Document generated on 04/16/2024 7:46 a.m.

Documentation et bibliothèques

DOCUMENTATION BIBLIOTHÈQUES

Synthèse vocale : une *technologie adaptée* au service des personnes ayant une déficience visuelle
Speech Synthesis: Technology Serving the Visually Impaired
Síntesis vocal: una tecnología adaptada al servicio de las personas con incapacidad visual

François Papik Bélanger

Volume 43, Number 4, October–December 1997

URI: https://id.erudit.org/iderudit/1032980ar
DOI: https://doi.org/10.7202/1032980ar

See table of contents

Publisher(s)

Association pour l'avancement des sciences et des techniques de la documentation (ASTED)

ISSN

0315-2340 (print)
2291-8949 (digital)

Explore this journal

Cite this article

Papik Bélanger, F. (1997). Synthèse vocale : une *technologie adaptée* au service des personnes ayant une déficience visuelle. *Documentation et bibliothèques*, *43*(4), 181–187. https://doi.org/10.7202/1032980ar

Article abstract

At a time when electronic information is available through computerised networks such as the Internet, barriers continue to exist between information and persons with visual disabilities. The rapid rise of new information technologies has allowed other technologies to emerge; these technologies will allow people with special needs to have access to information. The following article addresses one of these technologies, the voice synthesizer. Special consideration is given to the technological and social context that allowed its development, its operation, and its daily use, especially in libraries.

Tous droits réservés © Association pour l'avancement des sciences et des techniques de la documentation (ASTED), 1997

This document is protected by copyright law. Use of the services of Érudit (including reproduction) is subject to its terms and conditions, which can be viewed online.

https://apropos.erudit.org/en/users/policy-on-use/

érudit

This article is disseminated and preserved by Érudit.

Érudit is a non-profit inter-university consortium of the Université de Montréal, Université Laval, and the Université du Québec à Montréal. Its mission is to promote and disseminate research.

https://www.erudit.org/en/

Synthèse vocale: une *technologie adaptée* au service des personnes ayant une déficience visuelle

François Papik Bélanger*
Diplômé de l'École de bibliothéconomie et des sciences de l'information, Université de Montréal

Au moment où l'information sous forme électronique tend à se rapprocher des gens par l'entremise, par exemple, de réseaux informatiques comme Internet, des barrières continuent de se dresser entre l'information et les personnes ayant une déficience visuelle. Le développement fulgurant des nouvelles technologies de l'information a permis l'apparition de technologies adaptées *destinées notamment à accroître l'autonomie des personnes déficientes ayant des besoins spéciaux pour accéder à l'information. Le présent article s'attarde spécifiquement à une de ces technologies, la synthèse vocale, en mettant l'emphase sur le contexte technologique et social de sa création, sur son fonctionnement et ses composantes et finalement sur son application dans la vie courante et dans les milieux documentaires.*

Speech Synthesis : Technology Serving the Visually Impaired

At a time when electronic information is available through computerised networks such as the Internet, barriers continue to exist between information and persons with visual disabilities. The rapid rise of new information technologies has allowed other technologies to emerge; these technologies will allow people with special needs to have access to information. The following article addresses one of these technologies, the voice synthesizer. Special consideration is given to the technological and social context that allowed its development, its operation, and its daily use, especially in libraries.

Síntesis vocal : una tecnología adaptada al servicio de las personas con incapacidad visual

En el momento en el que la información electrónica se acerca más de la gente con diferentes redes informáticas como Internet, por ejemplo, ciertas barreras siguen elevandose entre la información y las personas con incapacidad visual. El desarrollo fulgurante de las nuevas tecnologías de la información ha permitido la aparición de tecnologías adaptadas destinadas especialmente a aumentar la autonomía de las personas con necesidades particulares para tener acceso a la información. Este artículo trata particularmente de una de estas tecnologías, la síntesis vocal, poniendo el énfasis sobre el contexto tecnológico y social de su creación, sobre su funcionamiento y sus componentes y finalmente sobre su aplicación en la vida corriente y en los medios documentales.

> *Les premières générations d'ordinateurs avaient été éduquées à partir de claviers semblables à ceux de machines à écrire et avaient répondu par l'entremise d'écrans et de téléscripteurs ultra-rapides. CARL pouvait opérer ainsi lorsque cela était nécessaire mais en général, il communiquait avec ses compagnons par la parole. Poole et Bowman s'adressaient à lui comme à un être humain et il leur répondait dans l'anglais idiomatique qui lui avait été enseigné durant son enfance électronique.*
>
> - Arthur C.Clark, *2001: l'odyssée de l'espace*, 1968, p. 129

CARL[1] est une créature informatique issue de l'imagination fertile de Arthur C. Clark. Pour l'instant, il appartient encore au monde de la science-fiction. Mais pour combien de temps encore? Sous la poussée de nouvelles technologies telles que les systèmes experts, les neurones artificiels ou encore la robotique, le mur de la fiction se lézarde, s'effrite et, inexorablement, il est appelé à s'écrouler. Déjà, en regardant à travers les fissures de ce mur, on peut se faire une idée de l'allure que prendront nos CARL du XXIe siècle. Regroupées sous l'appellation de *technologies de la parole*[2], la synthèse vocale, c'est-à-dire la reproduction synthétique de la parole, et la reconnaissance de la parole participent à ce mouvement vers demain. Déjà, la synthèse vocale permet à l'ordinateur de lire *à voix haute* un texte que l'être humain lui a transmis. Cette interface sonore joue un rôle de premier plan pour les personnes qui vivent avec une déficience visuelle et qui éprouvent des problèmes d'accès à l'information textuelle tant sur

* Les recherches pour cet article ont été effectuées dans le cadre du programme de maîtrise en bibliothéconomie et sciences de l'information de l'Université de Montréal. L'auteur tient à exprimer sa gratitude à James Turner, professeur à l'École de bibliothéconomie et des sciences de l'information pour ses nombreux commentaires et surtout pour son encouragement sans faille. Que Thomas Buffin-Bélanger soit également remercié pour son aide précieuse lors de la préparation du manuscrit.

1. CARL est l'acronyme de Cerveau Analytique de Recherche et de Liaison. Dans la version originale anglaise, l'ordinateur s'appelle HAL.

2. L'expression «technologies de la parole» est la traduction de «speech technologies» et de «voice technologies».

support papier que sur support électronique.

Le présent article se veut une introduction à la synthèse vocale, une technologie de l'information venant améliorer l'accès à l'information pour les personnes non voyantes ou semi-voyantes. Au moment même où l'information sous forme électronique tend à se rapprocher de plus en plus des gens, des barrières continuent en effet de se dresser entre l'information et les personnes ayant une déficience visuelle. Dans cette situation, il importe de savoir comment l'accès à cette information peut être amélioré et de mieux connaître la panoplie de technologies qui viennent en aide aux personnes ayant une déficience visuelle. Certains milieux documentaires sont déjà sensibles à cette problématique. Pensons, par exemple, aux centres de documentation s'adressant directement à la clientèle non voyante (Brûlé 1994) ou encore à la création au sein de grandes bibliothèques d'une section venant répondre aux besoins spécifiques des utilisateurs aux prises avec une ou plusieurs déficiences (Joanis 1992). Dans cet article, nous nous attarderons spécifiquement à la synthèse vocale en mettant l'emphase sur le contexte technologique et social de sa création, sur son fonctionnement et ses composantes et finalement sur son application dans la vie courante et dans le milieu des bibliothèques et centres de documentation.

Contexte technologique et social

Les technologies de la parole se trouvent à la croisée de deux grands axes de recherche et d'application. Tout d'abord, compte tenu que l'intelligence artificielle vise comme but ultime «*l'utilisation de l'ordinateur comme moyen de simulation des processus naturels ou comme support de capacités ordinairement attribuées à l'intelligence humaine*» (Vignaux 1992), il est normal que cette discipline intègre les technologies de la parole dans son champ de recherche. Autrement dit, si les chercheurs en intelligence artificielle s'intéressent à la parole, c'est parce que, fondamentalement, «*speech is inextricably intertwined with intelligence itself*» (Lange et al. 1991). Pour qu'un jour des ordinateurs véritablement intelligents puissent, à l'instar de CARL, converser normalement avec l'être humain, des recherches poussées en synthèse et reconnaissance de la parole apparaissent absolument essentielles. Par ailleurs, un autre point de vue sur les technologies de la parole, plus pragmatique celui-là, provient du domaine des *assistive technologies*[3] ou encore *adaptative technologies*[4].On pourrait caractériser ces technologies de la façon suivante: «*any software or hardware addition to the computer that renders it accessible to a person with special needs*» (McNulty and Suvino 1993) . La liste des appareils qui entrent dans la catégorie des technologies adaptées comprend, entre autres, des systèmes qui permettent de verbaliser ou de grossir un texte apparaissant à l'écran d'un ordinateur, des télécripteurs qui permettent aux personnes non entendantes ou malentendantes de communiquer par téléphone, des imprimantes qui produisent du texte en braille, des plages tactiles qui reproduisent en braille les lignes de texte apparaissant à l'écran, des claviers adaptés pour les personnes ayant des problèmes à contrôler leurs mouvements, des systèmes de reconnaissance vocale pour les personnes dont les membres supérieurs sont complètement paralysés, etc. La figure 1 présente un schéma du contexte technologique dans lequel s'inscrivent les technologies de la parole.

Le contexte social, sous-jacent à l'apparition de la synthèse vocale et des autres technologies adaptées, s'est trouvé fortement marqué, du moins en Amérique du Nord, par la volonté sans cesse croissante d'intégrer les diverses minorités dans les milieux de travail et d'éducation. D'aucuns estiment que cette attitude pourtant fort louable dérive parfois vers un simple sens de calcul politique ou provient d'un souci de rectitude politique. D'autres, heureusement plus nombreux, prétendent tout simplement qu'il s'agit d'assurer aux personnes handicapées les mêmes droits qu'aux autres citoyens. Aux États-Unis, les personnes handicapées ont reçu un très fort soutien politique, ce qui a accéléré le développement et la commercialisation des produits de technologies adaptées. Adoptée par le congrès américain en juillet 1990 et venant renforcer la *Rehabilitation Act* de 1973, la loi intitulée *Americans with Disabilities Act* vise essentiellement à interdire toute discrimination basée sur les déficiences physiques ou intellectuelles et à favoriser aux personnes handicapées l'accès aux institutions publiques d'information et d'éducation (McCormick 1994). Cette loi prévoit des subventions aux collèges et aux universités afin que ces institutions puissent offrir l'égalité d'accès à l'information. Ces subventions ont rendu possible, aux États-Unis, l'achat de systèmes de synthèse vocale dans des centaines de bibliothèques publiques et universitaires (Lazzaro 1993). Au Canada, un instrument juridique permet de garantir une certaine forme d'égalité des droits et tente d'enrayer la discrimination physique: la *Charte des droits et libertés*. En outre, au Québec, la *Loi assurant l'exercice des droits des personnes handicapées*, adoptée en 1978, a permis la création de l'*Office des personnes handicapées du Québec*. Cet organisme public a notamment reçu la mission de veiller à la coordination des services dispensés aux personnes handicapées, d'informer et de conseiller les personnes handicapées, de promouvoir leurs intérêts et finalement de favoriser leur intégration scolaire, professionnelle et sociale.

---

3. Dans le domaine des technologies de l'information, les «assistive technologies» constituent un phénomène relativement récent. Ce terme anglais ne se retrouve d'ailleurs pas dans le *Thésaurus : personne handicapée* publié par l'Office des personnes handicapées du Québec en 1996. Quelques traductions pourraient être proposées: «technologies palliatives» (palliatif = qui atténue les symptômes d'une maladie sans agir sur sa cause) ou encore «technologies auxiliaires» (auxiliaire = qui aide par son concours sans être indispensable). Pour ma part, j'utilise dans le présent article le terme «technologie adaptée» qui connaît déjà une certaine diffusion, au Québec, dans les milieux venant en aide aux personnes ayant des déficiences physiques ou intellectuelles.

4. À proprement parler, les «assistive technologies» sont destinées à accroître l'indépendance des personnes ayant des besoins spéciaux, tandis que les «adaptive technologies» ont pour but d'offrir au grand public des environnements plus ergonomiques et mieux adaptés (par exemple, filtre d'écran, souris qui épouse la forme de la main, appui-poignets, etc.). Dans la pratique toutefois, la distinction n'est pas toujours faite puisque bien souvent les produits des technologies adaptées seront aussi offerts au public en général. En conséquence dans la documentation de langue anglaise les deux termes sont souvent utilisés indistinctement : Linda, Wilson. *Assistive technology for the disabled computer user*. (1994), gopher://isaac.engr.washington.edu/academic/guides/irg-20.txt.2

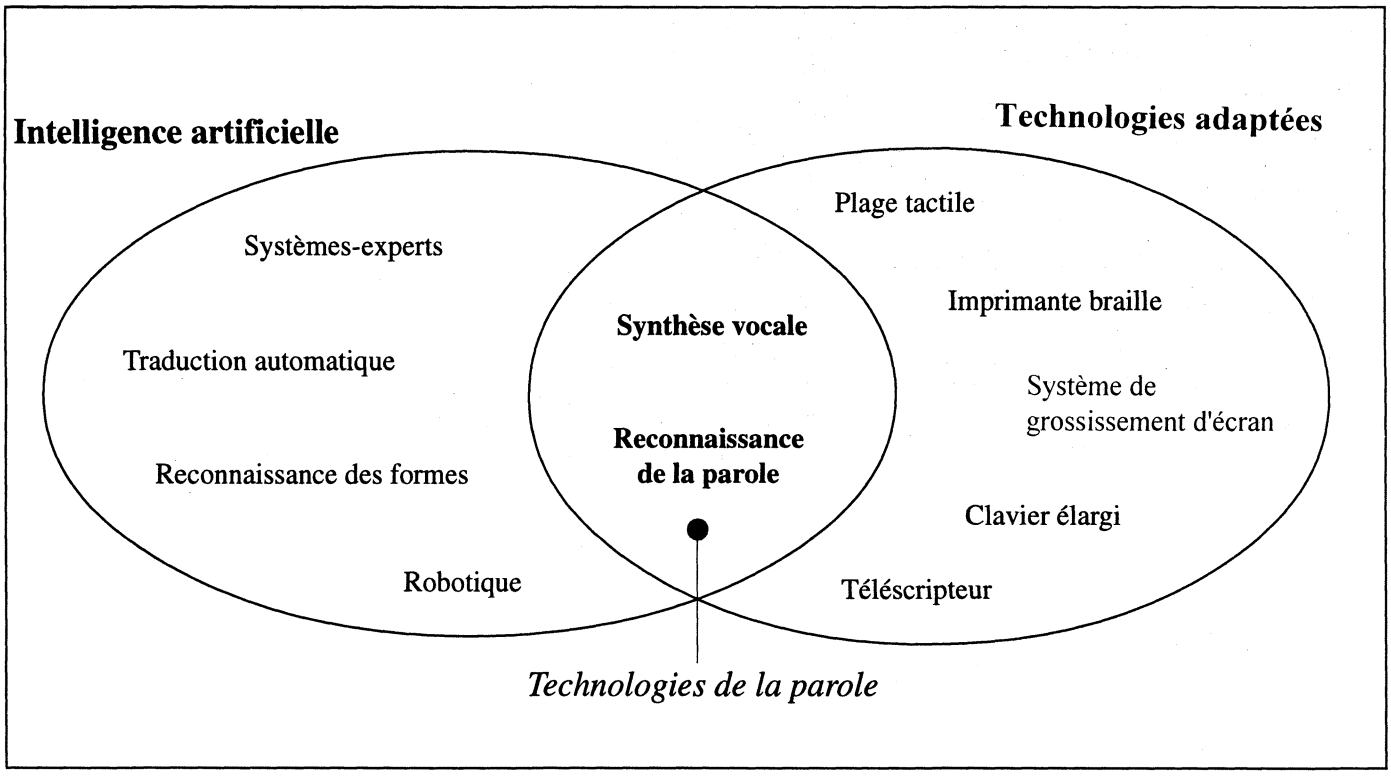

Figure 1. Contexte technologique

Considérée sous l'angle des technologies adaptées, la synthèse vocale peut venir en aide à deux catégories bien distinctes d'utilisateurs ayant une déficience. D'une part, cette technologie peut aider les personnes souffrant de mutisme ou ayant des troubles graves d'élocution. La synthèse vocale permet à ces personnes d'avoir une voix. D'autre part, la production synthétique de la parole peut aussi assister les personnes non voyantes ou semi-voyantes ainsi que les personnes analphabètes qui peuvent l'utiliser pour prendre connaissance de l'information représentée sous forme écrite. La synthèse vocale est particulièrement utile aux personnes ayant une déficience visuelle et ne connaissant pas le braille. Selon les statistiques de l'American Printing House for the Blind datant de1989, près de 31% des personnes non voyantes ou semi-voyantes ne pouvaient lire le braille (Rex 1989). La parution de ce résultat montrait d'ailleurs que l'utilisation du braille au sein de la population ayant une déficience visuelle connaissait un recul depuis quelques années. L'utilisation de la synthèse vocale par les personnes qui ne peuvent lire directement l'information imprimée est d'autant plus intéressante que l'on commence à la retrouver dans les milieux documentaires. La figure 2 présente un résumé schématique des technologies adaptées venant spécifiquement en aide aux personnes aveugles ou amblyopes.

Deux approches en synthèse vocale

En production synthétique de la parole, deux approches méthodologiques sont principalement utilisées (Alistair 1991; Lange 1993) : la synthèse de copie *(copy-synthesis)* et la synthèse par règles *(synthesis-by-rules)*. La première approche consiste à produire une voix synthétique à partir de fragments de parole (phrases, mots) préalablement enregistrés. Nous n'entrerons pas dans des considérations techniques en ce qui concerne les méthodes d'enregistrement et de conservation des fragments de parole. Qu'il nous suffise de dire de façon très générale que les ondes sonores sont numérisées et conservées dans la mémoire de l'ordinateur. Une fois sous forme numérique, les fragments sonores peuvent être reproduits ou concaténés. Ce genre de synthèse est notamment utilisé par les compagnies de téléphone pour l'automatisation de certains services. Par exemple, lorsque l'on compose le service d'assistance annuaire pour obtenir le numéro de téléphone de quelqu'un, une voix nous répond:«le numéro demandé est le... deux... deux... trois...». La chaîne vocale que l'on entend consiste en une concaténation de fragments, ce qui est d'ailleurs assez apparent pour peu que l'on y prête attention. D'autres exemples d'utilisation de la synthèse de copie se retrouvent dans différents gadgets électroniques comme les calculatrices et les montres parlantes.

La qualité globale du discours synthétique produit par copie, véritable reproduction de voix humaine, est le principal avantage de la synthèse de copie. Toutefois, il y a deux inconvénients majeurs. Le discours qui peut être ainsi généré est limité au corpus qui a été numérisé. Cela peut convenir lorsque l'on sait d'avance ce que la synthèse devra produire (par exemple, les chiffres de zéro à neuf pour le service d'assistance annuaire). Par contre, si on veut que la synthèse puisse reproduire toutes les phrases d'une langue, cette technique devient inadéquate.

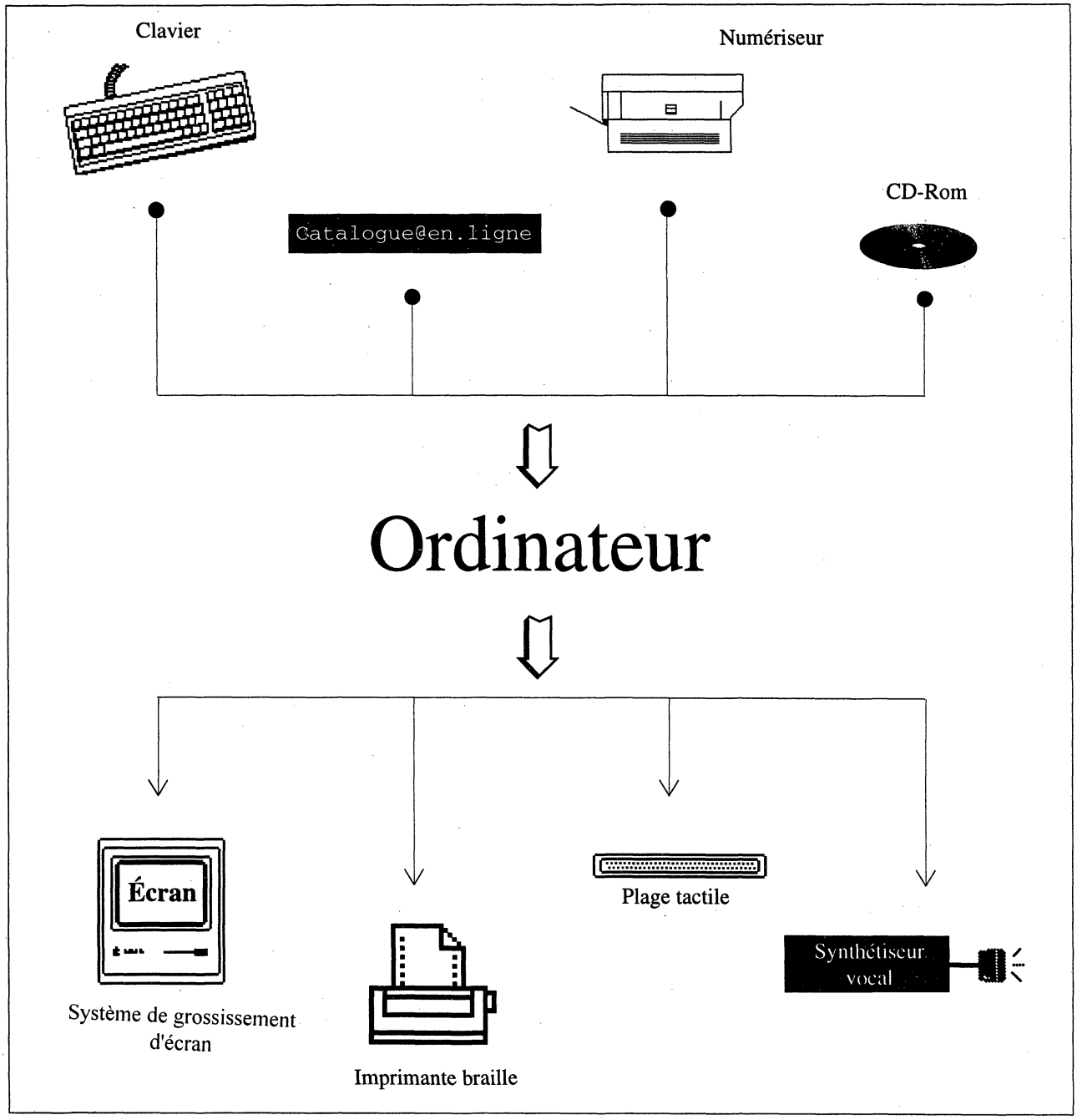

Figure 2. Technologies adaptées pour les personnes ayant une déficience visuelle

Le second inconvénient tient à l'espace de stockage qui est nécessaire à ce genre de synthèse. Par exemple, Pahwa (1995) souligne qu'une seconde d'intrant audio numérisé nécessite176 Ko. Une minute occupe alors plus de 10 000 Ko, soit environ 10 Mo. Si on numérisait un livre de 250 pages, lu à un rythme d'une page/min, on aurait besoin de 250 ´ 10 Mo = 2 500 Mo, soit 2,5 Go d'espace disque! Le développement des techniques de compression et des supports de stockage, notamment avec l'arrivée du DVD, pourra éventuellement apporter des solutions à ce point faible de la synthèse par copie.

La synthèse vocale réalisée à partir de règles représente la deuxième approche méthodologique. La mise en œuvre de cette approche sera plus approfondie dans la partie suivante. Il s'agit de générer, à partir d'un texte sous forme écrite, une représentation sonore en faisant appel à des règles phonétiques et prosodiques. Les règles peuvent ressembler à ceci: *Si la lettre* «t» *se trouve en finale de mot alors elle ne se prononce pas, sinon elle se prononce [t].* Ce système est beaucoup moins redondant que la synthèse par copie puisque les descriptions

numériques des sons qui sont gardées en mémoire se limitent aux sons distinctifs d'une langue, ce qu'on appelle les phonèmes d'une langue. En moyenne, les langues comportent de 20 à 40 phonèmes qui suffisent pour prononcer tous les mots; pour sa part le français en comporte 36. Reposant sur une analyse linguistique du texte, ce genre de synthèse vocale en constitue la forme la plus complexe mais aussi la plus flexible. Les synthétiseurs que nous présenterons utilisent tous cette méthode ou une adaptation de cette méthode.

Matériels et logiciels

Les synthétiseurs de parole se présentent sous une des deux formes suivantes: interne ou externe. Le synthétiseur interne est généralement constitué d'une carte et d'un haut-parleur (qui peut être interne ou externe) tandis que le synthétiseur externe consiste en un appareil périphérique externe relié à l'ordinateur par un port d'accès. Tant les synthétiseurs externes que les appareils internes possèdent leur propre mémoire, leur propre microprocesseur en plus d'un processeur de signal numérique et d'un convertisseur numérique-analogique (Lazzaro 1993, 46-50). Les synthétiseurs vocaux sur le marché offrent plusieurs fonctionnalités: choix du rythme de lecture qui peut aller de plusieurs dizaines à quelques centaines de mots/minute; choix de voix (voix féminine, voix masculine, voix d'enfant); choix du mode de lecture (lecture continue, lecture avec épellation de chaque mot). Les synthétiseurs les plus courants dans les milieux spécialisés sont le DECtalk, l'Echo PC, le Dolphin Apollo (Mates 1991, 161). De façon générale, il est reconnu que le synthétiseur vocal offrant la plus haute qualité de synthèse est le DECTalk de la firme Digital Equipment Corporation. Ce synthétiseur offrant une variété de neuf voix se vend à environ 1 200 $US (Digital Equipment Corporation 1994). Il est par ailleurs intéressant de noter que des applications des technologies de la parole commencent à faire leur apparition dans le marché grand public. C'est le cas par exemple du système ViaVoice d'IBM qui offre pour un peu moins de 140 $CAN le premier système de dictée vocale en mode *continu*. La version enrichie de ViaVoice (*ViaVoice Gold*) qui est offerte à environ 200 $CAN comporte en plus un module de synthèse vocale (IBM 1997).

Très brièvement, on peut expliquer le processus de synthèse à l'aide de trois paliers (Vitale 1993). Au premier palier, le texte codé en ASCII est converti en une chaîne d'unités phoniques. Pour réaliser la conversion de la chaîne de caractères ASCII en chaîne phonique, le système utilise diverses règles de correspondance lettre–phonème ainsi qu'un dictionnaire contenant des correspondances forme graphique–forme phonétique. Afin de bien comprendre cette étape, prenons un exemple très simple: l'expression «Mme Simone». Le synthétiseur commencera par vérifier si les deux mots se retrouvent dans son dictionnaire. Si le dictionnaire est complet, il retrouvera l'abréviation «Mme»avec la traduction phonétique correspondante [madam]. Quant au mot «Simone», il est fort probable qu'il ne soit pas dans le dictionnaire puisque c'est un mot régulier qui suit les règles normales. Le synthétiseur comprendra qu'il peut lui appliquer les règles et il produira la forme[simon]. Le dictionnaire du synthétiseur peut inclure plusieurs types d'éléments: abréviations, noms propres, mots d'origine étrangère dont la prononciation ne suit pas les règles normales. Certains systèmes comportent même une liste de morphèmes très courants dont la prononciation est régulière, par exemple: «-tion», «-able», «hexa-», «macro-». Cela permet au synthétiseur de reconnaître très rapidement des parties de mot et de les traduire en bloc.

Au deuxième palier, la chaîne phonétique se voit adjoindre divers paramètres qui rendent compte de la prosodie de discours, c'est-à-dire les liaisons entre les mots, le ton de la voix, l'accentuation sur les mots ou sur les syllabes, les pauses en finale de phrase, etc. Il faut souligner que ces paramètres sont principalement responsables de l'aspect naturel et convivial de la voix synthétique. C'est d'ailleurs à ce niveau que se distinguent les excellents synthétiseurs.

Finalement, au dernier palier, toutes les données phonétiques et prosodiques qui ont la forme de valeurs numériques sont envoyées au haut-parleur afin qu'il puisse transformer ces signaux en ondes sonores. Pour que l'auditeur ait l'impression d'une onde sonore continue, la fréquence des signaux envoyés au haut-parleur doit être extrêmement élevée (environ un signal par milliseconde). À la fin de ce processus, il est alors possible d'entendre une voix plus ou moins mécanique lisant le texte choisi. Ajoutons que la synthèse vocale n'est pas miraculeuse (du moins pas encore!): le résultat est parfois décevant et, à vrai dire, il faut avoir l'oreille habituée pour pouvoir suivre parfaitement ce qui est énoncé.

Un système opérationnel de synthèse vocale

Concrètement, un synthétiseur vocal ne produira rien du tout si on ne lui fournit pas du texte à traiter. Pour créer cet intrant textuel, il faut joindre au synthétiseur soit un numériseur, soit un logiciel de lecture d'écran.

Dans le premier cas, le numériseur crée une copie de l'image du texte qu'on lui présente, cette copie étant ensuite décodée à l'aide d'un logiciel de reconnaissance optique de caractères (ROC) afin de produire un fichier de caractères de type ASCII. Après cette étape, le fichier ASCII peut être transmis au synthétiseur pour la traduction sonore. Le premier système commercial de synthèse vocale, le Kurzweil Reading Machine (KRM), fonctionnait essentiellement de cette façon. Le système comprenait un ensemble de logiciels et de matériel qui pouvait fonctionner de façon autonome, c'est-à-dire qu'il incluait un scanner, un logiciel ROC, un bloc système, un clavier et un synthétiseur vocal. En dépit de son coût très élevé, le KRM a été très populaire aux États-Unis et, en quelque sorte, on peut considérer qu'il est responsable de la popularité de la synthèse vocale. En 1989, une nouvelle version de KRM, améliorée et plus compacte, a fait ses débuts sous le nom de Xerox-Kurzweil Personal Reader (KPR) qui utilise le synthétiseur DECtalk. Le système Reading Edge offre un autre exemple de système jumelant un système de reconnaissance optique de caractères et un synthétiseur vocal.

La deuxième façon de faire fonctionner un synthétiseur consiste à l'alimenter à l'aide d'un programme de lecture d'écran «Screen Reader Software», «Screen Access Software» ou encore «Software Driver for synthesizer». Ce type de logiciel prend connaissance de l'information qui est envoyée à l'écran et la transmet au

synthétiseur sous forme de chaîne de caractères ASCII. Comme l'écrit Barbara T. Mates, les lecteurs d'écran «*tell the synthesizer what to do and how to do it; they are to the speech synthesizer what the brain is to the mouth*» (Mates 1991, 49). Le logiciel lecteur d'écran est installé dans l'ordinateur en mode TSR (Terminate and Stay Resident). Il suffit d'activer le logiciel pour qu'il se mette à lire ce qui apparaît à l'écran. Par la suite, il est possible d'entrer dans diverses applications, comme des logiciels de traitement de texte par exemple.

Une fois le lecteur d'écran activé, l'utilisateur peut choisir à tout moment le type de lecture qu'il désire: lecture mot à mot, lecture ligne à ligne, lecture continue de la totalité de l'écran, épellation d'un mot. En outre, la plupart des logiciels d'accès à l'écran comportent une commande pour imposer le silence au synthétiseur quand les utilisateurs sont fatigués de l'entendre. Les systèmes de lecture d'écran les plus connus sont Arctic/Business Vision, IBM Screen Reader, JAWS, Vert Pro, Vert Plus, Soft Vert, Flipper, Vocal Eyes. Certains ensembles de lecteur d'écran offerts sur le marché contiennent à la fois un logiciel de lecture et un synthétiseur vocal (par exemple, les logiciels Vert Pro, Vert Plus sont offert dans un ensemble avec le synthétiseur DECtalk) tandis que d'autres se limitent au logiciel que l'on peut connecter au synthétiseur de son choix (c'est le cas d'IBM Screen Reader, FLIPPER et Soft Vert). Il faut aussi mentionner le cas du système Frank Audiodata livré avec un clavier spécial servant spécifiquement à commander le lecteur d'écran.

Comme on peut s'en douter, les systèmes de lecture d'écran sont assez bien adaptés pour travailler avec des interfaces dans lesquelles l'information se présente sous forme textuelle. Il semble toutefois plus difficile d'imaginer leur fonctionnement avec les GUI (Graphical User Interfaces), c'est-à-dire les environnements graphiques de type Windows ou Macintosh. Puisque ces interfaces mettent en scène des fenêtres, des menus, des icônes et l'utilisation d'une souris, il est facile de comprendre que la vision de l'utilisateur y joue un grand rôle et que l'information n'est pas transmise uniquement sous forme textuelle. Paradoxalement, les GUI qui ont été créés pour faciliter la vie des utilisateurs voyants et pour leur offrir un environnement plus convivial occasionnent bien des difficultés aux utilisateurs non voyants et rendent les lecteurs d'écran ordinaires beaucoup moins efficaces. Plusieurs recherches ont été menées afin de créer des lecteurs d'écran conçus pour décoder l'information d'un environnement graphique[5]. Récemment, sont apparus plusieurs lecteurs d'écran spécialement adaptés aux environnement graphiques (Lazzaro 1993). Un lecteur d'écran pour Apple Macintosh existe en version commerciale sous le nom de Out SPOKEN. En ce qui concerne l'environnement Windows, c'est le système Window Bridge qui est disponible. Une version du Screen Reader d'IBM, le ScreenReader/2, a été élaboré afin de pouvoir traiter l'environnement graphique du système OS/2. Il faut préciser que ces systèmes d'accès à l'écran ne reposent pas uniquement sur la synthèse vocale pour communiquer avec l'usager et ils peuvent recourir à plusieurs codes sonores pour véhiculer l'information. Par exemple, lorsque l'utilisateur fait glisser la souris sur la barre de menu, un son distinctif lui indique qu'il est au bon endroit.

Applications dans les milieux documentaires

Dans une bibliothèque moderne, les usagers voyants peuvent mener trois activités qui apparaissent beaucoup plus difficiles pour les usagers ayant une déficience visuelle: ils peuvent interroger le catalogue automatisé de la bibliothèque; ils peuvent consulter des ouvrages sur place et finalement ils peuvent consulter des bases de données qui sont disponibles soit en ligne, soit sur disque optique compact. Les applications de la synthèse vocale en milieu documentaire touchent ces trois activités et permettent d'offrir des solutions aux usagers non voyants ou semi-voyants.

La raison d'être d'une bibliothèque, c'est l'information contenue dans ses collections. Mettre cette information à la disposition des usagers apparaît donc, pour le personnel de la bibliothèque, comme un des objectifs principaux à atteindre. Pour les usagers non voyants, le contenu des imprimés est de prime abord inaccessible et c'est là qu'intervient la synthèse vocale. Un système comprenant un synthétiseur et un numériseur met à la portée des non voyants la presque totalité de la collection d'une bibliothèque. C'est sous cette forme que la synthèse vocale a fait son entrée dans de nombreuses bibliothèques en Amérique du Nord (Rosen et al. 1991; Mates 1992).

En ce qui concerne le catalogue automatisé, il est possible de le faire parler à l'aide d'un lecteur d'écran et d'un synthétiseur. C'est ce qu'ont essayé de réaliser les bibliothécaires de la M.C. Memorial Library de l'American Foundation for the Blind. Le catalogue de la bibliothèque est accessible à tout le monde via la base de données SearchMagic[6]. De plus, un logiciel de lecture d'écran, Screen Talk, et un synthétiseur vocal, DECtalk, ont été implantés afin de fournir un accès convivial au catalogue pour les utilisateurs non voyants. Une autre expérience de ce type a été réalisée avec succès à la Phoenix Public Library (Kneedler and Sizemore 1993). Le même équipement, lecteur d'écran et synthétiseur, permet également l'accès à diverses bases de données en ligne ou sur disque optique compact (Mates 1990), ce qui n'est toutefois pas toujours sans problème, entre autres à cause de certains problèmes de compatibilité entre le CD-ROM et le lecteur d'écran (Carter and Jackson 1992). En dernier lieu, mentionnons le projet prometteur implanté à l'Université d'État du Nouveau-Mexique (Thompson) L'objectif de ce projet était d'adapter tous les postes de travail des laboratoires destinés aux étudiants de façon à ce qu'ils puissent être utilisés sans restriction par des étudiants non voyants ou semi-voyants. Au terme du projet, tous les postes de travail ont été équipés d'une copie d'un logiciel de lecteur d'écran pour lequel une licence de site avait été obtenue et la New Mexico Commission for the Blind a fourni à chaque étudiant qui le désirait un synthétiseur vocal externe, léger et portatif, ainsi qu'un casque d'écoute. Le résultat de ce projet est qu'il n'y a pas de poste réservé strictement aux non

5. Pour un résumé de ces recherches, voir Edwards Alistair, *Speech synthesis: technology for disabled people*. (Londres: Paul Chapman Publishing, 1991).

6. SearchMagic est une version «search-only» et «user-friendly» de INMAGIC.

voyants et ceux-ci peuvent donc utiliser les mêmes postes que tout le monde. Ainsi, d'une manière presque invisible, les laboratoires sont devenus beaucoup plus conviviaux.

Pour sa part, l'Université de Montréal offre des services de synthèse vocale depuis quelque temps déjà et deux points de service ont été mis en place, la salle Gabrielle Fournier-Ethier située dans la Bibliothèque Samuel-Bronfman et l'Audiovidéothèque du pavillon Marie-Victorin. Parmi le matériel utilisé, mentionnons entre autres le système québécois IRIS (Numériseur + ROC + Synthétiseur) mis en marché par Visuaide 2000 de Longueuil. Bien que des projets soient en voie de réalisation, ni le catalogue Atrium de l'Université de Montréal, ni les bases de données sur disque optique compact ne sont encore accessibles sur place aux usagers ayant une déficience visuelle. Toutefois, le catalogue Atrium est accessible à distance par le moyen d'une connexion TELNET, ce qui permet aux usagers non voyants ou semi-voyants qui possèdent les logiciels et le matériel précédemment décrit de pouvoir prendre connaissance des informations bibliographiques.

Finalement, un dernier exemple d'application de la technologie vocale illustre merveilleusement bien les bénéfices que peuvent en retirer les personnes souffrant d'une déficience visuelle. Il y a quelques années, Norman Coombs, professeur d'histoire au Rochester Institute of Technology (État de New York), a donné un cours sur les droits civiques de la minorité afro-américaine à un auditoire très particulier: un tiers de ses étudiants était constitué de personnes malentendantes, un autre tiers de personnes aveugles ou amblyopes et le dernier tiers était composé d'étudiants sans déficience de vision ou d'audition. Autre particularité de ce cours: puisque les étudiants se trouvaient un peu partout aux États-Unis, le cours devait être donné en ligne par l'intermédiaire d'Internet. Cette situation extraordinaire n'aurait pu être possible sans l'aide des technologies adaptées. En effet, Norman Coombs est un professeur non voyant, et il utilise depuis quelques années un synthétiseur vocal pour se promener sur Internet, ce qui lui permet d'avoir accès à des catalogues en ligne, de consulter des ouvrages de référence, de«lire» des livres et des périodiques électroniques. En outre, le synthétiseur vocal lui permet d'enseigner et de répondre aux questions de ses étudiants en ligne! À propos de son expérience d'enseignement en ligne à l'aide de la synthèse vocale, le professeur a émis ce commentaire: «*We all were the same on the computer. Not only did it let us transcend physical distance and physical disabilities, but it also made other physical differences like race fade into the background and permit a really open and stimulating environment.*» (Combs 1993)

Au moment où l'industrie des technologies de l'information connaît une évolution fulgurante, il importe que les bibliothécaires et documentalistes soient sensibles au fait que la *disponibilité* de l'information n'en garantit pas automatiquement l'*accessibilité*. Nos sociétés, dit-on, se caractérisent de plus en plus par le clivage entre les infopauvres et les inforiches. Nous sommes convaincus que des technologies telle la synthèse vocale pourront contribuer à renverser cette tendance en permettant aux personnes ayant une déficience visuelle de ne pas être en plus des *handicapés de l'information*.

Sources consultées

Alistair, Edwards. 1991. *Speech synthesis: Technology for disabled people*. Londres: Paul Chapman Publishing, p.20-22.

Brûlé, Michelle. 1994. La Bibliothèque Braille ou une odyssée de 95 ans. *Documentation et bibliothèques* 40(2): 81-86.

Carter, Robert and Kathy Jackson. 1992. Speech synthesizer and screen reading software make CD-ROM databases accessible to visually impaired library users. *CD-ROM Professional* 5(1).

Coombs, Norman. 1993. *Bringing the mountain to Muhammad: Online services and the disabled.* gopher://gopher.cni.org/cniftp/miscdocs/farnet/stories/NY/story070

Digital Equipment Corporation. 1994. *Digital introduces DECtalk express speech synthesizer.* http://www.digital.com/info/pr-news/94071201PR.txt.html

IBM. 1997. http://www.software.ibm.com/voicetype/ us_vv.html

Joanis, Marc. 1992. Des services documentaires adaptés aux personnes handicapées à l'Université de Montréal. *Documentation et bibliothèques* 38(1): 51-58.

Kneedler, William H. and E.J. Sizemore. 1993. Speech synthesis + online library catalog = Talking catalog. *Library Hi Tech*11(1): 57-64.

Lange, Holley R. 1993. Speech synthesis and speech recognition: Tomorrow's human-computer interfaces? In *Annual review of information science and technology*. Medford (N.J.): Learned Information/ American Society for Information Science, p.156-158.

Lange, Holley R. et al. 1991. Voice technologies in libraries: A look into the future. *Library Hi Tech* 9(3): 96.

Lazzaro, Joseph J. 1993. *Adaptative technologies for learning & work environments*. Chicago: American Library Association.

Mates, Barbara T. 1990. CD-ROM: A new light for the blind and visually impaired. *Computer in Libraries* 10(3): 17-20.

_____. 1991. *Library technology for visually and physically impaired patrons*. Westport (Conn.): Meckler Publishing, p.161.

_____. 1992. Adaptative technology makes libraries people friendly. *Computer in Libraries* 12(10):20-24.

McCormick, John. 1994. *Computers and the American with Disabiblites Act: A manager guide*. New York: Windcrest/McGraw-Hill.

McNulty, Tom and Dawn M. Suvino. 1993. *Access to information: Materials, technologies, and services for print-impaired readers*. Chicago: American Library Association.

Pahwa, Ash. 1995. What is the capacity of CD-recordable media? *CD-ROM professional* 8(1): 100-102.

Rex, E. J. 1989. Issues related to literacy of legally blind learners. *Journal of visual impairment and blindness* 77(14): 306.

Rosen, Leslie, Robert B. Jaeggin and Peggy W. Ho. 1991. Enabling blind and visually impaired library users: INMAGIC and adaptative technologies. *Library Hi Tech* 9(3).

Thompson, Mike. *Invisible computer for blind clients: An international campus-wide model for adapting computer equipment for clients who are blind.*
gopher://gopher.inform.umd.edu/Student_Information_And_Resources/Diversity_Resources/Disability_Support/OutsideResources/Computers/Support/for-blind-clients

Vignaux, Georges. 1992. *Les sciences cognitives: une introduction.* Paris: Éditions La Découverte, p.13.

Vitale, Tony. 1993. *Hardware and software aspects of a speech synthesizer developed for persons with disabilities*. Digital equipment Corporation. gopher://val-dor.cc.buffalo.edu/software/.dec.speech