

# ***Corpus-based French linguistics in Japan: TUFS Corpus, PFC Corpus, PFC-EF, IPFC Corpus***

<sup>1</sup>Nori Kondo, <sup>2</sup>Sylvain Detey,  
<sup>3</sup>Isabelle Racine, <sup>1</sup>Yuji Kawaguchi

<sup>1</sup> Tokyo University of Foreign Studies, <sup>2</sup> SILS, Waseda University, <sup>3</sup> ELCF, Université de Genève

# Outline

- 1) Introduction
- 2) The TUFs French native corpus and POS
- 3) The PFC French native corpus and phonology
- 4) Language education: PFC-EF & Language Modules
- 5) Non-native phonology: IPFC
- 6) Conclusion

# 1) Introduction

- Corpus-based Linguistics: written vs oral  
e.g. *British National Corpus* : written and oral
- French oral corpora?
- In France: promotion by the DGLFLF (Délégation Générale à la Langue Française et aux Langues de France) – French Ministry of Culture and Communication
- Official inventory of French oral corpora and official website 'Corpus de la parole' (Speech Corpora)

# 1) Introduction

- French Linguistics in Japan: not so much based on corpora, and even less so on *oral* corpora
- One case: the research programmes at TUFs (Global COE – 21st Century & CBLLE) within a multilingual research framework.
- Two partnerships: one on POS with Aix-Marseille University; one on phonology (and beyond) with the international PFC project

## 2) The TUFS French native corpus and POS

***2000 : partnership between TUFS (Global COE – 21st Century) & University of Aix-Marseille – DELIC team (C. Blanche-Benveniste, J. Delulofeu, F. Sabio... Pioneering corpus-based work on spoken French)***

- ***Collection of French spoken in Aix-en-Provence (Southern variety of French) and Paris ('standard' French) in 2005, 2009, 2010, 2011 - around 1 000,000 words***

***→ The Aix-Marseille & Paris corpus of TUFS***



The screenshot shows the website for the French (Aix) Multilingual Spoken Corpus. At the top is a banner image of a university campus with the text "French (Aix) Multilingual Spoken Corpus". Below the banner is a navigation menu on the left with links: Top, Outline, Research Team, References, Corpus (highlighted), Related Sites, and Pictures. The main content area is titled "Corpus" and contains a list of four items, each with a globe icon: 1. Student affairs section, 2. Reception of the university, 3. About faculty, and 4. About health condition.

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

## 2) The TUFS French native corpus and POS

Orthographic transcription, synchronized reader and Japanese/English translation  
(→ possible pedagogical exploitation)

The screenshot displays the TUFS French native corpus and POS interface. On the left is a navigation menu with links: Top, Outline, Research Team, References, Corpus (selected), Related Sites, and Pictures. The main content area features a header image with the text "French (Aix) Multilingual Spoken Corpus". Below this is a media player with three green circular buttons (play, pause, stop) and a progress bar. To the right of the player is a settings panel with checkboxes for "French" (checked), "Japanese" (checked), "English" (unchecked), and "Scroll On/Off" (unchecked). Below the settings panel is a section titled "About faculty" with four audio sample entries, each showing the French transcription and its Japanese translation.

**French (Aix) Multilingual Spoken Corpus**

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) The TUFS French native corpus and POS

- Traditional work in Aix: syntactic analyses
  - Corpus analysis in Parts-of-Speech (POS)  
(Tagging with the *Tree Tagger* tool- Schmidt 1995)
  - Search engine: morphosyntactic analyses of the productions

## 2) The TUFS French native corpus and POS

### *The POS search engine*



 **POS SEARCH ENGINE**

[POS Search Engine](#)

■ **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/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://www.coelang.tufs.ac.jp/multilingual_corpus/fr2/index.php?contents_xml=gaisetsu&lang=ja)

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

## 2) The TUFS French native corpus and POS

Example: prenominal adjectives in French

- DET (definite article) + ADJ (adjective) + NOM (noun)

pos

lemma

token

☐ Ignore the second context

## 2) The TUFS French native corpus and POS

pos:     0 0

lemma:

token:

☐ Ignore the second context

☒ paris ☒ aix

45. 01Kathy.txt:2071 tu passes **une mauvaise nuit** non gros/ non franchement

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 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 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 reggaeton

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 barbeaux qui y en a encore un qui est sorti là Papa pinguin B (le petit le petit) (le papa le papa le papa pinguin le papa le papa le papa pinguin) (en chantonnant) (rires)

## 2) The TUFS French native corpus and POS

- Results from POS search Engine for Prenominal adjectives
- 635 examples : mauvais (bad), petit (small), grand (big), même (same), bon (good), certain (certain), vieux (old), nouvel (new), etc...

## 2) The TUFS French native corpus and POS

- Research in TUFS: not only syntax but also syntax/prosody interface : **oral** corpus (work by Nakata & Kawaguchi)

→ Towards phonology

→ Towards PFC...

### 3) The PFC French native corpus and phonology

**PFC: Phonologie du Français Contemporain: usages, variétés et structure**  
(*Phonology of Contemporary French: usages, varieties and structure*)

**A collaborative project coordinated by :**

- Jacques DURAND (Toulouse), Bernard LAKS (Paris), Chantal LYCHE (Oslo and Tromsø)

PFC is a project which aims at describing the pronunciation of French in its geographical, stylistic and social diversity.

The main objectives of PFC are to:

- give a better picture of spoken French in its unity and diversity (geographical, social and stylistic);
- test phonological and phonetic models from a synchronic, diachronic and variationist point of view;
- favor communication between phonological studies and speech specialists;
- allow for the preservation of spoken varieties of French 'conservation du patrimoine' (cultural heritage)
- provide new material for the teaching of French language and linguistic

### 3) The PFC French native corpus and phonology

**PFC**

=

a project, a database and a scientific network  
(phonologists, phoneticians, syntacticians, sociolinguists...)

169 collaborators over 10 years, 500 hours of recorded  
material with the same protocol, 50 areas both in  
France and outside France

**PFC:** a corpus of reference

If you want to know more:

All publications and information freely accessible and  
downloadable on the website of the project:

[www.projet-pfc.net](http://www.projet-pfc.net)

## 4) Language education: PFC-EF & Language Modules

- **PFC & TUFs** : a partnership through applied linguistics (language education)

Two steps: 1) corpus constitution, 2) corpus exploitation

→ PFC for French language education?

→ The PFC-EF programme (EF: enseignement du français / French teaching)

### **PFC useful in terms of**

- orality (phonology),
- variability (French-speaking world),
- accessibility (website),
- usability (transcriptions & tools)

## 4) Language education: PFC-EF & Language Modules

→ PFC-EF provides:

- **Linguistic resources on:**

- Spoken French (*vs. written French*)
- Varieties of French (*vs. « standard » French*)

- **Material:**

- Sound
- Text
- Alignment Text-Sound

- **Useful for:**

- Pronunciation learning
- Listening skills development
- Linguistic study of spoken French
- Etc.

→ **PFC-EF results:** online & offline resources

# 4) Language education: PFC-EF & Language Modules

**Online:** a freely accessible website

## ➤ Four sections:

- French Illustrated (samples)
- French explained (descriptions)
- Linguistic resources (thematic corpus and concordancer)
- Pedagogical resources (pedagogical frames and content)



[www.projet-pfc.net](http://www.projet-pfc.net)

## 4) Language education: PFC-EF & Language Modules

### Offline resources:

S, Detey, J. Durand, B. Laks, C. Lyche (eds) (2010)

### Les variétés du français parlé dans l'espace francophone : ressources pour l'enseignement (avec DVD)

The Varieties of Spoken French in the French-speaking world: resources for teaching (with DVD)

37 specialists (12 countries - 23 universities)

Descriptions of French spoken in: France (North & South), Belgium, Switzerland, Africa and Overseas Territories, North America

31 conversations (2h30) transcribed and analyzed from lexicon to phonology for French language and linguistics teaching.

→ a **UNIQUE** volume



## 4) Language education: PFC-EF & Language Modules

### ■ On the TUFS side: the 'Language Modules'

[http://www.coelang.tufs.ac.jp/modules/fr/index\\_en.html](http://www.coelang.tufs.ac.jp/modules/fr/index_en.html)

40 multimedia saynètes based on one speech act: 'Thank s.o.', 'give one's opinion', etc. → for beginners (A1-B1 Level on the CEFRL)

PFC-EF & TUFS : development of an international version (English interface) including several varieties of French:

- 4 varieties:
  - **Standard** French (online)
  - **Canadian** French (online)
  - **Swiss** French (online)
  - **Southern** French (online from Sept. 2012)
- Integrate **variation** in the teaching of spoken French

## 4) Language education: PFC-EF & Language Modules

# Illustrations

Swiss Version



Standard Version

01 : Saluer (Mode3)

M. Merlin Mlle Lettre

Contexte on/off



Choix des langues: Français Anglais

Rôle on/off

- Bonjour, ça va ?
- Bonjour... il fait froid !
- Ça c'est vrai !
- Et vous sortez avec votre chien ?
- Où oui, la promenade du chien, c'est sacré !
- Quel courage !
- Et votre mari, sa grippe ?
- Ça va maintenant.
- Il est à la maison, mais ça va.
- Dites-lui bonjour de ma part.
- D'accord, c'est gentil.
- Bon, j'y vais.
- Bonne journée !
- Bonne journée vous aussi !

Canadian Version

01 : Saluer (Mode3)



Rôle on/off

Mlle Tremblay Mlle Gagnon

Choix des langues: Français(Québec) Anglais Vocabulaire

Contexte on/off

- Bonjour Mme Gagnon!
- Comment ça va?
- Bonjour Mme Tremblay.
- Très bien merci.
- Et vous?
- Pas pire, pas pire.
- Il fait beau soleil aujourd'hui.
- Ça fait changement!
- Oui, peut-être que le printemps va arriver.
- Des fois la température, ça revire de b.
- On sait jamais!
- Bon, je vous quitte déjà.
- J'ai un rendez-vous séché à 9 heures.
- Alors à tout à l'heure Mme Tremblay,
- N'oubliez pas de barrer en partant.

## **5) Corpus-based non-native French phonology : IPFC**

**Language Education + Phonological research → L2 phonology research**

**PFC → IPFC = Interphonology of Contemporary French  
(coordinated by S. Detey, I. Racine, Y. Kawaguchi)**

**In French (and other languages): very few L2 corpus-based phonology research**

**→ IPFC : build a database similar to PFC's for non-native speakers of French (learners...)**

**→ One single multitask research protocol for all learners (several L1s)**

## 5) Corpus-based non-native French phonology : IPFC

### IPFC in short:

Launched in 2008: IPFC-Japanese (Detey & Kawaguchi)

Today: 8 international teams:

- German (Munich, Düsseldorf)
- English (Canada)
- Spanish (Spain)
- Greek (Cyprus)
- Italian
- Dutch
- Norwegian

In preparation: Portuguese (Brazil), Swedish and Croatian

Several Master and PhD students involved in the project

Several areas of study: segmental (e.g. nasal and rounded vowels) and suprasegmental (e.g. liaison, stress) levels, but also lexicon (work by Sugiyama)...

## 5) Corpus-based non-native French phonology : IPFC

To know more : check the website of the project

- <http://cblle.tufs.ac.jp/ipfc>

### IPFC

#### Interphonologie du Français Contemporain (IPFC)

##### IPFC

Actualité  
Cadre IPFC  
Participants  
Descriptif  
Corpus  
Références  
Thèses et Mémoires

##### Projets IPFC

IPFC-allemand  
IPFC-anglais canadien  
IPFC-espagnol  
IPFC-grec chypriote  
IPFC-italien  
IPFC-japonais  
IPFC-néerlandais  
IPFC-norvégien

##### Colloques

IPFC2011-Paris  
IPFC2011-Tokyo  
IPFC2010

##### Sites partenaires

Bienvenue sur le site du projet IPFC (Interphonologie du français contemporain), piloté par:

Sylvain Detey (Université Waseda & Université de Rouen)

Isabelle Racine (Université de Genève)

Yuji Kawaguchi (Tokyo University of Foreign Studies)

Jacques Durand (Université de Toulouse & IUF)

Ce projet est dédié à l'étude des systèmes phonético-phonologiques des locuteurs non-natifs du français, pour lesquels le français est une langue étrangère (FLE) ou seconde (FLS). Il s'agit donc de populations d'apprenants qui peuvent faire usage du français dans diverses situations et appartiennent de ce fait au monde francophone.

Par interphonologie, on désigne généralement le nouveau système (phonético-)phonologique des apprenants d'une langue étrangère en cours de construction ou dans un état stabilisé.

Par-delà l'interphonologie, le projet IPFC concerne tous ceux qui s'intéressent à la production (et la perception) orale en français langue étrangère, puisque, à terme, le corpus IPFC devrait pouvoir être, au moins en partie, exploité pour des analyses multi-niveaux (morphologie, lexique, syntaxe, pragmatique).

- Pour une vision globale des objectifs et des enjeux du projet, consulter la section [Cadre IPFC](#).
- Pour une vision plus précise de chacun des sous-projets de IPFC, consulter la section [Projets IPFC](#).

**Carte des enquêtes**

## 5) Corpus-based non-native French phonology : IPFC

- With samples from the Japanese corpus :

<http://cblle.tufts.ac.jp/ipfc/ipfcsearch/>

pe                      ef                      se  
Phonologie du Français Contemporain

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

Recherche

| 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) |
|----|--------|--------------------|-------------------|------|-------------|-------------------|--------------|------------------------------|------------|---------------------------------------|------------------------------------------------------|
| 1  | 275    | 1986               | Yamanashi         | F    | japonaise   | japonais          | linguistique | 60                           | 0          | 12                                    | 0                                                    |

LECTURE LISTE GÉNÉRIQUE

compagnie
médecin
baignoire

all

RÉPÉTITION LISTE SPÉCIFIQUE

peur
moule
la vase

all

LECTURE LISTE SPÉCIFIQUE

peur
moule
la vase

all

LECTURE TEXTE

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.

|   |     |      |          |   |           |          |              |    |           |    |   |
|---|-----|------|----------|---|-----------|----------|--------------|----|-----------|----|---|
| 2 | 271 | 1988 | Osaka    | F | japonaise | japonais | linguistique | 96 | APEF Jun1 | 1  | 6 |
| 3 | 267 | 1989 | Okinawa  | M | japonaise | japonais | linguistique | 24 | APEF 3    | 0  | 0 |
| 4 | 266 | 1943 | Shizuoka | M | japonaise | japonais | linguistique | 48 | APEF Jun1 | 6  | 0 |
| 5 | 263 | 1986 | Fukuoka  | F | japonaise | japonais | linguistique | 48 | APEF 2    | 10 | 0 |

## 6) Conclusion

The partnership between TUFs and PFC : a good example of international academic cooperation in the field of linguistics and applied linguistics (research and language education)

Acknowledged by the Ministry

→ New sub-project (2011-2014):

Grant-in-Aid for Scientific Research (B) n° 23320121

*A corpus-based longitudinal study of the interphonological features of Japanese learners of French*

S. Detey (leader), Y. Kawaguchi, M. Kondo (members)

I. Racine, L. Buscail, J. Eychenne, P. Ménetrey, H. Tsubaki (associates)

Linguistic research for academic and applied purposes

# Thank you for your attention

Thanks to:

- Jacques Durand, Chantal Lyche, Bernard Laks and DGLFLF
- MEXT & JSPS (Grant-in-Aid for Scientific Research (B)  
n° 23320121 – S. Detey)
- CBLLE programme (Y. Kawaguchi)

## Select references 1/3

- 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.
- 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.
- Detey, S., Durand, J., Laks, B. & Lyche, C. (éds) (2010a). *Les variétés du français parlé dans l'espace francophone. Ressources pour l'enseignement*. Paris : Ophrys.
- Detey, S., Durand, J., Laks, B. & Lyche, C. (dir.) (2010b). *PFC-EF : bilan d'un projet de recherche appliquée pour la diffusion du français parlé dans l'espace francophone via Internet*. Rapport de recherche dans le cadre du programme *Valorisation et usages de corpus oraux* 2008-2009, DGLFLF, Ministère de la culture et de la communication, 137p. Document téléchargeable à l'adresse : <http://www.projet-pfc.net/le-projet-pfc-ef>
- Detey, S., Durand, J. & Nespoulous, J.-L. (2005). Interphonologie et représentations orthographiques. Le cas des catégories /b/ et /v/ chez des apprenants japonais de Français Langue Etrangère », *Revue Parole* 34/35/36 (supplément) : 139-186.

## Select references 2/3

- Detey, S. & Kawaguchi, Y. (2008). Interphonologie du français contemporain (IPFC) : récolte automatisée des données et apprenants japonais », Colloque *Phonologie du français contemporain: variation, interfaces, cognition*, Paris, MSH.
- Detey, S. & Kawaguchi, Y. (2011) Corpus d'apprenants, interlangue et français parlé : éclairage à travers le projet IPFC. *Journée IPFC2011-Tokyo : corpus d'apprenants, interlangue et français parlé*. Tokyo, TUFS.
- Kawaguchi, Y., Minegishi, M. & Durand, J. (éds) (2009). *Corpus Analysis and Variation in Linguistics*. Amsterdam/Philadelphia : John Benjamins.
- Kawaguchi, Y., Zaima, S. & Takagaki, T. (éds) (2006). *Spoken Language Corpus and Linguistic Informatics*. Amsterdam/Philadelphia : John Benjamins.
- Kondo, N. (2010). Apprentissage de la liaison. À l'usage du corpus d'IPFC sur Moodle (en japonais). *Bulletin of Japanese Association of Foreign Language Education* 13: 78-90.
- Marushima, N., Detey, S., Kawaguchi, Y. (2010). Caractéristiques phonétiques des voyelles orales arrondies du français chez des apprenants japonophones. *Flambeau* 36, Tokyo: Université des Langues Étrangères de Tokyo, 53-72.
- Matsuzawa, M. (2010). Choix de deux temps du passé chez les apprenants japonophones au niveau débutant – Passé composé et imparfait (en japonais), *Flambeau* 36, Tokyo : Université des Langues Étrangères de Tokyo, 34-52.

## Select references 3/3

- Nakata, S. & Kawaguchi, Y. (2005). Traits prosodiques à la lumière de structures syntaxiques – Significations de phrases, intonation, durée syllabique et pause (en japonais). *Explorer le français. Problèmes de la linguistique française III*, Tokyo :Sanshûsha, 308-321.
- Sugiyama, K. (2010). Construction du corpus des apprenants du français et son analyse – Au tour de fillers et pronom neutre – (en japonais). *Working Papers in Corpus-based Linguistics and Language Education 4*, Tokyo: TUFS, 313-327.
- Sugiyama, K. (2011). Profil lexical de la production orale chez les étudiants japonais. *Journée IPFC 2011-Tokyo: Corpus d'apprenants, interlangue et français parlé*. Tokyo, TUFS, mars 2011.
- Sugiyama, K. (à paraître). Lexical Profile of French Learner Speech: The Case of Japanese University Students. In Tono, Y., Kawaguchi, Y. & Minegishi, M. (éds.), *Developmental and Crosslinguistic Perspectives in Learner Corpus Research*. Amsterdam/Philadelphia, John Benjamins.

# Acknowledgement

This work was supported by JSPS KAKENHI Grant Numbers 24320102, 24320106, 23320121

Grant-in-Aid for Scientific Research(B)

Responsible: Yuji KAWAGUCHI

Grant-in-Aid for Scientific Research(B)

Responsible: Sylvain DETEY

Grant-in-Aid for Scientific Research(B)

Responsible: Yasushi SEKIYA