Séminaire CARTEL

Analyse de commentaires sportifs en direct

Gilles Boyé, Anna Kupść, Catherine Mathon

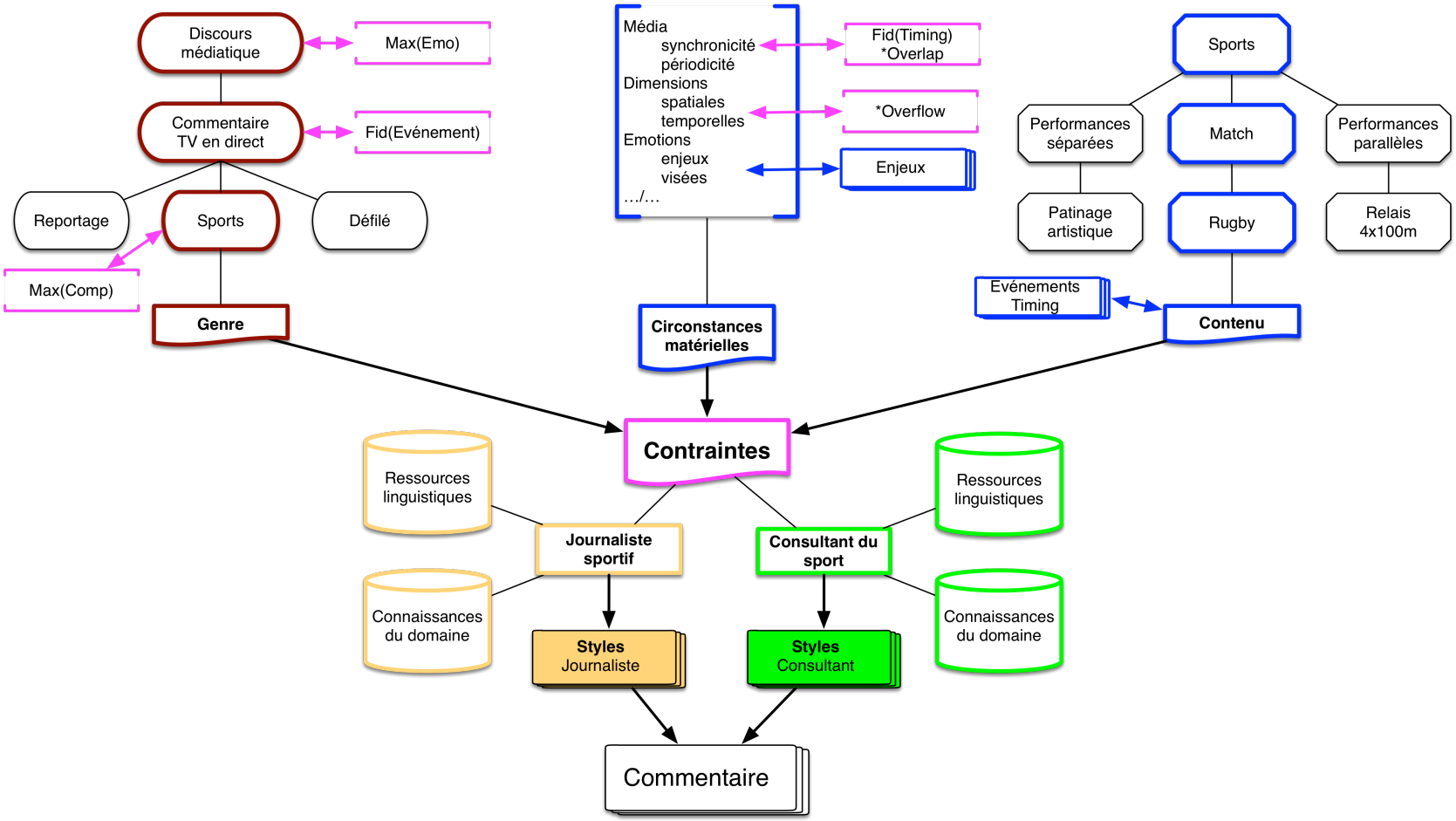
Université Bordeaux Montaigne
CLLE-ERSSàB

2 Novembre 2015

Motivation

- Interactions entre contexte et production linguistique
  - commentaire sportif $\Rightarrow$ contenu imposé
  - aspect visuel $\Rightarrow$ synchronisation contenu–vidéo
- Interactions entre participants
  - qui dit quoi quand comment?
  - et pourquoi?
- Beaucoup d'annotations nécessaires
  - contenu du match/vidéo et commentaire
  - analyse morphologique, syntaxique, prosodique
- **Un** match entièrement annoté à la main
  - ⋆ comment automatiser les annotations pour obtenir d'autres données?

Discours sous contraintes

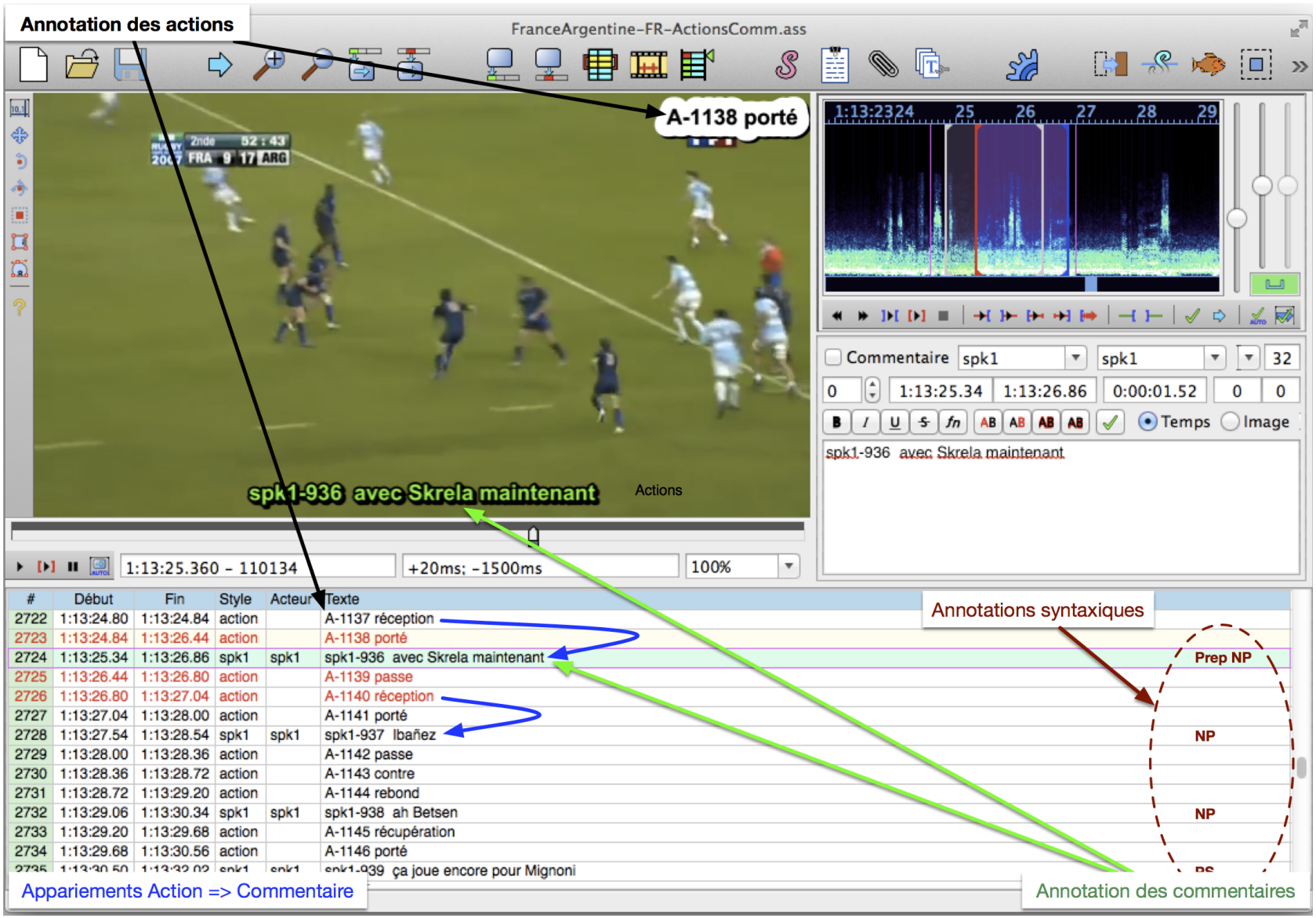


Corpus

Match de rugby France-Argentine, coupe du monde 2007

- commentaire sportif télévisé
- document : audio et vidéo, 1h48
- 3 locuteurs:
  - Thierry Gilardi, journaliste (39%)
  - Thierry Lacroix, spécialiste (24%)
  - Fabrice Landreau, terrain (1.8%)

Annotations : vue globale



#	Début	Fin	Style	Acteur	Texte
2722	1:13:24.80	1:13:24.84	action		A-1137 réception
2723	1:13:24.84	1:13:26.44	action		A-1138 porté
2724	1:13:25.34	1:13:26.86	spk1	spk1	spk1-936 avec Skrela maintenant
2725	1:13:26.44	1:13:26.80	action		A-1139 passe
2726	1:13:26.80	1:13:27.04	action		A-1140 réception
2727	1:13:27.04	1:13:28.00	action		A-1141 porté
2728	1:13:27.54	1:13:28.54	spk1	spk1	spk1-937 Ibañez
2729	1:13:28.00	1:13:28.36	action		A-1142 passe
2730	1:13:28.36	1:13:28.72	action		A-1143 contre
2731	1:13:28.72	1:13:29.20	action		A-1144 rebond
2732	1:13:29.06	1:13:30.34	spk1	spk1	spk1-938 ah Betsen
2733	1:13:29.20	1:13:29.68	action		A-1145 récupération
2734	1:13:29.68	1:13:30.56	action		A-1146 porté
2735	1:13:30.50	1:13:32.02	spk1	spk1	spk1-939 ça joue encore pour Mignoni

Annotations liées au discours

- Transcription orthographique
- Identification des locuteurs:
  - **spk1**: journaliste
  - **spk2**: spécialiste
  - **spk3**: sur le terrain
- Types de commentaire:
  - **play-by-play** : description synchrone des actions
  - **color-commentary** : informations complémentaires

Découpage du discours

Manuellement avec Transcriber

- découpage en tours de parole (groupe de souffle)
- transcription orthographique
  - avec les disfluences
  - sans ponctuation

Annotations liées au sport

- Annotation des actions du match sur la vidéo
  - sous-titres : Aegisub
- Actions: une liste prédéfinie
- Appariement actions/commentaires avec alignement sur la vidéo:
  - décalage actions/commentaires
  - durée/nombre des commentaires vs. actions

0:14:05.90	A-035 maul	Spk1-035, spk1-036, spk1-037	
0:14:01.92			spk1-033 proprement
0:14:03.28			spk1-034 du deuxième
0:14:04.48			spk1-035 des grosses
0:14:05.92			spk1-036 poussées défensives
0:14:06.26	A-036 écroulement		
0:14:07.70	A-037 récupération	Spk1-038	
0:14:07.20			spk1-037 du pack français
0:14:08.38	A-038 plaquage		
0:14:09.20			spk1-038 Ledesma encore
0:14:14.10	A-039 regroupement		
0:14:15.04			spk1-039 et vous entendez Monsieur Spreadbury qui fait de la prévention

Annotations liées à la syntaxe

- Annotation manuelle de chaque tour
- Au niveau global, le type de la structure syntaxique
  - distinction tour principal/subordonné
  - tours principaux :
    - PS (phrase simple), PC (phrase complexe)
    - X (nom propre), XQui (nom propre+relative), ...
    - PrepX, PrepXQui, ...
    - GN, GNQui, GNGPrep, ...
  - tours subordonnés :
    - ÉtiquettePrincipale+bis

Analyse des données

Actions et rythme du jeu: qui parle **quand**?

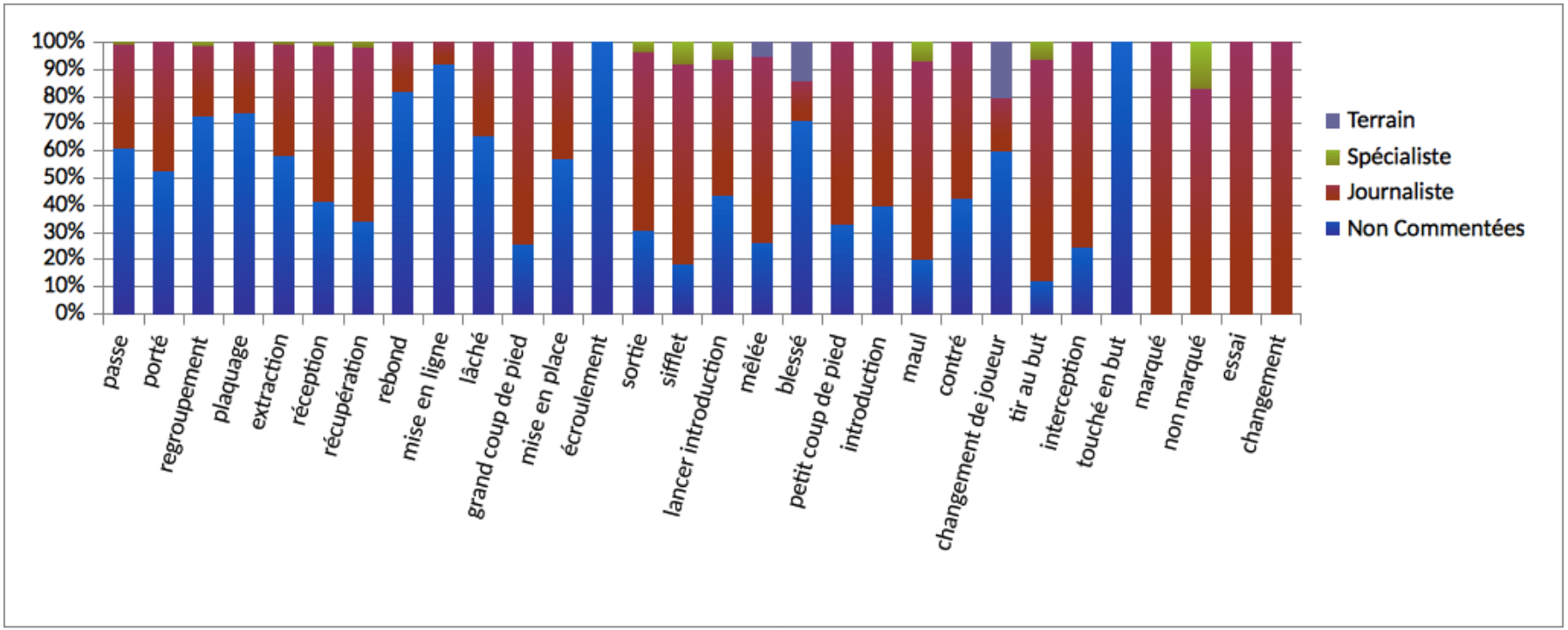
- distribution de commentaires par action
- rythme du jeu vs. commentaires:

plage 1 rythme lent: 1 à 3 actions/5 sec
plage 2 rythme moyen: 4 à 6 actions/5sec
plage 3 rythme rapide: 7 à 10 actions/5sec

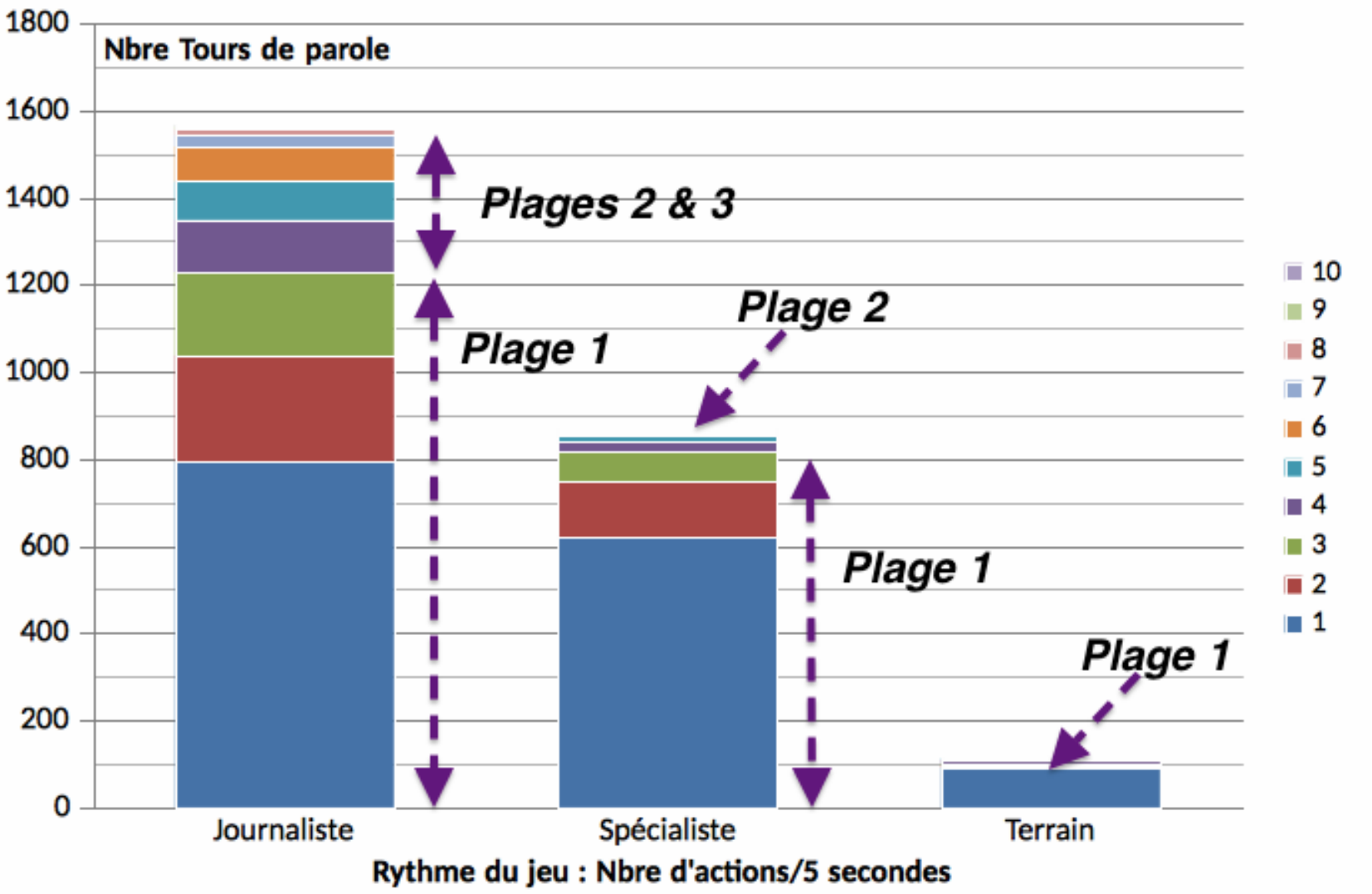
Texte: qui parle **comment** quand?

- nombre de mots par tours/locuteur/action
- distributions de catégories: noms et verbes
- distribution des structures syntaxiques

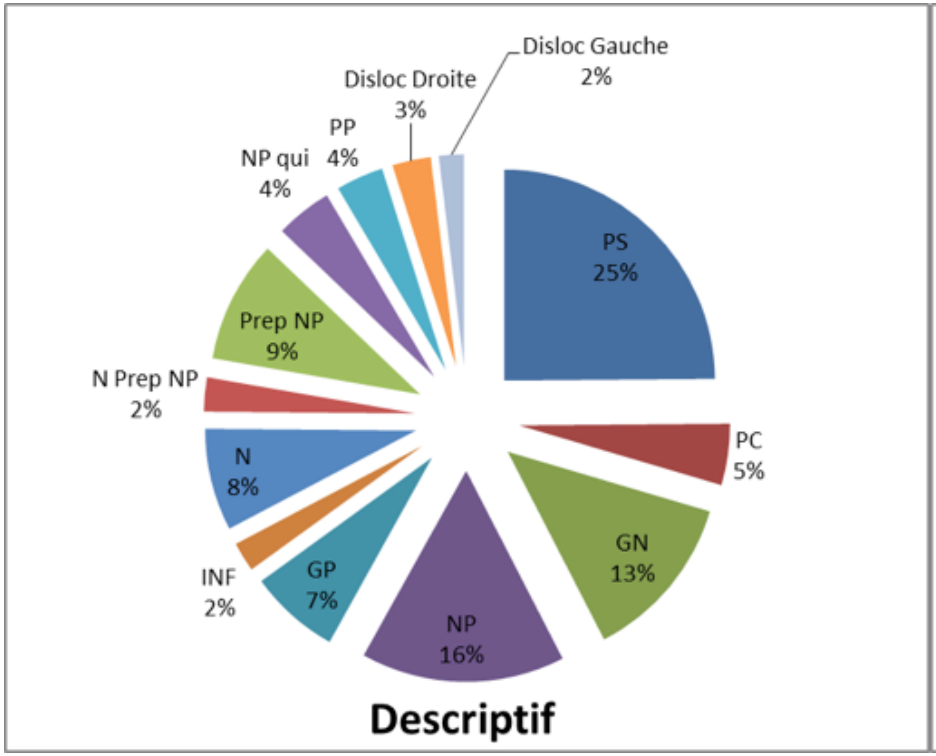
Commentaires vs. Actions: Qui parle quand?

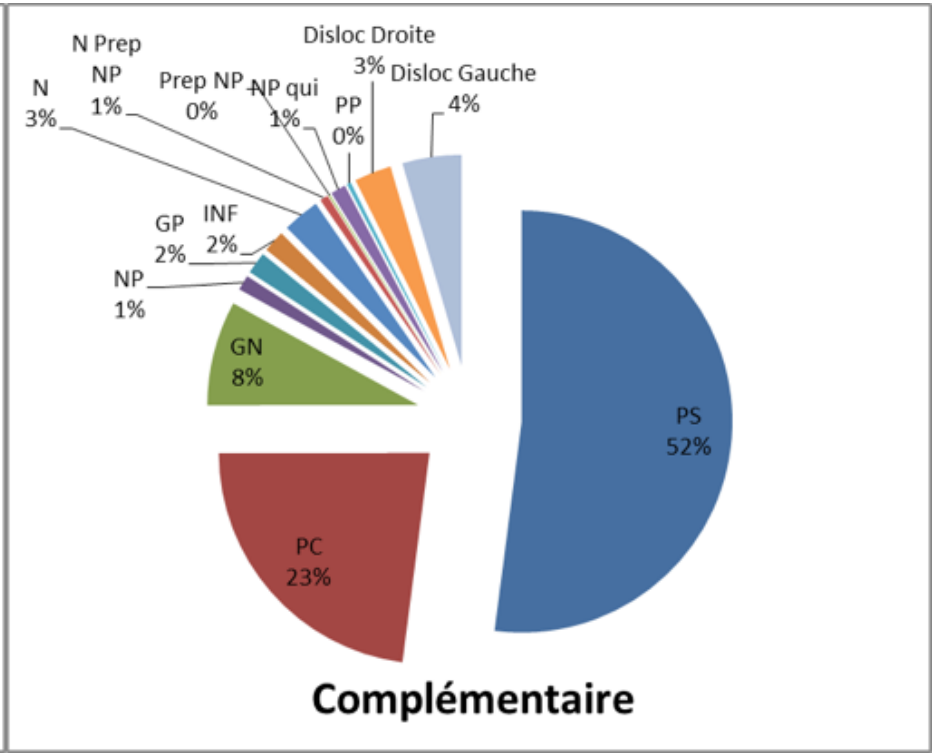


Rythme du jeu vs. Commentaires: Qui parle quand?



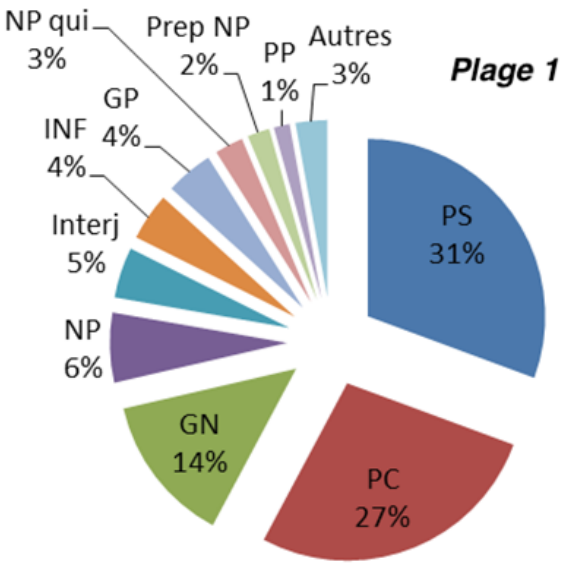
Syntaxe vs. Type de commentaire: Comment commenter quand?

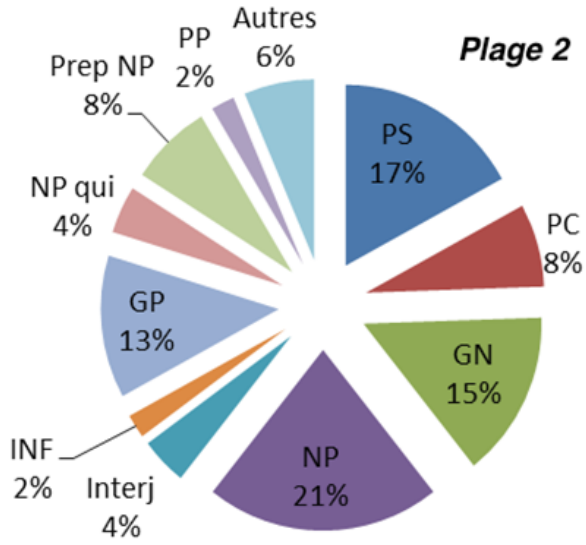


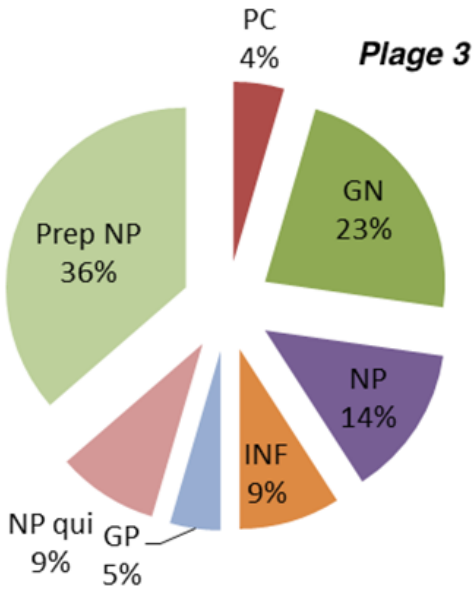


play-by-play **color-commentary**

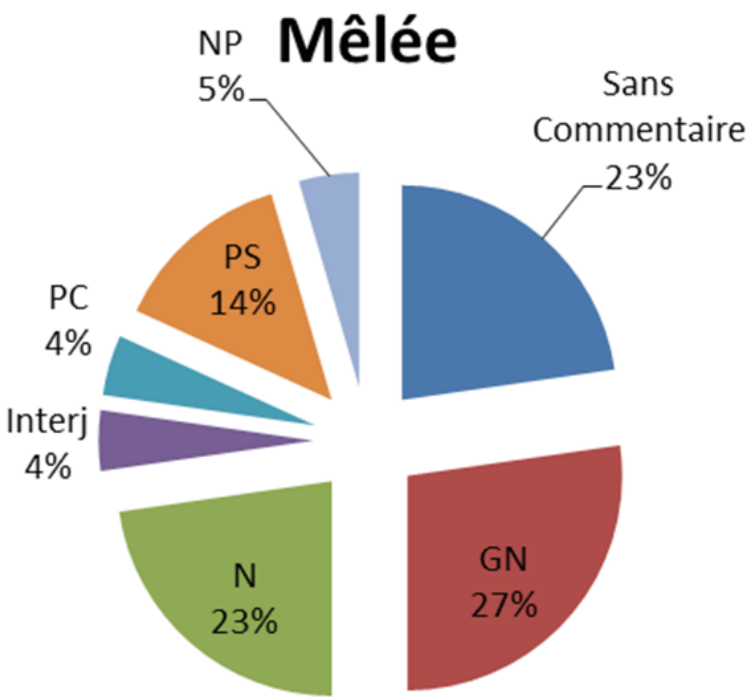
Syntaxe vs. Rythme

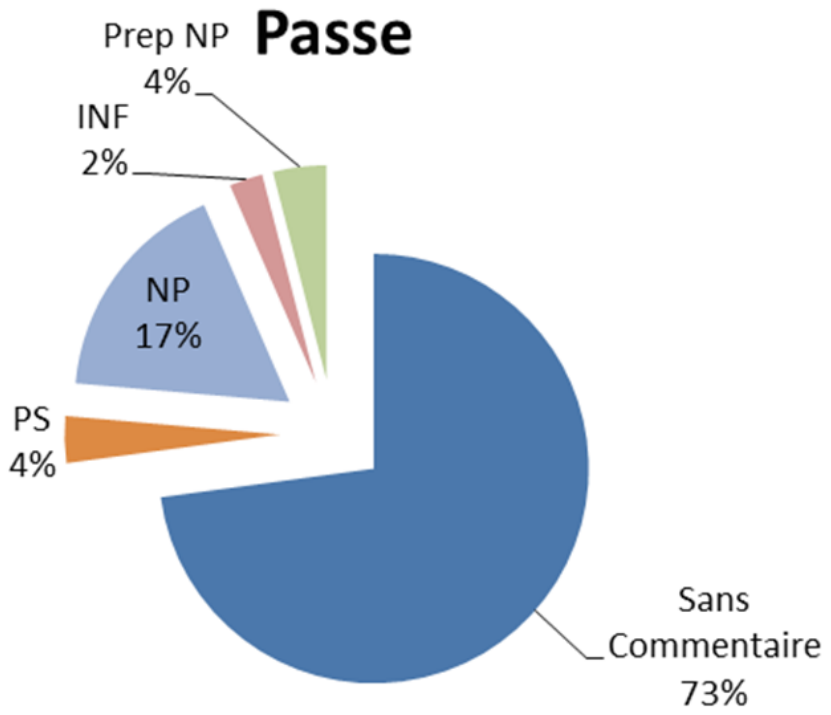






Syntaxe vs. Actions





Comment élargir le corpus ?

Analyse du sport

- annotation par les étudiants de STAPS
  - actions
  - correspondances actions/commentaires

Analyse linguistique :

- analyse syntaxique de l'oral
  - problèmes avec les disfluences
  - problèmes avec l'identification des phrases/énoncés
- analyse prosodique
  - repérage des variations de F0
  - repérage des syllabes
  - correspondances variations de F0/syllabes
- analyse lexicale
  - repérage des entités nommées (et termes référents)

Analyse syntaxique de l'oral

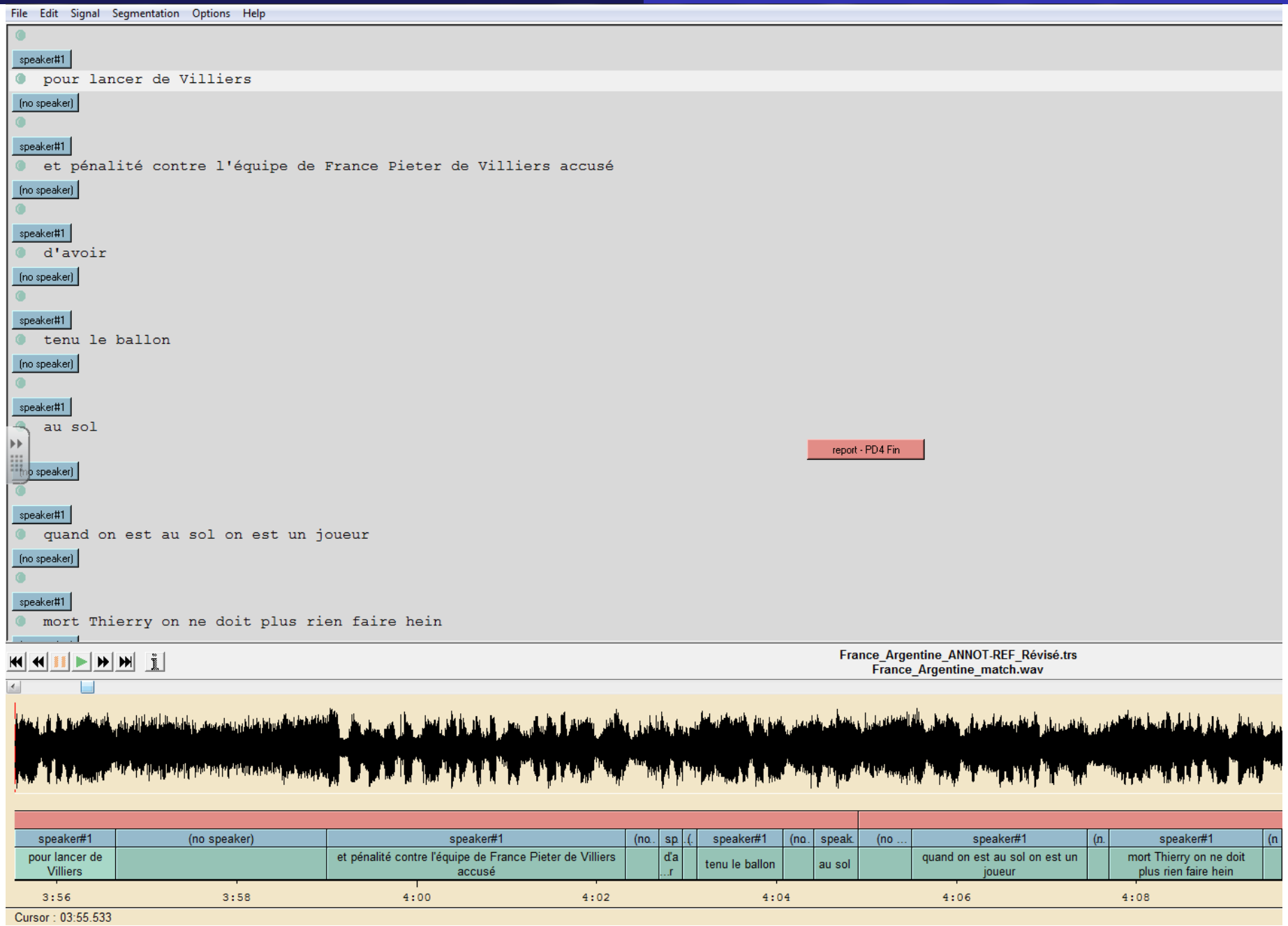
- Fillers:
  - allez de suite dans le camp **euh**
  - argentin
- Inachèvements:
  - Cédric Heymans en couverture
  - il va devoir **XXX**
- Auto-corrections:
  - sur **le le la le la** la cheville **de**
  - de Heymans et le déséquilibre

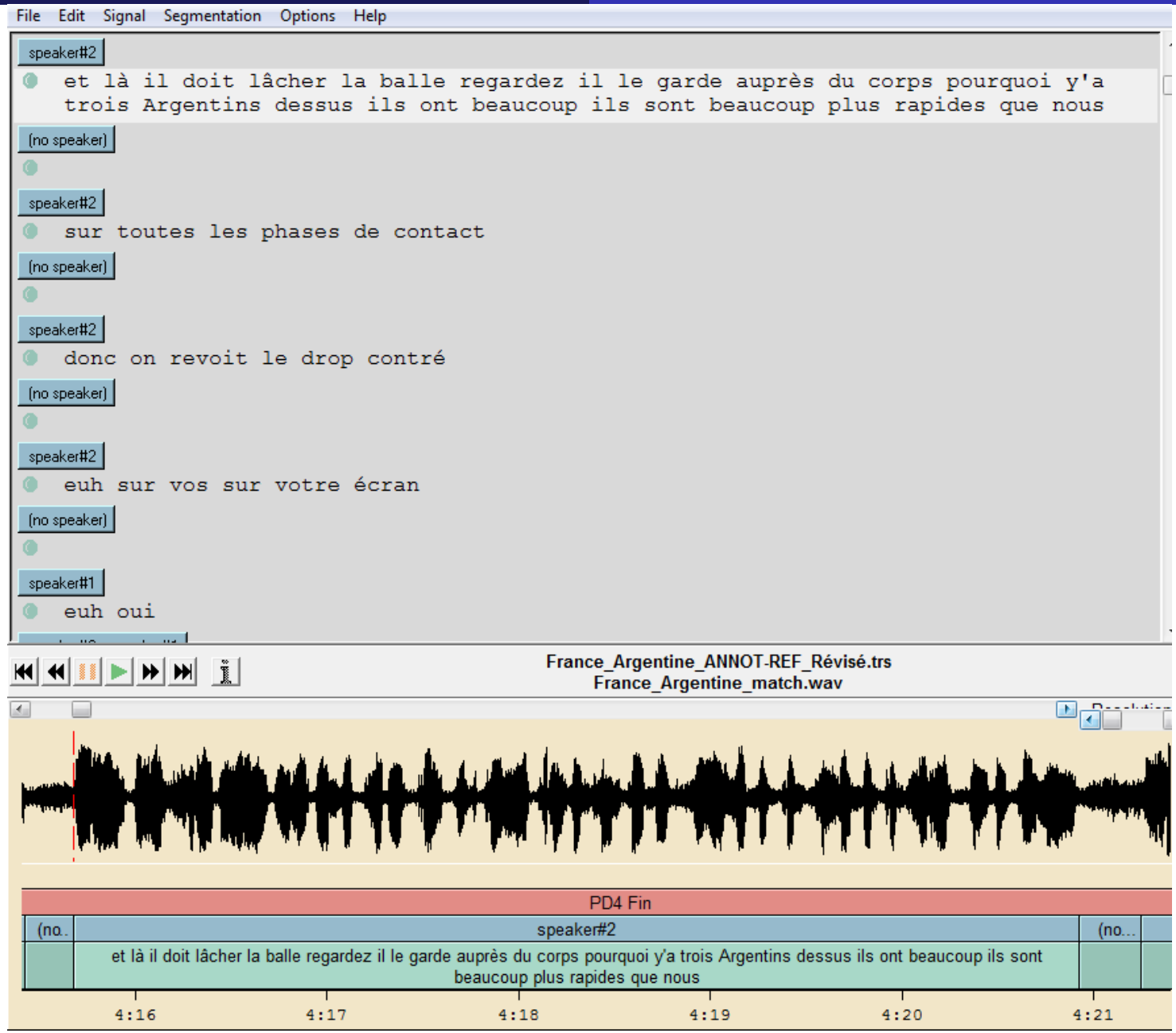
Des problèmes pour l'analyse syntaxique automatique qui se situe généralement au niveau phrase (p.e. Talismane)

Identification des phrases

- Phrases sur plusieurs tours
  - Pieter de Villiers accusé
  - d'avoir
  - tenu le ballon
  - au sol
- Tour avec plusieurs phrases
  - et là, il doit lâcher la balle. regardez, il le garde auprès du corps, pourquoi ? y'a trois argentins dessus, ils ont beaucoup ils sont beaucoup plus rapides que nous

Comment retrouver le niveau phrase au sein des tours ?


File Edit Signal Segmentation Options Help
speaker#1
pour lancer de Villiers
(no speaker)
speaker#1
et pénalité contre l'équipe de France Pieter de Villiers accusé
(no speaker)
speaker#1
d'avoir
(no speaker)
speaker#1
tenu le ballon
(no speaker)
speaker#1
au sol
report - PD4 Fin
speaker#1
quand on est au sol on est un joueur
(no speaker)
speaker#1
mort Thierry on ne doit plus rien faire hein
France_Argentine_ANNOT-REF_Révisé.trs
France_Argentine_match.wav
speaker#1 | (no speaker) | speaker#1 | speaker#1 | speaker#1 | speaker#1
pour lancer de Villiers | et pénalité contre l'équipe de France Pieter de Villiers accusé | tenu le ballon | au sol | quand on est au sol on est un joueur | mort Thierry on ne doit plus rien faire hein
3:56 3:58 4:00 4:02 4:04 4:06 4:08
Cursor : 03:55.533


File Edit Signal Segmentation Options Help
speaker#2
et là il doit lâcher la balle regardez il le garde auprès du corps pourquoi y'a trois Argentins dessus ils ont beaucoup ils sont beaucoup plus rapides que nous
(no speaker)
speaker#2
sur toutes les phases de contact
(no speaker)
speaker#2
donc on revoit le drop contré
(no speaker)
speaker#2
euh sur vos sur votre écran
(no speaker)
speaker#1
euh oui
France_Argentine_ANNOT-REF_Révisé.trs
France_Argentine_match.wav
PD4 Fin
(no..
speaker#2
(no...
et là il doit lâcher la balle regardez il le garde auprès du corps pourquoi y'a trois Argentins dessus ils ont beaucoup ils sont beaucoup plus rapides que nous
4:16
4:17
4:18
4:19
4:20
4:21

Ponctuation & Oral

- L'oral ne se réduit pas au texte
  - prosodie
  - rythme
- La ponctuation de l'écrit
  - donne des indications sur les groupes syntaxiques
  - oriente les interprétations
  - mais il faut un système linguistiquement motivé ($\neq$ écrit normé)
- $\star$ Ponctuation de l'oral

  $\Rightarrow$ Hidden, Kupść & Mathon, en préparation
  - durée des pauses ne donne pas un indicateur fiable
  - pour ce corpus $\Rightarrow$ frontières syntaxiques fortes

Annotations syntaxiques

- Des annotations syntaxiques pour :
  - caractériser les structures linguistiques
  - pas particulièrement au niveau **phrase**
  - mais une analyse robuste
- Solution envisagée :
  - identification des entités nommées (listes)
  - chunking (TreeTagger)
  - apprentissage supervisé des annotations syntaxiques

Apprentissage des annotations à partir des chunks

- Annotations principalement basées sur des chunks
  - plusieurs niveaux d'étiquetage

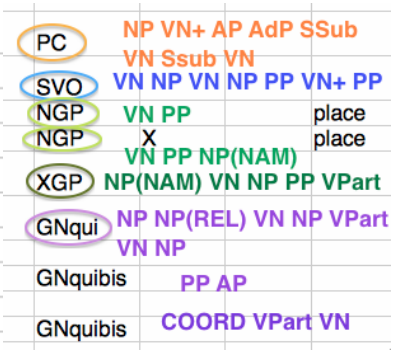
spk2	attention ça peut être dangereux hein des fois quand on tombe comme ça	PC	
spk2	y'a une clef aussi c'est les demis de mêlée on vous en parlera tout à l'heure	SVO	
spk2	place au jeu	NGP	
spk2	place à la touche Ledesma	NGP	X
spk1	Mario Ledesma super Mario pour cette remise en jeu	XGP	
spk1	ah les Argentins qui ont bien démarré en chipant un premier ballon	GN qui	
spk1	à l'alignement français	GN quibis	
spk1	et en s'emparant	GN quibis	

  - apprentissage supervisé $\Rightarrow$ déploiement sur corpus
  - ⋆ que faire avec la notion de tours subordonnés ?
    - annotation manuelle

Apprentissage des annotations à partir des chunks

- Annotations principalement basées sur des chunks
  - plusieurs niveaux d'étiquetage

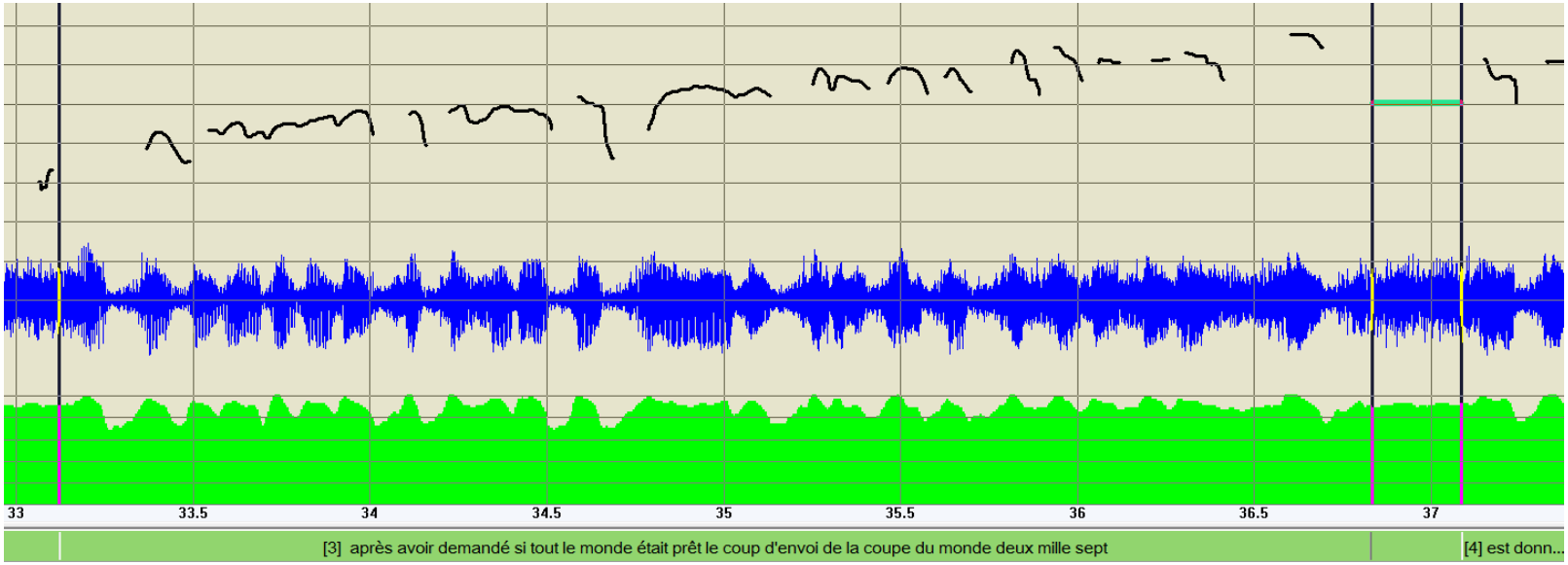
spk2	attention ça peut être dangereux hein des fois quand on tombe comme ça
spk2	y'a une clef aussi c'est les demis de mêlée on vous en parlera tout à l'heure
spk2	place au jeu
spk2	place à la touche Ledesma
spk1	Mario Ledesma super Mario pour cette remise en jeu
spk1	ah les Argentins qui ont bien démarré en chipant un premier ballon
spk1	à l'alignement français
spk1	et en s'emparant



  - apprentissage supervisé ⇒ déploiement sur corpus
  - ⋆ que faire avec la notion de tours subordonnés ?
    - annotation manuelle

Annotations prosodiques

- Découpage en syllabe avec les variations de F0
  - exportation des F0 avec WinPitch



  - découpage des syllabes potentielles
    - signal (apprentissage supervisé)
    - transcription (lexique phonétisé)
  - alignement semi-automatique syllabes/transcription

Annotations prosodiques

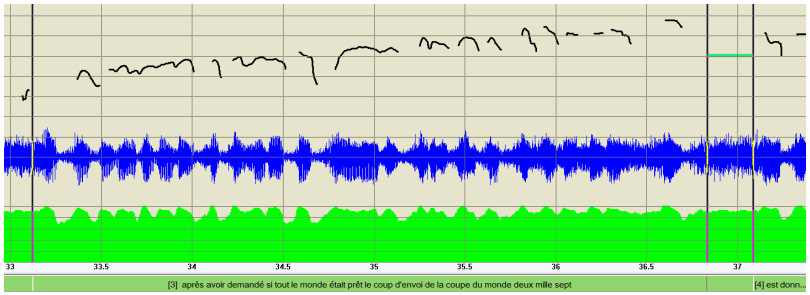
- Découpage en syllabe avec les variations de F0
  - exportation des F0 avec WinPitch

Time [s]	Type	FZero [Hz]	Intensity [dB]
33.140 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	168	46
33.160 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	182	47
33.180 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	183	47
33.200 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	183	49
33.220 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	178	50
33.240 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	180	48
33.260 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	178	44
33.280 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	162	35
33.300 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	158	37
33.320 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	154	37
33.340 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	144	39
33.360 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	0	43
33.380 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	0	48
33.400 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	106	49
33.420 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	106	47
33.440 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	0	45
33.460 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	0	39
33.480 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	0	37
33.500 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	187	43
33.520 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	187	48
33.540 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	191	47
33.560 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	191	45
33.580 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	191	42
33.600 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	195	45

  - découpage des syllabes potentielles
    - signal (apprentissage supervisé)
    - transcription (lexique phonétisé)
  - alignement semi-automatique syllabes/transcription

Annotations prosodiques

- Découpage en syllabe avec les variations de F0
  - exportation des F0 avec WinPitch



Time [s]	Type	FZero [Hz]	Intensity [dB]
33.140 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	168	46
33.160 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	182	47
33.180 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	183	47
33.200 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	183	49
33.220 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	178	50
33.240 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	180	48
33.260 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	178	44
33.280 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	162	35
33.300 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	158	37
33.320 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	154	37
33.340 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	144	39
33.360 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	0	43
33.380 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	0	48
33.400 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	106	49
33.420 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	106	47
33.440 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	0	45
33.460 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	0	39
33.480 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	0	37
33.500 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	187	43
33.520 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	187	48
33.540 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	191	47
33.560 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	191	45
33.580 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	191	42
33.600 s	après avoir demandé si tout le monde était prêt le coup d'envoi de la coupe du monde deux mille sept	195	45

  - découpage des syllabes potentielles
    - signal (apprentissage supervisé)
    - transcription (lexique phonétisé)
  - alignement semi-automatique syllabes/transcription
- Analyse des contours prosodiques
  - groupes syntaxiques
  - ponctuation orale
  - émotions

Analyse des référents

- Repérage des entités nommées
  - poursuivre le travail de Augendre & Mathon (2012)
- Variation lexicale des référents
  - David Skrela vs Skrela (Harinordoquy vs *Imanol Harinordoquy)
  - l'équipe de France vs les bleus
- Chaînes de références dans le discours

Perspectives

- Agrandir le corpus pour être représentatif
  - en taille (1 match $\Rightarrow$ échantillons de sports variés)
  - en niveaux d'annotation (sport, syntaxe, lexique, prosodie, rythme)
- Analyser la syntaxe de l'oral
  - analyse des structures textuelles
  - intégration des données prosodiques (disfluence, intonation, rythme)
- Automatiser pour formaliser
  - la macro-syntaxe, par exemple, donne des structures textuelles
  - le niveau d'explicitation n'est pas suffisant pour entraîner un tagger