



**HAL**  
open science

# Production orale de l'anglais L2 des francophones : une approche rythmique guidée de l'apprentissage peut-elle améliorer leur performance ? Étude de cas

Francis Picavet

## ► To cite this version:

Francis Picavet. Production orale de l'anglais L2 des francophones : une approche rythmique guidée de l'apprentissage peut-elle améliorer leur performance ? Étude de cas. Sciences de l'Homme et Société. 2012. dumas-00752578

**HAL Id: dumas-00752578**

**<https://dumas.ccsd.cnrs.fr/dumas-00752578v1>**

Submitted on 16 Nov 2012

**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.



**Production orale de l'anglais L2 des francophones : une approche rythmique guidée de l'apprentissage peut-elle améliorer leur performance ?**

-----  
**Étude de cas**

Nom : **PICAVET**  
Prénom : **Francis**

UFR :

**Département sciences du langage et français langue étrangère**

---

**Mémoire de master 2 recherche - 30 crédits – Mention : Sciences du langage**

Spécialité ou Parcours : **Langage - parole - variations - surdit **

Sous la direction de : **Mmes V ronique AUBERG  et Solange ROSSATO**

**Ann e universitaire 2011-2012**





**Production orale de l'anglais L2 des  
francophones : une approche rythmique  
guidée de l'apprentissage peut-elle améliorer  
leur performance ?**

-----  
**Étude de cas**

Nom : **PICAVET**  
Prénom : **Francis**

UFR :

**Département sciences du langage et français langue étrangère**

---

**Mémoire de master 2 recherche - 30 crédits – Mention : Sciences du langage**

Spécialité ou Parcours : **Langage - parole - variations - surdit **

Sous la direction de : **Mmes V ronique AUBERG  et Solange ROSSATO**

**Ann e universitaire 2011-2012**



## Remerciements

He had caught a great cold,  
Had he no other clothes to wear  
Than the skin of a bear  
Not yet killed.

*Thomas Fuller, 1608 – 1661*

Reprendre des études à l'âge de cinquante ans passés est un défi que certaines personnes se lancent ; d'autres considèrent cela comme une folie ; ont-ils raison ? Se retrouver sur les bancs d'une université, entouré de jeunes gens à l'esprit vif et à la mémoire exercée, est sans doute un défi encore plus grand pour un professeur qui a plutôt l'habitude d'être « de l'autre côté du bureau ». Avec humilité, beaucoup de courage et, surtout, le soutien de son entourage, l'expérience vaut vraiment la peine d'être tentée, me semble-t-il. Pourtant, le défi le plus important n'est pas là. Si nous ajoutons l'étiquette « recherche » à nos études, alors les choses peuvent prendre une toute autre dimension, à la fois fascinante et vertigineuse. L'important est, avant tout, de rencontrer les bonnes personnes.

Parmi toutes celles qui m'ont apporté leur aide, leurs conseils et leur soutien dans la préparation de ce mémoire, je tiens à remercier en premier lieu Solange Rossato, qui a bien voulu consacrer du temps aux idées de recherche, sans doute un peu saugrenues, que je proposais et qui, tant en cours qu'en séances de travail, m'a aidé à clarifier et à formaliser mes intuitions, tout en me faisant bénéficier de ses grandes compétences et de son expérience dans le traitement et l'analyse des sons de parole.

C'est elle qui m'a rapidement présenté à Véronique Aubergé, que je remercie chaleureusement aussi, pour toutes les heures passées à essayer de transformer le néophyte maladroit que j'étais en apprenti-chercheur efficace et rigoureux, me faisant profiter de son expertise dans le domaine complexe et si passionnant de la prosodie.

Sur leur proposition, j'ai eu l'occasion de participer à la conférence ML2P – *Methods in Second Language (L2) Prosody* de Padoue, en Italie, au printemps dernier, qui fut ma première participation à une rencontre de chercheurs dans le domaine de la prosodie ; ce fut un grand moment de partage, tant avec les chercheurs présents qu'avec les autres étudiants, à la fois participants et organisateurs.

Bien entendu, ce travail ne saurait exister sans la participation active et sympathique de mes étudiants de l'École nationale supérieure de l'énergie, de l'eau et l'environnement (Ense3), école d'ingénieurs de l'Institut polytechnique de Grenoble. L'intégralité de mes soixante-quinze étudiants du second semestre 2011-2012 s'est prêtée avec enthousiasme à la construction d'un corpus important, susceptible de servir de base à de futures recherches. Je leur en suis très reconnaissant, car sans eux, ce travail n'existerait pas. J'en profiterai également ici pour remercier l'administration de l'Ense3, qui m'a soutenue dans cette démarche de reprise d'études, notamment Olivier Métais, son Directeur, et Sylvie Rolland, sa Directrice administrative.

Ma participation aux RJCP 2011, 9èmes Rencontres des jeunes chercheurs en parole à Grenoble, fut un moment riche de rencontres et découvertes. Parmi les jeunes chercheurs rencontrés, les conseils que j'ai pu avoir de Sandra Cornaz, coordinatrice des RJCP et doctorante, intéressée tout comme moi par les rapports entre langue, chant et musique, et les approches didactiques qui peuvent en découler, m'ont été des plus précieux.

Que toutes les personnes qui, dans un esprit de partage des savoirs, des méthodes et des outils, mettent des ressources à la disposition d'autrui soient également remerciées ici. Parmi celles-ci, je suis tout particulièrement redevable à Paolo Mairano, maître en rythme linguistique, également rencontré aux RJCP, d'avoir développé et mis en ligne sur le site du Laboratoire de phonétique expérimentale 'Arturo Genre' de l'Université de Turin toutes ces ressources si utiles. De même, je remercie les auteurs Paul Boersma and David Weenink de l'université d'Amsterdam, qui mettent à notre disposition l'outil logiciel d'analyse acoustique PRAAT et qui font tout pour que nous puissions nous en servir le mieux possible. Ce paragraphe ne serait pas complet sans mentionner Nicolas Audibert, que je n'ai malheureusement pas eu l'occasion de rencontrer, mais qui m'a permis de gagner un temps précieux grâce à sa mise à disposition de *scripts* utilisables avec ce même logiciel PRAAT.

Enfin, je tiens à remercier mon épouse, pour sa patience à toute épreuve et son soutien logistique indéfectible et indispensable, ainsi que mes (grands) fils pour leur relecture critique et constructive de certaines parties de ce mémoire.



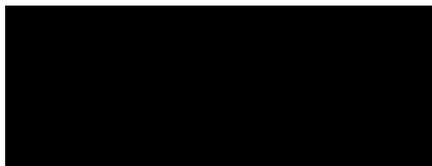
### DECLARATION

1. Ce travail est le fruit d'un travail personnel et constitue un document original.
2. Je sais que prétendre être l'auteur d'un travail écrit par une autre personne est une pratique sévèrement sanctionnée par la loi.
3. Personne d'autre que moi n'a le droit de faire valoir ce travail, en totalité ou en partie, comme le sien.
4. Les propos repris mot à mot à d'autres auteurs figurent entre guillemets (citations).
5. Les écrits sur lesquels je m'appuie dans ce mémoire sont systématiquement référencés selon un système de renvoi bibliographique clair et précis.

NOM : **PICAVET**

PRENOM : **Francis**

DATE : 27/09/2012



# Table des matières

Remerciements .....	4
Table des matières .....	4
Notes liminaires.....	6
Conventions typographiques et de notation.....	6
Variétés de l’anglais et modèles .....	6
Terminologie et abréviations d’origine anglaise .....	7
INTRODUCTION .....	8
<b>PARTIE I. - CADRE ET OBJET DE CETTE ETUDE.....</b>	<b>11</b>
Que recherchons-nous ? .....	12
Une démarche d’évaluation d’une tâche de lecture .....	12
L’objet de notre recherche : une étude de cas.....	14
Notre sujet d’étude : Jean .....	15
Accent et « stress ».....	15
« Stress » et focus.....	18
L’approche pédagogique utilisée : le métronome.....	20
<b>PARTIE II. - PARAMETRES D’ETUDE ET SYNOPSIS DE L’ETAT DE L’ART .....</b>	<b>21</b>
Paramètres retenus pour l’interprétation du rendu des phrases par notre sujet .....	22
Qualité vocalique .....	22
Quantité vocalique, quantité consonantique, quantité syllabique.....	24
Durée brute .....	24
Différence dS - dV.....	26
Vitesse instantanée de parole et accélération syllabique.....	27
Hauteur mélodique (F <sub>0</sub> ).....	28
Intensité sonore .....	29
<b>PARTIE III. - METHODOLOGIE EXPERIMENTALE .....</b>	<b>30</b>
Construction du corpus Ense3 2012 .....	31
Méthodologie utilisée pour l’analyse des phrases sous PRAAT : segmentation .....	32
Niveau orthographique.....	32
Niveau syllabique.....	32
Niveau syntaxique.....	33
Niveau vocalique .....	33
Remarques .....	34
<b>PARTIE IV. - ANALYSE, INTERPRETATION ET DISCUSSION DES RESULTATS .....</b>	<b>35</b>
Introduction.....	36
CHAPITRE 1. — ANALYSE DETAILLEE DES LECTURES DE LA PHRASE PH01 DU TEST PAR NOTRE SUJET ..	40
Niveaux de « stress » perçus : données perceptives-subjectives.....	40
Introduction.....	40
Phrase 01 .....	41
Qualité vocalique de la phrase Ph01 : données perceptives-subjectives.....	44
Phrase Ph01 : mesures objectives, analyse et interprétation.....	47
Introduction.....	47
Création d’un « test de référence ».....	48
Analyse succincte du graphique de la figure 6.....	49
Mesures de durées brutes de la phrase Ph01 .....	50
Introduction.....	50
Graphiques de durée brute : voyelles et syllabes.....	51
Conclusion : durée brute et « stress » .....	58
Différence dS – dV : quantité consonantique de la phrase Ph01 .....	59
Introduction.....	59
Analyse de la courbe de quantité consonantique de référence .....	60
Analyse des courbes de quantité consonantique issues de T <sub>init</sub> et T <sub>term</sub> .....	61
Conclusion : quantité consonantique de la phrase Ph01.....	66

Calcul de la vitesse de parole et de son accélération dans la phrase Ph01 .....	67
Calcul de la vitesse instantanée de parole dans la phrase Ph01 .....	68
Calcul de l'accélération syllabique dans la phrase Ph01 .....	72
Conclusion : vitesse de parole et accélération syllabique.....	75
Hauteur mélodique ( $F_0$ ) des voyelles de la phrase Ph01 .....	76
Introduction.....	76
Analyse de la hauteur mélodique ( $F_0$ ) de la phrase Ph01 .....	76
Conclusion sur la hauteur mélodique.....	82
Intensité sonore des voyelles de la phrase Ph01 .....	82
Introduction.....	82
Analyse de l'intensité moyenne des voyelles de la phrase Ph01 .....	82
Conclusion sur l'intensité sonore des voyelles de la phrase Ph01 .....	87
CHAPITRE 2. — ESSAI DE REPRESENTATION GRAPHIQUE SYNTHETIQUE DU <i>STRESS</i> PRODUIT POUR LA PHRASE 01 .....	88
Introduction.....	88
Représentation du <i>stress</i> : vues d'ensemble de la phrase Ph01 .....	89
Représentation du <i>stress</i> : analyse succincte de détails de la phrase Ph01 .....	93
Représentation dynamique du <i>stress</i> dans les graphiques à bulles.....	95
Conclusion de nos essais de représentation synthétique du stress.....	98
CHAPITRE 3. — ANALYSE COMPARATIVE DES VOYELLES DE 4 PHRASES DU TEST (PH01, PH03, PH05, PH10) .....	98
Introduction.....	98
Analyse comparative des voyelles de 4 phrases du test : diphtongues .....	102
Introduction à l'analyse des diphtongues .....	102
Analyse des graphiques de durée et d'intensité des diphtongues de Ph01, Ph03, Ph05 et Ph10 .....	102
Analyse comparative des voyelles de 4 phrases du test : voyelles longues .....	106
Introduction à l'analyse des voyelles longues.....	106
Analyse des graphiques de durée, de hauteur mélodique et d'intensité des voyelles longues de Ph01, Ph03, Ph05 et Ph10.....	106
Analyse comparative des voyelles de 4 phrases du test : voyelles courtes et voyelles faibles.....	109
Conclusion sur l'analyse comparative des voyelles de 4 phrases du test.....	113
CHAPITRE 4. — ANALYSE COMPARATIVE DE LA COMPOSITION SPECTRALE DE 4 VOYELLES LONGUES DANS LES 12 PHRASES DU TEST EN VERSIONS $T_{INIT}$ ET $T_{TERM}$ .....	115
Introduction.....	115
Formants $F_1$ et $F_2$ .....	115
Utilisation de données de référence à fin de comparaison.....	116
Mesures de $F_1$ et $F_2$ de 4 monophthongues $V_{long}$ de Jean et comparaison aux données de référence .....	117
Mesures de $F_1$ et $F_2$ pour le son /i:/ .....	118
Mesures de $F_1$ et $F_2$ pour le son /a:/ .....	119
Mesures de $F_1$ et $F_2$ pour le son /ɔ:/.....	121
Mesures de $F_1$ et $F_2$ pour le son /u:/ .....	122
Conclusion sur les mesures de la composition spectrale de 4 monophthongues de type « longues ».....	123
Conclusion.....	126
Bibliographie.....	129
Table des annexes.....	130
Table des illustrations (dans le texte) .....	138
Table des illustrations annexes.....	140
Sigles, abréviations, symboles et unités utilisés .....	146

## Notes liminaires

### *Conventions typographiques et de notation*

Notre sujet portant sur la langue anglaise, nous aurons fréquemment recours à cette langue. A fin d'efficacité et de clarté, nous utiliserons systématiquement les caractères italiques pour les mots en anglais dans le texte, et uniquement pour ceux-là ; ainsi le mot *accessible* (anglais) se différenciera de **accessible** (français) par la langue à laquelle il est fait référence.

Nos transcriptions phonémiques, de référence et perceptives, suivent les règles de l'Alphabet phonétique international (API simple) et les choix de transcription de John Wells, tels que nous les trouverons dans son *Pronunciation Dictionary* (Wells, 2008). Toutefois, nous préférons généralement matérialiser les frontières syllabiques par un point (·) plutôt que par un simple espace, pour la clarté de notre propos. Nous n'utiliserons qu'une transcription simplifiée de l'anglais et, parfois, de sons français utilisés par le locuteur francophone s'exprimant en anglais, quand cela s'avère nécessaire.

### *Variétés de l'anglais et modèles*

C'est le modèle « anglais britannique » que nous avons utilisé arbitrairement tout au long de ce travail, dans sa variété théorique actuelle généralement dénommée *Extended Received Pronunciation*, que nous abrègerons en « RP ». Dans la moitié sud de la Grande Bretagne, la prononciation RP est généralement considérée comme le modèle d'une « prononciation correcte » dans une situation de communication orale dans un registre soutenu. Cette variété de l'anglais a été autrefois appelée « anglais de la BBC ». L'anglais de prestige de la BBC<sup>1</sup> a souvent été pris comme exemple de référence dans un passé maintenant lointain. Depuis fort longtemps, la BBC ne sélectionne plus ses présentateurs sur la « pureté » de leur accent et tous les accents régionaux des Îles Britanniques, voire au-delà, sont susceptibles d'être représentés aujourd'hui. Les accents issus des publications sélectionnées<sup>2</sup> proviennent néanmoins de locuteurs s'exprimant avec un « accent standard

---

<sup>1</sup> *British Broadcasting Corporation*, organisme public de radio et télédiffusion britannique fondé en 1922.

<sup>2</sup> Cf. Partie IV, chapitre 4

du sud (de la Grande Bretagne) », la *Standard Southern British English Pronunciation*, qui a servi de base à la définition de la prononciation RP. Dans l'introduction de son dictionnaire, Wells rappelle que la prononciation RP n'est pas figée, qu'elle évolue dans le temps, notamment du fait d'une démocratisation de la société britannique, et qu'elle tolère une certaine variété interne (Wells, 2008).

Il est évident que, dans l'usage personnel que nous faisons de l'anglais et dans nos pratiques pédagogiques quotidiennes, nous abordons autant de variétés de l'anglais que nous le pouvons. Nous exposons nos étudiants à ces variétés, nous nous efforçons d'en décrire toute la richesse et d'en comprendre toutes les subtilités. Toutefois, dans le cadre restreint de ce travail, il nous a fallu faire un choix arbitraire nous permettant de ne pas multiplier les exemples ou les commentaires. Il existe, par exemple, un modèle similaire du dialecte américain de l'anglais connu sous le nom de *General American (GA)*, que nous aurions pu tout aussi bien employer. Les dictionnaires de qualité proposent généralement, au minimum, une version RP et une version GA de la prononciation de chacune de leurs entrées. Nous utiliserons occasionnellement les symboles £ (anglais britannique) et \$ (anglais américain) quand nous aurons à nous référer à l'une ou l'autre de ces variétés.

Enfin, si nous proposons des modèles à nos étudiants, nous ne leur demandons pas de les apprendre « par cœur », dans un psittacisme qui serait aveugle et stérile. Nous nous efforçons de faire en sorte qu'ils se les approprient, qu'ils en comprennent l'origine, la fonction et le fonctionnement, la subtilité, et la variation, omniprésente ; ceci avec le dessein de les rendre autonomes, adaptables et compétents.

### ***Terminologie et abréviations d'origine anglaise***

La langue utilisée tout au long de l'élaboration de ce mémoire a été principalement l'anglais ; tour à tour matière travaillée, outil de communication, sujet d'étude et langue de travail, elle est appréciée pour sa concision. Nous utiliserons un nombre limité d'abréviations d'origine anglaise issues de notre usage courant, qui sont toutes clairement définies lors de leur première occurrence dans le texte et qui sont rassemblées en fin de volume dans une table des **sigles, abréviations, symboles et unités utilisés**. Nous espérons que le lecteur se les appropriera rapidement.

## Introduction

Le travail de recherche présenté dans ce mémoire traite de l'expérimentation d'une approche rythmique de l'enseignement de l'anglais oral à des étudiants francophones. Il trouve son origine dans l'idée formulée et mise en pratique de longue date par les enseignants de la langue, mais peut-être encore insuffisamment explorée, qu'une approche musicale de l'anglais « langue étrangère » peut permettre une plus grande efficacité dans l'enseignement ou l'apprentissage à l'oral, en production comme en perception. L'approche rythmique mise en œuvre a consisté essentiellement en la réalisation de multiples tâches d'expression orale guidées par le mouvement pendulaire d'un métronome.

Nous pensons que les performances du locuteur non-natif (LnN), la qualité générale de la chaîne segmentale et la prosodie, inhérente à toute communication orale, peuvent être améliorées de façon importante par la mise en œuvre de ce type d'approche, notamment dans le cas de l'interaction avec des locuteurs natifs (LN). Parmi les éléments linguistiques ayant un lien avec la musique, tels que la mélodie, le rythme, les nuances ou le timbre – tous dignes d'intérêt – c'est le rythme que nous avons retenu pour la présente étude, à cause de sa place prépondérante dans la dichotomie anglais-français.

L'un des axes majeurs de notre étude réside dans la conviction qu'un enseignement des aspects prosodiques, et plus particulièrement du rythme, peut compléter et améliorer non seulement la qualité de la production orale au niveau suprasegmental, mais également au niveau segmental, ainsi qu'en perception. L'hypothèse de la prosodie comme vecteur fondamental de l'enseignement de la communication orale a déjà été mise en avant par des auteurs tels qu'Iván Fónagy (Fónagy, 1983). L'importance du rythme de parole avait été déjà été établie par d'autres comme Paul Fraise (Fraise, 1974) ou Louis-Jean Calvet (Calvet, 1979).

Ces aspects sont souvent problématiques pour les étudiants francophones s'exprimant en anglais, alors que le poids de l'iso-syllabisme imprégnant la langue maternelle freine la mise en œuvre de l'accentuation dynamique syllabique indispensable à la communication efficace en anglais. Cette classification entre langues isosyllabiques et iso-accentuelles a été proposée par Rebecca Dauer (Dauer, 1983), puis reprise, discutée et approfondie par Pier Marco Bertinetto (Bertinetto, 1989) et, plus récemment, par d'autres auteurs comme Paolo Mairano (Mairano, 2010). Pour discutable qu'elle soit, cette classification s'applique relativement bien à ces deux langues que sont l'anglais et le

français, qui bien que partageant un fonds lexical avoisinant les cinquante pour cent (cf. Walter, 2003), n'en ont pas moins des systèmes prosodiques et accentuels si différents que des locuteurs non-initiés des deux langues ne se comprennent pas.

En tant que professeur d'anglais « langue vivante étrangère » (et occasionnellement de « français langue étrangère ») depuis plus de vingt-cinq ans, en France et à l'étranger, l'auteur de ce mémoire a trop souvent entendu des étudiants rentrer de premiers séjours dans des pays anglophones dire que, bien que parfaitement capables de produire des énoncés « corrects » d'un point de vue morphosyntaxique, ils éprouvaient les plus grandes difficultés, tant à comprendre qu'à se faire comprendre, dans leur interaction avec des locuteurs natifs. Le chaînon manquant, ce n'est pas que la difficulté à relier le signifiant au signifié, dans les conditions exigeantes de l'interactivité en langue étrangère, c'est également, en tant que dimension indispensable à l'intercompréhension, la prosodie ; nombre d'auteurs l'ont très bien décrit et expliqué depuis longtemps (ex. Delattre, 1966, Fónagy, 1983). Peut-être n'a-t-elle pas reçue suffisamment d'attention pédagogique jusqu'à une période récente.

Voici quelques questions qui ont accompagné ma réflexion, préalablement à la décision d'entamer la présente étude. Qu'est-ce qui fait que parler l'anglais est si difficile pour les locuteurs de langues romanes, notamment les francophones ? Que pouvons-nous faire pour déclencher la conscience du rythme, de la dynamique, chez l'apprenant ? Par quels moyens pouvons-nous arriver à une production langagière intégrant aussi naturellement que possible ces éléments fondamentaux ? L'auteur de ce mémoire a voulu expérimenter l'application d'approches musicales simples dans son enseignement de l'anglais « langue seconde » (L2) et en mesurer les effets aussi précisément que possible, choisissant pour commencer de s'intéresser à la dimension rythmique de la prosodie. La méthode employée et ses tous premiers résultats, objet de ce travail de recherche et aujourd'hui intégrée à la pédagogie de l'auteur, est expliquée dans les pages suivantes, au moyen d'une étude de cas. Cette étude de cas pourra éventuellement servir à établir un protocole expérimental pour une étude de plus grande envergure, à partir des nombreuses données recueillies auprès de plusieurs dizaines de sujets.

Dans une première partie, nous détaillerons le cadre et l'objet de notre étude de cas. Dans un deuxième temps, nous présenterons les paramètres étudiés et la manière dont nous les avons abordés. Ensuite, nous décrirons notre méthodologie expérimentale. Enfin, nous

procèderons à l'analyse, à l'interprétation et à la discussion de nos résultats. En conclusion, nous proposerons quelques perspectives pour d'éventuels travaux de recherche à venir.

## **Partie I.**

-

### **Cadre et objet de cette étude**

### ***Que recherchons-nous ?***

Dans l'étude d'une langue, tous les aspects de l'utilisation de celle-ci sont potentiellement pertinents. Posons tout de suite en préambule que c'est dans le cadre restreint de **la communication orale** et de **l'évaluation de sa qualité en situation de production** que nous nous situons. Nous considérons **la parole**, émise ou perçue, comme étant un **continuum**. Le processus d'analyse de données isolées, recueillies et présentées en tableaux, conséquence du travail de segmentation de cette parole, pour autant indispensable qu'il soit, doit être considéré avec la plus grande circonspection. Ainsi, un anglophone (LN - locuteur natif) écoutant un locuteur non-natif (LnN) pourra parfaitement comprendre le message et/ou l'intention de ce dernier au niveau suprasegmental, dans la globalité du continuum, et ce malgré des inexactitudes lexicales, syntaxiques, phonologiques ou prosodiques inhérentes à son « interlangue », cet état intermédiaire de la langue étrangère en cours d'acquisition, variable et propre à chaque apprenant. Néanmoins, nous ferons le choix de nous atteler à cette tâche de segmentation, dans une démarche de l'élément vers l'ensemble (*atomistic* ou *bottom-up approach* en anglais). Nous nous efforcerons de vérifier nos hypothèses de travail, avec rigueur et vigilance quant à l'interprétation de nos résultats.

Le travail de segmentation nous permet l'analyse fine des phénomènes acoustiques ou articulatoires, afin d'en rendre compte de la meilleure façon possible et de proposer les évolutions utiles, à l'apprenant, à l'enseignant ou au chercheur, dans une démarche de progrès cognitif, notamment linguistique et communicationnel.

### ***Une démarche d'évaluation d'une tâche de lecture***

L'objectif recherché ici est de mesurer l'évolution des compétences et les progrès d'étudiants dans leur production orale de l'anglais « langue seconde » (L2) ou « troisième » (L3). Le travail effectué est basé sur un corpus construit à dessein à partir des enregistrements de douze phrases produites par des étudiants, avant et après un travail rythmique régulier sur 10 semaines, à raison d'une moyenne de trente minutes par séance, et de comparer les performances acoustiques initiales et terminales des étudiants.

Ces douze phrases sont lues à haute voix au cours d'un test **T**, identique en phase initiale **T<sub>init</sub>** et terminale **T<sub>term</sub>** de l'apprentissage.<sup>3</sup> Le choix d'une tâche de lecture nous éloigne d'emblée de toute situation de communication « naturelle » de type « discussion informelle ou formelle » entre LN et LnN. Ce sont avant tout des contraintes de praticité et de contrôlabilité qui nous poussent à faire ce choix d'une tâche « artificielle » de lecture. Conscient des limites de ce type de test, nous aurons à en tenir compte tout au long de l'étude, notamment lors de l'interprétation de nos résultats, ainsi que nous le rappelle A.C. Gimson (Gimson, 2008).

If, however, read texts are used [...] the artificiality of the procedure should be recognized and allowed for, since a certain unnaturalness of style is likely, whether it is a native speaker or a foreign learner who is reading aloud.<sup>4</sup>

Nous déterminerons subjectivement (écoute, perception) et objectivement (mesures) la qualité phonétique et prosodique de la production, dans une démarche comparative entre **T<sub>init</sub>** et **T<sub>term</sub>**, en faisant varier les paramètres et en rapprochant et en analysant les données selon les critères définis ci-dessous. L'oreille d'un enseignant expérimenté est capable d'apprécier le progrès en amont et en aval de la période de formation considérée, notamment en comparant les enregistrements sonores. Cependant, il est indispensable, dans une démarche qui se veut scientifique, de pouvoir quantifier et mesurer ces progrès avec la plus grande précision possible, de vérifier les hypothèses avancées et de confirmer, le cas échéant, notre perception subjective.

Notre approche est volontairement basée sur une référence. Le but déclaré est de proposer à l'étudiant de s'efforcer de se rapprocher d'une référence linguistique théorique telle que définie dans les dictionnaires (classiques, numériques ou sur l'Internet : Robert & Collins, Harrap's, Cambridge, Oxford, etc.) ou autres ouvrages de prononciation (ex. Gimson 2008, Wells 2008). Nous y trouvons la transcription phonétique conventionnelle au moyen de l'**alphabet phonétique international** (API) simple. Pour ce qui concerne le présent travail, et par soucis de clarté et d'efficacité, nous n'utiliserons systématiquement qu'une seule source de référence pour la transcription API, celle du *Pronunciation Dictionary* de J. C. Wells, dans sa dernière édition de 2008 (Longman).

---

<sup>3</sup> Le lecteur peut prendre connaissance des 12 phrases du test qui sont présentées dans la figure 1, dans l'introduction de la partie IV.

<sup>4</sup> « Si, toutefois, des textes lus sont utilisés [...] l'artificialité de la procédure doit être reconnue et prise en compte puisqu'une certaine artificialité de style est probable, que ce soit un locuteur natif ou un apprenant étranger qui lise à haute voix. » [Traduction de l'auteur]

Dans le domaine de l'apprentissage linguistique, l'exactitude absolue et le caractère répétable à l'identique des exercices et des expériences sont inconnus. Vouloir normaliser une langue est un exercice périlleux, en tout cas dans la volonté de représenter une réalité de terrain toujours mouvante, fonction de tant de paramètres et d'idiosyncrasies que la représentation en sera toujours imparfaite. La référence-outil, telle que nous la trouvons dans les dictionnaires de prononciation de qualité, pourra néanmoins rendre de grands services à l'apprenant, motivé par un modèle approximant à atteindre tout en étant conscient des limites de cette référence. L'un des problèmes rencontrés dans une approche référentielle est de définir le degré d'acceptabilité avec lequel nous évaluerons les productions des étudiants. En effet, quelle que soit la qualité d'un énoncé, même émis par un LN, il ne coïncidera jamais entièrement avec cette référence. Le danger est grand de considérer que tel ou tel résultat, forcément déviant, comme « non-acceptable » ou comme une « faute ». Notre position sera de considérer que tout énoncé enregistré en respectant les critères du test comme « correct » ou « acceptable » par défaut, sans émettre de jugement de valeur ; que l'énoncé enregistré est une réalisation instantanée de l'interlangue de l'étudiant à un moment donné ; que s'agissant d'une situation de lecture, l'extrapolation vers une situation de communication « authentique » devra se faire avec de grandes précautions ; qu'une évaluation de la capacité à communiquer ne pourrait se faire qu'en interaction avec un(e) ou plusieurs LN, dans des prises de parole libres et prolongées.

### ***L'objet de notre recherche : une étude de cas***

Tous les aspects de l'apprentissage d'une langue sont dignes d'intérêt. Le domaine qui nous intéressera tout particulièrement est celui de l'apprentissage d'une langue étrangère, l'**anglais**, langue seconde (L2) ou troisième (L3), par de jeunes adultes, majoritairement francophones, en fin de formation universitaire. Dans le strict cadre du présent travail de recherche, nous procéderons à l'étude approfondie d'un cas —que nous nommerons « Jean » par respect de son anonymat. Nous avons délibérément choisi une voix masculine dans cette approche initiale, celle-ci étant a priori relativement plus facile à analyser qu'une voix féminine, notamment pour notre partie d'analyse de la composition spectrale des énoncés.<sup>5</sup> Nous projetons d'étudier des voix féminines à l'avenir.

---

<sup>5</sup> Cf. Partie IV, chapitre 4

Ce cas, issu de notre **Corpus Ense3 2012**<sup>6</sup> est susceptible de servir d'échantillon-test pour une étude linguistique et statistique de plus grande envergure que la présente étude, basée sur tout ou partie du corpus, et éventuellement de permettre la définition d'un protocole de recherche. Dans cette **étude de cas**, nous nous concentrerons sur l'un des défis —de notre expérience personnelle et professionnelle, un des plus emblématiques et, surtout, des plus indispensables à la communication orale— que les francophones apprenant l'anglais L2 ou L3 doivent relever : **l'acquisition du rythme**. Cette étude de cas pourra éventuellement servir de base à l'élaboration d'un protocole pour des travaux de recherche plus approfondis, à partir du Corpus Ense3 2012, ou d'autres.

### ***Notre sujet d'étude : Jean***

Considérons brièvement le sujet de notre étude de cas, que nous avons appelé « Jean ». Nous avons fait le choix délibéré de prendre un étudiant qui soit représentatif de notre population étudiante masculine. Concernant son niveau d'anglais, Jean obtient des résultats qui se situent dans la moyenne de cette population. Ce n'est pas un étudiant « exceptionnel » dans cette matière.

Le questionnaire-type que Jean a rempli, comme tous les participants, se trouve en annexe<sup>7</sup> de ce mémoire. Nous ne rentrerons pas dans le détail de toutes les réponses fournies par Jean. En voici quelques extraits.

- C'est un étudiant français de 22 ans en 2<sup>ème</sup> année d'école d'ingénieurs (Bac +4)
- Il n'a séjourné que 2 semaines dans un pays anglophone et a donc appris l'anglais essentiellement dans le milieu scolaire, depuis l'âge de 12 ans.
- Il pratique l'espagnol en 3<sup>ème</sup> langue vivante (L3) depuis l'âge de 14 ans

Jean peut être considéré comme un représentant relativement typique des étudiants fréquentant l'école. Les 64 réponses détaillées aux questions de chacun des 75 étudiants pourront faire l'objet d'études statistiques à l'avenir.

### ***Accent et « stress »***

Parmi les différents paramètres du rythme linguistique, il y en a un sur lequel nous porterons tout particulièrement notre attention : **l'accent dynamique lexical**. Cet **accent**

---

<sup>6</sup> Cf. Partie III, chapitre 2 : construction d'un corpus.

<sup>7</sup> **Annexe 1** de ce mémoire

joue un rôle de premier ordre dans la langue anglaise orale. Nous en voulons pour preuve que son absence ou son déplacement compromet la bonne compréhension du message par un LN ; au mieux, il provoque une incompréhension partielle ou totale du LN, et ce même si une syntaxe et un lexique parfaitement appropriés sont mis en œuvre par le LnN ; au pire, il provoque un contresens. Pour les besoins de notre étude, nous considérerons que **l'accent dynamique** se décline en trois aspects, de natures semblables mais d'échelles différentes : **l'accent dynamique lexical**, se situant au niveau du segment syllabique, mettant en relief une syllabe à l'intérieur du mot ; **les accents dynamiques « syntagmatique »** et **« syntaxique »**, se situant au niveau d'un mot entier, ou d'un syntagme, parmi d'autres, dans la phrase. Pour ces deux dernières catégories, nous éviterons d'employer ces termes et les regrouperons sous la périphrase, moins commode mais plus exacte : tous les phénomènes accentuels qui émergent de l'organisation prosodique.

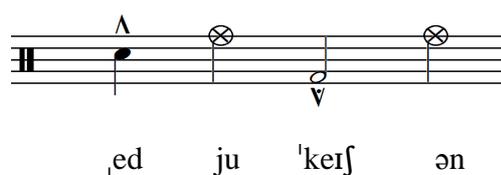
Pour ce qui concerne l'anglais, **au niveau lexical** — objet principal de notre étude — il est courant de considérer qu'un mot polysyllabique sera constitué d'une alternance de syllabes « accentuées » et « inaccentuées ». Dans les mots de 4 syllabes ou plus, les syllabes accentuées se déclineront généralement en « accent principal » et « accent secondaire ». Pour notre part, nous préfererons partir du principe que toutes les syllabes portent un accent : nous attribuerons la valeur « **1** » par défaut, aux syllabes couramment appelées « inaccentuées », de loin les plus fréquentes. En effet, un locuteur qui émet une phrase commençant par l'article indéfini *a* (prononcé [ə] dans l'immense majorité des cas) comme par exemple ce vers de John Keats « *a thing of beauty is a joy for ever* »<sup>8</sup>, passe du silence, valeur « **0** », à l'émission d'un son, de valeur « **1** ». Le fait que ce son initial rompe le silence est en lui-même un « accent », le début, le déclenchement, la mise en œuvre d'un processus phonatoire, l'**attaque** (*onset* en anglais acoustique ; *attack* en musique) d'un son. C'est dans la **relativité** des accents appliqués par le locuteur aux autres segments-syllabes constituants de la chaîne acoustique que se révéleront les accents de niveau **3** « très forts », ou les accents « forts » de niveau **2**. Les accents du syntagme *a thing of beauty* peuvent être transcrits par l'alternance **1-3-1-3-1** ou, en caractères phonétiques, par /ə·'θɪŋ·əv·'bju:t-i/. L'alternance de ces accents est à l'origine d'un **contraste** caractéristique

---

<sup>8</sup> **John Keats, *Endymion*, 1818** : A thing of beauty is a joy for ever: / Its loveliness increases; it will never / Pass into nothingness; but still will keep / A bower quiet for us, and a sleep / Full of sweet dreams, and health, and quiet breathing. [...]

de l'anglais oral et la source d'un **rythme**, une alternance de « temps forts » et de « temps faibles ». Dans les syntagmes incluant un adjectif (ex. *a sweet dream*), ou dans certains mots de quatre syllabes ou plus (ex. *education*), apparaîtra l'accent de niveau 2, le moins fréquent : /əˌswi:tˈdri:m/, soit la séquence 1-2-3 ; /,edˌjuːˈkeɪʃən/, soit 2-1-3-1. La transcription des accents 2 et 3 se fait, traditionnellement, à l'aide de signes diacritiques conventionnels de l'API, respectivement (ˌ) et (ˈ) placés avant la syllabe accentuée, comme dans l'exemple *education* ci-dessus. Conventionnellement, nous considérerons qu'un accent de niveau 2 ne peut exister que relativement à un accent de niveau 3. En termes de **qualité** de réalisation de phonèmes par les LN, il n'y a habituellement pas de différence majeure entre les niveaux accentuels 2 et 3 ; par contre, il y a généralement des différences remarquables entre le niveau 1 et les niveaux 2 et 3. *Alabama*, par exemple, se dira /,æɪəˈbæm.ə/£<sup>9</sup>, soit 2-1-3-1 : les « A » des niveaux 2 et 3 sont réalisés qualitativement de manière identique en £ : /æ/ ; contrairement à ceux de niveau 1, prononcés /ə/. Une transcription du rythme de *education* selon les conventions musicales pourrait aussi aisément être faite en notation de type « percussions » et joué sur une batterie où, par exemple :

- le niveau 1 serait représenté par la cymbale : 
- le niveau 2 serait représenté par la caisse claire : 
- le niveau 3 serait représenté par la grosse caisse : 
- le rythme de *education* serait transcrit ainsi :



,ed ju 'keɪʃ ən

- les symboles ci-dessous représentant respectivement des accents musicaux :

« fort »  $\Lambda$

« très fort » (*marcato*)  $\blacktriangle$

Ce type de transcription musicale du rythme, son rendu sonore —par des moyens purement acoustiques (instruments de percussion) ou numériques— et son utilisation à des

---

<sup>9</sup> Nous utiliserons, si nécessaire, ces symboles conventionnels pour différencier les prononciations britannique/RP (£) et américaine/GA (\$) ; s'agissant d'une référence à un état américain, il est utile de préciser que cette prononciation est britannique.

fins d'illustration, ou didactiques, pourra éventuellement faire l'objet d'autres de nos recherches à l'avenir.

Nous concluons cette partie par la présentation de notre choix terminologique en matière d'accent lexical. Le terme « accent » a en sa défaveur le fait d'être particulièrement ambigu, en français comme en anglais d'ailleurs. Il reçoit tour à tour les acceptions de signe diacritique graphique (accent aigu, grave, circonflexe, etc.) ; de diverses formes d'emphase sonore, par exemple d'une syllabe par rapport à une autre (accent de hauteur, accent tonique, etc.) ; de variétés orales régionales de langues ou de dialectes (l'accent du Midi, l'accent alsacien) ; de variétés orales sociales (l'accent bourgeois, l'accent populaire). A. C. Gimson rejette le terme courant de *stress*, suite, selon lui, à son utilisation ambiguë en phonétique et en linguistique; il lui préfère les termes « proéminence » (*prominence*) et retient celui de *accent* comme applicable « à ces syllabes s'élevant au dessus des autres, soit individuellement dans les mots, soit dans les énoncés plus longs » (notre traduction ; Gimson, 2008, p.24). J. C. Wells (Wells, 2008, p.783) opte pour une distinction nous semblant plus claire entre *stress* et *accent* :

**Stress:** a stressed syllable is one that carries a **rhythmic beat**. /.../ An **accent** is the placement of intonational pitch-prominence on a word.<sup>10</sup>

Ce sont ces définitions que nous utiliserons. Nous traduirons cet *accent* par **focus** en français et faisons le choix de conserver le terme concis et pratique de *stress* pour désigner cette pulsation rythmique typique, entre autres, de la langue anglaise. Nous utiliserons les abréviations **St1**, **St2** et **St3**, définies ci-après, pour désigner les accents lexicaux respectivement de niveaux 1, 2 et 3.

### « *Stress* » et *focus*

Le **focus** est une manifestation des phénomènes accentuels qui émergent de l'organisation prosodique et peut intervenir dans la dynamique de certaines syllabes ; *stress* et **focus** se superposent et interagissent. Le *stress*, accent lexical dynamique contrastif, se décline traditionnellement en 3 niveaux<sup>11</sup> (ex. Wells, 2008) chez les LN. Comme cela est

---

<sup>10</sup> *Stress* : une syllabe accentuée (sous *stress*) porte une pulsation rythmique. /... / *Accent* : placement d'une proéminence de hauteur mélodique intentionnelle sur un mot. [Traduction de l'auteur]

<sup>11</sup> Gimson constate l'occurrence d'un 4<sup>ème</sup> niveau qu'il qualifie de *minor prominence*, en présence d'une voyelle pleine ne portant pas d'accent (Gimson 2008, p.224). Ex. : la dernière syllabe de *hesitate* — /'hez ɪ teɪt/

généralement pratiqué, nous considérerons donc 3 niveaux de *stress*, avec toutefois la nuance décrite plus haut concernant le niveau 1 :

Niveau 1 : **Stress 1** ou **St1**. Accent « faible ». Accent de plus faible niveau (couramment : « syllabe inaccentuée / *unstressed syllable* ») ; également le plus fréquent

Niveau 2 : **St2**. Accent « secondaire », d'intensité marquée

Niveau 3 : **St3**. Accent « principal », de forte intensité.

Nous plaçant dans une approche référentielle, nous considérerons, a priori, que l'ensemble des LN applique les mêmes accentuations à l'intérieur de leur variété d'anglais (par ex. £ ou \$) ; ceci, bien sûr, est une simplification grossière de la réalité. Néanmoins, même si nos étudiants sont sensibilisés aux variations de la réalité du terrain, c'est vers un modèle théorique à atteindre que nous nous orientons.

Le **focus**, quant à lui, est généralement utilisé par le locuteur pour mettre en valeur un mot particulier de la phrase ; c'est une mise en relief de l'objet d'une discussion, par exemple. C'est le produit d'une deuxième articulation qui peut être à la base de toutes les fonctions communicatives. Nous donnerons au focus la valeur de **0** ou **+1**. Nous considérons que le focus se combine au *stress*. Généralement, cet ajout aura pour conséquence la perte d'un point de *stress* de la syllabe accentuée suivante. Par exemple, dans le mini-dialogue entre A. et B. ci-dessous, nous avons une combinaison article + adjectif + nom (St1 : pas souligné ; St2 : souligné simple ; St3 : souligné double) :

A — *I like **the blue car.***

Niv. :        **1**   **2**   **3**

B — *It's not a blue car, it's a black car!*

Niv. :        **1** **3**(=2+1) **2**(=3-1)        **1** **3**(=2+1) **2**(=3-1)

En B, au niveau sémantique, il y a contradiction et/ou correction de l'énoncé A. Cette adaptation due au focus appliqué à l'adjectif, illustre une sorte de hiérarchie accentuelle forte dans la phrase, une adaptation « automatique » découlant d'une quasi-incompatibilité de coexistence des accents de niveau 3 entre eux. Cette hiérarchie sera

utilisée tout naturellement par le LN pour différencier à l'oral le « professeur d'allemand » du « professeur allemand », par exemple : “a German teacher” ; “a German teacher”.<sup>12</sup>

### *L'approche pédagogique utilisée : le métronome*

Le principe de base de la méthode mise en œuvre a été d'inviter les étudiants, alors qu'ils étaient engagés dans des tâches d'expression orale ou qu'ils s'y préparaient, à concentrer une partie de leur attention aux battements rythmés d'un métronome, ajusté au tempo « confortable » de 60 à 70 pulsations par minute. Ainsi, leur production était guidée par le rythme des pulsations pendulaires.

Ce sont essentiellement des tâches de lecture qui furent mises en œuvre, parfois comme une fin en elles-mêmes, mais également comme préparation à des prises de parole libres, en continu ou en interaction avec d'autres étudiants.

Après avoir procédé à la recherche, puis au surlignage des accents principaux (de type St3) dans des phrases, celles-ci sont lues au tempo donné, selon le principe qu'une pulsation du métronome correspond toujours à une syllabe accentuée. Toutes les syllabes restantes, de type St1 ou St2 doivent pouvoir être placées entre 2 pulsations, de façon à « contraindre » les étudiants à faire ce que les locuteurs natifs font naturellement : accélérer la vitesse de parole entre deux pics d'accentuation, avec comme corolaire attendu, une réduction/centralisation vocalique, voire une neutralisation des voyelles.

Des métronomes sont disponibles « en ligne » sur l'Internet<sup>13</sup> ont été utilisés, car s'ils offrent une très grande précision, ils sont surtout d'une très grande souplesse d'emploi et permettent l'écoute au casque, ce qui offre de nombreux avantages lors d'un travail collectif notamment, ou d'un enregistrement. Des variations de tempo ont été soumises, de façon à mettre en relief l'effet de telle ou telle vitesse de parole sur la qualité de l'expression.

---

12 Nous noterons que le mot *German* passe du statut de **substantif** (une langue) à celui d'**adjectif** de nationalité : il ne s'agit pas de focus ici. L'écrit ne pouvant rendre les nuances accentuelles, une périphrase pourra être utilisée pour lever l'ambiguïté, si nécessaire : *a teacher of German*.

<sup>13</sup> Par ex. le site collaboratif : <http://bestmetronome.com/>

## **Partie II.**

-

### **Paramètres d'étude et synopsis de l'état de l'art**

## ***Paramètres retenus pour l'interprétation du rendu des phrases par notre sujet***

Comme vu auparavant, les quatre paramètres retenus classiquement comme pertinents dans notre analyse de la réalisation du contraste accentuel sont, dans notre ordre de travail :

- la qualité vocalique (*quality*)
- la quantité vocalique (*quantity*)
- la hauteur mélodique (*pitch*)
- l'intensité sonore (*intensity*)

Ces quatre paramètres sont développés ci-dessous.

### **Qualité vocalique**

La qualité de rendu vocalique sera considérée avant tout ici que dans son aspect de conséquence de l'application d'un niveau de *stress*, celui-ci ayant une influence déterminante sur la réalisation effective des voyelles en anglais. Dans le cadre de la présente étude, nous nous concentrerons sur la réalisation de certaines voyelles (quatre voyelles de l'anglais dites « longues ») par Jean, notre sujet, et essaierons d'apprécier la distance qui les sépare du modèle de référence de Wells (Wells, 2008) et d'un test de référence que nous appellerons  $T_{ref}$ <sup>14</sup>, pour les douze phrases des tests initial  $T_{init}$  et terminal  $T_{term}$ . Nous procéderons à une analyse formantique comparative de leur composition spectrale ( $F_1$  et  $F_2$ , essentiellement). Cette partie viendra en conclusion de notre travail d'analyse.<sup>15</sup>

Outre leur timbre, une autre caractéristique des voyelles qui nous intéressera tout particulièrement sera leur durée, qui sera traitée plus longuement ci-après dans notre partie « quantité vocalique », compte tenu de son importance. Remarquons simplement à ce stade que les syllabes sous *stress* de type **St2** et **St3** concentrent typiquement soit des sons riches, complexes : les diphtongues ( $V_{diph}$ ), de type /eə/, /əʊ/, /ɔɪ/, etc.) ; soit des sons longs, les voyelles « longues » ( $V_{long}$ ) de type /ɔ:/, /i:/, /u:/, etc.) ; soit des sons brefs, mais clairs, les

---

<sup>14</sup> Cf. Partie IV, Chapitre 1, p. 47

<sup>15</sup> Partie IV, Chapitre 4

voyelles « courtes » ( $V_{\text{short}}$ ) de type /e/, /i/, /ʊ/, etc. Les trois voyelles « faibles » ou « réduites » ( $V_{\text{weak}}$ ) /ə/, /ɪ/ et /u/ n'y seront jamais.

Tous les sons vocaliques peuvent se trouver dans les syllabes sous *stress* de type **St1** (les plus fréquentes). Toutefois, à cause du rythme et des pics de dynamique imposés par St2/St3, les syllabes sous St1 ont tendance à être raccourcies, parce que plus nombreuses et moins accentuées, et leurs sons vocaliques tendent à être neutralisés, réduits, par le phénomène de **centralisation-neutralisation** vocalique : tendance à être prononcés /ə/, son hypoarticulé se trouvant au centre du trapèze vocalique de l'API et le plus fréquent de la langue anglaise. Elles peuvent même disparaître : dans la langue orale, il n'est pas rare de constater l'apparition de « consonnes syllabiques » (syllabes sans voyelles ; cf. *colonel* ci-dessous) et, fréquemment, la **compression** de deux syllabes en une ou la disparition de syllabes entières : l'**élision**. Il en va de même de certains mots, quand ceux-ci sont très courants et évidents. Ce phénomène est certainement l'une des difficultés majeures de la maîtrise de l'anglais oral L2/L3, tant en compréhension qu'en expression. Par exemple, le mot *colonel*, de trois syllabes a priori, sera prononcé /'kɜ:nəl/ ; il y a une différence colossale entre l'orthographe en quatorze lettres du comté (*shire*) de la ville anglaise de Worcester, le Worcestershire, et sa prononciation très « compacte » (et, finalement, très facile) : /'wɔst.əʃə/, en sept symboles. Nous remarquerons donc que dans la plupart de ces cas, une liaison étroite entre qualité et quantité vocaliques existe, une qualité et une quantité qui varient simultanément en fonction du niveau de *stress*.

Notons ici que la classification phonémique des voyelles en  $V_{\text{long}}$  et  $V_{\text{short}}$  est toute théorique ou, au moins, relative. Gimson nous rappelle, à propos, que la durée d'une même voyelle ne sera jamais identique d'un énoncé à l'autre et qu'elle dépend de facteurs tels que le tempo de parole, le niveau d'accentuation de la syllabe, du voisement ou non de la consonne de coda (Gimson, 2008, p.24). Gimson conclue ce paragraphe par :

Such variations of length within the utterance constitute one manifestation of the rhythmic delivery which is characteristic of English and so is fundamentally different from the flow of other languages, such as French, where syllables tend to be of much more even length.<sup>16</sup>

---

<sup>16</sup> « De telles variations de durée à l'intérieur d'un énoncé constituent l'une des manifestations du rendu rythmique caractéristique de l'anglais ; ceci est fondamentalement différent du flux d'autres langues, telles que le français, où les syllabes ont tendance à être d'une durée beaucoup plus régulière. » [Traduction de l'auteur]

Dans nos mesures de quantités vocaliques, décrites ci-dessous, nous nous intéresserons à la durée des voyelles produites et, également, de celle des syllabes. Nous pourrions alors les comparer entre elles et effectuer des classements en fonction de leur durée objective.

### **Quantité vocalique, quantité consonantique, quantité syllabique**

#### *Durée brute*

La « quantité de voyelle » dans une syllabe peut se mesurer et s'interpréter de différentes manières. C'est probablement notre critère d'évaluation le plus important, le plus évident et l'un des plus pertinents ; c'est aussi l'un des plus facile à acquérir et à mesurer sous PRAAT, avec les précautions d'usage, bien entendu. Nous avons vu plus haut que les frontières voyelles-consonnes ne sont pas forcément clairement repérables, par exemple. Ce critère nous intéresse doublement : (1) les voyelles de l'anglais qui, notamment sous St2 ou St3, sont soit plutôt « (très) courtes » et « claquantes », soit plutôt « (très) longues », peuvent fournir un critère de comparaison des progrès d'un locuteur entre deux enregistrements de son interlangue, suite à un travail pédagogique par exemple ; (2) la qualité de l'expression orale d'un LnN francophone pourra être évaluée, entre autres, par sa capacité à jouer de façon appropriée sur cette durée, notamment à l'aide du contraste dynamique accentuel.

Dans son article sur les effets des variations de débit et tempo sur les syllabes parlées et chantées, N. Scotto Di Carlo (Scotto Di Carlo, 2007, §4) nous confirme que :

La durée moyenne des consonnes comme celle des voyelles augmente lorsque le débit ou le tempo ralentissent, mais alors que les durées vocaliques augmentent dans des proportions importantes, les durées consonantiques varient dans une limite réduite [ce qui] altère l'équilibre consonne/voyelle à l'intérieur de la syllabe.

Ainsi définit-elle « l'élasticité vocalique ». Les voyelles sont plus élastiques que les consonnes. Scotto Di Carlo explique cela par le fait que les voyelles, par définition, portent la voix, phénomène exemplifié à l'extrême dans la voix chantée ou criée (durée vocalique augmentée de 50%, durée consonantique réduite de 20% — Rostolland 1979, in Scotto Di Carlo 2007). Ces études portent sur la langue française ; sur cet aspect là, il en va de même, tout du moins dans le principe, pour la langue anglaise.

Le logiciel **PRAAT**<sup>17</sup> (cf. figure 1 ci-dessous) permet la mesure immédiate par lecture directe de la **durée brute** de la voyelle (ou de la syllabe), une fois la segmentation effectuée, ici manuellement ; ces données peuvent également être recueillies automatiquement à l'aide de traitements informatisés appropriés (*scripts*). Une fois les données de durée recueillies, une comparaison est relativement aisée à mettre en œuvre entre celles de nos tests  $T_{\text{term}}$  et de  $T_{\text{init}}$  ; nous analyserons alors la durée des voyelles. Dans le travail effectué en cours avec les étudiants sur le développement d'une plus grande énergie accentuelle, source de **contraste** entre les syllabes, la durée relative des voyelles  $V_{\text{short}}$  et  $V_{\text{weak}}$  devrait diminuer, accompagnée d'une réduction-centralisation vocalique pour les  $V_{\text{weak}}$ , et celle des  $V_{\text{long}}$  doit augmenter. Le symbole que nous utilisons pour la durée est **d**.

---

<sup>17</sup> Source : <http://www.fon.hum.uva.nl/praat/>

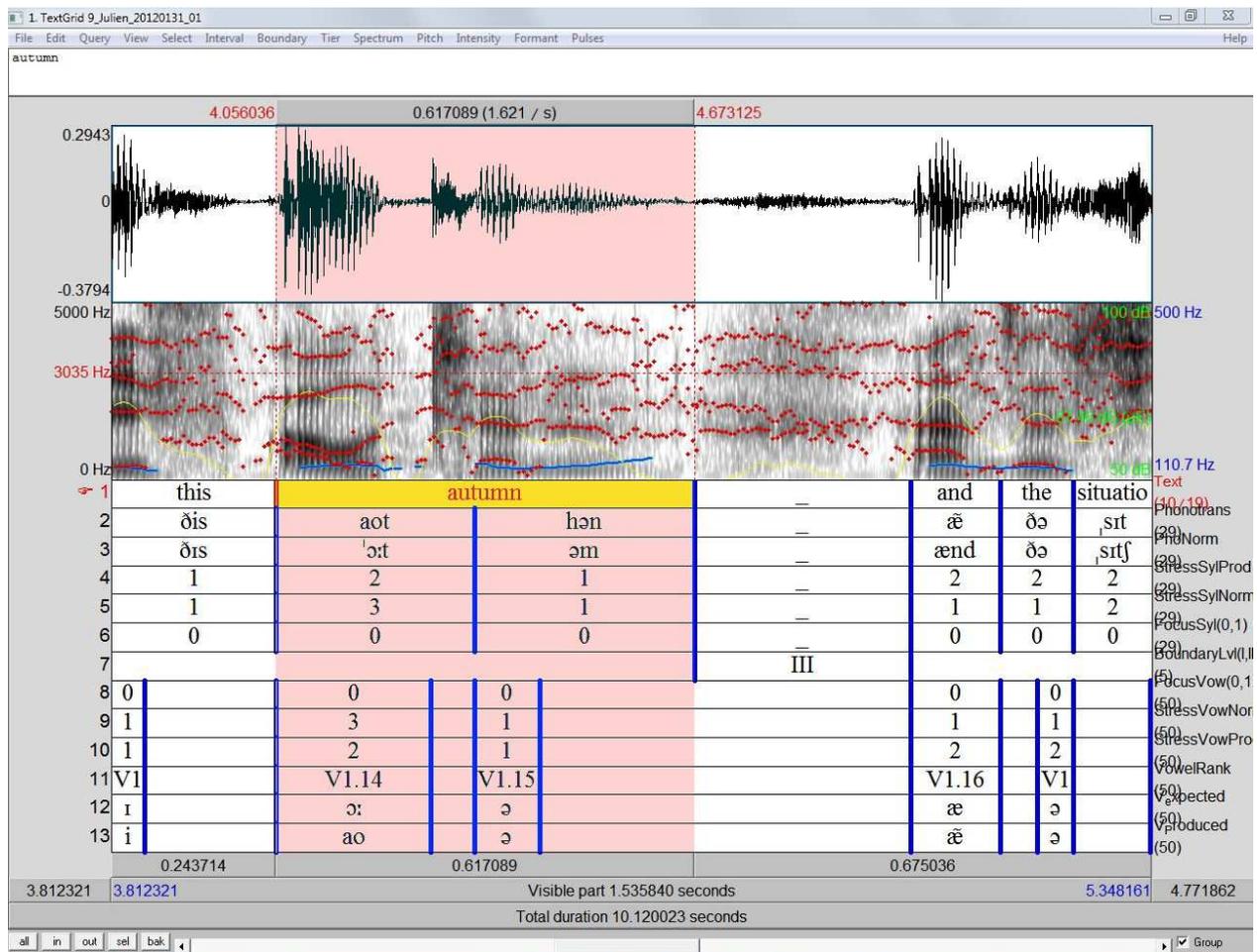


Figure 1. Exemple de fenêtre du logiciel PRAAT. Les « tiroirs » 1 à 6 concernent le découpage syllabique ; le tiroir 7 indique les frontières des pauses ; les tiroirs 8 à 13 concernent le découpage des sons vocaliques. Le découpage a été effectué manuellement.

### *Différence dS - dV*

Différence résultant de la soustraction de la durée de V à celle de S. Le résultat est le complément de la quantité vocalique : la **quantité consonantique** de la syllabe. Les consonnes, tout comme les voyelles, sont soumises à l'influence soit du *stress*, soit de son absence, en anglais oral, au travers des phénomènes de compression et d'élision. Un raccourcissement des valeurs dS - dV entre  $T_{\text{term}}$  et  $T_{\text{init}}$  pourra être interprété comme un progrès, quelque soit le niveau du *stress* : en effet, toutes les syllabes St1, toujours encadrées de syllabes proéminentes St2 ou St3, tendront à être écourtées ; sous l'impact de St2 ou St3, les syllabes avec  $V_{\text{short}}$  « claqueront » fortement, mais très brièvement (par conséquent, les consonnes initiales et/ou terminales avec), et dans les syllabes contenant le type  $V_{\text{long}}$  ou  $V_{\text{diph}}$  les consonnes serviront à déterminer la nature de l'attaque forte du son vocalique, long ou riche ; la coda conclura la syllabe, comme il se doit, plus ou moins

brièvement, selon la nature de la consonne : une consonne voisée aura tendance à être traitée différemment d'une consonne non-voisée.

Ainsi, une diminution de la durée consonantique de la syllabe entre  $T_{\text{init}}$  et  $T_{\text{term}}$  peut révéler une meilleure prise en compte du *stress* par notre sujet, durée théoriquement plus courte dans  $T_{\text{term}}$ , sous énergie accentuelle attendue comme plus forte.

### *Vitesse instantanée de parole et accélération syllabique*

Dans cette partie, nous nous intéresserons à la vitesse de parole de Jean, notre sujet. Parmi les paramètres du rythme, nous trouvons les aspects dynamiques (variations de la hauteur mélodique et de l'intensité, abordés dans les parties suivantes) et temporels (variations de la durée brute des syllabes et de ses composants, voyelles et consonnes, déjà traités). Étudions maintenant la vitesse de parole et son accélération, et leur probable variation, toujours à travers l'exemple de notre phrase Ph01. Dans le cadre de ce travail, la vitesse d'élocution et son influence ne nous intéresseront guère en eux-mêmes : ce ne sont pas à proprement parler des « variables ». Dans la tâche de lecture donnée à notre sujet, nous sommes dans une situation sans contraintes particulières, si ce n'est l'occurrence de la prochaine phrase. Nous laissons amplement le temps aux étudiants de lire celle-ci, ce qui peut être confirmé par la présence de « blancs » conséquents en amont et en aval de chaque enregistrement. Par contre, dans notre étude du rythme et de ses implications linguistiques, la **variation** de cette vitesse tout au long de la production de ces phrases nous intéressera tout particulièrement. En effet, selon la théorie isoaccentuelle, matérialisée à l'extrême dans le cas de l'utilisation d'un métronome, les accents de type St3 se présentent à intervalles réguliers. En fonction du nombre, variable, de syllabes présentes entre deux accents St3, le locuteur devra nécessairement accorder moins de temps à chaque syllabe intermédiaire.

En complément de notre réflexion sur l'accentuation des syllabes par notre sujet, nos mesures de durée nous permettront de calculer la **vitesse instantanée** pour chaque syllabe, puis l'**accélération syllabique**. Les modes de calcul respectifs seront détaillés dans la partie IV.

### Hauteur mélodique ( $F_0$ )

Gimson place les variations de **hauteur** mélodique (*pitch*) loin au premier rang des quatre facteurs, énoncés plus haut, qui marquent une syllabe accentuée, alors que Wells met en avant le terme, musical et séduisant, de « **pulsation rythmique** » (*rhythmic beat*) pour ce type syllabe, en mentionnant le fait que cette syllabe sera souvent caractérisée, entre autres, par « une proéminence de hauteur mélodique » (*pitch prominence*) avec son entourage (Gimson, 2008 ; Wells, 2008).

J. Vaissière rappelle que la hauteur perçue de  $F_0$  (**fréquence fondamentale**, ou  $F_0$  : fréquence la plus basse d'une onde périodique sinusoïdale, harmonique de 1<sup>er</sup> rang d'un son) du signal de source (le bourdonnement des plis vocaux), pour les sons périodiques, et la forme de ce signal, liée à la qualité de la voix, font partie des propriétés acoustiques de base des sons, ceux de parole compris (Vaissière, 2007). Ceci au même titre que leur intensité physique, dont nous parlerons ci-dessous, leur composition spectrale (répartition de l'énergie dans l'échelle des fréquences au niveau des harmoniques de rangs supérieurs  $F_1$  et  $F_2$  essentiellement), qui fait l'objet d'une partie à part entière dans le présent travail, ou, comme vu ci-dessus, leur durée.

Pour ce qui concerne notre objet d'étude, la dynamique accentuelle, nous noterons que celle-ci dépend autant de la hauteur mélodique absolue que de sa variation. En effet, en anglais, un simple glissement de cette hauteur (et/ou des autres paramètres) d'une syllabe à l'autre pourra, par exemple, faire passer un mot du statut de substantif à celui de verbe. Ainsi, le mot *contrast* sera perçu comme un nom avec le schéma accentuel 3-1, /'kɒn-trɑ:st/, et comme un verbe avec 1-3, /kən-'trɑ:st/. Dans le même ordre d'idée, l'adjectif *personal* (3-1, /'pɜ:sənəl/) aura un rendu sonore très différent du substantif *personnel* (2-1-3, /,pɜ:sənəl/), pourtant graphiquement proche. Une composante « fondamentale » de cette dynamique est, justement,  $F_0$  : un pic de la courbe de fréquence est normalement observé (cf. *pitch prominence* ci-dessus) sur la syllabe accentuée, de hauteur moyenne en St2 et élevé en St3, chez le LN. La réalisation —ou non— de cette variation chez l'apprenant doit être retenue comme critère pertinent essentiel dans notre évaluation.

Les valeurs moyennes de  $F_0$  de toutes les voyelles des phrases étudiées ont été extraites dans PRAAT à l'aide de traitements (*scripts*) appropriés<sup>18</sup> et recueillies classées dans des tableaux. Nous avons pu ainsi les trier et les présenter selon nos critères ou analyser les résultats et les présenter sous forme de graphiques.

### **Intensité sonore**

Pour Wells, la pulsation rythmique perçue sur la syllabe accentuée se caractérise systématiquement par un volume plus élevé (*loudness*), une intensité sonore (*intensity*) plus grande du locuteur lors de l'émission du son, alors que, comme nous l'avons vu, Gimson place les variations de  $F_0$  en premier (Gimson, 2008 ; Wells, 2008). Quoi qu'il en soit, nous retiendrons ce critère comme également pertinent dans notre étude.

Les syllabes accentuées en St2 et St3 de l'anglais présentent généralement une intensité plus élevée lors de leur émission qui, combinée à une montée simultanée de  $F_0$ , participera à la réalisation du *stress*, encore une fois nettement plus marquée en St3 qu'en St2. Gimson rappelle que l'intensité est portée principalement par les sons voisés, pour lesquels une plus grande amplitude de vibration des plis vocaux est renforcée par la résonance des cavités supra-glottales : c'est le résultat d'un effort de souffle et musculaire plus soutenu dans l'articulation du son. Toutefois, il met en garde le lecteur contre le fait de ne s'en remettre qu'à ce seul critère pour déterminer la nature ou la place d'un accent dynamique en anglais, la réalisation contrastive de l'intensité par un locuteur pouvant varier grandement en fonction de la situation et de l'attitude du locuteur. Les conditions dans lesquelles nos enregistrements ont été effectués (phrases lues) devraient nous prévenir de ce risque. Gimson nous invite à ne considérer que les variations d'intensité majeures.

Nous utiliserons l'intensité sonore, ainsi que tous les autres critères ci-dessus, avec discernement, seuls puis en combinaison entre eux, afin de nous approcher d'une description aussi objective et précise que possible de la réalité des faits de parole étudiés.

---

<sup>18</sup> Cf. **Annexe 2** pour un exemple de *script* utilisé

**Partie III.**

-

**Méthodologie expérimentale**

### ***Construction du corpus Ense3 2012***

L'une de nos premières démarches a été de nous employer à la construction d'un corpus, en vue de probables travaux de recherche à venir, entre janvier 2012 et mai 2012

Un total de soixante-quinze étudiants de 1<sup>ère</sup> et 2<sup>ème</sup> année de l'école d'ingénieurs Ense3<sup>19</sup> s'est prêté aux tâches pédagogiques dédiées au rythme, et a été enregistré.

La toute première démarche a été de faire remplir aux étudiants un formulaire de consentement de participation.<sup>20</sup> Puis ils ont rempli un questionnaire disponible en ligne, concernant leur passé linguistique.<sup>21</sup>

Ensuite, les étudiants ont passé une première fois le test de lecture<sup>22</sup> dans un laboratoire audionumérique (échantillonnage de haute qualité à 32kHz), sans aucune préparation, au début janvier de 2012. S'en est suivi un travail spécifique de quatre mois, soit environ dix séances de cours, à raison de trente minutes à une heure par cours, basé sur l'entraînement à la production orale et au cours duquel bon nombre de tâches d'expression étaient proposées avec un guidage du tempo par le métronome.

Pour terminer, les étudiants passent le même test de lecture de 12 phrases, mais en fin de semestre, sans métronome. L'idée est d'écouter et de mesurer ce qui reste du travail effectué une fois le métronome enlevé. Les productions des étudiants peuvent alors être analysées et comparées dans les versions initiales et terminales du test, notamment à l'aide de l'outil logiciel PRAAT, selon la méthode décrite ci-après, afin de déterminer s'il y a eu progrès.

Pour ce qui concerne le présent mémoire, nous nous sommes intéressés à l'étude d'un cas issu de ce corpus, celui de Jean, qui sera analysé en détail avec le but de servir de modèle pour le protocole d'une éventuelle étude de plus grande envergure.

Enfin, un groupe « témoin » de douze étudiants hors-méthode est constitué. Ce groupe a passé le même test dans les mêmes conditions et aux mêmes périodes, mais n'a effectué aucun travail avec le métronome.

---

<sup>19</sup> École nationale supérieure de l'énergie, de l'eau et l'environnement (Ense3), école d'ingénieurs de l'Institut polytechnique de Grenoble

<sup>20</sup> Cf. **Annexe 3** pour un exemplaire du formulaire.

<sup>21</sup> Cf. **Annexe 1** pour un exemplaire du questionnaire.

<sup>22</sup> Phrases du test visibles en figure 1

## ***Méthodologie utilisée pour l'analyse des phrases sous PRAAT : segmentation***

Quatre paramètres principaux sont à considérer dans la réalisation du contraste accentuel : la hauteur mélodique, *pitch* ; l'intensité, *loudness* ; la qualité, *quality* ; la quantité, *quantity* (Gimson, 2008). Nous retiendrons ces quatre paramètres comme pertinents et suffisants pour la portée du présent travail. Le sujet sera développé ci-après. Le logiciel d'analyse et de traitement du signal de parole PRAAT<sup>23</sup> nous permet d'appréhender ces 4 paramètres, soit par lecture directe, soit par l'intermédiaire de traitements informatiques appelés *scripts*, qui facilitent et/ou automatisent le recueil de données.

Chaque phrase considérée a fait l'objet d'une segmentation manuelle sous PRAAT, cette segmentation permettant, à l'aide des traitements *scripts*<sup>24</sup> appropriés d'extraire les données nécessaires à notre analyse sous forme de tableaux. Le tri et la réalisation de graphiques d'interprétation en sont grandement facilités. Les créateurs du logiciel, dans sa version en anglais, utilisent le terme imagé *tier* (couche, strate, niveau) pour désigner l'étage de travail créé par l'utilisateur. Nous utiliserons le terme également imagé « tiroir » en français, origine étymologique du mot *tier*, chaque tiroir contenant un type spécifique de données. Ainsi, chaque niveau de notre segmentation correspond à un tiroir spécifique créé dans la fenêtre PRAAT. Cette segmentation a été effectuée de la manière suivante, dans treize tiroirs répartis en quatre niveaux et selon l'axe du temps en abscisse :

### **Niveau orthographique**

#### **1. Tiroir Text**

- Segmentation du signal en lexèmes
- Transcription orthographique des phrases lues

### **Niveau syllabique**

Segmentation du signal du locuteur en syllabes ; chaque tiroir suit la même segmentation syllabique :

#### **2. Tiroir Phonotrans : transcription phonétique vraie (API simple) des syllabes**

---

<sup>23</sup> Logiciel d'analyse et de traitement du signal de parole PRAAT : *Praat: doing phonetics by computer*, téléchargeable à partir du site néerlandais : <http://www.fon.hum.uva.nl>. Auteurs : P. Boersma and D. Weenink, de l'Université d'Amsterdam.

<sup>24</sup> Cf. **Annexe 2** pour un exemple de *script* utilisé.

3. Tiroir **PhoNorm** : transcription phonétique théorique/de référence des syllabes
4. Tiroir **StressSylProd** : indication du niveau de *stress* perçu
5. Tiroir **StressSylNorm** : indication du niveau de *stress* théorique/de référence de la syllabe dans la phrase
6. Tiroir **FocusSyl(0,1)** : indique la présence (1)/absence (0) de focus perçu sur la syllabe

#### **Niveau syntaxique**

7. Tiroir **BoundaryLvl(I,II,III)** : indique la présence de frontières limitant des espaces sans parole à l'intérieur du signal de parole : silences, pauses, respirations ; une valeur subjective de durée (I, II ou III) leur est donnée selon leur importance relative par rapport à l'ensemble des phrases

#### **Niveau vocalique**

Segmentation du signal du locuteur en ne conservant que les voyelles ; chaque tiroir suit la même segmentation vocalique :

8. Tiroir **FocusVow(0,1)** : indique la présence (1)/absence (0) de focus perçu sur la voyelle (valeurs identiques au focus syllabique ci-dessus)
9. Tiroir **StressVowNorm** : indication du niveau de *stress* théorique/de référence sur la voyelle dans la phrase (valeurs identiques au niveau syllabique équivalent)
10. Tiroir **StressVowProd** : indication du niveau de *stress* perçu porté sur la voyelle dans la phrase (valeurs identiques au niveau syllabique équivalent)
11. Tiroir **VowelRank** : donne une référence unique à la voyelle en fonction de son emplacement (ex. V5.12 = 12<sup>ème</sup> voyelle de la 5<sup>ème</sup> phrase)
12. Tiroir **V\_expected** : indique le son vocalique théorique/de référence attendu (API simple)
13. Tiroir **V\_produced** : indique le son vocalique vrai perçu (API simple)

## Remarques

- Concernant le focus, il n'y en a eu que très peu d'occurrences perçues, ce qui est parfaitement normal en présence de phrases déclaratives lues. Nous avons néanmoins fait le choix de le faire figurer, ne serait-ce que pour mémoire, compte tenu de son importance en situation de communication réelle.
- Le référencement des voyelles du tiroir *VowelRank* pourra s'avérer utile lors de la comparaison des voyelles similaires issues de différentes phrases entre elles. Ceci permet de les replacer immédiatement dans leur contexte pour, par exemple, mieux analyser l'influence potentielle de leur environnement immédiat ou leur position dans la phrase.

La segmentation du signal de parole n'est pas toujours chose aisée. La pose de frontières (*boundaries*) dans PRAAT relève parfois de choix arbitraires ou subjectifs. Les difficultés rencontrées concernent notamment les transitions consonnes-voyelles, quand les consonnes sont voisées et, surtout, en présence des approximantes (*glide*) de type /j/ et /w/, fréquentes en anglais (ex. dans *education* /,ed·ju·'keɪʃ·ən/). Concernant la syllabation, elle est sujette à de nombreuses controverses. En l'occurrence, nous suivons le modèle de syllabation conventionnel de Wells<sup>25</sup> chaque fois que possible, dont nous résumons ici quelques principes parmi les plus représentatifs (adoptés sur le critère de la meilleure prédiction de la distribution des allophones):

- existence d'une frontière entre les deux éléments d'un mot composé
- une consonne ne peut appartenir à deux syllabes à la fois
- une consonne rejoindra toujours la voyelle adjacente la plus accentuée (celle de gauche, dans le cas de deux syllabes St1)
- les contraintes phonotactiques de l'anglais seront respectées, à quelques exceptions justifiées près.

---

25 Wells 2008 : *Syllabification*, Introduction p. xxvii

## **Partie IV.**

-

**Analyse, interprétation et discussion des résultats**

## Introduction

Les résultats recueillis de l'étudiant, sujet de notre étude de cas et à qui nous avons donné le pseudonyme de « Jean », sont analysés ci-après selon trois axes principaux : 1) une seule phrase, la première de notre test, sera analysée de façon détaillée, dans ses versions initiale  $T_{init}$  et finale  $T_{term}$ , à l'aide de l'enregistrement de référence  $T_{ref}$  ; 2) les voyelles de quatre des douze phrases du test seront analysées de manière comparative ; 3) la composition spectrale de 4 voyelles dites « longues », réparties dans les 12 phrases du test, dans les versions  $T_{init}$  et  $T_{term}$ , sera étudiée.

Le tableau de la figure 2 contient les 12 phrases utilisées dans le test. Tout comme la phrase Ph01, les 3 autres phrases étudiées en 2) (Ph03, 05, 10) apparaissent sur fond clair. Le choix de ces 3 phrases sera justifié dans la partie qui leur est dédiée ci-après.

N°	Phrases	Accents de référence St3	Syllabes
01	A re* <b>form</b> in high*er e*du* <b>cati</b> *on <b>took</b> place this <b>aut</b> *umn and the sit*u* <b>ati</b> *on has im* <b>proved</b> , at <b>last</b> .	7	26
02	A rec*ent <b>surv</b> *ey has <b>shown</b> that <b>more</b> and more fem*ale <b>teen</b> *agers <b>opt</b> for scien*tif*ic <b>stud</b> *ies.	6	22
03	Eng*i* <b>neer</b> *ing stud*ies in* <b>volve</b> a <b>lot</b> of <b>work</b> on <b>proj</b> *ects.	5	15
04	<b>Each</b> new <b>class</b> of <b>stud</b> *ents is sur* <b>veyed</b> in <b>terms</b> of ac*a* <b>dem</b> *ic a* <b>chieve</b> *ments.	7	19
05	The <b>rec</b> *ord of em* <b>ploy</b> *ment of <b>grad</b> *u*ates going <b>through</b> is <b>ver</b> *y good.	5	17
06	Un*i* <b>vers</b> *it*y au* <b>thor</b> *it*ies <b>make</b> con* <b>cert</b> *ed <b>eff</b> *orts on <b>stud</b> *ent <b>well</b> -be*ing.	7	21
07	Each <b>spring</b> , <b>those</b> who <b>have</b> a good <b>mus</b> *ic*al <b>lev</b> *el <b>give</b> a <b>conc</b> *ert.	7	16
08	At the <b>start</b> of the <b>year</b> , the di* <b>rect</b> *ors pro* <b>ject</b> an inf*or* <b>mati</b> *on <b>slide</b> -show.	6	19
09	This <b>graph</b> <b>shows</b> <b>how</b> our <b>stud</b> *ents' <b>or</b> *ig*in and <b>age</b> are re* <b>lat</b> *ed, in one <b>piece</b> of inf*or* <b>mati</b> *on.	9	24
10	The <b>year</b> be* <b>fore</b> had <b>been</b> quite <b>ord</b> *in*ar*y, but a <b>drop</b> is re* <b>cord</b> *ed this <b>month</b> , and it's <b>sharp</b> !	8	23
11	On <b>av</b> *er*age, we <b>lose</b> <b>three</b> per <b>cent</b> of <b>boys</b> per <b>ann</b> *um — that's <b>true</b> !	7	16

12	On that <b>chart</b> , we can <b>see</b> <b>how</b> they <b>choose</b> their <b>fut</b> *ure <b>place</b> of <b>stud</b> *y <b>too</b> .	8	17
----	--	---	----

Figure 2. Les 12 phrases étudiées selon l'ordre original dans les tests  $T_{init}$ ,  $T_{term}$  et  $T_{ref}$ . L'astérisque (\*) indique le découpage syllabique acoustique (approximatif, en alphabet courant) des mots (selon Wells, 2008). Les syllabes soulignées portent l'accent de référence St3. Les phrases étudiées de façon plus détaillée (01, 03, 05, 10) apparaissent non-grisées.

Considérons un instant la phrase Ph01 :

*A reform in higher education took place this  
autumn and the situation has improved, at last.*

Cette phrase faisant l'objet d'un travail approfondi ci-après, nous allons la considérer un instant. Pour ce qui concerne son origine, nous renvoyons le lecteur aux explications de notre partie précédente traitant de la construction de notre corpus. Du point de vue morphologique, nous remarquons que cette phrase présente une articulation principale en deux parties autour de la conjonction de coordination *and* et se termine par l'apposition *at last* séparée du reste de la phrase par une virgule. Du point de vue prosodique, ces deux marques majeures correspondront quasi-systématiquement à des pauses, peut-être à des respirations ; elles participeront directement au phrasé, dans le sens musical du terme. Les trois parties ainsi délimitées (nommons les A, B et C) constituent des entités qui auraient pu être indépendantes. Cette relative complexité délibérée de Ph01 nous permettra de mettre certains aspects de l'interlangue de notre sujet en relief.

A reform in higher education took place this autumn and the situation has improved, at last.
  
A B C

Sans vouloir procéder à une analyse syntaxique détaillée, arrêtons-nous brièvement sur la composition de la phrase et considérons-la du point de vue de la distribution de référence du *stress*. La partie A, par exemple, peut elle-même être subdivisée en syntagmes nominaux (A1 et A3) répartis de part et d'autre autour du syntagme verbal A2. Le syntagme nominal A1 est lui-même composé de 2 sous-syntagmes nominaux A1<sub>a</sub> [article + nom] et A1<sub>b</sub> [préposition + adjectif + nom] de composition différente. A3 se décline en [démonstratif + nom]. Le syntagme verbal A2 est un verbe composé, une expression verbale.

A reform | in higher education | took place | this autumn  
 A1<sub>a</sub>                      A1<sub>b</sub>                      A2                      A3

Considérons maintenant Ph01 du point de vue de la distribution du *stress*<sup>26</sup>, en l'absence de focus.

- Le groupe verbal, portera quasi-systématiquement un accent de type St3. Ici l'accent St3 tombera sur la partie purement verbale *took*, point remarquable puisqu'il s'agit d'une syllabe de type V<sub>short</sub>, moins courant que V<sub>long</sub> ou V<sub>diph</sub> sous ce niveau de *stress*.
- Un déterminant recevra un St1, sauf mise en relief exceptionnelle (focus) ; l'article indéfini *a* se prononcera /ə/.
- Un nom, un « substantif », en position isolée (sans la proximité immédiate d'autres noms) recevra par défaut un St3 (') sur l'une de ses syllabes.
- L'adjectif, sans focus particulier, recevra un St2 (,) en présence d'un nom.
- Les mots-outils (ex. coordinateur *and*, prépositions *in* ou *at*) recevront généralement le type St1.
- Pour les auxiliaires de conjugaison (d'ailleurs très fréquemment « contractés » en anglais), il en va le plus souvent de même : St1, en présence du verbe.
- Les syllabes saillantes, sous St3, sont alternativement de type V<sub>long</sub> (re'**form**) ou V<sub>diph</sub> (,edu'**ca**tion) ou V<sub>short</sub> ('**took** ,**place**), comme vu plus haut.
- Les syllabes sous St2 sont représentées par des V<sub>diph</sub> (,high**er** ; ,place) ou V<sub>short</sub> (,situ'**a**tion).
- Les syllabes St1 concentrent naturellement les V<sub>short</sub> et des V<sub>weak</sub>.
- Enfin, la présence de deux mots à structure accentuelle similaire, « education » et « situation », mais dans des environnements différentes (préposition + adjectif + nom // article + nom) pourra fournir une intéressante comparaison.

---

<sup>26</sup> Distribution du stress selon Wells (2008).

Voici le rendu accentuel et phonémique de référence de Ph01 :

A re'form in ,higher edu'cation 'took ,place this 'autumn (...)

ə ri 'fɔ:m in ,haɪ\_ə ,ed ju 'keɪʃ ən 'tɒk ,pleɪs ðɪs 'ɔ:t əm |

(...) and the ,situ'ation has im'proved, at 'last.

ənd ðə ,sɪtʃ u 'eɪʃ ən həz ɪm 'pru:vɪd | ət 'lɑ:st

Cette phrase de référence comprend donc 26 syllabes, avec 7 syllabes accentuées de type St3, et 3 de type St2, contenant la plus grande variété de types de voyelles. Elle est peut être considérée comme représentative a) de situations parmi les plus courantes en matière de rapports d'accentuation entre les différents éléments lexicaux et b) de la variété de la distribution phonémique sous les 3 niveaux de *stress*. Une analyse fine des productions d'anglophones/LN pourrait révéler des variations plus ou moins importantes de ce schéma accentuel théorique, bien entendu ; mais ce n'est pas l'objet de notre étude. Comparons et analysons maintenant en détail la phrase Ph01 produite par Jean, notre sujet, selon les différents angles d'approche et paramètres évoqués, dans les versions T<sub>init</sub> et T<sub>term</sub>.

## Chapitre 1. — Analyse détaillée des lectures de la phrase Ph01 du test par notre sujet

Dans un premier temps, nous étudierons de manière détaillée les données perceptives du *stress* de la première phrase (Ph01) dans ses deux versions issues des lectures de Jean puis, plus longuement, les données issues des mesures objectives. Nous déterminerons quelles sont les approches les plus pertinentes, notamment en matière de description et de mesure du progrès accompli par Jean et, par une extrapolation prudente, par des LnN apprenant l'anglais. Enfin, nous terminerons cette partie par l'analyse comparative de la composition spectrale de voyelles sélectionnées parmi les 12 phrases du test.

### *Niveaux de « stress » perçus : données perceptives-subjectives*

#### **Introduction**

Le graphique présenté ci-dessous (figure 3) rassemble les trois niveaux de *stress* : un « de référence » et deux perçus (Stress\_Init et Stress\_Term), dans les deux versions du test de lecture rythmique : initial ( $T_{init}$ ) et terminal ( $T_{term}$ ) pour cette première phrase Ph01.

Pour  $T_{init}$ , aucun conseil, ni consigne de type linguistique n'avaient été donnés quant à la lecture des phrases ; l'enregistrement s'est déroulé lors de la toute première séance de cours (janvier 2012). Pour  $T_{term}$ , il a eu lieu lors de la dernière séance du semestre (mai 2012), suite à un travail rythmique (entre trente minutes et une heure par semaine sur dix semaines, approximativement).

Les niveaux de *stress* de  $T_{init}$  de  $T_{term}$  sont ceux que l'auteur-enseignant a relevés perceptivement, à l'écoute des enregistrements. Ils ne correspondent à aucune mesure objective de niveau d'intensité, de hauteur mélodique ou de durée. Ils sont donc à apprécier avec toutes les précautions d'usage. Toutefois, nous prenons le parti de considérer qu'ils sont suffisamment fiables pour fournir des indicateurs utiles, dans le meilleur des cas, de valeurs de *stress* approximantes et, dans le pire, de tendance. Notre objectif sera par la suite de comparer ces valeurs perçues à celles des différentes composantes du *stress* que nous avons mesurées (durée, hauteur mélodique, intensité) et à en analyser les différentes convergences et divergences.

## Phrase 01

Voici le graphique de la perception de la réalisation syllabique du *stress* pour la phrase 01 (Ph01). Les 26 syllabes de cette phrase sont en abscisse ; les niveaux d'accentuation (St1 à St3) sont en ordonnée.

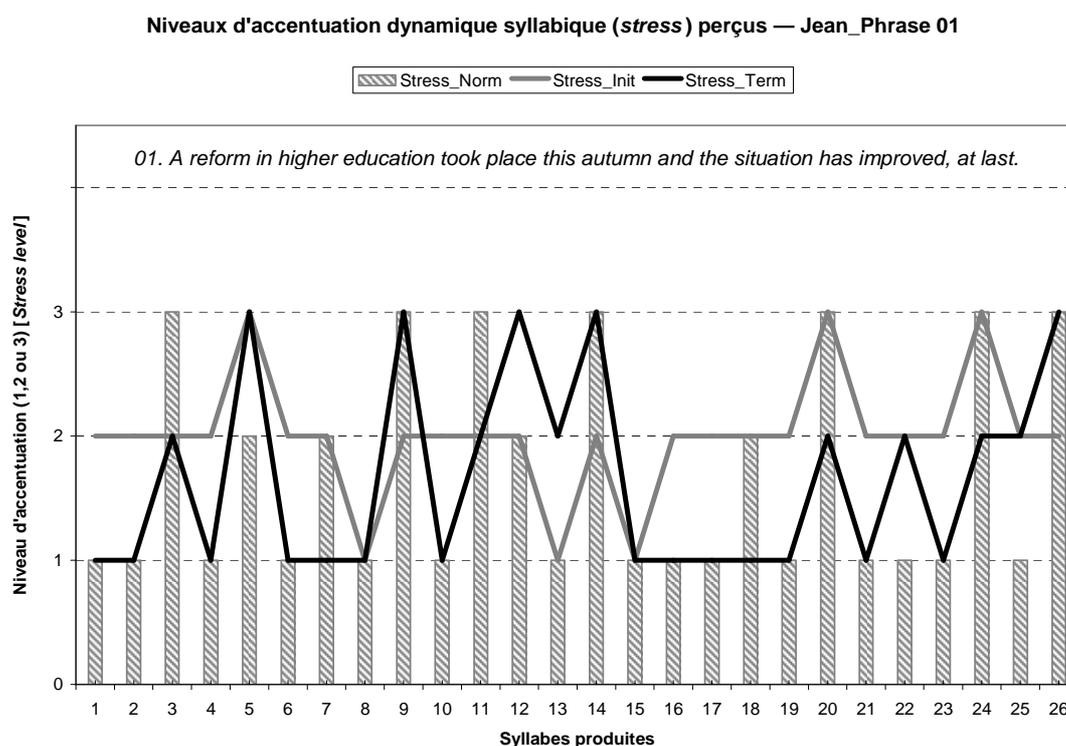


Figure 3. Niveaux de *stress* perçus de la phrase 01. L'historgramme indique le niveau de *stress* de référence.

À la lecture attentive de ce graphique reprenant les 26 syllabes de la phrase en abscisse, nous remarquons que le nombre théorique-de référence de St1 est sensiblement supérieur aux St2-St3 réunis, et donc très supérieur aux St3 seuls, comme cela est statistiquement le cas. Considérons brièvement l'historgramme ci-dessous (figure 4) qui résume de façon plus concise la distribution du *stress* pour les 3 versions de cette phrase :

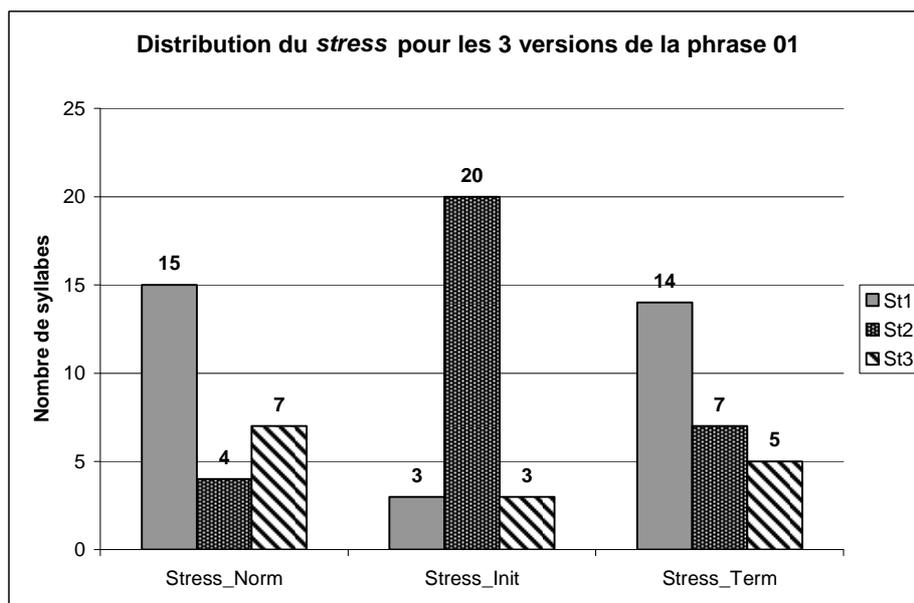


Figure 4. Niveaux de *stress* de référence et perçus pour les 26 syllabes de la phrase 01.

Nous trouvons d'emblée trois points remarquables dans le graphique de la figure 4 :

- groupe Stress\_Norm : les St1 de références dominent et dépassent en nombre le cumul St2 + St3
- groupe Stress\_Init : la distribution des *stress* initiaux est radicalement à l'opposé de celle de référence, avec une très forte majorité de St2
- groupe Stress\_Term : l'histogramme se rapproche fortement du groupe Stress\_Norm

La très forte dominante de St2 chez notre sujet en  $T_{init}$  pourra s'expliquer par la persistance de la prosodie de la langue française (« accent français ») à ce stade. En effet, le français, caractérisé par sa tendance à produire une isochronie syllabique, correspondra plus naturellement à une rythmique régulière, basée sur des accents homogènes et omniprésents, de type St2 de niveau « moyen ». Ceci est donc relativement classique pour l'interlangue d'un francophone en phase d'apprentissage.

Pour  $T_{term}$ , les résultats quantitatifs se rapprochent très fortement des valeurs de référence. Toutefois, au-delà de ces chiffres, il nous faut maintenant observer la répartition du *stress* dans la phrase (cf. figure 3) et voir si elle correspond à une logique, à une stratégie de la part de notre sujet, à un « progrès linguistique » : une occurrence du *stress*, quel qu'en soit le niveau, ne signifie pas qu'il tombe nécessairement au « bon endroit », c'est-à-dire à l'endroit où il apportera le sens voulu, à la fois par le locuteur et par les

contraintes communicatives de la langue (rappel : l'accent lexical en anglais n'est pas « libre »). Les effets d'un accent mal placé peuvent être au moins aussi négatifs qu'une absence totale de contraste.

Ainsi, la figure 3 ci-dessus nous permet de comparer instantanément, pour chaque syllabe, la réalisation du *stress* perçue à sa valeur de référence, en versions  $T_{init}$  et  $T_{term}$ . À ce stade, 2 niveaux de lecture peuvent se faire :

- une lecture « grossière », où nous considérerons l'existence, ou pas, d'un accent contrastif de niveau St2 ou St3 par rapport à St1 ; l'existence d'une tendance
- une lecture « fine », où nous considérerons la répartition précise du niveau de *stress*.

Concernant la courbe *Stress\_Init*, fortement centrée autour des St2, nous constatons que, hormis en début et en fin de phrase, elle suit approximativement les tendances de variation de *Stress\_Norm*, toutefois en atteignant très rarement les niveaux de contraste attendus : les St3 ne sont atteints à bon escient que 2 fois sur 7, en syllabes 20 et 24 (S1.20 ; S1.24) ; les St1, eux, ne sont réalisés que 3 fois sur 15, ce qui est très peu (S1.8, S1.13 et S1.15). Enfin, les réalisations initiales et terminales contredisent les valeurs et les tendances de variation de *Stress\_Norm*. Ceci peut avoir deux causes principales : le fait d'avoir affaire à une tâche de lecture peut modifier la prosodie ; l'influence de l'intonation du français déclaratif, généralement montante au début et descendante en conclusion de phrase : l'expression anglaise *at last* réclame une valeur de St3 sur *last* incontournable, qui n'est pas réalisée.

Pour ce qui concerne la courbe *Stress\_Term*, la première constatation porte sur le contraste : en première moitié de phrase, notamment, le contraste est très fort, avec une oscillation fréquente entre les extrêmes St1 et St3. Les tendances de variation sont toutes globalement respectées, sauf en S1.7, 18, 22. Concernant les valeurs, pour les St3, 4 sur 7 sont placés à l'endroit attendu, et 12 sur 15 pour les St1.

Pour conclure, dans la comparaison des courbes *Stress\_Init* et *Stress\_Term* :

- nous constatons des courbes relativement dissemblables, qui se retrouvent en variation 14 fois sur 26 (montée ou descente simultanée), mais qui ne se rejoignent que 7 fois sur 26 en valeur de *stress*

- les valeurs réalisées de St1 et St3 sont différentes entre les tests : elles doublent en Stress\_Term pour St3 (28% à 57% par rapport à Stress\_Norm) et elles quadruplent pour St1 (20% à 80% par rapport à Stress\_Norm)
- dans Stress\_Term le début et la fin de la phrase sont beaucoup plus cohérents avec l'accentuation de référence attendue
- la 2<sup>ème</sup> moitié de la phrase (hormis les 2 syllabes de fin) était cependant mieux réalisée en Stress\_Init

Ainsi, du point de vue perceptif de l'enseignant, nous pouvons dire qu'il y a eu de nets progrès en Stress\_Term, par rapport à Stress\_Init, dans cette première phrase : davantage de contraste accentuel, davantage de réalisations des niveaux de *stress* attendus, une meilleure corrélation entre les courbes de référence et terminale, qu'en initiale. Ci-après, nous allons considérer ce qu'il en est à partir de nos mesures objectives des durées, vocalique, consonantique et syllabique, de la hauteur mélodique (F<sub>0</sub>) et de l'intensité : nous comparerons alors nos résultats avec les données perceptives ci-dessus et, dans la mesure du possible, déterminerons les progrès de notre sujet, en qualité et en quantité. Auparavant, considérons brièvement nos transcriptions du rendu acoustique perceptif-subjectif de la lecture de cette phrase par Jean, dans ses trois versions, dans une analyse phonémique comparative succincte.

### ***Qualité vocalique de la phrase Ph01 : données perceptives-subjectives***

Le tableau ci-dessous (figure 5) reprend l'intégralité de la transcription simplifiée de la phrase 01 (n'apparaissent que les faits les plus marquants) figurant sur les axes des abscisses de nos graphiques. Ceux que nous commentons sont sur fond clair (éloignement de la référence) ou noir (rapprochement). Il s'agit plus de qualité vocalique que de durée ici ; nous avons vu qu'une relation forte existe entre qualité et durée en anglais, tant dans la réalisation des sons longs (sons riches, pleins) que dans les courts (neutralisation).

Dans T<sub>init</sub>, persisteront certains sons issus du français, L1<sup>27</sup> de Jean :

- #6 – *higher* prononcé ['haɪ\_œ]
- #16 – *and* prononcé [æ̃n]
- #18 – absence du /tʃ/ dans *situation*

---

<sup>27</sup> L1 = langue maternelle

- #22 – omission du /h/ dans *has*
- #3, 24 et 26 – sons intermédiaires ou courts [ə], [u] et [a] au lieu de sons longs dans *reform*, *improved* et *last*
- #22 et 25 – absence de réduction vocalique /ə/ dans *has* et *at*.

La plupart des sons de la L1 ont disparu dans  $T_{\text{term}}$ . Cependant, quelques uns y sont apparus alors qu'ils n'étaient pas dans  $T_{\text{init}}$  :

- #8 – [y] au milieu de *education*
- #9 et 20 – absence de diphtongues dans *education* et *situation*, par exemple.

Item#	Phonotrans_Init	PhoNorm	Phonotrans_Term
1	ə	ə	ə
2	ri	ri	ri
3	fəm	'fɔ:m	'fɔ:m
4	ɪn	ɪn	hɪn
5	'haɪ	'haɪ	'xai
6	æ	ə	ə
7	ed	ed	ed
8	y	ju	ju
9	'keɪf	'keɪf	'kef
10	ən	ən	ən
11	tɔk	tɔk	tu(k)
12	pleɪs	pleɪs	'pleɪs
13	ðɪs	ðɪs	ðɪs
14	aot	'ɔ:t	'xɔt
15	hən	əm	ən
16	æn	ænd	æn
17	ðə	ðə	ðə
18	sɪt	sɪtʃ	sɪtʃ
19	u	u	y
20	'eɪf	'eɪf	'eɪf
21	ən	ən	ən
22	az	həz	az
23	ɪn	ɪn	ɪn
24	'pru:vɔd	'pru:vɔd	'pru:vɔd
25	at	ət	at
26	last	'lɑ:st	'lɑ:st

Figure 5. Transcription phonémique des 26 syllabes de Ph01, en  $T_{init}$ ,  $T_{ref}$ ,  $T_{term}$ . Les items sur fond blanc s'éloignent de la référence, ceux sur fond noir s'en rapprochent.

Il nous faut également remarquer l'apparition de sons semblant issus de l'allemand ou de l'espagnol (L3 de Jean), surtout dans  $T_{term}$  : le [x] aux #5 et 14 et la prononciation [aøthen] de *autumn*, en  $T_{init}$ . Tout ceci est tout à fait normal chez un étudiant pratiquant deux langues étrangères à un niveau intermédiaire.

En revanche, de façon apparemment surprenante, le travail effectué en cours a déplacé tous les points de l'interlangue de Jean où il y a rapprochement de la référence : certains items qui étaient plutôt bien réalisés en  $T_{init}$  ne le sont plus (régression du sujet), alors que d'autres ont été nettement améliorés en  $T_{term}$  (progression). Ces items où la

réalisation tend vers la référence sont sur fond noirs dans notre tableau. Nous postulons ici qu'au cours de tout apprentissage, dont celui des langues, il faut parfois « désapprendre » pour progresser et que les ressources cognitives, mobilisées par l'objet en cours d'étude, peuvent être momentanément insuffisantes pour gérer l'ensemble des paramètres : en se focalisant sur des points particuliers, l'attention est accaparée par les nouveaux acquis et les anciens acquis sont momentanément écartés. Parler « correctement » une langue étrangère est un exercice extrêmement mobilisateur des ressources, tant que les réflexes langagiers ne sont pas acquis ; ceci peut en être une illustration. Comme nous l'avons mentionné au début de ce travail, l'un des fondements de notre approche consiste à déplacer la production linguistique du domaine de la réflexion vers celui du réflexe. Dans le cas présent, c'est le progrès réalisé dans le domaine rythmique qui est probablement à l'origine d'une partie des changements phonémiques.

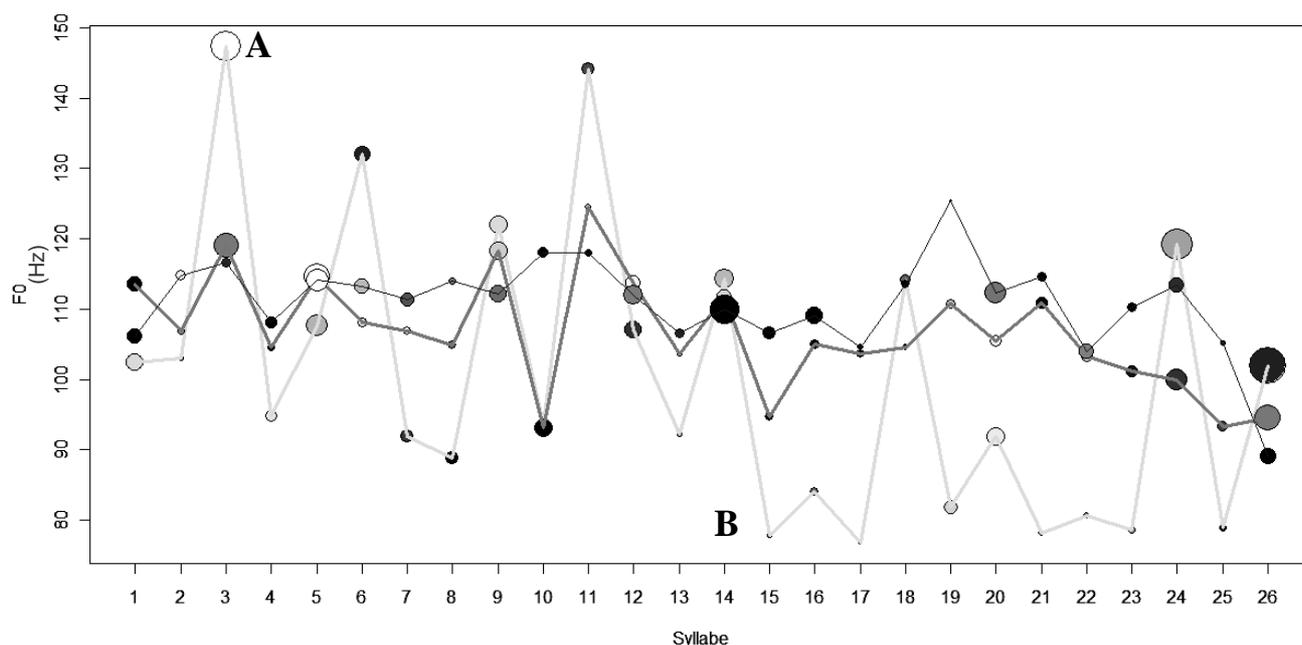
### ***Phrase Ph01 : mesures objectives, analyse et interprétation***

#### **Introduction**

Après une approche perceptive, étudions la phrase Ph01 à l'aide des mesures objectives effectuées : de durée, de hauteur mélodique et d'intensité. Cette étude approfondie nous permettra de mettre en relief les paramètres et les méthodes d'analyse les plus pertinents.

Le graphique synthétique ci-dessous (figure 6) permet l'observation simultanée des 3 paramètres influençant le *stress* décrits en introduction : durée (ms), hauteur mélodique ( $F_0$ , en Hz), et intensité (dB). Il cumule, sous la forme de trois courbes, les données fondamentales pour notre étude de Ph01 : celle du test initial ( $T_{init}$ ), du test terminal ( $T_{term}$ ) et du test de référence ( $T_{ref}$ ) expliqué ci-après, en fonction de chacune des 26 syllabes. Peut-être difficile à lire de prime abord, il nous est utile pour illustrer la complexité de la tâche que nous avons entreprise. Nous fixant pour objectif une analyse fine de nos résultats, il nous sera indispensable d'avoir recours, en complément, à une approche plus fragmentée.

Evolution de la fréquence, de la durée et de l'intensité au cours de la phrase 01



Ph01. A reform in higher education took place this autumn and the situation has improved, at last.

Légende :

- ○ Court    ○ Long — Durée V : taille proportionnelle à la durée
- ● Fort / ● Faible — Intensité V : niveau de gris proportionnel à l'intensité
- T<sub>init</sub>                      — T<sub>term</sub>                      — T<sub>ref</sub>

Figure 6. Evolution de la fréquence, de la durée et de l'intensité au cours de la phrase 01, pour les trois tests.

### Création d'un « test de référence »

Au cours de l'analyse de nos données, nous avons éprouvé le besoin d'introduire une référence nous permettant de pouvoir comparer et interpréter les résultats de notre sujet à un modèle, analysé selon les mêmes critères. Nous rappelons ici que cette démarche référentielle était déjà à la base de notre approche : les tâches d'apprentissage proposées aux étudiants étaient orientées vers un modèle de référence, celui du dictionnaire de prononciation de J.C. Wells (Wells, 2008).

Les 12 phrases composant le test présenté aux étudiants au début et à la fin de l'apprentissage ont été lues et enregistrées par leur professeur, dans les mêmes conditions techniques. Le modèle parfait et unique n'existant pas en matière de langues vivantes, que ce soit chez les LN ou les praticiens experts de l'anglais LnN, nous considérerons arbitrairement que le professeur constitue un modèle de référence acceptable, compte tenu

de ses qualifications, de ses compétences et de son expérience. Nous dénommerons ce test  $T_{ref}$ . Il ne s'agit donc pas ici à proprement parler d'un « test », puisque le professeur qui lit les 12 phrases n'est pas lui-même testé : l'abréviation  $T_{ref}$  est utilisée par simple commodité et cohérence.

### **Analyse succincte du graphique de la figure 6**

Considérons les points remarquables A et B figurant sur la courbe  $T_{ref}$  du graphique, à titre d'exemple. Ceux-ci permettent d'illustrer visuellement 2 extrêmes, sans toutefois les quantifier précisément à ce stade. Tous deux sont portés sur la courbe de la courbe  $T_{ref}$ , celle présentant les plus forts contrastes. Le point A (V1.3) cumule une fréquence fondamentale et une intensité élevées et une durée prolongée. À l'inverse, le point B (V1.15) concentrera une fréquence et une intensité faibles et une durée des plus réduites. Ainsi, cette représentation permet de mettre en évidence l'évolution du contraste accentuel de la phrase.

Si nous considérons maintenant le graphique dans son ensemble, nous remarquerons de prime abord que l'amplitude de la courbe de hauteur mélodique issue de  $T_{ref}$  est approximativement deux fois plus importante que pour  $T_{init}$  et  $T_{term}$  et que ces dernières sont elles-mêmes relativement dissemblables dans leurs variations. Ensuite, que l'intensité semble être distribuée de façon à peu près équitable sur l'ensemble des courbes, tout comme la durée ; les durées plutôt « longues » se trouvent cependant plutôt au dessus de la zone des 100 Hz de  $F_0$ .

Un examen un peu plus poussé révélera que sur certaines parties des courbes, l'évolution de  $T_{term}$  par rapport à  $T_{init}$  va dans le sens d'un rapprochement vers  $T_{ref}$ , ce nous considérerons comme un progrès de l'apprenant ; c'est le cas pour le début de la phrase par exemple, jusqu'à V1.5, ou de V1.7 à V1.17 où, même si l'amplitude de hauteur mélodique reste très différente, les courbes  $T_{term}$  et  $T_{ref}$  varient de la même manière. Lorsque les courbes  $T_{init}$  et  $T_{term}$  diffèrent, sans que  $T_{term}$  se rapproche de  $T_{ref}$  ou qu'elle la contredise (V1.20), cela pourra le plus souvent être considéré comme une tentative infructueuse, un essai de l'apprenant, ce qui est caractéristique de la phase d'apprentissage où l'on peut essayer de « bien faire » sans toutefois y parvenir. Notre sujet n'est pas « parfait ». Dans le cas du *stress*, le déplacement malencontreux de celui-ci sur une syllabe adjacente aura immédiatement un effet contre-productif sur la compréhension du message par un LN : au

pire, une incompréhension ; au mieux, une compréhension erronée (un exemple très courant est la différenciation des chiffres : à l'oral, ce qui fera par dessus tout la différence entre « 50 » et « 15 » sera le déplacement du *stress* (*stress shift*) de la première à la dernière syllabe, *fifty* /'fɪfti/ et *fifteen* /'fɪf 'ti:n/). De tels graphiques, dans une version simplifiée, pourraient être utilisables à des fins pédagogiques, dans le cadre de l'apprentissage de la langue. Procédons maintenant aux analyses des mesures de durée.

### ***Mesures de durées brutes de la phrase Ph01***

#### **Introduction**

Les graphiques de cette première partie concernent nos mesures de durées vocaliques et syllabiques. Nos observations montrent que la durée est la composante du *stress* capable de la plus grande variation, l'amplitude de nos valeurs relevées dans les différentes productions de Jean allant de presque 0 ms à plus de 300 ms pour les voyelles et à plus de 600 ms pour les syllabes. Rappelons que, lors de l'interprétation de la durée des voyelles dans l'étude du *stress*, nous devons être vigilant face à la coexistence possible de  $V_{\text{long}}$  et de  $V_{\text{short}}$  sous les niveaux St2 et St3 ; nous y reviendrons. Notre but principal, ici comme dans l'ensemble de nos autres analyses ci-après (hauteur mélodique, intensité) sera de déterminer si a) nous sommes en présence d'instances de *stress* ou pas ; b) en cas de présence, si celui-ci est contrastif (différents niveaux de *stress*) ; c) s'il est utilisé à bon escient, ou pas.

A priori, une analyse comparée des voyelles et des syllabes semble pertinente : la longueur de la voyelle est discriminante en anglais ; elle servira par exemple à faire la différence entre *bit* (morceau) et *beat* (pulsation). La voyelle est la partie de la syllabe la plus « élastique » : dans la voix criée ou chantée, la durée de la voyelle peut atteindre des valeurs extrêmement longues, alors que la consonne ne variera que très peu. Il pourra être néanmoins intéressant d'observer le comportement de la partie consonantique des syllabes. Nous déclinerons nos graphiques de durées en différentes versions de type  $T_{\text{init}}$ ,  $T_{\text{term}}$  et  $T_{\text{ref}}$ . Nous procéderons à des comparaisons de ces graphiques entre eux de façon à mettre en valeur les éventuelles variations, évolutions et corrections des productions de notre sujet. Enfin, nous procéderons à l'application de différentes formules de calcul simples destinés à mettre en valeur les stratégies potentiellement appliquées par Jean pour faire évoluer son interlangue, notamment du point de vue accentuel et, par conséquent, rythmique.

### Graphiques de durée brute : voyelles et syllabes

Les graphiques ci-dessous (figures 7 et 8) affichent la durée brute mesurée des 26 voyelles et des 26 syllabes de la phrase 01 (Ph01), enregistrée dans les deux versions du test, initial  $T_{init}$  avant tout travail rythmique en cours de langue, et terminal  $T_{term}$ , à l'issue d'une douzaine de séances de ce travail rythmique d'environ une heure par semaine. L'axe des abscisses reprend les phonèmes produits. En ordonnée, nous avons la durée exprimée en millisecondes (ms). La distance comprise entre les deux courbes à un point donné représente la durée de la part consonantique de la syllabe ; nous y reviendrons dans une prochaine partie.

Plusieurs aspects sont remarquables d'emblée sur le premier graphique (figure 14) ; quelques constations au niveau des courbes nous serviront pour les suivants :

1. La durée des voyelles (V) et des syllabes (S) n'évolue pas en parallèle : l'amplitude de la variation est beaucoup plus importante pour les syllabes. Les consonnes fluctuent donc elles aussi de façon importante
2. Dans sa variation, la courbe de V suit plutôt les tendances des mouvements de celle de S, mais avec des exceptions (ex. items #1 à 4).
3. Des points remarquables : les deux courbes S et V se rejoignent en 5 endroits, sur les items #1, 6, 8, 16 et 19 : la syllabe n'est composée que d'une voyelle, sans consonnes.
4. L'amplitude de variation de la durée entre V et S s'étend de 0 ms ( $V = S$ , cf. ci-dessus) à 404 ms, sur la dernière syllabe, #26.
5. Le contraste de variation des durées est relativement important, avec une  $V_{min}$  et une  $S_{min}$  autour de 20 ms ; une  $V_{max}$  à près de 230 ms et une  $S_{max}$  à plus de 550 ms.
6. La durée, de façon plus marquée pour la courbe S, varie en « dents de scie » le plus souvent, avec une alternance quasi-régulière par endroits
7. Hormis le #7 (valeur la plus basse) et le #26 (valeur la plus haute), l'ensemble des pics de la courbe S se situe dans une fourchette relativement étroite, entre les valeurs de pics de 270 ms (#20) et 400 ms (#24).

Durée brute\_Voyelles et syllabes\_Test initial — Jean\_Phase 01

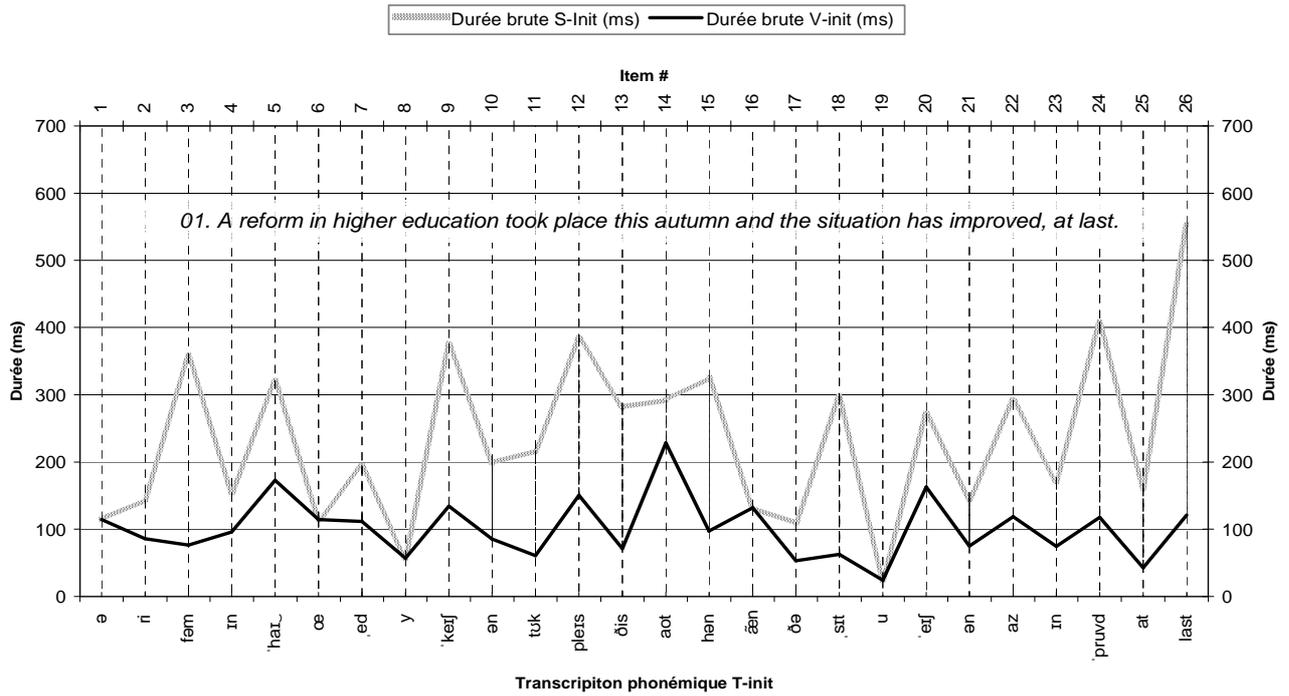


Figure 7. Durée brute des voyelles et des syllabes mesurée pour les 26 items de Ph01, en  $T_{init}$

Durée brute\_Voyelles et syllabes\_Test terminal — Jean\_Phase 01

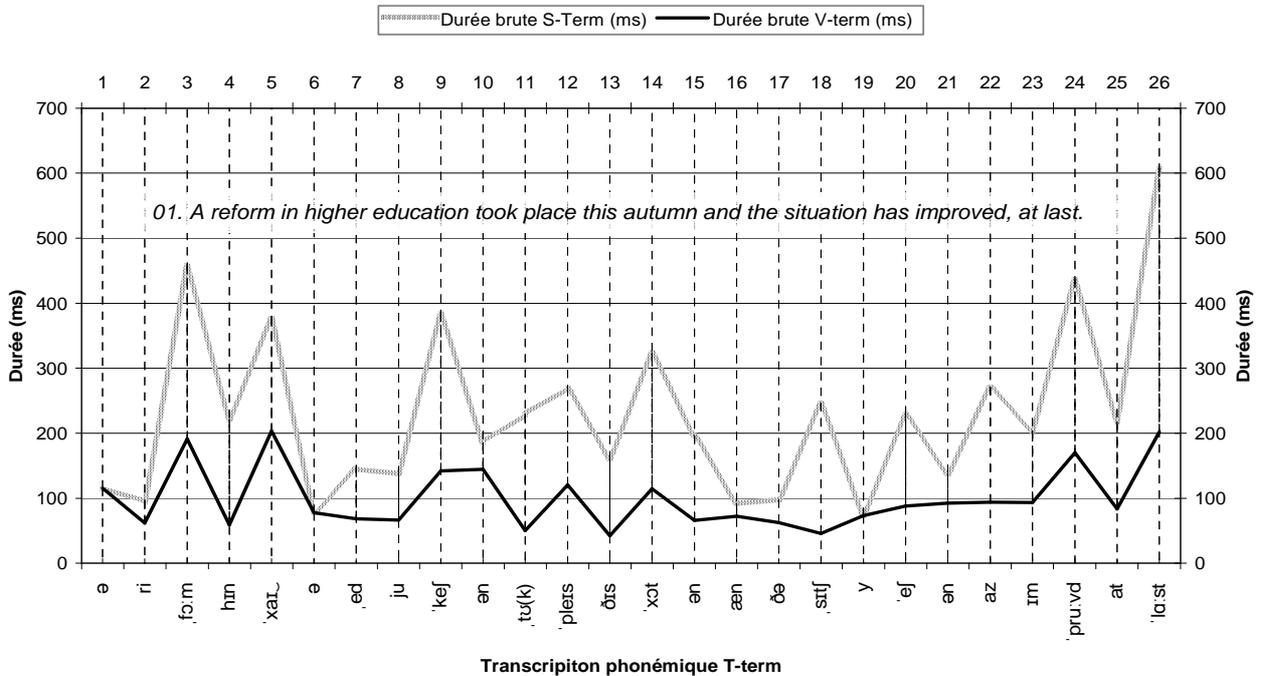


Figure 8. Durée brute des voyelles et des syllabes mesurée pour les 26 items de Ph01, en  $T_{init}$

Après ces premières constatations, comparons ce premier graphique de durée en  $T_{init}$  à deux autres, à la même échelle de valeurs (0-700 ms), dans les versions  $T_{term}$  et  $T_{ref}$ . La figure 8 représente le graphique similaire en version  $T_{term}$ . Les silhouettes des deux graphiques présentent des similitudes indéniables : c'est la même phrase qui est lue. Cette similitude s'avère toutefois beaucoup plus grande pour les syllabes que pour les voyelles. Pour les courbes S, les différences principales portent, dans le passage de  $T_{init}$  à  $T_{term}$  :

- sur l'amplitude positive : en début de phrase (items #3 et 5 : +80 à 100 ms) et en coda (#24 et 26 : + 50 à 60 ms) : une plus grande durée ira dans le sens d'une meilleure réalisation du *stress* de niveau St3, ainsi qu'un meilleur rendu des  $V_{long}$  et en S5-V5, sous St2, dans le sens d'une  $V_{diph}$  potentiellement améliorée.
- sur la réalisation du grand intervalle #6 à #15, avec amélioration de la prononciation sur le #8 par l'introduction de l'approximante /j/ : les courbes S et V se séparent
- dans ce même intervalle : allongement en #14 ( $V_{long}$  sous St3) ; raccourcissements importants en #12 (-110 ms : insuffisant, mais allant dans le sens d'une amélioration), #13 (-120 ms : progrès sur une  $V_{short}$  sous St1) et #15 (-130 ms ; progrès :  $V_{short}$  sous St1)

Les courbes des voyelles présentent, quant à elles, d'importantes différences :

- intervalle des items #1 à 6 : pics de durées importants, inexistantes en  $T_{init}$ , renforçant la durée des V3 et V5 : allongement en progrès sous St3 et St2
- #6 à 10 : séparation des courbes S et V (cf. ci-dessus)
- #18 à 23 : nivellement total des  $V_{term}$  : cohérent en présence du type  $V_{short}$  dans toutes ces syllabes
- #24 et 26 : allongement en progrès sous St3
- en coda, tendance à la baisse en  $V_{init}$ , à l'allongement en  $V_{term}$  : bonne adéquation avec la fin de la phrase en apposition (même sans présence de focus)

Pour poursuivre cette partie, nous proposons de comparer les deux graphiques et séries de commentaires ci-dessus avec le graphique de la phrase Ph01 dans sa version  $T_{ref}$  (figure 9 ci-dessous). Nous avons choisi de faire apparaître les valeurs de référence du *stress* des 26 syllabes de Ph01 sous la forme d'une pseudo-courbe dont l'intérêt principal est de montrer le parallèle qui existe entre les variations des niveaux de *stress* et les variations des durées. Ceci est patent pour la courbe des syllabes, mais encore plus flagrant pour celle des voyelles. Nous pouvons ainsi visualiser l'existence d'une corrélation entre les deux paramètres, qui varient quasi simultanément.

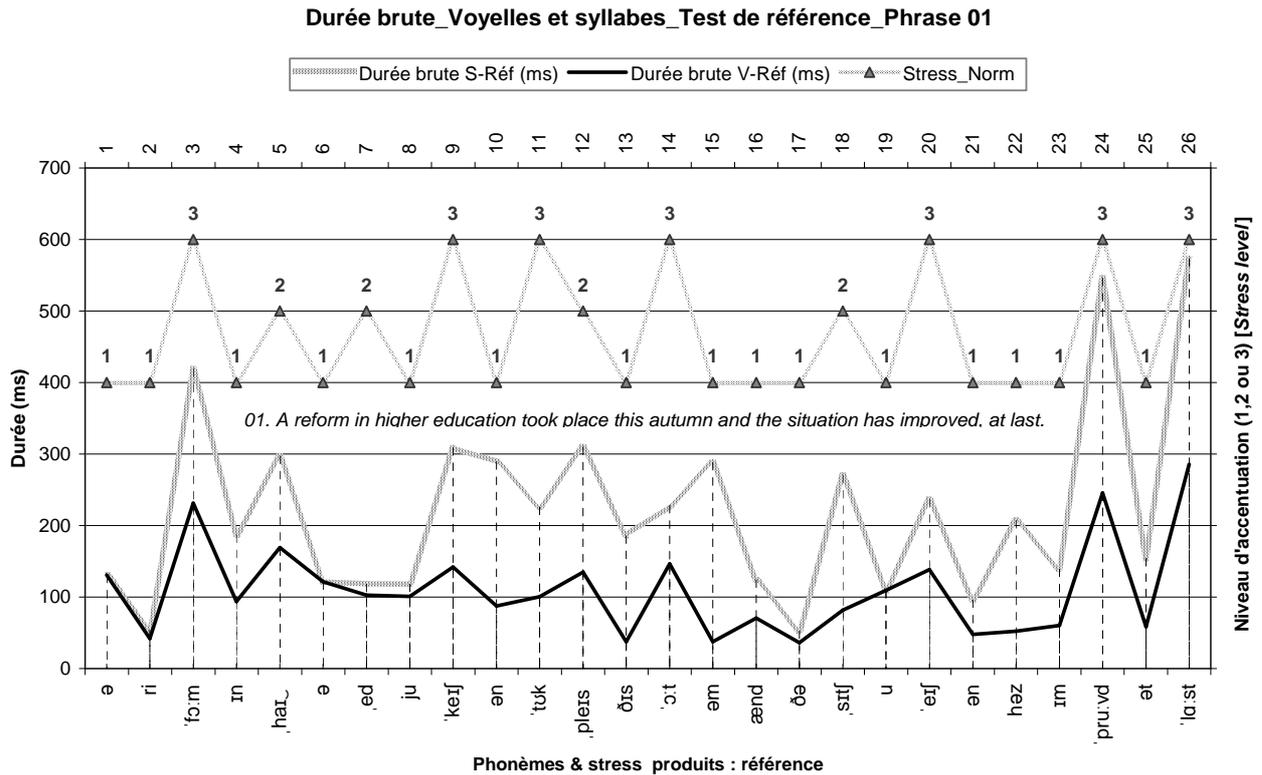


Figure 9. Durée brute des voyelles et des syllabes mesurée pour les 26 items de Ph01, en  $T_{ref}$

Comparons maintenant les courbes V et S de la phrase Ph01 en  $T_{ref}$  avec celles de  $T_{init}$  et  $T_{term}$ .

- Sur la partie des items #1 à 6, les pics et les creux sont nettement plus marqués en  $T_{term}$  et  $T_{ref}$  qu'en  $T_{init}$  : le contraste de durée est plus important (de 50-70 à 420-450 ms en  $T_{ref}/T_{term}$  contre 80 à 350 ms en  $T_{init}$ ) ; l'allure générale des deux courbes est beaucoup plus proche qu'en  $T_{init}$ .
- L'intervalle suivant, du #6 au #17 est un peu moins net et une analyse fine serait nécessaire
- Pour la fin de la phrase, du #17 au 26, les courbes se rapprochent de  $T_{ref}$  en  $T_{term}$ , sans toute fois atteindre la très grande variation de  $T_{ref}$  sur les 4 derniers items. Rappelons ici la divergence dans le sens de variation des courbes S et V en queue de phrase dans  $T_{init}$  : les courbes tendent à se séparer, ce qui n'est pas le cas dans les autres.

Pour terminer cette partie, nous finirons par l'examen de deux graphiques S et V séparés, dans lesquels nous avons superposé les trois courbes de durée, en  $T_{init}$ ,  $T_{term}$  et  $T_{ref}$ . Celui de la figure 10 ci-dessous représente les courbes de syllabes.

### Durée brute\_Syllabes\_Jean\_Phase 01

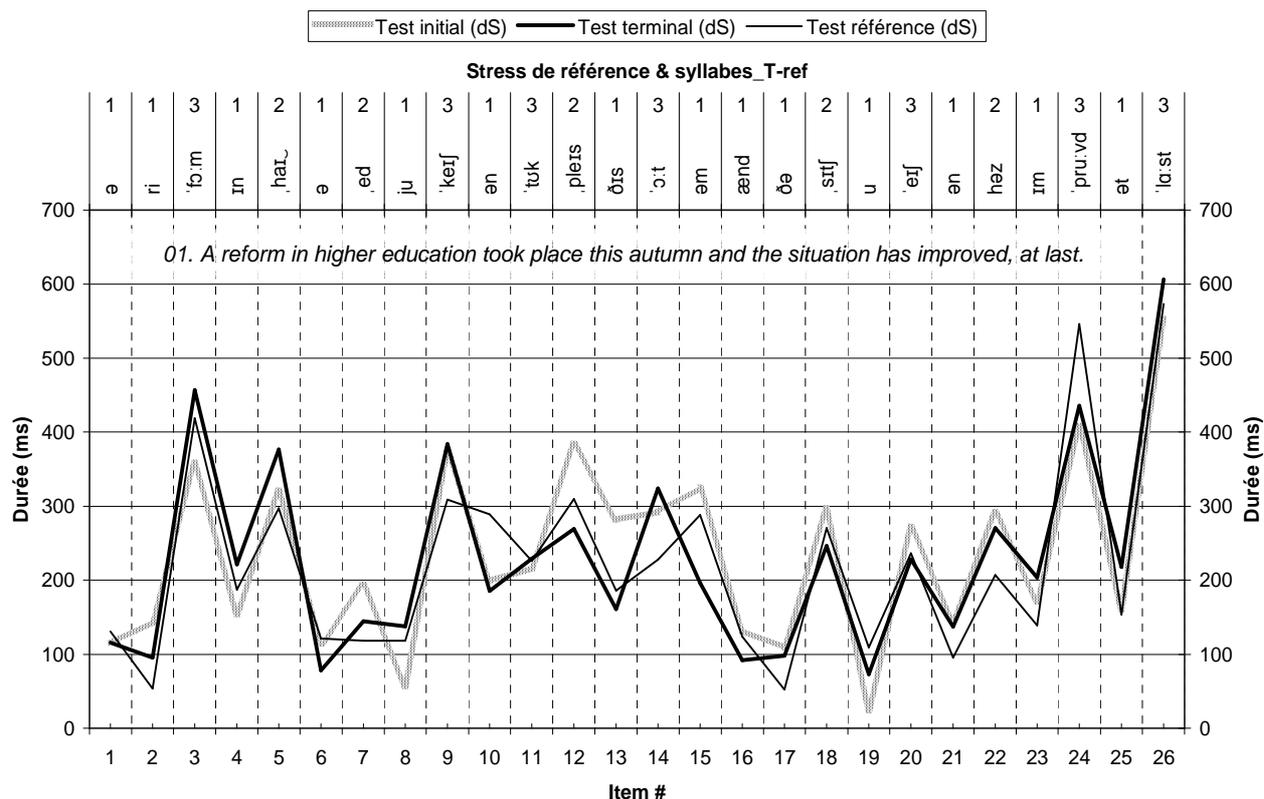


Figure 10. Durée brute des **syllabes** mesurée pour les 26 items de Ph01, en  $T_{init}$ ,  $T_{term}$  et  $T_{ref}$ .

Ce qui apparaît tout de suite dans ce graphique, c'est la grande similitude des trois courbes, qui s'entrecroisent tour à tour sans qu'a priori l'on puisse y déceler une tendance forte. Elles ne diffèrent que très peu, sauf en items #6 et à 8 et #9 à 15. Notons quelques points remarquables, néanmoins.

- L'évolution du contraste de durée dans la séquence #2 à 3, de  $T_{init}$  à  $T_{term}$ , puis de  $T_{term}$  à  $T_{ref}$  : l'amplitude augmente entre chaque test, vers le haut et vers le bas
- L'intervalle #6 à 8 présente quelques disparités.  $T_{init}$  est le plus dissemblable de  $T_{ref}$ .
- La gestion de l'intervalle du #11 à 15 : *took place this autumn* : différence prononcée entre  $T_{init}$  et  $T_{term}$ .  $T_{init}$  est plus en accord avec  $T_{ref}$  que  $T_{term}$ . Nous avons vu l'influence de la L3 espagnol pratiquée par Jean plus haut, de deux façons différentes, qui elles-mêmes diffèrent de l'anglais.

Voici maintenant les courbes superposées des trois tests dans un graphique similaire, mais restreint uniquement aux voyelles. Pour plus de lisibilité, nous avons séparé les courbes  $T_{init}$  et  $T_{term}$  et les présentons sur 2 graphiques séparément (figure 11 et 12). Si les 2 courbes se rapprochent en valeur et/ou en tendance (ce qui a plus d'intérêt ici, compte tenu de la présence de deux locuteurs différents) sur certains intervalles en  $T_{init}$ , nous constatons aisément que, dans son ensemble, la courbe  $T_{term}$  est globalement beaucoup plus proche de  $T_{ref}$ . Une analyse fine viendra nuancer ce propos.

- Intervalle #1-4 :  $T_{init}$  et  $T_{ref}$  varient à l'opposé, alors qu'ils sont « collés » en  $T_{term}$
- #4-6 : les 3 courbes varient quasi simultanément ; noter que la variation est supérieure au modèle en  $T_{term}$ , tout en restant très comparable en aspect
- #7-12 : entrecroisement des 3 courbes tour à tour
- #13-17 :  $T_{term}$  est proche, mais légèrement en deçà de  $T_{ref}$  en contraste ;  $T_{init}$  est relativement parallèle à  $T_{term}$ , mais va très au-delà en valeur.
- #18-23 :  $T_{init}$  présente beaucoup plus de contraste de durée que  $T_{term}$  et  $T_{ref}$ , mais ce contraste contredit souvent  $T_{ref}$  (sauf en #20-21).  $T_{term}$  présente une correction par rapport à  $T_{init}$ , mais il manque le contraste présent dans  $T_{ref}$ .
- #24-26 :  $T_{term}$  se rapproche beaucoup plus de  $T_{ref}$ , alors que  $T_{init}$  s'en éloigne.

Durée brute\_Voyelles\_T-init/T-ref — Jean\_Phase 01

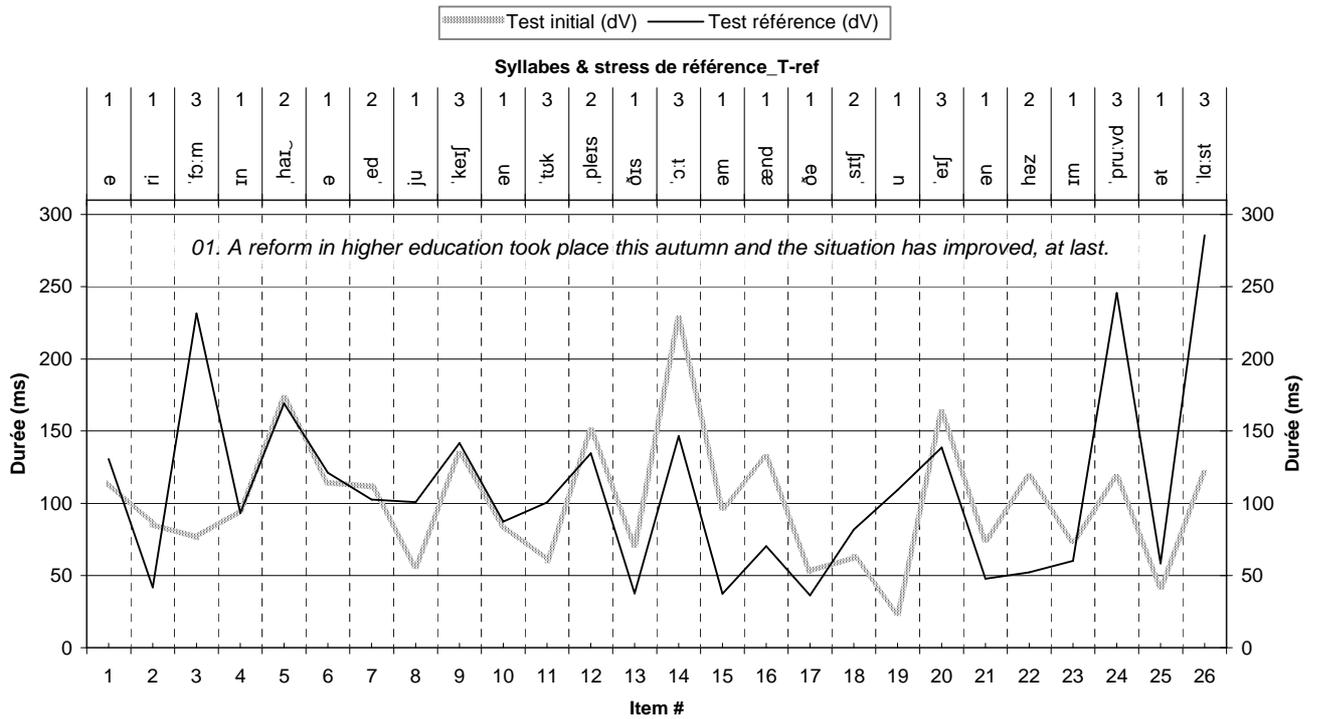


Figure 11. Durée brute des voyelles mesurée pour les 26 items de Ph01, en  $T_{init}$  et  $T_{ref}$ .

Durée brute\_Voyelles\_T-term/T-ref — Jean\_Phase 01

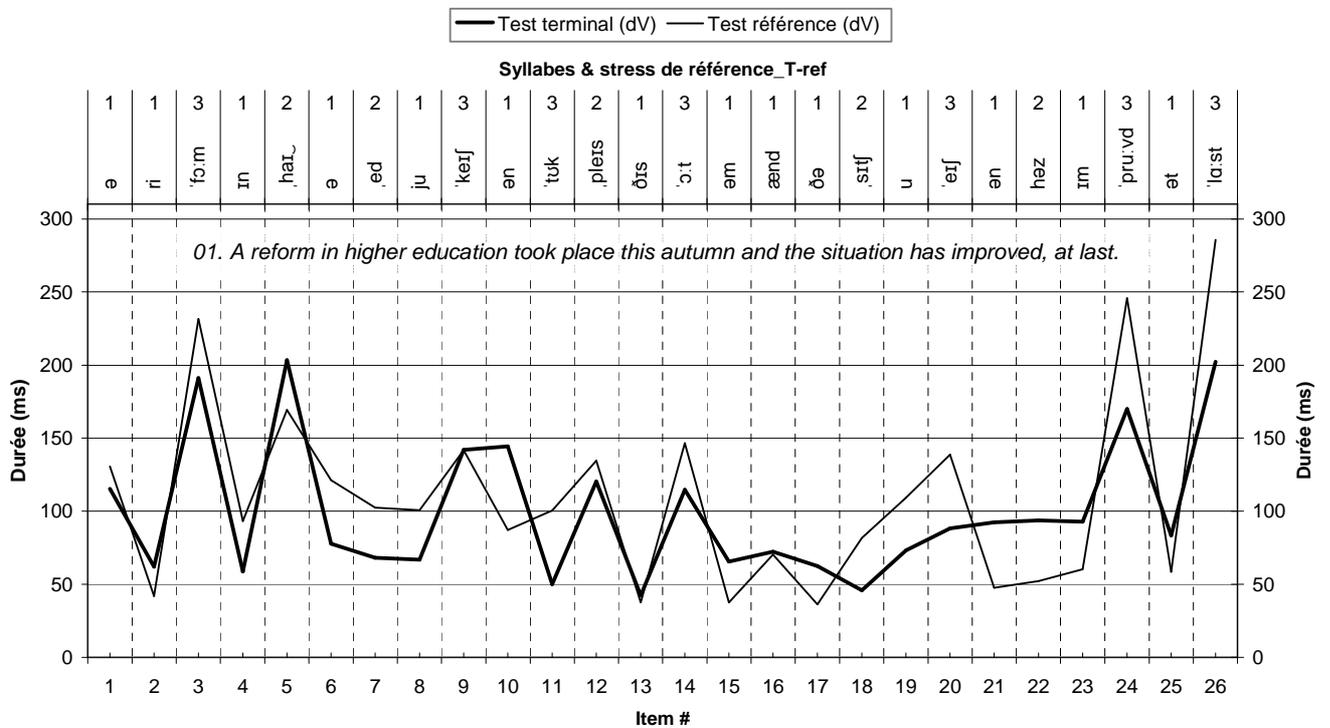


Figure 12. Durée brute des voyelles mesurée pour les 26 items de Ph01, en  $T_{term}$  et  $T_{ref}$ .

Ainsi, pour ce qui concerne les variations de durée, nous pouvons tirer deux enseignements essentiels des analyses ci-dessus. En premier lieu, si l'on compare les données relatives aux syllabes et aux voyelles, nous pouvons affirmer que ce sont bien les voyelles qui sont les plus susceptibles d'évolution : les courbes syllabiques se superposent alors que celles des voyelles sont disjointes, pour les mêmes items : les syllabes, en quelque sorte, « absorbent » les variations. Ceci met automatiquement en valeur le rôle joué par les consonnes, rôle qui n'est donc pas neutre ; nous y reviendrons plus loin. L'un de nos objectifs majeurs était de faire ressentir à notre étudiant la nécessité d'augmenter la durée des voyelles dites « longues », sous St3, par rapport à sa L1 pour se rapprocher du modèle de référence. C'est ce qui se produit effectivement dans cette phrase sur les items #3, #24 et #26. Note sujet apporte une correction bienvenue sur le #14 (*autumn*) en  $T_{\text{term}}$ , en réduisant la « surmodulation » due à sa prononciation très proche de son espagnol L3 en  $T_{\text{init}}$ .

Le deuxième terrain sur lequel nous visions une amélioration était celui du contraste de durée entre les voyelles « longues » ( $V_{\text{long}}$ ) et les « brèves » ( $V_{\text{short}}$ ), avec une production de  $V_{\text{long}}$  « plus longues » et  $V_{\text{short}}$  « plus brèves ». Sur ce terrain là, les réponses sont plus nuancées. Les points les plus positifs rejoignent naturellement les commentaires ci-dessus : c'est présence des  $V_{\text{long}}$  améliorées (#3, 24 et 26) que le contraste a le plus progressé. Nous notons ainsi un raccourcissement très marqué des  $V_{\text{short}}$  #2, 4, 6, 7, mais également des items #13, 15, 16 et 20. Toutefois, si le contraste de durée est parfois plus important en  $T_{\text{init}}$ , ce n'est pas toujours, comme nous l'avons vu, à bon escient.

#### **Conclusion : durée brute et « stress »**

Pour conclure cette partie, il nous faut répondre à la question de la corrélation entre *stress* et durée ou, dit autrement, de l'existence ou non d'un *stress* variant avec la durée, syllabique ou vocalique, ou encore, de l'existence d'un « *stress* de durée ». Nous pouvons confirmer sans ambages que la durée des syllabes, et des voyelles de ces syllabes joue un rôle déterminant dans le rendu de la phrase en anglais : l'apparence de nos courbes en dents de scie fait bien apparaître le contraste très grand qui existe entre les syllabes de type  $V_{\text{long}}$  et  $V_{\text{diph}}$  d'une part et  $V_{\text{short}}$  et  $V_{\text{weak}}$  d'autre part. La corrélation que nous avons mise en évidence dans le test  $T_{\text{ref}}$  entre la courbe accentuelle et celles des durées (figure 9 ci-dessus) est patente.

Jean, notre sujet, en faisant varier la durée des items syllabiques-vocaliques de la phrase Ph01, et en faisant évoluer cette variation, d'un point de vue global, vers un rapprochement du modèle de référence entre  $T_{init}$  et  $T_{term}$ , contribue de toute évidence à améliorer le rendu de son *stress*. Voyons maintenant ce qu'il en est de deux autres aspects dérivés de ces durées S et V, la quantité consonantique et la vitesse de parole et ses variations, avant d'aborder les deux autres paramètres fondamentaux que sont la hauteur mélodique et l'intensité.

### ***Différence dS – dV : quantité consonantique de la phrase Ph01***

#### **Introduction**

Après avoir traité de la durée des voyelles et des syllabes, nous abordons maintenant la durée des consonnes. Le procédé simple que nous avons retenu pour déterminer la quantité consonantique est celui de la soustraction de la durée de la voyelle à celle de la syllabe, puisque nous sommes déjà en possession de ces données. Regardons quel intérêt ce paramètre peut présenter dans l'analyse des productions de Jean, notre sujet.

Comme nous l'avons vu dans notre description des paramètres étudiés, la différence dS – dV nous renseigne sur la quantité consonantique de la syllabe (S) par soustraction de la durée de la voyelle (V). Cette quantité consonantique peut se répartir de plusieurs façons et être composée d'une ou plusieurs consonnes (C), contiguës ou pas. Nous l'avons déjà faite apparaître clairement dans les graphiques 7 à 9 de la partie précédente, alors que nous comparions syllabes et voyelles : c'est l'espace qui est créé entre la courbe des syllabes et celle des voyelles. Dans une approche simplifiée, nous considérerons que nous pouvons rencontrer quatre types majeurs de syllabes contenant des consonnes :

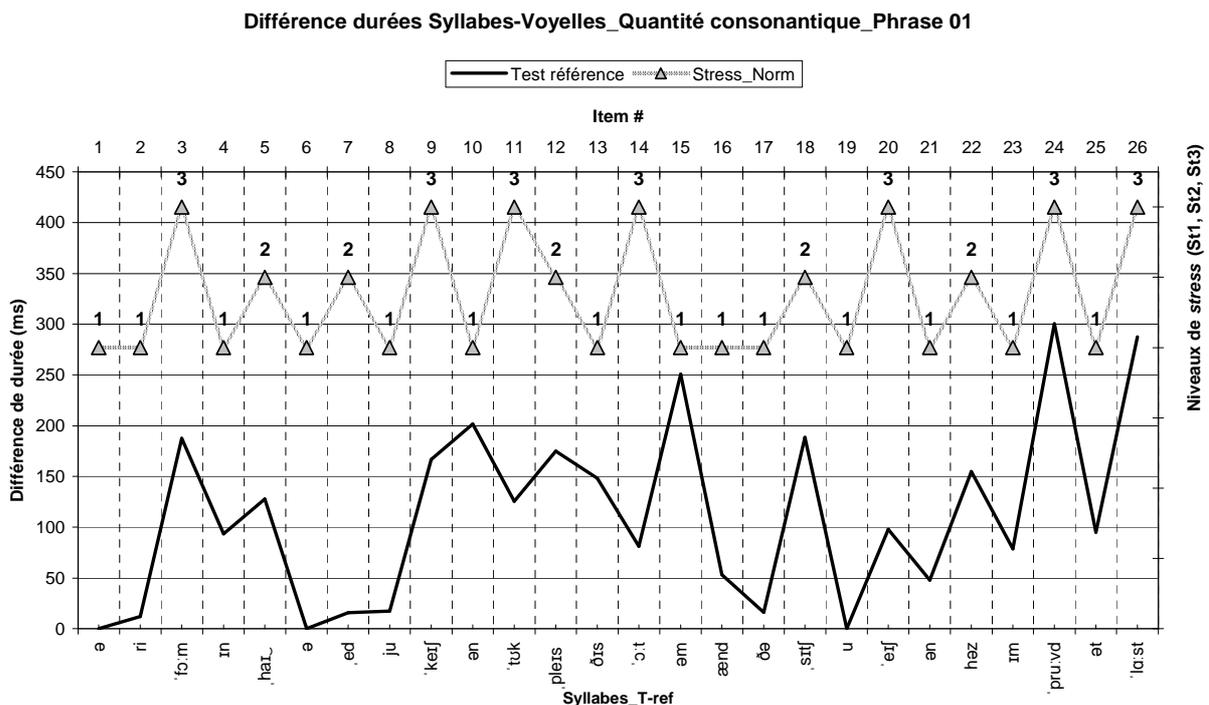
1. **CVC**
2. **CV**
3. **VC**
4. **C**

D'autres types plus complexes peuvent se rencontrer, notamment par l'accumulation de consonnes (*cluster*), par exemple la S24 /'pru:vd/ de la phrase Ph01 a pour structure CCVCC. A l'inverse, certaines syllabes ne contiennent pas de consonnes, ex. S1 ou S19 dans Ph01. Le cas #4 ci-dessus est particulier : généralement, les syllabes contiennent une voyelle ; toutefois, par le jeu de la neutralisation vocalique que nous avons déjà évoqué, en présence de débit rapide, de langue relâchée, ou de proximité de *stress* St3

particulièrement appuyé, il peut arriver que la voyelle sous St1 disparaisse totalement ; ce phénomène est relativement fréquent en anglais oral. Il nous intéresse tout particulièrement car il fait partie des difficultés rencontrées par nos étudiants et par les francophones en général, tant en perception qu'en expression. Dans le mot *situation* (S18 à S21) /,sɪtʃˈuːˈeɪʃən/, la S terminale /ən/ se réduira souvent à [n]. Certains dictionnaires (ex. Wells, 2008) notent cette possibilité par un phonème en position d'exposant, de la manière suivante /sɪtʃˈuːˈeɪʃˤn/.

### Analyse de la courbe de quantité consonantique de référence

Nous proposons la lecture du graphique  $T_{ref}$  (figure 13) pour commencer cette partie. Ce graphique superpose la courbe de quantité consonantique de référence avec la variation des niveaux de *stress* de référence pour chacune des 26 syllabes. Essayons de déterminer si une relation entre les deux représentations peut être établie.



01. A reform in higher education took place this autumn and the situation has improved, at last.

Figure 13. Quantité consonantique calculée pour les items #1 à 16 de Ph01, en  $T_{ref}$ , avec niveaux de *stress* de référence.

Les valeurs ne sont pas comparables entre elles ; par contre, un examen attentif de la variation montre une relation en tendances assez nette entre les deux courbes pour une majorité de syllabes. Seulement quatre syllabes contredisent apparemment cette variation commune : S10, 11, 14 et 15. Ces points d'intérêt s'expliquent aisément.

- S10 : présence de la nasale /n/ en coda du syntagme nominal précédant le syntagme verbal sous St3, articulation majeure de la phrase ; permet un voisement assez long (contrairement à ce que fait notre sujet, voir ci-après).
- S11 : chute de la durée sur une S pourtant de type CVC (2 consonnes) ; s'explique par la forte accentuation de niveau St3 combinée à la présence de la  $V_{\text{short}}$  /ʊ/ ; ce résultat avait été anticipé : c'est toute la syllabe qui est « raccourcie », pas seulement la voyelle
- S14 : chute de la durée sur une S de type VC ( $V_{\text{long}}$  + dentale non-voisée /t/) ; la consonne /t/ est sujette à brièveté en position médiane en anglais<sup>28</sup>
- S15 : idem S10 ci-dessus ; la nasale /m/ prolongée annonce l'articulation majeure de la phrase (coordination *and* en S16)

Ainsi, une corrélation entre la variation du *stress* et celle de la quantité consonantique se révèle exister pour le cas de cette phrase Ph01, les quelques contradictions pouvant être expliquées. Regardons maintenant comment les courbes de la production de notre sujet se comportent en présence de celle de  $T_{\text{ref}}$ .

#### **Analyse des courbes de quantité consonantique issues de $T_{\text{init}}$ et $T_{\text{term}}$**

Le graphique de la figure 14 représente la quantité consonantique issue de trois versions de la phrase 01, en provenance de nos enregistrements des tests  $T_{\text{init}}$ ,  $T_{\text{term}}$  et  $T_{\text{ref}}$ . La première constatation est que le graphique peut se diviser en 2 parties distinctes :

**A** — de S1 à S16 : les courbes suivent le même mouvement d'ensemble, mais de façon plutôt désordonnée, avec des croisements ;

**B** — de S17 à la fin : les courbes se rejoignent et varient quasi-simultanément.

De façon atypique, nous proposons de commencer en commentant la partie B, plus simple de par la cohérence des trois courbes ; nous nous intéresserons ensuite plus en détail à la partie complexe A.

---

<sup>28</sup> Elle est même fortement sujette à évolution vers [t̥], voire à élision dans cette position en anglais \$, notamment dans le registre de langue relâché.

### Différence durées Syllabes-Voyelles\_Quantité consonantique\_Phase 01

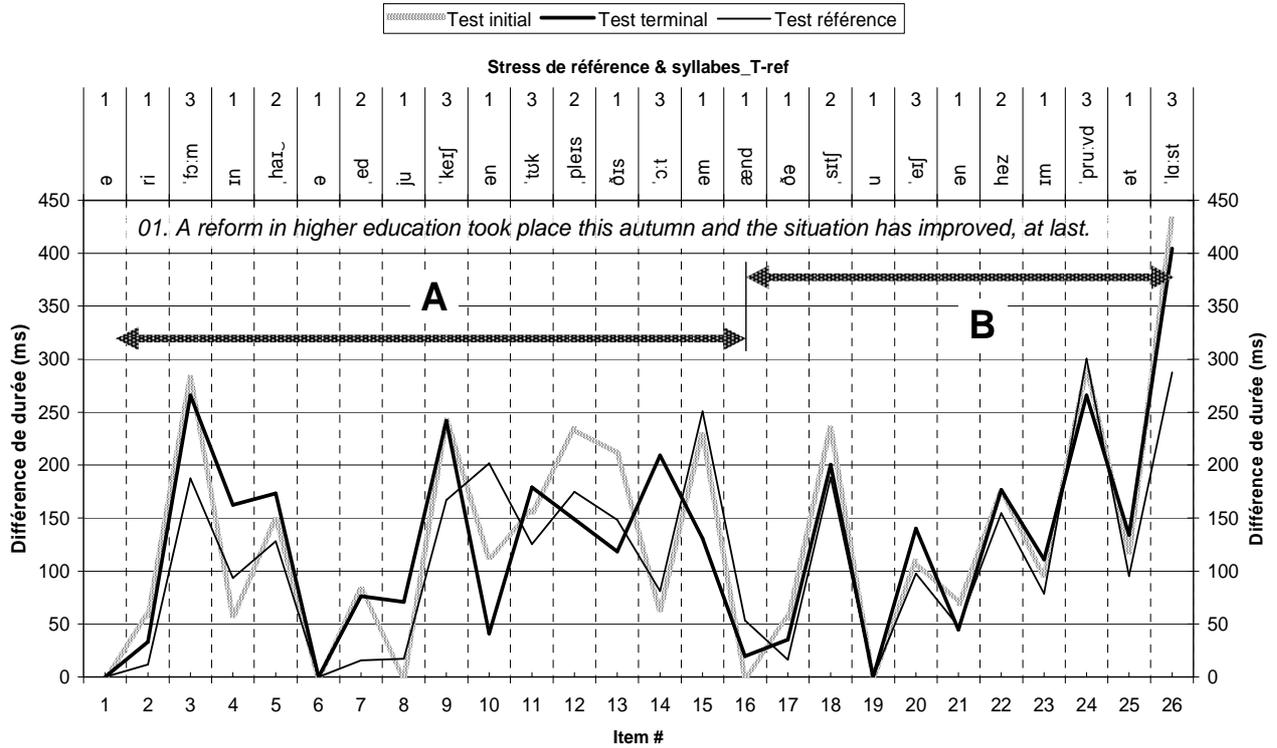


Figure 14. Quantité consonantique calculée pour les 26 items de Ph01, en  $T_{init}$ ,  $T_{term}$  et  $T_{ref}$ .

Le dernier tiers des trois courbes de quantité consonantique de Ph01, tel qu'il apparaît sur la figure 14 présente la particularité d'une cohérence quasi-parfaite entre elles. Un agrandissement de cette partie B du graphique (figure 15) nous permettra une lecture plus détaillée. Sur ce graphique 15, nous avons fait apparaître les niveaux de *stress* de référence de  $T_{ref}$ , à titre indicatif ; les niveaux de *stress* perçus en  $T_{init}$  et  $T_{term}$  apparaissent également sur les axes des abscisses.



Niveaux de *stress*\_Norm-Init-Term\_Ph01\_Items #17 à 26

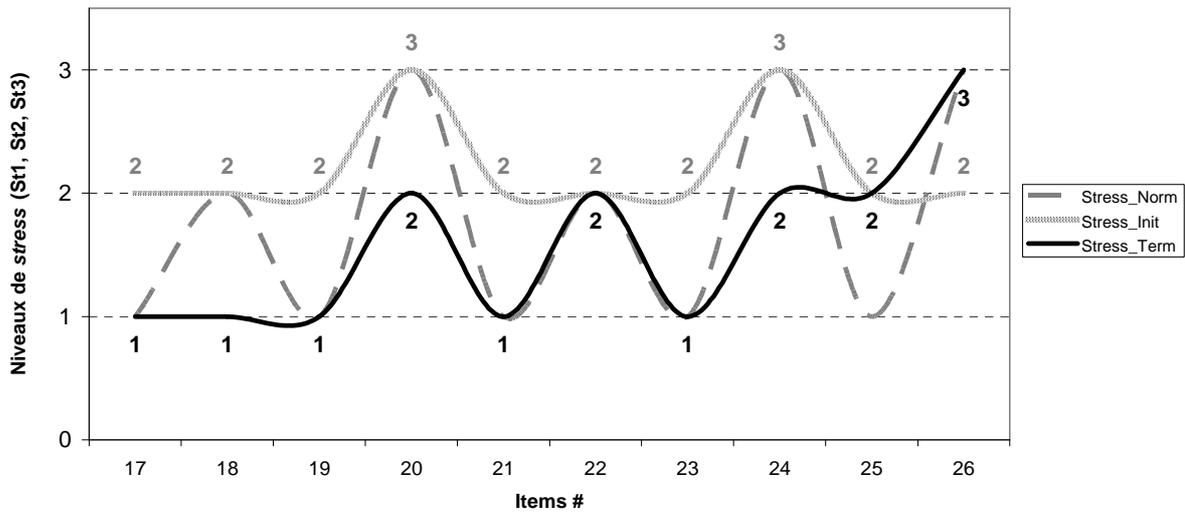


Figure 16. Niveaux de *stress* perçus pour les items #17 à 26 de Ph01, en  $T_{init}$ ,  $T_{term}$  et  $T_{ref}$ .

Considérons maintenant la première partie du graphique, de S1 à S16. Un changement d'échelle dans les figures suivantes permettra d'agrandir cette partie des courbes afin d'en faciliter la lecture. Nous commencerons par une superposition des courbes  $T_{init}$  et  $T_{ref}$  (figure 17).

### Différence durées Syllabes-Voyelles\_Phase 01\_Syllabes 1 à 16

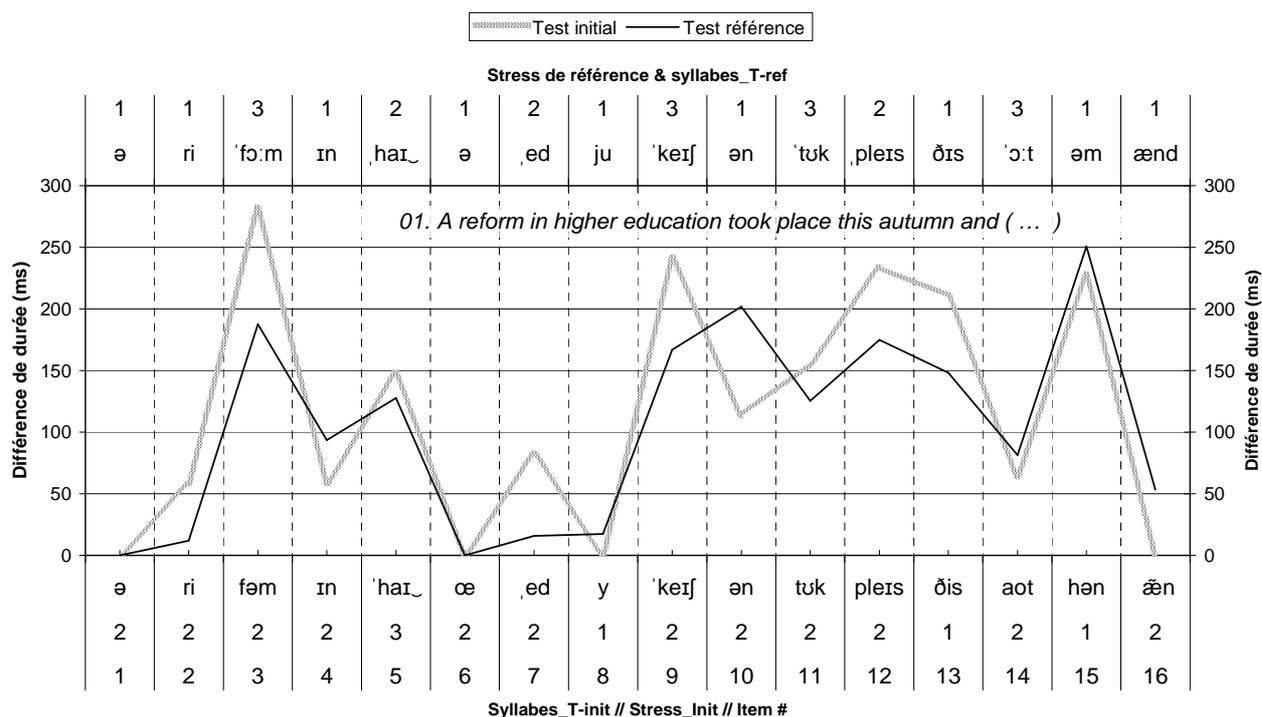


Figure 17. Quantité consonantique calculée pour les items #1 à 16 de Ph01, en  $T_{init}$  et  $T_{ref}$ .

Concernant l'intervalle S1-S8, les deux courbes varient de manière similaire, mais avec un contraste plus important pour  $T_{init}$  par rapport à  $T_{ref}$  : +100 ms pour S3 et -35 ms pour S4. Les courbes se rejoignent ensuite sur S5-S6, puis se séparent sur S7 : +65 ms pour  $T_{init}$ . Nous noterons en S8 l'absence de la « semi-consonne » approximante /j/ en  $T_{init}$  ( $dC8 = 0$ ). Sur la séquence S9-S11, les courbes se croisent par deux fois : il semblerait que notre locuteur produit des nasales [n] brèves : S4 = -35 ms ; S10 = -85 ms. Ce phénomène pourrait être dû à sa pratique de l'espagnol L3. Le phénomène n'est pas décelable en S15, puisque Jean a ajouté le son [h] très marqué en finale de *autumn*, d'où notre choix de transcription, préférée ici à /t<sup>h</sup>/<sup>29</sup>. Pour S11-S13, la variation est semblable, mais la durée des C de  $T_{init}$  est supérieure : +60 ms environ. Puis les courbes se rejoignent sur S14-S16. Examinons maintenant la même séquence S1-S16 dans sa version  $T_{term}$  (figure 18).

<sup>29</sup> La transcription simplifiée des dictionnaires de prononciation (ex. Wells 2008), par convention, ne différencie pas l'aspiration des consonnes comme /p/-/t/-/k/, généralement aspirées en initiales et non-aspirées à l'intérieur d'un mot.

### Différence durées Syllabes-Voyelles\_Phase 01\_Syllabes 1 à 16

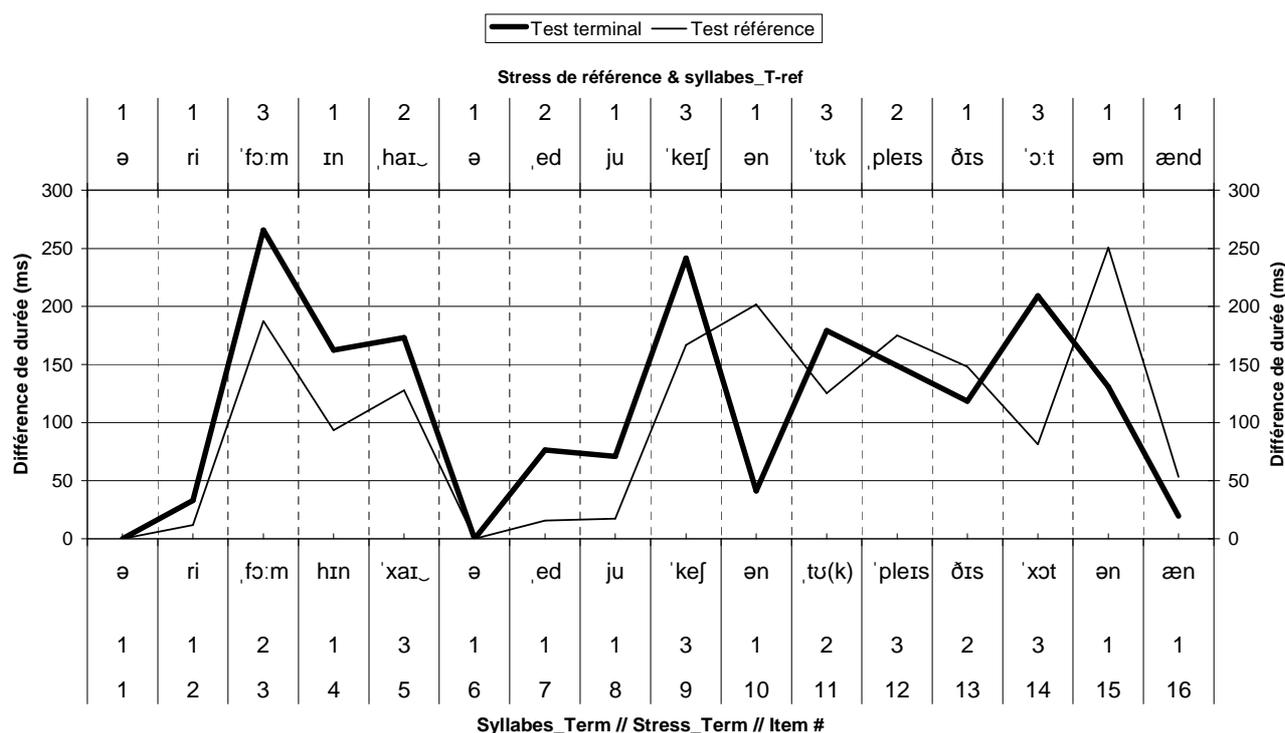


Figure 18. Quantité consonantique calculée pour les items #1 à 16 de Ph01, en  $T_{\text{term}}$  et  $T_{\text{ref}}$ .

Sur toute la séquence S1-S9, les courbes  $T_{\text{term}}$  et  $T_{\text{ref}}$  sont très semblables en variation, les valeurs de durée de Jean étant toutefois constamment supérieures à celles de référence d'environ 50 à 60 ms. Pour la séquence S9-S16, par contre, les courbes se contredisent constamment en variation, avec des croisements. Nous retrouvons la brièveté de la nasale /n/ en S10 de Jean par rapport à  $T_{\text{ref}}$ , déjà évoquée. La variation de sens opposé en S14 s'explique par l'utilisation du son consonantique fort [x]<sup>30</sup> en initiale du mot *autumn*.

#### Conclusion : quantité consonantique de la phrase Ph01

La quantité consonantique semble être un paramètre relativement moins déterminant dans notre travail sur l'amélioration de la dynamique accentuelle de notre sujet. Nous avons vu dans l'extrait illustré par la figure 15 ci-dessus que nous ne pouvions pas vraiment établir de relation directe entre les différents niveaux *stress* de Jean, tels que perçus par l'enseignant, et les variations du faisceau de courbes. L'analyse du début de la

<sup>30</sup> Rappelons ici pour mémoire que l'anglais est catégorisé dans les langues germaniques. D'ailleurs, le son /x/ est présent dans certaines variétés de l'anglais des Îles Britanniques, par ex. dans le mot « lac » : *loch* (Ecosse) ou *lough* (Irlande), prononcé /lɒx/.

phrase menée ensuite ne nous permet pas de dégager de tendances claires, ni de déterminer s'il y a eu progrès, ou pas. L'analyse de  $T_{ref}$  ci-dessus avait pourtant permis de dégager une apparente corrélation.

Un dernier aspect pourra être encore être observé cependant : si l'on fait la somme de l'ensemble de la durée consonantique (dC) de la Ph01 dans les trois versions du test, nous obtenons les valeurs suivantes :

Version du test – Ph01	dC – Ph01	Différence entre dC <sub>init</sub> et les autres tests
$T_{init}$	3,45 s	—
$T_{term}$	3,38 s	-0,07 s — -2 %
$T_{ref}$	2,92 s	-0,53 s — -18 %

Nous avons émis plus haut l'hypothèse qu'une dynamique accentuelle importante raccourcirait la part de la durée consonantique dans la phrase. Les résultats de tableau montrent que le raccourcissement est insignifiant. À titre indicatif,  $T_{ref}$  marque un raccourcissement un peu plus net de -0,53 s — -18 %, mais le locuteur est différent.

Nul doute que le *stress* a une influence sur la manière dont les consonnes sont réalisées, tant en qualité qu'en durée. Toutefois, l'examen attentif de la phrase 01 que nous avons mené sur les productions de Jean montrent que ce critère de durée est plus délicat à utiliser que les précédents pour notre propos, et ce malgré le bilan de notre analyse rapide de la courbe de référence  $T_{ref}$  qui suggère un lien probable entre *stress* et quantité consonantique. Les allongements/raccourcissements de durée, observés chez notre sujet, sont souvent dus à des phénomènes de production phonémique (incursion de la L3 de Jean notamment), étrangers aux phénomènes accentuels. De plus, ils ont demandé une analyse fine et point à point. Du point de vue pédagogique, un travail en cours avec cet étudiant sur ce phénomène identifié pourra être mis en œuvre.

### ***Calcul de la vitesse de parole et de son accélération dans la phrase Ph01***

Dans un premier temps, grâce à nos mesures de durée, nous calculerons la vitesse instantanée pour chaque syllabe. Ensuite, nous procéderons au calcul de l'accélération (ou du ralentissement) de l'élocution lors du passage d'une syllabe à l'autre. Ces paramètres pourront alors être utilement intégrés à notre réflexion sur l'accentuation des syllabes par note sujet.

### Calcul de la vitesse instantanée de parole dans la phrase Ph01

C'est avec les données du test de référence  $T_{ref}$  que nous proposons d'entamer cette partie de notre travail (figure 19). Nous avons obtenu ce graphique par le report du calcul de la vitesse instantanée de la parole de notre locuteur de référence pour chaque syllabe, à partir de sa durée. L'unité de vitesse que nous utilisons est la « syllabe par seconde » :  $S/s$  ou encore  $S \cdot s^{-1}$  où  $S$  est la syllabe et  $s$  l'unité « seconde ». La vitesse de parole est calculée selon la formule suivante :  $vS_i = 1/dS_i \times 1000$ , où  $dS_i$  est la durée de la syllabe  $i$  et  $vS_i$  la vitesse de parole instantanée pour la syllabe considérée. Le quotient  $1/dS_i$  est multiplié par 1000, car l'unité de nos mesures de durée est la milliseconde (ms). Il s'agit de vitesse instantanée ici, et non pas d'une vitesse de parole moyenne de notre locuteur, d'un tempo.

Le graphique de la figure 19 représente la vitesse instantanée pour chacune des 26 syllabes de la phrase Ph01, produites par le locuteur de référence (version  $T_{ref}$ ). Par ce graphique, nous avons cherché à déterminer si une relation pouvait être trouvée entre la variation des niveaux de *stress* et la vitesse de parole calculée à l'instant de chaque syllabe. Nous comparerons ensuite ces données aux productions initiale ( $T_{init}$ ) et terminale ( $T_{term}$ ) de Jean et chercherons à déterminer si a) nous pouvons relever par ce moyen des zones de *stress* dans les versions de Ph01 produites et b) s'il y a eu progrès de Jean, ou pas, entre  $T_{init}$  et  $T_{term}$ .

### Vitesse de parole instantanée\_Syllabes/seconde\_T-ref\_Ph01

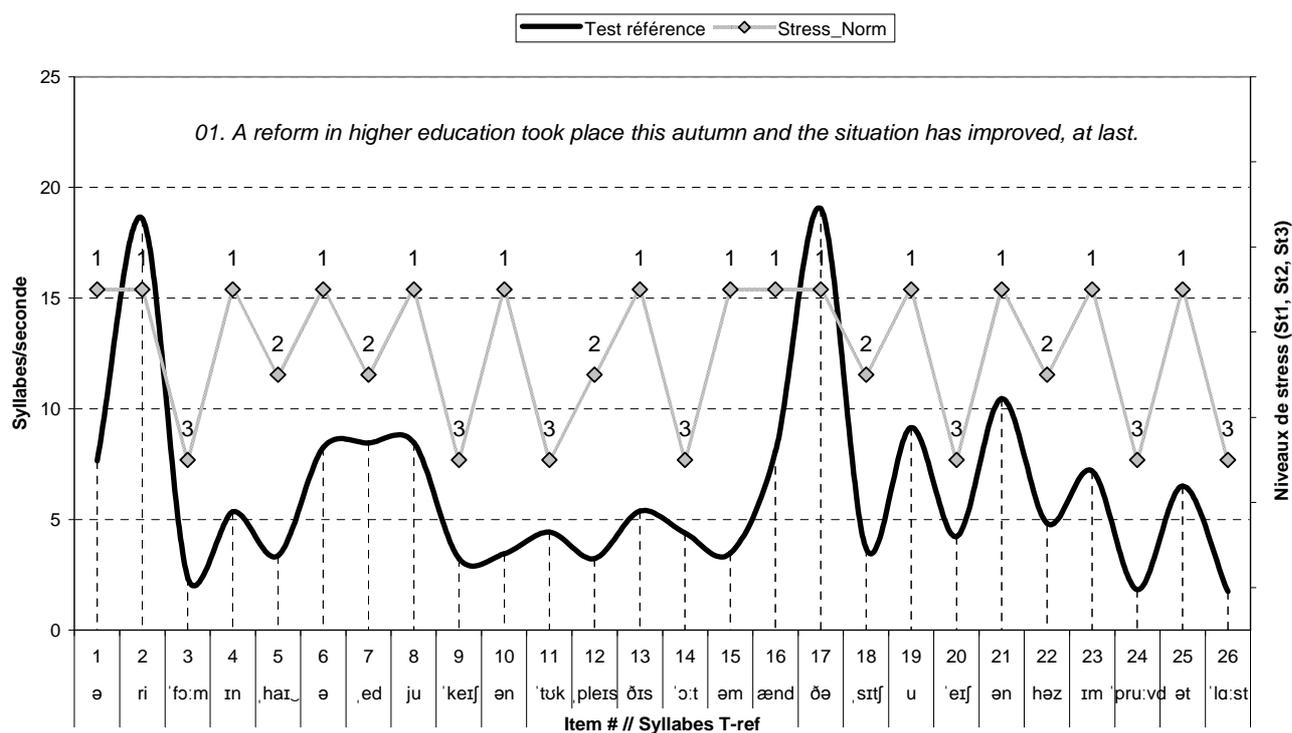


Figure 19. Vitesse de parole instantanée pour les 26 syllabes de Ph01, en  $T_{ref}$ , et niveaux de *stress* de référence.

À la courbe de vitesse instantanée se superposent les niveaux de *stress* de référence, comme nous le faisons pour d'autres parties de cette étude. Pour en faciliter la lecture, nous avons renversé la « courbe » des niveaux de stress : a priori, les syllabes de type St1 « passeront plus vite » que celles sous St3 et ce renversement, si nos présomptions se vérifient, mettra mieux en évidence le phénomène.

L'examen de ce graphique révèle que les tendances de variations se confirment en tous points<sup>31</sup>, à des degrés différents, bien entendu, les valeurs n'étant aucunement comparables. Nous constatons des pics de vitesse, parfois très élevés, signalant le passage des syllabes les plus courtes et les moins accentuées ; à l'inverse, les creux de la courbe de vitesse se trouvent sur les syllabes contenant le plus de consonnes et/ou les voyelles les plus longues. Les valeurs extrêmes vont de moins de  $2 \text{ S}\cdot\text{s}^{-1}$  pour les dernières à pratiquement  $20 \text{ S}\cdot\text{s}^{-1}$  pour les premières, soit un facteur d'évolution de vitesse de 10. Ceci

<sup>31</sup> Hormis S11 : syllabe à voyelle courte sous St3 ; et S15 : ralentissement à l'articulation majeure de la phrase. Nous renvoyons le lecteur à nos explications détaillées de la figure 13. Ceci ne contredit pas nos constatations.

nous conforte dans l'utilisation de  $T_{ref}$  comme courbe de référence. Confrontons maintenant ces observations aux données de notre sujet, Jean.

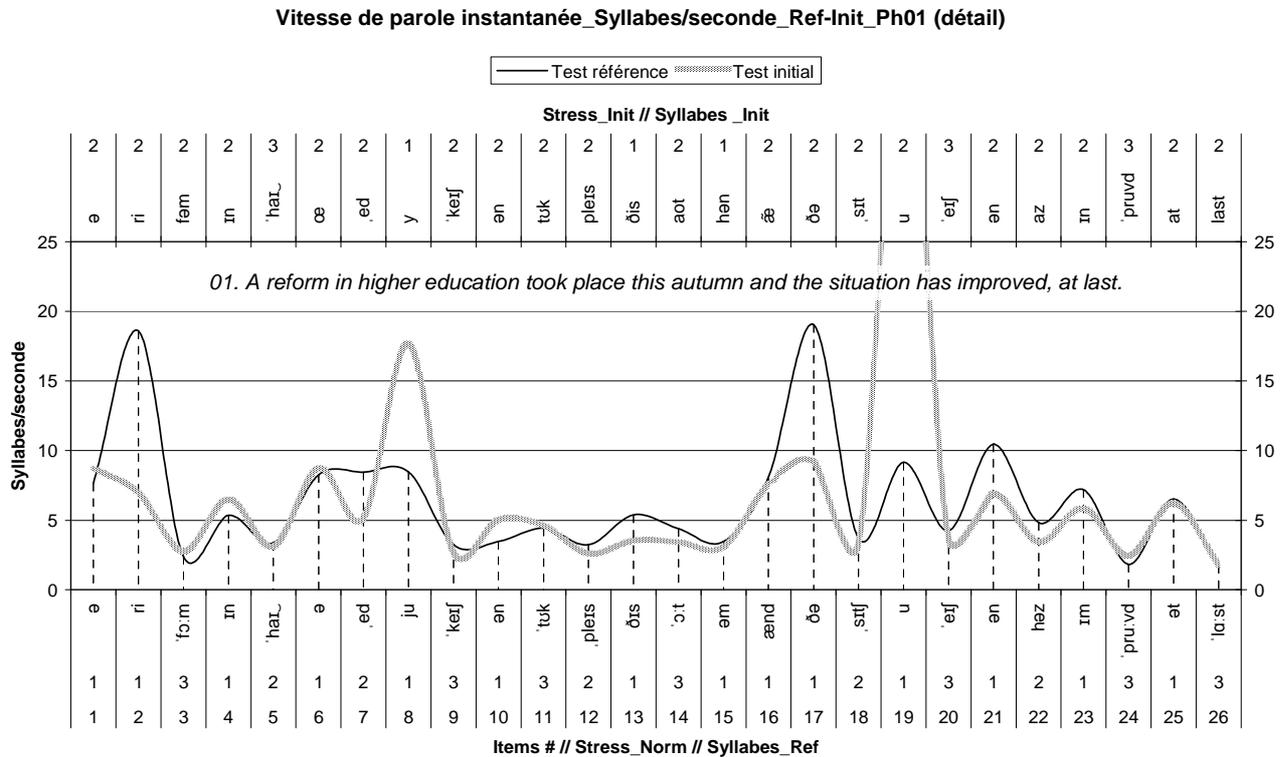


Figure 20. Vitesse de parole instantanée pour les 26 syllabes de Ph01, en  $T_{init}$  et  $T_{ref}$ .

Sur la figure 20 nous avons superposé les données issues du calcul de vitesse de la phrase Ph01 produite par Jean dans la version  $T_{init}$ , avec  $T_{ref}$ . Nous avons conservé la même échelle que pour la figure 19. Une valeur de  $T_{init}$  sort du champ (à  $42 \text{ S}\cdot\text{s}^{-1}$ ) cependant : nous avons vu que certaines voyelles ou syllabes sous St1 pouvaient être très réduites, voire disparaître ; dans ces conditions, la vitesse tend vers l'infini. Dans la pratique, une valeur avoisinant  $20 \text{ S}\cdot\text{s}^{-1}$  nous paraît suffisamment représentative du « maximum de brièveté » de nos syllabes. La lecture de ce graphique montre deux courbes relativement proches par endroits (proximité du modèle) et très divergentes en d'autres. Examinons brièvement ces points ou séquences de grande divergence : S1-3, S6-9, S16-18, S18-20.

- S1-3 : séquence très peu contrastée en  $T_{init}$  ; soit la  $V_{short}$  attendue en S2 n'est pas rendue et/ou le /r/ alvéolaire anglais est encore teinté de la fricative /ʁ/, résiduelle du français, de durée supérieure.
- S6-9 : série  $T_{init}$  comprenant des sons résiduels du français (cf. étiquettes du graphique) ; absence de l'approximante /j/ en S9.

- S16-18 : /ə/ de l'article *the* encore teinté de /œ/ français ; manque probable de dynamisme accentuel sur les syllabes adjacentes
- S18-20 : son V<sub>weak</sub> /u/ hypoarticulé en S19 (cf. explications graphiques à ce sujet ci-dessus) ; plutôt positif, mais la voyelle ne doit pas disparaître ici (tend vers ∞).

Comparons maintenant ce graphique à celui de la figure 21 pour la version T<sub>term</sub>.

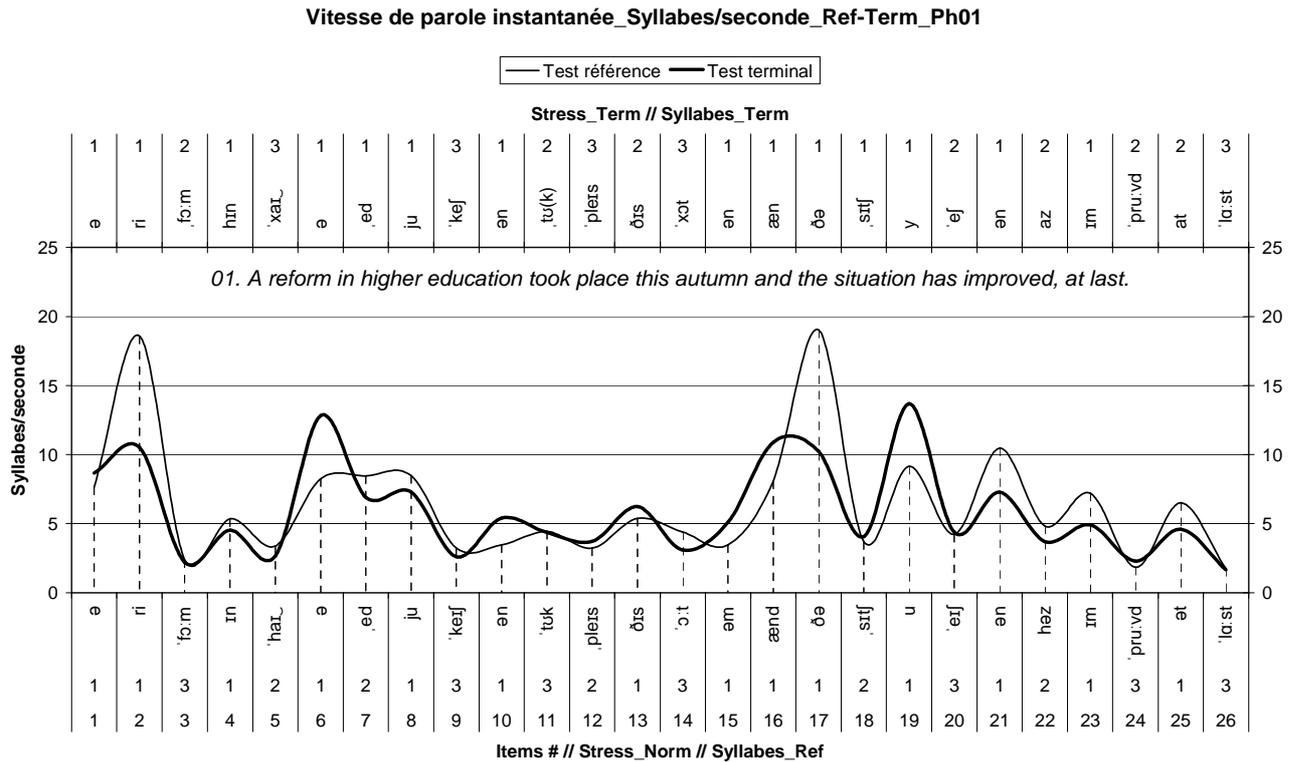


Figure 21. Vitesse de parole instantanée pour les 26 syllabes de Ph01, en T<sub>term</sub> et T<sub>ref</sub>.

Les disparités sont moins grandes ici : la courbe T<sub>term</sub> se rapproche très nettement de la référence en valeur, alternativement au dessus et au dessous de T<sub>ref</sub>, et la copie intégralement en tendance. La progression sur S18-26 est quasiment identique au modèle. Quelques disparités peuvent être signalées, mais elles ne sont pas très significatives :

- /ri/ en S2 peut être encore « raccourci »
- la plus grande brièveté du schwa [ə] en S6 va dans le « bon » sens
- une vitesse plus élevée en S16-17 serait possible

Ainsi, la vitesse de parole instantanée est un paramètre digne d'intérêt, permettant de mettre en lumière une stratégie de mise en œuvre, ou non, par notre sujet, d'un contraste

de durée dans sa production orale de l'anglais L2. Dans le cas la phrase Ph01, le progrès de Jean, dans le sens d'un rapprochement de son interlangue vers le modèle de référence, apparaît clairement entre  $T_{init}$  et  $T_{term}$ .

### Calcul de l'accélération syllabique dans la phrase Ph01

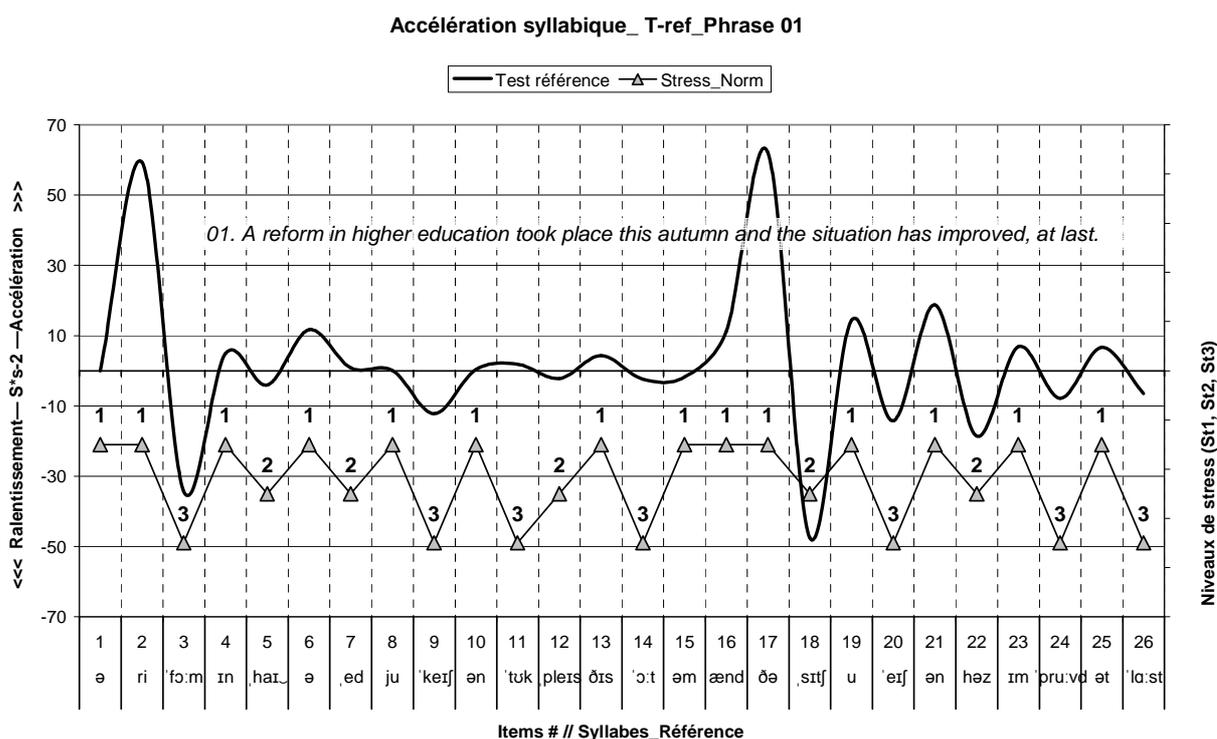


Figure 22. Accélération syllabique pour les 26 syllabes de Ph01, en  $T_{ref}$ .

Le graphique de la figure 22 affiche les valeurs de « l'accélération syllabique » calculées pour la phrase Ph01, dans sa version de référence  $T_{ref}$ . L'accélération syllabique est calculée comme une fonction dérivée de la vitesse. L'intérêt de ce calcul, et de la représentation des résultats par un graphique, réside dans le fait que nous pouvons représenter façon très claire les variations de la vitesse de parole. Comme cela a déjà été fait dans d'autres parties de ce travail, nous juxtaposons dans ce premier graphique la courbe  $T_{ref}$  et une représentation des variations du *stress* de référence. Pour calculer l'accélération syllabique, nous avons utilisé la dérivée de la vitesse précédemment calculée :

$$aS_{i-1}S_i = \frac{(vS_i - vS_{i-1})}{(dS_i + dS_{i-1})} \times 1000$$

- Où  $dS_i$  est la durée de la syllabe  $i$ ,  $vS_i$  la vitesse de parole instantanée de la syllabe  $i$ , et  $a$  l'accélération syllabique entre la syllabe  $i-1$  et la syllabe  $i$ .
- Le quotient est multiplié par 1000 car l'unité de nos mesures de durée est la milliseconde (ms).
- L'unité d'accélération que nous utilisons ici est « le nombre de syllabes par seconde, par seconde » ou « la vitesse par seconde » :  $vS/s$  ou  $vS \cdot s^{-1}$  ou  $S \cdot s^{-2}$ .
- Le dividende ( $vS_i - vS_{i-1}$ ) représente la variation de la vitesse entre la syllabe  $S_i$  et la syllabe  $S_{i-1}$ .
- Le diviseur ( $dS_i + dS_{i-1}$ ) représente l'espace de temps pendant lequel la vitesse a varié.
- Le résultat  $aS_{i-1}S_i$  représente l'accélération de la vitesse de parole d'une syllabe à l'autre.

La figure 22 illustre la relation qui existe entre l'accélération ou le ralentissement, souvent considérables, qui peuvent intervenir dans la vitesse de parole en anglais. Ces variations de vitesse, positives ou négatives apparaissent en deçà ou au-delà de l'axe des abscisses placé cette fois en position centrale du graphique : une accélération apparaîtra en haut, un ralentissement en bas. Pour une plus grande lisibilité, les niveaux de *stress* ont été renversés (faibles en haut, forts en bas), une décélération étant généralement attendue en présence d'un niveau de *stress* élevé.

L'examen de ce graphique fait apparaître une relation très étroite entre les niveaux de *stress* et les variations d'accélération d'une syllabe à l'autre, dans la chaîne de parole. Les deux zones contenant une succession de St1 (items #1-2 et #15-17) présentent, de très loin, les pics de variation maxima. Ceci est logique, puisque les syllabes sous St1 concentrent les durées les plus brèves. En revanche, dans les syllabes sous St2-St3, la courbe fait systématiquement un creux ; ce creux est vraiment très prononcé quand le St2-St3 suit une des séries de St1 évoquées ci-dessus. Les St2 sont eux-mêmes facilement repérables dans leur alternance avec les St1 ou les St3. Le seul endroit de la courbe contredisant le St3 se trouve, encore une fois, à l'item #11 : nous sommes en présence du cas particulier de l'initiale d'un syntagme verbal présentant une  $V_{short}$  sous St3, comme déjà abordé précédemment. Pour résumer notre lecture du graphique : si la courbe s'éloigne de l'axe des abscisses, si son amplitude augmente :

- l'accélération (ou la décélération) augmente
- le *stress* diminue
- les syllabes sont raccourcies

À l'inverse, si la courbe se resserre autour de l'axe :

- l'accélération (ou la décélération) diminue
- le *stress* augmente
- les syllabes sont plus longues

Comparons maintenant la courbe d'accélération  $T_{ref}$  avec celles en  $T_{init}$  et  $T_{term}$  produites par Jean. Nous retrouverons certaines similitudes avec les courbes de vitesse de parole de Jean étudiées précédemment ; néanmoins, la présentation de cette courbe dérivée de la vitesse nous permet l'appréhension immédiate et très utile du degré de la gestion de certains aspects du rythme par notre sujet. La courbe  $T_{init}$ , encore une fois, sort du champ de notre graphique ; nous avons volontairement privilégié la partie centrale. Pour information, voici les valeurs extrêmes de l'accélération de  $T_{init}$  n'apparaissant pas : item #19 :  $+120 S \cdot s^{-2}$  ; #20 :  $-130 S \cdot s^{-2}$ .

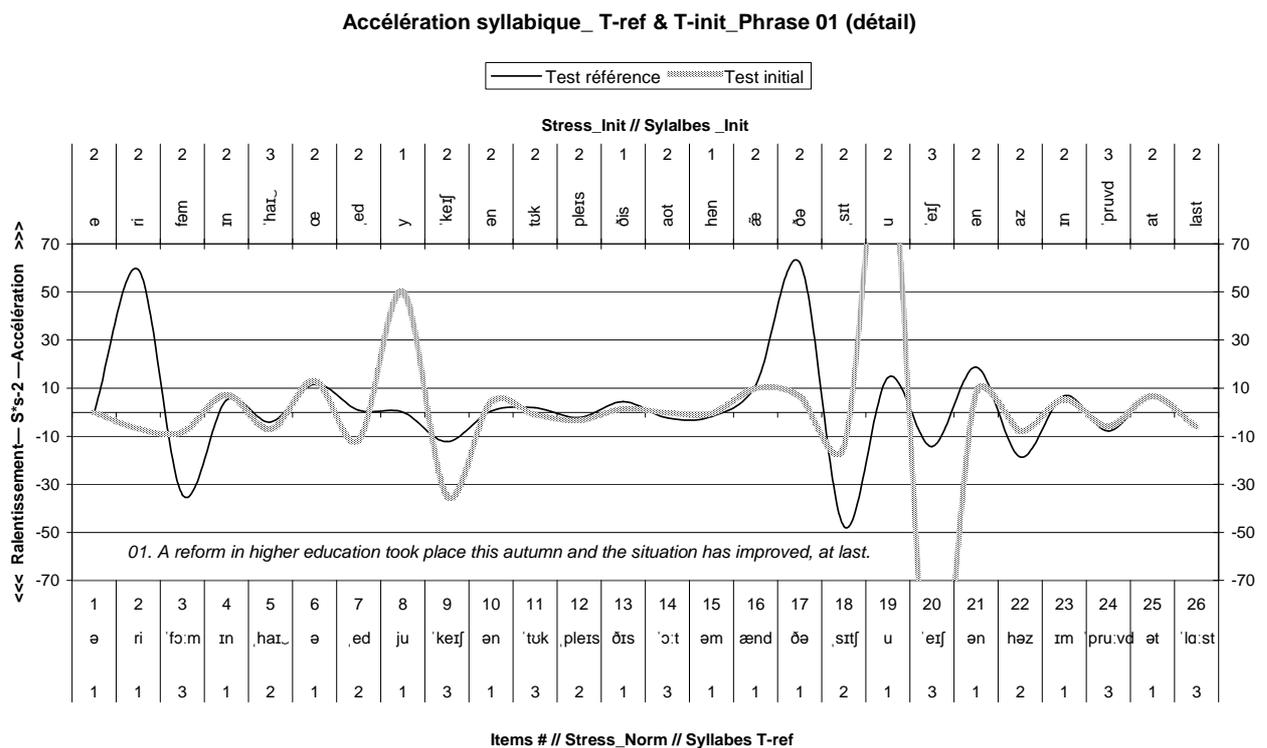


Figure 23. Accélération syllabique pour les 26 syllabes de Ph01, en  $T_{init}$  et  $T_{ref}$ .

Comme nous l'avons remarqué pour la vitesse instantanée, le courbe  $T_{init}$  est paradoxale : elle présente à la fois des disparités, parfois très importantes, avec  $T_{ref}$ , mais également des similitudes, parfois très proches de l'identique. Dans la séquence initiale des

items #1 à 3 les courbes se contredisent et dans les items #7 à 10 et #16-21 les courbes varient de façon semblable, mais à des valeurs extrêmement différentes. Ceci est, selon toute vraisemblance, l'illustration de la production d'un étudiant en phase d'acquisition d'une langue étrangère, avec parfois des passages plutôt réussis et d'autres qui sont en phase d'acquisition. Regardons maintenant si le travail effectué avec et par Jean en cours a eu une influence sur la qualité de sa production orale. La figure 24 montre les mêmes courbes, mais dans la réalisation terminale cette fois-ci.

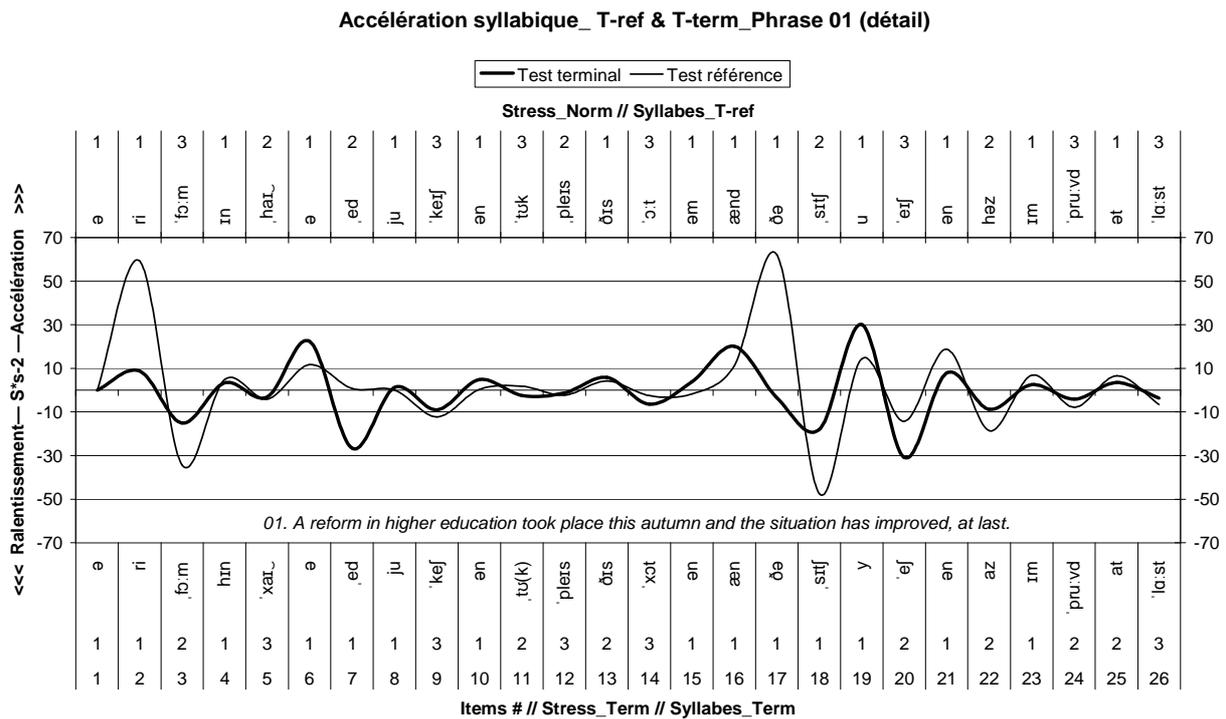


Figure 24. Accélération syllabique pour les 26 syllabes de Ph01, en  $T_{term}$  et  $T_{ref}$ .

Nous pouvons voir que Jean a apporté d'importantes corrections ici. La courbe  $T_{term}$  épouse la variation de  $T_{ref}$ , même si l'amplitude des variations de l'accélération est encore très en deçà de la réalisation de référence. Ceci, d'après notre expérience, est typique d'un francophone natif. Notre étudiant est en progrès, place ses accélérations au bon endroit, mais ne dynamise et ne contraste pas encore assez sa production.

### Conclusion : vitesse de parole et accélération syllabique

Dans les syllabes de niveau St3 (et, dans une moindre mesure, de niveau St2) dans lesquelles les voyelles  $V_{long}$  et  $V_{diph}$  sont généralement majoritaires, une syllabe accentuée correspond à un net ralentissement de la vitesse. À l'inverse, pour les syllabes de niveau St1, la vitesse de parole subit une forte accélération pour atteindre une vitesse beaucoup

plus élevée. Enfin, dans le cas avec  $V_{\text{short}}$  sous St3 (item #11), la question était plus délicate et il était intéressant de s'y attarder : ici, la consonne suit le mouvement de la voyelle dans une production brève : aucune variation de vitesse notable par rapport à la syllabe brève précédente.

Les méthodes employées et les analyses réalisées dans cette partie fournissent des résultats plus intéressants. En matière de rythme, les paramètres de vitesse et d'accélération sont fondamentaux, a fortiori pour une langue aussi dépendante de la dynamique accentuelle que l'anglais. Ceci est confirmé par les résultats que nous obtenons ici. En appliquant ces méthodes à l'étude du cas de Jean, nous avons pu mettre en valeur les endroits où il utilisait le *stress*, ou pas, et si ces endroits étaient ceux qui se rapprochaient de notre modèle de référence. Tout cela a été effectué jusqu'à présent au travers d'un seul facteur fondamental de l'expression orale en langue anglaise, et de ses dérivés: la durée. Nous allons maintenant nous intéresser aux deux autres composants fondamentaux du *stress* : la hauteur mélodique et l'intensité sonore.

### ***Hauteur mélodique ( $F_0$ ) des voyelles de la phrase Ph01***

#### **Introduction**

Comme nous l'avons vu, certains auteurs placent la hauteur mélodique au sommet des facteurs dont le *stress* dépend ; d'autres la considèrent comme essentielle, mais ni plus ni moins importante que la durée ou l'intensité. Ainsi que nous l'avons fait plus haut pour nos différentes approches de la durée, nous proposons d'analyser les données que nous avons recueillies au cours de nos enregistrements, de notre locuteur de référence pour commencer ( $T_{\text{ref}}$ ), puis de Jean, notre sujet, dans les deux versions  $T_{\text{init}}$  et  $T_{\text{ref}}$  de la phrase Ph01. Les valeurs de  $F_0$  que nous présentons ici sont toutes des moyennes des mesures de la fréquence fondamentale vocalique, acquises à l'aide du logiciel PRAAT.

#### **Analyse de la hauteur mélodique ( $F_0$ ) de la phrase Ph01**

Le graphique de la figure 25 représente l'évolution des valeurs moyennes de  $F_0$  pour les 26 voyelles de la phrase Ph01, dans sa version de référence  $T_{\text{ref}}$ . La variation des niveaux de *stress* de référence a été reportée sur le graphique au-dessous de la courbe de  $F_0$ , uniquement à fin de comparaison des tendances de variation, les valeurs n'étant pas de même nature.

L'observation de la seule courbe de  $F_0$  nous livre quelques données intéressantes. La variation de la courbe en amplitude semble relativement importante, avec une valeur maximum  $F_0^{\max}$  de presque 150 Hz en V3. À l'opposé, la valeur minimum  $F_0^{\min}$  descend à 76 Hz en V17, soit une amplitude  $F_0^{\max} - F_0^{\min} = 74$  Hz, ou environ 6 tons. Par ailleurs, au-delà de ces points extrêmes, la courbe présente presque partout de grandes variations d'amplitude toutes les 3 ou 4 voyelles en moyenne. Enfin, nous remarquerons que la deuxième moitié de la phrase utilise une étendue de fréquences globalement plus basses que la première moitié : les valeurs des pics de la première moitié ne sont jamais atteintes. Pour expliquer ce phénomène très marqué, nous pouvons émettre comme hypothèse que la structure même de la phrase a une influence sur la façon dont le locuteur gère l'information au moyen de sa  $F_0$ , en mettant plus d'emphasis sur le « thème » introductif de la phrase (amplitude plus grande, étendue de fréquences plus élevée) que sur le « rhème », commentaire complémentaire au thème<sup>32</sup>. Comparons maintenant les variations de cette courbe et l'évolution des niveaux de *stress* de référence tout au long de cette phrase. Les échelles ont été ajustées de manière à faciliter la comparaison des variations.

---

<sup>32</sup> Rappelons ici les termes très explicites de *focus* et de *comment* habituellement employés en anglais pour décrire ces deux parties d'un énoncé.

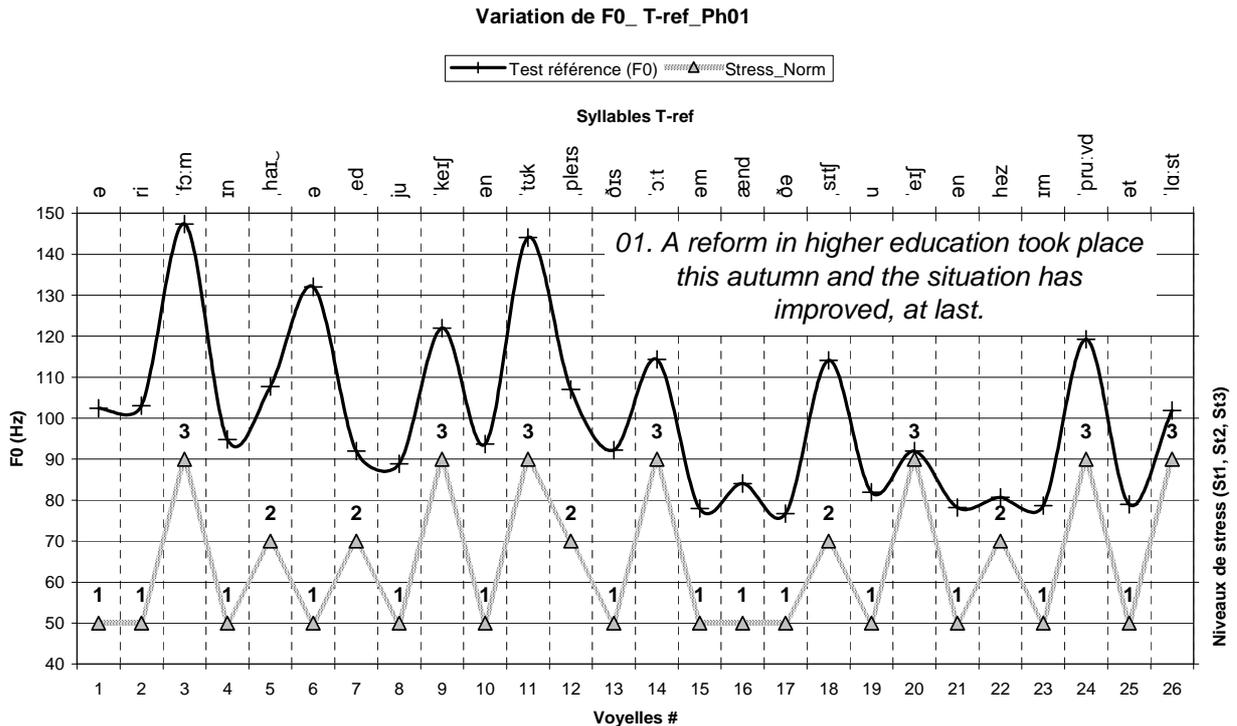


Figure 25. Evolution de F<sub>0</sub> pour les 26 voyelles de Ph01 en T<sub>ref</sub> et niveaux de *stress* de référence.

La correspondance que nous observons ici entre les variations du *stress* et celles de F<sub>0</sub> est quasiment parfaite. Hormis une contradiction de mouvement en V6 et une variation légèrement inversée en amplitude en V18 et V20 (*stress* 2 puis 3 de référence au lieu d'un *stress* plutôt 3 puis 2 observé) que nous allons commenter, les deux variations évoluent en parallèle. Pour ce qui concerne cette phrase, nous pouvons affirmer que F<sub>0</sub> et les variations du *stress* sont, à l'évidence, corrélées : les pics de *stress* de référence, de type St2 et St3, et la production du locuteur de référence correspondent en tous points en occurrence, et 9 fois sur 12 en hauteur/intensité ; la voyelle V6 mise à part, l'intégralité des St1 correspond à des creux de la courbe.

Concernant le premier des deux points déviants évoqués (contradiction *stress*/ F<sub>0</sub> en V6), ce phénomène pourrait être dû aux contraintes phonotactiques du passage très bref du schwa [ə] en V6, pris entre deux sons vocaliques de forte intensité, en l'absence totale de consonnes. La séquence V5-6-7 [ˌhaɪ\_əˌed] ne laisse pas le temps à la fréquence fondamentale de chuter entre les deux syllabes accentuées, ainsi que l'on s'y attendrait.

Pour le deuxième, l'explication que nous avançons est double. D'une part, le son [ɪ] en V18 peut pousser le locuteur vers la production d'une fréquence relativement élevée pour sa réalisation sous *stress* St2 ou St3, par rapport à la diphtongue [eɪ] en V20. D'autre

part, un niveau de *stress* de référence n'est qu'une tendance forte, pas la garantie absolue d'une réalisation effective.

Considérons maintenant les productions de Jean et comparons-les à notre modèle. La figure 26 superpose les courbes de Ph01 en version  $T_{ref}$  et  $T_{init}$ .

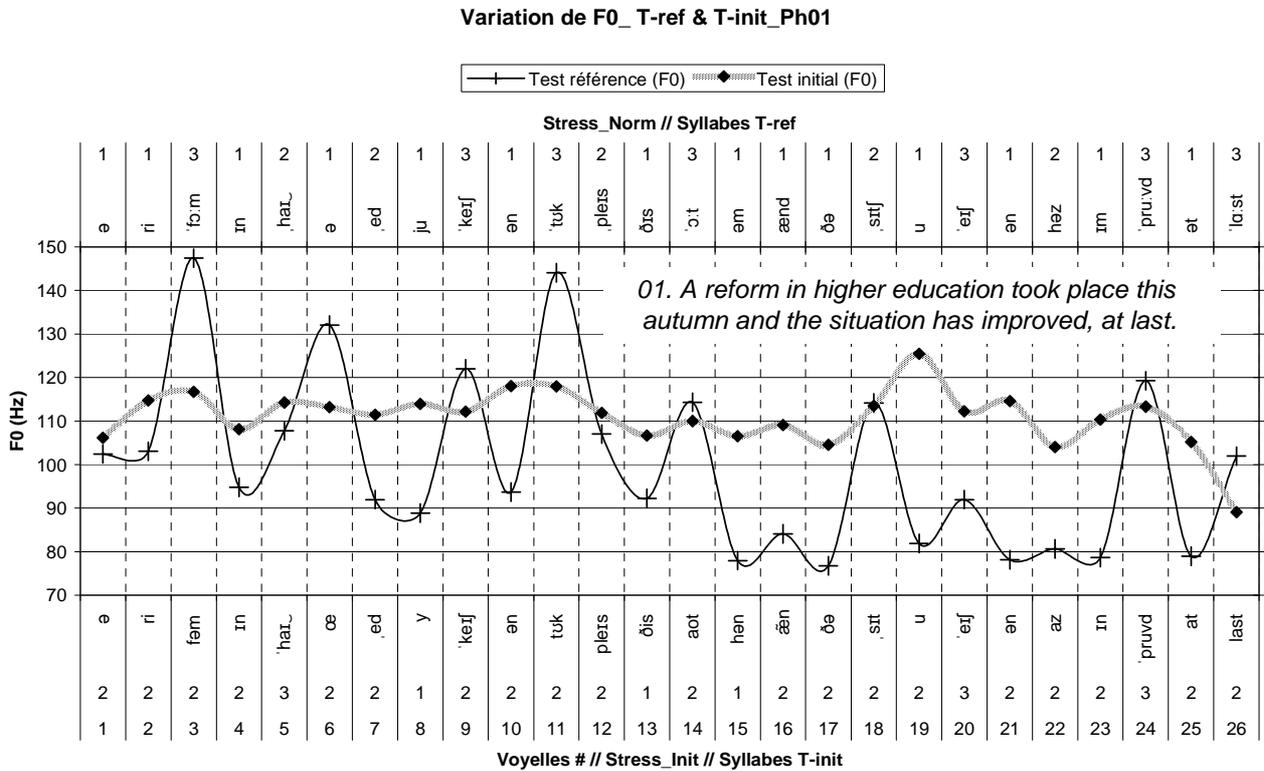


Figure 26. Evolution de  $F_0$  pour les 26 voyelles de Ph01 en  $T_{init}$  et  $T_{ref}$ .

La comparaison est édifiante et le contraste, surprenant : les courbes sont tellement dissemblables —si nous excluons l'anecdotique rencontre des points en V18— qu'il semblerait qu'elles sont issues de deux études sans aucun rapport, si ce n'est une fréquence moyenne à peu près similaire. En fait, cette superposition des courbes illustre la difficulté majeure à laquelle les francophones sont confrontés dans la communication orale en langue anglaise : l'absence de *stress*, de dynamique, rendue ici à travers l'une de ses composantes majeures : la fréquence fondamentale. La valeur  $F_0^{max}$  de Jean ici est de 125 Hz, en V19 ; sa valeur minimum  $F_0^{min}$  se trouve à 80 Hz en V26, soit une amplitude  $F_0^{max} - F_0^{min} = 45$  Hz ou environ  $3\frac{1}{2}$  tons.

Cette courbe quasiment horizontale, sans relief, correspond parfaitement aux valeurs de *stress* qui avaient été perçues pour  $T_{init}$ , à savoir la domination d'un niveau de *stress* moyen, autour d'un St2 omniprésent, proche de l'expression orale en langue

française : un « iso-syllabisme ». Nous remarquerons également le détail des courbes contradictoires en initiale et en finale, avec  $T_{init}$  proche de ce qu'on attendrait dans la lecture d'une phrase française. Regardons maintenant la version finale de la phrase Ph01 et comparons nos données à l'aide de la figure 27.

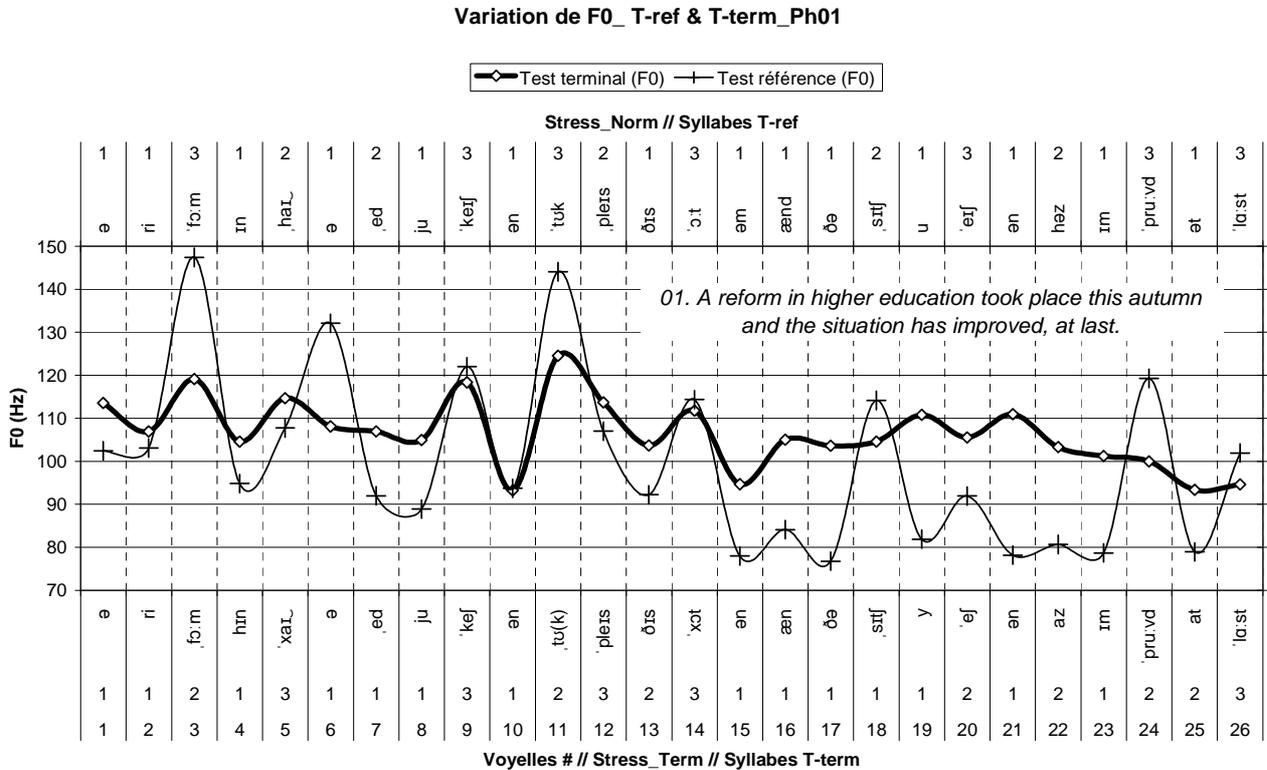


Figure 27. Evolution de F0 pour les 26 voyelles de Ph01 en  $T_{term}$  et  $T_{ref}$ .

La figure 27 révèle une évolution qui, sans tendre vers la « perfection », est indéniable : la première moitié de la courbe  $T_{term}$  est en accord avec la grande majorité des variations de  $T_{ref}$ , tout en restant bien en deçà de ses valeurs ; mais le dynamisme, même s'il est encore embryonnaire, est bien présent. Après V17, les courbes se séparent en tendance, les variations de  $T_{term}$  se resserrent autour d'une amplitude faible et contradictoire de  $T_{ref}$ , sans toutefois épouser les variations de  $T_{init}$ . Un point positif de cette deuxième moitié de phrase : la fin duplique timidement la variation de  $T_{ref}$ , contrairement à ce que faisait  $T_{init}$ . Nous noterons qu'il en va d'ailleurs de même pour le début de la phrase. La valeur  $F_0^{max}$  de Jean ici est de 124 Hz, en V11 ; sa valeur minimum  $F_0^{min}$  se trouve à 93 Hz en V10, soit une amplitude  $F_0^{max} - F_0^{min} = 31$  Hz ou environ  $2 \frac{1}{2}$  tons. Certes, ces valeurs sont a priori moins « bonnes » qu'en  $T_{init}$ , mais elles tombent sur des endroits beaucoup plus adéquats.  $F_0^{min}$  en  $T_{init}$  tombe sur l'ultime 26<sup>ème</sup> voyelle de la phrase, sur une

intonation typiquement francophone. Si nous excluons cette dernière valeur, les 25 autres oscillent (à peine) entre 104 Hz et 125Hz, soit 1½ ton.

Jean a-t-il progressé ? Pour répondre avec la plus grande précision possible à cette question, nous allons procéder à une manipulation de notre dernier graphique, celui de la figure 27. La figure 27<sup>bis</sup> reprend le même graphique que précédemment, mais l'axe secondaire des ordonnées utilisé pour  $T_{ref}$ , situé à droite, en caractères soulignés affiche une échelle compressée d'environ 130% par rapport à celle de  $T_{init}$ . Cette manipulation a pour objectif de ramener visuellement les variations de  $T_{ref}$  à approximativement la même amplitude que celles de  $T_{term}$ .

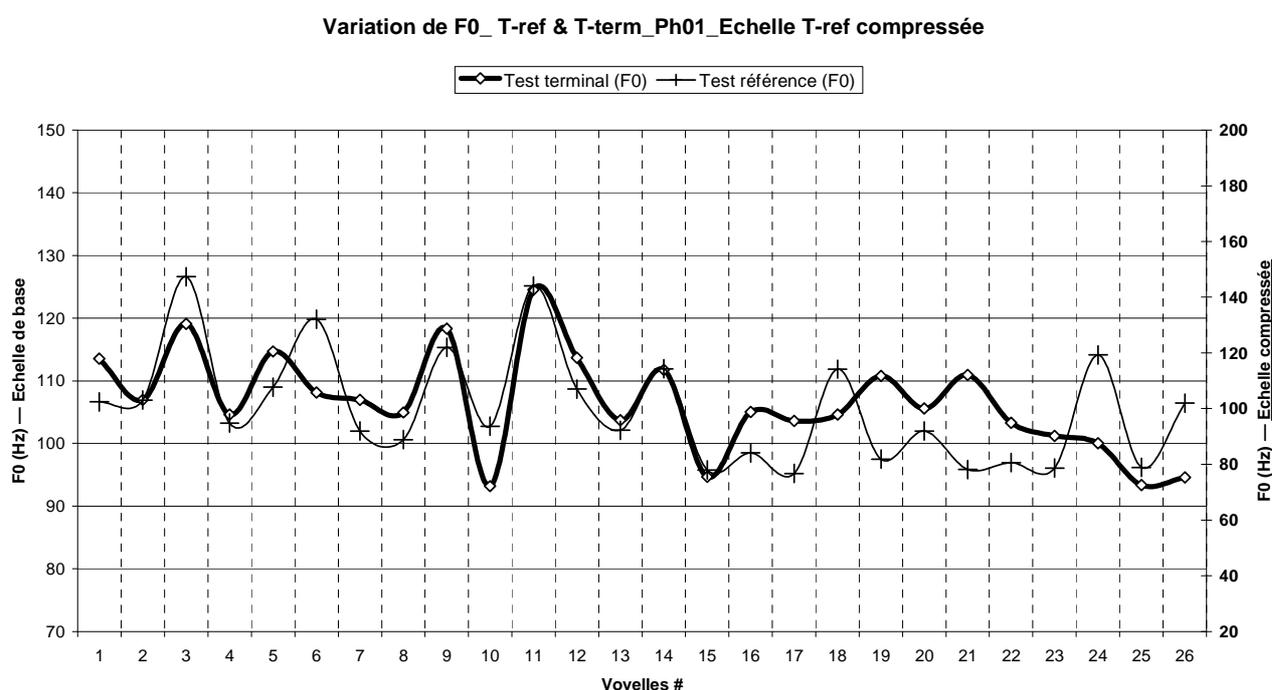


Figure 27<sup>bis</sup>. Evolution de  $F_0$  pour les 26 voyelles de Ph01 en  $T_{term}$  et  $T_{ref}$ . **Echelles dissemblables.**

Ce que nous observons peut se résumer principalement à deux choses :

- par dessus tout : la mise en valeur d'une très grande proximité des deux courbes sur la première moitié de Ph01, certaines parties se reproduisant à l'identique
- même césure de la phrase en deux parties distinctes, autour de la voyelle V16, à l'articulation majeure *and* : la phrase de Jean s'éloigne assez fortement et de manière plutôt contradictoire de la courbe  $T_{ref}$  ; toutefois nous noterons un point commun, positif : un (léger) affaissement en fréquence de la courbe sur cette partie, comme sur  $T_{ref}$ .

Nous voyons qu'avec une modulation mélodique plus importante, allant vers une amplitude plus grande de sa  $F_0$ , tout en se basant sur les mêmes points de repère, les mêmes lieux de creux et pics de la courbe, Jean a une bonne partie des atouts pour se rapprocher encore plus du modèle.

### **Conclusion sur la hauteur mélodique**

Nous pouvons dire sans conteste qu'il y a bien eu progrès de Jean ici. La courbe  $T_{\text{term}}$  se rapproche de  $T_{\text{ref}}$  de façon flagrante en certains endroits, ce qui la différencie fortement de  $T_{\text{init}}$ , qui n'avait quasiment aucun rapport avec notre référence. Jean est en progrès dans son interlangue. L'acquisition de ce paramètre du *stress* est en bonne voie, même s'il reste une marge importante de progrès, et même si ce progrès n'est pas uniforme sur l'ensemble de la phrase.

## ***Intensité sonore des voyelles de la phrase Ph01***

### **Introduction**

Pour conclure cette partie, nous abordons le dernier, et probablement le plus délicat à interpréter de nos paramètres caractérisant le *stress* : l'intensité sonore. Ainsi que nous l'avons vu au début de ce travail, l'intensité s'entend à partir de la production d'un son. La mise en relief d'une syllabe dans la chaîne de parole par le *stress* peut se faire à l'aide des différents paramètres déjà mentionnés (durée,  $F_0$  et, maintenant, intensité), indépendamment les uns des autres ou par leur combinaison à des degrés divers, et la perception de cette mise en relief est généralement teintée de beaucoup de subjectivité. Les valeurs que nous présentons ici sont les moyennes des valeurs de la partie vocalique de nos syllabes mesurées à l'aide du logiciel PRAAT. Notons que pour les valeurs en ordonnées nous avons choisi 3 dB comme unité de base : chaque incrément de 3dB correspond à un doublement de l'intensité sonore, ce qui nous donne une idée de l'échelle de variation.

### **Analyse de l'intensité moyenne des voyelles de la phrase Ph01**

La figure 28 superpose la courbe de l'intensité sonore de la phrase Ph01 en version de référence avec les niveaux de *stress* de référence, comme nous l'avons fait pour les autres paramètres, afin de faire apparaître d'éventuelles relations entre elles, non pas de valeurs mais de tendances.

Contrairement à la fréquence fondamentale, où la relation était très lisible, il faut prendre un peu plus de temps pour se rendre compte de correspondances. S'il n'y pas de variations directement liées, nous trouvons néanmoins des points de convergence assez réguliers et clairement identifiables : les pics de St3. Hormis la voyelle V11 —point particulier que nous avons déjà eu l'occasion d'évoquer :  $V_{short}$  sous St3 du syntagme verbal, tous les autres pics de stress (St3) correspondent nettement aux maxima d'intensité de la courbe.

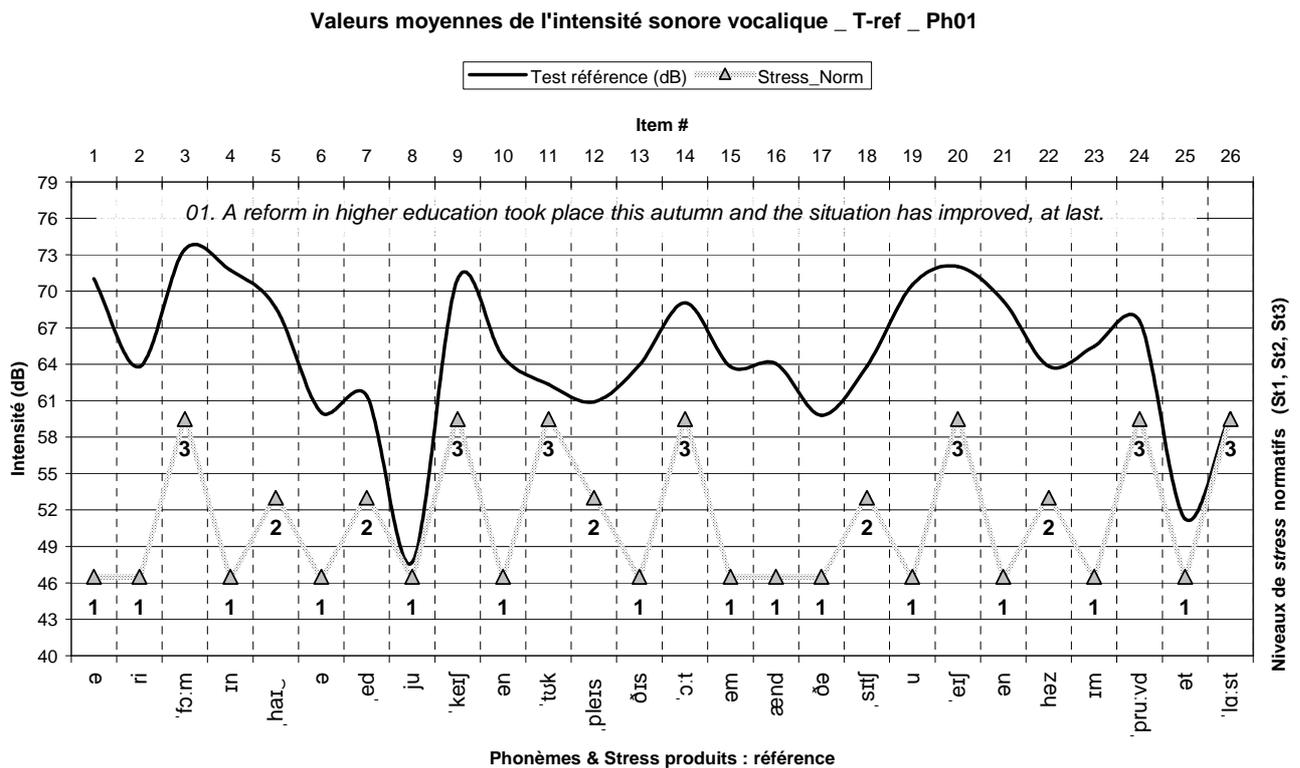


Figure 28. Evolution de l'intensité pour les 26 voyelles de Ph01 en  $T_{ref}$ , avec niveaux de stress de référence.

La tendance globale de la courbe d'intensité est de ne pas fluctuer à chaque variation du niveau de stress, tel que le fait celle de  $F_0$ . L'arrondi de la courbe sur la séquence S17-22 ne reflète pas du tout le zigzag de la fluctuation du *stress*, par exemple. Néanmoins, pour les St1, des points de convergence apparaissent très nettement en V8 et V25, d'autres moins marqués en V2, 6, 15-17. Certains St2, quant à eux, peuvent être identifiés assez clairement en V5, 7 et 18.

Nous noterons enfin la lecture très nette de la réduction vocalique des syllabes V8 et V25, où l'intensité chute jusqu'aux minima de la courbe, respectivement 47 et 51 dB. A l'opposé, l'intensité maximum dépasse 73 dB.

Ainsi, si des relations existent bien entre *stress* et intensité vocalique dans cette phrase, même si elles sont moins flagrantes que dans le cas de  $F_0$ . Comparons maintenant les courbes de production de Jean dans leurs phases initiales et terminales. Le graphique de la figure 29 superpose les courbes  $T_{ref}$  et  $T_{init}$ .

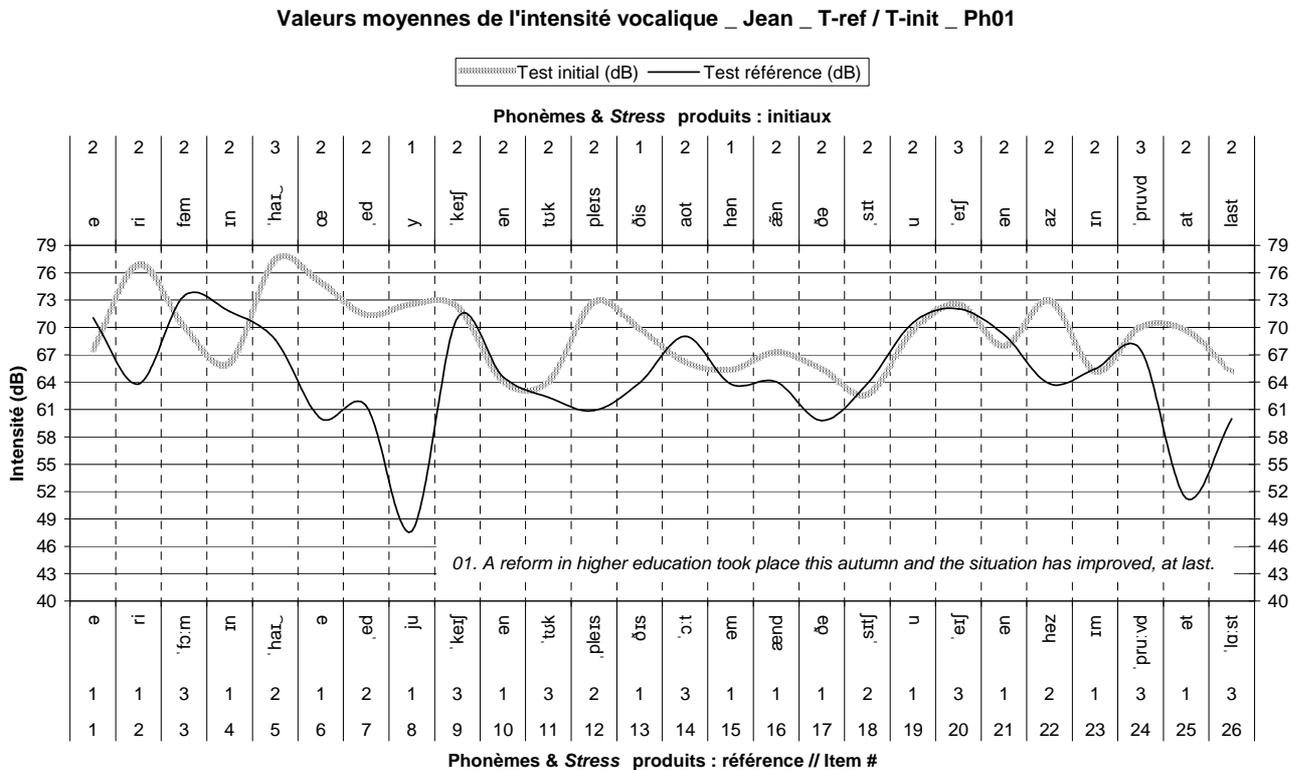


Figure 29. Evolution de l'intensité pour les 26 voyelles de Ph01 en  $T_{init}$   $T_{ref}$ .

Les courbes obtenues sont très dissemblables et contradictoires : beaucoup moins d'amplitude en  $T_{init}$ . Quelques rares zones présentent des similitudes grossières de forme et/ou de mouvement :

- de V5 à V9 : variation grossièrement similaire en forme, mais pas du tout en hauteur (beaucoup plus haute en  $T_{init}$ ), ni amplitude (très inférieure en  $T_{init}$ )
- V9-10 et V18-21 : très brefs passages à l'identique
- V15 à 17 : variation similaire en forme, mais pas en amplitude :  $T_{init}$  dépasse sensiblement  $T_{ref}$ , cette fois

Noter la contradiction, du début et de la fin de la phrase des deux courbes, qui est due vraisemblablement à l'influence de la langue maternelle française chez Jean. Les cas de contradiction totale des courbes sont très nombreux : V2, 4, 12, 22 et 25.

Insérons maintenant la version  $T_{term}$  dans notre graphique, visible en figure 30.

Valeurs moyennes de l'intensité vocalique \_ Jean \_ T-ref / T-term\_Ph01

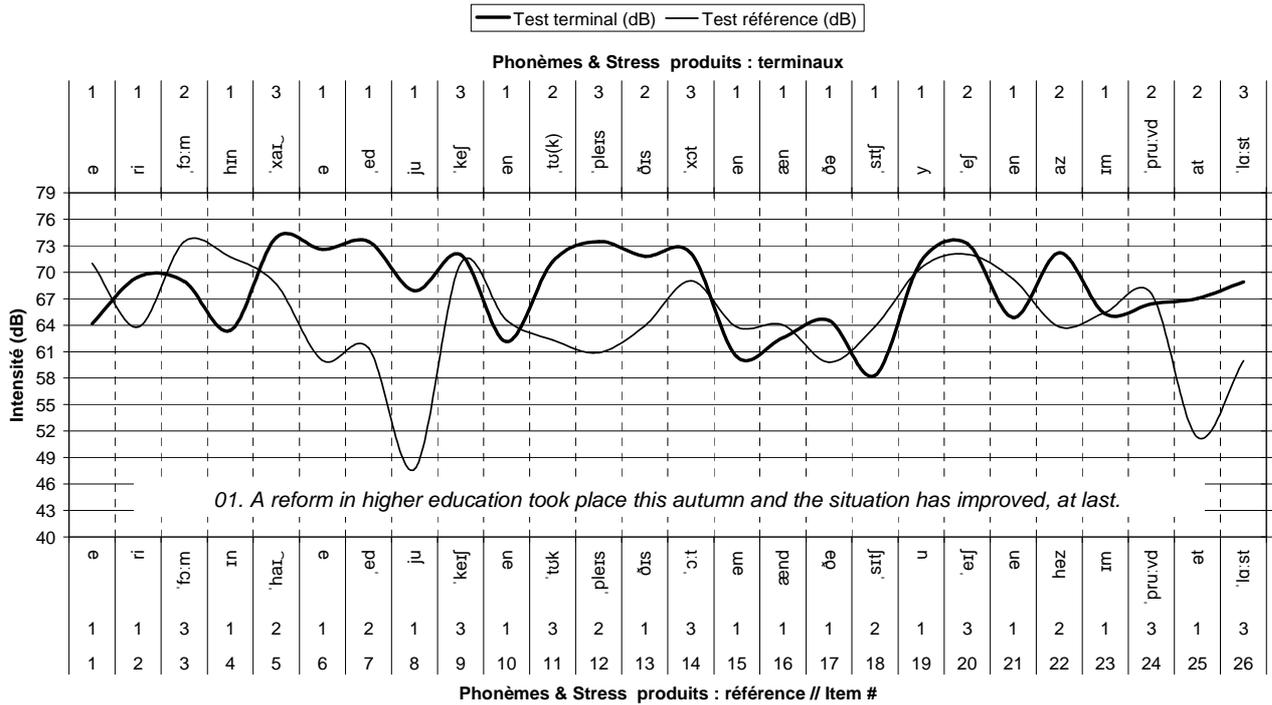


Figure 30. Evolution de l'intensité pour les 26 voyelles de Ph01 en  $T_{term}$  et  $T_{ref}$ .

Comme pour  $T_{init}$ , les courbes obtenues sont très dissemblables et contradictoires.

Quelques rares zones présentent des similitudes de forme et/ou de mouvement :

- de V5 à V9 : variation similaire en forme, mais pas du tout en hauteur (beaucoup plus haute en  $T_{init}$ ), ni amplitude (très inférieure en  $T_{init}$ .)
- V9-10 : très bref passage à l'identique
- V14 à 16 : variation similaire en forme, mais pas en amplitude :  $T_{init}$  dépasse  $T_{ref}$ .

Un dernier graphique de superposition (figure 30), de type similaire mais avec  $T_{init}$  et  $T_{term}$  permettra de mieux mettre en avant points communs et différences.

Valeurs moyennes de l'intensité vocalique \_ Jean \_ T-init / T-term\_Ph01

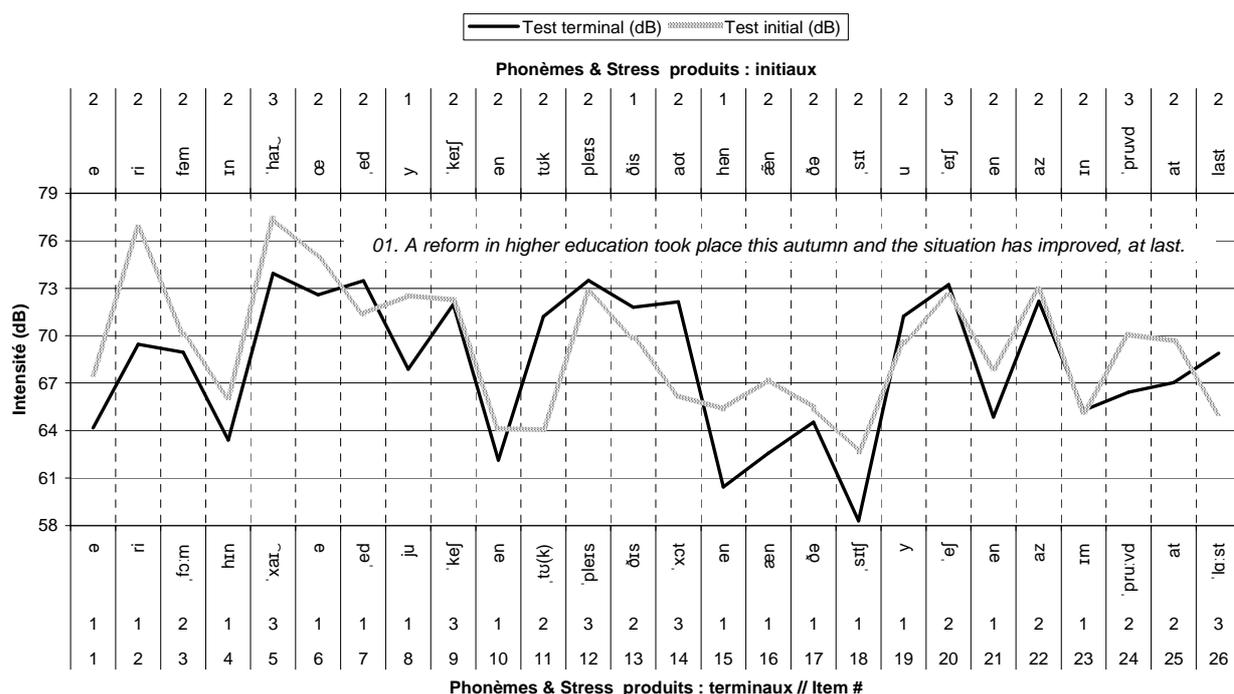


Figure 31. Evolution de l'intensité pour les 26 voyelles de Ph01 en T<sub>init</sub> et T<sub>term</sub>

La figure 31 montre deux courbes très similaires, finalement (n.b. : en vue d'une lisibilité accrue pour les besoins de l'analyse plus fine qui suit ci-après, nous avons modifié l'échelle des ordonnées et préféré ne pas lisser les courbes). Nous pouvons être au moins certains d'une chose ici : Jean n'a pas appliqué de changements majeurs dans le domaine de l'intensité. Ces courbes se croisent et se recroisent d'une syllabe à l'autre, mais sans jamais se démarquer franchement. De plus elles ne présentent pas de contrastes très marqués : l'étendue des valeurs va de 64 à 77 dB, soit une différence de 13 dB ; par contre, pour T<sub>ref</sub> : de 47 à 73 dB, soit un résultat de 26 dB, soit le double. Comme vu plus haut, l'intensité sonore double tous les 3dB, suivant une échelle logarithmique. Notre sujet fait varier la sienne d'un facteur légèrement supérieur à 4 ; le locuteur de référence utilise un facteur de 8,6.

Ceci dit, penchons-nous quand-même brièvement sur les points caractéristiques de ces courbes, ainsi que sur leurs convergences et leurs divergences.

- Sur la séquence V1-6, nous avons des mouvements de courbes parallèles. Toutefois, Jean a appliqué une correction utile de -7dB en V2-3 : déplacement de l'intensité de V2 vers V3. Les données perceptives confirment l'évolution du segment ['ri:fəm] vers le [ri:'fɔ:m] de référence. Noter l'intensité plus forte sur V5 : comme vu précédemment, Jean applique une aspiration très

prononcée de type germanique sur le /h/ de *higher*, jusqu'à le prononcer [x<sup>h</sup>] en T<sub>term</sub>.

- Correction, certes modeste, mais appropriée sur V7 : accent plus intense en T<sub>term</sub>. Assez bon traitement de *education* sur V7-10, même si le modèle requiert a priori une V9 plus intense que la V7.
- Bonne correction du syntagme verbal V11-12, mieux équilibré en intensité : on attend un accent sur les deux syllabes (St2+St3 ou St3+St2).
- Contraste renforcé sur *autumn*, avec une chute de -12dB entre V14 et V15 en T<sub>term</sub>, contre seulement -1dB en T<sub>init</sub>, mais avec l'intrusion inappropriée de l'affriquée vélaire [x].
- Bonne réalisation de la séquence V18-20 de *situation* pour les 2 courbes ; se épousent étroitement T<sub>ref</sub> toutes les deux.
- La réalisation du St3 attendu sur V24 était meilleure en T<sub>init</sub> qu'en T<sub>term</sub> ; en revanche, la finale en apposition V25-26 a été corrigée dans le sens du progrès en T<sub>term</sub>.

### **Conclusion sur l'intensité sonore des voyelles de la phrase Ph01**

Notre conviction est que l'intensité sonore reste un paramètre pertinent dans la production du rythme, et l'analyse du graphique de référence montre des relations indéniables. Il semblerait que ce soit un bon indicateur des valeurs extrêmes St1 et St3, et qu'elle ait plutôt tendance à correspondre à un nivellement des courbes sur les valeurs intermédiaires St2.

De toute évidence, Jean est très éloigné du modèle ici. L'amplitude de son contraste est deux fois plus faible que chez notre locuteur de référence et les variations de ses courbes d'intensité ne présentent pas de correspondances très marquées avec la courbe T<sub>ref</sub>. De plus, il n'y a pas eu d'évolution majeure entre T<sub>init</sub> et T<sub>term</sub>. Cependant nous nuancerons cela par le fait que l'analyse fine révèle de petits changements ou corrections dans T<sub>term</sub> qui, hormis en V24, vont tous dans une direction de progrès dans la gestion de l'intensité.

Nous avons décelé des progrès du *stress* de Jean plus marqués avec d'autres paramètres. À l'avenir, il faudra nous interroger de façon plus approfondie sur ce paramètre que le cadre restreint du présent travail nous le permet, afin de déterminer quelles approches pédagogiques peuvent être envisagées pour le faire progresser. Il nous faudra auparavant analyser les raisons qui font que la courbe de référence de l'intensité sonore, qui suit bien les contrastes majeurs du *stress*, ne les suit pas jusque dans le détail, comme le fait celle de F<sub>0</sub>. À regarder la phrase Ph01, l'intensité sonore semble moins « élastique » que la fréquence fondamentale, elle-même moins « élastique » que la durée.

## Chapitre 2. — Essai de représentation graphique synthétique du *stress* produit pour la phrase 01

### Introduction

Dans le cadre de notre approche didactique, liée à l'enseignement de l'anglais oral, il serait bon de pouvoir disposer d'un outil simple permettant de visualiser le *stress*, de le montrer dans toute la richesse de ses composants. En tirant les enseignements de l'étude approfondie du précédent chapitre, nous allons tenter de représenter ici la distribution *stress* dans la phrase 01, de la manière la plus lisible possible. Pour ce faire, nous avons utilisé le **graphique à bulles**, un outil prenant en compte les trois facteurs que nous venons d'étudier : durée vocalique, hauteur mélodique, intensité vocalique. Les graphiques à bulles utilisent un système de représentation à trois variables, que nous choisissons de faire correspondre à nos trois facteurs principaux composant chaque son vocalique : nous y trouvons :

- la fréquence fondamentale  $F_0$  en abscisse (70 Hz à 150 Hz)
- l'intensité sonore en ordonnée (40 dB à 85 dB ; incréments de 3 dB)
- la durée, dont dépend le diamètre de la bulle (30 à 300 ms)

Le mode de mise en œuvre de ces paramètres est illustré dans la figure 32.

Graphiques de distribution du stress dans PH01 : légende

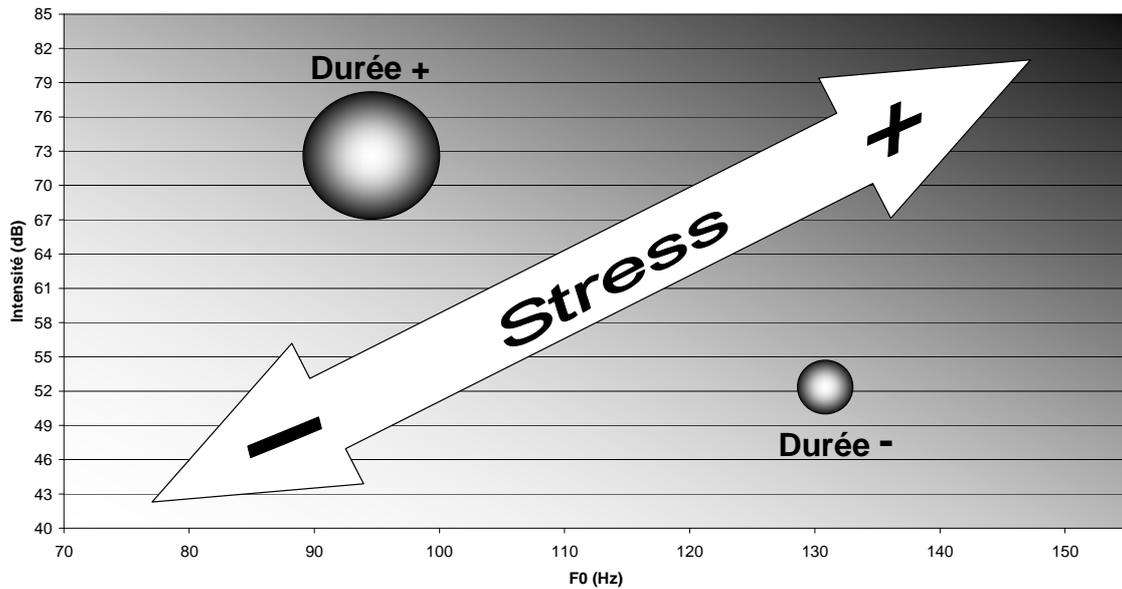


Figure 32. Légende du graphique de distribution du *stress* dans PH01

### Représentation du *stress* : vues d'ensemble de la phrase Ph01

Commençons par une série de trois figures mettant en œuvre ce système de représentation. Dans le but de permettre la meilleure comparaison possible, les trois graphiques ont tous la même échelle. Le premier graphique (figure 33) illustre la mise en relation des 3 facteurs du *stress* de la phrase 01 dans sa version initiale  $T_{init}$ . S'agissant d'un graphique intégrant l'ensemble des 26 phonèmes vocaliques de la phrase, nous n'avons pas fait figurer le texte, car il serait illisible du fait d'une trop grande superposition : c'est la silhouette de l'ensemble qui nous intéresse ici au premier chef.

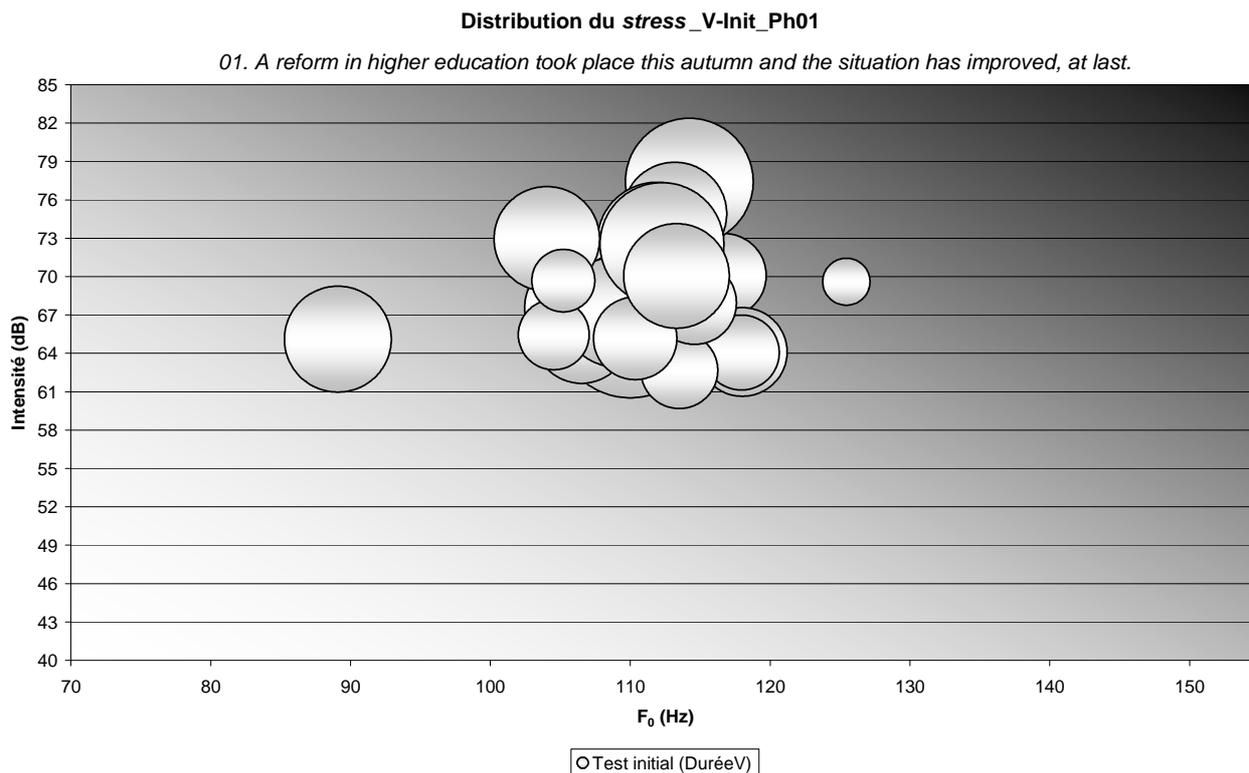


Figure 33. Représentation du *stress* pour les 26 voyelles de Ph01 en  $T_{init}$

Sur ce premier graphique, représentant la distribution du *stress* pour toutes les voyelles de la phrase, les items s'accumulent dans l'ordre chronologique : les premiers phonèmes (ici cachés) sont en dessous, les derniers au dessus. Nous constatons que l'ensemble des bulles (24 sur 26) se concentre sur une zone centrale comprise entre environ 105 Hz et 120 Hz en  $F_0$  (1 ton) et 63 dB et 78 dB (écart de 15 dB) en intensité (valeurs prises au centre des bulles), avec des bulles qui présentent une variation de taille très nette. Seules deux bulles sortent remarquablement du champ : celle de droite est plutôt brève (V19 : son [u] de [,sɪt·u·'eɪf·ən]) ; celle de gauche, dernière voyelle de la phrase, est plutôt longue : V26, son [a] de [last]. Mais notre propos n'est pas d'effectuer des mesures fines ici : il s'agit plutôt de comparer des tendances. Voici maintenant le graphique de la version  $T_{term}$ , en figure 34.

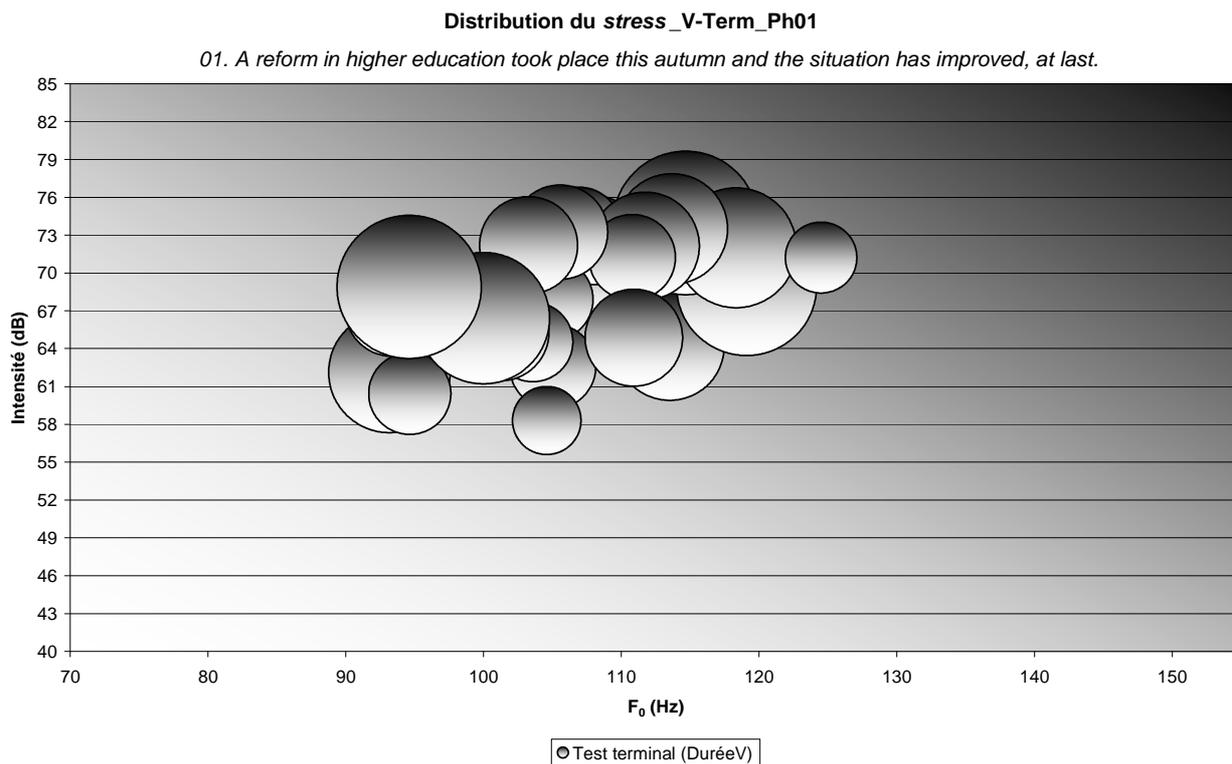


Figure 34. Représentation du *stress* pour les 26 voyelles de Ph01 en T<sub>term</sub>

Si nous comparons la figure 33 avec la figure 34, nous pouvons d'emblée faire trois constatations :

a) il n'y a plus de bulles isolées ;

b) le contraste de fréquence F<sub>0</sub> (étalement horizontal) est passé à un écart de 93 Hz à 125 Hz (2½ tons) ; l'ensemble des bulles s'est clairement étalé, mais la variation verticale de l'intensité est restée à peu près identique avec les valeurs extrêmes de 58 dB à 75 dB, soit un écart de 17 dB (valeurs lues au centre des bulles) ;

c) la taille des bulles est également plus contrastée, notamment avec un nombre plus grand de bulles de plus grand diamètre.

Considérons maintenant la figure 35, présentant les mêmes critères, mais dans la version de référence T<sub>ref</sub>.

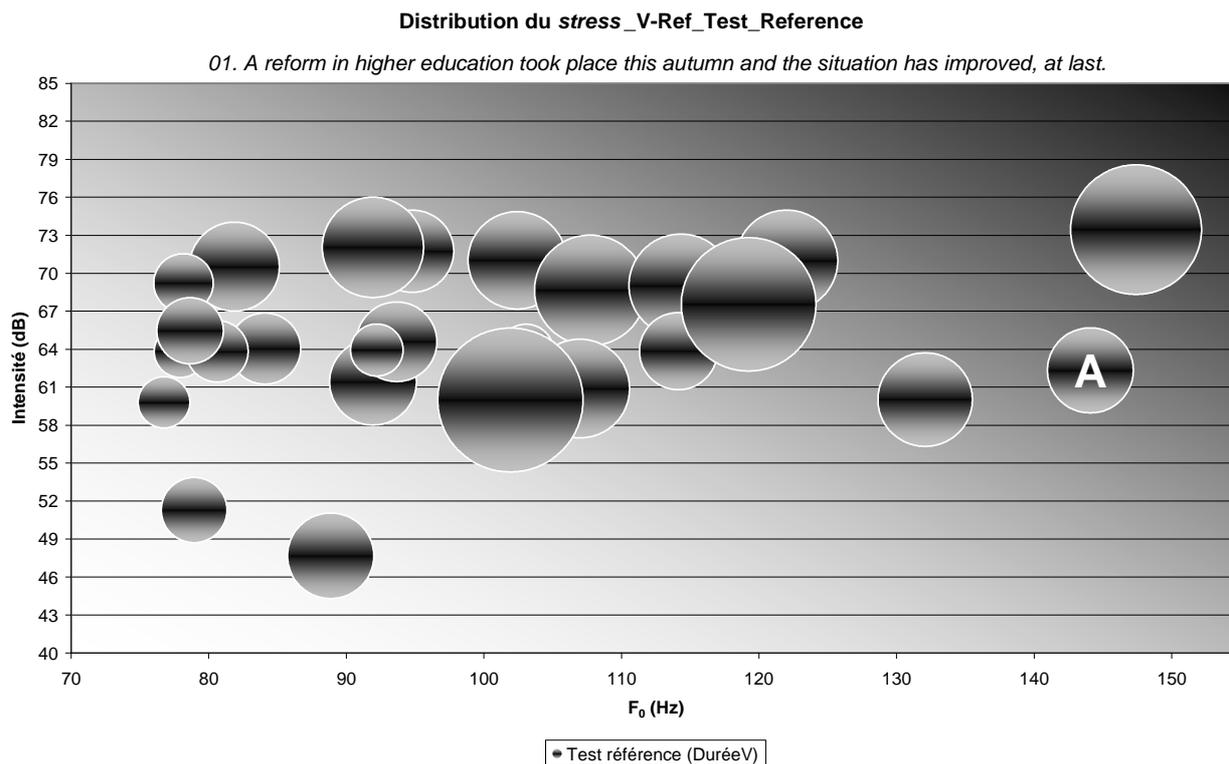


Figure 35. Représentation du *stress* pour les 26 voyelles de Ph01 en  $T_{ref}$

Les différences entre les deux graphiques précédents et celui-ci, illustrant le modèle de référence à approcher, apparaissent très clairement : les contrastes de fréquence (étalement horizontal), d'intensité (étalement vertical) et de durée (taille extrêmes des bulles, les petites notamment) sont tous plus importants.

Nous concluons cette première partie sur les graphiques à bulles en remarquant que, pour la phrase Ph01 au moins, ce mode de représentation permet de visualiser instantanément, d'une part, les progrès de Jean entre le test initial et le test terminal et, d'autre part, la distance importante qui le sépare encore du modèle de référence. Dans la figure 33, l'ensemble des phonèmes est concentré sur une zone étroite, typique de « l'accent français », cette tendance fortement isosyllabique à produire des accents réguliers égaux. La figure 35, au contraire, fait apparaître sans ambages, à travers « l'éclatement » des bulles, le contraste accentuel de la langue anglaise. Nous noterons que sur cette figure, la quasi-totalité des petites bulles se concentre sur le tiers gauche du graphique, la partie centrale et le tiers droit concentrant les bulles de moyenne et grande taille correspondant aux occurrences de *stress* les plus prononcées. Nous remarquerons cependant que notre locuteur de référence n'utilise pas systématiquement les trois paramètres du *stress* simultanément. Nous mentionnerons pour mémoire le point

remarquable **A** correspondant au cas d'une voyelle « courte »  $V_{\text{short}}$  sous  $St3$ , le verbe *took* : /'tok/. Son rendu apparaît logiquement dans la zone de *stress* fort, avec une bulle de taille plutôt petite correspondant à la brièveté relative de la voyelle.

### **Représentation du *stress* : analyse succincte de détails de la phrase Ph01**

Nous proposons maintenant de nous focaliser brièvement sur une analyse plus fine de ces graphiques synthétiques à travers certains syntagmes de la phrase Ph01 afin de déterminer si ces graphiques peuvent présenter un intérêt au-delà de celui que nous venons de mentionner. Du fait de la superposition des bulles, l'inconvénient majeur de ce choix de représentation est qu'il ne permet pas une lecture diachronique facile de la chaîne de parole, ce que permettait une animation (dépassant le cadre de ce travail écrit), par exemple. Un découpage arbitraire en syntagmes en facilitera la lecture, permettra la comparaison des trois versions de nos tests par superposition et autorisera un début de représentation de l'enchaînement diachronique sur les exemples choisis. Les graphiques ci-après utilisent ce découpage.

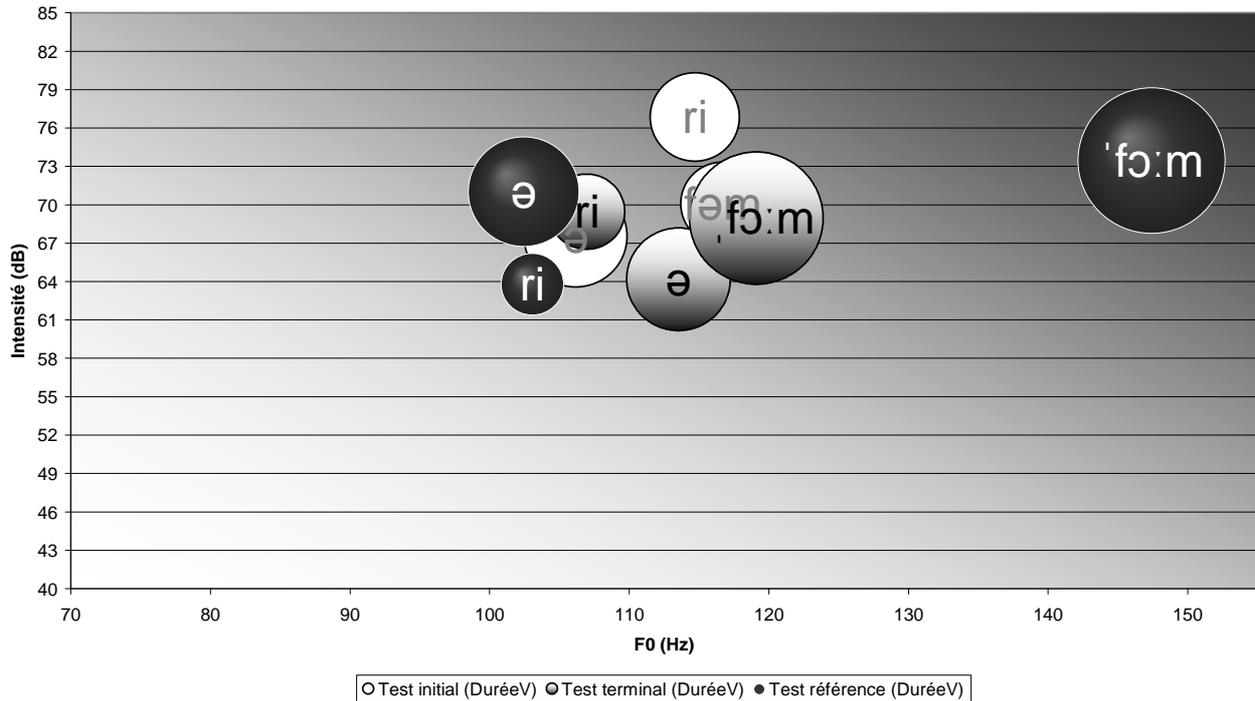
Nous proposons le découpage syntagmatique de la phrase Ph01 suivant :

1. *a reform*
2. a) *in higher*      b) *education*
3. *took place*
4. *this autumn*
5. a) *and the*      b) *situation*
6. *has improved*
7. *at last*

La figure 36 représente le premier syntagme de Ph01 dans les trois versions du test,  $T_{\text{init}}$ ,  $T_{\text{term}}$  de Jean et  $T_{\text{ref}}$  du locuteur de référence, soit les trois syllabes V1 à V3. L'échelle des graphiques précédents a été conservée. Le nombre de syllabes limité, combiné à la superposition des trois versions du test, permet une comparaison instantanée des productions.

Distribution du *stress* V-Init/V-Term/V-Ref\_Ph01\_Syntagme 1\_V1.1 à V1.3

01. *A reform in higher education took place this autumn and the situation has improved, at last.*



- La version  $T_{init}$ , en **blanc**, fait apparaître des bulles de tailles semblables et proches dans le plan. Nous devinons la petite taille de la bulle du phonème V3 [ə] de la prononciation initiale [ri·fɔ:m] pour *reform*. Les phonèmes V1 et V2 figurent dans des bulles de tailles supérieures à V3.
- La version  $T_{term}$ , en **gris**, fait apparaître un progrès léger dans l'étalement du syntagme le long de l'axe de  $F_0$ , mais une légère régression pour l'intensité. En revanche, la durée progresse dans le sens attendu, avec un gonflement de la bulle de V3, se combinant à une prononciation « correcte ». Noter la diminution de la V2 [ri] en réaction : nous mettons visuellement en évidence ici, de la façon la plus claire, le phénomène de réduction vocalique de la syllabe sous St1 précédent une syllabe sous St3. Le renversement en intensité du duo V2-V1 entre  $T_{term}$  devenant V1-V2 en  $T_{ref}$  peut s'expliquer par la persistance de l'intonation « classique » du français en début de phrase.
- La version  $T_{ref}$ , en **noir**, illustre le modèle de référence proposé. V3, sous St3 est projetée dans la zone de stress maximum, loin en haut à droite du plan, cumulant tous les maxima du syntagme, en durée, intensité et hauteur. Noter la réduction de V2, encore plus prononcée qu'en  $T_{term}$  et, bien entendu, le grand écart entre le duo rapproché V1-V2, et la voyelle puissante, mais isolée, V3.

Considérons maintenant brièvement un agrandissement de la zone de distribution vocalique de notre graphique, que nous obtenons grâce à un changement d'échelle de nos axes d'abscisse et d'ordonnée. Un agrandissement comme celui de la figure 36 bis permet bien évidemment une lecture encore plus fine des phénomènes décrits ci-dessus.

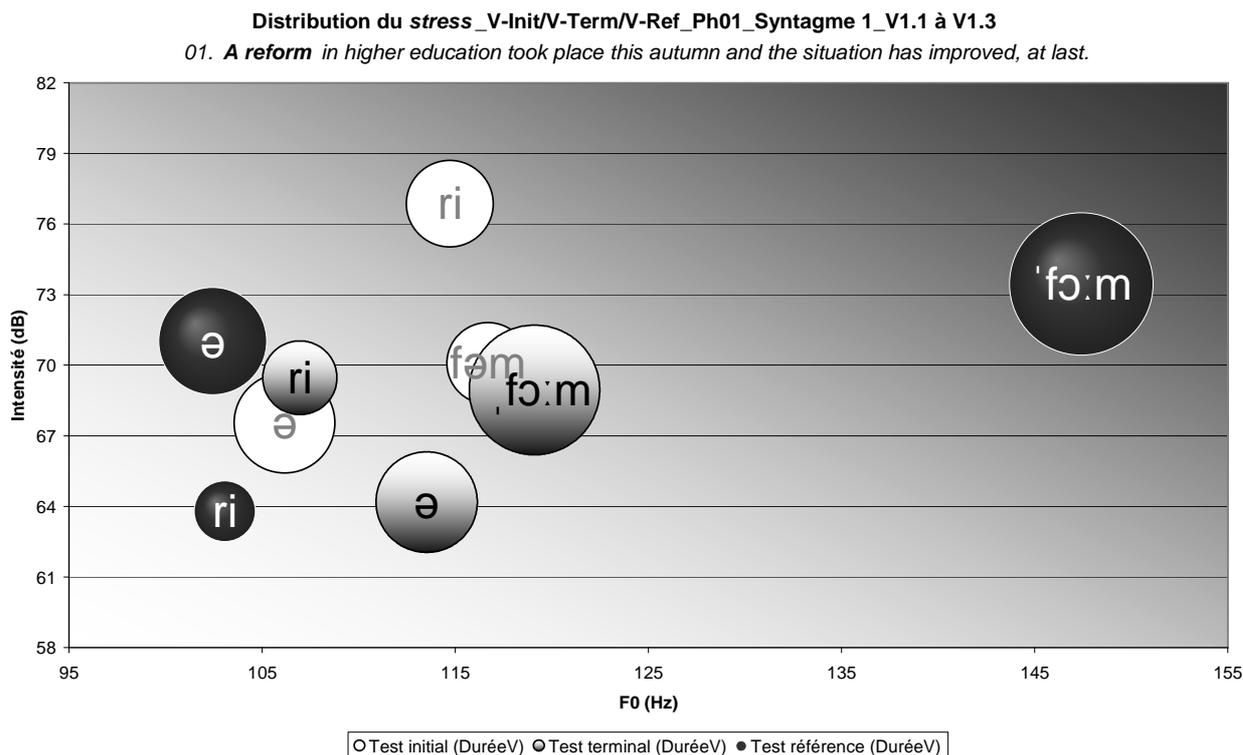


Figure 36 bis. Distribution du *stress* pour le syntagme #1 de Ph01 en  $T_{init}$ ,  $T_{term}$  et  $T_{ref}$ . Agrandissement.

Ainsi, nous pouvons dire que, pour ce syntagme, Jean a progressé dans la qualité de ses productions entre  $T_{init}$  et  $T_{term}$ . Nous en voulons pour preuve les éléments décrits ci-dessus. Nous constatons également qu'il lui reste une marge de progression conséquente pour approcher le modèle. La distance restant à franchir entre  $T_{term}$  et  $T_{ref}$ , la « marge de progression » peut aisément s'estimer en tendance pour chaque paramètre, phonème par phonème. Voici ce qui serait nécessaire dans le cas présent.

- V1 : plus d'intensité ; plus haute en fréquence
- V2 : moins d'intensité ; plus brève ; plus basse en fréquence
- V3 : plus d'intensité ; plus longue ; plus haute en fréquence

#### Représentation dynamique du *stress* dans les graphiques à bulles

Pour terminer ce chapitre, nous proposons d'améliorer notre méthode de représentation du *stress*. La figure 37 reprend le principe du précédent graphique, dans sa version « agrandie » pour en faciliter la lecture. Il présente le mot *education* de la phrase Ph01 dans les trois versions du test : initiale, terminale et de référence, pour les 4 voyelles V7 à V10 du mot.

Distribution du *stress* V-Init/V-Term/V-Ref\_Ph01\_Syntagme 2b\_V1.7 à V1.10

01. A reform in higher **education** took place this autumn and the situation has improved, at last.

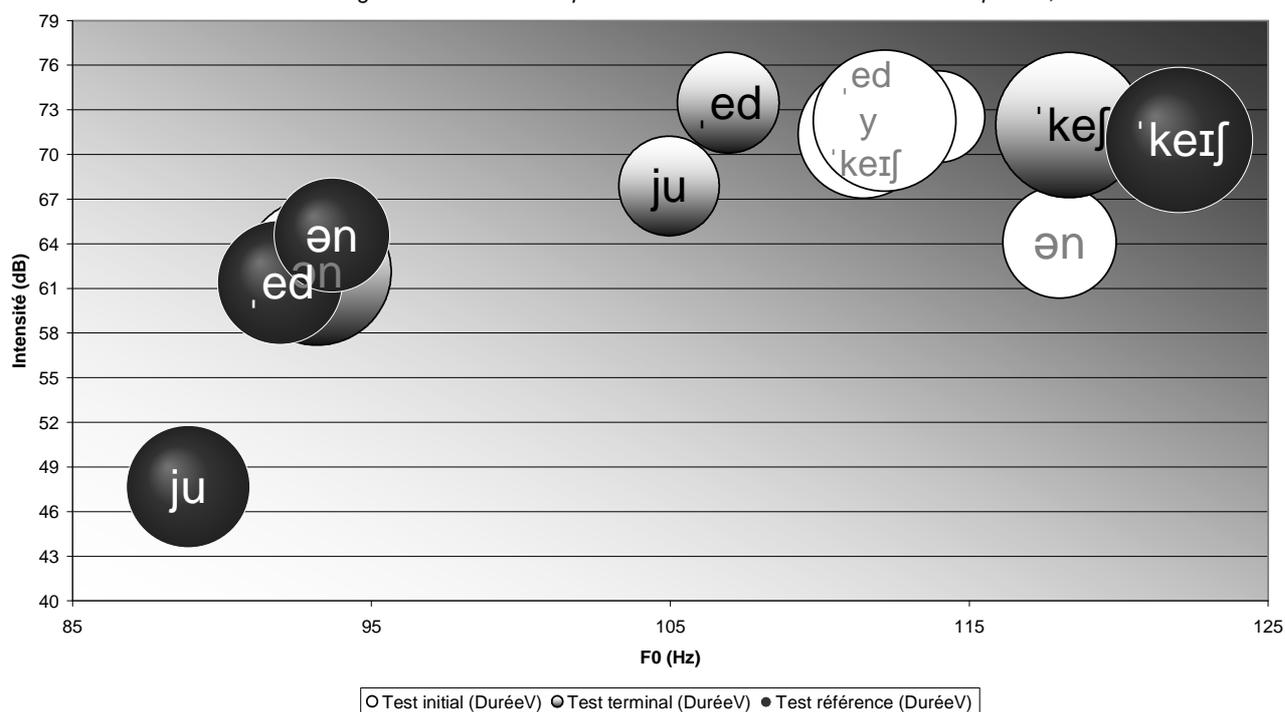


Figure 37. Distribution du *stress* pour le syntagme #2b de Ph01 en  $T_{init}$ ,  $T_{term}$  et  $T_{ref}$

Nous pouvons faire brièvement quelques remarques très similaires à celles déjà faites pour le syntagme #1. L'exemple est encore plus parlant ici :

- concentration forte, et même superposition des items en  $T_{init}$ , en intensité comme en fréquence, avec des bulles proches en taille, toutes caractéristiques fondamentales de l'iso-syllabisme
- progrès contrastifs très importants en  $T_{term}$ , dans les trois dimensions, et notamment en fréquence  $F_0$ ; une V9 sous St3 très proche du modèle, malgré que la  $V_{diph}$  /eɪ/ ne soit pas perçue comme réalisée ; une V10 proche de [œ:] en finale, qui « dure » un peu ici (bulle très gonflée pour une  $V_{short}$  supposée), pouvant découler de la présence du /n/ voisé en finale
- un modèle  $T_{ref}$  qui atteint les limites gauche, droite, haute et basse du plan graphique, soit un contraste maximum ; nous remarquerons toutefois que les trois  $V_{short}$  du mot ne sont pas très brèves (bulles moyennes), ceci pouvant s'expliquer :
  - V7 /e/ sous St2

- V8 /u/ en succession de l'approximante /j/ lui conférant presque une qualité de diphtongue
- V10 /ə/ suivie d'une nasale voisée en position finale du mot (notons que les V7 et V8 de Jean sont plus brèves que le modèle, alors que sa V10 est plus longue).

Une difficulté se pose toutefois : notre cerveau peut certes parvenir à une représentation mentale du cheminement d'une voyelle (ou syllabe) à l'autre, lors de notre lecture du graphique, mais la superposition des données des extraits de deux ou trois tests simultanément s'avère très rapidement délicate. Nous proposons dans la figure 38 de représenter ce cheminement par des flèches, qui permettront de suivre dynamiquement le « trajet vocalique » emprunté.

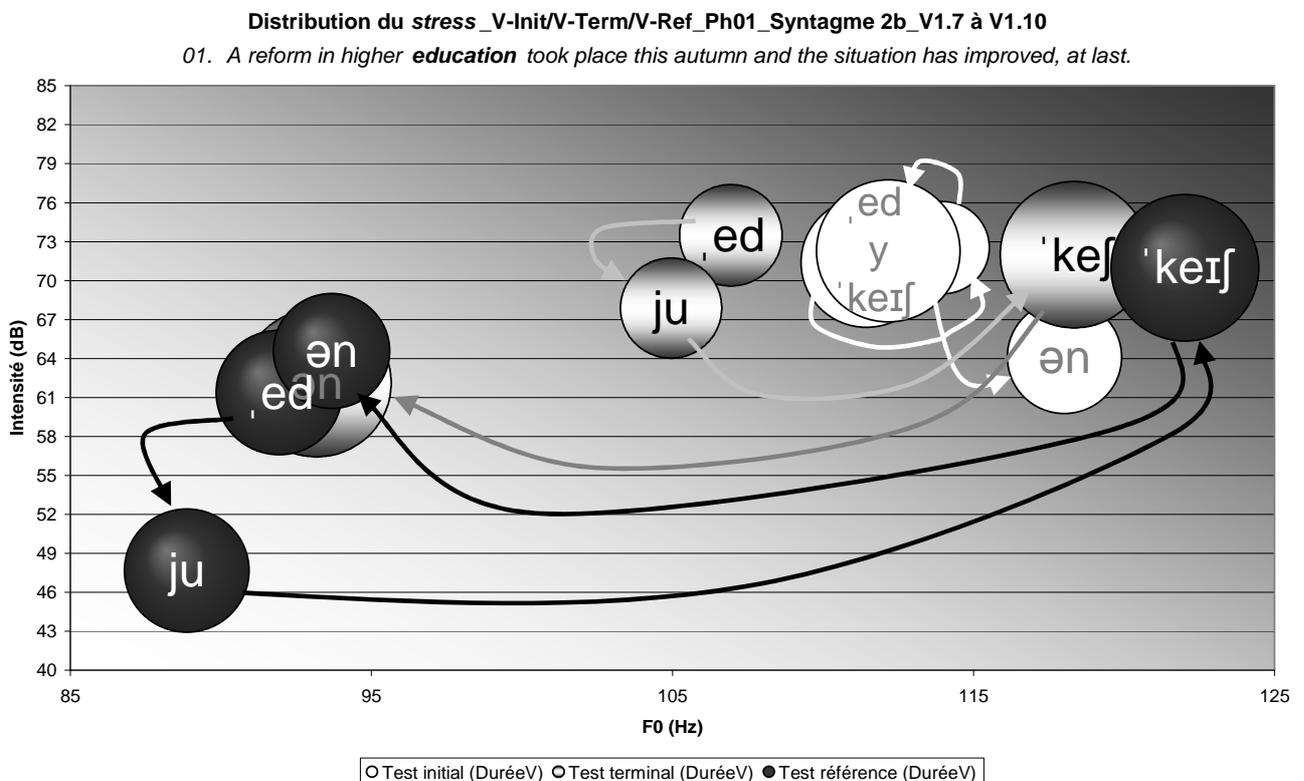


Figure 38. Distribution du *stress* pour le syntagme #2b de Ph01 en T<sub>init</sub>, T<sub>term</sub> et T<sub>ref</sub>. Représentation dynamique.

Le « cheminement » révélé en T<sub>init</sub> n'en est pas vraiment un : nos sons s'empilent, avant de s'écarter à peine en V10 finale, et dans la direction opposée à la référence. Par contre, T<sub>term</sub> adopte un cheminement vraiment très proche de T<sub>ref</sub>, mais avec un peu de contraste accentuel en moins, en début de mot. Notons le « bouclage » de la « boucle contrastive » de T<sub>ref</sub>, qui revient à son point de départ V7, en V10 finale. Le progrès de Jean est très important sur ce mot, cela « se voit » très explicitement.

### **Conclusion de nos essais de représentation synthétique du stress**

Nous représentons ainsi le *stress* comme une quantité tridimensionnelle, une densité d'énergie en un point à un instant donné. Il n'est pas surprenant que sa réalisation, pour le néophyte, ne soit pas simple ; a fortiori quand sa langue native « fonctionne » d'une façon très divergente. Jean a accompli des progrès certains dans cette phrase. Mais à partir de quel point seront-ils suffisants pour que l'interaction orale avec un locuteur natif se fasse sans gêne, sans perte de contenu ? Ce « degré d'efficacité interactive » est essentiel à l'évaluation des progrès d'un apprenant, mais il dépasse très largement le cadre du présent travail. Pour ce qui concerne l'anglais oral, le *stress* en est toutefois une composante fondamentale.

À l'avenir, s'ils le jugent utile ou nécessaire, l'enseignant et/ou l'étudiant pourront réfléchir aux stratégies didactiques les plus appropriées pour travailler l'un ou l'autre de ces aspects selon le degré de qualité de l'expression orale à atteindre. Le lecteur intéressé trouvera l'intégralité des graphiques à bulles des 9 syntagmes de la phrase Ph01 dans la **Table des illustrations annexes** en fin de volume. Cette méthode de représentation tridimensionnelle nous paraît « démystifier » le *stress* et elle a le mérite, du moins le croyons-nous, de poser clairement les données préalables à un apprentissage efficace. La version dynamique des graphiques à bulles permet de comprendre le cheminement et le traitement appliqué à la chaîne de parole par le locuteur anglophone, tout en affichant de façon limpide la « distance » que l'apprenant doit encore parcourir. Pour l'instant, ces graphiques restent limités par la superposition des données, ce qui en rend rapidement la lecture statique délicate, dès que le nombre d'items augmente.

## **Chapitre 3. — Analyse comparative des voyelles de 4 phrases du test (Ph01, Ph03, Ph05, Ph10)**

### ***Introduction***

Après avoir étudié la phrase Ph01 de façon assez approfondie, nous proposons d'élargir quelque peu notre champ d'investigation à trois autres phrases du test, de façon à déterminer si nous pouvons dégager des tendances fortes dans les progrès potentiels de Jean, notre sujet. Ne pouvant traiter de façon approfondie l'intégralité des enregistrements de Jean, dans tous les domaines susceptibles d'une évolution qualitative, dans le cadre

imparti, nous faisons ici le choix arbitraire d’approfondir l’étude de 4 phrases de notre test dans une approche globale. Il s’agit des phrases Ph01, Ph03, Ph05 et Ph10. Une autre approche aurait pu être de s’en remettre au hasard ou de choisir les plus dissemblables. L’une ou l’autre de ces approches aurait son intérêt. Nous expliquons notre choix ci-après. Voici les quatre phrases sélectionnées.

- Ph01.** *A reform in higher education took place this autumn and the situation has improved, at last.*
- Ph03.** *Engineering studies involve a lot of work on projects.*
- Ph05.** *The record of employment of graduates going through is very good.*
- Ph10.** *The year before had been quite ordinary, but a drop is recorded this month, and it’s sharp!*

Pour la présentation des ces phrases, nous renvoyons le lecteur au tableau de la figure 2, en introduction générale de cette partie de l’étude. Dans le présent chapitre, nous nous contenterons d’étudier les 3 paramètres composant le *stress* (durée,  $F_0$ , intensité)<sup>33</sup> au travers de l’ensemble des voyelles de ces quatre phrases, d’abord extraites, puis triées selon les critères de référence de voyelles diphtonguées  $V_{\text{diph}}$ , voyelles longues  $V_{\text{long}}$ , voyelles brèves  $V_{\text{short}}$  et voyelles faibles  $V_{\text{weak}}$ . Ce travail nous permettra de mettre en valeur l’évolution des voyelles, composantes des syllabes a priori les plus susceptibles de variation, entre les versions initiale  $T_{\text{init}}$  et terminale  $T_{\text{term}}$  de notre test. Les tableaux suivants (figures 39 et 40) montrent la distribution des phonèmes dans les quatre phrases.

Figure 39. Tableaux de distribution des niveaux de *stress* de référence et des phonèmes des phrases Ph01 et Ph03

Réf. Voyelle	Stress Norm.	Syllabe
V1.1	1	ə
V1.2	1	ri
V1.3	3	'fɔ:m
V1.4	1	ɪn
V1.5	2	'haɪ_
V1.6	1	ə
V1.7	2	,ed

Réf. Voyelle	Stress Norm.	Syllabe
V3.1	2	,endʒ
V3.2	1	ɪ
V3.3	3	'nɪər
V3.4	1	ɪŋ
V3.5	2	,stʌd
V3.6	1	ɪz
V3.7	1	ɪn

<sup>33</sup> NB : nous n’utiliserons pas le paramètre  $F_0$  pour les diphtongues, celui-ci étant par définition peu pertinent pour une voyelle dont le point d’articulation varie en permanence.

V1.8	1	ju
V1.9	3	'keɪf
V1.10	1	ən
V1.11	2	ˌtɒk
V1.12	3	'pleɪs
V1.13	1	ðɪs
V1.14	3	'ɔ:t
V1.15	1	əm
V1.16	1	ænd
V1.17	1	ðə
V1.18	2	ˌsɪtʃ
V1.19	1	u
V1.20	3	'eɪf
V1.21	1	ən
V1.22	1	həz
V1.23	1	ɪm
V1.24	3	'pru:vɪd
V1.25	1	ət
V1.26	3	lɑ:st

V3.8	3	'vɒlv
V3.9	1	ə
V3.10	3	'lɒt
V3.11	1	əv
V3.12	3	'wɜ:k
V3.13	1	ən
V3.14	3	'prɒdʒ
V3.15	1	ekts

Figure 40. Tableaux de distribution des niveaux de *stress* de référence et des phonèmes des phrases Ph05 et Ph10

Réf. Voyelle	Stress Norm.	Syllabe
V5.1	1	ðə
V5.2	3	'rek
V5.3	1	ɔ:d
V5.4	1	əv
V5.5	1	ɪm
V5.6	3	'plɔɪ
V5.7	1	mənt
V5.8	1	əv
V5.9	3	'grædʒ
V5.10	1	u_əts
V5.11	2	ˌkʌm
V5.12	1	ɪŋ
V5.13	3	'θru:
V5.14	1	ɪz
V5.15	3	'ver
V5.16	1	ɪ
V5.17	2	ˌgʊd

Réf. Voyelle	Stress Norm.	Syllabe
V10.1	1	ðə
V10.2	3	'jɪə
V10.3	1	bɪ
V10.4	3	'fɔ:
V10.5	1	həd
V10.6	3	'bi:n
V10.7	2	ˌkwɔ:t
V10.8	3	'ɔ:d
V10.9	1	ən
V10.10	1	ər_
V10.11	1	i
V10.12	2	ˌbʌt
V10.13	1	ə
V10.14	3	'drɒp
V10.15	1	ɪz
V10.16	1	ri
V10.17	3	'kɔ:d
V10.18	1	ɪd
V10.19	1	ðɪs
V10.20	3	'mʌntθ

V10.21	2	,ænd
V10.22	1	its
V10.23	3	'fɑ:p

Ces 4 phrases présentent le point commun d'avoir approximativement un rapport **Nombre de St3/Nombre de syllabes** d'environ 0,3, tout en ayant ce nombre de syllabes qui varie assez fortement : de 15 à 26 syllabes. La collecte de toutes les voyelles de ces quatre phrases permettra d'en avoir un nombre minimum pour chacune des quatre catégories, que nous pourrons ainsi comparer, ce que l'étude d'une seule phrase ne nous permet pas. L'étude de l'évolution du comportement des voyelles sous les différentes instances de *stress* (St1, St2 et St3) pourra présenter un intérêt. Notons qu'ici nous n'utiliserons pas d'enregistrements de référence, comme nous l'avons fait dans le chapitre 1 de cette partie, mais uniquement les valeurs de référence du *stress* et des phonèmes théoriques, comme moyen de comparaison. Nous observerons l'évolution des productions de Jean ente  $T_{init}$  et  $T_{term}$  à l'aune de ces points de repère.

Tous les graphiques que nous présentons ci-après ont les points communs suivants, de bas en haut (ex. figure 41) :

- axe des abscisses inférieur (principal) :
  - Référence des voyelles (ex. V5.11 = 11<sup>ème</sup> voyelle de la 5<sup>ème</sup> phrase)
  - *Stress* perçu et phonèmes produits par Jean, issus du test  $T_{term}$ , pour chaque voyelle
- axe des abscisses supérieur (secondaire) :
  - Phonèmes attendus, en anglais britannique *Received Pronunciation* (£-RP)<sup>34</sup>, selon le dictionnaire de prononciation de Wells (Wells, 2008)
  - *Stress* perçu et phonèmes produits par Jean, issus du test  $T_{init}$ , pour chaque voyelle
- zone de graphique :
  - En bas : niveaux de *stress* superposés dans les trois versions *Stress\_Norm* (en histogramme), *Stress\_Init* et *Stress\_Term*, selon la légende
  - En haut : données mesurées pour chaque phonème en version  $T_{term}$  (triangle blanc :  $\Delta$ ) et  $T_{init}$  (rond noir :  $\bullet$ ) ; les barres de hausse séparant les données  $T_{term}$  et  $T_{init}$  apparaissent en blanc quand  $T_{term} > T_{init}$  (données  $T_{term}$   $\Delta$  au dessus) et hachurées sombres quand  $T_{term} < T_{init}$  (données  $T_{init}$   $\bullet$  au dessus). L'ordre retenu pour la présentation des résultats est celui d'un

---

<sup>34</sup> C.f. nos **Notes liminaires/Variétés de l'anglais**, en introduction de cette étude.

classement effectué selon la plus grande valeur positive de la différence  $T_{\text{term}} - T_{\text{init}}$  vers la plus petite, sur l'axe des abscisses

- Tous les graphiques utilisent les mêmes échelles, de durée (0 — 310 ms), de  $F_0$  (70 — 150 Hz), d'intensité (50 — 83 dB, par incréments de 3dB).<sup>35</sup>

Une conclusion générale sur l'étude de l'ensemble des voyelles se trouve en fin de chapitre. N.B. : le fait que la référence requière une diphtongue (ou une voyelle longue, courte ou faible) ne garantit en rien que Jean ait produit une telle voyelle, bien entendu. Il arrive d'ailleurs assez fréquemment que la référence et la production de Jean diffèrent.

### ***Analyse comparative des voyelles de 4 phrases du test : diphtongues***

#### **Introduction à l'analyse des diphtongues**

Les diphtongues sont des voyelles et la qualité de leur réalisation présente un grand intérêt pour la démarche didactique qui nous anime mais, de par leur nature, elles sont plus difficiles à appréhender. Se définissant comme une voyelle dont le point d'articulation varie constamment, produit d'un glissement d'un son vers un autre (*glide*), elles sont relativement caractéristiques de la langue anglaise, même si elles existent dans d'autres langues. Leur réalisation, complexe, représente une difficulté majeure pour les LN francophones. Une diphtongue non réalisée sera souvent au mieux une source d'incompréhension, au pire une source de confusion (ex. *Stop shouting!* — /aʊ/ prononcé /u:/ — comme *Stop shooting!*). Cet aspect de la langue anglaise orale est d'une telle importance qu'il mérite un travail pédagogique spécifique en cours.

Par ailleurs, la nature particulière des diphtongues nous dispensera ici d'étudier leur  $F_0$ , par définition en constante variation. Pour très intéressante qu'elle soit, cette étude dépasserait le cadre du présent travail. Aussi nous contenterons nous des données phonémiques et accentuelles perçues par l'enseignant.

#### **Analyse des graphiques de durée et d'intensité des diphtongues de Ph01, Ph03, Ph05 et Ph10**

Les graphiques des figures 41 et 42 font ressortir les différences existant entre les deux productions de Jean, entre  $T_{\text{init}}$  et  $T_{\text{term}}$ , pour les critères de durée, de hauteur mélodique  $F_0$  et d'intensité appliqués aux  $V_{\text{diph}}$ . Ces critères s'appliquent ici à l'ensemble

---

<sup>35</sup> Rappel : cet incrément est utilisé pour sa praticité, l'intensité sonore doublant tous les 3dB.

des voyelles de référence de type  $V_{\text{diph}}$  présentes dans les quatre phrases. La même approche vaudra pour les  $V_{\text{long}}$ ,  $V_{\text{short}}$  et  $V_{\text{weak}}$ .

Les quatre phrases sélectionnées contiennent un total de 9  $V_{\text{diph}}$ . Pour que la palette des  $V_{\text{diph}}$  de référence élémentaires de l'anglais RP soit complète (nos phrases en contiennent 5 sur 8), il faudrait ajouter /əʊ/, /aʊ/ et /eə/, absentes ici. Notons que la  $V_{\text{diph}}$  /ʊə/ apparaît sous une forme atténuée /u\_ə/, voyelle (V5.10) de la syllabe finale du mot *graduates*. L'ordre de présentation des données dans les graphiques étant guidé par la soustraction entre  $T_{\text{term}}$  et  $T_{\text{init}}$ , il est normal qu'il soit différent d'un graphique à l'autre.

Ainsi, la figure 41 classe les dix  $V_{\text{diph}}$  selon leur durée réciproque  $T_{\text{init}}$  et  $T_{\text{term}}$ . Comme nous l'avons évoqué plus haut, la caractéristique première d'une  $V_{\text{diph}}$  est la complexité de sa réalisation, avant sa durée. Néanmoins, du fait de cette complexité et de son utilisation principalement sous *stress* fort de type St2 et St3, sa production demandera nécessairement à peu près autant de temps qu'un son simple long. Aussi retiendrons-nous le critère de durée comme pertinent pour l'analyse de la production de Jean, au côté de la qualité vocalique.

Durée brute\_V-diph\_Julien\_Phrases 01-03-05-10

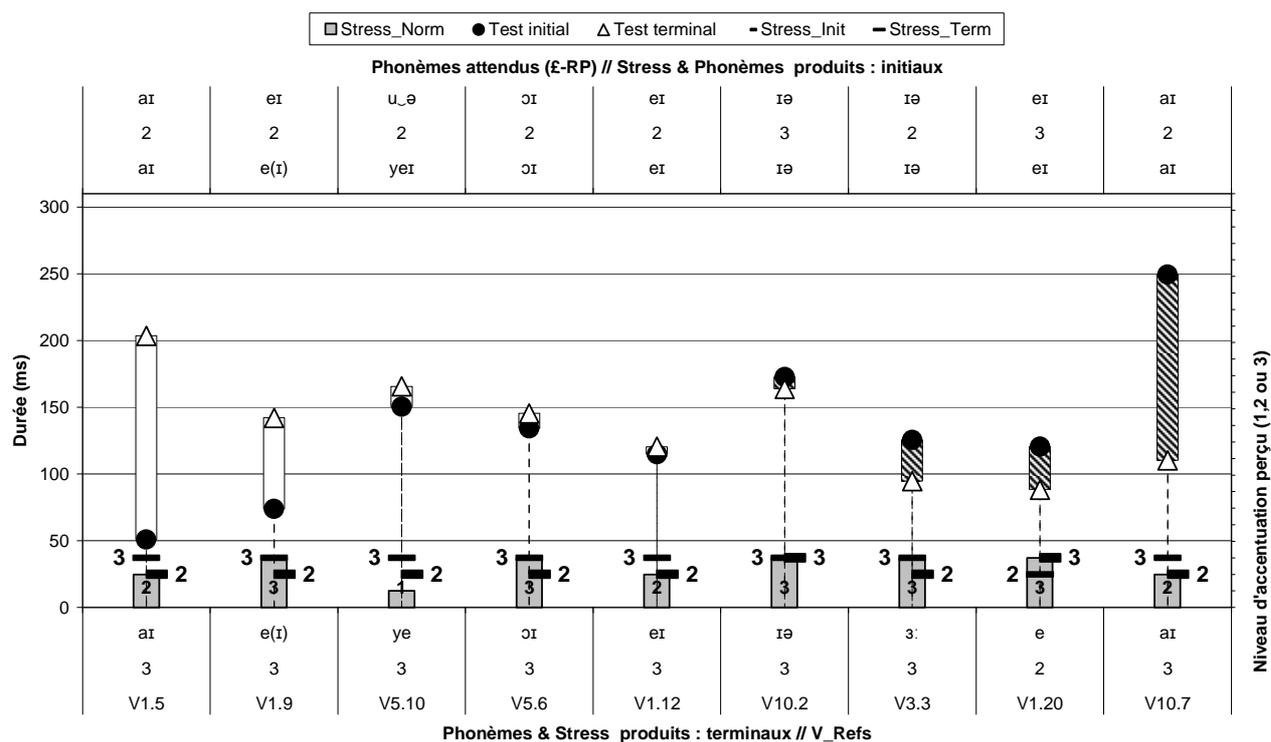


Figure 41. Durée brute des diptongues en  $T_{init}$  et  $T_{term}$ . Phrases 01-03-05-10.

La figure 41 fait apparaître une répartition quasi-symétrique des différences de durée, ce qui tend à montrer qu'il y a eu, en moyenne, autant de progrès vers la référence que de régression. Ceci est d'ailleurs très bien illustré par le fait que la première voyelle du graphique est identique à la dernière (son /ai/), mais dans un résultat inversé. Dans une moindre mesure, nous retrouvons le phénomène avec les trois sons /eɪ/ en V1.9, V1.12 (au centre) et V1.20.

Pour les autres sons de la moitié « positive », à gauche, nous remarquons que, hormis le V5.10 de *graduates* qui conserve le /y/ français en  $T_{init}$  et  $T_{term}$ , les diptongues sont bien réalisées (notons qu'il s'agit ici de la seule  $V_{diph}$  sous St1 de référence, mais qu'un *stress* fort a été perçu). En revanche, la moitié « négative » (droite), la qualité perçue par l'enseignant est meilleure (plus proche de la référence) en  $T_{init}$  qu'en  $T_{term}$  (V3.3, V1.20).

Rappelons ici que le travail effectué en cours portait en premier lieu sur la dynamique syllabique, la qualité vocalique venant dans un deuxième temps. Il peut être normal que pour un sujet dont les capacités cognitives sont mobilisées par la réalisation du *stress*, la réalisation qualitative des voyelles soit un temps perturbée. Quoi qu'il en soit,

nous ne pouvons pas dire que, ni dans le domaine de la durée des  $V_{\text{diph}}$ , ni dans celui de leur qualité, il y ait eu progrès. Considérons maintenant le paramètre de l'intensité sonore de ces mêmes  $V_{\text{diph}}$ , en figure 42.

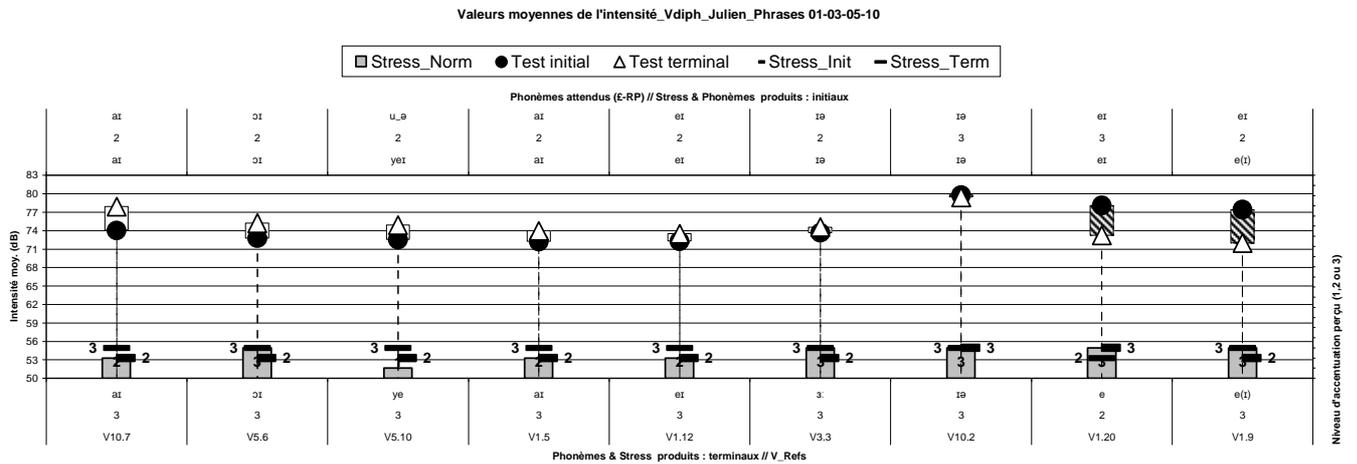


Figure 42. Intensité des diphtongues en  $T_{\text{init}}$  et  $T_{\text{term}}$ . Phrases 01-03-05-10.

Comme vu ci-dessus, huit de ces neuf  $V_{\text{diph}}$ , sont sous *stress* fort, de type St2 ou St3, pour lesquelles serait attendue une augmentation importante de l'intensité. Ce qui apparaît avant tout ici, c'est la petite amplitude des variations (les plus grandes sont dans la partie droite de « régression »), signe de différences faibles, presque non-significatives.

La seule  $V_{\text{diph}}$  sous St1 théorique (V5.10) affiche une intensité légèrement augmentée en  $T_{\text{term}}$  (elle a d'ailleurs été perçue comme étant sous un *stress* fort par l'enseignant en  $T_{\text{init}}$  comme en  $T_{\text{term}}$ ), ce qui ne va pas non plus dans le sens d'un progrès de Jean.

La difficulté de la réalisation de la diphtongue, notamment pour un francophone, pourrait expliquer ces résultats montrant un progrès stationnaire ou une régression. Considérons maintenant le cas des voyelles longues des quatre phrases sélectionnées.

## ***Analyse comparative des voyelles de 4 phrases du test : voyelles longues***

### **Introduction à l'analyse des voyelles longues**

Les voyelles longues ( $V_{long}$ ), de part leur simplicité, stabilité et longue durée relatives, peuvent constituer un terrain de recherche privilégié. Pour l'enseignant d'anglais L2, à l'instar des diphtongues, elles représentent un défi didactique. D'expérience, nous savons que l'idée de prolonger un son vocalique dans une L2 au-delà de sa durée coutumière dans sa langue maternelle ne va pas de soi. De notre point de vue, en anglais, l'allongement vocalique, tout comme l'abrègement d'ailleurs, est un corollaire du *stress* et, par conséquent du contraste de durée que nous recherchons. Ce contraste est tout autant le résultat d'un allongement de certaines voyelles, que de l'abrègement de certaines autres. Pour mémoire, les  $V_{long}$  se trouvent presque toujours sous *stress* de niveau St2 ou St3.

Comme cela a déjà été vu, la dynamique accentuelle permet la mise en valeur d'un son, par sa qualité, sa durée, son intensité et sa hauteur. Pour ce qui concerne les voyelles longues de l'anglais RP, au nombre de cinq, produire un son de la bonne qualité n'est pas très difficile pour un francophone : les sons /ɑ:/, /ɔ:/, /i:/, /u:/ et /ɜ:/ sont tous assez proches de phonèmes « standards » français. L'une des difficultés majeures est donc de produire cet allongement, là aussi indispensable à une interaction efficace avec un LN anglophone. En anglais oral, la quantité vocalique qui tranchera entre le fruit « pêche » (*peach*) et le « terrain de football » (*pitch*) n'est l'affaire que d'une poignée de centisecondes.

Mais il ne faudrait surtout pas que le terme de « voyelle longue » occulte les autres aspects du *stress*. S'agissant de voyelles presque toujours sous *stress* fort, il nous sera nécessaire de vérifier si une fréquence fondamentale  $F_0$  et une intensité élevées sont également présentes. Notons que le dernier chapitre de cette partie sera consacré à l'analyse de la composition spectrale de quatre des cinq voyelles longues dans les 12 phrases du test, en versions  $T_{init}$  et  $T_{term}$ . Notons également que les cinq sons  $V_{long}$  de base de l'anglais RP sont présents dans les quatre phrases choisies.

### **Analyse des graphiques de durée, de hauteur mélodique et d'intensité des voyelles longues de Ph01, Ph03, Ph05 et Ph10**

Pour ce qui concerne la durée des  $V_{long}$  de ces quatre phrases, nous pouvons voir sur la figure 43 que Jean a tout à fait intégré ce paramètre dans sa production terminale : sur les douze  $V_{long}$  que contiennent ces phrases, huit d'entre elles ont été allongées dans la version  $T_{term}$ , dans une fourchette allant d'environ +50 % à +150 %, ce qui est

considérable. Deux d'entre elles n'ont pas évolué. Deux ont régressé. Considérons brièvement ces deux dernières :

- V10.6 se trouve au cœur du syntagme verbal dans *the year before had been quite ordinary*. Ici, la référence voudrait qu'un son long /i:/ soit produit ; toutefois, dans la pratique courante, du fait de la très haute fréquence d'apparition du terme (verbe être *to be*), la prononciation /ɪn/ est très usitée et ne constitue aucunement un obstacle à la compréhension.
- V1.14 : le recours à une diphtongue peut-être issu de la pratique de l'espagnol L3 par Jean dans la première syllabe du mot *autumn*. Ce point a déjà été discuté lors de notre analyse de la Ph01. Jean a ajouté le phonème [x] en initiale du mot en T<sub>term</sub>, phonème qui n'a pas été retranscrit comme « voyelle » ici. Nous pouvons considérer ce point comme un épiphénomène, qui devra toutefois être pris en considération.

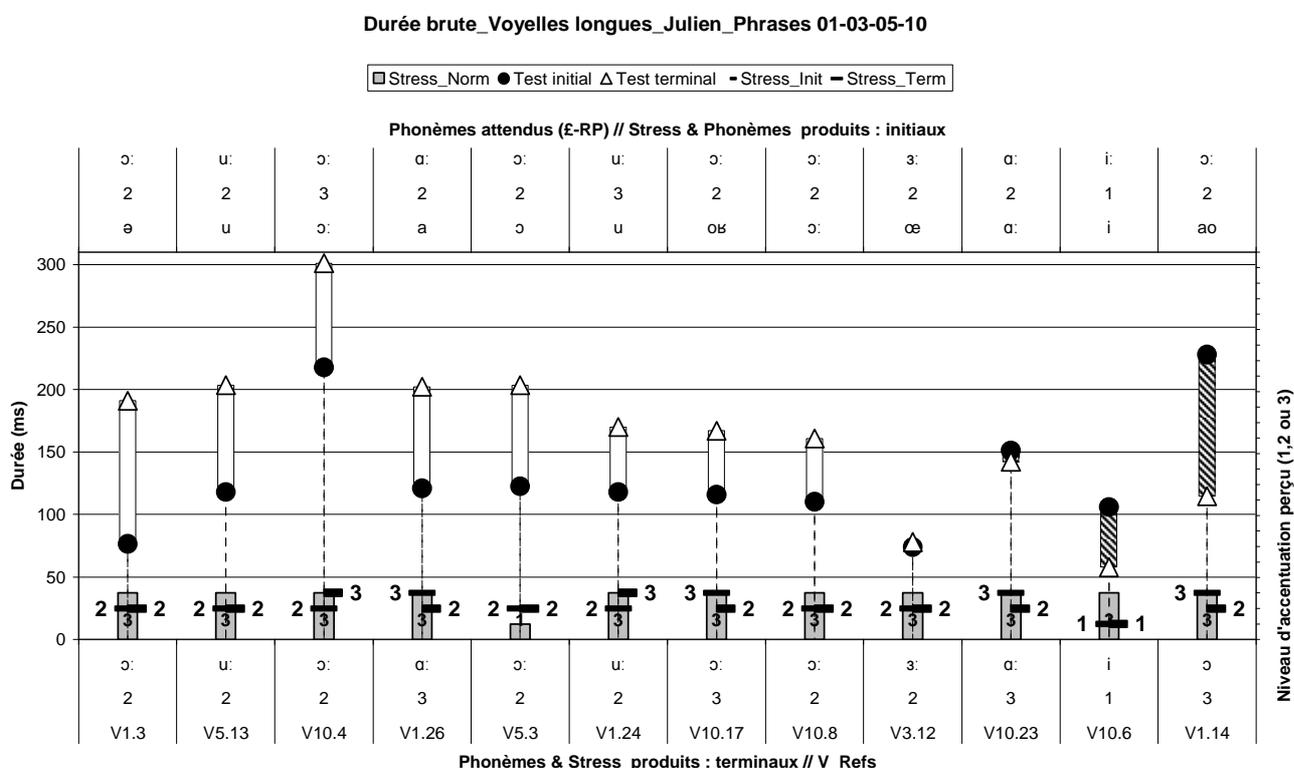


Figure 43. Durée brute des voyelles longues en T<sub>init</sub> et T<sub>term</sub>. Phrases 01-03-05-10.

Ainsi, pour ce qui concerne la durée vocalique des V<sub>long</sub>, nous pouvons avancer que Jean a bien intégré cette composante du *stress* dans ses productions terminales des quatre phrases. Voyons maintenant ce qu'il en est des deux autres paramètres : la hauteur mélodique et l'intensité. La figure 44 présente, selon le même principe, un classement des données mesurées de F<sub>0</sub> pour la même sélection de voyelles.

Valeurs moyennes de F0 des Vlong\_Julien\_Phrases 01-03-05-10

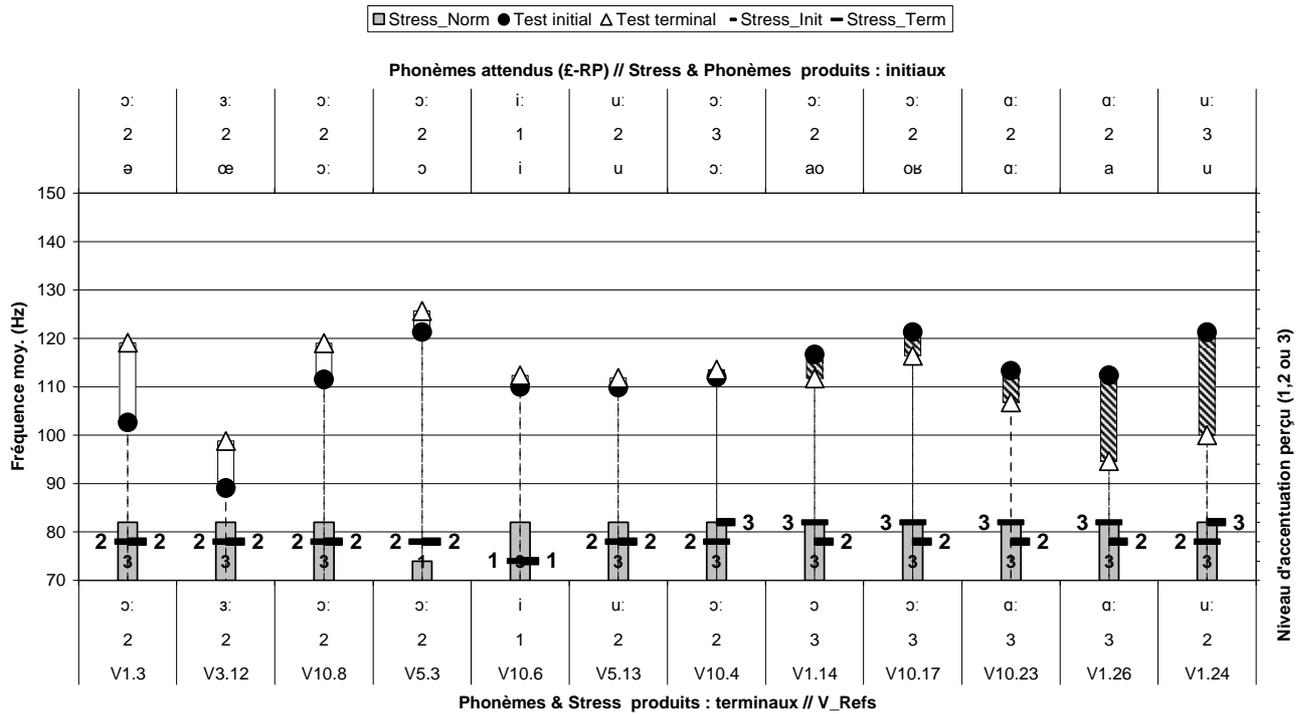


Figure 44. F<sub>0</sub> des voyelles longues en T<sub>init</sub> et T<sub>term</sub>. Phrases 01-03-05-10.

Nous remarquons d'emblée que l'équilibre du graphique penche plus vers la partie négative (droite). Nous attendions une progression vers des fréquences plus élevées en T<sub>term</sub> sur une majorité de voyelles. Sans plus nous attarder, nous concluons que l'objectif atteint avec le paramètre « durée » ne l'est pas pour F<sub>0</sub>. La figure 45 nous permet de voir ce qu'il en est de l'intensité sonore. Là encore, le résultat est plutôt décevant : même s'il est plus équilibré entre le positif et le négatif, nous ne pouvons pas en déduire une tendance forte vers un progrès de notre sujet.

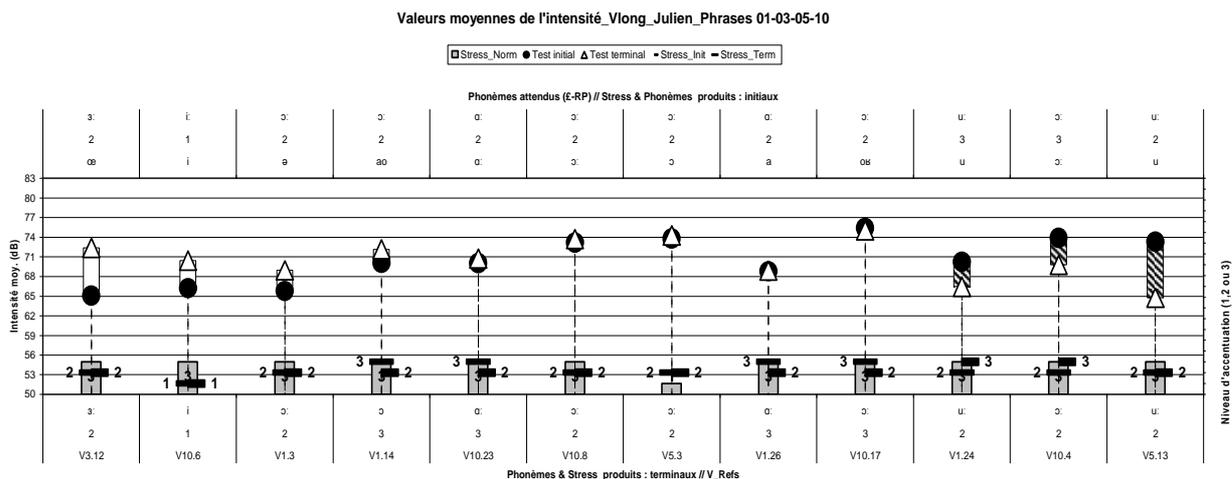


Figure 45. Intensité des voyelles longues en  $T_{init}$  et  $T_{term}$ . Phrases 01-03-05-10.

De ces trois graphiques, nous retiendrons que Jean a pris en compte de façon très positive le paramètre le plus évident de l'application du *stress* sur cette catégorie de voyelles longues : leur durée. De prime abord, nous aurions pensé que c'était l'intensité qui aurait été la plus facile à intégrer, le parler « plus fort » appliqué aux syllabes sous accent ; force est de constater que ce n'est pas le cas pour les exemples que nous avons retenus. Quant aux variations de la fréquence fondamentale, source de la mélodie de la langue, nous ne doutons pas qu'elles soient difficile à acquérir. Ces aspects seront à intégrer dans la conception de nos futurs cours et des supports pédagogiques correspondants.

### ***Analyse comparative des voyelles de 4 phrases du test : voyelles courtes et voyelles faibles***

Dans le cadre de la présente étude, nous avons effectué les mêmes mesures que ci-dessus ( $V_{long}$ ) pour les deux catégories de voyelles restantes des quatre phrases sélectionnées, à savoir les voyelles brèves  $V_{short}$  et les voyelles faibles  $V_{weak}$ . L'ensemble des graphiques correspondants sont présentés dans la partie « annexes » de ce volume. En effet, rappelons que notre but n'est pas l'exhaustivité dans l'exploitation de nos données, mais de tester des méthodes d'analyse en vue d'une éventuelle future exploitation plus approfondie du corpus dont notre sujet, Jean, est issu.

Nous avons vu que Jean avait effectué des progrès dans certains domaines et qu'il était resté stationnaire ou était en légère régression dans d'autres, dans sa réalisation du *stress* : son interlangue n'est pas fossilisée et, avec un travail approprié, il pourra encore progresser. Cette constatation vaut également pour ses réalisations des  $V_{short}$  et  $V_{weak}$ .

Le traitement des ces voyelles est plus complexe et plus long que celui des  $V_{diph}$  et, surtout, des  $V_{long}$ , les plus faciles à appréhender, comme nous l'avons vu. Cela vient du fait qu'elles sont beaucoup plus nombreuses que les autres et que, pour les  $V_{short}$ , elles peuvent se trouver sous tous les niveaux de *stress*. Pour ces dernières, ce sera avant tout la fréquence et l'intensité, plus que la durée, qui déterminera leur niveau de *stress* effectivement produit : St1, St2 ou St3. Pour les  $V_{weak}$ , les choses devraient être un peu plus simples, puisqu'elles ne seront réalisées a priori que sous *stress* de bas niveau.

Nous ne prendrons que quelques exemples choisis parmi toutes nos données pour illustrer notre propos et les progrès de Jean dans ce domaine, ce qui clora ce chapitre.

Les graphiques des figures 46 et 47 présentent, à eux deux, les données de durée brute pour l'intégralité des  $V_{short}$  des quatre phrases sélectionnées. Nous avons été obligé de les scinder en deux séries, à cause de leur nombre : série 1 : /ɪ/ et /e/, les plus fréquentes dans nos phrases et les plus courantes sous St1 ; série 2 : /æ/-/ɒ/-/ʌ/-/ʊ/, moins représentées, mais les plus courantes sous St2 et St3. Prenons soin de noter que, s'agissant de  $V_{short}$ , ces graphiques doivent être interprétés dans le sens inverse des graphiques de durée des  $V_{diph}$  et des  $V_{long}$  : en effet, c'est la brièveté de la voyelle qui est recherchée ici, non plus son extension.

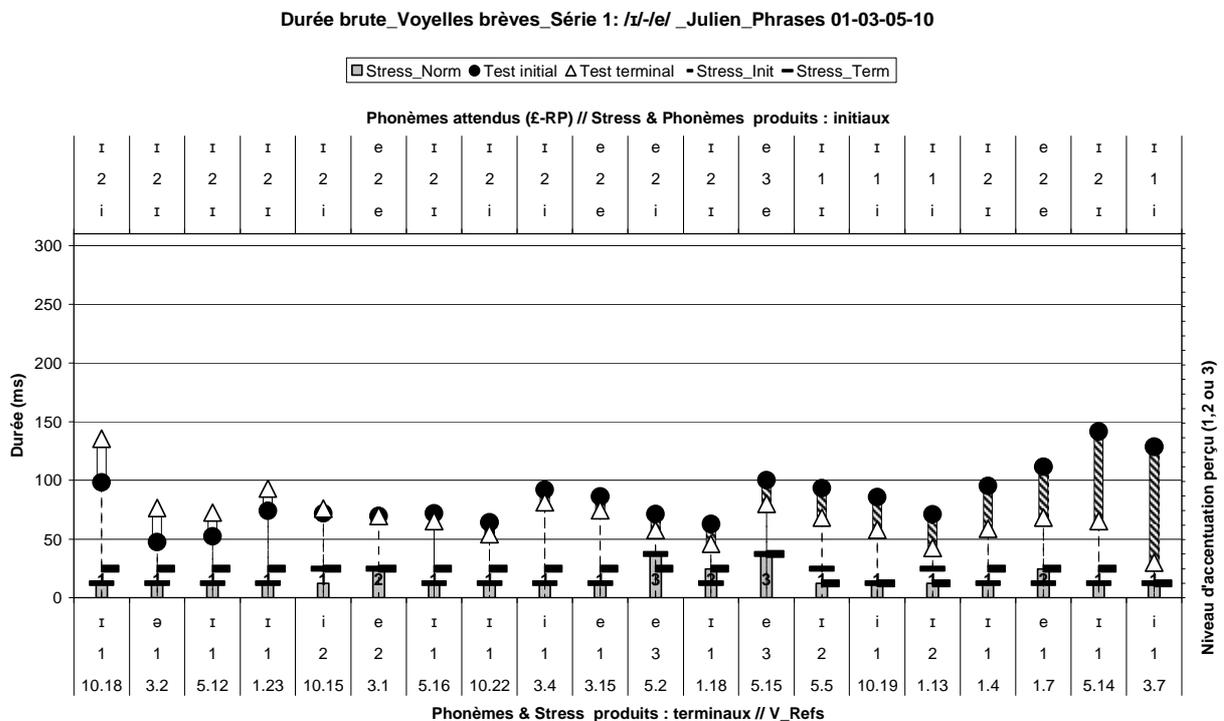


Figure 46. Durée brute des voyelles brèves /ɪ/ et /e/ en  $T_{init}$  et  $T_{term}$ . Phrases 01-03-05-10.

Durée brute\_Voyelles brèves\_Série 2 : /æ/-/ɒ/-/ʌ/-/ʊ/ \_Julien\_Phrases 01-03-05-10

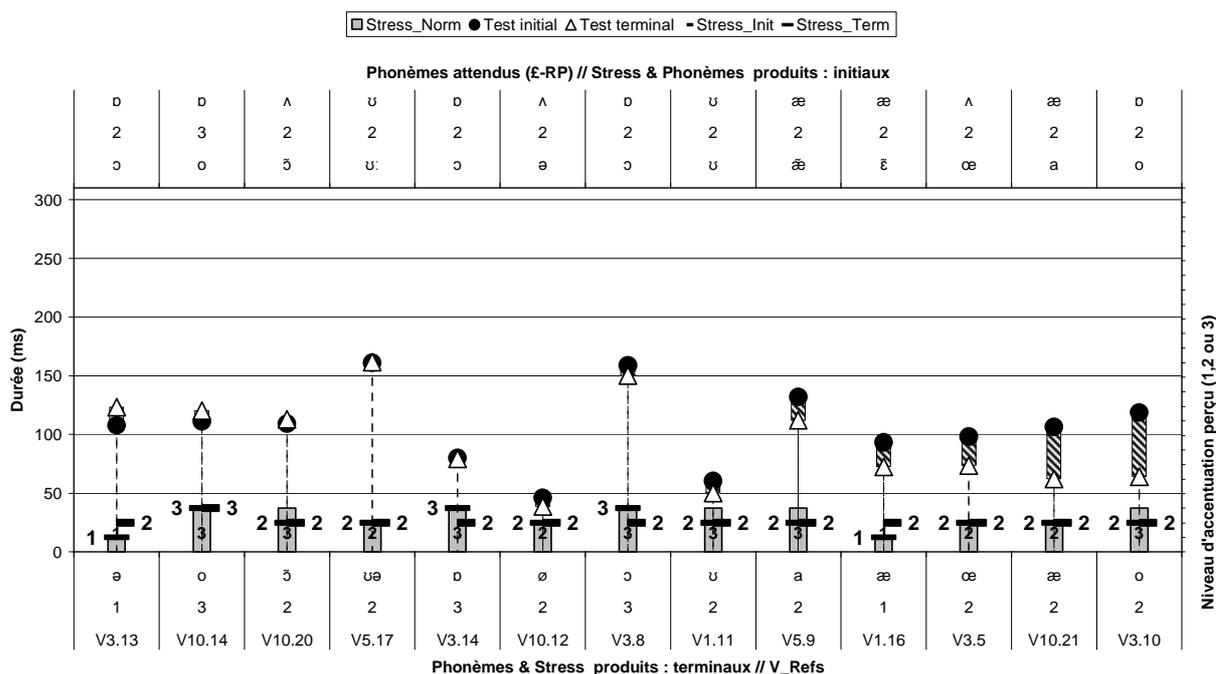


Figure 47. Durée brute des voyelles brèves /æ/-/ɒ/-/ʌ/-/ʊ/ en  $T_{init}$  et  $T_{term}$ . Phrases 01-03-05-10.

Il ressort de la lecture de ces deux graphiques que Jean, non seulement maintient une majorité des ses voyelles dans le bas du tableau, ce qui correspond effectivement à une durée brève sur l'axe des ordonnées, mais que simultanément il abrège une majorité des  $V_{short}$  davantage en  $T_{term}$  qu'en  $T_{init}$  ; et ce, indépendamment du niveau de *stress*, ce qui est l'un des but recherché, comme cela a été vu ci-dessus. Il y a bien un progrès dans la génération du contraste accentuel ici ; du moins, comme pour les  $V_{long}$ , pour la durée, mais dans le sens de l'abrègement et non de l'allongement.

Enfin, terminons maintenant sur deux graphiques dédiés aux voyelles faibles  $V_{weak}$ , en figures 48 et 49. Le premier représente la distribution de la hauteur mélodique selon le système de représentation commun à tous les graphiques de ce chapitre.

Moyennes de F<sub>0</sub> Voyelles faibles /i/-/ə/-/u/ \_Julien\_Phrases 01-03-05-10

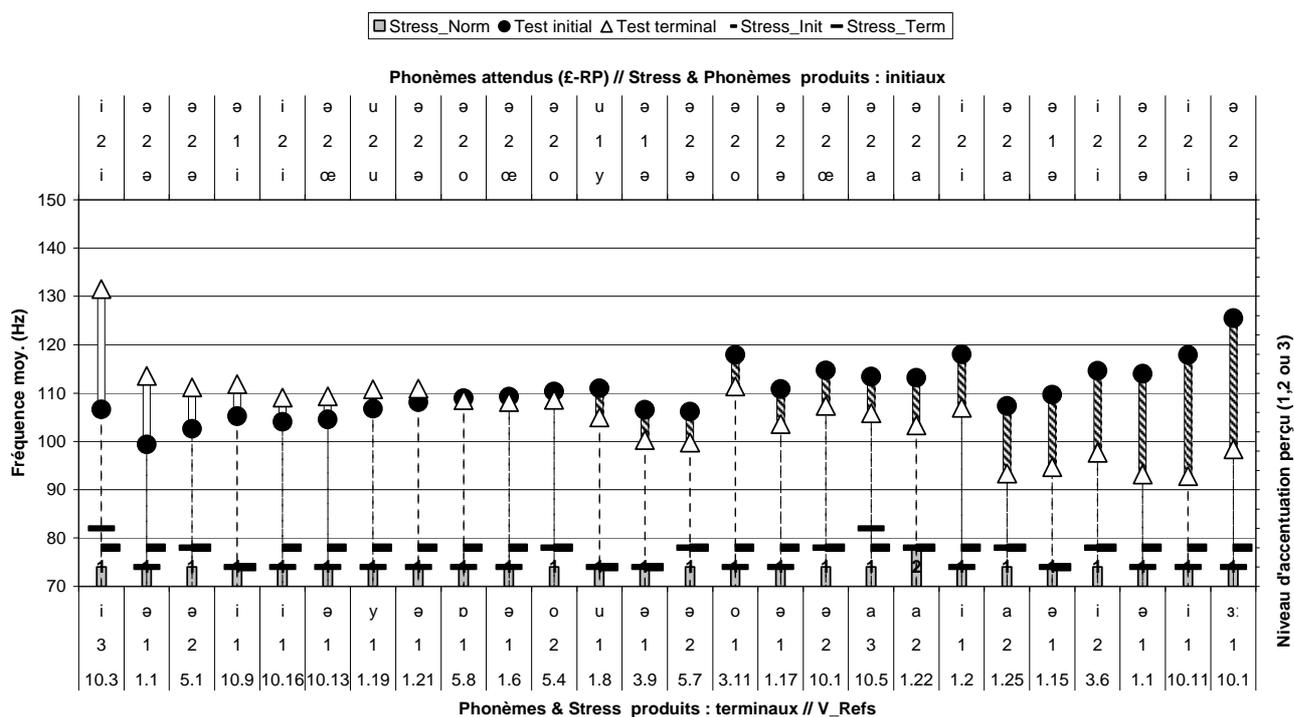


Figure 48. F<sub>0</sub> des voyelles faibles /i/-/ə/-/u/ en T<sub>init</sub> et T<sub>term</sub>. Phrases 01-03-05-10.

Le second graphique, en figure 49, traite de l'intensité. A des degrés divers et sur des échelles différentes, ces deux graphiques présentent le point commun d'avoir une répartition de l'équilibre horizontal en faveur de la partie droite du graphique, qui représente environ les deux tiers des voyelles V<sub>weak</sub>. Cela signifie que pour les deux paramètres de fréquence et d'intensité cette fois-ci, une similitude existe ici : Jean les a réduits dans la version terminale du test, ce qui peut être interprété, ici aussi, comme un progrès du contraste accentuel (par le bas), ce qui va dans la bonne direction.

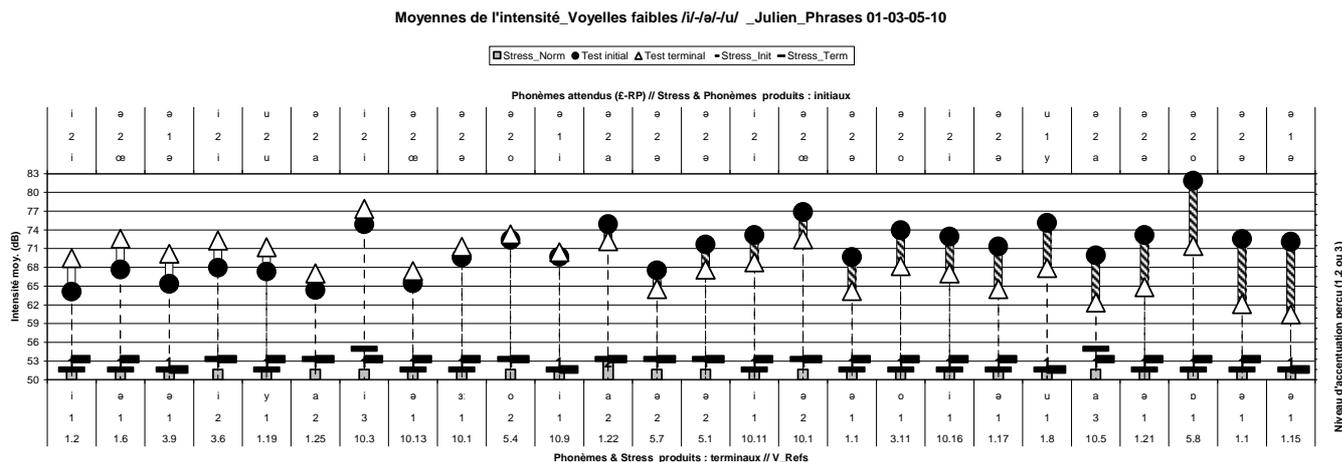


Figure 49. Intensité des voyelles faibles /i/-/ə/-/u/ en  $T_{init}$  et  $T_{term}$ . Phrases 01-03-05-10.

### ***Conclusion sur l'analyse comparative des voyelles de 4 phrases du test***

L'ensemble des graphiques décrits dans ce chapitre illustre la complexité de la tâche à accomplir à laquelle Jean, notre sujet d'étude, doit faire face lorsqu'il s'exprime en anglais L2 ; et cette partie ne concerne qu'un aspect de son expression orale, les voyelles. En abordant la variété de leurs formes, de leur traitement en fonction de leur position et de leur rôle dans la phrase, et de leur variation sous les effets du *stress*, nous espérons avoir montré (et peut-être démontré, si cela était nécessaire) que la question de la prise de parole en anglais pour un locuteur francophone était loin d'être chose facile. Si l'on se place dans une perspective de communication avec des locuteurs natifs, ce ne sera qu'à partir d'un certain niveau de prise en considération de tous les paramètres que nous avons décrits que cette communication pourra se faire avec un certain degré d'efficacité.

Dans bon nombre de nos analyses, nous avons vu que Jean avait progressé vers une meilleure prise en compte d'un ou de plusieurs des paramètres principaux composant le *stress*. Nous avons également constaté qu'il n'arrivait que rarement à intégrer ces paramètres simultanément, pour un *stress* plus efficace. Il lui arrive même aussi fréquemment de régresser (temporairement, souhaitons-le) dans l'un des paramètres alors qu'il progresse dans un autre, ce que nous avons parfaitement admis comme « normal » quand les ressources cognitives sont mobilisées par une tâche précise.

Tous ces résultats seront intégrés dans nos approches didactiques et nos pratiques pédagogiques à venir.

## Chapitre 4. — Analyse comparative de la composition spectrale de 4 voyelles longues dans les 12 phrases du test en versions $T_{\text{init}}$ et $T_{\text{term}}$

### *Introduction*

Pour achever cette partie de notre étude, nous proposons dans ce dernier chapitre de procéder à une comparaison de la qualité vocalique des productions de Jean dans les versions initiale et terminale du test. Nous pouvant procéder à l'intégralité de l'analyse des phonèmes produits par Jean dans le cadre imparti, nous proposons de nous limiter à l'analyse formantique des phonèmes qui sont probablement les plus influencés par l'application d'un stress fort en anglais : les voyelles dites « longues ». Ainsi que nous avons déjà eu l'occasion de le rappeler, ces voyelles sont traditionnellement au nombre de cinq en anglais RP<sup>36</sup>, soient les sons /i:/, /ɑ:/, /ɔ:/, /u:/ et /ɜ:/.

Nous nous limiterons arbitrairement ici à l'étude de quatre sons  $V_{\text{long}}$  attendus : /i:/, /ɑ:/, /ɔ:/ et /u:/. En effet, la position très centrale du son /ɜ:/ en RP dans le trapèze vocalique de l'API rend son étude plus délicate et ne permettra probablement pas de mettre en relief les variations des productions de notre sujet autant que les quatre autres sons, a priori situés aux quatre extrémités du trapèze. De par leur durée et leur relative stabilité théorique, ces voyelles sont les plus à mêmes de nous permettre d'obtenir des données relativement cohérentes et fiables. Nous présentons les productions de Jean ci-après dans les versions  $T_{\text{init}}$  et  $T_{\text{term}}$  de notre test.

### *Formants $F_1$ et $F_2$*

Dans les graphiques illustrant ce chapitre, nous ne ferons apparaître que les deux premiers formants  $F_1$  et  $F_2$ , correspondant aux zones d'harmoniques renforcées dans les fréquences basses et moyennes. Malgré tout l'intérêt que présenterait une analyse spectrale approfondie incluant  $F_3$ , notre but est avant tout d'analyser les conséquences de l'application, ou pas, d'une accentuation syllabique de type *stress* par notre sujet, sur la hauteur mélodique  $F_0$ , sur l'intensité, sur la durée et, dans ce chapitre, sur la qualité vocalique. Nous considérerons ici comme suffisant cette première ébauche d'analyse spectrale. Une étude plus approfondie présenterait un intérêt certain.

---

<sup>36</sup> Received Pronunciation. Cf. nos **Notes liminaires/Variétés de l'anglais**, en introduction de cette étude.

Nous présentons ci-après cinq graphiques (X, Y) selon les critères communs décrits ci-après. Les échelles des axes ont été inversées de façon à pouvoir représenter nos résultats selon la représentation traditionnelle du trapèze de l'API.

- Axe des abscisses (X) : formant  $F_2$ , en Hertz
- Axe des ordonnées (Y) : formant  $F_1$ , en Hertz
- Valeurs extrêmes pour  $F_2$ : 500 Hz à 2500 Hz
- Valeurs extrêmes pour  $F_1$  : 200 Hz à 750 Hz

NB : les graphiques utilisent des échelles différentes en  $F_1$  et en  $F_2$  selon les sons représentés, de manière à permettre une meilleure lisibilité des superpositions de données ; ils ne sont donc pas directement comparables entre eux en l'état. A fin de comparaison à échelle égale pour les quatre sons retenus, le graphique de la figure 54, en fin de chapitre, synthétise l'ensemble des données traitées sous la forme de trapèzes vocaliques, pour les différentes séries de données utilisées.

Concernant les productions de Jean, des références du type **i:2-02Eb** ou **i:2-02Lb** sont utilisées, par exemple.

- **i:2** – le son /i:/ dans sa 2<sup>ème</sup> occurrence dans le test
- **02** – le son se trouve dans la phrase Ph02
- **E** – version initiale de la phrase (*Early*)
- **L** – version terminale de la même phrase (*Late*)
- **b** – 2<sup>ème</sup> occurrence dans la phrase Ph02

### ***Utilisation de données de référence à fin de comparaison***

Quatre séries de données de référence ( $F_1$ ,  $F_2$ ) seront utilisées à fin de comparaison avec celles recueillies pour Jean, notre sujet, à l'aide du logiciel PRAAT. Trois d'entre elles sont issues d'études différentes, portant sur des locuteurs d'origine britannique tous différents et réalisées à des dates distinctes, mais relativement proches (Deterding, 1997, Hawkins et Midgley, 2001<sup>37</sup>, Gonet et Rózańska, 2003) ; leur point commun et leur intérêt résident dans le fait que leurs auteurs s'intéressent tous à la production de l'anglais

---

<sup>37</sup> NB : nous indiquons la date à laquelle l'étude a été réalisée (2001). Cette référence apparaît dans notre bibliographie à la date de publication de 2005

britannique RP par des adultes et/ou à son évolution diachronique et en fonction de l'âge des locuteurs. Dans leur étude, Hawkins et Midgley ont étudié les formants des monophthongues RP de 4 groupes de locuteurs d'âge différent, de 20 à 73 ans, dont la tranche 20-25 ans, ce qui nous intéresse tout particulièrement, puisque cela correspond aux âges des locuteurs de notre **Corpus Ense3 2012**<sup>38</sup> ; ces sont ces valeurs que nous avons retenues. Les valeurs (moyennées) issues de cette étude apparaissent sous la forme, par exemple : **i:-rp-2001\_20-25**, soit la valeur RP mesurée de /i:/, issue de cette étude de 2001, portant sur la tranche d'âge « 20 à 25 ans ».

Les deux autres études portent sur des locuteurs hommes et femmes issus de la BBC<sup>39</sup> (Deterding, 1997, Gonet et Róžańska, 2003). À chaque fois, nous n'avons retenu que les données issues de locuteurs masculins, par cohérence avec le locuteur masculin de notre étude de cas. Pour l'ensemble des données, nous avons utilisé les valeurs moyennes des formants, issues de séries de mots différents. Ainsi, par exemple **i:-rp-1997** correspondra à la valeur moyenne RP des données issues de l'étude de Deterding. Nous n'utiliserons ces données prudemment et uniquement à titre de tendances, sous forme de moyennes. Il est important de noter qu'elles portent sur des petits nombres de cas, respectivement six et deux.

La quatrième série de référence correspond à la production des 12 phrases du test par le professeur de Jean (T<sub>ref</sub>). Pour cette série de données, nous avons extrait les 6 occurrences des voyelles retenues et les avons comparées une à une dans nos graphiques (ex. **i:2-02\_Ref** correspond à la 2<sup>ème</sup> occurrence chronologique de /i:/ se trouvant dans la phrase Ph02). Nous avons également fait figurer la valeur moyenne de ces occurrences (ex. **i:-Ref-moy**).

### ***Mesures de $F_1$ et $F_2$ de 4 monophthongues $V_{long}$ de Jean et comparaison aux données de référence***

Dans les graphiques qui suivent, nous présentons nos résultats dans l'ordre /i:/, /ɑ:/, /ɔ:/, /u:/, puis nous terminerons par un graphique de synthèse de nos mesures. Nous commençons par la figure 50 représentant les valeurs mesurées pour la voyelle /i:/.

---

<sup>38</sup> Cf. Partie III : construction du corpus Ense3 2012

<sup>39</sup> *British Broadcasting Corporation*. Voir notre partie **Notes liminaires / Variétés de l'anglais** en début de ce mémoire.

## Mesures de $F_1$ et $F_2$ pour le son /i:/

Le graphique de la figure 50 montre la variation en  $F_1$  et  $F_2$  des 6 occurrences attendues du son /i:/ parmi les 12 phrases du test.

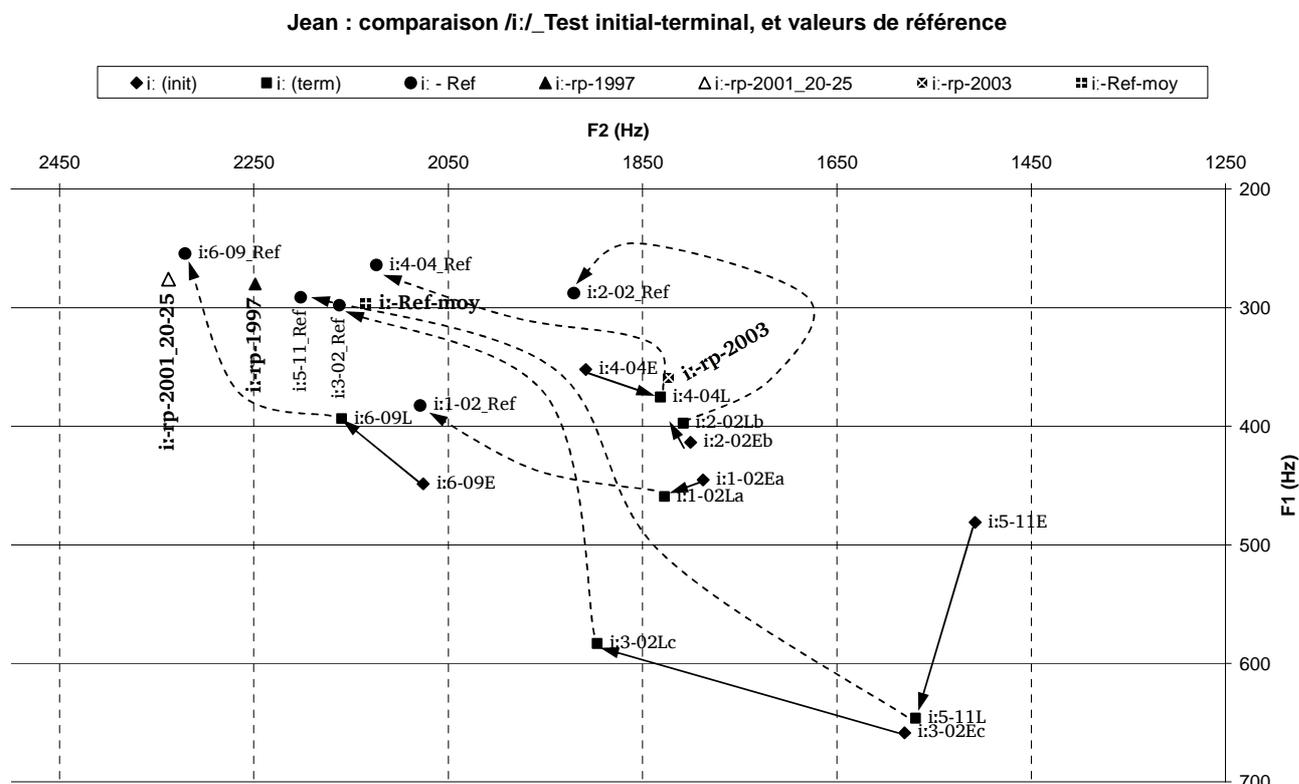


Figure 50. Valeurs mesurées de  $F_1$  et  $F_2$  du son /i:/ en 6 occurrences parmi les phrases Ph01 à Ph12, dans les deux versions du test  $T_{init}$  et  $T_{term}$ . Également : les valeurs moyennes de  $F_1$  et  $F_2$  des trois séries RP (**rp-1997** ; **rp-2001-20-25** ; **rp-2003**) et la moyenne des données de  $T_{ref}$  (**Ref-moy**). Les flèches pleines indiquent la variation entre  $T_{init}$  (◆) et  $T_{term}$  (■). Les flèches en pointillé indiquent la distance entre  $T_{term}$  (■) et  $T_{ref}$  (●) pour chaque occurrence. Légende identique dans les graphiques suivants.

Voici les quatre phrases contenant théoriquement ce son ; les syllabes retenues pour les besoins l'étude ont été grisées. Du point de vue morphologique, le son apparaît dans 3 adjectifs (stress de niveau St2-épithète ou St3-isolé, ex. Ph11, attendus) et 3 substantifs (St3 attendus).

Ph02	A recent survey has shown that more and more female teenagers opt for scientific studies.
Ph04	Each new class of students is surveyed in terms of academic achievements.
Ph09	This graph shows how our students' origin and age are related, in one piece of information.
Ph11	On average, we lose three per cent of boys per annum — that's true!

La figure 50 fait apparaître un regroupement assez net des prononciations RP de 1997 et 2001, ainsi que des phrases issues de  $T_{ref}$ , concentrées dans la zone de  $F_1$  basse et de  $F_2$  élevée (coin supérieur gauche). Notons la position un peu plus centrale de RP-2003, susceptible de révéler une évolution moderne de la prononciation de ces sons vocaliques vers une certaine centralisation.<sup>40</sup>

Concernant les productions de Jean, nous constatons deux phénomènes principaux : a) l'étalement de l'ensemble des données sur une zone relativement large du graphique et plutôt éloignée des valeurs de référence, toutes les données étant supérieures en  $F_1$  à 350 Hz ; b) des variations entre  $T_{init}$  et  $T_{term}$  de natures très différentes et souvent contradictoires, ne permettant pas de mettre en valeur de véritables tendances fortes.

Notons toutefois :

- qu'une partie des productions de Jean se trouve assez proche de RP-2003 : les items 1, 2 et 4, qui ne varient d'ailleurs guère entre  $T_{init}$  et  $T_{term}$
- les items 1, 3 et 6 varient en direction des productions de référence, tout en restant très éloignés

Nous relèverons a contrario que la valeur RP-2011 des 20-25 ans est la plus éloignée de toutes des productions de Jean.

#### **Mesures de $F_1$ et $F_2$ pour le son /ɑ:/**

La figure 51 reprend les mêmes éléments que le graphique précédent, mais pour le son /ɑ:/, avec des échelles  $F_1$  et  $F_2$  adaptées, pour une plus grande lisibilité. Voici les 6 occurrences du son, réparties dans 6 phrases différentes : 4 substantifs et 2 adjectifs, tous en position de stress de type St3 attendus.

---

<sup>40</sup> Tendence d'ailleurs soulignée dans les conclusions de l'étude RP-2003 (Gonet et Rózańska 2003)

Ph01	A reform in higher education took place this autumn and the situation has improved, at last.
Ph04	Each new class of students is surveyed in terms of academic achievements.
Ph08	At the start of the year, the directors project an information slide-show.
Ph09	This graph shows how our students' origin and age are related, in one piece of information.
Ph10	The year before had been quite ordinary, but a drop is recorded this month, and it's sharp!
Ph12	On that chart, we can see how they choose their future place of study too.

Ici, les données de référence se concentrent toutes dans le quart supérieur droit du graphique, alors que les données correspondant aux productions de Jean s'étalent sur toute la moitié inférieure du tableau. De plus, l'intégralité des 6 voyelles varie dans des directions opposées aux valeurs de référence. Notons que les productions de Jean sont plutôt plus proches des trois valeurs RP-1997, 2011 et 2003, que des valeurs de référence du professeur, notamment en initiales.

Jean : comparaison /a:/\_Test initial-terminal, et valeurs de référence

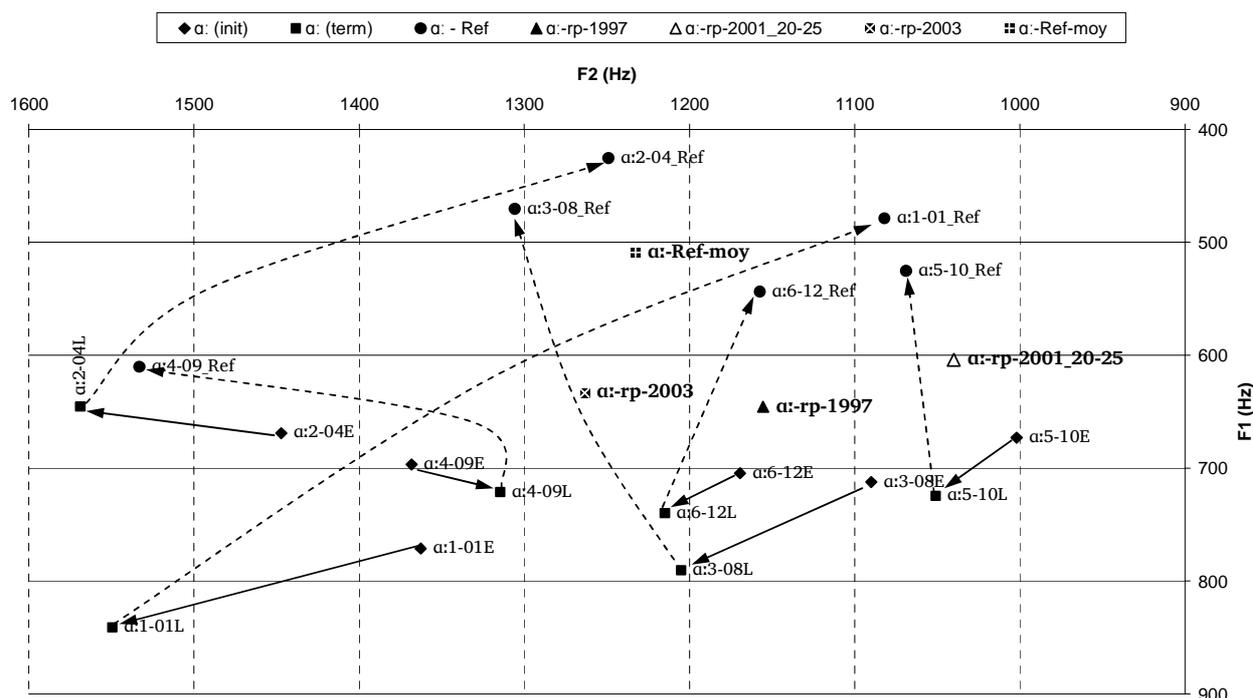


Figure 51. Valeurs mesurées de F<sub>1</sub> et F<sub>2</sub> du son /a:/ en 6 occurrences parmi les phrases Ph01 à Ph12, dans les deux versions du test T<sub>init</sub> et T<sub>term</sub>, les trois séries RP (rp-1997 ; rp-2001-20-25 ; rp-2003) et la moyenne des données de T<sub>ref</sub> (Ref-moy). Légendes : idem graphique précédent.

La figure 51 montre clairement une régression dans la qualité de production du son /a:/ par Jean : la distance séparant ses productions des valeurs de référence augmente dans tous les cas, allant parfois vers des variations assez importantes, en F<sub>2</sub> notamment.

### Mesures de F<sub>1</sub> et F<sub>2</sub> pour le son /ɔ:/

La figure 52 présente la même configuration que les précédentes, mais pour le son /ɔ:/. Voici les phrases retenues pour ce son.

Ph01	A reform in higher education took place this autumn and the situation has improved, at last.
Ph02	A recent survey has shown that more and more female teenagers opt for scientific studies.
Ph10	The year before had been quite ordinary, but a drop is recorded this month, and it's sharp!

Nous trouvons un peu plus de variété dans les morphèmes de ces phrases ici : substantifs, verbe, adjectif, déterminant du nom, marqueur temporel. Ils sont tous susceptibles de recevoir un accent de type St3.

Jean : comparaison /ɔ:/\_Test initial-terminal, et valeurs de référence

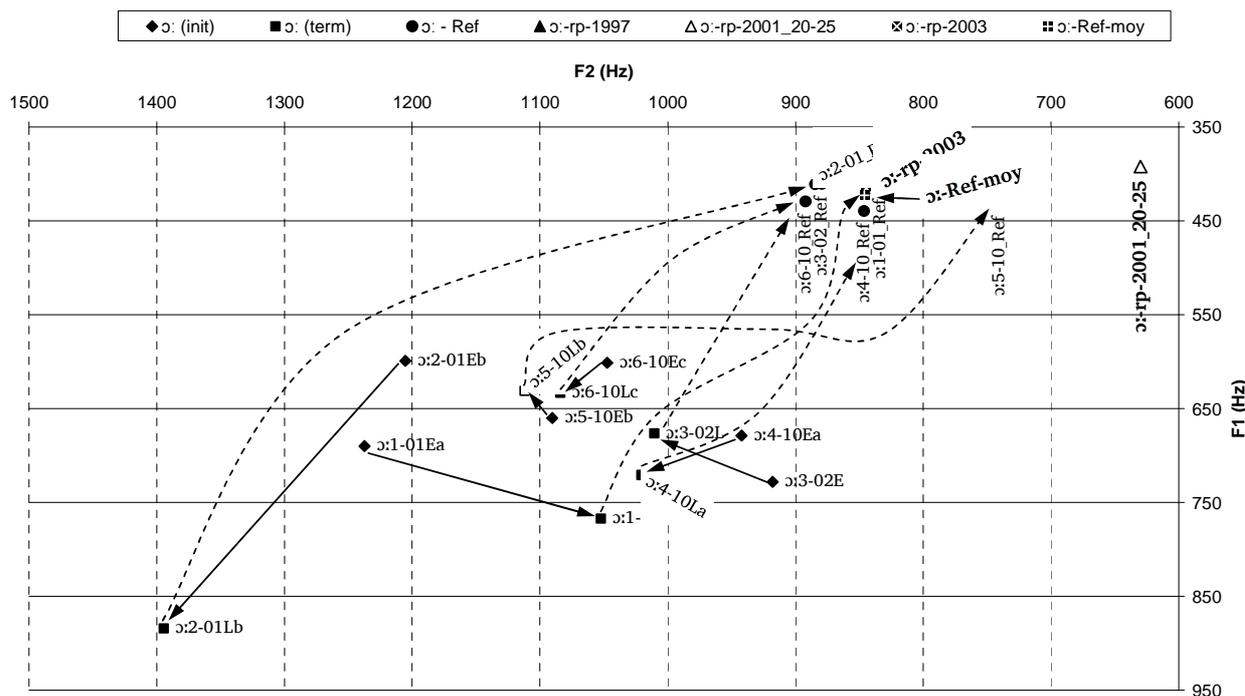


Figure 52. Valeurs mesurées de F<sub>1</sub> et F<sub>2</sub> du son /ɔ:/ en 6 occurrences parmi les phrases Ph01 à Ph12, dans les deux versions du test T<sub>init</sub> et T<sub>term</sub>, les trois séries RP (rp-1997 ; rp-2001-20-25 ; rp-2003) et la moyenne des données de T<sub>ref</sub> (Ref-moy). Légendes : idem graphiques précédents.

Ce graphique reprend la distribution en deux zones relativement distinctes des données, mais de façon plus resserrée. Les valeurs de référence sont toutes concentrées dans une zone très étroite. Notons le cas de la valeur RP-2001 qui, encore une fois, se retrouve à l'extrême opposé des productions de Jean. Ces dernières valeurs sont

relativement regroupées dans la partie centrale, hormis deux d'entre elles (items 1 et 2) qui subissent des variations importantes et n'allant pas dans le sens d'une progression vers les valeurs de référence.

#### Mesures de $F_1$ et $F_2$ pour le son /u:/

La figure 53 présente les mesures de  $F_1$  et  $F_2$  pour le son /u:/. Voici les phrases du test contenant ce son.

Ph01	A reform in higher education took place this autumn and the situation has improved, at last.
Ph05	The record of employment of graduates coming through is very good.
Ph11	On average, we lose three per cent of boys per annum — that's true!
Ph12	On that chart, we can see how they choose their future place of study too.

Nous retrouvons dans ce graphique la division nette entre les productions de Jean et les valeurs de référence, ces dernières étant toutes cantonnées cette fois ci en haut du graphique, avec une variation très importante de leur  $F_2$ , mais minime de leur  $F_1$ . La distance les séparant des valeurs de Jean est toujours importante.

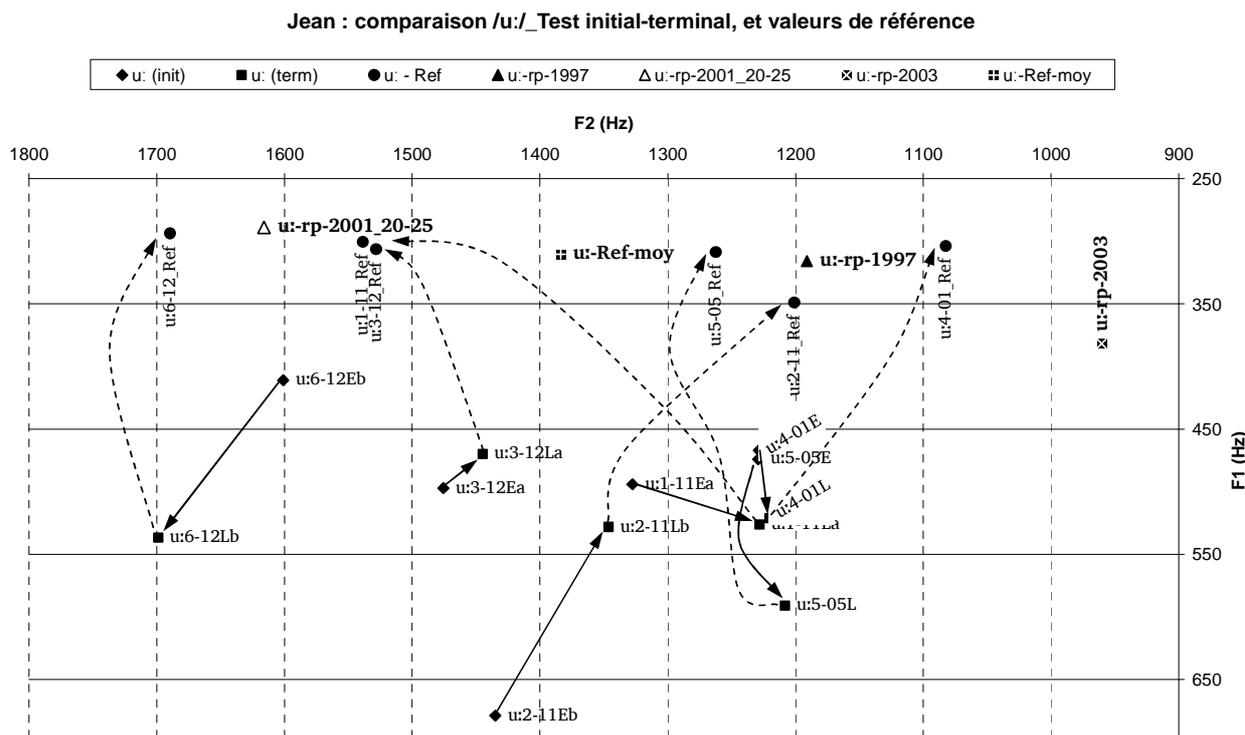


Figure 53. Valeurs mesurées de  $F_1$  et  $F_2$  du son /u:/ en 6 occurrences parmi les phrases Ph01 à Ph12, dans les deux versions du test  $T_{init}$  et  $T_{term}$ , les trois séries RP (**rp-1997** ; **rp-2001-20-25** ; **rp-2003**) et la moyenne des données de  $T_{ref}$  (**Ref-moy**). Légendes : idem graphiques précédents.

Encore une fois, les variations des valeurs des productions de Jean sont plutôt contradictoires entre elles et ne vont que rarement dans le sens d'un net « progrès », sauf pour l'item 2b.

### ***Conclusion sur les mesures de la composition spectrale de 4 monophthongues de type « longues »***

De toute évidence, les données reportées sur les quatre graphiques précédents ne montrent aucunement que notre locuteur progresse dans la qualité de ses productions vocaliques, du moins pour les quatre voyelles « longues » considérées ; nous avons même constaté en plusieurs occasions, une variation très aléatoire, voire une nette régression. Le dernier graphique, présenté ici en figure 54, propose une synthèse de nos mesures des productions de Jean et du report des différentes séries de données de référence choisies à titre de comparaison, sous la forme de trapèzes vocaliques, inspirée de la représentation traditionnelle des voyelles de l'API.

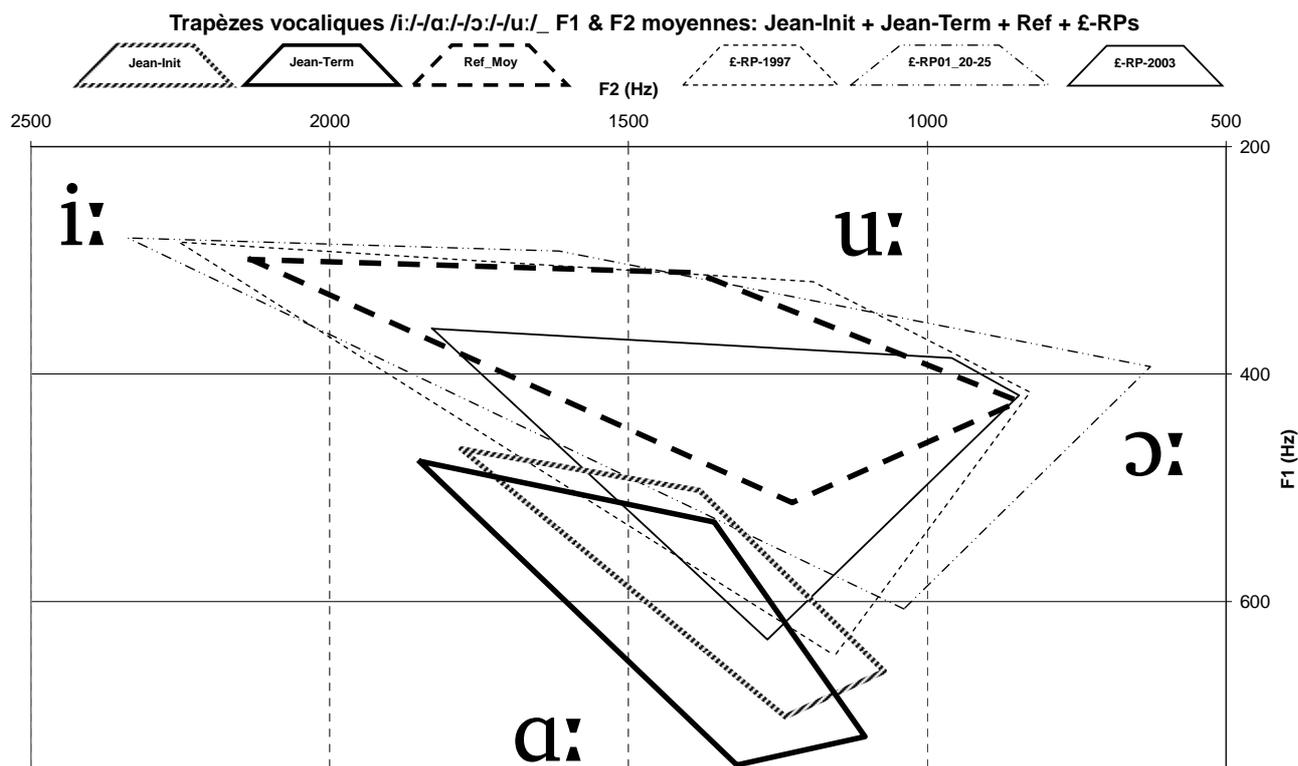


Figure 54. Synthèse des trapèzes vocaliques de type API pour /i:/-/a:/-/ɔ:/-/u:/ en valeurs moyennées de Jean (phrases Ph01 à Ph12) en  $T_{init}$  et  $T_{term}$  et de nos quatre séries de données de référence.

Nous voyons très clairement apparaître ici la distance qui sépare les productions initiales et terminales de Jean (en valeurs moyennées) de nos productions de référence. Nous noterons également l'éloignement de la version  $T_{term}$  du test, à la fois de  $T_{init}$  et, surtout, des valeurs de référence.

Pour ce qui concerne les valeurs de référence que nous avons utilisées, nous trouverons une certaine cohérence de forme et de taille entre trois d'entre elles. La dernière étude en date que nous ayons utilisé, celle de 2003, tendrait à montrer une tendance nette à la centralisation dans la prononciation des sons /i:/ et /u:/, décrite par les auteurs de l'étude (Gonet et Rózańska, 2003). Il faut toutefois nuancer ces propos, s'agissant d'une étude ne portant que sur un petit nombre de cas (deux). Enfin, notons que les valeurs de référence de la tranche des 20-25 ans se trouvent presque toujours à l'extrême opposé des valeurs des productions de Jean.

Ainsi, les résultats que nous avons obtenus dans cette partie de notre étude montrent très clairement que, si Jean a progressé dans certains des domaines évoqués dans les chapitres précédents, ce n'est pas du tout le cas dans la qualité vocalique de ses productions, en tout cas pour les formants fondamentaux que sont  $F_1$  et  $F_2$  des  $V_{long}$

considérées. La zone d'intersection entre nos trapèzes de référence et ceux de Jean, dans les deux versions du test, tend à diminuer. Toutefois, rappelons que la production de phonèmes d'une qualité formantique se rapprochant de celle de locuteurs natifs ne faisait pas partie de nos objectifs initiaux. C'est avant tout au rythme, aux accents syllabiques de type *stress* et à leurs conséquences directes en terme de durée, de hauteur mélodique et d'intensité que nous avons accordé toute notre attention, car de notre point de vue ils devraient constituer le point de départ de l'apprentissage de la langue anglaise. Ceci est encore très souvent ignoré ou oublié dans l'enseignement de l'anglais L2, même si c'est de moins en moins le cas. Nous sommes convaincu que les jeunes adultes, étudiants sur le seuil de la vie professionnelle, peuvent encore progresser de façon importante et rapidement dans leur compétence en expression orale, dans leur interaction avec des locuteurs natifs ou non natifs.

La qualité formantique des phonèmes est certes importante à maîtriser, mais elle n'intervient qu'après les caractéristiques rythmiques pour ce qui concerne l'efficacité dans l'encodage ou le décodage d'une information à l'oral. Enfin, la « régression » observée chez notre sujet dans ce domaine peut aussi s'expliquer, comme nous l'avons déjà remarqué, par la mobilisation de ses capacités cognitives pour réaliser la tâche rythmique proposée. Restons plutôt confiants dans ses capacités à inverser la tendance avec la pratique et l'acquisition de réflexes langagiers.

## Conclusion

Tout au long des tâches d'analyse réalisées sur les productions de Jean, nous nous sommes posé la question si, suite à l'utilisation du métronome en cours, il avait toujours une influence une fois enlevé ; si l'étudiant avait intégré et réutilisé à bon escient cette valeur fondamentale qu'est le *stress* en anglais, ce qui est le but recherché.

Nous avons vu, que pour certains paramètres tels que la durée vocalique, la hauteur mélodique, nous pouvions mettre en évidence un progrès certain, même s'il restait une marge de progression non-négligeable.

Pour d'autres paramètres, comme l'intensité ou la qualité vocalique, les résultats sont apparus plus mitigés, voire même allant vers une régression du sujet. Il n'est pas rare qu'un progrès dans un domaine s'accompagne temporairement d'une régression dans un autre, tant la demande sur les capacités cognitives peut être forte en expression orale dans une langue étrangère. Il conviendra d'analyser plus précisément la nature de ces stagnations/régressions et essayer de développer des approches permettant d'accomplir des progrès en matière de rythme sans engendrer de pertes dans d'autres domaines.

Le travail présenté dans ce mémoire nous aura permis de mettre en œuvre un certain nombre d'outils, utilisés dans différentes configurations. À travers cette étude de cas, et au-delà de l'analyse des résultats de Jean, notre sujet, nous avons cherché à déterminer lesquels de ces outils seraient les plus pertinents pour une analyse de plus grande envergure.

Les tâches de segmentation manuelle du signal de parole, au demeurant intéressantes, peuvent rapidement s'avérer longues et fastidieuses. Un traitement automatique de la segmentation et du recueil des données serait indispensable si nous voulions traiter, à grande échelle, tout ou partie des données de la production des soixante quinze étudiants du corpus. Dans ce même ordre d'idée, l'utilisation d'outils statistiques serait indispensable.

Avec de tels résultats sur une plus grande échelle, nous pourrions observer des tendances fortes, dont nous pourrions nous inspirer pour nos pratiques pédagogiques et les ajuster en conséquence, voire obtenir des retours d'information en temps réel pour pouvoir progresser plus efficacement. Des outils existent depuis longtemps déjà, mais ils offrent en

général une vision synthétique, globale du signal de parole. Le fait de pouvoir isoler un paramètre précis, ainsi que nous l'avons fait, est sans aucun doute un gage d'efficacité accrue.

Parmi ce que nous retiendrons, nous noterons également la mise en relief d'interférences entre la L1 (français) et la L2 (anglais) de Jean, mais également entre sa L2 et sa L3 (espagnol). Une connaissance plus approfondie des caractéristiques rythmiques et acoustiques de chacune des langues, tant chez les apprenants que chez les enseignants, permettrait sans aucun doute de gagner en efficacité.

Ainsi, nous proposons de valider la méthodologie d'observation mise en œuvre, tant pour ce qui concerne le mode expérimental, que l'analyse des données. Bien entendu, des améliorations restent à apporter, par exemple dans nos essais de représentation du *stress*. Si aucune correction phonétique n'a été constatée à ce stade, nous restons confiants dans l'utilisation de cette approche rythmique guidée. En ayant recours à des stratégies complémentaires spécifiques portant sur l'amélioration de la qualité phonémique, celles-ci devraient venir compléter l'amélioration de la dynamique rythmique et intonative constatée. Cela restera à vérifier sur tout ou partie des soixante-quinze élèves de notre corpus et à comparer avec le groupe témoin.

La référence à la production de l'enseignant ou d'un locuteur anglophone pourra éventuellement être gardée (en production post-métronomie), avec les précautions d'usage, car elle peut donner une idée de la distance, de la marge de progrès restante pour les étudiants.

Pour ce qui concerne le métronome, il sera intéressant d'observer l'ajustement pour chaque étudiant, selon son évolution personnelle. L'utilisation d'une sorte de « boîte-à-rythme » plutôt qu'un simple métronome, pourrait permettre un apprentissage plus fin, permettant de différencier le rôle de l'accent primaire St3 de l'accent secondaire St2. Des essais sur la variation du tempo pourraient permettre de renforcer davantage le contraste phonétique vers la réduction, voire la neutralisation vocalique, naturelle chez les locuteurs natifs.

Enfin, d'autres défis s'offrent tout naturellement. Tout au long de cette étude, nous avons utilisé l'anglais britannique « théorique » RP comme référence. En ce début de

21<sup>ème</sup> siècle, et pour la première fois, le nombre de locuteurs utilisant couramment l'anglais « langue seconde » a dépassé celui des locuteurs natifs (Jenkins, 2001) : quelles variétés d'anglais doivent être enseignées aujourd'hui ? Les modèles RP ou GA vont-ils disparaître ou évoluer vers un modèle d'anglais « international », ou de plusieurs ? L'accentuation de type *stress* va-t-elle conserver son rôle sémantique déterminant ? Ces questions sont tout à fait pertinentes. Il sera intéressant de comparer les productions de nos étudiants à de nouvelles références, reflétant probablement davantage la réalité en évolution constante et rapide du monde anglophone.

## Bibliographie

- Bertinetto, P. M. (1989). Reflections on the Dichotomy 'Stress' vs. 'Syllable-timing'. *Revue de Phonétique Appliquée*, 99-130. Mons, Belgique.
- Calvet, Louis-Jean (1979). *Langue, corps & société*. Paris : Payot
- Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics*, Vol. 11: 51-62. Amsterdam : Elsevier
- Dauer, Rebecca M. (1987). Phonetic and phonological components of language rhythm. *Proceedings of the 11th international congress of phonetic sciences*, 447-450. Tallinn, Estonia.
- Delattre, Pierre (1966). Les Dix Intonations de base du français. *The French Review* 40/1: 1-14. Carbondale, USA.
- Deterding, D. (1997). The Formants of Monophthong Vowels in Standard Southern British English. *Journal of the International Phonetic Association*, 27, 47-55. International Phonetic Association : United Kingdom
- Fónagy, I. (1983). *La vive voix*. Paris: Payot.
- Fraisse, P. (1974). *Psychologie du rythme*. Paris: Presses Universitaires de France.
- Gimson, A. C. (2008) & Cruttenden, A. (Ed.). *Gimson's Pronunciation of English*. USA : Oxford University Press.
- Gonet, W. & Róžańska, K. (2003). Vowel quality in the idiolects of four BBC World News presenters. *Speech and Language Technology. Volume 7*. UMCS : Pologne
- Hawkins S. & Midgley J (2005). Formant frequencies of RP monophthongs in four age groups of speakers. *Journal of the International Phonetic Association* 35/2. International Phonetic Association : United Kingdom
- Jenkins, J. (2001), *The Phonology Of English As An International Language*. Oxford : OUP
- Mairano, P. (2010). *Rhythm Typology : Acoustic and Perceptive Studies*. Thèse de doctorat, Université de Turin, Italie.
- Scotto Di Carlo, N. (2007), Speech-rate and tempo effects of the organization of spoken and sung syllables. *XVIIth International Congress of Phonetic Sciences*, Saarbrücken, Allemagne. Récupéré le 20 août 2012 à <http://aune.lpl.univ-aix.fr/lpl/personnel/scotto/scottopub.htm>
- Vaissière, J. (2007), *La phonétique. Que sais-je ?* Paris : PUF
- Walter, Henriette (2003). *Honni soit qui mal y pense : l'incroyable histoire d'amour entre le français et l'anglais*. Paris : Robert Laffont
- Wells, J. C. (2008). *Pronunciation Dictionary*. Harlow: Longman.

## **Table des annexes**

Annexe 1 Questionnaire employé dans la construction du corpus Ense3 2012 .....	131
Annexe 2 Exemple de traitements de type <i>script</i> employés sous PRAAT.....	134
Annexe 3 Exemple du formulaire de consentement de participation soumis aux étudiants .....	136

**Annexe 1**  
**Questionnaire employé dans la construction du corpus Ense3 2012**

---

**L2 English: your Language-and-Music History  
Questionnaire (Version 1.2)**

[List of questions]

---

Timestamp

Sample Question 1

Sample Question 2

1. Contact Information — Family name

2. First name

3. E-mail

4. Telephone

5. Today's date

6. Age

7. Sex

8. Education

9. Have you ever suffered/do you currently suffer from a handicap?

9.3 How have you learned English up to now: a MIXTURE of both classroom instruction AND interaction with people - or other?

10. Location - Your country of origin

11. Your country of current residence

12. Since?

13. How long have you lived in a foreign country where English is spoken?

14. Please list all the CULTURES (apart from your own) which appeal to you

15. Do you consider yourself a musician, today?

16. If YES, please detail your past and/or current musical skills and practice, instrument played, orchestral experience, public performances...

17. Do you sing? Do you play one or several musical instruments?

18. Do you perform music/singing in public?

19. Did you / Do you learn music...

20. Do you listen to recorded music or songs?

20.1 Please name your favourite singers/groups: those who sing in your native language

20.2 Please name your favourite singers/groups today: those who sing in ENGLISH

21. Do you know and understand the following musical terms: [Beat]

21. Do you know and understand the following musical terms: [Rhythm]

21. Do you know and understand the following musical terms: [Tempo]

21. Do you know and understand the following musical terms: [Pitch]

21. Do you know and understand the following musical terms: [Metronome]
22. Can you dance? Do you dance? (all types)
23. What type(s) of dance do you practice?
24. What is your native language?
25. Do you speak other languages than your native language and English? If YES, please specify language(s)
26. Did you start to learn English at home?
27. If YES, please specify age:
28. Did you start to learn English at school?
29. If YES, please specify starting age:
30. Did you start to learn English After arriving in an English-speaking country?
31. If YES, please specify the COUNTRY, DATE of ARRIVAL, LENGTH of STAY
32. How have you learned English up to now: through formal classroom instruction?
33. How have you learned English up to now: through interacting with people?
34. Other ways of learning English?
35. For English and other foreign languages that you know: rate your ability according to the scale below [English]
35. For other foreign languages that you know: rate your ability according to the scale below [Other foreign language 2]
- [Other foreign language 3]
- [Other foreign language 4]
- [Other foreign language 5]
36. For other foreign languages that you know: please specify "Other foreign language 2"
37. For other foreign languages that you know: please specify "Other foreign language 3"
38. For other foreign languages that you know: please specify "Other foreign language 4"
39. For other foreign languages that you know: please specify "Other foreign language 5"
40. Please list all the foreign languages you know in ORDER OF ACQUISITION and state STARTING AGE below
41. Please list all the CULTURES with which you identify or which appeal to you, in order of prominence.
42. Provide the age at which you were first exposed to English: LISTENING
43. Provide the age at which you were first exposed to English: SPEAKING
44. Provide the age at which you were first exposed to English: READING
45. Provide the age at which you were first exposed to English: WRITING
46. Among the languages that you speak on average per day, how often do you use your native language, in the 4 skills above combined?
47. Among the languages that you speak, how often do you use ENGLISH per day, in the 4 skills combined?
48. How often do you use other languages per day (other than your native language and English), in the 4 skills combined?
49. Are you engaged in the following activities for more than HALF AN HOUR per day on average?
  - [Listening to RADIO TALKS (includes web-based radios)]
  - [Watching TV or web-based videos with dialogues]
  - [Reading for pleasure]

- [Reading for studies/work]
- [Seeking/reading information on the Internet]
- [Writing: private (on paper or computer - includes e-mail)]
- [Writing: work/studies (on paper or computer - includes e-mail)]
- [Telephoning (either private or work/studies)]
- [Talking to people (either private or work/studies)]
- [Gaming (either live or online - only games requiring communication)]

50. Instruction language(s) (i.e. used by teachers) for each schooling level: PRIMARY/ELEMENTARY School

51. Instruction language(s) (i.e. used by teachers) for each schooling level: SECONDARY/MIDDLE School (or French "collège")

52. Instruction language(s) (i.e. used by teachers) for each schooling level: SECONDARY/HIGH School (or French "lycée")

53. Instruction language(s) (i.e. used by teachers, professors) for each schooling level: COLLEGE/UNIVERSITY

54. In which language do you usually COUNT or DO ARITHMETIC?

55. In which language do you usually DREAM?

56. In which language do you usually express ANGER / HUMOUR / AFFECTION?

57. When speaking, do you ever mix words from 2 or more languages?

58. In your own perception, how much of a foreign accent do you have when speaking in English?

59. Between your 2 best foreign languages, which one are you better at: READING?

60. Between your 2 best foreign languages, which one are you better at: WRITING?

61. Between your 2 best foreign languages, which one are you better at: SPEAKING?

62. Between your 2 best foreign languages, which one are you better at: UNDERSTANDING SPEECH?

63. Have you ever taken a proficiency test in English (e.g. BULATS, TOEIC, TOEFL, IELTS, Cambridge Tests, etc.)?

64. Please feel free to add any comment about your musical and/or linguistic background or music/language use that you feel appropriate.

**Thank you.**

— **Fin du questionnaire** —

## Annexe 2

### Exemple de traitements de type *script* employés sous PRAAT

#### Exemple de script d'extraction de la durée de voyelles

```
#####
## extrait_voyelles-durée.praat
##
## Extraction de la durée du phonème pour chaque voyelle.
## Ecriture des résultats dans un fichier texte.
##
## Le fichier d'annotations doit avoir le même nom que le fichier .wav,
avec l'extension .TextGrid
## Il doit comporter 1 seul champ de type (IntervalTier) avec les
intervalles des voyelles.
## Le symbole de la voyelle est réutilisé dans le fichier de résultats.
##
## Auteur : Nicolas Audibert, ICP, septembre 2007
## Modifié par Solange Rossato, LIG, juillet 2012
## Modifié par Francis Picavet, Grenoble INP/Université Stendhal, juillet
2012
#####

clearinfo

#####
# constantes à modifier à l'ouverture du script
#####
form extraction de durée
    word nomFichier 9_Jean_20120131_01
    word          chemin          I:\@_Univ\LMM-Research-
Prog_SOUNDFILES\Semester2_Language&Music_TEST-RECs\Early-
semester_TAKES\WAV-RECs_Semester2_Batch-EARLY-
EDITED\9_Jean_20120131\PRAAT-Wk_Julien_Early_SYLLABLES
endform
#####

# lecture des fichiers .wav et .TextGrid
Read from file... 'chemin$\'nomFichier$.wav
Read from file... 'chemin$\'nomFichier$.TextGrid

fichResultats$ = "'chemin$\RezVow_'nomFichier$.txt"
# on efface l'ancien fichier de résultats s'il existe
filedelete 'fichResultats$'
fileappend 'fichResultats$' VowelRank 'tab$' V_expected 'tab$' V_produced
'tab$' StressVowNorm 'tab$' StressVowProd 'tab$' FocusVow(0,1) 'tab$'
tDebutV(ms) 'tab$' tFinV(ms) 'tab$' DuréeV(ms) 'newline$'

select TextGrid 'nomFichier$'

numChampFocusVow = 8
numChampStressVowNorm = 9
numChampStressVowProd = 10
numChampVowelRank = 11
numChampV_expected = 12
numChampV_produced = 13
nbIntervalles = Get number of intervals... numChampVowelRank
```

```

i = 1
while i <= nbIntervalles
  select TextGrid 'nomFichier$'
  etiqInterv$ = Get label of interval... numChampVowelRank i
  etiqStressVowNorm$ = Get label of interval...
numChampStressVowNorm i
  etiqStressVowProd$ = Get label of interval...
numChampStressVowProd i
  etiqV_expected$ = Get label of interval... numChampV_expected i
  etiqV_produced$ = Get label of interval... numChampV_produced i
  etiqFocusVow$ = Get label of interval... numChampFocusVow i

  #on verifie qu'on est bien dans un intervalle non vide
  if length (etiqInterv$) > 0
    tDebut = Get starting point... numChampVowelRank i
    tFin = Get end point... numChampVowelRank i
    t1 = 1000*tDebut
    t2 = 1000*tFin
    duree = 1000*(tFin - tDebut)
    fileappend 'fichResultats$' 'etiqInterv$' 'tab$'
'etiqV_expected$' 'tab$' 'etiqV_produced$' 'tab$' 'etiqStressVowNorm$'
'tab$' 'etiqStressVowProd$' 'tab$' 'etiqFocusVow$' 'tab$' 't1' 'tab$'
't2' 'tab$' 'duree' 'newline$'
    endif

    i = i+1
  endwhile
fileappend 'fichResultats$' 'newline$'

println Fichier de résultats généré
:'newline$' 'chemin$' \ 'fichResultats$'

# nettoyage : on efface les objets créés
select TextGrid 'nomFichier$'
plus Sound 'nomFichier$'
Remove

```

**Annexe 3**  
**Exemplaire du formulaire de consentement de participation soumis aux étudiants**

**DOCUMENT D'INFORMATION et CONSENTEMENT DE PARTICIPATION**

*Remis aux personnes sollicitées pour participer aux travaux de recherche décrits ci-dessous.*

**Investigateur principal :** Francis PICAUVET  
Université Stendhal Grenoble 3 & Grenoble INP

**Responsables scientifiques :**

- ♦ Solange ROSSATO (Laboratoire d'Informatique de Grenoble / LIG & Université Stendhal)
- ♦ Véronique AUBERGÉ (CNRS)

**Intervenants :** Francis PICAUVET et divers professeurs du Centre des langues de Grenoble INP/Ense3

**Lieu principal d'enregistrements audio-visuels:**

Centre des langues de Grenoble INP/Ense3  
11, rue des Mathématiques - BP 46  
38402 Saint Martin d'Hères cedex France

**Cadre général :** Mémoire de Master 2 « recherche » en *Sciences du langage* à l'Université Stendhal et Grenoble INP, option Langue, Parole, Variation et Surdité (LPVS) :

***English: the Language-and-Music Match — Compétence en perception et production de l'anglais oral des francophones : une approche poético-musicale peut-elle améliorer les performances prosodiques en situation d'immersion?***

**Hypothèse de travail**

L'hypothèse servant de base à ce travail de recherche est qu'une approche poétique et musicale, appliquée à l'enseignement de la langue anglaise, notamment aux francophones, permet une progression qualitative plus rapide et plus efficace en perception et production orales qu'un enseignement classique dans l'acquisition et la production des caractères prosodiques suprasegmentaux et segmentaux, pour de meilleures performances en situation d'immersion, notamment grâce à un travail spécifique sur le rythme.

Pour ce faire, nous aurons besoin d'enregistrer votre parole et vos actes via des média audio-visuels et/ou informatiques (laboratoire de langue, notamment).

**Méthodologie :**

La partie de cette étude vous concernant directement comprendra notamment, outre le présent formulaire :

1. Le remplissage d'un questionnaire sur votre passé et présent linguistique.
2. Le double passage d'un même test de compétence, avant et après le travail spécifique proposé, à des fins d'évaluation uniquement scientifique (sans notation).
3. Un travail spécifique d'apprentissage linguistique, basé sur le rythme et la musicalité naturels de l'anglais, tout en abordant le sujet traité ordinairement en cours de langue anglaise à l'Ense3.
4. Une information sur les résultats de l'étude à laquelle vous aurez participé, lorsque celle-ci sera achevée (printemps 2012), si vous le souhaitez.
5. Toute publication des résultats de cette étude sera faite avec la garantie de votre anonymat.

**NOM :**

**Prénom :**

**Date de naissance :**

**Langue Maternelle :**

***J'ai lu et compris les consignes et je suis volontaire pour participer à cette expérience.  
Je donne mon consentement pour autoriser l'Université Stendhal et Grenoble INP à se servir de mon image et des données acoustiques enregistrées à des fins de recherche, sous le couvert de l'anonymat, sur une durée de 10 ans.***

**Date :**

**Signature**

## Table des illustrations (dans le texte)

Figure 1. Exemple de fenêtre du logiciel PRAAT. Les « tiroirs » 1 à 6 concernent le découpage syllabique ; le tiroir 7 indique les frontières des pauses ; les tiroirs 8 à 13 concernent le découpage des sons vocaliques. Le découpage a été effectué manuellement. ....	26
Figure 2. Les 12 phrases étudiées selon l'ordre original dans les tests $T_{init}$ , $T_{term}$ et $T_{ref}$ . ....	37
Figure 3. Niveaux de <i>stress</i> perçus de la phrase 01. L'histogramme indique le niveau de <i>stress</i> de référence. ....	41
Figure 4. Niveaux de <i>stress</i> de référence et perçus pour les 26 syllabes de la phrase 01.....	42
Figure 5. Transcription phonémique des 26 syllabes de Ph01, en $T_{init}$ , $T_{ref}$ , $T_{term}$ . Les items sur fond blanc s'éloignent de la référence, ceux sur fond noir s'approchent. ....	46
Figure 6. Evolution de la fréquence, de la durée et de l'intensité au cours de la phrase 01, pour les trois tests. ....	48
Figure 7. Durée brute des voyelles et des syllabes mesurée pour les 26 items de Ph01, en $T_{init}$ .....	52
Figure 8. Durée brute des voyelles et des syllabes mesurée pour les 26 items de Ph01, en $T_{init}$ .....	52
Figure 9. Durée brute des voyelles et des syllabes mesurée pour les 26 items de Ph01, en $T_{ref}$ .....	54
Figure 10. Durée brute des <b>syllabes</b> mesurée pour les 26 items de Ph01, en $T_{init}$ , $T_{term}$ et $T_{ref}$ .....	55
Figure 11. Durée brute des <b>voyelles</b> mesurée pour les 26 items de Ph01, en $T_{init}$ et $T_{ref}$ .....	57
Figure 12. Durée brute des <b>voyelles</b> mesurée pour les 26 items de Ph01, en $T_{term}$ et $T_{ref}$ . ....	57
Figure 13. Quantité consonantique calculée pour les items #1 à 16 de Ph01, en $T_{ref}$ , avec niveaux de <i>stress</i> de référence. ....	60
Figure 14. Quantité consonantique calculée pour les 26 items de Ph01, en $T_{init}$ , $T_{term}$ et $T_{ref}$ . ....	62
Figure 15. Quantité consonantique calculée pour les items #17 à 26 de Ph01, en $T_{init}$ , $T_{term}$ et $T_{ref}$ . ....	63
Figure 16. Niveaux de <i>stress</i> perçus pour les items #17 à 26 de Ph01, en $T_{init}$ , $T_{term}$ et $T_{ref}$ .....	64
Figure 17. Quantité consonantique calculée pour les items #1 à 16 de Ph01, en $T_{init}$ et $T_{ref}$ . ....	65
Figure 18. Quantité consonantique calculée pour les items #1 à 16 de Ph01, en $T_{term}$ et $T_{ref}$ .....	66
Figure 19. Vitesse de parole instantanée pour les 26 syllabes de Ph01, en $T_{ref}$ , et niveaux de <i>stress</i> de référence. ....	69
Figure 20. Vitesse de parole instantanée pour les 26 syllabes de Ph01, en $T_{init}$ et $T_{ref}$ . ....	70
Figure 21. Vitesse de parole instantanée pour les 26 syllabes de Ph01, en $T_{term}$ et $T_{ref}$ . ....	71
Figure 22. Accélération syllabique pour les 26 syllabes de Ph01, en $T_{ref}$ . ....	72
Figure 23. Accélération syllabique pour les 26 syllabes de Ph01, en $T_{init}$ et $T_{ref}$ . ....	74
Figure 24. Accélération syllabique pour les 26 syllabes de Ph01, en $T_{term}$ et $T_{ref}$ . ....	75
Figure 25. Evolution de $F_0$ pour les 26 voyelles de Ph01 en $T_{ref}$ et niveaux de <i>stress</i> de référence. ....	78
Figure 26. Evolution de $F_0$ pour les 26 voyelles de Ph01 en $T_{init}$ et $T_{ref}$ . ....	79
Figure 27. Evolution de $F_0$ pour les 26 voyelles de Ph01 en $T_{term}$ et $T_{ref}$ . ....	80
Figure 27 <sup>bis</sup> . Evolution de $F_0$ pour les 26 voyelles de Ph01 en $T_{term}$ et $T_{ref}$ . <b>Echelles dissemblables</b> . ....	81
Figure 28. Evolution de l'intensité pour les 26 voyelles de Ph01 en $T_{ref}$ , avec niveaux de <i>stress</i> de référence. ....	83
Figure 29. Evolution de l'intensité pour les 26 voyelles de Ph01 en $T_{init}$ $T_{ref}$ . ....	84
Figure 30. Evolution de l'intensité pour les 26 voyelles de Ph01 en $T_{term}$ et $T_{ref}$ . ....	85
Figure 31. Evolution de l'intensité pour les 26 voyelles de Ph01 en $T_{init}$ et $T_{term}$ . ....	86
Figure 32. Légende du graphique de distribution du <i>stress</i> dans PH01 ....	89
Figure 33. Représentation du <i>stress</i> pour les 26 voyelles de Ph01 en $T_{init}$ .....	90
Figure 34. Représentation du <i>stress</i> pour les 26 voyelles de Ph01 en $T_{term}$ .....	91
Figure 35. Représentation du <i>stress</i> pour les 26 voyelles de Ph01 en $T_{ref}$ ....	92
Figure 36. Distribution du <i>stress</i> pour le syntagme #1 de Ph01 en $T_{init}$ , $T_{term}$ et $T_{ref}$ .....	94
Figure 36 bis. Distribution du <i>stress</i> pour le syntagme #1 de Ph01 en $T_{init}$ , $T_{term}$ et $T_{ref}$ . Agrandissement. ....	95
Figure 37. Distribution du <i>stress</i> pour le syntagme #2b de Ph01 en $T_{init}$ , $T_{term}$ et $T_{ref}$ .....	96
Figure 38. Distribution du <i>stress</i> pour le syntagme #2b de Ph01 en $T_{init}$ , $T_{term}$ et $T_{ref}$ . Représentation dynamique. ....	97

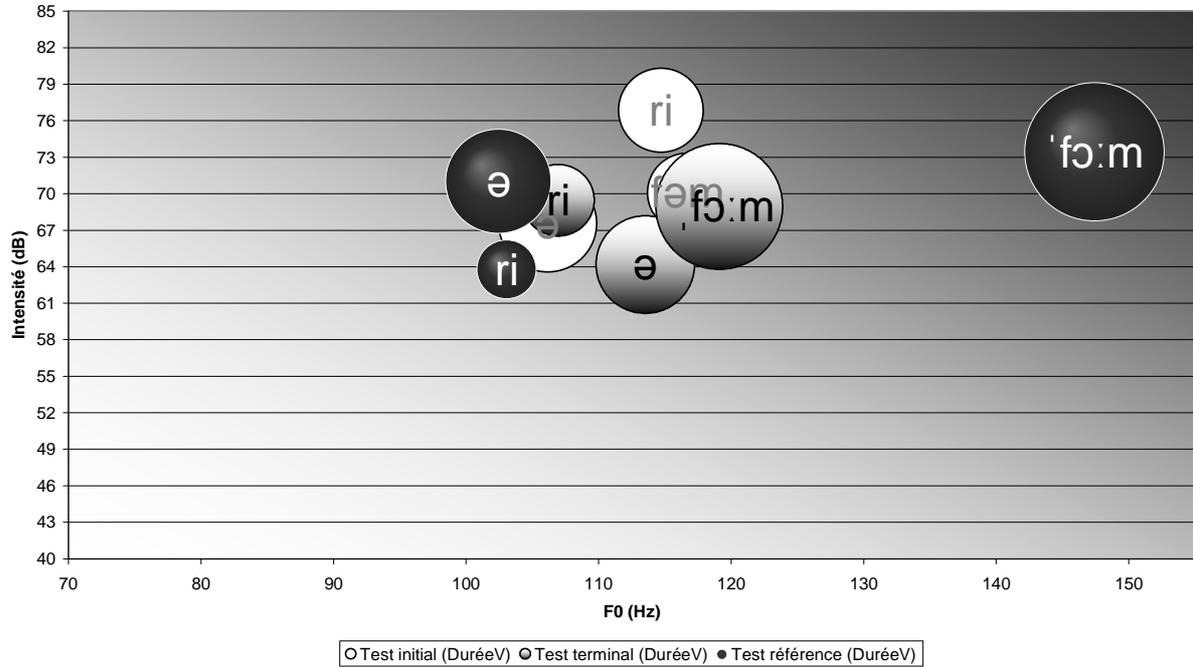
Figure 39. Tableaux de distribution des niveaux de <i>stress</i> de référence et des phonèmes des phrases Ph01 et Ph03.....	99
Figure 40. Tableaux de distribution des niveaux de <i>stress</i> de référence et des phonèmes des phrases Ph05 et Ph10.....	100
Figure 41. Durée brute des diphtongues en $T_{init}$ et $T_{term}$ . Phrases 01-03-05-10. ....	104
Figure 42. Intensité des diphtongues en $T_{init}$ et $T_{term}$ . Phrases 01-03-05-10. ....	105
Figure 43. Durée brute des voyelles longues en $T_{init}$ et $T_{term}$ . Phrases 01-03-05-10. ....	107
Figure 44. $F_0$ des voyelles longues en $T_{init}$ et $T_{term}$ . Phrases 01-03-05-10. ....	108
Figure 45. Intensité des voyelles longues en $T_{init}$ et $T_{term}$ . Phrases 01-03-05-10. ....	109
Figure 46. Durée brute des voyelles brèves /ɪ/ et /e/ en $T_{init}$ et $T_{term}$ . Phrases 01-03-05-10.....	110
Figure 47. Durée brute des voyelles brèves /æ/-/ɒ/-/ʌ/-/ʊ/ en $T_{init}$ et $T_{term}$ . Phrases 01-03-05-10. .	111
Figure 48. $F_0$ des voyelles faibles /i/-/ə/-/u/ en $T_{init}$ et $T_{term}$ . Phrases 01-03-05-10. ....	112
Figure 49. Intensité des voyelles faibles /i/-/ə/-/u/ en $T_{init}$ et $T_{term}$ . Phrases 01-03-05-10.....	113
Figure 50. Valeurs mesurées de $F_1$ et $F_2$ du son /i:/ en 6 occurrences parmi les phrases Ph01 à Ph12, dans les deux versions du test $T_{init}$ et $T_{term}$ . Également : les valeurs moyennes de $F_1$ et $F_2$ des trois séries RP ( <b>rp-1997</b> ; <b>rp-2001-20-25</b> ; <b>rp-2003</b> ) et la moyenne des données de $T_{ref}$ ( <b>Ref-moy</b> ). Les flèches pleines indiquent la variation entre $T_{init}$ (◆) et $T_{term}$ (■). Les flèches en pointillé indiquent la distance entre $T_{term}$ (■) et $T_{ref}$ (●) pour chaque occurrence. Légende identique dans les graphiques suivants.....	118
Figure 51. Valeurs mesurées de $F_1$ et $F_2$ du son /a:/ en 6 occurrences parmi les phrases Ph01 à Ph12, dans les deux versions du test $T_{init}$ et $T_{term}$ , les trois séries RP ( <b>rp-1997</b> ; <b>rp-2001-20-25</b> ; <b>rp-2003</b> ) et la moyenne des données de $T_{ref}$ ( <b>Ref-moy</b> ). Légendes : idem graphique précédent.....	120
Figure 52. Valeurs mesurées de $F_1$ et $F_2$ du son /ɔ:/ en 6 occurrences parmi les phrases Ph01 à Ph12, dans les deux versions du test $T_{init}$ et $T_{term}$ , les trois séries RP ( <b>rp-1997</b> ; <b>rp-2001-20-25</b> ; <b>rp-2003</b> ) et la moyenne des données de $T_{ref}$ ( <b>Ref-moy</b> ). Légendes : idem graphiques précédents.....	121
Figure 53. Valeurs mesurées de $F_1$ et $F_2$ du son /u:/ en 6 occurrences parmi les phrases Ph01 à Ph12, dans les deux versions du test $T_{init}$ et $T_{term}$ , les trois séries RP ( <b>rp-1997</b> ; <b>rp-2001-20-25</b> ; <b>rp-2003</b> ) et la moyenne des données de $T_{ref}$ ( <b>Ref-moy</b> ). Légendes : idem graphiques précédents.....	123
Figure 54. Synthèse des trapèzes vocaliques de type API pour /i/-/a:/-/ɔ:/-/u:/ en valeurs moyennées de Jean (phrases Ph01 à Ph12) en $T_{init}$ et $T_{term}$ et de nos quatre séries de données de référence. ....	124

## **Table des illustrations annexes**

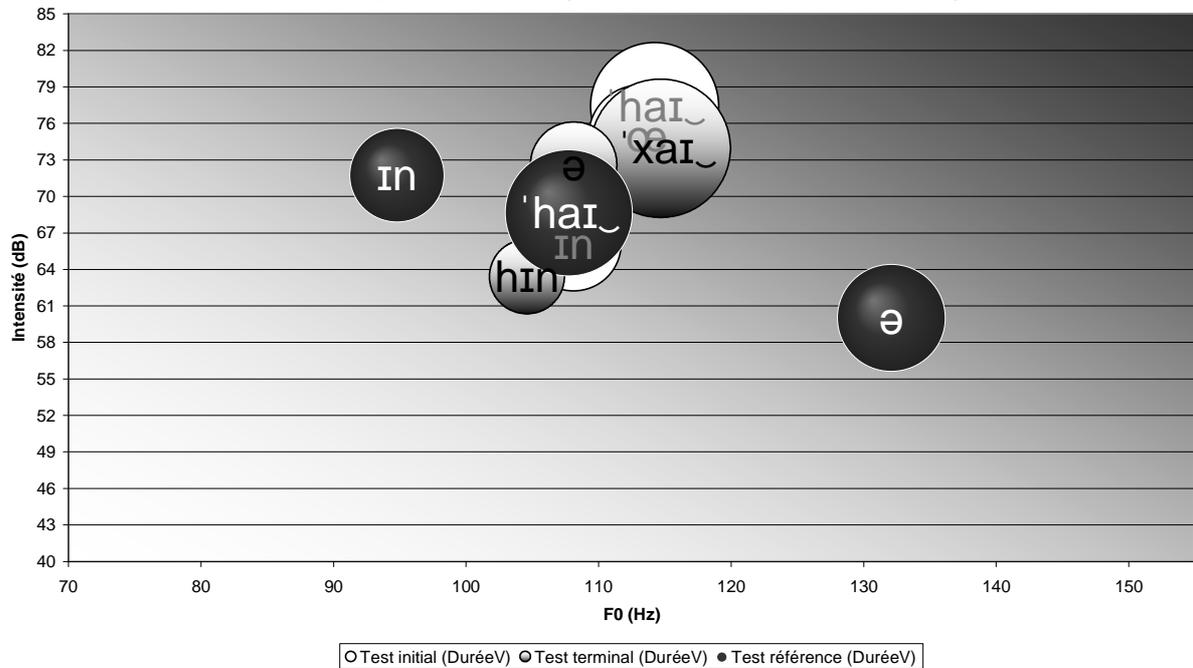
Illustration 1. Graphiques d'analyse syntagmatique comparative du stress des 26 voyelles de la phrase Ph01 .....	141
--	-----

# Illustration 1. Graphiques d'analyse syntagmatique comparative du stress des 26 voyelles de la phrase Ph01

Distribution du stress\_V-Init/V-Term/V-Ref\_Ph01\_Syntagme 1\_V1.1 à V1.3  
 01. A **reform** in higher education took place this autumn and the situation has improved, at last.

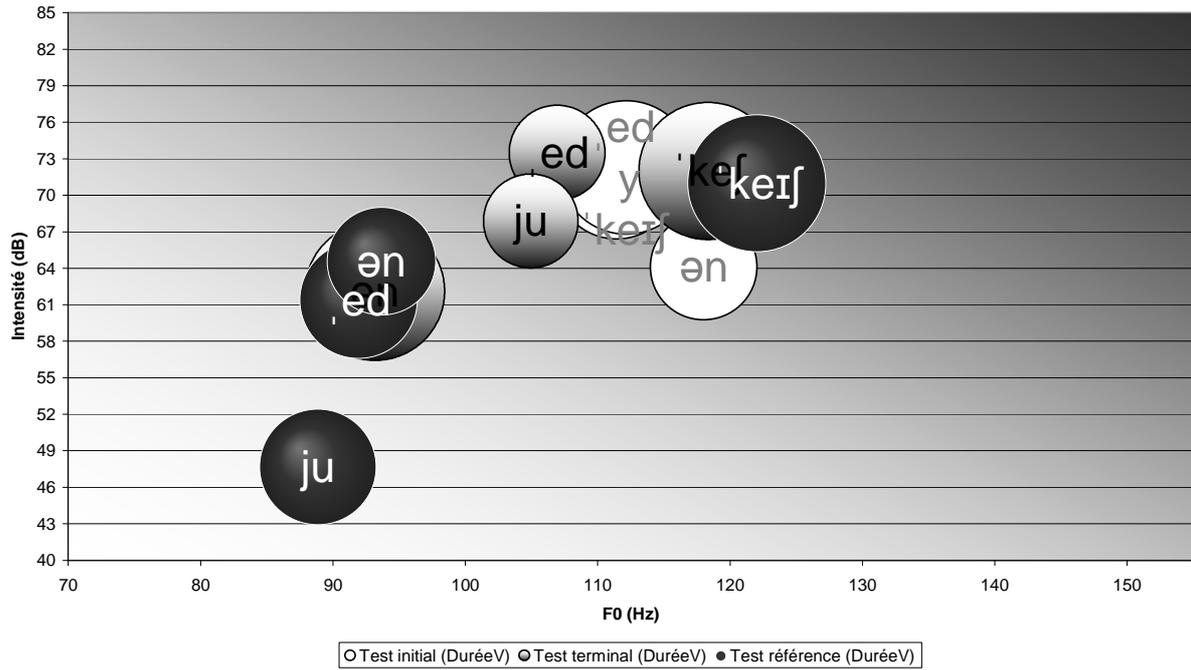


Distribution du stress\_V-Init/V-Term/V-Ref\_Ph01\_Syntagme 2a\_V1.4 à V1.6  
 01. A reform **in higher** education took place this autumn and the situation has improved, at last.



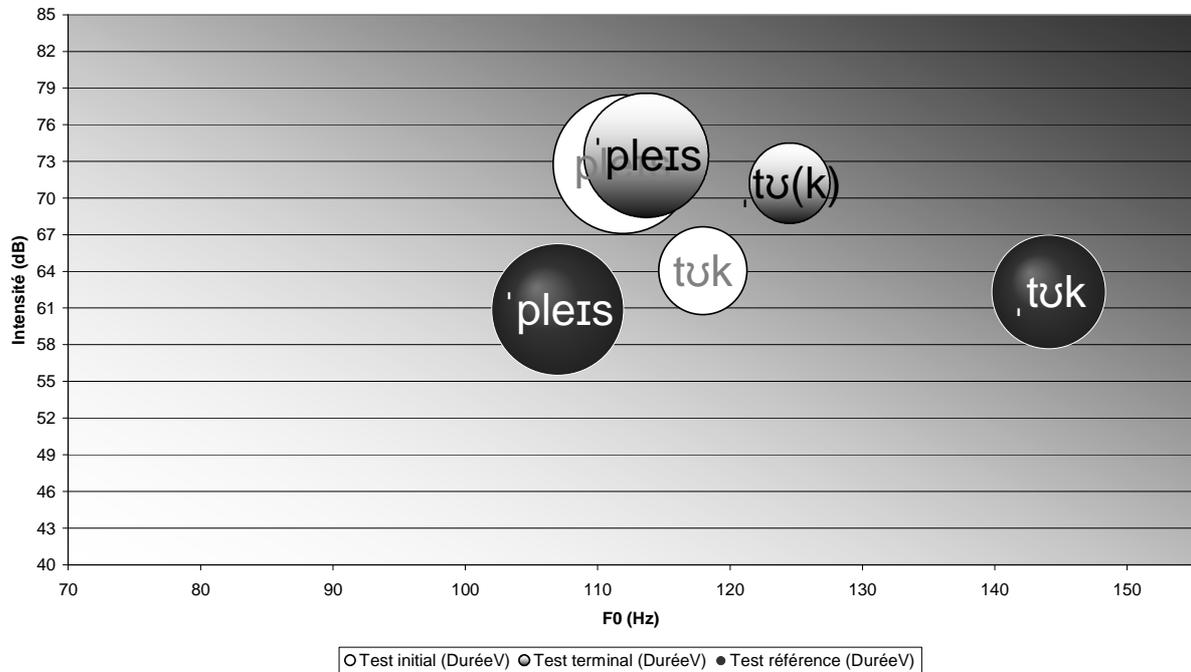
Distribution du stress\_V-Init/V-Term/V-Ref\_Ph01\_Syntagme 2b\_V1.7 à V1.10

01. A reform in higher **education** took place this autumn and the situation has improved, at last.



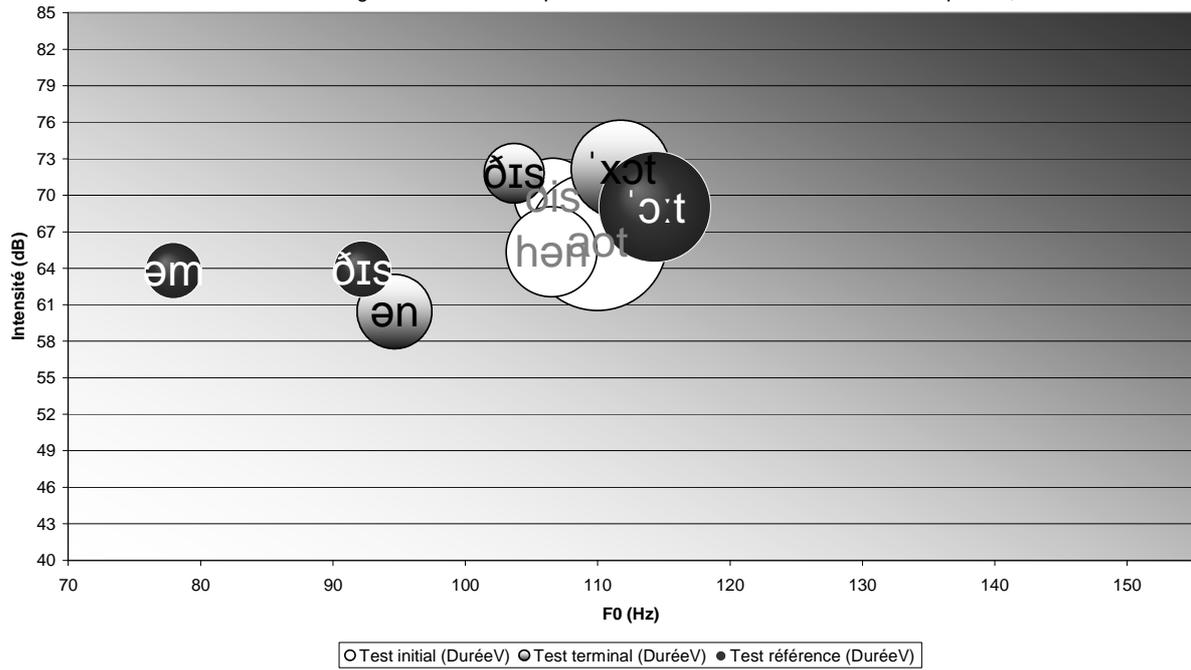
Distribution du stress\_V-Init/V-Term/V-Ref\_Ph01\_Syntagme 3\_V1.11 à V1.12

01. A reform in higher education **took place** this autumn and the situation has improved, at last.



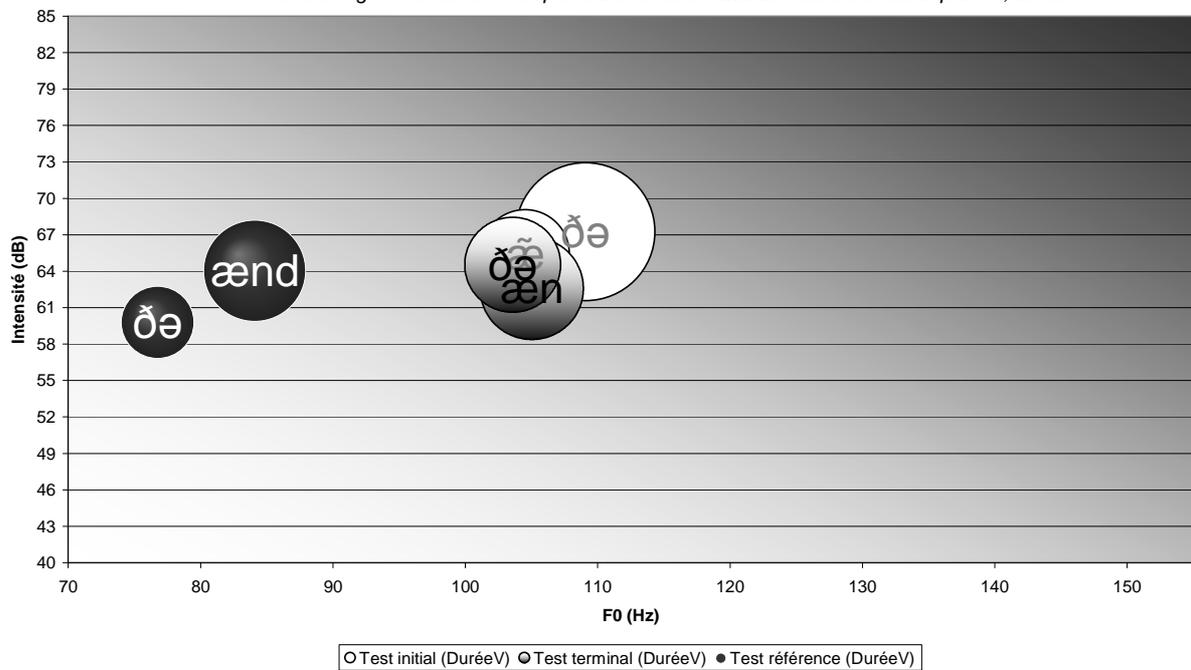
Distribution du stress\_V-Init/V-Term/V-Ref\_Ph01\_Syntagme 4\_V1.13 à V1.15

01. A reform in higher education took place **this autumn** and the situation has improved, at last.



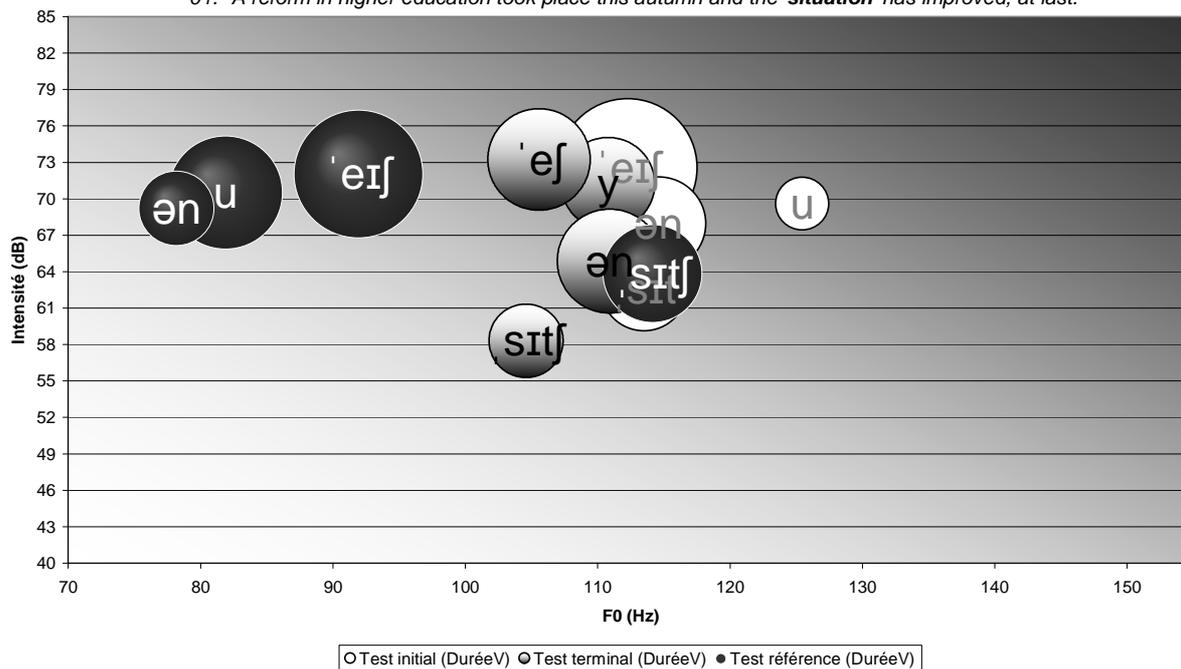
Distribution du stress\_V-Init/V-Term/V-Ref\_Ph01\_Syntagme 5a\_V1.16 à V1.17

01. A reform in higher education took place this autumn **and the** situation has improved, at last.



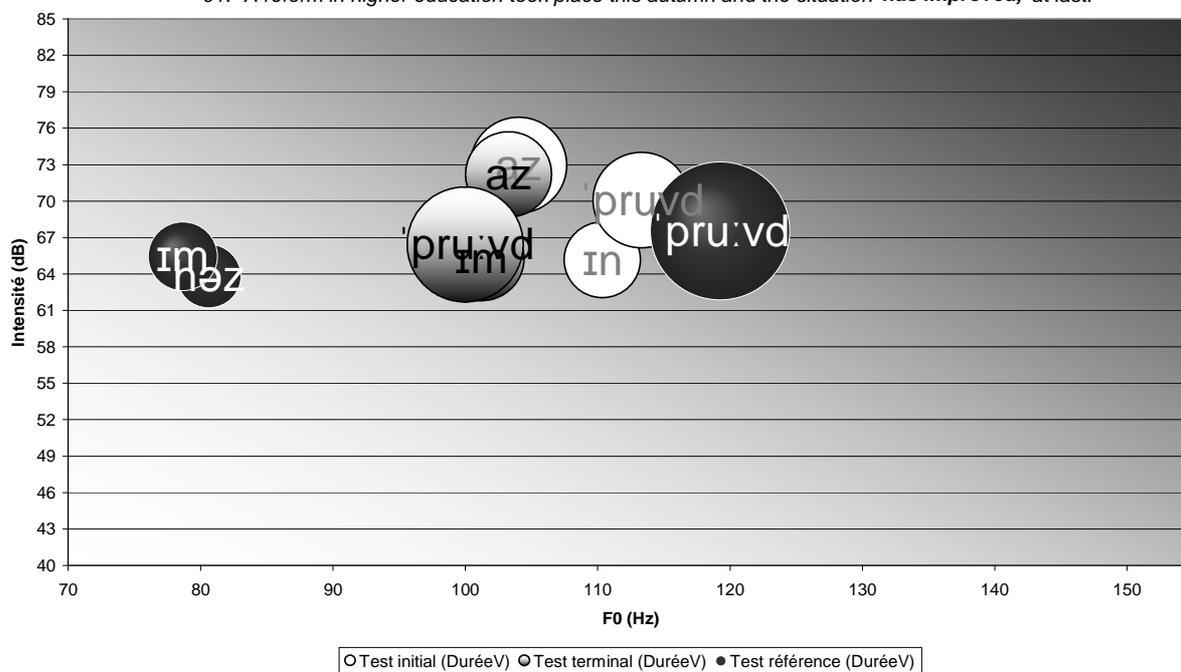
Distribution du stress\_V-Init/V-Term/V-Ref\_Ph01\_Syntagme 5b\_V1.18 à V1.21

01. A reform in higher education took place this autumn and the **situation** has improved, at last.



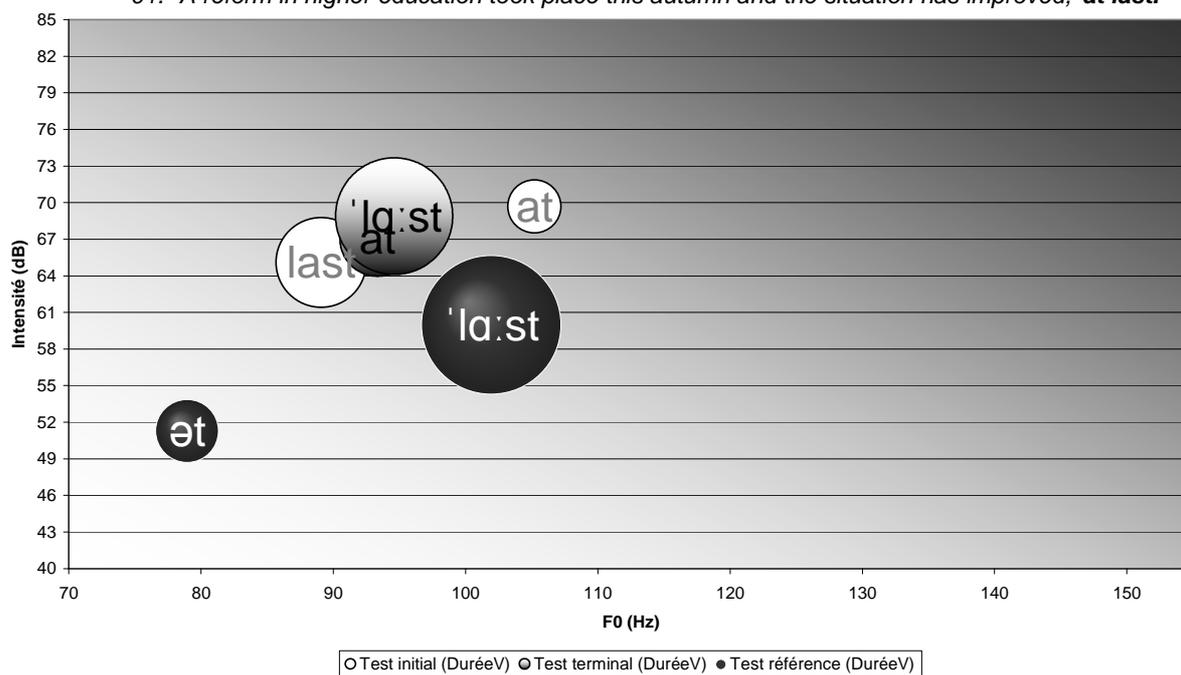
Distribution du stress\_V-Init/V-Term/V-Ref\_Ph01\_Syntagme 6\_V1.22 à V1.24

01. A reform in higher education took place this autumn and the situation **has improved**, at last.



Distribution du stress\_V-Init/V-Term/V-Ref\_Ph01\_Syntagme 7\_V1.25 à V1.26

01. A reform in higher education took place this autumn and the situation has improved, **at last**.



## Sigles, abréviations, symboles et unités utilisés

NB : ces abréviations sont classées par thèmes, puis par ordre alphabétique

#	Numéro ; notamment : n° d'item (ex. voyelle, syllabe, etc.)
API	Alphabet phonétique international (angl. : IPA)
Ex.	Exemple ; par exemple
GA	<i>General American</i> (anglais américain théorique de référence)
RP	<i>Received pronunciation</i> (anglais britannique théorique de référence, utilisé ici dans sa version étendue)
£	Anglais britannique
\$	Anglais américain
/ <sup>h</sup> Itəl/	Transcription : phonémique ; phonologique ; théorique ; de référence ; simple
[ <sup>h</sup> Itɹ]	Transcription : phonétique ; sons ; allophones ; « vraie ». Utilisée ici surtout pour différencier, si nécessaire, un son produit d'un son de référence.
C	Consonne
S	Syllabe
V	Voyelle
Ph03	Phrase 03 (référence de la phrase étudiée : 12 phrases au total, de 01 à 12)
V1.20	20 <sup>ème</sup> voyelle de la phrase 01 ; etc.
S5.4	4 <sup>ème</sup> syllabe de la phrase 05 ; etc.
d	Durée
I	Intensité
F <sub>0</sub>	Fréquence fondamentale ; fréquence de vibration des plis vocaux
F <sub>1</sub>	1 <sup>er</sup> formant (distribution de l'énergie du signal de parole : fréquences basses)
F <sub>2</sub>	2 <sup>ème</sup> formant (distribution de l'énergie du signal de parole : fréquences médium)
F <sub>3</sub>	3 <sup>ème</sup> formant (distribution de l'énergie du signal de parole : fréquences aigües)
T <sub>init</sub>	Test initial (1ère instance de recueil des données du corpus)
T <sub>term</sub>	Test terminal (dernière instance de recueil des données du corpus)
T <sub>ref</sub>	« Test » de référence (phrases du test lues par le professeur)
L1	Langue première, maternelle
L2	Langue seconde, étrangère ; « langue vivante 2 / LV2 » (système éducatif)
L3	3ème langue (2ème langue étrangère) ; « langue vivante 3 / LV3 » (système éducatif)
LN	Locuteur natif (ici : anglophone, le plus souvent)

<b>LnN</b>	Locuteur non-natif (ici : non-anglophone ; le plus souvent : francophone)
<b>St1</b>	<i>Stress</i> de niveau 1 (accent dynamique lexical)
<b>St2</b>	<i>Stress</i> de niveau 2 (accent dynamique lexical)
<b>St3</b>	<i>Stress</i> de niveau 3 (accent dynamique lexical)
<b>V<sub>diph</sub></b>	Diphthongue, <i>glide</i> (voyelle dont le point d'articulation varie constamment)
<b>V<sub>long</sub></b>	Voyelle longue
<b>V<sub>short</sub></b>	Voyelle courte
<b>V<sub>weak</sub></b>	Voyelle faible
<b>V<sub>max</sub></b>	Durée maximum de la voyelle
<b>V<sub>min</sub></b>	Durée minimum de la voyelle
<b>S<sub>max</sub></b>	Durée maximum de la syllabe
<b>S<sub>min</sub></b>	Durée minimum de la syllabe
<b>a</b>	Accélération ; ici : accélération syllabique en $S \cdot s^{-2}$
<b>v</b>	Vitesse ; ici : vitesse de parole instantanée
<b>s</b>	Seconde
<b>ms</b>	Milliseconde ; $10^{-3}$ seconde

...

**MOTS-CLÉS** : anglais, L2, métronome, rythme, accent, parole

## **RÉSUMÉ**

La dichotomie isoaccentuelle-isosyllabique entre les langues anglaise et française constitue un défi majeur, tant pour l'enseignant que pour l'apprenant de l'anglais « langue seconde ». L'accent dynamique lexical (*stress*), fixe en anglais, joue un rôle fondamental dans la compréhension (ou non) du message, et son absence ou son déplacement rendent souvent l'interaction avec un locuteur natif très difficile. Ceci ne pouvant être ignoré, l'expérimentation décrite dans ce mémoire relève d'une approche pédagogique originale pour l'apprentissage du rythme, basée sur l'utilisation fréquente d'un métronome lors de tâches de production orale en cours. Un test de lecture, utilisé en périodes initiale et terminale, a permis d'évaluer l'apport du métronome, un fois celui-ci retiré. Les productions de soixante-quinze étudiants sont enregistrées lors des tests ; des mesures sont effectuées et analysées afin de déterminer s'il y a eu progrès. Les résultats de l'étude d'un cas, issu du corpus créé, montrent des progrès dans certains paramètres (durée vocaliques, hauteur mélodique) et une stagnation dans d'autres (intensité, qualité vocalique). Cette étude préliminaire permet de jeter les bases d'une future étude de plus grande envergure basée sur le corpus construit.

**KEYWORDS** : English, L2, metronome, rhythm, stress, speech

## **ABSTRACT**

The stress-timed vs. syllable-timed dichotomy between English and French is at the origin of a major challenge for teachers and students of English L2 alike. Stress, not a free parameter in English words, plays a fundamental role in the good understanding of the message. Its absence, or its misplacement, makes the interaction with a native speaker extremely difficult. This cannot to be ignored: the experimentation described in this paper is based on an original pedagogical approach for the acquisition of rhythm using a metronome frequently, during oral production tasks in class. A reading test was used twice, before and after the work with the metronome, enabling the measuring and analysis of progress-related data. The productions of seventy five students were recorded during the tests; measures were made and the collected data analyzed in order to determine whether there was progress or not. First results of a case study show definite progress in some parameters (vowel duration, pitch) and stagnation in others (intensity, vowel quality). This preliminary study paves the way to a broader study based on the created corpus.