

## La liaison dans un corpus d'apprenants : Le projet « Interphonologie du Français Contemporain » (IPFC)

*Liaison in the Interphonology of Contemporary French (IPFC) learner corpus*

Isabelle Racine et Sylvain Detey

---



### Édition électronique

URL : <http://journals.openedition.org/corpus/3028>

ISSN : 1765-3126

### Éditeur

Bases ; corpus et langage - UMR 6039

### Édition imprimée

Date de publication : 15 octobre 2016

ISSN : 1638-9808

### Référence électronique

Isabelle Racine et Sylvain Detey, « La liaison dans un corpus d'apprenants : Le projet « Interphonologie du Français Contemporain » (IPFC) », *Corpus* [En ligne], 15 | 2016, mis en ligne le 15 janvier 2017, consulté le 08 septembre 2020. URL : <http://journals.openedition.org/corpus/3028>

---

Ce document a été généré automatiquement le 8 septembre 2020.

© Tous droits réservés

---

# La liaison dans un corpus d'apprenants : Le projet « Interphonologie du Français Contemporain » (IPFC)

*Liaison in the Interphonology of Contemporary French (IPFC) learner corpus*

Isabelle Racine et Sylvain Detey

---

*Nous tenons à remercier Marie-José Béguelin, Fédérica Diémoz et Mathieu Avanzi pour l'organisation du colloque « Corpus de français parlé et français parlés des corpus », à Neuchâtel, les 8 et 9 mai 2014, journées qui sont à la base de cette publication. Nous remercions également Yuji Kawaguchi qui codirige le projet IPFC avec nous, ainsi que Julien Eychenne pour le développement de Dolmen-IPFC. Le travail présenté ici n'aurait pas été possible sans le précieux concours de Marion Didelot, Tanjema Majeed, Cécile Mollet, Marie-Laure Sandoz et Roberto Paternostro, qui ont transcrit, codé et vérifié les données. Enfin, nous tenons à remercier également tous nos collègues impliqués dans IPFC pour leur dynamisme qui contribue à faire vivre ce projet, ainsi que tous les apprenants qui ont donné de leur temps pour la constitution de ce corpus.*

- 1 Si l'apport de la linguistique de corpus à l'étude du français parlé par des locuteurs natifs n'est plus à démontrer, comme en témoignent les différents chapitres qui constituent ce volume, il en va autrement lorsqu'il s'agit de la parole de locuteurs non natifs. Dans le domaine de l'acquisition du français langue étrangère ou seconde (ci-après L2)<sup>1</sup>, les études basées sur de grands corpus ne sont en effet pas légion<sup>2</sup>, et c'est encore moins le cas lorsque l'on se concentre sur les aspects phonétiques et phonologiques.
- 2 Les principaux modèles théoriques d'acquisition phonologique d'une L2 sont en effet le plus souvent fondés sur des données issues de parole « de laboratoire » qui sont trop restreintes tant au niveau du style de parole que du nombre de locuteurs et de structures étudiés (Zampini, 2008). Comme le mentionne Gut (2009, 2014), l'approche

par corpus permet en partie de dépasser ces limitations et d'élargir le regard porté sur l'acquisition phonologique d'une L2, en intégrant notamment la dimension de comparabilité inter-tâches et inter-apprenants d'une même L2 ou de L2 différentes.

- 3 Afin de combler ce retard en la matière, de nombreux projets de corpus L2 ont vu le jour ces dix dernières années et ce pour différentes langues : en néerlandais L2 (Neri *et al.*, 2006), en polonais L2 (Cylwik *et al.*, 2009), en allemand L2 et en anglais L2 dans un contexte européen (Gut, 2009) ou asiatique (Visceglia *et al.*, 2009), et ce, non seulement sur le plan segmental mais aussi suprasegmental (Trouvain & Gut, 2007 ; Meng *et al.*, 2009). Ce champ est toutefois resté peu exploité en ce qui concerne le français L2, et c'est ce qui a motivé le lancement, en 2008, du projet « Interphonologie du français contemporain » (Detey & Kawaguchi, 2008 ; Racine *et al.*, 2012 ; Detey & Racine, 2012, Detey *et al.*, à paraître<sup>3</sup>, ci-après IPFC), qui constitue l'une des extensions récentes du programme « Phonologie du français contemporain : usages, variétés et structure » (PFC ; Durand, Laks & Lyche, 2009, 2014 ; Racine, Durand & Andreassen, ce volume). Après sept ans d'existence, IPFC assure donc, conjointement à d'autres projets plus récents (le corpus *Coreil*, Delais-Roussarie & Yoo, 2010 et le corpus *Phodiflé*, Landron *et al.*, 2011), la présence du français aux côtés des travaux menés dans d'autres langues.
- 4 L'objectif du présent article est double : d'une part, il vise à offrir une brève présentation ainsi qu'un état des lieux du projet IPFC après sept ans d'existence et, d'autre part, à illustrer l'intérêt d'un tel corpus pour l'étude de l'acquisition de la phonologie du français L2, ce que nous nous proposons de faire par le biais d'un phénomène qui constitue un passage obligé dans l'enseignement du français langue étrangère (ci-après FLE), à savoir la liaison. Dans la deuxième partie de cet article, nous examinerons donc la question de l'acquisition de la liaison en français L2 et de ses enjeux pour l'enseignement. Nous expliquerons comment la liaison est traitée dans le projet IPFC avant de présenter une analyse des productions de liaison issues de la lecture du texte par des apprenants hispanophones et japonophones. Par ce biais, nous montrerons en quoi les données IPFC permettent d'enrichir la réflexion sur la liaison en L2, qui devrait permettre, à terme, un renouvellement du matériel pédagogique pour l'enseignement de cette question.

## 1. Le projet « Interphonologie du français contemporain »

- 5 L'objectif du projet IPFC est de constituer, analyser et mettre à disposition un corpus de recherche de FLE varié. La base de données consiste en un ensemble de points d'enquête, rassemblant des populations d'apprenants de profils linguistiques aussi similaires que possible (même variété de L1, même niveau de compétence linguistique en français d'après les descripteurs du *Cadre européen commun de référence pour les langues* (CECRL) et, si possible, parcours d'apprentissage similaire). Dans la lignée de PFC, un focus particulier a été mis sur l'analyse de la variation et sur la comparabilité des données. Le protocole de recueil de données est par conséquent identique pour toutes les L1 et est resté le plus proche possible de celui de PFC au niveau des tâches, sans exclure évidemment une adaptation de ses modalités à des populations d'apprenants (pour une description du protocole PFC, voir Racine, Durand & Andreassen, ce volume).

- 6 Le recueil de données est constitué de six tâches : 1) la répétition d'une liste spécifique de mots produits par un locuteur natif, 2) la lecture de la liste de mots PFC, 3) la lecture de la même liste spécifique qui a été répétée dans la première tâche, 4) la lecture du texte PFC, 5) un entretien guidé avec un natif, le plus souvent un enseignant, et 6) une interaction semi-contrainte entre deux apprenants.
- 7 La liste spécifique comprend à la fois des difficultés qui sont partagées par l'ensemble des apprenants, telles que les voyelles nasales ou antérieures arrondies, ainsi que des difficultés plus spécifiques à chaque population d'apprenants ciblée, comme par exemple les occlusives sonores en initiale de mots pour les apprenants suisses alémaniques ou les groupes consonantiques pour les japonophones. La tâche de répétition a été incluse de manière à contourner les effets orthographiques induits par une tâche de lecture (Detey, 2005 ; Vendelin & Peperkamp, 2006 ; Bassetti, 2006). L'entretien guidé, quant à lui, comporte des questions fermées et ouvertes, qui sont adaptées au niveau de l'apprenant ainsi qu'au contexte dans lequel se déroule son apprentissage (homoglotte vs hétéroglotte). L'interaction semi-contrainte débute par une brève présentation de l'un des apprenants sur un sujet spécifique, suivie d'une série de questions-réponses qui se transforme généralement en discussion libre entre les deux intervenants. Les rôles sont ensuite inversés. Un questionnaire biographique complète le protocole.
- 8 Si le projet IPFC a débuté avec la constitution d'un corpus d'apprenants japonophones et d'un corpus d'apprenants hispanophones, d'autres chercheurs s'y sont depuis associés, et le projet regroupe actuellement quinze équipes travaillant avec des apprenants de L1 différentes (allemand, anglais canadien, arabe, coréen, danois, espagnol, grec, italien, japonais, néerlandais, norvégien, portugais brésilien, russe, suédois et turc). De nouvelles équipes s'y adjoignent régulièrement. Ainsi, des projets avec des apprenants malais et tchèques sont par exemple en cours d'élaboration. En parallèle, le travail sur les L1 déjà représentées dans le projet s'élargit. C'est le cas pour l'italien où, outre les corpus de Milan, Trieste et Rome, un corpus avec des apprenants tessinois est en cours de constitution, de même que pour l'allemand, où des données d'apprenants autrichiens et suisses allemands sont également collectées depuis peu<sup>4</sup>.
- 9 L'ensemble du corpus ainsi recueilli représente environ une heure de données par apprenant. Les enregistrements sont transcrits orthographiquement avec alignement au signal – sous *Praat* (Boerma & Weenink, 2014) –, selon des conventions spécifiques adaptées à la parole en L2 (cf. Racine *et al.*, 2011), notamment concernant des cas d'ambiguïté au niveau morpho-lexical ou phonético-phonologique. Divers types d'analyse sont ensuite appliqués aux données. Pour les mots en isolé, une évaluation experte vs non experte, éventuellement complétées par des analyses acoustiques (pour des exemples concernant les voyelles nasales et arrondies, voir Detey *et al.*, 2010 ; Racine *et al.*, 2010 ; Racine, Detey & Kawaguchi, 2012 ; Racine, 2012). Ce type d'évaluation n'étant pas adapté à la parole continue – la dimension morpho-grammaticale est en effet susceptible de biaiser l'évaluation de la production de surface – nous avons choisi d'appliquer, dans la lignée de ce qui a été fait dans PFC pour la liaison et le schwa, une approche par le biais d'un codage alphanumérique de différents paramètres (Detey, 2012 ; Detey *et al.*, 2014a ; Detey *et al.*, 2014b ; Detey *et al.*, à paraître). Le code alphanumérique, spécifique à chaque phénomène ciblé, est inséré manuellement sur la base d'une évaluation perceptive, dans la transcription orthographique, dans des tires séparées (une par phénomène ciblé).

- 10 Une triple motivation a guidé ce choix : i) dans le quotidien des locuteurs apprenants, la parole non native est évaluée de manière catégorielle et non pas en termes de détails phonétiques fins. Les analyses acoustiques, qui fournissent des informations fines sur les caractéristiques acoustiques du signal, ne sont donc pas adéquates. Seule une évaluation perceptive humaine, en termes de catégories phonologiques, permet d'atteindre cet objectif. L'analyse du signal peut toutefois constituer une étape ultérieure en vue d'obtenir des informations plus précises sur un ensemble particulier de *stimuli*, par exemple des productions ayant été évaluées comme très éloignées de la cible visée ; ii) le recours à un codage alphanumérique ciblant certains paramètres précis permet d'éviter certains pièges liés à la transcription phonétique. En effet, s'il est possible d'évaluer le caractère nasalisé ou non d'une voyelle, l'adéquation de son timbre par rapport à une cible donnée ou la présence/absence d'un appendice consonantique postvocalique, il n'est pas toujours possible de sélectionner un symbole phonétique qui représente de manière adéquate la réalisation globale de cet élément par un apprenant ; iii) l'idée fondamentale de l'approche par corpus est de décrire, de manière aussi automatique que possible, sur la base d'un ensemble conséquent de données, les tendances globales qui se dégagent de la parole en L2, en prenant en compte le contexte de production et sans précatégoriser – et pré-interpréter – les productions interlangagières de l'apprenant sur la base des catégories phonémiques de la L2 ou de la L1, puisqu'il s'agit justement d'un système en construction dont les formes peuvent être potentiellement non identifiables dans les deux systèmes en contact, source et cible.
- 11 Une fois les données codées<sup>5</sup>, l'analyse s'effectue par le biais du concordancier *Dolmen*, développé par J. Eychenne (cf. Eychenne & Paternostro, à paraître), dans le cadre duquel des interfaces spécifiques pour les phénomènes ciblés dans IPFC ont été développées, à savoir, pour l'instant, les voyelles orales et nasales, les consonnes, les groupes consonantiques et la liaison. *Dolmen* permet d'obtenir de manière automatique des statistiques descriptives concernant certaines caractéristiques des productions non natives.
- 12 En conclusion, cette procédure de codage alphanumérique constitue, à nos yeux, une étape intermédiaire entre des analyses acoustiques fines et une catégorisation phonologique plus grossière, en termes de substitution, d'insertion ou d'effacement (Detey, 2012, 2014). La démarche que nous avons adoptée et développée dans IPFC a inspiré des travaux similaires d'analyse de corpus pour d'autres langues cibles (pour l'espagnol, voir Carranza, Cucchiarini, Llisterri, Machuca & Rios, 2014 ; pour l'anglais, voir Lacoste, Herry-Bénil & Kamiyama, 2013). Dans le cadre du projet IPFC, des analyses des productions de voyelles nasales et arrondies (Detey, Racine & Kawaguchi, 2014 ; Detey *et al.*, 2014a et 2014b) d'apprenants japonophones et hispanophones ont été conduites selon cette procédure. La question de l'acquisition de la liaison L2 par des apprenants de différentes L1 est également en cours d'examen par ce biais dans le projet (voir, par exemple, les contributions dans Racine et Detey (2015) pour les apprenants allemands (Pustka), canadiens anglophones (Tennant), espagnols (Racine), grecs chypriotes (Valetopoulos), italiens (Falbo, Janot, Murano & Paternostro), japonais (Detey, Kawaguchi & Kondo) et norvégiens (Andreassen & Lyche)). C'est à travers ce phénomène phonologique spécifique du français que nous avons choisi d'explicitier de manière plus concrète la démarche adoptée dans le projet IPFC.

## 2. La question de la liaison : enjeux pour le FLE<sup>6</sup>

- 13 La liaison est, comme le rappellent Durand & Lyche (à paraître), un phénomène de sandhi externe qui implique la présence – ou l'absence – d'une consonne entre deux mots produits conjointement, le premier mot étant appelé « mot liaisonnant ». Selon la littérature, les consonnes de liaison possibles sont au nombre de huit ([z, n, t, ʁ, p, v, k, ʒ])<sup>7</sup>, les trois premières couvrant la plus grande majorité des occurrences. La consonne de liaison peut apparaître lorsque le deuxième mot commence par une voyelle (p. ex. : « un escargot » [œ̃nɛskaʁʒo]), mais est absente lorsque le deuxième mot commence par une consonne (p. ex. : « un stylo » [œ̃stilo]). En outre, cette consonne se rattache le plus souvent – mais pas obligatoirement – à l'attaque de la syllabe suivante, qui appartient donc au deuxième mot. Lorsqu'il y a resyllabation de la consonne finale, la liaison est enchaînée (p. ex. : « un escargot » [œ̃nɛs-kaʁ-ʒo]). Elle est produite sans enchaînement lorsque cette resyllabation n'a pas lieu (p. ex. : « un escargot » [œ̃n-ɛs-kaʁ-ʒo]). Ainsi, lorsqu'elle est enchaînée, la liaison, qui reflète un état ancien de la langue où toutes les consonnes étaient prononcées (pour un historique, cf. Mallet, 2008), contribue à renforcer la tendance à la syllabation ouverte et la prédilection du français pour les syllabes de type CV (Delattre, 1946).
- 14 La réalisation de la liaison est soumise à variation et dépend, comme le rappellent Eychenne *et al.* (2014) et Durand & Lyche (à paraître) de facteurs phonologiques – y compris prosodiques –, lexicaux, (morpho)syntaxiques ainsi que sociolinguistiques. Parmi ces derniers, ces auteurs relèvent notamment le registre, l'âge, la provenance du locuteur, le niveau d'instruction ou encore le degré de connaissance de l'orthographe. En outre, lorsque la liaison est réalisée, elle peut parfois avoir un effet sur la voyelle précédant la consonne de liaison, soit en l'ouvrant (p. ex. : « premier élève » sera ainsi produit [pʁɛ̃mjɛvɛlɛv], avec une forme masculine ne se différenciant plus de la forme féminine « première »), soit en la dénasalisant (p. ex. : « bon après-midi » sera ainsi produit [bɔnapʁɛ̃midi], avec une forme masculine dont la prononciation est identique à sa contrepartie féminine « bonne »)<sup>8</sup>.
- 15 Si l'on se place du point de vue d'un apprenant, la question de la liaison est extrêmement complexe, notamment en raison du fait que sa réalisation va à l'encontre de certains principes que l'apprenant a dû – ou est en train – d'intégrer, comme le souligne Howard (2013). Il a ainsi dû apprendre à ne pas prononcer la consonne finale graphique d'un grand nombre de mots, ce qui constitue une difficulté accrue pour les apprenants dont la L1 dispose d'une très bonne correspondance entre prononciation et orthographe. Or, dans le cas de la liaison, cette consonne finale habituellement non prononcée peut l'être dans certaines conditions, de surcroît avec une réalisation qui ne correspond pas toujours à la consonne graphique, et le plus souvent en se resyllabant avec la voyelle initiale du mot suivant (p. ex. : « grand », avec une consonne finale non prononcée qui, si elle se réalise en liaison, l'est en [t], p. ex. : « grand animal » [ɡʁɑ̃-ta-ni-mal]).
- 16 La liaison en [n] renforce ce phénomène. En effet, l'acquisition des voyelles nasales du français constitue déjà en soi une difficulté majeure pour la plupart des apprenants. La dimension graphique ne leur simplifie pas la tâche puisque, outre le fait que les voyelles nasales possèdent des graphies multiples, l'apprenant doit acquérir des graphèmes complexes dans lesquels la consonne nasale ne se prononce pas (p. ex. : « plein » [plɛ̃]). Or, dans le cas de la liaison en [n], la consonne nasale graphique se

prononce, et la voyelle qui précède peut perdre sa nasalité (p. ex. : « plein hiver » [plɛn ivɛʁ]), mais pas de manière systématique (« rien à faire » [ʁjɛ̃nafɛʁ]).

- 17 À cela s'ajoute le fait que, comme mentionné précédemment, la liaison est soumise à variation, variation dont l'apprenant doit acquérir la maîtrise, et dont Wauquier (2009) distingue deux types : une variation conditionnée par le contexte d'abord, qui fait que certaines liaisons sont systématiquement réalisées, réalisées de manière variable ou pas réalisées – liaisons catégoriques, variables ou erratiques, selon la terminologie adoptée par Durand & Lyche (2008), appelées également obligatoires, facultatives ou interdites dans une perspective plus normative –, et, dans le cas des liaisons variables, une variation libre, qui reflète le choix du locuteur de réaliser ou non la liaison, généralement en fonction de facteurs d'ordre sociolinguistique.
- 18 Les difficultés posées par la liaison à l'apprenant sont donc multiples mais peuvent cependant, nous semble-t-il, être réparties en deux catégories distinctes : 1) les problèmes posés par le phénomène de liaison lui-même, qui sont de l'ordre de ce que l'on pourrait qualifier de *microplanification* : nature de la consonne de liaison à réaliser (en lien avec la graphie), placement de cette consonne dans la structure syllabique (avec resyllabation avec la voyelle initiale du mot suivant ou non), modification éventuelle de l'environnement immédiat (dénasalisation/ouverture de la voyelle précédente) ; 2) les problèmes posés par la variation inhérente à la liaison, qui relèvent davantage d'une planification à un niveau *macro*, cette fois-ci, et qui peuvent être résumés ainsi : identifier le contexte – sur la base de facteurs phonologiques, prosodiques, lexicaux, (morpho)syntaxiques et sociolinguistiques – afin de déterminer si la liaison doit – ou peut – ou non être réalisée.
- 19 L'essentiel des explications et des exercices consacrés à la liaison dans les manuels de FLE spécialisés dans le domaine de la prononciation se rapportent plutôt à ce deuxième type de difficultés, lié aux contextes de réalisation de la liaison. Les questions qui relèvent de la *microplanification* sont généralement mentionnées mais de manière sommaire, avec peu ou pas d'exercices ciblés et sans conscientisation des difficultés pour l'apprenant.
- 20 Les contextes de réalisation de la liaison sont généralement divisés en trois catégories – liaisons obligatoires, facultatives et interdites – et présentés sous forme de tableaux ou de listes basées sur des généralisations d'ordre morphosyntaxique (p. ex. : « la liaison est obligatoire entre le déterminant et le nom qui le suit »), suivies d'exemples. Les autres facteurs de variation – d'ordre sociolinguistique, lexical ou prosodique – ne sont généralement pas mentionnés, hormis la distinction, pour les liaisons facultatives, entre registre courant/standard et soutenu. On peut relever que les explications sont parfois opaques, que le nombre de règles énoncées est plutôt élevé (selon les manuels, entre 6 et 9 pour chacune des catégories « liaisons obligatoires » et « liaisons interdites »), avec des divergences d'un manuel à l'autre.
- 21 On constate également que certaines règles énoncées ne semblent pas correspondre à la réalité des usages actuels. Ainsi, par exemple, la liaison est présentée comme obligatoire après les prépositions monosyllabiques suivantes : « en », « dans », « chez », « sans », « sous » (Abry & Chalaron, 1994 : 110). Or, Eychenne *et al.* (2014 : 44), en se basant sur les données issues du corpus PFC, montrent que, si la liaison est quasi catégorique après « en » (taux de réalisation : 99.35 %), elle l'est moins après « sans » (93.13 %) et encore moins après « chez » (75.68 %). Le dernier cas illustre, selon les auteurs, le rôle primordial de la prosodie : la liaison est en effet catégorique dans le

contexte « chez + clitique », alors qu'elle est variable à 50 % dans le contexte « chez + syntagme nominal ». Eychenne *et al.* (2014 : 34) expliquent ce décalage entre les descriptions traditionnelles et les usages – mis en avant par les travaux sur la liaison basés sur des corpus oraux (entre autres Ågren, 1973 ; Encrevé, 1988 ; De Jong, 1994 ; Durand & Lyche, 2008)<sup>9</sup> – par le fait que, si de nombreux travaux ont été consacrés à la liaison dans la seconde moitié du XX<sup>e</sup> siècle, peu d'entre eux ont pris en compte les données authentiques. La plupart des analyses classiques en linguistique ont été élaborées sur la base de descriptions normatives destinées à faciliter l'apprentissage du phénomène par des apprenants étrangers – les auteurs citent notamment Delattre (1951) et Fouché (1959). Or, ce sont ces mêmes descriptions qui sont à la base des classements qui figurent généralement dans les manuels de phonétique. On peut ajouter que, hormis la question de l'authenticité des données sur lesquelles s'appuient ces descriptions, le décalage entre descriptions et usages réels peut également s'expliquer par le fait que l'on ne s'exprime plus aujourd'hui de la même manière qu'à l'époque où elles ont été élaborées. De Jong (1994) a en effet mis en évidence l'importance du facteur âge : les locuteurs jeunes de son corpus réalisent un taux de liaison significativement inférieur par rapport à leurs aînés, ceci étant particulièrement marqué entre la tranche des 18-29 ans et la tranche des 30-49 ans, ce qui semble indiquer qu'un changement majeur a eu lieu au milieu des années soixante<sup>10</sup>, les descriptions traditionnelles mentionnées – Delattre (1951) et Fouché (1959) – y étant donc antérieures.

- 22 Comme le soulignent Eychenne *et al.* (2014), les travaux basés sur des corpus oraux ont montré que l'on ne peut expliquer la liaison en se basant uniquement sur des principes d'association mécanique mais que d'autres facteurs, tels que par exemple la fréquence lexicale du mot liaisonnant et la prosodie, doivent être pris en compte. Le travail réalisé sur la liaison dans le cadre du projet PFC a permis d'affiner encore les descriptions des usages réels. Durand & Lyche (2008) peuvent ainsi réduire à quatre contextes les liaisons systématiquement réalisées dans les usages : déterminant + substantif (p. ex. : « un animal »), proclitique + verbe (« ils aiment », « ils y dorment souvent »), verbe + enclitique (p. ex. : « dit-il », « fais-en ») et expressions figées (p. ex. : « de mieux en mieux », « de temps à autre »). Les données PFC montrent également que la liaison variable est très peu réalisée en conversation spontanée, que les liaisons sont quasiment systématiquement produites avec enchaînement et que, du point de vue des facteurs externes, la liaison semble essentiellement influencée par l'âge et par l'origine géographique des locuteurs ainsi que par le registre (Eychenne *et al.*, 2014).
- 23 Si, d'un point de vue linguistique, comme le relèvent Eychenne *et al.* (2014 : 56), il semble qu'on ne peut plus aujourd'hui se satisfaire de données « fabriquées » et que l'on se doit d'intégrer les résultats des études sur corpus dans les modélisations phonologiques, la même conclusion semble s'imposer dans le domaine de l'enseignement du FLE : les descriptions des manuels de phonétique doivent être renouvelées afin de mieux correspondre aux usages réels en vigueur. À ce titre, on peut déjà mentionner les ouvrages de Lauret (2007) et d'Abry & Chalaron (2011), qui proposent quant à eux une description des contextes de liaison catégorique qui correspond aux quatre contextes décrits par Durand & Lyche (2008) où la liaison est systématiquement réalisée.
- 24 Pour compléter ce panorama de la liaison, il paraît important d'examiner également la question du point de vue des productions des apprenants, afin de déterminer ce qui



leur pose problème. Un certain nombre de travaux se sont penchés sur cette question et ont analysé les occurrences de liaison dans les productions d'apprenants de FLE (De Moras, 2011 ; Harnois-Delpiano *et al.*, 2012 ; Howard, 2005, 2013 ; Mastromonaco, 1999 ; Thomas, 2002, 2010). On peut tout d'abord relever l'hétérogénéité des travaux qui portent sur cette question et qui ont examiné les productions des liaisons par des apprenants – en grande majorité anglophones – de niveaux variés (débutants à avancés), dans des contextes d'apprentissage divers – notamment avec ou sans séjour dans une région francophone – et avec des tâches distinctes (production de séquences isolées, description d'images, texte lu, entretien guidé et parole spontanée). Les résultats sont néanmoins assez homogènes et montrent d'une part un taux généralement élevé de réalisation des liaisons obligatoires chez les apprenants avancés, avec – lorsqu'une dimension longitudinale est intégrée, comme par exemple dans l'étude de Howard (2013) – une progression dans le taux de réalisation de certaines de ces liaisons obligatoires, notamment pour celles entre déterminant et substantif et après les pronoms sujets. Certains contextes semblent par contre poser plus de problèmes. Howard (2005, 2013) relève en effet des taux moins élevés après les pronoms clitiques et un taux très faible de réalisation de la liaison entre un adjectif et le substantif qui le suit. Toutefois, il faut rappeler que cette dernière catégorie n'apparaît pas comme catégorique dans les données de corpus de Durand & Lyche (2008). La comparaison avec les natifs doit donc être considérée avec précaution car les études portant sur les apprenants n'ont généralement pas constitué de corpus de données natives comparables<sup>25</sup> et se basent donc sur la littérature dans le domaine. Or, comme le souligne De Moras (2011), les différentes catégories de liaison ne sont pas définies de manière uniforme dans la littérature. De plus, comme nous l'avons vu précédemment, la prise en compte des données de corpus, qui questionnent la classification traditionnelle en obligatoire/facultative/interdite et soulignent l'importance de certains facteurs, est encore relativement récente.

- 25 Les travaux qui ont examiné les productions des liaisons par des apprenants fournissent également des indications concernant les difficultés au niveau de ce que nous avons appelé la *microplanification*. Mastromonaco (1999), Harnois-Delpiano *et al.* (2012) et Thomas (2002) observent en effet des erreurs de type [gʁɑ̃dɑmi] pour « grand ami » ou [gʁɑ̃nɑm] pour « grand homme » qui semblent découler d'une prononciation orthographique et qui n'apparaissent pas chez les enfants natifs (Wauquier, 2009). Mastromonaco (1999) et De Moras (2011) mentionnent la difficulté particulière engendrée par les liaisons en [n] liées à la question des voyelles nasales. Un autre phénomène intéressant est le taux important de liaisons réalisées sans enchaînement par les apprenants, qui est de 7 % chez Mastromonaco (1999) – sans distinction de tâches (lecture, description et conversation) – de 8.5 % chez Thomas (2002) et de 8.85 % – pour les liaisons obligatoires uniquement – chez De Moras (2011), alors que, comme mentionné auparavant, ce type de réalisation est quasiment absent des données de conversation du corpus PFC (Eychenne *et al.*, 2014), et il semble également l'être des données d'acquisition en L1 (Wauquier, 2009). Pour Thomas (2002), cela peut refléter soit une hésitation devant un mot difficile, soit l'influence de la tendance à la syllabation fermée de l'anglais, la L1 des apprenants de son étude, hypothèse également avancée par De Moras (2011).
- 26 Il semble ainsi indispensable d'affiner davantage les connaissances dans le domaine de l'acquisition de la liaison en FLE, afin de déterminer si les difficultés observées dans les travaux portant sur des apprenants anglophones peuvent être généralisées aux

apprenants d'autres L1 ou sont induites par des caractéristiques spécifiques aux langues germaniques, telles que la tendance à la syllabation fermée. De nouvelles données d'apprenants de L1 plus diversifiées doivent être collectées et la comparabilité non seulement avec les productions natives mais également entre les différentes populations d'apprenants doit être améliorée, de même qu'en termes de tâche(s) effectuée(s).

### 3. Le traitement de la liaison dans IPFC

- 27 Sur la base de ce qui a été décrit dans la première partie de ce chapitre, le projet IPFC semble être en mesure de répondre aux besoins mentionnés ci-dessus et par conséquent constituer un cadre adéquat pour approfondir les recherches dans le domaine de l'acquisition de la liaison en FLE.
- 28 Sur le modèle de la méthodologie adoptée dans PFC, la liaison y est traitée par le biais d'un codage alphanumérique inséré dans une tire spécifique, sous *Praat* (Boersma & Weenink, 2014). Le code PFC pour la liaison a été modifié afin de s'adapter aux caractéristiques des productions non natives et aux difficultés générées par la liaison. Pour pouvoir repérer les liaisons « orthographiques » (p. ex. : [gʁɑ̃dɑmi] pour « grand ami » ou [gʁɑ̃nɔm] pour « grand homme ») observées dans les travaux antérieurs en FLE, il est en effet indispensable que la consonne cible soit codée, ce qui n'est pas le cas dans le code PFC, des erreurs de ce type n'apparaissant pas chez les natifs. Le code IPFC pour la liaison comprend donc sept champs. Les quatre premiers portent sur des éléments descriptifs : 1) nature de la consonne de liaison cible, 2) catégorie syntaxique du mot liaisonnant, 3) catégorie syntaxique du mot qui suit, 4) nombre de syllabes du mot liaisonnant et nature – orale ou nasale – de la voyelle du mot liaisonnant. Les trois derniers champs ciblent l'évaluation perceptive de la liaison : 5) réalisation de la liaison et, si oui, avec ou sans enchaînement, 6) nature et caractéristiques de la consonne de liaison (correspond à la cible ou non, etc.), 7) présence d'une pause, d'une hésitation ou d'un coup de glotte. Ces données sont ensuite décodées par le biais d'une interface spécifique du concordancier *Dolmen*, qui permet d'obtenir facilement des statistiques descriptives, telles que le nombre de liaison en [n] réalisées, le nombre de liaisons produites avec enchaînement, etc.
- 29 Cette procédure est actuellement appliquée à grande échelle dans le projet IPFC (voir Racine & Detey, 2015). Dans ce qui suit, nous l'illustrons en présentant les résultats d'une étude préliminaire portant sur des productions d'apprenants avancés hispanophones et japonophones.

## 4. Une illustration : la liaison en lecture de texte chez les apprenants hispanophones et japonophones

### 4.1 Méthode

- 30 Trente-sept apprenants avancés (B2-C1 du CECRL), 20 Espagnols et 17 Japonais, ainsi que 10 natifs suisses romands ont pris part à cette étude préliminaire. Les apprenants étaient répartis en deux groupes : avec séjour prolongé en milieu francophone vs sans séjour. Dans le groupe d'hispanophones, 10 apprenants (8 femmes et 2 hommes, âge

moyen : 27.0) étudiaient le français – ou en français – à l'Université de Genève et vivaient dans un environnement francophone depuis au moins douze mois. Les 10 autres (8 femmes et 2 hommes, âge moyen : 26.6) étudiaient le français à Madrid, à l'Université ou dans une école de langue officielle. Aucun d'entre eux n'avait effectué de séjour de plus de trois semaines en milieu francophone. Pour le groupe de japonophones, 7 apprenants (5 femmes et 2 hommes, âge moyen : 24.7) étudiaient le français à TUFs (Tokyo University of Foreign Studies) et avaient effectué pendant leurs études un séjour prolongé (de 6 mois à 1 an) en milieu francophone. Les 10 autres (8 femmes et 2 hommes, âge moyen : 20.2) étaient également étudiants de français à TUFs mais n'avaient effectué aucun séjour en milieu francophone. Ces données ont été comparées avec celles d'un groupe contrôle constitué de 10 francophones natifs suisses romands (8 femmes et 2 hommes, âge moyen : 27.3), enregistrés dans le cadre de PFC.

- 31 Les participants disposaient de 5 minutes pour prendre connaissance du texte PFC intitulé « Le Premier ministre ira-t-il à Beaulieu » et le préparer, avant de le lire à voix haute. Rappelons que, dans le protocole PFC, ce texte a été conçu spécifiquement pour étudier la liaison et le schwa et contient donc 35 sites potentiels de liaison (17 sites de liaisons obligatoires, 14 de liaisons facultatives et 4 de liaisons interdites selon la classification de Delattre, 1951). 1 295 sites potentiels de liaison ont ainsi été codés selon le code liaison IPFC par une locutrice native et vérifiés par un deuxième évaluateur<sup>12</sup>.

## 4.2 Résultats et discussion

- 32 Les analyses montrent tout d'abord que si les deux populations d'apprenants hispanophones ne se distinguent pas des natifs au niveau du taux global de réalisation des liaisons dans le texte (59.43 % pour les apprenants sans séjour, 58.83 % pour ceux avec séjour et 56 % pour les natifs, ns<sup>13</sup>), les deux populations d'apprenants japonophones réalisent moins de liaisons que les natifs (44.63 % pour les apprenants sans séjour, 47.93 % pour ceux avec séjour vs 56 % pour les natifs,  $p < 0.01$ ). Ce résultat, pour les hispanophones, contraste également avec les travaux portant sur des apprenants anglophones (Howard, 2005 ; Mastromonaco, 1999 ; Thomas, 2002), dans lesquels le taux de réalisation des apprenants était généralement inférieur à celui des natifs. Si l'on se base uniquement sur ce premier résultat, les hispanophones sembleraient donc *a priori* plus performants que les deux autres populations d'apprenants, au niveau du taux global de liaison réalisées.
- 33 Nous avons également analysé les productions en fonction des contextes obligatoires définis par Delattre (1951) ainsi que par Durand & Lyche (2008). Les résultats montrent que, si l'on se base sur la classification de Delattre (1951), le taux de réalisation de la liaison est significativement moins élevé dans les quatre populations d'apprenants que chez les natifs. Pour les hispanophones, on trouve 90.00 % de réalisation de la liaison pour les hispanophones sans séjour et 92.29 % pour ceux avec séjour, les deux groupes ne se différenciant pas entre eux mais effectuant moins de liaisons obligatoires que les natifs (99.41 %,  $p < 0.05$ ). Les deux groupes d'apprenants japonophones réalisent également moins de liaisons que les natifs, avec 72.48 % pour les apprenants sans séjour, contre 85.10 % pour ceux avec séjour ( $p < 0.01$ ). À noter que les japonophones sans séjour réalisent significativement moins de liaisons obligatoires que le groupe avec

séjour ( $p < 0.05$ ), le séjour en milieu francophone semblant donc avoir un effet positif au niveau de la réalisation des liaisons obligatoires.

- 34 Si l'on examine ces mêmes chiffres en fonction de la classification de Durand & Lyche (2008), établie sur la base du corpus PFC – et qui semble donc correspondre, comme mentionné précédemment, davantage aux usages actuels en vigueur –, on observe que seuls les japonophones se différencient des natifs avec un taux de réalisation de 75.28 % pour le groupe sans séjour et 86.03 % pour celui avec séjour, contre 99.17 % pour les natifs ( $p < 0.001$ ). Si l'on se base sur ces quatre contextes uniquement, les hispanophones ne se distinguent plus des natifs, avec 93.33 % pour le groupe sans séjour et 93.24 % pour celui avec séjour (ns). Les hispanophones semblent donc également, au niveau des liaisons catégoriques, plus performants que les japonophones.
- 35 Comme les apprenants anglophones, les hispanophones et les japonophones réalisent des liaisons avec une consonne inattendue, très souvent présente orthographiquement dans le mot (p. ex. : [gʁɑ̃nɔnɔɛʁ] « grand honneur » ou [siʁkɥitabitɥɛl]<sup>14</sup> « circuits habituels »). On observe 9.46 % de productions de ce type chez les hispanophones sans séjour (dont 72 % le sont avec [n]) et 7.22 % chez les apprenants avec séjour (dont 52 % en [n]). Ces chiffres sont beaucoup moins élevés chez les japonophones, avec seulement 5.20 % chez les apprenants sans séjour (dont 68 % en [n]) et 2.10 % (dont 0 % en [n]) chez ceux avec séjour. Les apprenants japonophones ayant effectué un séjour en milieu francophone se rapprochent donc davantage des natifs, qui, eux, ne réalisent aucune liaison avec une consonne inhabituelle.
- 36 Les apprenants hispanophones et japonophones produisent également des liaisons non enchaînées (p. ex. : [ɔ̃n-ɑ̃n-a-vy] « on en a vu », [tʁu-veʁ-o] « trouver au ») dans des taux comparables à ceux observés chez les apprenants anglophones (7 % chez Mastromonaco, 1999, 8.5 % chez Thomas, 2002 et 8.85 % chez De Moras, 2011). On trouve ainsi 7.98 % de liaisons non enchaînées chez les hispanophones sans séjour et 7.51 % chez ceux avec séjour. Si l'on observe un taux comparable pour les apprenants japonophones sans séjour, avec 6.93 %, ce chiffre diminue drastiquement chez les japonophones avec séjour, avec seulement 1.68 % de liaisons non enchaînées. Les japonophones avec séjour se rapprochent donc à nouveau des natifs, qui eux ne produisent aucune liaison sans enchaînement dans la lecture du texte. Pour les autres trois groupes, ces résultats concordent avec les observations effectuées sur les apprenants anglophones dans les travaux précédents et semblent donc montrer que les difficultés ne peuvent être expliquées par la tendance à la syllabation fermée des langues germaniques, l'espagnol et le japonais ayant, comme le français, une prédilection pour les syllabes ouvertes.
- 37 Nos données révèlent également parfois un cumul des deux phénomènes, consonne orthographique et liaison non enchaînée (p. ex. : [gʁɑ̃n-ɔ-nɔɛʁ] « grand honneur »), et ce plus fréquemment chez les hispanophones. Il paraît difficile toutefois dans ce cas – et ce même s'il s'agit d'apprenants avancés – de déterminer s'il s'agit réellement d'une liaison non enchaînée ou plutôt de la prononciation d'une consonne graphique, les apprenants produisant également en conversation des formes telles que [gʁɑ̃n-pʁɔ-blɛm] « grand problème ». Afin de mieux comprendre ces formes, une analyse parallèle des réalisations des consonnes graphiques finales semble donc indispensable, notamment pour les apprenants dont la L1 dispose d'une très bonne correspondance entre prononciation et orthographe, ce qui est le cas de l'espagnol. À noter également que la non-différenciation entre les deux groupes d'apprenants hispanophones et les

natifs, au niveau du taux global de réalisation des liaisons de l'ensemble du texte, qui avait été interprétée dans un premier temps comme une meilleure performance des hispanophones par rapport aux autres populations d'apprenants (japonophones et anglophones) pourrait également s'expliquer par le fait qu'il ne s'agit pas réellement de liaisons réalisées mais plutôt de consonnes graphiques finales prononcées. Ce résultat surprenant doit donc être considéré avec prudence et être complété avec une analyse des réalisations de ces dernières.

- 38 Ces résultats mettent également en évidence l'intérêt d'élargir les analyses des productions de liaison à un plus grand nombre de populations d'apprenants, puisque le comportement des apprenants hispanophones et japonophones n'est pas identique. En effet, si les apprenants hispanophones ne se différencient pas des natifs au niveau du taux global de liaisons réalisées ainsi que du taux de liaisons catégoriques réalisées (basé sur les 4 catégories de Durand & Lyche, 2008), ils réalisent, dans des proportions comparables aux apprenants anglophones des études précédentes, des liaisons avec une consonne orthographique et des liaisons non enchaînées, avec parfois un cumul des deux phénomènes. Le schéma n'est pas identique pour les apprenants japonophones, puisque l'on observe d'une part des taux de réalisation globale de la liaison ainsi que des taux de liaisons catégoriques significativement moins élevés que les natifs. D'autre part, les taux de liaisons avec une consonne orthographique ainsi que ceux de liaisons non enchaînées semblent globalement légèrement inférieurs à ceux des hispanophones, avec surtout une amélioration notable en fonction du séjour : les apprenants japonophones ayant effectué un séjour prolongé en milieu francophone s'approchent des taux observés chez les natifs pour ces deux phénomènes.

## 5. Conclusion

- 39 Sur le plan didactique, ces résultats préliminaires montrent que les questions liées à la *microplanification* ne sont pas anodines, y compris pour des apprenants dont la L1 partage la prédilection du français pour les syllabes ouvertes, et qu'il semble nécessaire de les traiter de manière approfondie, notamment en les conscientisant, dans l'enseignement du FLE. De Moras (2011) observe en effet que le taux de liaisons enchaînées dans la catégorie des liaisons obligatoires augmente sensiblement après que les apprenants ont suivi un cours sur la liaison.
- 40 Les différences observées entre les apprenants hispanophones et japonophones soulignent également la nécessité d'élargir les analyses à d'autres populations. Les futurs travaux devront également intégrer la dimension inter-tâches, afin de déterminer l'impact de la tâche effectuée. Une analyse récente des productions de douze apprenants hispanophones montre en effet que les taux de liaisons avec une consonne orthographique ainsi que ceux de liaisons non enchaînées sont beaucoup moins élevés en conversation par rapport à ceux observés dans la lecture du texte. Cette tendance semble confirmée par des observations similaires effectuées chez les apprenants italo-phones (12 débutants et 12 avancés) examinés dans cette même étude (voir Racine *et al.*, 2014). Les deux phénomènes observés (liaisons avec une consonne orthographique et liaisons non enchaînées) semblent donc intrinsèquement liés aux processus cognitifs impliqués dans la tâche de lecture, qui sont différents de ceux mis en œuvre dans la parole spontanée. Enfin, ajoutons encore que les différences observées (inter-populations et inter-tâches) dans le cadre de la liaison rejoignent les

observations effectuées sur les aspects segmentaux (voyelles nasales et voyelles arrondies).

- 41 Ces premières données montrent donc que la méthodologie adoptée dans le projet IPFC devrait permettre de disposer de données de liaison diversifiées, en termes de L1 et de tâches, rendant possible des comparaisons à grande échelle. Cet approfondissement et cette diversification des travaux portant sur l'acquisition de la liaison en FLE, combinés à une meilleure prise en compte des usages réels en vigueur, rendue possible par le travail conséquent issu de la linguistique de corpus devraient en outre, à terme, permettre un renouvellement adéquat du matériel pédagogique pour l'enseignement de la liaison.

---

## BIBLIOGRAPHIE

- Abry, D. & Chalaron, M. (1994). *350 Exercices de phonétique*. Paris : Hachette.
- Abry D. & Chalaron M. (2011). *Les 500 Exercices de phonétique*. Paris : Hachette.
- Ågren J. (1973). *Enquête sur quelques liaisons facultatives dans le français de conversation radiophonique*. Uppsala : Acta Universitatis Uspaliensis.
- Baralo M. (1999). *La adquisición del español como lengua extranjera*. Madrid : Arco libros.
- Bassetti B. (2006). « Orthographic input and phonological representations in learners of Chinese as a Foreign Language », *Written Language and Literacy* 9 (1) : 95-114.
- Boersma P. & Weenink D. (2014). *Praat : doing Phonetics by Computer*. <http://www.praat.org>.
- Carranza M., Cucchiarini C., Llisterri J., Machuca M. J. & Ríos A. (2014). « A corpus-based study of Spanish L2 mispronunciations by Japanese speakers », *Proceedings of Edulearn14, 6<sup>th</sup> International Conference on Education and New Learning Technologies*. Barcelone, Espagne : 3696-3705.
- Coquillon A.-L. & Durand J. (2010). « Le français méridional : éléments de synthèse », in S. Detey, J. Durand, B. Laks & C. Lyche (éd.) *Les Variétés du français parlé dans l'espace francophone : Ressources pour l'enseignement*, Paris : Ophrys, 185-197.
- Côté M.-H. (2012). « Laurentian French (Quebec) : Extra vowels, missing schwas and surprising liaison consonants », in R. Gess, C. Lyche & T. Meisenburg (éd.) *Phonological Variation in French : Illustrations from three continents*. Amsterdam/Philadelphia : John Benjamins, 235-274.
- 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.
- De Jong D. (1994). « La sociophonologie de la liaison orléanaise », in C. Lyche (éd.), *French Generative Phonology : Retrospective and Perspectives*. AFLS/ESRI : 95-130.
- De Moras N. (2011). *Acquisition de la liaison et de l'enchaînement en français L2 : Le rôle de la fréquence*. PhD. Dissertation, University of Western Ontario.

- Debrock M. & Flament-Boistrancourt D. (1996). « Le corpus LANCOM : Bilan et perspectives », *ITL - Review of Applied Linguistics* 111-112 : 1-36.
- Delais-Roussarie E. & Yoo H.-Y. (2010). « The COREIL corpus : a learner corpus designed for studying phrasal phonology and intonation », in K. Deziubalska-Kolaczyk, M. Wrembel & M. Kul (éd.) *Proceedings of New Sounds 2010*. Poznan, 100-105.
- Delattre P. (1946). « Pour imiter un disque de français parlé », *The French Review* 20 (1) : 43-48.
- Delattre P. (1951). *Principes de phonétique française à l'usage des étudiants anglo-américains*. Middlebury College.
- Detey S. (2005). *Interphonologie et représentation orthographiques. Du rôle de l'écrit dans l'enseignement/apprentissage du français oral chez des étudiants japonais*. Thèse de Doctorat, Université de Toulouse Le Mirail.
- Detey S. (2012). « Coding and L2 phonological corpus : from perceptual assessment to non-native speech models – an illustration with French nasal vowels », in Y. Tono, Y. Kawaguchi & M. Minegishi (éd.) *Developmental and Crosslinguistic Perspectives in Learner Corpus Research*. Amsterdam/Philadelphia : John Benjamins, 229-250.
- Detey S. (2014). « Vers une évaluation par codage perceptif sur corpus de la production des liquides françaises /R/ et /l/ des apprenants japonais en singleton et en groupe consonantique », *Flambeau* 40, 1-17.
- 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, 11-13 décembre 2008.
- Detey S. & Racine I. (2012). « Les apprenants du français face aux normes de prononciation : quelle(s) entrée(s) pour quelle(s) sortie(s) ? », *Revue française de linguistique appliquée* 17(1) : 81-96.
- Detey S., Racine I., Eychenne J. & Kawaguchi Y. (2014a). « Corpus-based L2 phonological data and semi-automatic perceptual analysis : the case of nasal vowels produced by beginner Japanese learners of French », *Proceedings of Interspeech 2014*. Singapore, 14-18 Septembre : 539-544.
- Detey S., Racine I. & Kawaguchi Y. (2014b). « Des modèles prescriptifs à la variabilité des performances non-natives : les voyelles nasales des apprenants japonais et espagnols dans le projet IPFC », in J. Durand, G. Kristoffersen & B. Laks avec la collaboration de J. Peuvergne (éd.) *La Phonologie du français : Des normes aux périphéries* (Festschrift pour Chantal Lyche). Paris : Presses universitaires de Paris-Ouest, 197-226.
- Detey S., Racine I., Kawaguchi Y. & Zay F. (à paraître). « Variation among non-native speakers : Japanese and Spanish learners of French », in S. Detey, J. Durand, B. Laks & C. Lyche (éd.) *Varieties of Spoken French*. Oxford : Oxford University Press.
- Detey S., Racine I., Kawaguchi Y., Zay F., Buehler N. & Schwab S. (2010). « Évaluation des voyelles nasales en français L2 en production : de la nécessité d'un corpus multitâches », in F. Neveu, J. Durand, T. Klingler, S. Prévost & V. Muni-Toké (éd.) *Actes de CMLF'10*. Paris : ILF, 1289-1301.
- Durand J., Laks B. & Lyche C. (2009). *Phonologie, variation et accents du français*. Paris : Hermès.
- Durand J., Laks B. & Lyche C. (2014). « French phonology from a corpus perspective : the PFC programme », in J. Durand, U. Gut & G. Kristoffersen (éd.) *The Oxford Handbook of Corpus Phonology*. Oxford : Oxford University Press, 486-497.
- Durand J. & Lyche C. (2008). « French liaison in the light of corpus data », *Journal of French and Language Studies* 18 (1) : 33-66.

- Durand J. & Lyche C. (à paraître). « Approaching variation in PFC : the liaison level », in S. Detey, J. Durand, B. Laks & C. Lyche (éd.) *Varieties of Spoken French*. Oxford : Oxford University Press.
- Encrevé P. (1988). *La liaison avec et sans enchaînement*. Paris : Seuil.
- Eychenne J., Lyche C., Durand J. & Coquillon A.-L. (2014). « Quelles données pour la liaison en français : la question des corpus », in C. Soum-Favaro, A.-L. Coquillon & J. P. Chevrot (éd.) *La Liaison : approches contemporaines*. Berne : Peter Lang, 33-60.
- Eychenne J. & Paternostro R. (à paraître). « Analyzing transcribed speech with Dolmen », in S. Detey, J. Durand, B. Laks & C. Lyche (éd.) *Varieties of Spoken French*. Oxford : Oxford University Press.
- Fouché P. (1959). *Traité de prononciation française*. Paris : Klincksieck. 2<sup>e</sup> édition.
- Gut U. (2009). *Non-native Speech : a Corpus-based Analysis of Phonological and Phonetic Properties of L2 English and German*. Wien : Peter Lang.
- Gut U. (2014). « Corpus phonology and second language acquisition », in J. Durand, U. Gut & G. Kristoffersen (éd.) *The Oxford Handbook of Corpus Phonology*. Oxford : Oxford University Press, 286-301.
- Harnois-Delpiano M., Cavalla C. & Chevrot J.-P. (2012). « L'acquisition de la liaison en L2 : étude longitudinale chez des apprenants coréens de FLE et comparaison avec enfants francophones natifs », in F. Neveu, V. Muni Toke, P. Blumenthal, T. Klingler, P. Ligas, S. Prévost & S. Teston-Bonnard (éd.) *Actes du Congrès mondial de linguistique française (CMLF 2012)*. Paris : Institut de linguistique française, 1575-1589.
- Howard M. (2005). « L'acquisition de la liaison en français langue seconde – Une analyse quantitative d'apprenants avancés en milieu guidé et en milieu naturel », *CORELA, Numéros thématiques, Colloque AFLS*.  
<http://corela.edel.univ-poitiers.fr/index.php?id=1127>.
- Howard M. (2013). « La liaison en français langue seconde : une étude longitudinale préliminaire », *Language, Interaction, and Acquisition 4 (2)* : 190-231.
- Lacoste V., Herry-Bénit N. & Kamiyama T. (2013). « The Interphonology of Contemporary English : An introduction ». Conférence annuelle du projet PAC : *Spoken English corpora : from annotation to interphonologies*. Aix-en-Provence, 30 mai-1<sup>er</sup> juin 2013.
- Laks B. (2013). « Diachronie de la liaison dans la parole publique », *Journées PFC 2013 Phonologie du français contemporain : Regards croisés sur les corpus oraux*. Paris, 5-7 décembre 2013.
- Landron S., Paillereau N., Nawafleh A. et al. (2011). « Le corpus PhoDiFLE : un corpus commun de français langue étrangère pour une étude phonétique des productions de locuteurs de langues maternelles plurielles », *CJC Praxiling2011*. Montpellier.
- Lauret B. (2007). *Enseigner la prononciation du français, questions et outils*. Paris : Hachette.
- Mallet G. (2008). *La liaison en français : Description et analyses dans le corpus PFC*. Thèse de Doctorat, Université Paris Ouest Nanterre la Défense.
- Mastromonaco S.M. (1999). *Liaison in French as a Second Language. Doctor of Philosophy*. PhD Dissertation, University of Toronto.
- Meng H., Tseng C.-Y., Kondo M., Harrison A. & Viscelgia T. (2009). « Studying L2 suprasegmental features in Asian Englishes : a position paper », *Proceedings of Interspeech 2009*. Brighton, UK.



- Myles F. & Mitchell R. (2007). *French Learner Language Oral Corpora (FLLOC)*. University of Southampton, 13.3 GB.
- Neri A., Cucchiari C. & Strik H. (2006). « Selecting segmental errors in L2 Dutch for optimal pronunciation training », *IRAL* 44 : 357-404.
- Perdue C. (1993). *Adult Language Acquisition : Crosslinguistic Perspectives*. Volume 1, Field Methods. Cambridge : Cambridge University Press.
- Racine I. (2012). « Spanish learners' productions of French close rounded vowels : a corpus-based perceptual study » in Y. Tono, Y. Kawaguchi & M. Minegishi (éd.) *Developmental and Crosslinguistic Perspectives in Learner Corpus Research*. Amsterdam/Philadelphia : John Benjamins, 205-228.
- Racine I. (2014). « Une approche par corpus de la liaison chez les apprenants hispanophones de français langue étrangère : quelles conséquences pour l'enseignement du FLE ? », *Flambeau* 40, 18-37.
- Racine I. & Detey S. (2012). « La liaison dans IPFC : premiers regards sur les données hispanophones et japonophones », Colloque *Du français et de l'anglais aux langues du monde : variation, structure et théorie du langage*. Montpellier, 28-30 juin 2012.
- Racine I. & Detey S. (2015). « L'apprentissage de la liaison en français par des locuteurs non natifs : éclairage des corpus oraux », *Bulletin VALS-ASLA* 102.
- Racine I., Detey S., Buehler 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 K. Deziubalska-Kolaczyk, M. Wrembel & M. Kul (éd.) *Proceedings of New Sounds 2010 - Sixth International Symposium on the Acquisition of Second Language Speech*. Poznan : Adam Mickiewicz University, 367-372.
- Racine I., Detey S. & Kawaguchi Y. (2012). « Les voyelles /y-u/ dans IPFC : évaluation perceptive de productions natives, hispanophones et japonophones », *Actes de JEP 2012*. Grenoble : 385-392.
- 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 A. Kamber & C. Skupiens (éd.) *Recherches récentes en FLE*. Berne : Peter Lang, 1-19.
- Racine I., Paternostro R., Falbo C., Janot P. & Murano M. (2014). « La liaison chez les hispanophones et les italophones : du texte lu à la conversation », *Rencontres FLORAL 2014 « Corpus oraux et enseignement de la prononciation en FLE & Interphonologie et corpus oraux »*. Paris, 8-9 décembre 2014.
- Racine I., Zay F., Detey S. & Kawaguchi Y. (2011). « De la transcription de corpus à l'analyse interphonologique : enjeux méthodologiques en FLE », in G. Col & S. N. Osu (éd.) *Travaux linguistiques du CerLiCo 24 (Actes du 24<sup>e</sup> colloque du CERLICO « Transcrire, écrire, formaliser », Université de Tours, juin 2010)*. Rennes : PUR, 13-30.
- Thomas A. (2002). « La variation phonétique en français langue seconde au niveau universitaire avancé », *AILE* 17 : 101-121.
- Thomas A. (2010). « La complexité en FLE2 universitaire avancé », in U. Paprocka-Piotrowska, C. Martinot & S. Gerolimich (éd.) *Actes du colloque La complexité en langue et son acquisition*. Paris : Université Descartes, 149-152.
- Trouvain J. & Gut U. (2007). *Non-Native Prosody. Phonetic Description and Teaching Practice*. Berlin/ New York : Mouton de Gruyter.

Visceglia T., Tseng C.-Y., Kondo M., Meng H. & Sagisaka Y. (2009). « Phonetic aspects of content design in AESOP (Asian English Speech cOrpus Project) », *Proceedings of Oriental-COCOSDA*. Urumuqi, Chine.

Vendelin I. & Peperkamp S. (2006). « The influence of orthography on loanword adaptations », *Lingua* 116 : 996-1007.

Wauquier S. (2009). « Acquisition de la liaison en L1 et L2 : stratégies phonologiques ou lexicales ? », in J.-Y. Dommergues (éd.), *Phonétique, bilinguisme et acquisition, Aile-Lia 2* : 93-130.

Zampini M. L. (2008). « L2 speech production research : Findings, issues, and advances. », in J. G. Hansen Edwards, & M. L. Zampini (éd.) *Phonology and Second Language Acquisition*. Amsterdam : John Benjamins, 219-249.

## NOTES

1. Bien que les termes « langue seconde » (L2) et « langue étrangère » (LE) ne soient pas synonymes (Baralo, 1999), ils sont utilisés ici de manière interchangeable, tout comme les termes « langue maternelle » et « langue première » (L1).
2. Parmi ceux-ci, on peut par exemple mentionner le corpus ESF (*European Science Foundation Second Language*, Perdue, 1993), le projet LANCOM (*LANgue et COMmunication*, Debrock & Flament-Boistrancourt, 1996) et le projet FLLOC (*French Learner Language Oral Corpora*, Myles & Mitchell, 2007).
3. [www.cbllle.tufs.ac.jp/ipfc/](http://www.cbllle.tufs.ac.jp/ipfc/)
4. Une description détaillée des différentes enquêtes ainsi que la liste des publications liées sont disponibles sur le site du projet : (<http://cbllle.tufs.ac.jp/ipfc/>).
5. Tous les codages sont effectués par des locuteurs natifs. Pour certains phénomènes, un double - voire triple - codage à l'aveugle est effectué. Une fonction de *Dolmen* permet de comparer aisément les codages effectués par des codeurs différents. Pour un exemple plus détaillé de cette procédure de double codage et du traitement des données, voir Detey *et al.* (2014a).
6. Les réflexions et analyses présentées dans cette section ont bénéficié des commentaires des participants à une conférence organisée à Montpellier, en 2012 (Racine & Detey, 2012), ainsi que de ceux des relecteurs anonymes du manuscrit de Racine (2014). Qu'ils en soient ici remerciés.
7. Ce nombre est toutefois soumis à la variation diatopique, le français laurentien connaissant par exemple des liaisons en [l] (Côté, 2012).
8. À noter, toutefois, que cette dénasalisation n'est pas systématique (dans « mon animal », la voyelle reste nasale, de même que dans « en effet » ou « un entrepôt ») et qu'elle est également soumise à la variation diatopique, étant en effet plus fréquente par exemple dans le français méridional (*cf.* Coquillon & Durand, 2010).
9. Pour un résumé récent des apports de ces différents corpus, voir Eychenne *et al.* (2014).
10. Selon Laks (2013), ce changement coïnciderait avec le mouvement de mai 68.
11. Certains auteurs ont tout de même constitué un groupe témoin de natifs qui sert de référence. C'est le cas de De Moras (2011) par exemple.
12. Les deux évaluateurs sont des linguistes, enseignants de FLE spécialisés en phonétique.
13. Une analyse ANOVA montre que cette différence n'est pas significative (ns).
14. Les transcriptions phonétiques ne cherchent pas à reproduire fidèlement la prononciation des apprenants, ce point n'étant pas le propos ici.

---

## RÉSUMÉS

Ce chapitre présente le projet « Interphonologie du français contemporain » (IPFC), qui vise à constituer et analyser une large base de données de français langue étrangère produit par des apprenants de diverses L1. Nous illustrons ensuite la méthodologie adoptée dans le projet à travers le phénomène de la liaison. Après avoir exposé les enjeux de la liaison pour le français L2, nous présentons une étude préliminaire des réalisations de liaisons par des apprenants hispanophones et japonophones en lecture de texte et montrons en quoi le projet IPFC est susceptible de fournir des données diversifiées en termes d'apprenants et de tâches, ce qui devrait, à terme, permettre un renouvellement des données pour l'enseignement de la liaison en français L2.

In this chapter, we present the project “Interphonology of Contemporary French” (IPFC). The aim of the project is to constitute and analyse a large database of L2 French, produced by learners of different origins. The methodology adopted in the project is then illustrated through the phenomenon of liaison. First we describe the challenges of liaison for L2 French learners and then we present a preliminary study of the realisations of liaison produced by Spanish and Japanese learners of French in the text reading task. We conclude by showing that the IPFC project provides an adequate framework to deepen our knowledge of L2 liaison, which in turn should help us to renew the pedagogical material used to teach liaison in L2 French.

## INDEX

**Mots-clés** : corpus, interphonologie, français parlé, français langue étrangère, acquisition phonologique, liaison, IPFC, apprenants hispanophones, apprenants japonophones

**Keywords** : corpus, interphonology, spoken French, French as a foreign language, phonological acquisition, liaison, IPFC, Spanish learners, Japanese learners

## AUTEURS

**ISABELLE RACINE**

ELCF, Université de Genève

**SYLVAIN DETEY**

SILS, Université Waseda