

Catalogue d'interférences anglais-français

*« Une étude comparative du système de la langue étrangère par rapport au système de la langue maternelle fournit les bases d'une meilleure prise de conscience par l'élève des différences entre les deux systèmes »
(Lanchec, 1976 : 162).*

Table du cinquième chapitre

1) Discussion méthodologique : choix et principes généraux utilisés

- 1.1. L'anglais comme L2
- 1.2. Une théorie de l'interférence : du connu vers l'inconnu
- 1.3. Nécessité d'une étude contrastive
- 1.4. Application de ces principes
 - 1.4.1. Catalogue d'interférences
 - 1.4.2. Batterie de tests de perception

2. Description comparée des systèmes prosodiques de l'anglais et du français

- 2.1. Rythme
 - 2.1.1. Localisation de l'accent
 - 2.1.2. Différence de réalisation acoustique de l'accent
 - 2.1.3. Différence de rythmique
 - 2.1.4. Accent de groupe vs accent de mot
 - 2.1.5. Langue non accentuée vs langue accentuée ?
- 2.2. Intonation
 - 2.2.1. L'accent de phrase
 - 2.2.2. Les tons mélodiques
 - 2.2.3. Description comparée de l'intonation déclarative en anglais et en français
 - 2.2.3.1. Finalité
 - 2.2.3.2. Continuation

3. Description comparée du système vocalique de l'anglais et du système vocalique du français

- 3.1. Caractéristiques générales
 - 3.1.1. Conséquences de la différence d'organisation rythmique sur le niveau segmental
 - 3.1.2. Différences d'organisation de l'espace articulatoire
 - 3.1.3. Conséquences pour un apprenant francophone

- 3.1.3.1. Prédiction d'une typologie des fautes découlant de ces différences
- 3.1.3.2. Paramètres importants nécessaires à la description de l'anglais pour un apprenant francophone
 - 3.1.3.2.1. Durée
 - 3.1.3.2.2. Timbre et stabilité de timbre
- 3.1.3.3. Nécessité d'une description plus poussée
- 3.1.4. Réalisation d'une étude descriptive
 - 3.1.4.1. Locutrices
 - 3.1.4.2. Matériel de parole
 - 3.1.4.3. Conditions d'enregistrement
 - 3.1.4.4. Analyses
- 3.2. Caractéristiques détaillées
 - 3.2.1. Les noyaux syllabiques simples de l'anglais et du français
 - 3.2.1.1. Durée
 - 3.2.1.1.1. Les voyelles brèves
 - 3.2.1.1.2. Les voyelles longues
 - 3.2.1.1.3. Les voyelles du français
 - 3.2.1.1.4. Interférences de durée du français vers l'anglais
 - 3.2.1.2. Timbre
 - 3.2.2. Les noyaux syllabiques complexes de l'anglais
 - 3.2.2.1. Définitions
 - 3.2.2.1.1. La diphtongue
 - 3.2.2.1.2. La triptongue
 - 3.2.2.2. Description des diphtongues et des triptongues de l'anglais en fonction des productions de la locutrice R.P. comparées à celles de la locutrice francophone.
 - 3.2.2.2.1. Les diphtongues
 - 3.2.2.2.2. Les triptongues

Bibliographie du chapitre V

1. Discussion méthodologique : choix et principes généraux utilisés

1.1. L'anglais comme L2

Nous avons choisi de travailler sur l'anglais, pour au moins deux raisons, la première étant d'ordre institutionnel et la seconde, d'ordre linguistique. On constate tout d'abord une prédominance de l'enseignement de l'anglais dans le premier degré. Pendant l'année scolaire 1999-2000, la proportion d'élèves apprenant l'anglais dans le primaire était de 79,8 % contre 15,2 % pour l'allemand, 2,4 % pour l'espagnol et 1,3 % pour l'italien. Qui plus est, la tendance est au recul pour ces trois dernières langues (respectivement 15,8 %, 2,8% et 1,5% en 99-2000¹). Par ailleurs, on trouve cette prédominance dans toutes les académies, sauf celles de Strasbourg et de Nancy. Ce pourcentage ne fait que confirmer la tendance nationale au niveau du secondaire et la tendance européenne (82,7 % des classes de l'Union étaient concernées par l'enseignement de l'anglais en 1991-92). Dans ces conditions, il était donc beaucoup plus simple de trouver des classes en apprentissage précoce de l'anglais. Mais le choix de l'anglais s'explique surtout par ses caractéristiques phonétiques. Si cette langue est relativement proche du français du point de vue du vocabulaire et de l'ordre des mots, il n'en va pas de même en ce qui concerne la prononciation. Les deux langues possèdent des systèmes très différents. Alors que la rythmicité de l'anglais repose sur une tendance accentuelle (Pike, 1946), celle du français repose sur une tendance syllabique. Cette différence de structuration rythmique engendre un régime de tension musculaire très différent, qui conditionne directement la prononciation des voyelles et en particulier la stabilité de leur timbre. Par ailleurs, on pourrait qualifier l'anglais de langue « musicale » car elle se caractérise par de fortes variations de hauteur et d'intensité et elle couvre une large tessiture (plus de deux octaves selon Ginésy, 1995 : 223). Du point de vue descriptif, Faure (1962) a souligné la richesse mélodique de

¹ INSEE (2001). Note d'Information 01-18. Edu050, avril.

l'anglais :

« cette richesse mélodique, à laquelle on a rarement accordé toute l'attention qu'elle mérite, est encore plus frappante lorsqu'on la compare aux ressources correspondantes d'autres langues modernes, comme par exemple, le français ou l'allemand ».

Cette musicalité frappe d'ailleurs l'oreille de l'apprenant francophone :

« nous avons de nombreuses fois entendu les réflexions de nos apprenants français qui, captant une multitude de contrastes sur les syllabes en anglais, ressentent la langue anglaise comme une langue qui « chante » et « accentue » » (Scott, 1987 : 135).

Enfin, le système vocalique de l'anglais est plus riche que celui du français en ce sens qu'il contient plus de contrastes (26 contre 16) et que certains de ces contrastes, les diphtongues et les triphthongues n'existent pas en français.

1.2. Une théorie de l'interférence : du connu vers l'inconnu

L'attitude qui consiste à analyser l'inconnu à partir du connu est tout à fait naturelle et nous portons des « préjugés », c'est-à-dire que nous appliquons les règles de fonctionnement de ce que l'on connaît à ce que l'on ne connaît pas. Dans un premier temps, nous avons besoin de ces préjugés, car ils nous permettent d'interpréter le monde à la lumière de ce que nous en savons et par la même, de nous adapter à lui. Le problème fondamental, c'est que nous n'avons pas conscience que la réalité que nous percevons est une construction établie à partir de la structure de notre propre système, autrement dit que cette réalité est filtrée par notre propre système perceptif. Ainsi, lorsqu'on se trouve confronté à une langue étrangère, on la compare intuitivement à sa LM. En effet, l'expérience a forgé un ensemble d'attentes perceptives (Lhote, 1995 : 48) et d'habitudes articulatoires qui ont permis l'acquisition de la LM, mais qui se montrent inadaptées au traitement et à la production d'une langue étrangère, surtout dans le cas où la langue cible est très éloignée de la LM (comme c'est le cas pour l'anglais et le français). Comme le souligne Troubetzkoy dès 1949 :

« l'homme s'approprie le système de sa langue maternelle. Mais s'il entend parler une autre langue, il emploie involontairement pour l'analyse de ce qu'il entend, le « crible phonologique »

de sa langue maternelle qui lui est familier. Et comme ce crible ne convient pas pour la langue étrangère entendue, il se produit de nombreuses erreurs et incompréhensions. Les sons de la langue étrangère reçoivent une interprétation phonologiquement inexacte, puisqu'on les fait passer par le « crible phonologique » de sa propre langue » (Troubetzkoy, 1949 : 54).

Cet effet de filtrage est particulièrement puissant, puisqu'il provoque chez l'apprenant une surdité sélective, une surdité dite « phonologique » aux sons de la langue étrangère.

Précisons que lorsque l'on parle de « surdité », on ne fait pas référence à un déficit d'audition, mais à une réorganisation perceptive en faveur de la LM. Tous les sons ne sont pas égaux devant cet « effet de filtrage » (Ostiguy et al., 1996 : 79) : certains sont assimilés purement et simplement aux sons de la LM en raison de leur similitudes de timbre ;

« on parlera alors de transfert négatif ou d'interférence lorsque l'apprenant va se servir en L2 d'un élément de L1 qu'il croit identique alors que celui-ci est différent » (Guimbretière, 1994 : 17).

D'autres sont analysés comme des exemplaires déviants, d'autres comme des sons différents, d'autres enfin comme des sons non linguistiques. Pour classer ces différences, Best propose un modèle d'assimilation perceptive, le PAM (« *Perceptual Assimilation Model* ») (1988, 1993), qui part du principe que, lorsque les différences entre les phonèmes de la langue étrangère et ceux de la LM sont claires, la discrimination est facilitée. Le modèle propose six cas de figure dans la discrimination d'une paire de phonèmes non natifs (les contrastes UU et UC ont été rajoutés au modèle en 1993) :

1. Les deux phonèmes sont assimilés à deux phonèmes différents de la LM. Dans ce cas, la discrimination est excellente (« *Two-Category, TC* »).
2. Les deux phonèmes sont assimilés à un seul phonème de la LM, mais divergent de l'« idéal natif » (l'un est acceptable et l'autre, considéré comme déviant). Comme ils sont différenciés, leur discrimination est modérée, voire bonne (« *Category-Goodness Difference, CG* »).
3. Les deux phonèmes sont assimilés à un seul phonème de la LM et divergent tous les deux de l'« idéal natif ». L'auditeur ne fait pas de différence entre les deux contrastes et la discrimination est très pauvre (« *Single-Category, SC* »).

4. Les deux phonèmes ne sont pas catégorisables par l'auditeur, car ils tombent en dehors de l'espace phonétique familier. Dans ce cas, la discrimination est modérée, voire très pauvre (« *Both Uncategorizable, UU* »).
5. L'un des deux phonèmes est assimilé à un phonème natif et l'autre tombe dans un espace phonétique non familier, en dehors des catégories natives. Leur discrimination est très bonne (« *Uncategorized vs. Categorized, UC* »).
6. Les deux phonèmes tombent en dehors du domaine du langage et sont considérés comme des sons n'appartenant pas à la parole. Leur discrimination est bonne, voire très bonne (« *Non-Assimilable, NA* »).

Les auditeurs adultes perçoivent donc les phonèmes étrangers par rapport à la structure des phonèmes appartenant à leur LM.

Ces phénomènes de crible sont d'une importance capitale pour la prononciation de la langue cible : en effet, on ne peut pas produire correctement des sons qu'on perçoit mal ou qu'on ne perçoit pas :

« ce qu'on appelle l'« accent étranger » ne dépend pas du fait que l'étranger en question ne peut pas prononcer un certain son, mais plutôt du fait qu'il n'apprécie pas correctement ce son. Et cette fausse appréciation des sons d'une langue étrangère est conditionnée par la différence existant entre la structure phonologique de la langue étrangère et celle de la langue maternelle du sujet parlant » (Troubetzkoy, 1949 : 56).

La notion de crible phonologique se réfère essentiellement au niveau segmental. En ce qui concerne le niveau prosodique, Delattre (1965) parle d'« interférences d'intonation » et De Bot (1986), de « transferts intonatifs ». Nous préférons parler de crible prosodique, de manière à englober les interférences d'ordre rythmique, qui vont également provoquer des erreurs graves,

comme nous le verrons plus loin (voir 2.2.1.). Les interférences prosodiques sont peut être les plus difficiles à combattre. En effet, l'organisation rythmique et intonative de la LM est tellement ancrée, tellement naturelle, car première dans l'ordre d'acquisition, qu'elle en devient « transparente » pour le locuteur : la plupart du temps, il n'a pas conscience de la façon dont il accentue et module ses énoncés, alors qu'il a en partie conscience de la façon dont il prononce les éléments segmentaux. Les locuteurs sont totalement imprégnés par la rythmique de leur LM et s'ils ne sont pas sensibilisés aux différences prosodiques dont nous venons de parler, ils appliqueront le patron rythmique et intonatif de leur LM à la LE, sans en avoir la moindre conscience. La surdit  prosodique est donc tr s puissante et il est tr s difficile de s'en affranchir, d' chapper   des automatismes d'autant plus install s qu'ils charpentent l'ensemble de la langue.

1.3. Nécessité d'une étude contrastive

Pour connaître les interférences qui se produiront entre deux langues, il est nécessaire de connaître le système de la langue cible et celui de la LM. Pour arriver à cette connaissance, on peut faire une analyse fine et détaillée de chacun de ces systèmes. La description prosodique et segmentale est actuellement poussée, comme en témoigne la sortie récente de plusieurs ouvrages et articles consacrés au français (Lacheret-Dujour et Beaugendre, 1999 ; Rossi, 1999 ; Di Cristo, 1999) et à l'anglais (Hirst, 1998 ; Montreuil, 2001). En effet, ces deux langues sont parmi les plus abondamment décrites. Cependant, une analyse comparative permettra de limiter la description, en mettant uniquement en lumière les points d'achoppement entre les deux langues. Ce type d'analyse « contrastive » (on parle également d'analyse différentielle) est né aux Etats-Unis dans les années 50. D'après Lado (1957), il repose sur le principe qu'en situation d'apprentissage, lorsque deux langues sont en contact, il peut y avoir deux sortes de transferts :

- soit un transfert positif, si les caractéristiques de la langue cible ressemblent à celles de la langue source ; dans ce cas, l'apprentissage de la langue cible sera facilité ;
- soit un transfert négatif, si les caractéristiques de la langue cible divergent par rapport à celles de la langue source ; dans ce cas, l'apprentissage de la langue cible sera gêné.

Du point de vue de l'apprenant, cette approche est très intéressante, car elle permet de prédire les erreurs qu'il pourrait produire. Le principal danger reste de relever plus d'erreurs potentielles que l'apprenant n'en produira réellement, notamment au niveau du lexique et de la grammaire. Cependant, l'analyse contrastive s'avère particulièrement adaptée à la prédiction des interférences de type phonétique. Par ailleurs, d'après Léon et Martin (1971 : 37), l'enseignement de l'intonation suppose une analyse préalable, comparative, de la langue de départ et de la langue d'arrivée.

Il existe des études comparatives pour l'anglais et le français, à commencer par celles de Delattre (1965) s'intéressant aux deux niveaux, prosodique et phonémique. D'autres qui ne s'occupent que d'intonation (Delattre, 1961 ; Pritchard, 1985 ; Grover et al., 1987) ou que de segmental. Certains manuels à destination des apprenants francophones sont fondés sur une analyse comparative de l'anglais par rapport au français pour la totalité du système (Faure, 1948 ; Adamczewski et al., 1973 ; Ostiguy et al., 1996).

1.4. Application de ces principes

1.4.1. Catalogue d'interférences

A partir de la littérature et de nos propres observations, nous réaliserons une analyse comparée du français et de l'anglais. Celle-ci nous permettra de prédire les interférences les plus graves qui se produiront entre ces deux langues et d'en faire le relevé systématique. Nous donnerons à cette liste d'interférences le nom de « catalogue d'interférences » (Dodane, 2001) et son élaboration respectera les principes suivants :

- à l'image de ce qui se passe lors de l'acquisition de la LM, nous accorderons la primauté à la prosodie ;
- la prosodie étant constituée du rythme et de l'intonation, nous les aborderons séparément, mais sans jamais perdre de vue leur interdépendance ;
- la première partie du catalogue concernera la prosodie et la seconde partie, l'organisation segmentale, sans jamais oublier que le premier niveau sous-tend le second ; nous ne nous intéresserons qu'aux voyelles car ces sons paraissent assumer un grand rôle pendant l'acquisition (cf. baby talk et berceuses, chapitre II, 2.4.) ; en outre, ce sont elles qui portent la plupart des paramètres prosodiques.

1.4.2. Batterie de tests de perception

Deux batteries de tests seront réalisées à partir des principales interférences relevées dans ce catalogue : une batterie « prosodie » et une batterie « contrastes vocaliques » (Volume II, annexe n°4.2.). Ces batteries serviront d'outil d'évaluation pour déterminer les capacités de discrimination en anglais et pour l'étude des interférences prosodiques et phonémiques entre le français et l'anglais. Nous avons donc réalisé une batterie de 8 tests portant sur :

- les spécificités prosodiques de l'anglais par rapport à celles du français (accent de mot, place du pivot au sein de l'énoncé, rythmicité accentuelle, tons mélodiques et schémas mélodiques).

- les contrastes vocaliques spécifiques de l'anglais par rapport aux français (oppositions vocaliques entre voyelles de timbres très proches en anglais ; oppositions vocaliques entre voyelles de timbres très proches en anglais et en français ; diphtongues et triphthongues de l'anglais).

2. Description comparée du système prosodique de l'anglais et du système prosodique du français

2.1. Rythme

Chaque langue alterne différemment les syllabes faibles et les syllabes fortes et « marque » le rythme différemment.

« Linguistiquement, nous sommes conditionnés par le rythme de notre langue maternelle. Apprendre une langue étrangère, c'est aussi, c'est avant tout, changer de rythme, subir une sorte de « recyclage » rythmique et intonatif » (Meschonnic, 1982 : 420).

Pour décrire la spécificité de cette rythmicité, nous commencerons par décrire la localisation des syllabes accentuées et observer comment se combinent les paramètres acoustiques pour réaliser l'accent. Ensuite, nous déterminerons le type de rythmique engendrée par l'alternance de ces accents.

2.1.1. Localisation de l'accent

L'accent primaire est désigné par de multiples noms, selon les auteurs. Pour souligner la différence entre les deux accents (au niveau de la localisation, de la réalisation acoustique, de la fonction...) et dans un but de clarté, nous parlerons d'accent final pour le français et d'accent tonique pour l'anglais (*« Nous préférons le désigner désormais sous le nom d'accent tonique, en signifiant par là que le ton en est le facteur essentiel »*, Faure, 1962 : 150).

2.1.1.1. Prédicibilité de l'accent final en français

Une opinion répandue parmi les phonéticiens, c'est que le français est une langue qui possède un seul accent et que cet accent affecte toujours la dernière syllabe du groupe rythmique. Dans ce sens, c'est un accent de groupe. Le groupe rythmique est un groupe de syllabes formant une unité sonore coïncidant souvent avec le syntagme grammatical.

Dans un groupe rythmique, les mots perdent leur individualité, car l'accent de mot disparaît pour se reporter à la fin du groupe. Cet accent, appelé également final, logique, objectif, tonique, normal ou encore interne, se réalise dans 100 % des cas sur la dernière syllabe (Delattre, 1965²), c'est-à-dire la pénultième du groupe rythmique, et ceci, quel que soit le nombre de syllabes composant le groupe rythmique. Pour cette raison, on parle d'oxytonie. Le français est donc une langue à accent fixe, tout comme le tchèque, le finnois, le hongrois et l'islandais (mais pour celles-ci, l'accent affecte l'initiale et non la finale). En français, l'accent primaire frappe donc la syllabe finale du groupe rythmique, engendrant ainsi un rythme de type iambique³.

2.1.1.2. Variabilité de l'accent tonique en anglais

En anglais, les mots conservent leur particularité, même s'ils sont mis en contexte au sein d'un énoncé. Tous les mots lexicaux (« *content words* » en anglais), c'est-à-dire les mots à sémantisme plein comme les verbes, les adjectifs, les substantifs, les démonstratifs et les adverbes portent un accent tonique. Ainsi, dans la phrase « *The workmen from Boston were leaving* », ce sont les mots « *workmen* », « *Boston* » et « *leaving* » qui sont frappés par l'accent. L'accentuation affectant les mots n'empêche bien sûr pas leur intégration dans une unité de niveau supérieur et il convient de respecter la distinction faite par Faure (1948 : 101) entre l'accent syllabique (accent de mot) et l'accent de phrase (pivot) :

« Chaque mot du groupe garde en général son accent propre, cet accent syllabique se doublant d'un accent de phrase qui porte sur les mots les plus importants de la phrase, dont l'accent syllabique se trouve ainsi renforcé ».

² Delattre (1965) : statistiques sur 2000 syllabes utilisées dans chaque langue (1000 en textes narratifs et 1000 en passages dramatiques)

³ En métrique ancienne, l'iambe désigne un pied composé de deux syllabes, la première brève et la seconde longue (*Dictionnaire Hachette Encyclopédique* sur CD, 2000).

L'accentuation anglaise est réputée complexe, car il est difficile pour un apprenant de savoir où placer l'accent tonique dans les mots multisyllabiques. La place de l'accent y est tellement mobile que certains phonéticiens ont affirmé qu'elle était non prédictible. Cependant, on sait maintenant qu'elle est gouvernée par des règles générales (décrites entre autres par Faure, 1948 ; Chomsky et Halle, 1968 ; Guierre, 1984 ; Ginésy, 1995 ; Hirst, 1998). Ces règles évitent à l'apprenant de mémoriser le schéma accentuel de chaque mot à partir du dictionnaire.

Par exemple, la loi d'alternance pousse les anglais à alterner les syllabes accentuées et les syllabes atones. Dans les mots de plus de deux syllabes, si la première syllabe est accentuée, la troisième portera un léger accent (appelé accent secondaire). Pour les mots de deux syllabes, trop courts pour être soumis à la loi d'alternance, c'est la loi de valeur qui se montre déterminante : ils seront accentués de préférence sur leur radical (« *Work* » dans « *workmen* », « *lea* » dans « *leaving* »).

Les statistiques données par Delattre (1965) montrent une tendance à l'accentuation de la syllabe initiale. Faure (1948 : 90) explique que les anglais ont

« tendance à accentuer dans un mot la syllabe la plus chargée de sens, c'est la raison pour laquelle ils font, en général, tomber l'accent sur la syllabe radicale, à l'exclusion des éléments accessoires (préfixes, suffixes, etc...). Or, en anglais, cette syllabe du radical se trouve très souvent au début des mots ».

Dans la phrase « *The workmen of Boston were leaving* », c'est la syllabe initiale de chaque mot lexical qui porte l'accent, soit « *work* », « *Bos* » et « *lea* », ce qui engendre un rythme de type trochaïque⁴. La localisation de l'accent tonique en anglais est variable, affectant de préférence la syllabe initiale des mots lexicaux. Les anglais sont donc habitués à un rythme dont

⁴ En métrique ancienne, le trochée ou le chorée désigne un pied composé de deux syllabes, une longue et une brève

l'élément acoustiquement fort est placé sur la syllabe initiale du mot lexical.

2.1.2. Différence de réalisation acoustique de l'accent

L'accent en français se matérialise avant tout grâce au paramètre acoustique de durée, les syllabes accentuées étant en moyenne deux fois plus longues que les syllabes atones. Pour cette raison et puisque cet accent est localisé en finale de groupe rythmique, on parle d'allongement final. La durée de la syllabe constitue le paramètre le plus important dans la perception de l'accentuation. Les écarts d'intensité sont en effet peu importants et, en ce qui concerne la hauteur, on ne constate aucune rupture mélodique, mais plutôt un « glissando », c'est-à-dire une variation de fréquence au cours de la production de la voyelle (Delattre, 1965).

(*Dictionnaire Hachette Encyclopédique* sur CD, 2000).

En anglais, l'accent se réalise avant tout grâce aux paramètres de hauteur et d'intensité : la syllabe accentuée est beaucoup plus forte, plus audible que les autres ([+intensité]), au détriment des autres syllabes qui deviennent presque inaudibles et qui disparaissent même. Mais surtout, la syllabe accentuée est prononcée sur une note différente des autres, ce qui crée une rupture tonale ([+mélodique]). Et selon Faure (1962 : 125), c'est ce paramètre qui se montre déterminant dans la mise en relief des syllabes accentuées en anglais :

« une syllabe réputée accentuée doit, à peu près toujours, l'essentiel de son relief acoustique au contraste mélodique qui l'oppose aux syllabes inaccentuées ».

Il pense que les différences de hauteur sont « beaucoup plus efficaces que les variations de durée et surtout d'intensité » (Faure, 1962 : 72). On parle d'ailleurs de « pitch accent ». Pour lui, l'augmentation d'intensité et l'augmentation de durée sont dépendantes de la variation de hauteur.

« L'élément essentiel de ce rythme, c'est-à-dire celui qui donne aux syllabes « fortes » leur plus haut degré de perceptibilité, est un facteur mélodique, généralement associé à des variations correspondantes d'intensité et de durée, dont le rôle nous paraît être beaucoup moins importants qu'on semble l'avoir pensé jusqu'ici ».

2.1.3. Différence de rythmique

2.1.3.1. La régularité syllabique du français

Nous avons vu que la syllabe finale de groupe rythmique est en moyenne deux fois plus longue que les syllabes non finales. Les syllabes atones ont donc toutes approximativement la même durée (du point de vue perceptif), ce qui donne une impression de régularité syllabique au français, d'isochronie. Pour cette raison, Pike (1946) parle de langue « syllable-timed ». C'est d'ailleurs le nombre de syllabes qui fonde notre versification, et non l'accentuation : seul est

retenu le nombre de syllabes ou pieds comme critère métrique. Ainsi, même avec un débit rapide, le français garde sa netteté articulatoire et individualise chaque syllabe. A cause de la place de l'accent en finale, l'articulation est très tendue, car l'élément important est toujours à venir. De facto, depuis Wenk et Wioland (1982), on parle maintenant plutôt de langue « trailer timed », c'est-à-dire à remorque (l'élément important figurant en arrière).

En fait, l'isochronie en français n'est pas si parfaite et il semble qu'il y aurait un allongement progressif des syllabes inaccentuées plus on se rapproche de la syllabe accentuée finale (Pasdeloup, 1990 : 224-225)

« La durée des syllabes inaccentuées tend à progresser jusqu'à la syllabe accentuée dans un mouvement général de ralentissement progressif ».

Wenk et Wioland (1982) se demandent cependant si c'est bien la répétition des syllabes successives qui donne sa spécificité rythmique au français. Le rythme ne serait-il pas créé par le retour de l'accent plutôt que par la régularité des syllabes atones ? S'ils ont raison, le français serait alors plutôt une langue à rythmicité accentuelle, et ils invoquent comme arguments contre l'isochronie du français que les syllabes accentuées n'ont pas la même longueur que les syllabes atones et que l'accent peut remplir une fonction distinctive dans le cadre du groupe rythmique. Mais le fait que l'accent puisse être distinctif en français ne va pas à l'encontre d'un écoulement régulier des syllabes atones. Quant aux syllabes accentuées, on ne peut pas dire que ce soit leur retour régulier qui assure la rythmique, à cause précisément de leur emplacement en finale. Leur apparition est conditionnée par la longueur du groupe rythmique et ne peut en conséquence pas être régulière. Il suffit d'essayer de reproduire l'écoulement des syllabes en frappant un crayon pour se rendre compte que la rythmicité syllabique est bien une réalité. Au lieu de classer les langues dans deux groupes opposés, « syllable timed » ou « stress timed », Dauer (1987) suggère que les différentes langues occupent différentes positions le long d'un axe allant du « syllable-timed » au « stress-timed » : l'emplacement de chaque langue le long de cet axe dépendrait dans chaque cas de la satisfaction d'un certain nombre de critères.

Depuis les travaux de Fonagy (1980) sur l'accent probabilitaire, on sait qu'une description de la rythmique du français contemporain ne peut plus se limiter à l'accent final et à la rythmicité syllabique. En effet, il existe une autre proéminence qui affecte de manière optionnelle la

syllabe initiale et de manière obligatoire la syllabe finale de chaque item lexical (Hirst et Di Cristo, 1984). Contrairement à l'accent final, cet accent ne se manifeste pas par un allongement de la durée, mais par un « pic » mélodique, beaucoup plus discret que celui affectant les syllabes accentuées en anglais.

Cet accent, qu'on appelle secondaire (Di Cristo, 1999 ; Astesano, 2001), permet de délimiter à droite une unité plus petite que le mot prosodique : le pied⁵ métrique, considéré par Di Cristo (1999 : 148) comme l'unité accentuelle minimale. Le système accentuel du français est donc beaucoup plus complexe, comme le montre Di Cristo (1999), qui tente d'établir le cadre accentuel du français contemporain. Il démontre notamment la présence de « *schèmes barytoniques* » et d'« *arcs accentuels* » (Di Cristo, 1999 : 185), dans lesquels seules les syllabes initiales et finales d'un groupe de mots reçoivent un accent (comme dans son exemple, « *les ONdées nuaGEUSES* »).

2.1.3.2. La régularité accentuelle de l'anglais

En revanche, la rythmique de l'anglais se caractérise par le retour régulier, non pas des syllabes atones, mais de l'accent tonique. Pour cette raison, Pike (1946 : 34) parle de langue « *stress-timed* », c'est-à-dire que

« les unités rythmiques ont tendance à se suivre de façon à ce que le laps de temps entre le début de leurs syllabes proéminentes soit uniforme ».

L'isochronie engendrée par le retour régulier de l'accent tonique est très similaire à un rythme de percussion, où les temps forts reviendraient à intervalles réguliers. C'est pourquoi, on parle de « *drum beat* », pour souligner la particularité de cette rythmique. L'important est de maintenir un intervalle analogue entre deux accents (1 seconde environ selon Faure, 1948 : 101). Peu importe le nombre de syllabes atones : si elles sont peu nombreuses, elles seront dilatées et si elles sont nombreuses, elles seront au contraire compressées ou même éliminées. Etant donné que chaque mot lexical est affecté d'un accent tonique, il en résulte que plus ces

⁵ Pied (Abercrombie, 1964) : notion empruntée à la poésie (« Chaque syllabe d'un vers dans la Métrique Ancienne ») ; « *dans notre acception, toute suite de syllabe terminée par un accent secondaire* » (Lacheret-Dujour et Baugendre, 1999 : 283).


mots sont nombreux au sein d'un énoncé, plus il y aura de syllabes accentuées et plus le rythme sera ralenti.

Cette organisation rythmique interfère directement avec le niveau segmental : à l'intérieur d'une syllabe compressée, la voyelle est réduite ; elle n'a pas le temps de se réaliser pleinement et son timbre se centralise pour s'approcher du [ə]⁶. Par ailleurs, l'inégalité de durée entre les syllabes introduit un rythme beaucoup plus complexe qu'en français.

Pour s'en rendre compte, il suffit d'essayer de le transcrire musicalement, comme l'a fait Faure (1948 : 102, voir figure n°54) pour l'énoncé « *Shut the door* ».



Figure n°54 : Faure (1948 : 102)

Il obtient ainsi la séquence musicale  suivante : avec comme référence la noire symbolisant la durée d'une syllabe accentuée.

« *Shut* » et « *door* » sont deux mots monosyllabiques à sémantisme plein, donc ils portent un accent. En revanche, le déterminant « *the* » n'est pas accentué, car c'est un mot outil. Pour indiquer la réduction de sa durée, Faure l'a transcrit à l'aide d'une double croche, c'est-à-dire par un élément trois fois plus court que la croche pointée initiale correspondant à la durée du mot « *shut* » et quatre fois plus court que la noire finale correspondant à la durée du mot « *door* ».

⁶ Voir paragraphe 3.1.1.1.

2.1.3.3. Régulation initiale versus régulation finale

En anglais, la syllabe accentuée tend à se placer en position initiale et elle est suivie de syllabes non accentuées. La régulation s'effectue donc à partir du début du groupe rythmique (régulation initiale). En français, au contraire, la syllabe accentuée se trouve en position finale et la régulation s'effectue en fonction de la fin du groupe rythmique (régulation finale). Cette différence de régulation engendre selon Wenk et Wioland (1982) un régime de tension musculaire différent dans les deux langues : en anglais, le régime de tension est décroissant car la syllabe accentuée apparaît en début de groupe rythmique ; les syllabes atones qui suivent sont prononcées de manière très relâchée et le timbre des voyelles se trouve réduit. En français au contraire, le régime de tension est croissant car la syllabe accentuée apparaît en fin de groupe rythmique. De ce fait, les syllabes atones, et en particulier le timbre de leurs noyaux, c'est-à-dire des voyelles, ne manifeste pas de réduction.

Par ailleurs, la montée mélodique marquant la syllabe accentuée peut se produire progressivement sans avoir recours à des mouvements brusques comme c'est le cas en anglais. La tension favorise un contrôle plus fin et indépendant des paramètres. En outre, l'augmentation de la durée ne s'accompagne pas d'augmentation de l'intensité.

2.1.3.4. Différences dans la durée d'acquisition de la rythmique anglaise et de la rythmique française

A cause de la mobilité de la place de l'accent en anglais, il est très difficile pour un apprenant d'inférer les règles qui lui permettront de bien placer les syllabes accentuées, surtout si celui-ci parle une langue où la place de l'accent est fixe, comme en français. Un francophone aura tendance à appliquer les règles d'accentuation de sa LM à l'anglais et donc à accentuer la syllabe finale. S'il n'est pas étonnant qu'un apprenant rencontre de telles difficultés dans l'apprentissage d'une langue étrangère, en raison des phénomènes de crible « prosodique » au niveau de la perception et d'automatismes articulatoires au niveau de la production, il semble que, lors de l'acquisition de leur LM respective, le « coût cognitif » soit beaucoup plus élevé pour les bébés anglais que pour les bébés français. Les travaux de Konopczynski (1990) montrent que l'allongement final est en place chez les bébés francophones, et ce de manière stable, dès l'âge de 16 mois. En effet l'acquisition du cadre rythmique est facilitée par le fait que la position de l'accent primaire est entièrement prédictible en français. Curieusement, le babillage des petits anglophones est également marqué par un allongement final pendant une certaine période, bien qu'il soit moins stable qu'en français. Konopczynski (1991 : 333) explique ces ressemblances par le fait que :

« l'enfant acquiert d'abord un rythme de type « universel » avant d'acquérir la structuration accentuo-temporelle propre à sa langue maternelle ».

Il se démarque ensuite graduellement de ce rythme « naturel » pour se rapprocher

progressivement de la rythmicité accentuelle de l'anglais, qui ne sera pas complètement mise en place qu'entre l'âge de 3 ans et 3 ans et demi. Ainsi, un bébé anglais met beaucoup plus de temps à maîtriser le schéma accentuel de sa LM, sans doute à cause d'une plus grande complexité. En effet, l'acquisition du rythme est généralement plus longue dans les langues à accent mobile, car la place de l'accent est moins facilement prédictible que dans les langues à accent fixe ou quasi fixe.

2.1.4. Accent de groupe vs accent de mot

2.1.4.1. Valeur de l'accent en français

En français, les mots perdent leur individualité au profit du groupe rythmique⁷ (nous préférons parler de groupe rythmique, car ce terme est plus simple à comprendre pour un apprenant).

« La langue française n'est pas une langue à accent de mot : cela veut dire que les différentes syllabes d'un mot ne sont pas hiérarchisées par une relation accentuelle » (Lhote, 1995 : 140).

Mackay (1987) définit l'accent de mot comme la proéminence qui affecte une ou plusieurs syllabes d'un mot multisyllabique. En effet, qu'est-ce qui peut indiquer à l'auditeur que le groupe « *La voiture jaune* » est composé de deux (« *La voiturejaune* ») ou de trois mots (« *La voiture jaune* ») ? Llorca (1987 : 54) souligne que la « *la spécificité du mot n'apparaît qu'à l'écrit et dans la conceptualisation* ». Une unité rythmique correspond en français à une unité de sens et n'est pas facilement identifiable pour des auditeurs non natifs, en particulier ceux dont la LM est une langue à accent de mots. En effet, alors qu'en anglais les mots sont nettement différenciés, l'interdépendance établie au niveau sonore entre les unités en français favorise leur interdépendance au niveau sémantique et réduit le poids de chacun de ses constituants (Llorca, 1987 : 53). Llorca (1987 : 52) donne l'exemple suivant : « *Il joue du violon* » prononcé d'une seule traite fonctionne comme unité simple équivalent à « *Il violonne* ». Les deux segments perdent donc quelques caractéristiques propres au profit d'une configuration globale qui les unit. Au contraire, si on dit « *Il joue / du violon* », en marquant une pause entre les deux, on attribue à chaque segment une configuration sonore propre. Pour

⁷ Lacheret-Dujour et Beaugendre (1999 : 45) définissent le groupe rythmique (ou groupe accentuel) comme un groupe « *correspondant à une suite de syllabes dont la dernière est marquée par un accent primaire* ». Di Cristo (1999 : 148) préfère le terme de mot prosodique qu'il définit comme le « *domaine de projection de l'accent final primaire* ».

préserver l'identité des deux segments, le locuteur oppose à l'enchaînement filé, la juxtaposition de deux signifiants distincts (« *Il joue du violon, du piano et de la guitare* »). L'opposition de la forme liée et de la forme comportant une rupture possède un caractère phonologique. La fusion entre les mots réalisée à cause de leur proximité est un phénomène spécifique au français. Certes, Rossi (1985) pense qu'un accent léger intervient dans l'identification et la démarcation du mot en français (en affectant la dernière syllabe de chaque mot lexical).

Cet accent interne assumerait donc bien selon lui une fonction morphologique. Di Cristo (1999 : 192-193) pense également que

« le mot, qui constitue l'unité accentuable minimale en français est doté dans la représentation sous-jacente d'une proéminence initiale et d'une proéminence finale. En d'autres termes, nous posons l'hypothèse que l'organisation métrique des mots accentogènes est assujettie à un principe de base que nous appellerons principe de bipolarisation ».

Toutefois, il insiste sur le fait qu'on ne peut pas parler d'accent lexical : « *le français ne possède pas d'accent lexical et cette remarque ne souffre aucune exception* » (1999 : 192). En effet, l'accent en français n'est pas distinctif, dans le sens où il ne sert pas à opposer des mots, comme c'est le cas en anglais.

En revanche, dans les langues à accent de mots comme l'anglais, les mots conservent leurs schémas accentuels, c'est-à-dire leur individualité, même s'ils sont mis en contexte. Bien sûr, il y a une réorganisation globale des unités dans une unité de niveau supérieur, mais le schéma qui s'applique à chacun de ces mots garde sa spécificité, à savoir son patron accentuel. Comme le souligne Llorca (1987 : 55), ces langues sont « *des langues à mots avant d'être des langues à accent de mots* ». Un anglais peut ainsi segmenter facilement les mots dans sa langue, même ceux qu'il ne connaît pas : il est guidé par la proéminence qui les affecte. Concernant la

production, la place de l'accent est même plus importante que la réalisation des phonèmes, car l'accent permet d'identifier les syllabes accentuées, c'est-à-dire les plus importantes au niveau du sens (par exemple, la syllabe portant le radical). Or, un accent mal placé au sein d'un mot en perturbera gravement la compréhension..

2.1.4.2. Valeur phonologique de l'accent lexical en anglais

Puisque l'accent en français est un accent de groupe, il n'a pas de valeur distinctive au niveau lexical. Il ne sert pas à opposer des mots, comme c'est le cas en anglais pour les deux versions du mot « *abstract* », qui se distinguent uniquement par leur accentuation (**extrait sonore n°23**). Comme adjectif, « *abstract* » se traduit en français par « *abstrait* » et comme substantif, par « *résumé* ». La transcription phonétique de ce mot [^ʰæbstrækt] indique que l'accent porte sur la première syllabe et en effet, on remarque une rupture tonale sur la figure n°55 (à gauche), marquée par un ton mélodique ascendant (226-251 Hz), tandis que la syllabe atone est prononcée avec un ton beaucoup plus bas et descendant (188-143 Hz).

On parle de ton mélodique quand la hauteur varie au cours d'une même syllabe.

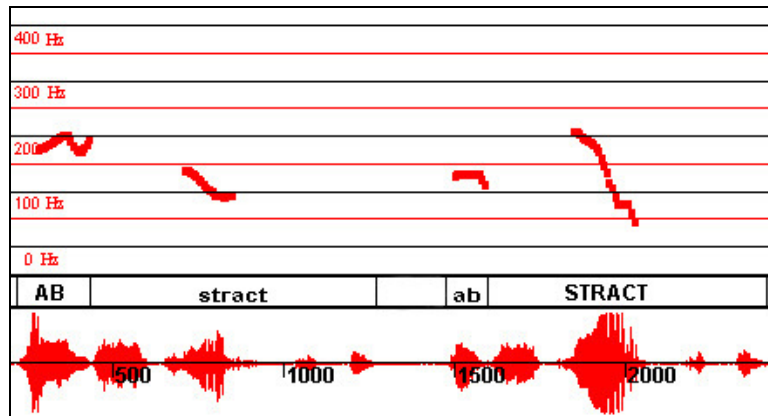


Figure n°55 : [*ˈæbstrækt*] et [*æbˈstrækt*] prononcés par la locutrice anglophone

Le mot suivant, le verbe « *abstract* » se traduit par « *soustraire, dérober qch à qn, détourner de l'argent* ». Selon sa transcription, l'accent porte sur la dernière syllabe [*æbˈstrækt*]. Effectivement, sur le tracé, on observe exactement l'inverse du premier mot : le décrochement mélodique se fait sur la seconde syllabe (258-126 Hz), alors que la première syllabe est « recto-tono » (178-180 Hz). Il apparaît encore une fois, au travers de cet exemple, que le paramètre mélodique est essentiel pour marquer l'accent de mot. On constate également une différence de durée substantielle entre la première syllabe de « *ˈABstract* » (221 ms) et la première syllabe de « *abˈSTRACT* » (126 ms), quasiment deux fois plus courte et une différence d'intensité (intensité plus grande sur la première syllabe dans « *ˈABstract* » et inversement pour « *abˈSTRACT* »). Mais selon Faure (1962 : 144),

*« la pression exercée sur notre jugement en matière d'accent par les éléments mélodiques, du fait de notre aptitude exceptionnelle à en saisir les plus petites nuances, fait que le seul changement de ton, à l'exception de toute altération de l'intensité et de la durée, suffit par exemple, à opposer « *ˈpermit* » au verbe « *perˈmit* » ».*

Le passage de l'un à l'autre mot provoque un changement de sens. Puisqu'en changeant la place

de l'accent au sein d'un mot, on change le sens de ce mot, l'accent a donc une valeur phonologique en anglais.

2.1.5. Langue non accentuée vs langue accentuée ?

La « faiblesse » acoustique de l'accentuation en français, sa régularité syllabique et son caractère non distinctif ont apporté de nombreux arguments à des chercheurs tels que Hjemsløv, Tøgeby et Pilch pour dire que le français était une langue sans accent.

Le synchronisme entre accent et intonation a été également avancé pour dire que l'accent n'avait pas de fonction propre, les syllabes accentuées étant également le point où le contour de hauteur est réalisé lors d'un groupe accentuel final dans une unité d'intonation. D'après Beckman (1992), cité par Di Cristo (1999 : 151) :

« les propriétés de l'organisation prosodique du français incitent à considérer l'accent final comme un indicateur de la limite du groupe rythmique (edge-marking). Il diffère donc fondamentalement, selon cet auteur, de l'accent de l'anglais et du suédois qui peut être interprété comme un indicateur de la tête du pied accentuel (head marking), complètement indépendant de tout effet d'allongement final de syntagme ». Selon Di Cristo, « ce point de vue tend à promouvoir l'idée que, en définitive, le français ne posséderait pas d'accent à proprement parler, mais seulement des indices démarcatifs des limites des mots ou des syntagmes ».

Ce point de vue est remis en question par Ladd (1996), qui pense au contraire que les syllabes finales de groupe rythmique sont bien en français des syllabes métriquement fortes. Ce point de vue est partagé par Di Cristo (1999 : 201) :

« l'accent initial, comme l'accent final peuvent être considérés comme les têtes métriques des pieds accentuels (ou unité rythmique minimale) en français ».

Par rapport à l'anglais, Vaissière pense que l'accent a bien une fonction en français puisqu'il permet de délimiter les groupes rythmiques les uns par rapport aux autres. Ainsi, elle parle de langue de frontière (par opposition à une langue accentuelle comme l'anglais). L'accent a donc

une valeur démarcative. Pour Llorca (1987), il a en outre une fonction unitaire, dans le sens où il assure la cohésion des éléments du groupe.

2.2. Intonation

2.2.1. L'accent de phrase

2.2.1.1. L'accent de phrase en anglais

Faure (1962 : 125) distingue l'accent de mot,

« celui qui détache une syllabe donnée dans un mot isolé » de l'accent de phrase qui permet « le renforcement de l'accent de mot, ou de l'accentuation volontaire d'une monosyllabe habituellement inaccentuée ».

Cet accent de phrase porte sur le mot le plus important de l'énoncé et dépend de l'information que le locuteur veut faire passer à son interlocuteur.

« On pourrait parler ici d'accent d'insistance : un mot reçoit un supplément d'intensité au détriment des autres mots de la phrase. Tout énoncé anglais peut ainsi recevoir presque autant de sens différents qu'il comporte de mots » (Keen et Adamczewski, 1973 : 189).

Pour désigner l'accent de phrase, on parle également de focus, notion définie par Rossi (1999 : 205) comme l'« *item porté au premier plan par l'accent de focalisation* ». Selon Hirst (1998 : 66), ce terme a remplacé le terme plus traditionnel d'emphase, bien que les deux termes ne désignent pas tout-à-fait la même réalité : tandis que « *focus* » est une notion syntagmatique, qui désigne un élément de la séquence, « *emphase* » est une notion paradigmatique, qui désigne chaque élément comme emphatique ou non emphatique. Cette notion s'applique mieux à l'anglais britannique selon Hirst.

Dans la phrase « *She was in the bedroom* », l'information apportée (le rhème) concerne le lieu où la personne dont on parle se trouve, c'est-à-dire la chambre à coucher. En revanche, dans la phrase « *She wasn't in the bedroom* », l'information primordiale n'est pas le lieu où la personne

se trouve, mais le fait qu'elle ne s'y trouve plus. En conséquence, c'est le mot « *wasn't* » qui porte l'accent de phrase. L'accentuation doit en effet être limitée au point nouveau portée par le message. L'accent de phrase se matérialise par un pic ou un sommet intonatif : dans la phrase « *She was in the bedroom* », ce pic correspond à la syllabe « *bed* » (figure n°56 et **extrait sonore n°24**) et dans la phrase « *She wasn't in the bedroom* », il correspond à la syllabe « *wasn't* » (figure n°57 et **extrait sonore n°25**).

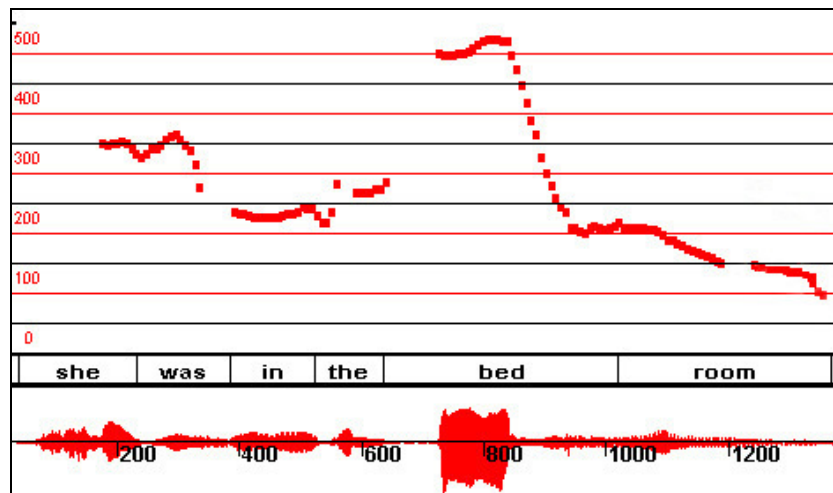


Figure n°56 : courbe mélodique de la phrase « *She was in the bedroom* » prononcée par une locutrice anglophone (logiciel Winsnoori).

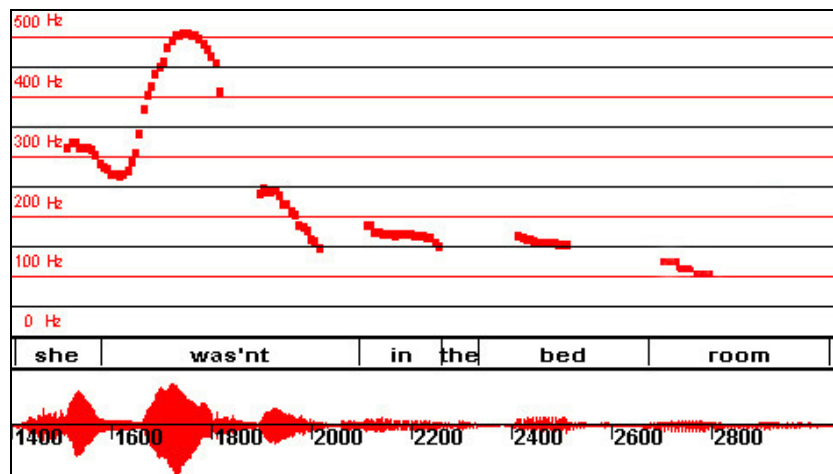


Figure n°57 : courbe mélodique de la phrase « *She wasn't in the bedroom* » prononcée par une locutrice anglophone (logiciel Winsnoori).

On constate également un accroissement de la durée de ces syllabes, mais l'indice le plus

saillant semble bien être l'indice mélodique.

2.2.1.2. L'accent de phrase en français

L'accent de phrase se manifeste différemment en anglais et en français, puisque nous avons vu qu'en anglais, il affecte le mot le plus important, le mot sur lequel il faut insister. Il se rapproche donc plutôt de la notion d'accent d'insistance en français. Mais, contrairement à l'anglais, dans un contexte neutre, le français ne privilégie pas une syllabe se trouvant à l'intérieur du groupe rythmique, mais bien la syllabe finale du groupe rythmique, qui sera marquée par un allongement de la durée et un glissando final. On observe pourtant un sommet intonatif, qui correspond au point le plus haut de l'énoncé.

Sa localisation n'est cependant pas dictée par le besoin de mettre en relief un mot particulier au sein de l'énoncé, mais bien par la structure de cet énoncé.

2.2.1.3. La notion de pivot

Le pivot ou le centre intonatif de l'énoncé est le point central de l'énoncé en anglais, son centre intonatif, appelé également « *nucleus* ». Le pivot se place sur le dernier ton accentué, c'est-à-dire sur la « *syllabe accentuée du dernier mot pleinement accentué de l'énoncé* » (Ginésy, 1995 : 204). Ce ton est en général mélodique ou constitué d'une combinaison de tons statiques suivie d'une remontée de la fréquence fondamentale sur les syllabes atones suivantes.

Cette suite de syllabes atones localisée après le pivot constitue la queue de l'énoncé ou « *tail* » en anglais. L'énoncé doit comporter un pivot pour être complet, sinon, l'auditeur a l'impression qu'il a été interrompu, qu'il est resté en suspens. La première syllabe accentuée de l'énoncé est appelée tonique de tête (ou « *head* ») en anglais. Elle doit supporter une rupture mélodique suffisamment importante pour être distinguée des syllabes atones voisines. Les syllabes atones précédant la tête constituent ce qu'on appelle la « *pre-head* » en anglais. Toutes les syllabes qui se trouvent entre la tonique de tête et le pivot constituent le corps de l'énoncé (« *body* », en anglais). Leur nombre est indéterminé, il peut même ne pas y en avoir.

2.2.2. Les tons mélodiques

2.2.2.1. La richesse musicale de l'anglais

Un ton mélodique consiste en une variation de la hauteur sur la même syllabe, associée à une augmentation de la durée et accessoirement, à un accroissement de l'intensité. La variation de hauteur est un élément essentiel pour assurer la mise en avant des syllabes accentuées en anglais. Ainsi, les tons mélodiques sont généralement utilisés pour mettre en relief le pivot. Il existe cinq tons mélodiques différents en anglais, dont trois qu'on pourrait qualifier de simples, car

leur mouvement mélodique consiste en une montée ou une descente. On les désigne par leur forme mélodique, c'est-à-dire « petite chute », « grande chute » (« High Jump ») et « montée » (voir figure n°58). La petite chute traduit le caractère final d'un énoncé et l'objectivité du locuteur par rapport à cet énoncé (ton de l'assertion).

Mais, lorsque cette chute est mise en relief par une pente beaucoup plus marquée, ce n'est plus l'objectivité qui est exprimée, mais l'implication ou la volonté d'insistance du locuteur : la grande chute est en effet utilisée pour l'exclamation, l'ordre, l'indignation... A l'inverse du pattern de descente, celui de montée traduit le caractère non final, suspensif d'un énoncé (inachèvement d'une phrase, appel à la réponse à une question de type oui ou non). Il est également utilisé pour exprimer un sous-entendu ou l'implication du locuteur par rapport à son interlocuteur (amitié et chaleur humaine, encouragement, politesse, courtoisie).

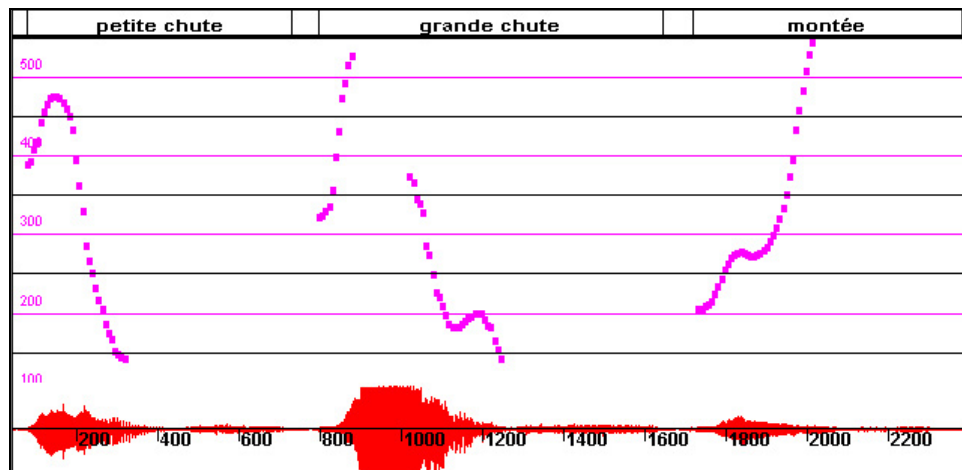


Figure n°58 : les 3 tons mélodiques simples de l'anglais sur le mot « Yes »
prononcée par une locutrice anglophone (logiciel Winsnoori)

La combinaison de ces traits donne naissance à des tons plus complexes, que sont le ton creusé (« *fall-rise* ») et le ton circonflexe (« *rise-fall* ») (voir figure n°59). Ces tons complexes se trouvent toujours en position terminale d'énoncé. Le ton creusé « *traduit d'abord une affirmation (au travers de la chute) suivie de réserves* » (Ginésy, 1995 : 213). Alors que la partie descendante est assertive, cette affirmation est nuancée par la remontée (implication, sous-entendu, non-dit). A l'inverse, le ton circonflexe

« marque d'abord une incertitude de la part du locuteur (c'est la montée), incertitude vite effacée par l'affirmation portée par la chute. Ce ton est plus rare que le fall-rise en anglais et marque

souvent l'irritation ou l'impatience ou l'ironie du locuteur. Mais il peut aussi marquer une chaleur, un intérêt véritables pour le sujet ou l'interlocuteur » (Ginésy, 1995 : 213).

L'**extrait sonore n°26** nous donne la suite des cinq tons mélodiques dans l'ordre où nous les avons présentés, réalisés par une locutrice anglophone.

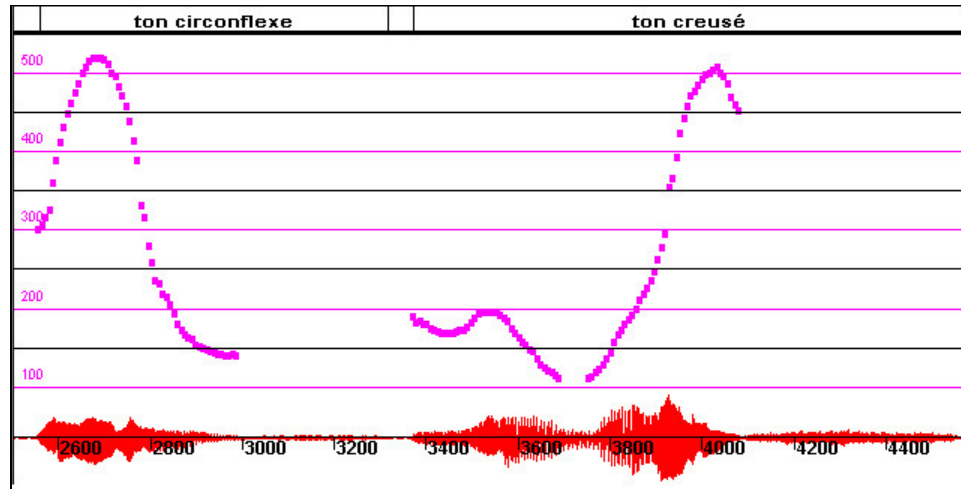


Figure n°59 : les 2 tons mélodiques complexes de l'anglais sur le mot « Yes » prononcée de manière très emphatique par une locutrice anglophone (logiciel Winsnoori).

2.2.2.2. Des tons mélodiques en français ?

Les locuteurs francophones modulent très rarement la hauteur sur une même syllabe et ont le plus souvent recours aux tons statiques, à cause de la grande tension articulatoire qui caractérise leur langue maternelle. Cette tension articulatoire empêche la modulation de la hauteur sur les syllabes, en plus de contraindre à la régularité syllabique. On observe ainsi une succession de tons statiques (paliers horizontaux) de même durée : pour une oreille étrangère, l'effet produit est proche du staccato en musique et ressemble au bruit d'une mitrailleuse. Pour cette raison, les anglais parlent de « machine gun effect » (effet mitrailleuse). Faure (1962) affirme qu'« on peut représenter toutes les nuances mélodiques en se limitant à quelques tons statiques et à trois tons mélodiques » :

Un ton mélodique descendant, comme dans « *Viens avec nous* ».

Un ton mélodique descendant, comme dans « *C’est absolument ridicule !* ».

Un ton mélodique ascendant, comme dans « *Il y a de l’argent !* ».

Il n’existe donc pas de tons complexes. Mais, en réalité, l’utilisation des tons mélodiques en français est surtout liée à l’accent d’insistance, qui marque l’emphase, l’affectivité (c’est le cas dans les deux dernières phrases). Le procédé de transcription adopté par Faure (1948 : 119) révèle de manière flagrante la différence entre les deux langues (différence de durée affectant les syllabes et tons mélodiques vs tons statiques) :

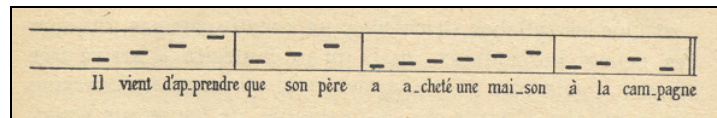


Figure n°60 : transcription de la phrase « Il vient d’apprendre que son père a acheté une maison à la campagne » (Faure, 1948 : 117).

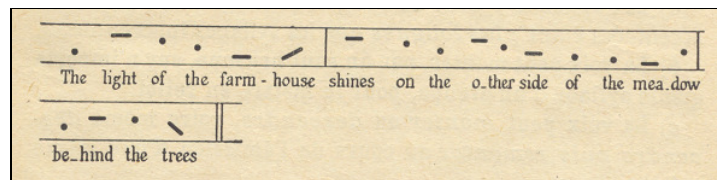


Figure n°61 : transcription de la phrase « The light of the farmhouse shines on the other side of the meadow behind the trees » (Faure, 1948 : 117).

2.2.3. Description comparée de l’intonation déclarative en anglais et en français :

Nous nous sommes limitée à la description du pattern d’intonation déclarative car le niveau en anglais des enfants est débutant. Dans cette étude, nous n’avons donc pas besoin de connaître toutes les ressources du système intonatif de l’anglais (se reporter aux descriptions de Faure (1962), Ginésy (1995), Hirst (1998)).

2.2.3.1. Finalité

En anglais, le schéma qui correspond à l’assertion est appelé « *glide down* ». Par exemple, dans la phrase « *I didn’t expect you to know her* » (phrase empruntée à Ginésy, 1995 : 208), le pivot «

know » porte l'accent mélodique (**extrait sonore n°27**). Le ton mélodique associé est la petite chute (voir figure n°62) : il exprime la neutralité et traduit souvent la certitude ou l'achèvement. C'est une intonation terminale utilisée dans l'affirmation simple, pour un contenu informatif, sans sous entendu, ni emphase.

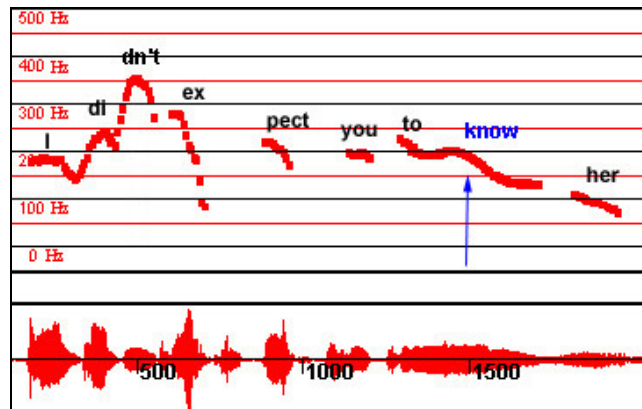


Figure n°62 : Modalité neutre en anglais : « petite chute » produite sur le pivot « know »

Le sommet intonatif de l'énoncé porte sur la première syllabe accentuée (« *didn't* » dans notre exemple), c'est-à-dire la tête de l'énoncé, mais aussi sur le mot qui porte l'information la plus importante. A sa suite, la ligne mélodique s'abaisse progressivement (« *glide* ») jusqu'au pivot, matérialisé par une petite chute (« *down* ») (effet global de déclinaison).

En français, la finalité est également traduite par une courbe mélodique descendante, mais sa structuration est différente.

« La phrase énonciative, objective, contient un sommet intonatif (le point le plus haut de la phrase autour duquel s'articulent les différents groupes qui la composent. La ligne générale de l'intonation est descendante, mais chaque groupe contient l'intonation suivante : ligne presque plate (légèrement montante par paliers statiques) et fin de groupe montante. [...] Les montées françaises ne se font pas à l'aide de tons mélodiques, mais par le biais d'un rejet vers le haut de la syllabe accentuée » (Ginésy, 1995 : 223).

Dans la phrase correspondante à celle de l'anglais, « *Je ne savais pas que tu la connaissais* », on observe un pattern circonflexe, constitué d'une montée et d'une descente (voir figure n°62). Selon Hirst et Di Cristo (1998), le pattern de hauteur typique d'une phrase déclarative finit par sur une hauteur basse finale, contenant une montée à la fin de chaque groupe accentuel.

2.2.3.2. Continuation

Nous avons vu dans le paragraphe 2.1.1.3. (la régularité syllabique du français) qu'il existait une autre proéminence en français, en dehors de l'accent final, c'est l'accent secondaire. Cet accent est marqué non point par l'allongement de la durée syllabique, mais par un indice mélodique et plus précisément par un sommet mélodique. Il permet de délimiter à droite une unité plus petite que le groupe rythmique, le pied métrique. Ainsi, notre phrase est constituée d'un seul groupe rythmique « *Je ne pensais pas que tu la connaissais* » mais de deux pieds

métriques. Le premier pied métrique, « *Je ne pensais pas* » est marqué par une ligne mélodique légèrement concave, suivie d'une montée sur la dernière syllabe « *pas* », qui correspond à l'accent secondaire. Le second pied métrique, « *que tu la connaissais* » repart après un décrochement et se termine par un mouvement mélodique descendant, qui traduit l'assertion (Lhote, 1995 : 139).

C'est donc le mouvement mélodique qui marque la continuation en français, ce qu'avait relevé Delattre dès 1961 : l'intonation de la dernière syllabe d'un groupe de sens indique la continuation en prenant essentiellement la forme ascendante en français. Dans le cas où l'énoncé n'est composé que d'un seul groupe accentuel, l'intonation est globalement descendante.

« Ce qui caractérise le plus le français, c'est la descente graduelle du deuxième groupe et l'inversion de pente entre les deux groupes » (Lhote, 1995 : 140).

Par ailleurs, le sommet intonatif de l'énoncé correspond à la dernière syllabe du dernier groupe accentuel interne de l'énoncé. Dans notre exemple, ce sommet correspond donc à la syllabe « *pas* ».

En anglais, le pattern mélodique qui exprime la continuation constitue une sorte de « négatif » de celui du français : la syllabe accentuée se trouvant en position initiale, la régulation s'effectue donc à partir du début du groupe rythmique (régulation initiale). Toute la tension nécessaire à la montée mélodique qui caractérise la syllabe accentuée en anglais se relâche lors du passage à la syllabe suivante. Après la rapidité de ce saut mélodique, la descente est toute aussi brusque. Le contour mélodique est donc descendant et de forme convexe : il ressemble à un S renversé (Delattre, 1965). Sur la figure n°63 (Delattre, 1965), on peut voir que ce pattern s'oppose à celui du français, ce dernier étant marqué par un contour montant de forme concave, qui aboutit sur un plateau élevé à la dernière syllabe.

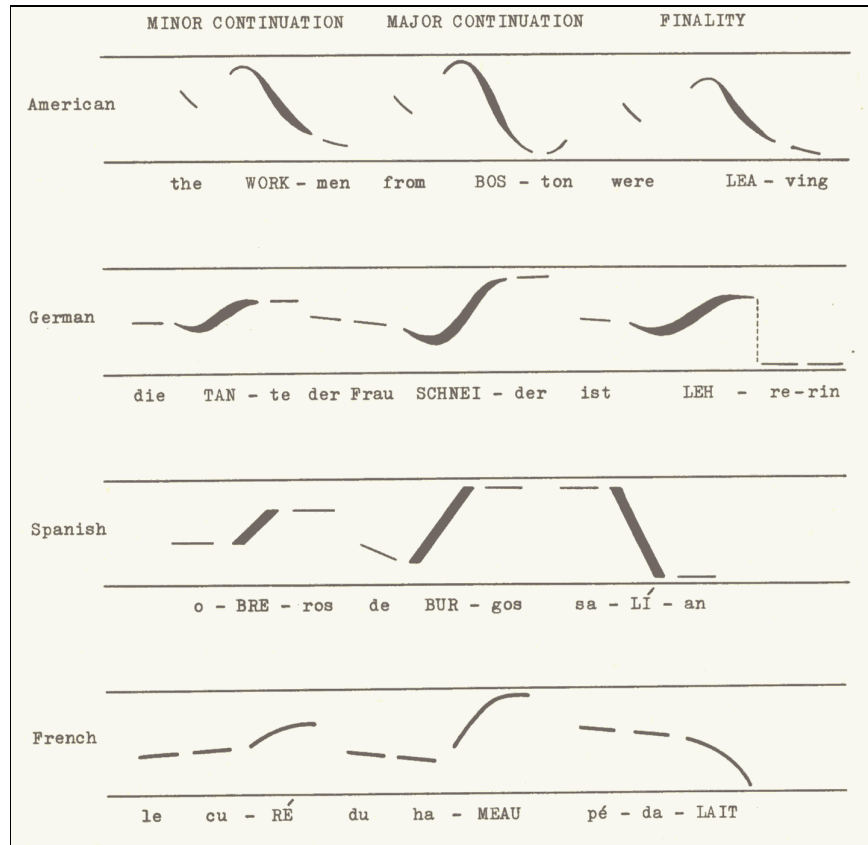


Figure n°63 : intonation caractérisant la continuation et la finalité en français et en anglais américain, Delattre (1965 : 25)

Selon Hirst et Di Cristo (1998), le paramètre de localisation de la tête (la syllabe métriquement la plus forte) dans le pied métrique est déterminant et permet de distinguer l'anglais du français sur la base d'une opposition : tête à gauche / tête à droite, respectivement pour l'anglais et le français.

Un tableau récapitulatif des principales différences prosodiques entre le français et l'anglais figure dans le Volume II, annexe n°3.1.

3. Description comparée du système vocalique de l'anglais et du système vocalique du français

3.1. Caractéristiques générales

3.1.1. Conséquences de la différence d'organisation rythmique sur le niveau segmental

Nous avons vu que l'anglais possédait une organisation rythmique radicalement différente de celle du français. Or, ce cadre rythmique influence directement la prononciation du niveau segmental, et plus particulièrement celle des voyelles.

En anglais, l'organisation rythmique se fait autour du retour régulier de l'accent tonique. La conséquence de cette rythmicité sur la prononciation des voyelles anglaises est déterminante. Les syllabes accentuées sont en effet beaucoup plus audibles que les syllabes atones, car elles sont mises en relief par les paramètres acoustiques que sont l'amplitude (leur intensité est plus forte), la fréquence fondamentale (correspondant à un ton mélodique ; à cause de ces sommets mélodiques, Faure (1962) va jusqu'à qualifier l'anglais, de façon volontairement provocante, de langue à tons) et quelquefois la durée (allongement, notamment en A.N.A.). Le timbre des voyelles « toniques » a le temps de se réaliser pleinement puisque ces voyelles constituent le noyau des syllabes accentuées.

En revanche, ce n'est pas le cas du timbre des voyelles « atones », constituant le noyau des syllabes non accentuées. N'ayant pas le temps de se réaliser, elles ont tendance à perdre leur articulation pleine et à se centraliser. Selon Delattre (1965), cette position centrale constitue la caractéristique distinctive de l'anglais, puisque 90 % des voyelles non accentuées sont proches de /ə/ et que la proportion de ces voyelles parmi toutes les voyelles de l'anglais est de 61 %. En résumé, le timbre d'une voyelle accentuée sera donc toujours pleinement réalisé tandis que celui d'une voyelle atone se centralisera ; une voyelle accentuée sera par ailleurs plus longue, plus intense et plus haute qu'une voyelle atone.

L'organisation rythmique de l'anglais a donc d'énormes conséquences sur la prononciation des voyelles. Elle est d'autant plus difficile à maîtriser pour un francophone que l'organisation rythmique de sa LM est opposée à celle de l'anglais. En français, la régularité n'est pas fondée sur le retour régulier de l'accent, mais sur le retour régulier de la syllabe (cf. 2.1.3.1.). En raison de cette régularité syllabique, le timbre des voyelles du français est toujours clair. Il y a bien un

[ə] caduc, mais

« le schwa a un statut de phonème en français et son occurrence n'est pas une simple conséquence d'un autre phénomène tel que l'accent » (Ostiguy et al, 1999 : 151).

A cause du régime de tension croissante du français (voir 2.1.3.3.), l'articulation est très tendue, et empêche la diphtongaison, typique de l'anglais. La diphtongaison s'explique en effet par une mobilité des articulateurs pendant l'émission. Lorsque les cavités supra-glottiques changent de configuration, elles provoquent un changement de la fonction de transfert, et en conséquence un changement de la qualité du timbre.

Le phénomène de diphtongaison dépend donc fortement de la durée. Les voyelles longues, nous le verrons, ont tendance à se diphtonguer. Quant aux noyaux vocaliques complexes que sont les diphtongues et les triphongues, ils figurent parmi les noyaux les plus allongés. Enfin, l'anglais se caractérise par une certaine lenteur à ouvrir et fermer les articulateurs se traduisant au niveau acoustique par de longues transitions entre les cibles vocaliques, ce qui n'est pas le cas du français.

Les problèmes de centralisation, de durée et de diphtongaison sont donc très dépendants de l'organisation rythmique de la langue et on ne peut étudier le système vocalique d'une langue sans avoir au préalable établi son cadre rythmique.

3.1.2. Différences d'organisation de l'espace articulatoire

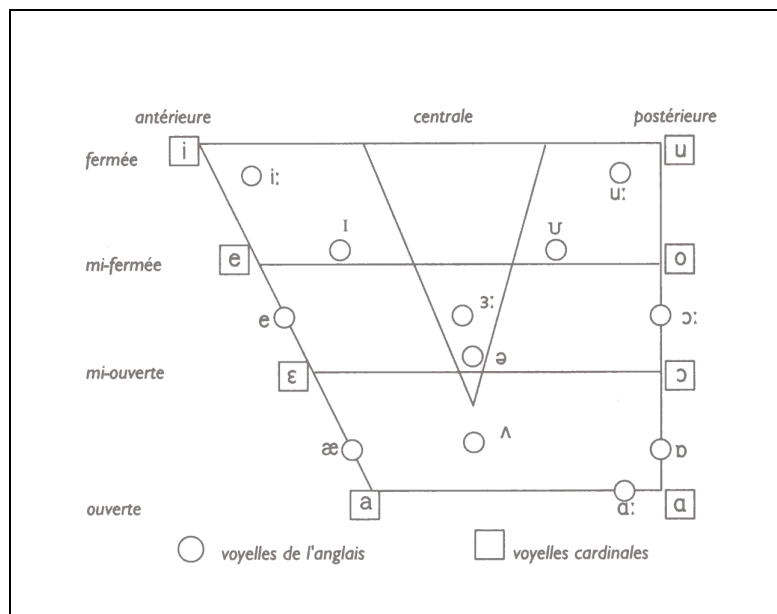


Figure n°64: trapèze vocalique comparatif des voyelles françaises et des monophongues anglaises (emprunté à Ginésy, 1995 : 15).

Un français qui va apprendre l'anglais sera confronté à un système vocalique beaucoup plus riche et plus complexe que celui de sa LM. En français, le système vocalique est composé de

seize monophthongues, dont quatre voyelles oralo-nasales.

En anglais britannique standard (désormais R.P.), on trouve vingt-cinq noyaux vocaliques, dont douze monophthongues (et parmi elles, cinq voyelles diphtonguées), huit diphtongues et cinq triphthongues.

A titre de comparaison, l'A.N.A. (désormais A.N.A.) ne possède que seize noyaux vocaliques (Lehiste et Peterson, 1960), dont dix monophthongues et six noyaux complexes (trois « glides » et trois diphtongues).

Au niveau de l'articulation, le français et l'anglais combinent différemment les traits articulatoires et leur système vocalique occupe différemment l'espace articulatoire (figure n°64, trapèze vocalique comparatif).

Le système vocalique du français, délimité à ses extrémités selon l'axe vertical par la voyelle fermée /i/ et la voyelle ouverte /a/ montre une grande amplitude d'aperture. En raison de leur « extrémité articulatoire », ces voyelles « cardinales » ont été utilisées comme points de référence par Daniel Jones pour établir des chartes vocaliques. En raison de l'asymétrie des organes articulatoires,

« la mâchoire s'écarte beaucoup plus pour les voyelles d'avant que pour les celles d'arrière puisqu'elle pivote sur un axe postérieur » (Léon, 1992 : 80).

Le français joue beaucoup de cette amplitude d'aperture car la plupart de ses voyelles sont articulées sur l'avant de la bouche : deux séries antérieures sur trois, soit dix voyelles sur seize (cette configuration donne au français une « *résonance vocalique claire* », Léon, 1992 : 82).

Au contraire, l'anglais se caractérise par une plus grande proportion de voyelles postérieures (deux séries sur trois, soit huit monophthongues sur douze) et donc une amplitude d'aperture plus réduite. Malgré cela, la proportion de voyelles dites ouvertes est plus importante qu'en français et le centre de gravité, plus bas (Delattre, 1965). On remarque également plus de centralisation (trois voyelles centralisées /əɜ:ʌ/).

Le français est dominé par l'articulation labiale, puisque deux tiers des voyelles sont labialisées, soit onze sur seize (Delattre, 1965). La proportion s'inverse en anglais où l'on ne trouve qu'une

série labialisé et pas de série antérieure labialisée, équivalente à la série /iyø/ en français. Par ailleurs, le degré de labialisation est beaucoup plus fort en français, les lèvres étant véritablement arrondies, alors qu'en anglais, on n'observe qu'une légère projection des lèvres.

Enfin, on trouve en français quatre phonèmes vocaliques très fortement nasalisés. On parle d'ailleurs de voyelles oralo-nasales, en raison d'un fort degré de couplage entre la cavité orale et la cavité nasale. En R.P, le trait de nasalité ne permet pas d'opposer des voyelles entre elles ; il n'est donc pas phonologique, mais phonétique. Pour cette raison, nous n'en ferons pas la description, car il ne peut pas y avoir d'interférences entre les deux langues à ce sujet.

3.1.3. Conséquences pour un apprenant francophone

3.1.3.1. Prédiction d'une typologie des fautes découlant de ces différences

L'anglais se démarque donc principalement du français par sa tendance à la diphtongaison, par l'utilisation qu'il fait de la durée et par son articulation plus postérieure, plus centralisée et moins labialisée. En fonction de ces différences, on peut prédire le type de fautes de prononciation qu'un francophone réalisera en anglais.

L'excès de tension articulatoire que le locuteur francophone transpose de sa LM à l'anglais est à l'origine de toute une série de fautes :

- elle va entraver la **tendance à la diphtongaison** des voyelles longues et la reproduction correcte des diphtongues et des triphthongues spécifiques de l'anglais ;
- elle va amener l'apprenant à produire des **transitions trop rapides entre les cibles vocaliques** qui composent une diphtongue, avec pour conséquence une tendance à produire une succession de plusieurs voyelles au timbre stable, plutôt qu'un glissement lent et progressif d'une cible vocalique à une autre. Or, une « vraie » diphtongue est perçue comme un glissement vers la seconde cible et non pas comme la succession de deux voyelles ;
- elle va l'amener à **raccourcir la durée des syllabes**, ce qui aggrave les deux défauts

précédents. En effet, il faut une certaine durée pour réaliser un son diphtongué ;

- elle va l'empêcher de reproduire correctement la **réduction vocalique qui caractérise les voyelles non accentuées de l'anglais**, car celui-ci se caractérise par un grand relâchement articulaire.

Les différences articulatoires entre le système vocalique du français et de l'anglais amènent l'apprenant à faire les erreurs suivantes :

- Le français étant une langue dont l'articulation est antérieure, l'apprenant aura le réflexe de **produire les voyelles de l'anglais trop en avant**, surtout si leur timbre est proche de certaines voyelles françaises plus antérieures. C'est le cas de [ɪ] par exemple qui aura tendance à être prononcé « à la française », comme [i] ou [e].
- Les voyelles françaises étant pour la plupart labialisées, l'apprenant aura tendance à **trop labialiser les voyelles anglaises** dont le timbre se rapproche de ces voyelles labialisées. Par exemple, la voyelle [ʊ] de « *book* », qui n'est que faiblement arrondie en anglais, aura tendance à être prononcée comme [u] français, et à être en conséquence, fortement labialisée.
- Les problèmes d'interférences naissent lorsque le timbre des sons non natifs est proche de celui de certaines voyelles de la LM (crible phonologique). Dans ce cas, elles sont analysées comme des exemplaires de sons natifs, éventuellement légèrement déviants. Bien que l'anglais possède un système vocalique très différent de celui du français, certaines de ses voyelles non diphtonguées ont un timbre proche de celui de certaines voyelles du français. Les catégories natives et non natives se chevauchant partiellement, il risque en conséquence d'y avoir **beaucoup d'interférences au niveau perceptif**.

3.1.3.2. Paramètres importants nécessaires à la description de l'anglais pour un apprenant

francophone**3.1.3.2.1. Durée**

Nous venons de voir que l'excès de tension articulatoire allait, entre autres choses, amener l'apprenant francophone à produire des noyaux vocaliques trop courts. Or, la durée des noyaux vocaliques est un facteur très important pour décrire les différents noyaux vocaliques de l'anglais. Ainsi, Lehiste et Peterson (1960) ont décrit le système vocalique de l'A.N.A.⁸ en fondant en grande partie leur classement sur ce critère de durée. Ils distinguent parmi les monophthongues, les noyaux simples et courts (monophthongues d'une durée de 180 à 200 ms), des noyaux simples et longs (monophthongues d'une durée, de 240 à 260 ms).

⁸ Analyse acoustique de 1263 monophthongues (70 mots prononcés par un locuteur de l'A.N.A.).

Il est probable que la différence entre ces deux catégories n'apparaîtra pas dans les productions d'un locuteur francophone, car les noyaux longs seront trop réduits. Par ailleurs, l'effet de crible phonologique risque de renforcer cette tendance. En effet, les voyelles anglaises auront tendance à être assimilées aux voyelles françaises de timbre « correspondant ». Par exemple le [i:] anglais aura tendance à être assimilé à la voyelle française [i] et produit comme un [i]. Or, la durée de cette voyelle est beaucoup plus courte que celle de son homologue anglais. Il apparaît donc pertinent de relever la durée absolue et la durée relative (par rapport à la durée totale de la syllabe) des noyaux vocaliques produits par un locuteur francophone en anglais. Elle fournira en effet des informations sur les effets du crible prosodique (et en particulier rythmique) et sur les effets du crible phonologique.

3.1.3.2.2. Structure interne du noyau vocalique

En plus de cette différence de durée intrinsèque qui affecte le noyau vocalique, les travaux de Lehiste et Peterson (1960) révèlent l'importance de la durée relative des événements composant ce noyau. Toute voyelle est constituée théoriquement d'une tête (« *on glide* »), d'une partie stable (« *target* ») et d'une queue (« *off glide* »), la relation entre les trois événements lors de la durée totale du noyau vocalique restant constante, indépendamment de la consonne suivante. Selon les auteurs, le temps mis pour produire un changement total de timbre (durée relative par rapport à la durée totale de la syllabe) pourrait constituer un critère pertinent pour dresser une typologie fine des différents types de noyaux vocaliques en anglais (1960 : 292). Or l'excès de tension articuloire qui caractérise les productions francophones va justement affecter la durée du changement de timbre et bouleverser la structure interne du noyau vocalique. En effet, la durée des phases de transition sera réduite, par rapport à la durée de la zone de stabilité de timbre, qui elle sera démesurée. Cette trop grande stabilité constitue un obstacle à la production des nombreux sons diphtongués de l'anglais, ainsi que des voyelles longues.

Il paraît donc pertinent d'étudier la structure interne des noyaux vocaliques produits par les locuteurs francophones en anglais. Ce critère pourrait servir d'indice pour évaluer la progression de leur maîtrise de la prononciation.

3.1.3.2.3. Timbre

Enfin, il est important de déterminer le timbre des noyaux vocaliques produits par les apprenants francophones en anglais pour savoir s'ils sont soumis au phénomène de crible phonologique et si oui, pour quels types de sons. Mais, comment mesurer le timbre d'une voyelle ? Traditionnellement, les chercheurs pensent que l'information essentielle pour déterminer la qualité des voyelles réside dans la localisation des maxima spectraux correspondant aux trois premiers formants (Lonchamp, 1983 ; Schwartz, 1987). Cette localisation correspond à la notion de cible vocalique. D'après Lehiste et Peterson (1960 : 290),

« the time interval within the syllable nucleus where the formants are parallel to the time axis has been considered as the extent of a vowel target ».

L'« extension » de la cible vocalique » dont il est question correspond à la partie stable de la voyelle et elle se mesure habituellement sur une représentation spectrographique en délimitant la zone où les formants sont parallèles à l'axe du temps. Cette stabilité spectrale reflète une position articuloire stable de la langue et des lèvres. Or, la notion de cible n'a de pertinence qu'au travers de la perception (Johnson et al., 1993 ; Ladefoged, 1997) : en effet, la cible qu'essaie d'atteindre le locuteur en produisant une voyelle est avant tout définie par des propriétés auditives. Diverses procédures de mesure permettent de relever les valeurs absolues de F1, F2 et F3 en Hz en un point précis, généralement au centre de la partie stable de la voyelle. Mais, est-il vraiment judicieux d'utiliser des valeurs absolues lorsqu'on travaille du point de vue de la perception ? En effet, la corrélation entre la localisation des trois premiers formants et sa couleur perceptive est atténuée par un certain nombre de facteurs, notamment la variabilité intra-individuelle (co-articulation, variations dans le discours selon les émotions, la situation...) et inter-individuelle (différences physiologiques, socioculturelles entre individus, stratégies individuelles de production...). Les résultats de Peterson et Barney (1952) sont très

souvent cités pour illustrer la variabilité inter-locuteurs, notamment entre les voix d'enfants, de femmes et d'hommes :

« the first formants for the children are seen to be about half an octave higher than those of the men, and the second and third formants are also appreciably higher » (Peterson et Barney, 1952 : 132).

Mais cette variabilité est en grande partie éliminée si les voyelles sont caractérisées, non plus en termes de fréquences absolues, mais en termes de rapports entre leurs formants. Ainsi, les rapports de fréquences entre le premier et le second formant, et entre le second et le troisième formant pourraient servir à éliminer les différences entre les locuteurs (Peterson, 1961). A la fin du XIXe siècle, Lloyd a formulé une théorie baptisée la « *formant-ratio theory* », relatée par Miller dans un article traitant de la normalisation perceptive (1989 : 2115) :

« Like articulations produce like perceptions of vowel qualities and like articulations produce like ratios of the formants. He called his theory « the relative resonance theory » and stated that the vowel quality depends on intervals between the resonances, not on their absolute values ».

Lloyd fonde notamment sa théorie sur le rôle central des rapports de fréquence en musique. Selon Miller (1989), des travaux paraissent de temps en temps depuis les travaux de Lloyd, mais ne reconnaissent que rarement la paternité de ce dernier. Potter et Steinberg (1950⁹), dans une étude concernant les voyelles de l'A.N.A. prononcées par 10 hommes, 10 femmes et 5 enfants, ont confirmé l'hypothèse que la différence inter-locuteurs était éliminée si les voyelles étaient caractérisées par leurs rapports entre leurs trois premiers formants. Sur la figure n°X, on peut voir que pour chaque catégorie de voyelle, les rapports entre F3 et F2 (M3/M2, ligne du dessus) et entre F2 et F1 (M2/M1, ligne du dessous) sont identiques. Ces rapports sont exprimés en mel, mais Potter et Steinberg (1950) soulignent que les rapports de fréquence éliminent la variabilité inter-locuteurs exactement de la même façon que le font les mel. Cependant, cette méthode ne permet pas de distinguer les voyelles [ɑ] et [ɔ] et les voyelles [ʊ] et [u].

⁹ Travaux cités par Miller (1989 : 2115-2116).

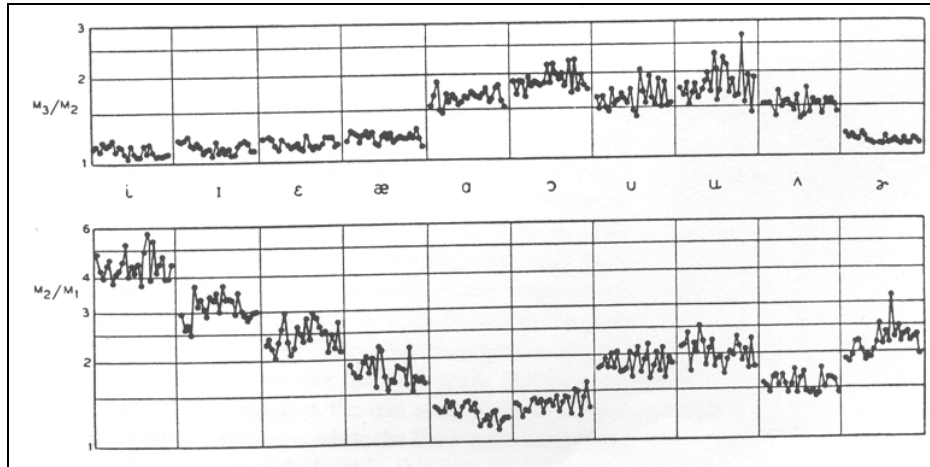


Figure n°65 : rapports $F3-F2$ (au dessus) et $F2-F1$ (en dessous) en équivalent mel (emprunté à Miller, 1989 : 2116, figure n°13 de Potter et Steinberg, 1950).

Cette hypothèse a été confirmée par Peterson (1961) pour les rapports de fréquence (qui correspondent aux intervalles musicaux). Pour cet auteur, la force majeure des rapports de formants réside dans leur capacité à éliminer ou à réduire de manière drastique les différences relatives au locuteur, à l'âge et au sexe dans la description acoustique des voyelles. En revanche, pour certaines voyelles, les rapports sont très similaires et il est possible que d'autres facteurs que les rapports interviennent pour permettre leur différenciation. Les travaux précurseurs de Lloyd et les travaux suivants suggèrent donc tous que les rapports entre les trois premiers formants d'une voyelle jouent un rôle déterminant dans son interprétation auditive et perceptive. Une charte vocalique n'étant pas conçue pour représenter les rapports entre les trois premiers formants, nous avons mis au point une nouvelle représentation, la « charte des rapports de formants » (pour une description de cette représentation, se reporter au Volume II, annexe n°3.2.).

3.1.3.3. Nécessité d'une description plus poussée

Nous ferons donc la description du système de l'anglais en fonction du timbre de ses noyaux vocaliques, mais également de leur durée (absolue et relative) ainsi que de leur structure interne. En ce qui concerne le timbre, nous disposons des données de Peterson et Barney (1952) qui

servent de référence en la matière (valeur standard des trois premiers formants pour 33 hommes, 28 femmes et 15 enfants), mais elles concernent l'A.N.A. Or, lors de l'apprentissage de l'anglais, le modèle de prononciation donné aux enfants est celui de la R.P., c'est-à-dire de l'accent britannique standard (« BBC English »).

En ce qui relève de la durée, les données de Lehiste et Peterson (1960) ne concernent malheureusement que l'A.N.A. Or, l'A.N.A. se différencie de la R.P., notamment par un système vocalique moins riche. Sauf pour le timbre, la littérature ne nous fournit pas de description suffisante correspondant aux indices que nous voulons étudier pour la R.P.

Nous avons donc décidé :

- a) d'établir un corpus qui nous servira dans un premier temps à décrire le système vocalique de la R.P. selon nos propres critères ;
- b) de réaliser une batterie de tests exploitant les principales interférences entre les systèmes vocaliques de l'anglais et du français (voir Volume II, annexe 4.2.).

3.1.4. Réalisation d'une étude descriptive

3.1.4.1. Locutrices

Deux locutrices ont été sélectionnées, une locutrice anglaise (native speaker), professeur d'anglais à l'Université, et une locutrice française, étudiante. La prononciation de la locutrice anglophone est standard (R.P.). Quant à la locutrice francophone, elle a été choisie en fonction de son absence d'accent régional, et ce, malgré son origine franc-comtoise. Sa prononciation est standard.

3.1.4.2. Matériel de parole

Pour décrire le système vocalique de l'anglais par rapport à celui du français, le matériel de parole suivant a été enregistré :

- Pour l'anglais, une liste de 23 mots monosyllabiques de type CVC et de type CV a été enregistrée par la locutrice anglophone. Ces mots ne diffèrent que par leur noyau vocaliques et couvrent la quasi totalité des noyaux vocaliques de l'anglais (8

diphthongues, 3 triphthongues et 11 monophthongues). La voyelle [ə] et les deux triphthongues [aʊə] et [əʊə] ayant une occurrence très faible dans la langue ont été écartées.

- Pour le français, une liste de 24 mots monosyllabiques de type CVC et de type CV a été enregistrée par la locutrice francophone. Là-aussi, les mots ne diffèrent que par leur noyau vocalique. Ils couvrent les 11 monophthongues orales du français. La voyelle [ə] et les 4 voyelles nasales (en raison de l'absence de nasales en anglais) ont été écartées.

En anglais et en français, les mots CVC et CV choisis commencent par la constrictive sourde [s] et les mots CVC finissent par l'occlusive sourde [t]. Nous avons retenus les deux types de syllabes afin de respecter les structures syllabiques de l'anglais : syllabes CVC pour les monophthongues et CV pour les diphtongues et les triphthongues. L'agencement des phonèmes respecte les règles combinatoires de l'anglais. En revanche, pour le français, la liste a été conçue en fonction des interférences possibles avec les voyelles anglaises de timbres proches. De fait, certaines combinaisons retenues n'existent pas en français (comme [set] par exemple).

3.1.4.3. Conditions d'enregistrement

Pour une qualité optimale, les enregistrements ont été réalisés en chambre sourde au Laboratoire de Phonétique de Besançon, avec un magnétophone D.A.T. « Aiwa HD-S1 » et un microphone « Aiwa ». Les locutrices ont pris tout le temps nécessaire pour se préparer et s'approcher de la prononciation la plus standard possible. Lorsqu'elles estimaient avoir échoué, les enregistrements étaient refaits. La consigne leur était donnée de garder le même schéma prosodique et de lire la liste de mots en gardant un débit constant (quelques mots supplémentaires figuraient en début de liste pour que la locutrice s'habitue et en fin de liste afin d'éviter la baisse de l'intonation correspondant à la fin de l'énonciation).

3.1.4.4. Analyses

Pour la description de la longueur des différents noyaux vocaliques produits par la locutrice francophone et par la locutrice francophone, la durée absolue de chaque noyau vocalique a été

relevée (en ms), ainsi que la durée absolue de la syllabe contenant le noyau (en ms). Cette dernière mesure permettra de calculer la durée relative du noyau vocalique au sein de la syllabe (exprimée en pourcentage).

Pour la description de la structure interne des noyaux vocaliques, 47 fichiers ont été édités¹⁰ pour chacun des noyaux vocaliques des 47 mots prononcés par les deux locutrices. Ces fichiers comportent le suivi des trois premiers formants sur la durée totale de chaque noyau vocalique. Ils ont ensuite été traités par le programme informatique réalisé dans le cadre de l'étude n°5 (pour une description de la méthode et une description détaillée de la conception du programme, se reporter au Volume II, annexe n°11.1.). Ce programme fournit la durée absolue (en ms) et relative (exprimée en pourcentage) de la durée des différents événements composant le noyau vocalique, c'est-à-dire la transition initiale, la zone de stabilité de timbre et la transition finale. Au préalable, chaque fichier a été vérifié manuellement, de façon à ce que les résultats fournis par le programme soient totalement fiables.

Enfin, pour l'étude du timbre, la fréquence des trois premiers formants a été prélevée par le programme au milieu de la zone de stabilité de timbre. Une fois ces valeurs transposées en huitièmes de tons¹¹, les intervalles entre F2 et F1 (F2-F1) et entre F3 et F2 (F3-F2) ont été calculés. Puisque le timbre des voyelles est présenté en terme de rapports de formants, la représentation classique de type « charte vocalique » ne pourra pas être utilisée.

3.2. Caractéristiques détaillées

¹⁰ Fichiers édités à l'aide du logiciel Winsnoori, logiciel distribué par Babel Technologies (<http://www.babeltech.com>), mis au point par Yves Laprie, LORIA, Nancy. Version utilisée v.1.2. (1998-1999). Actuellement, la dernière version du logiciel est la version 1.3.

¹¹ Selon la formule $S = (\log(X) - \log(220)) / (\log(2) / 48)$, X étant la valeur en Hz à convertir et 220, la fréquence de référence.

3.2.1. Les noyaux vocaliques simples de l'anglais et du français

3.2.1.1. Durée

En anglais R.P., il existe 12 monophthongues que nous qualifierons de voyelles simples (Ginésy, 1995 : 16). On peut les répartir en deux catégories en prenant pour critère leur durée intrinsèque : 7 voyelles brèves [ɪ], [e], [æ], [ʌ], [ʊ], [ɒ] (**extrait sonore n°28**) et 5 voyelles longues [i:], [ɜ:], [u:], [ɔ:] et [ɑ:] (**extrait sonore n°29**).

3.2.1.1.1. Les voyelles brèves

Lehiste et Peterson (1960) ont montré que les noyaux simples et « brefs » de l'A.N.A. ont une durée de 180 à 200 ms. Dans notre corpus, **la durée absolue des 7 noyaux étudiés n'excède pas 209 ms** pour la voyelle [æ], ce qui correspond aux données de Lehiste et Peterson.

Quant à leur **durée relative**, celle-ci est **courte** (de 25 à 29 % de la syllabe, voir tableau n°5).

En ce qui concerne la répartition des différents événements composant le noyau vocalique, c'est-à-dire la phase de transition initiale (tête), la partie stable et la phase de transition finale (queue), on remarque :

- que la **durée de la partie stable est longue** (de 75 % à 91 % de la durée du noyau, voir exemple de la voyelle [e], figure n°66), sauf dans le cas de la voyelle [ɒ], dont la partie stable, plus courte (58,33 %), occupe tout de même plus de la moitié du noyau. En outre, plus les voyelles sont postérieures, plus leur partie stable est réduite. C'est notamment le cas des voyelles [ɒ] et [ʊ] ;
- la **présence d'une tête dans tous les noyaux** avec une durée relative approximative de 20 %. Cette durée est cependant plus courte pour les voyelles [ɪ] (12,78 %) et [æ] (8,82 %) ;

- **l'absence de queue dans la plupart des noyaux**, sauf dans le cas des voyelles [ɪ] (9,02 %) et [ɒ] (16,11 %). Néanmoins, pour ces deux voyelles, la durée relative de la queue reste courte.

	Durée de la tête (DT)	Durée de la partie stable (DPS)	Durée de la queue (DQ)	Durée de la voyelle par rapport à la syllabe
[ɪ]	12,78 %	78,19 %	9,02 %	24,47 % (139 ms)
[e]	20 %	80 %	0 %	24,62 % (180 ms)
[æ]	8,82 %	91,17 %	0 %	29,52 % (209 ms)
[ʌ]	22,29 %	77,70 %	0 %	25,31 % (163 ms)
[ʊ]	25,15 %	74,84 %	0 %	25,11 % (169 ms)
[ɒ]	25,55 %	58,33 %	16,11 %	27,39 % (186 ms)

Tableau n°5 : durée relative des différents événements composant les 7 noyaux vocaliques brefs de la locutrice anglophone (exprimée en pourcentage).

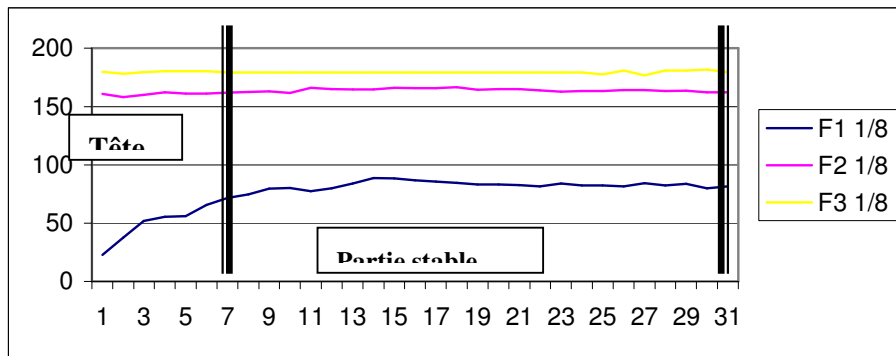


Figure n°66 : évolution des trois premiers formants de la voyelle brève [e] dans le mot « set » prononcé par la locutrice anglophone (exprimée en 1/8^{ème} de tons) ; la partie stable s'étend de la mesure n°7 à la mesure n°31.

En résumé, les voyelles brèves de notre corpus se caractérisent par une durée absolue (<210 ms) et relative réduite (25 à 29 % de la durée de la syllabe), une grande partie stable (plus de 50 % de la durée de la voyelle), la présence d'une tête (approximativement 20 % de la durée de la syllabe) et une queue presque toujours absente ou alors très réduite.

3.2.1.1.2. Les voyelles longues

D'après Lehiste et Peterson (1960), les noyaux vocaliques simples et longs de l'A.N.A. se caractérisent par une durée absolue comprise entre 240 et 260 ms. Dans notre corpus, la durée

absolue des 5 noyaux étudiés se répartit entre 233 et 325 ms (voir tableau n°6), la voyelle [i:] étant la plus courte (180 ms). La durée relative du noyau vocalique, exprimée en pourcentage de la durée totale de la syllabe, est plus importante pour les voyelles longues que pour les voyelles brèves (aux environs de 35 %), à cause de leur plus grande durée intrinsèque sauf dans le cas de la voyelle [i:]. Cette dernière se caractérise par une grande phase de stabilité (63,79 %) qui la rapproche des voyelles brèves. Cependant, elle se démarque de la voyelle brève la plus proche de timbre [ɪ] par une phase de transition initiale beaucoup plus longue (33,33 % contre 12,78 %). Bien que la voyelle [ɜ:] ait une durée plus longue que la voyelle [i:], elle se caractérise également par une grande phase de stabilité (59,42 %).

En revanche, les trois autres voyelles longues, c'est-à-dire [u:], [ɔ:] et [ɑ:] se caractérisent par une forte instabilité (respectivement 74,44 %, 72,15 %, 61,74 % de transition si on ajoute la phase de transition initiale à la phase de transition finale). La phase de transition finale constitue le segment le plus long dans les voyelles [ɑ:] et [u:] (respectivement 51,09 % et 56,38 %). À l'inverse, dans la voyelle [ɔ:] (voir figure n°67), c'est la phase de transition initiale qui couvre la durée la plus grande (48,71 %).

La grande instabilité qui caractérise ces voyelles s'explique par la tendance à la diphtongaison qui caractérise la production des sons vocaliques de longue durée en anglais. Cette tendance provient directement de l'organisation rythmique de l'anglais (cf. 3.1.1.). Cette instabilité semble plus particulièrement affecter les voyelles postérieures.

En résumé, il existe une assez grande différence entre les voyelles longues antérieures et les voyelles longues postérieures. Ces dernières se caractérisent par une grande phase d'instabilité (supérieure à 60 % de la voyelle), une grande durée absolue (de 279 à 325 ms) et une grande durée relative (de 32,63% à 42,09 % de la durée totale de la syllabe). Les caractéristiques des voyelles [i:] et [ɜ:] se caractérisent au contraire par une grande phase de stabilité et une répartition des événements qui les rapprochent des voyelles brèves.

	Durée de la tête (DT)	Durée de la partie stable (DPS)	Durée de la queue (DQ)	Durée de la voyelle par rapport à la syllabe
[i:]	33,33 %	63,79 %	2,87 %	27,90 % (180 ms)
[a:]	10,65 %	38,24 %	51,09 %	42,09 % (325 ms)
[ɔ:]	48,71 %	27,83 %	23,44 %	36,66 % (279 ms)
[u:]	18,06 %	25,55 %	56,38 %	32,63 % (233 ms)
[ɜ:]	23,96 %	59,42 %	16,61 %	35,68 % (319 ms)

Tableau n°6 : durée relative des événements composant les 5 noyaux vocaliques longs de la locutrice anglophone.

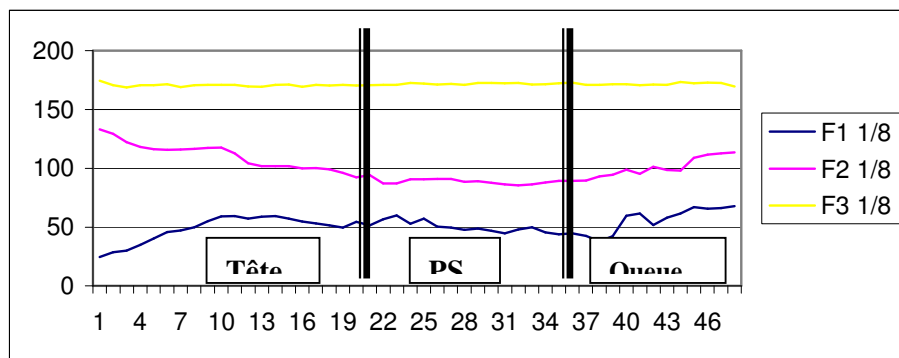


Figure n°67 : évolution des trois premiers formants (en 1/8^{ème} de ton) de la voyelle [ɔ:] dans le mot « sort » prononcé par la locutrice anglophone ; la partie stable s'étend de la mesure n°24 à la mesure n°37.

3.2.1.1.3. Les voyelles du français

A cause de sa rythmique, le français se caractérise par une grande tension articulatoire qui empêche toute tendance à la diphthongaison. En conséquence, les voyelles ont un timbre précis, c'est-à-dire qu'il ne change pas en cours d'émission, et ce, même si elles se trouvent dans une syllabe inaccentuée.

Si nous laissons de côté les voyelles oralo-nasales et la voyelle [ə], le français comporte 11 voyelles orales, [i], [e], [y], [ɛ], [a], [ø], [u], [o], [ɔ], [ɑ] et [œ].

La durée absolue du noyau vocalique n'excède pas 139 ms pour les voyelles les plus « brèves » du français, c'est-à-dire les voyelles [i], [e], [ɛ], [a], [ɔ], [u] et [œ]. Cette durée est plus courte que pour les voyelles brèves de l'anglais dont la durée s'étend jusqu'à 209 ms (tableau n°7). La durée relative du noyau au sein de la syllabe n'excède pas 21,95 %, ce qui est également plus court que pour les voyelles brèves de l'anglais (25 à 29 %). On note une grande stabilité de timbre (de 52,25 à 100 % ; voir pour la voyelle [ɔ], figure n°68), comme pour les voyelles brèves de l'anglais (de 58 à 91 %).

La durée absolue des voyelles fortement labialisées comme [o] et [ø] est plus longue (respectivement 180 et 203 ms), mais, même avec cette longueur, ces voyelles sont aussi longues que les voyelles brèves de l'anglais. Leur durée relative (respectivement 31,51 % et 35,97 %) est également beaucoup plus importante, comme celle des voyelles longues de l'anglais, mais elles sont beaucoup plus stables que ces dernières ; on relève une plus grande stabilité de timbre (zone de stabilité de 67,77 % pour [o] et de 91,13 % pour [ø]), en raison de l'articulation fortement tendue du français.

La voyelle [ɑ] est la plus longue des voyelles produites par la locutrice francophone, que ce soit du point de vue de sa durée absolue (301 ms), que de sa durée relative (46,79 % de la durée totale de la syllabe). On remarque cependant une zone de stabilité plus courte (42,52 % de la durée totale de la voyelle), au profit d'une grande transition initiale (51,82 %). On relève une répartition quasiment identique pour la voyelle [a] (55,17 % de stabilité et transition initiale de 55,17 %), bien que sa durée absolue et sa durée relative soient courtes (respectivement 116 ms et 18,59 % de la durée totale de la syllabe).

Mais, dans tous les cas, sauf pour [a] et [ɑ], la partie stable est toujours le segment le plus long du noyau vocalique. Des voyelles comme [i], [y], [o], [ø], [œ] se caractérisent notamment par une phase de stabilité qui couvre presque la totalité de leur durée (plus de 85 % de zone de stabilité par rapport à leur durée totale). Les voyelles orales du français se caractérisent donc par grande stabilité et une durée absolue généralement plus courte (< 203 ms) que les voyelles brèves de l'anglais, exception faite de la voyelle [ɑ].

Leur durée relative avoisine les 20 % ce qui est également plus court que les voyelles brèves de l'anglais. Les voyelles les plus fermées se caractérisent par une plus grande durée relative (de 31,51 à 46,79 % de la durée totale de la syllabe).

	Durée de la tête (DT)	Durée de la partie stable (DPS)	Durée de la queue (DQ)	Durée de la voyelle par rapport à la syllabe
[i]	0 %	100 %	0%	14,70 % (111 ms)
[e]	7,91 %	71,22 %	20,86 %	20,13 % (139 ms)
[ɛ]	8,27 %	57,14 %	34,58 %	21,95 % (133 ms)
[a]	55,17 %	35,34 %	9,48 %	18,59 % (116 ms)
[y]	6,89 %	93,10 %	0 %	15,70 % (87 ms)
[ø]	2,95 %	91,13 %	5,91 %	35,97 % (203 ms)
[œ]	6,89 %	86,20 %	6,89 %	18,19 % (87 ms)
[u]	0 %	68,18 %	31,81 %	18,68 % (110 ms)
[o]	16,11 %	67,77 %	16,11 %	31,51 % (180 ms)
[ɔ]	20,72 %	52,25 %	27,02 %	19,18 % (111 ms)
[ɑ]	51,82 %	42,52 %	5,64 %	46,79 % (301 ms)

Tableau n°7 : durée relative des événements composant les 11 monophthongues de la locutrice francophone.

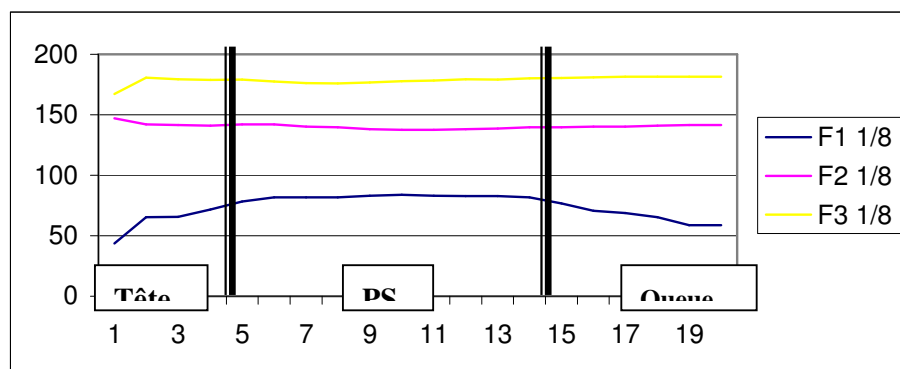


Figure n°68 : évolution des trois premiers formants (en 1/8^{ème} de ton) de la voyelle [ɔ] dans le mot « sothe » prononcé par la locutrice francophone ; la partie stable s'étend de la mesure n°5 à la mesure n°15.

3.2.1.1.4. Interférences de durée du français vers l'anglais

D'après nos résultats (pour un récapitulatif, voir tableau n°8), les voyelles brèves de l'anglais ont généralement une durée absolue et relative plus longue que celle des voyelles françaises. Par ailleurs, bien qu'elles se caractérisent par une stabilité certaine, cette stabilité reste moins marquée qu'elle ne l'est dans les voyelles françaises. Cependant, malgré ces différences, leurs structurations temporelles sont semblables et cette similitude risque de favoriser les phénomènes d'assimilation entre voyelles anglaises et voyelles françaises de timbre proche.

En cas d'assimilation, il est probable qu'un locuteur francophone produise les voyelles brèves de l'anglais avec la stabilité et la durée qui caractérisent les voyelles françaises. Dans ce cas, la mesure de la durée absolue et relative de la voyelle, ainsi que la mesure de la zone de stabilité fourniront des indices sur le poids du crible phonologique chez l'apprenant.

La durée absolue et la durée relative des voyelles longues de l'anglais sont proches de celles de la voyelle française la plus longue, la voyelle [ɑ]. Elles constituent donc des sons très longs pour un locuteur francophone, qui aura tendance à les raccourcir et de ce fait, à accroître leur stabilité. Le phénomène de crible devrait être moindre pour les voyelles longues que pour les voyelles brèves de l'anglais car ces sons ont des caractéristiques « non natives » plus saillantes, donc plus repérables par l'apprenant. Cependant, si malgré tout une assimilation entre voyelles de timbre correspondant se produit, les français vont transposer la tension articulatoire aux voyelles longues de l'anglais, qui se caractériseront par un excès de stabilité. Selon Faure (1948 : 79),

« les Français ont généralement le défaut de ne pas allonger assez les voyelles longues et les diphtongues, à la fin des mots ou devant une consonne sonore, et, donc, de prononcer des mots comme : sea, see, car, door, d'une façon trop brève [...]. Ils doivent tendre [...] à leur donner une longueur beaucoup plus considérable qu'ils ne le font d'ordinaire ».

Langues	Type de voyelles	Caractéristiques
Anglais	7 voyelles brèves [ɪəæʌʊə]	DA < 209 ms 25 % < DR < 29 % Grande PS (de 58 à 91 %) > T ou Q Petite tête (de 8,82 à 20 %) Absence de queue, sauf pour [ɪ] et [ə].
	5 voyelles longues ou « diphtonguées » [i:ɑ:ɔ:u:ɜ:]	233 < DA < 325 ms (180 ms pour [i:]) DR = 35 % (sauf pour [i:]) PS < T ou Q, sauf pour [i:] et [ɜ:]
Français	8 voyelles brèves [iɛɛɑʊœə]	DA < 139 ms DR < 21,95 % Grande PS (de 52,25 à 100 %) > T ou Q Tête, de à 20,72 % Queue, de à 34,58 % (plus longue pour [œʊə])

	3 voyelles longues [oø]	180 < DA < 203 ms 31,51% < DR < 35,97 % Tête et queue réduites.
	[ɑ]	DA=301 ms DR=46,79 % PS (42,52 %) < T (51,82 %), comme pour [ɑ] (PS=35,34 % et T=55,17 %).

Tableau n°8 : récapitulatif des caractéristiques de durée des différents types de voyelles de l'anglais et du français (DA pour durée absolue, DR pour durée relative, PS pour partie stable, T pour tête et Q pour queue).

3.2.1.2. Timbre

Les différents types d'interférences

Certains contrastes de la langue cible dits contrastes non-natifs (désormais NN) sont absents dans la LM, mais montrent une similitude de timbre avec certains contrastes natifs (désormais N). En raison du phénomène de crible phonologique (voir partie 1.2.), il sera très difficile pour l'auditeur de les différencier et il risque d'assimiler et de produire le contraste NN à un contraste N ayant un timbre proche.

Un autre type d'interférence peut intervenir, lorsque deux voyelles NN, absentes du système vocalique maternel montrent une similitude de timbre. L'auditeur aura alors beaucoup de difficultés à les différencier et les deux contrastes seront généralement assimilés l'un à l'autre, ou assimilés tous les deux à un contraste N de timbre proche (catégorie SC dans la typologie de Best, voir 1.2.).

La graphie peut, elle aussi, favoriser des interférences de timbre. Prenons un mot anglais comme « *lip* ». Dans ce mot, la graphie de la voyelle [ɪ] peut favoriser l'assimilation de ce son avec la voyelle française [i].

La similitude de forme graphique favorise également les interférences. Le mot « *orange* » s'écrit de manière identique et possède une signification similaire dans les deux langues. En fait, il ne diffère que par sa prononciation : [ɔʁɑ̃ːʒ] en français et [ˈɔrɪndʒ] en anglais. Dans cet exemple, non seulement la graphie favorise l'assimilation de la voyelle anglaise [ɔ] à la voyelle

française [ɔ] de la même façon que dans l'exemple du mot « *lip* », mais de plus, cette assimilation est renforcée par la similitude de forme entre les deux mots. Dans le cas d'« *orange* », cette similitude est totale, puisqu'il s'agit du même mot et l'apprenant aura tendance à prononcer le mot « *orange* » à la française.

Par ailleurs, deux voyelles NN peuvent se transcrire de la même manière. C'est le cas de la voyelle [ɑ:] dans le mot « *car* » et de la voyelle [æ] dans le mot « *cat* ». L'apprenant aura donc tendance à produire la même voyelle NN, soit [æ], soit [ɑ:] dans les deux mots ou alors il produira à la place la voyelle française [a], qui s'écrit de la même manière et qui possède une similitude de timbre avec les deux voyelles NN.

Enfin, les interférences avec les voyelles françaises peuvent être favorisées lorsque les mots anglais sont familiers et prononcés « à la française » par l'entourage de l'apprenant. Ce phénomène est accentué par le fait que l'anglais est une langue très répandue et les emprunts nombreux. Même les locuteurs francophones qui ne parlent pas anglais connaissent et utilisent des termes anglais. Ainsi, des mots « *book* », « *park* » (qu'on trouve dans le mot français « *parking* »), « *dog* », « *car* », « *shop* »...

Prédiction de ces interférences à partir de l'étude comparée des systèmes vocaliques de la locutrice anglophone et de la locutrice francophone

Les fréquences des trois premiers formants des voyelles composant le système vocalique de la locutrice anglophone et celui de la locutrice francophone ont été relevées (Volume II, annexe n°3.3., tableaux n° 1 et n°2). Après avoir calculé les intervalles séparant F2 et F1 et F3 et F2 pour chaque voyelle, une charte des rapports de formants a été établie pour chaque locutrice (figure n°69 et figure n°70). La comparaison de ces chartes a permis de déterminer les proximités de timbre entre voyelles anglaises et voyelles françaises et donc, les risques

d'assimilation pouvant intervenir entre les deux langues. La figure n°71 propose une vision synthétique de ces interférences.

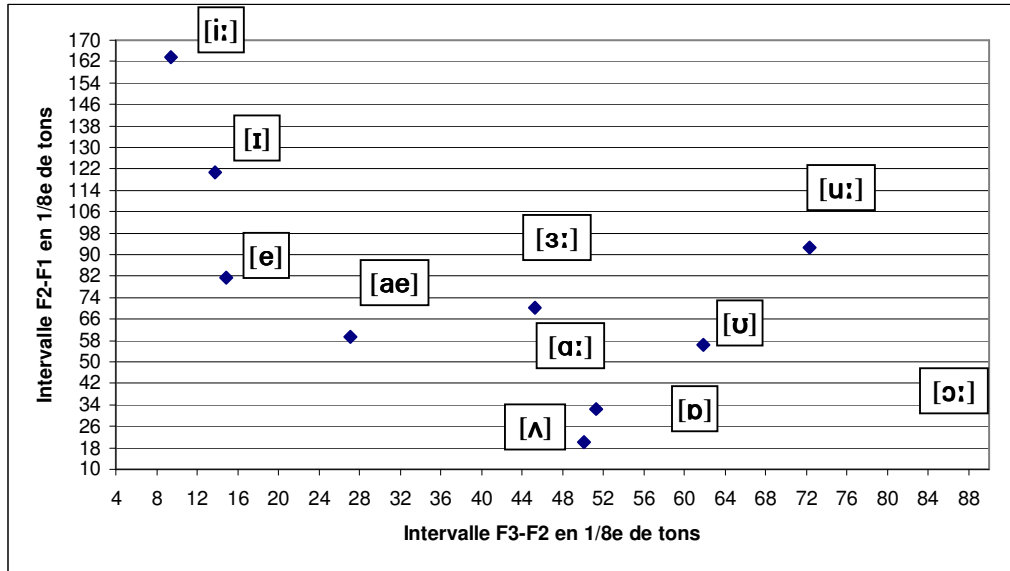


Figure n°69 : charte des rapports de formants de la locutrice anglophone.

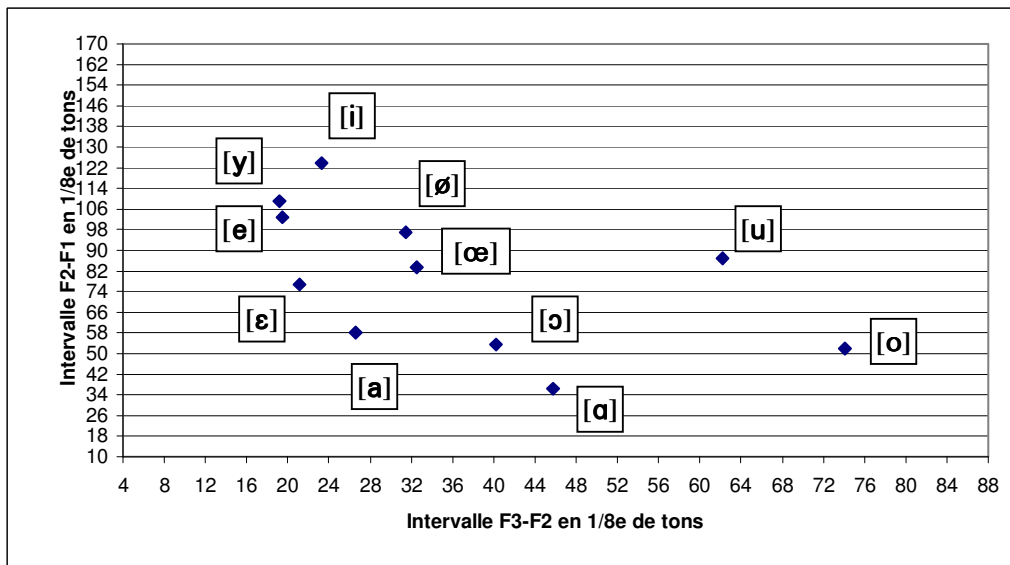


Figure n°70 : charte des rapports de formants de la locutrice francophone.

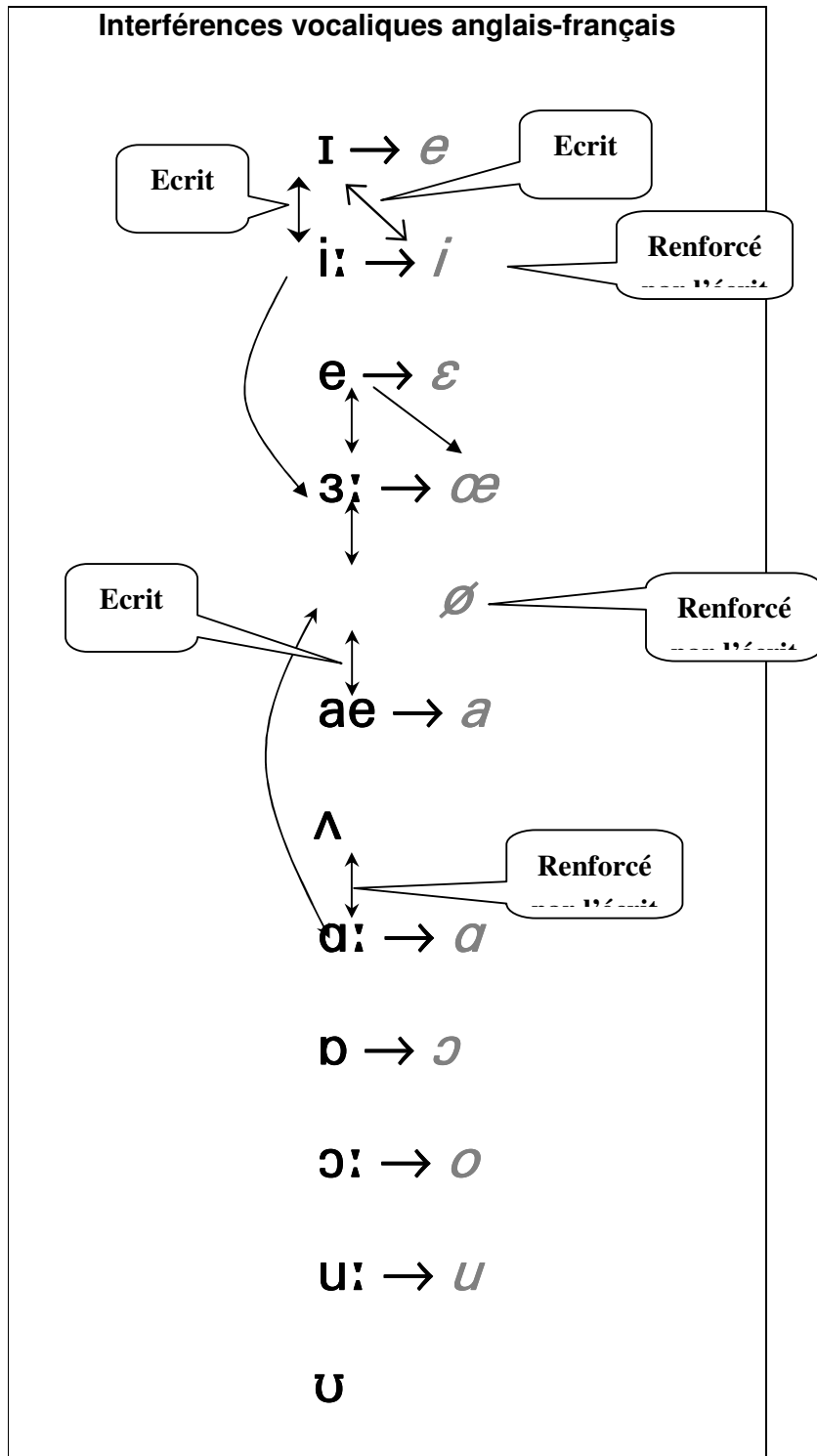


Figure n°71 : relevé des principales assimilations des contrastes non natifs de l'anglais à d'autres contrastes de l'anglais et aux contrastes natifs du français (en gris).

3.2.2. Les noyaux vocaliques complexes en anglais

Rappelons qu'en anglais R.P., il existe 13 noyaux vocaliques complexes dont 8 diphtongues et 5 triphthongues.

3.2.2.1. Les diphtongues

3.2.2.1.1. Définition

Au niveau de la production, on peut définir la diphtongue comme le noyau d'une syllabe contenant deux cibles vocaliques. Bien qu'elle soit transcrite en A.P.I. par un double symbole phonétique ([aɪ] par exemple), elle fonctionne comme une seule unité vocalique appartenant à la même syllabe (Gynésy, 1995 : 34) et les symboles phonétiques utilisés ne correspondent pas à la qualité de la première, ni de la seconde cible :

« They can be described as movements from one vowel to another [...]. Furthermore, contrary to the traditional transcriptions, the diphthongs often do not begin and end with any of the sounds that occur in simple vowels » (Ladefoged, 1975 : 69).

Sur une représentation spectrographique, une diphtongue se caractérise par un changement de trajectoire des trois premiers formants en cours d'émission. Les différents événements qui composent une diphtongue sont, dans l'ordre d'apparition, une phase de transition initiale (désormais TE pour « tête »), une première partie stable (désormais PS1) correspondant à la première cible, une zone de transition centrale (désormais TC), une deuxième partie stable (désormais PS2) correspondant à la deuxième cible et une zone de transition finale (désormais Q pour « queue »). Le tableau n°9 en schématise la structure. Les deux cibles sont clairement identifiables (voir figure n°71, diphtongue [ɔɪ]). Il est rare que les deux segments cibles aient les mêmes caractéristiques de durée et d'intensité. En anglais,

« *the first part of the diphthong is usually more prominent than the last. In fact, the last part is often so brief and transitory that is difficult to determine its exact quality* » (Ladefoged, 1975 : 69).

La diphtongue commence donc plutôt énergiquement, puis elle s'affaiblit. Pour désigner la partie « proéminente », Catford (1988 : 115-116) parle de « *stress pulse* ». Celui-ci se repère facilement si une représentation de la mélodie est superposée à une représentation spectrographique.

En effet, on observe une correspondance du pic mélodique du contour de Fo (voir figure n°72 pour la diphtongue [ɔɪ] où la flèche indique une correspondance entre le pic de Fo et le premier élément de la diphtongue). Lehiste et Peterson (1960 : 290) ont utilisé cet indice pour effectuer leurs prises de mesure :

« *the intonation pattern was often helpful in determining the position of the target, since the peak of the intonation contour associated with the CNC word usually coincided with the target of the syllable nucleus* ».

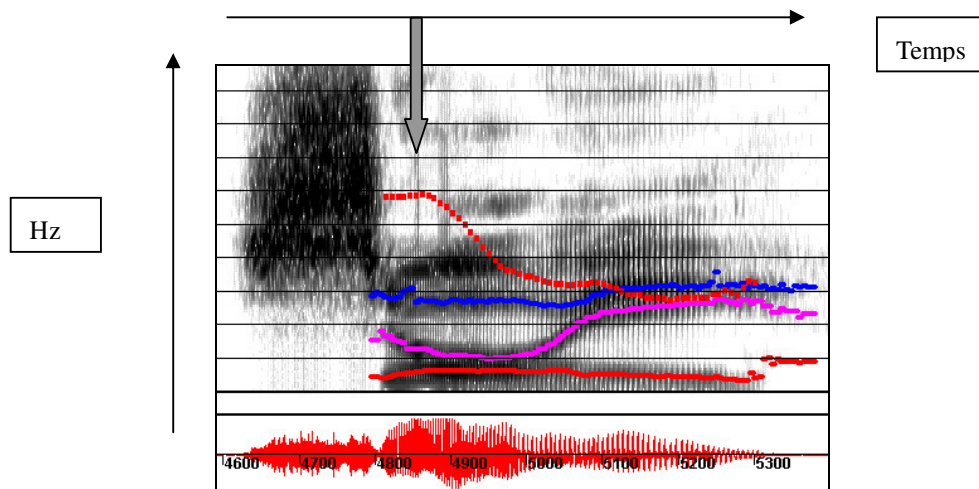


Figure n°72 : représentation spectrographique à bandes larges de la diphtongue [ɔɪ] du non mot « soy » prononcé par la locutrice anglophone avec superposition de l'évolution des trois premiers formants (lignes continues) et du contour de Fo (ligne en pointillés).

Du point de vue de la perception, dans une diphtongue, les deux cibles vocaliques, c'est-à-dire le point de départ et le point d'arrivée ne sont pas perçus comme des voyelles séparées, mais comme un glissement vers le second élément.

« La durée des segments initiaux et finaux et surtout l'angle que forme F2 en direction de la cible avec l'horizontale semblent constituer les deux facteurs de perception les plus importants pour la distinction des diphtongues » (Carton, 1974 : 44).

Dans une étude sur le français parlé à Montréal, Yaeger (1979) a identifié les principaux indices acoustiques permettant d'identifier des allophones diphtongués : une variation de F2 de plus de 200 Hz pour toutes les voyelles ou un changement d'au moins 100 Hz pour F1, même si F2 est stable pour les voyelles postérieures.

TE	PS1	TC	PS2	Q
-----------	------------	-----------	------------	----------

Tableau n°9 : les différents événements composant la structure d'une diphtongue (TE pour tête, PS1 pour première partie stable, TC pour transition centrale, PS2 pour deuxième partie stable et Q pour queue).

3.2.2.1.2. Description des diphtongues de l'anglais

Il existe différents types de diphtongues en anglais :

- 5 diphtongues fermantes (dont la trajectoire monte dans l'espace vocalique, c'est-à-dire que l'aperture se réduit au cours du temps. C'est le cas pour [aɪ], [eɪ], [ɔɪ], [əʊ] et [aʊ] (**extrait sonore n°31**) ;
- 3 diphtongues centralisantes dont la trajectoire commence dans la périphérie de l'espace vocalique et se dirige ensuite vers son centre. C'est le cas pour [ɪə], [ʊə] et [eə] (**extrait sonore n°32**). Ginésy (1995 : 34) souligne cependant que la diphtongue [eə] n'existe pratiquement plus en R.P.

Cibles et durées

Dans cette description, nous ne séparerons pas les caractéristiques de timbre de celles de durée comme nous l'avons fait pour les voyelles, pour la simple et bonne raison qu'elles sont intimement liées dans le phénomène de diphtongaison. En effet, la durée relative des événements composant une diphtongue dépend essentiellement de la cible à atteindre et tous les gestes articulatoires ne demandent pas le même effort. C'est donc en fonction de la plus ou moins grande proximité de la cible à atteindre, que les gestes articulatoires seront plus ou moins complexes à produire et que la durée nécessaire à leur réalisation sera plus ou moins longue.

Les diphtongues fermantes [aɪ], [eɪ], [ɔɪ], [əʊ] et [aʊ]

La trajectoire des diphtongues fermantes monte dans l'espace vocalique comme on peut le voir sur la charte des rapports de formant (figure n°73), à cause d'une réduction de l'aperture en cours d'émission. Au niveau articulatoire, l'aperture dépend de la position plus ou moins écartée de la mâchoire inférieure et de l'élévation subséquente de la langue par rapport au palais. Au niveau acoustique, une réduction d'aperture se manifeste par la baisse du premier formant (F1), provoquant une augmentation de l'intervalle entre F2 et F1.

La diphtongue [ɔɪ] se caractérise par la trajectoire la plus longue (figure n°73), une grande durée absolue (493 ms), ainsi qu'une grande instabilité (42,32 % de stabilité, en additionnant PS1 et PS2).

Cette instabilité peut s'expliquer par le fait qu'il faut du temps pour passer de la voyelle postérieure et ouverte correspondant à la première cible (proche de [ɔ:]) à la voyelle antérieure et fermée correspondant à la seconde cible (proche de [i:]), deux extrêmes sur la charte des rapports de formants. La diphtongue [eɪ] est également très instable (44,74 % de stabilité). Comme [ɔɪ], elle se caractérise par une grande durée absolue (510 ms) par rapport à des diphtongues comme [əʊ] (278 ms) ou [ʊə] (261 ms), dans le même contexte syllabique (CV).

En revanche, sa trajectoire est moins grande, car les deux cibles sont moins éloignées l'une de l'autre. On relève cependant la plus grande phase de transition centrale (TC) de toutes les diphtongues (31,08 %), après celle de la diphtongue [ɪɛ] (34,71 %). Alors que la diphtongue [ɔɪ] connaît un grand changement d'aperture et un grand changement de lieu d'articulation pendant son émission, la diphtongue [eɪ] ne connaît qu'un changement d'aperture (accompagné il est vrai d'une légère antériorisation, mais moindre par rapport à [ɔɪ]). Une transition centrale plus longue pourrait permettre à l'auditeur de bien faire la différence entre les deux cibles. Il est possible, par ailleurs, qu'il soit plus difficile, au niveau articulatoire, de produire un changement d'aperture non accompagné de changement de lieu d'articulation. Dans la diphtongue [aɪ], la transition centrale est au contraire très brève (6,41 %). Il y a peu de changement d'aperture et la trajectoire se caractérise plutôt par un changement de lieu articulatoire. C'est une diphtongue beaucoup plus stable (52,01 % de stabilité) que les deux précédentes, avec une PS1 plus grande (28,57 %).

Les diphtongues [aʊ] et [əʊ] sont encore plus stables que [aɪ] (respectivement 66,88 % et 80,94 % de stabilité). En effet, dans [aʊ] on relève une grande PS1 (51,01 %) et dans [əʊ], une grande PS2 (57,50 %). Leurs durées absolues sont moins grandes car elles apparaissent dans un contexte syllabique CVC (278 ms et 296 ms). Elles ont toutes les deux le même point de départ, une première cible très centralisée, puis la trajectoire de [aʊ] évolue vers une seconde cible très postérieure (proche de [ɔ:]), tandis que la trajectoire de [əʊ], beaucoup moins marquée, évolue vers seconde cible moins éloignée (proche de [u:]). Mais, dans les deux cas, il y a peu de changement d'aperture et comme ces deux diphtongues sont très stables, il semblerait que le taux d'instabilité soit lié au degré d'aperture.

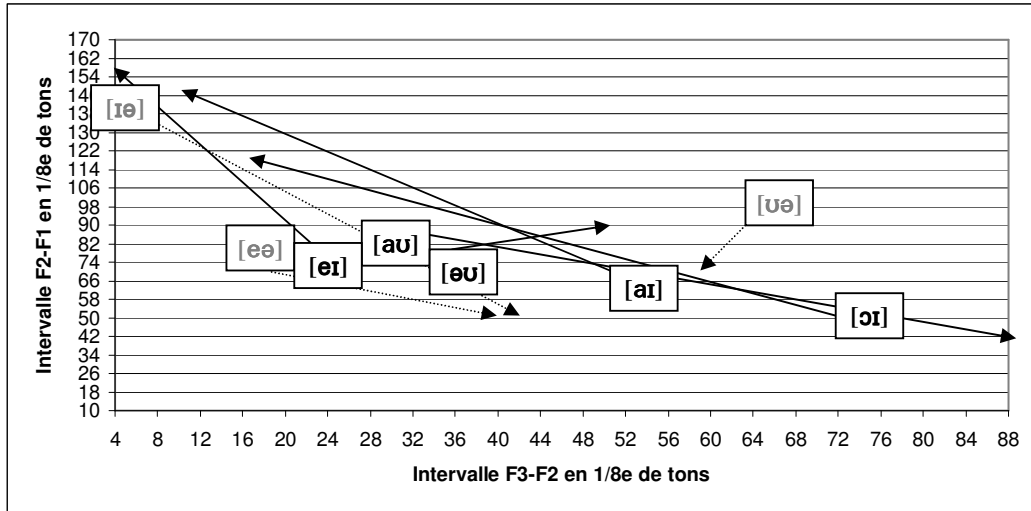


Figure n°73 : trajectoires des diphtongues de la locutrice anglophone

(en gris, diphtongues centralisantes et trajectoires en pointillé ; en noir, diphtongues fermantes et trajectoires en ligne continue).

Les diphtongues centralisantes [ɪə], [ʊə] et [eə]

La diphtongue [ɪə] se caractérise par une grande stabilité (65,28 %), mais également, comme nous l'avons déjà vu, par une forte TC entre les deux cibles (34,71 %). Sa trajectoire s'illustre notamment par un changement d'aperture plus grand que dans la diphtongue fermante [əɪ]. La diphtongue [eə] se caractérise par une très forte stabilité (77,53 %), notamment sur la première cible (49,19 %, ce qui représente quasiment la moitié de la durée totale de la diphtongue). La seconde cible occupe 28,34 % de la diphtongue. Entre ces deux cibles, la TC est réduite (7,08 %). La diphtongue [ʊə] se démarque en revanche par la plus forte instabilité de toutes les diphtongues étudiées (34,1 % de stabilité). Sa durée absolue est également la plus courte (261 ms) alors qu'elle est produite dans une syllabe CV. Sa trajectoire est également très courte et les zones de transition sont donc très marquées, et notamment la tête (20,39 %) et la queue, qui occupent presque la moitié de la durée totale de la diphtongue (43,13 %). En fait, ce son est très difficile à prononcer. La consonne le précédant est la constrictive dentale [s], qui s'articule à l'avant de la bouche. Pour atteindre la première cible de la diphtongue [ʊ], il faut aller à

l'opposé de l'espace vocalique, ce qui explique la longueur de la tête et le fait que la cible soit plus fermée que la voyelle correspondante.

	DT	DPS1	DTC	DPS2	DQ	Durée par rapport à la syllabe
[eɪ]	9,30 %	16,03%	31,08 %	28,71 %	14,85 %	75,78 % (510 ms) CV
[aɪ]	3,29 %	28,57 %	6,41 %	23,44 %	38,27 %	75,20 % (546 ms) CV
[ɔɪ]	13,11 %	23,77 %	15,57 %	19,05 %	28,48 %	73,25 % (493 ms) CV
[əʊ]	8,79 %	23,44 %	1,83 %	57,50 %	8,42 %	33,25 % (278 ms) CVC
[aʊ]	4,05 %	51,01 %	17,56 %	15,87 %	11,48 %	34,18 % (296 ms) CVC
[ɪə]	0 %	28,04 %	34,71 %	37,24 %	0 %	67,84 % (441 ms) CV
[eə]	15,38 %	49,19 %	7,08 %	28,34 %	0 %	72,35 % (500 ms) CV
[ʊə]	20,39 %	16,07 %	2,35 %	18,03 %	43,13 %	63,34 % (261 ms) CV

Tableau n°10 : durée relative des événements composant les 8 diphtongues prononcées par la locutrice anglophone (DT pour durée totale, DPS1 pour durée de la première partie stable, DTC pour durée de la transition centrale, DPS2 pour durée de la deuxième partie stable, DQ pour durée de la queue, CV pour syllabe consonne-voyelle et CVC, pour syllabe consonne-voyelle-consonne).

3.2.2.2. La triptongue

3.2.2.2.1. Définition

Du point de vue de la production, on peut définir une triptongue comme le noyau d'une syllabe contenant trois cibles vocaliques. Bien qu'elle soit transcrite en A.P.I. par un triple symbole phonétique (identifiant le point de départ, la cible à mi-chemin et la cible finale, comme dans [aɪə] par exemple), elle fonctionne comme une seule unité vocalique appartenant à la même syllabe. Les différents événements qui composent une triptongue sont, dans l'ordre d'apparition, une phase de transition initiale (TE), une première partie stable (PS1) correspondant à la première cible, une première zone de transition centrale (TC1), une deuxième partie stable (PS2) correspondant à la deuxième cible (partie stable souvent réduite, voire absente), une deuxième zone de transition centrale (TC2), une troisième partie stable

(PS3) correspondant à la troisième cible et une zone de transition finale (Q). Le tableau n°11 en schématise la structure.

Sur une représentation spectrographique, on observe une trajectoire complexe des trois premiers formants, avec un changement de direction à mi-course (voir figure n°74). Le premier élément de la triptongue, le plus proéminent, correspond à la localisation du pic mélodique du contour de Fo (voir figure n°74 où le pic de Fo concorde avec le premier élément de la triptongue [ɔɪə], à l'endroit indiqué par une flèche).

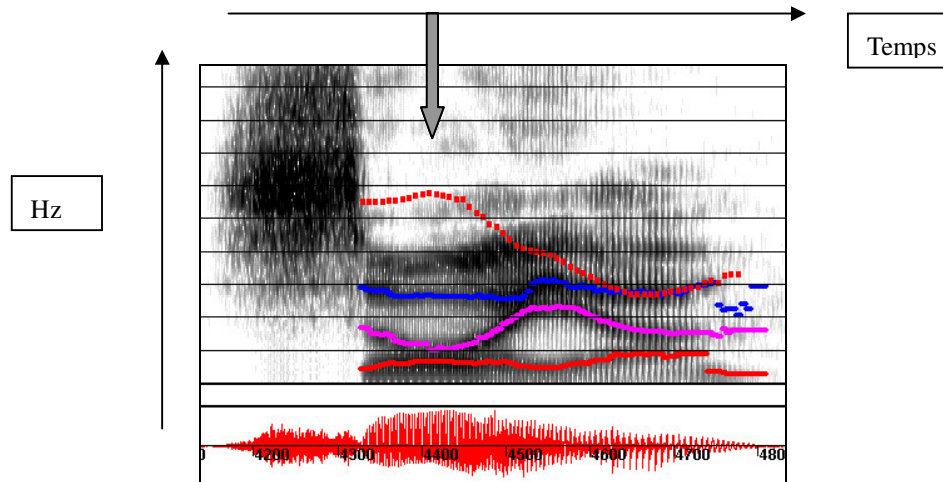


Figure n°74 : représentation spectrographique à bandes larges de la triptongue [ɔɪə] du non mot « soya » prononcé par la locutrice anglophone avec superposition de l'évolution des trois premiers formants (lignes continues) et du contour de Fo (ligne en pointillés).

Du point de vue de la perception, dans une triptongue, les trois cibles vocaliques, c'est-à-dire le point de départ, la cible à mi-chemin et le point d'arrivée ne sont pas perçus comme trois voyelles séparées, mais comme un glissement vers le second élément dans un premier temps et comme un glissement vers le troisième élément dans un second temps.

TE	PS1	TC1	PS2	TC2	PS3	Q
----	-----	-----	-----	-----	-----	---

Tableau n°11 : les différents événements composant la structure d'une triptongue (TE pour tête, PS1 pour première partie stable, TC1 pour première transition centrale, PS2 pour deuxième partie stable, TC2 pour deuxième transition centrale, PS3 pour troisième partie stable et Q pour queue).

3.2.2.2.2. Description des triptongues de l'anglais

Selon Ginésy (1995), les vraies triphthongues sont relativement rares en anglais. Elles apparaissent dans des mots indécomposables tels que « *fire* » [faiə], « *choir* » [kwaiə], « *shower* » [ʃaʊə], etc. En anglais R.P., on en dénombre 5 :

- 3 triphthongues ayant comme cible centrale [ɪ] : [aɪə], [eɪə] et [ɔɪə] (**extrait sonore n°33**) ;
- 2 triphthongues ayant comme cible centrale [ʊ] : [aʊə] et [əʊə] ; nous les décrivons pas car, comme nous l'avons déjà précisé, elles apparaissent très rarement en anglais R.P.

La tendance dominante est à la simplification et les triphthongues se réduisent souvent à une diphtongue, la cible centrale n'étant pas réalisée : « *les sons [ɪ] et [ʊ], qui sont les éléments centraux de ces triphthongues sont rarement atteints mais sont plutôt effleurés* » (Ginésy, 1995).

Il arrive même qu'elles soient réduites à une longue monophthongue.

Du point de vue de leur structuration temporelle, les deux triphthongues [eɪə] et [ɔɪə] sont semblables, bien que la trajectoire de [ɔɪə] soit beaucoup plus marquée (voir ci-dessous, sur la charte de rapports de formants, figure n°75). Cette différence s'exprime dans la durée de PS2, plus réduite pour [eɪə] (1,39 %), que pour [ɔɪə] (8,93 %) (voir tableau n°12). Dans [eɪə], la première cible [e] est proche de la seconde cible [ɪ], ce qui réduit le trajet articulatoire pour aller de l'une à l'autre. Une fois la seconde cible atteinte, la trajectoire s'infléchit immédiatement en direction de la troisième cible. En revanche, dans la triphthongue [ɔɪə], la distance articulatoire entre les deux premières cibles est plus élevée. La trajectoire s'attarde donc un peu plus longtemps sur la cible centrale, avant de s'infléchir en direction de la troisième cible. On relève pour ces deux triphthongues de grandes phases de transition entre les cibles (DTC1 : 21,67 % et DTC2 : 18,88 % pour [eɪə] ; DTC1 : 18,95 % et DTC2 : 15,25 % pour [ɔɪə]), ce qui rend ces sons très instables au niveau articulatoire, mais aussi en ce qui concerne leur timbre.

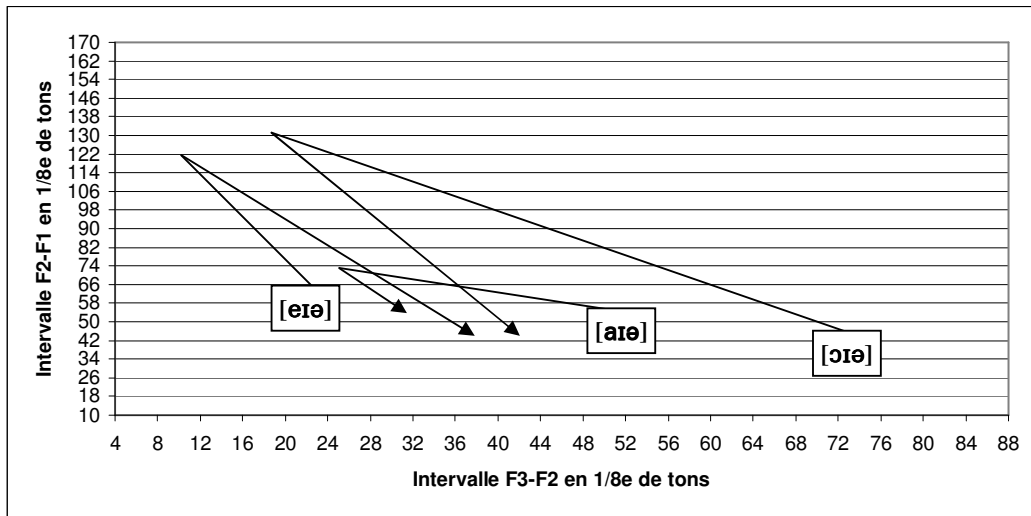


Figure n°75 : trajectoires des triptongues prononcées par la locutrice anglophone.

Comparée à ces deux diphtongues, la trajectoire de [aɪə] est plus simple. En fait, la cible centrale est « effleurée », pour reprendre les termes de Ginésy (1995 : 44), la locutrice n'étant pas allée assez loin dans l'espace articulatoire pour atteindre la cible (phénomène d'« undershooting »). En effet, [ɪ] est prononcé avec une ouverture trop grande, sa valeur de F1 étant identique à celle de [a] (714 Hz). En conséquence, la durée de la transition (TR1) entre la première et la seconde cible est réduite (1,39 %). En revanche, la proportion de stabilité pour cette triptongue est plus grande que dans les deux autres triptongues (52,89 % de stabilité, contre 35,18 % pour [eɪə] et 35,29 % pour [ɔɪə], la stabilité étant calculée en additionnant le pourcentage les trois PS), la locutrice s'attardant davantage sur chacune des cibles.

En ce qui concerne le timbre des « cibles », il coïncide avec celui des voyelles désignées par les symboles phonétiques utilisés dans la transcription phonétique des trois triptongues étudiées, à l'exception de [ɪ] dans la triptongue [aɪə].

	DT	DPS1	DTC1	DPS2	DTC2	DPS3	DQ	Durée par rapport à la syllabe
[eɪə]	12,12 %	8,15 %	21,67 %	1,39 %	18,88 %	25,64 %	12,12 %	75,78 % (429 ms) CV
[aɪə]	3,96 %	13,51 %	1,39 %	10,95 %	17,48 %	28,43 %	24,24 %	78,09 % (429 ms) CV
[ɔɪə]	16,55 %	5,01 %	18,95 %	8,93 %	15,25 %	21,35 %	13,94 %	74,00 % (459 ms) CV

Tableau n°12 : durée relative des événements composant les 3 triptongues prononcées

par la locutrice anglophone (DT pour durée totale, DPS1 pour durée de la première partie stable, DTC1 pour durée de la première transition centrale, DPS2 pour durée de la deuxième partie stable, DTC2 pour durée de la deuxième transition centrale, DPS3 pour durée de la troisième partie stable et DQ pour durée de la queue).

3.2.2.3. Difficultés rencontrées par les apprenants francophones dans la restitution des diphtongues et des triptongues de l’anglais

Les diphtongues et triptongues spécifiques de l’anglais sont très difficiles à reproduire par les francophones. La vraie diphtongaison s’explique physiologiquement par une inaptitude à tenir les organes en place pendant l’articulation d’une voyelle longue. Or, nous avons vu que le français se caractérisait par une grande tension articulaire. Un francophone a donc plutôt tendance à produire une succession de deux ou trois voyelles (avec des zones de transition centrale très réduites), plutôt qu’une diphtongue ou une triptongue, car son articulation tendue l’empêche de réaliser un glissement progressif : au lieu de passer graduellement d’un son à l’autre, il a tendance à sauter de l’un à l’autre, comme il le fait dans sa LM lorsqu’il rencontre un hiatus¹² (Faure, 1948 : 30). Il peut même aller jusqu’à simplifier et ne produire qu’une seule voyelle dans le cas de la diphtongue et deux voyelles dans le cas de la triptongue.

¹² Comme dans le mot « *pays* » par exemple. Faure (1948 : 30) parle d’ailleurs de « fausses diphtongues ».

La rythmique française appliquée à l'anglais renforce non seulement la tension articulatoire, mais elle empêche le locuteur francophone de produire correctement le fameux « stress pulse » dans les diphtongues, à cause de l'introduction d'un accent primaire sur la dernière syllabe (allongement final). Comme le souligne Faure (1948 : 30), « *le défaut de beaucoup de français consiste à donner trop de relief à la fin des diphtongues* ».

Dans les triptongues, l'allongement final a comme conséquence principale l'insertion d'une coupe syllabique après la première cible. Ainsi, un énoncé comme « soya » a tendance à être prononcé de la manière suivante : [sɔ-ɪə].

Enfin, le crible phonologique contribue également à altérer la production des diphtongues et des triptongues. Le timbre des cibles a tendance à être assimilé à celui des voyelles françaises de timbre correspondant. Or, si les cibles sont mal localisées, c'est toute la trajectoire de la diphtongue ou de la triptongue qui est bouleversée. Par ailleurs, dans les productions des francophones, la trajectoire a tendance à glisser trop loin, jusqu'à la semi-voyelle [j] dans les diphtongues et les triptongues contenant la cible [ɪ]. Dans les triptongues, cette tendance est renforcée par l'insertion d'une coupe syllabique après la première cible. Un récapitulatif de ces erreurs est exposé dans le Volume II (annexe n°3.4.).

Conclusion

A partir de la description comparée de l'anglais et du français présentée dans ce chapitre, un catalogue d'interférences a été réalisé dans lequel figure la liste des principales erreurs prosodiques (Volume II, annexe n°3.5.1.) et vocaliques (Volume II, annexe n°3.5.2.) produites par les francophones en anglais. Cette description nous a également permis de créer deux batteries de tests de perception qui nous serviront d'outil d'évaluation des enfants de notre corpus (chapitre VII) et elle constitue un document de référence qui nous permettra d'analyser leurs productions en anglais (chapitre VIII). Nous avons conscience que notre description du système vocalique de l'anglais en fonction de celui du français est fondée uniquement sur l'analyse de deux locutrices. De plus, les productions de la locutrice anglophone sont souvent exagérées, car elle est professeur d'anglais (l'exagération ayant une fonction pédagogique dans la « coopération exolingue » entre l'enseignant de langue et l'apprenant ; se reporter au chapitre II, 2.4.2.1.). Mais nous cherchions une production « modèle » sur laquelle fonder notre

description et nous avons pensé que celle d'une locutrice native et professeur d'anglais constituait le meilleur modèle qui soit, d'autant qu'il correspond au type de langage dans lequel sont baignés les apprenants (R.P., B.B.C. English). Par ailleurs, les enfants de notre corpus sont débutants en langue, d'où la nécessité de limiter la complexité de la description, notamment en ce qui concerne la prosodie (en début d'apprentissage, les productions des enfants se limitent le plus souvent à des mots et à quelques phrases courtes).

Bibliographie chapitre V

- Adamczewski, Henri et Keen, Denis (1973). *Phonétique et Phonologie de l'Anglais Contemporain*. Paris : Armand Colin, 253 p.
- Astesano, Corinne (2001). Rythme et Acentuation en Français : Invariance et Variabilité. Paris : L'Harmattan, 317 p.
- Baken, Ronald J. and Daniloff, Raymond G. (1991). *Readings in Clinical Spectrography of Speech*. San Diego: Singular, 566 p.
- Beckman, Mary (1992). Evidence for speech rhythms across languages. In Tohkura, Vatikiotis-Bateson and Sagisaka, 457-463.
- Best, Catherine, T., McRoberts Gerald W. and Sithole, Nomathemb N. (1988). The phonological basis of perceptual loss of nonnative contrasts: maintenance of discrimination among Zulu clicks by English-speaking adults and infants. In *Journal of Experimental Psychology: Human Perception and Performance*, 14, 345-360.
- Best, Catherine, T. (1993). Emergence of language-specific constraints in perception of non-native speech a window on early phonological development. In Boysson-Bardies, Schonen, Jusczyk, MacNeilage and Morton, 289-304.
- Bot, K. de (1986). The transfer of intonation and the missing data base. Kellerman, E. and Smith, M. S., Crosslinguistic Influence. In *Second Language Acquisition*, 110-119.
- Carton, Fernand (1974). *Introduction à la Phonétique Française*. Paris : Bordas (1ère édition), 253 p.
- Catford, John C. (1988). *A Practical Introduction to Phonetics*. Oxford University Press.
- Chomsky, Noam and Halle, Morris (1968). *The Sound Pattern of English*. New-York: Harper and Row, 470 p.
- Dauer, Rebecca M. (1983). Stress-timing and syllable-timing reanalysed. In *Journal of Phonetics*, 11, 51-62.
- Delattre, Pierre (1961). La leçon d'intonation de Simone de Beauvoir. Etude d'intonation déclarative comparée. Dans *French Review*, 35 (1), 59-67.
- Delattre, Pierre (1965). *Comparing the Phonetic Features of English, French, German and Spanish*. Heidelberg: Julius Gross Verlag, 118 p.
- Di Cristo, Albert (1999). Le cadre accentuel du français contemporain : essai de modélisation. Dans *Langues*, 3/2, 184-205.
- Dodane, Christelle (2001). Les nouvelles technologies multimédias dans l'apprentissage de la prosodie d'une langue étrangère : une révolution dans la pratique ? *Actes du VIIe Congrès International de Communication Sociale*, Santiago de Cuba, Vol. I, 176-182.
- Faure, Georges (1962). *Recherches sur les Caractères et le Rôle des Eléments Musicaux dans la Prononciation*

- Anglaise* (Essai de Description Phonologique). Thèse de Doctorat d'Etat. Paris : Didier, 380 p.
- Faure, Georges (1948). *Manuel Pratique d'Anglais Parlé*. Paris : Hachette, 211 p.
- Fonagy, Yvan (1980). L'accent en français : accent probabilitaire. Dans Fonagy et Léon, 123-233.
- Fonagy, Ivan and Léon, Pierre R. (1980). *L'Accent en Français Contemporain*. Studia Phonetica, 15.
- Ginesy, Michel (1995). *Mémento de Phonétique Anglaise*. Paris : Nathan Université, 288 p.
- Grover, C., Jamieson, D. G. and Dobrovolsky, M. B. (1987). *Intonations in English, French and German: Perception and Speech*, 30/3, 277-295.
- Guierre, Lionel (1984). *Drills in English Stress-Patterns*. Paris : A. Colin-Longman, 127 p.
- Guimbretière, Elisabeth (1994). *Phonétique et Enseignement de l'Oral*. Paris : Didier / Hatier, 96 p.
- Hirst, Daniel (1998). Intonation in British English. In Hirst and Di Cristo, 56-78.
- Hirst, Daniel and Di Cristo, Albert (1984). French intonation: a parametric approach. Dans *Die Neueren Sprachen*, 83/5, 554-569.
- Hirst, Daniel and Di Cristo, Albert (1998). *Intonation Systems: a Survey of Twenty Languages*. Cambridge: University Press, 487 p.
- Institut National de la Statistique et des Etudes économiques (2001). *Note d'Information 01-18*. Edu050, avril.
- Johnson K., Ladefoged P. and Lindau M. (1993). Individual differences in vowel production. In *Journal of the Acoustical Society of America*, 94, 701-714.
- Konopczynski, Gabrielle (1990). *Le Langage Emergent : Caractéristiques Rythmiques*. Hamburg : Buske Verlag, 362 p.
- Konopczynski, Gabrielle (1991). Acquisition de la proéminence dans le langage émergent. Dans *Actes du XIIème Congrès International des Sciences Phonétiques*. Aix-en-Provence, Vol I, 333-337.
- Lacheret-Dujour Anne et Beaugendre Frédéric (1999). *La Prosodie du Français*. Paris : C.N.R.S. Langage, 353 p.
- Ladd, Robert (1996). *Intonational Phonology*. Cambridge: Cambridge University Press.
- Ladefoged, Peter (1975). *A Course in Phonetics*. New York: Harcourt Brace Jovanovitch, Inc., 296 p.
- Ladefoged, Peter (1997). Linguistic Phonetic Descriptions. In Hardcastle and Laver, 589-618.
- Lado, Robert (1957). *Linguistics Across Cultures: Applied Linguistics for Language Teachers*. Ann Arbor, Michigan, 141 p.
- Lanchec, Jean-Yvon (1976). *Psycholinguistique et Pédagogie des Langues*. Paris : PUF, 164 p.
- Landercy, Albert et Renard, Raymond (1977). *Eléments de Phonétique*. Bruxelles : Didier, 214 p.
- Lehiste Ilse and Peterson Gordon E. (1960). Transitions, glides, and diphthongs. In Baken and Daniloff (1991), 286-295.
- Léon, Pierre (1992). *Phonétisme et Prononciation du Français*. Paris : Nathan, 192 p.

- Léon, Pierre et Martin, Philippe (1971). Linguistique appliquée et enseignement de l'intonation. Dans *Etudes de Linguistique Appliquée*, 3, 36-45.
- Lhote, Elizabeth (1987). *A la Découverte des Paysages Sonores des Langues*. Besançon : Annales Littéraires de l'Université, 215 p.
- Lhote Elisabeth (1995). *Enseigner l'Oral en Interaction*. Paris : Hachette, 158 p.
- Llorca, Régine (1987). Le paysage français pour un Français : les bases rythmiques du français. In Lhote, 47-58.
- Lonchamp, François (1983). On Vowel Normalisation. *Leeds Experimental Phonetics Symposium*, Leeds.
- Meschonnic, Henri (1982). *Critique du Rythme*. Paris : Verdier, 713 p.
- Miller J. D. (1989). Auditory-perceptual interpretation of the vowel. In *Journal of Acoustical Society of America*, 85, 2114-2134.
- Montreuil, Jean-Pierre (2001). *La Phonologie de l'Anglais*. Rennes : Presses Universitaires de Rennes, 168 p.
- Ostiguy, Luc, Sarrasin, Robert et Irons, Glenwood (1996). *Introduction à la Phonétique Comparée : les Sons. Le Français et l'Anglais Nord-Américain*. Laval : Presses de l'Université de Laval, 200 p.
- Pasdeloup Valérie (1991). *Modèles de Règles Rythmiques du Français Appliqué à la Synthèse de la Parole*. Thèse de doctorat, Université de Provence.
- Peterson Gordon E. (1961). Parameters of vowel quality. In *Journal of Speech and Hearing Research*, 4, 10-29.
- Peterson, Gordon E. and Barney, Harold L. (1952). Control Methods used in a study of vowels. In *Journal of the Acoustical Society of America*, 24, 175-184.
- Pike, Kenneth L. (1945). *The Intonation of American English*. Ann Arbor: University of Michigan Press, 203 p.
- Potter R. K. and Steinberg J. C. (1950). Towards the specification of speech. In *Journal of Acoustical Society of America*, 22, 807-820.
- Pritchard, R. M. O. (1985). *The teaching of French intonation to native speakers of English*. In *International Review of Applied Linguistics*, 23/2, 117-147.
- Rossi, Mario (1985). L'intonation et l'organisation de l'énoncé. Dans *Phonetica*, 42, 135-153.
- Rossi, Mario (1999). *L'Intonation, le Système du Français : Description et Modélisation*. Paris : Ophrys, 237 p.
- Schwartz J. L. (1987). *Représentation Auditive des Spectres Vocaliques*. Thèse de Doctorat d'Etat, Université de Grenoble.
- Scott, Thomas (1987). Le français face au paysage anglais. Dans Lhote, 131-142.
- Tohkura, Vatikiotis-Bateson and Sagisaka (1992). *Speech Perception, Production and Linguistic Structure*. Tokyo: IOS Press.
- Troubetzkoy, Nicolas S. (1949). *Principes de Phonologie*. Paris : Klincksieck, 396 p. Traduit par Cantineau, J. (1986).
- Wenk, and Wioland, François (1982). Is French really syllable-timed ? In *Journal of Phonetics*, 19.

Yaeger, Malcah (1979). *Context Determined Variation in Montreal French Vowels*. University of Pennsylvania, Ph.D, 353 p.