

DE L'INNOVATION TECHNOLOGIQUE APPLIQUEE A L'APPRENTISSAGE DES LANGUES

Avec l'avènement de l'informatique et du multimédia, l'apprentissage des langues est devenu de plus en plus facile et interactif.

Auralog s'est, dès sa création, largement investi dans cette aventure technologique et pédagogique, devenant ainsi une référence pour l'enseignement des langues sur CD/DVD-ROM et en solutions Internet.

Evolution de l'apprentissage des langues étrangères à l'oral

L'apprentissage des langues étrangères à l'oral a connu trois stades de développement majeurs.

Au tout début, l'apprenant avait un seul moyen à sa disposition pour se perfectionner à l'oral : reproduire le plus fidèlement possible la **voix de son professeur**.

Puis il a eu la possibilité d'**enregistrer** sa propre voix et de se réécouter. D'analogique l'enregistrement devint numérique, mais malgré une amélioration qualitative, cette technique restait limitée par la perception auditive du sujet.

Au début des années 90, Auralog introduit pour la première fois la **technologie de la Reconnaissance Vocale** dans ses méthodes, permettant ainsi à l'apprenant de dialoguer librement avec l'ordinateur, sans manipulation du clavier ni de la souris, et d'obtenir une évaluation automatique de sa prononciation.

Définition de la Reconnaissance Vocale

La Reconnaissance Vocale désigne l'ensemble des technologies permettant à une machine de reconnaître la parole. Les technologies de Reconnaissance Vocale modernes sont basées sur une analyse des phonèmes de la phrase prononcée.

Auralog fut la première entreprise au monde à adapter cette technologie à l'apprentissage des langues.

Classification de la Reconnaissance Vocale

Trois critères permettent de distinguer les moteurs de Reconnaissance Vocale :

1. Le premier critère est le type de discours reconnu ("word by word" ou "continuous speech") par le logiciel :

- soit les mots sont prononcés **isolément**, c'est-à-dire séparés par des périodes de silence.
- soit le discours se fait **en continu** et permet ainsi une prononciation naturelle.

2. Le deuxième critère distingue les moteurs de Reconnaissance Vocale fonctionnant uniquement sur les **mots**, de ceux fonctionnant sur les **phrases complètes**.

3. Le troisième critère est le degré de dépendance vis-à-vis du locuteur :

- Système monolocuteur : système adapté ou adaptable via un processus d'entraînement à un locuteur particulier ("**speaker dependent**").
- Système multilocuteurs ("**speaker independent**") : Reconnaissance Vocale fonctionnant pour tout un groupe de locuteurs, ce groupe pouvant à l'extrême intégrer n'importe quelle personne parlant la langue apprise (indépendamment de son âge, de son sexe, du timbre de sa voix...).

Les produits Auralog utilisent les moteurs de Reconnaissance Vocale les plus sophistiqués : l'apprenant peut perfectionner sa prononciation sur des mots ou des phrases complètes, au rythme qu'il désire. De plus, les logiciels de type "speaker independent" sont immédiatement utilisables par tous les locuteurs, sans entraînement préalable.

La Reconnaissance Vocale : 20 ans d'évolution technologique

La technologie de la Reconnaissance Vocale a évolué, passant du hardware (cartes) au software. Aujourd'hui, l'apprenant prononce librement une phrase et le logiciel note la qualité de la prononciation par une analyse complexe des phonèmes.

Voici la chronologie des différentes étapes du développement de la Reconnaissance Vocale sur les vingt dernières années :

Début des années 1980

Les premiers systèmes **monolocuteurs** de Reconnaissance Vocale sont appliqués sur des machines industrielles.

1985

C'est environ à cette date qu'apparaissent les premières **cartes** équipées de DSP (Digital Signal Processor), permettant d'implémenter une technologie de Reconnaissance Vocale sur un PC. Ces technologies sont très complexes à mettre en oeuvre (test sur un panel significatif de locuteurs). Ces cartes ne sont pas diffusées au grand public.

Fin des années 1980

Auralog commence à adapter des moteurs de Reconnaissance Vocale à différents marchés, puis plus spécifiquement au marché éducatif des langues.

1991

Auralog crée Aura-lang, première application d'apprentissage de langues fondée sur la Reconnaissance Vocale.

1994

C'est l'arrivée des cartes DSP **grand public** (sous Windows™).

1995

Des **solutions logicielles** issues de la technologie des cartes font leur apparition. Auralog lance **Talk to me**, logiciel grand public sous Windows™, avec un système de Reconnaissance Vocale multilocuteurs fonctionnant sur des phrases complètes. La Reconnaissance Vocale permet à **Talk to me** d'analyser la prononciation de l'apprenant et de l'évaluer grâce à un système de scoring sophistiqué.

1996

La Reconnaissance Vocale est exploitée différemment selon son utilisation : méthode de langues ou logiciel de dictée vocale.

*Parallèlement au marché des langues, les premiers logiciels de **dictée vocale** sont lancés sur le marché. Le principal éditeur est alors DragonSystems. La Reconnaissance Vocale se fait mot à mot.*

1998

Auralog lance **TELL ME MORE**, méthode complète de langues dans laquelle, pour la première fois, la Reconnaissance Vocale est appliquée aux exercices linguistiques.

*Sur le marché de la dictée vocale apparaissent les premiers logiciels autorisant la dictée **en continu** (principal logiciel : IBM ViaVoice).*

1999

Auralog met au point une innovation technologique révolutionnaire, **S.E.T.S.** et l'inclut dans la nouvelle version de **Talk to me**.

2003

Auralog introduit une nouvelle langue dans sa gamme, l'Arabe, qui travaille sur un moteur de reconnaissance Vocale spécifique.

2005

Auralog change de moteur de Reconnaissance Vocale pour un moteur plus performant - Nuance Vocon 3200 - et ajoute dans **TELL ME MORE** de nouvelles activités basées sur la reconnaissance vocale : des dialogues interactifs sur base de vidéos et non plus d'images, et le doublage de séries télévisées.

La technologie S.E.T.S.

La Reconnaissance Vocale fonctionnait de façon continue, sur des phrases complètes et indépendamment du locuteur. Une technologie complémentaire développée par Auralog permet de détecter automatiquement l'erreur de prononciation dans une phrase.

Cette technologie, baptisée **S.E.T.S. (Spoken Error Tracking System)**, a fait l'objet d'un dépôt de brevet par Auralog.

En repérant les points faibles dans la prononciation de l'apprenant, **S.E.T.S.** représente une avancée décisive dans l'apprentissage d'une langue étrangère sur ordinateur. Cette technologie équipe bien sûr les solutions d'apprentissage éditées par la société Auralog.

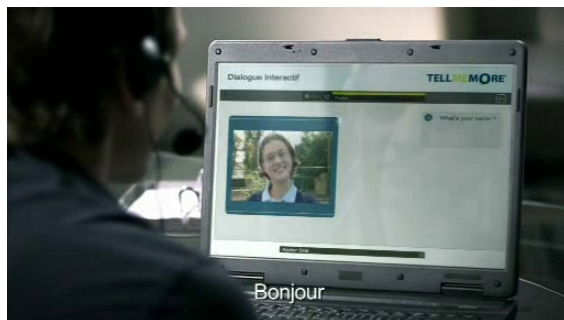
Exploitation de la Reconnaissance Vocale dans les méthodes de langues Auralog

1. Le dialogue interactif

Grâce à la technologie avancée de la Reconnaissance Vocale, l'utilisateur engage un véritable dialogue avec son PC. Un personnage à l'écran parle, l'apprenant répond. Selon son niveau il est amené :

- soit à choisir, parmi celles proposées, une phrase qui lui permettra d'orienter la conversation dans le sens souhaité
- soit à choisir, parmi celles proposées, la seule phrase, qui a un sens par rapport au dialogue.

Si l'ordinateur ne comprend pas la phrase prononcée, il le signale et encourage l'apprenant à travailler la réplique. L'apprenant peut paramétrer la Reconnaissance Vocale pour la rendre plus tolérante ou plus exigeante quant à la qualité de sa prononciation.



2. La prononciation de la phrase ou du mot

L'utilisateur s'entraîne à prononcer une phrase ou un mot et obtient un score lui permettant d'évaluer la qualité de son accent, de sa prononciation et de son intonation.



3. L'affichage de la phrase ou du mot prononcé

Le logiciel offre - en plus de l'affichage d'un score - la possibilité de visualiser précisément sa prononciation et son *intonation*. Deux types de représentations (graphe et courbe) sont disponibles.

Le graphe vocal indique l'amplitude de la voix en fonction du temps (notion d'énergie). Il représente l'intensité sonore de la voix et donne des indications sur la structure de la prononciation.

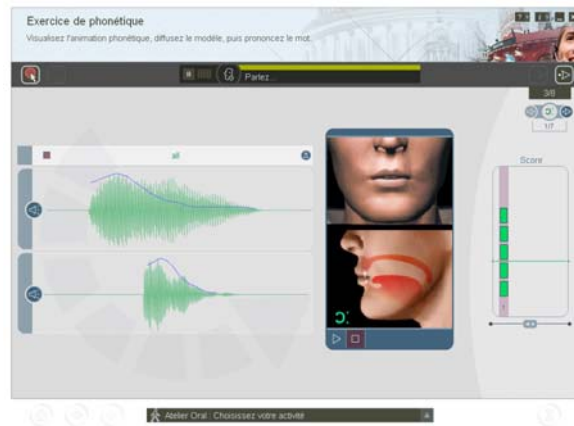
La courbe de la fréquence fondamentale reproduit les variations de la fréquence de la voix en fonction du temps (notion de hauteur). Cette courbe, complémentaire du graphe vocal, permet à l'apprenant de comparer précisément son intonation à celle du modèle (aigu/grave). Cet affichage unique est une innovation développée par Auralog.

Auralog est le seul éditeur à proposer des logiciels qui **évaluent** et permettent de visualiser la **prononciation** et l'**intonation** aussi bien pour des **phrases complètes** que pour des mots.

4. Les animations de phonèmes en 3D

Grâce à des animations phonétiques en images de synthèse, l'apprenant peut visualiser avec une grande précision les mouvements articulatoires pour mieux les reproduire.

Cette technologie d'exception aide à comprendre le fonctionnement articulatoire de la production de certains sons et à les mémoriser, afin de mieux les reproduire.



4. Les exercices qui utilisent la Reconnaissance Vocale :

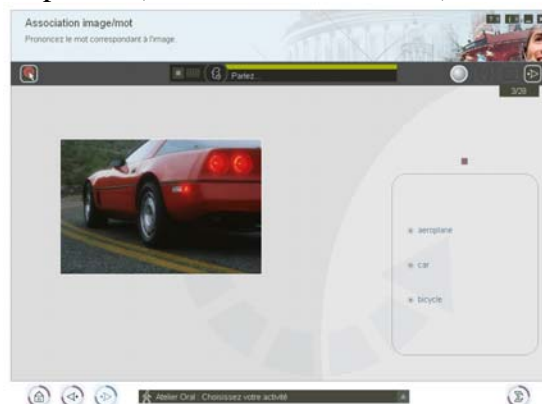
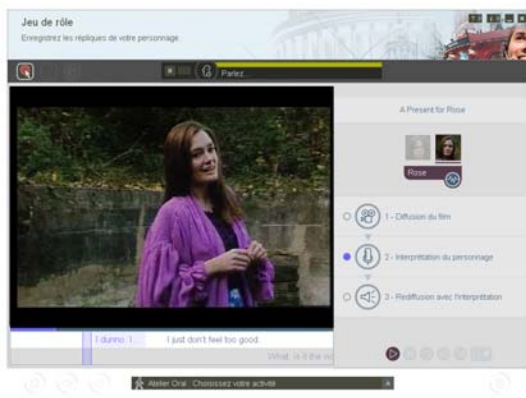
Au-delà des dialogues interactifs et des exercices de prononciation pure, l'utilisateur est amené à utiliser la Reconnaissance Vocale pour réaliser des exercices variés.

La dernière activité intégrée à la méthode est celle qui consiste à doubler des personnages de séries télévisées (activité « jeu de rôles » non présente sur les produits « online ») :

L'apprenant regarde la vidéo de la série : un dialogue entre 2 personnages. Il étudie ensuite les textes du dialogue, choisit un personnage, et travaille la prononciation des répliques. Puis, il peut enregistrer ses répliques, qu'il prononce en même temps que le texte s'affiche et que le personnage parle, comme pour un karaoké. Le film est enfin rediffusé avec la voix de l'apprenant mixée avec la voix du deuxième personnage.

Les autres exercices utilisant la Reconnaissance Vocale sont des exercices de vocabulaire et de grammaire. L'apprenant doit, par exemple :

- Prononcer le bon mot de vocabulaire correspondant à une image (« association image/mot »)
- Remettre oralement dans l'ordre les mots d'une phrase (« remise dans l'ordre »)



•

Les avantages pédagogiques de la Reconnaissance Vocale

Dans l'apprentissage des langues, la composante orale est celle qui est la plus difficile à maîtriser car elle demande beaucoup de pratique.

En proposant à l'utilisateur un bain linguistique où il peut prendre la parole et travailler sur sa prononciation autant qu'il le souhaite, **TELL ME MORE** lui permet d'acquérir plus de confiance en lui et de dépasser rapidement les appréhensions qu'il peut avoir à communiquer en situation réelle.



Glossaire

*Speech recognition / **Reconnaissance Vocale***

Ensemble des technologies permettant à une machine de reconnaître la parole.

*Difficulty level of the speech recognition / **Niveau de difficulté de la Reconnaissance Vocale***

Niveau à partir duquel l'ordinateur accepte un mot ou une phrase comme suffisamment bien prononcé. Les produits Auralog permettent à l'utilisateur de moduler ce niveau, afin d'adapter la difficulté de prononciation.

*S.E.T.S. (Spoken Error Tracking System) / **Technologie S.E.T.S.***

Technologie exclusive d'Auralog permettant de détecter automatiquement l'erreur de prononciation dans une phrase complète.

*Waveform / **Graphe vocal***

Représentation graphique de l'amplitude de la voix en fonction du temps. Ce graphe donne une indication sur la structure de la prononciation.

*Pitch curve / **Courbe de la fréquence fondamentale***

Représentation graphique permettant de visualiser les variations de la fréquence de la voix en fonction du temps et ainsi d'améliorer l'intonation.

*Continuous speech recognition / **Reconnaissance Vocale en continu***

Caractérise les moteurs de Reconnaissance Vocale les plus évolués. Le discours avec l'ordinateur peut se faire en continu.

*Word-by-word speech recognition / **Reconnaissance Vocale mot par mot***

Pour être reconnus par l'ordinateur, les mots doivent être prononcés isolément, c'est-à-dire entrecoupés de silences.

*Speaker-dependent system / **Système monolocuteur***

Système de Reconnaissance Vocale adapté ou adaptable via un processus d'entraînement à un locuteur particulier.

*Speaker-independent system / **Système multilocuteurs***

Reconnaissance Vocale fonctionnant pour tout un groupe de locuteurs, ce groupe pouvant à l'extrême intégrer n'importe quelle personne parlant la langue apprise.

*DSP (Digital Signal Processor) / **Processeur signal***

Processeur permettant d'implémenter une technologie *hardware* de Reconnaissance Vocale dans un PC. Il a été remplacé par des solutions logicielles de Reconnaissance Vocale.