

## Le schwa en français : pourquoi des corpus ?

Jacques DURAND, Julien EYCHENNE  
ERSS, UMR5610, C.N.R.S et Université de Toulouse II<sup>1</sup>

**Résumé :** L'objectif principal que nous poursuivons dans cet article est de montrer que pour un phénomène aussi complexe que le schwa, qui ne se laisse pas appréhender de manière triviale, la constitution de corpus s'appuyant sur une méthodologie rigoureuse et explicite est un élément indispensable de la construction des objets phonologiques. A cette fin, nous nous concentrons sur l'étude de la position finale, et plus spécifiquement de l'opposition /Cə#/ vs /C#/, à la lumière de trois enquêtes du projet « Phonologie du français contemporain (PFC) : usages, variétés et structure » : un corpus de Douzens (Aude), un corpus de Biarritz (Pays Basque) et un corpus de Treize-Vents (Vendée).

Après avoir resitué la problématique du schwa dans le débat général sur le traitement des positions finales, nous présentons le projet PFC, sa méthodologie et ses outils. Nous exposons enfin les résultats obtenus pour les trois corpus : le corpus de l'Aude représente un français du Midi très conservateur, où l'opposition /Cə#/ vs /C#/ est tenace malgré des variations entre générations ; le corpus du Pays Basque représente un français du Midi en évolution, où cette opposition tend à disparaître mais est résiduellement présente ; le corpus de Vendée (variété rarement décrite) se révèle plus proche du français de référence, et les données indiquent sans ambiguïté que l'opposition /Cə#/ vs /C#/ a été éliminée. Notre conclusion, si la démarche est correcte, est que ces résultats et les analyses qu'ils permettent (y compris l'étude des corrélations entre oral et écrit) seraient difficiles, voire impossible, à obtenir sur la base d'intuitions ou d'observations non contrôlées.

---

1. Nous remercions les membres du projet PFC, et tout particulièrement Géraldine-M. Mallet qui nous a fourni les données de l'enquête Vendée. Nous remercions également deux lecteurs anonymes qui nous ont fait part de leurs commentaires.

## 1. Introduction

Avec la liaison, le « h aspiré » et la nasalisation, la problématique du schwa constitue l'un des grands enjeux de la phonologie du français. Le terme de « schwa » n'est peut-être pas le plus approprié même s'il est déjà attesté (sous la forme *scheva*) dans la *Grammaire générale et raisonnée* d'Arnauld et Lancelot (1660 : 10-11). D'aucuns lui préfèrent les termes « e caduc » ou « e muet » mais, quelle que soit la terminologie adoptée, il faut rendre compte en français d'un ensemble de contextes où une voyelle qui correspond très souvent à un *e* graphique alterne avec zéro. Ces alternances peuvent se produire dans diverses positions dans la chaîne linéaire, en fonction du contexte. Même si un examen attentif de la structure interne des mots peut nous aider à mieux comprendre le phénomène du schwa, seules des études au niveau des énoncés peuvent permettre une analyse complète et cohérente de ces alternances (cf. Lyche et Durand, 1996), d'où l'importance de corpus qui constituent un instrument important pour l'examen du comportement des mots dans le discours.

Dans cet article, nous visons à démontrer, à travers l'étude de l'opposition /Cə#/ vs /C#/ dans trois corpus contrastés (Aude, Pays Basque et Vendée), qu'une étude empirique fondée sur des corpus est un élément indispensable de la construction des objets phonologiques. Il ne s'agit pas de sombrer dans un empirisme naïf qui réduirait les observables aux seuls observés d'enquêtes données : puisque le schwa présente l'alternance entre zéro et quelque chose, les objets à analyser ne se livrent pas si facilement à l'observation. Au-delà des contraintes linéaires, il est rare de trouver, même dans un corpus de taille conséquente, les mêmes items dans des environnements identiques ou strictement comparables. Néanmoins, si l'on accepte de comparer non seulement des items, mais aussi des classes d'items, la méthode extensionnelle que permet l'exploitation d'un corpus est indispensable.

Mais un corpus se construit. Ce n'est pas un objet inerte, omnibus, qui peut répondre à toutes les questions que se pose le linguiste pourvu qu'il se penche attentivement sur les données. Pour le schwa, nous présenterons plus loin (voir 3) la stratégie

*Le schwa en français : pourquoi des corpus ?*

suivie dans le projet PFC (*Phonologie du Français Contemporain : usages, variétés et structure*). Nous verrons que les corpus rassemblés dans ce projet sont des objets théorisés qui, doublés de méthodes de codage explicites, permettent des analyses quantitatives et qualitatives bien supérieures à la seule « intuition » (au sens de jugement de grammaticalité). Il ne s'agit pas de rejeter le travail théorique du linguiste mais de lui assigner sa place correcte dans les allers-retours entre théorie, observation et expérimentation qui fondent la démarche scientifique.

Pour les besoins de notre démonstration, nous ne nous intéresserons ici qu'aux positions finales de mot. Toutes les variétés du français ne présentent pas les mêmes caractéristiques quant aux finales autorisées et interdites. Le français méridional traditionnel, en maintenant une opposition entre schwa et zéro (*mère* /mɛrə/ vs *mer* /mɛr/), est censé représenter un stade plus archaïque de la langue que les variétés du nord de la Loire et d'autres régions francophones (comme la Belgique, le Québec ou la Suisse par exemple). Dans certains systèmes, on a pu défendre l'idée que les schwas historiques étaient recodés par la quantité vocalique (Martinet, 1969 : 177-178). Seules des études présentant une méthodologie uniforme permettent de mieux cerner la nature et l'ampleur des phénomènes en jeu.

Nos objectifs ici seront limités. Nous illustrerons la méthodologie du projet PFC en étudiant le schwa à partir de trois enquêtes : deux dans le sud de la France (Languedoc et Pays Basque) et une en Vendée. La Vendée a été choisie car elle représente une zone de transition entre le sud et le nord dont on sait finalement peu de choses. L'enquête en Languedoc (Douzens, Aude), quant à elle, offre un intérêt double : elle présente d'une part un accent du Midi conservateur, mais elle illustre une évolution lente de ce système vers un système plus standard, en particulier chez les jeunes locuteurs. Cet accent conservateur peut servir d'étalon par rapport à d'autres accents du Midi dont on se demande s'ils présentent toujours l'opposition entre présence et absence de schwa en position finale. Le Pays Basque est, de ce point de vue-là, extrêmement intéressant. Les grands traits de la phonologie de nos témoins

sont typiques d'un accent du Midi (Durand et Tarrier, 2003). En revanche, les oppositions entre schwa et zéro à la finale nous ont paru à l'écoute avoir basculé vers le système « nordique ». Mais entre l'écoute, même attentive, et les comparaisons et codages d'une enquête au format de notre projet, il y a un abîme que nous nous proposons de combler ici.

## 2. La position finale en français : diversité et unité

Etablir les séquences phonologiques finales (ou initiales ou internes) autorisées dans une langue donnée n'est pas chose aisée. On engage forcément la théorie des constituants syllabiques et sa relation avec les contraintes morphologiques. On sait, par exemple, que certaines théories traditionnelles qu'on pourrait qualifier de 'naïves' ou de 'surfacistes' calquent la structure de la syllabe sur les séquences observées dans les mots de la langue. Dans de tels cadres, l'existence d'un mot comme *strict* démontre qu'on peut avoir trois obstruantes en attaque de syllabe et deux en coda (Fudge, 1969 ; Clements & Keyser, 1983 ; Blevins, 1995). A l'opposé, un modèle comme la phonologie CVCV ou CV stricte (Lowenstamm, 1996 ; Scheer, 1996, 2000) pose que le patron syllabique est le même pour toutes les langues (à savoir CV) et que les séquences apparentes de consonnes finales observables dans diverses langues s'expliquent en termes de conditions sur la réalisation de noyaux vides. Nous ne donnerons pas à ces débats l'écho qu'ils méritent car nous nous situerons à un niveau plus fondamental, celui des observables, même si l'observation s'intègre toujours à un cadre de référence.

Les spécialistes du français s'accordent pour reconnaître qu'il a existé dans la variété standard une véritable opposition entre des finales en /C#/ et en /Cə#/. Cet état de langue est celui que reflète la graphie du français contemporain qui différencie *mer – mère, sel – selle, roc – roque*, etc. Fouché (1958 : 509 et suivantes) date le début de l'amuissement des 'e' finaux vers le milieu du XV<sup>ème</sup> siècle et, selon lui, le processus aurait été mené à bout au XVII<sup>ème</sup> siècle. L'absence d'une opposition /C#/ et /Cə#/ dans le français du nord de la Loire semble acquise pour les spécialistes jusqu'à l'avènement de la

*Le schwa en français : pourquoi des corpus ?*

grammaire générative, que les travaux soient pédagogiques, descriptifs ou théoriques. Martinet (1969 : 212) déclare par exemple :

« A la finale des mots, dans un usage majoritaire et qui s'impose, *-e* n'a pas laissé de trace : *patte* se prononce comme *pat*, *messe* comme *mess*, *coque* comme *coq*, *boxe* comme *box* ; à l'issue des mots *porte*, *garde*, *rustre*, *sable*, on peut entendre une détente qu'on pourrait vouloir noter comme une voyelle centrale, mais cette détente n'est ni plus ni moins caractérisée dans *acte* que dans *exact* (/ekgzakt/ en concurrence avec /egza/), dans *peste* que dans *Budapest*, et n'est donc pas à porter au crédit de 'e muet' ».

Dès ses premiers travaux, la phonologie générative classique a relancé le débat en posant pour le français de « référence » (cf. Morin, 2000) une différence au niveau sous-jacent entre mots se terminant par consonne(s) + schwa et mots à finale consonantique (cf. Schane, 1968 ; Dell, 1985). En effet, la liaison s'explique dans ce cadre par la présence d'une consonne latente qui est effacée devant consonne ou à la pause. Par exemple, *petit animal* se voit attribuer la forme sous-jacente /pətit#animal/ et la prononciation [pətitanimal]<sup>2</sup> découle naturellement de la forme de départ. En revanche, dans *petit cadeau* /pətit#kado/, le /t/ latent devra être effacé. Les mots morphologiquement féminins et diverses formes verbales ont tous un schwa « abstrait » qui protège la consonne finale en la déroband à la règle d'effacement (*petite camarade* = /pətitə#kamaradə/). En français standard, ce schwa est effacé mais seulement après l'application de la règle de troncation pour produire, en conjonction avec d'autres règles tardives, les formes de surface requises (par exemple, *petite camarade* = [pətitkamarad], [pœtitkamarad] ou encore [ptitkamarad]). Pour

---

2. Plus précisément [pœtitanimal] pour de nombreux locuteurs. Le symbole /ə/ a malheureusement une double valeur : vraie voyelle centrale ou indice de l'alternance entre voyelle et zéro. La voyelle en question est rarement un vrai schwa phonétique (voir Dell, 1985 ; Durand & Lyche, 1996 ; Durand *et alii*, 1987 pour le français du midi.)

justifier une telle analyse, des chercheurs comme Schane, Dell ou Selkirk (1972) pour ne citer qu'eux, invoquent à côté des généralisations établies le fait que les schwas abstraits postulés se retrouvent dans certains styles oratoires (la poésie, par exemple) ou dans la chanson. L'orthographe française, vue comme diabolique par les tenants de la transcription phonémique, se voit également justifiée comme reflétant des régularités plus profondes d'ordre morphophonologique ou phonosyntaxique. Enfin, la variation dialectale trouve aussi sa place dans la réflexion puisque le français du Midi est pris comme preuve vivante de la réalité de ces schwas sous-jacents (Schane, 1968). La différence entre un accent du Midi et un accent du nord de la Loire ne tiendrait alors qu'à des ajustements proches de la surface. L'accent standard nécessite tout simplement un plus grand nombre de règles d'effacement que l'accent du Midi, mais tous les deux partagent les mêmes formes profondes.

L'hypothèse d'un système commun aux divers dialectes d'une langue n'est pas inintéressante. Encrevé (1988), dans son travail classique sur la liaison non enchaînée, l'invoque à plusieurs reprises. Dans une communauté linguistique aussi intégrée que celle de la France hexagonale, les pratiques linguistiques entraînent probablement des structurations communes entre diverses zones dialectales. Nous ne nous contentons pas de produire du langage mais nous le percevons aussi (non seulement avec l'ouïe mais aussi avec la vue pour les sujets graphématisés, qui constituent la majorité des utilisateurs du français contemporain en France). Ainsi, Encrevé postule-t-il que la liaison non enchaînée et les mécanismes représentationnels qui lui sont associés font partie de la grammaire phonologique de tous les francophones de France même si le non enchaînement n'est vraiment présent que dans les parlers des professionnels de la parole (politiques, reporters, enseignants, etc.). Si cette solution est envisageable pour la liaison, elle l'est évidemment pour le schwa.

La nécessité de représentations communes pour les membres d'une communauté linguistique donnée reste cependant à démontrer. On sait que les locuteurs s'accommodent de différences fortes dans certains domaines

phoniques. En Belgique, on ne trouve habituellement que deux glissantes /w j/ face à trois /w j ɥ/ dans le français hexagonal. En français de référence, on pratique des oppositions entre /e/ et /ɛ/ qui ne se retrouvent pas dans d'autres variétés comme le français méridional. Est-il nécessaire, pour expliquer la communication entre membres de diverses aires dialectales ou sociales, de postuler un seul système sous-jacent commun aux locuteurs d'une même langue (d'autant plus que cette dernière est une construction plus idéologique que linguistique, voir Durand, 1996). Au moment de la parution de SPE (*The Sound Pattern of English*, Chomsky et Halle, 1968) et de son pendant pour le français *French Phonology and Morphology* (Schane, 1968), l'hypothèse de formes sous-jacentes pandialectales et panchroniques semblait également utile pour rendre compte d'alternances diverses (avec quelques dérives connues donnant la même forme sous-jacente à *bénir* et *bien* ou *serein* et *soir* ; cf. Schane, 1968 : ch. 2). Peu de linguistes aujourd'hui verraient cette tâche comme centrale. S'ils ont raison, l'hypothèse de formes sous-jacentes pandialectales (même fortement sous-spécifiées) n'est pas nécessaire pour rendre compte des échanges verbaux entre membres d'une même communauté. Le traitement du schwa ne saurait donc se régler par un simple alignement, au niveau sous-jacent, sur la variété apparemment la plus conservatrice, à savoir le français du Midi.

On sait par ailleurs que le traitement des positions finales en phonologie générative soulève des difficultés bien perçues depuis les premiers travaux. En effet, si, dans le cas du féminin et de la conjugaison verbale, on peut invoquer divers arguments morphologiques en faveur d'un schwa abstrait, peut-on véritablement distinguer des mots comme *patte* et *pat* ou *phare* et *pard* traités par Martinet et bien d'autres comme identiques en fait et en droit. Plusieurs solutions ont été envisagées.<sup>3</sup> On peut suivre la graphie et représenter par exemple *patte* avec un schwa final (soit /patə/) et *pat* sans (/pat/), et postuler un marquage d'exception à la règle de

---

3. Voir par exemple Schane (1967 : 46-49, 1968 : 8-9), Dell (1970 : 59-64), Selkirk (1972 : 326-333) pour une discussion de cette question dans les travaux fondateurs de la phonologie générative.

truncation pour tous les mots se terminant par une consonne (en l'occurrence *pat*). Mais n'est-ce pas faire porter à la graphie un poids démesuré quand on sait (a) qu'elle n'est pas maîtrisée de façon identique par tous les locuteurs, (b) que la liaison existe dans des dialectes sans tradition orthographique ; et (c) que les consonnes de liaison peuvent se stabiliser dans l'acquisition du langage chez des sujets alors même que la graphie n'est pas encore maîtrisée. Dell (1982), soulignant qu'une opposition entre des mots comme *patte* et *pat* est impossible à acquérir sans recours à ces systèmes métalinguistiques que sont l'orthographe et l'enseignement explicite, en conclut qu'on doit donner la même représentation aux deux classes de mots <XC#> et <XCe#>. Il revient donc dans la deuxième version (1985) de son ouvrage *Les règles et les sons* à la solution qu'il avait adoptée dans sa thèse de doctorat (cf. Dell, 1970 : 62-63) et qui traitait toutes les consonnes finales réalisées comme des attaques de syllabe à voyelle vide (schwa) : « *Pacte* et *tact* ont (à la consonne initiale près) des représentations phonologiques identiques, et la distinction entre les deux classes de mots est de nature purement orthographique. » (1985 : 236, note 39).

La question est alors de savoir comment justifier, en dehors du sauvetage de la règle de truncation, ce sur-codage massif des mots du français standard. On notera au passage que, si l'on adopte ce sur-codage posant un schwa à toute consonne finale, on perd alors l'unité supradialectale que permettait un traitement plus proche de l'orthographe. Par ailleurs, le postulat selon lequel tout mot se terminant en surface par une obstruante se ramène au type /Cə#/ au niveau sous-jacent a des conséquences parfois surprenantes. Ainsi, Dell (1985 : 189, note 20) se voit contraint de traiter comme exception une forme comme *sept* : « Si *sept* avait la représentation /setə/, on s'attendrait à ce que le schwa final soit maintenu devant un *h* aspiré. Or *sept housses* se prononce [setus] ou [setʔus], mais pas [setəus]. Force est donc d'admettre que certaines obstruantes sont des exceptions à la règle TRONC. »

Les phonologies plurilinéaires ont partiellement changé la donne en ce qui concerne le schwa. En effet, en traitant les consonnes de liaison comme flottantes, elles permettent de traiter les autres consonnes comme fixes, autrement dit comme

intégrées à la structure syllabique du mot sans nous forcer à postuler des schwas protecteurs. Mais la question n'a pas été close puisque même des travaux très empiriques et à orientation sociolinguistique comme ceux d'Encrevé maintiennent que dans certains cas une opposition est encore vivace. Ainsi propose-t-il une représentation phonologique différente pour les mots morphologiquement féminins du type <XCe> et les mots du type <XC> correspondants (*ourse* vs. *ours*, *patte* vs. *pat*, *nette* vs. *net*). Il déclare : « Nous faisons ce choix sur l'argument, peut-être pas décisif mais non dénué de poids, que, en dépit du fait que l'énoncé [lursəblø] ne permet pas de reconnaître s'il s'agit du masculin *l'ours bleu* ou du féminin *l'ourse bleue*, si l'on dit *non pas l'ours* [lurs] *mais l'ourse* [lursə], tout francophone comprendra que le [ə] désigne le féminin. C'est donc que quelque chose différencie dans la structure sous-jacente *ours* et *ourse* quant à la probabilité de l'épenthèse d'un schwa dans le contexte de la pause où elle est normalement exclue, distinction qui nous paraît bien traduite par l'opposition entre *vraie épenthèse* (celle qui peut affecter le masculin *ours*) et *fausse épenthèse* (celle qui peut affecter le féminin *ourse*). »

Nous pensons qu'il faut rester prudent sur le comportement des finales en français. Il serait surprenant que l'histoire du français n'ait pas légué aux systèmes synchroniques des traces d'oppositions antérieures en partie corrélées à la graphie. Ainsi, a-t-on pu nous faire remarquer qu'il existait des variétés non méridionales où les diverses formes en <e> d'un verbe comme *parquer* (cf. *je parque*, *tu parques*, *il parque*) seraient systématiquement différentes des réalisations du mot *parc*. On y opposerait *il parque ses moutons* [il parkə se mutō] à *un parc séduisant* [œ park sedɥizā] et non pas \*[œ parkə sedɥizā]. Mais la triple question du legs de l'histoire, de son statut en synchronie et du couplage graphie/phonie ne peut être établie à la volée, en comparant mentalement des formes entendues dans divers contextes. Pour établir de tels « faits », il nous semble que les deux meilleures stratégies sont les suivantes : soit on dispose de grands corpus qui permettent d'établir ces phénomènes de manière systématique, soit on construit des expériences contrôlées par des psycholinguistes qui ne consistent pas seulement à

demander à des sujets comment ils prononceraient tel ou tel mot dans tel ou tel contexte. Les corpus rassemblés dans le cadre du projet PFC ont plusieurs avantages. Tout d'abord, ils sont en partie construits pour vérifier diverses hypothèses sur la présence ou l'absence du schwa. Deuxièmement, ils permettent de comparer la lecture à haute voix et la conversation, et donc de vérifier diverses corrélations possibles entre prononciation et écriture. Enfin, à travers transcriptions et codages, ils se prêtent à des exploitations à la fois générales et locales. Cette stratégie est expliquée en (3) avant de nous pencher en (4) sur quelques résultats de ce travail.

### **3. Le projet PFC**

#### ***3.1. Généralités***

Le projet PFC vise à l'élaboration d'un vaste corpus oral couvrant plus de 400 heures d'enregistrements et un minimum de quarante points d'enquête. L'axe central de PFC est sociophonologique (Durand, Laks et Lyche, 2002 ; Durand et Lyche, 2003) sans suivre les contraintes d'une enquête sociologique *stricto sensu*. S'il se situe dans la continuité des grandes enquêtes phonologiques, PFC n'embrasse pas à la lettre une démarche sociolinguistique au sens technique du terme dans la mesure où l'échantillonnage dans chaque point d'enquête n'est pas déterminé par un nombre élevé de témoins sélectionnés aléatoirement et en équilibrant de façon rigoureuse tous les paramètres jugés essentiels (sexe, âge, classe sociale, provenance ethnique, etc.). Notre enquête, à la différence de nombreuses enquêtes sociolinguistiques, ne se focalise pas non plus sur quelques variables prises dans l'ensemble du système afin de les corrélérer à des facteurs stylistiques ou sociaux. Nos objectifs sont plus descriptifs et plus globaux. Il s'agit dans un premier temps d'envisager tous les volets phonologiques de la variation pour pouvoir par la suite proposer des analyses de phénomènes typiquement variables comme le schwa ou la liaison par exemple. La couverture géographique ne prétend évidemment pas à l'exhaustivité mais les points d'enquête quadrillent l'Hexagone et comprennent aussi d'autres pays

*Le schwa en français : pourquoi des corpus ?*

francophones comme le Canada, la Suisse, la Belgique et des départements d'outre-mer comme La Réunion.

Ce travail se rapproche, dans sa conception, de l'enquête préparatoire de Walter (Walter, 1982), travail important parce qu'il constitue une enquête sur les variétés du français qui s'appuie sur des méthodes partagées par PFC. H. Walter plaide par exemple pour des enquêtes linguistiques et sociolinguistiques qui fassent une place fondamentale aux idiolectes et note que :

« William Labov, dans *Sociolinguistique*, indique bien en effet qu'il a commencé son étude par une enquête préparatoire de 70 interviews (p. 95), mais à aucun moment il ne donne de précisions sur les différents systèmes de ses informateurs. Le lecteur n'a connaissance de cette enquête préliminaire que dans la mesure où elle a permis à l'auteur d'isoler les quelques variables phonologiques pouvant être mises en relations avec la différenciation sociale. » (1982 : 18).

Dans PFC, la description porte sur chaque locuteur enregistré avant de s'engager sur le terrain des généralisations pour une région donnée. Walter offre une discussion intéressante des paramètres qui peuvent influencer la variation : l'âge, le sexe, la catégorie socioprofessionnelle, le niveau culturel (ou niveau d'instruction), la situation géographique et le style. PFC est sensible à l'ensemble de ces paramètres et les note de façon rigoureuse. Cependant, la taille de nos échantillons, à savoir entre dix et vingt locuteurs par point d'enquête, ne permet pas d'étudier exhaustivement chacun de ces paramètres. PFC demande une répartition égale entre les sexes pour chaque point d'enquête, au moins deux classes d'âge, un minimum de scolarisation (vu la présence de lecture à haute voix) et, dans la mesure du possible, un certain échantillonnage des catégories socioprofessionnelles. En revanche, notre approche qui s'inspire fortement des travaux des Milroy (1987) en Grande-Bretagne consiste à travailler dans des réseaux denses (famille, amis, amis d'amis) afin d'obtenir des données aussi naturelles que possible. La dimension principale dans

cette approche de la variation est le facteur géographique couplé à la différenciation stylistique.

La force de PFC réside dans le rassemblement de données comparables pour tous les points d'enquêtes grâce aux quatre registres requis dont deux de lecture et deux de dialogue. Tous les témoins lisent une liste de 94 mots se terminant par des paires minimales, un petit texte, participent à un entretien semi-directif (que nous appelons « guidé ») et à une conversation non guidée ou « libre » dans la terminologie PFC. Dans l'entretien guidé le témoin est dans une situation d'interview où l'enquêteur n'est pas un proche mais un ami d'un ami (Milroy, 1987 ; Bourdieu, 1993) et idéalement, les enquêteurs évoluent par groupes de deux, où une personne seulement est un familier des témoins.

La personne étrangère au groupe est responsable de la conversation guidée alors que la conversation libre est du ressort du familier du groupe. Le recueil de données aussi naturelles que possible constitue l'objectif de la conversation libre même si le caractère toujours artificiel d'une telle situation (Gadet, 2000) est susceptible de fausser légèrement les données, ne serait ce que du fait même de la présence incontournable du magnétophone. Afin de minimiser le paradoxe de l'observateur, inévitable lors d'un enregistrement, deux ou trois personnes appartenant à un réseau dense sont mises en présence, et de préférence dans un lieu qui leur est familier. Dans cette situation de conversation libre, la présence de l'enquêteur n'est pas indispensable.

La dimension géographique, nous l'avons dit, est centrale dans PFC. Les points d'enquête ont été choisis en tenant compte de la variété dialectale du français déjà décrite et dans un souci de maintenir un certain équilibre entre zones rurales et zones urbaines, zones côtières ou non. Walter (1982) insiste aussi sur le caractère prévalent de cette dimension qui à son avis prime sur la dimension sociale :

« Il nous a semblé qu'il fallait d'abord faire porter nos recherches sur les facteurs géographiques. Mais donner la priorité à l'étude de la diversité sur le territoire ne signifie pas que l'on considère comme nulles les différences linguistiques liées aux milieux

### *Le schwa en français : pourquoi des corpus ?*

sociaux, à l'intérieur de chaque région. [...] Ce que l'on peut simplement avancer, c'est que toutes les enquêtes réalisées à ce jour montrent que les différences entre un lettré toulousain et un lettré parisien sont beaucoup plus considérables que celles que l'on peut constater entre un lettré toulousain et un ouvrier de la même ville, ou entre un lettré parisien et un ouvrier parisien. » (Walter, 1982 : 52).

Ce type d'hypothèse est précisément ce que nos enquêtes nous permettent de confirmer ou d'infirmier.

#### **3.2. *Le schwa et son codage***

L'enquête PFC vise à livrer des résultats d'une grande généralité sur la présence ou l'absence de schwa. Les données que nous recueillons sont de deux ordres : les données sur schwa dans le texte lu et les données dans la parole spontanée (conversations guidée et libre). Nous procédons à des codages à partir de la graphie, mais sans nous rendre esclaves de cette dernière, car elle fournit des renseignements précieux sur les sites possibles déclenchant la présence ou l'absence d'un schwa. De tels codages ne constituent qu'une étape préliminaire de toute analyse mais cette étape nous semble indispensable. Il s'agit de balayer largement le corpus et de recueillir des données statistiques qui permettront quelques grandes généralisations et qui serviront d'assise à des analyses plus fines. Une fois le corpus constitué, nous serons en mesure de comparer les prononciations d'un très grand nombre de locuteurs sur des bases similaires. La mise au point d'utilitaires informatiques dédiés à des tâches de comptage et de classement de ces phénomènes permet d'automatiser et d'uniformiser les observations. Les variétés de français n'ont pas, à notre connaissance, été abordées avec des principes de codage aussi explicites que les nôtres et nous pouvons donc espérer obtenir des données d'envergure comparables.

Le codage est effectué sur la base de la transcription orthographique dans PRAAT sur une tire spécifique. Une fois le codage terminé, notre utilitaire *classeur-codeur* (voir 3.3) permet d'effectuer un comptage et un classement des présences/absences de schwa (cf. annexe).

Avant de passer aux conventions de codage, nous considérerons un bref exemple concret. Soit l'énoncé : « Je vois vraiment pas comment ce petit peut deviner si vite. » Supposons qu'il est prononcé //§ J' vois vraiment pas comment ce p'tit peut deviner si vit' § // (dans une transcription informelle). Il sera codé : « Je0132 vois vraiment pas comment ce1112 pe0212tit peut de1212viner si vite0413 ».

L'avantage d'un tel codage est qu'il indique non seulement la présence ou l'absence d'un schwa mais aussi quels sont les contextes qui ont déterminé l'apparition ou la disparition de ce segment. Ainsi, le codage « Je0132 » indique que le schwa n'est pas prononcé (0), que le mot est monosyllabique (1), qu'il est situé en début de groupe intonatif (3) et que le mot suivant (à savoir *vois*) commence par une consonne (2).

### 3.2.1 Le système de codage

Le système de codage de schwa a pris comme assise les traitements traditionnels, en particulier Dell (1985). Un exposé complet de la méthodologie PFC dépasserait largement le cadre du présent article : nous renvoyons donc à Durand, Laks & Lyche (2001) et Durand & Lyche (2003 : 244-251 et *passim*) pour des descriptions exhaustives des codages et une justification de la méthodologie.

Pour chaque locuteur, nous codons le texte lu, trois minutes de conversation guidée et trois minutes de conversation libre. Ce codage, que nous faisons sur une base auditive, demande au transcripneur un important travail d'écoute et de concentration : il délivre entre 4000 et 6000 codages par point d'enquête. Il est inévitable que des erreurs se glissent dans les annotations et que des sites soient omis. Malgré les manques évidents, coder les sites de schwa à grande échelle nous semble important pour faire progresser nos connaissances à ce stade. Notons qu'un codage homogène permet une meilleure comparaison des données, que la masse des observations compense les erreurs éventuelles et que le décodage automatique qu'autorisent nos outils nous garantit un repérage simple des sites oubliés.

Nous partons *en général* du *e* graphique car il représente dans la plupart des cas une voyelle latente, c'est-à-dire que l'on observe un lien direct entre la présence de cette voyelle dans la graphie et la possibilité d'une réalisation acoustique de ladite voyelle. Nous reprenons la définition de Malmberg (1976): « Cet élément vocalique à prononciation facultative correspond au *e* de l'orthographe non suivi de consonne appartenant à la même syllabe ». Autrement dit, notre point de départ consiste à annoter tous les schwas graphiques, toutes les voyelles écrites « e », qu'elles soient prononcées ou pas (cf. Eychenne & Hambye, 2004, pour une discussion du domaine de codage). Cette prononciation peut être celle d'un schwa phonétique ou une voyelle moyenne antérieure, arrondie (la voyelle de *jeu*, ou de *peur*).

Il est ici important de souligner que le codage et l'analyse qui en résulte sont deux choses bien distinctes dans la stratégie PFC. Les codages ne renseignent que sur la présence / absence d'une réalisation vocalique dans un contexte donné, sans s'engager en aucune manière sur son statut phonologique et théorique. C'est bien à l'analyste, et non au codeur, qu'il incombe d'interpréter cette réalisation vocalique. Par ailleurs, nous insistons sur le fait que la stratégie PFC, si elle peut être considérée d'un certain point de vue comme minimaliste, n'est en rien limitative : à titre d'exemple, une personne qui considérerait qu'il y a un schwa dans des mots comme *monsieur* ou *faisait* pourrait tout à fait étendre le codage dans une nouvelle tire de l'outil Praat pour prendre en compte ces unités.

Tous les *e* graphiques sont donc soumis au codage, à deux exceptions près. Nous excluons, après de nombreuses préenquêtes, les *e* après voyelle (*vue*) même s'ils peuvent dans certaines variétés provoquer l'allongement de la voyelle précédente. Si tel est le cas, le transcripateur indiquera cet allongement dans ses commentaires généraux afin d'ouvrir la voie à une étude plus détaillée. En tout état de cause, si une variété présente des prononciations du type [vyə] pour *vue*, il est possible de l'indiquer explicitement et d'étendre le codage pour les prendre en compte.

Nous excluons par ailleurs les *e* graphiques après un groupe obstruante+liquide (*grenade, mercredi*) à l'intérieur des mots, mais on les codera en position finale (*entre*). En effet, les schwas après obstruante+liquide à l'intérieur des mots sont généralement considérés comme non sujets à effacement (cf. Morin, 1978 ; Morin, 1988 ; Durand *et alii*, 1987). Il faut à nouveau signaler deux choses. Tout d'abord, si des effacements sont attestés dans ces positions, notre système de codage permet là encore de les noter. Deuxièmement, s'il n'y a jamais effacement, comme c'est le cas dans les enquêtes que nous avons nous-mêmes réalisées, nous laissons à l'analyse le soin d'établir le statut phonologique des voyelles correspondant à ces 'e' graphiques. Par exemple, dans les variétés du midi que nous connaissons, la réalisation de 'e' dans un mot comme *brebis* est la même que la réalisation de 'eu' dans un mot comme *breuvage*, à savoir [ø]. En l'absence d'effacement ou d'alternance indiquant que le premier 'e' correspondrait à un 'schwa', nous ne voyons pour ces cas aucune raison de créer une opposition entre schwa (/brəbi/) et /ø/ (/brøvaʒø/), et nous attribuons à *brebis* une voyelle stable (/brøbi/<sup>4</sup> = [brøbi]).

En revanche, un schwa graphique après tout autre groupe consonantique est codé (ex : *gouvernement*), dans la mesure où la non réalisation est possible dans certaines variétés, et sous certaines conditions (voir par exemple Morin, 1983 : 87-88 pour le français de Saint-Etienne).

Sont aussi codées comme sites potentiels toutes les consonnes finales prononcées. En effet, un schwa final n'est pas exclusivement observé lorsque la graphie indique clairement la voyelle (*mon père*) ; il peut l'être aussi après une consonne graphique finale de mot comme dans *c'est pas net*[ə]. Pour permettre de mieux comprendre ce phénomène nous codons non seulement les *e* graphiques mais aussi toutes les consonnes finales prononcées (ex : *hasard, net*).

Le codage comprend toujours quatre chiffres qui sont placés immédiatement après le « e » graphique ou avant / après

---

4. Pour ne pas alourdir l'article avec des représentations trop complexes, nous opérons ici avec des représentations linéaires avec archiphonèmes, phonèmes et allophones.

*Le schwa en français : pourquoi des corpus ?*

la consonne concernée. Ces quatre chiffres ont pour objectif d'indiquer la présence ou l'absence de la voyelle et de délimiter les contextes de manière large. Ces chiffres vont de 0 à 5. Nous nous contenterons de résumer ici nos conventions. A titre d'illustration, nous donnons en annexe une partie du texte codée selon ces conventions :

<i>(a) Premier chiffre :</i>	
Schwa présent	1
Schwa absent	0
Schwa incertain	2
<i>(b) Deuxième chiffre :</i>	
Monosyllabe	1
Première syllabe de polysyllabe	2
2ème syll. et suiv. de polysyllabe	3
Dernière syllabe de polysyllabe	4
Métathèse	5
<i>(c) Troisième chiffre :</i>	
Voyelle à gauche (V(#)Cə)	1
Consonne à gauche (C(#)Cə)	2
Début de groupe intonatif	3
Schwa incertain à gauche	4
Groupe consonantique simplifié	5
<i>(d) Quatrième chiffre :</i>	
Voyelle à droite	1
Consonne à droite	2
Frontière intonative forte et/ou fin d'énoncé	3
Frontière intonative faible	4

*3.2.2 Le texte PFC et son codage*

Le texte PFC a été construit avec soin en ce qui concerne le schwa. Nous avons intégré dans le texte tous les contextes recensés dans le travail fondamental de Dell (1985) sur la question. Nous avons donc distingué 10 grands contextes dont nos codages tiennent compte :

(a) Schwa initial d'énoncé

<i>Le <u>v</u>illage</i>	§ Cə # CV
<i>De <u>p</u>lus</i>	§ Cə # OL
(b) Schwa initial de mot polysyllabique	
<i>ses <u>ch</u>emises</i>	V # Cə
<i>en <u>r</u>evanche</i>	V # Cə
(c) Schwa interne	
<i>bê<u>t</u>ement</i>	VCəCV
<i>gou<u>v</u>ernement</i>	VrCəCV
(d) Schwa final postconsonantique devant V	
<i>Minist<u>r</u>e ira</i>	
<i>commu<u>n</u>e au</i>	
(e) Schwa final postconsonantique devant C	
<i>villag<u>e</u> de</i>	
<i>Minist<u>r</u>e ne</i>	
(f) Schwa final postconsonantique d'énoncé	
<i>colè<u>r</u>e.</i>	
<i>Minist<u>r</u>e.</i>	
(g) Schwa final postvocalique marqueur du féminin (présent via la longueur de la voyelle ?)	
<i>tour<u>n</u>ée de</i>	
<i>ann<u>e</u>e.</i>	
(h) Schwa de monosyllabe	
<i>décid<u>e</u> de faire</i>	V#Cə#CV
<i>Il a l<u>e</u> sentiment</i>	V#Cə#CV
(i) Suites de schwas	
<i>sentim<u>e</u>nt de se trouver</i>	V#Cə#Cə#
<i>plutôt que de s<u>e</u> trouver</i>	V#Cə#Cə#Cə # CV
(j) Sites fréquents d'insertion	
<i>Marc<u>_</u>Blanc</i>	
<i>Ouest<u>_</u>Liberté</i>	

Pour illustrer le codage du texte PFC, nous donnons ci-dessous une réalisation possible de ce texte où nous avons indiqué par // des pauses ou des transitions rythmiques. Dans cette transcription informelle, l'absence de schwa est indiquée par une apostrophe, alors que sa présence est notée par le soulignement de la voyelle. Cette transcription est suivie du texte codé selon les conventions PFC.

*Prononciation du texte*

// Le sympathiqu' mair' de Beaulieu // ne sait plus à quel saint se vouer. // Il a le sentiment // de se trouver dans un' impass' stupid'. // Il s'est, // en désespoir de caus', // décidé à écrire' au Premier Ministr' // pour vérifier // si son villag' était vraiment un' étap' nécessair' dans la tournée prévue. // Beaulieu préfèr' êtr' inconnue et tranquill' // plutôt que d'\_se trouver au centre d'un' bataill' politiqu' dont, // par la télévision, // seraient témoins des millions d'électeurs. //

*Texte codé*

Le1132 sympathique0412 maire0412 de1122 Beaulieu ne1132 sait plus à quel0412 saint se1112 vouer. Il0411 a le1112 sentiment de1132 se1112 trouver dans une0411 impasse0412 stupide0413. Il0412 s'est, en désespoir0412 de1122 cause0413, décidé à écrire0411 au Premier Ministre0423 pour0412 vérifier si son village0411 était vraiment une0411 étape0412 nécessaire0412 dans la tournée prévue. Beaulieu préfère0411 être0421 inconnue et tranquille0413 plutôt que1112 de0112 se1122 trouver au centre1421 d'une0412 bataille0412 politique0412 dont, par0412 la télévision, se1232raient témoins des millions d'électeur0413s.

**3.3. Le classeur-codeur**

Le classeur-codeur, outil que nous avons utilisé pour obtenir les résultats présentés en (4), est un outil développé par Julien Eychenne, qui s'inspire des classeurs schwa et liaison de première génération d'Abderrahim Meqqori, dont il étend les fonctionnalités.<sup>5</sup> Cet outil permet d'extraire, de classer et d'interroger à grande échelle les codages PFC pour le schwa et la liaison. Le classeur-codeur fonctionne virtuellement sur toutes les plateformes qui disposent d'un interpréteur Perl, qui est un langage de programmation disponible sous une licence libre. Bien que le programme exige une installation préalable de Perl, il ne présuppose pas la connaissance de ce langage de

---

5. Les outils PFC et les manuels qui les accompagnent sont disponibles sur le site internet PFC : <http://infolang.u-paris10.fr/pfc/outils.htm>

programmation. En revanche, il n'interagit pas avec l'utilisateur au moyen d'une interface graphique classique comme celle de Windows mais nécessite le recours à des lignes de commande de type DOS sous Windows ou shell sous UNIX (Linux et Mac OS X). Nous présentons ci-dessous les principales fonctionnalités de l'outil :

- traiter un fichier pour le schwa : cette fonction permet d'extraire et de classer toutes les occurrences des codages schwa d'un fichier. Il donne par ailleurs des informations statistiques rudimentaires sur la présence/absence de réalisations vocaliques, qui permettent d'avoir une idée générale du comportement de la voyelle.
- Traiter un fichier pour la liaison : cette fonction permet d'extraire et de classer toutes les occurrences des codes liaison dans un fichier.
- Traiter un corpus : l'outil crée pour chaque locuteur un fichier « conversation » (correspondant aux entretiens libre et guidé, que l'on peut facilement comparer à la lecture de texte) et un fichier « global » qui regroupe les 2 entretiens et la lecture de texte. Il crée ensuite un « locuteur virtuel » correspondant à la concaténation de tous les locuteurs du corpus pour chaque tâche (lecture de texte, entretien libre, entretien guidé, « conversation » et « global »). Enfin, l'outil applique les fonctions de traitement du schwa et de la liaison à tous les fichiers obtenus.
- Rechercher un motif : cette fonctionnalité permet d'extraire un motif utilisant la syntaxe des expressions régulières Perl dans un fichier TextGrid. Par exemple, on peut chercher tous les polysyllabes qui n'ont pas de *e* graphique (motif [ ^e ] [ 0-2 ] 4 [ 1-5 ] [ 2-4 ]), ou encore « toutes les occurrences du mot *semaine* dont le premier schwa est réalisé mais dont le deuxième ne l'est pas et est suivi par une consonne », soit le motif `se12\d2maine0412`. Cette fonction, quoique très puissante, requiert l'apprentissage de la syntaxe des expressions régulières.

#### 4. Résultats des corpus

##### 4.1. Le corpus Douzens (Aude)

Comme nous l'avons souligné en introduction, l'enquête Douzens<sup>6</sup> est, du point de vue du schwa, représentative d'accents du Midi conservateurs, tels qu'ils ont pu être décrits dans la littérature (voir notamment Séguy, 1951 ; Borrell, 1975 ; Walter, 1982 ; Durand *et alii*, 1987).

Examinons les données : le corpus Douzens nous délivre 4805 codages schwa. Pour la position finale, à laquelle nous nous intéressons ici, nous disposons de 3136 codages.

La première chose que l'on essaie de vérifier est de savoir s'il existe, pour cette variété, une opposition /C#/ vs /Cə#/ et si elle correspond à une distinction <C> vs <Ce> dans la graphie (du type *mer* ~ *mère*). Si l'on considère tous les polysyllabes en contexte préconsonantique ou prépausal, on observe la distribution suivante :

(1) Distribution de schwa en contexte XC\_{#}{C ; #} à Douzens

##### *e graphique*

voyelle	occurrences	pourcentage (%)
absente	208	11.9
présente	1523	<b>86.8</b>
incertaine	23	1.3
total	1754	100

##### *pas de e graphique*

Voyelle	occurrences	pourcentage (%)
absente	730	96.3
présente	19	<b>2.5</b>
incertaine	9	1.2
total	758	100

Lorsqu'une consonne finale prononcée est suivie d'en *e* dans la graphie, on remarque qu'elle est suivie d'une réalisation

---

6. Douzens est un village viticole, situé approximativement entre Narbonne et Carcassonne.

vocalique dans 86.8% des cas. En revanche, lorsqu'une consonne finale n'est pas suivie d'un *e* dans la graphie, le taux de réalisation n'est alors que de 2.5%. Cet écart se révèle hautement significatif, et démontre clairement l'existence dans cette variété d'une opposition <C#> vs <Ce#>, qui se traduit au niveau phonologique par une opposition /C#/ vs /Cə#/. Nous donnons ci-dessous quelques exemples :

(2) Réalisation avec schwa (*e* souligné)

*une maladie génetique,*  
*en Bourgogne c'est sympa*  
*on habite le village*  
*c'est une persone sympathique*  
*mon autre grand frère.*  
*un litre d'alcool*  
*il faut être de gauche*  
*elle prenait ses livres*  
*moi je la marque,*  
*de lettres modernes*

(3) Réalisation sans schwa (notée par un astérisque)

*Marc\* [rkbl] Blanc*  
*tous les week-end\* [wikend].*  
*conseil\* municipal\*,*  
*c'est le top\*.*  
*elle a pris un box\* [ks##],*  
*mais j'ai habité, Moux\* [ks##],*  
*Je suis né à Douzens\*.*

Par ailleurs, on note une certaine asymétrie dans la distribution du schwa : alors qu'un schwa sous-jacent peut ne pas apparaître en surface (11.9% dans le corpus, ces cas étant essentiellement l'œuvre des jeunes locuteurs<sup>7</sup>, qui se révèlent les plus innovateurs de ce point de vue), un schwa n'apparaît jamais en surface s'il ne correspond pas à un schwa sous-jacent (autrement dit, il n'y a pas de schwa épenthétique).

---

7. Voir l'annexe pour des codages d'un locuteur conservateur pour la position finale.

*Le schwa en français : pourquoi des corpus ?*

Nous relevons cependant 19 occurrences, dont nous donnons la liste en (4), qui montrent que certains items semblent avoir un schwa qui n'est pas reflété par la graphie :

(4) occurrences présentant un schwa sans <i>e</i> graphique		
<i>donc</i>	[dɔ̃kə]	(x12)
<i>avec</i>	[avəkə]	(x2)
<i>vingt-deux</i>	[vɛ̃tədø]	(x3)
<i>vingt-neuf</i>	[vɛ̃tənœf]	(x1)
<i>circuits habituels,</i>	[sirkɥiabituelə]	(x1)

Les données en (4) montrent que les items concernés sont pour l'essentiel *donc*, *avec* et *vingt*. Durand et Tarrier (2003), s'appuyant sur Greimas (1969) soulignent que « *donc* vient historiquement d'un croisement entre *dumque* (alors) et *tunc* (alors) et [que] la graphie *donques* est attestée jusqu'au moins au XVII<sup>ème</sup> siècle » (p. 126). Notons aussi qu'en tant que particule énonciative, *donc* est très fréquemment suivi d'une pause/hésitation (*eu*h), dont le timbre se confond souvent avec schwa. En ce qui concerne *avec*, il est opportun de souligner que la graphie *avecque* est parfaitement établie, parallèlement à la forme *avec* (voir Vaugelas 1647, pp 311-315 qui prescrit de manière très détaillée les contextes où l'on doit préférer l'un à l'autre). Citons encore des alexandrins tels que :

*Lorsqu'avecque deux mots que vous daîgnâtes dire,  
Vous sûtes arrêter mes peines pour jamais,*

extraits du premier quatrain d'un poème de Vincent Voiture (1597-1648)<sup>8</sup>, et qui attestent de ce que la forme *avecque* avait bien un schwa prononcé. Il n'est donc pas impossible qu'une forme /avEkə<sup>9</sup>/ soit présente dans le lexique de certains de nos locuteurs douzenois. Pour la forme *vingt*, on peut défendre à la suite de Séguy (1951) et Durand *et alii* (1987) que la forme *vingt*+numéral (ex : *vingt-deux* [vɛ̃tədø]) a été refaite sur le paradigme des dizaines (*trente*, *quarante*, *cinquante*, etc.). Par

8. <http://poesie.webnet.fr/poemes/France/voiture/9.html>

9. Une majuscule telle que /E/ désigne une voyelle sous-spécifiée pour l'aperture ou un archiphonème (cf. note 4).

ailleurs, il ne fait aucun doute que la prononciation avec schwa, ou du moins avec [t] final réalisé, est fort répandue dans la plupart des usages du nord de la Loire. Enfin, l'occurrence *circuits habituels* [ə] provient de la lecture de texte par un jeune locuteur : cet hapax est suivi d'une longue pause et pourrait être un schwa prépausal ou une scorie.

S'il est vrai que, de manière générale, une opposition graphémique <C> vs <Ce> reflète une opposition phonologique /C/ vs /Cə/, on s'aperçoit, par l'observation rigoureuse de données massives, que l'adéquation n'est pas parfaite et que certains items, à la suite d'aléas historiques, peuvent ne pas traduire cette opposition. Ainsi, nous ne serions pas surpris si une partie des effacements de schwa constatés en fin de polysyllabe devaient en réalité être attribués à une absence de schwa phonologique final malgré un *e* dans la graphie. Par exemple, pour certains (jeunes) locuteurs du Midi qui maintiennent une opposition schwa vs absence de schwa, le mot *pote* est lexicalisé en /pOt/ (= [pɔt]), et ce malgré le *e* graphique.

Intéressons-nous maintenant aux groupes consonantiques complexes correspondant à la graphie <XCCe#C>. Puisqu'un schwa sous-jacent est typiquement réalisé, on s'attend dans ce contexte à ce qu'il le soit (presque) systématiquement. C'est effectivement le cas dans la grande majorité des cas (220 occurrences) : *liste de mots* ; *prendre le souffle* ; *risquent quoi* ; *double volant* ; *Charles par exemple* ; *organisme peut-être* ; etc. Nous relevons cependant quatre cas d'effacement : un seul cas concerne un groupe liquide+obstruante (locuteur TG1 : *ça march' bien*) et trois occurrences où un groupe consonantique est simplifié : *quat' kilomètr'* (loc. TG1) ; *activis' des communes voisines* (loc. TG1, lecture) ; *cent kilomèt' non, à peu près ?* (loc. MG1). Ces occurrences sont réalisées par les deux locuteurs les plus innovateurs du corpus (un frère et sa sœur). Il nous semble ici utile de relever que ces locuteurs vivent à l'écart du village et sont partis faire leurs études dans les villes alentour (Narbonne et Carcassonne), et l'on peut raisonnablement supposer que le brassage linguistique auquel ils ont pu être confrontés a

influencé leur système dans le sens d'une perte de certains schwas.

Lorsqu'il n'y a pas de schwa sous-jacent, en revanche, il n'y a jamais d'épenthèse (aucune occurrence de schwa ne correspondant pas à un *e* graphique dans ce contexte) : *Marc\* Blanc* ; *Ouest\* France*. Ce point doit être nuancé du fait que le contexte où un groupe consonantique final sans schwa suivi d'un mot à initiale consonantique est rare, dans le lexique français, et *a fortiori* dans le corpus : les occurrences que nous relevons ne proviennent donc que de la lecture, où les locuteurs gardent un certain contrôle sur leur élocution. Il est probable que ces mêmes contextes auraient connu plusieurs simplifications consonantiques (type *Oues'France*) dans un style moins surveillé. Néanmoins, nous soutenons que cette variété ne connaît pas de schwa épenthétique, contrairement à Treize-Vents, par exemple (cf. 4.3), où l'insertion d'un schwa en lecture est plus fréquente (cf. *Marc [ə] Blanc*). La simplification des groupes consonantiques, si elle doit avoir lieu, se fait à Douzens par suppression d'une partie du matériel consonantique, et non par insertion d'un noyau vocalique. Le fait que schwa n'est pas mobilisable pour l'épenthèse comme il l'est dans des variétés où sa réalisation en position finale s'est érodée est peut-être dû à son rôle actif dans la morphophonologie de cette variété (*e* du féminin, *net* ~ *nette* ; flexion verbale).

#### **4.2. Le corpus Biarritz (Pays Basque)**

Le corpus du Pays Basque met au jour une variété du Midi en évolution : bien que présentant un système vocalique « typique » du Midi, avec, outre le schwa, 7 voyelles orales /i y u E œ O a/, où les lettres capitales représentent des voyelles sous-spécifiées pour l'aperture, et 4 « voyelles nasales » /E<sup>N</sup>, œ<sup>N</sup>, O<sup>N</sup>, a<sup>N</sup>/ (cf. Durand et Tarrier, 2003), cette variété présente un comportement pour le schwa qui tend à se rapprocher d'un accent plus standard. Le schwa posttonique de position finale, position dont on sait qu'elle est faible, y est atteint au premier plan.

Nous disposons pour ce corpus de 6511 codages, dont 4191 pour la position finale (tous contextes confondus). Si nous

regroupons tous les polysyllabes en contexte prépausal ou préconsonantique, comme nous l'avons fait à Douzens, nous obtenons la répartition rapportée en (5) :

(5) Distribution de schwa en contexte XC\_# {C ; #} à Biarritz

*e graphique*

voyelle	occurrences	pourcentage (%)
absente	1048	48,2
présente	1023	<b>47,1</b>
incertaine	103	4,7
total	2174	100

*pas de e graphique*

voyelle	occurrences	pourcentage (%)
absente	998	91,5
présente	74	<b>6,8</b>
incertaine	19	1,7
total	1091	100

Dans ce contexte, il apparaît que, lorsqu'une consonne finale prononcée précède un *e* dans la graphie, elle est suivie d'un schwa dans 47,1% des cas. En revanche, si la consonne n'est pas suivie d'un *e* dans la graphie, le schwa n'est réalisé que dans 6,8% des cas. Ainsi, le taux de réalisation sur l'ensemble du corpus est presque sept fois supérieur s'il y a un *e* graphique, ce qui nous semble hautement significatif. Néanmoins, une fois sur deux, un *e* graphique ne correspond à aucune manifestation phonétique dans la chaîne parlée.

Pour pouvoir interpréter ces données de manière pertinente, il est nécessaire de passer à un niveau de détail plus fin : ces chiffres globaux donnent certes une idée du comportement de schwa dans notre corpus et peut-être, moyennant extrapolation, au Pays Basque, mais ils ne nous renseignent pas sur le système intégré par chacun des locuteurs. Ainsi, plusieurs pistes s'offrent *a priori* à nous pour expliquer ce taux d'effacement important : soit le corpus est relativement homogène, et l'on s'attend à ce que les locuteurs aient des taux d'effacement très proches (dans ce cas, on peut supposer qu'ils maintiennent tous une opposition phonologique, et il reste à élucider les facteurs favorisant l'effacement de la voyelle) ; soit

*Le schwa en français : pourquoi des corpus ?*

le corpus est hétérogène, avec d'un côté des locuteurs conservateurs maintenant le schwa et de l'autre des locuteurs « innovateurs » ayant perdu l'opposition /C#/ vs /Cə#/. Nous donnons ci-dessous (6a), pour les mots de type XVC\_# et toutes tâches confondues, le taux de réalisation de schwa en contexte prépausal et préconsonantique.

(6a) Taux de réalisation d'un schwa correspondant à un *e* graphique pour chaque locuteur de Biarritz en contexte VC\_#{C;#}

locuteur	e graphiques codés	schwas réalisés	réalisation (%)
Jl1	138	86	62,3
JM1	115	67	58,2
AB1	130	67	51,5
MA4	195	96	49,2
ST1	95	46	48,4
MA3	148	59	39,9
MA1	103	41	39,8
MA2	126	48	38,1
SL1	131	47	35,9
JC1	164	52	31,7
PL1	149	40	26,8
PI1	165	36	21,8

Les données traduisent une forte variabilité inter-individuelle, avec une amplitude allant de 21% à 62.3% de réalisation. Cette variété est donc peu homogène quant à la réalisation du schwa : les locuteurs qui présentent le système le plus conservateur (Jl1 et JM1) ont un taux d'effacement proche de 40% dans le contexte étudié, ce qui montre l'érosion du schwa final dans cette variété. En revanche, les locuteurs ayant le système le plus innovateur (PI1 et PL1) maintiennent toujours un schwa dans plus de 20% des cas. Nous interprétons ces chiffres comme la traduction d'une évolution lente d'un système du Midi traditionnel vers un système plus proche de la norme. Il nous est cependant difficile d'affirmer catégoriquement que les locuteurs innovateurs ont complètement perdu l'opposition /C#/ vs /Cə#/. Par exemple, alors qu'il réalise

21,8% de schwas correspondant à un *e* graphique dans le contexte VC\_# {C;#}, le locuteur PI1 ne réalise que 2,7% de schwas (soit 4 réalisations pour 146 codages) dans ce même contexte lorsqu'il n'y a pas de *e* graphique. Or, si le locuteur ne disposait plus que d'une seule classe /Cə#/, on devrait s'attendre à ce que les 2 classes s'équilibrent. Les occurrences de schwa pour les mots du type /C#/ sont par ailleurs limitées au contexte prépausal en conversation (*milieu rural* ə // ; *et alors* ə // ; *cinq jours* ə // ; *à Biarritz* ə //), et l'on ne doit pas exclure la possibilité que ce schwa soit comparable au schwa prépausal du français standard (cf. Berrit Hansen, 2003).

A l'appui de l'affirmation selon laquelle cette variété serait en évolution d'un système du Midi traditionnel vers un système plus « standard », nous donnons en (6b) la corrélation taux de réalisation/âge du locuteur. Nous avons classé les locuteurs par génération, en prenant pour point de référence le locuteur MA3, co-enquêteur avec JD :

- génération 1 : JM1 (grand-mère de MA3)
- génération 2 : MA1 (fille de JM1, mère de MA3) et MA4 (oncle de MA3)
- génération 3 : MA3, MA2 (soeur de MA3), et tous les autres locuteurs (connaissances de MA3).

(6b) corrélation taux de réalisation du schwa en contexte VC\_# {C;#} / âge du locuteur (au moment de l'enquête)

locuteur	âge	génération	réalisation (%)
J11	61	2	62,3
JM1	92	1	58,2
AB1	30	3	51,5
MA4	74	2	49,2
ST1	35	3	48,4
MA3	39	3	39,9
MA1	66	2	39,8
MA2	38	3	38,1
SL1	35	3	35,9
JC1	38	3	31,7
PL1	26	3	26,8
PI1	31	3	21,8

*Le schwa en français : pourquoi des corpus ?*

Idéalement, on devrait s'attendre à ce que les locuteurs les plus âgés (génération 1 et 2) soient ceux qui maintiennent le plus grand nombre de schwas, et qu'inversement les jeunes (génération 3) soient ceux qui l'effacent le plus. Une telle répartition irait dans le sens d'une évolution diachronique de cette variété.

Les données présentées en (6b) nous semblent globalement conforter cette hypothèse : ainsi, les locuteurs JI1 et JM1, tous deux « séniors », sont ceux qui maintiennent le plus le schwa, alors que les locuteurs JC1, PL1 et PI1, parmi les plus jeunes, sont aussi ceux qui affichent une nette tendance à l'effacement. La situation est néanmoins complexe, et certains locuteurs semblent contredire cette tendance. Considérons par exemple le réseau familial de MA3 qui, comme nous l'avons dit plus haut, représente les 3 générations. La locutrice la plus âgée, JM1 (la grand-mère de MA3), a un taux de réalisation relativement élevé (58,2%). Le locuteur MA4 (l'oncle de MA3), a un taux de réalisation nettement moindre (49,2%), mais MA1 (la mère de MA3), qui est de la même génération que MA4, et pour laquelle on pourrait par conséquent espérer un taux de maintien voisin, a un taux de réalisation (39,8%) presque identique à celui de ses enfants MA2 (38,1%) et MA3 (39,9%), lesquels représentent la troisième génération. Par ailleurs, au sein de la troisième génération, on observe des variations très importantes, allant de 21,8% (PI1) à 51,5% (AB1).

Notre interprétation de ces chiffres est que le système est réellement en évolution, mais que cette évolution est disparate : les jeunes locuteurs *tendent* à adopter un système plus standard, mais ils ne le font peut-être pas au même rythme et au même degré. Un élargissement de la base empirique, en particulier pour les générations 1 et 2, pourrait permettre de mieux cerner le rôle de l'âge dans l'évolution du français au Pays Basque.

Intéressons-nous maintenant aux groupes consonantiques complexes du type XCC(ə)#C. Ici comme ailleurs, l'empirie ne reflète pas exactement l'image idéale qu'en attend le linguiste. Ainsi, bien qu'un schwa se réalise dans la grande

majorité des cas (245 réalisations pour 351 occurrences, soit 72,4%), on relève 18 occurrences de non réalisations, très atypiques pour une variété du Midi, comme les exemples reproduits en (7) :

(7) non réalisation d'un schwa dans le contexte CC\_#C  
*D'un autr' côté,*  
*dix kilomètr's de*  
*Premier Ministr' ne cesse (x2)*  
*mes oncl's qui ont travaillé*  
*Premier Ministr' lassé*  
*par contr' moi*  
*il faut êtr' quand même*  
*notr' liberté*  
*peut-êtr' moins*  
*Je parl' basque,*  
*je parl' moi*  
*parl' bien (x3)*  
*l'agrégation extern' d'économie-gestion*  
*rappor't quand même*  
*l'accent circonflex' par exemple*

En ce qui concerne le statut « épenthétique » du schwa (i.e. des réalisations vocaliques ne correspondant à aucun *e* graphique), on en relève 9 occurrences dans le corpus (6 en lecture et 3 en conversation) :

(8) réalisation d'un schwa dans le contexte CC\_#C  
*Marc ə Blanc (x3)*  
*Ouest ə France (x3)*  
*film ə très bien (x2)*  
*concept ə de*

Ces exemples démontrent que cette variété est clairement en évolution par rapport à des accents plus conservateurs comme l'accent de Douzens présenté en (4.1). Il faut aussi se demander, et ce pour chaque locuteur, si un schwa est toujours présent dans des mots du type *quatorze*, *ministre*, ou si une restructuration plus poussée a eu lieu permettant des

codas consonantiques complexes comme /str/ à la finale de mot (ce qu'interdisent divers cadres théoriques). L'apparition d'un schwa se réduirait alors à une fonction de « lubrifiant phonétique » dans des jonctures complexes. Une étude de la corrélation statut lexical vs épenthétique du schwa dépasserait le cadre et la portée du présent article. Les outils en cours de développement, qui permettront une exploration plus fine des données, ainsi que les enquêtes en voie d'achèvement, devraient y apporter des éléments de réponse.

#### **4.3. Le corpus Treize-Vents (Vendée)**

Nous disposons pour le codage de l'enquête PFC à Treize-Vents en Vendée de 4117 codages, dont 2787 pour la position finale.

Considérons tout d'abord les 1902 codages, reportés en (9), pour la fin de polysyllabes du type graphique <XCe#> (*colère*) et <XC#> (*tunnel*) en contexte prépausal ou préconsonantique.

(9) réalisation d'un schwa en contexte XC\_{#}{C ; #} dans le corpus Treize-Vents

<b>Voyelle</b>	<b>occurrences</b>	<b>%</b>
Absente	1640	86.2
Présente	251	13.2
Incertaine	11	0.6
Total	1902	100

Aucun schwa n'est codé dans 86.2% des cas (soit 1640 codages négatifs) ; une voyelle réduite est perçue dans 6.6% des cas (soit 251 codages). Cette voyelle réduite, un schwa putatif, est le plus souvent présente dans des cas d'hésitation (précédant un *euh*) dans nos transcriptions orthographiques. Notre analyse est que la présence d'une voyelle de type schwa ne permet absolument pas d'opposer deux classes de mots comme c'était le cas dans le français de l'enquête Douzens (Aude). La grande disparité entre le nombre de codages présence (251 codages) vs. absence (1640 codages) va à l'encontre de l'hypothèse d'un équilibre entre deux comportements bien différenciés. En

délaissant les chiffres, on ne trouve pas de mots qui auraient un comportement stable dans chacune des classes examinées.

Pour un tel accent, les données confortent l'hypothèse que tous les mots sont de type /C#/ au niveau sous-jacent et que l'alternance entre [C#] et [Cə#] est liée à des facteurs d'ordre rythmique, dont le cas le plus prototypique est *BonjourE* (cf. Anita Berit Hansen, 2003, et les références qu'elle fournit). Bien sûr, le statut phonologique de ces séquences reste sujet à débat puisque certains cadres postulent que toute consonne phonétique finale correspond à une séquence consonne + voyelle vide (Scheer 1999, 2000, et dans une forme moins radicale, Angoujard, 1997). Néanmoins, si l'on postule que toute consonne finale est suivie d'un noyau vide, lequel peut, sous certaines conditions, se réaliser comme schwa, il faudra pouvoir distinguer ces vraies voyelles vides n'ayant de réalisation qu'épenthétique du schwa lexical tel qu'on le rencontre à Douzens et dans les accents du Midi traditionnel.

Si l'on s'intéresse maintenant au couplage graphie/phonie, on s'aperçoit là aussi que la présence d'une voyelle en surface n'est pas fortement corrélée à la graphie. Expliquons-nous. En français, le nombre de mots se terminant par une consonne graphique (correspondant à une consonne phonétique prononcée type *coq*) est nettement inférieur au nombre de mots se terminant par Consonne + e (type *coque*). En nous appuyant sur 170 731 items ayant une consonne finale prononcée<sup>10</sup>, nos calculs révèlent en effet que seuls 14 207 n'ont pas de *e* graphique, soit 8.3% du total. Ainsi, même s'il s'avérait qu'il y a davantage d'occurrences de type <XCe#> présentant une réalisation vocalique, cela ne démontrerait pas que l'apparition d'un schwa est favorisée par la graphie dans la mesure où le type <XCe#> en français est bien plus répandu que <XC#>. En effet, un schwa épenthétique après consonne à la finale a environ neuf fois plus de chances d'être attesté dans un contexte <XCe>. Fort de ce constat, l'examen des données ne nous semble que peu sujet à controverse :

---

10. Ces résultats ont été obtenus par examen de la base de données lexicales BDLex, développée à l'IRIT (Université Paul Sabatier, Toulouse).

(10) Corrélation graphie/phonie pour le contexte XC\_{#}{C ; #} à  
Treize-Vents

*e graphique*

<b>Voyelle</b>	<b>occurrences</b>	<b>%</b>
Absente	1014	83.9
Présente	183	<b>15.2</b>
Incertaine	11	0.9
total	1208	100

*pas de e graphique*

<b>Voyelle</b>	<b>occurrences</b>	<b>%</b>
Absente	626	90.2
Présente	68	<b>9.8</b>
Incertaine	0	0
total	694	100

On voit en (10) que les deux classes s'équilibrent (*a fortiori* si l'on néglige la fréquence textuelle ('token frequency') de certaines formes du type *donc*), puisqu'une voyelle est prononcée dans 15.2% des cas si elle correspond à un *e* graphique, et dans 9.8 % des cas si elle ne correspond à aucun corrélat graphique. La différence nous semble ici trop faible pour pouvoir être interprétée comme significative<sup>11</sup>.

Pour renforcer cette affirmation, il nous semble opportun d'exposer les résultats de la lecture et de la conversation : s'il existe une corrélation graphie/phonie, on est en droit de s'attendre à ce que la lecture favorise l'émergence de schwas qui sont par ailleurs « effacés » en conversation. Or, dans le contexte XC\_{#}C, nous relevons 16.3% de réalisation vocalique (soit 78 occurrences sur 478) en lecture contre 14.6% de réalisation en conversation (soit 105 occurrences sur 717). Ce faible écart (qui doit encore une fois être modulé par la fréquence textuelle de certains items en conversation) nous

---

11. On nous a fait remarquer, et ce à juste titre, qu'un test khi-carré permettait de déterminer de façon fiable si les deux classes s'équilibrent ou non. Nous restons toutefois réservés quant à l'applicabilité de ce test dans le cas présent, en raison de la fréquence élevée de certains items codés (*il, elle, donc...*).

paraît insignifiant et conforte d'après nous l'hypothèse selon laquelle la graphie ne reflète pas pour la variété en question un système qui opposerait des finales /C#/ vs /Cə#/ . Par ailleurs, rien dans nos données ne soutient l'idée qu'un mot féminin ou une terminaison verbale déclenche plus systématiquement une réalisation vocalique que la présence d'une simple consonne.

## 5. Conclusion

L'objectif que nous avons poursuivi tout au long de cet article est double : il s'agissait d'une part de présenter quelques traits du comportement du schwa en position finale dans plusieurs variétés contrastées, mais aussi et surtout de montrer la nécessité de constituer des corpus avec une méthodologie explicite telle que celle proposée par PFC.

Du point de vue du schwa, nous avons montré que le français de Douzens possédait une opposition claire entre finales /C#/ vs /Cə#/ , et que cette opposition avait complètement disparu en Vendée. Le schwa y existe toujours, mais notre corpus montre qu'il relève d'une autre dynamique. Entre ces deux « pôles », l'enquête du Pays Basque témoigne d'un parler en évolution, ce qui se traduit notamment par une grande variation inter-individuelle.

De tels faits, et en particulier ceux du Pays Basque, nécessitent le recours à des corpus, et ne peuvent être établis sur la base de l'écoute et/ou de l'intuition. La méthodologie PFC, bien qu'elle puisse être critiquée sur certains points, constitue selon nous une avancée réelle dans la description de la phonologie du français : si elle offre la possibilité d'exploiter des données massives et globales, à l'échelle du corpus, elle ne néglige pas les données plus fines et permet d'explorer les systèmes individuels tout autant que des questions de détail comme la corrélation graphie/phonie ou l'examen systématique des différentes positions dans le mot.

Bien entendu, les codages PFC ne représentent qu'un premier niveau de « déblayage », et ne sont pas en eux-mêmes une analyse. Ils ne donnent aucune interprétation, et celle-ci demeure à l'entière discrétion du linguiste, en fonction de ses orientations théoriques. Il est donc certaines questions

auxquelles les codages n'apportent pas de réponse directe : par exemple, l'examen des codages ne permet pas de déterminer l'influence de la nature des consonnes sur la chute ou le maintien de la voyelle. En revanche, nous pensons qu'il facilite grandement l'examen de ces questions en permettant de dégager les principaux contextes pertinents : les codages PFC ne donnent pas la réponse à des questions fines telles que l'influence du relâchement (ou non) des consonnes finales dans la simplification des groupes consonantiques (type *minis'de la justice*). En revanche, ils constituent un instrument qui aide le linguiste à y répondre.

Enfin, l'entreprise nous semblerait vaine qui se limiterait à constater les faits de langue. Il est donc de première importance que les phénomènes observés nourrissent la réflexion théorique, car ils doivent permettre de réfuter et d'étayer les hypothèses phonologiques. Loin d'un empirisme clos sur lui-même ou d'une approche hypothétique ayant l'intuition pour seule phénoménologie, le projet PFC espère, par un va-et-vient constant entre expérience et théorie, contribuer à améliorer notre appréhension du français et des objets sonores.

### **Références bibliographiques**

- Angoujard J.-P. (1997). *Théorie de la syllabe. Rythme et qualité*. Paris : Editions CNRS.
- Arnauld A. & Lancelot C. (1660). *Grammaire générale et raisonnée*. Paris : Republications Paulet (1969).
- Berrit-Hansen A. (2003). « Le contexte prépausal – un contexte dynamique pour le schwa dans le français parisien. », in J. Durand, B. Laks & C. Lyche (éds.), pp. 142-144.
- Borrell A. (1975). *Enquête sur la phonologie du français parlé à Toulouse*. Thèse de doctorat 3<sup>ème</sup> cycle. Université de Toulouse-Le Mirail.
- Bourdieu P. (1993). *La misère du monde*. Paris : Seuil.
- Blevins J. (1995). « The syllable in phonological theory », in J. A. Goldsmith (ed.). *The Handbook of Phonological Theory*. Oxford : Blackwell, pp. 206-244.

- Clements G. N. & Keyser S. J. (1983). *CV Phonology. A Generative Theory of the Syllable*. Cambridge, Massachusetts : MIT Press.
- Delais-Roussarie E. & Durand J. (eds.) (2003). *Corpus et variation en phonologie du français*. Toulouse : Presses Universitaires du Mirail.
- Dell F. (1970). *Les règles phonologiques tardives et la morphologie dérivationnelle du français*. PhD diss. MIT, inédit.
- Dell F. (1985). *Les règles et les sons. Introduction à la phonologie générative*. (Première édition : 1973). Paris : Hermann.
- Dell F. (1982). « On gaps undetectable for language learners », in J. Mehler, M. Garrett & E. Walker (eds.). *Perspectives in Mental Representation*. New York : Erlbaum, pp. 431-438.
- Delattre P. (1966). *Studies in French and Comparative Phonetics*. La Haye : Mouton.
- Durand J. (1993). « Sociolinguistic variation and the linguist », in C. Sanders (ed.) (1993). *French Today: Language in its Social Context*. Cambridge : Cambridge University Press, pp. 257-285.
- Durand J. (1996). « Linguistic purification, the French nation-state and the linguist », in C. Hoffman (ed.). *Language, Culture and Communication in the New Europe*. Clevedon : Multilingual Matters, pp. 75-92.
- Durand J. & Laks B. (eds.) (1996). *Current Trends in Phonology : Models and Methods*. Salford : European Studies Research Institute (ESRI).
- Durand J., Laks B. & Lyche C. (2002). « La phonologie du français contemporain : usages, variétés et structure », in C. Pusch & W. Raible (eds.) *Romanistische Korpuslinguistik- Korpora und gesprochene Sprache / Romance Corpus Linguistics - Corpora and Spoken Language*. Tübingen : Gunter Narr Verlag, pp. 93-106.

*Le schwa en français : pourquoi des corpus ?*

- Durand J. et Lyche C. (2003). « Le projet 'Phonologie du Français Contemporain' (PFC) et sa méthodologie », in E. Delais-Roussarie & J. Durand (eds.), pp. 213-276.
- Durand J., Slater C. & Wise H. (1987). « Observations on schwa in Southern French », *Linguistics* 25/2 : 983-1004.
- Durand J. & Tarrier J.-M. (2003). « Enquête phonologique en Languedoc (Douzens, Aude) », *La Tribune Internationale des Langues Vivantes* 33 : 117-127.
- Encrevé P. (1988). *La liaison avec et sans enchaînement. Phonologie tridimensionnelle et usages du français*. Paris : Seuil.
- Eychenne J. (2003). *Prolégomènes à une étude du schwa en français : aspects méthodologiques, empiriques et théoriques*. Mémoire de DEA. Université de Toulouse-Le Mirail.
- Eychenne J. & Hambye Ph. (2004). « La FAQ du codeur schwa », in J. Eychenne & G. Mallet (eds.) *Du segmental au prosodique : protocoles, outils, extensions et travaux en cours*. 29-34. Bulletin PFC n°3. ERSS, Toulouse-Le Mirail.
- Fouché P. (1958) *Phonétique historique du français. Vol. 2 : Les voyelles*. Paris : Klincksieck.
- Fudge E. (1969). « Syllables », *Journal of Linguistics* 5 : 253-286.
- Gadet F. (2000). « Derrière les problèmes méthodologiques du recueil des données », in M. Bilger (ed.) *Linguistique sur corpus : études et réflexions*. Actes du Colloque du 12 mars 1999, Perpignan. *Cahiers de l'Université de Perpignan* 33 : 59-75.
- Greimas A. (1969). *Dictionnaire de l'ancien français jusqu'au milieu du XIV<sup>ème</sup> siècle*. Paris : Larousse.
- Labov W. (1976). *Sociolinguistique*. Paris : Editions de Minuit.
- Lowenstamm J. (1996). « CV as the only syllable type », in J. Durand & B. Laks (eds.), pp. 419-441.

- Lyche C. & Durand J. (1996). « Testing government phonology ou pourquoi le choix du schwa », in J. Durand & B. Laks (eds.), pp. 443-47.
- Malmberg B. (1976<sup>5</sup>). *Phonétique française*. Malmö : Hermods.
- Martinet A. (1969). *Le français sans fard*. Paris : PUF.
- Milroy L. (1987<sup>2</sup>). *Language and Social Networks*. Oxford : Blackwell.
- Morin Y.-C. (1978). « The status of mute *e* », *Studies in French Linguistics* 1/2 : 79-140.
- Morin Y.-C. (1983). « Quelques observations sur la chute du *e* muet dans le français régional de Saint-Etienne », *La Linguistique* 19 : 71-93.
- Morin Y.-C. (1988). « De l'ajustement du schwa en syllabe fermée dans la phonologie du français », in S. P. Verlyuten (ed.). *La phonologie du schwa français*. Amsterdam : John Benjamins, pp. 133-190.
- Morin Y.-C. (2000). « Le français de référence et les normes de prononciation », *Cahiers de l'Institut de Linguistique de Louvain* 21 : 91-135.
- Schane S.A. (1968). *French Phonology and Morphology*. Cambridge, Massachusetts : MIT Press.
- Scheer T. (1999). « Aspects de l'alternance schwa-zéro à la lumière de 'CVCV' », *Recherches linguistiques de Vincennes* 28 : 87-114.
- Scheer T. (2000). « L'immunité de schwa en début de mot », *Langue française* 126 : 113-126.
- Séguy J. (1951). *Le français parlé à Toulouse*. Toulouse : Privat.
- Selkirk E. O. (1972). *The Phrase Phonology of English and French*. PhD Dissertation. MIT. (Publié 1980, Garland Publishing.)
- Selkirk E. O. (1982). « The syllable », in H. van der Hulst & N. Smith (eds.) *The Structure of Phonological Representations (Part II)*. Dordrecht : Foris.

*Le schwa en français : pourquoi des corpus ?*

Vaugelas C.F. de (1647). *Remarques sur la langue française utiles à ceux qui veulent bien parler et écrire*. Facsimile de 1970. Genève : Slatkine reprints.

Walter H. (1977). *La phonologie du français*. Paris : PUF.

Walter H. (1982). *Enquête phonologique et variétés régionales du français*. Paris : PUF.

**Annexe**

Locuteur conservateur du corpus Douzens (viticulteur à la retraite, 77 ans au moment de l'enquête).<sup>12</sup>

```
#####  
### FIN DE POLYSYLLABE ###  
#####  
  
*****  
*  
0411 ----> 43  
*****  
*  
0 : schwa absent  
4 : dernière syllabe de polysyllabe  
1 : voyelle à gauche  
1 : voyelle à droite  
*****  
*  
1 GM: mon arrièrel312-grand-mère0411 enfin euh, nous sommel412s de  
2 h c'est toujours1412s politique0411 il faut direl412 commel412 c'  
3 ctions en arrièrel413,GM: avec0411 euh une0411 amie, et un ami,  
4 rièrel413,GM: avec0411 euh une0411 amie, et un ami, on était tro  
5 municipalité est dell12 gauche0411 ici,GM: ça enregistrel422 là?  
6 GM: c'est essayer dell12 faire0411 unel413, GM: unel412 listel42  
7 1132 l'ai baratiné pendant une0411 heure0411 au moins hein.GM: E  
8 baratiné pendant une0411 heure0411 au moins hein.GM: Et lell12 m  
9 moins hein.GM: Et lell12 maire0411 il mell12 de/, répondait par0  
10 fin,GM: il m'a dit, 'ici, pour0411 ètre1422 sur0412 la liste0421  
11 urtout, c'était pour0412 faire0411 unel412 listel422 d'union, po  
12 nel412 listel422 d'union, pour0411 ètre1423,GM: parce0312 quel42  
13 ez à unel412 sépulturel412 par0411 exemplel424, jell132 sais pas  
14 direl414 euh, 'si on commence0411 à les prendrel423, ils commen  
15 as unel412 maison dell12 verre0411 hein, la mairie, on nell12 sa  
16 on, non mais moi ça m'est égal0411 hein cell12 quell12 jell12 di  
17 ell13, GM: dell132 briser cette0411 euh,GM: cettel412 scission là  
18 est à Douzens del212puis, sept0411 ou huit0411 ans maintenant he  
19 s del212puis, sept0411 ou huit0411 ans maintenant hein, ça fait  
20 31 lui ai dit, il y a soixante0411-et-quinze0411 ans quell12 j'y  
21 il y a soixante0411-et-quinze0411 ans quell12 j'y habitel413' (  
22 enfants euh, Gel232nel312viève0411 et Ma/, ils sont nés tous les  
23 on.GM: <JP: C'était,> ma soeur0411 aussi elle0411 a eu deux goss  
24 tait,> ma soeur0411 aussi elle0411 a eu deux gossel414s, <JP: ou  
25 h <GM: c'est primeur0412 cette0411 année non ? C'est un peu dell  
26 on ? C'est un peu dell12 bonne0411 heure2413 ?> oui, c'est un, c  
27 a ètre1422 ver0412s lell12 dix0411 alor0413s.JP: Oh certainement  
28 a en effet décidé dell12 faire0411 étapel314, dans cettel412 com  
29 blanc sec0413,ses chel212mise0411s en soie, un champion local04  
30 à pied, Louis Garret,quatrième0411 aux jeux olympique1412s dell1  
31 2s qui cou/ tournaient toujours0411s autour0412 des mêmel412s vil  
32 festations qui ont eu tendance0411 à sell12 multiplier lors des  
33 el412s officiellel413s.La côte0411 escarpée du Mont Saint-Pierre  
34 ont Saint-Pierrel414, qui mène0411 au villagel413,connait des ba  
35 2ment dell12 police1413, comme0411 on a vu à Jonquièrel413, et d  
36 vons lell12 soutien du village0411 entier'. Dell132 plus0413, que  
37 t Francel413, et Lell12 Nouvel0411 Observateur0413,indique1312ra
```

12. La présentation des données a dû être adaptée au format de la revue.

J. DURAND, J. EYCHENNE

```
38 s fanatique1411s auraient même0411 entamé un jeûne1412 prolongé
39 uell1412 saint sell12 vouer. I10411 a lell12 sentiment dell12 sel
40 2 impasse1414, stupidel1413. I10411 est en désespoir0412 dell12 c
41 112 cause1413, décidé à écrire0411 au Premier Ministre1423, pour
42 illage1413, était vraiment une0411 étape1412 nécessaire1414, dan
43 urnée prévue. Beaulieu préfère0411 être0421 inconnue et tranqui

*****
1411 ----> 2
*****
1 : schwa present
4 : derniere syllabe de polysyllabe
1 : voyelle a gauche
1 : voyelle a droite
*****

1 cher nous étions dell12 droite1411, et la municipalité est dell1
2 trel423. Quelque1422s fanatique1411s auraient même0411 entamé un

%%% contexte x411 %%%
- schwa est absent dans 95.55% des cas, soit 43 occurrence(s).
- schwa est present dans 4.44% des cas, soit 2 occurrence(s).

*****
0412 ----> 36
*****
0 : schwa absent
4 : derniere syllabe de polysyllabe
1 : voyelle a gauche
2 : consonne a droite
*****

1 cher euh, Joseph.GM: Cent pour0412 cent ma mère1413, mon gran/,
2 s municipale1413s.GM: euh pour0412 nell12 rien vous cacher nous
3 rêter?> non, non, ça m'est égal0412, ça mell12 gêne1412 pas <E: S
4 112 laisser passer.GM: Et alor0412s j'ai, jell134, j'ai rencontré
5 lell12 maire1413.GM: l'actuel0412 d'ailleurs0414s, et jell12 lui
6 présenter tous les trois avec0412 vous.GM: Voilà, prenez-nous,
7 l il mell12 de/, répondait par0412 monosyllable1414s, 'oui, non,
8 t, 'ici, pour0411 être1422 sur0412 la liste0421 il faut être1422
9 421 un hein, vous êtes1412s sûr0412 dell12 perdre1421 hein.GM: Vo
10 voulais surtout, c'était pour0412 faire0411 une1412 liste1422 d
11 hé, moi jell12 suis fâché avec0412 personnel1413, lell12 maire (X
12 à faire1412 là. GM: c'est pour0412 ça que1112 moi jell12 lui euh
13 ù on sel212ra trois au conseil0412 municipal0413.GM: on veut pas
14 quand jell12 vais lell12 voir0412 lell12 maire0414, jell12 vais
15 ai été étonné l'autre1422 jour0412 que1122 quelqu'un mell12 dise
16 i, (peu clair) ah non ma soeur0412 non, jell12 crois non, <JP: (r
17 it que1112 la sécurité sociale0412 n'était pas en déficit0413. (r
18 era ?JP: Oh <GM: c'est primeur0412 cette0411 année non ? C'est u
19 peu près.> ça va être1422 ver0412s lell12 dix0411 alor0413s.JP:
20 l'autre1422 jour0413, GM: ver0412s l'an dernier il y en a un j
21 cettel412 communel1413, au cour0412s dell12 sa tournée, dell12 la
22 fin d'année. Jusqu'ici les seul0412s titrel422s dell12 gloire1412
23 11s en soie, un champion local0412 dell12 course0421 à pied, Lou
24 dell12 Berlin en dix1312-neuf0412 cent trente0412-six0413, et p
25 n dix1312-neuf0412 cent trente0412-six0413, et plus récemment, s
26 r0414d, tout bête1312ment, car0412 lell12 Premier Ministre1424,
27 4, lassé des circuits habituel0412s qui cou/ tournaient toujours
28 tournaient toujours0411s autour0412 des même1412s ville1413s, veu
29 12s ville1413s, veut découvrir0412 cell122 qu'il appelle1412, 'la
30 or0413ds, manifestel422nt leuc0412 colèrel413. D'un autre1422 co
31 32ment prend contac0454t avec0412 la préfecture1412 la plus pro
32 412 que1112 tout est fait pour0412 lell12 protéger.Or0413, un gr
33 dell12 notrel422 liberté. S'il0412 faut monter euh, patte1412 bl
34 uh, patte1412 blanche1412 pour0412 circuler, nous mell12 répondr
35 dell12 sell12 trouver dans une0412 impasse1414, stupidel1413. I10
36 el413. I10411 est en désespoir0412 dell12 cause1413, décidé à éc

*****
1412 ----> 67
*****
1 : schwa present
4 : derniere syllabe de polysyllabe
1 : voyelle a gauche
2 : consonne a droite
*****

1 mère0411 enfin euh, nous somme1412s des,GM: des gens vraiment, d
```

*Le schwa en français : pourquoi des corpus ?*

2 a GM: Il y a euh c'est toujours<sup>1412s</sup> politique<sup>0411</sup> il faut dire<sup>14</sup>  
3 12s politique<sup>0411</sup> il faut dire<sup>1412</sup> commel<sup>1412</sup> c'est hein, il faut  
4 que<sup>0411</sup> il faut dire<sup>1412</sup> commel<sup>1412</sup> c'est hein, il faut pas,GM: s  
5 ell<sup>132</sup> cacher euh GM: les chose<sup>1412s</sup> parce<sup>0312</sup> que<sup>0421</sup> euh,GM: au  
6 m'est égal<sup>0412</sup>, ça mell<sup>112</sup> gêne<sup>1412</sup> pas <E: Si vous voulez que.>G  
7 M: Non non non, ça mell<sup>112</sup> gêne<sup>1412</sup> pas hein.GM: Vous pouvez lell  
8 112 faire<sup>0411</sup> unel<sup>1413</sup>, GM: unel<sup>1412</sup> listel<sup>1422</sup> d'union.GM: Euh, no  
9 x contre<sup>0421</sup> un hein, vous êtel<sup>1412s</sup> sûr<sup>0412</sup> dell<sup>112</sup> perdrel<sup>1421</sup> he  
10 c'était pour<sup>0412</sup> faire<sup>0411</sup> unel<sup>1412</sup> listel<sup>1422</sup> d'union, pour<sup>0411</sup> ê  
11 2 quel<sup>1424</sup>, dans lell<sup>112</sup> villagel<sup>1412</sup> vous savez quand vous allez à  
12 s savez quand vous allez à unel<sup>1412</sup> sépulture<sup>1412</sup> par<sup>0411</sup> exemple  
13 vous allez à unel<sup>1412</sup> sépulture<sup>1412</sup> par<sup>0411</sup> exemple<sup>1424</sup>, jell<sup>132</sup> s  
14 ), GM: comment ça se<sup>1112</sup> passel<sup>1412</sup> chez vous hein,GM: mais (rire  
15 uand <E: Une différence.> mêmel<sup>1412</sup> la scission, il y a la scissi  
16 lement.GM: Il y a rien à faire<sup>1412</sup> là. GM: c'est pour<sup>0412</sup> ça que  
17 parce<sup>0312</sup> qu'ici c'est pas unel<sup>1412</sup> maison dell<sup>112</sup> verre<sup>0411</sup> hein,  
18 où elles (XX).> ça mell<sup>112</sup> gêne<sup>1412</sup> pas.E: Bon on peut parler <GM  
19 peut parler d'autrel<sup>1422</sup> chose<sup>1412</sup> si vous voulez, moi, si vous  
20 briser cette<sup>0411</sup> euh,GM: cettel<sup>1412</sup> scission là voilà. GM: 'rende  
21 4, il faut, il faut vous faire<sup>1412</sup> connaitrel<sup>1424</sup> hein, c'est vra  
22 il m'a demandé, à la pétanque<sup>1412</sup> là il m'a dit, 'il y a longte  
23 (rires) .GM: Et j'y suis même<sup>1412</sup> né, enfin bon.E: Et toi papa  
24 M: Quand mêmel<sup>1413</sup> Oui. GM: Unel<sup>1412</sup> semaine<sup>1413?</sup> GM: [malerus] <J  
25 ent un peu à ça.GM: mais commel<sup>1412</sup> celle<sup>1312</sup>-là, j'en ai jamais  
26 rt d'abordGM: c'est la matièrel<sup>1412</sup> premièrel<sup>1413</sup>. <JP: la matièr  
27 GM: jell<sup>132</sup> suis passé dans unel<sup>1412</sup> vignel<sup>1412</sup> là,GM: euh il mell  
28 suis passé dans unel<sup>1412</sup> vignel<sup>1412</sup> là,GM: euh il mell<sup>112</sup> dit, 'tu  
29 il mell<sup>112</sup> dit, 'tu vois cettel<sup>1412</sup> vigne<sup>2414</sup>, on en fait deux ce  
30 Les deux cents hectos, hectarel<sup>1412s</sup> bon euh c'est bien,GM: mais,  
31 ité c'est (pâle), jell<sup>132</sup> pense<sup>1412</sup> pas <JP: c'est pas la qualité  
32 t-il à Beaulieu?Lell<sup>132</sup> villagel<sup>1412</sup> dell<sup>112</sup> Beaulieu est en grand  
33 aire<sup>0411</sup> étapel<sup>1314</sup>, dans cettel<sup>1412</sup> communal<sup>1413</sup>, au cour<sup>0412s</sup> dell  
34 0412s titrel<sup>1422s</sup> dell<sup>112</sup> gloire<sup>1412</sup> dell<sup>112</sup> Beaulieu, étaient son  
35 atrième<sup>0411</sup> aux jeux olympique<sup>1412s</sup> dell<sup>112</sup> Berlin en dix<sup>1312</sup>-neu  
36 , et plus récemment, son usinel<sup>1412</sup> dell<sup>112</sup> pâtel<sup>1412s</sup> italiennel<sup>141</sup>  
37 ent, son usinel<sup>1412</sup> dell<sup>112</sup> pâtel<sup>1412s</sup> italiennel<sup>1413s</sup>.Qu'est-ce<sup>1412</sup>  
38 1412s italiennel<sup>1413s</sup>.Qu'est-ce<sup>1412</sup> qui donc a valu à Beaulieu ce  
39 ujour<sup>0411s</sup> autour<sup>0412</sup> des mêmel<sup>1412s</sup> ville<sup>1413s</sup>, veut découvrirl<sup>04</sup>  
40 uvrir<sup>0412</sup> cell<sup>1122</sup> qu'il appelle<sup>1412</sup>, 'la campagne<sup>1412</sup> profondel<sup>141</sup>  
41 u'il appelle<sup>1412</sup>, 'la campagne<sup>1412</sup> profondel<sup>1413</sup>.Lell<sup>132</sup> maire<sup>141</sup>  
42 412 profondel<sup>1413</sup>.Lell<sup>132</sup> maire<sup>1412</sup> dell<sup>112</sup> Beaulieu, Marc<sup>0422</sup> Bl  
43 22 Blanc, est en rel<sup>1212</sup>vanche<sup>1412</sup> très inquiet.La cotel<sup>1412</sup> du P  
44 anhel<sup>1412</sup> très inquiet.La cotel<sup>1412</sup> du Premier Ministrel<sup>1422</sup> nell  
45 mier Ministrel<sup>1422</sup> nell<sup>112</sup> cesse<sup>1412</sup> dell<sup>112</sup> baisser del<sup>1212</sup>puis les  
46 112 multiplier lors des visite<sup>1412s</sup> officiel<sup>1413s</sup>.La côte<sup>0411</sup>  
47 illagel<sup>1413</sup>, connaît des barragel<sup>1412s</sup> chaquel<sup>1412</sup> fois quell<sup>112</sup> les  
48 onnait des barragel<sup>1412s</sup> chaquel<sup>1412</sup> fois quell<sup>112</sup> les opposants de  
49 D'un autrel<sup>1422</sup> coté, à chaquel<sup>1412</sup> voyagel<sup>1412</sup> du Premier Ministr  
50 1422 coté, à chaquel<sup>1412</sup> voyagel<sup>1412</sup> du Premier Ministrel<sup>1423</sup>, lell  
51 ac<sup>0454t</sup> avec<sup>0412</sup> la préfecture<sup>1412</sup> la plus proche<sup>1413</sup>, et s'assu  
52 a plus proche<sup>1413</sup>, et s'assure<sup>1412</sup> quell<sup>112</sup> tout est fait pour<sup>041</sup>  
53 t des vérifications d'Un jeune<sup>1412</sup> membrel<sup>1422</sup> dell<sup>112</sup> l'oppositio  
54 'il<sup>0412</sup> faut monter euh, pattel<sup>1412</sup> blanche<sup>1412</sup> pour<sup>0412</sup> circuler  
55 monter euh, pattel<sup>1412</sup> blanche<sup>1412</sup> pour<sup>0412</sup> circuler, nous nell  
56 cle<sup>1422s</sup> parus dans La Dépêchel<sup>1412</sup> du Centrel<sup>1423</sup>, L'Express<sup>0413</sup>,  
57 des activistel<sup>1422s</sup> des communel<sup>1412s</sup> voisinel<sup>1413s</sup>, préparel<sup>1412nt</sup>  
58 unel<sup>1412s</sup> voisinel<sup>1413s</sup>, préparel<sup>1412nt</sup> unel<sup>1412</sup> journée chaudiel<sup>1414</sup>, au Premie  
59 oisinel<sup>1413s</sup>, préparel<sup>1412nt</sup> unel<sup>1412</sup> journée chaudiel<sup>1414</sup>, au Premie  
60 aient même<sup>0411</sup> entamé un jeûnel<sup>1412</sup> prolongé dans l'églisel<sup>1412</sup> de  
61 ûnel<sup>1412</sup> prolongé dans l'églisel<sup>1412</sup> dell<sup>112</sup> Saint Marin/ martinvil  
62 inville<sup>1413</sup>.Lell<sup>132</sup> sympathique<sup>1412</sup> maire<sup>1412</sup> dell<sup>112</sup> Beaulieu, ne  
63 3.Lell<sup>132</sup> sympathique<sup>1412</sup> maire<sup>1412</sup> dell<sup>112</sup> Beaulieu, nell<sup>132</sup> sait  
64 ulieu, nell<sup>132</sup> sait plus à quell<sup>1412</sup> saint se<sup>1112</sup> vouer. Il<sup>0411</sup> a  
65 , était vraiment une<sup>0411</sup> étapel<sup>1412</sup> nécessaire<sup>1414</sup>, dans la journ  
66 r au centrel<sup>1422</sup> d'une bataille<sup>1412</sup> politique<sup>1412</sup> dont, par la té  
67 2 d'une bataille<sup>1412</sup> politique<sup>1412</sup> dont, par la télévision, sel2

\*\*\*\*\*

\*  
2412 ---> 1  
\*\*\*\*\*

\*  
2 : schwa incertain  
4 : dernière syllabe de polysyllabe  
1 : voyelle à gauche  
2 : consonne à droite  
\*\*\*\*\*

\*  
1 e0413,GM: jell132 vais lui dire<sup>2412</sup> 'quand mêmel<sup>1414</sup>, il y a des m

%%% contexte x412 %%%  
- schwa est absent dans 34.61% des cas, soit 36 occurrence(s).  
- schwa est présent dans 64.42% des cas, soit 67 occurrence(s).  
- schwa est incertain dans 0.96% des cas, soit 1 occurrence(s).

J. DURAND, J. EYCHENNE

\*\*\*\*\*  
\* 0413 ---> 17  
\*\*\*\*\*  
\* 0 : schwa absent  
4 : dernière syllabe de polysyllabe  
1 : voyelle à gauche  
3 : frontière intonative forte  
\*\*\*\*\*  
\* 1 gens vraiment, dell32 Douzens0413, voilà. (bruit).E: C'est un v  
2 trois au conseil0412 municipal0413, GM: on veut pas vous  
prendre1  
3 aire0414, jell32 vais lui dire0413,GM: jell32 vais lui dire2412  
parlons 4 ale0412 n'était pas en déficit0413. (rires) GM: Non mais  
5 2 ver0412s dell122 dix0411 alor0413s.JP: Oh certainement et même  
6 a semaine dernière/ <GM: malheur0413, j'y suis allé l'autre1422 jo  
7 lair) GM: Euh l'autre1422 jour0413, GM: ver0412s l'an dernier i  
8 ieu, étaient son vin blanc sec0413,ses che1212mise0411s en soie,  
9 2-neuf0412 cent trente0412-six0413, et plus récemment, son usine  
10 Beaulieu cell12 grand honneur0413.Lell131 hasar0414d, tout bétel  
11 les élections.Comment en plus0413, éviter les manifestations qu  
12 opposants dell12 tous les bor0413ds, manifestel422nt leur0412 c  
13 it pour0412 dell122 protéger.Or0413, un gros déclenche1312ment de  
14 llage0411 entier'. Dell32 plus0413, quelquel422s article1422s pa  
15 el412 du Centre1423, L'Express0413, Oues0452t France1413, et Lel  
16 Lell12 Nouvel0411 Observateur0413, indique1312traient quell12 des  
17 ns des millions de, d'électeur0413s. Le Premier Ministre, ira-t-  
i  
\*\*\*\*\*  
\* 1413 ---> 36  
\*\*\*\*\*  
\* 1 : schwa present  
4 : dernière syllabe de polysyllabe  
1 : voyelle à gauche  
3 : frontière intonative forte  
\*\*\*\*\*  
\* 1 GM: Cent pour0412 cent ma mère1413, mon gran/, ma grand-mère1413  
2 1413, mon gran/, ma grand-mère1413,GM: mon arrière1312-grand-mèr  
3 y a deux élections en arrière1413,GM: avec0411 euh une0411 amie  
4 senté aux élections municipale1413s,GM: euh pour0412 nell122 rien  
5 enregistré1422 là?GM: Ah mince1413. <E: [XX]> Ca fait rien? <E:  
6 4, j'ai rencontré dell122 maire1413,GM: l'actuel0412 d'ailleurs041  
7 t essayer dell122 faire0411 une1413, GM: une1412 listel422 d'unio  
8 il faut être1422 dell122 gauche1413. GM: Et ben, je0111 lui ai di  
9 vant la m/, avant la sépulture1413,GM: avant dell122 rentrer, vou  
10 2 suis fâché avec0412 personnel1413, dell132 maire (XXX). <E: Il y  
11 22 dell122 nez dans les affaire1413s, (XX) (bruit) GM: parce0312  
12 (euh), rien. GM: Il y a aucune1413,GM: alor0414s, [S@], quand je  
13 esponsabilité jell122 veux dire1413.> C'est anonyme. GM: Non mais  
14 nze0411 ans quell12 j'y habite1413' (rires) .GM: Et j'y suis mêm  
15 e, une semai/. <GM: Quand même1413> Oui.GM: Une1412 semaine1413?  
16 el413> Oui.GM: Une1412 semaine1413 ? GM: [malerus] <JP: A peu  
prè  
17 oli la vinifi/GM: Et ça compte1413. <JP: ah oui ça compte> GM: l  
18 c'est la matière1412 première1413. <JP: la matière première> Si  
19 en fait deux cents à l'hectare1413',GM: ça nell122 paraissait un  
20 pel314, dans cettel412 commune1413,au cour0412s dell122 sa tourné  
21 412 dell122 pâtel412s italienne1413s.Qu'est-ce1412 qui donc a val  
22 autour0412 des même1412s ville1413s, veut découvrir0412 cell122 q  
23 412, 'la campagne1412 profonde1413'. Lell32 maire1412 dell12  
Beau  
24 ors des visite1412s officielle1413s. La côte0411 escarpée du  
Mont  
25 el414, qui mène0411 au village1413, connaît des barragel1412s  
chaq  
26 anifestel422nt leur0412 colère1413. D'un autre1422 coté, à chaqu  
27 préfecture1412 la plus proche1413, et s'assure1412 quell122 tout  
28 éclenche1312ment dell122 police1413, comme0411 on a vu à Jonquièr  
29 comme0411 on a vu à Jonquièr1413, et des vérifications d'Un je  
30 'Express0413, Oues0452t France1413, et Lell12 Nouvel0411 Observa  
31 1422s des commune1412s voisine1413s, prépare1412nt une1412 journ  
32 ell122 Saint Marin/ martinville1413. Lell32 sympathique1412  
maire1  
33 s une0412 impasse1414, stupidel1413. Il0411 est en désespoir0412  
34 en désespoir0412 dell122 cause1413, décidé à écrire0411 au Premi  
35 , pour vérifier si son village1413, était vraiment une0411 étape  
36 tre0421 inconnue et tranquille1413, plutôt quell122 dell122 sell12

## Le schwa en français : pourquoi des corpus ?

```
*****
*
* 2413 ----> 1
* *****
*
* 2 : schwa incertain
* 4 : derniere syllabe de polysyllabe
* 1 : voyelle a gauche
* 3 : frontiere intonative forte
* *****
*
* 1 un peu dell12 bonne0411 heure2413 ?> oui, c'est un, c'est un p
*
* %%% contexte x413 %%%
* - schwa est absent dans 31.48% des cas, soit 17 occurrence(s).
* - schwa est present dans 66.66% des cas, soit 36 occurrence(s).
* - schwa est incertain dans 1.85% des cas, soit 1 occurrence(s).
* *****
*
* 0414 ----> 5
* *****
*
* 0 : schwa absent
* 4 : derniere syllabe de polysyllabe
* 1 : voyelle a gauche
* 4 : frontiere intonative faible
* *****
*
* 1 413,GM: l'actuel0412 d'ailleur0414s, et jell12 lui ai dit, 'nous
* 2 GM: il y a aucunel413,GM: alor0414s, [S@], quand jell12 vais lei
* 3 s lell12 voir0412 lell22 maire0414, jell32 vais lui dire0413,GM:
* 4 ça compte> GM: lell32 meilleur0414, <JP: ça compte> mais euh sur
* 5 grand honneur0413.Lell31 hasar0414d, tout bêtel312ment, car0412
* *****
*
* 1414 ----> 14
* *****
*
* 1 : schwa present
* 4 : derniere syllabe de polysyllabe
* 1 : voyelle a gauche
* 4 : frontiere intonative faible
* *****
*
* 1 répondait par0412 monosyllabe1414s, 'oui, non, oui, non', et à
* 2 bien entendu.GM: Parce0312 quel414, à deux contre0421 un hein, v
* 3 pas vous prendrel422 la place1414, hein', GM: mais (euh), ils o
* 4 M: mais (euh), ils ont dû dire1414 euh, 'si on commence0411 à le
* 5 vais lui dire2412 'quand même1414, il y a des mairies,'GM: jell
* 6 quell22 quelqu'un me1112 dise1414, il faut, il faut vous faire1
* 7 aussi elle0411 a eu deux gosse1414s, <JP: oui euh.> ils sont nés
* 8 a c'est sûr> d'autrel422 chose1414, c'est un sujet qui.JP: Non m
* 9 j'y suis allé l'autrel422 jour1414,> JP: oh oui attention hein
* 10 .GM: J'en ai vu hein, des cavel1414s, dans lell12 Bordel322lais e
* 11 escarpée du Mont Saint-Pierre1414, qui mène0411 au villagel413,
* 12 el412nt une1412 journée chaude1414, au Premier Ministrel423.Quel
* 13 2 trouver dans une0412 impasse1414, stupidel413. Il0411 est en d
* 14 t une0411 étapel412 nécessaire1414, dans la journée, dans la tou
* *****
*
* 2414 ----> 2
* *****
*
* 2 : schwa incertain
* 4 : derniere syllabe de polysyllabe
* 1 : voyelle a gauche
* 4 : frontiere intonative faible
* *****
*
* 1 dit, 'tu vois cettel412 vigne2414, on en fait deux cents à l'he
* 2 ll12 paraissait un peu bizarre2414, mais enfin, passons.GM: 'mai
*
* %%% contexte x414 %%%
* - schwa est absent dans 23.80% des cas, soit 5 occurrence(s).
* - schwa est present dans 66.66% des cas, soit 14 occurrence(s).
* - schwa est incertain dans 9.52% des cas, soit 2 occurrence(s).
* *****
*
* 0421 ----> 6
* *****
*

```

J. DURAND, J. EYCHENNE

```

0 : schwa absent
4 : derniere syllabe de polysyllabe
2 :consonne a gauche
1 : voyelle a droite
*****
*
1 : les chose1412s parce0312 que0421 euh,GM: au moment des électio
2 0411 être1422 sur0412 la liste0421 il faut être1422 dell12 gauch
3 rce0312 quel414, à deux contre0421 un hein, vous être1412s sûr041
4 habitez Douzens?', 'mon pauvre0421 ami, je0431 lui ai dit, il y
5 ampion local0412 dell12 course0421 à pied, Louis Garret,quatrièm
6 vue. Beaulieu préfère0411 être0421 inconnue et tranquille1413, p
*****
*
1421 ----> 1
*****
*
1 : schwa present
4 : derniere syllabe de polysyllabe
2 :consonne a gauche
1 : voyelle a droite
*****
*
1 être1412s sûr0412 dell12 perdre1421 hein.GM: Voilà, voilà cell112
%%% contexte x421 %%%
- schwa est absent dans 85.71% des cas, soit 6 occurrence(s).
- schwa est present dans 14.28% des cas, soit 1 occurrence(s).
*****
*
0422 ----> 1
*****
*
0 : schwa absent
4 : derniere syllabe de polysyllabe
2 :consonne a gauche
2 : consonne a droite
*****
*
1 ire1412 dell12 Beaulieu, Marc0422 Blanc, est en rel212vanche14
*****
*
1422 ----> 32
*****
*
1 : schwa present
4 : derniere syllabe de polysyllabe
2 :consonne a gauche
2 : consonne a droite
*****
*
1 toujours.GM: Ca peut paraître1422 oui.GM: Mais ça l'est pas (r
2 uche0411 ici,GM: ça enregistre1422 là?GM: Ah mince1413. <E: (XX)
3 411 unel413, GM: unel412 liste1422 d'union.GM: Euh, nous on veut
4 2 présenter à tout prix contre1422 vous,GM: on voudrait sell12 p
5 1 m'a dit, 'ici, pour0411 être1422 sur0412 la liste0421 il faut
6 0412 la liste0421 il faut être1422 dell12 gauchel413. GM: Et ben
7 senter, et on a chopé la veste1422 bien entendu.GM: Parce0312 qu
8 ur0412 faire0411 unel412 liste1422 d'union, pour0411 être1423,GM
9 3,GM: on veut pas vous prendre1422 la place1414, hein', GM: mais
10 ils commencent12ront à foutre1422 lell12 nez dans les affaire14
11 M: Oui, on peut parler d'autre1422 chose1412 si vous voulez, moi
12 oi, si vous voulez qu'on parle1422 d'autre1422 chose ça m'est ((
13 voulez qu'on parle1422 d'autre1422 chose ça m'est ((SOzame)) éga
14 t euh, j'ai été étonné l'autre1422 jour0412 quel122 quelqu'un me
15 lons <E: Ca c'est sûr> d'autre1422 chose1414, c'est un sujet qui
16 >GM: Vous quand est-ce0312 quel422 vous avez, vous (XX) <JP: Ben
17 ) <JP: A peu près.> ça va être1422 ver0412s lell22 dix0411 alor0
18 eur0413, j'y suis allé l'autre1422 jour1414,> JP: oh oui attenti
19 atière première> Si tu apporte1422s dell12 la salope1312rie, <JP
20 de la saloperie>GM: tu apporte1422s dell12 la salope1312rie. (JP
21 (JP peu clair) GM: Euh l'autre1422 jour0413, GM: ver0412s l'an d
22 .Jusqu'ici les seul0412s titre1422s dell12 gloire1412 dell12 Bea
23 a cote1412 du Premier Ministre1422 nell12 cesse1412 dell12 baiss
24 tous les bor0413ds, manifeste1422nt leur0412 colère1413. D'un a
25 eur0412 colère1413. D'un autre1422 coté, à chaque1412 voyage1412
26 ications d'Un jeune1412 membre1422 dell12 l'opposition aurait dé
27 in, on est jaloux dell12 notre1422 liberté. S'il0412 faut monter
28 ier'. Dell12 plus0413, quelquel422s article1422s parus dans La D
29 plus0413, quelquel422s article1422s parus dans La Dépêche1412 du
30 ltraient quell12 des activistel422s des commune1412s voisine1413
31 u Premier Ministre1423.Quelquel422s fanatique1411s auraient même

```

## Le schwa en français : pourquoi des corpus ?

```
32 e1112 se1112 trouver au centre1422 d'une bataille1412 politique1
%%% contexte x422 %%%
- schwa est absent dans 3.03% des cas, soit 1 occurrence(s).
- schwa est present dans 96.96% des cas, soit 32 occurrence(s).
*****
*
1423 ----> 6
*****
1 : schwa present
4 : derniere syllabe de polysyllabe
2 :consonne a gauche
3 : frontiere intonative forte
*****
*
1 ste1422 d'union, pour0411 etre1423,GM: parce0312 que1424, dans 1
2 on commence0411 à les prendre1423, ils commencent1312ront à foutr
3 voyagel412 du Premier Ministre1423, le1132 gouvernel322ment pren
4 dans La Dépêche1412 du Centre1423, L'Express0413, Oues0452t Fra
5 haudel414, au Premier Ministre1423.Quelquel422s fanatique1411s a
6 écrire0411 au Premier Ministre1423, pour vérifi/, pour vérifier
%%% contexte x423 %%%
- schwa est absent dans 0% des cas, soit 0 occurrence(s).
- schwa est present dans 100% des cas, soit 6 occurrence(s).
*****
*
0424 ----> 1
*****
0 : schwa absent
4 : derniere syllabe de polysyllabe
2 :consonne a gauche
4 : frontiere intonative faible
*****
*
1 pas parler dell112 ça peut-être0424 hein? <E: Si,>E: il y a pas d
*****
*
1424 ----> 6
*****
1 : schwa present
4 : derniere syllabe de polysyllabe
2 :consonne a gauche
4 : frontiere intonative faible
*****
*
1 411 etre1423,GM: parce0312 que1424, dans le1112 villagel412 vous
2 sépulturel412 par0411 exemple1424, je1132 sais pas (rire), GM:
3 faut vous faire1412 connaître1424 hein, c'est vrai que1112 c'es
4 teurs. Le1132 Premier Ministre1424, ira-t-il à Beaulieu?Le1132 v
5 d émoi.Le1132 Premier Ministre1424, a en effet décidé dell112 fai
6 ar0412 le1122 Premier Ministre1424, lassé des circuits habituel0
%%% contexte x424 %%%
- schwa est absent dans 14.28% des cas, soit 1 occurrence(s).
- schwa est present dans 85.71% des cas, soit 6 occurrence(s).
*****
*
0431 ----> 1
*****
0 : schwa absent
4 : derniere syllabe de polysyllabe
3 : debut de groupe intonatif
1 : voyelle a droite
*****
*
1 ens?', 'mon pauvre0421 ami, je0431 lui ai dit, il y a soixante04
%%% contexte x431 %%%
- schwa est absent dans 100% des cas, soit 1 occurrence(s).
- schwa est present dans 0% des cas, soit 0 occurrence(s).
*****
*
1434 ----> 1
*****
*
```

J. DURAND, J. EYCHENNE

```

1 : schwa present
4 : derniere syllabe de polysyllabe
3 : debut de groupe intonatif
4 : frontiere intonative faible
*****
*
  1 a des mairies, 'GM: jell134, je1434, jell132 di/, jell132 del212vra

%%% contexte x434 %%%
- schwa est absent dans 0% des cas, soit 0 occurrence(s).
- schwa est present dans 100% des cas, soit 1 occurrence(s).
*****
*
0452 ----> 1
*****
*
0 : schwa absent
4 : derniere syllabe de polysyllabe
5 : Simplification de groupe consonantique
2 : consonne a droite
*****
*
  1 entre1423, L'Express0413, Oues0452t France1413, et Lell12 Nouvel

%%% contexte x452 %%%
- schwa est absent dans 100% des cas, soit 1 occurrence(s).
- schwa est present dans 0% des cas, soit 0 occurrence(s).
*****
*
0454 ----> 1
*****
*
0 : schwa absent
4 : derniere syllabe de polysyllabe
5 : Simplification de groupe consonantique
4 : frontiere intonative faible
*****
*
  1 gouvernel322ment prend contac0454t avec0412 la prefecture1412 l

%%% contexte x454 %%%
- schwa est absent dans 100% des cas, soit 1 occurrence(s).
- schwa est present dans 0% des cas, soit 0 occurrence(s).

```