

Indifférenciation de la langue et de la musique au début de la vie

« The vocal production of musical elements like most other prerequisites of human language do occur isolated in one or another group of animals, but in none do they coincide to permit the explosive development of human language » (Marler, 1969, in Papousek & Papousek, 1981 : 169).

« La parole et la musique sont deux exemples d'une communication humaine formalisée qui passe par des signaux acoustiques. Tous deux doivent être destinés aux mêmes systèmes perceptifs et cognitifs. Les limitations et les capacités de ces systèmes doivent contribuer fortement, à la fois, au développement de la parole et de la musique » (Carlson et al., 1989 : 569).

Table du second chapitre

1. Préalables à la description du développement linguistique et musical de l'enfant

- 1.1. Perception et production
 - 1.1.1. La perception précède la production
 - 1.1.2. La perception et la production entretiennent des rapports étroits au moment de l'acquisition
- 1.2. La perception est globale avant d'être analytique
 - 1.2.1. Les principes gestaltistes et l'analyse de scènes
 - 1.2.2. Les processus de groupement sont opérants au début de la vie
 - 1.2.3. Traitement global versus traitement local

2. Les éléments musicaux, point d'entrée dans la LM et dans le SMM

- 2.1. Une oreille mature et ouverte sur la musique de l'environnement bien avant la naissance
 - 2.1.1. Prédominance des éléments musicaux dans le monde sonore intra-utérin
 - 2.1.2. Les fœtus et leur LM
 - 2.1.2.1. Les voix de son environnement
 - 2.1.2.2. Les éléments musicaux de la langue
 - 2.1.3. Des fœtus mélomanes
 - 2.1.3.1. Des réactions différentes selon le type de musique
 - 2.1.3.2. Une mémoire musicale réactivable après la naissance
 - 2.1.4. La « culture prénatale » du fœtus
- 2.2. Sensibilité du bébé aux éléments musicaux : le « trésor inépuisable » des ressources enfantines
 - 2.2.1. Sensibilité au rythme
 - 2.2.2. Sensibilité aux contours mélodiques
 - 2.2.2.1. Les contours mélodiques de la musique
 - 2.2.2.2. Les contours d'intonation de la langue
 - 2.2.3. Sensibilité à la prosodie affectant les mots
 - 2.2.4. Capacités commune de segmentation pour la LM et le SMM
- 2.3. L'amorçage de l'attention par les éléments musicaux
 - 2.3.1. Richesse et redondance du signal
 - 2.3.2. Les signaux d'amorçage dans la langue
 - 2.3.3. Une théorie de l'amorçage valable pour la musique ?
- 2.4. Ajustements vocaux inconscients adaptés à la sensibilité musicale du bébé
 - 2.4.1. Préférence de l'enfant pour le langage et la musique adressés
 - 2.4.1.1. Le langage adressé
 - 2.4.1.2. Le chant adressé
 - 2.4.2. Caractéristiques du langage adressé à l'enfant
 - 2.4.2.1. Une pédagogie « inconsciente »...
 - 2.4.2.2. ...qui évolue avec l'âge de l'enfant

- 2.4.3. Caractéristiques du chant adressé à l'enfant
 - 2.4.3.1. Les berceuses
 - 2.4.3.2. Les chants de jeux
- 2.4.4. Points communs entre le LAE et le CAE

3. D'une base commune à une différenciation progressive

- 3.1. Perception
 - 3.1.1. Jusqu'à 6 mois, l'enfant est un « super-discriminateur »
 - 3.1.2. Spécialisation progressive sur la LM aux dépens de la richesse perceptive de l'enfant
 - 3.1.2.1. Focalisation sur la structure vocalique de la LM à partir de 6 mois
 - 3.1.2.2. Focalisation sur la structure consonantique de la LM à partir de 10/12 mois
 - 3.1.3. Spécialisation progressive sur le SMM
 - 3.1.3.1. Définition de l'acculturation musicale
 - 3.1.3.2. Bons patterns ou intégration des régularités de l'environnement
 - 3.1.3.3. Evolution de la perception des gammes non natives
- 3.2. Production
 - 3.2.1. Maîtrise progressive de l'appareil phonatoire
 - 3.2.2. Le babillage canonique, étape de différenciation entre la langue et la musique ?
 - 3.2.3. Le langage émergent
 - 3.2.4. Le chant émergent

Bibliographie du chapitre II

Pendant la première année de la vie, l'acquisition de la langue maternelle est avant tout perceptive. Cette imprégnation progressive avec les sons de l'environnement commence de manière très précoce, puisque l'oreille est mature bien avant la naissance. Et il se trouve que l'information qui réussit à franchir les obstacles de la paroi abdominale et du liquide amniotique de la mère est avant tout... musicale. Donc, non seulement, la perception est première, mais les éléments musicaux sont perçus en premier par le fœtus. Il n'est donc guère étonnant que, après sa naissance, le bébé développe une forte sensibilité à la musique et aux caractéristiques prosodiques de sa langue maternelle. Comme s'ils devinaient cette sensibilité, les adultes qui entourent l'enfant se mettent à lui parler d'une façon spéciale, en exagérant les paramètres prosodiques. Cette musicalité du langage adressé à l'enfant le rapproche d'une certaine manière des « langues musicales » dont nous avons parlées dans le chapitre précédent. Et lorsque l'on découvre que l'entourage du bébé adapte également sa façon de chanter pour faire ressortir la mélodie et le rythme, on en vient à se demander si des processus communs ne sous-tendent pas l'acquisition de la langue et du système musical maternels, voire si le traitement de la langue et celui de la musique ne sont pas indifférenciés au début de la vie. Comme le souligne Fassbender (1995 : 98-99),

« il est légitime de supposer que [la] sensibilité [du bébé] se fonde partiellement sur la capacité à grouper et à segmenter l'information auditive sur la base des principes de la Gestalt. Au départ, la perception de la langue et de la musique font appel à une base identique. Ce qui n'exclut pas qu'elles puissent suivre des voies développementales distinctes dès l'instant où une signification est attachée à une information acoustique particulière à la faveur de l'interaction sociale découlant de l'intuition parentale ».

Dans les premières vocalisations infantiles, il est par ailleurs très difficile de faire la part entre ce qui relève du musical et du linguistique. Pendant les quelques pages qui vont suivre, empruntons au bébé ses oreilles et sa voix, pour découvrir le monde essentiellement musical qui est le sien pendant la première année de sa vie.

1. Préalables à la description du développement linguistique et musical de l'enfant

1.1. Perception et production

1.1.1. La perception précède la production

« C'est que, en définitive, la langue n'est ni dans l'onde sonore, ni dans les organes phonateurs mais dans la discrimination auditive, à savoir dans les habitudes de discrimination propres à chaque communauté. De façon imagée et paradoxalement, on peut dire que les êtres humains parlent avant tout avec leurs oreilles » (Martin, 1996 : 103).

Dans l'acquisition de la langue maternelle (désormais LM) et du système musical maternel (désormais SMM), la perception précède la production pour une raison avant tout physiologique : à la naissance, le tractus vocal du nourrisson n'est pas assez développé pour pouvoir produire autre chose que des vagissements et des cris et les capacités de l'enfant à produire des sons sont contraintes par l'état de développement de son organisme (configuration du conduit vocal en diagonale plus proche de celle du primate que celle de l'adulte, en coude, liée à la position debout ; la configuration infantine ne permet pas la maîtrise des articulateurs avant la seconde moitié de la première année). En revanche, les oreilles du nourrisson sont grandes ouvertes sur le monde puisque **l'oreille du fœtus est fonctionnellement mature dès le cinquième mois de gestation** (cf. 2.1.1.). En effet, chez l'être humain, le développement de la cochlée se fait in utero, contrairement à beaucoup d'autres espèces (Herbinet et Busnel, 1983). Elle s'enroule dès la 6^{ème} semaine de gestation et atteint son complet développement morphologique vers la 10^{ème} semaine (3 mm et enroulée sur 2 tours ½, comme une cochlée adulte). Elle entre en fonction vers 18 à 20 semaines, en dépit d'une immaturité cellulaire qui s'achèvera vers le huitième mois (Lecanuet, 1997 : 22). L'oreille interne est donc prête à fonctionner dès 4 mois ½, mais Shahidullah et Hepper (1992) ont observé une écoute réactive à des sons purs (250-500 Hz) encore plus précoce, dès la 16^{ème} semaine. Les fœtus de 30 à 36 semaines réagissent lorsqu'ils sont exposés à des sons intenses (< 90 dB de 2 à 5 secondes) (Hepper et Shaidullah, 1994). On observe les mêmes effets avec la musique (extraits de 5 minutes), bien que les fœtus âgés de 32-34 semaines réagissent moins que les fœtus de 35 à 37 semaines (Smith et al., 2000), révélant une différence de maturation auditive.

Certes, les seuils auditifs sont au départ élevés (100 dB), mais ils s'abaissent très rapidement pour permettre la discrimination des fréquences.

A sa naissance, le nourrisson est donc bien loin d'être une « ardoise vierge », totalement libre dans ses premiers jours de toute influence de son environnement linguistique et musical. La période de latence entre la perception et la production, rendue nécessaire par l'imaturité physiologique de son appareil vocal, joue un rôle primordial dans l'acquisition. En effet, pendant les premiers mois de vie, l'enfant s'imprègne de son environnement et notamment de sa LM et de son SMM ; il apprend progressivement à repérer les caractéristiques sonores saillantes au sein du flux sonore, et une fois qu'elles sont mémorisées, à les reconnaître. Selon Papousek et Papousek (1981 : 168), le décalage entre perception et production révèle la nécessité d'une mémoire à long terme au cours de l'acquisition.

1.1.2. La perception et la production entretiennent des relations étroites au moment de l'acquisition

Ce décalage ne signifie pas pour autant que perception et production soient deux processus totalement indépendants. Car si les activités de perception sont entravées par un système auditif déficient, la production de la LM et du SMM est gravement perturbée. Ainsi, on n'observe pas de babillage canonique chez les bébés sourds profonds (Vinter, 1992). Comme les autres sons de son environnement, les premières productions de l'enfant deviennent à leur tour un « input » pour son système perceptif, constituant une sorte de boucle où perception et production interagissent. On trouve de fortes empreintes de cette interaction dans le proto-langage¹ du bébé, qui possède déjà les caractéristiques prosodiques de base de la LM. Ces caractéristiques sont suffisamment saillantes pour que la LM d'un bébé de 8 mois soit identifiable par des auditeurs adultes appartenant à la même communauté linguistique (De Boysson Bardies, 1993). L'acculturation avec l'environnement va ainsi dans le sens du lien entre perception et production, puisque la production est influencée par la perception. Selon Jusczyk, perception et production sont très étroitement liées au moment de l'apparition des premiers mots.

¹ Proto-Langage ou pseudo langage : « *En situation d'interaction, l'enfant produit du « protolangage », discours construit comme une parole dépourvue de mots* » (Konopczynski et Vinter, 1989).

Alors que, dès 8-10 mois, le bébé est capable de reconnaître des mots en contexte (Benedict, 1979) et vers 11-12 mois, des mots hors contexte (Hallé et De Boysson-Bardies, 1994), il est remarquable qu'au niveau de la production, ses premiers mots n'émergent qu'aux alentours de 12 mois au sein du babillage mixte (cf. tableau récapitulatif « Les mots de la LM : reconnaissance, compréhension et production », Volume II, annexe 1.5.).

Cependant, dans les études portant sur l'acquisition du langage, perception et production sont souvent étudiées de façon isolée et la tendance reste majoritairement à la séparation, notamment à cause de la spécialisation des chercheurs dans l'un ou l'autre domaine. Pour Jusczyk (1997), ce « divorce » a été favorisé, parce que les informations collectées sur le développement de la perception et de la production l'ont été à des âges différents. Ainsi, beaucoup de travaux sur la perception se sont focalisés sur des enfants de 6 mois, voire plus jeunes, tandis qu'en production, les travaux se sont intéressés majoritairement aux enfants de plus d'un an. Cependant, les choses ont évolué depuis une vingtaine d'années et les productions précoces du bébé sont maintenant très bien décrites.

L'étude des rapports entre perception et production est d'une importance fondamentale pour comprendre comment les unités de représentation sont extraites du signal sonore. Cette conversion implique le passage d'un niveau acoustique, à un niveau abstrait, linguistique. La théorie qui pourrait décrire un tel processus de conversion serait capable non seulement d'éclairer la façon dont un enfant apprend sa LM, mais également d'unifier les deux niveaux. L'intérêt d'une théorie « unifiée » n'a bien sûr pas échappé aux chercheurs et sa nécessité a encore été soulignée lors des conférences plénières qui ont clôturé le dernier Congrès International des Sciences Phonétiques de San Francisco (1999). Il existe un certain nombre de théories sur l'acquisition qui tentent de faire interagir perception et production. La théorie motrice de la perception est l'une des plus anciennes et des plus connues (Liberman et al., 1967). Selon cette théorie, les sons sont perçus en fonction de leurs lieux d'articulation. Au moment du babillage, les bébés établiraient progressivement des relations entre les diverses articulations qu'ils produisent et les sons qui en résultent. Lindblom et al. (1984) proposent un modèle d'interaction entre les contraintes perceptives et articulatoires. Selon eux, la « discriminabilité » et la « prononcabilité » seraient des contraintes tout aussi importantes qu'écouter et parler. Selon leur modèle, les représentations perceptives des mots et leurs représentations articulatoires se combineraient en une généralisation plus abstraite et applicable aux deux systèmes.

Cette généralisation correspondrait au niveau phonologique. Une autre théorie, celle de Gathercole et Baddeley (1993) posent l'hypothèse que perception et production seraient coordonnées par l'intermédiaire d'une boucle phonologique, active chez l'enfant tout au long de sa première année. Cette boucle, en prise directe avec le flux sonore, permettrait de retenir l'information phonologique pendant une à deux secondes. Grâce à cette mémoire à court terme, l'enfant pourrait mettre constamment à jour son répertoire phonologique (capacité de base pour imiter).

1.2. La perception est globale avant d'être analytique

1.2.1. Les principes gestaltistes et l'analyse de scène

Au début du siècle dernier, les psychologues gestaltistes ont décrit un certain nombre de lois qui permettent au sujet, à partir d'un ensemble d'informations visuelles, de grouper les données provenant d'une même source (Wertheimer, 1923). Ces lois interviennent dans une étape préattentive (Neisser, 1967) et permettent la formation d'unités qui optimisent le traitement cognitif (Bregman, 1990). Elles permettent d'organiser au niveau perceptif le flot d'informations. Par exemple, si des éléments se déplacent en même temps et dans la même direction, ils seront groupés en une seule forme, obéissant ainsi à la loi de destin commun. Parmi ces formes, certaines sont plus prégnantes que d'autres ; on parle alors de « bonnes formes ». Une bonne forme satisfait à plusieurs de ces lois de groupement en raison de sa simplicité, de sa régularité et de sa symétrie. Le système perceptif « construit » donc une interprétation de son environnement en projetant des formes sur la réalité sensorielle. Initialement décrits pour la modalité visuelle, ces principes de groupement fonctionnent également pour l'audition (Deutsch, 1982 ; Bregman, 1990, 1994 ; Bigand, 1994). Lorsqu'un individu se trouve confronté pour la première fois à une mixture sonore, son système auditif doit extraire de cet ensemble des régularités acoustiques pour construire des formes. Par exemple, il regroupera les sons ayant une même évolution temporelle, avec un début et une fin simultanés. De la même manière, il groupera les éléments possédant le même timbre ou encore, proches du point de vue de leur fréquence. Pour désigner ce processus, Bregman (1990 : 38) parle d'« *analyse primitive de scènes auditives* ». L'analyse de scène est une méthode générale de partition d'une mixture en sources acoustiques distinctes à partir des régularités acoustiques de l'environnement sonore :

« L'analyse de scène désigne la stratégie par laquelle l'ordinateur regroupe l'ensemble des propriétés visibles – pourtours, textures de surface, couleurs, distances, etc. – d'un même objet. Ce processus semble déterminer la forme globale et les propriétés correctes d'un objet. Par analogie, l'analyse des scènes auditives est le processus réunissant en une unité perceptive l'ensemble des signaux provenant, dans une période de temps donnée, d'une seule source sonore de l'environnement » (Bregman, 1994 : 112).

Dans l'analyse de scène, l'attente et l'attention jouent un rôle primordial. En effet, la perception de l'auditeur sera différente s'il est prévenu qu'il va écouter un extrait de musique ou des sons de parole. Par exemple, il ne reconnaîtra pas l'objet « voyelle » à l'écoute des trois premiers formants synthétisés d'une voyelle, s'il n'a pas été prévenu au préalable que l'objet à percevoir était une voyelle. Dans un test de perception, les consignes sont donc d'une importance fondamentale, puisqu'elles vont « orienter » l'attention de l'auditeur et conditionner sa perception. Peu à peu, l'auditeur va mémoriser des formes dans son environnement et selon Bregman (1994), les activer soit de manière purement automatique (reconnaissance automatique), soit de manière volontaire en « cherchant à entendre » lorsqu'il sera confronté à une nouvelle mixture sonore (attention volontaire).

1.2.2. Les processus de groupement sont opérants au début de la vie

Si les processus de groupement décrits par les gestaltistes et l'analyse de scène sont opérants au début de la vie, ils devraient permettre à l'enfant, pendant la période de latence où il ne peut pas produire, de « rentrer » dans sa LM et dans son SMM, c'est-à-dire de détecter des formes à partir d'un tout vaste et indifférencié. Et il est effectivement apparu que certains principes de groupement sont déjà effectifs de façon très précoce chez le bébé et lui permettent un pré traitement du flux sonore, qu'il soit langagier ou musical.

Parmi ces principes, la **loi de proximité** permet au système perceptif de regrouper des éléments proches les uns des autres. Ainsi, pour la hauteur, si des notes successives se répartissent sur deux registres différents, l'auditeur adulte percevra deux lignes mélodiques distinctes évoluant parallèlement. Ce phénomène a été baptisé par Bregman « séparation des flux auditifs » (Bregman, 1978 ; Bregman, 1981). Les bébés sont capables de faire le même type de groupement, et ce de manière très précoce. Demany (1982) a testé **44 bébés âgés de 1 à 3 mois et demi, qui ont regroupé les stimuli auditifs en fonction de leur proximité de fréquence**. Cette capacité a également été relevée chez des bébés de 2 à 5 mois (Fassbender, 1993).

Demany (1977) a montré que **la loi de proximité fonctionnait également au niveau temporel, et ce, à partir de 2 mois ½**. Les bébés sont en effet capables d'organiser des sons successifs qui ne diffèrent pas par la hauteur en fonction des intervalles de temps qui séparent ces sons, les intervalles de grande longueur étant interprétés comme des pauses. **La loi de proximité s'applique également pour le contenu spectral**. En revanche, **les différences d'intensité ne constituent pas un indice auditif suffisant** pour permettre un groupement chez des bébés de 7 à 11 mois (Thorpe et Trehub, 1989).

Un autre principe, la **loi de similitude** permet au système perceptif de regrouper des éléments semblables plutôt que dissemblables. Dans une série d'études, Thorpe et ses collaborateurs (Thorpe et Trehub, 1989 ; Thorpe et al., 1988) utilisent un pattern de 6 tons séparés par un même intervalle de temps. Les trois derniers tons sont différents des trois premiers du point de vue de la Fo, de la structure spectrale ou de l'amplitude (XXXOOO). Une pause est ensuite introduite entre les deux groupes ainsi formés (XXX OOO) ou à l'intérieur d'un des deux groupes (XXXO OO). **Les bébés de 6 à 9 mois réagissent à la dernière séquence, car elle viole le principe de regroupement par similitude**. Si le changement respecte la structure de la séquence standard, il n'est pas remarqué. Ce faisant, **les bébés se comportent exactement comme les auditeurs adultes** (Thorpe, 1985). Le phénomène d'illusion de durée résulte de la loi de similitude. En effet, les auditeurs repèrent une pause entre les groupes, même si celle-ci n'existe pas matériellement dans le signal sonore. Selon Trehub (1989 : 308) :

« des discontinuités au niveau de la hauteur, du timbre ou de la sonie indiquent des aspects importants de la structure d'un pattern et déclenchent une illusion de durée qui guide les processus primitifs de segmentation ou de ségrégation. Plus la discontinuité ou le changement d'un de ces paramètres est grand, moins la détection des pauses entre les groupes est précise. De ce fait, les auditeurs ne tiennent pas compte de l'information temporelle disponible et imposent une structure temporelle correspondant aux aspects de la configuration générale ».

Cette illusion de durée fonctionne également dans le traitement de la langue, puisque l'auditeur regroupe les événements en fonction de leurs similarités. Selon Studdert-Kennedy (1975), cette illusion ne pourrait pas fonctionner dans une langue étrangère.

Les principes de groupements sont donc opérants pour la langue et pour la musique et cela, de manière très précoce. Les bébés sont capables d'organiser l'espace acoustique en regroupant les éléments qui sont proches les uns des autres ou qui partagent des caractéristiques communes (fréquentielles, temporelles, de timbre).

Ces aptitudes sont d'une importance capitale pour la segmentation du flux sonore, car elles relèvent d'une capacité générale à grouper et à segmenter, qui fonctionne pour la langue et la musique.

1.2.3. Traitement global vs traitement local

L'analyse de scène et les lois de type gestaltiste impliquent un traitement de type global ou holistique des données sensorielles. Dans un traitement de ce type, c'est la forme, l'ensemble qui prime sur les éléments. En effet, si on change un seul élément, c'est l'ensemble qui se modifie. En revanche, la forme ne sera pas modifiée si on change l'ensemble des éléments, tout en maintenant constants les rapports entre ces éléments. Dans la perception visuelle, l'information globale précède l'information locale (Kimchi, 1992 pour une revue) et elle est identifiée plus rapidement que l'information locale (Navon, 1977 ; Proverbio et al., 1998). Le traitement global est en effet nécessaire avant un traitement de l'information plus analytique qui permette de décomposer une forme en ses éléments constitutifs. Dans la perception des contours mélodiques, il est mis en jeu, que ce soit dans la langue ou dans la musique (traitement des patterns globaux de changement de direction de hauteur, forme des contours mélodiques). On le retrouve également dans la perception du registre de hauteur et dans l'approximation des intervalles mélodiques et harmoniques. Il permet la discrimination de deux patterns et la reconnaissance de ces patterns, mais en aucun cas leur catégorisation. Il est donc limité, car les signaux sont traités comme des touts qui ne sont pas décomposables et donc pas recombinaisons. C'est une capacité prélinguistique qui ne serait pas spécifique aux êtres humains, mais que l'on rencontrerait chez certains mammifères, les primates notamment.

En revanche, le traitement local qui est un traitement analytique, ne se rencontre que chez les êtres humains. Il permet d'isoler les éléments qui composent une forme, puis de comparer ces éléments pour ensuite les classer dans des catégories et leur attribuer une étiquette (processus de catégorisation). Selon Deutsch (1969), il intervient en musique dans la perception des phonèmes, des accords et des classes de hauteur des notes. Dans la langue, il intervient dans la perception des phonèmes. Le traitement analytique ne peut pas intervenir avant le traitement global. Cependant, le passage d'un type de traitement à l'autre est fortement conditionné par la difficulté relative de la tâche à réaliser par l'auditeur. Une tâche difficile conduit à un mode d'écoute synthétique (global) où l'auditeur manifeste peu de sensibilité pour les composantes individuelles du pattern.

En revanche, une tâche facile conduit à un traitement analytique, où l'auditeur se montre au contraire sensible aux composantes du pattern. Par ailleurs, il existe des contextes « facilitants » qui permettent d'orienter l'auditeur vers un traitement plus analytique. Si le contexte dans lequel apparaît un pattern donné respecte les principes utilisés dans l'environnement de l'auditeur (« bon contexte »), le pattern sera traité plus facilement, même s'il n'a jamais été rencontré auparavant.

Pour comprendre le passage d'un traitement global à un traitement analytique, prenons l'exemple de la reconnaissance d'un air familier. Chez les adultes, la reconnaissance de la mélodie d'un extrait musical est relative, elle dépend de la relation entre les différents sons successifs, c'est-à-dire les intervalles entre ces sons. Les intervalles sont très importants dans la reconnaissance des mélodies familières. Par exemple, c'est une tierce majeure qui marque le début de « *A la Claire Fontaine* ». Si les intervalles sont maintenus, mais que la hauteur de départ est changée, nous percevons les mélodies comme équivalentes. Le traitement par intervalles caractérise la perception d'une mélodie familière (traitement de type analytique). En revanche, le traitement est de type global lorsque nous percevons pour la première fois une mélodie. Dans ce cas, c'est le contour mélodique qui prime et notamment, les changements de sa direction (Bartlett et Dowling, 1980 ; Dowling, 1978). Cependant, si ces mélodies non familières ont été composées selon les principes musicaux familiers à l'auditeur (par exemple, la musique tonale), elles seront codées et mémorisées plus facilement (Francès, 1958 ; Zenatti, 1969). Selon Bartlett et Dowling (1980), si une oreille connaît les principes musicaux qui gouvernent la musique tonale, le traitement sera plus analytique, car le contexte est facilitant (stratégie de traitement par intervalles et non par contours). Les auditeurs musiciens qui ont acquis une meilleure connaissance de ces principes grâce au solfège et à une écoute intensive de la musique pourraient manifester un traitement plus analytique (voir chapitre III, 4.2.).

Au niveau du développement, l'apparition des capacités de traitement analytique dépend de la maturation corticale du cerveau de l'enfant. D'après LeCours (1975), la myélogénèse du cortex auditif n'est pas complètement achevée avant l'âge de 5 ans et avant cette maturation, le jeune enfant utiliserait un traitement auditif subcortical, plus adapté au traitement global (Whitaker, 1976). Jusqu'à 5 ans, la perception de l'enfant serait donc dominée par un traitement de type global qui expliquerait ses difficultés à intégrer les détails en un ensemble structuré.

Cette domination du global sur le local ne veut pas pour autant dire que l'enfant n'a pas de capacités de traitement analytique. Sans elles, il ne pourrait d'ailleurs pas distinguer et mémoriser les unités de la langue et de la musique. Il semblerait que des changements importants interviennent au niveau cognitif dès la fin de la première année de sa vie, qui lui permettent de développer des capacités plus analytiques. Ces « nouvelles » capacités lui permettraient de repérer et de mémoriser les contrastes qui constituent le stock d'unités de sa LM et de son SMM. Cette réorganisation expliquerait la perte des capacités de discrimination constatées à l'âge de 10-12 mois et l'ajustement à certaines catégories de sons (cf. partie 3.1.2.). Comme c'est le cas pour l'oreille musicienne, cette réorganisation pourrait intervenir à un moment où l'enfant a emmagasiné suffisamment de « connaissances » sur sa LM et son SMM pour que le traitement devienne plus analytique dans les contextes facilitants.

« Thus, in the early months, the infant auditory system may be better equipped to process relatively simple auditory patterns such as the melodic pitch patterns of motherese than the more complex auditory patterns of normal adult speech [...] Such holistic pattern recognition constitutes a prelinguistic auditory processing mode, for which the exaggerated pitch contours of mothers' speech may be a particularly appropriate form of acoustic stimulation » (Fernald, 1984 : 14).

Ces fameuses capacités dont le bébé fait preuve pour discriminer les contrastes de nombreuses langues (cf. 3.1.1.) mettraient en jeu un traitement de type global et non pas analytique, le traitement global permettant de discriminer deux patterns, mais non de les extraire et de les mémoriser.

2. Les éléments musicaux, point d'entrée dans la LM et le SMM

2.1. Une oreille mature et ouverte sur la musique de l'environnement bien avant la naissance

2.1.1. Prédominance des éléments musicaux dans le monde sonore intra-utérin

Les sons qui parviennent à la cochlée du fœtus sont altérés et transformés par un environnement bruyant. Des enregistrements intra-abdominaux (hydrophone placé dans le vagin après la rupture des membranes avant ou après l'accouchement) permettent de se placer dans les mêmes conditions d'audition que le fœtus et révèlent un bruit de fond dans les fréquences basses (500-700 Hz, Lecanuet, 1997 : 24) constitué des bruits internes de l'organisme maternel (battements de cœur, respiration, digestion et pour ce qui nous intéresse, mouvements laryngés). Cependant, ce bruit de fond est moins important que ce que l'on pensait au départ (72 à 96 dB SPL, Lecanuet, 1997 : 24) et en l'absence de bruits gastriques, l'utérus est relativement tranquille (Lecanuet, 1995). Les pressions sont comparables à celles du milieu extérieur dans un environnement silencieux, notamment pour les basses fréquences (<300-500 Hz, de longueur d'onde supérieure au diamètre maternel ; Lecanuet, 1997 : 25). A cause de la paroi abdominale et du liquide amniotique qui se comportent comme un filtre passe-grave, les sons arrivent de manière atténuée. Cette atténuation croît avec la fréquence du signal, mais **préserve les fréquences basses correspondant aux caractéristiques prosodiques**. Ainsi, chez la brebis, Armitage et al. (1980) ont observé peu d'atténuation de la Fo entre 175 et 500 Hz. Toutes les informations prosodiques (mélodie, rythme, pauses) atteignent donc la cochlée du fœtus, ainsi que quelques mots isolés et bien articulés, mais qui restent peu intelligibles (« *papa, maman, bonjour...* »). **Les voix, bien que déformées, émergent clairement du bruit de fond et, en particulier, la voix maternelle**. Celle-ci est moins atténuée (8 dB) que les voix externes (20 dB, Querleu et al., 1988) et selon Lecanuet (1997 : 25), elle est même supérieure de 5 dB au niveau d'émission externe, à cause de la double transmission des sons (conduction osseuse et transmission aérienne). **Ce sont donc les éléments musicaux qui sont favorisés dans la perception que le fœtus a de son environnement sonore.**

2.1.2. Les fœtus et leur LM

2.1.2.1. Les voix de son environnement

Grâce au cardiographe qui permet d'observer les modifications du rythme cardiaque du fœtus et à l'échographe, qui permet de visualiser ses mouvements, les chercheurs peuvent développer des protocoles d'habituation pour étudier les préférences du fœtus pour un type de stimulation sonore et ses capacités à discriminer différents stimuli. Ainsi, une fois qu'il s'est habitué à un stimulus, on lui fait écouter un nouveau stimulus. Si son rythme cardiaque se modifie ou s'il se met à bouger, c'est qu'il distingue le nouveau stimulus du précédent (la mémoire et l'apprentissage sont donc mis en jeu). C'est avec cette technique qu'il a été possible de démontrer la sensibilité des fœtus proches du terme à la **variation interlocuteurs**. Ils sont en effet capables de distinguer les caractéristiques vocales de deux personnes différentes, mais également les **voix de femmes des voix d'hommes**. La conversation exerce un effet calmant sur eux, mais ils deviennent particulièrement attentifs lorsqu'ils entendent la **voix de leur mère** (Gentil-Neymon, 1993). Après la naissance, le nouveau-né manifeste un certain nombre de préférences qui témoignent de son exposition prénatale : **il préfère les voix humaines par rapport aux autres sons** (DeCasper et Fifer, 1980). **Il préfère les voix féminines aux voix masculines** (Mehler et al., 1978) et, en particulier, **la voix de sa mère** (DeCasper et Fifer, 1980), même si celle-ci est filtrée et seulement après 12 heures de contact post-natal (Busnel, 1997 : 37). Si son père lui a parlé pendant la grossesse, il marquera une préférence pour sa voix (Pugin, 1983), sinon, il ne la reconnaîtra pas, même après un contact post-natal de 4 à 10 heures (DeCasper et Prescott, 1984). Selon Lecanuet (1995 : 34-35), l'expérience prénatale de l'enfant avec la voix de sa mère pourrait expliquer les résultats de Hepper et al. (1993), qui montrent que les nouveau-nés discriminent une voix produisant des contours prosodiques exagérés d'une voix normale parlant sans exagération, uniquement si cette voix est celle de leur mère.

2.1.2.2. Les éléments musicaux de la langue

En ce qui concerne les paramètres prosodiques, Nowik-Stern et al. (1998) ont montré que les prématurés regardent plus longtemps le haut parleur lorsqu'il diffuse une voix s'adressant à un bébé plutôt qu'à un adulte. **Ils préfèrent donc le langage adressé (LAE) dont les caractéristiques prosodiques sont fortement exagérées** (voir 2.4.1.1.).

Cette préférence n'est pas étonnante étant donné que la transmission de la prosodie est privilégiée pendant la vie fœtale. Une autre étude vient confirmer cette préférence : Busnel (1997) a montré, dans un travail concernant 200 mères dialoguant avec leur enfant avant et après la naissance, que la majorité des enfants se montrait plus sensible au langage leur étant adressé qu'au langage adulte. Dès 2 jours, des nourrissons préfèrent le LAE (Fernald, 1985 ; Cooper et Aslin, 1990). Par ailleurs, l'unité prosodique de base, la **syllabe est reconnue in utéro**. Dans une étude réalisée par Lecanuet et al. (1993), les fœtus se sont montrés capables de distinguer l'inversion « *biba* » / « *babi* ». Enfin, à 4 jours, ils ont pu **repérer les frontières prosodiques** qui différencient le groupe bi-syllabique « *mati* » dans « *panorama typique* » et « *mathématicien* » (Christophe et al., 1994).

S'ils font preuve d'une sensibilité envers les caractéristiques prosodiques de la parole, est-ce que les fœtus sont capables de mémoriser des extraits de parole auxquels ils ont été régulièrement exposés ? Mehler et son équipe ont mené une série d'expérimentation à Baudelocque entre 1986 et 1990. Ils ont montré que des bébés russes, français, anglais et italiens âgés de 4 jours auxquels leur mère avaient lu régulièrement une histoire dans leur LM pendant les cinq mois précédant l'accouchement s'arrêtaient de pleurer en entendant à nouveau l'histoire après leur naissance. Par ailleurs, l'histoire est reconnue même lorsqu'on lui applique un filtre passe-bas (Spencer et al., 1987). Et, lorsqu'elle est lue par une autre personne que la mère, elle est tout de même reconnue, ce qui implique que **le bébé se fie à des indices acoustiques et probablement prosodiques** (DeCasper et Spence, 1986). Par ailleurs, dans les expérimentations de Mehler, les bébés français ont réagi à leur LM pendant les phases d'habituation (alors que la voix n'était pas celle de la mère). Ils reconnaissent donc non seulement les caractéristiques sonores de l'histoire qui leur a été lue avant la naissance, **mais également leur LM**. Cette familiarité des nouveau-nés avec leur LM est confirmée avec des bébés « anglophones » capables de distinguer l'anglais de l'espagnol à 2 jours (Moon et al., 1993 ; Bertoncini, 1993). Ils se montrent même capables de **distinguer les grands types de langues sur la base de la prosodie** (anglais-hollandais / espagnol-italien). A deux jours, le nourrisson est également **capable de reconnaître les différentes émotions** (joie, tristesse, colère) qui affectent une voix féminine enregistrée, **avec toutefois une préférence pour la joie. Ces émotions ne sont reconnues que dans la LM de la mère et dans les deux langues pour les bébés bilingues.**

2.1.3. Des fœtus mélomanes

2.1.3.1. Des réactions différentes selon le type de musique

Comme ils préféraient les voix aux autres types de stimuli, **les nouveau-nés préfèrent les tons musicaux par rapport au bruit** (Levarie et Rudolph, 1978). Tout comme la parole, **la musique exerce un effet calmant sur les fœtus** (Lecanuet et al., 2000) : les mouvements de 43 fœtus à terme (37 semaines de gestation) deviennent moins variables à l'écoute des berceuses de Brahms (n°40). Chapman (1975) a fait écouter les mêmes berceuses à des prématurés pendant 5 minutes lors de séances intervenant six fois par jour et il a constaté chez eux une **prise de poids plus importante avec la musique qu'avec la voix**. Mosser (1990) a montré la même chose avec 104 fœtus de 36 semaines, mais avec un apaisement plus rapide avec les chants grégoriens et les chants traditionnels qu'avec les berceuses. Une étude en cours de Kisilevsky, avec l'équipe de Lecanuet confirme ces résultats mitigés pour les berceuses avec des fœtus âgés de 38 semaines en sommeil actif. A l'écoute des berceuses de Brahms, ils observent quelques effets sur les fœtus au début et à la fin de l'extrait musical, mais peu pendant les extraits. **Les fœtus ne réagissent donc pas de la même manière selon le type de musique.**

2.1.3.2. Une mémoire musicale réactivable après la naissance

En ce qui concerne la mémorisation d'extraits musicaux, **on constate le même effet d'apprentissage que pour la LM**. Les nourrissons sont **capables de reconnaître une mélodie** qui leur a été chantée par la mère pendant sa grossesse d'une autre mélodie, et ce, même sans les paroles, sur « *la, la, la* » (Panneton, 1985²). Puisque l'information segmentale ne change pas, ils ne peuvent s'appuyer que sur des indices mélodiques. Hepper (1991) a montré que les nourrissons de 2 à 4 jours **reconnaissent le générique de la série de télévision « Neighbours », série que leur mère regardait régulièrement pendant la grossesse** : à son écoute, ils s'arrêtent de bouger et leur rythme cardiaque décroît. Dans des études plus anciennes, Feijoo (1975, 1981) avait fait écouter à 23 femmes enceintes (24 séances de relaxation, intervenant de la 14^{ème} à la 36^{ème} semaine de gestation) une phrase musicale de 12 secondes, extraite de « *Pierre et le Loup* » de Prokofiev.

² Cité par Lecanuet (1995 : 35).

Cette phrase est jouée par le basson, instrument dont la bande de fréquence est inférieure à 2000 Hz. Feijoo avait ensuite réalisé deux tests avec 15 enfants, le premier étant prénatal et intervenant une semaine après la dernière stimulation, le second intervenant immédiatement après la naissance (6 mn au plus tard). **A l'écoute de la phrase musicale interprétée par le basson, les fœtus se sont mis à bouger.** Quant aux nouveau-nés, ceux qui criaient s'arrêtèrent de pleurer (9 sur 15), ceux qui avaient les yeux fermés les ont ouverts (6 sur 15) et ils ont effectué moins de mouvements cloniques (de 14 à 4). En revanche, lorsque la phrase était passée à l'envers ou qu'elle était interprétée par un violon, elle n'était pas reconnue. Un an après la naissance, Feijoo (1983) a testé 7 des 15 bébés qui avaient participé à l'expérimentation initiale. Sur les 7, 6 enfants se sont calmés à l'écoute de la phrase de Pierre et le Loup.

Les travaux de Lamont³ à l'Université de Leicester confirment que cette mémorisation prénatale est réactivable près d'une année après l'exposition musicale. Ainsi, **des bébés âgés de un an sont capables de reconnaître la musique à laquelle ils avaient été exposés pendant le dernier trimestre précédant leur naissance, et ce, quel que soit le type de musique** (musique classique, musique du monde, musique pop ou musique reggae) :

« After the babies were a year old, they heard the prenatal music and other music that was matched for style, key, pace and loudness. For example, a baby who was exposed to UB40's Many River to Cross before birth heard this piece with another slow reggae track, Freddie McGregor's Stop Lovin' You [...] » (Lamont³).

Les sujets contrôles par comparaison ne manifestent aucune préférence particulière pour l'une ou l'autre musique. **Le tempo de la musique semble jouer un rôle important** puisque, selon l'auteur, **les bébés qui ont été habitués à des musiques aux rythmes rapides montrent des préférences plus marquées que les bébés qui ont été habitués à des musiques « lentes » :**

« the style of the music is not important – the babies recognize UB40 just as much as they do Mozart. But the pace of the music seems to be influential --the babies with faster music like Five's If Y a Getting' Down or the start of Vivaldi's Four Seasons show stronger preferences than the babies with slower music like Mozart's Adagio for Wind » (Lamont³).

³ Voir site Web : <http://news.bbc.co.uk/1/hi/health/1432495.stm>.

Cette « mémoire musicale » expliquerait le sentiment de familiarité qu'éprouvent certains musiciens professionnels, eux-mêmes fils de musiciens à l'écoute d'une pièce spécifique et qui

« déclarent avoir éprouvé un sentiment de familiarité en entendant ou en commençant une pièce nouvelle... qui s'avérait – après enquête – avoir été maintes fois jouée par leur mère pendant qu'elle les portait » (Lecanuet, 1995 : 37).

Lecanuet souligne que

« quoi qu'il en soit, on peut dire que l'expérience musicale prénatale contribue – de la même façon, et certains souhaiteront dire « plus que » n'importe quelle exposition à une stimulation acoustique structurée – à modeler les capacités auditives, à développer des préférences à long terme ou une sensibilité générale au type de sons rencontrés in utero » (Lecanuet, 1995 : 37).

2.1.4. La culture prénatale du fœtus

De Casper et al. (1994) ont montré que la reconnaissance prénatale pouvait intervenir de manière très précoce, entre la 30 et la 37^{ème} semaine de gestation. La reconnaissance d'un extrait sonore implique au préalable la mémorisation de cet extrait. Le fœtus fait donc preuve de capacités très précoces de mémorisation et d'apprentissage. L'imprégnation de l'enfant avec son environnement commence quatre mois avant sa naissance et il y a une continuité entre la vie fœtale et la vie post-natale. Pour Busnel (1994), la période de gestation constitue pour le fœtus une véritable « culture prénatale » pendant laquelle il va apprendre et se familiariser avec son futur environnement sonore et pour ce qui nous intéresse, avec sa LM et son SMM. La durée d'exposition au stimulus est fondamentale dans la formation de cette « culture ». Le nouveau-né semble particulièrement sensible aux caractéristiques prosodiques de sa LM et à la musique, ce qui n'est finalement guère étonnant, puisque cette information lui parvient de manière privilégiée pendant la fin de la vie fœtale. Mais, cette sensibilité aux éléments musicaux se prolonge-t-elle après la naissance ?

2.2. Sensibilité du bébé aux éléments musicaux : le « trésor inépuisable » des ressources enfantines

« Il semble que le nourrisson se fonde sur l'information signalant les unités syntaxiques importantes pour décomposer les stimulations acoustiques verbales et musicales. L'environnement acoustique intra-utérin et l'expérience post-natale du discours adressé au bébé fournissent les contextes d'apprentissage des indices nécessaires à la segmentation et à la structuration du langage et de la musique. Mais, le nourrisson est aussi porteur d'importantes prédispositions à percevoir et à acquérir ces indices. Dès les tous premiers jours, il est extrêmement sensible aux variations mélodiques et temporelles du langage et de la musique. (Fassbender, 1995 : 98).

2.2.1. Sensibilité au rythme

Vers l'âge de 2 à 3 mois, **le bébé est capable de différencier des séquences ayant des structures rythmiques différentes** (Demany et al., 1977). Cependant, la perception du rythme chez les adultes est plus élaborée, puisqu'ils sont capables, à partir d'une séquence, d'abstraire une structure rythmique qui restera identique malgré les variations de tempo. Est-ce que les bébés sont capables de la même capacité de généralisation ? Trehub et Thorpe (1989) montrent qu'ils sont capables de différencier des séquences de 3 sons (X XX ou XX X) et de 4 sons (XX XX et XXX X), et ce, en dépit des variations de tempo et des variations de hauteur. Les résultats montrent cependant que cette tâche leur est difficile car le niveau de performance, bien que significatif, reste faible. Néanmoins, ils suggèrent que **chez le nourrisson, le traitement est relationnel au niveau temporel**. Cette capacité est centrale dans la perception de la parole, car elle permet à l'auditeur d'extraire le même contenu verbal en dépit des variations de débit. **Cette « normalisation » fonctionne également dans la perception des séquences musicales**. Cependant, ce type de traitement n'est pas spécifique à l'être humain, puisque les étourneaux (Hulse et al., 1984) et les pigeons (Miller et Liberman, 1979) sont également capables de normaliser des séquences rythmiques.

2.2.2 Sensibilité aux contours mélodiques

2.2.2.1. Les contours mélodiques de la musique

Nous l'avons vu dans la partie n° 1.2.2., le type de traitement appliqué à la mélodie par le système perceptif dépend fortement de l'expérience que le sujet a contractée au contact de son milieu sonore, qu'il soit langagier ou musical (traitement par contours ou par intervalles). Qu'en est-il des capacités du bébé à traiter les mélodies ? A quel type de stratégie a-t-il

recours ? Pour le savoir, Trehub et al. (1984) ont étudié la capacité de bébés de 6 à 7 mois et d'adultes à détecter des mélodies modifiées par rapport à une mélodie standard composée de 6 sons. Les trois variantes débutent et se terminent sur les mêmes notes, mais la première est une transposition de la mélodie standard dans une autre tonalité, la seconde comprend des changements de notes qui préservent le contour et la dernière, des changements qui ne préservent pas le contour (cf. figure n°47).

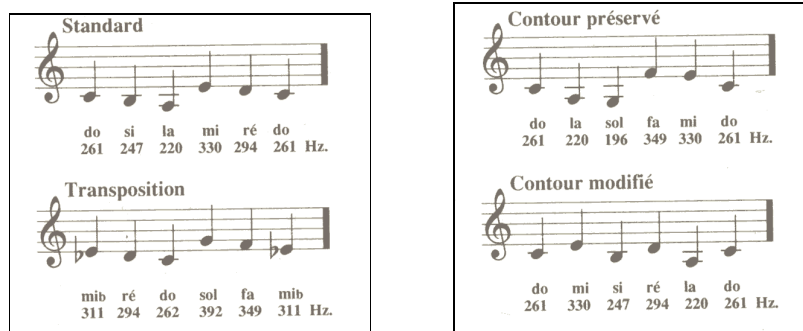


Figure n°47 : une mélodie standard, suivie des trois types de changement à détecter : transposition, changement préservant le contour et changement modifiant le contour (emprunté à Trehub et Trainor, 1994 : 313).

Comme les auditeurs adultes, les bébés se montrent aptes à détecter les changements affectant le contour, mais échouent à détecter les transpositions de la mélodie initiale et les changements préservant le contour. Ces deux types de modification passent inaperçus, comme si c'était la mélodie standard qui était reprise. Leur stratégie est donc similaire à celle des adultes pour le traitement des mélodies non familières. Ils traitent les contours de manière globale, en ignorant les détails de ces contours. Ces résultats montrent que **les bébés utilisent une stratégie de traitement global où l'information sur le contour domine celle qui concerne la hauteur absolue et les intervalles**. Pour Trehub et Trainor,

« ces résultats suggèrent que le traitement de contour de hauteurs fonctionne peut-être chez les nourrissons comme un important mécanisme d'organisation perceptif qui dirige leur segmentation des patterns auditifs complexes » (Trehub et Trainor, 1994 : 317).

2.2.2.2. Les contours d'intonation de la langue

Dès l'âge de 2 à 3 mois, **les bébés sont capables de discriminer des syllabes au contenu segmental identique, mais affectées de contours d'intonation différents**, montants et descendants (Morse, 1972 ; Kuhl et Miller, 1982). Menyuk (1994) obtient des résultats similaires avec des bébés plus jeunes, âgés de 6 à 7 semaines. Best (1993) montre que **les bébés anglophones âgés de 2 mois discriminent certes les contrastes prosodiques de l'anglais, mais seulement lorsque la parole est adressée aux enfants (LAE)**. Ils sont donc capables de différencier les contours prototypiques de leur LM et de répondre de manière adéquate aux messages qu'ils véhiculent, tels que l'incitation au réveil, à l'attention et au contact visuel (Papousek et al., 1990).

Plus tard, ces capacités s'étendent à d'autres langues : **les bébés de 6 à 8 mois sont capables de discriminer les contrastes prosodiques affectant la parole adressée à l'enfant dans deux langues différentes**, leur LM, l'anglais et une langue étrangère, l'espagnol (Best, 1993). Cependant, **vers 10-12 mois, cette capacité semble s'altérer**, puisque les bébés anglophones ne se montrent plus capables de discriminer les contrastes prosodiques de l'anglais, ni ceux de l'espagnol (Best, 1993). Ces résultats pourraient s'expliquer par le fait que le LAE utilisé dans l'expérimentation ne correspond pas au niveau des bébés de 10 à 12 mois, mais à celui de bébés plus jeunes. En effet, les caractéristiques du LAE évoluent avec l'âge et les capacités de l'enfant (Snow, 1977 ; voir 2.4.2.2.), et il eut peut-être fallu utiliser un autre type de LAE. Une autre étude confirme cette hypothèse, puisque, lorsque le LAE utilisé est adressé à un bébé de 10-12 mois, les bébés du même âge sont capables de discriminer les contours marquant l'affirmation et la question en anglais et en mandarin. Contrairement au niveau segmental (voir 3.1.), **il ne semble pas se produire de réorganisation perceptive du système prosodique de la LM au niveau des intonèmes**. Ces effets commenceraient à apparaître plus tard, vers 18 mois, au moment où l'enfant utilise la syntaxe de manière plus approfondie (au niveau de la phrase).

2.2.3. Sensibilité à la prosodie affectant les mots

Un certain nombre d'études ont mis en relief une sensibilité particulière des bébés à la prosodie affectant le mot, comme l'accentuation lexicale en anglais par exemple. Une étude récente a montré que des nouveau-nés étaient **capables de discriminer des stimuli multisyllabiques différant uniquement par la place de l'accent**, correspondant à des mots italiens (Sansavini et al., 1997). Cette capacité avait déjà été démontrée chez des bébés anglophones de 1 à 2 mois (Spring et Dale, 1977). Dès le premier mois, les bébés se montrent sensibles aux mots fortement accentués et à la terminaison des phrases (Menyuk, 1977). A 2 et 3 mois, ils ne font d'ailleurs **pas attention aux mots prononcés sans exagération de la prosodie** (Kagan et Lewis, 1965 ; Eimas, 1975). Cette sensibilité à la prosodie des mots va progressivement évoluer vers un traitement plus phonétique, comme le montre une étude de Jusczyk et al. (1993a). A 6 mois, les bébés américains sont capables de distinguer des mots anglais de mots néerlandais, et ce, même avec un filtrage passe-bas, mais uniquement sur la base de l'information prosodique. Par contre, à 9 mois, ils ont recours également aux différences phonétiques entre les mots. A 9 mois, les bébés anglophones marquent une préférence pour les mots anglais accentués sur la syllabe initiale (Jusczyk et al., 1993b). Ce type d'accentuation représente le pattern accentuel dominant dans cette langue (voir chapitre V, 2.1.1.2.). Parallèlement à cette sensibilité prosodique, la compréhension du sens des mots se développe, puisque les bébés sont capables vers 10-12 mois de comprendre une trentaine de mots en contexte. Et au même moment, apparaissent les premiers mots au sein du babillage mixte. Selon Bertoncini et De Boysson-Bardies (2000 : 104),

« au cours du développement perceptif, la prosodie peut aider à indexer différents niveaux d'unités fonctionnelles de la langue. En association avec les composants syllabiques, elle fournirait le format des mots, en offrant les premiers indices pour la segmentation et le groupement des syllabes en unités-mots ».

C'est donc la prosodie qui permet à l'enfant de se sensibiliser aux mots de sa LM, puis d'accéder progressivement à l'information segmentale. Elle joue le rôle d'un guide qui l'incite à porter son attention sur l'information importante.

2.2.4. Capacités communes de segmentation pour la LM et le SMM

Dans le premier chapitre, nous avons évoqué les préférences du bébé pour les extraits de parole segmentés correctement : ils écoutent plus longtemps les extraits interrompus aux limites des phrases, plutôt qu'à l'intérieur des phrases. Cette capacité a été relevée dès l'âge de 4 mois et demi (Jusczyk, 1989) et elle a également été attestée à 6 mois (Hirsh-Pasek et al., 1987 ; Kelmer Nelson et al., 1989 ; Nazzi et al., 2000). Dans une étude réalisée par Stern et al. (1983), des bébés de 7 à 10 mois marquent une **préférence pour le langage adressé à une fillette de 19 mois, lorsque les pauses sont insérées aux frontières des propositions**. Il est très important de souligner que cette préférence ne fonctionne que pour le LAE, car dans une autre étude concernant des bébés du même âge, les auteurs ne relèvent aucune préférence si le matériel utilisé concerne le langage adressé à l'adulte (Kelmer Nelson et al., 1989). Le LAE semble donc posséder des propriétés apparemment très attractives pour le petit enfant, et conserve ce pouvoir pendant un temps relativement long. Une autre étude réalisée par Jusczyk et al. (1992) montre que la préférence pour des extraits segmentés aux frontières linguistiques persiste chez les bébés de 9 mois et ce, même si les extraits sont filtrés pour ne laisser passer que l'information prosodique (à l'aide d'un filtre passe-bas).

De manière troublante, cette capacité de segmentation est similaire pour la musique et apparaît également de manière très précoce : **dès 4 mois et demi, les bébés préfèrent des extraits musicaux** tirés des menuets de Mozart, **segmentés aux limites des phrases musicales**, plutôt qu'à l'intérieur de ces phrases (Jusczyk et Krumhansl, 1993). Si l'extrait musical est inversé, la complexité du contour est conservée, ainsi que les hauteurs, mais la direction du contour et la durée des notes sont modifiées. Les bébés préfèrent la segmentation naturelle. Les auteurs en concluent donc que les changements de contour, et notamment son abaissement, et l'allongement des sons précédant les frontières sont des éléments d'importance pour diriger la segmentation chez le bébé.

2.3. L'amorçage de l'attention par les éléments musicaux

2.3.1. Richesse et redondance du signal

Nous avons vu dans la partie 1.2.2. que les bébés manifestaient des capacités générales de structuration des ensembles acoustiques. Ces capacités leur permettent de repérer des régularités structurelles à l'intérieur de la variabilité qui caractérise le flux sonore.

Or, le signal comporte lui-même un grand nombre d'informations qui vont permettre à l'enfant de guider son attention sur des événements importants. La langue et la musique sont organisées hiérarchiquement (chapitre I, partie n°3.2.) et les signaux saillants dans le signal correspondent à des frontières linguistiques et musicales importantes, qui délimitent notamment les différents constituants (voir paragraphe précédent, 2.2.4.). Dans le flux sonore, ces signaux constituent donc des éléments proéminents sur lesquels l'attention de l'enfant va pouvoir se fixer, un peu comme des prises sur un mur d'escalade, avant que son système perceptif procède à un traitement plus en profondeur. Et c'est justement grâce à la prosodie que ces signaux sont mis en proéminence. En effet, **la prosodie contient de nombreux indices sur les frontières linguistiques importantes** (Morgan et al., 1987). D'après la théorie dite du « bootstrapping prosodique », appelée également « bootstrapping du signal » (Gleitman et Wanner, 1982 ; Peters, 1983 ; Gleitman et al., 1988), **la prosodie guiderait ainsi l'attention sur les unités des différents niveaux d'organisation** (phrases, propositions, différents syntagmes, mots), ce qui permettrait à l'enfant d'élaborer son savoir sur l'organisation syntaxique de sa LM :

« an infant who is intentionally circumsuiped utterance segment as potential syntactic constituent would be at considerable advantage in learning the syntactic rules of his language »
(Read et Schreiber, 1982 : 26).

Cette hypothèse suppose que l'enfant ait des capacités perceptives suffisantes pour détecter les signaux prosodiques, mais les travaux évoqués dans les parties précédentes prouvent que la sensibilité du bébé à ce type de signaux est très forte. Elle suppose également que le signal de parole contienne des informations et plus celles-ci seront riches et nombreuses, moins les arguments en faveur de capacités « programmées » pour le traitement de la langue seront justifiés. Quels sont ces signaux qui orientent l'attention du bébé à l'intérieur du flux sonore ?

2.3.2. Les signaux d'amorçage dans la langue

Gleitman et Wanner (1982) pensent qu'il s'agit des **syllabes accentuées**, car pour les petits anglophones, l'apprentissage des syllabes atones est beaucoup plus long. Pye (1983) a découvert que les enfant Quiché Maya apprennent les inflexions accentuées avant les racines des verbes, qui sont inaccentuées. Mais, les syllabes accentuées ne se présentent pas acoustiquement de la même manière dans les différentes langues. En français par exemple, les syllabes atones ne subissent pas de réduction, contrairement à l'anglais (cf. chapitre V, différences d'accentuation entre le français et l'anglais).

Les signaux qui vont guider l'attention de l'enfant vont donc varier selon les langues et les spécificités de leur organisation prosodique. Ainsi, en anglais, la **frontière entre les propositions est marquée par une pause, une augmentation ou diminution de la fréquence fondamentale ou un allongement de la syllabe précédant la frontière de la proposition** (Klatt et Cooper, 1975 ; Luce et Charles-Luce, 1983 ; Kakatani et Dukes, 1977). Cooper et Paccia-Cooper (1980) ont relevé des **pauses plus longues aux frontières des propositions**, tandis que **les groupes intonatifs ont tendance à correspondre aux unités syntaxiques majeures telles que les propositions** (Bolinger, 1978 ; Cruttenden, 1986 ; Selkirk, 1984). Bernstein Ratner (1986) a relevé des allongements plus prononcés au niveau segmental aux frontières des propositions. Cependant, Fisher et Tokura (1996) n'ont pas identifié de marqueurs prosodiques plus petits que ceux relevés pour la proposition. Enfin, Jusczyk et al. (1992) ont remarqué une corrélation entre hauteur et durée dans la frontière entre le sujet et le prédicat. L'information sur les unités semble non pas portée par un seul type de signal, mais par une véritable constellation de signaux (intégration de signaux multiples mis en jeu dans l'apprentissage des catégories syntaxiques).

2.3.3. Une théorie de l'amorçage valable pour la musique ?

La théorie du bootstrapping a été élaborée pour la langue, mais les similitudes de segmentation constatées pour la langue et la musique (voir 2.2.4.) et la sensibilité de l'enfant aux éléments musicaux en général suggèrent la possibilité de l'étendre au traitement de la musique. On remarque en effet une forte analogie entre les signaux utilisés pour marquer la structure de la langue et celle de la musique : la chute de la hauteur avant les frontières musicales est similaire à celle que l'on retrouve à la fin des énoncés dans la parole. De la même manière, l'allongement de la durée des notes précédant les frontières musicales (Jusczyk et Krumhansl, 1993) se retrouve dans la parole (Scott, 1982). Cette analogie suggère des mécanismes communs de traitement de la langue et de la musique.

2.4. Ajustements vocaux inconscients adaptés à la sensibilité « musicale » du bébé

2.4.1. Préférences de l'enfant pour le langage et le chant adressé

2.4.1.1. Le langage adressé

Lorsqu'un adulte s'adresse à un enfant, et en particulier, la mère à son enfant, il adopte une façon de parler particulière, extrêmement simplifiée au niveau de la syntaxe, du sens et du répertoire phonologique : le langage adressé à l'enfant, également appelé « mamanais » (motherese ou baby talk en anglais). Or, selon la théorie de l'amorçage prosodique, la segmentation est opérée sur la base de l'information prosodique. Le LAE se caractérise justement par une remarquable musicalité (Papousek et Papousek, 1981 ; Fernald, 1989), qui s'exprime par l'intermédiaire de paramètres prosodiques fortement exagérés. Selon Papousek et Papousek (1981 : 171), le langage adressé constitue l'expérience musicale la plus précoce de l'enfant. Papousek ajoute dans un autre article (1995 : 104), que :

« les parents et l'entourage du nourrisson font preuve d'une propension intuitive à parler au nouveau-né et à lui prodiguer ce que l'on peut assimiler aux premières formes de l'éducation musicale ».

Nous savons que les fœtus préfère déjà le LAE dans le ventre de leur mère. Cette préférence persiste pendant la première année de vie. Les bébés de 2 jours préfèrent ce type de discours (Fernald, 1985 ; Cooper et Aslin, 1990), même si les contours sont synthétisés en ondes sinusoïdales (Fernald et Kuhl, 1987). Fernald (1985) a testé 48 bébés de 4 mois ; 33 tournent plus souvent la tête en direction du haut parleur qui diffuse du LAE. Papousek et Papousek (1981 : 173) soulignent les manifestations de joie et de plaisir suscitées chez le bébé à l'écoute du LAE. Fernald et Kuhl (1981) ont isolé les paramètres acoustiques qui permettent de faire la différence entre le LAE et la parole adressée à un adulte. Ils ont fait ainsi apparaître que la préférence pour le LAE était conservée uniquement pour la Fo, paramètre probablement hautement saillant pour les bébés, plus saillant que le rythme, les formants et l'amplitude (Fernald, 1984 ; Fernald et Kuhl, 1987 ; Katz et al., 1996). Ils préfèrent d'ailleurs un discours où la Fo varie (Fernald, 1985 ; Fernald et Kuhl, 1987), et les contours dont la Fo monte rapidement (Fernald, 1989).

A 8 mois ½, les bébés continuent à préférer le LAE, alors que les adultes ne manifestent aucune préférence pour ce type de langage (Kelmer Nelson et al., 1989). Cette préférence se prolonge jusqu'à 19 mois, âge où la progression de l'enfant dans l'acquisition de sa LM ne semble plus dépendre de ce style de parole.

2.4.1.2. Le chant adressé

De la même manière que pour le langage, les adultes adaptent leur façon de chanter aux petits enfants. Lorsqu'une mère chante à son enfant (« chants de nourrice » ou « nursery songs »), elle le fait en effet d'une manière particulière, très différente du chant « adulte ». Par analogie avec le LAE, on parle de chant adressé à l'enfant (désormais CAE). Il existe plusieurs types de chants adressés à l'enfant. Les **berceuses** sont destinées à endormir l'enfant et pour arriver à ce résultat, les contours descendants y sont prédominants. Elles peuvent se répéter infiniment jusqu'à ce que le bébé s'endorme. Un autre type de chant possède une fonction inverse à celle des berceuses : les **chansons de jeux** qui sollicitent l'activité motrice de l'enfant. Pour parvenir à ce résultat, les contours montants y sont majoritairement utilisés. Le chant adressé a beaucoup intéressé les chercheurs qui travaillaient sur le LAE, à cause des similitudes entre les deux types d'adaptation vocales. Le CAE semble en effet jouer un rôle aussi important dans l'acquisition du SMM, que le LAE pour la LM. L'étude de leurs similitudes pourrait contribuer à mettre à jour des processus généraux mis en œuvre dans l'acquisition valables pour la langue et pour la musique. Comme ils préfèrent le LAE, les enfants marquent une préférence pour le chant adressé à l'enfant (CAE) plutôt que pour le chant non adressé (Trainor, 1996) et notamment, ils préfèrent les berceuses aux chansons de jeux et aux chants d'adultes, ainsi que le chant de la mère plutôt que celui du père (Trehub et Henderson, 1994). Là aussi, la hauteur paraît jouer un rôle fondamental dans cette préférence : entre deux versions de la même chanson, les enfants préfèrent la version chantée avec une hauteur plus élevée (Trainor et Zacharias, 1999).

2.4.2. Caractéristiques du langage adressé à l'enfant

Les caractéristiques acoustiques du LAE ont été principalement décrites par Papousek et Papousek (1981), Stern et al. (1983) et Fernald (1989). Dans le LAE, la **hauteur de la voix** est **plus élevée** de 3 à 4 demi-tons. La mère parle comme si sa cavité vocale était de la même taille que celle de son bébé, avec une voix palatale.

On observe une grande **continuité dans l'évolution de la mélodie**, comparé à sa variabilité dans la parole adulte (Fernald et Simon, 1984) : il y a **beaucoup de glissandi** et peu de changements brusques. Les **contours sont très exagérés, mais leur configuration reste simple et unidirectionnelle** (de type haut-bas ou bas-haut) et ils sont **constamment répétés par la mère**. Cette continuité et cette simplicité les rendent particulièrement faciles à traiter par l'enfant. Ils lui permettent notamment de **localiser facilement la source**, c'est-à-dire la personne qui lui parle et de maintenir son attention fixée sur cette source :

« The exaggerated intonation of mothers' speech, with its greatly expanded pitch range and high degree of pitch continuity, thus maximizes both perceptual contrast, necessary for engaging and maintaining infant attention, and perceptual coherence, facilitating the task of following the voice of a single speaker » (Fernald, 1984 : 13).

Ces contours sont composés de « bons intervalles », c'est-à-dire **d'intervalles aux rapports fréquentiels simples**, et donc facilement décodables par le système perceptif (Burns et Ward, 1982 ; Rakowski, 1990) : les tierces, les quarts, les quintes et les octaves (Fernald, 1976) y sont prédominants (pour les non musiciens, se reporter à la grille d'intervalles). Le **degré de mélodicité du LAE est élevé**, ce qui lui confère une grande musicalité (voir chapitre 1, partie n°1.2.4.). Cinq ou six contours prototypiques reviennent régulièrement dans le LAE. Les contours montants servent à attirer l'attention du bébé, tandis que les contours en cloche, dont la tessiture est plus large, servent à maintenir son attention (voir figure n°48). Ils sont parfois utilisés dans des questions, même si les phrases ne sont pas interrogatives (Ferguson, 1977). La tessiture tend à être plus étroite dans les contours descendants, de manière à calmer l'enfant (Papousek et Papousek, 1981).

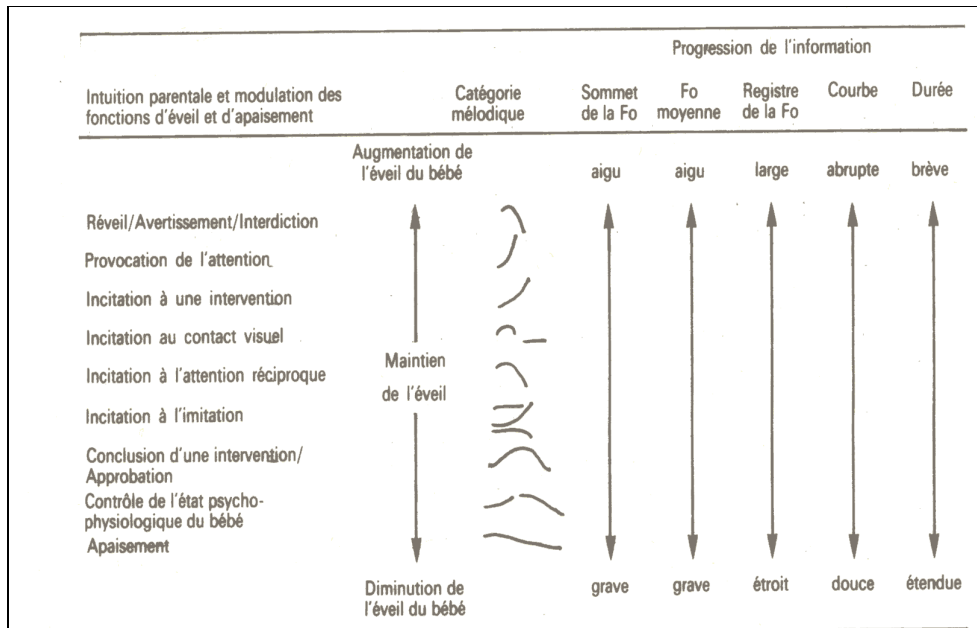


Figure n°48 : les mélodies du LAE (emprunté à Papousek, 1995 : 112)

Ces contours correspondent de façon remarquable aux principes de base de la perception de la forme, en particulier aux lois de proximité et de similitude (Fernald, 1984) et à la loi de destin commun (Fassbender, 1995). Dans le LAE, la plupart des énoncés sont courts, dépourvus de signification et incomplets d'un point de vue grammatical, mais ils tendent à être complets du point de vue intonatif (Broen, 1972). En ce qui concerne les autres paramètres prosodiques, les énoncés sont produits avec plus de régularité et plus de rythmicité. Le débit s'adapte à l'état du bébé : s'il est attentif, le débit sera plus lent et les énoncés plus rythmiques ; en revanche, le débit s'accélère lorsque le bébé n'est pas attentif. Il ralentit lorsque l'enfant est en train de s'endormir (Papousek et Papousek, 1981). Non seulement, on constate plus de pauses entre les différents segments, mais de plus, ces pauses sont prolongées. Enfin, le LAE se caractérise par de grands changements d'amplitude.

Le niveau segmental est complètement déformé au profit des exagérations produites au niveau de la prosodie. Les mots sont transformés et remplacés par des diminutifs, comme « *doggie* » pour « *dog* » en anglais et par des reduplications (« *chachat* » pour « *chat* » en français). Tout se passe comme si le niveau segmental était délaissé, car secondaire au début de l'acquisition du langage. En revanche, les voyelles sont prolongées comme elles le seraient dans une chanson. A l'intérieur de la syllabe, c'est la voyelle qui porte le contour mélodique. Or, dans le LAE, l'exagération des contours contribue à mettre les voyelles en relief. Par ailleurs, le grand nombre de monosyllabes de type V ou CV suggère une importance de ces sons dans l'acquisition (voir ce chapitre, partie n° 3.1.2.1.).

Kuhl et al. (1997) ont étudié les voyelles du langage adressé à des bébés américains, russes et suédois âgés de 2 à 5 mois. Il est apparu que les mères utilisaient la prosodie pour attirer l'attention de leur bébé sur les trois voyelles [i] [a] et [u] et consacrent du temps à obtenir une réponse. Hormis les travaux ponctuels de Robin, nous ne disposons pas encore de description complète du LAE pour le français.

2.4.2.1. Une pédagogie « inconsciente »...

L'ensemble de ces ajustements vocaux censé s'adapter aux capacités de traitement de l'enfant dans un but d'acquisition peut être qualifié de « pédagogie » à la différence près, qu'une pédagogie est un processus conscient alors que ces adaptations sont totalement inconscientes de la part des parents. Ces ajustements se font en effet de manière intuitive et naturelle et ils sont très rarement utilisés en dehors du contexte d'interaction avec un bébé (Fernald et Simon, 1984). Ce type de langage est également utilisé par des enfants âgés de 2 à 4 ans lorsqu'ils s'adressent à des enfants plus jeunes ou à des objets ersatz d'animés, tels que poupées ou peluches. Dans un contexte différent, on retrouve des ajustements vocaux similaires dans ce que Giacomi et Hérédia (1986) appellent la « coopération exolingue ». Cette coopération se met en place entre l'enseignant de langue et l'apprenant étranger. Comme la mère lorsqu'elle parle à son bébé, l'enseignant exagère l'intonation et marque le rythme lorsqu'il s'adresse à l'apprenant, sans doute pour faciliter sa compréhension et son apprentissage. Dans les deux cas, le locuteur adapte sa manière de parler aux capacités de son interlocuteur pour le « tirer » progressivement vers un savoir plus complexe. C'est donc un procédé pédagogique par excellence. En ce qui concerne le LAE, les adultes sont tout-à-fait capables de reconnaître ce type de langage d'après ses caractéristiques prosodiques amplifiées ; ils sont même capables de différencier un LAE produit en présence de l'enfant, d'un LAE produit en l'absence de l'enfant (langage adressé à une poupée par exemple). Et pourtant, lorsqu'à leur tour, ils produisent ce type de discours, ils n'ont pas conscience de produire ces ajustements. La mère en particulier, rencontre des difficultés à les reproduire en l'absence de l'enfant (Fernald, 1976) : « *more than a melody. It's a real tutorial and they didn't even know they were doing it* » (Kuhl et al., 1997).

2.4.2.2. ...qui évolue avec l'âge de l'enfant

Les caractéristiques du LAE ne restent pas statiques et la mère adapte son parler à mesure que l'enfant grandit : elle insère de plus en plus de mots isolés, de pauses et elle interrompt ses phrases beaucoup plus souvent (Broen, 1972). Ces caractéristiques se renforcent jusqu'à ce que le bébé atteigne l'âge de 4 mois, pour ensuite s'atténuer. C'est à ce moment justement que la dyade mère-enfant connaît une période de face à face très intense. La mère utilise les contours de hauteur montants pour attirer l'attention de son bébé lorsqu'il regarde ailleurs, puis utilise des contours en « cloche » pour maintenir le contact oculaire avec lui (Stern et al., 1982), et ce, jusque vers 6 mois. Plus tard, lorsque le bébé atteint 14 mois, elle accentue le prédicat, c'est-à-dire l'information nouvelle véhiculée par l'énoncé, grâce à une hauteur plus élevée et une plus grande intensité. Elle place les mots-clés en position finale (Fernald et Mazzie, 1991), de manière à ce que les enfants les repèrent plus facilement. Au même moment, les capacités de compréhension du bébé concernant les mots de sa LM se développent rapidement. Depuis l'âge de 8 mois, il est en effet capable de reconnaître les mots en dehors du contexte et à 11 mois, son attention se fixe plus longtemps sur les mots familiers, même s'ils sont altérés au niveau phonologique. Enfin, c'est vers 12 mois qu'apparaissent les premiers mots isolés, au sein du babillage mixte. Selon Kuhl et al. (1997), les voyelles et les consonnes sont prononcées par la mère de manière plus distincte lorsque l'enfant commence à parler (hyperarticulation), que lorsqu'il est plus jeune ou au contraire plus âgé (Malsheen, 1990 ; Ratner, 1984). Le parler « bébé » s'adapte donc intuitivement aux capacités perceptives du bébé et doit évoluer en fonction du développement :

« to be effective, the communicative signals of the mother must be well matched to the perceptual capabilities and limitations of the young » (Fernald, 1989 : 10).

Il se crée donc un jeu de miroir entre la mère et son enfant : la mère rend accessible à son enfant des informations sur les unités linguistiques importantes dans le signal de parole et continue à le faire jusqu'au moment où elle « ressent » que le bébé est prêt à passer à une étape plus difficile dans l'acquisition de sa langue.

2.4.3. Caractéristiques du chant adressé à l'enfant

Lorsque la mère chante à son enfant, elle le fait d'une manière beaucoup plus expressive (Trehub et al., 1997). Le ton est plus aimant (Trainor, 1996) et la qualité de la voix est émotionnellement plus engageante (Trehub et al., 1997). Cette différence d'interprétation est à ce point importante que des auditeurs adultes sont capables d'identifier le chant produit en présence de l'enfant de celui produit en son absence, tout comme ils étaient capables de le faire pour le LAE. Ainsi, dans une étude réalisée par Trehub et al. (1993a), des adultes américains sont capables d'identifier les chansons américaines interprétées en présence de l'enfant, avec un taux de réussite de 90 %. Ils sont également capables de la même identification pour une culture musicale et une langue différente, l'hindi, avec certes un taux de réussite moins élevé, mais toutefois supérieur au hasard (Trehub et al., 1993b). Même si les parents tentent de simuler la présence de l'enfant, la version produite reste identifiable par les auditeurs, même dans une culture étrangère (Trehub et al., 1997). Enfin, le CAE semble s'adapter au développement perceptif de l'enfant puisqu'une comparaison entre des chansons adressées à des bébés et des chansons adressées à des enfants d'âge préscolaire relève que les chansons pour les pré-scolaires sont articulées plus clairement (Bergeson et Trehub, 1999).

2.4.3.1. Les berceuses

La hauteur des berceuses est plus élevée que celle des autres chansons (Trainor et al., 1997). Bergeson et Trehub (1999) ont réalisé une étude acoustique comparative de chants adressés à des bébés et des chants adressés à des enfants d'âge préscolaire. Ils n'ont relevé aucune différence significative entre les deux types de chant concernant le tempo, la durée, les pauses entre les phrases et la durée des voyelles. En revanche, la différence de hauteur était très significative, les chants adressés aux bébés étant plus hauts de 3 à 5 demi-tons, avec une différence de hauteur modale d'un demi-ton. Ce registre semble s'accorder à la sensibilité des bébés qui préfèrent les chants dont le registre est plus élevé (Trainor et Zacharias, 1999). En revanche, la tessiture est étroite, la tonalité douce et les contours répétitifs. La principale fonction des berceuses étant d'endormir l'enfant, elles se caractérisent par une prédominance de contours descendants, un tempo lent et un rythme répétitif. Par ailleurs, les berceuses sont chantées pianissimo. Au niveau segmental, la mère privilégie l'effet sonore sur la signification et sa prononciation est relativement indistincte. Cependant, les voyelles sont prolongées. La mère redouble des mots et répète des séquences de mots familiers.

Est-ce que les adultes sont capables de reconnaître les berceuses d'autres chansons ayant un tempo et un style musical similaire ? Trehub et al. (1992) ont apparié 30 berceuses provenant de différentes cultures à 30 chansons équivalentes. Les adultes ont été capables de reconnaître les berceuses avec un taux de réussite de 63 %, et ce, même après l'adjonction d'un filtre passe-bas avec une fréquence de coupure de 500 Hz et même après la resynthèse de la ligne mélodique. D'après Unyk et al. (1992), la simplicité des contours et leur configuration majoritairement descendante constitueraient les caractéristiques responsables de leur identification.

2.4.3.2. Les chants de jeux

Les chants de jeux ont une autre fonction que les berceuses, puisqu'ils doivent susciter l'activité motrice du bébé. Pour cette raison, ils sont composés en majorité de contours montants. Par rapport au chant non adressé à l'enfant, la hauteur des chants de jeux est plus variable. Les syllabes accentuées sont plus longues et plus intenses et on constate plus d'énergie dans les basses fréquences. En ce qui concerne le rythme, le tempo est plus lent et les pauses plus longues (Trainor et al., 1997 ; Trehub et Trainor, 1998). Ces caractéristiques acoustiques permettent à l'auditeur de les identifier comme plus brillantes, plus rythmées et plus staccato (interprétées en détachant les notes les unes des autres), avec des consonnes plus proéminentes et chantées avec un ton de voix souriant (Trainor et Rock, 1997). Les auditeurs adultes sont capables de faire la différence entre les berceuses et les chants de jeux sans aucune difficulté. Ils reconnaissent en effet l'intention de la chanson avec un taux de réussite de 100 % (Trainor et Rock, 1997). Ces résultats ne sont guère surprenants car les différences formelles sont très grandes entre les deux types de chants. Ces différences ne sont d'ailleurs pas seulement acoustiques, mais également visuelles, car ils sont capables de les différencier seulement à partir de l'image de la mère (alors que le son est coupé), avec un taux supérieur au hasard (60 %).

2.4.4. Points communs entre le LAE et le CAE

A la croisée du langage adressé et du chant adressé, on trouve des caractéristiques acoustiques communes qui pourraient bien se révéler fondamentales dans le processus d'acquisition. Il semble que la hauteur joue le rôle le plus important. Outre leur hauteur plus élevée, le LAE et le CAE se caractérisent par des contours simples et unidirectionnels, continuellement répétés par l'adulte.

Dans les deux cas, ces contours sont utilisés pour les mêmes fonctions : soit solliciter l'activité motrice / l'attention du bébé, avec des contours montants, soit calmer / endormir le bébé, avec des contours descendants. Pour Fernald et al. (1989), ces contours sont universellement liés à des contextes de soins maternels et il est troublant de rencontrer des ajustements vocaux similaires au LAE dans un autre contexte de soin, celui des infirmières avec leur malades. En effet, celles-ci utilisent un registre plus élevé et une tessiture plus large lorsqu'elles s'adressent à leurs patients et ce type de parole est reconnu par les auditeurs (Journiac, 2002). Au niveau du rythme, le débit du LAE et le tempo du CAE sont plus lents. Quant au niveau segmental, il est déformé au profit de l'effet sonore. Enfin, les voyelles sont prolongées et mises en valeur par les contours mélodiques et on rencontre beaucoup de répétition au niveau des mots. Les enfants préfèrent le langage et le chant qui leur sont adressés. Les caractéristiques du LAE et du CAE sont tellement différentes lorsqu'ils sont produits en présence ou en l'absence de l'enfant, que les adultes reconnaissent la situation immédiatement. Enfin, dans les deux cas, les ajustements vocaux s'adaptent à l'âge de l'enfant, évoluant en même temps que lui. Les adultes identifient le chant et la parole adressés même dans une langue et dans une culture qui ne leur sont pas familières. Cette pédagogie « inconsciente » se retrouve en effet dans beaucoup de cultures différentes et de langues différentes (chinois mandarin, japonais, suédois, italien, allemand, anglais, français...). Est-ce que ces adaptations constitueraient des universaux dans le développement de l'être humain ?

3) D'une base commune à une différenciation progressive

3.1. Perception

Nous avons vu dans les parties 2.1. et 2.2. que le bébé manifestait de façon très précoce une grande sensibilité au rythme, aux contours mélodiques, à la prosodie affectant les mots de sa LM et enfin, aux indices « musicaux » correspondant aux frontières linguistiques et musicales majeures de sa LM et de son SMM. Ces étonnantes capacités de traitement se limitent-elles aux éléments musicaux ou concernent-elles également le niveau segmental ?

3.1.1. Jusqu'à 6 mois, l'enfant est un « superdiscriminateur »

Les sons de parole sont constitués par des gestes qui sont des actions coordonnées des mouvements du canal vocal. Best (1993) parle à leur sujet de gestes articulatoires (à la suite des travaux du physiologiste russe Berstein). Il est impossible de reproduire le même son exactement de la même manière à deux occasions différentes.

Il subsiste toujours des différences acoustiques entre deux sons parlés de la même catégorie ; ces différences sont dues aux mouvements d'ajustement des différents articulateurs vocaux. Mais elles sont inutiles au locuteur pour identifier les segments phonétiques. Il a en effet besoin de se concentrer sur un type de distinctions efficaces. Son système perceptif va donc recoder le signal de manière à préserver les seules distinctions qui sont utiles pour distinguer les mots, c'est-à-dire les oppositions phonologiques. Ce recodage, très rapide et très efficace, a été baptisé **perception catégorielle** par les chercheurs des Laboratoires Haskins (Liberman et al., 1967). Il consiste à identifier les événements sonores selon leur appartenance à des catégories, mais en ignorant les variations acoustiques au sein d'une même catégorie. En conséquence, l'auditeur ne peut discriminer deux sons appartenant à la même catégorie perceptive ; il ne peut le faire que s'ils appartiennent à deux catégories distinctes.

Or, en matière de discrimination, le nourrisson fait preuve de capacités bien supérieures à celles de l'adulte. Il est en effet capable, non seulement de discriminer les contrastes de sa LM, mais également une grande variété de contrastes appartenant à des langues qui lui sont inconnues, et pour certaines d'entre elles, rares et très éloignées de la LM (langues kikuyu, hindi, ntlakapmx, zoulou...). Il existe d'ailleurs très peu d'études où les enfants ont échoué (pour un récapitulatif des contrastes discriminés, voir Volume II, annexe 1.3.), ce qui a conduit Trehub (1981) à qualifier le bébé de « superdiscriminateur ». Ces études ont porté en majorité sur les consonnes, mais on sait que les bébés sont capables même plus tôt de discriminer les voyelles (Trehub, 1973), même non natives, dès 1 mois (Polka et Werker, 1994). Ces capacités ont stupéfié les chercheurs, à tel point que certains ont pensé qu'elles ne pouvaient résulter que d'une prédisposition de l'enfant au traitement des langues naturelles (Eimas, 1974). En effet, comment expliquer autrement le fait que les nouveau-nés montrent des capacités de discrimination supérieures à celles des adultes ? Par ailleurs, si cette prédisposition pour le traitement de la langue était avérée, on pourrait également supposer l'existence d'un traitement spécialisé pour la musique, ce qui militerait en faveur d'une séparation très précoce du traitement de la langue et de la musique.

Une étude de Kuhl et Miller (1975) est venue contredire cette vision « innéiste » des choses en montrant que les chinchillas perçoivent le voisement des sons de parole comme les êtres humains. Des résultats similaires ont été observés chez des macaques rhésus (Kuhl et Padden, 1982) et des cailles japonaises qui avaient été entraînées au préalable à percevoir le lieu d'articulation (Kluender et al., 1987).

Ce mode de traitement n'est donc pas réservé aux seuls êtres humains, mais partagé par les espèce du groupe des mammifères. Les recherches entreprises vers la fin des années 70 (Cutting et Rosner, 1974 ; Pisoni, 1977) sur les stimuli non-verbaux amènent un autre type de découverte : la perception catégorielle ne serait pas seulement un mode de traitement du mode langagier, mais plutôt un trait général de la manière dont tous les signaux complexes sont traités. Elle a été observée notamment pour la vision des couleurs (Bornstein et al., 1976) et pour la musique (Jusczyk et al., 1980). **La perception catégorielle est donc une capacité générale du système perceptif, valable entre autres pour la langue et la musique et commune à plusieurs espèces, parmi lesquelles figure l'être humain.**

3.1.2. Spécialisation progressive sur la LM aux dépens de sa richesse perceptive

Au contact de sa LM, l'attention de l'enfant va progressivement se fixer sur les représentations qui lui seront utiles pour son acquisition, augmentant sa performance et sa rapidité de traitement. Ce processus est nécessaire à l'acquisition de la LM, car il permet à l'enfant de se spécialiser sur les contrastes natifs à assimiler. Pour décrire ce phénomène de focalisation sur la LM, nous pourrions reprendre le terme d'acculturation, utilisé en psychologie sociale avec le sens de :

« processus d'apprentissage par lequel l'enfant reçoit la culture et l'ethnie ou le milieu auquel il appartient » (Fraise, 1958 : 108-109).

En ce qui concerne la langue, l'acculturation désignerait le processus d'apprentissage par lequel l'enfant « recevrait » sa LM de son environnement. Mais en se spécialisant, l'enfant perd une partie de ses capacités de discrimination initiales. Ainsi, alors qu'il pouvait encore discriminer un grand nombre de contrastes, à partir de 6 mois, on voit sa perception se restreindre progressivement aux seules distinctions utiles dans sa LM. Cette période de transition correspond à une baisse de la sensibilité du bébé aux contrastes non-natifs et à une augmentation simultanée de sa sensibilité aux contrastes natifs.

3.1.2.1. Focalisation sur la structure vocalique de la LM à partir de 6 mois

Les voyelles, des sons musicaux

Vers l'âge de 6 mois, l'enfant connaît une réorganisation perceptive graduelle en faveur des voyelles de sa LM. Au niveau segmental, pourquoi son attention commence-t-elle par se fixer sur les voyelles ?

Il semblerait que ces sons jouent un rôle particulièrement important au moment de l'acquisition. Les voyelles occupent une place privilégiée au sein de la syllabe dont elles forment le noyau. Au niveau perceptif, elles sont facilement repérables puisqu'elles portent l'information prosodique. Elles sont en outre majoritairement composées d'intervalles de quintes justes, intervalles considérés comme prototypiques, c'est-à-dire faciles à traiter par le système perceptif (voir partie n°3.1.3.). Le rapport qui correspond à un intervalle de quinte est un rapport simple (2/3), fortement présent dans les composantes des sons complexes de l'environnement naturel. Selon Best (1993), les propriétés acoustiques des voyelles les rendent plus faciles à traiter que les consonnes. Elles constituent des points de repère faciles à coder et à mémoriser, des **prototypes**. L'idée que certains objets pourraient fonctionner comme des points de référence cognitifs n'est pas nouvelle en psycholinguistique (Rosch, 1975). On pourrait définir le prototype comme une **représentation idéale formée par l'expérience et qui possède des caractéristiques invariantes. Il constitue le meilleur exemplaire ou l'instance centrale d'une catégorie**. L'idée fondamentale, c'est que les catégories ne sont pas composées de membres aux statuts équivalents, mais qu'elles comportent des membres qui sont de meilleurs exemplaires que d'autres :

« La notion de prototype est donc originellement reliée aux individus : le prototype est l'exemplaire qui est reconnu comme étant le meilleur par les sujets. Cette origine psychologique pose directement le problème de la variation individuelle. Puisque étant d'abord la meilleure instance que les individus ont à l'esprit, le prototype pourrait en théorie varier d'un individu à un autre.[...] Le point définitoire nouveau est que le prototype n'est vraiment considéré comme le meilleur exemplaire d'une catégorie que s'il apparaît comme étant celui qui est le plus fréquemment donné comme tel » (Kleiber, 1990 : 48-49).

Les stimuli prototypiques sont ceux qui incluent le maximum de caractéristiques de la catégorie (Grieser et Kuhl, 1989). Le prototype fonctionne donc comme référence à l'assimilation des structures d'événements perceptifs similaires.

Dans sa théorie dite « *native language magnet* » (1992, 1993, 1994), Kuhl pose l'hypothèse que l'espace acoustique initial est divisé en frontières psycho-acoustiques correspondant aux caractéristiques phonémiques de la LM. Kuhl appelle cet espace « carte perceptive ». Nous pouvons en voir trois exemples sur la figure n°49. Celle de gauche correspond à la carte perceptive du suédois, celle du centre, à la carte perceptive de l'anglais et celle de droite, à la carte perceptive du japonais. Les frontières délimitent les différentes catégories constituant le système vocalique de la langue concernée.

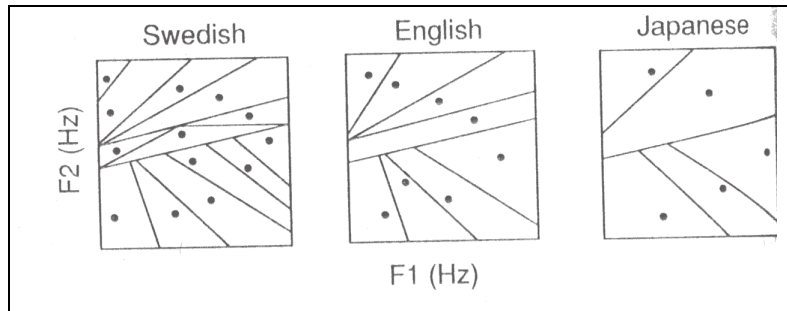


Figure n°49 : cartes perceptives du suédois (à gauche), de l'anglais (au centre) et du japonais (à droite) (emprunté à Kuhl, 1994 : 36).

Ces représentations internes, constamment générées par le cerveau, comparent leur propre structure aux formes qui se présentent dans l'environnement linguistique. L'enfant retient celles qui sont compatibles avec les contrastes, les structures syllabiques et les traits prosodiques de sa LM. Par exemple, le système japonais est composé de cinq voyelles, qui forment cinq catégories (cf. figure n°49). Le bon exemplaire de chaque voyelle (le prototype) se trouve au centre de ces cinq divisions ; sur les cartes, il est représenté par un point. Si un bébé japonais est confronté à une langue étrangère comme le suédois (13 catégories, cf. figure n°49), son système maternel va jouer le rôle d'un filtre. Le prototype fonctionne en effet comme un aimant pour les autres stimuli appartenant à la même catégorie. Lorsqu'un certain nombre de sons se trouvent autour d'un prototype dans une catégorie (figure n°50, au dessus), ils sont perceptivement attirés par ce prototype (« *perceptual magnet effect* », figure n°50, en bas). Les études sur les adultes suggèrent que ces limites ne disparaissent jamais réellement, même si on peut augmenter la performance de discrimination des contrastes NN par un apprentissage approprié (Flege 1989). Les catégories natives interfèrent donc avec la capacité à percevoir les distinctions phonétiques d'une nouvelle langue (cf. chapitre V, 1.2., apprentissage de l'anglais par des locuteurs francophones ; phénomène de crible phonologique).

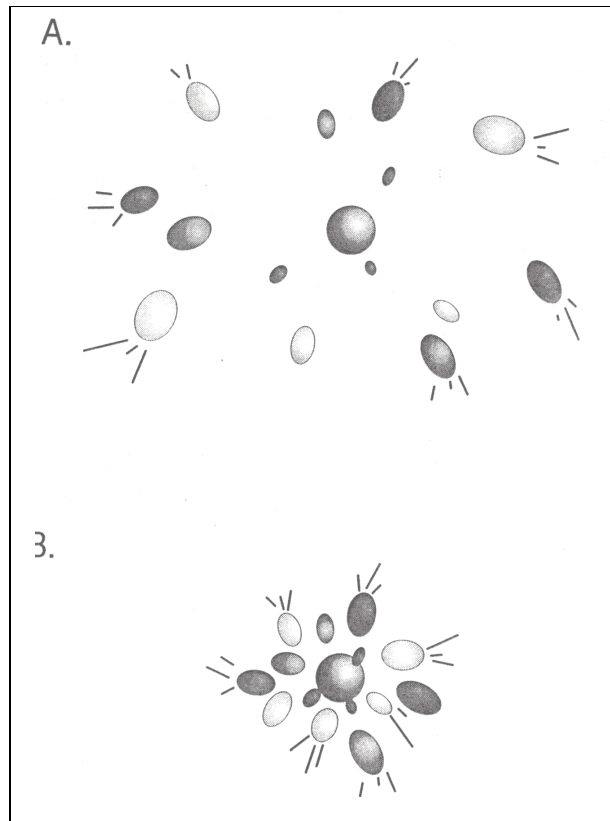


Figure n°50 : effet « magnétique » des prototypes de la LM (emprunté à Kuhl, 1994 : 124).

Kuhl et al. (1992) ont testé la discrimination de la voyelle [I] par des bébés américains et suédois afin de déterminer les effets potentiels des prototypes de cette voyelle en anglais et en suédois. Dans chacune de ces deux langues, le prototype de la voyelle n'est pas localisé au même endroit et la voyelle anglaise n'est pas perçue comme un prototype par un adulte suédois. Si la langue a déjà organisé à 6 mois une représentation de ces voyelles, il y aura une différence d'assimilation. Les résultats montrent qu'il existe bien une représentation de l'espace vocalique adaptée à la LM. En effet, les bébés américains se montrent sensibles uniquement à l'effet attractif de la voyelle anglaise et réciproquement pour les bébés suédois. Kuhl en conclut que « *by 6 months of age, exposure to the ambient language alters infants' perception of the phonetic units of language. This is the earliest age at which linguistic experience has been found to affect perception of the building blocks of language* » (Kuhl, 1994 : 126).

3.1.2.2. Focalisation sur la structure consonantique de la LM à partir de 10-12 mois

Après les voyelles, l'enfant se spécialise sur des constructions plus complexes que sont les consonnes de sa LM. Ces sons sont en effet produits plus rapidement que les voyelles et demandent plus de précision dans les gestes articulatoires. La focalisation sur les voyelles de la LM avait entraîné une perte de la sensibilité pour les voyelles non natives. Le même phénomène se produit vers l'âge de 10-12 mois pour les consonnes où l'attention de l'enfant se focalise sur les contrastes consonantiques de sa LM. Cette « spécialisation » entraîne une forte altération de ses capacités de discrimination pour les consonnes non natives, qu'il percevait pourtant bien à 2 mois, 4 mois et 6 mois. Cette **seconde réorganisation perceptive** le rapproche des capacités de discrimination d'un auditeur adulte pour le même type de contrastes. Elle serait liée à un traitement plus analytique qui permet à l'enfant, après un contact suffisant avec sa LM, de se spécialiser sur des constructions plus complexes et de les mémoriser. Ce type de traitement nécessite un traitement cognitif plus « lourd » de la part de l'enfant et ses capacités initiales à discriminer une grande variété de contrastes pourraient provenir d'un traitement global et donc, plus superficiel (cf. 1.2.2.).

L'étude menée par Werker et Tees (1984) constitue l'une des recherches les plus connues révélant les effets de cette réorganisation perceptive. Les auteurs ont testé les capacités de discrimination de bébés anglophones, à trois âges différents, 6-8 mois, 8-10 mois et 10-12 mois. Les contrastes non natifs testés appartenaient à deux langues très éloignées de la LM des bébés, l'hindi et le salish (langue thomson). Il est apparu que les bébés les plus jeunes discriminaient sans aucune difficulté les différents contrastes alors que les enfants les plus âgés échouaient dans cette tâche. Quant aux bébés de 8-10 mois, leurs résultats étaient mitigés. Cette étude révèle que l'acculturation à la LM est progressive. Best (1993) propose un modèle d'acquisition du langage, qui prend en compte cette réorganisation. Jusqu'à 6 mois, le bébé observe et perçoit les gestes articulatoires (« phonologie gestuelle »), sans être influencé par sa LM. A partir de 6 à 10 mois et jusqu'à 18 à 24 mois, il commence à reconnaître les gestes articulatoires natifs, mais pas encore leur valeur fonctionnelle. La transition s'effectue vers 10-12 mois pour les consonnes et de manière plus précoce (6 mois) pour les voyelles. C'est seulement à partir de 18 mois qu'il apprendra la fonctionnalité, la syntaxe et la morphologie (pour une description de son modèle d'assimilation perceptive, se reporter au chapitre V, partie 1.2.).

3.1.3. Spécialisation progressive sur le système musical

3.1.3.1. Définition de l'acculturation musicale

Si l'environnement linguistique modèle à ce point les capacités perceptives du bébé, il devrait se produire la même « acculturation » avec le SMM. Pour les petits occidentaux, ce SMM est celui de la musique tonale. Cela ne veut pas dire qu'ils ne peuvent pas être exposés à d'autres systèmes, mais que, dès la naissance, les auditeurs occidentaux sont majoritairement exposés à la musique tonale, notamment à cause de la très large diffusion de ce type de musique par les différents médias (musique de variétés notamment). Francès (1958 : 108-109) définit l'acculturation musicale comme

« un système d'habitudes perceptives qui s'enchaînent les unes aux autres, sans que les termes objectifs dont dépend leur acquisition soient consciemment posés par les sujets ».

Quelles sont ces habitudes perceptives et comment se mettent-elles en place ?

3.1.3.2. Bons patterns ou intégration des régularités de l'environnement ?

Les notes distantes d'une octave (rapport de fréquences 2 : 1) sont perçues comme équivalentes et portent d'ailleurs le même nom dans la plupart des cultures musicales. Ce principe, appelé **équivalence d'octave**, constitue une caractéristique musicale presque universelle selon Dowling et Harwood, (1986). Demany (1982) a montré que cette équivalence est perçue de façon très précoce, par des bébés de 2 mois et demi. Elle pourrait relever d'une contrainte de traitement du système perceptif. Dans la plupart des systèmes musicaux, l'octave est divisée en un nombre restreint de hauteurs distinctes (en général entre 5 et 7), qui composent les **gammes**. Dans la musique occidentale, la gamme la plus utilisée est la **gamme majeure** et les adultes sans formation musicale ont une connaissance tacite des différentes fonctions des notes de cette gamme (Krumhansl et Jusczyk, 1990 ; pour une description, voir chapitre I, 1.2.2.4.). La **triade majeure** occupe une place centrale au sein de la gamme majeure. Ainsi, Cohen et al. (1987) ont montré que, si les auditeurs adultes appliquent une stratégie de traitement par contours, habituelle pour une mélodie non familière, ils devraient échouer à détecter le changement d'intervalle si le contour est préservé (voir 1.2.2). Effectivement, les résultats montrent que les performances dans la tâche de discrimination sont beaucoup plus faibles pour les modifications affectant les intervalles, que pour les modifications affectant les contours.

Cependant, lorsque c'est la triade majeure qui sert de standard, les performances sont supérieures que pour la triade augmentée. Il est donc possible que la triade majeure soit un **bon pattern**, puisqu'elle permet aux auditeurs de dépasser leur stratégie habituelle de traitement par contours. Dans une autre étude, Trehub et al. (1990) ont montré que les mélodies occidentales conventionnelles, composées essentiellement à partir de triades majeures, facilitaient le traitement des mélodies chez des bébés de 6 mois. Les bébés seraient donc non seulement sensibles aux relations entre les notes à l'intérieur de la mélodie (hiérarchie intra-tonale, voir chapitre I, 1.2.2.4.), mais particulièrement sensibles à la triade majeure.

Les notes extrêmes de la triade sont en outre reliées par une **quinte juste** (do et sol, dans l'accord do-mi-sol en do majeur). Piston (1969) souligne le rôle central de cet intervalle, qui servirait de bonne forme autant pour les sons musicaux que pour les sons de la parole (voir ce chapitre, 3.1.2.1.). Cuddy et al. (1979) ont montré que les auditeurs adultes discriminent mieux les changements mélodiques lorsque la mélodie standard et la mélodie comparative sont reliées par un intervalle de quinte juste. Une autre étude révèle que la détection de changements d'un demi-ton dans la mélodie est plus aisée dans les mélodies composées à partir de gammes majeures, plutôt qu'à partir de gammes par tons, et ce, même pour des bébés âgés de 6 mois (Trehub, 1994).

Il reste cependant à déterminer si le statut particulier de la quinte juste, de la triade majeure et de la gamme majeure viennent du fait qu'elles soient des formes « intrinsèquement » bonnes, c'est-à-dire naturellement faciles à traiter ou du fait qu'elles apparaissent très souvent dans l'environnement musical du bébé. Pour Schenker (1906⁴), la triade majeure serait une forme consonante, stable et peut-être prototypique, alors que la triade augmentée serait dissonante et instable. En ce qui concerne la quinte juste, son rapport de fréquence simple (3 : 2) la rendrait particulièrement consonante (Burns et Ward, 1982 ; Rakowski, 1990). Kolinski (1967) relève que, dans la plupart des cultures, on trouve des adultes qui se mettent à entonner une chanson à la quinte ou à l'octave alors même qu'ils pensent être à l'unisson avec le groupe. Quant à la gamme majeure, elle posséderait, selon Balzano (1980), une certaine élégance structurelle. De là à dire que ces patterns sont « meilleurs » et plus « naturels » par rapport aux patterns d'autres cultures musicales, il y a un pas qu'il nous semble dangereux de franchir.

⁴ Cité par Trehub et Trainor (1994 : 324).

Il est d'autant plus dangereux que certains chercheurs comme Trehub par exemple (1994 : 336), considèrent les bébés de 6 mois comme des auditeurs naïfs de toute expérience avec leur SMM. Or la sensibilité du fœtus à la musique nous laisse entrevoir que l'acculturation est un long processus débutant bien avant la naissance (voir 2.1.3.). Le bébé de 6 mois est un auditeur déjà « aguerri », puisque bénéficiant d'une expérience longue de 10 mois avec son SMM. Ce point de vue oriente l'interprétation de leurs résultats et les amène à rejeter l'hypothèse d'une intégration progressive par l'enfant des régularités de son environnement. Pourtant, cette hypothèse n'entre pas en contradiction avec la présence de « bons patterns », qui faciliteraient le traitement de la musique (les fameuses « bonnes formes » de la Gestalt Theorie). Pour nuancer ce débat, il faudrait d'abord déterminer s'il se produit une acculturation en faveur du SMM, du même type que celle qui se produit pour la LM.

3.1.3.3. Evolution de la perception des gammes non natives

Il existe beaucoup moins de travaux sur l'acculturation musicale que sur l'acculturation avec la LM. Néanmoins, ces travaux suggèrent des différences de traitement entre les auditeurs adultes et les bébés. Alors que les bébés sont capables de détecter une note changée en dehors et à l'intérieur de la tonalité (Trainor et Trehub, 1992), les adultes détectent plus facilement ce changement en dehors de la tonalité (Cuddy et al., 1979 ; Trainor et Trehub, 1992). Ces résultats suggèrent que les bébés n'ont pas encore été assez exposés à leur SMM. Si la perception des bébés de 8 mois est différente de celles des adultes, il doit se produire à un moment donné une réorganisation perceptive en faveur du SMM.

Dans la musique tonale, l'intervalle de base est le demi-ton (plus petit dénominateur commun, rapport de 1.0595 ou 15/16). Tous les autres intervalles sont constitués à partir de lui et de ses multiples. Les gammes javanaises (appelées Pelog Scales) sont composées d'intervalles légèrement plus grands ou plus petits que les intervalles utilisés dans la musique tonale (cf. figure n°51). Ils ne sont pas fondés sur un rapport de fréquence constant et on ne peut pas les décomposer en unités stables. Ces intervalles possèdent donc une structure plus complexe, ce qui devrait les rendre plus difficiles à percevoir, même en faisant abstraction du poids de la culture musicale.

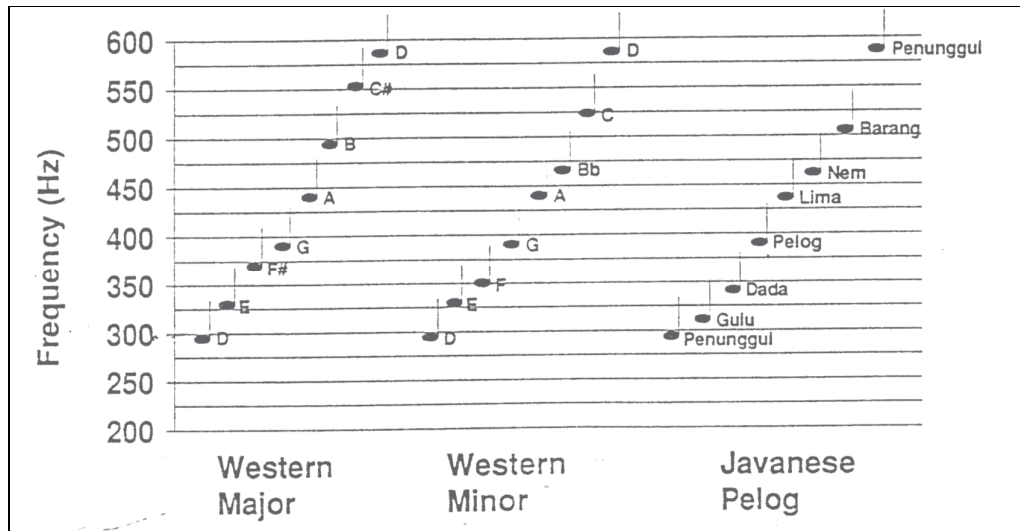


Figure n°51 : structuration des gammes majeures, mineures et javanaises (emprunté à Lynch et al., 1991 : 969).

Lynch et al. (1990) ont construit plusieurs mélodies de 7 notes, à partir de gammes majeures, de gammes mineures et de gammes « pelog ». La tâche consiste à discriminer la mélodie originale d'une mélodie dont la cinquième note est modifiée. Les résultats montrent que les bébés de 6 mois sont capables de discriminer les mélodies et leurs versions modifiées à partir des gammes majeures, mais également dans les mélodies construites à partir de gammes « pelog », ce qui n'est pas le cas chez les adultes. La performance similaire des bébés pour les gammes natives et non natives suggère que les gammes occidentales ne sont pas plus faciles à traiter que les gammes javanaises. En revanche, le poids d'une acculturation musicale est révélé par le fait que les adultes manifestent une meilleure performance pour les gammes natives que pour les gammes non natives. Selon Francès (1958), les auditeurs adultes ont intégré les intervalles spécifiques à leur culture. Ainsi, lorsque les auditeurs occidentaux écoutent une musique organisée selon une division différente de la gamme diatonique, comme c'est le cas des gammes javanaises par exemple, ils auront tendance, par un

« processus d'abstraction », à « réduire en notes une courbe vivante, et en intervalles inscrits sur une portée des intervalles qui en sont souvent fort écartés » (Francès, 1958 : 48-49⁵).

C'est le même phénomène que celui que nous avons appelé « crible tonal » dans le chapitre I (2.1.1.2.). Dans une autre étude de Lynch et Eilers (1992), les bébés de 6 mois se montrent meilleurs pour discriminer des changements intervenant dans des gammes majeures et gammes augmentées sans qu'il y ait beaucoup de différences entre les deux.

⁵ Cité par Zenatti (1994 : 330).

Ils ont également de bonnes performances pour les gammes pelog. En revanche, les performances des bébés de 10-12 mois sont mauvaises pour les gammes augmentées et les gammes pelog, alors qu'ils réussissent mieux avec les gammes majeures. Leurs résultats sont similaires à ceux des auditeurs adultes.

L'ensemble de ces résultats suggère que les bébés connaissent une période de réorganisation spécifiquement culturelle sur leur SMM. Cette réorganisation interviendrait entre l'âge de 6 et 12 mois. En effet, les bébés de 10-12 mois ne sont plus capables de discriminer les changements intervenant dans les gammes non natives, alors qu'ils étaient capables de le faire jusqu'à 6 mois. L'attention du bébé se fixe progressivement sur les schémas tonals spécifiques à son SMM. Ces schémas vont jouer le rôle de filtres en « modifiant » la perception des schémas musicaux non natifs en fonction de leur propre structure. Comme il existe un crible phonologique en faveur de la LM, il existe donc un « crible musical » en faveur du SMM (voir le récapitulatif sur la spécialisation progressive sur la LM et le SMM dans le Volume II, annexe 1.4.).

Alors que l'on peut remarquer les effets de l'acculturation linguistique car elle débouche sur l'acquisition de la LM, il n'en va pas de même pour l'acculturation musicale, dont les auditeurs n'ont souvent pas conscience. Pourtant, sans le savoir, ils ont intégré tous les principes qui leurs permettent de structurer la perception de leur SMM. On pourrait donc dire que tous les auditeurs ont une « connaissance » musicale car ils ont une connaissance intuitive des principes qui gouvernent leur SMM, comme ils ont une connaissance intuitive de la grammaire de leur LM.

3.2. Production

Chez l'être humain, les vocalisations se développent dans deux domaines différents et hautement structurés, les productions musicales et les productions verbales. Cependant, pendant les premiers mois de la vie, il paraît très difficile de différencier ce qui relève de la parole de ce qui relève de la musique et plus spécialement du chant.

Selon Papousek (1995 : 123-124) :

« Il n'est pas interdit de penser qu'il est impossible de discriminer les précurseurs du chant spontané des précurseurs du parler précoce. La communication préverbale pourrait bien être une route ontogénétique commune, le long de laquelle deux capacités extrêmement structurées et exclusivement humaines se développent : (a) le langage conduisant à la communication verbale et à la pensée, et (b) le chant autorisant les activités créatrices de la musique vocale. Ces deux capacités sont aussi intimement liées aux fonctions affectives du signal vocal et de la communication ».

Langue et musique sont d'autant plus liées qu'elles sont produites avec le même « générateur » de sons, c'est-à-dire l'appareil vocal :

« à l'évidence, la dualité de la fonction du tractus vocal humain témoigne de la liaison entre langage et musique : organe de la parole, il est aussi instrument de musique. C'est pourquoi les travaux touchant à la petite enfance accordent de plus en plus d'attention à la précocité de son fonctionnement » (Papousek, 1995 : 49).

Pendant ses premiers mois de vie, le bébé va progressivement explorer toutes les possibilités que lui offre cet instrument. A cette période de liberté vocale et d'intense improvisation, succèdera une période où l'enfant devra plier sa virtuosité aux contraintes imposées par son environnement. Il pourra alors s'engager dans la production différenciée de sa LM et de son SMM.

3.2.1. Maîtrise progressive de l'appareil phonatoire

Pendant la première année, l'émergence du langage de l'enfant est aujourd'hui très bien décrite et nous nous référerons le plus souvent aux travaux de Konopczynski (1990, 1991, 1994). Au cours du premier mois, le bébé va apprendre à contrôler ses mouvements respiratoires et le fonctionnement de son larynx. Cette première « prise en main » de son appareil phonatoire passe par des vocalisations réflexes, des cris et des sons végétatifs (**phase de phonation** : contrôle de la phonation et de la respiration). Peu à peu, il va entrer dans la **phase de résonance** qui va s'étendre jusqu'à son quatrième mois. Au cours de cette période, il va produire des syllabes primitives, composées de sons quasi-vocaliques et de contoïdes articulés à l'arrière de la gorge (il se trouve la plupart du temps en position couchée). Selon Papousek (1995), les véritables précurseurs communs à la musique et à la langue se trouvent dans ces sons quasi résonants.

En ce qui concerne la mélodie, le bébé commence à produire des contours descendants dès la fin de la première semaine, ce qui pourrait résulter de la baisse de la pression de l'air sous-glottique. L'apparition des premiers contours ascendants vers la sixième semaine marque le début du contrôle des cordes vocales. Ces contours vont commencer à se complexifier dès la dixième semaine. Vers 3 mois commence une **phase exploratoire** et **d'expansion vocale**. L'enfant va jouer avec son conduit vocal comme un musicien le ferait avec son instrument, en testant toutes les combinaisons articulatoires possibles et les sons résultant de ces combinaisons, ainsi que les différents modes de phonation. Il se sert de l'ensemble de ses cavités phonatoires et fait pleinement résonner sa voix, son développement physiologique lui permettant d'exercer un contrôle respiratoire, phonatoire et articulatoire croissant. Pendant cette phase, il apprend à maîtriser les différents paramètres, ce qui lui procure un plaisir évident (Papousek et Papousek, 1981). Oller (1980) a relevé une augmentation de la durée des vocoïdes entre 3 et 5 mois, allant de 700 ms à 1400 ms. Vers 3-4 mois, tous les schémas mélodiques sont représentés dans la production de l'enfant, bien que 80 % d'entre eux soient descendants. La tessiture s'élargit, le bébé passant quelquefois brusquement du craquement vocal dans le registre des graves (90 Hz) aux sons sur-suraigus (2000 Hz) et du hurlement au murmure et inversement (Konopczynski, 1991, 1994). A partir du cinquième mois, sa fréquence fondamentale moyenne (Fom) devient très variable alors qu'elle était jusque là relativement stable. C'est pendant cette période que la richesse mélodique est la plus remarquable. Lorsqu'il est en interaction avec un adulte, le bébé essaie d'imiter les sons produits par son « interlocuteur ». Ainsi, dès le troisième mois, le bébé est capable d'imiter une mélodie ou les sons émis par l'adulte si ces sons appartiennent à son répertoire vocal. Les parents encouragent d'ailleurs ce type d'imitation vocale et la récompensent affectivement lorsqu'elle est réussie. Les tours de parole s'installent progressivement, donnant naissance à de petits dialogues vocaux entre l'adulte et l'enfant qui s'imitent mutuellement.

3.2.2. Le babillage canonique : l'entrée différenciée dans le langage et la musique

Vers 6 mois, apparaissent les premières combinaisons de sons contoïdes et vocoïdes avec fermeture du tractus vocal. Ces « proto-syllabes » se caractérisent par une articulation relativement lâche et des transitions lentes entre les sons. Oller (1980) parle de « **marginal babbling** », terme auquel nous préférons celui de babillage rudimentaire (Vinter, 1994). C'est ce type de babillage qui va permettre l'apparition du babillage canonique.

Parallèlement, la position du bébé se verticalise progressivement, entraînant un changement de configuration du tractus vocal vers 7-8 mois et une augmentation de volume de ces cavités supra-glottiques.

Entre 6-8 mois, le **babillage canonique** se met en place de manière relativement brusque. Il se caractérise par une alternance rythmique de mouvements de fermeture et d'ouverture de la bouche, qui permet la production de syllabes bien formées, de type CV. Il est très important de noter que le babillage canonique n'apparaît qu'en situation d'interaction avec un adulte et, lorsque le bébé est seul, dans sa chambre par exemple, il continue ses jeux vocaux, jeux qui se prolongeront bien après l'apparition des premiers mots. Alors que le jeu solitaire possède un caractère fortement aléatoire (arythmie, syllabes rudimentaires, vocoïdes très longs, instabilité de hauteur maximale), le babillage canonique se caractérise à l'inverse par une grande structuration. Selon Konopczynski (1990, 1991), le babillage canonique ouvre la voie du langage propre. Il constitue en effet une étape primordiale pour que l'enfant puisse accéder ensuite au langage référentiel (correspondant aux premiers mots et la mise en place de la syntaxe) ; c'est un véritable « protolangage » selon l'auteur. Papousek explique que son émergence va déclencher un changement de comportement chez les parents :

« dès leur émergence, les syllabes en canon sont considérées comme des protomots potentiels : les parents leur attribuent en effet des significations et en font une utilisation déclarative pour désigner des personnes, des objets et des événements du micro-environnement du bébé » (1995 : 52).

En effet, les premières syllabes du babillage canonique sont très souvent composées d'occlusives labiales [β,μ] et de voyelles centrales [α,A] (Oller, 1980). Lorsque le bébé produit des combinaisons de syllabes du type [baba], [papa] ou [mama], elles sont immédiatement interprétées par les parents comme « papa » ou « maman » et ce d'autant plus spontanément que les parents attendaient ces « premiers mots » avec impatience. Le bébé va progressivement comprendre que ces séquences suscitent des réactions de joie de la part de ses parents et il est probable qu'il les reproduise pour leur faire plaisir. L'entrée dans la langue se fait donc dans l'interaction. Peu à peu, il va lui aussi leur attribuer un sens et c'est véritablement au moment où elles endosseront une signification pour le bébé qu'elles passeront du statut de « proto-mots » au statut de « mots » à part entière. Le lien entre la forme acoustique et le sens devient arbitraire, et c'est justement la nature de ce lien qui différencie le langage de la musique (voir chapitre I, 3.3.).

Il est possible que le bébé différencie déjà depuis longtemps la musique de la langue, en tout cas dans la parole des adultes (les situations d'émission et les attitudes sont très différentes), mais en revanche, pour l'adulte, il est peut être impossible de différencier clairement ce qui relève de la langue de ce qui relève de la musique dans les productions du bébé durant les premiers mois. D'ailleurs, avant le babillage canonique, l'adulte n'interprète pas les productions du bébé en termes linguistiques. Comme l'affirme Papousek (1995 : 53),

« au stade préverbal du bébé, il est extrêmement difficile de savoir si certaines interventions dues à l'intuition parentale encouragent plus particulièrement l'acquisition du langage, le jeu ou la musicalité. En revanche, lorsque la communication verbale s'installe, la spécificité de ces divers types de soutien s'affirme sous l'influence de facteurs culturels ».

Le babillage étant le véritable point d'entrée dans la communication verbale, il est probable que c'est à partir de son émergence que les parents vont pouvoir identifier facilement les productions linguistiques des productions musicales et en conséquence, apporter des stimulations différenciées à leur bébé. Cette différenciation va à son tour favoriser la séparation de la langue et de la musique dans les productions du bébé. Il est donc possible que l'entrée dans le SMM se fasse clairement à partir du babillage canonique.

3.2.3. Le langage émergent

Le protolangage contient déjà les caractéristiques rythmiques et intonatives de la LM, c'est-à-dire sa musique spécifique, ce qui le rend reconnaissable par des auditeurs de la même communauté linguistique (de Boysson-Bardies et al., 1984). Au niveau temporel, le bébé produit des syllabes ayant approximativement toutes la même durée. Selon Konopczynski (1990), cette durée (250 ms) est proche de la durée moyenne des syllabes adultes (200 ms) ; l'organisation de type isochronique est en place vers 8 mois et on la retrouve quelle que soit la langue de l'environnement. Mais, peu à peu elle va évoluer pour se rapprocher progressivement de la rythmique spécifique de la LM. Ainsi, pour un bébé français, l'organisation isochronique ne répond pas encore à la configuration rythmique du français, où la dernière syllabe du groupe rythmique est allongée (accentuation du français, cf. chapitre V, 2.1.1.1.). Cependant, très rapidement, les syllabes non finales vont se raccourcir, tandis que la durée des syllabes finales devient de plus en plus instable (trop courte ou trop longue, preuve d'un apprentissage par essai/erreur). L'allongement final qui caractérise le français se met progressivement en place pour se stabiliser entre 13 et 16 mois.

La durée des syllabes finales atteindra finalement le double de la durée des syllabes non finales, comme dans les productions adultes. Dans les langues où la place de l'accent est mobile comme l'anglais⁶, l'espagnol, le hongrois et le portugais, la rythmique est beaucoup plus difficile à acquérir.

Au niveau des paramètres mélodiques, le bébé contrôle sa voix en situation d'interaction (Fo moyen à 340 Hz), par rapport à la situation de jeu vocal. Les mélodies, jusque là riches et variées, se réduisent en nombre et leurs formes se stabilisent. Dès 9 mois, le bébé sait reproduire la configuration mélodique spécifique des énoncés interrogatifs, énonciatifs, exclamatifs, ainsi que celle des appels et des ordres. Ces productions sont correctement interprétées par les parents et même par des auditeurs extérieurs au cercle familial qui leur attribuent une fonction (phatique) ou une modalité précise (question, énonciation). Et c'est uniquement l'intonation qui permet cette différenciation (Konopczynski et Tessier, 1994), puisqu'il n'y a pas de couche verbale dans les énoncés du bébé. Ainsi, à la fin du dixième mois,

« un système premier est en place ; l'enfant dispose de quelques schèmes prosodiques, oppositifs, mais vides ; ces schèmes ou macro-segments forment un moule, un cadre prosodique qui va progressivement se remplir » (Konopczynski et Tessier, 1994 : 172).

Ces fameux « moules mélodiques » vont jouer un grand rôle dans la suite de l'acquisition, puisqu'ils vont recevoir le niveau segmental. Dans les stades suivants, l'enfant n'aura plus qu'à les remplir avec les premiers mots, au moment du babillage mixte (9-13 mois) et de la période du mot unique (12-18 mois) :

« au fur et à mesure qu'il s'approprie de nouvelles unités lexicales, il les introduit dans les moules prosodiques préexistants, sans – apparemment – rencontrer la moindre difficulté. Il connaît d'abord la musique de la phrase, il y met ensuite les paroles ; les éventuels trous ne le perturbent en aucune façon, ils sont comblés sans hésitation par du proto-langage » (Konopczynski et Tessier, 1994 : 174).

Plus tard, ces moules mélodiques formeront avec la rythmique spécifique de la LM (en place, rappelons le, dès le seizième mois chez le bébé français), la « structure d'accueil »⁷ de structures syntaxiques plus complexes (combinaison de 2 mots vers 18-20 mois).

⁶ Pour l'acquisition de la rythmique anglaise par le bébé, voir chapitre V, 2.1.3.4.

⁷ Terme emprunté à Konopczynski (1999 : 68).

3.2.4. Le chant émergent

Comparativement aux travaux sur l'acquisition du langage, il existe peu de travaux sur l'émergence des productions musicales et notamment sur l'imitation des mélodies et l'apparition des premiers chants spontanés. C'est d'autant plus dommage que les premières productions musicales et langagières sont indifférenciées durant les premiers mois et qu'il serait intéressant de mettre à jour le moment précis de leur divergence.

De ce point de vue, l'approche de Papousek et Papousek (1981) est particulièrement intéressante, car elle associe une approche linguistique à une approche musicologique. Ils ont établi un journal très détaillé des productions de leur fille Tanya de son premier à son seizième mois, en tentant d'identifier les précurseurs de la langue et les précurseurs de la musique. Ils ont ainsi relevé des capacités d'imitation musicale plus précoces que celles qui avaient été notées dans la littérature (7 mois pour Fridman, 1974⁸ et 14 mois pour Wing, 1941⁸) : dès 3 mois et demi, Tanya réussit à imiter des notes chantées ou jouées au piano. A 4 mois, elle essaie d'imiter des séquences plus longues et réussit à reproduire un pattern ressemblant au modèle adulte (même longueur, certains fragments mélodiques et rythmiques identiques). Ces capacités semblent toutefois communes à la langue, puisque Tanya est capable d'imiter les contours mélodiques de fragments de LAE dès 2 mois. En ce qui concerne ses productions spontanées, elles se caractérisent pendant les six premiers mois par des contours ondulants et continus. Les premières mélodies comportant des hauteurs discrètes apparaissent à 7 mois. La figure n°52 représente la transcription musicale d'une de ces mélodies. On peut y voir une suite de 8 tons de hauteur fixe, nettement délimités les uns des autres par des silences plus ou moins longs.

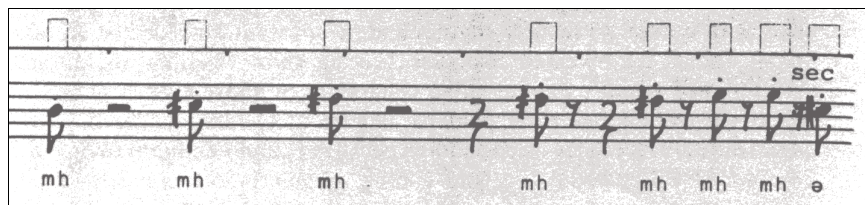


Figure n°52 : mélodie improvisée par Tanya à 7 mois (emprunté à Papousek et Papousek, 1981 : 190).

Selon Dowling (1994 : 172),

« les chansons sont distinctes de la parole en ce qu'elles évoluent à des niveaux discrets de hauteur, chaque voyelle reposant sur une hauteur plus ou moins stable ».

⁸ Cité par Papousek et Papousek (1981 : 193).

Nous prendrons le soin de préciser que cette définition s'applique à la musique tonale, où les hauteurs sont fixées et clairement délimitées les unes des autres (voir chapitre I, 1.2.2.). Les chansons enfantines sont particulièrement fidèles à cette configuration. Par rapport aux contours continus, la réalisation d'une telle séquence « staccato » nécessite une certaine maîtrise du tractus vocal et cette maîtrise coïncide avec l'apparition du babillage canonique (Papousek et Papousek, 1981 : 188).

C'est donc au moment du babillage canonique que les productions musicales deviennent clairement divergentes des productions linguistiques. Après l'apparition des premières mélodies, Papousek et Papousek réussissent clairement à distinguer les deux types de productions chez Tanya :

« in the further development of musical elements, two parallel lines become obvious, the speech-like intonation contours accompanying the linguistic differentiation, and the musical melodic contours in imitated, or invented songs » (1981 : 190).

Lewis (1936)⁹ considère le babillage comme une forme rudimentaire d'art et le début de l'usage esthétique du langage. Il est vrai que la rythmique du babillage canonique est très musicale en raison de sa régularité (écoulement régulier des syllabes à cause de leur isochronie). La mère pourrait la reprendre et lui associer une mélodie chantée, que l'enfant essaierait à son tour d'imiter. Lorsque le bébé est en situation d'interaction, il cherche en effet à

« s'harmoniser à la hauteur, à se mouvoir en synchronie avec la chanson ou à compléter la dernière note, à répéter certaines parties de l'air, voire à anticiper sur la phrase suivantes. De leur côté, les parents encouragent le chant commun en ralentissant le tempo ou en introduisant des silences afin de faciliter la participation du petit enfant par l'imitation ou la formation d'un duo joyeux » (Papousek, 1995 : 128).

L'entrée dans la musique, comme pour la langue, se fait donc dans l'interaction et « *le chant des parents est, semble-t-il, l'élément déterminant du sens musical du bébé* » (Papousek, 1995 : 129).

⁹ Cité par Papousek et Papousek (1981 : 212).

Ces observations s'accordent avec les données recueillies pour la perception. En effet, nous avons vu que le bébé connaissait une période de réorganisation perceptive sur son SMM entre 10 et 12 mois. A partir de cet âge, les petits occidentaux échouent à détecter des modifications intervenant dans une autre gamme que la gamme majeure. Ce processus de centration sur les structures musicales maternelles semble très efficace, puisque simultanément dans le cas de Tanya, on peut relever une préférence dans les productions spontanées pour les notes de la triade majeure. En revanche, le poids de la LM semble influencer les capacités perceptives de manière plus précoce que pour la musique, en raison peut-être d'une exposition plus massive à la LM qu'au SMM. Cependant, à l'instar de la musique, les productions du bébé semblent immédiatement refléter cette spécialisation. En effet, la date d'apparition du babillage canonique correspond approximativement à la période où l'attention du bébé se resserre sur les voyelles de sa LM.

Alors que, dans les premières productions du bébé, il était très difficile de différencier la langue de la musique, ses productions sont clairement influencées par sa LM et son SMM à la fin de la première année. A partir de cette différenciation, qui trouve sans doute son origine dans le babillage canonique et qui est renforcée dans les interactions avec l'entourage, les productions musicales et linguistiques vont diverger de manière radicale, comme nous allons le constater dans le chapitre suivant.

Bibliographie chapitre II

- Armitage, S.E., Baldwin, B.A. et Vince, M.A. (1980). The foetal sound environment of sheep. In *Science*, 208, 1173-1174.
- Balzano, G. J. (1980). The group-theoretic description of 12-fold and microtonal pitch systems. In *Computer Music Journal*, 4, 66-84.
- Bartlett, J. C. et Dowling, W. J. (1980). The recognition of transposed melodies: a key-distance effect. In *Journal of Experimental Psychology: Human Perception and Performance*, 6, 501-515.
- Benedict, H. (1979). Early lexical development: comprehension and production. In *Journal of Child Language*, 6, 183-200.
- Bergeson, Tonya R. et Trehub, Sandra E. (1999). Mothers' singing to infant and preschool children. In *Infant Behavior and Development*, 22, 1, 51-64.
- Bernstein Ratner, Nan (1986). Durational cues which mark clause boundaries in mother-child speech. In *Journal of Phonetics*, 14/2, 303-309.
- Bertoncini, Josiane (1993). Initial capacities for speech processing: infants' attention to prosodic cues for segmentation. In De Boysson-Bardies, De Schonen, Jusczyk, MacNeilage and Morton, 249-257.
- Bertoncini, Josiane et de Boysson-Bardies, Bénédicte (2000). La perception et la production de la parole avant deux ans. Dans Kail et Fayol, 95-136.
- Best, Catherine T., McRoberts, Gerald W. and Sithole, Nomathemb N. (1988). The phonological basis of perceptual loss of nonnative contrasts: maintenance of discrimination among Zulu clicks by english-speaking adults and infants. In *Journal of Experimental Psychology: Human Perception and Performance*, 14, 345-360.
- Best, Catherine T. (1993). Learning to perceive the sound pattern of English. In Rovee-Collier and Lipsitt.
- Best, Catherine T. (1994). The emergence of native-language phonological influences in infants: a perceptual assimilation model. In Goodman and Nusbaum, 167-224.
- Bigand, Emmanuel (1994). Contributions de la musique aux recherches sur la cognition auditive humaine. Dans Mc Adams et Bigand, 249-298.
- Bolinger, Dwight L. (1978). Intonation across languages. In Greeberg, Ferguson, and Moravcik, 2, 471-524.
- Bornstein, Marc H., Kessen, W. and Weiskopf, S. (1976). Color vision and hue categorisation in young human infants. In *Journal of Experimental Psychology: Human Perception and Performance*, 2, 115-129.
- Boysson-Bardies, Bénédicte de (1993). Ontogeny of language-specific syllabic productions. In Boysson-Bardies, Schonen de, Jusczyk, Mac Neilage and Morton, 353-363.
- Boysson-Bardies, Bénédicte de, Sagart, Laurent and Durand, Catherine (1984). Discernible differences in the babbling of infant according to target language. In *Journal of Child Language*, 11, 1-15.
- Boysson-Bardies, Bénédicte de, Hallé, Pierre A., Sagart, Laurent and Durand, Catherine (1989). A cross-linguistic investigation of vowel formants in babbling. In *Journal of Child Language*, 16, 1-17.
- Boysson-Bardies, Bénédicte de, Schonen, Scania, Jusczyk, Peter, MacNeilage and Morton (1993). *Developmental Neurocognition: speech and face processing in the first year of life*. Boston: Kluwer Academic Publishers.
- Bregman, Albert S. (1978). Auditory streaming is cumulative. In *Journal of Experimental Psychology: Human Perception and Performance*, 4, 380-387.
- Bregman, Albert S. (1981). Asking the "what for" question in auditory perception. In Kubovy and Pomerantz, 99-118.
- Bregman, Albert S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge : MIT Press.
- Bregman, Albert S. (1994). L'analyse des scènes auditives. Dans McAdams et Bigand, 11-39.
- Broen, P. A. (1972). The verbal environment of the language-learning child. In *American Speech and Hearing Association Monographs*, 17 (décembre).

- Burns, E. M. and Ward, W. D. (1982). Intervals, scales and tunings. In Deutsch, 241-269.
- Busnel, Marie-Claire (1994). Is there prenatal culture ? In Gardner, Chiarelli, Gardner and Plooij, 285-314.
- Busnel, Marie-Claire, Daffos, Fernand, Dolto-Tolitch, Lecanuet, Jean-Pierre et Negri, Romana (1997). *Que savent les Fœtus ?* Ramonville Saint-Agne : Erès, 103 p.
- Busnel, Marie-Claire (1997). Audition fœtale et réactivité prénatale à la voix maternelle “ adressée ”. In Busnel et al, 35-49.
- Caterette, E. C. and Friedman, M. P. (1976). *Handbook of Perception*, VII. New-York: Academic Press.
- Chapman, J. S. (1975). *The Relation between Auditory Stimulation of Short Gestation Infants and Their Gross Motor Limb Activity*. PHD, New-York University.
- Christophe, Anne, Dupoux Emmanuel, Bertoncini, Josianne and Mehler, Anne (1994). Do infants perceive word boundaries ? An empirical study of the bootstrapping of lexical acquisition. In *Journal of the Acoustical Society of America*, 95, 1570-1580.
- Cohen, Anthony J., Thorpe, Leigh A. and Trehub, Sandra, S. E. (1987). Infants’ perception of musical relations in short transposed tone sequences. In *Canadian Journal of Psychology*, 41, 33-47.
- Cohen, Antonie and Nooteboom, Sieb G. (1975). *Structure and Process in Speech Perception*. Berlin: Springer, 351 p.
- Cooper, W. E. and Paccia-Cooper (1980). *Syntax and Speech*. Cambridge: Harvard University Press.
- Cooper, R. P., & Aslin, Richard N. (1990). Preference for infant-directed speech in the first month after birth. In *Child Development*, 66, 1584-1595.
- Cruttenden, Alan (1986). *Intonation*. Cambridge: Cambridge University Press, 214 p.
- Cuddy, L. L., Cohen, Anthony J. and Miller, Joanne (1979). Melody recognition: The experimental application of musical rules. In *Canadian Journal of Psychology*, 33, 148-157.
- Cutting, James E. and Rosner, B. S. (1974). Categories and boundaries in speech and music. In *Perception and Psychoacoustics*, 16, 564-570.
- DeCasper, A. J. and Fifer, W. P. (1980). Of human bonding : Newborns prefer their mothers’ voices. In *Science*, 208, 1174-1176.
- DeCasper, A. J. and Prescott, P. A. (1984). Human newborns’ perception of male voices: preference, discrimination, and reinforcing value. In *Developmental Psychobiology*, 17, 481-491.
- DeCasper, A. J. and Spence, M. J. (1986). Prematernal speech influence newborns’ perception of speech sounds. In *Infant Behavior and Development*, 9, 133-150.
- DeCasper, A.J., Lecanuet, Jean-Pierre, Busnel, Marie-Claire, Granier-Deferre, C. et Maugeais, R. (1994). Fetal reactions to recurrent maternal speech. In *Infant Behavior and Development*, 17, 159-164.
- Delcloque, Philippe (1999). *Speech Technology Applications in CALL*. Dundee: University of Albertay Dundee, 87 p.
- Deliège, Irène et Sloboda, John, A. (1995). *Naissance et Développement du Sens Musical*. Paris : PUF, 307 p.
- Demany, Laurent, McKenzie, B. and Vurpillot, Eliane (1977). Rhythm perception in early infancy. In *Journal of Acoustic Society of America*, 76, 57-66.
- Demany, Laurent (1982). Auditory stream segregation in infancy. In *Infant Behavior and Development*, 5, 261-276.
- Deutsch, D. (1969). Music recognition. In *Psychological Review*, 76, 300-307.
- Deutsch, D. (1982). *The Psychology of Music*. New-York: Academic Press.
- Dowling, Jay W. (1978). Scale and contour: two components of a theory for melodies. In *Psychological Review*, 85, 341-354.
- Dowling, Jay W. (1994). La structuration mélodique. Dans Zenatti, 145-176.
- Dowling, Jay W. and Harwood, D. L. (1986). *Music Cognition*. Orlando: Academic Press.
- Eimas, Peter D. (1974). Auditory and linguistic processing of cues for place of articulation by infants. In *Perception and Psychophysics*, 16, 513-521.

- Eimas, Peter D. (1975). Auditory and phonetics coding of the cues for speech: discrimination of the r-l distinction by young infant. In *Perception and Psychophysics*, 18, 341-347.
- Fassbender, Christoph (1993). *Auditory grouping and segregation processes in infancy*. Norderstedt: Kaste Verlag.
- Fassbender, Christoph (1995). La sensibilité auditive du nourrisson aux paramètres acoustiques du langage et de la musique. In Deliège et Sloboda, 63-99.
- Feagans, C. Garvey, R. Golinkoff, M. T. Greenberg, C. Harding, & J. Bohannon (1984). *The Origins and Growth of Communication*. Norwood: New Jersey, Ablex Publishing Corporation.
- Feijoo, Jean (1975). Ut conscientia Noscatue. Dans *Cahiers de Soprologie*, 13, 14-20.
- Feijoo, Jean (1983). Le Fœtus, Pierre et le loup. In Herbinet, Etienne et Busnel, Marie-Claire, 193-206.
- Ferguson, Charles A. (1977). Baby talk as a simplified register. In Snow and Ferguson, 219-235.
- Fernald, Anne (1976). The mother's speech to the newborn. Paper presented at the Max-Plank Institute for Psychiatry, Munich (non published).
- Fernald, Anne (1984). The perceptual and affective salience of mothers' speech to infants. In Feagans, Garvey, Golinkoff, Greenberg, Harding & Bohannon, 5-29.
- Fernald, Anne (1985). Four-month-old infants prefer to listen to motherese. In *Infant Behavior and Development*, 8, 181-195.
- Fernald, Anne (1989). Intonation and communication intent in mother's speech to infants: is the melody the message ? In *Child Development*, 60, 1497-1510.
- Fernald, Anne and Kuhl, Patricia (1981). Fundamental frequency as an acoustic determinant of infant preference for " Motherese ". Boston: Paper presented at the Biennial Meeting of the Society for Research in Child Development.
- Fernald, Anne and Simon, Thomas (1984). Expanded intonation contours in mothers' speech to newborns. In *Developmental Psychology*, 20/1, 104-113.
- Fernald, Anne and Kuhl, Patricia (1987). Acoustic determinants of infant preference for motherese speech. In *Infant Behavior and Development*, 10, 279-293.
- Fernald, Anne and Mazzie, C. (1991). Prosody and focus in speech to infants mothers' speech to new-borns. In *Developmental Psychology*, 27, 209-221.
- Fisher, Cynthia L. and Tokura, H. (1996). Acoustic cues to grammatical structure in infant-directed speech: cross-linguistic evidence. In *Child Development*, 67, 3192-3218.
- Flege, James E. (1989). Using visual information to train foreign-language vowel production. In *Language Learning*, 38, 365-407.
- Francès, Robert (1958). *La Perception de la Musique*. Paris : Vrin, 408 p.
- Fridman, R. (1974). Los comienzos de la conducta musical. In *Biblioteca de Psicología Evolutiva Serie 2* (Vol. 13). Buenos Aires: Paidós.
- Gardner, R.A., Chiarelli, B.T., Gardner, B.T. and Plooi, F.X. (1994). *The Ethological Roots of Culture*. Kluwer Acad. Publish. Dordrecht Netherlands.
- Gathercole, Susan E. and Baddeley, Alan D. (1993). *Working memory and language*. Hove: Lawrence Erlbaum Associates.
- Gentil-Neymon, V. (1993). *Influence de la Voix de la Mère auprès des Prématurés en Service de Soins Intensifs*. Thèse de Doctorat de Médecine, Université Paris VII.
- Giacomi, A. et De Heredia, C. (1986). Réussites et échecs dans les communications linguistiques entre locuteurs francophones et locuteurs immigrés. In *Language*, 84, décembre.
- Gleitman, Lila R. and Wanner, E. (1982). *Language Acquisition: the State of the Art*. Cambridge. Cambridge: Cambridge University Press.
- Gleitman, Lila R., Gleitman, Harry, Landau, Barbara and Wanner, E. (1988). Where the learning begins: initial representations for language learning. In Newmeyer.
- Goodman, Judith C. and Nusbaum, Howard C. (1994). *The Development of Speech Perception: the Transition from Speech Sounds to Spoken Words*. Cambridge: MIT press.

- Greenberg, J.H., Ferguson, Charles A. and Moravcik, E. A (1978). *Universals of Human Language, Phonology*, 2. Standford: Stanford University Press.
- Grieser and Kuhl, Patricia K. (1989). Categorization of speech by infants: support for speech-sound prototypes. In *Developmental Psychology*, 25, 577-588.
- Halford, Brigitte K. and Pilch, Herbert (1994). *Intonation*. Tübingen: Gunter Narr Verlag.
- Hallé, Pierre A. and De Boysson-Bardies, Bénédicte de (1994). Emergence of an early receptive lexicon: infants' recognition of words. In *Infant Behavior and Development*, 17, 119-129.
- Hargreaves, David (1995). Développement du sens artistique et musical. Dans Deliège et Sloboda, 167-197.
- Hepper, P. G. (1991). An examination of fetal learning before and after birth. In *Irish Journal of Psychology*, 12, 95-107.
- Hepper, P. G., Scott, D. and Shahidullah, S. (1993). Newborn and fetal response to maternal voice. In *Journal of Reproductive and Infant Psychology*, 11, 147-153.
- Hepper and Shaidullah, (1994). Developmental of fetal hearing. In *Arch Dis Childhood*, 71, F81-F87.
- Herbinet, Etienne et Busnel, Marie-Claire (1983). *L'Aube des Sens*. Paris : Stock, 414 p.
- Hirsh-Pasek, Kathryn, Kemler, Nelson, Deborah, Jusczyk, Peter W., Wright, Cassidy, K., Druss, B. and Kennedy, L. (1987). Clauses are perceptual units for young infants. In *Cognition*, 26, 269-286.
- Hulse, S. H., Humpal, J. and Cynx, J. (1984). Discrimination and generalization of rhythmic and arhythmic sound patterns by European starlings. In *Music Perception*, 1, 442-464.
- Journiac, Claire (2002). *La Place de l'Autre dans la Relation Soignante : Contribution à l'Etude des Processus Relationnels à Travers la Voix*. Université de Bourgogne, 360p.
- Jusczyk, Peter W. (1989). Perception of cues to causal units in native and non-native languages. Paper presented at the biennial meeting of the *Society for Research in Child Development*, Kansas City, Avril.
- Jusczyk, Peter W. (1997). *The Discovery of Spoken Language*. Cambridge : MIT, 314 p.
- Jusczyk, Peter W. and Thompson, E. J. (1978). Perception of a phonetic contrast in multisyllabic utterances by two-month-old infants. In *Perception and Psychophysics*, 23, 105-109.
- Jusczyk, Peter W. and Krumhansl, Carol (1993). Pitch and rhythmic patterns affecting infants' sensitivity to musical phrase structure. In *Journal of Experimental Psychology: Human Perception and Performance*, 19/3, 627-640.
- Jusczyk, Peter W., Hirsh-Pasek, Kathryn, Kemler Nelson, D. G., Kennedy, L., Woodward, A. and Piwoz, J. (1992). Perception of acoustic correlates of major phrasal units by young infants. In *Cognitive Psychology*, 24, 252-293.
- Jusczyk, Peter W., Friederici, A. D., Wessels, J., Svenkerud, V. Y. and Jusczyk, A. M. (1993a). Infants' sensitivity to the sound patterns of native language words. In *Journal of Memory and Language*, 32, 402-420.
- Kagan, Jerome and Lewis, Michael (1965). Studies of attention in the human infant. In *Merill-Palmer Quarterly*, 11, 95-122.
- Kail, Michèle et Fayol, Michel (2000). *L'Acquisition du Langage : le Langage en Emergence de la Naissance à Trois ans*. Vol. I. Paris : PUF, 304 p.
- Katz, G. S., Cohn, J. F. and Moore, F. A. (1996). A combination of vocal fo dynamic and summary features discriminates between three pragmatic categories of infant-directed speech. In *Child Development*, 67, 205-217.
- Kavanagh and Cutting, James E. (1975). *The Role of Speech in Language*. Cambridge: MIT press.
- Kemler Nelson, D. G., Hirsh-Pasek, Kathryn, Jusczyk, Peter W. and Wright Cassidy, K. (1989). How the prosodic cues in motherese might assist language learning. In *Journal of Child Language*, 16, 53-68.
- Kimchi, R. (1992). Primacy of wholistic processing and global/local paradigm: a critical review. In *Psychological Bulletin*, 112, 24-38.
- Klatt, Dennis K. and Cooper, William E. (1975). Perception if segment duration in sentence contexts. In Cohen and Nooteboom, 69-86.
- Kleiber, Michel (1990). *La Sémantique du Prototype : Catégories et Sens Lexical*. Paris : P.U.F., 199 p.

- Kluender, K. R., Diehl, R. L. and Killeen, P. R. (1987). Japanese quail can learn phonetic categories. In *Science*, 237, 1195-1197.
- Kolinski, M. (1967). Recent trends in ethnomusicology. In *Ethnomusicology*, 11, 1-24.
- Konopczynski, Gabrielle (1990). *Le Langage Emergent : Caractéristiques Rythmiques*. Hamburg : Buske Verlag, 362 p.
- Konopczynski, Gabrielle (1991). *Le Langage Emergent II : Aspects Vocaux et Mélodiques*. Hamburg : Buske Verlag, 425 p.
- Konopczynski, Gabrielle (1999). L'acquisition du système prosodique de la langue maternelle et ses implications pour l'apprentissage d'une L2. Dans Delcloque Philippe, 62-70.
- Konopczynski, Gabrielle et Tessier, Sophie (1994). Structuration intonative du langage émergent. Dans Halford et Pilch, 157-192.
- Krumhansl, Carol L. and Jusczyk, Peter, W. (1990). Infant's perception of phrase structure in music. In *Psychological Science*, 1, 70-73.
- Kubovy and Pomerantz (1981). *Perceptual Organization*. Hillsdale: Erlbaum Assoc.
- Kuhl, Patricia K. (1992). Infants' perception and representation of speech: Development of a new theory. In Ohala, Nearey, Derwing, Hodge and Wiebe, 449-456.
- Kuhl, Patricia K. (1993). Innate predispositions and the effects of experience in speech perception: The native language magnet theory. In Boysson-Bardies de, Schonen, Jusczyk, MacNeilage and Morton, 259-274.
- Kuhl, Patricia K. (1994). Speech perception. In *Minifie*, 118-142.
- Kuhl, Patricia K. and Miller, J. D. (1975). Speech perception by the chinchilla: voiced-voiceless distinction in alveolar plosive consonants. In *Science*, 190, 69-72.
- Kuhl, Patricia K. and Miller, J. D. (1982). Discrimination of auditory target dimensions in the presence or absence of variation in a second dimension by infants. In *Perception and Psychophysics*, 31, 279-292.
- Kuhl, Patricia K. and Padden, D. M. (1982). Enhanced discriminability at the phonetic boundaries for voicing feature in macaques. In *Perception and Psychophysics*, 32, 542-550.
- Kuhl, Patricia K., Williams, K. A., Lacerda, F., Stevens, K. N. and Lindblom, Björn (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. In *Science*, 255, 606-608.
- Kuhl, Patricia K., Andruski, J.E., Chistovitch, I.A., Chistovitch, L.A., Kozhevnikova, E.V., Ryskina, V.L., Stolyarova, E.I., Sundberg, U. and Lacerda, F. (1997). Cross-language analysis of phonetic in language addressed to infants. In *Science*, 277, 684-686.
- Lecanuet, Jean-Pierre, Granier-Deferre, C., Jacquet, A. Y., Capponi, I. et Ledru, L. (1993). Prenatal discrimination of a male and a female voice uttering the same sentence. In *Early Development and Parenting*, 2, 217-228.
- Lecanuet, Jean-Pierre (1995). L'expérience auditive prénatale. In *Deliège et Sloboda*, 7-38.
- Lecanuet, Jean-Pierre (1997). Dans tous les sens... bref état des compétences sensorielles fœtales. In *Busnel et al.*, 17-34.
- Lecanuet, Jean-Pierre, Jacquet, A-Y., Hains S.M.J. et Kisilevsky, Barbara S. (2000). Effects of music on behaviour in term fetuses. *International Conference on Infant Studies* ; Brighton , Grande Bretagne, Juillet 3-6.
- Lecanuet, Jean-Pierre, Jacquet, Anne-Yvonne, Kisilevsky, Barbara, S. and Hains, Sylvia, M. J. (2000). Effects of music on behaviour in term fetuses. Poster, ICIS 2000 (International Conference on Infant Studies), Brighton.
- LeCours, A. R. (1975). Myelogenesis correlates of the development of speech and language. In *Lenneberg and Lenneberg*.
- Lenneberg, E. H. and Lenneberg, E. (1975). *Foundations of Language Development of Speech and Language*. New-York: Academic Press.
- Levarie, S. and Rudolph, N. (1978). Can newborn infants distinguish between tone and noise ? In *Perceptual and Motor Skills*, 47, 1123-1126.
- Lewis, Michael (1936). *Infant Speech: A Study of the Beginnings of Language*. New-York: Harcourt Brace.
- Liberman, Anatoly M., Cooper, F. S., Shankweiler, D. P. and Studdert-Kennedy, Michael G. (1967). Perception of the speech code. In *Psychological Review*, 74, 431-461.

- Lindblom, Björn, McNeilage and Studdert-Kennedy, Michael G. (1984).
- Lipsitt, Lewis P. and Rovee-Collier, Carolyn K. (1981). *Advances in Infancy Research*. Vol I : XXII. Norwood: Ablex, 267 p.
- Luce, Paul A. and Charles-Luce, Jan (1983). Contextual effects on the consonant/vowel ratio in speech production. Paper presented at the *Meeting of the Acoustical Society of America*, Cincinnati, May.
- Lynch, Michael P. (1993). Prototypical representations of music structure in infancy: theoretical exploration and a pilot study. In *Psychomusicology*, 12, 31-40.
- Lynch, Michael P., Eilers, Rebecca E., Oller, Kimbrough D. and Urbano, Richard C. (1990). Innateness experience and music perception. In *American Psychological Society*, I/4, 272-276.
- Lynch, Michael P. and Eilers, Rebecca, E. (1991). Children's perception of native and nonnative musical scales. In *Music Perception*, 9/1, 121-132.
- Lynch, Michael P., Eilers, Rebecca E., Oller, Kimbrough D. and Urbano, Richard C. (1991). Influences of acculturation and musical sophistication on perception of musical interval patterns. In *Journal of Experimental Psychology: Human Perception and Performance*, 17/4, 967-975.
- Lynch, Michael P. and Eilers, Rebecca, E. (1992). A study of perceptual development for musical tunings. In *Perception and Psychophysics*, 52/6, 599-608.
- McAdams, Stephen et Bigand, Emmanuel (1994). *Penser les Sons : Psychologie Cognitive de l'Audition*. Paris : P.U.F., 402 p.
- Malsheen, Bathsheb (1980). Two hypotheses for phonetic clarification in the speech of mothers-to-children. In *Yeni-Komshian, Kavanagh and Ferguson*, II, 173-184.
- Martin, Pierre (1996). *Eléments de Phonétique avec Application au Français*. Saint-Nicolas (Québec) : Presses Universitaires de l'Université de Laval, 253 p.
- Mehler, Jacques, Bertoncini, Josiane, Barrière, M., Jassik-Gershenfeld, Dora (1978). Infant recognition of Mother's voice. In *Perception*, 7, 491-497.
- Menyuk, Paula (1977). *Language and Maturation*. Cambridge: MIT Press.
- Menyuk, Paula (1994). Word acquisition: the important role of prosody. In *The Study of Sounds*, 23, 218-222.
- Miller, Joanne L. and Liberman, Anatoly M. (1979). Some effects of later occurring information on the perception of stop consonant and semivowel. *Perception & Psychophysics*, 25: 457-465.
- Minifie, Fred H. (1994). *Introduction to Communication Sciences and Disorders*. San Diego: Singular Publishing Group.
- Moog, H. (1976). *The Musical Experience of the Pre-school Child*. London: Scott.
- Moon, C., Cooper, R., and Fifer, W. P. (1993). Two-day-olds prefer their native language. In *Infant Behavior and Development*, 16, 495-500.
- Morgan, James L., Meier, R. P. and Newport, Elissa L. (1987). Structural packaging in the input to language learning: Contributions of prosodic and morphological marking of phrases to the acquisition of language ? In *Cognitive Psychology*, 19, 498-550.
- Morse, Philip A. (1972). The discrimination of speech and nonspeech stimuli in early infancy. In *Journal of Experimental Child Psychology*, 13, 477-492.
- Mosser, Christine (1990). *Effets physiologiques des Stimulations Sonores chez les Prématurés*. Thèse de Doctorat : Mondrini, Paris XII.
- Navon, D. (1977). Forest before trees: the precedence of global feature in visual perception. In *Cognitive Psychology*, 9, 353-383.
- Nazzi, Thierry, Kemler Nelson, D., Jusczyk, Peter and Jusczyk, A.M. (2000). Six-month-olds' detection of clauses embedded in continuous speech: Effects of prosodic well-formedness. In *Infancy*, 1/1, 123-147.
- Neisser, U. (1967). *Cognitive Psychology*. New-York: Appleton Century Crofts.
- Newmeyer, Frederick J. (1988). *Linguistics: The Cambridge Survey*. Cambridge: Cambridge University Press, III, 349 p.

- Nowik-Stern, Alisa, Marsha G. Clarkson, Mary K. Morris, Roger Bakeman (1998). The effect of premature infants' speech preferences on mother-preterm interaction. Poster, *11th Biennial International Conference on Infant Studies*, Atlanta, Georgia, April 2-5.
- Ohala, J. J., Nearey, T. M., Derwing, B. L., Hodge, M. M. and Wiebe, G. E. *Proceedings of the International Conference on Spoken Language Processing*. Edmonton, Alberta: University of Alberta.
- Oller, Kimbrough D. (1980). The emergence of the sounds of speech in infancy. In Yeni-Komshian, Kavanagh and Ferguson, I, 93-112.
- Panneton, R. K. (1985). *Prenatal auditory experience with melodies: Effects on postnatal auditory preferences in human newborns*. PhD, University of North Carolina at Greensboro.
- Papousek, Mechthild and Papousek, Hanus (1981). Musical elements in the infants vocalizations: Their significance for communication, cognition and creativity. In Lipsitt and Rovee-Collier, 163-224.
- Papousek, Mechthild, Bornstein, Marc H., Nuzzo, C., Papousek, Hanus and Symmes, D. (1990). Infant responses to prototypical melodic contours in parental speech. In *Infant Behavior and Development*, 13, 539-545.
- Papousek, Mechthild, Papousek, Hanus, & Symmes, D. (1991). The meaning of melodies in motherese in tone and stress languages. In *Infant Behavior and Development*, 14, 415-440.
- Papousek, Hanus (1995). Musicalité et petite enfance. Origines biologiques et culturelles de la précocité. In Deliège et Sloboda, 41-62.
- Papousek, Mechthild (1995). Le comportement parental intuitif, source cachée de la stimulation musicale dans la petite enfance. In Deliège et Sloboda, 101-130.
- Peters, Ann M (1983). *The Units of Language Acquisition*. Cambridge: Cambridge University Press.
- Pisoni, David B. (1977). Identification and discrimination of the relative onset of two component tones: Implications for voicing perception in stops. In *Journal of the Acoustical Society of America*, 61, 1352-1361.
- Piston (1969). *Harmony*. New-York: Norton.
- Polka, M. and Werker, Janet F. (1994). Developmental changes in perception of nonnative vowel contrasts. In *Journal of Experimental Psychology: Human Perception and Performance*, 20, 421-435.
- Proverbio, Alice Mado, Minniti, Alessia and Zani, Alberto (1998). Electrophysiological evidence of perceptual precedence of global vs. Local visual information. In *Cognitive Brain Research*, 6, 321-334.
- Pugin, Maryline (1983). *Les Stimulations Auditives et leur Importance dans la Genèse de la Relation Mère-Enfant*. Université de Besançon: DEA en Neurosciences.
- Pye, Clifton (1983). Mayan telegraphese: intonational determinants of inflectional development in Quiché Mayan. In *Language*, 59/3, 583-604.
- Querleu, Olivier, Renard, Xavier, Versyp, F., Paris-Delrue, L. and Crépin, Gilles (1988). Fetal hearing. In *European Journal of Obstetrics and Reproductive Biology*, 29, 191-212.
- Rakowski, A. (1990). Intonation variants of musical intervals in isolation and in musical contexts. In *Psychology of Music*, 18, 60-72.
- Read, C. and Schreiber, P. (1982). Why short subjects are hard to find ? In Gleitman and Wanner.
- Rovee-Collier, Carolyn K. and Lipsitt, Lewis P. (1993). *Advances in Infancy Research*. Ablex Publishing.
- Rovee-Collier, Carolyn K., Lipsitt, Lewis P. and Hayne, (1998). *Advances in Infancy Research*. Standford: Ablex.
- Sansavini, A., Bertoncini, Josiane and Giovanelli, G. (1997). Newborns discriminate the rhythm of multisyllabic stressed words. In *Developmental Psychology*, 33, 3-11.
- Schenker, H. (1906). *Harmony*. (ed. Jones, trans. Borgese from 1906 German ed.). Cambridge: MIT.
- Schellenberg, E. G. & Trehub, Sandra E. (1996). Natural musical intervals. In *Psychological Science*, 7 (5), 272-277.
- Schön D., Magne, C., Schrooten, M. et Besson, Mireille (2002). The music of speech: electrophysiological approach. In Bel and Marlien, *Actes de Speech Prosody*, 635-638.

- Scott, D. (1982). Duration as a cue to the perception of a phrase boundary. In *Journal of Acoustical Society*, 71, 996-1007.
- Selkirk, E. O. (1984). *Phonology and syntax: The relation between sound and structure*. Cambridge : MIT Press.
- Shahidullah, S. and Hepper, P. G. (1992). Hearing in the Fetus: Prenatal Detection of Deafness. In *International Journal of Prenatal and Perinatal Studies*, 4 (3/4), 235-240.
- Smith, Laura, Hains, Sylvia, M.J., Lecanuet Jean-Pierre et Kisilevsky, Barbara S. (2000). Maturation of behaviours elicited by music in 32-37 weeks GA fetuses. Poster, *ICIS 2000* (International Conference on Infant Studies), Brighton.
- Snow, Catherine E. (1977). The development of conversation between mothers and babies. In *Journal of Child Language*, 4, 1-22.
- Snow, Catherine E. and Ferguson, Charles A. (1977) *Talking to Children: Language Input and Acquisition*. Cambridge: Cambridge University Press, 219-235.
- Spencer, M. J. and DeCasper, A. D. (1987). Prenatal experience with low-frequency maternal-voice sounds influence neonatal perception of maternal voice samples. In *Infant Behavior and Development*, 10, 133-142.
- Spring, D. R. and Dale, P. S. (1977). Discrimination of linguistic stress in early infancy. In *Journal of Speech and Hearing Research*, 20, 224-232.
- Stern, Daniel N., Spiker, S. and MacKain, K. (1982). Intonation as signals in maternal speech to pre-linguistic infants. In *Developmental Psychology*, 18, 727-735.
- Stern, Daniel N., Spieker, Susan, Barnett, R. K. and Mac Kain, Kristine (1983). The prosody of maternal speech: infant age and context related changes. In *Journal of Child Language*, 10, 1-15.
- Studdert-Kennedy, Michael G. (1975). From continuous signal to discrete message: Syllable to phoneme. In Kavanagh and Cutting, 113-125.
- Thorpe, Leigh A. (1985). *Auditory-temporal organization: Developmental perspectives*. PhD, Université de Toronto.
- Thorpe, Leigh A., Trehub, Sandra E., Morrongiello, Barbara A. and Bull, Dale (1988). Perceptual grouping by infants and preschool children. In *Developmental Psychology*, 24, 484-491.
- Thorpe, Leigh A. and Trehub, Sandra E. (1989). Duration illusion and auditory grouping in infancy. In *Developmental Psychology*, 25, 122-127.
- Trainor, Laurel (1996). Infant preferences for infant-directed versus noninfant-directed playsongs and lullabies. In *Infant Behavior and Development*, 19, 83-92.
- Trainor, Laurel and Trehub, Sandra (1992). A comparison of infants' and adults' sensitivity to Western musical structure. In *Journal of Experimental Psychology: Human Perception and Performance*, ??.
- Trainor, Laurel and Rock, Adrienne, M.L. (1997). Distinctive messages in Infant-directed lullabies and playsongs. Poster, SRCDC, Washington DC, April.
- Trainor, Laurel, Clark, E.D., Huntley, A. and Adams, B.A. (1997). The acoustic basis of preferences for infant-directed singing. In *Infant Behavior and Development*, 20, 383-396.
- Trainor, Laurel and Zacharias, C.A. (1999). Infants prefer higher-pitched singing. In *Infant Behaviour and Development*.
- Trehub, Sandra E. (1973). Infant's sensitivity to vowel and tonal contrasts. In *Developmental Psychology*, 9, 91-96.
- Trehub, Sandra E. (1989). Infants' perception of musical sequences: Implications for language acquisition. In *Speech-Language Pathology and Audiology*, 13, 3-11.
- Trehub, Sandra E., Bull, Dale and Thorpe, Leigh A. (1984). Infants' perception of melodies: The role of melodic contour. In *Child Development*, 55, 821-830.
- Trehub, Sandra E. and Thorpe, Leigh A. (1989). Infants' perception of rhythm. Categorization of auditory sequences by temporal structure. In *Canadian Journal of Psychology*, 43, 217-229.
- Trehub, Sandra, Thorpe, Leigh A. and Trainor, L. J. (1990). Infants' perception of good and bad melodies. In *Psychomusicology*, 9, 5-19.

- Trehub, Sandra E. and Henderson, J. (1994). Caregivers' songs and their effect on infant listeners. In Deliège (Ed.). Proceedings of the 3rd International Conference for Music Perception and Cognition, Liège, Belgium, ESCOM, 47-48.
- Trehub, Sandra E. and Trainor, Laurel, J. (1994). Les stratégies d'écoute chez le bébé : origines du développement de la musique et de la parole. In McAdams et Bigand, 299-347.
- Trehub, Sandra, Trainor, Laurel, J. and Unyk, A. M. (1993a). Music and speech processing in the first year of life. In *Advances in Child Behavior and Development*, 24, 1-35.
- Trehub, Sandra, Unyk, A. M. and Trainor, Laurel J. (1993b). Material singing in cross cultural perspective. In *Infant Behavior and Development*, 16, 285-295.
- Trehub, Sandra E., Unyk, A. M., Kamanetsky, S.B., Hill, D.S., Trainor, Laurel J., Henderson, J.L. and Saraza, M. (1997). Mothers' and fathers' singing to infants. In *Developmental Psychology*, 33, 500-507.
- Trehub, Sandra E. and Trainor, Laurel, J. (1998). Singing to infants: lullabies and play songs. In Rovee-Collier, Lipsitt and Hayne, 43-77.
- Unyk, A. M., Trehub, Sandra, Trainor, Laurel, J. and Schellenberg, E. G. (1992). Lullabies and simplicity: a cross-cultural perspective. In *Psychology of Music*, ??.
- Vinter, Shirley (1992). *Mise en Place des Eléments Prosodiques dans le Langage Emergent de l'Enfant Sourde : Rôle des Stimulations Acoustiques et des Interactions Sociales*. Thèse de Doctorat, Besançon.
- Vinter, Shirley (1994). *L'Emergence du Langage de l'Enfant Déficier Auditif : des Premiers Sons aux Premiers Mots*. Paris : Masson, 147 p.
- Werker, Janet F. and Tees, R. C. (1984). Phonemic and phonetic factors in adult cross language speech perception. In *Journal of the Acoustic Society of America*, 75, 1866-1878.
- Wertheimer, M. (1923c) Untersuchung zur lehre von der gestalt II. In *Psychologische Forschung*, 4, 301-350.
- Whitaker, Harry A. (1976). Neurobiology of Language. In Caterandte and Friedman.
- Wing, H. D. (1941). *Musical Ability and Appreciation*. PhD Thesis, London University.
- Yeni-Komshian, G., Kavanagh, J. F. and Ferguson, Charles A. (1980). *Child Phonology: Production*, tome I. New-York: Academic Press.
- Yeni-Komshian, G., Kavanagh, J. F. and Ferguson, Charles A. (1980). *Child Phonology: Perception*, tome II. New-York: Academic Press, 255 p.
- Zenatti, Arlette (1969). *Le développement génétique de la perception musicale*. Monographies Françaises de Psychologie. Paris : CNRS (n°17).
- Zenatti, Arlette (1994). *Psychologie de la Musique*. Paris : PUF, 391 p.