



National Library
of Canada

Bibliothèque nationale
du Canada

Canadian Theses Service

Services des thèses canadiennes

Ottawa, Canada
K1A 0N4

CANADIAN THESES

THÈSES CANADIENNES

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30.

**THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED**

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30.

**LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE**

VERS UNE LOGIQUE DE LA CONNAISSANCE SONORE

Thèse de doctorat
présentée à
l'École des Études Supérieures
de l'Université d'Ottawa

par

Jean-Pierre Delage



Jean-Pierre Delage, Ottawa, Canada, 1987.

Permission has been granted to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film.

The author (copyright owner) has reserved other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without his/her written permission.

L'autorisation a été accordée à la Bibliothèque nationale du Canada de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur (titulaire du droit d'auteur) se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation écrite.

ISBN 0-315-36479-3



UNIVERSITÉ D'OTTAWA
UNIVERSITY OF OTTAWA

A mes parents.

L'immense voûte du ciel et tout ce qui orne la terre, en un mot tous les corps qui composent la puissante forme du monde n'ont de substance que dans notre esprit... Aussi longtemps qu'ils ne sont pas réellement perçus par moi ou n'existent pas dans mon esprit ou dans celui d'aucun esprit créé, on peut considérer qu'ils n'ont aucune existence ou subsistent seulement dans quelque Esprit éternel.

Berkeley

RÉSUMÉ

Ce travail porte sur la connaissance des sons et se veut une tentative d'explicitation des processus cognitifs qui sous-tendent cette connaissance. Nous commençons par prendre position quant au phénomène de la connaissance et au phénomène de la connaissance scientifique. Selon notre conception de la connaissance scientifique, la connaissance sonore représente une *entité problématique*. Cette entité est problématique parce que les éléments dont elle se compose sont en état de *carence attributive*. Il nous faut donc chercher à enrichir ces éléments constitutifs que sont le milieu, l'organisme et l'expérience sonore. Cette démarche nous amène à concevoir la connaissance sonore comme un processus de connaissance des différentes manifestations énergétiques du milieu. Ce processus s'effectue par le biais d'un organisme composé d'une hiérarchie de structures de caractérisation opérant des groupements temporels de différents types et produisant des entités sonores à tous ces niveaux de caractérisation de l'organisme. Ces groupements temporels peuvent être conçus comme des 'mises en relation temporelles' *invariantes* de 'ce qui se manifeste' à chaque niveau du processus de connaissance sonore. Les entités sonores (les invariants sonores) de tous les niveaux de connaissance de l'organisme sont dotées d'attributs de caractérisation et le résultat final de ce processus de connaissance s'exprime sous la forme des entités expérientielles sonores (les sons) qui se manifestent à notre conscience.

REMERCIEMENTS

Je tiens à exprimer ma gratitude à:

mon directeur de thèse, Claude Lamontagne Ph.D., qui a su initier, entretenir et diriger ce travail de recherche avec un effort de "résistance" soutenu et une infatigable passion pour la connaissance sous toutes ses formes,

- mes collègues, Jean-Roch Beausoleil et Pierre Tremblay, pour leurs multiples conseils de mathématicien, biologiste, informaticiens, psychologues et surtout pour leur présence amicale,

- ma compagne Sylvie Laurin pour avoir cru dans ce travail et pour m'avoir "supporté" durant cette dernière année de rédaction.

TABLE DES MATIERES

Introduction.....1

1. Connaissances.....4

 1.1 La connaissance.....6

 1.2 La connaissance scientifique.....11

 1.3 La connaissance scientifique de la connaissance sonore.....17

2. La connaissance sonore.....22

 2.1 Le milieu.....23

 2.2 L'organisme.....27

 2.2.1 Les structures d'interface.....29

 2.2.2 Les structures d'analyse spectrale.....35

 2.2.3 Les structures supérieures.....49

 2.3 L'expérience.....55

 2.4 Vers une logique de la connaissance sonore.....68

3. Le processus de connaissance sonore de l'organisme.....74

 3.1 Temps et connaissance sonore.....78

 3.2 Le premier niveau du processus de connaissance sonore.....88

 3.3 Le second niveau du processus de connaissance sonore.....95

 3.3.1 Le calcul des valeurs d'attribut des formes variationnelles..112

 3.3.2 La simulation computationnelle.....140

 3.4 Le troisième niveau du processus de connaissance sonore.....151

 3.5 Vers une logique du processus de connaissance sonore.....161

Conclusion.....170

Bibliographie.....173

Annexe 1:
 La simulation computationnelle de la structure d'analyse spectrale.....177

INTRODUCTION

La thèse est divisée en trois chapitres. Au premier chapitre, nous prenons position quant à la nature de la connaissance et quant à la nature de la connaissance scientifique, ce qui nous permet de "cadre" l'objet précis de ce travail: la connaissance sonore. Cette prise de position repose fortement sur la notion d'expérience. Nous posons l'expérience comme première et nous la concevons comme *une séquence d'entités expérientielles dotées d'attributs*. La connaissance devient ainsi *une entité expérientielle dotée d'attributs* et se caractérise, entre autres, comme une entité *problématique*, ce caractère problématique originant d'un état de "carence attributive" au niveau de ses trois entités constituantes: l'expérience, l'organisme et le milieu. La connaissance scientifique peut être conçue comme une tentative systématique de production d'entités expérientielles *non-problématiques*. La connaissance scientifique résulte de l'identification d'entités problématiques (les problèmes scientifiques) et de leur résolution en entités non-problématiques (les théories scientifiques). La *connaissance sonore* est enfin définie comme un problème scientifique qui appelle une solution sous la forme d'une théorie scientifique, problème auquel s'adressent les chapitres 2 et 3 de ce travail.

Plus spécifiquement, au deuxième chapitre de ce travail, nous étudions successivement les trois entités constituantes de l'entité 'connaissance sonore': le milieu, l'organisme et l'expérience sonore. Les attributs de ces trois entités sont successivement considérés dans la perspective de leurs interrelations respectives. Le milieu est défini comme le lieu d'origine de la chaîne

interactive qui conduit ultimement à l'expérience des sons. Certaines caractéristiques physiques du milieu sont mises en évidence pour nous amener à cerner plus précisément la tâche de connaissance qu'implique la connaissance de 'ce qui se manifeste' dans le milieu. Cette tâche de connaissance est l'apanage de l'organisme qui produit l'expérience des sons par le biais d'un ensemble hiérarchique de structures de caractérisation en interaction avec 'ce qui se manifeste' dans le milieu. Nous étudions les différentes caractéristiques des structures de caractérisation des organismes biologiques et artificiels et essayons de dégager leurs principes de fonctionnement. Ce deuxième chapitre se termine finalement par l'examen des différentes caractéristiques de l'expérience sonore dans les perspectives conjointes des travaux de la psychophysique et de la 'gestalt theorie'. Ceci nous amène à concevoir l'expérience sonore comme une hiérarchie d'entités sonores dotées d'attributs. Nous établissons alors un lien entre la hiérarchie de structures de caractérisation et la hiérarchie d'entités sonores. Les différentes considérations de ce deuxième chapitre forment les éléments de notre conception de la logique sous-jacente au phénomène de la connaissance sonore. Cette logique nous amène à concevoir le *processus de connaissance sonore de l'organisme* comme l'élément central de cette forme de connaissance, processus sur lequel nous nous penchons de façon plus exclusive au troisième chapitre.

Le troisième chapitre est donc consacré à l'élaboration de la logique sous-jacente au processus de connaissance sonore de l'organisme. Nous concevons ce processus comme étant orienté vers la production de ces *invariants* que sont les entités sonores des multiples niveaux de caractérisation de

l'organisme. Ces invariants sonores sont le résultat des différents types de groupements temporels que produisent les structures de caractérisation de chaque niveau de la hiérarchie structurale. Ces groupements temporels sont le reflet de 'mises en relation temporelles' invariantes de 'ce qui se manifeste' à chaque niveau du processus. Les résultats de ces groupements s'expriment sous la forme des attributs de caractérisation des entités sonores de chaque niveau. Les groupements temporels, les entités et les attributs de caractérisation de ces entités sont définis pour les deux premiers niveaux du processus de connaissance sonore et sont soumis à l'épreuve d'une simulation computationnelle. Nous présentons nos hypothèses sur la nature de ces mêmes éléments pour le troisième niveau du processus, mais ces hypothèses ne prennent toutefois pas la forme d'une simulation computationnelle. Nous terminons enfin ce troisième chapitre en considérant globalement le processus de connaissance sonore dans la perspective des trois premiers niveaux de caractérisation.

CHAPITRE 1

CONNAISSANCES

Ce premier chapitre est le fondement sur lequel repose l'ensemble des idées qui seront développées aux chapitres 2 et 3 de ce travail: nous y prenons position sur le phénomène de la connaissance. Cette prise de position s'articule essentiellement autour de la notion d'expérience, notion qui nous apparaît comme première, et ce, dans le sens précis d'une proposition acceptée et indémontrable. De la notion d'expérience, nous dérivons les concepts de 'connaissance' et de 'connaissance scientifique' à partir desquels nous situons finalement l'objet de ce travail de recherche: le problème scientifique de la connaissance sonore. Une telle réflexion, à caractère plus philosophique, peut paraître surprenante dans le contexte d'un travail en psychologie cognitive; elle nous est apparue, pour notre part, essentielle et doublement justifiable. Nous croyons premièrement que tout effort scientifique se doit de reposer sur des bases épistémologiques explicites, et en cela la psychologie ne fait pas exception. Mais à cette exigence générale s'en ajoute une autre qui est, à notre avis, beaucoup plus contraignante. S'il est possible, à la rigueur, de faire abstraction de la nature de la connaissance dans une démarche de recherche qui porte sur d'autres phénomènes que les phénomènes de connaissance, cela devient beaucoup plus difficile et inapproprié lorsque l'objet même de cette démarche de recherche est la connaissance et les processus qui la sous-tendent. Nous croyons que, pour

connaître la connaissance, il est nécessaire d'expliciter certaines positions philosophiques fondamentales. Plus encore, nous pensons que c'est à l'intérieur même de ces prises de position fondamentales que peuvent être trouvés certains éléments nécessaires à l'évolution de la problématique de la connaissance en général, et, plus spécifiquement, à l'évolution de la problématique de la connaissance sonore.

1.1 La connaissance

L'expérience est première et s'exprime sous la forme d'une succession d'entités expérientielles.

Nous vivons notre présence au monde sous la forme d'une séquence d'événements expérientiels de toutes natures: perceptions, idées, émotions, rêves etc.. Chacun de ces événements expérientiels peut être considéré comme une entité propre qui se distingue des autres entités par ses *attributs* spécifiques. Un son particulier a une intensité, une hauteur, un timbre; une entité visuelle a une forme, une couleur, un mouvement. Les émotions sont d'une certaine nature, agréables ou désagréables, violentes ou diffuses. Les idées portent sur la connaissance, la politique, l'art ou la science. Notre présence au monde est succession d'entités expérientielles mais elle est également succession d'entités expérientielles *différenciées*, d'entités qui se distinguent les unes des autres par leurs attributs spécifiques. Parmi l'ensemble de ces attributs, il en est un qui peut être rattaché à toutes les entités expérientielles: leur caractère *problématique* ou *non-problématique*. Une première rencontre avec une discipline jusqu'alors inconnue, une conversation dans une langue non-familière ou l'impression d'avoir un nom "sur le bout de la langue" sont autant d'exemples d'entités expérientielles problématiques, alors que l'observation du vol d'un oiseau, une émotion intense ou la reconnaissance d'une musique familière constituent plutôt des entités expérientielles à caractère non-problématique.

L'expérience: succession d'entités expérientielles? Mais comment expliquer ce flux, comment expliquer ce déroulement de l'expérience? L'expérience est expérience de quoi? L'expérience est expérience de qui? Poser ces questions, tenter d'établir des relations entre l'entité expérience et d'autres entités expérientielles, c'est rencontrer une entité expérientielle *problématique*: l'entité connaissance. Qu'est-ce que la connaissance? Quelle relation existe-t-il entre connaissance et expérience? Nous dirons de l'expérience qu'elle est un *produit* de la connaissance, le produit des interactions entre un milieu et un organisme. L'organisme, le milieu et les interactions qui les relient forment l'essence d'un *substrat expérientiel* qui "supporte" l'expérience: elles sont ce sans quoi la "réalité" de l'expérience ne saurait exister.

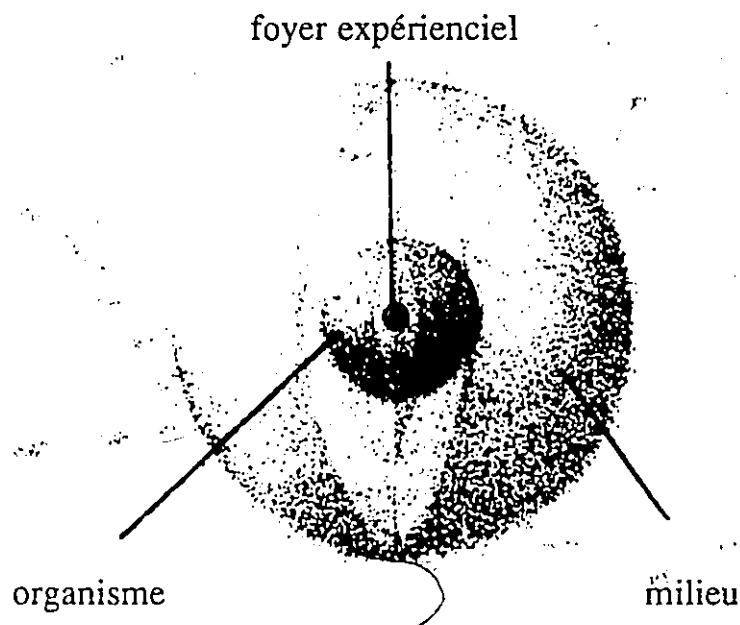


Fig. 1 Le substrat expérientiel

On peut se représenter l'organisme et le milieu comme deux sphères dont l'une, l'organisme, est comprise au centre de l'autre, le milieu, qui est plus étendue (Fig. 1). Au centre de la plus petite sphère se situe le "lieu" de l'expérience, lieu que nous désignerons comme étant le *foyer expérientiel*. Ce qui distingue fondamentalement les deux sphères, c'est, d'une part, la cohésion importante et momentanée des éléments du substrat expérientiel qui constituent l'organisme et c'est, d'autre part, l'emplacement du foyer expérientiel *dans* l'organisme. Ces deux éléments nous amènent à considérer la dynamique interactionnelle du substrat expérientiel sous une optique "polarisée" qui se caractérise par l'existence de "courants" interactionnels se dirigeant *vers* le foyer expérientiel et *hors* du foyer expérientiel. Le courant interactionnel dirigé *vers* le foyer expérientiel produit l'expérience sous la forme des entités expérientielles que nous appréhendons, alors que le courant interactionnel dirigé *hors* du foyer expérientiel produit des effets dans l'organisme et dans le milieu, effets qui peuvent être conséquents à "l'expression" d'une entité expérientielle au foyer expérientiel mais effets qui ne produisent pas d'entités expérientielles "en soi". Ces effets ne peuvent éventuellement être connus que sous la forme d'entités expérientielles qui sont produites par le biais d'un courant expérientiel qui se dirige *vers* le foyer expérientiel. Cela signifie que l'expérience est toujours le fruit d'une chaîne interactive dirigée *vers* le foyer expérientiel, chaîne interactive distincte de celle émergeant de ce même foyer. Dans cette optique, nous sommes en quelque sorte directement "inconscients" de nos actions et uniquement indirectement "conscients" de l'effet de ces actions sur le courant interactionnel qui se dirige *vers* le foyer expérientiel.

Les courants interactionnels qui sont caractéristiques du substrat expérientiel peuvent s'étendre au milieu environnant comme il peuvent se limiter à l'organisme de connaissance. 'Réfléchir', par exemple, est le fruit d'une chaîne interactive qui se limite à l'organisme, alors que 'déplacer un objet' est typique d'une chaîne interactive qui s'étend au milieu. Le déroulement de toute portion de la séquence expérientielle peut ainsi être conçue comme une succession ininterrompue de courants interactionnels à l'intérieur du substrat expérientiel: le vent forcé, le bateau gité, la lumière se réfléchit sur le paysage et sur le bateau; la lumière réfléchie interagit avec la rétine, le changement de position du bateau dans l'espace modifie l'état des récepteurs d'équilibre, le bruit du vent s'amplifie et atteint le tympan. L'oeil, l'oreille et les récepteurs d'équilibre interagissent avec le système nerveux périphérique qui interagit lui-même avec le système nerveux central. Celui-ci décrète des commandes motrices qui se propagent dans les différents membres; changement de position des jambes pour retrouver l'équilibre, déplacement des bras pour modifier la course, du voilier. Le bateau se redresse, le bruit du vent se modifie et la chaîne interactive se poursuit. Le foyer expérientiel est point de convergence et point de diffusion des courants interactionnels. Les manifestations expérientielles se situent à ce point de jonction entre la perception¹ et l'action. Elles sont l'émergence, l'excroissance de ce lieu précis où s'achèvent et où débutent les interactions entre les entités du substrat expérientiel.

¹ Nous entendons la perception dans un sens très large qui englobe percepts, affects et concepts.

C'est dans ce contexte que l'on peut dire de l'entité expérientielle 'connaissance' qu'elle se constitue de ces entités plus primitives que sont le milieu, l'organisme et l'expérience. Le caractère problématique de la connaissance origine d'un état de "carence attributive". Quels sont donc les attributs "déficients" de l'entité connaissance? Ce sont essentiellement ces trois entités plus primitives et les relations que ces trois entités entretiennent entre elles. Qu'est-ce que l'expérience? Qu'est-ce que l'organisme? Qu'est-ce que le milieu? Quelles relations existe-t-il entre ces trois entités? Voilà les questions auxquelles se doit de répondre une tentative systématique d'explicitation du phénomène de la connaissance.

1.2 La connaissance scientifique

La science peut se définir comme une entreprise de production d'entités expérientielles "vraies". Mais qu'est-ce que la vérité dans le contexte de cette conception de la connaissance que nous venons de décrire? Ce ne peut être une adéquation quelconque entre l'expérience et une "réalité" qui serait distincte de l'expérience puisque l'expérience est première et omniprésente. La réalité n'existe que sous la forme d'entités expérientielles. Il nous faut donc chercher un critère de vérité à l'intérieur même de ce contexte et plus spécifiquement dans les attributs des entités expérientielles. *Nous définirons une entité expérientielle "vraie" comme une entité expérientielle non-problématique.* Il nous faut toutefois distinguer deux sens au mot 'vérité': un sens large ou global où la vérité est la caractéristique des entités expérientielles non-problématiques, et un sens plus restreint, plus local, où le vrai et le faux décrivent certains attributs d'une entité vraie au sens global. A titre d'exemple, 'il pleut' constitue une entité "vraie" au sens global et "vraie" au sens local s'il pleut effectivement. S'il ne pleut pas, 'il pleut' constitue toujours une entité vraie au sens global, i.e. une entité non-problématique, mais fausse au sens local. Dans les deux cas, l'entité expérientielle résultante est non-problématique donc "vraie" au sens le plus global.

A la section précédente, nous avons brossé l'image d'un substrat expérientiel constitué d'entités et de courants interactionnels entre ces entités, courants qui produisent, entre autres, l'expérience (au foyer expérientiel). Considérons la situation suivante: nous regardons la mer et

apercevons sur l'eau une forme noire que nous avons de la difficulté à identifier: nous expérimentons une entité problématique. Pour résoudre cette entité problématique en entité non-problématique, nous décidons de nous approcher de l'objet en question. Nous posons une action et, ce faisant, les attributs de la forme noire se précisent: il s'agit d'un canard. Nous sommes passés d'une entité expérientielle problématique à une entité expérientielle non-problématique. De ce simple exemple, nous pouvons tirer quelques conclusions sur la nature et l'origine des entités problématiques. En premier lieu, ces entités se caractérisent par une certaine "carence" attributive. L'entité existe mais elle est insuffisamment caractérisée et c'est ce qui produit cette impression problématique. Deuxièmement, une entité problématique se résorbe par une modification du contexte interactionnel à l'intérieur du substrat expérientiel, en l'occurrence, un rapprochement de l'objet. Finalement, cette modification du contexte interactionnel *peut* être le résultat d'une action de l'organisme (nous nous approchons du canard) mais elle peut également être indépendante de l'action de l'organisme et se produire "au hasard" des interactions qui se produisent dans le milieu (le canard s'approche de nous). Ces modifications constantes du contexte interactionnel qui sont dues à l'organisme et au milieu nous suggèrent d'ailleurs l'image d'un substrat expérientiel qui se modifie et qui se transforme constamment sous l'influence des multiples interactions qui s'y produisent. Pour un organisme, ces modifications constantes du substrat expérientiel se reflètent par des phases successives d'équilibre et de déséquilibre expérientiel qui correspondent aux entités expérientielles problématiques et non-problématiques.

Mais quel est le rapport entre les modifications du contexte interactionnel, les entités problématiques et non-problématiques, et la connaissance scientifique? Nous croyons premièrement que le processus scientifique prend sa source dans la mise en évidence d'entités problématiques abstraites, les problèmes scientifiques, qui sont également, mais à un niveau plus abstrait, en état de "carence attributive". Nous croyons également que ce processus est orienté vers la production d'entités non-problématiques également abstraites: les théories scientifiques. Dans notre conception de la connaissance, les problèmes scientifiques correspondent à des entités abstraites (des concepts) qui mettent en évidence la nature problématique des relations (interactions) entre un certain nombre d'entités expérientielles (d'autres concepts ou des "phénomènes") dont sont tributaires ces entités problématiques. Les théories scientifiques sont également des entités abstraites, mais des entités qui représentent une tentative d'explicitation des relations entre ces mêmes entités qui produisent les problèmes scientifiques.

Les entités expérientielles (et leurs interrelations) qui font l'objet du processus scientifique peuvent être de différentes natures mais ce processus se caractérise par une forte tendance à l'explicitation de "phénomènes", ce que nous appelons, dans notre terminologie, des entités expérientielles perceptuelles. Or, pour nous, il n'existe pas de différence fondamentale entre les entités "concrètes" que sont les phénomènes et les entités plus abstraites que sont les problèmes et les théories scientifiques: elles sont toutes deux de nature expérientielle. Si on accorde tant d'importance, dans la démarche scientifique, à l'explication des entités expérientielles concrètes, cela est dû en partie, croyons-nous, à la nature hautement non-problématique de ces

entités qui originent d'un contexte interactionnel fortement équilibré². Mais cela n'en fait pas pour autant des entités qui résultent d'un contexte interactionnel distinct qui serait plus "valide" ou plus "véridique". Les études sur le développement perceptivo-moteur nous amènent d'ailleurs à penser que l'expérience perceptuelle peut être problématique à certains stades du développement ontogénétique et que c'est à la suite de déséquilibres successifs que se bâtit un contexte interactionnel équilibré à l'intérieur de l'organisme (Piaget, 1967). Mais un contexte équilibré ne signifie pas pour autant que l'expérience résultante sera à tout jamais non-problématique. Les illusions perceptuelles en sont un exemple pour les entités perceptuelles alors que la succession des différentes théories scientifiques³ représentent un exemple équivalent au niveau de l'expérience plus abstraite. Par ailleurs, à un niveau plus commun, chacun de nous a vécu l'expérience de la remise en question d'une croyance fortement ancrée qui a été déséquilibrée par un événement inattendu. Ce qu'il nous faut accepter, comme l'a si bien soutenu Karl Popper (1978), c'est la "quête sans fin" ("unended quest") que suppose la connaissance et l'impossibilité d'une vérité "absolue".

L'esprit du chercheur est donc confronté à des entités problématiques qui émergent de la mise en relation d'autres entités expérientielles et face à ces entités problématiques, il cherche à élaborer des théories qui en sont la

² Nous parlons ici du "contexte" interactionnel que représente l'ensemble des structures perceptivo-motrices. Concrètement, cela veut dire que les différents éléments de ces structures (organes sensoriels et moteurs, neurones et branchements synaptiques) se développent durant la gestation et les premiers mois de la vie et qu'ils conservent par la suite cette organisation.

³ Dont certaines, comme la théorie de Newton, étaient solidement "installées".

résolution en entités non-problématiques. Ces théories permettent au chercheur d'expérimenter, c'est à dire d'enclencher une série d'interactions à l'intérieur du substrat expérientiel, ce qui peut résulter en un nouveau problème ou dans le maintien de la théorie existante. Ce qu'il faut remarquer, c'est que ce processus peut se restreindre à l'organisme du chercheur, auquel cas nous parlerons "d'expérimentation abstraite", ou s'étendre au milieu par le biais de "l'expérimentation concrète". Lorsque le chercheur réfléchit à un problème, il agit sur son organisme en modifiant le contexte interactionnel de ses idées et, ce faisant, il peut en arriver à transformer une entité problématique en entité non-problématique (résoudre le problème). Lorsque le chercheur procède à de l'expérimentation concrète, il ne fait qu'étendre la boucle interactionnelle au milieu en essayant de modifier le contexte interactionnel en fonction de ses attentes⁴.

Ce qui devrait caractériser fondamentalement le processus scientifique, c'est la reconnaissance d'une telle dynamique et la volonté de la maximiser. Comment peut-on y arriver? En préconisant une méthode, c'est à dire une manière systématique pour accélérer le cycle problèmes --> théories --> problèmes. Cette méthode devrait reposer d'une part sur ce qui nous fait accepter le caractère non-problématique d'une entité abstraite (théorie) et s'appuyer d'autre part sur ce qui nous permet de reconnaître qu'il existe une entité problématique (problème). L'aspect le plus important de cette méthode nous semble être le recours au formalisme. Le recours à un formalisme approprié permet l'expression non-ambigüe des interactions entre

⁴ La plus ou moins grande "extension" de la boucle interactionnelle pourrait toutefois être considérée comme un critère de "solidité" des théories scientifiques.

différentes entités. Or les problèmes et les théories se constituent justement d'entités en interaction. L'expression formelle de ces interactions ne peut que faciliter la reconnaissance du caractère problématique ou non-problématique d'une entité abstraite qui se constitue de ces entités en interaction. Mais toutes les interactions envisageables ne se prêtent pas nécessairement à une tentative de formalisation, tentative qui doit plutôt être envisagée dans la perspective d'un but à atteindre et non comme une fin en soi qui exclut toute autre type d'explication. Le processus que nous venons de décrire donne un rôle privilégié au travail de connaissance scientifique qui se produit à l'intérieur d'un *seul* couple organisme-milieu. Il est évident toutefois que la démarche scientifique ne se restreint pas à un seul chercheur. Celui-ci s'inscrit dans un processus qui l'englobe à titre de producteur d'entités problématiques et non-problématiques, entités qu'il communique sous différentes formes à la communauté scientifique. En s'inscrivant dans ce contexte scientifique global, le chercheur alimente une boucle interactionnelle beaucoup plus étendue qui est à l'image, croyons-nous, du processus évolutif de l'ensemble du substrat expérientiel.

9

1.3 La connaissance scientifique de la connaissance sonore

Ces différentes considérations sur la connaissance et la connaissance scientifique nous amènent maintenant à l'objectif principal de ce travail: appliquer le processus scientifique que nous venons de décrire à *un* courant interactionnel particulier parmi l'ensemble des courants interactionnels qui sont caractéristiques d'un organisme de connaissance: celui qui est spécifique à l'entité expérientielle 'connaissance sonore'⁵. Dans une approche scientifique des processus cognitifs, connaissance, science et problème sont intimement liés puisque nous voulons appliquer un processus scientifique, un processus de connaissance donc, à un problème qui est également de nature cognitive: le problème de la connaissance sonore. L'omniprésence de ce "contexte cognitif" se reflétera donc à la fois dans la *nature* même de l'entité à connaître, la connaissance sonore, et dans la *démarche* de connaissance scientifique de cette entité.

A partir de notre prise de position sur la connaissance, la *nature* de la connaissance sonore sera définie par les trois entités (et les relations entre ces entités⁶) que sont:

⁵ A titre d'exemples, la connaissance visuelle ou la connaissance tactile représentent d'autres exemples de courants interactionnels qui sont orientés *vers* le foyer expérientiel alors que la connaissance motrice peut être considérée comme un courant interactionnel qui est orienté *hors* du foyer expérientiel, i.e vers le milieu.

⁶ On peut considérer les interrelations entre ces trois entités comme des attributs de chacune de ces entités.

- le milieu, c'est à dire les éléments du milieu qui sont impliquées dans la connaissance sonore,
- l'organisme, c'est à dire les éléments de l'organisme qui sont impliquées dans la connaissance sonore,
- l'expérience, c'est à dire les éléments de l'expérience qui sont impliquées dans la connaissance sonore.

La connaissance sonore est une entité problématique: elle est en état de carence attributive et cet état origine d'une caractérisation insuffisante des attributs constitutifs de la connaissance sonore. Quels sont ces attributs? Ce sont essentiellement les trois entités que nous venons de mentionner. Notre démarche de connaissance scientifique de la connaissance sonore visera donc à enrichir "attributivement" les caractéristiques du milieu, de l'organisme et de l'expérience qui sont impliquées dans le phénomène de la connaissance sonore. Nous procéderons en deux temps. Le chapitre 2 sera consacré à "l'enrichissement attributif" des trois composantes de la connaissance sonore. Pour ce faire, nous suivrons la direction "naturelle" du courant interactionnel sous-jacent à cette forme de connaissance. Nous étudierons tout d'abord les caractéristiques physiques du "milieu sonore". Nous considérerons par la suite les différentes composantes de "l'organisme sonore", dans le contexte de leur interaction avec ces caractéristiques du milieu physique, et nous tenterons finalement de cerner la nature particulière de "l'expérience sonore" en tant que produit de ces interactions entre le milieu physique et l'organisme. A la fin de ce deuxième chapitre, ceci nous amènera à concevoir le processus de connaissance sonore de l'organisme comme étant l'élément crucial de la connaissance sonore. Le chapitre 3 sera consacré à "l'enrichissement attributif" de ce processus de connaissance de

l'organisme. Nous y développerons un certain nombre de concepts théoriques qui nous permettront d'effectuer une simulation computationnelle d'une certaine partie de ce processus. Notre démarche de connaissance scientifique de la connaissance sonore sera donc théorique, au sens d'un cycle interactionnel abstrait qui se limite à notre organisme, mais elle se voudra également participante d'un cycle interactionnel concret, par le biais de l'expression computationnelle de ces différents concepts théoriques. Les machines computationnelles que sont les ordinateurs peuvent être en effet assimilées à des organismes malléables qu'il nous est possible de créer et de configurer⁷ en fonction de nos hypothèses sur les processus de connaissance sonore. Dans cette perspective, l'expression computationnelle de ces hypothèses s'insère dans un courant interactionnel qui est dirigé vers le milieu et dont l'effet de retour, sous la forme du comportement de cet organisme artificiel, peut être comparé aux résultats classiques d'une expérimentation traditionnelle.

L'ensemble de notre démarche de recherche sur la connaissance sonore se caractérise par un très grand "désir" de généralité. *Nous sommes en effet profondément convaincu que les différents contextes de connaissance sonore reflètent une seule logique de connaissance sonore.* C'est partiellement en fonction de ce désir de généralité que notre argumentation a été développée et continuera à être développée en utilisant une terminologie nouvelle et inusitée. C'est dans cette optique de généralité que nous essaierons d'inclure toutes les "formes" de connaissance sonore. Nous étudierons des structures

⁷ Par le biais de la programmation ("software") ou de la conception matérielle de machines computationnelles dédiées ("hardware").

cognitives biologiques comme la cochlée et le système nerveux auditif mais nous étudierons aussi des structures *cognitives* "artificielles" comme les analyseurs spectraux, les microphones et les structures mathématiques d'analyse spectrale. Le produit du processus de connaissance de ces structures physiques et mathématiques sera assimilé à des entités *cognitives* au même titre que le produit du processus de connaissance des organismes biologiques. Toute notre démarche d'explicitation visera à montrer que ces différents contextes de connaissance sonore représentent les différentes facettes d'une seule logique de connaissance sonore. Cette logique peut émaner d'un médium gazeux, liquide ou solide; elle peut être le fait d'êtres vivants ou d'organismes artificiels, s'incorporer dans une cochlée, un analyseur spectral ou un système nerveux mais elle reste toujours, croyons-nous, une *seule* logique.

Ceci complète le premier chapitre de ce travail: Nous avons vu que:

- l'expérience était première et pouvait être décrite comme une succession d'entités expérientielles dotées d'attributs,
- parmi ces attributs, il en est un qui peut être rattaché à toutes les entités expérientielles: leur caractère problématique ou non-problématique,
- la connaissance représente une entité problématique, ce caractère problématique originant d'une "carence attributive" au niveau de ses trois entités constituantes que sont le milieu, l'organisme et l'expérience,
- la connaissance scientifique peut être conçue comme une entreprise systématique de production d'entités "vraies", i.e. d'entités non-problématiques,
- la méthode scientifique procède de l'identification d'entités problématiques (les problèmes scientifiques) et de leur résolution en

entités non-problématiques, le recours au formalisme permettant la maximisation de ce processus,

- l'objectif de ce travail est d'appliquer le processus scientifique à une entité problématique: la connaissance sonore.

- nous essaierons de tout d'abord de définir la logique sous-jacente au phénomène de la connaissance sonore dans son ensemble en procédant à un "enrichissement attributif" de ses trois entités constituantes que sont les composantes "sonores" du milieu, de l'organisme et de l'expérience (chapitre 2),

- nous essaierons finalement de définir la logique sous-jacente au processus de connaissance sonore de l'organisme (chapitre 3).

CHAPITRE 2

LA CONNAISSANCE SONORE

L'entité problématique 'connaissance sonore' émerge de la mise en rapport des entités 'milieu', 'organisme' et 'expérience sonore'. L'objectif de ce chapitre est de contribuer à "déproblématiser" la connaissance sonore en décrivant les attributs de ces trois entités dans la perspective de leurs contributions respectives au phénomène de la connaissance sonore. Nous essaierons tout d'abord de définir les caractéristiques du milieu qui participent au phénomène de la connaissance sonore et nous nous situerons, pour ce faire, dans le contexte du discours physique sur les manifestations énergétiques qui se produisent dans un médium "sonore". Nous considèrerons par la suite les différentes composantes de l'organisme qui sont impliquées dans le phénomène de la connaissance sonore, et ce, dans la double perspective de l'insertion de l'organisme dans le milieu et de son rôle de médiateur de l'expérience sonore. Nous essaierons finalement de décrire l'expérience sonore en la considérant comme une production de l'interaction entre l'organisme et le milieu.

2.1 Le milieu

Quelle est la nature des événements qui se produisent dans le milieu et qui conduisent, par le biais de l'organisme, à l'expérience sonore? La physique nous décrit l'univers comme un continuum spatio-temporel qui est "habité" par de la matière en interaction, interaction qui s'exprime sous la forme de différentes manifestations énergétiques (gravitationnelles, chimiques, mécaniques, nucléaires, thermiques etc.)⁸. La forme énergétique qui est à l'origine de la chaîne interactive de la connaissance sonore est une manifestation énergétique de type mécanique. Dans un médium comme l'air, les particules aériennes rebondissent et se frappent les unes contre les autres. Plus le nombre de collisions est important, plus la *pression* du médium aérien augmente. Lorsqu'un corps plongé dans ce médium se meut, il provoque des variations de la pression ambiante autour de lui et ces variations de pression se propagent dans toutes les directions sous la forme d'ondes longitudinales ("longitudinal wave"). Ces ondes se propagent dans le médium en imprimant aux particules du médium une série de mouvements d'aller-retour qui sont le reflet du mouvement du corps. En chaque point de l'espace du médium, l'effet combiné de toutes les ondes longitudinales est additif et se reflète dans la pression locale. En d'autres termes, s'il existe plusieurs corps en mouvement dans le milieu sonore, les valeurs de pression en chaque point de l'espace seront égales à la somme locale de toutes les ondes qui proviennent de ces corps. En s'éloignant du corps-source d'origine, les ondes

⁸ Les définitions des concepts physiques de ce chapitre sont tirées de Van Nostrand's Scientific Encyclopedia (1983).

longitudinales s'affaiblissent graduellement⁹ et peuvent subir différentes déformations (diffraction, réflexion, absorption) du fait de leur rencontre avec d'autres corps physiques.

Ce ballet constant des particules aériennes est le reflet des différentes manifestations énergétiques qui se produisent dans le milieu: collision, déplacement, contraction, frottement, expansion d'objets physiques etc. Ces objets reçoivent de l'énergie sous différentes formes et émettent également de l'énergie sous différentes formes, dont la forme mécanique qui nous intéresse. Dans ce transfert énergétique, ces corps imposent, en quelque sorte, leurs structures particulières à l'énergie. C'est cette énergie "modulée" qui se transmet par le biais du mouvement des particules du médium de transmission, et qui provoque les variations concomitantes de pression locale. Mais qu'est-ce plus précisément que l'énergie et comment peut-on la mesurer? Celle-ci se définit habituellement comme "la capacité de faire un travail", c'est à dire la capacité d'appliquer une force sur une certaine distance. Le déplacement des particules aériennes est le reflet de toutes ces forces qui se déploient ponctuellement dans tous les "coins" du médium aérien et, comme nous venons de le voir, ce mouvement se traduit par les variations de pression du médium aérien. Or la valeur de l'énergie en chaque point de l'espace du milieu sonore à un instant donné, i.e. la *puissance instantanée*, est égale au carré de la pression en ce même point (voir, par exemple, Leshowitz (1978) pp. 89-93). En d'autres termes, la mesure instantanée de la pression en un point de l'espace nous donne accès à la

⁹ L'amplitude des variations de pression causées par le mouvement d'un corps-source est inversement proportionnelle à la distance de ce corps-source.

quantité d'énergie qui se manifeste à cet instant et en cet endroit précis où nous la mesurons. Les corps physiques occupant le milieu engendrent donc des variations de pression dans le médium aérien et ces variations de pression se propagent dans toutes les directions sous la forme d'ondes longitudinales. En chaque point de l'espace, l'effet combiné de toutes ces ondes correspond à la pression et la pression au carré est égale à la quantité d'énergie qui se manifeste en ce point de l'espace à cet instant précis. Puisque les ondes longitudinales se propagent dans toutes les directions, tous les objets physiques qui en "émettent" contribuent aux variations de pression locale. *La quantité d'énergie en chaque point de l'espace est donc la résultante de toutes les ondes longitudinales qui originent des objets physiques en mouvement dans l'environnement immédiat.* La nature omnidirectionnelle des ondes longitudinales fait en sorte que, sur un espace assez restreint¹⁰, les manifestations de l'énergie sonore ne sont pas très différentes d'un point spatial à un autre point spatial assez rapproché puisque chacun de ces points spatiaux "subit" approximativement les mêmes ondes longitudinales¹¹. Par ailleurs, comme ces ondes prennent un certain temps à voyager, il existe un décalage temporel entre les valeurs de pression qui se produisent en un point de l'espace à un instant donné et les valeurs de pression quasi-identiques qui se produiront en un point un peu plus éloigné quelques instants plus tard.

¹⁰ 1 pied cube, par exemple, ou approximativement la dimension d'une tête humaine.

¹¹ S'il n'existe pas d'obstacle entre ces deux points et dans un environnement où les phénomènes de réflexion, diffraction et d'absorption sont réduits à leur minimum.

Ceci complète ce bref survol des caractéristiques du milieu qui sont impliqués dans la chaîne interactive de la connaissance sonore.

Résumons:

- le milieu est "occupé" par des objets physiques qui réagissent à différentes formes de manifestations énergétiques,

- ces objets réagissent à ces manifestations en leur imposant leur structure particulière et en "émettant" à leur tour différentes formes de manifestations énergétiques, dont les manifestations de type mécanique qui sont caractéristiques de la chaîne interactive de la connaissance sonore,

- ces manifestations énergétiques se propagent dans le milieu sous la forme d'ondes longitudinales qui se propagent dans toutes les directions,

- en chaque point de l'espace, l'effet de ces ondes longitudinales est additif et se reflète dans la valeur de pression locale du médium de propagation,

- la valeur de la pression au carré correspond à la quantité de l'énergie qui se manifeste en ce point de l'espace, à cet instant du temps (puissance instantanée).

Ces différentes caractéristiques correspondent aux attributs de l'entité expérientielle 'milieu'. Ces caractéristiques peuvent avoir des connotations cognitives importantes lorsqu'on les considère dans la perspective de l'interaction d'un organisme avec le milieu. Nous allons maintenant étudier les attributs de l'entité expérientielle 'organisme' et essayer de voir comment ces attributs peuvent entrer en relation avec les attributs du milieu.

2.2 L'organisme

Quelle relation pouvons-nous supposer entre le milieu que nous venons de décrire et un organisme impliqué dans la chaîne interactive de la connaissance sonore? Nous définirons globalement cette relation de la manière suivante: *la fonction de l'organisme est de produire la connaissance des manifestations énergétiques "sonores" qui se produisent dans le milieu.* Le "problème" de l'organisme est un problème de *caractérisation* des *différentes* manifestations énergétiques qui se produisent dans l'environnement immédiat de l'organisme, alors que le "problème" du chercheur qui se penche sur le phénomène de la connaissance sonore est celui de la description de la manière *spécifique* dont l'organisme connaît (caractérise) ces manifestations énergétiques.

Nous entendons par 'organisme' toute organisation matérielle ou abstraite qui permet la connaissance des manifestations énergétiques "sonores". Ces organismes existent sous une grande diversité de formes dans la nature¹², diversité qui est le fait, croyons-nous, d'une seule *logique* de connaissance sonore et c'est dans la perspective de cette logique que nous regrouperons les différents éléments de ces organismes en trois grandes catégories de *structures de caractérisation* des manifestations énergétiques sonores:

- 1) la catégorie des *structures d'interface*, qui représente l'ensemble de ces structures biologiques et artificielles qui interagissent

¹² La nature comprend, à notre avis, les organismes artificiels qui sont le fruit de cet élément de la nature qu'est l'être humain.

directement avec le milieu, c'est à dire essentiellement les différents types de microphones et de tympan.

2) la catégorie des *structures d'analyse spectrale*, qui représente l'ensemble des structures biologiques et artificielles qui reçoivent la production des structures d'interface, tels la cochlée, les analyseurs spectraux, les filtres électro-acoustiques et les modèles mathématiques d'analyse spectrale, et finalement,

3) la catégorie des *structures supérieures*, qui représente l'ensemble des structures biologiques (essentiellement les structures nerveuses) et artificielles qui reçoivent la production des structures d'analyse spectrale.

Essayons maintenant de cerner plus précisément comment l'organisme connaît de manière sonore en considérant successivement ces différentes structures de caractérisation des manifestations énergétiques "sonores".

2.2.1 Les structures d'interface

Un organisme plongé dans le milieu, appréhende les manifestations énergétiques qui s'y produisent par le biais de la ou des structures d'interface qu'il possède. Quelles sont les caractéristiques de l'interaction entre les manifestations énergétiques du milieu et une structure d'interface? Les différents éléments de la famille des structures d'interface sont habituellement constitués d'une membrane élastique qui réagit aux variations de pression du médium sonore. La position d'équilibre de ces membranes correspond à la valeur normale ou non perturbée de la pression du médium sonore. Lorsque cette valeur de pression augmente, la membrane élastique se déforme dans un sens (vers "l'intérieur" de l'organisme) et lorsque cette valeur de pression diminue, la membrane retourne à sa position d'équilibre. Si la valeur de pression diminue plus bas que la valeur normale de pression du médium sonore, la membrane se déforme dans l'autre sens (vers "l'extérieur" de l'organisme)¹³. La position de la membrane reflète donc les variations de pression du médium sonore; elle reflète donc également la valeur de la puissance instantanée, c'est à dire la quantité d'énergie qui se manifeste à l'endroit spatial qu'occupe la structure d'interface à l'instant précis où nous considérons la position de la structure d'interface. Même si la membrane élastique occupe un certain espace et que son déplacement s'effectue sur un espace relativement étendu, on peut considérer

¹³ Les structures d'interface sont également dotées d'un mécanisme d'égalisation de la pression ambiante qui permet à l'organisme de s'ajuster aux variations de la pression ambiante (trompe d'Eustache)

"l'information" que représente la position de la structure d'interface comme une mesure des manifestations énergétiques en *un point de l'espace* du milieu. Or, comme nous l'avons vu à la section précédente, cette mesure énergétique est également la résultante de toutes les ondes longitudinales qui originent des objets physiques dans l'environnement immédiat de l'organisme. A ce *niveau de connaissance* que représente le processus de caractérisation de la structure d'interface, toutes les manifestations énergétiques se retrouvent "confondues" dans la valeur globale de la quantité énergétique qui est affichée par la position de cette structure dans l'espace. *La quantité énergétique que reflète la position d'une seule structure d'interface ne permet pas de différencier les manifestations énergétiques.* Dans ce contexte, comment l'organisme peut-il en arriver à connaître les *différentes* manifestations énergétiques qui l'atteignent?

Les organismes biologiques sont habituellement dotés de deux structures d'interface à chaque extrémité de l'axe transversal qui traverse la tête de l'animal. Se pourrait-il que la "*binauralité*" soit à l'origine de la caractérisation des différentes manifestations énergétiques? Avec une seule structure d'interface, il est effectivement beaucoup plus difficile de *localiser* un son dans l'espace puisqu'une seule membrane élastique n'opère pas ou peu

de différenciation spatiale¹⁴. Elle agit au contraire comme un intégrateur spatial de tous les heurts des particules aériennes qui l'atteignent sans réagir "différentiellement" à la direction de ces heurts. Cette différenciation spatiale est, semble-t-il, produite par la comparaison des manifestations énergétiques qui se produisent sur deux structures d'interface. Comme nous l'avons vu, il peut exister un certain décalage temporel entre les valeurs de pression qui se produisent en un point de l'espace à un instant donné et les valeurs de pression qui se produiront en un autre point assez rapproché de l'espace quelques instants plus tard. Si la position du corps physique qui est à l'origine de ces ondes longitudinales ne se situe pas dans un axe qui est parfaitement perpendiculaire à l'axe rejoignant les deux structures d'interface, les valeurs de pression qui atteindront les deux structures d'interface seront légèrement décalées dans le temps. En comparant ce décalage temporel, il devient alors possible d'opérer une certaine différenciation spatiale des manifestations énergétiques puisque des manifestations énergétiques identiques et synchroniques signifient que l'origine de ces manifestations se situe dans une direction qui leur permet d'atteindre les deux tympans au même instant alors qu'un décalage temporel entre les manifestations énergétiques reflètent une

¹⁴ Il faut noter toutefois qu'une seule structure d'interface peut avoir une certaine directionnalité spatiale. Les microphones unidirectionnels sont justement organisés pour "renforcer" les ondes longitudinales qui proviennent d'une certaine direction et atténuer les ondes qui proviennent des autres directions. Il semblerait également que le pavillon et le méat auditif de l'organisme humain servent aux mêmes fonctions. Mais pour qu'une différenciation spatiale s'effectue, il faut l'inscrire dans une représentation de l'espace. Cela n'est certainement pas le fait de la structure d'interface mais plutôt le résultat du travail de connaissance des structures cognitives qui agissent sur l'"output" de la structure d'interface et qui mettent en relation cet "output" avec une représentation de l'espace environnant.

position différente du corps-source d'origine. Chez les organismes biologiques, de plus, une orientation dissymétrique des deux structures d'interface peut résulter en deux successions de valeurs de pression très différentes. L'obstacle que constitue la tête de l'organisme, l'effet des conduits auditifs orientés différemment et la réflexion différente sur les obstacles environnants sont les causes de différences importantes dans les valeurs de pression qui se manifestent aux deux structures d'interface. Comme le soulignent Colburn et Durlach (1978):

"The component of binaural perception that is most uniquely binaural, as opposed to monaural, is the perception of differences between the stimuli presented to the two ears. In a natural environment, these differences result from differences in the propagation paths to the two ears and include, for each source in the environment, the differential effects on the two signals of the listener's body, head and pinna, as well as of other objects in the environment. It is precisely the perception of these differences that underlies the improved performance in localization and detection that occurs when two ears are used in place of a single ear." [p.368]).

La "binauralité" permettant de localiser les sons dans l'espace, elle est fort probablement à l'origine de la caractérisation des attributs spatiaux des manifestations énergétiques du milieu. Mais est-ce la "binauralité" qui nous permet d'identifier et de caractériser les différentes manifestations énergétiques qui contribuent à la valeur énergétique globale qui nous est fournie par la position de la structure d'interface? Nous ne le croyons pas. A l'exception d'un certain déficit dans la localisation des sons dans l'espace et d'une certaine difficulté à identifier tous les sons dans des situations énergétiques complexes ("cocktail party effect"), l'expérience sonore d'un organisme monaural reste riche et se constitue d'entités expérientielles

sonores qui se comparent aux entités expérientielles sonores générées par des organismes dotés de la binauralité. L'expérience sonore monaurale reste *essentiellement* sonore comme l'expérience visuelle monoculaire reste *essentiellement* visuelle. Si ce n'est pas dans la "binauralité" que réside l'essentiel du processus de connaissance sonore de l'organisme, nous pouvons donc poser le problème que représente ce processus à partir de la caractérisation énergétique qu'effectue l'unique structure d'interface d'un organisme monaural. Or, l'énergie qui se manifeste à une seule structure d'interface est totalement dénuée de "spatialité" puisque la valeur de l'énergie en ce point de l'espace qu'occupe la structure d'interface est la résultante d'ondes longitudinales qui proviennent de toutes les directions possibles. Si ce n'est pas dans la "spatialité" que réside l'"individualité" des manifestations énergétiques, ce ne peut être que dans la "temporalité" de ces manifestations. Nous en concluons que le processus de connaissance sonore qu'effectue l'organisme est essentiellement un processus de caractérisation des manifestations *temporelles* de l'énergie sonore et que c'est ce processus qui permet d'extraire les caractéristiques des différentes manifestations énergétiques qui se produisent dans l'environnement immédiat de l'organisme.

Quelle est la nature de la caractérisation des manifestations temporelles de l'énergie sonore qu'effectue la structure d'interface? *A chaque instant*, la structure d'interface produit une information qui représente l'effet combiné des "émissions" énergétiques qui proviennent du milieu environnant l'organisme et, instant après instant, ce "produit cognitif" représente toujours la somme de ces effets. Dans ce contexte, il est difficile de voir comment

l'organisme peut en arriver à connaître, c'est à dire à identifier et à caractériser, les *différentes* manifestations énergétiques (les différents sons) qui se produisent à *chaque instant et durant plusieurs instants* puisqu'il n'existe, à ce niveau, *qu'une caractérisation de l'effet total de toutes les manifestations énergétiques sonores*, c'est à dire la valeur totale de l'énergie à chaque instant. On peut voir dans cet argument une première justification intuitive au fait que le processus de connaissance sonore semble s'effectuer en plusieurs étapes chez les organismes "évolués" comme l'organisme humain. Pour en arriver à une connaissance du milieu qui puisse lui permettre de s'adapter, l'organisme doit "aller plus loin" que ce premier niveau de caractérisation. Il doit poursuivre le processus de connaissance sonore par le "biais" de structures additionnelles de caractérisation qui lui permettent de générer les "fiches signalétiques" des différentes manifestations énergétiques du milieu. Comment cela se produit-il? Quelles sont les caractéristiques de ces structures additionnelles de caractérisation et quel est leur rôle dans le processus de connaissance sonore? C'est ce que nous allons maintenant essayer de comprendre en considérant successivement les structures du deuxième niveau de caractérisation, i.e. les structures d'analyse spectrale; puis les structures des niveaux ultérieurs de caractérisation, i.e. les structures supérieures (essentiellement les structures nerveuses).

2.2.2 Les structures d'analyse spectrale

La séquence temporelle des positions de la structure d'interface, le *signal* sonore, "alimente" les structures d'analyse spectrale qui constituent le deuxième niveau de caractérisation. *Tous les processus de caractérisation des structures de ce second niveau de caractérisation fonctionnent à partir des différentes valeurs du signal sonore.* Ces structures procèdent à leur tour à une certaine caractérisation de ces valeurs, caractérisation qui s'exprime également sous la forme d'un "produit cognitif". Celui-ci (le produit cognitif des structures d'analyse spectrale) est souvent considéré comme une description complémentaire et essentiellement distincte du produit cognitif de la structure d'interface. Alors que ce dernier est couramment désigné comme une *représentation temporelle* ("time analysis") des manifestations énergétiques sonores, le produit des structures d'analyse spectrale correspond à une *représentation spectrale* ou "*fréquentielle*" ("spectral analysis" ou "frequency analysis") de ces mêmes manifestations. Les valeurs résultantes du processus de caractérisation des structures d'analyse spectrale sont en effet couramment interprétées comme le reflet de la présence plus ou moins importante de différentes fréquences ou groupes de fréquences dans le signal sonore¹⁵. Dans cette section, nous allons étudier les différentes incorporations de structures d'analyse spectrale et essayer de comprendre ce qui distingue la représentation temporelle de la représentation spectrale.

¹⁵ Le "spectre" des fréquences est une représentation bidimensionnelle de la présence plus ou moins forte d'une fréquence ou d'un groupe de fréquences. Un groupe de fréquences est également désigné sous le terme de région spectrale.

La grande famille des structures d'analyse spectrale peut être divisée en trois catégories: les structures d'analyse spectrale biologiques, physiques et mathématiques¹⁶. La catégorie des structures d'analyse spectrale biologiques regroupe l'ensemble des différents modèles cochléaires que l'on peut retrouver chez les organismes vivants¹⁷. Les caractéristiques du comportement de l'organe cochléaire ont fait l'objet de nombreuses études depuis la mise en évidence de son rôle de structure de caractérisation des manifestations énergétiques sonores. La cochlée étant un organe profondément enfoui dans la masse osseuse temporale, ce n'est que depuis la mise au point de techniques expérimentales spéciales (Bekesy, 1960; Rhode, 1970) que l'on a pu observer "de visu" les caractéristiques essentielles de son comportement en réaction aux changements de position de la structure d'interface. La cochlée effectue un type d'analyse spectrale qui se caractérise par l'apparition "d'ondes voyageuses" ("traveling waves") qui se propagent à travers l'organe cochléaire. Dans le cas de variations sinusoïdales de la position de la structure d'interface, ces ondes atteignent un maximum d'amplitude à des endroits différents de l'organe cochléaire, l'emplacement et l'amplitude de ces maximums d'amplitude étant fonction de la fréquence et de l'amplitude de la stimulation sinusoïdale. Ce comportement est, entre autres, le fait des propriétés physiques de la membrane basilaire qui s'étend à

¹⁶ Le terme "structures mathématiques" a, par ailleurs, un sens "mathématique" qui est assez différent du sens "cognitif" que nous lui donnons ici.

¹⁷ La cochlée "apparaît" chez les poissons à structure osseuse et se retrouve entre autres chez les reptiles, les oiseaux et les mammifères. L'évolution phylogénétique de la cochlée se caractérise surtout par l'allongement de l'organe et une innervation plus importante, ce qui laisse supposer une sensibilité spectrale accrue (Masterton & Diamond, 1978).

l'intérieur de la cochlée. Cette membrane est plus raide et plus étroite à l'entrée de la cochlée et devient progressivement de plus en plus lâche et plus large vers l'extrémité de la cochlée. Des fibres nerveuses spécialisées sont attachées tout au long de la cochlée et les mouvements de la membrane basilaire initient des potentiels d'action dans ces fibres nerveuses. Plus le mouvement vertical de cette membrane est ~~ampli~~, plus il y aura de potentiels d'action initiés sur un certain intervalle temporel. Le comportement cochléaire a donc une double dimension spatiale: la longueur "horizontale" de la cochlée et le déplacement "vertical" de chaque point qui se situe sur la dimension horizontale. A chaque instant, la courbe que forme la cochlée représente donc le travail de "caractérisation de cette structure d'analyse spectrale. Malgré son "ouverture" récente, il reste toutefois difficile d'observer la réaction de la cochlée à des stimuli plus complexes qu'une onde sinusoïdale. C'est pourquoi de nombreux modèles, tant mathématiques que physiques, ont été et sont encore élaborés pour essayer de simuler la réaction de cet organe à des signaux plus complexes qu'une onde sinusoïdale. (Bekesy, 1960; Duifhuis, 1972; Tonndorf, 1962; Steele, 1973; Weiss, 1966).

La deuxième catégorie de structures d'analyse spectrale se constitue de l'ensemble des différentes machines artificielles d'analyse spectrale: analyseurs spectraux à temps réel, récepteur hétérodyne, analyseur spectral à bande variable etc. La grande majorité de ces structures artificielles repose sur l'utilisation de filtres électroniques continus ou digitaux. Un filtre est une structure matérielle réagissant aux différentes fréquences qui sont présentes dans le signal sonore. On divise habituellement les filtres en filtres passe-haut, filtres passe-bas et filtres à bande passante selon leur réaction aux

différentes fréquences. Un filtre passe-haut ne "laissera passer" que les composantes énergétiques du signal sonore qui appartiennent aux fréquences qui sont plus élevées (plus rapides) qu'une certaine 'fréquence de coupure' et éliminera les composantes énergétiques qui appartiennent aux fréquences plus "basses" (plus lentes) de ce signal. De la même manière, un filtre passe-bas ne laissera passer que les composantes énergétiques des fréquences qui se situent en deçà d'une certaine fréquence de coupure et éliminera les composantes "hautes", alors qu'un filtre à bande passante réagira aux composantes énergétiques des fréquences du signal qui se situent entre deux fréquences de coupure, l'une qui délimite les composantes hautes qui seront éliminées par le filtre (la fréquence de coupure passe-bas), l'autre qui délimite les composantes basses qui seront éliminées (la fréquence de coupure passe-haut), ne laissant "subsister" que la contribution des fréquences qui sont spécifiques à la région spectrale qui se situe entre les deux fréquences de coupure.

Les structures d'analyse spectrale biologiques ainsi que les structures d'analyse spectrale physiques peuvent être conçues comme des "incorporations" spécifiques de la troisième catégorie de structures d'analyse spectrale: les structures mathématiques¹⁸. Ces structures mathématiques ont été élaborées dans les trois derniers siècles et sont, entre autres, le résultat du travail mathématique de Fourier (1822). Leur disponibilité ainsi que leur propension à la description des phénomènes temporels en ont fait naturellement, et depuis longtemps, les outils privilégiés d'une formalisation

¹⁸ De manière inverse, on peut concevoir les structures mathématiques comme des "idéalisations" des structures biologiques et physiques.

des processus de connaissance sonore, que ceux-ci soient artificiels ou biologiques. Cela ne signifie pas pour autant que l'interprétation "cognitive" de ces modèles mathématiques soit aisée. Comme le souligne Nordmark (1978):

"In addition, both the success and the limitations of the analytical approach is, to a large extent, bound up with its utilization of Fourier mathematics. Fourier analysis is of unique value if the time-varying quantity is a periodic function. If it is non-periodic, it can sometimes be represented by the Fourier series extension called a *Fourier integral*. But this kind of mathematics is ill-suited to describing the constantly changing sounds typical of our everyday experience. Also,..., the identification of the mathematical variables with physical quantities like time and frequency is not as straightforward as is customarily assumed... The complex task of describing a sound stimulus is thus intimately bound up with the analyzing procedure." [p.246-248]

Depuis les premiers travaux de Fourier, de nombreuses structures mathématiques ont été développées pour effectuer l'analyse spectrale du signal. Ces structures peuvent être regroupées en deux catégories selon le rationnel mathématique qui leur est sous-jacent (Beauchamp & Yuen, 1979; Schafer & Rabiner, 1974): d'une part, l'autocorrélation et la corrélation croisée qui sont des structures mathématiques basées sur la théorie probabiliste et, d'autre part, les séries de Fourier, la transformation de Fourier et l'analyse spectrale de courte durée ("short-time spectrum analysis") qui sont des structures mathématiques basées sur le calcul différentiel et intégral.

Les structures mathématiques qui s'inspirent de la théorie probabiliste sont basées essentiellement sur des mesures de corrélation. Comme son nom l'indique, la corrélation est une mesure de la relation qui existe entre deux

variables. Dans le cas de l'analyse spectrale corrélationnelle, on utilise les variables que représentent les valeurs de la position de la structure d'interface à différents instants. L'auto-corrélation et la corrélation croisée sont deux techniques dont le résultat se présente également sous la forme d'un spectre, mais cette fois-ci d'un spectre *corrélational* dont l'origine sur l'axe des X correspond à la corrélation de chaque valeur du signal avec elle-même, corrélation qui est donc toujours égale à 1. Les valeurs de chaque côté de l'origine indiquent la mesure de corrélation du signal avec des versions du signal qui sont temporellement de plus en plus *différées*. Dans ce contexte, on peut voir intuitivement qu'un signal aléatoire, qui ne comprend pas de concentration spécifique d'énergie dans l'une ou l'autre région de son spectre de fréquences, mais plutôt une répartition aléatoire de l'énergie sur tout le spectre, produira un "pattern" corrélationnel différent du pattern corrélationnel d'une sinusoïdale où la mesure de corrélation alternera d'une manière continue entre 0 et 1 (Beauchamp & Yuen, 1979).

Les séries de Fourier permettent de générer une série harmonique de coefficients sinusoïdaux pour des signaux *périodiques*, i.e. qui se répètent en théorie de $-\infty$ à $+\infty$, mais dont la période est comprise quelque part entre ces deux limites. La transformation de Fourier est une extension des séries pour des signaux *non-périodiques*. Dans la transformation de Fourier, la "période" du signal non-périodique est considérée comme s'étendant de $-\infty$ à $+\infty$. Les séries de Fourier produisent un spectre discret et infini des valeurs de phase et d'amplitude des coefficients sinusoïdaux harmoniques dont l'intervalle de discrétude est déterminé par la période T du signal ($1/T$), alors que la transformation de Fourier produit un spectre infini mais continu de

ces mêmes coefficients puisque l'extension de la période à l'infinité génère des "harmoniques" dont l'intervalle est infiniment petit ($1/\infty$). Dans la perspective mathématique des séries et de la transformation de Fourier, les valeurs de position de la structure d'interface qui sont représentatives d'une sinusoïdale pure sont parfaitement définies d'un point de vue spectral mais totalement indéfinies temporellement puisqu'elles recouvrent l'infinité de la dimension temporelle, alors qu'une *seule* valeur de position de la structure d'interface est parfaitement définie dans le temps mais complètement indifférenciée en termes de contenu spectral, son énergie se répartissant uniformément sur tout le spectre des fréquences. Pour obtenir une description spectrale 'à la Fourier', un "observateur" doit "embrasser" d'un seul coup l'infini de la dimension temporelle, ce qui revient à dire que, pour lui, les manifestations énergétiques sonores n'existent à aucun instant particulier, alors que si ce même observateur veut obtenir une représentation temporelle du même évènement, il doit se "concentrer" uniquement et successivement sur un seul instant, perdant du même coup toute "vision d'ensemble" sur le phénomène énergétique. Il est assez évident qu'une interprétation strictement 'à la Fourier' du travail des structures d'analyse spectrale ne correspond pas à notre appréhension intuitive du travail de connaissance sonore. Comme le souligne Leshowitz (1978):

"From the fundamental definition of the Fourier transform..., it is seen that in order to determine the frequency content of a time waveform, it is necessary to perform an integration over an infinite range of time. Quite obviously, a perceiving organism does not perform an integration over an infinite amount of time as required by the classical definition of *long-time* frequency spectrum. At best we carry out an integration within the limits $-\infty$ to the present time t ."

L'analyse spectrale de courte durée ("short-time energy spectrum") est une technique mathématique qui effectue justement ce type d'analyse spectrale. Cette technique a été développée dans le cadre de l'analyse des signaux sonores linguistiques et est à la base de la mise au point du spectrographe (Flanagan, 1972). Dans ce type d'analyse spectrale, on est intéressé à extraire les caractéristiques spectrales qui permettront de différencier entre les différents sons du langage (fricatifs, explosifs, vocaux, non-vocaux etc.). Pour mettre en évidence ces caractéristiques spectrales, on procède au calcul de la transformation de Fourier sur différentes "fenêtres" temporelles qui recouvrent des portions temporelles variables des valeurs *passées* du signal sonore jusqu'à l'instant présent¹⁹. On obtient alors une représentation spectrale "courante" (running spectrum) en affichant les valeurs successives de ce spectre sous la forme de la contribution des différentes régions spectrales.

Toutes les structures mathématiques de même que leur incorporation sous la forme de structures physiques ou biologiques effectuent donc, d'une certaine manière, le même type de travail: elles visent à caractériser les contributions de différentes fréquences dans un signal sonore. Mais il semble exister un "flou conceptuel" quant aux caractéristiques *temporelles* de ces contributions. Comme le souligne Gabor (1946):

"Hitherto communication theory was based on two alternative methods of signal analysis. One is the description of the signal as a function of time; the other is Fourier analysis. Both are idealizations,

¹⁹ C'est de ce calcul sur des fenêtres temporelles *limitées* que provient le nom d'analyse spectrale de *courte durée* puisque le calcul "normal" d'une transformation de Fourier s'effectue sur un signal qui s'étend théoriquement de + à - l'infini temporel.

as the first method operates with sharply defined instants of time, the second with infinite wave-trains of rigorously defined frequencies. But our everyday experiences-especially our auditory sensations-insist on a description in *both* time and frequency... Though mathematically this theorem is beyond reproach, even experts could not at times conceal an uneasy feeling when it came to the physical interpretation of results obtained by the Fourier method... The reason is that the Fourier-integral method considers phenomena in an infinite interval, *sub speciaē aeternitatis*, and this is very far from our everyday point of view. Fourier's theorem makes of description in time and description by the spectrum, two mutually exclusive methods. If the term "frequency" is used in the strict mathematical sense which applies only to infinite wave-trains, a "changing frequency" becomes a contradiction in terms, as it is a statement involving *both* time and frequency. [pp 429-431]

Il semble donc exister une ambiguïté fondamentale entre la représentation spectrale 'à la Fourier' et la représentation temporelle que constitue le signal sonore lorsqu'on les considère dans la perspective d'un organisme qui connaît *à travers le temps*. Comment pouvons-nous réduire cette ambiguïté et intégrer ces deux types de représentation dans une perspective cohérente du processus de connaissance sonore? A notre avis, il faut pour cela définir les "relations" qui existent entre la dimension temporelle, d'une part, et deux aspects de la connaissance sonore, d'autre part. Ces deux aspects sont:

- les *processus* de caractérisation de chaque niveau, et,
- les *produits* de ces processus de caractérisation.

Nous allons commencer par essayer d'établir les relations qui unissent *temps* et *processus* de caractérisation pour ensuite considérer les relations qui unissent *temps* et *produits* de ces processus de caractérisation.

Une certaine confusion provient tout d'abord de la nature temporelle des différents processus de caractérisation. Le processus de caractérisation de la structure d'interface s'effectue à chaque instant, *à travers le temps*, et c'est le résultat de ce processus, *temporellement polyvalent*, qui se reflète dans la 'représentation temporelle' qu'est le signal sonore. Les processus de caractérisation que sont les structures mathématiques des séries et la transformation de Fourier ainsi que les structures mathématiques de l'analyse corrélationnelle correspondent, quant à eux, à des processus de caractérisation qui sont *temporellement monovalents*: ils ne s'effectuent pas à chaque instant. Ils ne s'effectuent qu'une seule fois, à l'instant de leur calcul. Nous avons vu, par contre, que l'analyse spectrale de courte durée ainsi que les structures "concrètes" que sont les différents types d'analyseurs spectraux biologiques (cochlée) et artificiels (instrumentation physique), représentent, eux, des processus de caractérisation qui sont *temporellement polyvalents* car ils s'effectuent à travers le temps. Ces structures permettent d'obtenir une représentation spectrale "*courante*" en caractérisant successivement les nouvelles valeurs du signal sonore à chaque nouvel instant dans le temps. Au tout début de cette section, nous avons mentionné que l'on distinguait fréquemment le processus de caractérisation de la structure d'interface du processus de caractérisation des structures d'analyse spectrale en désignant le produit du premier processus sous l'appellation de 'représentation temporelle' (time analysis), alors que le produit du processus de caractérisation des structures d'analyse spectrale n'était pas affublé de cette connotation temporelle et était désigné couramment sous le terme de représentation spectrale. Nous pouvons voir que cette conception "non-temporelle" de la représentation spectrale prend son origine dans les

processus de caractérisation des structures mathématiques des séries et de la transformation de Fourier ainsi que dans ceux des structures corrélationnelles, puisque ces processus ne s'effectuent pas *à travers le temps*, i.e. ils ne s'effectuent qu'une *seule* fois. Par contre, dans la perspective des processus de caractérisation que sont l'analyse spectrale de courte durée et les analyseurs spectraux biologiques et artificiels, l'absence d'une connotation temporelle à la représentation spectrale devient beaucoup moins pertinente puisque ces processus de caractérisation s'effectuent justement *à travers le temps*, i.e. à tous les instants, et ce, exactement comme pour le processus de caractérisation de la structure d'interface. Les deux types de représentation, la représentation temporelle qui est produite par la structure d'interface et la représentation "temporelle" spectrale qui est produite par ces dernières structures spectrales "courantes", sont, toutes deux, de nature *temporelle polyvalente*: elles s'effectuent côte à côte, à chaque instant à travers le temps. Dans ce cas, le passage de la représentation temporelle à la représentation spectrale illustre en fait une transition entre *deux* représentations temporelles, celle de la structure d'interface et celle des structures d'analyse spectrale "courantes". C'est de cette "dualité" temporelle des processus de caractérisation du niveau de l'analyse spectrale que provient une partie de "l'ambiguïté temporelle" entre représentation temporelle et représentation spectrale. Lorsque l'on considère les structures d'analyse spectrale dans cette perspective dualiste, nous pouvons mieux saisir les relations qui existent entre les différents processus de caractérisation et la dimension temporelle, et cela nous permet de réduire cette ambiguïté.

Une autre source d'ambiguïté provient de la nature temporelle des différents *produits* cognitifs qui sont le fruit des niveaux successifs de caractérisation que sont les structures d'interface, les structures d'analyse spectrale et les structures supérieures. Quelle est la nature temporelle de ces produits cognitifs et, plus précisément, quelle est la nature temporelle *instantanée* de ces différents produits? Nous savons que la valeur instantanée du signal sonore représente la quantité totale d'énergie qui se manifeste, à cet instant et à cet endroit de l'espace qu'occupe la structure d'interface, mais que représentent, à chaque instant, les différentes valeurs du produit cognitif des structures d'analyse spectrale? Quelle est, en somme, la nature temporelle du produit cognitif qui résulte du deuxième niveau de caractérisation de l'organisme?

Nous croyons que les produits cognitifs de toutes les structures d'analyse spectrale, y compris ceux des structures "standard" de Fourier et ceux des structures corrélationnelles, expriment le résultat d'un processus de caractérisation des *relations temporelles* qui existent entre les différentes valeurs du signal sonore, c'est à dire du produit de la structure d'interface. Les produits de l'analyse spectrale de courte durée, le spectre discret des séries, le spectre continu de la transformation, le spectre corrélationnel et le produit des différentes structures matérielles (cochlée, analyseurs spectraux etc.) constituent de fait des représentations de ces relations temporelles qui existent entre les valeurs de position de la structure d'interface. Les structures 'temporellement monovalentes' (séries et transformations de Fourier, spectres corrélationnels) fournissent une seule représentation de ces relations alors que les structures 'temporellement polyvalentes' (les

structures matérielles et l'analyse spectrale de courte durée) fournissent une représentation "courante" de ces mêmes relations (une représentation à chaque "avancée" dans le temps, à chaque instant).

Nous avons vu, à la fin de la section 2.2.1, que la connaissance des différentes manifestations énergétiques du milieu ne pouvait être le fait que de la caractérisation des manifestations *temporelles* de l'énergie sonore. Au premier niveau de caractérisation de ces manifestations temporelles, le processus de caractérisation de la structure d'interface produit, à chaque instant, une mesure de la quantité totale d'énergie à cet instant. Au second niveau, le processus de caractérisation de ces mêmes manifestations temporelles produit, à chaque instant (dans le cas des structures d'analyse spectrale qui sont "courantes"), une mesure de la relation temporelle qui existe entre les produits successifs du premier niveau, c'est à dire une mesure de la relation entre les valeurs successives de quantité totale d'énergie, à chaque instant, jusqu'à l'instant présent de calcul.

Le processus de caractérisation des structures d'analyse spectrale soulève de nombreuses questions. Quels sont les caractéristiques précises du "produit" de ce deuxième niveau de caractérisation? Quelle est la nature exacte de la 'mise en relation temporelle' des valeurs successives de la structure d'interface et comment se réalise précisément cette 'mise en relation temporelle'? Finalement, comment "s'insère" ce processus de caractérisation dans l'ensemble des processus de caractérisation de l'organisme sonore? Nous essaierons de répondre à ces différentes questions dans le troisième chapitre de ce travail. Pour l'instant, nous allons suivre jusqu'au bout la "filière cognitive" de la connaissance sonore en considérant

successivement le travail de connaissance des structures supérieures et le produit final de ce travail: l'expérience sonore.

2.2.3 Les structures supérieures

Nous avons vu qu'il existait des structures de caractérisation *biologiques* et *artificielles* pour les deux premiers niveaux du processus de connaissance sonore. Les structures *supérieures* de caractérisation des organismes sont principalement de nature *biologique*. Les structures supérieures artificielles (modèles mathématiques et computationnels, appareillage électronique) qui sont le fruit des travaux sur la reconnaissance automatique de la parole (Pisoni, 1985) et des travaux psychophysiques sur la caractérisation des attributs de hauteur, d'intensité et de timbre des entités expérientielles sonores (Goldstein, 1973; Plomp, 1970; Wightman, 1973; Zwicker & Scharf, 1965) ne se comparent pas, à l'heure actuelle, à la sophistication atteinte par les structures biologiques de même niveau. Nous nous restreindrons donc, dans cette section, à l'étude des structures supérieures biologiques. Ces structures se constituent des différents parcours et "centres de traitement" du système nerveux sonore. Celui-ci peut être divisé en cinq niveaux structuraux distincts: le noyau cochléaire, le complexe de l'olive supérieure et le noyau lemniscal, le colliculus inférieur, le corps genouillé médian et finalement le cortex auditif. Tous ces niveaux structuraux possèdent des caractéristiques communes:

" It is not uncommon to describe the central auditory system as a pathway from ear to cerebral cortex. But it can be seen... that each central auditory structure, regardless of how "high" or how "low" it is stationed in the "pathway", is much more than a relay or stepping stone to cortex. Instead each structure in the auditory system is known to act in at least four diverse ways. *First*, as the usual view would suggest, each structure serves as a relay in the pathway from

the ear to higher structures in the system... *Second*, each structure in the auditory system transforms its neural input into an output which always differs from its input, and along at least some descriptive dimensions, the "neural code" transmitted by a center is entirely incommensurable with the "code" arriving at the center... *Third*, each structure in the auditory system serves to distribute its products to a variety of nonauditory structures... *Fourth*, each auditory structure serves to modulate its own input by efferent fibers that descend to lower auditory structures in feedback loops." (Masterton & Diamond, 1978) [pp. 428-429].

Si l'anatomie de ces structures est assez bien connue, leur physiologie reste par contre très obscure. Comme le souligne Evans(1971), "... the enormous complexity of interconnections between and within the nuclei of the auditory pathway almost defies functional investigation." [p.328]. Malgré ces difficultés d'investigation et d'interprétation, quelques caractéristiques fonctionnelles générales peuvent être identifiées. Premièrement, l'organisation tonotopique²⁰ des fibres du nerf auditif semble être conservée et projetée jusqu'au cortex. L' "entrée" du système nerveux, l'ensemble de fibres nerveuses spécialisées qui sont "branchées" sur toute la longueur de la cochlée, transmet l'amplitude des déplacements verticaux de la cochlée jusqu'au cortex. Comme le remarque Gacek (1972):

"... it is apparent that there is a definite orderliness to the projection of the end organ onto the cortex. To see this fact clearly, one need only look at the first and last neurons of the ascending chain. The bipolar cochlear neurons faithfully project the frequency-located points of the cochlear partition onto the ventral cochlear nucleus. The geniculocortical projection from the pars principalis to the primary cortex has been shown by retrograde degeneration methods to project the cochlea's frequency-ordered organization from the medial geniculate to the cortex." [p.251]

²⁰ En fonction de la fréquence.

Deuxièmement, le rôle des structures pré-corticales semble être essentiellement relié à la caractérisation spatiale des entités sonores et à la transmission de ces caractérisations aux systèmes moteurs et visuels à des fins d'orientation et de réponses réflexes. Cela se reflète par l'organisation anatomique particulière de ces structures et par un grand nombre de branchements synaptiques vers d'autres centres de traitement nerveux qui sont identifiés à de telles fonctions (Masterton & Diamond, 1978). Finalement, la projection tonotopique, le rôle subalterne des structures pré-corticales et le déséquilibre entre le nombre de neurones des structures pré-corticales et celui des structures corticales indiquent que ces structures corticales sont responsables des caractéristiques sophistiquées de l'expérience sonore. Chez le singe, le nombre de corps cellulaires se multiplie par *quatre* (4) entre le premier branchement synaptique (noyau cochléaire $\approx 88,000$) et le dernier relais au cortex (les corps genouillés $\approx 364,000$) ce qui suppose un certain traitement de l'information dans les voies afférentes, alors que ce dernier nombre se multiplie par *trente* (30) pour le cortex ($\approx 10,000,000$), mettant nettement en évidence que le rôle du cortex est primordial et que le travail descriptif de la chaîne ascendante qui va jusqu'au cortex est assez limité. Or le rôle et l'extension anatomique précise du cortex auditif sont fort mal connus. On lui assigne par défaut un rôle de caractérisation des attributs sonores sophistiqués puisque les recherches basées sur les techniques de micro-électrodes ont trouvé très peu de ressemblance entre les réponses des fibres nerveuses des structures pré-corticales (qui sont corrélés aux attributs sonores de hauteur et d'intensité) et les réponses des fibres des structures corticales. Contrairement aux fibres pré-corticales, les fibres du cortex sonore réagissent en effet à une très grande variété de stimulations sonores.

Près de 75% de ces fibres sont insensibles à des stimulations sinusoïdales simples et ne démontrent une certaine agitation que lorsque le signal est complexe. Comme le dit Evans (1971): "Many cortical neurons, then, at least in unanaesthetized preparations, appear to be less "interested" in the classical studied parameters of acoustic stimuli, namely, the frequency and intensity of continuous tones, than in certain less easily definable features of complex sounds." [p.137].

Le cortex auditif regroupe fort probablement de nombreuses structures de caractérisation. Quels sont les processus de caractérisation de ces structures et quels sont les "produits cognitifs" résultant de ces processus de caractérisation? Nous sommes loin d'avoir une réponse à ces deux questions. Les techniques actuelles d'investigation du fonctionnement du système nerveux (enregistrement par micro-électrodes, dégénérescence, "tracing", etc.) ne nous laissent pas envisager un accroissement rapide et *qualitatif* de notre compréhension de ces structures. Ceci se reflète d'ailleurs dans les tentatives d'élaboration de structures artificielles supérieures. Comme nous le mentionnions au tout début de cette section, la performance de ces structures artificielles ne se compare pas à la performance des structures biologiques. De fait, nous pouvons voir dans ce décalage entre les structures supérieures produites *par* l'organisme humain et les structures supérieures *des* organismes biologiques une mesure assez exacte de notre compréhension (ou de notre incompréhension) des processus de connaissance sonore qui sont caractéristiques de ces niveaux supérieurs de caractérisation de l'organisme. Les études sur le fonctionnement du système nerveux auditif ne nous permettent pas de saisir la logique des processus de connaissance sonore. Il

est très difficile de suivre les parcours nerveux et de comprendre les principes de fonctionnement qui sont sous-jacents aux multiples interactions qui se produisent au sein du système nerveux. Les études de micro-électrodes ne peuvent être utiles que dans le contexte d'une recherche précise de certaines caractéristiques des processus de connaissance sonore. Le cortex auditif est complexe; y plonger une micro-électrode et essayer de corréler la réaction de cette micro-électrode à des attributs que l'on ignore ne présente, à notre avis, que peu ou pas de perspectives d'explicitation à court terme. Comment pouvons-nous cerner alors la nature de ces processus 'de haut niveau'? Est-ce à dire qu'il nous faut attendre de nouveaux développements technologiques qui nous permettront de saisir la logique de ces processus? Nous ne le croyons pas. *S'il existe une logique des processus de connaissance sonore de l'organisme, celle-ci devrait déjà se refléter dans les premiers niveaux du processus de caractérisation. Si nous pouvions cerner cette logique, il nous suffirait de "l'étendre" à ces niveaux supérieurs de caractérisation. Cela pourrait nous permettre, d'une part, d'orienter les recherches sur ces structures biologiques de haut niveau et d'amorcer, d'autre part, la "construction" de structures artificielles équivalentes à ces structures biologiques de haut niveau.*

Ceci complète la section 2.2 de ce deuxième chapitre. Dans cette deuxième section, nous avons essayé de comprendre comment l'organisme contribue au phénomène global de la connaissance sonore en s'intégrant dans la chaîne interactive qui relie milieu, organisme et expérience sonore. Notre conception de cette contribution se reflète essentiellement dans la vision d'un organisme qui produit la connaissance des manifestations énergétiques du

milieu. Les caractéristiques générales de ce 'processus de production de connaissance' sont les suivantes:

- pour connaître les manifestations énergétiques du milieu, l'organisme procède à une caractérisation des aspects temporels de ces manifestations,

- ce processus de caractérisation s'effectue en plusieurs étapes; il est le fait de plusieurs structures de caractérisation, structures qui appartiennent à des niveaux hiérarchiquement emboîtés,

- à chaque niveau, le processus de connaissance s'exprime par un "produit cognitif" qui est le fruit d'une caractérisation du produit cognitif du niveau précédent, fruit qui "alimente" à son tour le processus de caractérisation de la structure du niveau suivant,

- la structure d'interface représente le premier niveau du processus de caractérisation et son "produit cognitif" correspond, à chaque instant, à la quantité totale d'énergie qui résulte des différentes manifestations énergétiques qui ont atteint l'endroit qu'occupe la structure d'interface,

- la structure d'analyse spectrale représente le deuxième niveau du processus de caractérisation et son produit cognitif correspond, à chaque instant, à l'expression de certaines relations entre les valeurs successives du produit de la structure d'interface,

- les structures supérieures représentent les niveaux successifs du processus de caractérisation. Nous ne savons à peu près rien sur la nature des produits cognitifs qu'elles génèrent mais nous pouvons supposer que ces produits sont également le fruit de "mises en relation" des produits des niveaux "inférieurs".



2.3 L'expérience sonore

L'expérience sonore représente le troisième et dernier élément du tryptique problématique de la connaissance sonore. Dans cette section du chapitre 2, nous allons décrire différentes caractéristiques de l'expérience sonore afin d'essayer de comprendre comment ce dernier élément contribue à la logique globale de la connaissance sonore. Notre expérience des sons peut être conçue comme le produit cognitif final des multiples niveaux de caractérisation de notre organisme. C'est ce produit final qui se manifeste au foyer expérientiel sous la forme d'entités expérientielles et ce sont ces entités "terminales" que nous désignons habituellement sous le terme de 'sons'.

Si l'expérience sonore est le fait des organismes biologiques, on peut également supposer qu'elle est le fait des organismes artificiels. L'expérience sonore des organismes *biologiques* résulte toutefois d'un processus de connaissance qui comprend les structures de caractérisation des deux premiers niveaux *plus toutes les structures de caractérisation des niveaux supérieurs du système nerveux auditif*, alors que l'expérience sonore des organismes artificiels ne résulte, pour l'essentiel, que du processus de connaissance *des deux premiers niveaux de caractérisation*. Ceci est apparent dans la richesse expérientielle des organismes biologiques, richesse expérientielle qui se reflète aussi bien dans l'expérience des différents bruits que dans l'expérience linguistique ou musicale. Mais cette richesse expérientielle des organismes biologiques est beaucoup plus difficile à cerner que la "pauvreté" expérientielle relative des organismes artificiels qui

s'exprime, elle, à l'intérieur de la métrique et de la terminologie physique. Comment pouvons-nous décrire l'expérience sonore des organismes biologiques? Comment pouvons-nous décrire ce qui se passe "dans notre tête" lorsque nous entendons?

La psychophysique est née de ce désir d'appréhender rigoureusement l'expérience sonore en essayant justement de "co-rélier" le résultat de la description physique qui origine des structures artificielles du premier et du second niveau de caractérisation au résultat de la description biologique qui origine de *tous* les niveaux de caractérisation des organismes biologiques. Le paradigme psychophysique est essentiellement le suivant: construire des situations de stimulation qui sont parfaitement définies dans les termes des paramètres des représentations temporelle ou spectrale²¹, présenter ces situations de stimulation à des organismes biologiques et évaluer l'expérience sonore, par le biais de la réponse de ces organismes. L'évaluation de ces réponses se fait traditionnellement à partir de trois méthodes (Green & Swets, 1966). Dans la première méthode (oui-non), le sujet doit dire s'il perçoit ou non une entité expérientielle ou un certain attribut d'une entité expérientielle²². Dans la deuxième méthode (choix forcé), le sujet doit choisir si l'entité expérientielle qui résulte de la situation de stimulation

²¹ Ces paramètres sont l'amplitude (positions de la structure d'interface) et le temps pour la "représentation temporelle"; ce sont les amplitudes, les fréquences et les phases des différents composants spectraux pour la représentation spectrale.

²² Les sujets biologiques mais non-humains sont préalablement conditionnés à effectuer un certain comportement dans le cas de la présence d'un certain attribut. On leur présente la situation de stimulation en modifiant les valeurs de cet attribut et on observe s'ils présentent toujours le comportement conditionné.

correspond à l'une ou l'autre de plusieurs entités expérientielles de référence qui sont le fait de situations de stimulation différentes. Dans la dernière méthode (échelle ordonnée), le sujet doit donner une évaluation quantitative, sur une échelle ordonnée, de certains attributs de l'entité expérientielle sonore. Plusieurs facteurs peuvent faire varier l'expérience sonore résultant de situations de stimulation physiquement identiques. Comme le mentionnent Green et Swets (1966):

"An even greater problem than physiological variability that had to be faced in the design of psychophysical methods was the problem of psychological variability. It was obvious from the start that the calculated value of the sensory threshold is heavily influenced by "nonsensory" factors. The history of psychophysics is in large part a history of analytical and experimental critique of methods centered about this problem. The basic methods have been refined again and again to minimize such extraneous contributions to the calculated value of the threshold as the subject's "timidity, warming up and anxiety", and his conscious and unconscious criteria for making a positive response." [p.118]

Au cours des années cinquante, ces problèmes ont été largement éliminés par la mise au point de la 'théorie de la détection du signal' (Green and Swets, 1966). Celle-ci offre un référentiel mathématique commun²³ pour les différentes procédures psychophysiques et permet de comparer les résultats en tenant compte de l'effet des différentes variables physiologiques et psychologiques. La théorie de la détection du signal unifie ces différents facteurs sous une seule bannière, le critère de réponse, et permet d'estimer ce critère tout en différenciant les réponses proprement sensorielles.

²³ Basé sur un rationnel probabiliste.

Le paradigme psychophysique a généré une quantité impressionnante de travaux et ces travaux peuvent être classifiés en fonction des attributs des entités expérientielles sonores qui sont produites par les différentes situations de stimulation. Les principaux *attributs de caractérisation* des entités expérientielles sonores qui ont été étudiés dans les recherches psychophysiques sont: la hauteur ("pitch"), l'intensité, le timbre, la durée et le rythme. Les recherches psychophysiques ont porté, en grande majorité, sur les attributs de hauteur et d'intensité d'une entité sonore ("tone" ou son élémentaire)²⁴. A cet égard, les recherches ont permis de "cartographier" les caractéristiques physiques de ces deux attributs en termes de seuils différentiels et de seuils absolus et ce, pour les différentes "variétés" d'organismes (Fletcher, 1929; Stevens and Davis, 1936). Ces recherches ont également révélé que les "métriques" des organismes biologiques s'écartaient d'une linéarité absolue: les seuils différentiels changeaient pour les différentes fréquences et les seuils absolus étaient plus bas pour la portion centrale des fréquences auxquelles étaient sensibles les différents organismes. Dans ces contextes de stimulation composés d'une seule fréquence, la relation entre les attributs biologiques de hauteur et d'intensité et les attributs artificiels (physiques) de fréquence et d'intensité est donc assez facile à établir. Mais la situation se complique énormément lorsqu'on essaie de corréler la hauteur et l'intensité d'une entité sonore avec un signal sonore composé de plusieurs fréquences à intensité différente. Pour interpréter l'expérience sonore résultant de ces signaux plus "complexes", la

²⁴ Cela est facilement compréhensible si l'on considère que ces deux attributs entretiennent une "relation privilégiée" avec les attributs physiques de fréquence et d'intensité.

psychophysique s'est longtemps servie de la loi d'Ohm (1843) qui avait été proposée par Helmholtz (1885) comme hypothèse de la fonction cochléaire. La loi d'Ohm énonce qu'un signal sonore est décomposé en ses différents constituants spectraux par la cochlée. Lors de l'établissement de la métrique biologique de hauteur et d'intensité, les travaux psychophysiques ont mis très rapidement en évidence que l'expérience sonore ne correspond pas à l'expérience qui devrait résulter d'un fonctionnement des structures cochléaires selon la loi d'Ohm. La résolution imparfaite des différentes composantes spectrales du signal en représente certainement l'évidence la plus frappante. Cette résolution imparfaite se traduit par "l'incapacité psychologique" de retrouver certaines composantes spectrales qui font partie de la description spectrale du signal sonore (Nordmark, 1978). De plus, l'existence d'une largeur de bande critique ("critical band") (Feldkeller et Zwicker, 1956), le phénomène des battements (Sauveur, 1701), celui des tons de combinaison ("combination tones") (Plomp, 1976) ainsi que le 'cas de la fondamentale manquante' ("missing fundamental") (Nordmark, 1978; Fletcher, 1929) vont tous à l'encontre de l'image simplifiée d'une simple décomposition spectrale du signal sonore par l'organisme sonore.

La caractérisation de la hauteur d'une entité sonore résultant d'un signal sonore complexe (plusieurs composantes spectrales) a donné lieu à une polémique d'envergure qui occupe, encore aujourd'hui, une proportion importante de la littérature. Cette polémique porte sur différentes hypothèses relatives aux processus de connaissance qui permettent de caractériser la hauteur d'un son complexe. Elle oppose les tenants d'une théorie *temporelle*, où les structures sonores supérieures analysent le "pattern" temporel des

signaux nerveux qui émergent de la cochlée (Schouten, 1940; Nordmark, 1978), aux tenants d'une théorie *spatiale* où les mêmes structures analysent plutôt le "pattern" spatial des signaux nerveux (Whitfield, 1978; Houtsma & Goldstein, 1972). L'argument central des tenants de la théorie temporelle est le suivant: si la fonction cochléaire est de décrire les composantes spectrales du signal sonore par l'emplacement spatial le long de la cochlée et par l'analyse spatiale conséquente des fibres nerveuses qui sont branchées sur la cochlée, comment est-il possible d'expliquer que nous pouvons attribuer une hauteur à un son complexe qui ne comprend pas d'énergie à la région spectrale correspondante? D'un côté comme de l'autre, les expériences ingénieuses se sont multipliées pour faire valoir un point de vue et, pour expliquer toutes ces données expérimentales, de nombreux modèles ont été élaborés par les différents chercheurs (Whitfield (1978); Goldstein (1973); Nordmark, (1978)). Mais la situation reste extrêmement confuse et aucun consensus ne semble se dégager. Pourquoi?

La psychophysique, comme son nom l'indique, repose largement sur les concepts et les techniques de la physique et de l'acoustique. Or, dans le contexte de la connaissance sonore, les concepts physiques n'ont de sens que lorsqu'on essaie de les interpréter dans une perspective cognitive. Les difficultés rencontrées par les travaux de recherche qui s'insèrent dans le paradigme psychophysique originent, à notre avis, d'une confusion entre les perspectives explicatives que sont les perspectives physique et cognitive. La psychophysique se situe dans une *perspective physique* et utilise des *concepts physiques* pour interpréter des *processus cognitifs* plutôt que de situer dans une *perspective cognitive*, d'utiliser des *concepts cognitifs* et des *concepts*

physiques pour interpréter des *processus cognitifs*. Cela devient de plus en plus évident à mesure que les recherches psychophysiques essaient de rendre compte d'attributs d'entités expérientielles sonores qui "s'éloignent" des attributs caractéristiques des "entités" produites par les structures artificielles (physiques) des deux premiers niveaux de caractérisation. En d'autres termes, comme il n'existe pas de structures de caractérisation artificielles correspondant aux structures supérieures des organismes biologiques, il n'existe pas non plus d'attributs "d'entités physiques" qui seraient produites par de telles structures artificielles et qui pourraient être corrélées aux attributs des entités expérientielles sonores qui sont le produit, elles, de structures biologiques supérieures. C'est dans ce contexte qu'il devient de plus en plus difficile d'interpréter l'expérience sonore dans une perspective physique et que la nécessité d'une perspective cognitive devient primordiale.

Cette perspective cognitive se reflète d'ailleurs dans les travaux de recherche qui portent sur ces attributs "de plus haut niveau" que sont le timbre (Plomp, 1971; Erickson, 1975; Risset, 1978), la durée et le rythme (Bregman & Campbell, 1971, Van Noorden, 1975, Deutsch, 1978), ainsi que dans les travaux qui portent sur l'expérience de la musique (Deutsch, 1978; Tenney & Polansky, 1980) et du langage (Sanders, 1977). Les concepts mis de l'avant par le mouvement de la 'psychologie de la forme' ("Gestalt theorie") (Ellis, 1938) et repris par l'approche contemporaine de la 'reconnaissance des formes' ("pattern recognition") (Dodwell, 1970; Reed, 1978) représentent l'essentiel de cette perspective cognitive. L'expérience sonore y est perçue comme le résultat et le reflet de processus cognitifs

organisateurs, résultat qui s'exprime sous la forme de *gestalts* ('bonnes formes') ou de "patterns".

"The given is itself in varying degrees "structured" ("*gestaltet*"), it consists of more or less definitely structured wholes and whole-processes with their whole-properties and laws, characteristic whole-tendencies and whole-determination of parts. "Pieces" almost always appear "as parts" in whole processes." (Wertheimer, 1922; cité dans Köhler, 1947)

"There is, in the first place, what is now generally called the organization of sensory experience... What we actually perceive are, first of all, specific entities such as things, figures, etc., and also groups of which these entities are members. This demonstrates the operation of processes in which the content of certain areas is unified, and at the same time segregated from its environment." (Köhler, 1947) [p.71]

Dans la perspective "gestaltiste", ce travail organisateur s'explique par un certain nombre de principes tels que:

- le principe figure-fond selon lequel toute expérience se constitue d'un sous-ensemble privilégié, la figure, et un fond qui sert de contraste organisateur pour la figure.
- les critères de l'organisation expérientielle en figure et en fond: bonne continuité, proximité, symétrie, fermeture, destinée commune etc.

Le principe figure-fond met en évidence le travail *sélectif* des organismes de connaissance. Ceux-ci doivent continuellement choisir des sous-ensembles dans la totalité de l'information qui atteint l'organisme à chaque instant et faire porter prioritairement le travail de connaissance sur ces sous-

ensembles. Pour l'expérience sonore, ce principe est apparent dans ce qui est considéré comme une situation sonore très complexe, le "cocktail party", situation où l'auditeur doit faire porter son attention sur un sous-ensemble sonore, la voix de son interlocuteur, qui est littéralement noyée dans un signal sonore composé de plusieurs manifestations énergétiques différentes. Ce principe est également apparent dans l'écoute musicale où l'auditeur sera amené à suivre une voix musicale particulière aux dépens des autres éléments de la pièce musicale.

"When we listen to music, we do not simply process each element as it arrives; rather we form sequential groupings out of combinations of elements. Once such groupings are formed, there is further a tendency for one to come to the foreground of our attention, while others are relegated to the background. The stability of such figure-ground organization depends on the type of music presented. Thus in some music, such as accompanied songs, one voice tends strongly to be heard as the figure and the other as the ground. In contrast, in contrapuntal music, such as canons and fugues, we attempt as much as possible to attend to all voices, which results in our fluctuating between alternative modes of figure-ground organization." (Deutsch, 1980) [p.201]

L'existence de figures et de fonds pose le problème de la constitution de ces figures et de ces fonds par l'organisme sonore et c'est là que les différents critères d'organisation expérientielle mentionnés plus haut entrent en ligne de compte. Le critère de proximité "fréquentielle" des différentes voix instrumentales qui composent une mélodie est utilisé par Deutsch (1980) pour expliquer la ségrégation expérientielle de ces mêmes instruments. Ce même critère peut être utilisé pour expliquer la "pseudopolyphonie" qui émerge d'une séquence rapide de notes d'un seul instrument jouée dans deux

régions spectrales différentes et il peut également expliquer les séquences expérientielles ("auditory streams") démontrées par Bregman & Campbell (1971). Le critère de 'bonne continuation' représente également une explication intéressante du fait que l'identification de l'ordre temporel d'une séquence de notes qui varient en fréquence est facilitée lorsque les séquences varient "unidirectionnellement", i.e. lorsqu'elles sont uniquement montantes ou uniquement descendantes (Divenyi & Hirsch, 1975).

Les entités expérientielles sonores peuvent donc être décrites comme des "gestalts" qui résultent d'un processus organisateur obéissant aux différents critères d'organisation expérientielle que nous venons de mentionner. Ces "gestalts", ces entités expérientielles, sont dotées d'un certain nombre d'attributs de caractérisation comme l'intensité, la hauteur, le timbre, la durée, le rythme ou la forme mélodique.

Mais ces attributs n'appartiennent pas tous aux mêmes *types* d'entités expérientielles: l'intensité, la hauteur, le timbre peuvent être associés à des entités sonores plus "élémentaires", i.e. les notes (ou les "tones"), alors que le rythme ou la forme mélodique sont plutôt les attributs de caractérisation d'entités sonores plus globales, i.e. les phrases musicales, qui sont une "gestalt" de notes. Dans le contexte de la connaissance musicale, Tenney et Polansky (1980) identifient ainsi plusieurs *niveaux* d'entités sonores qui représentent des 'unités de Gestalt temporelles' distinctes (temporal Gestalt units). Ces unités sont: les *éléments* (qui correspondent aux notes traditionnelles), les "*clangs*" (qui sont des groupements de notes), les *séquences* (qui sont des groupements de "clangs"), les *segments* (qui sont des groupements de séquences), et finalement les *sections* (qui sont des

groupements de segments). Ces auteurs ne sont pas restrictifs quant au nombre de niveaux qui sont impliqués dans l'expérience sonore. Ils remarquent, entre autres: "...it should be possible to extend the model "downward" to sub-elements levels... In the course of such a process, new "parameters" would emerge-pitch, and perhaps timbre." (Tenney & Polansky, 1980; [p.238-239]). Si nous généralisons les observations de ces auteurs à toutes les entités expérientielles sonores (linguistiques, musicales et autres), cela nous amène à concevoir l'expérience sonore comme une séquence d'*entités sonores*²⁵ "bien formées", de "gestalts", qui résultent de processus organisateurs obéissant à certains critères de sélection expérientielle. Cette séquence expérientielle se constitue d'entités qui appartiennent à différents niveaux de globalité et, à chacun de ces niveaux, les entités se constituent de l'organisation des entités du niveau précédent et servent à leur tour "d'éléments constitutifs" pour les entités du niveau suivant.

La "vision gestaltiste" de l'expérience sonore donne beaucoup d'importance aux caractéristiques hiérarchiques et organisatrices du processus de connaissance sonore de l'organisme. Les concepts de la "gestalt theorie" se situent dans cette perspective cognitive que nous soulevions plus haut, perspective cognitive qui est à la base, croyons-nous, d'une démarche puissante d'explicitation des phénomènes de connaissance. Ces concepts nous

²⁵ Les *entités sonores* ne se manifestent pas nécessairement au foyer expérientiel et c'est pourquoi nous ne les associons pas, du point de vue de la terminologie, aux entités *expérientielles* sonores. Une entité expérientielle sonore est nécessairement une entité sonore mais une entité sonore n'est pas nécessairement expérientielle.

permettent également de relativiser la référence absolue des mesures physiques en suggérant que ces mesures physiques sont également le fruit de processus organisateurs: ceux des instruments physiques qui les produisent. Nous croyons que ces processus organisateurs biologiques et artificiels relèvent d'une seule logique, logique que nous essaierons de développer au prochain chapitre.

Ceci complète la dernière section de ce deuxième chapitre. Nous y avons considéré le dernier élément constitutif de l'entité 'connaissance sonore': l'expérience sonore. Nous avons vu plus précisément que:

- l'expérience sonore se manifeste au foyer expérientiel sous la forme d'entités sonores,
- l'expérience sonore des organismes artificiels résulte uniquement du processus de connaissance des deux premières structures de caractérisation (structure d'interface et d'analyse spectrale), alors que l'expérience sonore des organismes biologiques résulte du processus de connaissance de ces deux premières structures *et* du processus de connaissance des structures supérieures (structures nerveuses),
- les entités sonores des organismes biologiques et artificiels sont dotées d'attributs de caractérisation comme la fréquence, la hauteur, l'intensité, le timbre etc.
- ces entités sonores peuvent appartenir à différents niveaux "expérientiels": notes et phonèmes, phrases mélodiques et parlées, "clangs", "séquences", "segments" etc.,
- dans la perspective gestaltiste, les entités sonores résultent des processus organisateurs de l'organisme: principe figure-fond et critères d'organisation expérientielle, et finalement,

- les entités sonores d'un niveau expérientiel semblent se constituer des entités du niveau précédent: une phrase mélodique se compose de notes, un mot de phonèmes etc.

2.4 Vers une logique de la connaissance sonore

Nous pouvons maintenant tenter de relier milieu, organisme et expérience en une vision globale de la logique sous-jacente à la connaissance sonore. L'organisme est "plongé" dans le milieu et sa fonction est de connaître les manifestations énergétiques qui originent des différents objets physiques qui l'environnent. Nous avons vu que les organismes artificiels et biologiques pouvaient être conçus comme une hiérarchie de structures de caractérisation des manifestations temporelles de l'énergie et nous avons décrit "l'output" de ces structures comme les produits cognitifs spécifiques à chaque niveau de caractérisation. Nous pouvons maintenant associer ces produits cognitifs aux entités sonores des différents niveaux expérientiels. *Nous croyons que les différents niveaux expérientiels qui se révèlent dans l'expérience sonore correspondent aux différents niveaux structuraux de l'organisme, les processus organisateurs qui président à la genèse des entités sonores étant reliés aux mécanismes spécifiques de chaque niveau structural.*

La hiérarchie structurale de l'organisme se reflète dans la hiérarchie expérientielle des différentes "gestalts" qui constituent notre univers expérientiel et les "gestalts" de chaque niveau se constituent (s'organisent) des relations entre les gestalts des niveaux "inférieurs". Nous avons vu que le produit cognitif des structures d'analyse spectrale pouvait être interprété comme une 'mise en relation temporelle' des valeurs successives du produit cognitif de la structure d'interface. Nous interprétons ce processus organisateur de mise en relation temporelle du produit d'une structure

inférieure, comme un exemple de "gestalt" (les valeurs spectrales) qui se constitue de "gestalt" du niveau inférieur (le produit de la structure d'interface). Nous croyons que de tels processus organisateurs se répètent à tous les niveaux de la hiérarchie structurale et qu'ils produisent, à tous ces niveaux, de nouvelles "gestalts" (de nouvelles entités sonores). Les processus organisateurs des deux premiers niveaux de caractérisation des structures biologiques et artificielles sont fort probablement similaires et c'est pourquoi il est possible et fructueux d'établir des "co-relations" entre les produits cognitifs (entités sonores) artificiels et biologiques (ce que fait la psychophysique). Mais le "travail" de connaissance des organismes biologiques va beaucoup plus loin que celui des organismes artificiels: il se poursuit dans les structures supérieures du système nerveux auditif. Ce sont ces processus de caractérisation additionnels qui permettent la genèse d'entités sonores de haut niveau telles que notes, "clangs", séquences etc. (et leurs attributs concomitants: timbre, rythme, "forme" mélodique etc.). Pour corréler ces attributs biologiques à des attributs artificiels (physiques), il faudrait disposer de structures artificielles qui génèrent des entités dotées de tels attributs. Or de telles structures artificielles "de haut niveau" n'existent pas et la psychophysique se retrouve dans une position où elle doit tenter de corréler ces attributs de haut niveau avec des attributs de bas niveau, ce qui est la cause, à notre avis, des principales difficultés de cette méthode d'investigation de l'expérience sonore.

Ces limitations "expérientielles" (ou cognitives) des structures artificielles (physiques) existantes se reflètent également dans les travaux de recherche

sur la reconnaissance 'automatique'²⁶ de la parole. Dans ce domaine, les chercheurs sont confrontés à l'absence d'*invariants* "expérientiels" dans la description artificielle des manifestations énergétiques qui sont caractéristiques du langage parlé. L'organisme humain peut reconnaître, par exemple, le mot 'bonjour' dans plusieurs situations énergétiques différentes (voix de femme ou d'homme, timbres de voix différents, mot prononcé isolément ou dans une phrase), alors que ces mêmes situations énergétiques "apparaissent" très différentes à un organisme artificiel comme le spectrographe. L'organisme humain perçoit, dans ces situations, un invariant expérientiel, alors que l'organisme artificiel n'en perçoit pas. Pisoni (1985) considère que l'absence d'invariants expérientiels aux deux premiers niveaux de caractérisation de l'organisme représente le problème fondamental de la perception du langage parlé:

"The lack of acoustic-phonetic invariance in speech is considered by all speech researchers to be *the* most important problem in the field; this was the major problem uncovered in the late 1940s after the invention of the sound spectrograph; it is still the central problem in the field of speech research today. Despite over 30 years of research on the problem, investigators have been unsuccessful in identifying acoustic segments and properties of the speech waveform that uniquely match the units derived from perceptual analysis.[p. 381]

Cette absence d'invariants expérientiels artificiels peut s'expliquer dans le contexte de nos précédentes remarques. Un invariant expérientiel comme 'bonjour' *n'apparaît pas* au niveau des entités sonores produites par les deux premières structures de caractérisation. Si l'organisme humain est capable de décréter de tels invariants, c'est probablement parce qu'il dispose de ces

²⁶ Nous dirions: reconnaissance *artificielle* de la parole.

structures additionnelles de caractérisation que sont les structures du système nerveux auditif, et l'on peut croire que c'est à ces niveaux d'entités sonores "plus élevés" *qu'apparaissent* les invariants sonores qui sont caractéristiques de notre expérience du langage parlé. Le "décret" d'invariants sonores représente une dimension fondamentale du processus de connaissance sonore et nous reviendrons sur cette notion au prochain chapitre.

Avant de passer à ce troisième chapitre, tentons maintenant de résumer l'essentiel de notre compréhension de la logique de la connaissance sonore. Au premier chapitre de ce travail, nous avons posé l'entité 'connaissance sonore' comme une entité problématique et nous avons défini l'objectif général de ce travail comme une tentative de transformation de l'état problématique de cette entité vers un état moins problématique. Ce deuxième chapitre a été consacré à la première étape de ce processus de résolution de problèmes: établir les attributs des trois entités constitutives de la connaissance sonore dans la perspective de leurs contributions respectives au phénomène de la connaissance sonore. Nous pouvons résumer ces attributs de la manière suivante:

1) Le milieu peut être décrit en termes de manifestations énergétiques qui originent des différents objets physiques qui l'occupent. — Certaines manifestations énergétiques de type mécanique sont reliées au phénomène de la connaissance sonore. Ces manifestations se propagent dans toutes les directions du médium sonore sous la forme d'ondes longitudinales et, en chaque point de l'espace de ce médium, ces ondes longitudinales s'additionnent pour provoquer des variations locales de pression.

2) L'organisme est plongé dans le milieu et sa fonction est de "produire" la connaissance de ces différentes manifestations énergétiques de type mécanique. Nous avons vu qu'un organisme monaural n'occupe qu'un seul point de l'espace du médium de transmission et "permet" tout de même la

plus grande partie de la richesse et de la diversité expérimentale d'un organisme "binaural". Et pourtant, l'énergie qui se manifeste en un seul point de l'espace est le résultat de l'ensemble des manifestations énergétiques qui se produisent dans le milieu. Cela nous a amené à conclure que le processus de connaissance sonore de l'organisme ne pouvait être le fait que d'une caractérisation des manifestations *temporelles* de l'énergie. Ce processus de connaissance s'actualise par le biais d'une hiérarchie de structures de caractérisation. Les organismes artificiels se constituent des structures des deux premiers niveaux de caractérisation (structure d'interface et structure d'analyse spectrale), alors que les organismes biologiques se constituent de ces deux premières structures *et* de structures qui appartiennent vraisemblablement à plusieurs autres niveaux de caractérisation: les structures supérieures du système nerveux auditif.

3) L'expérience sonore représente le résultat appréhendable de ce processus de connaissance des manifestations énergétiques du milieu. Les "produits cognitifs" des structures de caractérisation peuvent être assimilés à différentes entités sonores, alors que le "produit cognitif terminal" du processus de connaissance de toute la hiérarchie structurale d'un organisme peut être assimilé aux entités sonores expérimentales qui se manifestent au foyer expérimental. Les entités sonores de chaque niveau résultent du processus de caractérisation (du "travail" organisateur) qui est spécifique à ce niveau et le résultat de ce processus "s'exprime" dans les différentes valeurs d'attributs de ces entités. Ce processus de caractérisation s'applique aux entités sonores du niveau précédent et peut être conçu comme un processus de "mise en relation" de ces entités "plus primitives". Les entités sonores, qui sont produites à un niveau, servent à leur tour de "matière première" pour le processus de caractérisation du niveau suivant et ce processus de connaissance se poursuit jusqu'au niveau final que représente le foyer expérimental.

Cette conception de la logique de la connaissance sonore reportée, à notre avis, l'essentiel du caractère problématique "résiduel" de l'entité 'connaissance sonore' vers la nature du processus de connaissance sonore de l'organisme. Tous les éléments que nous avons soulevés au cours de ce deuxième chapitre nous ramènent au processus cognitif spécifique qui est "employé" par l'organisme dans sa tentative de connaissance du milieu. Nous savons maintenant que ce processus est hiérarchique, i.e. qu'il se produit en

plusieurs étapes. Nous savons également que ce processus s'effectue par le biais de structures de caractérisation des manifestations *temporelles* de 'ce qui se manifeste' à chaque niveau du processus, et nous savons finalement que ces structures de caractérisation produisent les entités sonores qui constituent notre expérience des sons. Mais quelle est la nature précise des caractérisations temporelles qu'effectue l'organisme? Comment s'effectuent ces caractérisations successives et quelles sont les entités et les attributs de ces différents niveaux de caractérisation? Quel rapport existe-t-il entre la notion d'invariant expérientiel et les différents niveaux du processus de connaissance sonore? Quelle est, en un mot, la *logique* sous-jacente au *processus* de connaissance sonore de l'organisme? C'est ce que nous essaierons de comprendre dans le dernier chapitre de ce travail.

CHAPITRE 3

LE PROCESSUS DE CONNAISSANCE SONORE DE L'ORGANISME

L'étude des trois éléments constitutifs de l'entité 'connaissance sonore' nous a conduit à une certaine conception de la logique sous-jacente à cette forme de connaissance. Ceci a pour effet de "déplacer" (tout en réduisant) le "poids" problématique de cette entité vers le processus de connaissance sonore de l'organisme. Nous avons vu que l'organisme ne peut connaître 'ce qui se manifeste' dans le milieu que par une caractérisation des aspects *temporels* de 'ce qui se manifeste' puisque son lieu d'interaction avec le milieu peut se restreindre à un seul point spatial. C'est donc au travers des manifestations temporelles de 'ce qui se manifeste' en ce point de l'espace que l'organisme "procède" à la connaissance sonore, c'est à dire à l'identification et à la caractérisation des entités sonores qui occupent notre champ expérientiel. Nous avons vu également que ce processus est le fait d'une hiérarchie de structures de caractérisation emboîtées: structure d'interface, structure d'analyse spectrale et structures supérieures. Pour aller plus loin dans notre compréhension de la logique de la connaissance sonore, il nous faut maintenant considérer plus en profondeur ce processus et tenter d'en extraire la logique sous-jacente. Quelle est la 'raison d'être' de cette hiérarchie structurale? Que représentent les entités sonores qui sont produites à chacun de ces niveaux de caractérisation? Comment s'effectue, à

chacun de ces niveaux, le processus de caractérisation des aspects temporels de 'ce qui se manifeste'?

Autant de questions auxquelles l'état actuel de la recherche ne nous permet pas de répondre. Il est vrai que nous connaissons suffisamment bien le processus des deux premiers niveaux de caractérisation pour en avoir élaboré une multitude "d'incorporations" artificielles, tant mathématiques que physiques (ou computationnelles). Mais jamais, à notre connaissance, n'a été posée la question de l'insertion de ces deux premiers niveaux de caractérisation dans la filiation cognitive de la séquence hiérarchique des structures de caractérisation. Quel est le rôle des structures d'interface et celui des structures d'analyse spectrale dans l'ensemble du processus de connaissance sonore des organismes biologiques? Comment ces deux premières étapes de caractérisation contribuent-elles à l'ensemble de ce processus? Comment pouvons-nous les interpréter, en somme, *dans le contexte d'une perspective cognitive?* Dans notre recension des travaux de recherche, nous n'avons jamais rencontré ces questions. Nous avons vu, au contraire, que les descriptions physiques des structures artificielles de ces deux premiers niveaux sont toujours posées comme des référents "objectifs" et absolus auxquels il nous faut absolument rapporter le processus de connaissance "plus flou" et éminemment "subjectif" des organismes biologiques. Par notre prise de position sur le phénomène de la connaissance (chapitre 1) et par notre conception de la logique de la connaissance sonore (chapitre 2), nous croyons avoir fait suffisamment la démonstration de l'inaptitude d'une telle approche.

Où pouvons-nous trouver alors une réponse à nos différentes questions? Est-ce dans les travaux de recherche qui portent sur les structures nerveuses de l'organisme biologique (neuro-anatomie)? Mais l'orientation dominante de ces travaux est également de retrouver les différents paramètres physiques (fréquence, intensité, phase) dans le "comportement" des structures nerveuses, et, là encore, l'absence d'une perspective cognitive élaborée se fait durement ressentir, lorsque ces paramètres disparaissent de la réponse des fibres nerveuses. Dans ces recherches comme dans les recherches psychophysiques sur les attributs des entités de haut niveau (timbre, rythme, forme mélodique etc.), les "explications cognitives" les plus intéressantes se ramènent toujours aux concepts "gestaltistes" ou à leur version plus moderne de la reconnaissance de formes (pattern and features). Or si les différents concepts "gestaltistes" (modernes ou anciens) nous permettent de comprendre globalement l'expérience sonore et la nature organisatrice des processus qui la génère, ils n'ont pas donné lieu encore à des conjectures intéressantes sur les structures qui permettent la connaissance sonore. Les gestaltistes utilisaient les modèles physiques de l'époque, en particulier le modèle du champ de force, pour essayer d'exprimer concrètement le fonctionnement des structures de caractérisation. Il était alors extrêmement difficile de tester expérimentalement ces modèles ou de les traduire en structures cognitives artificielles dont il aurait été possible d'analyser le comportement cognitif. Nous croyons que l'introduction récente des technologies de l'information a changé cet état de choses et que les machines computationnelles nous permettent maintenant d'exprimer concrètement nos hypothèses sur l'organisation des structures de connaissance sonore. Mais

encore faut-il que ces hypothèses soient "ancrées" dans cette perspective cognitive dont nous avons si souvent souligné l'importance.

Ce dernier chapitre portera sur l'élaboration et sur l'expression computationnelle de telles hypothèses. Dans une première section, nous essaierons de présenter une vision cognitive du processus de connaissance sonore en considérant les rapports étroits qui unissent connaissance sonore et temps. Nous essaierons par la suite d'identifier la nature temporelle des processus et des produits de caractérisation des deux premiers niveaux structuraux et nous exprimerons notre conception de ces processus sous la forme d'une simulation computationnelle. Cela nous conduira à présenter quelques hypothèses (non-formalisées) sur le processus et le produit des structures de caractérisation du troisième niveau de connaissance sonore. Nous conclurons finalement en essayant de dégager la logique sous-jacente à l'ensemble du processus de connaissance sonore de l'organisme.

3.1 Temps et connaissance sonore

Dans l'exposé qui suivra, nous aurons à discuter fréquemment d'événements qui se produisent dans la dimension temporelle. Nous nous doterons donc, tout d'abord, d'une modalité générale et uniforme de description des événements qui se produisent dans le temps. Nous situerons les différents événements qui relèvent de la connaissance sonore à partir d'une métrique temporelle que nous qualifierons de *courante* et de *discrète* (Fig. 2).

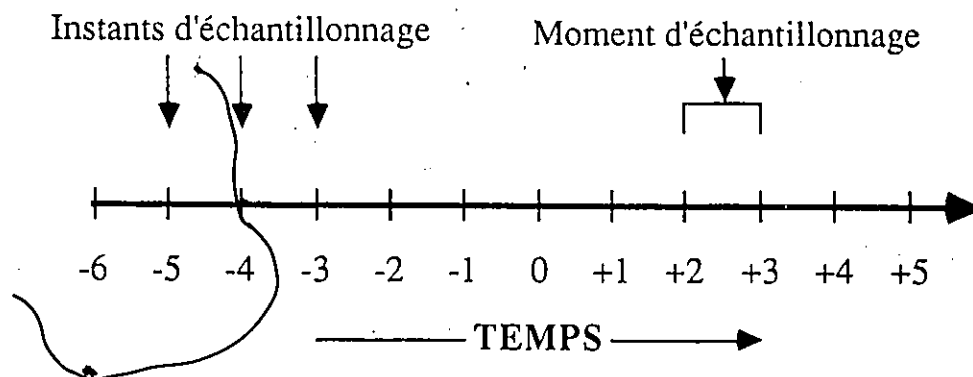


Fig. 2 Une métrique temporelle discrète

Nous entendons par 'courante', une métrique qui situe les événements dans le temps relativement à un événement temporel de référence qui se *déplace* lui-même dans la dimension temporelle et ce, par opposition à une métrique temporelle 'gelée' qui situerait les événements relativement à un événement

temporel de référence qui ne se déplace pas dans le temps. Situer les événements relativement à l'instant de notre naissance, par exemple, correspond à l'utilisation d'une métrique temporelle gelée puisque cet événement de référence (l'événement de la naissance) ne se déplace pas dans le temps: il s'est produit à cet instant particulier du temps et ne s'est jamais reproduit. Pour sa part, une métrique temporelle courante utilise, pour situer les événements dans le temps, un événement de référence qui se produit *constamment* à travers le temps et les événements que l'on veut situer dans le temps ont toujours une position temporelle qui est relative à cet événement de référence 'courant'. Un tel événement de référence est constitué par l'événement de la connaissance au foyer expérientiel. Nous connaissons de la naissance à la mort même si la nature de cette connaissance varie énormément à travers le temps. Dans une métrique temporelle courante, l'événement de la connaissance (l'événement de référence) se produit toujours à l'instant présent, instant présent qui se déplace continuellement dans le temps, instant présent donc qui se déplacerait constamment dans le futur d'un événement temporel de référence 'gelé'. De même, dans une telle métrique, les événements *passés* sont constitués par les événements qui se sont produits avant l'instant présent de connaissance et inversement pour les événements *futurs*. Ainsi, lorsque nous parlerons d'événements du passé, nous parlerons toujours des événements qui se sont passés jusqu'à l'instant présent de la connaissance et non pas d'événements qui se sont passés jusqu'à un autre événement figé dans le temps.

Une métrique temporelle discrète, quant à elle, représente les événements temporels sur des positions temporelles discrètes que l'on peut assimiler à des

"espaces" temporels infiniment petits. Nous désignerons ces positions temporelles discrètes comme des *instants d'échantillonnage* et nous assumerons que ces instants d'échantillonnage sont séparés par un espace temporel fixe que nous appellerons *moment d'échantillonnage*. Ainsi, par exemple, lorsque nous voudrions décrire la valeur de l'énergie à travers le temps, nous parlerons de la valeur de l'énergie à chacun de ces "espaces temporels" infiniment petits que sont les instants d'échantillonnage sans prendre en considération la valeur de l'énergie "entre" ces instants d'échantillonnage (durant un moment d'échantillonnage)²⁷.

Maintenant que nous disposons d'une métrique temporelle nous permettant de situer les événements dans la dimension temporelle, essayons de cerner les caractéristiques essentielles et temporelles des événements cognitifs sonores. Considérons la situation suivante. Nous sommes en train d'écouter une succession de "sons purs"²⁸ séparés par des intervalles de silence. Supposons que nous sommes dans un intervalle de silence et que nous "glissons" sur le temps. Durant un certain nombre d'instants d'échantillonnage, rien ne se produit: "nous entendons le silence"; puis soudain, à un instant d'échantillonnage donné, nous commençons à entendre un son pur et nous allons entendre ce son pendant un certain nombre d'instants d'échantillonnage, i.e. durant tous les instants d'échantillonnage que recouvre l'espace temporel correspondant à sa durée. Le temps passe et nous sommes maintenant rendus au dernier instant d'échantillonnage de la durée de ce son. A l'instant suivant, le silence sera "revenu". Figeons le temps et revenons sur

²⁷ Nous n'assignerons pas de valeur particulière à ces moments d'échantillonnage (secondes ou fraction de secondes).

²⁸ Un "son pur" est un son composé d'une seule fréquence.

notre expérience sonore des instants précédents. Supposons, par exemple, que la durée du son a été de cent (100) moments d'échantillonnage (Fig. 3).

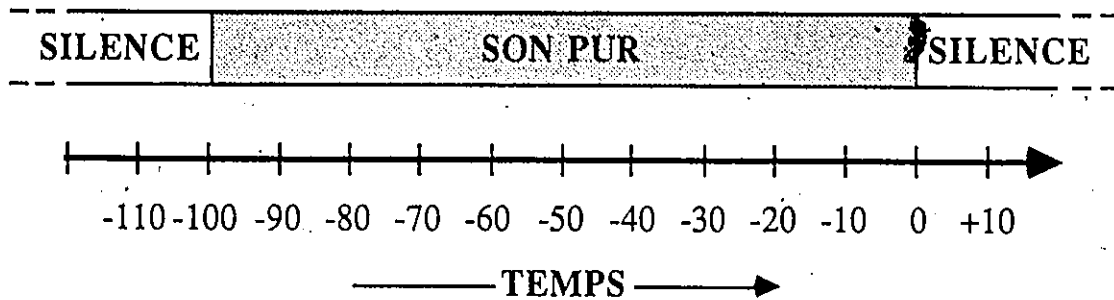


Fig. 3 Un son pur d'une durée de cent (100) moments d'échantillonnage

Il y a donc eu cent (100) entités expérientielles successives qui ont occupé le foyer expérientiel. La première chose à remarquer, c'est que *le processus de connaissance sonore s'est produit à chacun de ces instants d'échantillonnage*²⁹. A chacun de ces instants, nous pouvions nous concentrer sur notre expérience sonore et prendre conscience du son que nous entendions. En "glissant" à travers la dimension temporelle, l'organisme produit de la connaissance sonore à chaque instant de ce glissement. Mais si le *processus* de connaissance est instantané, est-ce que le *produit* de ce processus l'est également? Si nous demandions à un "bon" observateur de nous décrire

²⁹ Cela ne veut pas dire que le résultat du processus de connaissance sonore "occupe" nécessairement le foyer expérientiel de l'organisme à chaque instant d'échantillonnage. Nous ne tiendrons pas compte, dans notre discours, des modalités "attentionnelles" de l'organisme.

précisément ce son, il nous dirait quelque chose comme: "le son a duré quelques secondes; il était d'une certaine hauteur, d'une certaine intensité et possédait une certaine "qualité" (timbre)"³⁰. Qu'est-ce que cette description nous indique sur la nature temporelle du *produit* "instantané" du processus de connaissance sonore? Ce produit "instantané" a une nature temporelle qui s'étend sur un "espace temporel" plus étendu que l'espace temporel infiniment petit de l'instant d'échantillonnage. *Le processus de connaissance sonore s'effectue à chaque instant mais le produit de ce processus ne porte pas nécessairement sur ce qui se manifeste à chaque instant.*

Essayons de nous représenter le fonctionnement du processus de connaissance sonore à travers le temps de ce son. Au premier instant d'échantillonnage où le son commence (l'instant -99, si nous considérons que la fin du son s'est produite à l'instant 0), l'organisme génère une première entité au foyer expérientiel. Cette entité possède certains attributs de caractérisation. Au deuxième instant d'échantillonnage (instant -98), l'organisme génère une deuxième entité qui possède également certains attributs de caractérisation. Ce qui est remarquable, *c'est que le processus de connaissance sonore "décrète" que 'ce qui s'est manifesté' jusqu'au deuxième instant d'échantillonnage (instant -98) appartient au même phénomène que 'ce qui s'est manifesté' jusqu'au premier instant d'échantillonnage (instant -99) puisque la caractérisation de la seconde entité (l'entité de l'instant -98) porte sur 'ce qui s'est manifesté' durant les deux moments d'échantillonnage qui vont jusqu'au deuxième instant d'échantillonnage. Cette entité se*

³⁰ Cette description représente en fait la description de l'entité expérientielle de l'instant 0 (l'entité du dernier instant "d'existence" du son pur).

caractérise, en effet, par un attribut de durée qui a la valeur de *deux* moments d'échantillonnage³¹ (à cet instant, le son pur a duré deux moments). A l'instant d'échantillonnage suivant (le troisième depuis le début du son, l'instant -97), le processus de connaissance produit une autre entité expérientielle. Cette entité reflète encore que le processus de connaissance sonore "considère" que ce qui s'est manifesté jusqu'au troisième instant d'échantillonnage "appartient" également au même phénomène que ce qui s'est manifesté jusqu'au deuxième instant d'échantillonnage puisque l'entité de ce troisième instant comporte un attribut de durée de trois (3) moments. Pendant toute la durée du son, le processus de connaissance sonore va continuer à produire des entités dont l'attribut de durée augmente progressivement³². Ces entités sont le reflet d'un processus de connaissance qui possède cette caractéristique de décréter que 'ce qui se manifeste' à travers le temps appartient aux mêmes phénomènes.³ *Cette capacité de production d'invariants est une propriété fondamentale du processus de connaissance sonore. La production de ces invariants représente, à notre avis,*

³¹ L'espace temporel d'un moment d'échantillonnage jusqu'au "premier" instant d'échantillonnage plus l'espace temporel d'un moment d'échantillonnage entre ce "premier" instant et le deuxième instant d'échantillonnage.

³² Dans notre métrique discrète, nous considérons que de "nouvelles" entités sont produites à chaque instant d'échantillonnage puisque nous "brisons" le temps en unités discrètes. Par sa nature même, cette brisure discrète provoque une impression de discontinuité. Mais cette impression ne signifie pas pour autant que l'organisme "perçoit" une entité "autre" (ou différente) à chaque instant du flux expérientiel. Les entités "autres" (ou différentes) apparaissent plutôt lorsque le processus de connaissance sonore cesse de décréter que 'ce qui se produit' à un instant donné appartient au même phénomène et commence à décréter que 'ce qui se produit' à cet instant appartient à un nouveau phénomène.

la finalité intrinsèque de ce processus et nous croyons que tout effort d'explicitation du processus de connaissance sonore doit porter sur les mécanismes qui permettent la genèse de ces invariants.

Le processus de connaissance sonore est donc un processus qui s'effectue à chaque instant, et qui *relie* 'ce qui se manifeste' à chacun de ces instants à ce qui s'est manifesté dans le passé de l'instant présent d'échantillonnage. Ce processus de *mise en relation temporelle* des événements ('ce qui se manifeste') successifs se reflète dans les entités qui sont produites à chaque instant (et dans les attributs de ces entités). Or nous avons vu qu'il existe une hiérarchie dans les structures de caractérisation de l'organisme et dans les entités sonores qui sont produites par ces structures de caractérisation. Quelles relations existe-t-il entre cette hiérarchie de niveaux de caractérisation et l'existence d'invariants? Nous croyons que les structures de caractérisation représentent des mécanismes de "création" d'invariants aux différents niveaux de la hiérarchie structurale, et que les entités sonores des différents niveaux expérientiels (et leurs valeurs d'attributs) représentent "l'expression" successive de la création de plusieurs types d'invariants "sonores". Cela veut dire, entre autres, qu'il existe, à chaque instant d'échantillonnage, *plusieurs* entités sonores synchroniques (au même instant), qui sont chacune le reflet du décret d'invariants par une structure d'un niveau donné de caractérisation. Considérons à nouveau l'exemple des sons purs. Si le processus de connaissance sonore génère, à chaque instant, des entités sonores "au niveau des sons purs", il génère également, au même instant, des entités sonores qui appartiennent à des niveaux de caractérisation différents. Supposons que cette séquence de sons purs représente la chanson

"Frère Jacques". Nous savons que cette chanson se constitue d'une série de courtes phrases mélodiques qui sont constituées, elles-mêmes, des différents sons purs. Ces phrases mélodiques peuvent être considérées comme les entités sonores du niveau supérieur de caractérisation³³. Elles sont donc également le reflet du décret d'un invariant, mais d'un invariant qui n'est pas du même ordre (ou du même niveau) que l'invariant du niveau des sons purs. Cet invariant est alors la phrase mélodique qui se poursuit durant un certain temps et qui s'arrête, à un certain instant, pour céder la place à une nouvelle entité 'phrase mélodique'. On peut même se représenter la mélodie au complet comme un autre niveau d'entité sonore et imaginer qu'il existe d'autres entités sonores "sous" le niveau des sons purs. A chaque nouveau "pas en avant" de l'organisme dans la dimension temporelle, celui-ci produit conjointement plusieurs entités sonores (décrète plusieurs invariants) à tous les niveaux de caractérisation. Quelle relation existe-t-il entre toutes ces entités? Sont-elles produites en parallèle (indépendamment les unes des autres) ou est-ce que la production d'une entité à un niveau de caractérisation donné est "dépendante" de la production du niveau "inférieur"? Nous croyons que le processus de connaissance sonore s'effectue selon la deuxième hypothèse: les entités sonores des niveaux successifs de caractérisation se "construisent" à partir des entités des niveaux inférieurs. En termes d'invariants, cela signifie que le décret d'invariants de plus 'haut niveau' est dépendant du décret d'invariants de plus 'bas niveau'. Les deux éléments qui nous conduisent à cette prise de position sont:

³³ Dans la terminologie de Tenney & Polansky (1980), ces entités correspondent aux "unités de gestalt temporelles" des éléments et des "clangs".

- l'organisation séquentielle des structures de caractérisation (ces structures sont placées les unes à la suite des autres), et,

- l'impression générale que les entités sonores de plus 'haut niveau' se constituent des entités sonores de plus 'bas niveau' (les phrases mélodiques se constituent des notes, les mots du langage parlé des phonèmes etc.).

Le milieu est le point d'origine de la connaissance sonore d'un organisme. Le milieu est occupé par une multitude d'objets physiques qui imposent leur structure particulière aux différentes formes de manifestations énergétiques qui les atteignent. Ces objets réagissent à ces manifestations énergétiques en émettant à leur tour différentes manifestations énergétiques dont les manifestations de type mécanique qui sont impliquées dans le phénomène de la connaissance sonore. La fonction de l'organisme est de connaître ces différentes manifestations, et il y arrive par le biais d'un processus de connaissance qui repose sur une hiérarchie emboîtée de structures de caractérisation. Chaque structure de caractérisation produit des entités sonores dotées d'un certain nombre d'attributs de caractérisation et chaque entité sonore est le reflet d'une mise en relation temporelle particulière de 'ce qui se manifeste' à chaque instant, avec 'ce qui s'est manifesté' dans le passé. Cette hiérarchie cognitive se construit niveau après niveau et la connaissance qui est caractéristique d'un niveau repose sur la connaissance qui a été élaborée aux niveaux inférieurs. 'Ce qui se manifeste' à travers le temps et qui est le fait, originellement, des manifestations énergétiques du milieu, prend la forme, pour l'organisme, de plusieurs niveaux d'invariants. Ces 'mêmes phénomènes' diachroniques sont caractérisés: le processus de connaissance sonore leur attribue différentes caractéristiques aux différents

niveaux du "processing": hauteur, durée, intensité, timbre, forme mélodique, rythme etc. Finalement, parmi toutes les entités de ces différents niveaux, certaines sont acheminées au foyer expérientiel et constituent les entités expérientielles sonores de notre "conscience" au monde des sons.

Dans une telle conception de la connaissance sonore, le processus de connaissance sonore est un processus essentiellement *conjectural*. Les différents mécanismes cognitifs qui composent ce processus (les différentes structures de caractérisation ou encore les différents processus "organiseurs" de ce processus général) représentent une *tentative* de l'organisme pour générer une connaissance "adaptative" des manifestations énergétiques qui se produisent dans le milieu. Les entités expérientielles ne "sont" pas les objets physiques du milieu; elles "sont" le résultat d'une interaction entre le milieu et le processus de connaissance sonore. Si ces "représentations" du milieu permettent à l'organisme d'agir efficacement, c'est à dire d'engendrer des séquences interactionnelles qui lui permettent de survivre, nous pouvons considérer que le processus de connaissance sonore représente une "bonne" conjecture du milieu.

Nous allons maintenant tenter de donner une forme plus concrète à cette vision abstraite du processus de connaissance sonore en essayant d'interpréter les deux premiers niveaux de ce processus à partir des concepts qui ont été avancés jusqu'ici. Cette interprétation se voudra également plus formelle et se doublera d'une simulation computationnelle des deux premières structures de caractérisation.

3.2 Le premier niveau du processus de connaissance sonore

La première "prise de connaissance" de l'organisme s'effectue par le biais de la structure d'interface. Celle-ci est "plongée" dans le milieu et réagit à ce qui s'y manifeste. Nous avons vu à la section 2.2.1 que la structure d'interface était constituée d'une membrane élastique dont la position "au repos" correspond à la pression normale ou "non-perturbée" du médium. Cette position varie lorsque l'effet combiné des ondes longitudinales modifie les valeurs de pression locale. La membrane se déforme vers "l'intérieur" de l'organisme lorsque la pression augmente au delà de cette valeur de pression normale, et se déforme vers "l'extérieur" de l'organisme lorsque la pression diminue en deçà de cette même valeur. Nous avons vu également que la ~~valeur de la pression~~, et conséquemment la position de la structure d'interface, correspond à la quantité totale d'énergie (la puissance instantanée) qui se manifeste à cet endroit et à cet instant. La structure d'interface est une structure *cognitive*: elle permet la *connaissance* de 'ce qui se manifeste' dans le milieu et cette connaissance prend la forme d'une entité sonore qui est "produite" par la structure d'interface à chaque instant d'échantillonnage, entité qui ne possède qu'un seul attribut de caractérisation: l'intensité énergétique. Au niveau de cette structure, 'ce qui se manifeste' à travers le temps représente *un seul et même* phénomène, phénomène qui se modifie au fil du temps et qui se reflète dans les variations successives de la quantité énergétique. *Cette entité à attribut unique³⁴ peut être considérée*

³⁴ La durée de cette entité (de cet invariant) est infinie: elle existe depuis qu'existe la structure d'interface.

comme le seul invariant sonore de ce niveau du processus de connaissance sonore. Un organisme qui ne serait doté que d'une structure d'interface n'aurait accès qu'à la connaissance d'un seul "son" dont le seul attribut serait l'intensité. Dans le cas d'une manifestation énergétique de type "son pur" (une sinusoïdale), cet organisme entendrait le même son mais ce son changerait successivement (à chaque instant) de valeur d'intensité et n'aurait ni hauteur, ni timbre, ni aucun autre attribut.

Quelle est l'invariance sous-jacente à la production de cette entité? Qu'est-ce qui ne varie pas et qui décrète successivement cette entité unique? La structure d'interface possède des propriétés physiques particulières (masse, élasticité) qu'elle "oppose" aux manifestations énergétiques du milieu. Ces propriétés physiques déterminent plus particulièrement *le comportement temporel* de la structure d'interface en réaction aux manifestations énergétiques qui l'atteignent. Ce qui ne varie pas, ce qui est le fondement de l'invariance à ce niveau du processus de connaissance sonore, c'est ce comportement temporel. La structure d'interface "oppose" toujours le *même* comportement temporel aux manifestations énergétiques du milieu et cette "opposition" se reflète, à chaque instant, dans la production d'une seule entité à un seul attribut.. Comment pouvons-nous décrire ce comportement? Considérons l'exemple suivant de positions de la structure d'interface pour cinq instants d'échantillonnage³⁵ (Fig.4).

³⁵ Dans tous les exemples qui suivront, nous allons représenter la position "normale" de la structure d'interface par la position 0. Les positions de la membrane vers "l'intérieur" (pression plus élevée que la normale) seront représentées par des chiffres positifs et les positions de la membrane vers "l'extérieur" (pression plus basse que la normale) seront représentées par des chiffres négatifs.

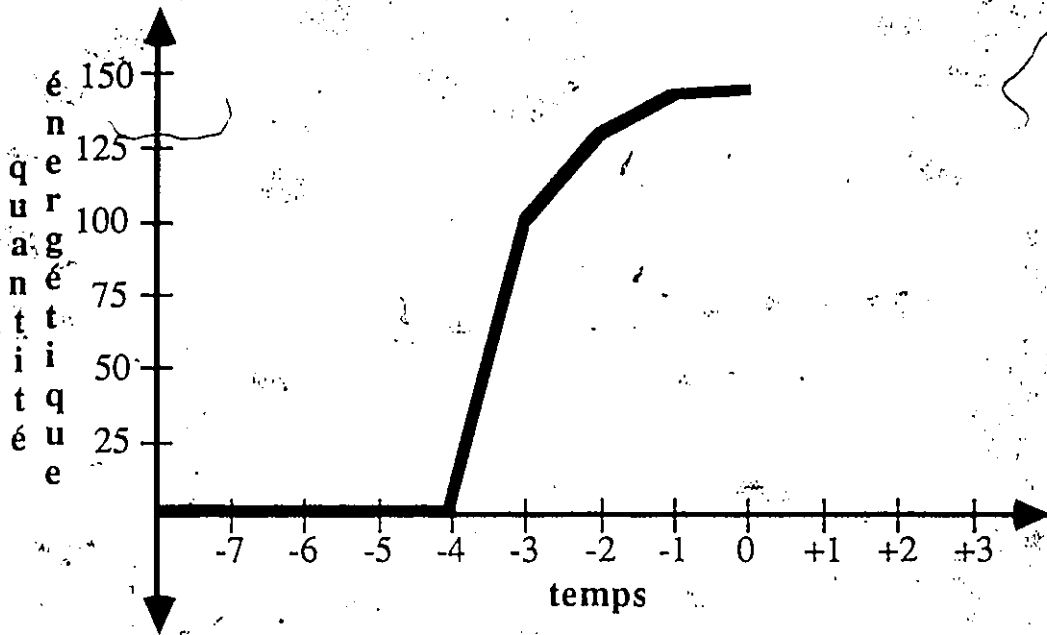


Fig. 4 Une séquence de positions de la structure d'interface

Supposons que la position de la structure d'interface a été de 0 depuis moins l'infini temporel jusqu'à l'instant -4 d'échantillonnage. De l'instant -4 à l'instant -3, cette position passe de 0 à +100. Durant le moment d'échantillonnage qui va de l'instant -4 à l'instant -3, la structure d'interface a été "confrontée" à ce qui s'est manifesté dans le milieu, y a opposé ses propriétés physiques particulières et a réagi en se déplaçant, en un moment, à la position +100. Nous savons que la position de la structure d'interface reflète la pression locale du médium ambiant et que cette pression locale constitue une mesure de la quantité d'énergie qui se manifeste à cet endroit. *La position de la structure d'interface reproduit exactement les variations d'intensité des manifestations énergétiques. Son comportement temporel est tel qu'il lui est possible de suivre exactement ces variations.* Essayons d'exprimer plus systématiquement ce comportement temporel. Nous

pourrions dire que la position de la structure d'interface à chaque instant dans le temps est égale à sa position de l'instant précédent plus la variation de la quantité énergétique durant le dernier moment d'échantillonnage. La position de la structure d'interface à l'instant -3 (PSI₋₃), par exemple, est égale à la position de la structure d'interface à l'instant précédent (instant -4) plus la variation de la quantité énergétique entre ces deux instants, variation qui correspond à la différence de position de la structure à l'instant -3 et à l'instant -4. Exprimé plus formellement, cela donne:

$$\text{PSI}_{-3} = \text{PSI}_{-4} + \text{variation énergétique sur le moment } (-3, -4)$$

$$\text{PSI}_{-3} = \text{PSI}_{-4} + (\text{PSI}_{-3} - \text{PSI}_{-4})$$

$$\text{PSI}_{-3} = 0 + (100 - 0)$$

$$\text{PSI}_{-3} = 100$$

Si nous exprimons, de la même manière, la valeur de position de l'instant suivant (instant -2), cela donne:

$$\text{PSI}_{-2} = \text{PSI}_{-3} + (\text{PSI}_{-2} - \text{PSI}_{-3})$$

$$\text{PSI}_{-2} = 100 + (130 - 100)$$

$$\text{PSI}_{-2} = 100 + 30$$

$$\text{PSI}_{-2} = 130$$

La nouvelle valeur de position de la structure d'interface correspond à l'addition des deux termes de droite dans l'équation. Le premier terme correspond à la valeur précédente de position (cette valeur est "transportée" à travers un moment d'échantillonnage) et le deuxième terme correspond à la



différence de position entre les deux instants. Le comportement temporel de la structure d'interface est présent dans les deux termes de l'équation même s'il n'est pas apparent. Nous pouvons le rendre apparent en multipliant les deux termes de l'équation par une *constante temporelle* (k) égale à 1, ce qui nous donne pour l'instant -3:

$$PSI_{-3} = (PSI_{-4} * k) + ((PSI_{-3} - PSI_{-4}) * k)$$

$$PSI_{-3} = (0 * 1) + ((100 - 0) * 1)$$

$$PSI_{-3} = 100$$

et pour l'instant -2:

$$PSI_{-2} = (PSI_{-3} * k) + ((PSI_{-2} - PSI_{-3}) * k)$$

$$PSI_{-2} = (100 * 1) + ((130 - 100) * 1)$$

$$PSI_{-2} = 100 + 30$$

$$PSI_{-2} = 130$$

La multiplication de ces deux termes par une constante temporelle égale à un (1) ne change rien au résultat final mais elle nous permet de mettre en évidence le comportement temporel de la structure d'interface en réaction aux manifestations énergétiques qui l'atteignent. Reprenons les deux termes de droite de l'équation. Le deuxième terme représente les variations de la quantité énergétique durant le moment d'échantillonnage qui vient de s'écouler. Ce terme est multiplié par la constante temporelle égale à un (1); cette variation est "prise" dans sa totalité par la structure d'interface. Le premier terme représente la position de la structure d'interface à l'instant

précédent et cette position équivaut à la quantité énergétique totale de l'instant précédent: elle est le résultat d'un processus récursif de calcul de *toutes* les variations énergétiques antécédentes 'sur un moment'. La valeur de ce terme est "transportée" de l'instant précédent à l'instant présent tout en étant également multipliée par la constante temporelle égale à un (1). Ce premier terme est donc également repris dans sa totalité et est ajouté au deuxième terme pour nous donner la nouvelle position de la structure d'interface à l'instant présent. La structure d'interface réagit à la *totalité* de la variation énergétique sur un moment (deuxième terme) et ajoute à cette réaction, sur un moment, la *totalité* de sa réaction sur tous les moments antécédents (premier terme). Si nous exprimons notre équation plus "concrètement", cela revient à dire que:

La position de la structure d'interface à l'instant présent (0)	≡	La totalité de la variation énergétique jusqu'à l'instant précédent (-1)	+	La totalité de la variation énergétique durant le dernier moment (-1 à 0)
---	---	--	---	---

Cette équation est une manière de décrire le comportement temporel de la structure d'interface. *Le comportement temporel de cette structure représente un comportement temporel parmi d'autres. Si nous changions la valeur de la constante temporelle, nous obtiendrions d'autres valeurs d'entité en réaction aux mêmes manifestations énergétiques mais ces valeurs d'entité ne correspondraient plus à celles de la structure d'interface.* La structure d'interface constitue une manière particulière de réagir à 'ce qui se manifeste' à travers le temps; elle effectue une *mise en relation temporelle* particulière des événements énergétiques et l'équation exprime cette mise en relation temporelle particulière. 'Ce qui se manifeste' est "relié" d'une

certaine manière à 'ce qui s'est manifesté' dans le passé. Nous désignerons cette 'mise en relation temporelle' comme un *groupement temporel*. La structure d'interface "groupe" (relie, met ensemble) temporellement, à chaque instant, les événements énergétiques qui se sont produits jusqu'à cet instant et le résultat de ce groupement s'exprime par la valeur d'attribut de l'entité sonore. C'est dans ce groupement temporel particulier que réside l'invariance sous-jacente à l'entité sonore de ce niveau: la structure d'interface groupe toujours de la *même* manière ces événements énergétiques.

La structure d'interface représente le premier niveau du processus de connaissance sonore. Ce premier niveau se caractérise par le décret d'un seul invariant: l'entité sonore produite à chaque instant par la structure d'interface. Cette entité existe à travers le temps, affichée des valeurs successives d'intensité et résulte du groupement temporel particulier qu'effectue la structure d'interface. Le processus de connaissance sonore se poursuit au niveau suivant de caractérisation: le niveau de l'analyse spectrale. A chaque instant d'échantillonnage, la structure d'analyse spectrale "reçoit" l'entité sonore du premier niveau et décrète de nouveaux invariants sonores sous la forme d'entités sonores multiples. Nous allons maintenant essayer de comprendre la nature de ces entités, celle de leurs attributs de caractérisation, ainsi que la nature des groupements temporels particuliers qui génèrent ces entités et qui constituent les invariants de second niveau du processus de connaissance sonore.

3.3 Le deuxième niveau du processus de connaissance sonore

L'entité sonore de la structure d'interface est le "matériau de base" du deuxième niveau du processus de connaissance sonore. Chez l'organisme humain, la position de la structure d'interface est transmise par la chaîne d'osselets de l'oreille moyenne, et le mouvement de la membrane élastique de la fenêtre ovale de l'oreille interne correspond au mouvement du tympan. La cochlée reçoit cette valeur d'attribut et applique un certain processus de caractérisation dont le résultat se reflète dans les différents déplacements de la membrane basilaire. Quelle est la nature de ce processus de caractérisation et quelle est, en général, la nature du processus de caractérisation des structures d'analyse spectrale? Nous venons de voir que la structure d'interface opère un groupement temporel particulier des manifestations énergétiques à travers le temps, groupement temporel qui représente l'invariance décrétée par cette structure. *Nous croyons que le processus de caractérisation des structures d'analyse spectrale est similaire: il vise à décréter de nouveaux invariants sonores dans 'ce qui se manifeste' à travers le temps mais en opérant, cette fois, plusieurs groupements temporels différents sur la position de la structure d'interface, groupements temporels qui donnent lieu à plusieurs entités sonores de second niveau.*

A chaque instant donc, l'entité sonore de premier niveau représente la quantité énergétique qui se manifeste à cet instant. Nous avons vu que la production de cette entité était rendue possible grâce aux propriétés physiques particulières de la structure d'interface. Ces propriétés physiques

permettent un groupement temporel particulier de ce qui se manifeste à travers le temps de telle manière que toutes les variations de la quantité énergétique sont "incorporées" dans le déplacement de la structure d'interface, ce qui permet justement la connaissance de la quantité énergétique totale à chaque instant. Mais si ce groupement temporel nous permet de connaître la quantité énergétique qui est le résultat de toutes ces variations, il ne nous permet pas, par contre, de connaître la nature exacte de ces variations. Comment pourrions-nous caractériser la séquence de positions successives qui "conduisent" à la position de la structure d'interface de chaque instant? Quelle est "l'histoire" variationnelle de l'entité sonore de premier niveau à chaque instant?

Nous croyons que le processus de caractérisation de la structure d'analyse spectrale permet la connaissance de cette histoire variationnelle en opérant une série de groupements temporels particuliers. Chacun de ces groupements reflète une "sensibilité" particulière à une certaine "forme" de variation de la quantité énergétique et produit la "connaissance" de cette forme, à chaque instant, en générant des valeurs successives sur de nouveaux attributs de caractérisation. La position de la structure d'interface est, à chaque instant, le résultat de *toutes* les formes de variation de la quantité énergétique, alors que chaque valeur spectrale (parmi l'ensemble de toutes les valeurs spectrales de chaque instant) n'est le résultat que d'*une seule forme* de variation. Mais qu'est-ce qu'une "forme" de variation?

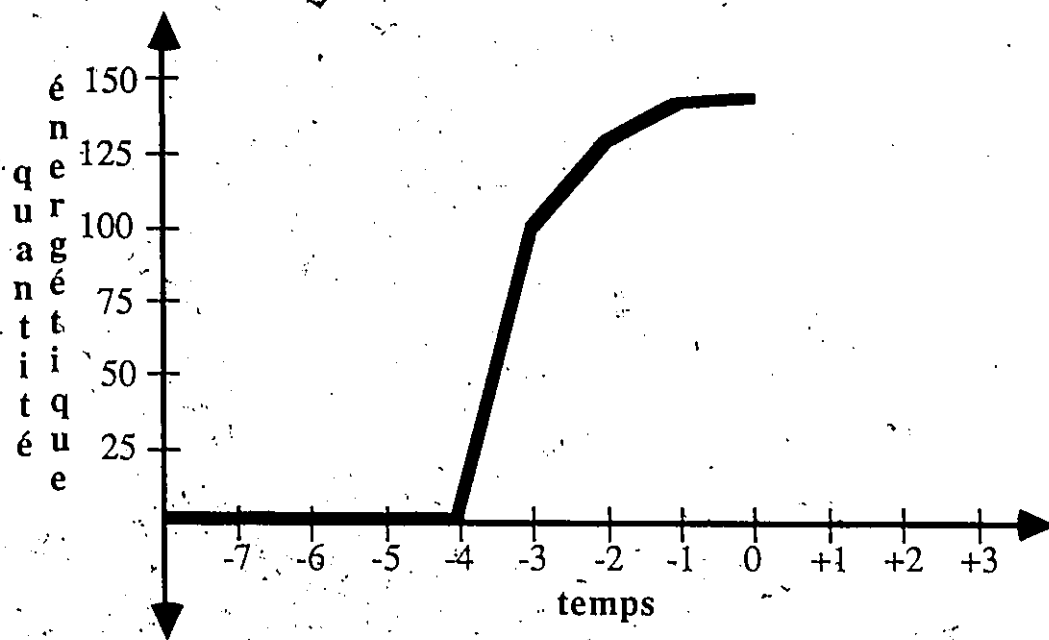


Fig. 4 Une séquence de positions de la structure d'interface

Considérons à nouveau l'exemple de la figure 4. A chaque instant, "l'histoire" variationnelle de la position de la structure d'interface se modifie. L'histoire variationnelle de cette position à l'instant -4 est certainement fort différente de l'histoire variationnelle de cette position à l'instant -3. Comment pourrions-nous décrire l'histoire variationnelle de chacune de ces positions, comment pourrions-nous, en d'autres termes, décrire la variation de positions "passées", qui est caractéristique de chaque position "présente" de la structure d'interface? Imaginons une série d'intervalles temporels qui "s'étendent" vers le passé de chaque position de la structure d'interface. Ces intervalles temporels peuvent être de différentes longueurs; ils peuvent s'étendre sur une fraction de moment d'échantillonnage comme ils peuvent s'étendre sur l'infinité du passé de chaque instant.(Fig. 5).

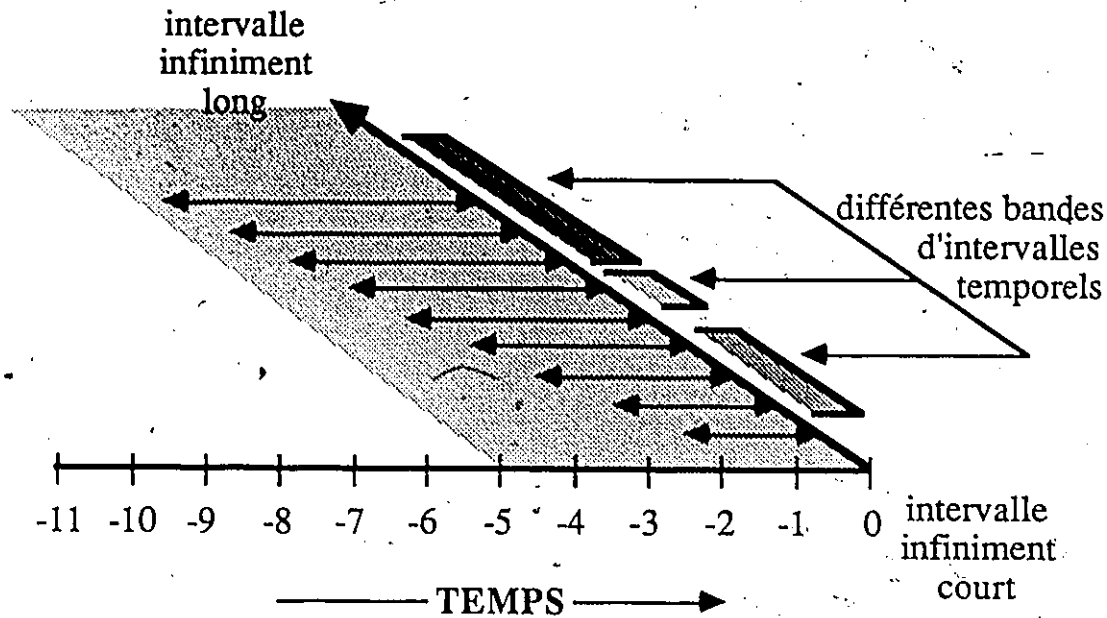


Fig. 5 Différents intervalles temporels et différentes bandes d'intervalles temporels

Imaginons de plus que tous ces intervalles temporels sont "collés" à l'instant d'échantillonnage *présent* et qu'ils se déplacent avec cet instant à chaque "avancée" de cet instant présent dans la dimension temporelle. L'instant présent "avance" à travers la dimension temporelle un peu comme une locomotive qui avance sur une voie ferrée en traînant derrière elle un train de longueur infinie. Chaque intervalle temporel correspond à une certaine "longueur" du train (exprimée en nombre de wagons et/ou de fractions de wagons). *Nous définirons une forme variationnelle comme étant une forme de variation spécifique à un intervalle temporel ou à une "bande" d'intervalles temporels.* Une "bande" d'intervalles temporels regroupe un

certain nombre d'intervalles temporels juxtaposés. De telles bandes peuvent se constituer, par exemple, de tous les intervalles temporels qui sont compris entre l'intervalle temporel infiniment court (la plus petite distance "derrière" le train) et une fraction de moment d'échantillonnage (1/10 de moment, par exemple)¹, ou encore de tous les intervalles temporels qui sont compris entre l'intervalle temporel d'un (1) moment d'échantillonnage et l'intervalle temporel de quinze (15) moments d'échantillonnage, ou encore de tous les intervalles temporels qui sont compris entre l'intervalle temporel de trois (3) moments d'échantillonnage et l'intervalle temporel d'une infinité de moments d'échantillonnage. A chaque forme variationnelle est associée un intervalle temporel ou une certaine bande d'intervalles temporels et, à chaque instant présent, l'histoire variationnelle de la position de la structure d'interface se caractérise par une mesure de toutes ces formes variationnelles *spécifiques* à ces intervalles temporels ou à ces bandes d'intervalles temporels. Cette mesure s'exprime par deux attributs de caractérisation: une *quantité de variation* et une *direction de variation*.

Considérons, par exemple, la séquence de positions de la figure 4. Nous avons supposé que la variation était nulle de moins l'infini temporel jusqu'à l'instant -4. L'histoire variationnelle de la position de la structure d'interface, à l'instant -4, devrait donc se caractériser par une série de valeurs nulles pour toutes les formes de variation. A l'instant suivant (instant -3), la position de la

¹ Cela voudrait dire que cette bande d'intervalles temporels se constituerait de cet intervalle temporel infiniment court *et* de l'intervalle temporel "un peu" plus long *et* de l'intervalle temporel "un peu" plus long *et* etc. jusqu'à l'intervalle temporel d'un dixième (1/10) de moment. Evidemment, il existe une infinité d'intervalles temporels pour chaque bande puisque nous pouvons diviser cette bande en une infinité de manières.

structure d'interface est à +100; l'histoire variationnelle de la position de la structure d'interface vient de se modifier et cela devrait se refléter dans les valeurs d'attribut des différentes formes de variation. En première approximation, nous pourrions dire que la variation, qui s'est produite jusqu'à l'instant -3, *est beaucoup plus caractéristique d'une variation sur des intervalles temporels courts, qu'elle n'est caractéristique d'une variation sur des intervalles longs*. En termes de formes de variation, cela revient à dire que cette variation se caractérise beaucoup plus par les formes de variation qui sont spécifiques à des intervalles ou à des bandes d'intervalles temporels courts, qu'elle ne se caractérise par les formes de variation qui sont spécifiques à des intervalles ou à des bandes d'intervalles temporels longs. Finalement, en termes de valeurs d'attribut de ces formes de variation, cela veut dire que la *quantité* de variation des formes de variation spécifiques aux bandes d'intervalles temporels courts devrait être plus importante que la quantité de variation des formes de variation spécifiques aux bandes d'intervalles temporels longs. Quant à la direction de variation de *toutes* les formes de variation, elle devrait être de même direction², *positive (+)*,

² Il existe deux directions possibles de variation. Nous désignerons une variation de la quantité énergétique qui va dans la "direction" d'un accroissement de la quantité énergétique comme une direction de variation positive et nous dénoterons cette direction par le signe + entre parenthèses (+). Nous désignerons une variation de la quantité énergétique qui va dans la "direction" d'une diminution de la quantité énergétique comme une direction de variation négative et nous dénoterons cette direction par le signe - entre parenthèses (-). Nous mettons les signes de variation entre parenthèses pour distinguer les mesures de *variation de la quantité énergétique* (structure d'analyse spectrale) des mesures de *quantité énergétique* (structure d'interface).

puisque l'on peut considérer que la quantité énergétique a augmenté sur tous les intervalles temporels imaginables.

A l'instant suivant (instant -2), la position de la structure d'interface est à +130. Nous pourrions dire que la variation, qui s'est produite jusqu'à l'instant -2, est un peu moins caractéristique d'une variation sur des bandes d'intervalles courts et un peu plus caractéristique d'une variation sur des bandes d'intervalles longs. En termes de formes de variation, cela signifie que l'histoire variationnelle de cet instant se caractérise un peu moins par les formes de variation spécifiques aux bandes d'intervalles temporels courts et qu'elle se caractérise un peu plus par les formes de variation spécifiques aux bandes d'intervalles longs. En termes d'attributs de caractérisation maintenant, cela revient à dire que la quantité de variation devrait diminuer chez les formes de variation spécifiques aux bandes d'intervalles temporels courts, alors que cette quantité de variation devrait augmenter chez les formes de variation spécifiques aux bandes d'intervalles temporels longs. Quant à la direction de variation de toutes ces formes variationnelles, elle reste toujours positive (+). La même "dynamique" variationnelle devrait se poursuivre durant les deux moments d'échantillonnage suivants. La quantité de variation des formes variationnelles spécifiques aux bandes d'intervalles temporels courts devrait continuer à diminuer, alors que la quantité de variation des formes variationnelles spécifiques aux bandes d'intervalles temporels longs devrait continuer à augmenter, la direction de variation demeurant toujours positive.

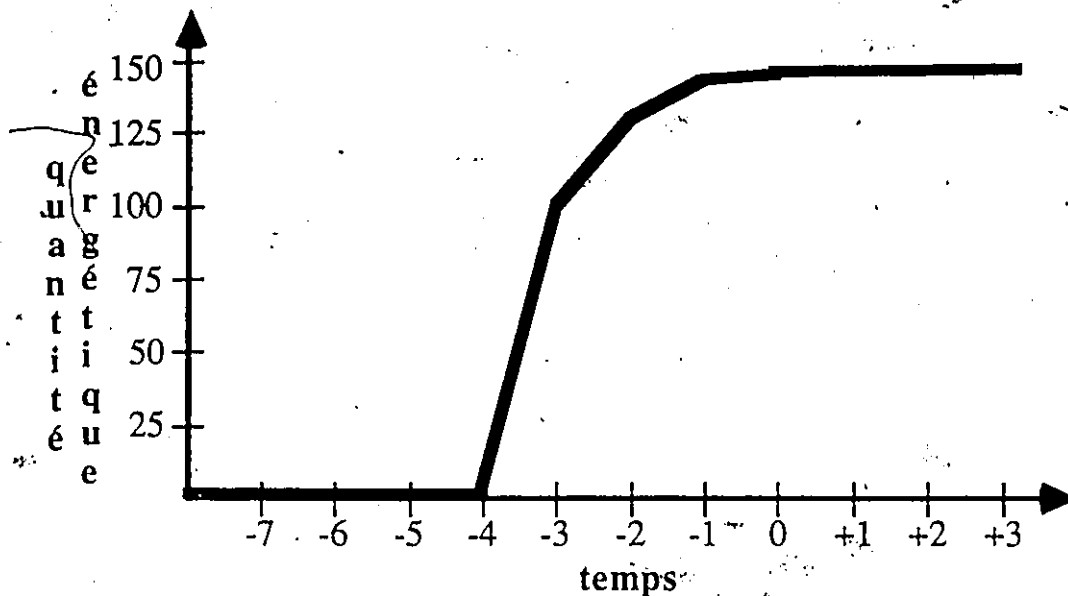


Fig. 6 Une séquence de positions qui se poursuit selon le même pattern variationnel que la séquence de la figure 4

Supposons maintenant que les valeurs futures de la position de la structure d'interface suivent sensiblement le même "pattern" variationnel durant une infinité d'instant d'échantillonnage futurs (Fig.6). Ces valeurs pourraient être, par exemple, de +143 à l'instant +1, de +141 à l'instant +2, de +140.5 à l'instant +3 etc.. Concentrons-nous sur une seule forme variationnelle que nous supposons spécifique à une bande d'intervalles temporels "moyens", et suivons le comportement de cette forme à travers le temps, de l'instant -4 jusqu'aux instants "futurs". A l'instant -4, la quantité de variation de cette forme est de 0. A l'instant -3, cette quantité de variation augmente et cette quantité de variation peut continuer à augmenter à l'instant -2, à l'instant -1 et durant un certain nombre d'instant, *mais il arrivera un instant où cette quantité de variation se remettra à diminuer pour se rapprocher progressivement de 0*. L'histoire variationnelle se définira pendant un certain

temps par une augmentation de la quantité de variation de cette forme variationnelle, mais il arrivera un instant où cette forme de variation atteindra un maximum de "contribution" à l'histoire variationnelle des positions de la structure d'interface. A l'instant suivant de ce maximum, la contribution (quantité de variation) de cette forme de variation se remettra à diminuer. Cela est le cas de toutes les formes variationnelles; ce qui les distingue, c'est l'instant précis où elles atteignent ce maximum de contribution. Cet instant est fonction de leur bande d'intervalles temporels. Plus cette bande se constitue d'intervalles temporels courts, plus elles atteignent ce maximum rapidement. C'est la raison pour laquelle nous disions plus haut que la quantité de variation des formes de variation spécifiques aux bandes d'intervalles temporels courts se remettait à diminuer dès l'instant -2. Ces formes variationnelles atteignent leur maximum à l'instant -3 et se remettent immédiatement à diminuer, alors que les formes variationnelles spécifiques aux bandes d'intervalles temporels plus longs atteignent leur maximum plus tard pour se remettre également à diminuer par la suite.

Supposons maintenant la même séquence initiale de positions de la structure d'interface jusqu'à l'instant -3: une absence de variation depuis moins l'infini temporel jusqu'à l'instant -4, et une position de +100 à l'instant -3. Mais contrairement à l'exemple de la figure 4, imaginons maintenant que la position de la structure d'interface reste à +100 pour une infinité d'instant d'échantillonnage (instant -2, -1, 0, +1, +2 etc.) (Fig. 7).

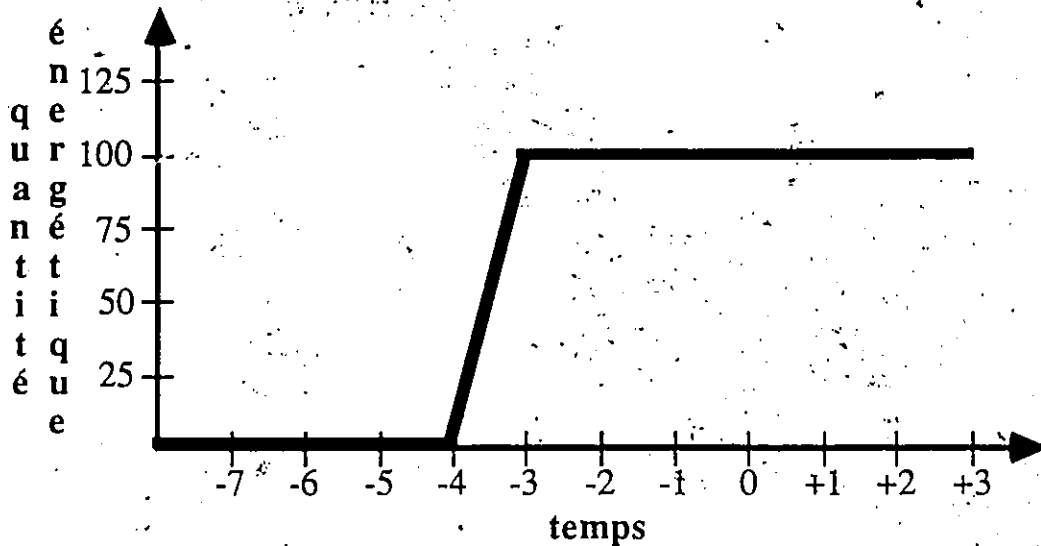


Fig. 7 Une séquence de positions de la structure d'interface qui reste à la position +100 pour une infinité de moments

Notre description initiale de l'histoire variationnelle de cette structure reste la même: elle se caractérise, à l'instant -4, par une quantité de variation nulle pour toutes les formes variationnelles, et, à l'instant -3, par une quantité de variation plus importante pour les formes variationnelles spécifiques à des intervalles courts que pour les formes variationnelles spécifiques à des intervalles longs. Que devient cette histoire variationnelle à l'instant -2? Que devient, plus particulièrement, la quantité de variation des formes variationnelles spécifiques à des intervalles temporels *plus courts* que le moment d'échantillonnage? On pourrait penser que cette quantité de variation devient nulle puisqu'il n'y a pas de variation durant le dernier moment d'échantillonnage (la position est à +100 à l'instant -3 et à l'instant -2). Cela est vrai pour ce moment d'échantillonnage particulier, mais

pouvons-nous dire que l'histoire variationnelle de la position de la structure d'interface, à l'instant -2, est complètement dénuée (quantité de variation = 0) de cette forme variationnelle spécifique à un intervalle temporel plus court que le moment d'échantillonnage? Non, puisque nous savons que cette forme variationnelle a participé (par une quantité et une direction de variation) à l'histoire variationnelle de la position de l'instant précédent et que l'histoire variationnelle de cette position précédente (celle de l'instant -3) fait, en quelque sorte, partie de l'histoire variationnelle de la position à l'instant -2. L'histoire variationnelle de la position de l'instant -2 se "constitue" de l'histoire variationnelle de l'instant-3 *et* d'un ajustement pour ce qui s'est produit entre l'instant -3 et l'instant -2. L'absence de variation entre l'instant -3 et l'instant -2 fait certainement diminuer la quantité de variation de cette forme variationnelle spécifique à un intervalle temporel plus court qu'un moment d'échantillonnage, puisqu'il n'y a pas de variation sur cet intervalle temporel pour ce moment d'échantillonnage. Mais *il y a eu* de la variation spécifique à cet intervalle temporel dans le passé et c'est ce qui nous suggère que la forme variationnelle spécifique à cet intervalle temporel se caractérise toujours à l'instant -2 par une certaine quantité de variation. A chaque instant donc, le calcul des valeurs d'attributs d'une forme variationnelle doit se faire en tenant compte de:

- la présence de cette forme variationnelle jusqu'à l'instant précédent,
et,
- de la présence de cette forme variationnelle durant le dernier moment d'échantillonnage.

Nous verrons plus loin (sous-section 3.2.1) un algorithme qui nous permet de calculer précisément les valeurs d'attributs de ces différentes formes

variationnelles. Cet algorithme se compose de deux calculs qui nous permettent d'évaluer les deux facteurs que nous venons d'énoncer. Pour l'instant, essayons de nous familiariser un peu plus avec la notion de forme variationnelle en considérant cette notion d'un point de vue différent, mais complémentaire, du point de vue 'intervalle temporel': le point de vue 'vitesse de variation'.

On pourrait dire que la variation de la quantité énergétique de l'exemple de la figure 4 se caractérise globalement par une variation initiale "rapide" (entre l'instant -4 et l'instant -3) et par un ralentissement subséquent de cette variation rapide. Le concept de 'forme variationnelle' entretient un rapport étroit avec le concept de 'vitesse'. Cela est compréhensible si nous considérons qu'une variation "*rapide*" se caractérise par une certaine quantité de variation sur une bande constituée d'intervalles temporels courts, alors qu'une variation "*moins rapide*" (mais de la même importance³) se caractérise par une certaine quantité de variation sur une bande constituée d'intervalles temporels plus longs. Nous pouvons utiliser le concept de 'vitesse' pour essayer de mieux comprendre le concept de 'forme variationnelle'. Une vitesse peut être nulle comme elle peut être infiniment élevée. Représentons-nous ces différentes *valeurs* de vitesse sur un axe de valeurs positives dont l'origine (la valeur 0) représente la valeur de vitesse nulle et dont les valeurs positives de plus en plus grandes représentent des

³ On peut associer "l'importance" de la variation au déplacement de la position de la structure d'interface. Une variation de 0 à +100 en un (1) moment d'échantillonnage est aussi importante qu'une variation de 0 à +100 en quinze (15) moments d'échantillonnage. Ce qui distingue ces deux variations, c'est l'intervalle temporel sur lequel elles se produisent.

valeurs de vitesses de plus en plus élevées, le point situé à $+\infty$ représentant une vitesse infiniment élevée.(Fig. 8)

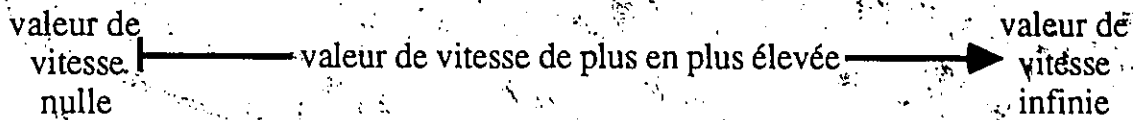


Fig. 8 L'axe des valeurs de vitesse

Nous pouvons diviser cet axe en plusieurs *bandes* de vitesses. Imaginons, par exemple, que nous divisons l'axe en *deux* bandes de vitesses en déterminant une valeur de vitesse *intermédiaire* qui sépare ces deux bandes de vitesses. La bande de valeurs de vitesse qui se situe à droite de cette valeur de vitesse regroupe l'ensemble des valeurs de vitesse qui vont de cette valeur de vitesse intermédiaire jusqu'à la valeur de vitesse infiniment élevée (la bande 'élevée' de valeurs de vitesse), alors que la bande de valeurs de vitesse qui se situe à gauche de cette vitesse intermédiaire regroupe l'ensemble des valeurs de vitesse qui vont de la valeur de vitesse nulle jusqu'à cette valeur de vitesse intermédiaire (la bande 'faible' de valeurs de vitesse)(Fig.9).

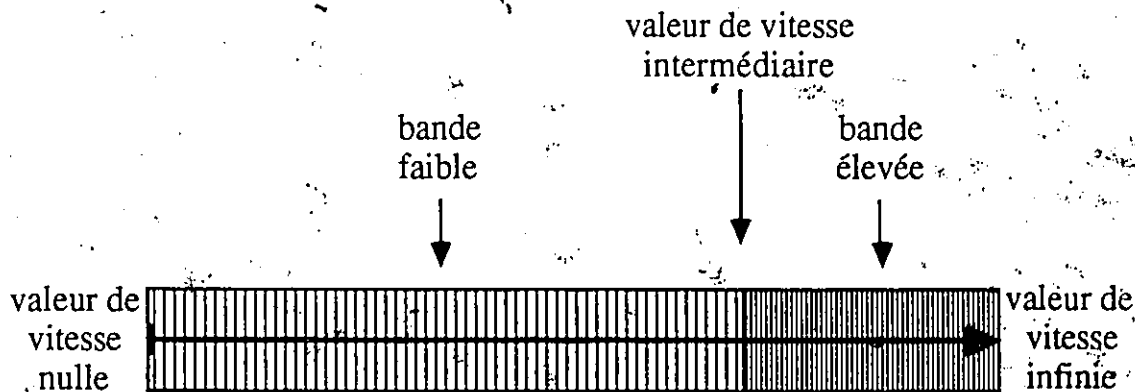


Fig. 9 Les bandes 'faible' et 'élevée' de valeurs de vitesse

Retournons maintenant aux bandes d'intervalles temporels et supposons maintenant que nous sommes à un instant précis du temps et que nous calculons la quantité de variation de *deux* formes de variation. Nous calculons tout d'abord la quantité de variation pour une forme variationnelle spécifique à une bande d'intervalles temporels qui va de l'intervalle temporel infiniment court jusqu'à un intervalle temporel *intermédiaire* d'une certaine longueur (la bande 'courte' d'intervalles temporels). Nous calculons ensuite la quantité de variation d'une forme variationnelle spécifique à une bande d'intervalles temporels qui va de cet intervalle temporel intermédiaire à un intervalle temporel infiniment long (la bande 'longue' d'intervalles temporels) (Fig.10).

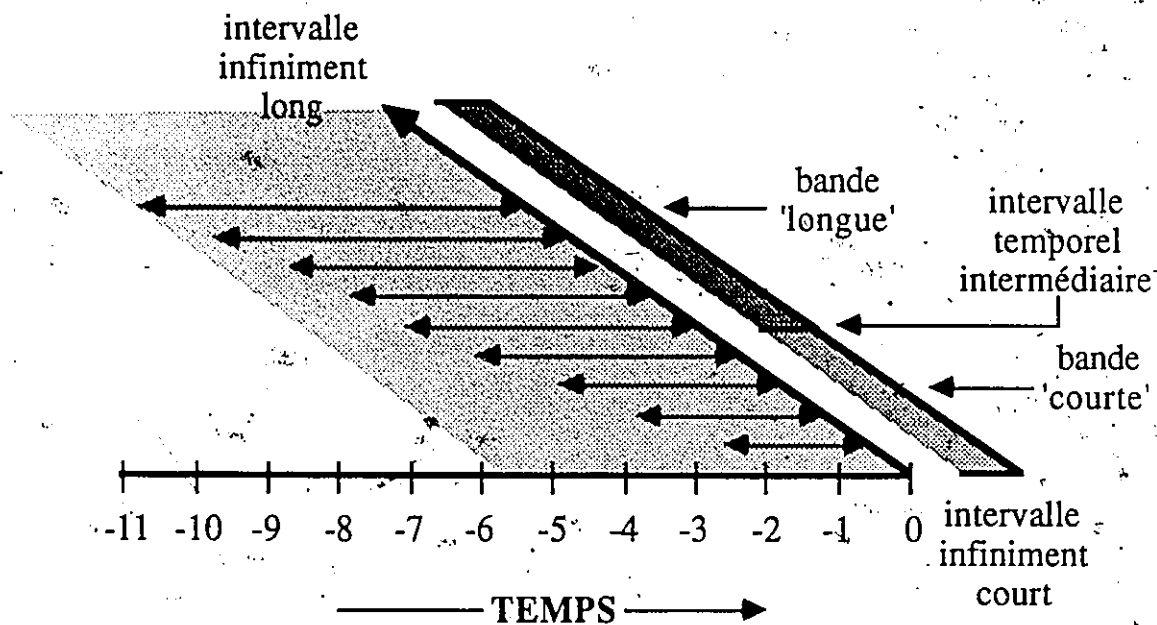


Fig. 10 Les bandes 'courte' et 'longue' d'intervalles temporels

La quantité de variation de chacune de ces deux formes variationnelles peut être associée à la contribution plus ou moins importante des deux bandes de vitesses dans l'histoire variationnelle de la position de la structure d'interface à cet instant. Nous pouvons associer la quantité de variation de la forme variationnelle 'longue' à la contribution de la bande de valeurs de vitesse 'faible', alors que la quantité de variation de la forme variationnelle 'courte' peut être associée à la contribution de la bande de valeurs de vitesse 'élevée'.

Il peut sembler inhabituel de qualifier l'histoire variationnelle de chaque position de la structure d'interface par la contribution de *plusieurs* bandes de vitesses. Nous utilisons habituellement le concept de vitesse pour quantifier le déplacement d'un objet, et cet objet a normalement une seule vitesse. Nous dirons d'une voiture qu'elle se déplace à 100 km à l'heure ou qu'une balle de fusil voyage à 300 mètres à la seconde. Mais une vitesse est une mesure de la quantité de variation sur un certain intervalle temporel de référence. Il devient plus acceptable de qualifier la variation à partir de *plusieurs* bandes de vitesses, lorsque nous considérons le fait que les formes variationnelles représentent une mesure de la variation spécifique à *plusieurs* bandes d'intervalles temporels. Dans les exemples de la voiture et de la balle, nous ne mesurons la vitesse que sur *un seul* intervalle temporel (l'heure ou la seconde); mais nous pourrions mesurer plusieurs vitesses, à chaque instant, en mesurant la quantité de variation (des formes variationnelles) de la position de la balle (ou de la voiture) sur *plusieurs* intervalles temporels. La quantité de variation d'une forme variationnelle peut être nulle ou plus ou moins importante. Lorsque nous associons une bande de vitesses à cette forme variationnelle, c'est la contribution de cette bande de vitesses (dans

l'histoire variationnelle de chaque position de la structure d'interface) que nous évaluons. Si la quantité de variation de cette forme variationnelle est faible, par exemple, cela signifie que les valeurs de vitesses de cette bande de vitesses ne se retrouvent presque pas dans l'histoire variationnelle (elles n'y contribuent à peu près pas). Si la quantité de variation de cette forme variationnelle est élevée, cela signifie que les valeurs de vitesses de cette bande de vitesses se retrouvent beaucoup plus dans cette histoire variationnelle. La contribution plus ou moins importante des valeurs de vitesse de cette bande de vitesses est toujours fonction de la quantité de variation de la forme variationnelle à laquelle est associée cette bande de vitesses.

Dans cette optique, nous pouvons concevoir la structure d'analyse spectrale comme une structure composée *d'éléments structuraux qui réagissent de manière distincte aux différentes valeurs de vitesse de la variation de la quantité énergétique*. Chacun de ces éléments est sensible à une bande particulière de valeurs de vitesse. Lorsque la vitesse de la variation est élevée, ce sont les éléments sensibles aux vitesses élevées qui réagissent maximalement; lorsque la vitesse de la variation est plus lente, ce sont les éléments sensibles à ces vitesses plus lentes qui réagissent maximalement. Nous pouvons assimiler concrètement ces éléments aux différentes portions de la membrane basilaire de la cochlée ou aux différents filtres électroniques d'un analyseur spectral. Ce sont les propriétés physiques particulières de ces différents éléments structuraux qui leur permettent, comme dans le cas de la

structure d'interface, d'effectuer les groupements temporels particuliers qui produisent la connaissance des différentes formes de variation⁴.

La structure d'analyse spectrale participe donc au processus de connaissance sonore en générant une nouvelle famille d'entités sonores: les formes variationnelles. Ces entités sonores sont l'expression du décret de nouvelles relations invariantes dans 'ce qui se manifeste' à travers le temps, et ces relations invariantes sont déterminées par les groupements temporels particuliers (des positions de la structure d'interface) qu'effectuent les différents éléments structuraux de la structure d'analyse spectrale. Ces nouvelles entités sonores sont dotées de trois attributs de caractérisation:

- la bande d'intervalles temporels de chaque forme variationnelle ou la bande de vitesses à laquelle est sensible cette forme variationnelle,
- la quantité de variation de chaque forme variationnelle et,
- la direction de variation de chaque forme variationnelle.

Nous avons tenté de donner un aperçu intuitif du processus de caractérisation du second niveau de connaissance sonore; nous allons maintenant essayer de lui donner un peu plus de consistance formelle en développant un algorithme qui nous permet de calculer précisément les valeurs d'attribut des différentes formes variationnelles à chaque instant d'échantillonnage.

⁴ Rappelons, à ce sujet, que la rigidité (stiffness) et la dimension de la membrane basilaire changent sur toute sa longueur.

3.3.1 Le calcul des valeurs d'attribut des formes variationnelles

Les différents éléments structuraux de la structure d'analyse spectrale effectuent donc des groupements temporels particuliers qui leur permettent de caractériser les différentes formes variationnelles qui sont présentes dans l'histoire variationnelle de chaque position de la structure d'interface. Il est possible d'exprimer ces groupements temporels à partir du même type d'équation que l'équation du groupement temporel de la structure d'interface. Rappelons cette équation:

$$PSI_0 = (PSI_{-1} * 1) + ((PSI_0 - PSI_{-1}) * 1)$$

La position de la structure d'interface à l'instant 0 est égale à la position de la structure d'interface à l'instant -1 multipliée par une constante temporelle de 1 plus la variation sur un moment d'échantillonnage multipliée également par une constante temporelle de 1. Le premier terme de droite de l'équation peut être considéré comme la somme de toutes les variations antécédentes sur un moment d'échantillonnage, i.e. comme la somme de toutes les variations passées sur un moment d'échantillonnage (variations passées qui ont été aussi prises dans leur totalité grâce à la constante temporelle de 1). Le deuxième terme de droite ajoute, à cette sommation récursive jusqu'à l'instant précédent, la totalité (constante temporelle de 1) de la variation du moment qui vient de s'écouler. C'est ce groupement temporel particulier qui permet à la structure d'interface d'afficher la totalité de la quantité énergétique à chaque instant. C'est également ce groupement temporel particulier qui intègre toutes les formes variationnelles qui constituent la variation de la

quantité énergétique. Le facteur déterminant de ce groupement réside dans la valeur de la constante temporelle des deux termes de droite de l'équation. *C'est en "jouant" avec cette valeur de constante temporelle que nous allons pouvoir générer de nouveaux types de groupements temporels qui reflèteront, eux, la contribution quantitative et directionnelle des différentes formes variationnelles.*

Que se produit-il si nous modifions la valeur de cette constante temporelle? Calculons, par exemple, les valeurs d'un groupement temporel où nous changeons la valeur de la constante temporelle à .1 dans les deux termes de droite. Pour l'instant -3 de l'exemple de la figure 4, cela donne:

$$FV_{(.1, -3)} = ((FV_{(.1, -4)}) * .1) + ((PSI_{-3} - PSI_{-4}) * .1)$$

$$FV_{(.1, -3)} = (0 * .1) + ((100 - 0) * .1)$$

$$FV_{(.1, -3)} = 0 + 10$$

$$FV_{(.1, -3)} = (+)10$$

Nous avons changé le nom de la variable principale pour indiquer que nous calculons maintenant la valeur de la contribution d'une forme variationnelle (FV), et nous avons ajouté, en indice de ce nouveau nom de variable, l'indication '1' ($FV_{(.1, -3)}$, $FV_{(.1, -4)}$) pour montrer qu'il s'agit ici de la contribution de la forme variationnelle '1'. L'autre indice de ces expressions ($FV_{(.1, -3)}$, $FV_{(.1, -4)}$) et les indices de l'expression ($PSI_{-3} - PSI_{-4}$) représentent toujours la valeur temporelle de l'instant d'échantillonnage. Nous calculons maintenant les valeurs d'une forme variationnelle. Nous avons indiqué, à la section précédente, que nous dénoterions la direction de

variation par les signes + et - entre parenthèses: (+) et (-). Il faudrait donc indiquer toutes les valeurs du premier terme de droite⁵ en utilisant les parenthèses puisqu'il s'agit d'une valeur de variation (la valeur de la forme variationnelle '.1' à l'instant précédent). Il faudrait également indiquer entre parenthèses la valeur du deuxième terme de droite de la troisième étape du calcul (10, dans ce cas-ci), puisqu'il s'agit également, à cette étape, d'une valeur de variation, et non plus de valeurs de quantités énergétiques comme à la deuxième étape. Pour ne pas surcharger l'équation de parenthèses, nous avons choisi de n'indiquer que le résultat final⁶ sous sa forme variationnelle ((+)10).

Si nous lisons l'expression, cela donne: la valeur de la forme variationnelle '.1' à l'instant d'échantillonnage -3 est égale à la valeur de cette forme à l'instant -4 *multipliée* par la constante temporelle de valeur '.1'(premier terme de droite), *plus* la variation totale sur le moment d'échantillonnage qui va de l'instant -4 à l'instant -3 *multipliée* également par une constante temporelle de valeur '.1'(deuxième terme de droite). Le résultat de ce calcul est égal à (+)10 à l'instant -3. Nous pouvons interpréter ce résultat en disant que l'histoire variationnelle de la position de la structure d'interface, à l'instant -3, se caractérise par la présence d'une forme variationnelle particulière (la forme '.1'), présence qui s'exprime par une quantité de variation (10) et par une direction de variation (+). La forme variationnelle qui est spécifique à une certaine bande d'intervalles temporels contribue à caractériser cette histoire selon les valeurs calculées (valeur d'attribut de

⁵ La valeur 0, dans ce cas-ci. La valeur 0 est un cas spécial puisque la direction de variation est inexistante. Elle n'est ni négative (-) ni positive (+).

⁶ Tous ces résultats sont arrondis à la première décimale.

quantité et valeur d'attribut de direction). Que représentent plus précisément ces deux valeurs d'attributs? C'est ce que nous verrons un peu plus loin. Nous verrons également comment il est possible d'interpréter ces valeurs en termes de contribution plus ou moins importante de bandes de vitesses. "Pour l'instant", calculons les valeurs de la forme variationnelle '1' pour les instants -2, -1, et 0. Pour l'instant -2:

$$FV_{(.1, -2)} = ((FV_{(.1, -3)}) * .1) + ((PSI_{-2} - PSI_{-3}) * .1)$$

$$FV_{(.1, -2)} = (10 * .1) + ((130 - 100) * .1)$$

$$FV_{(.1, -2)} = 1 + 3$$

$$FV_{(.1, -2)} = (+)4$$

Pour l'instant -1:

$$FV_{(.1, -1)} = ((FV_{(.1, -2)}) * .1) + ((PSI_{-1} - PSI_{-2}) * .1)$$

$$FV_{(.1, -1)} = (4 * .1) + ((140 - 130) * .1)$$

$$FV_{(.1, -1)} = .4 + 1$$

$$FV_{(.1, -1)} = (+)1.4$$

Et pour l'instant 0:

$$FV_{(.1, 0)} = ((FV_{(.1, -1)}) * .1) + ((PSI_0 - PSI_{-1}) * .1)$$

$$FV_{(.1, 0)} = (1.4 * .1) + ((145 - 140) * .1)$$

$$FV_{(.1, 0)} = .14 + .5$$

$$FV (.1, 0) = (+).64$$

L'équation de la forme variationnelle '.1' et l'équation de la structure d'interface sont très similaires, *la seule différence résidant dans la valeur de la constante temporelle*, cette valeur étant de 1 dans le cas de la structure d'interface et de .1 dans le cas de la forme variationnelle '.1'. On peut assimiler la structure d'interface et l'élément structural sous-jacent au calcul de la forme variationnelle '.1' à deux réservoirs. Le réservoir de la structure d'interface reste au même niveau à travers un moment d'échantillonnage (multiplication par 1) et ajoute ou soustrait, à ce niveau, la variation totale qui se produit durant un moment d'échantillonnage.

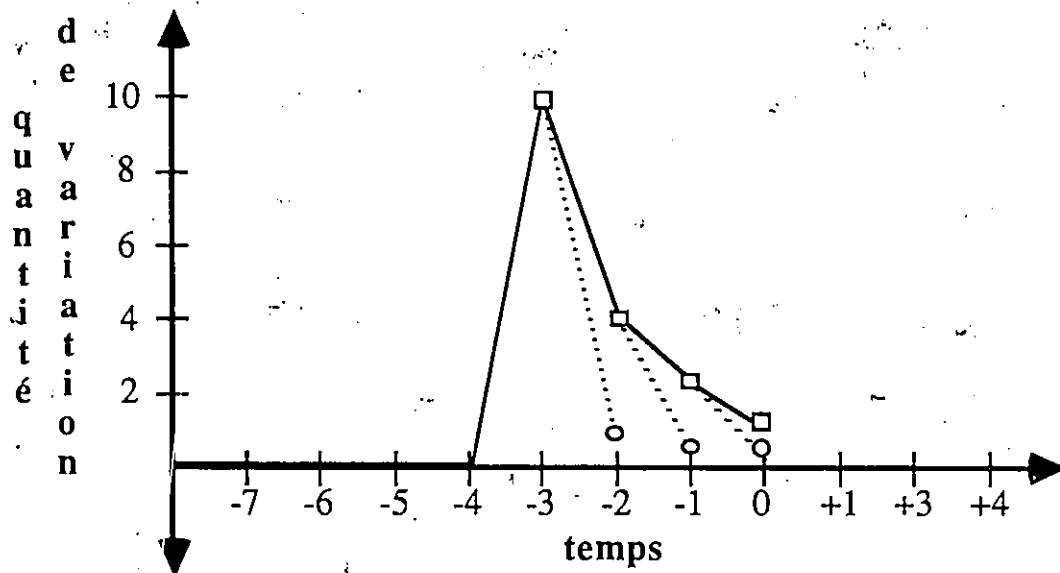


Fig. 11 Le comportement du réservoir de l'élément structural sous-jacent à la forme variationnelle '.1' pour la séquence de positions de la figure 4

Le réservoir de l'élément structural de la forme variationnelle '.1' a tendance, lui, à se rapprocher du niveau 0 à travers un moment d'échantillonnage (multiplication par .1) et ajoute ou soustrait à cette tendance un dixième (1/10) de la variation totale qui se produit durant un moment d'échantillonnage. La figure 11 illustre ce comportement du réservoir de la forme variationnelle '.1' pour l'exemple de la figure 4. Les symboles carrés représentent les valeurs réelles (calculées) de cette forme à chaque instant d'échantillonnage, valeurs réelles que nous avons reliées par des lignes pleines. Les symboles ronds et les lignes pointillées représentent ce qu'auraient été les valeurs de la forme variationnelle '.1' si la variation avait été nulle⁷ durant le moment d'échantillonnage *suivant* chaque calcul de la valeur réelle. Lorsque la constante est égale à 1 (dans le cas de la structure d'interface), le niveau du réservoir change en fonction de la variation totale sur un moment; lorsque la constante est égale à '.1', le niveau change en fonction de cette variation totale multipliée par la constante '.1' *et* en fonction de cette tendance à se rapprocher du niveau 0 du réservoir. Ainsi, le réservoir de la forme variationnelle '.1' se remplit jusqu'à la marque '(+)10' entre l'instant -4 et l'instant -3 ($100 * .1$). Si la variation était nulle (si la valeur de la position de la structure d'interface restait à 100 à l'instant -2), le réservoir se viderait jusqu'à la marque '(+)1', mais la variation existante ($100 \rightarrow 130$) ajoute une valeur de 3 ($30 * .1$) et le réservoir est à la marque '(+)4' ($3 + 1$) à l'instant -2. A l'instant -1, le réservoir de la forme variationnelle '.1' se serait vidé jusqu'à la marque '(+)4', mais la légère

⁷ Et non pas si la variation avait été nulle à partir de l'instant -3 jusqu'à l'instant 0 ($100 \rightarrow 100 \rightarrow 100$). Les symboles ronds et les lignes pointillées représentent en fait le calcul du premier terme de droite dans l'équation.

variation (130 \rightarrow 140) ajuste le niveau du réservoir à la marque '(+)1.4' ($.4 + (10 * .1)$). Finalement, à l'instant 0, le réservoir se serait vidé jusqu'à la marque '(+).14' mais la variation existante (140 \rightarrow 145) ajuste ce niveau à la marque '(+).64' ($.14 + (5 * .1)$).

Imaginons maintenant la même séquence de positions de la structure d'interface mais cette fois-ci, dans une direction de variation négative. La position a été de 0 depuis moins l'infini temporel jusqu'à l'instant -4. A cet instant, la quantité énergétique a commencé à varier mais dans la direction d'une *diminution* de la quantité énergétique (Fig. 12).

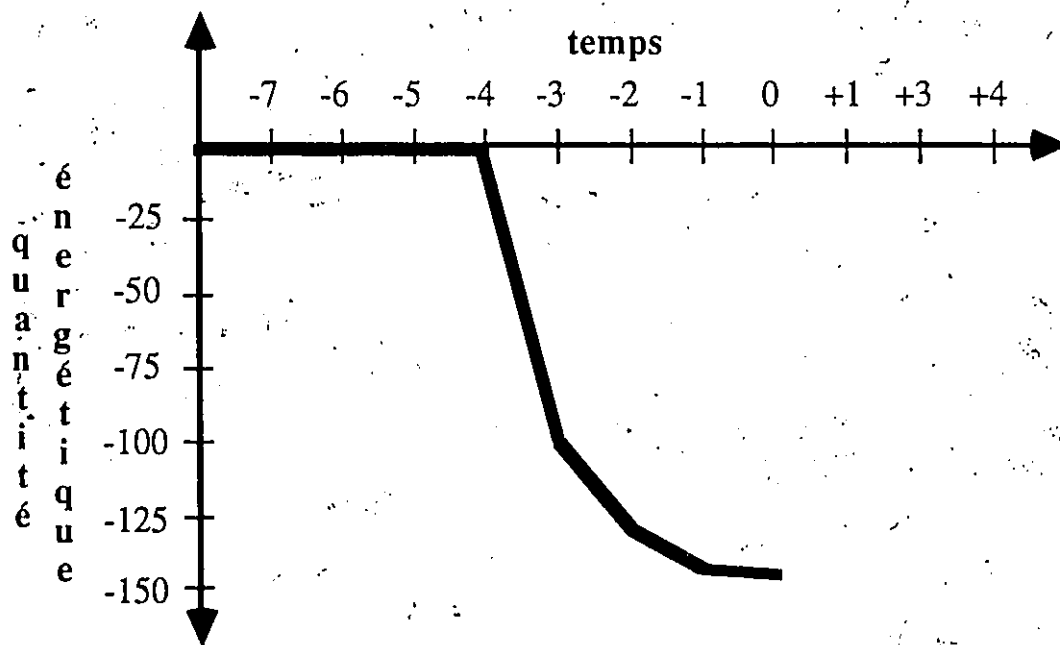


Fig. 12 Le même pattern de variation que celui de la figure 4 mais dans la direction d'une diminution de la quantité énergétique

La position de la structure d'interface est de -100 à l'instant -3, de -130 à l'instant -2, de -140 à l'instant -1 et de -145 à l'instant 0. Quelles sont les valeurs de la forme variationnelle '.1'? Elles sont exactement les mêmes que pour la séquence de variation positive (fig. 4); à l'exception de la direction de variation qui est maintenant négative. Les calculs de la forme variationnelle '.1' pour les instants -3, -2, -1 et 0 sont les suivants.

Pour l'instant -3:

$$FV_{(1, -3)} = ((FV_{(1, -4)}) * .1) + ((PSI_{-3} - PSI_{-4}) * .1)$$

$$FV_{(1, -3)} = (0 * .1) + ((-100 - 0) * .1)$$

$$FV_{(1, -3)} = 0 + (-10)$$

$$FV_{(1, -3)} = (-10)$$

Pour l'instant -2:

$$FV_{(1, -2)} = ((FV_{(1, -3)}) * .1) + ((PSI_{-2} - PSI_{-3}) * .1)$$

$$FV_{(1, -2)} = (-10 * .1) + ((-130 - (-100)) * .1)$$

$$FV_{(1, -2)} = -1 + (-3)$$

$$FV_{(1, -2)} = (-4)$$

Pour l'instant -1:

$$FV_{(1, -1)} = ((FV_{(1, -2)}) * .1) + ((PSI_{-1} - PSI_{-2}) * .1)$$

$$FV_{(1, -1)} = (-4 * .1) + ((-140 - (-130)) * .1)$$

$$FV_{(1, -1)} = -.4 + (-1)$$

$$FV_{(1, -1)} = (-)1.4$$

Et pour l'instant 0:

$$FV_{(1, 0)} = ((FV_{(1, -1)}) * .1) + ((PSI_0 - PSI_{-1}) * .1)$$

$$FV_{(1, 0)} = (-1.4 * .1) + ((-145 - (-140)) * .1)$$

$$FV_{(1, 0)} = -.14 + (-.5)$$

$$FV_{(1, 0)} = (-).64$$

Considérons cet exemple dans la perspective de nos deux réservoirs. Le réservoir de la structure d'interface se vide⁸ jusqu'à la marque '-100', puis jusqu'à la marque '-130', puis jusqu'à la marque '-140' et finalement jusqu'à la marque '-145'. Le réservoir de la forme variationnelle '.1', quant à lui, se vide jusqu'à la marque '(-)10' à l'instant -3. Si la variation était nulle au moment suivant, ce réservoir se re-remplirait jusqu'à la marque '(-)1'. Mais la variation négative de ce moment (-100 -> -130) ne fait remonter le niveau du réservoir que jusqu'à la marque '(-)4'. Le même phénomène se produirait à l'instant suivant si la variation était nulle. Le niveau du réservoir remonterait de la marque '(-)4' à la marque '(-).4', mais la variation existante (-130 -> -140) ne fait remonter le niveau du réservoir que jusqu'à la marque '(-)1.4' et il en est de même à l'instant 0. Si la variation était nulle, le niveau

⁸ Rappelons que la position 0 de la structure d'interface n'indique pas une absence d'énergie. Elle correspond à la quantité énergétique "normale" (à la valeur de pression normale. Lorsque cette position devient négative, cela signifie que la quantité énergétique diminue sous cette quantité normale.

du réservoir remonterait jusqu'à la marque '(-).14' mais la légère variation négative (-140 -> -145) ne fait remonter ce niveau que jusqu'à la marque '(-).64'. La figure 13 représente les valeurs de la forme variationnelle '1' pour la séquence de positions de la figure 12, ainsi que les valeurs qui auraient été calculées si la variation avait été nulle durant le moment d'échantillonnage suivant chaque calcul de la valeur réelle.

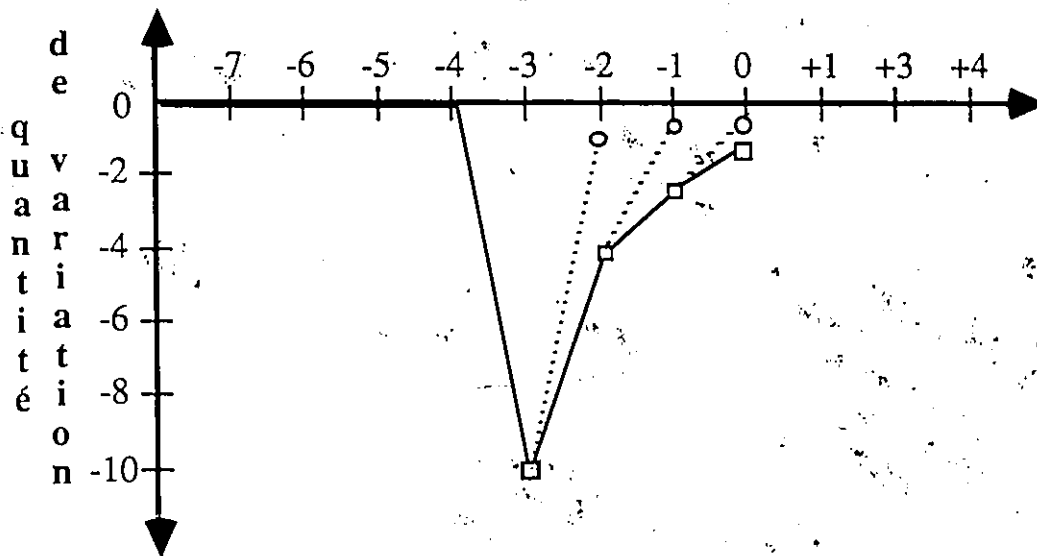


Fig. 13 Le comportement du réservoir sous-jacent à l'élément structural de la forme variationnelle '1' pour la séquence de positions de la figure 12

Dans notre mode de représentation des positions de la structure d'interface, nous savons que la position 0 correspond à la quantité

énergétique normale du médium de transmission. Lorsque ces valeurs deviennent négatives, par exemple -15, -300 ou -500, cela signifie que la quantité énergétique *diminue*. Plus cette valeur négative augmente, plus la quantité énergétique diminue. Dans le mode de représentation des valeurs de la forme variationnelle '1', les valeurs négatives n'ont pas la même signification. *Plus ces valeurs négatives augmentent, plus cette forme variationnelle contribue à l'histoire variationnelle de la position présente de la structure d'interface*. Nous devons interpréter les figures 11 et 13 de la manière suivante:

1) Les quadrants positif et négatif du plan vertical (ordonnée) représentent respectivement la direction de variation positive (+) et la direction de variation négative (-). Lorsque la valeur de la forme variationnelle '1' se situe, à un instant donné, dans le quadrant positif (en haut de 0), cela signifie que cette forme variationnelle contribue par une variation positive à l'histoire variationnelle de la position présente de la structure d'interface. Lorsque cette valeur se situe dans le quadrant négatif, cela signifie que cette forme variationnelle contribue par une variation négative à l'histoire variationnelle de la position présente de la structure d'interface.

2) Plus cette valeur est distante de 0, que ce soit vers $+\infty$ ou vers $-\infty$, plus la quantité de contribution (la quantité de variation) de cette forme variationnelle est importante.

Des valeurs de (+)10 et de (-)10, par exemple, indiquent la même *quantité* de contribution, mais deux contributions de *direction* différente. Des valeurs de (+)10 et de (+)100 indiquent deux contributions de même direction (positive) mais de quantité différente, la deuxième valeur indiquant une contribution plus importante que la première. Il en est de même pour des valeurs de (-)10 et de (-)100; il s'agit de deux contributions de même

direction (négative) mais la deuxième valeur indique une contribution plus importante que la première.

Nous avons vu qu'il était possible d'interpréter les quantités de variation de chaque forme variationnelle en termes de contributions plus ou moins importantes de différentes bandes de vitesses. Chaque forme variationnelle est spécifique à la variation sur une certaine bande d'intervalles temporels et chaque forme variationnelle est associée à une certaine bande de vitesses. Il y aurait donc une bande de vitesses et une bande d'intervalles temporels spécifiques à la forme variationnelle '.1'. Aux figures 11 et 13, la quantité de variation de cette forme variationnelle diminue (de l'instant -3 à l'instant 0); cela voudrait donc dire que les valeurs de vitesses (de la bande de vitesses associée à cette forme) contribuent de moins en moins à l'histoire variationnelle de chaque position de la structure d'interface des figures 4 et 12. Mais quelle est la bande de vitesses associée à la forme '.1'? Celle-ci se constitue d'un ensemble de valeurs de vitesse qui va d'une valeur de vitesse infiniment grande à une valeur de vitesse intermédiaire déterminée par la valeur de la constante temporelle '.1'. Si nous avons choisi une valeur de constante temporelle plus petite, '.05' par exemple, la bande de vitesse de cette autre forme variationnelle serait différente. Elle regrouperait un ensemble de vitesses qui partirait toujours d'une valeur de vitesse infiniment grande mais qui s'arrêterait, cette fois, à une valeur de vitesse intermédiaire *plus élevée (une plus grande vitesse)* que la vitesse intermédiaire de la forme temporelle '.1'. Par contre, si nous avons choisi une constante temporelle plus grande, '.5 par exemple, la bande de vitesses de cette autre forme regrouperait un ensemble de vitesses qui partirait toujours d'une vitesse

infiniment grande mais qui s'arrêterait, cette fois, à une valeur de vitesse intermédiaire *plus petite* que la vitesse intermédiaire de la forme temporelle '.1'. En termes de bandes d'intervalles temporels, cela signifie que la bande d'intervalles temporels qui est associée à la forme '.1' se constitue des intervalles temporels qui vont de l'intervalle temporel infiniment court à un intervalle temporel d'une longueur temporelle intermédiaire. Lorsque nous diminuons la valeur de constante temporelle (vers 0), nous raccourcissons la bande d'intervalles temporels, lorsque nous augmentons cette valeur (vers 1), nous allongeons cette bande. Le rapport entre la valeur de la constante temporelle et la dimension des bandes d'intervalles temporels et de vitesses est illustré à la figure 14 (page suivante). Plus la constante temporelle se rapproche de 1, plus les bandes de vitesses et d'intervalles s'élargissent. La bande de vitesse s'élargit en partant toujours de la valeur de vitesse infiniment grande, alors que la bande d'intervalles temporels s'élargit en partant d'un intervalle temporel de longueur 0. *Lorsque la constante est égale à 1, la bande de vitesses regroupe toutes les vitesses possibles et la bande d'intervalles regroupe tous les intervalles temporels possibles. Nous avons alors le comportement de la structure d'interface.* Le comportement de la structure d'interface reflète l'effet de toutes les formes de variation. On y retrouve la contribution d'une seule bande de vitesses qui va d'une valeur de vitesse nulle à une valeur de vitesse infiniment grande, et on y retrouve la variation spécifique à la bande d'intervalles temporels qui va de l'intervalle temporel de longueur 0 à l'intervalle temporel de longueur infinie. Les différents éléments structuraux de la structure d'analyse spectrale sont sensibles, eux, aux différentes formes variationnelles.

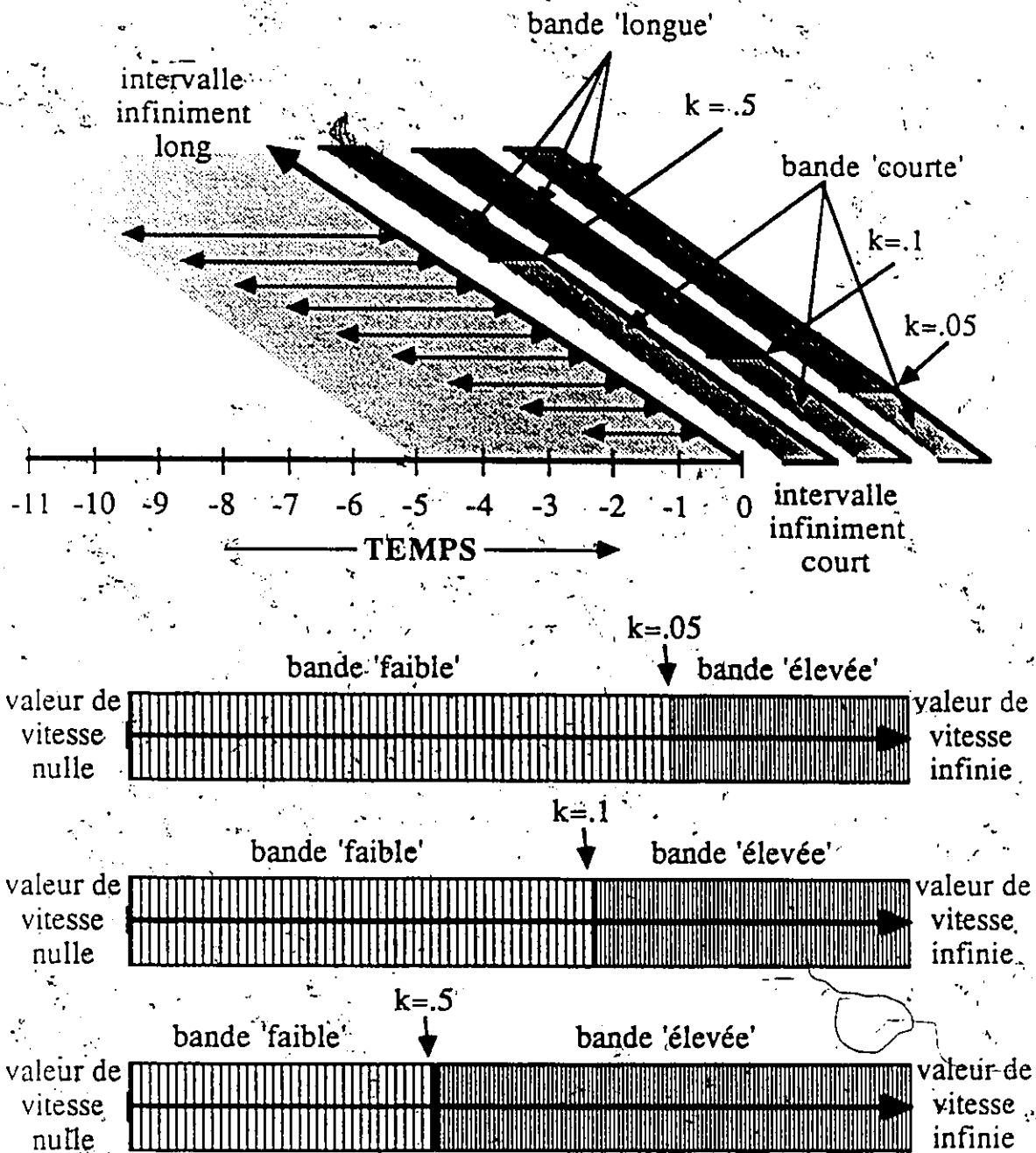


Fig. 14 L'effet de la valeur de la constante temporelle sur la largeur des bandes de vitesses et d'intervalles temporels

Chacun de ces éléments ne reflète que la contribution d'une seule bande de vitesses et n'est spécifique qu'à la variation sur une seule bande d'intervalles temporels. Si nous élargissons ou rapetissons la bandes de vitesses et d'intervalles temporels (en modifiant la valeur de constante temporelle) de notre première forme variationnelle, cela se traduirait par une fluctuation de la quantité de variation de cette forme variationnelle. En rapprochant la valeur de la constante temporelle de un (1), nous rendrions l'élément structural sous-jacent à la caractérisation de cette forme plus sensible à une large bande de valeurs de vitesses et cet élément deviendrait spécifique à la variation sur une plus grande bande d'intervalles temporels; ce qui se traduirait par une quantité de variation plus élevée. En rapprochant cette valeur de zéro (0), nous produirions l'effet inverse, et cela se traduirait par une quantité de variation plus faible. Dans le cas de l'exemple de la figure 4, un groupement temporel à constante temporelle de '.05' produirait, par exemple, des valeurs d'attribut de (+)5, (+)1.75, (+).5875, (+).28, alors qu'un groupement temporel à constante temporelle '.5' produirait des valeurs d'attribut de (+)50, (+)40, (+)25, (+)15.

En attribuant une valeur de constante temporelle à la forme variationnelle '.1', nous nous trouvons donc à définir une bande 'élevée' de valeurs de vitesse et une bande 'courte' de valeurs d'intervalles temporels. Aux exemples des figures 11 et 13, la quantité de variation de la forme '.1' diminue de l'instant -3 à l'instant 0. Si cette quantité de variation diminue, cela signifie que la quantité de variation sur la bande d'intervalles temporels 'courts' diminue. Cela signifie également que la contribution de la bande de vitesses 'élevées' diminue. *Si la quantité de variation de la forme '.1'*

diminue, cela signifie, de plus, que la quantité de variation de toutes les autres formes variationnelles augmente, puisque la position de la structure d'interface est le produit de toutes les formes variationnelles. A l'instant -3, par exemple (figure 4), la position de la structure d'interface est de +100 et la contribution de la forme variationnelle '.1' est de (+) 10. La quantité et la direction de variation de toutes ces autres formes variationnelles nous est donnée par la différence entre la position de la structure d'interface à chaque instant et la contribution de la forme variationnelle '.1' au même instant, c'est à dire, dans ce cas-ci, (+)90. *Autrement dit, si nous avions pu éliminer la forme variationnelle '.1' de la variation qui s'est produite jusqu'à l'instant -3, la position de la structure d'interface ne serait que de +90 à cet instant.* Nous pouvons appliquer ce raisonnement à tous les instants d'échantillonnage. A l'instant -2, la quantité de variation de la forme variationnelle '.1' passe à (+)4. La quantité et la direction de variation de toutes les autres formes variationnelles est donc de (+)126 (130-4): la quantité de variation de toutes les autres formes augmente (la direction restant la même). Cette quantité de variation va continuer à augmenter aux deux instants suivants: elle sera de (+) 138.6 (140 -1.4) à l'instant -1 et de (+)144.36 (145-.64) à l'instant 0.

Nous pouvons considérer la quantité et la direction de variation de toutes ces autres formes variationnelles comme la quantité et la direction d'une seule autre forme variationnelle: celle qui est spécifique à la bande d'intervalles temporels résiduelle (la bande 'longue') et celle qui est sensible à la bande de vitesses résiduelle (la bande 'faible'). La bande 'longue' d'intervalles temporels se constitue des intervalles temporels qui vont de l'intervalle temporel intermédiaire de la forme '.1' jusqu'à l'intervalle

temporel infiniment long, alors que la bande 'faible' de vitesses se constitue des valeurs de vitesse qui vont de la valeur de vitesse intermédiaire jusqu'à la valeur de vitesse zéro (0).

Une structure d'analyse spectrale permet habituellement la caractérisation de plus de *deux* formes variationnelles. La cochlée, par exemple, se compose d'un très grand nombre d'éléments structuraux sensibles à plusieurs bandes de vitesses et d'intervalles temporels⁹. Nous pouvons augmenter le nombre de formes variationnelles à trois (3) en divisant les bandes 'faible' et 'longue' de vitesses et d'intervalles temporels. *Pour effectuer cette division, nous n'avons qu'à décréter un nouveau groupement temporel sur la position résiduelle de la structure d'interface à chaque instant, c'est à dire sur la position originale de cette structure, à chaque instant, moins les valeurs d'attribut de la forme '.1' au même instant.* En effectuant ce nouveau groupement temporel, nous nous trouvons à caractériser une nouvelle forme variationnelle. Cette nouvelle forme variationnelle est spécifique à une nouvelle bande (la bande 'moyenne' d'intervalles temporels) d'intervalles temporels qui se situe entre la valeur de l'intervalle temporel long (l'ancien intervalle intermédiaire) de la bande de la forme '.1' et une nouvelle valeur d'intervalle intermédiaire. Cette nouvelle forme est également sensible à une nouvelle bande (la bande 'moyenne' de vitesses) de vitesses qui se situe entre la valeur de vitesse faible (l'ancienne valeur de vitesse intermédiaire) de la forme '.1' et une nouvelle valeur de vitesse intermédiaire (figure 15). La

⁹ Dans la perspective "continue" de la membrane basilaire, la cochlée se constitue d'une infinité d'éléments structuraux. Dans la perspective "discrète" des fibres nerveuses qui sont "branchées" sur la membrane basilaire, la cochlée se constitue d'environ 30000 éléments structuraux.

valeur de constante temporelle que nous choisirons pour ce nouveau groupement temporel déterminera la largeur de bande de ces deux bandes 'moyenne'. Plus la valeur de la constante temporelle sera proche de 0, plus ces nouvelles bandes de vitesses 'moyenne' seront étroites.

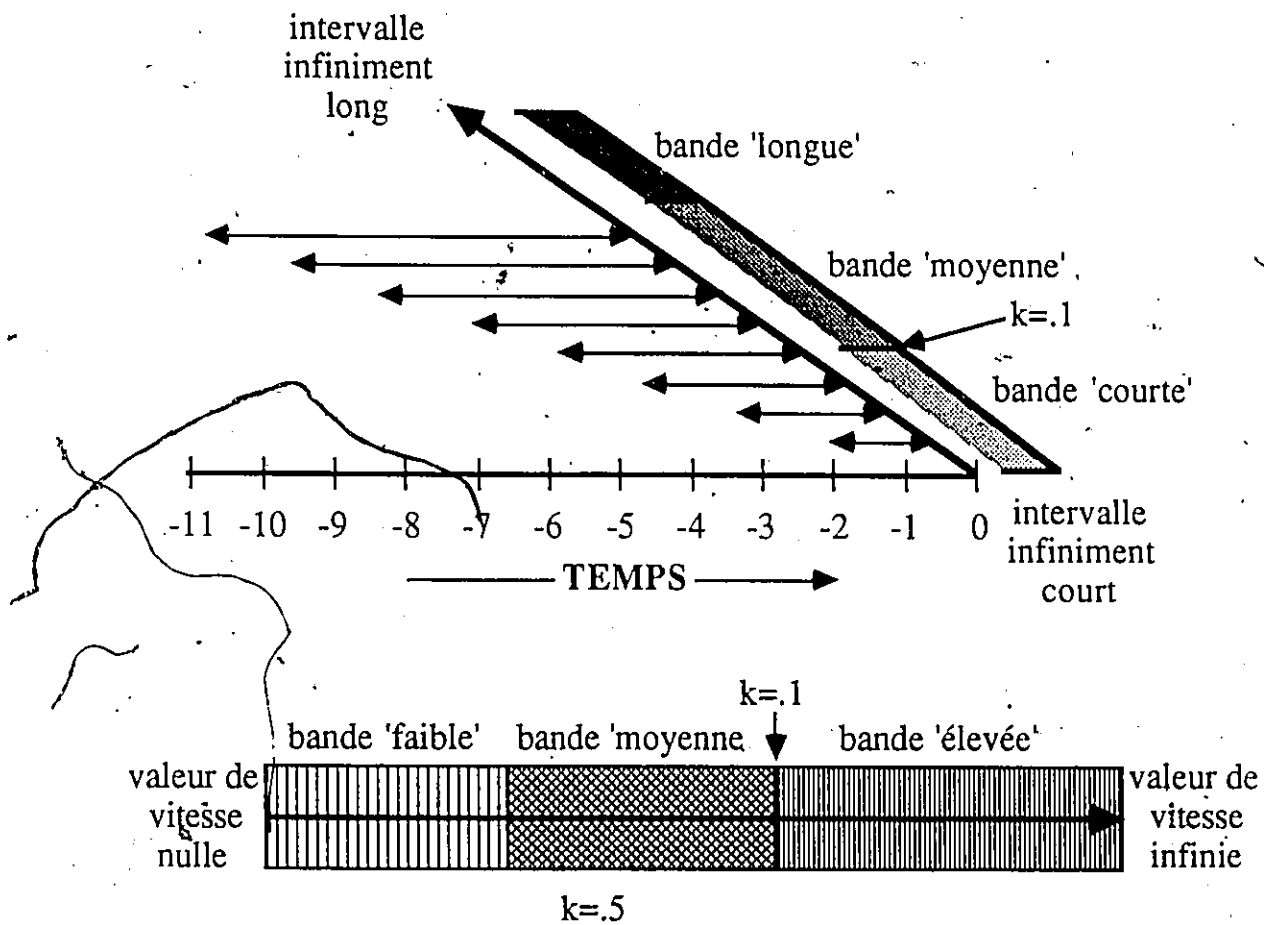


Fig. 15 La division des bandes de vitesses et d'intervalles temporels par l'application d'un nouveau groupement temporel sur la position résiduelle de la structure d'interface

Choisissons une valeur de constante temporelle de '.2' et effectuons le calcul de la contribution de cette nouvelle forme variationnelle pour l'instant d'échantillonnage -3 de la figure 4:

$$FV_{(.2(.1), -3)} = ((FV_{(.2(.1), -4)}) * .2) + ((PRSI_{(.1, -3)} - PRSI_{(.1, -4)}) * .2)$$

$$FV_{(.2(.1), -3)} = (0 * .2) + ((90 - 0) * .2)$$

$$FV_{(.2(.1), -3)} = 0 + 18$$

$$FV_{(.2(.1), -3)} = (+)18$$

Nous avons un peu changé notre terminologie. Ce groupement temporel s'effectue sur la position résiduelle de la structure d'interface (PRSI), position résiduelle qui correspond à la position de la structure d'interface à chaque instant *moins* la contribution de la forme variationnelle '.1' au même instant. Pour bien indiquer qu'il s'agit d'un groupement temporel sur une valeur (la position résiduelle) dont on a soustrait la contribution d'une première forme variationnelle '.1', nous avons ajouté la constante temporelle de cette première forme variationnelle en indice du symbole FV ($FV_{(.2(.1), -3)}$ et $FV_{(.2(.1), -4)}$). Nous avons également ajouté cette constante en indice du symbole PRSI ($PRSI_{(.1, -3)}$ et $PRSI_{(.1, -4)}$) pour indiquer qu'il s'agit de la position résiduelle¹⁰ de la structure d'interface dont on a soustrait la valeur de contribution de la forme temporelle '.1'.

Ce résultat de (+)18 représente la contribution quantitative et directionnelle de cette nouvelle forme variationnelle '.2(.1)' à l'instant -3.

¹⁰ Nous verrons plus loin qu'il peut y avoir d'autres positions résiduelles.

Elle correspond à la quantité et à la direction de variation spécifiques à cette nouvelle bande 'moyenne' d'intervalles temporels et elle correspond également à la contribution de cette nouvelle bande 'moyenne' de valeurs de vitesse. Voici les calculs de contribution de cette forme variationnelle pour les instants -2, -1 et 0. Pour l'instant -2:

$$FV_{(.2(.1), -2)} = ((FV_{(.2(.1), -3)}) * .2) + ((PRSI_{(.1, -2)} - PRSI_{(.1, -3)}) * .2)$$

$$FV_{(.2(.1), -2)} = (18 * .2) + ((126 - 90) * .2)$$

$$FV_{(.2(.1), -2)} = 3.6 + 7.2$$

$$FV_{(.2(.1), -2)} = (+) 10.8$$

Pour l'instant -1:

$$FV_{(.2(.1), -1)} = ((FV_{(.2(.1), -2)}) * .2) + ((PRSI_{(.1, -1)} - PRSI_{(.1, -2)}) * .2)$$

$$FV_{(.2(.1), -1)} = (10.8 * .2) + ((138.6 - 126) * .2)$$

$$FV_{(.2(.1), -1)} = 2.16 + 2.52$$

$$FV_{(.2(.1), -1)} = (+) 4.68$$

et pour l'instant 0:

$$FV_{(.2(.1), 0)} = ((FV_{(.2(.1), -1)}) * .2) + ((PRSI_{(.1, 0)} - PRSI_{(.1, -1)}) * .2)$$

$$FV_{(.2(.1), 0)} = (4.68 * .2) + ((144.36 - 138.6) * .2)$$

$$FV_{(.2(.1), 0)} = .936 + 1.152$$

$$FV_{(.2(.1), 0)} = (+) 2.088$$

Nous pouvons constater que la quantité de variation de cette forme variationnelle décroît progressivement comme pour la forme '.2(.1)'. La contribution de cette forme est toutefois plus élevée, à chaque instant, que la contribution de la forme '.1'. Ceci est dû au choix de la constante temporelle .2. Si nous avons choisi une constante plus petite que .12 (approx.), la valeur calculée aurait été plus petite que celle de la forme '.1'.

Nous disposons donc maintenant de trois formes variationnelles. En créant la forme variationnelle '.2(.1)', nous avons déterminé deux nouvelles bandes 'moyenne' d'intervalles temporels et de vitesses. Nous avons donc également déterminé deux nouvelles bandes 'faible' (de valeurs de vitesse) et 'longue' (d'intervalles temporels) et ces deux nouvelles bandes correspondent, encore une fois, aux formes variationnelles résiduelles. Nous pouvons toujours considérer ces formes variationnelles résiduelles comme une seule forme variationnelle et calculer les valeurs d'attribut de cette forme en additionnant, tout d'abord, les valeurs d'attributs des formes '.1' et '.2(.1)' puis en soustrayant ces valeurs de la position initiale de la structure d'interface. Cela donne (+)72 (+100 -(+)28) pour l'instant -3, (+)115.2 (+130 - (+)14.8) pour l'instant -2, (+)133.9 (+140 - (+)6.1) pour l'instant -1, et finalement (+)142.25 (+145 - (+)2.75), pour l'instant 0. Ces valeurs représentent également une autre position résiduelle de la structure d'interface. *Si nous avons pu éliminer les contributions conjointes des formes variationnelles '.1' et '.2(.1)', de la variation qui s'est produite jusqu'à ces différents instants d'échantillonnage, la position de la structure d'interface aurait été celle des valeurs d'attributs de la forme variationnelle résiduelle (+72, +115.2, +133.9, +142.25).*

Il est possible de créer une nouvelle forme variationnelle en subdivisant de nouveau les bandes 'faible' et 'longue' de vitesses et d'intervalles temporels. Pour ce faire, nous n'avons qu'à effectuer un nouveau groupement temporel sur cette *deuxième* position résiduelle de la structure d'interface. Nous obtenons alors quatre formes variationnelles qui sont le reflet de quatre paires de bandes de vitesses et d'intervalles temporels: les mêmes bandes 'élevée' et 'courte' de la forme '.1', les mêmes bandes 'moyenne' de la forme '.2(.1)', une nouvelle et deuxième paire de bandes 'moyenne' (nouvelle forme variationnelle) et une nouvelle et deuxième paire de bandes 'faible' et 'longue' (nouvelle forme variationnelle). Pour effectuer ce calcul, nous devons choisir une nouvelle valeur de constante temporelle. Avec une valeur de constante temporelle de .3, cela donne pour l'instant -3 de l'exemple de la figure 4:

$$FV_{(.3(.2), -3)} = ((FV_{(.3(.2), -4)}) * .3) + ((PRSI_{(.2,-3)} - PRSI_{(.2,-4)}) * .3)$$

$$FV_{(.3(.2), -3)} = (0 * .3) + ((72 - 0) * .3)$$

$$FV_{(.3(.2), -3)} = 0 + 21.6$$

$$FV_{(.3(.2), -3)} = (+)21.6$$

Cette fois-ci, les changements de notation sont minimaux. Nous avons modifié la valeur de constante temporelle à .3; l'indice¹¹ de cette nouvelle forme variationnelle est devenue $FV_{(.3(.2), -3)}$ et l'indice de la position résiduelle est maintenant $PRSI_{(.2, -3)}$. Les résultats de ce groupement temporel pour les autres instants d'échantillonnage sont les suivants:

Pour l'instant -2:

$$FV_{(.3(.2), -2)} = ((FV_{(.3(.2), -3)}) * .3) + ((PRSI_{(.2, -2)} - PRSI_{(.2, -3)}) * .3)$$

$$FV_{(.3(.2), -2)} = (21.6 * .3) + ((115.2 - 72) * .3)$$

$$FV_{(.3(.2), -2)} = 6.5 + 12.9$$

$$FV_{(.3(.2), -2)} = (+)19.4$$

Pour l'instant -1:

$$FV_{(.3(.2), -1)} = ((FV_{(.3(.2), -2)}) * .3) + ((PRSI_{(.2, -1)} - PRSI_{(.2, -2)}) * .3)$$

$$FV_{(.3(.2), -1)} = (19.4 * .3) + ((133.9 - 115.2) * .3)$$

$$FV_{(.3(.2), -1)} = 5.9 + 5.6$$

$$FV_{(.3(.2), -1)} = (+)11.5$$

¹¹ Il aurait été plus juste d'exprimer l'indice de la forme variationnelle et celui de la position résiduelle sous les formes suivantes: $FV_{(.3(.2(.1)), -3)}$ et $PRSI_{(.2(.1), -3)}$. Nous avons décidé d'utiliser une notation plus "légère" afin de ne pas surcharger les équations. Il faut toutefois garder à l'esprit que ce groupement temporel s'effectue sur une position résiduelle où le "résidu" variationnel découle des groupements temporels successifs des formes '.1' et '.2(.1)'.

et pour l'instant 0:

$$FV_{(.3(.2), 0)} = ((FV_{(.3(.2), -1)}) * .3) + ((PRSI_{(.2, 0)} - PRSI_{(.2, -1)}) * .3)$$

$$FV_{(.3(.2), 0)} = (11.5 * .3) + ((142.25 - 133.9) * .3)$$

$$FV_{(.3(.2), 0)} = 3.4 + 2.5$$

$$FV_{(.3(.2), 0)} = (+)5.9$$

Nous pouvons continuer à subdiviser ainsi les bandes 'faible' et 'longue' de vitesses et d'intervalles temporels en effectuant une série de groupements temporels sur les positions résiduelles successives de la structure d'interface. A chaque nouveau groupement temporel, nous définissons une nouvelle paire de bandes 'moyenne' de vitesses et d'intervalles temporels dont la largeur est fonction du choix de la constante temporelle. Le tableau 1 représente les valeurs d'attribut de dix (10) formes variationnelles à cinq (5) instants d'échantillonnage de la figure 4. Nous pouvons remarquer tout d'abord que le total des valeurs d'attribut de chaque forme variationnelle équivaut, à chaque instant, à la valeur de la position de la structure d'interface: cette position s'explique complètement par les valeurs d'attribut de toutes les formes variationnelles. Toutes les quantités de variation des différentes formes augmentent de l'instant -4 à l'instant -3. Par la suite, la quantité de variation des formes variationnelles '.1', '.2(.1)' et '.3(.2)' diminuent d'instant en instant. La quantité de variation de la forme variationnelle '.4(.3)' augmente de l'instant -3 ((+)20.2) à l'instant -2 ((+)26.2) pour se mettre à diminuer par la suite, ((+)21.2, (+)14). Le même phénomène se produit pour la forme variationnelle '.5(.4)' mais la diminution se produit un

instant plus tard ((+)15.1, (+)27.2, (+)29.5, (+)25.3). Les quantités de variation des formes variationnelles restantes augmentent, elles, progressivement, à chaque instant d'échantillonnage, mais on peut voir que cette augmentation est beaucoup plus prononcée pour la forme '6(.5)' que pour la forme '1(.9)'.

Formes variationnelles	Instant -4	Instant -3	Instant -2	Instant -1	Instant 0
'1'	0	(+)10	(+)4	(+)1.4	(+)0.6
'2(.1)'	0	(+)18	(+)10.8	(+)4.7	(+)2.1
'3(.2)'	0	(+)21.6	(+)19.4	(+)11.5	(+)5.8
'4(.3)'	0	(+)20.2	(+)26.2	(+)21.2	(+)14
'5(.4)'	0	(+)15.1	(+)27.2	(+)29.5	(+)25.3
'6(.5)'	0	(+)9	(+)21.8	(+)30.7	(+)33.6
'7(.6)'	0	(+)4.2	(+)13.2	(+)23.5	(+)32.2
'8(.7)'	0	(+)1.5	(+)5.6	(+)12.5	(+)21.1
'9(.8)'	0	(+)0.3	(+)1.6	(+)4.3	(+)8.6
'1(.9)'	0	(+)0.1	(+)0.2	(+)0.7	(+)1.7
	0	+100	+130	+140	+145

Tableau 1

Les valeurs d'attribut de dix (10) formes variationnelles

Chaque forme variationnelle est associée à une paire de bandes de vitesses et d'intervalles temporels spécifique. Ces bandes sont juxtaposées (côte à côte) et s'échelonnent des vitesses les plus élevées et des intervalles les plus courts (forme '.1') aux vitesses les plus faibles et aux intervalles les plus longs (forme '1(.9)'). La quantité et la direction de variation de chaque forme variationnelle représente l'importance et la direction de la variation spécifique à la bande d'intervalles temporels qui lui est associée, de même que la contribution (ou la présence) de la bande de vitesses qui lui est également associée. L'histoire variationnelle de la position de la structure d'interface, à l'instant -4, se caractérise donc par une quantité nulle et une absence de direction de la variation spécifique à ces différentes bandes d'intervalles temporels de même que par une contribution nulle de toutes les bandes de vitesse. L'histoire variationnelle de la position de l'instant -3 se caractérise par des formes variationnelles dont la quantité de variation se répartit sur une courbe en forme de cloche. La forme de cette cloche est asymétrique: la courbe du côté des formes de variation à bandes 'courtes' d'intervalles temporels et à bandes 'élevées' de valeurs de vitesse est plus prononcée que la courbe du côté des formes de variation à bandes 'longue' d'intervalles temporels et à bande 'faible' de valeurs de vitesse. Le mode de cette courbe se situe au niveau de la forme '.3(.2)' ((+)21.6). A l'instant suivant (instant -2), l'histoire variationnelle de la position de la structure d'interface se caractérise toujours par une courbe en forme de cloche, mais la forme de cette cloche s'est modifiée; elle est toujours asymétrique du côté 'court-élevée', mais le mode de cette courbe s'est déplacé: il se situe deux formes variationnelles plus loin, au niveau de la forme '.5(.4)' ((+)27.2). Ce mouvement se poursuit à l'instant suivant (instant -1): le mode de la courbe se

situe une forme variationnelle plus loin, au niveau de la forme '.6(.5)' ((+ 30.7). A l'instant 0 finalement, le mode est toujours au niveau de la forme '.6(.5)', mais on peut sentir que le mouvement de la courbe se poursuit, lorsque nous considérons la fluctuation des valeurs de quantité de variation des différentes formes variationnelles. La quantité de variation des formes variationnelles 'courtes-élevées' continue à diminuer, alors que la quantité de variation des formes 'longues-faibles' continue d'augmenter. De l'instant -3 à l'instant 0, l'histoire variationnelle de la position de la structure d'interface se transforme donc dans le sens d'une quantité de variation accrue pour les formes variationnelles qui sont spécifiques à des bandes d'intervalles temporels qui se constituent d'intervalles temporels de plus en plus longs. La transformation de l'histoire variationnelle est également caractéristique d'une contribution accrue des bandes de vitesses qui se constituent de valeurs de vitesses de plus en plus faibles. Cela correspond à notre caractérisation intuitive (et traditionnelle) de la vitesse de la variation de l'exemple de la figure 4. La structure d'analyse spectrale nous permet d'améliorer, de perfectionner cette caractérisation globale en l'exprimant sous la forme des contributions de multiples bandes de vitesses.

Cela termine cette sous-section où nous avons développé un algorithme pour calculer les valeurs d'attributs des différentes formes variationnelles. Cet algorithme repose sur une série de groupements temporels particuliers qui permettent de caractériser l'histoire variationnelle de chaque position de la structure d'interface. Ces groupements temporels s'effectuent en modifiant les valeurs de constante temporelle du groupement temporel qui est caractéristique de la structure d'interface. Cette dernière structure produit

une entité qui est le reflet de toutes les formes de variation, alors que la structure d'analyse spectrale produit des entités qui sont le reflet de différentes formes variationnelles.

3.3.2 La simulation computationnelle de la structure d'interface

L'algorithme que nous venons de présenter a fait l'objet d'une simulation computationnelle de la structure d'analyse spectrale¹². Dans cette simulation, la valeur de position de la structure d'interface est passée, à chaque instant, à la structure simulée d'analyse spectrale et celle-ci effectuée, à chacun de ces instants, le calcul de la quantité et de la direction de variation des différentes formes variationnelles. La simulation nous permet de produire ou de reproduire n'importe quelle séquence de valeurs de position de la structure d'interface.

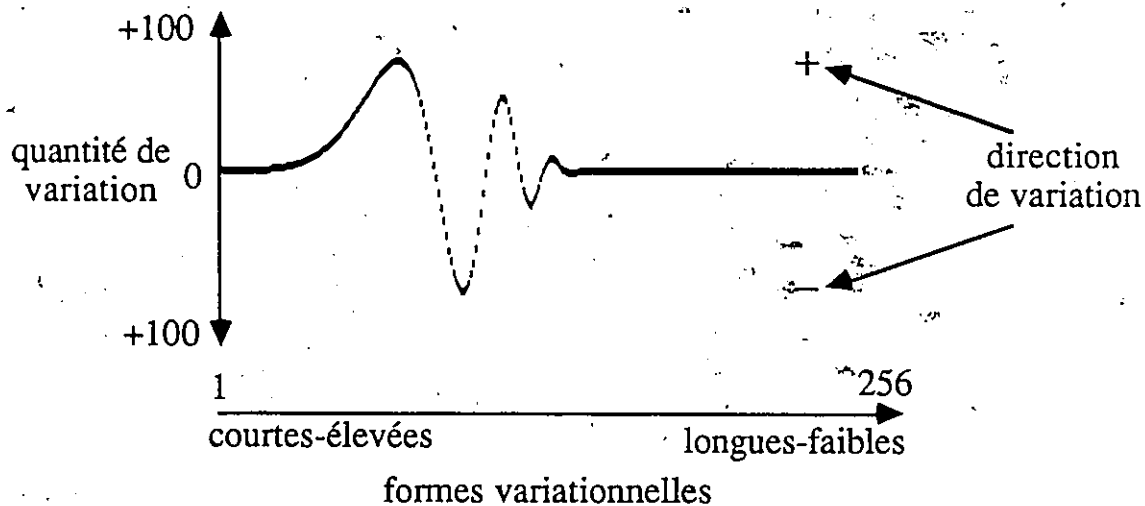


Fig. 16 La simulation computationnelle de la structure d'analyse spectrale (F= 32; A = 900)

¹² Le programme est présenté en détail à l'annexe 1.

La simulation nous permet également de déterminer le nombre de formes variationnelles ainsi que les valeurs de constante temporelle des groupements temporels qui sont sous-jacents à la production de ces formes (le nombre et la largeur des différentes bandes de vitesses et d'intervalles temporels). La figure 16 représente la réponse simulée de la structure d'analyse spectrale à un instant d'échantillonnage dans le temps. La séquence de positions de la structure d'interface de cet exemple représente une séquence sinusoïdale de fréquence 32 et d'amplitude 900 (la séquence de positions d'un son pur d'une certaine hauteur et d'une certaine intensité). Chaque point de la courbe¹³ représente les valeurs d'attributs d'une forme variationnelle différente. A chaque coordonnée d'abscisse correspond une forme variationnelle différente. La forme variationnelle de coordonnée d'abscisse '1' est la forme variationnelle spécifique à une bande de très courts intervalles temporels (la bande 'courte' qui se constitue des intervalles temporels qui vont de l'intervalle infiniment court à un intervalle intermédiaire); elle est également associée à une bande de vitesses très élevées (la bande 'élevée' qui va d'une valeur de vitesse infiniment grande à une valeur de vitesse intermédiaire). La forme variationnelle de coordonnée d'abscisse '2' est une forme variationnelle spécifique à une bande d'intervalles temporels un peu plus longs que la première forme, et cette forme est également associée à une bande de vitesses un peu plus faible. Il en est de même pour les formes variationnelles qui vont de la coordonnée d'abscisse 3 jusqu'à la coordonnée d'abscisse 256. Toutes ces formes sont spécifiques à des bandes d'intervalles temporels de plus en plus longs, et ces formes sont associées à des bandes de

¹³ Il y a deux cents cinquante-six (256) points dans cette figure (256 formes variationnelles).

vitesse de plus en plus faibles. La forme de la coordonnée d'abscisse 256 est la forme variationnelle spécifique à une bande d'intervalles temporels qui va jusqu'à un intervalle temporel infiniment long, et cette forme est associée à une bande de vitesses qui va jusqu'à la valeur de vitesse nulle.

La valeur d'ordonnée de chaque forme variationnelle nous donne accès à la quantité et à la direction de variation de cette forme variationnelle à cet instant d'échantillonnage. Les sections positive et négative de l'axe vertical (ordonnée) représentent respectivement la direction de variation positive (+) et la direction de variation négative (-). Lorsque le point correspondant à une forme variationnelle se situe, à un instant donné, dans la section positive (en haut de 0), cela signifie que cette forme variationnelle contribue par une variation positive à l'histoire variationnelle de la position présente de la structure d'interface. Lorsque le point correspondant à une forme variationnelle se situe dans la section négative, cela signifie que cette forme variationnelle contribue par une variation négative à l'histoire variationnelle de la position présente de la structure d'interface. Finalement, plus ce point est distant de 0, que ce soit vers le haut (vers $+\infty$) ou vers le bas (vers $-\infty$), plus la quantité de variation de cette forme variationnelle est importante. La figure 17 représente la réponse de la structure, à un instant, pour une séquence de positions de même fréquence mais de moindre amplitude, alors que la figure 18 représente la réponse instantanée de la structure à une séquence de positions de fréquence différente mais d'amplitude identique. Finalement la figure 19 représente la réponse de la structure à une stimulation composée de deux fréquences d'amplitudes différentes.

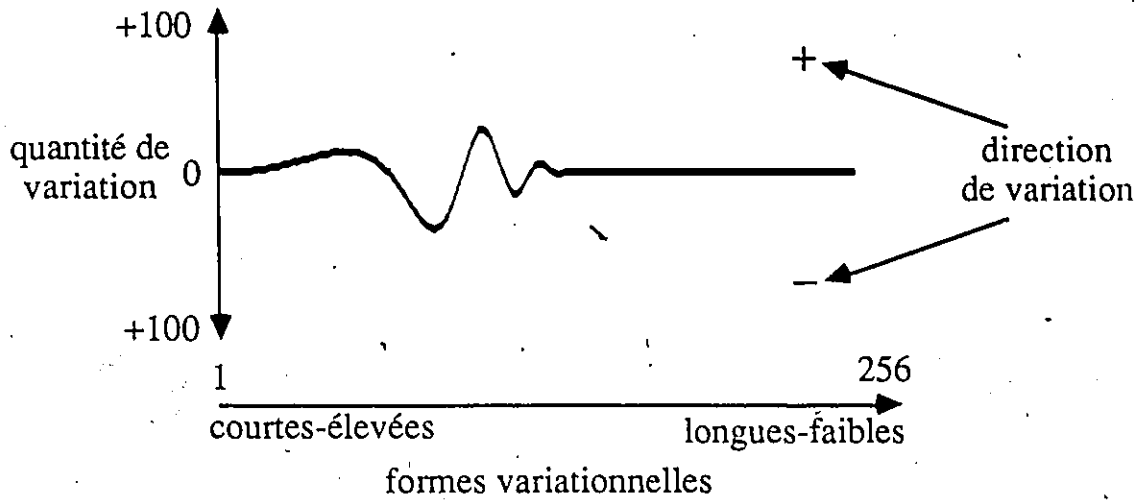


Fig. 17 La simulation computationnelle de la structure d'analyse spectrale ($F=32$; $A=20$)

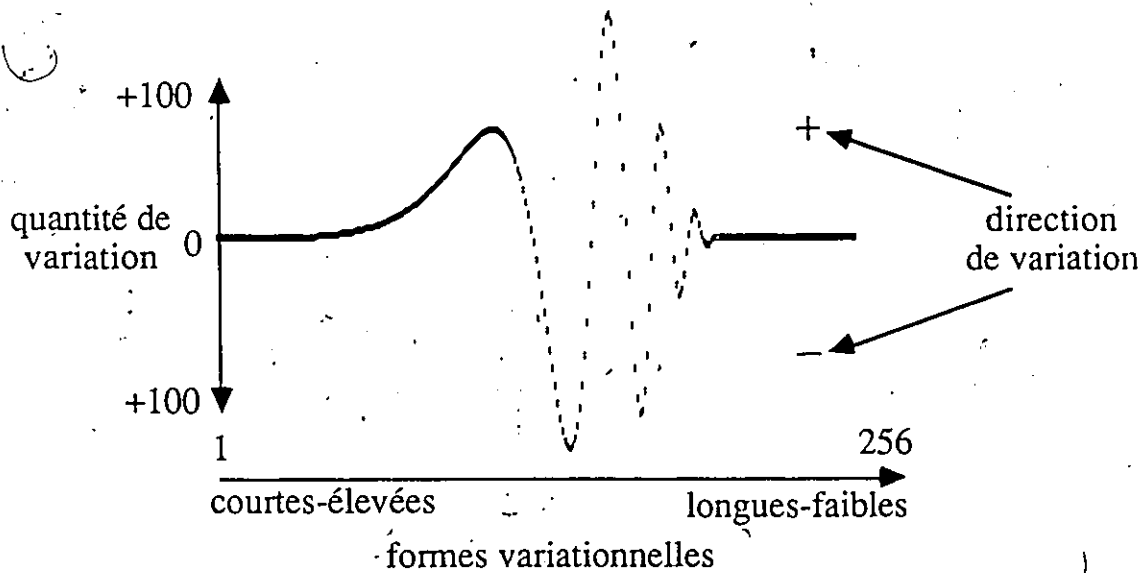


Fig. 18 La simulation computationnelle de la structure d'analyse spectrale ($F=8$; $A=900$)

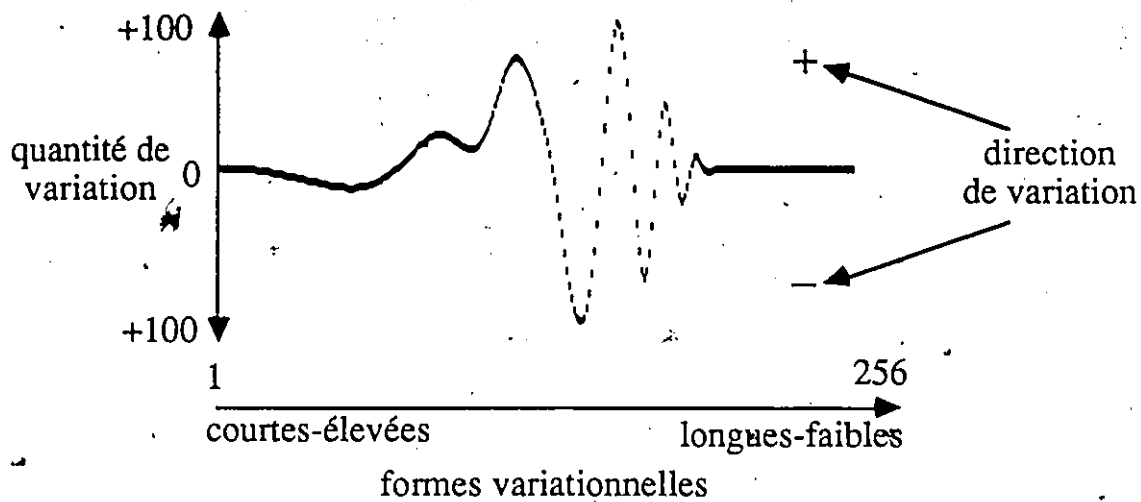


Fig. 19 La simulation computationnelle de la structure d'analyse spectrale (F= 32, A=25 et F=8, A=15)

Toutes ces figures représentent des "instantanés" ("snapshot") de l'histoire variationnelle de la position de la structure d'interface à un instant d'échantillonnage. Cette représentation synchronique peut être contrastée avec une représentation diachronique (à plusieurs instants d'échantillonnage) du processus de caractérisation de la structure d'analyse spectrale. A la figure 20, nous avons illustré cinq (5) réponses de la structure à la même stimulation que la figure 17, réponses qui sont espacées de deux moments d'échantillonnage¹⁴.

¹⁴ Nous avons choisi d'espacer les cinq réponses d'une distance temporelle de *deux* moments d'échantillonnage pour rendre plus apparent le changement diachronique des valeurs d'attribut des formes variationnelles.

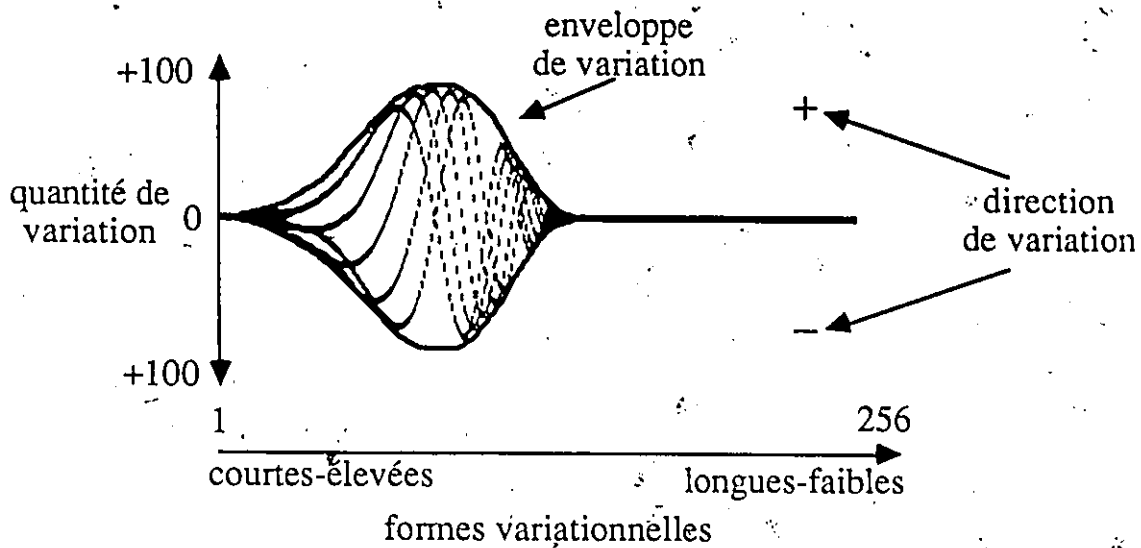


Fig. 20 La simulation computationnelle de la structure d'analyse spectrale ($F = 32$; $A = 30$)
Représentation diachronique de cinq (5) instants d'échantillonnage espacés de deux instants

Ce mode de représentation diachronique met en évidence l'un des aspects les plus intéressants de la simulation: l'existence d'ondes baladeuses (traveling waves) similaires à celles qui ont été observées sur la membrane basilaire de la cochlée (Bekesy, 1960; Rhode, 1970). Des crêtes et des creux se forment dans la partie gauche de la structure pour croître en intensité en se déplaçant vers la partie droite, atteindre un maximum et finalement décroître jusqu'à zéro. Dans le cas d'une stimulation sinusoïdale simple, ces crêtes et ces creux forment une *enveloppe de variation* globale qui se caractérise par une amplitude maximum., amplitude maximum dont l'emplacement sur l'axe des formes variationnelles (abscisse) varie en fonction de la fréquence de stimulation. Pour une même fréquence de stimulation, l'amplitude maximum

varie en hauteur (les quantités de variation positive et négative) si on fait varier l'amplitude de la stimulation. Dans le cas d'une stimulation harmonique à deux composantes sinusoïdales, on peut identifier deux séries de creux et de crêtes qui se chevauchent mutuellement et qui composent la réponse globale de la structure.

La simulation nous permet de choisir les constantes temporelles des différents groupement temporels. En faisant varier ces valeurs, nous pouvons obtenir différents "patterns" d'ondes baladeuses, mais tous ces patterns possèdent les caractéristiques que nous venons de décrire¹⁵. Chaque onde baladeuse effectue une sorte de "voyage" au travers des différentes bandes d'intervalles temporels et de vitesses. A la figure 20, nous pouvons identifier cinq ondes baladeuses qui voyagent à travers les bandes de vitesses et d'intervalles temporels. Il ne faudrait pas confondre ces cinq (5) ondes avec les cinq (5) courbes d'histoires variationnelles. Chacune de ces cinq (5) courbes est composée de trois (3) creux et de deux (2) crêtes. Ce sont ces creux et ces crêtes que nous associons aux cinq (5) ondes baladeuses. Les déplacements d'un creux ou d'une crête s'effectuent dans la portion positive ou dans la portion négative du plan directionnel. A la figure 20, deux (2) ondes voyagent dans la portion positive du plan directionnel (les deux crêtes de chaque courbe) et trois (3) ondes voyagent dans la partie négative de ce plan (les trois creux de chaque courbe). Chaque creux et chaque crête avance, d'instant en instant, dans la direction des bandes 'faible' et 'longue'. Au cours

¹⁵ Tous les exemples présentés ici utilisent les mêmes valeurs de constantes temporelles. Nous avons essayé de trouver une fonction qui correspond approximativement aux données expérimentales de la cochlée. De manière intéressante, cette fonction correspond à la dérivée d'une courbe normale.

de ce déplacement, les creux et les crêtes changent de hauteur et effectuent un parcours qui est défini par l'enveloppe de variation. Arrivées au terme de ce parcours, elles "meurent", c'est à dire que leur amplitude décroît jusqu'au niveau 0. La hauteur (cette hauteur peut être vers le bas) de chaque creux et de chaque crête est fonction de l'amplitude de la stimulation sinusoïdale: plus l'amplitude est élevée, plus la hauteur atteinte est importante. La distance parcourue sur l'axe des formes variationnelles est fonction de la fréquence de stimulation: plus cette fréquence est élevée, plus la distance parcourue est courte. Le déplacement de ces creux et de ces crêtes est plus particulièrement évident dans les séquences de positions à une seule composante sinusoïdale, mais on peut aussi constater ce phénomène dans les séquences de position à deux composantes sinusoïdales. Dans ce dernier cas, il est possible de détecter deux ensembles d'ondes voyageurs¹⁶ qui se différencient par la distance parcourue sur l'axe des formes variationnelles. Avec une séquence de positions composée de deux composantes sinusoïdales de fréquences 8 et 32 et d'amplitude identique, on observe, par exemple, qu'un ensemble de crêtes et de creux voyage jusqu'à une certaine valeur de forme variationnelle vers l'extrémité droite de la structure, alors qu'un second ensemble voyage beaucoup moins loin dans les bandes de vitesses et d'intervalles temporels: les ondes baladeuses de ce second ensemble "meurent" (leur amplitude décroît jusqu'au niveau 0) au niveau de formes variationnelles plus 'courtes-élevées' que les ondes baladeuses du premier ensemble.

¹⁶ Il est très difficile d'illustrer ces deux ensembles dans le mode de représentation visuel "statique" d'une figure.

On peut donc décrire le comportement diachronique global de la structure d'interface à partir de ce concept d'ondes baladeuses. Ces ondes se caractérisent par différents attributs: distance parcourue, hauteur atteinte, "durée de vie" etc. Ce qu'il faut réaliser, toutefois, c'est que ce concept d'onde baladeuse *n'appartient pas* au niveau de connaissance sonore de la structure d'analyse spectrale; il résulte d'une caractérisation "visuelle" que "nous" faisons du comportement de la structure d'analyse spectrale. Au niveau de la structure d'analyse spectrale, il n'y a que des formes variationnelles spécifiques à différentes bandes d'intervalles temporels et associées à différentes bandes de valeurs de vitesse. Au niveau de la structure d'analyse spectrale, il n'y a que des entités sonores de "deuxième niveau" avec leurs attributs de quantité et de direction de variation. Toute autre caractérisation, sous la forme "d'entités" comme les ondes baladeuses, ne peut être que le fait de niveaux supérieurs de caractérisation du processus de connaissance sonore.

Ceci complète cette sous-section sur la simulation de la structure d'interface. Cette simulation computationnelle représente une incorporation computationnelle des concepts qui ont été développés dans les trois premières sections de ce troisième chapitre. Nous croyons que cette simulation représente une conjecture intéressante du processus de caractérisation du second niveau du processus de connaissance sonore. Cette simulation représente également une source de réflexion et de "visualisation" intéressante (les ondes baladeuses) sur la nature du processus de connaissance sonore des niveaux supérieurs de caractérisation. Elle permet, enfin, une mesure précise des différentes valeurs d'attributs des formes variationnelles,

mesure précise qui pourra éventuellement être utilisée dans une simulation computationnelle des structures de caractérisation de ces niveaux supérieurs du processus de connaissance sonore.

Nous sommes maintenant rendus à la fin de la deuxième section, de ce troisième chapitre qui portait sur la nature du processus de connaissance sonore du second niveau de caractérisation de l'organisme. A la deuxième section de ce troisième chapitre, nous avons vu que le comportement de la structure d'interface pouvait être interprété comme un groupement temporel particulier de 'ce qui se manifeste' dans le milieu. Ce groupement temporel (à constante temporelle de 1) est équivalent à une mise en relation temporelle *invariante* des manifestations énergétiques qui atteignent l'organisme à travers le temps, mise en relation temporelle qui s'exprime par une première entité cognitive à un seul attribut de caractérisation: l'entité sonore du premier niveau et son attribut de quantité énergétique. A ce niveau du processus, tout 'ce qui se manifeste' dans le milieu est "connu" par l'organisme sous la forme d'une seule entité à un seul attribut.

Le rôle du second niveau du processus de connaissance sonore est de continuer ce processus de création d'invariants sonores en effectuant une série de groupements temporels "du deuxième type". Ces groupements temporels sont le fait des différents éléments structuraux de la structure d'analyse spectrale *et s'effectuent sur le produit du premier niveau de caractérisation (la position de la structure d'interface)*. La valeur d'attribut de l'entité sonore de premier niveau (la quantité énergétique) est transmise, par le biais de la position de la structure d'interface, à la structure d'analyse spectrale qui opère ses groupements temporels particuliers à partir de cette

valeur d'attribut. Le résultat de ces groupements s'expriment sous la forme de multiples entités sonores, les formes variationnelles.

Au *second* niveau du processus de connaissance sonore donc, l'organisme "connaît" 'ce qui se manifeste' dans le milieu sous la forme d'une multitude de formes variationnelles. Chaque forme variationnelle est le produit d'une mise en relation temporelle *invariante* des positions de la structure d'interface, et ces nouveaux invariants sonores que sont les formes variationnelles s'expriment, à chaque instant, par le biais de trois attributs:

- la bande d'intervalles temporels de chaque forme variationnelle ou la bande de vitesses à laquelle est sensible cette forme variationnelle,
- la quantité de variation de chaque forme variationnelle et,
- la direction de variation de chaque forme variationnelle.

Chez les organismes biologiques, le processus de connaissance sonore se poursuit aux niveaux supérieurs de caractérisation que sont les structures supérieures du système nerveux auditif. Quelles sont les mises en relation temporelles *invariantes* décrétées par ces structures? Quels sont les attributs de caractérisation de ces nouvelles entités? Quels sont les groupements temporels particuliers qui les génèrent? Autant de questions auxquelles se devrait de répondre une conjecture sur ces niveaux supérieurs de caractérisation. Nous n'en sommes malheureusement pas là. Dans la prochaine section, nous essaierons tout de même de donner un aperçu intuitif (non-formalisé) de notre conception du processus de caractérisation du troisième niveau de connaissance sonore.

3.4 Le troisième niveau du processus de connaissance sonore

Nous venons de voir que la structure d'analyse spectrale permet de produire un nouveau type d'entité sonore, la forme variationnelle, à partir d'un nouveau type de groupement temporel de ce qui est produit par la structure d'interface. Un organisme qui ne serait doté que d'une structure d'interface n'aurait pas accès aux formes variationnelles. Pour cet organisme, 'ce qui se manifeste' à chaque instant serait "connu" comme une seule entité dotée d'un seul attribut de caractérisation, car pour la structure d'interface, *il n'existe qu'un seul "objet" dans le milieu, et cet objet n'a, à chaque instant, qu'un seul attribut (la quantité énergétique)*. Un organisme doté d'une structure d'interface et d'une structure d'analyse spectrale accède, lui, à une nouvelle forme de connaissance. Pour un tel organisme, 'ce qui se manifeste' dans le milieu est, à la fois, l'entité sonore produite par la structure d'interface et la multitude d'entités sonores produites par la structure d'analyse spectrale, car pour cette dernière structure, *il existe plusieurs "objets" dans le milieu et ces objets ont, à chaque instant, trois attributs (forme "spécifique", quantité de variation, direction de variation)*. La connaissance sonore d'un organisme doté de ces deux premières structures de caractérisation est donc d'une double nature: une connaissance de niveau 'interface' (une seule entité à un seul attribut) et une connaissance de niveau 'spectral' (plusieurs entités à trois attributs). Nous avons vu que les organismes biologiques se constituaient d'autres structures de caractérisation qui se situent "au-dessus" de ces deux premières structures. Nous croyons que ces structures *supérieures* du système nerveux auditif des organismes

biologiques s'insèrent dans la *logique* du processus de connaissance des structures des deux premiers niveaux et qu'elles permettent la genèse de nouveaux types d'entités sonores (de nouveaux types "d'objets" sonores); donnant ainsi accès, à ces organismes, à plusieurs niveaux d'entités sonores. Nous voudrions maintenant présenter quelques hypothèses sur ce que pourraient être les entités sonores du *troisième* niveau de caractérisation de l'organisme.

Selon notre conception hiérarchique du processus de connaissance, les valeurs d'attribut des différentes formes variationnelles devraient représenter "l'entrée" du troisième niveau du processus de connaissance sonore. Ce devrait être à partir de ces valeurs que s'effectuent les groupements temporels caractéristiques de cette troisième structure de caractérisation. Ces groupements temporels "du troisième type" devraient également être de nouvelles 'mises en relation temporelles' *invariantes* de ce qui se manifeste à travers le temps et produire un nouveau type d'entités sonores dotées de nouveaux attributs de caractérisation. Que pourraient être ces nouvelles entités? Nous avons vu, à la section précédente, que les différentes formes variationnelles pouvaient être associées à la contribution de différentes bandes de vitesses; la quantité de variation d'une forme variationnelle correspondant à la présence plus ou moins importante d'une bande de valeurs de vitesse dans l'histoire variationnelle de chaque position de la structure d'interface. On peut donc considérer que ce qui est "passé" à la troisième structure de caractérisation représente différentes valeurs de vitesse. Nous avons vu, par ailleurs, que le comportement global de la structure d'analyse spectrale se caractérise par l'existence d'ondes baladeuses

et nous avons assimilé ces ondes baladeuses aux crêtes et aux creux qui "voyagent" à travers les différentes formes variationnelles. Dans le cas d'une stimulation sinusoïdale à une seule composante (son pur), nous pouvons dénombrer¹⁷ cinq (5) ou six (6) crêtes et creux qui se propagent à travers les formes variationnelles. Rappelons que ces ondes baladeuses se situent dans le plan positif *ou* dans le plan négatif de variation et que ces ondes changent de hauteur (quantité de variation) en suivant le parcours de l'enveloppe de variation. Or ces crêtes et ces creux qui voyagent à travers les différentes formes variationnelles sont en fait *des crêtes et des creux qui voyagent à travers différentes bandes de vitesses. Leur déplacement reflète des changements de vitesses: c'est à dire des accélérations ou des décélérations.*

Nous croyons que le troisième niveau du processus de connaissance sonore effectue des groupements temporels qui nous permettent de caractériser les différentes valeurs d'attribut des formes variationnelles en termes d'accélérations-décélérations. Prenons l'exemple d'une onde baladeuse en forme de crête. Cette crête se déplace vers la droite de l'axe des formes variationnelles. Or un déplacement vers la droite représente un déplacement vers des bandes de valeurs de vitesse *plus lentes*. Dans la partie *ascendante* de son parcours, la progression d'une crête reflète donc une contribution peu importante des bandes de valeurs de vitesse élevées et une contribution de

¹⁷ Le nombre de crêtes et de creux augmente si nous augmentons le nombre de formes variationnelles. Le nombre de formes variationnelles est assimilable à une forme de "résolution" de la structure d'analyse spectrale. Nous avons fait des "tests" avec une simulation de cinq cents douze (512) formes variationnelles et le nombre de crêtes et de creux augmentait à neuf (9) ou dix (10) pour la même situation de stimulation où nous n'identifions que cinq (5) ou six (6) creux pour une simulation de deux cents cinquante-six (256) formes.

plus en plus importante des bandes de valeurs de vitesse plus faibles, contribution qui atteint son maximum à une bande de valeurs de vitesse donnée. Une augmentation de la contribution des bandes de valeurs de vitesse plus faibles représente une *accélération négative (décélération)*, c'est à dire *une diminution de la vitesse*. Si l'on dit d'une voiture que son déplacement (sa variation) se caractérise par la présence *de plus en plus importante* de valeurs de vitesse *de plus en plus faibles*, c'est que cette voiture *décélère*. Dans la deuxième moitié de son parcours, cette crête reflète une contribution de moins en moins importante des bandes de valeurs de vitesse de plus en plus faibles. Or une diminution de la contribution des bandes de vitesses plus faibles peut être interprétée comme une *accélération*, c'est à dire une *augmentation de vitesse*. Si l'on dit d'une voiture que son déplacement se caractérise par la présence *de moins en moins importante* de valeurs de vitesses *de plus en plus faibles*, c'est que cette voiture *accélère*. Le déplacement d'une crête est donc caractéristique d'une décélération puis d'une accélération. Nous pouvons interpréter de manière similaire le déplacement d'un creux: une décélération suivie d'une accélération. Ce qui différencie les accélérations-décélérations d'un creux des accélérations-décélérations d'une crête, c'est la direction de variation des formes variationnelles dont se constituent ces accélérations ou décélérations. Une accélération-décélération positive se constitue des changements dans la quantité de variation associée à des formes variationnelles de direction positive alors qu'une accélération-décélération négative se constitue des changements dans la quantité de variation associée à des formes variationnelles de direction négative. On peut se représenter visuellement ces accélérations-décélérations positives et négatives à partir de la représentation

géométrique de fonctions paraboliques. A la figure 21, nous avons illustré deux fonctions paraboliques: $y = x^2$ et $y = -(x^2)$.

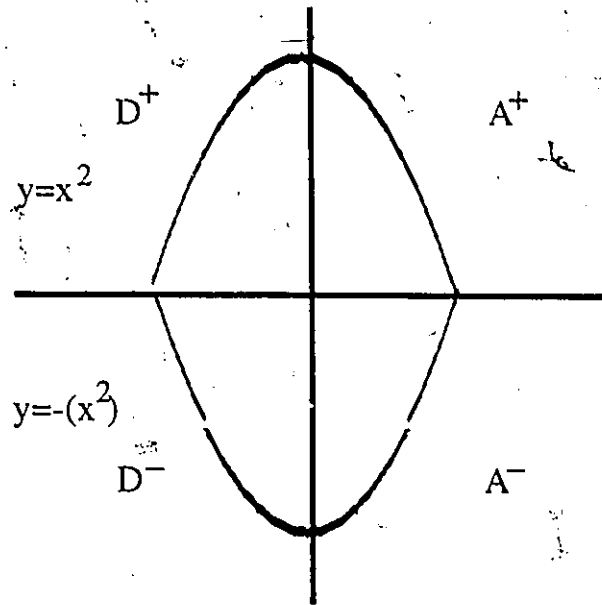


Fig. 21 Décélérations et accélérations positives et négatives

Les portions de courbes des quatre (4) quadrants du plan peuvent être associées aux accélérations-décélérations positives et négatives. Le quadrant D^+ représente une décélération positive; on peut l'associer à une crête ascendante. Le quadrant D^- représente une décélération négative; on peut

l'associer à un creux ascendant¹⁸. Le quadrant A⁺ représente une accélération positive; on peut l'associer à une crête descendante. Finalement, le quadrant A⁻ représente une accélération négative; on peut l'associer à un creux descendant. Les fonctions paraboliques entretiennent d'ailleurs une relation intéressante avec les concepts de variation, de vitesse et d'accélération. La dérivée d'une fonction parabolique (qui est une fonction du second degré) est égale à une ligne droite (fonction du premier degré). Cela signifie que la variation de la variation des valeurs d'une fonction parabolique s'exprime par une ligne droite. Une accélération (ou une décélération) représente justement une mesure de la variation d'une variation (la vitesse). On pourrait donc considérer les fonctions paraboliques comme l'expression la plus simple (accélération linéaire) d'une accélération ou d'une décélération.

Nous croyons que les entités sonores de ce troisième niveau de caractérisation peuvent être représentées par de telles fonctions paraboliques. Chaque entité sonore de ce niveau se constituerait d'un couple décélération-accélération. Il y aurait donc deux directions¹⁹ de "formes accélérationnelles"²⁰: la direction 'crête' et la direction 'creux'. La direction

¹⁸ Rappelons, encore une fois, qu'un creux "ascendant" descend "visuellement", mais qu'un tel creux représente, en fait, une ascension dans les valeurs de quantité de variation des formes variationnelles de direction négative dont il se constitue.

¹⁹ A ne pas confondre avec la direction de variation des formes variationnelles.

²⁰ Nous désignerons ces entités sous le terme général de "formes accélérationnelles" mais il faut comprendre que l'attribut "accélérationnel" regroupe aussi bien les accélérations que les décélérations inhérentes à chaque forme "accélérationnelle".

'crête' serait formée d'un couple ($D^+ \rightarrow A^+$) et la direction 'creux' d'un couple ($D^- \rightarrow A^-$). Quels pourraient être les autres attributs de caractérisation de ces formes "accélérationnelles"? Nous avons tenté de les illustrer aux figures 22, 23 et 24. Une forme "accélérationnelle" comme un creux ou une crête parcourt une certaine distance sur l'axe des formes variationnelles: les creux et les crêtes "naissent", "s'élèvent", "redescendent" et "meurent" sur une certaine distance. Cette distance constituerait un autre attribut de ces formes "accélérationnelles". A la figure 22, nous avons illustré quelques formes "accélérationnelles" de direction 'crête' dotées de valeurs d'attributs de 'distance' différentes.

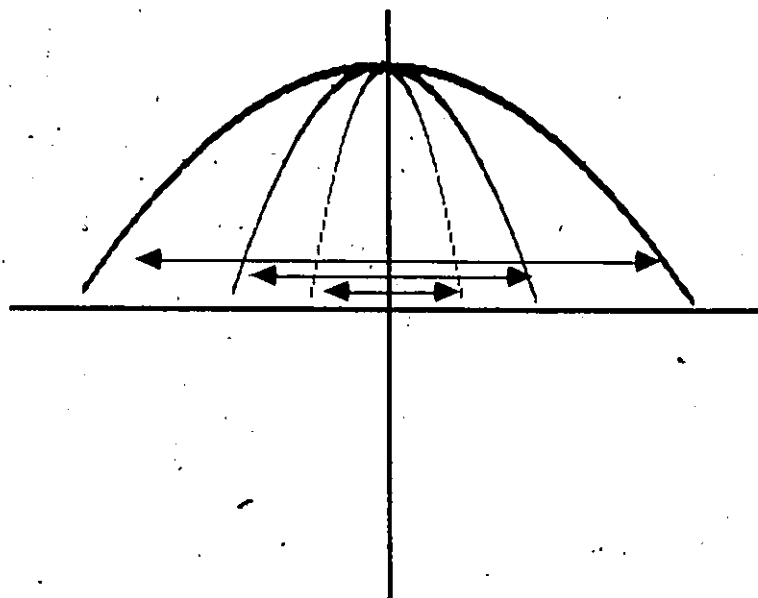


Fig. 22 Trois formes "accélérationnelles"
à valeurs de 'distance' différentes; direction 'crête'

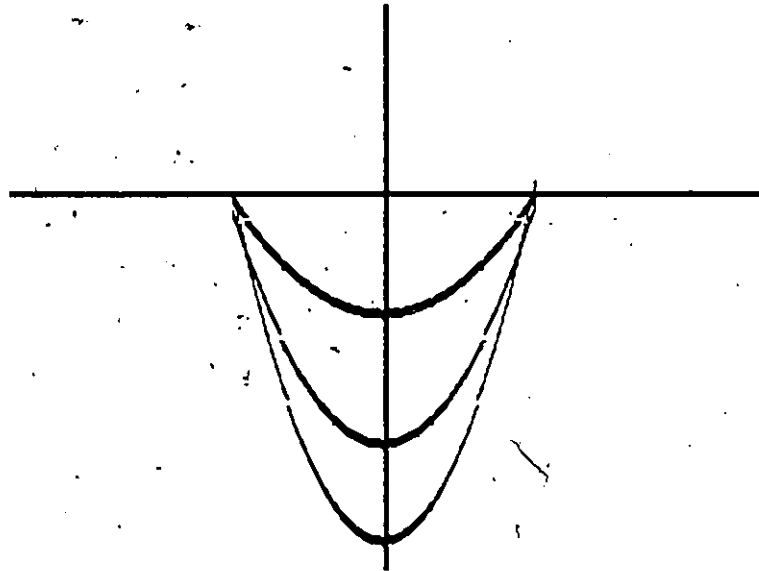


Fig. 23. Trois formes "accélérationnelles" à valeurs 'd'intensité' différentes; direction 'creux'

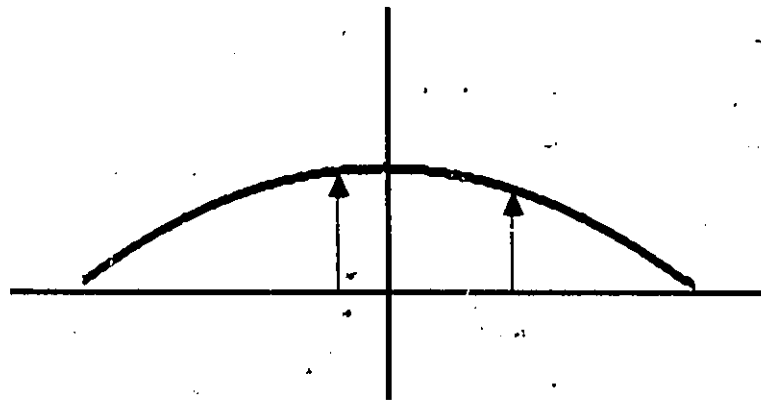


Fig. 24 Deux formes "accélérationnelles" à valeurs d'attributs de 'direction', 'd'intensité' et de 'distance' identiques mais à deux valeurs de 'position' différentes

Au cours de leur "parcours de vie", ces creux et ces crêtes atteignent différents maximums de hauteur. L'intensité (la hauteur atteinte) d'une forme "accélérationnelle" pourrait représenter un autre attribut de caractérisation. A la figure 23, nous avons illustré quelques formes "accélérationnelles" de direction 'creux' dotées de la même valeur d'attribut 'distance' mais de valeurs d'attribut 'intensité' différentes. Finalement, la position d'un 'creux' ou d'une 'crête' à un instant de son "parcours de vie" pourrait représenter le dernier attribut de caractérisation de ces formes "accélérationnelles". A la figure 24, nous avons illustré deux formes "accélérationnelles" de valeurs identiques sur les attributs de distance, de direction et d'intensité mais dotées chacune d'une valeur de position différente.

Les formes "accélérationnelles" seraient donc dotées de quatre attributs de caractérisation:

- une 'direction' "accélérationnelle" de valeur 'creux' ou 'crête',
- une 'distance' "accélérationnelle" de plusieurs valeurs possibles,
- une 'intensité' "accélérationnelle" de plusieurs valeurs possibles.
- une 'position' "accélérationnelle" de plusieurs valeurs possibles.

Ces entités sonores du troisième niveau de caractérisation seraient le fruit d'un nouveau type de groupement temporel qui prendrait les valeurs d'attributs des entités sonores du *deuxième* niveau de caractérisation (les formes variationnelles) et décrèterait sur ces valeurs un nouvel ensemble de 'mises en relation temporelles' *invariantes*. Au troisième niveau du processus de connaissance sonore donc, l'organisme connaîtrait les sons sous la forme d'entités "accélérationnelles" dotées de ces quatre (4) attributs de

caractérisation. Ces entités de troisième niveau seraient vraisemblablement, à leur tour, l'objet de groupements temporels de la structure de caractérisation du *quatrième* niveau du processus de connaissance sonore, structure qui produirait à son tour de nouvelles entités sonores qui seraient l'objet de groupements temporels d'une *cinquième* structure de caractérisation etc..

Ceci complète notre présentation intuitive du troisième niveau de caractérisation de l'organisme. Nous croyons que les hypothèses présentées dans cette section constituent un bon point de départ pour une recherche axée sur une expression plus formalisée (et computationnelle) des groupements temporels "du troisième type" et c'est dans ce sens que seront orientés nos travaux de recherche futurs. Nous venons de présenter notre conception des trois premiers niveaux de caractérisation de l'organisme. Nous allons maintenant tenter de faire une synthèse de la logique sous-jacente à l'ensemble du processus de connaissance sonore de l'organisme.

3.5 Vers une logique du processus de connaissance sonore

Au deuxième chapitre de ce travail, nous avons tenté de définir la logique *de l'ensemble* du phénomène de la connaissance sonore en tissant un certain nombre de relations entre le milieu, l'organisme et l'expérience sonore. Nous avons vu que le milieu est occupé par différents objets physiques qui réagissent aux manifestations énergétiques en "émettant" à leur tour des manifestations énergétiques "modulées" par leurs caractéristiques particulières. La fonction de l'organisme est de procéder à la connaissance (caractérisation) de ces manifestations énergétiques émanant des différents objets physiques environnants. Nous avons vu que ce processus s'effectue par le biais d'une hiérarchie de structures de caractérisation produisant des entités sonores à chaque niveau de la hiérarchie. Parmi les entités sonores qui sont produites aux différents niveaux, certaines se manifestent au foyer expérientiel et constituent notre expérience des sons. Cette conception de la logique d'ensemble du phénomène de la connaissance sonore nous a amené à conclure, à la fin du second chapitre, que l'essentiel du caractère problématique "résiduel" de ce phénomène se trouvait dans le processus de connaissance sonore de l'organisme, et dans ce troisième chapitre, nous avons étudié les trois premiers niveaux de ce processus afin d'essayer de définir les premiers éléments d'une logique *spécifique* au processus de connaissance sonore de l'organisme.

Quels sont les éléments de cette logique? Nous avons vu que la structure d'interface interagit directement avec les manifestations énergétiques qui

l'atteigne. Cette structure effectue un groupement temporel particulier de 'ce (l'énergie) qui se manifeste' à travers le temps et exprime le résultat de ce groupement sous la forme d'une seule entité dotée d'un seul attribut de caractérisation: la quantité énergétique. La structure d'analyse spectrale, quant à elle, interagit avec le produit de la caractérisation de la structure d'interface. Elle effectue une série de groupements temporels particuliers de 'ce (la quantité énergétique) qui se manifeste' à travers le temps et exprime le résultat de ces groupements sous la formes de plusieurs entités dotées de trois attributs de caractérisation:

- la spécificité de chaque forme variationnelle, i.e. les bandes de vitesses et d'intervalles temporels associées à chacune de ces formes variationnelles,
- une quantité de variation associée à chaque forme variationnelle, et,
- une direction de variation associée à chaque forme variationnelle.

La structure "accélérationnelle", quant à elle, interagit avec le produit de la caractérisation de la structure d'analyse spectrale. Elle effectue probablement une série de groupements temporels particuliers de 'ce (valeurs d'attributs des formes variationnelles) qui se manifeste' à travers le temps et exprime le résultat de ces groupements sous la forme de plusieurs entités "accélérationnelles" dotées de quatre attributs de caractérisation:

- une distance "accélérationnelle",
- une intensité "accélérationnelle",
- une position "accélérationnelle", et,
- une direction "accélérationnelle".

Les entités sonores de ces trois premiers niveaux de caractérisation peuvent être considérées comme *trois* types de formes sonores. Nous pouvons les associer à des gestalts, i.e. à des "bonnes formes" résultant des

processus organisateurs (les groupements temporels) de chaque niveau de caractérisation. Le mécanisme de genèse de ces trois niveaux de formes sonores (entités, gestalts) est similaire. Une structure de caractérisation d'un niveau donné reçoit 'ce qui se manifeste' (ou ce qui est produit) au niveau inférieur, applique un type de groupement temporel sur ces manifestations et produit en retour une ou plusieurs entités sonores. Les entités produites par cette structure de caractérisation sont à leur tour l'objet de groupements temporels de la structure suivante de caractérisation et ce processus se répète pour tous les niveaux de caractérisation de l'organisme. Nous pouvons supposer que les organismes biologiques les plus évolués (comme l'organisme humain) procèdent à la genèse d'un certain nombre d'autres formes sonores. L'entité sonore de la structure d'interface est la forme sonore la plus primitive. Les entités sonores de la structure d'analyse spectrale sont des formes plus "évoluées" qui se constituent de certaines 'mises en relation temporelles' des valeurs d'attribut de la forme sonore élémentaire de la structure d'interface et les entités "accélérationnelles" sont des formes sonores encore plus "évoluées" qui se constituent de certaines 'mises en relation temporelles' des valeurs d'attribut des formes variationnelles de la structure d'analyse spectrale. Les entités sonores de ces trois premiers niveaux de caractérisation ne se manifestent vraisemblablement pas au niveau du foyer expérientiel de l'organisme humain. Ce sont les entités résultantes des niveaux plus élevés de caractérisation qui constituent notre univers expérientiel sonore. Ces entités sont les notes, les bruits, les phrases musicales et parlées etc.

Nous croyons que les entités "tones"(ou les notes) apparaissent au quatrième niveau de caractérisation et qu'elles se constituent de groupements temporels sur les formes "accélérationnelles". Nous avons procédé à quelques "manipulations" de fonctions paraboliques et nous nous sommes aperçus qu'il était possible de générer des fonctions quasi-sinusoïdales en "groupant" des formes "accélérationnelles" de direction 'crête' et 'creux'. Si ces hypothèses (éventuellement formalisées et "computationnalisées") s'avéraient fondées, il s'agirait là d'une première théorie computationnelle des processus de caractérisation sous-jacents aux sons élémentaires (les "éléments" de Tenney et Polansky (1980) et les "tones" qui ont fait l'objet d'une multitude de recherches psychophysiques). Ces entités du quatrième niveau de caractérisation possèderaient un certain nombre d'attributs correspondant aux attributs "psychologiques" comme la hauteur, l'intensité, la durée et le timbre. Ces entités pourraient, à leur tour, être groupées sous la forme de séquences de notes ("clangs") en décrétant des 'mises en relation temporelles' invariantes de ce (les notes successives) qui se manifeste au niveau inférieur de caractérisation.

A chaque instant d'échantillonnage donc, l'organisme procède à la genèse d'une hiérarchie de formes sonores dotées d'un certain nombre d'attributs de caractérisation et, parmi ces formes sonores, certaines sont sélectionnées par l'organisme (processus attentionnel) et sont acheminées au foyer expérientiel. Les formes sélectionnées deviennent alors les figures sonores qui occupent le champ expérientiel. La figure 25 illustre ces différentes caractéristiques du processus de connaissance sonore pour les cinq (5) premiers niveaux et pour trois (3) instants d'échantillonnage.

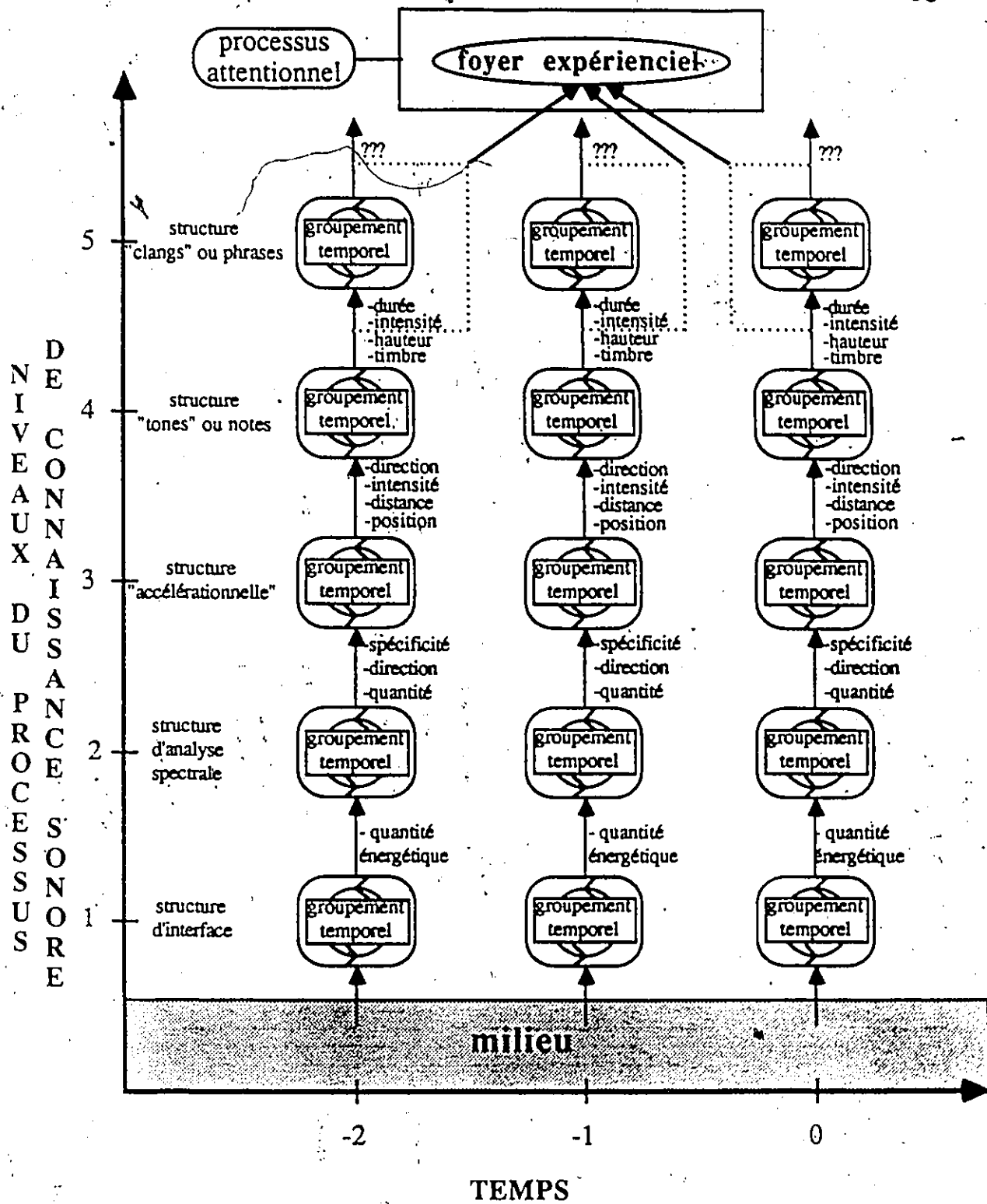


Fig. 25 Le processus de connaissance sonore de l'organisme

La fonction de l'organisme est de procéder à la connaissance du milieu. Ce processus s'effectue en plusieurs étapes et produit une hiérarchie de "connaissances" sous la forme des entités sonores de chaque niveau de caractérisation. Quelle est la 'raison d'être' de ces multiples niveaux? Qu'est-ce que l'organisme "gagne" à générer toute cette hiérarchie d'entités? On peut avoir une réponse "adaptative" à ces multiples questions et penser que cette organisation hiérarchique des structures de caractérisation (et les nouveaux invariants qui sont le produit de chacun de ces niveaux) est l'organisation qui a permis aux organismes biologiques actuels de s'adapter le plus efficacement à leur environnement. En d'autres termes, la logique du processus de connaissance sonore "fonctionne": les entités (formes) sonores que génère l'organisme sont le point de départ de courants interactionnels orientés *hors* du foyer expérientiel et ces courants interactionnels produisent en grande majorité les effets désirés (*répondre* au téléphone qui "sonne", *changer de vitesse* lorsque le régime du moteur devient "aigu" etc.), "confirmant" (ne réfutant point) de ce fait la "véracité" de la conjecture cognitive sonore. Les entités sonores sont, pour la plupart, hautement *non-problématiques*; elles proviennent d'un processus de connaissance *équilibré*.

Le nombre de niveaux de caractérisation d'un organisme pourrait être considéré comme une mesure de son degré d'évolution. Plus le nombre de niveaux augmente, plus l'organisme dispose de formes sonores "évoluées" qui lui permettent d'agir efficacement sur le substrat expérientiel. Cela est particulièrement évident dans l'univers sonore musical. L'entraînement à l'écoute musicale nous permet, par exemple, de développer notre "oreille"

(nos niveaux supérieurs de caractérisation) et de saisir des formes sonores que nous ne percevions pas auparavant. Les grands compositeurs sont probablement capables d'effectuer des 'mises en relation temporelles' à des niveaux hiérarchiques très élevés, et d'agencer à leur guise les formes sonores des niveaux "inférieurs" pour qu'elles produisent les formes sonores "supérieures" désirées. L'univers musical représente un "terrain de jeu" de prédilection pour celui qui désire explorer les "capacités" de groupements temporels de l'organisme humain. Mais ce n'est pas que dans cette activité "ludique" que se réalise le processus de connaissance sonore. La compréhension du langage parlé reflète également l'existence des formes sonores évoluées (mots, phrases, etc.), bien que dans ce cas, l'interaction avec des entités plus abstraites (concepts) soit probablement déterminante dans la sélection et la genèse de ces formes.

On pourrait donc comparer les entités sonores des niveaux successifs de caractérisation à des conjectures de plus en plus "évoluées" sur 'ce qui se manifeste' dans le milieu. Mais comment pourrions-nous cerner plus précisément le caractère évolutif accru de chacune de ces entités? Qu'est-ce qui distingue, en d'autres termes, l'entité sonore de la structure d'interface d'une entité sonore du quatrième niveau de caractérisation, par exemple? Nous avons vu que l'entité sonore de la structure d'interface nous donnait accès à la quantité d'énergie qui se manifeste à *chaque* instant d'échantillonnage. Nous pouvons assimiler cette première forme à une forme sonore *ponctuelle*. A chaque instant, le produit cognitif de ce premier niveau de caractérisation permet la connaissance d'un *point* temporel, et c'est la nature particulière du groupement temporel sous-jacent à la structure

d'interface qui produit ce *type* de forme sonore. Nous avons vu, par la suite, que les entités sonores de la structure d'analyse spectrale reflétait différentes 'mises en relation temporelles' (groupements) de ces points temporels que sont les entités sonores du premier niveau. Nous pouvons assimiler les formes sonores de ce deuxième niveau de caractérisation à des *lignes* sonores, ces lignes recouvrant différents espaces temporels (les bandes d'intervalles temporels) et étant de différentes orientations selon les valeurs de quantité et de direction de variation de chaque forme variationnelle. Les lignes sonores (formes variationnelles) se constituent d'un rapport "topologique" temporel entre les points temporels que sont les entités sonores successives de la structure d'interface, comme une ligne "visuelle" se constitue de certains rapports topologiques spatiaux entre différents points spatiaux. Nous pouvons considérer les formes "accélérationnelles" dans la même optique: elles représentent différentes "courbes" sonores qui se constituent d'un certain rapport topologique temporel entre les différentes lignes sonores. Il en serait de même pour les entités sonores du quatrième niveau (les "tones" ou notes) qui se constitueraient de rapports topologiques temporels entre les courbes "accélérationnelles" et qui formeraient des "courbes" plus complexes (à deux points d'inflexion comme une sinusoïdale). Plus nous progressons dans la hiérarchie des niveaux de caractérisation, plus les différentes entités sonores reflètent des formes sonores complexes, formes qui existent à *chaque instant d'échantillonnage*. C'est à cette "complexité" croissante que nous pouvons associer le caractère évolutif accru des entités sonores de chacun des niveaux successifs. En se donnant accès à des formes de plus en plus complexes (point, ligne, courbe du 1^{er} degré, courbe du 2^{ème} degré),

l'organisme augmente sa capacité d'adaptation et dispose d'une connaissance sonore de plus en plus sophistiquée du milieu dans lequel il est plongé.

Il est intéressant finalement de noter que l'organisation hiérarchique du processus de connaissance sonore de l'organisme semble être le fait des autres processus de connaissance de l'organisme. Cela est plus particulièrement évident dans le processus de connaissance visuelle où certaines structures du cortex visuel peuvent être assimilées à des structures de caractérisation de la linéarité visuelle (Hubel & Wiesel, 1979; Lamontagne & Beausoleil, 1982). Lamontagne & Beausoleil (1984) suggèrent d'ailleurs que cette organisation hiérarchique du processus de connaissance visuelle est une caractéristique de toutes les autres modalités perceptuelles (audition, olfaction, "proprioception", goût et perception tactile). Ils avancent également l'hypothèse intéressante que la "connaissance" motrice relève d'une organisation hiérarchique, mais d'une organisation hiérarchique où la "direction" du courant interactionnel est orientée des entités motrices les plus évoluées (les entités motrices 'de haut niveau') vers les entités motrices les moins évoluées (les entités motrices de 'bas niveau': les muscles ou effecteurs moteurs). Dans cette perspective, la logique du processus de connaissance sonore devient maintenant une facette particulière d'une logique *globale* des processus de connaissance perceptuels et moteurs, logique qui pourrait même inclure la logique des processus de connaissance conceptuelle.

connaissance sonore, que celui-ci soit le fait de différents médiums, de différents organismes (biologiques ou artificiels) ou de différentes entités sonores (mesures physiques ou attributs psychologiques). Cela nous a permis également de mettre en évidence la nature hiérarchique du processus de connaissance sonore et de "jeter un pont" entre la hiérarchie structurale et la hiérarchie expérientielle. Cette conception de la logique sous-jacente à l'ensemble du phénomène de connaissance sonore nous a amené à situer le caractère problématique résiduel de ce phénomène vers le processus de connaissance sonore de l'organisme. Le troisième et dernier chapitre a été consacré à l'élaboration d'une *logique spécifique à ce processus*. Nous y avons défini un certain nombre de concepts qui permettent là encore, croyons-nous, de saisir les isomorphismes existant entre les différents niveaux de caractérisation de l'organisme. Ces concepts ont donné lieu à une simulation computationnelle des deux premiers niveaux de caractérisation et ouvrent la voie à des tentatives computationnelles de simulation des niveaux supérieurs de caractérisation.

Dès le début de notre travail de recherche, nous avons été "frappés" par le contraste important entre la richesse de notre expérience des sons et la pauvreté relative de notre connaissance des processus cognitifs qui permettent cette connaissance. C'est pourquoi nous avons tenté continuellement d'orienter nos réflexions vers l'élaboration de concepts qui nous permettraient éventuellement d'aller plus loin que le niveau de caractérisation de la structure d'analyse spectrale. Pour ce faire, nous avons essayé, en premier lieu, de comprendre la logique des deux premiers niveaux de caractérisation, puis de généraliser cette logique à l'ensemble des niveaux

de caractérisation des organismes biologiques. Au terme de ce travail, la logique sous-jacente au processus de connaissance sonore nous apparaît comme *un* aspect de la logique sous-jacente à l'ensemble des processus de connaissance. La connaissance sonore est une "forme" de connaissance orientée vers la caractérisation *temporelle* de 'ce qui se manifeste' dans le milieu. Cette dimension temporelle est aussi présente dans les autres modalités de connaissance, mais elle "atteint" une absoluité inégalable dans le phénomène de la connaissance sonore. Le temps est omniprésent dans l'étude de la *genèse* des sons et il rend très difficile le travail de visualisation et de conceptualisation des processus qui permettent cette genèse. Mais si le temps rend le problème de la connaissance sonore plus ardu, il lui confère également une beauté formelle qui n'a jamais cessé de nous fasciner et qui nous motivera encore longtemps à tenter de faire évoluer cette problématique.

BIBLIOGRAPHIE

Beauchamp, K.G. & Yuen, K. (1979). *Digital methods for signal analysis*. London: G. Allen and Unwin.

Bekesy, G.V. (1960). *Experiments in hearing*. New York: Mac Graw Hill.

Bregman, A.S. & Campbell, J. (1971). Primary auditory stream segregation and perception of order in rapid sequence of tones. *Journal of experimental psychology*, 89, 244-249.

Colburn, H.S. & Durlach, N.I. (1978). Binaural phenomena. *Handbook of perception*, (Vol. 4: Hearing, pp. 365-455), London: Academic Press.

Deutsch, D. (1978). The psychology of music. *Handbook of perception*, (vol 10, pp. 191-223), London: Academic Press.

Divenyi, P.L. & Hirsch, I.J. (1974). Identification of temporal order in three-tone sequences. *Journal of the Acoustical Society of America*, 52, 166, (appendix).

Dodwell, P.C. (1970). *Visual pattern recognition*. New York: Holt, Rinehart and Winston.

Duifhuis, H. (1972). *Perceptual analysis of sound*. Unpublished Ph.D. dissertation, Eindhoven Tech. Highschool.

Ellis, W.D., (1938). *A source book of Gestalt Psychology*. London: Kegan Paul, Trench, Trubner and Cie.

Erickson, R. (1975). *Sound structure in music*. Berkeley: University of California Press.

Evans, E.F. (1971). Central mechanisms relevant to the neural analysis of simple and complex sounds. In O.J. Grusser and R. Klisáké (eds.) *Pattern recognition in biological and technical systems*. Heidelberg: Springer-Verlag.

Feldkeller, R. & Zwicker, E. (1956). *Das Ohr als Nachrichtenempfänger*. Stuttgart: Hirzel.

Flanagan, J.L. (1972). *Speech analysis, synthesis and perception*. Berlin: Springer-Verlag.

Fletcher, H. (1929). *Speech and hearing*. New York: Van Nostrand.

Fourier, J.B.J. (1822). *La théorie analytique de la chaleur*. Paris: Didot.

Gabor, D. (1946). Theory of communication. *Journal of Instrumental Electrical Engineers*, 93, 429-457.

Gacek, R.R. (1972). Neuroanatomy of the auditory system. In *Foundations of modern auditory theory*. (vol 2, pp. 241-262) ed. J.V. Tobias.

Goldstein, J.L. (1973). An optimum processor theory for the central formation of the pitch of complex tones. *Journal of the Acoustical Society of America*, 54, 6, 1496-1516.

Green, D.M. & Swets, J.A. (1966). *Signal detection theory and psychophysics*. New-York: Wiley.

Helmholtz, H.L.F von. (1885). *On the sensations of tone*. New York: Dover. (reprinted 1954).

Houtsma, A.S.M. & Goldstein, J.L. (1972). The central origin of the pitch of complex tones: Evidences from musical interval recognition. *Journal of the Acoustical Society of America*, 51, 520-529.

Hubel, D.H. & Wiesel, T.N. (1979). Brain mechanisms of vision. *Scientific American*, 241(3), 150-162.

Köhler, W. (1947). *Gestalt psychology*. Mentor Book, London: The New English Library.

Lamontagne, C. & Beausoleil, J.R. (1982). Achieving visual spatiality: towards a psychologically relevant, physiologically plausible and computationally efficient conjecture. *Cognition and Brain Theory*, 5, 343-365.

Lamontagne, C. & Beausoleil, J.R. (1984). Vers une représentation conceptuelle de la représentation préconceptuelle. *Communication Information*, 43, 169-194.

Leshowitz, B. (1978). Measurement of the auditory stimulus. *Handbook of perception*, (volume 4: Hearing, pp. 83-122), London: Academic Press.

Masterton, B. & Diamond, I.T., (1978). Hearing: Central neural mechanisms. *Handbook of perception*, (volume 3: Biology of perceptual systems, pp. 407-447), London: Academic Press.

Nordmark, J.O. (1978). Frequency and periodicity analysis. *Handbook of perception*, (volume 4: Hearing, pp. 243-282), London: Academic Press.

Ohm, G.S. (1843). Ueber die Definition des Tones, nebst daran geknüpfter Theorie der sirene und ähnlicher tonbildener Vorrichtungen. *Annalen der Physik und Chemie*, 135, 497-565.

Piaget, J., (1967). *Biologie et connaissance*. Paris: Gallimard.

Pisoni, D.B. (1985). Speech perception: Some new directions in research and theory. *Journal of the Acoustical Society of America*, 78, 381-388.

Plomp, R. (1971). Old and new data on tone perception. In *Contributions to sensory physiology*. ed. W.D. Neff, (Vol.5), New York: Academic Press.

Plomp, R. (1976). *Aspects of tone sensation*. New York: Academic Press.

Plomp, R. (1970). Timbre as a multidimensional attribute of complex tones. In R. Plomp and G.F. Smoorenburg (Eds.), *Frequency analysis and periodicity detection in hearing*. Leiden: Sijthoff.

Popper, K.R. (1978). *An intellectual autobiography: Unended quest*. Glasgow: Fontana-Collins.

Reed, S. K. (1978). Pattern processing. *Handbook of perception*, (volume 9: Perceptual processing), London: Academic Press.

Rhode, W.S. (1970). Observations on the vibration of the basilar membrane in squirrel monkey using the Mössbauer technique. *Journal of the Acoustical Society of America*, 49, 1218-1231.

Risset, J.C. (1978). Musical acoustics. *Handbook of perception*, (volume 4: Hearing, pp. 521-564), London: Academic Press.

Sanders, D.A. (1977). *Auditory perception of speech: an introduction to Principles and Problems*. Englewood Cliffs: Prentice-Hall.

Sauveur, J. (1701). Système général des intervalles du son. *Mémoires de l'Académie Royale des Sciences*.

Schafer, R.W. & Rabiner, L.R. (1974). Parametric representations of speech. In R. Reddy (ed.) *Speech recognition*. New York: Academic Press, pp. 100-148.

Shouten, J.F. (1940). The perception of pitch. *Philips Technical Review*, 5, 286-294(b).

Steele, C.R. (1973). Cochlear mechanics. *Handbook of sensory physiology*. (volume 5: The auditory system) W.D. Keidel and W.D. Neff (eds.), Springer-Verlag: Berlin.

Stevens, S.S. & Davis, H. (1936). Psychophysiological acoustics: Pitch and loudness. *Journal of the Acoustical Society of America*, 8, 1-15.

Tenney, J. & Polansky, L. (1980). Temporal gestalt perception in music. *Journal of music theory*, 24, 205-241.

Tonnard, J. (1962). Time/Frequency analysis along the partition of cochlear models: a modified place concept. *Journal of the Acoustical Society of America*, 34, 1337-1351.

Van Noorden, L.P.A. S. (1975). *Temporal coherence in the perception of tone sequence*. Unpublished doctoral thesis. Technische Hogeschool, Eindhoven, Holland.

Van Nostrand's Scientific Encyclopedia. (1983). New York: Van Nostrand Reinhold Company.

Weiss, T.F. (1966). A model of the peripheral auditory system. *Kybernetik*, 3, heft 4, 153-175.

Whitfield, I.C. (1978). The neural code. *Handbook of perception*, (volume 4: Hearing, pp. 163-183), London: Academic Press.

Wightman, F.L. (1973). The pattern-transformation model of pitch. *Journal of the Acoustical Society of America*, 54, 2, 407-416.

Zwicker, E. & Scharf, B. (1965). A model of loudness summation. *Psychological Review*, 72, 1, 3-26.

ANNEXE 1

**LA SIMULATION
DE LA STRUCTURE D'ANALYSE SPECTRALE**

La simulation computationnelle de la structure d'analyse spectrale a été écrite en langage C (Aztec Manx version 1.06G pour ordinateur MacIntosh). Nous présentons ici une version de la simulation où nous calculons les valeurs d'attributs de deux cents cinquante-six (256) formes variationnelles. Le programme est en caractères emphasés. Les commentaires sur les différentes sections du programme sont en caractères italiques non-emphasés. Nous avons restreint nos commentaires aux aspects "cognitifs" de la simulation et supposons une connaissance du langage C.

*Simulation computationnelle
de la structure d'analyse spectrale
(256 formes variationnelles)*

Fichiers externes: déclaration de constantes et de variables

Fichiers externes:-

```
#include <quickdraw.h>
#include <setjmp.h>
#include <memory.h>
#include <math.h>
#include <event.h>
#include <osutil.h>
```

Déclaration de constantes:

```
#define PI 3.1415926535897932
#define GrHdl &(gp->portRect)
```

Déclaration de variables:

```
int k,w,j,i,s,d,l,z,y;
double c,x;
int f1,a1,a2,f2;
int param[4];
char fr[5],amp[5];
int FVOUT[256], FVOUTMEM[256];
double FVTOTAL[256], FV1[256], PRSIPRE[256],
PRSIPAS[256];
double CONSTEMP[256];
```

Routine principale (Main)

```
main(argc,argv)
int argc;
char *argv[];
```

```
{
GrafPtr gp;
char *p;
```

Prise des paramètres:

Il est possible, dans cette version de la simulation, de produire des séquences harmoniques de positions de la structure d'interface. Ces séquences harmoniques peuvent avoir deux composantes sinusoïdales. Lorsque nous appelons le programme, nous devons lui donner quatre paramètres: les deux premiers paramètres sont la fréquence et l'amplitude de la première composante; les deux derniers paramètres sont la fréquence et l'amplitude de la seconde composante.

```
if(argc!=5){printf("liste d'arguments incomplete\n"); exit();}
w=0;
while(--argc)
{
p(++argv)[0];
cparg(p,argc);
```

```

while(*p)
{
    if(*p>'9' || *p<'0')
    {
        printf("valeurs entieres SVP\n");
        exit();
    }
    param[w]=10*param[w]+*p-'0';
    ++p;
}
++w;
}

```

f1=param[0]; Paramètre de fréquence de la première composante.
a1=param[1]; Paramètre d'amplitude de la première composante.
f2=param[2]; Paramètre de fréquence de la seconde composante.
a2=param[3]; Paramètre d'amplitude de la seconde composante.

Opération d'initialisation de l'écran vidéo:

```

GetPort(&gp);
OpenPort(gp);
EraseRect(GrHdl);

```

Boucle de calcul des constantes temporelles des deux cents cinquante-six (256) formes variationnelles:

Dans cette version de la simulation, nous utilisons, pour calculer la valeur des constantes temporelles, une fonction de calcul qui approxime la dérivée d'une courbe normale. C'est cette fonction qui nous semble produire le "comportement" spectral le plus rapproché du comportement cochléaire. La variable CONSTEMP est un vecteur de deux-cents cinquante-six (256) valeurs.

```

for (k=1;k<129;++k)
{
    c=k;
    CONSTEMP[k]=(c*c)/32768;
}
for (k=129;k<257;++k)
{
    CONSTEMP[k]=CONSTEMP[257-k];
}

```

```
CONSTEMP[k]=(-CONSTEMP[k])+1;
```

Boucle principale de la routine principale:

C'est dans cette boucle que sont calculées les valeurs d'attribut des différentes formes variationnelles. Si la simulation n'est pas arrêtée en cours de route (en appuyant sur le bouton de la "souris"), elle va continuer durant 512 moments d'échantillonnage (valeur de l'indice w).

```
for (w=1;w<513;++w)
```

```
{
  if(Button()) exit();
```

Calcul de la position de la structure d'interface:

La variable PRSIPRE est un vecteur contenant les deux cents cinquante-six positions résiduelles de la structure d'interface. La première position "résiduelle" (PRSIPRE[1]) représente la position originale de la structure d'interface, celle qui "inclut" toutes les formes de variation.

```
PRSIPRE[1]=((a1)*sin(2.0*PI*w*f1/512.0)+(a2)*sin(2.0*PI*w*f2/512.0));
```

Boucle de calcul du premier terme de droite de l'équation:

La variable FV1 est un vecteur contenant les deux cents cinquante-six (256) termes de droite des deux cents cinquante-six (256) équations de groupement temporels. Ce calcul s'effectue en multipliant les valeurs d'attribut de l'instant précédent (FVTOTAL) par la valeur de chaque constante temporelle. La variable FVTOTAL est un vecteur contenant les valeurs d'attributs des deux cents cinquante-six (256) formes variationnelles. A cette étape-ci du programme, ce vecteur contient les valeurs de l'instant précédent.

```
for (k=1;k<257;++k)
```

```
  FV1[k]=FVTOTAL[k]*CONSTEMP[k];
```

Boucle de calcul des nouvelles valeurs d'attribut des formes variationnelles:

Ces nouvelles valeurs vont remplacer les valeurs de l'instant précédent dans la variable FVTOTAL. Ces nouvelles valeurs sont égales au premier terme de droite (FV1) plus le deuxième terme de droite (CONSTEMP[k]*(PRSIPRE[k] - PRSIPAS[k])). La variable PRSIPRE[k] est le vecteur contenant les positions résiduelles présentes de la structure d'interface, alors que la variable PRSIPAS[k] contient les positions résiduelles passées de la structure d'interface. Ces nouvelles valeurs sont calculées en multipliant la constante temporelle d'une forme variationnelle par la variation entre les positions résiduelles passées et présentes de chaque "niveau" de forme variationnelle. La première opération de la boucle consiste à calculer une première nouvelle valeur de FVTOTAL. A la deuxième opération de la boucle, cette première "nouvelle" valeur est soustraite de la position résiduelle de la structure d'interface de ce "niveau" de forme variationnelle pour trouver la valeur de position résiduelle présente de la forme variationnelle du "niveau" inférieur. La troisième opération de cette boucle consiste à transférer les nouvelles valeurs de FVTOTAL dans une variable différente FVOUT qui nous permet d'afficher à l'écran toutes les valeurs de FVTOTAL sous la forme des exemples de la section 3.2.2.

```
for (k=1;k<257;++k)
{
  FVTOTAL[k]=FV1[k]+(CONSTEMP[k]*(PRSIPRE[k]-
  PRSIPAS[k]));
  PRSIPRE[k+1]=PRSIPRE[k]-FVTOTAL[k];
  FVOUT[k]=FVTOTAL[k];
}
```

Boucle de transfert "temporel" de la position résiduelle:

Cette boucle sert à "transférer" la position résiduelle présente de tous les "niveaux" de formes variationnelles à la position "passée" qui sera utilisée dans le prochain calcul des valeurs d'attributs (celui de l'instant suivant).

```
for (k=1;k<257;++k)
  PRSIPAS[k]=PRSIPRE[k];
```

Appel de la fonction d'affichage à l'écran:

```
affich(256,20,170,FVOUT,FVOUTMEM);
```

Boucle de transfert de la variable VOUT:

Cette boucle de transfert permet à la fonction d'affichage de s'effectuer plus rapidement.

```
for(k=1; k<257; ++k)
  FVOUTMEM[k]=FVOUT[k];
```

```
}
}
```

Fin de la routine principale.

Routine d'affichage

```
affich(l,x,y,v,vm)
int *v,*vm;
int l,y,x;
{
  int k;
  for(k=1; k<l; ++k)
  {
    if(v[k]==vm[k])
    {
      MoveTo((x+k-1),y-v[k]);
      LineTo((x+k-1),(y+2)-v[k]);
    }
    else
    {
      PenPat(white);
      MoveTo((x+k-1),y-vm[k]);
      LineTo((x+k-1),(y+2)-vm[k]);
      PenPat(black);
      MoveTo((x+k-1),y-v[k]);
      LineTo((x+k-1),(y+2)-v[k]);
    }
  }
}
```

Routine nécessaire à la prise de paramètres

```
cparg(p1,ct)
char *p1;
int ct;
{
    unsigned char no;
    char *strg,*strmem;

    if(ct==4) |
        strg=fr;
    else if(ct==3)
        strg=amp;
    else return;

    strmem=strg;
    ++strg;
    no=0;

    while(*strg++=*p1++)
        ++no;
    *strmem=no;
}
```

Fin du programme.
