

Perception du signal musical

par Michèle CASTELLENGO

Directeur du Laboratoire d'Acoustique Musicale (L.A.M.)
Unité de recherches associée au CNRS, à l'Université Paris VI
et au Ministère de la culture et de la francophonie
L.A.M. - Université Paris VI - Tour 66 - Boîte courrier 161
4, place Jussieu - 75005 Paris

Depuis un demi-siècle environ le son est devenu objet d'étude, et les diverses représentations du signal acoustique dont nous disposons maintenant permettent d'accéder à des mesures précises des grandeurs physiques : l'intensité (en dB), la fréquence (en Hz), la durée (en s., ms.). Mais dès que l'on souhaite interpréter les données pour mettre en correspondance l'analyse acoustique et le jugement perceptif, plusieurs ordres de difficultés surgissent du fait que l'auditeur humain capte le signal acoustique selon des modalités très particulières :

1 - Le signal effectivement reçu par le capteur auditif dépend des caractéristiques de sensibilité de l'ensemble de la chaîne. Or les «courbes de réponses» sont loin d'être plates, et varient d'un individu à l'autre.

2 - La variation d'une seule grandeur physique correspond le plus souvent à un effet perceptif complexe (changement de fréquence entraînant un changement d'intensité perçue ou de timbre).

3 - La réponse perceptive varie en fonction du contexte temporel. Elle s'élabore relativement aux sons perçus dans le passé (proche ou lointain), et conditionne, par anticipation, ceux que nous allons entendre.

Le système auditif humain, qui a des performances étonnantes pour traiter rapidement et globalement l'information des signaux complexes, semble paradoxalement inapte aux mesures quantitatives et reproductibles. Il est vrai que l'ouïe est avant tout un sens d'alerte, permettant de détecter un danger lointain. Aussi avons nous d'excellentes capacités pour repérer la direction des sons, et surtout reconnaître de quelle source il s'agit.

La plupart des connaissances que nous avons sur la perception auditive proviennent aujourd'hui encore d'expériences classiques menées en laboratoire à l'écoute de signaux normalisés : sons purs, impulsions et bandes de bruits, signaux fort éloignés des sons habituels de notre environnement. La démarche consiste le plus souvent à étalonner le système auditif, comme on le ferait d'un transducteur électro-acoustique : courbe de réponse, détection de seuils, etc. [2] [3]. Ces recherches sont complétées aujourd'hui par les enseignements issus de deux démarches. D'une part la synthèse sonore, qui permet d'imaginer des expériences avec des sons complexes et évolutifs que l'on maîtrise parfaitement ; d'autre part les expériences de manipulation du son dans les studios de production musicale et d'enregistrement sonore, qui ont fait progresser la connaissance de la perception spatio-temporelle.

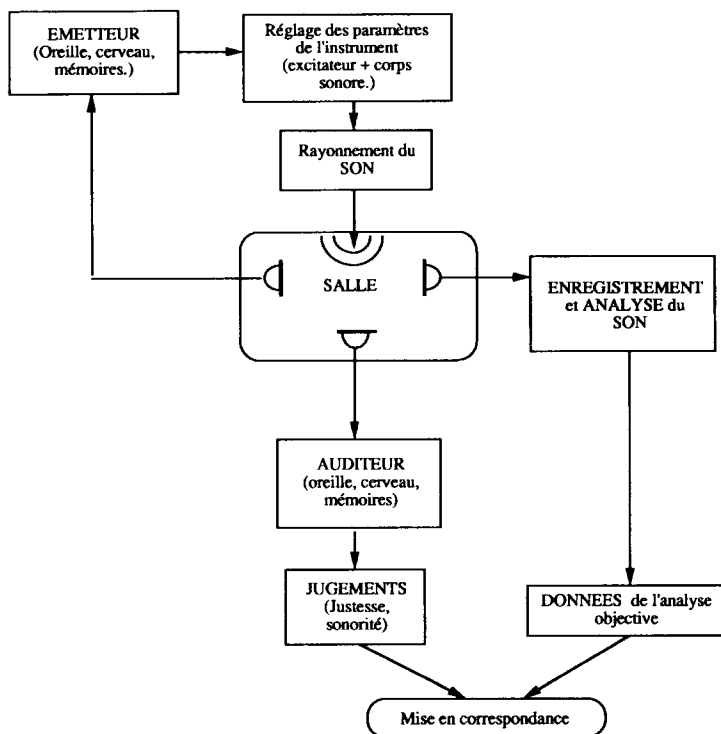


Figure 1 : Organigramme des éléments de la chaîne d'écoute, depuis le musicien jusqu'à l'auditeur.

LA «CHAÎNE D'ÉCOUTE», DU MUSICIEN À L'AUDITEUR

Le schéma classique de la figure 1 nous permettra de rappeler quelques points.

L'émetteur a généralement un «point d'écoute» très particulier sur son instrument (qu'on pense au violon, à la flûte, à la voix !), et c'est à partir de cette écoute qu'il règle la production du son. Les auditeurs sont nécessairement à des endroits différents, de même que les capteurs sonores (microphones) placés pour l'enregistrement et l'analyse acoustique. Les divergences seront d'autant plus marquées que le rayonnement de la source est variable avec la fréquence, et que la salle est hétérogène.

PARTICULARITÉS DU CAPTEUR OREILLE

Rappelons brièvement les caractéristiques les plus remarquables du système auditif, en priant le lecteur de se reporter aux ouvrages cités en fin de chapitre.

– Sensibilité auditive : elle dépend de la fréquence et présente un maximum dans une zone privilégiée, schématiquement entre 2000 et 4000 Hz. La sensibilité décroît très rapidement dans les basses fréquences (< 50 Hz) et les hautes fréquences (> 10 000 Hz). cf la courbe de seuil figure 2.

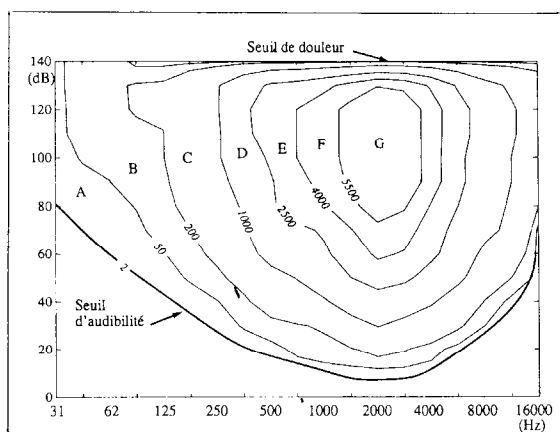


Figure 2 : Domaine d'audibilité et capacité de discrimination de l'oreille humaine. Le maximum de sensibilité (zone G) se situe aux alentours de 2000 Hz pour des intensités élevées, cf CASTELLENGO [11].

– La sensibilité auditive varie d'un individu à l'autre, se modifie selon l'état de santé et s'altère irrévocablement avec l'âge (figure 3). Prenons pour exemple la perte des aigus dont la limite qui est d'environ 20 000 Hz à quinze ans, tombe à 3000 Hz à quatre-vingt ans... pour fixer les idées.

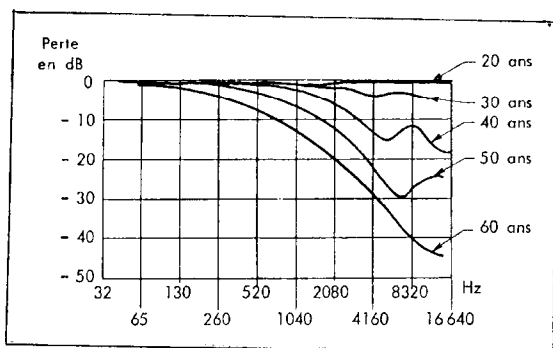


Figure 3 : Perte de la sensibilité auditive aux sons aigus, en fonction de l'âge ; moyenne statistique, cf LEIPP [1].

– La sensibilité différentielle de l'oreille est remarquable. Dans les meilleures conditions on perçoit : $\Delta F = 1/300^{\text{ème}}$ d'octave (soit 1 savart ou 4 cents) comme plus petit intervalle mélodique ; $\Delta I = 0,5$ dB pour les différences d'intensité ; $\Delta T = 1$ ms comme plus petit intervalle temporel entre deux sons impulsifs successifs. En audition binaurale un décalage de 0,5 ms entre les deux oreilles est parfaitement perçu : il permet de localiser les sources dans l'espace.

Le pouvoir de résolution de l'oreille, défini comme le nombre de sons élémentaires que nous pouvons discriminer (produit $\Delta F \times \Delta I$) varie grandement avec la fréquence et avec l'intensité. Très faible aux basses fréquences, il passe par un maximum vers 3000 Hz et 90 dB, puis décroît vers les hautes fréquences (figure 2).

Quelles sont les conséquences de l'inégale sensibilité de l'oreille aux diverses fréquences ?

1 - La seule mesure de l'intensité en dB ne peut rendre compte de l'intensité perçue puisque celle-ci dépend de la fréquence. Il faut donc tenir compte de la composition spectrale des sons. Ainsi lorsqu'on écoute deux sons de même fréquence fondamentale (500 Hz) et de même intensité globale (60 dB), l'un pur et l'autre possédant cinq harmoniques, le son complexe paraît beaucoup plus intense que le son

pur car une partie de l'énergie est dans la bonne zone de l'oreille (harmonique 4 à 2000 Hz et harmonique 5 à 2500 Hz).

Les musiciens exploitent - plus ou moins consciemment - ce phénomène. Lors d'un crescendo, on constate une augmentation importante de l'intensité des harmoniques situés dans la zone sensible de l'oreille* (figure 4). D'autre part, un chanteur dont la voix porte, développe un formant renforçant les harmoniques situés dans cette même zone (figure 5).

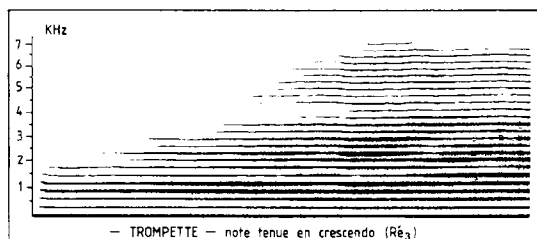


Figure 4 : Crescendo sur une note tenue (R₆₃) jouée à la trompette, et montrant l'enrichissement progressif en harmoniques jusqu'au *ff*. cf CASTELLENGO [8].

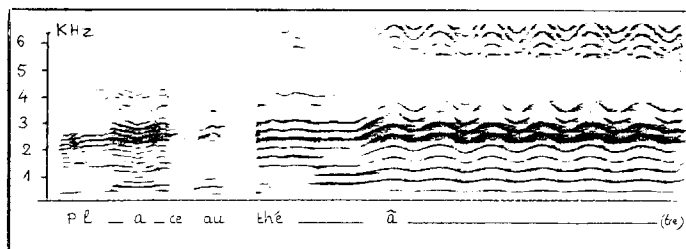


Figure 5 : Les chanteurs, les acteurs d'opéra, ainsi que certains orateurs sont capables de «placer» leur voix de sorte qu'une grande partie de l'énergie qu'ils émettent soit focalisée dans la bonne zone de l'oreille, ce qui leur assure une grande efficacité sur scène. L'exemple est particulièrement net sur cette figure, où l'on voit un formant fixe autour de 2500 Hz.

* NDLR : Pour le principe du sonagramme on pourra consulter la référence bibliographique n° 4 de l'éditorial.

Un sonagramme représente l'évolution du spectre d'un son en fonction du temps. L'amplitude d'une harmonique est d'autant plus grande que le trait correspondant est gras.

Autre exemple : l'analyse de la voix captée au téléphone montre aussi un renforcement du spectre à 3000 Hz (figure 6). Un tel système, s'il est efficace du point de vue de l'intensité affecte toutefois la qualité de la voix du locuteur. Comme nous l'avons annoncé au point 2 de l'introduction, les grandeurs perspectives ne sont pas indépendantes.

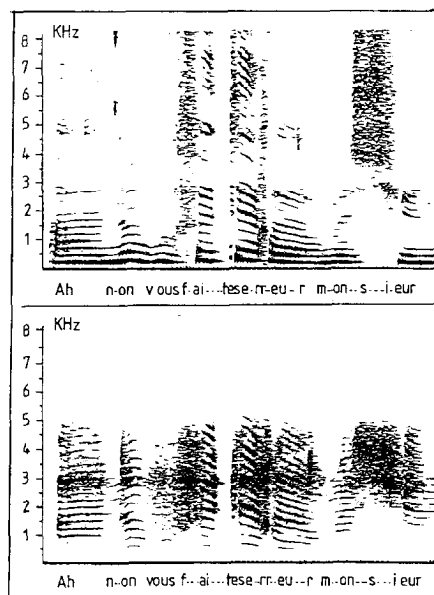


Figure 6 : Analyse au sonographe d'une phrase enregistrée simultanément près d'un locuteur (en haut) et à la sortie d'un récepteur de téléphone (en bas). Noter, outre les filtres des graves et des aigus, l'importante résonance à 3000 Hz.

2 - La seule mesure de la fréquence ne peut rendre compte de la perception de la hauteur. Celle-ci change selon la tessiture et dépend également du contenu spectral.

Trois sons synthétiques de même fréquence fondamentale mais de composition harmonique différentes peuvent donner des sensations de hauteur dissemblables. Généralement un son riche en harmoniques paraît plus aigu qu'un son pauvre. Ce phénomène intervient dans l'appréciation de la justesse mélodique d'un instrument à vent, par exemple lorsque certains doigtés produisent des sons de timbre très différents. Il permet également d'expliquer certains « paradoxes » comme le changement apparent de diapason des instruments jouant en coulisse [4].

Aux très basses fréquences (20 à 30 Hz), nous entendons plutôt une suite d'impulsions qu'un son continu. De plus, lorsqu'il s'agit de sons riches en harmoniques (contrebasson, clarinette basse), la hauteur spectrale due aux formants entre en conflit avec la hauteur fondamentale, et peut donner lieu, dans une succession de sons, à la perception d'une ligne mélodique fort différente de celle qui est produite par la fréquence fondamentale.

Pour les sons aigus supérieurs à 1000 Hz, la sensation d'intervalle ne correspond plus aux rapports numériques des fréquences : le fait de doubler la fréquence, en passant de 1000 à 2000 Hz donne la sensation d'un intervalle plus petit que l'octave. Cette particularité est prise en compte par les accordeurs de piano qui travaillent à l'oreille. Un relevé de mesures des fréquences montre que les notes de la dernière octave sont accordées avec un écart croissant par rapport à la justesse physique (figure 7). Pour la même raison, les notes aiguës d'une petite flûte doivent être accordées plus haut que la référence donnée par un accordeur électronique pour paraître justes à l'oreille.

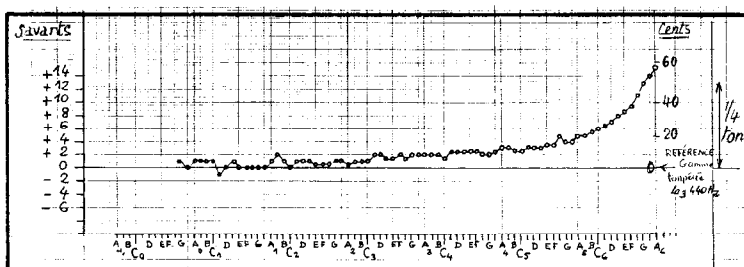


Figure 7 : Relevé de justesse d'un piano immédiatement après le passage d'un bon accordeur. Ici le diagramme montre clairement la montée progressive des notes aiguës, surtout dans la dernière octave.

RÔLE DU CONTEXTE TEMPOREL

1 - Le Passé - «On ne se baigne jamais deux fois dans le même fleuve», dit le proverbe. De la même façon, la sensation sonore que nous éprouvons à l'écoute d'un son, d'une séquence musicale, d'une œuvre, change d'une fois à l'autre. Tout jugement s'opère «par rapport» à une expérience antérieure. **Selon les éléments mémorisés dans le passé proche ou lointain l'avis que nous allons formuler pourra se modifier plus ou moins fortement.**

Prenons deux exemples :

– Demandons à un pianiste de jouer le même accord répété lentement avec la même nuance trois fois de suite, en alternant : *P, mF, F, mF, P*. La mesure de l'intensité, qui pour cet instrument et pour un accord répété, rend bien compte de la nuance réglée par le pianiste, montre d'étonnantes irrégularités. On constate que les nuances *mF* et *P* venant après le *F* ne sont pas au même niveau d'intensité que celles qui ont été réalisées dans le crescendo (figure 8). Ce fait est bien connu des chefs d'orchestre. Après un *FF* «éblouissant» l'oreille, il faut exagérer les écarts de nuance.

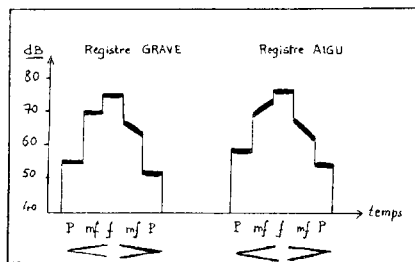


Figure 8 : Mesure de l'intensité d'un accord joué au piano, répété avec des nuances différentes. On voit nettement que la sensation d'intensité dépend du contexte temporel.

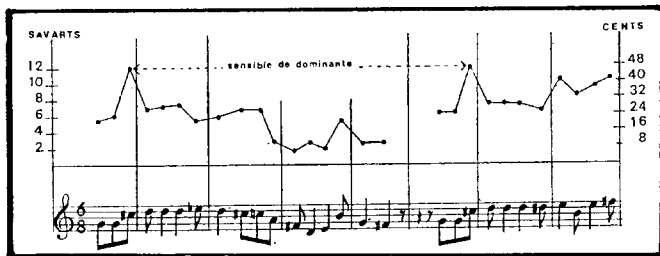


Figure 9 : Lors de l'exécution d'une mélodie au cor anglais (air du Roi de Thulé de La Damnation de Faust de Berlioz), on a mesuré l'écart des notes par rapport au tempérament égal. Le graphique met en évidence la façon dont le musicien adapte la grandeur des intervalles selon le contexte : notes sensibles, sens ascendant ou descendant de la mélodie. cf LEIPP [4].

– Tout musicien distingue sans peine un intervalle d'un demi ton de celui d'un quart de ton. Et pourtant, dans un contexte tonal, la grandeur perçue des intervalles devient quasiment élastique et change selon le mouvement mélodique. L'exemple de la figure 9 est significatif à cet égard. En position d'attraction la notre *Do#* est jouée beaucoup plus haute que les autres notes de la mélodie, de sorte que le triton *Sol Do#*

est agrandi et le demi-ton Do#Ré très réduit. Il est évident que cet exemple est apprécié comme étant très juste musicalement.

Nous n'avons parlé ici que du passé proche. Le passé lointain, mémoire de l'individu, conditionne ses réactions qualitatives : goût musical, style de jeu, appréciation de la justesse et de la sonorité [5].

2 - L'avenir - L'écoute active implique que l'auditeur, à chaque instant du temps présent, parie sur les événements sonores à venir, et adapte son écoute en fonction de la prévision qu'il fait. On sait maintenant que l'adaptation a des bases physiologiques.

– Adaptation du niveau sonore à l'entrée de l'oreille : l'articulation de la chaîne des osselets qui assure la transmission des vibrations par voie solide entre le tympan et l'oreille interne est réglée par deux muscles commandables par voie réflexe [5]. Selon que nous prévoyons d'écouter des sons faibles ou intenses nous pouvons ainsi adapter le niveau d'environ 30 dB. Cette adaptation mécanique s'accompagne de changements dans la perception de la sonorité.

– Les récentes découvertes de la physiologie cochléaire montrent que nous sommes capables, avant l'écoute d'un son, d'activer sélectivement certaines cellules de l'oreille interne, ce qui permettrait d'expliquer l'extraordinaire finesse d'analyse auditive dont font preuve les musiciens [10], ainsi que la capacité que nous avons «d'aller chercher» dans un ensemble complexe le bruit qui nous intéresse.

Finalement nous conseillons au lecteur de se reporter aux ouvrages sur la théorie de l'information [6], en soulignant l'importance de la prévisibilité dans l'attitude d'écoute : **c'est ce qui change qui est intéressant !**

PERCEPTION DES FORMES : TIMBRE ET SONORITÉ

Les principes de la théorie de la forme, développés surtout dans le domaine visuel, permettent de bien comprendre la plupart des caractéristiques de la perception sonore que nous venons d'évoquer et en particulier l'interdépendance des grandeurs perceptives. Une forme n'est pas définie par l'accumulation des parties qui la constituent, mais par les rapports, de grandeur et de situation, de ces parties [7]. **A la notion de mesure absolue se substitue celle de proportion.** Toute la difficulté, pour l'acousticien, consiste donc à établir une correspondance entre le monde physique fournissant la mesure des diverses

grandeurs du signal acoustique, et le monde perceptif qui nécessite une approche globale et dynamique.

Le cas de la perception du timbre est à cet égard exemplaire. Son étude est souvent absente des traités de psychoacoustique [2, 3], car c'est essentiellement un attribut relevant simultanément de toutes les dimensions sonores : fréquence, amplitude, durée. Ce que nous appelons communément le timbre comporte en fait deux aspects [8]. Le premier se réfère à l'identification de la source sonore (causalité) et le second aux changements de qualités que le signal peut subir selon l'espace ou le temps. Nous employons le terme «sonorité» pour l'aspect qualitatif, réservant celui de «timbre» pour l'aspect d'identification.

Les expériences de manipulation des sons et de synthèse ont montré l'importance de la composante temporelle (transitoires d'attaque, évolution dynamique) dans la caractérisation du timbre. Des bruits apparemment indésirables, comme le bruit de cavité de la caisse de la guitare, la percussion de la caisse du piano ou les bruits de sifflements à l'attaque du son de flûte font partie intégrante de la forme acoustique mémorisée et doivent être reproduits dans une synthèse crédible (figure 10).

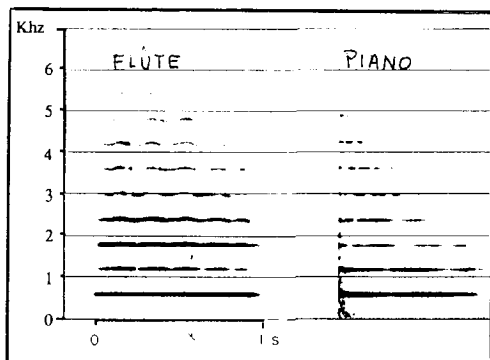


Figure 10 : L'analyse au sonagramme de la même note (Ré4) jouée sur une flûte et sur un piano permet de mettre en évidence les caractéristiques élémentaires du timbre comme forme spectro/temporelle. Le nombre d'harmoniques est similaire, mais l'évolution temporelle est radicalement différente. Observer l'établissement du son (rapidité et bruit de choc au piano) et l'extinction. Noter aussi le vibrato de la flûte.

Il n'est pas facile d'apprécier de manière objective la sonorité d'une source, en raison de la complexité du rayonnement dans l'espace. Supposons qu'un instrument produise une note parfaitement stable : le

signal que nous recueillons avec un microphone change considérablement selon les divers points de l'espace. Où faut-il placer le microphone ? Comment rendre compte à la fois de la position d'écoute du musicien et de celle de l'auditeur ? Le problème se complique si l'inhomogénéité de la salle d'écoute est grande. On conçoit combien il est difficile de comparer deux instruments : même joueur quasiment immobile, même stricte position des microphones, etc. Paradoxalement, l'écoute naturelle est binaurculaire (!) et de plus nous bougeons constamment la tête. Nous sommes donc capables d'élaborer, à partir des signaux différents arrivant sur les deux oreilles et fluctuant temporellement, une image sonore mentale unique et cohérente, à partir de laquelle nous formulons un jugement de qualité.

La deuxième difficulté est liée à la diversité de sonorité des sons d'un même instrument, donc de même timbre [11]. De même que tous les membres d'une famille possèdent des traits communs tout en étant des individus très différents, de même l'analyse acoustique de toutes les notes d'un instrument montre une étonnante diversité de spectres, même lorsque le musicien cherche à jouer de façon la plus homogène possible (figure 11). Quel degré de similarité est le plus apprécié, quelle variété introduire, comment caractériser l'inhomogénéité ? Les études en cours ne permettent pas encore de répondre à ces questions.

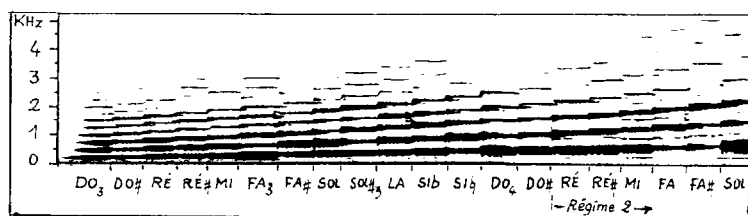


Figure 11 : Analyse au sonographe d'un fragment de gamme chromatique jouée sur une flûte traversière. Il s'agit d'une collection de sons appartenant bien à la même famille de timbre mais présentant d'importantes variations de sonorités (comparer par exemple, les spectres du Fa3 et du sol#3).

Le timbre et la sonorité sont sans doute les attributs les plus difficiles à traduire et les plus fragiles perceptivement. Ils dépendent plus ou moins consciemment de l'influence des autres sens : un instrument aux belles formes et bien décoré sonne «mieux» que son jumeau acoustique, brut de fabrication. Enfin, il est plus d'une salle de concert dont l'acoustique a été unanimement reconnue meilleure après que les sièges en ont été rendus plus confortables !

CONCLUSION

Nous avons exposé à grands traits quelques-unes des principales caractéristiques de la perception auditive pour des signaux musicaux. L'analyse de celles-ci va considérablement s'affiner dans les années qui viennent. Le développement des moyens de synthèse permettant de pratiquer des tests avec des signaux complexes, l'affinage des techniques de passation et de dépouillement de ces tests, l'étude des bruits de notre environnement et surtout les observations des musiciens pratiquant eux-mêmes la synthèse musicale laissent espérer d'importantes découvertes en psychoacoustique.

BIBLIOGRAPHIE

- [1] E. LEIPP - *Acoustique et musique*.
Masson Ed., Paris, 1971.
Notions de physique, et perception auditive ; étude du fonctionnement des instruments de musique. Ouvrage rédigé pour des musiciens.
- [2] E. ZWICKER et R. FELDTKELLER - *Psychoacoustique*.
Masson Ed., 234 pages, Paris, 1981.
Le point sur la psychoacoustique de laboratoire ; comportement de l'oreille à l'écoute de sons purs, de bruits blancs et d'impulsions.
- [3] M.-C. BOTTE, G. CANEVET, L. DEMANY et C. SORIN - *Psychoacoustique et perception auditive*.
Éditions INSERM/SFA, 144 pages, Paris, 1988.
Études psychophysiques de l'audition.
- [4] E. LEIPP et M. CASTELLENGO - *Du diapason et de sa relativité*.
La revue Musicale n° 294, Richard Masse Ed., Paris, 1977.
Vue d'ensemble des recherches effectuées au Groupe d'Acoustique Musicale sur les problèmes techniques et perceptifs afférents au diapason.
- [5] E. LEIPP - *La machine à écouter : essai de psychoacoustique*.
Masson Ed., 260 pages, Paris, 1977.
Étude du fonctionnement du système auditif et des modes de perception des messages sonores : interprétation de cas concrets (parole, musique).
- [6] A. MOLES - *Théorie de l'information et perception esthétique*.
Denoël Ed., 326 pages, Paris, 1972.

- Exposé des éléments d'une esthétique informationnelle applicable aux messages littéraires, musicaux et picturaux.
- [7] P. GUILLAUME - *La psychologie de la forme*.
Flammarion Ed., Coll. Champs, 260 pages, Paris, 1979 (première édition 1937).
Ouvrage de synthèse écrit par le spécialiste français de la théorie de la forme.
- [8] M. CASTELLENGO - *Les Sources acoustiques*.
Le livre des techniques du son, T.I., Ed. Fréquences, pages 45-70, Paris, 1986.
Production, rayonnement, timbre et sonorité des sources musicales.
- [9] J.-M. ARAN et coll. - *Physiologie de la cochlée*.
Éditions INSERM/SFA, 176 pages, Paris, 1988.
Le point des spécialistes sur la physiologie cochléaire. Lire l'avant-propos.
- [10] J.-Cl. RISSET - *Perception, environnement, musiques*.
Revue Inharmoniques n° 3, Christian Bourgois Ed., Paris, 1988, pages 10-42.
Réflexions approfondies d'un musicien, par ailleurs spécialiste de la synthèse sonore.
- [11] M. CASTELLENGO - *La perception auditive des sons musicaux*.
Psychologie de la musique, à paraître aux P.U.F., Paris, début 1994.
Perception des sons complexes réels ; exposé synthétique à l'usage des psychologues.
- [12] P. LINDSAY et D. NORMAN - *Traitement de l'information et comportement humain*.
Vigot Ed., 754 pages, Paris/Montréal, 1980.
Exposé très complet et très accessible des principales fonctions de la perception humaine.