



WASEDA University



UNIVERSITÉ
DE GENÈVE



東京外国語大学 Tokyo University of Foreign Studies

10 ans d'IPFC brève rétrospective et perspectives de développement

Sylvain Detey (SILS, Université Waseda, Japon),
Isabelle Racine (ELCF, Université de Genève, Suisse)
Yuji Kawaguchi (TUFS, Japon)



IPFC2018 – MSH (Paris) 26 & 27 novembre 2018



En décembre 2008 à la MSH à Paris...

- 2008: Pas de iPhone... pas de photo
- 2008: Premier CMLF
- 2008: PFC a presque 10 ans... naissance d'IPFC (Detey & Kawaguchi 2008)

Définition du projet, premier protocole, premières données japonophones Moodle, appel à collaboration...



Interphonologie du Français Contemporain (IPFC)
Récolte automatisée des données et apprenants japonais

Sylvain Detey
Université de Rouen

Yuji Kawaguchi
Tokyo University of Foreign Studies

2008 - 2009 - Collège Phonologique du Français Contemporain : variétés, interfaces, répertoires

Introduction

- PFC: image du français parlé dans son unité et sa diversité par des locuteurs natifs
 - Monoïngues (France...), bilingues (Amérique...), plurilingues (Afrique...)
 - Français L2? → volet didactique de PFC: PFC-EP (enseignement du français)
 - Apprenants de français (L2, L3...): interphonologie ("interlanguage phonology")
- IPFC: *Interphonologie du français contemporain*
image du français parlé dans sa diversité par des locuteurs non natifs (français langue étrangère - FLE)
- Plan
 - 1) Motivations de IPFC
 - 2) Le protocole IPFC
 - 3) Enquête pilote: IPFC Japon
 - 4) Perspectives et appel à collaboration

L'originalité du projet IPFC:

1. Focus sur la dimension phonético-phonologique (tout en recotant des informations sur la syntaxe, le lex, etc.).
2. Unicité du cœur du protocole qui assure la comparabilité des données.
3. Possibilité d'étudier de manière comparative la variation en L1 et en L2 grâce à la base PFC préexistante (données et outils).
4. Une exploitation pour la recherche (interphonologie) et la pédagogie (PFC-EP).
5. Un réseau international de collaborateurs issu du réseau PFC (Allemagne, Angleterre, États-Unis, Danemark, Italie, Japon, Norvège, Pays-Bas, etc.).

2) Protocole IPFC: première version

- Un protocole pas trop lourd et éprouvé: calqué sur le protocole PFC
- **Terrains possibles:** francophones (ex. Genève) et non francophones (ex. Tokyo)
- **Population:** adaptation du questionnaire sociolinguistique
→ biographie langagière, apprentissage du français...

Niveau CECR (productions orales)	B1 - B2
Certificat de compétence en français	Ex: DELF B1
Statut du français	Ex: 2 ^{ème} langue étrangère
Nombre d'années d'étude du français	Ex: 5 ans
Niveau d'étude universitaire (public/universitaire)	Ex: 2 ^{ème} année
Spécialité (public/universitaire)	Ex: Économie
Niveau de formation	Ex: Diplôme de 2 ^{ème} année universitaire

Detey, S., Kawaguchi, Y. 2008. *Interphonologie du Français Contemporain (IPFC): récolte automatisée des données et apprenants japonais. Journées PFC: Phonologie du français contemporain: variation, interfaces, cognition.* Paris, Dec. 11-13, 2008.

1) Motivations de IPFC: pourquoi / pour quoi ?

IPFC, pourquoi ?

- Trier parti de la structure PFC (protocole, outils, réseau...)
- Absence de corpus oraux pour documenter la phonologie du français en tant que langue étrangère

Le learner Corpus around the World

Validé avec le soutien de la Banque de France et de la Banque pour le Japon

SLABank

Les corpus oraux d'apprenants d'anglais:

Ex 1, *Corpus International Database of Spoken English Interlanguage* (Lindbæk, In *International Corpus of Learner English* (ICLE) (S. Granger, Université catholique de Louvain)

Ex 2, *Standard Speaking Test Corpus* (Japanese learners of English) (Y. Tomi, Tokyo University of Foreign Studies)

Ex 3, *Standard Speaking Test Corpus* (Japanese learners of English) (Y. Tomi, Tokyo University of Foreign Studies)

Procédure

- **Tâches obligatoires**
 - 1) Lecture de trois listes :
 - a) Liste de mots PFC
 - b) Liste de mots spécifique à chaque population
 - c) Liste de logatomes spécifique à chaque population
 - 2) Lecture du texte PFC
- **Tâches optionnelles**
 - 3) Un semestre après le premier enregistrement :
 - a) Lecture de la liste de mots PFC (pour comparaison longitudinale)
 - b) Répétition de la liste de mots PFC (pour comparaison inter-tâche)
 - 4) Conversation guidée (apprenants de niveau B1 et +) réponses à un questionnaire fixe (15 min.)
 - 5) Conversation libre (*matériaux de niveau C1 et +*)

Précisions

- **Lecture des mots :** Mots isolés, difficultés de lecture → enregistré après un semestre, lecture et répétition.
- **Lecture du texte :** Texte tiré en L1 ou dans une langue connue, afin de contextualiser la lecture.
- **Listes spécifiques (mots et logatomes) :** Élaboration en fonction des difficultés « typiques » des apprenants (ex. SVL pour les Japonais). Avec pauses marquées en fin de liste comme dans la liste PFC. Conception pour les logatomes : Liste la liste suivante comme des mots français.
- **Répétition de la liste PFC :** Liste PFC enregistrée par un locuteur français natif de prononciation proche du français de référence. Chaque mot répété trois fois de suite, puis une pause pour répétition par l'étudiant.
- **Conversation guidée :** Signet de niveau B1 et + : la conversation guidée de PFC devient une tâche de discussion entre deux apprenants sur un thème choisi par l'un d'eux (ex. Italie, Espagne, sans liste SVL). C'est un 1 sur 1, avec l'étudiant pose des questions et le premier étudiant anime la discussion pendant 4 minutes environ. Puis inversion des rôles. Env. 10 min.
- **Conversation libre :** Signet de niveau C1 et + : la conversation guidée de PFC devient une tâche de discussion entre deux apprenants sur un thème choisi par l'un d'eux (ex. Italie, Espagne, sans liste SVL). C'est un 1 sur 1, avec l'étudiant pose des questions et le premier étudiant anime la discussion pendant 4 minutes environ. Puis inversion des rôles. Env. 10 min.

2008-2009

- PFC2008: discussion animée le matin: codage/pas codage; logatomes/mots; longueur du protocole; valeur de l'approche perceptive...
- 2008: thèse d'I. Racine
- PFC2008: discussion animée le soir: développement du projet.
- 2009: AFLS2009: finalisation du protocole avec I. Racine et F. Zay, et ajout de données hispanophones (Racine, Detey, Zay & Kawaguchi 2012).

Racine, I., Detey, S., Zay, F. & Kawaguchi, Y. (2012). Des atouts d'un corpus multitâches pour l'étude de la phonologie en L2 : l'exemple du projet *InterPhonologie du français contemporain* (IPFC). In Kamber, A. & Skupiens, C. (éds), *Recherches récentes en FLE*, Berne, Peter Lang, 1-19.

2010: Premières Journées IPFC à Paris

16h15-16h30 : Chantal LYCHE & Unni LUNDSTEDT (U. de Tromsø)

IPFC-norvégien : protocole et contexte

16h30-16h45 : Jeff TENNANT & Nerissa TAYLOR (U. Western Ontario)

IPFC-anglais canadien : protocole et contexte

16h45-17h00 : Freiderikos VALETOPOULOS & Marina CHRISTOFI (U. de Poitiers)

IPFC-grec moderne : protocole et contexte

17h00-17h15 : Dominique NOUVEAU & Janine BERNS (U. Radboud de Nimègue)

IPFC-néerlandais : schwa en perspectives

17h15-17h30 : Juri CHERVINSKI & Elissa PUSTKA (LMU, München)

IPFC-allemand : une pré-enquête auprès de quelques étudiants munichois

17h30-17h45 : Trudel MEISENBURG & Majana GRÜTER (U. d'Osnabrück)

IPFC-allemand : aspects prosodiques

17h45-18h00 : Mito MATSUZAWA & Yuji KAWAGUCHI (Tokyo University of Foreign Studies)

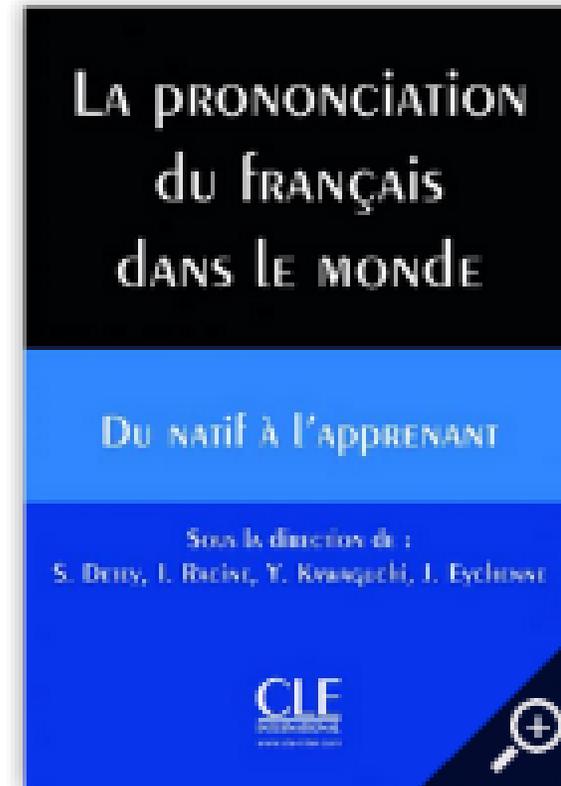
Les supports informatiques de IPFC : plateforme Moodle et site IPFC

Mercredi 8 décembre, avant les
journées PFC 9-11

Entre 2008 et 2018...

- Cadre IPFC Version envoyée en août 2018
- Equipes IPFC
- Site IPFC
- Sous-projets IPFC
- Dolmen (Julien Eychenne)
- Soutiens d'IPFC
- Colloques IPFC (& dîners y relatifs)
- Publications IPFC →
- Visibilité IPFC
- Nouvelles générations IPFC
- Redynamisation du champ de la prononciation en L2

Avec en 2016:



IPFC: objectifs et protocole

- Le protocole:
 - Répétition: Liste IPFC
 - Lecture: Liste PFC, Liste IPFC, Texte PFC
 - Interactions: avec natif, avec non-natif+ questionnaire biographique (22 questions)
+ consentement de participation

Attention, plus de passation possible via Moodle ⇨ passation manuelle

- L'approche:
 - Corpus: pas de pré-catégorisation
 - Perceptive: codage perceptif avant analyses acoustiques
 - Didactique: degré d'acceptabilité pour l'auditeur natif

Liste spécifique IPFC:

- 36 mots communs pour tous les apprenants
- ~ 30 mots avec difficultés spécifiques

Et sous-liste commune de 33 mots pour les langues germaniques (norvégien, allemand, anglais, etc.)

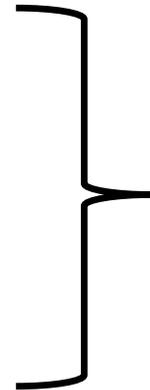
➔ Voir Detey (2012, 2014) et Racine & Detey (à paraître)

IPFC: des enregistrements, et après ?

- Transcription & codage

- Pérennisation: nécessité de vérification (transcription & codage) avant toute analyse et toute base de données

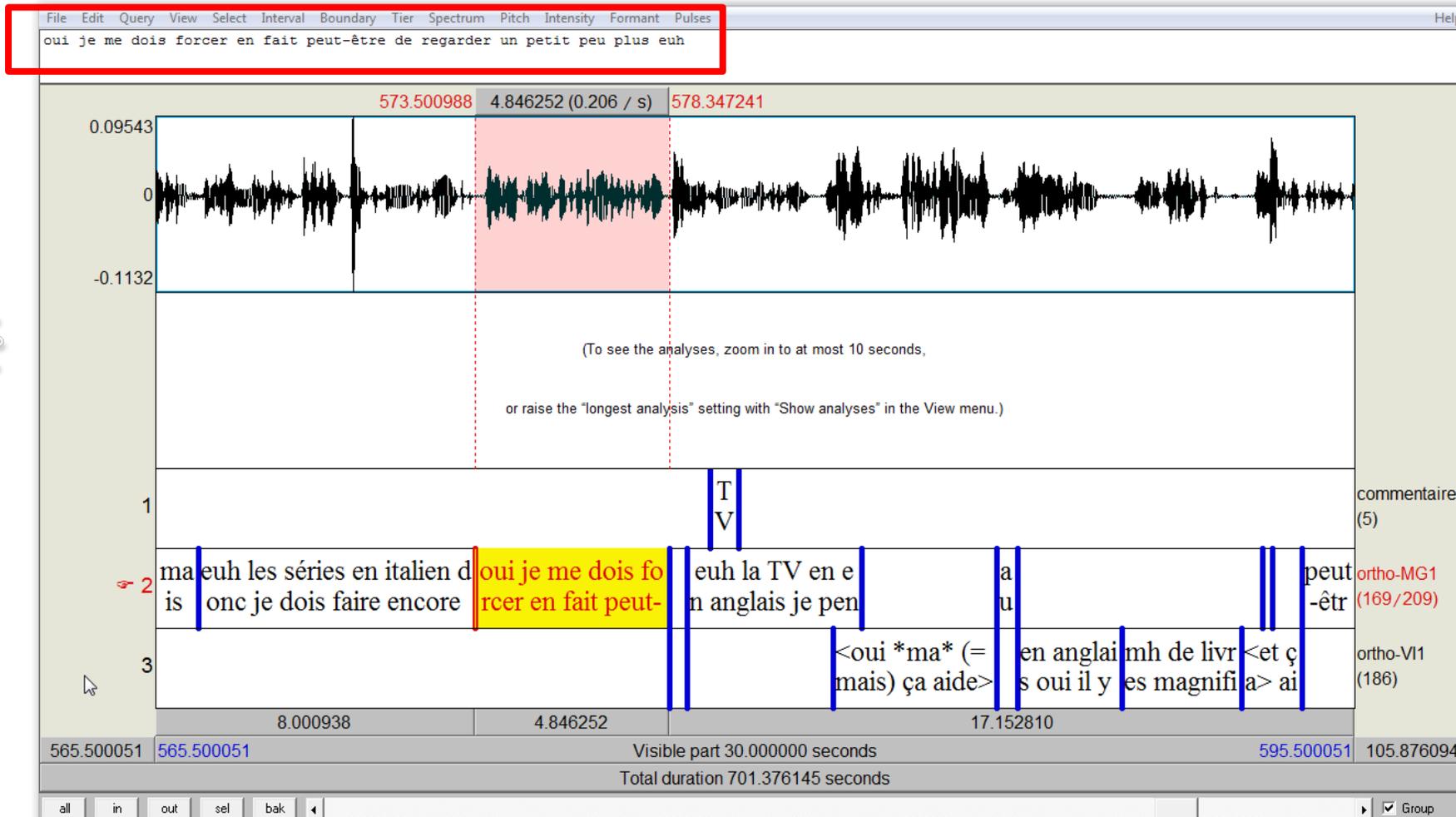
- Pourquoi une base ?
- Tension: Science ouverte vs PDP



**Voir
communication
tout à l'heure!**

IPFC: transcription et codage

- Première étape: transcription orthographique des données sous *Praat* (Boersma & Weenink, 2015) avec alignement sur le signal, selon des conventions de transcriptions adaptées pour la L2 (cf. Racine *et al.*, 2011).



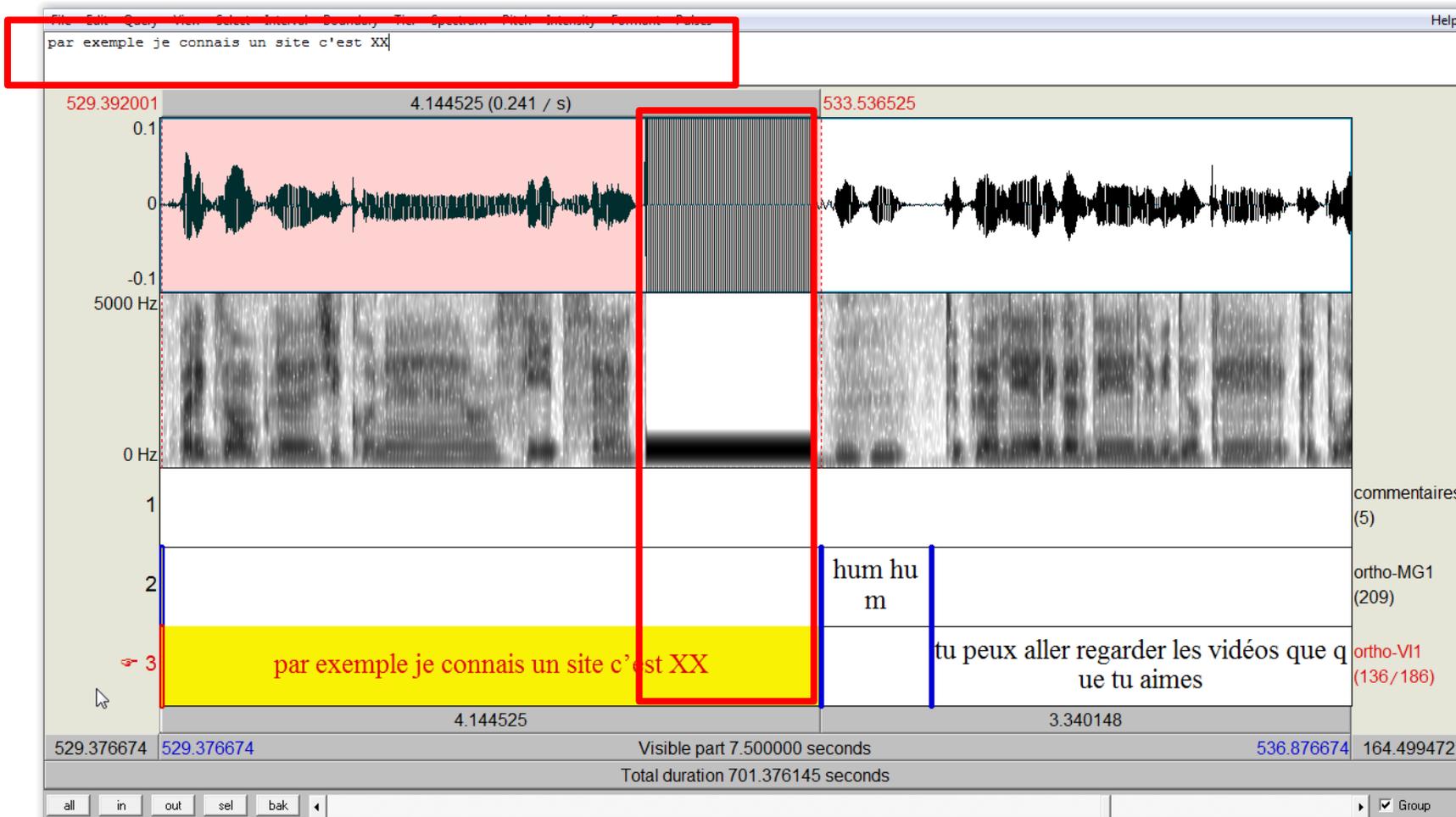
Exemple de transcription d'une conversation entre 2 apprenantes

2 tires



IPFC: transcription et codage

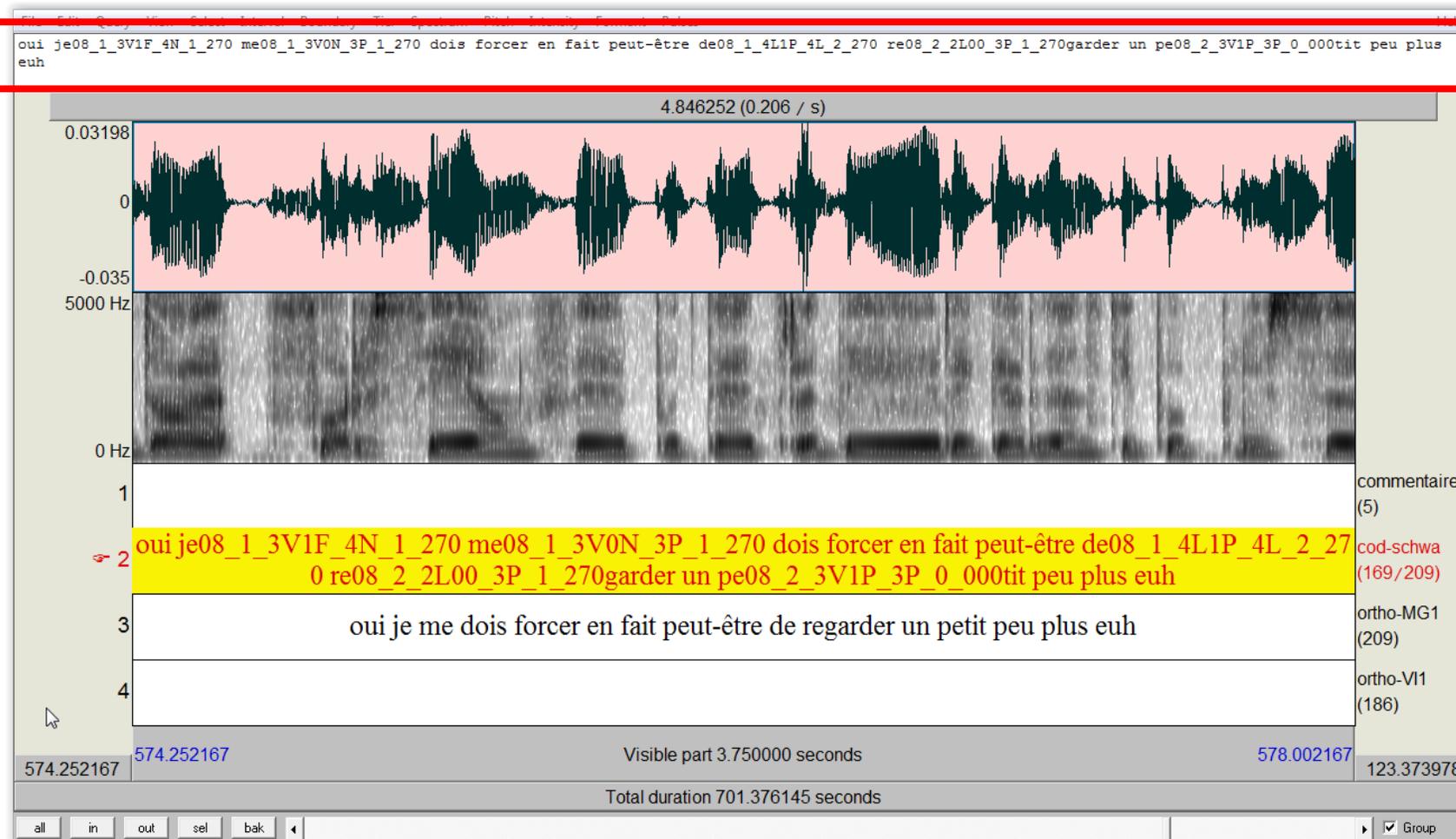
- Première étape: transcription orthographique des données sous *Praat* (Boersma & Weenink, 2015) avec alignement sur le signal, selon des conventions de transcriptions adaptées pour la L2 (cf. Racine *et al.*, 2011) avec anonymisation de certains éléments.



Exemple de transcription ANONYMISEE d'une conversation entre 2 apprenantes

IPFC: transcription et codage

- Deuxième étape: codage perceptif des données par un ou plusieurs codeurs, sous *Praat*, selon un code alphanumérique spécifique par phénomène (voyelles orales, voyelles nasales, consonnes, groupes consonantiques, liaison, schwa – *et codage prosodique en cours*)



Exemple de codage schwa de la CL d'une apprenante

1 tire cod-schwa



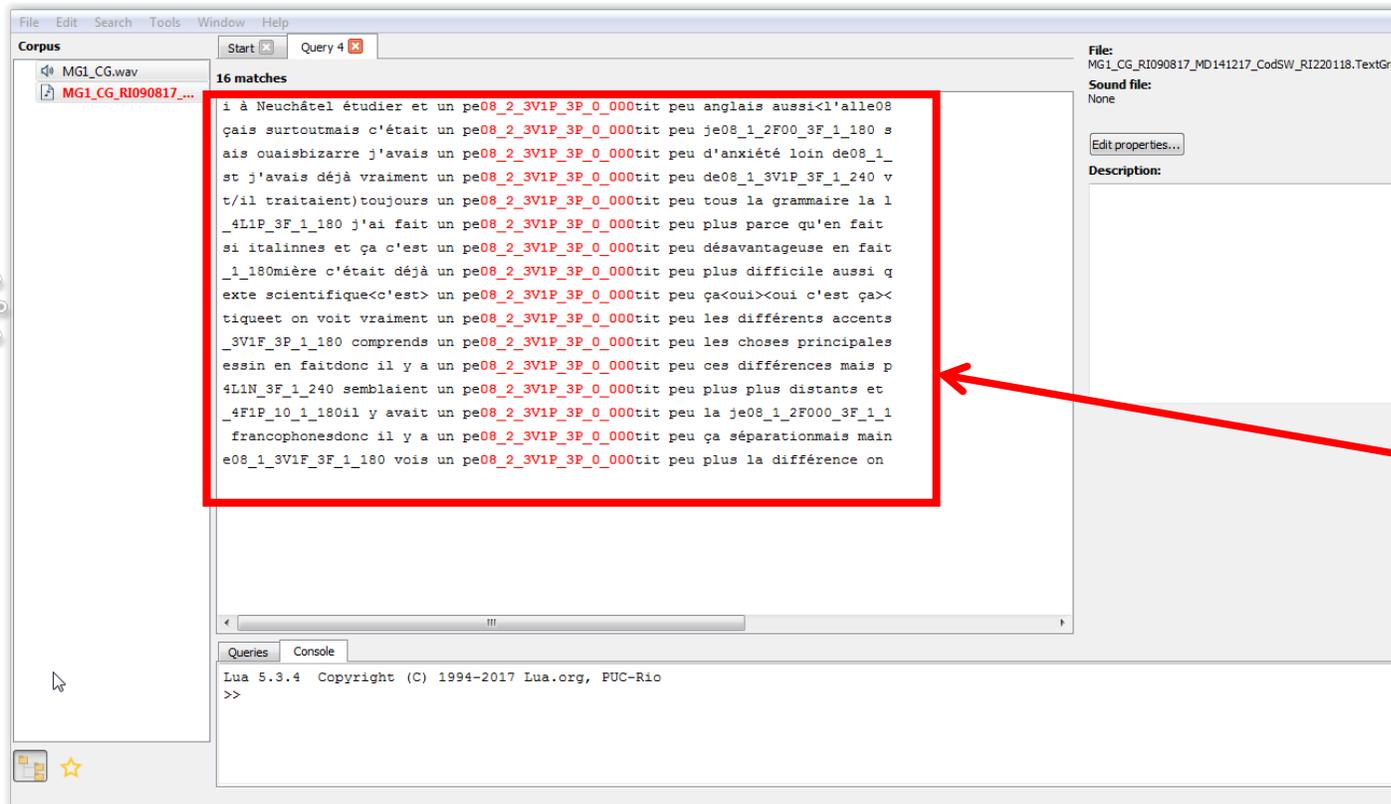
IPFC: transcription et codage

- Troisième étape: **décodage** via DOLMEN (J. Eychenne) données par un ou plusieurs codeurs, sous *Praat*, selon un code alphanumérique spécifique par phénomène (voyelles orales, voyelles nasales, consonnes, groupes consonantiques, liaison, schwa – *et codage prosodique en cours*)

Exemple d'analyse des productions de schwa en CG de cette apprenante

Résultats:

- 254 occurrences de schwa au total dans la CG
- 41 schwas en initiale de polysyllabes
- 16 chutes sur 41 = 39.02%
- Toutes dans la même séquence: «un petit peu»
- Ne chute pas ailleurs (devait, secondaire(ment), Genève, (de)vient, revient, retourner, refaire et petit)
- Parmi lesquels 12 sont réalisés avec un timbre non conforme, [e] ou [ɛ].



The screenshot shows the DOLMEN software interface. The main window displays a text transcription with phonetic codes. A red box highlights a segment of the text, and a red arrow points from the results box to this segment. The text is as follows:

```
i à Neuchâtel étudier et un pe08_2_3V1P_3P_0_000tit peu anglais aussi<l'alle08
çais surtoutmais c'était un pe08_2_3V1P_3P_0_000tit peu je08_1_2F00_3F_1_180 s
ais ouaisbizarre j'avais un pe08_2_3V1P_3P_0_000tit peu d'anxiété loin de08_1_
st j'avais déjà vraiment un pe08_2_3V1P_3P_0_000tit peu de08_1_3V1P_3F_1_240 v
t/il traitaient)toujours un pe08_2_3V1P_3P_0_000tit peu tous la grammaire la l
_4L1P_3F_1_180 j'ai fait un pe08_2_3V1P_3P_0_000tit peu plus parce qu'en fait
si italinnes et ça c'est un pe08_2_3V1P_3P_0_000tit peu désavantageuse en fait
_1_180mière c'était déjà un pe08_2_3V1P_3P_0_000tit peu plus difficile aussi q
exte scientifique<c'est> un pe08_2_3V1P_3P_0_000tit peu ça<oui><oui c'est ça>
tiqueet on voit vraiment un pe08_2_3V1P_3P_0_000tit peu les différents accents
_3V1F_3P_1_180 comprends un pe08_2_3V1P_3P_0_000tit peu les choses principales
essin en faitdonc il y a un pe08_2_3V1P_3P_0_000tit peu ces différences mais p
4L1N_3F_1_240 semblaient un pe08_2_3V1P_3P_0_000tit peu plus plus distants et
_4F1P_10_1_180il y avait un pe08_2_3V1P_3P_0_000tit peu la je08_1_2F000_3F_1_1
francophonnedonc il y a un pe08_2_3V1P_3P_0_000tit peu ça séparationmais main
e08_1_3V1F_3F_1_180 vois un pe08_2_3V1P_3P_0_000tit peu plus la différence on
```

IPFC: des enregistrements, et après ?

- Transcription & codage
- Pérennisation: nécessité de vérification (transcription & codage) avant toute analyse et toute base de données
- Pourquoi une base ?
- Tension: Science ouverte vs PDP

Développer IPFC

- PFC → PFC-EF → IPFC → IPFC-EF
- Nouvelles équipes/langues (BCMS, chinois, anglais irlandais + langues de la migration: apprenants érythréens, syriens, etc.)
- Nouvelles analyses/exploitations:
 - tester les modèles L2
 - Utiliser les codes IPFC et Dolmen
 - Développer le code
- Nouvelles tâches/modalités
- Nouveaux sous-projets
- Nouveaux étudiants
- Autres colloques (Ecole d'été « Phonologie de corpus », juillet 2019 en Suisse, NewSounds2019@Waseda...)

Dépasser IPFC

- Par-delà la phonologie: parole d'apprenant
- Par-delà la linguistique: sociolinguistique, didactique...
- Par-delà le français: IPAC, IPEC, IPJC...

Remerciements

- **Institutions:** Centre Franco-Norvégien, JSPS, FNS, Société Académique de Genève, Université de Genève
- **Personnes:**
 - Jacques, Bernard, Chantal et tous les collègues qui se sont lancés dans l'aventure avec nous, en particulier Julien Eychenne pour le développement de Dolmen-IPFC, et Kirstin!
 - Etudiants – et doctorants – qui ont été enregistrés ou enregistré, et ceux qui ont transcrit, et codé...

Pour citer IPFC

- Detey, S., Kawaguchi, Y. 2008. Interphonologie du Français Contemporain (IPFC): récolte automatisée des données et apprenants japonais. *Journées PFC: Phonologie du français contemporain: variation, interfaces, cognition*. Paris, Dec. 11-13, 2008.

ET

- Racine, I., Detey, S., Zay, F. & Kawaguchi, Y. (2012). Des atouts d'un corpus multitâches pour l'étude de la phonologie en L2 : l'exemple du projet *InterPhonologie du français contemporain* (IPFC). In Kamber, A. & Skupiens, C. (éds), *Recherches récentes en FLE*, Berne, Peter Lang, 1-19.