



HAL
open science

Rôle de la prosodie dans la fluence en lecture oralisée chez des apprenants de Français Langue Étrangère

Charlotte Alazard

► **To cite this version:**

Charlotte Alazard. Rôle de la prosodie dans la fluence en lecture oralisée chez des apprenants de Français Langue Étrangère. Linguistique. Université Toulouse le Mirail - Toulouse II, 2013. Français. NNT : 2013TOU20057 . tel-00944968

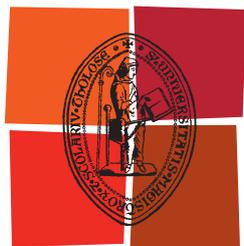
HAL Id: tel-00944968

<https://theses.hal.science/tel-00944968v1>

Submitted on 11 Feb 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Université
de Toulouse

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Université Toulouse 2 Le Mirail (UT2 Le Mirail)

Présentée et soutenue par :

Charlotte Alazard

Le vendredi 27 septembre 2013

Titre :

Rôle de la prosodie dans la fluence en lecture oralisée chez des apprenants de
Français Langue Étrangère

École doctorale et discipline ou spécialité :

ED CLESCO : Sciences du langage

Unité de recherche :

Unité de recherche interdisciplinaire OCTOGONE (EA 4156)

Directeur(s) de Thèse :

Michel Billières, PR, Université de Toulouse 2 - Le Mirail
Corine Astésano, MCF, Université de Toulouse 2 - Le Mirail

Rapporteurs :

Lorraine Baqué, PR, Université Autonome de Barcelone, Espagne
Bernard Harmegnies, PR, Université de Mons, Belgique

Autre(s) membre(s) du jury :

Jena-Luc Nespoulous, PR émérite, Université de Toulouse 2 - Le Mirail
Bruno Maurer, PR, Université de Montpellier 3 - Paul Valéry

Rôle de la prosodie dans la fluence en
lecture oralisée chez des apprenants de
Français Langue Étrangère

à Jean-Philippe,

Remerciements

Le travail de thèse est une entrée dans le monde de la recherche autant qu'un travail sur soi-même. Arrivée à la fin de ce parcours, et je l'espère au début de nouvelles aventures, je voudrais remercier toutes les personnes qui de près ou de loin ont contribué à l'aboutissement de ce travail.

Tout d'abord, je tiens à remercier chaleureusement Lorraine Baqué et Bernard Harmegnies d'avoir accepté d'être les rapporteurs de cette thèse. Nous nous sommes rencontrés au tout début de ma carrière de chercheuse et je suis fière de vous présenter aujourd'hui l'aboutissement de ce travail. Je remercie également Jean-Luc Nespoulous et Bruno Maurer d'avoir accepté de faire partie de mon jury de thèse. C'est un grand honneur pour moi que vous ayez accepté de juger mon travail.

Je tiens à remercier tout particulièrement mes deux directeurs de thèse, Michel Billières et Corine Astésano. Vous m'avez ouvert les portes d'un monde que je ne connaissais pas, vous m'avez poussée à m'affirmer en tant que chercheuse et je vous en suis très reconnaissante. Je me souviens de ma rencontre avec la Méthode Verbo-Tonale, en licence, et ce fut une révélation ! Merci à vous deux d'avoir cru en moi et de m'avoir aidée à concrétiser ce travail de recherche.

J'ai réalisé cette thèse au sein de l'unité interdisciplinaire Octogone, à l'Université de Toulouse Le Mirail et je souhaiterais remercier tous les membres de l'unité. Je tiens à remercier en particulier Michèle Guidetti, Barbara Köpke et Christiane Soum qui m'ont soutenue tant humainement que financièrement, Mélanie Jucla, pour ses conseils avisés, Pascal Gaillard, pour son aide pour les tests de perception, Vanda Marijanovic, pour les discussions FLE, Halima Sahraoui, pour les futurs projets sur l'enseignement de la prononciation et Julien Tardieu pour m'avoir aidée pendant les enregistrements des corpus de lecture oralisée. Au cours de mes années de thèse, j'ai eu la chance de présenter mes travaux à de nombreuses conférences, d'élaborer des projets de recherches

interdisciplinaires et de participer à la vie du laboratoire. Je remercie le laboratoire de m'avoir donné ces opportunités.

Je n'aurais pas pu réaliser cette recherche sans la participation d'apprenants anglophones qui ont accepté de prêter leur corps à la science et je tiens donc à remercier tous les sujets qui ont participé à mes études ainsi que les personnes qui m'ont aidé à les recruter, Chantal Domp martin et Cécilia Gunnarsson pour le Département d'Enseignement du FLE, Bozena Billerey et Annie Despatureau de l'Institut Catholique et toute l'équipe de Langue Onze. Je souhaite également remercier toute l'équipe de la Direction des Technologies de l'Information et le Communication pour l'Enseignement (DTICE) de l'Université du Mirail, qui ont filmé l'ensemble de nos cours et nous ont permis de constituer le corpus MULTIPHONIA. Je remercie tout particulièrement Bruno Bastard qui n'a ménagé ni son temps ni sa peine avec beaucoup de générosité.

Durant mes années de doctorat, j'ai eu l'occasion de séjourner plusieurs semaines au Laboratoire Parole et Langage d'Aix en Provence. Je tiens à remercier Noël Nguyen et Christine Meunier pour l'accueil chaleureux que j'ai reçu, Robert Espesser qui m'a fait découvrir les modèles linéaires mixtes en statistique, Brigitte Bigi pour SPPAS et les confitures, et Marion Tellier, la « FLE » du LPL, qui m'a permis d'ouvrir ce travail à de nouvelles perspectives. Je tiens également à remercier l'ensemble des doctorants et des jeunes docteurs du LPL, et en particulier Amandine Michelas et Francesco Cangemi pour leur gentillesse et leur accueil.

Au cours de ces années de thèse, beaucoup de doctorants m'ont apporté leur soutien et je les remercie tous du fond de mon cœur. Plus que des collègues, j'ai gagné parmi eux de véritables amis et je les remercie pour les fous rires, les conseils et les sessions de Partage Social des Émotions (PSE). Je tiens tout particulièrement à remercier Marie Lacabanne, Émilie Massa, Lionel Fontan, Nadia Boutechkil, Nicole Cantisano, Carine De Martin, Karine Aura et Aurore Berland, ainsi que deux jeunes docteurs du laboratoire CLLE-ERSS, Clémentine Adam et Marianne Verget-Couret.

Enfin, je tiens à remercier ma famille et mes amis, qui ont su me soutenir et m'entourer de tout leur amour. Je tiens tout d'abord à remercier mes parents, mes soeurs et mes grands-parents, qui m'ont toujours soutenue dans mes études, même s'ils ne voyaient pas trop où ça allait me mener. Giles Tilley, ma tante de cœur, et ma belle-famille qui ont su me guider et m'épauler dans les moments de doute. Mes amis, Camille et Jérémie, Samantha, David, Magali et Christophe, Marie, Mathieu, Sindia et Sophie qui m'ont toujours soutenue et qui ont su me pardonner mes absences. Merci de rendre la vie toujours plus belle ! Enfin (et surtout) je remercie Jean-Philippe, mon futur mari, qui a été mon roc pendant toutes ces années, qui a su m'entourer de beaucoup d'amour et apaiser mes peurs et mes doutes.

Jp, tu es ma force tranquille, tu sais me faire rire mieux que personne et c'est donc tout naturellement que je te dédie cette thèse.

Table des matières

Table des figures.....	1
Table des tableaux.....	7
Indices de lecture	9
<u>INTRODUCTION.....</u>	11
<u>1^{ERE} PARTIE : CADRES THEORIQUE ET METHODOLOGIQUE</u>	19
<u>CHAPITRE 1 : PROBLEMATIQUE DE L'ENSEIGNEMENT DE L'ORAL</u>	21
1.1 INTRODUCTION.....	25
1.1.1 Introduction au concept d'oral	25
1.1.2 L'oral dans la représentation des enseignants.....	28
1.2 DE L'ENSEIGNEMENT DE L'ORAL A L'ENSEIGNEMENT DE LA PRONONCIATION	31
1.2.1 Des méthodes traditionnelles à l'éclectisme : quelle place pour l'enseignement de l'oral ?	31
1.2.2 Présupposés à l'enseignement de la prononciation.....	35
1.2.3 Le rôle de la perception.....	44
1.3 COMPARAISON DE DEUX METHODES DE CORRECTION PHONETIQUE.....	48
1.3.1 La méthode articulatoire.....	49
1.3.2 La Méthode Verbo-Tonale d'intégration phonétique	54
1.4 SYNTHESE DU CHAPITRE 1	63
<u>CHAPITRE 2 : RYTHME ET APPRENTISSAGE D'UNE LANGUE ETRANGERE</u>	67
2.1 INTRODUCTION.....	71
2.1.1 La place du rythme dans la prosodie	72
2.1.2 Les constituants du rythme	75
2.2 RYTHME ET LANGUES.....	80
2.2.1 Principe d'isochronie : rythmicité syllabique et rythmicité accentuelle	80
2.2.2 Différences rythmiques du français et de l'anglais	84
2.3 CONSEQUENCES POUR L'APPRENTISSAGE DES LANGUES ETRANGERES	87
2.3.1 Transferts rythmiques et interlangue de l'apprenant.....	88
2.4 SYNTHESE DU CHAPITRE 2	91

CHAPITRE 3 : LA FLUENCE COMME MESURE DE LA COMPETENCE RYTHMIQUE EN LANGUE ETRANGERE 93

3.1	INTRODUCTION.....	97
3.2	LA FLUENCE EN LANGUE ETRANGERE	98
3.2.1	Première difficulté : définir la fluence en langue étrangère	99
3.2.2	Deuxième difficulté : mesurer la fluence.....	103
3.3	ÉCLAIRAGES COGNITIFS.....	107
3.3.1	Modélisation de la fluence dans une approche cognitive	107
3.3.2	La fluence comme mesure de la compétence rythmique ?	112
3.4	SYNTHESE DU CHAPITRE 3	113

CHAPITRE 4 ASPECTS METHODOLOGIQUES..... 115

4.1	OBJECTIFS DE LA THESE	119
4.1.1	Existe-t-il des transferts de compétences entre l'encodage de l'oral et de l'écrit en langue étrangère ? Quel rôle joue la prosodie dans ces transferts ?.....	120
4.1.2	Qu'est-ce que la compétence prosodique ? Comment l'enseigner ?	125
4.1.3	Hypothèses de recherche	128
4.2	ASPECTS METHODOLOGIQUES	130
4.2.1	Apport des études longitudinales à la didactique	130
4.2.2	Présentation des sujets constituer des groupes homogènes est-ce possible ?	131
4.2.3	Dessin expérimental des études	134
4.2.4	Comment mesurer la fluence de l'interlangue ?	141
4.2.5	Présentation du corpus MULTIPHONIA	150
4.3	SYNTHESE DU CHAPITRE 4.....	153

2^{EME} PARTIE : ANALYSES EXPERIMENTALES 155

CHAPITRE 5 : RESULTATS ETUDE LONGITUDINALE PILOTE ET ETUDE LONGITUDINALE 2 157

5.1	RESULTATS ETUDE 1	161
5.1.1	Résultats de l'analyse acoustique	161
5.1.2	Résultats de l'analyse perceptive.....	167
5.1.3	Discussion	169
5.2	RESULTATS ÉTUDE 2	173
5.2.1	Résultats de l'analyse acoustique	173

5.2.2	Résultats de l'analyse perceptive	189
5.2.3	Discussion	193
CONCLUSION ET PERSPECTIVES		197
BIBLIOGRAPHIE		205
ANNEXES		225
Annexe A	: Articles publiés sur les deux études longitudinales réalisées	227
Annexe B	: Exemple de test de niveau (étude 2)	235
Annexe C	: Niveaux communs de référence et de compétence du Cadre Européen Commun de Référence (CECR)	237
Annexe D	: Exemplaire du questionnaire donné aux participants de l'étude 2 (Auto-évaluation formative)	241
Annexe E	: Exemplaire du questionnaire bilan donné aux participants de l'étude 2 (Auto-évaluation sommative)	243
Annexe F	: Textes utilisés pour mesurer la fluence en lecture oralisée (étude 2).....	245
Annexe G	: Exemplaire du questionnaire donné aux auditeurs pour le test de perception (étude 2)	251
Annexe H	: Article publié sur la constitution du corpus MULTIPHONIA.....	253

Table des figures

Figure 1 Grille du ciep 2011 pour l'évaluation de l'oral pour le niveau A2	30
Figure 2 Schéma de la communication représentant les 4 compétences communicatives : la composante linguistique (être capable de faire des phrases), la composante discursive (être capable d'utiliser des phrases), la composante socio-culturelle (être capable de parler aux gens) et la composante référentielle (être capable de parler du monde) (d'après Defays, 2003)	33
Figure 3 Représentation des interactions entre les voyelles de la langue maternelle cercles gris (L1) et les voyelles de la langue étrangère cercles blancs (L2) dans le trapèze vocalique (en abscisse, fréquence du deuxième formant vocalique et en ordonnée, fréquence du premier formant). Les assimilations sont représentées par des flèches grises et les dissimilitons, entraînant la création de nouvelles catégories, par les flèches noires (Flege, 2005, http://jimflege.com/files/Vancouver_April_2005.pdf).....	38
Figure 4 Distribution des scores d'accentuation (grisé) et de compréhensibilité (blanc) donné par un groupe d'auditeurs sur un échantillon de phrases (Derwing & Munro, 1997 ; Munro, 2008 : 204)	42
Figure 5 cette figure montrant la mise en place progressive des frontières pour les voyelles du suédois, de l'anglais et du japonais (Kuhl, 1998 : 61)	45
Figure 6 Illustration de la différence entre réalité physique (le continuum sonore) et réalité perceptive (réorganisation du continuum avec les prototypes) pour l'anglais et le japonais (Kuhl, 1998 : 59).....	46
Figure 7 Exemple d'exercice articulatoire proposé pour la production des voyelles /i/ et /y/ dans un manuel (Charliac & Motron, 1998 : 46)	49
Figure 8 Exemple d'exercice articulatoire proposé pour la production des consonnes /z/ et /ʒ/ dans un manuel (Charliac & Motron, 1998 : 150)	50
Figure 9 Vue d'ensemble des principes de la méthode articulatoire (Billière, 2000 : 138).....	51

Figure 10 logatomes proposés filtrés (série A) et non filtrés (série B) (Renard, 1979:47).....	55
Figure 11 Exemple d'un appareil Suvag et ses accessoires http://www.suvag.com/fr/produits/suvag_it2.html	56
Figure 12 Représentation schématique des distances autour d'une voyelle non prototype et d'une voyelle prototype (représentées en noir) (Nguyen, 2005 : 8)	57
Figure 13 Classement des voyelles du français sur l'axe de la tension.....	60
Figure 14 Classement des consonnes du français sur l'axe clair-sombre.....	61
Figure 15 Représentation par un arbre phonologique de la structure prosodique d'un énoncé français, UI : Unité Intonative ; UR : Unité Rythmique; UT : Unité Tonale ; S : Syllabes (Di Cristo & Hirst, 1993 : 6)	75
Figure 16 Continuum rythmique (Dauer, 1983 : 60).....	82
Figure 17 Exemples des gabarits tonals du français et de l'anglais	84
Figure 18 Représentation d'une courbe rythmique (Wenk, 1983: 157)	85
Figure 19 Schéma des trois types de fluence d'après la description proposée par Segalowitz, 2010	108
Figure 20 représentation du modèle neurologique classique de la lecture (en haut) en comparaison au schéma parallèle et foisonnant (en bas) qui lui est aujourd'hui préférée (Dehaene, 2007 : 97)	123
Figure 21 Texte argumentatif utilisé pour cette étude.....	137
Figure 22 Exemple de texte utilisé pour le niveau débutant. Les cinq autres textes sont présentés en annexe (Annexe F).....	140
Figure 23 Exemple de phrase du corpus annotée manuellement sur Praat. On peut distinguer 4 niveaux d'annotation, les phrases (S1), les pauses (# : pauses vides, #* : pauses respiratoires, #AG pauses agrammaticales), les syllabes	

transcrites en SAMPA et les accents (AI : Accent Initial, AFmin : Accent Final mineur, AFmaj : Accent Final majeur).	144
Figure 24 Exemple de phrases extraites du corpus des débutants transcrites suivant les conventions de la TOE	145
Figure 25 Exemple de phrase du corpus alignée automatiquement puis manuellement sous Praat. On peut distinguer 7 niveaux d'annotation, les phrases (S1), les transcriptions via la TOE, les pauses (# : pauses vides), les tokens lexicaux, les syllabes transcrites orthographiquement selon les conventions de SPASS, les accents (A) et les niveaux de frontières (Fmin : accent final mineur, Fmaj : accent final majeur).	146
Figure 26 Extrait d'un cours de MVT avec le groupe débutant	151
Figure 27 Extrait d'un cours d'articulatoire niveau débutant	152
Figure 28 Durée (log ms) des pauses et hésitations vocales avant (before)/après (after) entraînement pour l'apprenant débutant du groupe test (MM) (# VAG = pauses silencieuses agrammaticales; hesvoc = hésitations vocales ; # = pauses vides; #* = pauses respiratoires)	164
Figure 29 Durée (log ms) des pauses et hésitations vocales avant (before)/après (after) entraînement pour l'apprenant débutant du groupe contrôle (SR) (# VAG = pauses silencieuses agrammaticales; hesvoc = hésitations vocales ; # = pauses vides; #* = pauses respiratoires)	164
Figure 30 Durée (log ms) des pauses avant (before)/après (after) entraînement pour l'apprenant avancé du groupe test (DR) (# = pauses vides ; #* = pauses respiratoires)	165
Figure 31 Durée (log ms) des pauses avant (before)/après (after) entraînement pour l'apprenant avancé du groupe contrôle (AM) (# = pauses vides ; #* = pauses respiratoires)	165
Figure 32 Moyenne des scores avant/après entraînement par groupe (ligne pointillée Groupe Test; ligne continue Groupe Contrôle	167

Figure 33 Histogrammes de l'effectifs des différents types de pauses (# : pauses vides ; #* : pauses respiratoires ; #AG : pauses agrammaticales ; #*AG : pauses respiratoires agrammaticales) et disfluences (FD : faux-départs ; Hes : hésitations ou pauses sonores ; Hes-lex : hésitations lexicales ; RE : répétions ; RA : reparans) par groupe (MVT ou ART) et par niveau (avancé ou débutant) en fonction des trois stades (en rouge s0 : avant, en vert s1 : pendant et en bleu s2 : après entrainement).....	174
Figure 34 Histogramme lissé de la densité de distribution des pauses agrammaticales (pag) et standard selon la méthode (MVT vs. ART) , le groupe de niveau (debutant vs.avancé) et le stade d'apprentissage (en rouge, s0, en vert s1 et en bleu s2)	175
Figure 35 Histogramme du nombre de pauses agrammaticales pour les locuteurs débutants du groupe MVT en fonction du stade d'apprentissage : s0 (en bleu), s1 (en rouge) et s2 (en vert).	177
Figure 36 Histogramme du nombre de pauses agrammaticales pour les locuteurs débutants du groupe ART en fonction du stade d'apprentissage : s0 (en bleu), s1 (en rouge) et s2 (en vert).	177
Figure 37 Boxplot de la durée des pauses (pag et pstand confondues) par niveau (avancé et débutant) et par méthode (ART et MVT) dans les trois stades (s0, s1 et s2).....	178
Figure 38 histogramme lissé de la densité de distribution des syllabes accentuées (A) et inaccentuées (IA) selon la méthode (MVT vs. ART) et les stades d'apprentissage (s0, s1 et s2).....	180
Figure 39 Boxplot des durées syllabiques en milisecondes, selon la situation accentuelle (A : syllabes accentuées ; INA : syllabes inaccentuées), le niveau (avancé vs. débutant), la méthode (ART vs MVT) et le stade de l'apprentissage (en blanc s0, en gris clair s1 et en gris foncé s2)	181

Figure 40 Moyenne et intervalle de confiance du nbre de syllabe entre 2 pauses , selon la methode (en blanc : ART ; en gris : MVT) , le niveau (avancé vs. débutant) et le stade de l'apprentissage (s0, s1 et s2)	185
Figure 41 Boxplot du taux d'articulation (exprimé en nombre de syllabes/secondes) selon la méthode (ART vs. MVT) le niveau (avancé vs. débutant) et le stade de l'apprentissage (en blanc s0, en gris clair s1 et en gris foncé s2)	187
Figure 42 Boxplot de la durée des IPU en secondes, selon la méthode (ART vs. MVT) le niveau (avancé vs. débutant) et le stade de l'apprentissage (en blanc s0, en gris clair s1 et en gris foncé s2)	188
Figure 43 Moyenne des scores en fonction de la méthode (en rouge ART ; en bleu MVT) et du stade d'apprentissage (s0, s1 et s2).....	190
Figure 44 Pente des scores obtenus en s1 par rapport à s0 par méthode (en rouge ART ; en bleu MVT)	191
Figure 45 Mesures des temps de réactions en fonction de la méthode (ART : en continu ; MVT : en pointillé) du niveau (AT : apprenants avancés ; DT : apprenants débutants) et des scores obtenus.....	192

Table des tableaux

Tableau 1 Les quatre savoir-faire visés en didactique des langues	34
Tableau 2 Représentation des paramètres locaux et globaux de la prosodie sur le plan acoustique et perceptif, les critères marqués par * ne s'appliquent pas au français	73
Tableau 3 Moyenne des durées syllabiques et des pauses dans les trois conditions de débit, (Pasdeloup, 2004 : 2).....	79
Tableau 4 Classification des langues selon leur classe rythmique (Tortel, 2009 : 44)	81
Tableau 5 Caractéristiques des langues Trailerd-timed et des langues leader-timed (Wenk, 1983 : 156)	85
Tableau 6 Mesures de la fluidité inspiré des 12 critères établis par Lennon (1990) (d'après Kormos, 2006)	101
Tableau 7 Phrases utilisées pour le test de perception	148
Tableau 8 : Nombre et Durée (log ms) des paramètres prosodiques pertinents par locuteur, avant/après entraînement (t0 = avant entraînement ; t1 = après entraînement). Les éléments significatifs sont en gras et marqués par *	162
Tableau 8 Moyenne des scores avant (t0) et après entraînement (t1) pour chaque locuteur. La significativité est indiquée par une ou plusieurs astérisques *.....	168

Indices de lecture

Les abréviations suivantes seront utilisées de façon récurrente dans ce manuscrit afin de rendre la lecture plus claire et plus synthétique.

LM	Langue Maternelle
LE	Langue Étrangère
FLE	Français Langue Étrangère
CECR	Cadre Européen Commun de Référence pour les Langues
MVT	Méthode Verbo-Tonale d'intégration phonétique
SGAV	Méthodologie Structuro Globale Audio-Visuelle
ART	Méthode Articulatoire
GR	Groupes Rythmiques
IPU	Unité Inter Pausale (Inter Pausal Unit)

INTRODUCTION

Cette recherche est le fruit d'aller-retour entre réflexions de terrain et éclairages théoriques pluridisciplinaires, en didactique, en phonétique et en psycholinguistique. En effet, notre expérience dans l'enseignement nous permet de penser qu'enseigner une langue étrangère ne peut se faire sans questionnements sur sa langue maternelle (LM), sur la langue étrangère (LE) mais aussi et surtout sur les différents processus d'apprentissage et les compétences mises en jeu dans l'acquisition et l'apprentissage d'une langue.

La didactique des langues étrangères a beaucoup évolué durant les dernières décennies. La communication orale s'est progressivement imposée comme une priorité de l'enseignement. Pourtant, l'enseignement de la prononciation est, paradoxalement, au pire négligé, au mieux considéré comme un perfectionnement optionnel, jamais comme une priorité. Tout se passe comme si la maîtrise des sonorités de la LE dépendait seulement d'un contact répété avec la langue cible. Ainsi, pour le Cadre Européen Commun de Référence pour les Langues (CECR), un apprenant de niveau B2 (Avancé) devra avoir acquis « *une prononciation et une intonation claires et naturelles* », alors que la question de la prononciation n'est pas abordée aux niveaux inférieurs.

Nos expériences d'enseignement, autant en France qu'à l'étranger, vont à l'encontre de ces présupposés : la majorité des apprenants de Français Langue Étrangère (FLE) de niveau B2 n'ont pas perdu leur accent étranger. En outre, on retrouve les mêmes erreurs de prononciation en parole spontanée et en lecture oralisée, non seulement au niveau segmental (phonèmes) mais aussi et surtout au niveau suprasegmental (accentuation, rythme et intonation). Il est intéressant de noter que l'élocution des apprenants est beaucoup moins fluide en lecture qu'en oral spontané. Pourtant, la lecture est souvent considérée comme une tâche de pratique orale simplifiée car le lecteur n'est pas en situation de planification du discours et le phrasé est en partie indiqué par la ponctuation.

Nos observations de classes qui remettent en question ce postulat nous ont amenées à interroger les compétences mises en jeu dans l'encodage et le

décodage de la langue : la maîtrise de l'oral est-elle un prérequis à l'apprentissage de la lecture en langue étrangère ? Existe-il des transferts de compétences entre l'oral et l'écrit ? Autrement dit, est-ce que maîtriser la fluence à l'oral permet d'améliorer l'encodage et le décodage de l'écrit ?

De récentes études en LM soutiennent cette idée. Fodor (2002) a montré que la prosodie serait utilisée en parole spontanée comme en lecture. Ainsi, même en lecture silencieuse, un contour prosodique par défaut serait projeté sur le texte écrit afin de découper ce dernier en groupes prosodiquement et syntaxiquement cohérents. Autrement dit, le lecteur ne partirait pas du texte mais ferait appel à des modèles prosodiques établis et mémorisés par la pratique orale. Si un parallèle entre LM et LE peut être établi, il est possible de penser que l'apprenant de LE projèterait les contours prosodiques de sa LM sur le texte de LE, handicapant ainsi considérablement le décodage et la compréhension du texte écrit. Le résultat serait une chute de la fluence en lecture oralisée pour l'apprenant de LE.

Des recherches développementales récentes en sciences cognitives, utilisant l'imagerie cérébrale, apportent une vision complémentaire à ce mécanisme (Dehaene, 2007). En effet, les caractéristiques anatomiques et fonctionnelles du cerveau indiqueraient un transfert possible des compétences auditives sur les compétences d'écriture et de lecture. Les aires visuelles et auditives sont non seulement interconnectées *avant* que l'enfant n'apprenne à lire, mais l'aire de Wernicke – où la compréhension du langage écrit et oral est traitée – est localisée à la croisée des aires auditives et visuelles. Par conséquent, tout processus linguistique appris *avant* le développement des compétences en lecture et en écriture est susceptible de 'nourrir' la modalité visuelle.

On sait que la prosodie est première dans l'ontogénèse ((Konopczynski, 1986 ; de Boisson-Bardies, 1996 ; Molino, 2000 ; Billières, 2002 ; Karmiloff-Smith & Karmiloff-Smith, 2003), se pourrait-il que la prosodie constitue un préalable nécessaire à l'apprentissage de la lecture en langue étrangère ?

Le but de ce travail est donc d'évaluer le rôle de la prosodie dans l'apprentissage d'une langue étrangère, autant à l'oral qu'en lecture. Nous pensons en effet, que les erreurs observées en parole lue et en parole spontanée dans nos pratiques de classes pourraient être dues à la non-maitrise de la prosodie de la langue cible.

Premièrement, notre recherche a pour objectif de tester l'incidence de la pratique linguistique à l'oral sur les compétences en lecture d'apprenants de FLE. Nous postulons que la fluence des apprenants étrangers, à la fois en oral spontané et en lecture oralisée, va dépendre de leur capacité à encoder la prosodie de la langue-cible. L'entraînement des paramètres prosodiques devrait avoir une incidence positive sur la fluence des apprenants et donc sur le décodage du texte écrit (*Transferts de compétences de l'oral vers l'écrit*). De plus, si les processus cognitifs qui sous-tendent l'apprentissage et l'acquisition du langage sont comparables, l'impact de l'entraînement linguistique devrait être plus important au début de l'apprentissage. Plus tôt cette pratique prosodique sera mise en place dans l'apprentissage (niveau A1 du CECR), meilleure devrait être la fluence des productions des apprenants.

Deuxièmement, nous souhaitons valider expérimentalement l'efficacité de la Méthode Verbo-Tonale (MVT) de correction phonétique, indissociable de la méthodologie audio-visuelle structuro-globale (SGAV). Dans cette méthodologie, la prosodie est un préalable à l'apprentissage des autres niveaux linguistiques et l'écrit est introduit après plusieurs dizaines d'heures exclusivement consacrées à l'apprentissage de l'oral. Cette méthode donne des résultats empiriques très prometteurs, notamment en orthophonie, mais n'a, à notre connaissance, jamais été validée expérimentalement.

Pour répondre à ces objectifs, nous avons mis en place deux études longitudinales sur une durée de huit semaines chacune avec des apprenants anglophones de FLE. Nous avons comparé la fluence des apprenants en parole lue à différents moments de l'apprentissage.

Dans l'étude pilote, nous avons voulu tester l'incidence d'un entraînement à la correction phonétique *via* la MVT dans l'amélioration de la fluence en lecture oralisée. Nous avons formulé deux hypothèses : l'hypothèse selon laquelle un apprentissage à l'oral centré sur la prosodie de la langue étrangère va améliorer la fluence des apprenants en parole lue (Hypothèse 1) et l'hypothèse selon laquelle les résultats seront meilleurs si l'apprentissage de la prosodie est proposé dès le niveau débutant (Hypothèse 2). Pour tester ces hypothèses, nous avons comparé des cours de correction phonétique *via* la MVT (*groupe test*) et des cours de FLE plus traditionnels, dans une approche communicative (*groupe contrôle*).

Afin d'approfondir les résultats de l'étude préliminaire et de valider l'efficacité de la MVT par rapport à d'autres méthodes de correction phonétique, nous avons comparé l'impact de la MVT et de la Méthode Articulaire dans la maîtrise de la fluence des apprenants. Nous avons également testé un des présupposés de la méthodologie SGAV selon lequel l'écrit ne devrait pas être introduit avant que les paramètres prosodiques ne soient maîtrisés. Ainsi, nous avons formulé deux nouvelles hypothèses : l'hypothèse selon laquelle seule la MVT aurait une incidence positive sur l'amélioration de la fluence en lecture oralisée des apprenants de LE. (Hypothèse 3) et l'hypothèse qui voudrait que l'introduction précoce de l'écrit (i.e. avant que la compétence rythmique ne soit installée) entraînerait une chute de la fluence (Hypothèse 4). Les séances de cours de la deuxième étude ont toutes été filmées afin de constituer le corpus MULTIPHONIA (96h d'enregistrements vidéo d'enseignement) qui sera prochainement mis à la disposition des enseignants et des chercheurs en FLE (Alazard *et al*, 2011).

Avant de décrire plus en détails chacune des deux études, le corpus MULTIPHONIA, et de présenter nos résultats, nous développerons les liens entre l'enseignement de l'oral, la prosodie et la fluence dans une première partie théorique.

Cette première partie, « *Partie 1 : Cadres théorique et méthodologique* », se compose de quatre chapitres :

- Dans le premier chapitre théorique, nous reviendrons sur la problématique de l'enseignement de l'oral. Nous discuterons les raisons qui ont provoqué l'abandon de l'enseignement de la prononciation puis nous présenterons deux méthodes de correction phonétique, la Méthode Verbo-Tonale et la Méthode Articulaire.
- Dans le deuxième chapitre théorique, nous illustrerons le rôle de la prosodie et plus précisément du rythme - noyau de la prosodie - dans l'apprentissage des langues. Nous parlerons des transferts rythmiques entre la LM et la LE. Nous terminerons cette partie par une réflexion sur la notion d'interlangue.
- Le troisième chapitre de cette partie théorique nous permettra de recentrer notre cadre théorique sur la notion de fluence qui se trouve être au cœur de la problématique de l'évaluation de l'oral. À partir d'une modélisation cognitive de la fluence, nous postulerons que la fluence pourrait constituer un critère de mesure de la compétence rythmique (ou plus globalement prosodique) des apprenants de LE.
- Enfin, dans notre quatrième chapitre méthodologique, nous développerons notre problématique et nos hypothèses de recherche. Nous présenterons le cadre méthodologique relatif aux deux études longitudinales que nous avons réalisées. Nous reviendrons sur les difficultés que nous avons rencontrées et les choix méthodologiques effectués. Nous clôturerons ce chapitre, par la présentation du corpus MULTIPHONIA que nous avons constitué.

La deuxième partie de la thèse, *Partie 2 : Analyses expérimentales*, sera l'occasion de présenter et de discuter les résultats des analyses que nous avons réalisées. Nous avons proposé deux types d'analyses complémentaires : des analyses quantitatives de la fluence produite (*i.e.* mesures acoustiques du nombre et de la durée des pauses, des syllabes accentuées et inaccentuées, des disfluences, du débit de parole, du taux d'articulation et de la durée moyenne des groupes rythmiques) et des analyses qualitatives de la perception des productions des apprenants par des natifs experts (enseignants ou futurs enseignants de FLE).

- Dans une première sous-partie, nous présenterons les résultats des analyses acoustiques et du test perceptif réalisé avec douze sujets de l'étude pilote. Les résultats nous permettront de vérifier l'existence de transferts de compétences entre le décodage de l'oral et de l'écrit en langue étrangère et de valider nos choix méthodologiques.
- Dans une deuxième sous-partie, nous présenterons les résultats des analyses acoustiques et du test perceptif réalisé avec vingt-deux sujets de la deuxième étude longitudinale. Les résultats nous permettront d'appuyer le rôle fondamental de la prosodie dans l'enseignement des langues étrangères et de valider l'efficacité de la Méthode Verbo-Tonale.

Enfin, en conclusion, nous reviendrons sur les objectifs de recherche que nous venons de présenter et nous ouvrirons ce travail à de nouvelles perspectives.

1^{ÈRE} PARTIE : CADRES THÉORIQUE ET MÉTHODOLOGIQUE

Chapitre 1 : Problématique de l'enseignement de l'oral

« Occupez vous du sens, les sons se débrouilleront bien eux-mêmes »

Lewis Carroll

Ce premier chapitre est consacré à un état des lieux sur l'oral et sur sa place dans l'enseignement des langues étrangères.

L'oral est aujourd'hui reconnu comme un véritable objet d'étude. Pourtant, si depuis la linguistique structuraliste on lui accorde en principe une place de choix, les recherches sur l'oral restent encore minoritaires par rapport à celles sur l'écrit.

En ce qui concerne la didactique des langues étrangères, l'oral est le plus souvent considéré comme simple outil au service de la communication et non pas comme objet d'enseignement. Ainsi, l'enseignement de la matière phonique (accentuation, rythme, intonation et phonèmes, les bases de l'oralité par excellence !) est le plus souvent laissé pour compte dans les manuels au prétexte que les aspects sonores « *se débrouilleront bien eux-mêmes* ».

Pour mieux appréhender ces paradoxes, nous reviendrons sur les limites de la dichotomie oral/écrit et les difficultés que rencontrent les chercheurs pour délimiter et analyser un objet d'étude aussi hétérogène et volatil que l'oral.

Nous proposerons une brève rétrospective de l'enseignement de l'oral en langue étrangère, puis nous recentrerons notre analyse sur l'enseignement de la prononciation. Nous illustrerons trois présupposés sur l'apprentissage de la prononciation auxquels nous apporterons un éclairage perceptif.

Enfin, nous clôturerons ce chapitre par la comparaison de plusieurs méthodes de correction phonétique, en nous recentrant principalement sur deux d'entre elles, la méthode articulatoire et la méthode verbo-tonale d'intégration phonétique.

1.1 INTRODUCTION

En dépit de la place accordée à l'oral dans la didactique, en langue maternelle (désormais LM) comme en langue étrangère (désormais LE), le définir comme objet d'enseignement ne va pas de soi.

« La didactique des langues a beau avoir fait reposer la plupart de ses grandes transformations méthodologiques sur l'affirmation d'une nécessaire prise en compte des formes orales dans la communication verbale (...) elle reste empêtrée dans les représentations inadéquates de cet oral » (Coste, 1994 : 7)

Pour Debaisieux (1997), il y a un décalage entre la représentation que l'on se fait de l'objet dans son ensemble (macro-représentation) et la représentation que l'on se fait des unités de base du discours oral (prosodie, constructions grammaticales propres à l'oral, hésitations *etc.*) (micro-représentation). Ainsi, avant de recentrer notre questionnement sur l'enseignement de l'oral, nous nous intéresserons dans un premier temps à la nature même de l'oral.

1.1.1 Introduction au concept d'oral

Les recherches portant sur l'oral sont relativement récentes. En effet, ces travaux, dans la perspective interactionniste comme en analyse du discours et en analyse conversationnelle, ne prennent leur essor qu'au début des années quatre-vingts du siècle dernier (cf. travaux pionniers notamment de Gadet, 1989 ; 1996 ; Blanche-Benveniste, 1997 ; Berrendonner, 1993).

Cela est dû non seulement à la place privilégiée qui était réservée jusque là à l'écrit, mais aussi et surtout au fait que la nature même de l'objet d'étude soulève de nombreux problèmes théoriques et méthodologiques (Nonnon, 1999 ; Detey, 2007) : est-il possible, par exemple, d'étudier l'oral en lui-même et pour lui-même ? Comment délimiter, analyser et décrire un objet d'étude aussi variable et éphémère que l'est l'oral ? Peut-on le normaliser ou le transcrire ? Les unités de l'écrit fonctionnent-elles à l'oral ? Autant de questions auxquelles les chercheurs travaillant sur l'oral sont automatiquement confrontés.

Concernant les macro-représentations que l'on peut avoir sur l'oral, par exemple, il est nécessaire de savoir si l'oral et l'écrit constituent deux systèmes distincts ou deux réalisations différentes d'un même système. Le deuxième point de vue est notamment soutenu par Blanche-Benveniste (1997) et Gadet (1996). Pour Blanche-Benveniste (1997) et Blanche-Benveniste & Bilger (1999), il faut se méfier des « mythes séparateurs » qui cherchent à opposer l'oral primaire et familier à l'écrit élaboré et littéraire.

« (...) ce que l'on va pouvoir entendre par 'oral spontané' finit par renvoyer, et ce de manière implicite, à un usage particulier et restreint de la langue orale, à savoir le parler familier ou relâché » (Blanche-Benveniste & Bilger, 1999 : 21).

Il est vrai, comme le souligne Gadet (1996), que contrairement à l'écrit « codifié, fixé, stabilisé, normé » l'oral se présente sous des formes très variées : variations diachroniques (temps), diatopiques (régionales), diatrasitiques (sociales), diaphasiques (situationnelles) mais aussi inter ou intra locuteurs. Nous noterons que la variation situationnelle ou stylistique ne constitue en réalité qu'une variation parmi d'autres et ne pourrait en aucun cas se substituer au médium oral lui-même. Toutefois, nous touchons ici une des principales difficultés à laquelle sont confrontés chercheurs et didacticiens de l'oral : la gestion de la variabilité. D'autant plus que les différentes catégories censées répertorier toutes les variations possibles de l'oral ne rendent pas toujours compte de la réalité de la langue ou des interactions pouvant exister entre elles. Que dire par exemple, de la comparaison entre un discours prononcé avec un accent marseillais (variante diatopique) dans le cadre d'une conférence à l'université et un discours prononcé avec l'accent parisien (variante diatopique) dans le cadre d'un repas en famille (variante diaphasique) ? Pour répondre à ce type de questions, des corpus, comme PFC (Phonologie du Français Contemporain) pour le français, ont pour objectif de décrire la prononciation d'une langue dans toutes ces variantes (Detey *et al*, 2009).

Concernant la micro-représentation de l'oral, il faut bien comprendre que les unités de l'écrit ne fonctionnent pas à l'oral. La phrase, unité de base de la grammaire de l'écrit, visuellement marquée par la présence d'une majuscule et d'un point, n'est pas représentative de la grammaire de l'oral. En effet, le langage oral est un supra-système où s'articulent trois systèmes complémentaires : le système verbal (phonèmes, morphèmes, lexique et syntaxe), le système vocal (prosodie et vocalisations) et le système gestuel. Ainsi, l'unité maximale de l'oral serait caractérisée par une dimension vocale (par exemple, par la chute de la fréquence fondamentale sur la dernière syllabe d'un groupe intonatif, Rossi, 1985). Même si la définition de l'unité maximale de l'oral est sujette à débat (Grobet & Simon, 2002 ; Avanzi & Martin, 2007), on ne peut nier le rôle fondamental de l'intonation dans la représentation de l'oral, dimension dont l'écrit ne peut pas rendre compte. Des recherches actuelles en synthèse de parole ont par exemple démontré l'insuffisance de la ponctuation pour représenter l'intonation et permettre d'oraliser correctement les textes (Mertens *et al*, 2001 ; Martin, 2006). De même, le développement des émoticônes dans les mails illustre bien les limites de la ponctuation et la nécessité de signaler (ici par une image) l'intention associée au message écrit afin d'éviter tout quiproquos.

À l'opposé de la linéarité temporelle et spatiale de l'écrit, l'oral est trop souvent vu comme un produit inachevé à cause de la présence d'hésitations, des retours-arrières, des répétitions ou des pauses. Longtemps négligées par les linguistiques car considérés comme des scories du discours, les pauses et les hésitations constituent pourtant des indices d'élaboration des énoncés et jouent un rôle déterminant dans la gestion des tours de paroles (Goldman-Eisler, 1968 ; Guaitella, 1996 ; Duez, 1993, 1997).

Autrement dit, il est nécessaire d'étudier l'oral dans toute sa globalité propre et, pour ce faire, la description et la transcription de l'oral ne sont possibles qu'à travers l'analyse de corpus. Ainsi, les données fournies par les corpus oraux ont notamment permis de se détacher des normes prescriptives héritées de l'écrit (Debaisieux, 1997 ; Blanche-Benveniste & Bilger, 1999 ; Detey

et al, 2009). Mais quelles sont alors les normes de l'oral ? Là, la question reste entière, en particulier en LE, où il n'existe encore que trop peu de corpus

Ainsi, on comprend mieux pourquoi les enseignants ont autant de difficultés à définir l'oral comme véritable objet d'enseignement et à l'évaluer en tant que tel (Nonnon, 2005 ; Weber, 2006; Detey, 2007).

1.1.2 L'oral dans la représentation des enseignants

Pour les enseignants, l'oral constitue un objet d'apprentissage, à travers la mise en place d'activités communicatives telles que les discussions, les débats ou les jeux de rôles, et non pas un objet d'enseignement.

« L'oral ne donne pas lieu à un enseignement ciblé et structuré (...) cette absence d'enseignement structuré et orienté de l'oral correspond à l'idée selon laquelle on apprend en communiquant » (Lauzon et al, 2009)

Pour Dolz et al (1998), cette difficulté à concevoir l'oral comme un véritable savoir est liée d'une part au peu de recherches portant sur la langue orale et à la nature même de l'oral qui le rend difficilement scolarisable, du moins dans une vision traditionnelle de l'enseignement. En effet, contrairement à l'écrit que l'on peut pratiquer avec de grands groupes et dans le silence, chaque apprenant pouvant travailler individuellement, l'oral suppose l'interaction entre au moins deux personnes dans une dimension sonore.

L'enseignement de l'oral reste en réalité au service de l'enseignement de l'écrit et les spécificités de l'oral ne sont pas soulignées :

« L'oral est bien présent, certes, mais un oral particulier, en quelque sorte 'scolaire', fonctionnant comme faire valoir au service de la structure formelle, bien sûr écrite de la langue » (De Pietro & Wirthner, 1996 : 38)

Cette idée est partagée par Maurer (2002) qui précise que le point de départ de l'enseignement de l'oral est malheureusement bien souvent le texte lu. Tout est fait comme si l'oral nécessitait automatiquement l'écrit. Dans ce travail,

nous développerons l'hypothèse inverse en appuyant le rôle de l'oral dans le décodage de l'écrit (cf. chapitre 4).

En effet, si l'oral tient aujourd'hui une place de choix en tant qu'outil de communication dans les manuels de langue, il n'y a pas de positionnement théorique ou méthodologique par rapport aux questions précédentes (cf. problèmes soulevés par Nonnon, 1999, Maurer, 2002 et Detey, 2007) et l'oral reste une notion floue, rarement définie. Les pédagogues contournent le problème en substituant à l'oral des quasi-synonymes tels que expression orale, communication ou interaction, bien incapables de rendre compte de sa globalité et de sa richesse.

Ces trois termes, étroitement liés, désignent des compétences communicatives globales : la capacité à s'exprimer dans des situations très diverses dans la langue cible, la capacité à prendre en compte les différents éléments d'une situation de communication et à réadapter son message en fonction du récepteur, et la capacité à co-construire son apprentissage par l'interaction (qui devient à la fois la fin et le moyen) (Cuq, 2003). Ainsi, l'enseignement de la communication s'est donc substitué à l'enseignement de l'oral.

Mais on peut se demander si lorsqu'on évalue ces compétences communicatives, on évalue vraiment l'oral.

« Pour un enseignant tenter d'évaluer l'oral demande de pouvoir se décentrer du jeu de la communication pour prendre en compte le langage réellement produit par chaque élève, ce sur quoi il bute et ce en quoi il évolue. Or la plupart des situations de langage sont des situations interactives où l'enseignant est impliqué en tant qu'interlocuteur, affectivement et linguistiquement » (Nonnon, 2005 : 169)

Dolz *et al* (1998) soulignent que même si les enjeux sociaux de la maîtrise de l'oral sont immenses (vie quotidienne, entretiens d'embauche *etc.*), il n'existe pas de critères fiables et objectifs pour évaluer l'oral. Le Cadre Européen Commun de Référence pour les langues (CECR) a beau envisager

l'oral à travers des activités variées, comme la réception, l'interaction ou la production, lorsqu'il s'agit d'évaluer l'oral, les critères retenus pour évaluer la production n'évaluent en réalité que la dimension interactive et restent à la fois subjectifs et très flous en ce qui concerne la maîtrise du système phonologique (figure 1). Que signifie, par exemple, « s'exprimer de façon suffisamment claire » ?

Grille d'évaluation de la production orale

1^{ère} partie – Entretien dirigé									
Peut établir un contact social, se présenter et décrire son environnement familial.	0	0,5	1	1,5	2	2,5	3		
Peut répondre et réagir à des questions simples. Peut gérer une interaction simple.	0	0,5	1						
2^{ème} partie – Monologue suivi									
Peut présenter de manière simple un événement, une activité, un projet, un lieu etc. liés à un contexte familial.	0	0,5	1	1,5	2	2,5	3		
Peut relier entre elles les informations apportées de manière simple et claire.	0	0,5	1	1,5	2				
3^{ème} partie – Exercice en interaction									
Peut demander et donner des informations dans des transactions simples de la vie quotidienne. Peut faire, accepter ou refuser des propositions.	0	0,5	1	1,5	2	2,5	3	3,5	4
Peut entrer dans des relations sociales simplement mais efficacement, en utilisant les expressions courantes et en suivant les usages de base.	0	0,5	1	1,5	2				
Pour l'ensemble des 3 parties de l'épreuve									
Lexique (étendue et maîtrise) Peut utiliser un répertoire limité mais adéquat pour gérer des situations courantes de la vie quotidienne.	0	0,5	1	1,5	2	2,5	3		
Morphosyntaxe Peut utiliser des structures et des formes grammaticales simples. Le sens général reste clair malgré la présence systématique d'erreurs élémentaires.	0	0,5	1	1,5	2	2,5	3	3,5	4
Maîtrise du système phonologique Peut s'exprimer de façon suffisamment claire. L'interlocuteur devra parfois faire répéter.	0	0,5	1	1,5	2	2,5	3		

Figure 1 Grille du ciep 2011 pour l'évaluation de l'oral pour le niveau A2

En prenant en compte l'oral uniquement dans sa dimension communicative, on néglige paradoxalement ce qui fait sa nature même, les traits de l'oralité comme la prosodie, la grammaire spécifique de l'oral *etc.* (Debaisieux, 1997 ; Harmegnies *et al*, 2005). Ainsi, même dans l'évaluation de la maîtrise du système phonologique (*cf.* figure 1), on est dans une évaluation globale et subjective.

Pour La Fontaine & Messier (2009), c'est un cercle vicieux : on ne sait pas évaluer l'oral car on ne sait pas l'enseigner et on ne l'enseigne pas parce qu'on ne sait pas l'évaluer.

Il n'y a donc plus d'échappatoire possible : si on veut enseigner et évaluer l'oral, il est nécessaire de s'interroger sur ces éléments qui font sa spécificité, les fameux traits d'oralité, comme la prononciation, la grammaire ou les ponctuations de l'oral.

Dans le cadre de ce travail, nous nous intéresserons seulement au premier aspect, la prononciation, sous ses aspects phonémiques mais aussi et surtout prosodiques qui constituent la base de l'oral, considéré comme un savoir-faire plutôt qu'un savoir (Harmegnies *et al*, 2005).

« le savoir-faire visé est celui qui consiste à traiter la matière phonique en situation d'expression orale d'une manière similaire à celle qui caractérise les natifs » (Harmegnies *et al*, 2005 : 285)

Avant de recentrer notre travail sur l'enseignement de la prononciation, nous proposons une brève rétrospective de la place de l'oral dans les méthodologies d'enseignement du Français Langue Étrangère (désormais FLE).

1.2 DE L'ENSEIGNEMENT DE L'ORAL A L'ENSEIGNEMENT DE LA PRONONCIATION

1.2.1 Des méthodes traditionnelles à l'éclectisme : quelle place pour l'enseignement de l'oral ?

L'enseignement des LE a longtemps été calqué sur les méthodes grammaire-traduction des langues anciennes (comme le latin et le grec), privilégiant l'écrit, littéraire de préférence. L'apprentissage était alors conçu comme la maîtrise d'un ensemble de règles et d'exceptions, par l'enseignement explicite de la grammaire et la traduction de textes de grands auteurs. À l'oral volatil, variable et jugé familier, les pédagogues préféraient l'écrit stable, normatif et littéraire.

Il a fallu attendre l'avènement des méthodes directes (fin du XIX^{ème} siècle) et surtout des Méthodologies Audio-Orales (MAO, dans les années 40) et Audio-Visuelles (MAV, dans les années 60) pour que soit enfin remis en question l'hégémonie de l'écrit. Ces méthodes/méthodologies accordaient une place prioritaire et prédominante à l'oral et à l'enseignement de la prononciation, par la répétition et l'imitation d'exercices structuraux dans les laboratoires de langues, ou la mise en place de session de correction phonétique. Autrement dit pour les MAO et les MAV l'enseignement de l'oral était très balisé (Girard, 1972). Notons qu'à la même époque en Français Langue Maternelle (désormais FLM) le terme oral ne figurait pas dans les instructions données aux professeurs. On lui préférait le « parler des élèves », sur lequel le professeur agissait en améliorant l'élocution (apprendre à parler en respectant la norme) ou en utilisant la parole pour apprendre des leçons (récitations). Dans cette logique, l'oral restait encore étroitement lié à l'écrit : « apprendre à écrire comme apprendre à parler c'est apprendre à penser » selon les instructions officielles de 1938.

La fin des années 70 marque un tournant avec l'arrivée des approches communicatives qui, tout en redonnant à l'écrit une place que les MAV très axées sur l'acquisition de la compétence orale avaient quelque peu négligé, relancent l'intérêt pour l'oral en explorant d'autres voies pédagogiques, telles que les jeux de rôles, les simulations globales etc. Le but de l'enseignement est alors de permettre à l'apprenant d'utiliser la langue dans toutes les situations de communication auxquelles il sera confronté, à travers des activités des jeux de rôles ou des simulations proches de la réalité. L'enseignement de compétences communicatives (figure 2) se substitue à l'enseignement de l'oral.

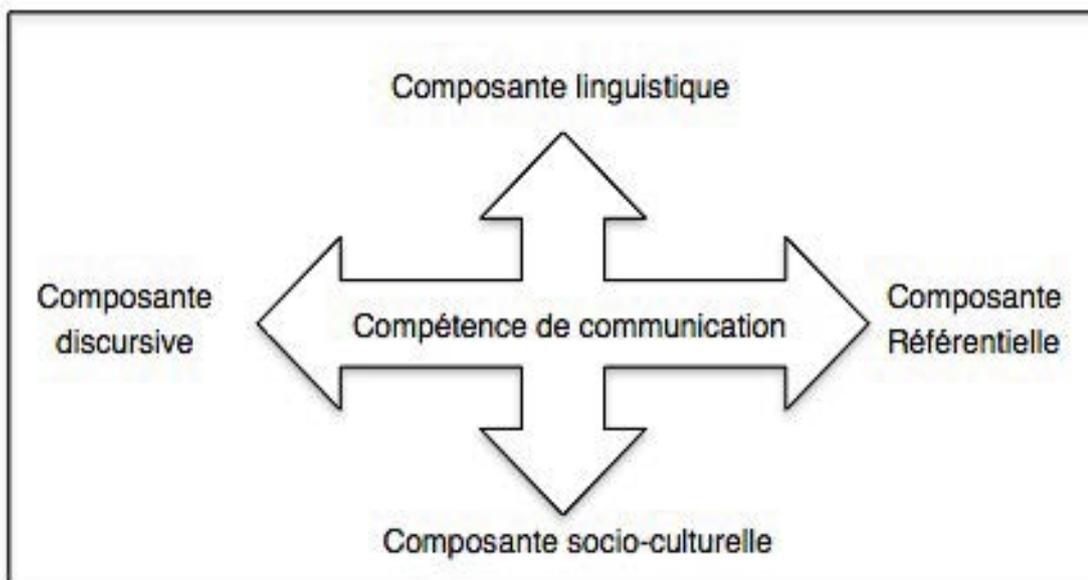


Figure 2 Schéma de la communication représentant les 4 compétences communicatives : la composante linguistique (être capable de faire des phrases), la composante discursive (être capable d'utiliser des phrases), la composante socio-culturelle (être capable de parler aux gens) et la composante référentielle (être capable de parler du monde) (d'après Defays, 2003)

Toutefois, ce concept ne fait pas l'unanimité. Les principales critiques qui lui sont adressées sont liées au manque de clarté dans l'articulation entre les différentes composantes et à l'absence de références faites aux compétences qui correspondent à chacune d'entre elles, pas plus en ce qui concerne la compréhension que l'expression, à l'oral comme à l'écrit (Moirand, 1982 ; Canale & Swain, 1981).

Autrement dit, au prétexte de la communication, on a non seulement contourné tout positionnement oral/écrit mais on a surtout soigneusement évité de définir l'oral (et l'écrit) en termes de compétences linguistiques. Les approches communicatives ont en fait été marquées par une mise au pilori de l'enseignement de la prononciation jugé trop lourd à mettre en place, demandant trop d'efforts de la part du professeur comme de l'élève pour des résultats souvent décevants (Billières, 2008).

À partir des années 80, on assiste également à une prise en compte de la dimension cognitive avec un intérêt marqué pour tout ce que recouvre le fait de comprendre, tant en modalité orale qu'écrite. Malheureusement en FLE, peu de chercheurs vont s'accommoder de la rigueur de ces approches et les approches communicatives vont dès lors régner en maître.

Aujourd'hui, depuis les années 90, le maître mot de la didactique est l'éclectisme (Puren, 1994). L'enseignant va piocher librement dans l'une ou l'autre des méthodes disponibles sur le marché, compte tenu de ses objectifs, des particularités du groupe qu'il a à former, des instructions éventuelles de l'institution, *etc.* L'enseignement s'articule autour de quatre grands savoir-faire (tableau 1) mais on ne sait pas comment hiérarchiser ces derniers : l'oral et l'écrit sont introduits simultanément sans apprentissage spécifique ; la lecture oralisée est devenue une activité de production orale, alors que comme nous le verrons plus loin (chapitre 4) elle est loin d'être une activité d'oralisation simplifiée ; l'enseignement de la prononciation a quasiment disparu de la salle de classe au profit de la communication, mot clef de l'enseignement des langues étrangères depuis les années 70.

Compétences	Compréhension	Production
Oral	écouter	parler
Écrit	lire	écrire

Tableau 1 Les quatre savoir-faire visés en didactique des langues

En réalité, on se trouve face à un grand flou méthodologique, il s'agit plus de recettes que de préceptes de travail et l'enseignant est bien souvent livré à lui-même.

Ce qui est sûr, c'est que si l'oral reste une priorité, toute la réflexion qui avait été amorcée sur l'oral, et en particulier, sur l'importance de la prononciation, s'est progressivement délitée, au prétexte qu'une prononciation défectueuse n'entrave pas la communication, et que l'accent étranger est un marqueur de l'identité culturelle des apprenants.

Depuis les années 60, les théories de l'apprentissage minimisent en effet l'importance d'une bonne prononciation en langue étrangère, en se basant sur trois présupposés réducteurs (Nonnon, 2002) :

- il n'est plus possible de maîtriser les sonorités d'une LE après une période donnée, située entre l'enfance et l'adolescence selon les auteurs (hypothèse de la période critique);

- l'accent étranger n'a pas d'incidence négative sur la compréhension des productions des apprenants ;
- comme pour l'acquisition de la LM, la prononciation en LE s'acquiert naturellement par les contacts répétés avec la langue cible ;

Mais que valent ces présupposés, par ailleurs très répandus dans l'enseignement ? Ont-ils déjà été testés ? Ont-ils été validés ?

Dans la partie suivante nous interrogerons chacun de ces trois présupposés à la lumière de recherches en didactique, en phonétique ou en pathologies du langage.

1.2.2 Présupposés à l'enseignement de la prononciation

Nous noterons que les deux premiers sont étroitement liés. En effet, l'accent étranger est l'aspect le plus soumis à l'hypothèse de l'existence d'une période critique dans l'apprentissage des langues. Ainsi, dans les deux sous-parties qui vont suivre, nous garderons présent à l'esprit que l'accent étranger se définit globalement par les écarts de prononciation commis par les apprenants vis à vis de la norme de prononciation attendue et partagée par les natifs d'une langue donnée.

1.2.2.1 Existe-t-il une période en deçà de laquelle il n'est plus possible de maîtriser une langue étrangère ?

L'hypothèse de l'existence d'une période critique dans l'apprentissage d'une LE est liée à l'âge d'acquisition de la LM. Pour Lenneberg (1967), un des pionniers des recherches sur la période critique, il ne serait plus possible de maîtriser une langue étrangère après l'adolescence, pour des raisons neurologiques liées à la perte de la plasticité cérébrale suite à la latéralisation de la zone du langage dans l'hémisphère gauche.

En effet, les zones du langage sont déterminées vers l'âge de sept ans. Ainsi, en ce qui concerne les pathologies du langage, il est prouvé que les enfants, contrairement aux adultes, peuvent récupérer leurs capacités

langagières de manière surprenante après une lésion cérébrale ou même une lobotomie. Si la lésion apparaît avant l'âge d'un an, la récupération sera même totale (Boysson-Bardies, 1996).

Mais la latéralisation de la zone du langage dans l'hémisphère gauche ne constitue pas la seule raison évoquée pour justifier l'existence d'une période critique. Ainsi, Birdsong (1999) fait également mention, entre autres, de la perte progressive de nos capacités d'apprentissage, de la perte progressive de nos capacités perceptives, des interférences entre la langue maternelle et la langue étrangère et des interférences entre les compétences de haut-niveau et les compétences de bas niveau. Concernant le dernier point, contrairement au proverbe « *qui peut le plus peut le moins* », il y aurait un ordre précis d'apprentissage du langage et l'acquisition de compétences de haut-niveau empêcherait l'acquisition de compétences de plus bas-niveau.

« (...) *maladaptive gain of processing capacity with maturation (...) is thought to be ill-suited to the narrow modularized task of acquiring language, hence the lack of success typically associated with adults second language learners* » (Birdsong, 1999 : 5)

Même si tous les travaux ne se limitaient pas à l'apprentissage de la prononciation, pour Scovel (1988) il s'agirait pourtant du paramètre le plus assujéti aux contraintes de la période critique parce que c'est le seul aspect de la performance linguistique qui repose sur une base neuromusculaire et qui a une réalité physique. Pour cet auteur, passé l'adolescence, l'apprenant de langue étrangère ne pourra plus prétendre à une prononciation semblable à celle des natifs.

Certaines études présentent pourtant des résultats en contradiction avec ces hypothèses. Les capacités perceptives seraient encore malléables à l'âge adulte (Best & Strange, 1992). Nos aptitudes à apprendre de nouvelles formes resteraient intactes tout au long de notre vie (Flege, 1981, 1993, 1999), même si l'absence d'accent étranger est plus rare chez des apprenants tardifs (Flege et al, 1995).

Il semblerait en réalité que les compétences ne soient pas perdues mais simplement moins accessibles chez un apprenant tardif, du fait d'une plus grande habitude et une plus grande pratique de la langue maternelle (Wode, 1994)

Dans son modèle de perception de la parole, *Speech Learning Model* (SLM), Flege (1995, 2005) explique que si les mécanismes qui permettent l'acquisition d'une langue restent intacts tout au long de la vie, il devient cependant de plus en plus difficile de former de nouvelles catégories correspondant aux sonorités de la langue étrangère au fur et à mesure que les catégories pour la langue maternelle se développent. Par exemple, si deux sons vocaliques sont très proches, par exemple le /i/ français présent dans le mot « ami » et le /ɪ/ anglais présent dans le mot « bit », ces deux sons seraient assimilés à une seule et même catégorie, celle de la LM. Autrement dit, un francophone apprenant l'anglais, ne fera pas la différence entre la voyelle antérieure /i/ qu'il possède dans sa langue maternelle, et la voyelle plus centrale /ɪ/ de l'anglais. Inversement, il serait plus facile pour un apprenant de créer de nouvelles catégories perceptives pour des sons qui ne s'apparentent pas à ceux de sa langue maternelle. Nous noterons que ce modèle traite les différences perceptives entre la langue maternelle et la langue étrangère sur la base d'un continuum, mesuré empiriquement et non prédit a priori. Toute modification des catégories existantes ainsi que toute création de nouvelles catégories entraînerait une réorganisation du continuum (figure 3).

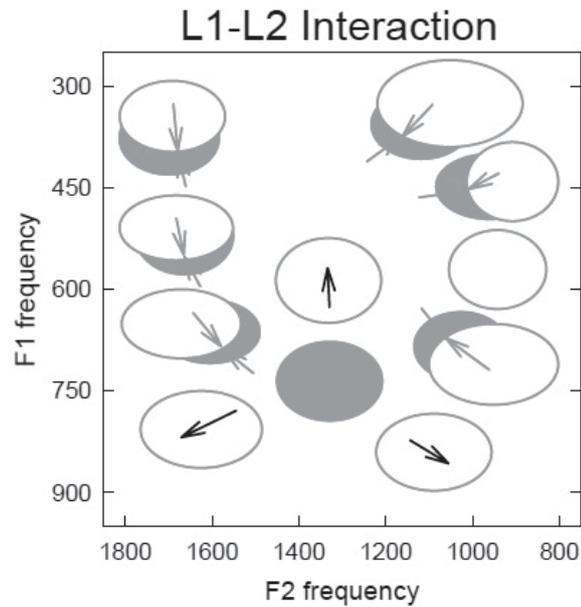


Figure 3 Représentation des interactions entre les voyelles de la langue maternelle cercles gris (L1) et les voyelles de la langue étrangère cercles blancs (L2) dans le trapèze vocalique (en abscisse, fréquence du deuxième formant vocalique et en ordonnée, fréquence du premier formant). Les assimilations sont représentées par des flèches grises et les dissimilations, entraînant la création de nouvelles catégories, par les flèches noires (Flege, 2005, http://jimflege.com/files/Vancouver_April_2005.pdf)

Ainsi, la maîtrise des éléments segmentaux (voyelles et consonnes) en LE n'est pas impossible après l'adolescence, bien qu'elle soit soumise à l'incidence des catégories déjà bien établies pour la LM.

Mais qu'en est-il pour les éléments suprasegmentaux (i.e. la prosodie constituée par l'accentuation, le rythme et l'intonation), par ailleurs premiers dans l'ontogénèse (Billières, 2002 ; Ramus et al. 2000 ; Karmiloff-Smith &, Karmiloff-Smith, 2003)? Sont-ils plus soumis à l'influence de la LM et à la période critique ?

Pour Kaglik & Boula de Mareüil (2009, 2010), ni les transferts entre la LM et la LE ni la période critique ne suffiraient à expliquer les difficultés liées à la maîtrise des éléments suprasegmentaux. En effet, l'analyse des productions d'apprenants polonais en Français Langue Étrangère (FLE) a montré que ces derniers utilisaient bien les stratégies de découpage prosodique de leur LM en français, mais que la fréquence d'utilisation de ces stratégies n'étaient pas la même que dans leur langue maternelle. Le coût cognitif serait plus élevé en LE car la production nécessiterait alors une double focalisation sur le fond et sur la

forme (Kaglik & Boula de Mareüil, 2009). Cette observation a amené ces auteurs à conclure que l'acquisition de la prosodie échapperait à l'hypothèse de la période critique car elle nécessiterait la mise en place de compétences de haut niveau:

« Mastering rules governing prosodic groupings and detachments as well as the nature of cues associated with these operations requires a very high level of competence and seems to raise difficulties even for very advanced speakers » (Kaglik & Boula de Mareüil, 2010 : 4)

Cette conclusion fait écho à l'hypothèse de Perdue & Gaonach' (2000) selon laquelle les coûts cognitifs ne seraient pas les mêmes en LM et en LE. De plus, considérant la variabilité des observations au sein de groupes monolingues, ces auteurs mettent également en avant les difficultés de transferts de compétences de la langue maternelle à la langue étrangère en raison du coût cognitif représenté par les processus de traitement et de production. Ces aspects seront discutés dans la suite de ce travail (chapitre 4) mais nous amènent dorénavant et déjà à considérer que l'apprenant n'est peut-être pas un « natif diminué » mais que sa production est le reflet de la mise en place de nouvelles connections en LE, d'une langue en construction.

En conclusion, la maîtrise de la prononciation ne serait pas inaccessible aux apprenants tardifs, d'où la remise en question de l'existence d'une période critique. Toutefois, les capacités perceptives seraient influencées par les catégories créées pour la langue maternelle et stabilisées par la pratique de celle-ci. Autrement dit, les difficultés d'apprentissage de la prononciation (pour le niveau segmental) sont plus imputables à l'influence de la LM qu'à l'âge d'acquisition à proprement parler. En ce qui concerne la prosodie, il semblerait au contraire, que les erreurs rythmiques et mélodiques dépassent la question de l'âge ou de l'incidence de la LM, mais qu'elles soient liées au coût cognitif que représente la focalisation à la fois sur le fond et sur le forme (nous y reviendrons).

La présence d'un accent étranger n'est donc pas le seul fait de l'âge. C'est pourquoi il nous semble intéressant de nous pencher sur la nature de cet accent étranger et le lien entre ce dernier, la prosodie et la compréhensibilité.

1.2.2.2 Quel est le lien entre la prosodie, l'accent étranger et la compréhensibilité ?

Comme nous l'avons déjà précisé, l'accent étranger se définit globalement par les écarts commis par les apprenants vis à vis de la norme de prononciation attendue et partagée par les natifs d'une langue donnée. Il ne s'agit donc pas d'une valeur palpable mais d'une impression d'authenticité évaluée perceptivement à l'aide d'une échelle de magnitude par des natifs (experts ou non).

Ce type de méthodologie a permis de dégager la pertinence de certains paramètres par rapport à d'autres. Il semblerait que les éléments suprasegmentaux jouent un rôle plus important que les éléments segmentaux dans la perception d'un accent étranger (Jilka, 2000 ; Birdsong, 2003 ; Vieru & Boula de Mareüil, 2006 ; Kaglik & Boula de Mareüil, 2009, 2010). De plus, plus la prosodie de la langue maternelle et de la langue étrangère sont éloignées plus les natifs vont être sensibles aux paramètres suprasegmentaux (Jilka, 2000).

Birdsong (2003) a montré que la prosodie était non seulement l'élément le plus important pour évaluer un accent étranger mais aussi que l'authenticité des segments était directement dépendante de l'authenticité du suprasegmental. Les analyses acoustiques et perceptives des enregistrements de parole lue (lecture de mots isolés et de phrases) en français langue étrangère par des apprenants anglophones ont permis à l'auteur de dégager les tendances suivantes : l'authenticité des éléments suprasegmentaux prédit l'authenticité des éléments segmentaux, alors que l'authenticité segmentale ne prédit pas l'authenticité suprasegmentale. La prosodie serait donc un élément clef dans la maîtrise de la prononciation d'une langue.

Mais comment expliquer son rôle et quelles incidences cela peut-il avoir sur le lien entre accent étranger et compréhension ?

Il faut savoir que l'acquisition du langage obéit à une chronologie étonnamment uniforme à travers les langues. Apprendre à parler, c'est d'abord maîtriser l'enveloppe prosodique de sa LM. En perception comme en production, la structuration prosodique se met en place avant l'acquisition des autres niveaux linguistiques (phonèmes, lexique, syntaxe *etc.*) (Konopczynski, 1986 ; de Boisson-Bardies, 1996 ; Molino, 2000 ; Billières, 2002 ; Karmiloff-Smith & Karmiloff-Smith, 2003). À la naissance, le nourrisson est déjà capable de distinguer sa LM d'une LE sur la base de leurs seules différences rythmiques (*cf.* test de la tétine, Ramus, 1999 ; Ramus *et al*, 2000). Le babillage, première étape de production de la parole, est structuré par le rythme et l'intonation de la langue maternelle.

De plus, la prosodie facilite la compréhension du discours. Elle participe en effet à la segmentation des énoncés en groupe de sens. Les premières études sur la parole de synthèse ont d'ailleurs illustré le fait qu'il est très difficile de comprendre un discours ne contenant pas les informations prosodiques (Grobet, 1997). De même, en langues étrangères, Rasier & Hiligsmann (2007) rappellent que des études montrent que les erreurs segmentales ont moins d'incidences négatives sur la compréhension et l'intelligibilité du message que les erreurs suprasegmentales (*cf.* Anderson-Hsieh *et al*, 1992 ; Munro & Derwing, 1995 ; Jilka, 2000 ; Hardisson, 2004).

Pourrait-on alors conclure puisque la prosodie constitue l'un des facteurs les plus marquants dans la perception d'un accent étranger et qu'elle joue un rôle dans la compréhension du message, que l'accent étranger et la compréhension seraient également étroitement corrélés ?

Il est encore difficile d'apporter une conclusion stricte à cette question. De récentes études semblent soutenir au contraire que l'accent étranger et la compréhension ne sont pas liés. Derwing & Munro (1997) ont étudié les relations entre accent étranger et compréhension du discours. Ces auteurs

ont montré qu'un discours peut-être compris malgré la présence d'un fort accent étranger, et inversement qu'un faible accent étranger ne garantit pas la compréhensibilité. Comme nous pouvons le voir sur la figure suivante (figure 4), l'évaluation de l'accent étranger représentée en gris (sur une échelle de 1= aucun accent à 9= accent très marqué) n'est pas corrélée à l'évaluation de la compréhensibilité représentée en blanc (sur une échelle de 1= compréhensibilité facile à 9= compréhensibilité impossible).

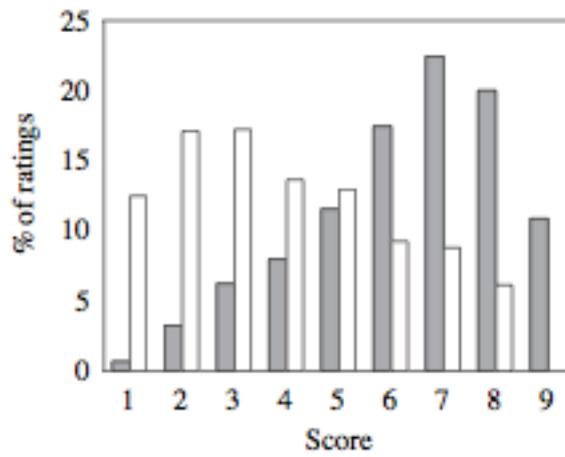


Figure 4 Distribution des scores d'accentuation (grisé) et de compréhensibilité (blanc) donné par un groupe d'auditeurs sur un échantillon de phrases (Derwing & Munro, 1997 ; Munro, 2008 : 204)

Pour Munro (2008), compréhensibilité et accent étranger feraient appel à des processus de traitement différents. Toutefois, l'auteur souligne que de plus amples recherches sont nécessaires pour comprendre le rôle entre l'accent étranger et la compréhensibilité du discours des apprenants de langue étrangère car les études actuelles se contredisent. En effet, faute de consensus sur la définition de l'accent étranger, les critères retenus ne sont pas les mêmes et amènent des conclusions différentes. Nous retiendrons néanmoins que Munro (2008) insiste sur le fait que la question n'est pas d'avoir ou pas un accent étranger mais d'avoir une production qui soit compréhensible et intelligible. Et, à ce niveau là, la prosodie jouerait un rôle déterminant.

Le dernier présupposé sur l'enseignement de la prononciation consistait à dire que la prononciation s'acquiert naturellement par simple contact avec la

langue cible et ne doit donc pas faire l'objet d'un apprentissage spécifique. Qu'en est-il réellement ?

1.2.2.3 *Peut-on acquérir la prononciation d'une langue étrangère par simple contact avec la langue cible ?*

Malgré les faits rapportés en supra, correspondant à des résultats d'expérimentations menées dans les laboratoires de recherche, la plupart des enseignants font aujourd'hui un parallèle (trop) rapide entre acquisition et apprentissage de la prononciation en LM et en LE. L'idée selon laquelle le simple fait d'être en contact avec la langue cible suffit à améliorer la production des apprenants, est aujourd'hui très répandue.

Seule l'expérience garantirait la maîtrise des sonorités de la langue étrangères. Selon le CECR, un apprenant de niveau B2 (Avancé) aura acquis « *une prononciation et une intonation claires et naturelles* », sans qu'aucune attention ne soit portée à la maîtrise de ces paramètres aux niveaux inférieurs. Au niveau B1, l'apprenant est intelligible malgré la présence d'un accent étranger et d'éventuelles erreurs de prononciation. Au niveau élémentaire (A1 et A2), la prononciation est censée s'améliorer progressivement malgré la présence d'un fort accent étranger et quelques difficultés d'inter-compréhension (nécessité pour l'apprenant de se répéter/de faire répéter son interlocuteur pour être compris/comprendre).

Nos expériences de classes, autant en milieu endolingue qu'en milieu exolingue, vont à l'encontre de ces présupposés : faute d'un enseignement spécifique, un apprenant de niveau B2 n'a pas acquis *une prononciation et une intonation claires et naturelles*. Au contraire, les productions des apprenants (tous niveaux confondus) sont très largement influencées par les règles de prononciation de leur LM (autant au niveau phonémique, qu'accentuel rythmique et intonatif) ce qui peut entraîner des problèmes de compréhension : un sinophone ne fera pas la différence entre les deux occlusives vélaires /k/ et /g/ dans les paires minimales « cale » et « gale » ; un russophone produira des questions ouvertes avec une intonation descendante, selon les règles

prosodiques de sa LM, alors que l'intonation descendante est associée en français, à une affirmation ou un ordre.

Inversement, des études comme celle de Bongaerts *et al* (2000) montrent que des apprenants néerlandais peuvent acquérir une prononciation perçue comme native en anglais grâce à un entraînement intensif en perception et en production des sonorités de la langue étrangère. Le bain linguistique ne suffirait pas à garantir l'apprentissage de la prononciation. Un apprentissage spécifique en production comme en perception des sonorités de la LE serait donc nécessaire.

Pour illustrer ce dernier point, nous présenterons deux modèles de perception de la parole celui de Kuhl (1991) et de Best (1994).

1.2.3 Le rôle de la perception

Nous savons que l'apprenant filtre les sonorités de LE via les catégories de sa LM (Flege, 1995 ; 2005). Certains auteurs comparent ce filtre à un crible phonologique qui rendrait l'apprenant sourd aux sonorités qui n'appartiennent pas à sa LM (Troubetzkoy, 1938 ; Renard, 1979 ; Billières, 2002).

« Le système phonologique d'une langue est semblable à un crible à travers lequel passe tout ce qui est dit. Seules restent dans le crible les marques phoniques pertinentes pour individualiser les phonèmes. [...] L'homme s'approprie le système de sa langue maternelle. Mais s'il entend parler une autre langue, il emploie involontairement pour l'analyse de ce qu'il entend le « crible phonologique » de sa langue maternelle qui lui est familier. Et comme ce crible ne convient pas pour la langue entendue, il se produit de nombreuses erreurs et incompréhensions » (Troubetzkoy, 1938)

Les travaux de Kuhl en perception de la parole (*théorie des aimants perceptifs*, Kuhl, 1991) renforcent la métaphore du crible phonologique. Dans le continuum sonore qui nous entoure, si nous sommes capables de reconnaître les sons de notre langue maternelle, c'est parce qu'au cours de l'acquisition de celle-ci, nous avons créé nos propres catégories perceptives. Autrement dit, le

cerveau de l'enfant va être modulé au contact avec sa langue maternelle, de manière à découper perceptivement le continuum sonore pour construire son propre crible (figure 5).

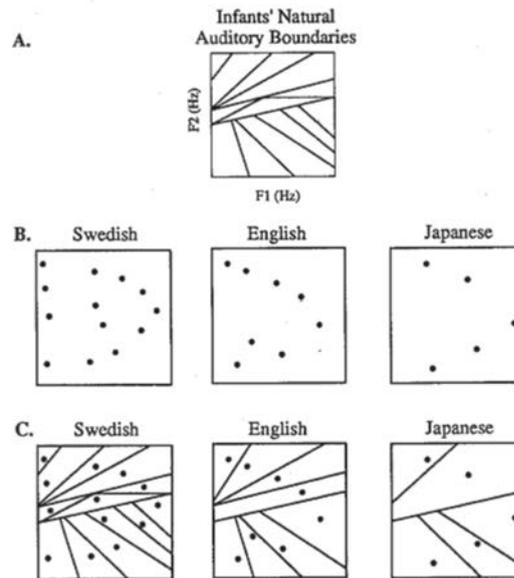


Figure 5 cette figure montrant la mise en place progressive des frontières pour les voyelles du suédois, de l'anglais et du japonais (Kuhl, 1998 : 61)

Comme nous pouvons le voir sur la figure précédente, le nouveau né posséderait potentiellement toutes les catégories de toutes les langues vu que les frontières ne sont pas encore spécifiées. Progressivement, l'enfant crée les catégories correspondant à sa LM (les espaces vocaliques du suédois, de l'anglais et du japonais sont représentés en b.). Kuhl *et al* (2006) ont montré que ces frontières se mettaient en place entre 10 et 12 mois et étaient nécessaire au bon développement de la LM. C'est ainsi que nous pouvons observer (en c.) des différences entre les enfants suédois, anglais et japonais : certaines frontières ont disparu et d'autres se sont maintenues, en fonction des langues cibles. Ainsi, le japonais ayant un espace vocalique réduit (seulement cinq voyelles), les catégories sont moins nombreuses que pour le suédois ou l'anglais mais recouvrent chacune un espace plus grand.

Ces catégories sont créées à partir de sons appelés des prototypes. Ces derniers agissent comme des aimants perceptifs et attirent à eux les stimuli qui leur ressemblent le plus. Les frontières entre catégories sont donc assujetties aux prototypes et non déterminées a priori (figure 6).

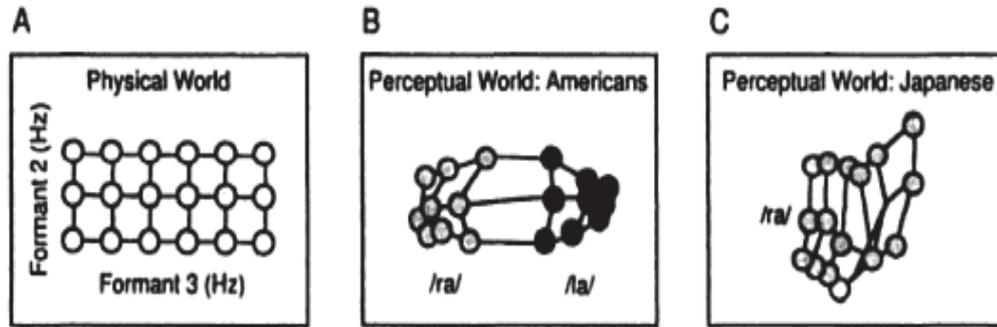


Figure 6 Illustration de la différence entre réalité physique (le continuum sonore) et réalité perceptive (réorganisation du continuum avec les prototypes) pour l'anglais et le japonais (Kuhl, 1998 : 59)

Dans la figure 6, l'image A représente la réalité physique. Dans cette réalité /r/ et /l/ sont équidistants l'un de l'autre. Dans la réalité perceptive d'un Américain (B), on constate que les éléments intermédiaires ont été regroupés pour appartenir à l'une ou à l'autre des deux catégories phonémiques et qu'une distance s'est installée entre les deux sons, marquant la frontière phonémique. La réalité perceptive du Japonais n'a rien à voir avec celle de l'Américain, on se distingue par de regroupement ni de frontière. Notre système perceptif serait donc modelé par notre LM. Mais que se passe-t-il lorsque nous sommes en contact avec les sons d'une LE?

Dans son modèle de perception des sons d'une langue étrangère, *Perceptual Assimilation Model* (PAM), Best (1994 ; Best *et al*, 2001) décrit plusieurs cas possibles de traitement des sons d'une LE.

Dans le premier cas, chaque son non-natif est assimilé à une catégorie native différente. Par exemple, les deux voyelles du français, l'antérieure mi-fermée arrondie /ø/ présente dans le mot « feu » et la postérieure mi-ouverte arrondie /ɔ/ présent dans le mot « porte », seront respectivement assimilées aux voyelles /e/ et /o/ en espagnol. Dans ce cas la discrimination est excellente (*Two-category discrimination*)

Dans le deuxième cas, les deux sons non natifs sont assimilés à la même catégorie native mais l'un des deux constituera une représentation plus acceptable que l'autre. C'est le cas, par exemple, des deux « i » de l'anglais, la

voyelle /ɪ/ présente dans le mot « bit » s'oppose à la voyelle /i:/ présente dans le mot « sheep » mais pourtant, en français ces deux sons vocaliques seront assimilés à la même catégorie « i ». Mais, pour des francophones, le second sera plus prototypique que le premier (*Category goodness difference*). La discrimination est modérée voir bonne.

Dans le troisième cas, les deux sons sont assimilés à une catégorie mais divergent tous les deux de l'idéal natif (*Single Category*). Pour des japonais, par exemple, les sons /l/ et /ɾ/ de l'anglais seront assimilés à une seule et même catégorie. La discrimination n'est pas possible.

Enfin, dans les dernier cas, les sons de la LE peuvent n'être rattachés à aucune catégorie existante ou même ne pas être considérés comme des sons de parole et associés à des bruits environnementaux (*Uncategorized*). C'est le cas, par exemple, de certains sons rares, comme les consonnes non pulmonaires appelées « clics » que l'on trouve dans quelques langues asiatiques ou africaines, qui vont être assimilés à des bruits environnementaux par des anglophones (Best *et al*, 1988).

Il ressort de ce dernier point, c'est que les erreurs de prononciation sont à mettre en relation avec les connaissances perceptives de la LM. On rejoint ainsi la conclusion de la partie sur la période critique, la maîtrise de la prononciation en langue étrangère n'est pas impossible mais un travail sur le crible phonologique reste nécessaire.

Il est possible d'imaginer que tout comme il existe un crible phonologique, il existerait un crible prosodique (Intravaia, 2000 ; Dodane, 2003). Faute d'entraînement spécifique, les apprenant de LE filtreraient la parole étrangère via les règles accentuelles, rythmiques et intonatives de leur LM, entraînant des erreurs de décodage et d'encodage. Delattre (1965) parle d'ailleurs d'interférences intonatives entre la LM et la LE. Nous pensons qu'il existe en effet des transferts prosodiques entre la LM et la LE et que ces transferts sont soumis à différentes contraintes, tels que l'influence de la LM ou le cout cognitif de la tâche. Nous reviendrons sur ces aspects dans le chapitre 4.

En conclusion les présupposés à l'enseignement de la prononciation ne sont pas validés expérimentalement. Il n'existe pas de période critique en LE, bien que les processus de perception soient influencés par les catégories de la LM. L'accent étranger est étroitement lié à la prosodie et de nouvelles recherches sont nécessaires pour déterminer la corrélation entre accent étranger, prosodie et compréhensibilité car les recherches actuelles se contredisent. En ce qui concerne l'acquisition de la prononciation en LE, il semblerait que cela ne soit possible sans un apprentissage spécifique, centrée sur la production mais également la perception des sonorités de la LE. En effet, les modèles de perception de la parole actuels montrent que la LM entraîne le découpage du continuum sonore en catégories spécifiques et que celui-ci est responsable des difficultés de discrimination des sons de la LE.

Avant de terminer ce premier chapitre, il nous paraît important, dans la mesure où nous avons mis en évidence la nécessité d'enseigner la prononciation, de présenter deux méthodes de correction phonétique utilisées en LE et de les discuter vis à vis des principes que nous venons d'énoncer.

1.3 COMPARAISON DE DEUX METHODES DE CORRECTION PHONETIQUE

Bien que peu utilisées dans les cours de langue actuels, il existe quelques méthodes permettant de corriger la prononciation d'apprenants de langue étrangère. Pour ce travail de thèse, nous nous intéressons plus précisément à deux de ces méthodes, la méthode articulatoire et la méthode verbo-tonale de correction phonétique. Tout d'abord ces méthodes sont comparables dans leur mise en place puisqu'elles ne nécessitent aucun matériel spécifique. De plus, contrairement aux logiciels de langue proposés sur le marché, elles se pratiquent dans une situation de classe standard en interaction enseignant/apprenant et ne s'appuient pas uniquement sur les capacités d'auto-correction des apprenants, comme c'est le cas de la plupart des logiciels disponibles sur le marché.

La description de ces deux méthodes est inspirée des travaux de Renard (1979), de Intravaia (2000), de Billières (1987, 1989, 1990, 1991, 1995, 2000, 2002, 2003, 2005), de Baqué (2004) et de Baqué & Estrada (2010).

1.3.1 La méthode articulatoire

La méthode articulatoire est la méthode de correction phonétique la plus ancienne et elle n'est rattachée à aucune méthodologie.

La méthode articulatoire considère que l'acte phonatoire implique la connaissance explicite de son fonctionnement : les sons sont décrits en termes de positionnement des organes phonatoires.

Pour produire le /y/ présent dans le mot « tulipe », on demandera aux apprenants de presser la pointe de la langue contre les incisives inférieures comme pour prononcer un /i/ et, en même temps, de projeter les lèvres comme pour siffler (figure 7).

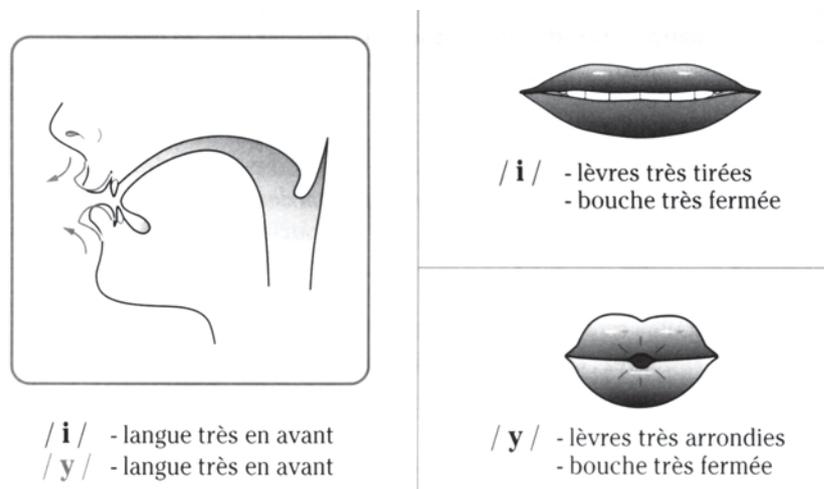


Figure 7 Exemple d'exercice articulatoire proposé pour la production des voyelles /i/ et /y/ dans un manuel (Charliac & Motron, 1998 : 46)

Pour produire les consonnes, par exemple, on sensibilise l'apprenant au mouvement de l'air (!) (cf. figure 8).

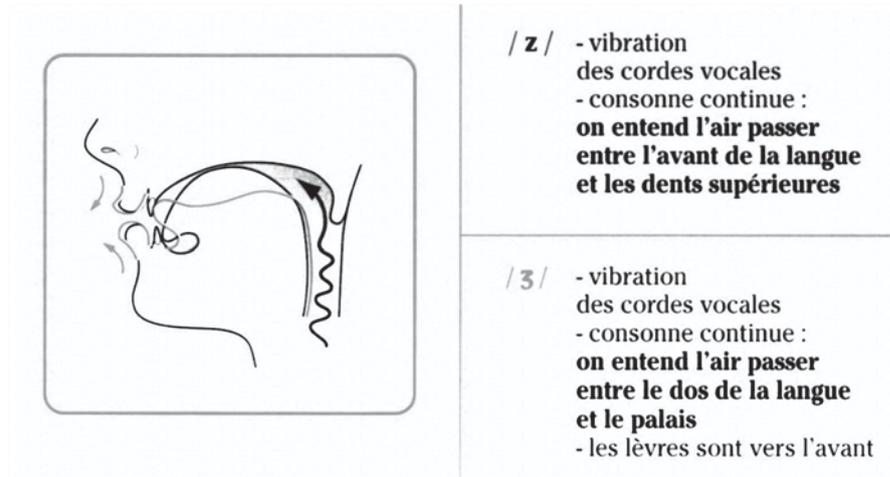


Figure 8 Exemple d'exercice articuloire porposé pour la production des consonnes /z/ et /ʒ/
dans un manuel (Charliac & Motron, 1998 : 150)

La correction passe par un processus d'intellectualisation de la façon dont il faut positionner la langue et les lèvres pour produire tel ou tel son : la première étape consiste à décrire le mouvement articuloire, éventuellement à l'aide de schéma (figures 7 et 8) avant de demander aux apprenants de le réaliser.

Ce principe d'intellectualisation peut être poussé à l'extrême. Pour prendre conscience des particularités articuloires des sons de la langue étrangère, l'apprenant peut avoir à suivre des cours de phonétique théoriques afin de pouvoir comparer les bases articuloires des deux systèmes.

Les exercices proposés respectent toujours l'ordre suivant : du plus simple au plus complexe, du son isolé au son en contexte de mot, puis en contexte de phrase. Ainsi, pour travailler le /y/, le son cible sera d'abord répété seul plusieurs fois, il sera ensuite opposé à un son voisin (dans notre cas, il peut s'agir du /u/ ou du /i/), puis introduit dans des mots isolés, mis en opposition avec un son voisin (« dessus » vs. « dessous ») et enfin, inséré dans un syntagme puis une phrase complexe (« le bouquet de tulipe » → « le bouquet de tulipe est sur la table »).

La démarche d'ensemble de cette méthode pourrait être résumée par la figure 9.

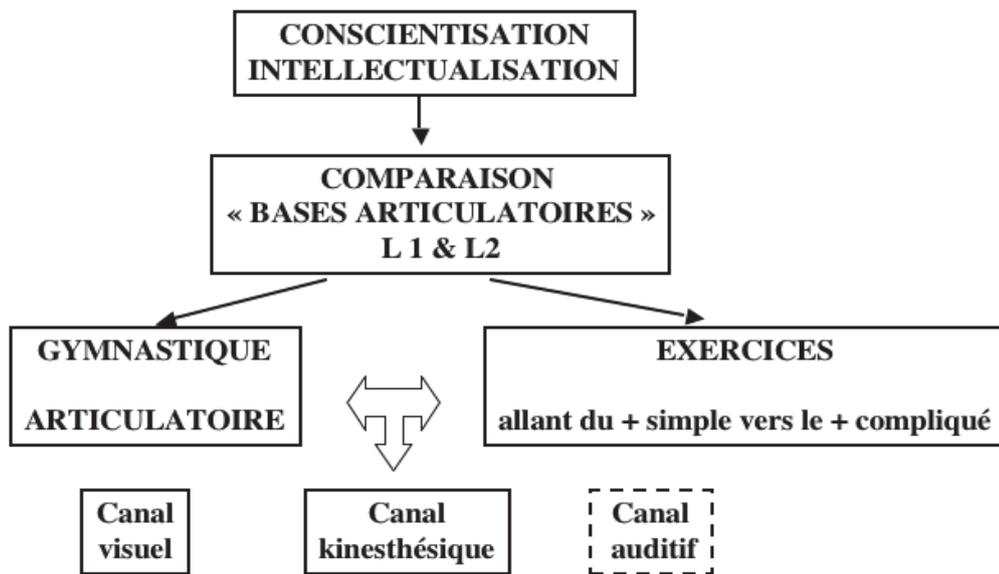


Figure 9 Vue d'ensemble des principes de la méthode articulatoire (Billière, 2000 : 138)

Comme nous pouvons le voir sur la figure précédente, la méthode articulatoire se base sur le fait qu'une bonne production implique la connaissance explicite et l'intellectualisation des procédés articulatoires. Pour cela, l'enseignant est invité à comparer les bases articulatoires de la LM (L1 dans le schéma) et de la LE (L2 dans le schéma). À partir de cette comparaison, il est possible de déduire les sons problématiques pour l'apprenant de LE et de proposer une gymnastique articulatoire adaptée (cf. figures 7 et 8) et de proposer des exercices du plus simple (son isolé) vers le plus compliqué (son cible intégré dans une phrase).

De plus, nous noterons que le canal auditif apparaît en pointillé car il n'est pas pris en compte par la méthode articulatoire qui néglige la dimension perceptive. Cette méthode néglige également les éléments suprasegmentaux, tels que l'accentuation, le rythme et l'intonation qui jouent pourtant un rôle déterminant dans la production et la perception de la langue. En ce qui concerne le canal kinesthésique, seuls les gestes de la main appuyant la description des schémas articulatoires (micro-motricité) sont pris en compte.

La figure 9 illustre que la méthode articulatoire considère la production de la langue dans sa forme la plus réduite, seuls les éléments segmentaux sont mis

en avant. Autrement dit, les différents paramètres de la communication, comprenant l'interaction des systèmes verbaux, vocaux et gestuels ne sont pas pris en compte.

A l'heure actuelle, on assiste à un regain d'intérêt pour cette méthode dans les manuels ou les formations de Français Langue Étrangère. Il est vrai que les exercices proposés sont intemporels et faciles à mettre en application, d'autant plus qu'ils ne nécessitent pas des connaissances très poussées en phonétique de la part de l'enseignant.

Bien que la méthode articulatoire ne soit rattachée à aucune théorie de la perception, son postulat central fait écho à la *théorie motrice* de Lieberman (Lieberman & Mattingly, 1985). Cette théorie soutient l'idée selon laquelle la parole serait perçue par l'intermédiaire des gestes articulatoires. En conséquence, la différence entre deux voyelles, produite par la modification graduelle du conduit vocal, serait perçue sur un mode continu, tandis que l'opposition entre deux consonnes, faisant intervenir des articulateurs différents, serait perçue sur un mode catégoriel. Cependant des auteurs comme Massaro & Cohen (1983) suggèrent que la perception catégorielle serait un artefact expérimental : la réponse catégorielle serait induite par le protocole expérimental.

Pour conclure, nous pensons que les principes de la méthode articulatoire soulèvent en réalité plusieurs questions.

Tout d'abord, la production d'un apprenant peut-elle se réduire aux différences entre LM et LE ? Jusque dans les années 60, on considérait que les compétences étaient transférables de la LM à la LE, sur la base de leurs seules ressemblances (transferts positifs) ou différences (transferts négatifs). Mais comme nous avons vu précédemment (cf. partie 1.2.2) d'autres processus (création de nouvelles catégories, coût cognitif, influence de la tâche *etc.*), seraient mis en jeu aussi bien en perception (Flege, 1995, 2005) qu'en production (Kaglik & Boula de Mareüil 2009, 2010).

Ensuite, la méthode articulatoire peut-elle fonctionner quel que soit le niveau des apprenants ? En effet, une démarche intellectualisée des processus articulatoires nous semble difficile à mettre en place avec des apprenants de niveau débutant. Cette méthode serait-elle alors plus appropriée pour des apprenants de niveau avancé ? Là encore, nous ne pensons pas qu'enseigner la prononciation dans un second temps, comme perfectionnement linguistique (comme c'est souvent le cas à l'heure actuelle) soit pertinent. En effet, l'apprenant risque d'avoir fossilisé ses erreurs de prononciation et d'avoir développé des stratégies de décodages inappropriées ou cognitivement plus coûteuses.

Nous avons d'ailleurs beaucoup insisté sur l'importance de la prosodie en illustrant le coût cognitif que représente une focalisation sur le fond et sur la forme en production quand les éléments suprasegmentaux ne sont pas maîtrisés. Peut-on alors négliger ces aspects dans l'apprentissage de la prononciation ? Nous ne le pensons pas. Pourtant la méthode articulatoire se focalise sur le phonème, unité minimale de perception mais pas de production. En production l'unité minimale est une unité rythmique de base, la syllabe (Massaro, 1972, 1974 ; Mehler *et al*, 1981). On sait que les phénomènes de coarticulation et de compensation entre les phonèmes sont particulièrement prégnants au sein de la syllabe. En ne tenant pas compte de cette réalité de production, la méthode articulatoire ne risque-t-elle pas d'enseigner des phonèmes éloignés de la réalité de production ?

De plus, si la prononciation est un savoir faire (Harmegnies *et al*, 2005) : ce type de savoir peut-il s'acquérir par un apprentissage explicite ? Pour Krashen (1985), les connaissances acquises et les connaissances apprises appartiennent à deux systèmes distincts et seules les premières permettraient une maîtrise de la LE comparable à celle de la LM. Autrement dit, un apprentissage explicite de la prononciation ne pourrait donc pas garantir l'automatisation et la procéduralisation des connaissances.

Ces questionnements se trouvent au cœur de notre problématique et seront développés plus en détails dans la suite de ce travail (chapitre 4). Dans la

partie suivante, nous présenterons une autre méthode de correction phonétique, la méthode verbo-tonale (désormais MVT) qui propose une approche complètement différente de celle que nous venons de décrire. Celle-ci nous permettra d'apporter dorénavant et déjà des éléments de réponses aux questions préalablement soulevées.

1.3.2 La Méthode Verbo-Tonale d'intégration phonétique

1.3.2.1 Présupposés théoriques

La MVT s'intègre dans la méthodologie d'enseignement-apprentissage *Structuro-Globale Audio Visuelle* (SGAV), développée par P. Guberina (Université de Zagreb) et P. Rivenc (ENS de Saint-Cloud) dans les années 60. La méthodologie SGAV est fondée sur plusieurs principes : la priorité à la communication orale, en prenant en compte les systèmes vocaux, verbaux et gestuels ; la construction graduelle et globalisatrice du sens ; le décalage du passage à l'écrit, un minimum de soixante heures d'enseignement exclusivement centré sur l'oral était requis avant d'aborder le code écrit et l'association d'enregistrements sonores à des séquences d'images (Renard, 1979 ; Cuq, 2003).

Les recherches de Guberina sur la réhabilitation de déficients auditifs vont l'amener à élaborer un nouveau système d'audiométrie, le système verbo-tonal ('verbo' se référant à la parole et 'tonal' aux fréquences auxquelles notre oreille est sensible). L'audiométrie verbo-tonale a pour objectif d'évaluer la zone de fréquences où l'individu structure le mieux son audition en utilisant des stimuli de parole comme les logatomes (messages phoniques sans signification). Les logatomes proposés sont fréquemment équilibrés, c'est à dire qu'ils sont perçus par l'oreille normale sur une seule octave. Deux types de logatomes sont utilisés : les premiers (série A) sont filtrés sur les bandes optimales de l'oreille normale et les autres (série B), non filtrés, permettent d'établir les seuils d'audition, de distinction et d'intelligibilité (figure 10).

Série A :	37,5	—	75 Hz	bru-bru
	75	—	150	mu-mu
	150	—	300	bu-bu
	300	—	600	vɔ-vɔ
	600	—	1 200	la-la
	1 200	—	2 400	ke-ke
	2 400	—	4 800	ʃi-ʃi
	4 800	—	9 600	si-si
Série B :	50	—	100 Hz	bru-bru
	100	—	200	mu-mu
	200	—	400	bu-bu
	400	—	800	vɔ-vɔ
	800	—	1 600	la-la
	1 600	—	3 200	ti-ti
	3 200	—	6 400	ʃi-ʃi
	6 400	—	12 800	si-si

Figure 10 logatomes proposés filtrés (série A) et non filtrés (série B) (Renard, 1979:47)

Les hypothèses qui accompagnent l'audiométrie verbo-tonale sont qu'une oreille malade ne va pas percevoir le logatome filtré par l'oreille saine et qu'un logatome non filtré sera mieux perçu que le filtré par l'oreille malade car cette dernière structure différemment. Autrement dit, une personne déficiente auditive n'est pas totalement « sourde » mais structure ce qu'elle entend en fonction des seules fréquences qu'elle perçoit. Par exemple, un malentendant qui présente des pertes de fréquences dans les fréquences graves entendra mieux [t] que [k], [p] que [m] etc. Le système verbo-tonal s'inscrit dans le courant de la *gestalt théorie*. Selon ce courant, chaque individu joue un rôle actif dans la manière de structurer et d'organiser les formes qui l'entourent.

L'objectif du praticien est donc de trouver comment l'oreille malade structure ce qu'elle entend, pour proposer un filtrage adapté, par le biais d'appareils comme le Suvaglingua (ou Suvag), appareil électronique équipé de deux filtres fréquentiels (bas et haut) permettant de supprimer certaines composantes d'un son pour ne conserver que les fréquences dites optimales (figure 11).



Figure 11 Exemple d'un appareil Suvag et ses accessoires
http://www.suvag.com/fr/produits/suvag_it2.html

Mais que sont ces fréquences optimales ? À quoi correspondent-elles en termes de phonétique acoustique et perceptive ?

1.3.2.2 Les optimales

Selon Guberina, il devrait exister une bande de fréquences, dite bande de fréquences optimales, qui assure l'intelligibilité maximale de chaque son.

Pour déterminer ces bandes optimales, on fait entendre un son cible à un sujet à l'aide d'un magnétophone (enregistrement d'une voix authentique). L'examineur va ensuite filtrer le son en utilisant successivement toutes les combinaisons proposées par l'appareil. Le sujet doit à chaque fois écrire ce qu'il entend. Dans une autre tâche, il est demandé au sujet de répéter ce qu'il a entendu. Enfin, le son est transmis sans aucun filtrage puis filtré et le sujet doit dire dans quel cas il croit reconnaître le son cible. Notons que les sons cibles ne sont pas toujours présentés isolément mais également intégrés dans des petites phrases (Renard, 1979).

Il faut savoir que la détermination des optimales dépendrait de la culture linguistique (notion de crible) et de l'âge. Guberina a montré qu'en français, l'optimale de [u] se situe à l'âge de 15 ans dans l'octave 300-600 Hz, à 25 ans dans l'octave 200-400Hz et à 40 ans dans l'octave 150-300Hz.

Des recherches ont été menées dans le but de déterminer les zones fréquentielles nécessaires et suffisantes pour qu'un francophone identifie tel ou tel phonème (Landericy & Renard, 1975, 1977). Ces recherches ont montré

que pour chaque phonème il existerait un filtrage assurant un score de reconnaissance au moins égal à celui du « canal direct ». Il serait également possible de déterminer les fréquences dont la suppression élimine certaines confusions. Ces travaux s'appuient sur l'idée selon laquelle la perception de la parole serait discontinue et que seuls les éléments optimaux des stimuli seraient traités par le cerveau.

La notion d'optimales suscite des réserves chez les phonéticiens. À notre connaissance, peu d'études se sont intéressées à ce phénomène en particulier. Il est cependant possible de rapprocher le questionnement sur la perception des optimales avec des théories actuelles sur la perception, telle que la *théorie des aimants perceptifs* proposée par Kuhl dans les années 90. L'auteur défend l'idée selon laquelle tous les membres d'une même catégorie phonémique ne sont pas considérés comme équivalents par celui qui les perçoit, il y a les « bons » et les « mauvais » exemplaires.

« Selon Kuhl, les « bons » exemplaires, ou prototypes, ont un rôle central à jouer dans la structuration perceptive de l'espace phonique, dans la mesure où c'est à partir de ces prototypes que s'opère la catégorisation des sons de parole » (Nguyen, 2005 : 7)

Les prototypes exerceraient également une forte attraction sur les sons qui les entourent (figure 12). Par conséquent, il sera plus difficile de distinguer deux sons proches du prototype qu'en périphérie.

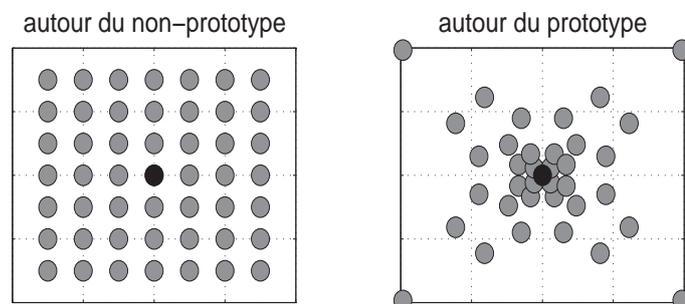


Figure 12 Représentation schématique des distances autour d'une voyelle non prototype et d'une voyelle prototype (représentées en noir) (Nguyen, 2005 : 8)

Les notions de *prototype* et d'*optimales* présentent toutes deux un point commun : l'existence de conditions perceptives idéales dont découlerait la catégorisation des sons de parole. Toutefois, la *théorie des aimants perceptifs* ne s'accorde pas avec l'idée d'une perception discontinue où seules les fréquences optimales seraient perçues, au contraire, selon cette théorie, l'auditeur serait capable de percevoir les bons comme les mauvais exemplaires dont la reconnaissance (ici discrimination entre deux sons) dépendrait de la position de l'exemplaire vis à vis du prototype.

La plupart des modèles de traitement de la parole tendent à s'accorder sur ce dernier point et considèrent en effet qu' « *un son de parole est associé à une valeur continue pour l'auditeur, selon que ce son représente plus ou moins bien la catégorie phonémique correspondante* » (Nguyen, 2005: 7).

En résumé, si les optimales se présentaient comme une découverte majeure dans les années 60 autant pour la rééducation des malentendants que pour l'apprentissage des langues étrangères (Guberina *et al*, 1965 ; Guberina, 1973) la validité scientifique de la théorie des optimales doit être discutée dans une perspective plus actuelle, à la lumière de nos connaissances en perception et en cognition. De plus amples recherches sont donc nécessaires pour pouvoir approfondir cette notion.

Mais revenons aux applications didactiques des recherches de Guberina via la description de la Méthode Verbo-Tonale (MVT) et de ses procédés de correction phonétique.

1.3.2.3 Applications didactiques

La MVT a pour origine des recherches appliquées à la rééducation des malentendants. Frappé par le caractère systématique des erreurs phonétiques commises par les apprenants de langue étrangère, Guberina va faire un parallèle entre le sourd qui apprend sa LM et l'apprenant de LE : chacun des deux structurerait la matière phonique en fonction de son propre système.

« (...) chaque apprenant/patient opère sa propre structuration de la matière phonique à partir de l'input qui lui est fourni par ses interlocuteurs mais également à partir de celui que lui renvoie sa propre boucle-audio phonatoire lorsqu'il parle » (Baqué & Estrada, 2010 : 55).

Si pour le système du patient est déterminé par des facteurs physiologiques (perte d'audition de certaines fréquences) dans le second de l'apprenant, il s'agit de son crible phonologique qui filtrerait les sons de la LE via les catégories de la LM. La métaphore du crible phonologique a été renforcée par les travaux de Kuhl (1991) et Best (1994 ; Best *et al*, 2001) en perception de la parole.

La MVT postule l'existence d'un lien étroit entre la production et la perception : une bonne perception précéderait et permettrait la production. Ainsi contrairement à la méthode articulatoire, la MVT est basée sur l'idée selon laquelle il serait nécessaire de « rééduquer » l'oreille de l'apprenant pour améliorer sa perception. Il est important de préciser que l'existence d'un lien direct entre la perception et la production en LE est toutefois à nuancer à la lumière de récentes études en perception de la parole (Billières *et al*, 2005 ; Dufour & Nguyen, 2010) : il est possible de produire des sons que l'on ne perçoit pas et inversement.

Néanmoins, la MVT est l'une des rares méthodes de correction phonétique à s'être intéressée au système perceptif de l'apprenant. De plus, les résultats spectaculaires observés aussi bien dans les rééducations des pathologies du langage qu'en didactique, et mérite qu'on s'intéresse de plus près à cette méthode.

1.3.2.4 La correction phonétique par la méthode verbo-tonale

Nous rappelons que l'objectif de la MVT est de rééduquer la perception des apprenants de LE. Le point de départ de cette rééducation est constitué par l'erreur de l'apprenant.

Selon la MVT, une erreur de prononciation peut être due à une erreur sur l'axe de la tension ou sur l'axe clair sombre. La tension se traduit par l'énergie neuromusculaire dépensée pour produire de la parole. Il peut arriver qu'un son soit produit avec une tension excessive (T+) ou inversement un relâchement trop important (T-) (figure 13).

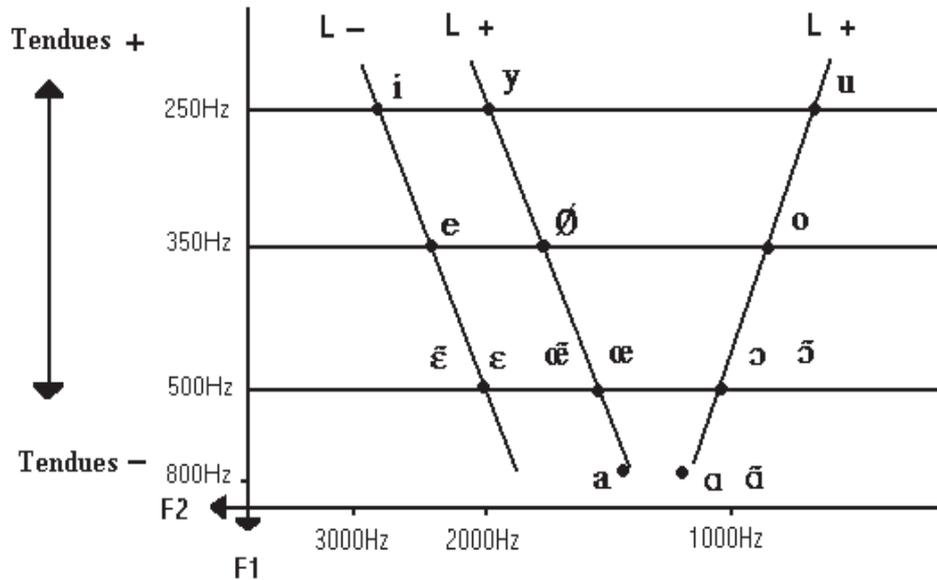


Figure 13 Classement des voyelles du français sur l'axe de la tension

L'axe clair-sombre correspond au timbre du son. Le timbre du son dépend de la répartition de l'énergie dans les différentes fréquences qui le composent. Ainsi, si les fréquences hautes reçoivent plus d'énergie, le timbre du son serait dit clair (C+), inversement si ce sont les fréquences basses qui reçoivent le plus d'énergie le timbre du son sera dit sombre (C-). La répartition de l'énergie sur certaines fréquences est déterminée par la taille de la cavité buccale. Une petite cavité va favoriser les fréquences élevées tandis qu'une grande cavité va favoriser les fréquences basses, comme pour les instruments de musique (figure 14).

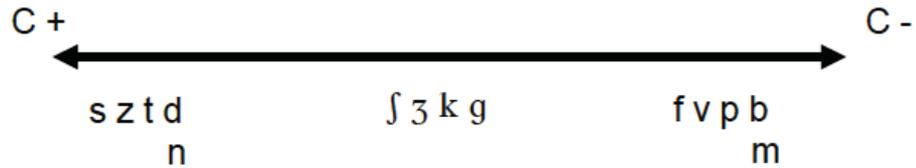


Figure 14 Classement des consonnes du français sur l'axe clair-sombre.

L'enseignant va donc corriger les erreurs des apprenants en favorisant l'un ou l'autre de ces axes.

Contrairement à la méthode articulatoire, la MVT propose une correction adaptée au système de fautes réel et non pas supposé de l'apprenant de LE. Ce système est lié non seulement aux caractéristiques de la LM de l'apprenant (*cf.* crible phonologique) mais aussi aux caractéristiques individuelles de la construction de l'interlangue. L'enseignant doit donc proposer une correction personnalisée et adaptée à chaque apprenant, dans le but d'amener ce dernier vers l'auto-corrrection. Ce fonctionnement renforce généralement la motivation des apprenants de LE.

La MVT repose sur des principes généraux comme la prise en compte du contexte sonore dans la correction phonétique, la priorité accordée à la prosodie, la prise en compte de la gestualité et le refus de l'intellectualisation

Selon la MVT, un son ne doit jamais être corrigé de manière isolée mais toujours dans un contexte sonore. Nous avons vu précédemment que l'unité de production étant la syllabe, les phonèmes sont particulièrement soumis aux phénomènes de coarticulation et de compensation à l'intérieur des syllabes. Ainsi un /u / prononcé après un /b/ (consonne C-) sera perçu avec un timbre plus sombre que s'il avait été prononcé avec un /s/ (consonne C+) On appelle ce phénomène une transition formantique : les formants de la voyelle sont orientés différemment selon la consonne qui la précède. La phonétique acoustique et la phonétique perceptive ont mis en évidence l'existence de ces variations contextuelles et leur rôle dans l'accès au lexique (Meunier, 2005). Pour corriger les phonèmes, il est possible d'avoir recours aux éléments suprasegmentaux : une intonation montante aidera à percevoir un son avec un

timbre plus clair alors qu'une intonation descendante favorisera l'assombrissement. De même un débit rapide augmentera la tension, tandis qu'un débit lent la diminuera. Padeloup (2004) a d'ailleurs mis en évidence le lien entre le débit et l'accentuation, l'accentuation influençant le timbre d'un son.

Les éléments suprasegmentaux constituent une priorité ; l'accentuation, le rythme et l'intonation sont des supports précieux de la mémoire auditive. Nous avons précédemment beaucoup insisté sur le rôle de la prosodie dans la production et la perception de la parole (partie 1.2.2). Certains auteurs pensent que tout comme il existe un crible phonologique, il existerait un crible prosodique (Intravaia, 2000 ; Dodane, 2003). La MVT se propose d'agir en priorité sur la correction des éléments suprasegmentaux. Pour corriger des erreurs de rythmes, par exemple, il existe plusieurs techniques comme la scansion ou le découpage régressif. L'intonation sera corrigée par le recours aux logatomes ou la visualisation de la courbe intonative par un geste de la main.

La MVT met ainsi en avant la relation corps/phonation, complètement laissée de côté dans la méthode articulatoire. Il existe en effet un lien entre la micro-motricité des organes de la phonation et la macro-motricité du grand corps (Fonagy, 1990). De même pour Intravaia (1993), les schémas rythmico-mélodiques trouveraient leur enracinement dans le système gestuel de la culture de la langue cible. Le recours aux gestes (modelage de la courbe intonative par le geste de la main) faciliterait la production d'un rythme, d'une intonation ou d'un son défectueux.

En MVT, l'oral est avant tout un comportement. Il s'agit donc d'intégrer les structures phonétiques du français au niveau inconscient. En effet, seule une compétence apprise de manière la plus implicite possible (on ne peut pas réellement parler d'implicite total dans un cours de langue) permettrait à l'automatisation et la procéduralisation (Paradis, 2004 ; Germain & Netten, 2005). De plus, le refus de l'intellectualisation permet de travailler dès le niveau débutant. Comme nous le soulignons dans la critique de la méthode articulatoire, il est important de proposer un travail de correction phonétique dès le début de l'apprentissage, pour éviter la mémorisation d'erreurs de

prononciation (phénomène dit de *fossilisation*). Selon les principes du SGAV, le passage à l'écrit doit également être retardé pour éviter que l'apprenant adopte une approche analytique de la langue.

En résumé, contrairement à la méthode articulatoire, la MVT offre une approche de l'enseignement de la prononciation dans un cadre méthodologique cohérent vis à vis de recherches en perception de la parole et en psycholinguistique. Elle tient compte des éléments fondamentaux de la production et de la perception de la parole comme les aspects suprasegmentaux ou la multimodalité (recours aux gestes facilitants). Et l'approche non intellectualisée de la prononciation rejoint les récentes hypothèses psycholinguistiques sur l'acquisition et le développement de connaissances procédurales.

En dépit de ces nombreux atouts, la MVT n'a jamais été validée expérimentalement. Dans le cadre de ce travail de thèse, nous tâcherons de remédier à ce manque en proposant une étude comparée des deux méthodes que nous venons de décrire dans des conditions expérimentales contrôlées. Nous testerons deux des principes fondamentaux de la MVT en lien avec les hypothèses psycholinguistiques sur les connaissances procédurales : a) l'apprentissage de la prononciation ne doit pas être intellectualisé et b) le passage à l'écrit viendrait perturber la mise en place des stratégies d'encodage et de décodage.

1.4 SYNTHÈSE DU CHAPITRE 1

Ce chapitre nous a permis d'illustrer le paradoxe actuel entre la place théorique accordée à l'oral dans la didactique et le fait que l'oral ne soit pas considéré comme un objet d'enseignement. Ainsi, les aspects propres à l'oralité comme la maîtrise des phonèmes et de la prosodie sont laissés de côté par les pédagogues. L'impact de l'enseignement de la prononciation en LE est minimisé aux prétextes que la maîtrise des sonorités de la LE n'est plus accessible aux apprenants tardifs ou bien que la prononciation s'acquiert par simples contacts avec la langue cible.

Ces présupposés ne sont pourtant pas validés expérimentalement. Au contraire, de récentes recherches en perception mettent en évidence que les mécanismes permettant l'acquisition du langage resteraient intacts tout au long de la vie (Flege, 1995 ; 2005). Toutefois les capacités perceptives seraient dépendantes des catégories créées pour la LM (Kuhl, 1995). La LM agirait comme un crible à travers lequel seraient filtrés les sons de la LE (Troubetzkoy, 1938 ; Renard, 1979 ; Billières, 2002). Les éléments segmentaux ne sont pas les seuls concernés dans la mesure où il existerait également un crible prosodique (Intravaia, 2000 ; Dodane, 2003) et nous savons que la prosodie joue un rôle déterminant dans la compréhensibilité du discours (Rasier & Hiligsmann, 2007).

Nos expériences de classes vont également à l'encontre de ces présupposés : faute d'un apprentissage spécifique, un apprenant de niveau avancé n'a pas acquis une bonne prononciation. De récentes études (Bongaerts *et al*, 2000) montrent que des apprenants de LE peuvent acquérir une prononciation perçue comme native via un entraînement intensif en perception et en production de la LE.

La Méthode Verbo-Tonale d'intégration phonétique prend en compte cette dimension perceptive dans la correction des sons de la LE. En effet, contrairement à la méthode articulatoire, plus ancienne et plus populaire en classes de langue, cette méthode est basée sur la modification progressive du crible phonologique et du crible prosodique de l'apprenant dans le but d'améliorer la production de ce dernier. La correction n'est pas proposée à priori mais en fonction des erreurs réellement produites par l'apprenant. De plus, cette méthode est l'une des rares méthodes de correction phonétique qui intègre les éléments fondamentaux de la production et de la perception comme la prosodie et la multimodalité. La MVT s'appuie enfin sur la non-intellectualisation de l'enseignement de la prononciation, considéré comme un savoir-faire. La méthodologie SGAV dans laquelle s'inscrit la MVT, postule également que l'apprentissage de l'écrit devrait être différé pour éviter qu'il ne vienne perturber l'acquisition de la prononciation.

En dépit de résultats empiriques très prometteurs la MVT n'a jamais été validée expérimentalement. L'un des objectifs de ce travail est donc de palier à ce manque en comparant la MVT et la méthode articulatoire dans des conditions expérimentales contrôlées.

Dans le chapitre suivant, afin d'appuyer notre choix méthodologique, nous développerons plus en détails le rôle de la prosodie dans l'acquisition du langage et plus particulièrement celui du rythme, prééminent à l'intonation et donc ossature de la prosodie.

Chapitre 2 : Rythme et apprentissage d'une langue étrangère

« Le rythme est dans le temps ce que la symétrie est dans l'espace »

Eugène d'Eichtal

Dans le chapitre précédent, nous avons vu qu'en dépit de la place accordée à l'oral, l'enseignement de traits de l'oralité était trop souvent négligé par la didactique des LE. Ainsi, l'enseignement de la prononciation est laissé de côté au prétexte que la maîtrise des sonorités de la LE s'acquiert naturellement par simple contact avec la langue cible.

Parmi les quelques méthodes de correction phonétique existant, rares sont celles qui, comme la Méthode Verbo-Tonale, prennent en compte les éléments suprasegmentaux - accentuation, rythme et intonation - dans l'apprentissage, alors que ces derniers sont à la base de la structuration de la langue (Di Cristo, 1981 ; Vaissière, 1997 ; Di Cristo, 2004).

Dans ce chapitre, nous nous intéresserons plus particulièrement au rôle du rythme dans l'apprentissage des LE. Nous verrons que celui-ci peut être considéré comme l'élément de base de la structuration prosodique (Di Cristo & Hirst, 1993).

Après avoir posé une définition du rythme et de ses différents constituants, nous développerons le concept d'isochronie qui a permis un classement des langues selon leurs différences rythmiques. Ce concept controversé sera discuté à la lumière de récentes recherches.

Nous clôturerons ce chapitre par la question de l'existence de transferts rythmiques entre la LM et la LE, et le rôle du rythme dans l'interlangue des apprenants.

2.1 INTRODUCTION

Proposer une définition exhaustive du rythme n'est pas aisé. En effet, si nous pouvons instinctivement caractériser le rythme comme un phénomène périodique, les différents éléments qui le composent sont en réalité très complexes et interviennent à différents niveaux.

Le rythme n'est d'ailleurs pas spécifique à la parole. On parle de rythme en médecine, rythme cardiaque ou respiratoire, en danse, en musique, pour la marche ou encore l'écriture. Le rythme verbal est lui-même un phénomène physique et psychobiologique avant d'être une construction linguistique : il y a une tendance biologique aux regroupements sur la base d'un tempo naturel (Fraisse, 1974 ; Lacheret-Dujour & Beaugendre, 1999).

Plus que tout autre phénomène de parole, le rythme est en réalité contraint par des négociations entre production et perception (Wioland, 1986 ; Lacheret-Dujour & Beaugendre, 1999). Pour Fraisse (1974), la structuration temporelle du rythme est un phénomène perceptif :

« L'action simultanée de l'assimilation et de la distinction a pour effet de simplifier le donné perceptif. Les différences entre des éléments voisins par la durée, la longueur, la forme sont diminuées ou supprimées par assimilation et les différences notables sont exagérées, ce qui supprime toute équivocité de la structure différente. Le nombre d'éléments différents est alors ramené à deux ou trois, c'est à dire que nous pouvons traiter sans difficulté » (Fraisse, 1974 : 111).

Autrement dit, la régularité rythmique est induite par nos capacités perceptives. Cette régularité implique la périodicité et le retour, à intervalles réguliers, d'un ou de plusieurs phénomènes identiques (ou du moins perçus comme tels). Mais de quels phénomènes s'agit-il ? Que structure le rythme ?

Pour Caelen-Haumont (1983), le rythme est considéré comme une structuration de durées. Pour Dell (1984), au contraire, la structure rythmique est composée par les schèmes accentuels. Di Cristo & Hirst (1993 ; 1997) considèrent le rythme comme « *l'organisation temporelle des proéminences* ».

La proéminence étant définie comme un phénomène perceptif permettant la mise en relief d'un ou plusieurs éléments d'un énoncé aux dépens des autres, quelque soit sa fonction dans l'énoncé ou sa réalisation (Vihanta, 1993).

La définition proposée par Di Cristo & Hirst permet de dépasser l'opposition entre structuration temporelle et structuration accentuelle – l'accent étant considéré comme un fait local de proéminence – et de hiérarchiser les constituants du rythme en posant le rythme comme révélateur de l'accentuation.

Mais comment sont hiérarchiser les autres constituants prosodiques, comme l'intonation? Quelle est la place du rythme dans la prosodie ?

2.1.1 La place du rythme dans la prosodie

La prosodie est constituée par différents éléments, aussi appelés suprasegmentaux, qui sont l'accentuation, le rythme et l'intonation.

Il existe différentes manières de définir les paramètres prosodiques, selon qu'on les considère sur le plan de la production, sur le plan acoustique ou sur le plan perceptif. Au niveau acoustique, l'étude des paramètres se base sur ce qui est mesurable. Il est donc possible d'étudier chacun des paramètres séparément. Inversement, il existe une forte interaction entre les différents paramètres en production et en perception (tableau 2).

Paramètres locaux	Acoustique	Perception
Accentuation (lexicale ou post-lexicale)* Ton* Quantité*	Fo, durée, amplitude Fo Durée des segments	Mélodie, longueur, intensité Mélodie Longueur
Paramètres globaux	Acoustique	Perception
<u>Rythme</u> : Accentuation (supralexicale) Tempo/ débit, pauses <u>Intonation</u> Segments tonals et contours mélodiques	Fo durée, amplitude Durée des événements et des silences Fo	Mélodie, longueur, intensité Longueur des événements et des silences Mélodie

Tableau 2 Représentation des paramètres locaux et globaux de la prosodie sur le plan acoustique et perceptif, les critères marqués par * ne s'appliquent pas au français

Lacheret-Dujour (2008) considère que de manière minimaliste, la prosodie s'articule autour de deux niveaux d'analyse : un niveau phonétique concret et un niveau phonologique abstrait.

- Au niveau phonétique, on trouve la **fréquence fondamentale (F0)**, qui correspond à l'estimation du son laryngien à partir du signal acoustique donné ; la **durée**, qui comprend le **débit de la parole** (nombre de syllabes par seconde) et les **pauses**, le **tempo** (accélération ou ralentissement du débit) et l'**intensité**, énergie contenue dans le signal de parole.
- Au niveau phonologique abstrait on trouve l'**accentuation** (processus local, portant sur le mot et la syllabe) et l'**intonation** (mécanisme global portant sur le groupe, phrase ou l'énoncé).

La définition du rythme adoptée par Di Cristo & Hirst (1993 ; 1997) permet en réalité de positionner le rythme comme élément prééminent non seulement à l'accentuation mais aussi à l'intonation.

« (...) il est nécessaire de ne pas concevoir le rythme seulement en tant que relation de 'proéminences accentuelles' mais aussi comme principe générateur des structures ou de groupements plus ou moins récurrents (...), ces derniers étant caractérisés à leur tour par leurs modes d'organisation interne » (Hirst & Di Cristo, 1993 : 1)

En effet, les proéminences (ou syllabes métriquement fortes) sont les points d'ancrage de l'intonation et des segments tonals. Ainsi, toute syllabe accentuée au moyen d'un changement de hauteur participe à la fois à manifestation du rythme et à celle de l'intonation. Autrement dit, l'intonation est basée sur une structure rythmique et le rythme est donné par l'accentuation. Intonation et accentuation correspondent donc à des niveaux prosodiques différents mais ne peuvent être dissociés.

L'approche paramétrique de l'intonation, adoptée par Di Cristo & Hirst (1993 ; 1997) permet de modéliser la hiérarchisation des différents constituants prosodiques sous la forme d'un arbre phonologique. L'idée centrale de l'approche paramétrique est que le rythme et l'intonation sont considérés comme les deux représentations d'une forme phonologique sous-jacente. Ainsi, dans toutes les langues du monde, les configurations acoustiques des paramètres prosodiques (la F0, la durée et l'intensité) seraient déterminées par une représentation phonologique qui comprendrait des segments tonals et une structure métrique (Di Cristo & Hirst, 1993).

Dans la version initiale du modèle, l'unité tonale (UT) correspondait aux groupes accentuels, c'est-à-dire à des séquences de syllabes bornées à droite par une syllabe accentuée (Di Cristo & Hirst, 1993 : 18-19). Mais comme il existe en français deux types d'accents rythmiques : un accent final (accent primaire) et un accent initial (accent secondaire) (Fònagy, 1979), Di Cristo & Hirst ont intégré à leur modèle une nouvelle unité, l'Unité Rythmique (UR) correspondant à une suite de syllabes délimitées par un accent primaire (Di Cristo & Hirst, 1993) (figure 15).

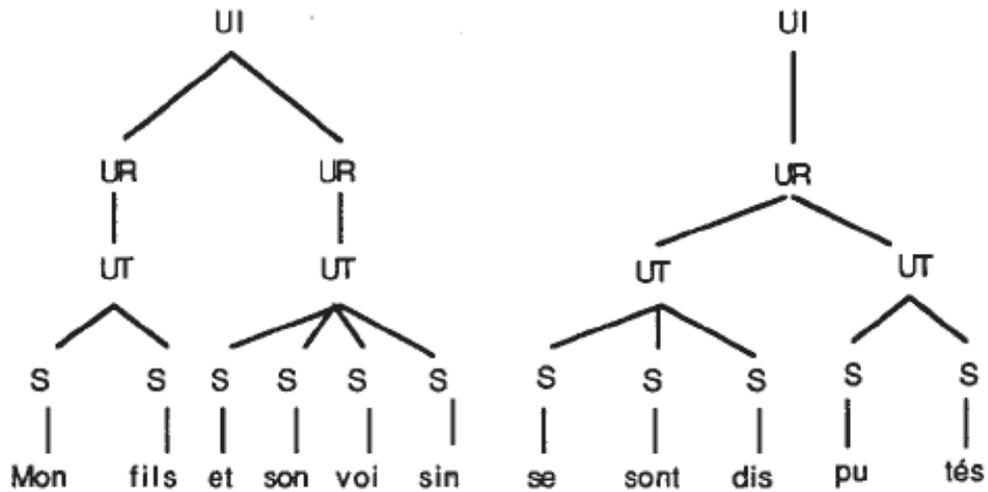


Figure 15 Représentation par un arbre phonologique de la structure prosodique d'un énoncé français, UI : Unité Intonative ; UR : Unité Rythmique; UT : Unité Tonale ; S : Syllabes (Di Cristo & Hirst, 1993 : 6)

Même si la représentation phonologique des éléments prosodiques ne constitue pas notre objet d'étude, il pourrait être intéressant dans de futures recherches d'appliquer ces modélisations à l'étude de l'interlangue (*cf.* partie discussion).

Dans le cadre de ce travail, nous adoptons la définition du rythme proposée par Di Cristo & Hirst (1993 ; 1997) et considérons le rythme comme élément prééminent à l'accentuation et à l'intonation et de fait, comme le noyau de la prosodie.

Nous recentrons donc notre cadre d'étude sur le rythme et ces différents constituants : l'accentuation, les pauses et le débit de parole.

2.1.2 Les constituants du rythme

2.1.2.1 L'accentuation

L'accent est la mise en valeur d'une syllabe par des variations acoustiques de fréquence, d'intensité ou de durée.

L'accent permet la structuration du message et a donc plusieurs fonctions : une fonction démarcative afin de marquer les frontières entre les

groupes rythmiques, une fonction contrastive pour mettre certaines unités en valeur, une fonction rythmique avec l'alternance des syllabes accentuées et inaccentuées et une fonction culminative puisque la force de l'accent indique la hiérarchie des constituants prosodiques.

Il existe des accents rythmiques primaires et secondaires dont la réalisation dépend de contraintes linguistiques et des accents dits sémantico-pragmatiques qui dépendent des choix du locuteur (accent nucléaire ou emphatique). Dans le cadre de ce travail nous attacherons seulement à la description des accents rythmiques primaires et secondaires en prenant l'exemple du français.

En français, l'accent primaire est un accent de durée. Une syllabe accentuée en français est deux fois plus longue qu'une syllabe inaccentuée (Delattre, 1966 ; Léon, 1992). Nous noterons que la durée n'a pas toujours constitué un paramètre pertinent pour déterminer l'accent car l'allongement devant une frontière prosodique est un phénomène universel (physiologique). Plusieurs facteurs peuvent avoir une influence sur la durée, nous n'en citerons que quelques uns : les facteurs extralinguistiques, syntaxiques, lexicaux (classe grammaticale du mot), phonologiques (nombre de phonèmes dans la syllabes, type de voyelle *etc.*) mais aussi pragmatiques (selon le type de discours, *cf.* Astésano, 2001).

Le français est décrit comme une langue oxytonique (l'accent primaire tombe sur la dernière syllabe du groupe rythmique) mais de nombreux chercheurs se sont attachés à démontrer l'existence d'un autre type d'accent, l'accent secondaire (initial) qui permet la formation d'un arc accentuel entre le début et la fin du groupe rythmique ou du mot (Fonagy, 1979). Cet accent initial secondaire est manifesté par une proéminence de hauteur. Cet accent participe également à l'organisation rythmique (Pasdeloup, 1990) et a une double fonction : une fonction d'indication de la structure prosodique et une fonction de démarcation des unités lexicales.

Nous noterons que l'accent secondaire initial tend à s'effacer au profit de l'accent primaire final (Di Cristo & Hirst, 1993).

2.1.2.2 Les pauses

Goldman-Eisler (1968) définit les pauses comme des interruptions du mouvement articulatoire. Elle détermine trois types de pauses : les pauses liées à des phénomènes articulatoires, comme par exemple lorsque deux occlusives se suivent, les pauses liées à des hésitations vocales et les pauses respiratoires. Les hésitations sont constituées par les pauses vides (ou pauses non sonores, d'une durée minimale de 250ms) et les pauses remplies (ou pauses sonores).

Les pauses ont plusieurs fonctions : elles participent à l'organisation et à la planification des énoncés et permettent la prise de souffle. Les pauses de souffle participent aussi à la programmation de l'énoncé puisqu'elles se situent à des moments clé de l'organisation du discours. C'est la structure morpho-syntaxique de la langue qui impose le rythme respiratoire de la parole et qui détermine du coup les groupes de souffle (Léon, 1992).

Pour Goldman-Eisler (1968) les pauses d'hésitations sont nécessaires au temps de programmation des énoncés. Les pauses remplies sont des interruptions momentanée, qui sont constituées soit par des pauses remplies par un silence et la suite sonore « *eu*h », soit par des syllabes allongées. Des chercheurs ont montré que les hésitations consistant à prolonger anormalement les syllabes en fin de mot, représentaient une caractéristique des langues à syllabation ouverte comme le français (Grosjean, 1980). De plus ces phénomènes de syllabes allongées ont en français presque la même ampleur que les hésitations vocales : elles représentent 43% des pauses sonores. Pour Guaïtella (1996) les hésitations participent à l'élaboration du système rythmique des productions spontanées et sont nécessaires en production comme en perception.

Grosjean et Deschamps (1972) ont montré que les trois quarts des pauses interviennent à des jonctures grammaticales, avec cependant une

différence selon le type de pauses : 38% des pauses sonores sont réalisées en dehors des points de ruptures grammaticaux contre 25 % des pauses non sonores. Ces résultats sont différents selon le type de parole. Des chercheurs ont montré qu'en parole spontanée seulement 55% des pauses se réalisent aux frontières grammaticales tandis que 45% se trouvent à l'intérieur de ces unités (Henderson *et al*, 1966). Les pauses en lecture oralisée se situent toujours aux frontières grammaticales. Les pauses peuvent marquer les différentes étapes de l'argumentation et aider à la compréhension du discours (Lacheret-Dujour & Beaugendre (1999)

« Ainsi, à la fonction syntaxique se superpose une fonction énonciative : la variation de durée des silences et largement dépendante de la structure thématique du discours et du temps requis pour une mise en relief des arguments » (Lacheret-Dujour & Beaugendre, 1999 :52)

Enfin les pauses jouent évidemment un rôle rythmique majeur, à différents niveaux : au niveau syntaxique (marquage de frontière) au niveau du discours (marquage de tours de parole, et des unités du discours) ; au niveau stylistique ; au niveau biologique (respiration) mais aussi psycholinguistique.

Pour certains chercheurs, le rythme de la parole serait même plus prégnant dans les moments de pauses car elles favoriseraient la perception du rythme (Astésano, 2001).

2.1.2.3 Le débit

Le débit correspond à la vitesse de parole exprimée en nombre de syllabes par seconde. Les variations du débit de parole serait en lien avec l'organisation temporelle des proéminences, autrement dit, avec les patterns rythmiques dépendant des variations de débit (Zellner, 1998 ; Astésano, 2001 ; Padeloup, 2004). Un débit rapide entraîne la réduction des degrés de frontières, voire la suppression de certaines frontières, et par conséquent, en français, des accents situés à ces frontières (Astésano, 2001).

Pasdeloup (2004) a étudié les contrastes entre la durée des syllabes accentuées et inaccentuées et la durée des pauses en fonction du débit de parole. L'auteur a enregistré un corpus de parole lu selon trois débits de parole : normal, très rapide et très lent. Les résultats de l'étude encore préliminaire indiquent qu'un débit lent renforce les contrastes de durée entre les syllabes accentuées et les syllabes inaccentuées et entraîne l'augmentation de la durée et du nombre des pauses vides et de pauses sonores. Toutefois, le contraste observé entre les différents types de syllabes en débit lent est beaucoup plus important qu'en débit normal ou rapide (tableau 3), ce qui amène l'auteur à conclure que la structuration temporelle d'un énoncé à un débit lent ne traduit pas un ralentissement général du même énoncé à un débit plus rapide. En effet, l'auteur observe une faible variation de la durée des syllabes inaccentuées entre les débits lents et rapides et postule donc que la syllabe inaccentuée jouerait un rôle déterminant dans l'organisation rythmique du français.

	Débit lent	Débit normal	Débit rapide
Syllabes inaccentuées	158ms +/- 38	137ms +/- 35	118ms +/- 32
Syllabes accentuées	351ms +/-142	270ms +/-109	194ms +/-57
Pauses	714ms +/- 692	431ms +/- 469	292ms +/-131

Tableau 3 Moyenne des durées syllabiques et des pauses dans les trois conditions de débit, (Pasdeloup, 2004 : 2)

Pour Dellwo (2006) toutes les langues ne seraient pas influencées de la même manière par les variations de débit, en fonction de la famille rythmique à laquelle elles appartiennent. Notons que cette notion de *famille rythmique* est sujette à débats.

Dans la partie qui va suivre, nous reviendrons sur les limites de la séparation traditionnelle entre les langues à rythmicité syllabique et les langues à rythmicité accentuelle, avant de nous intéresser plus précisément aux différences rythmiques de l'anglais et du français,

2.2 RYTHME ET LANGUES

2.2.1 Principe d'isochronie : rythmicité syllabique et rythmicité accentuelle

Pendant de nombreuses années, la majorité des études sur le rythme concernaient l'anglais. Les modèles proposés pour les autres langues étaient élaborés par rapport à la langue anglaise.

Pike (1945) est le premier à avoir distingué les langues à rythmicité accentuelle (comme l'anglais) dont la structure rythme est basée sur le retour de l'accent à intervalles de temps régulier, et les langues à rythmicité syllabique (comme le français) dans lesquelles le rythme est basé sur le retour de syllabes à intervalles de temps réguliers. Dans une langue dite accentuelle, la récurrence entre les syllabes accentuées serait donc constante. Inversement dans une langue syllabique, chaque syllabe serait isochrone, c'est-à-dire prononcée avec la même durée.

Abercrombie (1967) appuie les recherches de Pike et postule que ce principe d'isochronie est en lien avec la pulsation pulmonaire. Ainsi, il s'agirait d'un phénomène universel et toutes les langues du monde appartiendraient à l'une ou l'autre de ces catégories.

D'autres auteurs (Bloch, 1950 ; Ladegoged, 1975) ont proposé une troisième catégorie rythmique, les langues moraïques dont la rythmicité est associée à la périodicité des syllabes en fonction de leurs poids (mores).

Le tableau 4 illustre la classification des langues en fonction de ces trois rythmicités.

Langue accentuelles	Langues syllabiques	Langues moraïques
<p>Anglais</p> <p>Allemand</p> <p>Arabe</p> <p>Néerlandais</p> <p>Russes</p> <p>Polonais</p> <p>Thaï</p>	<p>Français</p> <p>Espagnol</p> <p>Italien</p> <p>Grec</p> <p>Catalan</p> <p>Telougou</p> <p>Yoruba</p> <p>Hindi</p>	<p>Japonais</p> <p>Tamoul</p>

Tableau 4 Classification des langues selon leur classe rythmique (Tortel, 2009 : 44)

Wenk & Wioland (1982) rejettent ce classement. Si on applique le principe de l'isochronie syllabique au français, cela voudrait dire qu'un francophone mettrait deux fois plus de temps pour articuler douze syllabes que six. Ces auteurs ont donc demandé à douze francophones natifs de lire la phrase suivante « *Il a sollicité ma collaboration, car Pierre aime toujours l'art* ». Selon la théorie de Pike, la première séquence devrait être plus longue que la seconde. Les résultats de cette étude montrent au contraire qu'un équilibre temporel se crée entre la première et la deuxième partie de la phrase. Il semblerait que ce soit les Groupes Rythmiques (désormais GR) et non pas les syllabes qui assurent l'équilibre temporel du français (*Ibid*). Léon (1993) définit les GR comme des syntagmes actualisés dans la parole et terminés par une syllabe accentuée C'est pourquoi, c'est le nombre de groupes rythmiques (donc d'accents primaires) qui est responsable de l'organisation rythmique du français.

Les syllabes françaises sont donc produites et perçues à l'intérieur des GR et non pas isolément. C'est pourquoi, Wenk & Wioland délaissent les termes d'isochronie et d'isosyllabité pour leur préférer les termes de *trailer-timed* (copochrone) et *leader timed* (codachrone) pour caractériser respectivement le français et l'anglais. En effet, si en français, c'est la durée de la dernière syllabe accentuée (trailer) qui permet de déterminer les GR, le rythme anglais est quand

à lui délimité par l'occurrence de syllabes métriquement fortes au début de chaque groupe (leader).

Plusieurs auteurs comme Dauer (1983) penchent plutôt pour l'existence d'un continuum rythmique (figure 16).



Figure 16 Continuum rythmique (Dauer, 1983 : 60)

Sur la base de mesures acoustiques, Dauer (1983) postulent que les différences observées entre les langues seraient la conséquence de la structuration interne de l'intervalle (i.e la complexité des composantes syllabique, l'absence de réduction vocalique et l'importance de l'accent). L'auteur distingue les langues plus ou moins syllabiques composées de syllabes simples, de taille stable, sans réduction vocalique, et les langues plus ou moins accentuelles qui présentent une composition syllabique plus complexe et variée, en particulier sous l'accent tonique.

De récentes études ont proposé de tester expérimentalement l'existence de ces différences rythmiques et d'un continuum.

Ainsi, Ramus *et al* (2003) proposent de tester la séparation entre rythmicité syllabique et rythmicité accentuelle ainsi que l'existence de langues dite intermédiaires, qui ne correspondraient ni à l'une ni à l'autre de ces distinctions. Les langues choisies étaient : l'anglais, l'allemand, l'espagnol, le catalan et le polonais. Les auteurs ont constitué un corpus de 20 phrases lues pour chaque langue, ont resynthétisé les phrases de manière à ce que seuls les éléments prosodiques soient perçus, puis ils ont soumis 168 francophones natifs à une tâche de catégorisation et une tâche de discrimination. Les sujets qui suivaient une phrase d'entraînement, pensaient devoir identifier et discriminer deux langues exotiques, le Sahatu et le Moltec. Les résultats ont montré que toutes les combinaisons de paires de langues ont été discriminées sans problème, sauf en ce qui concernait les oppositions anglais-allemand et

espagnol-catalan. L'opposition anglais-espagnol n'a posé aucun problème aux auditeurs, illustrant que les différences entre langues dites à rythmicité accentuelle et les langues à rythmicité syllabique seraient perçues. Le polonais a autant été différencié de l'anglais que de l'espagnol, appuyant l'hypothèse selon laquelle cette langue ferait partie des langues intermédiaires.

Grabe & Low (2002) apportent des preuves objectives, via des mesures de durée, à la classification des langues. Pour ces auteurs, la rythmicité est liée à la variabilité vocalique et intervocalique. La comparaison des mesures de variabilité avec le classement des langues leur a permis de mettre en évidence que l'anglais, l'allemand et le hollandais, classées comme des langues isoaccentuelles, présentaient des taux élevés de variabilité, tandis que le français et l'espagnol, langues isosyllabiques, présentaient des taux de variabilité bas.

Tortel (2009) formule l'hypothèse selon laquelle les différences rythmiques entre une LM syllabique et la LE accentuelle devraient se retrouver dans l'interlangue d'apprenants francophones d'anglais. L'auteur postule également qu'il serait possible de différencier les natifs des apprenants à l'aide d'une métrique rythmique basée sur des mesures de coefficients de variation, de variabilité vocalique et consonantique et de structures syllabiques. Les résultats concernant les productions sont en contradiction avec les hypothèses formulées vis à vis des résultats obtenus par Ramus et al (2003) et Grabe & Low (2002). Tortel observe toutefois que l'ensemble des paramètres rythmiques permet de classer correctement les locuteurs dans les groupes 'natif' ou 'apprenant' (score global de 69,5%). Les résultats de Tortel soulèvent la question de l'impact de la LM et de la LE dans l'interlangue des apprenants. Existerait-il un rythme propre à l'interlangue ? Nous y reviendrons.

Bien que notre travail n'ait pas pour objectif de tester la validité de l'opposition isochronie accentuelle et isochronie syllabique, nous ne pouvons pas faire l'impasse sur cette notion qui participe à considérer l'anglais et le français comme deux langues rythmiquement opposées.

2.2.2 Différences rythmiques du français et de l'anglais

Delattre (1966) a montré que l'anglais et le français sont les langues les plus opposées sur le plan phonétique : le rythme de l'anglais est caractérisé par une alternance de temps forts et de temps faibles, tandis que celui du français serait défini par une tendance à l'isochronie, non pas syllabique mais de groupe rythmique. De plus, l'accent de l'anglais est un accent d'intensité alors que celui du français est caractérisé par la durée (syllabe deux fois plus longue sous l'accent, Delattre 1966).

Notons que le gabarit tonal, lié à la structure métrique, est représenté en anglais par une variation « haut/bas » (HB) alors que celui du français est représenté par une variation « bas/haut » (BH) (figure 17).

<p>Gabarits tonals (ou rythmiques) de l'anglais 'Haut-Bas'</p> <p>« The youngmen from New-York were leaving »</p> <p>(H B) (H B) (H B)</p> <p>Gabarits tonals (ou rythmiques) du français 'Bas-Haut'</p> <p>« Les jeunes de New-York parlaient »</p> <p>(B H) (B H) (B H)</p>
--

Figure 17 Exemples des gabarits tonals du français et de l'anglais

Cette différence est en lien de cause à effet avec la distinction proposée par Wenk & Wioland (1982) entre les langues *leader-timed* et les langues *trailer-timed*. Le Tableau 5 représente l'ensemble des caractéristiques des syllabes accentuées et des syllabes inaccentuées pour chacune de ces deux catégories rythmiques.

	Codachronie (Trailer-Timing) (Français)	Copachronie (Leader-Timing) (Anglais)
Position	Fin de groupe	Initiale de groupe
Syllabes accentuées	[+articulation explicite] [++allongement] [- diminution de l'intensité] Changements de hauteur (pitch) décalés.	[+articulation explicite] [+allongement] [+ augmentation de l'intensité] Changements brusques de hauteur (pitch)
Syllabes non accentuées	Tendues Voyelles peu centralisées	Relachées Voyelles centralisées

Tableau 5 Caractéristiques des langues Trailer-timed et des langues leader-timed (Wenk, 1983 : 156)

Afin de mieux comprendre les différences entre ces deux langues, il nous semble nécessaire de revenir sur des aspects plus généraux de l'organisation rythmique. Le rythme, nous l'avons déjà dit, est l'organisation des proéminences. La courbe rythmique est organisée par l'alternance de cycles de tension et de relâchement musculaire. La figure 17 illustre ce phénomène.

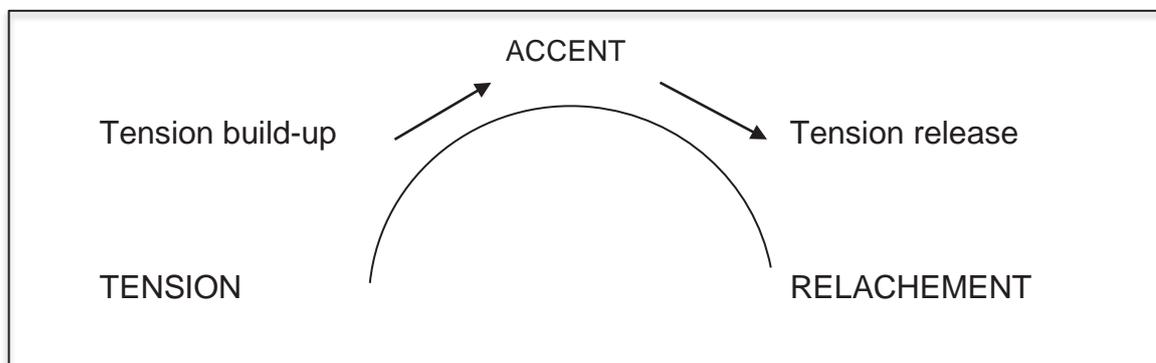


Figure 18 Représentation d'une courbe rythmique (Wenk, 1983: 157)

Les syllabes accentuées de l'anglais et du français correspondent à la tension musculaire maximale. Mais, dans les langues *leader-timed* les syllabes inaccentuées se situent dans la partie de relâchement de la tension musculaire, tandis que dans les langues *trailer-timed* elles se situent au début de la tension musculaire (Wenk & Wioland, 1982 ; Wenk, 1983). C'est pourquoi l'anglais a une tendance à l'instabilité vocalique (phénomène de diphtongaison) et à la

centralisation des voyelles pour les syllabes non accentuées, tendances que n'a pas le français.

On oppose aussi l'accentuation de l'anglais et du français en termes de fonction : en français l'accent est dit *démarcatif*, car il permet de délimiter -et donc d'identifier- les unités rythmiques et intonatives ; en anglais il est dit *culminatif* : l'accent se traduit par la mise en relief de la syllabe accentuée afin de faciliter l'intelligibilité de l'énoncé. En raison de son caractère non distinctif (au niveau lexical) et du syncrétisme entre l'accent et les frontières intonatives, il est difficile d'identifier l'accent primaire du français en lui-même et pour lui-même (Di Cristo, 1981). C'est pourquoi, les phonéticiens ont longtemps décrit le français comme une langue sans accent (Rossi, 1980).

En ce qui concerne les pauses, Grosjean & Deschamps (1972) ont comparé la structure des pauses en français et en anglais. Ces auteurs constatent que 85% des pauses non sonores en français sont en partie ou en totalité des pauses de prise de souffle, contre 34% en anglais. Cela s'explique par une tendance plus importante aux regroupements plus grands en français qu'en anglais.

D'autres part, en français les trois quarts des pauses interviennent à des jonctures grammaticales, avec cependant une différence selon le type de pauses vu que 38% des pauses sonores sont réalisées en dehors des points de ruptures grammaticaux contre 25% des pauses non sonores (Grosjean & Deschamps (1972). Goldman-Eisler (1968) propose une étude similaire avec anglais, en parole spontanée. Ses résultats montrent que seulement 55% des pauses anglaises se trouvent à des jonctions grammaticales. En effet, Grosjean (1980) met en avant une différence fondamentale entre la répartition des pauses en anglais et en français : en anglais le placement des pauses est plus libre qu'en français. En français, il est impossible de rencontrer des pauses à l'intérieur des syntagmes verbaux, comme c'est le cas en anglais (en raison du nombre important de verbes prépositionnels).

L'anglais aurait de plus tendance à produire plus de pauses que le français mais d'une durée inférieure.

Chaque langue possède donc ses spécificités rythmiques et il n'est pas possible de généraliser les résultats obtenus pour une langue à la description des autres langues, comme cela a longtemps été le cas pour l'anglais.

Il est vraisemblable que ces différences rythmiques auront des conséquences importantes pour l'apprentissage des LE et la construction de l'interlangue de l'apprenant.

2.3 CONSEQUENCES POUR L'APPRENTISSAGE DES LANGUES ETRANGERES

Rares sont les études s'intéressant aux conséquences de cette spécialisation rythmique sur l'apprentissage. Pourtant, comme le montrent Boysson-Bardies *et al* (1984), Levitt & Utman (1992) et Levitt (1993), les jeunes enfants montreraient déjà les spécificités rythmiques de leur langue maternelle dans leurs productions.

À propos des différences entre l'anglais et le français, Levitt (1993) s'est penchée sur l'acquisition des paramètres prosodiques de chacune de ces langues. Il a observé de nettes différences dans la réalisation des groupes rythmiques (GR) chez les enfants français et anglais : les premiers produisent des GR beaucoup plus longs (contenant au moins 4 syllabes) et plus réguliers que les seconds. En effet, Fletcher (1991) a mis en évidence que 56% des groupes rythmiques en français contenait au moins 4 syllabes contre seulement 6% pour l'anglais, qui aurait tendance à produire des GR beaucoup plus courts et moins réguliers.

Chez l'adulte, ces propriétés rythmiques constituent un paramètre suffisant pour que des sujets français soient capables de distinguer l'anglais du japonais. Dans son étude, Ramus (1997) a en effet resynthétisé des phrases en anglais et japonais, en préservant soit l'information phonétique, rythmique et intonative, soit l'information rythmique et intonative soit l'information rythmique ou l'information intonative seules. Les résultats ont montré que les sujets atteignaient des scores de reconnaissance comparables dans le premier et le troisième cas, alors que l'information rythmique était réduite au strict minimum.

Les spécificités rythmiques se mettent donc en place très tôt et seraient nécessaires et suffisantes à la discrimination des langues. Il est donc possible de s'interroger sur l'existence de transferts rythmiques entre la LM et la LE : Qu'est-ce qui détermine la nature de ces transferts ? L'interlangue de l'apprenant est-elle seulement le reflet des transferts entre LM et LE ?

2.3.1 Transferts rythmiques et interlangue de l'apprenant

Wenk (1985) s'est intéressé à l'interlangue des apprenants. Pour lui, lorsque les langues appartiennent à des familles rythmiques différentes (Wenk préfère l'opposition *trailer timed* et *leader timed*, à l'opposition syllabique et accentuelle), le passage d'un style rythmique à l'autre ne peut se faire sans une étape intermédiaire dans laquelle les productions portent les traces du rythme de la LM et de la LE.

White & Mattys (2007) se sont interrogés sur l'influence du rythme de la LM sur le rythme de la LE. Pour tester cela, ils ont comparé les productions en langue maternelle et en langue étrangère, en formulant l'hypothèse selon laquelle si le rythme de la langue maternelle et de la langue étrangère sont très différents (comme par exemple l'espagnol et l'anglais), le rythme des productions des apprenants devrait constituer une sorte de rythme d'interlangue : éloigné de celui de la LM mais sans correspondance avec celui de la LE. En revanche, si les deux rythmes sont proches (anglais et allemand), les auteurs s'attendent à observer dans les productions de LE un rythme correspondant à celui de la LM. Concernant la première hypothèse, les résultats sont plus controversés mais en ce qui concerne la seconde, il semblerait que les aspects communs ou similaires entre deux langues aient pour conséquence une production sur la base des règles rythmiques de la langue maternelle. On ne peut s'empêcher de faire le lien avec le modèle de la perception de la parole de Flege, selon lequel les éléments communs entre la LM et la LE seraient assimilés à des catégories déjà existantes de la LM.

Rasier & Hiligsmann (2007) partagent l'idée de Mennen (2006) selon laquelle les transferts prosodiques de la langue maternelle à la langue étrangère

touchent autant la perception que la production. En effet, pour Mennen (2006) les transferts prosodiques constituent des interférences à la fois phonétiques et phonologiques.

« The distinction between a phonetic and phonological component in intonation is important as it suggests that languages can differ at both these levels. As a results, the L1 and L2 intonation systems may influence one another both at the level of phonological representations as well as at the level of the phonetic implementation. A phonological influence would result from intonational differences in the inventory of phonological tune (Ladd, 1997). An example of phonological influence is the use of rises when native speakers would use falls and vice versa, found in many studies of L2 intonation (e.g Adams and Munro, 1978 ; Backman, 1979 ; Jenner, 1976 ; Lepetit, 1989 ; Willems, 1982). An example of phonetic influence is the finding of a different pitch range (e.g. Mennen, this paper) or a different slope of a rise (e.g. Ueyama, 1997) compared to the monolingual norm » (Mennen, 2006 : 4)

Dans leur étude, Rasier & Hiligsmann (2007) étudient le rôle de la LM dans l'acquisition des paramètres rythmiques comme le « pitch accent » chez des apprenants néerlandais et français. À l'instar de Selinker (1972), ils postulent que les difficultés d'apprentissage de la prosodie en langue étrangère ne peuvent se réduire à la comparaison des rythmes de la langue maternelle et de la langue étrangère ni à la description de l'interlangue des apprenants. Pour eux, il est nécessaire de comparer les productions des apprenants et des natifs afin de déterminer ce qui est correct ou non. En se basant sur des recherches antérieures, les auteurs formulent l'hypothèse selon laquelle une erreur courante des apprenants de langue étrangère serait d'emphatiser chaque mot de la phrase, car ils auraient des difficultés à distinguer les informations nouvelles des informations anciennes. Les différences accentuelles entre les deux langues expliqueraient les difficultés observées. L'accent serait en effet un des paramètres les plus difficiles à maîtriser en langue étrangère. Les résultats de l'étude montrent également un lien entre l'accentuation et les pauses, suggérant que les erreurs prosodiques sont cumulatives.

Mais certains rythmes seraient-ils plus faciles à acquérir que d'autres ?

Pour Konopczynski (1999) et Grabe *et al* (1999) la complexité de la structure rythmique aurait une incidence sur l'apprentissage de celle-ci. Une structure plus complexe et irrégulière serait plus difficile à acquérir. Pour ces auteurs, une structure isosyllabique comme le français serait par exemple plus facile à acquérir qu'une structure isoaccentuelle comme celle de l'anglais.

De même, les transferts seraient également liés à la tâche demandée aux apprenants. Pour Wenk (1985), le rythme de l'interlangue (combinaison de la LM et de la LE) varierait selon la tâche demandée puisque le syncrétisme des deux rythmes serait plus marqué en lecture oralisée qu'en répétition de mots ou en parole spontanée. Ce résultat rejoint la conclusion de Kaglik & Boula de Mareuil (2009 ; 2010) selon laquelle la prosodie (donc a fortiori le rythme) présenterait un coût cognitif important quand l'apprenant doit se focaliser sur le fond et sur la forme, comme c'est le cas en tâche de lecture oralisée.

En conclusion, nous retiendrons que plusieurs facteurs peuvent influencer les transferts rythmiques : les ressemblances entre la langue maternelle et la langue étrangère, leurs différences, la complexité de la structure de la langue ainsi que le coût cognitif de la tâche.

Mais la vraie question, pour la pédagogue que nous sommes, est de savoir s'il est possible ou non d'apprendre le rythme d'une LE. Il existe malheureusement encore trop peu d'études qui s'intéressent à l'évolution de l'interlangue de l'apprenant et les méthodes permettant d'agir sur celle-ci (nous citerons en exemple l'étude de Konopczynski *et al* (2004)).

C'est dans ce cadre là que se situe la problématique de ce travail.

2.4 SYNTHÈSE DU CHAPITRE 2

Dans ce chapitre, nous avons posé le rythme comme base de la prosodie, en tant qu'élément prééminent à l'intonation. Nous avons défini le rythme et ses différents composants : l'accentuation, les pauses et le débit.

Nous avons ensuite questionné la classification traditionnelle entre rythmicité syllabique et rythmicité accentuelle pour opposer des langues comme le Français et l'Anglais. Après avoir décrit le rythme de ces deux langues nous nous sommes intéressée aux conséquences que ces classifications pouvaient avoir dans l'apprentissage d'une LE.

Nous avons ensuite développé la question des transferts rythmiques entre la LM et la LE et leurs conséquences sur l'interlangue de l'apprenant. De récentes études ont mis en évidence plusieurs facteurs pour expliquer ces transferts: les différences entre les rythmes des deux langues (Rasier & Hiligsmann, 2007), leurs ressemblances (White & Mattys, 2007), la complexité de la structure rythmique de la LE (Konopczynski, 1999 ; Grabe *et al*, 1999) ou encore le coût cognitif de la tâche (Wenk, 1985 ; Kaglik & Boula de Mareüil, 2009, 2010). En ce qui concerne l'interlangue de l'apprenant, il semblerait que le rythme en construction soit marqué à la fois par le rythme de la LM et celui de la LE. Le rythme de l'interlangue ne serait donc pas la représentation exacte du rythme de l'une des deux langues concernées (Wenk, 1985). C'est pourquoi, il nous paraît nécessaire de placer l'interlangue au cœur des recherches sur la LE.

Peu d'études s'intéressent pourtant à l'apprentissage du rythme de la LE. Dans le cadre de ce travail, nous tâcherons d'apporter un éclairage nouveau sur les transferts rythmiques entre LM et LE et l'interlangue.

Avant de développer la problématique de ce travail, nous terminerons notre partie théorique en abordant la notion de fluence. En effet, cette notion interdisciplinaire se trouve être à la croisée de la phonétique et la didactique car si elle est utilisée comme mesure de la performance orale en cours de LE, nous

pensons qu'elle pourrait également permettre de rendre compte de la compétence rythmique d'un apprenant de LE.

Chapitre 3 : La fluence comme mesure de la compétence rythmique en langue étrangère

« La seule fluence des phrases dénouera sans effort la situation »

Paul Valéry

Ce troisième chapitre est consacré à un état des lieux sur la fluence en LE. En effet, souvent associée à l'aisance globale, aux capacités discursives et à la vitesse de parole, la fluence est utilisée depuis longtemps en didactique du FLE comme un critère d'évaluation des productions orales des apprenants (CECR).

Cependant, si être fluent constitue l'un des objectifs majeurs de l'apprentissage d'une LE, aussi bien à l'oral qu'en lecture, il n'existe paradoxalement pas de consensus sur ce que recouvre ce terme.

Dans ce chapitre, nous proposerons une rétrospective des travaux sur la fluence en LE pour mieux appréhender ce champ de recherches en plein essor. Nous verrons que la fluence recouvre une large gamme de définitions et de critères de mesure. Nous terminerons notre réflexion par l'apport des approches cognitives qui considèrent la fluence comme une compétence liée à l'automatisation des processus d'encodage de la langue.

À partir de cette définition de la fluence, nous proposerons de considérer la fluence comme une mesure de la compétence rythmique, résultant de l'automatisation des processus d'encodage et de décodage de la LE. Cette nouvelle approche nous permettra d'introduire notre problématique sur les transferts rythmiques entre l'oral et l'écrit et leurs incidences sur la fluence des productions d'apprenants anglophones de FLE.

3.1 INTRODUCTION

Depuis quelques décennies, la fluence constitue un champ de recherche à part entière en linguistique, dans les recherches sur la lecture (Hasbrouck, 1992 ; Fuchs et al. 2001) sur les pathologies du langage (voir par exemple l'étude des troubles de la fluence dans les cas de dysarthrie (Auriacombe *et al*, 1993), d'agrammatisme (Molinari *et al*, 1997), de troubles de la mémoire (Ruff, 1997), de dyslexie (Lefly & Pennington, 1991 ; Hook & Jones, 2002) ou de bégaiement (Hartsuiker *et al*, 2003)), et en langue étrangère (Sajavaara & Lehtonen, 1978 ; Sajavaara, 1987 ; Brumfit, 1984 ; Lennon, 1990 ; Ejzenberg, 1992 ; Segalowitz, 2000 ; Derwing *et al*, 2004 ; Wood, 2004 ; Wolf, 2008).

Ces études n'apportent pas pour autant de définition claire de la fluence, vue comme un phénomène temporel autant qu'une compétence globale à l'oral, en lecture, ou encore comme une compétence lexicale.

De plus, la fluence n'est pas qu'un phénomène de production. Concernant la perception, Hartsuiker *et al* (2003) ont montré que les bègues seraient plus sensibles aux écarts de productions (de types répétitions ou hésitations) - y compris dans des productions fluentes de parole spontanée chez les sujets sains - que les non-bègues. Cette hyper-vigilance perceptive serait une cause probable de leur bégaiement. D'autre part, il semblerait que disfluences produites par les sujets bègues ne soient pas comparables à celles produites par les sujets sains.

Pour Cole *et al* (2005), la perception et la production des répétitions, qui constituent un des types de disfluences le plus fréquent (environ 20%), seraient liées à la maîtrise de facteurs prosodiques (comme la variation de la F0) signalant à l'interlocuteur la présence d'une répétition. Ainsi, l'interlocuteur pourrait anticiper et ne pas traiter le mot répété. Nous ajouterons d'ailleurs qu'une des erreurs les plus fréquentes dans la transcription de corpus oraux est d'oublier de noter un mot répété (Pallaud, 2002).

Même si la fluence en LM ne constitue pas notre objet d'étude, on ne peut s'empêcher de faire un parallèle entre les études sur la fluence chez le

sujet pathologique et chez l'apprenant de LE. En effet, dans l'étude de Hartsuiker *et al* (2003) on pourrait imaginer que si le discours de sujets bègues est perçu comme plus disfluent alors que le nombre de répétitions est identique à celui d'un discours non bègue, c'est parce que les paramètres prosodiques permettant l'anticipation de l'interlocuteur ne sont pas respectés, comme c'est le cas dans les productions d'apprenants de LE (cf. chapitre 2).

Dans la partie qui suit, nous reviendrons sur les premières études sur la fluence en LE. Nous verrons que définir la fluence en la délimitant au cadre de la LE, est loin d'être acquis. Nous accentuerons le fait que l'interlangue pourrait offrir un bon cadre d'étude de la fluence en LE, en postulant que la fluence pourrait constituer un critère de mesure de l'évolution de la compétence rythmique des apprenants de LE.

3.2 LA FLUENCE EN LANGUE ETRANGERE

En ce qui concerne la fluence en LE, nous avons une idée générale de ce qu'elle pourrait représenter. Il n'est pas rare d'entendre des expressions, comme par exemple « *il est bilingue en anglais, il est totalement fluent* » où les termes fluence et expertise en LE sont associés, comme s'ils étaient synonymes.

Mais être fluent, est-ce réellement parler en LE comme dans sa LM? La compétence visée est-elle une compétence native? Nous y reviendrons (chapitre 4).

Plusieurs auteurs ont démontré que l'on pouvait être fluent en LE sans pour autant avoir développé des compétences grammaticales ou communicatives. Pour Schimdt (1992), la fluence et l'exactitude en langue étrangère sont deux choses différentes. Pour Sajavaara il n'y a pas forcément de corrélation entre une bonne fluence et une bonne compétence linguistique (Sajavaara, 1987), même si les ruptures (faux départs, hésitations etc.) dans le discours sont le plus souvent causées par des lacunes linguistiques (Sajvaara & Lehtonen, 1978).

Pawley & Syder (1983) se sont penchés sur la différence entre une production correcte et une production considérée comme native en termes de lexique (choix d'expressions ou utilisation de langage préfabriqué parmi des milliers de phrases correctes mais moins idiomatiques) et de fluence (production sans interruption d'une partie de discours, la fluence est ici également associée à la planification verbale). Pour ces auteurs, la maîtrise de l'authenticité (à la fois lexicale et du point de vue de la fluence) réside en réalité dans l'acquisition de patterns types en langue étrangère, via l'apprentissage des règles qui permettent de distinguer les phrases correctes de phrases idiomatiques. La fluence ne serait que la conséquence de cela.

Les conclusions de cette dernière étude nous amènent à nous interroger sur le type de patterns qu'il est nécessaire d'acquérir. S'agit-il de uniquement des phrases idiomatiques ou au contraire de patterns prosodiques ? Quelle est la part accordée à la dimension prosodique dans les études sur la fluence ? Lorsqu'on parle de compétence linguistique (*cf.* Sajavaara, 1987), de quel type de compétence s'agit-il ?

Nous allons voir que la fluence est associée à différentes compétences linguistiques. L'absence de consensus sur ce que recouvre la fluence pose problème pour comparer les différents résultats et proposer un modèle de la fluence en LE.

3.2.1 Première difficulté : définir la fluence en langue étrangère

Comme le rappelle Freed (1995) et Freed *et al* (2004), la fluence recouvre en réalité une large gamme de définitions.

Selon Fillmore (1979), un des pionniers des travaux dans ce domaine, la fluence est définie par quatre grandes compétences langagières :

- la capacité de parler avec un minimum de pauses (compétence phonétique)
- la capacité de parler en utilisant un contenu sémantiquement riche (compétence sémantique)
- la capacité d'utiliser les bons mots dans des contextes variés

Chapitre 3 : La fluence comme mesure de la compétence rythmique

(compétence pragmatique/discursive)

- la capacité d'utiliser le langage de manière créative (compétence lexicale)

Les définitions proposées par Fillmore recouvrent des compétences très variées, et il est bien difficile de dissocier alors la fluence d'une compétence orale globale en LE.

L'idée selon laquelle la fluence en LE serait associée à une *impression globale d'authenticité* a en effet servi de définition de base à de nombreuses études (Brumfit, 1984 ; Sajavaara, 1987; Riggenbach, 1991; Schmidt, 1992).

De même, pour Lennon (1990), la fluence est généralement utilisée dans deux acceptions : comme compétence orale globale ou comme la fluidité du discours. Dans ce dernier cas, elle est représentée par des critères d'ordre temporels (tableau 6).

Chapitre 3 : La fluence comme mesure de la compétence rythmique

Unités de mesure	Définition
(1) Débit de parole (syllabe/minute)	(1) Durée du nombre total de syllabes divisé par la durée totale de temps de parole (pauses comprises) exprimée en secondes
(2) Taux d'articulation (nbre de syllabes/minute)	(2) Durée du nombre total de syllabes divisé par la durée totale de temps de parole (pauses exclues) exprimée en secondes.
(3) Temps de phonation (pourcentage)	(3) Pourcentage du ratio de temps de parole par rapport au temps total de la durée de l'énoncé.
(4) Durée moyenne des séquences (nombre de syllabes)	(4) Nombre moyen de syllabes compris entre chaque pause (pauses supérieures ou égales à 250ms)
5) Nombre de pauses silencieuses /minutes	5) Durée totale du nombre de pauses silencieuses (supérieures ou égales à 250ms) divisée par la durée totale de temps de parole exprimée en secondes.
(6) Durée moyenne des pauses silencieuses (secondes)	(6) Durée moyennes des pauses (> 250 ms)
(7) Nombre de pauses remplies par minute	7) Durée totale du nombre de pauses remplies divisée par la durée totale de temps de parole exprimée en secondes.
(8) Nombre de disfluences par minute	8) Durée totale du nombre de disfluences divisée par la durée totale de temps de parole exprimée en secondes.
(9) 'Pace' (mots/minute)	(9) Nombre de mots accentués par minute
(10) 'Space' (ratio de mots accentués / total des mots)	(10) Proportion de mots accentués par rapport à la totalité des mots

Tableau 6 Mesures de la fluidité inspiré des 12 critères établis par Lennon (1990) (d'après Kormos, 2006)

En résumé, la fluence peut être considérée aussi bien comme le reflet d'une compétence sociolinguistique et communicative globale (Pawley & Syder, 1983 ; Ejzenger, 1992 ; Wood, 2004) que comme une variable physique et temporelle de production de parole (Grosjean, 1980 ; Lennon, 1990 ; Cucchiarini *et al*, 2000 ; Derwing *et al*, 2004).

Pour Kormos & Dénes (2004), la seule manière de définir correctement la fluence serait de prendre en compte aussi bien les aspects temporels que l'aisance globale, la créativité lexicale, la complexité grammaticale et les paramètres intonatifs. Ainsi, dans cette étude, les auteurs s'interrogent sur les

variables linguistiques et temporelles qui prédisent l'évaluation perceptive de la fluence chez des non-natifs par des enseignants de LE (natifs et non natifs). Les résultats montrent que les critères retenus par les enseignants, aussi bien natifs que non-natifs, sont la vitesse de parole, la durée moyenne des segments, la durée de phonation et le nombre de mots accentués par minute. Pour la plupart des enseignants, la précision serait également un critère non négligeable. Les auteurs concluent que les variables temporelles et linguistiques sont en réalité liées et que du coup la fluence serait le révélateur (dans les productions) de la mise en place d'automatismes linguistiques.

« The correlations between the temporal and linguistic variables also reveals that accuracy is positively related to temporal variables that are influential in fluency judgements. In other words, it seems that those students who were fluent in terms of speed and pace also produced accurate output. In psycholinguistics terms this means that one is only able to speak fluently if speech production mechanisms are largely automatic and if automatic sequences are memorised, retrieved and used accurately. » (Kormos & Dénes, 2004 : 160)

Plusieurs auteurs partagent ce point et soutiennent que la fluence est la manifestation des processus automatisés qui sous-tendent la planification de la parole par le locuteur (Segalowitz, 2000 ; Dewaele, 2002). Or, dans le chapitre précédent, nous avons mis en avant le rôle des constituants rythmiques, et en particulier des pauses dans la planification de la parole (partie 2.1.1). La fluence serait-elle associée à la compétence rythmique ?

Avant de développer ces aspects, nous nous intéresserons plus précisément aux critères choisis pour évaluer la fluence des apprenants dans différentes études. Nous verrons ainsi que la diversité de définitions n'explique qu'en partie les difficultés que nous avons à délimiter objectivement le cadre d'étude de la fluence. En effet, la diversité se retrouve également dans les méthodes utilisées et les unités de mesures choisies.

3.2.2 Deuxième difficulté : mesurer la fluence

Les critères utilisés pour mesurer la fluence sont très variés, et souvent très subjectifs. Pour évaluer la fluence, on peut mesurer aussi bien le débit, le nombre et le type d'hésitations, le nombre et le types de disfluences, le nombre de mots accentués, l'aisance globale, l'exactitude, la richesse du vocabulaire, la maîtrise de la grammaire, l'utilisation appropriée de phrases en contexte, les compétences communicatives *etc.* Même pour les critères temporels qui semblent a priori objectifs, il n'y a pas de consensus concernant par exemple, la durée minimale des pauses ou le types de disfluences choisies.

Ainsi, Riegenbach (1991) montre que parmi les critères influençant la perception d'une bonne ou d'une mauvaise fluence en LE par des natifs (ici enseignants de LE), on retrouve aussi bien le nombre d'hésitations, le nombre et le placement des pauses, le débit de parole que la capacité à initier des changements de sujets dans les conversations.

De même, Freed (1995) et Freed *et al* (2004) indiquent que les natifs (experts et non experts) évaluent la fluence d'apprenants de FLES (avant et après un séjour à l'étranger) en se basant principalement sur des appréciations qualitatives (*i.e.* richesse du vocabulaire, utilisation de structures grammaticales complexes, aisance de parole, confiance en LE, vivacité de la voix *etc*), même si les auditeurs sont également sensibles à l'amélioration du débit de parole ou la diminution du nombre d'hésitations.

Notons que ces travaux nuancent également la croyance selon laquelle il suffit de se rendre dans le pays de la langue étudiée pour améliorer sa production orale (cf. chapitre 1). En effet, dans la première étude, même si globalement les étudiants partis à l'étranger (*groupe test*) obtiennent de meilleurs scores après leur séjour (les apprenants ont un débit de parole jugé plus rapide, produisent moins de pauses agrammaticales, font preuve d'auto-correction et de plus d'aisance), certains étudiants du *groupe contrôle* (qui avaient suivi des cours de langues classiques) obtiennent des scores au moins égaux à ceux des apprenants du *groupe test*.

Chapitre 3 : La fluence comme mesure de la compétence rythmique

Dans la seconde étude, Freed *et al* (2004) ont comparé la progression de 28 apprenants répartis non plus en deux mais en trois groupes : un groupe parti étudier à l'étranger, un groupe resté dans le pays d'origine et un groupe ayant suivi des cours intensifs pendant l'été (cours basés sur l'oral et l'écrit). Les résultats montrent que ce sont les apprenants ayant suivi les cours intensifs qui obtiennent les meilleurs résultats. L'accent mis sur l'écrit pendant l'immersion amène les auteurs à s'interroger sur les transferts entre l'écrit et l'oral. Sur la base d'analyses complémentaires, les auteurs concluent que le nombre d'heures passées à renforcer l'écrit a eu une incidence sur l'amélioration de la fluence des productions des apprenants. Dans le cadre de ce travail nous formulons l'hypothèse opposée en postulant que l'oral servirait au décodage de l'écrit. Les résultats de Freed *et al* (2004) seront discutés vis à vis de nos propres résultats dans le chapitre 5.

Pour en revenir aux critères retenus, un des principaux problèmes, comme le soulève Segalowitz (2010), est que ce ne sont pas les hésitations en elles-mêmes qui font la disfluence mais leur caractère systématique. Si mesurer le nombre de pauses est en soi objectif (comparable si l'on a défini une durée de pause standard), il n'en va pas de même pour déterminer leur caractère intempestif qui renvoie plutôt à la perception de l'auditeur (*cf.* l'étude de Hartsuiker *et al* (2003) sur le bégaiement)

Est-il possible d'établir des critères de mesures objectifs ? Faute de quoi, l'appréciation de la fluence reste soumise à la subjectivité de l'auditeur.

Certaines études, comme celle de Lennon (1990) ont essayé de déterminer des critères objectifs et quantifiables de la fluence. L'auteur se proposait de comparer la fluence/fluidité des productions de quatre apprenants germanophones en anglais LE, avant et après un séjour de six mois en Angleterre, dans une tâche de narration. Pour ce faire, Lennon s'est basé sur douze mesures de performance orale qui pourraient se révéler être de bons indicateurs de fluence (mesures reprises et résumées par Kormos (2006), voir tableau 6). Tous les participants ont amélioré leur débit de parole et diminué le nombre de pauses remplies après leur séjour à l'étranger. Toutefois,

l'amélioration du débit de parole serait plus imputable à la diminution des pauses silencieuses qu'à l'augmentation de la vitesse de parole. Lennon conclut que le débit de parole et les pauses ne constituent pas en eux-mêmes des critères suffisants pour évaluer la fluence mais que c'est la combinaison de ces paramètres, associée à la complexité du discours, qui permet de l'évaluer. De même, les pauses remplies et les répétitions seraient étroitement liées et pourraient indiquer la planification du discours.

Olynyk *et al* (1987) ont également montré que la fréquence d'apparition des disfluences (pauses remplies, répétitions, réparations etc.) n'est pas suffisante pour évaluer la fluence dans une situation d'interaction. Les auteurs ont démontré que c'est leur place qui serait déterminante. Si les disfluences apparaissent avant qu'un mot soit produit dans sa totalité (disfluences progressives), elles perturberont moins la fluidité du discours que si elles apparaissent après le mot (disfluences régressives). Dans le premier cas, les disfluences respectent la linéarité de la parole et ne perturbent pas outre mesure la perception de l'interlocuteur qui n'a pas de réajustements à effectuer. Dans le second cas, au contraire, les erreurs de fluence requièrent des réajustements constants de la part de l'interlocuteur. Les apprenants commettant ce type de disfluences seront alors perçus comme étant moins fluents.

Dans le but de déterminer si l'appréciation de la fluence par des natifs peut être automatiquement prédite par des mesures temporelles, Cucchiarini *et al* (2000) ont testé la corrélation entre le jugement des natifs et sept critères objectifs de la fluence : le débit de parole, le taux d'articulation (« articulation rate »), le nombre et la durée des pauses, le nombre de disfluences, la durée moyenne des séquences (« runs »), et le rapport temps de phonation/ temps total de parole (« phonation/time ratio »). Les résultats ont montré que les notations des natifs étaient corrélées avec les variables choisies, en particulier avec le débit de parole, qui semblerait être le meilleur prédicteur de la fluence. Wennerstrom (2000) a également montré que la prosodie jouait un rôle important dans la perception de la fluence par des natifs. Le placement des pauses à des frontières syntaxiquement cohérentes et les variations de la F0 sont des facteurs déterminants dans l'évaluation d'une bonne fluence en LE.

Chapitre 3 : La fluence comme mesure de la compétence rythmique

Derwing *et al* (2004) ont également étudié la relation entre la perception des natifs et les mesures de débit ou de pauses. Les auteurs proposaient de tester la corrélation entre la fluence et les variables temporelles. Plus précisément, ils souhaitaient tester l'impact de la durée du message pour l'auditeur, celui de la tâche demandée et si la prosodie avait la même importance dans la perception de la fluence quelle que soit la tâche. Les auteurs ont également testé la corrélation entre la compréhensibilité, l'accent étranger et la fluence en demandant aux auditeurs d'évaluer la prégnance de l'accent étranger et la compréhensibilité du message, résultats qu'ils ont ensuite comparé avec ceux de la fluence. Leurs résultats nuancent le rôle attribué à la rapidité de parole (certains apprenants peuvent en effet parler très vite et donner l'illusion d'être fluents) et à la durée des pauses silencieuses dans la perception de la fluence par les auditeurs natifs. Les performances des apprenants ont été perçues comme étant meilleures sur les tâches de monologues et de dialogues mais, en ce qui concerne la notation de la prosodie, les auteurs n'observent pas de différence significatives selon les tâches demandées. En revanche, les auteurs observent une très forte corrélation entre la compréhensibilité et la fluence du discours, corrélation encore plus importante qu'entre fluence et accent étranger. Ainsi, un extrait où l'accent est très prononcé ne sera pas forcément jugé mauvais en compréhensibilité. Les résultats observés semblent en contradiction avec ceux des études précédentes, notamment en ce qui concerne l'impact du débit de parole et des pauses dans la perception de la fluence. Ces différences peuvent s'expliquer par plusieurs raisons : tout d'abord la durée des pauses silencieuses choisie était de 400ms or nous avons vu précédemment que les phonéticiens prennent en compte une durée nettement inférieure de 250ms (Goldman-Eisler, 1968). Kormos (2006) souligne d'ailleurs ce problème méthodologique : il n'y a pas de consensus concernant la durée minimale des pauses silencieuses choisie dans les études. De plus, comme nous l'avons vu dans les études précédentes, la fréquence d'apparition des pauses ne suffit pas à évaluer la fluence.

Enfin, les auteurs ont choisi des sujets de niveau débutant (plutôt que des apprenants de niveau intermédiaire ou avancé comme c'est le plus souvent le

cas) et il semblerait que le niveau des apprenants influence la perception de la fluence par les natifs (Cucchiaroni *et al*, 2002). Les critères choisis ne seraient pas les mêmes pour des niveaux débutants et avancés. Dans le premier cas, le temps de phonation serait le paramètre le plus déterminant tandis que dans le second il s'agirait de la longueur moyenne des segments.

Deux problèmes majeurs se dégagent de la description des études précédentes. Un problème méthodologique tout d'abord car faute de consensus sur les critères de mesures ou le dessein expérimental (tâche, choix des sujets *etc*), les résultats des différentes recherches ne sont pas comparables.

Un problème de modélisation ensuite car la plupart de ces études négligent les aspects cognitifs mis en jeu autant chez le locuteur que l'auditeur, alors que les critères prosodiques retenus – débit de parole ou pauses – permettent de rendre compte de la planification du discours. Il semblerait également que des facteurs tels que le feed-back verbal et non verbal (de type hochement de tête), donné par un interlocuteur puisse améliorer la fluence des productions d'apprenants de LE (Wolf, 2008).

La fluence serait-elle alors le reflet de processus complexes d'encodage verbal ? Existe-t-il une modélisation de la fluence qui permettrait de rendre compte des aspects cognitifs mis en jeu ainsi que de la production et de la perception ?

3.3 ECLAIRAGES COGNITIFS

3.3.1 Modélisation de la fluence dans une approche cognitive

Selon Segalowitz (2010), il n'existerait non pas une mais trois entités de fluences différentes mais complémentaires, recouvrant chacune des compétences et des critères de mesures spécifiques : la fluence cognitive, la fluence de production et la fluence perçue¹ (figure 19).

¹ Notre traduction de 'cognitive fluency', 'utterance fluency' et 'perceived fluency'

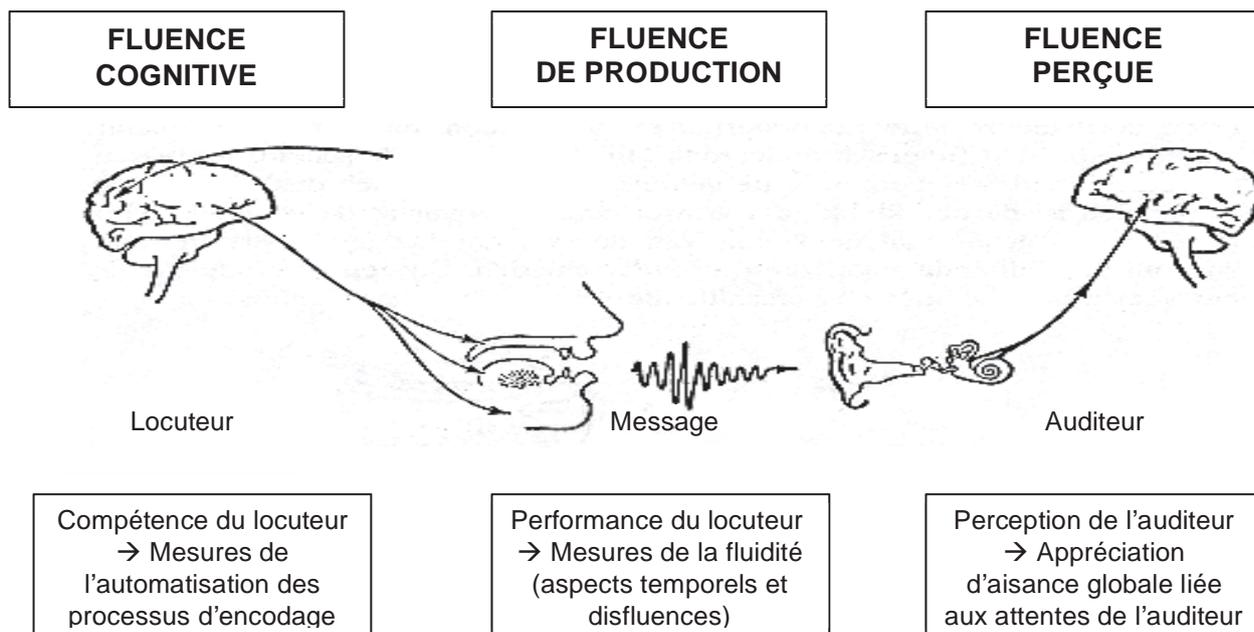


Figure 19 Schéma des trois types de fluence d'après la description proposée par Segalowitz, 2010

La fluence perçue (probablement l'aspect le plus étudié dans la littérature) reflète les inférences que les auditeurs font au sujet de la fluence cognitive, en se basant sur leur perception de la fluence produite. Mais nous savons que chaque individu organise et structure le monde qui l'entoure (*cf.* chapitre 1). L'évaluation de la fluence risque donc d'être influencée par le système perceptif de l'auditeur.

La fluence de production, quant à elle, est représentative de la performance du locuteur et se base sur des mesures temporelles de sa production. Elle est la représentation/ l'actualisation de la fluence cognitive.

La fluence cognitive, enfin, correspond à la compétence du locuteur. Elle se traduit par l'automatisation des différents processus d'encodage de la parole.

Cela rejoint les conclusions de Towell *et al* (1996) pour qui l'amélioration de la fluence serait le reflet d'un changement plus profond dans la manière dont la parole est générée. Pour ces auteurs, la fluence serait associée à la conversion d'un savoir déclaratif en un savoir procédural. Dans le but de déterminer comment se fait cette conversion et quels facteurs sont modifiés, les auteurs ont mesuré la fluence d'apprenants de FLE (tâche de narration) avant et

après un séjour d'un an dans un pays francophone. Les auteurs ont relevé des différences significatives entre ces deux périodes, en termes d'amélioration du taux d'articulation et d'augmentation de la longueur des segments entre deux pauses. En revanche, ils n'ont pas observé de différences significatives dans le temps de phonation ni dans la longueur moyenne des pauses. L'amélioration du débit de parole serait alors corrélée à l'augmentation de la durée moyenne des segments plutôt qu'à une diminution de la durée des pauses. Les auteurs concluent que l'amélioration de la fluence est associée à des changements complexes et profonds de la manière dont la langue étrangère est produite. Changement qui pour eux, serait associé à la procéduralisation de savoirs langagiers.

Pour aller plus loin dans la description de la fluence cognitive, Segalowitz (2000) propose de tester expérimentalement le rôle des processus automatiques et des processus contrôlés dans l'amélioration de la fluence cognitive, en lecture, chez des apprenants de LE. Dans cet article, l'auteur précise qu'il peut arriver, par exemple, qu'un apprenant améliore la rapidité de sa lecture sans pour autant avoir mieux compris le texte lu. L'apprenant est alors dans une fluence de production et pas de cognition. Par contre, si l'apprenant est capable d'être fluent dans n'importe quel contexte (multiples tâches différentes, avec ou sans distracteurs *etc.*), il s'agit bien d'une amélioration de la fluence cognitive.

L'auteur décrit plusieurs études menées dans le but de tester le rôle des processus automatiques et des processus contrôlés dans la fluence cognitive en lecture en langue étrangère (associée en lecture à la rapidité de reconnaissance lexicale).

La première étude (Favreau & Segalowitz, 1983) avait pour objectif de déterminer la corrélation entre la fluence en lecture et le degré d'automatisation de reconnaissance de mot chez des bilingues avancés (bilingues anglais/français). Sachant que les lecteurs en langue étrangère lisent plus lentement qu'en langue maternelle, les auteurs ont postulé que la répartition des mécanismes automatiques et contrôlés est différente dans les deux cas. Deux groupes ont

été constitués : des bilingues lisant aussi vite dans leur langue maternelle que dans leur langue étrangère (*equal group*) et des bilingues moins performant en lecture en langue étrangère (*unequal group*). La tâche demandée était une tâche de reconnaissance lexicale (« est-ce que ces lettres forment un vrai mot ou pas ? »), dans une condition attendue (« fruit » (prime) : « pomme » (mot cible)) ou inattendue (« fruit » (prime) : « chaise » (mot cible)). L'expérimentateur fait varier le temps d'exposition. Les résultats ont montré que les apprenants du groupe 'unequal' ont besoin de plus de temps pour traiter l'information (résultats moins bons lorsque temps d'exposition est réduit à 96ms) que les apprenants du groupe 'equal' dont les scores sont semblables à ceux obtenus en langue maternelle. Les auteurs concluent que les apprenants du groupe 'equal' ont en fait recours aux processus automatisés aussi en langue étrangère.

Les différences de fluence entre la LM et la LE, seraient associées à des différences dans l'automatisation des processus, ce qui rejoint l'hypothèse de Perdue & Gaonach' (2000) selon laquelle les coûts cognitifs ne seraient pas les mêmes en LM et en LE.

Dans la seconde étude, Segalowitz *et al* (1998) testent la corrélation entre un meilleur niveau de fluence en langue étrangère et un degré plus élevé de processus contrôlés (à savoir, par exemple, la capacité à désactiver les représentations inappropriées en lecture) chez des bilingues anglais/espagnol. Dans une tâche de reconnaissance lexicale, les sujets devaient dire si le mot cible était ou non un mot espagnol (les mots étaient présentés par paire avec un distracteur anglais et un mot (ou un non mot) espagnol (cible)). Les résultats montraient que les bilingues avancés avaient plus de facilités que les moins avancés à inhiber le distracteur quand celui-ci était la traduction du mot espagnol présenté dans la même paire. Quand le mot anglais était la traduction du mot espagnol de la paire suivante, il n'y a pas eu de différences quel que soit le niveau des bilingues. En conclusion, les apprenants qui ont un meilleur niveau seront plus capables que les autres d'activer ou de désactiver les informations pertinentes en temps réel. C'est pour cela qu'ils seront plus à l'aise en situation de communication car ils pourront s'adapter plus vite que les débutants.

Chapitre 3 : La fluence comme mesure de la compétence rythmique

L'acquisition de la fluence cognitive nécessiterait l'automatisation de processus permettant la reconnaissance de mots mais également le développement de mécanismes de contrôles nécessitant un niveau élevé de compétences (activation ou inhibition des fausses représentations).

Cette conclusion est très intéressante car elle rejoint l'hypothèse de Perdue & Gaonach' (2000) selon laquelle les coûts cognitifs ne sont pas les mêmes en langue maternelle et en langue étrangère. La langue étrangère nécessiterait à la fois le développement et l'automatisation des processus de décodage mais également la capacité à inhiber les fausses représentations ou les représentations de la LM.

Nous regrettons toutefois que les tâches proposées par les auteurs précédents (Favreau & Segalowitz, 1983 ; Segalowitz *et al* (1998)) se limitent à des compétences de bas niveau. Nous pensons que ces tâches ne peuvent pas rendre compte d'une compétence de décodage globale en langue étrangère.

D'autres auteurs postulent en effet qu'une lecture efficace implique non seulement la reconnaissance automatisée de mots mais aussi la vitesse de lecture et la compréhension globale du texte. Pour Hook & Jones (2002), l'identification ne peut se faire sans l'acquisition de stratégies phoniques (syllabation, connaissance des règles d'accentuation *etc.*). Pour ces auteurs, la fluence serait étroitement liée à la maîtrise des caractéristiques prosodiques de la LE.

« Not only do good readers read fluently with adequate speed, but when they read aloud, they use appropriate phrasing, intonation, and their oral reading mirrors their spoken language » (Hook & Jones, 2002 : 18)

Ainsi, la fluence cognitive serait associée à l'automatisation de processus de décodage liés à la maîtrise de l'accentuation et du rythme de la langue étrangère qui constituent en effet des éléments clefs de l'encodage (cf. chapitre 2). De même, Blanco *et al* (2006) soulignent que la prosodie jouerait un rôle déterminant à tous les niveaux de la représentation linguistique, morphologique, syntaxique et sémantique.

Il est donc regrettable que la modélisation proposée par Segalowitz (2010) n'intègre pas la dimension prosodique à tous les niveaux de fluence. Nous pensons en effet que les compétences de décodage de la LE pourraient être étroitement liées à la maîtrise de la prosodie de la langue cible et que fluence et prosodie seraient étroitement corrélées.

3.3.2 La fluence comme mesure de la compétence rythmique ?

Plusieurs auteurs ont montré que la fluence était associée à des mesures temporelles. L'occurrence et la durée des pauses (tous types confondus) seraient les deux facteurs les plus déterminants pour évaluer la fluence et/ou l'intelligibilité d'un discours (Cenoz, 2000 ; Cucchiari *et al*, 2002 ; Ullakonoja, 2008).

Pour Cenoz (2000), l'analyse des pauses et des hésitations permet de rendre compte des processus d'acquisition de la LE car ce sont des bons indicateurs des difficultés de planification et de production de la LE. Ainsi, les pauses constituent de bons indicateurs de la progression des apprenants.

Cucchiari *et al* (2002) ont testé la corrélation entre des mesures temporelles et objectives de la fluence et la perception de celle-ci en parole lue et en parole spontanée chez des locuteurs hollandais. L'objectif de l'étude était de déterminer si les mesures quantitatives sont de bons indicateurs de la fluence produite dans le but de créer des tests objectifs de mesure de la fluence. Les résultats indiquent que les mesures temporelles choisies (débit de parole, durée des pauses, durée moyenne des segments etc.) sont de bons indicateurs de la fluence, en particulier en lecture oralisée. Ce dernier point rejoint les conclusions de Goldman-Eisler (1972) selon lesquelles la structuration pausale est plus dépendante de la hiérarchie des structures syntaxiques en parole lue qu'en parole spontanée.

Ullakonoja (2008) a étudié la corrélation entre la fluence et les pauses en parole oralisée chez des finnois apprenants de russe au cours d'un séjour linguistique en Russie. L'auteur a demandé à trente enseignants de russe d'évaluer la fluence des productions apprenants sur une échelle de magnitude

Chapitre 3 : La fluence comme mesure de la compétence rythmique

de 1 à 5. Les analyses acoustiques de la fréquence, de la durée et du placement des pauses, révèlent une forte corrélation avec l'évaluation de la fluence par les enseignants.

Ces études renforcent l'idée selon laquelle fluence et rythme seraient étroitement corrélés. La fluence de production pourrait-elle rendre compte de la mise en place de transferts rythmiques entre la LM et la LE. ? Autrement dit, la fluence cognitive serait-elle associée au développement d'une compétence rythmique en langue étrangère ? De plus, si la fluence reflète la mise en place de compétences automatisées, ne serait-elle pas un bon indicateur de l'interlangue de l'apprenant ?

Ces questions n'ont à notre connaissance jamais été posées. Elles méritent pourtant que l'on s'y intéresse et c'est pourquoi nous développerons ces aspects dans la suite de ce travail.

3.4 SYNTHÈSE DU CHAPITRE 3

Dans ce chapitre, nous avons vu que définir la fluence, autant pour les pédagogues que les chercheurs en didactique des langues n'avait rien d'évident. La fluence recouvrirait en effet des compétences très (trop) variées et l'hétérogénéité des méthodologies ne permet pas de consensus sur l'objet d'étude.

Toutefois, il ressort que les paramètres temporels constituent des phénomènes clefs de la description de la fluence, autant en production qu'en perception. De récentes études cognitives apportent un éclairage nouveau, en proposant non plus une mais trois catégories de fluences, la fluence cognitive, la fluence de production et la fluence perceptive. La fluence cognitive serait le reflet de l'automatisation des processus d'encodage.

Nous regrettons toutefois que ces études sur la fluence cognitive n'intègrent pas la dimension prosodique, élément clé de l'encodage d'une langue et base de la compétence orale. Nous l'avons déjà critiqué au chapitre

Chapitre 3 : La fluence comme mesure de la compétence rythmique

1, la compétence orale reste définie comme une compétence globale où les aspects phoniques sont négligés.

Nous pensons que fluence et prosodie sont en réalité étroitement corrélées. La fluence pourrait rendre compte de la mise en place d'une compétence rythmique et donc prosodique en LE.

Chapitre 4 Aspects méthodologiques

*« On ne peut se passer d'une méthode pour se mettre en quête de la
vérité des choses »*

René Descartes

Dans les chapitres théoriques précédents, nous avons vu que la didactique négligeait les aspects phoniques - prosodiques et phonémiques - dans l'enseignement de l'oral, alors que ces derniers constituent la base de l'encodage et du décodage de la langue.

Nous avons particulièrement insisté sur le rôle de la prosodie dans l'apprentissage d'une langue et la perception d'un accent étranger. Nous avons défini le rythme comme l'organisation des proéminences et de fait, comme l'ossature de l'intonation. Nous avons mis en évidence les différences rythmiques entre les langues en prenant l'exemple du français et de l'anglais considérés comme les langues rythmiquement les plus éloignées. Nous avons ensuite développé la notion de transferts rythmiques entre la LM et la LE, et montré le rôle de ces transferts dans l'apprentissage des langues.

Dans le troisième chapitre, nous avons adopté une approche cognitive de la fluence en établissant un lien étroit entre la fluence et l'automatisation des processus de décodage, en particulier en lecture. Connaissant le rôle de la prosodie dans l'encodage et le décodage d'une langue, nous avons postulé que prosodie et fluence serait étroitement corrélées. En outre, les paramètres les plus représentatifs de la fluence sont aussi des critères de mesure du rythme : les pauses, le débit de parole et la taille des groupes rythmiques délimités par des accents métriques. Nous pensons que la fluence des productions de parole lue pourrait rendre compte de la compétence rythmique des apprenants de LE.

Ce dernier chapitre a pour objectif de poser la problématique et le cadre méthodologique de notre travail. Nous aborderons tout d'abord la question des transferts prosodiques en LE sous un angle d'approche original et novateur : la prise en compte de la prosodie dans le décodage de l'écrit en LE. Cet angle d'approche nous permet d'éviter d'étudier les transferts prosodiques uniquement dans une comparaison entre la LM et la LE, peut être apte à rendre compte de la construction de l'interlangue.

Nous présenterons ensuite le cadre méthodologique des deux études longitudinales que nous avons réalisées puis nous décrirons le corpus

MULTIPHONIA que nous avons constitué. Nous reviendrons sur les difficultés et les enjeux de la mise en place d'études longitudinales et d'analyses acoustiques pour les recherches en langue étrangère.

4.1 OBJECTIFS DE LA THESE

Ce travail de thèse a pour objectif d'évaluer la présence et le rôle des transferts de compétences dans l'encodage de l'oral et de l'écrit en LE.

Une des activités de production orale les plus utilisées en classe de LE est la lecture oralisée. On pourrait croire que le lecteur n'étant pas en situation de planification du discours et que le phrasé étant en partie indiqué par la ponctuation, la lecture permettrait un meilleur contrôle des transferts prosodiques entre la LM et la LE. Voire qu'elle faciliterait l'acquisition de la prononciation en LE. Autrement dit, on s'attendrait à ce que la fluence des productions soit meilleure en parole lue qu'en parole spontanée.

Or ce n'est pas ce que nous observons dans nos pratiques de classes. Quel que soit le niveau des apprenants, leur élocution est beaucoup moins fluide en lecture oralisée qu'en parole spontanée. En effet, on dénote la présence d'un plus grand nombre de pauses agrammaticales et d'hésitations, ainsi qu'un ralentissement global du débit de parole en parole lue. Il est vrai que la lecture représente une tâche cognitivement lourde. En effet, l'apprenant de langue étrangère se trouve face à plusieurs tâches complexes en situation de lecture: il doit non seulement décoder la chaîne graphique et la recoder en chaîne phonémique, mais aussi organiser simultanément le matériau linguistique en groupes sémantiquement et syntaxiquement cohérents.

Or nous savons que la prosodie participe à la structuration des énoncés (Di Cristo & Hirst, 1993 ; Vaissière, 1997 ; Di Cristo, 2004). La prosodie jouerait-elle un rôle déterminant pour l'encodage et le décodage de l'écrit ?

Il semblerait en effet que les informations prosodiques soient nécessaires au découpage du texte en groupe prosodiquement et syntaxiquement cohérents.

Nous savons par exemple que la ponctuation ne permet pas d'oraliser correctement les textes écrits. Les recherches actuelles en synthèse de parole ont démontré la nécessité d'ajouter au texte ponctué des marqueurs (balises correspondant entre autres aux pauses réelles ou aux schémas intonatifs) pour

signaler des aspects pragmatiques ou déclencher les formes prosodiques souhaitées (Mertens *et al*, 2001).

Mais lorsque nous lisons dans notre vie quotidienne ces informations n'apparaissent pas entre les lignes. Pourtant, nous arrivons à décoder et à oraliser sans peine le texte écrit. Serait-il alors possible que nous réutilisions des connaissances prosodiques acquises et mémorisées par notre pratique de la langue ? Existe-t-il des transferts prosodiques entre l'oral et l'écrit ?

Si c'est le cas, ces transferts ne pourraient-ils pas expliquer les difficultés que rencontrent les apprenants de LE en situation de lecture oralisée ? Serait-il également possible d'agir sur ces derniers pour améliorer la fluence des productions de parole lue ?

Ces questions se trouvent en fait au cœur de notre problématique. L'hypothèse de l'existence des transferts de compétences entre l'encodage de l'oral et de l'écrit en langue étrangère, n'a, à notre connaissance, jamais été posée. Nous développerons donc cet aspect en établissant un parallèle avec de récentes découvertes en psycholinguistique et en sciences cognitives en LM sur l'existence d'une prosodie implicite. Nous nous pencherons ensuite sur les moyens pouvant être mis en place pour favoriser l'apprentissage d'une compétence prosodique en LE.

4.1.1 Existe-t-il des transferts de compétences entre l'encodage de l'oral et de l'écrit en langue étrangère ? Quel rôle joue la prosodie dans ces transferts ?

De récentes études ont montré l'existence de transferts de compétences entre l'oral et l'écrit en LM (*cf.* Hypothèse de la prosodie implicite, Fodor, 2002).

« The Implicit Prosody Hypothesis (IPH) : In silent reading, a default prosodic contour is projected onto the stimulus, and it may influence syntactic ambiguity resolution. Other things being equal, the parser favors the syntactic analysis associated with the most natural (default) prosodic contour for the construction » (Fodor, 2002:2).

Ainsi, même en lecture silencieuse un contour prosodique par défaut serait projeté sur le texte. Le lecteur ne partirait pas du texte (démarche Bottom-up) mais ferait appel à des modèles prosodiques établis et mémorisés par la pratique orale (démarche Top-down) pour faciliter le décodage du code écrit. La prosodie ainsi projetée sur le texte écrit serait alors considérée comme un constituant du texte lui-même.

« Even in reading, prosody is present. Even in silent reading, and even if prosody-marking punctuation is absent. Prosody is mentally projected by readers onto the written or printed word string. And –the crucial point- it is then treated as if it were part of the input, so it can affect syntactic ambiguity resolution in the same way as overt prosody in speech does ». (Fodor, 2002 :1)

Pour Fodor (2002) la prosodie implicite serait donc responsable du choix de découpages syntaxiques effectués par le lecteur. Pour tester cette hypothèse, Fodor a comparé la résolution de phrases ambiguës à l'écoute et en lecture silencieuse dans plusieurs langues, en tenant compte des stratégies de découpages prosodiques de chaque langue.

Ainsi, dans le cas de phrases ambiguës comme « quelqu'un a tiré sur le serviteur de l'actrice qui était au balcon ² », nous ne savons pas qui était sur le balcon. Était-ce le serviteur ou l'actrice ? Les anglophones auront tendance à privilégier un attachement local et donc à associer la relative au nom qui précède, ici 'l'actrice'. Au contraire, d'autres langues comme le français ou l'espagnol favoriseront un attachement non local pour lequel la relative serait associée au nom le plus éloigné, ici le 'serviteur.' Dans ce cas, le deuxième nom, 'actrice', sera séparé de la relative par une pause vide. L'auteur a voulu comparer la résolution de ces phrases ambiguës dans deux tâches : en lecture et en compréhension orale. Les mesures des temps de réponses confirment que la prosodie est utilisée de manière comparable pour les deux tâches, pour toutes les langues étudiées. Cela amène Fodor à postuler que la prosodie

² Notre traduction de l'exemple donné par Fodor (2002) « Someone shot the servant of the actress who was on the balcony »

implicite joue le même rôle que la prosodie explicite dans la segmentation du langage. Il s'agirait de plus d'un phénomène naturel et universel.

D'autre part, l'hypothèse de la prosodie implicite suppose également qu'il existe un lien étroit entre le traitement des informations visuelles et celui des informations auditives.

De récentes recherches développementales en sciences cognitives utilisant l'imagerie cérébrale, corroborent cette hypothèse. En effet, les caractéristiques anatomiques et fonctionnelles du cerveau indiqueraient un transfert possible des compétences auditives sur les compétences d'écriture et de lecture. Les aires visuelles et auditives sont non seulement interconnectées *avant* que l'enfant apprenne à lire, mais l'aire de Wernicke – où la compréhension du langage écrit et oral est traité – est localisée à la croisée des aires auditives et visuelles. Par conséquent, tout processus linguistique appris *avant* le développement des compétences en lecture et en écriture est susceptible de 'nourrir' la modalité visuelle (Dehaene, 2007).

Comme nous pouvons le voir sur la figure 20, on sait aujourd'hui que la lecture implique de nombreuses interconnexions entre les différentes aires visuelles et les aires du langage.

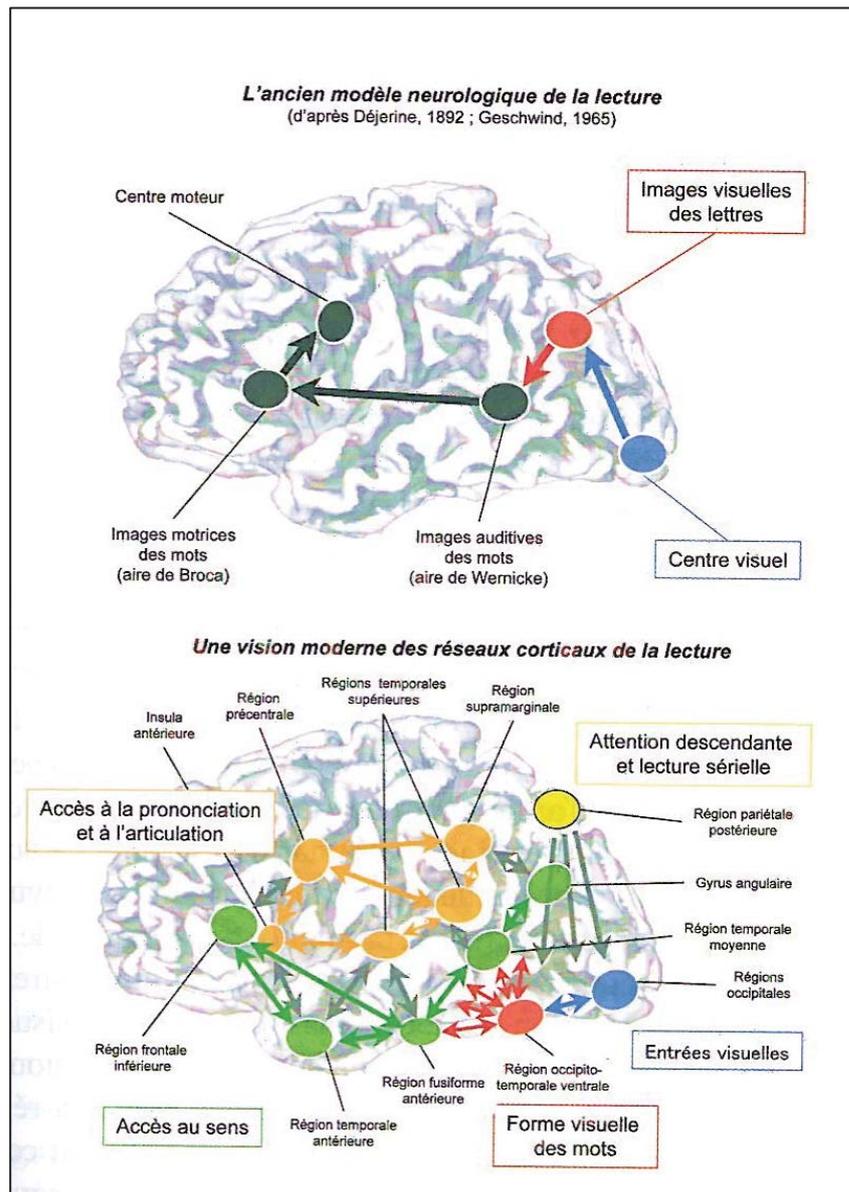


Figure 20 représentation du modèle neurologique classique de la lecture (en haut) en comparaison au schéma parallèle et foisonnant (en bas) qui lui est aujourd'hui préférée (Dehaene, 2007 : 97)

La vision moderne des réseaux corticaux de la lecture met donc en évidence de nombreuses interconnexions entre les aires visuelles, les aires auditives et les aires auditives. Il est intéressant de noter que par rapport à l'ancien modèle neurologique de la lecture, il ne s'agit plus de connexions unidirectionnelles. Les différentes aires sont au contraire interconnectées et comme le souligne l'auteur lui-même (Dehaene, 2007 : 97) nous ne connaissons pas encore en détail toute les connexions qui sont réalisées. Il se pourrait que la connectivité cérébrale sont encore plus foisonnante.

Il est possible d'imaginer que faute d'un apprentissage spécifique de la structure prosodique de la LE, l'apprenant risque d'utiliser les circuits neuronaux déjà mis en place et transférer la structure prosodique de sa LM sur le texte écrit de LE. Ce qui handicaperait considérablement le décodage et la compréhension du texte écrit. Cela pourrait expliquer pourquoi nous avons observé autant de dysfluences en lecture oralisée chez des apprenants, débutants comme avancés.

D'ailleurs en LE, plusieurs auteurs (Golder & Gaonac'h, 1998 ; Germain & Netten, 2005) postulent que l'oral influencerait l'écrit et que les connaissances acquises à l'oral seraient nécessaires à la lecture.

« (...) dans une perspective de développement de la littérarité, l'apprentissage de la lecture est étroitement dépendant, en quelque sorte, des compétences acquises tout d'abord à l'oral » (Germain & Netten, 2005 : 7).

Pour ces auteurs, tant que les formes ne sont pas maîtrisées à l'oral, il y a peu de chances qu'elles le soient à l'écrit :

« L'apprentissage de la lecture et de l'écriture sont deux types de compétences linguistiques qui sont étroitement dépendantes de l'état de développement antérieur des compétences de l'apprenant à l'oral » (idem : 8).

De même Golder & Gaonac'h (1998), postulent que les habiletés de décodage de l'écrit sont étroitement liées aux habiletés de compréhension globale à l'oral et à l'écrit.

Par contre, jusqu'à ce jour, aucun chercheur n'a, à notre connaissance, mis en relation les transferts entre l'oral et l'écrit en langue étrangère avec l'hypothèse de la prosodie implicite de Fodor.

C'est là toute la spécificité de ce travail. Nous postulons que la fluence utilisée par les pédagogues comme critère de mesure de la performance orale des apprenants de LE, reflète en réalité autant la performance prosodique (fluence de production) que la compétence prosodique des locuteurs (fluence cognitive). Nous pensons que cette compétence prosodique est en fait liée à la

mise en place de processus implicites et à leur automatisation. En d'autres termes, nous pensons que la fluence permet de mesurer l'automatisation des processus de décodage associée à la maîtrise de la prosodie de la LE.

Pour pouvoir tester cette hypothèse et mettre en place un protocole de recherche le plus écologique possible, il est nécessaire de s'interroger sur le type de connaissance que constitue la compétence prosodique et sur la manière de l'enseigner.

4.1.2 Qu'est-ce que la compétence prosodique ? Comment l'enseigner ?

À l'instar d'Harmegnies *et al* (2005), nous pensons que la compétence orale, et a fortiori la compétence prosodique, constituent des savoir-faire. En ce sens, nous considérons qu'il s'agit de connaissances procédurales et non déclaratives.

Toutefois, nous ne partageons pas l'idée selon laquelle la compétence visée est une compétence native. Nous rejoignons les questionnements de Hawkins & Towell (1991) : maîtrise-t-on jamais une LE comme sa LM? S'agit-il du même type de connaissances ? Doit-on/peut-on respecter le même ordre d'apprentissage en LM et en LE? Les mêmes méthodes d'apprentissage ?

Ainsi, pour pouvoir définir la compétence prosodique réelle d'un apprenant de LE, nous pensons qu'il est nécessaire de recentrer le champ d'étude sur l'interlangue de l'apprenant. Depuis Selinker (1972) nous savons en effet que les connaissances d'un apprenant de LE ne sont pas juste le résultat de la comparaison entre la LM et la LE, mais qu'il s'agit de connaissances en construction. A partir des années 70, des recherches se sont recentrés sur la l'interlangue de l'apprenant, impliquant des questionnements sur les stratégies individuelles d'apprentissage, sur les différentes sources de variabilité, sur les niveaux d'apprentissage, mais aussi et surtout sur la notion même de maîtrise d'une langue étrangère (Hawkins & Towell, 1991).

Nous nous inscrivons dans cette logique et nous postulons que la compétence prosodique n'est pas une compétence absolue vis à vis d'une norme attendue de prononciation, mais une compétence en construction qui ne peut pas être définie à priori mais doit être analysée à différentes étapes de l'apprentissage. Nous pensons que cette compétence traduit la mise en place de stratégies d'encodage et de décodage adaptées à la LE.

Mais comment enseigner cette compétence prosodique?

Krashen (1985) distingue deux types de connaissances : les connaissances acquises (implicites) et les connaissances apprises (explicites). Pour cet auteur ces connaissances appartiennent à deux systèmes distincts et seules les premières pourraient permettre une maîtrise de la LE comparable à celle de la LM.

Nous avons postulé dans le chapitre 3 que la compétence prosodique était étroitement liée à la fluence. Plus exactement, nous pensons que la fluence reflète la compétence prosodique. Nous avons de plus distingué trois types de fluence, la fluence cognitive, liée à l'automatisation des processus d'encodage, la fluence de production, et la fluence perçue. Si la fluence est le reflet de cette compétence prosodique, elle implique l'automatisation des performances rythmiques (ossature de la prosodie) pour faciliter le décodage et l'encodage de la LE. À l'instar de Paradis (2004) et Germain & Netten (2005) pour qui seule une compétence apprise de manière la plus implicite possible pourrait être automatisée et procéduralisée, nous postulons que l'automatisation de la compétence prosodique nécessite une méthodologie adaptée centrée sur une approche non pas totalement implicite (car peut-on réellement parler d'apprentissage implicite en classe de langues ?) mais le moins explicite possible.

Les deux méthodes que nous avons présentées dans le chapitre 1 offrent en réalité deux approches opposées de l'apprentissage de la prononciation : une approche intellectualisée et explicite pour la méthode articulatoire (MA), et

une approche sans métalangage et plus 'implicite' pour la méthode verbo-tonale (MVT).

Selon la MA, de loin la méthode la plus représentée, une bonne production implique en effet la connaissance métalinguistique de la manière d'articuler des sons, autrement dit, la mise en place de connaissances déclaratives. En conséquence, l'enseignant donne une description articulatoire des différents sons puis invite l'apprenant à répéter le geste articulatoire afin de produire le son cible. Par exemple, pour produire un /u/ l'enseignant dira à l'apprenant de placer sa langue à l'arrière de sa bouche et d'arrondir les lèvres. Dans cette méthode, on met l'emphase sur la production et la répétition de sons isolés, puis de mots isolés contenant le son cible et finalement de phrases. Les paramètres prosodiques sont globalement négligés bien qu'une description métalinguistique des caractéristiques intonatives de la langue cible puisse être envisagée dans les pratiques de classes, le plus souvent en fin de formation.

À l'opposé, la MVT se base sur une approche procéduralisée de l'enseignement de la prononciation. Cette méthode utilise la structure prosodique de la langue cible comme support à l'acquisition de compétences phonologiques en LE. Plus spécifiquement, les patterns rythmiques de la langue cible sont utilisés pour mettre en lumière les spécificités de la langue cible (accentuation, intonation et phonèmes). L'enseignant va aider l'apprenant à se familiariser avec la structure prosodique de la langue cible à travers la répétition de patterns prosodiques types à travers la répétition de logatomes (tels que « lalala » ou « dadada ») ou l'utilisation de gestes facilitateurs (par exemple un battement de la main pour souligner les prééminences). Dans une seconde phase, la structure prosodique est utilisée pour faciliter la perception et la reproduction des phonèmes (cf. chapitre 1). Par exemple, si l'apprenant assombrit le timbre du phonème cible, l'enseignant prononcera ledit phonème dans un contexte prosodique éclaircissant (c'est à dire à l'intérieur d'une syllabe accentuée) et demandera à l'apprenant de répéter le son dans ce nouveau contexte. Autrement dit, un contexte facilitateur aidera l'apprenant à percevoir le bon phonème et l'aidera à reproduire ses traits dans tous les autres contextes

prosodiques (Billières, 2005). La MVT se focalise donc sur un apprentissage non-explicite de la prononciation et la mise en place de connaissances procédurales des éléments prosodiques et segmentaux.

4.1.3 Hypothèses de recherche

Les points que nous venons de soulever, nous permettent de positionner l'apprentissage de la prononciation, et de la prosodie en particulier, comme une priorité en langue étrangère. Maîtriser la structure prosodique de la LE va permettre à l'apprenant de développer des stratégies de décodage et d'encodage adaptées à la LE, à l'oral comme en lecture³.

En lien avec les questions abordées ci-dessus, nous avons formulé plusieurs hypothèses de recherches.

Tout d'abord, nous avons voulu tester l'existence de transferts de compétences entre l'encodage de l'oral et de l'écrit en langue étrangère. Pour se faire, nous avons formulé une première hypothèse selon laquelle un apprentissage à l'oral centré sur la prosodie de la langue étrangère va améliorer la fluence des apprenants en parole lue (Hypothèse 1).

Deuxièmement, comme la prosodie est profondément ancrée dans notre comportement verbal et qu'elle est à la base des processus d'encodage et de décodage du langage (Di Cristo, 2004), nous postulons qu'elle se trouve au cœur de la pratique linguistique et de l'apprentissage d'une langue étrangère. Nous pensons que les résultats seront meilleurs si l'apprentissage de la prosodie est proposé dès le niveau débutant car les apprenants de niveau avancé auront fossilisé leurs erreurs (*cf.* travaux en perception de Wode, 1994) (Hypothèse 2)

Afin d'approfondir ces deux premières hypothèses, nous avons souhaité confirmer le lien entre compétence prosodique et fluence, ce qui impliquerait

³ Initialement, nous souhaitions également comparer la fluence des productions en parole lue et en parole spontanée (*cf.* lien entre prosodie implicite et explicite Fodor, 2002), en postulant que qu'il y aurait une progression similaire dans les deux cas. Nous n'avons pas pu tester cette hypothèse mais nous ne perdons pas de vue l'idée de comparer la parole lue et la parole spontanée dans le cadre de futures recherches (*cf.* chapitre discussion et perspectives).

que la compétence prosodique est bien une connaissance procédurale automatisée, et tester le type d'enseignement nécessaire à la mise en place de cette compétence.

Nous avons donc comparé l'incidence sur la lecture oralisée de deux approches de l'enseignement de la prononciation : une approche non intellectualisée centrée sur la prosodie, avec la Méthode Verbo-Tonale, et une approche métalinguistique de l'enseignement du segmental avec la Méthode Articulatoire.

Nous postulons que seule la MVT aura une incidence positive sur l'amélioration de la fluence en lecture oralisée des apprenants de LE. En effet, nous pensons que la MVT permet le développement d'une compétence prosodique en LE et que la prosodie et la fluence sont étroitement corrélées (Hypothèse 3).

À l'instar de Paradis (2004) et Germain & Netten (2005) nous pensons que seul un apprentissage 'implicite' permet l'automatisation et la procéduralisation des connaissances. Ainsi, nous émettons l'hypothèse selon laquelle l'introduction précoce de l'écrit (i.e. avant que la compétence rythmique soit installée), qui induit obligatoirement l'intellectualisation et l'analyse de la langue, entraîne une chute de la fluence verbale (Hypothèse 4).

Pour tester ces hypothèses nous avons mis en place deux études longitudinales sur une durée de huit semaines chacune, avec des apprenants anglophones de FLE de niveau débutant et avancé. Dans la première étude, étude pilote, nous avons testé nos deux premières hypothèses de recherche. Dans la seconde étude qui constitue véritablement le cœur de ce travail, nous avons testé les quatre hypothèses ci dessus. Ces deux études ont donné lieu à deux publications Alazard *et al* (2010) et Alazard *et al* (2012 a). Les articles sont présentés en annexes (annexe A).

La seconde étude nous a permis de constituer le corpus MULTIPHONIA qui comprend 96h d'enregistrements de cours de communication orale (séances de correction phonétique et activités de compréhension et d'expression orales).

Avant de présenter plus en détails chacune de ces deux études et les résultats obtenus, nous reviendrons sur les choix méthodologiques que nous avons effectués et les difficultés que nous avons rencontrées tout au long de ce travail.

4.2 ASPECTS METHODOLOGIQUES

4.2.1 Apport des études longitudinales à la didactique

Dans les chapitres précédents (chapitre 1 et 2) nous avons insisté sur la nécessité de prendre en compte l'interlangue de l'apprenant pour évaluer sa performance et sa compétence orales. À l'instar de Wenk (1985) nous pensons que seules les recherches sur l'interlangue des apprenants seront à même d'évaluer les compétences réelles d'apprenants en LE. Les transferts prosodiques abordés uniquement sous l'angle de la comparaison entre LM et LE, ne permettent pas de rendre compte de la compétence prosodique en LE. De plus, proposer une étude du rythme de l'interlangue pris à un moment donné (cf. Tortel 2009) ne nous satisfait pas car nous considérons la compétence rythmique comme une compétence en construction. Il est pour nous nécessaire de tester et de mesurer cette dernière par le biais d'études longitudinales.

C'est pourquoi, en dépit de la lourdeur méthodologique qu'implique le longitudinal, il nous a paru indispensable de mesurer la fluence des productions des apprenants à différents moments de leur apprentissage (avant, pendant et après entraînement), pour chacune des deux études réalisées.

Dans leurs recherches, Lennon (1990), Freed (1995) et Freed *et al* (2004) ont mis en place des études longitudinales de la fluence en LE sur une durée de six mois en moyenne, à raison de 2 à 4 heures d'entraînement par semaine. Il n'était pas possible d'envisager de mener deux études longitudinales sur une durée de six mois chacune dans le cadre d'un travail de thèse. Nous avons donc réduit la durée de l'expérimentation à deux mois pour chacune de nos études. Nous avons proposé 2 heures de cours par semaine aux apprenants de l'étude pilote et 3 heures de cours hebdomadaires aux

apprenants de l'étude 2. Nous avons choisi d'augmenter la durée des cours entre la première et la deuxième étude car dans la deuxième étude nous introduisons l'écrit et nous souhaitons que les apprenants aient au moins une dizaine d'heures de cours avant le passage à l'écrit (lors de la 4^{ème} semaine).

Même si la durée de nos études est inférieure à celles des études de Freed et Lennon, nous pensons que si nos hypothèses sont vérifiées, il sera possible d'observer de premières tendances en seulement quelques semaines. En effet, dans nos pratiques de classe nous avons pu observer que la MVT donnait des résultats spectaculaires dès les premières séances.

Lorsque l'on propose une étude hors laboratoire, quelle soit longitudinale ou non, on est également confronté au problème de la gestion des variables. Il devient difficile de contrôler toutes les variables qui rentrent en jeu. Ces dernières sont bien entendu démultipliées en longitudinal. Afin de contrôler le plus de variables possibles, nous avons contrôlé les conditions d'entraînement. Ainsi, contrairement aux études que nous venons de citer, nous avons nous-mêmes créé et enseigné tous les cours que nous avons proposés (groupes contrôles et groupes tests pour les deux études). Cela nous a permis de contrôler le contenu des cours et d'éliminer la variable « enseignant » de notre recherche. Par contre, comme nous avons réalisé notre recherche en milieu exolingue, nous avons dû gérer une grande hétérogénéité parmi les apprenants (variantes régionales de l'anglais, niveau, âge *etc.*).

Avant de discuter de nos choix méthodologiques pour mesurer la fluence de l'interlangue, nous reviendrons sur les difficultés que nous avons rencontrées pour constituer des groupes homogènes puis nous présenterons le dessin expérimental de nos deux études.

4.2.2 Présentation des sujets constituer des groupes homogènes est-ce possible ?

Contrairement à notre attente, il nous a été très difficile de réunir un grand nombre d'anglophones pour réaliser nos deux études. Il semblerait que les

apprenants anglophones ne soit pas la population la plus représentée à Toulouse. Nous avons donc du tolérer une certaine marge de variabilité.

Pour l'étude pilote, nous n'avons pu réunir que 4 sujets anglophones. Les participants étaient âgés de 18 à 40 ans (moyenne d'âge 29 ans). Tous les locuteurs étaient originaires du Royaume-Uni (3 Anglais et 1 Irlandaise) et résidaient en France au moment de l'étude. Ils avaient déjà reçu des cours de FLE, soit au cours de leur scolarité en Angleterre, soit à leur arrivée en France, dans des structures comme l'Alliance Française ou le département d'enseignement du FLE de l'université de Toulouse le Mirail. Lors de cette première recherche, aucun apprenant ne suivait de cours de français en dehors de ceux proposés dans le cadre de cette recherche. Cela constituait un avantage certain pour mesurer l'impact de l'entraînement sur la production des apprenants.

Pour l'étude 2, nous avons pu réunir 20 sujets anglophones⁴. Les participants étaient âgés de 20 à 60 ans (moyenne d'âge 32,5 ans). Il ne nous a malheureusement pas été possible de réunir vingt participants britanniques. Nos sujets étaient originaires de différents pays : l'Angleterre, les Etats-Unis, le Pays de Galles, l'Irlande, l'Écosse, l'Australie, Singapour et Hawaï (5 anglais, 7 américains, 2 Écossais, 1 galloise, 2 irlandais, 1 Australien, 1 Indonésienne et 1 Hawaïen). Nous avons conscience que ces différences peuvent entraîner des différences dans les productions des apprenants, nous savons par exemple que le rythme Singapourien et le rythme Hawaïen présentent tous deux des structures rythmiques différentes du rythme britannique (Bolinger, 1981 ; Deterding, 2001). Afin de limiter l'impact de cette variable, nous avons, dans la mesure du possible, répartis équitablement les apprenants de différentes nationalités au sein des deux méthodes de correction phonétique testées. Tous les locuteurs résidaient en France au moment de l'étude. Ils avaient tous préalablement suivis un enseignement en français, soit au cours de leur scolarité dans leur pays d'origine, soit à leur arrivée en France, dans des structures comme l'Alliance Française, l'Institut Catholique de Toulouse (ICT) ou

⁴ Les sujets ont répondu à un questionnaire d'auto-évaluation avant et après la formation. Un exemplaire vierge de ces questionnaires est présenté en annexes (annexes D et E).

le Département d'Enseignement du FLE (DEFLE) à l'université de Toulouse le Mirail. En parallèle des cours proposés dans le cadre de cette recherche, la grande majorité des apprenants suivaient des cours de français à l'ICT, au DEFLE, à l'association Langue Onze à Toulouse⁵, à l'association Plume et Parole à Castanet-Tolosan ou au centre social de Grenade (14/20 des apprenants étaient concernés soit plus de 70%). Afin de compenser au mieux cette variable, nous avons, dans la mesure du possible, équitablement répartis les étudiants parmi les deux méthodes testées.

Un autre problème auquel nous avons été confronté, est celui de l'évaluation du niveau des apprenants. En effet, nous nous sommes rapidement aperçu que les niveaux théoriques des apprenants ne correspondaient pas à leur niveau réel en production orale. Nous avons donc réévalué le niveau des apprenants sur la base d'un entretien oral semi-directif selon l'échelle proposée par le CECRL. Ce test de niveau, inspiré des épreuves d'évaluation de la production et de la compréhension orales du DELF-DALF, comprenait plusieurs activités : un entretien guidé, des répétitions de phrases et une activité de compréhension orale. Le test de niveau et les grilles d'évaluation du CECR sont présentés en annexe (respectivement Annexe B et Annexe C). Le problème que nous avons rencontré soulève la question de l'évaluation de la performance des apprenants en langue étrangère. Nous y reviendrons (*cf.* chapitre 5 et conclusion).

Pour l'étude pilote, deux apprenants ont été évalués à un niveau élémentaire (niveau A dans l'échelle proposée par le CECRL) et deux apprenants ont été évalués à un niveau avancé (niveau B du CECRL). Pour la seconde étude, dix apprenants ont été évalués à un niveau élémentaire (dont 4 ont été évalués à un niveau Introductif, niveau A1 du CECRL et 6 à un niveau Intermédiaire, niveau A2 du CECRL) et 10 apprenants ont été évalués à un niveau « avancé » (niveau B1 du CECRL). Nous précisons que deux apprenants de niveau débutant ont abandonné le protocole au bout de quelques semaines

⁵ Nous tenons à remercier ces organismes qui nous ont aidé à recruter nos sujets. Nous tenons tout particulièrement à remercier Mesdames Bozena Billerey et Annie Despatureaux de l'ICT, Mesdames Cécilia Gunnarson et Chantal Domp martin du DEFLE et toute l'équipe de Langue Onze.

pour raisons personnelles. Dans la suite de ce travail, nous présenterons donc les résultats de 18 apprenants (8 apprenants de niveau débutant et 10 apprenants dits de niveau avancé).

Afin de tester la corrélation entre l'amélioration de la fluence des productions et le niveau des apprenants (hypothèse 2), les apprenants ont été équitablement répartis au sein des groupes tests et des groupes contrôles pour les deux études.

Nous avons essayé de maîtriser l'impact de ces variables au maximum en tâchant de répartir équitablement les sujets dans les différents groupes. Nous restons consciente que des différences peuvent persister entre les groupes et nous amener à nuancer les résultats obtenus. D'autant plus que les deux apprenants débutants de l'étude 2 qui ont abandonné le protocole n'avaient pas le même niveau : l'apprenant du *groupe test* avait été évalué à un niveau A1 et l'apprenant du *groupe contrôle* avait été évalué à un niveau A2.

Néanmoins, la variabilité entre les apprenants, même au sein des groupes de niveau, est une réalité des pratiques de classes que les enseignants doivent gérer au quotidien. Ces études s'inscrivent donc dans une démarche écologique de recherche en didactique. C'est notamment cette dimension écologique qui fait la richesse du corpus MULTIPHONIA que nous avons constitué en filmant les séances de cours de la deuxième étude réalisée (cf. partie 4.2.4)

4.2.3 Dessin expérimental des études

4.2.3.1 Étude pilote

Dans l'étude pilote, nous avons testé nos deux premières hypothèses de recherche. Nous pensons qu'un apprentissage à l'oral centré sur la prosodie de la langue étrangère va améliorer la fluence des apprenants en parole lue (Hypothèse 1) et que cet apprentissage aura plus d'impact chez des apprenants de niveau débutant (Hypothèse 2). Pour vérifier la validité de ces hypothèses, nous avons comparé deux méthodes d'enseignement différentes : la

méthodologie *Structuro Globale Audio Visuelle* (SGAV) via la *méthode Verbo-Tonale* d'intégration phonétique (MVT), et les *approches communicatives*, qui prédominent l'enseignement du FLE depuis les années 1980. Même si elles reconnaissent toutes deux l'importance de l'oral dans l'enseignement des LE (cf. chapitre 1), ces deux méthodologies en proposent une approche sensiblement différente. Si l'oral et la maîtrise des paramètres prosodiques constituent une priorité pour le SGAV, les approches communicatives négligent quand à elles l'enseignement de la prononciation.

Le premier groupe (*groupe test*) a suivi des séances de correction phonétique centrées sur la sensibilisation à la structure prosodique du français : nous avons particulièrement travaillé le rythme et l'intonation, par la répétition de phrases types, les vires langues ainsi que des jeux de rôle. Le deuxième groupe (*groupe contrôle*) a suivi des cours de compréhension écrite 'classique' dans une approche communicative : nous avons proposé des activités mettant en jeux les quatre savoir-faire visés en didactique (cf. tableau 1 p. 32). Les sujets ont été équitablement répartis dans chaque groupe en fonction de leur niveau en français (cf. partie 4.2.1.1 pour la discussion sur le niveau des apprenants).

Nous avons testé les apprenants avant et après leur formation (respectivement t0 et t1). Les sujets ont été enregistrés individuellement dans une salle insonorisée du laboratoire Octogone-Lordat. Les enregistrements ont été réalisés à l'aide de micro-casques. Les conditions d'enregistrement étaient identiques pour les deux tests, avant et après les 8 semaines d'entraînement linguistique. Nous avons testé les apprenants dans trois tâches complémentaires : une activité de lecture oralisée d'un texte argumentatif, une activité d'argumentation orale à partir d'un texte écrit (production orale dirigée), et enfin, une improvisation sur un sujet de type « pour ou contre » (production orale semi-dirigée). À travers ces trois activités, nous avons voulu proposer une progression de la parole lue vers la parole spontanée. Initialement, nous souhaitons en effet comparer la progression des apprenants dans différents types de parole. Nous avons choisi d'utiliser l'argumentation comme fil rouge de notre test car le discours argumentatif implique l'utilisation de connecteurs

spécifique, du type '*parce que*', qui permettent de comparer un matériau linguistique similaire en oral spontané et en lecture oralisée. De plus, les connecteurs participent à la structuration du discours et « (...) *facilitent le repérage des frontières syntaxiques et sémantiques* » (Ziti, 1995). De fait, ce sont des candidats idéaux pour une analyse de l'interface prosodie/discours. Le discours argumentatif implique également un investissement du locuteur, ce qui facilite l'utilisation de figures prosodiques riches. Toutefois, il s'est avéré que les apprenants (en particulier les apprenants de niveau débutant) utilisaient peu de connecteurs en parole spontanée. Nous avons donc recentré nos analyses sur les corpus de lecture oralisée, recueillis en début et en fin de formation.

Le texte utilisé lors des deux tests est un texte argumentatif que nous avons créé en respectant la structure argumentative et thématique des autres activités proposées. Il se compose de onze phrases et est divisé en quatre paragraphes qui reprennent chacun un argument en faveur de la pratique du vélo en ville (cf. figure 21)

**Depuis dix ans, de plus en plus de personnes préfèrent le vélo à la voiture.
Pourquoi ce changement?**

De nombreux médecins recommandent le vélo parce qu'il permet de lutter contre les problèmes de poids. La pratique du vélo augmente la capacité respiratoire. Le vélo est donc conseillé aux asthmatiques mais aussi aux personnes souffrant d'arthrose et de rhumatismes, parce que, quand on fait du vélo, on étire sa colonne vertébrale, on assouplit les muscles et on supprime les maladies dues au manque d'exercice.

De plus, il est communément admis que le vélo est un moyen de locomotion facile et agréable, comme en témoigne Judith, étudiante : « Avant, je me déplaçais tout le temps en voiture, même pour de courts trajets. Mais, depuis quelques mois, je me suis mise au vélo, je trouve ça plus agréable. Je longe le canal et je n'ai plus l'impression d'être en centre ville! Et puis, le vélo, c'est quand même plus pratique que la voiture. On circule partout, on se gare facilement ... parce qu'en ville trouver un parking c'est pas facile... et surtout parce qu'à vélo on ne tombe pas en panne ! »

Enfin, le vélo est écologique, comme le constate un vendeur : « La vente de vélos a doublé ses dernières années parce que les gens se sentent concernés par l'environnement. En effet, comme on avance à la force de ses mollets, on ne pollue pas ! Les jeunes n'utilisent presque plus la voiture, sauf pour les longs trajets ».

Alors, dépêchez-vous parce que, avec l'arrivée des beaux jours et du soleil, on a envie de sortir, de se balader, et bientôt il n'y aura plus de vélos en magasins !

Figure 21 Texte argumentatif utilisé pour cette étude

De récentes études ayant montré l'importance de l'architecture textuelle dans les stratégies d'oralisation (Maurel 2004), la mise en page proposée a été adaptée au déroulement argumentatif.

Les sujets avaient pour consigne de s'approprier le texte par une ou plusieurs lectures silencieuses avant d'être enregistrés. Il leur était également demandé de rendre au mieux le déroulement argumentatif du texte. Nous précisons que ce type d'activité est tout à fait courant en cours de FLE, et de fait la tâche de lecture oralisée en elle-même ne semblait pas poser de problèmes aux apprenants, quel que soit leur niveau de langue.

4.2.3.2 Étude 2

Dans cette deuxième étude longitudinale, nous avons voulu confirmer le lien entre compétence prosodique et fluence. En effet, nous avons conscience que les tendances qui pourraient se dégager de l'étude pilote, ne permettraient pas de conclure directement que l'enseignement de la prosodie est responsable de l'amélioration de la fluence des apprenants.

C'est pourquoi, dans le cadre de cette deuxième étude, nous avons choisi de comparer non plus deux méthodologies d'enseignement des langues étrangères (SGAV vs. approches communicatives) mais deux méthodes de correction phonétique radicalement différentes: la méthode articulatoire (ART) qui considère que l'acte phonatoire implique la connaissance explicite de son fonctionnement et qui néglige les facteurs prosodiques et la MVT qui accorde la priorité à la prosodie et refuse l'intellectualisation. Nous formulons l'hypothèse selon laquelle seule la MVT aura une incidence positive sur l'amélioration de la fluence en lecture oralisée des apprenants de LE (Hypothèse 3). Nous avons constitué deux *groupes MVT* (niveau débutant et niveau avancé) qui ont donc suivi des cours de correction phonétique via la MVT et deux *groupes ART* (niveau débutant et niveau avancé) qui ont suivi des cours de correction phonétique par la méthode articulatoire.

Dans cette étude, nous avons également voulu évaluer l'incidence du passage à l'écrit sur la fluence de la parole lue. Nous pensons que l'intellectualisation nécessaire au passage à l'écrit risque de perturber la mise en place de la compétence prosodique si l'écrit est introduit trop tôt (Hypothèse 4).

Afin de tester également l'existence de transferts entre l'oral et l'écrit (Hypothèse 1), nous avons divisé cette seconde étude en deux temps. Pendant les trois premières semaines (6 séances de cours) nous avons proposé un travail centré uniquement sur l'oral (sans prise de note ou support écrit) pour tous les groupes (*MVT* et *ART*). Selon la méthodologie SGAV, l'écrit ne devrait être introduit qu'après 60h d'enseignement exclusivement consacrées à l'oral. Il est aujourd'hui difficile de trop retarder le passage à l'écrit car cela ne correspond pas aux habitudes d'apprentissage et les apprenants risquent de se démotiver rapidement. Néanmoins, nous pensons qu'il est nécessaire de différer l'écrit de quelques dizaines d'heures pour faciliter la mise en place et l'automatisation des transferts prosodiques entre l'oral et l'écrit en LE.

C'est pourquoi nous avons introduit l'écrit avec tous les groupes (*MVT* et *ART*) lors de la quatrième semaine de cours. Idéalement, nous aurions souhaité introduire l'écrit lors de la cinquième semaine de cours, afin de respecter un

équilibre temporel entre les semaines uniquement consacrées à l'enseignement de l'oral et les semaines où nous avons introduit l'écrit. Malheureusement, nous avons été soumise aux contraintes du calendrier universitaire. Nous avons dû interrompre cette étude pendant deux semaines après trois semaines de cours. Nous avons donc choisi de consacrer les trois premières semaines à l'enseignement de l'oral et d'introduire l'écrit après l'interruption des activités pédagogiques, soit durant la quatrième semaine de l'étude. Le passage à l'écrit en langue étrangère pose un vrai problème méthodologique : faut-il reprendre les bases de la correspondance graphème-phonème ? Faut-il proposer un apprentissage de la lecture selon la méthode globale ? Faut-il considérer, au contraire, que les capacités de décodages apprises en LM seront automatiquement réutilisées en LE ? Les enseignants de LE se cachent le plus souvent derrière la troisième possibilité. Il n'y a donc pas à l'heure actuelle de véritable méthodologie de l'enseignement de l'écrit en LE. À partir de nos connaissances sur l'apprentissage de la lecture (Dehaene, 2007 ; Ecalle & Magnan, 2003), nous avons proposé un enseignement du code écrit basé sur une approche à la fois globale et analytique du code écrit. Nous avons tenu compte de la philosophie des deux approches de correction phonétique que nous avons mis en place en adaptant le passage à l'écrit à chacune des deux méthodes. Ainsi, nous avons proposé un passage à l'écrit d'abord global puis analytique aux apprenants des groupes MVT et seulement analytique aux apprenants des groupes ART. Nous restons consciente de la nécessité d'approfondir la question de l'introduction de l'écrit en LE dans de nouvelles recherches.

Nous avons testé les apprenants en lecture oralisée à trois reprises, avant la formation (s0) après trois semaines d'entraînement exclusivement à l'oral (s1) puis à la fin des huit semaines (s2) dans des conditions expérimentales identiques. Les sujets ont passé les tests individuellement dans la salle d'enregistrement du Plateau d'Études techniques et de recherches en audition⁶ (PETRA), à l'université de Toulouse le Mirail. Les tests comprenaient chacun deux activités : une lecture oralisée d'un texte informatif inspiré d'un

⁶ Nous remercions Monsieur Julien Tardieu pour l'aide qu'il nous a apporté pour les enregistrements.

manuel de FLES et des questions de compréhension écrite, à l'oral, sur le texte lu. Nous n'avons pas analysé les corpus de compréhension écrite mais nous envisageons de développer ces aspects dans nos prochaines recherches (cf. partie conclusions et perspectives).

Contrairement à l'étude pilote, nous avons proposé trois textes différents à la lecture à chaque sujet. Afin d'éviter l'effet d'habituation, nous avons donc créé trois textes pour chaque niveau. Les textes ont été randomisés et présentés dans un ordre différent pour compenser les éventuels écarts de difficultés entre les trois textes.

Pour créer ces supports, nous nous sommes inspirés de textes extraits de manuels de FLES, *Alter Ego* ou *A propos* niveau A et B. Ces méthodes sont très utilisées dans les cours de langues, les apprenants étaient donc habitués à lire et à travailler ce type de document. Les textes extraits des manuels ont été remaniés pour les besoins de notre étude : nous avons ajouté des phrases complexes (longueur des phrases, énumérations, pourcentages, chiffres *etc.*), des mots transparents ('télévision'), des anglicismes ('fanatique') ainsi que des liaisons ('vingt heures') et des sons problématiques (/mwa/) pour des anglophones (cf. figure 22).

Les français et la télévision

La télévision occupe une place importante dans la vie quotidienne des français.

Vincent 36 ans, commercial, marié, deux enfants, témoigne : « Chez nous, toute la famille regarde la télé mais rarement les mêmes programmes. Le matin, à 7h30, les enfants prennent le petit-déjeuner devant le petit écran : ils regardent une émission pour les jeunes jusqu'à 8 heures, et ma femme et moi nous nous préparons tranquillement. Après l'école, les enfants se reposent une demi-heure devant la télé avec les dessins animés et, le soir, ils se couchent tôt. Avec ma femme, on regarde toujours le 20 heures. Après, le choix est difficile parce que moi, je suis un fanatique du foot et ma femme, elle préfère regarder les magazines de société ou les comédies romantiques avec Richard Gere! »

Figure 22 Exemple de texte utilisé pour le niveau débutant. Les cinq autres textes sont présentés en annexe (Annexe F).

Pour cette étude, nous avons choisi d'utiliser un texte informatif plutôt qu'un texte argumentatif pour ne pas ajouter de difficulté supplémentaire. Dans

l'étude précédente, nous nous sommes aperçue que l'argumentation pouvait poser problèmes aux apprenants débutants. En effet, l'apprentissage des schémas argumentatifs est le plus souvent proposé aux niveaux avancés.

Comme dans l'étude pilote, les sujets avaient pour consigne de s'appropriier le texte par une ou plusieurs lectures silencieuses avant d'être enregistrés. Après avoir pris connaissance du texte, les sujets lisaient une fois le texte à voix haute avant de répondre aux questions de compréhension posées par l'examineur.

4.2.4 Comment mesurer la fluence de l'interlangue ?

Comme nous l'avons soulevé au chapitre 3, il n'existe pas de consensus sur la définition de la fluence ou les outils de mesures utilisés. La majorité des études se concentrent soit sur la production des apprenants soit sur la perception qu'ont les natifs des dites productions. Rares sont les études qui s'intéressent à ce que la fluence induit comme changements cognitifs chez le locuteur.

Pourtant, plusieurs auteurs (Towell *et al*, 1996 ; Segalowitz, 2000 ; Dewaele, 2002 ; Derwing *et al*, 2004 ; Segalowitz, 2010) soutiennent l'idée selon laquelle la fluence serait liée à la mise en place de changements profonds dans l'encodage et le décodage de la langue, tels que la conversion d'un savoir déclaratif en un savoir procédural et l'automatisation des processus d'encodage. L'évaluation de la fluence à différents moments de l'apprentissage nous permettrait donc de mesurer la mise en place de ces changements et de rendre compte de la construction de l'interlangue des apprenants.

Pour réaliser cela, il nous a semblé primordial de prendre en compte à la fois les aspects cognitifs, la production et la perception dans l'évaluation de la fluence. La modélisation proposée par Segalowitz (2010) permet de rendre compte de ces trois dimensions. Segalowitz (2010) néglige la dimension prosodique dans sa modélisation. Nous postulons au contraire que la fluence cognitive est associée à la compétence prosodique. Sachant que la fluence de production est déterminée par l'évaluation de paramètres temporels (Olynyk *et*

al, 1987 ; Lennon, 1990 ; Towell *et al*, 1996 ; Cenoz, 2000 ; Cucchiarini *et al*, 2000 ; Cucchiarini *et al*, 2002 ; Ullakonoja, 2008), nous pensons que la fluence de production et la fluence cognitive sont étroitement liées et qu'elles permettent de rendre compte de la construction de la compétence prosodique en LE.

Pour les analyses acoustiques nous avons donc retenu principalement des paramètres prosodiques.

4.2.4.1 Analyses quantitatives de la fluence de production

En accord avec la littérature (*cf.* chapitre 3), les paramètres prosodiques retenus dans les deux études, car susceptibles de refléter au mieux la fluence des apprenants étaient :

- le nombre et la durée des différents types de pauses (pauses respiratoires et pauses silencieuses). Nous avons également tenu compte de la place des pauses (*cf.* Olynyk *et al*, 1987 ; Cucchiarini *et al*, 2000). Les pauses silencieuses (vides et respiratoires) ont donc été classées selon un critère de grammaticalité: les pauses situées à l'intérieur d'un mot ou d'un syntagme, et entre deux mots répétés ont été considérées comme agrammaticales.
- le nombre, la distribution (accent initial vs. accent final), la durée des syllabes accentuées et celle des syllabes inaccentuées. Les accents finaux ont également été étiquetés en fonction du niveau de frontière prosodique auquel ils étaient associés (mineur vs. majeur) (Astésano, 2001). Pour mesurer le débit, nous avons pris en compte la durée des syllabes inaccentuées (Pasdeloup, 2004).
- le nombre et la durée des différents types de disfluences : les hésitations sonores, les faux-départs et les répétitions. Dans l'étude 2, nous avons distingué les hésitations sonores lexicales et non lexicales et nous avons relevé les « reparans » qui correspondent à la reprise de l'énoncé (Paullaud, 2005) et sont à même d'indiquer la capacité d'autocorrection de l'apprenant.

Dans la deuxième étude, nous avons également pris en compte la durée moyenne des segments et le taux d'articulation (Cucchiarini *et al*, 2002). Pour la durée des segments, nous avons mesuré le nombre moyen de syllabes entre deux pauses > 220 ms, selon les valeurs de l'unité prosodique minimale de discours (*i.e* l'Unité Interpausale, désormais IPU), données par Bertrand & Espesser (2001). Nous avons mesuré le taux d'articulation à partir de la durée du nombre total de syllabes divisé par la durée totale de temps de parole, en excluant les pauses (tous types de pauses confondus).

Nous avons remarqué que les apprenants de FLE produisent un grand nombre d'accent non-standards, réalisés avec un contour intonatif de type 'demandes de confirmation'. Nous avons donc différencié les accents finaux produits avec des degrés de frontières attendus et les accents finaux réalisés avec des degrés de frontières inattendus (une excursion plus longue de la F0 et une durée plus longue de la syllabe accentuée), typiques d'une demande de confirmation en LE. Le degré de frontière attendu a été déterminé a priori, avant l'annotation, à partir de notre expertise de francophone natif.

Les données de lecture oralisée ont été recueillies à l'aide d'un logiciel d'enregistrement audio (Soundtrack-Pro ©; 16 bits/44kHz), puis étiquetées et analysées avec le logiciel *Praat* (v.5.3.39) (Boersma & Weenink, 2005).

Pour l'étude pilote, les 8 corpus de lecture recueillis (4 sujets * 2 tests, soit environ 20 minutes de corpus) ont été écoutés puis annotés manuellement par deux experts. La figure 23 illustre les différents niveaux d'annotation choisis : le niveau phrase, le niveau pause, le niveau syllabe et le niveau accent.

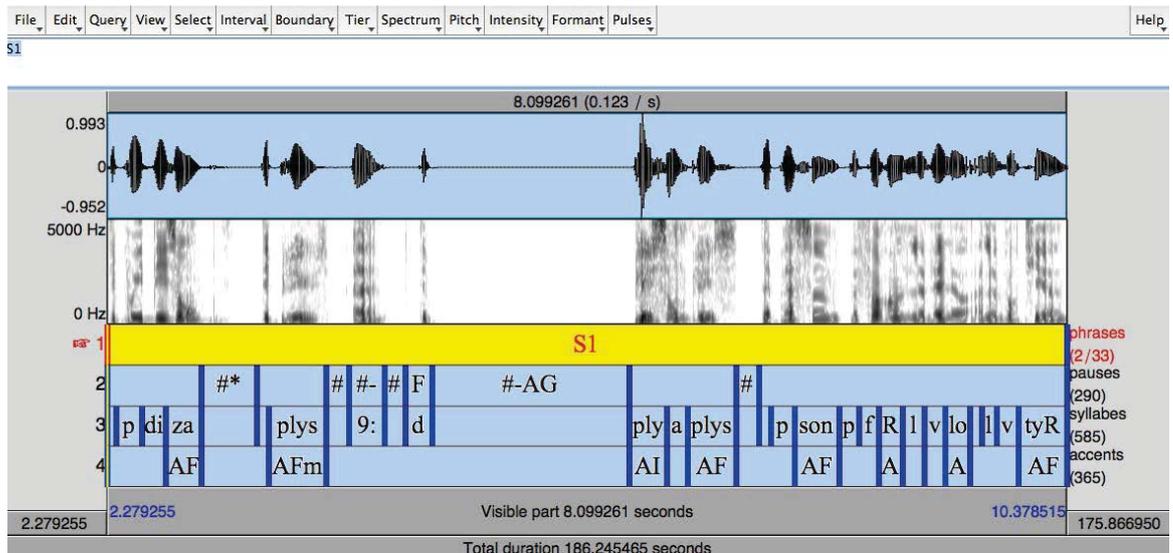


Figure 23 Exemple de phrase du corpus annotée manuellement sous Praat. On peut distinguer 4 niveaux d'annotation, les phrases (S1), les pauses (# : pauses vides, #* : pauses respiratoires, #AG pauses agrammaticales), les syllabes transcrites en SAMPA et les accents (AI : Accent Initial, AFmin : Accent Final mineur, AFmaj : Accent Final majeur).

L'étude 2 comprenait 54 corpus de lecture (18 apprenants* 3 lectures) soit presque 3 heures de parole à transcrire. En raison de la lourdeur de l'étiquetage manuel, nous avons préféré un alignement automatique⁷ des phrases à l'aide du logiciel Speech Phonetization Alignment and Syllabification (SPPAS) (Bigi & Hirst, 2012). Cet outil a pour objet de permettre la segmentation automatique de phrases, de mots, de syllabes et de phonèmes à partir d'un enregistrement audio et de sa transcription orthographique. SPPAS produit une transcription phonétique sur la base d'un dictionnaire phonétique. L'outil offre de plus la possibilité de sélectionner (manuellement ou automatiquement) parmi toutes les variantes phonémiques possibles. L'alignement se fait dans le cadre de chaque groupe rythmique (ou IPU) pour maximiser les effets.

Les corpus de lecture recueillis ont donc d'abord été transcrits phrase par phrase par nos soins selon les conventions de la Transcription Orthographique Enrichie (TOE) (voir Bertrand *et al*, 2008 pour les conventions de transcription). Nous avons donc noté les pauses silencieuses grammaticales (pauses > 220 ms) marquées par une frontière de tire, les pauses silencieuses agrammaticales (notées +) les hésitations (« ee »), les amorces (am-) et les écarts de

⁷ Nous remercions chaleureusement Madame Brigitte Bigi pour l'aide qu'elle nous a apportée.

prononciation ([exemple, egzemplé]) par rapport à la norme attendue de prononciation du français (cf. figure 24).

« Chez nous toute la [famille, famili] regarde la télé mais rarement les mêmes pro- + programmes »
« Les jeunes ont + plusieurs possibilités pour faire leurs ee [études, étudette] »
« Beaucoup vont a le + la + université pour préparer + une + licence »

Figure 24 Exemple de phrases extraites du corpus des débutants transcrites suivant les conventions de la TOE

La TOE offre l'avantage de proposer une transcription orthographique qui prend en compte les aspects spécifiques de la parole comme les phénomènes de pauses, d'hésitations, de faux-départs ou de répétitions. Cette transcription permet de rendre compte du discours tel qu'il a été produit et non pas d'un discours normalisé. Cela prend évidemment encore plus de sens en LE, aussi bien en parole spontanée qu'en parole lue, car il n'existe pas de normes de transcription de la LE. La TOE permet de générer automatiquement deux transcriptions : une transcriptions orthographique standard dont il est possible d'extraire des 'tokens orthographiques' pour des analyses sémantiques, syntaxiques ou discursives (voir par exemple Blache *et al*, 2009), et une deuxième transcription sélective dont on peut extraire des 'tokens phonétiques' pour une conversion graphème-phonème.

Les phrases transcrites ont ensuite été alignées automatiquement grâce au logiciel SPPAS. Afin de limiter le nombre d'erreurs de l'aligneur automatique, les alignements proposés ont été vérifiés un à un. Nous avons notamment constaté de nombreuses erreurs dans le découpage des consonnes occlusives sourdes. L'aligneur automatique place généralement les frontières des occlusives sourdes au milieu de la tenue de l'occlusive, sur la base d'un calcul moyen de la tenue des occlusives. Nous avons ensuite complété la transcription par l'ajout manuel des différents types d'accents, des différents types de pauses et des disfluences pour chaque locuteur (cf. figure 25).

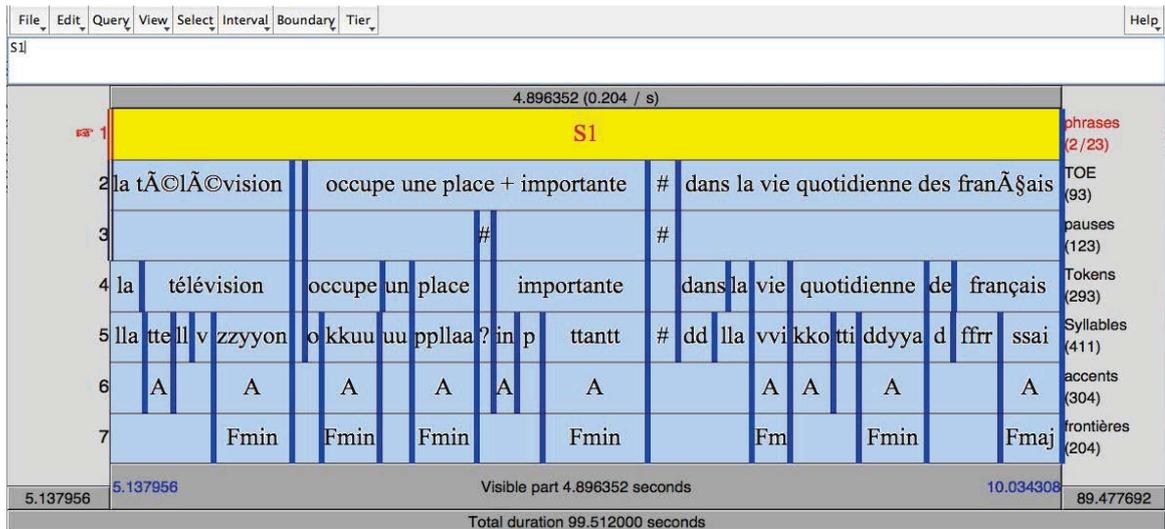


Figure 25 Exemple de phrase du corpus alignée automatiquement puis manuellement sous Praat. On peut distinguer 7 niveaux d’annotation, les phrases (S1), les transcriptions via la TOE, les pauses (# : pauses vides), les tokens lexicaux, les syllabes transcrites orthographiquement selon les conventions de SPASS, les accents (A) et les niveaux de frontières (Fmin : accent final mineur, Fmaj : accent final majeur).

Nous aurions pu nous arrêter là et proposer uniquement des analyses de mesures acoustiques de la fluence des productions des apprenants. Néanmoins, l’enseignant de langue étrangère évalue rarement la production des apprenants à l’aide de logiciels d’analyse de parole. L’évaluation proposée en classe est une évaluation perceptive. Nous avons conscience des limites de ce type d’évaluation (cf. réflexions sur le crible de l’enseignant Baqué, 2004). Toutefois, cette évaluation correspond à une situation de classe réelle et les aspects suprasegmentaux seraient également déterminants dans l’évaluation perceptive de la fluence (cf. Riegenbach, 1991 ; Freed, 1995 ; Cucchiariini *et al*, 2002 ; Freed *et al*, 2004 ; Ullakonoja, 2008). C’est pourquoi nous proposons également des analyses qualitatives de la perception des productions des apprenants par des natifs. Les auditeurs natifs choisis étaient enseignants ou futurs enseignants de FLE et nous les avons considérés comme experts de l’évaluation des productions orales des apprenants.

4.2.4.2 Analyses qualitatives de la fluence perçue

Pour l’étude pilote, nous avons demandé à douze futurs enseignants de FLES francophones natifs (experts) d’évaluer les productions des apprenants. Le test de perception a été réalisé avec le programme Presentation©. Les sujets

devaient écouter 264 phrases regroupées en trois séries de 88 phrases (11phrases*4 locuteurs*2 états*3 écoutes). Les phrases étaient présentées en mode aléatoire à l'intérieur de chaque série. Nous avons créé 6 séries afin de proposer une combinaison de trois séries de phrases différentes pour chaque auditeur. Le test durait environ une heure. Le temps indiqué incluait la phase d'entraînement, la phase de test et le temps de pauses (2 minutes entre chaque série).

Nous avons demandé aux auditeurs de juger la fluence des phrases sur une échelle de magnitude de 1 à 7 (1 : apprenant débutant – 7 : francophone natif). Cette échelle a été choisie de manière à reprendre les six niveaux du CECRL (A1 à C2). Nous avons ensuite ajouté une 7^{ème} valeur, correspondant à la production d'un natif.

La principale difficulté que nous avons rencontrée a été de poser une consigne claire mais qui n'induit pas de réponses attendues. Évaluer la fluence, bien que cela soit régulièrement demandé en cours de langue, pose problèmes aux enseignants. Pour être le plus précis possible sans pour autant orienter les réponses des auditeurs, nous avons donné la consigne suivante : **« Vous allez écouter des phrases lues par des apprenants de FLE. Vous allez ensuite juger la fluence de ces phrases (autrement dit est-ce que les mots s'enchaînent naturellement?) sur une échelle de 1 à 7 avec 1= apprenant débutant et 7= français natif »**. Nous avons conscience que l'évaluation de la fluence via un test perceptif rend compte d'une appréciation globale de la production des apprenants et qu'il est difficile de contrôler les critères qui se sont avérés pertinents pour l'évaluation.

Pour combler ce biais, lors de la deuxième étude, nous avons demandé aux auditeurs de préciser à l'issue du test les critères qu'ils avaient utilisés pour évaluer la fluence des productions.

Dans la seconde étude, nous avons demandé à vingt-deux enseignants et futurs enseignants de FLES francophones natifs (experts) d'évaluer la fluence des productions des apprenants. Contrairement à l'étude pilote, nous n'avons

pas pu proposer l'intégralité des textes à l'écoute car le texte perceptif aurait alors duré plus de 9 heures. Nous avons donc sélectionné une phrase par texte. Les phrases présentaient une structure similaire soit un ou deux détachement(s) en début ou en fin de phrase. À l'instar de Astésano *et al* (2008) nous considérons qu'un syntagme est dit détaché lorsqu'il apparaît à la périphérie gauche ou droite de la phrase. Ce détachement entraîne l'apparition d'un schéma prosodique type : possibilité d'augmentation de la durée du syntagme et présence d'une pause vide d'une durée supérieure à 220ms entre le syntagme détaché et le reste de la phrase. Chaque phrase présentait également au moins une liaison mais ne contenait pas de mots compliqués. Enfin la longueur des phrases était représentative du niveau des apprenants (voir tableau 4).

Niveau	Phrases sélectionnées
Niveau débutant (A)	De plus, les gratuits sont distribués à la sortie du métro ou aux arrêts de bus.
	En France, après le bac, les jeunes ont plusieurs possibilités pour faire leurs études.
	Chez nous, toute la famille regarde la télé mais rarement les mêmes programmes.
Niveau indépendant (B)	La presse gratuite est par contre toujours mal acceptée par les journaux payants parce qu'elle représente un concurrent de taille.
	On peut aussi préférer voyager et dans ces cas, il est également possible de préparer de nombreux diplômes à distance.
	En rentrant de l'école, les enfants se reposent une demi-heure devant la télé avec les dessins animés et, le soir, ils se couchent tôt.

Tableau 7 Phrases utilisées pour le test de perception

Par rapport à l'étude pilote, nous avons modifié deux aspects de la consigne. Nous avons tout d'abord réduit l'échelle de magnitude proposée car l'étendue de cette dernière s'était avérée problématique pour la plupart des auditeurs de l'étude précédente. Nous avons donc réduit l'échelle de 1 à 5 au lieu de 1 à 7. Nous perdons l'évaluation de la fluence par rapport aux niveaux proposés par le CECRL mais comme nous l'avons dit précédemment les niveaux établis par le Cadre suscitent des réserves (*cf.* sous partie 4.2.2).

De plus, dans cette étude nous avons pris en compte le temps de réponses des auditeurs, en supposant que la rapidité des temps de réponses serait corrélée avec la confiance des auditeurs dans leur réponse.

Ainsi, la consigne donnée aux auditeurs était la suivante : « **Vous allez écouter des phrases lues par des apprenants de FLE. Vous devez évaluer le niveau de fluence de chacune des productions sur une échelle de 1 à 5 (1 : mauvaise fluence ; 5 : bonne fluence). Donnez votre réponse *le plus rapidement possible*** ». Les experts devaient donc écouter les 162 phrases regroupées en trois blocs de 54 phrases (3 phrases*18 locuteurs*3 écoutes) et juger de la fluence des productions sur une échelle de magnitude de 1-5.

Nous avons réalisé ce test avec le logiciel *SuperLab*. Nous avons créé 6 blocs : un bloc 'instructions', 1 bloc 'entraînement', 3 blocs 'test' (contenant chacun les 54 phrases randomisées) et 1 bloc 'pause'. Les phrases présentées étaient randomisées non seulement à l'intérieur des blocs mais aussi par sujet. Les sujets pouvaient prendre une pause de quelques minutes entre chaque bloc. La durée du test était d'environ 45 minutes, incluant la phase d'entraînement, la phase test et les temps de pauses.

Les sujets ont passé le test individuellement dans une salle insonorisée du laboratoire Octogone-Lordat. Chaque sujet était équipé d'un casque et assis face à un ordinateur sur lequel était branché un boîtier réponses (USB Response Pad, RB-730). À l'issue du test, nous avons demandé aux auditeurs de remplir un questionnaire visant à préciser les critères qui leur avaient permis d'évaluer la fluence des productions des apprenants. Les questionnaires sont présentés en annexe (un questionnaire vierge est présenté en Annexe G et les questionnaires complétés sont présentés dans le DVD joint).

Les résultats des analyses acoustiques et perceptives de ces deux études seront présentés dans le chapitre 5.

4.2.5 Présentation du corpus MULTIPHONIA

On utilise des corpus dans l'apprentissage des langues depuis les années 90 mais la majorité de ces derniers portent sur l'écrit et non pas sur l'oral (Mauranen, 2004). Plus récemment, quelques travaux se sont intéressés à l'évaluation de l'interlangue d'apprenants de LE à l'oral en anglais langue étrangère, comme par exemple la base de données ENGLISH (Tortel, 2009), mais il n'existe pas, à notre connaissance, de corpus de ce type en FLE.

Nous pensons que la problématique de l'enseignement oral que nous exposons au chapitre 1, peut s'expliquer par le manque de données empiriques sur l'enseignement de l'oral en LE, et sur l'enseignement de la prononciation en particulier. En dépit du fait que nous savons aujourd'hui que la prosodie et la multimodalité sont non seulement les clefs de l'acquisition du langage mais qu'elles sont aussi nécessaires et irrépensables dans la communication de tous les jours (Di Cristo, 2004 ; Kendon, 2004 ; Mac Neill, 2005). De même, des études en LE ont montré que les gestes pédagogiques avaient un impact sur la recollection des items lexicaux en LE (Allen, 1995 ; Tellier, 2008). Pourtant, ces aspects sont négligés par la didactique.

Nous pensons que la didactique ne peut plus faire l'impasse sur cette réalité. Si on veut aujourd'hui délimiter et normaliser l'oral en LE, il est nécessaire de travailler à partir de corpus oraux d'interactions en classe de langue. Il est indispensable de s'intéresser à l'interlangue, qui constitue le cœur de l'évaluation et de la construction de la LE de l'apprenant.

Nous avons donc constitué un corpus d'interlangue d'apprenants⁸ anglophones de FLE, corpus MULTIPHONIA. Ce corpus a pour objectif de fournir une base de données multimodale de la construction phonétique de l'apprenant dans une situation écologique et interactive de classes en LE. Ce corpus permet de confronter des prédictions théoriques avec des situations de classes réelles, dans le but de favoriser l'enseignement de la phonétique en classe de langue étrangère (*cf.* partie discussion et perspectives).

⁸ Tous les participants ont signé une autorisation de droit à l'image. Les autorisations sont présentées dans le DVD joint.

MULTIPHONIA contient 96 heures d'enregistrements vidéo de cours de correction phonétique selon deux méthodes (la MVT vs. la méthode articulatoire) et de cours de communication orale. Les séances ont été enregistrées dans les mêmes conditions expérimentales, dans le studio vidéo professionnel de la Direction Technologique de l'Information et de la Communication pour l'Enseignement (DTICE)⁹ de l'université Toulouse II.

Afin d'être au plus près d'une situation de déroulement de cours classiques, une mise en scène pour chaque méthode a été choisie. Ces différentes mises en place impliquaient donc une organisation technique spécifique du plateau utilisé comme salle de cours.

Pour les séances de MVT, les participants¹⁰ étaient assis en cercle autour de l'enseignant (cf. figure 26).



Figure 26 Extrait d'un cours de MVT avec le groupe débutant

Comme on peut le voir sur la figure précédente, nous avons recourt à la prosodie et la multimodalité (geste de la main mimant la courbe intonative) pour aider les apprenants à mieux percevoir les sons de la langue cible. Les participants étaient libres de reproduire ou non les mouvements de main pendant qu'ils répétaient les phrases ou les courts-dialogues.

⁹ Nous tenons à remercier toute l'équipe de la DTICE et en particulier Monsieur Bruno Bastard qui a filmé l'ensemble du corpus et assuré la post-production des vidéos.

Pour les cours d'articulatoire, les participants étaient assis autour d'une table pendant que nous expliquions les gestes articulatoires à l'aide de mimes ou de schémas (cf. figure 27)



Figure 27 Extrait d'un cours d'articulatoire niveau débutant

Les détails techniques de la constitution de ce corpus sont présentés dans Alazard *et al* (2012 b) (article présenté en Annexe H).

MULTIPHONIA a été déposé dans la base de données SLDR du laboratoire Parole et Langage (<http://sldr.org/sldr000780/fr.>) et sera bientôt enrichi à de nombreux niveaux afin de permettre des analyses autant segmentales que prosodiques, morphosyntaxiques, syntaxiques, lexicales ou gestuelles en LE (cf. partie conclusion et perspectives). Les différents niveaux d'annotations automatiques seront faits dans un premier temps sur de courts extraits, puis confronté à des annotations manuelles par 3 experts avant d'étendre la procédure à de plus grandes parties du corpus. Des extraits bruts du corpus sont disponibles en annexes dans le DVD joint.

4.3 SYNTHÈSE DU CHAPITRE 4

Dans ce chapitre nous avons tout d'abord développé notre problématique et nos hypothèses de recherches. Nous postulons qu'il existe des transferts de compétences entre l'encodage de l'oral et de l'écrit en langue étrangère, liés à la maîtrise de la prosodie. Nous sommes parti du postulat que la fluence était étroitement liée à la mise en place d'une compétence prosodique en LE. Pour tester cela, nous avons formulé une première hypothèse selon laquelle un apprentissage à l'oral centré sur la prosodie de la langue étrangère va améliorer la fluence des apprenants en parole lue (transferts de compétences de l'oral vers l'écrit) (Hypothèse 1). Sachant que la prosodie est profondément ancrée dans notre comportement verbal et qu'elle est à la base des processus d'encodage et de décodage du langage (Di Cristo, 2004), nous postulons qu'elle se trouve au cœur de la pratique linguistique et de l'apprentissage d'une langue étrangère. Nous pensons que les résultats seront meilleurs si l'apprentissage de la prosodie est proposé dès le niveau débutant car les apprenants de niveau avancé auront fossilisé leurs erreurs (Hypothèse 2)

Afin d'approfondir ces deux premières hypothèses, nous avons souhaité confirmer le lien entre compétence prosodique et fluence, en comparant l'incidence sur la lecture oralisée de deux approches de l'enseignement de la prononciation : une approche non intellectualisée centrée sur la prosodie, avec la Méthode Verbo-Tonale, et une approche métalinguistique de l'enseignement du segmental avec la Méthode Articulaire. Nous postulons que seule la MVT aura une incidence positive sur l'amélioration de la fluence en lecture oralisée des apprenants de LE. (Hypothèse 3). Enfin, nous pensons que l'oral et la prononciation devraient être enseignés en priorité tandis que l'écrit ne devrait être abordé qu'après plusieurs heures consacrées à l'enseignement de l'oral. Nous émettons donc l'hypothèse selon laquelle l'introduction précoce de l'écrit (i.e. avant que la compétence rythmique soit installée), qui induit obligatoirement l'intellectualisation et l'analyse de la langue, entraîne une chute de la fluence verbale (Hypothèse 4). Pour vérifier la validité de ces hypothèses nous avons

mené deux études longitudinales sur une durée de huit semaines chacune, avec des apprenants anglophones de FLE.

Dans la suite de ce chapitre, nous avons donc présenté le cadre méthodologique de ces deux études. Nous avons appuyé l'intérêt de mener des études longitudinales pour évaluer la construction de l'interlangue. Nous avons présenté les apprenants (nationalités et niveaux) et nous avons mis en avant les difficultés que nous avons rencontrées pour constituer des groupes d'apprenants homogènes. Nous avons également soulevé le problème de l'évaluation des niveaux selon le CECR. Nous avons décrit le dessin expérimental de nos deux études puis justifié les deux types d'analyses choisis pour évaluer la fluence de l'interlangue : une analyse quantitative de la *fluence de production* et une analyse qualitative de la *fluence perçue* par des natifs experts. Les résultats de ces analyses seront présentés dans le chapitre 5.

Enfin, nous avons présenté le corpus MULTIPHONIA que nous avons constitué. Ce corpus de 96 heures d'enregistrements vidéos de cours de prononciation et de communication a été conçu dans le but est de fournir une base de données multimodale de la construction de l'interlangue de l'apprenant dans une situation écologique et interactive de classes en LE. Ce corpus sera bientôt enrichi et mis à la disposition des enseignants et des chercheurs de LE.

2^{ÈME} PARTIE : ANALYSES EXPÉRIMENTALES

Chapitre 5 : Résultats étude longitudinale pilote et étude longitudinale 2

Ce chapitre est dédié à l'analyse expérimentale des variables susceptibles de représenter la fluence des productions de parole lue par des apprenants anglophones de FLE. Nous pensons en que la fluence et la prosodie sont étroitement corrélées et que la fluence de production reflète la mise en place de la compétence rythmique des apprenants. La fluence permet donc d'évaluer les transferts prosodiques entre l'oral et l'écrit en langue étrangère.

En fonction de la littérature sur la fluence (chapitre 3), nous avons distingué deux types de fluence dans nos analyses, la *fluence de production*, mesurables par des données acoustiques, et la *fluence perçue*, qui rend compte de l'appréciation globale de l'authenticité des productions par des natifs. Nous pensons que ces deux types de fluence sont complémentaires car ils induisent deux types d'analyse : une analyse quantitative et une analyse qualitative des données.

Comme nous l'avons vu au chapitre précédent, les critères retenus pour évaluer la *fluence de production* sont : le nombre et la durée des différents types de pauses et de disfluences, le nombre, la distribution et la durée des syllabes accentuées, la durée des syllabes inaccentuées et la densité des groupes rythmiques. Pour évaluer la *fluence perçue*, nous avons utilisé des échelles de magnitude représentant l'authenticité des productions des apprenants. Pour la seconde étude, nous avons également demandé aux auditeurs de préciser les critères qui avaient influencé leurs choix.

En accord avec nos hypothèses (présentées au chapitre 4) nous pensons que seuls les apprenants qui auront suivi des cours de correction phonétique *via* la Méthode Verbo-Tonale amélioreront la fluence de leurs productions de parole lue car ils auront développé une compétence prosodique en LE. Nous nous attendons à ce que les apprenants des groupes tests (MVT) produisent donc des schémas prosodiques plus proches de ceux du français : groupes rythmiques plus longs, plus de contraste entre les syllabes accentuées et les syllabes inaccentuées, moins de pauses agrammaticales et augmentation du débit de parole. Nous pensons que les résultats seront meilleurs chez les

apprenants débutants car ils n'ont pas encore fossilisé leurs erreurs. Nous nous attendons donc à observer une progression similaire pour les apprenants des groupes test (MVT) dans les deux études. Dans la deuxième étude, nous avons également formulé l'hypothèse selon laquelle l'introduction précoce de l'écrit viendrait perturber la mise en place des transferts prosodiques entre l'oral et l'écrit. Nous prévoyons que la fluence des apprenants de l'étude 2 chutera après l'introduction de l'écrit, ce qui se traduirait notamment par un débit ralenti et une ré augmentation du nombre de pauses agrammaticales.

En ce qui concerne l'analyse qualitative, nous pensons que les résultats acoustiques et les résultats perceptifs seront corrélés. Nous savons en effet que les aspects rythmiques jouent un rôle essentiel dans l'évaluation de la fluence perçue par les natifs et que les éléments prosodiques sont déterminant dans la perception d'un accent étranger (*cf.* Jilka, 2000 ; Birdsong, 2003 ; Vieru & Boula de Mareüil, 2006 ; Kaglik & Boula de Mareüil, 2009, 2010).

Dans ce chapitre, nous présenterons et discuterons les résultats des analyses quantitatives de la *fluence de production* et des analyses qualitatives de la *fluence perçue* pour chacune des deux études longitudinales réalisées.

5.1 RESULTATS ETUDE 1¹¹

Cette étude est une étude pilote réalisée avec quatre sujets (2 débutants et 2 avancés) dont le but est de tester l'existence de transferts entre l'oral et l'écrit en langue étrangère (Hypothèse1) et l'impact du niveau des apprenants sur l'apprentissage de la prosodie (Hypothèse 2). Les sujets de cette étude ont été équitablement répartis en deux groupes : un *groupe test* qui a suivi des cours de correction phonétique via la MVT et un *groupe contrôle* qui a suivi des cours de FLE standards dans une approche communicative. Les apprenants ont été testés dans une tâche de lecture oralisée avant (t0) et après l'entraînement (t1). Les résultats de l'analyse acoustique et de l'analyse perceptive sont présentés ci-dessous.

5.1.1 Résultats de l'analyse acoustique

Tous les paramètres prosodiques initialement retenus ne se sont pas révélés pertinents pour évaluer la fluence de l'interlangue. En accord avec la littérature (cf. chapitre 3), nous n'avons observé aucun changement dans le nombre ou la durée des différents types de disfluences (faux-départ, hésitations, répétitions) entre t0 et t1. Inversement, l'entraînement a eu une incidence sur les pauses, les syllabes inaccentuées et les syllabes accentuées, appuyant le lien entre prosodie et fluence.

Afin de rendre compte de la significativité des différences entre le nombre et les durées des différents types de pauses et des différents types de syllabes (inaccentuées et accentuées) entre t0 et t1, nous avons réalisé des Analyses de variance simples (ANOVA) par types de syllabes et par types de pauses. La Durée (log ms) et l'Occurrence des événements constituaient les variables dépendantes, et l'Etat (Avant/Après entraînement de 8 semaines) la variable indépendante (facteur à deux niveaux). Nous précisons qu'en raison du faible nombre de sujets, les analyses statistiques présentées sont indicatives et ne permettent pas faire des généralisations sur les résultats obtenus.

¹¹ Nous tenons à remercier Monsieur Robert Espesser pour sa collaboration et ses conseils pour les analyses statistiques.

Le tableau 8 présente les paramètres qui révèlent une tendance, significative ou non, de l'effet d'entraînement, pour au moins un locuteur. Ce tableau représente donc les données pour l'effectif et la durée (en ms) des pauses vides (#), des pauses vides agrammaticales (#AG), des pauses respiratoires (#*), des syllabes inaccentuées (INA) et des accents finaux majeurs (Fmaj) et mineurs (Fmin), avant (t0) et après entraînement (t1). Les accents initiaux, qui n'ont subi aucun changement avant et après entraînement pour aucun des locuteurs ne sont pas répertoriés dans ce tableau.

	GROUPE TEST				GROUPE CONTROLE			
	Niveau A (MM)		Niveau B (DR)		Niveau A (SR)		Niveau B (AM)	
	Nbre	Durée (ms)	Nbre	Durée (ms)	Nbre	Durée (ms)	Nbre	Durée (ms)
# AG	t0=43 t1=40	t0=787 t1=401	t0=4 t1=3	t0=256 t1=195	t0=32 t1=34	t0=359 t1=385	t0=18 t1=15	t1=292 t0=363
#*	t0=31 t1=33	t0=901 t1=486***	t0=24 t1=28	t0=743 t1=661	t0=38 t1=39	t0=614 t1=634	t0=30 t1=26	t0=548 t1=572
#	t0=33 t1=24	t0=346 t1=250*	t0=23 t1=21	t0=347 t1=320	t0=8 t1=15	t0=367 t1=388	t0=12 t1=8	t0=176 t1=143
INA	t0=279 t1=275	t0=225 t1=202*	t0=224 t1=245	t0=156 t1=155	t0=225 t1=226	t0=172 t1=168	t0=226 t1=219	t0=201 t1=200
Fmaj	t0=61 t1=62	t0=321 t1=299	t0=57 t1=62	t0=280 t1=294	t0=58 t1=63	t0=350 t1=364	t0=67 t1=60	t0=341 t1=327
Fmin	t0=78 t1=65	t0=355 t1=304*	t0=74 t1=64	t0=238 t1=218	t0=72 t1=58	t0=313 t1=316	t0=61 t1=73	t0=337 t1=321

Tableau 8 : Nombre et Durée (log ms) des paramètres prosodiques pertinents par locuteur, avant/après entraînement (t0 = avant entraînement ; t1 = après entraînement). Les éléments significatifs sont en gras et marqués par *

Les analyses révèlent que la variable *Occurrence* des types de pauses et des types d'accents ne permet pas de distinguer significativement l'état avant/après entraînement, et ce pour les deux groupes de locuteurs (*groupe test* et *groupe contrôle*) ($p > 0.05$).

En revanche, la *Durée* des syllabes accentuées et inaccentuées, et plus particulièrement la *Durée* des types de pauses, nous permettent de comparer la

production des apprenants selon la méthode d'enseignement utilisée et selon leur niveau (A vs. B), avant et après entraînement.

5.1.1.1 Mesures des pauses

Nous présenterons les graphes illustrant l'évolution de la durée des différents types de pauses par niveau plutôt que par groupe (*groupe test* ou *groupe contrôle*). En effet, quelle que soit la méthode d'enseignement, les apprenants de niveau avancé (niveau B) produisent peu d'hésitations vocales et de pauses silencieuses agrammaticales, contrairement aux apprenants de niveau débutant qui en produisent énormément, aussi bien avant qu'après entraînement, comme nous pouvons le voir sur les figures 28 et 29 pour les résultats des apprenants de niveau débutant et les figures 30 et 31 pour les résultats des apprenants de niveau avancé.

Sur les figures 28 et 29, nous pouvons observer une nette différence entre la progression de l'apprenant débutant du *groupe test* (figure 28) et celle de l'apprenant débutant du *groupe contrôle* (figure 29). En effet, pour de dernier, la durée des différents types de pauses reste globalement inchangée entre le premier et le second test alors que les résultats de l'apprenant débutant du *groupe test* mettent en évidence non seulement la réduction de la durée des pauses vides et des pauses respiratoires, mais également une homogénéité plus importante de la durée globale de l'ensemble des pauses après entraînement.

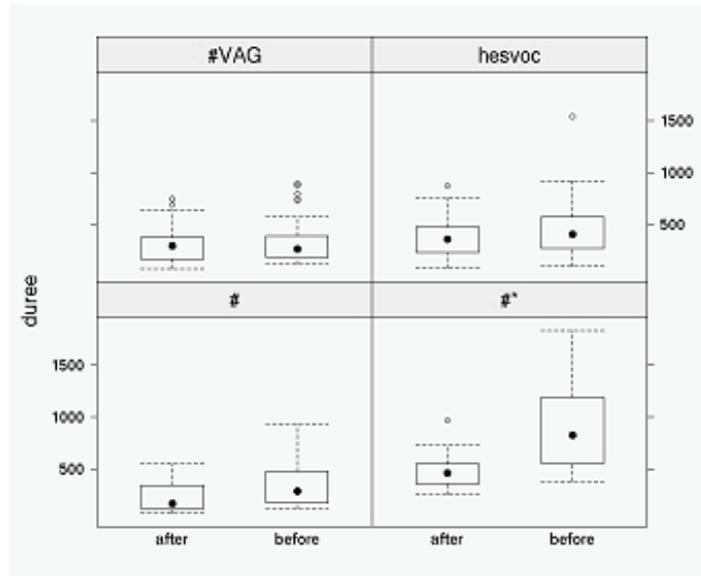


Figure 28 Durée (log ms) des pauses et hésitations vocales avant (before)/après (after) entraînement pour l'apprenant débutant du groupe test (MM) (# VAG = pauses silencieuses agrammaticales; hesvoc = hésitations vocales ; # = pauses vides; #* = pauses respiratoires)

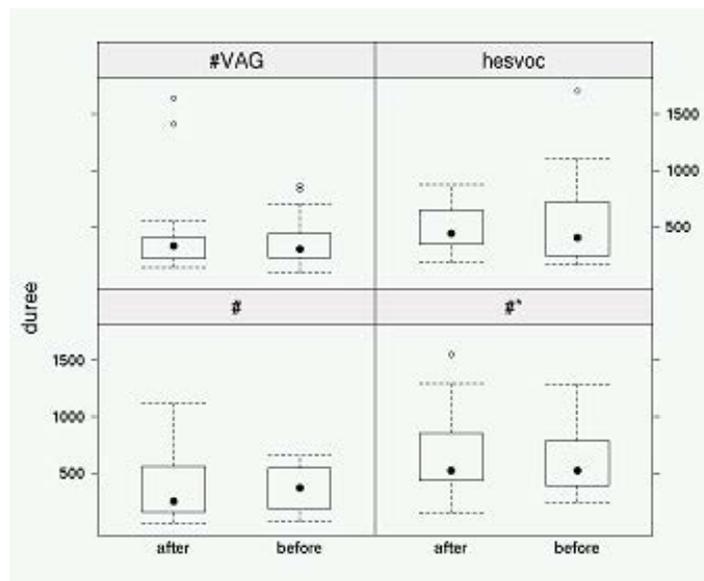


Figure 29 Durée (log ms) des pauses et hésitations vocales avant (before)/après (after) entraînement pour l'apprenant débutant du groupe contrôle (SR) (# VAG = pauses silencieuses agrammaticales; hesvoc = hésitations vocales ; # = pauses vides; #* = pauses respiratoires)

En ce qui concerne les apprenants de niveau avancé, nous observons une évolution similaire : la durée des pauses respiratoires est plus courte pour l'apprenant du *groupe test* après entraînement (figure 30), alors qu'il n'y a pas de modification de la durée des pauses (respiratoires ou vides) après entraînement pour l'apprenant du *groupe contrôle* (figure 31).

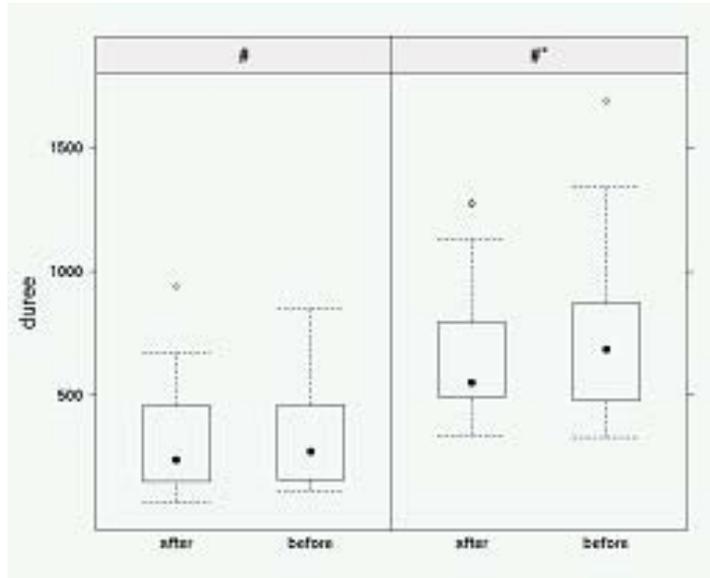


Figure 30 Durée (log ms) des pauses avant (before)/après (after) entraînement pour l'apprenant avancé du groupe test (DR) (# = pauses vides ; #* = pauses respiratoires)

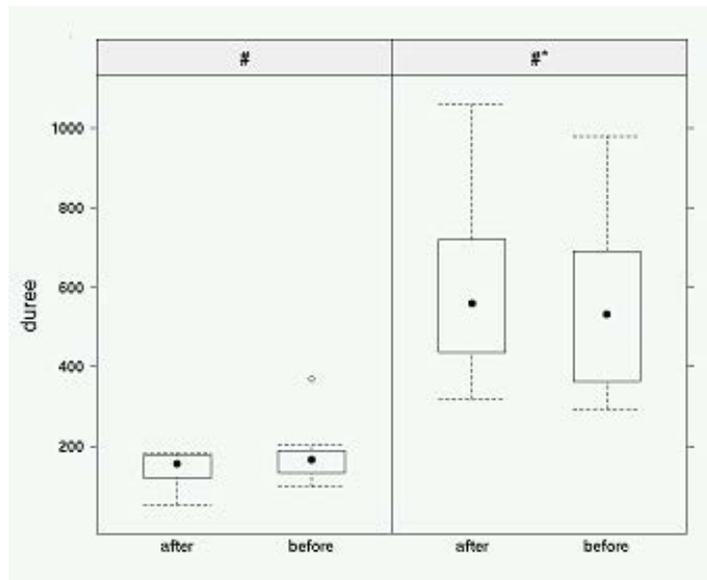


Figure 31 Durée (log ms) des pauses avant (before)/après (after) entraînement pour l'apprenant avancé du groupe contrôle (AM) (# = pauses vides ; #* = pauses respiratoires)

5.1.1.2 Mesures des syllabes

La Durée des types de syllabes (inaccentuées et types d'accents) et des types de pauses (incluant les pauses agrammaticales, les pauses respiratoires et vides) est globalement raccourcie pour les locuteurs du *groupe test* après l'entraînement, contrairement aux locuteurs du groupe *contrôle*. La variable Durée n'est néanmoins significativement raccourcie que pour un locuteur du *groupe test* (MM), et pour un sous-ensemble des phénomènes prosodiques susceptibles d'avoir un impact sur la fluence de la lecture oralisée.

La différence de durée des pauses respiratoires grammaticales est très significative et elle a été divisée par deux entre le premier et le second test [($F(1; 62) = 34.04; p < .001$)]. La durée des pauses silencieuses est elle aussi significativement raccourcie ($F(1; 55) = 5.17; p = 0.03$). Enfin, la durée des accents finaux mineurs et des syllabes inaccentuées a significativement diminué entre le premier et le second test ($F(1; 145) = 7.80; p = 0.006$, et $F(1; 500) = 13.65; p < .001$, respectivement). Cette diminution indique que le débit de parole de l'apprenant débutant du *groupe test* est plus rapide après entraînement.

Autrement dit, les résultats de l'apprenant débutant du *groupe test* montrent qu'il a développé des stratégies de lecture différentes entre le premier et le second test - soit à seulement huit semaines d'intervalles - en modifiant la durée des pauses respiratoires et des pauses vides grammaticales, ainsi que la durée des syllabes inaccentuées et des accents finaux situés à des frontières prosodiques mineures. Cet apprenant a donc développé des stratégies de d'encodage et de décodage différentes entre t0 et t1. Nous avons établi un lien entre la fluence et la compétence prosodique, nous pensons donc que MM a mis en place une compétence prosodique adaptée à la LE entre t0 et t1.

Nous pensons que ces changements seront également perçus par les enseignants de FLE, c'est pourquoi nous avons comparé les résultats de l'analyse acoustique à ceux de l'analyse perceptive.

5.1.2 Résultats de l'analyse perceptive

Comme nous pouvons le voir sur la figure 32, les jugements de fluence par les douze auditeurs natifs se concentrent autour du milieu de l'échelle de magnitude (scores moyens = 4), pour les deux groupes, aussi bien avant qu'après entraînement.

Toutefois, nous remarquons que la méthode choisie (MVT vs. Approches Communicatives) a eu une incidence sur l'amélioration de la fluence en lecture oralisée des apprenants. Le coefficient d'interaction révèle une amélioration très significative pour la MVT (beta = 0.336, $t = 3.52$, $p < .001$).

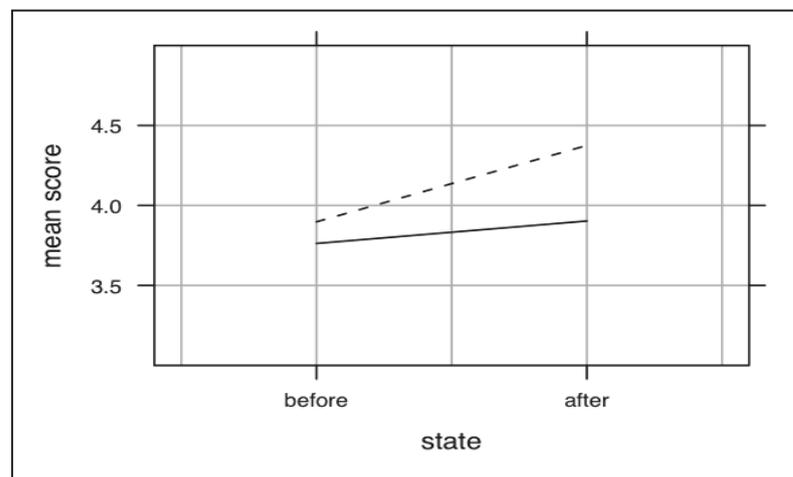


Figure 32 Moyenne des scores avant/après entraînement par groupe (ligne pointillée Groupe Test; ligne continue Groupe Contrôle)

Les scores de jugement des auditeurs font état d'une progression 0.336 supérieure pour les apprenants du groupe test par rapport aux apprenants du groupe contrôle, après les huit semaines d'entraînement. Sur la figure précédente, on note une légère différence de niveau entre les deux groupes. Les apprenants du *groupe test* obtiennent un score supérieur aux apprenants du *groupe contrôle* au temps t0. Même si la différence de niveau n'est pas significative ($p > 0.05$), cette observation renforce la critique que nous émettions précédemment (chapitre 4) par rapport à l'évaluation du niveau des apprenants selon les grilles du CECR.

Afin d'estimer l'influence de la méthode utilisée sur les jugements de la fluence des apprenants de FLE par les auditeurs natifs, nous avons utilisé un

modèle linéaire mixte avec un intercept aléatoire pour prendre en compte la variabilité inter-juges, et un intercept aléatoire pour prendre en compte la variabilité des réponses à travers les 11 phrases. Les réponses (de 1 à 7) constituaient la variable dépendante tandis que la méthode d'enseignement (*groupe contrôle vs. groupe test*) et l'Etat (Avant/Après entraînement de 8 semaines) constituaient les variables indépendantes (facteurs à deux niveaux).

Pour apprécier l'incidence du niveau (A vs. B) des apprenants sur les progrès réalisés en fonction des méthodes, nous avons réalisé des t-tests appariés pour chaque locuteur à partir de la moyenne des réponses de chaque auditeur, en fonction de l'Etat (Avant/Après entraînement).

Ces données sont résumées dans le tableau suivant qui illustre les scores obtenus pour chaque locuteur, avant et après entraînement, ainsi que les changements significatifs.

Groupes/ Méthodes	Niveau	Etat	
		t0	t1
Groupe Test (MVT)	Niveau A (MM)	2,46	3,08**
	Niveau B (DR)	5,22	5,67*
Groupe Contrôle (AC)	Niveau A (SR)	3,33	3,47
	Niveau B (AM)	4,20	4,34

Tableau 9 Moyenne des scores avant (t0) et après entraînement (t1) pour chaque locuteur. La significativité est indiquée par une ou plusieurs astérisques *.

Comme nous pouvons le voir dans le tableau 3, les analyses statistiques ne montrent pas de différence dans les scores de jugement de la fluence avant et après entraînement pour aucun des apprenants du *groupe contrôle*. En effet, la progression des apprenants du groupe contrôle (SR et AM) est non significative (progression de + 0.14 entre le premier et le second test pour

chaque locuteur (t -test = 0.5863; t = 0.5606; df = 11, pour l'apprenant de niveau A, SR; et t -test = 0.08; t = 1.9149; df = 11 pour l'apprenant de niveau B, AM)).

En revanche, les deux apprenants du *groupe test* ont amélioré la fluence de leurs productions. Nous pouvons observer une différence significative de la progression de l'apprenant de niveau débutant (+0.62) et de l'apprenant du niveau avancé (+0.45) après les classes de MVT (t -test = 0.0069; t = 3.3166; df = 11 pour l'apprenant de niveau A, MM; et t -test = 0.0116; t = 3.0225; df = 11, pour l'apprenant de niveau B, DR).

Les résultats de l'analyse perceptive corroborent les résultats de l'analyse acoustique : les apprenants du groupe test ont significativement amélioré la fluence (*fluence de production* et *fluence perçue*) de leur production après entraînement.

5.1.3 Discussion

En faisant le parallèle avec de récents travaux de Janet Fodor (2002) sur la prosodie implicite – il existerait des schémas prosodiques mémorisés par la pratique orale qui seraient ensuite projetés sur le texte et traités comme faisant partie de l'input par le lecteur – nous avons postulé l'existence de transferts de compétence entre l'oral et l'écrit en LE. Nous pensons que ces transferts sont dépendants de l'apprentissage d'une compétence prosodique liée à la maîtrise des processus d'encodage et de décodage de la LE. Dans les chapitres 3 et 4, nous avons appuyé l'idée que l'évaluation de la fluence des productions d'apprenant de LE, permettrait de rendre compte de la mise en place de la compétence rythmique en LE.

Les mesures quantitatives et qualitatives de la fluence des productions de parole lue soutiennent ces hypothèses. En effet, les mesures acoustiques indiquent une tendance nette de raccourcissement des pauses vides et respiratoires grammaticales et non grammaticales, ainsi qu'une tendance au raccourcissement de la durée des accents et des syllabes inaccentuées pour les locuteurs du *groupe test* uniquement.

Ces résultats montrent que seuls les locuteurs du *groupe test* parviennent à une meilleure fluence en lecture oralisée après l'entraînement à la correction phonétique par la MVT. La réduction de la durée des syllabes inaccentuées indique également un débit de parole plus rapide et donc plus fluide. De plus, nous savons que les phénomènes de pauses et d'allongement final reflètent les processus de planification, indiquant que les modifications observées dans la fluence de production pourraient révéler des modifications plus profondes de fluence cognitive (Hypothèse 1 validée)

Nous noterons également que les progrès réalisés ne sont statistiquement significatifs que pour le locuteur de niveau A (MM), ce qui renforce notre deuxième hypothèse, selon laquelle la correction phonétique est plus efficace à un stade précoce de l'apprentissage de la LE (Hypothèse 2 validée)

En contre partie, nos résultats acoustiques indiquent enfin que la durée des phénomènes de pause et des accents n'a pas varié pour les locuteurs du *groupe contrôle* avant et après l'entraînement. La fluence en lecture oralisée des deux apprenants du groupe contrôle ne s'est donc pas améliorée après les huit semaines d'enseignements.

Ces premières mesures quantitatives ont été corroborées par les résultats du test perceptif. En effet, les douze experts ont perçu une progression significative pour les apprenants du *groupe test* uniquement, en particulier pour l'apprenant de niveau débutant (validation des hypothèses 1&2). Il serait intéressant de connaître les critères qui ont permis aux experts d'évaluer la fluence des productions, et c'est pourquoi nous avons proposé aux auditeurs de l'étude 2 de remplir un questionnaire à l'issue du test. En nous basant sur de récentes études sur le lien entre authenticité et prosodie (cf. Jilka, 2000 ; Birdsong, 2003 ; Vieru & Boula de Mareüil, 2006 ; Kaglik & Boula de Mareüil, 2009, 2010), il est possible d'imaginer que les sujets ont été sensibles aux paramètres rythmiques que nous avons mesurés dans nos analyses acoustiques. Nous pensons en effet, que si l'on considère la fluence comme une

mesure de la compétence rythmique, *fluence de production* et *fluence perçue* sont étroitement corrélées.

En conclusion, les résultats de cette étude pilote confortent nos hypothèses de recherche car ils montrent une amélioration significative de la *fluence de production* et de la *fluence perçue* en lecture oralisée pour les apprenants du *groupe test* uniquement, et plus particulièrement pour l'apprenant débutant du *groupe test*. Ces résultats suggèrent que les transferts des caractéristiques prosodiques de la langue maternelle (anglais) sur la langue cible (français), comparables à travers les sujets lors du premier test, sont moins importants pour les apprenants du *groupe test* après entraînement à la correction phonétique : les caractéristiques prosodiques du français sont mieux maîtrisées en particulier pour l'apprenant de niveau débutant. Ces transferts ne seraient donc pas soumis à l'effet d'une période critique vis à vis de l'âge des apprenants (cf. chapitre 1) mais sembleraient liés au niveau d'apprentissage. Il serait ainsi plus facile de mettre en place ces transferts prosodiques au niveau débutant qu'au niveau avancé (*effet de fossilisation*), ce qui va à l'encontre des hypothèses de Kaglik & Boula de Mareuil (2009, 2010) qui supposent que la prosodie - représentant un coût cognitif très important - requiert un niveau élevé de compétence en LE pour être maîtrisée.

D'autre part, un travail centré sur la prosodie à l'oral aiderait l'apprenant à acquérir une meilleure fluence en lecture oralisée car il lui permettrait de se débarrasser de stratégies erronées d'accès au texte, et donc d'acquérir certains automatismes en langue-cible qui facilitent l'appréhension globale du texte (mise en place d'une bonne fluence cognitive // compétence prosodique cf. chapitre 3). Le décodage du code écrit devrait s'en trouver facilité pour les apprenants qui ont suivi un entraînement poussé de correction phonétique par la méthode verbo-tonale, par rapport à ceux qui ont suivi une formation plus 'traditionnelle' mêlant activités orales et activités écrites dans une approche communicative de la LE.

Cette étude pilote offre des résultats très encourageants et nous permet de valider nos choix méthodologiques. La méthode de correction phonétique

verbo-tonale semblerait en effet faciliter la mise en place d'une fluence cognitive en langue étrangère, liée à l'automatisation des processus d'encodage et de décodage.

Les résultats de cette étude pilote soulèvent néanmoins plusieurs questions méthodologiques et épistémologiques.

Premièrement, nous avons conscience que le nombre de sujets (la variable niveau étant confondue avec la variable locuteur) ne permet pas de généraliser les résultats obtenus. Les progrès observés dans cette étude pourraient-ils être dus à la mise en place de stratégies individuelles ? Pour répondre à cette question, nous avons reconduit cette étude pilote avec vingt sujets.

Deuxièmement, les progrès observés chez les apprenants du *groupe test* sont-ils directement imputables à la méthode utilisée ? On pourrait imaginer que d'autres méthodes de correction phonétique donneraient des résultats semblables. Le dessin expérimental de cette étude pilote ne permet pas de valider expérimentalement la MVT par rapport à une autre méthode.

Enfin, la fluence en lecture oralisée est-elle réellement associée à une compétence rythmique ? Ne s'agit-il pas simplement d'une compétence orale globale (phonèmes, rythme et intonation) ? Cette question loin d'être anodine soulève également la question du type de connaissances à enseigner. La mise en place de transferts prosodiques nécessite-t-elle le développement de connaissances procédurales ? Si on considère que la fluence est liée à l'automatisation des processus d'encodage et de décodage (chapitre 3), est-il possible d'acquérir une bonne fluence par le biais de connaissances déclaratives ?

Les résultats de la deuxième étude longitudinale vont nous permettre d'apporter des éléments de réponses à ces différentes questions.

5.2 RESULTATS ETUDE 2¹²

Dans la deuxième étude, nous avons testé nos quatre hypothèses de recherches avec dix-huit¹³ apprenants anglophones de FLE (8 débutants et 10 avancés) sur une durée de 8 semaines. Nous postulons qu'il existe des transferts prosodiques entre l'oral et l'écrit (Hypothèse1) et que les apprenants de niveau débutant auront plus de facilités à maîtriser la prosodie de la LE (Hypothèse 2). Nous pensons que ces transferts sont liés à la mise en place d'une compétence prosodique en LE et que seule la MVT, qui s'appuie sur un apprentissage sans intellectualisation de la prosodie, permet l'apprentissage de cette compétence (Hypothèse 3). Enfin, nous pensons que l'introduction précoce de l'écrit va venir court-circuiter la mise en place de ces transferts et entraîner une chute de la fluence des productions des apprenants (Hypothèse 4).

Les sujets ont été équitablement répartis en quatre groupes: deux *groupes tests* ont suivi des séances de correction phonétique par la MVT et deux groupes contrôles ont suivi des séances de correction phonétique selon la méthode articulatoire (2 groupes de niveau * 2 méthodes). Afin de tester la quatrième hypothèse, l'écrit n'a été introduit qu'en cours de formation. Les apprenants ont donc été testés dans une tâche de lecture oralisée avant (s0), pendant (i.e. avant introduction de l'écrit, s1) et après l'entraînement (s2). Les résultats de l'analyse acoustique et de l'analyse perceptive sont présentés ci-dessous.

5.2.1 Résultats de l'analyse acoustique

Cette nouvelle étude réalisées sur davantage de sujets vient confirmer les tendances de l'étude pilote. Les effectifs des différents types de disfluences (hésitations, hésitations lexicales, faux départ, répétitions et reparans) ne se sont pas révélés pertinents pour l'analyse de l'effet de l'entraînement. Comme on peut le voir sur l'histogramme ci-dessous (figure 33) : les hésitations (hes),

¹² Nous tenons à remercier Monsieur Robert Espesser pour sa collaboration et ses conseils pour les analyses statistiques.

¹³ Comme nous l'avons expliqué au chapitre précédent (chapitre 4), deux apprenants de niveau débutant ont abandonné le protocole au bout de quelques semaines. C'est pourquoi nous présentons les analyses des productions pour 18 sujets seulement.

les hésitations lexicales (hes-lex), les faux-départs (FD), les répétitions (RE) et les reparans (RA) ont peu d'occurrences. Les pauses agrammaticales non respiratoires et les pauses grammaticales (vides et respiratoires) sont les paramètres qui présentent le plus grand effectif pour tous les groupes. Nous noterons le plus grand nombre de pauses est observé dans les productions des apprenants avancés, quelque soit la méthode utilisée, ce qui s'explique par la longueur plus importante des textes utilisés à ce niveau (textes en moyenne 2 à 3 fois plus longs que pour les débutants).

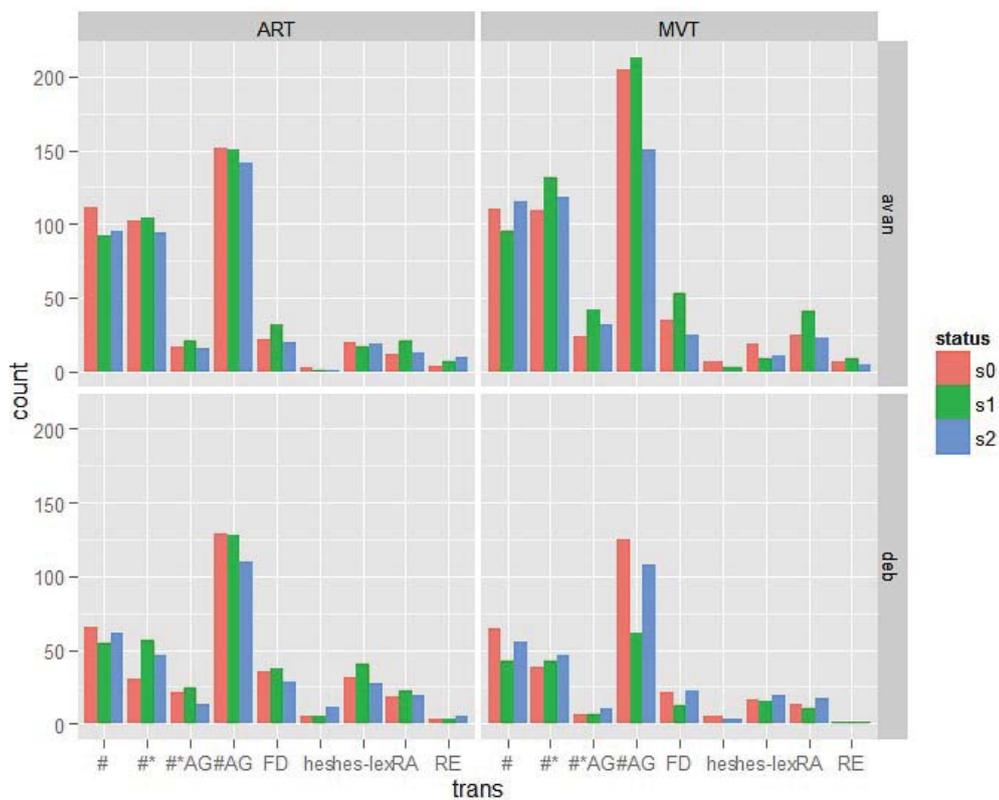


Figure 33 Histogrammes de l'effectifs des différents types de pauses (# : pauses vides ; #* : pauses respiratoires ; #AG : pauses agrammaticales ; #*AG : pauses respiratoires agrammaticales) et disfluences (FD : faux-départs ; Hes : hésitations ou pauses sonores ; Hes-lex : hésitations lexicales ; RE : répétitions ; RA : reparans) par groupe (MVT ou ART) et par niveau (avancé ou débutant) en fonction des trois stades (en rouge s0 : avant, en vert s1 : pendant et en bleu s2 : après entraînement)

Nous avons donc recentré nos analyses acoustiques sur les mesures de nombre et durée par types de pauses en regroupant les pauses vides et respiratoires grammaticales sous l'appellation pauses standards.

5.2.1.1 Mesures des pauses

Nous avons mesuré la densité des pauses agrammaticales (pag) et des pauses standards (pstand) selon la méthode, le niveau et la phase d'apprentissage. Ces mesures sont représentées sur l'histogramme lissé (figure 34).

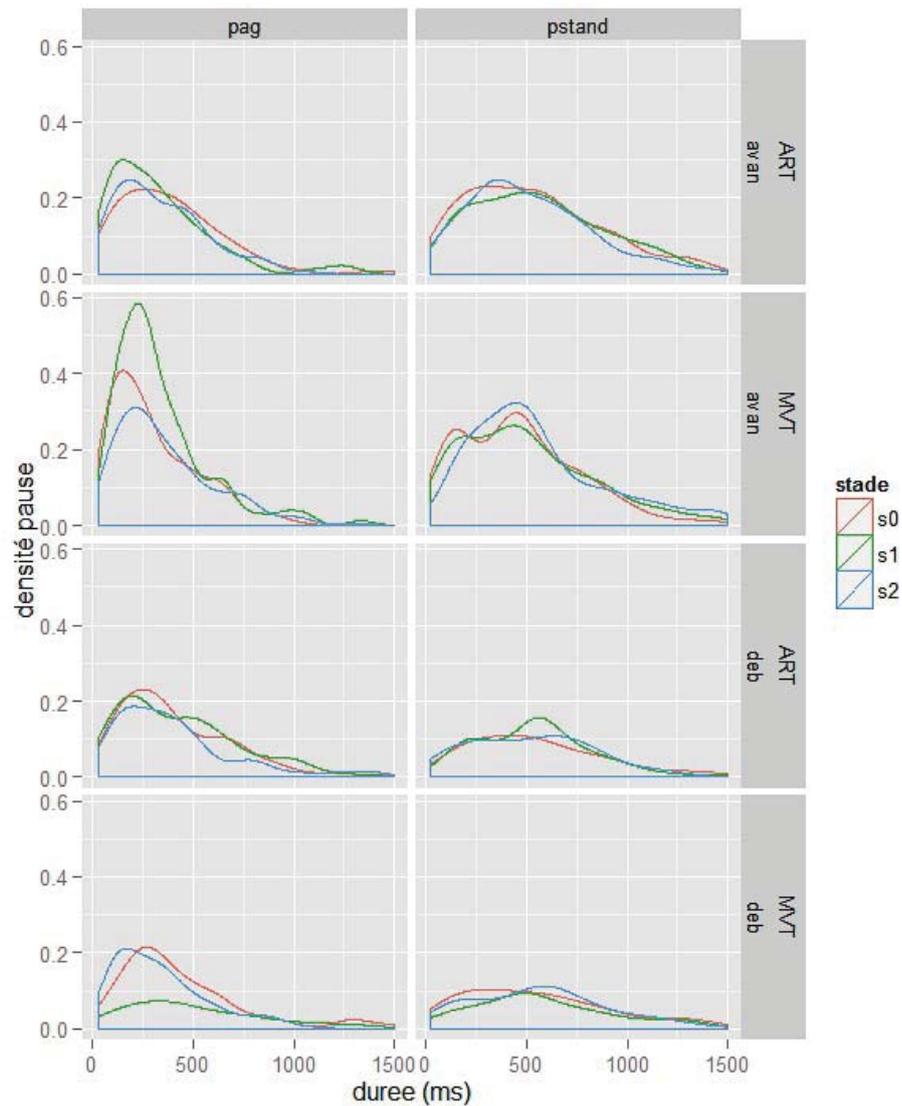


Figure 34 Histogramme lissé de la densité de distribution des pauses agrammaticales (pag) et standard selon la méthode (MVT vs. ART) , le groupe de niveau (debutant vs.avancé) et le stade d'apprentissage (en rouge, s0, en vert s1 et en bleu s2)

Sur cette figure, nous n'observons pas de changements visibles pour les apprenants qui ont suivi les cours de phonétique avec la Méthode Articulatoire (ART), quelque soit leur niveau (avancé (avan) ou débutant (deb)) au cours des

trois stades, ni pour les pauses agrammaticales ni pour les pauses standards. Les courbes des trois stades sont en effet quasiment superposées pour chaque type de pause.

Pour les apprenants avancés du *groupe MVT*, nous remarquons que le nombre de pauses agrammaticales produites semble en nette augmentation en s1 par rapport à s0 et diminue en s2. En ce qui concerne la durée des pauses agrammaticales, nous n'observons pas de changement.

En revanche, en ce qui concerne les apprenants débutants du *groupe MVT*, nous observons un net changement dans l'orientation de la courbe en s1 (pendant l'entraînement, courbe verte) pour l'effectif des pauses agrammaticales. Il semblerait que les apprenants de niveau débutant diminuent drastiquement le nombre de pauses agrammaticales produites en s1 par rapport à s0, indiquant une meilleure planification (*cf.* discussion étude pilote).

Après le passage à l'écrit (s2), le nombre de pauses agrammaticales augmente à nouveau. En ce qui concerne la durée des pauses agrammaticales, nous observons une nette diminution du nombre de pag < 700 ms entre s0 et s1 et leur ré augmentation entre s1 et s2. Cette observation va dans le sens de notre quatrième hypothèse selon laquelle un passage à l'écrit trop précoce entraîne une chute de la fluence.

Nous pensons que les résultats observés chez les apprenants avancés du *groupe MVT* peuvent s'expliquer d'une part par la longueur des textes lus et d'autre part, pas la fossilisation des stratégies erronées de décodage (*cf.* résultats étude pilote).

En fonction de ces résultats et des conclusions de l'étude pilote, nous nous avons comparé plus précisément le nombre de pauses agrammaticales produites par les apprenants débutants du *groupe MVT* et ceux du *groupe ART*. Les histogrammes suivants (figures 35 et 36) présentent le nombre de pauses agrammaticales par locuteur, par stade d'apprentissage (s0, s1 et s2), pour chaque méthode.

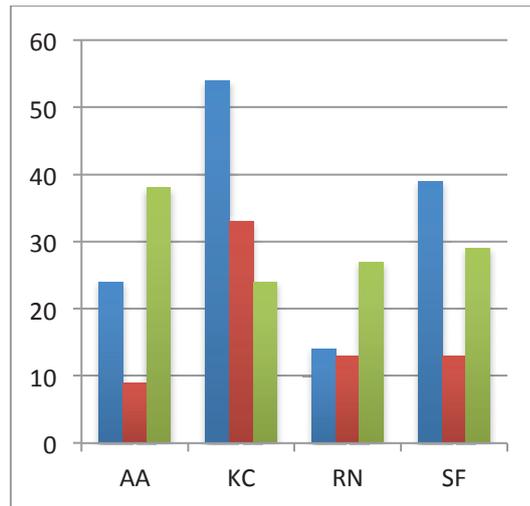


Figure 35 Histogramme du nombre de pauses agrammaticales pour les locuteurs débutants du groupe MVT en fonction du stade d'apprentissage : s0 (en bleu), s1 (en rouge) et s2 (en vert).

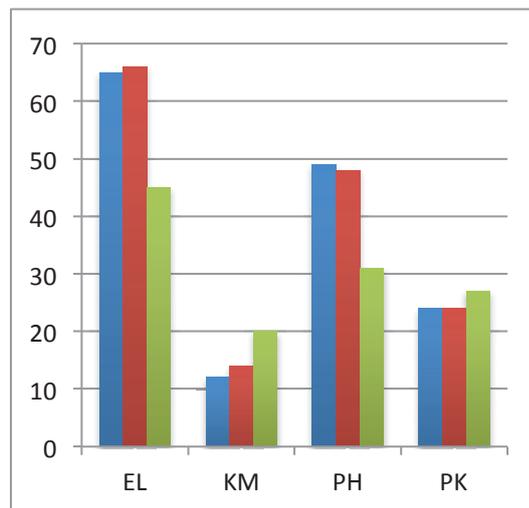


Figure 36 Histogramme du nombre de pauses agrammaticales pour les locuteurs débutants du groupe ART en fonction du stade d'apprentissage : s0 (en bleu), s1 (en rouge) et s2 (en vert).

Tous les apprenants débutants du *groupe MVT* produisent moins de pauses agrammaticales en s1. Le nombre de pauses agrammaticales diminue même drastiquement pour trois apprenants sur les quatre. En revanche, les apprenants débutants du *groupe ART* ne présentent pas de changement radical dans la production des pauses agrammaticales. En effet, deux apprenants sur quatre augmentent légèrement le nombre de pauses agrammaticales produites tandis que chez les deux autres, ce nombre diminue à peine.

Ces observations mettent en évidence l'homogénéité des effets de chaque méthode, indépendamment des variables individuelles, ce qui nous permet de faire valoir l'effet de la MVT dans l'amélioration de la fluence des apprenants. Conformément à nos hypothèses, nous n'avons observé d'amélioration dans la réduction du nombre de pauses agrammaticales que pour les apprenants débutants du *groupe MVT*.

Concernant la durée des pauses agrammaticales, la figure 37 montre une augmentation de la durée globale des pauses (pauses standards et pauses agrammaticales confondues) pour les apprenants débutants du *groupe MVT* et du *groupe ART*.

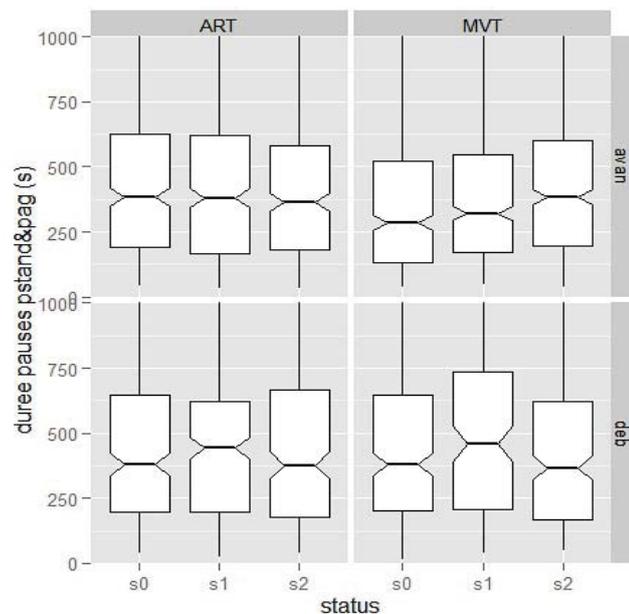


Figure 37 Boxplot de la durée des pauses (pag et pstand confondues) par niveau (avancé et débutant) et par méthode (ART et MVT) dans les trois stades (s0, s1 et s2).

Ces données vont à l'encontre des résultats obtenus dans l'étude pilote où nous avons observé une réduction significative de la durée des pauses agrammaticales après entraînement pour l'apprenant débutant du *groupe test* (MVT). Il est important de rappeler que dans l'étude pilote la fluence des productions a été mesurée à huit semaines d'intervalles. Dans cette étude, la différence entre s0 et s1 n'est que de trois semaines (pour les raisons évoquées dans le chapitre 4). Cette dimension temporelle peut expliquer les différences observées.

D'autre part, la nette diminution du nombre de pauses agrammaticales < 700 ms entre s0 et s1 (*cf.* figure 34) peut expliquer l'augmentation de la durée moyenne des pauses agrammaticales en s1. Cette augmentation serait donc non représentative des performances réelles des apprenants débutants du *groupe test*.

5.2.1.2 Mesures des syllabes accentuées et inaccentuées

Concernant les différents types d'accent initialement défini, nous avons trouvé très peu d'occurrence d'accents de demande de confirmation (appelés LMA), quelque soit le niveau ou le groupe des locuteurs. Pour les rares locuteurs qui produisent des LMA nous n'avons pas observé de différences dans le nombre ou la durée des accents selon les états (avant, pendant, après). Nous pensons que la tâche de lecture oralisée peut avoir une incidence sur le peu d'occurrences de ce type d'accents. Nous supposons également que ces derniers pourraient être révélateurs de variables plus psychologiques (besoin de l'apprenant d'être rassuré) que linguistiques.

Nous avons donc recentré nos analyses sur le nombre et la durée des syllabes accentuées et inaccentuées. La figure 38 représente la densité de distribution des syllabes accentuées (A) et inaccentuées (IA) selon la méthode (ART et MVT, tous niveaux confondus) et la phase d'apprentissage.

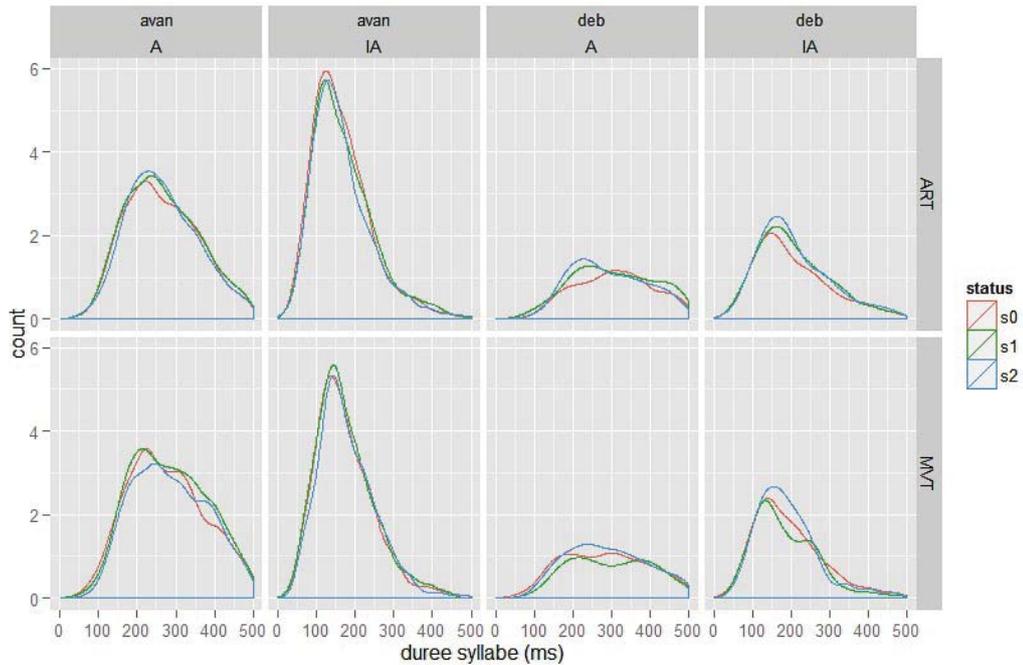


Figure 38 histogramme lissé de la densité de distribution des syllabes accentuées (A) et inaccentuées (IA) selon la méthode (MVT vs. ART) et les stades d'apprentissage (s0, s1 et s2)

Les courbes présentent un dessin assez uniforme quelque soit la nature de la syllabe, la méthode utilisée ou la phase d'apprentissage. Autrement dit, l'entraînement n'a pas eu d'effet sur le nombre de syllabes accentuées et inaccentuées ($s_0 \approx s_1 \approx s_2$) quelque soit la méthode utilisée

Mais qu'en est-il pour la variable durée ? La durée des syllabes accentuées et celle des syllabes inaccentuées avaient été nettement modifiées pour l'apprenant débutant du *groupe test* dans l'étude pilote. En particulier, l'apprenant débutant du *groupe test* avait significativement diminué la durée des syllabes inaccentuées produites, après entraînement.

La figure 39 illustre les mesures de durée des différentes syllabes (accentuées et inaccentuées) en fonction de la méthode (ART vs MVT), du niveau des apprenants (avancé et débutant) et des stades d'apprentissage (s0, s1 et s2).

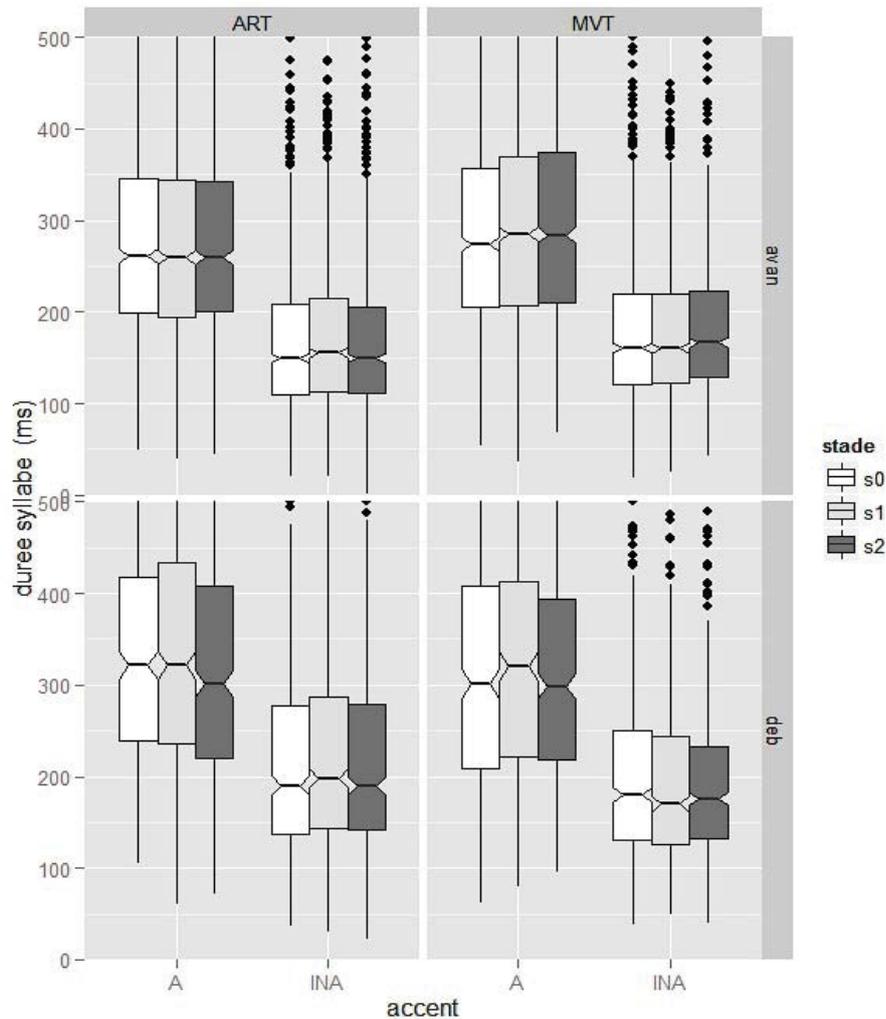


Figure 39 Boxplot des durées syllabiques en millisecondes, selon la situation accentuelle (A : syllabes accentuées ; INA : syllabes inaccentuées), le niveau (avancé vs. débutant), la méthode (ART vs MVT) et le stade de l'apprentissage (en blanc s0, en gris clair s1 et en gris foncé s2)

Sur cette figure, nous observons des différents par méthode et par niveau pour la durée des syllabes accentuées et des syllabes inaccentuées. Nous noterons tout d'abord que les boîtes des apprenants de niveau débutant sont plus étendues que celles des apprenants de niveau avancé, ce qui traduit une plus grande dispersion pour les débutants.

Pour les apprenants débutants du *groupe ART*, nous observons que l'écart de durée des syllabes accentuées entre s0 et s1 est le même ($s_0 \approx s_1$). La durée des syllabes accentuées diminue en s2 pour les apprenants débutants ($s_2 < s_1$). La durée des syllabes inaccentuées montre une augmentation en s1 et une réduction en s2 ($s_1 > s_0$ et $s_1 > s_2$). Les durées des syllabes inaccentuées en s2 sont similaires à celle obtenue en s0 ($s_0 \approx s_2$). Pour les apprenants

débutants du groupe MVT, nous observons le schéma opposé : augmentation de la durée des syllabes accentuées en s1 et réduction de celles-ci en s2 ($s1 > s0$, et $s1 > s2$). La durée des syllabes accentuées en s2 est semblable à la durée de celles-ci en s0 ($s2 \approx s0$). La durée des syllabes inaccentuées chute en s1 ($s1 < s0$) et augmente en s2 ($s2 > s1$ et $s2 \approx s0$).

Globalement, on retrouve les mêmes schémas par méthode pour des apprenants quelque soit leur niveau. En effet, pour les apprenants avancés du *groupe ART*, il semblerait que la durée des syllabes accentuées ne soit pas modifiée en fonction de l'apprentissage ($s0 \approx s1 \approx s2$). Pour les syllabes inaccentuées, on observe chez ses mêmes apprenants, une légère augmentation de la médiane de la durée en s1 et une réduction en s2 montrant une égalité dans la durée des syllabes inaccentuées entre s0 et s2. Pour les apprenants avancés du *groupe MVT*, on remarque une augmentation de la durée des syllabes accentuées en s1 et en s2. La durée des syllabes inaccentuées est identique en s0 et s1 mais augmente en s2. Pour les apprenants débutants du *groupe MVT*, nous observons une augmentation de la durée des syllabes accentuées en s1 et une réduction de celles-ci en s2 ($s1 > s0$, et $s1 > s2$). La durée des syllabes accentuées en s2 est semblable à la durée de celles-ci en s0 ($s2 \approx s0$). La durée des syllabes inaccentuées chute en s1 ($s1 < s0$) et augmente en s2 ($s2 > s1$ et $s2 \approx s0$).

Afin de tester la significativité des différences entre l'évolution des syllabes accentuées et inaccentuées, nous avons estimé un modèle linéaire mixte (LMM ; avec le logarithme de la durée des syllabes comme variable dépendante). Les prédicteurs sont l'état accentuel (A vs. INA), la méthode (MVT vs. ART) et la phase de l'apprentissage (facteur stade à 3 niveaux : s0, s1, s2). Le contraste « différences successives » (Venables & Ripley, 1999) a été associé à ce facteur. Afin d'éviter un modèle trop complexe, nous avons estimé un modèle pour chaque groupe de niveau. Etant donné ces 2 hypothèses simultanées, le seuil de significativité est alors égal à 0.025 (correction de Bonferonni).

Pour les groupes de niveau débutant (MVT et ART), l'analyse porte sur 4811 syllabes. Un intercept aléatoire rend compte de la variabilité des 8 locuteurs et un second rend compte de la variabilité des 457 séquences phonétiques de syllabe.

Les résultats montrent que le contraste syllabes accentuées (A) vs. inaccentuées (INA) n'est pas significatif pour les apprenants débutants du *groupe ART* (beta= 0.023, t=0.66, p=0.5).

Par contre, pour les apprenants débutant du *groupe MVT* la durée des syllabes inaccentuées diminue et celle des accentuées augmente, augmentant ainsi le contraste A vs. INA. Au passage de l'état s0 à s1, l'écart de contraste A vs. INA entre les méthodes est donné par un des coefficients de l'interaction double MVT/stade/syllabes. L'effet s'avère être marginalement significatif (beta= -0.11, t=-2.24, p= 0.029), soit un écart d'environ 11%. Entre s1 et s2, le contraste A vs INA diminue. Le second coefficient de l'interaction double n'est pas significatif (beta=0.039, t=0.79, p=0.46). L'absence de significativité des coefficients d'interaction ne permet pas d'être plus précis et d'affirmer que s2 est distinct de s1 ou de s0.

Pour les apprenants de niveau avancé, l'analyse porte sur 10710 syllabes. Un intercept aléatoire rend compte de la variabilité des 10 locuteurs, et un second rend compte de la variabilité des 612 séquences phonétiques de des syllabes.

Entre s0 et s1, le contraste A vs INA n'est pas significatif (beta= 0.012, t=0.16, p=0.43) pour les apprenants avancés du *groupe ART*.

Pour les apprenants avancés du *groupe MVT*, en revanche, le contraste est plus marqué : la durée des syllabes INA diminue, et celles des syllabes A augmente. Le coefficient de l'interaction double n'atteint toutefois pas la significativité (beta= -0.059, t=-1.97, p=0.06).

Entre s1 et s2, le contraste A vs INA diminue pour les apprenants avancés du *groupe MVT* mais le second coefficient de l'interaction double n'est

pas significatif ($\beta=0.029$, $t=0.98$, $p=0.37$). L'absence de significativité des coefficients d'interaction ne permet pas d'être plus précis et d'affirmer que s_2 est distinct de s_1 ou de s_0 .

Même si l'absence de significativité ne permet pas de déduire les variations de débit à partir des variations de la durée des syllabes inaccentuées, nous remarquons que les mesures de la durée des syllabes inaccentuées pour les apprenants débutants du *groupe test* rejoignent les résultats obtenus dans l'étude pilote : après entraînement à la correction phonétique par la MVT, la durée des syllabes inaccentuées chute (en s_1 de cette étude). De plus, les changements observés après l'introduction de l'écrit pour les apprenants du groupe test (augmentation de la durée des syllabes inaccentuées en s_2 par rapport à s_1) vont dans le sens de nos hypothèses (cf. Hypothèse 4)

5.2.1.3 Densité des groupes rythmiques

Nous avons estimé la densité des groupes rythmiques, en nous basant sur la définition des Unités Inter-Pausales (IPU) proposée par Bertrand & Espesser (2001). Pour cela nous avons mesuré le nombre de syllabes entre deux pauses d'une durée de 220ms. La figure 40 illustre les moyennes et les intervalles de confiance du nombre de syllabes comprises entre deux pauses selon la méthode (ART vs. MVT), le niveau des apprenants (avancé et débutant) et le stade d'apprentissage (s_0 , s_1 et s_2).

Nous remarquons que les apprenants avancés produisent en général des IPU contenant un nombre plus important de syllabes que ceux produits par les apprenants débutants (à l'exception des productions en s_1 des apprenants débutants du *groupe MVT*). Ainsi, ces mesures seraient de bons indicateurs du niveau des apprenants. Nous noterons également les écarts de niveau entre les apprenants avancés des *groupes ART* et *MVT* (écart estimé à environ 1 syllabe), ce qui rejoint notre réflexion sur les problèmes liés à l'évaluation du niveau des apprenants selon la grille donnée par le CECR.

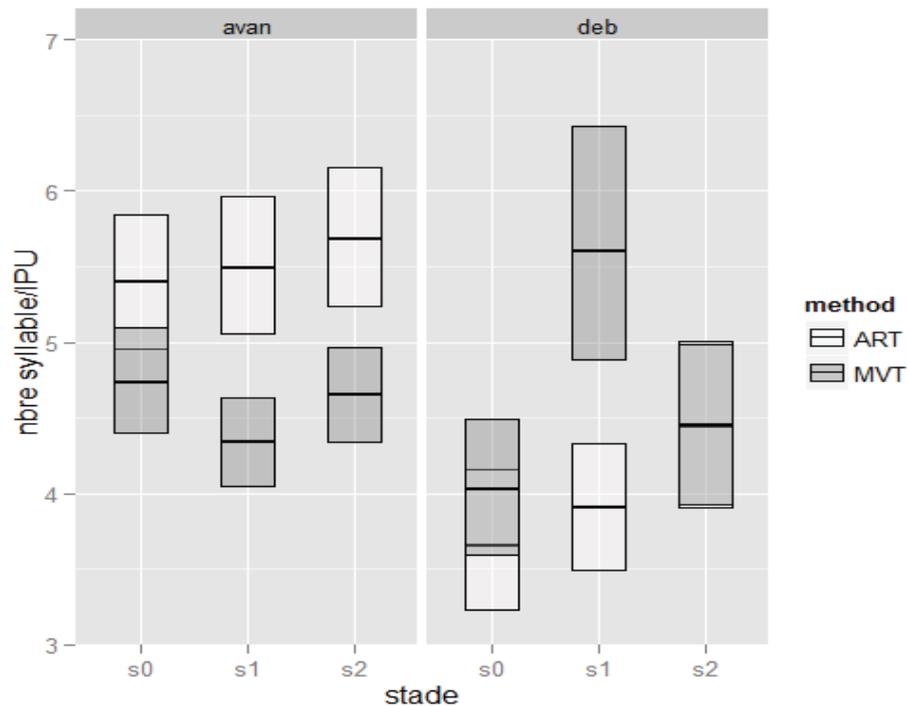


Figure 40 Moyenne et intervalle de confiance du nbre de syllabe entre 2 pauses , selon la methode (en blanc : ART ; en gris : MVT) , le niveau (avancé vs. débutant) et le stade de l'apprentissage (s0, s1 et s2)

Nous observons une évolution constante du nombre moyen de syllabes par IPU pour les apprenants débutants et avancés du *groupe contrôle* (ART) au cours des trois stades d'apprentissage ($s_0 < s_1 < s_2$).

Pour les apprenants des *groupes MVT*, nous observons de nettes différences entre les productions des apprenants de niveau avancé et les apprenants de niveau débutant. Les apprenants de niveau avancés produisent des IPU plus petits en s1 qu'en s0 (le nombre moyen de syllabes par IPU en s1 est inférieur au nombre moyen de syllabes par IPU en s0 ($s_1 < s_0$)). La densité des IPU réaugmente en s2 ($s_2 > s_1$) jusqu'à atteindre le niveau initial ($s_2 \approx s_0$).

Pour les apprenants débutants du *groupe MVT*, en revanche, nous remarquons une très nette augmentation du nombre moyen de syllabes par IPU en s1 (+ 1,5 syllabe) ($s_1 \geq s_0$). La densité moyenne des IPU chute ensuite en s2 ($s_1 > s_2 > s_0$). Cette nette augmentation pourrait faire état d'une amélioration du

débit de parole et donc de la fluence des productions en s1 pour les apprenants débutants du *groupe test*.

Afin de mesurer la significativité des différences entre les trois stades, nous avons estimé un modèle mixte de Poisson. La variable dépendante est constituée par le nombre de syllabe entre 2 pauses, les prédicteurs sont l'état (facteur à 3 niveaux, s0 /s1/s2) et la méthode (MVT vs ART). Le contraste « différences successives » a été associé au stade (s0, s1 et s2). Pour éviter un modèle trop complexe, nous avons estimé un modèle pour chaque groupe. Etant donné ces deux hypothèses simultanées, le seuil de significativité est égal à 0.025 (correction de Bonderont).

Pour les apprenants avancés des *groupes ART* et *MVT*, l'analyse porte sur 2230 observations, soit 2230 IPU. Un intercept aléatoire rend compte de la variabilité des 10 locuteurs. Les résultats ne montrent aucune différence significative en ART entre s1 et s0 (beta=-0.008, z=-0.16, p=0.87) ni entre s2 et s1 (beta=0.046, z=0.89, p=0.37). Le comportement des apprenants avancés du groupe MVT n'est pas différent de celui des apprenants avancés du groupe ART, les 2 coefficients d'interaction sont non significatifs (beta=-0.079, z =-1.14, p=0.25, beta=0.026, z=0.37, p=0.7).

Pour les débutants des *groupes ART* et *MVT*, l'analyse porte sur 1247 observations, soit 1247 IPU. Un intercept aléatoire rend compte de la variabilité des 8 locuteurs. Les résultats ne montrent pas de différence significative pour les apprenants du *groupe ART* entre s0 et s1 (beta=0.1, z=1.49, p=0.13) ni entre s1 et s2 (beta= 0.0099, z=0.145, p=0.88). En revanche, en ce qui concerne les apprenants débutants du *groupe MVT*, la différence du nombre de syllabe entre s1 et s0 est marginalement plus grande en MVT qu'en ART (beta=0.21, z=2.09, p=0.037). La différence du nombre de syllabe entre s2 et s1 est significativement plus faible en MVT qu'en ART (beta=-0.286, z=-2.7, p=0.005). Autrement dit, l'interaction stade/méthode est significative pour les apprenants débutants de MVT. Ce qui indique que les apprenants débutant du *groupe test* sont capables de produire des groupes rythmiques plus longs après entraînement car ils maîtrisent certains traits d'encodage prosodique du français.

En conclusion, le nombre de syllabes par IPU reste stable pour les avancés, quelque soit la méthode. Chez les débutants, en revanche, nous observons une tendance à l'augmentation du nombre de syllabes en s1 (passage de 4 à 5 syllabes), Ces résultats vont dans le sens de nos hypothèses. En effet, seuls les apprenants débutants du *groupe test* ont significativement augmenté la densité des IPU en s1, indiquent un changement dans leurs stratégies de découpage et d'encodage du texte lue, ce qui entraîne une amélioration de la fluence de la lecture oralisée. La différence observée en s2 (chute de la densité) renforce également notre 4^{ème} hypothèse de recherche en appuyant que les transferts prosodiques mis en place sont perturbés par l'introduction de l'écrit ce qui entraîne une chute de la fluence de la lecture oralisée.

5.2.1.4 Mesure du taux d'articulation et du débit de parole

Afin de rendre compte des variations du taux d'articulation, nous avons mesuré le nombre de syllabes par secondes dans les IPU (figure 41).

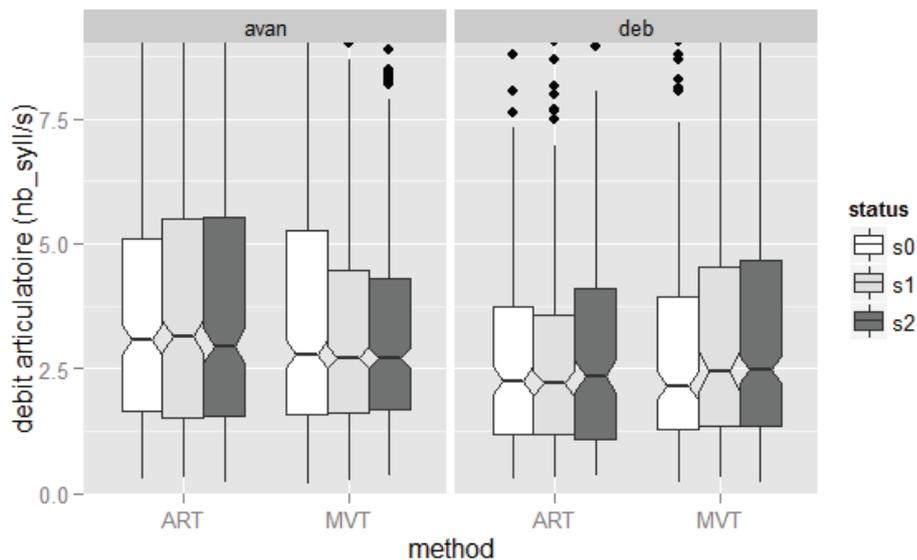


Figure 41 Boxplot du taux d'articulation (exprimé en nombre de syllabes/seconde) selon la méthode (ART vs. MVT) le niveau (avancé vs. débutant) et le stade de l'apprentissage (en blanc s0, en gris clair s1 et en gris foncé s2)

Sur cette figure on ne distingue pas d'évolution du taux d'articulation pour les apprenants avancés du *groupe ART* et du *groupe MVT* quel que soit le stade d'apprentissage.

Pour les apprenants débutants du groupe ART, on remarque une légère augmentation du taux d'articulation entre s1 et s2 (soit après l'introduction de l'écrit). Chez les apprenants débutants du groupe MVT, en revanche, on observe une nette augmentation du taux d'articulation entre s0 et s1, augmentation qui se stabilise entre s1 et s2. Nous savons que l'augmentation du taux d'articulation (cf. chapitre 3) est un des critères de l'amélioration de la fluence. Les apprenants débutants du *groupe MVT* ont une lecture plus fluente après 3 semaines d'entraînement que les apprenants débutants du *groupe ART*. L'introduction de l'écrit entraîne une légère amélioration du taux d'articulation pour les apprenants débutants du *groupe ART* et n'a pas d'incidence sur le taux d'articulation des apprenants débutants du *groupe MVT*.

Pour mesurer le débit de parole, nous avons comparé les variations de la durée des IPU (figure 42) au nombre de syllabes par IPU (figure 40).

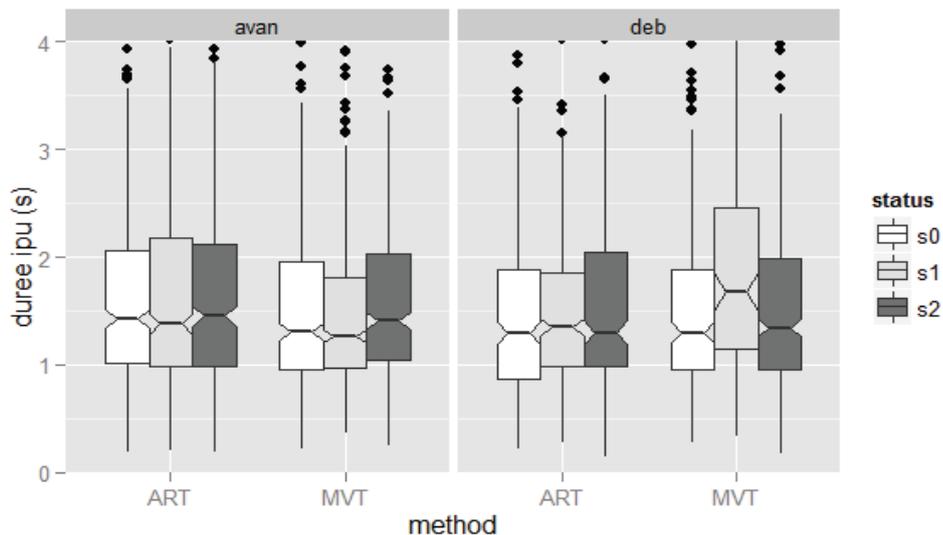


Figure 42 Boxplot de la durée des IPU en secondes, selon la méthode (ART vs. MVT) le niveau (avancé vs. débutant) et le stade de l'apprentissage (en blanc s0, en gris clair s1 et en gris foncé s2)

Sur la figure 42, on voit que l'évolution de la durée des IPU est comparable pour les apprenants avancés des *groupes ART* et *MVT* : la durée des IPU diminue légèrement en s1 et ré augmente en s2. Les analyses statistiques réalisées sur le nombre de syllabes par IPU (figure 40) ne montraient aucune différence entre les deux méthodes et les stades d'apprentissage.

Concernant les apprenants de niveau débutants, nous remarquons une différence importante dans l'évolution de la durée des IPU en fonction de la méthode. Les apprenants débutants du *groupe MVT* ont nettement augmenté la durée des IPU produits. L'augmentation des apprenants débutants du *groupe ART* est beaucoup plus faible. Cela est corrélé avec l'augmentation du nombre de syllabes observé dans la figure 40. Les résultats statistiques ne montraient pas de différence significative pour les apprenants du *groupe ART* entre les différents stades alors que la différence du nombre de syllabes entre s1 et s0 était marginalement significative pour les apprenants débutants du *groupe MVT*. L'augmentation du nombre de syllabes (figure 40) étant plus importante que l'augmentation de la durée des IUP (figure 41), il semblerait que le débit de parole ait augmenté pour les apprenants débutants du *groupe MVT* entre s0 et s1.

De plus, conformément à nos hypothèses, l'introduction de l'écrit vient modifier les tendances précédentes en faisant chuter les performances des apprenants (autant pour le nombre de syllabes par IPU que pour la durée des IPU) du *groupe MVT* entre s1 et s2.

5.2.2 Résultats de l'analyse perceptive

5.2.2.1 Analyse des scores de fluence perçue

Comme dans l'étude pilote, les jugements de fluence par les vingt deux auditeurs natifs se concentrent autour du milieu de l'échelle de magnitude (scores moyens = 3), pour les deux groupes, avant, pendant et après entraînement (*cf.* figure 44).

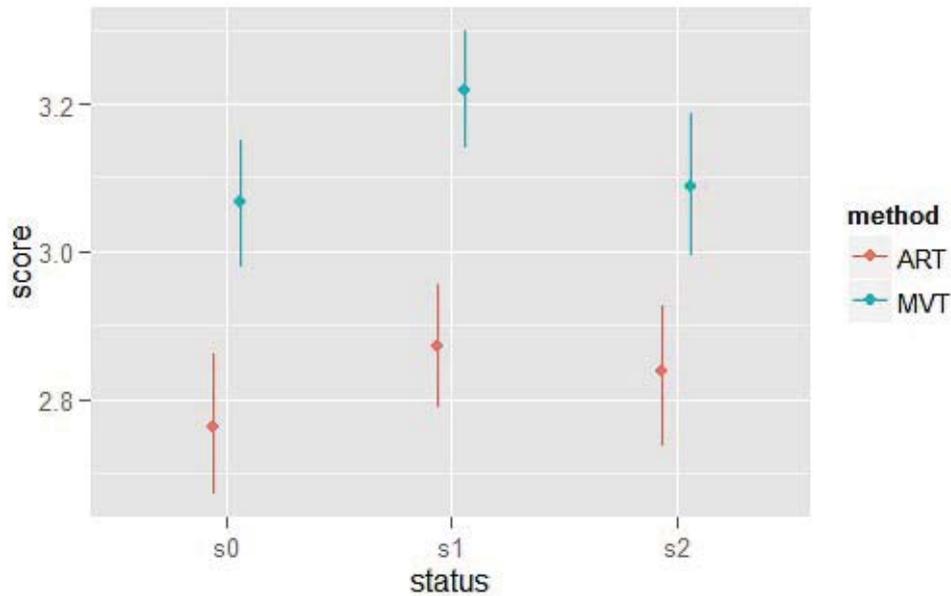


Figure 43 Moyenne des scores en fonction de la méthode (en rouge ART ; en bleu MVT) et du stade d'apprentissage (s0, s1 et s2).

La figure 44 indique que l'évolution des scores est similaire pour les apprenants des *groupes ART* et *MVT* (niveaux confondus) : les auditeurs ont perçu une amélioration de la fluence entre s0 et s1 et une chute de la fluence entre s1 et s2 pour l'ensemble des locuteurs. Sur cette figure, nous remarquons également un écart non négligeable entre les scores attribués aux locuteurs du *groupe MVT* et les scores attribués aux locuteurs du *groupe ART* en s0. Dans le chapitre précédent (chapitre 4, aspects méthodologiques) nous avons discuté de nos difficultés à constituer des groupes homogènes. Nous avons également soulevé le problème de l'évaluation du niveau des apprenants selon les grilles du CECR. Les limites que nous avons soulevées semblent donc se retrouver dans l'évaluation perceptive des natifs.

Afin de voir s'il y a des différences entre les stades s0 et s1 (différences les plus marquées visuellement) selon les méthodes, nous avons mesuré l'écart des pentes des scores obtenus dans les groupes MVT et dans les groupes ART (*cf.* figure 45)

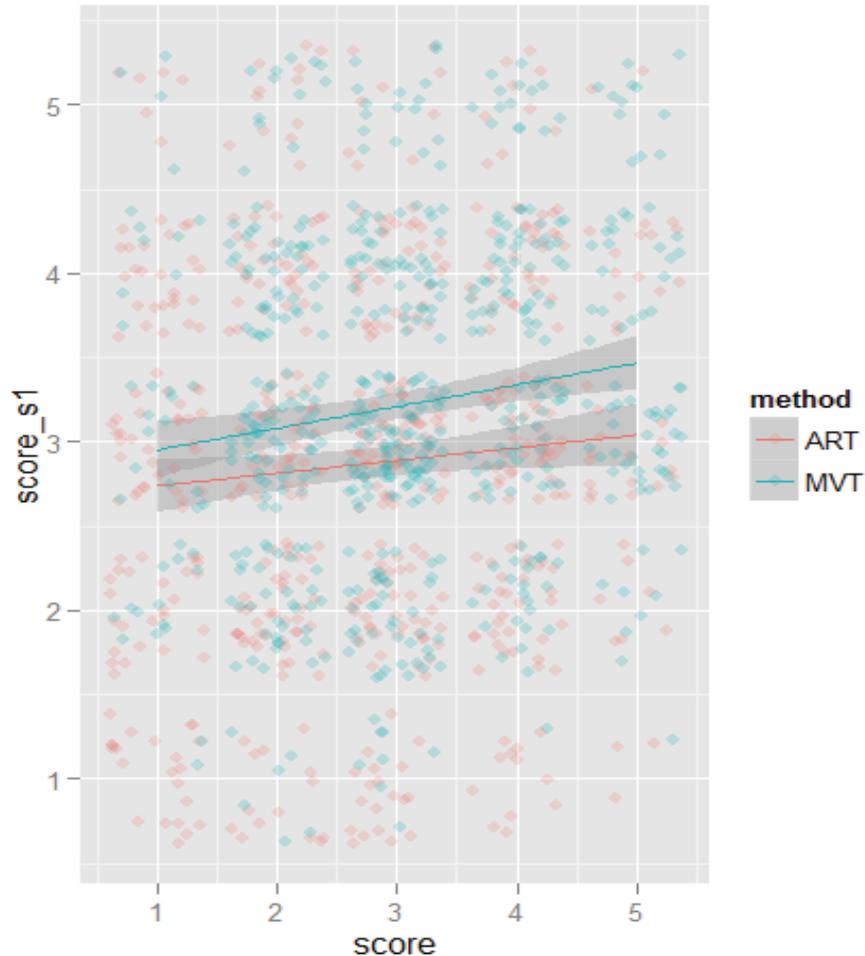


Figure 44 Pente des scores obtenus en s1 par rapport à s0 par méthode (en rouge ART ; en bleu MVT)

Globalement, les résultats ne montrent aucun changement dans l'orientation de la pente pour les apprenants des *groupes ART* ni pour les apprenants des *groupes MVT* ce qui indique qu'aucune amélioration n'a été perçue par les auditeurs après 3 semaines d'entraînement, quelle que soit la méthode.

5.2.2.2 Analyses des temps de réaction

Afin de rendre compte de la confiance des auditeurs dans leur réponse, nous avons aussi mesuré leurs temps de réaction. Ces données sont présentées dans la figure 46.

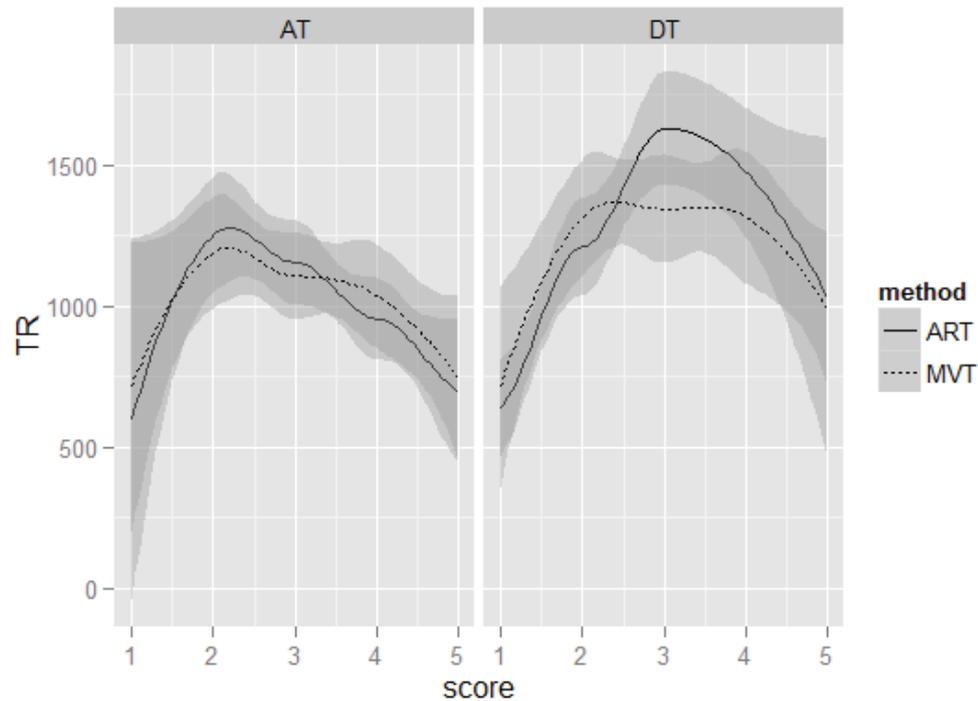


Figure 45 Mesures des temps de réactions en fonction de la méthode (ART : en continu ; MVT : en pointillé) du niveau (AT : apprenants avancés ; DT : apprenants débutants) et des scores obtenus.

Sur la figure 45, les scores à l'extrémité de l'échelle (1 ou 5) correspondent à des temps de réponses rapides. Les apprenants de niveau débutant ont des temps de réaction plus élevés que les apprenants de niveau avancé. Les temps de réaction des scores obtenus pour les apprenants de niveau avancé des groupes ART et MVT suivent globalement la même courbe.

Les scores obtenus pour les apprenants de niveau débutant du groupe ART ont des temps de réaction plus élevés que les scores obtenus pour les apprenants de niveau débutant du groupe MVT, en particulier pour les scores compris entre 3 et 4. Ces temps de réaction sembleraient indiquer que les auditeurs ont eu plus de difficultés à évaluer les productions des apprenants débutants du *groupe ART* que les apprenants du *groupe MVT*, ce qui rejoint les écarts de niveau observés dans la figure 44.

En conclusion, les résultats de cette analyse perceptive ne permettent pas de confirmer nos hypothèses de recherches. De plus, les analyses descriptives ont révélé d'importantes différences entre le niveau perçu des apprenants des *groupes MVT* et *ART* (cf. figure 44). Ces différences pourraient expliquer l'absence de résultats observés. Nous rappelons cependant que les auditeurs n'ont pas jugé l'ensemble du corpus mais seulement quelques extraits. Il se pourrait que les extraits choisis ne soient pas représentatifs de la fluence des productions des apprenants.

5.2.2.3 Questionnaires

Le dépouillement des questionnaires a montré que les enseignants et futurs enseignants de FLE, se référaient en priorité à des paramètres rythmiques pour évaluer la fluence des productions d'apprenants de LE. Le débit, le tempo, les pauses, l'intonation et les hésitations sont les critères les plus cités (ces critères ont été unanimement cités). En deuxième position, on trouve l'aisance, l'authenticité des productions et l'accent étranger. Certains sujets précisent que l'authenticité est pour eux associée autant aux aspects segmentaux que suprasegmentaux. Enfin, en troisième et dernière position, se trouve la compréhensibilité des productions. Les réponses données par les sujets soulignent l'importance des paramètres rythmiques dans l'évaluation de la *fluence perçue*.

Cela renforce le point méthodologique soulevé plus haut : les phrases choisies pour le test de perception ne seraient pas de bons exemples ou ne seraient pas en nombre suffisant pour évaluer la fluence des apprenants. Il est vrai que les auditeurs ont du juger la fluence des apprenants sur la base d'une seule phrase par locuteur et par état, ce qui n'est pas suffisant.

5.2.3 Discussion

Dans cette deuxième étude, nous avons tout d'abord voulu confirmer les résultats obtenus dans l'étude pilote avec un plus grand nombre de sujets. Les résultats des analyses acoustiques de cette étude vont dans le même sens que ceux de l'étude pilote.

Les résultats ne montrent aucune différences significatives pour les apprenants de niveau avancé, quelle que soit la méthode, ou pour les apprenants de niveau débutant du groupe articulatoire entre les trois stades d'apprentissage.

En revanche, des tendances significatives ont été observées chez les apprenants de niveau débutant du groupe MVT entre les stades s0 et s1. Nous remarquons une diminution du nombre de pauses agrammaticales, une augmentation du taux d'articulation et une augmentation du débit de parole pour ces apprenants, en seulement trois semaines d'entraînement. Ces paramètres étant de bons indicateurs de l'amélioration de la fluence (*cf.* chapitres 3 et 4) nous pouvons conclure que seuls les apprenants débutants du groupe MVT ont amélioré la fluence en lecture oralisée s0 et s1 car ils ont mis en place des stratégies de planification différentes. De plus, nous avons observé une nette diminution de la durée des syllabes inaccentuées et une augmentation de la durée des syllabes accentuées entre s0 et s1, indiquant une augmentation de la densité accentuelle et du contraste accentuées/inaccentuées. Ces variations de durée accentuelles indiquent que les apprenants débutants du *groupe MVT* produisent un schéma accentuel plus proche de celui du français après entraînement. De même, ces apprenants produisent des groupes rythmiques beaucoup plus longs après trois semaines d'entraînement *via* la MVT. Or nous savons que les francophones ont tendance à produire des groupes rythmiques beaucoup plus longs que les anglophones (Grosjean & Deschamps, 1972 ; Fletcher, 1991). Autrement dit, les apprenants débutants du *groupe MVT* ont mis en place des stratégies d'encodage adaptées à la LE *via* la maîtrise des patterns prosodiques du français. Ainsi, la fluence mesurée indique bel et bien la mise en place d'une compétence rythmique en langue étrangère.

Ces résultats confirment nos trois premières hypothèses et valident expérimentalement la Méthode Verbo-Tonale car elle permet de mettre en place des processus d'encodage liés à l'automatisation de la compétence prosodique en peu de temps.

De plus, conformément à notre quatrième hypothèse, l'introduction de l'écrit induit une chute des performances pour les apprenants débutant du groupe MVT (chute de la fluence entre s1 et s2). Même si les résultats ne donnent pas de différences significatives, nous remarquons que l'effet négatif de l'introduction de l'écrit se retrouve dans tous les paramètres. On constate en effet une chute du taux d'articulation, une chute de la densité des groupes rythmiques, une augmentation du nombre de pauses agrammaticales et une augmentation de la durée des syllabes inaccentuées. Les performances des apprenants en s2 sont comparables à celles qu'ils avaient avant l'entraînement. L'introduction de l'écrit, ou plutôt de l'intellectualisation que cela amène, a entraîné une régression des performances des apprenants. Les résultats observés sont en contradiction avec les résultats de Freed *et al*, 2004. En effet, dans cette étude, les auteurs ont montré que le renforcement de l'écrit a eu une incidence positive sur la fluence des productions des apprenants. Néanmoins, les activités écrites ayant été combinées à d'autres activités langagières, il est difficile de mettre uniquement en évidence l'incidence de l'apprentissage de l'écrit dans l'amélioration de la fluence. Les auteurs concluent d'ailleurs que les activités de production quelles qu'elles soient (*i.e.* orales ou écrites) amélioreraient la fluence des productions orales.

Les résultats de l'analyse perceptive ne font ressortir aucune progression significative quelle que soit la méthode utilisée, même si l'orientation des scores semblerait indiquer une amélioration des performances pour l'ensemble des apprenants entre s0 et s1 et une chute des performances entre s1 et s2.

Les questionnaires sur les critères d'évaluation choisis par les auditeurs natifs ont quant à eux mis en évidence la sensibilité de ces derniers aux aspects rythmiques.

Nous pensons que l'absence de résultats observés dans cette analyse qualitative peut s'expliquer par le fait que, contrairement à l'étude pilote, les sujets n'ont pas jugé la totalité du corpus mais seulement quelques phrases extraites (1 phrase par stade pour chaque locuteur). Dans le chapitre précédent (chapitre 4, aspects méthodologiques) nous avons expliqué ce choix par rapport

à la durée de passation du test de perception. Si les auditeurs avaient du juger tout le corpus, le test aurait en effet duré plus de 9 heures. Ce qui ressort de la comparaison entre les résultats de l'analyse acoustique et les données qualitatives, c'est que les phrases choisies ne sont peut-être pas les plus représentatives de la performance des apprenants ou qu'il n'est pas possible d'évaluer la fluence d'un apprenant sur un si petit corpus.

De plus, les scores relevés font état d'un écart de niveau important entre les groupes selon les méthodes (tous niveaux confondus). Plusieurs facteurs peuvent expliquer ces différences : premièrement, comme nous venons de le souligner, les phrases choisies ne sont peut-être pas les plus représentatives de la performance réelle des apprenants. Deuxièmement, comme nous l'avons précisé dans le chapitre 4, deux apprenants de niveau débutant (un apprenant de niveau A1 du groupe MVT et un apprenante de niveau A2 du groupe ART) ont abandonné l'expérimentation au bout de quelques semaines. L'équilibre qui avait préalablement établi entre les groupes a pu être perturbé par ce changement. Le niveau des apprenants semblerait en effet constituer un facteur important dans l'évaluation de la fluence par des natifs (Cucchiarini, 2002).

L'écart de niveau perçu, nous amène également à questionner la validité de l'évaluation de la production orale d'apprenants de FLE sur la base des grilles du CECR. Comme nous le disions au chapitre 1, il n'existe pas à l'heure actuelle de critères précis et objectifs pour évaluer la compétence orale en langue étrangère. Les aspects phoniques sont négligés ou réduits à des appréciations subjectives : que signifie, par exemple, « s'exprimer de façon suffisamment claire » ? (cf. figure 1) Nous espérons que des études comme les nôtres, qui traitent de l'interlangue des apprenants et la mesure à différents de l'apprentissage, permettront de dépoussiérer les grilles d'évaluation de l'oral, en proposant des critères objectifs de mesure de la performance des apprenants, comme le placements des pauses, la longueur des groupes rythmiques ou plus globalement le fait que l'apprenant soit capable de segmenter son discours en groupes prosodiquement cohérents.

CONCLUSION ET PERSPECTIVES

Ce travail avait pour objectifs de démontrer l'importance de l'enseignement de la prosodie dans l'apprentissage des langues étrangères et de valider expérimentalement une méthode de correction phonétique, la Méthode Verbo-Tonale d'intégration phonétique.

Ces objectifs étaient motivés par plusieurs questions de recherche :

- Existe-t-il des transferts de compétences entre l'encodage de oral et de l'écrit en langue étrangère ? Quel rôle joue la prosodie dans ces transferts ?
- Qu'est-ce que la compétence prosodique ? Comment l'enseigner ?

Dans le cadre de ce travail, nous nous avons choisi de nous situer au cœur de la construction de la compétence prosodique de l'apprenant, en mesurant la fluence de l'interlangue à différents moments de l'apprentissage. Nous pensons en effet que seuls des travaux de ce type permettent de rendre compte des compétences réelles des apprenants et pas de compétences linguistiques déterminées a priori en référence à une norme native ou à des niveaux linguistiques de références discutables (*cf.* discussion 5.2.3, chapitre 5).

À l'instar de l'hypothèse de la prosodie implicite développée par Fodor (2002), nous avons postulé qu'il existait des transferts de compétences entre l'encodage de l'oral et de l'écrit en LE, liés à la maîtrise de la prosodie de la langue cible (Hypothèse 1). Nous avons supposé que les résultats de l'apprentissage de la prosodie seraient meilleurs si celui-ci était proposé dès le niveau débutant (Hypothèse 2).

Nous avons avancé que la fluence des apprenants, à la fois en oral spontané et en lecture oralisée, dépendrait de leur capacité à encoder la prosodie de la langue-cible. Autrement dit, nous sommes partie du postulat que la fluence reflétait en réalité la compétence prosodique des locuteurs (fluence cognitive). En d'autres termes, nous pensons que la fluence permet de mesurer l'automatisation des processus d'encodage et de décodage associés à la maîtrise de la prosodie de la LE. À l'instar de Paradis (2004) et Germain & Netten (2005), nous avons considéré que l'automatisation de la compétence

prosodique nécessitait une méthodologie adaptée centrée sur une approche non intellectualisée de la matière sonore, *via* la Méthode Verbo-Tonale d'intégration phonétique (MVT). Nous avons postulé que seule la Méthode Verbo-Tonale aurait une incidence positive sur l'amélioration de la fluence en lecture oralisée des apprenants de LE. (Hypothèse 3). Enfin, à l'instar des prérequis de la méthodologie SGAV, nous avons supposé que l'oral et la prononciation devraient être enseignés en priorité car l'introduction précoce de l'écrit entraînerait une chute de la fluence verbale (Hypothèse 4).

Pour tester ces hypothèses de recherche, nous avons mis en place deux études longitudinales. Chacune de ces études a été menée sur une durée de huit semaines avec des apprenants anglophones de Français Langue Étrangère, de niveau débutant et avancé. Les résultats des analyses acoustiques de la fluence des productions de parole lue vont dans le sens de nos hypothèses.

Premièrement, nos résultats nous ont permis d'appuyer l'existence de transferts de compétences entre l'oral et l'écrit en langue étrangère (Hypothèse 1) et de valider l'efficacité de la méthode verbo-tonale (Hypothèse 3). En effet, seuls les apprenants qui ont suivi les cours de correction phonétique via la Méthode Verbo-Tonale ont amélioré la fluence de leurs productions en parole lue. Les caractéristiques prosodiques du français sont mieux maîtrisées après entraînement par la MVT, en particulier pour les apprenants de niveau débutant (hypothèse 2). Les apprenants des groupes MVT produisent moins de pauses agrammaticales et des groupes rythmiques plus longs. Ils ont également un débit plus rapide et montrent un meilleur contraste entre les syllabes accentuées et les syllabes inaccentuées, après entraînement. Ces résultats semblent indiquer que ces apprenants ont développé de meilleures capacités d'encodage et de décodage.

Les résultats des analyses acoustiques ont été corroborés par ceux de l'analyse perceptive pour l'étude pilote mais pas pour la deuxième étude. Dans les chapitres 4 et 5 nous avons soulevé les limites du corpus soumis à la perception des auditeurs dans la deuxième étude. Ces limites pourraient

expliquer pourquoi les résultats perceptifs de la deuxième étude ne permettent pas de conclure à l'efficacité de l'une ou l'autre des méthodes.

Deuxièmement, conformément à notre dernière hypothèse, le passage à l'écrit a entraîné une chute de la fluence en lecture oralisée chez tous les apprenants (Hypothèse 4). Les résultats indiquent que les performances des apprenants après l'introduction de l'écrit sont identiques à celles des apprenants avant la formation, ce qui va à l'encontre des résultats obtenus par Freed *et al* (2004). L'impact de l'introduction de l'écrit mériterait d'être étudié dans de nouvelles recherches pour mieux comprendre les résultats de notre étude et ceux de l'étude de Freed *et al* (2004), et proposer enfin une véritable méthodologie du passage à l'écrit, ce qui n'a jamais été fait en FLE.

Nous envisageons d'ores et déjà plusieurs perspectives de recherche à ce travail. Nous souhaiterions d'abord compléter et approfondir les résultats que nous avons obtenus.

Premièrement, nos données indiquent que les apprenants qui ont suivi les cours de MVT auraient développé de meilleures capacités de décodage. Plusieurs auteurs soutiennent en effet l'existence d'une interdépendance entre la capacité à oraliser correctement l'écrit (fluence) et la compréhension écrite (Golder & Gaonac'h, 1998 ; Hook & Jones, 2002 ; Fuchs *et al*, 2001). Il serait néanmoins intéressant de confirmer le lien entre fluence et compréhension écrite par un protocole plus adapté à contrôler uniquement ces deux facteurs. Le dépouillement des corpus de compréhension écrite pourrait nous donner des premières informations sur ce lien. Il serait également pertinent de compléter ces premières données par un test de résolution de phrases syntaxiquement ambiguës à l'aide de la prosodie. À l'instar de Fodor (2002), nous pensons en effet que la prosodie implicite jouerait le même rôle que la prosodie explicite. En reprenant le protocole proposé par Astésano *et al* (2007), nous projetons de mesurer les bonnes réponses et les temps de réactions dans des tâches de résolution de phrases ambiguës en production (lecture) et en perception (écoute de phrases), chez des apprenants de FLE qui auront suivi des cours de correction phonétique via la MVT, à différents moments de leur formation.

Deuxièmement, nous pensons qu'il pourrait être intéressant de reconduire le test perceptif sur un corpus plus important, en prenant par exemple trois phrases par texte et par état pour chaque locuteur. Nous pourrions également utiliser le protocole de transplantation prosodique développé par Boula de Mareüil *et al* (2004) afin de séparer le niveau suprasegmental et le niveau segmental et de contrôler objectivement les paramètres qui pourraient influencer la perception des auditeurs natifs.

Nous prévoyons ensuite d'exploiter le corpus MULTIPHONIA qui sera bientôt enrichi à de nombreux niveaux afin de permettre des analyses autant segmentales que prosodiques, morphosyntaxiques, syntaxiques, lexicales ou gestuelles en LE.

Nous projetons de tester deux hypothèses de recherche *via* ce corpus : l'idée selon laquelle la prosodie serait le support de la perception et de la production du segmental en langue étrangère, et l'idée selon laquelle les gestes faciliteraient l'apprentissage des caractéristiques prosodiques.

Pour évaluer le rôle de la prosodie dans la perception et la production du segmental en LE, nous extrairons des extraits audio significatifs qui seront annotés manuellement par 3 experts, utilisant le code élaboré par Bertrand *et al.* (2008) avec le logiciel Praat. Les annotations seront automatiquement phonétisées et alignées avec les extraits audio grâce à l'aligneur élaboré par Bigi & Hirst (2012) pour faciliter la détection automatique de voyelles. Nous prévoyons de mesurer et de comparer la répartition formantique des voyelles dans différents contextes de répétition et à différentes étapes d'entraînement. Nous faisons l'hypothèse que les voyelles produites sous l'accent seront mieux réalisées que les voyelles produites en position inaccentuée.

Pour montrer expérimentalement que les gestes pédagogiques ont un effet facilitateur sur la reproduction et la mémorisation d'indices prosodiques non natifs pertinents, nous analyserons toutes les phrases ou mots en français LE répétés par les apprenants anglophones au cours des 8 semaines de cours dans le groupe de MVT. Nous faisons l'hypothèse que les gestes vont aider à

l'ancrage des patterns prosodiques du français et vont aider à corriger les réalisations du système accentuel et segmental de la langue cible. Cela sera réalisé à travers l'analyse systématique des phrases ou mots en français LE répétés par les apprenants anglophones au cours des 8 semaines de cours dans le groupe de MVT. Le pattern accentuel français basique dans une phrase accentuelle (EP) est Accent Initial et Accent Final (cf. Di Cristo, 2000, Jun & Fougeron, 2002, pour une description). Nous faisons l'hypothèse que les gestes vont aider à l'ancrage des patterns prosodiques accentuels de type « AI-AF » et vont aider à corriger les réalisations des systèmes accentuel et segmental de la langue cible. Les annotations des gestes, en particulier les battements, les gestes métaphoriques et les gestes interactifs (cf. typologie de Mac Neil, 2005) seront faites en utilisant le code élaboré par Blache *et al* (2010) avec le logiciel ELAN.

En conclusion, ce travail de recherche nous a permis de replacer la prosodie au cœur de l'enseignement des langues étrangères. Nous avons démontré l'existence de transferts de compétences entre l'encodage de l'oral et de l'écrit en langue étrangère liés à la maîtrise de la prosodie de la langue cible. Nous avons mis en évidence que la mise en place d'une compétence prosodique en LE était le reflet de changements cognitifs profonds indispensables à l'encodage et au décodage de la langue étrangère. Nos données nous permettent de penser que la maîtrise des caractéristiques prosodiques constitue un prérequis au passage à l'écrit en langue étrangère. Ainsi, s'il paraît difficile de retarder le passage à l'écrit de plus de 60h comme le proposait la méthodologie SGAV, il semble néanmoins nécessaire de dissocier activités orales et activités écrites, aujourd'hui introduites simultanément, et d'accorder la priorité à l'enseignement de l'oral et de la prononciation via la MVT.

Convaincue de cette nécessité, nous avons créé, avec Michel Billières et Corine Astésano, une ressource pédagogique numérique sur « la correction phonétique et prosodique en FLE à l'usage des étudiants, des professeurs et

des formateurs », financée par l'Université Ouverte des Humanités (UOH). Cette ressource sera disponible à l'automne 2013.

Enfin, ce travail nous a également permis d'envisager de nombreuses perspectives de recherche pour appuyer et compléter nos premiers résultats. Nous espérons continuer à développer de tels projets dans un futur proche, notamment à travers l'exploitation du corpus MULTIPHONIA qui constitue une importante ressource pour les chercheurs en didactique et en psycholinguistique sur l'apprentissage des langues étrangères.

BIBLIOGRAPHIE

- Abercrombie, D. (1967). *Elements of General Phonetics*, Edimbourg: Edinburgh University Press.
- Alazard, C., Astésano, C., & Billières, M. (2010). The Implicit Prosody Hypothesis applied to Foreign Language Learning: From oral abilities to reading skills. *Paper presented at the 5th Speech Prosody, Chicago*. <http://www.speechprosody2010.illinois.edu/papers/100648.pdf>
- Alazard, C., Astésano, C., & Billières, M. (2011). MULTIPHONIA (MULTImodal database of PHONetics teaching methods in classroom InterActions). Oai :sldr.org:sldr000780. <http://sldr.org/sldr000780/fr>.
- Alazard, C., Astésano, C. & Billières, M. (2012a). From prosodic skills to L2 reading proficiency – An experimental approach to L2 phonetics teaching methods. *Paper presented at the Methods in L2 Prosody (ML2P) workshop, Padova*.
- Alazard, C., Astésano, C. & Billières, M. (2012b). MULTIPHONIA: a MULTImodal database of PHONetics teaching methods in classroom InterActions. *Paper presented at the 8th Conference on Language Ressources and Evaluation (LREC), Istanbul*.
- Allen, L. (1995). The effects of emblematic gestures on the development and access of mental representations of French expressions. *The Modern Language Journal* 79: 521-529.
- Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speaker judgements of non native pronunciation and deviance in segmentals, prosody and syllable structure. *Language Learning*, 42, 529-555.
- Astésano, C. (2001). *Rythme et accentuation du français, invariance et variabilité stylistique*. Paris: Harmattan.
- Astésano, C., Bard, E. G., & Turk, A. (2007). Structural influences on Initial Accent placement in French. *Language and Speech*, 50(3), 423-446.
- Astésano, C., Espesser, R., & Rossi-Gensane, N. (2008). Quelques cas particuliers de détachement à gauche ou la prosodie à l'aide de la syntaxe. *Actes des Journées d'Etude sur la Parole (JEP)*, 109-112.
- Auriacombe, S., Grossman, M., Carvell, S., Gollop, S., Stem, M., & Hurting, H. (1993). Verbal fluency deficits in Parkinson's disease. *Neuropsychology* 7(2), 182-192.

- Avanzi, M. & Martin, P. (2007). L'intonème conclusif : une fin (de phrase) en soi ? *Nouveaux cahiers de linguistique française*, 28, 247-258.
- Baqué, L. (2004). Analyse de l'interlangue de productions phoniques déviantes : un préalable pour la programmation de correction phonétique. In Suso López, Javier et López Carrillo, Rodrigo (coord.): *Le français face aux défis actuels. Histoire, langue et culture*, Granada. Vol. I, 293-302.
- Baqué, L. & Estrada, M. (2010). La notion de diagnostic dans le cadre d'applications de la méthode verbo-tonale à l'apprentissage d'une L2/LE par des bilingues et à la rééducation de patients aphasiques. *Glossa* 108, 53-68.
- Berrendonner, A. (1993). Périodes. *Temps et discours. Louvain: Presses universitaires*, 47-61.
- Best, C. T., McRoberts, G. W., & Sithole, N. N. (1988). The phonological basis of perceptual loss for non-native contrasts: Maintenance of discrimination among Zulu clicks by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance*, 14(3), 345-360.
- Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In Goodman, J., & Nusbaum, H. (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 167-224). Cambridge: MIT Press.
- Best, C. T., & Strange, W. (1992). Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of phonetics*, 20(3), 305-330.
- Best, C. T., Mc Roberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *The Journal of the Acoustical Society of America*, 109, 775-794.
- Bertrand, R. & Espesser, R. (2001). About Speech Overlaps: Prosodic Cues Contribution in Predicting a Change of Speaker. In *proceedings of Prosody 2000, Krakow*, 29-35.
- Bertrand, R., Blache, P., Espesser, R., Ferré, G., Meunier, C., Priego-Valverde, B. & Rauzy, R. (2008). Le CID – Corpus of Interactional Data. *TAL*, 49 (3), 105-134.

- Bigi, B. & Hirst, D. (2012). Speech phonetization alignment and syllabification (SPPAS): a tool for the automatic analysis of speech prosody. *Paper presented at the 6th Speech Prosody Conference, Shanghai.*
- Billieres, M. (1987), Le suivi des élèves en phonétique corrective, *Revue de Phonétique Appliquée*, 82-83-84, 15-28.
- Billieres, M. (1989). Non verbal, phonétique corrective et didactique des langues, *Revue de Phonétique Appliquée* 90, 1-16.
- Billières, M. (1990). L'impact du non verbal dans le processus d'intégration phonétique en langue étrangère. *Revue de Phonétique Appliquée*, 95-96-97, 83-87.
- Billières, M. (1991). Verbo-tonal et didactiques non conventionnelles. *Revue de Phonétique Appliquée*, 99-100-101, 151-158.
- Billieres, M. (1995). Didactique des langues et phonétique la place du verbo tonal, *Revue de phonétique appliquée*, 114, 43-63.
- Billières, M. (2000). *Didactique de l'enseignement de la prononciation. Aspects prosodiques, phonétiques, psycholinguistiques et méthodologiques.* Habilitation à Diriger les Recherches, Université de Toulouse 2, France.
- Billières, M. (2002). Le corps en phonétique corrective. In R. Renard (Ed.), *Apprentissage d'une langue étrangère/seconde 2. La phonétique verbo-tonale.* Bruxelles: De Boeck Université.
- Billieres, M. (2003). Des activités de classe aux activités cognitives en phonétique corrective. *Paper presented at the XIII^{ème} colloque international SGAV, Toulouse.*
- Billières, M. (2005). Les pratiques du verbo-tonal. Retour aux sources. In M. Berré (Ed.), *Linguistique de la parole et apprentissage des langues. Questions autour de la méthode verbo-tonale de P. Guberina* (pp. 67-87). Mons : CIPA.
- Billières, M. (2005). Codage phonologique et boucle articulatoire en mémoire de travail - Un support pour la facilitation de l'accès à l'oral et à la lecture pour des publics débutants en français langue étrangère. *Paper presented at the Colloque Association for French Language Study (AFLS), Chambéry.*
- Billieres, M. (2008). Le statut de l'intonation dans l'enseignement/apprentissage du FLE, *Le français dans le monde, série recherches et applications*, 43, 27-37.

- Billières, M., Gaillard, P., & Magnen, C. (2005). Surdit  phonologique et cat gorisation: perception des voyelles fran aises par les hispanophones. *Revue parole*, 33, 9-33.
- Birdsong, D. (Ed.). (1999). *Second language acquisition and the critical period hypothesis*. Routledge.
- Birdsong, D. (2003) Authenticit  de prononciation en fran ais L2 chez des apprenants tardifs anglophones: Analyses segmentales et globales. *Acquisition et Interaction en Langue  trang re*, 18, 17-36.
- Blache, P. Bertrand, R., & Ferr , G. (2008). Creating and exploiting multimodal annotated corpora. *Paper presented at the 6th Conference on Language Resources and Evaluation (LREC), Marrakech*.
- Blache, P., Bertrand, R., Bigi, B., Bruno, E., Cela, E., Espesser, R., Ferr , G., Guardiola, M., Hirst, D., Magro, E. P., Martin, J. C., Meunier, C., Morel, M. A., Murisasco, E., Nesterenko, I., Noc ra, P., Pallaud, B., Pr vot, L., Priego-Valverde, B., Seinturier, J., Tan, N., Tellier, M. & Rauzy, S. (2010). Multimodal Annotation of Conversational Data. In *proceedings of Linguistic Annotation Workshop 2010, Uppsala*, 186-191.
- Blanche-Benveniste, C. (1997). La notion de variation syntaxique dans la langue parl e. *Langue fran aise*, 115, 19-29.
- Blanche-Benveniste, C., & Bilger, M. (1999). Fran ais parl -oral spontan ; quelques r flexions. *Revue fran aise de linguistique appliqu e*, 4(2), 21-30.
- Blanco, X., Baqu , L., & Estrada, M. (2006). Observations sur le r le de la prosodie dans un mod le fonctionnel global de la langue. In Bogacki, Krzysztof; Miatluk, Anna (2006): *Semantic Relations in Language and Culture*, Wydawnictwo Uniwersytetu w Białymstoku, Białystok (Pologne). 35-62.
- Bloch, B. (1950). Studies in Colloquial Japanese IV: Phonemics, *Language*, 26, 86-125.
- Boersma, P., & Weenink, D. (2005). Praat: doing phonetics by computer, <http://www.praat.org>
- Bolinger, D. (1981). *Two Kinds of Vowels, Two Kinds of Rhythm*, Bloomington : IULC Publications, 68 p.
- Bongaerts, T., Mennen, S., & Slik, F. V. D. (2000). Authenticity of pronunciation in naturalistic second language acquisition: The case of very advanced

- late learners of Dutch as a second language. *Studia linguistica*, 54(2), 298-308.
- Boula de Mareüil, P., Brahim, B., & Gendrot, C. (2004). Role of segmental and suprasegmental cues in the perception of Maghrebian-accented French. In proceedings of *8th International Conference on Spoken Language Processing Jeju Island, Corée*, 2885-2889.
- Boysson-Bardies, B., Sagart, L., & Durand, C. (1984). Discernible differences in the babbling of infants according to target language. *Journal of child language*, 11(1), 1-15.
- Boysson-Bardies, B. (1996). *Comment la parole vient aux enfants: de la naissance jusqu'à deux ans*. Odile Jacob.
- Brumfit, C. (1984). *Communicative methodology in language teaching*. Cambridge: Cambridge University Press.
- Caelen-Haumont, G. (1983). Le rythme dans la parole: une revue des études portant sur le français. In *Actes du 4^{ème} colloque Langages et Signification, Albi*, 233-245.
- Canale, M., & Swain, M. (1981). A Theoretical Framework for Communicative Competence. In Palmer, A.S., Groot, P.J.M., & Tropper, G.A. (Eds.), *The Construct Validation of Tests of Communicative Competence*. (pp. 31-35). Washington DC : TESOL Publications.
- Cenoz, J. (2000). Pauses and hesitation phenomena in second language production. *ITL. Institut voor Toegepaste Linguistik*, 127-128, 53-69.
- Charliac, L. & Motron, A. C. (1998). *Phonétique progressive du français*. Paris: CLE international.
- Cole, J., Hasegawa-Johnson, M., Shih, C., Kim, H., Lee, E. K., Lu, H.Y. (2005). Prosodic parallelism as a cue to repetition disfluency. *Paper presented at the Disfluency in Spontaneous Speech Workshop (Diss-05)*, Aix-en-Provence.
- Conseil de l'Europe, (2000). *Cadre Européen Commun de Référence pour les Langues*. Paris: Didier.
- Coste, D. (1994). *Vingt ans dans l'évolution de la didactique des langues (1968-1988)*. Les Editions Didier.
- Cucchiaroni, C., Strik, H., & Boves, L. (2002). Quantitative assessment of second language learners' fluency: comparisons between read and spontaneous

- speech. *The Journal of the Acoustical Society of America* 6(111), 2862–2873.
- Cucchiaroni, C., Strik, H., & Boves, L. (2000). Quantitative assesment of second language learners' fluency by means of automatic speech recognition technology. *Journal Acoustical Society of Americca*, 107, 989-999
- Cuq, J.P. (2003). *Dictionnaire de didactique du français langue étrangère et seconde*. CLE international.
- Dauer, R. (1983). Stress-timing and syllable-timing reanalyzed, *Journal of Phonetics*, 11(1), pp. 51-62.
- Debaisieux, J. M. (1997). Pour une approche micro et macro du français parlé dans la formation des enseignants de FLE. *Mélanges Pédagogiques*, 23, 27-53.
- Dehaene, S. (2007). *Les neurones de la lecture*. Paris : Odile Jacob.
- Delattre, P. (1965). *Comparing the phonetic features of English, German, Spanish and French: An interim report*. Heidelberg: Julius Groot.
- Delattre, P. (1966). *Studies in French and comparative phonetics*. La Haye: Mouton.
- Dell, F. (1984) L'accentuation dans les phrases en français. In Dell, F., Hirst, D. J., & Vergnaud, J. R. (Eds.), *Formes sonores du langage* (pp. 65-122). Paris : Herman.
- Dellwo, V. (2006). Rhythm and speech rate: A variation coefficient for Δ C. Language and Language Processing. In *proceedings of the 38th Linguistic Colloquium, Piliscsaba, Hongrie*, 231-241.
- De Pietro, J. F., & Wirthner, M. (1996). Oral et écrit dans les représentations des enseignants et dans les pratiques quotidiennes de la classe de français. *Tranel*, 25, 29-49.
- Deterding, D. (2001). The measurement of rhythm: A comparison of Singapore and British English. *Journal of Phonetics*, 29(2), 217-230
- Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility. *Studies in second language acquisition*, 19(01), 1-16.
- Detey, D., Lyche, C., Tchobanov, A., Durand, J., & Laks, B. (2009). Resources phonologiques au service de la didactique de l'oral: Le projet PFC-EF. *Mélanges CRAPEL*, 31, 223-236.

- Detey, S. (2007). Transcription, translittération et didactique de l'oral en FLE au Japon: katakana, rōmaji et orthographe française¹. *Revue japonaise de didactique du français*, 2(1).
- Derwing, T. M., Rossiter, M. J., Munro, M. J., & Thomson, R. I. (2004). Second Language Fluency : Judgments on Different Tasks. *Language Learning*, 54(4), 655-679.
- Dewaele, J. (2002). Individual differences in L2 fluency: The effects of neurobiological correlates. In V. Cook (Ed.), *Portraits of the L2 user* (pp. 221-249). Clevedon, UK: Multilingual Matters.
- Di Cristo, A. (1981). Aspects phonétiques et phonologiques des éléments prosodiques. *Modèles Linguistiques Villeneuve d'Ascq*, 3(2), 24-83.
- Di Cristo, A., & Hirst, D. (1993) Rythme syllabique, rythme mélodique, et représentation hiérarchique de la prosodie du français. *Travaux de l'Institut de Phonétique d'Aix 1993-1994*, 9-24.
- Di Cristo, A., & Hirst, D. (1997). Un procédé d'alignement automatique de transcriptions phonétiques sans apprentissage préalable. *Paper presented at the 4 Congrès Français d'Acoustique, Marseille*.
- Di Cristo, A. (2000). Interpréter la prosodie. *Actes des 23èmes Journées d'Etude sur la Parole*, Aussois, 13-29.
- Dodane, C. (2003). *La langue en harmonie : Influences de la formation musicale sur l'apprentissage précoce d'une langue étrangère*. Thèse de doctorat, Université de Franche-Comté, France.
- Di Cristo, A. (2004). La prosodie au carrefour de la phonétique, de la phonologie et de l'articulation formes-fonctions, *Travaux Interdisciplinaires du Laboratoire Parole et Langage*, 23, 67-211.
- Dolz, J., Schneuwly, B., & De Pietro, J. F (1998). *Pour un enseignement de l'oral: initiation aux genres formels à l'école*. ESF.
- Duez, D. (1993). Acoustic correlates of subjective pauses. *Journal of Psycholinguistic Research*, 22(1), 21-39.
- Duez, D. (1997). Acoustic markers of political power. *Journal of Psycholinguistic Research*, 26(6), 641-654.
- Dufour, S., Nguyen, N., & Frauenfelder, U. H. (2010). Does training on a phonemic contrast absent in the listener's dialect influence word recognition?. *The Journal of the Acoustical Society of America*, 128 (1), EL43-EL48.

- Ejzenberg, R. (1992). *Understanding nonnative oral fluency: The role of task structure and discourse variability*. Ann Arbor : University Microfilms International.
- Favreau, M., & Segalowitz, N. (1983). Automatic and controlled processes in the first and second language reading of fluent bilinguals. *Memory and Cognition*, 11, 565-574.
- Fillmore, C. J. (1979). On fluency. In Kempler, D., & Wang, W. S. Y. (Eds.), *Individuals differences in language ability and language behavior* (pp. 85-102). New York: Academic Press.
- Flege, J. E. (1981). The Phonological Basis of Foreign Accent: A Hypothesis*. *Tesol Quarterly*, 15(4), 443-455.
- Flege, J. E. (1993). Production and perception of a novel, second-language phonetic contrast. *The Journal of the Acoustical Society of America*, 93(3), 1589-1608.
- Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. *Speech perception and linguistic experience: Issues in cross-language research*, 233-277.
- Flege, J. E. (1999). Age of learning and second language speech. In D. Birdsong (Ed.), *Second language acquisition and the Critical Period Hypothesis* (pp. 101–131). Mahwah, NJ: Erlbaum.
- Flege, J. E. (2005). Origins and development of the Speech Learning Model. *Retrieved December, 13, 2005*.
- Flege, J. E., Munro, M. J., & Mackay, I. R. A. (1995). Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America*, 97(5), 3125-3134.
- Fletcher, J. (1991). Rhythm and final lengthening in French. *Journal of Phonetics*, 19(2), 193-212.
- Fodor, J. D. (2002). Prosodic disambiguation in silent reading. *In proceedings of North Eastern Linguistic Society*, 32, 113-132.
- Fonagy, I. (1979). L'accent français, accent probalitaire ? *Studia Phonetica*, 15, 123-233.
- Fonagy, I. (1990). Les langages de la voix. In Aby, J. M, Ales, C., Sansoy, P. (dir) *L'esprit des voix. Étude sur la fonction vocale*. Grenoble: La pensée sauvage.

- Fraisse, P. (1974). *Psychologie du rythme*. Paris: Presses Universitaires de France.
- Freed, B. F. (1995). What Makes Us Think that Students Who Study Abroad Become Fluent? *Second Language Acquisition in a Study Abroad Context*, 9, 123-145.
- Freed, B. F., Segalowitz, N., & Dewey, D. (2004). Contexts of learning and second language fluency in French: Comparing regular classrooms, study abroad, and intensive domestic programs. *Second Language Acquisition*, 26, 275-301.
- Fuchs, L. S., Fuchs, D., Hosp, M., K., & Jenkins, J., R. (2001). Oral Reading Fluency as an Indicator of Reading Competence: A Theoretical, Empirical, and Historical Analysis. In J. Kame'enui, D., Simmons, C., & Williams, J. (Eds.), *Scientific Studies of Reading* (Vol. 5, pp. 239-256). London: Lawrence Erlbaum Associates.
- Gadet, F. (1989). *Le français ordinaire*. Paris: Armand Colin.
- Gadet, F. (1996). Une distinction bien fragile : oral/écrit. *Tranel*, 25, 13-27.
- Germain, C., & Netten, J. (2005). Place et rôle de l'oral dans l'enseignement/apprentissage d'une L2. *Babylonia*, 2(05), 7-10.
- Girard, D. (1972). Méthodes directes et méthodes audio-visuelles. In R. Filipovic.
- Golder, C., & Gaonac'h, D. (2009). *Lire et comprendre. Psychologie de la lecture*, Paris: Hachette.
- Goldman-Eisler, F. (1968). *Psycholinguistics: Experiments in spontaneous speech*. New York: Academic Press.
- Goldman-Eisler, F. (1972). Pauses, clauses, sentences. *Language and Speech*, 15(2), 103-113.
- Grabe, E., Gut, U., Post, B., & Watson, I. (1999). The acquisition of rhythm in English, French and German. In *proceedings of the 14th International Congress of Phonetic Science, San Francisco*, 1201-1204.
- Grabe, E., & Low, E.L. (2002). Durational variability in speech and the rhythm class hypothesis, In Gussenhoven, C., & Warner, N. (Eds.), *Laboratory Phonology 7* (pp. 515-546). Berlin : Mouton De Gruyter.

- Grobet, A. (1997). La ponctuation prosodique dans les dimensions périodique et informationnelle du discours. *Cahiers de linguistique française*, 19, 83-123.
- Grobet, A., & Simon, A. C. (2002). Différents critères de définition des unités prosodiques maximales. *Cahier de Linguistique Française*, 23, 143-163.
- Grosjean, F., & Deschamps, A. (1972). Analyse des variables temporelles du français spontané. *Phonética*, 26(3), 129-156.
- Grosjean, F. (1980). Temporal variables within and between languages. *Towards a cross-linguistic assessment of speech production*, Lang, Frankfurt, 39-53.
- Guaïtella, I. (1991). *Rythme et Parole : Comparaison critique du rythme de la lecture oralisée et de la parole spontanée*. Thèse de Doctorat, Université de Provence I, France.
- Guaïtella, I. (1996). Analyse prosodique des hésitations vocales: *propositions pour un modèle rythmique*. *Revue de Phonétique Appliquée*, 118-119, 113-144.
- Guberina, P., Gospodnetić, Y., Pozojević, M., SKarić, I., & Vuletić, B. (1965). Correction de la prononciation des élèves qui apprennent le français. *Revue de phonétique appliquée*, 1, 81-94.
- Guberina, P. (1973). Les appareils Suvag et Suvag lingua, *Revue de phonétique appliquée*, 27-28, 7-16.
- Harmegnies, B., Delvaux, V., Huet, K., & Piccaluga, M. (2005). Oralité et cognition: pour une approche raisonnée de la pédagogie du traitement de la matière phonique. *Revue PArole*, 34-35-36, 265-336.
- Hartsuiker, R. J., Corley, M., Lickley, R., & Russell, M. (2003). Perception of disfluency in people who stutter and people who do not not stutter: results from magnitude estimation. *Paper presented at the Disfluency in Spontaneous Speech Conference (Diss-03), Sweden*.
- Hasbrouck, J. E., & Tindal, G. (1992). Curriculum-based oral reading fluency norms for students in grades 2 through 5. *Teaching Exceptional Children*, 24(3), 41-44.
- Henderson, A., Goldman-Eisler, F., & Skarbek, A. (1966). Sequential temporal patterns in spontaneous speech, *Language and Speech* 9, 207-216.
- Hook, P. E., & Jones, S. D. (2002). The importance of automaticity and fluency for efficient reading comprehension. *Perspectives*, 28(1), 9-14.

- Intravaia, P. (1993). Pour une étiologie approfondie de l'erreur phonétique. *Revue de phonétique appliquée*, 108-109-110, 239-266.
- Intravaia, P. (2000). *Formations des professeurs de langue en phonétique corrective*. Mons: CIPA.
- Jilka, M. (2000). *The contribution of intonation to the perception of foreign accent*. Thèse de doctorat, Université de Stuttgart, Allemagne.
- Jun, S. A., & Fougeron, C. (2002). Realizations of accentual phrase in French intonation. *Probus* 14,147-172.
- Kaglik, A., & Boula de Mareüil, P. (2009). Perception d'un accent étranger et part de la prosodie selon l'âge de première exposition à la L2: transfert ou phénomène universel en acquisition? *Paper presented at the 6èmes journées d'études en linguistiques, Nantes*.
- Kaglik, A., & Boula de Mareüil, P. (2010). Polish-accented French prosody in perception and production: transfer or universal acquisition process? *Paper presented at the 5th Speech Prosody, Chicago*.
- Karmiloff-Smith, K., & Karmiloff-Smith, A. (2003). *Comment les enfants entrent dans le langage*. Retz.
- Kendon, A. (2004). *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.
- Konopczynski, G. (1986). *Du prélangage au langage : acquisition de la structure prosodique*. Thèse de doctorat, Université de Strasbourg, France.
- Konopczynski, G. (1999). L'acquisition du système prosodique de la langue maternelle et ses implications pour l'apprentissage d'une L2. *Speech Technology Applications in CALL, Eurocall*, 99, 62-70.
- Konopczynski, G., Dodane, C., & Dautry, C. L. (2004). «Wie schade»!, ou l'enseignement de la prosodie d'une L 2 par des méthodes non conventionnelles. *Les signes dans tous leurs états: Hommages à Yves Gilli*, 269.
- Kormos, J., & Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System*, 32(2), 145-164.
- Kormos, J. (2006). *Speech production and Second Language Acquisition*. New York, London: Routledge.

- Krashen, S. (1985). *The Input Hypothesis: Issues and Implications*. London: Longman.
- Kuhl, P. K. (1991). Human adults and human infants show a “perceptual magnet effect” for the prototypes of speech categories, monkeys do not. *Perception & psychophysics*, 50(2), 93-107.
- Kuhl, P. K. (1998). Language, culture and intersubjectivity: The creation of shared perception. *Intersubjective communication and emotion in early ontogeny*, 3, 297-315.
- Kuhl, P. K. Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental science*, 9(2), F13-F21
- Lacheret-Dujour, A & Beaugendre, F. (1999). *La prosodie du français*. Paris : CNRS Editions.
- Lacheret-Dujour, A. (2008). Petit parcours prosodique de Grenoble à Aix-en-Provence : pionniers, dissidents et fédérateurs. In Boë, L. J., & Vilain, C. E. (Eds.), *Un siècle de phonétique expérimentale : Histoire et développement de Théodore Rosset à John Ohala* (pp. 277-297). Lyon : ENS Editions.
- Ladefoged, P. (1975). *A Course in Phonetics*. New York : Harcourt Brace Jovanovich.
- Lafontaine, L., & Messier, G. (2009). Représentations de l’enseignement et de l’évaluation de l’oral chez des enseignants et des élèves du secondaire en français langue d’enseignement. *Revue du Nouvel-Ontario*, 34, 119-144.
- Landercy, A., & Renard, R. (1975). Zones fréquentielles et reconnaissance des voyelles françaises. *Revue de Phonétique Appliquée*, (33-34), 51-79.
- Landercy, A., & Renard, R. (1977). *Éléments de phonétique*. Bruxelles : Didier.
- Lefly, L., & Pennington, B. F. (1991). Spelling errors and reading fluency in compensated adult dyslexics. *Annals of Dyslexia* 41, 141-162.
- Lenneberg, E.H. (1967). *Biological foundations of language*. New york: Wiley.
- Lennon, P. (1990). Investigating Fluency in EFL: A Quantitative Approach. *Language Learning*, 40, 387-417.
- Léon, P. (1992). *Phonétisme et prononciations du français*. Paris : Nathan.

- Léon, P. (1993). *Précis de phonostylistique : parole et expressivité*. Paris : Nathan.
- Levitt, A. G., Utman, J., & Aydelott, J. (1992). From babbling towards the sound systems of English and French: a longitudinal two-case study. *Journal of Child Language*, 19(01), 19-49.
- Levitt, A. G. (1993). The acquisition of prosody: Evidence from French-and English-learning infants. In de Boysson-Bardies, B., de Schonen, S., Jusczyk, P., McNeilage, P., & Morton, J. (Eds.), *Developmental neurocognition: Speech and face processing in the first year of life* (pp. 385-398). Dordrecht: Kluwer.
- Liberman, A.M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, 21(1), 1-36.
- McNeill, D. (2005). *Gesture and Thought*. Chicago: Chicago University Press.
- Martin, P. (2006). Intonation du français : Parole spontanée parole lue. *Estudios de Fonética Experimental*, 15, 133-162.
- Mauranen, A. (2004). Speech Corpora in the classroom. In Aston, G., Bernardini, S., & Stewart, D. (Eds.), *Corpora and Language Learners* (Vol. 17) (pp. 95-211). Amsterdam : John Benjamins.
- Maurel, F. (2004). *Transmodalité et multimodalité écrit/oral : modélisation, traitement automatique et évaluation de stratégies de présentation des structures « visuo-architecturales » des textes*. Thèse de doctorat, Université Toulouse III, France.
- Massaro, D. (1972). Preperceptual images, processing time and perceptual units in auditory perception. *Psychological Review*, 79, 124–145.
- Massaro, D. W. (1974). Perceptual units in speech recognition. *Journal of experimental Psychology*, 102(2), 199-208.
- Massaro, D. W., & Cohen, M. M. (1983). Categorical or continuous speech perception: A new test. *Speech Communication*, 2(1), 15-35.
- Maurer, B. (2002). Contribution à une histoire de l'enseignement de l'oral en primaire par une étude du discours des manuels. *Études de linguistique appliquée*, 1, 53-67.
- Mehler, J., Dommergues, J. Y., Frauenfelder, U., & Segui, J. (1981). The syllable's role in speech segmentation. *Journal of verbal learning and verbal behavior*, 20(3), 298-305.

- Mennen, I. (2006). Phonetic and phonological influences in non-native intonation: an overview for language teachers. *QMUC Speech Science Research Centre Working Papers* WP-9.
- Mertens, P., Auchlin, A., Goldman, J-P., Grobet, A. & Gaudinat, A. (2001). Intonation du discours et synthèse de la parole : premiers résultats d'une approche par balises, *Cahiers de Linguistique Française*, 23, 189-209.
- Meunier, C. (2005). Invariants et variabilité en phonétique. In Nguyen, N., Wauquier-Gravelines, S. & Durand, J. (Eds.) *Phonologie et phonétique: Forme et substance* (pp. 349-374). Paris: Hermès.
- Moirand, S. (1982). *Enseigner à communiquer en langue étrangère*. Paris, Hachette, 1982, 188p (coll. Recherches/applications).
- Molinari, M., Leggio, M., & Silveri, M. (1997). Verbal fluency and agrammatism. In J. Schmahmann (Ed.), *The cerebellum and cognition* (Vol. 41, pp. 325-339). San Diego: Academic Press.
- Molino, J. (2000). *Toward an evolutionary theory of music and language*. Wallin, N.L.
- Munro, M. J., & Derwing, T. M. (1995). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech*, 38, 289-306.
- Munro, M. J., & Derwing, T. M. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 49(s1), 285-310.
- Munro, M. (2008). Foreign accent and speech intelligibility. In Hansen Edward, J. G., & Zampini, M.L. (Eds.), *Phonology and Second Language Acquisition* (pp. 193-218). Amsterdam: John Benjamins.
- Nguyen, N. (2005). La perception de la parole. In N. Nguyen, S. Wauquier-Gravelines & Durand, J. (Eds.), *Phonologie et phonétique: Forme et substance* (pp. 425-447). Paris: Hermès.
- Nonnon, E. (1999). L'enseignement de l'oral et les interactions verbales en classe. *Revue Française de pédagogie*, 129, 87-131.
- Nonnon, E. (2002). Des interactions entre oral et écrit : notes, canevas, traves écrites et leurs usages dans la pratique orale. *Pratiques*, 115 (116), 73-92.

- Nonnon, E. (2005). Entre description et prescription, l'institution de l'objet : qu'évalue-t-on quand on évalue l'oral ? *Repères 31, L'évaluation en didactique du français : résurgence d'une problématique*, 161-188.
- Olynyk, M., D'Anglejan, A., & Sankoff, D. (1987). A quantitative and qualitative analysis of speech markers in the native and second language speech of bilinguals. *Applied Linguistics 8*, 121-136.
- Pallaud, B. (2002). Erreurs d'écoute dans la transcription de données orales. *Revue PArôle, 22-24*, 267-294
- Pallaud, B. (2005). The re-adjustment of word-fragments in spontaneous spoken French. *Paper presented at the Disfluency in Spontaneous Speech Workshop, Aix en Provence*.
- Paradis, M. (2004). *A neurolinguistic theory of bilingualism*. Amsterdam: John Benjamins.
- Pasdeloup, V. (1990) *Modèle de règles rythmiques du français appliqué à la synthèse de la parole*. Thèse de doctorat, Université de Provence, France.
- Pasdeloup, V. (2004). Le rythme n'est pas élastique : étude préliminaire de l'influence du débit de parole sur la structuration temporelle. *Actes des Journées d'Etudes sur la Parole, Fès*.
- Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Native like selection and native fluency. In Richards, J. C., & Schimdt, R. W. (Eds.), *Language and Communication* (pp. 191-226). New York: Longman.
- Perdue, C., & Gaonac'h, D. (2000). Acquisition des langues secondes. In Kail, M. & Fayol, M. (Eds.), *Acquisition du langage* (pp. 215-246). Paris: Presses Universitaires de France.
- Pike, K. (1945). *The Intonation of American English*. Ann Arbor, MI, USA: The University of Michigan Press, 203p.
- Puren, C. (1994). *La didactique des langues étrangères à la croisée des méthodes: essai sur l'éclectisme*. Paris: Didier.
- Ramus, F. (1997). Le rôle du rythme pour la discrimination des langues. *Actes des Journées Internationales d'Orsay des Sciences Cognitives*, 225-229.
- Ramus, F. (1999). La discrimination des langues par la prosodie: Modélisation linguistique et études comportementales. *Paper presented at the 1ère journée d'étude sur l'identification automatique des langues, Lyon*.

- Ramus, F., Hauser, M.D., Miller, C, Morris, D. & Mehler, J. (2000). Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science*, 288, 349-351.
- Ramus, F., Dupoux, E., & Mehler, J. (2003). The psychological reality of rhythm classes: perceptual studies. In *proceedings of the 15th International Congress of Phonetic Sciences, Barcelona*, 337-342.
- Rasier, L., & Hiligsmann, P. (2007). Prosodic transfer from L1 to L2. Theoretical and methodological issues. *Nouveaux cahiers de linguistique française*, 28, 41-66.
- Renard, R. (1979). *La méthode Verbo-Tonale de correction phonétique*. Mons : Didier-Bruxelles.
- Riggenbach, H. (1991). Toward an understanding of fluency: A microanalysis of nonnative speaker conversations. *Discourse Processes*, 14(4), 423-441.
- Rossi, M. (1980). Le français, langue sans accent ? In Fónagy, I., Léon, P., & Carton, F. (Eds.), *L'accent en français contemporain (Studia Phonetica)*, Vol. 15 (pp. 13-51). Montréal : Didier.
- Rossi, M. (1985). L'intonation et l'organisation de l'énoncé. *Phonetica*, 42(2-3), 135-153.
- Ruff, R. M. (1997). The psychological construct of word fluency. *Brain and Language*, 57(3), 394-405.
- Sajaavara, K., & Lehtonen, J. (1978). Spoken Language and the concept of fluency. In Lautamatti, L. & Lindqvist, P. (Eds.), *Focus on spoken language* (pp. 23-57). Jyväskylä: Language Center for Finnish University.
- Sajaavara, K. (1987). Second language speech production: Factors affecting fluency. In Riggenbach, H. & Raupach, M. (Eds.), *Psycholinguistic models of production* (pp. 45-65). Norwood, NJ: Ablex.
- Schmidt, R. (1992). Psychological Mechanisms Underlying Second Language Fluency. *Studies in Second Language Acquisition*, 14, 357-385.
- Segalowitz, S. J., Segalowitz, N. S., & Wood, A. G. (1998). Assessing the development of automaticity in second language word recognition. *Applied Psycholinguistics*, 19(01), 53-67.
- Segalowitz, N. (2000). Automaticity and attentional skill in fluent performance. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 200-219). Ann Arbor : University of Michigan Press.

- Segalowitz, N. (2003). Automaticity and second languages. In Doughty, C. & Long, M. (Eds.), *The handbook of second language acquisition* (pp. 382-408). Oxford: Blackwell Publishers.
- Segalowitz, N. (2010). *Cognitive bases of Second language fluency*. New York: Routledge.
- Scovel, T. (1988). *A time to speak: A psycholinguistic inquiry into the critical period for human speech*. Cambridge : Newbury House.
- Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics*, 10, 209-231.
- Tellier, M. (2008). The effect of gestures on second language memorisation by young children. In Gullberg, M., & de Bot, K. (Eds.) *Special issue Gestures in language development. Gesture*, 8(2), 219-235.
- Tortel, A. (2009). Evaluation qualitative de la prosodie d'apprenants français: apport de paramétrisations prosodiques. Thèse de doctorat, Université de Provence, France.
- Towell, R., Hawkins, R., & Bazergui, N. (1996). The Development of fluency in Advanced Learners of French. *Applied Linguistics*, 17(1), 84-119.
- Trubetzkoy, N.S. ([1938] 1976). *Principes de phonologie*, Paris: Klincksieck.
- Ullakonoja, R. (2008). Pausing as and Indicator of Fluency in the Russian of Finnish Learners. In *proceedings of the 4th Speech Prosody Conference, Campinas, Brésil*, 339-342.
- Vaissière, G. (1997). Langue, prosodie et syntaxe. *Revue Traitement automatique des Langues, numéro spécial Prosodie et syntaxe. ATALA*, 38, 53-82.
- Venables W.N., & Ripley B.D. (2001). *Modern Applied Statistics with S-Plus*. New York: Springer-Verlag (3rd ed.).
- Vihanta, V. V. (1993). Focalisations et autres proéminences en français lu et spontané. *Turun yliopiston julkaisuja. Sarja B: Humaniora*, 202, 258-289.
- Vieru-Dimulescu, B., & Boula de Mareüil, P. (2006). Identification perceptive d'accents étrangers en français. *Paper presented at the 26èmes Journées d'Etude sur la Parole*.
- Weber, C. (2006). Pourquoi les Français ne parlent-ils pas comme je l'ai appris ? *Le Français dans le monde*, 345, 31-33.

- Wenk, B.J., & Wioland, F. (1982). Is French really syllable-timed? *Journal of Phonetics*, 10(2), 193-216.
- Wenk, B. J. (1983). Effets de rythme dans le français parlé. *Recherches sur le français parlé*, 5, 147-162.
- Wenk, B. J. (1985). Speech rhythm and second language acquisition, *Language and Speech*, 28(2), 157-175.
- Wennerstrom, A. (2000). The role of intonation in second language fluency. In H. Riggensbach (Ed.), *Perspectives on fluency* (pp. 102-127). Ann Arbor: University of Michigan Press.
- White, L., & Mattys, S.L. (2007). Calibrating rhythm: First language and second language studies. *Journal of Phonetics*, 35(4), 501-522.
- Wode, H. (1994). Nature, nurture, and age in language acquisition. *Studies in second language acquisition*, 16(3), 325-345.
- Wolf, J.-P. (2008). The effects of backchannels on fluency in L2 oral task production. *System*, 36, 279-294.
- Wood, D. (2004). An Empirical Investigation into the Facilitating Role of Automatized Lexical Phrases in Second Language Fluency Development. *Journal of Language and Learning*, 2(1), 27-50.
- Zellner, B. (1998). Fast and Slow Speech Rate: a Characterisation for French, In *proceedings of the 5th International Conference on Spoken Language Processing (ICSLP)*, Sydney, 3159-3163.
- Ziti, A. (1995). Effets des connecteurs sur le traitement du texte en situation de lecture-compréhension. *Revue de Phonétique Appliquée*, 115-116-117, 385-401.

ANNEXES

Annexe A : Articles publiés sur les deux études longitudinales réalisées

Alazard, C., Astésano, C., & Billières, M. (2010). The Implicit Prosody Hypothesis applied to Foreign Language Learning: From oral abilities to reading skills. *Paper presented at the 5th Speech Prosody, Chicago*. <http://www.speechprosody2010.illinois.edu/papers/100648.pdf>

The Implicit Prosody Hypothesis applied to Foreign Language Learning: From oral abilities to reading skills

Charlotte Alazard, Corine Astésano, Michel Billières

Université de Toulouse - Octogone-Lordat, EA 4156

alazard@univ-tlse2.fr, astesano@univ-tlse2.fr, billieres@univ-tlse2.fr

Abstract

This study investigates the positive influence of *oral skills* training on *reading* abilities for learners of French as a foreign language. We hypothesize that teaching prosody, especially at an early stage of the learning process, will not only improve students' speech fluency and pronunciation skills, but also dramatically improve their strategies to decode written speech.

We conducted a longitudinal study over eight weeks with four English students of French, split into a *Control Group*, where the teacher focused on reading comprehension, and a *Test Group*, where she emphasized phonetic correction and prosodic abilities.

Both acoustic and perception data indicate an improvement in reading fluency within the *Test Group* only, especially for the beginner student.

Index Terms: Prosody, transfer of oral abilities on reading skills, Didactic of French as a Foreign Language, "Verbo-Tonal" Method.

1. Introduction

Prosody is the first element one acquires when learning one's native language [14]. The prosodic structures, the rhythmic and intonational characteristics of one's language, are the first linguistic features a child acquires before any other linguistic levels (phonemes, lexicon, syntax etc...). Recent studies have shown that newborns are able to distinguish their mother tongue from any other language relying on rhythmic cues only [15].

Prosody is thus the most deep-rooted linguistic element in one's system, which makes it hard for learners of a foreign language to get rid of. Foreign language teachers and teaching theoreticians constantly point out the difficulty for students to learn a foreign prosodic structure, which highly contributes to their foreign accent [16].

1.1. The place of prosody in foreign language teaching methods

Although prosody is a fundamental parameter in both language acquisition and language learning, traditional methods focused largely on written texts as a means to learn a foreign language. Only in the mid 60's did foreign language theoreticians take into account its role in language acquisition and its potential impact on foreign language learning. At this time, a teaching method called "*Méthode Structuro-Globale Audio-Visuelle*" or "*SGAV*" [11] was developed. It was the first method to suggest that phonetics, both segmental and supra-segmental, should be taught prior to any other linguistic aspects in a foreign language, especially at an early stage in the learning process. Oral skills are thus taught for 6 weeks before any written element is introduced. This proposition stems from an unusual view on auditory perception processes. Indeed, Guberina [11] made a parallel between hearing-impaired persons and learners of a foreign language: both will

structure the sound background with their own system. In the first case, the system is biased by pathological factors, in the second case the system depends on the learner's own language system. Language learners will thus perceive the prosody and the phonemes of a foreign language through their own prosodic and phonemic system (*cf.* the notion of "*crible phonologique*", ("*phonological sieve*") [12], [17]).

By recommending that oral abilities should be taught first, the SGAV emphasizes the phonetic features of languages. However, unlike the articulatory method which focus on teaching the proper articulatory gestures to language learners, the SGAV uses the prosodic structure of the target language as the 'shell' for pronunciation skills' improvement. More specifically, this prosodic-driven methodology (called the "*Verbo-Tonal Method*", henceforth *MVT*) uses the rhythmic pattern of the target language to bring to light the phonemic specificities of the target language. The third author uses this methodology with learners of French as a Foreign Language (henceforth *FFL*) of various linguistic backgrounds [2]. The teacher first helps the learners familiarize themselves with the prosodic structure of the target language through the repetition of prosodic patterns using logatoms (/dadada/). He then uses the prosodic structure to facilitate phoneme perception. For example, if the learner darkens the timbre of a target phoneme, the teacher will pronounce the phoneme in a prosodically brightening context (accented syllable) and have the learner repeat it in the same context, on the basis that there is a phonological loop between the production and the perception of phonemic features. Namely, a facilitating production context will help the learner perceive the proper phonemic features of the target language and thus it will help them to be able to correctly re-produce in any other prosodic contexts. The *MVT* thus focuses on implicit prosodic learning.

This teaching method was, however, gradually abandoned in the 80's with the arrival of the Communicative Approaches, stipulating that prosodic contours - as well as phonemes - are acquired naturally in the course of the learning process. Furthermore, and contrary to what the SGAV recommends, oral practice is not prioritized anymore: writing and reading activities are re-introduced at the very beginning of the learning process.

We argue that such theoretical choices are questionable, not only in light of our own teaching practice, but also when looking at some issues raised by research in cognitive sciences.

1.2. Questioning the practical experience of class

According to the *Common European Framework of Reference for Languages* (hereafter, *CEFR*), advanced level students (level B) should speak without any accent, thanks to mere repeated contacts with the foreign language. The first author's own experience of teaching *FFL* to English students, however, is in contradiction to such a statement. Even though advanced, students still transfer the prosodic strategies of English onto French: pauses are shorter and more numerous than they

should be in French, and may appear in the middle of verbal phrase, which is non-grammatical in French but not in English [9]. The distribution and acoustic characteristics of accents are those of English rather than French. The length of rhythmic groups is also more erratic than it would be in native French [1].

Although there is a similarity in the quality of prosodic errors in spontaneous speech and reading aloud, the students' performances are noticeably worse in reading than in spontaneous speech activities. This may seem quite surprising considering that students are provided with a script and no planning is involved. Their decoding of the script is also facilitated by punctuation, which provides some structural indication, even though punctuation does not encompass the richness of the prosodic structure ([10]; [13]).

In the Communicative Approaches, reading and writing activities are introduced at the beginning of the learning process, without any preparation. It is as if higher-level competence (such as grapheme/phoneme correspondence or lexical access) is expected to be automatically transferred from one language to the other. Lower-level competence (such as phonetic skills) is neglected, when the use of the phonologic code is actually the most important basis for good reading [8]. Billières [3] poses that an audio-phonatory practice introduced early in the learning process will help students access the phonological conscience of the target language. Once this phonological code is mastered, the acquisition of the other levels of linguistic competence will more easily fall into place.

1.3. Looking for cognitive insights

This proposition echoes some recent findings on native language writing decoding strategies. It's been argued that prosody is used in silent reading (*Implicit prosody hypothesis* [6]; [7]). More specifically, readers project a prosodic structure onto the text, indicating that prosody is treated as part of the input of the written material.

Thus, in language learning, one could argue that if the prosodic structure of the target language is not mastered, the learner will project the prosodic structure of its native language on the text, hindering considerably the decoding process of the written text and its comprehension.

Recent data in cognitive sciences using brain imaging bring an additional insight to this issue [4]. Anatomical and functional characteristics of the brain may hint to a possible transfer of auditory skills onto writing and reading skills: not only are visual and auditory areas interconnected *before* children learn to read, but also the Wernicke's area - where oral and written language comprehension is allegedly processed - is located at the crossroads of those two areas. Hence, any linguistic process that has been learned prior to the development of reading and writing skills could 'feed' the visual modality.

Because prosody is so deep-rooted in our linguistic experience, and since it allegedly governs speech encoding and decoding processes [5], we argue that prosody is at the center of language learning abilities.

1.4. Research questions

On the basis of these observations, we argue that foreign students' fluency in both spontaneous *and* read speech will depend on their capacity to encode the prosody of the target language. The implicit training of prosodic parameters via the *MVT* could have a positive incidence on their speech fluency, and thus on the decoding of written speech. Moreover, should the cognitive processes underlying language learning and language acquisition be comparable, the impact of this training

should be more important at an early stage of language learning,

2. Material and procedure

2.1. Material and experimental design

This pilot study is based on the analysis of the reading of a short argumentative text by 4 English speakers (3 females; mean age: 29; age range: 18-40). They were living in France but were not attending any French class at the time of the experiment. They all had at least school basic knowledge of French, and some of them had attended classes at the French Alliance in Toulouse. A short oral interview allowed the first author to evaluate their level in French according to the *CEFR*. Speakers MM and SR were judged to have an elementary level in French (level A), whereas speakers DR and AM had an advanced level in French (level B).

To test whether a specific training in prosody prior to any reading/writing activity is a condition to better fluency in decoding a written text in the target language, we chose to compare the impact of two radically different teaching methods. The speakers were divided into two groups of classes, both of which lasted 2 hours a week over a period of 8 weeks. In the first class, based on the Communicative Approaches, the teacher (first author) focused on reading activities: global text comprehension, reading out loud and creative writing (*Control Group*). In the second class, based on the *MVT*, she emphasized on phonetic correction and prosodic abilities, through oral practice exclusively: repetition of tongue-twisters and rhymes, rhythmically balanced sentences, improvisation and role games etc. (*Test Group*).

In order to test the level at which the type of training is most beneficial, each class was composed of an A-level and a B-level student. Hence, speakers SR and AM were in the *Control Group*, and speakers MM and DR were in the *Test Group*.

The speakers were tested twice - before and after the 8 weeks' training - on the same text material. The text was created specifically for this study around an argumentative theme (the benefits of bikes over cars in cities). It was adapted from an exercise found in a *FFL* teaching method. It contains 11 sentences organized in 4 paragraphs. The argumentative structure of the text is meant to help the students get involved in their rendering of the out-loud reading.

Speakers were asked to read the text for themselves as many times as they needed before reading it aloud to be recorded. The tests were taken in the same conditions in the two recording sessions 8 weeks apart.

2.2. Experimental procedure

Speakers were tested individually and recorded with head microphones using an audio-recording software (Soundtrack-Pro ©; 16 bits/44kHz).

Recordings were analyzed both acoustically and perceptively. The acoustic analysis allowed us to determine the prosodic parameters that are most representative of *FFL* speakers' fluency before and after training. The perceptual analysis allowed us to complement the acoustic analysis with a qualitative appreciation of their fluency, more representative of a practical class evaluation.

2.2.1. Acoustic analysis

The first two authors listened to the recorded texts and labeled manually the syllables and pauses for each speaker, using

Praat. The prosodic parameters chosen to best reflect speakers' fluency are the following:

- Number, distribution and duration of the different types of pauses (breathing and empty pauses, vocal hesitations and false starts). Break phenomena that were found inside a word or phrase were labeled as non-grammatical, whatever their type.
- Number, distribution (initial vs. final) and duration of perceived accents, and duration of unaccented syllables. Final accents were labeled according to the strength of the associated prosodic boundary (minor vs. major) [1].

Statistics: For each speaker, one-way Anovas were conducted on each type of accents and each type of breaks separately, with the Log Duration and Occurrences as dependent variables, and the State (before/after training) as the independent variable (two-level factor).

2.2.2. Perception test

12 native French subjects, all prospective French teachers, took part in the experiment. They were presented with 264 sentences in 3 different runs: 11 sentences * 4 speakers * 2 States (before/after training), randomized within each run. Six different randomized runs were created and each participant listened to a different combination of three runs. They were asked to judge the fluency of the speakers' production on a magnitude scale from 1 to 7 (1 = poor fluency or beginner; 7 = good fluency or native speaker). The perception test was run using Presentation© and lasted approximately 50 minutes, including a training phase and a two minute pause between each run.

Statistics: To test for the influence of the training method on subjects' judgment of speaker's fluency, we used a linear mixed model with a random intercept to account for the variability across the 12 judges and a random intercept to account for the variability across the 11 sentences. The dependent variable was the response (1 to 7), with the training method (*Control vs. Test Group*) and the State (before vs. after 8 weeks) as factors. To test for the incidence of the training method according to the speakers' level in French, we performed paired t-tests for each speaker, with the mean response of each listener according to State (before and after training).

3. Results

3.1. Acoustic results

Statistical analyses revealed no significant difference for Occurrences of types of accents and types of break for both groups of speakers.

The Duration of all accent and break types (including non-grammatical breathing and silent pauses) was shortened for the *Test Group* speakers after training, but not for the speakers of the *Control Group*. The Duration variable however reached significance for one speaker of the *Test Group* (MM) only, and for a subset of prosodic phenomena. More specifically, this speaker exhibited significantly different strategies before and after training for the durational control of grammatical breathing and silent pauses, as well as Final accents at minor boundaries, and unaccented syllables. The duration of breathing pauses was divided by two between the 2 tests ($F(1; 62) = 34.04; p < .001$), and significantly shortened for silent pauses ($F(1; 55) = 5.17; p = 0.03$). Minor final accents and unaccented syllables were also significantly shortened (respectively, $F(1; 145) = 7.80; p = 0.006$, and $F(1; 500) = 13.65; p < .001$).

3.2. Perception test's results

Participants' responses were concentrated around the middle of the magnitude scale (mean scores = 4) in their judgment on speakers' sentences before and after training in both groups. However, the interaction coefficient was highly significant ($\beta = 0.336, t = 3.52, p < .001$), indicating that the score for the *Test Group* was 0.336 superior to the score of the *Control Group* after training (cf. Figure 1).

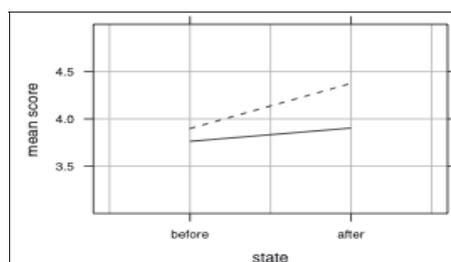


Figure 1 Mean score before/after training in both groups (dotted line = *Test Group*; full line = *Control Group*)

As shown in Table 1, the mean scores for the *Control Group's* speakers did not significantly improve before and after training, whatever their level (+ 0.14 for both speakers: $t\text{-test} = 0.5863; t = 0.5606; df = 11$, for the A-Level speaker SR; and $t\text{-test} = 0.08; t = 1.9149; df = 11$ for the B-level student AM). For the speakers of the *Test Group*, conversely, the mean scores were significantly improved, especially for the beginner speaker (+0.62: $t\text{-test} = 0.0069; t = 3.3166; df = 11$ for the A-Level speaker MM; and + 0.45: $t\text{-test} = 0.0116; t = 3.0225; df = 11$, for the B-Level speaker DR, respectively).

Table 1: Mean scores before/after training for each speaker

Training Methods	Level	State	
		Before	After
Test Group	A-Level (MM)	2,46	3,08
	B-Level (DR)	5,22	5,67
Control Group	A-Level (SR)	3,33	3,47
	B-level (AM)	4,20	4,34

4. Discussion

The aim of this study was to evaluate the impact prosodic training on reading abilities of *FFL* learners, using the *MVT* training method. The main hypothesis was that a training exclusively dedicated to oral skills' development, with emphasis on the prosodic characteristics of the target language, would facilitate the decoding of the written code of the target language.

The prosodic structure can be seen as the 'shell' of a given language, structuring the verbal material at the phonemic, lexical, syntactic and pragmatic levels of the linguistic system. Ontogenetically, the prosodic structure is the first linguistic level the child acquires before all other levels. Because it is so deep-rooted into the speaker's linguistic experience, it is also the most difficult feature to get rid of when speaking a foreign language.

Moreover, following Fodor's *Implicit Prosody Hypothesis*, postulating that the prosodic structure is mapped onto the text to encode the written word-string in silent reading, it is

probable that prosody helps access higher-level competencies by creating encoding automatisms in the target language. Thus, the earlier the FL learner is familiarized with the prosodic structure of the target language, the easier it will be for him to access other levels of the linguistic system. This facilitation of the decoding process will have direct incidence on the reading fluency in the target language, with less occurrences of break phenomena in the speech string, and shorter duration, as well as less accentual density and shorter accent duration.

We ran a pilot study on four English learners of French, divided into two groups: a *Control Group* where the language training was based on traditional teaching methods, and a *Test Group* based on a prosodic oriented teaching method (*MVT*).

Globally, our acoustic data revealed limited impact of language training on speakers' fluency. Speakers exhibited the same number of non-grammatical pauses, false starts and vocal hesitations before and after training, whatever the training method. As far as the Duration variable was concerned, there was a clear tendency for a shortening of grammatical and non-grammatical breathing and silent pauses for the speakers of the *Test Group* only, as well as a tendency to shorten the duration of accents and unaccented syllables. Because pausing and final lengthening are indicative of the planning process, these results suggest a better fluency of the reading of the *Test Group's* speakers. The reduction of unaccented syllables' duration also indicates a faster, more fluent speech rate. These prosodic parameters however reached significance for the A-Level speaker only, indicating that the *MVT* training is most efficient at an early stage of the FL learning. For the *Control Group*, the duration of breaks and accents did not vary before and after training. Their reading fluency was thus not improved.

The perception test confirmed these results insofar as the fluency of the *Control Group's* speakers was not better scored after than before training. Conversely, the perceptual test revealed a significant improvement for the *Test Group's* speakers, with a higher score (stronger improvement) for the beginner student. Globally, the *MVT* improved both speakers' fluency, but more significantly so for the A-Level student. This result supports the second part of our hypothesis postulating that prosody should be taught at an early stage of the learning process.

5. Conclusion

Our results tend to confirm our hypotheses: implicit training of the prosodic structure in the target language improves reading fluency. Indeed, only the two speakers of the *Test Group* showed significant improvement in their reading fluency revealing a better planning and understanding of the written text. This was particularly true for the A-Level, beginner student. Following up on Fodor's *Implicit Prosody Theory*, these results indicate that prosody helps the reader access higher-level competencies by creating encoding automatisms in both native and foreign language - especially when prosody is introduced before any writing or reading activity. This could also have implications for teaching reading to native children. Namely, it could well be that working on oral skills could improve children's reading skills, and help prevent reading difficulties such as grapheme/phoneme correspondence for example.

The follow-up of this study is planned at two levels: first, we'd like to confirm these results on larger groups of English *FFL* learners, since we are conscious that, at present, our Level Variable is confounded with individual speakers. The training phase will also take place over a longer period of

time. More prosodic parameters will also be tested, such as F0 patterns on accented syllables. Also, the syntagmatic distribution of break and accent phenomena will be closely investigated. Second, we plan to test the incidence of oral skills' training on *silent reading comprehension*, hypothesizing that mastering prosody of the target language will not only help the FL students be more fluent, but also help them understand the written text better and quicker.

6. Acknowledgements

The authors wish to thank Dr. Robert Espesser, Laboratoire Parole & Langage, Aix-en-Provence, for his help in acoustic data extraction, and in the statistical analyses of the acoustic and perception data in an earlier version of this study.

7. References

- [1] Astésano, C., *Rythme et accentuation en français: invariance et variabilité stylistique*, Paris : L'Harmattan, 2001.
- [2] Billières, M., "La genèse du crible prosodique. Conséquences sur l'enseignement de la prononciation en langue étrangère". *Cahiers du Centre Interdisciplinaire des Sciences du Langage*, 16: 51-64, 2002.
- [3] Billières, M., "Codage phonologique et boucle articulaire en mémoire de travail. Un support pour la facilitation de l'accès à l'oral et à la lecture pour les publics débutants en français langue étrangère", *CORELA*, Numéro Spécial: Colloque AFLS, 2005.
- [4] Dehaene, S., *Les neurones de la lecture*, Paris : Odile Jacob, 2007.
- [5] Di Cristo, A., "La prosodie au Carrefour de la phonétique, de la phonologie et de l'articulation formes-fonctions", *Travaux Interdisciplinaires du Laboratoire Parole & Langage* 23: 67-211, 2004.
- [6] Fodor, J., "Prosodic Disambiguation in Silent Reading", *Proceedings of NELS 32*, M. Hirotani (Ed.), Amherst, MA: GLSA, University of Massachusetts, 2002a.
- [7] Fodor, J., "Psycholinguistics cannot escape prosody", *Speech Prosody ISCA Workshop*, Aix-en-Provence, 83-88, 2002b.
- [8] Golder, C. & D. Gaonac'h, *Lire et comprendre*. Paris: Hachette, 1998.
- [9] Grosjean, F. & A. Deschamps., "Analyse contrastive des variables temporelles de l'anglais et du français: vitesse de la parole et variables composantes, phénomènes d'hésitation", *Phonetica* 31: 144-184, 1975.
- [10] Guaitella, I., "Analyse prosodique des hésitations vocales : propositions pour un modèle rythmique", *Revue de Phonétique Appliquée*, 118-119 : 113-144, 1996.
- [11] Guberina, P., "La Méthode audio-visuelle structuro-globale", *Revue de Phonétique Appliquée* 1: 35-64, 1965.
- [12] Intravaia, P., *Formation des professeurs de langue en phonétique corrective*. Paris, Didier Erudition; Mons, CIPA, 2000.
- [13] Mertens, P., Auchlin, A., Goldman, JP., Grobet, A., & A. Gaudinat., "Intonation du discours et synthèse de la parole : premiers résultats d'une approche par balises", *Cahiers de Linguistique Française* 23: 189-209, 2001.
- [14] Molino, J., "Toward an evolutionary theory of music and language", In N.L. Wallin, B. Merker, S. Brown (Eds.). *The Origins of Music*. MIT Press: Cambridge, MA: 165-176, 2000.
- [15] Ramus, F., *Rythme des langues et acquisition du langage*. Thèse de doctorat, EHESS, Paris, 1999.
- [16] Renard, R., 1979. *La méthode Verbo-Tonale de correction phonétique*, Mons: Didier-Bruxelles.
- [17] Troubetzkoy, N.S., *Principes de phonologie*, Paris: Klincksiek, 1976 [1938].

Alazard, C., Astésano, C. & Billières, M. (2012). From prosodic skills to L2 reading proficiency
An experimental approach to L2 phonetics teaching methods. *Paper presented at the Methods in
L2 Prosody (ML2P) workshop, Padova.*

From prosodic skills to L2 reading proficiency – An experimental approach to L2 phonetics teaching methods –

Charlotte Alazard, Corine Astésano, Michel Billières

Université de Toulouse - Octogone-Lordat, EA 4156

alazard@univ-tlse2.fr; astesano@univ-tlse2.fr; billieres@univ-tlse2.fr;

ABSTRACT

This paper aims at testing the positive incidence of *oral* skills' training on *reading* abilities for English learners of French as a Foreign Language. A recent preliminary study indicates that emphasis on the assimilation of the L2 prosodic system - especially at an early stage of the learning process - not only improves students' oral fluency but also allows them to develop L2 (prosodic) strategies to decode written speech [2]. In the present study, we follow up on these results on a larger-scale longitudinal study, designed to systematically compare the impact of two phonetics' teaching methods on reading skills for L2 learners of French: the Articulatory Method (AM) and the Verbo-Tonal Method (VTM).

Our data, still under statistical analyses, already show the same tendencies as in the previous study, which is a strong correlation between L2 reading fluency and L2 prosodic skills.

Keywords: prosodic transfers, reading proficiency, L2 acquisition, Verbo-Tonal Method.

1. INTRODUCTION

It is now well established that prosody plays a fundamental role in first language acquisition and structuring [8]. It is the first element one acquires when learning one's native language [15] and one of the last elements to be lost in some cases of aphasiology (*cf.* phonemic jargon, [13]). Prosody is such a central aspect of language that it has been recently shown to play a role in decoding read scripts [10]. Indeed, in silent reading, default prosodic contours are projected onto the script and help encoding and decoding the written text. Namely, prosody operates along with other linguistic levels such as semantics and syntactic structure to maximize the comprehension of the sentence [9]. Hence, not only does the reader use bottom-up processes to decode the visual information, but s/he will also use implicitly memorized prosodic structures (top down processes), as well as previous syntactic or semantic knowledge. Those 'silent' prosodic contours are thus treated as part of the input.

In the case of Second Language Acquisition, it is questionable whether the skills involved in the combined processing of different linguistic levels are available to L2 learners during the reading process. More particularly, empirical observations show that phonological transfers (prosodic and segmental) still largely impact advanced students' speech, and even more so in reading out loud, resulting in non-fluent, hesitant speech. Recently, a

psycholinguistic study also showed that Level B L2 readers are still influenced by their L1 prosodic phonology in a silent reading disambiguation task [7]. These observations contradict the statement of the CEFRL (Common European Framework of Reference for Languages), according to which advanced level students should have '*a clear and natural intonation*' and read with fluency. No reference is however made to the teaching of prosodic - or more globally phonetic - skills even though these are now known to impact on both reading and spontaneous speech skills. On the contrary, the CEFRL seems to imply that those abilities should be naturally acquired thanks to mere repeated contacts with the L2.

As a whole, phonetics is rarely taught in L2 classes, mainly due to its perceived high degree of technicality. Although some L2 segmental aspects may be taught through explicit articulatory exercises, almost no time is dedicated to the teaching of the L2 prosodic system. Therefore, it comes as no surprise that L1 prosodic structure should interfere in the decoding process of written texts. To follow up on the role of silent prosody in decoding written texts in L1, more time should be dedicated to prosodic features' teaching so that proper L2 prosodic structuring could be properly used to support reading decoding.

2. PREVIOUS WORK AND FOLLOW-UP

In order to test the positive incidence of *oral* skills' training on *reading* abilities for L2 learners, we ran a pilot longitudinal study [2] on four English

learners of French (2 advanced and 2 beginners), equally divided into two groups: a *Control Group* where the language training was based on traditional teaching methods, and a *Test Group* based on a prosodic oriented teaching method (*Verbo-Tonal Method, hereafter VTM*) (see below for a description of the method).

The speakers were tested twice - before and after the 8 weeks' training - on the same text material, in the same experimental conditions.

We measured the number distribution and duration of the different types of pauses, the number, distribution (initial *vs.* final) and duration of perceived accents, and the duration of unaccented syllables. Recordings were analyzed both acoustically and perceptively.

Our acoustic and perceptive analyses confirmed that only the two speakers of the *Test Group* (prosodic training) had significantly improved their reading fluency between the 2 tests: they showed significant shortening of breathing and silent pauses duration, as well as shortening of accented and unaccented syllables' duration, revealing a better planning process and a faster speech rate.

This study confirmed that the emphasis on the assimilation of the L2 prosodic system - especially at an early stage of the learning process - not only improves students' oral fluency but also allows them to develop L2 (prosodic) strategies to decode written speech.

In the present study, we follow up on these results on a larger-scale longitudinal study, designed to systematically compare the impact of two phonetics' teaching methods on reading skills for L2 learners of French: the Articulatory Method (hereafter AM) and the Verbo-Tonal Method. According to AM, by far the most widespread method, a good production implies the metalinguistic knowledge of sounds' articulation. The teacher thus provides explicit articulatory descriptions of the different L2 sounds, and then prompts the student to repeat the correct articulatory gestures in order to produce the target sound. No focus is put on prosodic parameters such as rhythm and intonation. VTM, on the contrary, uses the prosodic structure of the target language as the 'shell' for pronunciation skills' improvement. More specifically, the rhythmic pattern of the target language is used to bring to light the phonetic specificities of the target language. The teacher first helps the learners familiarize with the prosodic structure of the target language through the repetition of prosodic patterns using logatoms (/dadada/) or the use of facilitating gestures (for example rising hand movement for salient syllables). In a second phase,

the prosodic structure is used to facilitate phoneme perception (clearer spectral characteristics in stressed syllable) and re-production, on the basis that there is a phonological loop between the production and the perception of phonetic features [3] and [11]. VTM is grounded on the strong assumption that mastering the L2 prosodic system allows for the automatization of lower-level processes, hence facilitating the learning of higher-level processes. Thus, this method suggests that phonetics should be taught prior to any other linguistic aspects in a foreign language.

Despite extremely positive empirical results both in didactics and speech therapy, VTM is still unknown and quite confidential. Our goal is to experimentally validate its benefits on L2 speech fluency, particularly in reading.

3. METHODOLOGY

2.1. Material and experimental design

In order to achieve this goal and to test the positive influence of oral skill's training on reading abilities for L2 learners, we ran a larger longitudinal study over eight weeks. We recorded phonetics' teaching classes, with twenty participants, all English Speakers (15 female; mean age: 32; age range: 20-60). An oral interview allowed us to evaluate their level in French according to the CEFRL: ten of the participants were judged to have an elementary level in French (level A) and ten were judged to have an advanced level in French (level B).

The participants were equally divided into four groups: two groups per method according to their level. Each group received two pronunciation trainings per week - lasting one hour and a half each - for eight weeks.

Both methods were taught by the same teacher - the first author - and recorded in the same experimental conditions. All these recordings constitute a large multimodal database (96 hours of audio-video recording) that will be made shortly available for L2 researchers [1].

The training phase was divided into two parts for both methods. During the first three weeks, the teacher focused on oral practice only: the AM group had explicit articulatory exercises, while emphasis was put on prosodic exercises in the VTM group. Written activities were only introduced after the first three weeks.

Participants undertook three different reading tests: before the training, after three weeks -before the introduction of written activities - and at the end of the eight weeks.

All the texts were informative and adapted from reading activities from FFL's teaching methods. Their content and their complexity in terms of vocabulary and sentence construction were adapted

to each level (level A and B). The texts were presented in a different order for each speaker as they were randomized for each group throughout the 3 different recording sessions.

In each level, speakers were asked to read the text for themselves as many times as they needed before reading it aloud to be recorded. The tests were taken in the same conditions in the three recording sessions.

2.2. Experimental procedure

Speakers were tested individually and recorded in a double-walled sound-isolation booth using Soundtrack-Pro © (16 bits/44kHz).

To begin with, the data were automatically and manually annotated. The acoustic analyses will then allow us to determine the prosodic parameters that are most representative of *FFL* speakers' fluency.

2.2.1 Annotation of the database

The phonetic annotations have been done with the Speech Phonetization Alignment and Syllabification (SPPAS) tool [5]. The aim of this tool is to provide automatic segmentation annotation of utterances, words, syllables and phonemes from a speech recording transcription.

SPPAS produces a phonetic transcription based on a phonetic dictionary. The program offers the possibility to select (automatically or manually) among all the phonemics variants that are proposed.

To complement SPPAS' automatic annotation, the first two authors listened to the recorded texts and labeled manually the different types of accents and the different types of pauses for each speaker, using Praat.

According to our previous results (*i.e.* significant duration shortening of breathing and silent pauses, accented and unaccented syllables) and the literature (for a review of the literature see [12] and [16]), the prosodic parameters chosen to best reflect speakers' fluency were the following:

- Number, distribution and duration of the different types of pauses (breathing and empty pauses, vocal hesitations and false starts). Break phenomena that were found inside a word or phrase were labeled as non-grammatical, whatever their type. Glottal stops replacing obligatory liaisons were also labeled as non-grammatical pauses.
- Number, distribution (initial vs. final) and duration of perceived accents, and duration of unaccented syllables.
- Final accents were labeled according to the strength of the associated prosodic boundary

(minor vs. major). However, *FFL* student do produce a lot of non-standard accents, especially on final accents that are realized with an interrogative intonative contours, characteristic of a confirmation request. We have thus differentiated final accents produced with the expected degree of boundary and final accents realized with the non-expected degree of boundary (longer F0-excursion and longer syllable's duration) typical of a L2 learner's confirmation request. The expected degree of boundaries was determined for each sentence before the annotation, using our expertise of native French speakers.

- Speech and articulation rates (syllables/min)
- Phonation-time ratio (percentage ratio of time speaking to time to take the whole speech sample)
- Mean length of runs (average number of syllables between pauses > 220 ms)
- Dysfluencies per minute (times total number of dysfluencies (false starts, repetitions...) divided by the total time speaking in seconds)
- The number of self-corrections [6]
- The duration ratio between accented and unaccented syllables
- The local F0 patterns on linguistic relevant points [14] and [4]. We compared the peak alignment and the pitch range on accented syllables at the three different stages (before the training, after three weeks and at the end of the eight weeks training) for each speaker.

2.2.2 First tendencies

Our data, still under statistical analyses, already show the same tendencies as in the previous study. Globally, only the two *Test Groups* (advanced and beginner, VTM) showed a reduction of the duration of pausing, accented and unaccented syllables, and a better speech and articulation rate between the first and the second test, indicating a better planning process and a better reading fluency.

However, the differences between the *Test* and *Control groups* appear to be less obvious between the second and the third test, that is after the introduction of reading activities. The introduction of reading activities could have interfered with the automatization of the prosodic transfers. Indeed, the VTM postulated that reading and writing activities should not be introduced before the L2 phonetics system (both segmental and suprasegmental) is acquired, as reading activities (especially the explicit introduction of grapheme/phoneme correspondence) may slow

down the acquisition of the L2 prosodic system and the establishment of the L2 implicit prosodic transfers.

Nevertheless, we noticed a higher utilization of rising F0 patterns on continuative phrases (more similar to native French production) and a higher utilization of self-corrections for the students of the *Test Groups*, between the two last tests (especially for the advanced students), indicative of a higher degree of L2 proficiency for the students of the *Test Groups* after the VTM training.

4. CONCLUSION

Our preliminary results tend to confirm our hypotheses: the acquisition of the L2 prosodic system improves students' reading fluency. Indeed, both the advanced and the beginner students of the *Test Groups*, who received VTM classes, acquired several French prosodic characteristics and were able to transfer and to use them when reading. Indeed, their readings appeared to be more fluent (higher speech rate, reduction of the duration of pauses, accents and unaccented syllables, less disfluencies and more accurate self-repairs), indicating a better planning process.

Therefore, a non-explicit training of the prosodic structure seems to facilitate the automatization of lower-level processes, as both the segmental and the prosodic systems are closer to those of the L2.

The follow-up of this study is planned at two levels : first we want to compare these acoustic results with a perceptual analysis. Forty native French Speakers (20 experts and 20 non experts) will be asked to evaluate the reading fluency and the degree of perceived foreign accents on significant reading extracts randomized by speaker and by time of reading (before the classes, after three weeks and at the end of the 8 week classes). As in our preliminary work, we expect to see a significantly better judgment of readers' fluency for the Test Group (VTM), especially before the introduction of the reading activities.

Then, we would like to compare the measures of the local (accentual) F0 patterns of our English participants with measures of local F0 patterns of native French speakers in order to determine if our results can really be interpreted as positive L2 prosodic transfers or if an 'interlanguage' prosodic system is still partly present after the training. We hypothesize that the 'Test Groups' students will have closer acoustic prosodic cues to the target L2 language than the Control Group.

Our results will ultimately be interpreted in favor of a late introduction of the teaching of grapheme-phoneme correspondence, which seems to hinder our Test Group fluency performances.

5. REFERENCES

- [1] Alazard, C., Astésano, C., & Billières. M. (2012). "MULTIPHONIA: a MULTImodal database of PHONetics teaching methods in classroom InterActions". In proceedings of the 8th LREC 2012 (Istanbul, Turkey) <http://www.lrecconf.org/proceedings/lrec2012/index.html>
- [2] Alazard, C., Astésano, C., & Billières. M. (2010). "The Implicit Prosody Hypothesis applied to Foreign Language Learning: From oral abilities to reading skills". In proceedings of the 5th Speech Prosody 2010 (Chicago, USA) <http://www.speechprosody2010.illinois.edu/papers/100648.pdf>
- [3] Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language-learning device. *Psychol Rev*, 105(1), 158-173.
- [4] Bartkova, K., Bonneau, A., Colotte, V. & M. Dargnat. (2012). "Production of "continuation contours" by French Speakers in L1 (French) and L2 (English)". Proceedings of the 6th Speech Prosody 2012 (Shanghai, China).
- [5] Bigi, B., and Hirst, D. (2012). "Speech Phonetization Alignment and Syllabification: A tool for the automatic analysis of Speech Prosody". Proceedings of the 6th Speech Prosody 2012 (Shanghai, China).
- [6] De Jong, N. H & Wempe, T. (2009). Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Method*, 41, 385-39à.
- [7] Dekydtspotter, L., Edmonds, A., Liljestrånd Fultz, A., & Renaud, C. (2010). Modularity of L2 Sentence Processing: Prosody, Context, and Morphology in Relative Clause Ambiguity in English-French Interlanguage. In proceedings of the 2009 Mind/Context Divide Workshop, ed. Michael Iverson et al. 13-27.
- [8] Di Cristo, A. (2004). La prosodie au Carrefour de la phonétique, de la phonologie et de l'articulation formes-fonctions. *Travaux Interdisciplinaires du Laboratoire Parole & Langage* 23, pp. 67-211.
- [9] Dodane, C., & Brunellière, A. (2006). "Lecture silencieuse et oralisée des phrases relatives: le rôle de la prosodie". In proceedings of the XXVI Journées d'études sur la parole, Dinard, France.
- [10] Fodor, J. (2002). Psycholinguistics cannot escape prosody. In proceeding of *Speech Prosody ISCA Workshop*, Aix-en-Provence: 83-88.
- [11] Jacquemot, C., & Scott, S. K. (2006). What is the relationship between phonological short-term memory and speech processing? *Trends Cogn Sci*, 10(11), 480-486.
- [12] Kormos, J. (2006). *Speech production and second language acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates
- [13] Louis, M., Di Cristo, A., Habib, M., & Hirst, D. (2003). « Transcriptions segmentale et suprasegmentale d'un jargon phonémique ». *Revue Parole*, vol. 23/24, 245-266.
- [14] Mennen, I. (2007). Phonological and phonetic influences in non-native intonation. In Trouvain, J. & Gut, U. (Eds.) *Non-native Prosody: Phonetic Descriptions and Teaching Practice (Nicht-muttersprachliche Prosodie: phonetische Beschreibungen und didaktische Praxis)*. Mouton De Gruyter, 53 - 76.
- [15] Molino, J. (2000). "Toward an evolutionary theory of music and language", in N.L. Wallin, B. Merker, S. Brown (Eds.). *The Origins of Music*. MIT Press: Cambridge, MA. 165- 176.
- [16] Segalowitz, N. (2010). *Cognitive bases of Second language fluency*. New York: Routledge.

Annexe B : Exemple de test de niveau (étude 2)

1. Entretien guidé (5 minutes)

- Pouvez-vous vous présenter ?
- Depuis combien de temps êtes-vous à Toulouse? Pour combien de temps ?
- Pourquoi êtes-vous venus ici ?
- Quels sont vos projets pour la suite ?
- Pourquoi avez-vous choisi d'apprendre le français ?

2. Répétition de trois phrases

Il fait chaud.

Il fait très chaud aujourd'hui.

Qu'est ce qu'il fait chaud aujourd'hui, c'est étouffant.

3. Répétition d'une comptine

Fais dodo Cola mon petit frère

Fais dodo t'auras du lolo

Maman est en haut qui fait des gâteaux,

Papa est en bas

Qui fait du nougat.

4. Compréhension orale (à partir d'un document audio de type dialogue ou interview sur un sujet d'actualité adapté au niveau des apprenants)

- De quel type de document s'agit-il ?
- Quel est le sujet de ce document ?
- Qui parle ?
- Pouvez-vous décrire les personnes ?

Terminer par une question ouverte pour amener l'apprenant à développer son opinion par rapport au sujet de l'extrait audio.

**Annexe C : Niveaux communs de référence et de compétence du
Cadre Européen Commun de Référence (CECR)**

UTILISATEUR EXPERIMENTÉ	C2	Peut comprendre sans effort pratiquement tout ce qu'il/elle lit ou entend. Peut restituer faits et arguments de diverses sources écrites et orales en les résumant de façon cohérente. Peut s'exprimer spontanément, très couramment et de façon précise et peut rendre distinctes de fines nuances de sens en rapport avec des sujets complexes.
	C1	Peut comprendre une grande gamme de textes longs et exigeants, ainsi que saisir des significations implicites. Peut s'exprimer spontanément et couramment sans trop apparemment devoir chercher ses mots. Peut utiliser la langue de façon efficace et souple dans sa vie sociale, professionnelle ou académique. Peut s'exprimer sur des sujets complexes de façon claire et bien structurée et manifester son contrôle des outils d'organisation, d'articulation et de cohésion du discours.
UTILISATEUR INDEPENDANT	B2	Peut comprendre le contenu essentiel de sujets concrets ou abstraits dans un texte complexe, y compris une discussion technique dans sa spécialité. Peut communiquer avec un degré de spontanéité et d'aisance tel qu'une conversation avec un locuteur natif ne comportant de tension ni pour l'un ni pour l'autre. Peut s'exprimer de façon claire et détaillée sur une grande gamme de sujets, émettre un avis sur un sujet d'actualité et exposer les avantages et les inconvénients de différentes possibilités.
	B1	Peut comprendre les points essentiels quand un langage clair et standard est utilisé et s'il s'agit de choses familières dans le travail, à l'école, dans les loisirs, etc. Peut se débrouiller dans la plupart des situations rencontrées en voyage dans une région où la langue cible est parlée. Peut produire un discours simple et cohérent sur des sujets familiers et dans ses domaines d'intérêt. Peut raconter un événement, une expérience ou un rêve, décrire un espoir ou un but et exposer brièvement des raisons ou explications pour un projet ou une idée.
UTILISATEUR ÉLÉMENTAIRE	A2	Peut comprendre des phrases isolées et des expressions fréquemment utilisées en relation avec des domaines immédiats de priorité (par exemple, informations personnelles et familiales simples, achats, environnement proche, travail). Peut communiquer lors de tâches simples et habituelles ne demandant qu'un échange d'informations simple et direct sur des sujets familiers et habituels. Peut décrire avec des moyens simples sa formation, son environnement immédiat et évoquer des sujets qui correspondent à des besoins immédiats.
	A1	Peut comprendre et utiliser des expressions familières et quotidiennes ainsi que des énoncés très simples qui visent à satisfaire des besoins concrets. Peut se présenter ou présenter quelqu'un et poser à une personne des questions la concernant – par exemple, sur son lieu d'habitation, ses relations, ce qui lui appartient, etc. – et peut répondre au même type de questions. Peut communiquer de façon simple si l'interlocuteur parle lentement et distinctement et se montre coopératif.

Niveaux communs de référence. Échelle globale (CECR, 2000 : 25)

		A1	A2	B1						
C O M P R E N	Écouter	Je peux comprendre des mots familiers et des expressions très courantes au sujet de moi-même, de ma famille et de l'environnement concret et immédiat, si les gens parlent lentement et distinctement.	Je peux comprendre des expressions et un vocabulaire très fréquent relatifs à ce qui me concerne de très près (par exemple moi-même, ma famille, les achats, l'environnement proche, le travail). Je peux saisir l'essentiel d'annonces et de messages simples et clairs.	Je peux comprendre les points essentiels quand un langage clair et standard est utilisé et s'il s'agit de sujets familiers concernant le travail, l'école, les loisirs, etc. Je peux comprendre l'essentiel de nombreuses émissions de radio ou de télévision sur l'actualité ou sur des sujets qui m'intéressent à titre personnel ou professionnel si l'on parle d'une façon relativement lente et distincte.						
		D R E	Lire	Je peux comprendre des noms familiers, des mots ainsi que des phrases très simples, par exemple dans des annonces, des affiches ou des catalogues.	Je peux lire des textes courts très simples. Je peux trouver une information particulière prévisible dans des documents courants comme les publicités, les prospectus, les menus et les horaires et je peux comprendre des lettres personnelles courtes et simples.	Je peux comprendre des textes rédigés essentiellement dans une langue courante ou relative à mon travail. Je peux comprendre la description d'événements, l'expression de sentiments et de souhaits dans des lettres personnelles.				
				P A R L	Prendre part à une conversation	Je peux communiquer, de façon simple, à condition que l'interlocuteur soit disposé à répéter ou à reformuler ses phrases plus lentement et à m'aider à formuler ce que j'essaie de dire. Je peux poser des questions simples sur des sujets familiers ou sur ce dont j'ai immédiatement besoin, ainsi que répondre à de telles questions.	Je peux communiquer lors de tâches simples et habituelles ne demandant qu'un échange d'informations simple et direct sur des sujets et des activités familiers. Je peux avoir des échanges très brefs même si, en règle générale, je ne comprends pas assez pour poursuivre une conversation.	Je peux faire face à la majorité des situations que l'on peut rencontrer au cours d'un voyage dans une région où la langue est parlée. Je peux prendre part sans préparation à une conversation sur des sujets familiers ou d'intérêt personnel ou qui concernent la vie quotidienne (par exemple famille, loisirs, travail, voyage et actualité).		
						E R	S'exprimer oralement en continu	Je peux utiliser des expressions et des phrases simples pour décrire mon lieu d'habitation et les gens que je connais.	Je peux utiliser une série de phrases ou d'expressions pour décrire en termes simples ma famille et d'autres gens, mes conditions de vie, ma formation et mon activité professionnelle actuelle ou récente.	Je peux m'exprimer de manière simple afin de raconter des expériences et des événements, mes rêves, mes espoirs ou mes buts. Je peux brièvement donner les raisons et explications de mes opinions ou projets. Je peux raconter une histoire ou l'intrigue d'un livre ou d'un film et exprimer mes réactions.

Niveaux communs de compétence. Grille pour l'auto-évaluation (CECR, 2000 : 26)

B2	C1	C2
<p>Je peux comprendre des conférences et des discours assez longs et même suivre une argumentation complexe si le sujet m'en est relativement familier. Je peux comprendre la plupart des émissions de télévision sur l'actualité et les informations. Je peux comprendre la plupart des films en langue standard.</p>	<p>Je peux comprendre un long discours même s'il n'est pas clairement structuré et que les articulations sont seulement implicites. Je peux comprendre les émissions de télévision et les films sans trop d'effort.</p>	<p>Je n'ai aucune difficulté à comprendre le langage oral, que ce soit dans les conditions du direct ou dans les médias et quand on parle vite, à condition d'avoir du temps pour me familiariser avec un accent particulier.</p>
<p>Je peux lire des articles et des rapports sur des questions contemporaines dans lesquels les auteurs adoptent une attitude particulière ou un certain point de vue. Je peux comprendre un texte littéraire contemporain en prose.</p>	<p>Je peux comprendre des textes factuels ou littéraires longs et complexes et en apprécier les différences de style. Je peux comprendre des articles spécialisés et de longues instructions techniques même lorsqu'ils ne sont pas en relation avec mon domaine.</p>	<p>Je peux lire sans effort tout type de texte, même abstrait ou complexe quant au fond ou à la forme, par exemple un manuel, un article spécialisé ou une œuvre littéraire.</p>
<p>Je peux communiquer avec un degré de spontanéité et d'aisance qui rende possible une interaction normale avec un locuteur natif. Je peux participer activement à une conversation dans des situations familières, présenter et défendre mes opinions.</p>	<p>Je peux m'exprimer spontanément et couramment sans trop apparemment devoir chercher mes mots. Je peux utiliser la langue de manière souple et efficace pour des relations sociales ou professionnelles. Je peux exprimer mes idées et opinions avec précision et lier mes interventions à celles de mes interlocuteurs.</p>	<p>Je peux participer sans effort à toute conversation ou discussion et je suis aussi très à l'aise avec les expressions idiomatiques et les tournures courantes. Je peux m'exprimer couramment et exprimer avec précision de fines nuances de sens. En cas de difficulté, je peux faire marche arrière pour y remédier avec assez d'habileté pour que cela passe inaperçu.</p>
<p>Je peux m'exprimer de façon claire et détaillée sur une grande gamme de sujets relatifs à mes centres d'intérêt. Je peux développer un point de vue sur un sujet d'actualité et expliquer les avantages et les inconvénients de différentes possibilités.</p>	<p>Je peux présenter des descriptions claires et détaillées de sujets complexes, en intégrant des thèmes qui leur sont liés, en développant certains points et en terminant mon intervention de façon appropriée.</p>	<p>Je peux présenter une description ou une argumentation claire et fluide dans un style adapté au contexte, construire une présentation de façon logique et aider mon auditeur à remarquer et à se rappeler les points importants.</p>

Niveaux communs de compétence. Grille pour l'auto-évaluation (CECR, 2000 : 27)

**Annexe D : Exemple du questionnaire donné aux participants de
l'étude 2 (Auto-évaluation formative)**

Prénom :

Nationalité :

Age :

Essayez de répondre aux questions en faisant des phrases

1. Avez-vous fait de la musique ? Si oui, pendant combien de temps ?

.....
.....

2. Quelles langues parlez-vous ?

.....
.....

3. Quand avez-vous commencé à apprendre le français ?

.....
.....

4. Suivez-vous des cours de français actuellement (en dehors des cours de prononciation) ? Si oui, dans quel cadre (Alliance Française, Institut Catholique, association de FLE)

.....
.....

5. Comment avez-vous appris le français ? A l'oral ou à l'écrit ?

6. Ces méthodologies vous convenaient-elles ? Quels étaient selon vous leurs avantages et leurs inconvénients ?

7. Avez-vous déjà suivi des cours de prononciation? Si oui, de quels types (laboratoire de langue, cours de phonétique *etc.*)

8. Quelles sont vos principales difficultés en français ?

9. Sur une échelle de 1 à 10 où situez-vous votre niveau de français ? (1 : débutant complet ; 10 : expert)

1 2 3 4 5 6 7 8 9 10

10. Est-ce la première fois que vous venez en France ?-----

11. Remarques éventuelles

**Annexe E : Exemple du questionnaire bilan donné aux participants
de l'étude 2 (Auto-évaluation sommative)**

Prénom :

Nationalité :

Essayez de répondre aux questions en faisant des phrases

1. Quelles sont vos habitudes de lecture dans votre langue maternelle ?

2. Quelles sont vos habitudes de lecture en français ?

3. Que pensez-vous des cours de phonétique que vous avez reçus?

4. Avez-vous le sentiment de mieux parler français? Pourquoi ?

5. Avez-vous le sentiment de lire plus facilement? Pourquoi ?

6. Avez-vous le sentiment de comprendre plus facilement? Pourquoi ?

7. Sur une échelle de 1 à 10 où situez-vous votre niveau de français ? (1 : débutant complet ; 10 : expert)

1 2 3 4 5 6 7 8 9 10

8. Remarques éventuelles

**Annexe F : Textes utilisés pour mesurer la fluence en lecture oralisée
(étude 2)**

Textes proposés au niveau débutant (D)

Texte 1D

Les jeunes et les gratuits

Les journaux quotidiens sont chers et les jeunes ont l'habitude d'accéder gratuitement à l'information, à la télévision, à la radio ou sur Internet. Les journaux gratuits ont beaucoup de succès auprès de la jeune génération: 40% du lectorat de *20 Minutes* et 49% de celui de *Métro* est âgé de 15 à 34 ans.

Les gratuits plaisent aux jeunes parce qu'ils ont un format pratique et contiennent des informations courtes. De plus, les gratuits sont distribués à la sortie du métro ou aux arrêts de bus. Les gens n'ont plus besoin d'aller acheter leur journal dans des kiosques ou de prendre un abonnement.

Questions de compréhension texte 1 D:

- -De quoi parle ce texte ?
- Quel est l'âge moyen des lecteurs des journaux gratuits ?
- Pourquoi les gratuits ont autant de succès ?

Texte 2D

Étudier après le Bac

En France, après le Bac*, les jeunes ont plusieurs possibilités pour faire leurs études. Beaucoup vont à l'université pour préparer une licence (6 semestres), puis éventuellement un master (4 semestres après la licence) et enfin un doctorat (6 semestres après le master). D'autres préfèrent des formations plus courtes et essaient d'entrer dans les IUT (Institut Universitaire de Technologie, 2 années d'études) ou préparent un BTS (Brevet de technicien supérieur, 2 ans). Enfin, certains s'inscrivent dans des classes préparatoires aux grandes écoles pour préparer les concours d'entrée aux écoles de commerce et d'ingénieur : HEC, Polytechnique Ecole des Mines, etc.

Il est aussi possible de préparer de nombreux diplômes à distance, et, quand on travaille, on peut bénéficier de la formation continue pour compléter ses connaissances ou se réorienter professionnellement tout au long de sa carrière.

* *school leaving certificate*

Questions de compréhension texte 2 D:

- -De quoi parle ce texte ?

- Que fait la majorité des jeunes après le bac ?
- - Peut-on se former toute la vie en France ?

Texte 3 D

Les Français et la télévision

La télévision occupe une place importante dans la vie quotidienne des français.

Vincent 36 ans, commercial, marié, deux enfants, témoigne :

« Chez nous, toute la famille regarde la télé mais rarement les mêmes programmes. Le matin, à 7h30, les enfants prennent le petit-déjeuner devant le petit écran : ils regardent une émission pour les jeunes jusqu'à 8 heures, et ma femme et moi nous nous préparons tranquillement. Après l'école, les enfants se reposent une demi-heure devant la télé avec les dessins animés et, le soir, ils se couchent tôt. Avec ma femme, on regarde toujours le 20 heures. Après, le choix est difficile parce que moi, je suis un fanatique du foot et ma femme, elle préfère regarder les magazines de société ou les comédies romantiques avec Richard Gere. »

Questions de compréhension texte 3 D:

- De quoi parle ce texte ?
- Qui regarde la télévision ?
- Est-ce que toutes les personnes regardent les mêmes programmes.

Textes proposés au niveau avancé (A)

Texte 1 A

Les jeunes et les gratuits

Les quotidiens sont chers et le budget des jeunes est sollicité par des dépenses croissantes et nouvelles, par exemple, dans la téléphonie, les jeux, etc. Or, en matière d'information, les jeunes sont habitués à y accéder gratuitement, à la télévision, à la radio ou sur Internet. D'où le succès des journaux gratuits auprès de la jeune génération. Les chiffres sont impitoyables : 40% du lectorat de *20 Minutes* et 49% de celui de *Métro* sont âgés de 15 à 34 ans.

Les gratuits conquièrent les jeunes grâce à leur format pratique et leur contenu. Ils traitent l'actualité sur un modèle peu répandu jusqu'alors : infos brèves, factuelles, sans parti pris ni recul. Le nombre de points de vente des journaux ne cesse de diminuer, car l'essentiel de la diffusion des gratuits d'information se réalise dans, ou près des lieux de transit, à la sortie du métro ou aux arrêts de bus. Ils ont inventé un modèle de diffusion : aller vers le lecteur au lieu d'attendre qu'il vienne à eux dans les kiosques ou qu'il prenne un abonnement. Grâce aux gratuits, les jeunes retrouveraient même le goût de la lecture. En effet, la presse gratuite se remet au niveau des gens. Les gratuits répondent de plus en plus aux attentes des lecteurs, par des sujets les concernant, des maquettes dynamiques et colorées et, ils sont distribués dans des lieux de diffusion étudiés.

La presse gratuite est par contre toujours mal acceptée par les journaux payants parce qu'elle représente un concurrent de taille. Un quotidien créerait pourtant un gratuit du soir très prochainement.

Questions de compréhension texte 1 A :

- Quel est le thème de ce texte ?
- Quel est le public visé par les gratuits ?
- Pourquoi sont-ils si appréciés par la jeune génération ?
- Qu'en pensent les journaux payants et pourquoi ?

Texte 2 A

Étudier après le bac

En France, après le bac, les jeunes ont plusieurs possibilités : l'université, pour préparer une licence (6 semestres), puis éventuellement un master (4 semestres après la licence) et enfin un doctorat (6 semestres après le master), les IUT (Institut Universitaire de Technologie, 2 années d'études), les BTS (Brevet de Technicien Supérieur) ou les classes préparatoires aux Grandes Ecoles : HEC, Polytechnique Ecole des Mines, etc.

On aurait tort de croire que les « meilleurs élèves » choisissent tous les élitistes classes préparatoires pour intégrer les grandes écoles. La moitié des diplômés sont passés par les bancs de l'université, par un IUT ou un BTS et ont intégré les grandes écoles via les "admissions parallèles", des concours qui ouvrent de plus en plus de places au fil des ans, mais qui restent sélectifs, car de plus en plus convoités.

On peut aussi préférer voyager et dans ces cas, il est également possible de préparer de nombreux diplômes à distance. Enfin, une fois rentré dans la vie active, on peut se aussi tout au long de sa carrière, c'est ce qu'on appelle la formation continue.

Questions de compréhension texte 2A :

- Quel est le thème de ce texte ?
- Quelles sont les filières préférées des français ?
- Est-il nécessaire de faire une classe préparatoire pour intégrer une grande école ?
- Existe-il d'autres types de formation ?

Texte 3 A

Les Français et la télévision

La télévision occupe une place importante dans la vie quotidienne des français.

Vincent 36 ans, commercial, marié, deux enfants, témoigne :

« Chez nous, c'est un rituel : toute la famille regarde la télé mais par contre, nous regardons rarement les mêmes programmes. Le matin, à 7h30, les enfants prennent le petit-déjeuner devant le petit écran : ils regardent une émission pour les jeunes jusqu'à 8 heures pendant que ma femme et moi nous nous préparons tranquillement.

En rentrant de l'école, les enfants se reposent une demi-heure devant la télé avec les dessins animés et, le soir, ils se couchent tôt. Nous, avec ma femme, on regarde toujours le 20 heures. Après, le choix est difficile parce que moi, je suis un fanatique du foot et ma femme, elle préfère regarder les magazines de société ou les comédies romantiques avec Richard Gere ! »

Tout comme Vincent, la quasi-totalité des personnes de 15 ans ou plus regarde la télévision au moins une fois par jour. Ils ne sont que 5% à la regarder seulement une à deux fois par semaine. La durée moyenne d'écoute va de 2 heures par jour chez les 15-19 ans à 3H40 chez les plus de 75 ans. C'est la seule pratique de loisir qui diminue quand les autres augmentent : les personnes qui font très souvent du sport ou des activités culturelles regardent la télévision moins de deux heures par jour !

Question de compréhension texte 3 A

- Quel est le thème de ce texte ?
- Quels programmes regardent les membres de cette famille ?
- Quelle est la durée moyenne d'écoute chez des gens de cette âge ?
- Est-ce que cette famille regarde bcp la télé ?

Annexe G : Exemple du questionnaire donné aux auditeurs pour le test de perception (étude 2)

Prénom :

Âge :

Niveau d'études :

Problèmes auditifs ? Oui Non

Si oui, précisez-----

Pratique musicales (précisez l'instrument et le nombre d'années de pratique) : -----

Langue(s) parlée(s) (pour chacune, précisez votre niveau)

Expérience(s) à l'étranger supérieure à une durée de 3mois

Expérience(s) dans l'enseignement du FLES (pour chaque expérience, précisez, le nombre d'heures de cours, le type de cours, les groupes linguistiques concernés et le niveau des apprenants)

À compléter à l'issue du test

Précisez le ou les critères qui vous ont permis d'évaluer le niveau de fluence des productions de parole lue des apprenants.

Remarques éventuelles

Alazard, C., Astésano, C. & Billières, M. (2012). MULTIPHONIA: a MULTImodal database of PHONeticis teaching methods in classroom InterActions. *Paper presented at the 8th Conference on Language Ressources and Evaluation, Istanbul.*

MULTIPHONIA: a MULTImodal database of PHONeticis teaching methods in classroom InterActions.

Charlotte Alazard, Corine Astésano, Michel Billières

U.R.I Octogone-Lordat, EA 4156, Université Toulouse II

5 allées Antonio-Machado, F-31058 Toulouse Cedex 1

E-mail: alazard@univ-tlse2.fr, astesano@univ-tlse2.fr, billieres@univ-tlse2.fr

Abstract

The *Multiphonia Corpus* consists of audio-video classroom recordings comparing two methods of phonetic correction (the 'traditional' articulatory method, and the Verbo-Tonal Method) This database was created not only to remedy the crucial lack of information and pedagogical resources on teaching pronunciation but also to test the benefit of VTM on Second Language pronunciation. The VTM method emphasizes the role of prosody cues as vectors of second language acquisition of the phonemic system. This method also provides various and unusual procedures including facilitating gestures in order to work on spotting and assimilating the target language prosodic system (rhythm, accentuation, intonation). In doing so, speech rhythm is apprehended in correlation with body/gestural rhythm. The student is thus encouraged to associate gestures activating the motor memory at play during the repetition of target words or phrases. In turn, pedagogical gestures have an impact on second language lexical items' recollection (Allen, 1995; Tellier, 2008). Ultimately, this large corpus (96 hours of class sessions' recordings) will be made available to the scientific community, with several layers of annotations available for the study of segmental, prosodic and gestural aspects of L2 speech.

Keywords: database, multimodality, phonetic correction, second language acquisition.

1. Introduction

Prosody and multimodality are not only the key to language acquisition but also necessary and irreplaceable in everyday communication (Di Cristo, 2004; Kendon, 2004 ; Mac Neill, 2005). The number of studies in L2 prosody is however rather limited compared to the amount of work carried out on L2 segmental aspects.

The lack of database in didactics, particularly in French as a Second Language (hereafter FSL), could explain the lack of experimental researches in this field. Even if more focus is now put on communication in foreign language teaching methods, some main aspects of oral communication, such as phonetics and prosody, remain remarkably left out

The aim of our research is thus to provide a multimodal database of real classroom interactions in FSL. It is also aimed at confronting theoretical predictions and real class situations, in order to favour phonetics teaching in foreign language courses.

It is expected by the *Common European Framework of Reference for Languages* (hereafter, *CEFR*) that advanced level students (level B) should have 'a clear and natural intonation' and read with fluency. No mention is made of pronunciation training and oral skills' mastering. It is as if good fluency and prosody in L2 came naturally to advanced L2 students. Experience in teaching L2 however contradicts this view, insofar as advanced students still transfer the prosodic characteristics of their L1 onto the L2, in both unscripted and read speech. In other words, foreign accent remains persistent at an advanced level.

This apparent contradiction has various roots: first of all, even if recent researches have shown the importance of phonetic training in the improvement of speaking fluency in spontaneous and read speaking skills (Freed, 1995; Freed *et al.*, 2004 ; Alazard *et al.*, 2009, 2010), the

idea that pronunciation will improve naturally thanks to mere repeated contacts with the foreign language is persistent. Secondly, in spite of the recognized role of prosody in both first and second language acquisition (Di Cristo, 2004), L2 traditional teaching methods - in the rare cases where pronunciation is taken into account - focus exclusively on the segmental level. It is worth noticing at this point that L2 teachers very rarely perform phonetic correction in their classroom activities, due to their lack of expertise in this discipline. The emphasis is rather on grammatical and lexical aspects of the target language; very rare moments are dedicated in L2 classes to phonetics and segmental pronunciation correction. Finally, and because of this lack of phonetics teaching practice in the classroom, phonetics correction methods have never been experimentally tested or validated.

In order to confront this last point, we propose to question and to experimentally test two different pronunciation teaching methods - the Articulatory Method (hereafter AM) and the Verbo-Tonal Method (hereafter VTM) - and to make these teaching-methods recordings available for researchers through our database.

According to AM, by far the most widespread method, a good production implies the metalinguistic knowledge of how we articulate sounds. Thus, the teacher will provide an articulatory description of the different sounds of the foreign language, then prompt the student to repeat the correct articulatory gestures in order to produce the target sound. For example, to produce [u] the teacher will tell the student to place the tongue at the back of the mouth and to round the lips, in opposition to the fronted [i] for which the lips should be stretched. In this method, the emphasis will first be put on the production and repetition of isolated sounds, then isolated words containing the target sound and finally sentences. There will be no real focus on prosodic parameters such as

rhythm and intonation. The AM thus focuses on explicit learning of phonetic articulatory gestures.

The VTM, on the contrary, uses the prosodic structure of the target language as the 'shell' for pronunciation skills' improvement. More specifically, the rhythmic pattern of the target language is used to bring to light the phonetic specificities of the target language. The teacher first helps the learners familiarize themselves with the prosodic structure of the target language through the repetition of prosodic patterns using logatoms (/dadada/) or the use of facilitating gestures (for example rising hand movement for salient syllables). In a second phase, the prosodic structure is used to facilitate phoneme perception and re-production, on the basis that there is a phonological loop between the production and the perception of phonetic features. (Borrell, 1996) For example, if the learner darkens the timbre of a target phoneme, the teacher will pronounce the phoneme in a prosodically brightening context (accented syllable) and have the learner repeat it in the same context. Namely, a facilitating production context will help the learner perceive the proper phonemic features of the target language and thus help them correctly re-produce these features in any other prosodic contexts (Billières, 2005). The VTM thus focuses on non-explicit prosodic learning. Despite extremely positive results both in didactics and speech therapy, teachers are wary of this method as it implies a different teaching approach and an expertise in phonetics and prosody. Furthermore, because this method remains confidential to a small group of international experts, its validity remains to be demonstrated to a larger audience in order to, one day, be included in comprehensive L2 teaching methods.

The originality of our database is thus to record and compare for the first time these two different methods in an ecological classroom situation, and to propose an enrichment of this database for future L2 researches.

2. Collection of the database

The database consists of a longitudinal recording of classroom teaching of phonetics over eight weeks, with twenty participants, all English Speakers (15 female; mean age: 32; age range: 20-60). An oral interview allowed us to evaluate their level in French according to the CEFRL: ten of the participants were judged to have an elementary level in French (level A) and ten were judged to have an advanced level in French (level B).

The participants were equally divided into four groups: two groups per method according to their level. Each group received two pronunciation trainings per week - lasting one hour and a half each - for eight weeks.

Both methods were taught by the same teacher – the first author - and recorded in the same experimental conditions. All class sessions were recorded in the professional audio-visual recording studio of the Direction of Information Technology and Communication for Teaching service (DTICE) of the University of Toulouse II.

Depending on the methodological approach, the

classroom stage was reorganized as follows: for the AM classes, the participants were sitting around an U-shaped table while the teacher was explaining articulatory features, using oro-facial and vocal tract gestures or diagrams (Figure1). The participants were then asked to repeat one by one the target sound, presented in isolation or in single words. During the second part of the class, they would listen to a dialogue or an authentic document, before answering a few questions on the audio documents.



Figure 1: Example of an AM classroom layout.

Participants are sitting at a table around the teacher who uses diagrams and meta-discourse to describe the articulation of the different phonemes.

For the VTM classes, the participants were sitting around the teacher, with no table, while the teacher used prosody and multimodality to help participants perceive the target sound and the prosodic features of the target language (Figure2). Participants were free to reproduce or not the hands movement (following the variation of intonation) made by the teacher, while repeating the sentences or the short dialogues. The second part of the course would be identical in both methods.



Figure 2: Example of a VTM classroom layout, with no table between the participants and emphasis on body language.

These different set-ups imply a specific technical organization of the classroom stage. The studio combines three video cameras (BVP50 Sony) – an overall shot of the teacher, two overall shots of the students – and six microphones (half-track AKG hanging down from the ceiling) – one microphone for the teacher and the other 5 for the students.

The classes were recorded in an artificial lighting: one spotlight of 800 w for the reserve angle on the teacher, 4 spotlight of 800w for the students and one

spotlight for the backlighting.

The stage controller used a Panasonic video mixer, a Tascam eight tracks audio mixer, a TL 12 Coyllins light controller and dvcam 4/3 DSR 45 Sony of 184mm for each class.

We used dvcams for the production rushes and avid media-composer for the post-production. The data were transferred onto dvd for the trimming. In order to be used on the Internet, the masters were encoded with Adobe cs5.

The database is constituted of ninety-six hours (3h/week*4groups*8 weeks) of classroom recording.

3. Enrichment of the database

This multimodal database constitutes an important resource for Second Language Acquisition's (hereafter SLA) researchers.

Hence, MULTIPHONIA will be enriched at many different levels, to allow for segmental, prosodic, morphosyntactic, syntactic, lexical and gestural analyses of L2 speech. The different levels of automatic annotation will be done on short excerpts of the database at first, and confronted to manual annotations by three experts, before extending the procedure to larger parts of the corpus (see section 4- below).

The automatic annotation of such a corpus, consisting of interactional speech in a classroom environment and recorded with several multi-directional microphones on a single sound-track, represents an interesting challenge for the automatic tools' developers to test for the portability of their tools to more challenging speech corpora.

3.1 Transcription (cf. Bertrand et al, 2008).

3.1.1 Segmentation in Interpausal Units

Before any annotation, some significant audio extracts of the recording will be automatically segmented in Interpausal Unit (hereafter IPU). IPU are constituted of blocks of speech separated by 200 ms silent pauses. The IPU segmentation has been commonly used for large corpora as it then facilitates sound and transcription alignment

3.1.2 Enriched Orthographic Transcription (EOT)

The advantage of the EOT is to provide an orthographic transcription as well as specifying natural speech production phenomena such as pauses, false starts or repetitions. According to the EOT transcription convention, the transcribers will be asked to annotate silent pauses, filled pauses, elisions, false starts, word truncation, liaisons (absence of a standard liaison, presence of an unusual liaison), assimilation phenomena and specific phenomena, such as, in our case, deviant pronunciations or code switching from L2 to L1 (those phenomena will be labelled as 'interlanguage phenomena') (see Bertrand et al, 2008 for transcription conventions).

Two transcriptions will then be automatically generated from the EOT.

First, a standard orthographic transcription from

which the orthographic tokens are extracted for semantics, syntax or discourse analyses (see Blache et al. 2009 for example)

Second, a selective transcription from which the phonetic tokens are extracted for grapheme-phoneme conversion.

3.2 Phonetic annotations

3.2.1 Phonetization

The phonetic annotations will be done with the Speech Phonetization Alignment and Syllabification (SPPAS) tool (Bigi and Hirst, 2012) on the significant extracts. The aim of this tool is to provide automatic utterance, words syllables and phonemes segmentations annotations from a speech recording and its transcription.

SPPAS produces a phonetic transcription based on a phonetic dictionary. The program offers the possibility to select (automatically or manually) among all the phonemics variants that are proposed.

3.2.2 Alignment

The phonetic alignment will consist of an automatic temporal matching between a speech utterance and its phonetic representation. The alignment will be done in the frame of each IUP, to maximize the alignment.

In order to evaluate the errors rate of the aligner the automatic alignment will be compare with the manual alignment of two experts.

3.2.2 Syllabification

The syllabification will be done according to two main principles: (1) a syllable contains only one vowel and (2) a pause signals a syllable boundary, as previously described in Bigi *et al.* (2010).

3.2.3 Disfluencies

Disfluencies can be prosodic (abnormal lengthening, agrammatical silent and filled pauses, truncated speech runs, etc.) or lexicalized (word or phrase truncation, repetitions, etc.). The prosodic ruptures of the speech flow will be annotated according to the quantitative measures detailed in Kormos (2006) while, for the lexicalized disfluencies, we will annotate three distinctive parts of the disfluencie: the Reparandum (that precedes the interruption point), the Break Interval (optional event between the Reparandum and the Reparans) and the Reparans (the part following the break) (Blanche-Benveniste, 1987).

3.3 Prosodic annotations

The annotation of prosody is very complex and cannot yet be done automatically. Bertrand *et al.* (2007) & Blache *et al.* (2009) propose to distinguish different levels of relevant prosodic information to annotate in order to account for discourse prosodic structuring:

- Prosodic phrasing, with two hierarchical levels consisting of the Intonational Phrase (IP) and the Accentual Phrase (AP). These units are marked by a Final

Accent and a (potential) Initial Accent.

Intonation contours, which can be characterized as follows: minor contour (m), major continuative rising contour (RMC), major list rising contour (RL), major falling contour (F), major terminal rising contour (RT), major questioning rising contour (RQ), rising-falling contour (RF1), falling from the penultimate contour (RF2), no melodic variation (f1)

3.4 Gestural annotations

The annotations of gestures can be done manually with the ELAN or ANVIL softwares, according to Mac Neill's typology (MacNeill, 2005): metaphoric (gestures representing an abstract idea), iconic (gestures representing an action or concrete object), deictic (pointing gestures) and beating (gestures accompanying rhythm) gestures, are differentiated.

For the purpose of our own research on L2 (see section 4- below), significant extracts of the teacher's and students' gestures in the two teaching classroom environments will be annotated and compared. We will annotate interactive gestures (gestures addressed to the interlocutor in order to manage the interaction) and aborted gestures (half-made gestures), quite typical of L2 speech interaction (Tellier and Stam, 2010). Head movements, body position, gaze directions and facial expression, will be encoded as well as hand movements.

4. First steps towards the exploitation of the MULTIPHONIA database

4.1 Prosody supporting L2 segmental perception and production

At the segmental level, we are planning on measuring the influence of prosody on the acquisition of the phonemic features. The VTM hypothesizes indeed that the prosodic structure will facilitate phoneme perception and thus will also facilitate phoneme production.

In order to test this hypothesis, we will measure formants' repartition of L2 vowels in different repetition contexts and at different stages of the training. We will extract significant audio sequences that will be annotated manually by three experts using the coding elaborated in Bertrand *et al.* (2008) with the PRAAT software (Boersma, & Weenink, 2005), as described in section 3. The annotation will then be automatically phonetized and aligned with the audio extracts using the aligner elaborated by Bigi & Hirst (2012), to facilitate automatic vowel detection.

4.2 Gestural impact on prosodic characteristics' learning

Two different phonetics teaching methods imply two different uses of pedagogical gestures. The AM puts the accent on the gestures of the articulators only, implying central or peripheral gestural spaces, whereas the VTM focuses on prosodic guiding gestures, implying peripheral and upper gestural spaces (Figure 3).

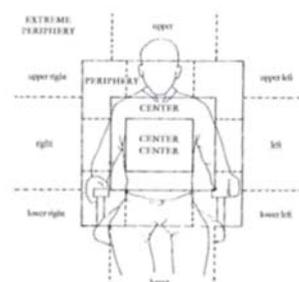


Figure 3: The gestural space (McNeil, 1992)

More specifically, with the VTM, the teacher helps apprehending speech rhythm in correlation with body/gestural rhythm, as well as perceiving prosodically salient syllables and segmenting the speech flow (Billières, 2002). The teacher thus helps developing the perception of 'rhythmic phrases' through the perception of rhythmic prominences (Initial Accent (IA) and Final Accent (FA)). Rhythmic saliences are seen as anchor points for the speech flow segmentation into smaller units (3 to 4 syllables), necessary for ultimate processing in the working memory.

In parallel to the pedagogic work on the verbal continuum, the student is thus encouraged to associate gestures activating the motor memory at play during the repetition of target words or phrases. Indeed, research in cognitive psychology has demonstrated the impact of the motor modality on sentence recall (Cohen & Otterbein, 1992). By the same token, Second Language Acquisition studies have shown that pedagogical gestures have an impact on second language lexical items' recollection (Allen, 1995; Tellier, 2008).

Because prosodic cues help access the lexicon and segment the speech flow in the native language but are poorly explored in a second language (Snijders *et al.*, 2007), we wish to experimentally demonstrate that pedagogical gestures have a facilitating impact on the reproduction and memorization of relevant non native prosodic cues. This will be achieved through the systematic analysis of SL French words or phrases repeated by the English learners with the use of both gestural and proper French accentual patterns throughout the eight weeks' classes in the VTM group. The basic accentual pattern for French in the Accentual Phrase (AP) is Initial Accent and Final Accent (hereafter /IA-FA/; see Di Cristo, 2000, and Jun & Fougeron, 2002, for a description). We hypothesize that gestures will help anchor the /IA-FA/ prosodic pattern, and help correct accurate realization of the target language accentual and segmental systems.

The annotation of gestures – specifically beating metaphors and interactive gestures - using the coding elaborated in Blache *et al.* (2010) will be done with the ELAN software.

5. Conclusion

This paper presents a MULTImodal database of PHONetics teaching methods in classroom InterActions (MULTIPHONIA), consisting of 96 hours of audio-video classroom recordings of two methods of phonetic correction (the 'traditional' articulatory method, and the Verbo-Tonal Method).

If this database primarily constitutes a rich supply for pedagogical resources, it also provides multimodal resources for SLA researchers interested in various aspects of L2 learning, via the annotation of segmental, prosodic, morphosyntactic, syntactic, lexical and gestural levels of L2 interactional speech.

Annotated extracts of this MULTIPHONIA database are going to be shortly available on line.

(http://crdo.up.univ-aix.fr/voir_depot.php?lang=fr&id=000780).

6. Acknowledgements

The authors wish to thank the DTICE and especially Mr Bruno Bastard, the audio-visual technician of the DTICE, for his help in the realization of this database

7. References

- Alazard, C., Astésano, C., and Billières. M. (2009). "Rôle de la prosodie dans la structuration du discours - Proposition d'une méthodologie d'enseignement de l'oral vers l'écrit en Français Langue Etrangère-". *Proceedings of the Interfance Discours and Prosodie conference 2009* (Paris, France)
http://makino.linguist.jussieu.fr/id09/actes_fr.html
- Alazard, C., Astésano, C., and Billières. M. (2010). "The Implicit Prosody Hypothesis applied to Foreign Language Learning: From oral abilities to reading skills". *Proceedings of the 5th Speech Prosody 2010* (Chicago, USA).
<http://www.speechprosody2010.illinois.edu/papers/100648.pdf>
- Allen, L. (1995). The effects of emblematic gestures on the development and access of mental representations of French expressions. *The Modern Language Journal* 79: 521-529.
- Bigi, B., and Hirst, D. (2012). "Speech Phonetization Alignment and Syllabification: A tool for the automatic analysis of Speech Prosody". *Proceedings of the 6th Speech Prosody 2012* (Shanghai, China).
- Billières, M. (1993). Théorie et pratique du rythme parolier en phonétique corrective. *Cahiers du Centre Interdisciplinaire des Sciences du Langage* 9: 3-32.
- Billières, M. (2002). Le corps en phonétique corrective. In R. Renard (Ed.), *Apprentissage d'une langue étrangère/seconde. La Phonétique verbo-tonale*. Bruxelles: De Boeck Université, pp. 37-70.
- Billières, M. (2005). Les pratiques du verbo-tonal. Retour aux sources. Berré, M (Eds), *Linguistique de la parole et apprentissage des langues. Questions autour de la méthode verbo-tonale de P. Guberina*. Centre International de Phonétique Appliquée, Mons, pp. 67-87.
- Blache P., Bertrand R. and Ferré, G. (2009). Creating and Exploiting Multimodal Annotated Corpora: The ToMA. In M. Kipp et al. (Eds): *Multimodal Corpora*, LNAI 5509, pp. 38-53.
- Blache P., Bertrand R., Bigi B., Bruno E., Cela E., Espesser R., Ferré G., Guardiola M., Hirst D., Magro E. -P., Martin J. -C., Meunier C., Morel M. -A., Muriasco E., Nesterenko I., Nocéra P., Pallaud B., Prévot L., Priego-Valverde B., Seinturier J., Tan N., Tellier M. and Rauzy, S. (2010). 'Multimodal Annotation of Conversational Data'. *Proceedings of Linguistic Annotation Workshop 2010* (Uppsala, Sweden), pp. 186-191.
- Blanche-Benveniste C. (1987). "Syntaxe, choix du lexique et lieu de bafouillage". In *DRLAV* 36-37
- Bertrand, R., Blache, P., Espesser, R., Ferré, G., Meunier, C., Priego-Valverde, B. and S. Rauzy. (2008). Le CID – Corpus of Interactional Data. *TAL*, 49 (3), pp105-134.
- Boersma, P. and Weenink, D. (2005). Praat: doing phonetics by computer, <http://www.praat.org>
- Borrell, A. (1996). Parallèle entre perception et production: complexité du lien entre reconnaissance et production des unités phonético-phonologique. *La linguistique* 32 (2)
- Cohen, R. L. and Otterbein, N. (1992). The mnemonic Effect of speech Gestures: Pantomimic and Non-Pantomimic Gestures compared. *European Journal of Cognitive Psychology*, 4 (2), pp 113-139.
- Di Cristo, A. (2000). Vers une modélisation de l'accentuation en français. Deuxième partie : le modèle. *Journal of French Language Studies*, 10, pp. 27-44.
- Di Cristo, A. (2004). La prosodie au Carrefour de la phonétique, de la phonologie et de l'articulation formes-fonctions. *Travaux Interdisciplinaires du Laboratoire Parole & Langage* 23, pp. 67-211.
- Freed, B. F. (1995). What Makes Us Think that Students Who Study Abroad Become Fluent? In B. F. Freed (Eds), *Second Language Acquisition in a Study Abroad Context*, pp. 123-145.
- Freed, B. F., Segalowitz, N., and Dewey, D. (2004). Contexts of learning and second language fluency in French: Comparing regular classrooms, study abroad, and intensive domestic programs. *Second Language Acquisition*, 26, pp. 275-301.
- Jun, S. -A. and Fougeron C. (2002). Realizations of accentual phrase in French intonation. *Probus* 14, pp147-172.
- Kendon, A. (2004). *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.
- Kormos, J. (2006). *Speech production and Second Language Acquisition*. New York, London: Routledge.
- McNeill, David (2005). *Gesture and Thought*. Chicago: Chicago University Press.
- Snijders, T. M., Kooijman, V., Cutler, A., and Hagoort, P. (2007). Neurophysiological evidence of delayed segmentation in a foreign language. *Brain Research*, 1178, pp. 106-113.
- Tellier, M. (2008). The effect of gestures on second language memorisation by young children. In Gullberg, M., & de

- Bot, K. (Eds.) *Special issue Gestures in language development. Gesture*, 8(2), pp. 219-235.
- Tellier, M and Stam, G. (2010). Découvrir le pouvoir de ses mains: La gestuelle des futurs enseignants de langue. Proceedings of the conference *Specificités et diversités des interaction didactiques: disciplines, finalités, contextes*, 2010 (Lyon, France)