



**HAL**  
open science

## Perception de la hauteur chromatique chez le nourrisson.

Françoise Armand

► **To cite this version:**

Françoise Armand. Perception de la hauteur chromatique chez le nourrisson.. Psychologie. Université René Descartes - Paris V, 1984. Français. NNT: . tel-00338825

**HAL Id: tel-00338825**

**<https://theses.hal.science/tel-00338825>**

Submitted on 14 Nov 2008

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université de Paris V (René DESCARTES) - U.E.R. de Psychologie

**PERCEPTION DE LA HAUTEUR  
CHROMATIQUE CHEZ LE NOURRISSON**

Thèse présentée en vue de l'obtention du Doctorat de  
Troisième Cycle par Françoise ARMAND.

Sous la direction d'Eliane VURPILLOT, Professeur à Paris V.

Paris 1984

C'est au professeur Eliane VURPILLOT que vont mes premiers remerciements, car c'est elle qui a permis à cette thèse de voir le jour.

J'adresse aussi mes remerciements à Laurent DEMANY pour ses conseils et son aide tout au long de ce travail.

Le Laboratoire de Psychologie Expérimentale, que dirigeait le Professeur Georges NOIZET, m'a accueillie et m'a fourni les moyens de réaliser cette thèse. Je remercie en particulier Anne Yvonne JACQUET, Christine PIERANDREI et Marie-Claire BOTTE.

Jean-Pierre ARMAND, Fabrice ARNAUD, Didier RONCIN et plus spécialement Didier DEMANY ont contribué à résoudre un certain nombre de problèmes techniques, je leur en suis très reconnaissante.

Je tiens à remercier également Mr ORTEGA, Directeur du Conservatoire de Musique de Pantin, ainsi que Mr SEQUERA, pour nous avoir permis de réaliser au mieux les expériences avec des musiciens du Conservatoire.

Remerciements aussi au Docteur HAZEMANN, à Mlle TRUFFY et à l'ensemble du personnel du Centre P.M.I. Anatole France de Choisy-le-Roi (Val de Marne) où ont eu lieu toutes les expériences sur le bébé.

Merci au professeur Marc BLANCHETEAU pour la confiance qu'il m'a témoignée au début de ce travail.

## TABLE DES MATIERES

INTRODUCTION .....	1
1. Hauteur d'un son .....	2
1.1. Hauteur des sons périodiques .....	2
1.2. Hauteur chromatique et hauteur tonale .....	3
2. Etude de l'équivalence chromatique .....	5
2.1. Sujets tout venant .....	5
2.2. Sujets tout venant sélectionnés .....	16
2.3. Sujets musiciens .....	20
3. Extraction de la hauteur d'un son .....	24
3.1. Hauteur d'un son pur .....	24
3.2. Hauteur d'un son complexe .....	27
3.2.1. Modèle du résidu .....	27
3.2.2. Modèle de reconnaissance de patterns .....	27
4. Explications possibles de l'équivalence chromatique .....	32
4.1. L'équivalence chromatique et le modèle de Terhardt ....	32
4.2. L'équivalence chromatique et le modèle temporel d'Ohgushi .....	34
4.3. L'octave subjective .....	36
4.4. L'équivalence chromatique et le phénomène de l'ac- culturation .....	38
 PROBLEMATIQUE ET TECHNIQUE EXPERIMENTALE	
1. Problématique .....	42
2. Choix des sujets .....	43
3. Choix de la méthode .....	43
4. Choix de la technique expérimentale/choix de la réponse ....	44

5. Appareillage - Procédure (expériences 1, 2 et 3) .....	45
6. Déroulement de l'expérience .....	48
7. Conditions d'élimination des sujets .....	50
8. Stimulus .....	51
8.1. Choix de l'intensité .....	52
8.2. Choix de la fréquence .....	54
8.3. Seuils différentiels en fréquence .....	54
8.4. Discrimination de sons séparés par une ou plusieurs octaves .....	56

## EXPERIMENTATION

1. Expérience 1 .....	57
1.1. Problématique .....	57
1.2. Pré-expérience 1 .....	58
1.2.1. Plan expérimental .....	58
1.2.2. Les stimulus .....	58
1.2.3. Sujets .....	60
1.3. Résultats .....	62
1.3.1. Pré-expérience 1 .....	62
1.3.2. Expérience 1 .....	64
2. Expérience 2 .....	67
2.1. Les stimulus .....	67
2.2. Procédure .....	70
2.3. Sujets .....	71
2.4. Résultats de l'expérience 2 .....	71
3. Expérience 3 .....	74
3.1. Les stimulus .....	74
3.2. Les sujets .....	76
3.3. Résultats .....	76
4. Expérience 4 .....	79
4.1. Les stimulus .....	80
4.2. Méthode .....	80
4.3. Appareillage et déroulement de l'expérience .....	80

4.4. Procédure .....	82
Première partie .....	82
Deuxième partie .....	84
Troisième partie .....	86
Quatrième partie .....	86
4.5. Sujets .....	87
4.6. Résultats .....	87
4.6.1. Résultat de la première partie .....	
4.6.2. Résultat de la deuxième partie .....	91
4.6.3. Résultat de la troisième partie .....	91
4.7. Discussion .....	93
DISCUSSION - CONCLUSION .....	94
BIBLIOGRAPHIE .....	99

## INTRODUCTION

La hauteur d'un son peut se définir comme le caractère grave ou aigu du son : plus la fréquence est élevée, plus le son est aigu, plus la fréquence est basse, plus le son paraît grave.

La hauteur est la sensation sonore liée à la fréquence du son. Elle se définit selon deux dimensions : la hauteur tonale - fonction monotone de la fréquence, et la hauteur chromatique qui est la qualité des sons séparés par un intervalle d'octave.

D'un point de vue musical, deux notes séparées par une octave sont considérées comme fonctionnellement équivalentes.

D'un point de vue perceptif, on sait que ces deux notes sont perçues comme similaires : c'est ce qu'on appelle l'équivalence chromatique.

Percevoir la hauteur chromatique des sons, c'est-à-dire percevoir une similarité entre 2 sons séparés par une octave, est une capacité que l'on trouve très fréquemment chez les musiciens, par contre la plupart des sujets non musiciens, capables "d'entendre" une différence de hauteur entre 2 notes séparées par un 1/2 ton, ne perçoivent pas de similarité chromatique.

Certains auteurs, ayant étudié cette question, ont avancé des théories explicatives sur l'origine de ce phénomène. Le présent travail se situe directement dans ce courant de recherche, en posant la question suivante : peut-on mettre en évidence l'équivalence chromatique chez des sujets très jeunes (3 mois) n'ayant qu'une expérience auditive réduite.

## I - HAUTEUR D'UN SON

L'échelle musicale utilisée dans les pays occidentaux est construite à partir d'un intervalle de base : l'octave.

En effet, chaque note se répète à une octave d'intervalle, comme ceci :

... la<sub>1</sub> si<sub>1</sub> do<sub>2</sub> ré<sub>2</sub> mi<sub>2</sub> fa<sub>2</sub> sol<sub>2</sub> la<sub>2</sub> si<sub>2</sub> do<sub>3</sub> ré<sub>3</sub>...  
  
 une octave

Ces deux notes (do<sub>2</sub> = do<sub>3</sub>) ont des hauteurs tonales très différentes, donc elles sont parfaitement discriminées, mais elles sont équivalentes pour un musicien.

On sait, en effet, qu'en musique, si on parle d'une mélodie dont la tonalité est en do majeur, elle se terminera obligatoirement par un do, mais n'importe quel do.

### 1.1. Hauteur des sons périodiques

Outre les bruits (sons non périodiques) qui n'ont pas de hauteur bien définie, il existe deux types de sons périodiques dans notre environnement sonore, les sons purs et les sons complexes.

purs

Les sons sont composés d'une seule onde sinusoïdale. Ils sont peu familiers à l'oreille humaine, puisqu'ils n'existent pratiquement pas à l'état naturel. La fréquence de l'onde sonore du son pur donne sa hauteur au son. Par contre, les sons complexes, qui nous sont plus connus (instruments de musique, parole, cris d'animaux ...) sont la somme de plusieurs sons purs dont les fréquences se superposent. Cette onde sonore plus complexe que celle d'un son pur, peut se décomposer en une série de sinusoides correspondant chacune à une "harmonique" du complexe.

La décomposition d'un complexe se fait de la façon suivante (analyse de FOURRIER) :

si  $f_0$  : fréquence de la première harmonique ou fondamentale,  
la fréquence  $f_n$  de chaque harmonique supérieure est égale à :

$$f_n = n \times f_0 \quad \text{Unité : le Hertz (Hz)}$$

Le rapport de fréquence entre deux harmoniques successives est donc toujours le même, quelque soit la fréquence de la fondamentale.

On voit d'ores et déjà que la hauteur d'un son complexe ne peut pas se déterminer de la même façon que la hauteur d'un son pur.

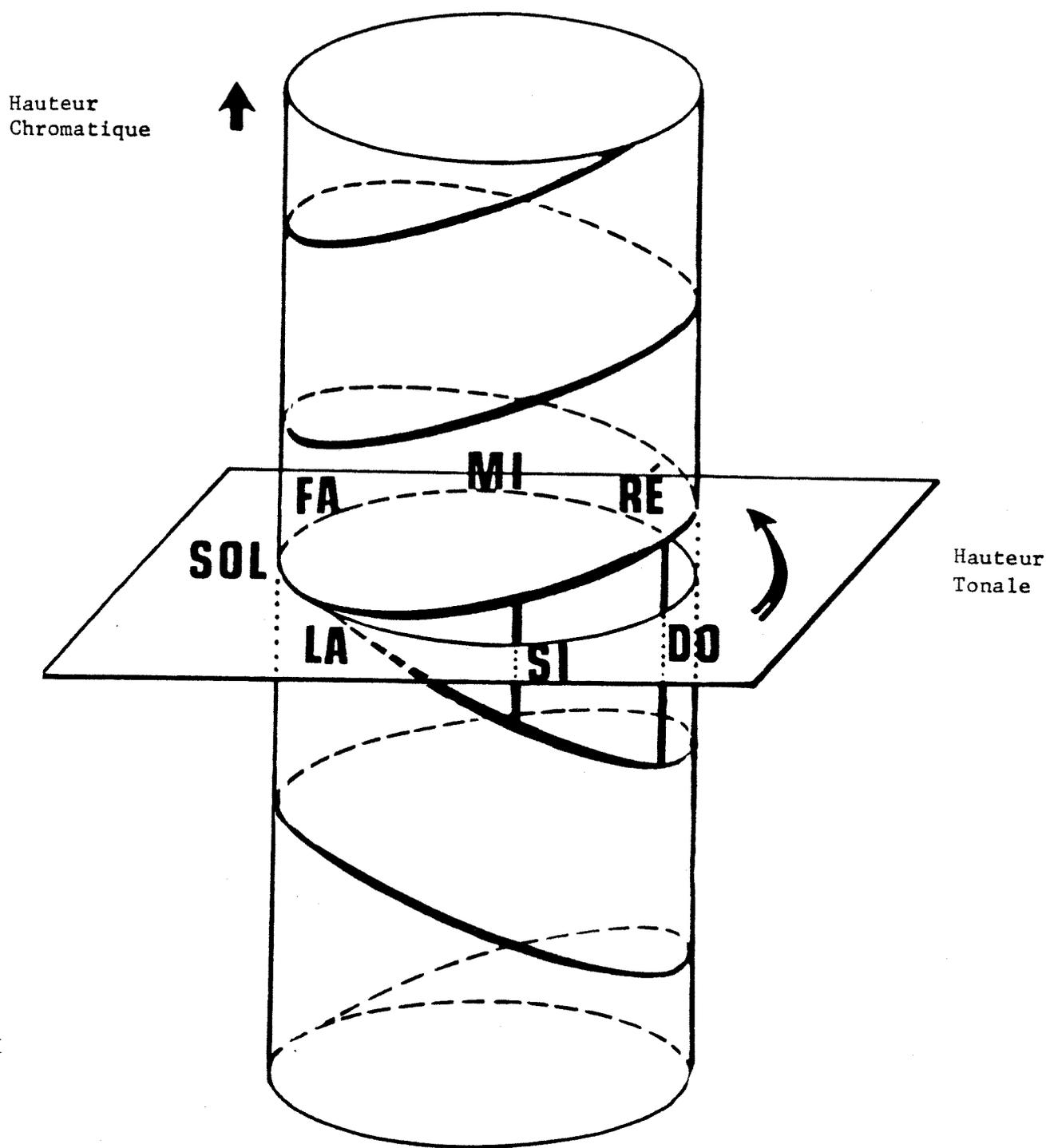
### 1.2. Hauteur chromatique et hauteur tonale

Pour les sons périodiques, la hauteur est déterminée à partir de la (ou des) fréquence(s) composant le son. On la représente selon deux dimensions : la hauteur tonale et la hauteur chromatique (schéma page suivante).

Il semblerait que ce soit BRENTANO qui, le premier, ait établi le concept de hauteur définie selon deux composantes qu'il a appelées : clarté et qualité. C'est en 1890 que STUMPF a rapporté les idées communiquées oralement par BRENTANO. La qualité est ce qu'on a défini comme la hauteur chromatique, c'est-à-dire la ressemblance perçue entre deux sons à l'octave; c'est l'élément chromatique de la hauteur.

BACHEM (1937) a introduit les expressions de hauteur tonale (tone height) et hauteur chromatique (tone chroma) pour désigner respectivement le composant unidimensionnel et le composant cyclique de la hauteur. Il a été l'un des premiers à étudier la relation perceptive entre des sons à l'octave appelée généralement équivalence chromatique ou similarité chromatique.

TEPLOV (1966), contestant le terme d'identité entre des sons à l'octave employé par BRENTANO, suggère le terme de similitude; cette distinction marque bien que deux sons à l'octave ne sont pas identiques, mais plutôt équivalents ou ressemblants.



Représentation bidimensionnelle de la hauteur  
D'après RISSET (1978)

## 2 - ETUDE DE L'EQUIVALENCE CHROMATIQUE

L'ensemble des expériences faites sur ce sujet peut se diviser en 3 groupes, suivant le niveau d'éducation musicale des sujets testés : sujets "tout venant", sujets "tout venant", ayant subi un test de sélection, et sujets musiciens.

### 2.1. Sujets "tout venant"

Ces sujets sont recrutés au hasard. Parmi eux, on trouve aussi bien des sujets sans aucune expérience musicale, que des sujets musiciens.

### 2.2. Sujets "tout venant" sélectionnés

Ce sont aussi des sujets "tout venant", mais que l'on a sélectionnés avant l'expérience par un test qui permet de conserver les sujets les plus capables de passer cette expérience.

### 2.3. Sujets musiciens

Ces sujets ont, selon les expériences, entre 3 et 10 ans d'expérience musicale.

### 2.1. Sujets "tout venant"

Une des expériences les plus anciennes a été faite sur des animaux, des rats blancs. BLACKWELL et SCHLOSBERG (1943) ont utilisé une méthode de conditionnement opérant d'évitement. On entraîne les rats à courir le long d'un couloir en leur présentant des sons de 5 fréquences différentes (3, 5, 7, 8 et 10 Khz). La réponse (course) est renforcée négativement par un choc électrique, pour une fréquence de 10 Khz.

Les auteurs notent alors une extinction de la réponse pour les fréquences de 3, 7 et 8 Khz, alors que pour la fréquence de 5 Khz, l'extinction se fait plus lentement (la réponse se maintenant bien entendu pour la fréquence renforcée). Or, 5Khz est la fréquence de la sub-octave du son renforcé de 10 Khz. Les auteurs concluent que la réponse s'est généralisée pour les sons séparés par une octave du son renforcé. Cela voudrait dire que des animaux peuvent être sensibles au rapport de similarité entre les notes à l'octave. Cette expérience est souvent citée comme la preuve que l'équivalence chromatique est innée, mais on l'a aussi beaucoup

critiquée, en particulier à propos des mesures de distorsion du son qui n'ont été faites que sur les signaux électriques et non sur les signaux acoustiques perçus par les rats. De plus, les fréquences utilisées sont presque toutes supérieures à 5 Khz, or on sait (WARD, 1954) qu'à partir de cette fréquence, l'équivalence chromatique n'est plus perçue.

On peut simplement regretter que cette expérience n'ait jamais été refaite. En effet, la plupart des expériences récentes utilisent des sujets humains adultes et sont basées sur des méthodes totalement différentes : comme par exemple, la reconnaissance de mélodies. Cette méthode est employée dans les trois expériences suivantes.

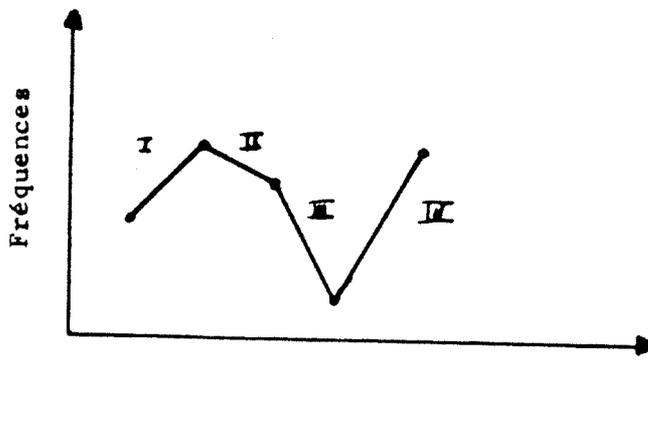
Voyons d'abord une expérience de DEUTSCH (1972). Une mélodie composée des 28 premières notes d'une chanson connue ("Yankee Doodle"), est transformée suivant 5 modalités :

- a) la mélodie originale est jouée dans une certaine tonalité (en  $fa_2$ ).
- b) la mélodie est jouée dans la même tonalité qu'en a, c'est-à-dire dans la tonalité de fa, mais une octave plus haut (en  $fa_3$ ).
- c) la mélodie est jouée cette fois une octave en-dessous de  $fa_2$  (condition a), c'est-à-dire en  $fa_1$ .
- d) la mélodie est jouée en  $fa_2$ , comme dans la condition a, mais cette fois, chaque note est transposée indépendamment à une octave choisie parmi les 3 octaves employées dans les conditions a, b, et c. Dans cette dernière condition, l'information chromatique est préservée, c'est-à-dire que chaque note conserve sa tonalité d'origine (un fa reste un fa, un do un do, mais à une ou plusieurs octaves près).
- e) la séquence mélodique composée jusqu'ici de 28 sons purs devient une série de "clic" (bruit bref sans hauteur bien définie). Cela revient à dire que seule l'information rythmique est conservée, l'information donnée par la hauteur des notes est totalement supprimée.

Les 51 sujets pris pour cette expérience entendent tous les 5 transformations a, b, c, d et e. A chaque fois, ils doivent dire le titre de la mélodie qu'ils ont reconnue.

Les résultats montrent que les mélodies transposées selon les modalités a, b, c sont parfaitement reconnues; il n'y a aucune erreur.

Cela semble normal, puisque dans ces trois cas, toute la mélodie est transposée. Par contre, DEUTSCH note beaucoup d'erreurs dans les deux dernières conditions : d, 17,65% d'erreurs - e, 18,75% d'erreurs, et aussi beaucoup de non réponses : 70,59% pour d, 62,5% pour e. Les mauvaises performances obtenues par les sujets dans la condition d font conclure à l'auteur que la préservation de l'information chromatique ne suffit pas pour reconnaître la mélodie originale. DEUTSCH remarque enfin que le nombre d'erreurs et de non-réponses est pratiquement équivalent pour les conditions d et e; elle en conclut que le fait de préserver l'information chromatique indépendamment pour chaque note, n'améliore pas plus la reconnaissance des mélodies que lorsqu'on conserve uniquement l'information rythmique (condition e). Cela tendrait à prouver qu'aucune similarité chromatique n'a été perçue par les sujets. Il faut nuancer cette conclusion, en précisant que dans la condition d, le contour de la mélodie est modifié. Par contour, nous voulons parler de la succession du sens des intervalles. Prenons comme exemple la mélodie ci-dessous; l'intervalle I est ascendant, II et III sont descendants et IV ascendant :



Transposer la mélodie, comme dans les conditions b et c de l'expérience précédente, revient à déplacer entièrement la mélodie à un degré différent de l'échelle musicale, ce qui implique 2 choses : premièrement la taille des intervalles reste inchangée, deuxièmement le contour de la mélodie transposée n'est pas modifié. Par contre, si l'on transpose chaque note indépendamment à une ou plusieurs octaves de la note originale, on modifie obligatoirement la taille des intervalles, et éventuellement le contour de la mélodie. Il est donc nécessaire de bien contrôler le

contour des mélodies transformées car une modification de ce genre entraîne des difficultés supplémentaires au moment de la reconnaissance des mélodies; c'est ce qu'ont fait IDSON et MASSARO (1978) et KALLMAN et MASSARO (1979).

Chez les premiers, la procédure est la suivante : 5 mélodies composées de sons purs constituent les stimulus de base. On vérifie d'abord que ces mélodies sont bien connues des sujets, puis on fait entendre à chacun une des 5 mélodies, soit sous sa forme originale, soit sous une forme modifiée. Il y a 4 transformations possibles :

- OVC (Octave Violated Contour) : chaque note est transposée indépendamment à une ou plusieurs octaves de la note originale. Le contour est modifié.
- OPC (Octave Preserved Contour) : chaque note est transformée comme dans la condition OVC, mais ici, le contour est préservé.
- PC (Preserved Contour) : chaque note est transposée indépendamment des autres, avec la contrainte que toutes les notes de la mélodie soient comprises dans un seul intervalle d'octave. Dans ce cas, le contour est préservé, alors que la fréquence de chaque note est modifiée ainsi que la taille relative des intervalles successifs.
- LT (Linear Transformation) : Transformation linéaire du contour c'est-à-dire que la taille de chaque intervalle est réduite de moitié. Ici, le contour est encore préservé, mais cette fois, la taille relative des intervalles successifs n'est pas modifiée. Dans ces deux dernières conditions, PC et LT, la hauteur tonale et la hauteur chromatique de toutes les notes sont changées.
- O : dans cette condition, la mélodie est présentée dans sa forme originale, sans aucune transformation.

On fait entendre au sujet une des 5 mélodies, selon une des 5 modalités que nous venons de voir (OVC - OPC - PC - LT - O); il doit ensuite choisir parmi la liste des 5 titres qu'on lui propose, le titre qu'il pense avoir reconnu. Il est donc placé dans une tâche de choix forcé. Les résultats recueillis sur 36 sujets sont donnés sous forme de pourcentage de bonnes réponses pour chacune des 6 conditions expérimentales : O : 80%, OPC : 78%, LT : 60%, PC : 55%, OVC : 42% . Remarquons tout d'abord l'importance de l'effet des modifications du contour mélodique : OPC : 78% contre OVC : 42% (ce qui fait une diminution de près de la moitié

des bonnes réponses). Examinons à présent le reste des résultats. Les mauvaises performances obtenues dans la condition OVC (42% d'identifications correctes de la mélodie) permettent de penser que l'information chromatique, si elle est préservée, n'arrive pas à compenser l'effet perturbant des distorsions du contour. Par contre, dans la condition LT (60%) où le contour de la mélodie originale est maintenue, mais pas l'information chromatique, le nombre de bonnes réponses est malgré tout moins important que pour OPC (78%) où là, l'information chromatique n'a pas été supprimée. Dans la condition OPC, la hauteur chromatique permet une meilleure reconnaissance des mélodies distordues. Enfin, il faut noter la différence entre les résultats de la condition PC et ceux de la condition LT. Dans les deux cas, le contour mélodique reste inchangé par rapport à l'original, mais la transformation opérée dans la condition LT (où la taille relative des intervalles successifs ne varie pas, puisqu'elle est réduite de moitié pour chacun d'eux) perturbe moins la reconnaissance des mélodies que celle de la condition PC (où la taille des intervalles varie constamment par rapport à la mélodie originale). La transformation PC affecte plus la reconnaissance des séquences mélodiques originales que la transformation LT. Si l'on compare les résultats des conditions LT et OPC, on peut alors savoir si l'information chromatique a permis de mieux reconnaître les mélodies présentées. En effet, dans la condition OPC, le pourcentage d'identifications correctes est supérieur (78%) à celui de LT (60%), cette différence ne peut être due qu'à la conservation de l'information chromatique. Les résultats de cette expérience ont été confirmés dans une expérience de KALLMAN et MASSARO (1979). Dans cette expérience, où les sujets ne sont pas dans une situation de choix forcé comme précédemment, on utilise 3 mélodies très connues sous la forme d'une suite de sons purs. Il y a 3 transformations possibles de la mélodie originale :

- OPC : chaque note est transposée à une ou plusieurs octaves au-dessus ou au-dessous du degré original; le contour reste ici inchangé.
- OVC : la transformation est la même que dans la condition OPC, l'information chromatique est donc maintenue, par contre, le contour mélodique est modifié par rapport à l'original.
- PCS : la hauteur tonale et la hauteur chromatique de chaque note sont modifiées, mais le contour reste intact.

Chaque sujet entend les 3 mélodies chacune selon une modalité différente; ce qui donne par exemple, pour un groupe expérimental : mélodie 1 version OVC, mélodie 2 version PCS, mélodie 3 version OPC. Les sujets sont prévenus des transformations possibles des mélodies sans savoir de quelles transformations il s'agit. A chaque fois que le sujet entend une mélodie, il doit écrire sur une feuille de papier le nom de la mélodie qu'il pense avoir reconnue. (Il semble, d'après ce que dit l'auteur, qu'on ne vérifie pas si les sujets connaissent vraiment toutes les mélodies, on le suppose simplement puisque ce sont des mélodies comme "Yankee doddle" par exemple).

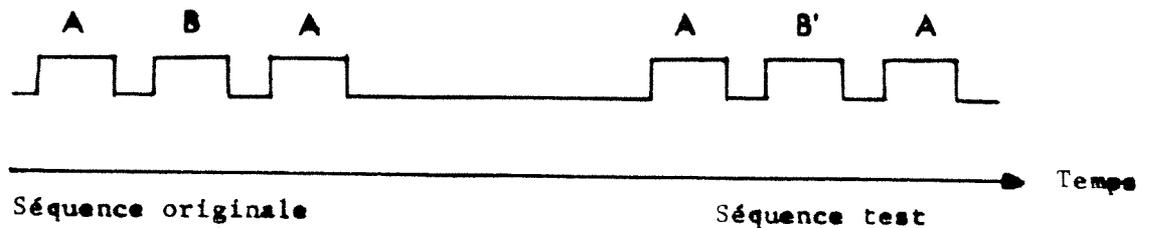
Pour chaque modalité de transformation, on compte le nombre d'identifications correctes : OPC : 72%, PCS : 11%, OVC : 8%. Cela veut dire que lorsque le contour est modifié, le fait de conserver l'information chromatique ne suffit pas pour reconnaître les mélodies (OPC : 72% contre PCS : 11%). Le pourcentage d'identifications correctes, lorsque la hauteur chromatique des notes ne change pas. Ce résultat va donc dans le sens de la conclusion de l'expérience précédente (IDSON et MASSARO p. 8). KALLMAN, MASSARO et KELLY (1980) ont confirmé ces résultats avec le même type de tâche, c'est-à-dire la reconnaissance de mélodies, mais en employant cette fois des mélodies créées pour l'expérience, donc totalement inconnues des sujets.

Ces expériences montrent surtout que le contour est un facteur primordial de reconnaissance de mélodies, et que s'il est modifié, l'équivalence chromatique ne permet pas de reconnaître des mélodies où des notes ont été transposées à l'octave.

Malheureusement, ces études opposent le plus souvent l'effet des modifications du contour à l'effet de la conservation de l'information chromatique, alors que des comparaisons entre des transformations du type OPC (chaque note transposée séparément à une ou plusieurs octaves) et du type LT (chaque note est transposée indépendamment à un degré différent de l'octave : la seconde, la tierce, la quarte ...) nous renseignent de façon beaucoup plus précise sur le phénomène de l'équivalence chromatique. C'est ce type de comparaisons qui a été fait par KALLMAN dans une récente expérience (1982). On demande à des sujets d'évaluer le degré de similarité qu'ils peuvent percevoir entre 2 séquences sonores.

KALLMAN utilise ici une méthode d'estimation, alors que dans les expériences que nous venons de citer, on demandait aux sujets de reconnaître une mélodie.

La partie de l'expérience qui nous intéresse se présente ainsi : le sujet entend 6 sons purs :



Pour former l'intervalle AB', on prend l'intervalle AB auquel on rajoute un intervalle de 0 à 18 demi-tons, par pas d'un demi-ton. Il y a donc 19 conditions expérimentales différentes. A chaque essai, le sujet entend les 2 séquences mélodiques : séquence originale, puis séquence-test; il doit ensuite donner une note de similarité entre les 2 intervalles AB et AB'. Cette note est donnée sur une échelle de 11 éléments discrets numérotés de 0 à 1.

Les résultats obtenus sur 20 sujets montrent que le degré de similarité décroît au fur et à mesure que la différence entre AB et AB' s'accroît. On remarque cependant une très légère augmentation du degré de similarité entre les 2 intervalles lorsque AB' diffère de AB, de 12 demi-tons, soit une octave. Cette légère augmentation est statistiquement significative à  $P < .05$ , par rapport à une différence de 11 ou 13 demi-tons entre les 2 intervalles. L'auteur note néanmoins de grandes différences interindividuelles, certains sujets étant beaucoup plus sensibles que d'autres à l'équivalence chromatique. Dans une deuxième partie de l'expérience, KALLMAN utilise la même procédure, mais cette fois, l'intervalle AB' est égal à l'intervalle AB augmenté de 10 à 14 demi-tons par pas de un 1/2 ton. Les résultats obtenus sur 25 sujets confirment très nettement la tendance mise en évidence dans la première partie de l'expérience.

Enfin, une expérience plus ancienne de HOUSE (1977) arrive à des conclusions plus tranchées. Comme dans les autres expériences, sauf celle de KALLMAN, le sujet doit reconnaître des mélodies connues, mais cette fois cette tâche est associée à une méthode de choix forcé. L'expérience se déroule ainsi : le sujet entend une mélodie que l'on a transformée par rapport à l'originale. Il y a 10 mélodies connues parmi lesquelles le sujet choisit le titre de la mélodie reconnue.

- Il y a 5 transformations possibles des mélodies originales :
- 1 - les notes sont transposées indépendamment à des degrés pris au hasard parmi 3 octaves successives (nous les appellerons : octave supérieure, octave centrale et octave inférieure).
  - 2 - les notes sont alternativement transposées à l'octave centrale et à l'octave supérieure.
  - 3 - les notes sont alternativement transposées à l'octave centrale et à l'octave inférieure.
  - 4 - les notes sont alternativement transposées à la tierce (quarte ou quinte) centrale et inférieure.
  - 5 - les notes sont alternativement transposées à la tierce (quarte ou quinte) centrale ou supérieure.

L'auteur note d'abord qu'il n'y a pas de différences entre les résultats des conditions 1, 4 et 5, où le nombre d'identifications correctes est faible. Dans les conditions 2 et 3, par contre, le taux de reconnaissance des mélodies est plus élevé (HOUSE ne précise pas de combien) que pour les 3 autres conditions. L'auteur conclut qu'il y a eu une assimilation entre des sons à l'octave, ce qui a permis aux sujets de mieux reconnaître les mélodies modifiées des conditions 2 et 3. Malheureusement, les résultats de HOUSE sont incomplets, de plus il ne dit rien des modifications éventuelles du contour mélodique lors des transformations, et nous avons vu à quel point elles intervenaient dans les résultats. Il faut donc prendre ce résultat avec quelques réserves.

La dernière expérience que nous citerons dans ce chapitre sur les sujets "tout-venant" nous intéresse à plus d'un titre, puisque les sujets sont des bébés de 8 à 11 mois. TREHUB, BULL et THORPE<sup>o</sup> ont utilisé une méthode de conditionnement opérant dans une tâche de discrimination de mélodies. La réponse que l'on va conditionner est l'orientation de la tête vers la source sonore. Le renforcement se fait en présentant, dans le champ visuel du bébé un objet lumineux. Les séquences mélodiques sont composées de 6 sons purs de 200 msec. chacun, séparés par des silences de 200 msec. L'intervalle de silence entre l'émission de 2 séquences sonores successives est de 800 msec. Il y a 3 mélodies originales différentes, le niveau sonore des sons de ces séquences originales est à 72 dB. Ces 3 mélodies subissent 4 types de transformation :

<sup>o</sup> (communication personnelle) 1982.

- 1 - Transposition : la fréquence absolue de chaque son est changée mais la taille de 2 intervalles successifs est la même.
- 2 - Contour préservé : la fréquence absolue des sons et la taille des intervalles sont modifiées, mais pas le contour.
- 3 - Changement d'octave/contour préservé : plusieurs sons dans la mélodie sont déplacés d'une ou plusieurs <sup>octaves</sup>, le contour étant préservé. Dans ce cas, l'information chromatique est conservée.
- 4 - Changement d'octave/contour modifié : plusieurs sons sont déplacés d'une ou plusieurs octaves, mais cette fois, le contour est modifié. L'expérience se déroule de la façon suivante :

Phase 1 : on présente au bébé une des 3 mélodies originales, et ce plusieurs fois de suite jusqu'à ce qu'il soit bien calme et qu'il regarde droit devant lui.

Phase 2 : (phase d'apprentissage) on fait entendre au bébé une deuxième mélodie, cette mélodie est soit "contrastée" (groupe contrôle) c'est-à-dire que la fréquence des notes ainsi que le contour sont changés par rapport à la mélodie originale, avec une intensité sonore supérieure de 5 dB à celle de l'originale, soit "transformée" : la mélodie est transformée selon l'une des 4 transformations ci-dessus 1, 2, 3 ou 4.

Dans le groupe contrôle, on renforce la mélodie contrastée, dans le groupe expérimental, c'est la mélodie transformée qui est renforcée. Le renforcement se fait de la façon suivante : si le bébé tourne la tête vers la source sonore durant la mélodie-test (c'est-à-dire celle qui sera renforcée : la mélodie contrastée ou la mélodie transformée) ou durant l'intervalle de silence de 800 msec., la réponse est enregistrée et renforcée par un jouet lumineux, qui s'allume et s'anime pendant 4 secondes. Lorsque le critère d'apprentissage, fixé par les auteurs, est atteint, on passe à la phase-test.

Phase 3 : (phase-test) cette phase comporte 20 essais (un essai étant constitué par la présentation de la mélodie originale suivie de la mélodie-test) : 10 essais où la mélodie-test est la mélodie contrastée - nous les appellerons les essais-contrôles - 10 essais avec la mélodie transformée, ce sont les essais expérimentaux. L'analyse des résultats se fait à partir du décompte du nombre de bonnes réponses (orientation de la tête vers la source sonore lors de la présentation de la mélodie-test) durant

les 20 essais.

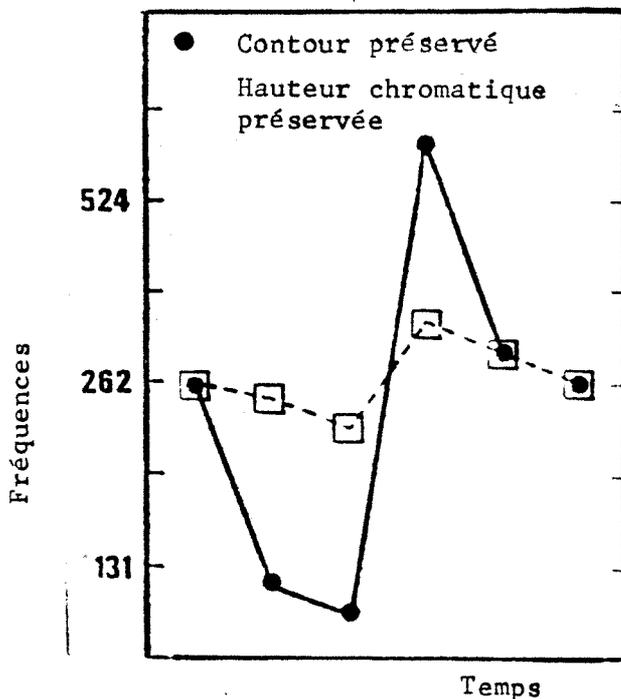
Les bonnes réponses permettent de voir si la mélodie-test et la mélodie originale ont bien été discriminées.

Les résultats obtenus sur 92 bébés (53 garçons et 39 filles) sont nets et peuvent se résumer ainsi : toutes les mélodies transformées, quelque soit la modalité de transformation 1, 2, 3 ou 4, ont été parfaitement discriminées des mélodies originales. La donnée qui nous intéresse le plus, concerne plus particulièrement les conditions 3 et 4. En effet, la discrimination entre les 2 mélodies est meilleure lorsque les notes sont déplacées d'une ou plusieurs octaves que dans les 2 autres conditions (1 et 2). On peut donc dire que les mélodies transformées dans les conditions 3 et 4 sont très différentes de la mélodie originale.

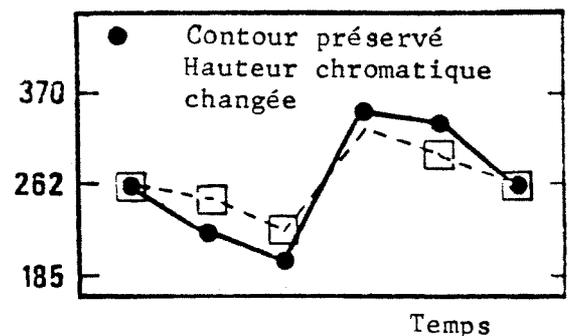
Les auteurs concluent donc qu'une mélodie ne peut être confondue avec sa transposition à l'octave. Dans ces conditions, l'équivalence chromatique n'a pu être mise en évidence.

Un point un peu obscur semble cependant remettre en question la conclusion des auteurs. Cette remarque concerne les conditions 2 et 3 ; regardons ce que donnent ces 2 transformations sur les 2 schémas suivants :

Condition 2



Condition 3



□ - - □ Mélodie originale  
● — ● Mélodie transformée

Stimulus utilisés par TREHUB et al. (1982)

Conditions 2 et 3

On voit tout de suite que la mélodie de la condition 3 subit une transformation très importante par rapport à l'autre mélodie. Pour que les 2 conditions soient effectivement comparables, il aurait fallu égaliser la taille des intervalles pour les 2 conditions. En voyant ces 2 schémas, on peut difficilement soutenir que la mélodie de la condition 3 a été mieux discriminée de la mélodie originale que la mélodie de la condition 2 uniquement parce que la hauteur chromatique des notes a été conservée. La modification des fréquences absolues ne s'est pas faite dans le même rapport pour les 2 conditions, or la conclusion des auteurs est basée sur cette comparaison.

Citons enfin une dernière expérience qui résume assez bien l'ensemble des conclusions sur la perception de la hauteur chromatique chez des sujets tout-venant. THURLOW et ERCHUL (1977) ont testé 11 adultes à qui on présente alternativement 2 sons purs : l'un d'une fréquence de 400 Hz, l'autre d'une fréquence de 500 (550, 1000, 1100, 2000 ou 2200 Hz). La tâche du sujet est d'augmenter la fréquence du deuxième son, grâce à un potentiomètre afin d'obtenir la "même qualité de hauteur" (selon les termes de la consigne) que pour le premier son de 400 Hz. Sur 11 sujets, un seul a choisi pour chaque ajustement (il y en a 48) un son dont la fréquence est à une ou plusieurs octaves du son de 400 Hz, un deuxième sujet a fait 33 ajustements à l'octave sur 48. Pour les 9 sujets restants aucun résultat significatif n'apparaît. Ces résultats peuvent paraître décevants pour qui s'attendait à trouver un effet de la similarité chromatique dans les ajustements demandés.

ATTNEAVE et OLSON (1971), eux, ont utilisé des mélodies de 3 notes.

On demande aux sujets de reproduire cette mélodie en leur donnant comme point de départ une des 3 notes mais en présentant celle-ci à un degré de la gamme différent de celui de l'originale. Il faut donc transposer la mélodie de 3 notes; les résultats montrent que lorsque la transposition doit se faire à l'octave, les performances ne sont pas meilleures que pour les autres transpositions. ATTNEAVE et OLSON attribuent ces résultats en partie au fait que les stimulus sont pauvres.

Il est difficile après toutes ces expériences de donner une conclusion catégorique, on peut simplement dire que dans un groupe de sujets pris au hasard dans la population, certains sont capables dans

certaines conditions expérimentales de percevoir une similarité entre des sons à l'octave. Le problème est de savoir ce que veut dire "un groupe de sujets tout-venant", autrement dit, est-ce que les sujets qui, dans ce groupe, sont sensibles à la hauteur chromatique des sons, ont eu ou non une expérience musicale. En général, le degré d'éducation musicale n'est pas notifié pour chaque sujet dans les expériences citées, la sélection des sujets permettrait donc de préciser les conclusions concernant l'équivalence chromatique.

## 2.2. Sujets tout-venant sélectionnés

DEUTSCH (1973) en utilisant une procédure expérimentale tout-à-fait différente de celle que nous avons vue dans ces précédents travaux, a obtenu des résultats intéressants. L'expérience est faite sur 36 sujets non-musiciens, mais ayant un minimum de capacités auditives (sélectionnés par un petit test). L'expérience se déroule ainsi : les sujets entendent d'abord un son pur S, suivi d'une série de 6 sons purs, puis un dernier son pur C. Après avoir entendu ces 8 sons, le sujet doit dire si S et C sont les mêmes ou sont différents. On va mesurer l'effet de la séquence de 6 sons sur la mémorisation du son S. Il y a 2 conditions expérimentales : soit S et C ont la même fréquence, soit S et C diffèrent d'un 1/2 ton. Pour chacune de ces 2 conditions, il y a 4 façons de composer la série de 6 sons purs entre S et C.

Conditions	Pourcentages d'erreurs
<u>1 - S et C sont différents</u>	
1.a/ Aucun son dont la hauteur est égale à celle de C ou de C <sup>+</sup> une octave n'est inclus dans la séquence .....	4,6
1.b/ Un son identique à C est inclus dans la séquence ....	26,7
1.c/ Un son égal à C <sup>+</sup> une octave est inclus dans la séquence .....	20,2
1.d/ Un son égal à C <sup>-</sup> une octave est inclus dans la séquence .....	12,1
<u>2 - S et C sont identiques</u>	
2.a/ Aucun son égal à S (ou C) $\pm$ 1/2 ton, ou S $\pm$ 1/2 ton - 1 octave n'est inclus dans la séquence.....	5,6
2.b/ S + 1/2 ton et S - 1/2 ton son inclus dans la séquence .....	24,4
2.c/ S + 1/2 ton + 1 octave et S - 1/2 ton + 1 octave sont inclus dans la séquence .....	21
2.d/ S + 1/2 - 1 octave et S - 1/2 ton - 1 octave sont inclus dans la séquence .....	11,3

En comparant les pourcentages d'erreurs (colonne de droite) pour chacune des 8 conditions expérimentales, on voit que c'est dans les cas 1b et 2b que la mémorisation de S a été le plus perturbée. Les conditions 1a et 2a servent en quelque sorte de référence, c'est là que l'on trouve le plus faible pourcentage d'erreurs, lorsque la série de 6 sons ne contient pas de son ayant la même fréquence que S et C (+ 1 octave). Dès que S et C se retrouvent dans la série intermédiaire des 6 sons (1b et 2b) alors le taux d'erreurs devient beaucoup plus important; cet effet se maintient lorsqu'on augmente d'une octave les fréquences de S et de C (1c et 2c): le pourcentage d'erreurs dans ces 2 conditions se rapproche en effet de celui des conditions 1b et 2b. Ceci voudrait dire que la présence des sons S et C (+ 1 octave) au sein de la série intermédiaire des 6 sons perturbe la mémorisation du son S presque autant que lorsqu'on insère dans cette série les sons S et C. On peut donc supposer que les sujets ont été sensibles à la hauteur chromatique du son S. On peut regretter une seule chose; pourquoi ne pas avoir fait des groupes contrôles où l'on aurait inséré dans la série intermédiaire des 6 sons, non plus S et C + une octave, mais S et C + une tierce, (4 1/2 tons) une quarte (5 1/2 tons) ou une quinte (7 1/2 tons). Ces intervalles ont un statut privilégié en musique, tout comme l'octave, il aurait donc été intéressant de compléter cette expérience en utilisant ces différents intervalles.

Une question se pose concernant les conditions 1d et 2d. On remarque en effet, que les pourcentages d'erreurs ont diminué par rapport à ceux des conditions 1c et 2c. A priori, il ne devrait pas y avoir de différences aussi marquées entre les conditions c et d, puisque S-1 octave diffère de S d'une octave, tout comme S+1 octave diffère de S d'une octave. L'auteur ne donne malheureusement aucune explication à ce résultat qu'on peut simplement noter pour mémoire. Néanmoins, cette expérience reste intéressante, et originale par la procédure employée; en effet, la plupart des travaux que nous allons citer utilisent des méthodes d'identification d'intervalles musicaux ou de mélodies, ou des jugements de similarité. C'est le cas d'une expérience de THURLOW et ERCHUL (1977). Les stimulus ici ne sont plus des sons purs, mais des sons complexes (sons de piano). Les sujets entendent 3 notes successivement :

Dans un premier cas, la première note est  $sol_0$ , la deuxième  $sol_1$ ,  $sol_2$ ,  $sol_3$ ,  $sol_4$  ou  $sol_5$  (les indices correspondent à 6 octaves successives,

0 à 5), la troisième note est le la ou le si de la même octave que le sol précédent : ce qui donne par exemple la séquence suivante :

sol<sub>0</sub>, sol<sub>4</sub>, la<sub>4</sub>, ou bien, sol<sub>0</sub>, sol<sub>2</sub>, si<sub>2</sub>. Dans un deuxième cas, la première note est toujours sol<sub>0</sub>, les deuxième et troisième notes sont interverties par rapport au cas précédent, on a donc par exemple la séquence suivante :

sol<sub>0</sub>, la<sub>3</sub>, sol<sub>3</sub> ou encore sol<sub>0</sub>, si<sub>5</sub>, sol<sub>5</sub>.

On demande au sujet quelle note lui semble la plus similaire à la note sol<sub>0</sub> : est-ce la deuxième ou la troisième de la séquence?

Dans ce test, seuls 3 sujets sur 9 ont choisi un des multiples de l'octave de sol<sub>0</sub> (c'est-à-dire sol<sub>1</sub>, sol<sub>2</sub>, sol<sub>3</sub>, sol<sub>4</sub> ou sol<sub>5</sub>) comme étant la note la plus similaire à sol<sub>0</sub> (résultat significatif à .05). Certains sujets, peu nombreux, perçoivent une similarité chromatique entre les sons présentés. Malgré le faible effectif des sujets dans cette expérience, on peut dire qu'il n'y a qu'une faible tendance (3 sur 9) à choisir plus particulièrement les notes à l'octave comme étant les plus similaires à la note-test sol<sub>0</sub>.

Citons enfin une expérience de DEUTSCH (1979). La procédure est la suivante : le sujet entend 2 mélodies, une mélodie originale composée de 6 sons purs et une mélodie de comparaison qui est une transformation "juste" ou "fausse" de la mélodie d'origine. Pour la moitié des essais, la transformation est "juste", c'est-à-dire qu'elle correspond à une transposition exacte à 4 demi-tons au-dessus de la mélodie originale; pour l'autre moitié, la transformation "fausse" est une transposition à 4 demi-tons, comme pour la transformation "juste", mais en plus, on permute 2 notes dans la série (exceptées la première et la dernière.) Dans ce dernier cas, l'auteur note que le contour de la mélodie est toujours préservé.

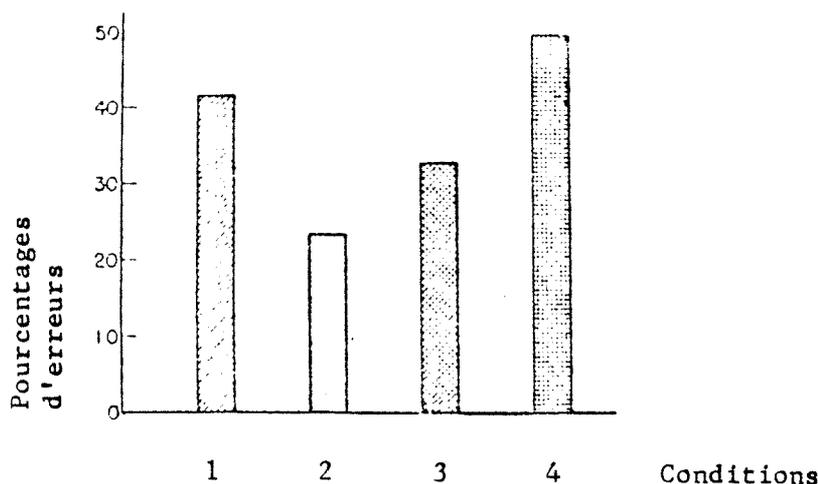
Il y a 4 conditions expérimentales :

- 1 - une présentation de la mélodie originale, puis une présentation de la mélodie de comparaison (soit la transformation "juste", soit la transformation "fausse").
- 2 - 7 présentations de la mélodie originale, puis une fois la mélodie de comparaison.
- 3 - une présentation de la mélodie originale, puis 6 présentations de cette même mélodie, mais transposée à l'octave (3 fois à l'octave supérieure, 3 fois à l'octave inférieure), enfin une présentation

de la mélodie de comparaison.

- 4 - une présentation de la mélodie originale, puis 6 présentations de cette même mélodie où chaque note est transposée alternativement à l'octave supérieure et à l'octave inférieure, et enfin une présentation de la mélodie de comparaison.

A chaque fois le sujet doit dire "pareil" ou "différent". Les réponses correctes sont: "pareil" lorsqu'on présente la transformation "juste", et "différent" pour la transformation "fausse". Pour chaque condition, il y a 12 essais : 6 avec une transformation "juste", 6 avec une transformation "fausse". On comptabilise ensuite le nombre des erreurs commises, pour chacune des 4 conditions.



Pourcentage d'erreurs dans une tâche de reconnaissance de mélodies

D'après DEUTSCH(1979)

Comme on pouvait s'y attendre, la répétition de la mélodie originale dans la condition 2 a un effet facilitateur sur la mémorisation de la mélodie (26% d'erreurs seulement). On remarque que lorsqu'il n'y a qu'une seule répétition de la mélodie originale (condition 1) le taux d'erreurs est plus élevé que dans la condition 2 où la mélodie originale est présentée 7 fois :

la différence entre les deux conditions est significative à  $\underline{p} < .01$ . Si l'on compare maintenant les conditions 1 et 3, la différence entre les pourcentages d'erreurs est encore significative à  $\underline{p} < .01$ . La présentation répétée de la mélodie originale transposée à l'octave (condition 3) améliore sensiblement la reconnaissance de la mélodie originale. Par contre, dans la condition 4, le taux d'erreurs atteint 50%, mais il semble évident que ce type de transformation a de grandes chances de modifier le contour de la mélodie; l'auteur ne dit rien à ce sujet, on peut donc difficilement comparer les résultats obtenus dans la condition 4 avec le reste des résultats. Dans l'ensemble, les sujets semblent avoir été sensibles à la hauteur chromatique, c'est du moins ainsi qu'on peut expliquer les différences observées entre les conditions 3 et 1.

On peut encore une fois regretter le manque de situations contrôles où on aurait utilisé des intervalles différents de l'octave. Rappelons que ces résultats sont obtenus sur des sujets non musiciens, que l'on a sélectionnés en vue de passer cette expérience.

Ce qui caractérise les résultats de ces sujets, comme ceux des sujets "tout-venant", c'est la dispersion des données. En effet, les auteurs notent à chaque fois des différences interindividuelles importantes, ces différences s'amenuisent dans le groupe des sujets musiciens.

### 2.3. Sujets musiciens

Dans une partie d'une expérience que nous avons déjà citée, p. 15, THURLOW et ERCHUL (1977) testent, avec la même procédure, des sujets musiciens (le nombre d'années d'études musicales n'est pas précisé). On procède ainsi : les 9 sujets entendent 2 sons purs (durée d'un son : 2,5 sec.) alternativement. Le premier son a une fréquence de 200 Hz, le deuxième a une fréquence initiale de 325 Hz, 650 Hz, 1300 Hz et 2600 Hz, cette fréquence est variable. Le sujet a pour tâche d'augmenter cette fréquence, afin d'avoir la "même qualité de hauteur" que pour le son de 200 Hz. Il y a donc 4 ajustements à faire à partir des 4 fréquences initiales ci-dessus. On obtient les résultats suivants :

- 6 sujets sur 9 passent de la fréquence 325 Hz à la fréquence 400 Hz : c'est-à-dire que la fréquence est augmentée jusqu'à l'octave supérieure du premier son de 200 Hz.

- 5 sujets sur 9 augmentent jusqu'à 800 Hz, la fréquence initiale de 650 Hz (800 Hz correspond à la double octave du son de 200 Hz).
- 3 sujets sur 9 augmentent jusqu'à 1600 Hz la fréquence initiale de 1300 Hz (1600 Hz correspond à la troisième octave supérieure du son de 200 Hz).

Ces résultats concernent donc le nombre de sujets qui répondent en fonction de la hauteur chromatique des sons. On voit que plus on s'éloigne de la fréquence du son de référence (200 Hz) moins les sujets sont sensibles à la similarité chromatique : 6 sujets sur 9 pour la première octave, contre 3 sur 9, pour l'ajustement à la troisième octave.

L'expérience se poursuit avec la même tâche, mais le sujet va devoir faire des ajustements, en diminuant la fréquence initiale. Le premier son est toujours à 200 Hz, les fréquences initiales des 4 sons de fréquence variable sont égales à 6000, 3000, 1500 et 750 Hz. Le sujet doit diminuer la fréquence de chacun de ces 4 sons, jusqu'à obtenir la "même qualité de hauteur" que pour le son de 200 Hz. Dans ce cas, il n'y a aucun ajustement fait aux octaves supérieures. Cela voudrait dire que dans cette condition, les sujets ne perçoivent pas la hauteur chromatique; ce qui paraît un peu étonnant. Les auteurs expliquent ce résultat par le fait que les sujets, qui sont musiciens, accordent leurs instruments en formant des intervalles mélodiques ascendants. Ils concluent que les ajustements se sont faits selon la grandeur des intervalles et non à partir de la similarité des sons à l'octave : dans ce dernier cas en effet, le nombre d'ajustements corrects (à l'octave) aurait le même <sup>été</sup> dans les 2 parties de l'expérience - ajustements ascendants et ajustements descendants. Mais, si on y regarde de plus près, on voit qu'il est impossible de comparer les 2 types d'ajustement; prenons un exemple :

1ère partie : on présente le son 325 Hz, pour l'ajustement à l'octave supérieure de 200 Hz, c'est-à-dire 400 Hz, il faut augmenter la fréquence variable de 75 Hz (400 Hz - 325 Hz).

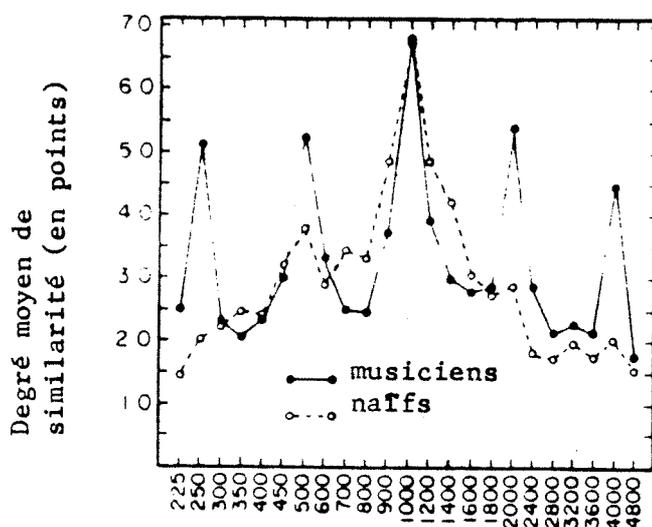
2ème partie : on présente un son de 750 Hz, il faut donc diminuer la fréquence de 350 Hz pour arriver à un son de 400 Hz. Pour les 4 fréquences proposées, l'ajustement descendant oblige le sujet à balayer une gamme de fréquences beaucoup plus large que pour l'ajustement ascendant, ce qui rend la tâche plus difficile lorsqu'on fait un ajustement descendant; ceci peut expliquer la différence

entre les résultats et redonner plus de poids à l'hypothèse de l'équivalence chromatique. Reste à savoir pourquoi la similarité chromatique diminue au fur et à mesure qu'on s'éloigne de la fréquence de 200 Hz, en effet, lorsqu'on arrive à la troisième octave (1600 Hz) il n'y a plus que 3 sujets sur 9 qui font un ajustement à l'octave contre 6 sur 9 pour la première octave (400 Hz). L'expérience ne permet pas de donner une explication à ce résultat.

Une expérience faite sur un nombre de sujets un peu plus nombreux (20 sujets) donne des résultats plus nets tout en utilisant des stimulus simples.

ALLEN (1967) a pris 2 groupes de sujets : 10 musiciens, 10 sujets naïfs (non musiciens). Chaque sujet entend la séquence suivante : un son de 1000 Hz - silence - un deuxième son dont la fréquence varie de 225 Hz à 4800 Hz.

La tâche du sujet est d'évaluer le degré de similarité entre les 2 sons présentés, et de donner sa réponse sur une échelle en 7 points. Les résultats montrent que pour les sujets musiciens, des sons à l'octave ont un très fort degré de similarité. On voit aussi apparaître une très nette différence entre les sujets musiciens et les non musiciens. (voir ci-dessous) .



Fréquence du 2ème son  
présenté (en Hz)

Degré de similarité perçu  
entre 2 sons

D'après ALLEN (1967)

Pour les non musiciens, en effet, la sensibilité à la hauteur chromatique semble pratiquement inexistante. Néanmoins, on peut essayer d'expliquer les performances des sujets non musiciens par le fait qu'ALLEN n'utilise que 2 sons. Des auteurs tels que ATTNEAVE et OLSON (1971) et THURLOW et ERCHUL (1977) ont noté à quel point il est difficile de mettre en évidence l'équivalence chromatique avec des stimulus aussi pauvres. Pour leur part, les musiciens n'ont pas eu l'air très gêné par ces stimulus. Cette expérience est intéressante parce qu'elle oppose nettement 2 groupes de sujets : les musiciens et les non musiciens.

Une ancienne expérience de FRANCES (1958) apporte elle aussi des renseignements sur l'importance de l'éducation musicale dans la perception de la hauteur chromatique. FRANCES a été l'un des premiers à s'intéresser de près à la perception de transpositions de mélodies. Son expérience est simple. On utilise des mélodies composées de 12 sons complexes (sons de piano). Les sujets sont divisés en 2 groupes :

- musiciens ("exécutants moyens") avec 3 à 8 ans d'expérience musicale.
- non musiciens ("analphabètes musicaux").

Tous les sujets entendent des couples de mélodies : une mélodie originale suivie de sa transformation. Cette transformation peut être de deux sortes :

- transposition : soit à l'octave, soit à un degré différent de l'octave sur l'échelle musicale.
- Imitation : la mélodie est d'abord transposée à un degré quelconque de la gamme, puis quelques intervalles musicaux sont modifiés, c'est-à-dire qu'on change la fréquence de certaines notes au sein de cette nouvelle mélodie.

D'après les quelques exemples que donne l'auteur, il semble que le contour des mélodies soit préservé. La tâche du sujet est de dire si la deuxième mélodie entendue est une "transposition" ou une "imitation". Deux résultats sont à retenir à l'issue de cette expérience :

- 1 - Les musiciens font beaucoup d'erreurs, dans cette tâche que les non-musiciens.
- 2 - Les musiciens ne font pratiquement aucune erreur lorsqu'on présente une transposition à l'octave, ils répondent automatiquement "transposition", alors qu'il y a plus d'ambiguïté lorsque

la transposition se fait à un autre degré que l'octave. FRANCES note que ce résultat se retrouve, mais à un degré moindre chez les sujets non musiciens. On peut dire en conclusion, que la plupart des musiciens et une partie des non musiciens sont sensibles à la similarité chromatique.

De l'ensemble des expériences que nous avons citées, nous pouvons tirer la conclusion suivante :

l'équivalence chromatique peut être mise en évidence dans certaines conditions expérimentales (THURLOW et ERCHUL, IDSON et MASSARO ...). La capacité qui permet de percevoir la hauteur chromatique d'un son semble beaucoup plus développée chez les sujets ayant une certaine éducation musicale que chez des sujets ayant peu ou pas d'expérience musicale.

Cette différence laisserait supposer que le développement de cette capacité perceptive est lié au degré d'éducation musicale des sujets. (ALLEN, FRANCES).

### 3 - EXTRACTION DE LA HAUTEUR D'UN SON

L'existence même du phénomène de l'équivalence chromatique peut laisser penser que le mécanisme d'extraction de la hauteur d'un son se fait de façon particulière, peut-être simplifiée, lorsque l'on présente des sons à l'octave par rapport à des sons séparés par un autre intervalle. Pour comprendre l'origine de ce phénomène perceptif, il faut en premier lieu s'intéresser aux mécanismes qui permettent de déterminer la hauteur d'un son.

#### 3.1 Hauteur d'un son pur :

Un son pur se compose d'une seule fréquence, la hauteur perçue varie donc selon cette fréquence unique. HELMHOLTZ (1877) a été l'un des premiers à soutenir que le système auditif fait une analyse séquentielle des sons au niveau de l'organe de Corti (récepteur auditif périphérique). Plus tard, VON BEKESY (1942) étudie les mouvements de la membrane basilaire en faisant varier la fréquence stimulante; il s'aperçoit que la membrane vibre avec un maximum d'amplitude à une place donnée. Ces vibrations sont transformées ensuite en activité électro-chimique dans l'organe de Corti.

° Cité par MOORE (1982).

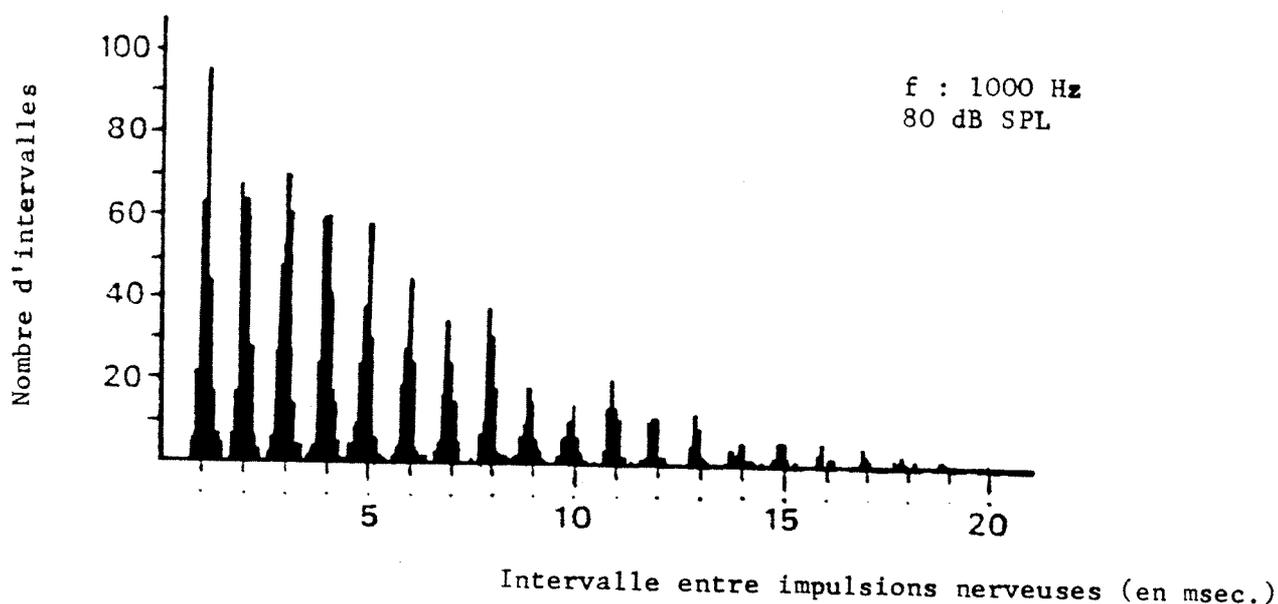
Lorsqu'on entend un son pur d'une fréquence donnée, le maximum d'oscillations de la membrane basilaire apparaît dans une seule région limitée de la membrane, région qui dépend de la fréquence du son. Sur la membrane basilaire, on détermine ainsi pour chaque fréquence, une région de sensibilité maximale, ou "région de résonance". En 1960, VON BEKESY a pu établir une représentation tonotopique de la membrane basilaire. La première sensation de hauteur est donc donnée par la position  $x$  de l'excitation sur la membrane basilaire. Dans ce cas, on admet que la hauteur d'un son est entièrement déterminée par cette position  $x$ , où l'on trouve le maximum d'amplitude de la vibration pour une fréquence donnée.

La première sensation de hauteur serait donc donnée par la position  $x$  sur la membrane basilaire; un changement de la fréquence du son pur entraînerait un changement de position de la région activée ( $x$ ). Ce changement entraînerait un changement de la hauteur du son présenté. Dans ce cas, l'extraction de la hauteur d'un son pur est basée sur l'analyse spatiale du son.

D'autre part, on sait qu'un son est transformé au sein de l'oreille interne en impulsions nerveuses acheminées ensuite le long des fibres du nerf auditif. Ces impulsions sont généralement synchronisées avec l'onde sonore du son stimulant de fréquence basse.

WEVER (1949) avec sa "théorie de la volée" a été le premier à dire que les intervalles de temps entre les impulsions nerveuses évoquées par un son jouaient un rôle dans la détermination de la hauteur.

L'histogramme (p.26) montre la distribution temporelle de ces impulsions nerveuses sous la forme d'un ensemble de raies. Plusieurs auteurs ont obtenu ce type de résultats (KIANG et al. 1965, ROSE et al. 1967) en expérimentant sur des animaux.



Cet intervalle permet de mesurer la taille de l'intervalle de temps entre deux pics nerveux successifs. Ici, on observe plusieurs raies (maximum) correspondant chacune à une valeur différente de l'intervalle entre deux spikes.

Ces valeurs sont approximativement égales à 1 msec, 2 msec, 3 msec, 4 msec ... C'est-à-dire des multiples entiers de la période du son (égale à 1 msec.).

Dans ce cas, l'extraction de la hauteur d'un son pur se fait selon une analyse temporelle du stimulus sonore.

Sans nous attarder sur la perception de l'extraction de la hauteur d'un son pur, voyons comment est perçue la hauteur d'un son complexe.

(Voir page 27).

### 3.2. Hauteur d'un son complexe

Pour les sons purs, la chose est "simple" : la hauteur ne peut correspondre qu'à la fréquence unique qui compose ce son.

Pour les sons complexes, le problème est différent puisqu'un tel son comporte plusieurs fréquences. La hauteur d'un son complexe est égale à la hauteur de la fondamentale, c'est-à-dire l'harmonique la plus basse. On a montré que même lorsque la fondamentale du complexe n'est pas présentée, on perçoit malgré tout la hauteur qui correspond à cette fondamentale : c'est le phénomène de la fondamentale absente ("missing fundamental"). C'est en 1841 que SEEBECK a mis en évidence expérimentalement ce phénomène. Plus tard, SCHOUTEN (1940) cité par MOORE (1982) p. 123 et LICKLIDER (1954) ont confirmé ce résultat avec différentes méthodes. Si la fondamentale n'est pas nécessaire à l'extraction de la hauteur, quelles sont les harmoniques "utiles" à cette extraction?

#### 3.2.1. Modèle du résidu

Dans un son complexe, on ne perçoit pas séparément toutes les harmoniques qui le composent. Les plus basses (7 à 9 premières harmoniques) sont séparées au niveau de la cochlée, alors que les harmoniques les plus hautes ne sont pas résolues par l'oreille (PLOMP, 1975, p. 209-210). Le groupe d'harmoniques hautes non résolues est appelé le résidu. La périodicité de cette onde résiduelle formée par ces harmoniques hautes est la même que celle de la fondamentale du complexe.

SCHOUTEN (1940) se base sur ce résultat pour établir son modèle de périodicité (ou modèle de résidu) : selon lui, la hauteur d'un son est extraite à partir de cette onde résiduelle. Ce modèle permet d'expliquer le phénomène de la fondamentale absente puisque l'extraction de la hauteur d'un complexe peut se faire uniquement à partir des harmoniques non résolues; dans ce cas, la fondamentale peut être absente, la hauteur du son sera malgré tout perçue.

#### 3.2.2. Modèle de reconnaissance de patterns

Les signaux nerveux évoqués sur la cochlée lors d'une stimulation correspondent aux premières sensations de hauteur issues de chacun des composants du son pris individuellement. Plusieurs expériences montrent

° PLOMP (1967).

que la hauteur d'un son complexe est étroitement liée à la hauteur de chacune de ses harmoniques. WALLISER (1969, cité par PLOMP, 1975) a trouvé que, dans certaines conditions, le changement de la hauteur perçue d'un groupe d'harmoniques est corrélé à un changement de la hauteur du son pur qui coïncide avec l'harmonique la plus basse de ce groupe, mais pas avec le son pur qui coïncide avec la fondamentale (absente). Ce qui veut dire que si l'on change la fréquence d'une des harmoniques d'un complexe (autre que la fondamentale) on change la hauteur du complexe lui-même. TERHARDT (1969-1970) montre après WARD (1954) que le rapport de 2 fréquences correspondant aux 2 hauteurs des 2 sons purs successifs composant un intervalle d'octave, est supérieur à 2:1<sup>1</sup>; c'est ce qu'on appelle le phénomène de l'élargissement d'octave ("Oktavspreizung"). Ce phénomène s'explique par une augmentation de la hauteur du son le plus aigu du fait de la présence du son plus grave; c'est ce qu'il appelle le déplacement de hauteur ("Tonhöhenverschiebung"). TERHARDT (1972) note que ce phénomène peut être mis en évidence avec des sons simultanés. Cela montre qu'il existe une interaction entre les différentes harmoniques d'un son, ce qui provoque des changements de hauteur des harmoniques et par conséquent du complexe tout entier. En outre, il montre (1972) que pour un son complexe contenant une fondamentale et un certain nombre d'harmoniques, la hauteur du composant fondamental a légèrement baissé, alors que la hauteur des harmoniques supérieures (à partir de la seconde) a augmenté. Ces expériences montrent donc que la hauteur de chaque harmonique d'un complexe se modifie (par une interaction avec les autres harmoniques).

D'après ces auteurs, (WALLISER, TERHARDT) l'extraction de la hauteur se fait à partir du pattern spatial de l'activité nerveuse de la membrane basilaire lors d'une excitation. Ce mécanisme d'extraction serait, selon ce modèle, un processus de reconnaissance des patterns spatiaux évoqués au niveau de la cochlée.

---

1 - Rappelons que le rapport des fréquences des 2 sons formant un intervalle d'octave est égal à 2:1, par exemple :  $f_1 = 200$  Hz,  $f_2 = 2f_1 = 400$  Hz; avec  $f_2/f_1 = 2/1$ .

Le modèle de reconnaissance de patterns de TERHARDT (1974) part d'un principe connu depuis HELMHOLTZ (1957) : l'analyse d'un son complexe se fait selon deux modes; d'abord un traitement analytique dont le but est de déterminer la hauteur de chacune des harmoniques (hauteur spectrale), et puis un traitement synthétique qui permet d'attribuer une hauteur globale au complexe (hauteur virtuelle). La hauteur spectrale est étroitement liée à la place de l'excitation dans la cochlée; alors que la hauteur virtuelle correspond à un mécanisme plus complexe où intervient justement le processus de reconnaissance de patterns. Selon le modèle de TERHARDT, le processus d'extraction de la hauteur se fait de la façon suivante :

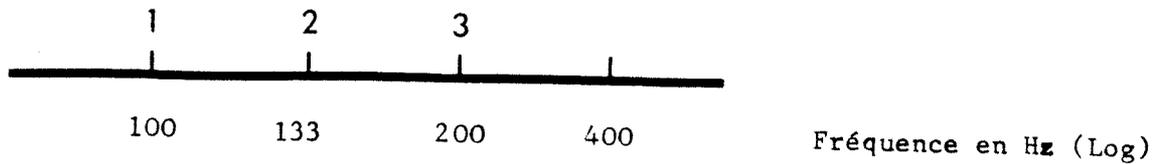
il y a d'abord une phase d'apprentissage puis une phase de reconnaissance. La phase d'apprentissage se déroule très tôt au cours du développement. Cet apprentissage se fait simplement à partir des sons naturels de l'environnement sonore du sujet, en particulier à partir des sons de la voix. Cette idée d'apprentissage à partir des sons de la voix a déjà été avancée par THURLOW (1957) et WHITFIELD (1969-70, cité par MOORE, (1982), p. 124.

A chaque harmonique correspond une information spectrale (traitement analytique) ainsi lorsqu'on perçoit un son complexe, le système auditif reçoit un grand nombre d'informations spectrales simultanées correspondant chacune à des harmoniques du complexe.

L'écoute répétée des sons complexes de la voix crée des liens virtuels entre les harmoniques d'un même complexe. On apprend donc à associer à un son une série harmonique. Cette phase d'apprentissage permet donc de mémoriser un ensemble d'harmoniques liées à une fondamentale.

La phase de reconnaissance se déroule ensuite ainsi : chaque harmonique d'un son complexe va en quelque sorte devenir une fondamentale à laquelle est associée une série de hauteurs virtuelles correspondant aux informations spectrales qui lui sont associées. Ces hauteurs virtuelles sont celles des sub-harmoniques de chacune des harmoniques du son complexe. Prenons comme exemple l'harmonique 4 d'un son complexe de 100 Hz; cette harmonique a donc une fréquence égale à 400 Hz; à cette fréquence correspondent les sub-harmoniques virtuelles suivantes :

(Figure  
à revoir)



Trois sub-harmoniques (1, 2 et 3) d'un son de 400 Hz

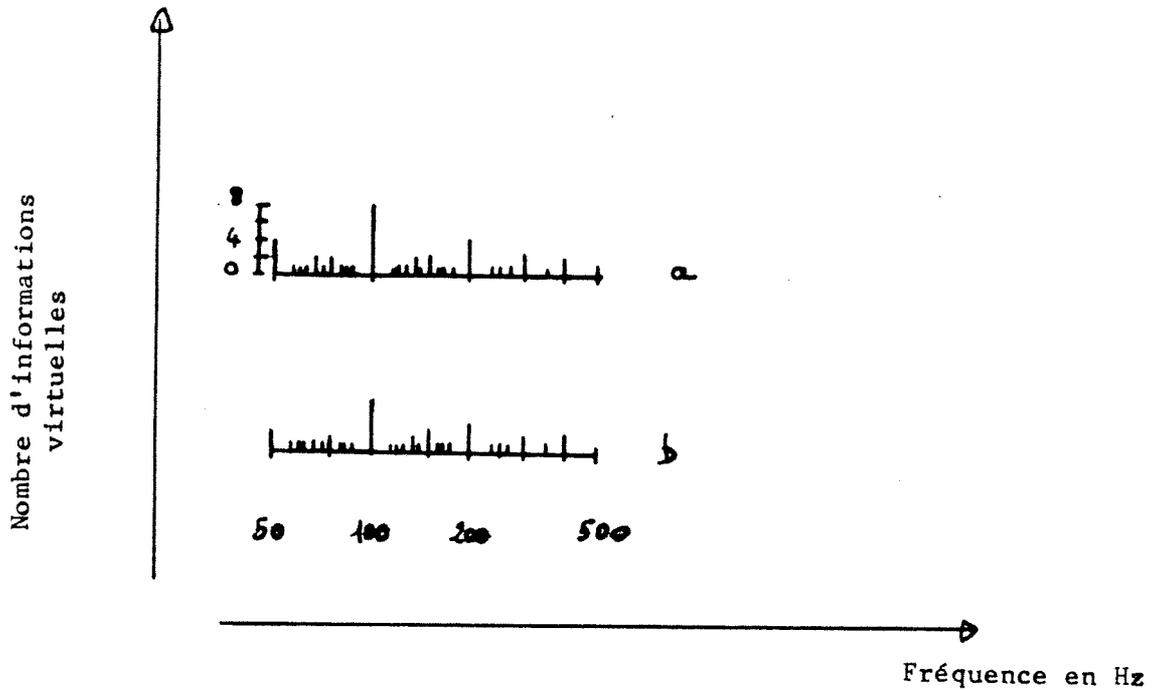
(Dans cet ensemble, nous avons ignoré les déplacements de hauteur)

A la hauteur spectrale de l'harmonique 4 (à 400 Hz) correspond donc selon TERHARDT une série de 8 hauteurs virtuelles des sub harmoniques de 400 Hz. A chaque harmonique sont associées 8 hauteurs virtuelles. La hauteur virtuelle du son complexe se détermine alors à partir des informations spectrales de la façon suivante :

les informations virtuelles de toutes les harmoniques d'un son s'ajoutent pour une fréquence donnée; la hauteur du complexe est alors déterminée à partir de la fréquence où coïncident le plus grand nombre d'informations virtuelles.

Sur ce schéma, on voit que le maximum d'informations virtuelles se trouve sur la fréquence 100 Hz, dans les 2 cas a et b, la hauteur virtuelle du son correspond à la fréquence 100 Hz.

(Voir page suivante).



Position du maximum de coïncidences  
pour une fréquence donnée.

Distribution temporelle des informations permettant  
de déterminer la hauteur virtuelle d'un son de 100 Hz.

Cas a : 10 harmoniques du son sont présentées (de 1 à 10)

Cas b : 5 harmoniques du son présentées (de 3 à 8).

D'après TERHARDT, 1974.

Le mécanisme d'apprentissage qui permet d'imprimer des "traces" de ces corrélations entre hauteurs spectrales se limite, toujours selon TERHARDT, à une bande de fréquence qui va approximativement de 50 à 500 Hz; c'est dans cette gamme de fréquences que l'on trouve toutes les fréquences fondamentales des sons complexes de la voix humaine, c'est-à-dire des sons qui permettent d'apprendre à extraire la hauteur de n'importe quel son complexe. Un modèle tel que celui-ci est basé sur la résolution en fréquence des composantes du son par le système auditif. C'est ce qu'on retrouve dans le modèle de GOLDSTEIN (1973) : ici le processus de reconnaissance de pattern est envisagé de façon différente. GOLDSTEIN introduit en effet la notion d'un processeur central faisant une sorte d'estimation optimale de la fréquence fondamentale à partir des harmoniques résolues par l'oreille. Il y a une différence essentielle entre le modèle de GOLDSTEIN et celui de TERHARDT. Pour ce dernier, l'extraction de la hauteur passe par l'analyse du pattern spatial de l'activité nerveuse observée au niveau de la membrane basilaire lors d'une excitation; alors que GOLDSTEIN fait l'hypothèse que le mécanisme de reconnaissance du pattern d'excitation se ferait à partir d'informations temporelles (rythme de décharge nerveuse dans les neurones récepteurs).

#### 4 - EXPLICATIONS POSSIBLES DE L'EQUIVALENCE CHROMATIQUE

Nous allons nous intéresser plus particulièrement aux théories permettant d'expliquer le phénomène de l'équivalence chromatique.

##### 4.1. L'équivalence chromatique et le modèle de TERHARDT

Nous venons de voir comment TERHARDT décrit le processus d'extraction de la hauteur d'un son complexe par apprentissage : comment à partir de ce processus peut-on expliquer le phénomène de l'équivalence chromatique?

On a vu qu'à chaque hauteur spectrale (i.e. la hauteur de chaque harmonique d'un complexe) correspondait un ensemble de hauteurs virtuelles. Pour un son complexe comportant par exemple 8 harmoniques (de la première à la huitième), il y a donc 8 groupes de sub-harmoniques à partir desquels la hauteur virtuelle du complexe est extraite. Cette hauteur virtuelle correspond au maximum de coïncidences<sup>1</sup> entre toutes les

---

1 - "Coïncidence" désigne un ensemble de sub-harmoniques virtuelles ayant la même hauteur.

hauteurs des sub-harmoniques. Mais il se trouve qu'il existe plusieurs sub-harmoniques pour lesquelles le nombre de coïncidences se rapproche de ce maximum qui donne la valeur de la hauteur du son. Le schéma de la page 31 montre bien qu'il existe plusieurs points où le nombre de coïncidences est élevé, par exemple, pour les fréquences 50 Hz, et 200 Hz; or ces fréquences sont celles des octaves supérieures et inférieures du son de 100 Hz. La similarité des sons à l'octave serait donc basée sur une ambiguïté de la hauteur virtuelle. RITSMA et ENGEL (1974, cité par TERHARDT, 1974) et RITSMA (1966) ont en effet observé que la hauteur d'un son complexe pouvait être confondue avec d'autres hauteurs différant d'une octave ou d'un cinquième (5 demi-tons) de la hauteur vraie. Pour un son de 100 Hz, par exemple, on peut trouver des hauteurs à des fréquences de 50 Hz (octave inférieure) de 100 Hz (octave supérieure) ou de 300 Hz (cinquième).

On a aussi montré que cette ambiguïté entre des sons à l'octave existait aussi chez des sujets ayant l'oreille absolue ("absolute pitch")<sup>1</sup>. Grâce à cette capacité perceptive, ces sujets ne font pratiquement aucune erreur lorsqu'on leur fait entendre une note et qu'on leur demande de la nommer. Les seules exceptions sont les "erreurs d'octave" (RUCKMICK, 1929; LUNDIN et ALLEN, 1962; WARD, 1963; et LOCKHEAD et BYRD, 1981).

Ainsi, si l'on demande à un sujet ayant l'oreille absolue d'identifier un la à 440 Hz, il va peut-être répondre la (220 Hz) ou la (880 Hz). Ces erreurs d'octave seraient, selon TERHARDT, une des conséquences de l'ambiguïté de la hauteur d'un son complexe. D'après les résultats des expériences sur l'équivalence chromatique, on sait que ce sont les sujets musiciens qui sont le plus sensibles au rapport de similarité; cela voudrait dire que ces sujets sont plus sensibles à l'ambiguïté de la hauteur virtuelle d'un son complexe que des sujets naïfs. En effet, s'il n'y a pas d'ambiguïté, il ne peut y avoir d'assimilation ou de confusion entre des sons à l'octave.

---

1 - L'oreille absolue est la capacité qui permet de nommer ou de reconnaître une note sans avoir de référence sonore extérieure.

Ainsi, certains auteurs (THURLOW, 1954; TERHARDT, 1974; HALL et PETERS, 1980-1982) soutiennent l'hypothèse d'un apprentissage associatif à partir des sons naturels. La sensibilité à la hauteur chromatique serait donc une capacité perceptive acquise à partir des sons complexes de la nature. Considérons à présent ce phénomène du point de vue d'un modèle temporel.

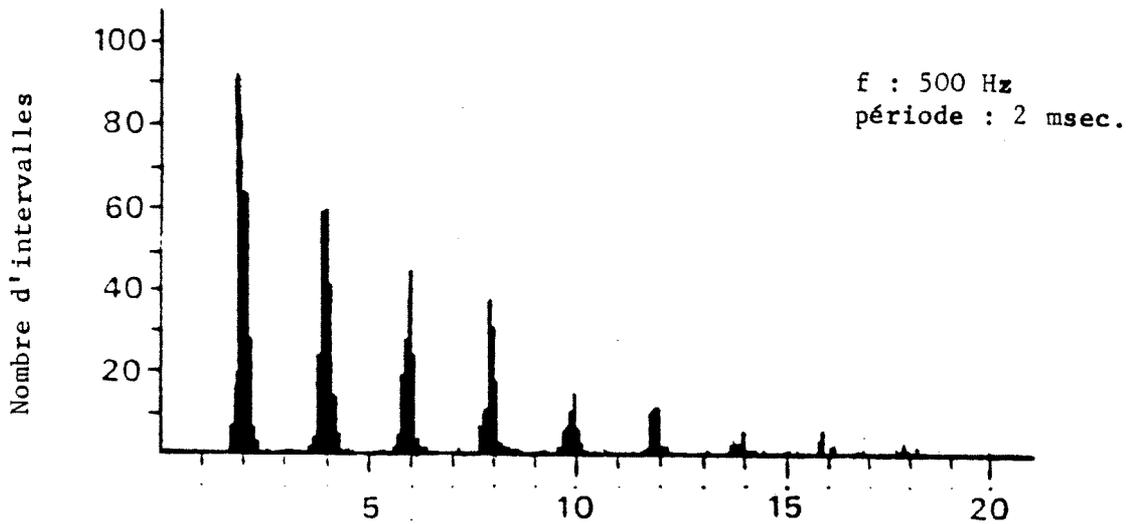
#### 4.2. L'équivalence chromatique et le modèle temporel d'OHGUSHI

A partir des résultats de ROSE et al (1967) (voir hauteur d'un son pur), OHGUSHI (1983) a expliqué le phénomène de l'équivalence chromatique. (Voir schéma page suivante).

Pour l'histogramme 1, la fréquence du son de stimulation est égale à 500 Hz, pour l'histogramme 2, la fréquence est de 1000 Hz. Si l'on compare la distribution des raies de ces 2 histogrammes, on voit qu'il existe une grande similarité entre les 2; en effet, toutes les raies de l'histogramme 1 sont présentes dans l'histogramme 2. Ainsi, lorsqu'on double la fréquence du son stimulant, c'est-à-dire lorsqu'on passe à l'octave supérieure, la fréquence du rythme de décharge nerveuse double elle aussi. Pour la fréquence de 500 Hz, le rythme de décharge correspond à la période de ce son, c'est-à-dire 2 msec, pour le son de 1000 Hz (octave supérieure), le rythme de décharge correspond à la valeur 1 msec, c'est-à-dire la période du son stimulant.

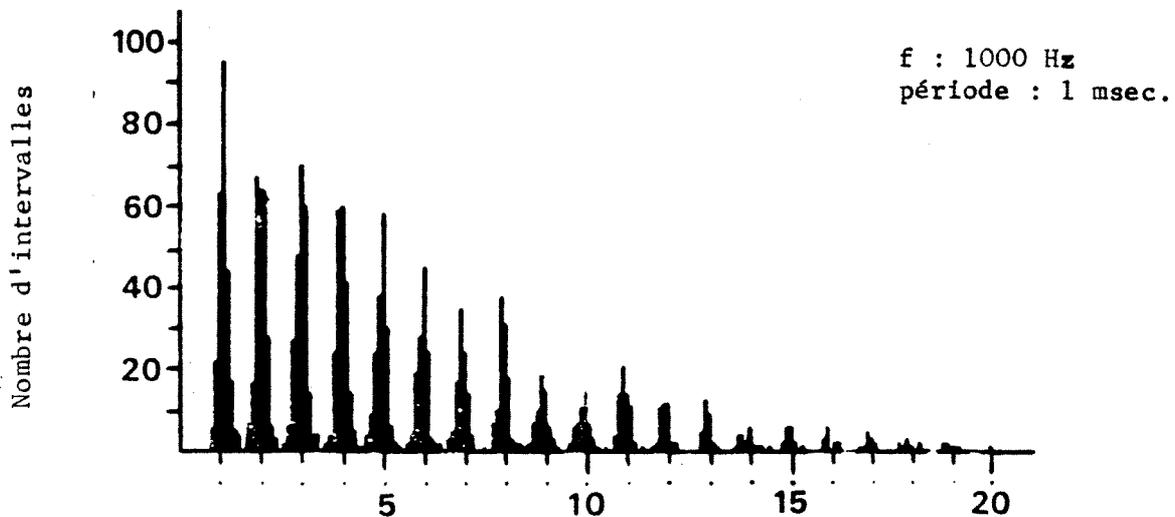
Ces 2 schémas mettent clairement en évidence la similarité de l'activité nerveuse évoquée par 2 sons séparés par une octave. Pour OHGUSHI, cette similarité est l'explication que l'on peut donner à l'équivalence chromatique. Ce mécanisme ne dépend donc pas d'un apprentissage, mais serait lié à la structure même du système auditif.

OHGUSHI explique enfin que si la limite de sensibilité à la hauteur chromatique se trouve à 5 KHz, c'est à cause de la désynchronisation qui se produit à cette fréquence entre le rythme de décharge nerveuse et la période du son de stimulation. Cette désynchronisation pourrait s'expliquer aussi la perte de qualité musicale des sons à partir de 5 KHz (ATTNEAVE et OLSON, 1971).

HISTOGRAMME 1

Intervalle entre impulsions nerveuses

(Cet histogramme ne correspond pas à des valeurs réelles).

HISTOGRAMME 2

Intervalle entre impulsions nerveuses

Distribution des intervalles entre impulsions nerveuses  
sur une fibre d'un nerf auditif d'un singe-écureuil.

D'après ROSE et al. 1967.

Jusqu'à présent, nous avons parlé de l'intervalle d'octave sans préciser que la valeur de l'octave subjective (perçue par le sujet) n'est pas la même que la valeur de l'octave physique.

#### 4.3. L'octave subjective

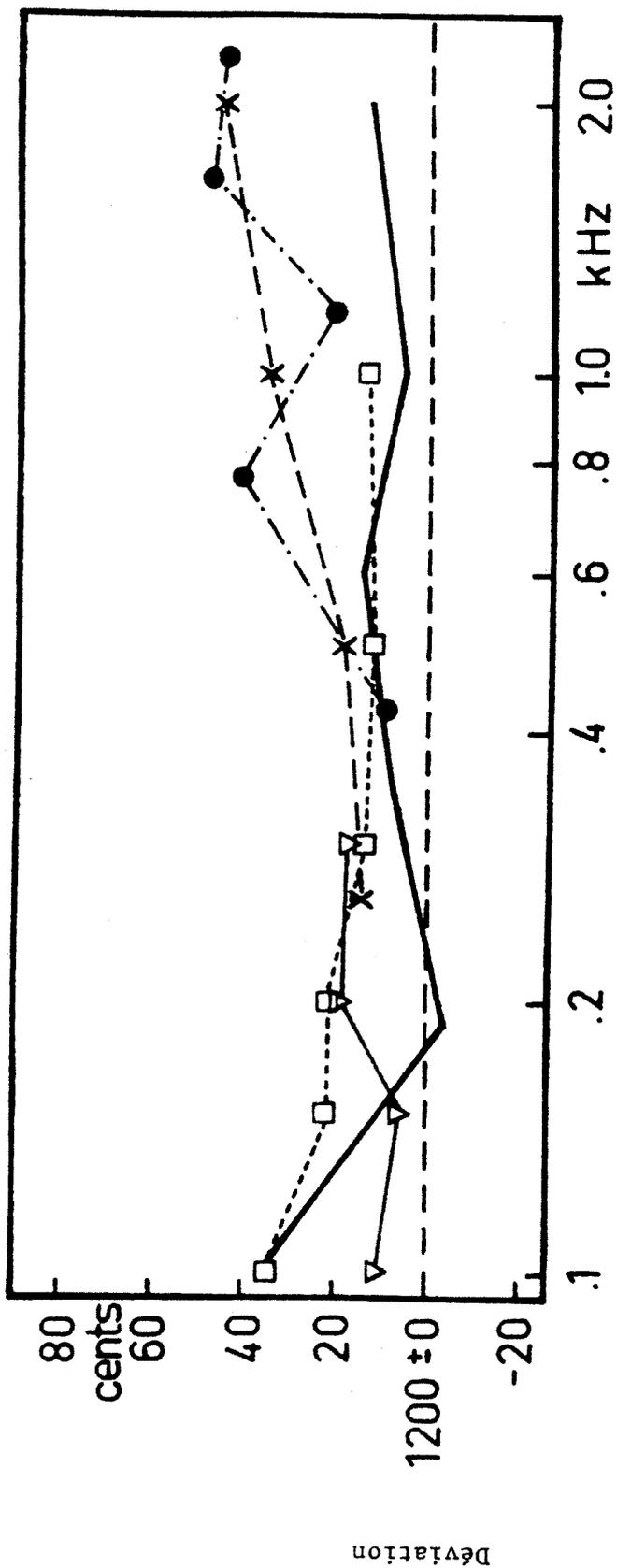
Lorsqu'on demande à un sujet d'égaliser un son de fréquence variable  $f_b$ , à l'octave supérieure d'un son de fréquence  $f_a$ , les sons étant présentés successivement, la valeur de  $f_b$  que l'on trouve lorsque l'ajustement est fait, est légèrement supérieure à la valeur physique attendue, c'est-à-dire  $f_b = 2 f_a$ .

De nombreux auteurs se sont intéressés à ce phénomène de l'octave subjective. WARD (1954) a donné un certain nombre d'indications concernant ce phénomène :

- les valeurs de la déviation qui existe entre l'octave subjective et l'octave physique varient selon les sujets.
- la valeur de la déviation augmente en fonction de la fréquence.
- les jugements intrasujets varient significativement et au hasard d'un jour à l'autre.
- il existe des différences entre l'oreille droite et l'oreille gauche. WARD note enfin que lorsqu'on présente les 2 sons simultanément, une légère différence apparaît entre octave subjective et octave physique; cette différence est significativement moins importante qu'en présentation successive.

De nombreux auteurs ont fait des expériences sur ce sujet. SUNDBERG et LINQVIST (1972-73) ont fait une comparaison entre leurs résultats et ceux obtenus par ces auteurs (schéma page suivante).

Les données montrent cependant une légère différence entre les sons purs et les sons complexes; dans ce dernier cas, en effet, la déviation entre octave subjective et octave <sup>physique</sup> complexe est moins importante. SUNDBERG et LINQVIST notent que les valeurs de la déviation varient non seulement en fonction de la fréquence, mais aussi de l'amplitude des sons. Ils remarquent enfin que l'échelle musicale de certains instruments est étirée, c'est-à-dire que les sons graves (en-dessous de 200 Hz environ) sont plus bas, et les sons aigus (au-dessus de 800 Hz environ) plus aigus, que sur une échelle tempérée où des intervalles musicaux auraient été calculés à partir de valeurs physiques (MARTIN et WARD, 1961).



Fréquence du son le plus grave

SUNDBERG-LINQVIST, 1973 : moyenne des résultats obtenus (sons complexes)

- : (sons purs)
- × : (sons purs)
- ▲ : (sons purs)
- : (sons complexes)

Déviatiion moyenne (octave subjective/octave objective)

Note : 1200 cents = 1 octave

Les déplacements de la hauteur des harmoniques des sons de la voix par exemple, entraînent un élargissement de l'octave perçue, et d'après TERHARDT, c'est cet intervalle élargi que l'on mémorise.

Voyons à présent, comment le modèle d'OHGUSHI explique le phénomène de l'élargissement de l'octave.

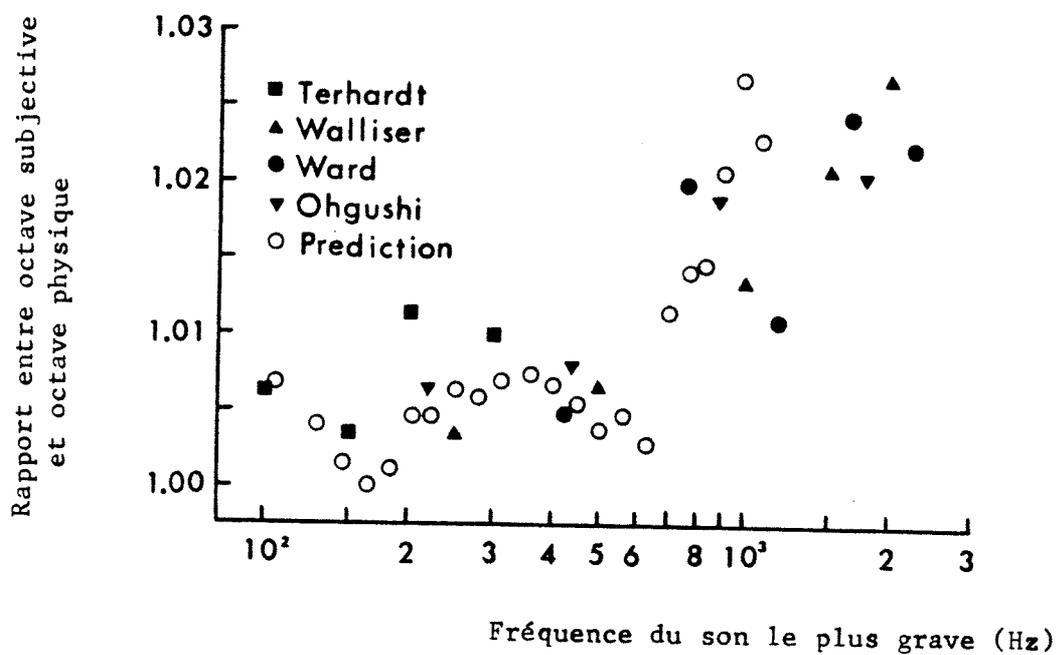
A partir des données neurophysiologiques de ROSE et al. (1967), OHGUSHI montre qu'il existe un léger décalage entre le rythme de décharge nerveuse recueillie sur les neurones récepteurs et le rythme de décharge prévu à partir de la période du son. Les raies sur l'histogramme sont déplacées par rapport à la place idéale directement liée à la période du son. Selon OHGUSHI, ce léger décalage est dû au fait qu'à partir de 2 Khz, il devient difficile pour le neurone de décharger 2 impulsions successives correspondant à 2 pics successifs de l'onde stimulante et ce à cause de la période réfractaire du neurone, c'est-à-dire du temps pendant lequel le neurone, après avoir répondu à une stimulation, demeure inexcitable; cette période terminée, le neurone redevient sensible aux stimulus présentés. OHGUSHI suggère que l'information temporelle est modifiée dans les fibres nerveuses. Cette modification peut être évaluée par le calcul. Ce qui permet alors de prédire la valeur de la différence entre l'octave subjective et l'octave physique. Ces précisions s'accordent avec les données expérimentales de plusieurs auteurs dont OHGUSHI lui-même (voir page suivante).

Ainsi, selon le modèle d'OHGUSHI, la similarité chromatique est innée, la hauteur d'un son est extraite à partir des informations temporelles recueillies par le système auditif périphérique.

#### 4.4. L'équivalence chromatique et le phénomène de l'acculturation

En 1865, BASEVI (FRANCES, p.123) énonce l'idée d'un parallélisme entre "l'histoire des institutions harmoniques et l'évolution du mode de perception dont cette histoire témoigne". L'acculturation se définit donc comme un processus interactif entre les caractéristiques d'un système musical donné et la perception auditive. Ce processus crée un certain nombre d'habitudes perceptives à partir d'une "grammaire" musicale donnée. La musique occidentale que nous entendons le plus souvent est une musique tonale, l'acculturation "tonale" définit le sentiment de tonalité que l'on peut expliquer ainsi :

lorsqu'on entend une mélodie en do majeur, qui commence par exemple par do<sub>1</sub>, on s'attend à ce qu'elle se termine par un do (c'est la tonique).



Déviations entre octave subjective et octave physique :

- Données expérimentales : TERHARDT, WALLISER, WARD, OHGUSHI
- Prédiction faite par le calcul : OHGUSHI

D'après OHGUSHI, 1982.

Selon le phénomène de l'acculturation tonale, une habitude perceptive va se créer à partir de cette caractéristique de la musique tonale. L'attente de la note finale représente l'étape secondaire après la période de familiarisation qui naît de la fréquence d'apparition de ce schéma "tonal". L'acculturation crée un certain nombre d'habitudes perceptives; ainsi, dans la musique tonale, une note est souvent associée à la note qui se trouve à l'octave inférieure et supérieure, ce qui établit une sorte de lien entre 2 notes séparées par une octave, et permet d'expliquer l'équivalence chromatique.

L'acculturation est un phénomène essentiellement génétique, et l'acculturation "tonale" n'est donc qu'une étape de ce phénomène; elle semble être établie entre 6 et 12 ans (voir FRANCES, p. 107). Certains auteurs tels que FRANCES (1958) et TEPLOV (1966) ont utilisé une méthode où la tâche du sujet était de compléter une mélodie en donnant la dernière note; on peut ainsi mesurer ainsi l'effet du "sentiment de tonalité" sur chaque sujet. Le sujet se familiarise donc au fur et à mesure de son développement avec certains aspects prégnants de la musique environnante. Selon le phénomène de l'acculturation tonale, l'équivalence chromatique est une capacité perceptive apprise (BURNS, 1974; BURNS et WARD, in *Psychology of Music*,<sup>°</sup> Autrement dit, la similarité entre des sons à l'octave est liée à certaines caractéristiques de notre musique tonale occidentale et n'est perçue qu'à partir de la huitième année (environ) de la vie (IMBERTY, p. 44, 1969).

Le processus d'acculturation permet d'expliquer la différence qui existe entre les musiciens et les non musiciens : le niveau d'acculturation étant différent pour chacun, les capacités perceptives d'un sujet musicien se développeront différemment de celles d'un sujet qui ne pratique pas un instrument ou le chant. TEPLOV (1966, p. 122) va même jusqu'à dire que la sensibilisation à la hauteur se développe sous l'effet de la pratique musicale. L'acculturation peut se définir comme un "apprentissage culturel"; ici le mot apprentissage est pris dans un tout autre sens que chez TERHARDT où l'apprentissage se fait à partir de sons naturels.

Chacune des trois interprétations que nous venons de voir répond de façon différente à la question "comment se développe la

<sup>°</sup> DEUTSCH, 1982.)

capacité d'entendre la hauteur tonale des sons?".

Les capacités perceptives du jeune enfant lui permettent-elles de percevoir cette hauteur tonale: c'est la question que nous nous sommes posée.

## PROBLEMATIQUE ET TECHNIQUE EXPERIMENTALE

### 1 - PROBLEMATIQUE

Aucune des expériences décrites précédemment n'a apporté de certitude quant à l'origine de l'équivalence chromatique. Néanmoins il reste trois hypothèses pour expliquer ce phénomène, hypothèses qui semblent à priori aussi fondées les unes que les autres. Chacune d'entre elles envisage de façon différente la façon dont apparaît et se développe cette capacité perceptive.

1. Bien que TERHARDT ne donne pas de précisions sur l'âge auquel doit se produire "l'apprentissage précoce" dont il parle, on peut penser que cet apprentissage se fait au cours des premiers mois de la vie, à partir de sons naturels de l'environnement sonore.

2. OHGUSHI pour sa part, défend une théorie innéiste, où la perception de la hauteur tonale ne serait pas liée aux stimulus de notre environnement sonore, mais simplement à une caractéristique propre au système auditif, l'équivalence chromatique serait perçue dès la naissance.

3. Enfin, l'acculturation prédit qu'un phénomène tel que celui de l'équivalence chromatique ne peut être mis en évidence sur des sujets âgés de quelques mois seulement.

La question est pour nous de rechercher quand et comment apparaît cette capacité perceptive : nous avons donc choisi d'expérimenter sur de très jeunes enfants, suffisamment jeunes pour n'avoir qu'une expérience musicale très réduite. Une telle expérience doit permettre de tester principalement l'hypothèse de l'acculturation.

Notre problématique peut se résumer ainsi : peut-on montrer l'existence de l'équivalence chromatique chez des sujets sans expérience musicale?

Selon TERHARDT et OHGUSHI, la réponse est oui, selon la thèse de l'acculturation, la réponse est négative.

## 2 - CHOIX DES SUJETS

Dans nos trois expériences (1, 2, et 3) portant sur le nourrisson, nous avons travaillé avec des bébés âgés de 1 à 3 mois. Cette tranche d'âge nous permet à la fois de tester des sujets très jeunes et d'utiliser une technique expérimentale (voir plus loin) bien adaptée à notre problématique et que nous connaissons bien.

On sait que des bébés âgés de quelques heures ont une certaine expérience auditive puisque le fœtus perçoit des stimulations sonores dans le ventre de sa mère. (FORBES et FORBES (1927), in CARMICHAEL (1960), BUSNEL et GRANIER-DEFERRE (1983)).

Disposant d'une technique expérimentale parfaitement adaptée à des sujets de 1 à 3 mois, il était préférable de s'en tenir à cette tranche d'âge.

## 3 - CHOIX DE LA METHODE

Pour toutes les expériences sur le bébé, nous utilisons une même méthode : habituation à un stimulus puis présentation d'un nouveau stimulus.

Nous n'avons pas utilisé le paradigme classique d'habituation puisqu'en fait la phase où l'on présente au bébé le premier stimulus est en réalité une phase de familiarisation (phase F). Dans le cas de l'habituation on attend que la réponse du bébé s'habitue, c'est-à-dire que l'attention qu'il porte au stimulus diminue et ça n'est qu'à ce moment-là que l'on présente le deuxième stimulus. Ici on fait simplement entendre plusieurs fois le stimulus au bébé : c'est la phase de familiarisation, ensuite on lui présente le deuxième stimulus : c'est la phase de test (phase T). On peut ainsi mesurer la réaction du sujet au changement de stimulus : c'est ce que nous appelons la réaction à la nouveauté. Cette réaction peut être nulle si le bébé n'a pas discriminé les 2 stimulus, ou positive (ou négative) dans le cas où il les a différenciés. La réaction à la nouveauté est une mesure de l'attention que porte le bébé aux stimulations présentées.

#### 4 - CHOIX DE LA TECHNIQUE EXPERIMENTALE/CHOIX DE LA REPONSE

L'attention du bébé est mesurée grâce à une technique expérimentale utilisée par BOYD (1974) et DEMANY et al. (1977).

Le principe général de cette technique est le suivant : le bébé est placé face à une cible visuelle; chaque fois qu'il la regarde, on lui fait entendre une séquence sonore. Dès que ses yeux se détournent de la cible visuelle, la séquence sonore s'arrête. Le temps d'écoute du stimulus sonore correspond donc aux temps de fixation de la cible. La fixation du regard, c'est-à-dire la réponse du sujet est le médiateur de l'attention portée au stimulus auditif. Cette technique s'apparente à une sorte de conditionnement opérant où la séquence sonore jouerait le rôle de renforcement. Le nombre de fixations durant une phase est fixé par l'expérimentateur, alors que le temps de fixation est propre à chaque bébé, puisque c'est lui qui la détermine.

L'indice de la réaction au moment du changement de stimulus est calculé à partir des durées de fixation.

#### Calcul de l'indice de réaction à la nouveauté

Ce calcul se fait en réalité à partir des périodes d'observation (P.O.). On distingue deux types de P.O. :

- . P.O. simple : correspondant à une seule fixation de la cible.
- . P.O. composée : correspondant à la somme de plusieurs fixations successives séparées par un intervalle de silence inférieur ou égal à une seconde.  
Le nombre de P.O. est donc inférieur ou égal au nombre de fixations que fait le bébé.

Cette façon de comptabiliser les fixations a été utilisée par DEMANY (1978) qui avait constaté que certains bébés faisaient parfois des fixations extrêmement rapprochées dans le temps. Ce court laps de temps peut correspondre à un simple clignement d'yeux ou à un brusque mouvement de la tête. Si la tête et/ou les yeux reviennent immédiatement à leur position précédente, les 2 fixations sont très rapprochées, l'expérimentateur n'a matériellement pas le temps de

marquer précisément le début de la deuxième fixation. Il y a donc une certaine ambiguïté sur les durées recueillies dans ces conditions; le découpage en P.O. lève une grande partie de cette ambiguïté.

On recueille donc les durées de chaque fixation pour les phases F (Familiarisation) et T (Test).

L'indice de réaction à la nouveauté permet de comparer la durée des P.O. à la fin de la phase F, à la durée de la première P.O. de la phase T, c'est-à-dire juste après le changement de stimulus. Le problème est de savoir à quoi correspond "la fin de la période de la familiarisation". Le nombre de P.O. en phase F est compris entre 12 (ou 13) - nombre maximum de fixations, et 6 - nombre minimum, en dessous duquel on supprime les résultats obtenus.<sup>1</sup> Il s'agit de savoir quelles sont les P.O. qui "représentent" le mieux la fin de la phase de familiarisation. Nous avons choisi de calculer l'indice à partir des 3 dernières P.O. de la phase F

#### Calcul de l'indice

$$\text{INDICE} = \frac{\text{(durée de la 1ère P.O.)}}{\text{(de la phase T)}} - \frac{\text{(moyenne des 3 dernières)}}{\text{(P.O. de la phase F.)}}$$

Si le bébé a perçu le changement de stimulus entre F et T, la durée de la première P.O. en phase T sera significativement plus longue que la moyenne des trois dernières P.O. de la phase F. L'indice sera positif, ce qui marque qu'il y a eu une réaction à la nouveauté. Dans le cas où il n'y a pas de réaction, l'indice est nul.

#### 5 - APPAREILLAGE-PROCEDURE (EXPERIENCES 1, 2 et 3)

Durant l'expérience, le bébé et la personne l'accompagnant, généralement sa mère, ainsi que deux expérimentateurs se trouvent dans la pièce expérimentale. Au début de l'expérience, le bébé est

---

1 - Voir plus loin "Conditions d'élimination".

assis sur une chaise inclinée montée sur un trépied à roulettes. On place la chaise à l'intérieur de la cabine d'observation. Cette cabine en bois a trois côtés<sup>1</sup>. Elle est supportée par quatre pieds de métal de 80 cm de haut. La cabine est fermée sur le haut et est éclairée de façon diffuse par deux ampoules de 45 W. Le bois de la cabine est recouvert d'un tissu bleu. Sur le fond de la cabine (face au bébé) se trouve un petit trou d'observation ("peep hole") de 7 mm de diamètre (Schéma p 47 ). Ce petit trou permet à l'expérimentateur, placé derrière la cabine, d'observer le reflet cornéen d'une cible visuelle sur l'oeil du bébé et de déterminer quand ce dernier regarde la cible. Cette cible se trouve à 20° sur la gauche du bébé; elle est faite d'un rond de tissu blanc de 13 cm de diamètre.

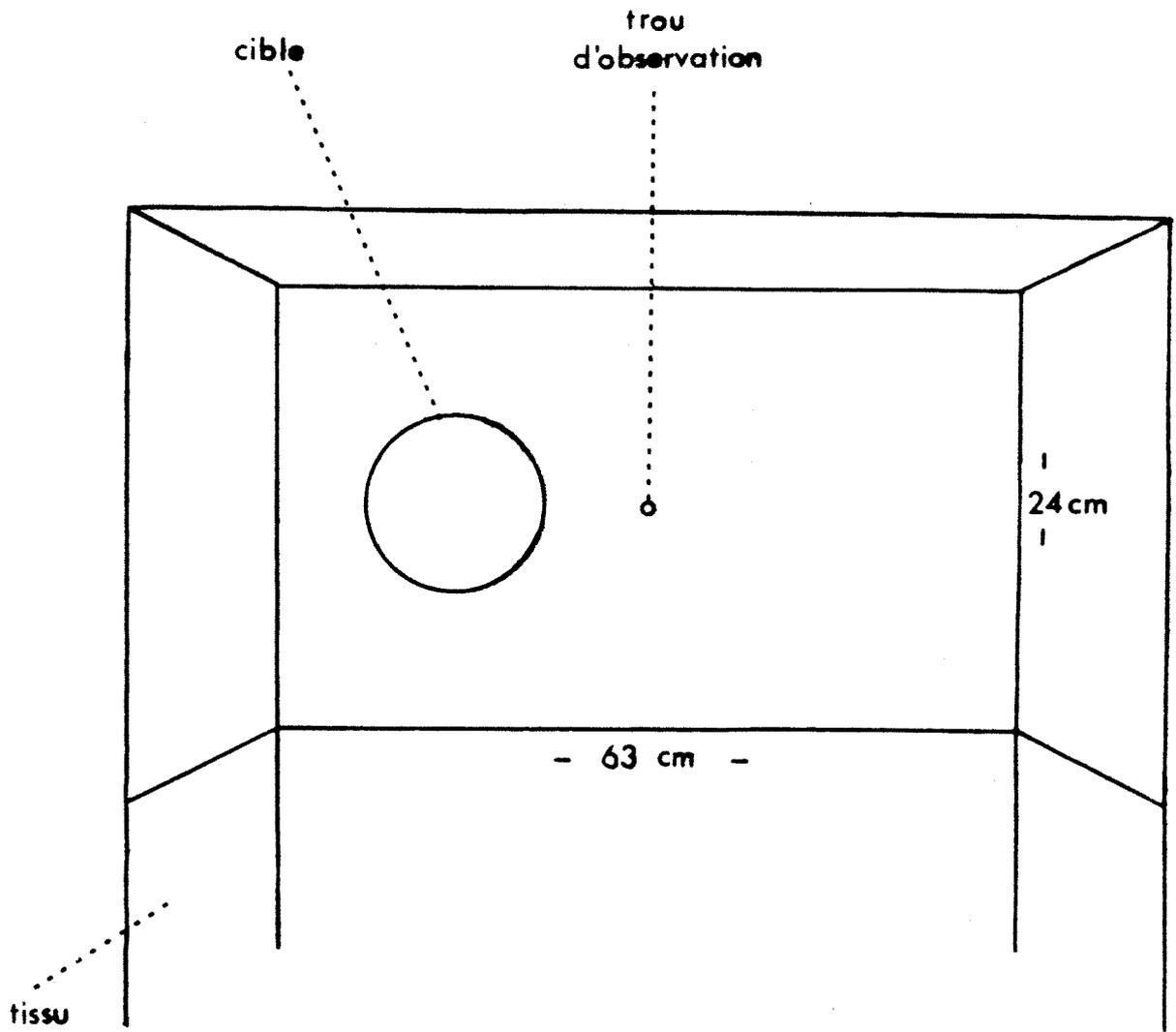
la cabine,  
L'expérimentateur se place derrière face au peep hole; il voit le bébé. A sa gauche est fixé un commutateur ON/OFF, à sa droite un haut-parleur (AUDAK HD 13 B - 12 cm de diamètre) est encastré dans la planche en bois. Ce haut-parleur est caché : du côté du bébé, par la cible visuelle. Enfin, l'expérimentateur porte un casque sur lequel on envoie de la musique permettant de masquer les stimulus sonores sortant du haut-parleur.

Les sons qui composent les séquences mélodiques sont synthétisés par un micro-ordinateur (TRS 80 - Radio Shack avec un convertisseur digital/Analogique). Le bouton ON/OFF est relié à la fois au micro-ordinateur et à un enregistreur à plumes (Sefram) qui permet de recueillir sur papier gradué les durées des fixations du regard du bébé sur la cible et les durées des intervalles entre 2 fixations successives.

Les sons synthétisés par le TRS 80 sont envoyés sur le haut-parleur encastré dans la cabine d'observation via un amplificateur (AKAI AM 2400) et un atténuateur programmable piloté par le TRS 80. Enfin, on a placé un filtre passe-bas (fréquence de coupure : 1000 Hz) afin de supprimer la distorsion créée par l'ordinateur. Un deuxième

---

1 - Utilisée par FANTZ et NEVIS (1967), VURPILLOT et al. (1977), DEMANY et al. (1977).



Ce que voit le bébé.

expérimentateur s'occupe du matériel et surveille le bébé pendant l'expérience. (Voir schéma p. 49 ).

## 6 - DEROULEMENT DE L'EXPERIENCE

Au début de l'expérience, le bébé sur sa chaise est avancé au milieu de la cabine, face au peep hole; la cible est sur sa gauche. Les deux ampoules sont déjà allumées et l'enregistreur à plume est mis en marche. L'expérience peut commencer. Dès que le bébé porte son regard sur la cible, l'expérimentateur met le commutateur ON/OFF en position ON, lorsque la fixation est terminée, le bouton est remis en position OFF.

Le déroulement de l'expérience se décompose en 3 phases :

### - Phase S (ou fixation silencieuse) :

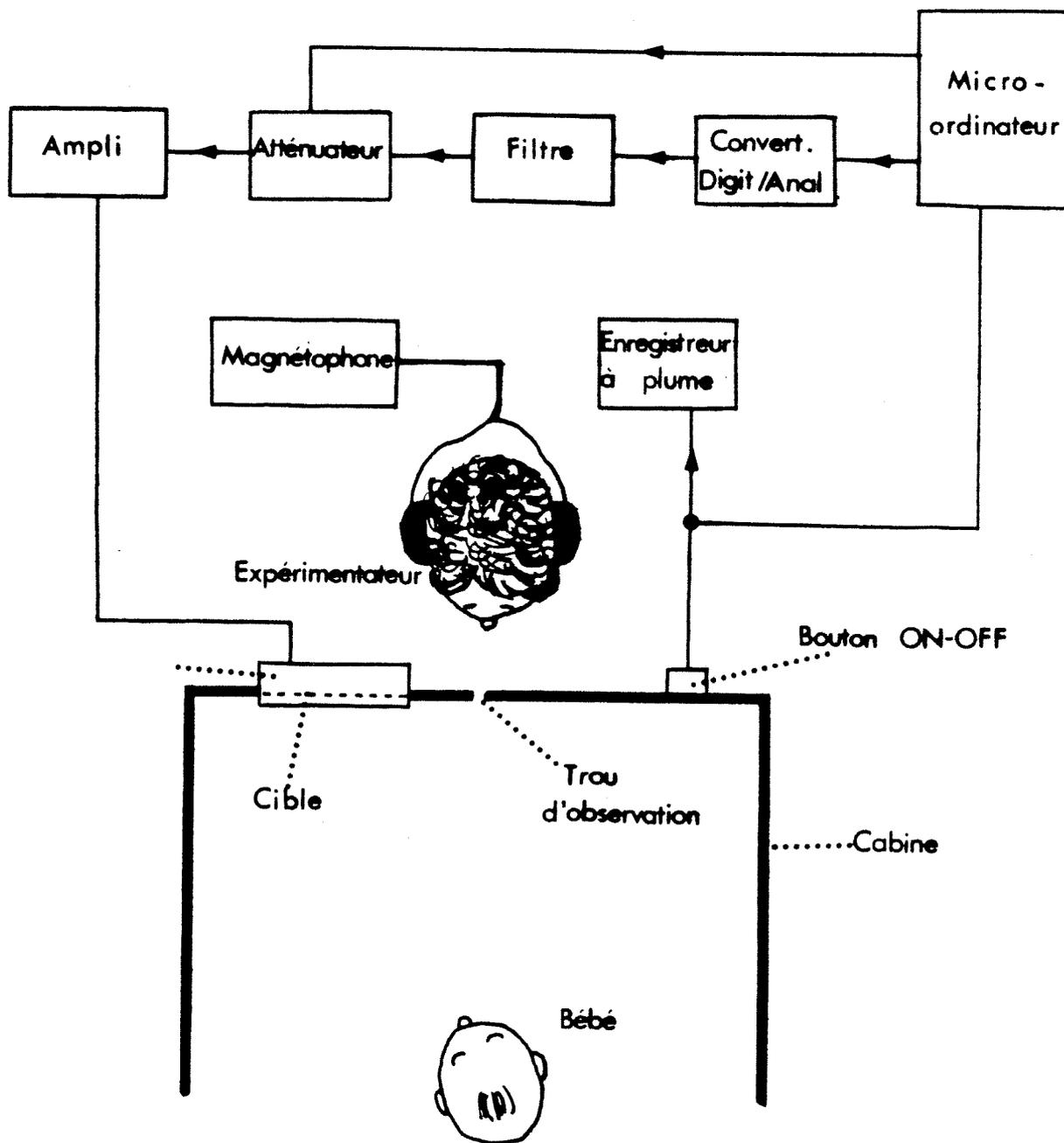
Dès que le bébé regarde une première fois la cible, le bouton est mis sur ON mais aucun son n'est émis, contrairement à ce qui se passe pour les deux autres phases. Lorsque le bébé détourne les yeux, le commutateur est remis sur OFF; c'est ce que nous appelons une fixation silencieuse. Cette première fixation (ce premier essai "à vide" permet au sujet de se familiariser un peu avec son environnement visuel. La durée de cette fixation correspond à l'attention portée à la cible seule au début de la session expérimentale.

### - Phase F (phase de familiarisation) :

La première séquence sonore débute lorsque le bébé fixe la cible et s'arrête dès que le bébé ne regarde plus la cible. Après une douzaine d'essais/de fixations (nous préciserons ce chiffre (12 ou 13) pour chacune des expériences) où le bébé a entendu toujours la même séquence sonore, on remplace cette séquence par une nouvelle séquence différente de la première.

### - Phase T (phase-test) :

Cette phase se déroule comme la phase F : le bébé entend une séquence mélodique dès qu'il regarde la cible. Outre le fait que la séquence n'est pas la même que dans la phase F, il y a une seconde différence entre les 2 phases. En effet, la phase T ne compte que 6 essais, étant donné qu'on ne s'intéresse qu'à ce qui se passe juste après le changement de stimulus, nous n'avons pas besoin de fatiguer trop



DISPOSITIF EXPERIMENTAL

longtemps le bébé en multipliant les essais dans cette phase T.

En résumé, il y a 3 phases différentes :

1. Phase S : une fixation "silencieuse".
2. Phase F : une douzaine d'essais au cours desquels le bébé se familiarise avec le premier stimulus présenté sous la forme d'une séquence mélodique.
3. Phase T : 6 essais avec un stimulus différent de celui de la phase F mais toujours présenté sous la forme d'une séquence mélodique.

La fin de la phase T marque la fin de l'expérience.

#### 7 - CONDITIONS D'ELIMINATION DES SUJETS

Toutes les données n'ont pas été retenues, nous avons déterminé six conditions dans lesquelles les résultats du sujet seront rejetés.

1. Le bébé pleure durant la phase F :  
l'expérience est arrêtée si les pleurs ne s'arrêtent pas.
2. Le bébé se met à pleurer au moment du changement :  
(entre les phases F et T).
3. Un intervalle de silence entre deux fixations a une durée supérieure à deux minutes.

Le paradigme d'habituation mettant en jeu la mémoire du sujet, on peut penser qu'une "pause" de deux minutes au sein de la période F peut avoir un effet inhibiteur sur la mémorisation et sur la familiarisation avec le stimulus présenté.

De plus, BRIDGER (1961) montre que le nombre d'essais nécessaires à l'habituation augmente proportionnellement avec la longueur des intervalles de silence entre deux présentations d'un stimulus.

En éliminant les résultats des sujets qui font au moins un intervalle de silence supérieur à deux minutes, on homogénéise la phase de familiarisation pour tous les sujets.

4. Le bébé fait seulement six P.O. (périodes d'observation) en phase F.

On distingue deux types de P.O. :

- 1 - P.O. simple : correspondant à une seule fixation.
- 2 - P.O. composée : égale à la somme de plusieurs fixations successives séparées par un intervalle de silence inférieur ou égal à une seconde.

Le nombre de P.O. est inférieur ou égal au nombre de fixations que fait le bébé. Durant la phase F, le nombre de fixations (d'essais) est égal à 12 ou 13; si le nombre de P.O. est inférieur à 6, on élimine les résultats du sujet. Cette condition permet de supprimer les sujets qui ont une trop courte période de familiarisation.

#### 5. Erreurs de procédure :

Au cours de l'expérience, il peut se produire des erreurs de manipulation de la part des deux expérimentateurs, ou des dysfonctionnements des appareils utilisés : dans ces cas, l'expérience est considérée comme nulle.

Si au moins une de ces six situations se produit, les résultats obtenus ne sont pas pris en compte.

#### 8 - STIMULUS

Tous les sons présentés dans nos quatre expériences sont des sons purs, synthétisés par un micro-ordinateur. Nous avons préféré les sons purs aux sons complexes pour éviter que les sujets n'entendent des harmoniques, et donc des intervalles d'octave dans les stimulus présentés.

Le choix des sons a été délicat, étant donné qu'on ne connaît que très peu de choses sur les capacités auditives des bébés. D'après EISENBERG (1976), la fonction de la cochlée peut être mise en évidence dès le 5ème mois de la vie foetale. A cet âge, en effet, l'oreille interne et l'oreille moyenne ont atteint leur structure adulte ainsi que l'oreille externe, qui elle, va continuer à se développer jusqu'à l'âge de 9 ans (voir NORTHERN et DAVIS, p.26; 1978). On peut penser que lorsqu'il vient au monde, le bébé a déjà une certaine expérience auditive.

Pour le choix des stimulus, il faut déterminer :

- le niveau d'intensité sonore (en dB SPL)
- et la fréquence de chacun des sons utilisés.

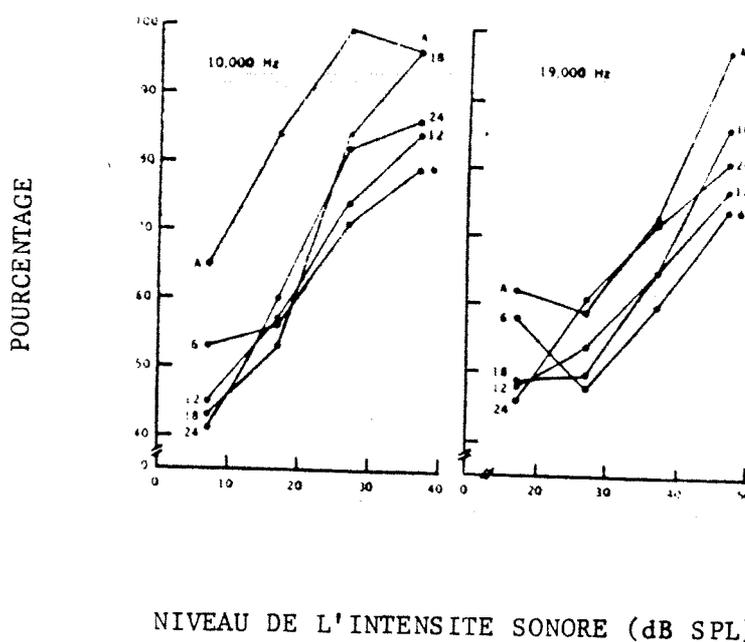
Enfin, si on utilise deux sons proches en fréquence, il faut connaître le seuil différentiel de nos sujets pour n'employer que des sons bien discriminés.

### 8.1. Choix de l'intensité

Voyons d'abord un résultat de BARTOSHUK (1964) obtenu sur des nouveaux-nés âgés de 24 à 100 heures.

On leur fait entendre des sons de 1000 Hz, avec une amplitude de : 38, 47.5, 52 ou 67.5 dB. Pour ces quatre valeurs, on mesure les modifications du rythme cardiaque du bébé. Les résultats indiquent que le seuil auditif est compris entre 38 dB et 47.5 dB. TAGUSHI (1969), cité par SCHNEIDER et al. (1979) emploient des sons intermittents de 500, 1000 et 2000 Hz. Pour des nouveaux-nés de 0 à 2 jours, le seuil auditif d'intensité sonore se situe entre 55 et 58 dB. Ce qui fait une différence importante avec les résultats de BARTOSHUK. Pour des bébés de 4 jours, TAGUSHI et al. notent que la valeur du seuil diminue (40 dB), puis continue à chuter (30 dB pour des bébés de 5 mois).

Une expérience plus récente de SCHNEIDER et al. (1980) donne des résultats sur des sujets de 6, 12, 18 et 24 mois, plus un groupe de sujets adultes. La réponse est la même pour les cinq groupes : l'orientation de la tête vers la source sonore. Les stimulus sont deux bandes de bruit d'une demi-octave de large, centrées sur 10 KHz pour l'une et 19 KHz pour l'autre. L'intensité sonore des stimulus prend 4 valeurs différentes : 7, 17, 27 et 37 dB pour 10 KHz, et 17, 27, 37 et 47 dB pour 19KHz. La courbe, page 53, montre d'abord que la différence entre les bébés (de 6 à 24 Mois) et les adultes n'est sensible que pour 10 KHz.



Pourcentage d'orientations correctes de la tête en fonction du niveau sonore des 2 stimulus pour les 4 groupes de sujets "Bébés" et le groupe de sujets "Adultes" (A). SCHNEIDER et al. (1980).

Pour le stimulus de 10 KHz, le seuil absolu des bébés est plus élevé que celui des adultes. Néanmoins, il faut noter que pour des intensités sonores inférieures à 10 dB, ce sont les sujets les plus jeunes (6 mois) qui obtiennent les meilleurs scores au sein du groupe "Bébés" aussi bien pour le stimulus de 10 KHz que pour celui de 19 KHz.

Il est relativement difficile de tirer une conclusion précise à partir de ces différentes données; on peut avancer néanmoins que des sons à un niveau sonore compris entre 50 dB et 80 dB sont nettement perçus par des bébés de 3 mois.

### 8.2. Choix de la fréquence

Pour ce choix, nous pouvons retenir deux données principales. La première montre que des nouveaux-nés perçoivent des sons de 125 Hz; pour cette expérience, les auteurs (HUTT et al. 1968) ont utilisé la réponse cardiaque. Cette donnée ne représente pas le seuil inférieur de la perception de la fréquence d'un son, il est probable en fait que le bébé entend des sons beaucoup plus graves encore. Pour le seuil supérieur en fréquence, nous allons revenir à l'expérience de SCHNEIDER et al. (1980). Ces résultats montrent que des bébés de 6 mois perçoivent les stimulus de 10 KHz à 20 dB. Ces 2 valeurs (125 Hz et 10 KHz) donnent une idée de ce que doit être l'aire d'audition d'un bébé de quelques mois. Malheureusement, la plupart des expériences faites jusqu'ici ne donnent que des valeurs approchées des seuils absolus.

Cette remarque s'applique aussi aux recherches faites sur le seuil différentiel en fréquence.

### 8.3. Seuils différentiels en fréquence

En 1925, KASATKIN et LEVIKOVA ont étudié deux groupes de bébés (2 mois 1/2 et 3 mois 1/2) en employant une méthode de conditionnement de réflexe alimentaire (taux de sécrétion salivaire). Par cette méthode, ils ont montré que les sujets discriminaient deux sons complexes séparés par 11 1/2 tons, c'est-à-dire un peu moins d'une octave.

BRIDGER (1961) obtient une discrimination entre deux sons purs de 400 Hz et 1000 Hz, mais uniquement pour 15 nouveaux-nés (1 à 5 jours) sur 50. Il utilise une technique d'habituation de la réponse cardiaque.

LEVENTHAL et LIPSITT (1964) par contre, n'arrivent pas à mettre en évidence la discrimination chez des nouveaux-nés entre deux sons purs de 200 Hz et 500 Hz, avec une autre technique où la réponse observée est le rythme de succion. En utilisant une méthode de conditionnement du rythme de succion, WORMITH et al. (1975) trouvent que des bébés de 35 jours distinguent deux sons purs de 200 et 500 Hz. En fait, ces résultats ne témoignent pas vraiment du seuil différentiel

réel en fréquence car l'écart entre les sons à discriminer est considérable et les auteurs n'ont pas fait varier systématiquement cet écart. Le résultat le plus précis que nous connaissons est malheureusement un résultat négatif et donc difficile à interpréter. L'expérience (TREHUB, 1977) a été faite sur des bébés de 5 mois, avec une méthode d'habituation de la réponse cardiaque. Les stimulus sont des fréquences sonores composées de six sons purs choisis au hasard dans une gamme de fréquences allant de 262 Hz à 1046 Hz. La durée d'un son est égale à 200 msec. ainsi que la durée du silence entre deux sons. Les sujets sont répartis en deux groupes : un groupe expérimental, un groupe contrôle. Dans le premier, on présente une séquence de 6 sons purs (séquence originale) 15 fois de suite, c'est la période de familiarisation, puis 4 fois cette même séquence, mais transposée  $3 \frac{1}{2}$  demi-tons au-dessus de l'originale pour la première moitié du groupe, et  $3 \frac{1}{2}$  demi-tons au-dessous pour la deuxième moitié (c'est la période de test). Pour le groupe contrôle, la période d'habituation est la même que pour le groupe expérimental, la séquence-test, elle, est composée des mêmes sons que la séquence originale, mais placés dans un ordre différent, exception faite du premier son de la séquence qui conserve sa place. L'enregistrement du rythme cardiaque montre qu'il n'y a pas de réaction au changement de stimulus dans le premier groupe, alors que les bébés du groupe contrôle ont bien discriminé les deux séquences présentées. D'après ce résultat, il semblerait que des bébés de 5 mois ne différencient pas deux sons séparés par un intervalle de trois demi-tons (c'est-à-dire un intervalle de tierce mineure). Mais on remarque dans cette expérience que l'intervalle de silence entre la présentation des deux séquences est de 15 secondes, ce qui laisse peut-être le temps au bébé d'oublier les fréquences absolues de chaque son, alors que le contour de la séquence (modifié pour le groupe contrôle) a toutes les chances d'avoir été mémorisé. Pour le groupe expérimental, l'information sur le contour restant inchangée, et la taille relative des intervalles successifs étant préservée; la modification de la tonalité de départ peut paraître un facteur négligeable. Sans être confondues, les deux séquences mélodiques présentées dans le groupe expérimental ont pu sembler équivalentes; ce qui pourrait expliquer le fait qu'il n'y ait pas eu de réaction au changement. Bien que cette expérience apporte un résultat difficile à analyser, elle est

intéressante dans la mesure où elle n'utilise pas des sons isolés comme dans la plupart des expériences, mais des séquences mélodiques qui constituent peut-être des stimulus plus attractifs pour de très jeunes enfants.

La seule expérience, à notre connaissance, qui utilise des sons proches en fréquence, est celle de DEMANY (1982), où l'on montre que des bébés de 2 à 3 mois discriminent des sons carrés séparés par un intervalle de un ton.

#### 8.4. Discrimination de sons séparés par un ou plusieurs intervalles d'octave

A ce sujet, plusieurs résultats sont à noter. Le premier est déjà ancien (STUBBS, in CARMICHAEL, 1960). Dans cette expérience, on observe l'activité motrice générale et l'ouverture des yeux chez des sujets nouveaux-nés, en leur présentant des sons purs de fréquences différentes (128, 256, 1024 et 4096 Hz). On remarque que les trois dernières fréquences correspondent aux première, troisième et cinquième octaves supérieures du son de 128 Hz. Ces quatre sons sont donc séparés par un nombre entier d'intervalles d'octave. STUBBS n'arrive pas à montrer que ces quatre fréquences sont distinguées les unes des autres. Si l'on suppose que dans cette expérience, les sons étaient purs, sans distorsion, alors on peut dire, soit que les capacités discriminatives du nouveau-né ne lui permettent pas de distinguer ces quatre fréquences, soit qu'une grande similarité a été perçue entre les sons à l'octave, ce qui voudrait dire que les bébés sont sensibles à l'équivalence chromatique. Malheureusement, rien dans l'expérience ne nous permet de trancher entre ces deux hypothèses. L'expérience de BENCH (1969) montre pourtant que deux sons séparés par deux octaves (500 Hz et 2000 Hz) sont bien discriminés par des nouveaux-nés. Enfin, l'expérience de TREHUB et al. (1982) - voir page 12 - montre que des sons à l'octave sont bien discriminés par des bébés de 8 à 11 mois, mais, nous avons vu que les résultats obtenus étaient à prendre avec prudence.

Ainsi, la question de la discrimination des sons à l'octave et de la similarité chromatique chez des sujets plus jeunes reste posée.

° Cité par SCHNEIDER et al. (1979).

## EXPERIMENTATION

La partie expérimentation se divise en deux :

- expériences 1, 2 et 3 sur des bébés de 2 à 3 mois.
- expérience 4 sur des adultes.

La méthode et la procédure expérimentale que nous avons déjà décrites s'appliquent aux trois premières expériences ; nous allons voir maintenant en détails comment se déroule chacune de ces trois expériences.

### 1 - EXPERIENCE 1

#### 1.1. Problématique

Cette première expérience porte sur la discrimination d'intervalles.

On présente à des bébés de 3 mois deux intervalles sous forme de séquences mélodiques. Dans un cas, les deux intervalles sont différents de l'octave, dans l'autre cas, l'un des intervalles est une octave et l'autre pas.

Dans ce deuxième cas, on s'attend à ce que la réaction au changement de stimulus soit plus grande que dans le premier cas, si on admet l'hypothèse de l'existence de l'équivalence chromatique, après avoir construit les stimulus que nous voulions utiliser, nous nous sommes aperçus que la pièce expérimentale dans laquelle nous allions travailler, introduisait des différences d'intensité entre nos divers sons (mesures faites au sonomètre Brüel et Kjaer 2209).

Il a donc fallu égaliser les intensités subjectives de tous les sons utilisés en modifiant séparément l'amplitude de chacun.

Nous avons profité du temps qu'il nous a fallu pour résoudre ce problème (construction d'un atténuateur programmable) pour commencer notre expérience 1, tout en sachant que les sons différaient à la fois par la hauteur et par l'intensité. L'expérience 1

a véritablement débuté au moment où nous avons disposé de sons égaux en sonie. Nous avons pensé qu'il était instructif de présenter les résultats de ce que nous appellerons la pré-expérience 1.

## 1.2. Pré-expérience 1

### 1.2.1. Plan expérimental

Au départ, il avait été prévu deux conditions expérimentales :

- . une condition Octave (où l'on donne à comparer un intervalle d'octave et un intervalle différent de l'octave -  $\pm$   $3/4$  de ton)
- . une condition non Octave (où l'on donne à comparer deux intervalles différents de l'octave; l'un de  $+ 3/4$  de ton, l'autre de  $- 3/4$  de ton) avec deux groupes de sujets dans chaque condition. Le plan n'a pas été totalement réalisé; seuls les deux groupes de la condition Octave ont été examinés.

Chaque groupe entend deux séquences mélodiques différentes (une séquence durant la phase de familiarisation, une séquence durant la phase-test) (schéma p. 61). Les séquences mélodiques sont composées de deux sons qui alternent continûment.

### 1.2.2. Les stimulus

On appellera stimulus, le groupe de deux notes qui forme l'intervalle de base sur lequel chaque séquence sera construite.

#### a/ Fréquence des sons

Trois séquences différentes ont été utilisées à partir des quatre sons suivants :

- a. 480 Hz
- c 262 Hz
- d. 240 Hz
- c' 220 Hz (Voir schéma p. 61)

Nous avons désigné chaque séquence par le nom de l'intervalle qui la compose.

Entre les séquences de familiarisation a - c et a - c' et la séquence-test a - d, la différence est de  $3/4$  de ton. Ce choix est arbitraire dans la mesure où aucune expérience

(sauf peut-être DEMANY, 1978) ne permet d'être sûr que le seuil différentiel des bébés âgés de 3 mois soit aussi bas.

#### b/ Intensité des sons

On a vu que l'on pouvait fixer l'intensité de nos sons à une valeur située dans une plage allant de 50 à 90 dB pour être sûrs que les sons soient nettement perçus par nos sujets. Ce choix ne posait qu'un seul problème lié aux sonies (intensités subjectives) de sons n'ayant pas la même fréquence. En effet, on sait (FLETCHER et MUNSON, 1927) que deux sons de fréquences différentes, mais d'amplitudes égales peuvent avoir des sonies différentes : c'est-à-dire que l'un des deux sons paraîtra plus fort que l'autre. Le son a (480 Hz) a une intensité sonore de 76 dB SPL, les sons c, d et c' ont une intensité de 82,5 dB SPL.

#### c/ Durée des sons

La durée d'une séquence sonore (i.e. la durée d'une fixation de la cible) dépend du bébé, mais il faut déterminer la durée de chaque son et la durée des intervalles de silence entre les sons. La durée des sons (et des silences) doit répondre à une double contrainte :

1. Les sons doivent être suffisamment séparés dans le temps pour ne pas être confondus les uns avec les autres.
2. Les sons et les silences doivent être suffisamment courts pour que le sujet entende le plus de stimulus possible au cours d'une seule fixation.

Pour ce choix, nous nous sommes basés principalement sur une recherche de DEMANY (1978) où des sons de 40 ms sont séparés par des intervalles de silence de 97 ms.

Dans notre expérience 1, tous les sons ont la même durée :  $437,1 \pm 1$  ms (ce qui nous place très au-delà des valeurs de DEMANY,<sup>1</sup>

(Le temps de montée et de descente de chaque son est égal à 5,05 msec., la durée de la période stable est égale à  $427 \pm 1$  ms). La durée de chaque silence entre deux sons est fixée à  $73 \pm 1$  ms.

---

1 - et ce qui fait que nos sons seront parfaitement perçus par les sujets).

#### d/ Les séquences

Les séquences sonores utilisées sont toutes constituées sur le même modèle : elles sont basées sur l'alternance de deux notes, par exemple ici A et B :

B - A - B - A - B - A - B - A ...

A chaque fixation de la cible par le bébé, la séquence commence par le son d (c ou c'). Alors que chaque séquence doit se terminer par le son a, le son commun à toutes les séquences. Ceci veut dire que si la fixation d'un bébé s'achève alors qu'il entend le son d (c ou c'), la séquence se poursuit jusqu'au son a suivant. Toutes les séquences entendues ont donc la même forme.

Il y a trois séquences différentes (cinq séquences avaient été prévues dans le plan initial) - voir schéma p.61.

Il y a une séquence de familiarisation d-a et deux séquences-test c-a et c'-a; pour le seul groupe testé (groupe octave).

Sujets : tous les sujets pris pour cette expérience sont des nourrissons nés à terme avec un poids minimum de 3 Kg à la naissance. Ils sont âgés de 84 jours en moyenne (écart-type : 28 jours).

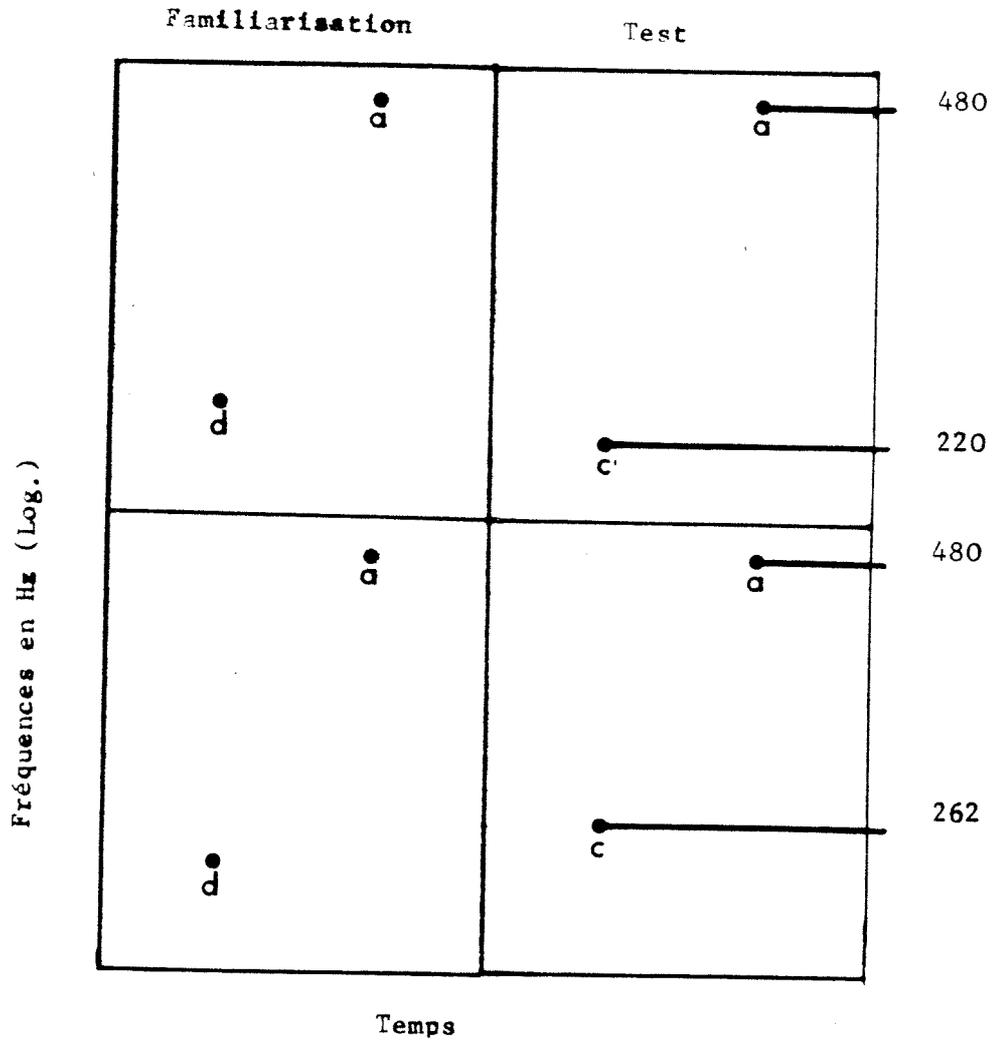
#### 1.2.3. Sujets

Comme nous l'avons vu (p. 57) l'expérience se divise en deux parties. En tout 43 sujets ont été testés : 17 pour la pré-expérience 1, 26 pour l'expérience proprement dite.

##### a/ Pré-expérience 1

Dans cette partie, sur les 17 sujets testés, 5 seulement ont été gardés. Parmi les 12 non retenus, 5 ont été éliminés pour avoir fait un intervalle de silence supérieur à deux minutes pendant la phase F - 4 pour pleurs - 1 pour difficultés d'observation - 2 pour des erreurs de procédure.

Tous les sujets ont été pris dans le groupe Octave, c'est-à-dire qu'ils ont entendu une séquence de familiarisation a - d et une séquence-test a - c pour une moitié des sujets, a - c' pour l'autre moitié.



STIMULUS UTILISES POUR LES SEQUENCES  
DU GROUPE OCTAVE

Expérience 1

### b/ Expérience 1

Pour l'expérience 1, 26 sujets ont été testés, 12 ont été retenus. Sur les 14 sujets non retenus, 10 l'ont été pour des pleurs, 2 pour avoir fait un intervalle supérieur à deux minutes pendant la phase F, 2 pour des erreurs de procédure.

#### Procédure

La phase de familiarisation compte 12 fixations successives de la cible, c'est-à-dire que le bébé entend 12 fois la séquence mélodique de familiarisation a - d .

A partir de la 13<sup>ème</sup> fixation, la séquence mélodique change; c'est la phase-test qui commence.

Cette phase compte au minimum six P.O.

L'expérience s'arrête soit lorsque l'expérimentateur le décide pour ne pas fatiguer inutilement le bébé, soit parce que le bébé se met à pleurer.

Pour chaque bébé on recueille sur papier gradué les durées des fixations et des silences entre deux fixations successives.

Lorsque le regroupement en P.O. est fait, on calcule l'indice de réaction à la nouveauté pour chaque sujet.

$$I = \frac{(\text{durée de la première})}{(\text{P.O. de la phase T})} - \frac{(\text{moyenne des 3 dernières})}{(\text{P.O. de la phase F})}$$

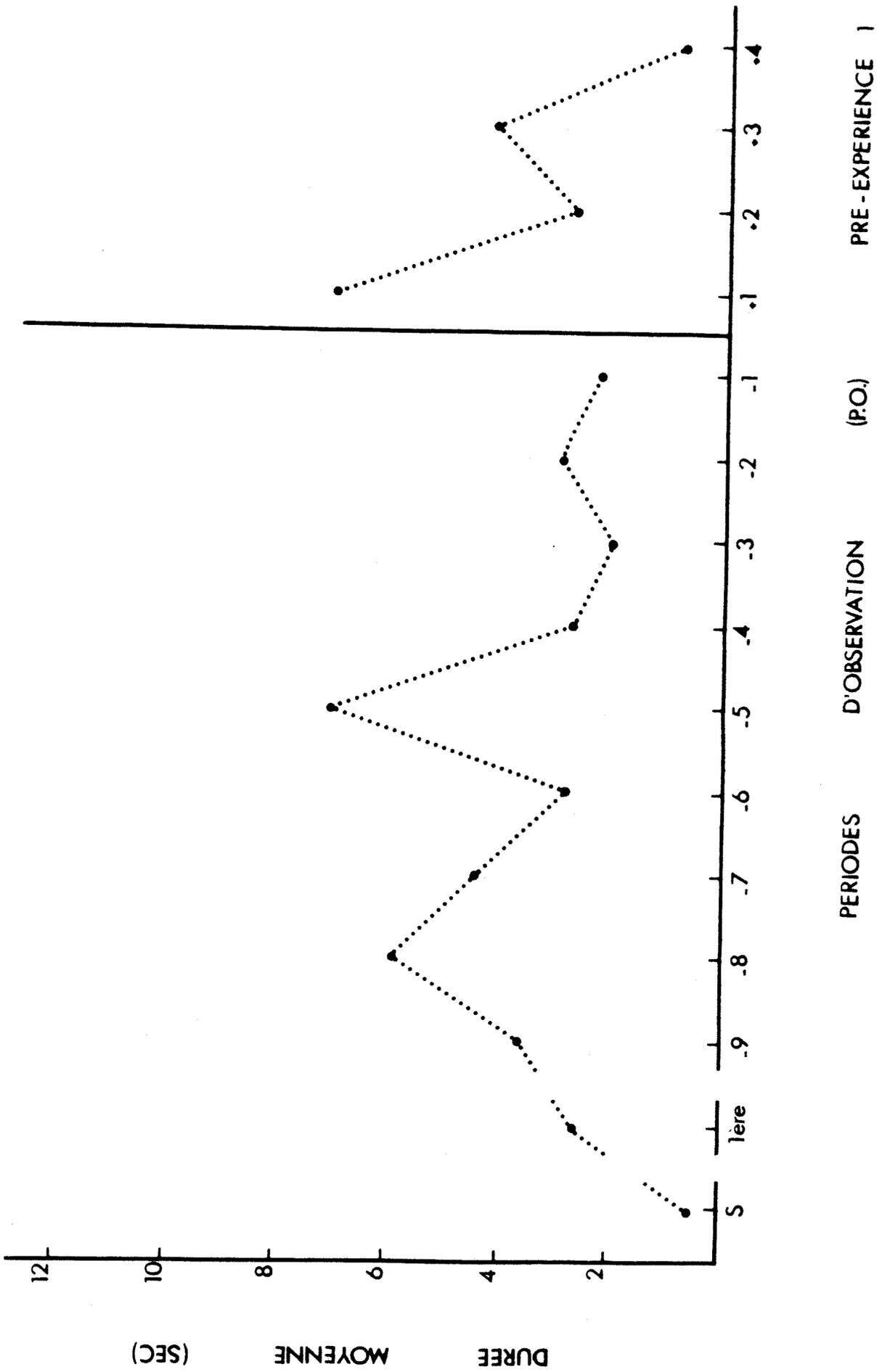
### 1.3. Résultats

Avant de nous intéresser aux données de l'expérience 1 proprement dite, nous allons regarder les résultats de ce que nous avons appelé la pré-expérience 1.

#### 1.3.1. Pré-expérience 1

Dans cette partie de l'expérience, les stimulus entre la phase F et la phase T différaient à la fois en fréquence et en intensité. Les 5 sujets retenus sont passés dans la condition "Octave" : 3 dans le groupe où l'on présente la séquence d - a puis c - a , 2 dans le deuxième groupe Octave avec d - a puis c' - a . La différence d'intensité entre les sons d et c est de 3 dB; entre d et c' , elle est de 6 dB.

La courbe de la page 63 permet de voir deux choses : d'abord il y a une importante réaction à la nouveauté, au



changement de stimulus ( $t(4) = 2.5$ ;  $P < .05$  - test unilatéral). Les sujets ont donc discriminé des sons qui différaient à la fois par la fréquence (3/4 de ton) et par l'intensité (3 et 6 dB). On ne peut pas dire si la discrimination se fait à cause des différences d'intensité ou des différences de fréquences.

La deuxième chose à remarquer concerne la phase de familiarisation : la courbe montre qu'il n'y a pas d'habituation de la réponse, il y a en effet une remontée de l'attention à la P.O. 5 (durée de cette P.O. : 8 sec.).

A la fin de cette phase, l'attention du bébé chute brutalement (la durée des 4 dernières P.O. est égale environ à 3 sec.).

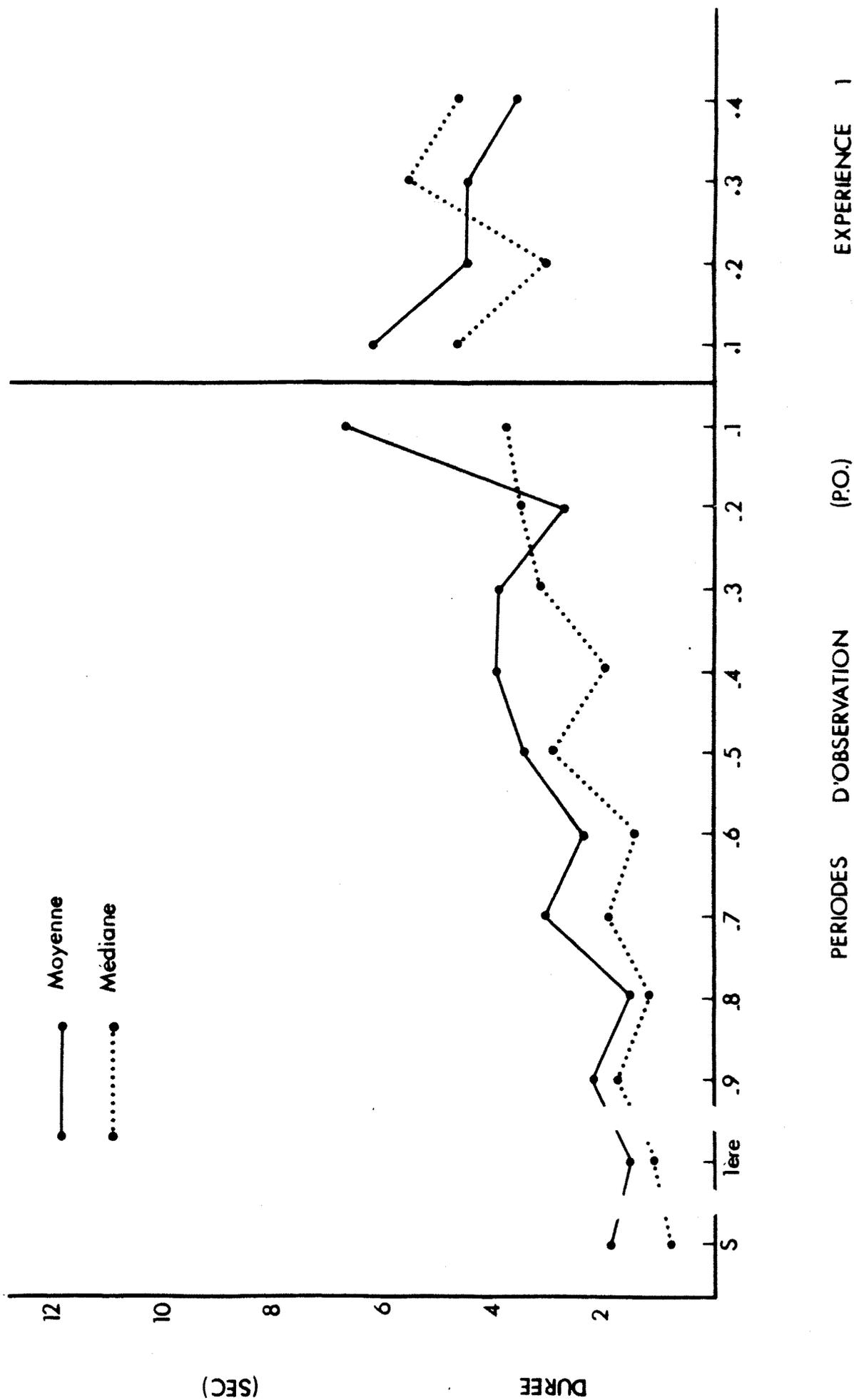
Etant donné le faible nombre de sujets, nous nous en tiendrons à ces deux remarques concernant la pré-expérience 1.

Voyons maintenant les résultats de l'expérience 1 faite après égalisation des sonies de tous les sons utilisés.

### 1.3.2. Expérience 1

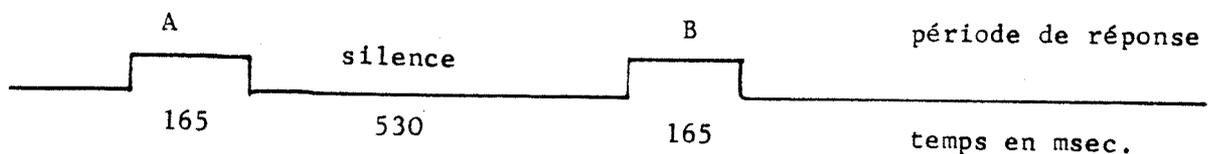
La courbe (p. 65) montre nettement qu'il n'y a pas de réaction à la nouveauté lors du changement de stimulus. Cela voudrait donc dire que les sujets n'ont pas fait de différences entre les deux séquences. La première conclusion que l'on peut tirer est que les bébés de 3 mois n'ont pas discriminé deux sons séparés par 3/4 de ton : ainsi les deux séquences présentées leur sont apparues comme rigoureusement identiques. Si on regarde les moyennes des durées des P.O. dans la phase F, on s'aperçoit qu'il n'y a pas d'habituation, plus encore, il y a une augmentation importante de la durée des périodes d'observation de la première à la dernière pendant la phase F. On peut penser que cette montée de l'attention a pu "masquer" en quelque sorte une éventuelle réaction à la nouveauté après le changement de stimulus.

En examinant la courbe des médianes (schéma p. 65), on remarque que l'augmentation de la durée de fixation n'apparaît plus, ce qui voudrait dire que cette augmentation ne s'est produite que chez quelques sujets. En effet, si on regarde de plus près les données individuelles, cette hypothèse est confirmée : 3 sujets



sur 12 ont un temps de fixation extrêmement long pour la P.O. - 1 (11,5 - 13 et 24,6 secondes), alors que les temps pour les 9 autres sujets sont compris entre (0,8 sec. et 6,45 sec.). La montée de l'attention ne semble être qu'un comportement déviant pour quelques sujets, s'il y a une réaction à la nouveauté, elle ne peut qu'être masquée par l'effet de ce comportement.

Si l'on admet cette interprétation et que l'on supprime les données des 3 sujets "déviants", il faut avouer que cela ne suffit pas à mettre en évidence une réaction à la nouveauté. Il faut alors chercher une troisième interprétation pour expliquer ce résultat. Si on suppose que les bébés discriminent bien deux sons séparés par  $3/4$  de ton, alors on doit admettre que les stimulus utilisés n'ont pas permis d'éveiller l'attention des sujets. Certains auteurs ont montré en travaillant sur l'équivalence chromatique que des facteurs tels que la procédure, le type de stimulus employé, le contexte musical dans lequel se trouve le stimulus qui sert de test (CUDDY, 1979) pouvaient influencer sur les résultats. Rappelons qu'ici, dans l'expérience 1, nous n'utilisons que deux sons alternés. Or, dans des études antérieures, (ATTNEAVE et OLSON, 1971 - THURLOW et ERCHUL, 1977), citées précédemment, le même type de stimulus a été utilisé (intervalles musicaux à reconnaître et à reproduire), les auteurs ont noté qu'avec de tels stimulus, répétition de deux notes successives, il était très difficile de mettre en évidence le phénomène de l'équivalence chromatique. KALLMAN (1982) a récemment montré à quel point il est difficile pour des adultes de percevoir une similarité entre deux sons isolés dans de telles conditions. Les stimulus qu'il utilise sont très semblables à ceux que nous employons dans l'expérience 1. L'expérience de KALLMAN (postérieure à la nôtre) se déroule de la façon suivante : le sujet entend deux sons purs successifs A et B



La tâche du sujet est d'estimer un degré de similarité entre ces deux notes; pour cela il dispose d'un potentiomètre linéaire qu'il règle en fonction de la similarité qu'il a perçue. Les sons A et B sont séparés par un intervalle qui va de 0 à 28 demi-tons par pas de 1/2 ou 1 ton. Les résultats ne mettent pas en évidence une relation particulière lorsque les deux sons sont séparés par un intervalle d'octave. Le degré de similarité varie de façon monotone, proportionnellement à la taille de l'intervalle : plus l'intervalle est petit, plus la similarité est grande. Dans une autre partie de cette même expérience, les auteurs utilisent des séquences de trois sons : dans ce cas, un son reste inchangé; pour les deux autres sons, cela se passe comme dans la première partie où il n'y avait que deux sons. Le sujet doit à nouveau estimer un degré de similarité entre deux sons au sein de la séquence présentée.

Avec ce nouveau stimulus, KALLMAN arrive à mettre en évidence le phénomène de l'équivalence chromatique.

De ce résultat, on peut tirer un enseignement précieux pour la suite de nos expériences : en ce qui concerne l'expérience, on peut supposer que la simplicité des séquences (2 sons alternés) a été un handicap pour mettre en évidence le phénomène de la similarité chromatique chez nos sujets.

Tout en gardant la procédure utilisée, nous pouvons poursuivre cette expérience, mais en employant cette fois des séquences sonores plus élaborées. Le fait que beaucoup de bébés aient pleuré au cours de l'expérience l'appuie notre décision de changer de stimulus. Dix sujets sur 26 ont en effet été éliminés pour s'être mis à pleurer au cours de la période de familiarisation : il semble assez probable que nos stimulus soient mal adaptés aux sujets et à la problématique.

## 2 - EXPERIENCE 2

Cette expérience est menée de la même façon que la précédente; seuls les stimulus sont différents.

### 2.1. Les stimulus

Comme l'a fait KALLMAN, nous avons modifié la séquence

sonore en rajoutant un son au stimulus de base. Le nouveau stimulus de base se compose maintenant de trois sons. La séquence mélodique est de la forme A-B-C-A-B-C ... .

Rajouter un quatrième son aurait amené un certain nombre de contraintes sur la durée des sons. En effet, nous ne contrôlons pas le temps d'une fixation de la cible par le bébé; il faut donc prévoir un stimulus le plus court possible pour qu'il y ait le plus de stimulus c'est-à-dire de triolets A-B-C entendus au cours d'une seule fixation.

Pour un même sujet, la séquence mélodique de la période de familiarisation est de la forme A B C A B C, ... , alors que la séquence de la phase de test est construite de la façon suivante

A B' C' A B' C' ... , ce qui veut dire qu'on conserve la première note du triolet A et qu'on transpose uniquement les deux notes de l'intervalle B-C. Cette transposition se fait à l'octave pour le groupe "Octave", et à un intervalle différent mais proche de l'octave pour le groupe "Non-Octave" (pour la moitié de ce groupe, l'intervalle choisi est plus grand que l'octave, pour l'autre moitié, l'intervalle est plus petit.) (Voir schéma p. 69).

#### 2.1.1. Durée des sons

Tous les sons ont la même durée fixée à 281,1 ms (le temps de montée et le temps de descente sont égaux à 5,05 ms. La durée du silence entre deux sons est égale à 94,4 ms.)

Le stimulus de base que nous définirons comme l'ensemble de trois sons (plus deux silences) a une durée d'environ 1030 ms, à peine plus d'une seconde.

#### 2.1.2. Fréquence des sons

Les quatre séquences sont composées à partir de neuf sons de fréquences différentes :

A : 737 Hz

B : 487 Hz      B-7 : 273 Hz      B-8 : 244 Hz      B-9 : 218

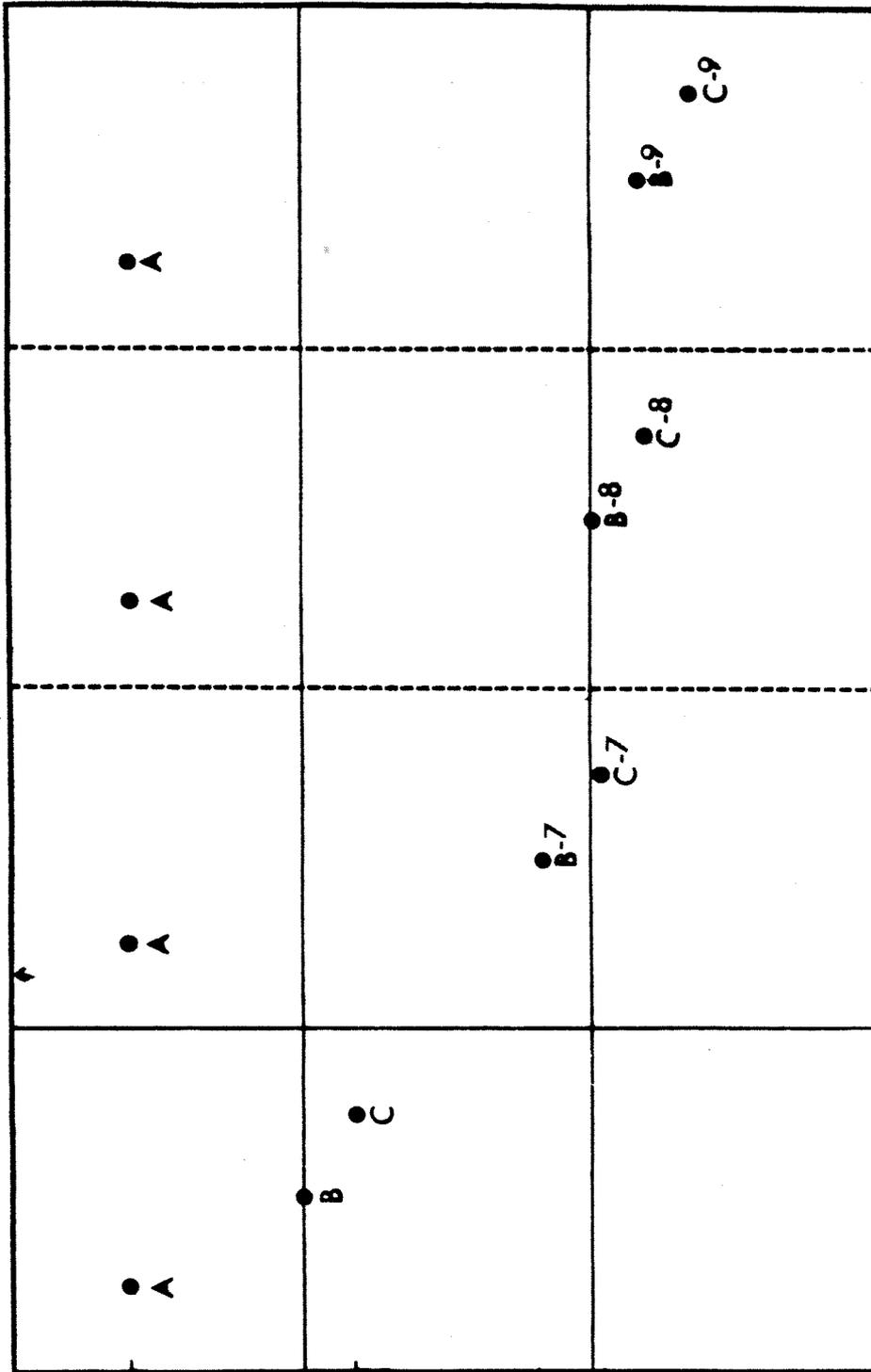
C : 428 Hz      C-7 : 240 Hz      C-8 : 214 Hz      C-9 : 192

L'intervalle BC est égal à un intervalle musical de un ton et demi. Les intervalles [B-7, C-7], [B-8, C-8] et [B-9, C-9]

Familiarisation

Test

Groupe 7      Groupe 8      Groupe 9



Temps

STIMULUS UTILISES POUR LES 4 SEQUENCES SONORES

Expérience 2

sont des transpositions de l'intervalle BC respectivement à la septième, à l'octave, et à la neuvième.

### 2.1.3. Intensité des sons

Comme pour l'expérience 1, nous avons été obligés d'égaliser les intensités des sons A, B et C d'une part, et des sons B-7, C-7, B-8, C-8, B-9 et C-9 d'autre part. L'intensité des sons A, B et C est fixée à 80 dB; celle des six autres sons à 85 dB.

Les mesures de distorsion acoustique ont été faites avec un analyseur de spectre (Nicolet, 444 A). On trouve pour tous les sons que les harmoniques sont à 40 dB en-dessous de la fondamentale, pour B-8 et C-8, les harmoniques sont à 49 dB en-dessous de la fondamentale.

### 2.1.4. Les séquences

Il y a quatre séquences différentes : une séquence de familiarisation composée à partir des trois sons A, B et C, et trois séquences-tests.

La séquence 7 est constituée à partir des trois sons [A, B-7, C-7], la séquence 8 à partir de [A, B-8, C-8], la séquence 9 à partir de [A, B-9, C-9]. Il y a trois groupes expérimentaux, qui entendent tous la même séquence de familiarisation au cours de la phase F.

Au cours de la phase-test, le groupe 7 entend la séquence [A, B-7, C-7], le groupe 8, la séquence [A, B-8, C-8], et le groupe 9, la séquence [A, B-9, C-9].

Le groupe 8 est le groupe Octave puisque l'intervalle [B-8, C-8] est la transposition à l'octave de l'intervalle B C de la séquence de familiarisation.

Par opposition, les groupes 7 et 9 sont appelés groupes Non-Octave.

## 2.2. Procédure

Dans ses grandes lignes, la procédure est identique à celle de l'expérience 1; notons cependant deux différences essentielles :  
1/ Le nombre d'essais dans la phase de familiarisation est égal à 12 pour la moitié des sujets, et 13 pour l'autre moitié.

Cette condition expérimentale permet d'éviter un quelconque biais dû au fait qu'on utilise un nombre constant de 12 essais au cours de la phase F. (Remarquons cependant que cette précaution n'est pas cruciale dans la mesure où la durée de chaque fixation dépend du bébé et ne lui est donc pas imposée par les conditions expérimentales).

- 2/ Contrairement à l'expérience 1, le son qui ne change pas d'une séquence à l'autre (ici A) est toujours présenté en première position (Voir CHANG et TREHUB, 1977).

Tous les essais (toutes les séquences entendues) commencent par le son A et se terminent par le son C (ou  $C^{-7}$ ,  $C^{-8}$ ,  $C^{-9}$  selon le cas), le son B (ou  $B^{-7}$ ,  $B^{-8}$ ,  $B^{-9}$ ) étant toujours placé entre les 2 sons précédemment cités. Chacun des trois groupes (7, 8 et 9) est divisé en deux sous-groupes selon le nombre d'essais de la phase F (12 ou 13).

Il y a donc six groupes expérimentaux.

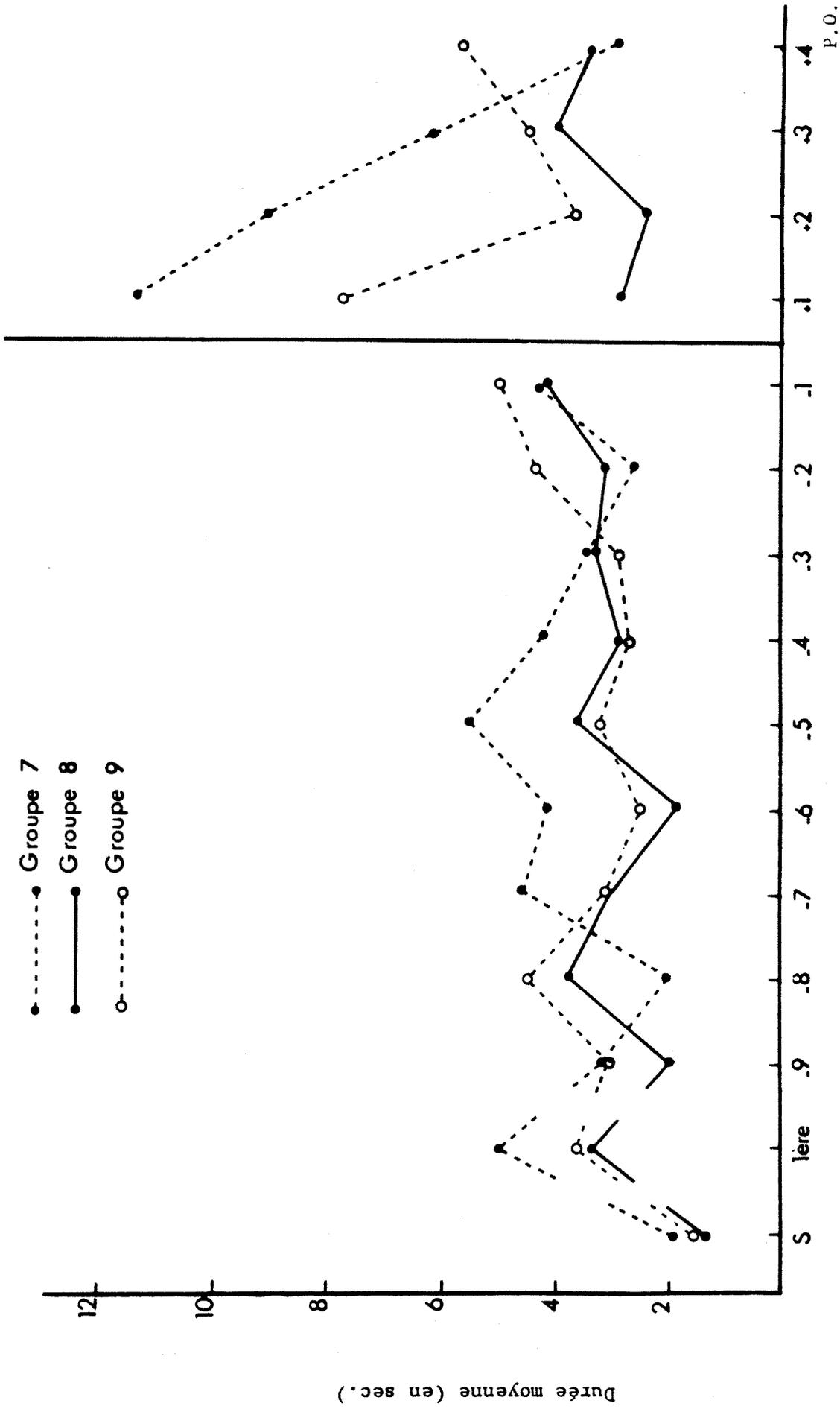
### 2.3. Sujets

Des nourrissons de 90 jours ( $\pm$  20 jours) ont été testés pour cette expérience 2. Ils étaient tous nés à terme et étaient en bonne santé au moment de l'expérience. Sur 57 sujets vus, 17 ont été éliminés pour les causes suivantes : 6 pour pleurs, 4 pour avoir fait un intervalle de silence supérieur à deux minutes, 2 pour difficultés d'observation, 5 pour des erreurs de procédure.

Les résultats de 40 sujets ont été gardés : 16 dans la condition "Octave" (groupe 8) - 24 dans la condition "Non-Octave" (groupes 7 et 9). Parmi ces 40 sujets il y a 18 filles et 22 garçons répartis de la façon suivante : 14 garçons et 10 filles dans les groupes 7 et 9, 9 garçons et 7 filles dans le groupe 8.

### 2.4. Résultats de l'expérience 2

Un rapide coup d'oeil sur les courbes (p. 72) montre un résultat très net : il n'y a pas de réaction à la nouveauté dans le groupe 8 (le groupe Octave) alors que les bébés ont une forte réaction dans les groupes 7 et 9. Le résultat du groupe 8 pourrait s'expliquer simplement par le fait que les bébés n'ont effectivement pas perçu de différence entre les deux séquences sonores entendues pendant les phases F et T. Ceci signifierait que l'intervalle B C et sa transposition



Durées moyennes des P.O.

EXPERIENCE 2

à l'octave inférieure sont perçus comme identiques par nos sujets.

L'analyse de variance montre que la valeur de l'indice de réaction à la nouveauté dans le groupe Non-Octave (groupes 7 et 9) est significativement plus grande que celle de l'indice du groupe 8 à  $P < .005$  ( $F(1-34) = 9.28$ ). (La différence entre les deux groupes 7 et 9 n'étant pas significative ( $F(1-20) = 1,79$ ,  $P > .10$ ), nous avons regroupé les deux résultats).

Outre ce résultat sur lequel nous reviendrons plus loin, plusieurs choses sont à noter.

1/ Pour les trois groupes expérimentaux (7, 8 et 9) on remarque qu'il n'y a pas d'habituation de la réponse du sujet, c'est-à-dire que la durée de la fixation du regard sur la cible ne diminue pas avec la répétition du stimulus.

Si on regarde uniquement les résultats du groupe 8, on pourrait penser que l'absence d'habituation est la cause principale de l'absence de réaction au moment du changement de stimulus. Cet argument ne tient pas, puisque dans les groupes 7 et 9, il y a une réaction à la nouveauté importante, alors qu'il n'y a pas plus d'habituation que dans le groupe 8.

Enfin, on aurait pu s'attendre à ce que la réaction à la nouveauté soit plus importante pour le groupe 9, en effet, l'intervalle de transposition entre [B C] et  $[B^{-9}, C^{-9}]$  est plus grand que dans le groupe 7 entre [B C] et  $[B^{-7}, C^{-7}]$ .

La différence entre les deux groupes n'étant pas significative, nous ne poursuivons pas plus loin cette discussion. Revenons-en plutôt au résultat principal de cette expérience 2.

- 2/ Le fait qu'il n'y ait aucune réaction au changement de stimulus dans le groupe Octave (groupe 8) peut signifier 2 choses :
- a. soit les bébés discriminent bien deux sons à l'octave, mais les perçoivent comme tellement similaires qu'ils les considèrent comme équivalents. On peut dire alors que les bébés de 3 mois sont extrêmement sensibles à la hauteur chromatique des sons.
  - b. soit les bébés ne discriminent pas du tout 2 sons séparés par une octave alors qu'ils différencient des sons séparés par des intervalles plus petits. Dans ce cas, les 2 séquences mélodiques sont perçues comme identiques par les sujets, ce qui pourrait expliquer le manque de réaction à la nouveauté : il faut donc

d'abord vérifier que des bébés de 3 mois sont capables de discriminer des sons à l'octave .

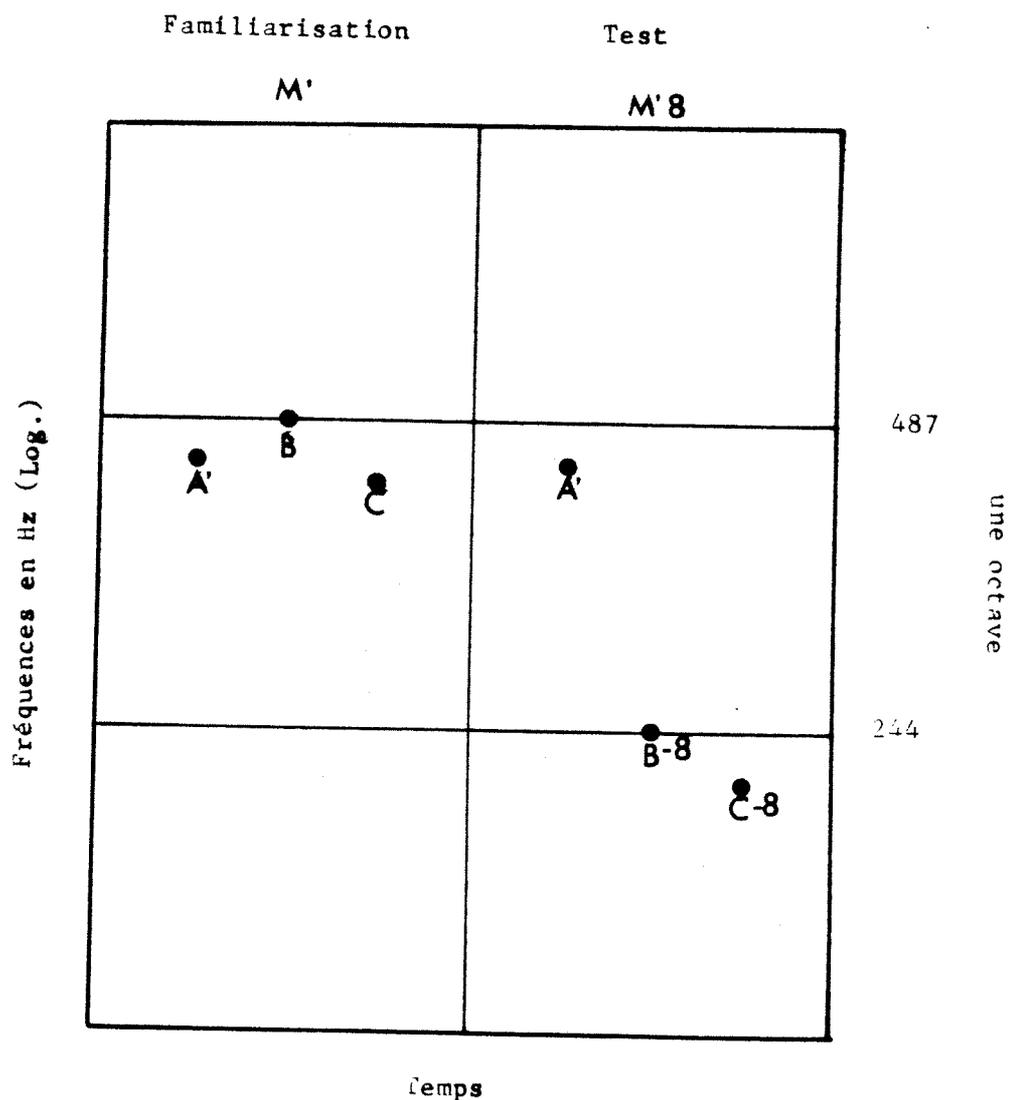
L'expérience suivante a pour objet cette vérification.

### 3 - EXPERIENCE 3

Pour cette expérience, nous avons conservé la même méthode, la même procédure et le même appareillage que dans l'expérience 2 : seuls les stimulus sont modifiés.

#### 3.1. Les stimulus

(Voir ci-dessous). Nous avons utilisé deux nouvelles séquences mélodiques sur la base de trois notes consécutives qui se répètent continûment.

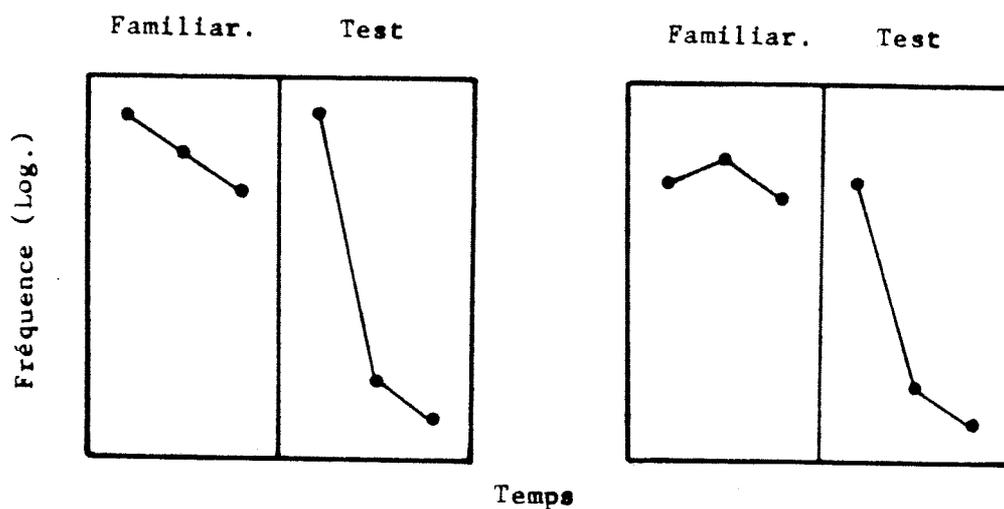


STIMULUS UTILISES POUR LES DEUX  
SEQUENCES SONORES PRESENTEES

Ces deux séquences (une pour la phase F, séquence M'; une pour la phase T, séquence M'8) ressemblent aux deux séquences du groupe Octave de l'expérience 2; on a simplement remplacé le son A (737 Hz) par le son A' (459,13 Hz). Ce son A' se trouve à équidistance de B et C sur l'échelle logarithmique des fréquences (3/8 de ton). Ce changement est important dans la mesure où le contour de la séquence mélodique n'est pas le même pour la phase F et pour la phase T; alors que dans l'expérience 2, le contour de toutes les séquences sonores est le même.

Même contour

Contours différents



Groupe 2

Groupe 3

Dans l'expérience 2, les 2 contours mélodiques sont les mêmes (2 intervalles descendants). Dans l'expérience 3, le stimulus de la phase F a un contour différent (un intervalle ascendant suivi d'un intervalle descendant) de celui du stimulus de la phase T (2 intervalles descendants). Dans les 2 cas, l'intervalle de transposition de BC est toujours égal à une octave.

Si les sons à l'octave ne sont pas discriminés, les 2 séquences mélodiques M' et M'8 seront perçues comme identiques, malgré la différence de contour entre les 2 séquences sonores.

Cette expérience 3 nécessite un seul groupe expérimental. Après la phase S (première fixation silencieuse), on présente la séquence M' pendant 12 essais pour la moitié du groupe, et 13 essais pour l'autre moitié. Ensuite, on présente la séquence M'8 pendant 6 essais au moins. Le déroulement de l'expérience 3 se fait de la même façon que dans l'expérience 2. La variable indépendante reste le temps de fixation du regard du bébé sur la cible.

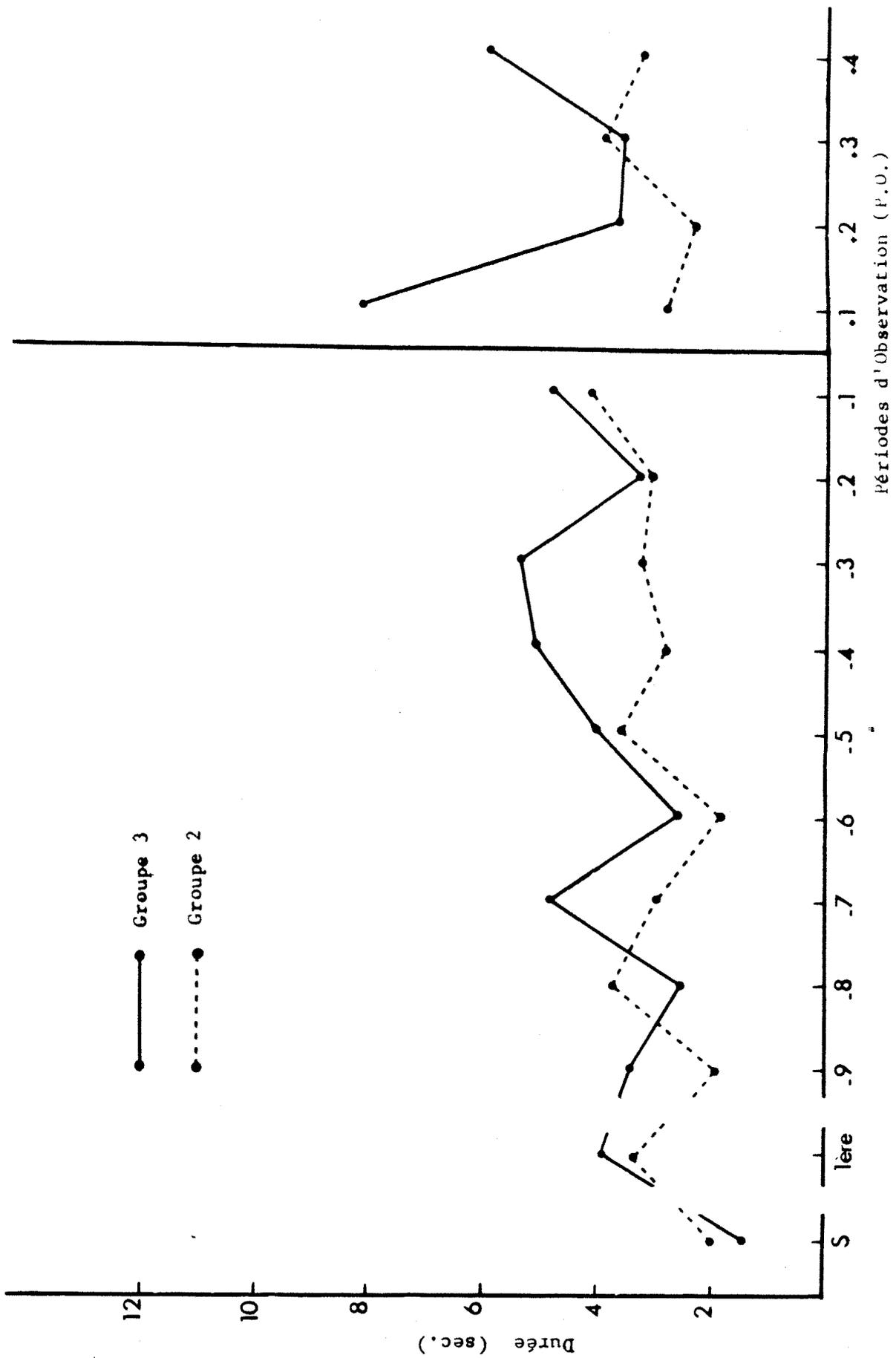
### 3.2. Les sujets

Pour faciliter la comparaison avec le groupe Octave de l'expérience 2, nous avons retenu le même nombre de sujets soit 16. Les bébés sont âgés de 91 jours en moyenne (70 à 110 jours). Sur 29 bébés testés, 13 ont été éliminés. Parmi les 16 retenus, on trouve 10 filles et 6 garçons.

### 3.3. Résultats

Nous appellerons groupe 2, le groupe Octave de l'expérience 2, et groupe 3, le groupe expérimental de l'expérience 3. Le schéma de la page 77 présente les durées moyennes de chaque P.O. pour les groupes 2 et 3. En comparant les 2 courbes, on voit immédiatement que les bébés ont une forte réaction à la nouveauté au changement de stimulus entre les 2 séquences M' et M'8 ( $t$  de student :  $t(15) = 2,61$ ;  $P < .01$  - test unilatéral). Ce résultat signifie que les sujets ont différencié les deux séquences; donc qu'ils sont bien capables de discriminer des sons séparés par un intervalle d'octave.

L'analyse de variance des groupes 2 et 3 montre que l'indice de réaction à la nouveauté est significativement plus élevé pour le groupe 3 que pour le groupe 2 ( $F(1,28) = 5.51$ ;  $P < .05$ ). Donc, si les bébés du groupe 2 n'ont pas de réaction au changement de stimulus, ça n'est pas parce que les intervalles [ B C ] et [ B<sup>-8</sup>, C<sup>-8</sup> ] sont perçus comme identiques puisque nous venons de voir qu'ils sont parfaitement discriminés. Il faut donc admettre que les deux



DUREE MOYENNE DE CHAQUE PERIODE D'OBSERVATION

Expérience 3



la condition e, mais au cours de la phase-test, on présente un autre stimulus (9 croix disposées en carré).

Les résultats montrent que les bébés de 2 et 4 mois répondent différemment à des éléments identiques (9 croix) disposées selon un carré ou selon une croix (condition c), mais répondent de la même façon à 9 éléments-croix et 9 éléments carrés, lorsque ceux-ci sont disposés de la même manière, selon une croix (condition e).

La perception du tout prime sur la perception des éléments.

Si on admet ce point de vue gestaltiste, l'expérience 3 ne suffit pas à prouver que les sons à l'octave sont discriminés. La différence entre les contours est telle que les deux stimulus ont pu être distingués même si certains éléments les composant ont été perçus comme très similaires. Nos données ne nous permettent pas de trancher sur ce point car l'analogie entre stimulus visuels et stimulus sonores n'est pas parfaite dans la mesure où les éléments sonores sont présentés les uns après les autres, alors que dans l'expérience de VURPILLOT et al., les éléments composant le stimulus visuel sont présentés en une seule fois. Le résultat de cette expérience est assez délicat à analyser.

A l'issue de l'expérience 3, nous avons été curieux de savoir si avec les stimulus que nous avons employés, il était possible de mettre en évidence le phénomène de l'équivalence chromatique chez des sujets adultes (musiciens et non-musiciens). La quatrième expérience de ce travail faite sur des adultes est un "lien" entre nos résultats sur les nourrissons et les expériences que nous avons décrites dans la première partie; elle sert en quelque sorte à tester sur des adultes les stimulus choisis pour les expériences 2 et 3.

#### 4 - EXPERIENCE 4

Pour cette expérience, nous allons tester deux groupes de sujets : des musiciens et des non-musiciens, faisant l'hypothèse que les sujets musiciens seront plus sensibles à la hauteur chromatique des sons que les non-musiciens.

#### 4.1. Les stimulus

Pour les expériences 2 et 3, nous avons composé six séquences mélodiques différentes qui vont être réutilisées ici. Les 4 séquences sonores de l'expérience 2 sont appelées M, M7, M8 et M9 (M correspond à la séquence entendue pendant la phase F; M7 correspond à la phase T du groupe 7; M8 à la phase T du groupe 8; M9 à la phase T du groupe 9) - Voir schéma p 81.

Aux deux séquences de l'expérience 2, que nous appelons à présent M' et M'8, il faut rajouter 2 nouvelles séquences M'7 et M'9 faites sur le modèle de M7 et M9, mais où l'on a remplacé le son A par le son A' comme pour M' et M'8 (voir schéma p 81). Il y a donc en tout 8 séquences sonores pour cette expérience 4.

#### 4.2. Méthode

Chaque sujet va entendre les deux séries de quatre séquences mélodiques M, M7, M8, M9 - M', M'7, M'8, M'9. Chaque mélodie est entendue séparément, la tâche du sujet est de classer les trois mélodies M7, M8 et M9 - (M'7, M'8 et M'9), selon leur degré de ressemblance avec la mélodie de référence M - (M').

#### 4.3. Appareillage et déroulement de l'expérience

Dans un premier temps, les sons des séquences mélodiques sont synthétisés par le micro-ordinateur TRS 80 (comme dans les expériences précédentes). Ensuite, chaque séquence mélodique est enregistrée séparément sur une piste d'un magnétophone à cassettes quatre pistes, pendant une durée d'environ 1/4 d'heure. Sur chaque piste on a une séquence ininterrompue. L'intensité subjective des sons a été abaissée, étant donné que l'écoute se fait cette fois au casque, et que cette écoute doit rester confortable pour le sujet.

Pour cette expérience, on dispose donc d'un magnétophone quatre pistes, à cassettes (TEAC 144 - vitesse de déroulement : 9,5 cm/sec.). Ce magnétophone permet d'avoir les quatre séquences mélodiques d'une même série enregistrées en parallèle. Le sujet peut ainsi sélectionner l'une ou l'autre de ces quatre mélodies grâce à quatre interrupteurs placés sur un boîtier. Les quatre séquences

Mélodie de  
référence

FREQUENCES

	M	M7	M8	M9
	a	a	a	a
	b			
	c			
		b		
		c		
			b	
			c	
				b
				c

mélodie de  
référence

	M'	M'7	M'8	M'9
	a'	a'	a'	a'
	b			
	c			
		b <sup>-7</sup>		
		c <sup>-7</sup>		
			b <sup>-8</sup>	
			c <sup>-8</sup>	
				b <sup>-9</sup>
				c <sup>-9</sup>

TEMPS

Triolets utilisés pour les 6 séquences sonores

EXPERIENCE 4

se déroulent simultanément et à la même vitesse durant toute l'expérience.

Le boîtier qui permet de sélectionner une des quatre mélodies enregistrées fonctionne de la façon suivante : au départ tous les interrupteurs sont en position OFF, le sujet n'entend rien. S'il veut écouter une mélodie, il met un interrupteur en position ON, s'il veut l'arrêter, il faut repasser en position OFF. Ce dispositif permet de ne jamais arrêter le déroulement de la bande magnétique tout au long de l'expérience. Les quatre interrupteurs ON/OFF sont repérés par quatre couleurs différentes : la mélodie de référence M dans un cas, M' dans l'autre est toujours repérée par la couleur grise les trois mélodies-test M7, M8, M9 ou M'7, M'8, M'9 correspondent aux trois couleurs : rose-vert-blanc. La correspondance entre les mélodies-test et les couleurs peut être permutée : il y a six permutations possibles (la tâche du sujet étant de classer les trois mélodies-test, nous avons évité de les repérer par des chiffres ou des lettres, afin de n'induire aucun classement). Le sujet porte un casque (Sennheiser HD430) sur lequel sont envoyées les séquences mélodiques via un amplificateur (Pioneer SA 6003). (Schéma p. 83 ).

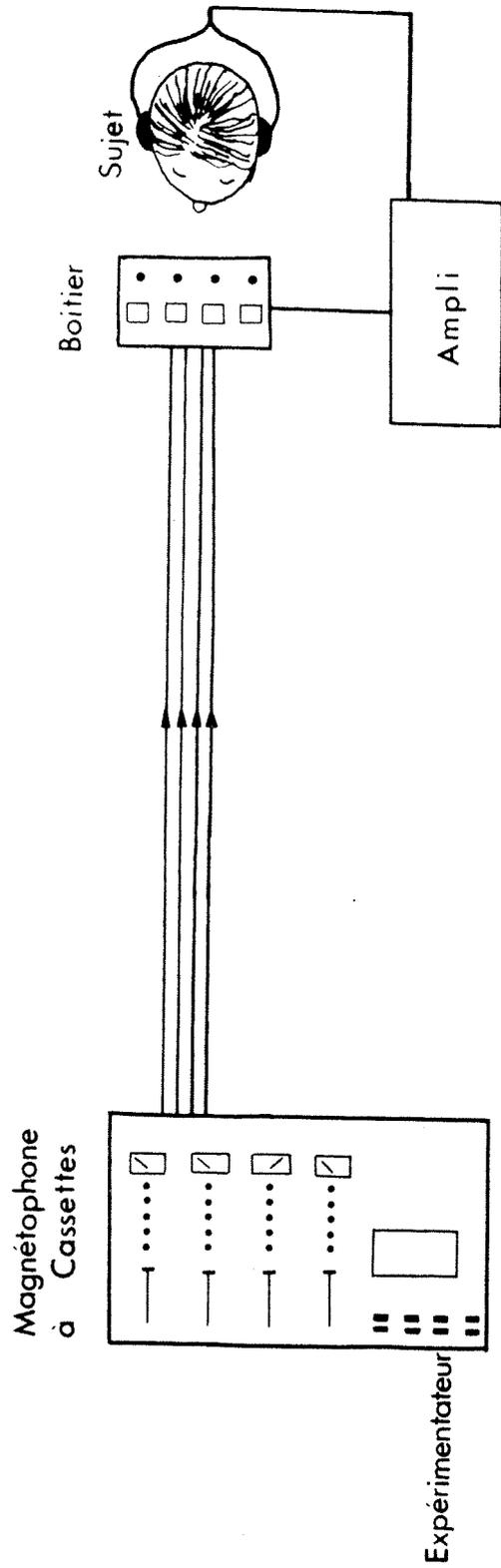
#### 4.4. Procédure

L'expérience se déroule dans une pièce calme où seuls sont présents le sujet et l'expérimentateur. Avant que l'expérience ne débute, on demande au sujet : son âge, son niveau d'études et sa profession. Enfin, on note éventuellement les problèmes d'audition du sujet. Pour le groupe des sujets musiciens, on précise le nombre d'années de pratique musicale (voix, solfège, instruments). L'expérience se décompose en 4 parties :

##### Première partie :

Classement en degré de similitude avec la séquence M, des séquences M<sup>7</sup>, M<sup>8</sup>, M<sup>9</sup>.

Le sujet, assis face à l'expérimentateur, a devant lui le boîtier avec les 4 interrupteurs ON/OFF. On lui donne la consigne suivante : "quatre mélodies, sous la forme de trois sons qui se répètent continûment, sont enregistrées sur cette cassette. Vous allez pouvoir écouter chacune de ces mélodies séparément en actionnant l'un ou l'autre



DISPOSITIF

EXPERIMENTAL

EXPERIENCE

4

de ces quatre boutons. Parmi ces quatre mélodies, la mélodie M (repérée par la couleur grise) va jouer le rôle de mélodie de référence, une fois que vous l'aurez écoutée, vous écouterez successivement les trois autres, dans l'ordre qui vous convient. Votre tâche consiste à classer ces trois mélodies (verte-blanche-rose) selon le degré de ressemblance que vous pourrez lui trouver avec la mélodie de référence M".

On précise encore deux points importants :

1/ le sujet peut écouter chacune des quatre séquences autant de fois qu'il le désire, avant de donner un classement définitif. Il a pour consigne de n'écouter qu'une seule mélodie à la fois, c'est-à-dire qu'il doit obligatoirement passer par la position OFF de l'interrupteur correspondant à la séquence qu'il entend avant de mettre un autre interrupteur sur ON pour écouter une nouvelle séquence.

2/ Enfin, on demande au sujet d'être attentif et de ne pas donner une réponse au hasard. S'il ne sait pas répondre, il doit le dire. Pour l'encourager dans cette voie, on lui dit qu'il n'y a pas "une bonne réponse" à trouver; ce qui importe c'est sa réponse.

Lorsque la consigne est donnée, et que le sujet a bien compris ce qu'il doit faire, il met le casque. L'expérimentateur vérifie que les quatre boutons du boîtier sont en position OFF, et met le magnétophone en marche. Le sujet dispose alors de tout le temps qu'il désire pour écouter toutes les séquences. Lorsqu'il a terminé, il donne le classement des trois mélodies-test : de la mélodie la plus ressemblante à la moins ressemblante. Le classement est donné sous la forme de trois couleurs, par exemple : rose-vert-blanc; pour le sujet, la mélodie "rose" est celle qui ressemble le plus à la mélodie M. Cette première partie de l'expérience s'achève lorsque le sujet a donné son classement.

#### Deuxième partie :

Le dispositif expérimental reste le même, mais cette fois, c'est l'expérimentateur qui prend le boîtier devant lui, de façon à ce que le sujet ne le voit plus. Dans cette partie, nous n'allons utiliser que les mélodies-test M7, M8 et M9. Voici la nouvelle consigne donnée au sujet : "Vous allez entendre successivement deux séquences mélodiques; votre tâche est de dire si elles sont identiques ou différentes. Après

chaque paire de mélodies (il y en a six en tout) vous direz simplement "pareil" si les mélodies vous semblent identiques ou "différent" si les mélodies ne sont pas les mêmes".

On a donc deux types de paires de comparaison : les paires identiques (comme M7 et M7) et les paires différentes (comme M7 et M8). Il y a trois paires identiques (M7 et M7, M8 et M8, M9 et M9) et six paires différentes (M7 et M8, M8 et M7, M8 et M9, M9 et M8, M7, et M9, M9 et M7). Normalement, nous aurions donc dû présenter neuf paires de mélodies à chaque sujet -

Cette procédure étant trop longue, nous avons décidé de ne présenter que 3 paires différentes : M7 et M8, M8 et M9, M7 et M9.

Pour ce choix, nous nous sommes fixés deux contraintes : M7 est toujours présentée avant M8 et M9, et M8 est toujours présentée avant M9. L'ordre dans lequel les paires sont présentées au sujet est tiré au hasard.

Les couples de mélodies à comparer sont présentés de la façon suivante : la première séquence est entendue pendant 7 secondes, puis après un silence de 3 secondes, on fait entendre la deuxième séquence encore pendant 7 secondes. Le sujet doit alors dire "pareil" ou "différent". L'expérimentateur note cette réponse en s'employant à faire le même geste à chaque fois, de façon à laisser à tous les sujets le même laps de temps entre deux comparaisons. Ce temps doit permettre aussi au sujet d'oublier les séquences qu'il vient d'entendre, afin qu'il n'y ait pas de confusions avec les deux séquences mélodiques suivantes. Dès que la réponse du sujet a été notée, l'expérimentateur annonce la prochaine comparaison, et fait entendre deux autres séquences au sujet : 7 secondes pour la première, 3 secondes de silence, 7 secondes pour la deuxième. Enfin, le sujet donne sa réponse.

On procède de la même façon pour les six comparaisons à faire : trois comparaisons "pareil", et trois comparaisons "différent". Cette partie de l'expérience est en fait un "test de discrimination" qui permet de vérifier que les sujets ont bien discriminé les trois séquences mélodiques qu'ils avaient à classer dans la première partie de l'expérience. Si le sujet ne fait aucune erreur, il poursuit l'expérience avec la troisième partie, mais s'il fait au moins une erreur,

il "saute" cette troisième partie, et passe directement à la quatrième partie que nous verrons plus loin.

Troisième partie :

(Réservée aux sujets n'ayant fait aucune erreur au test de discrimination)

Tout se déroule comme dans la première partie, mais avec la mélodie M' et ses variantes. La mélodie de référence M' est toujours repérée par la couleur grise. Chaque mélodie-test (M'7, M'8, M'9) est successivement repérée par l'une des trois couleurs vert-rose-blanc.

La tâche du sujet est la même, on lui donne donc la même consigne que dans la partie 1, en lui précisant toutefois que l'expérience est un peu plus difficile que la précédente. Ceci lui permettant de ne pas être trop surpris par le changement de contour entre la mélodie M' et les 3 autres. On lui rappelle qu'il ne doit pas donner de réponse au hasard; s'il n'arrive pas à classer les trois mélodies-test, c'est-à-dire s'il ne trouve pas que l'une ressemble plus que l'autre à la mélodie M', il doit le dire. Il ne donne donc aucun classement.

Ces précautions étant prises, le sujet reprend le boîtier avec les quatre interrupteurs devant lui, le casque sur ses oreilles. On vérifie que les interrupteurs sont en position OFF, puis on met le magnétophone en marche.

Comme dans la partie 1, le sujet n'est pas limité dans le temps, il peut écouter toutes les mélodies autant de fois qu'il le désire. L'expérience s'arrête lorsque le classement définitif est donné.

Quatrième partie :

Tous les sujets doivent subir cette dernière épreuve, même ceux qui n'ont pas réussi le test de discrimination (deuxième partie de l'expérience).

Cette dernière épreuve consiste simplement pour le sujet à répondre aux deux questions suivantes :

1. Citez 3 intervalles musicaux consonnants.
2. Donnez un exemple d'accord parfait.

Ces deux questions ne sont là que pour tester les connaissances théoriques des sujets. Ceci permet de vérifier d'une part que les sujets naïfs le sont bien, et d'autre part, que les sujets musiciens ont un minimum de culture générale en matière de théorie de la musique.

L'expérience se termine par ces deux questions.

#### 4.5. Sujets

54 sujets ont été testés (36 femmes et 18 hommes). Ils sont âgés de 19 à 35 ans. Ils se divisent en deux grands groupes, les musiciens et les non-musiciens, selon les critères suivants : les musiciens sont ceux qui ont fait au moins cinq ans d'études musicales. Ces sujets sont tous des élèves du conservatoire de Pantin. Les sujets non-musiciens, ou naïfs, ne devaient pas avoir fait plus de trois ans d'études musicales. Tous les sujets naïfs sont des étudiants en Psychologie.

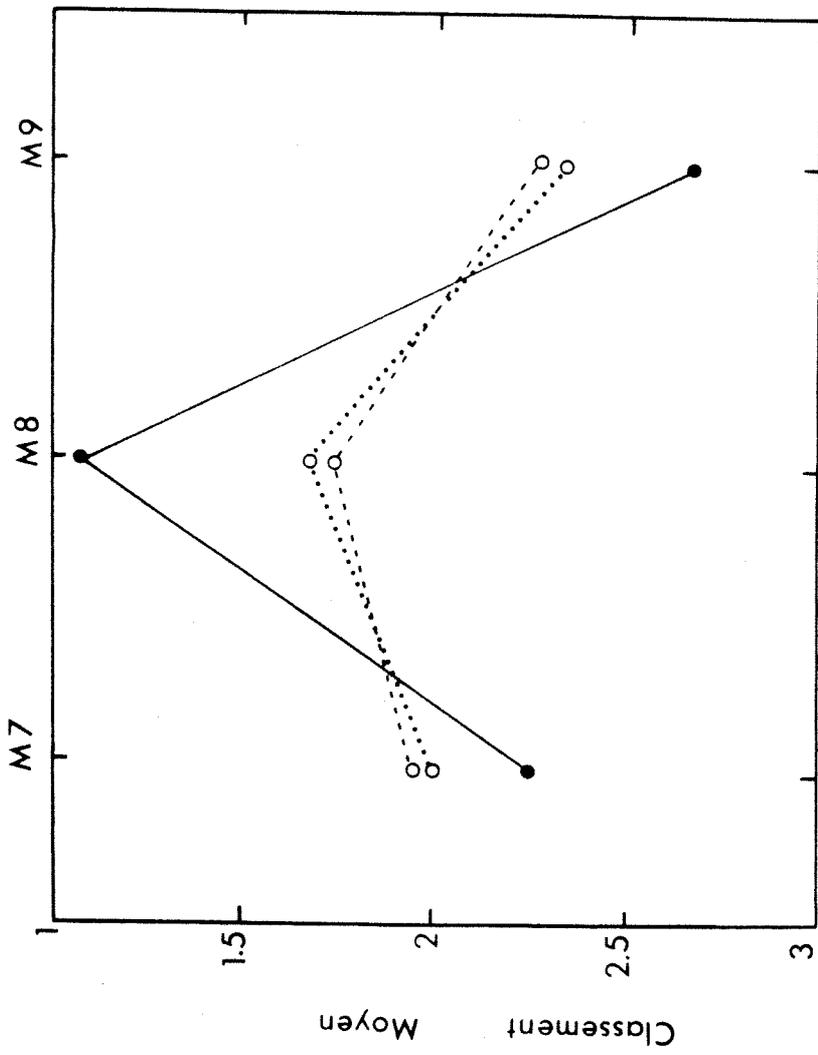
Parmi les 54 sujets testés : on trouve 42 sujets naïfs, 12 sujets musiciens. Parmi les sujets naïfs, 3 n'ayant pu donner de classement, ont été éliminés. Les résultats concernent donc 51 sujets (39 naïfs et 12 musiciens).

Le groupe des sujets naïfs a été divisé en deux : d'un côté les sujets "naïfs sélectionnés", c'est-à-dire ceux qui n'ont pas fait d'erreur au test de discrimination, de l'autre, les sujets "naïfs non-sélectionnés", ceux qui ont passé la première partie de l'expérience, mais qui ont fait au moins une erreur dans le test de discrimination. Parmi les 39 sujets naïfs, il y a 30 sujets "sélectionnés" et 9 "non-sélectionnés".

#### 4.6. Résultats

##### 4.6.1. Résultats de la première partie

Pour chaque groupe de sujets nous avons déterminé le rang moyen obtenu par chacune des trois mélodies-test M7, M8, M9. Le rang 1 correspond à un degré de ressemblance maximum avec la mélodie de référence M. Sur les courbes (p. 88) on voit tout de suite qu'il y a peu de différences entre les deux sous-groupes : "naïfs sélectionnés" et "naïfs non-



●——● Musiciens  
 ○- - -○ Sujets Naïfs  
 ○.....○ Sujets Naïfs Sélectionnés

RESULTATS

EXPERIENCE 4

sélectionnés"; nous les regrouperons sous le nom "sujets naïfs".

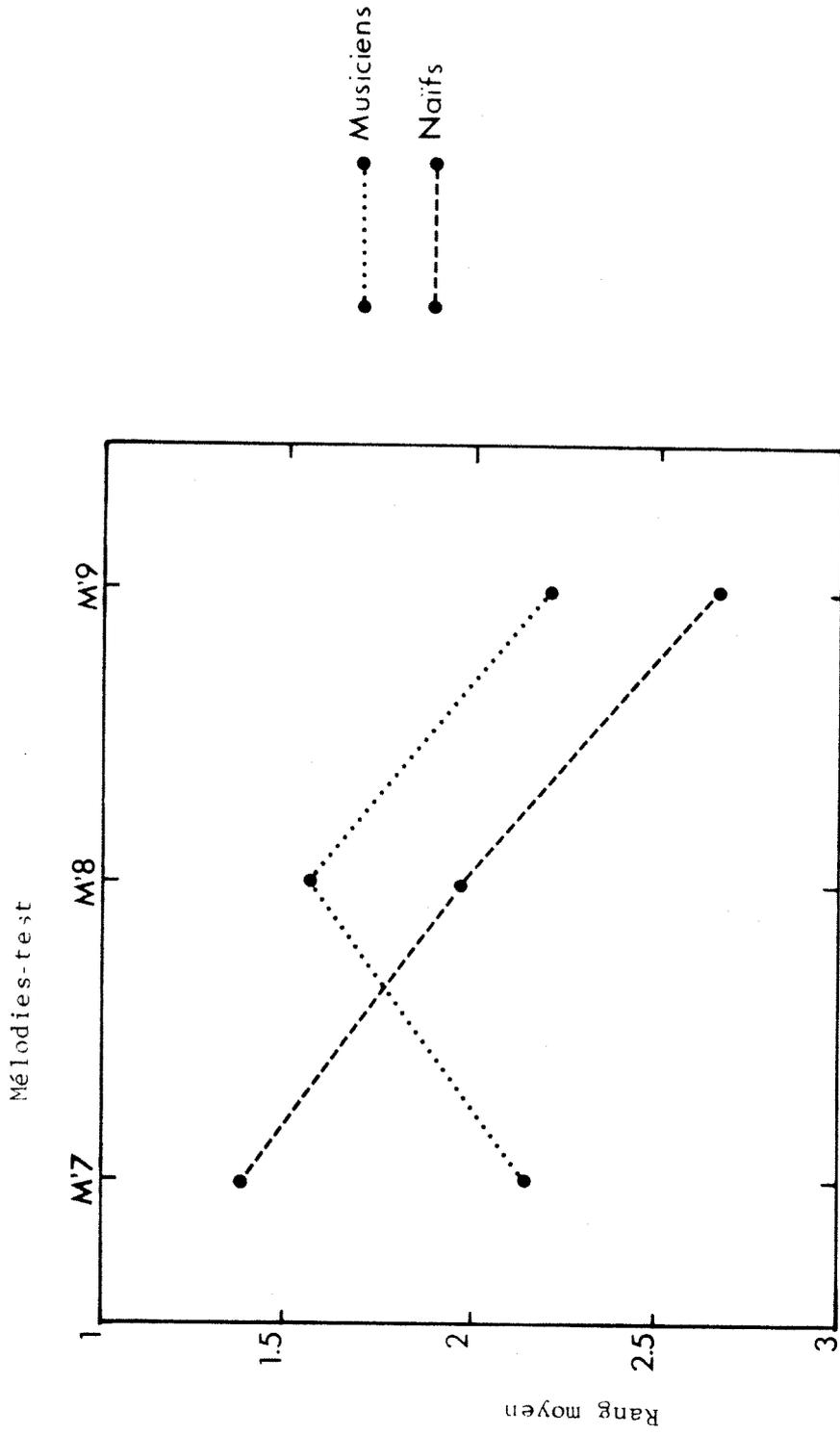
Pour cette première partie de l'expérience, une nette opposition apparaît entre les deux groupes : pour les sujets naïfs, aucune des trois mélodies M7, M8, M9 ne ressemble particulièrement à la mélodie de référence M. Seulement la moitié des sujets sélectionnés (15/30) ont classé la mélodie M8 comme étant la plus ressemblante à la mélodie de référence. (Voir tableau p 91 ).

Les trois mélodies sont classées à un rang moyen (autour de 2). Par contre, les sujets musiciens ont presque tous (11/12) classé la mélodie M8 comme étant la plus ressemblante à la mélodie M.

Si on compare les deux proportions 15/30 et 11/12, on trouve que la seconde est significativement plus importante que la première ( $p < .05$ , test de chi 2).

On se rappelle que dans la mélodie M8, l'intervalle  $B^{-8} C^{-8}$  est la transposition, une octave en-dessous, de l'intervalle B C (de la mélodie M). Cela confirme bien les résultats déjà connus : les musiciens sont plus sensibles à la hauteur chromatique que des sujets non-musiciens, ce qui met encore en lumière le rôle de l'éducation musicale dans ce domaine.

Mais il faut noter que la proportion 15/30 est supérieure au hasard, puisqu'à priori les trois mélodies M7, M8, et M9 avaient la même chance d'être classées au rang 1; la proportion pour que M8 soit classée en 1 est donc de 10/30. On peut expliquer ce résultat par le fait que certains sujets non-musiciens sont capables de percevoir la hauteur chromatique. (Voir page 91).



Classement moyen de chacune des trois mélodies-test, selon leur degré de similarité avec la mélodie M'.

EXPERIENCE 4

<u>MUSICIENS</u>			
Mélodies Classement	M7	M8	M9
1	1	11	0
2	6	1	5
3	5	0	7

N = 12

<u>NON MUSICIENS</u>			
Mélodies Classement	M7	M8	M9
1	9	15	6
2	12	10	8
3	9	5	16

N = 30

Classement de chacune des 3 mélodies  
M7, M8 et M9 pour les 2 groupes :  
musiciens et non-musiciens.

#### 4.6.2. Résultats de la deuxième partie : test de discrimination.

Pour ce test de discrimination, entre les trois mélodies-test M7, M8, M9, aucun musicien n'a commis d'erreur, par contre, sur 39 sujets naïfs, 9 ont commis au moins une erreur, c'est-à-dire presque un quart de la population testée. Or, on sait que des sujets adultes perçoivent sans difficulté une différence d'un ton et demi entre deux sons. Les erreurs commises à ce test de discrimination peuvent être attribuées à la difficulté de la tâche, mettant en jeu la capacité de mémoriser les hauteurs relatives des sons présentés.

#### 4.6.3. Résultats de la troisième partie

Rappelons que dans cette partie, le contour de la mélodie de référence est différent de celui des trois mélodies-test (revoir p.81 ). En ce qui concerne les sujets musiciens, l'allure générale de la courbe (p. 90) donnant le classement moyen obtenu par chacune des trois mélodies M'7, M'8 et M'9, reste la même que dans la partie 1; ce qui veut dire que la

M'8 a été perçue comme étant la plus ressemblante à la mélodie de référence M', mais de façon moins nette que dans la première partie. Le changement de contour a été un handicap pour les sujets musiciens : 5 sujets n'ont pu donner de classement. La similarité chromatique a donc été beaucoup moins évidente pour eux; seulement 4 sujets sur 7 ont classé la mélodie M'8 en premier. Pour les sujets naïfs, il y a un renversement de situation :

la mélodie M'7 est classée très nettement devant M'8 et M'9; c'est donc elle qui pour eux ressemble le plus à la mélodie M'.

Cette partie 3 de l'expérience nous apporte une précision puisqu'on voit que certains sujets musiciens continuent à classer la mélodie M'8 en premier (4/7) alors que dans les classements des sujets naïfs, c'est la mélodie M'7 qui est classée nettement en tête.

Le degré de similarité décroît avec l'intervalle de transposition, mais à l'inverse de KALLMAN, on ne note aucune particularité pour la mélodie M'8, c'est-à-dire, la mélodie à l'octave. Les sujets naïfs se sont basés uniquement sur la taille de l'intervalle de transposition. ATTNEAVE et OLSON (1971) avaient noté à ce propos, que les sujets non-musiciens se représentaient les intervalles comme des quantités, des grandeurs ou des distances, alors que les musiciens se les représentent en termes quantitatifs ou catégoriels.

La modification du contour a joué un rôle perturbant à la fois pour les sujets naïfs et pour les sujets musiciens.

Cette expérience qui se rapproche par le paradigme utilisé (jugement de similarité) de celles de ALLEN (1967) THURLOW et ERCHUL (1977) et KALLMAN (1982) semble montrer que les sujets naïfs sont plus sensibles à la hauteur chromatique que ne le laissaient penser ces trois expériences.

Notre principale conclusion porte sur la différence importante qui existe entre les sujets musiciens, très sensibles à la hauteur chromatique, et les sujets sans expérience musicale.

#### 4.7. Discussion

La plupart des expériences faites sur la perception de la hauteur chromatique chez les adultes concluent que la similarité chromatique est mieux perçue par les sujets ayant une expérience musicale. Ce que les expériences ne précisent pas, c'est la manière dont l'éducation musicale exerce une influence. Est-ce qu'elle permet d'apprendre ce qu'est la hauteur chromatique, ou bien de maintenir une capacité perceptive qui serait fonctionnelle chez le bébé dès l'âge de 3 mois?

En effet, si on regarde les résultats des expériences 2 et 3, et ceux de l'expérience 4, on peut admettre que de très jeunes enfants perçoivent la hauteur chromatique d'un son et que cette capacité perceptive tend à disparaître si elle n'est pas "renforcée" par l'éducation musicale. Ceci expliquerait la différence que l'on trouve toujours entre les sujets ayant une expérience musicale, et ceux qui n'en ont pas.

Reste à savoir si l'éducation musicale joue un rôle dans l'acquisition de cette capacité perceptive qui permet de percevoir la hauteur chromatique des sons ou si ce rôle se limite à maintenir fonctionnelle une capacité innée ou acquise très tôt au cours du **développement**. L'ensemble de nos expériences sur les bébés ont apporté un début de réponse.

## DISCUSSION - CONCLUSION

Plusieurs interprétations ont été proposées pour expliquer le phénomène de l'équivalence chromatique, chacune d'entre elles amène une hypothèse différente sur le développement de la perception du chroma.

Selon le phénomène de l'acculturation, la perception de la hauteur chromatique des sons serait acquise à partir des caractéristiques de la musique environnante que nous entendons habituellement en Occident. Ces caractéristiques, en interaction avec le système auditif, créent des habitudes perceptives dont par exemple le sentiment de tonalité qui semble être acquis vers 6-8 ans par l'écoute et l'étude de notre musique tonale. Des arguments en faveur de cette théorie ont été apportés par FRANCES, IMBERTY, BURNS et WARD.

Pour TERHARDT, la hauteur chromatique est perçue par l'écoute répétée des sons naturels complexes et plus spécialement des sons de la voix humaine. Cet apprentissage que nous avons appelé associatif "habitude" le sujet entendant à assimiler tous les sons ayant la même hauteur chromatique.

Enfin, OHGUSHI suggère que la perception de la hauteur chromatique est innée. Le fonctionnement du système auditif expliquerait à lui seul cette capacité perceptive.

A partir de ces trois théories, on peut faire diverses prévisions sur l'âge auquel apparaîtra l'équivalence chromatique. Pour la théorie de l'acculturation, ce phénomène ne peut être mis en évidence avant 6-7 ans. Dans ce cas, le développement de ce phénomène dépend du niveau de culture musicale du sujet. TERHARDT semble penser, sans que cela soit clairement dit, que la hauteur chromatique serait acquise beaucoup plus tôt. Pour OHGUSHI enfin, cette capacité auditive serait innée.

Dans cette dernière hypothèse, il faut préciser que la manifestation du phénomène est liée au développement du système auditif du bébé. Ce qui veut dire que la capacité à percevoir la hauteur tonale peut être

latente chez le bébé et qu'elle n'apparaîtra réellement que lorsque la maturation du système nerveux le permettra.

En prenant comme sujets des bébés de quelques mois, nous avons mis à l'épreuve les différentes prévisions de ces auteurs quant à l'âge auquel peut apparaître l'équivalence chromatique. La manifestation de celle-ci chez des enfants aussi jeunes apporterait des arguments en faveur des théories de TERHARDT et d'OHGUSHI.

#### Bilan expérimental

L'expérience 2 nous apporte un résultat assez surprenant; en effet, elle montre que les bébés assimilent parfaitement des sons séparés par une octave.

La réaction à la nouveauté dans le groupe 8 (transposition à l'octave) est extrêmement faible, alors que pour les deux groupes où la transposition se fait à un autre intervalle que l'octave, la réaction au changement de stimulus est importante. Cependant, le phénomène n'a pu être mis en évidence chez tous les sujets.

Au vu de ce seul résultat, on pourrait imaginer que deux sons à l'octave sont parfaitement semblables pour nos sujets.

A quoi attribuer ce résultat ? En réalité, les auteurs qui se sont intéressés aux problèmes de l'équivalence chromatique, ont fait peu d'hypothèses sur l'explication du mécanisme en lui-même. Nos expériences ne nous permettent pas plus de faire une hypothèse à ce sujet, mais les résultats de l'expérience 2 montrent que les bébés perçoivent plus qu'une simple similarité entre des sons à l'octave.

Ce résultat nous conduit à envisager une quatrième hypothèse qu'on pourrait appeler hypothèse des harmoniques aurales.

On sait que, pour un certain niveau d'intensité, un son pur produit dans l'oreille une série de sons dont les fréquences sont des multiples de la fréquence du son présenté (VIEMEISTER 1972 - cité par MOORE 1982). Ainsi, si l'on entend deux sons purs à l'octave et qu'on perçoive la série d'harmoniques aurales de ces deux sons, toutes les harmoniques aurales du son le plus aigu sont présentes dans la série harmonique du son à la sub-octave. On comprend alors que deux sons à l'octave paraissent semblables.

Cette hypothèse permettrait de comprendre le résultat obtenu chez nos sujets, mais rien ne nous autorise à dire que des bébés de 3 mois perçoivent des harmoniques aurales. Il faut donc vérifier ce point.

Le résultat de l'expérience 2 peut aussi s'expliquer tout simplement par le fait que les bébés de 3 mois ne discriminent pas des sons séparés par une octave (tout en discriminant des sons séparés par des intervalles plus petits).

Dans ce cas, la mélodie dans laquelle on a transposé l'intervalle BC à l'octave serait perçue en tout point indentique à la mélodie originale non transposée.

Du point de vue de la discrimination des sons à l'octave, l'expérience 3 n'est pas décisive, car dans le groupe 3 on modifie à la fois les éléments et la configuration générale (le contour) de la séquence mélodique.

VURPILLOT et al. (1977) ont montré que des bébés de 2/4 mois ne répondent pas à un changement de la forme des éléments constitutifs d'une configuration simple alors qu'ils discriminent d'une part la même modification d'un de ces éléments isolés, d'autre part un changement de forme de la configuration. Il semblerait que les propriétés structurelles d'un ensemble d'éléments discrets, visuels ou sonores, soient perçus très précocément et déterminent en priorité sa réponse. En ce cas, pour le groupe 3 de l'expérience 3, on peut conclure que 2 sons à l'octave sont traités comme équivalents par les bébés puisqu'on a introduit une modification de la configuration sonore présentée (i.e. changement de contour entre la mélodie de familiarisation et la mélodie-test).

L'hypothèse, bien que peu probable, de la non discrimination des sons à l'octave, par de tout-jeunes enfants reste une explication possible de nos résultats.

Par contre, les données obtenues dans l'expérience 3 montrent que l'hypothèse des harmoniques aurales ne tient plus pour expliquer les résultats de l'expérience 2.

Si les bébés étaient réellement capables de percevoir une série d'harmoniques aurales à partir d'un son pur, on n'obtiendrait pas une forte réaction à la nouveauté dans le groupe 3 (expérience 3) car,

bien que le contour soit modifié entre la mélodie de familiarisation et la mélodie-test, la transposition de l'intervalle BC se fait toujours à l'octave, donc la mélodie-test devrait sembler similaire à la mélodie de familiarisation de la même façon dans les deux groupes (2 et 3) de l'expérience 3.

Puisque la réaction des bébés est différente dans les deux cas, l'hypothèse des harmoniques aurales semble très improbable.

A l'issue de ce bilan expérimental, il nous est impossible de ne retenir qu'une seule des trois explications de l'équivalence chromatique que nous avons vues. Nos résultats permettent simplement de dire que ce phénomène perceptif peut être mis en évidence, dans certaines conditions, chez des sujets âgés de 3 mois : cela tendrait à accréditer l'idée que l'équivalence chromatique n'est pas due à un apprentissage culturel. Le phénomène d'acculturation décrit par IMBERTY et FRANCES d'une part, et plus récemment par BURNS et WARD (1982) d'autre part, ne peut expliquer nos données. Par contre, les hypothèses de TERHARDT et d'OHGUSHI sont corroborées par notre résultat principal dans ce travail : soit la perception de la hauteur chromatique est innée, soit elle se développe avec les premiers stimulus sonores (sons complexes de la voix) perçus par le tout petit bébé.

Si TERHARDT a raison alors on peut dire que l'apprentissage associatif sur lequel il fonde sa théorie est achevé chez certains sujets dès le troisième mois post-natal.

On peut alors se demander pourquoi un apprentissage aussi rapide à partir de stimulus naturels, se maintiendrait chez si peu de sujets adultes, (les musiciens et quelques sujets non musiciens).

En effet, les résultats montrent aussi que cette capacité n'existe, à l'âge adulte, pratiquement que chez les musiciens; ce qui voudrait dire qu'il y a une régression de cette capacité auditive. Non entretenue par une pratique musicale régulière, la perception de la hauteur chromatique des sons se dégraderait au cours du développement d'un sujet.

On retrouve le même type de remarque, mais cette fois à propos d'une autre capacité auditive : l'oreille absolue.

Pour certains auteurs (ABRAHAM 1901, WATT 1917 - cités par WARD 1963), l'oreille absolue, c'est-à-dire la capacité de reconnaître une note sans référence extérieure serait innée et universelle, mais serait inhibée au cours du développement.

La théorie du "développement par perte" défendue par MEHLER (1974) et bien connue des biologistes et des spécialistes du développement, peut expliquer ce type de phénomène.

On peut alors penser que certains facteurs tels que la perception du contour mélodique d'une suite de sons (notes, syllabes, ...) se développent et peuvent inhiber en cela d'autres capacités perceptives moins utiles telles que la perception de la hauteur chromatique des sons ou l'oreille absolue.

(Cette hypothèse nous étant suggérée à la fois par les résultats de l'expérience 3 et ceux de l'expérience 4 : dans les deux cas, en effet, on voit qu'une modification du contour de la mélodie perturbe la perception de la hauteur chromatique).

Certaines potentialités du système auditif, mises en évidence très tôt chez le jeune enfant, évoluent dans le cadre de l'interaction avec l'environnement; dans le cas d'un environnement musical riche, elles peuvent se maintenir jusqu'à l'âge adulte, ou régresser dans un autre environnement. C'est dans cet ensemble de capacités perceptives que semble se placer la perception de la hauteur chromatique des sons.

BIBLIOGRAPHIE

---

ALLEN D.

Octave discriminability of musical and non-musical subjects,  
1967, Psychonomic Sci., 7 (12).

ATTNEAVE F. et OLSON R.

Pitch as a medium : a new approach to psychophysical scaling, 1971,  
American Journal of Psychol. , 84, 147-166.

BACHEM A.

Various types of absolute pitch. J. Acoust. Soc. Am., 1937, 9, 146-151.

BARTOSHUK A.K.

Human neonatal cardiac acceleration to sound : habituation and dishabituation,  
Perceptual and Motor Skills, 1962, 15, 15-27.

VON BEKESY G.

Experiments in hearing. 1960, New-York : Mc Graw Hill.

BLACKWELL H.R., SCHLOSBERG H.

Octave generalization, pitch discrimination and loudness thresholds in the  
white rat, 1943, Journal of Experi. Psychol., 33, 407-419.

BOYD E.F.

Visual fixation and voice discrimination in 2 month - old infants.  
In F.D. HOROWITZ (Ed.) : Visual attention, auditory stimulation, and  
language discrimination in young infants, Monog. of the Society for Res.  
in child devt., 1974, 39, 63-77.

BRIDGER W.H.

Sensory habituation and discrimination in the human neonate, 1961,  
American Journal of Psychiatry, 117, 991 - 996.

BURNS E.M.

Octave adjustment by non-western musicians - Nov. 1974, 88th meeting  
Acoustical Society of America, Vol. 56, Supplement M.3.

BURNS E.M., WARD W.D.

Intervals, Scales and Tuning.  
in The Psychology of Music, Deutsch D. (Ed.) New-York : Academic Press,  
1982, ch. 8

BUSNEL M.C. et GRANIER-DEFERRE C.

Aspects de l'audition foetale, 1983, 23èmes journées néonatales, Paris.

CARMICHAEL L. (Ed.)

Manual of Child Psychology (2nd Ed.), 1960, Wiley et Sons, N.Y. Londres.

CHANG H.W., TREHUB S.E.

Auditory processing of relational information by young infants, 1977,  
J. Of Expe. Child Psychol., 24, 324-331.

CUDDY L.L., COHEN A.J.

Melody recognition : the experimental application of musical rules, 1979,  
Canad. J. of Psychol., 33, 148-156.

DEMANY L., Mc KENZIE B., et VURPILLOT E.

Rythm perception in early infancy, 1977, Nature, 266, 718-719.

DEMANY L.

La perception de forme temporelle chez le nourrisson, 1978, Paris,  
Thèse de Psychologie, Paris V.

DEUTSCH D.

Octave generalization and tune recognition, 1972, Percept and Psychophy., 11 (6), 411-412.

DEUTSCH D.

Octave generalization and the consolidation of melodic information, 1979, Canad. J. Psychol./Rev. Canad. Psychol., 33 (3).

EISENBERG R.B.

Auditory competence in early life, 1976, Baltimore, University Park Press.

FANTZ R.L. et NEVIS S.

The predictive value of changes in visual preferences in early infancy. In J. HELLMUTH (Ed.) : The exceptional infant (vol. 1). 1967, New-York, Brunner Masel, 349-414.

FORBES H.S. et FORBES H.B.

Fetal sense reaction : hearing, Journal of Comparative Psychology, 1927, 7, 353-355.

FRANCES R.

La Perception de la musique. 1958, Paris, J.Vrin, 190-201.

GOLDSTEIN J.L.

An optimum processor theory for the central formation of the pitch of complex tones, 1973, J. Acoust. Soc. Am. 54, 1496-1516.

HALL J.W., PETERS R.W.

Pitch : it is learned ?, 1980, J.A.S.A., Sup 1, 68, S110, 2: 30.

HELMHOLTZ H.

On the sensations of tone as a physiological basis for the theory of music. 1954, Dover Publications, Inc., New-York, 2nd ed.

HOUSE W.S.

Octave generalization and the identification of distorted melodies, 1977, Percept and Psychology, 21 (6), 586-589.

- HUTT S.J., HUTT C., LENARD H.G., BERNUTH H. et MUNTJEWERFF W.J.  
Auditory responsivity in the human neonate, 1968, Nature, 218, 888-890.
- IDSON W.L. MASSARO D.W.  
A bidimensional model of pitch in the recognition of melodies, 1978, Percep and Psychophy, 24 (6), 551-565.
- IMBERTY M.  
L'acquisition des structures tonales chez l'enfant, 1969, Paris, Klincksieck, Ed. C.N.R.S.
- KALLMAN H.J.  
Similarity judgments of pitch : A test of octave equivalence, 1982, J.A.S.A., 102nd meeting of the Acoustical Society.
- KALLMAN H.J., MASSARO D.W.  
Tone chroma is functional in melody recognition, 1979, Percep. and Psychophy., 26 (1), 32-36.
- KASATKIN N.I., LEVIKOVA A.M.  
On the development of early conditioned reflexes and differentiations of auditory stimuli in infants. J. of Expe. Psychol., 1935, 18, 1-19.
- KIANG N.Y.S., SACHS H.B., PEAKE N.T.  
Shapes of tuning curves for single auditory nerve fibers, 1967, I.A.S.A., 42, 1341-1342.
- LEVENTHAL A.S., LIPSITT L.P.  
Adaptation pitch discrimination and sound localisation in neonate, 1964, Ch. Dvpt, 35, 759-767.
- LICKLIDER J.C.R.  
"Periodicity" pitch and "place" pitch, 1954, J.A.S.A., 26, 945 (A).
- LOCKHEAD G.R., BYRD R.  
Practically perfect pitch, 1981, J.A.S.A., 70 (2), 387-389.

LUNDIN R.W., ALLEN J.D.

A technique for training perfect pitch, 1962, The Psychol. Report, 12, 139-146.

MARTIN D.W., WARD W.D.

Subjective evaluation of musical scale temperament in Pianos, 1961, J.A.S.A., 33 (5), 582-588.

MASSARO D.W., KALLMAN H.J., KELLY J.L.

The role of tone height, melodic contour, and tone chroma in melody recognition, 1980, J. of Expe. Psychology : Human Learning and Memory, Vol.6 (1), 7-90.

MEHLER J.

"Connaître par désapprentissage"- "A propos du développement cognitif" in MORIN E. et PIATELLI-PAIMARINI M. Eds. : L'Unité de l'Homme (Centre Royaumont pour une Science de l'Homme), 1974, Paris, Seuil.

MOORE B.C.J.

An introduction to the Psychology of Hearing, 1982, Londres, Academic Press (2nd édition).

NORTHERN J.L. et DAVIS M.P.

Hearing in Children (2nd Edition), 1978, Baltimore, The Williams and Wilkins Compagny.

OHGUSHI K.

The origin of tonality and a possible explanation of the octave enlargement phenomenon, 1983, J.A.S.A., 73 (6).

PLOMP R.

Auditory psychophysics, 1975, Annual Review of Psychology, 26, 207-232.

PLOMP R.

Pitch of complex tones, 1967, J.A.S.A., 41, 1526-1533.

PLOMP R.

Aspects of tone sensation, 1976, Academic Press, Londres

RISSET J.C.

Hauteur et timbre des sons, 1978, Rapports IRCAM, 11/78.

RITSMA R.J.

The "Octave Deafness" of the Human Ear, IPO, Annual Progress Report, 1966.

POSE J.E., BRUGGE J.F. ANDERSON D.J., and HIND J.

Phase-locked response to low frequency tones

in Single Auditory nerve fibers of the squirrel monkey, 1967, J. Neurophysiol., 30, 769-783.

RUCKMICK C.A.

A new classification of tonal qualities, 1929, Psychol. Rev., 36, 172-180.

SALAPATEK P.

Pattern perception in early infancy. In L.B. COHEN and P. SALAPATEK (Eds), Infant Perception : from sensation to cognition, Vol. 1 Basic processes, New-York, San Francisco, London, Academic Press, 1975, p. 133-248.

SCHNEIDER B.A., TREHUB S.E., BULL D.

The development of basic auditory process in infants, 1979, Canad. J. of Psychol., 33 (4), 306-319.

SCHNEIDER B, TREHUB S.E., BULL

High frequency sensitivity in infants, 1980, Science, 207 (29).

STUMPF C.

Tonpsychologie, 1890, Hirzel, Leipzig, Vol. 2.

SUNDBERG J.E.F., LINDQUIST J.

Musical octaves and pitch, 1973, J.A.S.A., 54 (4).

TEPLOV B.M.

Psychologie des aptitudes musicales, 1966, Paris, P.U.F.

TERHARDT E.

Oktavspreizung und Tonhöhenverschiebung bei Sinustönen, 1969/70, Acustica, 22, 345-351.

TERHARDT E.

Zur Tonhöhenwahrnehmung von Klängen, 1972, Acustica, 26, 173-199.

TERHARDT E.

Pitch, consonance and harmony, 1974, J.A.S.A., 55 (5).

THURLOW W.R.

An auditory figure-ground effect, 1957, Am. J. Psychol., 70, 653-654.

THURLOW W.R., ERCHUL W.P.

Judged similarity in pitch of octave multiples, 1967, Percept and Psychophy., 22 (2), 177-182.

TREHUB S.E., BULL D., THORPE L.A.

Infant's perception of melodies : the role of melodic contour, 1982, Communication personnelle.

VURPILLOT E., RUEL J. et CASTREC A.Y.

L'organisation perceptive chez le nourrisson : réponse au tout ou à ses éléments, 1977, Bulletin de Psychologie, 30, 396-405.

WARD W.D.

Subjective Musical Pitch, 1954, J.A.S.A., 26 (3).

WARD D.W.

Absolute pitch, 1963, Part I - Sound, 2 (3), Part II - Sound, 2 (4).

WEVER E.G.

Theory of hearing, New-York, Wiley and sons, 1949.

WORMITH S.J., PANKHURST D. and MOFFITT A.R.

Frequency discrimination by young infants, 1975, Ch. Dvpt, 46, 272-275.