

Cinq contributions pour comprendre la lecture
journées d'études des 25 – 26 – 27 février 1980

CONFÉRENCE DE JEAN-SYLVAIN LIÉNARD
Maître de recherche au Laboratoire Informatique
pour la Mécanique et les Sciences de l'Ingénieur
(LIMSI - CNRS – Orsay)

LA LECTURE DES ORDINATEURS
ÉCRITS ET INTELLIGENCE ARTIFICIELLE

Cinq contributions pour comprendre la lecture

journées d'études des 25 – 26 – 27 février 1980

Je m'occupe d'analyse et de synthèse de la parole et je crains un peu que Messieurs FOUCAMBERT et VIOLET n'aient surestimé l'apport que je peux faire, parce que mes travaux sont relativement distincts des vôtres.

On m'a demandé d'exposer mes travaux, sans trop tenter de les appliquer à votre domaine, mais j'imagine que c'est pour établir une sorte de parallélisme entre le processus de lecture humain, à voix haute et ce que l'on sait faire actuellement avec les ordinateurs en matière de synthèse de la parole à partir du texte.

Évidemment, il y a des choses communes ; dans les deux cas on part d'une chaîne de symboles orthographiés, de graphèmes, et on aboutit à un signal de parole, un signal acoustique et objectif.

Les différences on les verra au fur et à mesure de l'exposé ; c'est notamment le fait que, lorsque je lis un papier, j'ai un interlocuteur - la machine, a priori n'en a pas, en tous cas, elle n'en a pas conscience, il y a par ailleurs des problèmes de lecture optique...

Le processus lui-même, vous le verrez, est foncièrement différent. Mais, avant d'aboutir à ce qui fera en quelque sorte notre conclusion, et peut-être l'amorce d'une discussion, je voudrais vous exposer les grandes lignes de ma causerie, à commencer par un peu d'analyse de la parole.

Je vous montrerai ensuite ce que l'on peut entendre par synthèse de la parole et les diverses étapes fonctionnelles qu'on peut distinguer dans ce processus. Je parlerai en particulier de synthèse acoustique, de commande phonétique, tout ceci pour aboutir aux problèmes que nous nous posons actuellement lorsque nous faisons parler des machines, c'est-à-dire les problèmes d'intelligibilité, de qualité et de naturel.

Puis je vous parlerai de la prosodie ; je vous dirai en quoi la communication par la parole et la communication par l'écriture me semblent foncièrement différentes. Enfin, je vous dirai quelques mots de la reconnaissance automatique et des systèmes de communication parlée entre hommes et machines.

C'est un programme un peu ambitieux. Je vais essayer d'aller très vite et de ne pas trop vous surcharger avec des notions qui vous sont, à priori, étrangères.

Cinq contributions pour comprendre la lecture

journées d'études des 25 – 26 – 27 février 1980

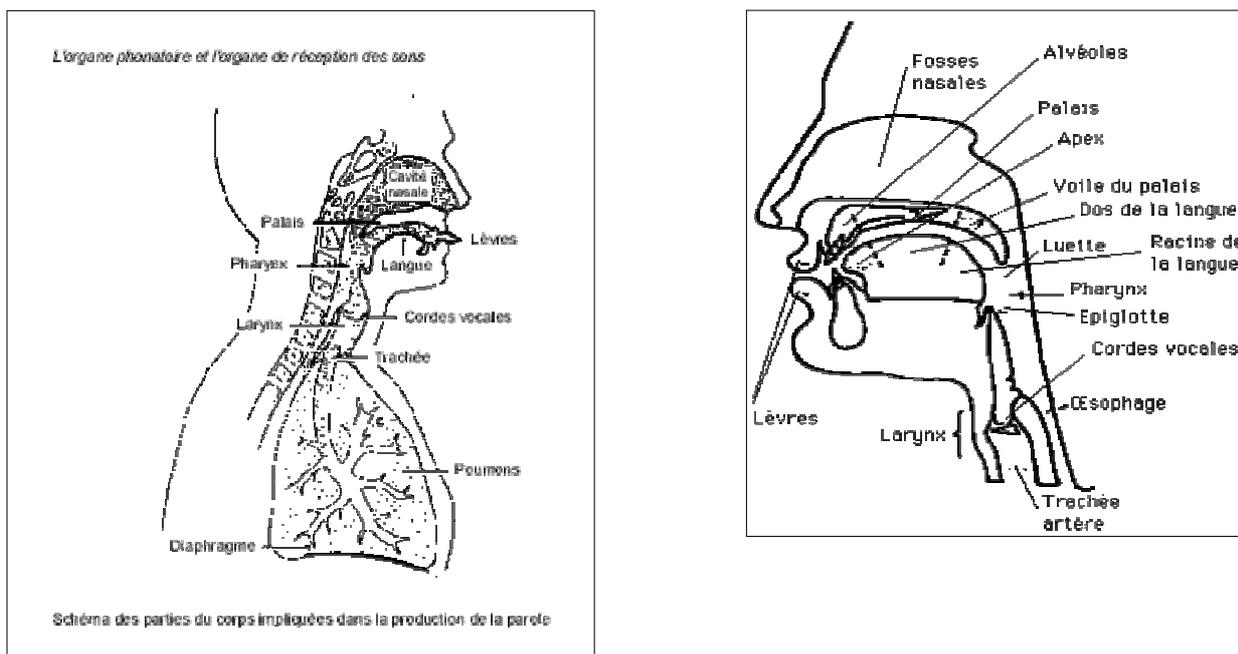


Fig. 1 ANATOMO-PHYSIOLOGIE DE L'APPAREIL VOCAL

Pour l'analyse de la parole, il nous faut quand même quelques notions de base :

Sur ce schéma vous avez un individu, un être humain en coupe et vous voyez qu'il y a une machine à parler, constituée d'un soufflet - les poumons - d'un instrument qui vibre - les deux cordes vocales (qui ne sont pas des cordes d'ailleurs) - et un résonateur, plus exactement un tuyau acoustique qui introduit certaines résonances dans le son émis par les cordes vocales. Ce tuyau acoustique est un peu complexe. Il est séparé en deux : une partie orale modelée par le pharynx, le voile du palais, les dents, les lèvres, qui donne des résonances variables, plus ou moins aiguës, et une partie nasale mise en circuit par l'abaissement du voile du palais et qui donne des résonances fixes.

Alors, qu'est-ce que la parole ? Eh bien, la parole, c'est le résultat de l'évolution de ce système extrêmement complexe, puisqu'il ne comporte pas moins d'une soixantaine de muscles fonctionnant de manière synchrone.

La parole, c'est en somme l'image acoustique de ces mouvements.

Je dois dire tout de suite ce sur quoi j'insisterai plus tard : s'il n'y avait pas mouvement de tout cet appareil, il n'y aurait pas perception de parole. C'est ce qui change qui est perçu, c'est ce qui change qui a de l'importance (ceci se vérifie d'ailleurs dans tout le domaine perceptif).

Décrivons les divers fonctionnements du conduit vocal :

- lorsque les cordes vocales vibrent, le son est modulé par le conduit vocal,
- quand la langue se rapproche suffisamment du voile du palais, l'écoulement d'air produit un bruit qui se mélange éventuellement avec le son des cordes vocales.

Cela fait des "che", des "je", etc.

- pour le "s", le rétrécissement a lieu entre le bout de la langue et les dents.

Cinq contributions pour comprendre la lecture

journées d'études des 25 – 26 – 27 février 1980

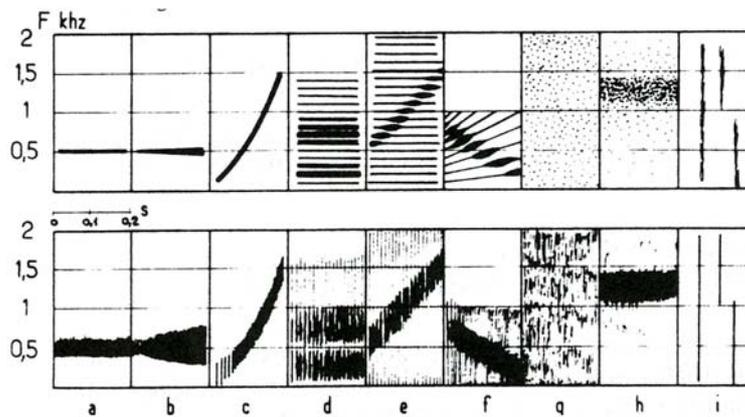


Fig 2 Quelques spectrogrammes schématiques établis selon une analyse en bande étroite (en haut) et en bande large (en bas)

► La vue suivante est un petit alphabet d'interprétation, des sonagrammes, documents fournis par un appareil d'analyse du son, le sonographe, que les phonéticiens utilisent également, et qui permet une visualisation de la parole (le premier nom de l'appareil était "visible speech").

En abscisse, de la gauche vers la droite, vous avez le temps qui se déroule et, en ordonnée la fréquence. C'est un peu comme une partition musicale plus un son est grave et plus il est vers le bas.

Sur cette image, vous avez tout à fait à gauche, un son simple, c'est à dire une simple sinusoïde, par exemple : un simple sifflet oral, à 1000 hz, qui est une fréquence assez aiguë.

Sur la deuxième vue, (en b), ce même son avec une intensité croissante.

En (d) vous avez un son simple, une sinusoïde dont la fréquence descend.

En (e) vous avez des impulsions, des chocs. Dans ce cas-là, vous avez toutes les fréquences. Un son comme cela est "bref", il dure quelques milli-secondes, mais il y a du grave et de l'aigu dedans, c'est pour cela que le trait s'étend du bas vers le haut.

Ensuite, en (f), vous avez un «bruit blanc» dans lequel il y a toutes les composantes fréquentielles pendant un temps indéfini.

En (g), vous avez un son harmonique comparable à celui qu'émettraient les cordes vocales si... on coupait la tête juste au-dessus des cordes vocales. La plupart des sons musicaux, des sons qui ont une hauteur perceptible, ont cette structure.

Il y a un fondamental qui est la première des raies à partir du bas et il y a des harmoniques qui se trouvent à deux fois, trois fois, n fois la fréquence du fondamental.

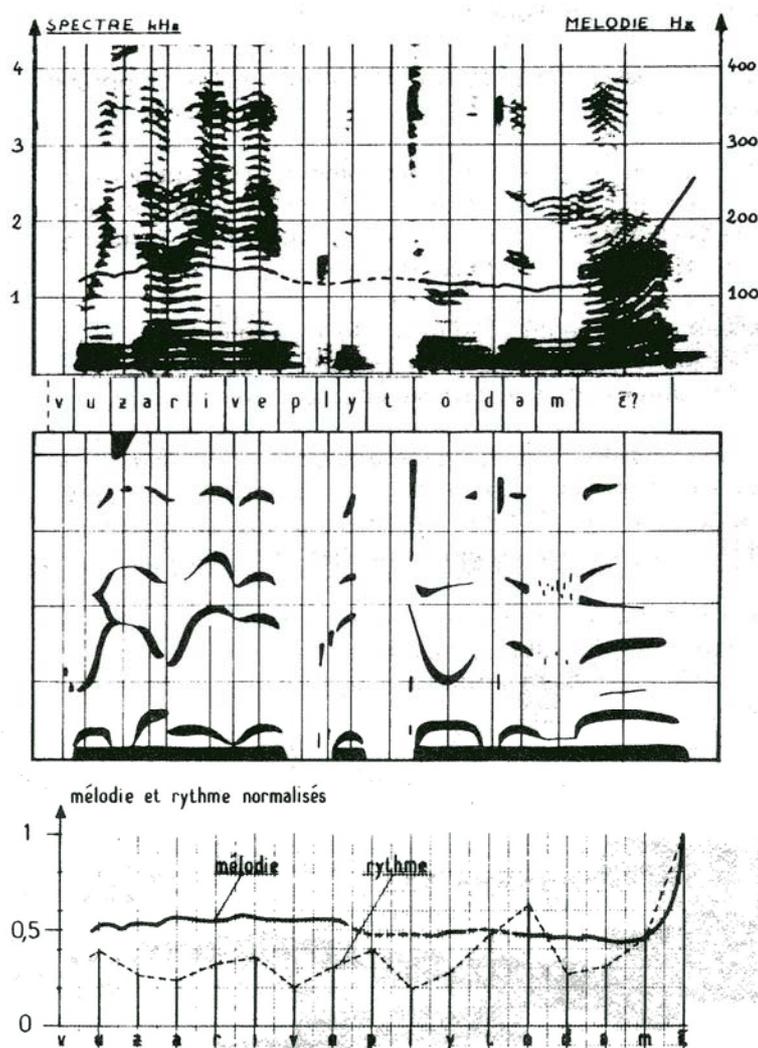
Dans l'image suivante (g) vous avez des zones du spectre qui sont plus fortes que d'autres. Ces zones sont renforcées ; ce sont elles qui donnent une certaine coloration aux sons, sinon ils paraîtraient complètement neutres. En matière de parole, on connaît bien ces zones. On les appelle des "formants". Elles correspondent aux résonances qu'applique le conduit vocal sur le son émis par les cordes vocales.

En (h), vous avez un son harmonique émis par les cordes vocales et un "formant", cette zone fréquentielle favorisée, qui se déplace du grave vers l'aigu puis qui redescend un petit peu.

Ces "formants" sont des choses importantes, qui reflètent les mouvements du conduit vocal, qui transportent une bonne partie de l'information phonétique de la parole.

Cinq contributions pour comprendre la lecture

jours d'études des 25 - 26 - 27 février 1980



3 Exploitation d'un spectrogramme. La phrase "vous arrivez plus tôt demain ?" est prononcée par un locuteur masculin avec une nuance interrogative. De haut en bas : spectrogramme (bande étroite) : l'évolution mélodique est indiquée par le 10^{ème} harmonique; limites et identification des segments phonétiques; squelette phonétique; courbes prosodiques normalisées.

► Cette nouvelle vue permet de comprendre que la parole peut être analysée, avec cet appareil dont je vous ai montré quelques éléments, quelques représentations élémentaires.

La parole est un signal extrêmement complexe, parce qu'il y a non seulement un fondamental dont vous voyez la structure de raies qui reflète l'évolution du fondamental, mais aussi des tâches qui se déplacent et qui reflètent l'évolution du formant donc du conduit vocal.

Il faut bien dissocier ces deux notions sinon on ne comprend rien à la suite.

Au milieu, on a représenté uniquement l'évolution de ce qu'on pourrait appeler le squelette phonétique de cette phrase, c'est-à-dire uniquement les évolutions des formants, ce qui tient à l'évolution du conduit vocal.

En bas, on a représenté l'évolution du fondamental de la parole, (fréquence mélodique ou "hauteur") et également l'évolution des durées des divers sons composant cette phrase.

La phrase en question c'est "vous arrivez plus tôt demain ?". La nuance interrogative se traduit par la montée du fondamental de la parole à la fin de la phrase. Ce qui est interprété, dans la langue française comme un indice de l'interrogation.

Il n'y a pas que le fondamental qui donne cette indication mais également les durées relatives des divers sons et de manière très secondaire l'intensité de la parole à un instant donné.

Cinq contributions pour comprendre la lecture

journées d'études des 25 – 26 – 27 février 1980



Fig 4 Spectrogramme des voyelles orales françaises prononcées par un locuteur masculin en voix laryngée et classées de l'aigu au grave selon la fréquence du 2^{ème} formant.

► Cette nouvelle vue sert à montrer que les diverses voyelles sont définies par diverses fréquences de formants. En général, on considère les deux ou trois premières résonances du conduit vocal qui sont caractéristiques des sons, des voyelles notamment. Tout à fait à gauche, vous avez en tireté les 4 premiers formants. Vous voyez qu'il y a un jeu entre ces diverses fréquences : chaque voyelle correspond à une combinaison originale de ces deux ou trois premières fréquences.

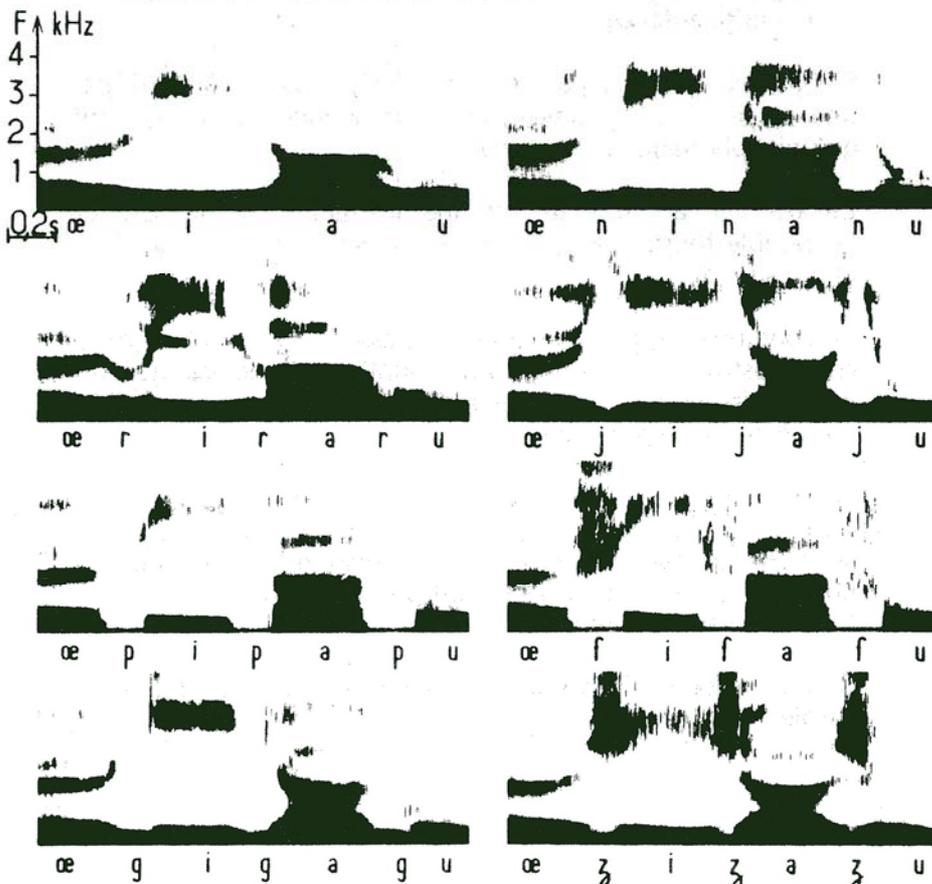


Fig. 5 Enchaînement de diverses consonnes /n/ /r/ /l/ /h/ /p/ /f/ /g/ /z/ avec les quatre voyelles principales /œ/ /i/ /a/ /u/ par rapport à l'enchaînement de ces voyelles entre elles.

Cinq contributions pour comprendre la lecture

jours d'études des 25 - 26 - 27 février 1980

► Sur la vue suivante, on montre, une fois qu'on a défini les voyelles, qu'on peut également définir la plupart des consonnes par les mouvements des formants aux alentours de ces consonnes. La série en haut à gauche, c'est "e, j, a, ou" avec un certain rythme. Ensuite, en descendant on a prononcé "e, ri, ra, ron", "e, pi, pa, pon" etc. Ceci pour montrer qu'à chaque passage d'une voyelle à une consonne est associée une transition originale qui définit la consonne.

D'une manière générale, on sait que les phonèmes n'ont pas d'existence réelle dans la matière sonore. Les phonèmes, c'est déjà quelque chose d'abstrait reconstruit mentalement à partir de l'apprentissage de la langue, vraisemblablement depuis la toute petite enfance.

Si l'on prend le "r" par exemple, il n'y a pas un instant, pas une partie de sonagramme, qui corresponde au "r" en tant qu'unité phonétique autonome.

Le "r", en tant que tel est une abstraction. Il provient de l'ensemble fourni par les évolutions des voyelles qui précèdent et qui suivent.

Ce code phonétique qui permet de passer de la réalité sonore à cette abstraction linguistique que sont les phonèmes, nous ne le connaissons pas encore, et si on le connaissait, beaucoup de problèmes seraient résolus...

► Cette nouvelle vue (fig. 6), simplement pour vous montrer que la même phrase prononcée par 3 personnes différentes, un homme, une femme et un enfant de 5 ans, donne des images foncièrement différentes.

Ce qui ne simplifie pas les problèmes de ceux qui traitent la parole.

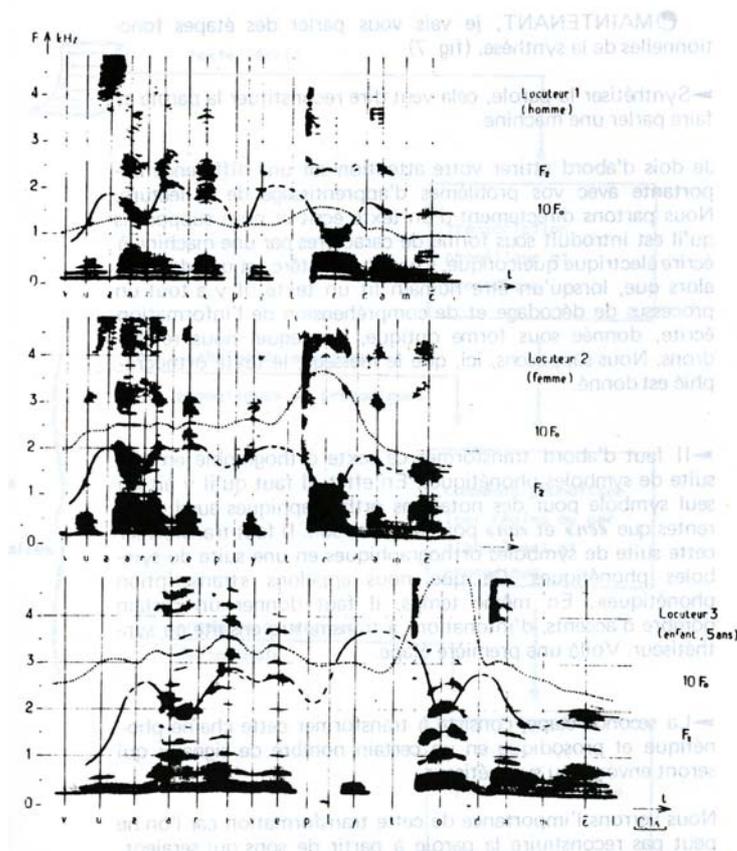


Fig 6 : La même phrase ("Vous arrivez plus tôt demain ?") prononcée par un homme, une femme et un enfant de 5 ans.

Cinq contributions pour comprendre la lecture

journées d'études des 25 – 26 – 27 février 1980

Maintenant, je vais vous parler des étapes fonctionnelles de la synthèse. (fig. 7)

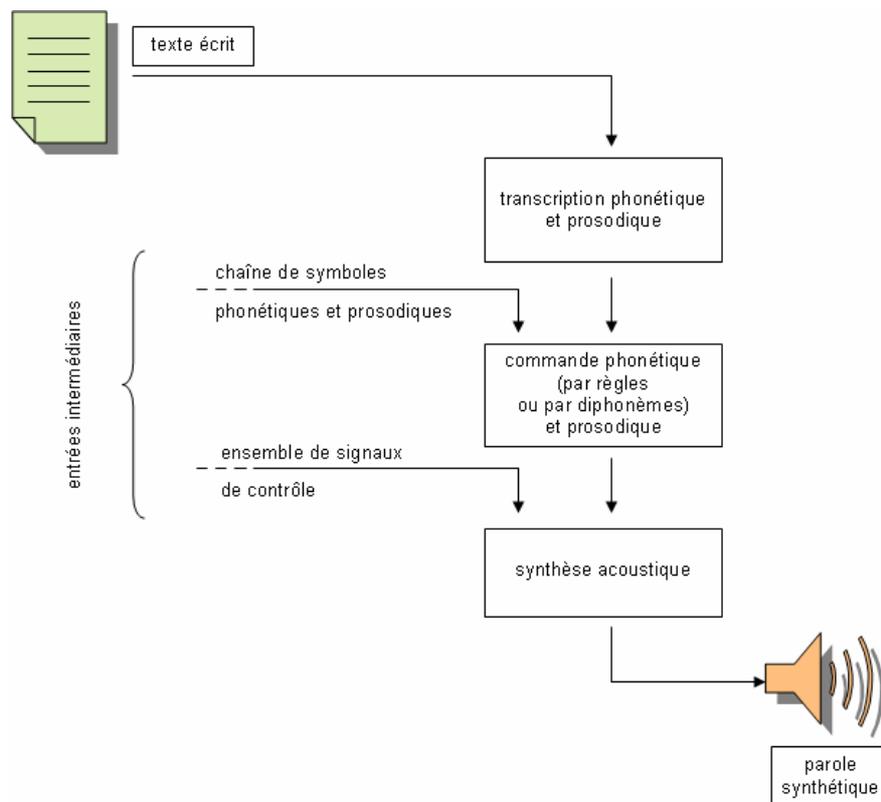


Fig 7 : Les trois étapes fonctionnelles de la synthèse à partir du texte.

► Synthétiser la parole, cela veut dire reconstituer la parole et faire parler une machine.

Je dois d'abord attirer votre attention sur une différence importante avec vos problèmes d'apprentissage de la lecture. Nous partons directement d'un texte écrit et nous supposons qu'il est introduit sous forme de caractères par une machine à écrire électrique quelconque. Chaque caractère est une donnée, alors que, lorsqu'un être humain lit un texte, il y a tout un processus de décodage et de compréhension de l'information écrite, donnée sous forme optique, sur lequel nous reviendrons. Nous supposons, ici, que le message, le texte orthographié est donné.

► Il faut d'abord transformer ce texte orthographié en une suite de symboles phonétiques. En effet, il faut qu'il y ait un seul symbole pour des notations orthographiques aussi différentes que "en" et "an" pour le même son. Il faut transformer cette suite de symboles orthographiques en une suite de symboles phonétiques. Ce que nous appelons "transcription phonétique". En même temps, il faut donner un certain nombre d'accents, d'intonations à transmettre ensuite au synthétiseur. Voilà une première étape.

► La seconde étape, consiste à transformer cette chaîne phonétique et prosodique en un certain nombre de signaux qui seront envoyés au synthétiseur.

Nous verrons l'importance de cette transformation car l'on ne peut pas reconstruire la parole à partir de sons qui seraient, chacun, associés à un phonème ou à un symbole phonétique ; il faut un intermédiaire qui élabore des signaux de contrôle du synthétiseur à partir d'une suite de symboles phonétiques.

► Tout à fait au dernier stade, arrive la synthèse acoustique, qui consiste à passer de ces signaux à un signal acoustique de parole.

Voilà donc les 3 principales étapes que nous allons maintenant détailler.

► Je vais commencer par la dernière. Pour faire comprendre ce qu'est la synthèse acoustique, je vais vous

Cinq contributions pour comprendre la lecture

Journées d'études des 25 - 26 - 27 février 1980

parler d'un appareil qui a été inventé, il y a une quarantaine d'années, et qui s'appelle le "vocoder" (voix coder, codeur de voix) qui est né dans les années 36, 37 pour les besoins du téléphone, parce qu'on voulait faire passer plusieurs communications sur une même ligne téléphonique. Il y avait donc un intérêt économique important à inventer un appareil qui réduise l'encombrement informationnel du signal.

Le "vocoder" est essentiellement un appareil qui sépare le fondamental de la parole (fréquence de vibration des cordes vocales) du spectre correspondant aux résonances apportées par le conduit vocal.

Comment fonctionne cet appareil, du moins dans sa phase primitive ?

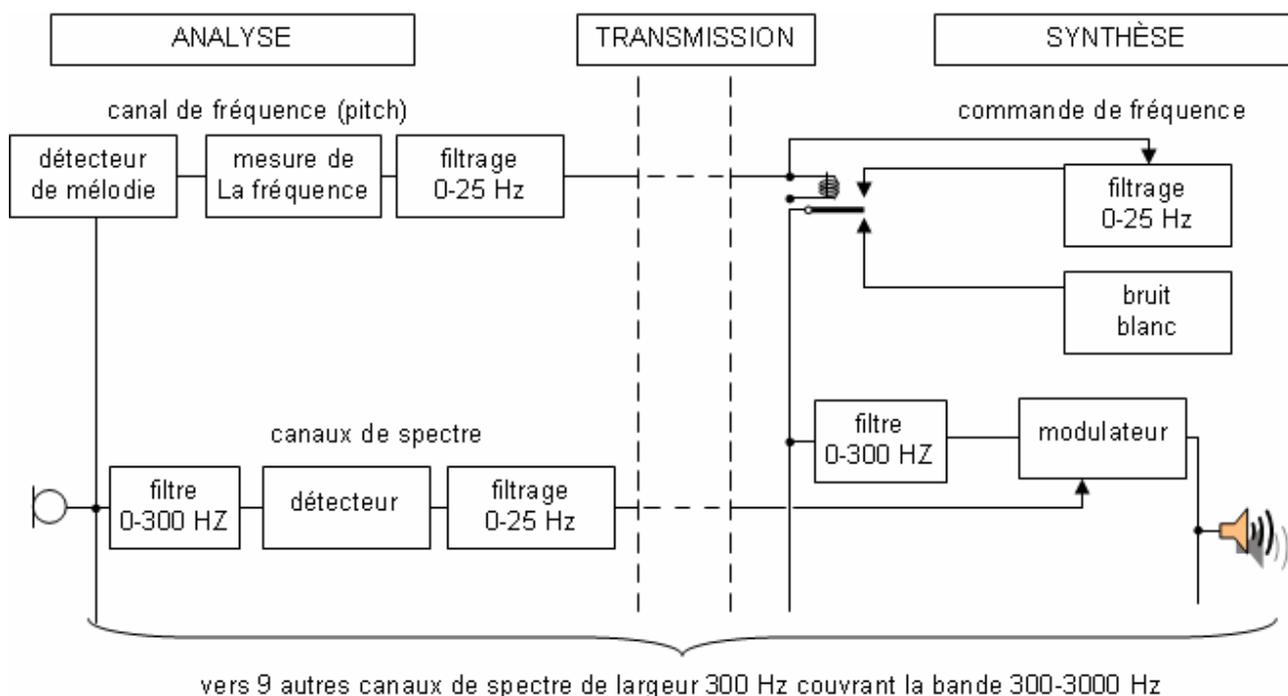


Fig : principe du VOCODER à canaux de DUDLEY

→ Le son qui vient d'un microphone passe dans un certain nombre de filtres. Le filtre, c'est une passoire. Celui que j'ai représenté, 0-300 hz, ne laisse passer que l'extrême grave, un harmonique ou deux du signal de parole ; ensuite, il y a un autre filtre qui laissera passer entre 300 et 600 hz, les harmoniques 2 et 3 de la parole. Et ainsi, une dizaine de filtres permettent de couvrir la bande fréquentielle qui correspond à l'évolution des premiers formants.

Donc on mesure, avec un certain nombre d'astuces électroniques, la force du son dans chacune de ces bandes filtrées. (On sait ce que c'est qu'un filtre, un "equalizer" en jargon Hi-fi actuel).

Il y a un autre canal, celui du haut, qui lui, va transmettre la fréquence du fondamental de la parole et la nature - bruit ou son harmonique - du signal de parole à chaque instant.

Ces informations sont transmises séparément, et codées, et elles se trouvent beaucoup moins encombrantes en matière d'information que le signal original, car on limite fortement la rapidité de leurs variations temporelles.

À l'extrémité, on a un autre appareil, qui est à peu près symétrique du précédent. C'est le synthétiseur qui, à partir des 11 signaux transmis, va reconstituer quelque chose qui ressemble beaucoup à de la parole et est en général compréhensible même s'il y a quelques erreurs, quelques distorsions : on fabrique une suite d'impulsions électriques, un son qui a la même fréquence que les cordes vocales et l'on envoie ce son dans une série de résonateurs ou de filtres analogues à ceux de l'analyse.

Au lieu de partir de l'analyse d'une voix réelle, on peut tenter de redresser à la main (on a l'aide d'un ordinateur, ce qui n'est pas plus facile), les signaux de contrôle qui alimentent la partie synthétiseur de

Cinq contributions pour comprendre la lecture

journées d'études des 25 - 26 - 27 février 1980

l'appareil.

C'est de cette manière qu'on va faire la synthèse de la parole. On ne part plus d'une voix réelle mais de quelque chose qu'on a inventé, dessiné ou élaboré d'une manière ou d'une autre.

La partie de droite de cet appareil est le prototype du synthétiseur de parole. Il y en a toutes sortes de versions qui font toutes à peu près la même chose, même si le travail ne porte pas exactement sur les mêmes paramètres.

→ Voici par exemple, (Fig. 9) un appareil qui est fonctionnellement assez voisin du synthétiseur, du vocoder, et aussi du sonagramme, de la représentation visuelle de la parole, qu'on a vu tout à l'heure.

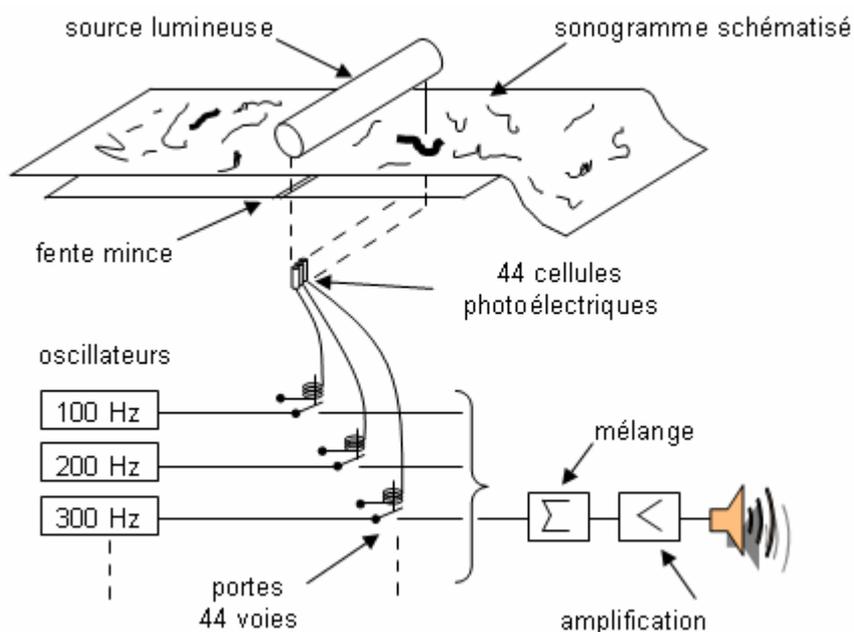


Fig 9 : L'icophone à lecture optique de Leipp, Castellengo, Liénard et Sapaly (1967).

C'est un appareil qu'on a appelé icophone parce qu'il permet de passer d'une image, d'un sonogramme, d'une analyse de parole, extrait d'une voix réelle ou fait à la main, à du son fabriqué par l'intermédiaire d'un certain nombre d'appareils électroniques (oscillateurs).

Ce qui donne de la parole intelligible que vous aurez l'occasion d'entendre tout à l'heure.

Il y a d'autres types de synthétiseurs. Par exemple, certains permettent d'imiter de manière très détaillée, le fonctionnement acoustique du conduit vocal. On considère que c'est un tuyau acoustique avec une onde plane qui se propage d'un bout à l'autre, et on imite chacune des portions de ce tuyau, etc.

Si vous allez dans les magasins de gadgets éducatifs, vous pourrez acheter un appareil qui s'appelle "speak and spell" qui doit peser moins d'un kilo et coûter 3 ou 400 francs. Il est fait en principe pour apprendre l'orthographe de l'anglais - en fait pour être vendu et faire des bénéfices... - et il propose un mot anglais et vous devez en frapper l'équivalence orthographique, si vous avez bien orthographié le mot, ce qui prouve que vous l'avez perçu sinon compris, - l'appareil dit : "bravo, vous avez gagné !", ou quelque chose d'équivalent en anglais.

Cet appareil ne fait que reproduire des mots qui ont été prononcés par une voix réelle, codés selon un processus un peu plus complexe que celui du vocoder à canaux, et enregistrés dans des mémoires extrêmement performantes - c'est une électronique merveilleuse.

Il ne s'agit donc pas de synthèse véritable, mais de reconstitution de parole codée ; la synthèse - ce mot est ambigu - n'est effectuée qu'au niveau acoustique, ce qui est relativement facile.

Voyons maintenant le dernier maillon en suivant notre schéma de tout à l'heure depuis le bas ; c'est la

Cinq contributions pour comprendre la lecture

Journées d'études des 25 - 26 - 27 février 1980

commande phonétique.

Pourquoi faut-il un intermédiaire avant de synthétiser les sons de la parole ?

Vous pouvez vous-même faire l'expérience suivante. Vous enregistrez sur un magnétophone à bandes un "ch" prolongé puis un peu plus loin un "a" prolongé. Vous coupez un morceau de chaque et vous les montez avec un morceau de ruban adhésif. Eh bien en aucun cas, ça ne vous donnera le mot chat, ça vous donnera un "ch", un clic, et puis "a". Dans la parole, on ne peut pas reconstituer un mot complet en collant bout à bout des états stables qui correspondent aux divers symboles phonétiques. La correspondance entre les symboles phonétiques et la matière sonore n'est pas une simple équivalence. Donc, il faut bien rendre compte de cette correspondance par des processus plus complexes. C'est l'objet de ce maillon de commande phonétique.

→ La vue suivante (fig. 10) montre le résultat de cette expérience sous forme imagée avec la même représentation que tout à l'heure.

Vous avez à gauche le collage du "ch" et du "a", et à droite la syllabe "cha" prononcée d'un seul trait et vous voyez que dans la syllabe prononcée de manière continue, il y a une évolution interne, une transition qui est spécifique de cette syllabe. D'ailleurs, à la limite, on peut enlever les états stables et ne laisser que la transition, et la syllabe sera encore perçue.

→ Dans cette nouvelle vue (fig. 11), il s'agit de sonagramme schématisé, pour l'icophone. Vous voyez la représentation de la syllabe "oui". Si je prononce "oui" ou si je fais prononcer à mon synthétiseur la syllabe "oui", il y a cette évolution qui est à peu près continue, qui est très différente de la juxtaposition d'un "ou" stable et d'un "i" stable. Par contre si je ne garde que la transition qui est tout à fait à droite et que je la rejoue sur mon appareil, il y a alors de grandes chances pour que vous perceviez la syllabe "oui". En d'autres termes, c'est le changement dans notre environnement sonore qui porte l'information. C'est là que doit être cherchée la parole, et non dans les états stables du signal.

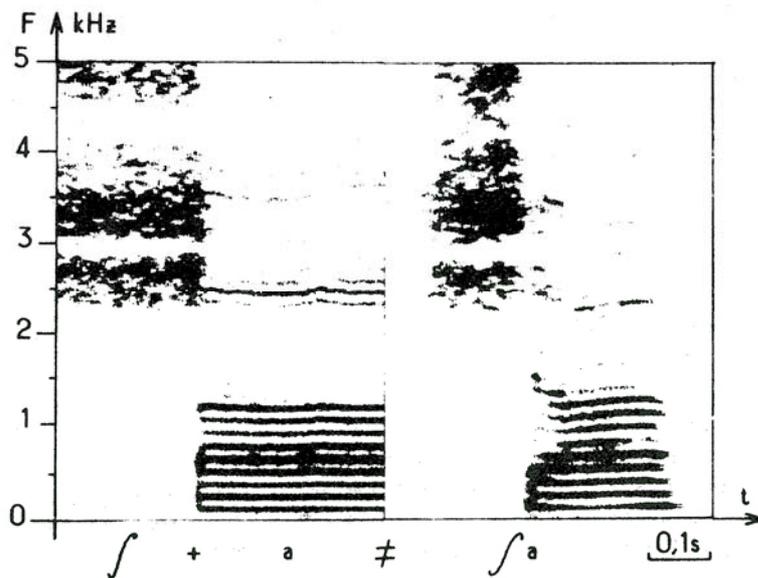


Fig. 10 Réalités des transitions phonétiques dans la parole. La juxtaposition de [ʃ] et [a] par collage de bande magnétique ne fournit pas la syllabe [ʃa] : CH + A = CHA

Cinq contributions pour comprendre la lecture

jours d'études des 25 - 26 - 27 février 1980

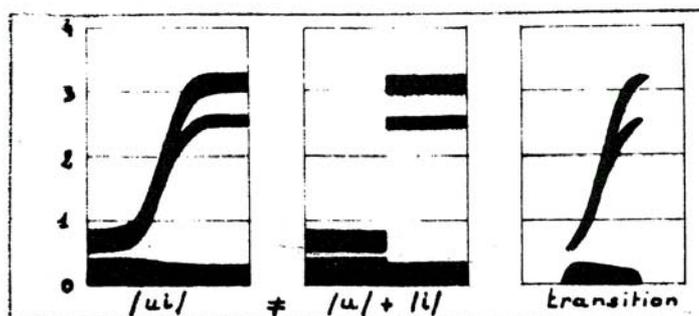


Fig. 11 - Pour constituer "oui", il suffit de la transition entre "ou" et "i"

C'est normal d'ailleurs. Nous avons un appareil, non seulement du point de vue de la perception, mais également du point de vue de la production, qui est très lent. La bouche, la langue, constituent une mécanique qui est extrêmement imprécise, lente, et il faut bien que la nature tire parti de cette imprécision. En somme, la transition reflète la continuité du mouvement de l'appareil vocal.

Il est donc indispensable de rendre compte de ces transitions si on veut synthétiser la parole. Pour cela il y a deux techniques, deux grandes approches.

► La première est connue sous le nom de synthèse par règles. Très brièvement, étant donné un certain synthétiseur qui imite plus ou moins l'appareil vocal humain, on calcule les transitions d'un état stable au suivant, selon des règles qui reflètent l'évolution, la dynamique de l'appareil vocal.

Il peut agir de règles purement articulatoires. Par exemple, la langue a telle inertie ; quand on va du "a" au "j", elle se déplace de telle manière, elle dépasse un petit peu son objectif ou elle ne l'atteint pas tout à fait, etc.

Voilà donc des ensembles de règles et l'on essaye de fabriquer de la parole avec cela.

Le problème, c'est comment découvrir les règles. La radio-cinématographie, par exemple, ne donne pas d'indications suffisamment précises. J'ai dit tout à l'heure que l'appareil vocal est très lent mais il est encore trop rapide pour la radio cinématographie. Ce n'est donc pas très commode mais c'est pourtant une voie de recherche fondamentale sur laquelle travaillent beaucoup de gens. On fait des photos, des films, pour comprendre le fonctionnement de l'appareil vocal.

► Une autre technique, beaucoup plus expéditive, consiste à dire : si on ne sait pas comment sont engendrées les transitions entre l'état stable et la parole, on va tout simplement mémoriser toutes ces transitions.

Contrairement à ce qu'on pourrait croire, cela n'est pas considérable : songez qu'en français comme dans la plupart des langues occidentales, il faut entre 30 et 40 symboles phonétiques pour noter la langue de manière convenable. Si l'on considère les doublets, c'est-à-dire les transitions d'un symbole phonétique à un autre, cela ne fait jamais qu'environ un millier (30 X 30 égal 900)... de transitions à reproduire.

Nous avons donc fait une sorte de catalogue de tous les "diphonèmes", de toutes les transitions d'un phonème au suivant que l'on peut trouver dans la langue française.

D'ailleurs, on peut estimer qu'un millier de diphonèmes est plus que suffisant, car l'on peut faire des économies à droite ou à gauche. Certains ne sont jamais utilisés selon mes statistiques, on n'utilise jamais que 5 à 600 transitions. Certains sont symétriques dans les temps et peuvent être utilisés en les retournant etc.

Cinq contributions pour comprendre la lecture

jours d'études des 25 - 26 - 27 février 1980

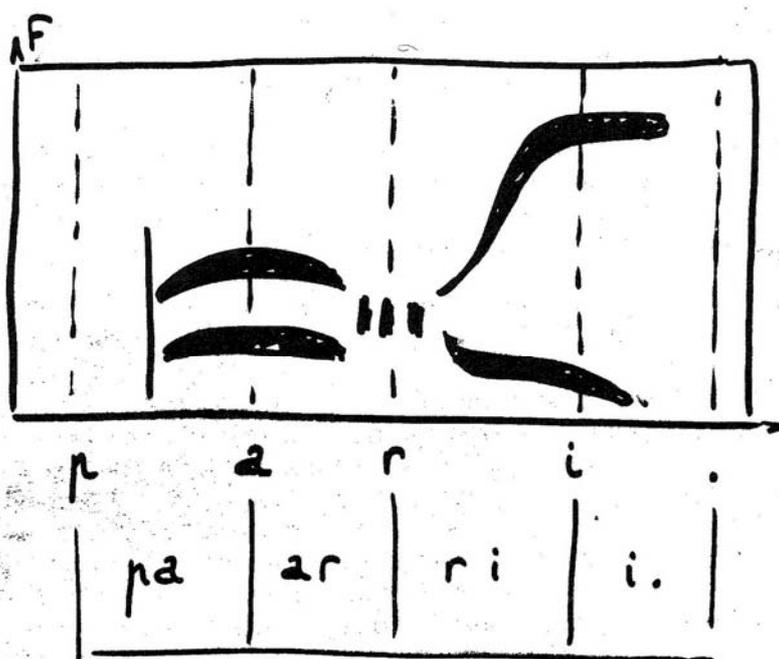


Fig. 12 - Décomposition en diphonèmes du mot !

→ Pour reconstituer un mot quelconque en français, on le décompose en "diphonèmes". Le mot Paris sera reconstitué avec la transition "pa" à laquelle on associe la transition "ar", puis la transition "ri". Avec ces 3 transitions, on reconstitue le squelette phonétique, l'image que je vous ai montrée tout à l'heure, le squelette de la parole, et ceci sera tout à fait intelligible. Voilà donc la technique.

► Maintenant, je termine par la 3^{ème} étape de la synthèse (en fait c'est la première quand on part du texte) : c'est celle qui permet de passer du texte écrit à une suite de symboles phonétiques. Cela peut paraître trivial, ou très compliqué... Il y a des règles qui permettent par exemple, de dire que p suivi de h se prononce [f] ce qui est relativement élémentaire et ne souffre pas beaucoup d'exceptions, mais prenez le cas du e, s'il est suivi de u cela fait [Ø] s'il est suivi de i, donne [ε], etc. En français le graphisme E peut apparaître dans une trentaine de contextes différents et dans chaque cas il faut savoir comment se prononce le groupe de lettres dans lequel il est inclus.

Les règles de prononciation sont à la fois très simples, - on les apprend à l'école primaire - et en même temps très nombreuses et relativement compliquées. Elles doivent tenir compte d'un contexte assez large, et pas seulement de la lettre du graphisme qui suit ou qui précède. De plus les règles admettent des exceptions, en général au niveau du mot. Par exemple *monsieur* devrait selon les règles élémentaires être prononcé *mon sieur*. Il faut donc que le programme qui passera automatiquement de l'écrit au phonétique connaisse toutes ces règles et toutes ces exceptions.

Ce programme est relativement facile à écrire. Il y en a maintenant plusieurs en France qui marchent plus ou moins bien. On arrive de cette manière à élaborer des règles extrêmement astucieuses dont on n'a même pas idée quand on réfléchit superficiellement. Dans mon laboratoire on a mis au point un programme dont le taux d'erreur est de l'ordre de 1/1000. Par exemple sur 1000 lettres, soit une demi page imprimée, on va faire 1,2 ou 3 erreurs de transcription. Donc c'est minime. Il se trouve que nous avons une langue assez régulière quoi qu'on en dise. Un tel programme pour l'anglais est beaucoup plus difficile.

Il y a des problèmes qui tiennent à la syntaxe. Devant la phrase : "*Les poules du couvent couvent.*", on ne peut pas lever l'ambiguïté phonétique entre les homographes sans faire une analyse syntaxique. On est donc obligé de prendre une option "ent" se prononcera [ã]. Il y a des astuces qui permettent d'aller

Cinq contributions pour comprendre la lecture

journées d'études des 25 – 26 – 27 février 1980

rechercher l'indice d'un pluriel dans la partie de phrase qui précède, et dans ce cas "enl" ne se prononce pas. Évidemment cela ne va pas très loin et il vaudrait mieux faire une analyse syntaxique complète, mais c'est une opération très lourde et très compliquée, Il y a donc quelques ambiguïtés qui restent "couvent/président/violent/fier/est, etc." qui ne sont pas vraiment gênant dans la mesure où elles arrivent très rarement et sont facilement rétablies par l'auditeur grâce au contexte.

Il y a un problème de sémantique également, mais je n'en connais qu'un : c'est l'ambiguïté entre des fils (de coton) et des fils (des enfants).

Il y a, en français : le problème des liaisons. Certaines sont obligatoires *des - hommes*, d'autres sont interdites *un soldat - américain*, d'autres sont facultatives *je voudrais parler - à untel*. Un certain nombre de considérations de surface permettent de faire toutes les liaisons obligatoires et d'éviter la plupart des liaisons interdites.

► Maintenant, je vais vous présenter de la parole obtenue de cette manière : (voix de l'ordinateur)

«Bonjour mesdames, bonjour mesdemoiselles, bonjour messieurs. Je suis une machine parlante, je peux prononcer n'importe quel texte écrit en français. Ma voix peut vous sembler étrange, mais on s'y habitue très vite. Et puis, je prends des cours de diction. Essayer par exemple, de répéter après moi : les chaussettes de l'archiduchesse sont sèches, archi-sèches, un chasseur sachant chasser sans son chien. Vous voyez, je parle aussi bien que le commun des mortels. En tous cas, je vous remercie de m'avoir écoutée. »

► Voilà donc un exemple. Cette voix est totalement impersonnelle mais nous y sommes tellement habitués dans notre laboratoire qu'elle nous semble tout à fait familière.

Disons que jusqu'à ces dernières années, l'objectif principal de la synthèse de la parole était de fabriquer une voix intelligible et il est vrai qu'à partir du moment où on a affaire à une machine, on peut se demander s'il est vraiment souhaitable de rendre cette machine pseudo-humaine. Je ne crois pas qu'il soit utile à tout prix d'imiter l'homme. Imaginez le cas où, téléphonant à un centre de renseignements et discutant gaiement avec l'opératrice ou l'opérateur, vous vous rendez compte au bout de cinq minutes que c'est une machine : ne serait-ce pas frustrant ?

L'intelligibilité est une propriété de la parole de synthèse qu'on peut mesurer. Il existe des tests, je n'insisterai pas là-dessus, vous devez d'ailleurs, en tant que pédagogues, en connaître, tels que les logatomes, les mots aléatoires, les paires minimales, c'est à dire des ensembles de mots qui ne diffèrent que par une consonne, etc.

► Maintenant, on voudrait faire de la parole synthétique qui ait une certaine qualité et un certain naturel, deux notions passablement différentes. La **qualité**, c'est en somme l'absence de distorsions et de parasites, c'est un timbre normal. Le naturel lui, s'adresse à un autre niveau psychologique. Ma voix est naturelle si on ne met pas en doute son origine humaine. Évidemment, la voix que je vous ai fait entendre est hautement artificielle.

Le **naturel**, pour moi, tient essentiellement à un facteur que j'ai jusqu'à présent laissé de côté, dans mon exposé, et qu'on appelle la prosodie.

Dans notre jargon, nous entendons par prosodie l'ensemble des facteurs que sont

- le rythme (la durée des divers sons de la parole).
- l'évolution de la **fréquence fondamentale**,
- accessoirement l'**intensité** de la parole (ou le timbre qui y est corrélé).

Ces 3 facteurs sont évidemment liés les uns aux autres. On parle quelquefois **d'accent**. Dans certaines langues, pas tellement en français, on doit marquer un **accent tonique** ou un **accent de mot**.

La notion d'accent est unique, et pourtant elle se manifeste par le biais de 3 facteurs qui sont fondamentalement différents. Je n'insisterai pas là-dessus.

Cinq contributions pour comprendre la lecture

journées d'études des 25 – 26 – 27 février 1980

► ► SIMPLEMENT, je voudrais vous signaler que la prosodie s'adresse à divers niveaux structurés du message.

Il y a une prosodie qui est liée au **niveau phonétique**. Par exemple les phonéticiens mettront en évidence le fait que si je dis "ada", ou "ama", ou "ala", dans ces 3 cas, le fondamental de la parole peut avoir une micro-évolution qui est fonction de la consonne. À ce niveau-là, la prosodie complète la phonétique.

Il y a aussi une prosodie de niveau lexical. Pas tellement en français, mais dans certaines langues, si on ne met pas l'accent tonique où il faut, on n'est pas compris.

Il y a une prosodie que j'appellerai **syntaxico-sémantique** qui est le reflet de certaines structures syntaxiques ou sémantiques de la phrase. (Je me refuse à trancher - je crois que la structure sémantique est plus importante et que la syntaxe n'en est que la partie la plus apparente et la plus structurée.)

Et puis, surtout, le 4^{ème} niveau qui se manifeste dans la prosodie, c'est ce qu'on appelle le niveau "**pragmatique**", c'est ce qui tient à la situation du locuteur en face de son auditeur, tous deux plongés dans un certain contexte. C'est à ce niveau que je me réfère si je pose une question ou si j'énonce une affirmation.

► ► LES TRAVAUX ACTUELS en matière de synthèse visent à augmenter la qualité et le naturel de la parole. Je peux vous présenter quelques exemples du travail en cours dans mon laboratoire, dans lesquels nous faisons évoluer l'intonation de cette parole synthétique à partir du texte. Nous essayons de regrouper les mots en fonction de considérations très simples sur la structure de surface de la phrase écrite. Nous découpons la phrase écrite en un certain nombre de groupes prosodiques et à chacun d'eux, nous attachons un certain profil, c'est-à-dire une certaine évolution de l'un des trois paramètres dont je vous parlais tout à l'heure. Se superpose à ceci, une évolution qui est liée à l'ensemble de la phrase. Par exemple si la phrase a une structure déclinante (qui baisse du début vers la fin sauf si elle se termine par une question, auquel cas elle remonte à la fin), c'est une considération que l'on peut mettre assez facilement dans un programme de production de la prosodie.

► Cette vue (fig. 13) montre comment on fabriquait les profils prosodiques en coupant la phrase par des considérations de longueur de mots. Il y a des coupes de la phrase qui sont permises, d'autres qui ne le sont pas.

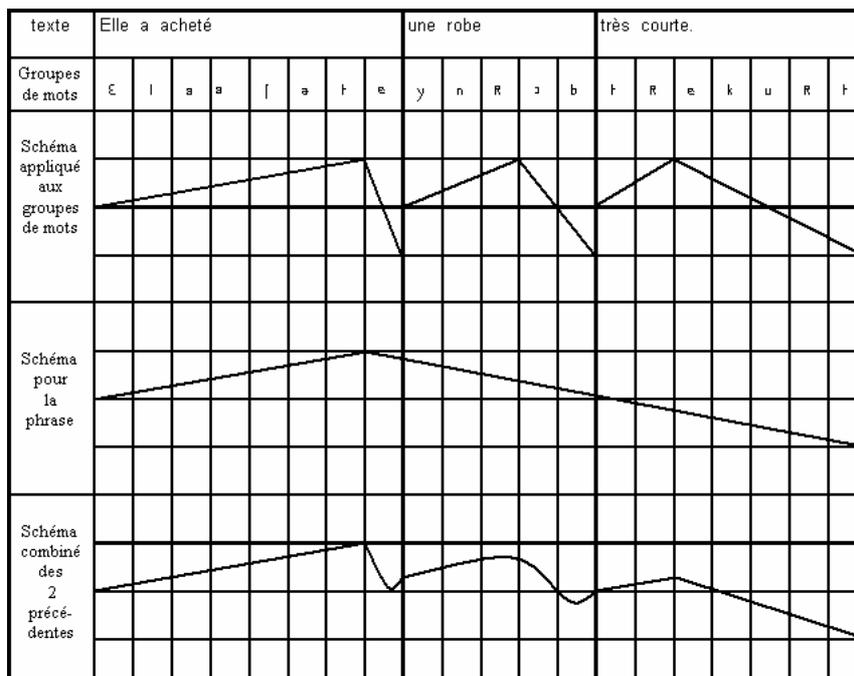


Fig. 13 - Calcul de l'intonation d'une phrase simple.

Cinq contributions pour comprendre la lecture

journées d'études des 25 – 26 – 27 février 1980

Dans "Elle a acheté une robe très courte." on a représenté 2 coupes internes permises. Il est difficile d'imaginer qu'on puisse dire "Elle a acheté une / robe très courte." Rien n'est véritablement déterministe dans des phrases comme cela parce qu'il nous manque le contexte. À chacun de ces groupes de prosodie on associe une évolution des paramètres, et on combine à cette suite d'évolutions une évolution d'ensemble de la phrase.

Voici un exemple : (voix de l'ordinateur)

"Bonjour mesdames, bonjour mesdemoiselles, bonjour messieurs. La synthèse de la parole que vous entendez est réalisée automatiquement à partir du texte écrit. Voici quelques exemples de phrases synthétisées avec prosodie automatique :

- *le petit chien a mordu le gros chat noir,*
- *la radio des voisins fait trop de bruit,*
- *à bon entendeur, salut !*
- *les médicaments sont vendus en pharmacie,*
- *le facteur fait sa tournée à bicyclette,*
- *le programmeur travaille dans son bureau,*
- *est-ce que la voix de la machine est agréable ?*
- *dans 3 heures, il fera nuit !*
- *comment allez-vous ?"*

Ce sont des essais ; vous pourriez me dire que cette voix est presque plus stupide que la première qui elle, au moins, ne prétendait pas imiter ou singer un homme.

Voilà le problème actuel. Il existe dans le monde d programmes notamment au M.I.T. à Boston qui, pour l'anglais font de la parole synthétique à partir du texte qui est meilleur que cela, qui choque moins. Mais le problème reste toujours le même, cette voix n'est ni intelligente, ni expressive. Ceci nous amène à des considérations d'un autre niveau.

Cinq contributions pour comprendre la lecture

Journées d'études des 25 – 26 – 27 février 1980

► ► DANS LA VUE SUIVANTE, (fig.14) j'ai essayé de résumer la différence entre la communication par écrit et la communication par la parole.

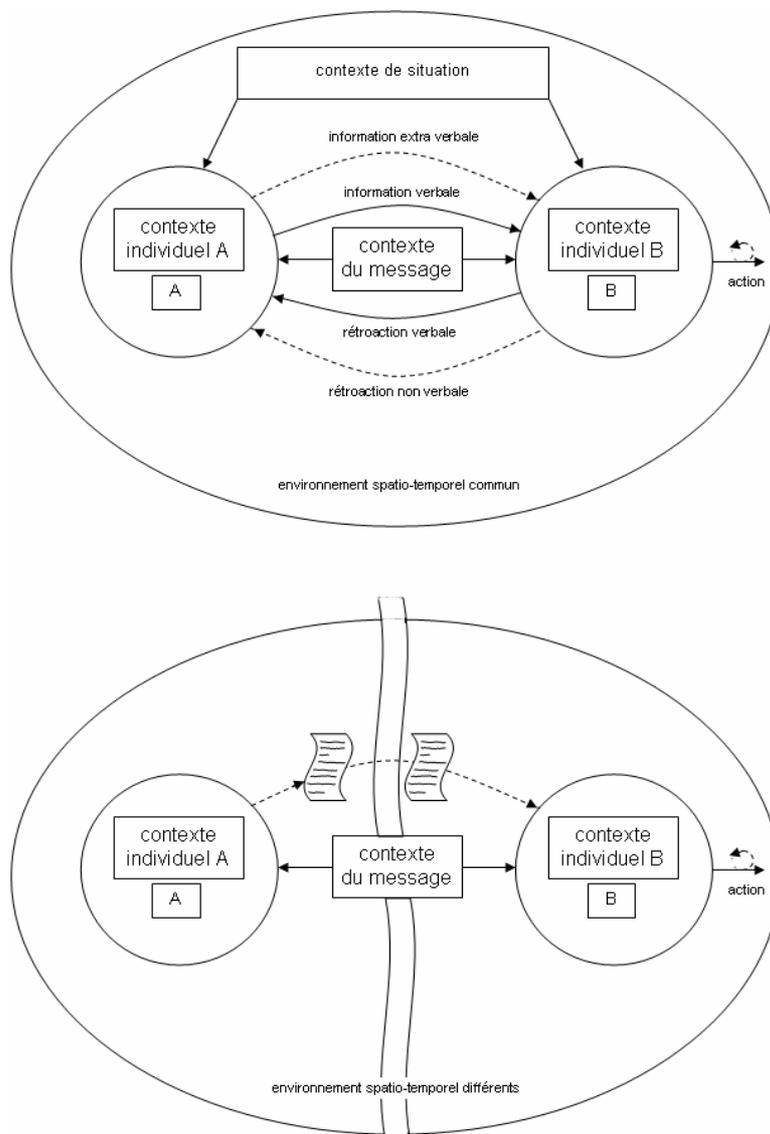


Fig 14 Communication parlée et communication écrite.

► Dans la communication écrite, vous avez un émetteur qui est A qui écrit, et un récepteur qui est l'individu B. Ces deux individus ne sont pas forcés de communiquer dans le même monde (à des années, des siècles, des milliers de km de distance). Ils sont dans des contextes complètement séparés. Évidemment, le message lui-même possède une certaine information. Il y a certaines conventions qui sont valables pour les 2 individus et qui se traduisent par une information qui passe et un contexte du message lui-même.

A priori, il n'y a pas de réaction, il y a juste transmission unidirectionnelle d'une information. Par contre, lorsqu'on communique par la parole, du moins jusqu'à l'invention du magnétophone, on est dans le même contexte, à portée de voix, et généralement à portée du regard. On est plongé dans une même situation, on partage un certain nombre d'actions à décider, etc.

Il y a alors une information verbale qui passe de A vers B, il y a une information extra-verbale et surtout, il y a cette réaction de l'auditeur sur le locuteur qui est immédiate et qui se traduit ou bien par la parole ou par des gestes etc.

Cinq contributions pour comprendre la lecture

journées d'études des 25 - 26 - 27 février 1980

Ceci permet de comprendre pourquoi le message écrit doit être totalement structuré, soumis à un code absolument explicite pourquoi les mots écrits doivent être séparés, les phrases soumises à une grammaire qui fait l'objet d'un code explicite. Mais tout cela est beaucoup moins nécessaire dans la communication parlée parce qu'elle est essentiellement interaction.

En d'autres termes, la parole sert à interagir avec nos semblables en fonction de l'environnement, alors que l'écrit sert à transmettre la connaissance.

Ce sont des fonctions extrêmement différentes et évidemment, vouloir faire de la synthèse de parole à partir du texte, c'est vouloir mettre en équivalence deux choses qui ne sont pas de même nature.

Qu'est-ce qui se passe dans la lecture à voix haute ? Nous verrons cela tout à l'heure.

Je vais vous dire un mot très bref de l'opération inverse qui est la reconnaissance automatique de la parole, ceci pour arriver à voir un peu le système complet.

Le problème de la reconnaissance de la parole est un peu différent. Il s'agit de parler à une machine et non plus de la faire parler. On s'adresse à la machine et on s'efforce que d'une certaine manière la machine comprenne ou transcrive ce que l'on dit. C'est un problème bien plus difficile que la synthèse. Tout d'abord, la parole est continue dans un texte écrit, les mots sont séparés les uns des autres ainsi que les phrases et les paragraphes.

Dans le message parlé, vous entendez des mots distincts mais **si vous regardez les sonagrammes et les analysez, vous ne voyez aucun système de pause physique entre les mots, entre les phrases, entre les groupes d'idées.** Un autre problème est que les voix sont extrêmement différentes, parce qu'on ne connaît pas encore le code qui permet de passer de l'acoustique au phonétique.

Une troisième source de difficultés, c'est que dans la parole il y a toutes sortes de niveaux mélangés, en vrac, des considérations acoustiques, phonétiques, lexicales, pragmatiques (informations sur le contexte et sur la situation). Tout cela se mélange terriblement et concourt à faire de la reconnaissance de la parole un problème sinon insoluble, du moins extrêmement difficile sous sa forme générale. Actuellement, on simplifie le problème on demande à la machine de reconnaître un mot choisi parmi une liste de mots séparés. Nous construisons à Orsay des machines de ce type : un certain nombre de mots, prononcés par un locuteur donné, isolés, séparés par une petite pause d'1/2 seconde sont introduits dans la machine lors d'une phase d'apprentissage. Ensuite le même locuteur prononce devant le micro l'un des mots de la liste (phase de reconnaissance) et la machine reconnaît de quel mot il s'agit.

→ Et même ce problème-là n'est pas tellement facile parce que quand on prononce plusieurs fois de suite le même mot, on ne produit jamais un message physiquement identique.

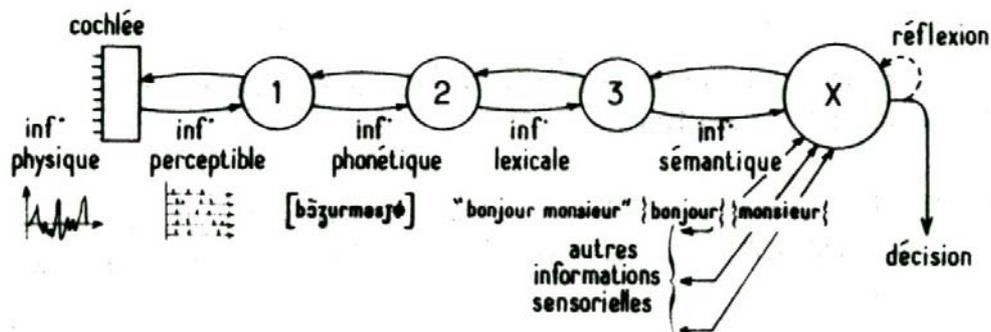


Fig 15 Les étapes de la compréhension de la parole. Chaîne phonétique

Le problème général de reconnaissance de parole n'est pas résolu. Il fait intervenir divers niveaux de reconnaissance et on est à peu près d'accord maintenant pour admettre que la reconnaissance est un processus actif, c'est-à-dire, qu'il n'y a pas simplement une suite de décodages d'informations acoustiques en informations phonétiques puis lexicales puis sémantiques, etc. (**processus ascendant**), mais que ces

Cinq contributions pour comprendre la lecture

jours d'études des 25 - 26 - 27 février 1980

transformations sont faites en prenant en compte les informations qui circulent en sens inverse (**processus descendant**, correspondant à une certaine **prévisibilité** à chaque niveau). C'est un processus actif dans la mesure où ce niveau X que j'ai représenté dans la partie supérieure de notre compréhension, qui ne traite que des objets conceptuels, envoie des informations vers la périphérie. Ce que nous appelons la compréhension ou la reconnaissance à chacun de ces niveaux que j'ai spécifiés et qui en réalité sont fondus les uns dans les autres, utilise des informations venant du bas, du support matériel à la parole en l'occurrence, et également venant du haut, c'est-à-dire notre connaissance des structures syntaxiques ou sémantiques etc.

La prévisibilité de tel mot, de tel ou tel phonème par tel ou tel contexte, tout cela fait partie d'une information descendante. Donc la compréhension est un processus actif.

Je vous détaille tout cela parce que tout à l'heure, on établira un parallèle avec la lecture visuelle, car je crois que c'est un processus assez comparable.

► ► LA VUE SUIVANTE montre comment nous essayons d'harmoniser dans nos laboratoires, la communication homme/machine par la parole. (fig. 16A et fig. 16B)

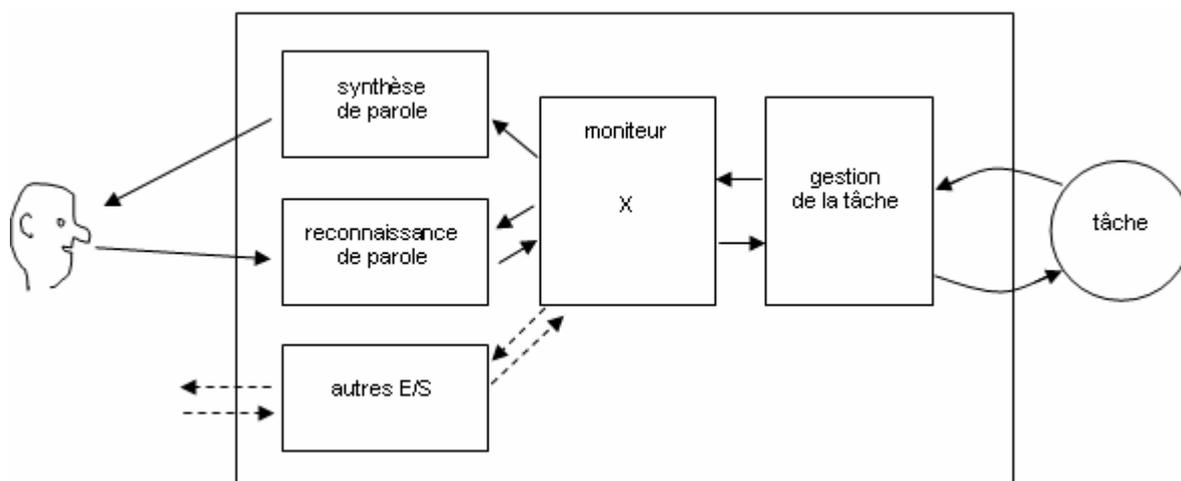


Fig 16A Les chaînes de synthèse et de reconnaissance - compréhension

En vérité on communique entre un opérateur et une certaine tâche, la machine n'est pas là que comme un intermédiaire pour assurer la gestion des messages qui vont vers l'opérateur et vers la tâche.

En bas, il y a une chaîne de reconnaissance, une analyse acoustique, un décodage phonétique, un décodage lexical et il y a la prise en compte de la prosodie, tout ce qui aboutit à une certaine représentation de nature conceptuelle du message prononcé par l'opérateur. Ce que j'ai appelé le moniteur, (c'est le niveau X de tout à l'heure) va également être capable d'engendrer des messages à partir d'une représentation conceptuelle, c'est-à-dire de fabriquer de la parole.

La tâche du moniteur est de gérer les échanges d'information entre la tâche qui peut poser des questions à l'opérateur ou lui demander des informations, et l'opérateur qui peut également envoyer des commandes vers la tâche ou encore se faire entendre du programme de gestion.

Cinq contributions pour comprendre la lecture

journées d'études des 25 – 26 – 27 février 1980

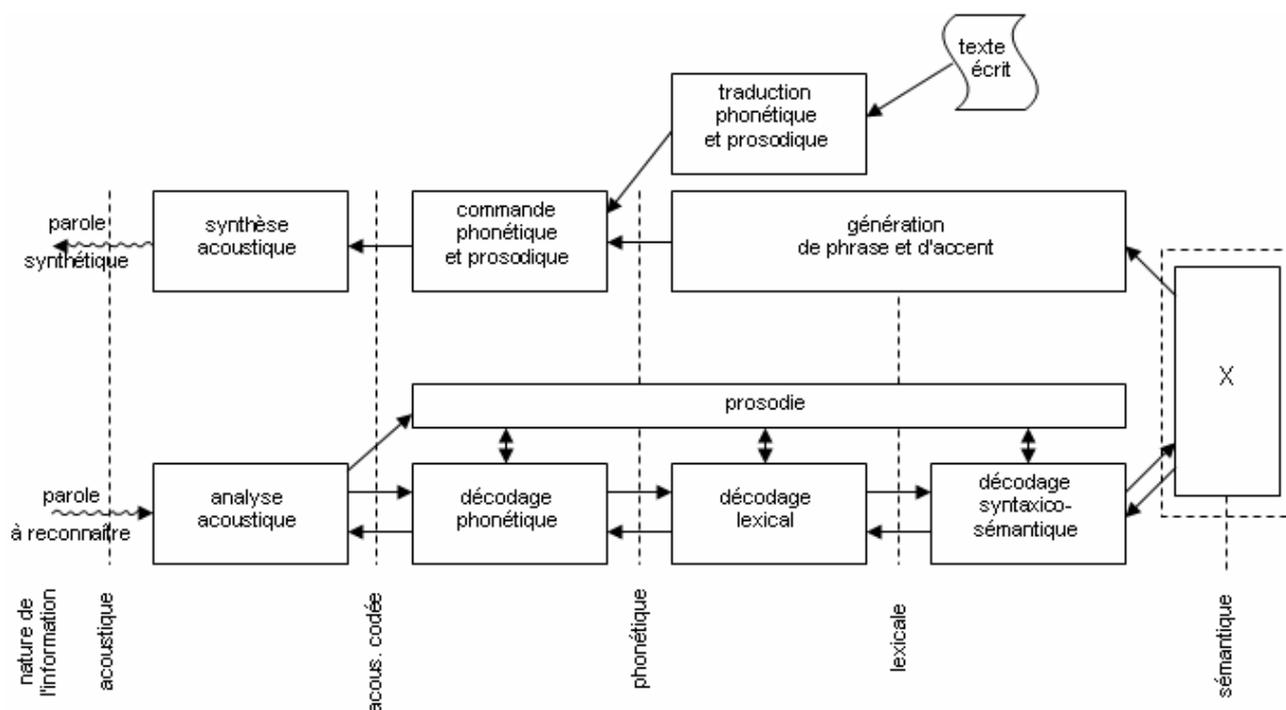


Fig 16A Schéma de la communication parlée Homme – Machine

Tout ceci suppose des programmes "intelligents". Tout ce qui est présenté là n'est pas complètement réalisé ; c'est seulement un canevas pour des travaux qui sont en train de se développer, et dans lesquels on essaie de respecter les aspects spécifiques de la communication parlée, c'est-à-dire la spontanéité, la rétroaction, la prise en compte du contexte.

Dans un schéma comme celui-ci, le système doit être capable de surveiller en permanence ce qu'on lui dit et de réagir en permanence.

Nous développons au LIMSI un programme capable de passer d'une représentation de type logico-sémantique comme aimer (*enfant, jeune, soupe*) dont les termes ne représentent pas des mots, mais des concepts, à une séquence de phonèmes correspondant à la phrase :

"Le jeune enfant aime la soupe."

et qui est capable de transformer également cette séquence en parole, en alimentant un synthétiseur avec les symboles phonétiques et les marques prosodiques calculés.

Si, à la même formule, on ajoute interrogation, négation et pluriel, le programme va sortir

"Est-ce que les jeunes enfants n'aiment pas la soupe ?"

C'est un programme que nous envisageons d'utiliser pour que le système puisse s'adresser à l'opérateur mais à partir d'une certaine représentation conceptuelle de sa tâche et de l'opérateur.

Un module comme celui que j'ai représenté par X sur les schémas doit avoir un certain modèle de la tâche et un certain modèle physiologique et psychologique de l'opérateur.

On ne peut pas imaginer d'opération intelligente à ce niveau sans que ce programme ait des informations de cette nature.

Cinq contributions pour comprendre la lecture

journées d'études des 25 – 26 – 27 février 1980

► ► JE VOUDRAIS maintenant revenir sur ce parallélisme entre la lecture à voix haute et la synthèse à partir de texte.

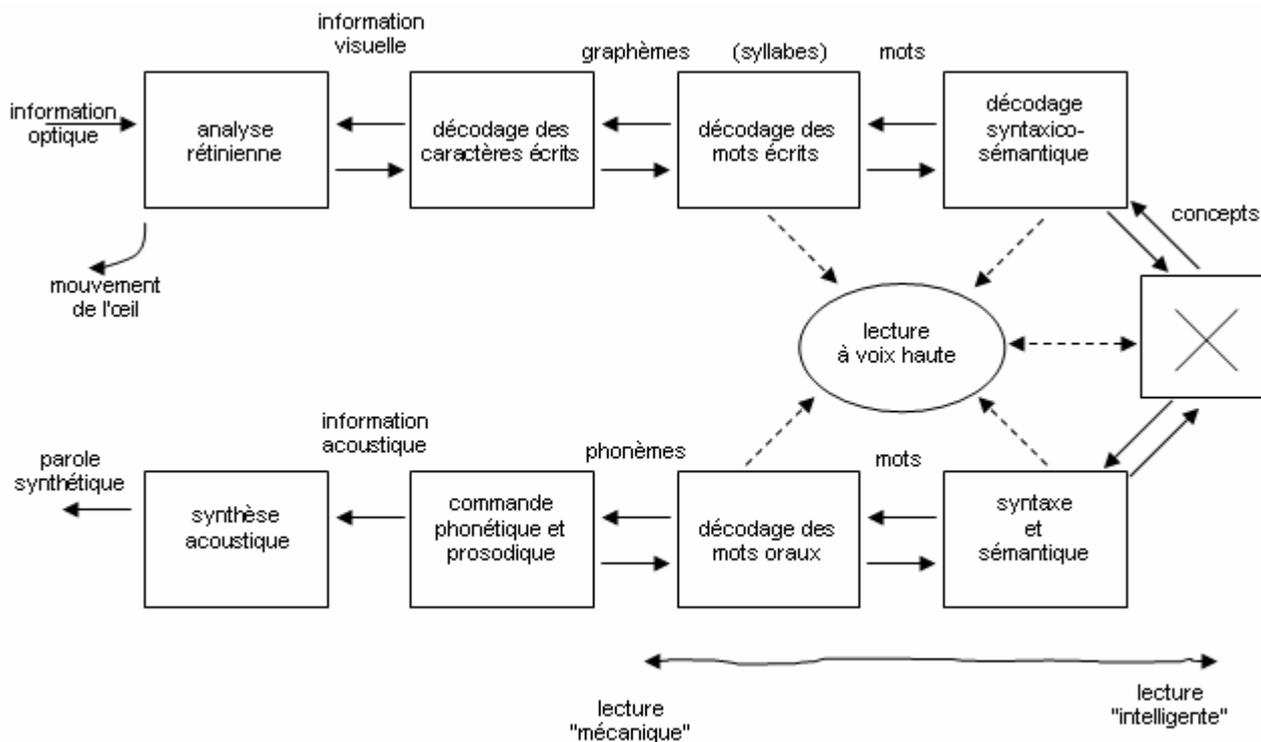


Fig 17 Schéma (hypothétique) de la "lecture à voix haute".

Je crois, en réalité, que la lecture à voix haute est d'abord un ensemble de deux processus, qui sont passablement différents. D'une part, un processus de lecture optique, lecture visuelle, d'un message orthographié, qui fait appel à un certain nombre de niveaux de décodages et qui est un processus actif comme la reconnaissance de la parole : une information optique arrive sur la rétine, elle est décodée de manière active en caractères écrits, puis en mots peut-être par l'intermédiaire de syllabes et ceci aboutit à une représentation syntaxique et puis sémantique, ce qui fait qu'au niveau supérieur de la compréhension on obtient une représentation de caractère conceptuel.

Le fait de retraduire ce message en parole fait appel à un processus de génération de message phonétique avec les étapes mentionnées plus haut.

La lecture à voix haute est donc un processus qui peut concerner l'un quelconque des niveaux à partir des syllabes. C'est une espèce de super fonction qui va mettre en correspondance ces deux chaînes, l'une de lecture et décodage du texte et l'autre de fabrication du message oral. Si on lit bêtement, on va lire syllabe par syllabe ou mot par mot sans faire intervenir aucune compréhension si on lit de manière un peu plus intelligente, on va faire intervenir la structure syntaxique, qui est quand même bien explicite dans le message orthographique si on lit très intelligemment, on va comprendre le texte avant de le reproduire sous forme de mots parlés.

Cette préparation de lecture à voix haute peut donc être effectuée à divers niveaux et puis, elle est toujours fonction d'un certain auditeur et d'un certain contexte.

Lorsque je lis un texte, à voix haute, même de la manière la plus neutre possible, je l'interprète d'une certaine manière pour mes auditeurs. Lire à voix haute, c'est déjà interpréter, c'est en particulier s'adresser à un interlocuteur ou un auditeur, présent ou non. On suppose en fait que quelqu'un est là pour vous entendre à une certaine distance, qu'il peut avoir tel ou tel type de réaction, qu'il est influençable ou non, qu'il est intelligent ou non et on module sa lecture en fonction de ces suppositions.

Cinq contributions pour comprendre la lecture

jours d'études des 25 - 26 - 27 février 1980

Nous sommes bien loin des machines à synthétiser la parole et il y a justement dans tous ces éléments que je viens d'évoquer, tout ce qui nous manque pour faire une synthèse de parole qui paraisse naturelle et intelligente.

* * *

Nota : Pour plus de détails sur la synthèse de la parole, on pourra consulter les ouvrages suivants :

- J.S. LIÉNARD - Les processus de la communication parlée, introduction à l'analyse et à la synthèse de la parole - Paris, MASSON 1977.
- R. CARRÉ, J.P. HATON, J.S. LIÉNARD - Reconnaissance et synthèse de la parole ; état de la recherche et du développement - IRIA, Collection "Les synthèses du SESORI", 1979.

DÉBAT - DISCUSSION

APRÈS L'INTERVENTION DE JEAN-SYLVAIN LIÉNARD

Le débat est ouvert sur une intervention préalable de Michel VIOLET : Les organisateurs ayant eu "*conscience qu'il n'a pas été répondu à des questions venues de la salle, écrites ou orales, dans les débats précédents, ont estimé qu'il était souhaitable de se donner des règles de moins en moins strictes pour qu'il soit traité de questions qu'on sent présentes, implicites et qui traitent du problème de la double articulation, du déchiffrement*".

Mais cette volonté de tenir compte des préoccupations, ou de la température de la salle, a empêché d'une manière regrettable que les membres du Panel ne tentent de relever dans l'exposé de M. LIÉNARD les différents éléments susceptibles d'apporter un éclairage nouveau au problème de la lecture. D'autant plus que les deux seules questions écrites n'allaient guère dans ce sens. La publication des travaux des trois journées n'ayant pas pour seul but de rendre compte fidèlement de leur déroulement, il nous a semblé utile de remédier très brièvement à la carence signalée.

Devraient en effet nous poser question :

- le fait que Jean-Sylvain LIÉNARD a constaté que l'association d'une même consonne avec les phonèmes-voyelles, constitue une série de faits physiques tout à fait différents les uns des autres et en tous cas fort éloignés de la construction bien connue de /b - a : ba/, /b - o : bo/ qu'il considère comme une abstraction sans rapport avec la réalité.
- le fait qu'il soit nécessaire de demander à la machine de prélever des indices. Ce qui, par anticipation, permettra de désambiguïser la graphème /ent/ quand il s'agit d'un pluriel de verbe, par exemple.

Le débat, on va le constater, va être centré, très vite, sur le problème : N'est-ce pas le rôle de l'école de "permettre" la réflexion de l'enfant sur le code écrit ? Un participant au Panel a posé cette question avec beaucoup d'insistance...

Nous avons choisi de ne pas résumer le débat, parce qu'il nous semble que cette insistance exprimait assez les "questions implicites de la salle" et qu'elle a permis, progressivement, des mises au point utiles.

1^{ère} question écrite : *À quoi peut servir une machine capable de transformer l'écrit en parole. Quel est le but de vos recherches ?*

Jean-Sylvain LIÉNARD - Une machine capable de transformer l'écrit en parole, à la limite, c'est déjà bien amusant... La seconde réponse c'est que ça peut rendre quelques services, par exemple aux gens qui sont aveugles. Il existe aux États-Unis un constructeur, aveugle lui-même, ancien du MIT, Kurts WEILL qui a mis au point une machine à lire le texte imprimé. Il y a donc une lecture optique de caractères (toutes les étapes dont je vous ai parlé) et une voix synthétique qui lit le texte. Vous avez entendu que cette voix n'est pas particulièrement charmeuse. Cela dit, cette machine pourra rendre des services aux non-voyants.

À côté de cela, il y a un immense marché : celui de la synthèse vocale à partir des ordinateurs. Imaginez qu'un ordinateur tienne un fichier... et vous pourrez très bien l'interroger au moyen d'un téléphone à touches, dont paraît-il nous seront bientôt munis. Il y a aussi des données beaucoup plus prosaïques, comme la gestion de stocks, la réservation de places. On économisera des intermédiaires, on pourra interroger une machine dans toutes les situations, de nuit, à distance, etc.

Quel est le but de mes recherches ?

C'est l'étude de la communication parlée dans son ensemble, c'est à dire pas seulement simulée par les machines, et, ce qui m'intéresse peut être le plus, c'est le fonctionnement de certains

Cinq contributions pour comprendre la lecture

journées d'études des 25 – 26 – 27 février 1980

processus cérébraux, de compréhension, de pensée, de perception de formes.

2^{ème} question écrite (adressée à MM. LIÉNARD et FOUCAMBERT) : *Si la prosodie est un élément parfois nécessaire à la désambiguïsation du message parlé, comment peut-elle être neutralisée dans un processus de lecture strictement idéo-visuelle, en tant qu'étape du processus ?*

Jean-Sylvain LIÉNARD - Il est certain que la prosodie est un élément qui ne contribue pas directement à l'intelligibilité de la parole au sens phonétique du terme mais qui contribue certainement à l'intelligibilité du discours. Elle permet de mettre en valeur des mots ou des portions de phrase ; de séparer les idées, de regrouper les mots, etc.

"Comment peut-elle être neutralisée dans un processus de lecture idéovisuelle ?". Est-ce qu'il y a réellement besoin de la neutraliser ? Je crois qu'on lit à voix haute pour un auditeur : s'il n'y a pas d'auditeur, il n'y a pas besoin de lire à voix haute... Pourquoi faire intervenir la prosodie qui est quelque chose d'exclusivement verbal ?...

Jean Foucambert - Je pense que cette question exprime le trouble des gens qui se posent des fausses questions sur la lecture. L'idée qui est derrière est que le message écrit n'a pas de sens et que la prosodie qu'on pourrait créer (et je me demande comment), pour quelque chose qui n'a pas de sens donnerait du sens à ce quelque chose. Je pense que le message écrit contient tous les éléments de signification. Il n'est pas nécessaire que cette prosodie existe pour trouver le sens du message écrit, puisqu'il contient précisément les éléments qui permettront de donner la prosodie.

Jean-Sylvain LIÉNARD - Je crois que ce qui est écrit et ce qui est parlé, relèvent de deux mondes différents. Pour moi, la parole c'est l'échange, c'est la relation, c'est l'instantané, et la prosodie sert essentiellement à piloter ces interactions entre interlocuteurs. S'il n'y a pas d'interlocuteurs, il y a seulement transmission en sens unique et donc, il n'y a plus besoin de prosodie. Toute l'information est explicite dans le message.

Question (portant sur l'ironie). - *...On peut lire un message et comprendre exactement le contraire.*

Jean Foucambert - Si l'on ne comprend pas l'ironie écrite, je ne vois pas comment l'on pourrait mettre une intonation qui révélerait que ce qu'on est en train de lire est ironique.

...J'ai l'impression qu'on se débat pour maintenir quelque chose dont on s'aperçoit qu'il n'a pas de raison d'être par rapport à la lecture et la réflexion sur la lecture. On essaie de s'accrocher à cette existence de l'oral qui serait nécessaire pour comprendre l'écrit. Par tous les bouts, il faut essayer de le réintroduire et je crois que, de moins en moins, on peut voir de raisons de passer par l'oral pour attribuer du sens à l'écrit. Il faut peut-être se résigner à ce que ce soit comme ça ?... Et c'est à ce prix là qu'on devient lecteur...

François Ters, auteur de la question, précise : Je pense tout simplement que, dans une certaine mesure, ce que disait M. LIÉNARD (la prosodie est nécessaire à certaines désambiguïsations syntaxiques) est en contradiction avec l'affirmation que la prosodie est fondamentalement liée à l'image acoustique et n'intervient pas en tant qu'élément du processus de la lecture. C'est là-dessus que j'aurais voulu une explication.

Jean-Sylvain LIÉNARD - Je ne suis pas sûr de bien saisir votre propos... Parmi les gens qui s'occupent de synthèse de parole et de phonétique, il existe une sorte de querelle pour savoir quels sont les facteurs de la prosodie de la parole. Beaucoup de mes collègues pensent que les facteurs principaux sont d'ordre syntaxique. C'est-à-dire qu'on va chercher dans le texte (il s'agit toujours de passer du texte écrit à de la parole) des fonctions syntaxiques, la fonction grammaticale de chaque mot et puis on cherche des évolutions prosodiques caractéristiques de ces fonctions grammaticales. Il y a beaucoup de théories qui me semblent à la fois justes et fausses parce que je crois que la syntaxe est un reflet du sens et que la prosodie est essentiellement liée au sens et à la relation avec l'auditeur.

La prosodie, à la lecture d'un texte parlé, n'a pas grand chose à voir avec la prosodie dans l'ordre d'un échange oral entre deux interlocuteurs.

Cinq contributions pour comprendre la lecture

journées d'études des 25 – 26 – 27 février 1980

Dans le premier cas la prosodie ne tient pas tant au texte qu'à l'idée que le lecteur s'en fait.

Nicole Tenesi - Je voudrais intervenir en tant que praticienne qui refuse de s'attacher à un système qui serait universel, à savoir : quand on lit, on n'oralise jamais. Prenez un cas concret comme la lecture d'une recette. On lit pour faire du couscous... Il faut... et on connaît ou non le mot qui vient. Ou bien l'enfant connaît le mot "*mouton*", ou bien il a mémorisé une forme voisine qui serait "*bouton*". Se rendant compte que bouton ne convient pas, l'énoncé de "*bouton*" déclenchera le déclic du rapprochement.

Mais s'il ne reconnaît pas du tout le mot globalement ? Que peut-il faire ? Ou dire qu'il ne sait pas ou en venir à ce que vous appelez les unités distinctives. C'est-à-dire que l'enfant va s'attacher à des parties du mot. Par exemple, il va peut-être reconnaître le *on* et je me dis que le *on*, il va l'oraliser "*pour faire du couscous il faut du ...on*" et on va peut-être provoquer le déclic - *mouton* ! -. Il me semble qu'il va passer peut-être partiellement par une oralisation.

J'essaie d'entrevoir des démarches différentes, de comprendre comment cela se passe dans la tête de l'enfant.

Michel Violet - Est-ce que la reconnaissance de cette forme "*on*" passe obligatoirement par "*ça c'est on*", prononcé ? Pourquoi l'enfant le dirait-il. Pourquoi, lorsqu'on se met à penser à l'identification d'une unité inférieure au mot songe-t-on obligatoirement au passage par le son ? Autrement dit, c'est comme si le gosse, n'identifiant pas le mot globalement, ne peut qu'avoir des repères sonores dans des unités inférieures au mot. Quand vous reconnaissez le mot *thym*, où est la forme sonore distinctive ? C'est le *h* l'indice visuel qui me permet de discriminer entre les différents homophones. Or, ce n'est pas un indice sonore.

Jean Foucambert - Surtout que vous vous donnez la partie belle en disant qu'il voit "*on*". Supposez que ce soit "*to*" qu'il reconnaisse et qu'il cherche un mot qui contienne "*to*". Avant qu'il soit arrivé à mouton, il faudra qu'il fasse le tour d'un certain nombre de condiments !

Nicole Tenesi - Je ne dis pas qu'il le fait toujours, mais qu'il peut y avoir un recours à l'oralisation, qui peut être aidante...

Jean Foucambert - Il peut avoir recours à d'autres processus. L'enfant peut chercher dans un dictionnaire, il peut demander à quelqu'un. Pourquoi penser que, prioritairement, il va trouver tout seul quelque chose qu'il ne sait pas ? Surtout que l'on sait que, dans le processus de lecture, en permanence, le lecteur qui ne comprend pas un mot ou qui n'est pas sûr d'un mot, saute ce mot. C'est-à-dire que jamais il ne s'arrête dessus. Que le meilleur conseil qu'on peut donner à un apprenti lecteur...

Nicole Tenesi - C'est quand même important pour le couscous de savoir qu'il faut du mouton ! Si on passe...

Jean Foucambert - C'est pour cela qu'il existe de nombreuses stratégies qui permettent d'attribuer du sens à l'écrit et qui sont beaucoup plus efficaces dans l'exemple que vous donnez. On n'est jamais tout seul pour fabriquer un couscous. À l'école on est tout seul devant un texte parce qu'on n'est pas là pour fabriquer un couscous mais pour lire une recette, sous prétexte de fabriquer un couscous. Mais si on est dans une situation réelle de fabrication de couscous, on se trouve avec des individus qui ont un savoir différent, on a des dictionnaires, des fichiers, etc., et je ne vois pas pourquoi on attribue une valeur particulière à ce tâtonnement qui, à partir de la reconnaissance de "*t-on - ton*", permet de trouver mouton...

Nicole Tenesi - À ce moment-là pourquoi apprendre à lire ? Mon intervention n'a pas bien été interprétée. Je ne mets pas comme stratégie première l'oralisation, du tout !

Jean Foucambert - Vous dites "*entre...autres stratégies*", et je dirais, à la limite "*en dernier recours*". Ce qui veut dire que si le gamin, voyant cette forme écrite, n'est pas allé chercher un dictionnaire, n'a pas demandé à son voisin, si personne ne peut lui dire, alors il faut qu'il se débrouille tout seul... et je dis que cette situation, sauf à l'école, est extrêmement rare.

M. Roche - Je voudrais revenir sur le problème du déchiffrement à partir de deux réflexions faites pendant le colloque. Une, de Jean Foucambert lorsqu'il a parlé de l'acquisition du

Cinq contributions pour comprendre la lecture

journées d'études des 25 – 26 – 27 février 1980

langage oral, où il a dit «*l'enfant crée une grammaire de l'oral et ne reproduit pas textuellement l'oral de l'adulte. Il construit un système à partir de celui auquel il est confronté, il le questionne, il inter-agit*». Une seconde réflexion faite par de nombreux participants au colloque qui est : l'enfant est questionneur d'écrit.

À ce sujet, je me suis dit que questionner l'écrit, ce peut-être une recherche d'information et alors c'est la lecture qui permet d'accéder au sens. Mais ça peut être aussi, pourquoi pas, dans la mesure où l'enfant est confronté au code écrit, l'envie de questionner le fonctionnement de la langue, le code écrit lui-même.

Autrement dit, **le recours au déchiffrement**, que je vois régulièrement dans les classes, qui peut être relativement spontané chez un enfant, ne correspondrait-il pas à un besoin de comprendre le code lui-même ? Le recours au déchiffrement ne serait pas un moyen ou un support pour l'apprentissage de la lecture, mais quelque chose qui correspondrait au besoin de questionner l'écrit, le code. Ce qui me gêne un peu, c'est quand on a une démarche d'apprentissage de la lecture qui est trop globale, on risque de condamner l'enfant à apprendre sans comprendre. Il apprend à lire bien sûr, il comprend ce qu'il lit, mais à côté de cela, je crois qu'en plus de la compréhension par la lecture, il y a compréhension parallèle du code. Pourquoi ne pas permettre à l'enfant ce besoin, cette volonté de savoir comment est construit ce code ?

Couramment dans les classes, avec des maîtres qui pratiquent des méthodes très différentes, (y compris ceux qui s'inspirent des thèses de Jean FOUCAMBERT, et même surtout dans ces classes-là) on voit souvent des enfants qui font des remarques de ce genre : "*C'est pareil que*", "*ça commence comme*", etc.

C'est une conséquence du savoir lire, à mon avis, et là, je suis peut-être d'accord avec Jean FOUCAMBERT, mais ça ne suffit pas. L'enfant ne veut pas apprendre sans comprendre, il veut comprendre comment ça marche.

Est-ce qu'on va étouffer cette volonté de savoir, est-ce qu'au nom de ce principe sacro-saint : pas de déchiffrement avant le CE2 ! Pas de déchiffrement pendant le cycle GS-CP-CE1 ! On va refuser de favoriser ces prises de conscience et de répondre à cette interrogation de l'enfant qui me paraît tout à fait légitime ?...

Jean Foucambert - Je suis beaucoup plus préoccupé que beaucoup d'autres personnes, de permettre à l'enfant de questionner l'écrit et de faire des hypothèses dessus, et, je suis content que vous me permettiez de préciser ce point.

Ce que je crois tout à fait dangereux, c'est la confusion, le saut logique que vous avez fait dès le début. Vous avez dit : "*il est tout à fait important que l'enfant questionne et vous avez parlé de "recours au déchiffrement"*". Or, ce sont deux choses tout à fait différentes, d'être observateur du fonctionnement de l'écrit et d'être devant l'écrit en train de tenter de lui attribuer du sens en l'oralisant à travers un mécanisme de déchiffrement.

M. Roche - Je n'ai jamais dit que le déchiffrement permettait d'accéder au sens. J'ai dit simplement que de façon spontanée, l'enfant avait recours à ce qu'on peut appeler une sorte de déchiffrement, mais que ce déchiffrement ne se faisait pas dans le but d'arriver à une accession au sens mais dans le but de questionner le code écrit, d'essayer de voir comment ça fonctionne, son organisation, les lois sous-jacentes...

Jean Foucambert - Qu'est-ce que c'est le déchiffrement dans ce cas-là ?

M. Roche - C'est arriver à l'oralisation d'une suite de signes. Ça ne veut pas dire déchiffrement pour arriver au sens. Je n'ai jamais dit cela, d'ailleurs je suis d'accord là-dessus.

Jean-Sylvain LIÉNARD apporte les précisions suivantes : Si on a cherché un système de correspondance formelle entre les lettres et les sons, c'était au début, sans trop y croire, parce qu'on pensait qu'il y avait des choses beaucoup plus compliquées, c'était en ignorant qu'on s'attaquait à ce problème-là. On se disait : tiens ce serait amusant de faire prononcer un texte écrit, de fabriquer un programme qui transforme l'écrit en parole ! Mais sans aucune illusion quant à la précision de ce qu'on obtiendrait, car on sait bien qu'il y a des "*poules qui couvent au couvent*"... En fait, il se trouve que nous avons une langue relativement simple et régulière,

Cinq contributions pour comprendre la lecture

journées d'études des 25 – 26 – 27 février 1980

malgré une orthographe à première vue assez touffue. Cela dit, l'ordinateur n'a rien à voir avec un cerveau. S'il est vrai qu'un avion vole, ce n'est pas pour cela qu'il imite le vol d'un canard, qu'il permet de comprendre le vol d'un canard.

Question - *Donc la façon dont vous construisez votre programme ne doit pas être transposée dans la stratégie d'un humain ?*

Jean-Sylvain LIÉNARD - Certes non, je ne pense pas être allé dans ce sens-là. Je vous présente ce programme avec un certain étonnement, par le fait que notre langue est assez régulière pour que ça marche dans 99 % des cas. Maintenant, ce n'est pas au niveau de la transcription des graphèmes en phonèmes qu'on a des problèmes, c'est au niveau de l'élaboration de la prosodie et là, c'est un tout autre problème. On ne peut pas donner une prosodie convenable à un texte qu'on lit sans le comprendre alors qu'on peut le prononcer convenablement.

Raymond Millot - Il y a quelque chose qui me gêne beaucoup quand j'entends dire : "*Je vois chaque jour des enfants qui s'interrogent sur le fonctionnement de la langue*". L'enfant devient une entité. Pourtant, on sait que 75 %, au moins, des enfants qui passent par l'école ne savent pas se servir de la lecture. Or, ce sont ces enfants-là qui nous importent beaucoup. C'est pourquoi il me semble très important que ces enfants-là puissent avoir une expérience de la langue passant par un usage fonctionnel pour qu'à un moment donné, ils puissent avoir, eux aussi, envie de jouer et de s'interroger sur le fonctionnement de la langue. Toute polarisation, sur ce fonctionnement, ou sur le déchiffrement constitue pour eux une cause d'échec. Il me semble que cette question devrait vous interpeller...

M. Roche - Je veux bien essayer de vous répondre. Je trouve que l'un n'empêche pas l'autre. On peut très bien avoir des pratiques fonctionnelles de lecture, telles que je pense que vous en avez dans votre école, et qui existent aussi ailleurs, sans exclure...

Jean Foucambert - Selon le dosage «un cheval, une alouette» !

M. Roche - Si vous faites les réponses pour moi... Je reprends. L'un n'exclut pas l'autre, il ne s'agit pas d'imposer le questionnement du code à tout le monde, mais si certains enfants manifestent ce questionnement, on peut l'exploiter. Pourquoi ne pas répondre à leur désir de savoir...

Raymond Millot - De là à en faire le modèle scolaire, il n'y a pas loin ! C'est cela qu'un certain nombre de personnes attendent comme conclusion car c'est beaucoup plus rassurant pour un enseignant d'avoir à enseigner que d'apporter une aide à l'apprentissage...

M. Roche - Je n'admets pas votre raisonnement, car alors on supprime complètement la réflexion sur le fonctionnement de la langue de crainte que cela soit un modèle. Les gens sont majeurs, ils sont capables d'avoir une réflexion à ce niveau qui leur permet de mettre les choses à leur exacte place. Il peut y avoir lecture fonctionnelle, un travail systématique d'aide à l'apprentissage et aussi aide à la réflexion sur le code à ceux qui le demandent, et même la favoriser !

Jean Foucambert - Je me demande si je me suis bien fait comprendre. Vous êtes en train de dire que la position qu'on essaie de défendre conteste cette réflexion sur la langue.

Alors qu'on dit depuis le début qu'elle est fondamentale. Ce qu'on dit, c'est qu'abusivement, on fait semblant de croire que la réflexion sur la langue est la même chose que le déchiffrement et que celui-ci est une stratégie de lecture. Il est évident que lorsque l'enfant a une pratique de l'oral depuis qu'il est tout petit, parallèlement à cette pratique de l'oral, il construit - vous m'avez cité - une grammaire, un système de l'oral, un ensemble de règles qu'il réinvestit, qui se réajuste, qui se transforme, etc. Et vous semblez d'accord avec la description de ce fonctionnement. Ce qui est très grave, c'est de dire ensuite : lorsque l'enfant parle il met bout à bout des phonèmes au sens que la parole serait un jeu de combinaison de phonèmes. Parallèlement, vous êtes d'accord pour ne pas le dire. De la même manière, cette réflexion sur la langue peut amener à isoler un certain nombre de choses, à tirer des règles de composition de ces choses. L'enfant fait cela à 4 ans bien avant d'être capable d'oraliser, il le fait encore à 15 ans ! Je cite toujours l'exemple de cet enfant de 5 ans 1/2 qui voulait justifier à

Cinq contributions pour comprendre la lecture

journées d'études des 25 – 26 – 27 février 1980

son camarade qu'il était devant le mot "*vient*", cachait toutes les lettres pour ne montrer que les 2 lettres "e-n" et disait : c'est *vient* puisque dedans il y a "*en*". L'enfant était bien en train d'avoir une activité de réflexion sur la langue, mais qui n'aboutissait pas du tout à quelque chose permettant un déchiffrement, puisque précisément il ne donnait pas aux éléments isolés la valeur phonétique correcte. Autrement dit, on peut distinguer la réflexion sur la langue, qui est absolument fondamentale, mais, pour qu'elle soit possible, il ne faut surtout pas l'enseigner ! Enseigner un système qui évite à l'enfant de faire des hypothèses sur le fonctionnement de la langue...

M. Roche - Pas l'enseigner, mais permettre des aides !

Jean Foucambert - Il faut favoriser cette réflexion mais elle n'est pas du tout comparable avec une activité de déchiffrement. Il ne faut pas sous ce prétexte réintroduire le déchiffrement alors que celui-ci n'existe pas dans l'activité de lecture. Je suis tenté de dire un peu méchamment ceci : Alors que le déchiffrement a été, pendant très longtemps incontesté par tous les pédagogues, depuis 10 ans, il y a un certain nombre de coups de boutoir qui, au nom de la description de ce qu'est la lecture, ont fait voler en éclat la justification du déchiffrement dans la pédagogie de la lecture. Pendant un certain temps, il y a eu de la résistance. Maintenant on plie. Les gens disent : "*absolument d'accord, le déchiffrement ne sert plus à rien dans la lecture*" mais, comme il faut que l'enfant réfléchisse sur la langue, ce que personne ne conteste, alors on réintroduit par ce biais le déchiffrement qui n'a rien à voir avec la réflexion.

En conclusion, le déchiffrement est la description d'un processus faux de lecture. La réflexion sur la langue est nécessaire mais c'est une activité méta-lexique tout à fait différente de l'activité lexicale.

Yves Parent - Je me méfie beaucoup de l'éclectisme quand on dit : "*on doit trouver ceci et cela, pratique fonctionnelle et réflexion sur la langue*". Ce n'est pas une plaisanterie quand on dit "*un cheval, une alouette*". Le cheval, tout le monde sait que dans l'école française, c'est le fonctionnement de la langue et que, c'est pour introduire des pratiques fonctionnelles qu'il faut "pousser"...

M. Roche - On ne doit pas abandonner le reste !

Y. Parent - Je crois que la langue, ça permet d'abord de communiquer, c'est comme la voiture (je sais mieux l'utiliser que réparer le moteur). J'utilise un instrument et après, éventuellement, j'essaie d'en comprendre le fonctionnement.

M. Roche - C'est très discutable votre théorie, parce qu'alors, vous êtes un infirme ! Prenons l'exemple de la bande dessinée. Il n'y a rien de mieux pour la comprendre que d'en créer soi-même, c'est pareil pour l'affiche. Il faut comprendre le fonctionnement du code, sinon...

Jean Foucambert - C'est très grave de dire cela ! J'avais raison de "tiquer" sur le mot "permettre" que vous avez employé. Je connais des méthodes d'enseignement qui **empêchent** les enfants de penser, mais je crois que, quand on met les enfants dans des situations de recours fonctionnels à quelque chose, **on ne peut pas les empêcher** de théoriser leurs pratiques. Il ne peut pas y avoir décalage entre le fait d'avoir un certain usage et celui de théoriser un certain nombre de règles rendant possible cet usage, sauf si l'on enseigne un fonctionnement sans usage. Alors là, on est sûr que l'enfant acquiert un savoir qu'il ne peut pas théoriser puisqu'il n'en a pas la pratique ! Puisque M. ROCHE est attaché à cette réflexion, il faut absolument qu'il se persuade que le déchiffrement n'étant pas du tout une stratégie de lecture, **on peut aider les enfants à commettre des actes de lecture et à réfléchir sur ceux-ci.**

M. Roche - Quand on regarde les classes, on s'aperçoit qu'elles fonctionnent tout à fait différemment de ce que vous venez de dire. C'est pourquoi je suis heureux d'avoir obtenu cette précision, qui n'est peut-être pas évidente dans votre livre, et que ce soit enregistré. Beaucoup de collègues feront peut-être une remise en cause de leur pratique...

Jean-Sylvain LIÉNARD - Je ne suis pas sûr qu'il y ait un accord aussi net entre les deux thèses...

Evelyne Andreewsky - Je veux vous rappeler qu'il y a des gens qui "lisent", qui semblent

Cinq contributions pour comprendre la lecture

journées d'études des 25 – 26 – 27 février 1980

comprendre le code écrit et qui sont incapables d'énoncer une lettre ou un quelconque logatome, ou non-mot. Ils semblent complètement incapables d'utiliser les mécanismes de déchiffrement. Il existe donc des mécanismes qui permettent, complètement indépendamment de la transformation écrit/oral, de comprendre un texte écrit, en tout cas en partie.

Ces mécanismes, s'ils existent, il faut les développer. Issus de la compréhension, ils semblent beaucoup plus rapides que les mécanismes de déchiffrement. Je crois qu'il faut les développer chez les enfants. Je ne dis pas qu'il faut éliminer tout autre mécanisme.

Je ne pense pas qu'il faut donner à un gosse d'un an les mécanismes du français, lui enseigner la grammaire de la langue, lui conjuguer les verbes, (après tout, c'est le "fonctionnement de la langue"), pour lui apprendre à comprendre le français. Je pense qu'effectivement, il va fabriquer des règles. Par exemple, l'enfant qui dit : "*j'ai perdu*", dans une famille d'intellectuels, n'a jamais entendu cela autour de lui. Il a fabriqué des règles à partir de ce qu'il a déjà entendu. De même, l'enfant qui va apprendre à lire pour comprendre, va se fabriquer, implicitement, des règles, de grammaire, de conversion entre les graphèmes et les phonèmes, et ce n'est peut-être pas indispensable de lui enseigner cela.

Jean Foucambert - Je pense que l'argument que vous utilisez doit être démonté. Vous dites que le système de correspondance est un élément marquant du code, et qu'il n'y a pas de raison de le court-circuiter. Il y a bien d'autres éléments caractéristiques et pour savoir distinguer ceux qui sont utilisés à ce moment donné de ceux qui ne le sont pas, ce n'est pas le linguiste qui peut le dire. **Ce n'est pas en décrivant la langue comme un objet qu'on peut savoir quelles sont les caractéristiques de cet objet qui, à un moment donné jouent un rôle dans quelque chose qui sera la lecture ou la parole.**

Ce qui permet de faire des choix dans ce qui est important ou non dans les caractéristiques de l'objet, ça n'est pas le linguiste qui peut le savoir, mais le psycho-linguiste, c'est-à-dire celui qui au lieu d'étudier les caractéristiques de l'objet regarde ce qui se passe quand un individu utilise cet objet.

Quand l'étude des comportements de lecteur fait apparaître qu'il n'y a pas usage de cet élément que vous considérez comme déterminant, il faut en tirer les conséquences !

Je donnerai un exemple caricatural. Un nageur, c'est quelqu'un qui se déplace dans l'eau. Un chimiste ou un physicien pourrait dire : "*c'est quand même terrible qu'on ne tienne pas compte des caractéristiques chimiques de l'eau, pour parler de la natation !*". Comment choisir quelles sont les caractéristiques chimiques de l'eau qui sont déterminantes pour l'apprentissage de la natation ? Ce n'est pas en étudiant l'eau qu'on le saura, mais en étudiant la natation. Effectivement, il y a des caractéristiques de l'eau qui sont déterminantes dans l'apprentissage de la natation : la résistance, la densité... d'autres ne jouent aucun rôle.

Il en est exactement de même pour la lecture : on a écrit énormément de choses à son sujet, non pas à partir de l'observation de ce qu'est le comportement de lecture mais à partir d'une description "à plat", froide, exhaustive, des caractéristiques de la langue écrite. Il faut peut-être admettre qu'un certain nombre de caractéristiques sont parfaitement vraies mais qu'on pourrait fort bien discuter de la réalité de la "double articulation" et de différents problèmes de ce genre. Ce qui semble important c'est de se demander si ces caractéristiques sont pertinentes dans l'activité de lecture. C'est dans ce sens qu'il faut discuter. Le fait qu'elles soient caractéristiques ne semble pas un élément probant pour dire que c'est cela qu'il faut enseigner d'abord pour accéder au code écrit.

* * *