

PHONETIQUE ACOUSTIQUE APPLIQUEE EN AUDIOPROTHESE

par Benoît Virole

INTRODUCTION

L'audition est la voie naturelle privilégiée du langage et de la communication humaine. Elle n'est pas strictement essentielle au langage dans la mesure où l'exemple des sourds utilisant la langue de signes montre bien l'indépendance des structures profondes du langage du support perceptif et effecteur. La vision et la motricité gestuelle peuvent ainsi tout aussi bien servir de vecteur au langage et il est probable que le choix de l'évolution humaine du langage phonétique ne tient qu'à des contingences externes existant à l'aube de l'humanité (communication de nuit, à distance etc.). Toujours est-il que le langage s'est étayé sur la fonction auditive et s'est modelé sur ses contraintes et possibilités. La parole, comme phénomène acoustique, se déploie ainsi dans des limites énergétiques, spectrale et temporelles compatibles avec les capacités des capteurs cochléaires. L'intensité moyenne de la parole, aux alentours de 60 dB avec une variation de plus ou moins 15 dB selon les circonstances de l'énonciation, comme ses plages fréquentielles entre 500 et 4000 Hz sont ainsi prédéterminées par les capacités sélectives optimales de la cochlée et des voies auditives. Mais il s'agit là de la parole comme objet purement acoustique. La particularité et le grand mystère du langage est que sur cet objet physique sont extraits par des processus encore mal compris, des éléments catégorisateurs, nommés indices phonétiques, qui permettent de transformer le flux physique du signal de parole en un flux phonologique permettant la génération du sens et donc l'intelligibilité de la parole.

C'est souvent à partir du moment où la perte auditive altère cette intelligibilité de la parole dans les circonstances habituelles de la communication que le sujet déficient auditif adresse à l'audioprothésiste une demande d'autant plus exigeante qu'elle engage non seulement la qualité de sa vie mais l'ensemble des relations avec les autres et bien souvent l'histoire interne de sa vie. Evaluer l'intelligibilité de la parole de patients souffrant de déficits auditifs a ainsi été depuis le début de l'audioprothèse une préoccupation majeure des audioprothésistes qui ne pouvaient se satisfaire des seules données de l'audiométrie tonale. En effet celle-ci est réalisée à partir de sons purs et ne rend ainsi pas compte des capacités phonétiques des sujets.

L'audiométrie vocale prothétique est née de cette nécessité. Elle consiste à proposer aux sujets des éléments de parole à des intensités d'émissions différentes et à compter le nombre d'erreurs sur ces éléments. Ramenés en pourcentage, ces scores donnent une indication sur l'intelligibilité globale de la parole et sur la compatibilité avec l'audiométrie tonale et les possibilités éventuelles de réadaptation prothétique. Rappelons que l'intelligibilité de la parole désigne la qualité de compréhension d'un message linguistique contenu dans un signal de parole. Elle se définit par un seuil d'intensité sonore où le sujet entend et comprend l'unité de message proposée (généralement des mots). En dessous de ce seuil existent un premier seuil de détection où le sujet perçoit quelque chose mais ne peut le reconnaître suivi d'un seuil d'audibilité où le sujet commence à reconnaître le son mais n'en comprend pas la signification.

Les différents tests utilisés en audioprothèse sont généralement des instruments utilisés en audiométrie vocale clinique, mais un certain nombre de tests ont été créés pour les besoins particuliers en audioprothèse. Les tests de phrases sont composés de phrases proposées au sujet testé qui doit les répéter. On compte le nombre de phrases justes et on peut chiffrer ainsi une capacité globale d'intelligibilité. On est ici à un niveau très éloigné du décodage acoustico-phonétique et les facteurs sémantiques et syntaxiques déterminent en grande partie la compréhension globale de la phrase (suppléance mentale). Il n'en reste pas moins que pour qu'il y ait une répétition correcte, il est nécessaire que des éléments acoustiques soient catégorisés correctement sur le plan phonétique. Ces tests sont utilisés chez les personnes pour qui la passation de tests plus difficiles (test de mots ou test phonétique) est rendue impossible du fait de leur fatigabilité. Les tests de mots (par exemple: les listes de Fournier) ne présentent pas l'inconvénient d'être fortement dépendants de ces facteurs, bien que leurs éléments soient des mots dissyllabiques porteurs de sens. Leur validité en audioprothèse est très bonne, et ce sont d'ailleurs les premiers à être utilisés dans l'histoire de l'audiométrie vocale (Rousselot, 1886-1924) mais leur usage ne permet pas le recueil des confusions phonétiques. Ils ne sont pas phonétiquement équilibrés et la projection d'un résultat sur une seule liste à l'ensemble de l'intelligibilité de la parole peut être considérée comme douteuse. Leur intérêt en audioprothèse est dans la mesure de l'intelligibilité aux différentes intensités d'émission. Si en augmentant l'intensité d'émission, on améliore pas l'intelligibilité alors le clinicien en déduit que le système auditif peut être saturé soit du fait d'une perte audiométrique trop importante pour être palliée, soit du fait de troubles qualitatifs de type distorsions cochléaires (comme le recrutement, augmentation pathologique de la sensation d'intensité) ou des troubles d'intégration du message phonétique. La représentation des résultats est une courbe tracée sur un repère où l'abscisse est l'axe des intensités d'émission en dB SPL et l'ordonnée le pourcentage de mots correctement répétés. Normalement, un sujet sans troubles auditifs arrive à 100 % d'intelligibilité à 20 dB SPL d'intensité alors qu'à 5 dB SPL il est proche de 0 %. La dynamique de l'intelligibilité de la parole s'effectue donc dans une plage de 20 dB SPL.

Le test des listes <<cochléaires>> du Prof. J.C. Lafon est très utilisé en audiophonologie ainsi qu'en audioprothèse. Elles sont censées évaluer les distorsions effectuées par les cochlées pathologiques sur le message phonétique. Ces listes sont phonétiquement équilibrées et sont étalonnées depuis longtemps sur une population qu'on peut considérer comme une bonne représentation de la population française (métropolitaine) (Lafon, 1964). Le test cochléaire comporte 20 listes de 17 mots de 3 phonèmes, correspondant à des mots monosyllabiques. On a donc une condensation entre l'unité syllabique et l'unité lexicale. En théorie, la fréquence d'occurrence des phonèmes dans le français parlé est la même à l'intérieur de chaque liste. Chaque liste peut ainsi être considérée comme un <<modèle réduit>> de la distribution des phonèmes dans la langue, ou plutôt des <<segments phonétiques>> comme l'a relevé B. Malmberg dans ses relata de la thèse de J.C. Lafon (Malmberg, 1964). On peut comparer les listes entre elles ainsi que leurs scores. Le sujet devant répéter les mots, le test présente l'inconvénient d'être tributaire des capacités articulatoires. Il doit donc en théorie être proposé aux enfants de plus de sept ans sans difficulté articulatoire. Il est demandé au testeur de noter les phonèmes substitués en API (Alphabet Phonétique International) permettant ainsi en théorie une analyse phonétique des erreurs. Ce test permet de mettre en évidence trois formes de distorsions ; celles qui sont inhérentes au seuil tonal que le Professeur Lafon a appelé distorsions liminaires et qui sont facilement corrigibles par une amplification prothétique ; celles qui persistent alors que toutes les fréquences sont perceptibles, appelées distorsions spatiales parce qu'elles dépendent du champ auditif ; celles qui augmentent avec l'intensité d'émission que le Professeur Lafon a appelé distorsions spatiales aggravées.

J.P. Dupret a construit un test similaire en ce qui concerne la composition phonologique des éléments de liste et des listes mais faits de logatomes. On peut légitimement arguer que l'usage de logatomes en enlevant effectivement un facteur sémantique rajoute par contre des facteurs neuropsychologiques - comme l'attention sélective - qui demandent en toute rigueur à nouveau un contrôle (Dupret, 1980). L'analyse des confusions dans le test phonétique du Professeur Lafon consistant à noter les phonèmes substitutifs ne tient pas compte de la dépendance de ces erreurs du contexte phonétique dans lequel elles se produisent. Ce n'est pas le cas d'autres tests dont une des particularités est le contrôle strict du contexte phonétique. Parmi ceux-ci, le développement du test de Voiers et de Rossi constitue un bon exemple de transfert entre la phonétique et l'audiologie. Voiers (1968) a travaillé sur la discrimination des indices consonantiques élémentaires à partir de segments phonétiques perturbés par différents bruits de masque. Il a élaboré le <<Diagnostic Rhyme Test>> (DRT) pour tester l'intelligibilité de la parole dans des conditions d'écoute différente. Peckels et Rossi (1972) ont adapté à la langue française le DRT de Voiers afin d'évaluer la qualité d'une parole synthétique. Leur test utilise des paires minimales et dans l'interprétation des résultats le contexte phonémique est pris en compte. Leur test permet de détecter les indices acoustiques pertinents dans les oppositions consonantiques et établit un diagnostic sur le fonctionnement des systèmes de synthèse vocale. Cartier et Rossi en 1973, modifie le test en faisant un test à choix ternaire et en normalisant les conditions de passation. M. Rossi enfin a l'idée à la suite de la connaissance des premiers travaux du Professeur Lafon sur le Test Phonétique de créer une application de ce test de diagnostic par paires minimales à l'étude des surdités (1975). Fontanez et Beraha l'utilisèrent alors en audioprothèse pour l'effet des distorsions prothétiques (1977).

Ces tests sont construits sur l'unité trait distinctif et sont donc plus propices à l'analyse phonétique des confusions. Mais ils sont par contre plus longs et plus difficiles à utiliser en clinique. De plus, les réponses à choix forcés ne peuvent être totalement satisfaisantes pour une analyse phonétique car on conditionne l'éventail des réponses. On voit que chaque test, ou famille de tests, possède des qualités et des limites intrinsèques qui les rendent tous complémentaires et qui offrent à l'audioprothésiste une palette d'outils d'investigation dans lequel il peut choisir selon les circonstances les mieux adaptés aux particularités cliniques de son patient et à sa propre méthodologie d'appareillage.

Ce rapide parcours, non exhaustif, autour des tests vocaux en audioprothèse illustre ainsi la variabilité des instruments et la nécessité de comprendre leur structure et leurs fonctions sur le plan phonétique. L'importance de la compréhension de la parole dans la demande audioprothétique implique ainsi une certaine familiarité de l'audioprothésiste avec les sciences du langage et en particulier la phonétique acoustique. Nous décrirons ici les notions de base en phonétique acoustique nécessaires à l'audioprothésiste et présenterons les grandes lignes des théories expliquant le processus encore mystérieux de la perception phonétique avant de conclure par l'exposé de quelques grands problèmes pratiques de l'audiophonologie impliquant la catégorisation phonétique.

I. UNITES DE PHONETIQUE ACOUSTIQUE

1.1 Les phonèmes

Chaque mot, unité de première articulation entre signifiant et signifié, est composé de la succession dans le temps d'unités minimales de deuxième articulation appelées phonèmes. La substitution d'un phonème par un autre dans un mot suffit à faire changer ce mot soit vers une forme qui n'appartient pas au lexique soit vers un autre mot. Ainsi, dans le mot <<bile>> /b/ est un phonème car il distingue <<bile>> de <<ville>>. Ce n'est donc pas la valeur absolue d'un phonème qui compte mais sa différence avec tous les autres phonèmes du système phonologique auquel il appartient. Le phonème est donc d'abord une unité distinctive avant d'être une unité pourvue d'une matérialité acoustique. Certaines langues possèdent plus de phonèmes que d'autres mais toutes ont en commun le fait qu'elles construisent tous leurs mots à partir d'un nombre limité de phonèmes (36 pour le français).

1.2 Les traits phonétiques

Chaque phonème de chaque système phonétique est constitué par un faisceau de traits distinctifs binaires pouvant chacun prendre une valeur positive ou négative. Les traits distinctifs acoustiques sont les correspondants au niveau acoustique des traits articulatoires pour décrire la production de la parole. Un phonème peut donc être décrit soit en traits distinctifs articulatoires, tels les traits <<occlusif>> <<vélaire>>, ou en traits acoustiques, tels les traits <<discontinu>> <<compact>>.

Du point de vue perceptif, la description en traits acoustiques présente plus de légitimité que les traits articulatoires, même si en pratique et par commodité de langage on désigne parfois les phonèmes par leurs traits articulatoires. Ainsi, il est plus en usage de parler de <<occlusive sourde dentale>> que de phonème <<discontinu, non voisé, aiguë >> alors qu'il s'agit du même phonème. La conception des phonèmes comme étant non plus uniquement des unités phonétiques distinctives mais comme composés de faisceaux de traits distinctifs date du congrès de Prague en 1930 (Cf. Trubetzkoy, 1939). Les traits distinctifs acoustiques ont été ensuite décrits et formalisés par les travaux de Jakobson (1958). Le système de Jakobson a été très critiqué pour son binarisme et sa ressemblance avec les systèmes moléculaires <<atomiques>>. De nombreux auteurs y voient une métaphore séduisante mais simplificatrice. Il est vrai que la bivalence des traits simplifie fortement la description phonétique. Ainsi un phonème X est soit nasal soit oral mais ne peut pas être les deux à la fois, puisque nasal et oral constituent les deux valeurs opposées d'un même trait distinctif, celui de la nasalité. La validité linguistique des traits distinctifs pour décrire la langue et expliquer à un niveau profond les règles de phonologie générative (Chomsky, 1968) s'est vue redoublée par l'existence pour certains chercheurs d'une validité fonctionnelle. Ainsi, pour Fant, la structure en traits des phonèmes découle des nécessités de codage: <<Le besoin d'efficacité fait que la perception des indices s'appuie naturellement sur les différentiels polaires, en sorte que le décodeur natif peut choisir entre les deux membres de chaque opposition binaire au moyen d'une simple décision par oui ou non. Ainsi, se trouve satisfaite la nécessité d'une simplicité maximale dans les stratégies ordinaires de ceux qui utilisent le langage, d'autant que le nombre d'oppositions différentes dans une langue quelconque est toujours préétabli et strictement limité>> (Cité in Jakobson, 1980). Les travaux de Miller et Nicely sur les confusions entre consonnes anglaises ont attesté l'indépendance de chaque trait vis-à-vis des autres des lors que l'on fait une analyse de covariance des traits impliqués dans les confusions perceptives (Miller & Nicely, 1955).

Les traits distinctifs nécessaires et suffisants pour décrire un système phonologique sont les suivants :

VOCALIQUE / NON VOCALIQUE (Voc /NVc). Un phonème est vocalique si sa source phonatoire est unique, périodique, de faible amortissement et d'attaque non abrupte. Les voyelles sont des cas typiques, mais certaines consonnes peuvent être aussi vocaliques comme les liquides par exemple /L/ /R/. Par exemple, /i/ /a/ /u/ sont des voyelles. /p/ /t/ /k/ sont des consonnes.

CONSONANTIQUE / NON CONSONANTIQUE (C /NCc). Un phonème est consonantique si la fonction de transfert entre l'excitation glottale de la source sonore et l'énergie acoustique à l'ouverture des lèvres comporte des zéros et des antirésonances. Ce qui signifie la présence de pôles des bruits non périodiques et ce qui exclu toutes les voyelles.

CONTINU / DISCONTINU (Cu / Dcu). Une attaque abrupte distingue les occlusives qui sont discontinues des fricatives qui sont continues. Ce trait distingue aussi /L/ qui est continu de /R/ qui est discontinu. Il est redoublé dans le système des traits distinctifs de Rossi pour le français par le trait interrompu / non interrompu qui est déterminé par la vitesse des transitions (Rossi, 1977). Ce redoublement de cette opposition permet en fait de décrire des événements acoustiques produits par des techniques articulatoires différentes qui peuvent ou non être employées ensemble pour un même phonème. En fait, le trait interrompu / non interrompu permet de distinguer le caractère occlusif des nasales de celui des consonnes plosives orales. /p/ /t/ /k/ sont des occlusives et présentent donc le trait discontinu.

GLOTTALISE / NON GLOTTALISE . Un phonème est glottalisé lorsque le flux d'air est interrompu par la fermeture de la glotte. Ce trait caractérise certaines occlusives des langues amérindiennes, africaines, caucasiennes.

STRIDENT / MAT. La stridence est liée à la turbulence et est donc le trait caractéristique des fricatives. Dans le système des consonnes, l'écart est maximal entre les occlusives mates et les fricatives stridentes. Ce trait est peu utilisé pour la description des phonèmes du français.

VOISE / NON VOISE (Vo/Nvo). Le trait de voisement est caractérisé par la présence de la vibration des cordes vocales qui se visualise sur les spectrogrammes par la barre de voisement. En fait, il existe d'autres indices du voisement qui seront décrits plus loin. En tous cas, le trait de voisement se traduit fréquemment par le renforcement des fréquences graves autour de 250 Hz. Par exemple /b/ est voisé, mais /p/ ne l'est pas.

NASAL / ORAL (N/O). Le trait de nasalité est caractérisé par la mise en résonance nasal. Les effets sur le spectre acoustique de la nasalisation ont longtemps divisé les chercheurs en phonétique. De nombreuses langues ne possèdent pas de phonèmes marqués par le trait de nasalité. Pour les consonnes, l'opposition minimale de la nasalité est entre /n/ et /L/.

COMPACT / DIFFUS (Cp/D). Le trait de compacité est de nature spectrale et exprime une accumulation d'énergie acoustique dans la région centrale du spectre où se regroupent les formants des voyelles compactes et se dirigent les transitions F2 et F3 des consonnes compactes.

AIGU / GRAVE (A/G). Ce trait de tonalité concerne la dominance de la partie haute ou basse du spectre sur l'autre. Pour les voyelles, c'est la position fréquentielle du F2 qui le contrôle. Pour les consonnes, ce sont la position fréquentielle du pôle de bruit ou turbulence et/ou la direction des transitions de F2 qui déterminent la valence aiguë ou grave de ce trait de tonalité. Pour /k/ et /g/, la valence de ce trait ne peut être déterminée qu'en fonction de la connaissance de l'environnement vocalique.

BEMOLISE / NON BEMOLISE (b/Nb). Le trait de bémolisation est également un trait de tonalité qui se caractérise par un déplacement vers le bas des formants et singulièrement en français par le déplacement du deuxième formant des voyelles F2. Ainsi /i/ bémolisé devient /y/, par le transfert vers le bas d'un demi-ton du deuxième formant.

DIESE / NON DIESE. Ce trait constitue l'inverse du trait précédent, il correspond à un déplacement vers le haut du deuxième formant. Il s'obtient au niveau articulaire par une palatisation et dilatation du larynx.

TENDU / LACHE. Le trait *tu* est caractérisé par un son plus long et pour les occlusives par une force plus grande du bruit d'explosion. Pour les voyelles, il oppose en Français le /a/ de pâte qui est tendu à celui de patte qui est lâche.

COMPLEXE / NON COMPLEXE (Cx / NCx). Ce trait a été proposé par B. Malmberg pour décrire le système des voyelles du français. Il permet de créer des oppositions entre les voyelles à l'intérieur des séries de voyelles antérieures et postérieures (degré d'ouverture). Ce trait est peu usité.

Les traits distinctifs peuvent être ou non pertinents pour un phonème donné. En français, certains de ces traits ne sont pas pertinents. Ainsi, le trait de glottalisation n'est pas utilisé pour différencier deux phonèmes alors que dans les langues africaines les systèmes phonétiques comportent des consonnes glottales. Par contre dans des rééducations de voix pathologiques, la distinction par ce trait de deux phonèmes assimilés est parfois utilisée. Plusieurs phonèmes présentent des indéterminations dans l'établissement de leur identification en traits distinctifs qui ne peut être levée que par la connaissance du contexte. La neutralisation est la perte d'une opposition distinctive dans un contexte phonétique déterminé. Là où l'opposition est neutralisée apparaît alors une seule unité, un archiphonème, qui coiffe les deux unités neutralisées. En français, et pour les listes de mots utilisées dans les épreuves d'audiométrie phonétique, le problème de la neutralisation se pose surtout au niveau de la complémentarité partielle qui exerce une contrainte phonologique sur l'éventail des confusions potentielles. L'usage des traits distinctifs pour l'analyse des confusions phonétiques est intéressant par le petit nombre d'unités, et par leur binarisme. Par contre les traits distinctifs présentent de nombreux désavantages ; l'imprécision de leurs corrélats acoustiques et leur absence de corrélation fine avec les phénomènes de coarticulation sont les principaux défauts de ce niveau d'analyse qui doit être complété par une description plus proche du niveau acoustique.

1.3 Les indices acoustiques

Les indices acoustiques sont des éléments du substrat acoustique de la parole qui sont investis dans le processus de reconnaissance de la parole d'une fonction discriminative. Ils sont présents de façon physique dans le signal de parole tel que l'on peut le visualiser par sonagramme. En extrayant de ces sonagrammes, les objets sonores ayant une valeur catégorielle, on peut établir une cartographie des indices acoustiques. On s'aperçoit que leur nombre est largement redondant, ce qui explique que la parole puisse être perturbée dans de larges proportions, par du bruit masquant par exemple, sans que l'intelligibilité puisse en pâtir. Ainsi le trait distinctif voisé n'est pas uniquement déterminé par la perception du fondamental laryngé mais par tout un faisceau d'indices. En théorie, on peut concevoir que la perception d'un seul de ces indices suffit à déclencher le processus de catégorisation voisé/non voisé et donc aboutir à l'identification correcte. La liste des indices acoustiques du français a été établie par Delattre (1958) et remise à jour par les nombreux travaux actuels concernant la synthèse et la reconnaissance de la parole.

A - VOYELLES

A.1 Hauteur formantique

Les voyelles sont les phonèmes les plus simples à décrire en termes d'indices acoustiques car elles sont des états stationnaires présentant ainsi des <<formants>>, régions de fréquence de plus grande intensité repérables sur les spectrogrammes de parole. En fait, il faudra décrire également les amplitudes de chaque formant et non uniquement sa hauteur fréquentielle. Mais comme les contraintes articulatoires redent les amplitudes dépendantes des fréquences, il est possible de restreindre la description des indices acoustiques des voyelles aux seules hauteurs fréquentielles des formants. En effet, on sait depuis Fant (1960) et son analyse de l'acoustique de la production de la parole que le redoublement d'un formant à sa fréquence double augmente de 6 dB SPL le niveau d'intensité des formants supérieurs. L'amplitude des formants est ainsi complètement définie par la distance entre les fréquences formantiques et l'on peut négliger l'intensité des formants pour les caractériser et se contenter de la notation de la fréquence. Cependant, l'intensité du premier formant sera une donnée importante pour la spécification de la nasalité vocalique. En ce qui concerne la modification du timbre par les variations prosodiques ou mélodiques de la fréquence fondamentale, il ressort que les positions des frontières phonétiques entre les voyelles mesurées à partir de la position présumée du formant auditif semblent se décaler légèrement vers les hautes fréquences lorsque la fréquence fondamentale augmente (Lonchamp, 1988). Les voyelles sont identifiées par trois formants F1, F2, F3. Mais F3 ne joue un rôle appréciable dans l'identification que pour les voyelles qui ont F2 et F3 très rapprochés, donc les voyelles antérieures (Delattre, 1958). Lorsque F2 et F3 sont rapprochés, comme dans les voyelles antérieures, la perception de leur somme équivaut à peu près à la perception d'un seul formant dont la fréquence sera intermédiaire entre F2 et F3.

A.2 Formants de nasalité

Les voyelles nasales sont plus complexes à décrire du fait des effets de la mise en résonance de la cavité nasale. La nasalité des voyelles est marquée selon Delattre par la faiblesse énergétique du premier formant et par l'existence de trois formants de nasalité : FN1 à 250 Hz, FN2 à 1000 Hz, FN3 à 2000 Hz. Ces indices acoustiques de la nasalité vocalique ont été très discutés (Lonchamp, 1988). Il ressort néanmoins que l'abaissement du voile du palais provoque une perturbation du spectre de la voyelle. La validité perceptive de ces indices acoustiques de la nasalité vocalique a été étudiée par Maeda (1982). Il ressort que cette perturbation spectrale est relativement audible pour les voyelles fermées mais moins pour les voyelles ouvertes. Dans ce dernier cas, c'est surtout la faiblesse d'intensité du premier formant qui sera l'indice catégorisateur de la nasalité. Une nasalisation adéquate des voyelles ouvertes en simulation par des modèles articulatoires nécessite un plus grand passage vélo-pharyngal que pour les voyelles fermées (Maeda, 1982).

A.3 Indice de compacité

Les hauteurs fréquentielles des formants sont liées au volume des cavités de résonance mises en jeu dans l'articulation. Le rapport entre F1 et F2 définit ce qu'on appelle l'indice de compacité qui donne une assez bonne information sur l'intensité spécifique de chaque voyelle. Les voyelles les plus compactes sont les plus intenses. Le rôle du troisième formant est particulier. En effet sa présence est indispensable, en reconnaissance de parole synthétique, pour le /i/ mais est inutile pour la reconnaissance du /y/.

B - CONSONNES

B.1 Turbulence ou pôle de bruit.

En tant que phénomènes aperiodiques, les consonnes sont caractérisées par la présence de bruits sans définition périodique précise. On les définit alors comme « turbulence », et on les caractérise par une plage fréquentielle de résonance maximale, une intensité et une durée. Certains auteurs pensent que c'est la limite inférieure du bruit de friction (discontinuité bruit/silence) qui constitue l'indice discriminatif, ce qui expliquera pourquoi il est possible de reconnaître des /s/ avec un bande passante coupant à 3500 Hz comme au téléphone. En fait, pour les turbulences longues caractérisant les fricatives, on considère actuellement que c'est le noyau de densité énergétique maximale, centre de gravité de la turbulence (zone de résonance maximale) qui constitue le vecteur discriminatif. Leur définition comme « turbulence » ouvre également la question de leur structure « dissipative » éventuellement « chaotique », et du lien de cette structure interne avec leur fonction discriminative. Les fricatives présentent un bruit produit par les turbulences de l'air au niveau d'un rétrécissement du conduit vocal. Il sera accompagné ou non d'une source périodique (voisement). L'allure du spectre est continu, d'une durée supérieure à 35 ms et d'hauteur variable selon les lieux d'articulation.

B.2 Locus et directions de transitions formantiques

Les transitions formantiques sont des éléments essentiels de la reconnaissance phonétique. Les travaux de synthèse ont en effet montré que les transitions, à elles seules, suffisaient à transmettre les « squelettes sémantiques » des mots et permettre une intelligibilité globale satisfaisante de la parole (Liénard, 1972, p.69). Par contre, la transmission uniquement des bruits et parties stables des voyelles ne permet pas l'intelligibilité de la parole. Elles sont constituées de la variation de fréquences entre l'attaque consonantique de la syllabe et l'état stationnaire vocalique. Delattre les a décrites en caractérisant pour chaque consonne le lieu fréquentiel fictif (locus) vers lequel ces mouvements formantiques se dirigent. Le locus permet de définir l'invariant acoustique correspondant aux différents lieux d'articulation. La critique contre la cartographie de Delattre s'est souvent portée à l'encontre de la théorie des loci des consonnes. Ainsi pour Liénard, par exemple, la théorie du locus est d'une systématisation excessive et sans portée pratique pour l'analyse acoustique « car l'imprécision inhérente à la mesure de fréquence des formants est telle, en parole normale, qu'il n'est guère possible de définir leur point de rencontre » (Liénard, 1972 p. 62). De l'avis général cependant, cette théorie a permis de mettre l'accent sur l'importance des transitions phonétiques. Pour Delattre, les T2 (transitions de deuxième formant) sont sans doute les plus puissants indices de distinction entre les lieux d'articulation. Excepté chez /k/ devant voyelle arrondie, ils sont plus effectifs que les explosions, ce qui est compréhensible, car ils ont, comme les formants vocaliques, une beaucoup plus grande audibilité que les bruits sourds d'explosion (Delattre, 1966, p.260). Les transitions de F2 chez les occlusives durent en moyenne 50 ms et sont plus courtes pour les labiales et plus longues pour les dentales devant les voyelles postérieures.

B.3 joints périodiques

Les joints périodiques relient par une zone formantique traversant l'espace de la consonne les formants des voyelles environnantes. Les joints périodiques bas et discontinus caractérisent les consonnes nasales. La nasalité consonantique se manifeste aussi par l'existence de formants de nasalité supérieurs autour de 1000 Hz et 2500 Hz qui occupent la tenue de la consonne. Les consonnes nasales diffèrent des occlusives par la présence d'un murmure nasal à basse fréquence pendant l'occlusion. Son intensité est un indice important de 8 à 15 dB SPL au dessous de la voyelle suivante.

B.4 voisement

Le voisement est caractérisé par un faisceau d'indices. Pour les occlusives l'indice principal est le « VOT », (voice onset time), qui est l'intervalle séparant le début de l'explosion du début de la vibration des cordes vocales (Lisker & Abramson, 1964). S'il est négatif ou nul, le voisement de l'occlusive sera perçu, sinon la consonne sera identifiée comme non voisée. En fait, la durée du VOT est soumise à d'importantes variations contextuelles. Pour Delattre, le non voisement se caractérise dans les occlusives par la coupure (cut back) du premier formant qui débute après les deux autres. La barre de voisement est présente et caractérise les autres consonnes voisées à côté d'autres indices secondaires. Pour Delattre, c'est la représentation du fondamental à laquelle s'ajoute une plus ou moins forte intensité du second harmonique selon les sujets. Les tests d'identification perceptive montrent que la durée de la tenue constitue un indice important pour la perception du trait de voisement. Cependant il ne peut à lui seul déterminer de manière non ambiguë le caractère non voisé ou voisé de la consonne occlusive (Serniclaes, 1973). Plus la durée relative de la voyelle est longue, plus la consonne est perçue comme voisée. L'aspiration se traduit de façon acoustique par l'omission de la transition de F1 et la présence d'une turbulence dans les 50 ou 60 premiers ms. Elle est un indice faible de non voisement. L'intensité relative des bruits aperiodiques est un facteur jouant sur la perception du voisement, mais de façon faible. Les consonnes voisées présentent un bruit plus faible que les consonnes non voisées.

II. LA PERCEPTION PHONÉTIQUE CATEGORIELLE

L'identification de la parole nécessite la sélection inconsciente d'unités discrètes et opposées dans un ensemble de possibles constitué par la matrice phonologique d'une langue donnée. C'est donc effectuer un processus de catégorisation assimilable à une véritable procédure de décision. Un exemple simple de perception phonétique catégorielle est illustrée par la variation continue et descendante du deuxième formant du /i/ qui est reconnue comme /i/ par les sujets testés jusqu'à une certaine valeur du deuxième formant, différente d'ailleurs selon les sujets, à partir duquel ceux-ci reconnaissent un /y/. Pour les consonnes, on sait que la variation progressive de la pente du deuxième formant d'une voyelle de synthèse fait à un moment donné passer brusquement la perception d'un /b/ à un /d/. Pour un continuum de VOT (voice onset time) allant de /ba/ à /pa/, le seuil différentiel est de 30 ms pour les anglophones et correspond à la frontière d'identification de ces consonnes en anglais (Lisker & Abramson, 1970 in Calliope, 1989). Il existe ainsi dans le continuum acoustique des zones frontières dont le franchissement fait basculer l'identification d'un phonème vers un autre phonème. Deux stimuli voisins seront discriminés de la même façon tant qu'ils seront situés du même côté d'une interface séparant deux catégories identifiables adjacentes. Le processus de discrimination est donc ordonné à l'identification, elle-même définie comme effet de catégorisation. Ce phénomène d'«effet de frontières phonétiques» a été mis en évidence autant pour la perception des voyelles que pour la perception des consonnes même s'il n'est pas sûr que pour ces catégories de phonèmes les processus en jeu soient strictement identiques. Il semble en effet que la perception des voyelles isolées soit plus continue que catégorielle, puisqu'elle concerne l'identification plus que la discrimination catégorielle (Stevens, 1969). Cependant de façon globale, la compréhension de la perception phonétique nécessite un modèle qui rend compte du passage entre le niveau physique continu et le niveau de la discrétisation phonétique. De plus, on sait depuis les travaux de Pisoni en 1977 que la perception catégorielle existe aussi pour des sons non phonétiques et produits par d'autres sources que la parole (Pisoni, 1977). La spécificité des processus de décodage de la parole par rapport aux autres processus auditivo-perceptifs est donc également problématique. De même que sa dépendance des lois générales de la perception puisqu'on retrouve cette notion de catégorie dans bien d'autres domaines que la perception de la parole. Ainsi, la capacité de la mémoire immédiate limitée à 7 plus ou moins 2 éléments, les échelles subjectives d'intensité sonore, les jugements de durée etc., tous ces phénomènes obéissent aux lois catégorielles de la perception des formes.

La tâche des modèles de la perception phonétique est de rendre compte de façon satisfaisante du passage entre le niveau physique du signal de parole et le niveau de discrétisation phonétique. Ils doivent donc proposer une explication du phénomène de catégorisation phonétique observable dans la perception de la parole et que l'on peut résumer comme suit: la modification continue d'une ou plusieurs dimensions du signal physique n'a pas de conséquence sur l'identification du phonème jusqu'à une certaine valeur critique à partir de laquelle l'identification bascule vers la reconnaissance d'un autre phonème. Pour les tenants des théories motrices de la perception de la parole, (Joos, 1948) (Liberman, 1967), les seuils critiques de l'identification phonétique sont la trace sur la perception phonétique des positions articulaires dont l'auditeur, à l'écoute d'un message phonétique émanant d'un locuteur, effectue, dans un espace proprioceptif, le geste articulaire. Ainsi pour Repp (1978), l'intégration perceptive des indices acoustiques ne peut reposer sur des principes purement auditifs et doit être expliquée par «l'intention articulaire».

Sur le plan expérimental, Lane en 1965 a développé une importante critique extensive des théories motrices en montrant qu'il était possible de faire discriminer après apprentissage des indices acoustiques de parole dans des signaux non produits par la parole (Lane, 1965). Mais les objections à ce modèle ont pour la plupart trait au fait trivial que les enfants décodent la parole avant de pouvoir articuler. Les insuffisances des théories motrices ont amené à construire des modèles «sensoriels» rendant compte de la perception phonétique sans référence à l'articulation. La perception catégorielle sera alors une propriété générale de la perception qui sera déductible des principes psychophysiques généraux. Ainsi, selon Jusczyk (1980), les universaux univers psychophysiques de chaque espèce, dus à la structure de l'appareil auditif des mammifères supérieurs, déterminent des propriétés de saillance perceptive des stimuli acoustiques et le système phonétique s'accrochera à ces saillances à travers une pondération des transformant en traits phonologiquement distinctifs. Pour Miller J.D, le processus de catégorisation phonétique est issu des effets de seuils différentiels masqués (Miller J.D, 1976). En dessous de ces seuils, les variations des paramètres du signal n'ont aucun effet jusqu'au dépassement de ce seuil où la traduction électrophysiologique diffère qualitativement. Cet argument théorique a reçu une confirmation sur le plan clinique par l'étude du rapport entre le seuil différentiel de discrimination fréquentielle et les pertes auditives. La discrimination fréquentielle dépend de la fréquence et du degré de perte sur chaque fréquence. Cette acuité fréquentielle peut être l'objet d'un apprentissage. Pour les fréquences graves, les effets d'apprentissage sont les mêmes chez les normo-entendants que chez les déficients auditifs. Les variations de seuil différentiel en fréquence induites par les lésions cochléaires responsables de la déficience auditive peuvent venir modifier les processus de catégorisation phonétique.

L'hypothèse des détecteurs de traits est née de la nécessité de concilier la distinction des langues et l'unicité de la perception de la parole. Selon ce modèle, les traits distinctifs phonologiques encodés dans les indices acoustiques seraient décodés par les récepteurs neurosensoriels (complexes de neurones) qui répondraient spécifiquement à certains paramètres physiques du signal de parole. Il y aura ainsi entre les mécanismes perceptifs visuels (hypothèse des détecteurs visuels de traits) et les mécanismes perceptifs auditifs une véritable parenté. En extrapolant ces données issues des expérimentations animales chez l'homme, on pouvait émettre l'hypothèse que les processus de catégorisation en traits distinctifs acoustiques seraient comparables aux processus de détection des sons complexes spécifiques de l'espèce et seraient donc localisés pour certains d'entre eux dès le corps genouillé médian puis ensuite dans les aires auditives primaires. Les indices acoustiques de la parole, mis en évidence par les phonéticiens, seraient dans ce modèle codés sous forme de patterns neuronaux spécifiques dès la sortie de la cochlée. Selon ce type de modèle, il existe des détecteurs de traits phonétiques spécialisés dans la reconnaissance de tel ou événement acoustique se déployant dans les limites des bassins d'attraction des différents détecteurs.

L'hypothèse des détecteurs de traits peut gagner en puissance et se dégager du postulat de l'existence d'une carte cognitive réalisant d'hypothétiques processus logiques en faisant appel aux modèles connexionnistes issus de l'Intelligence Artificielle (IA). Dans les couches profondes d'un réseau neuromimétique comportant chacune un nombre fixé de neurones logiques se définissent, au fil des présentations des stimuli pendant la phase d'apprentissage, des noyaux denses (attracteurs) correspondant aux traits structuraux des objets présentés. En phase d'exploitation, les formes présentées en entrée du réseau sont attirées vers ces attracteurs et ainsi reconnues, à la condition que les coordonnées du stimulus ne sont pas trop éloignées de celles acquises par apprentissage et qu'elles restent incluses dans le bassin d'attraction de chaque attracteur. Bien que ces modèles connexionnistes ne soient pas encore d'un usage courant en audiophonologie, ils présentent cependant un fort intérêt en particulier du fait des rapprochements suivants. Les effets de frontières catégorielles mis en évidence par les études des systèmes de confusions (Virole, 1994) paraissent isomorphes aux effets d'apprentissage de configurations neuronales de type réseaux compétitifs où les effets de renforcement et d'inhibition aboutissent à reproduire l'organisation topographique des formes d'entrée sans supervision d'apprentissage (Kohonen, 1984). Une fois le réseau de neurones stabilisé, une de ses propriétés est de posséder une résistance remarquable à toutes perturbations des formes présentées en entrée, à l'instar de la parole. Ainsi des phonèmes peuvent être identifiés malgré une altération considérable de leur indices acoustiques due au couple masquage et pertes auditives. Par contre, il existe des troubles dits «<d'intégration>>» où la dégradation de la discrimination phonétique ne peut être attribuée à la déficience auditive périphérique mais à des perturbations neuropsychologiques marquées spécifiquement par une sensibilité accrue au bruit perturbant et tout particulièrement aux sons de paroles concurrentiels (effet de «<cocktail party>>»). Enfin, les modèles utilisant les réseaux neuronaux s'avèrent très proches des modèles mathématiques morphodynamiques ainsi que l'ont montré les travaux de Jean Petitot (Cf. aussi Virole 1996).

III. REHABILITATION AUDIOPHONOLOGIQUE

Les modèles théoriques censés rendre compte de l'identification de la parole sont cependant pour la plupart issus d'expérimentations sur des sujets ne présentant pas de troubles de décodage. L'application de ces modèles en pathologie a souvent été réalisée, mais la démarche inverse, à savoir l'utilisation de données cliniques pour la construction d'un modèle a rarement retenu l'attention ou s'est heurtée à des difficultés méthodologiques. En effet, la circulation des savoirs et des données entre recherche et clinique est encore particulièrement difficile aujourd'hui dans le domaine de la parole. Ceci a pour conséquence que les modèles et les méthodologies élaborés dans le champ de la recherche expérimentale sont souvent connus avec retard par les cliniciens et qu'inversement les données et les connaissances cliniques concernant les troubles de l'identification de la parole sont souvent méconnues par les chercheurs. Les résultats issus du champ de la clinique de la parole, que ce soit en audiologie en neuropsychologie où bien encore dans les domaines de la rééducation comme en orthophonie ou en audioprothèse sont pourtant d'une grande valeur informative pour la recherche sur la perception de la parole et viennent compléter les données issues du champ expérimental.

L'intérêt principal des déficiences auditives pour la connaissance des processus de catégorisation phonétique réside ainsi dans le fait qu'elles modifient l'entrée du système auditif, et que leurs effets se traduisent jusqu'à l'identification phonétique. Elles peuvent donc être assimilées à des perturbations dont on connaît un certain nombre de paramètres (importance du déficit, forme de la courbe audiométrique, date d'apparition ...) posées devant le signal de parole et dont on recueille les modifications en sortie sous la forme de systèmes de confusions phonétiques. En faisant varier les perturbations d'entrée, donc les types de déficiences auditives, on peut observer, moyennant un contrôle rigoureux des différentes variables (type de tests, contexte, contraintes phonologiques, distributionnelles, lexicales, sémantiques ..) les variations des systèmes de confusions entre phonèmes qui sont eux-mêmes une image des modifications des réseaux de frontières catégorielles de l'espace phonétique. Ces confusions phonétiques sont d'une importance cruciale en audioprothèse car elles constituent souvent les seuls indicateurs des troubles qualitatifs de la perception de la parole et peuvent servir de marqueurs des bénéfices apportés par les prothèses auditives.

III.1 Les confusions phonétiques

Les confusions phonétiques perceptives peuvent être ainsi définies comme la substitution à un phonème proposé par un locuteur d'un autre phonème répété par un auditeur. En recherchant les éléments communs et ceux qui diffèrent entre des

deux phonèmes, on peut déterminer la nature de l'erreur et en inférer sa détermination à la condition de prendre en considération l'ensemble des facteurs influençant la perception phonétique. De nombreux facteurs doivent donc être contrôlés pour pouvoir réellement approcher les liens entre la nature quantitative et qualitative de la déficience auditive et la forme des systèmes de confusions phonétiques. On doit citer ainsi :

1. L'environnement contextuel des phonèmes erronés et substitués. Ce facteur est princeps car la valeur discriminative des indices encodés dans le signal de parole et permettant l'identification phonétique est tributaire d'autres indices des phonèmes environnants.
2. La structure phonologique du stimulus de test. On sait par exemple que l'identification des syllabes Voyelle (V) + Consonne occlusive est meilleure que l'identification C occlusive + V. De même certaines oppositions sont rares à l'initiale des mots et d'autres ne peuvent apparaître dans tous les contextes ; /k-/t/ par exemple ne peut se réaliser de façon minimale que dans un contexte de voyelle aiguë.
3. La longueur syllabique des unités de test. Plus le mot est long et contient de syllabes, plus la suppléance sémantico-lexicale peut influencer la réponse et pallier aux difficultés d'identification des phonèmes constitutifs.
4. La fréquence d'occurrence des phonèmes et des mots dans la chaîne parlée, bien que la fréquence d'occurrence du mot n'a virtuellement aucun effet sur l'intelligibilité lorsque les choix offerts à l'auditeur sont prédéfinis pour chaque stimulus (choix forcé).
5. La composante lexico-sémantique des mots. Le lexique impose des contraintes de sélection et de disponibilité. Il n'est pas rare en clinique de rencontrer des patients qui lors des tests d'audiométrie vocale répondent de façon erronée en donnant des mots appartenant à des champs sémantiques préférentiels.
6. Les informations prosodiques ou esthétiques jouant sur l'identification phonématique. Ces informations ont les plus résistances à la dégradation du signal sur le plan perceptif. Elles sont véhiculées par les variations d'intonation de la voix et sa mélodie.
7. Les conditions initiales de l'émission vocale et les déformations éventuelles du message par l'articulation du locuteur (en particulier la vitesse d'articulation). La voix normale du testeur peut en effet être plus ou moins chargée de bruit du à l'écoulement aérien au niveau de la glotte ou de clics parasites dus à la viscosité de la salive.
8. Le sexe du testeur et donc les caractéristiques de la voix.
9. La vitesse d'émission du flux de parole (débit).
10. Les contextes acoustiques de propagation du signal vocal et en particulier les déformations dues au masquage par le bruit de fond (rapport signal/bruit) ou dues aux conditions de réverbération. On sait en effet que la réverbération, en générant des traînées sonores sur certaines composantes du signal altère la perception de la parole.
11. Le traitement du signal par des systèmes d'amplification et de codage analogiques ou digitaux.
12. L'ensemble des facteurs neuropsychologiques, tels l'attention sélective et la mémoire à court terme, qui peuvent influencer, en cas de pathologie, la discrimination phonétique.
13. L'ensemble des facteurs cliniques et spécifiques, tels l'âge du sujet testé, sa conscience phonologique, les particularités phonologiques régionales, etc.

L'interprétation de vastes corpus de confusions phonétiques recueillies nécessite ainsi un traitement informatique pour pouvoir accéder à une validité statistique et une puissance de traitement importante. L'article princeps sur l'analyse des confusions phonétiques est celui de Miller et Nicely (1955). On peut consulter aussi le logiciel ANAPRO (Virole, 1994) qui permet d'apposer une grille d'analyse en traits et indices acoustiques sur les systèmes de confusions et de calculer automatiquement les taux d'erreurs sur chacune des unités d'analyse. Afin de prendre en considération, les variations contextuelles sur la valeur des traits distinctifs de chaque phonème, ce logiciel permet de modifier les valences de traits selon le contexte. Il utilise le système de traits de Jakobson. Pour les indices acoustiques, il utilise le système de Delattre qui est, de par sa présentation systématisée et binarisée, très propice à l'informatisation. Le calcul des taux d'erreurs sur chaque élément d'analyse phonétique, trait ou indice, tient compte de l'écart à la probabilité de confusion due au hasard. Cette méthodologie nous a permis d'interpréter les systèmes de confusions comme résultantes principalement d'effet de seuil critique. Nous avons essayé de montrer dans notre thèse en sciences du langage (Virole, 1994)

que vu sur plan phonétique catégorielle, le champ auditif n'est pas un espace homogène mais hétérogène à l'intérieur duquel certains points sont réguliers et d'autres singuliers au sens topologique de ces termes. Ces points singuliers constituent les interfaces de catégorisation phonétique et sont l'objet d'une discrimination maximale, au contraire des points réguliers. L'un de ces points singuliers, et sans doute le point générique de la morphologie de l'ensemble des interfaces phonétiques est le 2 kHz à 60 dB HTL. Une exploration de l'audition, normale ou pathologique, phonétiquement éclairée, consiste selon nous à raisonner en termes de franchissement de seuils de catégorisation.

III.2 Distorsions phonétiques générées par les aides auditives

Marie Thérèse Vie a réalisé en 1988, une étude sur la transmission de la parole au travers des prothèses auditives en analysant les modifications apportées par les systèmes d'amplification et de compression à des signaux artificiels sensés simuler la parole. Ces résultats expérimentaux montrent que les prothèses auditives peuvent atteindre un niveau de qualité sonore élevé, mais que simultanément elles génèrent certaines déformations du message sonore qui nuisent à l'intelligibilité de la parole. En particulier, en régime impulsif et avec la mise en oeuvre de systèmes de compression, il existe des modifications importantes de l'enveloppe temporelle surtout lorsque les impulsions d'entrée ont une durée inférieure à 20 ms. La compression de dynamique peut ainsi agir sur les sons impulsifs du message vocal et entraîner « une modification de timbre et l'apparition d'un signal bruité » (Vie, 1988). Ces résultats ont été suivis d'une autre étude à partir de véritables sons de parole. Marie-Sophie Guibert-Blanchard a en effet réalisé en 1992 une étude sur la transmission des indices acoustiques de la parole par les prothèses auditives en analysant physiquement les signaux de sortie d'une prothèse auditive recevant en entrée des sons de parole d'un unique locuteur. D'après ces conclusions (Guibert-Blanchard, 1992) :

- 1) La transmission des indices acoustiques des voyelles se fait parfaitement à travers la prothèse auditive. Les modifications spectrales correspondent à l'effet de la courbe de réponse. Elles peuvent permettre des renforcements dans les aigus.
- 2) Les indices de voisement sont transmis différemment au travers de la prothèse, selon leur nature. L'indice d'ordre spectral (F_0) est d'autant moins transmis que la courbe de réponse montre une faible amplification dans les graves. L'indice énergétique correspondant à l'intensité du F_0 est mal respecté par la prothèse. L'indice temporel (V.O.T) est bien transmis et devrait donc assurer la reconnaissance perceptive.
- 3) Le bruit d'explosion des plosives semble mal transmis par la prothèse. Compte tenu des règles perceptives de compensation, il est alors nécessaire que la transition du F_2 soit bien transmise. Dans le cas de fonctionnement linéaire, cet indice est bien transmis. Par contre, si la montée rapide du niveau de signal de la voyelle adjacente déclenche les systèmes de compression ou de limitation de niveau de sortie alors, il est possible que cet indice soit altéré physiquement et que la plosive ne soit pas discriminée perceptivement.
- 4) Les plosives sont mieux respectées par les appareillages d'amplification linéaire. De façon générale, les prothèses auditives transforment les fricatives en plosives par manque d'amplification correcte des bruits de friction aiguë. Les transitions, plus énergétiques, fournissent alors les principales informations distinctives. Ces données illustrent la nécessité du double paramétrage indiciaire (bruits et transitions) pour une identification phonétique correcte de ces consonnes.
- 5) Dans les cas de forte compression, le signal est altéré dans sa dimension temporelle, et les prothèses génèrent des confusions intermodes dont la plus fréquente est /s/ confondu en /p/.
- 6) Les confusions intermodes comme /s/ confondu en /t/ illustrent une erreur de catégorisation due au déplacement d'une interface paramétrée temporellement (durée des bruits).

III.3 Systèmes de compensation indiciaire et implants

Dans les déficiences auditives sévères et profondes de l'enfant, le champ auditif est si réduit que les informations phonétiques se trouvent limitées. Cette limitation de transmission de l'information auditive due aux lésions cochléaires a amené à essayer de coder l'information de telle façon qu'on puisse la transmettre dans les zones d'audition résiduelles (Lafon, 1961). Différents algorithmes de transposition ont été utilisés avec des succès globaux divers, mais tous sont particulièrement intéressants car ils tentent d'imposer une nouvelle cartographie de frontières phonétiques dans un espace auditif plus restreint. Un certain nombre de ces systèmes existent dans les centres de rééducation pour enfants sourds. Il convient de citer les amplificateurs à sons de compensation élaborés par le Pr. Lafon qui appliquent sur le plan rééducatif les princi-

pes de l'audition impulsionnelle. Des sons artificiels, modulés par l'amplitude du signal dans les fréquences aiguës de la parole sont transposés dans la zone d'audition résiduelle des sourds profonds. Le système EMILY (Cf. Lefèvre, 1984, 1985) fonctionne sur un principe différent, mais toujours basé sur la même idée force d'une translation d'éléments informatifs dans la zone d'audition résiduelle (méthode de la confrontation indiciaire de F. Lefèvre). Il s'agit dans ce cas d'un boîtier portable connecté à des prothèses portables classiques.

La limitation des résultats des systèmes spéciaux à translation de fréquence a amené l'audiophonologie à tourner ses espoirs vers les implants cochléaires. L'application aux enfants sourds congénitaux pose un ensemble de questions d'ordre technique et éthique très complexes.. Cependant sur le plan de la perception phonétique le développement actuel des implantations cochléaires ouvre incontestablement une nouvelle voie dans la compréhension des mécanismes de la perception de la parole. Ce champ clinique présente la singularité de court-circuiter véritablement l'étape du traitement cochléaire naturel de la parole en tentant de faire parvenir à une cochlée lésée des stimulations électriques intégrables plus ou moins directement au niveau neuronal. Issues du traitement du signal physique de parole par microprocesseur, ou par circuits analogiques spéciaux, ces stimulations électriques encodent, selon une forme plus ou moins prédéterminée par les réglages, certains indices acoustiques, et les transforment sous la forme d'indices neuronaux permettant la catégorisation et la reconnaissance phonétique. Le succès, ou l'échec, de ces stratégies artificielles fournissent des informations sur leur validité et leur congruence avec les processus naturels centraux de la perception phonétique. La comparaison de ces résultats obtenus avec ceux issus des réhabilitations prothétiques classiques des déficiences auditives moins importantes est également à même de fournir de façon différentielle des indications sur le traitement de la parole par la cochlée. Un certain nombre de résultats (pour les sujets adultes post linguaux) nous paraissent dignes d'attention sur le plan des théories de la perception de la parole. En particulier, il apparaît que pour certains patients, l'apport de l'implant aboutit au début à une chute brutale de la discrimination avec lecture labiale comme si l'implant perturbait la discrimination usuelle. Pour d'autres patients, il existe une amélioration sur certains traits, mais d'autres sont au contraire perturbés. Il sera intéressant d'étudier de façon systématique ce phénomène, car il peut constituer une voie d'accès aux processus de synthèse entre les informations visuelles et celles issues de l'implant et véhiculées par la voie auditive. Il devra être utile par exemple de comparer ces processus d'interférence ou de fusion existant chez les implantés avec ceux présents chez des déficients auditifs appareillés par des prothèses traditionnelles. Cette comparaison implique au sein des équipes et réseaux d'implantation la collaboration avec les audioprothésistes ayant l'expérience de l'analyse qualitative des erreurs phonétiques et de leur application dans les choix techniques de l'appareillage.

III.4 L'audition binaurale et l'appareillage stéréophonique

La spécialisation interauriculaire pour le décodage de la parole est maintenant bien connue d'après les travaux de Kimura (1961). Chez le sujet "droitier", l'oreille gauche est dominante pour les indices phonétiques de durée longue, les voyelles, alors que les indices phonétiques brefs tels les consonnes sont décodés préférentiellement par l'oreille droite. Une amélioration de la localisation des sources sonores grâce aux différentiels énergétiques des fronts d'onde acoustique parvenant aux oreilles et aux effets de masque de la tête. Cette amélioration contribue au sentiment subjectif d'avoir une audition plus naturelle. Une amélioration de l'intelligibilité de la parole dans le bruit grâce à la possibilité de faire émerger le signal sur le bruit par des processus d'écoute focalisée rendue possible à nouveau par le différentiel des deux oreilles. A ces deux bénéfices s'ajoute celui moins connu de la fusion binaurale pour l'intégration des traits phonétiques.

Ces données ont été mises en rapport par Kimura en 1967 avec le plus fort contingent de fibres nerveuses pour l'oreille controlatérale et avec le fait que ces voies croisées inhibent les efférences homolatérales. De plus, sur le plan neuropsychologique, il est possible que les informations phonétiques arrivant à l'hémisphère droit ne soient traitées qu'après la transmission à l'hémisphère gauche par la voie calleuse et donc après un certain délai. A ces raisons d'ordre neuroanatomique, se sont rajoutées des interprétations d'ordre neurolinguistique ; la dominance de l'oreille droite est en rapport avec la complexité de l'encodage au niveau cortical (Cf. Cutting 1974). Cette complexité de l'encodage est liée à l'intrication

temporelle de plusieurs indices acoustiques nécessaires à la catégorisation du phonème. La dominance de l'oreille droite est donc:

- forte pour les occlusives
- faible pour les fricatives et les semi-voyelles
- nulle pour les voyelles (fortement énergétique)

pour les liquides (L et R), il n'y a dominance que si elles sont en position initiale dans la syllabe. Ce qui explique pourquoi sur le plan phonologique, les liquides sont proches des occlusives en position initiale et proche des voyelles en position finale (Cf. Petitot, 1984).

On a pu ainsi mettre en évidence (Cutting) que du fait de la spécialisation interauriculaire due aux différences des voies auditives et des modes de traitement, une des deux oreilles, en l'occurrence la "gauche" chez le droitier encodait préférentiellement (c.a.d plus rapidement) les transitions de formants, variations des mouvements formantiques des voyelles que les bruits consonantiques, événements très courts et faiblement énergétique. Dès lors, il était possible de construire un test (PHONO-SCAN) qui mesure cet effet en contrôlant la composition indiciaire des stimuli phonétiques et en les proposant sur l'oreille droite avec prothèse à 60 dB SPL input, puis sur l'oreille gauche toujours avec prothèse à 60 dB input, et enfin en binaural avec prothèses cette fois à 55 dB afin de corriger le gain de sonie et de mieux mettre en évidence le gain de discrimination de l'audition binaurale. L'ensemble de ces données invitent à valider à nouveau, mais cette fois sur le plan neuropsychologique, la conclusion bien connue des audioprothésistes (Cf. les travaux de Monsieur Dehaussy) : l'appareillage stéréophonique est une nécessité fonctionnelle. Ne pas appareiller les deux oreilles pour des raisons d'ordre économique ou esthétique est un non-sens pour la réhabilitation de l'audition. Ceci a été souvent souligné en ce qui concerne la localisation des sources sonores ou l'intelligibilité de la parole dans le bruit. Il convient aujourd'hui de rajouter à ces raisons le fait que la perception phonétique catégorielle nécessite l'apport qualitatif des informations venant des deux oreilles. Celles-ci présentent en effet des caractéristiques fonctionnelles différentes et seule la fusion de leurs informations permet à la fonction de perception de la parole de se réaliser pleinement.

III.5 La lecture labiale

La compréhension de la parole n'est pas uniquement déterminée par la réception d'indices acoustiques présents dans le signal de parole. Il existe une perception visuelle de la parole par la lecture labiale qui est particulièrement mise en évidence par les troubles auditifs. L'existence de cette lecture labiale et donc d'indices visuels de la parole pose le problème de la fusion de ces indices avec les indices acoustiques pour la catégorisation phonétique. L'importance de la lecture labiale pour la catégorisation phonétique peut ainsi être considérée comme un argument majeur en faveur des théories motrices dans la mesure où on peut considérer que le regard porté sur le mouvement des lèvres active un écho proprioceptif articulaire (McGurk, 1976) qui joue, malgré son inhibition motrice, un rôle dans la catégorisation phonétique. Mais on sait aussi que des informations distinctives indépendantes de l'articulation peuvent tout aussi bien aider les enfants sourds à pallier aux insuffisances des informations auditives. De même, les aides manuelles à la lecture labiale de type LPC (Langage Parlé complété) : système de clefs gestuels pour lever les ambiguïtés de la lecture labiale (Cf. La thèse de Charlier B.L., 1994) ne sont en rien consubstantielles de l'articulation. Elles améliorent cependant de façon efficace la perception phonétique chez les enfants ayant déjà des informations acoustiques conséquentes, mais ne sont que de moindre effet sur la production articulaire. De nombreuses études, et l'observation clinique courante montrent que la lecture labiale est optimale lorsque le labiolecteur voit le visage du locuteur avec un angle de 45 degrés. Il est ensuite capable de retenir les éléments visuels informatifs de façon séquentielle en mémoire à court terme avant la combinaison intégrative avec les éléments issus du canal auditif. Il est connu que tout un chacun utilise de façon inconsciente la lecture labiale qui favorise l'intelligibilité, en particulier quand les conditions de transmission du message acoustique sont troublées. La perception auditivo-visuelle donne des résultats meilleurs que la somme théorique des scores théoriques auditifs et visuels et cette interaction est d'autant plus que les conditions d'écoute sont plus défavorables (Siegenthler, 1969).

L'amélioration de la synthèse auditivo-visuelle est ainsi proportionnelle au rapport signal sur bruit (Neely, 1956). La lecture labiale seule, sans information auditive, est par contre difficile et ne permet qu'une compréhension limitée du message (Guilliams, 1987). Chez les personnes adultes atteintes de surdité post linguale très profonde voire totale et qui se font appareillées ou implantées, la réhabilitation est largement tributaire de la qualité de la lecture labiale qui dans certains cas peut s'avérer remarquablement informative. Dans pratiquement tous ces cas, c'est bien la lecture labiale le vecteur principal d'informations permettant la catégorisation phonétique. On a pu montrer (Walden, 1975) sur une population d'individus sourds que la contribution des codes visuels dans l'identification auditivo-visuelle pour l'identification des consonnes était constante (70 % par rapport à 30 % pour la modalité auditive seule). Cependant d'autres études montrent une assez grande variabilité dans les compétences en lecture labiale des sujets sourds, variabilité dépendante de facteurs cliniques généraux et neuropsychologiques en particulier (Cf. Peslon & Prather, 1974).

Théoriquement on est en droit de supposer que les informations auditives devraient venir compenser les manques de la lecture labiale, d'autant plus que les traits non discriminés par la lecture labiale correspondent grossièrement aux fréquences graves de l'audition résiduelle. En comparant les scores en discrimination avec et sans lecture labiale mais avec l'apport de la modalité auditive, on peut ainsi approcher l'apport différentielle de la lecture labiale sur chacun des traits. Pour les consonnes, les gains maxima concernent le trait distinctif aigu/grave ainsi que la polarité continu du trait de

temporalité opposant les occlusives aux fricatives. La lecture labiale améliore donc sensiblement plus la discrimination des fricatives aiguës que des occlusives aiguës. Ces résultats sont en accord avec ceux de la lecture labiale seule. Par contre la nasalité consonantique est parfois perturbée par la fusion lecture labiale et audition. Pour les voyelles, la bémolisation (labialisation sur le plan articulatoire) est le trait dont la discrimination est la plus améliorée par la lecture labiale. Les voyelles les mieux aidées par la lecture labiale sont les voyelles aiguës, diffuses et bémolisées, soit les voyelles de la série antérieure du triangle des voyelles. Les voyelles postérieures ne sont pas discriminées par la lecture labiale seule (graves et compactes) de même que les nasales. L'audition fusionnée avec la lecture labiale ne fournit que peu d'éléments informatifs sur la nasalité consonantique et vocalique. Pour les consonnes, l'audition fournit les indices nécessaires à la catégorisation du trait de tonalité. Pour les voyelles, le gain est plus marqué sur le trait de compacité (globalement corrélé à un paramètre énergétique) qu'au trait de tonalité.

CONCLUSION.

En conclusion, nous aimerions insister sur le fait qu'il existe manifestement une non linéarité entre les pertes audiométriques tonales et les potentialités de catégorisation phonétique. Ainsi, il ne sert souvent à rien, sur le plan prothétique, d'essayer d'amplifier à tous prix une zone fréquentielle si cette amplification ne peut permettre de franchir un seuil de catégorisation. Par contre 5 dB de plus sur cette zone peuvent permettre parfois de franchir un seuil de catégorisation et entraîner ainsi une amélioration franche de la discrimination du phonème et par extension une amélioration de l'intelligibilité. Cependant au-delà de certaines valeurs de seuils, la récupération auditive des potentialités de catégorisation phonétique est, à notre avis, illusoire malgré les systèmes d'amplification, y compris ceux contenant des systèmes de codage par translation de fréquences. Mais à l'intérieur de ces valeurs de seuils (délimitant en fait le champ auditif des déficiences moyennes, sévères et profonde 1er groupe selon la classification du BIAP), il s'avère que l'on peut encore mettre en évidence un certain nombre de seuils critiques dont le dépassement permet la bonne catégorisation de tel ou tel trait. Ceci peut s'expliquer par l'hypothèse que, malgré un fort filtrage en entrée due à la perte audiométrique, des structures spécifiques du décodage acoustico-phonétique peuvent compenser ce filtrage en utilisant des indices de parole très altérés. L'expérience clinique des surdités et en particulier des implants cochléaire, montre que ces détecteurs peuvent être activés et déclencher des processus de catégorisation en recevant en entrée des indices artificiels dont le rapport avec les indices acoustiques naturels de la parole est assez éloigné.

On ne peut donc les considérer comme des structures spécialisées spécifiques de formes précises du signal sonore, mais comme des structures plus souples et fortement adaptatives. Ce critère d'adaptation permet de concevoir que les détecteurs de traits phonétiques seraient assimilables à des sortes d'attracteurs fonctionnant en sortie en tout ou rien (décharge ou non de l'assemblée neuronale) et recevant en entrée les indices neuronaux de la parole encodés aux différentes étapes des voies auditives ascendantes. Chacun de ces détecteurs projetterait ainsi sur l'espace du substrat auditif une cartographie de frontières correspondant à son <<bassin d'attraction>> si l'on continue à utiliser cette métaphore inspirée de la dynamique qualitative. Si un indice est <<capté>> par le bassin d'attraction, il permet l'activation du détecteur. L'étendue du bassin permet à des indices altérés d'activer néanmoins le détecteur à la condition que cette altération reste dans certaines limites. On voit ici l'importance pour l'audioprothèse d'une meilleure connaissance de ces strates de catégorisation phonétique. Enfin, l'arrivée récente des aides auditives numériques sur le marché de l'audioprothèse offre de nouvelles perspectives intéressantes pour l'utilisation pratique de ces analyses phonétiques.

Références

- Appaix A., Decroix G., Olivier J.C La prothèse auditive, Arnette, 1974.
- Botte M.C., Canévet G., Demany L., Sorin C. (1988) Psychoacoustique et Perception auditive, Paris, INSERM, Editions Médicales Internationales, 1988.
- Beraha J.P. (1978) "Recherche sur la pertinence des indices acoustiques de la parole chez le malentendant appareillé" Bulletin d'Audiophonologie, Université de Besançon, 5, 1978.
- Caelen J. (1974) Un modèle mathématique de la cochlée et son application : l'analyse du signal vocal , Thèse, Université Paul Sabatier, 1974.
- Charlier B.L. (1994) Le développement des représentations phonologiques chez l'enfant sourd : étude comparative du Langage Parlé Complété avec d'autres outils de communication. Bruxelles : Université Libre de Bruxelles (thèse de doctorat, inédite), 1994.
- Cutting J.E. Two left-hemisphere mechanisms in speech perception Perception and Psychophysics, 16, 3, 1974, 601-612.
- Delattre P.C. (1958) Les indices acoustiques de la parole, *Phonetica*, 1,2, 1958.
- Delgutte B. (1984) Codage de la parole dans le nerf auditif , Thèse de Doctorat d'Etat, Université Paris VI, 1984.
- Dorman M.F, Lindholm J.M, Hannley M.T, LeeK M.R (1987) <<Vowel Intelligibility in the Absence of the Acoustic Reflex : Performance-Intensity Characteristics>> , J. Acoust. Soc. Am. 81 (2), 1987, pp. 562-564.
- Dupret J.P. (1980) Tests de mots sans signification. Mémoire du collège national d'Audioprothèse.
- Eddington D.K (1980) <<Speech Discrimination in Deaf Subjects with Cochlear Implants>> J. Acoust. Soc. Am. 22, 1980, pp. 740-753.

- Fontanez F. (1983) "Modélisation du champ auditif en termes perceptif par ordinateur. Ann. Franç. Des microtechniques et de chronométrie, tome 37, 1.
- Fraysse B. (ed) (1989) *Cochlear Implant, Acquisitions and Controversies*, Toulouse, B. Fraysse Ed. 1989.
- Guberina P. (1956) - L'audiométrie verbo-tonale, *Revue d'ORL*, 1-2, pp ; 20-58.
- Guibert-Blanchard M.S. - Transmission des indices acoustiques de la parole par la prothèse auditive : approche d'une méthode d'essais techniques. Thèse 1992, Université de Montpellier I, Faculté de Pharmacie.
- Guilliams I. (1987) <<ALLAO, un système d'apprentissage et d'évaluation de la lecture labiale à partir d'un vidéodisc interactif>>, *Bulletin d'Audiophonologie*, Vol 3, 5, 1987, pp. 89-614.
- Joos M. (1948) *Acoustic phonetics*. *Language* 24, (suppl. 2).
- Juscik P.N (1980) - Is there a phonetic basis for speech perception during infancy ? Perspectives on mental representation (M/ Garret, J. Mehler, E. Walker). L. Erlbaum ass. Hillsdale, New Jersey.
- Jakobson R. (1963) *Essais de linguistique générale* - Editions de Minuit, 1963.
- Kimura D. Cerebral dominance and the perception of verbal stimuli *Canad. J. Psychol.* 15, 166-171. 1961.
- Kimura D. Fonctionnal asymmetry of the brain in dichotic listening *Cortex*, 3, 163-168, 1967.
- Kohonen T., et al. (1984) <<Phonotopic Mapsinsightful Representation of Phonological Features for Speech Recognition>>, in *Proc. of the Seventh Int. Conf. on Pattern Recognition*, Montreal, Canada, July 1984, pp. 182-185.
- Kiang N.Y.S, Eddington D.K, Delgutte B. (1979) <<Fundamental Considerations in Designing Auditory Implants>> *Acta Otolaryngol* 87, 1979, pp. 204-218.
- Lafon J.C. (1961) <<Principe de la transmission de la parole dans une zone fréquentielle limitée et son application à l'appareillage du sourd>>, *C.R Acad. Sc.*, 1961, pp. 252-327.
- Lafon J.C. (1964) *Le test phonétique et la mesure de l'audition*, Eindhoven, Ed. Centrex, 1964.
- Lane H. (1965) <<The Motor Theory of Speech Perception : A Critical Review>>, *Psychol. Rev.* 72 : pp. 275-309. 1965
- Lane H. (1970) *Nouvelles perspectives en phonétique*, Presses Universitaires de Bruxelles, 1970.
- Liberman A.M. (1974) <<Perception of the Speech Code>>, *Psychological Review*, 74, 6.
- Lefèvre F. (1985) *Une méthode d'analyse auditive des confusions phonétiques : la confrontation indiciaire*. Doctorat d'Université. Université de Franche-Comté.
- Lienard, J.S. (1972) *Analyse, synthèse et reconnaissance de la parole*, Thèse de Doctorat d'Etat es Sciences Appliquées, Université Paris VII 1972.
- Mc Gurck H., MacDonad J. (1976) <<Hearing Lips and Seeing Voices>>, *Nature*, 264, 1976, pp. 746-748.
- Marslen-Wilson W.D., Welsh. A. (1982) << Processing Interactions and Lexical Access During Word Recognition in Continuous Speech>>, *Cognitive Psychology* 10, 1982, pp. 29-63.
- Miller A. & Nicely P. E. (1955) <<Analyse de confusions perceptives entre consonnes anglaises>> *J. Acous. Soc. Am.* 27, 2, (trad Française, Mouton, 1974 in Melher & Noizet, textes pour une psycholinguistique).
- Moles A., Vallancien B. (1966) *Phonétique et Phonation*, Masson, Paris, 1966.
- Montandon P., Pellizzone M., Kasper A. (1988) <<Le point sur les implants cochléaires>>, *Journal français d'Oto-rhino-laryngologie* 37, 6 1988, pp.278-282.
- Neely K.K. (1956) << Effect of Visual Factors on the Intelligibility of Speech>>, *J. Acoust. Soc. Am.*, 28, 1956, pp. 1275-1277.
- Pastore R.E. (1977) <<Common Factor Model of Categorical Perception>>, *J. Exp. Psych.* 3,4, 1977.
- Petitot J. (1984) *Les catastrophes de la parole*, Paris, Maloine, Collection Interdisciplinaire, 1984.
- Pickett J.M. (1971) *One Pattern of Speech Feature Reception by Deaf Listeners*, *International Symposium on Speech Communication Ability and Profound Deafness* (G. Fant ed.) Alexander G. Bell Association for the deaf, Washington D.C, 1971, pp. 119-133
- Pimonow L. (1969) *Introduction à la bioinformatique*, Thèse, Ecole Pratique des Hautes Etudes, Paris, 1969.
- Pisoni D.B. (1977) <<Identification and Discrimination of the Relative Onset of two Components tones : Implications for the Perception of Voicing in Stops>>, *J. Acoust. Soc. Am.* 61, 1977, pp. 1352-1361.
- Pisoni D.B. (1979) <<On the Perception of Speech Sounds as Biologically Signifiant Signals>>, *Brain Behav. Evol.*, 16, 1979, pp.330-350.
- Rossi M., Peckels J.P (1975) <<Le test de diagnostic par paires minimales>>, *Rev. d'acoustique*, 27, pp. 245-262, 1975.
- Serniclaes W. (1973) *La simultanéité des indices dans la perception du voisement des occlusives*, *Rapport d'activité de l'institut de phonétique*, R.A 7/2, Université libre de Bruxelles, 1973.
- Siegenhaler B.M., Gruber V. (1969) <<Combining Vision and Audition for Speech Reception>>, *Journal of Speech and Hearing disorders*, 34, 1, 1969
- Tyler R. & coll (1982) <<Natural Vowel Perception by Patients with the Inneraid Cochlear Implant>>, *Audiology*, 31, 1992, pp. 228-239.
- Vie M.T. (1988) *Contribution à l'étude de la transmission de la parole par les prothèses auditives*. Doctorat de Sciences Pharmaceutiques de l'Université de Montpellier I.

Virole B. (1991) << Données cliniques pour un modèle de la perception phonétique au travers des implants cochléaires >> , Les cahiers de l'audition Vol 5, 3, 1991, pp. 12-19.

Virole B. (1992) << Modèles connexionnistes et troubles phonétiques >>, ANAE 3, vol. 4, 1992, pp. 124-130.

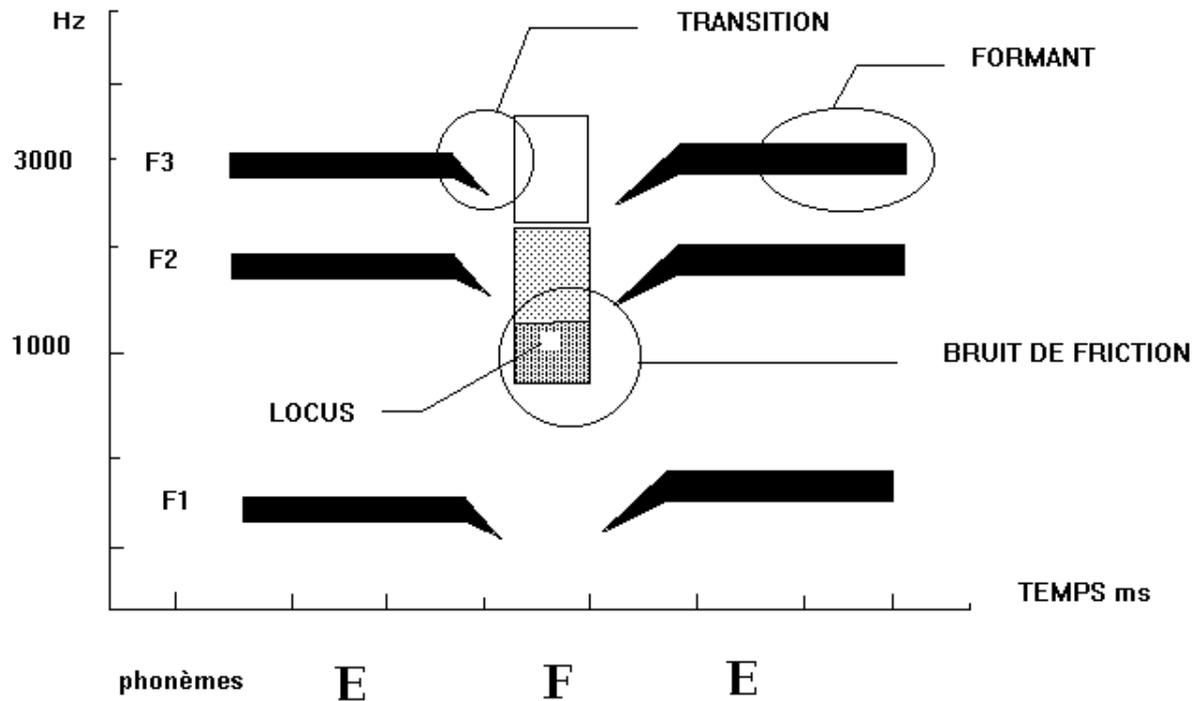
Virole B. et al. (1996) *Psychologie de la surdit * - DeBoeck Universit , 1996

Walden B.E, Prosek R.A, Worthington O.W. (1975) << Auditory and Audiovisual Feature Transmission in Hearing Impaired Adults >> , Journal of Speech and Hearing Research, 18 (2), 1975, pp. 272-279.

Le cheval court dans le pr�	Graph�mes
L ɸ θ ɸ v a L k u R d ˜ a L ɸ p R e	Phon�mes

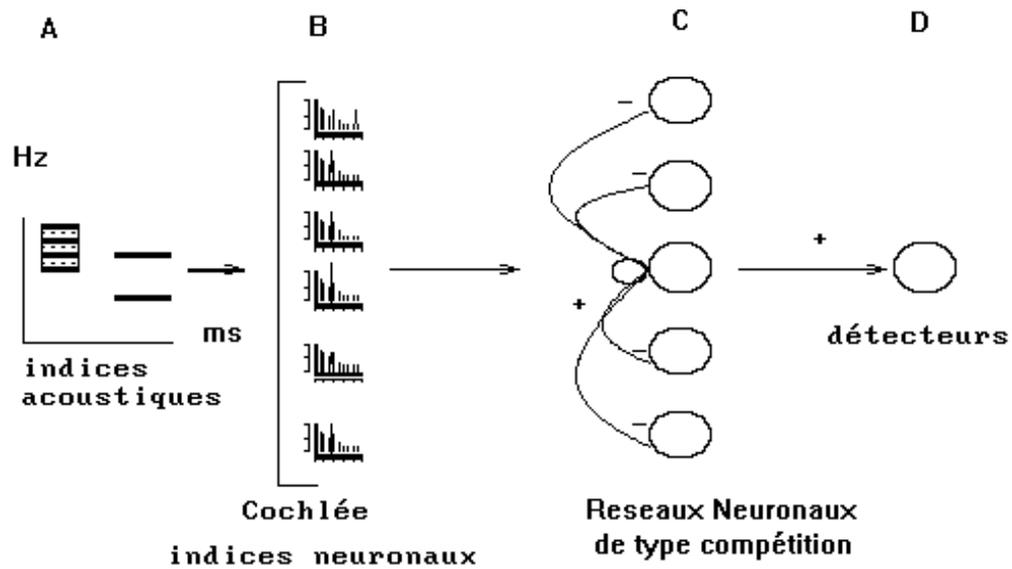
Figure 1

Exemple de conversion d'écriture entre une phrase comportant des graph mes (lettres)
et sa transcription en code phon tique.



Représentation d'un sonagramme stylisé représentant les principaux indices acoustiques d'un triplet phonématique /E/F/E/ et faisant abstraction des éléments acoustiques non pertinents. Au centre, le bruit de friction du /F/ avec ses plages de plus ou moins grande intensité. Le locus correspond au point (virtuel) de l'articulation. Les transitions de formant présentent un certain angle de pente et sont ainsi saillantes sur le plan perceptif.

Figure 2



Représentation des différentes étapes de la perception phonétique catégorielle.

En A - B, la cochlée traduit les indices acoustiques encodés dans le signal de parole dans des indices neuronaux qui sont ensuite (C -D) traités par des réseaux neuronaux avant d'activer les détecteurs de traits phonétiques (modèle théorique des détecteurs de traits).

Figure 3

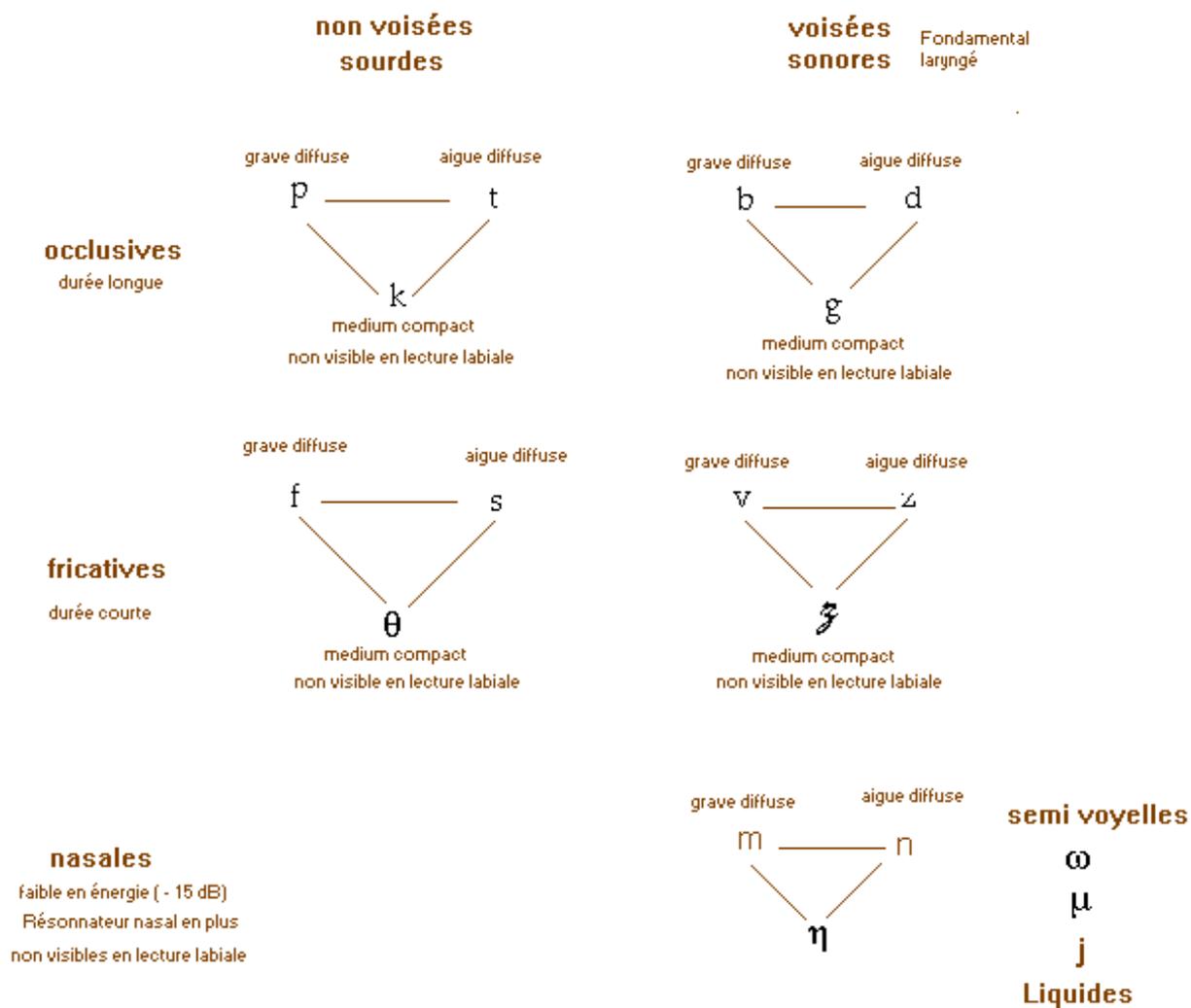


Tableau 1

Agencement systématique des consonnes du français

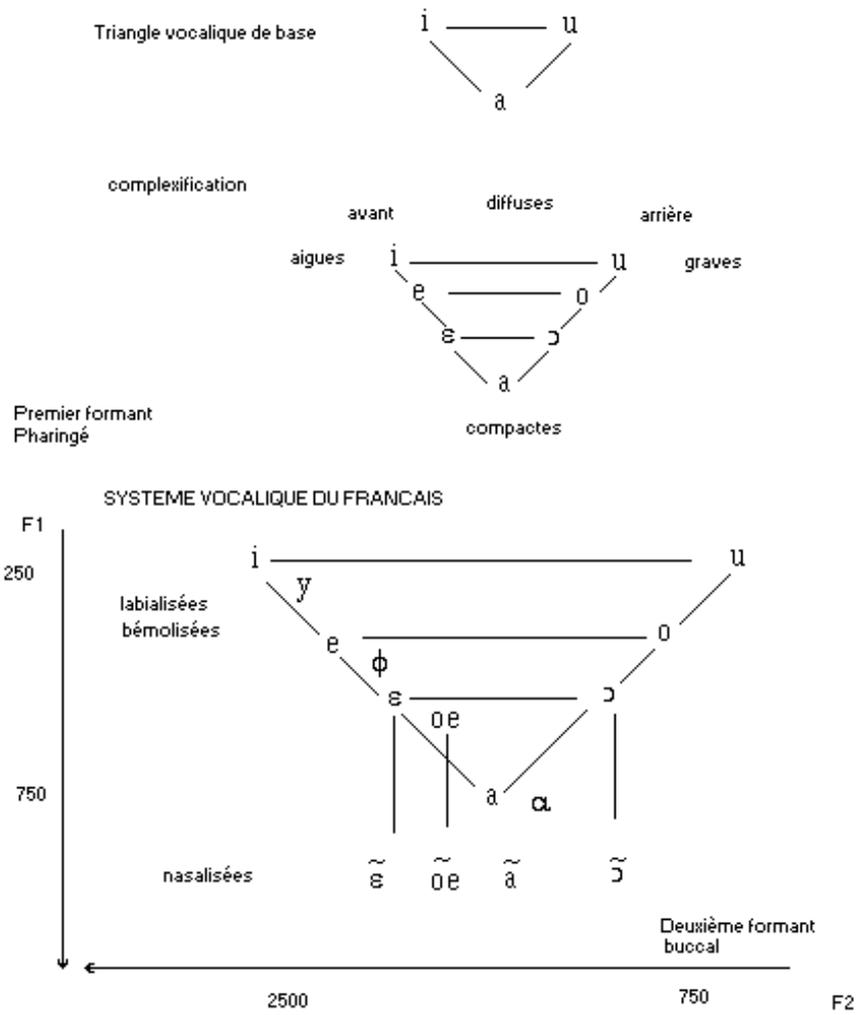


Tableau 2
Système vocalique