

Vers un Nouvel Atlas Linguistique de la France : aspects de méthodologie sociolinguistique et dialectologique

Hall, Damien

University of Kent (Royaume-Uni)
d.hall@kent.ac.uk

1 Résumé

Cet article décrit un projet de dialectologie qui a pour but de produire une étude phonétique – phonologique des accents urbains dans le français du nord de la France. Le projet s'intitule en anglais « Towards A New Linguistic Atlas of France » (TANLAF), ce qui peut se traduire « Vers un Nouvel Atlas Linguistique de la France ». ¹ Des projets semblables n'étant habituels ni en dialectologie ni en sociolinguistique françaises, nous commencerons par un compte-rendu des raisons du projet, et ferons par la suite une description de sa méthodologie. Nous terminerons par un appel aux suggestions de la part de la communauté des linguistes qui travaillent sur le français et la France.

2 La linguistique du projet

2.1 Pourquoi faire une telle étude ?

Ce projet représente certes un départ de l'approche traditionnelle de la dialectologie et la sociolinguistique en France. Sauf exceptions, la sociolinguistique française est beaucoup plus interactionnelle que variationniste, se situant beaucoup plus au niveau macrosociolinguistique qu'au niveau micro : elle fait appel aux interactions entre une langue, ses locuteurs et la/les société(s) où elle est parlée, beaucoup plus qu'à la variation linguistique entre individus. Le présent projet, qu'on pourrait appeler de la dialectologie sociolinguistique, cible davantage la variation linguistique entre individus. Nous nous pencherons un moment sur nos raisons de penser qu'une étude variationniste de grande échelle complètera la littérature sociolinguistique française de façon utile. Ensuite, dans les sections 2.2 *et seq.*, nous approfondirons l'explication du projet en examinant de près la description qui en est énoncée ci-dessus : « une étude phonétique – phonologique des accents urbains dans le français du nord de la France ».

Nous pensons que cette étude enrichira la littérature sur la langue française parce que, d'une certaine façon, l'étude se trouve dans la droite ligne de la tradition dialectologique française. Depuis l'*Atlas Linguistique de la France* (Gilliéron et Edmont 1902-10), la France est connue à juste titre comme le berceau de la dialectologie moderne. La tradition de cet atlas monumental continue jusqu'à ce jour dans la série des *Atlas Linguistiques de la France par régions*. Cela va sans dire que ces atlas sont d'une valeur inestimable aux linguistes et ethnologues qui s'intéressent à la France ; leurs techniques ne sont pourtant pas les nôtres. Au cours des cent derniers ans s'est établie la légende (véritable ou non : Brun-Trigaud *et al.* 2005 : 19-20) du dialectologue Edmont qui parcourait les campagnes françaises à vélo, à la recherche de ses prononciations et de ses dialectes, qu'il notait soigneusement ; telle est aussi (plus ou moins) la technique des dialectologues plus récents qui font les recherches des atlas par régions. Nous continuons cette tradition – nous dresserons le plan de la variation phonétique et phonologique que nous trouverons dans les villes d'intérêt – mais nous le ferons avec l'aide des techniques modernes d'enregistrement et d'analyse phonétique, qui n'étaient pas disponibles – ou du moins n'étaient pas faciles – pour les dialectologues des années passées.

Une étude à grande échelle de la phonologie de la France, menée par l'analyse d'enregistrements, nous révélera peut-être aussi des choses que nous n'aurions pas pu déceler sans l'aide de logiciels phonétiques. Il existe des différences tellement subtiles que l'oreille n'est pas consciente de les avoir perçues, tellement subtiles que nous ne sommes certainement pas en mesure de les décrire, même après les avoir entendues, à moins de les entendre plusieurs fois et de nous faire indiquer où se trouve la différence, pour que nous puissions nous concentrer là-dessus. La section 2.6, ci-dessous, donne quelques détails de l'histoire de ce type de recherche en dialectologie hors de France.

2.2 « Une étude phonétique – phonologique ... »

Tout comme la technique de l'*Atlas Linguistique de la France*, la technique de notre étude consiste aussi en un parcours des sites qui nous intéressent, mais à ces différences près :

- nous enregistrons nos entretiens
- (pour la plupart des voyelles) nous les analyserons phonétiquement par la suite, et constituerons à partir de ces analyses des comptes-rendus des systèmes phonologiques des locuteurs.

Une part des analyses sera donc faite à l'oreille, comme les classifications des atlas faits jusque maintenant – il s'agira des voyelles dont nous ne connaissons pas encore assez bien les attributs phonétiques (les nasales, par exemple) ; mais les voyelles orales et quelques consonnes d'intérêt seront analysées par ordinateur.

Par ces différences, notre projet part de la tradition dialectologique française, mais il se réunit, par contre, avec trois autres lignes de recherche linguistiques :

- la recherche phonologique sur l'organisation des espaces vocaliques, commencée notamment par Martinet (1955, 1960, 2008)
- la tradition de phonétique instrumentale de laboratoire en France (pour ne rien dire de la tradition instrumentale nord-américaine, dont fait partie le Canada)
- la tradition labovienne de sociolinguistique et dialectologie instrumentales.

Le travail de Martinet (entre autres) a fait valoir l'importance de l'équilibre dans les systèmes phonologiques :

on peut [...] supposer que [les éléments distinctifs d'une langue] tendront à être aussi différents les uns des autres que le permettent les organes qui contribuent à leur production [...] On dira qu'un système évoluera tant que ne sera pas établie l'**équidistance entre les phonèmes** qu'il comporte.

Martinet 2008 : 202 (emphasis dans le texte d'origine)

Nous notons que l'idée d'équilibre dans un système (en l'occurrence vocalique, dans notre projet) n'exclut pas de différentes configurations de l'espace vocalique. Plus loin à la même page, Martinet poursuit son point par l'exemple d'un système à trois voyelles antérieures, /i e a/, avec une « case vide » à la position mi-ouverte : « le phonème /é/ ne tarderait pas à se placer à égale distance de /i/ et de /a/ ». ² Un tel système, à trois voyelles antérieures, se trouve en espagnol (comme le note Martinet à la même page) et en français méridional (comme il a été noté par beaucoup de linguistes, dont Martinet aussi). Il s'oppose à un système à quatre voyelles antérieures (/i e ε a/) – équilibré lui aussi – tel qu'on trouve, par exemple, en français de référence.

Les processus de variation phonétique peuvent pourtant aboutir autrement qu'en équilibre, notamment par la fusion de deux phonèmes. Si nous continuons à parler des phonèmes /e ε/ du français de référence, de nos jours, il y a des variétés autres que le français méridional qui ne font pas de différence phonémique entre eux. Lyche (2010 : 150-1) synthétise la situation :

L'opposition /e/-/ε/ en syllabe accentuable ouverte semble menacée chez certains locuteurs même si les instituteurs âgés par exemple, la maintiennent [...]. De

nombreuses descriptions rapportent soit l'existence d'une voyelle intermédiaire, soit le respect de la loi de position [...].

La voyelle intermédiaire n'est pas non plus le seul résultat possible. Par exemple, Hall (2008 : ch.4) trouve une fusion de /e/ et /ɛ/ en Normandie, et non vers une réalisation intermédiaire, mais vers une réalisation haute, plutôt près de /e/.

Au-delà du français, aussi, les langues du monde ont connu beaucoup de fusions phonémiques historiques, et continuent à en connaître de nos jours. Labov (1994 : 115-154) passe en revue beaucoup de fusions historiques au cours de sa revue des développements phonologiques diachroniques. Nous pouvons nous tourner vers le travail de Labov aussi pour des comptes-rendus de beaucoup de fusions variables synchroniques en anglais d'Amérique du Nord, et vers celui de Wells pour des fusions en anglais britannique. La plus célèbre des fusions nord-américaines est sans doute la soi-disant « fusion de *cot* et *caught* » (de /a/ et /ɔ:/, LOT et THOUGHT selon le système de Wells, fusionnés vers [a]), qu'on trouve en anglais du Canada, de l'Ouest des États-Unis et de plusieurs autres aires dialectales. D'autres fusions s'opèrent seulement devant certains autres phonèmes (/ej/ et /ɛ/ devant /l/, comme dans *fail* « rater » et *fell* « tomber » (passé simple), prononcés tous les deux [fɛl] par les locuteurs qui font cette fusion). Pour la Grande-Bretagne, beaucoup de locuteurs écossais font aussi la fusion de LOT et THOUGHT, mais /ɒ/ (LOT) et /ɔ:/ (THOUGHT) pour eux sont réalisés [ɔ:] (donc une fusion vers la voyelle de THOUGHT, l'inverse de la situation pour la plupart des Américains qui font cette fusion : Wells 1982). Des locuteurs de la région de Liverpool, G.-B., font la fusion de /ʌ:/ (NURSE) et /ɛ:/ (SQUARE), mais la voyelle qu'ils produisent pour les deux phonèmes peut être [ʌ:] ou [ɛ:], dépendant de leur région précise d'origine (Watson & Mattock 2009). L'intention de tous ces exemples est de souligner qu'il est possible que les espaces vocaliques soient organisés différemment selon les accents, ce que l'Hexagone ne contredit pas, mais aussi que ces différences d'organisation peuvent être plutôt subtiles et variées. Jusqu'à date, l'on n'a pas trouvé beaucoup de variation phonologique à ce niveau de subtilité en français, mais l'un des buts de ce projet est de la chercher.

Pour ce qui est de la phonétique, les contributions françaises à la recherche phonétique de laboratoire, des jours de Delattre (1948) à nos jours, montrent que la France compte beaucoup d'expertise en la technique que nous utiliserons pour faire nos observations. Nous comptons ainsi contribuer à cette expertise par nos travaux, qui viennent pourtant d'une optique différente. En sociolinguistique et dialectologie, les travaux de Labov et ses collègues sur bientôt cinquante ans – mais surtout l'*Atlas of North American English* (Labov, Ash et Boberg 2006) – montrent que les analyses presque strictement instrumentales peuvent déceler des régularités phonétiques et phonologiques au sein de grandes communautés linguistiques. Des cartes des régularités phonétiques mesurées peuvent même être plus régulières que des cartes faites selon des transcriptions auditives de linguistes. Les linguistes en question ont beau être bien formés, mais, si les oreilles de deux linguistes différents leur indiquent deux transcriptions différentes pour un même son, la carte reflètera cette différence, qu'elle existe réellement ou non.

La technique phonetico-phonologique de Labov et ses collègues a été usitée en France, mais peu. Elle a néanmoins prouvé son utilité dans plusieurs travaux sur des français régionaux : citons à titre d'exemple Lennig (1978), la première étude labovienne d'un français régional, qui a décelé un changement en chaîne des voyelles dans l'espace vocalique parisienne dont l'existence n'aurait pas été soupçonnée sans des mesures instrumentales, et Arnaud (2006), qui a documenté les mouvements des voyelles des natifs d'une ville du Jura. Il est donc à espérer que nos conclusions apporteront des données à l'étude instrumentale des voyelles du français de France.

2.3 « ... des accents urbains ... »

Les pratiquants de la dialectologie traditionnelle de la France diraient qu'une enquête sur les accents urbains du nord de la France apporterait peu : les enquêtes traditionnelles sur la variation linguistique en

France se sont concentrées sur les représentants les plus enracinés des variétés rurales, les « langues de France » anciennes, pour faire des comptes-rendus de ces langues avant qu'elles ne s'éteignent, et les villes sont censées renfermer une variation linguistique comparativement moindre. Nous ne nions pas que d'importantes variations seraient décelées par un parcours en France rurale, mais au cours d'un tel parcours nous risquerions aussi de tomber sur des différences au sein non du français mais des variétés romanes autochtones locales des régions (picard, normand, champenois ...). Tel est l'objet des *Atlas Linguistiques de la France par régions*, mais notre objet est différent : nous ciblons les différences au sein du français même. Nous souhaitons donc éviter toute influence qui pourrait mener à croire qu'une différence linguistique décelée par nos études serait due au fait qu'un locuteur parle aussi une variété locale. Pour éviter autant que possible ces variétés locales, il nous semble que c'est aux villes qu'il faut aller.

Deuxièmement, les changements phonétiques et phonologiques montrés par l'œuvre labovienne en Amérique du Nord commencent le plus souvent dans les villes et se répandent autour de celles-ci par la suite. Les raisons linguistiques et sociologiques pour lesquelles ces différences ont *débuté* ne sont pas connues, pour la plupart d'elles. Par exemple, le fait qu'un /æ/ devienne plus tendu, plus haut, devant certaines consonnes, dans l'anglais de la plupart de l'Amérique du Nord, mais non pas (autant) en anglais britannique (Wells 1982 : 477-9), ne peut pas s'expliquer par l'origine de la population, les populations de début des deux endroits étant plus ou moins les mêmes (c'est-à-dire des Britanniques). Nous ne cherchons donc pas à expliquer les variations que nous trouvons au sein des villes, mais seulement à en faire le plan. Si une explication évidente se présente, nous la noterons, mais à ce stade les explications ne seront pas notre premier but.

Nous avons bien remarqué que les raisons des *débuts* de beaucoup de changements linguistiques ne sont pas connues. Mais nous connaissons bien la raison de base de la *propagation* des changements : le contact humain. Plus une personne ou une communauté est en contact avec une autre, plus ces deux personnes ou communautés en viendront à parler de la même manière (Labov 2001 : xv et 2001 : 19 donnent un aperçu de l'histoire de cette idée). Les gens sont évidemment en contact plus étroit au sein de communautés urbaines denses (plus particulièrement quand ils travaillent ensemble) qu'entre communautés rurales plus éparpillées ; c'est donc dans des communautés urbaines que nous devons chercher des accents régionaux unifiés. La structure sociale et économique de la France renferme plusieurs types de communauté urbaine ; la section 3.1.1 considère le type de communauté à prendre comme base d'un projet de ce type.

2.4 « ... dans le français du nord de la France »

La section 2.3 établit que le présent projet cible non les langues régionales de France, mais les accents régionaux des Français *quand ils parlent en français*.³ Pour certains, donc, le fait que nous nous concentrons de prime abord sur le français du nord de la France présenterait un autre problème pour notre étude, car le français des classes moyennes des villes du nord de la France n'est pas censé beaucoup varier géographiquement (c'est-à-dire à l'intérieur du nord) : si variation il y a, elle est censée être plutôt sociale que géographique. Tel est le consensus de la plupart des recherches sur le français du nord de la France ; pour un survol de beaucoup d'études qui prennent ce point de vue, voir Armstrong et Pooley (2010).

Il y a pourtant des études qui ne prennent pas ce point de vue. Parmi les plus importantes, par ordre chronologique, nous pouvons citer :

- Martinet (1945 / 1971), fait par auto-déclaration de prononciation ;
- Martinet et Walter (1973), fait par évaluation des prononciations des enquêtés par les linguistes André Martinet et Henriette Walter ;
- Walter (1982), fait de la même façon par Henriette Walter seule ;
- le programme PFC (Durand, Laks et Lyche 2009 ; <http://www.projet-pfc.net>), qui étudie depuis presque dix ans les prononciations régionales de toute la Francophonie, par le biais d'enregistrements.

Armstrong et Pooley disent explicitement (2010 : 150) qu'ils rejettent la contention de Walter (1982), selon laquelle la variation géographique l'emporterait sur la sociale en France. Nous ne nions pas que la variation sociale au nord de la France soit considérable, mais nous croyons que les méthodes de la plupart des études françaises ne sont pas à même de dévoiler toute la variation géographique, subtile et phonétique, qui peut exister. Qui plus est, au cours de nos entretiens dans diverses villes du nord de la France, les interviewés eux-mêmes nous ont exprimé leur surprise face à la contention qu'il n'existe pas beaucoup de variation parmi les villes du nord de la France. Ils s'imaginent mal que diverses villes du nord de la France n'aient pas leurs accents particuliers.

2.5 La dialectologie auditive en France

Comme il est noté ci-dessus, la grande tradition française de dialectologie se fonde uniquement sur des rapports auditifs faits par des spécialistes. Ce point n'est pas fait pour insinuer que ces études ont moins de valeur – loin de là : leurs résultats et leur tradition, valorisés par les plus grands spécialistes depuis cent ans, montrent la valeur durable de ce travail.

Pourtant, avec tout le respect qu'on doit aux linguistes qui ne croient pas que le français des classes moyennes de la France septentrionale varie beaucoup : les quatre études citées dans la section 1.4 montrent bien que, même quand toute autre source potentielle de variation sociolinguistique est exclue, de la variation purement géographique persiste. C'est surtout le programme PFC qui le montre pour les générations présentes, mais Martinet l'a montré pour des officiers de l'Armée Française de la deuxième guerre mondiale ; Martinet et Walter l'ont montré pour des Parisiens d'adoption, venant de divers endroits en France, dans les années 1970 ; et Walter l'a montré pour des Français dans les diverses régions de France en 1982. S'il n'y avait pas eu de variation géographique des accents français, ces trois projets n'auraient rien apporté, mais nous voyons bien qu'ils ont des résultats : les cartes de Martinet (1945 / 1970) et les chartes des voyelles de Walter (1982), qui varient selon la région de l'enquête, nous en donnent la preuve.

Martinet (1945) a divisé la France en douze aires – le Midi, et onze aires au nord du Midi, donc dans la région des langues d'oïl – et ses sujets lui ont bien rapporté des différences à l'intérieur de l'aire d'oïl. Ces différences ont été relevées au cours d'une étude parmi des membres d'une classe sociale dont on dit que leur français ne varie pas ; ou du moins on dit qu'il ne varie pas quand ils ne parlent pas de manière très familière. C'étaient des officiers de l'Armée Française, à une époque quand les forces armées puisaient leurs officiers dans les classes moyennes et supérieures, parmi des personnes censées être bien formées ; il n'y avait donc pas beaucoup de chances qu'un soldat d'un rang inférieur soit promu. Les personnes interrogées étaient tous aussi des hommes d'un certain âge, étant donné qu'ils auraient tous eu une carrière militaire de quelques années. Les interrogés de Martinet (1945) étaient donc tous assez semblables des points de vue de classe sociale, de sexe et d'âge ; si nous trouvons quand même des différences linguistiques entre eux, ces similarités en réduisent les explications, et nous forcent à prendre au sérieux les différences géographiques dans l'échantillon. L'échantillon est contrôlé au maximum, ce qui représente des conditions optimales pour l'uniformité linguistique, si uniformité il y a ; mais des différences sont quand même apparentes, et, puisque ces différences ne peuvent pas être imputées à des différences de classe sociale, de sexe et d'âge, nous devons les imputer à des accents régionaux.

Nous ne devrions pas quand même oublier que les données de Martinet ont été collectionnées il y a soixante-dix ans, et qu'elles sont composées d'auto-déclarations de prononciations, en réponse à des questions comme « Quelle prononciation vous paraît la plus naturelle [...] de *j'me dis*, *je m'dis*, *je me dis* ? » et « Si vous faites une différence entre *patte* et *pâte*, est-ce une différence a) de timbre ? b) de longueur ? » (les questions 5a et 8 respectivement du questionnaire : Martinet 1971 : 11). Nous savons indépendamment que l'auto-déclaration peut ne pas être fiable ; des descriptions de phénomènes ont été basées sur la différence entre l'auto-déclaration de prononciation et la réalité phonétique (la quasi-fusion : Labov 1994 : 20, 349 *et seq.*). Ce manque de fiabilité mis à part, nous devons avouer que l'étude de Martinet a bien révélé des différences géographiques de prononciation à l'intérieur de l'aire des langues d'oïl.

Les trente années qui ont suivi la publication de Martinet (1945) ont vu publier comparativement peu d'études du français de France à l'échelle nationale (mettons à part les auteurs des quelques études qui ont paru : Gaston Tuaille, Georges Straka, Gérard Taverdet et Léon Warnant, par exemple). Il faut quand même noter que ces années ont été riches en études du français de régions bien définies, en particulier Paris ; et elles ont été riches aussi en études des variétés romanes autochtones de la France, en particulier celles du Midi. Mais nous nous attarderons en plus de détail sur l'œuvre d'Henriette Walter, surtout sur son livre de 1982, *Enquête Phonologique et Variétés Régionales du Français*. La plupart des données de cette étude ont été collectionnées dans les années soixante-dix. Avant le livre de 1982, nous connaissions déjà le style du travail de Walter, surtout par *La Dynamique des Phonèmes dans le Lexique du Français Contemporain* (1976) et par le livre produit en collaboration avec André Martinet, *Dictionnaire de la Prononciation du Français dans son Usage Réel* (1973). L'on a critiqué ces livres pour leurs échantillons de locuteurs, limités et en termes des nombres – le dictionnaire de 1973 en avait dix-sept – et en termes de géographie : le même dictionnaire a utilisé uniquement des Parisiens. Ces faits ne limitent pas du tout la valeur de leurs études, surtout parce que l'un de leurs buts était de démontrer que Paris constitue un *melting-pot* où, justement, tout le parler n'est pas uniforme : « le grand creuset où se mêlent et se fondent les caractéristiques particulières à chaque région » (Walter 1982 : 10). Nous devons pourtant avouer que, même si Martinet et Walter (1973) et Walter (1976) utilisent des locuteurs dont les *racines* se trouvaient partout en France, ils n'ont pas fait figurer des locuteurs qui *venaient* de partout en France et *y vivaient toujours*.

Enquête Phonologique et Variétés Régionales du Français (1982) a donné à ces locuteurs toute leur place dans la littérature sur la variation en français. Cette étude a divisé l'Europe francophone en 35 régions – dont 28 en France – représentées par 111 informateurs, dont 99 en France. Ce livre constitue la première étude de l'ensemble de la population francophone de l'Europe, qui avait pour but d'interroger des personnes qui avaient grandi et vivaient dans les régions d'intérêt. L'âge moyenne des informateurs était pourtant toujours assez élevée, et beaucoup d'entre eux habitaient en milieu rural ; donc (commentaire qui ne veut pas du tout dénigrer l'étude novatrice et cruciale de Walter) on peut dire que le résultat regardait vers l'histoire de la langue plutôt que vers son avenir. Le résultat est quand même beaucoup plus informé par la sociolinguistique que n'avait été aucun des études précédentes qui ont couvert l'ensemble de la France.

Dans toute étude de la variation géographique en français, le programme « Phonologie du Français Contemporain » (Durand, Laks et Lyche 2009) mérite une mention, surtout à cause de son ampleur et sa systématisme. Les dirigeants du programme ont donné la meilleure synthèse de ses buts quand ils ont présenté le projet en 2002 :⁴

PFC est un projet qui vise à décrire la prononciation du français dans sa diversité géographique, sociale et stylistique. A partir d'un protocole d'enquête uniforme, un groupe international d'une trentaine de chercheurs et leurs étudiants est impliqué dans la constitution d'un vaste corpus de français parlé à travers le monde. Grâce à des méthodes d'analyse et à des outils communs, le projet a pour ambition d'offrir une vision globale de la phonologie du français. [...] La volonté du projet PFC est de créer une base de données comportant entre 300 à 500 locuteurs étudiés à partir d'une méthodologie commune et permettant des comparaisons intéressantes sur un échantillon représentatif de variétés du français contemporain.

Durand, Laks et Lyche (2002 : 5-6)

Le lecteur qui connaît quelque peu les activités du programme PFC se posera pourtant une question : pourquoi avons-nous inclus ces activités dans la section « dialectologie auditive », alors que le programme avait au moins les deux nouveautés d'enregistrer tous ses entretiens, et d'analyser des parties des enregistrements par des logiciels faits sur mesure ? Nous ne nions pas du tout que PFC soit novateur de ces deux façons-là et de beaucoup d'autres, et nous ne critiquons pas du tout le programme – d'autant moins que nous y avons contribué et nous avons l'intention de poursuivre cette contribution à l'avenir. Mais nous incluons le programme PFC sous la rubrique « dialectologie auditive » parce que le noyau de l'analyse qu'il mène grâce aux logiciels est fait à l'oreille par le chercheur, et codifié à l'aide des logiciels : il s'agit du codage de l'usage de schwa et des liaisons. Le programme fait exprès pour dire que

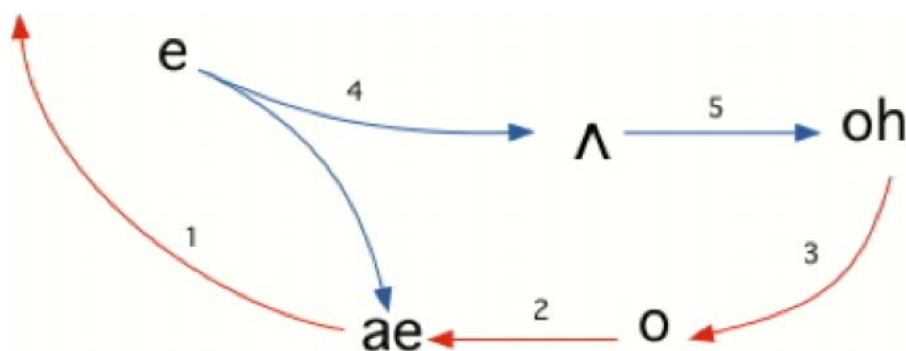
des chercheurs individuels peuvent utiliser les enregistrements PFC pour mener des analyses plus poussées sur certains aspects du parler, avec des codages ou à l'oreille ou par ordinateur, et quelques-uns l'ont fait (voir par exemple Hall (2007) pour le système vocalique d'une Rouennaise, et beaucoup d'autres articles téléchargeables à la même adresse, surtout ceux qui traitent de la prosodie). Nous espérons que d'autres chercheurs voudront en faire autant à l'avenir. Qui plus est, le programme PFC est bien géographique, mais il n'est pas encore systématiquement géographique : c'est-à-dire qu'une grande couverture géographique est ciblée, mais il n'y a pas de plan systématique pour celle-ci. Le corpus PFC en ligne, qui compte actuellement 394 locuteurs à 36 sites dans 11 pays (PFC 2004-2008), s'est constitué à partir d'enregistrements faits par des chercheurs volontaires où qu'ils se trouvent, et cela fait que nous puissions bien en faire de la dialectologie, mais nous risquons de trouver qu'un tel site où nous voudrions faire une comparaison n'a pas encore été couvert.

Si donc ces études, et en particulier Martinet (1945 / 1971) et Walter (1982), ne nous ont pas fourni les « données de base » sur la variation géographique à travers la France, nous devons nous demander à quoi ressembleraient ces données. Nous pensons que les données dont nous disposons déjà constituent une partie essentielle de ce qu'il faut, mais il y a une deuxième partie importante : l'obtention de données acoustiques comparables de partout en France, et leur analyse phonétique. De telles données instrumentales – des voyelles dont les fréquences constitutrices sont mesurées, afin de les décrire de façon scientifiquement précise – viendront bien s'ajouter à ce que nous avons déjà, à cause d'un type de variation géographique que nous révèle le travail phonétique de la tradition de Labov.

2.6 La dialectologie instrumentale

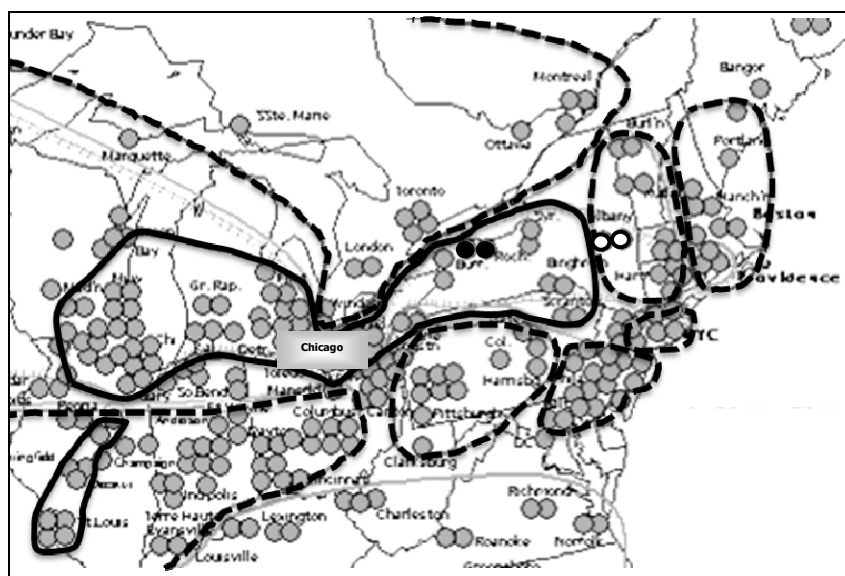
À part les études de variation phonétique / phonologique qui se fondent sur l'observation auditive, donc, il existe une autre catégorie d'évidence de variation phonétique : la variation qui se trouve lors de l'analyse phonétique *mais qui n'est pas évidente à l'oreille à la première écoute*. Une fois que l'analyse phonétique (par ordinateur) a révélé ce type de variation, elle devient souvent plus facile à repérer à l'oreille, car l'analyse par ordinateur nous a montré « quoi chercher ». Des études de l'anglais d'Amérique du Nord trouvent depuis longtemps ce type de variation.

Dans la région américaine autour des Grands Lacs (villes de Chicago, Detroit, Buffalo, Rochester ...), il existe un changement en chaîne, le « Northern Cities Shift », marqué par plusieurs changements vocaliques qui semblent liés (Graphique 1 ; Labov, Yaeger & Steiner 1972 ; Labov 1994 : 184 *et seq.*)⁵ (« Northern Cities Shift » signifie « Changement en chaîne (vocalique) des Villes du Nord » ; nous nous y référons par le sigle anglais « NCS ».)



Graphique 1
La phonologie du « Northern Cities Shift »





Des régions limitrophes de la région du plein changement (notamment au centre de l'État de New York) ont une version réduite du plein changement, par lequel les voyelles impliquées dans le plein changement ne changent pas toutes, et celles qui changent le font d'un moindre degré. Par exemple, dans la ville d'Albany, limitrophe de la région NCS, /æ/ s'avance dans l'espace vocalique par moins d'Hertz qu'il ne le fait à Rochester, au centre de la région NCS. À Rochester, donc, *cat* 'chat' (phonémiquement /kæt/) se dit [kæt], alors qu'à Albany le mot se dit plutôt [kæ̃t]. Dinkin (2009) nous donne plus de détails phonétiques de la variation entre la région du plein changement et celles d'un changement réduit ; le graphique 2 montre l'aire dialectale du NCS, ainsi que les aires limitrophes de celle-ci.⁶



Graphique 2

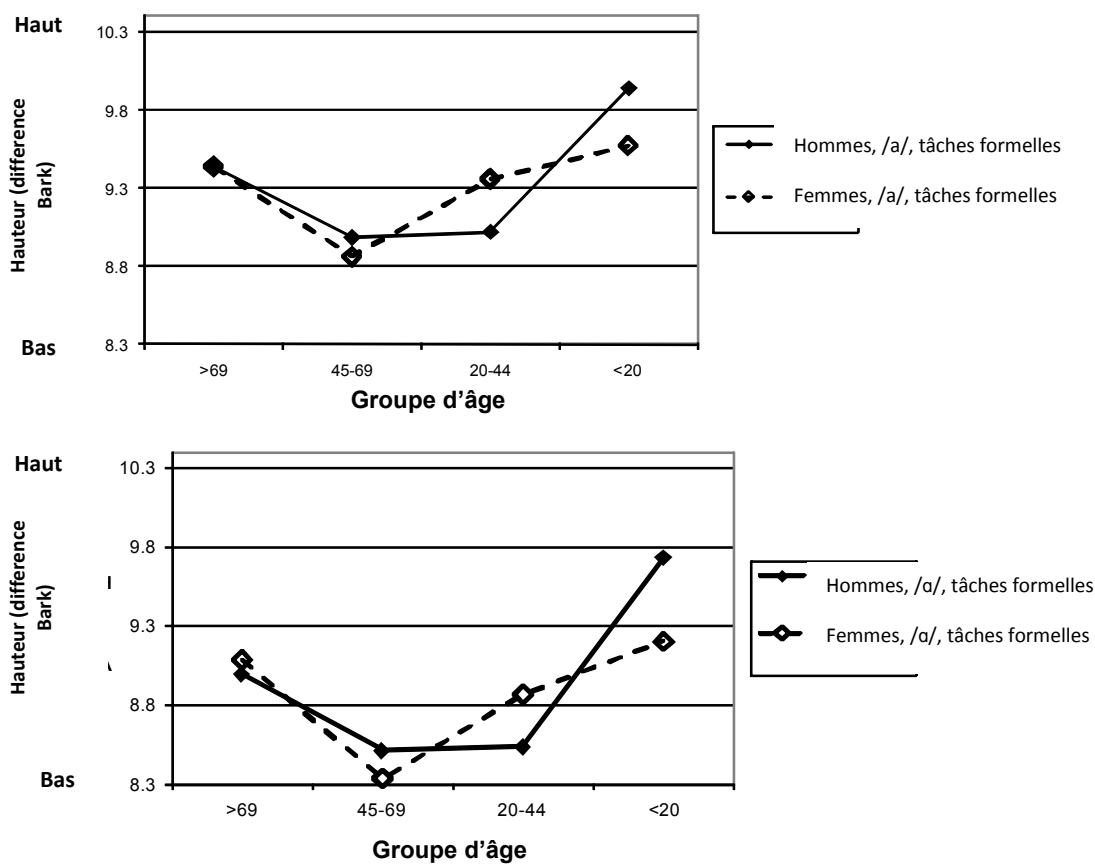
L'aire dialectale du « Northern Cities Shift » et les aires limitrophes
(d'après Labov *et al* 2006 : 140, Carte 11.12)

Légende

-  Aire du NCS
-  Aires limitrophes (de Sud en Nord): Centre (« Midland »), Pennsylvanie de l'Ouest, Côte Atlantique Centrale (« Mid-Atlantic »), Ville de New York (« NYC »), Nouvelle-Angleterre de l'Ouest, Nouvelle-Angleterre de l'Est, Canada
-  Points d'enquête à Rochester, État de New York (qui montre pleinement tous les éléments du NCS)
-  Points d'enquête à Albany, État de New York (qui montre les changements du NCS à un moindre degré)

Nous trouvons un autre exemple du même mouvement subtil mais mesurable à Philadelphie : la voyelle /ej/ (comme dans *snake* 'serpent') monte dans l'espace vocalique pour s'approcher à [i:] : *snake* /snejk/ se dit donc plutôt [sni:k], comme *sneak* 'faire furtivement'. Ce changement est très répandu à Philadelphie : il est évident à l'oreille si l'on sait qu'il est possible, mais il n'aurait pas été soupçonné d'abord sans l'usage des mesures acoustiques. Labov (2001 : 136 *et seq.*) donne plus de détails de ce changement.

Hall (2008) a également trouvé un développement de ce type en Normandie (c'est-à-dire un changement phonétique très subtile, pas évident à la première écoute, mais répandu de façon systématique dans la communauté – donc perceptible à un certain niveau). À Darnétal, limitrophe de Rouen, les voyelles /a/ (comme dans *pattes*) et /ɑ/ (*pâtes*) sont séparées dans l'espace vocalique (à la différence donc de beaucoup de communautés du nord de la France), mais *toutes les deux voyelles montent vers le centre de l'espace vocalique chez les locuteurs jeunes et âgés (pas chez ceux d'entre 20 et 70 ans)*. Cette différence n'était pas évidente à l'oreille lors du codage des données, mais se voit clairement sur un graphique des voyelles /a/ α pour tout l'échantillon darnétalais (Graphique 3).



Graphique 3: /a/ et /ɑ/ à Darnétal

haut: hauteur de /a/ (*pattes*), tâches formelles, par sexe et groupe d'âge

bas: hauteur de /ɑ/ (*pâtes*), tâches formelles, par sexe et groupe d'âge

Adaptation de la Graphique 3-10 (p129) de Hall (2008)

Dans ce graphique, plus la ligne est basse, plus la voyelle est basse : c'est-à-dire, plus la ligne est basse, plus la voyelle ressemble à la voyelle du français de référence ; moins la ligne est basse, plus la phonétique de la voyelle est innovatrice. Le but de notre projet est donc de chercher des effets de ce type en enquêtant dans toutes les villes du nord de la France au-delà d'une certaine taille.

3 La structure sociale du projet

L'objet de notre étude est donc clair : nous souhaitons examiner la variation géographique dans le français des villes du nord de la France. Pour ce faire, et pour nous assurer que toute variation que nous trouverons sera bien de la variation géographique, non imputable à la variation des classes sociales ou des âges (par exemple), il a fallu structurer notre échantillon d'une certaine façon.

3.1 Les villes de l'échantillon

Aire Urbaine	Population	Aire Urbaine	Population
Amiens	292.047	Lorient	212.317
Arras	128.558	Maubeuge	131.760
Beauvais	123.680	Metz	389.603
Béthune	365.513	Mulhouse	281.520
Boulogne-sur-Mer	134.244	Nancy	434.202
Brest	311.735	Paris	12.089.098
Caen	396.959	Quimper	123.963
Calais	126.618	Reims	313.818
Charleville-Mézières	106.747	Rennes	654.478
Cherbourg-Octeville	118.673	Rouen	649.291
Colmar	126.302	Saint-Brieuc	166.040
Creil	117.825	Saint-Quentin	111.595
Douai-Lens	544.143	Strasbourg	757.609
Dunkerque	259.566	Thionville	134.012
Évreux	110.528	Troyes	188.320
Le Havre	293.851	Valenciennes	366.781
Lille	1.150.530	Vannes	144.353

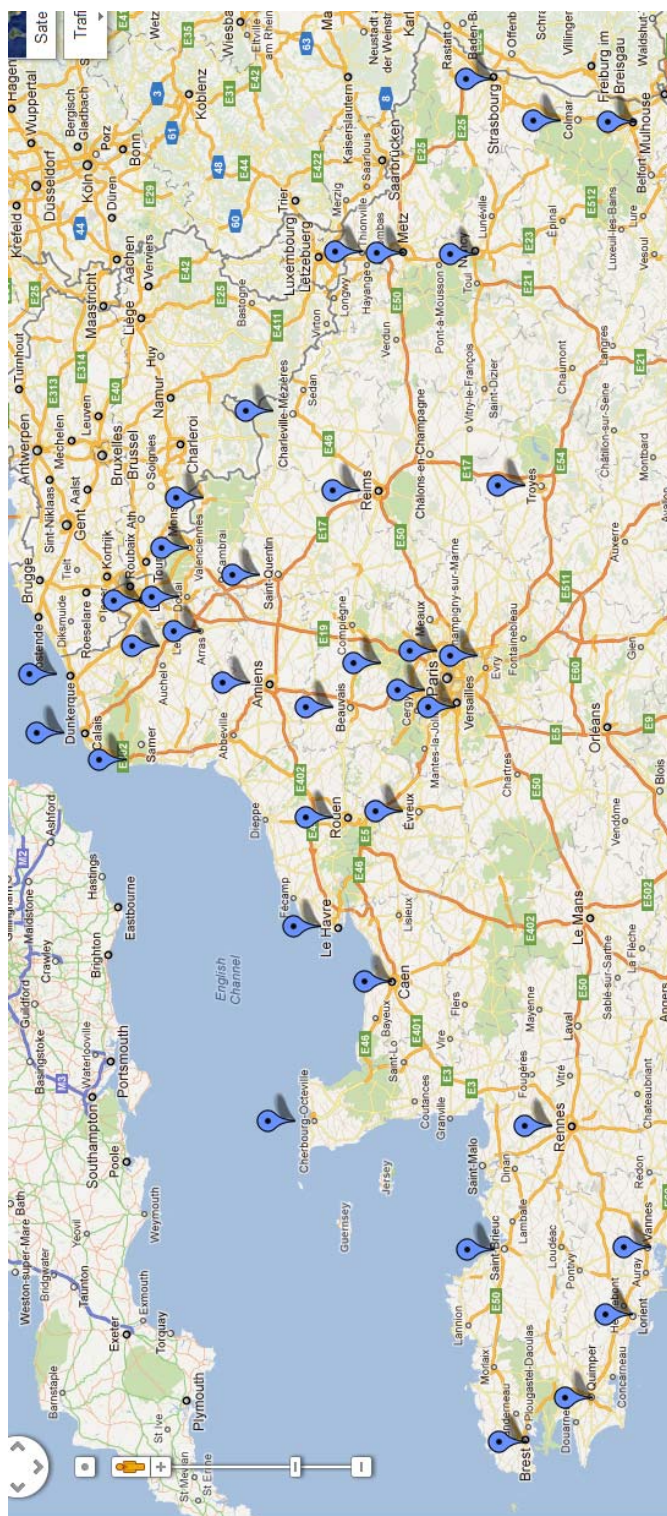
Tableau 1
Villes enquêtées dans TANLAF

Nous enquêtons dans toutes les villes du tiers septentrional de la France qui ont 100.000 habitants ou plus dans leur aire urbaine (selon les chiffres de l'INSEE de 2008, les plus récents qui soient disponibles pour les aires urbaines).⁷ Il y en a trente-quatre dans un croissant qui va de la Bretagne à l'Alsace en passant par Paris. Les tailles des aires urbaines vont d'un peu plus de 100.000 habitants (Charleville-Mézières, Lorraine et Saint-Quentin, Aisne) à plus de douze millions d'habitants (Paris) - voir Tableau 1 et Graphique 4. Par ce choix de zone d'étude, nous n'impliquons pas que les grandes villes au sud de notre zone (Lyon, Marseille) ne soient pas dignes d'étude : il s'agit simplement d'avoir dû nous arrêter quelque part, et ces villes seront couvertes dans des étapes ultérieures de notre étude. Dans chaque ville étudiée, nous analyserons l'accent d'un nombre égal d'hommes et de femmes, répartis comme suit :

- dans les villes (aires urbaines) de 100.000 à 199.999 habitants, un homme et une femme ;
- dans les villes de 200.000 habitants et plus, deux hommes et deux femmes ;
- sauf Paris, où nous enquêterons dans jusqu'à quatre sites dans l'aire urbaine, avec deux hommes et deux femmes par site, pour un total de seize personnes.

3.1.1 Pourquoi l'aire urbaine ?

Pour nous assurer que nos résultats seraient, sinon vérifiables, du moins répétables par d'autres, nous avons souhaité baser notre étude le plus possible sur des critères publics et non idiosyncratiques. Ainsi, dans le domaine du social, avons-nous souhaité prendre comme unité territoriale de base l'une des unités utilisées par l'Institut National de la Statistique et des Études Économiques (l'INSEE) pour ses données.



Graphique 4
Carte des villes enquêtées dans TANLAF
(© Google Maps 2012)

L'INSEE utilise plusieurs unités administratives territoriales de différentes tailles (INSEE 2012a), de la commune à la région ; nous nous attarderons brièvement ici sur quelques-unes d'entre elles, pour montrer pourquoi l'aire urbaine est le niveau d'analyse le plus approprié à un projet comme TANLAF.

Quand nous parlons de ce projet en général, afin que les nuances statistiques et analytiques n'obscurcissent pas le point du projet, nous parlons non des « aires urbaines » où nous allons enquêter mais des « villes ». Mais ce désir de simplicité apporte en soi un problème : en France en particulier, quand on parle des « villes », cela fait penser aux *communes*. La commune est « la plus petite subdivision administrative française mais [...] aussi la plus ancienne » (INSEE 2012b *sv* « commune » ; dans cette section, toutes les définitions d'unités administratives citées viennent de ce glossaire). De ce fait, et du fait que la commune est la seule unité administrative territoriale que rencontrent la majorité des Français au cours de la vie de tous les jours, vient le constat que c'est envers les communes que beaucoup de gens ont un sentiment de loyauté ; c'est donc à sa commune qu'on pense comme réponse à la question « De quelle ville venez-vous ? ». Du point de vue sociolinguistique, aussi, le choix de la commune comme unité territoriale de base aurait pu paraître logique d'une certaine façon, car les communes, étant relativement petites pour la plupart, ont souvent une certaine unité sociale, sans grands écarts de statut socioéconomique d'un bout de la commune à l'autre.

Le choix de la commune comme unité de base pour cette étude aurait quand même eu un grand inconvénient : tout le territoire français est découpé en communes, sans interstices. De nos jours le découpage (historique pour la plupart) n'a rien à voir avec les caractéristiques sociales du territoire découpé. Il est donc tout à fait possible, voire probable, qu'il n'y ait pas de grande différence entre le statut socioéconomique des personnes qui habitent juste d'un côté d'une frontière intercommunale et celui des habitants de la zone qui se trouve juste de l'autre côté. Si tel est le cas, cela équivaut à dire que le placement exact d'une frontière intercommunale à l'intérieur d'une zone urbaine ou d'une zone rurale n'a sans doute pas beaucoup d'importance sur les plans économique et démographique, même si les habitants de ces zones attribuent aux frontières exactes une importance locale et sociale qui ne prend pas en compte l'économie.

Nous ne disons pas du tout que la loyauté locale et sociale n'ait aucune importance en ce qui concerne la variation linguistique. Loin de là : cette importance est reconnue depuis au moins les travaux de Labov (1963) à Martha's Vineyard.⁸ Plus récemment, et focalisés plus étroitement sur l'identité, nous avons Kiely *et al.* (2000), Llamas (2007), et le projet « Accent and Identity on the Scottish-English Border » (« Accent et Identité sur la Frontière Écossaise-Anglaise » ; publications depuis 2009).⁹ Mais le projet TANLAF n'est pas un projet sociolinguistique à ce point-là : il ne cherche pas à expliciter les différences sociolinguistiques à l'échelle micro. Il vise plutôt les plus grossières divisions linguistiques entre les ensembles urbains, donc, les frontières intercommunales auraient été à une échelle trop petite pour cette étude.

Si la commune est donc trop petite pour servir d'unité géographique de base à cette étude, nous retenons quand même le principe que l'étude doit être basée sur la possibilité de différences entre zones d'habitation, comme les zones d'habitation sont les endroits où les gens ont le plus de contact entre eux, donc où les changements linguistiques que nous recherchons ont le plus de chances de se développer. Après la commune, donc, nous en venons à l'*unité urbaine*, « une zone de bâti continu [...] qui compte au moins 2.000 habitants ». 2.000 habitants est quand même toujours assez petit quand on a affaire à la dialectologie d'une zone qui compte plusieurs millions d'habitants ; nous avons donc pensé que l'unité urbaine était toujours trop petite pour cette étude.

Comme unités basées explicitement autour des villes, au début de cette étude l'INSEE utilisait aussi la *communauté urbaine* et l'*espace urbain*. À la différence de la commune et de l'unité urbaine, la communauté urbaine et l'espace urbain sont trop grands pour les besoins de cette étude : la communauté urbaine doit compter au moins 500.000 habitants (nous prenons en compte des aires urbaines qui en ont 100.000 et plus). L'espace urbain, que l'INSEE n'utilise plus depuis octobre 2011 (voir la définition du concept dans INSEE 2012b), était composé de plusieurs aires urbaines et des communes urbanisées qui les reliaient. Comme exemple de l'échelle de l'espace urbain, celui de Paris couvrait une grande partie du

centre de notre aire d'étude, et englobait au moins trois des aires urbaines que nous étudierons (à part celle de Paris même : INSEE 1997). L'espace urbain était donc trop grand et pas assez détaillé pour notre étude.

Nous avons remarqué à la section 2.3 qu'un contact particulièrement étroit entre les locuteurs est assuré quand ils travaillent ensemble ; nous avons donc voulu privilégier le travail comme motivation de l'unité territoriale que nous prendrions comme base de ce projet. Les gens voyagent et se réunissent pour travailler souvent en ville, et l'INSEE a une unité administrative ciblée exactement sur l'emploi : la **zone d'emploi**, « un espace géographique à l'intérieur duquel la plupart des actifs résident et travaillent ». LA zone d'emploi présente néanmoins un problème pour cette étude, celui de couvrir toute la France sans interstices (comme la commune). Si nous voulons voir s'il y a des différences entre les villes, de prime abord il nous a paru nécessaire de prendre comme unités d'analyse des unités où il y avait au moins la possibilité de décalages entre elles.

Face à ce choix d'unités, dont les unes semblaient trop petites et les autres trop grandes, l'**aire urbaine** nous a semblé être le choix du juste milieu :

une aire urbaine [...] est un ensemble de communes, d'un seul tenant et sans enclave, constitué par un pôle urbain (unité urbaine) de plus de 10.000 emplois, et par des communes rurales ou unités urbaines (couronne périurbaine) dont au moins 40 % de la population résidente ayant un emploi travaille dans le pôle ou dans des communes attirées par celui-ci.

La définition de l'aire urbaine se fait donc par référence aux emplois, comme nous le souhaitons, mais l'aire est automatiquement assez grande, contrairement à la zone d'emploi, qui n'a pas de limites (ni haute ni basse) d'étendue.

Une fois notre unité d'analyse sélectionnée, nous étions très conscients que sa population pouvait être socialement variable – et que, si nous cherchions à caractériser « le français » d'une région, nous devrions aussi spécifier les groupes sociaux que nous étudions. Les caractéristiques sociales des villes et des quartiers seront aussi prises en compte dans la sélection des interviewés, bien entendu ; voir la section 3.2.

3.2 Les interviewés

Pour nous assurer de trouver le bon type de variation si elle existe, il nous faut chercher des interviewés d'un certain type : hommes et femmes, jeunes, des classes moyennes, assez instruits (mais pas trop), ne travaillant pas dans l'éducation, et ne parlant pas une variété locale.

La nécessité de prendre des jeunes vient des conclusions de l'œuvre labovienne. Des centaines d'études dans cette tradition dans tout le monde concluent que s'il y a un changement linguistique, il se trouvera le plus vraisemblablement parmi des jeunes, et surtout parmi de jeunes femmes (Labov 1990 ; Labov *et al* 2006 : 28). En faisant nos entretiens avec des personnes d'entre 20 et 30 ans (idéalement), donc, nous décelerons des changements linguistiques naissants s'ils existent. Nous prenons des personnes des « classes moyennes » pour éviter l'« accent de banlieue » que les ouvriers de toutes les villes du nord de la France sont censés partager (potentiellement), et également pour éviter les accents trop « pointus » que des personnes des classes plus hautes risqueraient de prendre.

Pour cette raison, également, nous voulons éviter de faire des entretiens avec des personnes qui travaillent dans l'éducation ou qui ont poursuivi leurs études après la licence, parce que le fait d'être dans le monde éducatif pourrait amener une personne à modifier son accent quand elle se trouve dans une situation qui semble « public », tel un entretien enregistré avec un microphone. On pourrait penser que le même facteur affecterait aussi des personnes qui étaient toujours dans le système éducatif, même des étudiants de licence, que beaucoup de jeunes qui ont la vingtaine risquent d'être. Tel est certainement le point de vue des études qui ont été faites dans l'optique labovienne, surtout dans le monde anglo-saxon. Dans le monde anglo-saxon, les étudiants ont tendance à quitter leur ville de naissance pour faire leurs études et peut-être à ne pas avoir beaucoup de contacts avec celle-ci pendant leurs études. Nous avons pourtant

discuté de la question avec des éducateurs et des sociologues français lors de colloques récents, et on nous a proposé que le monde éducatif français n'est peut-être pas pareil, car (sauf exception) les étudiants français de licence ont plus tendance à rester près de chez eux, peut-être même dans le foyer familial, ou à rentrer souvent dans la ville de leur foyer familial s'ils l'ont quittée. Si c'est possible au sein de la présente étude, nous allons donc faire la comparaison entre les voyelles des étudiants de la vingtaine et les voyelles de jeunes de la vingtaine qui ont terminé leurs études. La nécessité de travailler avec des personnes qui ont un certain niveau d'éducation implique sans doute aussi la nécessité de travailler dans des quartiers d'un certain type, pas défavorisés, avec un bas niveau de chômage.

Enfin, nous voulons éviter de faire d'entretiens avec des personnes qui parlent une variété locale qui n'est pas le français (qu'il s'agisse d'une variété romane comme le picard ou le normand, ou d'une variété plus différente du français, comme le breton ou l'alsacien). Il est fort possible que nous décèlerons des effets linguistiques qui ont leur origine dans ces variétés locales, et si tel est le cas nous le noterons. Par contre, nous voulons nous assurer par-dessus tout que les variations que nous trouvons sont *des variations au sein du français local*, qui pourraient être partagés par n'importe quel locuteur de ce français, *non pas des variations individuelles* qui paraissent dans le français d'un interviewé donné parce que celui-ci parle une autre variété locale.

4 Quelles seront nos méthodes ?

Nous ferons des entretiens sociolinguistiques avec des locuteurs natifs des 34 villes qui nous intéressent, et les analyserons (pour la plupart) phonétiquement, à l'aide du logiciel gratuit Praat (Boersma et Weenink 1992-2012). De prime abord, quand le corpus est assez grand pour permettre une première analyse mais n'est pas encore complet, nous nous pencherons sur toutes les voyelles de l'espace vocalique francophone, ainsi que sur quelques consonnes peut-être intéressantes ; dans un deuxième temps, quand toute la centaine d'entretiens est faite, nous nous restreindrons aux voyelles qui, pendant la première étape, ont montré un décalage intéressant par rapport à la phonétique du français dit « de référence ».

Les entretiens de ce projet sont faits selon une version étendue du protocole d'enquête PFC (Durand, Lyche et Laks 2002), pour que les éléments pertinents de nos entretiens puissent s'intégrer au corpus PFC par la suite. Cette intégration permettra un double usage de nos données dans la science linguistique du français : elles figureront et dans nos analyses, et (nous le souhaitons) dans celles des autres qui travaillent sur le corpus PFC. Durand, Lyche et Laks (2002) – ainsi que tout le numéro du *Bulletin PFC* où cet article est publié – donne le détail du protocole. En sommaire, l'entretien contient les éléments suivants.

- liste de mots
- texte à lire
- conversation guidée (questions spécifiques, pour remplir la fiche signalétique)
- conversation libre (thèmes non prédéfinis)

Tout entretien doit comporter ces quatre éléments, mais leur ordre n'est pas fixé, puisque l'objet est de faire des entretiens qui seront aussi naturels que possible, pour être aussi près que possible du parler spontané des enquêtés.

La liste des phonèmes d'intérêt dans la première étape de ce projet figure dans le Tableau 2.¹⁰ La classification des voyelles orales en « familles » est celle des matériaux du programme PFC. Pourtant, comme TANLAF s'intéresse à plus de phonèmes et de positions phonologiques que PFC, la liste de mots de TANLAF est plus longue : la liste de mots PFC contient 94 mots, qui sont aussi les 94 premiers mots de la liste TANLAF, mais la liste TANLAF contient quelques 220 mots supplémentaires, pour un total de 314 mots. TANLAF ne modifie pas du tout le texte à lire de PFC, et les conversations guidée et libre sont du même style aussi. Les entretiens sont faits par l'enquêteur avec un ou deux enquêtés à la fois, enregistrés par un microphone attaché aux vêtements de l'enquêté, aussi près que possible de la bouche.

Variables PFC figurant aussi dans TANLAF

Vocoïdes (tous en syllabe accentué)

« famille a »	Maintien ou non du contraste entre /a/ (comme dans <i>pattes</i>) et /ɑ/ (comme dans <i>pâtes</i>)
« famille é / è »	Maintien ou non du contraste entre /e/ (<i>fée</i>) et /ɛ/ (<i>fais</i>)
« famille eu »	Maintien ou non du contraste entre /œ/ (<i>jeune</i>) et /ø/ (<i>jeûne</i>)
« famille o »	Maintien ou non du contraste entre /o/ (<i>côte</i>) et /ɔ/ (<i>cote</i>)
nasales	Contrastes entre /ɔ̃/ (<i>blond</i>), /ɑ̃/ (<i>blanc</i>), /ɛ̃/ (<i>brin</i>) et /œ̃/ (<i>brun</i>): quatre voyelles séparées, fusion de /ɛ̃/ et /œ̃/, et/ou fusion de /ɔ̃/ et /ɑ̃/
glissantes	Contrastes entre /ɥ/ (<i>huer</i>), /j/ (<i>yé-yé</i>) et /w/ (<i>ouais</i>)

Variables ne figurant pas dans PFC mais figurant dans TANLAF

Autres voyelles en syllabe accentué

i	Réalisation phonétique de /i/ : tendue ou relâchée ?
u	Réalisation phonétique de /u/ : tendue ou relâchée ?

Consonnes

r	Prononciation de /ʁ/ en fin de mot : r dit « grasséyé », approximante, pharyngéalisée ...
k	Aspiration de /k/ (surtout en début de mot)

Tableau 2
Variables de TANLAF et PFC

Avec un seul enquêté, les entretiens durent plus ou moins une heure ; avec deux, plus ou moins une heure et demie.

Il ne sera sans doute pas possible de faire une analyse acoustique de tous les phonèmes d'intérêt, faute de techniques acoustiques assez fiables, mais au moins les voyelles orales et l'aspiration de /k/ devraient être mesurables acoustiquement. Les méthodes d'analyse feront l'objet de publications ultérieures.

Les mesures ainsi obtenues feront par la suite l'objet de classifications en aires dialectales, comme elles sont bien connues dans les atlas linguistiques. Les conclusions de ces classifications seront exposées d'abord dans des articles, ensuite dans un atlas si le temps et le financement le permettent.

5 Qu'est-ce que nous nous attendons à trouver ?

Par l'analyse acoustique, donc, nous nous attendons à trouver des aires dialectales peut-être plus détaillées que jamais en France du nord. Tout comme les trente ans de recherches menant à l'*Atlas of North American English* (Labov *et al* 2006) ont décelé des aires dialectales phonétiques très détaillées, nous espérons que nos recherches mèneront à un portrait détaillé de la variation vocalique du nord de la

France. Nous ne voulons pas dire que nos conclusions seront le dernier mot sur la variation vocalique de notre aire – loin de là : il y aura beaucoup de régions urbaines plus petites, entre celles où nous enquêterons, où les possibles changements linguistiques resteront à expliciter. Nous souhaitons donc que la suite de nos recherches prenne deux formes (qu'elles soient réalisées par nous ou par d'autres) :

- des recherches plus détaillées dans la région où nous enquêtons, et dans des villes pas encore étudiées, et en plus de détail dans des villes où nous enquêterons ;
- des recherches du même type dans le reste de la France (c'est-à-dire le centre et le sud), et dans d'autres pays francophones, de façon à constituer (à la longue) un ou des atlas phonétique(s) et phonologique(s) du français moderne de la France et de la francophonie.

Nos conclusions sortiront donc d'abord sous la forme d'articles, et dans des revues linguistiques traditionnelles, et sur l'Internet ; nous espérons qu'à la longue il y aura un atlas physique et électronique. Celui-ci est également le parcours qu'ont pris les recherches de l'*Atlas of North American English*.

6 Sommaire et conclusion

Nous décrivons cette étude phonétique – phonologique des accents urbains dans le français du nord de la France comme de la dialectologie sociolinguistique. Il s'agit bien de la dialectologie, car l'étude se concentrera sur la distribution géographique de la variation phonétique et phonologique en France du nord ; mais cette dialectologie sera informée par la sociolinguistique, de façon à chercher la variation linguistique là où elle se trouvera le plus vraisemblablement. Comme nous cherchons de la variation

- en français
- parmi les jeunes
- en région urbaine,

la présente étude représente un départ des objectifs de la dialectologie française traditionnelle. Quand même, nous espérons que cela va sans dire que nous ne souhaitons d'aucune façon supplanter la dialectologie française traditionnelle – loin de là ; nous souhaitons ajouter un aspect à nos connaissances de la variation linguistique en France.

Ce projet est un projet en cours, que nous espérons utile à la communauté linguistique française et francophone. Nous accueillons donc vivement les suggestions, les commentaires et les critiques des collègues ! Nous sommes joignables le plus facilement par mail à d.hall@kent.ac.uk.

Références bibliographiques

- Armstrong, N. et Pooley, T. (2010). *Social and Linguistic Change in European French*. Basingstoke, Royaume-Uni : Palgrave Macmillan.
- Arnaud, V. (2006). *La Dimension Variationniste du Français en Usage à Saint-Claude (Haut-Jura) : une étude acoustique des voyelles orales des « gens d'en haut »*. Thèse doctorale non publiée, Université Laval.
- Boersma, P. et Weenink, D. (1992-2012). *Praat : doing phonetics by computer*. Logiciel gratuit, téléchargeable de <http://www.praat.org>.
- Brun-Trigaud, G. ; Le Berre, Y. ; & Le Dù, J. (2005). *Lectures de l'Atlas linguistique de la France de Gillieron et Edmont : Du temps dans l'espace*. Paris : Éditions du Comité des travaux historiques et scientifiques.
- Calvet, L.-J. (1993). *La Sociolinguistique*. Série « Que Sais-Je ? ». Paris : Presses Universitaires Françaises.
- Calvet, L.-J. (1994). *Les Voix de la Ville : introduction à la sociolinguistique urbaine*. Paris : Payot et Rivages. Consulté dans l'édition de poche (2011, Payot & Rivages).
- Delattre, P. (1948). Un triangle acoustique des voyelles orales du français. *French Review*, 21(6), 477-84.

- Dinkin, A. (2009). *Dialect Boundaries and Phonological Change in Upstate New York*. Thèse doctorale non publiée, Université de Pennsylvanie.
- Durand, J. ; Laks, B. ; & Lyche, C. (2002). Synopsis du projet PFC : la Phonologie du Français Contemporain : Usages, Variétés et Structure. *Bulletin PFC* 1 : 5-6. Téléchargeable de http://www.projet-pfc.net/bulletins-et-colloques/cat_view/918-bulletins-pfc.html.
- Durand, J. ; Laks, B. ; & Lyche, C. (2009). *Phonologie, Variation et Accents du Français*. Paris : Hermès.
- Durand, J. ; Lyche, C. ; & Laks, B. (2002). Protocole d'enquête (mai 2002). *Bulletin PFC* 1 : 7-20. Téléchargeable de http://www.projet-pfc.net/bulletins-et-colloques/cat_view/918-bulletins-pfc.html.
- Gilliéron, J. et Edmont, E. (1902-10). *Atlas Linguistique de la France*. Paris : Champion.
- Hall, D. (2007). Le français de Rouen (Haute-Normandie, France): une variété de français qui représente bien sa ville. *Bulletin PFC* 7: 171-191. Téléchargeable de http://www.projet-pfc.net/bulletins-et-colloques/cat_view/918-bulletins-pfc.html.
- Hall, D. (2008). *A Sociolinguistic Study of the Regional French of Normandy*. Thèse doctorale non publiée, Université de Pennsylvanie.
- Institut National de la Statistique et des Études Économiques [INSEE]. (1997). Pôles Urbains et Périurbanisation. *INSEE Première* 516. Téléchargeable de http://www.insee.fr/fr/ffc/docs_ffc/ip516.pdf.
- Institut National de la Statistique et des Études Économiques [INSEE]. (2011a). Le Nouveau Zonage en Aires Urbaines de 2010. *INSEE Première* 1374. Téléchargeable de http://www.insee.fr/fr/themes/document.asp?ref_id=ip1374.
- Institut National de la Statistique et des Études Économiques [INSEE]. (2011b). *Données Complémentaires*. Tableur gratuit, téléchargeable de http://www.insee.fr/fr/themes/document.asp?ref_id=ip1374.
- Institut National de la Statistique et des Études Économiques [INSEE]. (2011c). *Résultats du Recensement de la Population – 2008*. Base de données en ligne, accessible par une interface: <http://www.recensement.insee.fr/home.action>.
- Institut National de la Statistique et des Études Économiques [INSEE]. (2012a). *Code Officiel Géographique, zonages d'études*. Page de synthèse des zonages utilisés par l'INSEE aux différents niveaux géographiques. <http://www.insee.fr/fr/methodes/default.asp?page=zonages/liste-zonages.htm>.
- Institut National de la Statistique et des Études Économiques [INSEE]. (2012b). *Dictionnaire de Définitions de l'INSEE*. Tableur gratuit, téléchargeable de <http://www.insee.fr/fr/methodes/default.asp?page=definitions/liste-definitions.htm>. (La version la plus récente est toujours consultable aussi en ligne à la même adresse ; version du 14 février 2012 consulté le 20 avril 2012.)
- Kiely, R. ; McCrone, D. ; Stewart, R. ; et Bechhofer, F. (2000). Debatable land: national and local identity in a border town. *Sociological Research Online* 5(2). <http://www.socresonline.org.uk/5/2/kiely.html>.
- Labov, W. (1963). The Social Motivation of a Sound Change. *Word* 19: 273-309.
- Labov, W. (1972). *Sociolinguistic Patterns*. Philadelphie, États-Unis: University of Pennsylvania Press. Publication en français: Labov, W. (1976). *Sociolinguistique*. Paris: Les Éditions de Minuit.
- Labov, W. (1990). The Intersection of Sex and Social Class in the Course of Linguistic Change. *Language Variation and Change*, 2, 205-54.
- Labov, W. (1994). *Principles of Linguistic Change, Volume 1: Internal Factors*. Malden, MA, États-Unis et Oxford, Royaume-Uni : Blackwell.
- Labov, W. (2001). *Principles of Linguistic Change, Volume 2: Social Factors*. Malden, MA, États-Unis et Oxford, Royaume-Uni : Blackwell.
- Labov, W. ; Ash, S. ; et Boberg, C. (2006). *The Atlas of North American English: Phonetics, Phonology and Sound Change: a Multimedia Reference Tool*. Berlin, Allemagne et New York, NY, États-Unis : Mouton de Gruyter.

- Labov, W. ; Yaeger, M. ; et Steiner, R. (1972). *A Quantitative Study of Sound Change in Progress*. Philadelphia, PA, États-Unis : The US Regional Survey.
- Lennig, M. (1978). *Acoustic Measurement of Linguistic Change: the modern Paris vowel system*. Thèse doctorale, Université de Pennsylvanie, États-Unis ; Pennsylvania Dissertation Series, no.1.
- Llamas, C. (2007). « A place between places »: language and identities in a border town. *Language in Society*, 36(4), 579-604.
- Lyche, C. (2010). Le français de référence : éléments de synthèse. In Detey, S. ; Durand, J. ; Laks, B. ; et Lyche, C. *Les Variétés du Français Parlé dans l'Espace Francophone : Ressources pour l'enseignement*. Paris : Éditions Ophrys.
- Martinet, A. (1945 / 1971). *La Prononciation du Français Contemporain: témoignages recueillis en 1941 dans un camp d'officiers prisonniers*. Première édition Paris : Librairie E. Droz ; deuxième édition Genève : Librairie E. Droz.
- Martinet, A. (1955). *L'Économie des Changements Phonétiques*. Berne, Suisse : Francke.
- Martinet, A. (1960 / 2008). *Éléments de Linguistique Générale*. Paris : Armand Colin. Première édition 1960 ; cinquième édition 2008.
- Martinet, A. et Walter, H. (1973). *Dictionnaire de la Prononciation Française dans son Usage Réel*. Paris : France Expansion.
- Programme « Phonologie du Français Contemporain » [PFC]. (2004-2008). *Recherche : base de données en ligne*. Accessible à <http://www.projet-pfc.net/pfc-recherche.html>.
- Walter, H. (1976). *La Dynamique des Phonèmes dans la Lexique du Français Contemporain*. Paris : France Expansion.
- Walter, H. (1982). *Enquête Phonologique et Variétés Régionales du Français*. Paris : Presses Universitaires de France.
- Watson, K. et Mattock, K. (2009). *The effect of speakers' geographical mobility on their production and perception of a phonological merger: evidence from NURSE and SQUARE in Liverpool English*. Louvain, Belgique: communication au colloque Perception, Production & Attitude, 2-3 avr. 2009.
- Wells, J.C. (1982). *Accents of English*. Cambridge, Royaume-Uni ; New York, NY, États-Unis ; et Melbourne, Australie : Cambridge University Press.

¹ Nous remercions The Leverhulme Trust (<http://www.leverhulme.ac.uk/funding/ECF/ECF.cfm>) pour leur soutien financier de ce projet, et nos deux rapporteurs anonymes pour leurs commentaires sur cet article (ainsi que d'autres collègues pour leurs commentaires sur le projet). Nous espérons en avoir tenu compte dans la rédaction finale.

« Vers un Nouvel Atlas Linguistique de la France » n'est qu'une simple traduction du titre primaire du projet, qui est en anglais, puisque le projet est financé par une agence britannique (The Leverhulme Trust) et basé dans une université britannique (l'Université de Kent). La traduction n'est pas censée se référer explicitement au projet des *Nouveaux Atlas Linguistiques de la France par régions*, les buts linguistiques des deux projets n'étant pas les mêmes, malgré certaines similarités d'approche.

² Martinet (2008) utilise le symbole « /é/ » pour noter le phonème dont la représentation en API est /e/. Pour la citation, nous gardons la notation de Martinet, tout en utilisant l'API conventionnel ailleurs dans l'article.

³ Nos études précédentes ont montré que, chez certains vieux locuteurs (surtout ruraux), il peut être difficile de trancher sur la langue qu'ils parlent à un moment donné (français ou variété locale autochtone). Chez les jeunes citadins qui font l'objet de ce projet, par contre, cette question ne se posera sans doute pas.

⁴ Le texte préféré pour faire la présentation du programme PFC de nos jours est Durand, Laks et Lyche (2009), mais cette citation, qui date du début du programme en 2002, en est la meilleure présentation courte. Nota : la variation entre « programme PFC » et « projet PFC » n'est pas un lapsus. Au début des activités PFC, le nom officiel de l'ensemble de celles-ci était bien « projet PFC » ; au fur et à mesure que ces activités ont augmenté, avec plusieurs volets d'études non conçus au début, les dirigeants ont décidé que l'ampleur des activités méritait qu'on les appelle le « programme PFC ».

⁵ Graphique basé sur Labov *et al.* (2006 : 121), avec l'autorisation des auteurs. À noter :

- Les éléments du changement en chaîne sont numérotés : le mouvement de /æ/ (comme dans TRAP) vers le haut est le mouvement le plus répandu au nord des États-Unis et semble avoir débuté le premier, suivi de l'antériorisation du soi-disant 'o court' (phonémiquement /a/, comme dans LOT), et ainsi de suite.
- Les symboles utilisés dans l'œuvre de Labov pour les voyelles de l'anglais n'ont pas toujours la même signification que dans l'Alphabet Phonétique International (l'intention des symboles laboviens étant à l'origine d'être plus facilement interprétables pour les ordinateurs et plus mnémoniques par rapport à l'anglais américain). Les symboles de la Graphique 1 sont à interpréter comme suit.

Symbole labovien	Valeur API	Classe vocalique selon le schéma de Wells (1982)
æ	/æ/	TRAP
o	/a/	LOT ('cot' dans beaucoup d'écrits de la tradition labovienne)
oh	/ɔ:/	THOUGHT ('caught')
ʌ	/ʌ/	STRUT
e	/e/	DRESS

⁶ Carte modifiée avec l'autorisation des auteurs.

⁷ Les chiffres de population viennent des sources suivantes :

- aires urbaines parmi les 50 plus grandes de France : INSEE 2011b (Amiens, Béthune, Brest, Caen, Douai-Lens, Dunkerque, Le Havre, Lille, Lorient, Metz, Mulhouse, Nancy, Paris, Reims, Rennes, Rouen, Strasbourg, Troyes et Valenciennes)
- autres aires urbaines : INSEE 2011c (les aires urbaines ayant une population de plus de 100.000 mais ne figurant pas parmi les 50 plus grandes aires urbaines de France : Arras, Beauvais, Boulogne-sur-Mer, Calais, Charleville-Mézières, Cherbourg-Octeville, Colmar, Creil, Évreux, Maubeuge, Quimper, Saint-Brieuc, Saint-Quentin, Thionville et Vannes)

Les territoires des aires urbaines ont été redéfinis en 2010 (INSEE 2011a) ; le chiffre de population pour chaque aire est issu du Recensement de la Population 2008, qui, malgré son nom, est le « cumul des cinq enquêtes annuelles de recensement réalisées de 2006 à 2010 [et décrit] une situation moyenne au début 2008 » (INSEE 2011a : 4).

⁸ Publié aussi en anglais dans Labov (1972 : 1-42), et en français dans Labov (1976 : 45-93). Voir aussi les courtes présentations de l'essentiel de cette étude par Calvet (1993 : 62-64 et 1994 : 98-101). Nota : il y a une erreur d'impression dans la transcription phonétique de la centralisation dans les travaux cités de Calvet. Une phrase à 1993 : 62 et 1994 : 99 dit que « le premier élément [des diphtongues /ay/ et /aw/], le /a/, a une tendance à être « centralisé » chez les Vineyardais, c'est-à-dire à prendre une prononciation plus proche du /e/ » ; à la dernière partie de la phrase, il faudrait plutôt lire « une prononciation plus proche du /ə/ [schwa] ».

⁹ Une liste des publications du projet « Accent and Identity on the Scottish-English Border » se trouve à <http://www.york.ac.uk/res/aiseb/>.

¹⁰ Le lecteur averti verra que le projet TANLAF n'inclut pas deux variables qui sont justement très variables en France, et qui sont couverts par le programme PFC : variabilité de prononciation de schwa, et usage variable de la liaison. Nous avons pris cette décision parce que nous souhaitions cibler l'analyse phonétique des voyelles et des consonnes dans ce projet, pour voir si notre terrain d'enquête y montrait les types de variabilité phonétique que l'on a trouvé dans les pays anglophones (voir la section 2.6). Le protocole d'enquête de TANLAF inclut l'intégralité de celle de PFC, pourtant ; à la longue, les enregistrements TANLAF pourront donc être inclus dans le corpus PFC, et l'analyse du schwa et de la liaison dans ces enregistrements pourra donc se faire ultérieurement.