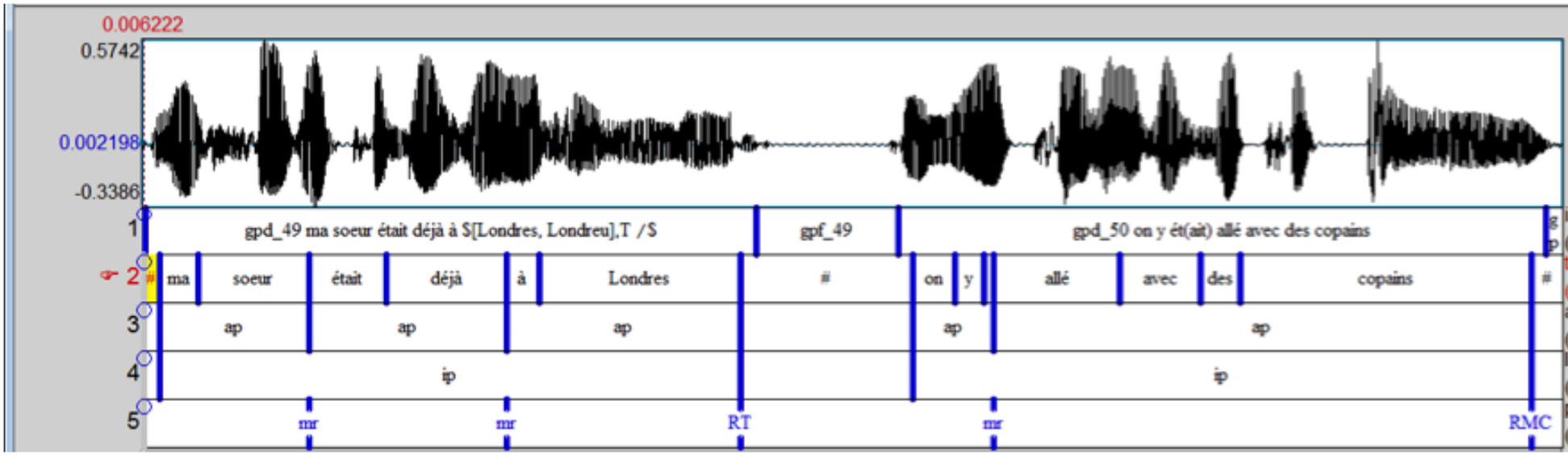
symmetry:	{Both hands symmetrical, Both hands asymmetrical, ...}
phase:	{Preparation, Stroke, Hold, Retraction, ...}
gesture:	{Adaptor, Iconic, Metaphoric, Deictic, Emblem, ...}
orientation:	{ Palm up, Palm down, Palm towards self, Palm away from self,
region:	{ Center center, Center, Periphery, Extreme periphery}
coordinates:	{ Right, Left, Upper, Lower, Upper right, ...}
contact:	{ Forehead, Hair, Cheek, Chin, Eyes, Eyebrow, ...}
trajectory:	{ Upper, Lower, Right, Left, Upper right, lower right, ...}
velocity:	{ Normal, Fast, Slow}
amplitude:	{ Small, Medium, large}

Gestures: hands



Stand-off hierarchical encoding

- When editing, no distinction features vs. hierarchy
- Prosody: intonation, contour are features vs. AP $\in$ IP



TFS representation: XML scheme

$$\begin{bmatrix} \text{LABEL IP} \\ \text{CONSTITUENTS list}(ap) \\ \text{CONTOUR} \begin{bmatrix} \text{DIRECTION } \textit{string} \\ \text{POSITION } \textit{string} \\ \text{FUNCTION } \textit{string} \end{bmatrix} \end{bmatrix}$$

```
<xs:complexType name="IntonationalPhrase">
  <xs:complexContent>
    <xs:extension base="ProsodicPhrase">
    <xs:sequence>
      <xs:element name="constituents">
        <xs:complexType>
          <xs:sequence>
            <xs:element name="accentual\_phrase" typ
          </xs:sequence>
        </xs:complexType>
      </xs:element>
      <xs:element name="contour" type="Contour"/>
    </xs:sequence>
    </xs:extension>
  </xs:complexContent>
</xs:complexType>
```

XML representation

```
  intervals [2]:
  xmin = 0.78
  xmax = 1.7559754641684542
  text = "ip"
...
  intervals [5]:
  xmin = 2.6703535937364578
  xmax = 3.329971301020408
  text = "ip"
...
  class = "TextTier"
  name = "at_ctr"
  xmin = 0
  xmax = 3573.6
  points: size = 2118
  points [1]:
  time = 1.7559754641684542
  mark = "RT"
...
  points [3]:
  time = 3.329971301020408
  mark = "F"
```

```
<IntonationalPhrase index=0>
   <localisation start=0.78 end=1.7559 />
   <contour type=RT time=1.7559  />
</IntonationalPhrase>
   ...
<IntonationalPhrase index=5>
   <localisation start=2.6703 end=3.3299 /
   <contour type=F time=3.3299 />
</IntonationalPhrase>
```

Distribution: the CID at SLDR