

EIIDA'S SECOND LIFE : PRÉPARER UN CORPUS MULTILINGUE ORAL À DES TRANSPOSITIONS DIDACTIQUES FUTURES

**Achille Falaise, Adrien Méli, Alice Henderson & Nicolas
Ballier**

LLF, Paris 7; CLILLAC-ARP Uni de Paris; LIDILEM, UGA; CLILLAC-ARP Uni de Paris



RADELAS, UGA, 20-21 février

Plan

- Le corpus EIIDA
- Questions de recherche: Pourquoi une deuxième vie?
- Aspects techniques: Comment?
- Aspects didactiques
- Conclusion

Corpus EIIDA

Projet 2012-17

Shirley Carter-Thomas & Jeanne-Marie Debaisieux

LATTICE (CNRS/ ENS
/Paris 3)

300k mots,
approx. 20h
d'enregistrements

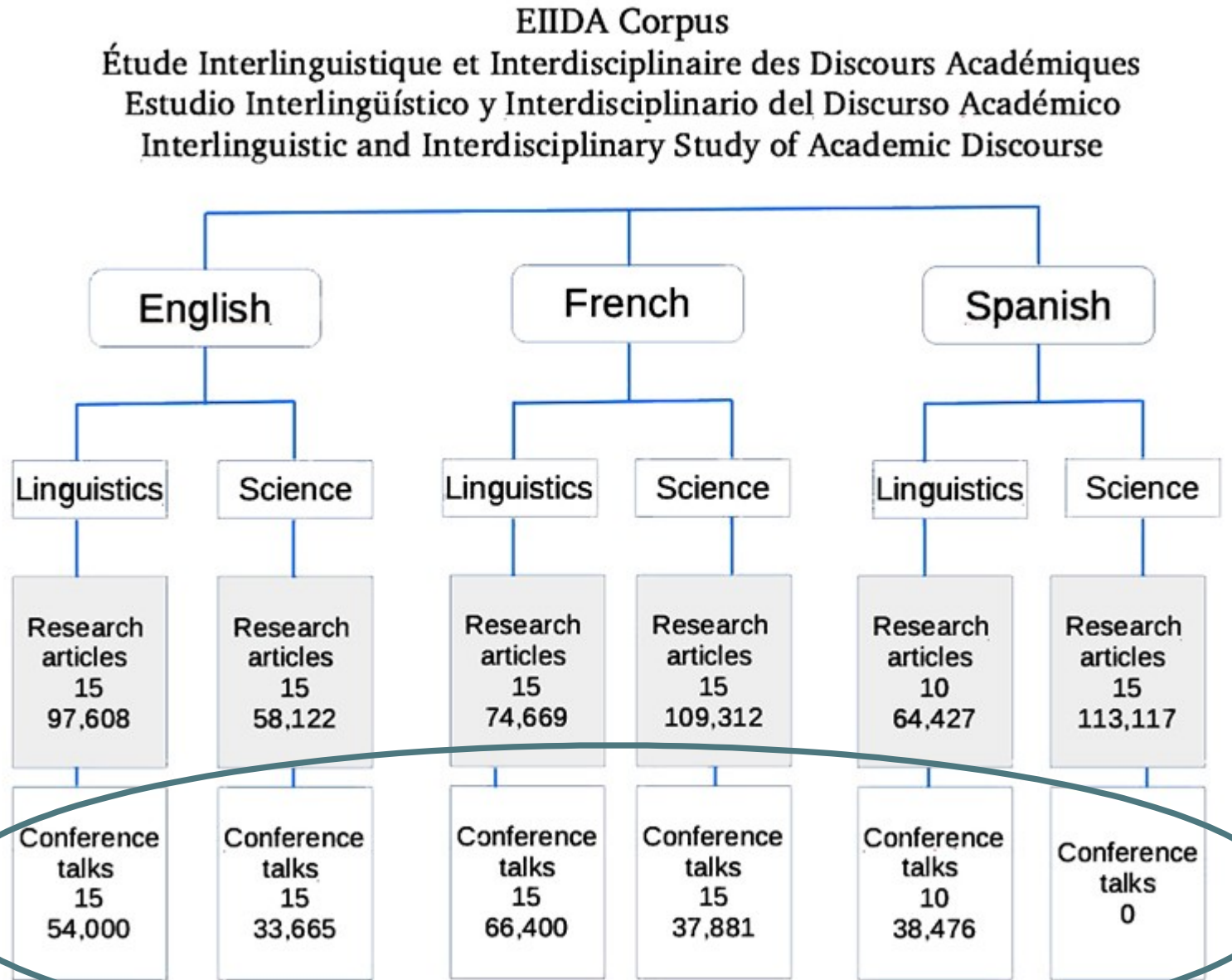


Image de:

<http://www.transfers.ens.fr/eiida-etudes-interdisciplinaires-et-interlinguistiques-du-discours-academique?artpa>

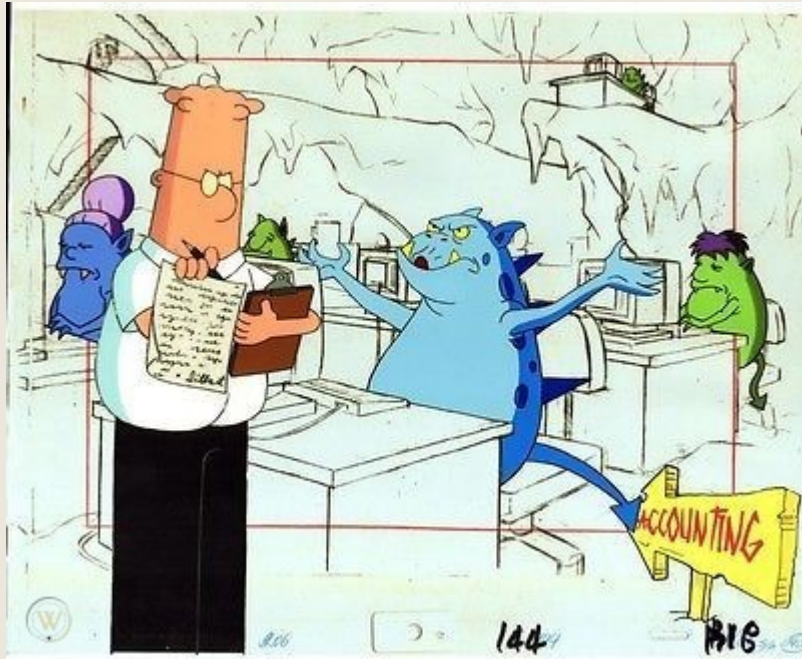
Questions de recherche:

Pourquoi ?

- La prosodie = pont entre lexique – syntaxe – org. du discours
 - *I think we need... ≠ i think WE need... ≠ i THINK we need...*
- Est-ce qu'un phonostyle existe qui serait caractéristique du genre « communication à un congrès scientifique » ?
- Si oui, comment exploiter cela avec des apprenants?
- Phonostyles
 - *des styles sonores « tels qu'ils sont perçus en tant que caractéristiques d'un individu (jeune, vieux, homme, femme), d'un groupe social (prolétaire, bourgeois), ou d'une circonstance particulière (discours politique, sermon, etc.). » (Léon, 1993 : 3).*

Financement

- CORLI = financement de l'alignement & l'annotation par Adrien Méli (Consortium 'Corpus, Langues, Interactions')
- *Via* le CNRS et le LLF



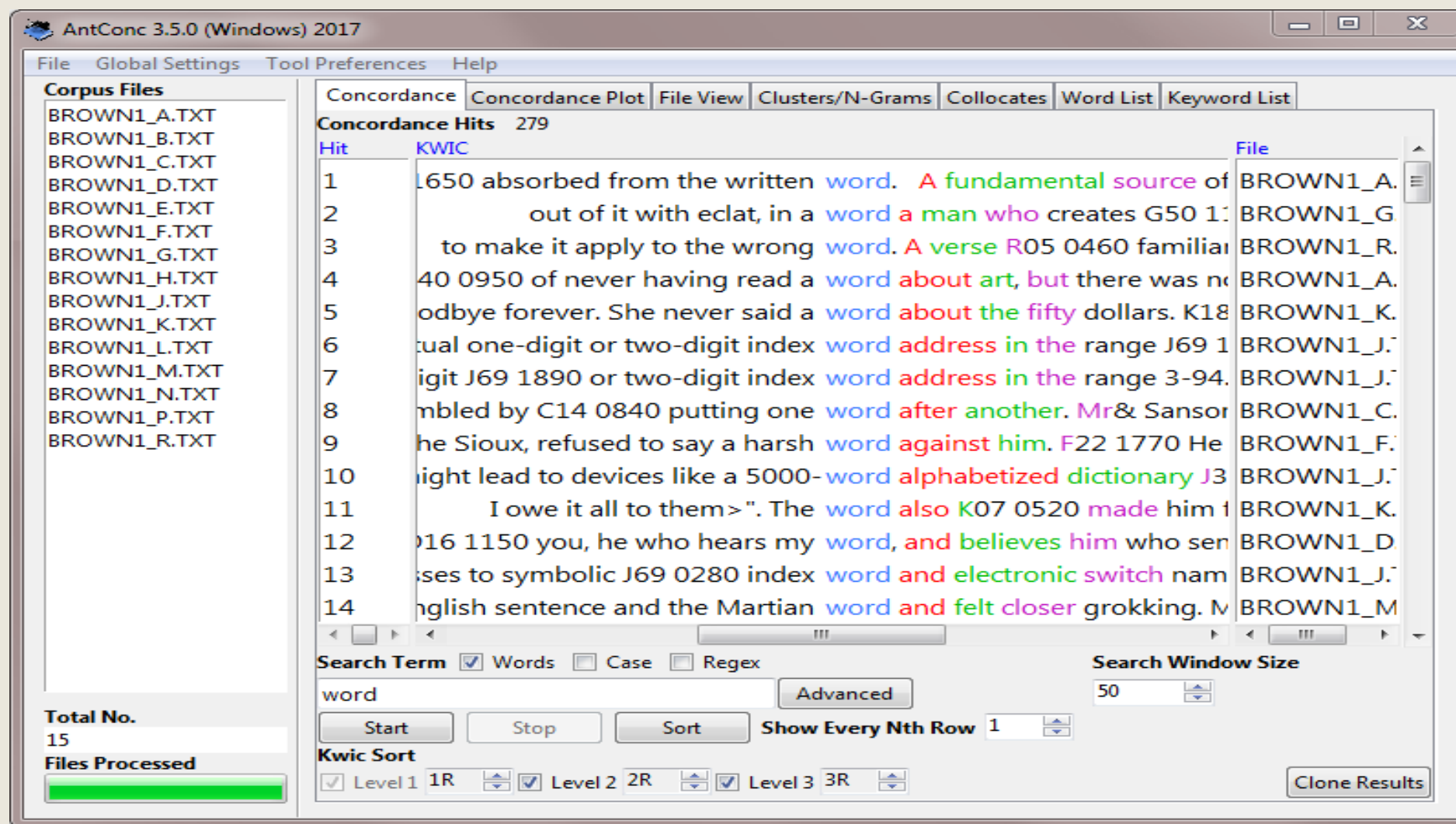
Chronologie

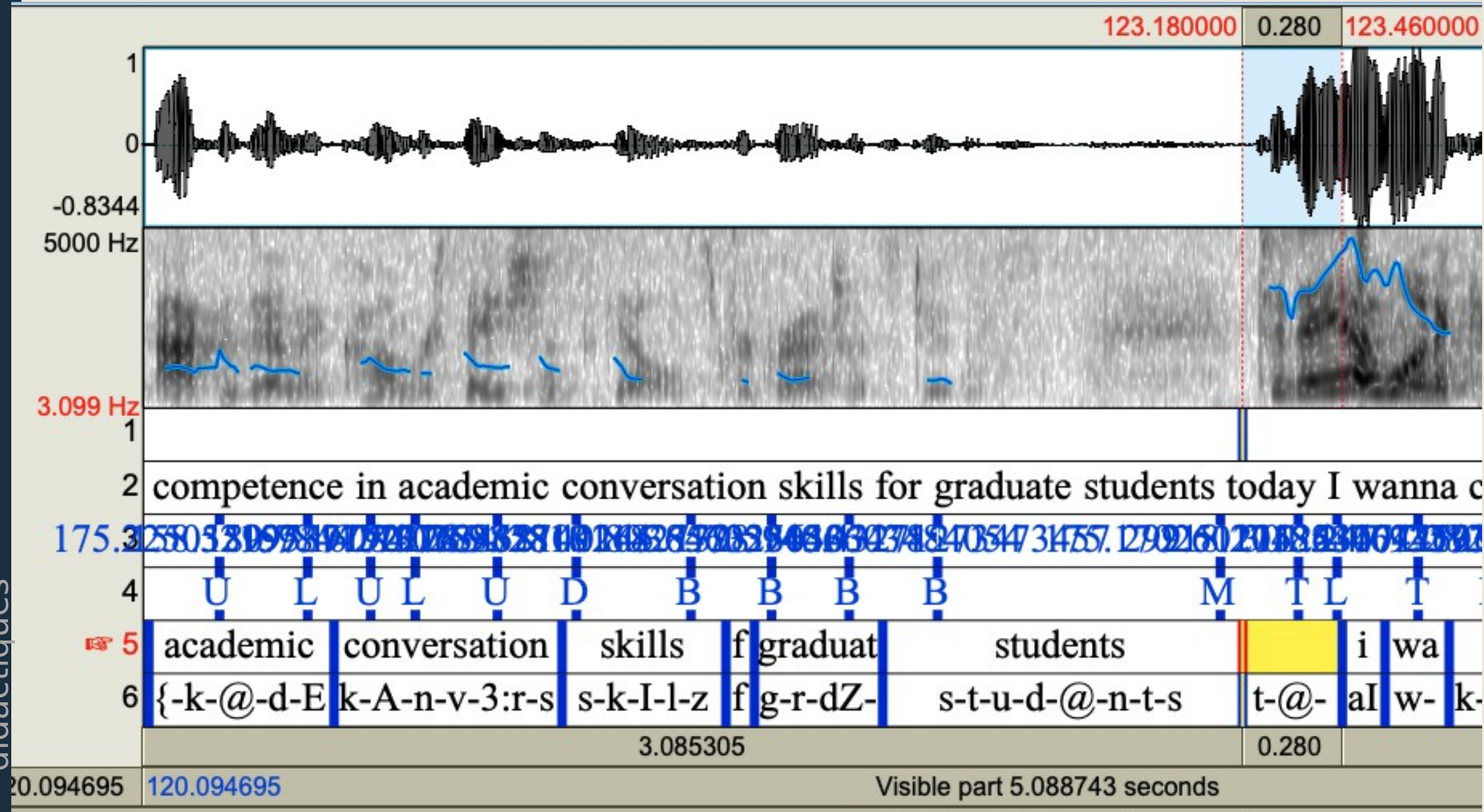
- 1. Collecte (2012-17)
 - 2. Transcription (2016-17)
 - 3. Normalisation (Doriane Simonnet – 2017)
 - 4. Annotation POS + ScienQuest (Achille Falaise – 2018)
 - 5. Alignement
 - 6. Annotation prosodique
 - 7. ScienQuest (Achille Falaise – 2020)
- (Adrien Méli 2019-2020)
- EIIDA's first life
- EIIDA's second life

Focalisation sur l'oral: *tier vs token*

- Différences entre un corpus écrit & un corpus oral
- Impact sur requêtes possibles / « facilitées » par le format

Focalisation sur l'oral: *token* vs *tier*





Au départ de cette seconde vie

<div type="lecture">

<u who="#s1" PID="JMS"> our second speaker of this morning is Jane Freeman who is <w type="ono">um</w> director of the writing in English language support services of University of Toronto, <w type="ono">uh</w> she has the best EAP website in the world

<vocal who="#su-f">

<desc>laugh</desc>

</vocal>

and <w type="ono">uh</w> so i'm very pleased that she's able to come back and talk to us again Jane </u>

<u who="#s2" PID="JF"> thank you. i'm pleased to be here it's maybe the first time in the history of conference presentations that a speaker gets to go early instead of late.

<vocal>

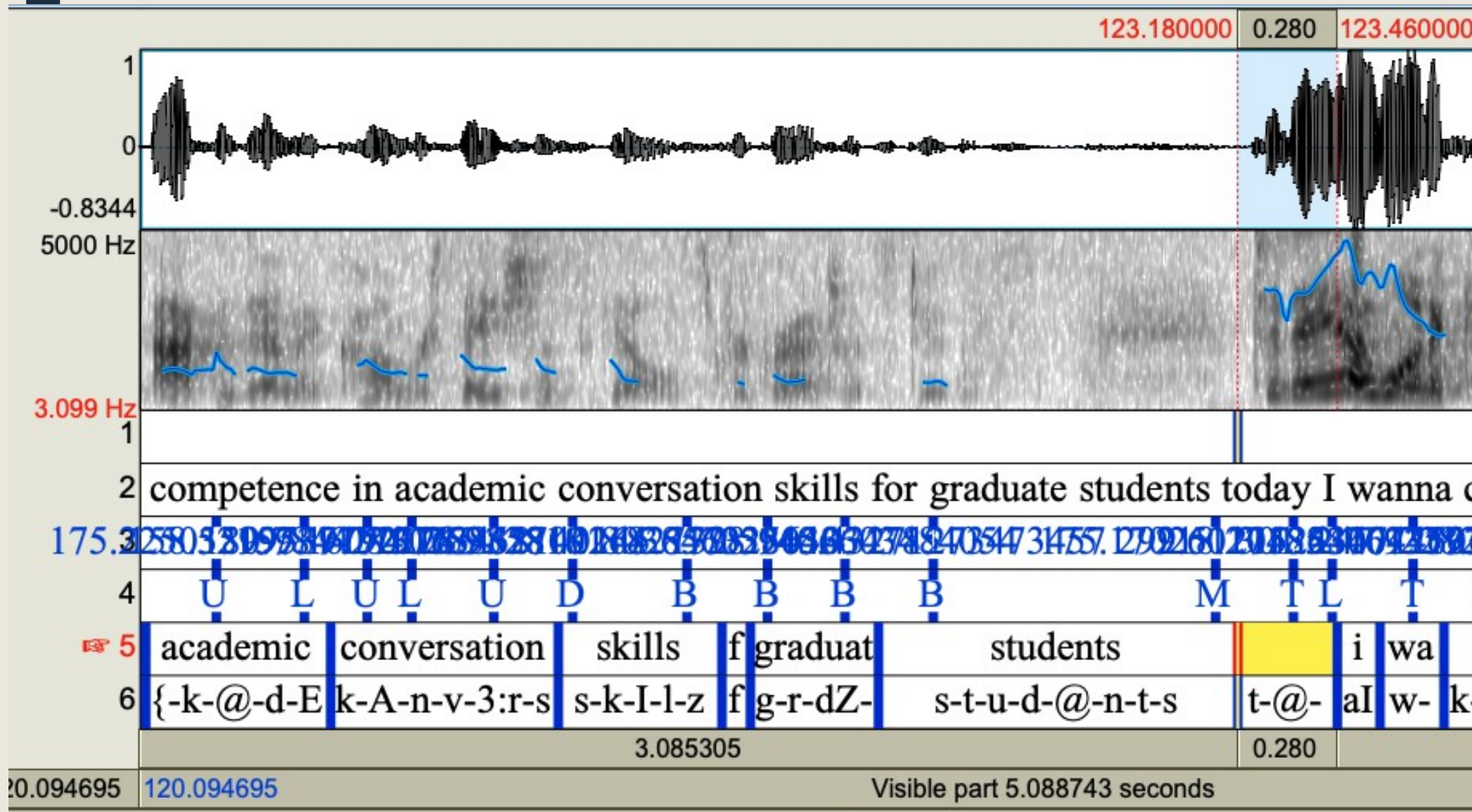
<desc>laugh</desc>

</vocal>

uh i'm talking about making the tacit explicit helping graduates students to enter discourse community as experts with other experts. now clearly in this room i don't need to persuade anyone that concepts of effective academic writing and speaking vary from one discourse community to another. and in the increasingly multicultural interdisciplinary world of our grad- of our graduate schools graduate students frequently encounter both professors and colleagues whose tacit culturally shaped assumptions are different from their own. when students enter grad school in a new country or in a new discipline, they anticipate they'll have to learn certain linguistic and discourse level conventions in order to demonstrate competence in the discourse community of graduate disciplines. but there's often a significant gap i'm discovering between perceived competence the set of _competences_ they believe they need and already have and situated competence the competences they'll actually need and may not have, i posed that the gap between perceived and situated competence in grad school is exacerbated by what i'm gonna call academic doxa i'm picking at that last term from Bourdieu a set of practices so fundamental to members of a community they never discussed questioned or explained. i'd like to acknowledge too that the terms perceived and situated competence are terms i first used in a paper i co-wrote with Dr Li Xi Huan and presented last fall <w type="ono">um</w> at the TESOL conference and there we were looking specifically at the problems of perceived and situated competence in academic conversation skills for graduate students. today i wanna connect three ideas, i'll begin ex- by examining the role of consciousness in the development of competence, and then i'd like to link that examination of the stages of competence development <w type="ono">um</w> to situation to discourse communities <w type="ono">uh</w> and my goal in connecting these two is to consider what happens when people who are very expert in their fields connect. so rather than focusing on people who are just entering <w type="ono">uh</w> a discourse community i'm looking at people who've been in for a long time but _the_ fact of being experts and moving as experts might cause complications for them. now i'd like to start <w type="ono">um</w> with one possible mechanism through which a problematic gap develops between perceived and situated competence...um, many years ago Louis Robert Robinson vice-president of industrial training for international correspondence schools described competence as something that develops in four stages, the first stage is uncompe unconscious incompetence



Ce qu'on espère à l'arrivée

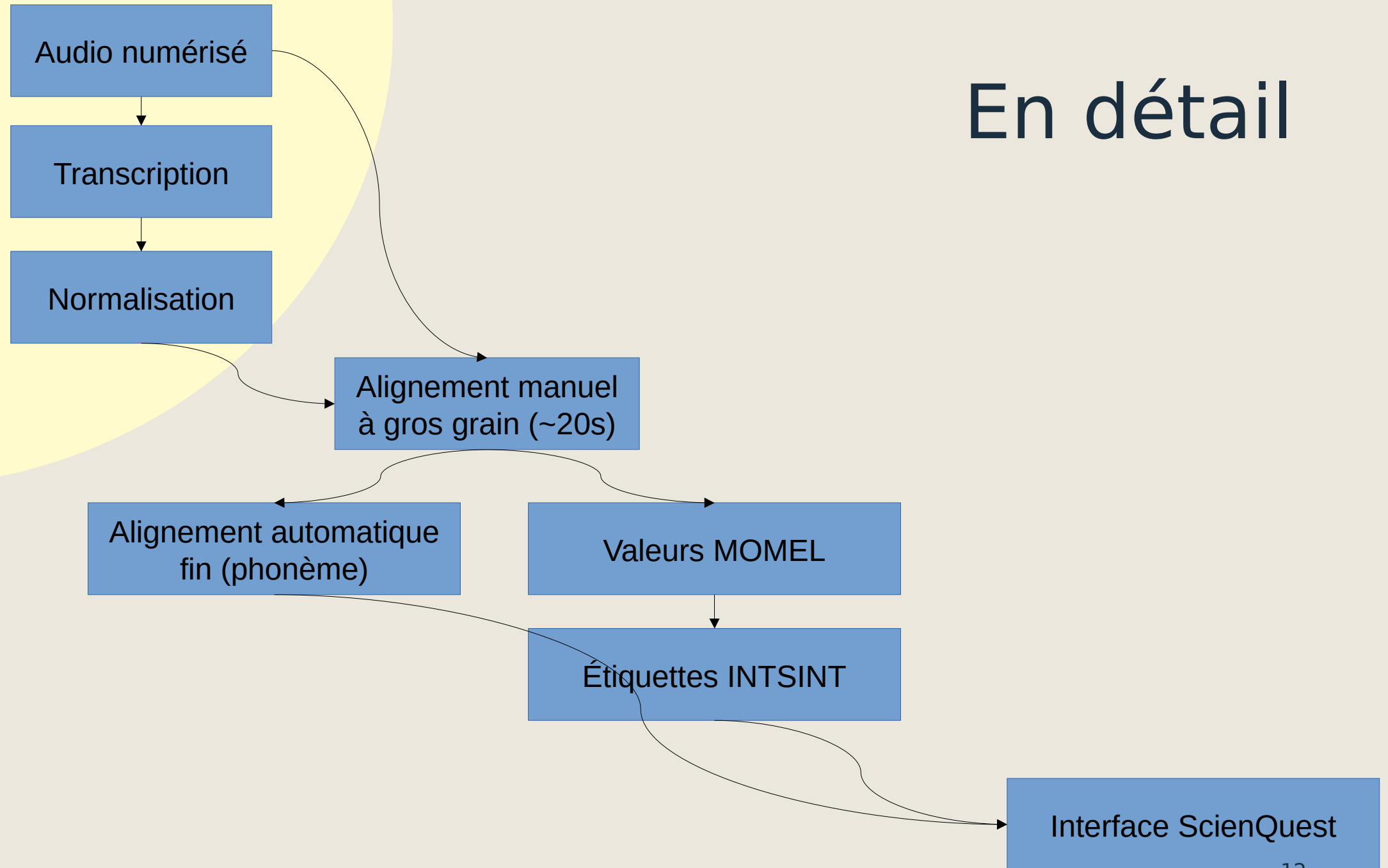


annotation
prosodique

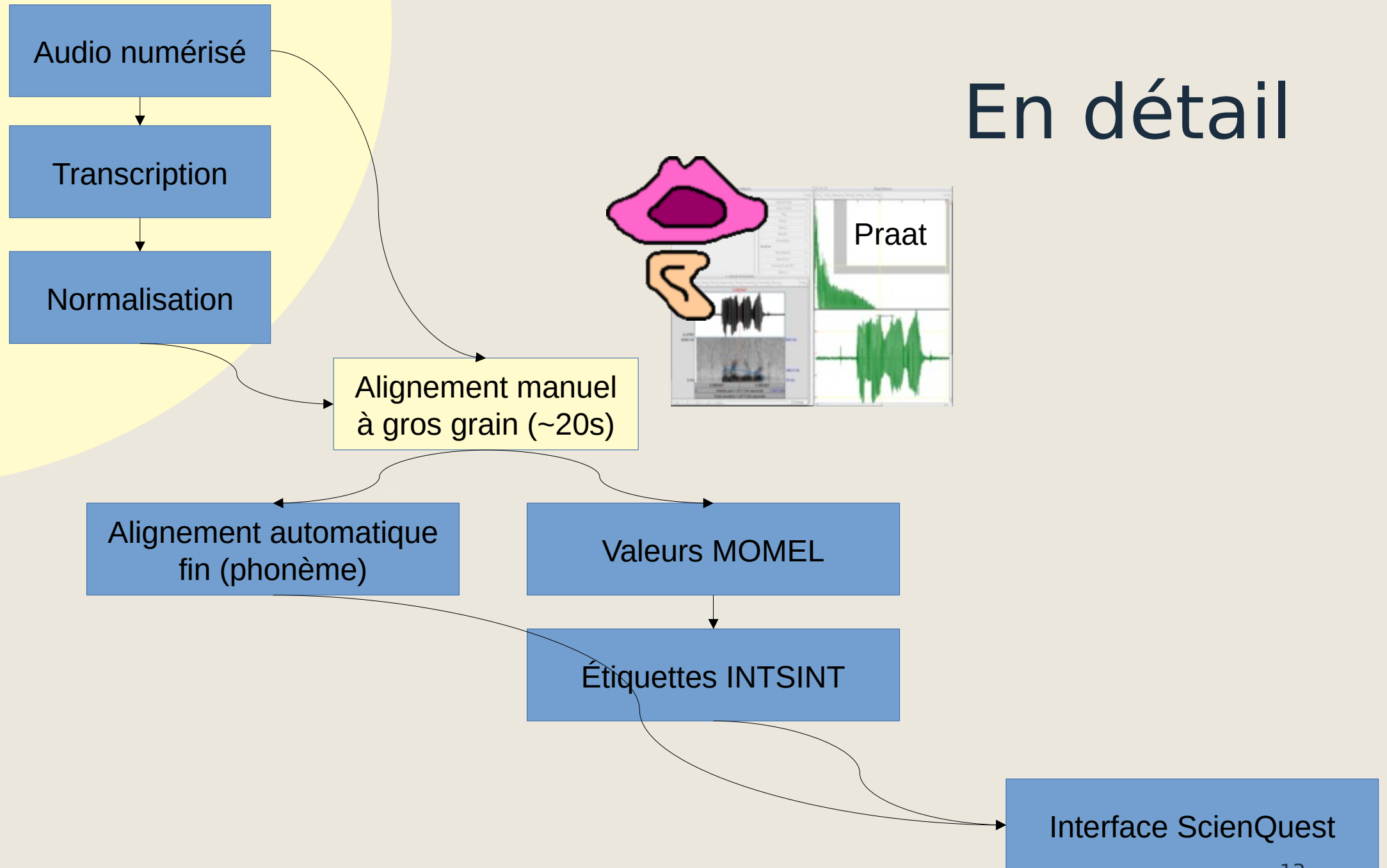
alignement

En détail

1ère vie



En détail



Alignement manuel

File Edit Query View Select Interval Boundary Tier Spectrum Pitch Intensity Formant Pulses Help

multidisciplinary

2186.440000

0.880000 (1.136 / s)

2187.320000

0.7065

0

-0.485

5000 Hz

0 Hz

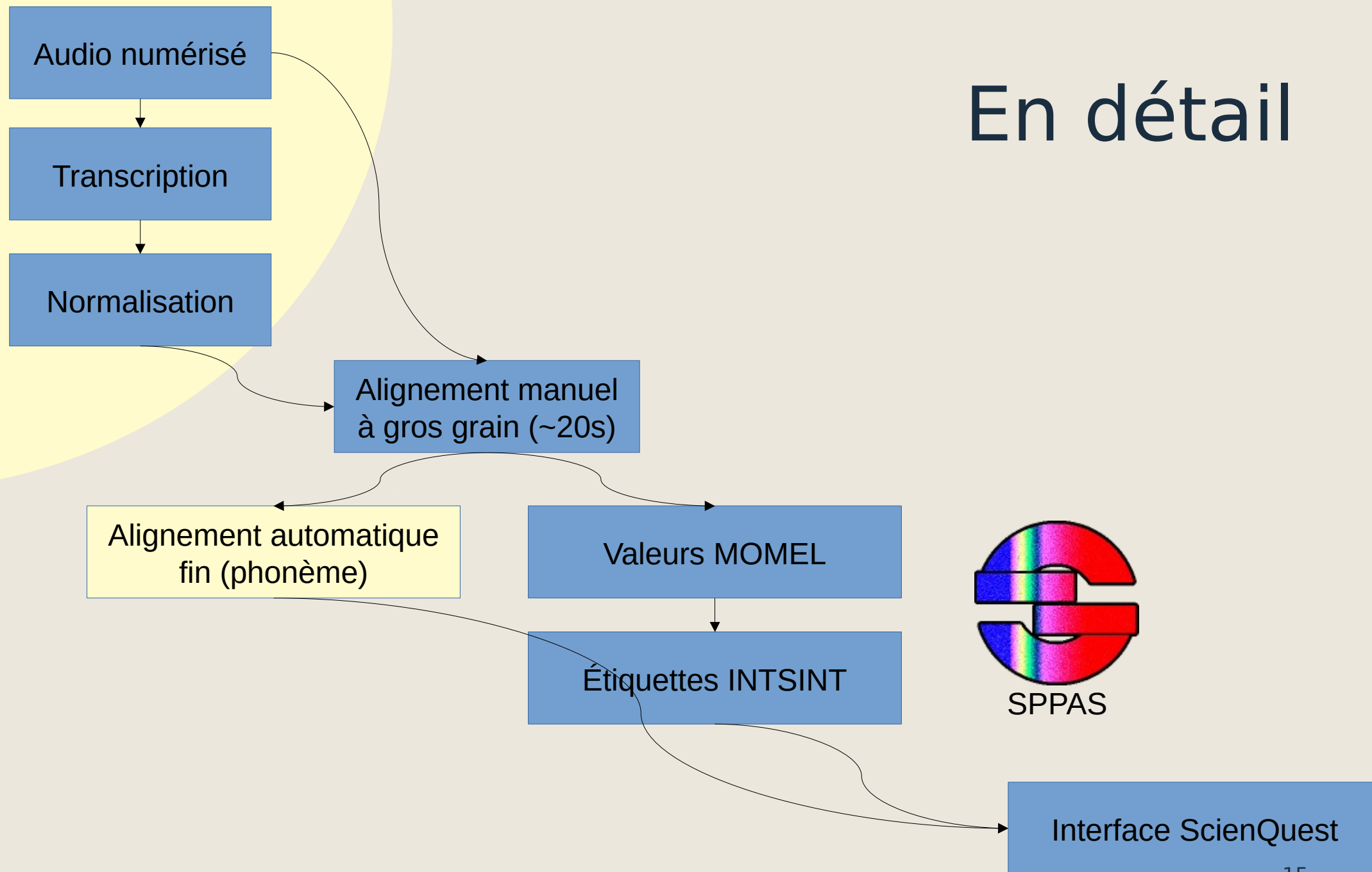
500 Hz

176 Hz

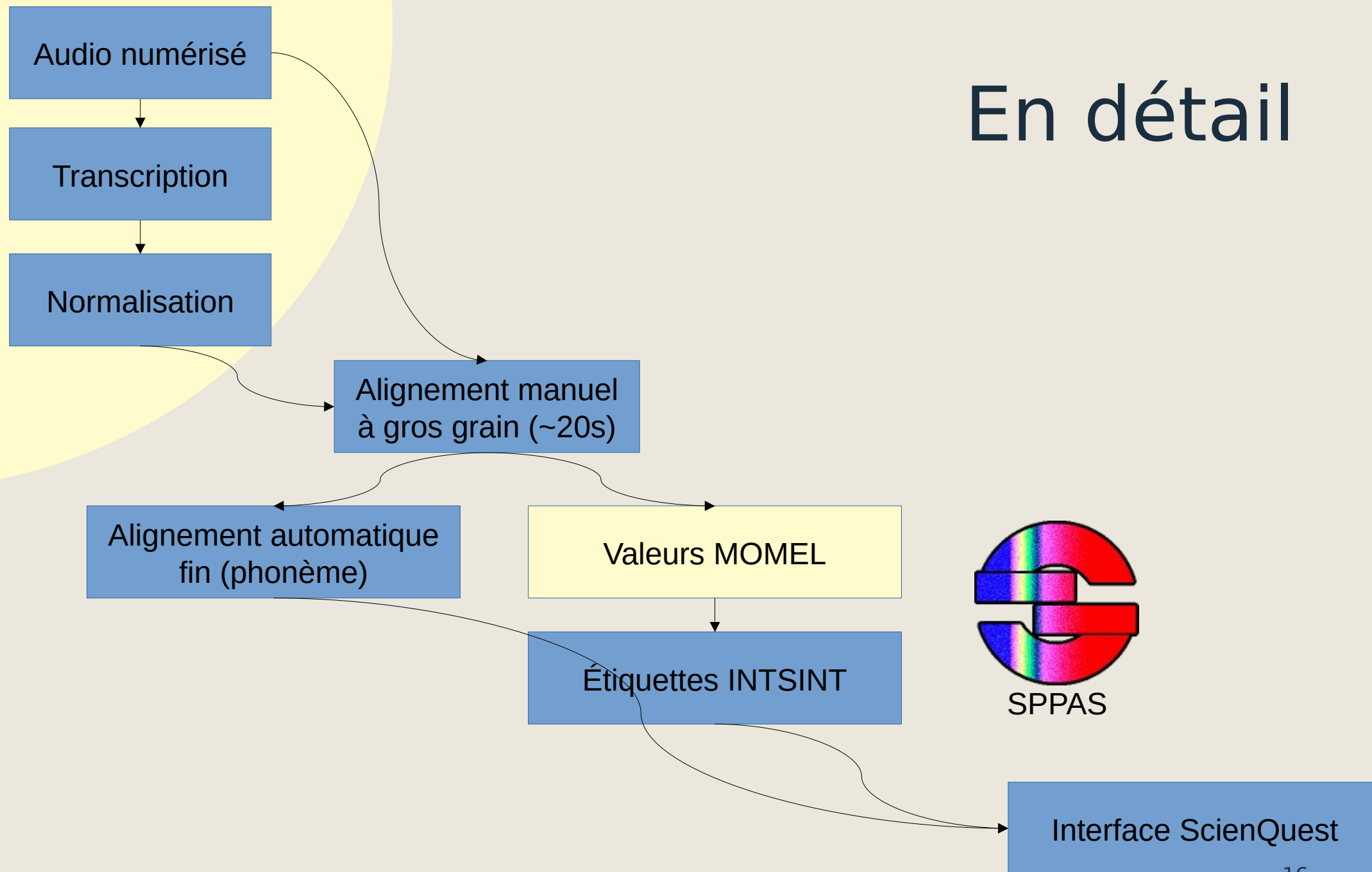
75 Hz

1 paragraph of a proposal were talking about servo valves because in their way of even though the proposal had said the grant said you have to write for a multidisciplinary audience their way of demonstrating expertise is we know exactly what we're gonna buy so that because our needs are that precise we're we're that far along in our planning but for for the medical practitioners who were reading the document and having to decide

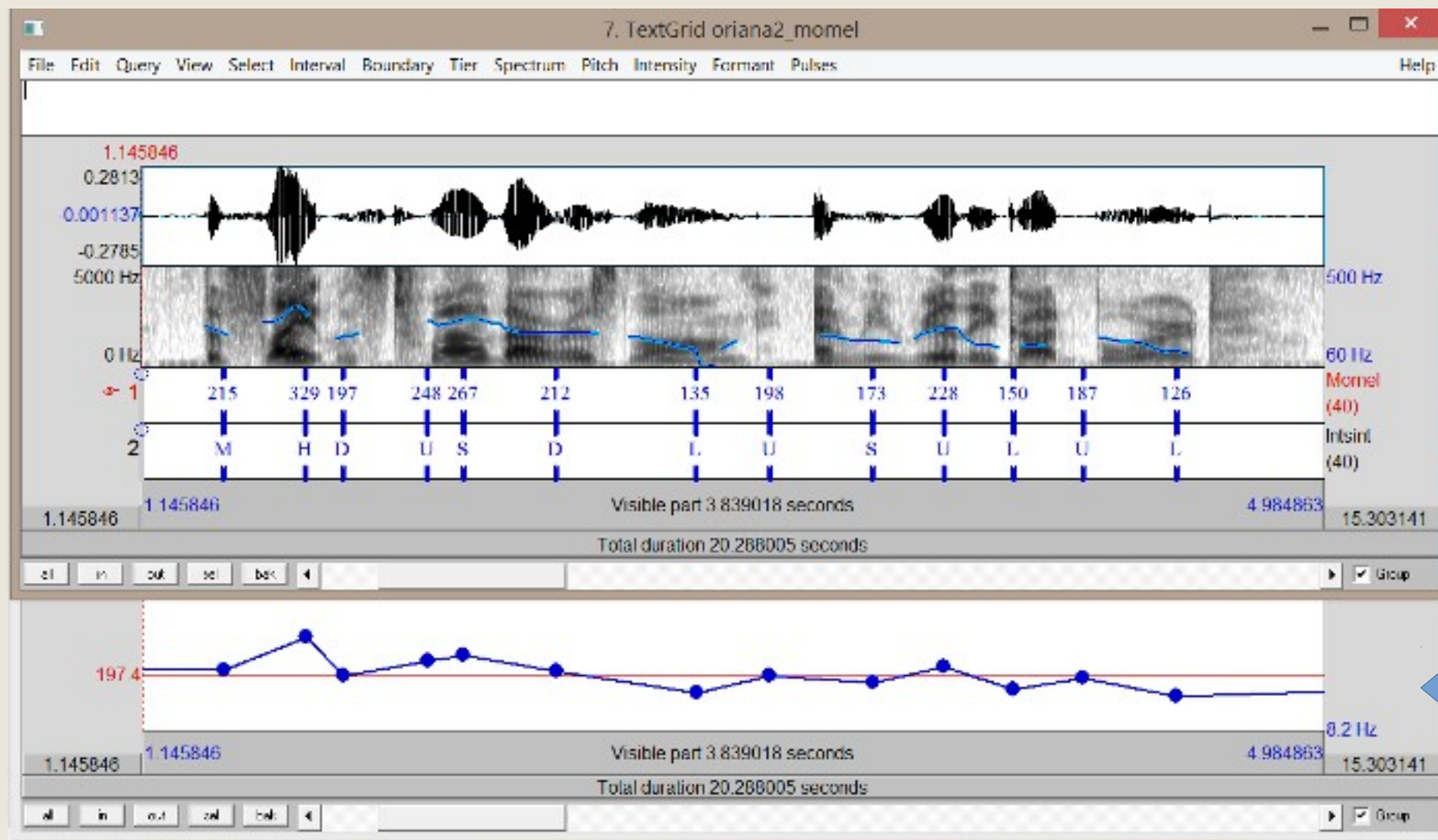
En détail



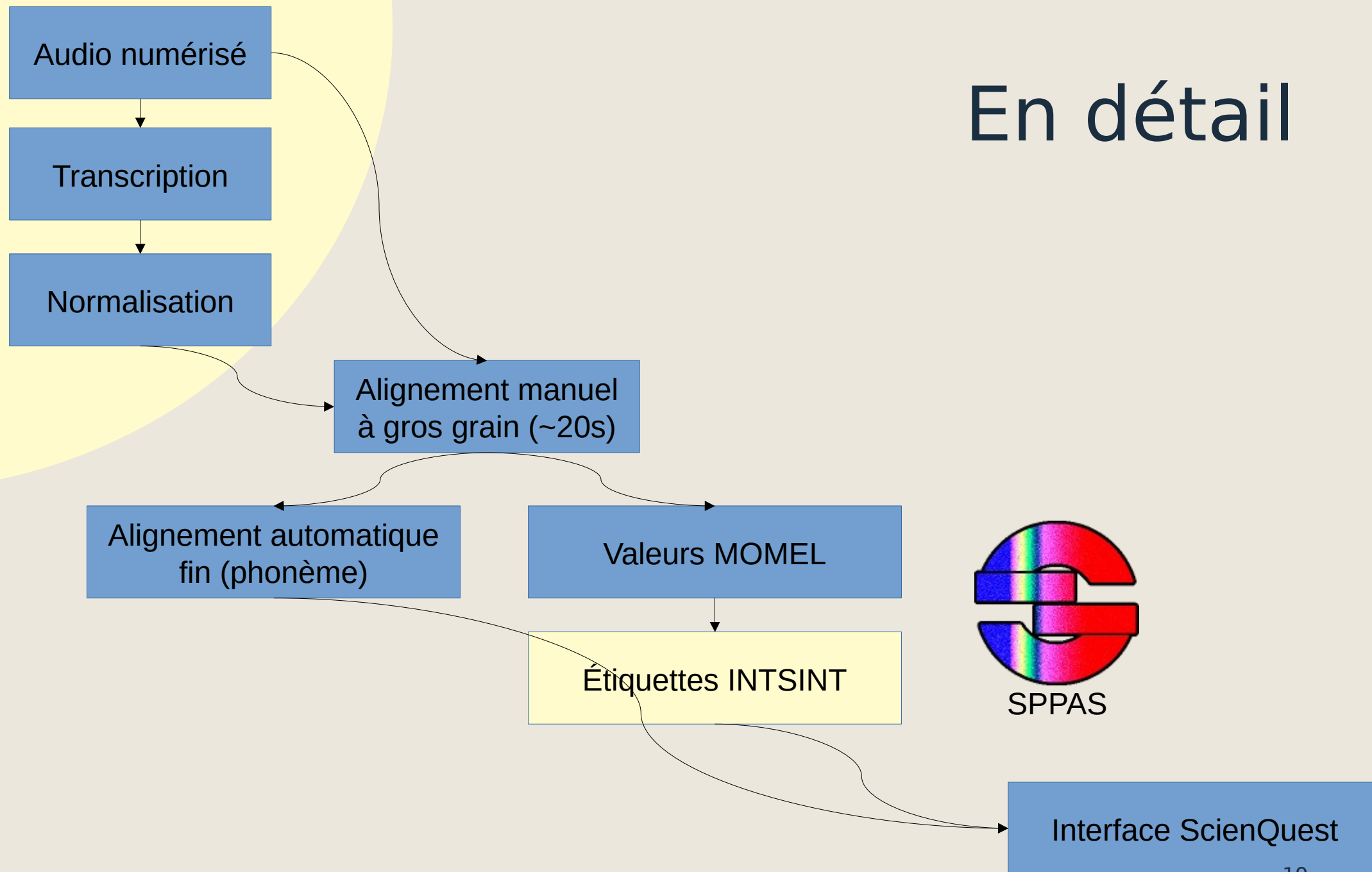
En détail



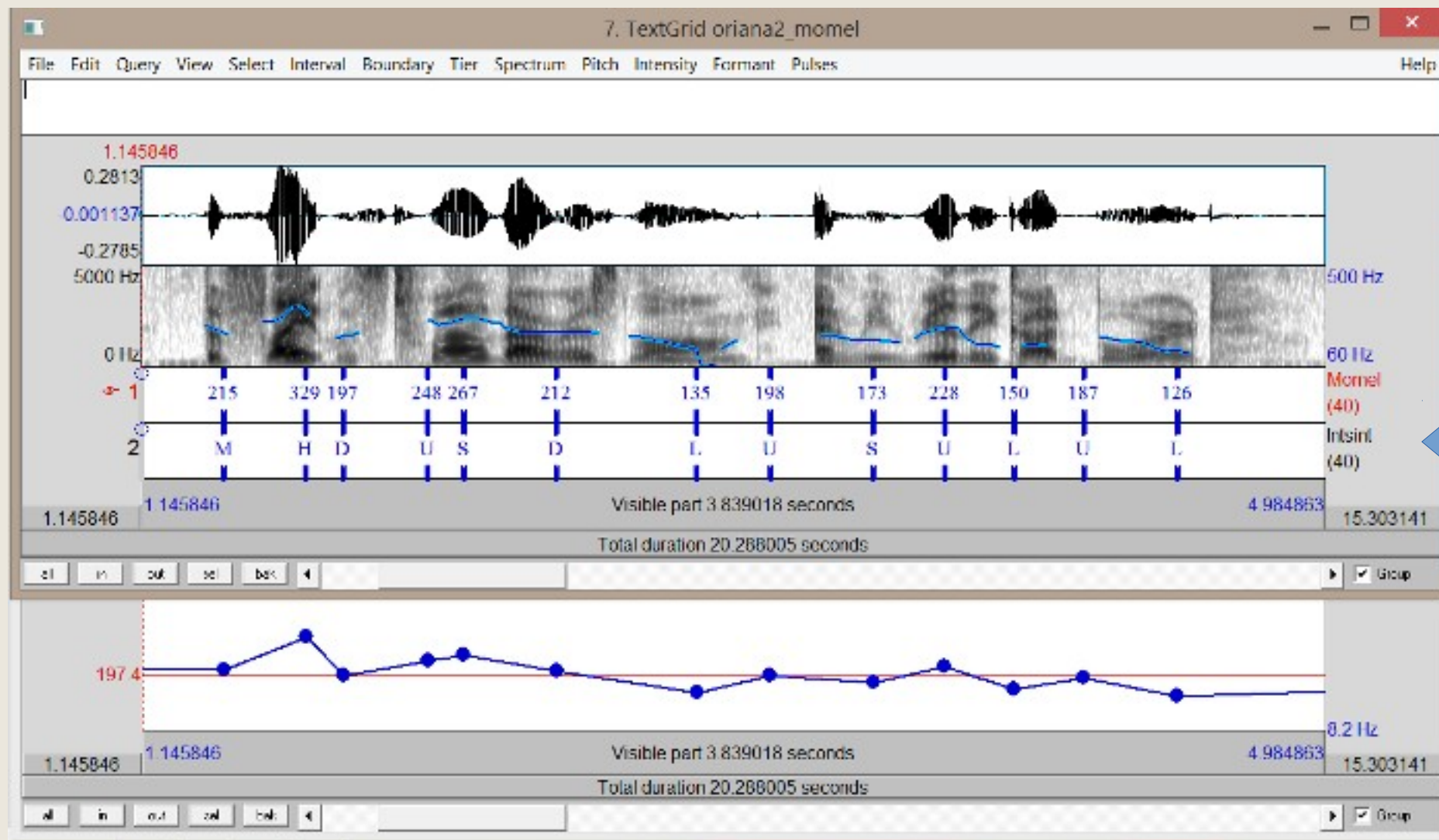
Valeurs MOMEL



En détail

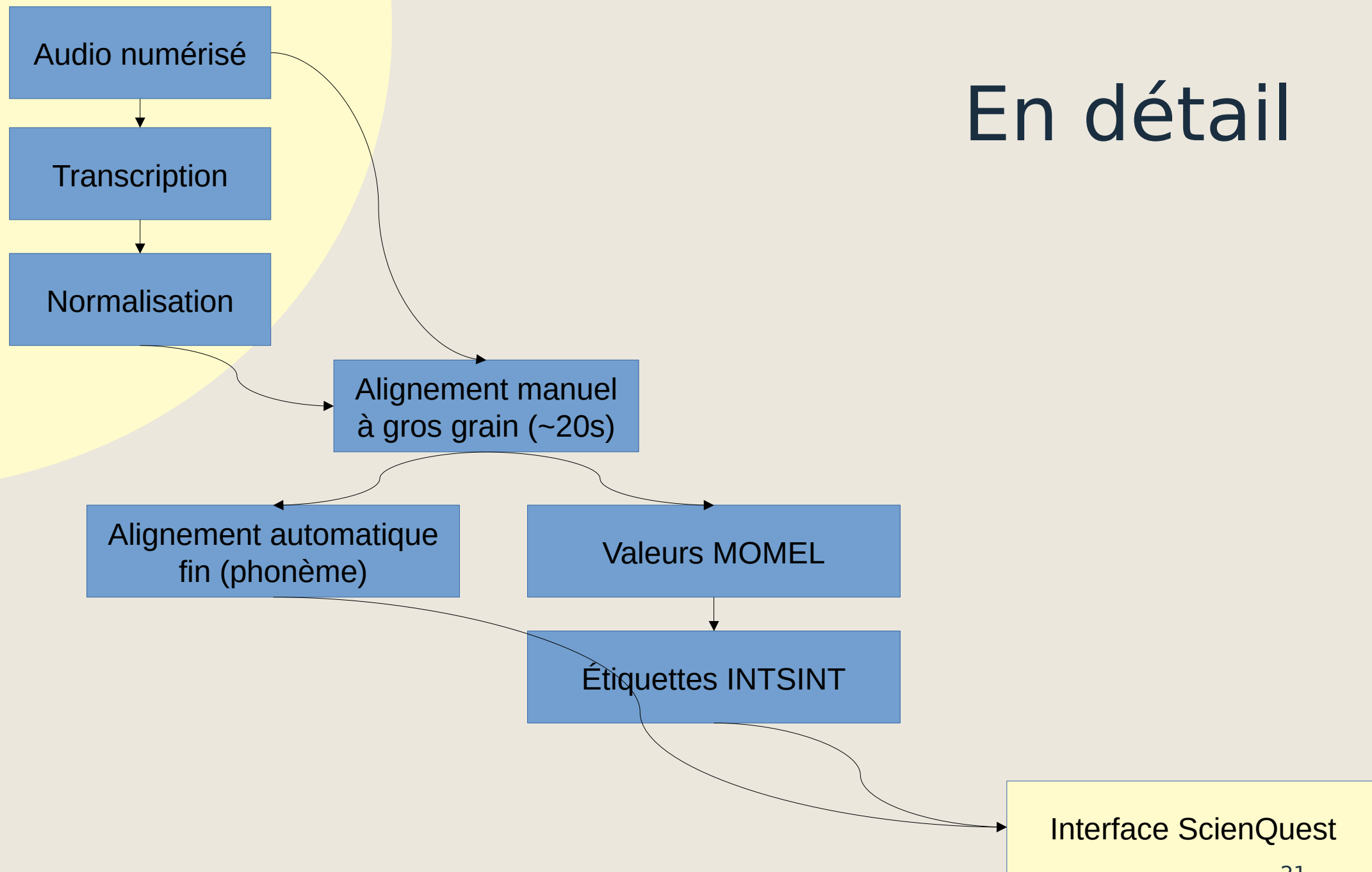


Étiquettes INTSINT



Top
Mid
Bottom
Higher
Lower
Same
Upstepped
Downstepped

En détail



Perspective : intégration dans ScienQuest

- La communauté des corpus de parole a déjà ses propres outils pour étudier la prosodie.
- Comment rendre accessible l'information prosodique à la communauté des corpus écrits ?
 - Retour à l'audio
 - Recherche à partir de symboles « visuels » : 

 - Visualisation des courbes MOMEL

Perspective : intégration dans Sciences

Mot 1

Forme ?

Lemme ?

Catégorie ?

Chercher 8 occurrences

☒ #	Contexte gauche	Occurrences	Contexte droit	Texte
☒ 1	can be applied to the development of more effective teaching	methodology	methodologies and materials for international business students . thank you	#22 - Linguistique - Parole - Conférence
☒ 2	a given experiment merely through mention of the subject and	methodology	and therefore did not need an explicit purpose statement .	#23 - Linguistique - Parole - Conférence
☒ 3	especially in describing	methodology	. while academic journal articles written by experts hardly use	#23 - Linguistique - Parole - Conférence
☒ 4	so we are still looking for a	methodology	in order to understand what is happening in professional practices	#25 - Linguistique - Parole - Conférence
☒ 5	as an example of that i 'm gonna talk about	methodology	i find my graduate students don 't get confused about	#29 - Linguistique - Parole - Conférence
☒ 6	across for experts across uh disciplinary boundaries because they know	methodology	is something that needs to be examined and and tried	#29 - Linguistique - Parole - Conférence
☒ 7	we set up we used the same sort of uh	methodology	uh we set up certain modeling goals where we would	#40 - Science - Parole - Conférence
☒ 8	diversity of bacteria okay but they did not have the	methodology	to assess their importance in the ocean context okay with	#45 - Science - Parole - Conférence

Résultats de 1 à 8 sur un total de 8

Télécharger [\[CSV\]](#) [\[HTML\]](#) [\[XLSX\]](#) – Sauvegarder [\[Concordances\]](#) – Restaurer [\[Concordances\]](#)

Perspective : intégration dans Sciences

Mot 1

Forme ⓘ

Lemme ⓘ

Catégorie ⓘ

8 occurrences

☒ #	Contexte gauche	Occurrences	Contexte droit	Texte
☒ 1	can be applied to the development of more effective teaching	methodology	methodologies and materials for international business students . thank you	#22 - Linguistique - Parole - Conférence

Afficher

☒ Contexte

☐ Arbre syntaxique

going on to our hypotheses uh because we didn 't actually find any clear differences between the two corpora in either of the two features investigated , it would seem that these persuasive strategies are actually characteristic of the genre and transcend medium of interaction . on the other hand we did have some interesting features that emerged along the way that would uh warrant further investigation for example face - to - face face - to - face presentations tended to be much longer , we had some jokes we had digressions , they also had a more interpersonal slant uh for example some engagement strategies to involve other members of the team , we 've had things like as Tom mentioned but also to involve the live audience . these features were almost completely lacking in the teleconference presentation which incurs in secret draft . so finally just to wrap up , uh let 's see ... just to conclude with some pedagogical implications , um a better understanding of these persuasive strategies used in authentic financial presentations , both face - to - face and teleconference , can be applied to the development of more effective teaching **methodology** methodologies and materials for international business students . thank you . .

Perspective : intégration dans Scier

Mot 1

Forme

methodology

Lemme

Catégorie

+

Chercher

8 occurrences

☒ #	Contexte gauche	Occurrences	Contexte droit	Texte
☒ 1	can be applied to the development of more effective teaching	methodology	methodologies and materials for international business students . thank you	#22 - Linguistique - Parole - Conférence

Afficher

☒ Contexte
 ☐ Arbre syntaxique

the genre and transcend medium of interaction . on the other hand we did have some interesting features that emerged along the way that would uh warrant further investigation for example face - to - face face - to - face presentations tended to be much longer , we had some jokes we had digressions , they also had a more interpersonal slant uh for example some engagement strategies to involve other members of the team , we 've had things like as Tom mentioned but also to involve the live audience . these features were almost completely lacking in the teleconference presentation which incurs in secret draft . so finally just to wrap up , uh let 's see ... just to conclude with some pedagogical implications , um a better understanding of these persuasive strategies used in authentic financial presentations , both face - to - face and teleconference , can be applied to the development of more effective teaching methodology methodologies and materials for international business students . thank you .

▶

0:00 / 0:00

🔊

Perspective : intégration dans Scien

Mot 1

Forme

Lemme

Catégorie

Prosodie



Chercher

8 occurrences

☒ #	Contexte gauche	Occurrences	Contexte droit	Texte
☒ 1	can be applied to the development of more effective teaching	methodology	methodologies and materials for international business students . thank you	#22 - Linguistique - Parole - Conférence

Afficher

☒ Contexte☐ Arbre syntaxique

the genre and transcend medium of interaction . on the other hand we did have some interesting features that emerged along the way that would uh warrant further investigation for example face - to - face face - to - face presentations tended to be much longer , we had some jokes we had digressions , they also had a more interpersonal slant uh for example some engagement strategies to involve other members of the team , we 've had things like as Tom mentioned but also to involve the live audience . these features were almost completely lacking in the teleconference presentation which incurs in secret draft . so finally just to wrap up , uh let 's see ... just to conclude with some pedagogical implications , um a better understanding of these persuasive strategies used in authentic financial presentations , both face - to - face and teleconference , can be applied to the development of more effective teaching methodology methodologies and materials for international business students . thank you .

 0:00 / 0:00 

Aspects didactiques: Exploitation

« Les langues de spécialité doivent être décrites en tant que compétence si l'on veut en faciliter la transmission aux apprenants. » (RADELAS, AaC).

La prosodie: indicateur potentiel de compétence pragmatique en LSP (voir par exemple Taguchi, N, 2019; Taguchi & Roeever, 2017))

Aspects didactiques: Objectifs larges

Favoriser une prise de conscience de l'impact ...

- des spécificités du contexte;
- de la variation de la parole et/ou du mode oral:
 - Bousculer la représentation d'une langue & de la norme.

Aspects didactiques: Objets prioritaires

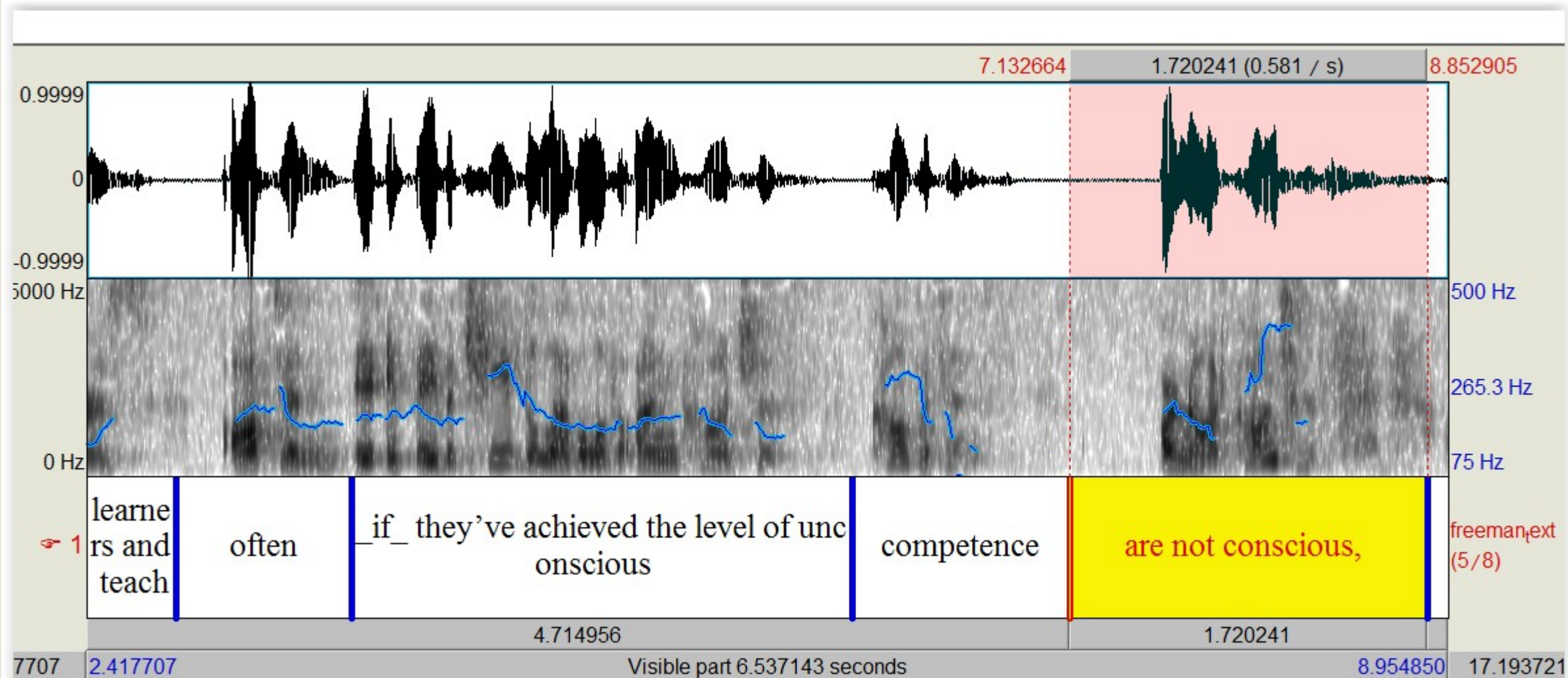
- Traits utiles = rendant plus facile la compréhension du message
- Traits apprenables & maîtrisables
 - Groupes de souffle
 - Débit de parole
 - + rapide dans styles formels
 - Débit d'articulation: - préparé, + de pauses
 - les journaux parlés: DA le +élevé (6,13 et 5,87 syl./s)
 - le discours politique: DA le +lent (4,85 syl./s)
- Traits apprenables mais plus difficilement maîtrisables
 - les proéminences *chaque langue les marque différemment
- Traits enseignables?

Application I: Groupes de souffle

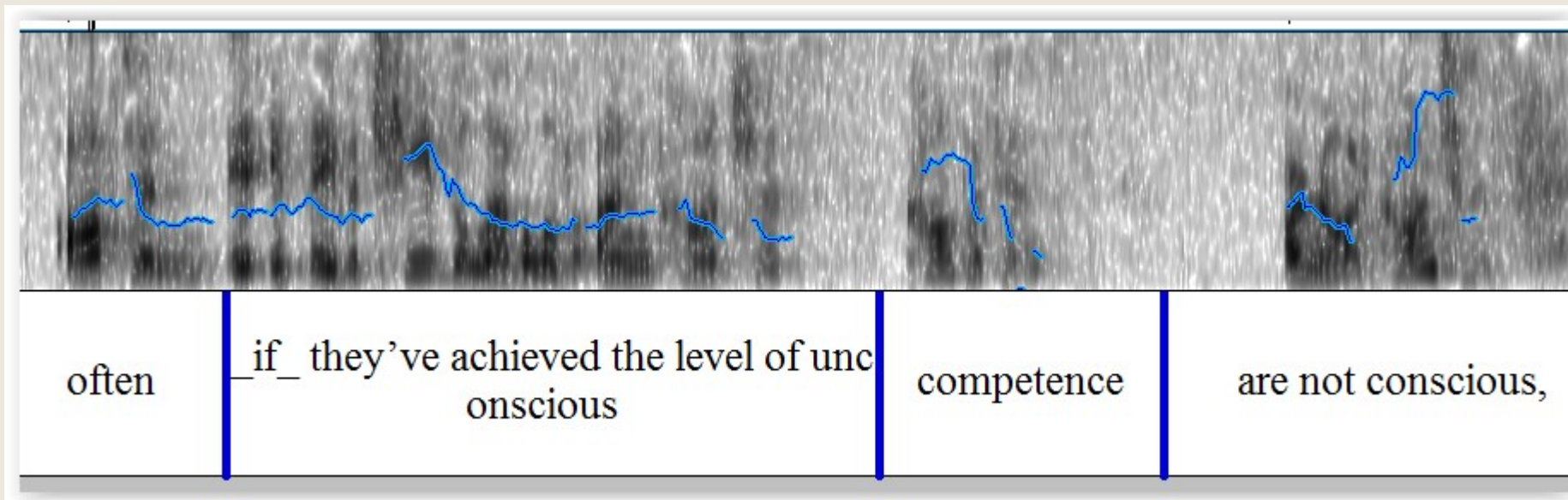
- learners and teachers
- often
- _if_ they've achieved the level of unconscious competence
- are not conscious, of what they're expecting and that therein lies the problem.

- Chaque ligne = 1 groupe de souffle
- Variation
- Dysfluences

Application I: Groupes de souffle



Application II: Proéminence



« du carbone pardon en effet non seulement elles participent à la contre-() à **la pompe** du carbone via la photosynthèse mais également à **la contre-pompe** du carbone via la biocalcification en effet ... » (FROG_11)

Conclusion

- La prosodie dans un contexte spécialisé = la communication scientifique
- Extension du corpus EIIDA: pour
 - *livrer des outils à la communauté* (EIIDA via Orféo + l'oral dans ScienQuest)
 - *permettre des analyses prosodiques des LSPs*
 - *améliorer l'accessibilité & l'utilité des outils*
 - *améliorer l'enseignement de la prononciation des LPSs*

Merci / Thank you / Gracias

Références

- Bigi, B. (2015). SPPAS - Multi-lingual Approaches to the Automatic Annotation of Speech. *The Phonetician*, n° 111-112 / 2015-I-II. pages 54-69, ISSN 0741-6164.
- Carter-Thomas, S. et M-P Jacques. (2017). Interdisciplinary and interlinguistic perspectives on Academic Discourse: the mode variable. Introduction to the special issue on the French EIIDA. *Chimera: Romance Corpora and Linguistic Studies*, 4 (1), 1-11. <https://revistas.uam.es/index.php/chimera/article/view/7810>.
- Hirst, D., Di Cristo, A., & Espesser, R. (2000). Levels of representation and levels of analysis for the description of intonation systems. In *Prosody: Theory and experiment*, Dordrecht :Springer. 51-87. ISBN 978-90-481-5562-0
- Léon, P. (1993). *Précis de phonostylistique : Parole et expressivité*. Paris : Nathan Université. ISBN : 2-09-190065-6.
- Moreau, M.-L. 1997. *Sociolinguistique. Concepts de base*. Mardaga, Liège.
- Simon, A-C., Auchlin, A., Avanzi, M. et J-P Goldman. (2010). Les phonostyles : une description prosodique des styles de parole en français. Dans Abecassi, M. et G. Ledegen, (dirs.) *Les voix des Français. En parlant, en écrivant*, Berne : Peter Lang. 71-88. ISBN : 978-3-0343-0170-1.
- Simon, A. C., Avanzi, M., Goldman, J.-Ph. 2008, "La détection des proéminences syllabiques. Un aller-retour entre l'annotation manuelle et le traitement automatique", *Actes du Congrès Mondial de Linguistique française*, papier n° 151, Paris. [doi:10.1051/cmlf08256].
- Taguchi, N. (2019). *The Routledge Handbook of Second Language Acquisition and Pragmatics*, Routledge: Oxon, England. ISBN 9780815349761.
- Taguchi, N. & C. Roever (2017). *Second Language Pragmatics*. Oxford University Press: Oxford. ISBN 978 019 4200585