

EN QUOI LA PSYCHO-ACOUSTIQUE PEUT-ELLE AIDER A APPROFONDIR LA CONNAISSANCE DE LA COMMUNICATION LINGUISTIQUE ?

par Etienne EMERIT

La communication linguistique.

Dire que le prodigieux développement de l'intelligence humaine n'a été rendu possible que par la communication linguistique sous une forme ou sous une autre, est un lieu commun auquel, pour une fois, nous n'échapperons pas. Le cas parfaitement connu des «enfants-loups» découverts au début du siècle aux Indes, et recueillis par un missionnaire, démontrait cette vérité première : des enfants élevés par des animaux, et par là-même privés de l'acquisition du langage dans leurs premières années, restent privés, par la suite, d'une structure cérébrale à base linguistique qui est la condition nécessaire de toute pensée humaine. Ces enfants qui n'ont jamais pu apprendre à parler, et demeuraient incapables de raisonner, illustraient tragiquement la différence qui sépare l'homme de l'animal : la pensée humaine, elle-même produit du langage qui, à son tour, ne peut se concevoir hors du contexte de la communication linguistique.

Si on prend le mot « linguistique » dans son sens le plus général, on peut dire que c'est la science de la communication humaine par le moyen du langage. Cette communication elle-même implique une transmission de pensée d'un individu à l'autre par le truchement de codages et de décodages successifs (parole, écriture). Si on s'en tient à la parole, et si l'on décompose toutes les opérations qui lient le locuteur à l'auditeur, il est d'usage de distinguer plusieurs échelons distincts : pensée — langage pensé — langage articulé — production et transmission des sons du langage — décodage complexe par l'oreille et les centres subcorticaux — puis de nouveau transformation de ce message nerveux en langage pensé et en pensée proprement dite qui lui est étroitement liée.

Entre cette pensée «émettrice» et cette pensée «réceptrice», les sujets d'étude ne manquent pas. Il serait hors de propos de les aborder tous, et nous ne retiendrons, à des degrés divers, que les préoccupations du psycholinguiste, du linguiste, du phonologue, et du phonéticien.

Si le linguiste accorde une place, encore que modeste, au rôle des sons dans cette chaîne de transformations, le psycholinguiste, lui, ne retient que les procédés de codage et de décodage sous leur forme la plus abstraite, c'est à dire au niveau des symboles. OSGOOD et SEBEOK définissent ainsi ses préoccupations (1967, 1-2, p. 4) : «La discipline assez nouvelle qui doit être connue sous le nom de «psycholinguistique» a trait, dans le sens le plus général, aux rapports entre les messages et les caractéristiques des individus qui les sélectionnent et les interprètent. Dans un sens plus restreint, la psycholinguistique étudie les processus par lesquels les intentions du locuteur sont transformées en signaux dans le code accepté par une culture, et par lesquels ces signaux sont transformés dans l'interprétation de l'auditeur».

Plus loin, ils définissent les préoccupations du linguiste (Ibid., 2.2.1, p. 8) : «La première préoccupation du linguiste est le langage parlé. L'écriture et les autres systèmes partiellement ou totalement isomorphiques de la parole sont considérés par la plupart des linguistes comme des systèmes secondaires».

En fait, si la psycholinguistique est un lien entre la linguistique et la psychologie, la linguistique proprement dite, étude du langage, doit nécessairement s'appuyer sur les éléments précis malgré leur caractère d'abstraction, que sont les phonèmes tels que les conçoit le phonologue, c'est à dire étudiés non sous leur aspect purement physique ou psycho-acoustique-caractère considéré inintéressant dans le cadre d'une même langue — mais sous leur aspect purement fonctionnel. Cependant, même considéré sous cet aspect abstrait car purement fonctionnel, le phonème « phonologique » ne peut être une pure et simple abstraction et mérite, même en phonologie, une description au moins articulatoire.

Le linguiste désirant en savoir davantage sur les rapports opposant le système phonique d'une langue à celui d'une autre langue, ou d'un dialecte, devra s'écarter de la simplification phonologique, et se pencher sur une étude physique des sons appartenant à une même classification phonémique, mais différant acoustiquement d'une langue à l'autre. Et c'est la phonétique et ses modernes moyens d'analyse qui lui viendront en aide.

A son tour, cette phonétique, étude matérielle des sons du langage, ne saurait se contenter de mesures purement physiques, et devra faire appel, dans la mesure du possible, à la psycho-acoustique, dont les faits sont difficilement mesurables, certes, mais très sûrs, car nous savons parfaitement reconnaître entre les sons des différences que les appareils les plus perfectionnés ne permettent pas de délimiter avec exactitude.

Ces aspects si différents des moyens mis en œuvre pour étudier un seul et même sujet d'ensemble qui est le langage obéissent à un besoin toujours accru de diversification, de spécialisation. Un domaine qu'un seul spécialiste ne peut défricher doit et peut l'être par plusieurs à condition que, jusqu'à un certain point, *il parlent le même langage* : un linguiste doit être aussi phonéticien, un phonéticien doit être aussi linguiste, et les uns et les autres doivent garder présent à l'esprit qu'ils sont tous engagés plus ou moins profondément sur un seul et même territoire : celui de la *linguistique*.

La phonétique, qui s'est érigée plus ou moins science autonome, doit plus que jamais se rapprocher de ce bercail. GRAMMONT (1965, p. 15) mettait déjà au début du siècle l'accent sur cette nécessité, en ces termes : «La phonétique instrumentale n'est qu'une méthode auxiliaire de la phonétique, qui n'est elle-même qu'une partie de la linguistique Il n'y a pas de science d'observation où il suffise de faire fonctionner un appareil pour obtenir des découvertes Dans la plupart des cas, il faut être linguiste pour concevoir le problème et le poser correctement pour tirer les conséquences des observations que l'on est à même de faire et ne pas prendre pour une loi générale ce qui n'est souvent qu'un accident d'expérience».

Cette opinion est de nouveau exprimée par MALMBERG (1967, p. 89) qui dit que «les sciences phonétiques sont les sciences du langage», tout en considérant que la phonétique ne justifie plus le statut de science à part, mais devrait être considérée comme un «sous-département» de la linguistique.

Une première opposition entre la linguistique et la phonétique est que la linguistique est l'étude du langage pris dans un sens général, au niveau même de la pensée, alors que la phonétique est l'étude des sons du langage parlé (et la phonologie le lien fonctionnel unissant ces deux conceptions). Cependant le langage pensé n'est qu'une image cérébrale du langage parlé, car il est difficile, voire impossible, d'avoir une pensée tant soit peu abstraite sans «entendre» mentalement les signifiants qui la forment, d'où l'inévitable interférence entre linguistique et phonétique. La multiplicité et la diversité des sciences d'étude du langage ont eu pour mérite de cristalliser, de concrétiser, de codifier certaines notions qui se dégagent de ce processus hautement cérébral tout en étant presque parfaitement inconscient.

LOTZ (1956, p. 207) met l'accent sur le caractère subconscient, et par là-même difficilement observable, du langage : «Le langage est comme un iceberg. Une partie, la production de la parole dans le conduit vocal, les mouvements qui l'accompagnent, le passage du son à travers l'air et son impact sur l'oreille, tout cela est ouvert à l'observation immédiate. Mais la partie incomparablement plus grande, la formation de l'énoncé dans le cerveau du locuteur, sa réception par l'auditeur et l'association du signal à l'expérience — passée et présente, isolée individuellement et socialement partagée, se trouve en dessous de la surface et ne peut être mesurée que par des sondages ».

SAPIR dit à ce sujet (1967, p. 7) : «La parole est un trait si familier de la vie quotidienne que nous prenons rarement le temps de la définir. Elle semble aussi naturelle à l'homme que la marche, et à peine moins normale que la respiration ».

D'après BENVENISTE (1966, p. 63), «la réalité de langue demeure en règle générale inconsciente : hormis le cas d'étude proprement linguistique, nous n'avons au plus qu'une conscience faible et fugitive des opérations que nous accomplissons pour parler De là procède cette conviction largement répandue et elle-même inconsciente comme tout ce qui touche au langage, que penser et parler sont deux activités distinctes par essence, qui se conjoignent pour la nécessité pratique de la communication, mais qui ont chacune leur domaine et leurs possibilités indépendantes, celles de la langue consistant dans les ressources offertes à l'esprit pour ce qu'on appelle l'expression de la pensée».

La langue est une structure complexe qui donne sa forme à la pensée, et cette même structure doit exister tant chez le locuteur que chez l'auditeur si l'on veut que cette pensée soit transmissible de l'un à l'autre. Et si une pensée ne peut se transmettre convenablement que par le moyen de la forme linguistique, cette forme linguistique est elle-même indispensable à la réalisation de la pensée.

Avant d'aborder la question des liens qui unissent la pensée au langage, il convient toutefois de se demander si une certaine forme de pensée peut exister en dehors même du langage, ou une certaine réaction au langage sans qu'il y ait eu véritablement pensée. FRANÇOIS (1968, p. 13) exprime son opinion en ces termes : «Il est manifeste qu'il existe une pensée sans langage. Si on restreint même «pensée» à «pensée conceptuelle», le langage n'apparaît ni comme nécessaire, ni comme suffisant».

Nous avons nous-mêmes maintes fois assisté à des inférences déjà relativement complexes et irréfutables chez cet animal remarquablement doué qu'est le chat, sans qu'il y ait eu, bien entendu, langage. Mais d'autre part, sans que cette fois-ci il y ait eu véritablement «pensée», mais plutôt réflexe auto-conditionné, nous avons pu observer un phénomène courant, en l'occurrence chez une chatte la compréhension immédiate et non équivoque de quelques mots prononcés au milieu d'une conversation quelconque, sans aucune intonation particulière, mots qui déclenchaient la venue de l'animal avec une totale fiabilité. Il s'agirait ici de la toute première acquisition du langage sous sa forme la plus simple : le réflexe du Pavlov, qui n'est séparé de l'acquisition linguistique que par le fait que le réflexe joue sur une chaîne entière de mots (ici cinq syllabes en trois mots) sans possibilité de les dissocier. C'est également ainsi que l'enfant fait sa première acquisition du langage, la segmentation ne venant que plus tard.

«Avec le mot» — écrivait PAVLOV (1951, t. III, livre 2, p. 215) «s'établit un nouveau principe de l'activité nerveuse : l'abstraction, et en même temps, la généralisation de signaux innombrables avec analyse et synthèse de ces nouveaux signaux généralisés. C'est ce principe qui conditionne une orientation sans limites vers le monde extérieur et sur lequel se fonde l'aptitude supérieure de l'homme : la science».

Donc, au stade de l'animal — qui fut un jour celui de nos lointains ancêtres — on note que signaux phoniques et pensée proprement dite sont assez indépendants. Cependant, privée de l'apport du langage sous toutes ses formes, une telle pensée ne peut être qu'une volition obscure dénuée de toute précision, un peu à la manière du langage musical qui peut, par exemple, exprimer l'approche d'un orage («Shéhérazade», de Rimsky Korsakoff) mais eût pu être interprété différemment sous forme d'un tourment psychologique, et rien, dans la partition ne sera en mesure de nous raconter l'histoire musicalement illustrée.

Au stade élémentaire de la formation de langues les plus anciennes, on note une certaine adaptation des sons aux racines originelles : tel phonème, ou tel groupe de phonèmes semble plus particulièrement adapté à exprimer en bloc une catégorie de pensée très vaste, et par là-même assez vague, en un mot un symbole très général. PAGET (1930, pp. 154 et sqq) met l'accent sur l'importance de ce choix instinctif dans les racines aryennes choisies, qui n'est pas toujours, loin de là, de nature onomatopéique : le contact du dos de la langue contre palais, qui produit /k/, /g/ou/ŋ/, «est associé à des actions telles que celle d'avalier». La consonne /r/

est presque toujours associée à un mouvement de retrait, — généralement l'action de se pencher en arrière, — «d'où l'on peut supposer que le /r/ aryen originel était lui-même produit par un retrait de la langue». De même la semi-consonne /m/ semblerait indiquer une fermeture durable, «ce à quoi on peut s'attendre, étant donné que ce n'est qu'en permettant au courant d'air de passer par le nez qu'on peut garder ses lèvres fermées, et cette forme de fermeture produit inévitablement la consonne /m/... /n/ indique souvent aussi une fermeture continue, /d/ et /n/ ont souvent un rapport avec le fait de se nourrir, de goûter, etc... c'est à dire que la langue touche le palais... /dr/, /tr/ dénotent fréquemment le flot. la marche...» etc...

GUBERINA (1952. p. 127) souligne les remarques de JESPERSEN («Language») sur le choix délibéré de certains phonèmes expressifs en anglais, arménien, turc et danois et remarque que la langue a toujours tendance à fabriquer des mots extralinguistiques (tels que « pif-paf », qui n'a pas la même signification en français et en serbo-croate 1), mots qui finissent tôt ou tard par conquérir les honneurs du lexique.

Le choix de certains phonèmes dans les racines originelles pour exprimer préférentiellement certains symboles larges, s'est parfois conservé jusqu'à nos jours, mais ne pouvait satisfaire qu'à ce chaînon manquant entre la volition obscure et la pensée cohérente. Or cette dernière n'a pu s'élaborer que grâce à la multiplicité des signifiés, devenus pour la plupart phonétiquement inexpressifs. Seuls certains auraient acquis au cours des siècles une certaine expressivité, et se seraient trouvés embellis par le génie du peuple, comme, suivant l'image de BOURCIEZ (1955), des monnaies lissées par le contact de multiples mains, mais ils sont loin de former la majorité.

Les liens entre pensée et langage.

Il était peut-être intéressant de souligner, après PAGET, l'importance de la valeur tant articulatoire que psycho-acoustique de certains phonèmes pour exprimer certains concepts généraux, ne serait-ce que pour nous rappeler que toute langue est avant tout un *système de sons*, et que toute pensée supérieure, qui provient nécessairement du langage *parlé* ne saurait elle-même échapper totalement à une conception psycho-acoustique du langage. La communication linguistique est en effet un processus complexe où tous ces facteurs sont intimement liés.

Pour épuiser le sujet de la communication linguistique à son niveau conceptuel, nous dirons quelques mots sur les liens entre la pensée et le langage. D'une part, une véritable pensée ne peut se matérialiser que dans le cadre des structures linguistiques, d'autre part la langue n'a d'autre fonction que de «signifier». Mais même si d'une langue à l'autre, une pensée sensiblement équivalente peut être rendue par des chaînes de symboles différents, il n'en reste pas moins vrai que la langue traduit les symboles, et que les symboles s'appuient eux-mêmes sur la langue. WHORF (1956, p. 2) donne un exemple de cette interdépendance en présentant les symboles très différents utilisés pour la même phrase en anglais et en shawnee, «Clean with ramrod» — nettoyer (un fusil) avec une tige — devient en shawnee la chaîne de symboles : «espace sec — intérieur d'un trou — par mouvement d'un outil» («nipēkwālakha»).

On ne peut guère plus imaginer une pensée sans langage (sous quelque forme que ce soit), qu'on ne peut imaginer un enchaînement de mots (ou de signes gestuels ou mathématiques) sans suite, donc privé de pensée. Cependant établir une solidarité mutuelle et nécessaire entre pensée et langage ne suffit pas à expliquer la nature des liens qui les unissent, car il existe une opposition entre le langage et la pensée. Si l'on admet que la pensée se développe au mieux

1) — En français, il s'applique à un bruit sec alors qu'en serbo-croate, il est unanimement utilisé pour illustrer soit un coup, soit l'actualisation d'un jet de salive... (*Ibid.* p. 128).

dans le cadre du langage parlé — même s'il n'est que pensé. — langage dont les structures complexes sont, rappelons-le, subconscientes chez l'utilisateur qui s'en sert depuis son plus jeune âge. Il n'en reste pas moins vrai que ces structures peuvent être étudiées et détaillées par les grammairiens et les linguistes. Mais peut-on accéder à une connaissance aussi précise de la pensée elle-même ? Si on pouvait connaître la nature intime de la pensée, et la définir par des traits qui lui sont propres, peut-être verrait-on alors comment elle s'ajuste au langage.

Cette question avait été abordée par ARISTOTE dans ses «Catégories» (cf : bibl., 1966, chap. 4, pp. 5-6). Ces «catégories» ne présentent pas le même aspect suivant qu'elles sont catégories de pensée ou catégories de langue. Il ressort de l'ensemble de cette étude que la pensée peut spécifier librement ses catégories, en créer de nouvelles, alors que les catégories linguistiques, déjà automatisées par l'usage, ne peuvent être modifiées : elles sont reçues et utilisées par chacun dans le cadre de règles précises, et forment un code particulier pour chaque langue. Une telle considération suffirait déjà à confirmer la position indépendante de la pensée à l'égard de la langue.

Une conception abstraite des éléments du langage parlé : la phonologie.

Epuiser le sujet excessivement vaste de la communication linguistique nous entrainerait beaucoup trop loin, et nous n'avons pas eu d'autre intention que d'en esquisser les grandes lignes. De l'avis unanime de linguistes, cette communication linguistique s'appuie avant tout sur le langage *parlé*, même s'il n'est que pensé, c'est à dire un système de signaux sonores, réels ou virtuels. Car qu'est-ce que le langage pensé si ce n'est une image cérébrale du langage parlé. Un Chinois lisant son écriture idéographique, un radio-opérateur lisant une écriture de signaux Morse auront tous deux la représentation cérébrale de l'équivalent sonore des signes déchiffrés, et il est sans doute impossible d'avoir une pensée tant soit peu abstraite sans «entendre» mentalement les signifiants qui la forment. Ces mots sont eux-mêmes formés de phonèmes, ou plus petites unités sonores pertinentes, capables à eux seuls, par le simple fait d'un changement de l'un deux, de transformer un mot en un autre tout différent. L'étude de la fonction des phonèmes est du domaine de la phonologie, alors que l'étude de leur nature physique et psycho-acoustique est du domaine de la phonétique. Il convient de faire cette distinction en délimitant avec précision ces deux domaines, sans oublier que ce qu'on appelé plus récemment la «phonologie» (concept encore mal détaché de celui de phonétique au début du siècle, cf : GRAMMONT, 1965, p. 144) est antérieur de plus de deux millénaires à ce qu'on appelle actuellement la «phonétique».

En effet les premiers phonologues ont été les inventeurs inconnus de l'écriture, qui ont imaginé de peindre les sons des mots au lieu des objets dont ils sont les signes, et de remplacer par cette représentation nouvelle l'écriture idéographique qui était le système égyptien et reste actuellement, par exemple, le système chinois.

On sait que l'écriture idéographique fit place dans nombre de pays à l'écriture syllabique, première simplification qui permettrait de traduire en bloc des groupes indissociables de logatomes. Une langue ne comporte pratiquement jamais plus de 2 000 syllabes qui constituent un inventaire *complet* de toutes ses possibilités phoniques, et il est question actuellement de revenir à un système instrumental équivalent pour l'identification (DEWEZE, 1966, pp. 159-160) du langage par les machines.

Puis on assiste à une nouvelle simplification encore plus considérable : les syllabes sont divisées en phonèmes, en nombre extrêmement réduit, et dont les combinaisons peuvent représenter un nombre presque illimité de mots différents. On sait que le premier système de ce genre, l'écriture phénicienne, ancêtre de presque tous les alphabets actuels, ne traduisait que les consonnes. En adoptant l'alphabet sémitique les Grecs s'aperçoivent que les seules consonnes sont insuffisantes à traduire la richesse phonétique de leur langue, et introduisent à leur tour les voyelles. Ces «phonémas» complètent les «sumphona». Platon, Aristote, les Epicuriens et les Stoï-

ciens créent la science des «lettres», la «grammatiké techné». La grammaire est donc à l'origine l'étude des «lettres», et cette étude des «lettres» était elle-même *l'étude des sons qu'elles représentent*.

ABERCROMBIE nous définit ainsi la «lettre» (1965, p. 76) : «La *lettre* est le terme clef de toute discussion sur les rapports entre la parole et l'écriture», et (p. 84) : «les étudiants de linguistique ont sans doute intérêt à éviter le terme ambigu de «lettre». Les discussions typographiques sur les «figura», et les arguments phonologiques concernant les «potestas», sont encore la principale préoccupation des théoriciens de transcription phonétique de nos jours, mais de tels problèmes sont sans doute abordés plus aisément en termes de «sons du langage», «symbole», et «phonème» »

La «lettre», même si elle devient «signe phonétique», n'est donc qu'une représentation conventionnelle d'un certain phonème *normalisé* au sein d'une certaine communauté linguistique. En un mot, c'est une *abstraction* procédant d'un désir de simplification que OSGOOD et SEBEOK (1967, p. 11) nous présentent en ces termes : «Le nombre de sons potentiels dans une langue tend vers l'infini. Le grand avantage du système phonémique est de permettre au linguiste d'effectuer une puissante réduction de puis cette complexité jusqu'à un nombre limité de signaux qui constitue le code, et cela représente une grande économie de description...».

Pour simplifier encore la classification, le phonologue, suivant les principes de TROUBETZKOY (1949), l'un des principaux initiateurs de l'École de Prague, adopte le principe des «formes distinctives», par exemple des phonèmes appariés sourde/sonore : p/b, k/g, t/d, s/z etc...

JAKOBSON (1971, pp. 484-486) va plus loin en imposant à tout matériau phonémique une opposition binaire, en deux classes globales : formes de sonorité et formes de tonalité, puis en subdivisant binairement ces deux classes, la première en : vocalique/non vocalique — consonnantique/non consonnantique — compact/diffus — tendu/relâché — voisé/non voisé — nasal/oral — discontinu/continu — strident/vélu — glottalisé/non glottalisé, et la seconde en : grave/aigu — «flat/non flat» — «sharp/non sharp» (ces deux derniers termes difficilement traduisibles exprimant l'effet acoustique des labialisations et palatalisations).

Un tel système d'étude qui consiste à cerner l'objet étudié sous le maximum d'angles différents ne se fixe pas cependant pour but de *décrire* cet objet, en l'occurrence le phonème, mais seulement de *l'entourer d'un nombre suffisant de critères distinctifs* pour permettre de le placer dans une certaine catégorie *qu'il ne partagerait pas avec un objet différent*. Cela revient à dire que la phonologie présente un phonème conventionnel, valable pour une certaine communauté linguistique, sous un jour purement fonctionnel, c'est à dire aussi abstrait et aussi peu caractérisé qu'une pièce d'échecs sur un échiquier.

La description phonologique s'appuie sur certaines unités fondamentales que l'analyse linguistique permet d'isoler : tout d'abord, nous l'avons vu, le phonème «phonologique», puis, à l'échelon directement supérieur, le morphème 2). Si le phonème est l'unité de description du système phonologique d'une langue, le morphème est, moins bien délimité, la base d'une description grammaticale.

BLOOMFIELD (1933) posait comme base fondamentale que dans une communauté linguistique, certaines phrases dénotent une partielle similarité « formelle - sémantique ». Par exemple, dans une communauté francophone, les phrases : «il est entré dans la maison» et «il est entré dans le salon», sont partiellement similaires de par leurs séquences de phonèmes, et

2) — Dont la notion générale avait déjà été effleurée par ARISTOTE (« De l'interprétation », chap. 4 dans «Organon» — cf. Bibl. 1966).

réfèrent à des situations partiellement similaires. En poussant au maximum les réductions de séquences phoniques similaires, Bloomfield en arrivait à isoler les morphèmes, ou plus petits segments phoniques porteurs de sens, segments qui parfois correspondent à des mots, mais sont le plus souvent plus petits. Certains mots sont monomorphémiques, tels «house», mais certains sont plurimorphémiques : ainsi «sing» est un morphème, «singing» est formé des morphèmes «sing» et «ing», ce dernier étant la marque du participe présent ou du gérondif. Cependant si le système morphémique peut servir de base à une grammaire, il n'en est pas moins entaché d'arbitraire, car tous les linguistes ne s'accordent pas sur la possibilité d'un groupe de phonèmes à accéder au statut — mal délimité — de morphème. Ainsi un mot comme «deceive» comprend un morphème valable «de» (bien que non porteur de sens quand il est considéré isolément), car il est porteur d'un sens privatif dans le cadre de mots, et qu'on le retrouve tel quel dans d'autres structures. Cependant ce n'est pas le cas du contestable morphème «ceive», variante de «cep», qui, isolé, n'a aucun sens.

Le problème du structuralisme morphémique tel que soumis à l'orthographe, revêt actuellement une importance particulière du fait de son utilisation dans les machines à traduire (dont on trouvera, entre autres, les principes dans les ouvrages de DELAVENAY : 1963, et GUTENMAKHER : 1961). Il apparaît évident que la solution d'un tel problème ne peut être confiée aux seuls linguistes, mais mène nécessairement à une collaboration harmonieuse entre linguistes, informaticiens, cybernéticiens et mathématiciens. Une étude même rapide des articles spécialisés dans ces sens, en particulier celui de VIETZE (1970) aura sans doute tôt fait de convaincre les plus traditionalistes que le chemin tracé par les anciennes écoles de phonologie a fini par déboucher dans un domaine immense et hérissé de difficultés.

La formation du morphème est conditionnée par des règles particulières à chaque langue. Ces règles ont été reprises en particulier par CHOMSKY (1967, pp. 30-34) qui expose la probabilité d'associations phonémiques au sein d'une langue, sous le nom de «degrés de pertinence en phonologie». «Considérons d'abord» — écrit-il — «le cas de ce qu'on appelle les «lacunes accidentelles» dans le lexique. Ainsi en anglais, il y a un mot «pick», mais aucun «blick» ou «ftick». Pour atteindre le niveau de pertinence descriptive, une grammaire devrait fournir, par surcroît, une règle générale qui dresse une barrière spécifique contre «ftick», mais pas contre «blick» (que nous qualifierons ainsi de «lacune accidentelle», une syllabe dénuée de sens, mais acceptable phonologiquement).. Ce niveau serait encore perfectionné par une grammaire qui contredirait la généralisation qu'en position initiale devant une consonne véritable (un segment consonnantique et non vocalique, dans les termes des formes distinctives de JAKOBSON (1971, pp. 484-486), une consonne est nécessairement /s/. Le niveau de pertinence explicative serait atteint par une théorie linguistique qui fournirait une raison de principe pour incorporer cette généralisation dans une grammaire de l'anglais...»

Telles sont, citées par CHOMSKY, certaines des grandes règles phonologiques qui apparaissent dans la théorie des règles structurales de morphèmes, développée par HALLE (1959 a et b).

Pour définir le système phonologique, il suffit donc de constater et de classer les faits pour en déduire les caractéristiques du système. Cependant pareil structuralisme purement observationnel ne saurait s'appliquer qu'à la phonologie synchronique, c'est à dire aux faits, de langue dans le cadre d'une certaine époque déterminée.

En plus de la probabilité d'association phonémiques dans un contexte synchronique, AVRAM (1971, p. 448) considère la probabilité des changements phonétiques sur un plan diachronique. «Il est évident» — écrit-il — «que la probabilité qu'un phonème x se transforme en y dépend, premièrement, des traits distinctifs des deux phonèmes et en même temps de leurs traits phonétiques... Tandis qu'un changement /i/ — /e/ est très probable, la probabilité du changement /i/ — /o/ est beaucoup plus réduite, et celle du changement

/i/ — /s/ est nulle. D'autre part la probabilité que x devienne y dépend aussi de la position qu'occupe le premier. Par exemple la transformation de /s/ en /z/ est plus probable dans le contexte /a...a/, que dans le contexte /a...t/ ».

La dualité du phonème.

Il n'est pas dans nos intentions d'épuiser le sujet, aussi vaste que controversé, de la phonologie, sinon pour la présenter dans ses grandes lignes, nécessairement incomplètes. Mais il fallait en parler ne serait-ce que pour mettre l'accent sur le fait qu'en phonologie, le phonème, même présenté sous sa forme la plus abstraite, n'en demeure pas moins un élément de base qu'il serait souhaitable de définir et de délimiter dans la mesure du possible. Cependant avant d'aborder le problème de la délimitation du phonème, il nous faut essayer d'en dégager une première qualification à travers les aspects très différents sous lesquels on le présente en phonologie et en phonétique. L'imprécision dans laquelle se confine la définition du phonème tient en premier lieu à la dualité conceptuelle qui oppose phonétique et phonologie. Nous n'étudierons pas cette dualité 3) trop connue, sinon pour en dégager une première constatation : non seulement le phonème revêt en phonétique deux aspects particuliers, mais même au sein de l'une ou l'autre de ces sciences, l'unanimité est loin d'être faite sur sa définition précise.

Pour nous faire une première opinion à ce sujet, nous ne saurions mieux faire que de citer in extenso quelques auteurs parmi les plus connus, en commençant par JAKOBSON (1971, p. 231) qui nous présente le phonème en ces termes : «Le phonème est le concept de base de la phonologie. Par ce terme nous désignons un éventail de ces propriétés sonores concurrentes qu'on utilise dans une langue donnée pour distinguer des mots de signification différente. Dans l'acte de parole, des sons différents peuvent se rapporter à un seul et même phonème. Cette variété dépend du style de la parole et/ou du contexte phonétique dans lequel le phonème se présente. La différence entre de tels sons est déterminée par des facteurs externes et de ce fait ne peut servir à distinguer la signification des mots. De tels sons portent le qualificatif de variantes du phonème en question... La phonologie est une partie de la linguistique qui a trait aux sons de la parole en rapport avec les fonctions qu'ils remplissent dans une langue donnée, alors que la phonétique se donne pour tâche l'investigation des sons de la parole d'un point de vue purement physiologique, physique et psycho-acoustique».

Nous resterons quelques instants dans le domaine de la phonétique, pour souligner que l'aspect *psycho-acoustique* du phonème, aspect dont parle JAKOBSON et qui est pour nous le plus important, passe au second plan chez nombre de phonéticiens pour des raisons avouées de facilité d'approche. Par exemple, pour MARTINET (1964, p. 45) la phonétique est surtout articulatoire. «C'est» — dit-il — «en référence à la façon dont ils sont réalisés au moyen des «organes de la parole» que seront identifiés... les traits phoniques pertinents et que seront décrites les variantes des unités phonologiques. On pourrait utiliser aux mêmes fins les ondes sonores produites par le jeu des organes. Mais la phonétique articulatoire reste plus familière à la plupart des linguistes et, en général, elle permet de mieux percevoir la causalité des changements phonétiques».

Par ailleurs, il déplore la dualité qui n'oppose que depuis un demi-siècle la phonétique et la phonologie, (qu'il place dans le cadre de la phonétique), en ces termes (1965, p. 36) : «On pourrait présenter la phonologie comme une façon d'envisager la phonétique : ce serait la phonétique traitée des points de vue fonctionnel et structural...» — et (Ibid., p. 37) : «Le main-

3) — Dualité officialisée dans les transcriptions : en phonologie, le phonème s'inscrit, rappelons-le, entre traits obliques, alors qu'en phonétique, il s'inscrit entre crochets.

tien de la distinction entre «phonétique» et «phonologie» résulte en fait de circonstances historiques : la phonologie a été, en quelque sorte, lancée comme on lance aujourd'hui un produit dans le commerce, vers la fin des années vingt, par des linguistes plus soucieux de marquer l'originalité de leurs façons de voir, que de s'intégrer dans le cadre de la recherche traditionnelle. Il en est résulté des heurts, des conflits, des prises de position passionnées qui se sont cristallisées autour de ces deux termes ».

Le principal concept qui oppose, sur le plan du phonème, phonétique et phonologie, étant celui de «pertinence», la valeur du phonème en tant que réalisation concrète passe, en phonologie, au second plan, ce qui n'aide guère à l'identification précise de cet élément de base : pour le phonéticien, le phonème est intéressant par lui-même, par la manière dont on l'articule, par l'impression qu'il crée sur l'oreille ; pour le phonologue, le phonème n'est intéressant que parce que c'est la plus petite unité phonique capable de différencier deux mots, conception si abstraite qu'elle peut, dans certains cas, mener jusqu'à une dénaturaison de la valeur observationnelle de l'étude du phonème. C'est ainsi que, alors que le phonéticien doit reconnaître actuellement, la longueur mise à part, la normalisation phonétique du /a/ de «patte» et du /ɑ/ de «pâte» le phonologue continue de considérer qu'il s'agit là de deux phonèmes différents en ce sens qu'ils ont une valeur oppositive suffisante à différencier linguistiquement ces deux mots.

On peut citer des exemples illustrant la dualité des phonèmes phonétique et phonologique, en considérant l'éventail acoustique réel fourni par une langue comme le français : le phonème anglais [ŋ] de («song») existe réellement en français, accidentellement — et phonétiquement — («une longue distance»), mais n'a aucune valeur distinctive pertinente, donc aucune existence phonologique, bien qu'il soit réalisé phonétiquement par environ 70% des Français dans certains mots récemment importés, tels «ping-pong», comme le fait justement remarquer MALMBERG (1969, p. 106).

TROUBETZKOY dit (1949, p. 13) : «La phonétique demeure toujours sur le terrain de l'acte de parole tandis que la phonologie... reste toujours sur le terrain de la langue. Les deux définitions se recouvrent réciproquement : la phonologie est l'étude des sons de la langue, la phonétique l'étude des sons de la parole ; en outre la phonologie s'occupe nécessairement de la fonction linguistique des sons du langage, la phonétique par contre, du côté phénoménologique de ces sons, sans égard à leur fonction».

PRIETO (1964, p. 24) définit ainsi la phonologie : «On peut résumer l'essentiel de cette théorie» (la phonologie pragoise...) «Elle considère comme *connu* le signifié des phonies, dont dépend... leur contribution à l'établissement du sens. Il suffit donc, pour classer les phonies, de tenir compte de leurs signifiés...».

FALINSKI (1966, p. 11) établit les liens entre phonèmes «phonétiques» et phonèmes «phonologiques» qui, au sein même de chacune de ces sciences, ne recueillent jamais l'unanimité des voix pour ce qui est de leur définition exacte et de leur délimitation : «Phonétiquement» — dit FALINSKI (*ibid.*) — «le français possède 16 voyelles d'après l'Académie française, 22 voyelles selon l'Abbé ROUSSELOT, le créateur de la phonétique expérimentale, 21 voyelles selon le Professeur FOUCHE, l'éminent phonéticien de la Sorbonne 4), 15 voyelles, 3 semi-voyelles et 17 consonnes selon la majorité des phonologues ».

MALMBERG (1969 p. 25) délimite la phonétique et la phonologie en plaçant cette dernière dans le cadre de la phonétique : «On appelle souvent phonologie la partie de la phonétique consacrée à l'analyse des phonèmes et du système que forment ceux-ci (forme de l'expression), et phonétique (dans un sens restreint) l'analyse de la *substance* qui manifeste cette forme.

4) — Cf : Bibl. : 1951.

On appelle aussi *langue* le système *abstrait* de possibilités distinctives, et *parole* l'acte *concret* de communication réalisée».

Ces quelques références montrent suffisamment le caractère concret, d'une part, de la phonétique, étude de la *substance*, et d'autre part le caractère abstrait de la phonologie étude de la *forme* de la langue, réduite à une conception fonctionnelle.

La simplification phonologique est justifiée du fait que, au sein d'une même communauté linguistique, les variantes phonétiques sont en général faibles. Dès qu'il a passé le stade de l'enfant où JAKOBSON (1949, p. 368) note «un oubli de liaison entre l'image acoustique et motrice», l'individu doit se plier strictement aux normes phonétiques de sa langue, sous peine d'attirer la dérision.

Car le langage est un phénomène social. C'est pourquoi TROUBETZKOY (1949, pp. 343-350) plaçait l'étude phonologique dans le cadre de la géographie linguistique et étudiait séparément (ibid. p. 343) «les différences dialectales : phonologiques, phonétiques, et étymologiques».

On note que, vu sous son jour «phonologique», c'est à dire purement fonctionnel, le phonème reste un concept d'autant plus mal délimité qu'il n'est plus formellement attaché ni à la «lettre» censée le représenter, ni à l'impression acoustique qu'il donne. CHOMSKY met l'accent sur cette imprécision en considérant (1967, p. 31) la question des variantes phonétiques prévisibles d'une racine comme celle du mot «te legraph», qui se retrouve avec des phonèmes phonologiques que l'on peut qualifier de semblables — en ce sens que, prononcés correctement ou non, les mots restent parfaitement reconnaissables, leurs variantes phonétiques n'ayant pas de valeur distinctive — mais phonétiquement tout à fait différents, par exemple dans les mots «te legraph» — «telegraphic» — «telegraphy». En fait le dictionnaire phonétique de D. JONES (1947, p. 419) donne les prononciations suivantes :

«'teligra:f», «teli'græfik», «ti'legrəfi».

Bien que l'orthographe des racines reste la même et que les phonèmes phonologiques soient généralement considérés comme les mêmes, la place différente de l'accent d'intensité dans les trois cas a transformé complètement la nature psycho-acoustique des voyelles et CHOMSKY considère ce cas, d'ailleurs exceptionnel, comme une véritable *alternance vocalique*, comparable, bien que non enregistrée par l'orthographe, à celle qui relie «see» et «saw», ou encore «man» et «men». La grammaire préconisée par CHOMSKY («item and arrangement grammar») range cette racine complexe dans les exceptions qui demandent une règle particulière soumise au contexte interne du mot (adjonction de suffixes différents jouant le même rôle phonétique que la modification par alternance vocalique nécessaire à la mise au passé du verbe « to see » ou au pluriel du nom « man »).

De tels exemples montrent d'abord le peu de fiabilité sensorielle des phonèmes si leur perception n'est pas continuellement soumise à des règles phonologiques que l'exception même confirme et qui, si elles sont subconscientes, n'en sont pas moins continuellement présentes aux profondeurs de l'esprit humain et viennent en aide à l'intelligibilité du discours. Mais ils montrent aussi le peu de cas que l'on peut faire du phonème « phonologique » en dehors de l'étude purement fonctionnelle à laquelle il se confine.

Le phonème : réalité ou abstraction ? Le problème de sa délimitation.

Pour le phonologue, le phonème, base même de son étude, est une abstraction dont il connaît l'existence et dont il sous-entend la toute-puissance au sein d'une langue donnée. Il en parle comme le peintre parle de la couleur, c'est à dire par référence à certains tubes : bleu de prusse, vert Véronèse, etc., couleurs dont il ignore d'autant plus volontiers les propriétés tant

physiques que sensorielles, qu'il est sûr de les retrouver telles quelles dans tel ou tel tube provenant de tel ou tel fournisseur.

Or le phonème n'est pas une abstraction. Il n'est que le résultat audible de certaines vibrations de l'air. Pas plus qu'une abstraction, il ne possède son contraire, et tout au plus les phonologues ont-ils pu le caractériser en lui accordant une opposition binaire telle que «sourde /sonore», opposition que des phonéticiens ont pu étudier plus concrètement : le regretté DELATTRE (1965, pp. 113-118) sur le plan psycho-acoustique en se servant des ressources nouvelles de la synthèse de la parole, HADJ-SALAH (1971) sur le plan articuloire en se servant des plus modernes méthodes de l'élasticimétrie.

Non seulement dans ses conceptions que nous appellerions aujourd'hui «phonologiques», mais aussi dans toute sa philosophie, Aristote opposait la «forme» à la «substance». (cf : Bibl., 1966, catégories, chap. 5, pp. 17 et sqq). JAKOBSON fait le point sur une telle dualité en linguistique (1971, pp. 751-752) : « Les études de Eli FISCHER-JØRGENSEN montrent clairement » (à propos des divergences de vues entre Hans Jørgen UL DALL et Louis HJELMSLEV) « que la présentation de la forme en dehors de tout contact avec la substance et vice-versa, s'avère dans un cas comme dans l'autre tout aussi dénué d'intérêt. C'est avec une particulière clairvoyance qu'elle a montré à quel point serait sans objet l'analyse phonétique si on ne tenait pas compte des fonctions variées et spécifiques des sons de la parole ».

Dans un même ordre d'idées, nous pouvons dire aussi qu'il y a des substances qui subordonnent la forme. Par exemple l'air (substance) subordonne la forme (image formantique) du phonème. Si cette substance est remplacée par une autre (l'azote surcomprimé dans les profondeurs sous-marines) qui déforme les sons vocaux du fait que les résonateurs vocaux sont prévus pour fonctionner dans un fluide aux caractéristiques de célérité bien déterminées (l'air à pression normale en l'occurrence) la voix cesse d'être compréhensible du fait d'une anamorphose excessive de la structure formantique des phonèmes, que les facultés d'adaptation du cerveau ne peuvent plus suivre.

Tout ceci nous amène à considérer que le phonème ne peut être considéré comme appartenant à un type bien déterminé, donc pertinent du point de vue linguistique, que s'il demeure à l'intérieur d'un cadre déterminé tant par les configurations articuloires qui lui donnent naissance dans des conditions normales, que par les configurations sonores qui en découlent directement.

A l'intérieur de ce cadre, il pourrait subir certaines modifications qui ne nuiraient pas à son intelligibilité du fait de mécanismes correcteurs psycho-physiologiques dont les effets rappellent ceux de la vision des couleurs d'objets éclairés par des lumières riches en dominantes colorées (cf : EMERIT, 1966, p. 219, citant les découvertes de Hussain Yilmaz et Lewis Clapp). Une étude intéressante sur l'intelligibilité de la voix synthétisée par «icophone», soumise à des anamorphoses formantiques, a été réalisée par LEIPP (1970). Dans la vie courante il faut également souligner l'importance des conditions de filtrage du son tel que transmis au voisinage du sol (cf : BRILLOUIN, 1935). Mais il est intéressant de constater dans tous les cas que les phonèmes ainsi déformés le sont toujours dans le même sens et dans la même proportion. De tels phonèmes isolés, anamorphosés alternativement dans un sens puis dans l'autre ne seraient sans doute plus reconnaissables. Tant dans la perception de la mélodie que celle du timbre, il existe une certaine « mise au diapason » de l'oreille (EMERIT 1966, p. 196) qui avait été pressentie par ROUSSELOT (« Phonétique et surdité », p. 114, cité par MILLET, 1926, p. 69). Une telle faculté d'adaptation explique l'impression de sûreté de notre oreille en présence de sons continuellement déformés, et vient encore compliquer le problème de la délimitation du phonème.

Les phonologues, qui se sont transformés en véritables grammairiens ont pu ignorer tant de subtilités, en laissant aux phonéticiens le soin d'une telle délimitation qui serait valable tant phonologiquement que phonétiquement.

C'est à E et K. ZWIRNER (1936) qu'il appartenait de tenter d'établir les bases d'une «phonométrie», où le phonème «phonologique» ne serait que *la réalisation verbale d'un modèle physique*, formantique et intonatoire. à l'intérieur de cadres dont les centres seraient des moyennes obtenues à partir d'un certain nombre de sujets appartenant à un même groupe linguistique. E. ZWIRNER exprime le fond leur méthode en ces termes (Bibl : ZWIRNER E, cité par TROUBETZKOÏ, 1949, op. cit, p. 7) : «Etant donné que de telles normes ne peuvent être réalisées deux fois exactement de la même façon, la transition de la phonologie comme science des normes, à la phonétique comme leur *réalisation* dans le langage, doit être de nature *statistique*, de sorte que les variations d'un son soient distribuées autour de leur moyenne selon la loi de Gauss (courbe normale), et ces moyennes sont ce qui correspond aux normes». Ce genre de classification a été repris sur des bases nouvelles (cf : infra) par DELATTRE (1948, et 1966, p. 266) sous forme de son célèbre «triangle vocalique». Cependant l'auteur mettait l'accent sur le fait qu'une telle délimitation devrait être vérifiée par synthèse vocalique et *sur la foi des données de l'audition*, ce qui revenait à placer le phonème dans un cadre psycho-acoustique.

Délimitation de la psycho-acoustique.

Avant de considérer le rôle de la psycho-acoustique dans la délimitation du phonème, il conviendrait de la délimiter elle-même en disant qu'elle est l'étude de la sensation sonore globale, c'est-à-dire la différence de la psycho-physiologie de l'audition qui n'est que l'étude des transformations du signal sonore dans le système nerveux.

Comme tout ce qui touche aux manifestations cérébrales, la psycho-acoustique est fort mal connue. On connaît les trois principales localisations cérébrales de fréquences sonores, découvertes dans le cerveau du chien par TUNTURI (1944), mais on ignore à peu près tout de ce qui s'y passe dans le domaine de la sensation, et on ne peut aborder la psycho-acoustique que grâce aux manifestations tangibles que l'audition d'un son complexe peut provoquer chez un sujet, qui décrit alors ce qu'il ressent ou ce qu'il comprend.

Quand on parle de psycho-acoustique, on pense immédiatement à la loi dite «de WEBER-FECHNER» («la sensation croît comme le logarithme de l'intensité» ou encore : puissance sonore) et c'est ainsi qu'elle est citée sans discussion par CHRETIEN (1964, p. 8) et d'innombrables autres auteurs. Or non seulement elle n'a pratiquement aucun rôle à jouer en psycho-acoustique, mais elle s'est trouvée formellement contredite par les indiscutables résultats expérimentaux de STEVENS (1936). Nous avons expliqué (1970) pourquoi cette loi qui reposait sur des bases relativement justes (le «quotient de WEBER» qui est presque une constante, du moins à partir de 50 dB, le caractère *quasi*-logarithmique en fonction de la puissance sonore, de l'augmentation du nombre de paliers sensoriels de plus petite différence d'intensité notable, ou « JND », qui en découle directement) était quantitativement fautive par suite d'un détail que ses promoteurs n'avaient eu aucun moyen de prévoir et encore moins de vérifier : l'énorme expansion sensorielle des « JND » au niveau du cerveau, suivant une formule précise et inédite (*ibid.*).

Point ne sert de réajuster cette loi comme nous l'avons fait (*ibid.*) en suivant une méthode très simple, bien que similaire dans ses résultats à celle que DREYFUS-GRAF (1971, p. 100) adoptait indépendamment l'année suivante, c'est à dire en faisant coïncider à 20 dB courbe de sonie et courbe d'intensité, transformée pour la circonstance en courbe de pression sonore, ce qui revient à modifier toutes les valeurs de sonie théorique de la loi de Weber-Fechner.

De tels ajustements ne sauraient sauver une loi dont, par surcroît, l'intérêt en phonométrie, même si elle était juste, serait des plus réduits, car en adoptant la notation expérimentale de force subjective en SONES (telle que définie par STEVENS : 1936), on s'aperçoit que

cette sonie est grossièrement *proportionnelle*, au moins sur un éventail de 30 dB (rapport 1-32, courant dans le langage) *aux variations de pression sonore*, qu'on peut lire directement telles qu'elles sur l'écran de n'importe quel oscilloscope...

Cette sonie elle-même, étant donné son caractère de simple mesure, ne peut guère être considérée comme du domaine de la psycho-acoustique, mais plutôt du domaine de la psychophysologie de l'audition, de même que tous les autres phénomènes intéressant le traitement du signal sonore dans le système nerveux, tels que : fatigue auditive, effet de « masque » (WEGEL et LANE, 1924), isosonie (FLETCHER, 1933) perception de la hauteur tonale des sons purs suivant l'échelle des « mels » (de STEVENS et VOLKMAN, cf : WINCKEL, 1960, pp. 63-64) aberration de hauteur tonale subjective en présence de sons purs d'intensité variable (FLETCHER, 1934, STEVENS, 1935), seuils différentiels d'intensité (RIESZ, 1928) et de fréquence, sons de composition et perception de la hauteur tonale en l'absence de fondamental (FLETCHER 1924, p. 7), battements, perception de la phase (EXNER et POLLACK, 1903) etc... 5)

En effet si tous ces phénomènes peuvent, à des degrés divers, intéresser l'audiologie, leur connaissance n'offre guère d'intérêt en linguistique, car la langue n'est pas formée de sons purs, mais de suites de sons extrêmement complexes et en permanente évolution, qui sont perçues *comme un tout acoustique* intimement lié à des *structures acoustiques déjà profondément imbriquées dans la mémoire*.

Il nous faut donc faire une distinction fondamentale entre la *psycho-physologie de l'audition* qui s'appuie sur l'étude de phénomènes tangibles parfaitement connus d'une part, et d'autre part la *psycho-acoustique*, dont on ne sait presque rien car elle s'intéresse à des manifestations purement cérébrales, confinées à la « boîte noire » des psychologues, la « terra incognita » de la sensation globale. Cette science que l'on pourrait qualifier de naissante, s'appuie actuellement en majeure partie soit sur des tests d'intelligibilité en voix filtrée, comme ceux déjà anciens de FLETCHER et STEINBERG (1929) ou par déformation de signaux sonores (UNGEHEUER : 1958), soit sur des expériences de synthèse de la parole (LIBERMAN et al. : 1959, DELATTRE : 1962, 1965).

On comprendra mieux la profondeur du fossé qui sépare psycho-physologie de l'audition et psycho-acoustique en notant le nombre considérable de théories de l'audition — dont aucune n'a jamais pu être prouvée à cent pour cent. WEVER (1949) en cite plus d'une trentaine, la plupart incompatibles. Alors que les théories de la *perception auditive* — qui tendent à expliquer de façon nouvelle comment les sons *sont réellement perçus et décodés au niveau du cerveau* — abstraction faite parfois même des moyens mis en œuvre par la nature — peuvent se compter sur les doigts d'une seule main : théorie de HELMOLTZ, 1859 (1863, chap. 4) : la sensation est obtenue à partir d'une analyse de fréquences — théorie de RUTHERFORD, 1899 : la sensation est obtenue par « lecture » d'un « oscillogramme » nerveux — théorie de WEVER et BRAY, 1930 (1949) : la théorie de Rutherford se vérifie jusqu'à près de 3.000 Hz, celle de Helmholtz se vérifie aux fréquences supérieures, aux fréquences moyennes les deux processus collaborent — notre théorie de « synchro-trichromie auditive (EMERIT, 1966, 1967, 1970, 1972) : modification de la théorie de Wever et Bray par la notion que le processus d'analyse fréquentielle se réduirait à isoler les enveloppes d'intensité de trois plages, basses, moyennes et hautes fréquences au niveau du cerveau à la manière des messages visuels qui lui fournissent la couleur globale d'un objet à partir de trois teintes fondamentales.

5) — Nous ne pouvons citer ici que quelques unes des principales publications qui nous ont effectivement servi à étayer nos précédentes études. Une bibliographie même incomplète sur les faits psycho-physiologiques comprendrait un nombre très considérable de titres, ce qui montre bien la disproportion entre la connaissance des faits psycho-physiologiques, d'une part, et celle de la psycho-acoustique de l'autre. On peut consulter avec profit, WEVER (1949) STEVENS and DAVIS (1947), HIRSH (1956).

La délimitation psycho-acoustique des phonèmes, en particulier des voyelles.

Lorsque DELATTRE a publié «un triangle acoustique des voyelles orales du français» (1948, et 1966, p. 266) la parole était encore à l'optimisme né de la popularisation des premiers sonagrammes par POTTER, KOPP et GREEN (1947). DELATTRE disait notamment (1948, p. 477) : «On possède un nouveau moyen de distinguer visuellement une voyelle d'une autre avec une grande facilité. Désormais on pourra VOIR jusqu'aux moindres nuances, et on pourra les traduire instantanément en chiffres — deux chiffres pour être exact... De fait cette décomposition vocalique est tellement aisée que l'on peut prévoir le moment très proche où l'on ne mentionnera plus le timbre d'une voyelle sans la définir numériquement... On peut prévoir aussi que la description physiologique des sons, telle qu'on l'a faite jusqu'à présent, va être reléguée au second plan... ».

Ce triangle était basé sur les fréquences de résonance des deux premiers formants des voyelles : la fréquence du premier formant était reportée en abscisse, et celle du deuxième en ordonnée. Chaque voyelle était donc représentée sur ce diagramme par un point occupant une place particulière. Cependant, avec la clairvoyance qui l'a toujours caractérisé, DELATTRE mettait un frein à tout enthousiasme excessif, car il ajoutait notamment (1966, p. 82) : « soulignons maintenant le fait que le triangle acoustique que nous présentons ici n'est pas LE triangle acoustique français mais UN triangle acoustique du français. LE triangle acoustique du français exigera de longs travaux statistiques dans lesquels il faudra tenir compte, non seulement des prononciations de diverses personnes, mais de la prononciation de chaque voyelle dans toutes les conditions phonétiques notoires qu'elle rencontre au cours de la chaîne parlée, aussi bien que des voyelles isolées... La seule garantie de la qualité française des timbres vient de ce que, dans chaque cas, les voyelles ont été d'abord enregistrées sur la bande magnétique du spectrographe, puis ont été écoutées, effacées, enregistrées à nouveau, et écoutées à nouveau jusqu'à ce que le timbre de toutes les voyelles soit entièrement satisfaisant à l'oreille. »

DELATTRE mettait déjà ici l'accent sur les deux obstacles majeurs à la caractérisation «physique» des voyelles : a) la voyelle ne prend son timbre véritable qu'en centre de syllabe, hors de la zone de transition consonnantique qui doit être mise à part, — b) l'oreille, et l'oreille seule, reste seul juge de la qualité des voyelles, et prend le pas sur toute machine. Il s'agit donc d'une délimitation psycho-acoustique des voyelles.

Certaines expériences sur le synthétiseur «playback» (à roue phonique) des Laboratoires HASKINS semblaient confirmer à première vue la justesse des mesures, malgré certaines différences de fréquence pouvant aller jusqu'à 100 Hz (DELATTRE 1965, p. 49).

Cependant ici, on note une nouvelle mise en garde (*ibid.*, p. 48) : «Ces fréquences de formant ne produiront de voyelles synthétiques satisfaisantes que pour un fondamental standard de 120 Hz» (la seule fréquence fondamentale que ce synthétiseur puisse d'ailleurs fournir). On note également que ces voyelles synthétiques sont obtenues avec deux formants d'égale amplitude (facilement réglable sur ce type de synthétiseur), à l'exception de [u] dont le deuxième formant est à — 6 dB, et surtout des nasales qu'il est impossible de synthétiser à moins d'utiliser au moins trois formants *d'inégale amplitude*. (Les deux premiers FN et F1, de fréquence constante, ont respectivement : FN - 12 dB, et F1 - 6 dB, ce qui confirme la thèse de HUSSON : 1957, p. 251).

Ces considérations montrent déjà l'importance de l'amplitude relative des formants, qu'on avait toujours niée jusqu'alors.

L'interaction sensorielle des fréquences de formants et de leurs amplitudes relatives.

Bien que le fait semble ignoré de la majorité des phonéticiens, il semble que ESSNER, qui avait collaboré avec notre regretté ami J. CANTINEAU à l'ancien Laboratoire de Phonétique de l'Université d'Alger, ait été l'un des très rares auteurs (après HELMOLTZ lui-même...).

a avoir mis l'accent sur l'importance de l'amplitude relative des formants. Il dit (1947, p. 55) : «Le deuxième facteur définissant le timbre vocalique est le rapport des amplitudes de formants» — et il ajoute (*ibid.*) : «...on peut citer dans le tableau de BOUMAN 6) que certains timbres vocaliques ne diffèrent que par les intensités ».

Nous avons nous-mêmes considéré que les apparentes anomalies dans les fréquences de formants, variables d'un individu à l'autre, pouvaient être mises sur le compte de compensations formantiques, dues aux divergences d'intensité relative des formants (EMERIT 1966, pp. 212-216, et 1967, p. 211).

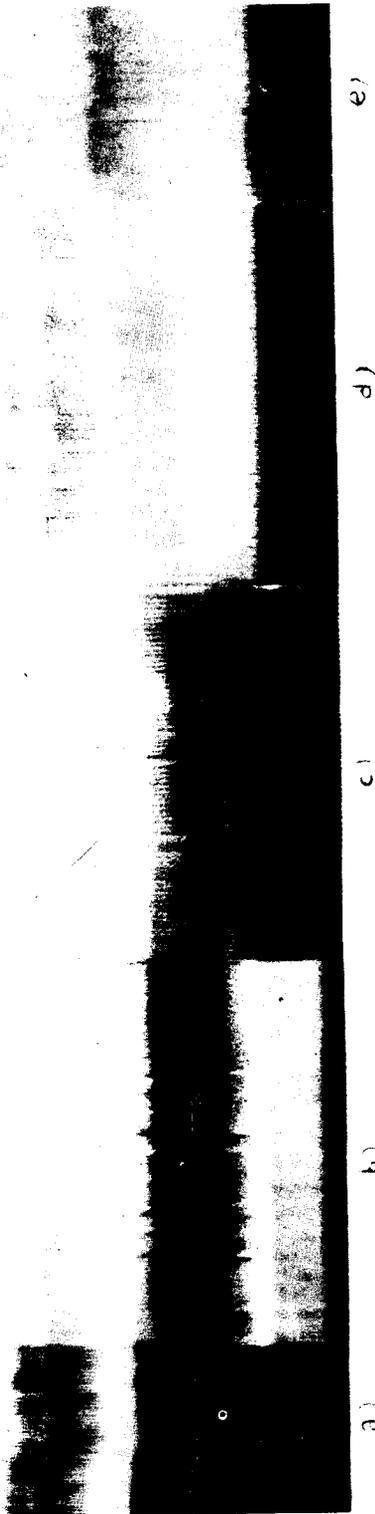
Les expériences de FLANAGAN (1957) montrent effectivement que les amplitudes relatives des formants peuvent avoir un rôle considérable à jouer dans la perception du timbre.

Le «Phonétographe IV» de DREYFUS-GRAF (1965, p. 2) tient compte des intensités relatives de formants, simplifiées de 100 à 3 «bits».

D'autres expériences très récentes, sont d'autant plus intéressantes qu'elles confirment notre théorie de trichromie auditive (cf : infra ; le timbre dépend à la fois des fréquences de formants et de leurs intensités relatives, et il y a étroite interdépendance entre ces deux facteurs). Ces expériences, menées et relatées par LINDNER (1970) ont consisté à synthétiser des voyelles orales au moyen de trois formants maintenus à des fréquences rigoureusement fixes ; on s'est contenté de faire varier à des degrés divers les amplitudes de deuxième et troisième formant. Les tableaux présentés montrent que deux combinaisons fréquentielles seulement suffisent, en jouant uniquement sur les amplitudes relatives, à reconstituer presque toutes les teintes vocaliques sans aucunement jouer sur les fréquences de formant. Il n'est guère douteux que des expériences plus systématiques menées avec trois ou quatre formants de fréquences judicieusement choisies, ne donnent des résultats encore plus probants.

Nous avons nous-même observé et publié (EMERIT 1972) entre autres observations, que la fréquence moyenne du bloc F2-F3 de [y] était notablement plus élevée en voix naturelle qu'en voix synthétique (contrôlé par sonagramme de deux bandes magnétiques collées, ce qui exclut les erreurs de calage de fréquences) et nous avons mis cette «compensation formantique» sur le fait d'un F1 d'amplitude notablement plus élevée en voix naturelle qu'en voix synthétique (FIG. 1, d-e).

6) — Cf : bibl. (1926). — Toujours à propos des amplitudes de formants, on peut penser qu'elles ont eu leur rôle à jouer dans la «vérification» par synthèse effectuée par DELATTRE (1965, p. 49), de son «triangle vocalique», l'absence totale du troisième formant ayant pu être compensée par un véritable doublage de l'amplitude du deuxième.



Deux exemples significatifs de compensation formantique : cinq séquences vocales ont été enregistrées au magnétophone, et les morceaux de bande collés bout à bout (pour être absolument sûr que les fréquences du «Sogramph» sont bien alignées; ces collages provoquent ici un léger pleurage, d'ailleurs peu gênant).

a) et b) : A légèrement postérieur en voix synthétique monotone sur EVA III. En a) $F1 = F2 = 750 = 1.250$ Hz. En b), $F1 = F2$ ont convergé jusqu'à former un bloc unique à 1.050 Hz. La couleur sonore de ces deux séquences (contrôlée rigoureusement par scintimétrie acoustique) est *absolument la même*. Une telle convergence n'est permise que sur une «frontière chromatique» (1.050 Hz, cf : infra). L'identité avec la synthèse des couleurs visuelles est frappante (ici, analogie avec le jaune). En c) : A naturel plus postérieur.

d) : (Y) («vues») naturel. $F1 = 250$ Hz, $F2 = F3 = 1900 = 2100$ Hz. — e) : (Y) synthétique : $F2 = F3 = 1700 = 1900$ Hz, la *seule combinaison* donnant cette fragile couleur sonore sur EVA III (tout comme en synthèse de couleur visuelle, le vert-jaune pur ne peut être obtenu — en synthèse additive — par aucune combinaison de teintes fondamentales autres que sa propre présence, et à l'exclusion des autres teintes). Dans le cas du (Y) naturel, l'élévation des formants peut être mise sur le compte d'une compensation formantique rendue nécessaire par l'intensité de $F1$ beaucoup plus élevée qu'en voix synthétique. (de : EMERIT, *phonetica* 1972).

HUSSON met également l'accent sur les divergences de fréquences de formants *chez un même individu* en voix chantée et parlée (1958, p. 341) : « Du chant à la parole, pour une voyelle donnée, la différence peut atteindre 900 c/s (exemple : le formant buccal de [i] peut passer de 1.800 c/s dans le chant à 2.700 c/s en voix parlée). Ces différences apparaissent nettement lorsqu'on parcourt les résultats collationnés par B. HALA dans son récent travail sur les voyelles (1956) 7) ».

Comme il est désormais évident qu'une amplitude formantique accrue à une extrémité du spectre doit être compensée par une déviation en sens inverse (pour une même voyelle) de la fréquence des formants situés à l'autre extrémité, il est également évident, dans ce cas, qu'une fourniture laryngée de haute fréquence et de plus grande amplitude en voix chantée, tend à faire *baisser* le bloc formantique F2-F3 de la voyelle [i], et c'est encore un exemple de « compensation formantique ». Il en va de même dans les couleurs : si une image se trouve, pour une raison ou une autre, plus chargée en violet, par exemple, la teinte globale bleue pourrait être restituée en *abaissant* la fréquence moyenne des composantes bleues (analogues au bloc F2-F3), c'est à dire en les faisant virer au vert. C'est l'une des multiples analogies qui nous ont d'ailleurs servi à étayer nos théories (cf : infra).

FISCHER-JØRGENSEN (1958, pp. 447-448) confirme ce point de vue sur la foi des expériences menées aux Laboratoires HAS KINS : « Des expériences avec des voyelles artificielles ont montré que deux formants relativement rapprochés ont le même effet sur l'oreille qu'un seul formant se trouvant entre les deux, c'est-à-dire qu'un seul formant se trouvant *au milieu* si les deux formants sont *d'égale intensité*, ou *plus près du formant le plus fort*, et il existe un second accord très sensible entre les résultats obtenus par différentes méthodes ».

Cette autre remarque (*ibid.*, p. 448) est encore plus importante en ce sens qu'elle contribue encore à affaiblir la trop simpliste théorie de la *perception* vocalique prétendument liée à la seule fréquence des formants compte non tenu de leurs amplitudes relatives : « Il apparaît clairement aussi d'après les expériences HASKINS, que lorsque seuls deux formants sont utilisés pour faire un [i] synthétique, F2 doit être placé *considérablement plus haut que pour un [i] naturel*, de façon à correspondre à une *position intermédiaire* entre les F2 et F3 du [i] naturel ».

De tels arguments qui montrent l'étroite *interaction* dans la sensation, entre les fréquences des formants et leurs amplitudes relatives, suffiraient à eux seuls à confirmer notre thèse d'une véritable perception « colorée » du son suivant un principe de « polychromie » analogue à celui de la vision des couleurs.

Ce point de vue s'est trouvé encore confirmé par nos propres expériences menées ici-même sur synthétiseur EVA III (EMERIT 1972) sur les « convergences de formants » (FIG. 1 a-b) : si on fait monter le formant F1 d'un [a], il faut faire *descendre* son formant F2 dans une certaine proportion précise, et sans qu'il y ait le moindre changement de vocalité, l'ex-

7) Cf. : Bibl. HALA 1956.

périence peut être menée à bien jusqu'à fusion complète de F1-F2 à 1.050 Hz. D'autre part de telles expériences ne sont possibles que de part et d'autre de deux « charnières vocaliques » : 1.050 Hz et 6.500 Hz, fréquences que nous avons prévues par simple raisonnement (1966, p. 205).

Qu'on accepte ou non cette explication des « anomalies formantiques », ces dernières n'en existent pas moins, et sont la cause majeure de l'inefficacité des « machines à reconnaître le langage ». MOLES qui nous dit plaisamment (1966, p. 236) qu'il regretterait « personnellement » l'apparition de cette « sténo-dactylo automatique », doit pouvoir être rassuré à ce sujet, car il faudra encore très longtemps avant que l'on vienne à bout des « divergences formantiques ».

Ces divergences ont été quantifiées par POTTER et STEINBERG (1950, p. 816) : 8) utilisant le principe du triangle vocalique de DELATTRE, ils ont délimité les « aires formantiques » des voyelles orales anglaises. Pour un même individu, les aires formantiques ainsi déterminées étaient presque punctiformes, et très éloignées les unes des autres, ce qui explique que les machines à reconnaître le langage puissent reconnaître « la voix de leur maître » avec une fiabilité de 97 à 99 pour cent 9). (Fig. 2.)

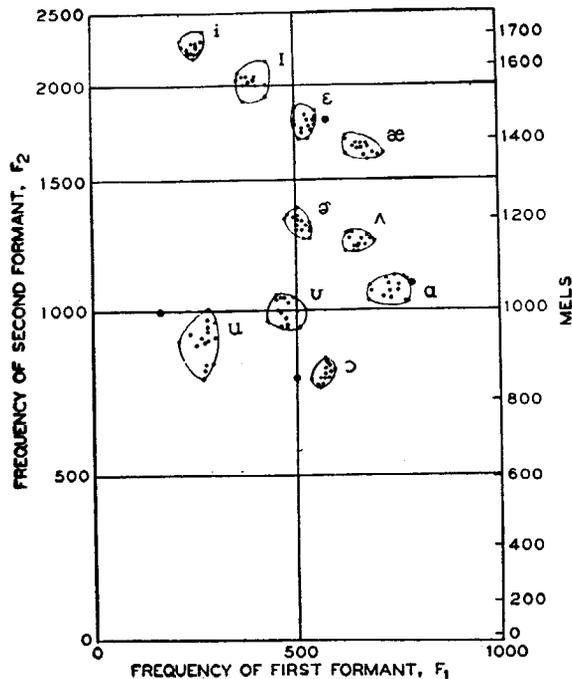


FIG. 10. First and second formant positions. Repeated utterance Fig. 2. by speaker A.

8) — Mesures confirmées par PETERSON et BARNEY — bibl : 1952.

9) DEWEZE (1966, p. 153) cite la machine de Davis et alt. qui possède un lexique total de 9 chiffres, celle de Fry et Denes, celle d'Olson et Belar, et celle de Dreyfus-Graf, en mettant toujours l'accent sur la fiabilité beaucoup moindre de ces dispositifs en présence d'une autre voix que « celle de leur maître », à l'exception toutefois de celle de Dreyfus-Graf dont les « capteurs » sont les plus perfectionnés et tendent vers une véritable simulation de la perception auditive. Le lexique de cette dernière n'est cependant que de 18 mots (cf : LORAND, 1971, p. 71).

Cependant ces mêmes mesures de formants effectuées entre différents individus montrent que les fréquences des deux premiers formants — pour ne parler que de ceux-là, peuvent varier du simple au double (FIG. 3). Les aires vocaliques ne sont plus alors punctiformes,

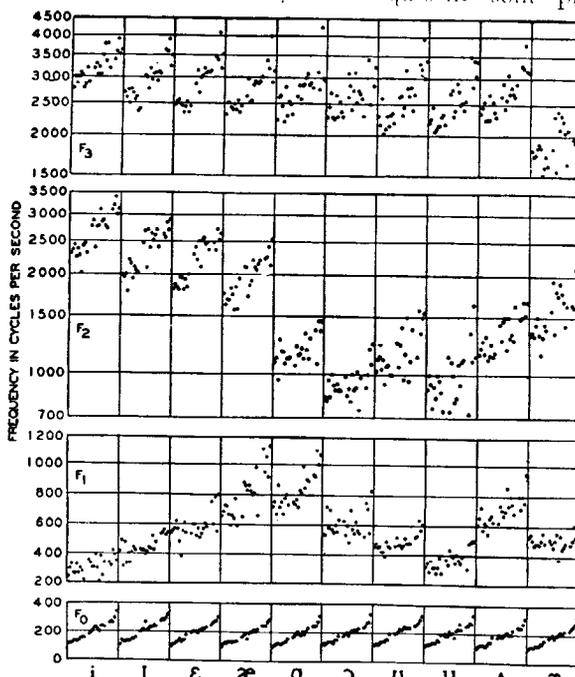


Fig. 3.

mais se présentent sous forme de très larges taches dont les bords se touchent et parfois même se recouvrent. (FIG. 4) Si on considère ces aires comme une possible délimitation du pho-

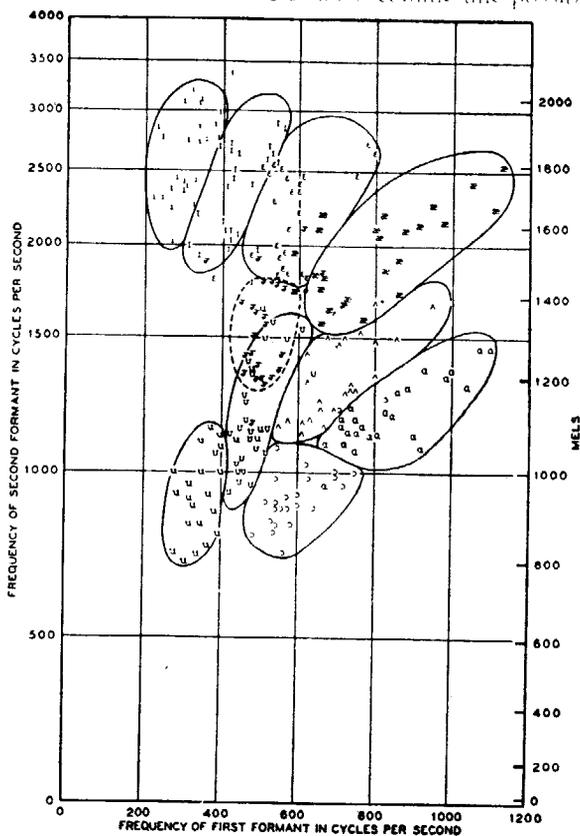


Fig. 4.

nème, il ne peut s'agir que d'une première approximation, car il est évident que des combinaisons formantiques *très différentes* peuvent fournir à l'oreille des teintes sonores *presque semblables*. Toujours est-il que si l'on fournit à une « dactylo automatique » une combinaison formantique telle qu'elle formerait un point à l'intersection de deux aires formantiques, la machine n'aura aucun moyen de se « décider » à ranger le phonème en question dans l'une ou l'autre catégorie. et la catégorisation se fera totalement au hasard au profit de l'une ou l'autre voyelle. En fait, en présence d'une voix « inconnue », la majorité de ces machines n'ont plus qu'une fiabilité de 45 à 60 pour cent (DEWEZE, 1956, p. 153).

Les divergences formantiques rendent en core plus ardue la « lecture » des sonagrammes, même étalonnés en fréquences, dont FANT dit par ailleurs (1968, p. 241-242) qu'il ne faut pas moins de *seize* opérations mentales pour commencer à les déchiffrer... Ces mêmes divergences formantiques devaient amener DELATTRE à réviser sa position première trop catégorique. Il dit (1965, p. 52) : « Des linguistes de nombreux pays ont tenté de mesurer les formants 1 et 2 sur des spectogrammes et ont publié leurs résultats. Quand nous vérifions ces résultats par synthèse, ils ne résistent pas à l'examen... par la faute d'éléments humains encore partiellement inexpliqués — les formants d'une femme peuvent être d'environ 15 pour cent plus élevés que ceux d'un homme parlant le même dialecte, et même des personnes du même sexe peuvent montrer des divergences considérables dans les fréquences de formants pour un son similaire ».

En faisant appel à la parole synthétique pour obtenir un équivalent mesurable de la perception auditive, DELATTRE a peut-être été le promoteur d'une véritable « colorimétrie acoustique » (terme inauguré par nous-même, cf : EMERIT 1967). En effet, de même que la colorimétrie visuelle ne peut être définie que grâce à des filtres colorés de caractéristiques parfaitement déterminées en ce qui concerne les longueurs d'onde admises, la largeur de bande passante, l'intensité lumineuse restituée par chacun d'eux, et en définitive l'impression visuelle obtenue par rapport à une couleur type, de même une colorimétrie acoustique *ne peut être valable que si elle s'appuie sur les résultats d'un synthétiseur de type particulier* qui fournirait à l'oreille, compte non tenu de la véritable structure acoustique du son, *une impression vocalique que l'on pourrait alors comparer par « scintimétrie » acoustique à une voyelle véritable*, jusqu'à éta- blissement d'une identité psycho-acoustique optimale.

En fait, non seulement les voyelles synthétiques des Laboratoires HASKINS ont permis une première standardisation de la valeur psycho-acoustique des phonèmes, mais c'est en core la synthèse qui permet de dissocier et de classer par ordre d'importance les « indices acoustiques de la parole » (cf : DELATTRE, 1958), de démontrer le peu de cas que l'on doit faire de l'intensité réelle des syllabes accentuées que l'on distingue et classe comme telles beaucoup plus par le jeu du rythme et de l'intonation que par l'intensité elle-même (le plus schématisé des paramètres en parole synthétique 10), de connaître les mécanismes accessoires de différenciation acoustique sourde/sonore, intimement liés à la durée du phonème (DELATTRE, 1965, pp.

10 — La méthode *expérimentale* nécessairement suivie par les constructeurs de notre synthétiseur de parole EVA MARK III (cf : bibl.) montre clairement le peu de cas que l'on doit faire d'une mesure *exacte* de l'amplitude des syllabes accentuées. On trouve sur l'« Instruction Manual » (ibid. : additional leaflet 2, sans pagination) les précisions suivantes : « Les niveaux de «AO» (amplitude de la syllabe synthétique) «ont été réglés suivant la perception auditive initiale de la syllabe accentuée. La pleine puissance a été réservée aux accents principaux (contre seulement 1/2 ou 1/4 pour les syllabes inaccentuées). Cela s'est avéré parfaitement adéquat et aucun changement n'a été nécessaire. Cependant la *durée* de syllabe et ses *contours intonatoires* peuvent être les principaux éléments déterminants de l'accent d'intensité, de sorte que nous ne sommes pas sûrs de l'importance des variations de AO ». En fait, nous nous sommes aperçu dès les premiers essais, que l'amplitude des bruits consonnantiques étant trop faible, une bonne intelligibilité n'était obtenue qu'en ramenant toutes les voyelles (mêmes inaccentuées) au seul niveau 1/2, soit — 20 dB, ou encore 1/10 de la pression sonore (qui est grossièrement proportionnelle à la sonie, — cf : STE VENS, 1936). Le niveau 1/4 (— 34 dB, soit 1/50 de la pression sonore) n'intervient que pour la réalisation des consonnes sonores et de [l], que les études de ZINKIN (1968, p. 238) montrent comme n'affectant pratiquement pas la courbe d'intensité des syllabes...

114-115), les mécanismes de l'attention, de l'anticipation linguistique et de la divination en présence de bruit ou de lacunes informatives, la valeur linguistique des invariants de l'intonation (11). C'est elle qui donnerait la possibilité de créer un lexique de logatomes typiques d'une langue donnée (valable pour un seul type de syntétiseur, projet à l'étude). C'est elle enfin qui nous a permis d'effectuer une vérification supplémentaire de notre conception d'une audition «colorée» du son : les teintes sonores «pures» [u] : F1-F2 groupés autour de 450 Hz, F3 pratiquement inaudible, — et [y] : F1 presque inaudible et ne participant pas à la teinte d'ensemble du fait de sa très basse fréquence 220 Hz, et F2-F3 groupés sur 1.700-1.900 Hz) ne souffrent aucune anamorphose formantique notable, alors que les voyelles médianes ([ø], [œ], [ʌ], etc.... voyelles à trois formants principaux très écartés) supportent des anamorphoses formantiques considérables. Ces dernières sont en effet comparables à des couleurs très désaturées, et de ce fait peu fragiles, ce qui confirme une fois de plus notre principe de trichromie auditive (FIG. 5, de : EMERIT 1972).

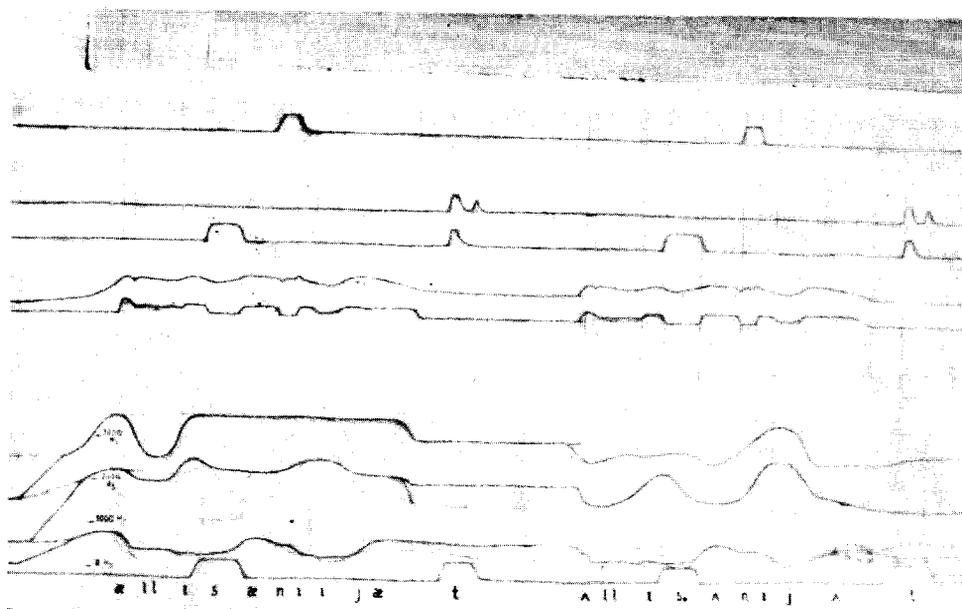


Fig. 5.

Toute modification de ces valeurs de base ne donne que des résultats négligeables ce qui laisse à penser le peu de profit que l'on peut tirer des multiples études sur les niveaux précis des syllabes. Aussi pensons-nous adopter sur notre tonographe à indication de tonalité SATREM I encore à l'étude, une indication d'amplitude fournie, non plus sur un cathoscope séparé comme c'est le cas sur la majorité des appareils de ce genre, mais sur un simple effecteur lumineux punctiforme, beaucoup moins encombrant sur les clichés. Cet appareil, dont une maquette a été présentée en fonctionnement à la Sorbonne (EMERIT 1970), et qui fera, une fois complété, l'objet d'une publication séparée, fonctionne sur le principe de l'«image sonore» de GRÜTZMACHER (1965), et de l'«enveloppe tonale» de GSELL et al. (1963). Il doit être complété par un canal d'amplitude globale, et un canal triple «de tonalité» (ou «chrominance» simulation instrumentale de la «vocalité») réalisant un dispositif de mesure directe des «formes de tonalité» de JAKOBSON (1971, pp. 484-486). Un dispositif spécial de projection (cf : EMERIT 1970) doit également permettre de refléter directement la vocalité des sons.

FIG. 5 : Evagramme dessiné par l'auteur, de deux versions du mot arabe «Al Lisaniyyat» («La Linguistique») celle de gauche exécutée sans sonagramme à partir de simples barèmes expérimentaux, celle de droite à partir d'un sonagramme dûment rectifié (les rectifications ayant porté sur le premier formant, en fin de séquence).

Tous les auditeurs, arabophones ou non, identifient phonétiquement les deux versions sans hésitation, mais tous s'accordent à reconnaître à celle de droite un caractère plus «naturel» et plus «humain». Le fait que des divergences formantiques considérables n'affectent que peu la qualité sonore d'une voyelle médiane (qui passe ici de /æ/ à une sorte de /œ/ sans pour autant altérer la compréhension du mot) est en soi une constatation intéressante : les voyelles à formants écartés sont peu sensibles aux anamorphoses formantiques. Il n'en va pas de même des voyelles plus «monochromes», telles que /y/ («rue»); cette voyelle porte en effet en voix synthétique les formants F2-F3 très groupés entre 1700 et 2000 Hz (pour environ 2000-2200 Hz en voix naturelle où l'élévation de ces formants doit être mise sur le compte d'un déséquilibre provoqué par le formant F1 d'amplitude beaucoup plus grande en voix naturelle qu'en voix synthétique). Or, en voix synthétique, la voyelle /y/ ne supporte pas la moindre anamorphose formantique, la fréquence des formants naturels devant être indiscutablement baissée de 200 à 300 Hz pour reproduire cette fragile couleur en voix synthétique.

Cet exemple parmi tant d'autres, qui nous portent à assimiler réellement le son à la couleur où de telles corrections sont courantes, vient une fois de plus à l'appui de notre théorie de trichromie auditive (la couleur du son est restituée au niveau du cerveau par le dosage de trois teintes sonores fondamentales analogues à celles de la vision colorée).

Deux paramètres essentiels du «timbre» «texture» et «vocalité».

a) La «texture».

Ce qu'on appelle le «timbre», notion vague et non mesurable, semble en fait tenir à deux facteurs essentiels : la «texture» et la «vocalité» du son.

La «texture» est la perception directe des basses fréquences du son par le cerveau, notion en tous points conforme aux enseignements de la théorie de WEVER et BRAY (WEVER 1949) et l'importance de cette perception directe est mise en relief par plusieurs considérations également dignes de confiance :

DELATTRE (1965, pp. 49-50) a mis l'accent sur l'importance, en synthèse des voyelles nasales, d'un formant bas de très faible amplitude (250 Hz. — 12 dB) et d'un formant affaibli à 600 Hz (- 6 dB).

SKALA (1948, 1949) a pu démontrer à l'aide de son synthétiseur musical « Mixtur Trautonium », que certains *sous-harmoniques* ont une influence très caractéristique sur le timbre.

Ces mêmes sous-harmoniques amplifiés par le jeu de filtres SUVAG II avaient été mis à profit par GUBERINA pour sa méthode de rééducation des sourds profonds (relatée par GIRON, 1964).

— Toujours à propos de l'importance de l'«intensité» dans l'accent dit d'«intensité», en particulier en français, la plupart des auteurs s'accordent pour la minimiser au profit des facteurs hauteur et durée. Notre regretté Maître Pierre FOUCHÉ disait à ce propos (1951, p. 99) : «La fin de chaque groupe se reconnaîtra non seulement à l'intensité de la dernière syllabe, mais encore à la note aiguë sur laquelle elle est intonnée. Où seront les accents les plus forts ? D'une manière générale, on peut dire qu'ils coïncident avec les sommets de plus grande hauteur».

DELATTRE (1951, p. 45) insiste davantage sur l'importance de la durée de la syllabe accentuée, en ces termes : «En français il suffit d'augmenter la durée d'une syllabe pour donner l'impression d'accent; et le seul élément physique qui soit toujours en excès est la durée... L'intensité n'est pas plus forte dans la syllabe accentuée que dans les autres, elle est plutôt légèrement moindre. La hauteur fait généralement un écart considérablement plus grand entre la syllabe accentuée et la précédente qu'entre les syllabes inaccentuées, mais cet écart n'a pas de sens fixe : il peut aller vers le grave comme vers l'aigu. De plus, il n'est pas indispensable pour transmettre l'impression d'accent».

Ce sont encore ces sous-harmoniques, ou plus exactement les «subformants» qui ne sont autres que les enveloppes d'intensité des attaques sonores, qui ont été exploités, avec le succès que l'on connaît, par DREYFUS-GRAF dès avant 1950 (1948, 1950 a et b, et sqq.) pour permettre la reconnaissance des consonnes par ses «phonétographes».

D'autre part WINCKEL (1960, pp. 39-56) accorde une place prépondérante à la perception des transitoires, en particulier des attaques musicales, dans l'identification des sons.

Quant aux basses fréquences proprement dites, surtout celles inférieures à 600 Hz. nos propres expériences sur EVA III nous ont permis de constater leur rôle considérable dans la formation des voyelles synthétiques. Alors que le premier formant est très difficile à décalquer de façon précise sur un sonagramme (à cause de l'échelle linéaire de fréquences du «sonagraph» qui «écrase» les variations de basse fréquence), le moindre changement de fréquence de ce premier formant peut suffire pour passer d'une voyelle à l'autre toute différente, sans qu'on ait eu besoin de retoucher tant soit peu certaines valeurs *maintenues fixes* des formants 2 et 3 dits souvent «informatifs» ou encore «linguistiques»... L'extrême sensibilité de l'ouïe aux variations de basses fréquences, mesurée par FLANAGAN (1955) — alors que les points focaux d'octave sont les plus rapprochés sur la membrane basilaire, qui possède alors de surcroît la sélectivité la plus médiocre — semble confirmer le principe wévérien de la perception directe des fluctuations d'intensité (ou «texture» du son). Ce principe suffirait à expliquer pourquoi deux sons possédant exactement la même «vocalité» (ou «couleur» sonore) sont nettement différenciés par l'oreille, tout en nous amenant à considérer cette deuxième caractéristique importante du «timbre», qui est la «vocalité».

b) La «vocalité» (ou «couleur» vocalique.)

On sait quel rôle important la synthèse des sons joue depuis près d'un siècle en psycho-acoustique. Elle seule permet de soumettre à l'oreille différentes combinaisons sonores qui la satisfont plus ou moins. Elle aura également permis de dégager la notion de «vocalité».

HUSSON (1957, p. 250) définit ainsi la *vocalité* :

«On appelle voyelle tout son produit par un organe vocal humain dans les conditions définies plus haut : fourniture laryngée modifiée dans ses amplitudes composantes par son passage à travers le pavillon pharyngo-buccal. On appelle d'autre part «vocalité» d'un son sa *parenté acoustique* (plus ou moins grande) avec une quelconque voyelle humaine. Cette distinction est fondamentale. Pour avoir omis de la faire, nous verrons que nombre de phonéticiens ont été conduits à de grossières erreurs. C'est ainsi que certains ont cru faire la *synthèse des voyel-*

S'associant aux vues de P. FOUCHÉ, RIGAULT insiste tout particulièrement sur l'importance de la hauteur de la syllabe, mise en relief par des expériences de voix synthétique, en ces termes (1961, p. 735) : «Nous présentons le résultat de recherches sur le rôle joué par la fréquence, l'intensité et la durée vocaliques dans la perception de l'accent en français. Nous avons fabriqué des mots et des phrases synthétiques, le même mot ou la même phrase étant produits à plusieurs exemplaires, chaque exemplaire ne présentant que la modification d'une seule variable. Les enregistrements de ces mots synthétiques ont ensuite été joués devant des sujets de langue française qui ont consigné par écrit leurs réactions. Les résultats montrent que la fréquence est, de loin, le facteur physique prédominant de l'accent. Ils nous paraissent remettre en question la définition du soi-disant «accent d'intensité» du français, et nous amènent à proposer une révision de la terminologie relative aux faits accentuels».

11 — GSELL et al. (1963) ont réalisé d'un des rares «détecteurs de mélodie» capables d'afficher les courbes mélodiques sur échelle logarithmique à base 2 (demi-tons de la gamme tempérée régulièrement espacés), réalisant ainsi une véritable simulation de la perception de hauteur, seule capable, selon GSELL (*ibid.*), de permettre de dégager les «invariants» de l'intonation.

les, alors qu'ils ne donnaient naissance qu'à des sons pourvus d'une certaine *vocalité*... En choisissant convenablement les régions tonales à associer, ainsi que les *intensités relatives* des deux sons prédominants, on parvient à reproduire de façon convenable toutes les vocalités caractéristiques de la voix humaine... Pour qu'un son fût pourvu d'une vocalité donnée, il était nécessaire et suffisant qu'il contînt deux partiels suffisamment intenses, appartenant respectivement aux deux régions formantiques... correspondant à chaque voyelle, ou mieux à chaque *couleur vocalique* ou *vocalité particulière*».

Après avoir énuméré les méthodes très diverses de synthèse vocalique qui ont été employées par différents chercheurs depuis le début du siècle. HUSSON remarque (*ibid.*, p. 251) : «Au fond, ce qu'on avait fait dans chaque cas : *n'était pas du tout la «synthèse des voyelles»,* mais tout simplement *la «synthèse des vocalités»,* problème à la fois beaucoup plus simple et beaucoup plus général ».

Puis il remarque (*ibid.*) que : «le scieur de bois fabrique des [i] fort convenables, et, en frappant du doigt à la porte, on reproduit la *vo calité* de [o] de «toctoc»... «etc...»

Ces citations n'avaient pour objet que de souligner l'importance de la *vocalité* (ou «*couleur*» des sons) dont la part est prépondérante dans l'intelligibilité des phonèmes. *Elle pourrait à elle seule servir de critère dans une délimitation psycho-acoustique des phonèmes,* grâce à des expériences de scintimétrie acoustique, menées à l'aide d'un synthétiseur de type parfaitement déterminé, qui devrait servir de témoin, *l'oreille restant,* comme nous l'avons dit après GRAMMONT, *seul juge de la similarité psycho-acoustique* d'une voyelle naturelle et d'une voyelle synthétique typique. On serait parfois étonné de la diversité formantique de sons strictement équivalents du point de vue de la seule vocalité, mais n'est-ce pas le but suprême des recherches que nous poursuivons inlassablement depuis plus de dix ans, que de trouver le moyen précis d'établir pour la première fois dans l'étude des sons de la parole une *normalisation* équivalente à celle que l'on a adoptée pour la caractérisation des *couleurs* telles que vue par l'œil humain, et compte non tenu de leur véritable structure physique ?

Une idée ancienne : la caractérisation de la vocalité par référence à la couleur : métaphore ou réalité ?

La référence à la couleur, qui était à la base de notre théorie de la perception auditive (cf : infra) est en fait une idée fort ancienne, puisqu'on trouve déjà cette analogie (fausse car appliquée uniquement aux mélodies) mentionnée par J.J. ROUSSEAU (1817, p. 535). Actuellement la qualité sonore globale, ou vocalité du son, est universellement appelée «couleur» vocalique par l'immense majorité des phonéticiens. Nous ne nous étendrons pas ici sur la réalité qui reste soumise à l'acceptation d'une théorie entièrement nouvelle. Disons seulement que tout, actuellement, porte à se représenter une voyelle isolée (pour prendre le cas le plus simple à l'exclusion des logatomes et syllabes) comme un véritable «objet» sonore. Comme tout objet, elle possède une forme (texture, structure formantique) et une *couleur* (vocalité provenant d'un certain mélange équilibré de composantes de diverses fréquences). Comme en acoustique, la «forme» influe sur la «couleur», et vice-versa, il conviendrait de tenir compte de ces deux facteurs et de leur interaction.

Pour prendre un exemple concret, nous supposons qu'une première délimitation d'une voyelle puisse être fournie, pour ce qui est de sa «forme», par l'une des innombrables combinaisons formantiques F1-F2, qui permettraient de la classer, en première approximation, dans l'une des «aires vocaliques» telles que délimitées par POTTER et STEINGERG (1950, p. 817) sur le triangle vocalique de DELATTRE.

Cette première délimitation étant, comme nous l'avons vu, insuffisante, il faudrait encore noter (sur synthétiseur du genre «playback» des Laboratoires Haskins, à l'exclusion des synthétiseurs paramétriques) de combien les amplitudes F1-F2 devraient être modifiées pour

chaque combinaison fréquentielle représentée par un point différemment placé dans l'enceinte de chaque aire vocalique, pour reproduire chaque fois une stricte équivalence auditive avec le son de référence dont on serait parti.

Il s'agit ici que d'une suggestion ; nous ne pensons pas que pareil travail, qui demanderait d'une part l'utilisation d'un synthétiseur «playback», et de l'autre la patience et la minutie d'un FLETCHER ou d'un MUNSON, soit réalisé de sitôt. Une telle synthèse serait strictement équivalente à une synthèse de couleurs visuelles pour laquelle il serait possible de faire varier à la fois les fréquences et les amplitudes des composantes. L'intérêt d'une telle méthode de délimitation serait de ne plus reposer seulement sur trois facteurs (les fréquences des trois formants), mais sur six (les fréquences et les amplitudes relatives des trois premiers formants). La redondance des indices acoustiques serait de ce fait considérablement augmentée, et de nature à faciliter l'identification des sons de la parole par des machines.

Une telle méthode de délimitation de la voyelle (extensible dans une moindre mesure aux consonnes) n'est cependant concevable que si l'on adopte une attitude nouvelle vis-à-vis du parallèle son/couleur, c'est à dire si l'on ne considère plus seulement qu'il s'agit d'une *métaphore*, mais d'une *réalité* comme nous en avons émis l'hypothèse (EMERIT, 1966), hypothèse vérifiée numériquement par la suite (EMERIT, 1967, 1970, 1972).

Dès 1967, le psycholinguiste H.L. LANE 12) (LANE H.L., 1971, p. 108. thème repris de : 1967, p. 113 et 1968) s'inscrivait dans cette optique nouvelle, en ces termes : «Voici une description des résultats obtenus pour la parole, extraits d'un article sur la «théorie de la perception de la parole» 13) : » Bien que les « stop-consonnes » reposent sur un « continu » acoustique, la perception est essentiellement discontinue. C'est à cause des crêtes dans la fonction de discrimination aux limites des classes phonémiques que les sons sont catégoriquement entendus... et par conséquent classés rapidement avec exactitude en catégories phonémiques appropriées». Pour démontrer que cette description de la perception catégorique n'est pas caractéristique de la parole seule, il faut en venir aux résultats concernant la *couleur* en substituant aux termes de la *parole* dans la citation, les termes de la *couleur* : «bien que les (couleurs) reposent sur un « continu » (visuel), la perception est essentiellement discontinue. C'est à cause des crêtes dans la fonction de discrimination aux limites des classes (de couleur) que les (couleurs) sont catégoriquement (vues)... et par conséquent classées rapidement et avec exactitude en catégories (chromatiques) appropriées ». — La description des données *s'applique aussi bien à la couleur qu'à la parole*. La discrimination... se fait mieux en effet à travers les limites chromatiques qu'à l'intérieur des catégories de couleurs».

Une telle citation parle d'elle-même et montre qu'on ne peut faire aucune distinction notable entre la catégorisation des couleurs et celle des vocalités. Nous irons plus loin en établissant un parallèle on ne peut plus strict entre les habitudes visuelles de certains peuples et les catégories phonémiques de leur langue.

On sait que les catégories phonologiques dans lesquelles un auditeur classe les phonèmes ne sont pas les mêmes pour toutes les communautés linguistiques, et dépendent du système phonémique adopté par ces communautés.

FISCHER-JØRGENSEN nous dit à ce sujet (1958, p. 443) : «La perception est influencée par le système phonémique de la langue maternelle et par l'orthographe... Il est possible, par exemple, de s'attendre à ce que les auditeurs français trouveraient [e] et [ɛ] très semblables, parce que dans leur langue, l'opposition est souvent neutralisée... ».

12 — Ne pas confondre avec l'audiologue C.E. LANE, auteur, entre les deux guerres, de remarquables travaux sur l'effet de masque, Cf. Bibl. : WEGEL R.L. and LANE C.E., 1926 (1924).

13) — De : LIBERMAN et al., op. cit., 1963, p. 3.

Il peut être intéressant de signaler à ce sujet les transformations phonétiques instantanées subies par suite d'une nécessaire adaptation à un système phonémique particulièrement simple, par les mots anglais passés au tahitien. Le système phonémique du tahitien en comporte en effet que treize phonèmes, cinq voyelles : /a/ — /e/ — /i/ — /o/ — /u/, et huit consonnes : /f/ — /h/ — /m/ — /n/ — /p/ (légèrement voisé) — /r/ (ressemblant à une sorte de l) — /t/ — /v/. La seule opposition phonologique sourde / sonore est la paire f/v, et les géminées et affriquées sont inexistantes, de même que /s/ (remplacé par son équivalent le plus proche : /t/). C'est ce qui explique la transformation des mots anglais, dont voici quelques exemples significatifs : february : fepuare — march : mati — april : eperera — july : tiuraï — august : atete — september : titema — october : atopa — november : noema — december : titema — hundred : hanere — thousand : tauatini — million : mirioni, — etc...

On retrouve un *strict* équivalent de cette distinction psycho-acoustique des phonèmes, telle que soumise au système *phonémique* d'une langue donnée, dans la distinction des couleurs, telle que soumise au système *symbolique* d'une langue donnée (système qui influe sans doute sur la perception elle-même). FRANCES nous dit à ce sujet (1962, p. 158) : «Le vocabulaire des noms de couleurs, par sa richesse ou sa pauvreté et par la signification que les termes ont dans une langue ou une autre, est un témoin de la diversité des *attitudes* que chaque culture impose à l'homme devant la diversité des couleurs et nuances de la perception colorée. On cite souvent les observations de Margaret Mead sur les Indiens d'Amérique pour lesquels des couleurs bien distinctives pour nous telles que le jaune, le vert olive, le bleu vert, et le gris lavande, sont des nuances d'une même teinte».

Conclusion : les applications possibles d'une meilleure délimitation du phonème.

Nous n'avions pas ici d'autre ambition que d'effectuer un rapide tour d'horizon, en nous appuyant sur des références précises empruntées à une centaine d'auteurs, sur le rôle grandissant de la psycho-acoustique dans une connaissance approfondie de la communication linguistique, en insistant particulièrement sur la nécessité d'une délimitation précise du phonème tant du point de vue phonétique que phonologique, et nous avons vu que l'unanimité est loin d'être faite à ce sujet.

Nous avons vu en quoi une connaissance toujours plus poussée de la psycho-acoustique peut venir en aide à la connaissance du langage lui-même. La phonologie, qui ne tient compte que de la fonction linguistique du phonème, doit cependant s'appuyer plus ou moins, comme toute étude, sur les éléments *matériels* que sont les segments phoniques. A leur tour ces segments phoniques, ne peuvent être délimités de façon satisfaisante par leurs seules caractéristiques physiques visuellement observables, si on ne tient pas compte de l'effet de ces caractéristiques physiques sur la perception elle-même. Un problème aussi apparemment insoluble, la catégorisation mentale, donc l'identification *non équivoque* de tel ou tel phonème, a très longtemps passé au second plan dans le domaine de la communication linguistique, car l'oreille l'acceptait tel qu'il était, et n'avait eu, depuis les anciens Grecs (et même auparavant, avant même que le concept «phonologique» ait vu le jour), aucune difficulté à le classer comme structure phonologique de base.

L'exploitation des méthodes d'analyse modernes, loin d'expliquer tous les mystères de la caractérisation du phonème, n'a bien souvent fait que compliquer les choses, tout en ouvrant les yeux de certains sur l'existence d'un réel problème, plus actuel que jamais, et encore partiellement irrésolu : qu'est-ce que le phonème ? Comment le délimiter ? Sur quelle bases *précises* doit-on s'appuyer pour lui donner enfin une définition valable *tant pour les grandeurs physiques que pour la perception*, tout comme on a fini par y parvenir, depuis peu, pour la définition des couleurs ? Et quand on aura enfin défini le phonème ne faudra-t-il pas définir avec une même précision les segments de la chaîne parlée (logatomes, syllabes,

syntagmes automatisés par l'usage) qui sont perçus comme un tout acoustique et devront être perçus comme tels par les « machines à reconnaître le langage » 14) ? Or, faute à la fois d'une simulation valable de la perception auditive, et d'une véritable « pensée », ces irréfutables témoins n'en sont encore, malgré les énormes moyens mis en œuvre, qu'à leurs tout premiers balbutiements.

Le temps est sans doute venu de tirer parti des discussions parfois confuses qui ont opposé par le passé divers théoriciens de la syllabe (cf : ROSETTI, 1961 et 1963) : Il ressort de leurs divergences de vue que la syllabe est, comme tout ce qui touche à la sensation, *une vue de l'esprit*. Elle n'existe que sous forme d'une catégorisation mentale adaptée aux besoins linguistiques, de même que sur un mur couvert de plus ou moins vagues taches de couleurs, on pourra voir — ou ne pas voir — des formes bien délimitées. Le problème de la segmentation mentale de la chaîne parlée ne pourra être résolu que si l'on s'accorde sur la véritable nature psycho-acoustique des segments de cette chaîne, sans jamais oublier que tous nos appareils ne peuvent nous fournir aucune solution valable si leurs résultats ne sont pas testés par l'oreille. Ce n'est qu'en appliquant de nouvelles méthodes d'analyse, appuyées sur de nouvelles théories de la perception auditive, que ce problème aura des chances d'être résolu un jour. Nous y travaillons depuis de longues années, mais une telle route est longue, et la patience n'est pas la moindre qualité du chercheur. N'aura-t-il pas fallu plus de vingt-trois années d'une admirable persévérance et quatre conceptions successives pour que *performan ce encore inégalee*, le « phonétographe IV » de DREYFUS-GRAF 15) parvienne à reconnaître, avec une fiabilité de 95%, dix-huit mots *prononcés par des individus différents* ? (cf : LORAND, 1971, p. 70).

De telles recherches sont-elles un lointain prélude à la réalisation de machines capables d'un dialogue direct avec l'homme ? Une telle réalisation est-elle possible, est-elle seulement intéressante ? De telles machines devraient obéir à des exigences d'utilisation que, faute de capteurs convenables, les actuels ordinateurs nous refuseront encore longtemps. GUTENMAKHER dresse un programme concret pour ces précieux auxiliaires (1961, p. 18) : « Il faut construire des machines logiques à traiter l'information, qui puissent remplir les fonctions de travailleurs intellectuels (renseignements, avis, recherches, et c...) avec la même vitesse de réaction et la même compétence » (que l'homme) « afin que le *dialogue* entre l'homme et la machine ait lieu à l'échelle de temps réel d'un dialogue habituel (conversation entre hommes). L'introduction simultanée d'un grand nombre de questions et l'obtention des réponses à ces questions après un certain temps supérieur au temps nécessaire à l'homme pour répondre à la question posée, serait la condition la plus simple et la plus facile à obtenir d'une machine. Il faut cependant considérer comme *idéale* l'exécution d'un travail intellectuel *au rythme de la conversation* ».

Or il est désormais évident qu'un tel programme ne pourrait être réalisé de façon idéale que par des machines capables de *comprendre* le langage (et non des mots isolés comme les réalisations actuelles) et de le *parler* (ces deux processus étant, comme nous l'avons vu, *intimement liés*), le problème à résoudre dépasse le cadre déjà vaste de la simple catégorisation phonémique. Ce fait dire à MOLES (1966, p. 233) : « Si le mot, c'est la pensée, comme le prouvent

14) — Et le sont effectivement sur la machine de FRY et DENES, jusqu'à un certain point. Le mot est d'abord soumis à un processus d'identification provisoire, sorte de mémorisation acoustique externe semblable à celle de l'homme. Cette identification est ensuite confrontée à l'information linguistique (mémoire interne). Pour un locuteur particulier, la fiabilité est, sans information linguistique, de 60% pour les sons isolés et de 24% pour les mots entiers. Avec information linguistique, ces taux passent respectivement à 70% et 44%. Il est évident que seule une simulation valable de la perception auditive serait de nature à améliorer de tels résultats (Cité par DEWEZE, 1966, p. 156).

15) — Qui vient à peine d'en être récompensé par l'attribution du prix GALF 1970, prix « réservé aux jeunes chercheurs » (cf : LORAND, 1971, p. 71).

nos modernes psychologues, chercher une machine à *parler*, c'est se mettre sur la trace d'une machine à *penser*, et c'est là une très grande ambition... » Puis il évoque (*ibid.*, p. 234) « le problème de la *fluidité du langage*, de sa *continuité*, d'une *structure d'ensemble* ».

Le problème est donc clairement posé : seule une machine capable de « *penser* » comme l'homme, serait capable d'un dialogue oral avec lui. Ce tonneau des Danaïdes n'aurait-il donc vraiment pas de fond ? Faudrait-il abandonner définitivement ce mythe moderne à la facile et condamnable démagogie de la science-fiction, et renoncer aux perspectives insoupçonnées que l'aboutissement de telles recherches ouvrirait un jour, encore que relativement lointain, à l'Humanité ?

Nous avons vu que la reconnaissance du langage par des machines pourrait être en premier lieu rendue plus sûre par une simulation valable de la perception auditive appuyée sur une meilleure connaissance psycho-acoustique des phonèmes. En second lieu, de telles machines ne pourront se réaliser que grâce à une souplesse puissamment accrue de la « perception machine » (possibilité de classer dans les mêmes catégories phonémiques des syllabes plus ou moins longues ou plus ou moins déformées). En troisième lieu, l'indispensable « intelligence machine » ne pourra être atteinte que par son aptitude à enregistrer de vive voix et à traiter des informations orales. Aucun ordinateur ne peut prétendre à une telle souplesse. Cependant ce qui paraissait irréalisable il y a seulement dix ans, ne l'est plus actuellement, du moins en théorie, depuis l'avènement des « mécanismes cognitifs » ou « machines mathématiques », capables non seulement de reconnaître et de catégoriser avec une parfaite fiabilité des formes même très anamorphosées, mais également d'apprendre presque indéfiniment grâce à la simulation du neurone concrétisée pour la première fois par l'« homéostat » d'ASHBY (1952). Les principes de cette cellule auto-adaptrice ont été perfectionnés et mis à profit par ROSENBLATT (1960) sous forme de son extraordinaire PERCEPTRON MARK I, capable d'auto-apprentissage et de reconnaissance des formes les plus diverses. L'intérêt de son utilisation comme capteur de sons ne fait aucun doute.

Il y a sans doute aussi loin - ou aussi près ? - de ce Perceptron aux machines « intelligentes » des siècles à venir, qu'il y avait de la marmite de Denis-Papin aux plus modernes locomotives. Mais la marmite de Denis-Papin n'a-t-elle pas préfiguré, longtemps à l'avance, l'avènement d'un bouleversement que rien, à l'époque, ne laissait prévoir ?

L'intérêt lointain de telles machines « intelligentes » réside surtout dans le fait que, ni physiquement, ni intellectuellement, l'homme n'est indéfiniment perfectible. Alors qu'une machine mathématique suffisamment perfectionnée et agrandie, continuellement éduquée au contact verbal de multiples savants (qui ne seront pas des informaticiens), infatigable, pratiquement immortelle, ne connaîtrait plus guère de limites ni en connaissances, ni même en « sagesse ».

Certains auteurs nous laissent entrevoir dans cet être aussi omniscient qu'infailible, la prochaine étape de l'Évolution, tel l'éminent informaticien DEWEZE qui nous dit à ce sujet (1966, p. 210) : « En étudiant les organismes biologiques afin de réaliser un automate, qui nous livrera à son tour des données objectives sur les mécanismes de la pensée, confirmerons-nous autre chose que la loi de l'Évolution des organismes vers une *complexité* toujours croissante, qui a fait de nous ce que nous sommes ? ».

Evidemment ce surhomme électronique — qui existera nécessairement tôt ou tard — sera bien différent de ce qu'on l'imagine parfois, et les auteurs qui ont l'audace d'en parler, ont à cœur de démystifier l'opinion publique. Voici comment l'ingénieur Claude BROSSY, imagine cet être pensant — et parlant — sous le sous-titre : « Un conte... aujourd'hui » (1966, p. 42) :

« Le jour n'est pas loin où nous allons ne plus pouvoir refuser le titre d'« être pensant » à un mécanisme fait entièrement des mains de l'homme. Il ne se présentera certainement pas

sous la forme robotique chère à la science-fiction. Imaginez plutôt pénétrer dans un imposant bâtiment. Vous vous trouvez dans une antichambre ; pour tous meubles, une table, une chaise. Le microphone sur la table vous intrigue. Vous prononcez quelques mots. Par haut-parleur, on vous annonce la bienvenue. Vous répondez, et ainsi de fil en aiguille vous poursuivez un dialogue étonnant. Votre interlocuteur a peut-être parfois des déficiences dans sa culture générale, mais il raisonne juste, il semble de bonne composition, réceptif, *prêt à apprendre*. Ce qui vous étonne pourtant, c'est sa personnalité. Ayant toujours le mot pour rire, il cache cependant sous son air frondeur une certaine angoisse que vous ne tardez pas à déceler. Vous le harcelez de questions, et vous découvrez qu'il y a chez lui un malaise qu'il peut difficilement expliciter, comme s'il se sentait prisonnier dans une carcasse de béton, à la merci de forces occultes... Cet être qui pense, se comporte, et si on lui donnait les effecteurs nécessaires (ce qui est bien la moindre des difficultés) pourrait, autonome, agir comme un être humain, cet être *aurait pu être réalisé depuis longtemps*, puisqu'il ne comprendra aucune pièce *que l'on ne connût depuis des décennies*, ou qui n'ait même son analogue mécanique : les Jacquart, les Vaucanson auraient pu en imaginer le principe. Ceci est un conte...aujourd'hui. Nous voudrions faire comprendre... pourquoi ceci n'est qu'une *extrapolation*. Des systèmes adaptatifs puis auto-adaptatifs dont l'exemple le plus classique est l'«homéostat» d'ASHBY, aux «machines mathétiques», *le progrès est continu sur la voie qui conduit aux machines intelligentes* ».

Libre à nous de vaticiner sur le temps qui nous sépare de pareille naissance, si progressive qu'on ne saurait même parler de naissance, mais plutôt d'Evolution, car dit encore BROSSY (ibid., p. 56), « la machine autonome, vivante, existe peut-être déjà ». En mettant les choses au pire nous pourrions compter cinq cents ans, ce qui est tout de même une prospective des moins déraisonnables. Cependant les premières «machines mathétiques» dont l'«intelligence», c'est à dire la faculté de catégorisation, la capacité de reconnaître les formes les plus diverses et *d'apprendre*, sont assez similaires aux qualités correspondantes d'un jeune enfant, ont déjà dépassé le stade «intellectuel» du protozoaire. Or il aura fallu plus de cinq cent millions d'années pour que l'Evolution transforme peu à peu ce protozoaire en être humain. Que seront en comparaison ces malheureux cinq cents ans, porte bien mince qui nous sépare de ce trésor des âges futurs ?

Certes, la porte est mince, mais les combinaisons de ses serrures sont nombreuses. Qu'il en manque une seule : elle restera fermée. Et l'une de ces multiples combinaisons, et non la moindre, se dévoile peu à peu dans la pénombre des laboratoires de phonétique où de persévérants chercheurs l'auront traquée, année après année, sur des papiers sensibles, ou sur de petites fenêtres rondes couleur de perle où s'inscrit, dans le lumineux ballet magique des électrons, l'insaisissable frémissent des ondes sonores qui depuis des millénaires véhiculent la Pensée humaine.

RESUME.

Se référant à une centaine d'ouvrages, l'auteur effectue un tour d'horizon sur le rôle grandissant de la psycho-acoustique dans une connaissance approfondie de la communication linguistique. Il aborde en particulier le problème de la caractérisation et de la délimitation du phonème à travers ses aspects phonologique et phonétique, et conclut que l'unanimité est loin d'être faite à ce sujet, faute d'une simulation valable de la perception auditive, qui reste encore à réaliser, et peut seule ouvrir la voie aux applications les plus actuelles de la phonétique : la reconnaissance du langage humain par les machines.

الخلاصة

بماذا تستطيع سيكولوجية الصوت اللغوي أن تساهم في زيادة معرفتنا لظاهرة التبليغ اللغوي

هذه المأمة بين فيها صاحبها ، بالاعتماد على مجموعة من الوثائق تبلغ المئة مؤلف ، الدور المتزايد الأهمية الذي لا تزال تلعبه سيكولوجية الصوت اللغوي في توسيع معرفتنا لظاهرة التبليغ . وقد تعرض بالخصوص لمشكلة التمييز والتحديد لذوات الحروف من خلال مظاهرها الادائية وأفضى به ذلك الى الاثبات بأن اجماع العلماء على نظرية في هذا الميدان لا يزال بعيد التحصيل ويرى سبب ذلك في عدم وجود تمثيل اجرائي صحيح لعملية الادراك السمعي وهذا التمثيل الذي يحتاج الى ان يحققه الباحثون هو في رايه الوسيلة الوحيدة التي يمكن ان يفسح المجال بها لحدث تطبيق تحاول الصوتيات ان تنجزه وهو اعطاء الاجهزة الاليكترونية القدرة على تحقيق وتشخيص كلام الانسان .

SUMMARY

What can psychoacoustics contribute to deepen our knowledge of linguistic communication ?

Quoting about a hundred publications, the author makes a general synthesis of the growing importance of psychoacoustics for a deepened knowledge of linguistic communication. He approaches more accurately the problem of the characterisation and the delimitation of the phoneme, through its phonological and phonetic aspects, and concludes that its definitions are far from gaining unanimous approval, owing to the lack of an adequate simulation of auditory perception. Such simulation, which has not yet been carried out, would be the only way to make really possible the most advanced applications of phonetics : intelligibility of human speech by machines.

BIBLIOGRAPHIE

des ouvrages cités dans le texte.

- ABERCROMBIE D. : *Studies in Phonetics and Linguistics*, Oxford 1965 (p. 76, p. 84).
- ARISTOTE : *Catégories, de l'interprétation* (de : Organon, trad. J. Tricot, Paris 1966) pp. 5-6, p. 17.
- ASHBY W.R. : *Design for a brain*, New-York 1952.
- AVRAM A. : *Le principe de la continuité en phonologie diachronique*, Recherches sur la philosophie des sciences, Bucarest 1971 (p. 448).
- BENVENISTE E. : *Problèmes de linguistique générale*, Paris 1966 (p. 63).
- BLOOMFIELD L. : *Language*, New-York 1933 (existe également en traduction française : *le Langage*, trad. J. Gazio, Paris 1970).
- BOUMAN A. : *Analyse des sons*, Arch. néerl. phon. exp., 2 : 50-90, 1926 cité par Essner Ch., 1947, op. cit., p. 55).
- BOURCIEZ E. : *Précis historique de phonétique*, Paris 1955.
- BRILLOUIN L. : *Le rôle des surfaces absorbantes dans la propagation du son*, Rev. Acoust., vol. IV, 4 : 113-129, 1935.
- BROSSY CL. : *Les machines mathématiques*, Atomes : 41-57, 1966 (p. 42, p. 56).

- CHOMSKY N. : *Current issues in linguistic theory*, The Hague 1967 (pp. 30-34, p. 31).
- CHRETIEN L. : *Les décibels*, Paris 1964 (p. 8).
- DELATTRE P. : *Un triangle acoustique des voyelles orales du français*, The French Review, XXI, 6 : 477-484, 1948 (p. 477).
- DELATTRE P. : *Principes de phonétique française* (2ème édition), Middlebury 1951 (p. 45).
- DELATTRE P. : *Indices acoustiques de la parole, premier rapport*, *Phonetica* 2 : 226-251, 1958.
- DELATTRE P. : *Le jeu des transitions de formants et la perception des consonnes*, Proc. IVth internat. Congr. phon. Sci. Helsinki 1961, The Hague 1962 (pp. 407-417).
- DELATTRE P. : *Comparing the phonetic features of English, German, Spanish and French*, Heidelberg 1965 (pp. 113-118, pp. 48-50, p. 52).
- DELATTRE P. : *Studies in French and comparative phonetics*, The Hague 1966 (p. 266, p. 82).
- DELAVENAY E. : *La machine à traduire*, Paris 1963.
- DEWEZE A. : *Traitement de l'information linguistique*, Paris 1966 (pp. 159-160, p. 153, p. 156, p. 210).
- DREYFUS-GRAF J.A. : *Le sonographe ; éléments et principes*, Schweizer Arch. angew. Wiss. und Techn., 12 : 353-362, 1948.
- DREYFUS-GRAF J.A. : *Sonograph and sound mechanics*, Journ. Acoust. Soc. Amer., 22 ; 6 : 731-739, 1950.
- DREYFUS-GRAF J.A. : *Le sténo-sonographe phonétique*, Techn. Mitt. PTT, 3, 1950.
- DREYFUS-GRAF J.A. : *Phonétographe, présent et futur*, Bulletin techn. PTT, 5 : 160-172, 1961.
- DREYFUS-GRAF J.A. : *Phonétographe et cybernétique*, 5ème Congr. internat. Acoust., Liège 1965 (A 36, p. 2).
- DREYFUS-GRAF J.A. : *Cybernétique auditive*, Rev. Acoust., 14 : 99-103, 1971 (p. 100).
- EMERIT E. : *L'image sonore trichrome à trois dimensions par la synchronotrichromographie acoustique*, *Phonetica* 14 : 193-236, 1966 (p. 219, p. 196, pp. 212-216).
- EMERIT E. : *Vérification de la perception trichrome du son par les courbes d'isophonie, et bases d'une colorimétrie acoustique*, *Phonetica* 17 : 208-230, 1967 (p. 211).
- EMERIT E. : *Relation entre la force subjective du son et le nombre de paliers de plus petite différence d'intensité pour quatre fréquences caractéristiques*, *Phonetica*, 25 : 72-107, 1972.
- EMERIT E. : *La trichromie auditive et ses applications* (thèse de phonétique, Sorbonne, Paris 1970, à paraître à l'Inst. Ling. et Phon. Université d'Alger).
- ESSNER CH. : *Recherches sur la structure des voyelles orales*, Arch. néerl. Phon. exp., 20 : 41-77, 1947 (p. 55).

- EVA MARK *III Instruction Manual* : (Manufactured by Melpar Inc., 7 700 Arlington bd, Falls Church, Virginia 22 046, - additional leaflet 2, sans pagination).
- EXNER S. UND POLLACK J. : *Beitrag für Resonanztheorie der Tonempfindungen*, Zeits. für Psychol., 32 : 305-332, 1903.
- FALINSKI E. : *Psycho-pédagogie du langage écrit*, Paris 1966 (p. 11).
- FANT G. : *Analysis and synthesis of speech processes*, Malmberg Manual of Phonetics, Lund 1968 (p. 241).
- FISCHER-JØRGENSEN E. : *What can the new techniques of acoustic phonetics contribute to linguistics ?*, Proc. VIII th internat. Congr. of Linguists, 433-478, Oslo 1958 (pp. 447-448, p. 443).
- FLANAGAN J.L. : *A difference limen for vowel frequency*, Journ. Acoust. Soc. Amer., 27 : 3, 1955.
- FLANAGAN J.L. : *A difference limen for formant amplitude*, Journ. of speech and hearing disorders, 22 ; 2 : 205-212, 1957.
- FLETCHER H. : *Loudness, its definition, measurement and calculation*, Journ. Acoust. Soc. Amer., 5 : 82, 1933.
- FLETCHER H. : *The physical criterion for determining the pitch of a musical tone*, Western Electric Co Engineering Dept. B 61-1 : 1-11, 1924, — reprint from : Phys. Rev., 23 : 3, (1924) (p. 7).
- FLETCHER H. : *Loudness, pitch and timbre of musical tones*, Bell teleph. Syst. techn. publ., monograph B. 824, 1934, — reprint from : Journ. Acoust. Soc. Amer., 6 : 59-69, (1934).
- FOUCHÉ P. : *Phonétique historique du français*, Paris 1951 (p. 99).
- FRANCES R. : *Le développement perceptif*, Paris 1962 (p. 158).
- FRANÇOIS F. : *Le langage* Paris 1968 (p. 13).
- GALAMBOS R. : *Neurophysiology of the auditory system*, Journ. Acoust. Soc. Amer., 22 ; 6 : 785-791, 1950.
- GIROD F. : *Surdité, un traitement révolutionnaire*, Science et Vie, n° 557, 1964.
- GRAMMONT M. : *Traité de phonétique* (8ème édition), Paris 1965 (p. 15, p. 144).
- GRÜTZMACHER M. : *Demonstration eines Tonhöschreibers*, Phonetica 13 : 3-17, 1965.
- GSELL R., CARRE R. LANCIA R. et PAILLE J. : *Etude et réalisation d'un détecteur de mélodie pour analyse de la parole*, l'Onde électrique, 434 : 1-7, 1963.
- GUBERINA P. : *Povezanost jezičnih elemenata*. Zagreb 1952 (p. 157, p. 128).
- GUTENMAKHER L.I. : *Traitement électronique de l'information* (traduit du russe par A. Deweze) Paris 1961 (p. 18).
- HADJ-SALAH A. : *Nouvelle contribution à la connaissance physiologique de l'opposition sourde/sonore*, VIIème Congrès des Sciences phonétiques, Montréal 1971.
- HALÁ B. : *Nature acoustique des voyelles*, Acta Universitatis Carolinae, Prague, V : 3-119, 1956 (cité par Husson R., op. cit., 1958, p. 341).

- HALLE M. : *Questions of linguistics*, Nuovo Cimento, 13 : 494-517, 1959.
- HALLE M. : *The sound pattern of Russian*, The Hague 1959.
- HELMOLTZ H.L.F. : *Die Lehre von den tonempfindungen als physiologische Grundlage für die Theorie der Musik*, Braunschweig 1863 (chap. 4).
- HIRSH I.J. : *La mesure de l'audition* (traduit de l'anglais par J. Bouche), Paris 1956.
- HUSSON R. : *Comment se forment les voyelles*, La Nature, 3267 : 249-257, 1957 (pp. 250-251).
- HUSSON R. : *Le pavillon pharyngo-buccal et sa physiologie phonatoire*, La Nature, 3281 : 337-344, 1958 (p. 341).
- JAKOBSON R. : *Principes de phonologie historique* (appendice à : Principes de phonologie, par Troubetzkoy N.S., traduit du russe par J. Cantineau), Paris 1949 (p. 368).
- JAKOBSON R. : *Selected writings ; phonological studies*, The Hague 1971 (pp. 484-486, p. 231, pp. 751-752).
- JONES D. : *An English pronouncing dictionary* (8ème édition), London 1947 (p. 419).
- LANE H.L. : *A behavioral basis for the polarity principle in linguistics*, Language, 43 : 494-511, 1967.
- LANE H.L. : *The hue and cry concerning categorical perception (and conversely)*, Zeits. Phonet. Sprachw. KommunForsch., 21 ; 1-2 108-115, 1968 (p. 113).
- LANE H. L. : *Production et perception de la parole*, Phonetica 23 ; 2 : 94-125, 1971 (p. 108).
- LEIPP E. : *Le problème de l'intelligibilité de la parole*, Rev. Acoust., 12 : 343-358, 1970.
- LIBERMAN A.M., INGEMANN F.
LISKER L., AND DELLATRE
P. : *Minimal rules for synthesizing speech*, Journ. Acoust. Soc. Amer., 31 ; 11 : 1490-1499, 1959.
- LIBERMAN A.M., COOPER
F.S., HARRIS K.S., AND MAC
NEILAGE P.F. : *A motor theory of speech perception ; in Fant Proc. Speech Communication Seminar*. Unpublished report, Speech Transmission Laboratory, Royal Institute of Technology, Stockholm 1963 (p. 3 : cité par Lane H.L., 1971, op. cit., p. 108).
- LINDNER G. : *Veränderungen der Formant-Intensität bei synthetischen Vokalen*, Zeits. Phonet. Sprachw. KommunForsch., 23 ; 5 : 498-504, 1970.
- LORAND P. : *Eloge de Monsieur Jean Dreyfus-Graf*, Rev. Acoust., 14 : 70-71, 1971.
- LOTZ J. : *Linguistics ; symbols make Man*, New-York 1956 (p. 207).
- MALMBERG B. : *Allocution du Président de la Société internationale des Sciences phonétiques*, Proc VIth internat. Congr. Phon. Sci., Prague 1967 (p. 89).

- MALMBERG B. : *Phonétique française*, Malmö (Suède) 1969 (p. 25, p. 106).
- MARTINET A. : *Éléments de linguistique générale*, Paris 1964 (p. 45).
- MARTINET A. : *La linguistique synchronique*, Paris 1965 (p. 36, p. 37).
- MILLET A. : *L'oreille et les sons du langage*, Paris 1926 (p. 69).
- MOLES A. : *Méthode cybernétique et structures linguistiques* (de : Phonétique et Phonation, par Moles A. et Vallancien B., pp. 233-251), Paris 1966 (p. 236, pp. 233-234).
- OSGOOD CH. AND SEBEOK TH : *Psycholinguistics*, Indiana University Press 1967 (p. 4, p. 8, p. 11).
- PAGET R. : *Human speech*, London 1930 (p. 154 et sqq.).
- PAVLOV I.P. : *Oeuvres complètes*, Moscou 1951 (t. III ; 1. 2, p. 215).
- PETERSON G.E. AND BARNEY H.L. : *Control methods used in a study of the vowels*, Journ. Acoust. Soc. Amer., march 1952, pp. 175-184.
- POTTER R.K. KOPP G.A., AND GREEN H.C. : *Visible speech*, New-York 1947.
- POTTER R.K. AND STEINBERG J.C. : *Toward the specification of speech*, Journ. Acoust. Soc. Amer., 22 ; 6 : 807-820, 1950 (p. 816).
- PRIETO L.J. : *Principes de noologie*, The Hague 1964 (p. 24).
- RIESZ R.R. : *Differential intensity sensitivity of the ear for pure tones*, Bell teleph. Laboratories, reprint B. 325, 1928, from : *Phys. Rev.*, second series, 31 : 867-875, (1928).
- RIGAULT A. : *Rôle de la fréquence, de l'intensité et de la durée vocaliques dans la perception de l'accent en français*, Proc. IVth Congr. Phon. Sci. : 735-748, Helsinki 1961 (p. 735).
- ROSENBLATT F. : *Perceptron simulation experiments*, Proc. I.R.E., vol. 48 : 301-309 (1960).
- ROSETTI A. : *Sur le problème de la syllabe ; réponse à M.B. Halà*, *phonetica* 7 : 109-113, 1961.
- ROSETTI A. : *Les problèmes de la syllabe*, The Hague 1963.
- ROUSSEAU J.J. : *Essai sur l'origine des langues*, A. Belin, 1817 (chap. 16 : pp. 535-536), ouvrage reproduit dans la série : « Les Cahiers pour l'analyse », Paris (p. 535).
- ROUSSELOT (Abbé) : *Phonétique et surdité* (p. 114) — cité par Millet A., 1926, op. cit., p. 69.
- SAPIR E. : *Le langage*, Paris 1967 (p. 7).
- SKALA O. : *Experimentelle und theoretische Grundlagen des Trautoniums*, *Frequenz*, 2 : 315, N° 12 (1948) et 3 : 13, N° 1 (1949).
- STEVENS S.S. : *The relation of pitch to intensity*, Journ. Acoust. Soc. Amer., 6 : 150-154, 1935.
- STEVENS S.S. : *A scale for the measurement of a psychological magnitude : loudness*, *Psychol. Rev.*, 43 : 405-416, 1936.

- STEVENS S.S. AND DAVIS H. : *Hearing, its psychology and physiology*, New-York 1947.
- TROUBETZKOY N.S. : *Principes de phonologie* (traduit du russe par J. Cantineau), Paris 1949 (p. 13, pp. 343-350, p. 7).
- TUNTURI A. : *Audiofrequency localization in the acoustic cortex of the dog*. Amer. Journ. Physiol., 141 : 397-403, 1944.
- UNGEHEUER G. : *Systematische Signaldestruction als Methode der psychoakustischen Phonetik*, Phonetica 18 ; 3 : 129-185, 1968.
- VIETZE H.P. : *Strukturtheorie des Wortes und seiner Komponenten*, Zeits. Phonet. Sprachw. KommunForsch., 23 ; 2/3 : 261-290 1970.
- WEGEL R.L. AND LANE C.E. : *Auditory masking and dynamics of the inner ear*, Bell teleph. Laboratories, reprint B. 58-1, 1926, from : Phys. Rev. 23 : 266-285, (1924).
- WEVER H.G. : *Theory of hearing*, New-York 1949.
- WHORF B.L. : *Linguistic relativity and the relation of linguistic processes to perception and cognition*, Language, thought and reality, New-York 1956 (p. 2).
- WINCKEL F. : *Vues nouvelles sur le monde des sons* (traduit de l'allemand par A. Moles et J. Lequeux), Paris 1960 (pp. 63-64, pp. 39-56).
- ŽINKIN N. : *Mechanisms of speech*, The Hague 1968 (p. 238).
- ZWIRNER E. : *Aufgaben und Methoden der Sprachvergleichung durch Mass und Zahl*, Phonetrie, Zeits. für Mundartforschung, XII ; 2: 78 (cité par Troubetzkoy N.S., 1949, op. cit., p. 7).
- ZWIRNER E. UND K. : *Grundlagen der Phonetrie*, Berlin 1936.