



WASEDA University



東京外国語大学 Tokyo University of Foreign Studies

Corpus d'apprenants, interlangue et français parlé: éclairage à travers le projet IPFC

Sylvain DETEY¹ & Yuji KAWAGUCHI²

¹SILS, Waseda University (Japan) & E.A. 4305 LiDiFra, Université de Rouen (France)

²Tokyo University of Foreign Studies (Japan)

Plan

- 1) Introduction
- 2) Corpus oraux et français parlé
- 3) Corpus d'apprenants et interlangue
- 4) Interphonologie: le projet IPFC
- 5) Illustration japonaise: le site IPFC
- 6) IPFC2011-Tokyo

1) Introduction

● **Objectif de cette journée d'étude:**

Corpus d'apprenants, interlangue et français parlé

Corpus d'apprenants:

- Majorité de corpus écrits
- Encore peu de corpus oraux en français (vs. anglais)
- Peu de corpus "internationaux" (plusieurs L1)

Interlangue:

- En FLE: corpus ESF, InterFra, FLLOC
- Peu de corpus dédiés à la phonétique-phonologie
- Peu de lien entre phonologie de corpus et psycholinguistique en L2

Français parlé:

- Une dynamique constante depuis les 70's (GARS) mais un objet qui reste à appréhender hors du cercle des spécialistes

Géographie/Typologie:

- Langues européennes

→ Corpus, d'apprenant, interlangue et FLE en zone Asie?

2) Corpus oraux et français parlé

70 s GARS (*Groupe Aixois de Recherche en Syntaxe*)

90 s France en retard (*Corpus de Portugêês Fundamental, British National Corpus*)

2005 Inventaire des corpus oraux (DGLFLF)

- *Corpus de Référence du Français Parlé* (CRFP, Provence)
- *Corpus de Langues Parlées en Interaction* (CLAPI, Lyon II).
- ...

2007 Site internet officiel *Corpus de la parole*

En cours:

- *Corpus International et Ecologique de la Langue Française* (CIEL-F)
- *Français Contemporain en Afrique et dans l'Océan Indien* (CFA)

2) Corpus oraux et français parlé

Les dimensions non-phonologiques:

Le corpus d'Aix-Marseille et de Paris de TUFs

- **Collecte du français parlé à Aix-en-Provence et Paris effectuée dans les années 2005, 2009, 2010 environ 300,000 mots**



The screenshot shows the website for the French (Aix) Multilingual Spoken Corpus. The header features a landscape image with the text "French (Aix) Multilingual Spoken Corpus". A navigation menu on the left includes links for Top, Outline, Research Team, References, Corpus (highlighted), Related Sites, and Pictures. The main content area, titled "Corpus", lists four categories:

1. Student affairs section
2. Reception of the university
3. About faculty
4. About health condition

Collaboration: TUFs (Global COE – 21st Century) & Université Aix-Marseille équipe (ex-) DELIC (C. Blanche-Benveniste, J. Delofeu, F. Sabio...)

http://www.coelang.tufs.ac.jp/multilingual_corpus/fr/index.html?contents_xml=corpus&menulang=en



- Top
- Outline
- Research Team
- References
- Corpus**
- Related Sites
- Pictures



- French
- Japanese
- English

- Scroll On/Off

About faculty

French; moi j'aime beaucoup l'Italie +
Japanese; 私はイタリアが好き

French; L1 mh mh +
Japanese; うん

French; L2 mais bon c'est enfin je je sais pas si je pourrais aller vivre ailleurs comme ça +
Japanese; でも、なんていうか、他の場所で生活できるかどうかわからないけど

French; bon déjà j'ai beaucoup d'attaches ici bon tout ce qui est famille amis tout ça +
Japanese; もうここにたくさんのつながりができたし、家族も友達もここにいるし

2) Corpus oraux et français parlé

Les dimensions non-phonologiques:

- **Moteur de recherche selon les parties du discours**
(étiquetage avec Tree Tagger – collaboration avec l'équipe d'Aix)



■ What's New
• (218366 words)

■ Outline & Purpose
CbLLE POS Research Engine (Spoken French) is to search part-of-speech tags of the *Multilingual Spoken Language Corpora (French)* developed by the 21st Century COE Program "Usage-Based Linguistic Informatics". Part-of-speech (POS) tagging is based on the system called *Tree Tagger*, cf. <http://www.ims.uni-stuttgart.de/projekte/corplex/TreeTagger/>. The purpose of this research engine is largely to publish the research results from the Global COE Program "Corpus-based Linguistics and Language Education (CbLLE) and to analyze various grammatical phenomena in spoken French.

■ Corpus Description
The current corpora of Spoken French are composed of two different datas.
1. Twenty dialogs recorded during July 2005 with the cooperation of the University of Aix-Marseille I. The total length is around seven hours and the total number of word tokens exceeds 100.000. They cover various subjects such as office life, work at the university, recent news about one's family and improvised sketches.

References
http://www.coelang.tufs.ac.jp/multilingual_corpus/fr/index.php?contents_xml=gaisetsu&lang=ja
http://www.coelang.tufs.ac.jp/multilingual_corpus/fr2/index.php?contents_xml=gaisetsu&lang=ja

<http://cblle.tufs.ac.jp/tag/fr/index.php?menulang=en>

- **Confrontation des corpus du français parlé et des apprenants japonophones aux niveaux lexicaux et morpho-syntaxiques**

Ex: adjectifs antéposés

The screenshot shows a web-based interface for analyzing French text. At the top, there are search filters for 'pos' (part of speech) with dropdown menus for 'DET:ART', 'ADJ', and 'NOM'. There are also input fields for 'lemma' and 'token', a 'Search' button, and a 'Reset' button. On the right, there are checkboxes for 'paris' and 'aix', and a 'HELP' button with a 'search sample' link. Below the filters, there is a list of text examples, each with a line number and a snippet of text. In these snippets, certain adjectives are highlighted in red, indicating they are the focus of the analysis. The examples include phrases like 'un gros bruit', 'les même problèmes', 'le vrai bruit', 'un bon rythme', 'une nouvelle forme', 'le même truc', 'le même genre', 'la même musique', 'des jolies voix', 'des supers chansons', 'une belle voix', 'une autre génération', 'un nouveau style', 'des petits gâteaux', 'un petit gâteau', 'un bon public', 'un bon exemple', and 'une certaine façon'.

pos DET:ART ADJ NOM Search 0 0 ja HELP
lemma
token
 Ignore the second context

46. 01Kathy.txt:2425 mais c' est venu comme ça du jour au lendemain mais pas pour **un gros bruit** ou quoi que ce soit
47. 01Kathy.txt:2466 et c' est ce qu' il se dit il disait il a même pensé à devenir sourd et tout mais il a quelqu'un sur internet qui disait qu' il avait **les même problèmes** de de machin
48. 01Kathy.txt:2509 tu vois à la limite je préfère **le vrai bruit** de tu sais y avait un concert une fois mais je crois que c' était heu en Angleterre ou quelque chose comme ça ils donnaient des concerts où ils tapaient sur des boute des
poubelles heu
49. 01Kathy.txt:2530 mais ils sont peut-être quinze hein genre heu ils sont habillés en éboueurs ou en truc comme ça avec du balai ils font des mais ça fait un **un bon rythme** tu vois
50. 01Kathy.txt:2534 c' est **une nouvelle forme** de musique
51. 01Kathy.txt:2535 ouais exactement c' est **une nouvelle forme** de musique quoi ils inventent plein des trucs hein
52. 01Kathy.txt:2540 non mais y a plein de de de styles de musique si tu regardes ça ça reste dans **le même truc** tu vois pourquoi regarde pourquoi j' aime pas la la Dancehall et tout parce que ils prennent ils appellent ça des Riddim
ils prennent **le même son** et les gens en fait ils changent juste de paroles donc si tu veux tu as **le même genre** de musique c' est la même exactement **la même musique** et derrière
53. 01Kathy.txt:2543 toujours un peu **le même rythme** quoi
54. 01Kathy.txt:2544 non mais non le zouk c' est **le même rythme** mais c' est pas la même dans la
55. 01Kathy.txt:2545 **la même musique**
56. 01Kathy.txt:2546 c' est vraiment pas **la même musique** d' une à l' autre tu vois vraiment
57. 01Kathy.txt:2548 mais la Dancehall tu vois vraiment c' est **la même musique** hein
58. 01Kathy.txt:2559 c' est toujours **le même rythme** ouais voilà c' est un peu chiant quoi
59. 01Kathy.txt:2564 ça me vraiment c' est un truc surtout qui enfin je sais pas pour moi enfin j' y tu sais vu que je connais quand même pas mal de morceaux vraiment y a une différence à chaque fois y a **la petite pointe** qui fait la
différence
60. 01Kathy.txt:2567 souvent c' est **des jolies voix** et tout quoi
61. 01Kathy.txt:2626 mais Wallen en fait le problème de Wallen c' est qu' elle est pas vachement commercialisée je veux dire tu la vois pas forcément beaucoup à la télé mais elle a fait franchement **des supers chansons**
62. 01Kathy.txt:2657 elle a **une belle voix** c' est vrai franchement dommage que tu connaites pas tu connais pas Wallen non
63. 01Kathy.txt:2658 non pas du tout pas du tout je crois que je j' appartiens à **une autre génération** (rire)
64. 01Kathy.txt:2704 ahh c' est ça m' énerve ça c' est c' était **un nouveau style** le reggeaton
65. 01Kathy.txt:2869 ouais c' est **un nouvel artiste** on va dire enfin
66. 01Kathy.txt:2875 Raphaël tu veux **des petits gâteaux** là
67. 01Kathy.txt:2882 je vais prendre **un petit gâteau**
68. 01Kathy.txt:2939 c' est **la nouvelle forme** de de fans quoi enfin les nouveaux heu
69. 01Kathy.txt:2943 mes petites cousines à fond sur les elles achètent les heu pis non ce qui marche en fait où ils ont **un bon public** quand ils ont des publics des petits c' est que les petits ils veulent pas les CD gravés ils veulent le
CD le vrai
70. 01Kathy.txt:2947 tu peux pas leur pourtant t' as beau leur dire c' est **la même chose**
71. 01Kathy.txt:2952 ouais moi aussi Mickael Jackson à la limite le personnage iii quoi que si il danse bien quand même ouais c' est pas **un bon exemple** Mickael Jackson
72. 01Kathy.txt:2994 ouais ouais c' était lui rendre hommage d' **une certaine façon**
73. 01Kathy.txt:3011 t' achètes le single c' est bon elles le mettent en boucle en boucle mais si tu remarque moi j' ai remarqué **le nouveau style** là en ce moment t' as pas remarqué c' est heu les chansons avec les clips où où ce sont
des hebdomadaires qui en a encore un qui est certifié Papa pisserie là («le petit le petit») («le papa le papa le papa pisserie le papa le papa le papa pisserie») («le chantant») (rire)

→ possibilités d'analyses syntaxiques

2) Corpus oraux et français parlé

→ *La phonologie... et plus: le corpus PFC*

Le projet PFC: un corpus de référence, par-delà la phonologie

- Projet fédératif décentralisé coordonné par:
 - Jacques Durand (ERSS, Université de Toulouse-Le Mirail)
 - Bernard Laks (MODYCO, Université de Paris X-Nanterre)
 - Chantal Lyche (Universités d'Oslo et de Tromsø)
- Objectifs initiaux:
 - vaste corpus de français parlé à travers le monde
 - protocole d'enquête uniforme
 - méthodes d'analyse et à des outils communs
 - décrire la prononciation du français dans sa diversité géographique, sociale et stylistique
 - vision globale de la phonologie du français

- Exploitation au-delà de la phonétique et de la phonologie:
le français oral
- Syntaxe, particules énonciatives, tours de parole, variation lexicale, français familier...
- Pour l'enseignement du français et de la linguistique française (2010):



Consulter le site: www.projet-pfc.net

PARUTION
« Les variétés du français parlé dans l'espace francophone : ressources pour l'enseignement » (avec DVD) (S. Detey, J. Durand, B. Laks & C. Lyche eds) dans la collection L'Essentiel français, aux Editions Ophrys [Lire la présentation](#)

RECHERCHE
Si vous êtes chercheur, universitaire ou étudiant en linguistique française. Utilisez la base de données schwa, liaison et prosodie pour vos recherches.

ENSEIGNEMENT DU FRANÇAIS
Vous êtes enseignant de français (FLE/FLS/FLM), concepteur pédagogique ou étudiant de français? Allez au module PFC Enseignement du Français.

PRÉSENTATION DU FRANÇAIS
Vous vous intéressez au français oral contemporain dans l'espace francophone en général ? Cette rubrique est pour vous...

ACTUALITÉ
Prochain rendez-vous PFC: [Appel à communications Journée PFC 2011 London, Ontario](#)
Ecoutez le texte PFC en ch'ti [ici](#)

Base de données sur le français oral contemporain dans l'espace francophone
Le projet international PFC (Phonologie du Français Contemporain), codirigé par Jacques Durand (ERSS, Université de Toulouse-Le Mirail), Bernard Laks (MoDyCo, Université de Paris X) et Chantal Lyche (Universités d'Oslo et de Tromsø) s'adresse à un triple public, susceptible de s'intéresser au français oral dans ses usages attestés et dans sa variation au sein de l'espace francophone : chercheurs, enseignants/apprenants de français et grand public.
Il offre une base de données qui, à terme, constituera la plus grosse base de données orales portant sur le français et l'une des plus grosses bases toutes langues confondues. Cette base peut être utilisée dans le cadre de la recherche (phonétique, phonologie, syntaxe, pragmatique, sociolinguistique, analyse conversationnelle, etc.), de l'enseignement / apprentissage du français (langue étrangère, maternelle ou seconde) et de la diffusion des savoirs

Et les nombreuses références:

Phonologie, variation et accents du français,
Durand, Laks, Lyche (eds) (2009) Hermès

3) Corpus d'apprenants et interlangue

● **Corpus oraux en L2:**

- Synthèse de S. Rule (JFLS, 2004)
- Répertoire des corpus d'apprenants (écrits et oraux) du *Center for English Corpus Linguistics* de l'Université Catholique de Louvain (S. Granger & D. Goosens)

The screenshot shows the website of the Centre for English Corpus Linguistics at the University of Louvain. The page title is "Learner corpora around the world". Below the title, there is a paragraph explaining that the list is comprehensive and invites researchers to submit new corpora. A table follows, listing four corpora with columns for Corpus, Target language, First language, Medium, Text type/ task type, Proficiency level, Size in words, and Project director.

Corpus	Target language	First language	Medium	Text type/ task type	Proficiency level	Size in words	Project director	Availability
The Pilot Arabic Learner Corpus	Arabic	English	written		Intermediate and advanced	c. 9,000	Ghazi Abuhakema Reem Faraj Anna Feldman Eileen Fitzpatrick Montclair State University, USA	
Leerdercorpus Nederlands als Vreemde Taal	Dutch	French	written				Liesbeth Degand Université catholique de Louvain, Belgium	
Asao Kojiro's Learner Corpus Data	English	Japanese	written	Essays and stories written or reproduced by Japanese college students.			Asao Kojiro	Te
The Barcelona English Language Corpus (BELC)	English	Spanish Catalan	spoken and written	4 tasks: Written composition Oral narrative Oral interview Role-play Longitudinal data (children and young adults learning English)			Carmen Muñoz University of Barcelona, Spain	Sp av

● 103 corpus

(dont 64 corpus exclusivement consacrés à l'apprentissage de l'anglais)

Corpus de français langue étrangère :

The <i>Chy-FLE</i> (Cypriot Learner Corpus of French)	French	Modern Greek (and Cypriot Greek)	written	Argumentative and descriptive essays	From intermediate to advanced	c. 250,000 (under development)	Freiderikos Valetopoulos Université de Poitiers, France In collaboration with the University of Cyprus
The COREIL corpus PDF	French English		spoken				Elisabeth Delais-Roussarie Hiyon Yoo Université Paris-Diderot, France
The "Dire Autrement" corpus	French (Second Language)	Mainly L1 speakers of English	written	Narrative, injunctive, persuasive and informative texts		48,114	Marie-Josée Hamel Jasmina Milicevic Dalhousie University, Canada
French Interlanguage Database (FRIDA)	French	various	written				Sylviane Granger Centre for English Corpus Linguistics Université catholique de Louvain, Belgium
French Learner Language Oral Corpora (FLLOC)	French	various	spoken	See description of the 7 corpora	various		Florence Myles Newcastle University Rosamund Mitchell University of Southampton, UK
The <i>Interfra</i> corpus	French	Swedish	spoken	Interviews, retellings of video clips and picture stories	various		Inge Bartning Stockholm University, Sweden. interfra@fraitu.su.se
The "Interphonologie du Français Contemporain" (IPFC) corpus	French	Japanese Spanish ?	spoken	Reading aloud, repeating words, guided interviews, interactions between two learners.	various	under development	Sylvain Detey Waseda University, Japan Université de Rouen, France Isabelle Racine Université de Genève, Switzerland Yuji Kawaguchi Tokyo University of Foreign Studies, Japan
The LCF corpus (Learner Corpus French)	French	Dutch	written	Argumentative essays Informative texts Journalistic texts Formal letters Summaries Written compositions by Flemish students of French	From intermediate to advanced	490,000	K.U.Leuven Campus Kortrijk, UGent and Lessius Hans Paulussen
The Lund CEFLÉ Corpus (Corpus Écrit de Français Langue Étrangère)	French	Swedish	written	Descriptive and narrative essays; picture-based stories.	various	100,000	Malin Ågren Lund University, Sweden
The UWi (University of the West Indies) learner corpus PDF	French	English and Jamaican Creole	spoken	Conversations during oral exams and in informal contexts	various		Hugues Peters University of New South Wales, Sydney, Australia

● Français:

- 10 corpus
- 5 corpus oraux :
 - FLLOC (R-U),
 - Interfra (Suède),
 - COREIL (France),
 - UWI (Jamaïque) et...
 - IPFC (International).

(autres corpus construits (LANCOM) ou en cours (Pillot-Loiseau et al), etc.)

● *Exploitation des corpus en SLA:*

- *Klein et Perdue (80's) dans le projet ESF*
- *S. Granger au CECL (Université de Louvain)*
- *Exemple: les corpus suédois (InterFra et Lund) et les itinéraires et stades acquisitionnels (morphologie verbale, pronoms, etc.) (Bartning & Schlyter 2004)*
- *revue de Myles (2005) dans Second Language Research*
- *Etc.*

→ *Peu de travaux en phonétique-phonologie du FLE*

4) Interphonologie: le projet IPFC

● Travaux sur corpus en phonétique / phonologie en L2?

- Néerlandais (Neri, Cucchiarini & Strik, 2006), polonais (Cylwik, Wagner & Demenko, 2009), allemand (Gut, 2009), anglais (Gut, 2009; Visceglia, Tseng, Kondo, Meng & Sagisaka, 2009),
- Aspects segmentaux et suprasegmentaux (Trouvain & Gut, 2007 ; Meng, Tseng, Kondo, Harrison & Visceglia, 2009)

⇒ Absence de données similaires pour le français...

→ **IPFC:**

InterPhonologie du Français Contemporain:

usages, variétés, structures

Historique:

Début en 2008: IPFC-japon

Travaux en cours: espagnol, norvégien, néerlandais, allemand, grec chypriote, anglais (Canada), italien...

Protocole: adapté de PFC

quasi-identique pour toutes les L1s ⇒ données comparables

Protocole: comprend 6 tâches

- Répétition d'une liste de mots spécifique à la L1
- Lecture de la liste de mots PFC
- Lecture de la liste de mots spécifique
- Lecture du texte PFC
- Entretien guidé avec un natif
- Interaction semi-contrainte entre deux apprenants

•**Récolte des données:** *Manuelle* ou *Semi-automatique* (Moodle pour 4 tâches)

•**Transcription des données:** sous *Praat* (Boersma & Weenink, 2009)

Pour plus de détails: → voir *Cadre IPFC*

5) Illustration japonaise: le site IPFC

- Les premiers corpus d'IPFC:
 - IPFC-japonais (Kawaguchi & Detey, TUFS)
 - IPFC-espagnol (Racine, U. Genève)
- IPFC-japonais:
 - Quelques difficultés typiques des apprenants japonais

Français	Japonais
/r/ ([ʀ ʁ]) - /l/	/r/ ([ɾ])
/b/ - /v/	/b/ ([b, β])
/f/	/h/ ([h, ɸ])
/y, u, ø, œ/	/u/ ([ɯ])
Voyelles nasales /ã, õ/	Pas de voyelles nasales
Groupes consonantiques (OBLI, sC, CCCC)	Pas de groupes consonantiques

- Taille: 100 pour répétition & lecture; 12 pour protocole complet = 120 heures

● Le site IPFC: <http://cblle.tufs.ac.jp/ipfc>



ACCUEIL

PFC RECHERCHE
base de données en ligne

PFC ENSEIGNEMENT
le français expliqué

PFC PRÉSENTATION
le français oral dans le monde

IPFC

PFC PUBLICATIONS
bulletins, colloques, logiciels

Année de naissance Études de français (en mois) Séjours en pays francophone (en mois) Sexe

No	UserId	Année de naissance	Lieu de naissance	Sexe	Nationalité	Langue maternelle	Spécialité	Études de français (en mois)	Certificat	Séjours en pays francophone (en mois)	Séjours en pays étrangers non-francophone (en année)
----	--------	--------------------	-------------------	------	-------------	-------------------	------------	------------------------------	------------	---------------------------------------	--

<input type="checkbox"/>	1	275	1986	Yamanashi	F	japonaise	japonais	linguistique	60	0	12	0
--------------------------	---	-----	------	-----------	---	-----------	----------	--------------	----	---	----	---

LECTURE LISTE GÉNÉRIQUE			RÉPÉTITION LISTE SPÉCIFIQUE			LECTURE LISTE SPÉCIFIQUE		
<input type="button" value="♪"/> compagnie	<input type="button" value="♪"/> médecin	<input type="button" value="♪"/> baignoire	<input type="button" value="♪"/> peur	<input type="button" value="♪"/> moule	<input type="button" value="♪"/> la vase	<input type="button" value="♪"/> peur	<input type="button" value="♪"/> moule	<input type="button" value="♪"/> la vase
<input type="button" value="all♪"/>			<input type="button" value="all♪"/>			<input type="button" value="all♪"/>		
LECTURE TEXTE								
<input type="button" value="♪"/> Le Premier Ministre ira-t-il à Beaulieu? « Le village de Beaulieu est en grand émoi. Le Premier Ministre a en effet décidé de faire étape dans cette commune au cours de sa tournée de la région en fin d'année.								

<input type="checkbox"/>	2	271	1988	Osaka	F	japonaise	japonais	linguistique	96	APEF Jun1	1	6
<input type="checkbox"/>	3	267	1989	Okinawa	M	japonaise	japonais	linguistique	24	APEF 3	0	0
<input type="checkbox"/>	4	266	1943	Shizuoka	M	japonaise	japonais	linguistique	48	APEF Jun1	6	0
<input type="checkbox"/>	5	263	1986	Fukuoka	F	japonaise	japonais	linguistique	48	APEF 2	10	0

Problèmes:

● Liste PFC

- *Compagnie*: voyelle nasale
- *Médecin*: aperture du /e/, traitement du schwa, voyelle nasale
- *Baignoire*: aperture du /e/, C+w, /R/ en coda pré-pausale

(→ problèmes de décodage grapho-phonémique)

● Liste IPFC-japonais (répétition et lecture)

- *Peur*: aperture vocalique, /R/ en coda pré-pausale
- *Moule*: aperture vocalique, /l/ en coda pré-pausale
- *La vase*: /v/, /z/ en coda

● Texte PFC

- *Le Premier Ministre ira-t-il à Beaulieu? Le village de Beaulieu est en grand émoi.*

Perspectives

● Développement

- Technique et méthodologique
- Protocole
- Corpus

● Exploitation

- Comparaisons (inter-tâches, -L1, natif/n-natifs)
- Analyses
- Pédagogique (formation d'enseignants, intercompréhension, etc.)

→ IPFC: projet d'actualité

- Projets apparentés en anglais: Gut (2009); *Wildcat Corpus of Native- and Foreign-accented English* (Van Engen et al. 2010), *Rated L2 speech corpus* (Yoon et al. 2009)
- Projets émergents: en France (Paris 7, Paris 3, ANR) et en Suisse (FNRS)
- Visibilité au plan international: *NewSounds2010*, *CMLF2010*, ...
- Réseau IPFC en extension (italien, mandarin, danois...)

6) IPFC2011-Tokyo

IPFC2010-Paris: interphonologie, corpus et FLE

- Première rencontre fédératrice de ce type en France
- IPFC-norvégien, anglais canadien, grec chypriote, néerlandais, allemand (10 pays, 15 universités, 25 participants...)

IPFC2011-Tokyo: corpus d'apprenants et français parlé

- Première rencontre en Asie
 - *IPFC-japonais, IPFC-espagnol et accentuation*
 - *Les apprenants coréens et la prononciation du français*
 - *La liaison et son enseignement chez les apprenants japonais*
 - *Caractéristiques du français parlé: production lexicales et vocaliques par les apprenants japonais, catégorie et distribution des accents sur corpus*

IPFC2011-Paris: 6-7 décembre 2011 (& PFC-EF)

- ***IPFC: Voyelles nasales, voyelles antérieures arrondies, liquides et liaison : traitement interphonologique en français langue étrangère ou seconde.***
- ***PFC-EF: Corpus oraux et enseignement : références, outils, pratiques.***

**Merci de votre
attention!**

Quelques références IPFC

- Detey, S. (2009). Phonetic input, phonological categories and orthographic representations: a psycholinguistic perspective on why oral language education needs oral corpora. The case of French-Japanese interphonology development. In Kawaguchi, Y., Minegishi, M. & Durand, J. (eds), *Corpus Analysis and Variation in Linguistics*. Amsterdam/Philadelphia: John Benjamins, 179- 200.
- Detey, S., Durand, J. & Nespoulous, J.-L. (2005). Interphonologie et représentations orthographiques. Le cas des catégories /b/ et /v/ chez des apprenants japonais de FLE. *Revue PArôle* 34/35/36 (supplément), 139-186.
- Detey, S. & Kawaguchi, Y. (2008). Interphonologie du français contemporain (IPFC) : récolte automatisée des données et apprenants japonais. *Phonologie du français contemporain : variation, interfaces, cognition*. Paris.
- Detey, S. & Nespoulous, J.-L. (2008). Can orthography influence L2 syllabic segmentation? Japanese epenthetic vowels and French consonantal clusters. *Lingua* 118 (1), 66-81.
- Detey, S., Racine, I., Kawaguchi, Y., Zay, F., Bühler, N., Schwab, S. (2010). Evaluation des voyelles nasales en français L2 en production : de la nécessité d'un corpus multitâches. In Neveu, F., Muni-Toké V., Durand, J., Klingler, T., Mondada, L. et Prévost S. (éds.), Actes de CMLF'10, « Phonétique, phonologie et interfaces », Paris : ILF, 1289-1301.
- Racine, I., Detey, S., Zay, F. & Kawaguchi, Y. (2009). Interphonologie du français contemporain : réflexions méthodologiques et premières données d'apprenants hispanophones et japonophones. *AFLS 2009, Langue française en contextes*, Université de Neuchâtel (Suisse), Sept. 2009.
- Racine, I., Detey, S., Bühler, N., Schwab, S., Zay, F., Kawaguchi, Y. (2010). The production of French nasal vowels by advanced Japanese and Spanish learners of French : a corpus-based evaluation study. In Dziubalska-Kolaczyk, K, Wrembel, M. & Kul, M. (éds.), Proceedings of New Sounds 2010, Poznan: Adam Mickiewicz University, 367-372.
- Racine, I., Detey, S., Zay, F. & Y. Kawaguchi (à paraître). Des atouts d'un corpus multitâches pour l'étude de la phonologie en L2 : l'exemple du projet « Interphonologie du français contemporain » (IPFC). In A. Kamber et C. Skupiens (éds). *Recherches récentes en FLE*. Berne : Peter Lang.
- Racine, I., Zay, F., Detey, S. & Kawaguchi, Y. (à paraître), « De la transcription de corpus à l'analyse interphonologique: enjeux méthodologiques en FLE ». In *Travaux Linguistiques du CerLiCO* 24, Rennes: PUR.

Références présentation

- Répertoire des corpus oraux en L2: Granger, S. & Goosens, D. <http://www.uclouvain.be/en-cecl-icWorld.html>
- Bartning, I. & Schlyter, S. (2004). Itinéraires acquisitionnels et stades de développement en français L2. In F. Myles & R. Towell (ed) "The acquisition of French as a Second Language", *Journal of French Language Studies*, 14, 281-299.
- Boersma, P. & Weenink, D. (2009). *Praat : doing phonetics by computer* (version 5.0), <http://www.praat.org>.
- Cylwik, N., Wagner, A., Demenko, G. 2009. The EURONOUNCE corpus of non-native Polish for ASR-based Pronunciation Tutoring System. *Proceedings of SlaTE 2009 – 2009 ISCA Workshop on Speech and Language Technology in Education*. Birmingham, UK.
- Detey, S., Durand, J., Laks, B. & Lyche, C. (éds). (en préparation). *Varieties of Spoken French : a source book*. Oxford : Oxford University Press.
- Gut, U. (2009). *Non-native Speech: A Corpus-based Analysis of Phonological and Phonetic Properties of L2 English and German*. Wien: Peter Lang.
- Meng, , Tseng, , Kondo, , Harrison, & Viscelgia, (2009). Studying L2 suprasegmental features in Asian Englishes: a position paper. *Proceedings of Interspeech 2009*, Brighton, R-U.
- Myles, F. (2005). Interlanguage corpora and second language acquisition research. *Second Language Research*, 21(4), 373-391.
- Neri, A., Cucchiaroni, C. & Strik, H. (2006). Selecting segmental errors in non-native Dutch for optimal pronunciation training. *IRAL - International Review of Applied Linguistics in Language Teaching*, 44, 357-404.
- Rule, S. (2004). French interlanguage oral corpora: recent developments. In F. Myles & R. Towell (ed) "The acquisition of French as a Second Language", *Journal of French Language Studies*, 14, 342-356.
- Trouvain, J. & Gut, U. (eds) (2007). *Non-Native Prosody. Phonetic Description and Teaching Practice*. Berlin/New York: Mouton de Gruyter.
- Visceglia, Tseng, Kondo, Meng & Sagisaka (2009). Phonetic aspects of content design in AESOP (Asian English Speech cOrpus Project). *Proceedings of Oriental-COCOSDA*, Urumuqi, Chine.
- Van Engen K., Baese-Berk M., Baker R.E., Choi A., Kim M., Bradlow A.R. (2010). The *Wildcat Corpus* of Native- and Foreign-Accented English: communicative efficiency across conversational dyads with varying language alignment profiles. *Language and Speech*. published online Sept. 1st 2010.
- Yoon S.-Y., Pierce L., Huensch A., Juul E., Perkins S., Sproat R. & Hasegawa-Johnson M. (2009). Construction of a rated speech corpus of L2 learners? speech. *CALICO Journal* 26(3): 662-673.