

UNE AIDE POUR LA DICTION  
DES VERS FRANÇAIS VERS 1700 :  
NOTE SUR UN PROGRAMME  
DE SYNTHÈSE DE PAROLE

Philippe CARON  
Université de Poitiers, Forell (EA 3816)

INTRODUCTION <sup>1</sup>

Le projet dont on fait état dans cet article est déjà dans une phase d'exécution malgré son extrême complexité. Il s'agit là d'une de ces utopies dont l'avenir dira s'il a porté des fruits et atteint son but. Cela dit, il est des échecs qui ne déshonorent pas la science.

En termes d'épistémologie, on peut en discuter le bien-fondé, puisqu'il s'agit de restituer un objet à jamais perdu sous sa forme originale. En effet, la prononciation haute <sup>2</sup> du français vers 1700 n'offre plus aucun locuteur natif et, par malheur, il n'y a pas eu de transmission orale de cette diction dans des lieux comme l'Opéra ou le Théâtre Français. Il y a donc réelle solution de continuité. À partir de là, on peut adopter plusieurs attitudes :

- le pyrrhonisme total qui aboutit évidemment au renoncement ;
- l'aventurisme qui consiste parfois à se lancer dans une tentative d'exotisme phonétique sans véritable enquête profonde ;
- le travail de fourmi qui récolte, traite, recoupe des données du temps et essaye de les projeter *en action*, tout en maintenant

1. Ce projet est au départ une initiative de Philippe Caron (EA Forell, Université de Poitiers) et Michel Morel (EA Crisco, Université de Caen). Le signataire de ces pages écrit en son nom propre, mais reconnaîtra, en son lieu, la part vitale de son partenaire Michel Morel dans la rédaction de cet article.

2. Voir Caron (2008), pour un exposé plus approfondi et *infra*, pour une définition cavalière de cette notion englobante.

sans cesse un environnement scientifique qui en présente les limites.

Nous avons choisi quant à nous la troisième voie, estimant qu'aboutir à un portrait-robot de cette diction, c'est-à-dire à une modélisation de ses grands traits, apporterait pour le moins quelque éclaircissement et au plus les ingrédients d'une restitution plausible.

Avant d'entrer plus avant en propos, il est nécessaire de dire ici que nous ne sommes pas des puristes forcenés de la conservation. Cette position est d'autant plus intenable qu'à cette époque les artistes (anachronisme ici assumé) n'ont pas une vision figée de leur art. En musique, Bach transcrit pour le clavier des concertos de Marchello ou de Vivaldi ; une voix manquante dans un groupe vocal est remplacée par une viole de gambe par exemple. En somme, les timbres changent selon les besoins et l'objet n'est pas encore inscrit dans un marbre intangible. C'est donc avec cette position souple – et, nous semble-t-il, intelligente – que la présente tentative se fraye un chemin vers un objet perdu.

### 1. QUEL OBJET ?

Dans le brouhaha des idiolectes perdus, nous sommes vers 1700 obligés de nous rendre à l'évidence : certains sociolectes populaires et bourgeois, certaines pratiques régionales de l'idiome ne sont plus accessibles ; essentiellement parce qu'ils ne méritent pas, en leur temps, l'attention du public et, qui plus est, ne remontent à la surface qu'à l'occasion d'une remarque corrective qui ne fait que stigmatiser un aspect du phonétisme ou de la prosodie.

En revanche, vers cette fin du XVII<sup>e</sup> siècle, un climat normatif est à l'œuvre, favorisant l'éclosion d'une belle floraison d'ouvrages susceptibles d'éclairer tout ou partie de ce qu'on peut appeler la diction haute. Cette diction n'est pas une : il s'agit à vrai dire d'un feuilletage d'habitus qui va de la parole oratoire, en passant par la lecture publique, la diction théâtrale jusqu'au chant, lequel présente évidemment certaines spécificités dues à l'émission de la voix chantée et aux codes musicaux spécifiques. Ces habitus nous intéressent, parce qu'ils concernent encore aujourd'hui les artistes lyriques et les acteurs qui voudraient tenter une restitution de l'objet sonore. Nous sommes ici autant dans l'histoire d'un passé perdu que dans l'esthétique d'un objet encore présent au théâtre ou à l'opéra.

Cet objet, nous pouvons le décomposer, d'une façon un peu arbitraire, en plusieurs strates :

- La phonologie du temps, c'est-à-dire l'ensemble des articulations effectivement distinctives dans ce français classique précis.
- La phonétique normative du temps, c'est-à-dire la chair de la réalisation segmentale moyenne de ces oppositions, au-delà des particularités idiolectales. Dans ce domaine, on peut s'interroger, par exemple, sur la place de l'articulation des /o/ à l'arrière de la bouche : ces timbres étaient-ils plus postérieurs qu'aujourd'hui ? Il existe bien des indices, mais cela vaut-il la peine ?
- La prosodie, qui est probablement la partie la plus problématique parce qu'elle échappe à la vigilance des contemporains, qui n'éprouvent pas le besoin d'en parler directement en tant que telle. C'est donc par des détours laborieux et risqués que nous pouvons néanmoins accéder aux linéaments prosodiques de cette diction<sup>3</sup>.

## 2. À QUELLE ÉPOQUE ?

Nous avons drastiquement limité notre propos au règne personnel de Louis XIV, considérant que les données récoltées autour de 1700 ne pouvaient guère valoir en amont, au-delà de 1660, et en aval, au-delà de 1730 environ. En amont, en effet, des esthétiques contradictoires de la diction versifiée perturbent certainement le modèle que nous essayons de dessiner. En aval, des réorganisations phonologiques se font déjà sentir, qui aboutissent à modifier notamment les voyelles de cavité moyenne dans leur timbre et leur longueur. Une durée de soixante ans environ est déjà beaucoup...

## 3. À LA RECHERCHE DE L'EXHAUSTIVITÉ

Une telle entreprise n'a pas de valeur si elle ne parvient pas à une certaine exhaustivité. C'est bien là la difficulté, et pour l'affronter nous avons eu recours à un outil de torture : la synthèse de parole.

La proposition paraît paradoxale et pourtant elle est vraie : viser une synthèse de parole est le meilleur moyen de laisser une trace de ce travail de recherche, une trace mobilisable à tout moment par des usagers qui souhaitent entendre une base phonético-prosodique de ce qu'ils ont à prononcer. Mais c'est aussi à tout moment un excellent moyen de savoir ce qui manque encore au puzzle ! Nous

3. Voir Caron & Morel (2007) pour une hypothèse forte concernant le rôle du récitatif lullien dans la reconstitution probable de la prosodie du vers grave. Voir aussi Caron (2008b) sur un distique du *Saint-Genest* de Rotrou.

avons donc une double motivation pour utiliser la synthèse de parole : une motivation heuristique et une motivation de vulgarisation du savoir. À terme, cette synthèse de parole sera implantée, par exemple, sur un site comme [www.musicanet.org](http://www.musicanet.org), qui permet aux chanteurs du monde entier de télécharger du matériel de chant choral. Un logiciel de synthèse de parole, qui permettrait d'entendre un texte de Cambert ou Lully oralisé lentement, apporterait une ressource appréciable aux chefs de chœur ou aux chanteurs solistes, et, pourquoi pas, à l'acteur qui viserait une prononciation restituée comme un objet esthétique perdu.

#### 4. LES SOURCES

Elles proviennent d'une multiplicité d'ouvrages : grammaires publiées en France ou à l'étranger, manuels de civilité linguistique, dictionnaires de langue et de rimes, remarques de langue, manuels d'orthoépique, discours savants prononcés dans les Académies du Royaume ou de l'étranger, manuels de français langue étrangère, tentatives de modernisation phonétisante de l'orthographe. Toutes offrent néanmoins des difficultés fort nombreuses, car les données, faut-il le dire, ne nous sont pas destinées et nécessitent toujours une contextualisation délicate : ainsi, un dictionnaire de rimes est destiné à des versificateurs, et non à une approche phonétique des timbres vocaliques. Certaines rimes peuvent être proscrites, même si elles sont homophones<sup>4</sup>. Un réformateur graphique vise l'optimisation de l'écriture en son temps, il ne cherche pas à nous renseigner sur le phonétisme de son époque. S'il ne met qu'un signe graphique pour /e/ et /ɛ/<sup>5</sup>, il n'en reste pas moins que ces timbres sont phonologiquement pertinents. Et lorsqu'un dictionnaire de prononciation de l'anglais, tel celui de Walker (1797), donne des timbres français comme approchant ceux des voyelles anglaises<sup>6</sup>, il faut manipuler ces équivalences avec la plus grande précaution. La règle est alors celle du doute méthodique, du recoupement des données, pour s'assurer qu'un passage équivoque tiré de l'une est

4. Par exemple les terminaisons en *-é* et les infinitifs en *-er* ne riment pas, même si nous avons des attestations nettes selon lesquelles elles sont homophones à la pause : « Cinquièmement, Que quelque muette que soit l'*r* des infinitifs en *er* placez au bout du Vers, il n'est jamais permis de les faire rimer avec les Noms ou les Participes qui ne portent point d'*r*, étant terminez par un *e* fermé : Comme congé / songer, vangé / changer, armez / charmer, assez / passer » (Mourgues 1724 : 72 ; v. aussi Hindret 1698 : II, 3, 425-427).

5. Pour la Renaissance, c'est le cas de Ramus qui dans sa *Gramere* de 1562 n'a que la lettre *e* pour rendre les deux voyelles d'avant de cavité moyenne /e/ et /ɛ/.

6. Dans le principe 559 de son *Critical Pronouncing Dictionary* de 1797.

partiellement discriminé par le recours à une autre et confirmé par une troisième ; en espérant qu'elles ne se copient pas les unes les autres...

Dans ce domaine, comme toujours en Histoire, *testis unus testis nullus*<sup>7</sup>.

##### 5. DEUX EXEMPLES DE COMPLEXITÉ

Lorsqu'une donnée a un domaine de pertinence précis, il arrive qu'on veuille la faire sortir de ce champ, avec le risque d'une erreur. Ainsi Bacilly (1679), professeur de chant, écrit-il à propos de l'articulation de la voyelle nasale /ã/ en musique vocale que toute la note doit être tenue sur le timbre oral [a] avant d'articuler un segment consonantique en fin d'émission, [m] ou [n]. S'appuyant sur cette spécificité de la diction vocale, Green en a déplacé la zone de pertinence vers le domaine de la diction orale. Ainsi les acteurs du *Bourgeois gentilhomme*<sup>8</sup> ont-ils des articulations méridionales pour les voyelles nasales d'avant qui, pour participer d'un exotisme sans doute recherché, n'en sont pas moins erronées. En effet, ce déplacement entre en contradiction avec un des développements les plus solides, les plus argumentés de cette période en matière de description de l'oral : celui du discours de l'abbé de Dangeau à l'Académie française sur les voyelles du français (Dangeau 1754 : 16-17). Ce discours argumente très fortement et très longuement en faveur de la monophthongaison parfaite des voyelles nasales, notamment en montrant que tous les ornements que la voix peut faire sur des voyelles orales (trilles, ports de voix) sont également réalisés sur les voyelles nasales.

Cet exemple nous semble représentatif des erreurs qu'un maniement peu soigneux des données du temps peut occasionner. Un autre exemple de lecture controversée nous est également offert par E. Green (2001 : 89) qui essaye de comprendre un passage de J. Palsgrave :

Quand on lit chez Palsgrave que le *e* suivi d'une consonne nasale dans les mots *embler*, *amendrir*, et *humblement* se prononce « comme un *a* italien et quelque chose dans le nez », et qu'on constate chez Bacilly, un siècle et demi plus tard, qu'un chanteur qui articule la dernière syllabe de *charmante*, ou *seulement* doit faire toute l'émission sur la voyelle *a*, et qu'il ne doit « prononcer l'*n* que lors [il] est prest de finir », en précisant toutefois que

7. « Un seul témoin, pas de témoin. »

8. Molière, *Le Bourgeois gentilhomme*, mise en scène Benjamin Lazar, DVD EDV 1857.

« quand je dis qu'il ne faut prononcer l'n qu'à la fin [...] je ne veux pas dire frapper, mais seulement effleurer », en conclure qu'aux XVI<sup>e</sup> et XVII<sup>e</sup> siècles on prononçait la voyelle [ã] (*a* nasal) en deux temps (*a* ouvert oral, suivi de nasalisation), comme on entend encore dans le sud de la France, est peut-être une preuve d'ignorance présomptueuse et de mystification, mais je continue à croire, naïvement, que c'est une manifestation de bon sens.

Or voici ce que dit le texte de Palsgrave :

If m or n folowe nexte after e, all in one syllable, than e shall be sounded lyke an Italian a, and some thyng in the nose, so that for these wordes thus written, *embler*, *amendrir*, *endementiers*, *humblement*, and suche lyke: in redyng and spekyng they sounde *ambler*, *amendrir*, *andementiers*, *humblemant*. (Palsgrave 1530, f. i v<sup>o</sup> et f. ij r<sup>o</sup>)<sup>9</sup>

Palsgrave n'est pas du tout univoque. Il veut indiquer avant tout que le timbre nasal des mots en *-en-* ou *-em-* a fusionné avec celui des mots en *-am-* ou *-an-*. Il n'indique pas par là qu'il y a dans cette diction un segment vocalique oral [a] puis « quelque chose dans le nez » qui serait une consonne résiduelle subséquente. Quoi qu'il en soit, E. Green ne saurait conclure de la diction chantée (qui effectivement pratique cette méthode d'après Bacilly) sur la diction parlée.

#### 6. DE LA MODÉLISATION SAVANTE VERS LA SYNTHÈSE DE PAROLE<sup>10</sup>

Une fois définies les règles de prononciation et de prosodie, avec une part intuitive inévitable que l'on tentera de réduire par la suite en approfondissant la recherche, il est possible à un locuteur averti et entraîné de simuler la lecture de textes de l'époque choisie.

L'intérêt de cette simulation, de cette reconstitution de corpus oral, est double :

1. proposer une diction plausible pour le chant et le théâtre, à destination des amateurs et des professionnels ;
2. présenter un objet concret, à la portée du public, illustrant l'élocution de cette époque.

Mais comme les textes disponibles représentent une quantité de corpus importante (plusieurs milliers d'heures de lecture), il est

9. La graphie en gothique est rendue par le romain et le romain par de l'italique.

10. Pour cette partie et les sections suivantes, je cite presque littéralement avec sa permission la participation de Michel Morel à notre article commun (Caron & Morel 2010), les choses n'ayant pas substantiellement changé depuis.

exclu d'enregistrer la lecture humaine sans une part significative de ceux-ci. La synthèse vocale est alors un moyen d'étendre cette simulation à l'ensemble des textes, par la création d'un modèle à partir de morceaux choisis (250 000 mots) et de seulement 30 minutes de lecture d'un corpus fabriqué pour la circonstance. Bien qu'imparfait, le résultat doit être suffisamment performant pour fournir une prononciation correcte de toutes les phrases et reproduire pour l'essentiel l'intonation du locuteur en respectant au maximum le rythme, le naturel et l'expressivité.

Pour modéliser une langue parlée – actuelle ou ancienne –, nous devons d'abord sélectionner un corpus représentatif, puis l'oraliser ; le corpus oral étant la base de la synthèse vocale. En pratique, nous devons disposer des ressources suivantes :

- un corpus écrit, numérisé en format texte, d'environ 250 000 mots, constitué d'extraits représentatifs des productions écrites de l'époque concernée ;
- un dictionnaire grammatical de la plupart des mots du corpus écrit ;
- un dictionnaire de prononciation de ces mêmes mots ;
- un corpus oral d'environ 30 minutes, enregistré par un locuteur simulant la diction de l'époque sur des extraits sélectionnés pour la variété des prononciations d'une part et celle des unités prosodiques d'autre part.

Ces ressources de base sont à fabriquer, à partir des corpus dont nous disposons, de nos connaissances et de nos hypothèses. Il s'agit plus d'un travail de recherche que d'un travail de développement. Nous avons décidé de nous appuyer pour l'oralisation du vers (qui constitue notre objectif numéro un) sur un corpus de passages les plus nettement récitatifs de deux tragédies lyriques de Lully : *Atys* et *Armide*, dont les livrets sont dus à la collaboration de Philippe Quinault.

## 7. LA SYNTHÈSE DE PAROLE

Le synthétiseur de parole est un programme – un moteur en quelque sorte – dont le fonctionnement ne dépend pas de la langue à synthétiser. Pour chaque langue, il utilise des ressources linguistiques et phonétiques spécifiques sous une forme qu'il est capable d'interpréter et de généraliser. Le travail de développement consiste à fabriquer ces ressources modélisées à partir des ressources de base, soit :

- un dictionnaire pragmatique (probabilité de mise en relief),

- un dictionnaire des catégories grammaticales,
- un dictionnaire des lemmes verbaux,
- un dictionnaire des flexions verbales,
- un jeu de règles de désambiguïsation des catégories grammaticales,
- un jeu de règles de groupement et de mise en relation des unités prosodiques,
- un jeu de règles de transcription graphème-phonème,
- une base de diphtongues,
- un dictionnaire de contours prosodiques.

Une fois ces ressources modélisées intégrées au logiciel ©KALI, celui-ci est capable de prononcer toute phrase du corpus complet. La lecture du corpus écrit par la synthèse permet de cerner les erreurs, les omissions, les défauts et de les corriger en amont. La synthèse de la parole permet aussi, par sa souplesse, de faire avancer la recherche en stimulant une démarche fondamentale et expérimentale riche (essais, comparaisons, possibilité d'options...). Le logiciel ©KALI est utilisé dans l'industrie, les applications multimédia, les serveurs vocaux, le handicap de la vue ou de la parole. Il reçoit un texte en entrée et fournit de la parole en sortie (Morel 2001). Celle-ci doit être intelligible et, dans la mesure du possible, agréable, naturelle et expressive. Pour la restitution d'une langue ancienne, la parole doit être aussi fidèle que possible au modèle, notamment sur les plans de l'articulation et du rythme.

Le texte à prononcer traverse successivement les modules suivants :

- L'analyse syntactico-prosodique fournit un découpage du texte en constituants prosodiques ainsi que leurs relations de dépendance. Pour les besoins de la prosodie, l'analyse syntactico-prosodique fusionne, en un même constituant prosodique, les syntagmes élémentaires ou tronçons qui sont fortement liés entre eux ; par exemple, pronom personnel + verbe ou *être* ou *avoir* + groupe nominal. L'analyse pragmatique du texte (c'est-à-dire l'organisation informationnelle et l'expressivité) complète le modèle prosodique, en repérant les constituants prosodiques à mettre en relief, ainsi que leur modalité affective ; améliorant ainsi le naturel et facilitant l'accès au sens.
- Pour prononcer correctement un texte, il faut ensuite appliquer de nombreuses règles de prononciation. Il suffit, pour s'en faire une idée, de comparer les prononciations des mots : « *choco-*

lat » [ʃɔkɔla] vs « chorégraphe » [kɔʁegʁaf], « les options » [lezɔpsjɔ̃] vs « nous options » [nuzɔptjɔ̃], « aliment » [alimã] vs « animent » [anim(ə)], « écarter » [ekɑʁte] vs « carter » [kɑʁtɛʁ]... Le module de transcription est chargé de fournir un texte phonétique à partir du texte alphabétique (Morel 1998). Si le français comporte environ mille règles, l'anglais en comporte plusieurs milliers. De plus, les mots d'emprunt et les noms propres nécessitent plusieurs milliers de règles supplémentaires.

- À partir de la voix d'un locuteur, on extrait une base de diphtongues, c'est-à-dire 1 000 à 2 000 segments de signal, qui seront ensuite concaténés pour former le signal acoustique de parole. Chaque segment va de la moitié d'un phonème à la moitié du suivant. L'interface entre les segments est le phonème ; ce qui limite le nombre de segments nécessaires à la fabrication de n'importe quel énoncé.
- Les diphtongues étant prélevés dans des contextes différents, il est nécessaire de les modifier pour que leur raccordement ne présente pas de discontinuités. Entrent alors en jeu des traitements complexes à base de transformées de Fourier et de fenêtres de Hanning<sup>11</sup>.
- Le rôle de la prosodie est de donner à la lecture un rythme et une intonation naturels, agréables, expressifs, et de faciliter l'accès au sens, en découpant l'énoncé en groupes de mots fortement liés entre eux, plus ou moins mis en relief. Pour l'heure, la prosodie de ©KALI a été schématisée en traits principaux : déclinaison, pauses, allongements, accents. Un modèle utilisant un dictionnaire de contours prosodiques naturels est en cours de réalisation. En partant du principe que la forme de chaque patron prosodique émerge de la combinaison de contraintes multiniveaux (rythmiques, syntaxiques, pragmatiques) (Lacheret 2002), la méthode utilisée consiste à associer une clé à chaque contour entré dans le dictionnaire, c'est-à-dire une représentation mathématique de ces contraintes. Le principe général consiste à coder des domaines (de taille et de nature linguistique variable, du groupe accentuel au paragraphe), des distributions (initiale, médiane, finale du domaine) et des fonctions (modale, stylistique) associées à ces domaines. Il suffit qu'un

11. La transformée de Fourier est un outil mathématique permettant de décrire un signal par son spectre de fréquence (Brucher 2007). Le fenêtrage (comme celui de Hanning, le plus utilisé) est une technique servant à découper une section des données à mesurer, afin de minimiser ou pondérer les distorsions (Phénomènes de Gibbs) provoquant une fuite spectrale de la transformée de Fourier (Furon s.d.).

des paramètres change pour qu'une autre clé soit produite. La phase de génération, quant à elle, consiste à calculer, sur la base de l'analyse automatique du texte (v. *supra*), la clé associée à chaque segment, puis à la comparer aux clés en mémoire, grâce à une fonction de coût (distance entre les deux clés). Cette méthode permettra ainsi de sélectionner dans le dictionnaire la forme prosodique la plus pertinente et de l'appliquer à l'unité en cours de traitement. Le module générateur de parole convertit la chaîne phonétique en un signal de parole par concaténation des diphtongues de la base. Les contours prosodiques sélectionnés sont concaténés et les valeurs acoustiques sont interpolées sur toute la phrase, puis appliquées en temps réel. Ainsi enrichi, le signal de parole est envoyé en continu vers les haut-parleurs à fréquence d'échantillonnage constante.

#### 8. IMPLÉMENTATION DU PROGRAMME SOUS ©KALI

Le développement de ce projet, fondé sur la synthèse vocale ©KALI, implique la réalisation des points suivants <sup>12</sup> :

- Collecte et numérisation d'un corpus écrit, aussi varié et exhaustif que possible.
- Définition d'un jeu de phones (réalisations réelles des phonèmes) du français aux environs de 1700, à partir des connaissances acquises.
- Fabrication d'un jeu de règles de transcription graphème-phonème (environ 1 500 règles), en mode « déclamation » (réalisation des schwas, des liaisons et de la synérèse selon les règles en vigueur).
- Phonétisation automatique d'un corpus écrit aussi large et varié que possible (« grand corpus »).
- Fabrication d'une liste de phrases à partir de ce corpus, par sélection des parties les plus productives en termes de diphtongues différents. Des phrases complémentaires seront fabriquées, afin de réaliser des diphtongues rares risquant d'être rencontrés dans des textes non encore numérisés <sup>13</sup>.
- Enregistrement d'un corpus oral à partir de ces phrases, pour le prélèvement des diphtongues. L'oralisation sera effectuée avec le souci d'une diction phonétique soignée, prioritairement à la prosodie.

12. Les trois premiers points sont déjà réalisés.

13. Il va de soi que ces 150 à 200 phrases hors contexte ne constituent qu'un outil de travail, en aucun cas un corpus de démonstration.

- Prélèvement des diphtongues du corpus oral (environ 1 000 à 1 500) et fabrication de la voix.
- Construction de dictionnaires et de règles pour l'analyse syntaxique et pragmatique. Toutes les variantes orthographiques pouvant être rencontrées dans les textes devront être prévues ; ce qui nécessitera un prétraitement.
- Construction d'un petit corpus écrit pour le prélèvement de contours prosodiques, par sélection des parties les plus productives du grand corpus en termes de variabilité syntaxique et rythmique.
- Oralisation de ce petit corpus (« corpus prosodie »), avec le souci d'une prosodie soignée, prioritairement à la diction phonétique.
- Prélèvement des contours prosodiques de plusieurs centaines d'unités prosodiques du « corpus prosodie ».
- Intégration dans ©KALI d'une voix en français classique.

#### CONCLUSION : AD AUGUSTA PER ANGUSTA

Le développement des ressources linguistiques soulève de nombreux problèmes en lien avec la recherche fondamentale. La nécessaire exhaustivité du travail de développement pousse le chercheur à affiner ses hypothèses et parfois à prendre parti sur des points qui sont en discussion. C'est le cas, notamment, pour l'allongement et la quantité vocalique, en relation avec l'accent lexical, ainsi que pour l'oralisation des consonnes finales à la pause, qui constituent pour nous deux des domaines les plus cruciaux (Caron 2010).

Ce programme ambitieux (trop ambitieux ?) est pourtant assez bien circonscrit dans ses prétentions (habitus, période) et assez documenté pour qu'on puisse parvenir à une certaine exhaustivité ; condition nécessaire pour une synthèse de parole. Sous réserve de ne vouloir restituer qu'une base phonético-prosodique, une « image robot » (pour prendre une métaphore tirée de la police judiciaire), privée de la chair que représentent les habitudes articulatoires du temps (hélas sans doute perdues pour la chair acoustique) et des effets dramatiques locaux, la réalisation pourra prétendre à une certaine pertinence.

#### RÉFÉRENCES

DE BACILLY Bénigne, 1679, *L'Art de bien chanter de M. de Bacilly. Augmenté d'un discours qui sert de réponse à la critique de ce traité et d'une plus ample instruction pour ceux qui aspi-*

- rent à la perfection de cet Art*, Paris, Chez l'auteur / Claude Bageart, 3<sup>e</sup> éd., rééd. Genève, Minkoff, 1971.
- BRUCHER Matthieu, 2007, *La Transformée de Fourier rapide et d'une de ses dérivées* (en ligne).
- CARON Philippe, 2008a, « Pouvons-nous reconstituer la diction haute du français vers 1700 ? À propos du *Bourgeois gentilhomme* en DVD », *17th Century French Studies*, 30-2, p. 182-195.
- CARON Philippe, 2008b, « La diction des vers à propos d'un distique de Rotrou. Introduction à la reconstitution de la diction haute au XVII<sup>e</sup> siècle », dans Bénédicte Louvat-Molozay et Dominique Moncond'huy (éds), *Plus sur Rotrou : Écrire le théâtre*, Neuilly, Atlante, collection « Clef/Concours », p. 37-56.
- CARON Philippe 2010, « La diction composite de la comédie. À propos du *Bourgeois gentilhomme* », *Annales de l'Association pour un Centre de Recherche sur les Arts du Spectacle aux XVII<sup>e</sup> et XVIII<sup>e</sup> siècles*, 4 (Actes du colloque Restitution et création dans la remise en spectacle des œuvres des XVII<sup>e</sup> et XVIII<sup>e</sup> siècles – Versailles et Nantes, 29-31.05.2008), p. 185-194.
- CARON Philippe et MOREL Michel, 2007, « Note récapitulative sur le programme de reconstitution de la diction haute du français vers 1700 », *Annales de l'Association pour un Centre de Recherche sur les Arts du Spectacle aux XVII<sup>e</sup> et XVIII<sup>e</sup> siècles*, 1, p. 17-24.
- CARON Philippe et MOREL Michel, 2010, « Vers une synthèse vocale du français classique à l'aide du logiciel ©kali », CORELA - Numéros thématiques : Parole (en ligne).
- DE COURCILLON DE DANGEAU Louis (abbé), 1700, *Essais de Grammaire contenus en trois discours*, Paris, Coignard (réédité dans *Opuscules sur la langue française par divers académiciens*, Paris, Brunet, 1754 et Genève, Slatkine, 1969).
- FURON Patrick, s.d., « Rôles des fenêtres de pondération », *Initiation au traitement numérique du signal*, ENSI Caen (en ligne).
- HINDRET Jean, 1687, *L'Art de bien prononcer et de bien parler la langue française*, Paris, Laurent d'Houry, réédition Genève, Slatkine, 1973.
- HINDRET Jean, 1696, *L'Art de prononcer parfaitement la langue française*, Paris, Laurent d'Houry, réédité par la Bibliothèque de l'Association pour un Centre de Recherche sur les Arts du Spectacle aux XVII<sup>e</sup> et XVIII<sup>e</sup> siècles, 2007.
- LACHERET-DUJOUR Anne, 2002, *Modélisation prosodique du français parlé : analyse de la substance, représentation formelle, in-*

- terprétation symbolique*, HDR de l'Université Paris X Nanterre.
- LACHERET-DUJOUR Anne et BEAUGENDRE Frédéric, 1999, *La Prosodie du français*, Paris, CNRS Éditions.
- MOREL Michel et LACHERET-DUJOUR Anne, 1998, « Utilisation d'une structure arborescente pour une hiérarchisation fine des règles de transcription graphème-phonème », *Actes des XXII<sup>es</sup> Journées d'Études sur la Parole (15-19.06.1998) à Martigny (Suisse)*, p. 151-154.
- MOREL Michel et LACHERET-DUJOUR Anne, 2001, « Le logiciel de synthèse vocale ©kali : de la conception à la mise en œuvre », *Traitement Automatique des Langues*, 42, p. 193-221.
- MOURGUES Michel, 1724, *Traité de la poésie française*, Paris, Jacques Vincent, 2<sup>e</sup> éd.
- PALSGRAVE John, 1530, *L'Esclarcissement de la langue françoise* [Londres].
- RÉGNIER-DESMARAIS François Seraphin (abbé), 1706, *Traité de la grammaire française*, Paris, Coignard.
- TALLEMANT LE JEUNE Paul, 1698, *Remarques et décisions de l'Académie française, recueillies par M. L. T.*, Paris, Coignard ; réédition Genève, Slatkine, 1972.
- THOULIER D'OLIVET Pierre-Joseph (abbé), 1738, « Prosodie française » dans *Remarques sur la langue française*, Paris, Barbou ; réédition de l'éd. de 1771 : Genève, Slatkine, 1968.
- VAUDELIN Gilles, 1713, *Nouvelle maniere d'ecrire comme on parle en France*, Paris, Vve Jean Cot et J. B. Lamesle.
- VAUDELIN Gile, 1715, *Instructions cretiennes mises en ortografe naturelle, pour faciliter au peuple la lecture de la sience du salut*, Paris, J. B. Lamesle.
- WALKER John 1797, *A Critical Pronouncing Dictionary* [...], London, Robinson, Cadell & Davies.