

Le modèle musical et ses caractéristiques : Motivation, pulsation et hauteur

John Bispham

INTRODUCTION

De plus en plus souvent, la musique se retrouve au centre des débats consacrés à l'évolution, ainsi que dans la littérature traitant du même sujet (*cf.* Wallin *et al.*, 2000; Morley, 2003; Balter, 2004; McDermott & Hauser, 2005; Cross & Morley, à paraître). Un facteur-clé de l'apparition de ces débats est constitué par l'acceptation de plus en plus répandue de la valeur représentée par le champ de la psychologie de la musique, que ce soit intrinsèquement ou en termes de l'importance qu'elle revêt par rapport à des investigations scientifiques plus générales. Les spécialistes de la musique ont énormément bénéficié des perspectives pluridisciplinaires (*cf.* Clayton *et al.*, 2003; Miell *et al.*, 2005; Peretz & Zatorre, 2005), et la plupart des chercheurs sont actuellement d'accord pour dire que la musique et la psychologie de la musique ouvrent des pistes et offrent des moyens inestimables pour mieux comprendre tout ce qui concerne la cognition et la socialité. Une perspective évolutionnaire de la musique est intéressante en elle-même et par elle-même; elle présente une valeur supplémentaire, celle de fournir un terrain commun à de multiples perspectives. Selon Cross (2003), la grande force de ce point de vue est constituée par le fait qu'en se concentrant sur les rapports phylogénétiques et sur les attributs permettant éventuellement à un gène, à un comportement, un organisme *ou* à une dynamique interpersonnelle ou de groupe d'être fonctionnels, l'évolution offre, et c'est unique, un cadre intégré permettant de comprendre la musique à la fois au niveau biologique et culturel, dans le contexte plus large du comportement humain. En bref, il existe une convergence de plus en plus

grande sur la notion de musique et de musicalité en tant que caractéristiques principales de ce que nous, les *Homo sapiens*, sommes, tant au niveau biologique que culturel.

Toutefois, on ne trouve pas, dans la littérature spécialisée, un modèle de la musique et de la musicalité qui soit à la fois bien clair, aux niveaux psychologique et physiologique, fondé et consensuel. Souvent, il n'est pas aisé de déterminer si les aspects d'un effet observé ou d'une fonctionnalité évolutionnaire supposée sont spécifiquement dus à des caractéristiques « musicales » ou, au contraire, à d'autres facettes de l'engagement en question. Un modèle psychologique et physiologique de la musicalité qui soit biologiquement et culturellement viable constitue, dans cette matière comme dans toutes les investigations pluridisciplinaires qui impliquent la musique, une étape essentielle pour atteindre la clarté et la compréhension. Depuis quelque temps, un argument revient régulièrement, selon lequel ces problèmes de définitions sont de faux problèmes, puisque tout le monde sait ce dont il parle (et donc, par extension, ce dont son interlocuteur parle) et qu'on ferait mieux de consacrer du temps à des recherches empiriques, plus productives (*cf.* Hauser & McDermott, 2006; Mithen, 2006). Toutefois, à la lumière de mes lectures dans ce domaine spécifique, je pense pouvoir dire que, trop souvent, ce n'est pas le cas. Je considère que, loin de soustraire un temps précieux à la recherche empirique, la définition d'un modèle à la fois exact et viable pourrait servir de base à l'inspiration et la formulation d'hypothèses vérifiables par expérimentation et ainsi augmenter grandement l'efficacité et l'applicabilité des investigations empiriques.

L'objectif principal du présent article est de satisfaire une partie de ce besoin en présentant un concept et une discussion des caractéristiques de base de la musicalité considérées ici comme étant spécifiques de l'espèce, spécifiques du contexte de l'engagement musical, et universellement présentes : il s'agit de la motivation musicale, de la pulsation musicale et de la hauteur. En d'autres termes, cette contribution vise à décrire les « caractéristiques spécifiques » évolutionnaires de la musique, c'est-à-dire les caractéristiques psychologiques et comportementales qui la distinguent génériquement d'autres formes de communication humaine et animale et dont on peut donc considérer qu'elles ont évolué spécifiquement pour la musique.

MUSIQUE, COMMUNICATION ET AFFECT

La musique, comme toute autre activité de communication, existe à l'intérieur d'une culture et d'interactions, tout en dépendant d'elles pour les significations et les fonctionnalités qu'elle revêt. Si on fait exception du monde occidental moderne et industrialisé, dans lequel elle est parfois considérée comme un objet physique commercialisable et à valeur ajoutée, la musique constitue, à la base, un comportement social actif et interactif. Toutefois, un paradoxe apparaît, inhérent aux faits largement acceptés que, d'une part, les significations des diverses musiques résultent essentiellement de processus culturels et de contextes sociaux (*cf.* Blacking, 1995 ; Bohlman, 2000), et que, d'autre part, nous semblons néanmoins capables de nous engager par rapport à *certain*s aspects de musiques extérieures à notre propre culture, et/ou d'y réagir (*cf.* Balkwill & Thompson, 1999 ; Krumhansl *et al.*, 2000 ; Nan *et al.*, 2006). Même certaines espèces non humaines réagissent à une série de caractéristiques structurelles génériques de la musique (McDermott & Hauser, 2005). Ce paradoxe disparaît toutefois si nous acceptons la notion d'un engagement musical opérationnel à des niveaux de signification et d'affect (niveaux bien distincts,

mais pouvant néanmoins se chevaucher) biologiques, sociaux et ancrés culturellement (*cf.* Cross, ce volume).

On attribue généralement une fonction spéciale de la musique à son association avec l'« émotion ». En dépit de toute une série de preuves anecdotiques selon lesquelles l'émotion musicale est en quelque sorte « unique » ou « spéciale », la littérature spécialisée n'a jusqu'à présent pas été capable de déterminer les raisons de cet état de choses. Les preuves d'un potentiel d'expériences émotionnelles particulièrement fortes dans un contexte musical (*cf.* Gabrielsson, 2001) sont constituées par les évocations d'une « expérience très forte » (peak experience) (Maslow, 1968), d'un « flux » (*cf.* Csikszentmihalyi, 1990) ou encore de « frissons » et de « tremblements » (chills) (Panksepp, 1995). Toutefois, ces expériences sont relativement rares et ne constituent pas forcément la meilleure base pour une perspective comparative générique des émotions musicales. Des « affects de vitalité » dynamiques (*cf.* Stern, 1985) ou la création intrinsèque de tensions et d'attentes (*cf.* Meyer, 1956 ; Steinbeis *et al.*, 2005) peuvent engendrer certaines réactions émotionnelles spécifiques à la musique. Cependant, il n'existe pas, à ma connaissance, de preuves solides selon lesquelles les émotions musicales se différencient d'une quelconque manière d'autres formes d'émotions. En termes d'expression de l'émotion, une méta-analyse d'études (Juslin & Laukka, 2003 ; *cf.* aussi Thompson & Balkwill, 2006) confirme que l'« expression » vocale et musicale fait intervenir des indices structurels similaires tels la F_0 générale, la hauteur du son, le rythme et l'intensité.

Alors que la majorité des études consacrées au sujet se sont focalisées sur l'expression de l'émotion en musique, l'usage que la société fait peut-être le plus souvent de la musique est celui d'un moyen de modifier l'humeur d'un individu ou d'un groupe (Bailey & Davidson, 2005 ; DeNora, 2001 ; Gomart & Hennion, 1999). Quoique aucune théorie clairement définie ne vienne la soutenir, la capacité que possède la musique de fonctionner à ce niveau est à

ce point reconnue, dans plusieurs domaines relevant de la psychologie, que de nombreux paradigmes expérimentaux ont régulièrement recours à la musique en tant que moyen d'induire des humeurs particulières chez les participants à l'expérience concernée, simplement parce que cela fonctionne (*cf.* Albersnagel, 1988 ; Gerards-Hesse *et al.*, 1994). Il est à noter que, bien que d'autres formes de communication puissent également revêtir cette fonctionnalité, la musique semble être à ce niveau d'une efficacité particulière (DeNora, 2001). Une possibilité est que l'attention soutenue et régulière au matériel affectif fourni par la musique soit particulièrement apte à remettre à zéro des états affectifs de base (c'est-à-dire les humeurs, *cf.* Davidson, 1994). De nouveau, la recherche s'est concentrée sur l'expérience de l'auditeur (d'où le terme « induire », *cf. supra*). Cependant, en élargissant la pertinence aux contextes interactifs génériques, il semble correct de dire que, dans l'engagement musical, chaque participant régule les états affectifs des autres participants, et que cette interaction constitue une composante essentielle de la motivation d'un individu à s'engager par rapport à la musique (Swaine, communication personnelle).

LA MOTIVATION MUSICALE

On peut globalement décrire la motivation comme « une influence modulatrice et coordinatrice sur la direction, la vigueur et la composition d'un comportement. Cette influence provient d'une large variété de sources internes, environnementales et sociales, et se manifeste à de nombreux niveaux de l'organisation comportementale et neurale » (Shizgal, 1999, p. 566). Chez les musicologues, les discussions génériques consacrées à la motivation ne semblent pas très nombreuses. C'est seulement dans la littérature neurophysiologique consacrée à la musicothérapie que nous trouvons une référence au besoin d'un modèle de motivation pour la musique (Unkefer & Thaut, 2002). Toutefois, même dans ce cas, on ne dispose pas encore d'un modèle explicite. Le fait de ne pas prendre en compte les

facteurs de motivation dans l'engagement par rapport à la musique exerce de puissantes implications à de nombreux niveaux. Tout d'abord, on ne peut comprendre totalement la nature affective de la musique sans au moins disposer d'un modèle conceptuel des forces motivationnelles qui engendrent et soutiennent l'engagement en question (Thaut, 2002). Ensuite, les perspectives comparatives sont forcément incomplètes si on ne considère pas la motivation. Il est d'une importance cruciale de bien considérer que la signification fonctionnelle/adaptative d'un ensemble de comportements donné dépend au moins autant des forces biologiques et/ou culturelles qui motivent une espèce à les adopter et à les maintenir que des capacités qui les soutiennent.

La littérature psychologique consacrée à la motivation en identifie deux formes principales : l'intrinsèque et l'extrinsèque. Dans Ryan & Deci (2000), on trouve cette définition de la motivation intrinsèque : « il s'agit de l'inclination naturelle à l'assimilation, la maîtrise, l'intérêt spontané et l'exploration, inclination essentielle pour le développement cognitif et social et qui représente une source principale de jouissance et de vitalité tout au long de la vie (Csikszentmihalyi & Rathunde, 1993) ». C'est particulièrement évident chez les jeunes enfants en bonne santé : même en l'absence de gratification, ils sont curieux de tout et jouent tout le temps (Harter, 1978) ; c'est encore plus évident lorsque les enfants sont attachés de manière sûre à un parent (Bowlby, 1979). La motivation intrinsèque est « de plus en plus limitée par les pressions sociales à la réalisation d'activités qui ne sont pas intéressantes, et à la prise en charge de toute une série de responsabilités nouvelles » (Ryan & LaGuardia, 2001, p. 71). Par contre, l'expression « motivation extrinsèque » fait référence au fait de pratiquer une activité afin d'atteindre un résultat séparé qui, à son tour, peut être considéré comme une recherche de la gratification (acceptation) ou une motivation à éluder une punition (évitement) (Norman & Shallice, 1980 ; Elliott & Covington, 2001). Bien sûr, faire de la musique peut comporter,

à certains moments, une motivation extrinsèque à sa base (*cf.* dans le cas d'une audition, ou lorsqu'un artiste recherche la gloire, *etc.*). Cependant, la grande majorité des participations à des activités musicales de par le monde ont un caractère de « jeu » et sont socialement exploratoires (*cf.* Cross, 1999). En musique, les valences négatives (et, logiquement, la diminution de motivation qui s'ensuit) sont presque toujours le résultat de pressions extrinsèques (Stephoe, 2001); les personnes à qui on pose la question rapportent avec une belle constance que leurs moments musicaux les plus agréables et les plus importants, elles les ont vécus lorsqu'elles sentaient qu'elles pouvaient « se laisser aller ».

Les caractéristiques spécifiques de la musicalité, et j'y reviendrai en détail plus loin, fournissent un cadre temporel et/ou basé sur les hauteurs aux interactions interpersonnelles synchrones. En tant que tel, l'engagement musical implique la conception et l'entretien d'intentions partagées (*cf.* Miell *et al.*, 2005). Tomasello et ses collègues (2005) soutiennent, de manière très globale, que l'intentionnalité conjointe constitue une caractéristique spécifique et définissante de la psychologie humaine; ils décrivent les actions intentionnelles partagées comme des objectifs partagés et des plans d'action coordonnés. De plus, ils considèrent la régulation des actions intentionnelles en termes de principes de contrôle cybernétique servant à ajuster les actions afin de réduire le fossé existant entre, d'une part, les réalités des personnes qui interagissent et, d'autre part, leurs objectifs.

Si on considère les diverses facettes du phénomène et le fait que la musique est opérationnelle dans la stimulation et la régulation de l'affect (*cf. supra*), il semble probable que l'intersubjectivité¹ constitue un objectif premier de l'interaction musicale² et que cette interaction, dans un cadre mutuel et synchronisé, motive les musiciens à partager des

états psychologiques convergents. Pour Swaine (à paraître), le glissement d'une motivation extrinsèque vers une motivation intrinsèque de partager des états psychologiques, ainsi que la génération de correspondances entre l'état réel des personnes en interaction et l'état final intersubjectif souhaité dans l'engagement musical vocal servent à fixer l'attention sur les actions associées à l'objectif proximal d'intersubjectivité. Selon lui, c'est par cela que l'engagement musical promeut l'affect positif (*cf.* Bailey & Davidson, 2005). Il faut noter que, bien qu'elle ne soit pas encore établie en psychologie de la musique, cette idée que la musique est issue d'un objectif d'intersubjectivité et de régulation de l'affect apparaît dans la littérature consacrée au développement de l'enfant ainsi que dans les applications thérapeutiques du domaine concerné. Dans ces engagements proto-musicaux et musicaux (le « musicking », comme Small l'appelle, 1998), la co-régulation de l'affect et des états motivationnels, ainsi que la maturation, le développement et la réparation de l'intersubjectivité sont considérés comme inextricablement liés (Papoušek, 1996; Bunt & Pavlicevik, 2001; Trevarthen & Aitken, 2001; Beebe *et al.*, 2005).

PULSATION ET HAUTEUR

Une bonne partie de la difficulté de définir la musique et la musicalité, que ce soit au niveau psychologique ou en termes de corrélats physiques, est constituée par le fait que ses manifestations culturelles, ses rationalisations et ses contextualisations sont éminemment variables (Nettl, 2000). Elles sont même tellement variables que la comparaison directe de deux genres différents est susceptible de révéler fort peu de recouvrements descriptibles; pourtant, on reconnaît bien ces deux genres comme de la « musique ». Par exemple, essayons de

1. Pour Stern (1985), le terme « intersubjectivité » couvre à la fois l'inter-attentionnalité (attention conjointe), l'inter-intentionnalité (intention partagée) et l'accordage affectif.

2. Comme Swaine le note (communication personnelle), ceci survient très probablement à l'intérieur d'une hiérarchie d'objectifs.

comparer les chants monocordes tibétains aux Taiko japonais ou, encore plus difficile, essayons d'expliquer à des Martiens pourquoi nous considérons ces deux pratiques comme de la musique, alors que nous considérons les interactions mère-enfant comme «proto-musicales» et que, de manière générale, nous ne considérons même pas le pant hoot du chimpanzé comme de la musique. Une approche du problème des caractéristiques universelles en musique a été jusqu'à présent de se concentrer sur les prédispositions de l'enfant à la traiter (Trehub, 2000; Trehub & Hannon, 2006). Selon Trehub, «les caractéristiques universelles du traitement du schéma musical présentent des parallèles frappants avec les caractéristiques universelles ou quasi universelles de la structure musicale» (p. 427). C'est pour cela qu'il semble fort probable de pouvoir identifier les caractéristiques universelles basiques qui constituent les fondements d'une capacité inhérente d'interaction musicale sur lesquels la variation culturelle se construit et/ou auxquels elle est restreinte.

On considère généralement que les caractéristiques fondamentales et distinctives (aux niveaux descriptif et neurophysiologique, cf. Peretz & Zatorre, 2005) de la musique sont constituées par une organisation temporelle de l'action (c'est-à-dire la pulsation et le rythme) et une organisation basée sur les fréquences (c'est-à-dire la mélodie et l'harmonie). Afin d'incarner complètement une capacité basique inhérente et observable dans toutes les musiques, nous devons néanmoins considérer le fait qu'elles n'ont pas toutes recours à une pulsation temporelle observable, que certaines d'entre elles ne se basent absolument pas sur les hauteurs, et que le recours musical à des hauteurs n'implique pas forcément une mélodie, une harmonie ou l'utilisation de gammes (c'est, par exemple, le cas du chant monocorde). En gardant cela présent à l'esprit, les caractéristiques spécifiques de la musicalité peuvent être décrites comme une capacité spécifique de l'espèce et de son contexte de s'engager musicalement dans des configurations de pulsation et/ou de hauteurs (décrites plus loin). Que l'un ou l'autre de ces

deux composants soit absent ou pas, la pulsation et la hauteur semblent fournir un cadre à l'interaction interpersonnelle synchrone – une sorte de stratégie de coordination.

LA PULSATION

Dans un article récent (Bispham, 2006b), j'ai présenté une perspective comparative inter- et intra-espèces du comportement rythmique musical, accompagnée de ses fondements psychologiques et physiologiques. L'objectif principal que je poursuivais était d'identifier les caractéristiques uniques du rythme musical, celles qui le définissent. Dans la section qui va suivre, on trouvera un résumé des points principaux développés dans cet article, mais le lecteur intéressé se référera au document original et à un autre article, plus récent (Cross *et al.*, 2008) pour de plus nombreux détails et une discussion sur la manière dont tout cela peut influencer les perspectives évolutionnaires générales relatives au rythme et à la synchronisation.

La caractéristique principale du comportement rythmique musical (caractéristique dont on peut dire qu'elle le définit) est constituée par le fait que la structuration temporelle des actions se base sur une pulsation contenant des éléments réguliers permettant à des individus, à l'intérieur des limites temporelles d'un présent psychologique (Fraisse, 1984; Clarke, 1999), d'interagir en temps réel par un processus de synchronisation (entraînement, Clayton *et al.*, 2003). La synchronisation survient à partir du moment où les individus utilisent les régularités du signal afin de prédire et de diriger leurs ressources attentionnelles sur le timing d'événements futurs (Jones, 1976); elle peut être considérée comme une composante essentielle de tout engagement écologique inter taxons structuré temporellement. On a également avancé que la synchronisation est un mécanisme qui fournit à l'interaction mère-enfant et aux interactions linguistiques un cadre temporel mutuellement manifeste et basé sur la pulsation (Webb, 1972; Jaffe &

Anderson, 1979 ; Auer *et al.*, 1999). En matière de développement humain, on a établi que même le nouveau-né démontre une capacité de réagir aux mouvements et aux sons produits par un adulte s'occupant de lui (Trevarthen, 1999). Néanmoins, il apparaît que certaines caractéristiques psychologiques et/ou physiologiques uniques de la synchronisation musicale suivent des formes non musicales ontogénétiquement et peut-être également phylogénétiquement. Selon Bahrick & Lickliter (2004), les enfants aux alentours de cinq mois ont besoin de signaux multimodaux afin de détecter les modifications rythmiques, tandis qu'à l'âge de huit mois ils sont capables de recourir à de simples modalités sensorielles pour exécuter ce type de tâches. Phillips-Silver & Trainor (2005) ont bien montré qu'à moins que des enfants âgés de sept mois ne soient bougés en respectant la métrique d'une œuvre musicale, ils sont incapables de faire la différence entre divers groupements métriques. Il semble que la capacité de se synchroniser par rapport à un matériel musical n'apparaisse pas totalement avant le premier anniversaire (Hannon & Johnson, 2005 ; McAuley *et al.*, 2006).

Comme expliqué dans Bispham (2006b), il existe deux manières opposées d'interpréter la différence entre, d'une part, le type de pulsation opérationnel en musique et, d'autre part, celui qui est présent dans d'autres formes d'interaction. Tout d'abord, on peut dire que toute interaction humaine se compose de formes de synchronisations interpersonnelles dont la gamme part a) d'une utilisation relâchée, inconsciente de la pulsation qui forme le cadre d'interactions interpersonnelles pouvant changer de direction : c'est le cas, par exemple, des interactions linguistiques ou des interactions mère-enfant (Wilson & Wilson, 2005), pour en arriver b) au fait de «coller» strictement à la pulsation (groove) dans les comportements musicaux de groupe et à la synchronicité de l'output : les participants sont conscients du cadre formé par la pulsation et souhaitent maintenir un certain degré de stabilité temporelle et de coordination dans le groupe (*cf.* la musique et la danse). Une autre possibilité est

que l'apparence d'une pulsation à l'intérieur d'une interaction non musicale ne dépende pas de mécanismes de synchronisation similaires à ceux utilisés en musique mais soit plutôt le résultat de l'organisation d'actions en relation avec des pulsations et des attentes de courte durée, interrompues perpétuellement, qui se basent sur de l'expérience et des indices temporels. Quelle que soit la version correcte, je soutiens qu'en matière musicale, il existe à tout le moins des caractéristiques du rythme qu'on peut décrire comme contextuellement et mécaniquement distinctes et dont, dès lors, on ne peut expliquer qu'elles ont évolué exclusivement par rapport à des comportements non musicaux.

Au contraire d'interactions non musicales et de comportements supposés analogues dans d'autres espèces, la pulsation musicale se maintient à travers le temps, implique la conscience d'un cadre basé sur la pulsation (Repp, 2001) et est perçue sans ambiguïté ou à des niveaux hiérarchiques associés (London, 2004) par des individus ayant adopté les schémas comportementaux d'une culture donnée (Stobart & Cross, 2000). De plus, la réaction à une pulsation musicale implique une telle activation du système moteur qu'elle rend l'individu capable de gérer un contrôle temporel (fin comme grossier) dans ses mouvements balistiques ou en douceur (Thaut *et al.*, 1997). Ceci semble suggérer l'implication de mécanismes oscillatoires internes périodiques en chevauchement avec la coordination motrice (Bispham, 2003). De plus, le fait d'atteindre l'objectif désiré de maintenir un cadre temporel interactif «en temps réel» ainsi qu'une synchronisation sensorimotrice implique la mise en œuvre de mécanismes de correction qui se basent sur les outputs propres ou ceux des autres. Ils sont essentiels pour expliquer les écarts au niveau moteur (Wing & Kristofferson, 1973) ainsi que les modulations de tempo et de microtimings expressives et créées structurellement, qu'elles soient conscientes ou subconscientes (Palmer, 1997 ; Collier & Collier, 2002 ; Iyer, 2002). Des études récentes faisant intervenir l'imagerie neurologique (Stephan *et al.*, 2002), de même qu'une abondante littérature

psychologique, identifient deux mécanismes de correction qui interagissent et dont il est généralement accepté qu'ils opèrent indépendamment l'un de l'autre en musique : il s'agit de la correction de phase et de la correction de période (Repp, 2005)³. Au contraire de la correction de période, les mécanismes de correction de phase sont très vraisemblablement communs à toutes les activités dans lesquelles l'attention est tournée vers le futur et dans lesquelles les attentes sont continuellement ajustées afin de tenir compte des divergences existant entre les pulsations attentionnelles séparées et les événements-stimuli. Toutefois, la correction de période est, presque par définition, spécifiquement fonctionnelle dans le cadre d'une pulsation musicale soutenue. Selon Repp (2004), il est probable que la correction de période soit une capacité spécifiquement humaine et qu'elle constitue la manifestation de la capacité plus générale de l'homme de battre le tempo d'une activité rythmique. Les mécanismes de correction de période présentent une dernière caractéristique d'importance cruciale : de nouveau, en opposition avec les mécanismes de correction de phase, les premiers semblent impliquer la prise de conscience (Repp, 2001) et sont affectés par les manipulations de l'intention, de l'attention et de la prise de conscience (Repp & Keller, 2004).

LA HAUTEUR

En matière de musique, certaines caractéristiques universelles relatives à l'organisation des hauteurs ont été proposées (Justus & Husler, 2005 ; McDermott & Hauser, 2005) ; parmi elles, la présence

d'intervalles répartis de manière inégale (Burns & Ward, 1999), une présence majoritaire d'intervalles de faibles valeurs entières (et une préférence pour ce type d'intervalles⁴) (Dowling & Harwood, 1986 ; Schellenberg & Trehub, 1996) et l'existence de hiérarchies tonales (Castellano *et al.*, 1984 ; Krumhansl, 1990). Je suis d'accord pour dire que ces caractéristiques sont presque des caractéristiques universelles. Toutefois, les gammes et les hiérarchies tonales sont toutes deux construites culturellement et ne constituent pas des prérequis absolus pour qu'une interaction basée sur les hauteurs soit qualifiée de « musicale ». Afin d'incarner l'ensemble des types de hauteurs en musique (en y reprenant, par exemple, le chant monocorde) tout en envisageant la question en termes de sa qualité de capacité émergente de musicalité, il convient de creuser à un niveau encore plus profond. J'avance l'idée que le système des hauteurs, quelle qu'en soit la forme, se construit sur base d'une capacité primaire de produire et de recourir à une fréquence fondamentale stable et soutenue, ainsi que sur la capacité de créer ou de traiter certains rapports entre les hauteurs. Je pars de l'hypothèse selon laquelle, partout dans le monde, la structure des hauteurs (qu'elle soit monocorde ou très ornée) peut se caractériser par son organisation relationnelle par rapport à des zones tonales soutenues, mais néanmoins variables (McAllester, 1971). Ceci peut éventuellement soulever la controverse, mais n'implique (et c'est important) aucune forme de tonalité structurée ; cette théorie suggère simplement que les hauteurs présentent à tout moment (ou dans les limites d'une phrase donnée) une organisation relationnelle dans le cadre d'une région (ou de régions) de hauteur dominante⁵.

3. La correction de phase permet d'ajuster les asynchronies existant entre la dernière réaction et les événements stimuli qui ne génèrent pas de modification de période, tandis que la correction de période modifie l'intervalle-cible suivant sur base des différences existant entre l'intervalle de référence et le dernier (ou les quelques derniers) intervalle(s) inter-stimuli, ce qui change la période de la pulsation musicale attentionnelle.

4. Cette idée de « préférence » pose un problème dans la mesure où « consonance » et « dissonance » sont des notions relatives, qui varient en fonction des styles musicaux. Ce qui est génériquement important, c'est l'interaction entre la « consonance » et la « dissonance », ainsi que la création de « tensions » et de « résolutions ».

5. Cette idée est quelque peu renforcée par la suggestion selon laquelle l'une ou l'autre forme de drone est presque universellement présente dans les divers répertoires culturels (Mâche, 2000). Comme c'est le cas dans de nombreuses

La production d'une fréquence fondamentale soutenue, ainsi que celle d'autres formes plus complexes de hauteurs s'appuie sur la capacité de produire des sons de puissance, hauteur (F_0) et timbre pouvant varier de manière indépendante. Ceci peut paraître évident au premier abord, mais il semble que cette capacité (en tout cas en ce qui concerne la production vocale) n'est pas présente chez d'autres primates qui témoignent, au niveau de la vocalisation, d'un rapport linéaire étroit entre F_0 et intensité avec modifications de la pression sous-glottique (Demolin, 2006). La capacité de maintenir une F_0 stable et de faire varier de manière indépendante les corrélats acoustiques de la vocalisation apparaît au fur et à mesure du développement de l'enfant (Wermke & Mende, 2008) et s'achève par le contrôle indépendant des nombreux muscles qui entrent en jeu dans la production vocale : les muscles adducteurs et abducteurs de la glotte servent à ouvrir et fermer les plis vocaux qui régulent les niveaux de pression sous-glottique et influencent par là la F_0 et l'intensité ; les muscles constricteurs du pharynx contrôlent la largeur du pharynx, modifiant en cela le rapport entre la F_0 et les harmoniques les plus élevées (c'est-à-dire les corrélats de timbre) ; les muscles extrinsèques du larynx contrôlent le positionnement vertical du larynx dans la gorge ; les muscles intrinsèques du larynx modulent, quant à eux, la longueur, la rigidité et l'épaisseur des plis vocaux par la contraction des muscles cricothyroïdes et thyroaryténoïdes, et ils permettent, ce qui est crucial, un découplage de l'intensité et de F_0 (Sundberg, 1987 ; Titze, 1994). Il est vraisemblable que le contrôle volontaire de ces muscles (et donc la capacité résultante de faire varier de manière indépendante la F_0 et l'intensité) soit unique, ou particulièrement développé, chez l'être humain. Il est intéressant de noter (mais il n'y a eu jusqu'à présent que peu d'études comparatives à ce sujet) que les muscles thyroaryténoïdes de l'être humain sont les seuls à présenter des fibres

musculaires à faible tonicité qui, à l'opposé de la plupart des muscles, ne se contractent pas et soutiennent des contractions prolongées en y résistant de manière stable et contrôlée précisément (Han *et al.*, 1999). Le contrôle de la respiration constitue un autre point d'importance capitale, et il est significatif qu'on ait évoqué l'augmentation de l'innervation du thorax au cours de l'évolution des hominidés afin de démontrer le meilleur contrôle fin de la respiration chez les humains modernes et les Néandertaliens (Maclarnon & Hewitt, 2004).

Bien que le paragraphe précédent se soit centré sur la production vocale, on peut généraliser les principes présentés en avançant la notion selon laquelle toute activité musicale implique un « chant intérieur » (inner singing) (Janata, 2001 ; Kalakoski, 2001). Ce n'est qu'une supposition, mais elle est appuyée par une étude neurologique récente qui a montré que les zones de l'opercule rolandique (en particulier celles susceptibles de représenter une articulation du larynx et du pharynx) entraînent en activité lors du traitement émotionnel de stimuli musicaux instrumentaux (Koelsch *et al.*, 2006). Elle est également cohérente par rapport aux écoles (*cf.* Kodály) selon lesquelles la voix humaine constitue l'instrument de musique naturel. Une hypothèse réaliste consiste à dire que le découplage de F_0 et de l'intensité/du volume est, de manière générique, plus prononcé dans la musique que dans la parole. Ce qui expliquerait peut-être certaines conclusions comparatives interpellantes selon lesquelles les modifications de hauteur exercent des effets opposés sur la valence dans la musique et la parole et, chez les sujets affectés, sur la mesure de l'éveil/énergie uniquement pour la parole (Ilie & Thompson, 2006). Selon Gussenhoven (2002), il se pourrait que ce découplage différencie radicalement le « code de production » (aux racines biologiques) de la musique et celui de la parole et de la communication vocale considérée de manière plus générale.

musiques africaines, dans lesquelles la pulsation est totalement implicite, et ressentie par des individus partageant la même culture sans qu'elle soit explicite au niveau auditif, il est possible que les zones dominantes de hauteurs deviennent simplement plus explicites lorsque des drones entrent en jeu.

Un centre de hauteur soutenu et stable est en quelque sorte similaire à une pulsation soutenue en ce qu'il fournit un cadre mutuellement manifeste à l'interaction. On peut également faire l'analogie avec le mariage des hauteurs, car celui-ci implique nécessairement des mécanismes de correction basés sur la production propre ou d'autrui et sur le désir d'atteindre et de maintenir certaines relations entre les fréquences. Même s'il est vrai que les spécialistes ne s'accordent pas sur la question de savoir si c'est le traitement absolu ou relationnel qui représente l'état inné (Saffron & Griegentrog, 2001 ; Trehub, 2003 ; Platinga & Trainor, 2005), il semble à peu près certain que le traitement relatif de la hauteur représente partiellement une capacité émergente de musicalité. Si on fait exception d'un exemple bien connu et très inhabituel (Will, 1997), les systèmes de hauteurs qui existent dans les diverses cultures sont construits sur base de connexions de fréquences relatives plutôt qu'absolues, et compris en ces termes⁶. Par contre, McDermott & Hauser (2005) établissent qu'à l'exception d'une étude qui mentionne les équivalences d'octaves chez les macaques Rhésus (Wright *et al.*, 2000), il est régulièrement démontré que d'autres espèces, de manière très naturelle, encodent les stimuli musicaux soit en termes de hauteur absolue, soit en termes de leurs contenus en fréquences absolues.

À l'heure actuelle, on ne comprend pas encore très bien les mécanismes de correction qui entrent en jeu dans le traitement des hauteurs. Toutefois, sur base d'enquêtes réalisées en recourant au paradigme de feedback vocal avec glissement des fréquences⁷ sur le contrôle auditif de la F_0 de la voix (Donath *et al.*, 2002), nous pouvons commencer à formuler quelques suppositions. Dans les études qui font intervenir une vocalisation continue, il existe

quelques preuves selon lesquelles les réactions opposées surviennent au terme d'une latence de courte durée (100-150 ms) (Burnet *et al.*, 1997), tandis que les réactions suivantes surviennent au terme d'une latence plus grande (250-600 ms) (Larson *et al.*, 1998)⁸. Hain *et al.* (2000) confirment la présence des deux types de réactions : ils ont découvert que la direction de la deuxième réaction (mais pas de la première) peut être modifiée par des instructions. Selon Donath *et al.* (2002), « la première réaction indique un système de feedback négatif qui stabilise automatiquement la F_0 de la voix, tandis qu'il se pourrait que la deuxième réaction reflète un mécanisme volontaire qui ajuste la F_0 de la voix afin qu'elle corresponde à une (supposée) référence externe » (p. 1587). Il est intéressant de constater que les études consacrées à la parole (Natke & Kalveram, 2001 ; Donath *et al.*, 2002) n'ont fait apparaître que des réactions opposées survenant avec une latence d'à peu près 160 ms due au feedback auditif avec glissement de fréquence. Ceci indique que, dans le cas de la parole, seules les premières réactions involontaires surviennent, et suggère (en tenant compte des autres preuves) qu'un mécanisme spécifique à la musique et contrôlé par la volonté (partiellement analogue aux corrections de période en matière de pulsation), est à l'œuvre dans le traitement des hauteurs. Il conviendra de tester cela empiriquement dans le cadre d'un paradigme musical d'accordage fin et de correspondance de l'output par rapport à des références externes.

FONCTIONNALITÉ DE LA MUSIQUE

La fonctionnalité proximale des caractéristiques supposées décrites précédemment semble assez claire. La motivation musicale, la pulsation et la

6. Contrairement à ce que Mithen (2005) soutient, nous ne croyons absolument pas que ce soit l'acquisition ou le maintien d'une hauteur absolue qui rende l'homme ou le Néandertalien « plus apte à la musique » (voir Bispham, 2006a).

7. Dans ce paradigme, on demande aux sujets de porter des écouteurs et d'écouter leur propre production vocale en temps réel. En général, les chercheurs introduisent des perturbations dans les fréquences entendues (ou au niveau d'autres caractéristiques acoustiques), et étudient les effets que ces modifications amènent dans la production.

8. Une réaction opposée est une réaction dans laquelle la F_0 produite par le sujet va dans le sens opposé de celle de la manipulation subie par le feedback auditif. Les réactions suivantes dénotent de manière évidente le contraire.

hauteur constituent un cadre de coordination pour l'interaction affective interpersonnelle. L'être humain présente un besoin inné d'interactions sociales (Baumeister & Leary, 1995); les capacités basiques perceptuelles et de production décrites permettent une organisation du temps et de la hauteur qui peut faire de la musique un moyen particulièrement efficace pour faciliter ces interactions. À l'opposé d'autres cadres interactifs opérationnels, par exemple dans l'interaction linguistique, la pulsation et la hauteur musicales permettent à des individus de partager un cadre plus ou moins commun et synchronisé, et favorisent les interactions affectives groupales. Conformément à cette caractéristique potentiellement définissante, on peut globalement considérer que la musique contribue à atteindre la confluence socio-affective (Graham, 2007; Bisham, 2007). Considérer de cette manière la fonctionnalité de la musique paraît sensé, étant donné le recours universel à la musique dans les rituels cérémoniels : le souhait d'atteindre, par le biais de l'engagement social, des états de motivation convergents dans de grands groupes y semble particulièrement affirmé (*cf.* Rappaport, 1999).

Il faut toutefois clarifier un point important : bien qu'il soit relativement prudent de dire que les deux choses sont en rapport, la fonctionnalité proximale d'une activité ne reflète pas forcément son importance en termes d'évolution. Bien qu'on le présente intuitivement, la manière dont la convergence d'états sociaux et affectifs (tant dans la musique que dans le rituel) peut exercer un impact sur la valeur d'adaptation évolutive de l'individu ou du groupe n'apparaît pas clairement. La littérature spécialisée regorge de modèles possibles relatifs à l'adaptation évolutive des capacités musicales. On peut schématiquement les catégoriser en : attrac-

tion du partenaire (*cf.* Darwin, 1871; Miller, 2000; Merker, 2000); signalement de coalition (*cf.* Hagen & Bryant, 2003; Hagen & Hammerstein, ce volume); altricialité et interaction mère-enfant (*cf.* Dissanayake, 2000; Falk, 2004); et enfin, cohésion du groupe (*cf.* McNeill, 1995; Roederer, 1984; Brown, 2000a; Brown, 2000b; Freeman, 2000; Cross, ce volume). Chacune de ces perspectives compte de nombreux partisans, et il semble correct de dire que la musique englobe, et a englobé, une multitude de fonctions.

Cet article n'a pas pour objectif d'entrer dans le débat autour des importances relatives qui devraient être attribuées aux divers scénarios envisagés par la littérature spécialisée⁹. Toutefois, selon moi, il faut que les débats en la matière partent de la compréhension des spécificités de l'interaction et de l'engagement musical. Ce sont ces spécificités qui doivent nourrir le débat. À l'instar de toute capacité complexe, la musique constitue le résultat de toute une série d'adaptations et/ou exaptations évolutives combinées (Foley, 2004). Des capacités musicales génériques et complètes comprennent des compétences et des mécanismes dont on pense qu'ils étaient présents chez les premiers vertébrés dotés de mâchoires il y a quelque 500 millions d'années (Chase, 2001). Il est donc forcément insuffisant de discuter de la fonctionnalité adaptative de la musique ou de son statut en tant qu'adaptation, exaptation ou conséquence secondaire sans définir au départ le sujet du débat et ses caractéristiques potentiellement uniques. Nous courons également le risque constant d'attribuer erronément certaines fonctionnalités à la « musique », fonctionnalités qu'on pourrait plus précisément attribuer à des catégories de pertinence plus générales.

Traduction Stéphane Renard

9. Il vaut toutefois la peine de noter que, dans la mesure où les caractéristiques décrites permettent l'interaction dans un groupe, l'approche prônée ici va plutôt dans le sens des hypothèses qui décrivent des processus actifs au niveau du groupe.