

---

# Accents étrangers et régionaux en français

## Caractérisation et identification

**Philippe Boula de Mareuil — Bianca Vieru-Dimulescu — Cécile Woehrling — Martine Adda-Decker**

LIMSI-CNRS

BP 133, F-91403 Orsay CEDEX

{mareuil,bianca,woehrlin,madda}@limsi.fr

---

*RÉSUMÉ.* Cette recherche est dédiée à la caractérisation et à l'identification d'accents en français. À la fois pour des accents étrangers et régionaux, nous sommes partis d'expériences d'identification perceptive, nous avons mesuré des traits phonétiques qui peuvent caractériser ces accents en utilisant l'alignement automatique en phonèmes, et nous avons hiérarchisé les traits les plus discriminants en utilisant des techniques de classification. Les indices suivants sont perceptibles, et se montrent relativement robustes pour l'identification automatique : dévoisement des occlusives sourdes, mouvement du /e/ vers [i], confusions /b/~v/ et /s/~z/, « r roulé », antériorisation ou fermeture du schwa pour les accents anglais, allemand, arabe, espagnol, italien et portugais en français ; antériorisation du /O/ dans la moitié nord de la France, maintien du schwa et dénasalisation des voyelles nasales dans le sud de la France, pour les accents régionaux du français.

*ABSTRACT.* This research focuses on the identification and characterisation of accents in French. For both foreign and regional accents, we started with perceptual identification experiments, we measured phonetic features which may characterise these accents using automatic phoneme alignment, and we ranked the most discriminating features by using classification techniques. The following features are perceivable and happen to be quite robust for automatic identification: devoicing of voiced stop consonants, shift of /e/ toward [i], /b/~v/ and /s/~z/ confusions, “rolled r”, schwa fronting or raising for German, English, Arabic, Spanish, Italian and Portuguese accents in French; /O/ fronting in the North of France, schwa pronunciation and denasalised nasal vowels in the South, for regional French accents.

*MOTS-CLÉS :* accents, identification perceptive et automatique, alignement en phonèmes, classification, analyse de données.

*KEYWORDS :* accents, perceptual and automatic identification, phoneme alignment, classification, data mining.

---

## 1. Introduction

Ce travail vise à la caractérisation et à l'identification d'accents étrangers et régionaux en français, à travers des indices phonétiques que l'on est capable de mesurer automatiquement. Entre accents étrangers et régionaux, l'histoire des sujets parlants n'est certes pas la même (Flege *et al.*, 1995 ; Léon et Léon, 1997). Mais, quelle qu'en soit la cause, quelle que soit l'importance relative de la langue d'origine et du substrat régional, les questions de diagnostic sont en partie les mêmes pour les accents étrangers et régionaux. Combien d'accents peut-on reconnaître, lesquels et pourquoi ? Autrement dit, avec quel degré de granularité peut-on distinguer divers accents ? Quels sont les indices qui permettent de reconnaître la langue maternelle ou l'origine géographique d'un locuteur ? La machine peut-elle contribuer à modéliser la perception humaine ? La même méthodologie peut ainsi être appliquée. À la fois pour des accents étrangers (allemand, anglais, arabe, espagnol, italien ou portugais) et des accents régionaux (français du Nord ou du Sud, mais également français de Suisse romande), nous sommes partis d'expériences perceptives d'identification, nous avons mesuré les traits phonétiques qui peuvent caractériser ces accents en tirant profit notamment de l'alignement automatique en phonèmes, et nous avons hiérarchisé les traits les plus discriminants en utilisant des techniques de classification.

Qu'entend-on par *accent* ? Ce terme relève aussi bien de la terminologie savante que du langage courant. Il est défini ainsi dans l'ouvrage de référence *Dialectology* (Chambers et Trudgill, 2004, p. 5) : « *'Accent' refers to the way in which a speaker pronounces, and therefore refers to a variety which is phonetically and/or phonologically different from other varieties.* » Mais les auteurs ajoutent quelques lignes plus loin : « *Dialects and accents frequently merge into one another without any discrete break.* » Dans le *Dictionary of Linguistics and Phonetics* (Crystal, 2003) la rubrique *accent* renvoie à : « *the cumulative auditory effect of those features of pronunciation which identify where a person is from regionally and socially* ». Si l'on comprend *regionally* au sens large, la définition couvre également ce qu'en « linguistique naïve » (*folk linguistics*) on désigne par « accent étranger ». Il est cependant des cas où l'on perçoit un accent (étranger ou autre) sans que l'on n'arrive à cerner ou à définir cet accent avec précision. La reconnaissance d'un accent fait appel à un double processus (Fries et Deprez, 2003) : celui du repérage (évaluation d'une divergence plus ou moins marquée vis-à-vis d'une norme) et celui de la catégorisation (identification proprement dite).

Une autre question qui se pose est : qu'est-ce qui caractérise un accent ? Est-ce que cela repose davantage sur la suite de phonèmes et l'instanciation des faits de coarticulation, sur la qualité de voix, sur les clichés mélodiques, sur les profils de durées, sur des erreurs de placement de l'accent lexical ou autres *schibboleths*, qui peuvent être interprétés comme étrangers ou régionaux ? Parfois, un simple élément comme un « *r* roulé » ou un [ɔ] ouvert dans, par exemple, « une autre chose jaune ou rose » peut colorer tout un énoncé et faire basculer la perception. Plus généralement, est-ce la combinaison de traits segmentaux et prosodiques qui aboutit

à la perception d'un accent ? Les réponses peuvent bien sûr dépendre des langues et des accents.

Parmi les indices qui contribuent à une impression d'accent espagnol en anglais, sont cités des facteurs affectant la structure syllabique, le timbre des voyelles, les consonnes (en particulier /s/~z/ et /b~/v/) ainsi que l'accent lexical (Magen, 1998). En anglais également, les occlusives peuvent trahir un accent arabe ou français (Flege, 1984). Pour l'accent allemand en anglais et l'accent anglais en allemand (Jilka, 2000) ainsi que pour l'accent espagnol en italien et l'accent italien en espagnol (Boula de Mareuil et Vieru-Dimulescu, 2006), la prosodie joue un rôle important. Comment utiliser ces résultats en perception pour identifier automatiquement différents accents ?

Dans le domaine de la reconnaissance de la parole, les études se focalisent sur les taux d'erreurs plus que sur l'identification des accents. Citons toutefois les travaux conduits sur les accents mandarin, allemand et turc en anglais (Arslan et Hansen, 1997) ou sur l'accent arabe et vietnamien en anglais (Kumpf et King, 1997). D'autres études, plus récentes, ont cherché à générer des variantes de prononciation qui sont spécifiques à l'anglais avec accent étranger (Livescu et Glass, 2000 ; Silke *et al.*, 2004). Des études d'orientation plus linguistique existent (Schaden, 2004 ; Bartkova et Jouviet, 2004) : elles quantifient des diminutions de taux d'erreur de reconnaissance, mais n'essaient pas d'expliquer comment identifier l'origine d'un accent étranger.

L'identification perceptive de dialectes, variétés ou accents régionaux a également retenu l'intérêt de phonéticiens, linguistes et sociolinguistes, en français (Bauvois, 1996 ; Armstrong et Boughton, 1998 ; Sobotta, 2006), en norvégien, en néerlandais et en anglais (Heeringa, 2004 ; van Bezooijen et Gooskens, 1999 ; Clopper et Pisoni, 2004) notamment. La reconnaissance de la parole a été exploitée pour l'identification de variétés de néerlandais (ten Bosch, 2000) ; et la dégradation des performances pour la reconnaissance de variétés non prévues au départ a été à l'origine d'études acoustiques sur des variétés d'anglais (Yan et Vaseghi, 2002 ; Hansen *et al.*, 2004). À notre connaissance, de tels travaux n'ont pas été menés sur le français régional.

Après une présentation du corpus et de la méthode utilisés (section 2), nous traiterons successivement des accents étrangers (section 3) et régionaux (section 4). Ces sections 3 et 4 adoptent le même plan : identification perceptive, caractérisation phonétique, hiérarchisation des traits de prononciation pour une identification automatique. Enfin, la section 5 conclut et ouvre quelques perspectives.

## **2. Corpus et méthode**

### **2.1. Corpus**

La présente étude repose essentiellement sur le corpus PFC (« Phonologie du Français Contemporain »). Ce projet (Durand *et al.*, 2003), dans le sillage de

travaux antérieurs (Martinet, 1945 ; Walter, 1982), a entrepris de collecter des enregistrements couvrant un vaste territoire francophone à travers une dizaine de locuteurs par point d'enquête. Pour chaque locuteur, on dispose des quatre styles contextuels distingués par (Labov, 1972), allant du plus formel au moins formel : la lecture d'une liste d'une centaine de mots et d'un texte d'une vingtaine de phrases, ainsi que 10 minutes d'entretien guidé et de conversation libre (30–60 minutes par locuteur au total).

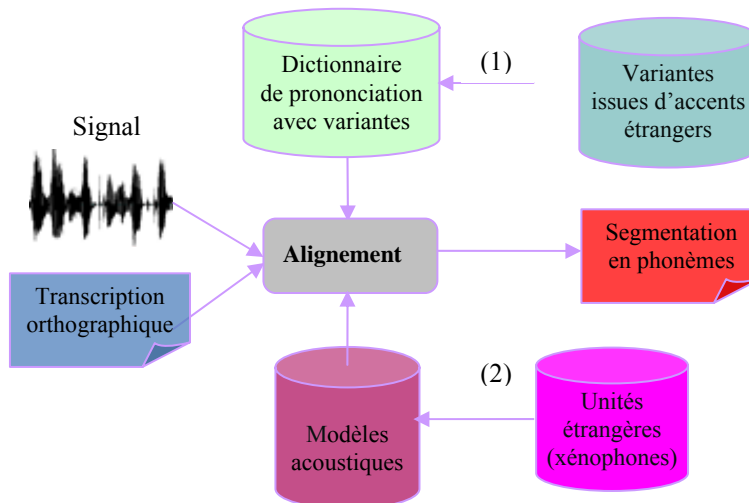
Pour les variétés régionales, nous nous appuyons ici sur douze points d'enquête : six dans la moitié nord de la France (Brécey, Brunoy, Dijon, Lyon-Villeurbanne, Roanne, Treize-Vents), un en Suisse romande (le canton de Vaud) et cinq dans le sud de la France (Biarritz, Douzens, Lacaune, Marseille, Rodez). Malgré un substrat franco-provençal, la Suisse romande sera, sauf à la fin de cet article, comptée comme nord, car sa variété de français est très peu perçue comme méridionale. Totalisant des dizaines d'heures de parole, ce matériel est constitué de plus de 100 locuteurs, autant d'hommes que de femmes, de tranches d'âge équilibrées, de niveaux d'études et de milieux socioprofessionnels variés, qui sont nés et ont passé la plus grande partie de leur vie en un même lieu. Nous disposons en outre d'un gros corpus de parole spontanée (environ 500 locuteurs) incluant également des variétés nord et sud de la France, qui pourra être utilisé à des fins de comparaison. Ce corpus est constitué de 367 conversations téléphoniques entre des locuteurs appelant de sept grandes régions françaises : Centre, Est, Nord, Ouest, Paris (comptés comme nord), Sud-Est et Sud-Ouest (comptés comme sud). Chaque région comprend environ 70 locuteurs, dont un tiers d'hommes. La durée moyenne d'une conversation est de 14 minutes.

Pour les accents étrangers, nous avons enregistré 72 locuteurs, répartis en deux groupes de 36 : l'un « d'entraînement » pour émettre des hypothèses, l'autre pour tester ces hypothèses sur un corpus non vu et mener des expériences d'identification automatique. Ces 72 locuteurs ont lu le texte ainsi que la liste de mots PFC (ici enrichie de mots internationaux), et pour le test le texte « La bise et le soleil », de l'Association Phonétique Internationale (API), qu'ils ont de plus traduit dans leur langue maternelle. Nous disposons également de parole spontanée (issue d'entretiens semi-directifs avec l'expérimentateur), qui a été utilisée dans une première expérience perceptive. Ces locuteurs étaient natifs de six langues : allemand, anglais, arabe, espagnol, italien et portugais. Le choix de ces langues a été établi en croisant des statistiques sur le tourisme et l'immigration, afin de s'assurer que des auditeurs français soient suffisamment familiers des accents correspondants. Les locuteurs retenus pour l'entraînement (autant d'hommes que de femmes, tous étudiants) étaient en moyenne âgés de 24 ans, vivaient en France (dans la région parisienne) depuis 16 mois et avaient commencé à étudier le français à l'âge de 17 ans. Les hispanophones n'étaient ni catalans ni latino-américains, et les arabophones venaient de différents pays du Maghreb – mais une étude antérieure a montré la difficulté à discriminer les origines possibles, algérienne, marocaine ou tunisienne, de locuteurs parlant français (Boula de Mareuil *et al.*, 2004). Les

locuteurs retenus pour le test (également européens ou originaires de pays arabes) étaient des étudiants d'une moyenne d'âge de 27 ans, arrivés en France depuis 21 mois et ayant commencé à apprendre le français à l'âge de 15 ans en moyenne. En plus de ces deux groupes à bien des égards comparables, nous avons enregistré dans les mêmes conditions 12 jeunes locuteurs de la région parisienne, natifs du français et sans accent marqué, pour référence. Nous reviendrons sur cette « absence » d'accent.

## 2.2. Méthode

Une dizaine d'heures de parole avec des accents étrangers et plus d'une centaine d'heures avec des accents régionaux ont été transcrites orthographiquement et segmentées en phonèmes à l'aide de l'alignement automatique dérivé du système de reconnaissance du LIMSI (Gauvain *et al.*, 2005 ; Lamel et Gauvain, 2003). La méthode a été utilisée dans plusieurs études qui ont permis de la valider (Adda-Decker et Lamel, 1999 ; Adda-Decker *et al.*, 2005 ; Gendrot et Adda-Decker, 2005). À partir d'un signal de parole et de sa transcription orthographique, en utilisant des modèles acoustiques (ici indépendants du contexte) ainsi qu'un dictionnaire de prononciation qui peut inclure des variantes, le décodeur produit la séquence de phonèmes réalisée la plus probable. La figure 1 montre le diagramme bloc de l'alignement automatique dans lequel, pour étudier les accents étrangers, nous avons ajouté des variantes de prononciations qui, comme les confusions /s/~z/, sont liées aux différents accents (1), et introduit des unités comme le « r roulé », provenant de modèles acoustiques établis pour d'autres langues que le français (2).



**Figure 1.** Diagramme bloc de l'alignement automatique (voir le texte)

Nous avons ici tiré profit de l'alignement automatique en phonèmes pour tracer des triangles vocaliques, mesurer des (ratios de) durées, calculer des taux de voisement, extraire des patrons mélodiques et quantifier des réalisations de phonèmes tels que /s/~z/ ou un /R/ éventuellement roulé. Nous avons ensuite cherché à hiérarchiser les traits les plus discriminants au moyen de techniques d'analyse ou fouille de données (*data mining*): *clustering*, *scaling*, analyse discriminante, arbres de décision, etc.

### 3. Accents étrangers

#### 3.1. Identification perceptive

Des expériences perceptives ont été menées pour déterminer dans quelle mesure des auditeurs français natifs sont capables d'identifier les accents étudiés. Après une phase de familiarisation, la tâche des auditeurs consistait entre autres choses à identifier la langue maternelle des locuteurs et à évaluer le degré d'accent sur une échelle graduée de 0 à 5. Les degrés proposés étaient paraphrasés de la façon suivante : (0) pas d'accent, (1) petit accent, (2) accent modéré, (3) assez fort accent, (4) fort accent, (5) très fort accent.

Une première expérience avec les locuteurs retenus pour l'entraînement d'expériences d'identification automatique (enTrain) impliquait des extraits d'environ 10 secondes de parole spontanée, avec un choix forcé (sans distracteur ni classe rejet) entre six possibilités : allemand, anglais, arabe, espagnol, italien et portugais (Vieru-Dimulescu et Boula de Mareuil, 2006). Les échantillons avaient été sélectionnés selon les critères suivants : absence de référence culturelle et d'erreur morphosyntaxique typique d'une origine donnée, pas trop d'hésitations et cohérence thématique de l'énoncé. Les auditeurs, munis d'un microphone, étaient de plus invités à réagir verbalement à l'écoute de chaque stimulus (en l'imitant voire en le caricaturant) ou à écrire leurs commentaires dans une fenêtre de texte. Les consignes suggéraient simplement de préciser quels traits leur semblaient marquants dans la prononciation et l'intonation du locuteur, et sur quels indices ils fondaient leurs décisions.

Une deuxième expérience (Test) avec les locuteurs retenus pour le test d'expériences d'identification automatique, dont 6 locuteurs natifs du français impliquait la lecture du texte de l'API (125 mots dans sa traduction française représentant environ 1 minute de parole) et un choix forcé entre sept possibilités : le français en plus des six origines ci-dessus. Chaque expérience, qui durait environ 30 minutes, a été conduite auprès de 25 auditeurs de la région parisienne, de langue maternelle française, sans problèmes d'audition connus. Les tests étaient réalisés à travers une interface conviviale, qui permettait entre autres d'entrer des informations sur la familiarité avec les différents accents et de saisir les réponses. Les stimuli (préalablement égalisés) étaient présentés dans un ordre aléatoire

différent pour chaque auditeur. Chaque stimulus pouvait être réécouté, arrêté au milieu ou repris à partir d'un certain point, mais il était impossible de revenir en arrière une fois passé au stimulus suivant.

La majorité des sujets, avant le test, s'estimaient capables de reconnaître les accents arabe, allemand et anglais, mais ils étaient en moins grand nombre à penser pouvoir reconnaître les accents espagnol, portugais et italien. Ces tendances ne vont pas de pair avec les connaissances des auditeurs dans les langues correspondantes, également autoévaluées par les sujets : presque tous, par exemple, déclaraient qu'ils n'avaient pas ou que peu de connaissances en arabe, alors que presque tous se sentaient capables de reconnaître un accent arabe en français.

Dans chacune des deux expériences, il ressort que le degré d'accent des locuteurs non natifs a été jugé moyen (cf. tableau 1) : il a été évalué à 2,7 sur 5 en moyenne. Il était comparable entre les différents groupes linguistiques, sauf pour les Arabes retenus pour le test, qui avaient un accent plus léger (1,5).

| Expérience enTrain |           |           |           |           |           |           |
|--------------------|-----------|-----------|-----------|-----------|-----------|-----------|
| rép<br>orig        | Al        | A<br>n    | Ar        | Es        | It        | Po        |
| Al                 | <b>63</b> | 15        | 6         | 3         | 5         | 8         |
| An                 | 28        | <b>49</b> | 9         | 9         | 3         | 3         |
| Ar                 | 6         | 1         | <b>77</b> | 2         | 5         | 8         |
| Es                 | 3         | 3         | 5         | <b>59</b> | 19        | 11        |
| It                 | 5         | 3         | 7         | 34        | <b>40</b> | 10        |
| Po                 | 17        | 8         | 17        | 21        | 12        | <b>25</b> |
| Degré              | 2,2       | 3,0       | 2,4       | 2,9       | 3,1       | 2,4       |

| Expérience Test |           |           |           |           |           |           |           |
|-----------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| rép<br>orig     | Al        | A<br>n    | Ar        | Es        | It        | Po        | Fr        |
| Al              | <b>65</b> | 15        | 3         | 2         | 5         | 9         | 1         |
| An              | 15        | <b>73</b> | 3         | 3         | 2         | 3         | 0         |
| Ar              | 14        | 10        | <b>36</b> | 9         | 15        | 7         | 10        |
| Es              | 1         | 0         | 2         | <b>67</b> | 15        | 15        | 0         |
| It              | 3         | 0         | 3         | 22        | <b>46</b> | 23        | 3         |
| Po              | 11        | 5         | 11        | 19        | 19        | <b>34</b> | 1         |
| Fr              | 2         | 0         | 1         | 0         | 1         | 0         | <b>96</b> |
| Degré           | 2,9       | 3,1       | 1,5       | 3,0       | 2,4       | 3,0       | 0,6       |

**Tableau 1.** Matrices de confusion pour les expériences enTrain et Test entre les accents allemand, anglais, arabe, espagnol, italien, portugais et français natif (%). La dernière ligne indique les degrés d'accents moyens correspondants, évalués sur une échelle de 0 à 5

La langue maternelle des locuteurs a été correctement identifiée à plus de 50 %, ce qui est beaucoup mieux que le hasard. Le taux d'identification correcte était de 52 % dans l'expérience enTrain, de 60 % (ou 54 %) dans l'expérience Test (selon qu'on considère ou non les locuteurs français, qui ont été bien reconnus à 96 %). Des tests de khi-deux ( $\chi^2$  révèlent que pour chaque langue on est significativement au-dessus du seuil de hasard. À chaque fois, la réponse majoritaire était la bonne pour une origine donnée, ce qui reste vrai pour la plupart des locuteurs (25 locuteurs sur 36 dans l'expérience enTrain, 28 non natifs et les 6 natifs dans l'expérience Test). Dans les deux expériences, les confusions les plus fréquentes portaient sur les

accents espagnol/italien (*cf.* tableau 1). Des techniques d'échelonnement multidimensionnel (scaling) et de clustering implémentées dans le logiciel R<sup>1</sup>, comme on le verra au paragraphe 4.1, permettent de visualiser ce fait. Fournissant une sorte de distance perceptive entre les différents accents, elles rassemblent les Anglais et les Allemands dans un groupe, et dans un autre groupe les locuteurs de langue romane – non français, lesquels se voient bien isolés dans la deuxième expérience. L'accent le moins bien reconnu a toujours été le portugais. Cependant, si l'accent arabe a été le mieux identifié dans l'expérience enTrain, c'est l'origine anglaise (après la française) qui a été la mieux reconnue dans l'expérience Test. Cette différence s'explique par le degré d'accent moindre des Arabes retenus pour le test (1,5). Quant à l'accent portugais, le stéréotype chuintant qui lui est souvent associé et qui est loin de la réalité peut expliquer le fait qu'il soit mal reconnu.

Les auditeurs de l'expérience enTrain ont déclaré s'être appuyés sur des indices dont on peut vérifier la pertinence par des mesures acoustiques. Parmi les indices segmentaux, nos 25 auditeurs ont relevé : le *r*, qu'il soit « roulé » ou prononcé « à l'anglaise » (93 fois) ; [i] à la place de /e/ dans le cas des locuteurs de langue maternelle arabe (31 fois) ; *yé* à la place de *je*, [v] à la place de /b/ et [s] à la place de /z/ pour les Espagnols (38 fois) ; [z] à la place de /s/ pour les Allemands (24 fois) ; [u] à la place de /y/ ou l'inverse, ainsi qu'une mauvaise réalisation des nasales (37 fois), sans rapprochement avec une origine particulière, mais plutôt signe d'un accent étranger en général. Parmi les traits suprasegmentaux notés par les sujets, on peut citer : des phrases « chantantes » qui seraient typiques des Italiens ou une « précipitation » sur certains mots. Aussi les analyses acoustiques suivantes ont-elles été entreprises. L'influence du système phonologique de la langue maternelle (L1) sur la perception et la production d'une langue seconde (L2) est traitée dans nombre d'études psycholinguistiques sur la parole non native : le fait que, par exemple, l'allemand possède un /y/ dans son inventaire de phonèmes, alors qu'en anglais il n'a pas de rôle fonctionnel – même si un son proche peut être entendu dans un mot comme *due*. Si d'autres dimensions comme le rythme sont des catégories phonologiques, elles devraient permettre de déceler une langue maternelle différente de la langue parlée (Freland-Ricard, 1996). La part des indices segmentaux et suprasegmentaux sera examinée dans la sous-section suivante.

### 3.2. Caractérisation phonétique

Des analyses phonétiques ont été conduites sur la lecture du texte PFC (400 mots) par les 36 locuteurs retenus pour l'entraînement ainsi que 6 locuteurs natifs du français, de la même tranche d'âge (Vieru-Dimulescu *et al.*, 2007a). Ce matériau commun, moins restreint que nos échantillons de parole spontanée, facilite l'analyse et se prête mieux aux comparaisons entre locuteurs. Même si en

---

<sup>1</sup> [www.r-project.org/](http://www.r-project.org/)



contrepartie il reflète moins bien la façon naturelle de parler, ce corpus permet la comparaison directe, sur le même texte, avec les natifs français du projet PFC.

Sur cette base, nous avons d'abord mesuré les formants des voyelles, la durée et le voisement des consonnes, ainsi que des indices prosodiques. Ensuite, nous avons introduit des variantes de prononciations combinant des unités acoustiques françaises et étrangères (xénophones).

Les commentaires des auditeurs enregistrés lors de la première expérience perceptive (enTrain) ainsi que nos connaissances linguistiques nous ont conduits à faire l'hypothèse qu'un certain nombre de traits pouvaient être caractéristiques des accents étudiés : formants de voyelles, durée et voisement de consonnes, variantes de prononciation, indices prosodiques liés au schwa final... Les fréquences des formants ont été mesurées sur les voyelles (plus de 500 par locuteur) grâce au logiciel Praat<sup>2</sup>. Les valeurs des formants et de la fréquence fondamentale ( $F_0$ ) ont été prélevées toutes les 10 ms en utilisant les options standard de Praat. Des filtres ont été prévus pour les formants (adaptés à chaque voyelle, hommes et femmes distingués) afin d'écartier les valeurs aberrantes hors d'un intervalle tolérant de  $\pm 500$  Hz en moyenne par rapport à des valeurs de référence (Calliope, 1989 ; Gendrot et Adda-Decker, 2005). Seulement 5,5 % des voyelles ont ainsi été rejetées. Les formants ont ensuite été normalisés à l'aide de diverses procédures décrites par (Adank, 2003), pour minimiser les différences dues aux caractéristiques physiologiques des locuteurs. Nous avons, pour cette étude, retenu la procédure log-mean de Nearey (Disner, 1980). Quant au taux de voisement, il a été défini pour chaque phonème comme le nombre de mesures voisées divisé par le nombre total de mesures.

Les résultats des analyses phonétiques mettent en évidence des prononciations qui, dans une large mesure, confirment à la fois les prédictions linguistiques et les jugements des auditeurs :

- antériorisation du /u/ chez les Anglais, phénomène largement décrit de /u/-*fronting* (Harrington *et al.*, 2000) qui s'observe également ici dans l'accent anglais en français ;
- postériorisation du /y/ chez les Espagnols et les Italiens ;
- rapprochement de /e/, /i/ et /y/ chez les Arabes, qui n'ont pas ces distinctions au moins dans le système à trois voyelles phonologiques (/a u i/) de l'arabe classique ;
- schwa antériorisé chez les Espagnols et les Italiens, et fermé chez les Portugais, se rapprochant du /i/ haut central de leur système phonologique (Veloso, 2007) ;
- /p/, /t/, /k/ longs (aspirés) chez les Anglais et les Allemands, alors que les débits de parole sont comparables entre les locuteurs non natifs (*cf.* tableau 3) ;

---

2. [www.fon.hum.uva.nl/praat](http://www.fon.hum.uva.nl/praat)

- /b/ et /v/ courts chez les Espagnols (qui n’ont pas le contraste /b/~v/ dans leur L1) ;
- /ɸ/ court chez les Italiens ;
- dévoisement de /b/, /d/, /g/ chez les locuteurs germaniques ;
- dévoisement de /v/ et /ʒ/, également, chez les Anglais ;
- dévoisement de /z/ chez les Espagnols, qui n’ont pas de fricatives voisées dans leur inventaire phonologique.

REMARQUE. — Il est intéressant de noter que la réalisation du /y/ français suit un mouvement contraire entre les locuteurs espagnols ou italiens notamment (chez qui elle est plus proche du [u]) et les locuteurs arabes (chez qui elle tend vers le [i]). Les uns privilégient le trait [+ arrondi], les autres le trait [+ antérieur]. Ce phénomène souvent caricaturé est connu (Rochet, 1995) : on peut le retrouver dans des transcriptions ludiques telles que *tou m’as toué* ou bien *Itats-Inis*. Il est bien mis en évidence par *scaling* ou *clustering* à partir d’une caractérisation de chaque locuteur par les coordonnées moyennes de son /y/ dans le plan F1/F2 : on obtient – et de même si l’on introduit le troisième formant (F3) – une représentation où les locuteurs arabes d’une part (6 sur 6) et les locuteurs italiens et espagnols d’autre part (5 + 4 sur 12), sont assez bien regroupés. Quant au premier formant, les valeurs rééchelonnées les plus basses sont obtenues pour le /e/ chez les Arabes (370 Hz) et pour le /ə/ chez les Portugais (350 Hz). Nous reviendrons plus loin sur la fermeture de ces phonèmes.

En utilisant un dictionnaire de prononciation dans lequel à chaque entrée correspond une ou plusieurs prononciations du français « standard », nous pouvons donc mesurer des réalisations phonétiques de locuteurs non natifs qui dévient de la forme canonique. Nos analyses acoustiques suggèrent que ces déviations peuvent être communes aux locuteurs d’une même origine. Sur cette base et comme mentionné plus haut, nous avons défini un ensemble de 20 variantes de français non natif : /b/→[v], /v/→[b], /s/→[z], /z/→[s], /ʒ/→[j], /ʒ/→[ʃ], /ʃ/→[tʃ], /b/→[p], /d/→[t], /g/→[k], /v/→[f], /ɸ/→[l], /ɸ/→[w], /l/→[w], /o/→[ɔ], /e/→[ɛ], /y/→[u], /y/→[i], /e/→[i], en plus des voyelles nasales – /ã/, /ẽ/ et /õ/ éventuellement dénasalisées et suivies d’un appendice nasal ([n] ou [m] avant *p/b*). Dans chaque paire (par exemple /b/→[v]), le premier phonème pouvait être reconnu comme le second, noté conventionnellement entre crochets puisqu’il renvoie à une réalisation phonétique. Pour chaque variante, un dictionnaire de prononciation spécifique a été construit, et un alignement a été produit en conséquence.

Ces variantes permettent d’évaluer des confusions entre phonèmes français, faites par des locuteurs non natifs. Nous avons également inclus des phonèmes ou allophones comme [r], [ɾ], [ʀ], [ʊ], [ʂ], [j] et [β], qui, motivés par des mécanismes linguistiques, peuvent être particuliers aux locuteurs de certaines origines (Delattre, 1965). Les xénophones [r], [ʂ], [j] et [β] introduits dans les modèles acoustiques français ont été empruntés à l’espagnol, langue pour laquelle nous disposons de

modèles entraînés sur de grandes quantités de données. Rappelons que l'espagnol n'a pas /v/ dans son inventaire de phonèmes (Delattre, 1965) : un [b] est réalisé après pause ou nasale, un [β] apparaît ailleurs. Quant aux xénophones [ɹ], [ʔ] et [ʊ], ils venaient de l'anglais. Le [ʔ] sombre, par exemple, est un allophone du /l/ en anglais mais également en portugais (Delattre, 1965). Les variantes retenues, incluant des xénophones, sont présentées avec les résultats qu'elles fournissent dans le tableau 2.

|                                     |            | Al        | An        | A<br>r    | Es        | It        | Po | Fr | Moyenne |
|-------------------------------------|------------|-----------|-----------|-----------|-----------|-----------|----|----|---------|
| Alternances entre phonèmes français | /b/ → [v]  | 32        | 30        | 8         | <b>60</b> | 8         | 22 | 3  | 23      |
|                                     | /v/ → [b]  | 14        | 17        | 2         | 23        | <b>28</b> | 2  | 5  | 13      |
|                                     | /s/ → [z]  | 4         | 3         | 1         | 4         | <b>12</b> | 7  | 1  | 5       |
|                                     | /z/ → [s]  | 32        | 47        | 26        | <b>79</b> | 31        | 19 | 24 | 37      |
|                                     | /ʒ/ → [j]  | 7         | 11        | 7         | <b>29</b> | 1         | 7  | 4  | 9       |
|                                     | /ʒ/ → [ʃ]  | 14        | <b>26</b> | 11        | 25        | 5         | 6  | 12 | 14      |
|                                     | /ʃ/ → [tʃ] | 2         | 7         | 5         | <b>15</b> | 6         | 3  | 3  | 6       |
|                                     | /b/ → [p]  | 42        | <b>55</b> | 8         | 31        | 3         | 6  | 6  | 21      |
|                                     | /d/ → [t]  | 30        | <b>59</b> | 9         | 30        | 6         | 9  | 12 | 22      |
|                                     | /g/ → [k]  | 59        | <b>67</b> | 36        | 43        | 13        | 30 | 20 | 38      |
|                                     | /v/ → [f]  | 8         | <b>28</b> | 9         | 15        | 5         | 12 | 12 | 13      |
|                                     | /ʁ/ → [l]  | 7         | 32        | 4         | 7         | <b>46</b> | 6  | 6  | 15      |
|                                     | /ʁ/ → [w]  | 4         | 12        | 2         | 5         | <b>14</b> | 3  | 2  | 6       |
|                                     | /l/ → [w]  | 2         | <b>8</b>  | 1         | 1         | 3         | 5  | 3  | 3       |
|                                     | /o/ → [ɔ]  | 16        | 56        | 18        | <b>70</b> | 32        | 38 | 45 | 39      |
|                                     | /e/ → [ɛ]  | 15        | <b>50</b> | 17        | 47        | 39        | 26 | 19 | 30      |
|                                     | /y/ → [u]  | 5         | 21        | 8         | 32        | <b>35</b> | 21 | 3  | 18      |
|                                     | /y/ → [i]  | 34        | 34        | 32        | <b>36</b> | 26        | 30 | 26 | 31      |
|                                     | /e/ → [i]  | <b>18</b> | 15        | 15        | 11        | 8         | 7  | 9  | 12      |
| V nasale → V +<br>nasale            | 28         | 41        | 22        | <b>69</b> | 63        | 46        | 7  | 39 |         |
| Xénophones                          | /ʁ/ → [r]  | 14        | 33        | 7         | 9         | <b>62</b> | 12 | 8  | 21      |
|                                     | /ʁ/ → [ɹ]  | 6         | 21        | 3         | 4         | <b>22</b> | 4  | 2  | 9       |
|                                     | /l/ → [ʔ]  | 3         | <b>10</b> | 2         | 7         | 7         | 8  | 3  | 6       |
|                                     | /u/ → [ʊ]  | 12        | <b>56</b> | 16        | 38        | 15        | 26 | 12 | 25      |
|                                     | /s/ → [ʃ]  | 30        | 31        | 29        | <b>56</b> | 43        | 36 | 10 | 34      |
|                                     | /Z/ → [j]  | 34        | 35        | 41        | <b>55</b> | 40        | 45 | 23 | 39      |
|                                     | /b/ → [β]  | 16        | 26        | 5         | <b>43</b> | 9         | 23 | 9  | 19      |
|                                     | /v/ → [β]  | 26        | 19        | 8         | <b>43</b> | 36        | 5  | 5  | 20      |

**Tableau 2.** Pourcentages de variantes alignées en utilisant (1) des alternances entre phonèmes français, (2) des xénophones

Résumons ce qu'engendrent ces alignements autorisant des alternances entre phonèmes français ou des xénophones (tableau 2) : à l'exception notable du rapprochement /e/~y/~i/ qui apparaît pour les Arabes dans les triangles vocaliques (non représentés ici), la plupart des résultats sont cohérents avec les mesures à base de formants, durées et  $F_0$ . Les Italiens voient leur /y/ aligné avec [u] et leur /ʁ/ avec une liquide : le *r* est le plus souvent roulé – comme parfois pour les Anglais (ce que nous avons vérifié perceptivement), mais pas les Espagnols, qui préfèrent approximer le /ʁ/ français par une sorte de *jota*. Les occlusives sonores des Anglais sont le plus souvent alignées avec leurs contreparties françaises sourdes, et leur /u/ avec le [u] centralisé. Les Espagnols tendent à prononcer [s] pour /z/, [v] ou [β] pour /b/, [j] ou [j̄] pour /ʒ/ et [tʃ] pour /ʃ/. Espagnols et Italiens montrent 10 fois plus d'appendices nasaux après voyelle nasale que les natifs du français.

De manière synthétique, des indices pertinents pour les accents étrangers, influencés par les langues maternelles correspondantes, sont d'après les analyses de formants, de la durée et du voisement des consonnes, des variantes de prononciation : les occlusives aspirées ou dévoisées chez les Allemands et les Anglais, le mouvement de /e/ vers [i] chez les Arabes, le [ʃ] (se substituant éventuellement à /z/) et les confusions /b/~v/ chez les Espagnols, le [r] roulé chez les Italiens, et la prononciation fermée du schwa chez les Portugais (Vieru-Dimulescu *et al.*, 2007b). Ces résultats concernent des indices segmentaux. Rappelons pourtant que les sujets de notre première expérience perceptive (enTrain), sans recourir au jargon technique des spécialistes, ont mentionné des indices qui pouvaient être perçus ou se comprendre comme prosodiques. Des indices prosodiques ont pu être mesurés : ils figurent dans le tableau 3, avec notamment la différence (en demi-tons) entre la  $F_0$  moyenne de la voyelle précédant un schwa final et celle de ce schwa final prononcé. On observe que dans les mots polysyllabiques terminés par un schwa, les Italiens réalisent le plus grand nombre de schwas (23 % de 123 schwas potentiels). Ceci peut s'expliquer par le fait qu'en italien, terminer un mot plein par une consonne est extrêmement rare. En outre, les Italiens allongent la voyelle pénultième et montrent une baisse de  $F_0$  sur le schwa final. Les Allemands, quant à eux, montrent une augmentation de  $F_0$  sur le schwa final.

En plus de ces indices liés au schwa, nous avons mesuré la variation en durée des intervalles vocaliques en termes d'écart type ( $\Delta V$ ) comme proposé par (Ramus, 1999) et l'indice PVI (*Pairwise Variability Index*) de (Grabe et Low, 2002) sur les voyelles. Ces deux paramètres sont plus élevés pour les Italiens, en accord avec leur tendance à allonger les syllabes accentuées par rapport aux syllabes inaccentuées.

|                | Al          | An  | Ar  | Es  | It         | Po  | Fr  | Moyenne |
|----------------|-------------|-----|-----|-----|------------|-----|-----|---------|
| %s chwa        | 14          | 15  | 10  | 11  | <b>23</b>  | 15  | 11  | 14      |
| durée(pén/fin) | 1,9         | 1,6 | 2,0 | 2,1 | <b>2,4</b> | 2,2 | 1,9 | 2,0     |
| $\Delta F_0$   | <b>-0,1</b> | 0,9 | 1,3 | 1,0 | 2,2        | 0,7 | 1,1 | 1,0     |
| Durée ph.      | 89          | 90  | 91  | 89  | 91         | 94  | 73  | 88      |
| $\Delta V$     | 59          | 63  | 51  | 59  | <b>69</b>  | 67  | 39  | 58      |
| PVI            | 54          | 62  | 49  | 54  | <b>62</b>  | 59  | 46  | 55      |

**Tableau 3.** *Schwa final et indices prosodiques : durée(pén/fin) – respectivement  $\Delta F_0$  – représente le ratio de durée – respectivement la différence de  $F_0$  en demi-tons – entre la voyelle pénultième et le schwa final. Durée ph. représente la durée moyenne des phonèmes,  $\Delta V$  l'écart type de la durée des intervalles vocaliques et PVI le Pairwise Variability Index sur les voyelles (en ms). Comme dans d'autres tableaux, les valeurs extrêmes sont en gras*

### 3.3. Hiérarchisation des traits

Dans la sous-section précédente, nous avons décrit des traits qui peuvent contribuer à la caractérisation d'accents étrangers en français. Le but de cette sous-section est de vérifier si les tendances dégagées restent valides pour de nouveaux locuteurs et un nouveau corpus, de déterminer automatiquement les attributs les plus pertinents dans l'identification des accents, et de comparer l'importance relative des voyelles, des consonnes et de la prosodie en utilisant différentes techniques. Des techniques d'analyse de données implémentées dans le logiciel Weka (Witten et Frank, 2005) ont été utilisées pour sélectionner les traits les plus discriminants et classer les locuteurs selon leur accent. Au nombre de 87, ces traits englobent les deux premiers formants des voyelles orales ( $2 \times 10$ ), la durée et le voisement des consonnes ( $2 \times 17$ ), les alternances entre variantes françaises (20), les xénophones (8) et les indices prosodiques (5) – ceux que rapporte le tableau 3 hormis *Durée ph.*

Comme certains traits peuvent être non pertinents ou redondants, nous avons eu recours à la sélection automatique sur nos données d'entraînement (la lecture du texte PFC par les locuteurs de l'expérience perceptive enTrain et 6 natifs français, comme au paragraphe 3.2). Cette sélection vise à supprimer des traits inadéquats et à améliorer la performance des algorithmes d'apprentissage (Guyon et Elisseeff, 2003), tous les attributs étant également importants dans cette sélection.

Nous avons appliqué 7 algorithmes de sélection implémentés dans Weka (ex. : *Information Gain*, *Principal Components Analysis*). Comme ces algorithmes fournissent différents classements et nombres de traits sélectionnés, nous avons employé un type de ROVER (*Recognizer Output Voting Error Reduction*) pour trier les traits qui obtiennent le meilleur score moyen (voir l'équation ci-dessous [1]).

$$score = \frac{p}{n} \sum_{i=1}^n (C - rang_i) \quad [1]$$

où  $rang_i$  est le rang obtenu par le trait avec l'algorithme  $i$ ,  $p$  est le nombre d'algorithmes qui sélectionnent ce trait,  $n$  est le nombre d'algorithmes utilisés (ici 7) et  $C$  est une constante fixée à 100.

Les meilleurs traits, sélectionnés par la plupart des algorithmes, sont : les deux premiers formants de /ə/, le deuxième formant du /e/ et du /a/, les pourcentages d'appendices nasaux et d'alignements /z/→[s], /b/→[v], /b/→[p], /d/→[t], /ʁ/→[l] et /ʁ/→[r].

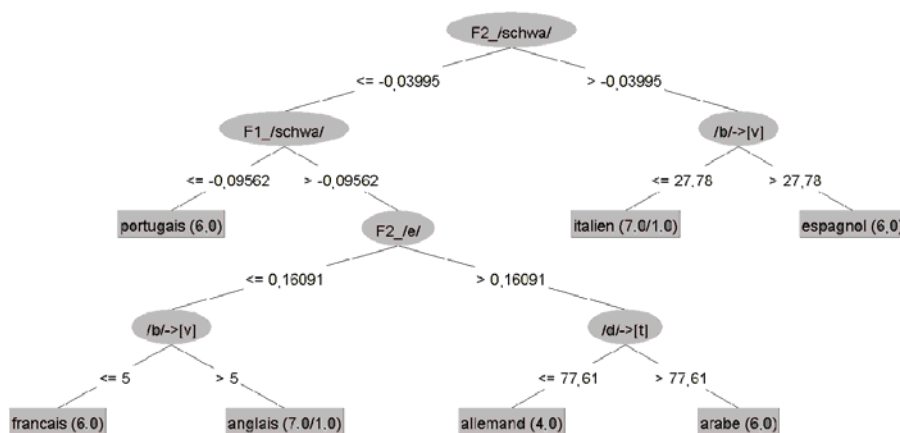
Des scores d'identification ont été calculés dans trois configurations incluant des données non vues. Dans l'expérience PFC, les locuteurs utilisés pour l'entraînement sont différents de ceux du test, mais ils ont lu le même texte. Dans l'expérience API, non seulement les locuteurs mais aussi les textes sont différents. Dans l'expérience L1O (*leave-one-out*), une méthode de validation croisée est utilisée sur l'ensemble des données (84 locuteurs lisant les textes PFC et API) : un seul locuteur à la fois est utilisé pour le test ; le reste dont on a retiré les deux textes du locuteur en question est utilisé pour l'apprentissage. Vingt algorithmes de classification compatibles avec nos données ont été appliqués : *Bayesian Networks*, *Logistic Regression Models*, *Multilayer Perceptrons*, *Support Vector Machines* (SVM), C4.5 (J48), *Random Forest*, etc. Comme les performances diffèrent dans une large mesure selon le type d'expérience et l'ensemble de traits, nous avons décidé de moyenner les taux d'identification correcte dans le tableau 4. En outre, moyenner les résultats rend ceux-ci plus comparables avec les résultats des tests perceptifs impliquant 25 auditeurs.

|                           | Expérience PFC | Expérience API | Expérience L1O |
|---------------------------|----------------|----------------|----------------|
| Meilleurs traits (12)     | 48             | 37             | 61             |
| Tous les traits (87)      | 47             | 35             | 64             |
| Voyelles (20)             | 36             | 27             | 45             |
| Consonnes (34)            | 39             | 19             | 46             |
| Prosodie (5)              | 26             | 16             | 18             |
| Variantes françaises (20) | 36             | 32             | 60             |
| Xénophones (8)            | 37             | 30             | 44             |

**Tableau 4.** Taux d'identification correcte (%) fondés sur les 12 meilleurs traits sélectionnés automatiquement, tous les traits ou des sous-ensembles linguistiques (formants des voyelles, durée et voisement des consonnes, prosodie, variantes françaises et xénophones)

Les résultats obtenus avec les meilleurs traits sont consignés dans le tableau 4. Ces résultats sont en moyenne stables au-delà de 15 traits – les performances n’augmentent pas proportionnellement au nombre de traits. En utilisant une méthode de validation croisée et en retenant une douzaine de traits pour les six accents étrangers étudiés ici et le français natif, on obtient une moyenne de 61 % d’identification correcte. Les meilleurs traits sélectionnés automatiquement (par exemple la confusion /b/~/v/ pour les Espagnols parlant français) sont robustes au changement de corpus et font sens au regard de connaissances linguistiques.

L’algorithme d’arbre de décision J48 n’est pas le meilleur – dans l’ensemble, les plus robustes semblent être les modèles de régression logistique et les SVM – mais ses résultats apparaissent parmi les meilleurs et sont directement interprétables (cf. figure 2). Il est à noter que l’arbre de décision utilise d’abord les deux premiers formants (normalisés) du /ə/ pour isoler l’accent portugais, et ceci que l’on retienne 12, 15 traits ou tous les attributs. Peu d’indices, en effet, caractérisent le français avec accent portugais, car cet accent est relativement peu marqué (cf. § 3.1). L’antériorisation du schwa permet d’isoler les Espagnols et les Italiens, lesquels sont départagés par le pourcentage de réalisation /b/→[v]. La fermeture-antériorisation de /e/ est quant à elle commune aux Arabes et aux Allemands, lesquels sont départagés par le pourcentage de dévoisement /d/→[t].



**Figure 2.** Arbre de décision obtenu en utilisant 12 traits (les chiffres séparés par des barres obliques désignent à gauche le nombre total de locuteurs classés sous le nœud en question, à droite le nombre de locuteurs classés par erreur sous ce nœud)

Le tableau 5 donne à gauche la matrice de confusion obtenue sur le texte API en moyennant les résultats de 20 algorithmes utilisant 12 traits. On est ici dans les mêmes conditions que lors de l’expérience perceptive Test, ce qui permet de

comparer identifications humaine et automatique toutes choses égales par ailleurs. Si le taux d'identification automatique est moindre (37 %, ce qui reste plus de deux fois supérieur au seuil de hasard), la réponse majoritaire est la bonne pour pratiquement chaque origine (à plus de 50 % pour les Espagnols et les Italiens, ici mieux distingués et identifiés que par les auditeurs de l'expérience perceptive Test). Les seules exceptions sont les origines portugaise et arabe, ce qui est aisément compréhensible d'après ce que nous avons vu : accent portugais difficile à cerner et faible degré d'accent des locuteurs arabes retenus pour le test (*cf.* § 3.1).

Dans le tableau 5, la matrice de confusion correspondant à l'expérience L10 montre des taux d'identification correcte plus élevés. Même si un biais est induit par les textes communs à l'entraînement et au test, le gain est notamment appréciable pour les locuteurs espagnols, caractérisés par plus d'un trait robuste. Il est intéressant également de noter les confusions entre Anglais et Allemands, qui demeurent malgré tout bien identifiés.

| Expérience API |           |           |    |           |           |    |           | Expérience L10 |           |           |           |           |           |           |           |
|----------------|-----------|-----------|----|-----------|-----------|----|-----------|----------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| rép<br>orig    | Al        | An        | Ar | Es        | It        | Po | Fr        | rép<br>orig    | Al        | An        | Ar        | Es        | It        | Po        | Fr        |
| Al             | <b>23</b> | 17        | 20 | 8         | 11        | 7  | 15        | Al             | <b>48</b> | 15        | 14        | 6         | 5         | 8         | 4         |
| An             | 16        | <b>36</b> | 6  | 7         | 16        | 4  | 16        | An             | 20        | <b>50</b> | 9         | 5         | 4         | 7         | 7         |
| Ar             | 16        | <b>28</b> | 27 | 3         | 2         | 2  | 23        | Ar             | 18        | 8         | <b>38</b> | 2         | 3         | 10        | 21        |
| Es             | 3         | 23        | 6  | <b>50</b> | 8         | 3  | 7         | Es             | 7         | 7         | 3         | <b>77</b> | 1         | 4         | 1         |
| It             | 6         | 10        | 11 | 4         | <b>51</b> | 2  | 17        | It             | 4         | 7         | 4         | 2         | <b>64</b> | 11        | 8         |
| Po             | 11        | <b>31</b> | 9  | 12        | 9         | 18 | 10        | Po             | 10        | 9         | 14        | 7         | 11        | <b>46</b> | 3         |
| Fr             | 0         | 25        | 17 | 1         | 0         | 4  | <b>53</b> | Fr             | 3         | 5         | 24        | 1         | 2         | 2         | <b>63</b> |

**Tableau 5.** Matrices de confusion obtenues dans les expériences API (test) et L10 (validation croisée) en moyennant les résultats de 20 algorithmes utilisant 12 traits.

Nous nous sommes enfin intéressés à la contribution des classes suivantes dans la classification des accents : les deux premiers formants des *voyelles*, la durée et le voisement des *consonnes*, la *prosodie* (incluant les phénomènes liés au schwa final), les alternances entre *variantes françaises* et les *xénophones*. Les dernières lignes du tableau 4 montrent les résultats en termes d'identification correcte, moyennés sur 20 algorithmes, en fonction des classes de traits utilisées. On voit dans ce tableau combien les classes *voyelles*, *consonnes* et *prosodie* sont dépendantes du corpus. Les alternances entre *variantes françaises* et les *xénophones* semblent plus robustes, même si les formants de certaines voyelles peuvent être utiles à l'identification automatique. La contribution de la *prosodie* reste modeste, y compris dans l'expérience L10, mais le but ici est moins d'atteindre des scores d'identification élevés que d'acquérir des connaissances linguistiques. Ce résultat est intéressant au vu d'études antérieures (Boula de Mareuil et Vieru-Dimulescu,



2006), la plupart des études en matière d'accents étrangers se concentrant sur le segmental (Lauret, 1998 ; Flege *et al.*, 2003).

#### 4. Accents régionaux

##### 4.1. Identification perceptive

Les accents étant des marqueurs d'identité, notamment les accents régionaux, la question de leur identification nous semble particulièrement importante. Elle pose des problèmes extrêmement complexes, auxquels s'ajoute celui du passage de la production à la perception. Toute identification perceptive, nécessairement par rapport à un prototype, implique un traitement *bottom-up* (à partir de l'input qu'est le signal acoustique) et un traitement *top-down* (à partir des représentations linguistiques). Les prototypes représentent des connaissances partagées qui, sans être universelles, sont utiles comme points de repère pour toute évaluation, pour faire des prédictions, pour organiser notre perception (Rosch, 1975 ; Desrosières et Thévenot, 2002). Un prototype existe-t-il pour toute variété de langue ? Un accent peut être stigmatisé, dévalorisé et générateur de ségrégation, ou au contraire revendiqué pour affirmer son identité, sa loyauté, son intégration à une communauté et se démarquer d'un autre groupe (Bourdieu, 1982). L'appréciation d'un accent est dictée par le statut social auquel celui-ci est attaché, d'où les difficultés. Aussi un accent peut-il être plus ou moins marqué, plus ou moins masqué. Une double force s'exerce en effet sur le langage, oscillant entre ce qui permet d'afficher une différence et ce qui favorise la communication. Concernant les accents régionaux, un mouvement de standardisation peut affecter les couches les plus jeunes de la population avec le développement des échanges, des médias, des transports. Dans ce contexte, la reconnaissance de la parole peut donner des indications précieuses de par leur objectivité. L'alignement automatique, en particulier, à travers les proportions de réalisations de variantes de prononciation qu'il permet de dégager, peut être vu comme un instrument de mesure de distance par rapport à une norme.

Des expériences perceptives dont le protocole est similaire à celui de la section précédente ont été répliquées avec des auditeurs des environs de Paris et d'Aix-Marseille, qui devaient localiser l'origine géographique de six accents régionaux du français (Woehrling et Boula de Mareuil, 2006). Les tests étaient fondés sur des enregistrements de locuteurs de six points d'enquête du projet PFC, représentant autant d'aires dialectales différentes : Brécey (Normandie), Treize-Vents (Vendée), le canton de Vaud (Suisse romande), Biarritz (Pays basque), Lacaune (Languedoc) et Marseille (Provence). Dans chacun de ces points d'enquête, 6 locuteurs ont été sélectionnés (3 hommes, 3 femmes), à l'intérieur de trois tranches d'âge : 15-30 ans, 30-60 ans, 60 ans et plus. Pour chacun des locuteurs, on avait la lecture d'une longue phrase tirée du texte PFC (25 mots) et un échantillon de parole spontanée tiré d'entretiens guidés : un énoncé assertif d'une durée équivalente à celle de l'extrait lu (une dizaine de secondes), avec un débit comparable, et sans référence à

un endroit introduisant un biais dans l'identification. Les stimuli, préalablement égalisés, étaient randomisés dans un ordre différent pour chaque sujet.

Les expériences d'identification perceptive de l'accent régional ont chacune été soumises à 25 auditeurs, tous de langue maternelle française, à travers une interface conviviale. Les auditeurs de la région parisienne étaient membres d'un laboratoire d'informatique (le LIMSI) ; ceux de la région marseillaise étaient membres d'un laboratoire de sciences du langage (le LPL). En moyenne, les auditeurs étaient âgés de 32 ans et avaient passé 20 ans dans leur région de résidence. Ils se disaient quasiment tous familiers des accents de Marseille et de Suisse romande, quasiment tous non familiers des autres accents. Lors d'un test mené auprès de 25 autres auditeurs de la région parisienne, un degré d'accent devait également être attribué à chaque extrait, sur une échelle à six degrés graduée de 0 (pas d'accent) à 5 (très fort accent). Chaque expérience, sans retour possible une fois passé au stimulus suivant, durait environ 20 minutes.

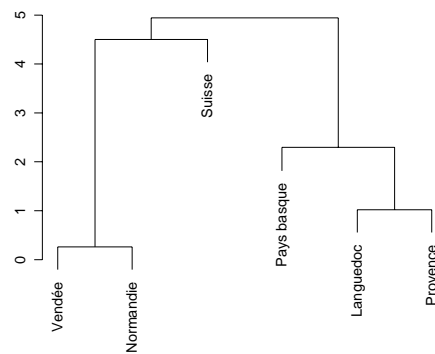
Le tableau 6 donne les résultats obtenus au cours des expériences d'identification perceptive menées dans les régions parisienne et marseillaise, ainsi que les degrés d'accent attribués. Avec 43 % d'identification correcte, ce qui est malgré tout significativement mieux que le hasard d'après des tests de  $\chi^2$ , ces expériences suggèrent qu'une discrimination fine entre accents régionaux est plus difficile qu'entre accents étrangers – pour lesquels dans une tâche similaire les taux d'identifications correctes étaient de 10 % supérieurs. Des expériences sur l'anglais britannique ou américain ont montré de même des performances assez modestes (van Bezooijen et Gooskens, 1999 ; Clopper et Pisoni, 2004). Le test d'évaluation du degré d'accent indiquait un accent moyen de 2 sur 5.

| Exp. menée en région parisienne |     |           |           |     |           |           | Exp. menée en région marseillaise |    |           |           |           |           |           |
|---------------------------------|-----|-----------|-----------|-----|-----------|-----------|-----------------------------------|----|-----------|-----------|-----------|-----------|-----------|
| rép orig                        | Ve  | No        | SR        | PB  | La        | Pr        | rép orig                          | Ve | No        | SR        | PB        | La        | Pr        |
| Ve                              | 37  | <b>48</b> | 5         | 2   | 7         | 1         | Ve                                | 34 | <b>47</b> | 8         | 5         | 5         | 1         |
| No                              | 37  | <b>51</b> | 2         | 3   | 6         | 1         | No                                | 38 | <b>46</b> | 3         | 6         | 6         | 2         |
| SR                              | 11  | 15        | <b>71</b> | 1   | 2         | 1         | SR                                | 13 | 13        | <b>73</b> | 1         | 0         | 0         |
| PB                              | 7   | 6         | 1         | 36  | <b>38</b> | 12        | PB                                | 10 | 5         | 2         | <b>41</b> | 32        | 10        |
| La                              | 6   | 2         | 0         | 25  | 29        | <b>38</b> | La                                | 6  | 5         | 0         | 28        | <b>33</b> | 29        |
| Pr                              | 12  | 8         | 2         | 20  | <b>30</b> | 29        | Pr                                | 10 | 13        | 1         | 13        | 27        | <b>36</b> |
| Degré                           | 1,1 | 0,8       | 2,5       | 2,5 | 3,4       | 2,0       |                                   |    |           |           |           |           |           |

**Tableau 6.** Matrices de confusion des expériences perceptives menées en région parisienne et en région marseillaise, entre Vendée, Normandie, Suisse romande, Pays basque, Languedoc et Provence (%). La dernière ligne indique le degré d'accent évalué sur une échelle de 0 à 5 par 25 autres auditeurs de la région parisienne.

Des six régions, la Suisse romande est la mieux reconnue (à plus de 70 %), mais les confusions sont fréquentes entre les trois variétés méridionales, de sorte que leur taux d'identification moyen est de 34 % – même si les accents du Sud ont été légèrement mieux reconnus par les auditeurs de la région marseillaise que par les auditeurs de la région parisienne (*cf.* tableau 6). Quarante-six auditeurs sur 50 s'estimaient pourtant capables, avant le test, de reconnaître l'accent de Marseille parmi les six proposés. Par ailleurs, l'origine des locuteurs les plus âgés (ceux qui, jugés par des auditeurs de la région parisienne ont le plus fort accent) est mieux reconnue que celle des plus jeunes. On observe pour les catégories de locuteurs des plus au moins âgés 47 %, 44 % et 38 % de réponses correctes, pour des degrés d'accent en moyenne de 2,7, 2,1 et 1,4.

Dans l'ensemble, deux ou trois accents émergent : français du Nord (incluant le suisse romand) *vs* français du Sud (à partir de techniques de *clustering*) ; français du Nord, français du Sud et suisse romand (à partir de techniques de *scaling* multidimensionnel). Ces techniques d'analyse de données fournissent des représentations graphiques des résultats, sous la forme respectivement de dendrogrammes et de plans à deux dimensions (Manning et Schütze, 1999). Les dendrogrammes résultant du *clustering* hiérarchique sont des arbres dans lesquels la distance verticale entre deux feuilles est fonction de la distance qui les sépare dans une matrice d'observations – la matrice de confusion dont chaque ligne permet de définir un vecteur caractérisant la région correspondante, dans le cas de la figure 3.



**Figure 3.** Dendrogramme issu du clustering hiérarchique agglomératif utilisant une distance euclidienne (à partir des réponses des 50 auditeurs)

Plusieurs algorithmes avec différentes distances ont été utilisés. Dans presque toutes les configurations, que ce soit pour les auditeurs de la région parisienne, ceux

de la région marseillaise ou l'ensemble des auditeurs, le *clustering* donne une bipartition qui oppose le Nord au Sud. Ce résultat corrobore en partie le fait que « la plupart des gens ne distinguent pas très bien les accents régionaux sauf s'ils sont très marqués. La seule perception générale est celle des accents du Midi, opposés aux accents du Nord de la France » (Léon et Léon, 1997, p. 102). Nous nous sommes donc focalisés dans ce qui suit sur les accents du nord de la moitié nord de la France et du sud, où une forme nivelée de français méridional semble émerger.

#### 4.2. Caractérisation phonétique

Sur les douze points d'enquête mentionnés en section 2, nous avons examiné des traits phonétiques tels que la réalisation des voyelles orales et nasales (Boula de Mareuil *et al.*, 2007). Nous avons quantifié des tendances connues et moins connues : en particulier l'avancée du /ɔ/ vers [œ] dans le Nord (Martinet, 1945, 1969 ; Malderez, 1995 ; Landick, 1995 ; Fagyal *et al.*, 2002), le maintien du schwa et la réalisation des voyelles nasales dans le Sud (Martinet, 1945 ; Walter, 1982 ; Carton *et al.*, 1983). Des changements linguistiques en cours au-dessous du seuil de la conscience peuvent affecter ces voyelles.

Dans un article célèbre, « C'est jeuili, le Mareuc ! », Martinet (1969) analysait cette avancée du /ɔ/ vers [œ] en termes de rendement fonctionnel (relativement faible et sans grande incidence sur la compréhension pour l'opposition /ɔ/~ /œ/). Déjà pendant la Seconde Guerre mondiale, à partir des témoignages d'officiers dans un camp de prisonniers, il avait observé l'émergence de cette variante antériorisée ou centralisée du /ɔ/ chez les locuteurs non méridionaux (Martinet, 1945). Il ouvrait ainsi des pistes pour des études empiriques et théoriques.

Si l'on s'en rapporte à l'histoire, d'ailleurs, un mot latin comme *florire* a naturellement donné le français *fleurir* ; le verbe *florir* (d'où *florissant*) n'est qu'un archaïsme littéraire. On a d'autre part en synchronie le doublet *senior~seigneur* (d'où *seigneurial* alors que l'adjectif correspondant à *directeur* est *directorial*), des alternances morphologiques comme *mort~meurt*, des erreurs comme *\*je vous serais gré* pour *je vous saurais gré*, *\*contreverse* pour *controverse* et *petit rond* pour *potiron*, qui appartient aujourd'hui au langage enfantin. Nous avons comparé cette tendance à l'antériorisation avec un phénomène symétrique : la postériorisation du schwa que des fautes d'orthographe d'enfants permettent également de dégager, reflétant des confusions voire une neutralisation partielle de l'opposition entre *re-* et *ro-* dans des mots comme *reblochon* (Malderez, 1995).

Pour quantifier la fréquence d'occurrence de ces variantes (/ɔ/→[œ], schwa et voyelles nasales), les dictionnaires de prononciation ont été adaptés séparément. Dans des mots tels que *sol* et *entier*, par exemple, les réalisations suivantes sont autorisées : [sɔl, sœl, sol], [ãtje, ãntje, antje]. Nous avons étudié l'influence de la région Nord/Sud, de l'âge et du sexe des locuteurs, du type de parole (lu/spontané), de la fréquence des mots et du contexte phonétique gauche/droit. Dans ce qui suit,

nous entendrons par « mots fréquents » les 610 mots les plus fréquents de notre corpus (5 % du vocabulaire). Le corpus analysé compte 12 000 mots distincts, 15 000 occurrences de /ɔ/, plus de 72 000 schwas et 60 000 voyelles nasales potentiels.

Comme on le voit dans le tableau 7, nous avons pu mettre en évidence 30 % d'avancée du /ɔ/ en [œ] dans le Nord, phénomène qui ne s'observe que dans 10 % des occurrences dans le Sud où la « loi de position » est mieux respectée (Durand et Lyche, 2004) – dans le Sud, on mesure 52 % de /ɔ/→[o]. Les locuteurs avaient des débits de parole comparables (10,1 ; 10,2 phonèmes/seconde) qui ne suffisent pas à expliquer ces différences.

|                  | % Alignement         | Nord | Sud |
|------------------|----------------------|------|-----|
| /ɔ/              | [ɔ]                  | 48   | 38  |
|                  | [œ]                  | 30   | 10  |
|                  | [o]                  | 22   | 52  |
| Schwa            | élision              | 63   | 49  |
|                  | maintien             | 29   | 39  |
|                  | postériorisation     | 8    | 11  |
| Voyelles nasales | voyelle nasale       | 80   | 47  |
|                  | V nasale + appendice | 15   | 18  |
|                  | V orale + appendice  | 5    | 35  |

**Tableau 7.** Réalisation du /ɔ/, du schwa et des voyelles nasales dans le Nord et dans le Sud (pourcentages issus de l'alignement).

On n'observe pas de différence majeure entre les hommes et les femmes d'une part, entre les locuteurs de moins de 30 ans et de plus de 60 ans d'autre part. Les mots les plus fréquents avec 'o' ne sont pas plus antériorisés que les autres, et la position en syllabe finale/non finale n'a que peu d'effet. En revanche, on a davantage d'antériorisation en parole spontanée qu'en lecture : plus de 10 % de différence entre la conversation libre et le texte PFC, que ce soit pour tous les locuteurs ou seulement ceux de la moitié nord de la France. En outre, les consonnes d'avant, en contexte gauche et droit, favorisent l'antériorisation du /ɔ/ (ex. : *personne, socialisme).*

On note également sans surprise, dans le tableau 7, que plus d'un schwa sur deux tombent dans le Nord. Surtout marquée en parole spontanée, la tendance des jeunes Méridionaux à élider davantage de schwas que leurs aînés est également confirmée (Armstrong et Unsworth, 1999). Mais une certaine postériorisation du /ə/ est également notable. Elle est en proportion équivalente dans le Nord et dans le Sud, entre les hommes et les femmes, jeunes et vieux, que ce soit en lecture ou en parole spontanée : elle touche 20 % des 'e' réalisés. C'est essentiellement un

contexte gauche en ‘r’ qui la conditionne (ex. : *revanche*, *relier*, *religion*, *rebelle*, *brebis*). Ces deux derniers exemples sont intéressants car ils sont souvent cités comme des cas d’école de ‘e’ graphique stable, ne pouvant tomber à l’oral (Hansen, 1994).

L’alignement automatique permet également de quantifier une tendance bien documentée : là où le français standard utilise des voyelles nasales, le français méridional prononce souvent des voyelles partiellement nasalisées et suivies d’un élément consonantique nasal bien audible (Durand, 1988). On constate effectivement davantage d’appendices nasaux dans le Sud que dans le Nord. Les appendices nasaux sont même majoritaires dans deux points d’enquête du Languedoc au parler conservateur. Les femmes ne nasalisent pas plus que les hommes, contrairement aux prédictions de (Thomas, 1991). En revanche, les Méridionaux de plus de 60 ans produisent près de 10 % d’appendices nasaux de plus que ceux de moins de 30 ans.

On observe légèrement moins d’appendices nasaux en parole spontanée et dans les mots les plus fréquents (avec dans les deux cas une différence de 5 %). Tandis qu’une consonne nasale en contexte gauche renforce la nasalisation, la prononciation d’un appendice nasal est favorisée avant une occlusive sonore (ex. : *vendre*). En proportion, ce qui correspond dans le Sud au ‘in’ du français standard est la voyelle, qui est le plus souvent suivie d’un appendice nasal. En outre, quand un appendice nasal apparaît, la voyelle est majoritairement dénasalisée, mais ceci est rare pour le ‘on’.

Les résultats attendus concernant la variation régionale et stylistique suggèrent que la méthode à base d’alignement est appropriée, et permet d’éclairer de nouveaux phénomènes plus ou moins motivés phonétiquement. En particulier, l’antériorisation du /ɔ/ est patente en français du nord de la Loire, notamment (mais pas seulement) avant ou après consonne d’avant. Si en parole spontanée on a tendance à moins dénasaliser les voyelles nasales dans le Sud et à davantage antérioriser le /ɔ/ dans le Nord, on peut risquer l’hypothèse qu’on est face à un changement linguistique qui gagne du terrain. L’antériorisation du /ɔ/ pourrait devenir une variable discriminant le Nord du Sud mieux que les traditionnelles voyelles nasales, pour les locuteurs les plus jeunes. Nous allons tenter de le vérifier dans ce qui suit.

### 4.3. Hiérarchisation des traits

L’approche à base d’alignement manipule des classes symboliques qui sont, à notre sens, intéressantes pour des interprétations catégorielles en phonologie. Nous avons complété cette approche par une étude à base de triangles vocaliques, pour valider la méthodologie, au moins sur les voyelles orales, et hiérarchiser la contribution de ces dernières à l’identification Nord/Sud. Une question se pose dès lors : quelle est la fiabilité de l’estimation automatique des formants des voyelles ?



Corroborant l'étude à base d'alignement, l'antériorisation du /ɔ/ dans la moitié nord de la France (avec des valeurs de F2 supérieures à 1 100 Hz pour les hommes et 1 200 Hz pour les femmes) est de loin le trait le plus discriminant d'après des arbres de décision appliqués sur les formants des voyelles orales. Chaque locuteur étant représenté par les coordonnées moyennes dans le plan F1/F2 de ses voyelles orales, ce trait de prononciation donne à lui seul 73 à 93 % d'identifications correctes Nord/Sud sur ces corpus. Il rivalise avec les prononciations mieux connues du schwa et des voyelles nasales dans le Sud, si l'on ajoute pour chaque locuteur le pourcentage d'antériorisation du /ɔ/, le taux de schwas et d'appendices nasaux comme dans le tableau 7. Le F2 du /ɔ/ suffit même à l'algorithme d'arbre de décision pour classer correctement (entre Nord et Sud) 32 locuteurs sur 33 parmi les plus jeunes du corpus PFC. Nous ne reproduisons pas ici les arbres de décision, car avec seulement deux branches ils sont peu informatifs.

Un travail est en cours pour essayer de caractériser l'accent suisse, si bien identifié par les auditeurs lors de nos expériences perceptives. Les premiers résultats semblent indiquer une tendance à l'aperture du /e/ dans le mot *et* en particulier, qui serait plus discriminante que des indices prosodiques tels que la tendance à l'accentuation initiale. Cependant, nos données sont trop déséquilibrées pour en tirer des conclusions à ce stade du travail.

## 5. Conclusion

En définitive, les mêmes questions se posent, au moins en partie, pour les accents étrangers et régionaux. Dans cette étude, nous avons examiné dans quelle mesure des auditeurs humains et des algorithmes de classification sont capables d'identifier des accents étrangers et régionaux en français. Pour ce faire, nous avons combiné les résultats d'une série d'expériences perceptives avec des techniques de traitement automatique de la parole (en particulier l'alignement en phonèmes) et d'analyse/fouille de données pour déterminer les indices qui contribuent à identifier l'origine d'un accent donné. Nous avons pu distinguer six accents étrangers et deux grands accents régionaux, sans que le style de parole ne semble affecter les résultats du *clustering*.

Sur plusieurs centaines de locuteurs au total, nous avons évalué l'importance relative des formants de voyelles, de variantes de prononciation et d'indices prosodiques pour identifier des accents allemand, anglais, arabe, espagnol, italien, portugais (ceux avec lesquels nous avons le plus de chances d'être exposés) et des accents régionaux en français. Ainsi, nous avons pu identifier des indices caractéristiques et indépendants du corpus qui permettent de reconnaître correctement différents accents. Il semble que l'information portée par les segments (l'articulation des phonèmes) permette d'identifier l'origine d'un accent en français mieux que la prosodie, qui ne fait pas ressortir des tendances bien tranchées.



Des différences, au demeurant, sont notables entre nos approches des accents étrangers et régionaux. La première tient au succès avec lequel les auditeurs de nos expériences perceptives ont pu discriminer six accents étrangers. Ce sont donc six accents étrangers que nous avons cherché à identifier automatiquement, et de fait nous avons pu mesurer nombre de traits opératoires : d'où l'importance d'algorithmes de sélection et de classification. Les résultats des expériences de classification sont encore en deçà des performances humaines, dans les mêmes conditions. En même temps, les taux d'identification obtenus par validation croisée, supérieurs à 60 %, nous encouragent à espérer des scores meilleurs avec des quantités de données moins limitées. L'instrument de mesure est peut-être à perfectionner, pour affiner les traits et en saisir de nouveaux. Suite aux expériences perceptives sur les accents régionaux, nous nous sommes limités à deux grands accents, ceux du Nord et du Sud, pour les variétés natives de français, d'où il résulte que moins de traits ont été étudiés, mais que ceux-ci sont remarquablement robustes, y compris sur la parole téléphonique. La dernière différence avec les accents étrangers réside en effet dans le type et la taille des corpus d'accents régionaux dont nous disposons. Le volume de données accessible nous a conduits à effectuer d'autres traitements à grande échelle.

D'autres travaux sont nécessaires pour valider ou invalider les diverses tendances dégagées. Un travail est en cours pour mesurer des traits subsegmentaux plus fins comme le VOT (*Voice Onset Time*) des occlusives sourdes, espérant que cette information sera utile en identification automatique pour distinguer les locuteurs allemands, anglais et arabes parlant français. Une comparaison L1/L2 plus générale est également envisagée avec les productions en langue maternelle de nos locuteurs. De même, une étude diachronique en temps réel de l'antériorisation du /ɔ/ est prévue, à travers des enregistrements de la deuxième moitié du XX<sup>e</sup> siècle. Plus généralement, une modélisation des accents est envisagée, à travers le traitement automatique de la parole et des corrélations entre perception et mesures acoustiques. Une extension à d'autres accents (belges, alsacien, etc.) a été ébauchée, ainsi qu'une application à la reconnaissance de la parole, afin de diminuer le taux d'erreurs sur la parole avec accents. La synthèse de la parole mériterait également d'être mise à profit : mieux que des imitateurs humains qui ont trop tendance à caricaturer certains traits, cet outil est un bon instrument de simulation. Enfin, nous espérons que ce travail pourra être utile à l'enseignement du français dans la variété de ses usages.

#### Remerciements

Ce travail a été en partie financé par les projets CNRS TCAN VarCom et ANR PFC-Cor. Nous remercions Noël Nguyen, Jacques Durand, Bernard Laks et Chantal Lyche, qui ont rendu possible cette étude. Nous exprimons également notre gratitude envers celles et ceux qui ont enregistré et transcrit les enquêtes, celles et ceux qui ont donné de leur temps pour les tests perceptifs. Enfin, nous sommes

reconnaisants à Gilles Adda, Éric Bilinski, Jean-Luc Gauvain et Lori Lamel pour leurs conseils et leur aide si précieuse.

## 6. Bibliographie

- Adank P.M., Vowel normalization: a perceptual-acoustic study of Dutch vowels, Thèse de doctorat, Radboud University Nijmegen, 2003.
- Adda-Decker M., Lamel L., « Pronunciation variants across system configuration, language and speaking style », *Speech Communication*, vol. 29, 1999, p. 83–98.
- Adda-Decker M., Boula de Mareuil P., Adda G., Lamel L., « Investigating syllabic structures and their variation in spontaneous French », *Speech Communication*, vol. 46, n° 2, 2005, p. 119–139.
- Armstrong N., Boughton Z., « Identification and evaluation responses to a French accent: some results and issues of methodology », *Revue PArôle*, vol. 5–6, 1998, p. 27–60.
- Armstrong N., Unsworth S., « Sociolinguistic variation in southern French schwa », *Linguistics*, vol. 37, n° 1, 1999, p. 127–156.
- Arslan L.M., Hansen J.H.L., « A study of temporal features and frequency characteristics in American English foreign accent », *Journal of the Acoustical Society of America*, vol. 102, 1997, p. 28–40.
- Bartkova K., Jouvét D., « Foreign accent processing in automatic speech recognition », *SPECOM*. Saint-Pétersbourg. 2004, p. 22–28.
- Bauvois C., « Parle-moi, et je te dirai peut-être d'où tu es », *Revue de Phonétique Appliquée*, vol. 121, 1996, p. 291–309.
- Boula de Mareuil P., Brahimi B., Gendrot C., « Role of segmental and suprasegmental cues in the perception of Magrebian-accented French », *8<sup>th</sup> International Conference on Spoken Language Processing*, Jeju, 2004, p. 341–344.
- Boula de Mareuil P., Vieru-Dimulescu B., « The contribution of prosody to the perception of foreign accent », *Phonetica*, vol. 63, 2006, p. 247–267.
- Boula de Mareuil P., Adda-Decker M., Woehrling C., « Analysis of oral and nasal vowel realisation in northern and southern French varieties », *16<sup>th</sup> International Congress of Phonetic Sciences*, Sarrebruck, 2007, p. 2221–2224.
- Bourdieu P., *Ce que parler veut dire. L'économie des échanges linguistiques*, Paris, Fayard, 1982.
- Calliope, *La parole et son traitement automatique*, Paris, Masson, 1989.
- Carton F., Rossi M., Autesserre D., Léon P., *Les accents des Français*, Paris, Hachette, 1983.
- Chambers J., Trudgill P., *Dialectology*, Cambridge, Cambridge University Press, 2004.
- Clopper C.G., Pisoni D.B., « Some acoustic cues for the perceptual categorization of American English regional dialects », *Journal of Phonetics*, vol. 32, 2004, p. 111–140.
- Crystal D., *A dictionary of linguistics and phonetics*, Malden, Blackwell, 2003.
- Delattre P., *Comparing the phonetic features of English, French, German and Spanish*,

Heidelberg, Julius Groos Verlag, 1965.

- Desrosières A., Thévenot, L., *Les catégories socioprofessionnelles*, Paris, Éditions La Découverte, 2002.
- Disner S.F., « Evaluation of vowel normalization procedures », *Journal of the Acoustical Society of America*, vol. 67, p. 253–261, 1980.
- Durand J., « Phénomènes de nasalité en français du midi. Phonologie de dépendance et sous-spécification », *Recherches linguistiques*, vol. 17, 1988, p. 29–54.
- Durand J., Laks B., Lyche C., « Le projet “Phonologie du Français Contemporain” (PFC) », *La Tribune Internationale des Langues Vivantes*, vol. 33, 2003, p. 3–9.
- Durand J., Lyche C., « Structure et variation dans quelques systèmes vocaliques du français : l’enquête *Phonologie du français contemporain (PFC)* », in A. Coveney et C. Sanders (eds), *Variation et francophonie*, Paris, L’Harmattan, 2004, p. 217–240.
- Fagyal Z., Nguyen N., Boula de Mareüil P., « From *dilation* to coarticulation: is there vowel harmony in French? », *Studies in the Linguistic Sciences*, vol. 32, n° 1, 2002, p. 1–21.
- Flege J.E., « The detection of French accent by American listeners », *Journal of the Acoustical Society of America*, vol. 76, 1984, p. 692–707.
- Flege J.E., Munro M.J., MacKay I.R.A., « Effects of age of second-language learning on the production of English consonants », *Speech Communication*, vol. 16, 1995, p. 1–26.
- Flege J.E., Schirru C., MacKay I.R.A., « Interaction between the native and second language phonetic subsystems », *Speech Communication*, vol. 40, 2003, p. 467–491.
- Freland-Ricard M., « Organisation temporelle et rythmique chez les apprenants étrangers. Étude multilingue », *Revue de Phonétique Appliquée*, vol. 118–119, 1996, p. 61–91.
- Fries S., Deprez C., « L’accent étranger : identification et traitement social en France et aux États-Unis », *Cahiers du français contemporain*, vol. 8, 2003, p. 89–105.
- Gauvain, J.-L., Adda, G., Adda-Decker, M., Allauzen, A., Gendner, V., Lamel, L., Schwenk, H., « Where are we in transcribing French broadcast news? », *9<sup>th</sup> European Conference on Speech Communication and Technology*, Lisbonne, 2005, p. 1665–1668.
- Gendrot C., Adda-Decker M., « Impact of duration on F1/F2 formant values of oral vowels: an automatic analysis of large broadcast news corpora in French and German », *9<sup>th</sup> European Conference on Speech Communication and Technology*, Lisbonne, 2005, p. 2453–2456.
- Grabe E., Low E.L., « Durational variability in speech and the rhythm class hypothesis », in C. Gussenhoven et N. Warner (eds), *Laboratory phonology 7*, Berlin, Mouton de Gruyter, 2002, p. 515–546.
- Guyon I., Elisseeff A., « An introduction to variable and feature selection », *Journal of Machine Learning Research*, vol. 3, 2003, p. 1265–1287.
- Hansen A., « Étude du E caduc — stabilisation en cours et variations lexicales », *French Language Studies*, vol. 4, 1994, p. 25–54.
- Harrington J., Palethorpe S., Watson C., « Does the Queen speak the Queen’s English? », *Nature*, vol. 408, 2000, p. 927–928.

- Heeringa W., Measuring dialect pronunciation differences using Levenstein distance, Thèse de doctorat, Rijksuniversiteit, Groningen, 2004.
- Jilka M., The contribution of intonation to the perception of foreign accent, Thèse de doctorat, Université de Stuttgart, 2000.
- Kumpf K., King R.W., « Foreign speaker accent classification using phoneme-dependent accent discrimination models and comparisons with human perception benchmarks », *5<sup>th</sup> European Conference on Speech Communication and Technology*, Rhodes, 1997, p. 2323–2326.
- Labov W., *Sociolinguistic patterns*, Philadelphia, University of Pennsylvania Press, 1972.
- Lamel, L. et Gauvain, J.-L., « Speech recognition », in R. Milroy (ed.), *The Oxford Handbook of Computational Linguistics*, New York, Oxford University Press, 2003, p. 305–322.
- Landick M., « The mid-vowels in figures: hard facts », *French Review*, vol. 68, n° 1, 1995, p. 88–102.
- Lauret B., Aspects de phonétique expérimentale contrastive : « l'accent » anglo-américain en français. Thèse de doctorat, Université Paris III, 1998.
- Léon P., Léon M., *La prononciation du français*, Paris, Fernand Nathan, 1977.
- Livescu K., Glass J., « Lexical modeling of non-native speech for automatic speech recognition », *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Istanbul, 2000, p. 1683–1686.
- Magen H.S., « The perception of foreign-accented speech », *Journal of Phonetics*, vol. 26, 1998, p. 381–400.
- Malderez I., Contribution à la synchronie dynamique du français contemporain : le cas des voyelles orales arrondies, Thèse de doctorat, Université Paris VII, 1995.
- Manning C.D., Schütze H., *Foundations of statistical natural language processing*, Cambridge, The MIT Press, 1999.
- Martinet A., *La prononciation du français contemporain*, Paris, Droz, 1945.
- Martinet A., « C'est jeuli, le Mareuc ! », in A. Martinet (éd.), *Le français sans fard*, Paris, Presses Universitaires de France, 1969, p. 345–355.
- Ramus F., Rythme des langues et acquisition du langage, Thèse de doctorat, EHESS, 1999.
- Rochet B.L., « Perception and production of second-language speech sounds by adults », in W. Strange (ed.), *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues*, York, York Press, 1995, p. 379–410.
- Rosch E., « Cognitive reference points », *Cognitive Psychology*, vol. 7, 1975, p. 532–547.
- Schaden S., « CrossTowns: Automatically generated phonetic lexicons of cross-lingual pronunciation variants of European city names », *4<sup>th</sup> International Conference on Language Resources and Evaluation*, Lisbonne, 2004, p. 1395–1398.
- Silke G., Stefan R., Ralf K., « Generating non-native pronunciation variants for lexicon adaptation », *Speech Communication*, vol. 42, n° 1, 2004, p. 109–123.
- Sobotta E., Phonologie et migration — Aveyronnais et Guadeloupéens à Paris, Thèse de doctorat, Universités Paris X – Nanterre et LMU de Munich, 2006.

- ten Bosch L., « ASR, dialects and acoustic/phonological distances », *6<sup>th</sup> International Conference on Spoken Language Processing*, Beijing, 2000, p. 1009–1012.
- Thomas A., « Évolution de l'accent méridional en français niçois : les nasales », *XXI<sup>e</sup> Congrès International des Sciences Phonétiques*, Aix-en-Provence, 1991, p. 194–197.
- van Bezooijen R., Gooskens G., « Identification of language varieties. The contribution of different linguistic levels », *Journal of Language and Social Psychology*, vol. 18, n° 1, 1999, p. 31–48.
- Veloso J., « Schwa in European Portuguese: The phonological status of [ɨ] », *5<sup>es</sup> Journées d'Études Linguistiques*, Nantes, 2007, p. 55–60.
- Vieru-Dimulescu B., Boula de Mareüil P., « Perceptual identification of 6 foreign accents in French », *9<sup>th</sup> International Conference on Spoken Language Processing*, Pittsburgh, 2006, p. 411–414.
- Vieru-Dimulescu B., Boula de Mareüil P., Adda-Decker M., « Characterizing non-native French accents using automatic alignment », *16<sup>th</sup> International Congress of Phonetic Sciences*, Sarrebruck, 2007, p. 2217–2220.
- Vieru-Dimulescu B., Boula de Mareüil P., Adda-Decker M., « Identification of foreign-accented French using data-mining techniques », *International Workshop on Paralinguistic Speech*, Sarrebruck, 2007, p. 47–52.
- Walter H., *Enquête phonologique et variétés régionales du français*, Paris, Presses Universitaires de France, 1982.
- Witten I.H., Frank E., *Data mining: Practical machine learning tools and techniques*, San Francisco, Morgan Kaufmann, 2005.
- Woehrling C., Boula de Mareüil P., « Identification d'accents régionaux en français : perception et analyse », *Revue PArole*, vol. 37, 2006, p. 25–65.
- Woehrling C., Boula de Mareüil P., « Comparing Praat and Snack formant measurements on two large corpora of northern and southern French », *8<sup>th</sup> Annual Meeting of the International Speech Communication Association*, Anvers, 2007, p. 1006–1009.
- Yan Q., Vaseghi S., « A comparative analysis of UK and US English accents in recognition and synthesis », *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Orlando, 2002, p. 413–416.