

Les progrès de l'informatique, tant en ergonomie qu'en puissance et rapidité, mettent à disposition du public des logiciels de reconnaissance vocale de plus en plus performants, parfois directement intégrés dans l'arsenal de logiciels des nouveaux systèmes d'exploitation. L'occasion a été donnée à Thierry Opillard d'observer plusieurs de leurs utilisateurs et d'en tirer quelques réflexions.

LA DICTÉE À L'ORDINATEUR

Thierry OPILLARD

Le slogan d'un vendeur de ces logiciels affirme que son produit « transforme la dictée continue en texte ». C'est en effet spectaculaire de voir un habitué parler dans un micro, d'observer simultanément une fenêtre provisoire qui affiche les mots prononcés, de les voir s'y organiser, et enfin les voir rapidement portés sur la page à la suite d'analyses qui nous échappent.

En dehors de quelques imperfections que le logiciel est sans cesse en train « d'apprendre » à corriger au fur et à mesure qu'il se « familiarise » avec la voix et la structure langagière du locuteur, le résultat est très correct. La parole, prouesse technique, est bien retranscrite. Au point que le fabricant annonce un « taux de reconnaissance de 99% ».

LES LOGICIELS DE RECONNAISSANCE VOCALE

Le principe : on dicte dans un microphone relié à un ordinateur. Celui-ci est équipé d'un logiciel de reconnaissance vocale qui retranscrit par écrit ce qu'il reconnaît : des mots et des groupes de mots, probablement des structures syntaxiques, car le logiciel est plus efficace quand on dicte en continu que mot par mot.

Une « période d'apprentissage » débute l'utilisation du logiciel. Il faut lire un texte imposé par le fabricant, pendant 10 à 20 minutes, pendant lesquelles le logiciel fait coïncider les particularités vocales du nouvel usager à une série de répertoires oraux, eux-mêmes déjà couplés à un répertoire écrit. Cette période est finalement très courte, et on le comprend aisément quand on se remémore le peu de mots employés couramment et leurs fréquences.

Le nouvel usager peut aussi lui lire ou lui intégrer ses propres textes pour qu'il s'imprègne des structures syntaxiques qu'il utilise habituellement. Le fabricant vend, chèrement, des bibliothèques professionnelles spécialisées, pour le lexique le moins fréquent ; on est sûrement là au cœur de « l'intérêt » de la chose... C'est cette période d'apprentissage qui induit en erreur : on dicte du texte, il s'écrit. On en conclurait vite fallacieusement qu'il suffit de parler pour avoir aussitôt du texte, qu'il y aurait homothétie en oral et écrit.

Un autre argument du vendeur : la vitesse. Donc, le temps gagné (sous-entendu : la rentabilité... puisqu'on peut enregistrer numériquement sa voix où que l'on soit, pour réinjecter le fichier son dans le logiciel qui retranscrit ensuite). Il annonce en effet 120 à 160 mots-minute, c'est-à-dire 7 200 à 9 300 mots-heure ; on reconnaît dans ce dernier chiffre la vitesse maximale de débit de la parole. Cet argument semble imparable quand on connaît les normes en vigueur dans les bureaux : en moyenne une secrétaire débutante doit saisir au moins 25 mots/minute. Une secrétaire confirmée, pour laquelle les travaux de frappe sont la mission principale, doit saisir environ 40 à 45 mots/minute.

Les 400 ou 500 mots-heure produits par des experts lors de l'activité d'écriture, comme l'ont montré les études que nous avons menées lors de la recherche *Genèse de Texte*, sembleraient faire pâle figure.

Pour résumer, les mots s'alignent sur l'écran 3 à 6 fois plus vite que si on les tape, et 20 fois plus vite que si on doit réfléchir pour écrire. S'il y a bien quelque chose qui a disparu dans la situation de dictée à l'ordinateur, c'est probablement la réflexion.

UN FORMIDABLE OUTIL... DE PRISE DE CONSCIENCE

Si le texte que l'on produit sortait directement de notre cerveau, si on dictait directement du texte, si la lenteur apparemment constatée des experts en écriture relevait uniquement d'une lenteur du poignet qui tient le stylo ou des doigts sur le clavier, ces logiciels constitueraient un progrès extraordinaire.

Pourquoi donc les écrivains, quand ils ont percé et en ont les moyens, ne se sont-ils pas tous déjà dotés d'une secrétaire très performante qui leur permettraient de produire 6 à 7 fois plus vite ? C'est peut-être qu'il y a dans l'écriture une activité qu'il n'y a pas dans la production orale. C'est peut-être qu'il s'y joue autre chose. C'est certainement qu'il n'y a pas homothétie entre l'oral et l'écrit : pas d'*opération qui consiste à faire correspondre à tout élément d'un ensemble A un élément d'un ensemble B et un seul*. Le producteur de l'oral n'est jamais dans la **position** du producteur de l'écrit. Leurs fonctions, leurs fonctionnements, sont différents.

Quel que soit ce que l'on dicte, il y a toujours quelque chose à reprendre, à corriger. Imaginons un scripteur très expert qui dicterait quasiment du texte : malgré tous les progrès réalisés par les correcteurs automatiques, jusque dans la complexité des accords de participe passé et de choix participe/infinif, il en reste toujours que le logiciel manque. Il y a toujours des ambiguïtés homophoniques qu'il ne peut lever, des problèmes textuels insolubles, parfois une ponctuation défectueuse impossible à rétablir. Pour cela, ces logiciels offrent tout un arsenal de corrections possibles, commandées à la voix : les opérations d'ajout, d'effacement, de remplacement, de déplacement sont possibles, ainsi que toutes celles concernant les mises en formes de mots et paragraphes, la ponctuation, etc., bien qu'assez laborieuses et susceptibles de « faire tomber la rentabilité ». Et l'on sait que quand on commence à retoucher, c'est-à-dire à relire, regarder du point de vue du lecteur, la spirale de la réécriture s'enclenche. L'insatisfaction chronique du lecteur-scripteur trouve à s'investir.

Quand on observe la dictée d'un utilisateur lettré non expert en écriture, on constate qu'il est fréquemment emporté par le flot discursif de son oral. Il constate rapi-

dement « ce qui ne va pas » et les opérations de travail du texte prennent alors une importance grandissante qui grève largement les « avantages » précédemment vantés.

Toutefois, si l'on classe les enseignants dans cette catégorie, on voit l'intérêt qu'il peut y avoir à ce qu'ils utilisent ce type d'outil. Il devient un redoutable outil de prise de conscience qu'écrire, produire du texte, ne relève aucunement de l'oral, qu'écrire commence dans ce travail, cette reprise du matériau. Il est intéressant de voir comment cette réécriture nécessitée par l'oralité du retranscrit redonne la main à la main, rend le façonnage à la souris efficace : la complexité des opérations de restructuration plus ou moins profondes, de reconfiguration par un autre angle d'attaque en reprenant une bonne partie des mots mais en les agençant autrement, les opérations de resyntaxisation, de mise en cohérence, de polissage, d'entrelacement, d'élagage, tout cela ne peut être assumé par les moyens de correction proposés, à moins d'y passer un temps trop important et de perdre le fil.

Quant à l'utilisateur peu lettré, il est manifeste qu'il produit de l'oral, que celui-ci se transcrit, qu'il a peu de moyens d'y remédier, et qu'il peut voir renforcée sa croyance dans la redondance des deux systèmes de communication. La fonction de sonorisation de son « texte » peut peut-être lui apporter le recul pour une prise de conscience ; cela saute à nos oreilles, il faudrait encore le vérifier pour les siennes.

Revenons sur la « lenteur » de l'écriture, aussi bien manuscrite que tapuscrite : elle a comme vertu de favoriser la réflexion. Quand on tape ou qu'on écrit à la main, c'est suffisamment lent pour qu'on pense en même temps. Pour qu'on ait le temps de penser aux niveaux micro et macro structurels du texte, à la globalité de son projet, aux réseaux sémantiques, aux effets produits, etc. On réécrit déjà même plusieurs fois pendant qu'on écrit parce qu'on explore les possibles, les impossibles, qu'on « exerce la gestualité de l'invention ». Il y a donc dans l'écriture une non soumission au flux temporel de la parole, au débit qu'impose la situation dialogique directe. Cette libération du temps permet d'exercer un autre mode de pensée, celui de l'écriture, de l'organisation spatiale de matériaux visuels, *vs* celui de la parole, flux temporel d'immatérialité sonore.

CONCLUSION

Nous l'avons vu, il n'est pas difficile de montrer que ces logiciels ne « *transforme[nt] [pas] la dictée continue en texte* ». Ils transcrivent le flux de paroles en reconnaissant des mots et des structures syntaxiques. Leur utilisation sur plusieurs heures est annoncée comme aisée tant il paraît « naturel » de pouvoir soutenir une conversation au long cours. Mais, dans une démarche d'écriture, et donc de réécriture, et non de dictée d'éléments pré-écrits ou régurgités automatiquement comme des lettres types ou des rapports très cadrés et formatés, ils s'avèrent beaucoup plus fatiguant que la tenue du stylo.

Leur utilisation est l'occasion spectaculaire de prendre conscience de la nature fondamentalement différente de l'oral et de l'écrit.

■ **Thierry OPILLARD**

Dans le monde de la textualité numérique, les discours ne sont plus inscrits dans des objets, qui permettent de les classer, hiérarchiser et reconnaître dans leur identité propre. C'est un monde de fragments décontextualisés, juxtaposés, indéfiniment recomposables sans que soit nécessaire ou désirée la compréhension de la relation qui les inscrit dans l'œuvre dont ils ont été extraits.

Roger CHARTIER, *L'avenir numérique du livre*. *Le Monde*. 27.10.09

