

Déméter

ISSN : 1638-556X

Publisher : Université de Lille

12 | Été | 2024

Du réparable et de l'irréparable : les pratiques artistiques confrontées aux limites de la réparation

L'interaction musicale révèle la musique comme langage incarné

Alessandro DELL'ANNA, Marc LEMAN and Annamaria BERTI

Translated by David CAULET, Félix EMERY, Joséphine GÉDÉON-GONÇALVES, Fanny HERMANT, Julie LABEQUE, Matilde NORTIER, Paul VENNIN and Christian HAUER

 <https://www.peren-revues.fr/demeter/1803>

DOI : 10.54563/demeter.1803

Electronic reference

Alessandro DELL'ANNA, Marc LEMAN and Annamaria BERTI, « L'interaction musicale révèle la musique comme langage incarné », *Déméter* [Online], 12 | Été | 2024, Online since 29 janvier 2025, connection on 03 février 2025. URL : <https://www.peren-revues.fr/demeter/1803>

Copyright

CC-BY-NC

L'interaction musicale révèle la musique comme langage incarné

Alessandro DELL'ANNA, Marc LEMAN and Annamaria BERTI

Translated by David CAULET, Félix EMERY, Joséphine GÉDÉON-GONÇALVES,
Fanny HERMANT, Julie LABEQUE, Matilde NORTIER, Paul VENNIN and
Christian HAUER

OUTLINE

Première partie : la musique comme langage incarné : un cadre conceptuel

Action conjointe

Cognition incarnée : sa composante « étendue »

Le codage prédictif : sa composante « sensorimotrice »

Cognition musicale incarnée

Les bases sensorimotrices neurales de l'interaction musicale

Un cadre pour la musique en tant que langage incarné

Deuxième partie : données expérimentales

Le temps : entraînement et incarnation dans une interaction de tapping

Espace : reconfiguration de l'espace péripersonnel dans une interaction
en jazz

Qualité : timing expressif et action concertée dans l'interaction à l'œuvre
lors de l'exécution d'un hoquet

Application du cadre théorique

Conclusion

EDITOR'S NOTES

Traduction de l'article de Alessandro Dell'Anna, Marc Leman et Annamaria Berti,
« Musical Interaction Reveals Music as Embodied Language », *Frontiers in
Neuroscience*, vol. 15 (2021). DOI : <https://doi.org/10.3389/fnins.2021.667838>
[consulté le 21 janvier 2025]

Traduction de David Caulet, Félix Emery, Joséphine Gédéon-Gonçalves, Fanny
Hermant, Julie Labeque, Matilde Nortier, Paul Vennin et Christian Hauer, dans le
cadre d'un séminaire du Master Arts de l'Université de Lille, année universitaire
2022-2023.

Cette traduction est publiée avec l'autorisation de la revue et des auteurs.

TEXT

« ... la domanda se la musica sia o non-sia un linguaggio è una domanda mal posta alla quale non ha senso dare una risposta affermativa o negativa ; mentre potremmo trovare interessante considerare la musica alla luce della molteplicità di aspetti presenti nell'analogia in essa suggerita. »

Giovanni Piana, *Filosofia della Musica*, Milano, Edizioni Angelo Guerini & Associati, 1991.

- 1 Selon Darwin, les êtres vivants interagissent avec leur écosystème pour créer un environnement qui permet leur propre répliation ou (dans le langage post-darwinien) la répliation de leur génotype (en opposition à leur phénotype). D'un autre côté, les êtres humains interagissent de manière si complexe que leur environnement, plus que celui des autres animaux, transcende leurs impératifs génétiques au profit des impératifs liés à celui-ci (Denett, 2017). La musique, comme le langage, peut être exactement positionnée entre ces deux impératifs, dans la mesure où elle est constituée d'une composante biologique partagée par tous les êtres humains à travers le monde, ainsi que d'une composante culturelle différenciant les communautés humaines réparties sur la Terre. Les deux composantes sont généralement appelées « musicalité » et « musique » :

La musicalité peut être définie dans toute sa complexité par un ensemble de caractères naturels et au développement spontané, fondé sur notre système cognitif et biologique et contraint par celui-ci. La musique dans toute sa diversité peut être définie comme une construction sociale et culturelle fondée sur cette même musicalité. (Honing *et al.*, 2015, p. 2, cf. aussi Huron, 2001)

- 2 Les sciences de la vie et les sciences sociales se concentrent souvent sur la nature sociale de la musique (et du langage). En biologie, par exemple, les trois principales hypothèses d'évolution concernant la musique, à savoir la sélection sexuelle (Miller, 2000 ; Fitch, 2006), le lien parent-enfant (Dissanayake, 2008 ; Malloch & Trevarthen, 2009) et la cohésion de groupe (Freeman, 2000 ; Dunbar, 2012), mettent en lumière son caractère intrinsèquement social. La neurobiologie met ainsi l'accent sur les fondements neuronaux et hormonaux liés à la musicalité (Chanda & Levitin, 2013 ; Salimpoor *et al.* (<https://www.frontiersin.org/articles/10.3389/fnins.2021.667838/full#B119>), 2015 ; Mehr *et al.* (<https://www.frontiersin.org/articles/10.3389/fnins.2021.667838/full#B89>), 2019). Conformément à ces approches, l'objectif de cet article est de démontrer que la manière adéquate de saisir la nature sociale interactive de la musique (et, avant cela, de la musicalité) consiste à la concevoir comme un langage incarné, enraciné dans des structures cérébrales s'adaptant au contexte culturel. Nous pouvons proposer ici une définition de travail de la musique considérée comme un langage incarné : **un moyen de coordination, s'appuyant sur le sens d'action concertée, induit par les actions motrices suggérées par les sons.** Cette proposition répond à celle de Ian Cross d'étudier la musique comme un « processus de communication interactif » plutôt que comme « une manifestation de formes sonores » (Cross, 2014), en insistant sur ses aspects incarnés et prédictifs (de codage) (Vuust & Witek, 2014 ; Lesaffre *et al.*, 2017). Une approche similaire a récemment été proposée par van der Schyff et Schiavio lorsqu'ils écrivent que leur « approche bioculturelle considère la cognition (musicale) comme une propriété émergeant d'une activité incarnée et située dans un environnement socio-matériel en évolution » (2017, p. 7). Nous nous accordons sur une telle proposition, qui soutient fondamentalement que les niveaux corporels, neuronaux et environnementaux participent, depuis les premières sociétés, à la mise en place de comportements musicaux. **Notre propos n'est pas d'envisager le traitement de la musique comme étant identique à celui du langage, même si les racines des deux processus pourraient être les mêmes** (Mithen, 2005). Cependant, pour éviter les malentendus, nous insistons sur la façon métaphorique d'utiliser le mot « langage » dans notre proposition⁴. La musique, en particulier, n'est pas dotée d'une capacité référentielle claire qui lui permettrait de référer au monde

comme peut le faire le langage (la musique a une « intentionnalité flottante », comme le dit Cross, 2014). Néanmoins, en tant que langage, la musique est dotée d'un aspect syntaxique, sémantique et pragmatique. Nous nous concentrons sur ce dernier aspect et maintenons que la métaphore de la musique, en tant que langage incarné, pourrait non seulement mieux identifier la musique comme un phénomène social unique, mais aussi encourager une étude plus approfondie du terrain commun qu'elle partage avec la compétence linguistique elle-même.

- 3 Pour résumer, cet article a pour objectif de : (i) proposer un cadre considérant la musique sous l'aspect du langage incarné en s'appuyant sur une revue de littérature qui définit les principaux concepts de l'action musicale conjointe, en mettant particulièrement l'accent sur la cognition musicale incarnée et le traitement prédictif, ainsi que sur certains fonctionnements neuronaux correspondants ; (ii) passer en revue trois expériences réalisées dans nos laboratoires (et récemment publiées), dont les résultats à la fois appuient l'hypothèse du nouveau cadre conceptuel proposé dans (i) et peuvent être interprétés en fonction de celui-ci. Dans la première partie, nous préparons le terrain pour notre cadre, en définissant différents concepts issus à la fois des sciences cognitives musicales et des neurosciences. Notre cadre s'appuie sur la littérature consacrée à l'action conjointe et peut être vu comme le développement des approches traitant de la musique incarnée et du codage prédictif de la musique dans un tel cadre. Depuis que nous nous intéressons tout particulièrement aux fondements neuronaux de la pratique de la musique d'ensemble, nous faisons un état des lieux des études pertinentes les plus récentes, sans pour autant les développer davantage. Notre objectif principal est de présenter un cadre pour la musique comme langage incarné, qui soit biologiquement plausible et qui rassemble les multiples résultats obtenus jusqu'à présent dans le champ d'étude du « musiquer » ensemble (Small, 1998). Certains de ces résultats proviennent directement de nos laboratoires et sont donc exposés dans la seconde partie de cet article. En se concentrant sur le temps, l'espace et la qualité des interactions musicales au sein de duos de musiciens et de sujets non-musiciens, l'objectif des trois expériences présentées ici est de corroborer la conception de la

musique en tant que langage incarné fondé sur le concept central d'action concertée.

Première partie : la musique comme langage incarné : un cadre conceptuel

Action conjointe

- 4 L'action conjointe a été largement étudiée dans le domaine des sciences cognitives depuis plus d'une décennie. Une définition, proposée par Sebanz *et al.* (2006, p. 70), stipule que l'action conjointe correspond à « toute forme d'interaction sociale qui amène au moins deux individus à coordonner leurs actes dans le temps et l'espace pour engendrer un changement dans l'environnement ». Alors que soulever un objet à plusieurs est un cas d'action conjointe largement étudié (Marsh *et al.*, 2019), le changement d'environnement mentionné dans la définition ci-dessus peut être dans le même temps plus subtil et plus profond comme quand, par exemple, deux personnes échangent leurs regards dans le but de saisir leurs intentions mutuelles (Becchio *et al.*, 2018). En fait, le besoin de comprendre le phénomène de l'interaction sociale a été récemment réinvesti par un certain nombre de neuroscientifiques, soulignant que le « mode social » est sans doute le mode par défaut du cerveau de *l'homo sapiens*, sans oublier les autres espèces sociales et les mammifères en général (Caccioppo *et al.*, 2010 ; Schilbach *et al.*, 2013 ; Hari *et al.*, 2015). De ce fait, il est capital pour les études portant sur le cerveau de développer des méthodologies appropriées pour prendre en compte non seulement l'observation de l'action (à l'image des paradigmes classiques des neurones miroirs), mais également les contextes dans lesquels deux sujets ou plus modulent mutuellement leurs comportements sur le moment, que ce soit pour la compétition ou la coopération. Il est bien connu que le système des neurones miroirs est un réseau neuronal qui est impliqué de la même manière, non seulement pendant le mouvement de production de l'action, mais aussi durant son observation, ce qui suggère son implication dans la compréhension de l'action (Rizzolatti & Sinigaglia, 2010) et l'imita-

tion (Iacoboni *et al.*, 1999). Pour franchir les limites d'un paradigme « spectatorial » (Reddy & Uithol, 2015), une première découverte révélatrice est venue de Newman-Norlund *et al.* (2007) : l'activation du système BOLD⁵ est plus importante dans les aires fronto-pariétales (qui doivent correspondre au système humain des neurones miroirs) lors des phases de planification d'action complémentaire que lors des phases de planification d'action seulement simulée (et ce, avec une intensité ou une précision exigée pour la préhension d'un objet). Ces auteurs ont découvert que le même réseau neuronal responsable de la compréhension passive des actions observées est actif (en fait, il est plus actif) pour (la préparation de) l'interaction possible. Les échanges mère-enfant représentent l'essence même de l'interaction sociale. En effet, cette condition met en évidence le fait que l'observation est toujours intégrée dans les processus dynamiques d'adaptation, de réaction, d'incitation, etc., bien avant toute prise de conscience du contexte du côté du nourrisson, selon ce que De Jaegher et Di Paolo (2007) appellent le « faire-sens participatif » (« *participatory sense-making* ») (cf. ci-dessous). L'hyperscanning⁶, c'est-à-dire la production simultanée de données cérébrales de deux sujets ou plus, offre une possibilité intéressante d'explorer l'interaction sociale, dans la mesure où il prend en compte plus d'un individu en même temps, bien que les résultats impliquent des interprétations qui sont loin d'être simples (Konvalinka & Roepstorff, 2012 ; Babiloni & Astolfi, 2014 ; Hari *et al.*, 2015).

Cognition incarnée : sa composante « étendue »

- 5 Une idée centrale qui guide cette recherche est la cognition incarnée, un paradigme théorique à multiples facettes qui questionne depuis trois décennies les principes de base des sciences cognitives, en particulier la nature computationnelle de la représentation de l'esprit humain (Varela *et al.*, 1991 ; Clark, 1997 ; Thompson & Varela, 2001 ; Noë, 2004 ; Chemero, 2009 ; Gallagher, 2017). La cognition incarnée met l'accent sur le fait que le corps, l'environnement et les composantes sociales sont inextricablement liés, ce qui est appelé « 4E » (Newen *et al.*, 2018), c'est-à-dire les composantes incarnées (« *embodied* »), situées (« *embedded* »), éenactives (« *enactive* ») et

étendues (« *extended* ») de l'esprit et de la cognition. Cet article ne peut avoir pour ambition de développer chacun de ces aspects⁷, mais certains d'entre eux nécessitent une explication dans la perspective d'un essai théorique visant à combiner deux cadres de la recherche musicale apparemment opposés, qui sont l'interaction musicale incarnée (Leman, 2007, 2016) et le codage prédictif (Vuust & Witek, 2014 ; Koelsch *et al.*, 2019), pour les faire ainsi converger vers le concept de langage incarné.

6 Introduire le corps dans le cadre implique que l'esprit et la cognition ne sont plus considérés comme étant à l'origine de représentations du monde extérieur par le biais de computations neuronales, mais plutôt comme des processus guidant l'action dans/sur le monde, y compris pour des parties du monde qui sont particulièrement importantes pour les humains (et les animaux en général) : les conspécifiques. Cela met en évidence deux des 4E sur lesquels notre attention doit se porter de manière plus précise : la nature incarnée et, plus particulièrement, la nature étendue de l'esprit. Pour la première, nous pouvons nous concentrer sur ce que Hurley (1998) appelle « le modèle du sandwich » (« *the sandwich view* »), selon lequel le « microprocesseur » de l'esprit se situe entre la perception et l'action, c'est-à-dire dans ces processus computationnels se produisant après les stimuli sensoriels, mais avant les réponses motrices. Les approches incarnées ont contesté ce point de vue, soulignant que ce qu'un organisme perçoit dépend de la façon dont il bouge et que, vice versa, la manière dont un organisme bouge dépend de ce qu'il perçoit (comme Merleau-Ponty et Gibson l'avaient déjà dit). De telles boucles sensorimotrices dépendent non seulement des processus cérébraux, mais aussi de la morphologie et du fonctionnement du corps, dans la mesure où des corps différents peuvent être réceptifs à des affordances extérieures différentes (cf. ci-dessous). C'est l'interaction entre le cerveau, le corps et le monde (Clark, 1997) qui permet à l'esprit et à la cognition d'émerger, ce qui va à l'encontre de l'approche computationnelle centrée sur le cerveau de la (neuro)science cognitive classique.

7 La composante « étendue » de la théorie des 4E est typiquement l'une des plus controversées (Menary, 2010), puisqu'elle implique, dans sa version forte, l'inclusion (de parties) du monde extérieur dans le processus computationnel dont un organisme fait usage pour

résoudre un problème cognitif donné (cf. « le principe de parité » de Clark & Chalmers, 1998, selon lequel le cerveau n'a pas de priorité cognitive sur l'environnement). L'usage d'outils est un exemple classique. Lorsqu'un aveugle touche le bord d'un bâtiment avec sa canne pour s'orienter et tourner dans la bonne direction, il manifeste le fait que sa canne devient une partie de son corps, comme si ses propres doigts analysaient l'environnement. Maintenant, considérons une action conjointe telle que faire du vélo tandem. Non seulement cette action est impossible à réaliser par une personne seule, mais le degré de synchronisation nécessaire pour l'accomplir est si élevé que la coordination des actions individuelles peut engendrer une sorte de système « super-ordonné », c'est-à-dire un système étendu composé de deux agents (ou plus) interactifs (Müller *et al.*, 2018). Ce sont deux manières d'étendre l'esprit, par l'utilisation d'outils dans le premier cas et, par la coordination, avec un conspécifique dans le second. Tous deux sont pertinents pour la musique. Quels que soient les arguments philosophiques au sujet de l'inclusion de telles extensions dans la machine computationnelle de l'esprit, le phénomène précédent (en particulier, l'utilisation d'outils) a été étudié en profondeur dans les travaux récents en neurosciences cognitives et sera brièvement présenté dans ce qui suit.

- 8 1) Quels processus cognitifs sont-ils censés être étendus par l'utilisation d'outils ? Bien que les philosophes aient également attiré l'attention sur la mémoire et la pensée, les neurosciences se sont surtout focalisées sur la perception du corps et de l'espace péripersonnel. Rizzolatti *et al.* (1981) ont découvert des neurones bimodaux⁸ visuo-tactiles dans le cortex prémoteur ventral (PMv), le putamen et le sillon intrapariétal (SIP) des macaques, tous ces neurones déchargeant lorsqu'un objet semble proche du corps et lorsqu'il le touche. Dans la mesure où de tels neurones sont centrés sur des parties du corps (« *body-part centered* »), codifiant l'espace de la main, de la tête et du torse *mais aussi* ce qui se passe dans l'espace qui les environne⁹, ils peuvent être considérés comme les corrélats neuronaux de l'espace corporel (l'espace proprioceptif et tactile) et péripersonnel (l'espace multisensoriel à portée de main). En ce qui concerne l'espace corporel, Graziano (1999), Graziano et Gandhi (2000) ont démontré que ces champs neuronaux réceptifs sont activés par des objets à proximité d'une main factice (alors que la

vraie main du singe lui est dissimulée) et par la position de la fausse main après qu'elle a été incarnée grâce à une stimulation synchrone de la main factice et de la vraie main (cachée). Il s'agit d'un phénomène bien connu appelé « illusion de la main en caoutchouc » (Botvinick & Cohen, 1998), dans lequel une main factice est perçue par un individu comme étant sa vraie main, si elle est placée dans une position en adéquation avec son propre corps et si elle est touchée en même temps que la vraie main (cachée) au moyen d'une brosse. Par conséquent, le sentiment d'appartenance du corps, le sentiment qu'éprouve un individu qu'une partie du corps appartient à son propre corps, est modulé par la position, la forme et le mouvement de la main factice. De même, il a été démontré que l'espace péripersonnel est un phénomène plastique. En effet, Iriki *et al.* (1996, cf. aussi Marativa & Iriki, 2004) ont montré qu'après que le macaque s'est entraîné à ramasser des objets avec un râteau, ses neurones bimodaux visuo-tactiles SIP ont étendu leurs champs réceptifs pour couvrir toute la longueur du râteau. En d'autres termes, alors qu'avant de s'entraîner avec des outils ces neurones ne déchargeaient que lorsqu'un stimulus était émis à proximité de la main ou de l'épaule ou lorsqu'il les touchait, ils déchargeaient également, après l'activité, pour des stimuli émis dans l'espace éloigné : aussi éloigné que la longueur du râteau. Une telle reconfiguration de l'espace proche en espace lointain trouve son équivalent chez les humains. Par exemple, les patients atteints de négligence visuelle suite à un AVC montrent des signes de dissociation des espaces proche et lointain, la négligence ne se manifestant que pour l'espace proche, ainsi que les résultats de tâches de bissection le démontrent¹⁰. Cependant, si la bissectrice était réalisée avec un bâtonnet, plutôt qu'un crayon optique, allongeant ainsi la longueur du bras, la négligence visuelle se manifestait également dans l'espace lointain (Berti & Frassinetti, 2000 ; Neppi-Modona *et al.*, 2007).

- 9 2) Plus récemment, les neurosciences ont abordé la possibilité que l'interaction sociale ait également une certaine influence sur les processus cognitifs comme ceux liés à la perception du corps ou de l'espace péripersonnel. Soliman *et al.* (2015) avancent que, pendant et après une action conjointe, comme le sciage en binôme d'une bougie avec une corde, il se développe chez les participants un « schéma corporel conjoint » (« *joint body-schema* »), mesurable au moyen

d'une tâche d'intégration multisensorielle visuo-tactile. Cette expérimentation consiste à évaluer le temps de réaction à un stimulus tactile délivré sur le pouce ou sur l'index pendant qu'un stimulus visuel apparaît, soit près du pouce/index du participant, soit près du pouce/index du partenaire (pour les détails, cf. Maravita *et al.*, 2003). Au cours de la condition conjointe, contrairement à la condition solo, l'incongruence (par exemple, pouce touché/index vu) a eu une incidence sur les temps de réaction, en les ralentissant, ce qui indique qu'une interdépendance du schéma corporel des deux sujets a vu le jour, en raison de l'action conjointe qui vient d'être accomplie. Profitant d'un paradigme d'intégration multisensorielle différent, Teneggi *et al.* (2013) démontrent qu'une interaction coopérative, comparée à une interaction non coopérative, peut moduler l'espace péripersonnel d'une personne au sein d'un duo. En effet, après la condition coopérative, les sujets ont réagi plus rapidement à un stimulus tactile sur leurs mains, à l'écoute d'un stimulus auditif émis non seulement près d'eux, mais également plus loin, près du partenaire coopératif (pour aller plus loin, cf. Canzoneri *et al.*, 2012). Puisqu'une réponse à un stimulus tactile est facilitée par un stimulus auditif émis dans l'espace péripersonnel, grâce aux neurones bimodaux mentionnés ci-dessus, ce résultat est interprété comme la preuve que l'espace péripersonnel s'est étendu après l'interaction coopérative.

- 10 À ce stade, il convient également de souligner que Thompson et Varela (2001), deux des principaux théoriciens de la cognition incarnée, avaient déjà fait valoir que l'une des trois dimensions de l'incarnation est l'interaction intersubjective (ainsi que ce qu'ils appellent « autorégulation corporelle » et « couplage sensorimoteur »). Comme nous le verrons, l'une de nos expériences aborde l'une des deux précédentes caractéristiques « étendues » constitutives du cadre de la cognition incarnée, explorant l'espace péripersonnel multisensoriel de musiciens après une interaction musicale (jazz) coopérative/non coopérative (cf. ci-dessous). Cependant, afin de préparer le terrain pour notre proposition théorique et pour chacune de nos expériences, nous devons examiner l'approche du codage prédictif et la façon dont les problématiques liées à l'incarnation, dont nous venons de discuter, peuvent être transposées en termes musicaux.

Le codage prédictif : sa composante « sensorimotrice »

- 11 Les boucles sensorimotrices décrites ci-dessus comme une caractéristique cruciale de l'approche incarnée ressemblent fortement à la « causalité circulaire sensorimotrice » (« *circular sensorimotor causality* ») mentionnée par Friston (2013) dans sa présentation de l'approche du codage prédictif (cf. aussi Clark, 2016). Dans ce processus inférentiel, la causalité circulaire sensorimotrice implique que

les états externes provoquent des changements dans les états internes, via des états sensoriels, tandis que les états internes se couplent aux états externes par le biais d'états actifs – de telle sorte que les états internes et les états externes s'entraînent réciproquement. Cette causalité circulaire serait donc une architecture causale fondamentale et omniprésente pour l'auto-organisation. (Friston, 2013, p. 2-3)

- 12 Selon le cadre prédictif, le cerveau a pour fonction de donner sens au monde extérieur en minimisant les erreurs résultant de la comparaison entre la prédiction des causes d'un ressenti et ce ressenti lui-même. Supposons que l'on observe quelqu'un qui se saisit d'un scalpel (Kilner *et al.*, 2007). Le cerveau peut utiliser sa connaissance du contexte, disons un hôpital, comme un préalable à comparer à l'action observée, en émettant l'hypothèse que le scalpel a été saisi pour soigner un patient donné. Si le scalpel est utilisé pour frapper la tête du patient, on peut considérer que le cerveau a commis une erreur de prédiction importante. Au contraire, si le scalpel est placé dans une boîte de stérilisation, l'erreur est moindre, et même nulle s'il est réellement utilisé pour opérer le patient. Dans tous les cas, l'erreur de prédiction permet de mettre à jour les prédictions antérieures à l'événement (qui, une fois mises à jour, deviennent postérieures), dans un processus circulaire de prédictions fondé sur la sensorimotricité. Le monde est ainsi modélisé selon des termes bayésiens comme une « hiérarchie de systèmes au sein desquels des causes supra-ordonnées induisent et modèrent les changements dans les causes subordonnées », offrant « une orientation contex-

tuelle vers la cause la plus probable de l'information sensorielle » (Kilner *et al.*, 2007, p. 163).

- 13 De telles boucles sensorimotrices peuvent également être qualifiées d'inférences actives, en ce sens que le corps entier, plutôt que le cerveau seul, enclenche activement le processus inférentiel (prédictif), en échantillonnant activement l'environnement (externe ou interne), y compris les comportements des conspécifiques. Ainsi, si les mécanismes inférentiels sont envisagés en termes sensorimoteurs plutôt qu'en termes computationnels-représentationnels, l'approche bayésienne peut coexister et enrichir l'approche incarnée (cf. Maes, 2016 ; Gallagher & Allen, 2016, pour des croquis de propositions de synthèse similaires, et ci-dessous). Un modèle de musique correspondant aux principes du codage prédictif a récemment été proposé par Koelsch *et al.* (2019), qui affirment que, en écoutant de la musique (même sans la jouer), nous pourrions générer des prédictions sensorimotrices concernant les caractéristiques rythmiques, couplées à des actions motrices comme taper des mains, secouer la tête ou danser, en particulier lorsque la musique « *groove* ». De telles prédictions sont continuellement actualisées, en les confrontant à l'environnement sonore réel. En d'autres termes, « bouger » sur la musique aide à désambigüiser certaines de ses caractéristiques au moyen d'une prédiction incarnée qui peut être décrite en termes bayésiens. En effet, dans ce modèle, que les auteurs qualifient explicitement d'« énactif » (Koelsch *et al.*, 2019, p. 74), l'appréciation musicale n'est pas simplement guidée par la minimisation des erreurs, mais plus encore par les fluctuations de l'incertitude des prédictions. Après avoir passé en revue quelques études pertinentes sur la cognition incarnée musicale, nous reviendrons sur ces concepts afin de les intégrer dans un cadre unique, global, autour de l'idée de la musique considérée comme langage incarné.

Cognition musicale incarnée

- 14 Le point de vue désincarné considère typiquement la cognition musicale comme une reconstruction computationnelle d'une organisation hiérarchique de la musique de manière récursive, allant des stimuli acoustiques de base à la vaste structure formelle d'une composition donnée : à la manière, plutôt, d'une grammaire générative de la

cognition du langage (Lerdahl & Jackendoff, 1983). Au contraire, la cognition musicale incarnée tire parti des boucles sensorimotrices susmentionnées pour en faire une caractéristique cruciale du fonctionnement du cerveau, afin de mettre en évidence le rôle du corps dans la perception et la production de la musique (Leman, 2007). Cela est étayé par des études de synchronisation et d'entraînement, de désambiguïsation et d'externalisation du « timing »¹¹ (Maes, 2016).

- 15 Premièrement, considérons le phénomène d'entraînement qui amène le rythme du corps à se synchroniser avec le rythme d'une musique (Clayton, 2012 ; Phillips-Silver & Keller, 2012 ; Moens & Leman, 2015). Les mécanismes de prédiction et d'adaptation sensorimoteurs sont pris en charge par des réseaux neuronaux situés dans le lobe pariétal postérieur, le cortex prémoteur, le cervelet et les ganglions de la base, faisant naître le phénomène du « groove » (Janata *et al.*, 2012), ce qui suggère que les mêmes processus qui provoquent le mouvement du corps sont impliqués dans la perception du rythme musical. Comme l'écrit Todd :

Si la forme spatio-temporelle de certains stimuli [sensoriels] correspond à la dynamique du système moteur, alors ils peuvent évoquer un mouvement d'une représentation interne, ou image motrice, des éléments synergiques correspondants au système musculo-squelettique, *même si le système musculo-squelettique lui-même ne bouge pas.* (Todd, 1999, p. 120).

- 16 Iyer (2002) souligne que la musique, selon son tempo, peut évoquer différentes actions humaines, comme la respiration, la marche et la parole¹² (avec des fréquences comprises respectivement entre 0,1 et 1 Hz, 1 et 3 Hz, 3 et 10 Hz), mais l'inverse est également vrai. En effet, de nombreuses compositions musicales se rangent dans cette catégorie de tempo, ce qui laisse entendre que les résonateurs corporels ont en quelque sorte modelé la façon dont les humains créent de la musique (van Noorden & Moelants, 1999).
- 17 Deuxièmement, le mouvement peut également désambigüiser une structure métrique. Dans des expériences menées par Phillips-Silver et Trainor (2005) sur un motif rythmique ambigu¹³, des bébés ont été bercés en rythme, puis, dans un deuxième temps, il a été demandé à des adultes de plier leurs genoux en rythme. Leurs mouvements

étaient réglés de manière à marquer la deuxième ou la troisième pulsation, donnant ainsi un mètre binaire ou ternaire, ainsi que le montrèrent par la suite leurs réponses lorsqu'on leur demanda de reconnaître sur laquelle des deux versions rythmiques ils avaient bougé (tandis que les adultes répondaient verbalement, on observa les nourrissons à l'écoute de leur version préférée).

- 18 Troisièmement, le timing n'est souvent pas une question de comptage des pulsations, mais plutôt de mouvement, soit par l'utilisation de stratégies d'externalisation par lesquelles les membres se mettent en mouvement, soit par des chorégraphies maintenues dans des mouvements itératifs ne nécessitant pas d'attention cognitive. Su et Pöppel (2012) ont montré que les non-musiciens s'appuient plus que les musiciens sur leurs propres mouvements afin de sentir la pulsation d'une séquence rythmique, ce qu'ils n'arrivent pas à faire lorsqu'il ne leur est pas permis de se mouvoir. Cependant, les musiciens peuvent aussi se fier à leur horloge interne pour comprendre une séquence sans bouger, mettant ainsi en évidence l'importance du mouvement corporel, en particulier lorsque l'expertise est absente. De plus, il est utile de rappeler qu'il a été observé que les neurones miroirs dépendent également d'une telle expertise sensorimotrice. Par exemple, il a été montré que les aires frontales inférieures et pariétales, typiquement impliquées dans l'activation des neurones miroirs, sont plus actives (d'après un scanner IRMf) chez les pianistes, comparés à des sujets non-musiciens, et ce, en observant les mouvements de doigts jouant du piano, par rapport à des mouvements de doigts ne jouant pas de piano (Haslinger *et al.*, 2005, cf. aussi Herholz & Zatorre, 2012).

- 19 Un cadre proposé par Leman (2007) soutient que :

Le corps humain peut être vu comme un médiateur biologique qui transfère l'énergie physique jusqu'à un niveau de significations orientées vers l'action, à un niveau mental dans lequel les expériences, les valeurs et les intentions constituent les composantes de base de la signification musicale. Le processus inverse est également possible : que le corps humain transfère une idée, ou une représentation mentale, dans une forme matérielle ou dynamique. (p. XIII)

- 20 L'énergie physique est la surface acoustique de la musique et la représentation mentale correspondante est l'intention attribuée par l'auditeur/producteur à cette musique, « sur la base d'une simulation de l'action perçue dans la propre action du sujet » (Leman, 2007, p. 92, cf. aussi Koelsch *et al.*, 2019 comme résumé ci-dessus). Autrement dit, à partir d'un répertoire d'actions motrices (à la fois transitives et intransitives, c'est-à-dire des gestes), le corps reproduit des caractéristiques musicales, telles que le rythme, les contours mélodiques, les intensités, les tempi, etc., pour mieux les comprendre et les apprécier. Alors que Schiavio et Menin (2013) interprètent cette approche comme un dualisme, la proposition de Leman peut être comprise différemment : avec l'esprit (impliquant une certaine forme d'attention et une réflexion sur la signification) considéré comme émergeant des processus liés au corps (Broeckx, 1981). En effet, la médiation entre le mental et le physique intervient à un niveau conscient de traitement, généralement impliqué dans la construction réfléchie du sens, alors qu'une grande partie du traitement peut être conçue de manière strictement sensorimotrice, suspendant toute proposition sur la nature de ce dont le corps est supposé être le médiateur. Considérons à nouveau le processus de désambiguïsation rendu possible par le mouvement d'une partie du corps suivant un rythme binaire ou ternaire sur une pulsation isochrone. Il n'est pas nécessaire d'attribuer des propriétés physiques aux pulsations que nous n'entendons ni, d'un autre côté, des propriétés mentales à notre expérience subjective, dans la mesure où la perception de ces sons est couplée aux mouvements corporels nécessaires à leur désambiguïsation. Ce qui importe dans une approche incarnée de la cognition (musicale) est que de tels mouvements itératifs sensorimoteurs, plutôt que des opérations abstraites, constituent précisément la cognition (musicale). Surtout, le mécanisme sensorimoteur dont il est question ici est double. D'un côté, il s'agit de la morphologie du corps, le fait que le corps humain peut effectuer des actions différentes de celles des autres corps animaux ; par exemple, comme nous l'avons vu ci-dessus, la synchronisation autour de plages de fréquences spécifiques, en fonction de l'action motrice impliquée. D'un autre côté, les mécanismes sensorimoteurs sont dotés d'une correspondance neuronale spécifique, bien représentée, aussi bien chez les humains que chez les singes, par le système miroir (Rizzolatti & Sinigaglia, 2008, et ci-dessous).

- 21 Les neurones miroirs permettent de mieux comprendre les interactions musicales entre plusieurs personnes comme un processus incarné qui ne peut être réduit à des processus internes de lecture de l'esprit ou de simulation de cerveaux en interaction (Thompson & Varela, 2001). Selon De Jaegher et Di Paolo (2007), on pourrait plutôt parler de « faire-sens¹⁴ participatif » (cf. Schiavio & De Jaegher, 2017, pour une application musicale de ce concept), soulignant ainsi le caractère incarné d'une interaction entre deux sujets (ou plus) nourrie par l'échange continu de paramètres spatio-temporels. De Jaegher et Di Paolo (2007) attirent notre attention sur une action conjointe très élémentaire comme passer à deux en même temps par une porte qui est trop étroite pour laisser passer ensemble deux personnes sans qu'elles ne soient obligées de se pencher et d'adapter leur taille et leur position au corps de l'autre. Il est à noter que, si l'on demandait à ces personnes de répéter cette action plusieurs fois, elles le feraient probablement chaque fois d'une manière légèrement différente, ce qui rendrait manifeste le fait qu'une dynamique d'ajustement mutuel légèrement différente s'est développée, bien que parvenant au même résultat (passer par la porte). Si nous appliquons ce scénario à un contexte de pratique musicale d'ensemble, certaines caractéristiques peuvent émerger, qui sont étroitement liées à la connexion temporelle entre les corps et aux relations sonores dans les espaces.
- 22 Chez Walton *et al.* (2015), par exemple, un système de capture de mouvement a enregistré les mouvements de l'avant-bras et de la tête de deux pianistes improvisant soit sur un bourdon (fondé sur une alternance uniforme de deux accords), soit sur un ostinato (fondé sur une progression complexe de quatre accords). Grâce à la transformation en ondelettes croisées (« *cross-wavelet transform* »), la série temporelle de ces mouvements a révélé différentes périodicités, en fonction des caractéristiques de la musique. En effet, il s'est avéré que la figure de base de l'ostinato, répétée toutes les quatre secondes, permet aux mouvements des musiciens de se coordonner selon des multiples de quatre secondes. Au contraire, le bourdon n'a révélé aucune périodicité spécifique, ce qui est probablement dû à sa structure plus simple, qui a offert plus de variété de mouvements aux musiciens (et, par conséquent, de possibilités musicales ; mais l'inverse est également vrai). Les auteurs en ont conclu que les inter-

actions expressives ne sont pas seulement guidées par des processus cérébraux, mais également par des dynamiques corporelles émergeant spontanément, conformément avec l'un des principes de l'approche incarnée de la cognition (cf. aussi Walton *et al.* (<https://www.frontiersin.org/articles/10.3389/fnins.2021.667838/full#B119>), 2018).

Les bases sensorimotrices neurales de l'interaction musicale

23 Au cours de la dernière décennie, un certain nombre de travaux ont étudié les circuits neuronaux à l'œuvre lors de l'action musicale conjointe. La preuve de l'implication du cortex primaire moteur (M1) dans l'observation de l'action a été fournie par la recherche pionnière de Fadiga *et al.* (1995), dans laquelle il a été démontré que l'activation cortico-spinale d'un sujet était renforcée pendant l'observation d'une action motrice transitive (saisie) d'une autre personne sur un objet, par rapport au simple fait d'observer cet objet (pour un équivalent auditif, cf. Aziz-Zadeh *et al.* (<https://www.frontiersin.org/articles/10.3389/fnins.2021.667838/full#B119>), 2004). Au moins deux études ont corroboré de tels résultats dans le domaine de la musique d'ensemble. Dans la première étude, Novembre *et al.* (2012) ont laissé un groupe de pianistes répéter quelques compositions avant de leur demander de jouer leur partie mélodique avec la main droite, soit seule, soit pendant que la partie de la main gauche était jouée par un partenaire caché (enregistrement). La stimulation magnétique transcrânienne (SMT) à pulsation-unique du M1 de la main/du bras gauche inactive/inactif a montré des potentiels évoqués moteurs (PEM)¹⁵ plus élevés dans la situation d'ensemble, mettant en lumière que des représentations motrices peuvent apparaître en réponse à une interaction sociale potentielle. Dans une seconde étude, Novembre *et al.* (2014) ont testé un autre groupe de pianistes, dont la moitié avait travaillé un morceau donné, l'autre non. Lorsqu'il leur a été demandé d'adapter le tempo de leur main droite au tempo changeant de la main gauche du partenaire, après la stimulation magnétique transcrânienne à double pulsation du M1 de la main gauche, le dernier groupe s'est montré plus précis que le premier. C'est-à-dire que la SMT à double pulsation n'a perturbé que les processus reposant sur la stimulation sensorimotrice de la partie répétée jouée par le parte-

naire, un mécanisme qui est clairement mobilisé dans la coordination en temps réel des actions générées par soi et par le partenaire.

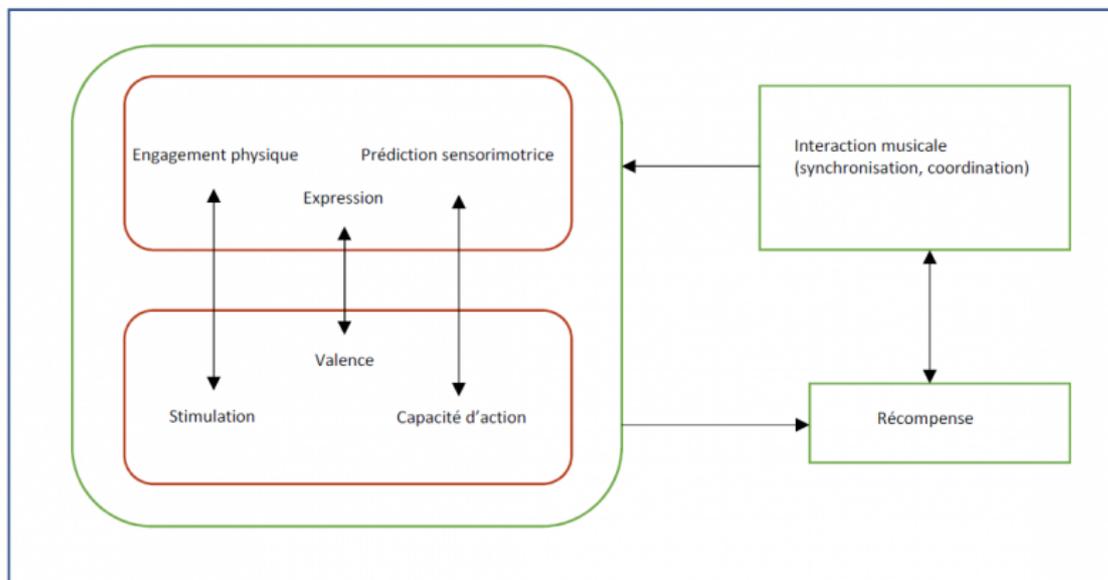
- 24 Étant donné que la musicalité s'appuie sur une dimension biologique, il peut être attendu que nous trouvions également chez les non-musiciens des mécanismes sensorimoteurs, comme ceux décrits précédemment. Par exemple, Gordon *et al.* (2018) ont récemment trouvé une facilitation¹⁶ cortico-spinale au niveau des premiers muscles interosseux dorsaux (IOD)¹⁷ des non-musiciens pendant qu'ils regardaient une séquence de trois notes au piano comportant un retard de 200 ms entre le son et la vidéo, condition expérimentale comparée à la condition correcte (sans retard) d'un stimulus audiovisuel ou à des conditions unimodales correctes (soit uniquement auditive, soit uniquement visuelle). Les auteurs ont conclu que les modèles prédictifs sensorimoteurs sont ici en jeu, plutôt que les mécanismes de type simulation, étant donné que seule la violation du résultat sensoriel attendu entraîne une augmentation de l'excitation cortico-spinale. D'un autre côté, si l'on apprend à des non-musiciens à exécuter des mélodies simples au piano, l'excitation cortico-spinale des premiers muscles interosseux augmente en écoutant ces mélodies, quelques millisecondes avant même qu'elles ne commencent, ce qui montre ainsi la différence entre un entraînement moteur et le simple fait d'écouter (Stephan *et al.*, 2018). Ce résultat est cohérent avec Candidi *et al.* (2012), qui ont montré qu'à chaque fois qu'un doigté incorrect était observé, les pianistes ayant l'expérience motrice d'une pièce musicale donnée affichaient des PEM des doigts plus élevés que les pianistes ayant uniquement l'expérience visuelle de cette pièce.
- 25 Là encore, ces résultats peuvent être interprétés comme des processus sensorimoteurs qui appuient une approche incarnée du traitement de la musique associée à une approche de traitement prédictif bayésien (cf. le paragraphe suivant). Les neurones sensorimoteurs multimodaux sont probablement le substrat de tels processus, en particulier dans des aires, telles que les STG et les STS¹⁸, qui répondent plus fortement aux stimuli auditifs et visuels qu'aux stimuli auditifs ou visuels séparément (Beauchamp *et al.*, 2004 ; cf. Kohler *et al.*, 2002, pour les neurones miroirs). Cependant, comme nous l'avons indiqué ci-dessus, les méthodes interactives de recherche sur le cerveau ouvrent la voie pour surmonter certaines des contraintes qui caractérisent les neurosciences sociales depuis la

découverte des neurones miroirs. Babiloni *et al.* (2012), dans une étude pionnière, ont exploré la performance musicale de trois différents quatuors de saxophones au moyen d'EEG simultanés, découvrant que les rythmes alpha dans les zones frontales (Bas 44/45) sont corrélés avec les scores d'empathie des musiciens qui observent leur propre performance (à propos de « l'hyperscanning » musical, cf. aussi Osaka *et al.*, 2015 ; Pan *et al.*, 2018).

Un cadre pour la musique en tant que langage incarné

- 26 La composante émotionnelle-motivationnelle, qui implique le circuit de la récompense (Salimpoor *et al.*, 2015), devrait compléter les éléments théoriques présentés ci-dessus, fondés sur l'incarnation et le traitement prédictif. En tant que tel, chaque type d'interaction avec la musique (que ce soit en l'écoutant ou en la jouant, en solo ou en groupe) peut être compris comme constitué par une boucle cognitive-motivationnelle qui engendre, chez les sujets impliqués, un sentiment de récompense et de réalisation personnelle (« *empowerment* »). Par conséquent, le modèle récent de Leman (2016) comprend l'engagement physique, le traitement prédictif et l'expression (sous forme de signaux biosociaux) en parallèle avec l'activation physiologique, la capacité d'action et les attitudes pro-sociales (Figure 1).

Figure 1



L'hypothèse d'interaction-récompense (Leman, 2016) postule que le processus de récompense via l'interaction en musique (au travers de la coordination et de la synchronisation) se fonde sur un engagement physique, un contrôle cognitif et une gestuelle expressive qui fonctionnent de pair avec la stimulation, la capacité d'action et les valeurs prosociales.

- 27 Nous verrons un peu plus loin si et comment ce modèle peut être intégré à celui de Keller *et al.* (2014 ; Keller, 2008), mais, pour l'instant, nous devons nous pencher sur une seule de ses composantes : l'orientation prosociale induite par le sentiment d'action¹⁹ (induit, à son tour, par les prédictions sensorimotrices qui lui sont propres) dans l'interaction avec la musique, puisque c'est un point crucial pour notre définition de la musique comme langage incarné. Le sentiment d'action, un phénomène largement étudié dans les neurosciences cognitives, est le sentiment de contrôle d'un individu donné sur une action donnée qu'il est en train de réaliser (Haggard & Eitam, 2015). Dans la vie de tous les jours, c'est un sentiment implicite, qui devient manifeste si quelque chose ne va pas, comme lorsque l'on s'apprête à appuyer sur un interrupteur, mais que la lumière s'allume juste avant d'appuyer : ce n'est pas nous qui allumons la lumière, mais quelqu'un d'autre, d'où un sentiment d'action faible (voire inexistant). D'autre part, étant probablement construit sur la prédiction des conséquences de nos actions, plutôt que sur leurs conséquences sensorielles réelles (Berti & Pia, 2006), un sens illusoire de sentiment d'action peut s'ensuivre.

- 28 Les prédictions sensorimotrices (fondées sur les inférences bayésiennes susmentionnées) sont capables d'induire la sensation qu'une musique donnée a été produite par notre action motrice, ce qui rappelle le concept de causalité de Hume (Leman, 2016). Une telle sensation serait (consciemment) trompeuse dans les cas où l'on bouge sur la musique sans la jouer, comme lorsqu'on court, danse, ou même simplement tape sur le rythme de la musique, mais elle serait vraie chaque fois que nous jouons vraiment de la musique. Néanmoins, dans les deux cas, un sentiment de récompense et de réalisation personnelle peut s'ensuivre, dû aussi à un élément prosocial (valence) qui (au moins partiellement) explique le pouvoir expressif des interactions musicales. Cette idée est en phase avec les études qui mettent l'accent sur la capacité de la musique à faire en sorte que les individus soient (Overy & Molnar-Szakacs, 2009) ou restent (McNeill, 1995 ; Hove & Risen, 2009) en rythme ensemble, développant un sens de l'action concertée, un concept sur lequel le philosophe Pacherie a récemment travaillé (cf. ci-dessous). Probablement, ce qui manque encore à une telle théorie (comme à de nombreuses propositions dans la littérature des neurosciences et de la musicologie) est une description plus détaillée, dans l'interaction musicale, de la relation entre la qualité expressive et les aspects prosociaux. Des indications sur ce type de lien sont mises en évidence dans le modèle *Shared Affective Motion Experience (SAME)*²⁰ d'Overy et Molnar-Szakacs (2009, p. 492), qui « suggère que le son musical est perçu non seulement en termes de signal auditif, mais aussi en termes de séquences intentionnelles et hiérarchiquement organisées d'actes moteurs expressifs à l'origine du signal ». Sans surprise, ces auteurs invoquent la mobilisation du réseau des neurones miroirs dans la mise en œuvre neurale de telles expériences avec la musique. En outre, ils utilisent le concept de « sentiment d'action » (« *sense of agency* ») (terme utilisé dans un sens différent de l'usage courant) pour souligner le sentiment d'interaction humaine qui se trouve au cœur de l'expérience musicale, « un sentiment de la présence d'une autre personne, de ses actions et de ses états affectifs » (Overy & Molnar-Szakacs, 2009, p. 494 ; cf. aussi Clarke, 2005 ; Livingstone & Thompson, 2009 ; Windsor & de Bézenac, 2012).
- 29 C'est précisément l'idée qu'une personne se cache derrière un son musical qui mène à la possibilité de concevoir la musique comme un

langage incarné. Cette idée résonne avec la proposition de Leman, pour qui « l'expression musicale est plus qu'une simple habitude ou qu'une pratique établie. L'expression est ancrée dans la biologie du fonctionnement humain propre aux interactions sociales, elles-mêmes étroitement liées à des états et des comportements affectifs », (2016, p. 49). Contrairement au langage naturel, la musique permet de coordonner en temps réel des comportements au sein de grands groupes, comme par exemple dans les chants de stades, de guerre et de travail, et la manière dont ces comportements renforcent les identités collectives et donc l'appartenance culturelle est bien connue, en particulier dans les études ethnologiques (Freeman, 2000 ; Nettl, 2005 ; Clarke *et al.*, 2015). Si les facteurs biologiques de la musicalité (qui sous-tend la musique) sont probablement présents dans l'encodage des hauteurs, la perception de la pulsation et l'encodage métrique du rythme (Honing *et al.*, 2015), nous pouvons supposer qu'ils sous-tendent le caractère de communication de la musicalité et, par conséquent, de la musique en tant que processus de communication interactif incarné (Mithen, 2005 ; Malloch & Trevarthen, 2009 ; Cross, 2014). On peut donc s'attendre à trouver ces facteurs universellement répandus parmi les humains, quel que soit leur niveau d'expertise musicale, constituant ainsi les conditions préalables à cette expertise, plutôt que le résultat de celle-ci (Mehr *et al.*, 2019).

- 30 Afin de montrer comment un langage incarné peut fonctionner, considérons à nouveau la musique d'ensemble, une forme sophistiquée d'action conjointe qui, peut-être sans surprise, a permis une étude équilibrée, depuis environ une décennie, entre des conditions expérimentales contrôlées, d'une part, et des conditions en situation réelle, d'autre part (D'Ausilio *et al.*, 2015). Selon le modèle de Keller (Keller, 2008 ; Phillips-Silver & Keller, 2012 ; Keller *et al.*, 2014), la coordination interpersonnelle dans un ensemble musical repose sur une combinaison de processus cognitifs d'ordre supérieur, comme le partage d'une idée globale de la pièce musicale en train d'être exécutée (qui, à son tour, dépend des conventions socioculturelles) et de compétences cognitives et motrices d'ordre inférieur, comme le timing adaptatif mutuel²¹, la participation intégrative priorisée²² et les images de transformation²³. Ces processus peuvent en quelque

sorte caractériser tout type d'action commune (Vesper *et al.*, 2010), mais dans un contexte musical, ils reviennent au fait que :

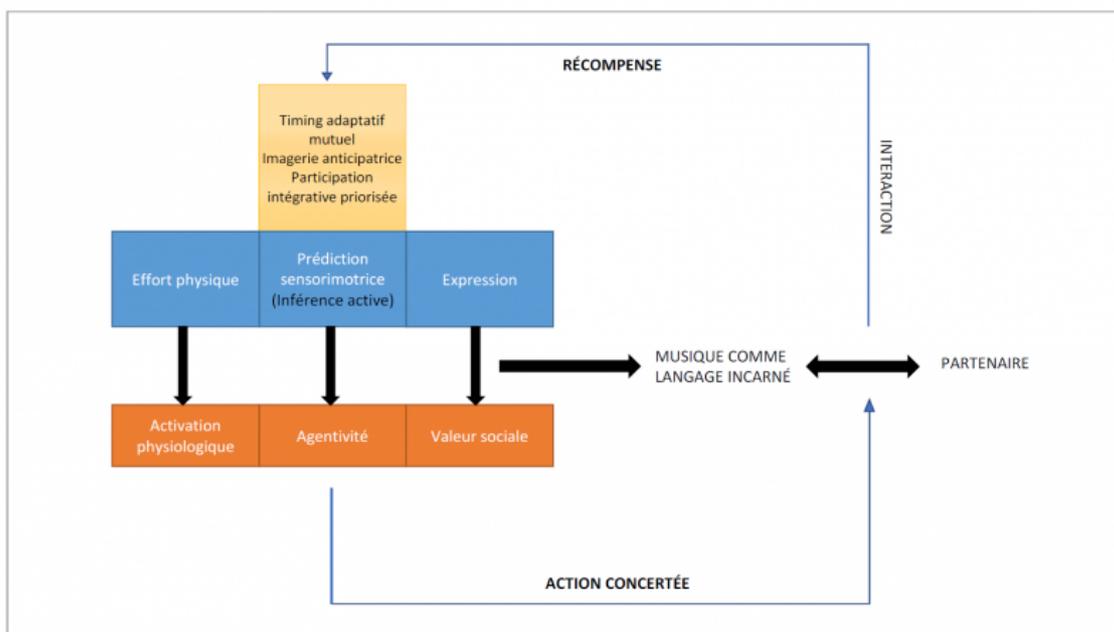
- 31 1) deux sujets ou plus qui jouent ensemble ont besoin d'une coordination temporelle si précise et si souple qu'ils puissent faire face, d'une part, à des micro-perturbations involontaires du timing, dues à la variabilité intrinsèque aux actions humaines, et, d'autre part, à des variations volontaires de timing dues à des effets expressifs (*accelerando/ritardando*). La correction de phase et la correction de période sont deux mécanismes mis en avant pour expliquer de telles compétences (Repp & Su, 2013)²⁴.
- 2) Un musicien doit être attentif non seulement à ce qu'il joue, mais aussi à ce que joue l'ensemble, en priorisant ses ressources pour le premier aspect, sans perdre de vue le second. La capacité interne à garder la pulsation est alors sollicitée pour suivre la structure multicouche de la musique d'ensemble, assez souvent composée de sections rythmiques, de lignes mélodiques entrelacées et, plus généralement, selon la pièce musicale, de différentes parties (London, 2004).
- 3) S'ils veulent que leur performance reste stable et cohérente, les musiciens doivent dans une certaine mesure anticiper le jeu de leurs partenaires. Keller et Appel (2010) ont démontré, par exemple, que les duos de piano les plus synchronisés étaient ceux formés par des pianistes qui, dans une tâche d'interprétation musicale sans retour auditif, avaient une plus grande faculté de création d'images mentales.
- 32 Le concept du codage prédictif propose une approche qui semble capable de fédérer les trois aspects précédents, puisque la musique est dotée d'une structure intrinsèquement hiérarchique, à la fois du point de vue mélodique (cellules insérées dans des phrases insérées dans des sections) et rythmique (rythmes insérés dans des cellules insérées dans des métriques), que l'inférence bayésienne peut traiter de manière appropriée (Salimpoor *et al.*, 2015 ; Koelsch *et al.*, 2019). De plus, l'interaction musicale peut, elle aussi, bénéficier d'un tel cadre. En effet, les boucles sensorimotrices nécessaires à la réalisation d'une action individuelle, prédisant le résultat d'une action donnée et ajustant celle-ci en cas de retour sensoriel inapproprié, peuvent être traduites en termes d'interaction sociale (Wolpert *et al.*, 2003 ; Kilner *et al.*, 2007 ; Friston & Frith, 2015 ; Volpe *et al.*, 2016 ;

Brattico & Vuust, 2017). Dans ce dernier cas, nous pouvons prédire les conséquences d'une de nos actions sur un partenaire (par exemple, accepter d'être embrassé), tandis que la rétroaction sensorielle serait fournie par la réaction du partenaire (par exemple, nous éviter), qui, à son tour, permet un ajustement de notre action (par exemple, faire semblant d'atteindre quelque chose derrière le partenaire) pour minimiser les erreurs de prédiction. Dans un contexte musical, considérons que l'action est l'attaque du thème du standard de jazz *Autumn Leaves* après sept mesures introductives. Le musicien jouant le thème doit s'adapter au tempo fixé par la section rythmique (disons, le piano, la basse et les percussions), en assurant sa propre partie sans négliger celles des autres et en anticipant une exécution correcte. Après avoir joué les deux premières notes de la huitième mesure, le soliste se rend compte que ni la basse, ni le piano n'ont changé l'accord menant à la véritable première mesure du thème, il ajuste donc sa trajectoire en transformant ces deux notes en une sorte d'ornement précédant le thème, dont le départ est retardé d'une mesure. Il convient de souligner que de tels processus n'ont pas à être pleinement conscients, puisque les modèles internes sont censés fonctionner selon une hiérarchie imbriquée, allant de niveaux très bas (proches des réflexes) à des niveaux conscients (très proches de la pensée propositionnelle, cf. Friston & Frith, 2015). Comme nous le verrons, deux de nos expériences étudient le timing adaptatif à la recherche, d'une part, de compétences protomusicales de la pratique d'ensemble chez des non-musiciens et de la manière dont elles sont modulées par la façon dont le partenaire incarne son geste de la main, d'autre part, de marqueurs dynamiques dans une performance chantée, en fonction de la manière dont l'action concertée est expérimentée (cf. ci-dessous).

- 33 Le concept de musique et de musicalité en tant que langage incarné, autour duquel s'articule le présent travail, peut maintenant être résumé au moyen d'un diagramme (Figure 2). L'ambition d'un tel cadre théorique est d'associer le codage prédictif et les approches incarnées (par exemple, Leman, 2007, 2016 ; Keller, 2008 ; Vuust & Witek, 2014 ; Keller *et al.*, 2014 ; Koelsch *et al.*, 2019). Ce nouveau cadre implique trois composantes nécessaires pour jouer de la musique ensemble. Une première, qui inclut les trois compétences sensorimotrices de Keller, déploie une inférence active²⁵ au cours d'une inter-

action musicale, que ce soit lors d'une action individuelle ou conjointe, que ce soit simplement en écoutant ou en faisant de la musique. Une seconde composante implique l'agentivité²⁶ comme conséquence des prédictions sensorimotrices incarnées. Une dernière composante, résultante des deux premières, implique l'activation physiologique.

Figure 2



Un cadre théorique de la musique comme langage incarné. Le rôle crucial est joué par l'action conjointe, qui est une conséquence du système de prédiction sensorimotrice, ainsi que l'activation physiologique (l'action conjointe et l'activation physiologique participant l'une comme l'autre à l'effet de récompense). Le système prédictif s'appuie sur l'inférence active et inclut les trois compétences sensorimotrices qui, pour Keller, sont impliquées dans la pratique de la musique d'ensemble. Quelques preuves expérimentales (qui concernent par exemple le *tapping*, le jazz, le hoquet et l'utilisation d'outils) peuvent être interprétées à travers ce modèle (cf. ci-dessus).

34 L'action concertée est la façon dont nous voyons l'agentivité dans un contexte musical. En effet, compte tenu de la valeur prosociale de la musicalité (et donc de la musique) en tant que langage incarné (Wiltermuth & Heath, 2009 ; Kokal *et al.*, 2011), l'agentivité dans de tels contextes n'est pas simplement le sentiment d'être en contrôle d'une action individuelle donnée, comme dans la saisie d'un objet pour une action ordinaire, mais elle se doit d'être dotée d'une dimension « concertée » impliquant la présence plus ou moins manifeste d'un ou

plusieurs sujets interagissant musicalement. La version simplifiée de ce cadre est aisément applicable à une performance d'ensemble réelle, comme les interactions dyadiques que nous allons explorer dans les expériences suivantes, mais une version plus développée pourrait identifier une composante sociale également dans les interactions individuelles avec la musique. En effet, si la musique (et avant elle la musicalité) est le produit bioculturel d'interactions sociales, on pourrait considérer qu'une trace de son origine sociale est toujours présente, quel que soit le type d'interaction musicale en jeu. L'exemple le plus marquant est celui de la musique d'ensemble, mais on peut établir une échelle mesurant la présence de moins en moins manifeste d'une autre personne dans l'écoute de la musique (moins, s'il s'agit de musique live, plus, si ce n'est pas le cas), et, éventuellement, dans le fait de jouer seul. D'après ce point de vue, jouer seul peut être comparé à parler seul, qu'il s'agisse de répéter un monologue ou de penser à voix haute. Le sentiment d'action concertée qui en résulte diminuerait en conséquence (cf. aussi Clarke, 2005 ; Livingstone & Thompson, 2009 ; Overly & Molnar-Szakacs, 2009 ; Windsor & de Bézenac, 2012, pour des considérations similaires). Concernant le troisième élément du réseau, nous considérons simplement les aspects suivants. Une expérience telle que celle de Fritz *et al.* (2013) a montré que dès qu'une personne contrôlait certains paramètres de la musique à l'aide de divers appareils de sport, sa sensation d'effort était moindre par rapport à une condition passive, dans laquelle elle écoutait simplement de la musique tout en faisant du sport²⁷. En outre, la force de la stimulation engendrée par la musique a été démontrée dans plusieurs expériences portant sur la vitesse de la marche en écoutant de la musique par rapport à un métronome (Styns *et al.*, 2007) ou sur différents genres musicaux (Leman *et al.*, 2013), ce qui a permis d'identifier des genres apparaissant comme plus stimulants que d'autres. Dans tous ces cas, il semble que se produise un transfert d'énergie sonore en énergie motrice (cf. aussi Tarr *et al.*, 2014).

Deuxième partie : données expérimentales

- 35 Trois de nos propres expériences décrites dans les prochaines sections fournissent des données à l'appui du cadre proposé et peuvent être interprétées à la lumière de celui-ci. Elles se concentrent respectivement sur le temps, l'espace et la qualité de l'interaction musicale. Dans une première expérience, nous montrons que les non-musiciens peuvent aussi communiquer protomusicalement. Nous avons étudié le timing lorsqu'ils tapaient alternativement des doigts de façon conjointe et si et comment ce mouvement est modulé par la position du partenaire (avec l'activation cortico-spinale relative mesurée au moyen d'une stimulation magnétique transcrânienne (SMT) à pulsation-unique). Dans une deuxième expérience, nous montrons que l'espace péripersonnel de deux musiciens de jazz en interaction peut être modulé selon le caractère coopératif ou non coopératif d'une telle interaction, un tel espace se mesurant au moyen du paradigme d'intégration multisensorielle visuo-tactile qui nous a également permis de comparer les temps de réaction des musiciens et des non-musiciens. Enfin, dans la dernière expérience, nous nous sommes concentrés sur le concept d'action concertée dans des duos chantant en hoquet, corrélant ce paramètre subjectif avec une mesure objective et dynamique de la qualité de leur timing, conformément aux principes bayésiens.

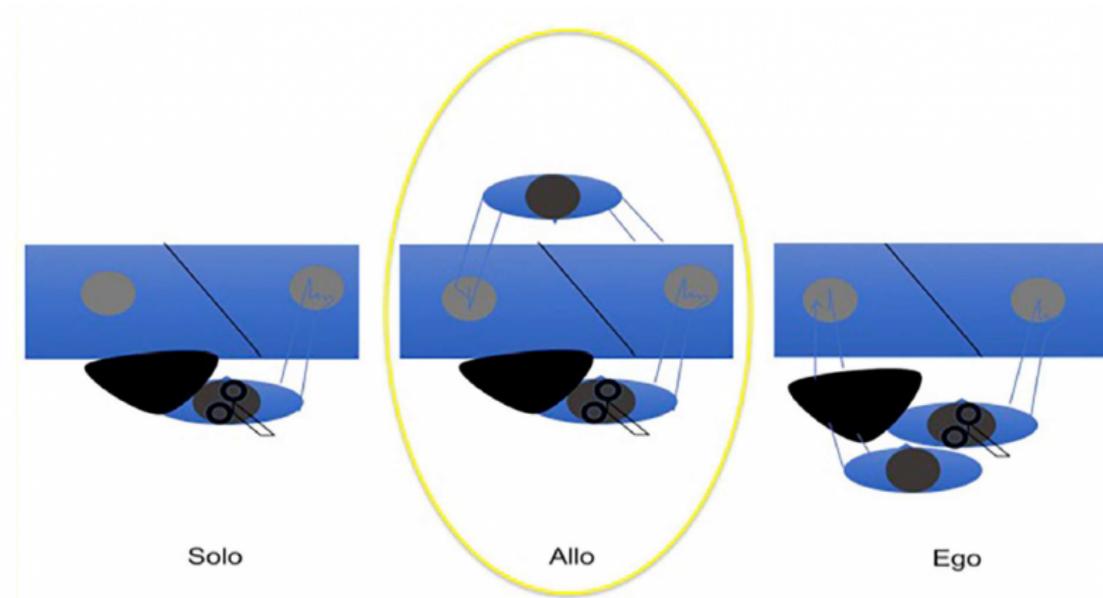
Le temps : entraînement et incarnation dans une interaction de tapping²⁸

- 36 Une façon simple d'étudier le timing adaptatif mutuel (TAM) est le *tapping*, une action motrice protomusicale permettant aussi aux non-musiciens de synchroniser le mouvement d'une partie de leur corps avec la pulsation de la musique. Des expériences précédentes ont montré que les musiciens sont capables d'adapter leur timing à celui du *tapping* d'un partenaire qui tape en opposition de phase (Nowicki *et al.*, 2013), et que les non-musiciens sont capables d'en faire de même dans une tâche de *tapping* en phase (Konvalinka *et al.*, 2010)²⁹. Selon l'hypothèse d'une musicalité innée, à l'instar

de Konvalinka *et al.* (2010, cf. aussi Koelsch *et al.*, 2000), nous avons montré (Dell'Anna *et al.*, 2018) que les non-musiciens sont aussi capables de s'adapter au timing de leur partenaire dans une tâche de *tapping* conjoint alterné (c'est-à-dire en opposition de phase) par rapport à une battue métronomique de référence qui marque tous les demi-cycles (c'est-à-dire le début et le milieu de chaque cycle). Nous avons utilisé la corrélation des asynchronies comme méthode pour mesurer l'entraînement. La tâche de *tapping* a été effectuée dans trois conditions : (i) seul avec le métronome, (ii) avec un partenaire en face du sujet et (iii) avec un partenaire à côté du sujet, dans une position congruente à son corps, de manière à ce que le partenaire tape avec sa main gauche et le sujet avec sa main droite (Figure 3). La dernière condition utilise les modèles dans lesquels des mains « étrangères » peuvent être utilisées, aussi bien chez des sujets en bonne santé (où le sujet, moyennant des manipulations et contraintes particulières, perçoit une main en caoutchouc comme étant la sienne : l'illusion de la main en caoutchouc, Botvinick & Cohen, 1998, cf. aussi ci-dessus), que chez ceux souffrant de lésions cérébrales (où le sujet ressent et considère la main réelle d'un autre individu comme étant la sienne, Garbarini *et al.*, 2014). Nous nous attendions à ce que l'excitabilité cortico-spinale soit plus élevée dans la condition (ii), comparée aux conditions (i) et (iii), en raison des mécanismes miroirs que l'action partagée devrait activer. Lorsque la distinction entre le soi et les autres faiblit, comme dans la condition (iii), alors le mécanisme miroir ne fonctionne pas, comme s'il n'y avait plus de partenaire avec qui interagir et, par conséquent, l'excitabilité cortico-spinale sera similaire entre (i) et (iii). Les résultats de notre expérience montrent que le timing s'adapte mutuellement dans les conditions (ii) et (iii), mais pas dans la condition (i). De surcroît, les conditions (ii) et (iii) diffèrent par le sentiment d'appropriation, car, dans la dernière, la main étrangère est perçue comme la propre main du sujet, avec un sentiment d'agentivité concernant le *tapping*. Dans la condition (iii), lorsque le sujet incarne une main étrangère, l'excitabilité cortico-spinale a tendance à décroître par rapport à la condition (ii) c'est-à-dire lorsqu'il y a un partenaire en face du sujet (Schutz-Bosbach *et al.*, 2006 ; Della Gatta *et al.*, 2016). Les résultats peuvent être interprétés comme si un contexte d'interaction engageait le système moteur, tandis que l'incarnation de la main ou du bras d'un partenaire n'entraîne aucune interaction sociale. L'enregistrement des potentiels

évoqués moteurs (PEM) au moyen de la stimulation magnétique transcrânienne (SMT) sur les premiers muscles interosseux dorsaux (IOD) de la main a confirmé cette idée dans notre tâche protomusicale. Quand le sujet tapotant a incarné le bras de son partenaire (selon les rapports subjectifs d'agentivité et de sentiment d'appropriation), l'excitabilité corticale ne différait pas de la condition seule. Au contraire, lorsque le sujet tapait en face de son partenaire, le caractère social du contexte entraînait une excitabilité cortico-spinale plus élevée, en accord également avec la littérature traitant des neurones miroirs (Fadiga *et al.*, 1995 ; Novembre *et al.*, 2012). En tant que composante d'un langage incarné, la dimension rythmique fournie par le métronome a entraîné mutuellement les actions motrices de base des deux sujets en interaction, avant toute prise de conscience du processus de leur part.

Figure 3



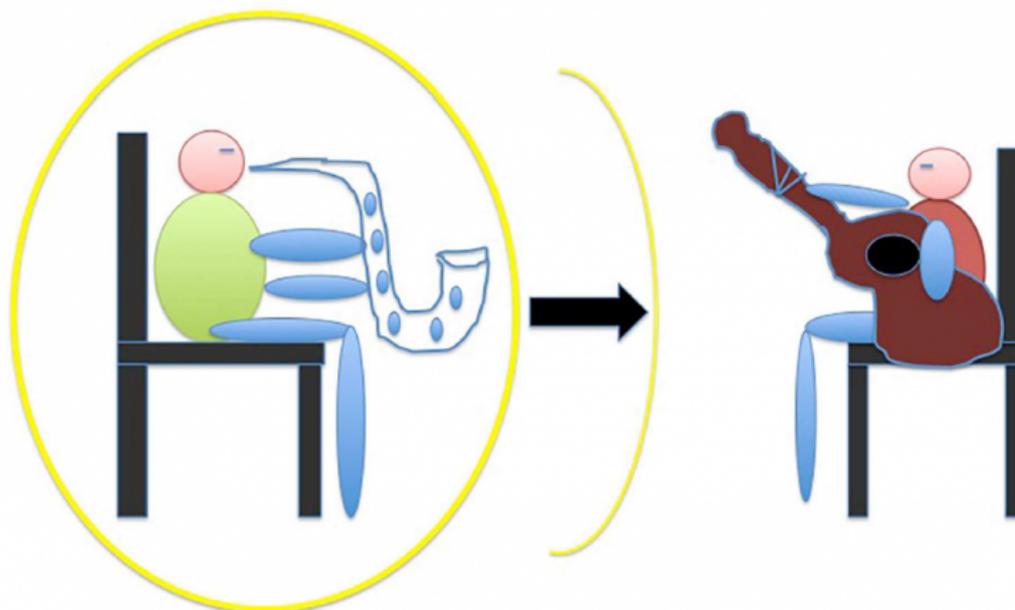
La condition allocentrique induit une excitabilité cortico-spinale plus élevée que la condition égocentrique (mais une corrélation conjointe des asynchronies comparable), ce qui constitue la meilleure configuration pour l'action concertée.

Espace : reconfiguration de l'espace péripersonnel dans une interaction en jazz

- 37 Comme nous l'avons mentionné, l'espace péripersonnel, qui est la représentation multisensorielle de l'espace entourant immédiatement le corps, s'est récemment révélé être sensible non seulement à l'utilisation d'outils (Iriki *et al.*, 1996 ; Berti & Frassinetti, 2000), mais aussi à l'interaction sociale (Patané *et al.*, 2016 ; Pellencin *et al.*, 2018). En particulier, il a été démontré que l'espace péripersonnel s'étend après un échange qui s'inscrit dans un cadre structuré coopératif par rapport à un échange de même type non coopératif (Teneggi *et al.*, 2013). De même, nous avons laissé jouer des duos de musiciens, dont l'un des partenaires jouait la séquence harmonique d'un standard de jazz de manière soit correcte soit incorrecte, en faisant l'hypothèse que seule la première condition provoquerait une extension de l'espace péripersonnel des musiciens (Figure 4). Afin de mesurer l'espace péripersonnel suite aux deux conditions expérimentales (la condition harmonique coopérative et la condition harmonique non coopérative), nous avons utilisé une tâche d'intégration audiotactile conçue par Serino *et al.* (2007, cf. aussi Canzoneri *et al.*, 2012), qui ont montré qu'un son dont la source d'émission est proche du sujet, comparé à un son éloigné, facilite les temps de réaction à un stimulus tactile cooccurrent³⁰. Un son éloigné est donc conditionné par ce qui est subjectivement expérimenté comme éloigné. En revanche, il est apparu (Dell'Anna *et al.*, 2020b) que seule la condition non coopérative affectait l'étendue de l'espace péripersonnel, le faisant disparaître, comme si le sujet se désolidarisait du partenaire non coopératif. Nous avons interprété ce résultat comme une preuve que, dans la mesure où la musique et la musicalité sont des langages incarnés intrinsèquement sociaux, une interaction musicale a un impact mesurable sur la perception de l'espace entre deux sujets ou plus. Ce modèle nous a également permis de comparer notre échantillon de musiciens à un échantillon de non-musiciens. En accord avec une découverte récente (Landry & Champoux, 2017), nous avons eu confirmation que les musiciens réagissent plus rapidement aux stimuli audiotactiles que les non-musiciens, quelle que soit la distance du stimulus auditif, probablement en raison de l'entraîne-

ment sensorimoteur des musiciens avec leur instrument et (dans une moindre mesure) le chant, qui entraîne des réorganisations cortico-sous-corticales bien connues (Munte *et al.*, 2002 ; Zimmerman & Lahav, 2012).

Figure 4



Nous nous attendions à une extension de l'espace péripersonnel du musicien après l'interaction coopérative, mais nous avons obtenu une sorte de disparition de celui-ci après l'interaction non coopérative. Dans tous les cas, nous pouvons conclure que le sentiment d'action concertée a été altéré par le partenaire non coopératif.

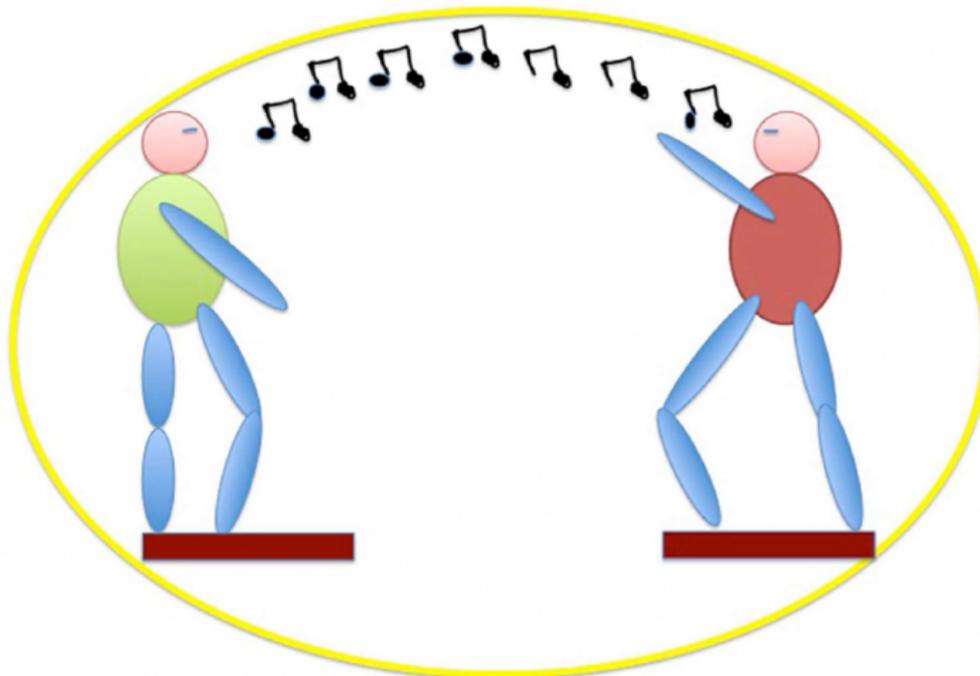
Qualité : timing expressif et action concertée dans l'interaction à l'œuvre lors de l'exécution d'un hoquet³¹

38 Les expériences décrites jusqu'à présent, bien que se rapportant à l'interaction musicale, sont axées sur une approche centrée sur l'analyse de faits musicaux. Un objectif central d'une autre de nos expériences était de concevoir un moyen de saisir la dynamique d'un duo de chant, afin d'évaluer la qualité de l'interaction pendant l'exécution d'un hoquet, et ce, en se concentrant sur le timing (Dell'Anna *et al.*, 2020a). Alors que la majeure partie des études sur le timing chez un duo de musiciens ont utilisé une certaine forme

de corrélation des asynchronies ou des asynchronies signées moyennes (Goebel & Palmer, 2009 ; Clayton *et al.*, 2019)³², les méthodes restent centrées sur la musique elle-même car les faits qui se produisent correspondent à des occurrences temporelles régulières, qui peuvent être interprétées comme des événements subséquents. Nous avons essayé ici de mettre au point une méthode qui, indépendamment d'une référence fixe, serait capable de prendre en compte la variabilité intrinsèque du comportement humain dans le temps. En effet, compte tenu du caractère alterné du chant en hoquet, la référence s'exprime de manière latente (comme un tempo qui émerge et qui peut éventuellement changer avec le temps). Pour tenir compte de l'interaction, nous avons choisi l'écart temporel de déclenchement entre chaque succession de deux notes (chantées de manière alternée par les deux chanteurs) et calculé une erreur de durée en termes bayésiens, par rapport au tempo latent qui fluctue dans le temps et que nous avons utilisé comme prédicteur de la durée. Cette approche, dans laquelle le tempo latent est une sorte de moyenne évolutive utilisée en tant que prédicteur pour mesurer l'écart temporel de déclenchement subséquent, a conduit à une mesure dynamique de la précision temporelle, que nous avons appelée erreur de fluctuation. Puisque nous nous intéressions également à l'expérience subjective d'une interaction musicale, nous avons corrélié cette mesure temporelle avec une autoévaluation, par les chanteurs, de la qualité de la performance et du sentiment d'action concertée (Figure 5). Le concept d'action concertée a récemment fait l'objet d'un débat intense. Selon Pacherie (2012), il existe un sentiment d'action concertée PARTAGÉ (« SHARED ») et un sentiment d'action concertée d'un NOUS (« WE »), le premier étant le sentiment de contrôler une partie de l'action concertée, le second étant le sentiment de constituer une seule entité avec le partenaire tout en accomplissant cette action. En effet, un duo de chant peut être considéré comme une *Gestalt* dynamique dont les éléments contraignent le déroulement de l'exécution de l'un et de l'autre au moyen de ce langage incarné qu'est la musique (Walton *et al.*, 2015 ; Müller *et al.*, 2018). La manière dont nous avons conçu notre partition de hoquet aurait pu engendrer l'action concertée d'un NOUS, mais c'est en fait une action concertée PARTAGÉE qui a été observée. De plus, nous avons constaté, quant aux erreurs de durées, une corrélation plus

marquée avec l'autoévaluation qu'avec les valeurs mesurées de l'action concertée.

Figure 5



L'action concertée est renforcée chaque fois que les écarts temporels de déclenchement (ETD) subséquents sont correctement prédits par le système bayésien dans lequel le duo de chant se construit.

Application du cadre théorique

- 39 Faisant écho à la proposition de Cross (2014) d'étudier la musique en tant que « processus de communication interactif » et de combiner le codage prédictif avec des descriptions incarnées de la musique, nous formulons maintenant une interprétation des trois expériences précédentes à la lumière de notre cadre de la musique en tant que langage incarné. Dans la première expérience, c'est la musicalité, plutôt que la musique, qui est prise en compte, tant que seuls la pulsation isochrone du métronome et le *tapping* – entraîné par celle-ci – du duo sur les *drum pads* constituent ensemble le modèle sonore : alors il y a interaction protomusicale, telle que nous l'avons nommée. Le caractère incarné du langage musical s'est traduit ici par sa capacité à coordonner à un niveau microtemporel les actions

motrices des deux sujets en interaction, malgré leur manque d'expertise musicale. Il est important de noter que, selon notre cadre, à la fois la prédiction des (quelques) prochains tapotements du partenaire et le sentiment d'action concertée qui s'ensuit font partie de cette boucle cognitivo-motivationnelle générée par l'interaction musicale. Au contraire, la composante « concertée » de l'action ne semble jouer aucun rôle dans les modèles de Keller ou de Leman mentionnés ci-dessus. La corrélation que nous avons trouvée entre les asynchronies des partenaires peut être interprétée comme une première évaluation d'un marqueur temporel d'un tel langage incarné, c'est-à-dire d'une caractéristique qui identifie un échange expressif plus ou moins réussi entre deux sujets (ou plus). En revanche, nous pensons avoir trouvé un autre marqueur temporel, et plus intéressant, dans notre étude sur la qualité du chant du hoquet. L'intérêt de ce marqueur réside dans sa nature dynamique, qui tire parti de l'ajustement réciproque continu entre les déclenchements sonores des deux chanteurs sur la base de prédictions à court terme. En particulier, dans cette étude, nous nous sommes concentrés sur une caractéristique temporelle : l'écart de déclenchement entre les sons émis successivement par les deux chanteurs, qui est intrinsèquement intersubjectif, appliquant ainsi déjà en quelque sorte le concept de langage incarné. Contrairement aux deux autres expériences, l'étude sur le hoquet a aussi questionné explicitement le paramètre de l'action concertée, trouvant une corrélation entre celui-ci et le marqueur dynamique de la qualité du timing de l'exécution, ce qui corrobore la conception de la musique comme langage incarné fondé sur l'action concertée. Le résultat de l'étude sur la modulation de l'espace péripersonnel, après une interaction en jazz, peut là encore être compris comme l'effet d'un manque d'action concertée. La condition non coopérative a en effet altéré « l'incorporation mutuelle » (Fuchs & De Jaegher, 2009) nécessaire pour coordonner une action musicale conjointe non seulement dans le temps, mais aussi dans l'espace (péripersonnel), empêchant ainsi le sentiment de récompense et de réalisation personnelle d'émerger. Il s'agit là d'un exemple clair de communication (incarnée) manquée, qui brise d'une manière ou d'une autre le cadre de coordination qui se met en place chaque fois qu'une interaction se produit au sein d'un ensemble musical (et provoquant, métaphoriquement parlant, l'explosion de la bulle temporaire enveloppant les musiciens, mais cf. Bufacchi & Iannetti, 2018, pour une critique

d'une telle métaphore). Pour conclure, bien que les trois compétences sensorimotrices postulées par Keller (2008, Keller *et al.*, 2014), à savoir le timing adaptatif mutuel, l'imagerie anticipatrice et la participation intégrative priorisée, et les composantes du modèle de Leman (2016), à savoir la prédiction, l'engagement physique et l'expression, sont certainement à l'œuvre dans nos trois scénarios expérimentaux³³, nous soulignons le rôle crucial d'un autre facteur : l'action concertée. Le sentiment d'un contrôle partagé sur une action (musicale) donnée ou l'unité totale avec un ou plusieurs conspécifiques rendue possible par une activité musicale sous forme de rites, de cérémonies ou simplement d'échanges mère-enfant, constitue une dimension tellement fondamentale de la musique et de la musicalité en tant que langages incarnés que des recherches supplémentaires sont certainement nécessaires pour en démêler toutes les complexités.

- 40 Comme nous l'avons dit, alors que l'application de notre cadre à des contextes sociaux semble assez aisée, sa version la plus aboutie devrait également prendre en compte les interactions individuelles avec la musique. L'étude suivante, actuellement en cours dans nos laboratoires, en est un exemple. Étant donné que dans notre première expérience, nous avons utilisé la stimulation magnétique transcrânienne à pulsation unique (SMT), afin de vérifier la sociabilité de la condition allocentrique, et que dans notre deuxième expérience nous avons utilisé l'intégration multisensorielle audiotactile (MSI) comme moyen pour mesurer l'espace péripersonnel, dans une nouvelle étude, nous combinons ces deux méthodes. Notre objectif est de déterminer si l'activation cortico-spinale d'un instrumentiste à vent est renforcée par le fait de toucher une trompette tout en écoutant des sons de trompette (par rapport au fait de toucher des ciseaux et/ou d'écouter des bruits blancs), en partant de l'hypothèse que son expertise l'inciterait à se sentir prêt à jouer avec un autre musicien (virtuel), si la condition audiotactile congruente se produisait (cf. Schulz *et al.*, 2003 ; Yamaguchi *et al.*, 2014 pour des données empiriques stimulantes). À cette fin, nous comparons un groupe d'instrumentistes à vent avec un groupe de non-musiciens, dans la mesure où seul le premier groupe est censé présenter une telle forme de marqueur d'action concertée, en raison des compétences spécifiques requises par le cadre expérimental et, probablement, par les

réseaux de neurones miroirs sous-jacents (cf. aussi ci-dessus). En d'autres termes, l'association des stimuli auditif et tactile pourrait être une conséquence de l'expertise acquise, étant donné que, pour le musicien, le fait de s'engager avec la trompette est au départ nécessairement associé au son de cet instrument.

Conclusion

- 41 En résumé, nous nous sommes appuyés dans cet article à la fois sur la musicologie cognitive et les neurosciences pour esquisser un cadre théorique global sur l'interaction en musique, en tirant parti de plusieurs aspects de la pratique musicale en duo, depuis une action protomusicale très élémentaire, comme le *tapping*, jusqu'à des actions plus élaborées, comme le fait de jouer un standard de jazz et de chanter une mélodie en hoquet. Notre cadre combine des approches du codage incarné et du codage prédictif, centrées sur le concept d'action concertée. Si l'interaction sociale est le mode par défaut par lequel les humains communiquent avec leur environnement (Hari *et al.*, 2015), la musique et la musicalité, considérées comme un langage incarné, peuvent sans doute offrir une voie vers sa réalisation. Le caractère métaphorique de l'analogie que nous proposons entre la musique et le langage devrait encourager, à notre avis, une exploration plus poussée de la nature sociale de tout type d'interaction avec la musique. En outre, il pourrait inviter à une analyse approfondie d'aspects autres que l'aspect pragmatique que nous avons mis en évidence dans le présent article, en partant, par exemple, de la littérature sur les neurones miroirs relative aux processus linguistiques, pour faire ressortir des liens plus profonds entre la musique et le langage (Rizzolatti & Arbib, 1998 ; Arbib, 2013).
- 42 Tout d'abord, étant donné l'intérêt récent des neurosciences pour la compréhension des interactions sociales, nous avons exploré certaines hypothèses dans le contexte de la recherche sur l'action conjointe, dans lequel les approches incarnées et prédictives pourraient être mieux articulées entre elles. Nous avons donc mis l'accent sur les composantes incarnées et étendues de la cognition, étant donné qu'elles pourraient être les principaux éléments susceptibles de favoriser l'intégration des deux approches susmentionnées dans le domaine musical. Ensuite, la composante sensorimotrice du modèle

du codage prédictif a été mise en évidence, dans la mesure où elle peut être considérée comme la plus naturellement proche du cadre incarné. Une incursion dans le domaine intensément débattu de la cognition musicale incarnée a été envisagée comme une étape nécessaire vers la formulation de notre synthèse, juste avant un bref aperçu des résultats les plus récents des neurosciences cognitives concernant l'interaction sociale dans la musique. Un cadre théorique de la musique en tant que langage incarné a été esquissé, qui vise finalement à rendre justice à la nature intrinsèquement interactive de l'expérience musicale, indépendamment de l'interaction sociale réelle qui pourrait être impliquée. L'action concertée, principale caractéristique de notre approche envisageant la musique comme un langage incarné, est présentée comme le cœur conceptuel autour duquel les approches incarnées et prédictives peuvent converger.

43 Le principal intérêt de notre proposition réside dans la volonté de fédérer différents champs de recherche, qui ont été fortement débattus au cours des vingt dernières années, et d'en opérer une nouvelle synthèse dans le domaine de la cognition musicale. Encore une fois, concevoir la musique comme un langage incarné, c'est prendre au sérieux la neurobiologie actuelle lorsqu'elle insiste sur l'importance de l'interaction sociale dans l'émergence de l'esprit humain et de la cognition (Caccioppo *et al.*, 2010 ; Dennett, 2017), dont une part est constituée du formidable phénomène de la musique. Si tel est le cas, les futures voies de recherche devraient tenir compte du fait que la meilleure façon de définir la musique passe par l'interaction sociale, même si nous traitons de paramètres apparemment neutres comme le timbre, le rythme, le profil mélodique, etc. (cf. aussi McDermott, 2009 ; Bryandt, 2012).

44 Cependant, un certain nombre de limites subsistent dans le présent travail. Premièrement, plutôt qu'un modèle, c'est un cadre de la cognition musicale que nous présentons, d'où la difficulté d'émettre des hypothèses plus ciblées. En particulier, des recherches empiriques supplémentaires sont nécessaires pour tester la portée du concept d'action concertée ; par exemple, s'il peut aussi, comme nous le postulons, réellement jouer un rôle dans les interactions individuelles avec la musique. Deuxièmement, il semble souhaitable et possible de prendre en compte de manière plus poussée l'intégration de nombreux autres aspects des approches incarnées et prédictives

que nous n'avons pas abordés ici. Les partisans du modèle des 4E, par exemple, nous invitent à inclure la cognition incarnée dans le paradigme plus large de « l'énaction », tel que proposé initialement par Varela *et al.* (1991) et précisé plus récemment par Gallagher (2017) ou Newen *et al.* (2018). Il serait intéressant de se pencher sur une telle possibilité, étant donné qu'une approche énative dans le domaine musical a été continuée avec persistance ces dernières années par Schiavo et Altenmüller (2015), Schiavo et De Jaegher (2017) ou Van der Schyff et Schiavo (2017). Enfin, bien que le codage prédictif ait parfois été présenté comme étant compatible avec l'approche incarnée, même par certains de ses partisans (Friston & Frith, 2015 ; Koelsch *et al.*, 2019), l'essentiel de ses applications est centré sur le cerveau, tirant parti des hiérarchies neurales facilement identifiables dans le cerveau. Nous avons juste commencé à montrer comment cette approche peut être « étendue » à l'environnement, non seulement en utilisant des outils (musicaux), mais aussi (et principalement) au moyen d'interactions sociales. De nombreuses recherches sont nécessaires pour compléter ce tableau. Étant donné la nature complexe du phénomène du « musiquer » (« *musicking* »)³⁴ ensemble (Small, 1998), qui englobe des aspects biologiques et culturels, il ne sera pas surprenant de voir se multiplier, dans un avenir proche, des travaux interdisciplinaires associant des biologistes de l'évolution, des neuroscientifiques, des psychologues, des musicologues, des philosophes ainsi que des musiciens. Le présent travail n'est qu'une goutte d'eau dans cette mer dont les limites restent inexplorées.

BIBLIOGRAPHY

Arbib, M. (ed.), *Language, Music, and the Brain: A Mysterious Relationship*, Cambridge, MA : MIT Press, 2013.

Aziz-Zadeh, L., Iacoboni, M., Zaidel, E., Wilson, S., and Mazziotta, J., « Left hemisphere motor facilitation in response to manual action sounds », *Eur. J. Neurosci*, 19, 2004, p. 2609-2612. doi : 10.1111/j.0953-816x.2004.03348.x

Babiloni, C., Buffo, P., Vecchio, F., Marzano, N., Del Percio, C., Spada, D., et al., « Brains in concert: frontal oscillatory alpha rhythms and empathy in professional musicians », *NeuroImage* 60, 2012, p. 105-116. doi 10.1016/j.neuroimage.2011.12.008

Babiloni, F., and Astolfi, L., « Social neuroscience and hyperscanning

- techniques: past, present and future », *Neurosci. Biobehav. Rev.* 44, 2014, p. 76-93. doi : 10.1016/j.neubiorev.2012.07.006
- Beauchamp, M. S., Argall, B. D., Bodurka, J., Duyn, J. H., and Martin, A., « Unravelling multisensory integration : patchy organization within human STS multisensory cortex », *Nat. Neurosci.* 7:1190, 2004. doi : 10.1038/nn1333
- Becchio, C., Koul, A., Ansuini, C., Bertone, C., and Cavallo, A., « Seeing mental states: an experimental strategy for measuring the observability of other minds », *Phys. Life Rev.* 24, 2018, p. 67-80. doi : 10.1016/j.plprev.2017.10.002
- Berti, A., and Frassinetti, F., « When far becomes near : remapping of space by tool use », *J. Cogn. Neurosci.* 12, 2000, p. 415-420. doi : 10.1162/089892900562237
- Berti, A., and Pia, L., « Understanding motor awareness through normal and pathological behavior », *Curr. Dir. Psychol. Sci.* 15, 2006, p. 245-250. doi : 10.1111/j.1467-8721.2006.00445.x
- Bolt, N. K., Poncelet, E. M., Schultz, B. G., and Loehr, J. D., « Mutual coordination strengthens the sense of joint agency in cooperative joint action », *Conscious. Cogn.* 46, 2016, p. 173-187. doi : 10.1016/j.concog.2016.10.001
- Bolt, N., and Loehr, J., « The predictability of a partner's actions modulates the sense of joint agency », *Cognition* 161, 2017, p. 60-65. doi : 10.1016/j.cognition.2017.01.004
- Botvinick, M., and Cohen, J., « Rubber hands “feel” touch that eyes see », *Nature* 391:756, 1998. doi : 10.1038/35784
- Brattico, E., and Vuust, P., « Brain-to-brain coupling and culture as prerequisites for musical interaction », in *Routledge Companion to Embodied Music Interaction*, eds M. Lesaffre, P. J. Maes, and M. Leman, New York, NY : Routledge, 2017, p. 249-257. doi : 10.4324/9781315621364-28
- Broeckx, J. L., Muziek, *Ratio en Affect : Over de Wisselwerking van Rationeel Denken en Affectief Beleven bij Voortbrengst en Ontvangst van Muziek*, 1981, Antwerp: Metropolis.
- Bryandt, G. A., « Animal signals and emotion in music : coordinating affect across group », *Front. Psychol.* 4:990., 2012. doi : 10.3389/fpsyg.2013.00990
- Bufacchi, R. J., and Iannetti, G. D., « An action field theory of peripersonal space », *Trends Cogn. Sci.* 22, 2018, p. 1076-1090. doi : 10.1016/j.tics.2018.09.004
- Caccioppo, J., Bernston, G., and Decety, J., « Social neuroscience and its relationship to social psychology », *Soc. Cogn.* 28, 2010, p. 675-685. doi : 10.1521/soco.2010.28.6.675
- Candidi, M., Maria Sacheli, L., Mega, I., and Aglioti, S. M. , « Somatotopic mapping of piano fingering errors in sensorimotor experts : TMS studies in pianists and visually trained musically naives », *Cereb. Cortex* 24, 2012, p. 435-443. doi : 10.1093/cercor/bhs325
- Canzoneri, E., Magosso, E., and Serino, A., « Dynamic sounds capture the boundaries of peripersonal space representation in humans », *PLoS One* 7:e44306. 2012. doi : 10.1371/journal.pone.0044306
- Chemero, A., *Radical Embodied Cognitive Science*, Cambridge, MA : MIT

Press, 2009.

Clark, A., *Being There : Putting Brain, Body and World Together Again*, Cambridge, MA : MIT Press. 1997.

Clark, A., *Surfing Uncertainty*, Oxford : Oxford University Press, 2016.

Clark, A., and Chalmers, D., « The extended mind », *Analyses* 58, 1998, p. 7-19.

Clarke, E. F., *Ways of Listening. An Ecological Approach to the Perception of Musical Meaning*, New York, NY : Oxford University Press, 2005.

Clarke, E. F., De Nora, T., and Vuoskoski, J., « Music, empathy and cultural understanding », *Phys. Life Rev.* 15, 2015, p. 61-88. doi : 10.1016/j.plrev.2015.09.001

Clayton, M., « What is entrainment ? Definition and applications in musical research », *Empir. Musicol. Rev.* 7, 2012, p. 49-56. doi : 10.18061/1811/52979

Clayton, M., Jakubowski, K., and Eerola, T., « Interpersonal entrainment in Indian instrumental music performance : synchronization and movement coordination relate to tempo, dynamics, metrical and cadent structure », *Music. Sci.* 23, 2019, p. 304-331. doi : 10.1177/1029864919844809

Cross, I., « Music and communication in music psychology », *Psychol. Music* 42, 2014, p. 809-819. doi : 10.1177/0305735614543968

D'Ausilio, A., Novembre, G., Fadiga, L., and Keller, P. E., « What can music tell us about social interaction ? », *Trends Cogn. Sci.* 19, 2015, p. 111-114. doi : 10.1016/j.tics.2015.01.005

De Jaegher, H., and Di Paolo, E., « Participatory sense- making an

enactive approach to social cognition », *Phenomenol. Cogn. Sci.* 6, 2007, p. 485-507. doi : 10.1007/s11097-007-9076-9

Dell'Anna, A., Buhmann, J., Six, J., Maes, P. J., and Leman, M., « Timing markers of interaction quality during semi-hockey singing », *Front. Neurosci.* 14:619, 2020a. doi : 10.3389/fnins.2020.00619

Dell'Anna, A., Fossataro, C., Burin, D., Bruno, V., Salatino, A., Garbarini, F., et al., « Entrainment beyond embodiment », *Neuropsychologia* 119, 2018, p. 233-240. doi : 10.1016/j.neuropsychologia.2018.08.017

Dell'Anna, A., Rosso, M., Bruno, V., Garbarini, F., Leman, M., and Berti, A., « Does a musical interaction in a jazz duet modulates peripersonal space ? », *Psychol. Res.* online ahead of print, 2020b. [Cet article a été publié en juillet 2021, dans *Psychology Research*, 85 (1-2), doi : 10.1007/s00426-020-01365-6 (<http://link.springer.com/article/10.1007/s00426-020-01365-6>)]

Della Gatta, F., Garbarini, F., Puglisi, G., Leonetti, A., Berti, A., and Borroni, P., « Decreased motor cortex excitability mirrors own hand disembodiment during the rubber hand illusion », *Elife* 5:e14972, 2016.

Dennett, D., *From Bacteria to Bach and Back*, London : Penguin Books, 2017.

Dewey, J. A., Pacherie, E., and Knoblich, G., « The phenomenology of controlling a moving object with another person », *Cognition* 132, 2014, p. 383-397. doi : 10.1016/j.cognition.2014.05.002

Dissanayake, E., « If music is the food of love, what about survival and reproductive success ? », *Music. Sci.* 12, 2008, p. 169-195. doi : 10.1177/1029864908012001081

- Dunbar, R. I. M., « On the evolutionary function of song and dance », in *Music, Language and Human Evolution*, ed. N. Bannan, Oxford : Oxford University Press, 2012, p. 201-214. doi : 10.1093/acprof:osobl/9780199227341.003.0008
- Fadiga, L., Fogassi, L., Pavesi, G., and Rizzolatti, G., « Motor facilitation during action observation : a magnetic stimulation study », *J. Neurophysiol.* 73, 1995, p. 2608-2611. doi : 10.1152/jn.1995.73.6.2608
- Fitch, W. T., « The biology and evolution of music : a comparative perspective », *Cognition* 100, 2006, p. 173-215. doi : 10.1016/j.cognition.2005.11.009
- Freeman, W. J., « A neurobiological role of music in social bonding », in *The Origins of Music*, eds N. L. Wallin, B. Merker, and S. Brown Cambridge, MA : The MIT Press, 2000, p. 411-424,
- Friston, K., « Life as we know it », *J. R. Soc. Interface* 10:20130475, 2013.
- Friston, K., and Frith, C. D., « A duet for one », *Conscious. Cogn.* 36, 2015, p. 390-405,
- Fritz, T., Hardikar, S., Demoucron, M., Niessen, M., Demey, M., Giot, O., et al., « Musical agency reduces perceived exertion during strenuous physical performance », *Proc. Natl. Acad. Sci. U.S.A.* 110, 2013, p. 17784-17789. doi : 10.1073/pnas.1217252110
- Fuchs, T., and De Jaegher, H., « Enactive Intersubjectivity: participatory sense-making and mutual incorporation », *Phenomenol. Cogn. Sci.* 8, 2009, p. 465-486, doi : 10.1007/s11097-009-9136-4
- Gallagher, S., *Enactive Interventions*, Oxford : Oxford University Press, 2017.
- Gallagher, S., and Allen, M., « Active inference, enactivism and the hermeneutics of social cognition », *Synthese* 195, 2016, p. 2627-2648, doi : 10.1007/s11229-016-1269-8
- Garbarini, F., Fornia, L., Fossataro, C., Pia, L., Gindri, P., and Berti, A., « Embodiment of others' hands elicits arousal responses similar to one's own hands », *Curr. Biol.* 24, 2014, R738-R739. doi : 10.1016/j.cub.2014.07.023
- Goebel, W., and Palmer, C., « Synchronization of timing and motion among performing musicians », *Music Percept.* 26, 2009, p. 427-438. doi : 10.1525/mp.2009.26.5.427
- Gordon, C., Iacoboni, M., and Balasubramaniam, R., « Multimodal music perception engages motor prediction : a TMS study », *Front. Neurosci.* 12:736, 2018. doi : 10.3389/fnins.2018.00736
- Graziano, M. S. A., « Where is my arm ? The relative role of vision and proprioception in the neuronal representation of limb position », *Proc. Natl. Acad. Sci. U.S.A.* 96, 1999, p. 10418-10421. doi : 10.1073/pnas.96.18.10418
- Graziano, M. S. A., and Gandhi, S., « Location of the polysensory zone in the precentral gyrus of anesthetized monkeys », *Exp. Brain Res.* 135, 2000, p. 259-266. doi : 10.1007/s002210000518
- Haggard, P., and Eitam, B., *The Sense of Agency*, Oxford : Oxford University Press, 2015.
- Hari, R., Henriksson, L., Malinen, S., and Parkkonen, L., « Centrality of social interaction in human brain function », *Neuron* 88, 2015, p. 181-193. doi : 10.1016/j.neuron.2015.09.022

- Haslinger, B., Erhard, P., Altenmüller, E., Schroeder, U., Boecker, H., and Ceballos-Baumann, A. O., « Transmodal sensorimotor networks during action observation in professional pianists », *J. Cogn. Neurosci.* 17, 2005, p. 282-293. doi : 10.1162/0898929053124893
- Herholz, S. C., and Zatorre, R. J., « Musical training as a framework for brain plasticity : behaviour, function and structure », *Neuron* 76, 2012, p. 486-502. doi : 10.1016/j.neuron.2012.10.011
- Honing, H., ten Cate, C., Peretz, I., and Trehub, S. E., « Without it no music : cognition, biology and evolution of musicality », *Philos. Trans. R. Soc. B* 370:20140088, 2015. doi : 10.1098/rstb.2014.0088
- Hove, M. J., and Risen, J. L., « It's all in the timing: interpersonal synchrony increases affiliation », *Soc. Cogn.* 27, 2009, p. 949-960. doi : 10.1521/soco.2009.27.6.949
- Hurley, S., *Consciousness in Action*, London : Harvard University Press, 1998.
- Huron, D., « Is music an evolutionary adaptation ? », *Ann. N.Y. Acad. Sci.* 930, 2001, p. 43-61. doi : 10.1111/j.1749-6632.2001.tb05724.x
- Iacoboni, M., Wood, R. P., Brass, M., Bekkering, H., Mazziotta, J., and Rizzolatti, G., « Cortical mechanisms of human imitation », *Science* 286, 1999, p. 2526-2528. doi : 10.1126/science.286.5449.2526
- Iriki, A., Tanaka, M., and Iwamura, Y., « Coding of modified body schema during tool use by macaque postcentral neurones », *Neuroreport* 7, 1996, p. 2325-2330. doi : 10.1097/00001756-199610020-00010
- Iyer, V., « Embodied mind, situated cognition, and expressive micro-timing in African-American music », *Music Percept.* 19, 2002, 387-414. doi : 10.1525/mp.2002.19.3.387
- Janata, P., Tomic, S. T., and Haberman, J. M., « Sensorimotor coupling in music and the psychology of the groove », *J. Exp. Psychol. Gen.* 141, 2012, p. 54-75. doi : 10.1037/a0024208
- Keller, P. E., « Joint action in music performance », in *Enacting Intersubjectivity : A Cognitive and Social Perspective to the Study of Interactions*, eds F. Morganti, A. Carassa, and G. Riva, Amsterdam : IOS Press, 2008, p. 205-221.
- Keller, P. E., and Appel, M., « Individual differences, auditory imagery, and the coordination of body movements and sounds in musical ensembles », *Music Percept.* 28, 2010, p. 27-46. doi : 10.1525/mp.2010.28.1.27
- Keller, P. E., Novembre, G., and Hove, M. J., « Rhythm in joint action : psychological and neuro-physiological mechanisms for real-time interpersonal coordination », *Philos. Trans. R. Soc. B* 369 :20130394, 2014. doi : 10.1098/rstb.2013. 0394
- Kilner, J. M., Friston, K. J., and Frith, C. D., « Predictive coding : an account of the mirror neuron system », *Cogn. Process.* 8, 2007, p. 159-166. doi : 10.1007/s10339-007-0170-2
- Koelsch, S., Gunter, T., Friederici, A., and Schroeger, E., « Brain indices of music processing : "nonmusicians" are musical », *J. Cogn. Neurosci.* 12, 2000,

p. 520-541. doi :

10.1162/089892900562183

Koelsch, S., Vuust, P., and Friston, K., « Predictive processes and the peculiar case of music », *Trends Cogn. Sci.* 23, 2019, p. 63-77. doi : 10.1016/j.tics.2018.10.006

Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese, V., and Rizzolatti, G., « Hearing sounds, understanding actions : action representation in mirror neurons », *Science* 297, 2002, p. 846-848. doi : 10.1126/science.1070311

Kokal, I., Engel, A., Kirschner, S., and Keysers, C., « Synchronized drumming enhances activity in the caudate and facilitates prosocial commitment – if the rhythm comes easily », *PLoS One* 6:e27272, 2011. doi : 10.1371/journal.pone.0027272

Konvalinka, I., and Roepstorff, A., « The two-brain approach : how can mutually interacting brains teach us something about social interaction ? », *Front. Neurosci.* 6:215, 2012. doi : 10.3389/fnhum.2012.00215

Konvalinka, I., Vuust, P., Roepstorff, A., and Frith, C. D., « Follow you, follow me : continuous mutual prediction and adaptation in joint tapping », *Q. J. Exp. Psychol.* 63, 2010, p. 2220-2230. doi : 10.1080/17470218.2010.497843

Landry, S. P., and Champoux, F., « Musicians react faster and are better multisensory integrators », *Brain Cogn.* 111, 2017, p. 156-162. doi : 10.1016/j.bandc.2016.12.001

Leman, M., *Embodied Music Cognition*, Cambridge MA : MIT Press, 2007.

Leman, M., *The Expressive Moment*, Cambridge MA : MIT Press, 2016.

Leman, M., Moelants, D., Varewyck, M., Styns, F., van Noorden, L., and Martens, J.- P., « Activating and relaxing music entrains the speed of beat synchronized walking », *PLoS One* 8:e67932, 2013. doi : 10.1371/journal.pone.0067932

Lerdahl, F., and Jackendoff, R., *A Generative Theory of Tonal Music*, Cambridge, MA : MIT Press, 1983.

Lesaffre, M., Maes, P. J., and Leman, M. (eds), *Routledge Companion to Embodied Music Interaction*, New York, NY : Routledge, 2017.

Livingstone, S. R., and Thompson, W. F., « The emergence of music from the theory of mind », *Music. Sci.* 13, 2009, p. 83-115. doi : 10.1177/1029864909013002061

London, J., *Hearing in Time: Psychological Aspects of Musical Meter*, New York, NY : Oxford University Press, 2004.

Maes, P. J., « Sensorimotor grounding of musical embodiment and the role of prediction : a review », *Front. Psychol.* 7:308, 2016. doi : 10.3389/fpsyg.2016.00308

Malloch, S., and Trevarthen, C., *Communicative Musicality*, Oxford : Oxford University Press, 2009.

Maravita, A., and Iriki, A., « Tools for the body (schema) », *Trends Cogn. Sci.* 8, 2004, p. 79-86. doi : 10.1016/j.tics.2003.12.008

Maravita, A., Spence, C., and Driver, J., « Multisensory integration and the body schema : close to hand and within reach », *Curr. Biol.* 13, 2003, R531-R539.

Marsh, K., Richardson, M., and Schmidt, R., « Social connection through joint

- action and interpersonal coordination », *Top. Cogn. Sci.* 1, 2009, p. 320-339. doi : 10.1111/j.1756-8765.2009.01022.x
- McDermott, J. H., « What can experiments reveal about the origins of music ? », *Curr. Dir. Psychol. Sci.* 18, 2009, p. 164-168. doi : 10.1111/j.1467-8721.2009.01629.x
- McNeill, W. H., *Keeping Together in Time*, Cambridge, MA : Harvard University Press, 1995.
- Mehr, S. A., Singh, M., Knox, D., Ketter, D. M., Pickens-Jones, D., Atwood, S., et al., « Universality and diversity in human song », *Science* 366:eaax0868, 2019.
- Menary, R. (ed.) (2010), *The Extended Mind*, Cambridge, MA : MIT Press.
- Miller, G. F., « Evolution of human music through sexual selection », in *The Origins of Music*, eds N. L. Wallin, B. Merker, and S. Brown, Cambridge, MA : MIT Press, 2000, p. 329-360.
- Mithen, S., *The Singing Neanderthals : The Origins of Music, Language, Mind and Body*, London : Weidenfeld and Nicolson, 2005.
- Moens, B., and Leman, M., « Alignment strategies for the entrainment of music and movement rhythms », *Ann. N.Y. Acad. Sci.* 1337, 2015, 86-93. doi : 10.1111/nyas.12647
- Müller, V., Delius, J. A. M., and Lindenberger, U., « Complex networks emerging during choir singing », *Ann. N.Y. Acad. Sci.* 1431, 2018, p. 85-101. doi : 10.1111/nyas.13940
- Munte, T., Altenmüller, E., and Jäncke, L., « The musician's brain as a model of neuroplasticity », *Nat. Rev. Neurosci.* 3, 2002, p. 473-478.
- Neppi-Mòdona, M., Rabuffetti, M., Folegatti, A., Ricci, R., Spinazzola, L., Schiavone, F., et al., « Bisecting lines with different tools in right brain damaged patients : the role of action programming and sensory feedback in modulating spatial remapping », *Cortex* 43, 2007, p. 397-410. doi : 10.1016/s0010-9452(08)70465-9
- Nettl, B., *The Study of Ethnomusicology : 31 Issues and Concepts*, Urbana-Chicago, IL : University of Illinois Press, 2005.
- Newen, A., De Bruin, B., and Gallagher, S. (eds), *The Oxford Handbook of 4E Cognition*, Oxford : Oxford University Press, 2018.
- Newman-Norlund, R. D., van Schie, H. T., van Zuijlen, A. M., and Bekkering, H., « The mirror neuron system is more active during complementary compared with imitative action », *Nat. Neurosci.* 10, 2007, p. 817-818. doi : 10.1038/nn1911
- Noë, A., *Action in Perception*. Cambridge, MA : MIT Press, 2004.
- Novembre, G., Ticini, L. F., Schütz-Bosbach, S., and Keller, P. E., « Motor simulation and the coordination of self and other in real-time joint action », *Soc. Cogn. Affect. Neurosci.* 9, 2014, p. 1062-1068. doi : 10.1093/scan/nst086
- Novembre, G., Ticini, L. F., Schütz-Bosbach, S., and Keller, P. E., « Distinguishing self and other in joint action. evidence from a musical paradigm », *Cereb. Cortex* 22, 2012, p. 2894-2903. doi : 10.1093/cercor/bhr364

- Nowicki, L., Prinz, W., Grosjean, M., Repp, B. H., and Keller, P. E., « Mutual adaptive timing in interpersonal action coordination », *Psychomusicolog.* 23, 2013, p. 6-20. doi : 10.1037/a0032039
- Osaka, N., Minamoto, T., Yaoi, K., Azuma, M., Minamoto Shimada, Y., and Osaka, M., « How two brains make one synchronized mind in the inferior frontal cortex : fNIRS-based hyper-scanning during cooperative singing », *Front. Psychol.* 6:1811, 2015. doi : 10.3389/fpsyg.2015.01811
- Overy, K., and Molnar-Szakacs, I., « Being together in time: musical experience and the mirror neuron system », *Music Percept.* 26, 2009, p. 489-504. doi : 10.1525/mp.2009.26.5.489
- Pacherie, E., « The phenomenology of joint action : self-agency vs. joint-agency », in *Joint Attention: New Developments*, ed. A. Seemann, Cambridge, MA: MIT Press, 2012, p. 343-389.
- Pan, Y., Novembre, G., Song, B., Li, X., and Hu, Y., « Interpersonal synchronization of inferior frontal cortices tracks social interactive learning of a song », *NeuroImage* 183, 2018, p. 280-290. doi : 10.1016/j.neuroimage.2018.08.005
- Patané, I., Iachini, T., Farnè, A., and Frassinetti, F., « Disentangling action from social space : tool-use differently shapes the space around us », *PLoS One* 11:e0154247, 2016. doi : 10.1371/journal.pone.0154247
- Pellencin, E., Paladino, M. P., Herbelin, B., and Serino, A., « Social perception of others shapes one's own multisensory peripersonal space », *Cortex* 104, 2018, p. 163-179. doi : 10.1016/j.cortex.2017.08.033
- Phillips-Silver, J., and Keller, P. E., « Searching for roots of entrainment and joint action in early musical interactions », *Front. Hum. Neurosci.* 6:26, 2012. doi : 10.3389/fnhum.2012.00026
- Phillips-Silver, J., and Trainor, L. J., « Feeling the beat: movement influences infant rhythm perception », *Science* 308:1430, 2005. doi : 10.1126/science.1110922
- Piana, G., *Filosofia Della Musica*, Milano : Guerini, 1991.
- Reddy, V., and Uithol, S., « Engagement : looking beyond the mirror to understand action understanding », *Br. J. Dev. Psychol.* 2015, 34, p. 101-114. doi : 10.1111/bjdp.12106
- Repp, B. H., and Su, Y. H., « Sensorimotor synchronization : a review of recent research (2006-2012) », *Psychon. Bull. Rev.* 20, 2013, p. 403-452. doi : 10.3758/s13423-012-0371-2
- Rizzolatti, G., and Arbib, M. A., « Language within our grasp », *Trends Neurosci.* 21, 1998, p. 188-194. doi : 10.1016/s0166-2236(98)01260-0
- Rizzolatti, G., and Sinigaglia, C., *Mirrors in the Brain. How our Minds Share Actions and Emotions*, Oxford : Oxford University Press, 2008.
- Rizzolatti, G., and Sinigaglia, C., « The functional role of the parieto-frontal mirror circuit : interpretations and misinterpretations », *Nat. Rev. Neurosci.* 11, 2010, p. 264-274. doi : 10.1038/nrn2805

- Rizzolatti, G., Scandolara, C., Matelli, M., and Gentilucci, M., « Afferent properties of periarculate neurons in macaque monkeys. II. visual responses », *Behav. Brain Res.* 2, 1981, p. 147-163. doi : 10.1016/0166-4328(81)90053-x
- Salimpoor, V. N., Zald, D. H., Zatorre, R. J., Dagher, A., and McIntosh, A. R., « Predictions and the brain : how musical sounds become rewarding », *Trends Cogn. Sci.* 19, 2015, p. 86-91. doi : 10.1016/j.tics.2014.12.001
- Schiavio, A., and Altenmüller, E., « Exploring music-based rehabilitation for Parkinsonism through embodied cognitive science », *Front. Neurol.* 6:217, 2015. doi : 10.3389/fneur.2015.00217
- Schiavio, A., and De Jaegher, H. « Participatory sense-making in joint musical practice », in *Routledge Companion to Embodied Music Interaction*, eds M. Lesaffre, P. J. Maes, and M. Leman, New York, NY : Routledge, 2017, p. 31-39. doi : 10.4324/9781315621364-4
- Schiavio, A., and Menin, D., « Embodied music cognition and the mediation technology », *Crit. Rev. Psychol. Music* 41, 2013, p. 804-814.
- Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., et al., « Toward a second-person neuroscience », *Behav. Brain Sci.* 36, 2013, p. 393-414. doi : 10.1017/s0140525x12000660
- Schulz, M., Ross, B., and Pantev, C., « Evidence for training-induced cross modal reorganization of cortical functions in trumpet players », *NeuroReport* 14, 2003, p. 157-161. doi : 10.1097/00001756-200301200-00029
- Schutz-Bosbach, S., Mancini, B., Aglioti, S. M., and Haggard, P., « Self and other in the human motor system », *Curr. Biol.* 16, 2006, p. 1830-1834.
- Sebanz, N., Bekkering, H., and Knoblich, G., « Joint action: bodies and minds moving together », *Trends Cogn. Sci.* 10, 2006, p. 70-76. doi : 10.1016/j.tics.2005.12.009
- Serino, A., Bassolino, M., Farnè, A., and Làdavas, E., « Extended multisensory space in blind cane users », *Psychol. Sci.* 18, 2007, p. 642-648. doi : 10.1111/j.1467-9280.2007.01952.x
- Small, C., *Musicking: The Meanings of Performing and Listening*, Middletown, CT : Wesleyan University Press, 1998.
- Soliman, T. M., Ferguson, R., Dexheimer, M. S., and Glenberg, A. M., « Consequences of joint action : entanglement with your partner », *J. Exp. Psychol. Gen.* 144, 2015, p. 873-888. doi : 10.1037/xge0000089
- Stephan, M. A., Lega, C., and Penhune, V. B., « Auditory prediction cues motor preparation in the absence of movements », *NeuroImage* 174, 2018, p. 288-296. doi : 10.1016/j.neuroimage.2018.03.044
- Styns, F., Van Noorden, L., Moelants, D., and Leman, M., « Walking on music », *Hum. Mov. Sci.* 26, 2007, p. 769-785.
- Su, Y.-H., and Pöppel, E., « Body movement enhances the extraction of temporal structures in auditory sequences », *Psychol. Res.* 76, 2012, p. 373-382. doi : 10.1007/s00426-011-0346-3
- Tarr, B., Launay, J., and Dunbar, R. I. M., « Music and social bonding : “self-other” merging and neurohormonal

- mechanisms », *Front. Psychol.* 5:1096, 2014. doi : 10.3389/fpsyg.2014.01096
- Teneggi, C., Canzoneri, E., di Pellegrino, G., and Serino, A., « Social modulation of peripersonal space boundaries », *Curr. Biol.* 23, 2013, p. 406-411. doi : 10.1016/j.cub.2013.01.043
- Thompson, E., and Varela, F., « Radical embodiment : neural dynamics and consciousness », *Trends Cogn. Sci.* 5, 2001, p. 418-425. doi : 10.1016/s1364-6613(00)01750-2
- Todd, N. P. M., « Motion in music : a neurobiological perspective », *Music Percept.* 17, 1999, p. 115-126. doi : 10.2307/40285814
- Van der Schyff, D., and Schiavio, A., « Evolutionary musicology meets embodied cognition : biocultural coevolution and the enactive origins of human musicality », *Front. Neurosci.* 11:519, 2017. doi : 10.3389/fnins.2017.00519
- Van Noorden, L., and Moelants, D., « Resonance in the perception of musical pulse », *J. N. Music Res.* 28, 1999, p. 43-66. doi : 10.1076/jnmr.28.1.43.3122
- Varela, F., Thompson, E., and Rosch, E., *The Embodied Mind*. Cambridge, MA : MIT Press, 1991.
- Vesper, C., Butterfill, S., Knoblich, G., and Sebanz, N., « A minimal architecture for joint action », *Neural Netw.* 23, 2010, p. 998-1003. doi : 10.1016/j.neunet.2010.06.002
- Volpe, G., D'Ausilio, A., Badino, L., Camurri, A., and Fadiga, L., « Measuring social interaction in music ensembles », *Philos. Trans. R. Soc. B* 371:20150377, 2016. doi : 10.1098/rstb.2015.0377
- Vuust, P., and Witek, M. A. G., « Rhythmic complexity and predictive coding : a novel approach to modelling rhythm and meter perception in music », *Front. Psychol.* 5:1111, 2014. doi : 10.3389/fpsyg.2014.01111
- Walton, A., Richardson, M. J., Langland-Hassan, P., and Chemero, A., « Improvisation and the self-organization of multiple musical bodies », *Front. Psychol.* 6:313, 2015. doi : 10.3389/fpsyg.2015.00313
- Walton, A., Washburn, A., Langland-Hassan, P., Chemero, A., Kloos, H., and Richardson, M. J., « Creating time : social collaboration in music improvisation », *Top. Cogn. Sci.* 10, 2018, p. 95-119. doi : 10.1111/tops.12306
- Wiltermuth, S. S., and Heath, C., « Synchrony and cooperation », *Psychol. Sci.* 20, 2009, p. 1-5. doi : 10.1111/j.1467-9280.2008.02253.x
- Windsor, W. L., and de Bézenac, C., « Music and affordances », *Music. Sci.* 16, 2012, p. 102-120.
- Wolpert, D. M., Doya, K., and Kawato, M., « A unifying computational framework for motor control and social interaction », *Philos. Trans. R. Soc. B* 358, 2003, p. 593-602. doi : 10.1098/rstb.2002.1238
- Yamaguchi, K., Nakamura, K., Oga, T., and Nakajima, Y., « Eating tools in hand activate the brain systems for eating action : a transcranial magnetic stimulation study », *Neuropsychologia* 59, 2014, p. 142-147. doi : 10.1016/j.neuropsychologia.2014.05.003
- Zimmerman, E., and Lahav, A., « The multisensory brain and its ability to learn music », *Ann. N.Y. Acad. Sci.* 1252,

2012, p. 179-184. doi : 10.1111/j.1749-6632.2012.06455.x

NOTES

- 1 Le concept d'action conjointe est un domaine de recherche actif dans le domaine des sciences cognitives depuis une dizaine d'années. Sebanz *et al.* (2006, p. 70) proposent une définition pratique : une action conjointe correspond à « toute forme d'interaction sociale qui amène au moins deux individus à coordonner leurs actes dans le temps et l'espace pour engendrer un changement dans l'environnement ». (NdT)
- 2 « *Tapping* » : action de tapoter alternativement des doigts ou des mains. Par souci de clarté, nous conserverons dans cette traduction le terme en anglais. (NdT)
- 3 Le principe du hoquet est de faire alterner strictement les différentes notes d'une ligne mélodique entre plusieurs voix. (NdT)
- 4 Nous remercions deux relecteurs pour leur insistance quant à l'utilisation inadéquate du mot « langage » de manière littérale plutôt que métaphorique. Cf. par ailleurs notre citation liminaire de Piana (1994).
- 5 « Le signal BOLD (de l'anglais *blood-oxygen-level dependent*, “dépendant du niveau d'oxygène sanguin”) est le signal qui reflète les variations locales et transitoires de la quantité d'oxygène transporté par l'hémoglobine en fonction de l'activité neuronale du cerveau. » (URL : https://fr.wikipedia.org/wiki/Signal_BOLD [consulté le 29 novembre 2022]). (NdT)
- 6 « L'hyperscanning est une technologie issue des neurosciences qui permet de mesurer simultanément l'activité cérébrale de deux individus au cours d'une interaction sociale, à l'aide d'une imagerie par résonance magnétique fonctionnelle (IRMf) ou d'un double enregistrement en électroencéphalographie. » (URL : <https://www.fondation-fondamental.org/hyperscanning> [consulté le 16 mars 2023]). (NdT)
- 7 Les composantes situées et énaactives ne seront pas traitées ici, car elles exigeraient un autre article, qui mériterait d'être écrit à l'avenir, comme nous l'évoquerons dans la conclusion.
- 8 « Il existe des neurones bimodaux visuo-tactiles, voire multisensoriels, qui répondent aux stimuli visuels, auditifs et tactiles », Arlette Streri,

[alis.fr/encyclopedie/perception-tactile/](https://www.alis.fr/encyclopedie/perception-tactile/) [consulté le 29 novembre 2022]. (NdT)

9 Pour le dire autrement, ce sont ces neurones qui permettent par exemple de fournir une représentation *visuelle* de l'espace proche de ces parties du corps. (NdT)

10 « Le *Line Bisection Test* est une mesure rapide pour détecter la présence d'une négligence spatiale unilatérale (NSU). Pour compléter le test, il suffit de mettre une marque avec un crayon à travers le centre d'une série de lignes horizontales. Généralement, un décalage de la marque de bissection vers le côté de la lésion cérébrale est interprété comme un symptôme de négligence. » URL : <https://strokengine.ca/fr/assessments/line-bisection-test/> [consulté le 29/11/2022]. (NdT)

11 Le « *timing* » envisagé comme la pulsation intérieure, elle-même considérée comme un ressenti temporel subjectif, qui détermine alors le ressenti temporel musical. (NdT)

12 « Les phénomènes musicaux pourraient évoquer les vagues dynamiques associées à la respiration, le tempo régulier associé à la marche et les figures rythmiques rapides associées à la parole », Iyer, 2002, p. 392. (NdT)

13 Ce motif musical rythmique ambigu est composé de six pulsations, comprenant des sons de caisse claire et de bâton claqueur, jouées sans accentuation particulière. (NdT)

14 Traduction littérale du terme « *sense-making* » qui s'inscrit dans le contexte théorique général de la prise de décision et de la manière dont les individus font sens des actions qui les entourent. Cf. l'article de De Jaegher et Di Paolo (2007), qui donne des éléments de compréhension de la notion. (NdT)

15 « En physiologie, un potentiel évoqué se définit comme la modification de l'activité électrique du système nerveux en réponse à une stimulation extérieure qu'elle soit visuelle, auditive, sensitive ou motrice. » (URL : <https://www.ffn-neurologie.fr/grand-public/explorations-neurologiques/potentiels-evoques> [consulté le 17/01/2023]). (NdT)

16 « Facilitation » : « Phénomène bioélectrique aboutissant à rendre plus efficace la transmission du message nerveux. » (*Larousse*. URL : <https://www.larousse.fr/dictionnaires/francais/facilitation/32584> [consulté le 17 janvier 2023]). (NdT)

- 17 « *First dorsal interosseous (FDI)* » : premiers muscles interosseux dorsaux (IOD) de la main. (NdT)
- 18 STG, gyrus temporal supérieur ; STS, sillon temporal supérieur. (NdT)
- 19 À envisager dans le sens de la sensation interne d'avoir la capacité à agir. (NdT)
- 20 Littéralement : expérience motrice et affective partagée. (NdT)
- 21 Le *timing* est ici envisagé comme la pulsation intérieure, elle-même considérée comme un ressenti temporel subjectif, qui détermine alors le ressenti temporel musical. (NdT)
- 22 « La participation intégrative priorisée s'appuie à la fois sur la capacité d'une personne à répartir son attention entre les différentes sources sonores et sur les compétences d'attention conjointe du groupe (dans la mesure où plusieurs personnes participent à l'élaboration de la structure globale qui résulte de leurs actions coordonnées) (Jessica Phillips-Silver et Peter E. Keller, « Searching for roots of entrainment and joint action in early musical interactions », *Frontiers in human neuroscience*, vol. 6 (2012). DOI : <https://doi.org/10.3389/fnhum.2012.00026> [consulté le 21 janvier 2025]). (NdT)
- 23 « Piaget distingue les images *statiques*, qui concernent la représentation d'états, et les images de *transformation*, qui portent sur les étapes d'une modification lors de la représentation d'un processus de changement » (Magali Bovet et Daphn Voelin, « Le rôle de l'image mentale dans le raisonnement opératoire : auxiliaire ou structurant ? », *Enfance*, vol. 55 (2), 2003. DOI : <https://doi.org/10.3917/enf.552.0175> [consulté le 21 janvier 2025]). (NdT)
- 24 Ces deux mécanismes, qui correspondent respectivement aux deux cas précités, sont indépendants : l'un, automatique, n'affecte pas le timing ; l'autre, intentionnel, l'affecte (cf. l'article cité, notamment p. 407). (NdT)
- 25 « L'inférence active est le cadre computationnel (<https://fr.wikipedia.org/wiki/Computationalisme>) qui décrit comment un agent (comme une personne ou une intelligence artificielle) peut interagir avec son environnement pour atteindre un résultat souhaité. » (Wikipédia, « Principe d'énergie libre », URL : [https://fr.wikipedia.org/wiki/Principe_de_l%27%C3%A9nergie_libre#Inf%C3%A9rence_active_et_th%C3%A9orie_\(des_jeux\)_de_la_d%C3%A9cision_optimale](https://fr.wikipedia.org/wiki/Principe_de_l%27%C3%A9nergie_libre#Inf%C3%A9rence_active_et_th%C3%A9orie_(des_jeux)_de_la_d%C3%A9cision_optimale) [consultée le 07 février 2023]). (NdT)

26 « En éducation à l'environnement, Hayward (2012) la définit pour sa part comme une capacité à développer une pensée indépendante et une capacité à choisir librement d'agir en fonction de ses idées. Il s'agit donc d'une habileté, d'une capacité ou d'une capacité propre à un individu. » (Émilie Morin, Geneviève Therriault et Barbara Bader, « Le développement du pouvoir agir, l'agentivité et le sentiment d'efficacité personnelle des jeunes face aux problématiques sociales et environnementales : apports conceptuels pour un agir ensemble », *Éducation et socialisation, Les Cahiers du CERFEE*, vol. 51 (2019). DOI : <https://doi.org/10.4000/edso.5821> [consulté le 21 janvier 2025]). (NdT)

27 Il s'agissait, en l'occurrence, d'appareils de fitness, et la tâche active pour les participants consistait à utiliser ces appareils avec un retour musical lié aux mouvements qu'ils étaient en train d'effectuer. (NdT)

28 « Entraînement » doit être entendu ici comme l'interaction entre des processus rythmiques indépendants les uns des autres (cf. Clayton, 2012). (NdT)

29 C'est-à-dire, respectivement, en alternance et en synchronie. (NdT)

30 C'est-à-dire qui se produit simultanément au stimulus sonore. (NdT)

31 Il s'agit de l'adaptation dans le style du hoquet d'une chanson de Michael Jackson, *Billy Jean*. (NdT)

32 Les asynchronies signées, mesurées en millisecondes, se rapportent à la précision de la synchronisation, et sont soit négatives (lorsqu'il y a anticipation par rapport au stimulus), soit positives (lorsqu'il y a retard par rapport au stimulus) (cf. Matthieu Ghilain, *Synchronisation au rythme de la musique et effet du contexte social dans la maladie d'Alzheimer et le vieillissement physiologique*, Thèse de doctorat non publiée en Psychologie, Université de Lille, Villeneuve d'Ascq, 2019). (NdT)

33 Cf. Figure 1. (NdT)

34 Traduction du terme par Jedediah Skower. Cf. Christopher Small, *Musiquer. Le sens de l'expérience musicale*, Paris, Cité de la musique-Philharmonie de Paris, « la rue musicale », (1998) 2019, trad. par Jedediah Sklower. Voici la définition du terme formulée par Small : « musiquer, c'est participer, de quelque manière que ce soit, à une performance musicale, en jouant, en écoutant, en répétant ou en pratiquant, en fournissant un matériau pour une performance (ce qu'on appelle composer), ou en dansant » (p. 33). (NdT)

ABSTRACTS

Français

Les sciences de la vie et les sciences sociales se concentrent souvent sur la nature sociale de la musique (et du langage). En biologie, par exemple, les trois principales hypothèses d'évolution concernant la musique (c'est-à-dire la sélection sexuelle, le lien parent-enfant et la cohésion de groupe) mettent en lumière son caractère intrinsèquement social (Honing *et al.*, 2015). Depuis plus de deux décennies, la neurobiologie étudie donc les fondements neuronaux et hormonaux liés à la musicalité (Chanda & Levitin, 2013 ; Salimpoor *et al.* (<https://www.frontiersin.org/articles/10.3389/fnins.2021.667838/full#B119>), 2015 ; Mehr *et al.* (<https://www.frontiersin.org/articles/10.3389/fnins.2021.667838/full#B119>), 2019). Conformément à ces approches, l'objectif de cet article est de démontrer que la manière adéquate de saisir la nature sociale interactive de la musique (et, avant cela, de la musicalité) consiste à la concevoir comme un langage incarné, enraciné dans des structures cérébrales s'adaptant au contexte culturel. Cette proposition répond à celle de Ian Cross d'étudier la musique comme un « processus de communication interactif » plutôt que comme « une manifestation de formes sonores » (Cross, 2014), en insistant sur ses aspects incarnés et prédictifs (de codage) (Clark, 2016 ; Lemm, 2016 ; Koelsch *et al.*, 2019). Dans cet article, notre objectif est de : (i) proposer un cadre considérant la musique sous l'aspect du langage incarné en s'appuyant sur une revue de littérature qui définit les principaux concepts de l'action musicale conjointe¹, en mettant particulièrement l'accent sur la cognition musicale incarnée et le traitement prédictif, ainsi que sur certains fonctionnements neuronaux correspondants ; (ii) résumer trois expériences réalisées dans nos laboratoires (et récemment publiées), dont les résultats à la fois appuient l'hypothèse de ce nouveau cadre conceptuel et peuvent être interprétés en fonction de celui-ci. Ce faisant, nous nous appuyons à la fois sur les sciences cognitives musicales et sur les neurosciences pour dessiner un cadre conceptuel global consacré à l'interaction musicale. Celui-ci explore plusieurs aspects de la pratique musicale en duo, depuis une action protomusicale très basique, comme le *tapping*², jusqu'à des actions plus complexes³, comme jouer un standard de jazz et chanter une mélodie en hoquet³. Notre cadre propose de considérer le concept d'action concertée en y intégrant les notions d'incarnation et de prédiction (Pacherie, 2012 ; Keller *et al.*, 2016 ; Bolt & Loehr, 2017). Si l'interaction sociale est le « mode par défaut » via lequel les cerveaux humains communiquent avec leur environnement (Hari *et al.*, 2015), la musique et la musicalité envisagées comme un langage incarné peuvent sans doute y contribuer.

English

Life and social sciences often focus on the social nature of music (and language alike). In biology, for example, the three main evolutionary hypo-

theses about music (i.e., sexual selection, parent-infant bond, and group cohesion) stress its intrinsically social character (Honing *et al.*, 2015). Neurobiology thereby has investigated the neuronal and hormonal underpinnings of musicality for more than two decades (Chanda & Levitin, 2013; Salimpoor *et al.*, 2015; Mehr *et al.*, 2019). In line with these approaches, the present paper aims to suggest that the proper way to capture the social interactive nature of music (and, before it, musicality), is to conceive of it as an embodied language, rooted in culturally adapted brain structures (Clarke *et al.*, 2015; D'Ausilio *et al.*, 2015). This proposal heeds Ian Cross' call for an investigation of music as an "interactive communicative process" rather than "a manifestation of patterns in sound" (Cross, 2014), with an emphasis on its embodied and predictive (coding) aspects (Clark, 2016; Leman, 2016; Koelsch *et al.*, 2019). In the present paper our goal is : (i) to propose a framework of music as embodied language based on a review of the major concepts that define joint musical action, with a particular emphasis on embodied music cognition and predictive processing, along with some relevant neural underpinnings; (ii) to summarize three experiments conducted in our laboratories (and recently published), which provide evidence for, and can be interpreted according to, the new conceptual framework. In doing so, we draw on both cognitive musicology and neuroscience to outline a comprehensive framework of musical interaction, exploring several aspects of making music in dyads, from a very basic proto-musical action, like *tapping*, to more sophisticated contexts, like playing a jazz standard and singing a hocket melody. Our framework combines embodied and predictive features, revolving around the concept of joint agency (Pacherie, 2012; Keller *et al.*, 2016; Bolt & Loehr, 2017). If social interaction is the "default mode" by which human brains communicate with their environment (Hari *et al.*, 2015), music and musicality conceived of as an embodied language may arguably provide a route toward its navigation.

INDEX

Mots-clés

cognition musicale incarnée, action musicale conjointe, codage prédictif, musique comme langage, sens de l'action concertée

Keywords

embodied music cognition, musical joint action, predictive coding, music as language, sense of joint agency

AUTHORS

Alessandro DELL'ANNA

Department of Art, Music, and Theatre Sciences, IPEM, Ghent University, Ghent, Belgium & SAMBA Research Group, Department of Psychology, University of Turin, Turin, Italy

Marc LEMAN

Department of Art, Music, and Theatre Sciences, IPEM, Ghent University, Ghent, Belgium

Annamaria BERTI

SAMBA Research Group, Department of Psychology, University of Turin, Turin, Italy

TRANSLATORS

David CAULET

Félix EMERY

Joséphine GÉDÉON-GONÇALVES

Fanny HERMANT

Julie LABEQUE

Matilde NORTIER

Paul VENNIN

Christian HAUER