
SWANS, un système auteur de synchronisation et d'annotation pour un apprentissage multimodal des phénomènes accentuels en langue vivante 2.

Partie 1 : enjeux et positionnement théorique

Aryel Beck[?], Nicole Décuré[?], Pascal Gaillard[?], Nabil Kabbaj[?], Anne Péchou[?], Anthony Stenton[?], Saïd Tazi[?], Antoine Toma^{*}, André Tricot[•] et Christine Vaillant-Sirdey[?]

[?] Université Toulouse I, UFR Informatique ; Place Anatole France – 31 042 Toulouse Cedex aryelbeck@yahoo.fr et Kabbaj@insa-tlse.fr

[?] Laboratoire Interuniversitaire de Recherche en Didactique des Langues (LAIRDIL) Université Toulouse III, Département des Langues et Civilisations; Université Toulouse I Place Anatole France – 31 042 Toulouse Cedex {anthony.stenton, anne.pechou}@univ-tlse1.fr

[?] « Hypermédias et apprentissages » ERT 34 - IUFM de Midi-Pyrénées et Laboratoire Jacques Lordat, IFR 96 « Sciences du Cerveau » et Université de Toulouse le Mirail, 5 Allées Antonio Machado - 31 058 Toulouse cedex pascal.gaillard@univ-tlse2.fr

[•] « Hypermédias et apprentissages » ERT 34 et Laboratoire Travail et Cognition, UMR 5551 CNRS – EPHE – Université de Toulouse le Mirail andre.tricot@toulouse.iufm.fr

[?] LAAS-CNRS Avenue Colonel Roche 31077 Toulouse Cedex 4, et Université Toulouse I. tazi@laas.fr

^{*} LAIRDIL, Université Toulouse III, Département du Monde Anglophone, Université de Toulouse le Mirail, toma@univ-tlse2.fr

[?] LAIRDIL, IUT Université Toulouse III, 115 Route de Narbonne -31077 Toulouse Cedex 4. christine.vaillant-sirdey@iut-tlse3.fr

RÉSUMÉ : Une des difficultés de l'enseignement de l'anglais concerne la prosodie, notamment l'accent tonique. Les apprenants français éprouvent des difficultés à encoder cette information sur le mot, peu pertinente pour eux. Ils ont par conséquent des difficultés à produire cet accent tonique quand ils parlent. L'objet de cet article est d'analyser les processus qui sous-tendent

cette difficulté et de recenser les pratiques didactiques et les outils informatiques qui tentent d'y remédier. Cette analyse nous permet de proposer quelques principes qui guident le développement d'un système auteur que nous appelons SWANS. Ce système doit permettre à des enseignants d'anglais de générer semi-automatiquement des documents où la syllabe accentuée est codée visuellement et auditivement.

MOTS-CLÉS : accent tonique, prosodie, apprentissage de l'anglais, didactique des langues, double-codage, multimédia, multimodalité, système auteur

1. Introduction

Les langues anglaise et française partagent un important lexique ; dans plus de 70 % des cas, les formes orthographiques d'un même mot dans les deux langues sont proches. Pourtant, les systèmes accentuels mis en place par les deux langues pour ces mots les rendent opaques à l'oral pour les apprenants. En écoutant des locuteurs natifs, francophones ou anglophones, on observe que certaines syllabes sont plus aisément « audibles » que d'autres. On parle dans ce cas là de syllabes « accentuées ». Mais, alors que le français comporte un accent fixe portant sur la fin du mot, l'anglais dans le cas de mots polysyllabiques, le place de façon irrégulière. A l'accent principal, fortement marqué, peut aussi s'ajouter un accent secondaire, moins marqué. Corrélativement, les syllabes non-accentuées du français ne sont pas neutralisées, en ce sens qu'elles gardent une longueur certaine, un timbre constant et une intensité régulière, alors qu'en anglais on trouve dans de nombreuses syllabes inaccentuées une faible intensité et un timbre réduit comme contrepoint aux syllabes accentuées.

L'apprenant francophone se trouve ainsi confronté à deux difficultés : percevoir, lors de la phase d'écoute, les syllabes accentuées et les syllabes non-accentuées, et reproduire, lors de la phase de production, un contraste suffisant entre syllabes accentuées et non-accentuées. Le but de notre recherche est d'améliorer la perception, la mémorisation et la production des syllabes en anglais à l'aide d'un système auteur que nous appelons SWANS, qui permet de générer de façon semi-automatique des documents où la syllabe est codée deux fois : visuellement et auditivement.

L'analyse des schémas accentuels d'une langue présente des problèmes fondamentaux de segmentation. Comment distinguer l'accent lexical de l'accent de phrase ? Les deux phénomènes sont imbriqués dans un flux sonore qu'il convient de décortiquer soigneusement. Nous visons une séparation des phénomènes pour les besoins didactiques, mais nos hypothèses sont forcément construites à la fois au niveau segmental et supra-segmental.

Dans cet article, nous ne traiterons que de l'accent lexical. Nous présentons les fondements didactiques, linguistiques et psycholinguistique de notre démarche. Nous recensons les outils existants, pour aboutir aux propositions qui ont guidé le développement de SWANS. Dans un second article, nous présentons l'architecture

informatique de SWANS et une première évaluation de l'efficacité sur la mémorisation de l'accent tonique par des apprenants.

2. Difficultés de perception et de production de l'accent tonique chez le locuteur français

Les étudiants français, lors de la réalisation d'exposés oraux peuvent, malgré une langue assez correcte sur le plan lexical et syntaxique, faire preuve de sérieuses lacunes aux niveaux phonétique et prosodique. L'absence de discrimination voyelle/diphthongue et les déplacements d'accents toniques rendent certains mots méconnaissables, de sorte que ces exposés seraient probablement difficiles à comprendre pour des non-francophones. En bref, ces prestations orales ne présentent qu'une ressemblance lointaine avec la langue cible. Des études empiriques contrôlées confirment ce que les enseignants observent de façon quotidienne. [GUPTA & MERMELSTEIN 82] ont montré dans une étude expérimentale que des canadiens anglophones reconnaissaient moins bien des mots isolés prononcés par un canadien francophone que par un canadien anglophone. Les auteurs attribuent cette différence à un défaut d'accentuation. [HAWKINS & WARREN 94] ont montré que ce type d'information phonétique, sur certains mots-cibles, peut jouer un rôle plus important dans la compréhension du discours que le fait même d'avoir été exposé ou non à ces mots-cibles.

[PEPERKAMP 01] qui a testé les contrastes accentuels parmi les locuteurs de plusieurs langues - hongrois, finlandais et polonais - parle de la « surdité accentuelle » des français qui, contrairement aux utilisateurs de langues à accentuation irrégulière, n'ont pas besoin de stocker d'informations concernant l'accentuation des mot, dans leur langue maternelle, celle-ci étant pour eux prévisible. Sans doute, l'utilisation du mot « surdité » est excessive. [DUPOUX et al. 97] montrent en effet que si les Français, par rapport aux Espagnols, ont des difficultés à percevoir l'accent tonique, ils n'en ont pas à distinguer la différence entre un mot trisyllabique sans accent (*fidape*) et le même mot avec accent (*fidàpe*). Le niveau de performance est le même que dans une tâche où ils doivent distinguer deux mots qui diffèrent d'un phonème (*fidape*, *lidape*). Ces auteurs ont élaboré une autre expérience dans laquelle les participants devaient juger si deux mots d'une langue étrangère étaient identiques, sans tenir compte de l'accent tonique. Les Espagnols ne parvinrent pas à en faire abstraction, tandis que les français lignoraient sans problème. Nous formulons, avec Dupoux et ses collègues, l'hypothèse selon laquelle le traitement en question n'est pas sensoriel, mais se situe plutôt au niveau de la mémoire de travail : il s'agit en quelque sorte d'une « négligence » lors de l'encodage de l'information. Les étudiants français qui apprennent l'anglais négligent de traiter l'accent tonique, et par conséquent ils ne stockent pas cette information qui a peu de valeur dans leur langue maternelle. Les preuves ne manquent pas en psychologie de la perception pour attester qu'une information sans valeur n'est, la plupart du temps,

pas encodée par les êtres humains, comme l'ont montré [MYLES-WORSLEY et al. 88] dans le domaine du traitement d'images radiologiques par des radiologues experts.

Les étudiants français ont d'ailleurs du mal à comprendre la pertinence de cet accent lorsqu'on la leur explique, car la notion même d'accent tonique leur est étrangère. Lors des exposés, ils placeront donc l'accent tonique de façon aléatoire sur l'une des syllabes du mot anglais *development* indiquant ainsi non pas une « surdité » ou un problème de production, mais une négligence au niveau de l'encodage et une absence de stockage de la place de cet accent tonique.

En résumé, les francophones (a) n'encoderaient pas l'accent tonique en anglais. Ils ont par conséquent du mal (b) à mémoriser la place de l'accent tonique et (c) à le produire ou à le reproduire, ce qui (d) a un effet négatif sur la compréhension de leur discours par des interlocuteurs anglophones.

Amener les étudiants à distinguer, encoder, mémoriser et produire l'accent tonique est donc un enjeu didactique majeur pour l'enseignement de l'anglais tout au long du parcours scolaire, et spécialement à l'université, qui constitue en quelque sorte la dernière chance.

On peut raisonnablement faire l'hypothèse que les processus attentionnels jouent ici un rôle fondamental. L'attention recouvre en effet un ensemble de processus qui organisent les traitements perceptifs, mettant en exergue tel signal présent dans l'environnement, inhibant tel autre. Notre problème didactique est donc d'amener les étudiants français à encoder de façon contrôlée et attentive des caractéristiques phonétiques du mot anglais qu'ils traitent de façon automatique, sans y faire attention, dans leur propre langue à partir du moment où ils savent que la nature même de ces caractéristiques (variations d'intensité, de timbre et de longueur des syllabes à l'intérieur du mot) est différente en anglais et en français.

3. Problèmes d'apprentissage de la représentation phonologique du mot

L'apprentissage de la place de l'accent tonique dans les langues où sa place est irrégulière commencerait au cours des 6 premiers mois de la vie et, dès l'âge de 10 mois, l'enfant semble distinguer l'accent tonique d'un mot de sa langue de celui d'une langue étrangère [CURTIN et al. 04]. Dans ce type de langue, la place de l'accent tonique ne peut être apprise à partir de règles mais avec le mot lui-même, si bien que certains mots peu fréquents de la langue entraînent des erreurs d'accentuation chez des adultes natifs [BURANI & ARDUINO 04]. Ainsi, dans les langues à accentuation irrégulière, l'apprentissage de la place de l'accent tonique serait précoce, inhibiteur, implicite et lexical, de type strictement déclaratif.

Selon les modélisations classiques de l'organisation lexicale en mémoire chez le sujet bilingue [DE GROOT 98], nous disposerions de deux registres pour stocker la forme des mots : le lexique L1 et le lexique L2. Chacun de ces deux registres stockerait les aspects phonologiques et orthographiques des mots. Ces deux lexiques seraient

associés à un registre conceptuel, où serait stockée la signification des mots. Selon le modèle de représentation subordonnée, qui correspondrait à l'organisation du lexique chez l'apprenant débutant, le lexique L1 domine le lexique L2, de sorte qu'un accès au registre conceptuel se fait par la médiation du lexique L1. Selon le modèle de médiation conceptuelle, qui correspondrait à l'organisation du lexique chez le bilingue, l'accès au registre conceptuel pourrait être réalisé directement depuis L2, les significations servant de médiateurs de L2 vers L1 et réciproquement. En début d'apprentissage, des liens locaux commenceraient à se développer au sein de L2, item par item, puis classe par classe. Viendrait ensuite une phase d'autonomisation progressive du second avec un développement syntaxique (voir [PERDUE & GAONOAC'H 00] pour une synthèse récente).

Pour l'apprenant français de la langue anglaise, la transparence orthographique du lexique anglais pourrait constituer une sorte de piège. La représentation orthographique du mot anglais peut fréquemment être appariée à sa représentation en français et déclencher l'activation de la représentation conceptuelle pertinente. Ainsi, dans un enseignement où les mots écrits sont très présents, l'apprenant peut se passer sans encombre, pendant les premières années d'apprentissage, de l'élaboration d'une représentation phonologique correcte, que ce soit pour les phonèmes qui composent le mot où l'accent tonique. C'est après des années, quand l'élève avancé se trouve en situation de compréhension ou de production de discours oraux « véritables » que les lacunes phonologiques constituent une gêne.

Ces modèles classiques ont été discutés, remis en cause et finalement profondément modifiés dans le cadre des études psycholinguistiques du bilinguisme. Nous ne discuterons pas ici de ces modèles, mais nous référerons à l'enseignement.

Par exemple, la méthode des mots clés (*keywords*) en enseignement d'une L2 utilise la technique suivante : on propose à l'apprenant d'associer l'item cible en L2 à un item homophone ou quasi homophone en L1 (le mot clé) ainsi qu'à une image représentant ce mot clé avec une définition de l'item cible en L2. Imaginons par exemple un francophone apprenant l'anglais. L'item cible *beach* (la plage) pourrait être associé au mot clé français *biche*, l'image représentant une biche se promenant sur une plage. Cette méthode semble incompatible avec le modèle de médiation conceptuelle, puisque le mot clé semble proposer une double association conceptuelle dépourvue de sens, et avec le modèle de représentation subordonnée puisque l'item cible en L2 est lié avec un mot clé en L1 qui ne lui correspond pas au niveau sémantique et ne peut donc pas activer la bonne représentation conceptuelle. Dès les années 70, la méthode des mots clés a été critiquée, preuves empiriques à l'appui. Des auteurs comme par exemple [ATKINSON 75] ont montré qu'elle ne fonctionnait bien qu'avec des apprenants débutants et pour des apprentissages explicites. Pourtant, même si les résultats de cette méthode ne sont bons qu'avec des apprenants débutants, le modèle de la représentation subordonnée semble bien incapable de les expliquer. Une recherche récente [LAWSON & HOGBEN 98] montre l'efficacité de la méthode des mots clés avec des apprenants avancés, notamment quand les apprenants doivent eux-mêmes créer leurs mots clés et leurs images. Pour

autant, ces résultats sont incompatibles avec le modèle de médiation conceptuelle. Ils sont en revanche compatibles avec les notions d'élaboration [ANDERSON 00] et de profondeur d'encodage [SCHACTER 99], que l'on trouve dans les modèles récents de la mémoire : un item est d'autant plus facilement rappelé qu'il a fait l'objet d'un codage riche et multiple (et non seulement répété).

Ainsi, pour élaborer la représentation complète d'un mot anglais, un apprenant français doit mettre en relation trois représentations : conceptuelle, orthographique et phonologique. Dans les situations d'apprentissages par immersion, cet apprentissage est réalisé à l'oral et de façon implicite, sans doute parce que cet apprentissage est adaptatif. Dans les situations d'apprentissage par enseignement, nous avons émis l'hypothèse que la présence des mots écrits permet à l'apprenant de se passer de l'élaboration d'une représentation phonologique correcte. L'enseignant a alors deux grandes familles de solutions : amener l'élève à élaborer ou l'amener à enrichir l'encodage. L'élaboration est habituellement promue par des situations où l'élève doit comprendre, élaborer des hypothèses, produire des inférences, pour élaborer par lui-même une représentation de la connaissance. L'enrichissement de l'encodage est la solution que nous avons choisi dans le projet que nous présentons ici.

Nous examinons rapidement différents types de situations didactiques et d'outils pour faire le point sur la façon dont le problème est habituellement traité.

4. Apport de l'ordinateur dans l'apprentissage de la prononciation : quelques solutions existantes

4.1. Le cadre didactique classique

Dans le cas d'un mot présenté à la fois à l'écrit et à l'oral, par un enseignant en cours ou par un dictionnaire vocal comme *Cambridge* (voix humaines), *Dicovox* ou *Babylon*, (voix de synthèse) afin d'en donner une représentation conceptuelle, l'attention de l'apprenant francophone va le plus souvent se porter sur la représentation orthographique du mot. L'apprenant débutant pouvant se passer d'une représentation phonologique correcte, l'enseignement de la phonétique qui jusqu'à présent est la seule façon précise de transcrire sur papier les sons d'une langue au moyen d'un alphabet spécifique international établi, l'IPA (*International Phonetic Alphabet*) est délaissé dans le secondaire. La correction systématique des erreurs en production étant difficile pour des effectifs supérieurs à 15 élèves ou étudiants, il s'ensuit un renforcement des erreurs très préjudiciable à un bon niveau en compréhension orale. En effet lorsque les mots de l'anglais correctement prononcés sont entendus, ils ne sont pas reconnus.

Certains étudiants après 10 ans d'anglais (en Maîtrise Langues Etrangères Appliquées par exemple) ne maîtrisent pas la prononciation de mots qui paraissent

élémentaires comme *who, women, minute, chocolate, village, low, allow, law, son, sun, would, should, could*. Les séries telles que *there're, aren't, weren't, were, where* ou les mots graphiquement proches tels que *tough, trough, though, through, thought, thorough* posent d'énormes problèmes de mémorisation à l'oral.

4.2. Les laboratoires informatiques multimédias

Conscients de ces difficultés, un certain nombre de chercheurs [BERTIN 96 ; CAZADE 95 ; HOCHART 99 ; PERRIN 05 ; STENTON 91 ; TOMA 93] ont essayé depuis le début des années 1990 (date de l'introduction des cartes-son dans les ordinateurs) de développer des fonctions informatiques d'apprentissage donnant à l'apprenant les moyens de trouver les solutions par lui-même. Ce n'est que parce que celui-ci saura qu'il est en possession des outils qui lui permettront d'atteindre ce but qu'il fera l'effort nécessaire. Lorsqu'un didacticiel peut conduire l'apprenant à travers les différents étapes qui l'aideront à passer du son entendu au sens à comprendre, l'effort, librement consenti, sera tel que la mémorisation de la forme orale des mots qui lui ont posé problème se fera au niveau le plus profond. La reproduction de ces mots en sera d'autant meilleure.

Les fonctions réellement didactiques développées excluent dans un premier temps tout recours à la graphie. S'il n'est pas le lieu, ici, de les décrire en détail, il paraît néanmoins utile d'en citer quelques-unes :

– séquençage (ou segmentation) automatique du document sonore ou vidéo en fonction des pauses observées dans le discours¹. Plus la durée choisie pour la pause est courte, plus il y a de séquences créées pour le même document et plus ces séquences sont courtes. C'est parce qu'elles sont courtes que l'attention peut être maximale, ce qui favorise la comparaison des sons entendus avec les sons déjà mémorisés par l'apprenant.

– écoute éclatée avec longueur du temps de pause pour la prise de notes proportionnelle à la durée de chaque séquence. Ce temps de pause permet également la répétition de la séquence en temps limité.

– aides multimodales (soit par coloration-décoloration des séquences, soit par visualisation de la courbe d'amplitude²).

– aides récursives temporelles : la séquence qui pose problème peut être rejouée instantanément puisqu'elle est visualisée ; un clic dans la courbe avant le mot

¹ Seul le logiciel Lavac [TOMA 93] permet, pour l'instant, cette fonction. Il permet également les fonctions suivantes de manière entièrement automatisée.

² Cette courbe qui représente la force du son permet de visualiser avec une précision suffisante les groupes de sons constituant des mots, même si des plosives peu accentuées ont parfois une représentation oscillographique exagérée.

incompris entraîne automatique la lecture du passage à partir de l'endroit du clic ; cette lecture peut de plus être réitérée *ad libitum*.

- mémorisation automatique des passages difficiles propres à chaque apprenant par inscription d'une marque d'index dans le numéro d'une séquence chaque fois que celle-ci est jouée.

- modification de la vitesse du débit de parole. Une écoute inaudible à vitesse normale le devient souvent à 70% de la vitesse. Le temps de traitement de l'information est un facteur trop souvent négligé dans les didacticiels actuels.

- comparaison séquentielle, puis simultanée des enregistrements maître et élève avec visualisation des courbes respectives. Les défauts que l'élève relèvera dans sa production sont dus à un mauvais encodage de l'information sonore et une telle fonction peut favoriser la remédiation du problème grâce la prise de conscience des différences entre les deux productions.

Il est possible de plus dans un deuxième temps de proposer des aides visuelles liées aux différentes séquences. Ces aides de type phonétique, déductif, graphique, ou de traduction peuvent être proposées progressivement et dans cet ordre. Ces fonctions ont toutes pour but de permettre à l'apprenant de mobiliser ses propres ressources pour résoudre par lui-même les problèmes de compréhension qui se posent à la première écoute afin d'atteindre un niveau correct d'appropriation du contenu informationnel. Le temps passé sur un document avant d'accéder à la solution est significatif de l'effort de l'apprenant qui se pose comme un défi le projet de comprendre. Ces efforts permettent un encodage des formes orales à un niveau profond, ce qui entraîne en retour une reproduction beaucoup plus précise de ces formes tant au niveau accentuel que phonétique. L'encodage se fait ici « en interne » par une relation cognitive directe du son entendu au sens exprimé. Les résultats ont été validés par plusieurs années d'expériences collectives [VAILLANT 98].

Cependant si une minorité d'apprenants a la chance de travailler en utilisant ces outils, plus de 90% des apprenants sont dans un cadre didactique classique. Partant du constat que la mémorisation des formes orales des mots a été mal ou pas effectuée pour les raisons déjà citées, il apparaît indispensable de travailler sur une autre voie, complémentaire aux fonctions déjà décrites, mais qui aura cette fois-ci pour but « d'extérioriser » la forme orale d'un mot en supposant son sens déjà connu, et ce sans recours à l'outil phonétique.

4.3. Les outils informatiques de prononciation

En ce qui concerne la phonétique acoustique, les progrès des représentations graphiques sur écran sont spectaculaires depuis quelques années. S'il existe toujours un consensus au sujet de l'apport limité des oscillogrammes (*raw waveforms*) de logiciels tels que *Speaker*, *Tell Me More*, *English Plus* ou *VoiceBook* [CAZADE 99], l'utilisation des courbes de fréquence fondamentale (F0) est devenue plus simple (*WinPitch*, *Speech Analyzer*, *Praat*, *Prosodia*). Elle offre un alignement du texte et des courbes nettement plus précis. Dans une courbe fondamentale les

passages non-voisés ne s'affichent pas, ce qui contraint les concepteurs à les modifier, ou à les lisser, pour les rendre plus faciles à interpréter. Certains programmes très prometteurs ont été abandonnés [DENOEUDE et al. 96] tandis que d'autres ont une utilisation trop restreinte en raison de leur complexité (*WinPitch*, *Praat*) ou à cause d'erreurs d'ergonomie (caractères trop petits, *Prosodia*).

Il s'agit à présent de faciliter l'affichage d'informations mémorables et utiles pour améliorer la production et la segmentation automatique sur la base d'unités reconnaissables. Nos propres recherches dans ce domaine incluent le programme *Sounds Right* [PECHOU & STENTON, 01, 02] qui se base sur les courbes fondamentales pour dessiner des courbes intonatives simplifiées au moyen de flèches extensibles qui s'affichent en dessous le texte. (Fig. 1)

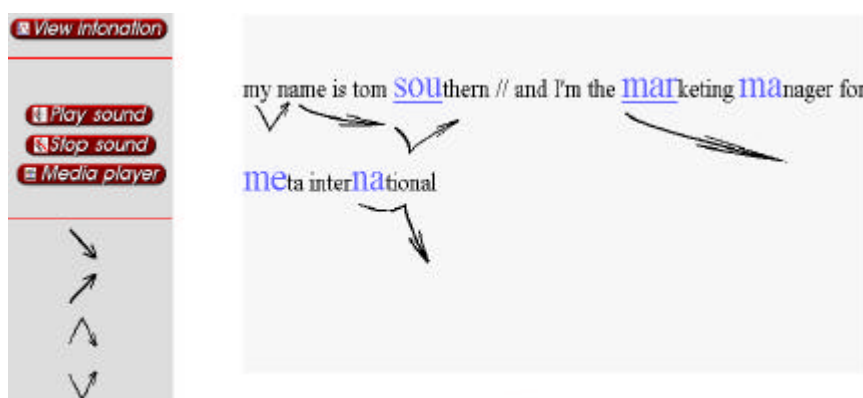


Figure 1. Le système auteur *Sounds Right*. Les flèches extensibles sont traînées en place à partir de la colonne à gauche et ensuite modifiées pour obtenir une taille appropriée. L'annotation textuelle est effectuée par menu. La courbe fondamentale (F0) générée avec *Speech Analyzer* peut être visionnée en cliquant sur le bouton « view intonation ».

L'étude microscopique de l'accentuation et de la prosodie de la langue grâce à ces logiciels se prête à une objection didactique récurrente car, en travaillant sur l'amélioration de la perception auditive, il n'est pas donné d'avance que l'amélioration concernera également la production orale. Dès lors que la production orale est plus aisée et plus confiante, l'exactitude de la prononciation et l'efficacité de la communication prennent le pas parmi les préoccupations pédagogiques. Cette évolution coïncide avec d'autres progrès importants. Les travaux sur la phonétique et la phonologie [GUIERRE 87], [DESCHAMPS 94] et [GINESY 00] ont permis de définir un ensemble de règles d'accentuation et de prononciation et de mettre en évidence ce qui est régulier et ce qui ne l'est pas. Lorsque la pédagogie oscille entre fluidité et précision, le risque de perdre de vue la nécessité de pratiquer la langue courante est bien réel. Les exercices hors contexte atomisent les compétences et ne

rendent pas compte de la maîtrise globale. L'enjeu, en ce qui concerne la didactique de la production orale, reste donc de créer des activités vraiment et non faussement authentiques où la compréhension globale du sens ne soit point négligée.

La difficulté d'interprétation des courbes complexes et celle inhérente à l'incorporation des images de l'écran à un parcours d'activités d'apprentissage authentiques expliquent en partie leur utilisation limitée en cours de langue dans les établissements scolaires et à l'université.

Notre conception d'une solution à la représentation graphique du flux sonore implique donc, non pas les courbes acoustiques, mais plutôt l'utilisation de modifications typographiques. Le débat didactique à ce sujet remonte à plusieurs siècles. La méthode Robertson [HAMILTON 57] au milieu du 19^{ème} siècle avait déjà adopté des annotations qui s'affichaient au-dessus du texte pour indiquer la qualité des voyelles, des sons longs et brefs et la présence de diphtongues. A l'époque, de telles annotations étaient préférées aux modifications typographiques car celles-ci « défigurent » l'orthographe et ralentissent la reconnaissance et la lecture. L'utilisation de textes électroniques constitue cependant une évolution importante car, contrairement à l'environnement noir et blanc des documents papiers, la couleur intervient pour permettre des annotations facilement lisibles, tandis que, selon nos observations, la synchronisation du texte et du son modifie les stratégies traditionnelles de lecture. L'activité d'annotation n'est plus réservée aux experts mais peut devenir en elle-même un processus d'apprentissage utile dans la formation des enseignants aussi bien que celle des apprenants. En fait, l'introduction de la synchronisation fine du texte et du son annonce une métamorphose de l'espace d'apprentissage linguistique, et le système auteur que nous examinons dans le second article constitue une réponse à une double problématique : maîtriser les techniques de synchronisation fine sur ordinateur (trop limitées pour être utiles jusqu'ici, [BICKERTON et al 97]) pour mettre ensuite ces nouvelles possibilités à la portée de non-programmeurs dans le domaine des langues.

5. Conclusion

Apprendre la place de l'accent tonique est un enjeu didactique important. Les solutions visuelles informatiques semblent constituer une bonne solution. Le repérage visuel aide à la discrimination des parties du discours où se posent des problèmes de perception et de compréhension. Le son prend alors une forme visuelle observable dans le temps à l'inverse de sa matérialité première constituée par d'éphémères vibrations de l'air. Voir ce que l'on entend est loin d'être un gadget [CAZADE 99] comme il était permis de le supposer de prime abord. Mais il faut parvenir à ce que cette visualisation soit synchrone, peu coûteuse à générer par des utilisateurs comme les enseignants, mais aussi non interférente et peu coûteuse pour les apprenants. Notre projet est donc de (a) créer un système auteur qui permette à tout enseignant de générer les documents où l'accent tonique est marqué visuellement (b) valider expérimentalement cette approche.

6. Bibliographie

- [ANDERSON 00] Anderson, J.R., *Learning and memory, an integrated view* (2nd ed.), Wiley, New York, 2000.
- [ATKINSON 75] Atkinson, R.C., Mnemotechnics in second-language learning, *American Psychologist*, vol. 30, 1975, p. 821-828.
- [BERTIN 96] Bertin, J.-C., *Learning Lab*, Le Havre, 1996 [application informatique commercialisée, société Learning Labs, Le Havre].
- [BICKERTON et al. 97] Bickerton D., Ginet, A., Stenton, A., Temmerman M., Vaskari T., Final Report of the RAPIDO Project, 91 p., February 1997, University of Plymouth, UK.
- [BURANI & ARDUINO 04] Burani C., Arduino, L.S., Stress regularity or consistency? Reading aloud Italian polysyllables with different stress patterns, *Brain and Language*, vol. 90, 2004, p. 318–325.
- [CAZADE 95] Cazade, A., *Help Yourself*, Paris, 1995 [application informatique non-commercialisée].
- [CAZADE 99] Cazade, A. De l'usage des courbes sonores en apprentissage des langues. *ALSIC*, vol. 2, n° 2, 1999, p. 3-32.
- [CURTIN et al. 04] Curtin, S., Mintz, T.H., Christiansen, M.H. Stress changes the representational landscape: evidence from word segmentation, *Cognition*, sous presse, 2004, p. 1-30.
- [DE GROOT 98] de Groot, A.M.B., La représentation lexico-sémantique et l'accès lexical chez le bilingue, *Psychologie Française*, vol. 43, n° 4, 1998, p. 297-312.
- [DENOEUDE et al 96] Denoeud B., Raquillet M., Foti A., Aider à l'Apprentissage des phénomènes prosodiques en situation d'autoformation, *Colloque Ergo IA*, Biarritz, 1996.
- [DESCHAMPS 94] Deschamps, A. *De l'écrit à l'oral et de l'oral à l'écrit. Phonétique et orthographe de l'anglais*, Orphrys, Paris, 1994.
- [DUPOUX et al. 97] Dupoux, E., Pallier, C., Sebastian, N., & Mehler, J., A destressing "deafness" in French? *Journal of Memory and Language*, n° 36, 1997, p. 406-421.
- [GINESY 00] Ginesy, M., *Phonétique et phonologie de l'anglais*, Ellipses, Paris, 2000.
- [GUIERE 87] Guierre, L. *Règles et exercices de prononciation anglaise*, Longman – A. Colin, Paris, 1987.
- [GUPTA & MERMELSTEIN 82] Gupta, V., Mermelstein, P., Effects of speaker accent on the performance of a speaker-independent, isolated-word recognizer, *Journal of the Acoustical Society of America*, n° 71, 1982, p. 1581-1587.
- [HAMILTON 57] Hamilton H., *The Vicar of Wakefield*, quatre chapitres avec la prononciation figurée d'après la méthode Robertson, Derache, Paris, 1857.

- [HAWKINS & WARREN 94] Hawkins, S., Warren, P., Phonetic influences on the intelligibility of conversational speech, *Journal of Phonetics*, n° 22, 1994, p. 493-511.
- [HOCHART 99] Hochart, J-J., *Genexi*, 1999 [gratuitiel, <http://membres.lycos.fr/jjhochart/>].
- [LAWSON & HOGBEN 98] Lawson, M.J., Hogben, D., Learning and recall of foreign-language vocabulary: Effects of a keyword strategy for immediate and delayed recall. *Learning and Instruction*, vol. 8, n°2, 1998, 179-194.
- [MYLES-WORSLEY et al. 88] Myles-Worsley, M., Johnston, W.A., Simons M.A, The influence of expertise on X-Ray image processing, *Journal of Experimental Psychology: Learning, Memory & Cognition*, vol. 14, 1988, p. 553-557.
- [PECHOU & STENTON 01] Péchou, A., Stenton, A., Encadrer la médiation – le cas de la prononciation, *Groupe d'études et de recherches en anglais de spécialité* (Geras), Bordeaux, Mars 2001.
- [PECHOU & STENTON 02] Péchou, A., Stenton, A., Encadrer la médiation – le cas de l'intonation, *Colloque Compréhension et Hypermédia*, Albi, Octobre 2002.
- [PEPERKAMP 01] Peperkamp, S., Typologie des langues accentuelles : perspectives développementales et données comparatives. *Troisièmes journées internationales du GDR 1954 Phonologie*, Nantes, 30 mai, 1er juin 2001.
- [PERDUE & GAONAC'H 00] Perdue, C., Gaonac'h, D., Acquisition des langues secondes. In M. Kail & M. Fayol (Eds.), *L'acquisition du langage*. Vol. 2. PUF, Paris, 2000.
- [PERRIN 05] Perrin, M., *Didalog* [application informatique, mise en service en 2005].
- [SCHACTER 99] Schacter, D.L. *Searching for memory: The brain, the mind, and the past*. trad. fr. A la recherche de la mémoire, De Boeck, Bruxelles, 1999.
- [STENTON 91] Stenton, A. *Smart Alex*, 1991 [application informatique non-commercialisée] Formation par recherche, rapport sur l'utilisation du logiciel Smart Alex pour l'apprentissage de l'anglais, *Actes des premières journées scientifiques hypermédia et apprentissages*, Chatenay Malabry, INRP, 1991.
- [TOMA 93] Toma, T. *Lavac*, 1993 [application informatique, C puissance 3 Informatique, Morgan Multimedia]
- [VAILLANT 98] Vaillant, C. Les pratiques des étudiants en laboratoire informatique multimédia, enquête réalisée pour l'Association des directeurs des IUT, 1998.

6.1. Références sur le WEB

Babylon : www.babylon.com

Cambridge Advanced Learner's Dictionary (2004) : dictionary.cambridge.org

Dicovox : www.latl.unige.ch

Praat : www.praat.org

Prosodia : www.lpl.univ-aix.fr

Speaker Neuroconcept : www.neuroconcept.com

Speech Analyzer : www.sil.org/computing/speechtools/speechanalyzer.htm

Tell Me More, Auralog : www.auralog.fr

VOICEbook, Englishhear System : www.voicebook.com

Winpitch : www.winpitch.com